

# INAPPROXIMABILITY OF COUNTING HYPERGRAPH COLOURINGS

ANDREAS GALANIS, HENG GUO, JIAHENG WANG

**ABSTRACT.** Recent developments in approximate counting have made startling progress in developing fast algorithmic methods for approximating the number of solutions to constraint satisfaction problems (CSPs) with large arities, using connections to the Lovász Local Lemma. Nevertheless, the boundaries of these methods for CSPs with non-Boolean domain are not well-understood. Our goal in this paper is to fill in this gap and obtain strong inapproximability results by studying the prototypical problem in this class of CSPs, hypergraph colourings.

More precisely, we focus on the problem of approximately counting  $q$ -colourings on  $K$ -uniform hypergraphs with bounded degree  $\Delta$ . An efficient algorithm exists if  $\Delta \lesssim \frac{q^{K/3-1}}{4^K K^2}$  (Jain, Pham, and Vuong, 2021; He, Sun, and Wu, 2021). Somewhat surprisingly however, a hardness bound is not known even for the easier problem of finding colourings. For the counting problem, the situation is even less clear and there is no evidence of the right constant controlling the growth of the exponent in terms of  $K$ .

To this end, we first establish that for general  $q$  computational hardness for finding a colouring on simple/linear hypergraphs occurs at  $\Delta \gtrsim Kq^K$ , almost matching the algorithm from the Lovász Local Lemma. Our second and main contribution is to obtain a far more refined bound for the counting problem that goes well beyond the hardness of finding a colouring and which we conjecture is asymptotically tight (up to constant factors). We show in particular that for all even  $q \geq 4$  it is NP-hard to approximate the number of colourings when  $\Delta \gtrsim q^{K/2}$ . Our approach is based on considering an auxiliary weighted binary CSP model on graphs, which is obtained by “halving” the  $K$ -ary hypergraph constraints. This allows us to utilise reduction techniques available for the graph case, which hinge upon understanding the behaviour on random regular bipartite graphs that serve as gadgets in the reduction. The major challenge in our setting is to analyse the induced matrix norm of the interaction matrix of the new CSP which captures the most likely solutions of the system. In contrast to previous analyses in the literature, the auxiliary CSP demonstrates both symmetry and asymmetry, making the analysis of the optimisation problem severely more complicated and demanding the combination of delicate perturbation arguments and careful asymptotic estimates.

## 1. INTRODUCTION

Constraint satisfaction problems (CSPs), such as  $q$ -colourings and  $k$ -SAT, are perhaps the most-well studied problems in computer science. We consider the case where the number of appearances of variables (also called degrees) is bounded by some absolute constant. For this class of CSPs, the Lovász local lemma [EL75] is a classical tool in combinatorics that provides sharp degree thresholds under which the existence of solutions to CSPs is guaranteed. After a long line of research [Bec91, Alo91, MR98, CS00, Sri08, Mos09], Moser and Tardos [MT10] showed that, under the same conditions as the local lemma, an efficient algorithm exists to find a solution. One remarkable aspect of this algorithm in the case of the bounded-degree  $k$ -SAT problem is that it gives up to lower order terms [GST16] the location of the algorithmic threshold for finding solutions [KST93], as the degree varies.

A related computational problem that has been intensively studied recently is to efficiently sample or approximately count solutions,<sup>1</sup> instead of merely finding one. Under local lemma type conditions, there are some barriers, such as the connectivity barrier, for traditional approaches to approximate counting

---

(Andreas Galanis) DEPARTMENT OF COMPUTER SCIENCE, UNIVERSITY OF OXFORD, WOLFSON BUILDING, PARKS ROAD, OXFORD, OX1 3QD, UNITED KINGDOM. E-mail: [andreas.galanis@cs.ox.ac.uk](mailto:andreas.galanis@cs.ox.ac.uk)

(Heng Guo, Jiaheng Wang) SCHOOL OF INFORMATICS, UNIVERSITY OF EDINBURGH, INFORMATICS FORUM, EDINBURGH, EH8 9AB, UNITED KINGDOM. E-mail: [hguo@inf.ed.ac.uk](mailto:hguo@inf.ed.ac.uk), [jiaheng.wang@ed.ac.uk](mailto:jiaheng.wang@ed.ac.uk)

This project has received funding from the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation programme (grant agreement No. 947778).

<sup>1</sup>Approximate counting and sampling are often computationally equivalent, for example in the so-called “self-reducible” settings [JVV86]. The local lemma setting is typically not self-reducible. However, reductions still exist [FGYZ21, JPV21b] between approximate counting and sampling without degradation of the parameters.

and sampling. Recently, there have been some surprising developments that managed to bypass these barriers, making great progress towards a sampling version for the local lemma [Moi19, GJL19, GLLZ19, FGYZ21, FHY20, JPV21b, JPV21a, HSW21]. However, there is no sharp threshold established in the sampling setting yet.

A major difference between searching and sampling is that sampling can be computationally harder in the local lemma settings. For example, for  $K$ -CNF formulas where each variable appears at most  $\Delta$  times, if  $\Delta \leq \frac{2^K}{eK} + 1$ , then there must be a satisfying assignment, and it can be efficiently found; yet if  $\Delta \geq 5 \cdot 2^{K/2}$ , there is no algorithm to sample or approximately count satisfying assignments unless  $\text{NP} = \text{RP}$  [BGG<sup>+</sup>19], even when all variables appear positively (the monotone case). For monotone  $K$ -CNFs, the threshold  $2^{K/2}$  is sharp up to constants, because Hermon, Sly, and Zhang [HSZ19] showed a complementary efficient algorithm for  $\Delta \leq c2^{K/2}$  where  $c$  is a constant.

Our goal in this paper is to detail how this “sampling-is-computationally-harder” phenomenon manifests into local-lemma-type hypergraph problems with non-boolean domain and which are not necessarily monotone, and to make progress towards carving the computational thresholds for sampling problems in the local lemma setting. Among the more recent algorithmic developments, the most promising one to establish a computational transition is the problem of counting hypergraph colourings, which was the original setting where the local lemma was developed [EL75]. For a hypergraph  $H = (V, E)$ , a proper  $q$ -colouring  $\sigma : V \rightarrow \{1, 2, \dots, q\}$  assigns a colour to each vertex, such that no hyperedge is monochromatic under  $\sigma$ . Suppose further that  $H$  is  $K$ -uniform and  $\Delta$  is the maximum degree of  $H$ . In this setting, an efficient sampling algorithm exists if  $\Delta \lesssim \frac{q^{K/3-1}}{4^K K^2}$  [JPV21a, HSW21], where  $\lesssim$  (and similarly,  $\gtrsim$ ,  $\asymp$ ) hides logarithm or other lower order terms.

Somewhat surprisingly, despite being a canonical and well-studied problem, not much is known regarding the computational hardness for bounded-degree hypergraph colourings, even for the search version. Thus we first show that it is NP-hard to find a proper hypergraph colouring if  $q \geq 2$ ,  $K \geq 2$  (but not  $q = K = 2$ ), and  $\Delta \gtrsim Kq^K$  (see Theorem 2.2), and to approximately count if  $q \geq 2$ ,  $K \geq 4$ , and  $\Delta \gtrsim Kq^{K-1}$  (see Theorem 2.4). These bounds almost match the algorithmic Local Lemma threshold. In fact, Theorem 2.2 and Theorem 2.4 still hold when restricting to simple<sup>3</sup> hypergraphs. If we restrict monotone  $K$ -CNFs to simple hypergraphs, the condition of the aforementioned algorithmic result [HSZ19] improves to  $\Delta \lesssim \frac{2^K}{K^2}$ . In view of this result and the searching algorithm [MT10], it seems reasonable to conjecture that the sharp hardness threshold (for both approximate counting and searching) in simple hypergraphs is  $\Delta \gtrsim q^{K-1}$ , up to some polynomial factors in  $K$ . Our hardness result almost matches it.

Our second and main contribution is a more refined hardness result for approximate counting and sampling, stated as follows.

**Theorem 1.1.** *Let  $q \geq 4$  be even,  $K \geq 4$  be even, and  $\Delta \geq 5q^{K/2}$ . It is NP-hard to approximate the number of proper  $q$ -colourings in  $n$ -vertex  $K$ -uniform hypergraphs of maximum degree at most  $\Delta$ , even within a factor of  $2^{c_n}$  for some constant  $c(q, K) > 0$ .*

A few remarks are in order.

- First, the threshold in Theorem 1.1 is far more refined than the corresponding theorem (Theorem 2.2) for the searching problem. The exponent of the sampling threshold we achieve is roughly half of that of the threshold for the searching problem, which is analogous to the aforementioned (monotone)  $K$ -CNF example [BGG<sup>+</sup>19]. Interestingly, and in contrast to the  $K$ -CNF case, our colourings threshold is getting close to matching the constant in the exponent in the algorithmic threshold of  $\Delta \lesssim \frac{q^{K/3-1}}{4^K K^2}$  [JPV21a, HSW21]. We conjecture that our hardness threshold is asymptotically tight (up to constant factors), namely that for all  $q \geq 2$  and  $K \geq 2$ , there is an efficient algorithm to approximately count the number of proper  $q$ -colourings in  $K$ -uniform hypergraphs whenever  $\Delta \lesssim q^{K/2}$ .

<sup>2</sup>Note that in [JPV21a, HSW21], their  $\Delta$  is the degree bound for the dependency graph, which is at most  $\Delta(K-1)$  in our setting.

<sup>3</sup>A hypergraph is called *simple* (or *linear*) if the intersection of any two hyperedges has size at most 1.

- Second, our result applies to only even  $K$  for  $K$ -uniform hypergraphs. This is due to a particular halving construction we use in the reduction. The hardness results for (monotone)  $K$ -CNF [BGG<sup>+</sup>19] allow hyperedges with sizes at least  $K$ . This is a stronger assumption and our hardness bound would still apply without changing the order. In fact, we expect a slight variant of our construction to work for odd  $K$  to achieve a threshold of the same order. (See Remark 3.2.) As we explain soon, the details for even  $K$  are already very complicated, so for clarity and simplicity we did not pursue the odd  $K$  case.
- Lastly, our result applies to an even number of colours  $q$ , which is analogous to hardness results for counting in the graph colouring setting [GŠV15]. It was left as an open problem in [GŠV15] to handle odd  $q$  (see also the recent work [CGSV21]), and we met the same difficulty in our setting as well. Our hardness proof for counting builds on ideas from [GŠV15], and we focus on the challenges needed to refine them in the hypergraph setting (rather than addressing the parity of  $q$ ). We expect that substantial new ideas are required to resolve the parity of  $q$ , even in the graph setting.

In order to show Theorem 1.1, we first reduce from an auxiliary weighted binary CSP, namely a “spin system” in graphs. Basically, we replace each vertex of the graph by a cluster of  $K/2$  vertices in the hypergraph, and an edge by a hyperedge of size  $K$ . This construction is identical to the one in [BGG<sup>+</sup>19], via which one reduces from weighted independent set in graphs to hypergraph independent sets. However, in order to reduce to the hypergraph  $q$ -colouring problem, the variables of the weighted binary CSP take  $q + 1$  possible values. There are  $q$  values that correspond to “pure” colours, and one special value that corresponds to a “mixed” state. The interactions among these  $q + 1$  states are dictated by the hypergraph colouring problem, and the mixed state behaves very differently from the pure colours; roughly, the pure colours behave symmetrically (as in the graph case) but the mixed state causes asymmetry.

Our next and main step is to show the desired hardness result for this spin system. We follow an established route of establishing inapproximability for spin systems [DFJ02, MWW09, Sly10, CCGL12, SS14, GŠV16], and in particular [GŠV15], where the key is to understand the system on random regular bipartite graphs which are used as gadgets in the reduction. More precisely, we need to analyze what are the most likely configurations of the system on random regular bipartite graphs, the so-called dominant phases (given by the normalised counts of the colours on each side of the graph). It was shown in [GŠV15] that these are captured by a certain matrix norm of the interaction matrix. These norms are in general very hard to penetrate analytically and it was already a major difficulty in the perfectly symmetric setting of [GŠV15]. For us, the presence of a special spin together with  $q$  symmetric spins makes our spin system very different from all of the spin systems analyzed before and the mixture of symmetry and asymmetry make the analysis substantially harder. For example, in [GŠV15], to show that the two parts of the graph are unbalanced, a simple Hessian calculation suffices, whereas in our setting, there are Hessian stable balanced phases due to the presence of this special spin (that can be favoured against the others). Also, being Hessian stable means that this phase is locally maximal, making perturbation arguments hard to carry out. What we do instead is to directly compare this phase with the dominant phase via a careful interpolation path and a sequence of delicate estimates. This reflects the main difference between our work and previous works, namely that our estimates and perturbation arguments are significantly more delicate in order to rule out the local-maxima.

The main open problem left is to close the gap between the algorithm of [JPV21a, HSW21] and our hardness threshold, Theorem 1.1, although we expect that any progress towards a computational transition threshold now would come from the algorithmic side. Another open problem is to handle the odd  $q$  case in Theorem 1.1 and similarly in [GŠV15]. When  $q$  is odd, the current perturbation-based analysis cannot determine, among a few candidates, which phases are dominant. New ideas would be required to handle the odd  $q$  case.

## 2. HARDNESS FOR FINDING COLOURINGS IN SIMPLE HYPERGRAPHS

In this section we show hardness results for finding hypergraph colourings for parameters beyond the local lemma condition. The key is to find instances that do not have proper colourings.

We will use a configuration model to construct random regular hypergraphs. With constant probability, the resulting hypergraph is simple [CFMR96, PP19]. Frieze and Mubayi [FM13] showed that if  $q > c \left( \frac{\Delta}{\log \Delta} \right)^{\frac{1}{K-1}}$  for some constant  $c = c(K)$  that only depends on  $K$ , then any simple  $K$ -uniform hypergraph with maximum degree  $\Delta$  is  $q$ -colourable. In particular, their condition holds if  $\Delta \leq cKq^{K-1} \ln q$  for some constant  $c = c(K)$ . Our next lemma complements their result by showing as an intermediate result that if  $\Delta > Kq^{K-1} \ln q + 1$ , we can find a  $K$ -uniform hypergraph with maximum degree  $\Delta$  which is not  $q$ -colourable. For our reductions, we use such hypergraphs to obtain a “disequality” gadget, as detailed in the lemma below.

**Lemma 2.1.** *Let  $q, K \geq 2$  be integers. Then, for all integers  $\Delta > Kq^{K-1} \ln q + 1$ , there exists a  $q$ -colourable  $K$ -uniform simple hypergraph  $H$  with maximum degree  $\Delta$  and two distinct vertices  $u, v$  such that the degree of  $u$  is 1, the degree of  $v$  is at most  $\Delta$ , and for every  $q$ -colouring  $\sigma$  of  $H$  it holds that  $\sigma(u) \neq \sigma(v)$ .*

*Proof.* We first argue that for all  $\Delta > Kq^{K-1} \ln q$  there is a  $\Delta$ -regular hypergraph  $H_0$  such that  $Z_{col}(H_0) = 0$ , where  $Z_{col}(H)$  denotes the number of  $q$ -colourings in  $H$ .

Let  $n$  be such that  $m = n\Delta/K$  is an integer. We sample a  $K$ -uniform  $\Delta$ -regular hypergraph  $H$  according to the following pairing model (see [PP19]). Start with a bipartite graph with the points  $[n] \times [\Delta]$  on the left and the points  $[m] \times [K]$  on the right, and pair the two sides using a uniformly random perfect matching; the vertex set of the final hypergraph  $H$  is obtained in the natural way by projecting the set  $[n] \times [\Delta]$  onto  $[n]$ . Note that it will be convenient to view the hyperedges of  $H$  for now as ordered tuples rather than sets; this does not make any difference when considering colourings of  $H$  due to the symmetry among possible ordering of the colours within the hyperedge. It is a well-known fact, see for example [CFMR96, Lemma 2] or [PP19, Theorem 2.4 & Appendix A.4], that the probability that  $H$  is simple is bounded away from zero for all sufficiently large  $n$ .<sup>4</sup>

For a colouring  $\sigma : [n] \rightarrow [q]$ , a colour  $i \in [q]$  and a  $K$ -tuple of colours  $\mathbf{i} = (i_1, \dots, i_K) \in [q]^K$ , let  $n\alpha_i$  be the number of vertices with colour  $i$ , and  $m\beta_{\mathbf{i}}$  be the number of hyperedges whose vertices are coloured according to  $\mathbf{i}$  (i.e., the  $j$ -th vertex of the hyperedge takes the colour  $i_j$ ). Let  $\boldsymbol{\alpha} = \{\alpha_i\}_{i \in [q]}$  and  $\boldsymbol{\beta} = \{\beta_{\mathbf{i}}\}_{\mathbf{i} \in [q]^K}$ , and note that  $(\boldsymbol{\alpha}, \boldsymbol{\beta}) \in S_q$ , where  $S_q$  is the space of all pairs of vectors in  $\mathbb{R}^q \times \mathbb{R}^{q^K}$  satisfying

$$(1) \quad \begin{aligned} \sum_{i \in [q]} \alpha_i &= 1, \quad \sum_{\mathbf{i} \in [q]^K} t_{\mathbf{i}, i} \beta_{\mathbf{i}} = K\alpha_i \text{ for } i \in [q] \\ \alpha_i &\geq 0 \text{ for } i \in [q], \quad \beta_{\mathbf{i}} \geq 0 \text{ for } \mathbf{i} \in [q]^K, \quad \beta_{(i, i, \dots, i)} = 0 \text{ for } i \in [q], \end{aligned}$$

where for  $i \in [q]$  and  $\mathbf{i} \in [q]^K$  we denote by  $t_{\mathbf{i}, i}$  the number of occurrences of colour  $i$  in the tuple  $\mathbf{i}$ . Then, we have

$$\mathbb{E}[Z_{col}(H)] = \frac{1}{(\Delta n)!} \sum_{(\boldsymbol{\alpha}, \boldsymbol{\beta}) \in S_q; n\boldsymbol{\alpha} \in \mathbb{Z}^q, m\boldsymbol{\beta} \in \mathbb{Z}^{q^K}} \binom{n}{\alpha_1 n, \dots, \alpha_q n} \binom{m}{\beta_1 m, \dots, \beta_{q^K} m} \prod_{i \in [q]} (\Delta \alpha_i n)!,$$

since a term in the sum corresponding to  $(\boldsymbol{\alpha}, \boldsymbol{\beta})$  accounts for the number of ways to choose  $\sigma$  and  $H$  with vertex-colour frequencies given by the vector  $\boldsymbol{\alpha}$  and edge-colour frequencies given by the vector  $\boldsymbol{\beta}$ . Using Stirling’s approximation  $(2\pi k)^{1/2} (k/e)^k \leq k! \leq ek^{1/2} (k/e)^k$  that holds for all integers  $k \geq 1$ , we obtain by expanding the terms inside the sum (note that there are at most  $n^{q^K+q}$  of them) that

$$(2) \quad \mathbb{E}[Z_{col}(H)] \leq n^{O(1)} \exp \left( n \max_{(\boldsymbol{\alpha}, \boldsymbol{\beta}) \in S_q} F(\boldsymbol{\alpha}, \boldsymbol{\beta}) \right),$$

where  $F(\boldsymbol{\alpha}, \boldsymbol{\beta}) = -(\Delta - 1)h(\boldsymbol{\alpha}) + \frac{\Delta}{K}h(\boldsymbol{\beta})$  and  $h(\cdot)$  is the entropy function (here, we adopt the usual convention that  $0 \ln 0 = 0$  which makes  $h$  and  $F$  continuous and therefore the maximum in (2) well-defined).

For  $(\boldsymbol{\alpha}, \boldsymbol{\beta}) \in S_q$ , we have that  $\alpha_i = \frac{1}{K} \sum_{\mathbf{i} \in [q]^K} t_{\mathbf{i}, i} \beta_{\mathbf{i}}$  for  $i \in [q]$ , and hence

$$F(\boldsymbol{\alpha}, \boldsymbol{\beta}) = h(\boldsymbol{\alpha}) + \frac{\Delta}{K}G(\boldsymbol{\alpha}, \boldsymbol{\beta}) \text{ where } G(\boldsymbol{\alpha}, \boldsymbol{\beta}) = h(\boldsymbol{\beta}) - \sum_{i \in [q]} \ln(\alpha_i) \sum_{\mathbf{i} \in [q]^K} t_{\mathbf{i}, i} \beta_{\mathbf{i}}.$$

<sup>4</sup>We remark that the term “simple” has different meanings across the literature. A simple hypergraph in this paper actually corresponds to a configuration without 4-cycles in the context of [PP19, Theorem 2.4] (where one should plug in  $\ell = 2$ ), or a hypergraph without 2-cycles in [CFMR96].

Note that for a fixed vector  $\alpha$ , the function  $G_\alpha(\beta) := G(\alpha, \beta)$  is concave and the method of Lagrange multipliers yields that the maximum of  $G_\alpha$  happens at  $\beta^* = \{\beta_i^*\}_{i \in [q]^K}$  that satisfies

$$\beta_i^* = \frac{\prod_{i \in [q]} (\alpha_i)^{t_{i,i}} \prod_{i \in [q]} 1_{i^* \neq (i, i, \dots, i)}}{1 - \|\alpha\|_K^K} \text{ for } i \in [q]^K, \quad G_\alpha(\beta^*) = \ln(1 - \|\alpha\|_K^K).$$

It follows that

$$(3) \quad F(\alpha, \beta) \leq h(\alpha) + \frac{\Delta}{K} \ln(1 - \|\alpha\|_K^K) \leq \ln \left( q \left( 1 - \frac{1}{q^{K-1}} \right)^{\Delta/K} \right),$$

where the last inequality follows from  $h(\alpha) \leq \ln q$  and  $\|\alpha\|_K^K \geq 1/q^{K-1}$ , both of which are simple applications of Jensen's inequality. For  $\Delta > Kq^{K-1} \ln q$ , the r.h.s. of (3) is negative and therefore  $\max_{(\alpha, \beta) \in S_q} F(\alpha, \beta) < 0$ . From (2), we conclude that  $Z_{col}(H) = 0$  with probability  $1 - \exp(-\Omega(n))$ . By a union bound, we obtain a simple  $\Delta$ -regular hypergraph  $H_0$  with  $Z_{col}(H_0) = 0$ , as claimed.

To obtain the final hypergraph  $H$  with the desired property, we invoke an argument in [GG16, Lemma 28] (which in turn was inspired by [KST93]). We give the details here for completeness. Given  $H_0 = (V, \mathcal{E})$  with  $Z_{col}(H_0) = 0$ , we can remove hyperedges from  $\mathcal{E}$  one by one until removing any more hyperedge makes  $Z_{col}(H) > 0$ . Call the resulting hypergraph  $H'_0 = (V, \mathcal{E}'_0)$ . Clearly  $H'_0$  is simple and has at least one hyperedge.

Choose an arbitrary hyperedge  $e \in \mathcal{E}'_0$ . Let  $S \subseteq e$  be the set of vertices with non-zero degree in  $H'_0 - e$ . If  $S = \emptyset$ , then  $e$  is disconnected from the rest of the graph. Thus as  $H'_0$  is not  $q$ -colourable, removing  $e$  would not make the hypergraph  $q$ -colourable. This contradicts to the minimality of  $H'_0$  and thus  $S \neq \emptyset$ . Denote the vertices in  $S$  by  $v_1, \dots, v_i$ , and the vertices in  $e \setminus S$  by  $v_{i+1}, \dots, v_K$ . We construct  $i$  simple hypergraphs  $H_1, \dots, H_i$  where for  $1 \leq j \leq i$ , in  $H_j$  we introduce new vertices  $u_1, \dots, u_j$  and replace the hyperedge  $e$  by  $e_j := \{u_1, \dots, u_j, v_{j+1}, \dots, v_K\}$ . By minimality of  $H'_0$  again,  $Z_{col}(H_i) > 0$  as  $e_i$  is disconnected from the rest of  $H_i$ . Thus we can find the smallest  $j \geq 1$  such that  $Z_{col}(H_j) > 0$  and  $Z_{col}(H_{j-1}) = 0$  (or  $Z_{col}(H'_0) = 0$  if  $j = 1$ ). For any proper colouring  $\sigma$  of  $H_j$ , if  $\sigma(u_j) = \sigma(v_j)$ ,  $\sigma$  would be a proper colouring of  $H_{j-1}$ , contradicting to the above. Thus it must hold that for any colouring  $\sigma$  of  $H_j$ ,  $\sigma(u_j) \neq \sigma(v_j)$ . This is the hypergraph required by the lemma, with  $u = u_j$  and  $v = v_j$ . Moreover, the degree of  $u_j$  is 1, and the degree of  $v_j$  is at most  $\Delta$ .  $\square$

Lemma 2.1 leads to the following hardness result, where we lose a factor  $2q$  in the degree bound due to the reduction. We note that for  $q = 2, K = 3, \Delta = 4$ , and simple hypergraphs, NP-hardness is known [DD20]. However, the main point of the next theorem is that there is a degree bound that scales roughly as  $q^K$  and makes the problem NP-hard.

**Theorem 2.2.** *Let  $q, K \geq 2$  be integers with  $(q, K) \neq (2, 2)$ . Then, it is NP-hard to find a  $q$ -colouring on a  $K$ -uniform simple hypergraph of maximum degree at most  $\Delta$ , when  $\Delta \geq 2Kq^K \ln q + 2q$ .*

*Proof.* For  $q > 2$ , we reduce from the problem of finding  $q$ -colourings in graphs whose degrees are bounded by  $2q$ . The latter problem is shown to be NP-hard by [EHK98]. Given a graph  $G$ , we replace each edge  $(u, v)$  of  $G$  by the hypergraph in Lemma 2.1, where  $u$  and  $v$  are identified with the special vertices in the hypergraph. Then each such hypergraph is effectively a disequality for the colours of  $u$  and  $v$ . Call the resulting hypergraph  $H$ . Thus  $G$  is  $q$ -colourable if and only if  $H$  is  $q$ -colourable. The maximum degree of  $H$  is  $2q(Kq^{K-1} \ln q + 1)$ .

For  $q = 2$  and  $K > 2$ , using two copies of the hypergraph from Lemma 2.1, we build an ‘equality’ gadget, i.e., a simple hypergraph  $H$  of maximum degree  $\Delta \leq 2(Kq^{K-1} \ln q + 1) = K2^K \ln 2 + 2$  with distinct vertices  $u, v$  which both have degree 1 such that for every  $q$ -colouring  $\sigma$  it holds that  $\sigma(u) = \sigma(v)$ . It is well-known that finding 2-colourings of  $K$ -uniform simple hypergraphs is NP-hard (or we can use for example [DD20]), and using the equality gadget  $H$ , for any  $K$ -uniform simple hypergraph  $F$ , we can construct a  $K$ -uniform simple hypergraph  $F'$  of maximum degree  $\Delta$  such that  $F$  is 2-colourable if and only if  $F'$  is 2-colourable. One possible way to do so is replacing each degree- $d$  vertex  $w$  of  $F$  with a cycle of length  $d$  and then replacing each edge  $e$  of the cycle with a distinct copy of the hypergraph  $H$  using  $u, v$  for the endpoints of the edge  $e$ ; then, for each hyperedge of  $F$  that uses  $w$ , in  $F'$  we use instead one of the  $d$  vertices of the cycle.  $\square$

Note that the result of Frieze and Mubayi [FM13] is also algorithmic. Thus Theorem 2.2 is sharp for simple hypergraphs up to a factor  $cq$  where  $c = c(K)$  is a constant depending only on  $K$ . For general hypergraphs, the algorithm of Moser and Tardos [MT10] applies in this setting when  $\Delta \leq \frac{q^{K-1}}{e^{(K-1)}}$ , in which case Theorem 2.2 almost matches the algorithmic result, up to a factor of  $cK^2q \ln q$  where  $c$  is a constant.

For approximate counting, we can avoid the loss of the factor  $q$  when  $q \geq 2$  and  $K \geq 4$ . For a  $q$ -by- $q$  matrix  $B = \{B_{ij}\}_{i,j \in [q]}$ , the partition function for the  $q$ -spin system with interaction matrix  $B$  in a graph  $G = (V, E)$  is given by

$$(4) \quad Z_B(G) := \sum_{\sigma: V \rightarrow \{1, \dots, q\}} \text{wt}(\sigma),$$

where  $\text{wt}(\sigma) := \prod_{(u,v) \in E} B_{\sigma(u)\sigma(v)}$  is the weight of an assignment  $\sigma : V \rightarrow \{1, \dots, q\}$  of the  $q$  spins to the vertices of  $G$ . In particular, the  $q$ -state antiferromagnetic Potts model corresponds to the case where  $B$  is the matrix whose off-diagonal entries are equal to 1, whereas the diagonal entries equal to some parameter  $B < 1$  (note,  $B = 0$  corresponds to  $q$ -colourings).

We will use the following hardness result about the Potts model. A *fully polynomial-time randomized approximation scheme* (FPRAS) is an algorithm that takes the accuracy  $\varepsilon$  as an extra input, outputs an  $\varepsilon$ -approximation, and runs in time polynomially bounded by both the instance size and  $1/\varepsilon$ .

**Lemma 2.3.** *There is a constant  $C_1 > 5$  such that, for any integers  $q \geq 2$ ,  $\Delta \geq 2C_1q \ln q$ , and  $B < 1 - \frac{C_1q \ln q}{\Delta}$ , there is no FPRAS to approximate the  $q$ -state antiferromagnetic Potts partition function  $Z_B$  in graphs with bounded degree  $\Delta$ , unless  $\text{NP} = \text{RP}$ .*

The proof of Lemma 2.3 is quite a detour from the problems we focus on, so we postpone it to Section 5.1. We note that Lemma 2.3 is weaker than the inapproximability result in [GŠV15, Theorem 1.2], which achieves  $B < 1 - \frac{q}{\Delta}$  but only holds for even  $q$ . We want to deal with general  $q$ , and thus settle with this weaker version.

**Theorem 2.4.** *There is a constant  $C_1 > 5$  such that, for any integers  $q \geq 2$ ,  $K \geq 4$ , and  $\Delta \geq C_1Kq^{K-1} \ln q$ , unless  $\text{NP} = \text{RP}$ , there is no FPRAS for the number of  $q$ -colourings in  $K$ -uniform simple hypergraphs of maximum degree at most  $\Delta$ .*

*Proof.* We reduce the partition function of the  $q$ -state antiferromagnetic Potts model with  $B = 1 - \frac{1}{q^2 - 3q + 3}$  in graphs with bounded degree  $\Delta$  to the problem of counting  $q$ -colourings in  $K$ -uniform simple hypergraphs of maximum degree at most  $\Delta$ . Note that if  $K \geq 4$  and  $\Delta \geq C_1Kq^{K-1} \ln q$ , where  $C_1$  is from Lemma 2.3, then  $B < 1 - \frac{C_1q \ln q}{\Delta}$ . Thus the reduction implies the theorem via Lemma 2.3.

The reduction goes as follows. For each edge  $(u, v)$  in a  $\Delta$ -regular graph  $G = (V, E)$ , we replace it by a gadget using the hypergraph  $H$  in Lemma 2.1, whose degree bound is  $\Delta_0 = Kq^{K-1} \ln q + 1$ . To be more specific, we introduce new vertices  $w_1$  and  $w_2$ . We add three copies of the hypergraph  $H$  with special vertices  $(u, w_1)$ ,  $(w_2, w_1)$ , and  $(v, w_2)$ , respectively. Do this for all edges in  $G$ . Then, the degrees of  $u$  and  $v$  are still  $\Delta$ , the degrees of  $w_1$ 's are at most  $2\Delta_0 < \Delta$ , and the degrees of  $w_2$ 's are at most  $\Delta_0 + 1 < \Delta$ . All other newly introduced vertices have degrees at most  $\Delta_0 < \Delta$ . Thus, the degree requirement is met. Call the resulting hypergraph  $H_G$ .

To finish the reduction, we claim that

$$Z_{\text{col}}(H_G) = C^{|E|} Z_B(G),$$

where  $C$  is a constant depending only on  $H$ . First notice that for any pair of colours  $i$  and  $j$ , the number of colourings  $\sigma$  of  $H$  such that  $\sigma(u) = i$  and  $\sigma(v) = j$  is a constant, due to the symmetry among colours. Denote this constant by  $C_0$ . Thus, in the gadget above, when the two endpoints  $u$  and  $v$  have different colours, the number of possible colourings for the gadget is  $((q-2)^2 + (q-1))C_0^3$ ; when the two endpoints  $u$  and  $v$  have the same colour, the number of possible colourings for the gadget is  $(q-1)(q-2)C_0^3$ . The claim holds with  $C = ((q-2)^2 + (q-1))C_0^3$ .  $\square$

In Theorem 2.4, we could avoid the large constant  $C_1$  in the degree bound by using [GŠV15, Theorem 1.2] as the starting point of our reduction, but doing so will restrict the result to even  $q$  only.

For  $K$ -CNFs on simple hypergraphs, Hermon, Sly, and Zhang [HSZ19] showed an efficient approximate counting and sampling algorithm if  $\Delta \leq \frac{c2^K}{K^2}$ , where  $c$  is a constant. In view of their result, Theorem 2.2 and Theorem 2.4 are potentially sharp for simple hypergraphs, up to some polynomial factor in  $K$ .

### 3. REFINED INAPPROXIMABILITY FOR APPROXIMATE COUNTING

In this section we show our main theorem, Theorem 1.1, namely a refined inapproximability result for counting. As mentioned earlier, we will do this by first relating it to a multi-spin system on graphs with “antiferromagnetic” interaction matrix  $B$ , and then establishing inapproximability results. It is tempting to pursue a strategy similar to that of Lemma 2.3 to show hardness for the spin system defined by  $B$ . However, that strategy relies on hardness of finding the maximum weight configuration, and somewhat surprisingly, as we shall see soon, that problem for  $B$  is trivial. Instead, we need sharper tools from [GSV15].

To define the spin system on graphs we will be interested in, we only need to specify its interaction matrix  $B$  (recall (4)). We use  $[q]$  to denote  $\{1, \dots, q\}$  and  $[\bar{q}]$  to denote  $\{0, 1, \dots, q\}$ . Let  $t := (q^k - q)^{1/\Delta}$ , where  $k := K/2$ , and  $B = \{B_{ij}\}_{i,j \in [\bar{q}]}$  be the matrix with block form

$$(5) \quad B = \begin{bmatrix} t^2 & t1^T \\ t1 & J \end{bmatrix},$$

where  $J$  is the  $q \times q$  matrix with 0s on the diagonal and 1s elsewhere, and  $1$  is the  $q \times 1$  vector with all ones. In the language of [GSV15], the matrix  $B$  is antiferromagnetic and ergodic.<sup>5</sup>

Let  $H$  be a  $K$ -uniform hypergraph, where  $K = 2k$  is even, and recall that we use  $Z_{col}(H)$  to denote the number of proper  $q$ -colourings of  $H$ . For any given  $\Delta$ -regular graph  $G = (V, E)$ , let  $H_G$  be the hypergraph where every vertex  $v \in V$  is replaced by  $k$  new vertices  $v_1, \dots, v_k$ , and each edge  $(u, v)$  is replaced by a hyperedge  $\{u_1, \dots, u_k, v_1, \dots, v_k\}$  of size  $2k$ . Then  $H_G$  is  $2k$ -uniform and  $\Delta$ -regular. This construction has been used in [BGG<sup>+</sup>19], and yields the following lemma in our case.

**Lemma 3.1.** *Let  $G = (V, E)$  be a  $\Delta$ -regular graph, and  $H_G = (V', E')$  be the  $2k$ -uniform hypergraph constructed as above. Then,  $Z_B(G) = Z_{col}(H_G)$ .*

*Proof.* Let  $\Omega_B$  be the set of all assignments  $\sigma$  of  $G$  whose weights are non-zero. Let  $\Omega_{col}$  be the set of all proper  $q$ -colourings  $\tau$  of  $H$ . We will construct a surjective mapping  $\varphi$  between  $\Omega_{col}$  and  $\Omega_B$ , such that for any  $\sigma$ ,  $|\varphi^{-1}(\sigma)| = \text{wt}(\sigma)$ . This implies the lemma.

The mapping  $\varphi$  is as follows. Given  $\tau : V' \rightarrow \{1, 2, \dots, q\}$ , let

$$\varphi(\tau)(v) := \begin{cases} i & \text{if } \tau(v_1) = \tau(v_2) = \dots = \tau(v_k) = i \text{ for some } 1 \leq i \leq q, \\ 0 & \text{otherwise.} \end{cases}$$

We first show that  $\varphi$  is surjective. Let  $\sigma \in \Omega_B$  and we construct  $\tau \in \Omega_{col}$  such that  $\varphi(\tau) = \sigma$ . For any  $v$  such that  $\sigma(v) \neq 0$ ,  $\tau(v_i) = \sigma(v)$  for any  $1 \leq i \leq k$ . If  $\sigma(v) = 0$ , then let  $\tau(v_i) = 1$  for any  $1 \leq i \leq k-1$ , and  $\tau(v_k) = 2$ . It is easy to verify that  $\tau$  is a proper  $q$ -colouring and  $\varphi(\tau) = \sigma$  for this construction.

Next we calculate  $|\varphi^{-1}(\sigma)|$ . Let  $n_0(\sigma)$  be the number of vertices assigned 0 under  $\sigma$ . Then

$$|\varphi^{-1}(\sigma)| = (q^k - q)^{n_0(\sigma)}.$$

On the other hand, since  $G$  is  $\Delta$ -regular,

$$\text{wt}(\sigma) = t^{\Delta n_0(\sigma)} = (q^k - q)^{n_0(\sigma)} = |\varphi^{-1}(\sigma)|,$$

which verifies the properties of  $\varphi$ . □

<sup>5</sup>Antiferromagnetism amounts to checking that  $B$  has all but one of its eigenvalues negative; it is not hard to see that  $B$  has  $-1$  as an eigenvalue by multiplicity  $q-1$ , and therefore using trace/determinant we see that the other two eigenvalues have sum equal to  $q-1+t^2$  and product  $-t^2$ . Ergodicity amounts to the fact that  $B$  is irreducible and aperiodic.

**Remark 3.2.** For odd  $K$ , we may consider a similar construction, but in addition to clustering half of each hyperedge as a single vertex, we leave one vertex in the middle which appears only in this single hyperedge. The resulting spin system would have a different matrix  $B'$ , but the difference between  $B'$  and the current  $B$  is not too much in the sense that the zeros would be replaced by small constants. We expect that we may obtain a hardness result for  $B'$  for  $\Delta$  of a similar order. However, since the details are already getting very complicated, we will only handle  $B$  in the rest of this paper.

Given Lemma 3.1, all we need to show is that the spin system with interaction matrix  $B$  is hard to approximate on  $\Delta$ -regular graphs, with  $\Delta$  in the desired range. For this, we will use a result by Galanis, Vigoda, and Štefankovič [GŠV15, Theorem 1.5] which gives a sufficient condition in terms of studying a certain function (that can be formulated in terms of an induced norm of  $B$ ). Note that since  $t > 1$  the corresponding optimization problem related to  $B$  is trivial. Thus, we cannot use a strategy similar to that of Lemma 2.3 to show hardness for the spin system defined by  $B$ .

The main construction in the gadget to show the hardness is the bipartite random  $\Delta$ -regular graph. Let  $(\alpha, \beta)$  be a pair of vectors such that for  $i \in [\bar{q}]$ ,  $\alpha_i$  and  $\beta_i$  denotes the fraction of vertices with colour  $i$  on the left and right sides of the bipartite random regular graph. If we draw a sample  $\sigma$  proportional to its weight  $\text{wt}(\sigma)$ , then with high probability over the choice of the random graph, the fraction of colours  $(\alpha, \beta)$  will be from one of the dominant phases, for all but an exponentially small probability. Analyzing these dominant phases lies in the heart of [GŠV15, Theorem 1.5].

Let  $\mathcal{G}_n$  denote the family of  $\Delta$ -regular bipartite graphs with  $n$  vertices on each side. For a bipartite graph  $G$  uniformly drawn from  $\mathcal{G}_n$  and probability vectors  $\alpha = \{\alpha_i\}_{i \in [q]}$ ,  $\beta = \{\beta_i\}_{i \in [q]}$ , we use  $Z_B^{\alpha, \beta}(G)$  to denote the total weights of assignments whose fractions of colours on the two sides are given by  $\alpha, \beta$  respectively. Consider the function  $\Psi_1$  that captures the exponential growth of the expectation of  $Z_B^{\alpha, \beta}(G)$ , i.e.,

$$(6) \quad \Psi_1(\alpha, \beta) := \lim_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{E}_{\mathcal{G}_n} [Z_B^{\alpha, \beta}(G)].$$

The function  $\Psi_1$  has a relatively explicit form (see [GŠV15, Section 2]) using entropy-style functions though the exact details are not going to be important and we will in fact use a surrogate function later on (see Section 4).

Before stating the main result of [GŠV15], we need some further terminology. A *dominant phase*  $(\alpha, \beta)$  is a maximizer of the function  $\Psi_1(\alpha, \beta)$  and captures the most likely configurations for the spin system with interaction matrix on a random  $\Delta$ -regular graph. A dominant phase is called *Hessian dominant* if the Hessian of  $\Psi_1$  is negative definite. Finally, two dominant phases  $(\alpha_1, \beta_1)$  and  $(\alpha_2, \beta_2)$  are *permutation symmetric* if there is a permutation matrix  $P$  such that  $B = PBP^T$  and  $(\alpha_1, \beta_1) = (P\alpha_2, P\beta_2)$  or  $(\alpha_1, \beta_1) = (P\beta_2, P\alpha_2)$ . Now we can state [GŠV15, Theorem 1.5].<sup>6</sup>

**Proposition 3.3** ([GŠV15]). *Let  $\Delta \geq 3$  be an integer, and suppose that  $B$  is an ergodic interaction matrix of an antiferromagnetic spin system. Suppose further that the dominant phases  $(\alpha, \beta)$  satisfy  $\alpha \neq \beta$ , are permutation symmetric and Hessian dominant. Then, it is NP-hard to approximate the partition function  $Z_B(G)$  on  $n$ -vertex triangle-free  $\Delta$ -regular graphs  $G$ , even within a factor of  $2^{cn}$  for a constant  $c(B, \Delta) > 0$ .*

The key ingredient in the conditions of Proposition 3.3 is the condition that  $\alpha \neq \beta$ ; this enables a reduction in [GŠV15] to the MAX-CUT problem; the Hessian dominance and the permutation symmetry condition are more on the technical side, but is one of the main reasons that complicates the overall arguments (this was already prevalent in [GŠV15]).

The main challenge to show our inapproximability results is to establish the conditions of Proposition 3.3 for  $B$  and the relevant range for  $\Delta$ , which is the scope of the following lemma.

**Lemma 3.4.** *Let  $q \geq 4$  be even,  $k \geq 2$ , and  $\Delta = 5q^k + 1$ . Then the dominant phases of the spin system with interaction matrix  $B$  (defined in (5)) satisfy the conditions of Proposition 3.3.*

Theorem 1.1 follows from Lemma 3.1, Proposition 3.3, and Lemma 3.4. It remains to analyse the dominant phases of  $B$  and establish Lemma 3.4, which is the focus of Section 4.

<sup>6</sup>Technically, [GŠV15, Theorem 1.5] demands the assumption  $\text{NP} \neq \text{RP}$  but that is merely to exclude randomised algorithms, the reduction itself is deterministic.

#### 4. ANALYSIS OF THE DOMINANT PHASES

In this section we analyze the dominant phase. We will state the main lemmas in this section and in Section 4.1. However, because the calculations are often very heavy, many of the lemmas are not immediately proved. The sections in which their proofs appear can be found in Table 1 at the end of Section 4.1.

Let  $q, \Delta \geq 3$  be integers. To prove Lemma 3.4, we need to analyse the function  $\Psi_1$  from (6). The function  $\Psi_1$  turns out to be inconvenient to work with, but there is a simpler surrogate function  $\Phi$  from [GŠV15] that we can use. For vectors  $\mathbf{r} = \{R_i\}_{i \in [\bar{q}]}$  and  $\mathbf{c} = \{C_i\}_{i \in [\bar{q}]}$  with nonnegative entries, let

$$(7) \quad \Phi(\mathbf{r}, \mathbf{c}) := \Delta \ln \frac{\mathbf{r}^T \mathbf{B} \mathbf{c}}{\|\mathbf{r}\|_p \|\mathbf{c}\|_p}, \text{ where } p = \Delta/(\Delta - 1).$$

It is not hard to see that for the matrix  $\mathbf{B}$  defined in (5), the critical points of  $\Phi$  satisfy the following equations:<sup>7</sup>

$$(8) \quad \begin{aligned} R_0 &\propto t^d \left( tC_0 + \sum_{j \in [q]; j \neq i} C_j \right)^d, & R_i &\propto \left( tC_0 + \sum_{j \in [q]; j \neq i} C_j \right)^d & \text{for } i \in [q]; \\ C_0 &\propto t^d \left( tR_0 + \sum_{i \in [q]; i \neq j} R_i \right)^d, & C_j &\propto \left( tR_0 + \sum_{i \in [q]; i \neq j} R_i \right)^d & \text{for } j \in [q], \end{aligned}$$

where  $t = (q^k - q)^{1/\Delta}$  and  $d := \Delta - 1$ . The equations in (8) are often called the “tree recursions”, because they are the same as the recursion for marginal probabilities on an infinite  $d$ -ary tree. Note that  $1 \leq t \leq 1.0312$  for any  $q \geq 4, k \geq 2$  and  $d \geq 5q^k$ . The connection between the functions  $\Psi_1$  and  $\Phi$  is detailed in the following result from [GŠV15], applied to our setting.

**Proposition 4.1** ([GŠV15, Theorem 4.1]). *Let  $q, \Delta \geq 3$  be integers, and let  $p = \Delta/(\Delta - 1)$ . Then, the local maxima of  $\Phi$  and  $\Psi_1$  happen at critical points, i.e., there are no local maxima on the boundary. The transformation  $(\mathbf{r}, \mathbf{c}) \mapsto (\boldsymbol{\alpha}, \boldsymbol{\beta})$  given by  $\alpha_i = R_i^p / \|\mathbf{r}\|_p^p$  and  $\beta_i = C_i^p / \|\mathbf{c}\|_p^p$  for  $i \in [\bar{q}]$  yields a one-to-one correspondence between the critical points of  $\Phi$  and  $\Psi_1$ . Moreover, for the corresponding critical points  $(\mathbf{r}, \mathbf{c})$  and  $(\boldsymbol{\alpha}, \boldsymbol{\beta})$  it holds that  $\Psi_1(\boldsymbol{\alpha}, \boldsymbol{\beta}) = \Phi(\mathbf{r}, \mathbf{c})$ .*

The function  $\Phi$  is still multi-dimensional ( $2q$ ), but fortunately we can reduce its dimensions significantly down to 11 by studying the structure of fixpoints to the system (8). A first observation is that  $R_i < R_j$  implies  $C_i > C_j$ , and  $R_i = R_j$  implies  $C_i = C_j$ , where  $i, j \neq 0$ . The next lemma is similar to [GŠV15, Lemma 7.6].

**Lemma 4.2.** *Let  $(R_0, R_1, \dots, R_q, C_0, C_1, \dots, C_q)$  be a positive fixpoint of (8). Then the number of distinct values in  $\{R_i\}_{1 \leq i \leq q}$  and  $\{C_i\}_{1 \leq i \leq q}$  is at most 3.*

*Proof.* Let  $R := \sum_{i=1}^q R_i$  and  $C := \sum_{i=1}^q C_i$ . Suppose all variables are normalized so that  $R_0 + R = C_0 + C = 1$ . Then for any  $i \in [q]$ , we have that

$$\frac{R_i}{R_0} = t^{-d} \left( \frac{(t-1)C_0 + 1 - C_i}{(t-1)C_0 + 1} \right)^d = t^{-d} \left( \frac{(t-1) + C_0^{-1} - C_i/C_0}{(t-1) + C_0^{-1}} \right)^d = t^{-d} \left( 1 - \frac{1}{t^d C'} \left( 1 - \frac{R_i}{R_0 R'} \right)^d \right)^d,$$

where  $C' = (t-1) + C_0^{-1}$  and  $R' = (t-1) + R_0^{-1}$ . Let  $x = (R_i/R_0)^{1/d}$  and note that  $x \in [0, 1]$ . Then the above equation becomes  $f(x) = 0$ , where  $f(x) := t^{-1} \left( 1 - \frac{1}{t^d C'} \left( 1 - \frac{x^d}{R'} \right)^d \right)^d - x$ . We have that

$$f'(x) := (g(x))^{d-1} - 1, \text{ where } g(x) := \left( \frac{d^2}{t^{d+1} R' C'} \right)^{1/(d-1)} \left( 1 - \frac{x^d}{R'} \right) x.$$

Note that  $g(x) > 0$  on the interval  $[0, 1]$  because  $x^d = \frac{R_i}{R_0} < (t-1) + R_0^{-1} = R'$ . Using that  $(g(x))^{d-1} - 1 = (g(x) - 1)(g(x)^{d-2} + \dots + 1)$ , we therefore obtain that the roots of  $f'(x) = 0$  can only come from the roots of  $g(x) - 1$ , which has at most two roots by the Descartes' rule of signs. Hence  $f'(x)$  changes its sign at most twice in the interval of  $[0, 1]$  and  $f(x)$  has at most 3 roots over  $[0, 1]$ , showing that the

<sup>7</sup>Here, and elsewhere, we use the notation  $x_i \propto y_i$  for  $i \in [\bar{q}]$  to denote that  $x_i = Ay_i$  for  $i \in [\bar{q}]$ , for some arbitrary  $A$ .

$R_i$ 's for  $i \in [q]$  can only be supported on three different values. The statement for the  $C_i$ 's follows by an analogous argument.  $\square$

The above lemma motivates the following definition.

**Definition 4.3.** Let  $(R_0, R_1, \dots, R_q, C_0, C_1, \dots, C_q)$  be a positive fixpoint. We call the fixpoint  $m$ -supported, if the number of distinct values in  $\{R_i\}_{1 \leq i \leq q}$  is  $m$ , where  $m \in \{1, 2, 3\}$ . We call the fixpoint is of type  $(q_1, q_2, q_3)$  where  $q_1 + q_2 + q_3 = q$ , if the multiplicities of different numbers in  $\{R_i\}_{1 \leq i \leq q}$  are  $q_1, q_2, q_3$  respectively.<sup>8</sup> In case that the fixpoint is 2 or 1-supported, let one or two of  $q_i$ 's take zero respectively.

From now on we may also abuse the notation  $R_i$  (also  $C_i$ ,  $i = 1, 2, 3$ ) by absorbing all the same values, and hence  $R_1$  stands for the value that  $q_1$  of  $R$ 's (except  $R_0$ ) take, rather than the value of  $R$  on the first index in the fixpoint.

The main lemma of this section can be stated as follows.

**Lemma 4.4.** Suppose  $q \geq 4$  is even,  $k \geq 2$  and  $d = 5q^k$ . The maximum of  $\Psi_1$  over  $(q_1, q_2, q_3)$ -type fixpoints is attained uniquely, when  $(q_1, q_2, q_3) = (q/2, q/2, 0)$ .

We also need to prove that 2-maximal triples  $(q/2, q/2, 0)$  yield unique  $r$  and  $c$  (up to scaling and permutation), and that the corresponding maxima are Hessian dominant.

**Lemma 4.5.** Suppose  $q \geq 4$  is even,  $k \geq 2$  and  $d \geq 5q^k$ . The fixpoints of type  $(q/2, q/2, 0)$  are unique up to scaling and permutation symmetric. In addition, they are Hessian dominant maxima of  $\Psi_1$ .

Lemma 3.4 follows immediately by combining Lemmas 4.4 and 4.5.

**4.1. Restricting to three values.** In order to prove Lemma 4.4, we need to determine which type of fixpoints maximizes  $\Psi_1$ . By using Proposition 4.1, the value of  $\Psi_1$  corresponding to such a fixpoint in (8) can be given by the matrix norm (7), which can be seen to be equal to

$$\begin{aligned} \overline{\Phi^S}(q, r, c) := \\ (9) \quad & (d+1) \ln \left( R_0 C_0 t^2 + \left( \sum_{i=1}^3 q_i R_i \right) C_0 t + \left( \sum_{i=1}^3 q_i C_i \right) R_0 t + \left( \sum_{i=1}^3 q_i R_i \right) \left( \sum_{i=1}^3 q_i C_i \right) - \left( \sum_{i=1}^3 q_i R_i C_i \right) \right) \\ & - d \ln \left( R_0^{(d+1)/d} + \sum_{i=1}^3 q_i R_i^{(d+1)/d} \right) - d \ln \left( C_0^{(d+1)/d} + \sum_{i=1}^3 q_i C_i^{(d+1)/d} \right). \end{aligned}$$

Where we define the vector  $r = (R_0, R_1, R_2, R_3)$  and  $c = (C_0, C_1, C_2, C_3)$ . It is worth noting that this function is scale-free with respect to  $r$  and  $c$ , as this property will be used intensively in our later proofs.

The discrete optimization of (9) over all fixpoints of the tree recursion (8) is difficult to cope with. Instead, we then try to maximize (9) over all nonnegative  $q$  and  $\sum_{i=1}^3 q_i = q$ , wishing the maximum to be taken at integer  $q$ . This is the main reason such approach can only deal with even  $q$ .

For all  $q = (q_1, q_2, q_3)$  with  $q_1 + q_2 + q_3 = q$ ,  $q_i \geq 0$ , define

$$(10) \quad \overline{\Phi}(q) := \max_{r, c} \overline{\Phi^S}(q, r, c)$$

where the maximum is taken over  $r = (R_0, R_1, R_2, R_3)$ ,  $c = (C_0, C_1, C_2, C_3)$  satisfying

$$(11) \quad \begin{aligned} & R_0 C_0 t^2 + \left( \sum_{i=1}^3 q_i R_i \right) C_0 t + \left( \sum_{i=1}^3 q_i C_i \right) R_0 t + \left( \sum_{i=1}^3 q_i R_i \right) \left( \sum_{i=1}^3 q_i C_i \right) - \left( \sum_{i=1}^3 q_i R_i C_i \right) > 0, \\ & R_i, C_i \geq 0, i = 0, 1, 2, 3. \end{aligned}$$

Our first step is to verify the maximum in (10) is well defined, and moreover, the maximum in  $\max_q \overline{\Phi}(q)$  can also be taken. This is formalized by the next lemma.

**Lemma 4.6** ([GŠV15, Lemma 7.10]). The maximum in (10) is well-defined. In addition,  $\max_q \overline{\Phi}(q)$  can be attained in the region where  $q_1 + q_2 + q_3 = q$ ,  $q_i \geq 0$ .

*Proof.* The argument is verbatim the same as in [GŠV15], the only difference is that the function has slightly different form, but still accounts for the relevant parameters  $q_1, q_2, q_3$ .  $\square$

<sup>8</sup>Any permutation over  $q_1, q_2, q_3$  is considered equivalent. E.g.,  $(q/2, q/2, 0)$  and  $(q/2, 0, q/2)$  are regarded as the same type.

The next trouble we may encounter later is that we are now dealing with all possible  $r, c$  conditioned on (11), instead of just fixpoints of (8). The good news is that, in contrast to [GŠV15], we can rule out fairly easily that the maximizer in (10) is at the boundary.

**Lemma 4.7.** *For all triples  $q = (q_1, q_2, q_3)$ , any maximizer in (10) satisfies (a)  $R_0, C_0 > 0$ , (b) for any  $i$  such that  $q_i > 0$ , it holds that  $R_i, C_i > 0$ , and (c) for distinct  $i, j$  such that  $q_i, q_j > 0$ , it holds that  $R_i = R_j$  if and only if  $C_i = C_j$ .*

The problem in [GŠV15] that also appears in our setting is that it might be that  $q_i, q_j > 0$ , but  $R_i = R_j$  and  $C_i = C_j$ . For example, imagine we are now strengthening the restriction (11) by adding  $R_1 = R_2$  and  $C_1 = C_2$ . Then  $\bar{\Phi}(q_1 + q_2, 0, q_3) \leq \bar{\Phi}(q_1, q_2, q_3)$ . Such degenerate case makes it difficult to compare between different  $q$  triples because the equality can be taken. This motivates the next definition.

**Definition 4.8.** Let  $m = 2, 3$ . A triple  $q$  is called *m-maximal*, if exactly  $m$   $q_i$ 's in  $q$  are non-zero, and there exists  $r, c$  maximizing (10) such that,  $q_i, q_j > 0$  and  $i \neq j$  imply that  $R_i \neq R_j$  and  $C_i \neq C_j$ . We also call  $q$  *maximal* if it is either 2- or 3-maximal.

Now we connect  $m$ -maximal triples with fixpoints in (8).

**Lemma 4.9.** *Suppose a triple  $q$  is  $m$ -maximal. Then there exists  $r, c$  achieving the maximum in (10) and specifying an  $m$ -supported fixpoint of tree recursion (8) of type  $q$ .*

For 2 and 3-maximal triples, the key is the next lemma.

**Lemma 4.10.** *Suppose  $q \geq 4$  is even. Then the following statements hold:*

- (a) *There does not exist any 3-maximal triple that maximizes (10).*
- (b) *The only possibility of a 2-maximal triple to maximize (10) is  $(q/2, q/2, 0)$  or its permutations, with  $R_i/R_j = C_j/C_i$ , where  $i \neq j$  are the two indices such that  $q_i, q_j = q/2$ .*

The above lemma is not yet enough to finish the proof of Lemma 4.4 because we have to rule out degenerate cases of all triples, i.e., the triple  $(q, 0, 0)$ . This is the main difference with the colour-symmetric setting of [GŠV15]. Instead, we have the special colour corresponding to  $(R_0, C_0)$ , which makes the system behave like a 2-spin system when all “pure” colours take the same fraction. What is worse is that, it is possible for the 2-spin system to have three fixpoints (two of them being symmetric), when the tree recursion lies in the so-called “non-uniqueness” region (see Section 4.4). Therefore, we need to discuss such fixpoints by two different cases.

Before continuing the discussion, let us state another useful result from [GŠV15]. A fixpoint  $x$  of a mapping  $f$  is *Jacobian stable* if the Jacobian of  $f$  at  $x$  has spectral radius less than 1.

**Proposition 4.11** ([GŠV15, Theorem 4.2]). *A fixpoint of the tree recursion (8) is Jacobian stable if and only if it corresponds to a Hessian dominant local maximum of  $\Psi_1$ .*

The first kind of fixpoints satisfy  $R_0/R_1 \neq C_0/C_1$ . As stated in the next lemma, such a fixpoint is Jacobian stable, and hence it is a possible candidate to be the maximizer in  $\max_q \bar{\Phi}(q)$ . Though the proof of stability is not necessary for our main theorem, we still leave it here for future references.

**Lemma 4.12.** *Suppose  $d \geq 5q^k$ . The fixpoint corresponding to triple  $(q, 0, 0)$  and  $R_0/R_1 \neq C_0/C_1$  is unique up to scaling and swapping  $R$  and  $C$ . Moreover, it is Jacobian stable.*

For the reason above, we can only go through a very detailed calculation to rule out this case. The equality in  $d = 5q^k$  from the next lemma is for the sake of simplification in calculation.

**Lemma 4.13.** *Suppose  $d = 5q^k$ . Any fixpoint corresponding to triple  $(q, 0, 0)$  and  $R_0/R_1 \neq C_0/C_1$  does not maximize (10).*

On the other hand, when  $R_0/R_1 = C_0/C_1$ , things become easier as such fixpoints are not Jacobian stable. Thus, by Proposition 4.11, these fixpoints do not correspond to local maxima of  $\Psi_1$ .

**Lemma 4.14.** *Suppose  $d \geq 5q^k$ . Any fixpoint corresponding to triple  $(q, 0, 0)$  and  $R_0/R_1 = C_0/C_1$  is Jacobian unstable.*

Now we are ready to prove Lemma 4.4, which given the above ingredients can be done by following closely a related argument in [GŠV15]. The main compicacy in the proof is that when we find a maximizer  $q$  of  $\bar{\Phi}(q)$ , the corresponding  $r$  (or  $c$ ) values are not necessarily distinct. We need to carefully rule out these degenerate cases.

*Proof of Lemma 4.4.* Denote  $MAX := \max_q \bar{\Phi}(q)$ . We first claim that  $MAX$  is attained at  $\hat{q} = (q/2, q/2, 0)$ , and  $\hat{q}$  is maximal. Assuming the claim, Lemma 4.9 yields that there exist  $\hat{r}, \hat{c}$  with  $\bar{\Phi}(\hat{q}) = \bar{\Phi}^S(\hat{q}, \hat{r}, \hat{c})$ , specifying a  $(q/2, q/2, 0)$ -type fixpoint of the tree recursion (8). Hence  $MAX = \max \Psi_1$ . To show that  $\hat{q}$  is the unique type of fixpoint achieving the maximum of  $\Psi_1$ , consider an arbitrary  $q^*$ -type fixpoint achieving the maximum of  $\Psi_1$ , say  $(r^*, c^*)$ . Then  $q^*$  must also achieve the maximum in  $\max_q \bar{\Phi}(q)$ . By Lemma 4.13,  $q^* \neq (q, 0, 0)$  and hence it is maximal according to Definition 4.8 (using  $(r^*, c^*)$  as the maximizers; Recall Definition 4.3 that  $R_i \neq R_j, C_i \neq C_j$  for  $i \neq j$  and  $q_i, q_j > 0$ ). Therefore we can apply Lemma 4.10 and obtain that  $q^* = \hat{q}$ .

It remains to prove the claim above. Let  $q^*$  be any maximizer of  $\max_q \bar{\Phi}(q)$ .

- (1)  $q^*$  has at least two positive entries. This is a consequence of Lemmas 4.13 and 4.14 (after using Proposition 4.11).
- (2) In case  $q^*$  has exactly two positive entries, then  $q^*$  must be maximal. Otherwise, suppose  $q^* = (q_1, q_2, 0)$  and the maximizer in (10) is achieved at  $r^*, c^*$  where  $R_1 = R_2$  or  $C_1 = C_2$ . By Lemma 4.7 (c), both equalities are true and hence  $\bar{\Phi}(q^*) = \bar{\Phi}((q, 0, 0))$ , contradicting item (1).
- (3) In case  $q^*$  has exactly two positive entries, it must holds that  $q^* = \hat{q}$ . This is from item (2), and Lemma 4.10 (b).
- (4) If  $q^*$  has all positive entries, then it cannot be 3-maximal. This is from Lemma 4.10 (a).
- (5) If  $q^*$  has all positive entries, then  $\bar{\Phi}(q^*) = \bar{\Phi}(\hat{q})$ . This can be proved by the following argument. Let  $r^*, c^*$  be the maximizer corresponding to  $q^*$ . By item (4),  $q^*$  is not 3-maximal, and using the argument of item (2), there exist distinct  $i, j \geq 1$  such that  $R_i = R_j$  and  $C_i = C_j$  in  $r^*, c^*$ . Let  $k \geq 1$  be the remaining index.
  - If  $R_i = R_j = R_k$ , then by Lemma 4.7 (c),  $C_i = C_j = C_k$ , and hence  $\bar{\Phi}(q^*) = \bar{\Phi}(q, 0, 0)$ , contradicting item (1).
  - If  $C_i = C_j = C_k$ , then by Lemma 4.7 (c),  $R_i = R_j = R_k$ , and hence  $\bar{\Phi}(q^*) = \bar{\Phi}(q, 0, 0)$ , contradicting item (1).
  - If  $R_i \neq R_k$  and  $C_i \neq C_k$ , we can “merge” the indices  $i, j$  to get a new triple  $q' := (q_i + q_j, q_k, 0)$ . Let  $r' := (R_0, R_i, R_k, 0)$ ,  $c' := (R_0, C_i, C_k, 0)$ . Then

$$\bar{\Phi}(q^*) = \bar{\Phi}^S(q^*, r^*, c^*) = \bar{\Phi}^S(q', r', c') \leq \bar{\Phi}(q').$$

This means that  $q'$  is also a maximizer of  $\max_q \bar{\Phi}(q)$  since  $q^*$  is a maximizer. However,  $q'$  has exactly two positive entries. Hence by item (3),  $\bar{\Phi}(q^*) = \bar{\Phi}(q') = \bar{\Phi}(\hat{q})$ .

The above arguments imply that for any maximizer  $q^*$ , it holds that  $\bar{\Phi}(q^*) = \bar{\Phi}(\hat{q})$ , which means that  $\hat{q}$  is indeed a maximizer. This also indicates all items above apply to  $q^* = \hat{q}$ , and from item (3), we obtain that  $\hat{q}$  is 2-maximal. This concludes the proof.  $\square$

Before diving into the proofs of all the lemmas above, we want to mention the following observation. The partial derivatives  $\partial \bar{\Phi}^S / \partial q_i$ , conditioned on  $r$  and  $c$  achieving the maximum in (10), can be written as follows. (Note that it applies to all triples  $q$ , including non-maximal ones.) Based on these partial derivatives, we can argue the non-optimality by perturbing  $q_i$ 's.

**Lemma 4.15.** *Suppose  $r, c$  achieve the maximum in (10). Then for any  $i \in \{1, 2, 3\}$  such that  $q_i > 0$ , it holds that*

$$\frac{\partial \bar{\Phi}^S}{\partial q_i} = \frac{R_i C_0 t + R_0 C_i t + (d-1) R_i C_i + R_i \left( \sum_{j=1}^3 C_j q_j \right) + C_i \left( \sum_{j=1}^3 R_j q_j \right)}{R_0 C_0 t^2 + \left( \sum_{j=1}^3 C_j q_j \right) R_0 t + \left( \sum_{j=1}^3 R_j q_j \right) C_0 t + \left( \sum_{j=1}^3 R_j q_j \right) \left( \sum_{j=1}^3 C_j q_j \right) - \left( \sum_{j=1}^3 R_j C_j q_j \right)}.$$

Moreover, if there exists  $i, j$  such that  $q_i, q_j > 0$  and  $i \neq j$  and satisfies  $\partial \bar{\Phi}^S / \partial q_i - \partial \bar{\Phi}^S / \partial q_j \neq 0$ , then the maximum in (10) is not achieved.

Proposition/Lemma	Section
Lemma 4.10	Section 4.2
Lemma 4.5	Section 4.3
Lemma 4.12, Lemma 4.14	Section 4.4
Lemma 4.13	Section 4.5
Lemma 4.9, Lemma 4.15	Section 5.2
Lemma 4.7	Section 5.3

TABLE 1. The sections where the lemmas are proved

Unproved propositions and lemmas in this subsection can be found later. We make a list of where they are proved.

**4.2. 2, 3-maximal Triples.** Let  $q$  be a maximal triple and let  $I = \{i \mid q_i > 0\}$ . From Lemma 4.7 (a) and (b), by taking partial derivatives of  $\overline{\Phi^S}$  with respect to non-zero  $R_i$  and  $C_i$ 's and setting them to 0, we get that the maximizer of  $\overline{\Phi^S}$  satisfies

$$(12) \quad R_0^{1/d} \propto C_0 t^2 + (q_1 C_1 + q_2 C_2 + q_3 C_3)t, \quad R_i^{1/d} \propto C_0 t + q_1 C_1 + q_2 C_2 + q_3 C_3 - C_i \text{ for } i \in I;$$

$$(13) \quad C_0^{1/d} \propto R_0 t^2 + (q_1 R_1 + q_2 R_2 + q_3 R_3)t, \quad C_i^{1/d} \propto R_0 t + q_1 R_1 + q_2 R_2 + q_3 R_3 - R_i \text{ for } i \in I.$$

First assume  $q$  is 3-maximal, for any  $i \neq j$  it holds that  $R_i \neq R_j$  and  $C_i \neq C_j$ . From Lemma 4.7 (a) and (b), we may assume the following strict ordering

$$(14) \quad R_1 > R_2 > R_3 > 0 \quad \text{and} \quad 0 < C_1 < C_2 < C_3.$$

**Lemma 4.16.** *Suppose  $R_i$ 's and  $C_i$ 's satisfy (12), (13) and (14). We have the following:*

(a) *If  $R_1/R_3 \neq C_3/C_1$ , then  $\partial \overline{\Phi^S} / \partial q_1 - \partial \overline{\Phi^S} / \partial q_3 \neq 0$ .*

(b) *If  $R_1/R_3 = C_3/C_1$ , then  $\partial \overline{\Phi^S} / \partial q_1 - \partial \overline{\Phi^S} / \partial q_2 \neq 0$ .*

For the sake of convenience, we further set

$$r_0^d := R_0/R_3, r_1^d := R_1/R_3, r_2^d := R_2/R_3, \text{ and } c_0^d := C_0/C_1, c_2^d := C_2/C_1, c_3^d := C_3/C_1.$$

which means

$$(15) \quad r_1 > r_2 > 1 \quad \text{and} \quad c_3 > c_2 > 1.$$

We will need these notations in later sections too. With them, from (12) and (13), we obtain that

$$(16) \quad r_0 = \frac{c_0^d t^2 + (q_1 + q_2 c_2^d + q_3 c_3^d)t}{c_0^d t + q_1 + q_2 c_2^d + (q_3 - 1)c_3^d}, \quad c_0 = \frac{r_0^d t^2 + (q_1 r_1^d + q_2 r_2^d + q_3)t}{r_0^d t + (q_1 - 1)r_1^d + q_2 r_2^d + q_3}.$$

$$(17) \quad r_1 = \frac{c_0^d t + q_1 - 1 + q_2 c_2^d + q_3 c_3^d}{c_0^d t + q_1 + q_2 c_2^d + (q_3 - 1)c_3^d}, \quad c_3 = \frac{r_0^d t + q_1 r_1^d + q_2 r_2^d + q_3 - 1}{r_0^d t + (q_1 - 1)r_1^d + q_2 r_2^d + q_3},$$

*Proof of Lemma 4.16.* From (12) and (13), we get

$$(18) \quad \frac{r_1 - 1}{r_2 - 1} = \frac{c_3^d - 1}{c_3^d - c_2^d}, \text{ yielding that } r_2 = \frac{r_1 c_3^d - 1 - c_2^d(r_1 - 1)}{c_3^d - 1}$$

Similarly, we obtain that

$$(19) \quad \frac{c_3 - 1}{c_2 - 1} = \frac{r_1^d - 1}{r_1^d - r_2^d}, \text{ yielding that } r_2^d = \frac{r_1^d c_3 - 1 - c_2(r_1^d - 1)}{c_3 - 1},$$

From (12) and (13), we have that  $r_2 = \frac{c_0^d t + q_1 + (q_2 - 1)c_2^d + q_3 c_3^d}{c_0^d t + q_1 + q_2 c_2^d + (q_3 - 1)c_3^d}$  which combined with (18) gives that

$$(20) \quad \frac{c_0^d t + q_1 + (q_2 - 1)c_2^d + q_3 c_3^d}{c_0^d t + q_1 + q_2 c_2^d + (q_3 - 1)c_3^d} = \frac{r_1 c_3^d - 1 - c_2^d (r_1 - 1)}{c_3^d - 1}.$$

Symmetrically we obtain that

$$(21) \quad \frac{r_0^d t + q_1 r_1^d + (q_2 - 1)c_2^d + q_3}{r_0^d t + (q_1 - 1)r_1^d + q_2 c_2^d + q_3} = \frac{r_1^d c_3 - 1 - r_2^d (c_3 - 1)}{r_1^d - 1}.$$

We can view (20) and (21) as a linear system in  $q_1$  and  $q_3$  after clearing the denominators, which yields that

$$(22) \quad q_1 \cdot (r_1^d c_3^d - 1) = c_0^d t + q_2 c_2^d + \frac{1 - r_1 c_3^d}{r_1 - 1} - c_3^d \left( r_0^d t + q_2 r_2^d + \frac{1 - c_3 r_1^d}{c_3 - 1} \right),$$

$$(23) \quad q_3 \cdot (r_1^d c_3^d - 1) = r_0^d t + q_2 r_2^d + \frac{1 - r_1^d c_3}{c_3 - 1} - r_1^d \left( c_0^d t + q_2 c_2^d + \frac{1 - r_1 c_3^d}{r_1 - 1} \right),$$

From (12) and (13), we also obtain that

$$(24) \quad r_0^d t = \frac{r_1^d - 1}{c_3 - 1} - (q_1 - 1)r_1^d - q_2 r_2^d - q_3, \quad c_0^d t = \frac{c_3^d - 1}{r_1 - 1} - q_1 - q_2 c_2^d - (q_3 - 1)c_3^d.$$

We can now show the following:

$$(25) \quad \text{if } r_1 = c_3, \text{ then (i) } r_2 = c_2, \text{ (ii) } q_1 = q_3, \text{ and (iii) } r_0 = c_0.$$

The proof of (i) in (25) is identical to that in [GŠV15, Lemma 7.20] using the expressions for  $r_2, r_2^d$  in (18) and (19), respectively. From (i) and the assumption that  $r_1 = c_3$ , we obtain from (22) and (23) that

$$(26) \quad q_3 - q_1 = (r_0^d - c_0^d) \cdot \frac{t}{r_1^d - 1}.$$

Furthermore, the equations in (16) can also be regarded as a linear system in  $q_1$  and  $q_3$ . Using the assumption  $r_1 = c_3$  and  $r_2 = c_2$ , we obtain that

$$q_3 - q_1 = \frac{t \left[ (r_0^d - c_0^d)t^2 + (r_0 c_0^d - c_0 r_0^d + c_0^{1+d} - r_0^{1+d})t + (r_0^d - c_0^d)r_0 c_0 - (r_0 - c_0)r_1^d \right]}{(r_0 - t)(c_0 - t)(r_1^d - 1)},$$

which, together with (26), implies  $r_0 = c_0$  and hence  $q_1 = q_3$ . This finishes proving (25).

We are now ready to give the proof of Lemma 4.16. For part (a), Lemma 4.15 yields that

$$\begin{aligned} \frac{\partial \overline{\Phi^S}}{\partial q_1} - \frac{\partial \overline{\Phi^S}}{\partial q_3} &= \frac{1}{S} \times \left[ (r_1^d - 1)(c_0^d t + q_1 + c_2^d q_2) - (c_3^d - 1)(r_0^d t + q_3 + r_2^d q_2) \right. \\ &\quad \left. + (d - 1)(r_1^d - c_3^d) + r_1^d c_3^d (q_3 - q_1) + r_1^d q_1 - c_3^d q_3 \right]. \end{aligned}$$

where  $S > 0$ . Then plug in the expression of  $q_1$  and  $q_3$  in (22) and (23), we get

$$(27) \quad \frac{\partial \overline{\Phi^S}}{\partial q_1} - \frac{\partial \overline{\Phi^S}}{\partial q_3} = -\frac{g(r_1, c_3)}{S}$$

where  $g(r_1, c_3) := (r_1 - c_3)(r_1^d - 1)(c_3^d - 1) - d(r_1 - 1)(c_3 - 1)(r_1^d - c_3^d)$ . This quantity was shown to have the same sign as  $r_1 - c_3$  (see Equation (123) in the proof of Lemma 7.19 in [GŠV15]), and specifically, non-zero when  $r_1 \neq c_3$ , concluding part (a).

Now we prove part (b) of Lemma 4.16. From (25), the assumption  $r_1 = c_3$  implies  $r_2 = c_2$ ,  $q_1 = q_3$  and  $r_0 = c_0$ . Applying Lemma 4.15 based on these, we get

$$\begin{aligned} \frac{\partial \bar{\Phi}^S}{\partial q_1} - \frac{\partial \bar{\Phi}^S}{\partial q_2} &= q_1(1 + r_1^d)(1 + r_1^d - 2r_2^d) + r_2^d(q_2 - (2q_2 + d - 1)r_2^d - 2r_0^d t) + r_0^d t + r_1^d(r_0^d t + q_2 r_2^d + d - 1) \\ &= -\frac{(d-1)(r_1-1)r_2^{2d} + 2(r_1^{d+1}-1)r_2^d - (r_1^{2d+1} + dr_1^{d+1} - dr_1^d - 1)}{r_1 - 1}, \end{aligned}$$

where in the second line we use (24). This quantity was shown to be non-zero in the proof of Lemma 7.19 in [GŠV15] (from Equation (124) onwards) under (18), concluding part (b).  $\square$

Now we assume  $q = (q_1, q_2, q_3)$  is a 2-maximal triple, and assume  $q_2 = 0$  without loss of generality. The result here is analogous to Lemma 4.16 (a).

**Lemma 4.17.** *Under the assumption that  $q_2 = 0$ , suppose  $R_i$ 's and  $C_i$ 's ( $i \neq 2$ ) satisfy (12), (13) and (14). For any  $q_1, q_3 > 0$ , it holds that  $\partial \bar{\Phi}^S / \partial q_1 - \partial \bar{\Phi}^S / \partial q_3 \neq 0$ , unless  $q_1 = q_3$  and  $R_1/R_3 = C_3/C_1$ .*

*Proof.* First, note that the values of  $R_2$  and  $C_2$  do not affect the value of derivatives  $\partial \bar{\Phi}^S / \partial q_1$  and  $\partial \bar{\Phi}^S / \partial q_3$  when  $q_2 = 0$ . In addition, the expressions of  $q_1$  and  $q_3$  in (22) and (23) still hold for  $q_2 = 0$ . Therefore, one can carry out the proof of Lemma 4.16 (a) once again for this case, showing  $\partial \bar{\Phi}^S / \partial q_1 - \partial \bar{\Phi}^S / \partial q_3 = 0$  only when  $R_1/R_3 = C_3/C_1$ . Assuming this, one can show  $q_1 = q_3$  by going through the proof of (25).  $\square$

We conclude this subsection with Lemma 4.10.

*Proof of Lemma 4.10.* This comes after Lemma 4.16, Lemma 4.17 and the second part of Lemma 4.15.  $\square$

**4.3. Stability of Maximal  $(q/2, q/2, 0)$  Fixpoints.** In the next two subsections, we focus on the (in)stability of candidate fixpoints that may maximize  $\Psi_1$ . The condition of Jacobian stability is given by the following Lemma.

**Lemma 4.18** (cf. [GŠV15, Lemma 4.16]). *Suppose  $(R_0, R_1, \dots, R_q, C_0, C_1, \dots, C_q)$  is a fixpoint of the tree recursion (8). Let  $\alpha_i := \sum_{j=0}^q B_{ij}R_iC_j$  and  $\beta_j := \sum_{i=0}^q B_{ij}R_iC_j$ . Define the matrix  $A := (a_{ij})_{0 \leq i, j \leq q}$  as  $a_{ij} = B_{ij}R_iC_j / \sqrt{\alpha_i \beta_j}$ , and the matrix  $L := \begin{bmatrix} 0 & A \\ A^\top & 0 \end{bmatrix}$ . Then  $L$  has symmetric real spectrum (symmetry means if  $a$  is an eigenvalue then so is  $-a$ ), and  $\pm 1$  is a pair of its eigenvalues. The condition for the fixpoint to be stable is that the second largest eigenvalue of  $L$  is less than  $1/d$ .*

We will also need the following lemma which is proved in Section 5.4.

**Lemma 4.19.** *For any  $q \geq 4, k \geq 2$  and  $d \geq 3q^k$ , the function*

$$h(x) := \left( \frac{x^{d+1} - 1}{x^d - 1} \right)^d t^{d+1} - \frac{x^d - 1}{x - 1} + q' + (q' - 1)x^d$$

*has exactly one root in the region  $(1, \infty)$ .*

We are now ready to prove Lemma 4.5.

*Proof of Lemma 4.5.* Define  $q' := q/2$ . We first prove the uniqueness of 2-maximal  $(q', q', 0)$  fixpoint (up to scaling). According to the proof of Lemma 4.17, fixpoints of type  $(q', q', 0)$  maximize  $\bar{\Phi}$  only when  $r_1 = c_3$ . Now denote  $x := r_1 = c_3$ . To prove the first part of this lemma, we show there exists exactly one possible  $x > 1$  when  $d \geq 3q^k$ . By (16) and (17), we get

$$\frac{r_0/t - r_1}{r_1 - 1} = \frac{1}{c_3^d - 1}.$$

Combining this with (24),  $x > 1$  satisfies  $h(x) = 0$ , where

$$h(x) := \left( \frac{x^{d+1} - 1}{x^d - 1} \right)^d t^{d+1} - \frac{x^d - 1}{x - 1} + q' + (q' - 1)x^d.$$

By Lemma 4.19,  $h(x)$  has exactly one root  $x > 1$ .

Next, we construct the matrices  $A$  and  $L$ . Note that both matrices are scale-free with respect to  $R_i$  and  $C_i$ . Directly plug in the formula in Lemma 4.18 to get

$$A := \begin{bmatrix} c^2 & bc1^T & ac1^T \\ ac1 & abJ & a^2J' \\ bc1 & b^2J' & abJ \end{bmatrix}.$$

where  $a := \sqrt{x^{d-1} \frac{x-1}{x^d-1}}$ ,  $b := \sqrt{\frac{x-1}{x^d-1}}$  and  $c := \sqrt{\frac{x^{d+1}-1-q'(x-1)(x^{d+1})}{x^{d+1}-1}}$ ,  $J$  is the  $q' \times q'$  matrix with zeros on the diagonal and ones elsewhere,  $J'$  is the  $q' \times q'$  matrix with ones everywhere, and  $1$  is the  $q' \times 1$  matrix. The eigenvalues of  $L = \begin{bmatrix} 0 & A \\ A^T & 0 \end{bmatrix}$  consist of  $\pm ab$  (each by multiplicity  $q-2$ ) and  $\pm\lambda_1, \pm\lambda_2, \pm\lambda_3$ , where  $\lambda_1, \lambda_2, \lambda_3$  are the zeros of the following cubic function

$$f(z) = z^3 - (q'a^2 + q'b^2 + c^2)z^2 + (2q' - 1)a^2b^2z + a^2b^2c^2.$$

We claim that  $ab$  is the second largest eigenvalue. To prove this, recall that  $1$  is the eigenvalue of  $L$ . We can assume  $\lambda_1 = 1$  (because  $ab < 1$ , which means  $1$  must be among  $\lambda_{1,2,3}$ ) and hence  $f(1) = 0$ . In addition,  $f(z)$  is monic and  $f(0) > 0$ . This means it suffices to show  $f(-ab) \leq 0$  and  $f(ab) \leq 0$ , which are true since

$$f(ab) = -a^2b^2q(a-b)^2 < 0, \quad f(-ab) = -a^2b^2q(a+b)^2 < 0.$$

It remains to prove  $ab = x^{(d-1)/2} \frac{x-1}{x^d-1} < 1/d$  which follows from  $\frac{x^d-1}{x-1} = x^{d-1} + \dots + 1 > dx^{(d-1)/2}$ , where the last inequality is an application of the AM-GM inequality when  $x > 1$ .  $\square$

**4.4. (In)stability of  $(q, 0, 0)$  Fixpoints.** Set  $x := R_0/R_1$  and  $y := C_0/C_1$ . Then by rewriting the tree recursion, one can see  $x, y$  satisfies the system

$$(28) \quad x = t^d \left( \frac{ty + q}{ty + q - 1} \right)^d, \quad y = t^d \left( \frac{tx + q}{tx + q - 1} \right)^d.$$

Before analysing the stability of the original  $(q+1)$ -spin system, we first need to study this 2-spin system. By replacing  $\beta := t/q$ ,  $\gamma = (q-1)/t$  and  $\lambda = q^d$ , the system above is actually the tree recursion of a general anti-ferromagnetic Ising model with parameter  $(\beta, \gamma, \lambda)$ . It follows that such system has either one solution  $(Q^*, Q^*)$  (uniqueness) or three solutions  $(Q^+, Q^-)$ ,  $(Q^*, Q^*)$ ,  $(Q^-, Q^+)$  (non-uniqueness) where  $Q^+ > Q^* > Q^-$  (see [MSW07, Section 6.2] or [GSV16, Theorem 7]). First and foremost, if  $d \geq 5q^k$ , the system (28) is actually the latter case.

**Lemma 4.20.** *When  $q \geq 4, k \geq 2$  and  $d \geq 5q^k$ , the system (28) lies in non-uniqueness region.*

One way to prove Lemma 4.20 is to verify the non-uniqueness condition in [LLY13]. However, in our case, that would cause pages of tedious calculation, and we could not get crucial quantitative information about solutions, which is the key to the stability of the original  $(q+1)$ -spin system. Hence, we show the non-uniqueness by locating the solutions directly, as in the next two lemmas. Also note that, when  $x = R_0/R_1 = C_0/C_1$ , the two-step recursion (28) can be simplified into the following one-step recursion

$$(29) \quad x = \left( \frac{t^2x + qt}{tx + q - 1} \right)^d.$$

**Lemma 4.21.** *Let  $(x, x)$  be the solution of (28) i.e.,  $x$  be the solution of (29). When  $q \geq 4, k \geq 2$  and  $d \geq 5q^k$ , it holds that  $tx + q - 1 < d$ .*

**Lemma 4.22.** *When  $q \geq 4, k \geq 2$  and  $d \geq 5q^k$ , there exists a solution  $(x, y)$  to (28) satisfying (a)  $x > y$ , and (b)  $x > \frac{d}{q^k - q} \cdot d$ .*

We give the proof of Lemmas 4.21 and 4.22 in Section 5.5.

*Proof of Lemma 4.20.* This directly follows from Lemma 4.21 and Lemma 4.22.  $\square$

Now we are ready to analyse the stability of  $(q, 0, 0)$ -type fixpoints. In the following it will be convenient to let  $J$  be the  $q \times q$  matrix with 0s on the diagonal and 1s elsewhere, and  $1$  to be the  $q \times 1$  vector with all ones.

*Proof of Lemma 4.12.* Let  $x = R_0/R_1$  and  $y = C_0/C_1$  be the solution of (28) with  $x > y$ . Set  $a := \sqrt{\frac{1}{tx+q-1}}$ ,  $b := \sqrt{\frac{1}{ty+q-1}}$ ,  $r := \sqrt{\frac{ty}{tx+q}}$  and  $s := \sqrt{\frac{tx}{ty+q}}$ . By applying the formula in Lemma 4.18, the  $(q+1) \times (q+1)$  matrix  $A$  can be written in the block form

$$A = \begin{bmatrix} rs & as1^T \\ br1 & abJ \end{bmatrix}.$$

The eigenvalues of  $L = \begin{bmatrix} 0 & A \\ A^T & 0 \end{bmatrix}$  consist of  $\pm ab$  (with multiplicity  $q-1$  respectively) and  $\pm\lambda_1, \pm\lambda_2$ , where  $\pm\lambda_1, \pm\lambda_2$  are the zeros of the following biquadratic function

$$f(z) = z^4 - ((q-1)^2 a^2 b^2 + q b^2 r^2 + q a^2 s^2 + r^2 s^2) z^2 + a^2 b^2 r^2 s^2.$$

Again, we assume  $\lambda_1 = 1$  (note that  $ab \neq 1$ ). By Vieta's formula,  $\lambda_2 = abrs$ . Since  $rs < 1$ , this means  $ab$  is the second largest eigenvalue. Now it suffices to prove  $ab < 1/d$ , which is equivalent to showing  $(tx+q-1)(ty+q-1) > d^2$ . Note that  $ty > t^{d+1} = q^k - q$ , and Lemma 4.22 gives  $x > d \frac{d}{q^k - q}$ . Therefore  $(tx+q-1)(ty+q-1) > txy > d^2$ .  $\square$

**Remark 4.23.** It is worth noting that the Jacobian stable fixpoints of the system (28) do not necessarily induce  $(q, 0, 0)$ -type Jacobian stable fixpoints of the original  $(q+1)$ -spin system. This is because the eigenvalue  $ab$  from the  $(q+1)$ -spin system is missing in the 2-spin system. Interestingly, by directly applying results over 2-spin system (e.g., [GŠV16, Lemma 8]), what we get is  $abrs < 1/d$  instead of  $ab < 1/d$ . There is an interval of  $d$  such that the former holds but the latter does not. Thus here we cannot only analyze the simplified 2-spin system.

*Proof of Lemma 4.14.* According to the formula in Lemma 4.18, we construct the following  $(q+1) \times (q+1)$  matrix  $A$  with block form

$$A = \begin{bmatrix} b & \sqrt{ab}1^T \\ \sqrt{ab}1 & aJ \end{bmatrix}$$

where  $a := \frac{1}{q-1+tx}$ ,  $b := \frac{tx}{tx+q}$ , and  $x$  is the solution of equation (29). Because  $A$  is symmetric, the spectral radius of  $L = \begin{bmatrix} 0 & A \\ A^T & 0 \end{bmatrix}$  is the same as that of  $A$ . It is not hard to see that  $-a$  is an eigenvalue of  $A$  by multiplicity  $q-1$ . From Lemma 4.21, we have that  $1/d < a$ , and  $a < 1$  from  $q \geq 2$  and  $x > 0$ . Therefore, the fixpoint is unstable.  $\square$

**4.5.  $(q, 0, 0)$  Fixpoint Is Not Maximal.** Let  $q_1 = q, q_2 = q_3 = 0$  and  $R_0/R_1 \neq C_0/C_1$ . Due to stability, it is difficult to analyse this kind of fixpoint's global optimality (recall that it corresponds to a local maxima of  $\Psi_1$ ). However, observe that changing the value of  $R_3$  and  $C_3$  will not affect the value of  $\Phi^S$ . Therefore, we can force  $R_3$  and  $C_3$  to be subject to (12) and (13). As we will show later, doing so allows us to reuse some lemmas we have utilized in our argument regarding 2-maximal fixpoints, among which the most important one is the perturbation argument. We define  $r_0, r_1, c_0, c_3$  analogously, and without loss of generality, suppose  $r_1, c_3 > 1$ .

The next proposition shows how we choose  $r_1$  and  $c_3$ .

**Lemma 4.24.** *Let  $x = r_1$  and  $y = c_3$  be a pair of solutions to the following system*

$$(30) \quad \begin{aligned} f_1(x, y) &:= (x-1) \left( \left( 1 + \frac{x^d(y-1)}{x^d-1} \right)^d t^{d+1} + q - y^d \right) - y^d + 1 = 0; \\ f_2(x, y) &:= (y-1) \left( \left( 1 + \frac{y^d(x-1)}{y^d-1} \right)^d t^{d+1} + qx^d - x^d \right) - x^d + 1 = 0, \end{aligned}$$

with  $x, y > 1$ . Then there exists  $r_0$  and  $c_0$  such that (16) and (17) are satisfied for  $q_1 = q, q_2 = q_3 = 0$ .

*Proof.* The  $r_0$  and  $c_0$  we choose are defined by

$$(31) \quad r_0/t := \frac{r_1 - 1}{c_3^d - 1} + r_1, \quad c_0/t := \frac{c_3 - 1}{r_1^d - 1} + c_3.$$

Combining (31) with the expression of  $f_2(r_1, c_3) = 0$ , it holds that

$$c_0^d t + q - c_3^d - \frac{c_3^d - 1}{r_1 - 1} = 0,$$

which is exactly (24), and is equivalent to the expression for  $r_1$  in (17). The same argument holds for the  $c_3$  expression in (17). In addition, plugging (31) back into (17) yields the expressions for  $r_0, c_0$  in (16).  $\square$

Be cautious that we do not assume  $R_0/R_1 = C_0/C_1$  in Lemma 4.24. Even if we managed to find a pair of solutions  $r_1 > c_3 > 1$  to (30), it does not imply that we can find  $R_3$  and  $C_3$  for the case  $R_0/R_1 \neq C_0/C_1$ , because it is possible for such a pair to correspond to the other case  $R_0/R_1 = C_0/C_1$ . We will handle this in Lemma 4.33 after finding a special solution to (30).

To study the solution of the system (30), we need to look into the properties of both functions. To clarify the intuition of our approach, we plot both functions for the case  $q = 6, k = 3, d = 5q^k$  (see Figure 1a). In this setting, the two functions have three intersections in the region  $(1, +\infty)^2$ : one above  $y = x$ , one near  $y = x$  (but still below  $y = x$ ; see Figure 1b) and one far below  $y = x$ . Experimentally, only the first two intersections correspond to the case  $R_0/R_1 \neq C_0/C_1$ . Hence we would only be interested in them. Moreover, as we will see at the end of this subsection, a solution such that  $x > y$  is required. For this purpose, the rest of the subsection endeavours to prove the existence of the intersection near  $y = x$  before finishing the proof of Lemma 4.13. Doing so also avoids the need of fully characterising the shape of both curves  $f_i(x, y) = 0$ .

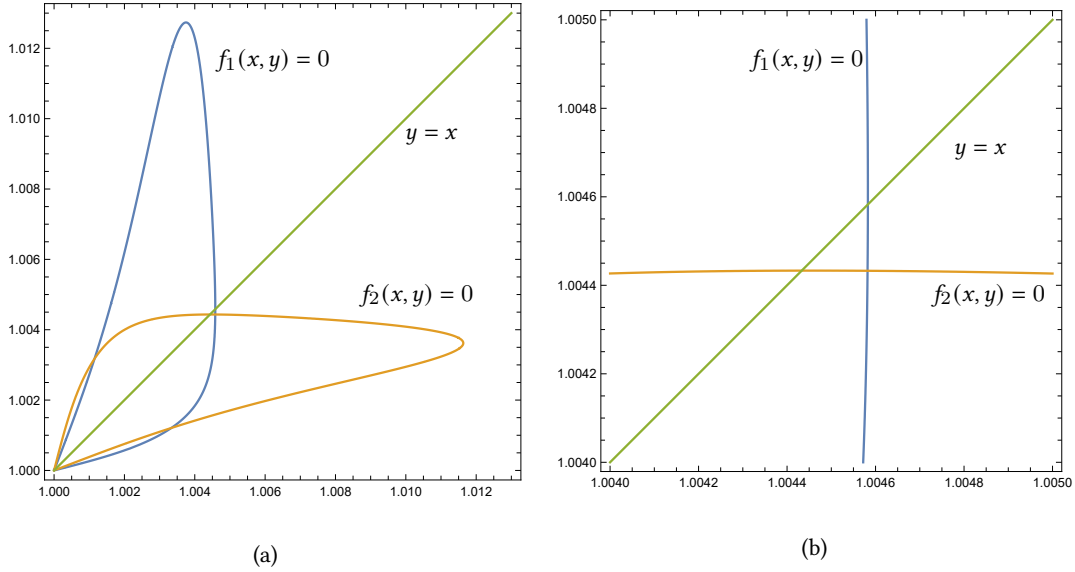


FIGURE 1. (a): Shape of the curve  $f_1(x, y) = 0$ ,  $f_2(x, y) = 0$ , and  $y = x$ . (b): Zoom in on the intersection near  $y = x$ .

Now we formalize our argument. Note that, by mimicking the proof of Lemma 4.19, one can show  $f_2(x, x) = 0$  has exactly one solution  $x^{**} > 1$ . Moreover, for any  $x \in (1, x^{**})$ ,  $f_2(x, x) < 0$ , and for any  $x > x^{**}$ ,  $f_2(x, x) > 0$ . A detailed proof is given in Section 5.4.

**Lemma 4.25.** *For any  $q \geq 4, k \geq 2$  and  $d \geq 3q^k$ , the function*

$$h_2(x) := \left( \frac{x^{d+1} - 1}{x^d - 1} \right)^d t^{d+1} - \frac{x^d - 1}{x - 1} + (q - 1)x^d$$

*has exactly one root  $x^{**}$  in the region  $x > 1$ .*

For  $f_1$ , we do not need the uniqueness of its intersection with the line  $y = x$ .

**Lemma 4.26.** For any  $q \geq 4, k \geq 2$  and  $d \geq 3q^k$ , the function

$$h_1(x) := \left( \frac{x^{d+1} - 1}{x^d - 1} \right)^d t^{d+1} - \frac{x^d - 1}{x - 1} + q - x^d$$

has at least one root in the region  $x > 1$ . Let  $x^*$  be its smallest root. Then  $h_1(x) < 0$  for  $x \in (1, x^*)$ . Moreover,  $x^* > x^{**}$ , and consequently  $h_2(x^*) > 0$ .

*Proof.* The first part of the lemma is similar to the proof of Lemma 4.25 and Lemma 4.19, by computing  $\lim_{x \rightarrow 1} h_1(x) < 0$  and  $\lim_{x \rightarrow +\infty} h_1(x) = +\infty$ . To prove the second part, note that  $h_2(x) > h_1(x)$  for all  $x > 1$ .  $\square$

The next property will be useful later.

**Proposition 4.27.** If  $f_1(x, y) = 0$ , then  $x < 1 + \frac{1}{t^{d+1}-1}$ . If  $f_2(x, y) = 0$ , then  $y < 1 + \frac{1}{t^{d+1}-1}$ .

*Proof.* Suppose  $x \geq 1 + \frac{1}{t^{d+1}-1}$ . Then

$$\begin{aligned} f_1(x, y) &\geq \frac{1}{t^{d+1}-1} \left( \left( 1 + \frac{x^d(y-1)}{x^d-1} \right)^d t^{d+1} + q - y^d \right) - y^d + 1 \\ &> \frac{1}{t^{d+1}-1} (y^d t^{d+1} + q - y^d) - y^d + 1 = \frac{q}{t^{d+1}-1} + 1 > 0. \end{aligned}$$

A similar argument holds for  $f_2$ .  $\square$

Then we study the shape of  $f_1$  below the line  $y = x$ .

**Lemma 4.28.** Let  $g(x) := \frac{(x^d-1)^d}{(x^{d+1}-1)^{d-1}(x-1)}$  and assume that  $d \geq 3q^k$ . Then

- (a) there is a unique  $x_0 \in (1, \infty)$  such that  $g(x_0) = t^{d+1}$ ;
- (b) for any  $1 < x < x_0$ ,  $\frac{\partial f_1}{\partial y} < 0$  for  $y \in (1, x]$ ; and
- (c)  $x_0 > x^*$ , where  $x^* > 1$  is the smallest solution to  $f_1(x, x) = 0$  (see Lemma 4.26).

Moreover, for any  $1 < x < x_0$ ,  $f_1(x, y)$  is decreasing for  $y \in (1, x]$ .

*Proof.* We first show that  $g(x)$  is decreasing for  $x > 1$ . By direct calculation,

$$g'(x) = \frac{(x^d - 1)^{d-1}}{x(x-1)^2(x^{d+1}-1)^d} (x^d d^2 (x-1)^2 - x(x^d - 1)^2) < 0,$$

where the last inequality has already been shown in the proof of Lemma 4.5 for  $x > 1$ . Notice that  $\lim_{x \rightarrow 1} g(x) = \frac{d^d}{(d+1)^{d-1}}$  and  $\lim_{x \rightarrow \infty} g(x) = 1$ . As  $\frac{d^d}{(d+1)^{d-1}} > \frac{d}{e} > q^k > t^{d+1} = q^k - q > 2$ , there is a unique  $x_0$  such that  $g(x_0) = t^{d+1}$  and for  $x \in (1, x_0)$ ,  $g(x) > t^{d+1}$ . This shows part (a).

For part (b), we have  $\frac{\partial f_1}{\partial y} = -x d y^{d-1} + \frac{(x-1) d x^d t^{d+1} (x^d y - 1)^{d-1}}{(x^d - 1)^d}$  and thus,  $\frac{\partial f_1}{\partial y} < 0$  is equivalent to

$$\left( x^{d+1} - \frac{(x^d - 1)^{d/(d-1)}}{(x-1)^{1/(d-1)} t^{\frac{d+1}{d-1}}} \right) y < x.$$

As the range of  $y$  we consider is  $(1, x]$ , we only need to show that

$$x^{d+1} - \frac{(x^d - 1)^{d/(d-1)}}{(x-1)^{1/(d-1)} t^{\frac{d+1}{d-1}}} < 1,$$

which, after rearranging, is equivalent to  $g(x) > t^{d+1}$ . This is guaranteed by part (a) of the lemma.

To prove the third part, by Lemma 4.26, it suffices to show  $h_1(x_0) = f_1(x_0, x_0) > 0$ . Note that  $x_0$  satisfies

$$x_0^{d+1} - \frac{(x_0^d - 1)^{d/(d-1)}}{(x_0 - 1)^{1/(d-1)} t^{\frac{d+1}{d-1}}} = 1, \text{ or equivalently, } \left( \frac{x_0^{d+1} - 1}{x_0^d - 1} \right)^{d-1} = \frac{x_0^d - 1}{t^{d+1}(x_0 - 1)}.$$

By multiplying this with  $\frac{x_0^{d+1}-1}{x_0^d-1}$ , we have  $\left(\frac{x_0^{d+1}-1}{x_0^d-1}\right)^d t^{d+1} = \frac{x_0^{d+1}-1}{x_0-1}$  and plugging into the expression for  $f_1(x, x)$  we get  $f_1(x_0, x_0) = q(x_0 - 1) > 0$ , yielding part (c).  $\square$

By Lemma 4.28 (b) and (c), the partial derivative  $\partial f_1 / \partial y \neq 0$  at all points  $(x, y)$  such that  $f_1(x, y) = 0$  and  $1 < y \leq x \leq x^*$ . Applying the implicit function theorem,  $f_1$  yields a continuous function between  $x$  and  $y$  in the region  $1 < y \leq x \leq x^*$ .

**Corollary 4.29.** *The set  $\mathcal{P}_1^+ := (1, 1) + \{(x, y) : f_1(x, y) = 0, x \geq y > 1, x \leq x^*\}$  forms a continuous curve from  $(1, 1)$  to  $x^*, x^*$ , where  $x^* > 1$  is the smallest solution to  $f_1(x, x) = 0$ .*

Regarding the shape of  $f_2$ , we have the next lemma.

**Lemma 4.30.** *For any  $1 < y < 1 + \frac{1}{q-1}$ , there are at most two  $x > 1$  such that  $f_2(x, y) = 0$ . Moreover, if  $1 < y < x^{**}$ , where  $x^{**} > 1$  is the unique value such that  $f_2(x^{**}, x^{**}) = 0$  (see Lemma 4.25), then there is exactly one  $x > y$  such that  $f_2(x, y) = 0$ .*

*Proof.* The crucial idea of this proof is to study the sign of  $f_2(x, y)$  at its critical points w.r.t.  $x$  (i.e.,  $x'$  such that  $\partial f_2(x, y) / \partial x = 0$  at  $x = x'$ ).

Fix  $y$  in the range and define  $g(x) := f_2(x, y)$ . By direct calculation, if  $g'(x) = 0$ ,  $x'$  satisfies

$$t^{d+1} \left(1 + \frac{y^d(x' - 1)}{y^d - 1}\right)^d = \frac{x'^{d-1}(y - q(y' - 1))(x'y^d - 1)}{(y - 1)y^d}.$$

Plugging it back to  $g$ , we get

$$g(x') = 1 - \frac{x'^{d-1}(y - q(y - 1))}{y^d}.$$

Because  $y - q(y - 1) > 0$ , for any critical point  $x'$  of  $g$ ,

- (a) if  $x' < \chi$ , then  $g(x') > 0$ ;
- (b) if  $x' = \chi$ , then  $g(x') = 0$ ;
- (c) if  $x' > \chi$ , then  $g(x') < 0$ ,

where  $\chi$  is defined by  $\chi := \left(\frac{y^d}{y - q(y - 1)}\right)^{1/(d-1)}$ .

As  $g(x) = f_2(x, y)$  for the fixed  $y$ ,  $g(x)$  is a polynomial in  $x$ . Moreover,  $g(1) = (y - 1)(t^{d+1} + q - 1) > 0$  and  $\lim_{x \rightarrow +\infty} g(x) = +\infty$ . It implies that  $g(x)$  must have an even number of roots. If  $g(x)$  does not have any root greater than 1 then we are done. Otherwise, let  $x_1 > 1$  be the smallest root and  $x_2$  be the largest root.

- If  $x_1 < \chi$ , this means the next critical point  $x' \geq x_1$  cannot be  $x_1$ , or otherwise,  $g(x') = 0$ , contradicting with item (a) above. Therefore,  $g(x') < 0$ , which means  $x' > \chi$ . If there exists another zero  $x' < x_3 < x_2$ , then either  $g'(x_3) < 0$  or  $g'(x_3) > 0$  (otherwise, it contradicts with item (b)). In the former case, there must exist another critical point  $x''$  such that  $\chi < x' < x'' < x_3$  and  $g(x'') > 0$ , which contradicts to item (c). In the latter case, there must exist another critical point  $x'''$  such that  $x_3 < x''' < x_2$  and  $g(x''') > 0$ , violating item (c) as well.
- If  $x_1 > \chi$ , this means all the critical points  $x'$  in  $[x_1, x_2]$  must have function value  $g(x') < 0$ , which implies there is not any other root in  $(x_1, x_2)$ .
- If  $x_1 = \chi$  and  $x_1$  is not a critical point, then  $g'(x_1) < 0$  and the argument of the previous case still applies.
- If  $x_1 = \chi$  and  $x_1$  is a critical point, then for any other critical point (if exists)  $x' > x_1$ , it must holds that  $g(x') < 0$ . Namely once  $g(x)$  becomes positive as  $x$  increases, the sign of  $g'(x)$  will not change. It implies that  $x_2$  is the only root larger than  $x_1$  in this case. If no critical point  $x' > x_1$  exists, then  $x_1 = x_2$  is the only root.

In all cases,  $g(x)$  has at most two roots greater than 1. This finishes the first part of the lemma.

For the second part, notice that if  $y < x^{**}$  then  $g(y) < 0$ , and recall  $\lim_{x \rightarrow \infty} g(x) = \infty$ . The number of zeros larger than  $y$  must be odd, and by the first part, it must be unique. Proposition 4.27 guarantees that  $x^{**} < 1 + \frac{1}{t^{d+1}-1} < 1 + \frac{1}{q-1}$ .  $\square$

We then argue there is a point on  $\mathcal{P}_1^+$  (except  $(1, 1)$ ) such that  $f_2$  takes zero. To establish this, we first find a point  $E$  with  $f_2(x_E, y_E) = 0$  such that it lies to the right of  $\mathcal{P}_1^+$  (with some extra conditions, and later we will apply Lemma 4.30). To simplify the calculation, we only consider the case  $d = 5q^k$ . The proof of the next lemma consists some detailed calculations, which we postpone till Section 5.6.

**Lemma 4.31.** *Suppose  $d = 5q^k$ . There exists a point  $E$  with  $f_2(x_E, y_E) = 0$  such that it lies to the right of  $\mathcal{P}_1^+$ . More specifically, (a)  $y_E = 1 + \frac{0.5}{t^{d+1}-1}$ ; (b)  $y_E < x^{**}$ ; and (c)  $x_E > 1 + \frac{1}{t^{d+1}-1}$ .*

This yields the following lemma.

**Lemma 4.32.** *Suppose  $d = 5q^k$ . The system (30) has a solution  $(x, y)$  such that  $x > y > y_E$ .*

*Proof.* Consider the following point  $M$  on  $\mathcal{P}_1^+$ :  $y_M = y_E$ , and  $x_M$  is the largest one such that  $(x_M, y_M) \in \mathcal{P}_1^+$ .<sup>9</sup> Lemma 4.31 (b) asserts that  $y_E < x^{**}$ , which allows us to invoke Lemma 4.30: for any  $y_E < x < x_E$ , we have  $f_2(x, y_E) < 0$ . More specifically,  $f_2(x_M, y_M) < 0$  because  $x_M < x^* < 1 + \frac{1}{t^{d+1}-1} < x_E$ , where the second and third inequalities come from Proposition 4.27 and Lemma 4.31 (c) respectively.

Now consider the path of  $\mathcal{P}_1^+$  between the point  $M$  and  $(x^*, x^*)$ . It is continuous and bounded away from both  $x = 1$  and  $y = 1$ , and the function  $f_2(x, y)$  is continuous over  $(1, +\infty) \times (1, +\infty)$ . This means as one walks along the path, the value of  $f_2$  changes continuously; otherwise, it violates the continuity of  $f_2$  by a simple  $\varepsilon$ - $\delta$  argument. Moreover, by the second part of Lemma 4.26,  $f_2(x^{**}, x^{**}) > 0$ . This means there must be a point  $(x, y)$  on the path such that  $f_2(x, y) = 0$ . Moreover, by the choice of  $x_M$ , it must hold that  $y > y_M = y_E$ .  $\square$

Now we argue that the solution we find actually satisfies  $R_0/R_1 \neq C_0/C_1$ .

**Lemma 4.33.** *If  $r_1 > c_3 > 1$  yields  $R_0/R_1 = C_0/C_1$ , then it must hold that  $c_3 < y_E$ .*

*Proof.* In this case,  $C_0/C_1$  is the solution to the one-step recursion (29). Define  $u := c_0/t = (C_0/C_1)^{1/d}/t$ . By rewriting (29), one can see that  $u$  is the unique solution to the following equation

$$h(u) := u - \left(1 + \frac{1}{t^{d+1}u^d + q - 1}\right) = 0.$$

Note that  $u > 1$ . Using this notation, (31) yields that  $c_3 = \frac{u(r_1^d - 1) + 1}{r_1^d}$ , giving that  $c_3 < u$ .

It remains to show  $u < 1 + \frac{0.5}{t^{d+1}-1}$ . Note that the system (29) has a unique fixpoint, namely that  $h(u)$  has a unique solution over  $u > 1$ . Because  $h(1) < 0$  and  $\lim_{u \rightarrow \infty} h(u) = \infty$ , it suffices to prove  $h(1 + \frac{0.5}{t^{d+1}-1}) > 0$ . After plugging in the expression and clearing the denominator, it turns out to be equivalent to

$$1 + 3q - 2q^k + (q^k - q) \left(1 + \frac{1}{2q^k - 2(q+1)}\right)^{5q^k} > 0$$

which is true for any  $q \geq 4$  and  $k \geq 2$ .  $\square$

We can finally conclude Lemma 4.13.

*Proof of Lemma 4.13.* Lemma 4.32 guarantees the existence of  $r_1 > c_3 > y_E$  satisfying (30) with  $x = r_1$  and  $y = c_3$ . By Lemma 4.24, given  $r_1$  and  $c_3$ , we can choose  $R_0, R_1, R_3, C_0, C_1, C_3$  to satisfy (12) and (13). Lemma 4.33 implies that for this choice,  $R_0/R_1 \neq C_0/C_1$ . Moreover, the 2-spin system regarding  $R_0/R_1$  and  $C_0/C_1$  lies in non-uniqueness region, and hence the values of  $R_0/R_1$  and  $C_0/C_1$  are unique up to the swap of  $R$  and  $C$  (see Section 4.4).

Because  $R_3$  and  $C_3$  are subject to (12) and (13), the first part of the proof in Lemma 4.15 still holds, even when  $q_3 = 0$  (since we only require (32)). Therefore the expression of  $\partial \overline{\Phi^S} / \partial q_3$  still applies. Based on this, by going through the proof of Lemma 4.16 (a), we can see (27) still holds, i.e.,

$$\text{sgn} \left( \frac{\partial \overline{\Phi^S}}{\partial q_1} - \frac{\partial \overline{\Phi^S}}{\partial q_3} \right) = -\text{sgn}(r_1 - c_3).$$

<sup>9</sup> $x_M$  is well-defined. This follows from the fact that  $(x^d - 1)^d f_1(x, y_M)$  is a non-zero polynomial with respect to  $x$ , thus having a finite number of zeros.

Hence under this choice,  $\partial \overline{\Phi^S} / \partial q_1 - \partial \overline{\Phi^S} / \partial q_3 < 0$ . Now consider a new  $q$  vector  $(q - \varepsilon, 0, \varepsilon)$ . When  $\varepsilon$  is small enough, the value of  $\overline{\Phi^S}$  increases, and feasibility in (11) still holds. Because the value of  $\overline{\Phi^S}$  at  $q = (q, 0, 0)$  is irrelevant to  $R_3, C_3$ , and the value of  $\overline{\Phi^S}$  is the same for all fixpoints of type  $(q, 0, 0)$  and  $R_0/R_1 \neq C_0/C_1$ , it means  $\overline{\Phi}$  does not take the maximum at fixpoints of such type.  $\square$

**Remark 4.34.** Our approach in fact jumps out of the local area around the fixpoint. Intuitively, the argument considers a new “imaginary” fixpoint where  $\varepsilon$  portion of the  $q$  entries  $R_1$  (resp.  $C_1$ ) is changed into  $R_3$  (resp.  $C_3$ , recall that  $R_3$  and  $C_3$  are bounded away from  $R_1$  and  $C_1$ ), and compares its value of the original induced matrix norm with the one of  $(R_0, R_1, \dots, R_1, C_0, C_1, \dots, C_1)$ . This is another reason why optimizing  $\overline{\Phi^S}$  over all nonnegative  $q$ ’s instead of integer  $q$ ’s helps a lot.

## 5. REMAINING PROOFS

**5.1. Proof of Lemma 2.3.** We will consider the following computational problem. Given a graph  $G = (V, E)$ , for a  $q$ -colouring  $\sigma : V \rightarrow \{1, \dots, q\}$ , let  $\text{Mono}(G, \sigma)$  be the number of monochromatic edges under  $\sigma$ .

*Name:* MAX- $q$ -CUT

*Instance:* A undirected graph  $G = (V, E)$

*Output:*  $\max_{\sigma: V \rightarrow \{1, \dots, q\}} \{|E| - \text{Mono}(G, \sigma)\}$

Let MAX-CUT be the  $q = 2$  version of MAX- $q$ -CUT. Alimonti and Kann [AK00] showed the following.

**Proposition 5.1.** *There is a constant  $\delta_0 > 0$  such that, there is no randomized polynomial-time approximation algorithm for MAX-CUT in cubic graphs with relative error  $\delta_0$  unless  $\text{NP} = \text{RP}$ .*

Furthermore, Kann, Khanna, Lagergren, and Panconesi [KKLP97] showed the following reduction.

**Proposition 5.2.** *For any  $0 \leq \delta \leq 1$ , if MAX- $q$ -CUT in  $\left(\frac{\Delta(q+1)}{2} + \frac{q-1}{2}\right)$ -regular graphs can be approximated within relative error  $\frac{\delta}{2(q+1)}$  in polynomial-time, then MAX-CUT can be approximated within  $\delta$  in polynomial-time for  $\Delta$ -regular graphs.*

The original reduction in [KKLP97, Theorem 1] works for only even  $q$  and gives relative error lower bound  $\frac{\delta}{2(q-1)}$  instead. For odd  $q$  they used a different reduction to achieve the same lower bound but it does not keep the degrees bounded. Here we briefly describe how to modify the reduction in [KKLP97, Theorem 1] such that it works for odd  $q$  as well, albeit with a slightly worse relative error lower bound  $\frac{\delta}{2(q+1)}$ . For odd  $q$ , given an instance  $G = (V, E)$  for MAX-CUT, we replace each vertex  $v \in V$  by a clique  $C_v$  of size  $\frac{q+1}{2}$  (instead of  $\frac{q}{2}$  in the original reduction), and replace each edge  $(u, v) \in E$  by a bipartite complete graph between  $C_v$  and  $C_u$ . Moreover, give weight  $\frac{q+1}{q-1} d_G(v)$  for edges inside  $C_v$  (instead of  $d_G(v)$ ) and keep weight 1 for all other edges. It can be verified straightforwardly that the proof still works, except that the parameters  $\alpha$  and  $\beta$  changed from  $(\frac{q(q-1)}{2}, \frac{2}{q})$  to  $(\frac{(q+1)^2}{2}, \frac{2}{q+1})$ , which leads to the worse lower bound  $\frac{\delta}{2(q+1)}$ . Finally, Crescenzi, Silvestri, and Trevisan [CST01] showed that for a general class of combinatorial optimization problem, including MAX- $q$ -CUT, the weighted and unweighted versions have the same approximation complexity.

**Lemma 5.3.** *There is a constant  $0 < C_0 < 1$  such that for any  $q \geq 2$ , there is no FPRAS for the  $q$ -state Potts model with weights  $B < q^{-1/C_0}$  in  $(2q+1)$ -regular graphs unless  $\text{NP} = \text{RP}$ .*

*Proof.* Let  $C_0 := \frac{5\delta_0}{24}$ , where  $\delta_0$  is from Proposition 5.1. We claim that for any  $q \geq 2$ , an FPRAS for  $Z_B(G)$  with weight  $B < q^{-1/C_0}$  in graphs with degree bound  $2q+1$  implies an efficient approximation of MAX- $q$ -CUT within relative error  $\varepsilon_0 := \frac{\delta_0}{2(q+1)}$  in graphs with the same degree bound. Then Proposition 5.1 and Proposition 5.2 (with  $\Delta = 3$ ) imply the lemma.

Given an instance  $G = (V, E)$  to MAX- $q$ -CUT, assume the maximum value of  $q$ -cut is Opt. Let  $n := |V|$  and  $m := |E|$ . Then  $m = \frac{(2q+1)n}{2}$ . If we had an FPRAS for the  $q$ -state Potts model, then we can efficiently

sample a colouring proportional to its weight. (In the local lemma setting, one such reduction is given in [JPV21b].) The probability that the cut value of the colouring is less than  $(1 - \varepsilon_0)\text{Opt}$  is at most

$$\frac{q^n B^{m-(1-\varepsilon_0)\text{Opt}}}{B^{m-\text{Opt}} + q^n B^{m-(1-\varepsilon_0)\text{Opt}}}.$$

In particular, this probability is at most  $1/2$  if

$$B^{m-\text{Opt}} \geq q^n B^{m-(1-\varepsilon_0)\text{Opt}},$$

which is equivalent to  $B^{-\varepsilon_0\text{Opt}} \geq q^n$ . On the other hand, notice that a uniformly at random colouring achieves cut value  $(1 - \frac{1}{q})m$  in expectation. Thus,  $\text{Opt} \geq (1 - \frac{1}{q})m = \frac{(2q+1)(q-1)n}{2q}$ . Consequently, for any  $q \geq 2$ ,  $B^{-\varepsilon_0\text{Opt}} \geq B^{-\delta_0 n \frac{(2q+1)(q-1)}{4q(q+1)}} \geq B^{-C_0 n}$ , since  $C_0 = \frac{5\delta_0}{24} \leq \frac{(2q+1)(q-1)\delta_0}{4q(q+1)}$  for  $q \geq 2$ . Thus if  $B < q^{-1/C_0}$ ,  $B^{-\varepsilon_0\text{Opt}} \geq q^n$  as desired. Standard methods can boost the success probability from  $1/2$  to arbitrarily close to 1.  $\square$

Now we are ready to show Lemma 2.3.

*Proof of Lemma 2.3.* Given a  $(2q+1)$ -regular graph  $G = (V, E)$ , we replace each edge by  $s := \lfloor \frac{\Delta}{2q+1} \rfloor$  parallel edges to get a new graph  $G'$  whose degree is at most  $(2q+1)s \leq \Delta$ . As  $q \geq 2$ ,  $C_1 \geq 5$ , and  $\Delta \geq 2C_1 q \ln q$ ,  $s \geq \frac{\Delta}{2q+1} - 1 > 0.63 \frac{\Delta}{2q+1}$ .

If we have a Potts model with edge weight  $B$  on  $G'$ , then effectively, this is a Potts model on  $G$  with  $B' = B^s$ . Thus if  $B < 1 - \frac{C_1 q \ln q}{\Delta}$  for  $C_1 = 5/C_0$ , where  $C_0$  is from Lemma 5.3, then

$$B^{-sC_0} > \left(1 + \frac{C_1 q \ln q}{\Delta}\right)^{sC_0} \geq e^{\frac{0.8sC_0 C_1 q \ln q}{\Delta}} \geq e^{\frac{0.8 \cdot 0.63 C_0 C_1 q \ln q}{2q+1}} > e^{\frac{C_0 C_1 q \ln q}{2(2q+1)}} \geq q^{0.2C_0 C_1} \geq q,$$

where in the first line we used  $1+x \geq e^{0.8x}$  for  $x \leq 0.5$ . Thus this parallel construction can reduce from the Potts model satisfying the conditions of Lemma 5.3, which is NP-hard to approximate.  $\square$

## 5.2. Proof of Lemmas 4.15 and 4.9.

*Proof of Lemmas 4.15 and 4.9.* We first prove Lemma 4.15. Let

$$S := R_0 C_0 t^2 + \left(\sum_{j=1}^3 C_j q_j\right) R_0 t + \left(\sum_{j=1}^3 R_j q_j\right) C_0 t + \left(\sum_{j=1}^3 R_j q_j\right) \left(\sum_{j=1}^3 C_j q_j\right) - \left(\sum_{j=1}^3 R_j C_j q_j\right),$$

$$R := R_0^{(d+1)/d} + \left(\sum_{j=1}^3 R_j^{(d+1)/d} q_j\right), \quad C := C_0^{(d+1)/d} + \sum_{j=1}^3 C_j^{(d+1)/d} q_j$$

By direct calculation,

$$\frac{\partial \overline{\Phi^S}}{\partial q_i} = \frac{(d+1)}{S} \left[ R_i C_0 t + R_0 C_i t - R_i C_i + R_i \left(\sum_{j=1}^3 C_j q_j\right) + C_i \left(\sum_{j=1}^3 R_j q_j\right) \right] - d \left( \frac{R_i^{(d+1)/d}}{R} + \frac{C_i^{(d+1)/d}}{C} \right).$$

Note that if  $q_i > 0$  and  $R_i \neq 0$ , then it must hold that  $\partial \overline{\Phi^S} / \partial R_i = 0$ , and hence (12) applies, which gives

$$(32) \quad R_0^{(d+1)/d} \propto R_0 (C_0 t^2 + (q_1 C_1 + q_2 C_2 + q_3 C_3) t), \quad R_i^{(d+1)/d} \propto R_i (C_0 t + q_1 C_1 + q_2 C_2 + q_3 C_3 - C_i).$$

Therefore,  $\frac{R_i^{(d+1)/d}}{R} = \frac{R_i C_0 t + R_i \left(\sum_{j=1}^3 C_j q_j\right) - R_i C_i}{S}$  and, similarly,  $\frac{C_i^{(d+1)/d}}{C} = \frac{C_i R_0 t + C_i \left(\sum_{j=1}^3 R_j q_j\right) - R_i C_i}{S}$ . Note that these two equations also hold trivially when  $R_i = 0$  or  $C_i = 0$ , respectively. Putting these together yields the desired expression for  $\frac{\partial \overline{\Phi^S}}{\partial q_i}$  in Lemma 4.15.

For the second part of Lemma 4.15, without loss of generality, suppose  $q_1, q_2 > 0$  and  $\partial \overline{\Phi^S} / \partial q_1 - \partial \overline{\Phi^S} / \partial q_2 > 0$ . Take a positive  $\varepsilon$  and consider  $(q_1 + \varepsilon, q_2 - \varepsilon, q_3)$ . When  $\varepsilon$  is small enough, the entries  $q_1 + \varepsilon$  and  $q_2 - \varepsilon$  are positive, the value of  $\overline{\Phi^S}$  increases, and feasibility in (11) still holds. Hence  $(q_1, q_2, q_3)$  does not maximize  $\overline{\Phi}$ .

Finally we prove Lemma 4.9. Here we have an extra condition that  $q$  is  $m$ -maximal. This means there exists a maximizer  $r, c$  such that for every  $i \neq j$  such that  $q_i, q_j > 0$ , it holds that  $R_i \neq R_j$  and  $C_i \neq C_j$ . From (12) and (13), we obtain that  $r, c$  specify an  $m$ -supported fixpoint of the tree recursion (8).  $\square$

### 5.3. Proof of Lemma 4.7.

*Proof of Lemma 4.7.* We first show that the maximum in (10) cannot be achieved at  $R_0 = 0$  or  $C_0 = 0$ . Assume otherwise. If  $R_0 = 0$ , we have that

$$\left. \frac{\partial \Phi^S}{\partial R_0} \right|_{R_0=0} = \frac{(d+1)t}{S} \cdot (C_0 t + q_1 C_1 + q_2 C_2 + q_3 C_3) > 0$$

where  $S > 0$ . Therefore, increasing  $R_0$  by a sufficiently small amount increases also the value of  $\Phi^S$ , contradiction. An analogous argument applies for  $C_0$ .

Next, we show that at least one of  $R_1, R_2, R_3, C_1, C_2, C_3$  are non-zero. Assume otherwise, then

$$\left. \frac{\partial \Phi^S}{\partial R_1} \right|_{R_1=0} = \frac{d+1}{S} \cdot (C_0 t + (q_1 - 1)C_1 + q_2 C_2 + q_3 C_3) = \frac{d+1}{S} C_0 t > 0,$$

and therefore we obtain a contradiction as above.

Consider now a triple  $(q_1, q_2, q_3)$  with positive entries, and assume w.l.o.g. that the maximum is taken when  $R_1 = 0$ . We claim that  $C_1 > 0$ . Otherwise, by the first part of Lemma 4.15, we have  $\partial \Phi^S / \partial q_1 = 0$ , and  $\partial \Phi^S / \partial q_i > 0$  for some  $i \in \{2, 3\}$  since we cannot have  $R_2 = R_3 = C_2 = C_3 = 0$ . This yields a contradiction to the second part of Lemma 4.15, and therefore  $C_1 > 0$ . Observe also that

$$\left. \frac{\partial \Phi^S}{\partial R_1} \right|_{R_1=0} = \frac{d+1}{S} \cdot (C_0 t + (q_1 - 1)C_1 + q_2 C_2 + q_3 C_3),$$

so by the argument above we conclude that  $C_0 t + (q_1 - 1)C_1 + q_2 C_2 + q_3 C_3 \leq 0$  and therefore  $q_1 < 1$  (since  $C_0, C_1 > 0$ ). This yields that

$$C_1 \geq \frac{1}{1 - q_1} (C_0 t + q_2 C_2 + q_3 C_3) > C_0.$$

On the other hand, since both of  $C_0, C_1$  are nonzero, to achieve the maximum, (13) must hold for  $i = 1$ , which gives  $C_0 > C_1$ , contradiction. Therefore we have  $R_1 > 0$  for triples with positive entries.

Exactly the same argument works for triples of type  $(q_1, 0, q_3)$  with  $q_1, q_3 > 0$ . For the case  $(q, 0, 0)$ , note that  $q \geq 4 > 1$ , which means the partial derivatives with respect to both  $R_1$  and  $C_1$  are positive at  $R_1 = 0$  and  $C_1 = 0$  respectively, and hence the maximum cannot be taken at either  $R_1 = 0$  or  $C_1 = 0$ .

To prove the final part of the lemma, suppose that  $q_i, q_j > 0$ . Since  $R_i, C_i, R_j, C_j > 0$ , we have that (12) and (13) apply, which yields that  $R_i = R_j$  iff  $C_i = C_j$ .  $\square$

### 5.4. Proof of Lemma 4.19 and Lemma 4.25.

*Proof of Lemma 4.19.* We put the expression of  $h$  here for convenient reference.

$$(33) \quad h(x) := \left( \frac{x^{d+1} - 1}{x^d - 1} \right)^d t^{d+1} - \frac{x^d - 1}{x - 1} + q' + (q' - 1)x^d.$$

We have that  $h$  is continuous over  $x \in (1, +\infty)$  and  $\lim_{x \rightarrow +\infty} h(x) = +\infty$ . Using that  $t^{d+1} = t^\Delta = q^k - q$ , we have that

$$\lim_{x \downarrow 1} h(x) = \left( \frac{d+1}{d} \right)^d t^{d+1} - d + q - 1 < eq^k - eq - d + q - 1 < eq^k - d < 0.$$

This implies the existence of  $x$  with  $h(x) = 0$ . To prove the uniqueness of the root, we will show that for any root  $x > 1$  of  $h'(x)$ , it holds that  $h(x) < 0$  (note if such  $x$  does not exist then we are already done), using the fact that  $h$  is differentiable and its derivative is continuous. To see the reason why it is sufficient, note that the number of roots of  $h(x)$  over  $x > 1$  must be odd (because any critical point of  $h$  has value less than zero). Assuming towards contradiction, let  $x_2 > x_1 > 1$  be the smallest two roots. Then  $h'(x_1) > 0$  and  $h'(x_2) < 0$ , indicating there must be some  $x^* \in (x_1, x_2)$  such that  $h'(x^*) = 0$ . However, in this case  $h(x^*) > 0$ , which leads to contradiction.

Next we prove our claim. Take the derivative of  $h$  and let it be zero:

$$h'(x) = d(q' - 1)x^{d-1} - \frac{dx^{d-1}}{x-1} + \frac{x^d - 1}{(x-1)^2} + \frac{dt^{d+1}x^{d-1} \left( \frac{x^{d+1}-1}{x^d-1} \right)^{d-1} (d - dx + x(x^d - 1))}{(x^d - 1)^2} = 0,$$

or equivalently,

$$(34) \quad \left( \frac{x^{d+1} - 1}{x^d - 1} \right)^d t^{d+1} = \frac{(x^d - 1)(x^{d+1} - 1)(x - x^d(d(q'(x-1) - x)(x-1) + x))}{d(x-1)^2 x^d (d - dx + x(x^d - 1))}.$$

Combining (33) and (34), we obtain that for any  $x$  such that  $h'(x) = 0$ , it holds that

$$h(x) = \frac{g(x, d, q')}{d(x-1)^2 x^{d-1} (d - dx + x(x^d - 1))}$$

where

$$(35) \quad g(x, d, q') := dq'(x-1)^2(x+1)(x^d-1)x^{d-1} - (x^d-1)^2(x^{d+1}-1) - d^2x^{d-1}(x-1)^2(1-x^{1+d}+q'(x-1)(x^d+1)).$$

It is not hard to see that  $d - dx + x(x^d - 1) > 0$  for any  $x > 1$ , so, to show  $h(x) < 0$ , it suffices to prove  $g(x, d, q') < 0$  for all  $x > 1$ . This will follow by showing that

$$(36) \quad g(x, d, 0) < 0 \text{ and } g(x, d, q') \text{ is decreasing in } q', \text{ for any } x > 1 \text{ and } d \geq 3,$$

We have  $g(x, d, 0)/(x^{d+1} - 1) = (d^2(x-1)^2 x^{d-1} - (x^d - 1)^2)$ ; the last quantity has been shown negative for all  $x > 1$  in the proof of Lemma 4.5. To prove the monotonicity w.r.t.  $q'$  note that

$$\frac{\partial g}{\partial q'} = -d(x-1)^2 x^{d-1} \left( -(x+1)x^d + d(x-1)(x^d + 1) + x + 1 \right) =: dx^{d-1}(x-1)^2 g_1(x)$$

where  $g_1(x) := -(x+1)x^d + d(x-1)(x^d + 1) + x + 1$ . Note that

$$g'_1(x) = (d+1)(x^{d-1}(d+x-dx) - 1) < 0 \text{ for } x > 1$$

Since  $g_1(1) = 0$ , we obtain  $g_1(x) < 0$  for all  $x > 1$ , proving (36) and concluding the proof of Lemma 4.19.  $\square$

*Proof of Lemma 4.25.* Recall that  $h_2(x) := \left( \frac{x^{d+1}-1}{x^d-1} \right)^d t^{d+1} - \frac{x^d-1}{x-1} + (q-1)x^d$ . We adopt the same idea as the proof of Lemma 4.19 by showing that  $h_2$  takes negative values at critical points. The estimation of  $\lim_{x \rightarrow 1} h_2(x)$  is the same as we did in Lemma 4.19.

Taking the derivative of  $h_2$  and setting it to zero, we get

$$dqx^{d-1} + \frac{dt^{d+1} \left( \frac{x-1}{x^d-1} + x \right)^{d-1} (x(x^d-1) + d(-x) + d)x^{d-1}}{(x^d-1)^2} - \frac{(d+1)x^d}{x-1} + \frac{x^{d+1}-1}{(x-1)^2} = 0,$$

or equivalently,

$$\left( \frac{x^{d+1}-1}{x^d-1} \right)^d t^{d+1} = \frac{x^{-d}(x^d-1)(x^{d+1}-1)(dqx^d - 2dqx^{d+1} + dqx^{d+2} + dx^{d+1} + x^{d+1} - dx^{d+2} - x)}{d(x-1)^2(-x^{d+1} + dx - d + x)}.$$

By plugging this back into the expression for  $h_2(x)$  and simplifying, we obtain that for any  $x$  such that  $h'_2(x) = 0$  it holds that  $h_2(x) = \frac{g(x, d, q)}{d(x-1)^2(d-dx+x(x^d-1))}$ , where

$$g(x, d, q) := -d^2(x-1)^2(x^d(q(x-1) - x) + 1) + dq(x-1)^2(x^d-1) - (x^d-1)^2(x^{d+1}-1)x^{1-d}.$$

Since  $d - dx + x(x^d - 1) > 0$  for any  $x > 1$ , it remains to prove that  $g(x, d, q) < 0$ . Note that  $\frac{\partial g(x, d, q)}{\partial q} = d(x-1)^2(x^d(d(-x) + d + 1) - 1) < 0$  for  $x > 1$  and therefore  $g(x, d, q) < g(x, d, 0)$ . We also have that  $g(x, d, 0)/(x^{d+1} - 1) = (d^2(x-1)^2 - x^{1-d}(x^d - 1)^2) < 0$ , where the inequality follows from the argument below (36). Therefore  $g(x, d, 0) < 0$  for all  $x > 1$ , as desired, finishing the proof.  $\square$

**5.5. Proof of Lemma 4.21 and Lemma 4.22.** We will use the following inequality.

$$(37) \quad \exp\{a\} > \left(1 + \frac{a}{b}\right)^b > \exp\left\{\frac{ab}{a+b}\right\} \quad \text{for all } a, b > 0.$$

*Proof of Lemma 4.21.* Let  $p := tx + q - 1$  and assume for the sake of contradiction that  $p \geq d$ . Let  $w := p/q^k$  and  $c := d/q^k$ , so tha the assumptions of the lemma imply that  $w \geq c \geq 5$ . (29) gives

$$p = q - 1 + t^{d+1} \left(1 + \frac{1}{p}\right)^d q - 1 + t^{d+1} \exp\left\{\frac{d}{p}\right\} < q - 1 + q^k \exp\left\{\frac{c}{w}\right\}.$$

Therefore,  $w < \frac{q-1}{q^k} + \exp\left\{\frac{c}{w}\right\} < \frac{1}{q^{k-1}} + e < 3$ , contradicting  $w \geq 5$ .  $\square$

*Proof of Lemma 4.22.* For any solution  $(x, y)$  of (28),  $x$  satisfies the two-step recursion  $f(x) = 0$ , where

$$f(z) := t^d \left(1 + \frac{1}{t \cdot t^d \left(1 + \frac{1}{tz+q-1}\right)^d + q - 1}\right)^d - z.$$

Take  $x$  as the largest root of  $f$ . Define  $c := d/q^k$ . Because  $\lim_{x \rightarrow \infty} f(x) = -\infty$ , to show (b), it suffices to prove  $f\left(c^2 q^k \frac{q^k}{q^k - q}\right) > 0$ , or equivalently,

$$(38) \quad \left(1 + \frac{1}{(q^k - q)D + q - 1}\right)^d > tc^2 \left(\frac{q^k}{q^k - q}\right)^2 \quad \text{where} \quad D := \left(1 + \frac{1}{t \left(c^2 q^k \frac{q^k}{q^k - q}\right) + q - 1}\right)^d.$$

Because  $D < \exp\left\{\frac{d}{c^2 q^k}\right\} < \exp\left\{\frac{1}{c}\right\} < 1.2215$ ,

$$\text{LHS of (38)} > \left(1 + \frac{1}{1.2215(q^k - q) + q - 1}\right)^{cq^k} > 2.2674^c,$$

where the last inequality follows from (37). Moreover, for any  $q \geq 4, k \geq 2, d \geq 5q^k$ , we have  $(q^k/(q^k - q))^2 < 1.7778$  and  $t < 1.0312$ . Therefore, RHS of (38)  $< 1.8332c^2$ , which is smaller than  $2.2674^c$  whenever  $c \geq 5$ . This concludes (b). Part (a) follows from (b) and Lemma 4.21.  $\square$

## 5.6. Proof of Lemma 4.31.

*Proof of Lemma 4.31.* Define  $s := \frac{d}{t^{d+1}-1}$ . By Proposition 4.27, any point on  $x = 1 + \frac{s}{d}$  must be on the right of  $\mathcal{P}_1^+$ . Therefore we are interested in the point  $(x, y)$  where  $x = 1 + \frac{s}{d}$  and  $y = 1 + \frac{s}{2d}$ . Specifically, we will show  $f_2(x, y) < 0$ , which, together with the fact that  $\lim_{x \rightarrow +\infty} f_2(x, y) = +\infty$  for any fixed  $y > 1$ , implies the existence of  $x_E > x$  such that  $f_2(x_E, y) = 0$ . However, in order to apply Lemma 4.30, we further need to show  $y < x^{**}$ . The latter can be done by proving  $f_2(y, y) < 0$  due to Lemma 4.25.

We deal with the latter one first. Assume  $q \geq 4, k \geq 3$ , or  $q \geq 12, k \geq 2$ . Then  $5 < s < 5.4962$ ,  $\frac{q^k - q}{2(q^k - q - 1)} < 0.5085$  and  $\left(1 - \frac{q-1}{2(q^k - q - 1)}\right) > 0.9580$ . Set  $D := \left(1 + \frac{s}{2d}\right)^d$ . By using (37), one can show  $D > \exp\{5/2\} > 12.1824$ . Therefore,

$$\begin{aligned} f_2\left(1 + \frac{s}{2d}, 1 + \frac{s}{2d}\right) &= 1 + \left(1 + \frac{s \left(1 + \frac{1}{-1+D}\right)}{2d}\right)^d \frac{q^k - q}{2(q^k - q - 1)} - D \left(1 - \frac{q-1}{2(q^k - q - 1)}\right) \\ &< 1 + \exp\left\{\frac{s}{2} \left(1 + \frac{1}{-1+D}\right)\right\} \frac{q^k - q}{2(q^k - q - 1)} - D \left(1 - \frac{q-1}{2(q^k - q - 1)}\right) \\ &< 1 + 0.5085 \exp\left\{\frac{5.4962}{2} \left(1 + \frac{1}{-1+D}\right)\right\} - 0.9580D < 0, \end{aligned}$$

where in the last inequality we use the fact that the function is decreasing in  $D$ . The cases  $(q, k) = (4, 2), (6, 2), (8, 2), (10, 2)$  also holds by directly computing  $f_2$ .

The first one can be handled similarly. Denote  $E := (1 + \frac{s}{d})^d$ . Then  $D > E^{1/2}$ . By using (37) again,  $E > \exp\{5\}$ . Consider the case  $q \geq 8, k \geq 3$ , or  $q \geq 28, k \geq 2$ . Then  $5 < s < 5.1921$ ,  $\frac{q^k - q}{2(q^k - q - 1)} < 0.5010$  and  $(1 - \frac{q-1}{2(q^k - q - 1)}) > 0.9821$ . Therefore,

$$\begin{aligned} f_2\left(1 + \frac{s}{d}, 1 + \frac{s}{2d}\right) &= 1 + \left(1 + \frac{s(1 + \frac{1}{-1+D})}{d}\right)^d \frac{q^k - q}{2(q^k - q - 1)} - E\left(1 - \frac{q-1}{2(q^k - q - 1)}\right) \\ &< 1 + \exp\left\{s\left(1 + \frac{1}{-1+D}\right)\right\} \frac{q^k - q}{2(q^k - q - 1)} - E\left(1 - \frac{q-1}{2(q^k - q - 1)}\right) \\ &< 1 + \exp\left\{s\left(1 + \frac{1}{-1+E^{1/2}}\right)\right\} \frac{q^k - q}{2(q^k - q - 1)} - E\left(1 - \frac{q-1}{2(q^k - q - 1)}\right) \\ &< 1 + 0.5010 \exp\left\{5.1921\left(1 + \frac{1}{-1+E^{1/2}}\right)\right\} - 0.9821E < 0, \end{aligned}$$

where in the last inequality we use the fact that the function is decreasing in  $E$ . The remaining cases  $(q, k) = (4, 3), (6, 3), (4, 2), (6, 2), \dots, (26, 2)$  also holds by directly computing  $f_2$ .  $\square$

#### REFERENCES

- [AK00] Paola Alimonti and Viggo Kann. Some APX-completeness results for cubic graphs. *Theor. Comput. Sci.*, 237(1-2):123–134, 2000.
- [Alo91] Noga Alon. A parallel algorithmic version of the local lemma. *Random Struct. Algorithms*, 2(4):367–378, 1991.
- [Bec91] József Beck. An algorithmic approach to the Lovász local lemma. I. *Random Struct. Algorithms*, 2(4):343–366, 1991.
- [BGG<sup>+</sup>19] Ivona Bezáková, Andreas Galanis, Leslie Ann Goldberg, Heng Guo, and Daniel Štefankovič. Approximation via correlation decay when strong spatial mixing fails. *SIAM J. Comput.*, 48(2):279–349, 2019.
- [CCGL12] Jin-Yi Cai, Xi Chen, Heng Guo, and Pinyan Lu. Inapproximability after uniqueness phase transition in two-spin systems. In *COCO*, volume 7402 of *Lecture Notes in Computer Science*, pages 336–347. Springer, 2012.
- [CFMR96] Colin Cooper, Alan Frieze, Michael Molloy, and Bruce Reed. Perfect matchings in random  $r$ -regular,  $s$ -uniform hypergraphs. *Combinatorics, Probability and Computing*, 5(1):1–14, 1996.
- [CGSV21] Zongchen Chen, Andreas Galanis, Daniel Stefankovic, and Eric Vigoda. Sampling colorings and independent sets of random regular bipartite graphs in the non-uniqueness region. *arXiv*, abs/2105.01784, 2021.
- [CS00] Artur Czumaj and Christian Scheideler. Coloring nonuniform hypergraphs: A new algorithmic approach to the general Lovász local lemma. *Random Struct. Algorithms*, 17(3-4):213–237, 2000.
- [CST01] Pierluigi Crescenzi, Riccardo Silvestri, and Luca Trevisan. On weighted vs unweighted versions of combinatorial optimization problems. *Inf. Comput.*, 167(1):10–26, 2001.
- [DD20] Andreas Darmann and Janosch Döcker. On a simple hard variant of Not-All-Equal 3-SAT. *Theor. Comput. Sci.*, 815:147–152, 2020.
- [DFJ02] Martin E. Dyer, Alan M. Frieze, and Mark Jerrum. On counting independent sets in sparse graphs. *SIAM J. Comput.*, 31(5):1527–1541, 2002.
- [EHK98] Thomas Emden-Weinert, Stefan Hougardy, and Bernd Kreuter. Uniquely colourable graphs and the hardness of colouring graphs of large girth. *Combin. Probab. Comput.*, 7(4):375–386, 1998.

- [EL75] P. Erdős and L. Lovász. Problems and results on 3-chromatic hypergraphs and some related questions. In *Infinite and finite sets (Colloq., Keszthely, 1973; dedicated to P. Erdős on his 60th birthday)*, Vol. II, pages 609–627. Colloq. Math. Soc. János Bolyai, Vol. 10. 1975.
- [FGYZ21] Weiming Feng, Heng Guo, Yitong Yin, and Chihao Zhang. Fast sampling and counting  $k$ -SAT solutions in the local lemma regime. *J. ACM*, 2021. to appear.
- [FHY20] Weiming Feng, Kun He, and Yitong Yin. Sampling constraint satisfaction solutions in the local lemma regime. *arXiv*, abs/2011.03915, 2020.
- [FM13] Alan Frieze and Dhruv Mubayi. Coloring simple hypergraphs. *J. Combin. Theory Ser. B*, 103(6):767–794, 2013.
- [GG16] Andreas Galanis and Leslie Ann Goldberg. The complexity of approximately counting in 2-spin systems on  $k$ -uniform bounded-degree hypergraphs. *Information and Computation*, 251:36–66, 2016.
- [GJL19] Heng Guo, Mark Jerrum, and Jingcheng Liu. Uniform sampling through the Lovász local lemma. *J. ACM*, 66(3):18:1–18:31, 2019.
- [GLLZ19] Heng Guo, Chao Liao, Pinyan Lu, and Chihao Zhang. Counting hypergraph colorings in the local lemma regime. *SIAM J. Comput.*, 48(4):1397–1424, 2019.
- [GST16] Heidi Gebauer, Tibor Szabó, and Gábor Tardos. The local lemma is asymptotically tight for SAT. *J. ACM*, 63(5):43:1–43:32, 2016.
- [GŠV15] Andreas Galanis, Daniel Štefankovič, and Eric Vigoda. Inapproximability for antiferromagnetic spin systems in the tree nonuniqueness region. *J. ACM*, 62(6):50, 2015.
- [GŠV16] Andreas Galanis, Daniel Štefankovič, and Eric Vigoda. Inapproximability of the partition function for the antiferromagnetic Ising and hard-core models. *Combin. Probab. Comput.*, 25(4):500–559, 2016.
- [HSW21] Kun He, Xiaoming Sun, and Kewen Wu. Perfect sampling for (atomic) Lovász local lemma. *arXiv*, abs/2107.03932, 2021.
- [HSZ19] Jonathan Hermon, Allan Sly, and Yumeng Zhang. Rapid mixing of hypergraph independent sets. *Random Struct. Algorithms*, 54(4):730–767, 2019.
- [JPV21a] Vishesh Jain, Huy Tuan Pham, and Thuy Duong Vuong. On the sampling Lovász local lemma for atomic constraint satisfaction problems. *arXiv*, abs/2102.08342, 2021.
- [JPV21b] Vishesh Jain, Huy Tuan Pham, and Thuy Duong Vuong. Towards the sampling Lovász local lemma. In *62nd IEEE Annual Symposium on Foundations of Computer Science, FOCS 2021, Denver, CO, USA, February 7-10, 2022*, pages 173–183. IEEE, 2021.
- [JVV86] Mark Jerrum, Leslie G. Valiant, and Vijay V. Vazirani. Random generation of combinatorial structures from a uniform distribution. *Theoretical Computer Science*, 43:169–188, 1986.
- [KKLP97] Viggo Kann, Sanjeev Khanna, Jens Lagergren, and Alessandro Panconesi. On the hardness of approximating max  $k$ -cut and its dual. *Chic. J. Theor. Comput. Sci.*, 1997, 1997.
- [KST93] Jan Kratochvíl, Petr Savický, and Zsolt Tuza. One more occurrence of variables makes satisfiability jump from trivial to NP-complete. *SIAM Journal on Computing*, 22(1):203–210, 1993.
- [LLY13] Liang Li, Pinyan Lu, and Yitong Yin. Correlation decay up to uniqueness in spin systems. In *SODA*, pages 67–84. SIAM, 2013. Full version from arXiv at abs/1111.7064.
- [Moi19] Ankur Moitra. Approximate counting, the Lovász local lemma, and inference in graphical models. *J. ACM*, 66(2):10:1–10:25, 2019.
- [Mos09] Robin A. Moser. A constructive proof of the Lovász local lemma. In *STOC*, pages 343–350. ACM, 2009.
- [MR98] Michael Molloy and Bruce A. Reed. Further algorithmic aspects of the local lemma. In *STOC*, pages 524–529. ACM, 1998.
- [MSW07] Fabio Martinelli, Alistair Sinclair, and Dror Weitz. Fast mixing for independent sets, colorings, and other models on trees. *Random Structures & Algorithms*, 31(2):134–172, 2007.
- [MT10] Robin A. Moser and Gábor Tardos. A constructive proof of the general Lovász local lemma. *J. ACM*, 57(2):11, 2010.
- [MWW09] Elchanan Mossel, Dror Weitz, and Nicholas Wormald. On the hardness of sampling independent sets beyond the tree threshold. *Probability Theory and Related Fields*, 143(3-4):401–439,

2009.

- [PP19] Konstantinos Panagiotou and Matija Pasch. Satisfiability thresholds for regular occupation problems. In *ICALP*, volume 132 of *LIPICs*, pages 90:1–90:14. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2019.
- [Sly10] Allan Sly. Computational transition at the uniqueness threshold. In *FOCS*, pages 287–296. IEEE Computer Society, 2010.
- [Sri08] Aravind Srinivasan. Improved algorithmic versions of the Lovász local lemma. In *SODA*, pages 611–620. SIAM, 2008.
- [SS14] Allan Sly and Nike Sun. Counting in two-spin models on  $d$ -regular graphs. *Ann. Probab.*, 42(6):2383–2416, 2014.