SIAM J. COMPUT.
Vol. 48, No. 3, pp. 1098–1119

THE LOVÁSZ THETA FUNCTION FOR RANDOM REGULAR GRAPHS AND COMMUNITY DETECTION IN THE HARD REGIME*

JESS BANKS†, ROBERT KLEINBERG‡, AND CRISTOPHER MOORE§

Abstract. We derive upper and lower bounds on the degree d for which the Lovász ϑ function, or equivalently sum-of-squares proofs with degree two, can refute the existence of a k-coloring in random regular graphs $G_{n,d}$. We show that this type of refutation fails well above the k-colorability transition, and in particular everywhere below the Kesten–Stigum threshold. This is consistent with the conjecture that refuting k-colorability, or distinguishing $G_{n,d}$ from the planted coloring model, is hard in this region. Our results also apply to the disassortative case of the stochastic block model, adding evidence to the conjecture that there is a regime where community detection is computationally hard even though it is information-theoretically possible. Using orthogonal polynomials, we also provide explicit upper bounds on $\vartheta(\overline{G})$ for regular graphs of a given girth, which may be of independent interest.

Key words. sum-of-squares proofs, semidefinite relaxation, graph coloring, random graphs, community detection, planted partitions

AMS subject classifications. 05C80, 05C15, 90C22, 90B15

DOI. 10.1137/18M1180396

1. Introduction.

Many constraint satisfaction problems have *phase transitions* in the random case: as the ratio between the number of constraints and the number of variables increases, there is a critical value at which the probability that a solution exists, in the limit $n \to \infty$, suddenly drops from one to zero. Above this transition, most instances are too constrained and hence unsatisfiable. But how many constraints do we need before it becomes easy to *prove* that a typical instance is unsatisfiable? When is there likely to be a short refutation, which we can find in polynomial time, proving that no solution exists?

For a closely related problem, suppose that a constraint satisfaction problem is generated randomly, but with a particular solution "planted" in it. Given the instance, can we recover the planted solution, at least approximately? For that matter, can we tell whether the instance was generated from this planted model, as opposed to an unplanted model with no built-in solution? We can think of this as a statistical inference problem. If there is an underlying pattern in a dataset (the planted solution) but also some noise (the probabilistic process by which the instance is generated), the question is how much data (how many constraints) do we need before we can find the pattern, or confirm that one exists.

Here we focus on the k-colorability of random graphs, and more generally the community detection problem. Let $G = G(n, p = d/n)$ denote the Erdős–Rényi graph with n vertices and average degree d. A simple first moment argument shows that

*Received by the editors April 12, 2018; accepted for publication (in revised form) March 11, 2019; published electronically May 21, 2019.
http://www.siam.org/journals/sicomp/48-3/M118039.html
Funding: The third author was supported by the John Templeton Foundation, the Army Research Office under grant W911NF-12-R-0012, and NSF grant BIGDATA-1838251.
†Department of Mathematics, University of California-Berkeley, Berkeley, CA 94720 (jess.m.banks@berkeley.edu).
‡Department of Computer Science, Cornell University, Ithaca, NY 14853 (robert.kleinberg@cornell.edu).
§Santa Fe Institute, Santa Fe, NM 87501 (moore@santafe.edu).

with high probability G is not k-colorable if

$$d \geq d_{\text{first}} = 2k \ln k - \ln k \,. \tag{1}$$

(We say an event E_n on graphs of size n holds with high probability if $\lim_{n\to\infty} \Pr[E_n] = 1$, and with positive probability if $\liminf_{n\to\infty} \Pr[E_n] > 0$.) Sophisticated uses of the second moment method [8, 24] show that this is essentially tight, and that the k-colorability transition occurs at

$$d_c = d_{\text{first}} - O_k(1) \,.$$

Now consider the planted coloring model, where we choose a coloring σ uniformly at random and condition G on the event that σ is proper. If $d > d_c$, then $G(n, d/n)$ is probably not k-colorable, while graphs drawn from the planted model are k-colorable by construction. Thus, above the k-colorability transition, we can tell with high probability whether G was drawn from the planted or unplanted model by checking to see if G is k-colorable. However, searching exhaustively for k-colorings would take exponential time.

A similar situation holds for the stochastic block model, a model of graphs with community structure also known as the planted partition problem (see [52, 1] for reviews). For our purposes, we will define it as follows: Fix a constant τ, and say a partition σ of the vertices into k groups is "good" if a fraction τ/k of the edges connect vertices within groups. Equivalently, if G has m edges, σ is a multiway cut with $(1 - \tau/k)m$ edges crossing between groups. Generalizing the planted coloring model where $\tau = 0$, the block model chooses σ uniformly, and conditions G on the event that σ is good. The cases $\tau > 1$ and $\tau < 1$, where vertices are more or less likely to be connected to others in the same group, are called *assortative* (or *ferromagnetic*) and *disassortative* (or *antiferromagnetic*), respectively.

Two natural problems related to the block model are *detection*, i.e., telling with high probability whether G was drawn from the block model or from $G(n, d/n)$, and *reconstruction*, finding a partition which is significantly correlated with the planted partition σ. (This is sometimes called *weak* reconstruction to distinguish it from finding σ exactly, which becomes possible when $d = \Theta(\log n)$ [16, 2, 3, 31, 32, 9, 54].) Both problems become information-theoretically possible at a point called the condensation transition [40, 22, 21], and the first and second moment methods [12] show that this scales as

$$d_c \sim \frac{2k \log k}{(\tau - 1)^2} \,, \tag{2}$$

where $\sim$ denotes equality up to $(1 \pm o_k(1))$ factors. As in k-coloring this is roughly the first moment bound above which, with high probability, no good partitions exist in $G(n, d/n)$. However, the obvious algorithms for detection and reconstruction, such as searching exhaustively for good partitions or sampling from an appropriate Gibbs distribution [6, 4], require exponential time.

In fact, conjectures from statistical physics [42, 25, 26] suggest this exponential difficulty is sometimes unavoidable. Specifically, these conjectures state that polynomial-time algorithms for detection and reconstruction exist if and only if d is above the *Kesten–Stigum threshold* [36, 37],

$$d_{\text{KS}} = \left(\frac{k-1}{\tau - 1} \right)^2 . \tag{3}$$

Several polynomial-time algorithms are now known to succeed whenever $d > d_{\mathrm{KS}}$, including variants of belief propagation [53, 5] and spectral algorithms based on nonbacktracking walks [56, 41, 47, 17]. Moreover, for $k = 2$ we know that the information-theoretic and Kesten–Stigum thresholds coincide [55]. Comparing (2) and (3) we see that for any $\tau \neq 1$ we have $d_c < d_{\mathrm{KS}}$ for sufficiently large k, and in fact this occurs for some $\tau < 1$ when $k = 4$ and more generally when $k \geq 5$ [6, 4, 12].

Thus in the regime $d_c < d < d_{\mathrm{KS}}$, detection and reconstruction are information-theoretically possible, but are conjectured to be computationally hard. In particular, this conjecture implies that there is no way to refute the existence of a coloring, or of a good partition, whenever $d < d_{\mathrm{KS}}$, even when d is large enough so that a coloring or partition probably does not exist. Our goal in this paper is to rule out spectral refutations based on the Lovász theta function, or equivalently sum-of-squares proofs of degree two.

For technical reasons, we focus on random d-regular graphs, which we denote $G_{n,d}$. A series of papers applying the first and second moment methods in this setting [50, 7, 35, 20] have determined the likely chromatic number of $G_{n,d}$ for almost all d, showing that the critical d for k-colorability is $d_c = d_{\mathrm{first}} - O(1)$ just as for $G(n, d/n)$. (There are a few values of d and k where $G_{n,d}$ could be k-colorable with probability strictly between 0 and 1, so this transition might not be completely sharp.)

We define the d-regular block model by choosing a planted partition σ uniformly at random and conditioning $G_{n,d}$ on the event that σ is good. Equivalently, we choose G uniformly from all d-regular graphs such that a fraction τ/k of their $m = dn/2$ edges connect vertices within groups. We claim that our results also apply to the regular block model proposed in [55] where d-regular graphs are chosen with probability proportional to $\tau^{\#\text{ within-group edges}}((k-\tau)/(k-1))^{\#\text{ between-group edges}}$: in that case, the fraction of within-group edges fluctuates, but is $\tau/k + o(1)$ with high probability.[1] We again conjecture that refuting the existence of a coloring or a good partition is exponentially hard below the Kesten–Stigum bound. Since the branching ratio of a d-regular tree is $d - 1$, in the regular case this becomes

$$d < d_{\mathrm{KS}} = \left(\frac{k-1}{\tau-1}\right)^2 + 1\,.$$

Main results. The Lovász ϑ function, which we review below, gives a lower bound on the chromatic number which can be computed in polynomial time. In particular, if $\vartheta(\overline{G}) > k$, this provides a polynomial-time refutation of G's k-colorability. We first prove that this type of refutation exactly corresponds to sum-of-squares proofs of degree two in a natural encoding of k-colorability as a system of polynomials; this is intuitive, but it does not seem to have appeared in the literature. We then show the following bounds on the likely value of $\vartheta(\overline{G})$ when G is a random d-regular graph.

THEOREM 1. *Let d be constant. For any constant $\epsilon > 0$, with high probability*

$$\frac{d}{2\sqrt{d-1}} + 1 - \epsilon \leq \vartheta(\overline{G_{n,d}}) \leq \frac{d}{2\sqrt{d-1}} + 2 + \epsilon\,.$$

As a consequence, the Lovász ϑ function cannot refute k-colorability with high probability

[1] These models are not to be confused with a stricter model, where for some constants q_{rs} each vertex in group r has exactly q_{rs} neighbors in group s [18, 23, 58, 15]. Our model only constrains the total number of edges within or between groups.

if

$$k > 2 + \frac{d}{2\sqrt{d-1}}\,, \tag{4}$$

and in particular if d *is below the Kesten–Stigum threshold.*

Rearranging, no refutation of this kind can exist when

$$d < 2(k-2)\left((k-2) + \sqrt{(k-2)^2 - 1}\right) = (4 - o_k(1))d_{\mathrm{KS}}\,.$$

Our lower bound on $\vartheta(\overline{G_{n,d}})$ follows easily from Friedman's theorem [29] on the spectrum of $G_{n,d}$. For the upper bound, we first use orthogonal polynomials to derive explicit bounds on $\vartheta(\overline{G})$ for arbitrary regular graphs of a given girth—which may be of independent interest—and then employ a concentration argument for $G_{n,d}$.

We also relate the Lovász ϑ function to the existence of a good partition in the disassortative case of the block model, giving the following theorem.

THEOREM 2. *Fix* $\tau < 1$ *and say a partition is* good *if a fraction* τ/k *of its edges connect endpoints in the same group. Then sum-of-squares proofs of degree two cannot refute the existence of a good partition in* $G_{n,d}$ *if*

$$\frac{k-\tau}{1-\tau} > 2 + \frac{d}{2\sqrt{d-1}}\,.$$

Thus degree-two sum of squares cannot distinguish the regular stochastic block model from $G_{n,d}$ until d is roughly a factor of 4 above the Kesten–Stigum threshold.

Related work. The distributions of $\vartheta(\overline{G})$ for the Erdős–Rényi graph $G = G(n,p)$ and the random d-regular graph $G = G_{n,d}$ were studied in [19]. In particular, that work showed that when d is sufficiently large, with high probability $\vartheta(\overline{G_{n,d}}) > c\sqrt{d}$ for a constant $c > 0$. Our results tighten this lower bound, making the constant c explicit, and provide a nearly matching upper bound.

The power of semidefinite programs (SDPs) for distinguishing with high probability the sparse Erdős–Rényi and random regular distributions from their respective two-group block models, as opposed to refuting the existence of colorings or sparse cuts, was examined in [51]. They construct a feasible solution using the covariances of a certain Gaussian process on the vertices of the graph, designed so that the correlations between variables depend only on their shortest path distance in the graph. Our approach shares the same spirit, but the SDP they use is less constrained than that for the Lovász ϑ, and our analysis differs substantially from theirs. In particular, we construct a sequence of feasible solutions conditioned on the girth being sufficiently large, each of which is based on a deterministic function of graph distance derived from a family of orthogonal polynomials.

Our results on the power of degree-two sum-of-squares refutations for k-colorability contribute to a recent line of work on refutations of random constraint satisfaction problems (CSPs), which we briefly survey. If we define the density of a CSP as the ratio of constraints to variables—which for coloring equals half the average degree of the graph—then the conjectured hard regime for k-coloring corresponds to a range of densities bounded below and above by constants (i.e., depending on k but not n). For CSPs such as k-SAT and k-XOR, there is again a satisfiability transition at constant density, but with high probability sum-of-squares refutations with constant degree do not exist unless the density is much higher, namely, $\Omega(n^{k/2-1})$ [60], a result which was recently extended to a more general class of CSPs. It is shown in [38] that there is no constant degree sum-of-squares refutation until the clause density is $\Omega(n^{k/2-1})$, whenever the constraint predicate—the Boolean function on subsets of variables that decides whether a given clause is satisfied—supports a $(k-1)$-wise uniform distribution. That is, there is a distribution on the set of satisfying assignments for each clause with the property that every collection of at most $k-1$ variables is uniformly distributed.

Conversely, if a predicate does not support a t-wise uniform distribution, then [10] shows that there is an efficient sum-of-squares refutation when the density is $\tilde{O}(n^{t/2}-1)$. For coloring, the constraint predicate is the Boolean function querying whether the two colors on the endpoints of an edge are the same, and this does not support a 2-wise uniform distribution. This gives refutations at roughly constant density; our contribution makes this a nearly precise constant in the special case of degree-two sum of squares on random regular graphs.

The hidden clique problem also has a conjectured hard regime. It is well known that the random graph $G(n, 1/2)$ has no cliques larger than $O(\log n)$ [28], but it is conjectured to be computationally hard to distinguish $G(n, 1/2)$ from a graph with a planted clique of size $o(n^{1/2})$. A sequence of progressively stronger sum-of-squares lower bounds for this problem [27, 33, 49] have culminated in the theorem that with high probability the degree-d sum-of-squares proof system cannot refute the existence of a clique of size $n^{1/2-c(d/\log n)^{1/2}}$ in $G(n, 1/2)$ for some constant $c > 0$ [13].

In contrast to the aforementioned work on refuting random k-CSPs and planted cliques, our result pertains to a much more specific pair of problems, namely, k-coloring and the stochastic block model, and only to degree-two sum-of-squares refutations; but it attains a sharp bound, within an additive constant, on the density at which these refutations become possible. We conjecture that sum-of-squares refutations of any constant degree do not exist below the Kesten–Stigum threshold, but it seems difficult to extend our current techniques to degree higher than two.

2. Colorings, partitions, and the Lovász ϑ function.

2.1. Background on sum of squares. One type of refutation which has gained a great deal of interest recently is sum-of-squares proofs; see [14] for a review. Suppose we encode our variables and constraints as a system of m polynomial equations on n variables, $f_j(x_1, x_2, \ldots, x_n) = 0$ for all $j = 1, \ldots, m$. One way to prove that no solution $\boldsymbol{x} \in \mathbb{R}^n$ exists—in algebraic terms, that this variety is empty—is to find a linear combination of the f_j which is greater than zero for all $\boldsymbol{x}$. Moreover, the *positivstellensatz* of Krivine [39] and Stengle [63] shows that a polynomial is nonnegative over $\mathbb{R}^n$ if and only if it is a sum of squares (SOS) of rational functions. Thus, clearing denominators, it is sufficient to find $g_1, \ldots, g_m$ and $h_1, \ldots, h_t$ and a constant $\epsilon > 0$ (which we can always scale to 1 if we like) such that

$$\sum_{j=1}^{m} g_j(\boldsymbol{x}) f_j(\boldsymbol{x}) = S + \epsilon, \quad \text{where} \quad S = \sum_{\ell=1}^{t} h_\ell(\boldsymbol{x})^2 \,. \tag{5}$$

This proof technique is complete as well as sound. That is, there is such a set of polynomials $\{g_j\}$ and $\{h_\ell\}$ if and only if no solution exists.

Even when the f_j are of low degree, the polynomials g_j and h_ℓ might be of high degree, making them difficult to find. However, we can ask when a refutation exists where both sides of (5) have degree δ or less. As we take $\delta = 2, 4, 6, \ldots$ we obtain the *SOS hierarchy.* The case $\delta = 2$ is typically equivalent to a familiar semidefinite relaxation of the problem. More generally, a degree-δ refutation exists if and only if a certain SDP on $O(n^\delta)$ variables is feasible; thus we can find degree-δ refutations, or confirm that they do not exist, in time $\text{poly}(n^\delta)$ [61, 57, 59, 43].

To see why, note that there is a natural identification between $\mathbb{R}[\boldsymbol{x}]_{\le\delta}$, the vector space of polynomials of degree at most δ, and the bilinear forms on $\mathbb{R}[\boldsymbol{x}]_{\le\delta/2}$. Given such a bilinear form $\mathcal{S}$, written down in the basis of monomials $x^{(\alpha)} = \prod_i x^{\alpha_i}$, we can form the polynomial

$$S(\boldsymbol{x}) = \sum_{\alpha,\alpha'} \mathcal{S}(\alpha,\alpha')\, x^{(\alpha)} x^{(\alpha')} ,$$

and similarly any polynomial can be expressed (often nonuniquely) as such a bilinear form. A positive semidefinite bilinear form $\mathcal{S}$ can be decomposed as $\mathcal{S} = \sum_{\ell\in[t]} w_\ell \otimes w_\ell$ for some collection of vectors $w_1, \dots, w_t$, so we can write its associated polynomial as

$$S(\boldsymbol{x}) = \sum_{\alpha,\alpha'} \mathcal{S}(\alpha,\alpha')\, x^{(\alpha)} x^{(\alpha')} = \sum_{\alpha,\alpha'} \sum_{\ell\in[t]} w_\ell(\alpha) w_\ell(\alpha')\, x^{(\alpha)} x^{(\alpha')} = \sum_{\ell\in[t]} \left(\sum_\alpha w_\ell(\alpha) x^{(\alpha)} \right)^2 .$$

Finally, the constraint that $S = \sum_j g_j f_j - \epsilon$ for some $\{g_j\}$ and some $\epsilon > 0$ corresponds to a set of affine constraints on the entries of $\mathcal{S}$.

The dual object to a degree-δ refutation is a *pseudoexpectation.* This is a linear operator $\tilde{\mathbb{E}}$ on polynomials of degree at most δ with the properties that

$$\tilde{\mathbb{E}}[1] = 1, \tag{6}$$

$$\tilde{\mathbb{E}}[f_j q] = 0 \quad \text{for all } j \text{ and every polynomial } q \text{ of degree at most } \delta - \deg f_j\, , \tag{7}$$

$$\tilde{\mathbb{E}}[p^2] \ge 0 \quad \text{for any polynomial } p \text{ of degree at most } \delta/2\, . \tag{8}$$

If we write $\tilde{\mathbb{E}}$ as a bilinear form on monomials $x^{(\alpha)}$, then (6) and (7) are linear constraints on its entries, and (8) states that this matrix is positive semidefinite. Under some mild conditions on the polynomial equations we are trying to refute, the resulting SDP is dual to the SDP for refutations, so each of these SDPs is feasible precisely when the other is not. Thus there is a degree-δ refutation if and only if no degree-δ pseudoexpectation exists, and vice versa.

We can think of a pseudoexpectation as a way for an adversary to fool the SOS proof system. The adversary claims there are many solutions—even if in reality there are none—and offers to compute the expectation of any low-degree polynomial over the set of solutions. As long as (6) and (7) hold, this appears to be a distribution over valid solutions, and as long as (8) holds, the SOS prover cannot catch the adversary in an obvious lie like the claim that some quantity of degree $\delta/2$ has negative variance.

2.2. Colorings, partitions, and sum of squares. For a given graph G with adjacency matrix A, we can encode the problem of k-colorability as the following system of polynomial equations in kn variables $\boldsymbol{x} = \{x_{i,c}\}$, where $i \in [n]$ indexes vertices and $c \in [k]$ indexes colors:

$$\text{The } x_{i,c} \text{ are Boolean:} \quad p^{\text{bool}}_{i,c} \triangleq x_{i,c}^2 - x_{i,c} = 0 \qquad \forall i, c\, . \tag{9}$$

$$\text{Each vertex has one color:} \quad p^{\text{sing}}_{i} \triangleq -1 + \sum_c x_{i,c} = 0 \qquad \forall i\, . \tag{10}$$

$$\text{The coloring is proper:} \quad p^{\text{col}}_{ij} \triangleq \sum_c x_{i,c}\, x_{j,c} = 0 \qquad \forall (i,j) \in E\, . \tag{11}$$

Then G is k-colorable if and only if (9)–(11) has a solution in $\mathbb{R}^{kn}$. We can encode the stochastic block model similarly: Fix τ, and recall that a partition of G into k

groups is *good* if a fraction τ/k of the edges have endpoints in the same group. If G has m edges, we can replace constraint (11) with

$$\text{(12)} \qquad \text{Good partition:} \qquad p^{\text{cut}} \triangleq -\frac{\tau}{k} + \frac{1}{2m}\sum_{i,j} A_{ij} \sum_{c} x_{i,c}\, x_{j,c} = 0\,.$$

A degree-δ sum-of-squares refutation of (9)–(11) is an equation of the form

$$\text{(13)} \qquad \sum_{i,c} b_{i,c} p_{i,c}^{\text{bool}} + \sum_{i} s_i p_i^{\text{sing}} + \sum_{(i,j)\in E} g_{ij} p_{ij}^{\text{col}} = S + \epsilon\,,$$

where $b_{i,c}, s_i, g_{ij}$ are polynomials over $\boldsymbol{x}$, S is a sum of squares of polynomials, ϵ is a small positive constant which we will omit when clear, and the degree of each side is at most δ. Such an equation is a proof that no coloring exists. Replacing $\sum_{i,j} g_{ij} p_{ij}^{\text{col}}$ with $g_{\text{cut}} p^{\text{cut}}$ gives a refutation of the system formed by (9), (10), and (12), proving that no good partition exists. We focus on refutations of degree two, which as we will see are related to a classic relaxation of graph coloring.

2.3. The Lovász ϑ function. An *orthogonal representation* of a graph G with n vertices is an assignment of a unit vector $u_i \in \mathbb{R}^n$ to each vertex i such that $\langle u_i, u_j \rangle = 0$ for all $(i,j) \in E$. The Lovász function [45], denoted $\vartheta(\overline{G})$ by convention, is the smallest κ for which there is an orthogonal representation $\{u_i\}$ and an additional unit vector $\mathfrak{z} \in \mathbb{R}^n$ such that $\langle u_i, \mathfrak{z} \rangle = 1/\sqrt{\kappa}$, that is, such that all the u_i lie on a cone[2] of width $\cos^{-1}(1/\sqrt{\kappa})$.

The Gram matrix $P_{ij} = \langle u_i, u_j \rangle$ of an orthogonal representation is positive semidefinite with $P_{ii} = 1$ and $P_{ij} = 0$ for $(i,j) \in E$. Adding an auxiliary row and column for the inner products with $\mathfrak{z}$, we can define ϑ in terms of an SDP,

$$\text{(14)} \qquad \vartheta(\overline{G}) = \min_{P} \kappa > 0 \qquad \text{such that} \qquad \begin{pmatrix} 1 & \mathbf{1}/\sqrt{\kappa} \\ \mathbf{1}/\sqrt{\kappa} & P \end{pmatrix} \succeq 0\,, \quad P_{ii} = 1 \quad \forall i\,, \quad P_{ij} = 0 \quad \forall (i,j) \in E\,,$$

where $\mathbf{1}$ is the n-dimensional vector whose entries are all 1's. The dual of this program can be written [45] as

$$\text{(15)} \qquad \vartheta(\overline{G}) = \max_{D} \langle D, \mathbb{J} \rangle \qquad \text{such that} \qquad D \succeq 0\,, \quad \operatorname{tr} D = 1\,, \quad D_{ij} = 0 \quad \forall (i,j) \notin E\,,$$

where $\mathbb{J}$ is the matrix of all 1's and $\langle A, B \rangle = \operatorname{tr}(A^\dagger B) = \sum_{i,j} A_{ij} B_{ij}$ denotes the matrix inner product.

If G is k-colorable, then $\vartheta(\overline{G}) \le k$, since we can use the first k basis vectors $e_1, \ldots, e_k$ as an orthogonal representation and take $\mathfrak{z} = (1/\sqrt{k}) \sum_{t=1}^{k} e_t$. Thus if $\vartheta(\overline{G}) > k$, the Lovász function gives a polynomial-time refutation of k-colorability. As stated above, degree-two sum-of-squares proofs typically correspond to well-known semidefinite relaxations, and the next theorem shows that this is indeed the case here.

[2]To see that this definition of ϑ is equivalent to the more common one that $\langle u_i, \mathfrak{z} \rangle \le 1/\sqrt{\kappa}$ for every i, i.e., where the u_i can be in the interior of this cone, simply rotate each u_i in the subspace perpendicular to its neighbors until $\langle u_i, \mathfrak{z} \rangle$ is exactly $1/\sqrt{\kappa}$.

THEOREM 3. *There is a degree-2 SOS refutation of k-colorability for a graph G if and only if $\vartheta(\overline{G}) > k$.*

We prove this in the appendix, where we show that any orthogonal representation of G that lies on an appropriate cone lets us define a pseudoexpectation for the system (9)–(11). This will also allow us to modify the SDPs for refutations and pseudoexpectations, and work with simplified but equivalent versions.

2.4. Good partitions and a relaxed Lovász function. The reader may have noticed that while the coloring constraint (11) fixes the inner product $\sum_c x_{i,c}x_{j,c} = \langle x_i, x_j \rangle$ to zero for each edge $(i,j) \in E$, the "good partition" constraint (12) only fixes the sum of all these inner products. This suggests a slight relaxation of the Lovász ϑ function, where we weaken the SDP (14) by replacing the individual constraints on P_{ij} for all $(i,j) \in E$ with a constraint on their sum. In other words, we allow a vector coloring where neighboring vectors are orthogonal on average. We denote the resulting function $\hat{\vartheta}$:

$$\hat{\vartheta}(\overline{G}) = \min_P \kappa > 0 \qquad \text{such that} \qquad \begin{pmatrix} 1 & \mathbf{1}/\sqrt{\kappa} \\ \mathbf{1}/\sqrt{\kappa} & P \end{pmatrix} \succeq 0\,, \quad P_{ii} = 1 \quad \forall i\,, \quad \langle P, A \rangle = 0\,. \tag{16}$$

The dual SDP tightens (15) by requiring that the matrix D take the same value on every edge. Thus D is a multiple of A plus a diagonal matrix,

$$\hat{\vartheta}(\overline{G}) = \max_{\eta,\boldsymbol{b}} \langle D, \mathbb{J} \rangle \qquad \text{such that} \qquad D \triangleq \eta A + \operatorname{diag} \boldsymbol{b} \succeq 0\,, \quad \operatorname{tr} D = \langle \boldsymbol{b}, \mathbf{1} \rangle = 1\,. \tag{17}$$

Since $\hat{\vartheta}$ is a relaxation of ϑ, we always have $\hat{\vartheta}(\overline{G}) \le \vartheta(\overline{G})$.

This modified Lovász function $\hat{\vartheta}$ is equivalent to degree-two SOS for good partitions in the disassortative case of the block model in the following sense.

THEOREM 4. *If $\tau < 1$, there exists a degree-two SOS refutation of a partition of G where a fraction τ/k of the edges are within groups if and only if*

$$\hat{\vartheta}(\overline{G}) > \frac{k-\tau}{1-\tau}\,. \tag{18}$$

Once again we leave the proof to the appendix. Note that the SDP (16) for $\hat{\vartheta}$ contains no information about k or τ; this relaxed orthogonal representation has the uncanny capacity to fool degree-two SOS about an entire family of related cuts of different sizes and qualities.

2.5. Upper and lower bounds. With these theorems in hand, we can set about producing degree-two sum-of-squares refutations and pseudoexpectations for our problems; throughout this section we will refer to these simply as "refutations" and "pseudoexpectations." In fact, the same construction will give us refutations and pseudoexpectations for both the coloring and partition problems.

To warm up, we have the following simple construction of a refutation, which we will phrase in terms of the Lovász theta function and its relaxed version.

LEMMA 1. *Let G be a d-regular graph, and let $\lambda_{\min}$ be the smallest eigenvalue of its adjacency matrix A. Then*

$$\vartheta(\overline{G}) \ge \hat{\vartheta}(\overline{G}) \ge 1 + d/|\lambda_{\min}|\,. \tag{19}$$

Proof. Denoting by $\mathbb{1}$ the identity matrix, we can construct a feasible solution D to the dual SDP (17) by taking

$$D \triangleq \frac{1}{n}\left(\mathbb{1} + \frac{1}{|\lambda_{\min}|} A\right),$$

and we use the fact that $\langle A, \mathbb{J} \rangle = dn$. ∎

By invoking Friedman's theorem [29] that (as $n \to \infty$) the smallest eigenvalue of a random d-regular graph is with high probability larger than $-2(1+\epsilon)\sqrt{d-1}$ for any $\epsilon > 0$, we obtain the following.

COROLLARY 1. *When $G = G_{n,d}$, for any $\epsilon > 0$, with high probability*

$$\vartheta(\overline{G}) \geq \hat{\vartheta}(\overline{G}) > 1 + \frac{d}{2\sqrt{d-1}} - \epsilon \,. \tag{20}$$

Putting this together with Theorems 3 and 4 gives the following.

COROLLARY 2. *If $G = G_{n,d}$ and $\tau < 1$, with high probability there exists a refutation of a partition with a fraction τ/k of within-group edges when*

$$\frac{k-\tau}{1-\tau} < 1 + \frac{d}{2\sqrt{d-1}} \,. \tag{21}$$

Setting $\tau = 0$, a refutation of k-colorability exists with high probability when

$$k < 1 + \frac{d}{2\sqrt{d-1}} \,.$$

Note that for large k, the minimum value of d satisfying (21) is a factor of four above the Kesten–Stigum threshold in both the coloring and partition problems.

Our construction for this lower bound on ϑ is quite simple, but remarkably we find that for both the coloring and partition problems, it is asymptotically optimal in d and k. In particular, we have the following theorem.

THEOREM 5. *For any d-regular graph G with girth at least γ, we have*

$$\hat{\vartheta}(\overline{G}) \leq \vartheta(\overline{G}) < 1 + \frac{d}{2\sqrt{d-1} - \epsilon_\gamma} \,, \tag{22}$$

where ϵ_γ is a sequence of constants which decrease to zero as $\gamma \to \infty$.

Since for any constant γ a random regular graph has girth γ with positive probability [65, Theorem 2.12], we rely on the following result showing that $\vartheta(\overline{G_{n,d}})$ is concentrated in an interval of width one. The proof is essentially the same as that of [7] for the chromatic number, and is given in the appendix.

LEMMA 2. *Let $\theta \geq 3$. If $\vartheta(\overline{G_{n,d}}) \leq \theta$ with positive probability, then $\vartheta(\overline{G_{n,d}}) \leq \theta + 1$ with high probability.*

COROLLARY 3. *If $G = G_{n,d}$, with high probability there does not exist a degree-two refutation of a partition with a fraction τ/k of within-group edges when*

$$\frac{k-\tau}{1-\tau} > 2 + \frac{d}{2\sqrt{d-1}} \,. \tag{23}$$

Setting $\tau = 0$, with high probability no degree-two refutation of k-colorability exists when

$$k > 2 + \frac{d}{2\sqrt{d-1}} \,.$$

Thus for both problems, no degree-two sum-of-squares refutation exists until d is roughly a factor of 4 above the Kesten–Stigum threshold.

3. Constructing a pseudoexpectation with orthogonal polynomials. We now prove Theorem 5 by constructing a feasible solution to the primal SDP (14), that is, unit vectors $\{u_i\}$ such that $\langle u_i, u_j \rangle = 0$ for every edge (i,j), and a unit vector $\mathfrak{z}$ so that $\langle u_i, \mathfrak{z} \rangle = 1/\sqrt{\kappa}$ for all i. Recall that such a collection exists if and only if $\vartheta(\overline{G}) \le \kappa$.

It is convenient to instead define a set of unit vectors $\{v_i\}$ such that $\langle v_i, v_j \rangle = -1/(\kappa - 1)$ for every edge (i,j). We claim that such a set exists if and only if $\vartheta(\overline{G}) \le \kappa$. In one direction, given $\{u_i\}$ and $\mathfrak{z}$ with the above properties, if we define

$$v_i = \sqrt{\frac{\kappa}{\kappa - 1}}\, u_i - \frac{1}{\sqrt{\kappa - 1}}\, \mathfrak{z}\,,$$

then the v_i are unit vectors with $\langle v_i, v_j \rangle = -1/(\kappa - 1)$ for $(i,j) \in E$. For instance, if the u_i are k orthogonal basis vectors, then the v_i point to the corners of a k-simplex. In the other direction, given $\{v_i\}$ we can take $\mathfrak{z}$ to be a unit vector perpendicular to all the v_i and define

$$u_i = \sqrt{\frac{\kappa - 1}{\kappa}}\, v_i + \frac{1}{\sqrt{\kappa}}\, \mathfrak{z}\,.$$

Then $\langle u_i, u_j \rangle = 0$ for $(i,j) \in E$, and $\langle u_i, \mathfrak{z} \rangle = 1/\sqrt{\kappa}$ for all i. This means that we can characterize the Lovász ϑ function with a slightly different SDP, which uses the Gram matrix of the $\{v_i\}$:

$$\vartheta(\overline{G}) = \min_P \kappa > 1 \quad \text{such that} \quad \begin{aligned} &P \succeq 0\,, \\ &P_{ii} = 1 && \forall i\,, \\ &P_{ij} = -1/(\kappa - 1) && \forall (i,j) \in E\,. \end{aligned} \tag{24}$$

This SDP is not of our invention; it appears throughout the literature (e.g., [34]), often under the name "vector coloring," and is related to the classic SDP relaxation of MAX k-CUT [30].

We will show that for any d-regular graph G with girth at least γ, this SDP has a feasible solution with

$$\kappa = 1 + \frac{d}{2\sqrt{d-1} - \epsilon_\gamma}\,,$$

where ϵ_γ depends only on γ and tends to zero as $\gamma \to \infty$. Therefore, there is a pseudoexpectation that prevents degree-two SOS from refuting k-colorability for any $k \ge \kappa$. We will construct this pseudoexpectation by taking a linear combination of the "nonbacktracking powers" of G's adjacency matrix A. Our strategy is indebted to the proof in [44] that d-regular graphs of large girth have eigenvalues arbitrarily close to $-2\sqrt{d-1}$, and is in the same spirit as the "Gaussian wave" construction in [51].

Denote by $A^{(t)}$ the matrix whose i,j entry is the number of nonbacktracking walks of length t from i to j, that is, walks which may freely wander the graph so long as they do not make adjacent pairs of steps $a \to b \to a$ for any vertices a, b. There is a simple two-term recursion for these matrices: to count nonbacktracking walks of length t, we first extend each walk of length $t-1$ by one edge, and then subtract those that backtracked on the last pair of steps. There are d such pairs of steps that could be added to a path of length $t-2$, but if $t \geq 3$ one of these would fail to produce a nonbacktracking path of length $t-1$. This gives

$$\begin{aligned} A^{(0)} &= \mathbb{1}\,, \\ A^{(1)} &= A\,, \\ A^{(2)} &= A^2 - d\mathbb{1}\,, \\ A^{(t)} &= A \cdot A^{(t-1)} - (d-1)A^{(t-2)}, \quad t \geq 3\,. \end{aligned} \tag{25}$$

Thus we can always write $A^{(t)} = q_t(A)$ for some polynomials $q_t(z) \in \mathbb{R}[z]$ of degree t satisfying the same three-term recurrence as above. That is,

$$\begin{aligned} q_0(z) &= 1\,, \\ q_1(z) &= z\,, \\ q_2(z) &= z^2 - d\,, \\ q_t(z) &= z q_{t-1}(z) - (d-1) q_{t-2}(z)\,, \quad t \geq 3\,. \end{aligned} \tag{26}$$

The polynomials q_t are orthogonal with respect to the Kesten–McKay measure

$$\mu_d(z) = \frac{d}{2\pi} \frac{\sqrt{4(d-1)-z^2}}{d^2 - z^2}\, 1_{|z| \leq 2\sqrt{d-1}}\,. \tag{27}$$

That is, if we define the inner product

$$\langle f, g \rangle = \int f(z) g(z)\, \mu_d(z)\, \mathrm{d}z\,,$$

then $\langle q_s, q_t \rangle = 0$ if $s \neq t$ and, in our scaling,

$$\|q_t\|^2 = \langle q_t, q_t \rangle = \begin{cases} 1\,, & t = 0\,, \\ d\,, & t = 1\,, \\ d(d-1)^{t-1}\,, & t > 1\,. \end{cases} \tag{28}$$

These facts are implicit in [48] and remarked on without proof in [11]. The relevant calculations can be found, for instance, in [44] or [62, Proposition 2.3]. The reader should be aware that these references often orthonormalize the q_t or rescale the support of the measure μ_d to $[-1,+1]$ when convenient to their calculations. Although it is incidental to our proof, we note that for $t \geq 1$ we can write q_t explicitly as

$$q_t = (d-1)^{t/2}\, U_t\left(\frac{z}{2\sqrt{d-1}}\right) - (d-1)^{t/2-1}\, U_{t-2}\left(\frac{z}{2\sqrt{d-1}}\right)\,, \tag{29}$$

where U_t is the tth Chebyshev polynomial of the second kind,

$$U_t(\cos\theta) = \frac{\sin((t+1)\theta)}{\sin\theta}\,, \tag{30}$$

which is a polynomial in $\cos\theta$ of order t.

If the girth of the graph is at least γ, then whenever $\ell + m \leq \gamma - 1$ it is impossible for a pair of vertices to be connected by nonbacktracking walks of lengths ℓ and m simultaneously, and thus the matrices $A^{(\ell)}$ and $A^{(m)}$ have disjoint sets of nonzero entries. In particular, a nonbacktracking path of length $\gamma - 2$ or less cannot return to its starting point or to a neighbor of its starting point. We can therefore satisfy the diagonal and edge constraints of (24) by considering linear combinations of $A^{(t)}$ up to $t = \gamma - 2$, or equivalently polynomials in A of degree $\gamma - 2$, whose 0th and 1st coefficients ensure that P has 1's on its diagonal and $-1/(\kappa - 1)$ on the edges of G,

$$P = \mathbb{1} - \frac{1}{\kappa - 1} A + \sum_{t=2}^{\gamma-2} a_t A^{(t)}$$

$$= \sum_{t=0}^{\gamma-2} a_t \, q_t(A) \ , \qquad \text{where } a_0 = 1 \text{ and } a_1 = \frac{-1}{\kappa - 1} \,, \tag{31}$$

$$\triangleq f(A) \ , \qquad f \in \mathbb{R}[z], \ \deg f \leq \gamma - 2 \,. \tag{32}$$

Our job is to optimize the coefficients a_t for $1 \leq t \leq \gamma - 2$ so as to minimize a_1, and hence κ, while ensuring that $P \succeq 0$.

The eigenvalues of the matrix $f(A)$ are of the form $f(\lambda)$, where λ ranges over the spectrum of A. Thus, to ensure that our construction $P = f(A)$ is positive semidefinite for every d-regular graph, it is sufficient to require

$$f(z) \geq 0 \qquad \text{for all } \ |z| \leq d \,. \tag{33}$$

We will see that the optimal choice of f is in fact positive on all of $\mathbb{R}$. By orthogonality and (28), we can write the coefficients a_t as inner products,

$$a_t = \frac{\langle q_t, f \rangle}{\|q_t\|^2} \,.$$

Collecting everything, construction of our pseudoexpectation reduces to the problem

$$\begin{aligned} \min \quad & \langle q_1, f \rangle \\ \text{such that} \quad & \langle q_0, f \rangle = 1 \,, \\ & f(z) \geq 0 \qquad \forall z \in [-d, d] \,, \\ & \deg f \leq \gamma - 2 \,. \end{aligned} \tag{34}$$

All that is now required to prove Theorem 5 is a clever choice of a feasible polynomial f with $\langle q_1, f \rangle = (-2\sqrt{d-1} + \epsilon_\gamma)$. In fact, when the degree $\gamma - 2$ of f is even, we can obtain such a polynomial by solving (34) explicitly. Set $m = \gamma/2$, and let $r_1 > \cdots > r_m$ be the roots of q_m in decreasing order; it follows from standard arguments about orthogonal polynomials that these are all simple roots contained in the support of μ_d, i.e., in the interval $(-2\sqrt{d-1}, 2\sqrt{d-1})$, and that the leftmost root r_m tends to $-2\sqrt{d-1}$ as $m \to \infty$ [64, Theorems 3.3.1 and 6.1.1]. Consider the following polynomial of degree $2(m-1) = \gamma - 2$,

$$s(z) = \frac{1}{\zeta} \prod_{j=1}^{m-1} (z - r_j)^2 \,, \tag{35}$$

where

$$\zeta = \left\langle q_0, \prod_{j=1}^{m-1} (z - r_j)^2 \right\rangle$$

is a normalizing factor to ensure that $\langle q_0, s \rangle = 1$. We claim that $s(z)$ is the optimum of (34). To prove this, we begin with a general lemma on orthogonal polynomials and quadrature. The proof is standard (e.g., [64, Theorem 3.4.1-2]) but we include it in the appendix for completeness.

LEMMA 3. *Let $\{p_t\}$ be a sequence of polynomials of degree t which are orthogonal with respect to a measure ρ supported on a compact interval I. Then the roots $r_1, \ldots, r_t$ of p_t form a quadrature rule which is exact for any polynomial u of degree less than $2t$, in that*

$$\int_I u(z)\,\mathrm{d}\rho = \sum_{i=1}^{t} \omega_i\, u(r_i)$$

for some positive weights $\{\omega_1, \ldots, \omega_t\}$ independent of u.

Now let $g(z) = z - r_m$. In view of Lemma 3, for any polynomial $f(z)$ of degree at most $\gamma - 2$, the inner product $\langle g, f\rangle$ can be expressed using the roots $r_1, \ldots, r_m$ of q_m as a quadrature,

$$\langle g, f\rangle = \int (z - r_m) f(z)\,\mathrm{d}\mu = \sum_{j=1}^{m} \omega_j\,(r_j - r_m) f(r_j) = \sum_{j=1}^{m-1} \omega_j\,(r_j - r_m) f(r_j)\,.$$

Note that $\omega_j(r_j - r_m) > 0$ for every $1 \le j \le m-1$, since r_m is the leftmost root. If we impose the constraints that $f(r_j) \ge 0$ for all $j = 1, \ldots, m-1$, then $\langle g, f\rangle \ge 0$. If we also impose the constraint $\langle f, q_0\rangle = 1$, then

$$\begin{aligned} \langle q_1, f\rangle &= \langle z, f\rangle \\ &= \langle g, f\rangle + r_m\langle q_0, f\rangle \\ &\ge r_m\,, \end{aligned} \tag{36}$$

with equality if and only if $f(r_j) = 0$ for all $j = 1, \ldots, m-1$. Since $s(z)$ obeys this equality condition, we have

$$\langle q_1, s\rangle = r_m\,,$$

and this is the minimum possible value of $\langle q_1, s\rangle$ subject to the constraints that $\langle q_0, f\rangle = 1$ and $f(r_j) \ge 0$ for $j = 1, \ldots, m-1$. Moreover, $s(z) \ge 0$ on all of $\mathbb{R}$, so $s(z)$ in fact obeys the stronger constraint (33). This completes the proof that $s(z)$ is the optimum of (34).

Referring back to (32) gives

$$\frac{-1}{\kappa - 1} = c_1 = \frac{\langle q_1, s\rangle}{\|q_1\|^2} = \frac{r_m}{d}\,,$$

and so

$$\vartheta \ge \kappa = 1 + \frac{d}{-r_m}\,.$$

Finally, we obtain (22) by defining $\epsilon_\gamma = r_m + 2\sqrt{d-1}$ and recalling that $r_m \to -2\sqrt{d-1}$ as $m \to \infty$. This completes the proof of Theorem 5.

4. Discussion. We close with two observations on the construction of Theorem 5. First, in the limit of large d the Kesten–McKay measure (27) approaches the semicircle law

$$\mu(z) = \frac{\sqrt{4(d-1) - z^2}}{2\pi(d-1)}\, 1_{|z| \le 2\sqrt{d-1}}\,, \tag{37}$$

and (29) shows that q_m approaches a scaled Chebyshev polynomial,

$$q_m = (d-1)^{m/2}\, U_m\!\left(\frac{z}{2\sqrt{d-1}}\right). \tag{38}$$

From (30) the roots r_j then become

$$\frac{r_j}{2\sqrt{d-1}} = \cos\left(\frac{j\pi}{m+1}\right)$$

for $1 \le j \le m$, and indeed

$$\frac{r_m}{2\sqrt{d-1}} = \cos\frac{m\pi}{m+1} = -1 + O(1/m^2)\,.$$

Now recall the constraint $\langle q_0, s\rangle = 1$ in (34), or equivalently $\int s(z)\mu_d(z)\,\mathrm{d}z = 1$. Along with $s(z) \ge 0$, this implies that $s(z)\mu_d(z)$ is a probability distribution on $[-2\sqrt{d-1}, 2\sqrt{d-1}]$. The inner product $\langle q_1, s\rangle = \int s(z)\mu_d(z)\,z\,\mathrm{d}z$ is the expectation of z under this distribution. We can minimize $\langle q_1, s\rangle$ by moving as much mass as possible to the left, and as $m \to \infty$ this distribution approaches a delta function at the left end of the support:

$$s(z)\mu_d(z) \to \delta\Big(z - 2\sqrt{d-1}\Big)\,,$$

where δ is the Dirac delta function. (We are being deliberately vague as to what sort of convergence $\to$ represents here.) The coefficients a_t then approach

$$a_t = \frac{\langle q_t, s\rangle}{\|q_t\|^2} = \frac{1}{\|q_t\|^2}\int q_t(z)s(z)\mu_d(z)\,\mathrm{d}z \to \frac{1}{\|q_t\|^2}\,q_t\Big(-2\sqrt{d-1}\Big)\,.$$

Using (28), (29), and the identity $U_t(-1) = (-1)^t(t+1)$, this gives for $t \ge 1$

$$\begin{aligned} a_t &\to \frac{(-1)^t}{d(d-1)^{t-1}}\Big((d-1)^{t/2}(t+1) - (d-1)^{t/2-1}(t-1)\Big) \\ &= \frac{(-1)^t}{(d-1)^{t/2}}\left((t+1) - \frac{2t}{d}\right)\,. \end{aligned}$$

Fixing t and taking d to be large, this becomes

$$a_t \to (d-1)^{-t/2}\,U_t(-1) = (-1)^t(t+1)(d-1)^{-t/2}\,, \tag{39}$$

which we could also obtain from (38).

Now recall the (pseudo)probabilistic meaning of these coefficients. They describe the optimal solution P to the SDP (24) as a sum over nonbacktracking paths of length t,

$$P = \sum_{t=0}^{\gamma} a_t A^{(t)}.$$

As discussed in the appendix, this matrix can be translated into a degree-two pseudo-expectation $\tilde{\mathbb{E}}$ for the coloring problem: a linear operator that claims to give the joint distribution of colors at each pair of vertices i and j. The reader will find there that P_{ij} is related to the "pseudocorrelation" between vertices i and j, by

$$\frac{k}{k-1}\,(P_{ij} - 1/k) = \widetilde{\Pr}[i \text{ and } j \text{ are the same color}]\,.$$

Our expansion of P in terms of nonbacktracking paths means that if i and j are sufficiently close or sufficiently far, this pseudoexpectation depends only on the shortest path distance $d(i,j)$ between them. Specifically, whenever $d(i,j) = t < \gamma/2$ the shortest path is unique, and we have $P_{ij} = a_t$, while if $d(i,j) > \gamma - 2$, then $P_{ij} = 0$ since $s(z)$ is a polynomial of degree $2(m-1) = \gamma - 2$.

One might think that in the limit of large γ, the optimal pseudoexpectation would behave as if these shortest paths were colored uniformly at random, ignoring correlations with the remainder of the graph. An easy calculation shows that that would give

$$a_t = (1-k)^{-t} \,. \tag{40}$$

However, the optimal coefficients behave rather differently, especially near the refutation threshold: setting $d - 1 \approx 4d_{\mathrm{KS}} = 4(k-1)^2$ in (39) gives

$$a_t \approx (t+1)2^{-t}(1-k)^{-t} \,,$$

so the pseudocorrelations decay roughly 2^{-t} faster than random colorings of a path suggest. This suggests that the best strategy for an adversary to fool the SOS proof system is in fact not to act as if the path between two vertices is colored in a uniformly random way.

Appendix A. Proof of Theorems 3 and 4. We prove Theorems 3 and 4 by directly simplifying the SDP that defines feasible degree-two pseudoexpectations. The first step is a broad result on the structure of these objects that applies to any set of constraints which includes the Boolean (9) and single-color (10) constraints and is suitably symmetric; we then specialize to the coloring and partition problems.

Recall that a degree-two pseudoexpectation for a system of polynomials $f_j(\boldsymbol{x}) = 0$ is a linear operator $\tilde{\mathbb{E}} : \mathbb{R}[\boldsymbol{x}]_{\le 2} \to \mathbb{R}$ which satisfies

- $\tilde{\mathbb{E}}[1] = 1$,
- $\tilde{\mathbb{E}}[f_j q] = 0$ for any polynomials f_j and q such that $\deg f_j q \le 2$,
- $\tilde{\mathbb{E}}[p^2] \ge 0$ for any polynomial p with $\deg p^2 \le 2$.

We can identify such objects with positive semidefinite $(nk+1) \times (nk+1)$ matrices of the form

$$\tilde{\mathbb{E}} = \begin{pmatrix} 1 & \boldsymbol{\ell}^\dagger \\ \boldsymbol{\ell} & \mathcal{E} \end{pmatrix}, \tag{41}$$

where $\ell_{i,c} = \tilde{\mathbb{E}}[x_{i,c}]$ and $\mathcal{E}_{(i,c),(j,c')} = \tilde{\mathbb{E}}[x_{i,c}\, x_{j,c'}]$. It is useful to think of $\mathcal{E}$ as a block matrix, with a $k \times k$ block $\mathcal{E}_{ij}$ corresponding to each pair of vertices i,j. Consistency with the Boolean and single-color constraints (9), (10) then controls the diagonal elements and row and column sums of each of these blocks,

$$\mathcal{E}_{(i,c),(i,c)} = \tilde{\mathbb{E}}[x_{i,c}^2] = \tilde{\mathbb{E}}[x_{i,c}] = \ell_{i,c} \qquad \forall i, c \,, \tag{42}$$

$$\sum_{c'} \mathcal{E}_{(i,c),(j,c')} = \sum_{c'} \tilde{\mathbb{E}}[x_{i,c}\, x_{j,c'}] = \tilde{\mathbb{E}}[x_{j,c}] = \ell_{i,c} \qquad \forall i, j \,. \tag{43}$$

Moreover, each of our constraints is fixed under permutations of the colors, and $\tilde{\mathbb{E}}$ inherits this symmetry. That is, the matrix carries with it a natural S_k action that simultaneously permutes $\tilde{\mathbb{E}}[x_{i,c}] \to \tilde{\mathbb{E}}[x_{i,\sigma(c)}]$ and $\tilde{\mathbb{E}}[x_{i,c}\, x_{j,c'}] \to \tilde{\mathbb{E}}[x_{i,\sigma(c)}\, x_{i,\sigma(c')}]$. This action preserves the spectrum of $\tilde{\mathbb{E}}$ as a matrix, as well as every hard constraint. By convexity, we may assume that $\tilde{\mathbb{E}}$ is stabilized under it by beginning with an arbitrary pseudoexpectation and averaging over its orbit.

This assumption substantially constrains and simplifies $\tilde{\mathbb{E}}$. In particular we are free to (i) assume that $\ell_{i,c} = \tilde{\mathbb{E}}[x_{i,c}] = 1/k$, and (ii) assume that each $k \times k$ block in $\mathcal{E}$ has only two distinct values: one on the diagonal and the other off the diagonal. In other words, the pseudoexpectation claims that the marginal distribution of each vertex is uniform, and that joint marginal of any two vertices depends only on the probability that they have the same or different colors. As a result, for each i, j we can assume that $\mathcal{E}_{ij}$ is a linear combination of the identity matrix $\mathbb{1}_k$ and the matrix $\mathbb{J}_k$ of all 1's, and that the row and column sums of $\mathcal{E}_{ij}$ are all $1/k$. In that case for each i, j we can write

$$\mathcal{E}_{ij} = \frac{1}{k-1}\left(P_{ij} - \frac{1}{k}\right)\left(\mathbb{1}_k - \frac{\mathbb{J}_k}{k}\right) + \frac{\mathbb{J}_k}{k^2} \tag{44}$$

for some P_{ij}, or equivalently that

$$\mathcal{E} = \frac{1}{k-1}(P - \mathbb{J}_n/k) \otimes \left(\mathbb{1}_k - \frac{\mathbb{J}_k}{k}\right) + \frac{\mathbb{J}_{nk}}{k^2} \tag{45}$$

for some $n \times n$ matrix P. Note that

$$\operatorname{tr} \mathcal{E}_{ij} = P_{ij} \,,$$

so (42) requires that $P_{ii} = 1$ for all i.

Since the pseudoexpectation (41) consists of $\mathcal{E}$ with an additional row and column, we consider the following lemma. We leave its proof as an exercise for the reader.

LEMMA 4. *For any matrix X, vector $\boldsymbol{v}$, and scalar $b > 0$,*

$$\begin{pmatrix} b & \boldsymbol{v}^\dagger \\ \boldsymbol{v} & X \end{pmatrix} \succeq 0$$

if and only if $X - (1/b)\boldsymbol{v} \otimes \boldsymbol{v} \succeq 0$.

Since $\boldsymbol{\ell}$ is the nk-dimensional vector whose entries are all $1/k$, we have $\boldsymbol{\ell} \otimes \boldsymbol{\ell} = \mathbb{J}_{nk}/k^2$. Thus (45) and Lemma 4 imply that $\tilde{\mathbb{E}} \succeq 0$ if and only if

$$(P - \mathbb{J}_n/k) \otimes (\mathbb{1}_k - \mathbb{J}_k/k) \succeq 0\,.$$

Since $\mathbb{1}_k - \mathbb{J}_k/k$ is a projection operator, this in turn occurs if and only if

$$P - \mathbb{J}_n/k \succeq 0\,.$$

To summarize, finding a pseudoexpectation is equivalent to finding a positive semidefinite matrix $P \in \mathbb{R}^{n \times n}$ with $P_{ii} = 1$ for all i, such that P remains positive semidefinite when we subtract the rank-one matrix $\mathbb{J}_n/k$. However, we have thus far only reasoned about the Boolean and single-color constraints, and including either the coloring or cut constraint places an additional restriction on P. In the case of coloring, we demanded that

$$\sum_c \mathcal{E}_{(i,c),(j,c)} = \sum_c \tilde{\mathbb{E}}[x_{i,c}\, x_{j,c}] = 0 \tag{46}$$

for every edge (i, j). This implies that $\operatorname{tr} \mathcal{E}_{ij} = 0$, and so $P_{ij} = 0$ for each edge. Collecting these observations, a pseudoexpectation for coloring exists exactly when

$k > \vartheta(\overline{G})$, where

$$\vartheta(\overline{G}) \triangleq \min_P \kappa > 0 \quad \text{such that} \quad P - \mathbb{J}_n/\kappa \succeq 0\,, \quad P_{ii} = 1 \quad \forall i\,, \quad P_{ij} = 0 \quad \forall (i,j) \in E\,. \tag{47}$$

Finally, note that $\mathbb{J}_n/\kappa = v \otimes v$, where $v = \mathbf{1}_n/\sqrt{\kappa}$. Applying Lemma 4 again then gives exactly the positive semidefinite (14) for the Lovász ϑ function, thus completing the proof of Theorem 3.

In the case of good partitions, we required that

$$\sum_{i,j} A_{ij} \sum_c \mathcal{E}_{(i,c),(j,c)} = \sum_{i,j} A_{ij} \sum_c \tilde{\mathbb{E}}[x_{i,c}\, x_{j,c}] = (\tau/k)dn\,, \tag{48}$$

but this means that

$$\sum_{i,j} A_{ij} \operatorname{tr} \mathcal{E}_{ij} = \sum_{i,j} A_{ij} P_{ij} = \langle P, A\rangle = (\tau/k)dn\,.$$

Following the path above, a degree-two pseudoexpectation exists for community detection when $k > \hat{\vartheta}_\tau(\overline{G})$, where

$$\hat{\vartheta}_\tau(\overline{G}) \triangleq \min_{P_\tau} \kappa_\tau \quad \text{such that} \quad P_\tau - \mathbb{J}_n/\kappa_\tau \succeq 0\,, \quad (P_\tau)_{ii} = 1 \quad \forall i\,, \quad \langle P_\tau, A\rangle = (\tau/\kappa_\tau)dn\,. \tag{49}$$

Note that $\kappa_\tau \geq 1$, since in order for $P_\tau - \mathbb{J}_n/\kappa_\tau \succeq 0$ it cannot have negative entries on its diagonal. A priori, it seems that we may need to solve a different SDP for each value of τ, but a bit more work shows that this is not the case. Lemma 4 lets us transform the SDP (16) for $\hat{\vartheta}$ to the following problem:

$$\hat{\vartheta}(\overline{G}) \triangleq \min_P \kappa \quad \text{such that} \quad P - \mathbb{J}_n/\kappa \succeq 0\,, \quad P_{ii} = 1 \quad \forall i\,, \quad \langle P, A\rangle = 0\,. \tag{50}$$

The following lemma then shows us how to relate optima of (50) to those of (49) for any τ in the disassortative range $\tau < 1$, thus completing the proof of Theorem 4.

LEMMA 5. *For any $\tau < 1$,*

$$\hat{\vartheta}(\overline{G}) = \frac{\hat{\vartheta}_\tau(\overline{G}) - \tau}{1 - \tau}\,. \tag{51}$$

Proof. We show how to translate back and forth between solutions of (49) and (50). Given a matrix P, define

$$P_\tau = (1 - \tau/\kappa_\tau)P + (\tau/\kappa_\tau)\mathbb{J}_n\,.$$

It is easy to check that $P_{ii} = 1$ if and only if $(P_\tau)_{ii} = 1$, and $\langle P_\tau, A\rangle = (\tau/\kappa_\tau)dn$ if and only if $\langle P, A\rangle = 0$. Finally, if we set

$$\kappa = \frac{\kappa_\tau - \tau}{1 - \tau}\,, \tag{52}$$

then

$$P_\tau - \mathbb{J}_n/\kappa_\tau = (1-\tau/\kappa_\tau)\,(P - \mathbb{J}_n/\kappa)\ ,$$

so, using $\tau < 1$ and $\kappa_\tau \geq 1$, we see that $P_\tau - \mathbb{J}_n/\kappa_\tau \succeq 0$ if and only if $P - \mathbb{J}_n/\kappa \succeq 0$. Thus (50) is feasible for κ if and only if (49) is feasible for κ_τ. Since $\hat{\vartheta}(\overline{G})$ and $\hat{\vartheta}_\tau(\overline{G})$ are the smallest κ and κ_τ respectively, for which this is the case, (52) implies (51). ◻

Appendix B. Proof of Lemma 3. It is immediate that there is such a quadrature rule for polynomials of degree strictly less than t, since the space of linear functionals on such polynomials has dimension t and is thus spanned by the t linearly independent functionals which evaluate at the roots x_i. Now let $\deg u < 2t$. We can divide u by p_t to write $u(z) = a(z)p_t + b(z)$, where $\deg a, \deg b < t$. We have

$$\int_I u(z)\,\mathrm{d}\rho = \int_I \big(a(z)p_t(z)+b(z)\big)\,\mathrm{d}\rho = \langle p_t, a\rangle + \int_I b(z)\,\mathrm{d}\rho = 0 + \sum_{i=1}^t \omega_i b(r_i) = \sum_{i=1}^t \omega_i u(r_i),$$

since p_t is orthogonal to all polynomials of degree less than t and has roots r_i. This verifies exactness of the quadrature rule for polynomials of degree smaller than $2t$.

To show that the weights $\{\omega_i\}$ are positive, let $i \in \{1,\ldots,t\}$ and let $v_i(z) = (p_t(z)/(z-r_i))^2$ be the polynomial with double roots at every root of p_t save r_i. Since v_i is everywhere nonnegative and is a polynomial of degree $2t-2 < t$, we have

$$0 < \int_I v_i(z)\,\mathrm{d}\rho = \sum_{j=1}^t \omega_j v_i(r_j) = \omega_i v_i(r_i)\,,$$

but since $v(z)$ is nonnegative, ω_i must be positive.

Appendix C. Proof of Lemma 2. The proof closely follows [7, Theorem 4], which shows that the chromatic number of $G_{n,d}$ is concentrated on two adjacent integers, and which is in turn based on the proof in [46] of two-point concentration for $G(n,p)$ with $p = O(n^{-5/6-\epsilon})$. Recall the configuration model [65], where we make d "copies" of each vertex corresponding to its half-edges, and then choose uniformly from all

$$(dn-1)(dn-3)(dn-5)\cdots = \frac{(dn)!}{2^{dn/2}(dn/2)!}$$

perfect matchings of these copies. If we denote the set of such matchings by $\mathcal{P}_{n,d}$ and condition the corresponding multigraphs on having no self-loops or multiple edges, the resulting distribution is uniform on the set of d-regular graphs and occupies a constant fraction of the total probability of $\mathcal{P}_{n,d}$. Thus any property which holds with high probability for $\mathcal{P}_{n,d}$ holds with high probability for $G_{n,d}$ as well.

If P, P' are two perfect matchings in $\mathcal{P}_{n,d}$, we write $P \sim P'$ if they differ by a single swap, changing $\{(a,b),(c,d)\}$ to $\{(a,c),(b,d)\}$ or $\{(a,d),(b,c)\}$. The following martingale inequality [65, Theorem 2.19] shows that a random variable which is Lipschitz with respect to these swaps is concentrated.

LEMMA 6. *Let c be a constant, and let X be a random variable defined on $\mathcal{P}_{n,d}$ such that $|X(P) - X(P')| \leq c$ whenever $P \sim P'$. Then*

$$\Pr[|X - \mathbb{E}[X]| > t] \leq 2\mathrm{e}^{-\frac{t^2}{dnc}}\,.$$

Now fix θ and define X as the minimum number of edge constraints $P_{i,j} = 0$ in the SDP (14) violated by an otherwise feasible solution with $\kappa = \theta$. Given any such

P, if we perform an edge swap on the underlying graph, at most two new constraints are violated, so X meets the Lipschitz condition with $c = 2$. By assumption $X = 0$ with positive probability. Lemma 6 then implies that (say) $\mathbb{E}[X] \le (1/2)\sqrt{n \log n}$, in which case $X < \sqrt{n \log n}$ with high probability.

Let S denote the set of endpoints of the violated edges. Then there is an orthogonal representation $\{u_i\}$ of the subgraph induced by $V \setminus S$ and a unit vector $\mathfrak{z}$ such that $\langle u_i, \mathfrak{z} \rangle = 1/\sqrt{\theta}$ and $\langle u_i, u_j \rangle = 0$ if $(i,j) \in E$ and $i, j \notin S$. Our goal is to "fix" $\{u_i\}$ on the violated edges, and if necessary on some additional vertices, to give an orthogonal representation $\{v_i\}$ for all of G.

As in [7, 46], we inductively build a set of vertices $S = U_0, U_1, \ldots, U_T = U$ as follows. Given U_t, let $U_{t+1} = U_t \cup \{i, j\}$, where $i, j \notin U_t$, $(i,j) \in E$, and i and j each have at least one neighbor in U_t. We define T as the step at which there is no such pair i, j and this process ends. Let I denote U's neighborhood, i.e., the set of vertices outside U which have a neighbor in U. Then I is an independent set, since otherwise the process would have continued. We make the following claim.

LEMMA 7. *With high probability, the subgraph induced by U is 3-colorable.*

Proof. For all $0 \le t \le T$ we have $|U_t| = 2t + |S|$. Moreover, the subgraph induced by U_t has at least $3t + |S|/2 = (3/2)|U_t| - |S|$ edges and thus average degree at least $3 - 2|S|/|U_t|$. On the other hand, a crude union bound shows that for any d and any $\beta > 2$, there is an $\alpha > 0$ such that, with high probability, all induced subgraphs of G containing αn or fewer vertices have average degree less than β. Since $|S| = o(n)$ with high probability, this implies that $|U_t| \le (2 + o(1))|S|$ for all t, and in particular that $|U| = o(n)$.

The same union bound then implies that with high probability the subgraph induced by $|U|$, and all its subgraphs, have average degree less than 3. But this means that this subgraph has no 3-core; that is, it has at least one vertex of degree less than 3, and so will the subgraph we get by deleting this vertex, and so on. Working backwards, we can 3-color the entire subgraph by starting with the empty set and adding these vertices back in, since at least one of the three colors will always be available to them. ∎

To define our orthogonal representation, let w be a unit vector such that $\langle \mathfrak{z}, w \rangle = \langle u_i, w \rangle = 0$ for all $i \notin S$; such a vector exists since $|S| \ge 2$. Then define

$$\mathfrak{z}' = \sqrt{\frac{\theta}{\theta+1}}\, \mathfrak{z} + \frac{1}{\sqrt{\theta+1}}\, w\,.$$

Then $|\mathfrak{z}'|^2 = 1$, and $\langle w, \mathfrak{z}' \rangle = \langle u_i, \mathfrak{z}' \rangle = 1/\sqrt{\theta+1}$ for all $i \notin S$. Moreover, there exist three mutually orthogonal unit vectors y_1, y_2, y_3 such that $\langle y_j, \mathfrak{z}' \rangle = 1/\sqrt{\theta+1}$ and $\langle y_j, w \rangle = 0$ for all $j \in \{1,2,3\}$. This follows from the fact that the following matrix is PSD whenever $\theta \ge 3$, in which case it can be realized as the Gram matrix of $\{y_1, y_2, y_3, w, \mathfrak{z}'\}$:

$$\begin{pmatrix} 1 & 0 & 0 & 0 & \frac{1}{\sqrt{\theta+1}} \\ 0 & 1 & 0 & 0 & \frac{1}{\sqrt{\theta+1}} \\ 0 & 0 & 1 & 0 & \frac{1}{\sqrt{\theta+1}} \\ 0 & 0 & 0 & 1 & \frac{1}{\sqrt{\theta+1}} \\ \frac{1}{\sqrt{\theta+1}} & \frac{1}{\sqrt{\theta+1}} & \frac{1}{\sqrt{\theta+1}} & \frac{1}{\sqrt{\theta+1}} & 1 \end{pmatrix}.$$

Finally, let $\sigma(i) \in \{1,2,3\}$ be a proper 3-coloring of the subgraph induced by U. Then the following is an orthogonal representation of G:

$$v_i = \begin{cases} u_i\,, & i \in V \setminus (U \cup I)\,, \\ w\,, & i \in I\,, \\ y_{\sigma(i)}\,, & i \in U\,, \end{cases}$$

and $\langle v_i, \mathfrak{z}' \rangle = 1/\sqrt{\theta+1}$ for all i. This gives a feasible solution to the SDP (14) with $\kappa = \theta + 1$, implying that $\vartheta(\overline{G}) \leq \theta + 1$.

Acknowledgments. We are grateful to Charles Bordenave, Emmanuel Abbe, Amin Coja-Oghlan, Yash Deshpande, Marc Lelarge, and Alex Russell for helpful conversations. Part of this work was done while C.M. was visiting École Normale Supérieure. Part of this work was done while R.K. was a researcher at Microsoft Research New England.

REFERENCES

[1] E. Abbe, *Community detection and stochastic block models: Recent developments*, J. Mach. Learn. Res., 18 (2018), 177.

[2] E. Abbe, A. S. Bandeira, and G. Hall, *Exact recovery in the stochastic block model*, IEEE Trans. Inform. Theory, 62 (2016), pp. 471–487.

[3] E. Abbe and C. Sandon, *Community detection in general stochastic block models: Fundamental limits and efficient algorithms for recovery*, in Proc. 56th Annual Symposium on Foundations of Computer Science, FOCS, 2015, pp. 670–688.

[4] E. Abbe and C. Sandon, *Detection in the Stochastic Block Model with Multiple Clusters: Proof of the Achievability Conjectures, Acyclic BP, and the Information-Computation Gap*, preprint, https://arxiv.org/abs/1512.09080, 2015.

[5] E. Abbe and C. Sandon, *Achieving the KS threshold in the general stochastic block model with linearized acyclic belief propagation*, in Proc. Neural Information Processing Systems, NIPS, 2016, pp. 1334–1342.

[6] E. Abbe and C. Sandon, *Crossing the KS threshold in the stochastic block model with information theory*, in IEEE International Symposium on Information Theory, ISIT, 2016, pp. 840–844.

[7] D. Achlioptas and C. Moore, *The chromatic number of random regular graphs*, in Proc. 8th International Workshop on Randomization and Computation, RANDOM, 2004, pp. 219–228.

[8] D. Achlioptas and A. Naor, *The two possible values of the chromatic number of a random graph*, Ann. of Math. (2), 162 (2005), pp. 1335–1351.

[9] N. Agarwal, A. S. Bandeira, K. Koiliaris, and A. Kolla, *Multisection in the Stochastic Block Model Using Semidefinite Programming*, Springer, 2017, pp. 125–162.

[10] S. R. Allen, R. O'Donnell, and D. Witmer, *How to refute a random CSP*, in IEEE 56th Annual Symposium on Foundations of Computer Science, FOCS, 2015, pp. 689–708.

[11] N. Alon, I. Benjamini, E. Lubetzky, and S. Sodin, *Non-backtracking random walks mix faster*, Commun. Contemp. Math., 9 (2007), pp. 585–603.

[12] J. Banks, C. Moore, J. Neeman, and P. Netrapalli, *Information-theoretic thresholds for community detection in sparse networks*, in Proc. 29th Conference on Learning Theory, COLT, 2016, pp. 383–416.

[13] B. Barak, S. B. Hopkins, J. Kelner, P. Kothari, A. Moitra, and A. Potechin, *A nearly tight sum-of-squares lower bound for the planted clique problem*, in Proc. 57th Annual IEEE Symposium on Foundations of Computer Science, FOCS, 2016, pp. 428–437.

[14] B. Barak and D. Steurer, *Sum-of-squares proofs and the quest toward optimal algorithms*, Electronic Colloquium on Computational Complexity (ECCC), 21 (2014), 59.

[15] P. Barucca, *Spectral Partitioning in Random Regular Blockmodels*, preprint, https://arxiv.org/abs/1610.02668, 2016.

[16] P. J. Bickel and A. Chen, *A nonparametric view of network models and Newman–Girvan and other modularities*, Proc. Natl. Acad. Sci. USA, 106 (2009), pp. 21068–21073.

[17] C. Bordenave, M. Lelarge, and L. Massoulié, *Non-backtracking spectrum of random graphs: Community detection and non-regular Ramanujan graphs*, in Proc. 56th Annual Symposium on Foundations of Computer Science, FOCS, 2015, pp. 1347–1357.

[18] G. Brito, I. Dumitriu, S. Ganguly, C. Hoffman, and L. V. Tran, *Recovery and rigidity in a regular stochastic block model*, in Proceedings of the Twenty-Seventh Annual ACM-SIAM Symposium on Discrete Algorithms, SODA, 2016, pp. 1589–1601, https://doi.org/10.1137/1.9781611974331.ch108.
[19] A. Coja-Oghlan, *The Lovász number of random graphs*, Combin. Probab. Comput., 14 (2005), pp. 439–465.
[20] A. Coja-Oghlan, C. Efthymiou, and S. Hetterich, *On the chromatic number of random regular graphs*, J. Combin. Theory Ser. B, 116 (2016), pp. 367–439.
[21] A. Coja-Oghlan, C. Efthymiou, N. Jaafari, M. Kang, and T. Kapetanopoulos, *Charting the replica symmetric phase*, Comm. Math. Phys., 359 (2018), pp. 603–698.
[22] A. Coja-Oghlan, F. Krzakala, W. Perkins, and L. Zdeborová, *Information-theoretic thresholds from the cavity method*, Adv. Math., 333 (2018), pp. 694–795.
[23] A. Coja-Oghlan, E. Mossel, and D. Vilenchik, *A spectral approach to analysing belief propagation for* 3*-colouring*, Combin. Probab. Comput., 18 (2009), pp. 881–912.
[24] A. Coja-Oghlan and D. Vilenchik, *Chasing the K-colorability threshold*, in 54th Annual IEEE Symposium on Foundations of Computer Science, FOCS, 2013, pp. 380–389.
[25] A. Decelle, F. Krzakala, C. Moore, and L. Zdeborová, *Asymptotic analysis of the stochastic block model for modular networks and its algorithmic applications*, Phys. Rev. E, 84 (2011), 066106.
[26] A. Decelle, F. Krzakala, C. Moore, and L. Zdeborová, *Inference and phase transitions in the detection of modules in sparse networks*, Phys. Rev. Lett., 107 (2011), 065701.
[27] Y. Deshpande and A. Montanari, *Improved sum-of-squares lower bounds for hidden clique and hidden submatrix problems*, in Proc. 28th Conference on Learning Theory, COLT, 2015, pp. 523–562.
[28] P. Erdős, *Some remarks on the theory of graphs*, Bull. Amer. Math. Soc., 53 (1947), pp. 292–294.
[29] J. Friedman, *A Proof of Alon's Second Eigenvalue Conjecture and Related Problems*, American Mathematical Society, 2008.
[30] A. Frieze and M. Jerrum, *Improved approximation algorithms for MAX k-CUT and MAX BISECTION*, Algorithmica, 18 (1997), pp. 67–81.
[31] B. Hajek, Y. Wu, and J. Xu, *Achieving exact cluster recovery threshold via semidefinite programming*, IEEE Trans. Inform. Theory, 62 (2016), pp. 2788–2797.
[32] B. Hajek, Y. Wu, and J. Xu, *Achieving exact cluster recovery threshold via semidefinite programming: Extensions*, IEEE Trans. Inform. Theory, 62 (2016), pp. 5918–5937.
[33] S. B. Hopkins, P. Kothari, A. H. Potechin, P. Raghavendra, and T. Schramm, *On the integrality gap of degree-*4 *sum of squares for planted clique*, in Proc. of the 27th Annual ACM-SIAM Symposium on Discrete Algorithms, SODA, 2016, pp. 1079–1095, https://doi.org/10.1137/1.9781611974331.ch76.
[34] D. Karger, R. Motwani, and M. Sudan, *Approximate graph coloring by semidefinite programming*, J. ACM, 45 (1998), pp. 246–265.
[35] G. Kemkes, X. Pérez-Giménez, and N. Wormald, *On the chromatic number of random d-regular graphs*, Adv. Math., 223 (2010), pp. 300–328.
[36] H. Kesten and B. P. Stigum, *Additional limit theorems for indecomposable multidimensional Galton–Watson processes*, Ann. Math. Statist., 37 (1966), pp. 1463–1481.
[37] H. Kesten and B. P. Stigum, *Limit theorems for decomposable multi-dimensional Galton–Watson processes*, J. Math. Anal. Appl., 17 (1967), pp. 309–338.
[38] P. K. Kothari, R. Mori, R. O'Donnell, and D. Witmer, *Sum of squares lower bounds for refuting any CSP*, in Proc. 49th Annual ACM Symposium on Theory of Computing, STOC, 2017, pp. 132–145.
[39] J. L. Krivine, *Anneaux préordonnés*, J. Anal. Math., 12 (1964), pp. 307–326.
[40] F. Krzakala, A. Montanari, F. Ricci-Tersenghi, G. Semerjian, and L. Zdeborová, *Gibbs states and the set of solutions of random constraint satisfaction problems*, Proc. Natl. Acad. Sci. USA, 104 (2007), pp. 10318–10323.
[41] F. Krzakala, C. Moore, E. Mossel, J. Neeman, A. Sly, L. Zdeborová, and P. Zhang, *Spectral redemption in clustering sparse networks*, Proc. Natl. Acad. Sci. USA, 110 (2013), pp. 20935–20940.
[42] F. Krzakala and L. Zdeborová, *Hiding quiet solutions in random constraint satisfaction problems*, Phys. Rev. Lett., 102 (2009), 238701.
[43] J. B. Lasserre, *Global optimization with polynomials and the problem of moments*, SIAM J. Optim., 11 (2001), pp. 796–817, https://doi.org/10.1137/S1052623400366802.
[44] W.-C. W. Li and P. Solé, *Spectra of regular graphs and hypergraphs and orthogonal polynomials*, European J. Combin., 17 (1996), pp. 461–477.

[45] L. Lovász, *On the Shannon capacity of a graph*, IEEE Trans. Inform. Theory, 25 (1979), pp. 1–7.

[46] T. Łuczak, *A note on the sharp concentration of the chromatic number of random graphs*, Combinatorica, 11 (1991), pp. 295–297.

[47] L. Massoulié, *Community detection thresholds and the weak Ramanujan property*, in Proc. 46th Annual ACM Symposium on Theory of Computing, STOC, 2014, pp. 694–703.

[48] B. D. McKay, *The expected eigenvalue distribution of a large regular graph*, Linear Algebra Appl., 40 (1981), pp. 203–216.

[49] R. Meka, A. Potechin, and A. Wigderson, *Sum-of-squares lower bounds for planted clique*, in Proc. 47th Annual ACM on Symposium on Theory of Computing, STOC, 2015, pp. 87–96.

[50] M. Molloy and B. A. Reed, *The Chromatic Number of Sparse Random Graphs*, Master's thesis, University of Waterloo, 1992.

[51] A. Montanari and S. Sen, *Semidefinite programs on sparse random graphs and their application to community detection*, in Proceedings of the Forty-Eighth Annual ACM Symposium on Theory of Computing, STOC '16, ACM, 2016, pp. 814–827.

[52] C. Moore, *The computer science and physics of community detection: Landscapes, phase transitions, and hardness*, Bull. Eur. Assoc. Theor. Comput. Sci. EATCS, 121 (2017), pp. 25–61.

[53] E. Mossel, J. Neeman, and A. Sly, *Belief propagation, robust reconstruction and optimal recovery of block models*, in Proc. 27th Conference on Learning Theory, COLT, 2014, pp. 356–370.

[54] E. Mossel, J. Neeman, and A. Sly, *Consistency thresholds for the planted bisection model*, in Proc. Forty-Seventh Annual ACM on Symposium on Theory of Computing, STOC, 2015, pp. 69–75.

[55] E. Mossel, J. Neeman, and A. Sly, *Reconstruction and estimation in the planted partition model*, Probab. Theory Related Fields, 162 (2015), pp. 431–461.

[56] E. Mossel, J. Neeman, and A. Sly, *A proof of the block model threshold conjecture*, Combinatorica, 38 (2018), pp. 665–708.

[57] Y. Nesterov, *Squared functional systems and optimization problems*, High Performance Optimization, 33 (2000), pp. 405–440.

[58] M. E. J. Newman and T. Martin, *Equitable random graphs*, Phys. Rev. E, 90 (2014), 052824.

[59] P. A. Parrilo, *Structured Semidefinite Programs and Semialgebraic Geometry Methods in Robustness and Optimization*, Ph.D. thesis, California Institute of Technology, 2000.

[60] G. Schoenebeck, *Linear level Lasserre lower bounds for certain k-CSPs*, in Proc. 49th Annual IEEE Symposium on Foundations of Computer Science, FOCS, 2008, pp. 593–602.

[61] N. Z. Shor, *An approach to obtaining global extremums in polynomial mathematical programming problems*, Cybernetics, 23 (1987), pp. 695–700.

[62] S. Sodin, *Random matrices, nonbacktracking walks, and orthogonal polynomials*, J. Math. Phys., 48 (2007), 123503.

[63] G. Stengle, *A nullstellensatz and a positivstellensatz in semialgebraic geometry*, Math. Ann., 207 (1974), pp. 87–97.

[64] G. Szegő, *Orthogonal Polynomials*, American Mathematical Society, 1939.

[65] N. Wormald, *Models of random regular graphs*, in Surveys in Combinatorics, Cambridge University Press, 1999, pp. 239–298.