Extremes of the internal energy of the Potts model on cubic graphs

Davies, Ewan; Jenssen, Matthew; Perkins, Will; Roberts, Barnaby

DOI:
10.1002/rsa.20767

Document Version
Peer reviewed version

Citation for published version (Harvard):
Davies, E, Jenssen, M, Perkins, W & Roberts, B 2018, 'Extremes of the internal energy of the Potts model on cubic graphs', *Random Structures and Algorithms*, vol. 53, no. 1, pp. 59-75. https://doi.org/10.1002/rsa.20767

EXTREMES OF THE INTERNAL ENERGY OF THE POTTS MODEL ON CUBIC GRAPHS

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

Abstract. We prove tight upper and lower bounds on the internal energy per particle (expected number of monochromatic edges per vertex) in the anti-ferromagnetic Potts model on cubic graphs at every temperature and for all $q \geq 2$. This immediately implies corresponding tight bounds on the anti-ferromagnetic Potts partition function.

Taking the zero-temperature limit gives new results in extremal combinatorics: the number of q-colorings of a 3-regular graph, for any $q \geq 2$, is maximized by a union of $K_{3,3}$'s. This proves the $d = 3$ case of a conjecture of Galvin and Tetali.

1. The Ising and Potts models

The Potts model is a probabilistic model of interacting *spins* on a graph. Here, we use the term *color* instead of spin to highlight a connection to extremal combinatorics which we cover in Section 2. Let $G = (V, E)$ be a graph and $\sigma \in [q]^{V(G)}$ a coloring (not necessarily proper) of the vertices of G with q possible colors. Let $m(\sigma)$ denote the number of monochromatic edges of G under σ. Then the q-color Potts model partition function is:

$$Z_G^q(\beta) = \sum_{\sigma \in [q]^{V(G)}} e^{-\beta m(\sigma)}.$$

The parameter β is the *inverse temperature* and the model is *antiferromagnetic* if $\beta > 0$ and *ferromagnetic* if $\beta < 0$. (For general statistical physics terminology, we refer the reader to Chapter 2 of [17], for example). The Potts partition function also plays an important role in graph theory as it is an evaluation of the Tutte polynomial of G.

The Potts model [19] (with no external field) is a random q-coloring σ of $V(G)$ chosen according to the distribution

$$\sigma \mapsto \frac{e^{-\beta m(\sigma)}}{Z_G^q(\beta)}.$$

Date: May 18, 2017.

Key words and phrases. Potts model, partition function, graph colorings, graph homomorphims, Ising model.

Thus the partition function $Z_G^q(\beta)$ is the normalizing constant that makes this a probability distribution. When β is positive, the model prefers colorings with fewer monochromatic edges, and the effect is intensified as β gets large. When β is negative (the ferromagnetic Potts model), the distribution is biased towards colorings with more monochromatic edges.

The Potts model generalizes the *Ising model* [10] (the case $q = 2$). See [25] for a survey of the Potts model.

The negative of the logarithmic derivative of Z_G^q with respect to β gives the expected number of monochromatic edges, or the *internal energy* of the model. If we scale by the number of vertices, this gives the *internal energy per particle*, $U_G^q(\beta)$:

$$\begin{aligned} U_G^q(\beta) &:= -\frac{1}{|V(G)|}\frac{d}{d\beta}(\log Z_G^q(\beta)) \\ &= -\frac{1}{|V(G)|}\frac{1}{Z_G^q(\beta)}\frac{d}{d\beta}(Z_G^q(\beta)) \\ &= \frac{1}{|V(G)|}\frac{\sum_{\sigma\in[q]^{V(G)}} m(\sigma)e^{-\beta m(\sigma)}}{Z_G^q(\beta)} \\ &= \frac{1}{|V(G)|}\mathbb{E}_\sigma[m(\sigma)]. \end{aligned}$$

When $\beta = 0$ there are no interactions in the model, and $Z_G^q(0) = q^{|V(G)|}$ for all G. Starting from here, we can integrate the internal energy per particle to obtain the scaled logarithm of the partition function, or the *free energy per particle*, $F_G^q(\beta) := \frac{1}{|V(G)|}\log Z_G^q(\beta)$.

$$F_G^q(\beta) = \frac{1}{|V(G)|}\log Z_G^q(\beta) = \log q - \int_0^\beta U_G^q(t)\,dt. \tag{1}$$

In this paper we derive tight bounds on $U_G^q(\beta)$ for cubic (3-regular) graphs in the anti-ferromagnetic ($\beta > 0$) regime. From (1) these bounds immediately imply corresponding tight bounds on the free energy of the Ising and Potts models and hence the respective partition functions. We determine, for every q, the maximum and minimum of both the internal energy and the free energy per particle as well as the family of graphs that achieve these bounds.

Recall that $K_{d,d}$ is the complete d-regular bipartite graph on $2d$ vertices and K_{d+1} is the complete graph on $d+1$ vertices.

Theorem 1. *For any cubic graph G, any $q \geq 2$, and any $\beta > 0$,*

$$U_{K_{3,3}}^q(\beta) \leq U_G^q(\beta) \leq U_{K_4}^q(\beta).$$

Furthermore, the respective equalities hold if and only if G is a union of $K_{3,3}$'s or a union of K_4's. As a corollary via (1)*, we have*

$$F_{K_4}^q(\beta) \leq F_G^q(\beta) \leq F_{K_{3,3}}^q(\beta).$$

We conjecture that these bounds extend to higher regularity d (the case $d = 2$ is simply a calculation, see Section 5).

Conjecture 1. *For any d-regular graph G, any $q \geq 2$, and any $\beta > 0$,*

$$U^q_{K_{d,d}}(\beta) \leq U^q_G(\beta) \leq U^q_{K_{d+1}}(\beta),$$

and in particular,

$$F^q_{K_{d+1}}(\beta) \leq F^q_G(\beta) \leq F^q_{K_{d,d}}(\beta).$$

If we restrict ourselves to bipartite regular graphs, then Galvin (building on [9, 11]) proved that the maximizer of the free energy is $K_{d,d}$.

Theorem 2 (Galvin [6])**.** *For any d-regular bipartite G, any β and any $q \geq 2$,*

$$F^q_G(\beta) \leq F^q_{K_{d,d}}(\beta).$$

For lower bounds on the ferromagnetic Ising and Potts free energy over regular, bipartite graphs, where the infimum is the 'Bethe free energy', or the free energy of the infinite d-regular tree, see Ruozzi [20, 21] and Csikvári [2].

A bound such as Theorem 2 was known without the bipartite restriction in one case previously: in the anti-ferromagnetic Ising model. An extension of Galvin's result by Zhao [26] using the 'bipartite swapping trick' gives the following.

Theorem 3 (Zhao [26])**.** *For any d-regular graph G, $\beta > 0$, and $q = 2$ (the Ising model),*

$$F^q_G(\beta) \leq F^q_{K_{d,d}}(\beta).$$

Zhao's method does not work for $q \geq 3$, and in the ferromagnetic phase Galvin's result cannot be extended to all G; $K_{d,d}$ is not the maximizer. The clique K_{d+1} has a higher free energy for any d when $\beta < 0$. It is natural to conjecture that K_{d+1} is in fact extremal in this case, and also that Galvin's result can be extended to triangle-free graphs.

Conjecture 2. *For any d-regular G, any $q \geq 2$, and any $\beta < 0$,*

$$U^q_G(\beta) \leq U^q_{K_{d+1}}(\beta),$$

and in particular,

$$F^q_G(\beta) \leq F^q_{K_{d+1}}(\beta).$$

Moreover, if in addition G is triangle-free, then for any β

$$U^q_G(\beta) \leq U^q_{K_{d,d}}(\beta),$$

and in particular,

$$F^q_G(\beta) \leq F^q_{K_{d,d}}(\beta).$$

The main contribution of this paper is to prove Theorem 1. We do so by considering the following experiment. Fix q, β and a d-regular graph G. Choose a vertex v uniformly from $V(G)$ and independently sample a coloring $\sigma \in [q]^{V(G)}$ from the Potts model. Now for each neighbor u of v, record the number of its 'external' neighbors (neighbors outside $v \cup N(v)$) receiving each color; record also any edges within $N(v)$ (see Figure 1 for examples). This gives a *local view* of σ from v. Note that although we have sampled a coloring σ of the whole graph, the colors of v and its neighbors do not form part of the local view. In fact it is best to think of these colors as having not been revealed. An important part of the method is that, conditioned on the local view, the distribution of colorings of v and its neighbors can be determined. For fixed d and q there are only a finite number of possible local views; call this set of local views $\mathcal{L}_{d,q}$. Each d-regular graph G and inverse temperature β induces a probability distribution on $\mathcal{L}_{d,q}$.

Not all probability distributions on $\mathcal{L}_{d,q}$ can arise from a graph; there are certain consistency conditions that must hold. For example, the expected number of monochromatic edges incident to v must equal the expected number of monochromatic edges incident to a uniformly chosen neighbor of v. Moreover, we can compute both of these expectations given a probability distribution on $\mathcal{L}_{d,q}$; in fact they are both linear functions of the probabilities. For $d = 3$ this constraint is sufficient, but for larger d more are required. Another family of consistency conditions are that for every multiset S of size d from q colors, the probability $N(v)$ is colored by S must be the same as the probability $N(u)$ is colored by S for a uniformly chosen neighbor u of v. Finally, the quantity we wish to optimize, $U_G^q(\beta)$ is also a linear function of the probabilities in the distribution on $\mathcal{L}_{d,q}$.

So instead of maximizing or minimizing $U_G^q(\beta)$ over all d-regular graphs, we relax the problem and instead optimize over all probability distributions on $\mathcal{L}_{d,q}$ that satisfy the above consistency conditions. This is simply a linear program over $|\mathcal{L}_{d,q}|$ variables. For some values of d, q and β we know this linear program is not tight although we conjecture it to be tight whenever $q \geq d + 1 \geq 3$ and $\beta > 0$.

This method builds on previous work on independent sets and matchings [4, 5, 18] and the Widom-Rowlinson model [1], but here we generalize the previous approach in two ways: 1) we deal with q-spin models instead of 2-spin models; 2) we deal with soft and hard constraints instead of just hard constraints. The family of linear programs in [4] for matchings was an infinite family of LP's indexed by two parameters - the vertex degree d and a fugacity parameter $\lambda > 0$ - and the entire family could be solved analytically with a single proof via LP duality. Here the situation is worse: we have an infinite family of LP's indexed by d, q, β. Moreover, while the number of constraints for the matching LP grew linearly in d, here the number of constraints needed can grow like the integer partition number of d. In this paper we solve the program for $d = 3$ where there are 35 variables.

In Section 3 we solve the LP (both the maximization and minimization problem) for $d = 3$, $q \geq 2$, and all $\beta > 0$, and we solve it in a somewhat mechanical way that does not reveal much about generalizations to higher d.

Nevertheless, we suspect that the LP is tight for a much wider set of parameters, including for $q \geq d + 1$ and $\beta > 0$. For some other parameter values the constraints described above are not enough to solve the internal energy minimization problem for all d, q, $\beta > 0$. It is easy to find values of β so that if $d \geq 4$ and $q \leq d$, the minimizer of the LP is smaller than $U^q_{K_{d,d}}(\beta)$. Two challenges for future work are:

(i) Solve the infinite family of LP's with $d \geq 4, q \geq d + 1$, $\beta > 0$ that we conjecture is tight.
(ii) Find additional consistency conditions (constraints) that can further tighten the LP for $q \leq d$.

It is worth comparing this method to the entropy method used in [6, 9, 11, 26]. The entropy method has the great virtue of generality - the theorems encompass many models all at once. The method we employ here has the virtue that the results are stronger—we show optimality on the level of the internal energy, which is both a physical observable and the derivative of the free energy—and that it can be used to prove tight bounds in a wide variety of situations, albeit with much model-specific work required.

In another direction, extremal bounds in the Potts model on graphs of maximum degree d are given by Sokal [22], proving that all zeros of the chromatic polynomial, $P_G(q)$, have modulus bounded by $7.963907d$. The complete graph K_{d+1} shows that a linear bound in d is best possible. Given our results here, it would be interesting to find the sharp constant in the upper bound; that is, over all d-regular graphs G what is the supremum of $R(G)$ where

$$R(G) = \max\{|r| : r \in \mathbb{C}, P_G(r) = 0\}.$$

In particular, for $d \geq 4$, do we have $R(G) \leq R(K_{d,d})$? And for $d = 3$ do we have $R(G) \leq R(K_4) = 3$? For more, see the discussion in Section 9 of [23].

2. Maximizing the number of q-colorings of d-regular graphs

If we take $\beta \to \infty$ in the Potts model, we bias more and more against monochromatic edges, and thus if a *proper* q-coloring of G exists, the 'zero-temperature' anti-ferromagnetic Potts model is simply the uniform distribution over proper q-colorings of G. The limit of the partition function $\lim_{\beta\to\infty} Z^q_G(\beta) = C_q(G)$, the number of proper q-colorings of G. Maximizing $C_q(G)$ over different families of graphs has been the study of much work in extremal combinatorics. Linial [15] asked which graph on n vertices with m edges maximizes $C_q(G)$. After a series of bounds by Labeznik and coauthors [12, 13, 14], Loh, Pikhurko, and Sudakov [16] gave a complete answer to this question for a wide range of parameters q, n, m, using the

Szemerédi's Regularity Lemma to reduce the maximization problem over graphs to a quadratic program in $2^q - 1$ variables.

A similar question in a very different setting is to ask which d-regular, n-vertex graph maximizes the number of q-colorings; or, given that C_q is multiplicative when taking disjoint unions of graphs, which d-regular graph maximizes $\frac{1}{|V(G)|} \log C_q(G)$? Although neither question specifies the sparsity of the graph, one can think of Linial's question as a question about dense graphs and this question as one about sparse graphs (and the techniques of [16] and this paper reflect this: the Regularity Lemma primarily concerns dense graphs, while statistical mechanics is primarily concerned with sparse, regular graphs).

For regular graphs, Galvin and Tetali [9] conjectured that $K_{d,d}$ maximizes the normalized number of q-colorings over all d-regular graphs.

Conjecture 3 (Galvin–Tetali [9])**.** *For any $q \geq 2$, $d \geq 1$, and all d-regular graphs G,*

$$C_q(G)^{1/|V(G)|} \leq C_q(K_{d,d})^{1/(2d)}. \tag{2}$$

In the same paper they prove that (2) holds for all d-regular, *bipartite* G. In the language of graph homomorphisms, $C_q(G)$ counts the number of homomorphisms from G into K_q, and their results holds for the number of homomorphisms of a d-regular bipartite G in to *any* target graph H.

Before this work, Conjecture 3 was not known for any pair (q, d) apart from the trivial cases $d = 1$, $d = 2$, and $q = 2$ (see Section 5). However, significant partial progress was made in addition to the bipartite case. Employing the bipartite swapping trick, Zhao [26] showed that for $q \geq (2n)^{2n-2}$, the bipartiteness restriction could be removed for graphs on n vertices. Galvin [7] then reduced the lower bound on q, showing that $q > 2^{\binom{nd/2}{4}}$ suffices. Dependence on n in the number of colors is of course not ideal, as it does not prove Conjecture 3 for any pair (q, d), but it does restrict the class of possible counterexamples. In another direction of partial progress on Conjecture 3, Galvin [8] gave an upper bound on $C_q(G)$ for all d-regular G, that is tight, asymptotically in d, on a logarithmic scale.

The following lemma relates bounds on internal energy per particle to bounds on the number of q-colorings.

Lemma 1. *Fix d, q. If for all d-regular G, and all $\beta > 0$, $U^q_G(\beta) \geq U^q_{K_{d,d}}(\beta)$, then* (2) *holds.*

Proof. Let G be any d-regular graph. If $C_q(G) = 0$ then (2) clearly holds. Otherwise, we take logarithms and write

$$\begin{aligned}\frac{1}{|V(G)|}\log C_q(G) &= \lim_{\beta\to+\infty}\frac{1}{|V(G)|}\log Z_G^q(\beta)\\ &= \log q - \int_0^\infty U_G^q(\beta)\,d\beta\\ &\le \log q - \int_0^\infty U_{K_{d,d}}^q(\beta)\,d\beta\\ &= \frac{1}{2d}\log C_q(K_{d,d})\,.\end{aligned}$$

□

As a corollary of Theorem 1 and Lemma 1, we prove Conjecture 3 for $d = 3$ and all q.

Corollary 1. *For any 3-regular graph G, and any $q \ge 2$,*

$$C_q(G)^{1/|V(G)|} \le C_q(K_{3,3})^{1/6}\,,$$

with equality if and only if G is a union of $K_{3,3}$'s.

We remark that in a similar fashion, Theorem 1 gives that $C_q(G)^{1/|V(G)|} \ge C_q(K_4)^{1/4}$ for all cubic graphs G, but this result was recently proved for all d by Csikvári (see [27]).

Theorem 4 (Csikvári)**.** *For all d, all $q \ge 2$, and all d-regular G,*

$$C_q(G)^{1/|V(G)|} \ge C_q(K_{d+1})^{1/(d+1)}.$$

Csikvári and Lin [3] also proved that for any d-regular, bipartite G,

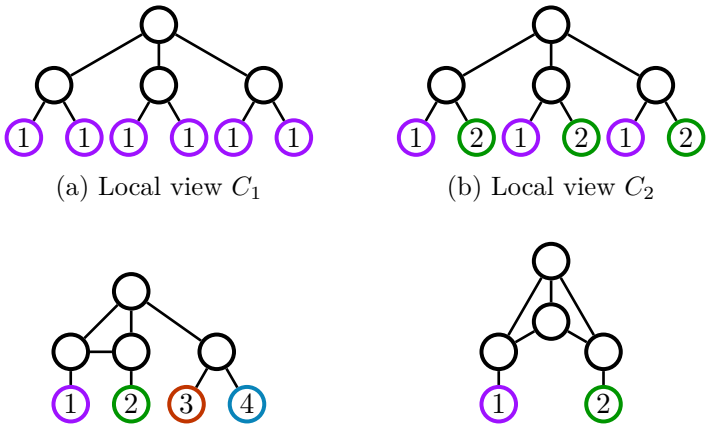
$$C_q(G)^{1/|V(G)|} \ge q\left(\frac{q-1}{q}\right)^{d/2}\,,$$

a result that, for q large enough as function of d, is tight asymptotically for a sequence of bipartite graphs of diverging girth.

3. Proof of Theorem 1

In this section we prove Theorem 1 by formulating and solving the LP described in the introduction. For brevity we drop the superscripts in notation for partition functions and internal energy of a graph G, writing Z_G and U_G for these quantities.

Recall the experiment which defines the local view: draw a coloring $\sigma \in [q]^{V(G)}$ according to the q-color Potts model with inverse temperature β and independently, uniformly at random choose a vertex $v \in V(G)$. The local view consists of the induced subgraph of G on $v \cup N(v)$, together with, for each $u \in N(v)$, the multiset of colors that appears in $N(u) \setminus (\{v\} \cup N(v))$. Four examples are pictured in Figure 1. As noted in the introduction, our calculations depend only on the *number* of 'external neighbors' of vertices $u \in N(v)$ which receive each color, and not the graph structure between



(a) Local view C_1 (b) Local view C_2

(c) Local view with 1 triangle (d) Local view with 2 triangles

Figure 1. Example local views. Figures (a) and (b) show, up to permutations of the colors, the only local views which can arise in $K_{3,3}$. The colored, numbered vertices are representations of the multiset of colors that appear in the boundary $N(u) \setminus (\{v\} \cup N(v))$ for $u \in N(v)$.

these external neighbors. For clarity we draw the vertices themselves. Let $\mathcal{C}_q$ denote all possible local views for the q-color Potts model on cubic graphs.

As q grows larger the number of possible local views grows like $q^{d(d-1)}$. However, if we consider equivalence classes of the local views under permutations of the colors (as detailed below), the number of possible local views is bounded in terms of d. This makes the LP finite for any fixed d. For a complete list of local views (up to equivalence) for $d = 3$ see Appendix A.

Suppose that the local view C arises from selecting the coloring σ and the vertex v. We refer to the colored 'external neighbors' at distance two from v as the *boundary*, and write V_C for the set of uncolored vertices in C, so that the set of q-colorings of these vertices is $[q]^{V_C}$. The coloring σ induces a *local coloring* $\chi \in [q]^{V_C}$ that, by the Markov property of the Potts model, is distributed according to the Potts model on C. For $\chi \in [q]^{V(C)}$, write $m(\chi)$ for the total number of monochromatic edges in C (including any monochromatic edges between V_C and the boundary), and given a vertex $u \in V_C$ write $m_u(\chi)$ for the number of monochromatic edges in C incident to u. Then, with the local partition function defined as

$$Z_C(q,\beta) := \sum_{\chi \in [q]^{V_C}} e^{-\beta m(\chi)},$$

a local coloring $\chi \in [q]^{V_C}$ is distributed according to

$$\chi \mapsto \frac{e^{-\beta m(\chi)}}{Z_C(q,\beta)}.$$

This fact means that we can interpret the internal energy per particle as an expectation over the random local view C and local coloring χ. Each edge of G is incident to exactly two vertices, hence

$$\begin{aligned} U_G(\beta) &= \frac{1}{|V(G)|}\mathbb{E}_\sigma[m(\sigma)] \\ &= \frac{1}{2|V(G)|}\sum_{v\in V(G)}\sum_{u\in N(v)}\mathbb{P}_\sigma(uv \text{ monochromatic}) \\ &= \frac{1}{2}\mathbb{E}_C\Big[\sum_{u\in N(v)}\mathbb{P}(uv \text{ monochromatic}|C)\Big] \\ &= \frac{1}{2}\mathbb{E}_{C,\chi}[m_v(\chi)]\,. \end{aligned}$$

Moreover, since G is regular, a neighbor of v chosen uniformly at random is distributed uniformly over $V(G)$, giving

$$U_G(\beta) = \frac{1}{2}\mathbb{E}_{C,\chi}\Big[\frac{1}{3}\sum_{u\in N(v)} m_u(\chi)\Big]\,.$$

Given these observations, for a local view C we define

$$\begin{aligned} U_C^v &:= \frac{1}{2Z_C}\sum_{\chi\in[q]^{V_C}} m_v(\chi)e^{-\beta m(\chi)}\,, \\ U_C^N &:= \frac{1}{6Z_C}\sum_{\chi\in[q]^{V_C}}\sum_{u\in N(v)} m_u(\chi)e^{-\beta m(\chi)}\,, \end{aligned}$$

so that

$$U_G(\beta) = \mathbb{E}_C[U_C^v] = \mathbb{E}_C[U_C^N]\,,$$

giving us a *constraint* on probability distributions on local views that holds for all distributions arising from graphs.

We can now define the two LP's for the q-color Potts model,

$$\begin{aligned} \{U^{\min}, U^{\max}\} = \{\text{Minimum, Maximum}\} \text{ of } & \sum_{C\in\mathcal{C}_q} p_C U_C^v \qquad \text{subject to} \\ & p_C \geq 0\ \forall C\in\mathcal{C}_q\,, \\ & \sum_{C\in\mathcal{C}_q} p_C = 1\,, \\ & \sum_{C\in\mathcal{C}_q} p_C(U_C^v - U_C^N) = 0\,, \end{aligned}$$

The final constraint is an example of a *consistency condition* that holds for all distributions on local views which arise from the experiment on a graph defined at the beginning of this section.

3.1. **Minimizing.** For the minimization problem, the dual program (with variables Λ, Δ) is

$$\begin{aligned} U^{\min} = \max \Lambda \qquad & \text{subject to} \\ \Lambda + \Delta(U_C^v - U_C^N) \leq U_C^v \qquad & \forall C \in \mathcal{C}_q \,. \end{aligned}$$

For a given $\beta > 0$ and $q \geq 2$, to show that $U^{\min} = U_{K_{3,3}}$ via linear programming duality, we must find Δ^* so that the assignment $\Lambda = U_{K_{3,3}}, \Delta = \Delta^*$ is *feasible* for the dual. That is,

$$U_{K_{3,3}} + \Delta^*(U_C^v - U_C^N) \leq U_C^v \tag{3}$$

for all $C \in \mathcal{C}_q$.

In fact it suffices to show (3) on a subset of $\mathcal{C}_q$. We say $C, C' \in \mathcal{C}_q$ are equivalent if

$$U_C^v \equiv U_{C'}^v \text{ and } U_C^N \equiv U_{C'}^N$$

as functions of q and β. For instance if C is obtained from C' by a permutation of the q-colors, then C and C' are equivalent by symmetry. This equivalence relation partitions $\mathcal{C}_q$ into equivalence classes. Call this set of equivalence classes $\mathcal{C}_q'$. We always choose a representative member of the equivalence class that has an initial segment of the colors $[q]$ on its boundary, and write q_C for the total number of colors on the boundary of a local view C. For $d = 3$ and arbitrary q there are 35 non-isomorphic equivalence classes which we list in Appendix A.

We give the local views that arise in the optimizing graphs names, writing C_1 and C_2 (see Figure 1) for representatives of the only two equivalence classes of local views are that can appear with positive probability when $G = K_{3,3}$. In K_4, the only local view that can arise is isomorphic to K_4 itself.

To find a value of Δ^* we solve the dual constraint corresponding to local view C_1 (Figure 1a) to hold with equality when $\Lambda = U_{K_{3,3}}$:

$$U_{K_{3,3}} + \Delta^*(U_{C_1}^v - U_{C_1}^N) = U_{C_1}^v \,.$$

Writing $\lambda = e^{-\beta}$ (so that $0 < \lambda < 1$), we find

$$\begin{aligned} Z_{C_1} &= (\lambda^3 + q - 1)^3 + (q-1)(\lambda^2 + \lambda + q - 2)^3 \,, \\ U_{C_1}^v &= \frac{3}{2Z_{C_1}} \left(\lambda^3(1+\lambda^3)^2 + (q-1)\lambda^2(\lambda^2 + \lambda + q - 2)^2 \right) \\ U_{C_1}^N &= \frac{1}{2Z_{C_1}} \left(3\lambda^3(1+\lambda^3)^2 + (q-1)(\lambda + 2\lambda^2)(\lambda^2 + \lambda + q - 2)^2 \right) \end{aligned}$$

and hence

$$U_{C_1}^N - U_{C_1}^v = \frac{1}{2Z_{C_1}} \lambda(1-\lambda)(q-1)(\lambda^2 + \lambda + q - 2)^2 \,.$$

Then we calculate

$$Z_{K_{3,3}} = q(\lambda^3+q-1)^3 + 3q(q-1)(\lambda^2+\lambda+q-2)^3 + q(q-1)(q-2)(3\lambda+q-3)^3\,,$$

$$U_{K_{3,3}} = \frac{3q}{2Z_{K_{3,3}}}\Big(\lambda^3(\lambda^3+q-1)^2 + \lambda(q-1)(q-2)(3\lambda+q-3)^2 + (q-1)(2\lambda^2+\lambda)(\lambda^2+\lambda+q-2)^2\Big)\,,$$

(4)

$$\begin{aligned}\Delta^* = -\frac{3q(1-\lambda)^2}{2(\lambda^2+\lambda+q-2)^2 Z_{K_{3,3}}}\Big(&2\left(\lambda^4+9\lambda^3+22\lambda^2+27\lambda+13\right)(\lambda-1)^6 + \\ &2\left(\lambda^2+2\lambda-1\right)q^5 + \left(\lambda^5+3\lambda^4+20\lambda^3-41\lambda+17\right)q^4 + \\ &\left(8\lambda^4+31\lambda^3+91\lambda^2+47\lambda-57\right)(\lambda-1)^2q^3 + \\ &\left(25\lambda^4+82\lambda^3+146\lambda^2+22\lambda-95\right)(\lambda-1)^3q^2 + \\ &\left(2\lambda^5+36\lambda^4+87\lambda^3+93\lambda^2-31\lambda-79\right)(\lambda-1)^4q\Big)\,.\end{aligned}$$

Claim 1. *For all $q \geq 2$ and all $\beta > 0$, the function*

$$SLACK(C) = U_C^v + \Delta^*(U_C^N - U_C^v) - U_{K_{3,3}}$$

with Δ^* given by (4) is identically 0 for $C \in \{C_1, C_2\}$ and strictly positive for all other $C \in \mathcal{C}_q'$.

Claim 1 immediately proves that $U_G^q(\beta) \geq U_{K_{3,3},}^q(\beta)$. To show uniqueness, observe that strict positivity of the slack function implies via complementary slackness that the support of any distribution achieving the optimum must be contained in $\{C_1, C_2\}$; $K_{3,3}$ is the only connected graph whose distribution satisfies this. To see this note that, for any other connected cubic graph, there exists a vertex v with two neighbors u_1, u_2 such that the external neighborhoods of u_1 and u_2 are distinct. Then there exists a coloring such that the external neighbors of u_1 are monochromatic, whilst those of u_2 are not. This means a local view not isomorphic to C_1 or C_2 appears with positive probability.

In order to prove Claim 1, we change variables and multiply the slack by a positive scaling factor, carefully chosen to result in a polynomial with positive coefficients. Write $r = q-3$ and $t = e^\beta - 1 = 1/\lambda - 1$, so that for any $q \geq 3$ and $\beta > 0$ we have $r \geq 0$ and $t > 0$. It then suffices to show that the following scaling of the slack is non-negative:

$$\tilde{S}_C = \frac{4(1+t)^{17}(r(1+t)^2+t^2+3t+3)^2}{(3+r)t^2} Z_{K_{3,3}} Z_C \cdot SLACK(C)\,. \tag{5}$$

In fact something stronger is true:

Claim 2. *For all $C \in \mathcal{C}'_q$, $\tilde{S}_C$ is a bivariate polynomial in r and t with all coefficients positive. The polynomial is identically 0 if and only if $C \in \{C_1, C_2\}$.*

For the case $q = 2$ we do something slightly different.

Claim 3. *For all $C \in \mathcal{C}'_2$ (which necessarily use at most 2 colors on the boundary), evaluating $\tilde{S}_C$ at $r = -1$ yields a polynomial in t with positive coefficients. The polynomial is identically 0 if and only if $C \in \{C_1, C_2\}$.*

Claims 2 and 3 are proved by simply computing the functions $\tilde{S}_C$ for each of the 35 equivalence classes in $\mathcal{C}'_q$, simplifying and collecting coefficients.

We include in Appendix B an explanation of a computer program used to compute $\tilde{S}_C$ for each local view C. A version of the program in the Sage mathematical language and its output are available as ancillary files in the arXiv submission of this paper. Each of the steps the program performs are readily achievable by hand, though the number of steps and the size of the polynomials involved make this unappealing.

The output shows a list of $\tilde{S}_C$ for all 35 non-isomorphic local views, demonstrating that it is zero for C_1 and C_2 and a non-zero polynomial in r and t with non-negative coefficients for all other C. It also shows $\tilde{S}_C$ evaluated at $r = -1$ for local views C which use at most 2 colors on the boundary, yielding a non-zero polynomial in t with non-negative coefficients for all such C except C_1 and C_2, as desired.

3.2. **Maximizing.** To show that K_4 is the unique maximizer of the LP is somewhat more straightforward. Since the distribution yielding K_4 as a local view with probability one is feasible in the LP, it suffices to show that $U^v_{K_4} > U^v_C$ for all $C \neq K_4$.

Claim 4. *Let*

$$D^v_C = 2(1+t)^{14} Z_{K_4} Z_C t^{-2} \left(U^v_{K_4} - U^v_C\right). \tag{6}$$

Then for all $C \in \mathcal{C}_q$, D^v_C is a polynomial $t = e^\beta - 1 = 1/\lambda - 1$ and $s = q - \max\{3, q_C\}$ with all positive coefficients, and identically 0 if and only if $C = K_4$.

Since local views with $q_C > q$ cannot occur, for $q \geq 3$ and $\beta > 0$ we have $s \geq 0$ and $t \geq 0$ and hence Claim 4 implies $U^v_{K_4} > U^v_C$ for all $C \neq K_4$. The quantity D^v_C is listed for all 35 non-isomorphic local views in Appendix A. Again, for $q = 2$ we must do more; for $C \in \mathcal{C}'_2$ we list D^v_C evaluated at $q = 2$, observing that it is a polynomial in t with non-negative coefficients, except for K_4 where it is zero.

As with the computations for $\tilde{S}_C$, we use a computer to multiply polynomials and obtain D^v_C for each local view, see Appendix B.

4. Extensions to $d \geq 4$?

How might we extend Theorem 1 to graphs of larger degree? The minimization program defined above in Section 3 is not tight in general: we can in fact see that it is underconstrained by comparing the number of constraints (2) to the number of equivalence classes of local views in the support of the distribution induced by $K_{d,d}$, which is the partition number of $d-1$ when $q \geq d-1$, and always more than 2 if $d \geq 4$ and $q \geq 2$.

There is a large family of constraints that we can add to the program. Let $\mathcal{S}_{q,d}$ be the set of all *q-partitions* of size d; that is, partitions of d into at most q parts which we represent by vectors of length q with non-negative integer entries that sum to d, written in non-decreasing order. Any q-coloring χ of d vertices induces a q-partition; for instance if χ assigns the colors $\{1,4,2,2,1,2\}$, then the q-partition $H(\chi) = \{3,2,1,0\} \in \mathcal{S}_{4,6}$. Our family of constraints will be that for every $S \in \mathcal{S}_{q,d}$, the probability that the neighbors of v receive a coloring with q-partition S equals the average probability of the same for a neighbor of v.

Both of these probabilities can be computed as expectations over the random local view. For a local view C and a q-partition $S \in \mathcal{S}_{q,d}$ we define

$$\gamma_C^{v,S} := \frac{1}{Z_C} \sum_{\chi \in [q]^{V_C}} \mathbf{1}_{\{H(\chi(N(v)))=S\}} \cdot e^{-\beta m(\chi)}\,,$$

$$\gamma_C^{N,S} := \frac{1}{d}\frac{1}{Z_C} \sum_{\chi \in [q]^{V_C}} \sum_{u \in N(v)} \mathbf{1}_{\{H(\chi(N(u)))=S\}} \cdot e^{-\beta m(\chi)}\,.$$

Observe that for any graph and any q-partition S, we must have

$$\mathbb{E}_C[\gamma_C^{v,S}] = \mathbb{E}_C[\gamma_C^{N,S}]\,.$$

Our minimization program becomes

$$\begin{aligned} U^{\min} = \text{Minimum of } & \sum_C p_C U_C^v \qquad \text{subject to} \\ & p_C \geq 0 \ \forall C\,, \\ & \sum_C p_C = 1\,, \\ & \sum_C p_C(\gamma_C^{v,S} - \gamma_C^{N,S}) = 0 \ \text{ for all } S \in \mathcal{S}_{q,d}. \end{aligned}$$

This program is at least as strong as the one used in Section 3: the q-partition constraints together imply the constraint $\mathbb{E}_C[U_C^v] = \mathbb{E}_C[U_C^N]$.

We can solve this program for small values of d and fixed β, which leads us to the following conjecture.

Conjecture 4. *The above minimization LP is tight for $d \geq 3, q \geq d+1$ and all $\beta > 0$, and shows that*

$$U^q_{K_{d,d}}(\beta) \leq U^q_G(\beta)$$

for all d-regular G.

However we can also find values of β for $d \geq 4, q \leq d$ so that $U^{\min} < U^q_{K_{d,d}}$, and so we believe that this program is not tight in these cases.

5. 2-regular graphs and other easy cases

Theorem 1 shows that $K_{3,3}$ is optimal on the level of internal energy per particle in the Potts model, and by Corollary 1 it maximizes $\frac{1}{|V(G)|}\log C_q(G)$ over cubic graphs G. For arbitrary d, in the case $q = 2$, the fact that $K_{d,d}$ maximizes $\frac{1}{|V(G)|}\log C_q(G)$ over d-regular G follows simply from the observation that $K_{d,d}$ is the smallest bipartite d-regular graph. Indeed for $q = 2$, if G is not bipartite then $C_q(G) = 0$ and if G is bipartite $C_q(G) = 2^{\#\text{ connected components of } G}$.

For $d = 2$, the only d-regular connected graphs are cycles, and there is an explicit formula for the q-color Potts partition function of the n-cycle. In the language of statistical physics the 1-dimensional Potts model (including the 0-temperature Potts model) is *exactly solvable*:

$$Z^q_{C_n}(\beta) = (q-1)(e^{-\beta}-1)^n + (e^{-\beta}+q-1)^n.$$

One way to obtain this formula is to use the mapping of the Tutte polynomial $T(x,y)$ to the Potts partition function, given in e.g. [24], and then using the formula $T_{C_n}(x,y) = \frac{x^n-x}{x-1} + y$.

Taking the logarithmic derivative gives:

$$U^q_{C_n}(\beta) = \frac{e^{-\beta}}{e^{-\beta}-1} \cdot \frac{\left(1+\frac{q}{e^{-\beta}-1}\right)^{n-1} + q - 1}{\left(1+\frac{q}{e^{-\beta}-1}\right)^{n} + q - 1} \tag{7}$$

Proposition 1. *If $\beta > 0$ then*

$$U^q_{C_n}(\beta) > U^q_{C_{n+2}}(\beta) \text{ for } n \geq 3 \text{ odd},$$
$$U^q_{C_n}(\beta) < U^q_{C_{n+2}}(\beta) \text{ for } n \geq 4 \text{ even}.$$

If $\beta < 0$ then

$$U^q_{C_n}(\beta) > U^q_{C_{n+1}}(\beta) \text{ for all } n \geq 3.$$

Proof. Let $\beta > 0$ and suppose that $n \geq 3$ is odd. Let $x := 1 + \frac{q}{e^{-\beta}-1}$. By (7), we then have that $U^q_{C_n}(\beta) > U^q_{C_{n+2}}$ if and only if

$$x^n + x^{n+1} > x^{n-1} + x^{n+2}. \tag{8}$$

Since n is odd, (8) holds if and only if $x + x^2 > 1 + x^3$ which holds since $x < -1$. For even $n \geq 4$, the proof is the same.

Suppose now that $\beta < 0$ and $n \geq 3$. Defining x as before, we have that $x > 0$. In this case, the inequality $U^q_{C_n}(\beta) > U^q_{C_{n+1}}(\beta)$ simply reduces to the inequality $x^{n-1}(1-x)^2 > 0$. □

Letting $\mathbb{Z}_1$ denote the infinite line we have that

$$U^q_{\mathbb{Z}_1}(\beta) = \frac{e^{-\beta}}{e^{-\beta}+q-1} = \lim_{n\to\infty} U^q_{C_n}(\beta). \tag{9}$$

Taking the limit here is justified as the Potts model on $\mathbb{Z}_1$ is in the Gibbs uniqueness regime for all $q, \beta > 0$.

Corollary 2. *Conjectures 1, 2, and 3 hold for $d = 2$. Moreover, If $\beta > 0$ then*

$$U^q_{C_n}(\beta) > U^q_{\mathbb{Z}_1}(\beta) \text{ for } n \geq 3 \text{ odd},$$
$$U^q_{C_n}(\beta) < U^q_{\mathbb{Z}_1}(\beta) \text{ for } n \geq 4 \text{ even}.$$

If $\beta < 0$ then

$$U^q_{C_n}(\beta) > U^q_{\mathbb{Z}_1}(\beta) \text{ for all } n \geq 3.$$

Proof. This follows from Proposition 1 and (9). □

Acknowledgements

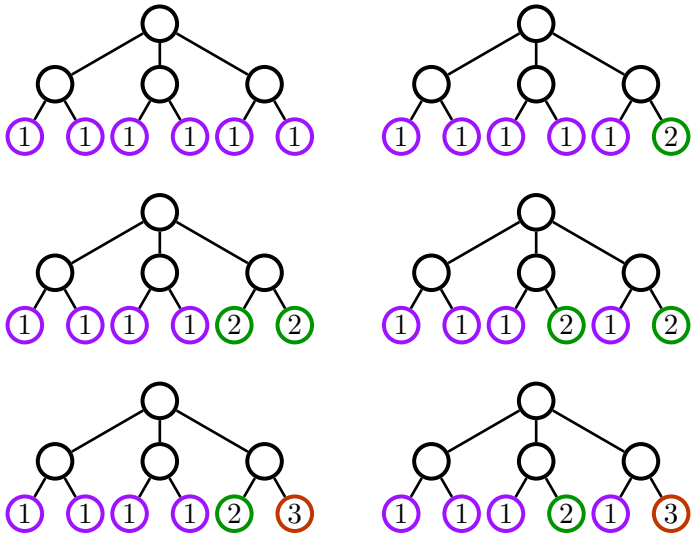
We thank Emma Cohen, Prasad Tetali, and Yufei Zhao for enjoyable and enlightening conversations about graph homomorphism problems.

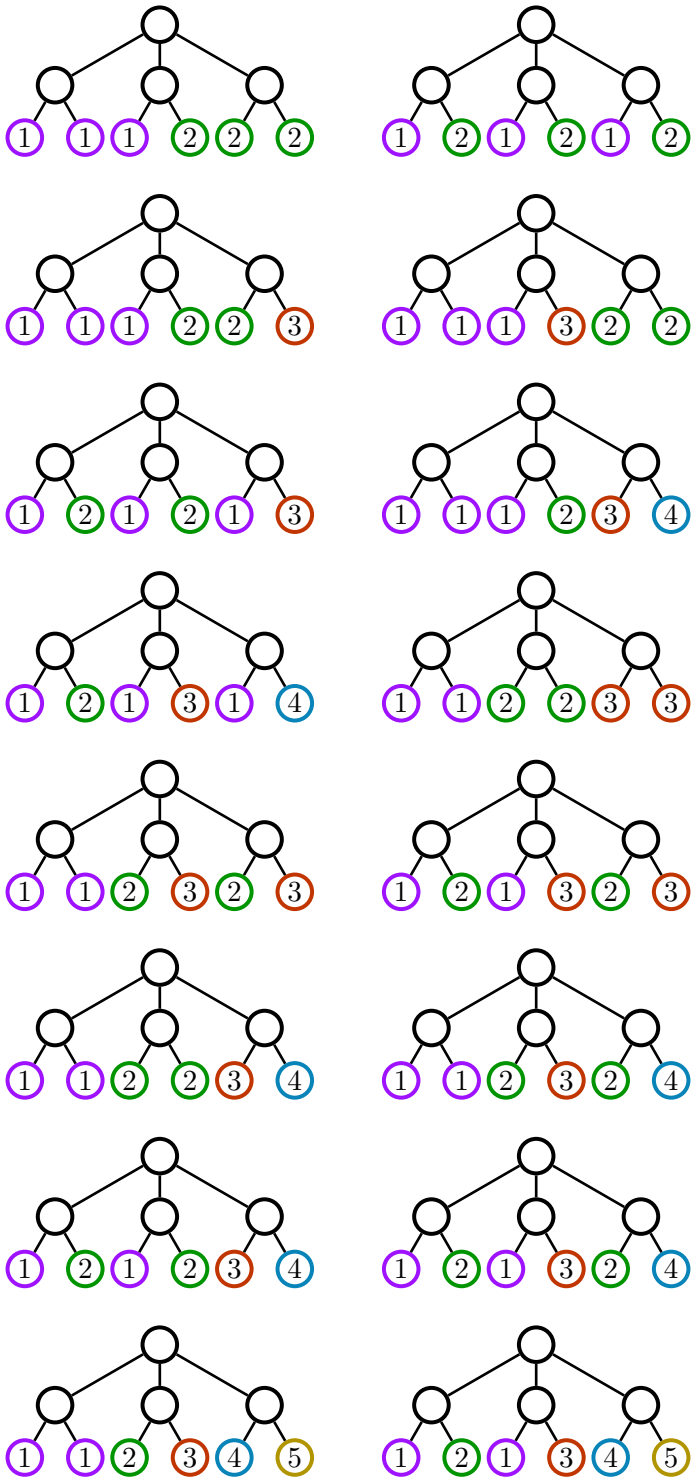
References

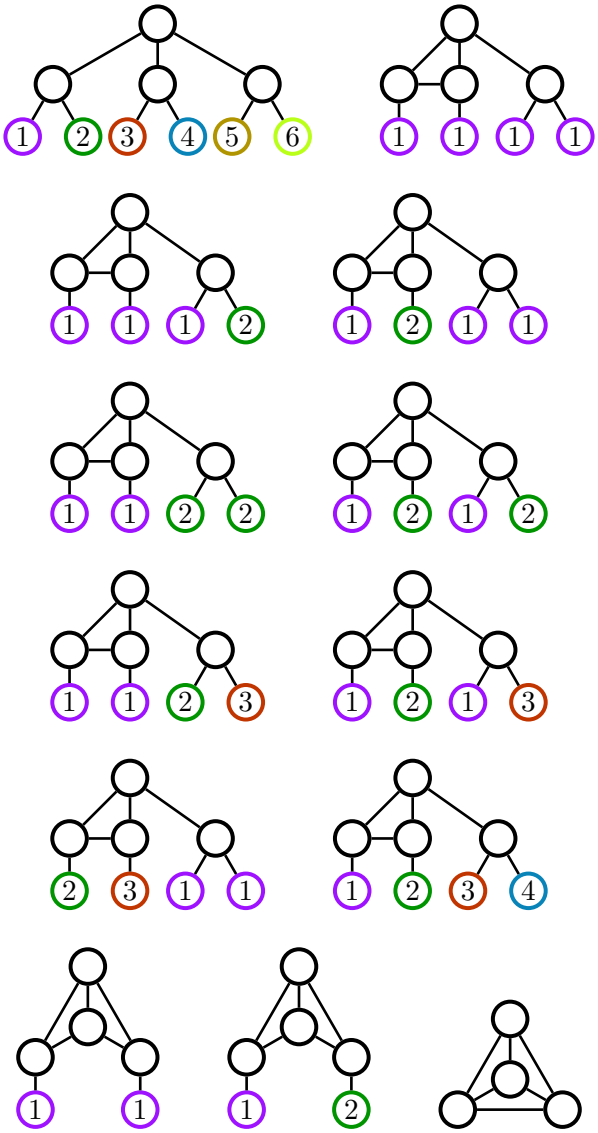
[1] E. Cohen, W. Perkins, and P. Tetali. On the widom–rowlinson occupancy fraction in regular graphs. *Combinatorics, Probability and Computing*, 26(2):183–194, 2017.
[2] P. Csikvári. Extremal regular graphs: the case of the infinite regular tree. *arXiv preprint arXiv:1612.01295*, 2016.
[3] P. Csikvári and Z. Lin. Sidorenko's conjecture, colorings and independent sets. *Electronic Journal of Combinatorics*, to appear.
[4] E. Davies, M. Jenssen, W. Perkins, and B. Roberts. Independent sets, matchings, and occupancy fractions. *arXiv preprint arXiv:1508.04675*, 2015.
[5] E. Davies, M. Jenssen, W. Perkins, and B. Roberts. On the average size of independent sets in triangle-free graphs. *Proceedings of the American Mathematical Society*, to appear.
[6] D. Galvin. Bounding the partition function of spin-systems. *Electronic Journal of Combinatorics*, 13(3):R72, 2006.
[7] D. Galvin. Maximizing H-colorings of a regular graph. *Journal of Graph Theory*, 73(1):66–84, 2013.
[8] D. Galvin. Counting colorings of a regular graph. *Graphs and Combinatorics*, 31(3):629–638, 2015.
[9] D. Galvin and P. Tetali. On weighted graph homomorphisms. *DIMACS Series in Discrete Mathematics and Theoretical Computer Science*, 63:97–104, 2004.
[10] E. Ising. Beitrag zur theorie des ferromagnetismus. *Zeitschrift für Physik A Hadrons and Nuclei*, 31(1):253–258, 1925.
[11] J. Kahn. An entropy approach to the hard-core model on bipartite graphs. *Combinatorics, Probability and Computing*, 10(03):219–237, 2001.
[12] F. Lazebnik. On the greatest number of 2 and 3 colorings of a (v,e)-graph. *Journal of Graph Theory*, 13(2):203–214, 1989.
[13] F. Lazebnik. New upper bounds for the greatest number of proper colorings of a (V, E)-graph. *Journal of Graph Theory*, 14(1):25–29, 1990.

[14] F. Lazebnik, O. Pikhurko, and A. Woldar. Maximum number of colorings of $(2k, k^2)$-graphs. *Journal of Graph Theory*, 56(2):135–148, 2007.
[15] N. Linial. Legal coloring of graphs. In *Foundations of Computer Science, 1983., 24th Annual Symposium on*, pages 470–472. IEEE, 1983.
[16] P.-S. Loh, O. Pikhurko, and B. Sudakov. Maximizing the number of q-colorings. *Proceedings of the London Mathematical Society*, 101(3):655–696, 2010.
[17] M. Mezard and A. Montanari. *Information, physics, and computation*. Oxford University Press, 2009.
[18] G. Perarnau and W. Perkins. Counting independent sets in cubic graphs of given girth. *arXiv preprint arXiv:1611.01474*, 2016.
[19] R. B. Potts. Some generalized order-disorder transformations. *Mathematical proceedings of the Cambridge Philosophical Society*, 48(01):106–109, 1952.
[20] N. Ruozzi. The Bethe partition function of log-supermodular graphical models. In *Advances in Neural Information Processing Systems*, pages 117–125, 2012.
[21] N. Ruozzi. Beyond log-supermodularity: lower bounds and the Bethe partition function. In *Proceedings of the Twenty-Ninth Conference on Uncertainty in Artificial Intelligence*, pages 546–555. AUAI Press, 2013.
[22] A. D. Sokal. Bounds on the complex zeros of (di)-chromatic polynomials and Potts-model partition functions. *Combinatorics, Probability and Computing*, 10(01):41–77, 2001.
[23] A. D. Sokal. The multivariate Tutte polynomial (alias Potts model) for graphs and matroids. *Surveys in combinatorics*, 327:173–226, 2005.
[24] D. J. Welsh and C. Merino. The Potts model and the Tutte polynomial. *Journal of Mathematical Physics*, 41(3):1127–1152, 2000.
[25] F.-Y. Wu. The Potts model. *Reviews of modern physics*, 54(1):235, 1982.
[26] Y. Zhao. The bipartite swapping trick on graph homomorphisms. *SIAM Journal on Discrete Mathematics*, 25(2):660–680, 2011.
[27] Y. Zhao. Extremal regular graphs: independent sets and graph homomorphisms. *arXiv preprint arXiv:1610.09210*, 2016.

Appendix A. List of local views




1 1 1 2 2 2
1 2 1 2 1 2
1 1 1 2 2 3
1 1 1 3 2 2
1 2 1 2 1 3
1 1 1 2 3 4
1 2 1 3 1 4
1 1 2 2 3 3
1 1 2 3 2 3
1 2 1 3 2 3
1 1 2 2 3 4
1 1 2 3 2 4
1 2 1 2 3 4
1 2 1 3 2 4
1 1 2 3 4 5
1 2 1 3 4 5



Appendix B. Computing properties of local views

The verification that, for each of the local views shown in Appendix A, the scaled slack (5) and the scaled difference (6) are non-negative (and zero where required) was done with the aid of a computer. Here we describe some additional considerations required to perform this verification for arbitrary q and β.

As noted in Section 3, the number of equivalence classes of local views we must consider is bounded independently of q, and we only consider representatives of each equivalence class that use an initial segment of colors from $\{1, \dots, 6\}$ on the boundary. In order to compute the partition function and other properties of a local view C, one is required to consider the q^4

possible local colorings of V_C. This too can be done in a way that is bounded independently of q by considering equivalence classes of local colorings.

Let C be a local view (that uses an initial segment of $\{1,\dots,6\}$ on the boundary) and recall that q_C is the largest color appearing on the boundary of C. Then given a local coloring χ of V_C, we can only see at most q_C+4 colors. After permuting colors not used on the boundary, we may assume that χ consists only of colors in $[q_C]$ and initial segment of $\{q_C+1,\dots,q_C+4\}$ (which may be empty). This means we are considering equivalence classes of local colorings and choosing a representative $\tilde{\chi}$ of each class such that, together with the colors on the boundary, we only ever color C with an initial segment of $[q_C+4]$.

Then for arbitrary q it suffices to consider at most $q_C+4\le 10$ colors in the calculations for Z_C, U_C^v, and U_C^N. Given the set $Q_C\subset[q_C+4]^{V_C}$ of *representative* local colorings $\tilde{\chi}$ such that $\tilde{\chi}$ uses an initial segment (which may be empty) of the colors $\{q_C+1,\dots,q_C+4\}$, and writing ℓ for the largest color used in $\tilde{\chi}$, the Potts model on C induces the distribution

$$\tilde{\chi}\mapsto\begin{cases}\frac{e^{-\beta m(\tilde{\chi})}}{Z_C}\binom{q-q_C}{\ell-q_C} & \text{if } \ell>q_C\\ \frac{e^{-\beta m(\tilde{\chi})}}{Z_C} & \text{otherwise}\end{cases}$$

on $\tilde{\chi}\in Q_C$.

The consideration of these equivalence classes of local colorings means that for any β, q, and C, the quantities Z_C, U_C^v, and U_C^N may be computed by summing over Q_C whose size is bounded independently of q. Using this simplification, we used a *SageMath*[1] computer program (hosted with the arXiv version of this paper) to compute the scaled slack function $\tilde{S}_C$ and the scaled difference D_C^v for each of the 35 local views. The program can be used to generate a document (included as a supplementary file) containing all the required polynomials so that the reader may verify the proof. In addition the program can check itself for non-negative coefficients and print these observations on request.

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

University of Birmingham
E-mail address: math@willperkins.org

[1] http://www.sagemath.org