LSE Research Online

THE LONDON SCHOOL
OF ECONOMICS AND
POLITICAL SCIENCE ■

Ewan Davies, Matthew Jenssen, Will Perkins, Barnaby Roberts

Independent sets, matchings, and occupancy fractions

**Article (Accepted version)
(Refereed)**

http://eprints.lse.ac.uk

INDEPENDENT SETS, MATCHINGS, AND OCCUPANCY FRACTIONS

EWAN DAVIES, MATTHEW JENSSEN, WILL PERKINS, AND BARNABY ROBERTS

ABSTRACT. We prove tight upper bounds on the logarithmic derivative of the independence and matching polynomials of d-regular graphs. For independent sets, this theorem is a strengthening of Kahn's result that a disjoint union of copies of $K_{d,d}$ maximizes the number of independent sets of a bipartite d-regular graph, Galvin and Tetali's result that the independence polynomial is maximized by the same, and Zhao's extension of both results to all d-regular graphs.

For matchings, this shows that the matching polynomial and the total number of matchings of a d-regular graph are maximized by a union of copies of $K_{d,d}$. Using this we prove the asymptotic upper matching conjecture of Friedland, Krop, Lundow, and Markström.

In probabilistic language, our main theorems state that for all d-regular graphs and all λ, the occupancy fraction of the hard-core model and the edge occupancy fraction of the monomer-dimer model with fugacity λ are maximized by $K_{d,d}$. Our method involves constrained optimization problems over distributions of random variables and applies to all d-regular graphs directly, without a reduction to the bipartite case.

Using a variant of the method we prove a lower bound on the occupancy fraction of the hard-core model on any d-regular, vertex-transitive, bipartite graph: the occupancy fraction of such a graph is strictly greater than the occupancy fraction of the unique translation-invariant hard-core measure on the infinite d-regular tree.

1. INDEPENDENT SETS

Let G be a graph. The independence polynomial of G is

$$P_G(\lambda) = \sum_{I \in \mathcal{I}} \lambda^{|I|}$$

where $\mathcal{I}$ is the set of all independent sets of G. By convention we consider the empty independent set to be a member of $\mathcal{I}$. The *hard-core model* with fugacity λ on G is a random independent set I drawn according to the distribution

$$\Pr_\lambda[I] = \frac{\lambda^{|I|}}{P_G(\lambda)}.$$

$P_G(\lambda)$ is also called the partition function of the hard-core model on G.

In the hard-core model, the quantity $\alpha_G(\lambda) = \frac{1}{|V(G)|} \frac{\lambda P_G'(\lambda)}{P_G(\lambda)}$ is the *occupancy fraction*: the expected fraction of vertices of G belonging to the random independent set I. In particular,

$$\alpha_G(\lambda) = \frac{1}{|V(G)|} \sum_{v \in G} \Pr[v \in I] = \frac{1}{|V(G)|} \frac{\sum_{I \in \mathcal{I}} |I| \lambda^{|I|}}{P_G(\lambda)} = \frac{1}{|V(G)|} \frac{\lambda P_G'(\lambda)}{P_G(\lambda)} = \left(\frac{1}{|V(G)|} \log P_G(\lambda) \right)'.$$

We write $K_{d,d}$ for the complete bipartite graph with d vertices in each part. If $2d$ divides n, let $H_{d,n}$ denote the d-regular, n-vertex graph that is the disjoint union of $n/(2d)$ copies of $K_{d,d}$. Kahn [18] showed that $H_{d,n}$ maximizes the total number of independent sets over all d-regular, n-vertex bipartite graphs, and then showed [19] that in fact $K_{d,d}$ (or $H_{d,n}$) maximizes $\frac{1}{|V(G)|}\log P_G(\lambda)$ for $\lambda \geq 1$ over all d-regular bipartite graphs. The log partition function result generalizes the counting result as the latter can be recovered by setting $\lambda = 1$. Galvin and Tetali [14] then gave a broad generalization of Kahn's result to counting homomorphisms from a d-regular, bipartite G to any graph H. The case of H formed of two connected vertices, one with a self-loop, is that of counting independent sets. Via a modification of H and a limiting argument, they proved that in fact $\frac{1}{|V(G)|}\log P_G(\lambda)$ is maximized for any $\lambda > 0$ over d-regular bipartite graphs by $K_{d,d}$. Zhao [31] then removed the bipartite restriction in these results for independent sets by reducing the general case to the bipartite case, in particular proving that $H_{d,n}$ has the greatest number of independent sets of any d-regular graph on n vertices.

Here we prove a strengthening of the above results for independent sets.

Theorem 1. *For all d-regular graphs G and all $\lambda > 0$, we have*

$$\alpha_G(\lambda) \leq \alpha_{K_{d,d}}(\lambda) = \frac{\lambda(1+\lambda)^{d-1}}{2(1+\lambda)^d - 1}.$$

The maximum is achieved only by unions of copies of $K_{d,d}$.

In particular Theorem 1 states that the derivative of $\frac{1}{|V(G)|}\log P_G(\lambda)$ is maximized over d-regular graphs for all λ by $K_{d,d}$, which when integrated, immediately implies that the normalized log partition function, $\frac{1}{|V(G)|}\log P_G(\lambda)$, is maximized. Even more, it says that the difference $\frac{1}{2d}\log P_{K_{d,d}}(\lambda) - \frac{1}{|V(G)|}\log P_G(\lambda)$ is strictly increasing in λ for any d-regular graph G that is not $H_{d,n}$. Note that $\frac{1}{n}\log P_{H_{d,n}}(\lambda) = \frac{1}{2d}\log P_{K_{d,d}}(\lambda)$ for any n divisible by $2d$.

In Section 8 we observe that the above bound on the partition function gives new upper bounds on a related problem: maximizing the number of independent sets of a given size in d-regular graphs.

Next, let $\alpha_{T_d}(\lambda)$ be occupancy fraction of the unique translation invariant hard-core measure on the infinite d-regular tree T_d at fugacity λ; that is, $\alpha_{T_d}(\lambda)$ is the solution of the equation

$$\frac{\alpha}{\lambda(1-\alpha)} = \left(\frac{1-2\alpha}{1-\alpha}\right)^d$$

(see e.g. [4]).

Using a variant of the method used to establish Theorem 1, we prove a lower bound on the occupancy fraction in any d-regular, vertex-transitive, bipartite graph G.

Theorem 2. *For any d-regular, vertex-transitive, bipartite graph G,*

$$\alpha_G(\lambda) > \alpha_{T_d}(\lambda).$$

The corresponding statement for the normalized log partition function (the integrated version of Theorem 2) holds without the condition of vertex transitivity [25]. Theorem 2 itself may not hold without vertex transitivity (see Section 5 of [9] for a related discussion about matchings). For $\lambda \leq \lambda_c(T_d) = \frac{(d-1)^{d-1}}{(d-2)^d}$ (the uniqueness threshold of the hard-core

model on T_d), the bound in Theorem 2 is asymptotically tight for this class of graphs. From the results of Weitz [30], any sequence of graphs G_n that converges locally (in the sense of Benjamini-Schramm [3]) to T_d has occupancy fraction $\alpha_{T_d}(\lambda) + o(1)$ as $n \to \infty$; for example, we can take a sequence of bipartite Cayley graphs of large girth.

2. MATCHINGS

The matching polynomial of a graph G is
$$M_G(\lambda) = \sum_{H \in \mathcal{M}} \lambda^{|H|}$$
where $\mathcal{M}$ is the set of all matchings of G (including the empty matching) and $|H|$ is the number of edges in the matching H. Just as in the hard-core model above we can define a probability distribution over matchings:
$$\Pr_{\lambda}[H] = \frac{\lambda^{|H|}}{M_G(\lambda)}.$$

This defines the monomer-dimer model from statistical physics [15]: dimers are edges of the random matching H and monomers the unmatched vertices.

The *edge occupancy fraction*, or the dimer density, is the expected fraction of the edges of G in such a random matching:
$$\alpha_G^M(\lambda) = \frac{1}{|E(G)|} \sum_{e \in G} \Pr[e \in H] = \frac{1}{|E(G)|} \frac{\lambda M_G'(\lambda)}{M_G(\lambda)}.$$

Our next result is an upper bound on the edge occupancy fraction of any d-regular graph:

Theorem 3. *For all d-regular graphs G and all $\lambda > 0$, we have*
$$\alpha_G^M(\lambda) \le \alpha_{K_{d,d}}^M(\lambda).$$
The maximum is achieved only by unions of copies of $K_{d,d}$.

This states that the normalized logarithmic derivative of $M_G(\lambda)$ is maximized by $K_{d,d}$, and hence via integration that $K_{d,d}$ (and thus also $H_{d,n}$) maximizes $\frac{1}{|E(G)|} \log M_G(\lambda)$ for any $\lambda > 0$. This resolves Conjecture 7.1 in [13]. Brégman's theorem [6] says that the number of perfect matchings of a d-regular, n-vertex bipartite graph is maximized by $H_{d,n}$, and this was extended by Kahn and Lovász to all d-regular graphs (see [13] for a full discussion). Our result on $M_G(\lambda)$ extends this: letting $\lambda \to \infty$ recovers the result for perfect matchings, while setting $\lambda = 1$ shows that $H_{d,n}$ maximizes the total number of matchings of any d-regular graph on n vertices.

In Section 8 we use Theorem 3 to give new upper bounds on the number of matchings of a given size in d-regular graphs. We then use these bounds to prove the 'asymptotic upper matching conjecture' of Friedland, Krop, Lundow, and Markström [11].

3. RELATED WORK

The results of Kahn [18, 19], Galvin and Tetali [14], and Zhao [31] culminating in the fact that $\frac{1}{|V(G)|} \log P_G(\lambda)$ is maximized over d-regular graphs by $K_{d,d}$ are based on the entropy method, a powerful tool for the type of problems we address here. Apart from the results

mentioned above, see [24] and [13] for surveys of the method. A direct application of the method requires the graph G to be bipartite. Zhao [32] showed that in some, but not all cases, this restriction can be removed by using a 'bipartite swapping trick'. An entropy-free proof of Galvin and Tetali's general theorem on counting homomorphisms was recently given by Lubetzky and Zhao [22]. Our method also does not use entropy, but in contrast to the other proofs it works directly for all d-regular graphs, without a reduction to the bipartite case. The method deals directly with the hard-core model instead of counting homomorphisms and seems to require more problem-specific information than the entropy method; a question for future work is to extend the method to a more general class of homomorphisms.

The technique of writing the expected size of an independent set in two ways (as we do here) was used by Shearer [27] in proving lower bounds on the average size of an independent set in K_r-free graphs and then by Alon [1] for graphs in which all vertex neighborhoods are r-colorable. The idea of bounding the occupancy fraction instead of the partition function comes in part from work of the third author [23] in improving, at low densities, the bounds on matchings of a given size in Ilinca and Kahn [16] and independent sets of a given size in Carroll, Galvin, and Tetali [7]. The use of linear programming for counting graph homomorphisms appears in Kopparty and Rossman [20], where they use a combination of entropy and linear programming to compute a related quantity, the homomorphism domination exponent, in chordal and series-parallel graphs.

For matchings, Carroll, Galvin, and Tetali [7] used the entropy method to give an upper bound of $\frac{1}{2} \log(1 + d\lambda)$ on $\frac{1}{|V(G)|} \log M_G(\lambda)$ over d-regular graphs. It was previously conjectured (eg. [12, 13]) that $K_{d,d}$ maximizes $\frac{1}{|V(G)|} \log M_G(\lambda)$ over all d-regular graphs. This is an implication of our Theorem 3.

In [8], Csikvári proved the 'lower matching conjecture' of [12] and in [9] gave a new lower bound on the number of perfect matchings of d-regular, vertex-transitive, bipartite graphs, in both comparing an arbitrary graph with the infinite d-regular tree (see also the recent extension by Lelarge [21] to irregular graphs). Proposition 2.10 in [9] states that the edge occupancy fraction of any d-regular, vertex-transitive, bipartite graph is at least that of the infinite d-regular tree; in Theorem 2 we prove an analogous result for independent sets. Csikvári's techniques in the two papers are different than the methods of this paper, but similar in that he bounds the occupancy fraction instead of directly working with the partition function. His results rely on an elegant interplay between the Heilman-Lieb theorem [15] and Benjamini-Schramm convergence of bounded-degree graphs.

In statistical physics, the analogue of the occupancy fraction in a general spin system is called the *mean magnetization*; on general graphs it is $\#P$-hard to compute the magnetization in the ferromagnetic Ising model, the monomer-dimer model, and the hard-core model [28, 26].

4. The Method

To introduce our method, we start by proving Theorem 1 under the assumption that G is triangle-free. In what follows I will denote the random independent set drawn according to the hard-core model with fugacity λ on a d-regular, n-vertex graph G.

We say a vertex v is *occupied* if $v \in I$ and *uncovered* if none of its neighbors are in I: $N(v) \cap I = \emptyset$. Let p_v be the probability v is occupied and q_v be the probability v is uncovered. The idea of considering q_v appears in Kahn's paper [18].

We will show that for every $\lambda > 0$ and any triangle-free G, $\alpha_G(\lambda)$ is maximized by $K_{d,d}$. (It is easy to see by linearity of expectation or by manipulating the partition function that the occupancy fraction is the same for any number of copies of $K_{d,d}$).

Letting $\alpha = \alpha_G(\lambda)$, we write

$$\alpha = \frac{1}{n}\sum_{v \in G} p_v$$

$$(1) \qquad = \frac{1}{n}\sum_{v \in G}\frac{\lambda}{1+\lambda}q_v$$

$$(2) \qquad = \frac{\lambda}{1+\lambda}\cdot\frac{1}{n}\sum_{v \in G}\sum_{j=0}^{d}\Pr[j \text{ neighbors of } v \text{ are uncovered}]\cdot(1+\lambda)^{-j}$$

$$= \frac{\lambda}{1+\lambda}\cdot\mathbb{E}[(1+\lambda)^{-Y}]$$

where Y is the random variable that counts the number of uncovered neighbors of a uniformly chosen vertex from G, with respect to the random independent set I. Y is an integer valued random variable bounded between 0 and d. Equation (1) follows since v must be uncovered if it is to be occupied, and conditioning on being uncovered v is occupied with probability $\frac{\lambda}{1+\lambda}$. Equation (2) is similar: conditioned on the event that $u_1, \ldots, u_j$, neighbors of v, are all uncovered, the probability that none are occupied is $(1+\lambda)^{-j}$. This is where we use the fact that G is triangle-free: there are no edges between neighbors of v.

We also have

$$\mathbb{E}Y = \frac{1}{n}\sum_{v \in G}\sum_{u \sim v}q_u = d\cdot\frac{1+\lambda}{\lambda}\alpha$$

since each u appears in the double sum exactly d times as G is d-regular. This gives the identity

$$\mathbb{E}Y = d\cdot\mathbb{E}[(1+\lambda)^{-Y}].$$

Now let

$$\alpha^* = \frac{\lambda}{d(1+\lambda)}\cdot\sup_{0 \le Y \le d}\{\mathbb{E}Y : \mathbb{E}Y = d\cdot\mathbb{E}[(1+\lambda)^{-Y}]\}$$

where the sup is over all distributions of random variables Y bounded between 0 and d.

For any λ and d there is a unique distribution Y supported only on 0 and d that satisfies the constraint $\mathbb{E}Y = d\cdot\mathbb{E}[(1+\lambda)^{-Y}]$. We claim that the sup is uniquely achieved by this distribution. The claim follows from convexity, but we defer details to the proof of a more general statement in Section 5. Since the distribution Y associated to $H_{d,n}$ satisfies the constraint and is supported on 0 and d, it must maximize α. Since unions of copies of $K_{d,d}$ are the only graphs whose associated distribution is supported on 0 and d, they uniquely achieve the maximum.

To recap, the method is the following:

(i) Define a random variable Y using randomness in the hard-core model on G and in choosing a random vertex of G. In the proof above, Y was the number of uncovered neighbors of a random vertex.

(ii) Write α as the expectation of a function of Y.

(iii) Add constraints that the random variable Y must satisfy for any graph G in our class. In the case above, the constraints were that the two ways of writing α are equal and that $0 \leq Y \leq d$.

(iv) Relax the optimization problem from random variables Y induced by graphs to all random variables Y that satisfy the constraints. Show that the unique maximizer of α is the distribution associated to the extremal graph, and therefore α is maximized by the extremal graph.

In Section 5 we give the full proof of Theorem 1. We prove the lower bound, Theorem 2, in Section 6. We turn to matchings and Theorem 3 in Section 7 before giving new bounds on the number of independent sets and matchings of a given size in Section 8.

5. Proof of Theorem 1

For a vertex $v \in G$ and an independent set I, we define the *free neighborhood* of v to be the subgraph of G induced by the neighbors of v which are not adjacent to any vertex in $I \setminus N(v)$. We use the convention $v \notin N(v)$. The vertices in the free neighborhood may be uncovered or covered, but if they are covered it must be from another vertex in the free neighborhood. In a triangle-free graph the free neighborhood is always a set (possibly empty) of isolated vertices. Note that if $v \in I$, then the free neighborhood of v is necessarily empty.

Let C be the random free neighborhood of v when we draw I according to the hard-core model and choose vertex v uniformly at random from G. For any graph F, let p_F be the probability that C is isomorphic to F. Also let $P_C = P_C(\lambda)$ be the independence polynomial of C at fugacity λ. Then we can write α in two ways:

$$(3) \qquad \alpha = \frac{\lambda}{1+\lambda} \mathbb{E}\left[\frac{1}{P_C(\lambda)}\right]$$

and

$$(4) \qquad \alpha = \frac{\lambda}{d} \mathbb{E}\left[\frac{P_C'(\lambda)}{P_C(\lambda)}\right]$$

where in both equations the expectations are over the random free neighborhood C. Equation (3) follows since v itself is uncovered if and only if all vertices in its free neighborhood are unoccupied. Given that the free neighborhood is isomorphic to C, all vertices in the free neighborhood are unoccupied with probability $\frac{1}{P_C(\lambda)}$. Equation (4) follows by counting the expected number of occupied neighbors of v and dividing by d: only vertices in the free neighborhood can be occupied, and, given C, the expected number of occupied vertices in the free neighborhood is $\frac{\lambda P_C'(\lambda)}{P_C(\lambda)}$.

Now let

$$(5) \qquad \alpha^* = \frac{\lambda}{1+\lambda} \cdot \sup\left\{ \mathbb{E}\left[\frac{1}{P_C(\lambda)}\right] : \frac{d}{1+\lambda} \cdot \mathbb{E}\left[\frac{1}{P_C(\lambda)}\right] = \mathbb{E}\left[\frac{P_C'(\lambda)}{P_C(\lambda)}\right]\right\}$$

where the sup is over all distributions of random free neighborhoods C supported on graphs of at most d vertices. From (3) and (4), the distribution obtained from G satisfies the constraint above.

We claim that for any $\lambda > 0$, α^* is achieved uniquely by a distribution supported only on the empty graph and the graph consisting of d isolated vertices, $\overline{K_d}$. The theorem follows since disjoint unions of copies of $K_{d,d}$ are the only graphs for which the free neighborhood

can only be the empty set or $\overline{K_d}$, and since there is a unique distribution with this support satisfying the constraint. To prove this claim we use the language of linear programming, see e.g. [5].

The linear program. Let p_C be the probability of a given free neighborhood C, and let $\mathcal{C}_d$ be the set of all graphs on at most d vertices, including the empty graph. Equation (5) defines a linear program with the decision variables $\{p_C\}_{C \in \mathcal{C}_d}$. We write the linear program in standard form as

$$\alpha^* = \max \frac{\lambda}{2(1+\lambda)} \sum_{C \in \mathcal{C}_d} p_C(a_C + b_C) \text{ subject to}$$

$$\sum_{C \in \mathcal{C}_d} p_C = 1$$

$$\sum_{C \in \mathcal{C}_d} p_C(a_C - b_C) = 0$$

$$p_C \geq 0 \ \ \forall C \in \mathcal{C}_d$$

where $a_C = \frac{1}{P_C(\lambda)}$ and $b_C = \frac{(1+\lambda)P'_C(\lambda)}{dP_C(\lambda)}$. We can calculate $a_\emptyset = 1$, $b_\emptyset = 0$, $a_{\overline{K_d}} = (1+\lambda)^{-d}$, $b_{\overline{K_d}} = 1$. The solution $p_\emptyset = \frac{1-(1+\lambda)^{-d}}{2-(1+\lambda)^{-d}}$ and $p_{\overline{K_d}} = \frac{1}{2-(1+\lambda)^{-d}}$ is the unique feasible solution supported only on $\emptyset$ and $\overline{K_d}$, and gives the objective value $\frac{\lambda(1+\lambda)^{d-1}}{2(1+\lambda)^d-1}$. Our claim is that this is the unique maximum.

The dual linear program is

$$\alpha^* = \min \frac{\lambda}{2(1+\lambda)} \Lambda_1 \text{ s.t.}$$

$$\Lambda_1 + \Lambda_2(a_C - b_C) \geq a_C + b_C \ \ \forall C \in \mathcal{C}_d$$

where Λ_1, Λ_2 are the decision variables.

Guided by the candidate solution above we set $\Lambda_1 = \frac{2}{2-(1+\lambda)^{-d}}$ and $\Lambda_2 = 1 - \Lambda_1$. With these values, the dual constraints corresponding to $C = \emptyset, \overline{K_d}$ hold with equality, and the objective value is $\frac{\lambda}{2(1+\lambda)}\Lambda_1 = \frac{\lambda(1+\lambda)^{d-1}}{2(1+\lambda)^d-1}$. To finish the proof we claim that Λ_1, Λ_2 are feasible for the dual program, which means showing that

$$\Lambda_1 + \Lambda_2(a_C - b_C) \geq a_C + b_C$$

for all $C \in \mathcal{C}_d$. We will show that in fact the inequality holds strictly for all $C \in \mathcal{C}_d \setminus \{\emptyset, \overline{K_d}\}$. Substituting our values of Λ_1, Λ_2, this inequality reduces to

$$(6) \qquad \frac{\lambda P'_C(\lambda)}{P_C(\lambda) - 1} < \frac{\lambda d(1+\lambda)^{d-1}}{(1+\lambda)^d - 1}.$$

The LHS of (6) is the expected size of the random independent set from the hard-core model on C conditioned on it being non-empty. The RHS is the same quantity for $\overline{K_d}$.

Inequality (6) follows directly from the observation that, over all $C \in \mathcal{C}_d$, the graph $\overline{K_d}$ maximizes the ratio of subsequent terms in the polynomial P_C. Let $t_i = \binom{d}{i}$, the coefficient of λ^i in $P_{\overline{K_d}}$, and write $P_C = 1 + \sum_{i=1}^{d} r_i \lambda^i$. We have $(i+1)t_{i+1} = (d-i)t_i$ and $(i+1)r_{i+1} \leq (d-i)r_i$ by counting independent sets of size $i+1$.

To verify (6) we show that for each $1 \leq k \leq d$ the coefficient c_k of λ^k in the polynomial $(\lambda P'_{\overline{K}_d})(P_C - 1) - (\lambda P'_C)(P_{\overline{K}_d} - 1)$ is non-negative. We have

$$
\begin{aligned}
s_k &= \sum_{i=1}^{k-1} i t_i r_{k-i} - \sum_{i=1}^{k-1} i t_{k-i} r_i \\
&= \sum_{i=1}^{\lfloor k/2 \rfloor} (k - 2i)(t_{k-i} r_i - t_i r_{k-i}).
\end{aligned}
$$

Observe that term-by-term the above sum giving s_k is non-negative by comparing the ratio of successive coefficients in $P_{\overline{K}_d}$ and P_C. Furthermore, if $P_C \neq P_{\overline{K}_d}$ then at least one s_k must be positive, which completes the claim.

To see the optimizer is unique note that strict inequality in the dual constraints corresponding to configurations besides $\emptyset$ and $\overline{K}_d$ implies by complementary slackness that any optimal solution is supported on these two configurations, and there is a unique such distribution.

6. Proof of Theorem 2

To prove Theorem 2 we will use the fact that occupancies of vertices on the same side of a bipartite graph are positively correlated:

Lemma 1. *Let G be a bipartite graph with bipartition $\mathcal{E} \cup \mathcal{O}$. For any $r \geq 2$, let $u_1, u_2, \ldots, u_r \in \mathcal{E}$. Then*

$$
\Pr[\{u_1, \ldots, u_r\} \subseteq I] \geq \prod_{i=1}^{r} p_{u_i}
$$

in the hard-core model for any λ. Similarly, let U be the random set of uncovered vertices of G. Then

$$
\Pr[\{u_1, \ldots, u_r\} \subseteq U] \geq \prod_{i=1}^{r} q_{u_i}
$$

Moreover, the inequalities are strict when $\lambda > 0$ and at least two of the u_i's are in the same connected component of G.

The first part of the lemma follows by induction on r from the fact that $\Pr[u_1, u_2 \in I] > \Pr[u_1 \in I] \cdot \Pr[u_2 \in I]$ when u_1, u_2 are in the same connected component and in the same part of the bipartition of G. In [29] this is shown to be a consequence of the FKG inequality; see also [10] and Corollary 1.5 of [2]. An intuitive reason for this fact (which can be turned into a rigorous argument using Weitz's tree [30]), is that conditioning on the event that a vertex v is occupied forbids its neighbors from being in the independent set; conditioning on the event that v is not occupied increases the probability each of its neighbors are occupied, and these effects propagate through the bipartite graph.

To prove the second part of the lemma, note that $p_{u_i} = \frac{\lambda}{1+\lambda} q_{u_i}$, and for $u_1, \ldots, u_r \in \mathcal{E}$, $\Pr[\{u_1, \ldots, u_r\} \subseteq I] = (\frac{\lambda}{1+\lambda})^r \Pr[\{u_1, \ldots, u_r\} \subseteq U]$, since there are no edges between the u_i's. Then the desired inequality follows from the first part of the lemma.

Proof of Theorem 2. By vertex transitivity, for all v, $p_v = \alpha$ and $q_v = \frac{1+\lambda}{\lambda}\alpha$. Fix a vertex v and let Y be the number of uncovered neighbors of v. For $u \sim v$ let Y_u be the indicator that

u is uncovered.

$$\alpha = \frac{\lambda}{1+\lambda}\mathbb{E}[(1+\lambda)^{-Y}]$$

$$= \frac{\lambda}{1+\lambda}\mathbb{E}[(1+\lambda)^{-\sum_{u\sim v}Y_u}]$$

$$= \frac{\lambda}{1+\lambda}\left(\alpha + (1-\alpha)\mathbb{E}[(1+\lambda)^{-\sum_{u\sim v}Y_u}|v\notin I]\right), \text{ hence}$$

$$\frac{\alpha}{\lambda(1-\alpha)} = \mathbb{E}[(1+\lambda)^{-\sum_{u\sim v}Y_u}|v\notin I].$$

Now for $u \sim v$, let $\tilde{Y}_u$ be the indicator that u is uncovered, conditioned on the event $\{v \notin I\}$. For each u, $\tilde{Y}_u$ has a Bernoulli(p) distribution, where $p = \frac{1+\lambda}{\lambda}\frac{\alpha}{1-\alpha}$, and by Lemma 1 applied to $G \setminus v$, the $\tilde{Y}_u$'s are positively correlated. This gives

$$\frac{\alpha}{\lambda(1-\alpha)} = \mathbb{E}[(1+\lambda)^{-\sum_{u\sim v}\tilde{Y}_u}] > \prod_{u\sim v}\mathbb{E}[(1+\lambda)^{-\tilde{Y}_u}] = \left(1-p+\frac{p}{1+\lambda}\right)^d = \left(\frac{1-2\alpha}{1-\alpha}\right)^d.$$

The function $\frac{\alpha}{\lambda(1-\alpha)}$ is increasing in α, the function $\left(\frac{1-2\alpha}{1-\alpha}\right)^d$ is decreasing in α, and the two functions are equal at $\alpha = \alpha_{T_d}(\lambda)$, so we conclude that $\alpha > \alpha_{T_d}(\lambda)$. $\square$

7. Proof of Theorem 3

Recall that we use the notation $M_G(\lambda)$ for the matching polynomial of a graph G, and let H be a matching drawn from the monomer-dimer model at fugacity λ.

We refer to an edge as covered if an incident edge is in the random matching H. Let e be an edge of G chosen uniformly at random, with an arbitrary left/right orientation chosen at random. In applying the method to matchings we introduce a subtle change of presentation. We now define the *free neighborhood* C to be the subgraph of G containing all the incident edges to e that are not covered by edges outside of both e and its incident edges. When considering independent sets, the free neighborhood was empty if the random vertex v was in the independent set. Here the presence or absence in the matching of e or an edge adjacent to e does not affect C. Given e and C, we use the term *externally uncovered neighbor* to refer to an edge of C incident to e.

The possible free neighborhoods C are completely defined by three parameters: $L, R, K \in \{0, 1, \ldots, d-1\}$, counting the number of left and right neighboring edges in C with an endpoint of degree 1, and the number of triangles formed by e and C. An example is pictured below.
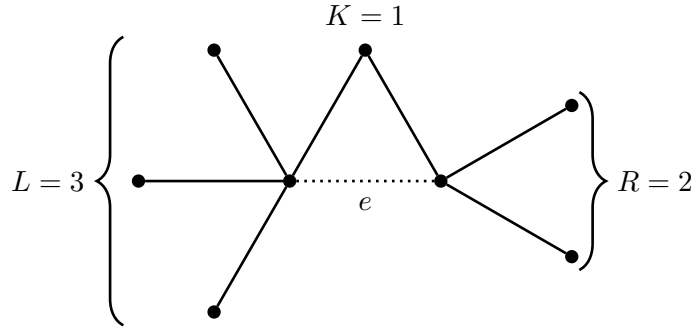
Let $q(i, j, k) = \Pr[L = i, R = j, K = k]$, and denote the matching polynomial for such a free neighborhood by $M_{i,j,k}$, where we can compute

$$M_{i,j,k}(\lambda) = 1 + (i + j + 2k)\lambda + \left[k^2 + k(i + j - 1) + ij\right]\lambda^2.$$

Conditioned on the event that the free neighborhood of e is C, the random matching H restricted to e and its incident edges is distributed according to the monomer-dimer model on the graph C with the edge e added; the partition function of this model is $\lambda + M_C(\lambda)$, with the term λ corresponding to the event that $e \in H$.

We can write $\alpha_M := \alpha_G^M(\lambda)$ as the expected fraction of edges incident to e that are in the matching, as each edge in a d-regular graph is incident to exactly $2(d-1)$ other edges:

$$\alpha_M = \frac{2}{dn} \sum_e \sum_{f \sim e} \frac{1}{2(d-1)} \Pr[f \in H].$$

For an edge e, conditioned on having free neighbourhood given by parameters i, j and k, the expected number of occupied incident edges to e is

$$\frac{\lambda M'(i, j, k)}{\lambda + M(i, j, k)}.$$

This is because in order to have occupied incident edges the edge e must be unoccupied which happens with probability

$$\frac{M(i, j, k)}{\lambda + M(i, j, k)},$$

and conditioned on e being unoccupied the expected number of occupied edges incident to e is

$$\frac{\lambda M'(i, j, k)}{M(i, j, k)}.$$

From this we can see that

$$\alpha_M = \sum_{i,j,k} q(i, j, k) \frac{\lambda M'_{i,j,k}(\lambda)}{2(d-1)(\lambda + M_{i,j,k}(\lambda))}.$$

If we write the expected fraction of occupied neighbors of e in a configuration as $\overline{\alpha}_M(i, j, k) = \frac{1}{2(d-1)} \frac{\lambda M'_{i,j,k}}{\lambda + M_{i,j,k}}$, the above expression can be written $\alpha_M = \sum_{i,j,k} q(i, j, k)\overline{\alpha}_M(i, j, k)$.

The linear program for matchings. We now introduce additional constraints before optimizing α_M over distributions of free neighborhoods. We could write multiple expressions for α_M, equate them, and solve the maximization problem as we did for independent sets. Using three expressions for α_M we were able to prove Theorem 3 for the case $d = 3$, in which the optimal distribution is supported on only three values: $q(0, 0, 0)$, $q(1, 1, 0)$, $q(2, 2, 0)$. But in general we need at least $d - 1$ constraints (in addition to the constraint that the $q(i, j, k)'s$ sum to one) as the distribution induced by $K_{d,d}$ is supported on d values.

Instead, we write, for all t, two expressions for the marginal probability that the number of uncovered neighbors on a randomly chosen side of a random edge is equal to t. We find the two expressions by choosing uniformly: a random edge e, a random side left or right, and f, a random neighboring edge of e from the given side. We first calculate the probability that e has t uncovered neighbors on the side containing f, then we calculate the probability that f has t uncovered neighbors on the side containing e.

Given a free neighborhood C with $L = i$, $R = j$, and $K = k$, e can have $0, 1, i + k - 1$, or $i + k$ uncovered left neighbors; an edge f to the left of e can have $0, 1, i + k - 2, i + k - 1, i + k$, or $i + k + 1$ uncovered right neighbors (depending on whether f itself is in the free neighborhood C).

Let $\gamma_{i,j,k}^e(t) = \Pr[e$ has t uncovered left neighbors $|L = i, R = j, K = k]$ and $\gamma_{i,j,k}^f(t) = \Pr[f$ has t uncovered right neighbors $|L = i, R = j, K = k]$, where f is a uniformly chosen left neighbor of e.

Claim 1. *Let $\beta_t = 1 + t\lambda$. Then we have*

$$(7) \qquad \gamma_{i,j,k}^e(t) = \frac{1}{\lambda + M_{i,j,k}} \Big(\mathbf{1}_{t=0} \cdot \lambda + \mathbf{1}_{t=1} \cdot [i\lambda\beta_{j+k} + k\lambda\beta_{j+k-1}]$$

$$+ \mathbf{1}_{t=i+k} \cdot \beta_j + \mathbf{1}_{t=i+k-1} \cdot k\lambda \Big)$$

$$(8) \qquad \gamma_{i,j,k}^f(t) = \frac{1}{(d-1)(\lambda + M_{i,j,k})} \Big(\mathbf{1}_{t=0} \cdot [i\lambda\beta_{j+k} + k\lambda\beta_{j+k-1}]$$

$$+ \mathbf{1}_{t=1} \cdot [(d-1)\lambda + (d-2)(i\lambda\beta_{j+k} + k\lambda\beta_{j+k-1})]$$

$$+ \mathbf{1}_{t=i+k-2} \cdot [(i+k-1)k\lambda] + \mathbf{1}_{t=i+k-1} \cdot [(d-i-k)k\lambda + (i+k)j\lambda]$$

$$+ \mathbf{1}_{t=i+k} \cdot [(d-1-i-k)j\lambda + (i+k)] + \mathbf{1}_{t=i+k+1} \cdot [d-1-i-k] \Big).$$

Proof. To compute the functions $\gamma_{i,j,k}^e(t)$ we consider the following disjoint events: 1) no left edge and no right edge from a triangle is in the matching 2) e is in the matching 3) a left edge is in the matching 4) no left edge is in the matching, but a right edge from a triangle is in the matching. These events happen with probability $\frac{\beta_j}{\lambda + M_{i,j,k}}$, $\frac{\lambda}{\lambda + M_{i,j,k}}$, $\frac{i\lambda\beta_{j+k} + k\lambda\beta_{j+k-1}}{\lambda + M_{i,j,k}}$, and $\frac{k\lambda}{\lambda + M_{i,j,k}}$ respectively. Under these events the number of uncovered neighbors of e is $i + k, 0, 1$, and $i + k - 1$ respectively. This gives (7).

To compute the functions $\gamma_{i,j,k}^f(t)$ we refine the above events to include the possible choices of f: f can be an edge outside the free neighborhood with probability $(d - 1 - i - k)/(d - 1)$; an edge in the free neighborhood but not in a triangle with probability $i/(d - 1)$; in the free neighborhood and in a triangle with probability $k/(d - 1)$. If a left edge is in the matching we choose it as f with probability $1/(d - 1)$, and if a right edge in a triangle is in the matching we choose f adjacent to it with probability $1/(d - 1)$. Computing the number of uncovered neighbors of f in each case gives (8). $\qquad\square$

We now define a linear program with constraints imposing that the two different ways of writing the marginal probabilities are equal. The marginal probability constraint for $t = d - 1$ is redundant and we omit it. To account for the equal chance that f is chosen from the left

side of e and the right side of e, we average $\gamma^f_{i,j,k}(t)$ and $\gamma^f_{j,i,k}(t)$, and $\gamma^e_{i,j,k}(t)$ and $\gamma^e_{j,i,k}(t)$.

$$\alpha^*_M = \max \sum_{i,j,k} q(i,j,k)\overline{\alpha}_M(i,j,k) \quad \text{subject to}$$

$$q(i,j,k) \geq 0 \ \forall \ i,j,k$$

$$\sum_{i,j,k} q(i,j,k) = 1$$

$$\sum_{i,j,k} q(i,j,k)\frac{1}{2}\left[\gamma^f_{i,j,k}(t) + \gamma^f_{j,i,k}(t) - \gamma^e_{i,j,k}(t) - \gamma^e_{j,i,k}(t)\right] = 0 \ \forall \, t = 0, \ldots, d-2 \,.$$

Disjoint unions of copies of $K_{d,d}$ are the only graphs that induce a distribution $q(i,j,k)$ supported on triples with $i = j$ and $k = 0$. This gives us a candidate solution to the linear program.

The dual program is

$$\alpha^*_M = \min \Lambda_p \quad \text{subject to}$$

$$\Lambda_p - \overline{\alpha}_M(i,j,k) + \sum_{t=0}^{d-2} \Lambda_t \frac{1}{2}\left[\gamma^f_{i,j,k}(t) + \gamma^f_{j,i,k}(t) - \gamma^e_{i,j,k}(t) - \gamma^e_{j,i,k}(t)\right] \geq 0 \ \forall \ i,j,k \,.$$

To show that $K_{d,d}$ is optimal, we find values for the dual variables $\Lambda_0, \ldots, \Lambda_{d-2}$ so that the dual constraints hold with $\Lambda_p = \alpha^M_{K_{d,d}}(\lambda)$. To find such values, we solve the system of equations generated by setting equality in the constraints corresponding to $i = j$ and $k = 0$ and solve for the variables Λ_t, $t = 0, \ldots, d-2$.

With this choice of values for the dual variables, we start by simplifying the form of the dual constraints with a substitution coming from equality in the $(i,j,k) = (0,0,0)$ constraint. The $(0,0,0)$ dual constraint has the simple form

$$\Lambda_0 - \Lambda_1 = \alpha^M_{K_{d,d}} \,.$$

Moreover, observe that from the $\mathbf{1}_{t=0}$ and $\mathbf{1}_{t=1}$ terms in $\gamma^e_{i,j,k}(t)$ and $\gamma^f_{i,j,k}(t)$, every dual constraint contains the term

$$\left[\overline{\alpha}_M(i,j,k) - \frac{\lambda}{(\lambda + M_{i,j,k})}\right](\Lambda_0 - \Lambda_1) = \left[\overline{\alpha}_M(i,j,k) - \frac{\lambda}{(\lambda + M_{i,j,k})}\right]\alpha^M_{K_{d,d}} \,.$$

With this simplification, we multiply through by $2(d-1)(\lambda + M_{i,j,k})$ and expand $\overline{\alpha}_M(i,j,k)$ terms to obtain the following form of the dual constraints:

$$
\begin{aligned}
(9) \qquad \alpha_{K_{d,d}}^{M}&\big[\lambda M'_{i,j,k} + 2(d-1)M_{i,j,k}\big] - \lambda M'_{i,j,k} \\
&+ \Lambda_{i+k-2} \cdot (i+k-1)k\lambda \\
&+ \Lambda_{i+k-1} \cdot [(d-i-k)k\lambda + (i+k)j\lambda - (d-1)k\lambda] \\
&+ \Lambda_{i+k} \cdot [(d-1-i-k)j\lambda + i+k - (d-1)\beta_j] \\
&+ \Lambda_{i+k+1} \cdot (d-1-i-k) \\
&+ \Lambda_{j+k-2} \cdot (j+k-1)k\lambda \\
&+ \Lambda_{j+k-1} \cdot [(d-j-k)k\lambda + (j+k)i\lambda - (d-1)k\lambda] \\
&+ \Lambda_{j+k} \cdot [(d-1-j-k)i\lambda + j+k - (d-1)\beta_i] \\
&+ \Lambda_{j+k+1} \cdot (d-1-j-k) \geq 0 \,.
\end{aligned}
$$

The $(i,i,0)$ equality constraints now read

$$
(10) \qquad \alpha_{K_{d,d}}^{M}\beta_i\Big(\beta_i + \tfrac{i\lambda}{d-1}\Big) - \tfrac{i\lambda\beta_i}{d-1} + \Lambda_{i-1}\tfrac{i^2\lambda}{d-1} - \Lambda_i\tfrac{d-1-i+i^2\lambda}{d-1} + \Lambda_{i+1}\tfrac{d-1-i}{d-1} = 0\,.
$$

With this we can write Λ_{i+k+1} in terms of Λ_{i+k} and Λ_{i+k-1}, and similarly for Λ_{j+k+1}. Substituting this into (9) and dividing by λ we derive the simplified form of the dual constraints:

$$
\begin{aligned}
(11) \qquad \lambda\big[(i-j)^2 &+ 2k\big]\big(1 - d\alpha_{K_{d,d}}^{M}\big) \\
&+ \Lambda_{i+k-2}(i+k-1)k + \Lambda_{i+k-1}[k + (i+k)(j-i-2k)] \\
&+ \Lambda_{i+k}(i+k)(i+k-j) \\
&+ \Lambda_{j+k-2}(j+k-1)k + \Lambda_{j+k-1}[k + (j+k)(i-j-2k)] \\
&+ \Lambda_{j+k}(j+k)(j+k-i) \geq 0 \,.
\end{aligned}
$$

Write $L(i,j,k)$ for the LHS of this inequality.

The marginal constraint for $t = d-1$ was omitted, but we nonetheless introduce $\Lambda_{d-1} := 0$ in order to simplify the presentation of the argument. The $(d-1, d-1, 0)$ equality constraint gives Λ_{d-2} directly:

$$
\Lambda_{d-2} = \frac{1}{(d-1)\lambda}\left[\lambda + (d-1)\lambda^2 - \alpha_{K_{d,d}}^{M}\beta_{d-1}\beta_d\right]\,.
$$

With Λ_{d-1}, Λ_{d-2}, and the recurrence relation (10) the dual variables are fully determined. We do not give a closed-form expression for Λ_t as the values are used in an induction below. Using Λ_{d-1}, Λ_{d-2}, and (10) suffices for the proof.

We now reduce the problem of showing that the dual constraints (11) corresponding to triples (i,j,k) with $k > 0$ or $i \neq j$ hold with strict inequality to showing that a particular function is increasing. We go on to prove this fact in Claims 2 and 3.

Putting $k = 0$ into (11) gives:

$$
\begin{aligned}
\frac{L(i,j,0)}{(j-i)} &= \lambda(j-i)(1 - d\alpha_{K_{d,d}}^{M}) + i\Lambda_{i-1} - i\Lambda_i - j\Lambda_{j-1} + j\Lambda_j \\
&= F_d(j) - F_d(i)
\end{aligned}
$$

where

$$
(12) \qquad F_d(t) := t\left[\lambda(1 - d\alpha_{K_{d,d}}^{M}) + \Lambda_t - \Lambda_{t-1}\right]\,.
$$

From (11) we obtain

$$L(i-1, j-1, k+1) - L(i, j, k) = F_d(i+k) - F_d(i+k-1) + F_d(j+k) - F_d(j+k-1).$$

Therefore if $F_d(t)$ is strictly increasing, we have $L(i, j, 0) > 0$ for $i \neq j$, and $L(i-1, j-1, k+1) > L(i, j, k) > \cdots > L(i+k, j+k, 0) \geq 0$.

We first find an explicit expression for $F_d(t)$. Recall that we write $M_{K_{t,t}}$ for the matching polynomial of the graph $K_{t,t}$.

Claim 2. *For all $d \geq 2$ and $1 \leq t \leq d-1$,*

$$(13) \qquad\qquad F_d(t) = \frac{t(d-1)}{M_{K_{d,d}}} \sum_{\ell=t-1}^{d-2} \frac{(d-1-t)!}{(\ell+1-t)!} \lambda^{d-\ell} M_{K_{\ell,\ell}}.$$

Proof. We will use the following two facts:

$$(14) \qquad\qquad M_{K_{d,d}} - \beta_{2d-1} M_{K_{d-1,d-1}} + (d-1)^2 \lambda^2 M_{K_{d-2,d-2}} = 0$$

$$(15) \qquad\qquad \alpha^M_{K_{d,d}} = \frac{\lambda M_{K_{d-1,d-1}}}{M_{K_{d,d}}}.$$

The first is a Laguerre polynomial identity, verifiable by hand; the second is a short calculation. The equality dual constraint (10) implies:

$$(16) \qquad (d-1-t)F_d(t+1) = (t+1)[t\lambda F_d(t) + (d-1)\lambda - (d-1)\alpha^M_{K_{d,d}} \beta_{d+t}].$$

We first show that the right hand side of (13) satisfies the above recurrence relation. Using (15) this amounts to showing that the following expression is equal to zero for all $d \geq 2$ and $1 \leq t \leq d-1$:

$$\Phi_d(t) := (d-1-t)! \left(\sum_{\ell=t}^{d-2} \frac{\lambda^{d-\ell} M_{K_{\ell,\ell}}}{(\ell-t)!} - t^2 \sum_{\ell=t-1}^{d-2} \frac{\lambda^{d+1-\ell} M_{K_{\ell,\ell}}}{(\ell+1-t)!} \right) - \lambda(M_{K_{d,d}} - \beta_{d+t} M_{K_{d-1,d-1}}).$$

We proceed by induction on d. Note that when $d=2$, $\Phi_2(1)$ is easily verified to be zero. Note that

$$\Phi_{d+1}(t) = \lambda\Big((d-t)\Phi_d(t) - M_{K_{d+1,d+1}} + \beta_{2d+1} M_{K_{d,d}} - d^2\lambda^2 M_{K_{d-1,d-1}} \Big).$$

By the induction hypothesis and (14) the result follows. To complete the proof of the claim it suffices to show that (13) holds for $t = d-1$. Recalling that

$$\Lambda_{d-1} = 0$$

$$\Lambda_{d-2} = \frac{1}{d-1} + \lambda - \frac{\alpha^M_{K_{d,d}}}{(d-1)\lambda}\beta_d\beta_{d-1},$$

substituting into (12), and using (14) and (15) we have

$$F_d(d-1) = (d-1)\left[\lambda(1-d\alpha^M_{K_{d,d}}) - \frac{1}{d-1} - \lambda + \frac{\alpha^M_{K_{d,d}}}{(d-1)\lambda}\beta_d\beta_{d-1}\right]$$

$$= \frac{\alpha^M_{K_{d,d}}}{\lambda}\beta_{2d-1} - 1$$

$$= \frac{1}{M_{K_{d,d}}}\left[\beta_{2d-1}M_{K_{d-1,d-1}} - M_{K_{d,d}}\right]$$

$$= \frac{(d-1)^2\lambda^2 M_{K_{d-2,d-2}}}{M_{K_{d,d}}},$$

verifying (13) for $t = d-1$. $\qquad\square$

Using Claim 2 we prove the following.

Claim 3. $F_d(t)$ *is strictly increasing as a function of* t.

Proof. To prove that $F_d(t)$ is increasing, we show that

$$R_d(t) := \frac{M_{K_{d,d}}}{(d-1)} \cdot \frac{F_d(t+1) - F_d(t)}{(d-2-t)!}$$

$$= (t+1)\sum_{\ell=t}^{d-2}\frac{\lambda^{d-\ell}}{(\ell-t)!}M_{K_{\ell,\ell}} - t(d-1-t)\sum_{\ell=t-1}^{d-2}\frac{\lambda^{d-\ell}}{(\ell+1-t)!}M_{K_{\ell,\ell}}$$

is positive for each t with $1 \le t \le d-2$. We do this by fixing t and inducting on d from $t+2$ upwards. A useful inequality will be $M_{K_{t,t}} > t\lambda M_{K_{t-1,t-1}}$ which comes from only counting matchings of $K_{t,t}$ that use a specific vertex. Iterating this inequality we obtain

$$(17) \qquad\qquad M_{K_{t,t}} > \frac{t!}{\ell!}\lambda^{t-\ell}M_{K_{\ell,\ell}} \text{ for } 0 \le \ell \le t-1.$$

For the base case of our induction, $d = t+2$, we have $R_d(d-2) = \lambda^2\left[M_{K_{d-2,d-2}} - (d-2)\lambda M_{K_{d-3,d-3}}\right]$ which by (17) is positive.

For the inductive step we have

$$R_{d+1}(t) = \lambda\left[R_d(t) + \frac{\lambda}{(d-1-t)!}M_{K_{d-1,d-1}} - \sum_{\ell=t-1}^{d-2}\frac{t\lambda^{d-\ell}}{(\ell-t+1)!}M_{K_{\ell,\ell}}\right]$$

and so it is sufficient to show

$$(18) \qquad\qquad \sum_{\ell=t-1}^{d-2}\frac{t\lambda^{d-\ell}}{(\ell+1-t)!}M_{K_{\ell,\ell}} < \frac{\lambda}{(d-1-t)!}M_{K_{d-1,d-1}}.$$

We use the inequality (17) in each term of the sum to see that the LHS of (18) is less than

$$\sum_{\ell=t-1}^{d-2}\frac{t\ell!\lambda}{(\ell+1-t)!(d-1)!}M_{K_{d-1,d-1}}$$

and so

$$\sum_{\ell=t-1}^{d-2} \frac{t\lambda^{d-\ell}}{(\ell+1-t)!} M_{K_{\ell,\ell}} < \sum_{\ell=t-1}^{d-2} \frac{t\ell!\lambda}{(\ell+1-t)!(d-1)!} M_{K_{d-1,d-1}}$$

$$= \frac{\lambda M_{K_{d-1,d-1}}}{(d-1-t)!} \cdot \sum_{\ell=t-1}^{d-2} \frac{t\ell!(d-1-t)!}{(\ell+1-t)!(d-1)!}$$

$$= \frac{\lambda M_{K_{d-1,d-1}}}{(d-1-t)!} \cdot \binom{d-1}{t}^{-1} \cdot \sum_{\ell=t-1}^{d-2} \binom{\ell}{t-1}$$

$$= \frac{\lambda M_{K_{d-1,d-1}}}{(d-1-t)!},$$

therefore (18) holds as required.   □

This completes the proof of dual feasibility and shows our candidate solution to the primal program is optimal. The uniqueness of the solution follows from two facts. First, strict inequality in the dual constraints outside of the $(i,i,0)$ constraints implies, by complementary slackness, that the support of any optimal solution in the primal is contained in the set of $(i,i,0)$ configurations. Second, the distribution induced by $K_{d,d}$ is the unique distribution satisfying the constraints with such a support. This follows from the fact that Λ_i is uniquely determined by (10) where we have set the $(i,i,0)$ dual constraints to hold with equality, which in turn shows that the relevant $d \times d$ submatrix of the constraint matrix is full rank. This proves Theorem 3.

8. Independent sets and matchings of a given size

Let $i_k(G)$ be the number of independent sets of size k in a graph G, and $m_k(G)$ the number of matchings of size k. Kahn [18] conjectured that $i_k(G)$ is maximized over d-regular, n-vertex graphs by $H_{d,n}$ for all k (when $2d$ divides n), and Friedland, Krop, and Markström [12] conjectured the same for $m_k(G)$. Previous bounds towards these conjectures were given in [7, 16, 23]; for d fixed and k linear in n, all previous bounds were off the conjectured values by a multiplicative factor exponential in n. Here we adapt the method of Carroll, Galvin, and Tetali (and use the above result on the matching polynomial) to give bounds for both problems that are tight up to a factor of $2\sqrt{n}$, for all d and all k.

Theorem 4. *For all d-regular graphs G on n vertices (where $2d$ divides n),*

$$i_k(G) \leq 2\sqrt{n} \cdot i_k(H_{d,n})$$

and

$$m_k(G) \leq 2\sqrt{n} \cdot m_k(H_{d,n}).$$

We start with a fact about the independence and matching polynomials of $H_{d,n}$.

Lemma 2. *For all $1 \leq k \leq n/2$, there exists a λ so that*

$$\frac{i_k(H_{d,n})\lambda^k}{P_{H_{d,n}}(\lambda)} = \Pr_{H_{d,n}}[|I| = k] > \frac{1}{2\sqrt{n}}$$

and a λ so that

$$\frac{m_k(H_{d,n})\lambda^k}{M_{H_{d,n}}(\lambda)} = \Pr_{H_{d,n}}[|H| = k] > \frac{1}{2\sqrt{n}} \, .$$

Proof. The distribution of the size of a random independent set I drawn from the hard-core model on $H_{d,n}$ is log-concave; that is,

$$\Pr_{H_{d,n}}[|I| = j]^2 > \Pr_{H_{d,n}}[|I| = j+1] \cdot \Pr_{H_{d,n}}[|I| = j-1]$$

for all $1 < j < n/2$. This follows from two facts: the size distribution of the hard-core model on $K_{d,d}$ is log-concave, and the convolution of two log-concave distributions is again log-concave. The first fact is simply the calculation

$$\binom{d}{j}^2 > \binom{d}{j-1}\binom{d}{j+1}.$$

Now choose λ so that $\Pr_{H_{d,n}}[|I| = k] = \Pr_{H_{d,n}}[|I| = k+1]$. Log-concavity then implies that $\Pr_{H_{d,n}}[|I| = k]$ is maximal. Some explicit computations for the variance for a single $K_{d,d}$ give that the variance of $|I|$ is at most $n/8$; then via Chebyshev's inequality, with probability at least $2/3$ the size of I is one of at most $\frac{4}{3}\sqrt{n}$ values, and thus the largest probability of a single size is greater than $\frac{1}{2\sqrt{n}}$.

The proof for $m_k(H_{d,n})$ is the same: the variance of the size of a random matching is also at most $n/8$ (see, e.g. [17]), and log-concavity of the size distribution on $K_{d,d}$ is verified via the inequality

$$\binom{d}{j}^4 j!^2 > \binom{d}{j-1}^2 (j-1)! \binom{d}{j+1}^2 (j+1)! \qquad \square$$

Proof of Theorem 4. Assume for sake of contradiction that $m_k(G) > 2\sqrt{n} \cdot m_k(H_{d,n})$. Choose λ according to Lemma 2. We have:

$$M_G(\lambda) \geq m_k(G)\lambda^k > 2\sqrt{n} \cdot m_k(H_{d,n})\lambda^k > M_{H_{d,n}}(\lambda),$$

but this contradicts Theorem 3. The case of independent sets is identical. $\square$

The above proof is essentially the same as the proofs in Carroll, Galvin, and Tetali [7] with the small observation that λ can be chosen so that k is the most likely size of a matching (or independent set) drawn from $H_{d,n}$. The factor $2\sqrt{n}$ in both cases can surely be improved by using some regularity of the independent set and matchings sequence of a general d-regular graph; we leave this for future work.

As a consequence, we prove the asymptotic upper matching conjecture of Friedland, Krop, Lundow, and Markström [11]. Fix d and consider an infinite sequence of d-regular graphs $\mathcal{G}_d = G_1, G_2, \dots$ where G_n has n vertices. For any $\rho \in [0, 1/2]$, the ρ-monomer entropy is

$$h_{\mathcal{G}_d}(\rho) = \sup_{\{k_n\}} \limsup_{n \to \infty} \frac{\log m_{k_n}(G_n)}{n},$$

where the supremum is taken over all integer sequences $\{k_n\}$ with $k_n/n \to \rho$. Let $h_d(\rho) = \lim_{n \to \infty} \frac{\log m_{\lfloor \rho n \rfloor}(H_{d,n})}{n}$, where the limit is taken over the sequences of integers divisible by $2d$. Then the conjecture states that for all $\mathcal{G}_d$ and all $\rho \in [0, 1/2]$, $h_{\mathcal{G}_d}(\rho) \leq h_d(\rho)$.

To prove this, first assume $\rho > 0$ since for $\rho = 0$ the result is trivially true. Assume for the sake of contradiction that $\limsup \frac{\log m_{k_n}(G_n)}{n} > h_d(\rho) + \epsilon$ for some $\epsilon > 0$. Take N_0 large

enough that for all $n_1 \geq N_0$, divisible by $2d$, $\frac{\log m_{\lfloor \rho n_1 \rfloor}(H_{d,n_1})}{n_1} < h_d(\rho) + \epsilon/2$. Now take some $n \geq N_0$ with $\frac{\log m_{k_n}(G_n)}{n} > h_d(\rho) + \epsilon$, and let $n_1 = 2d \cdot \lceil n/(2d) \rceil$. By Lemma 2, we choose λ so that $m_{\lfloor \rho n_1 \rfloor}(H_{d,n_1})\lambda^{\lfloor \rho n_1 \rfloor} > \frac{1}{2\sqrt{n_1}} M_{H_{d,n_1}}(\lambda)$. Note that since $\rho > 0$, such λ is bounded away from 0 as $n_1 \to \infty$. Then we have

$$
\begin{aligned}
\frac{\log M_{G_n}(\lambda)}{n} \geq \frac{\log m_{k_n}(G_n)\lambda^{k_n}}{n} &> \frac{k_n}{n} \log \lambda + h_d(\rho) + \epsilon \\
&= \rho \log \lambda + h_d(\rho) + \epsilon + o(1) \text{ as } n \to \infty
\end{aligned}
$$

and

$$
\begin{aligned}
\frac{\log M_{K_{d,d}}(\lambda)}{2d} = \frac{\log M_{H_{d,n_1}}(\lambda)}{n_1} &< \frac{\log \left(2\sqrt{n_1} \cdot m_{\lfloor \rho n_1 \rfloor}(H_{d,n_1})\lambda^{\lfloor \rho n_1 \rfloor} \right)}{n_1} \\
&< \frac{\log(2\sqrt{n_1})}{n_1} + \frac{\lfloor \rho n_1 \rfloor}{n_1} \log \lambda + h_d(\rho) + \epsilon/2 \\
&= \rho \log \lambda + h_d(\rho) + \epsilon/2 + o(1),
\end{aligned}
$$

but this contradicts Theorem 3. With the same proof, the analogous statement for independent set entropy holds.

9. Conclusions

To recap, our method consists of writing down a set of constraints on local probabilities in the hard-core or monomer-dimer model that hold for every d-regular graph, then optimizing an expression for the occupancy fraction in terms of local probabilities over all distributions that satisfy the constraints. Verifying that our desired graph is the optimizer involves constructing a feasible solution to the dual linear program. This method allowed us to prove tight bounds on the logarithmic derivative of the partition function in both models. In the case of independent sets the result is a strengthening and an alternate proof of the fact that the independence polynomial is maximized by $K_{d,d}$; in the case of matchings, the corresponding statement about the matching polynomial was itself previously unknown.

In both cases our results are neither implied by nor imply conjectures that the numbers of independent sets [18] and matchings [12] of each given size are maximized by $H_{d,n}$; while we improve the known bounds in both cases, these conjectures remain open. Here we give even stronger conjectures:

Conjecture 1. *Let G be a d-regular, n-vertex graph, where $2d$ divides n. Then for all k, the ratio $\frac{i_k(G)}{i_{k-1}(G)}$ is maximized by $H_{d,n}$.*

Conjecture 2. *Let G be a d-regular, n-vertex graph, where $2d$ divides n. Then for all k, the ratio $\frac{m_k(G)}{m_{k-1}(G)}$ is maximized by $H_{d,n}$.*

Conjecture 1 also appeared in a draft of [23]. These conjectures are stronger than Theorems 1 and 3 and imply the conjectures of [18] and [12]. The relation to the work here is that Conjectures 1 and 2 can be stated as follows: the expected number of neighbors of uniformly random independent set (matching) of size k is minimized by $H_{d,n}$. Theorems 1 and 3 show that such a statement is true when the random independent set (matching) is chosen according to the hard-core model instead of uniformly over those of a given size.

Acknowledgements

References

[1] N. Alon. Independence numbers of locally sparse graphs and a ramsey type problem. *Random Structures and Algorithms*, 9(3):271–278, 1996.

[2] F. Bencs. Christoffel-Darboux type identities for independence polynomial. *arXiv preprint arXiv:1409.2527*, 2014.

[3] I. Benjamini and O. Schramm. Recurrence of distributional limits of finite planar graphs. In *Selected Works of Oded Schramm*, pages 533–545. Springer, 2011.

[4] N. Bhatnagar, A. Sly, P. Tetali, et al. Decay of correlations for the hardcore model on the d-regular random graph. *Electronic Journal of Probability*, 21, 2016.

[5] S. Boyd and L. Vandenberghe. *Convex optimization*. Cambridge university press, 2004.

[6] L. Bregman. Some properties of nonnegative matrices and their permanents. In *Soviet Math. Dokl*, volume 14, pages 945–949, 1973.

[7] T. Carroll, D. Galvin, and P. Tetali. Matchings and independent sets of a fixed size in regular graphs. *Journal of Combinatorial Theory, Series A*, 116(7):1219–1227, 2009.

[8] P. Csikvári. Lower matching conjecture, and a new proof of Schrijver's and Gurvits's theorems. *arXiv preprint arXiv:1406.0766*, 2014.

[9] P. Csikvári. Matchings in vertex-transitive bipartite graphs. *Israel Journal of Mathematics*, 215(1):99–134, 2016.

[10] J. A. Fill and M. Machida. Stochastic monotonicity and realizable monotonicity. *Annals of probability*, pages 938–978, 2001.

[11] S. Friedland, E. Krop, P. H. Lundow, and K. Markström. On the validations of the asymptotic matching conjectures. *Journal of Statistical Physics*, 133(3):513–533, 2008.

[12] S. Friedland, E. Krop, and K. Markström. On the number of matchings in regular graphs. *the electronic journal of combinatorics*, 15(1):R110, 2008.

[13] D. Galvin. Three tutorial lectures on entropy and counting. *arXiv preprint arXiv:1406.7872*, 2014.

[14] D. Galvin and P. Tetali. On weighted graph homomorphisms. *DIMACS Series in Discrete Mathematics and Theoretical Computer Science*, 63:97–104, 2004.

[15] O. J. Heilmann and E. H. Lieb. Theory of monomer-dimer systems. *Communications in Mathematical Physics*, 25(3):190–232, 1972.

[16] L. Ilinca and J. Kahn. Asymptotics of the upper matching conjecture. *Journal of Combinatorial Theory, Series A*, 120(5):976–983, 2013.

[17] J. Kahn. A normal law for matchings. *Combinatorica*, 20(3):339–391, 2000.

[18] J. Kahn. An entropy approach to the hard-core model on bipartite graphs. *Combinatorics, Probability and Computing*, 10(03):219–237, 2001.

[19] J. Kahn. Entropy, independent sets and antichains: a new approach to dedekind?s problem. *Proceedings of the American Mathematical Society*, 130(2):371–378, 2002.

[20] S. Kopparty and B. Rossman. The homomorphism domination exponent. *European Journal of Combinatorics*, 32(7):1097–1114, 2011.

[21] M. Lelarge. Counting matchings in irregular bipartite graphs and random lifts. In *Proceedings of the Twenty-Eighth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 2230–2237. SIAM, 2017.

[22] E. Lubetzky and Y. Zhao. On replica symmetry of large deviations in random graphs. *Random Structures & Algorithms*, 2014.

[23] W. Perkins. Birthday inequalities, repulsion, and hard spheres. *Proceedings of the American Mathematical Society*, 144(6):2635–2649, 2016.

[24] J. Radhakrishnan. Entropy and counting. *Computational Mathematics, Modelling and Algorithms*, page 146, 2003.

[25] N. Ruozzi. The bethe partition function of log-supermodular graphical models. In *Advances in Neural Information Processing Systems*, pages 117–125, 2012.

[26] L. J. Schulman, A. Sinclair, and P. Srivastava. Symbolic integration and the complexity of computing averages. In *Foundations of Computer Science (FOCS), 2015 IEEE 56th Annual Symposium on*, pages 1231–1245. IEEE, 2015.

[27] J. B. Shearer. On the independence number of sparse graphs. *Random Structures & Algorithms*, 7(3):269–271, 1995.

[28] A. Sinclair and P. Srivastava. Lee–Yang theorems and the complexity of computing averages. *Communications in Mathematical Physics*, 329(3):827–858, 2014.

[29] J. Van den Berg and J. Steif. Percolation and the hard-core lattice gas model. *Stochastic Processes and their Applications*, 49(2):179–197, 1994.

[30] D. Weitz. Counting independent sets up to the tree threshold. In *Proceedings of the thirty-eighth annual ACM symposium on Theory of computing*, pages 140–149. ACM, 2006.

[31] Y. Zhao. The number of independent sets in a regular graph. *Combinatorics, Probability and Computing*, 19(02):315–320, 2010.

[32] Y. Zhao. The bipartite swapping trick on graph homomorphisms. *SIAM Journal on Discrete Mathematics*, 25(2):660–680, 2011.

Department of Mathematics, London School of Economics, London, UK.

E-mail address: {e.s.davies,m.o.jenssen,b.j.roberts}@lse.ac.uk

School of Mathematics, University of Birmingham, Birmingham, UK.

E-mail address: w.f.perkins@bham.ac.uk