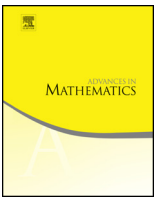
Advances in Mathematics 430 (2023) 109212

Contents lists available at ScienceDirect

Advances in Mathematics

journal homepage: www.elsevier.com/locate/aim



Homomorphisms from the torus ☆

Matthew Jenssen [a,*], Peter Keevash [b,*]

[a] *King's College London, United Kingdom*
[b] *University of Oxford, United Kingdom*

ARTICLE INFO

Article history:
Received 27 October 2021
Received in revised form 11 April 2023
Accepted 3 July 2023
Available online 24 July 2023
Communicated by Benjamin Sudakov

MSC:
primary 05C30, 05C31, 82B20

Keywords:
Cluster expansion
Polymer model
Graph homomorphism
Discrete spin system
Discrete torus

ABSTRACT

We present a detailed probabilistic and structural analysis of the set of weighted homomorphisms from the discrete torus $\mathbb{Z}_m^n$, where m is even, to any fixed graph: we show that the corresponding probability distribution on such homomorphisms is close to a distribution defined constructively as a certain random perturbation of some dominant phase. This has several consequences, including solutions (in a strong form) to conjectures of Engbers and Galvin and a conjecture of Kahn and Park. Special cases include sharp asymptotics for the number of independent sets and the number of proper q-colourings of $\mathbb{Z}_m^n$ (so in particular, the discrete hypercube). We give further applications to the study of height functions and (generalised) rank functions on the discrete hypercube and disprove a conjecture of Kahn and Lawrenz. For the proof we combine methods from statistical physics, entropy and graph containers and exploit isoperimetric and algebraic properties of the torus.

Contents

☆ Research supported in part by ERC Consolidator Grant 647678.
* Corresponding authors.
E-mail addresses: matthew.jenssen@kcl.ac.uk (M. Jenssen), keevash@maths.ox.ac.uk (P. Keevash).

https://doi.org/10.1016/j.aim.2023.109212

1. Introduction

1.1. Spin models

A central notion at the intersection of combinatorics and statistical physics is that of a *graph homomorphism*. Let G and H be graphs (possibly with loops). We call a function $f : V(G) \to V(H)$ a *homomorphism* (or *H-colouring*) if the map preserves edges, that is, $f(e) \in E(H)$ for all $e \in E(G)$. We write $\mathrm{Hom}(G, H)$ for the set of homomorphisms from G into H. From a combinatorial perspective, graph homomorphisms provide a unifying framework for a number of important graph theory concepts. For the statistical physicist, homomorphisms arise in the study of *spin models* and their critical phenomena. Tools from each field have enriched the other and recent years have witnessed an explosion of literature at their interface.

A *weighted graph* is a pair (H, λ) where H is a graph (possibly with loops) and $\lambda : V(H) \to \mathbb{R}_{>0}$ is a function assigning each vertex of H an 'activity' λ_v. Weighted graphs provide a rich set of probability distributions on the space $\mathrm{Hom}(G, H)$. Indeed, given a weighted graph (H, λ), we may define a probability measure $\mu_{H,\lambda}$ on $\mathrm{Hom}(G, H)$ given by

$$\mu_{H,\lambda}(f) \coloneqq \frac{\prod_{v \in V(G)} \lambda_{f(v)}}{Z_G^H(\lambda)}\,, \tag{1.1}$$

for $f \in \mathrm{Hom}(G, H)$ where

$$Z_G^H(\lambda) \coloneqq \sum_{f \in \mathrm{Hom}(G,H)} \prod_{v \in V(G)} \lambda_{f(v)}\,. \tag{1.2}$$

This type of probability distribution is called a *spin model* (with hard constraints) and the normalising factor $Z_G^H(\lambda)$ is the *partition function* of the model. One may think of the vertices of G as sites and the vertices of H as a set of particles (or spins) that can occupy these sites. The edges of G represent bonds between sites and the edges of H represent constraints on which pairs of particles can occupy bonded sites. A spin model is then a probability distribution on the set of legal spin configurations on the sites of G. Two of our main motivating examples will be the following:

Example 1. *The q-colouring model (zero-temperature anti-ferromagnetic Potts model):* $H = K_q$, $\lambda \equiv 1$. In this case $\mu_{H,\lambda}$ is the uniform measure over proper q-colourings of G and $Z_G^H(\lambda)$ is the number of such colourings.

Example 2. *The hard-core model:* $H =$  , that is, an edge $\{v_{\text{in}}, v_{\text{out}}\}$ with a loop at vertex v_{out}. We let $\lambda(v_{\text{out}}) = 1$ and $\lambda(v_{\text{in}}) = x$ for some fixed $x > 0$ called the *fugacity*. $Z_G^H(\lambda)$ is the hard-core model partition function (also known as the independence polynomial). In particular when $x = 1$, $Z_G^H(\lambda)$ is the number of independent sets in G.

In the statistical physics literature, spin models are traditionally studied on the integer lattice $\mathbb{Z}^n$, a setting where the phenomenon of *phase coexistence* can be rigorously studied. We mention two landmark results in this field. Galvin and Kahn [24] establish phase coexistence for the hard-core model on $\mathbb{Z}^n$ where the fugacity is allowed to tend to 0 with n (see also [45]) and Peled and Spinka [50] prove phase coexistence for the q-colouring model. Informally these results show that in large subregions of $\mathbb{Z}^n$, a typical sample from $\mu_{H,\lambda}$ (with the appropriate choice of (H, λ)) exhibits long-range order by correlating with some dominant phase.

In this paper we study spin models on the n-dimensional discrete torus $\mathbb{Z}_m^n$ where m is a fixed even integer and n is large. In this setting we establish the phase coexistence phenomenon in a strong form: we show that on $\mathbb{Z}_m^n$, the measure $\mu_{H,\lambda}$ can be closely approximated by a measure defined constructively as a random perturbation of a random dominant phase. The random perturbation has the distribution of a *polymer model* with convergent *cluster expansion*. Via the cluster expansion, we are able to gain an essentially complete probabilistic description of the measure $\mu_{H,\lambda}$ and hence a precise structural description of the set $\text{Hom}(\mathbb{Z}_m^n, H)$. Establishing convergence of the cluster expansion is a non-trivial task and requires a combination of entropy tools, the method of graph containers and algebraic and isoperimetric properties of the torus. Our synthesis of container and entropy methods has its roots in the work of Peled and Spinka [50] and later appears in the work of Kahn and Park [38] (see Section 4 for a more detailed discussion of the use of entropy tools and the container method in this context).

In order to formally state our main result (Theorem 1.2 below) we require some preliminary notation and language from the theory of polymer models which we will now introduce.

1.2. The defect polymer model

The n-dimensional discrete torus $\mathbb{Z}_m^n$ is the graph on vertex set $\{0,\dots,m-1\}^n$ where two vertices x, y are adjacent if and only if there exists a coordinate $i \in [n]$ such that $x_i = y_i \pm 1 \pmod m$ and $x_j = y_j$ for all $j \neq i$. The case $m = 2$ returns the n-dimensional hypercube for which we use the more familiar notation Q_n. We note that when m is even, $\mathbb{Z}_m^n$ is a bipartite graph and we denote its vertex classes by $\mathcal{E}, \mathcal{O}$. We note also that $\mathbb{Z}_m^n$ is a d-regular graph where $d = (1 + \mathbb{1}_{m>2})n$.

Given a weighted graph (H, λ) and $A, B \subseteq V(H)$, we write $A \sim B$ if $\{a, b\} \in E(H)$ for all $a \in A$ and $b \in B$. We call such a pair (A, B) a *pattern*. Letting $\lambda_X \coloneqq \sum_{v\in X} \lambda_v$ for $X \subseteq V(H)$ we define

$$\eta_\lambda(H) \coloneqq \max\{\lambda_A \lambda_B : A, B \subseteq V(H), A \sim B\}. \tag{1.3}$$

Following Engbers and Galvin [14], we call a pattern (A, B) *dominant* if $\lambda_A\lambda_B = \eta_\lambda(H)$, and we let $\mathcal{D}_\lambda(H)$ denote the collection of all dominant patterns. In Example 1 of the previous section, the dominant patterns are all pairs (A, B) where $V(K_q) = A \cup B$ and $\{|A|, |B|\} = \{\lfloor q/2\rfloor, \lceil q/2\rceil\}$. In Example 2, there are two dominant patterns: $(\{v_{\text{out}}\}, \{v_{\text{in}}, v_{\text{out}}\})$ and $(\{v_{\text{in}}, v_{\text{out}}\}, \{v_{\text{out}}\})$. We call a homomorphism $f \in \text{Hom}(\mathbb{Z}_m^n, H)$ a *dominant colouring* if $f(\mathcal{E}) \subseteq A$, $f(\mathcal{O}) \subseteq B$ for some dominant pattern (A, B).

In the following, we let G denote $\mathbb{Z}_m^n$ and we fix a weighted graph (H, λ). We say that a subset $S \subseteq V(G)$ is G^2*-connected* if the graph $G^2[S]$ is connected (here G^2 denotes the square of the graph G). We will define a family $\mathcal{P}$ of G^2-connected subsets of $V(G)$ the elements of which we call *polymers* (see Definition 3.1 for a full description of $\mathcal{P}$). For now we can safely think of $\mathcal{P}$ as the set of *all* G^2-connected subsets of $V(G)$.

Given a colouring $f \in \text{Hom}(G, H)$ and a pattern (A, B), we say that f *agrees* with (A, B) at $v \in V(G)$ if $v \in \mathcal{O}$ and $f(v) \in A$ or if $v \in \mathcal{E}$ and $f(v) \in B$. We say that f disagrees with (A, B) at v otherwise. For a subset $S \subseteq V(G)$, let $\chi_{A,B}(S)$ be the set of $f \in \text{Hom}(G, H)$ such that f *disagrees* with (A, B) at each $v \in S$ and agrees with (A, B) at each $v \in V\backslash S$. We define the *weight* of S (with respect to (A, B)) to be

$$w_{A,B}(S) = \frac{\sum_{f\in\chi_{A,B}(S)} \prod_{v\in V} \lambda_{f(v)}}{\eta_\lambda(H)^{m^n/2}}. \tag{1.4}$$

We say two polymers γ_1, γ_2 are *compatible* if the graph distance between γ_1, γ_2 in G is > 2; that is if $\gamma_1 \cup \gamma_2$ is not G^2-connected. Otherwise we say that γ_1 and γ_2 are *incompatible*. We let Ω denote the family consisting of all sets of mutually compatible polymers. The weights $w_{A,B}$ allow us to construct a probability distribution $\nu_{A,B}$ on Ω. For $\Gamma \in \Omega$, let

$$\nu_{A,B}(\Gamma) = \frac{\prod_{\gamma\in\Gamma} w_{A,B}(\gamma)}{\Xi_{A,B}}, \tag{1.5}$$

where the normalising factor

$$\Xi_{A,B} = \sum_{\Gamma \in \Omega} \prod_{\gamma \in \Gamma} w_{A,B}(\gamma) \tag{1.6}$$

is the *polymer model partition function.* We thus have a polymer model for each dominant pattern (A, B). We call $\nu_{A,B}$ the (A, B) *polymer model.* Using these polymer models, we construct a probability distribution $\hat{\mu}_{H,\lambda}$ which will serve as a very good approximation to the measure $\mu_{H,\lambda}$.

Definition 1.1. For $f \in \mathrm{Hom}(G, H)$, let $\hat{\mu}_{H,\lambda}(f)$ denote the probability that f is selected by the following three-step process:

1. Choose $(A, B) \in \mathcal{D}_\lambda(H)$ with probability proportional to $\Xi_{A,B}$.
2. Choose a random polymer configuration $\Gamma \in \Omega$ from the distribution $\nu_{A,B}$.
3. Letting $S = \bigcup_{\gamma \in \Gamma} \gamma$, select a colouring $f \in \chi_{A,B}(S)$ with probability proportional to $\prod_{v \in V} \lambda_{f(v)}$.

We may now state our main theorem. For two probability measures μ and $\hat{\mu}$, we let $\|\hat{\mu} - \mu\|_{TV}$ denote their total variation distance.

Theorem 1.2. *For $m \geq 2$ a fixed even integer and (H, λ) a fixed weighted graph, the measures $\mu_{H,\lambda}$ and $\hat{\mu}_{H,\lambda}$ on $\mathrm{Hom}(\mathbb{Z}_m^n, H)$ satisfy*

$$\|\hat{\mu}_{H,\lambda} - \mu_{H,\lambda}\|_{TV} \leq e^{-\Omega(m^n/n^2)}.$$

Given a colouring $f \in \mathrm{Hom}(\mathbb{Z}_m^n, H)$ we may identify the dominant pattern (A, B) which agrees most with f (breaking ties arbitrarily if necessary) and call the vertices at which f disagrees with (A, B) the *defect vertices.* We can think of Step 1 in the definition of $\hat{\mu}_{H,\lambda}$ as selecting the dominant pattern (A, B) which will agree most with f. Step 2 then identifies the defect vertices. Theorem 1.2 shows that up to very small error, the defect vertices of a sample from $\mu_{H,\lambda}$ have the distribution of a mixture of the polymer models $\nu_{A,B}$. The power of Theorem 1.2 stems from the fact that each of these polymer models admits a *convergent cluster expansion* allowing us to obtain an essentially complete understanding of the measures $\nu_{A,B}$ and therefore of $\mu_{H,\lambda}$ also.

In the remainder of this introduction we discuss the various consequences of Theorem 1.2.

1.3. The structure of $\mathrm{Hom}(\mathbb{Z}_m^n, H)$

Theorem 1.2, combined with a large deviation result for the measures $\nu_{A,B}$ (see Theorem 11.1), can be used to prove a detailed structural description of the set $\mathrm{Hom}(\mathbb{Z}_m^n, H)$.

For a weighted graph (H, λ) and dominant pattern $(A, B) \in \mathcal{D}_{\lambda(H)}$, we say that an H-colouring of $\mathbb{Z}_m^n$ is λ-*balanced* with respect to (A, B) if for each $k \in A$, the proportion of vertices of $\mathcal{O}$ coloured k is within $1/n$ of λ_k/λ_A and for each $\ell \in B$, the proportion of vertices of $\mathcal{E}$ coloured ℓ is within $1/n$ of λ_ℓ/λ_B.

The following theorem provides a structural decomposition of the set $\mathrm{Hom}(\mathbb{Z}_m^n, H)$ showing in particular that, with respect to the measure $\mu_{H,\lambda}$, almost all elements of $\mathrm{Hom}(\mathbb{Z}_m^n, H)$ are λ-balanced with respect to some dominant pattern. In what follows we use $a \pm b$ to denote a quantity that lies in the interval $[a - b, a + b]$.

Theorem 1.3. *Fix a weighted graph (H, λ) and $m \geq 2$ even. There is a partition of $\mathrm{Hom}(\mathbb{Z}_m^n, H)$ into $|\mathcal{D}_\lambda(H)| + 1$ classes*

$$\mathrm{Hom}(\mathbb{Z}_m^n, H) = F(0) \cup \bigcup_{(A,B)\in\mathcal{D}_\lambda(H)} F(A, B)$$

with the following properties. With $\mu = \mu_{H,\lambda}$,

1. *$\mu(F(0)) \leq e^{-\Omega(m^n/n^3)}$.*
2. *For $(A, B) \in \mathcal{D}_\lambda(H)$, each $f \in F(A, B)$ is λ-balanced with respect to (A, B).*
3. *For each $(A, B) \in \mathcal{D}_\lambda(H)$,*

$$\mu(F(A, B)) = \frac{1}{Z_G^H(\lambda)} \eta_\lambda(H)^{\frac{m^n}{2}} \Xi_{A,B} \left(1 \pm e^{-\Omega(m^n/n^3)}\right).$$

Theorem 1.3 is inspired by a decomposition result of Engbers and Galvin [14, Theorem 1.2]. We in fact prove a strengthening of Theorem 1.3 (Theorem 13.2), which directly strengthens [14, Theorem 1.2] and resolves conjectures of the authors in a strong form [14, Conjectures 6.1, 6.2, 6.3]. We defer a more detailed discussion of these conjectures to Section 13.

It is natural to wonder to what extent the measures of the classes $F(A, B)$ in Theorem 1.3 can differ. We remark that, in contrast to the examples of q-colouring and the hard-core model, it is easy to construct (H, λ) and $(A, B), (A', B') \in \mathcal{D}_\lambda(H)$ such that $\mu(F(A', B')) = o(\mu(F(A, B)))$ (see the discussion after the proof of Theorem 13.2 for more).

As mentioned in the previous section, we will be able to gain a precise understanding of the measures $\nu_{A,B}$ and therefore also of the partition functions $\Xi_{A,B}$. Via conclusion 3 of Theorem 1.3, we can therefore obtain detailed asymptotic expressions for the partition function $Z_G^H(\lambda)$. In particular this yields a plethora of asymptotic formulae for combinatorial quantities such as the number of proper q-colourings and the number of independent sets in $\mathbb{Z}_m^n$.

1.4. Asymptotic enumeration and the cluster expansion

In order to state our approximation results for partition functions $Z_G^H(\lambda)$ we first introduce the cluster expansion formally.

For a multiset Γ of polymers, the *incompatibility graph*, I_Γ, is the graph with vertex set Γ and an edge between any two incompatible polymers. A *cluster* Γ is a finite *ordered* multiset of polymers so that I_Γ is connected. We let $\mathcal{C}$ denote the set of all clusters.

For a cluster $\Gamma \in \mathcal{C}$ and pattern $(A, B) \in \mathcal{D}_\lambda(H)$, we let

$$w_{A,B}(\Gamma) := \phi(I_\Gamma) \prod_{\gamma \in \Gamma} w_{A,B}(\gamma)\,, \tag{1.7}$$

where $\phi(F)$ is the *Ursell function* of a graph F, defined by

$$\phi(F) := \frac{1}{|V(F)|!} \sum_{\substack{E \subseteq E(F) \\ \text{spanning, connected}}} (-1)^{|E|}\,. \tag{1.8}$$

The *cluster expansion* of the partition function $\Xi_{A,B}$ (as defined in (1.6)) is the formal power series

$$\ln \Xi_{A,B} = \sum_{\Gamma \in \mathcal{C}} w_{A,B}(\Gamma)\,. \tag{1.9}$$

We refer the reader to Section 2 for a more detailed introduction to the cluster expansion.

As mentioned above, establishing the convergence of the expansion in (1.9) is a non-trivial task and requires a careful combination of methods. We use the graph container method of Korshunov and Sapozhenko [39,52] to replace polymers with 'approximate polymers' that are less numerous than the polymers themselves. We then use entropy tools and isoperimetric properties of the torus to bound both the total number and total weight of polymers that have a fixed approximation. Crucially, we make an extra saving in the count on polymers by exploiting an algebraically constructed vertex partition of $\mathbb{Z}_m^n$ with good 'covering properties' (Lemma 8.1).

Equipped with the convergence of the cluster expansion we obtain the following approximation result obtained by truncating the cluster expansions of $\Xi_{A,B}$. The *size* of a cluster Γ is $\|\Gamma\| := \sum_{\gamma \in \Gamma} |\gamma|$. For $k \geq 1$ we define

$$L_{A,B}(k) := \sum_{\substack{\Gamma \in \mathcal{C}: \\ \|\Gamma\| = k}} w_{A,B}(\Gamma)\,. \tag{1.10}$$

Theorem 1.4. *For m an even integer, $k \in \mathbb{N}$ and (H, λ) a weighted graph, we have*

$$Z_G^H(\lambda) = \eta_\lambda(H)^{\frac{m^n}{2}} \sum_{(A,B) \in \mathcal{D}_\lambda(H)} \exp\left\{ \sum_{j=1}^{k-1} L_{A,B}(j) + \varepsilon_k \right\}$$

where $\varepsilon_k = O(m^n n^{2(k-1)}\delta^{nk})$ *for some* $\delta \in [0,1)$ *depending only on* (H,λ). *In particular there exists a constant* K *depending only on* (H,λ) *and* m *such that* $\varepsilon_K = o(1)$.

We emphasise that $\varepsilon_K = o(1)$ for some constant K in Theorem 1.4 and so to compute an explicit asymptotic formula (correct up to a $(1+o_n(1))$ factor) for $Z_G^H(\lambda)$ one only needs to compute a constant number of the terms $L_{A,B}(j)$. In Section 16 (see Theorem 16.1) we show that for fixed (H,λ) and m, one can obtain an explicit expression for $L_{A,B}(j)$ as a function of n in time $e^{O(j\ln j)}$. In particular obtaining an explicit asymptotic formula for $Z_G^H(\lambda)$ is a finite task. We can of course go further and obtain yet more detailed approximations to $Z_G^H(\lambda)$ by computing more of the terms $L_{A,B}(j)$.

To give a concrete example of the type of expression that arises in Theorem 1.4, we note that for a dominant pattern $(A,B) \in \mathcal{D}_\lambda(H)$,

$$L_{A,B}(1) = \left[\frac{1}{\lambda_A(\lambda_B)^d}\sum_{v\in A^c}\lambda_v(\lambda_{N(v)\cap B})^d + \frac{1}{\lambda_B(\lambda_A)^d}\sum_{v\in B^c}\lambda_v(\lambda_{N(v)\cap A})^d\right]\frac{m^n}{2} \quad (1.11)$$

(where we recall that $d = (1+\mathbb{1}_{m>2})n$, the degree of a vertex in $\mathbb{Z}_m^n$).

Theorem 1.4 contains a variety of asymptotic formulae for combinatorial quantities such as the number of proper q-colourings and the number of independent sets in $\mathbb{Z}_m^n$. As a special case we resolve a conjecture of Kahn and Park [38, Conjecture 5.2] on the number of proper q-colourings of the hypercube Q_n for $q \in \{5,6\}$ (see Corollary 1.6).

Specialising to the case where $m=2$, $H=K_q$ and $\lambda \equiv 1$, we obtain an estimate for $c_q(Q_n) \coloneqq |\mathrm{Hom}(Q_n,K_q)|$, the number of proper q-colourings of Q_n.

Theorem 1.5. *For* $q \geq 3$,

$$c_q(Q_n) = (1+\mathbb{1}_{\{q\ odd\}})\left\lfloor\frac{q}{2}\right\rfloor^{2^{n-1}}\left\lceil\frac{q}{2}\right\rceil^{2^{n-1}}\binom{q}{\lfloor q/2\rfloor}\cdot\exp\{f(n)+o(f(n))\}\ ,$$

as $n\to\infty$ *where*

$$f(n) = \frac{\lceil q/2\rceil}{2\lfloor q/2\rfloor}\left(2-\frac{2}{\lceil q/2\rceil}\right)^n + \frac{\lfloor q/2\rfloor}{2\lceil q/2\rceil}\left(2-\frac{2}{\lfloor q/2\rfloor}\right)^n .$$

In Section 17 we will strengthen Theorem 1.5 considerably, providing a detailed picture of the approximations to $c_q(\mathbb{Z}_m^n)$ contained in Theorem 1.4.

Galvin [20] proved the $q=3$ case of Theorem 1.5 showing that $c_3(Q_n) \sim 6e2^{2^{n-1}}$ and recently Kahn and Park [38] established the case $q=4$: $c_4(Q_n) \sim 6e2^{2^n}$. Kahn and Park also tentatively conjecture that the error term, $o(f(n))$, in Theorem 1.5 (a conjecture at the time) is negative. We will show that this is in fact not the case (see Theorem 17.4), this error term is positive for all q.

The expression in Theorem 1.5 was obtained by calculating $L_{A,B}(1)$ for a dominant pattern $(A,B)\in\mathcal{D}_\lambda(H)$ where $H=K_q, \lambda\equiv 1$. Theorem 1.5 does not give an asymptotic formula for $c_q(Q_n)$ (correct up to a $(1+o_n(1))$ factor) except in the cases $q=3$

and $q = 4$. However, there is no reason for us to stop at a calculation of $L_{A,B}(1)$. We collect some asymptotic formulae that arise from calculating more terms in the expansion of Theorem 1.4. The following are all easily obtained by hand and the reader may enjoy deriving some further examples of their own. We let $i(G)$ denote the number of independent sets in the graph G.

Corollary 1.6.

$$c_5(Q_n) \sim 20\sqrt[3]{e} \cdot 6^{2^{n-1}} \exp\left\{\left(\frac{4}{3}\right)^n\right\}$$

$$c_6(Q_n) \sim 20 \cdot 3^{2^n} \exp\left\{\left(\frac{4}{3}\right)^n\right\}$$

$$c_7(Q_n) \sim 70 \cdot e^{n/2} 12^{2^{n-1}} \exp\left\{\left(\frac{3}{2}\right)^{n-1} + \frac{1}{2}\left(\frac{4}{3}\right)^{n-1} + \frac{n^2-n-54}{108}\left(\frac{9}{8}\right)^{n-1}\right\}$$

$$c_8(Q_n) \sim 70 \cdot 4^{2^n} \exp\left\{\left(\frac{3}{2}\right)^n + \frac{n^2+41n-54}{108}\left(\frac{9}{8}\right)^n\right\}$$

$$i(\mathbb{Z}_m^n) \sim 2^{m^n/2+1} \exp\left\{\frac{m^n}{2^{d+1}}\right\} \text{ for } m = 2, 4, 6, 8, 10, 12, 14$$

$$i(\mathbb{Z}_m^n) \sim 2^{m^n/2+1} \exp\left\{\frac{1}{2}\left(\frac{m}{4}\right)^n + 2n(2n-1)\left(\frac{m}{16}\right)^n\right\} \text{ for } m = 16, 18, \ldots, 62.$$

The first two expressions of Corollary 1.6 resolve a conjecture of Kahn and Park [38, Conjecture 5.2]. We also remark that the case $m = 2$ in the fifth expression of Corollary 1.6 is the asymptotic $i(Q_n) \sim 2\sqrt{e} \cdot 2^{2^{n-1}}$, a classical result of Korshunov and Sapozhenko [39]. See also [23], [34] and [2] for further results on independent sets and the hard core model in the hypercube.

The expressions for $c_q(Q_n)$ in Corollary 1.6 were obtained by calculating $L_{A,B}(1)$, $L_{A,B}(2)$ for a dominant pattern (A, B). In Section 17 we show that more generally, in order to obtain an asymptotic formula for $c_q(\mathbb{Z}_m^n)$, one necessarily has to compute the terms $L_{A,B}(1), \ldots, L_{A,B}(k-1)$ where k is the least integer for which $m(1 - 1/\lceil q/2 \rceil)^{(1+\mathbb{1}_{m>2})k} < 1$.

1.5. The defect distribution

Theorem 1.3 shows that with probability $1 - e^{-\Omega(m^n/n^3)}$, a sample from $\mu_{H,\lambda}$ agrees with a dominant colouring on all but a $O(1/n)$ fraction of the vertices of $\mathbb{Z}_m^n$. Recall that we call the set of vertices at which an element $f \in \mathrm{Hom}(\mathbb{Z}_m^n, H)$ differs from its closest dominant colouring (breaking ties arbitrarily if necessary) the *defect vertices* of f. It is natural to ask for the distribution of the number of defect vertices in a sample from $\mu_{H,\lambda}$. By Theorem 1.2 this amounts to understanding the distribution of the number of vertices of $\mathbb{Z}_m^n$ contained in a sample from the polymer measures $\nu_{A,B}$.

We say that an event $\mathcal{A}_n$ occurs *with high probability* (whp) if $\mathbb{P}(\mathcal{A}_n) \to 1$ as $n \to \infty$. We use '$\xrightarrow{d}$' to denote convergence in distribution as $n \to \infty$. We let $\mathrm{Pois}(\rho)$ denote the Poisson distribution with mean ρ and let $N(0,1)$ denote the standard Normal distribution with mean 0 and variance 1.

We have the following central limit theorem for the measures $\nu_{A,B}$.

Theorem 1.7. *Fix $m \geq 2$ even, (H,λ) a weighted graph and $(A,B) \in \mathcal{D}_\lambda(H)$. Let $\mathbf{\Gamma}$ be a random configuration drawn from the distribution $\nu_{A,B}$. Letting $L := L_{A,B}(1)$ and $\|\mathbf{\Gamma}\| := \sum_{\gamma \in \mathbf{\Gamma}} |\gamma|$,*

1. *if $L \to 0$ as $n \to \infty$, then $\|\mathbf{\Gamma}\| = 0$ whp,*
2. *if $L \to \rho$ for some constant $\rho > 0$, then $\|\mathbf{\Gamma}\| \xrightarrow{d} Pois(\rho)$,*
3. *if $L \to \infty$, then*

$$\frac{\|\mathbf{\Gamma}\| - L}{\sqrt{L}} \xrightarrow{d} N(0,1)\,.$$

We record the following corollary of Theorem 1.7 in the case of the q-colouring model on $\mathbb{Z}_m^n$.

Corollary 1.8. *Let X denote the number of defect vertices in a uniformly chosen q-colouring of $\mathbb{Z}_m^n$.*

- *If $m \in \{2,4\}$ and $q \in \{3,4\}$ then $X \xrightarrow{d} Pois(1)$,*
- *otherwise $\frac{X - f(n)}{\sqrt{f(n)}} \xrightarrow{d} N(0,1)$, where*

$$f(n) = \frac{m^n}{2}\left[\frac{\lceil q/2\rceil}{\lfloor q/2\rfloor}\left(1 - \frac{1}{\lceil q/2\rceil}\right)^d + \frac{\lfloor q/2\rfloor}{\lceil q/2\rceil}\left(1 - \frac{1}{\lfloor q/2\rfloor}\right)^d\right].$$

Specialising to the case $m = 2$, Corollary 1.8 shows that q-colourings of Q_n exhibit a sharp change in behaviour from $q = 4$ to $q = 5$ where a typical q-colouring goes from having a constant number of defect vertices to an exponential number. This transition indicates why it is more difficult to enumerate q-colourings of Q_n when $q \geq 5$.

1.6. Generalised rank functions and height functions

A *rank function* on the Boolean lattice $2^{[n]}$ is a function $f : 2^{[n]} \to \mathbb{N}$ such that $f(\emptyset) = 0$ and

$$0 \leq f(A \cup \{x\}) - f(A) \leq 1$$

for all $A \subseteq [n]$ and $x \in [n] \backslash A$. In an early application of entropy methods in combinatorics, Kahn and Lawrenz [37] showed that the number of rank functions on $2^{[n]}$ is at

most $2^{(1+o(1))2^{n-1}}$, thus resolving a conjecture of Athanasiadias [1]. The authors in fact consider a natural generalisation of the notion of rank function: a function $f : 2^{[n]} \to \mathbb{N}$ is *k-bounded*, if $f(\emptyset) = 0$ and

$$0 \leq f(A \cup \{x\}) - f(A) \leq k$$

for all $A \subseteq [n]$ and $x \in [n] \backslash A$.[1] Let $\mathcal{B}_k(n)$ denote the set of all k-bounded functions. Kahn and Lawrenz [37] show that

$$\ln |\mathcal{B}_k(n)| \leq (1 + O(n^{-1/2}))2^{n-1} \ln \left(\left\lfloor \frac{k}{2} + 1 \right\rfloor \left\lceil \frac{k}{2} + 1 \right\rceil \right)$$

and conjectured that in fact $\ln |\mathcal{B}_k(n)| \leq 2^{n-1} \log \left(\left\lfloor \frac{k}{2} + 1 \right\rfloor \left\lceil \frac{k}{2} + 1 \right\rceil \right) + O(1)$, [37, Conjecture 1]. We disprove this conjecture, providing an asymptotic expression for $|\mathcal{B}_k(n)|$ which shows that the conjecture holds if and only if $k = 1$ or 2.

Theorem 1.9. *For fixed $k \geq 1$,*

$$|\mathcal{B}_k(n)| = (1 + \mathbb{1}_{k \text{ odd}}) \left(\left\lfloor \frac{k}{2} + 1 \right\rfloor \left\lceil \frac{k}{2} + 1 \right\rceil \right)^{2^{n-1}} \exp \left\{ \frac{(1 + \mathbb{1}_{k \text{ even}})}{\lfloor k/2 + 1 \rfloor} \left(\frac{2\lceil k/2 \rceil}{\lceil k/2 + 1 \rceil} \right)^n (1 + o(1)) \right\}.$$

For the proof we construct a bijection between the set $\mathcal{B}_k(n)$ to a set of 'rooted' homomorphisms in $\mathrm{Hom}(Q_n, H)$ for an appropriately chosen Cayley graph H. We then apply Theorem 1.4 (in particular more precise asymptotics than those recorded in Theorem 1.9 are available to us). In the special case of rank functions, such a bijection was already well-known: as observed by Mossel (see [36]), there is a bijection between the set of rank functions $\mathcal{B}_1(n)$ and the set of *height functions* (also known as 'cube-indexed random walks')

$$\mathcal{F}(n) \coloneqq \{f : V(Q_n) \to \mathbb{Z} : f(\mathbf{0}) = 0 \text{ and } u \sim v \implies |f(u) - f(v)| = 1\}$$

(here $\mathbf{0}$ denotes the vertex $(0, \ldots, 0)$ of Q_n). Randall (see [20]) observed that there is bijection between the set $\mathcal{F}(n)$ and the set of proper 3-colourings of Q_n which assign the vertex $\mathbf{0}$ a fixed colour.

For $f \in \mathcal{F}(n)$, let

$$R(f) = |\{f(x) : x \in V(G)\}|,$$

[1] For notational convenience, we have modified the definition in [37] where $f : 2^{[n]} \to \mathbb{N}$ is said to be k-bounded if $f(\emptyset) = 0$ and $0 \leq f(A \cup \{x\}) - f(A) < k$ for all $A \subseteq [n]$ and $x \in [n] \backslash A$.

the size of the range of f. In Section 18 we sketch how a large deviation inequality for our polymer models (Theorem 11.1, a key ingredient in the proofs of Theorems 1.2 and 1.3) can be used to recover the following special case of a result due to Peled [44].

Henceforth we let log denote the base 2 logarithm and let ln denote the natural logarithm. The *binary entropy function* is the map $H : [0,1] \to \mathbb{R}$ given by

$$H(p) = -p \log p - (1-p) \log(1-p)$$

where we interpret $0 \log 0$ as 0.

Theorem 1.10. *If $t \in (0,1]$, and f is chosen uniformly at random from $\mathcal{F}(n)$, then*

$$\log \log \left(1/\mathbb{P}(R(f) \geq tn)\right) \sim H(t/2)n\,.$$

The upper bound on $\mathbb{P}(R(f) \geq tn)$ in Theorem 1.10 is a special case of [44, Theorem 2.1] which applies to a general class of tori including $\mathbb{Z}_m^n$ (where m is allowed to be large with respect to n) and provides strong bounds on $\mathbb{P}(R(f) \geq k)$ for arbitrary k. The lower bound in Theorem 1.10 seems to be new, though it is not difficult and follows from a simple construction.

Another well-studied class of functions in this context are Lipschitz functions (see for example [44,47,46]). We note that the tools of this section extend naturally to the analysis of Lipschitz functions on Q_n (and more generally $\mathbb{Z}_m^n$ for m even) but we do not pursue the details here.

For more background, we refer the reader to Section 18.2.

1.7. Torpid mixing

In this section we discuss an algorithmic implication of Theorem 1.2. We establish torpid mixing for a natural class of Markov chain algorithms (including Glauber dynamics) designed to sample from the space $\mathrm{Hom}(\mathbb{Z}_m^n, H)$.

The main tool we use for deriving algorithmic results from Theorem 1.2 is the notion of *conductance* of a Markov chain. Conductance was first introduced to the field of Markov chain algorithms by Jerrum and Sinclair [54].

Let $\mathcal{M}$ be an ergodic (i.e. connected and aperiodic) Markov chain on a finite state space Ω, with transition probabilities $P(\omega, \omega')$, $\omega, \omega' \in \Omega$. Let π denote the stationary distribution of $\mathcal{M}$. For $\omega_0 \in \Omega$, we denote by $P_{t,\omega_0}(\omega)$ the probability that the system is in the state ω at time t given that ω_0 is the initial state. The mixing time of the Markov chain $\mathcal{M}$ is defined as

$$\tau_{\mathcal{M}} := \min \left\{ t_0 : \max_{\omega_0 \in \Omega} \|P_{t,\omega_0} - \pi\|_{TV} \leq \frac{1}{e}\,,\ \forall t > t_0 \right\}$$

where $\|\cdot\|_{TV}$ denotes total variation distance.

The Markov chains we consider will have state space $\Omega \subseteq \mathrm{Hom}(\mathbb{Z}_m^n, H)$ for some fixed graph H. We call a Markov chain *β-local* if f_1, f_2 differ at $\leq \beta m^n$ vertices whenever $P(f_1, f_2) \neq 0$. In other words, the chain updates the colours of $\leq \beta m^n$ vertices at each step. We only consider Markov chains that are ergodic.

If H is connected and bipartite, with vertex partition $V(H) = V^+ \cup V^-$, then any $f \in \mathrm{Hom}(\mathbb{Z}_m^n, H)$ satisfies either (i) $f(\mathcal{O}) \subseteq V^+, f(\mathcal{E}) \subseteq V^-$ or (ii) $f(\mathcal{E}) \subseteq V^+, f(\mathcal{O}) \subseteq V^-$. In particular, if $\beta < 1$, a β-local Markov chain on $\mathrm{Hom}(\mathbb{Z}_m^n, H)$ cannot be connected. In this case, we let $\mathrm{Hom}^+(\mathbb{Z}_m^n, H)$ denote the set of f satisfying (i) and let $\mathrm{Hom}^-(\mathbb{Z}_m^n, H)$ denote the set of f satisfying (ii). Further, we let $\mu_{H,\lambda}^+$ denote the measure $\mu_{H,\lambda}$ conditioned on $\mathrm{Hom}^+(\mathbb{Z}_m^n, H)$ and define $\mu_{H,\lambda}^-$ similarly.

Finally, let us call a non-bipartite weighted graph (H, λ) *trivial* if (H, λ) has only one dominant pattern. We say that (H, λ) is non-trivial otherwise. We call a bipartite (H, λ) trivial if it only has two dominant patterns. Note that any bipartite (H, λ) has at least two dominant patterns since if (A, B) is dominant, then (B, A) is a distinct dominant pattern.

Theorem 1.11. *For (H, λ) a fixed non-bipartite, non-trivial, weighted graph, there exists $\beta > 0$ such that for $m \geq 2$ even, any β-local ergodic Markov chain $\mathcal{M}$ on $\mathrm{Hom}(\mathbb{Z}_m^n, H)$ with stationary distribution $\mu_{H,\lambda}$ satisfies*

$$\tau_{\mathcal{M}} = e^{\Omega(m^n/n^2)}.$$

For (H, λ) bipartite, the analogous statement holds with $\mathrm{Hom}(\mathbb{Z}_m^n, H)$ replaced with $\mathrm{Hom}^{\pm}(\mathbb{Z}_m^n, H)$ and $\mu_{H,\lambda}$ replaced with $\mu_{H,\lambda}^{\pm}$.

In the case of the hard-core model, Theorem 1.11 follows from arguments of Galvin and Tetali [28] who establish torpid mixing for the Glauber dynamics of the hard-core model on a class of bipartite expander graphs. In the regime where n is fixed and m is large, analogues of Theorem 1.11 have been proved for specific choices of (H, λ) [8,22,26]. We return to this topic in the concluding remarks.

1.8. Outline

The outline for the remainder of this paper is as follows. In Section 2 we introduce abstract polymer models and the cluster expansion and in Section 3 we introduce the concrete polymer models that we work with in this paper and explore some of their basic properties.

In Section 4, we give an overview of the proof of our main theorem, Theorem 1.2 and show how it follows from two key lemmas (Lemmas 4.2 and 4.3). We also show how the asymptotic approximations to $Z_G^H(\lambda)$ in Theorem 1.4 follow from these key lemmas and the convergence of the cluster expansion.

In Section 5 we collect some standard tools based on entropy.

In Sections 6 and 7 we establish some isoperimetric results in the discrete torus and introduce some notions from the theory of graph containers.

Sections 8, 9 and 10 are then dedicated to proving the convergence of the cluster expansion of our polymer models.

In Section 11 we show how the convergence of the cluster expansion can be used to prove a large deviation result for the polymer measures $\nu_{A,B}$. In Section 12, we show how this large deviation result and an entropy argument can be used to establish the key lemmas, Lemmas 4.2 and 4.3, thus concluding the proof of our main theorem, Theorem 1.2 and Theorem 1.4.

In Section 13 we show how Theorem 1.2 can be used to give a precise structural description of the set $\mathrm{Hom}(\mathbb{Z}_m^n, H)$ thus resolving conjectures of Engbers and Galvin [14, Conjectures 6.1, 6.2, 6.3].

In Section 14 we prove the torpid mixing result, Theorem 1.11.

In Section 15 we prove Theorem 1.7 and Corollary 1.8 on the asymptotic distribution of the number of defect vertices.

In Section 16 we provide an algorithm for computing the terms of the cluster expansion. In Section 17 we specialise the discussion to the q-colouring model and provide a detailed picture of the cluster expansion in this setting.

In Section 18 we prove Theorem 1.9 and then Theorem 1.10 on k-bounded functions and height functions.

In Section 19 we end with some concluding remarks and directions of future research.

1.9. Notation and terminology

Here we gather some notation for ease of reference.

For $k \in \mathbb{N}$, we let $[k]$ denote the set $\{1, \dots, k\}$. For a set X, we let 2^X denote the power set of X and for $k \in \mathbb{N}$ let $\binom{X}{k}$ denote the collection of subsets of X of size k.

Given a graph $G = (V, E)$ and $x \in V$, we let $N(x) = \{y \in V : x \sim y\}$ and for $X \subseteq V$, we let $N(X) = \bigcup_{x \in X} N(x)$. For $X \subseteq V$ we let $\partial(X) = N(X) \backslash X$, the *vertex boundary* of X. We let $X^+ = X \cup \partial(X)$. For $v \in V$, we let $d(v) = |N(v)|$ and for a subset $Y \subseteq V$ we let $d_Y(v) \coloneqq |N(v) \cap Y|$. For two vertices $u, v \in V$ we write $d_G(u, v)$ for the graph distance between u, v i.e. the length of the shortest path in G with endpoints u, v. For a set $S \subseteq V$, we let $G[S]$ denote the graph with vertex set S and edge set $E \cap \binom{S}{2}$. For $k \in \mathbb{N}$ we let G^k denote the kth power of G, that is, the graph with vertex set $V(G)$ such that $u \sim v$ if and only if $d_G(u, v) \leq k$. We say a set $S \subseteq V$ is G^k-*connected* if $G^k[S]$ is connected.

Throughout this paper we use log to denote the base 2 logarithm and we use ln to denote the natural logarithm.

Henceforth we can assume that we have fixed a weighted graph (H, λ) and an even integer $m \geq 2$. Standard asymptotic notation such as O, Ω, o, Θ, is used under the assumption $n \to \infty$ and the implicit constants may depend on (H, λ) and m.

2. Abstract polymer models and the cluster expansion

In this section we describe two classic tools from statistical physics, abstract polymer models [40] and the cluster expansion. Both tools have been used extensively to study phase diagrams of lattice spin models. Following the breakthough work of Helmuth, Perkins and Regts [32], these tools (along with contour models from Pirogov-Sinai Theory) have found a wealth of applications in an algorithmic setting: namely the design of efficient approximate counting and sampling algorithms for spin models on graphs [6,9,19,31–33,42].

We have already encountered the terms 'polymer' and 'cluster' in the previous section. Indeed, the polymers from the introduction are concrete examples of a more general notion which we introduce now.

Let $\mathcal{P}$ be a finite set whose elements we call 'polymers'. We equip $\mathcal{P}$ with a complex-valued weight $w(\gamma)$ for each polymer $\gamma \in \mathcal{P}$ as well as a symmetric and reflexive incompatibility relation between polymers. We write $\gamma \nsim \gamma'$ if polymers γ and γ' are incompatible. Let Ω be the collection of pairwise compatible sets of polymers from $\mathcal{P}$, including the empty set of polymers. Then the polymer model partition function is

$$\Xi(\mathcal{P}) = \sum_{\Gamma \in \Omega} \prod_{\gamma \in \Gamma} w(\gamma)\,,$$

where the contribution from the empty set is 1.

For a multiset of polymers Γ, we define the *incompatibility graph* I_Γ to be the graph on vertex set Γ where γ_1 is adjacent to γ_2 if and only if γ_1, γ_2 are incompatible. A *cluster* is an *ordered* multiset of polymers whose incompatibility graph is connected. Let $\mathcal{C}$ be the set of all clusters. The *cluster expansion* is the formal power series in the weights $w(\gamma)$

$$\ln \Xi(\mathcal{P}) = \sum_{\Gamma \in \mathcal{C}} w(\Gamma)\,,$$

where

$$w(\Gamma) = \phi(I_\Gamma) \prod_{\gamma \in \Gamma} w(\gamma)\,,$$

and ϕ is the Ursell function as defined in (1.8). In fact the cluster expansion is simply the multivariate Taylor series for $\ln \Xi(\mathcal{P})$ in the variables $w(\gamma)$, as observed by Dobrushin [12]. See also Scott and Sokal [53] for a derivation of the cluster expansion and much more.

A sufficient condition for the convergence of the cluster expansion is given by a theorem of Kotecký and Preiss.

Theorem 2.1 *(Kotecký and Preiss [40]). Let $f : \mathcal{P} \to [0,\infty)$ and $g : \mathcal{P} \to [0,\infty)$ be two functions. Suppose that for all polymers $\gamma \in \mathcal{P}$,*

$$\sum_{\gamma' \nsim \gamma} |w(\gamma')| e^{f(\gamma')+g(\gamma')} \leq f(\gamma)\,, \tag{2.1}$$

then the cluster expansion converges absolutely. Moreover, if we let $g(\Gamma) = \sum_{\gamma\in\Gamma} g(\gamma)$ and write $\Gamma \nsim \gamma$ if there exists $\gamma' \in \Gamma$ so that $\gamma \nsim \gamma'$, then for all polymers γ,

$$\sum_{\substack{\Gamma\in\mathcal{C}\\ \Gamma\nsim\gamma}} |w(\Gamma)|\, e^{g(\Gamma)} \leq f(\gamma)\,. \tag{2.2}$$

As a preview of one of the applications of the above theorem, we remark that (2.2) can be used to give tail bounds on the cluster expansion. This will allow us to show that certain truncations of the cluster expansion serve as good approximations to the logarithm of the partition function.

3. Concrete polymer models

In this section we formally introduce the concrete polymer models that will be the subject of study in this paper and explore some of their properties. These polymer models are a generalisation of those used by the current authors and Perkins [33] to design efficient approximate counting and sampling algorithms for the hard-core model and q-colouring model on bipartite expander graphs.

We repeat some of the notation and definitions from the introduction for the reader's convenience. Throughout this paper we let G denote the n-dimensional discrete torus $\mathbb{Z}_m^n$, the graph on vertex set $\{0,\dots,m-1\}^n$ where two vertices x, y are adjacent if and only if there exists a coordinate $i \in [n]$ such that $x_i = y_i \pm 1 \pmod m$ and $x_j = y_j$ for all $j \neq i$. We consider only the case where $m \geq 2$ is even, in which case G is a bipartite graph with vertex classes

$$\mathcal{E} := \left\{ x : \sum_{i=1}^n x_i \equiv 0 \pmod 2 \right\} \text{ and } \mathcal{O} := \left\{ x : \sum_{i=1}^n x_i \equiv 1 \pmod 2 \right\}.$$

Note that we have $|\mathcal{O}| = |\mathcal{E}| = m^n/2$. We also let $V = V(G) = \mathcal{O} \cup \mathcal{E}$. When $m = 2$, the graph $\mathbb{Z}_m^n$ is the familiar n-dimensional hypercube and we use the more standard notation Q_n in this case. We note that Q_n is an n-regular graph, whereas $\mathbb{Z}_m^n$ is a $2n$-regular graph when $m > 2$. To avoid having to reiterate this distinction, we will simply say that $\mathbb{Z}_m^n$ is d-regular, where $d = (1 + \mathbb{1}_{m>2})n$.

Recall that given a subset $S \subseteq V$, a colouring f and a pattern (A, B), we say that f *agrees* with (A, B) at $v \in V$ if $v \in \mathcal{O}$ and $f(v) \in A$ or if $v \in \mathcal{E}$ and $f(v) \in B$. We say that f disagrees with (A, B) at v otherwise. Let $\chi_{A,B}(S)$ be the set of colourings f such

that f *disagrees* with (A,B) at each $v \in S$ and agrees with (A,B) at each $v \in V \backslash S$. For each pattern (A,B) we have a weight function $w_{A,B} : 2^V \to \mathbb{R}$ where

$$w_{A,B}(S) = \frac{\sum_{f \in \chi_{A,B}(S)} \prod_{v \in V} \lambda_{f(v)}}{\eta_\lambda(H)^{m^n/2}}. \tag{3.1}$$

For a colouring $f \in \mathrm{Hom}(G,H)$ and subset $S \subseteq V$, we let f_S denote the restriction of f to S. A more convenient (from a computational perspective) form for the weight function $w_{A,B}$ is the following: let $\hat{\chi}_{A,B}(S)$ be the set of colourings f_{S^+} where $f \in \chi_{A,B}(S)$. Given any $h \in \hat{\chi}_{A,B}(S)$, we may extend h to an element of $\chi_{A,B}(S)$ by arbitrarily assigning vertices of $\mathcal{O} \backslash S^+$ colours from A and arbitrarily assigning vertices of $\mathcal{E} \backslash S^+$ colours from B. Thus

$$w_{A,B}(S) = \frac{\sum_{f \in \hat{\chi}_{A,B}(S)} \prod_{v \in S^+} \lambda_{f(v)}}{\lambda_A^{|S^+ \cap \mathcal{O}|} \lambda_B^{|S^+ \cap \mathcal{E}|}}. \tag{3.2}$$

To illustrate the point, let us calculate $w_{A,B}(S)$ in the case where $S = \{v\}$ for some $v \in \mathcal{O}$. In this case $\hat{\chi}_{A,B}(S)$ is the set of homomorphisms $f : G[v \cup N(v)] \to H$ such that $f(v) \in A^c$ and $f(N(v)) \subseteq B$. Using (3.2), we have

$$w_{A,B}(S) = \frac{1}{\lambda_A \lambda_B^d} \sum_{v \in A^c} \lambda_v \lambda_{N(v) \cap B}^d .$$

We now define our collection of polymer models and partition functions. Recall that we say that $S \subseteq V(G)$ is G^2-*connected* if the graph $G^2[S]$ is connected (here G^2 denotes the square of the graph G).

Definition 3.1. We say a subset $\gamma \subseteq V$ is a *polymer* if it is G^2-connected and $|N(\gamma \cap \mathcal{E})|, |N(\gamma \cap \mathcal{O}))| < (1-\alpha)m^n/2$ where $0 < \alpha < 1$ is some constant (dependent only on (H,λ)) which will be specified later (see (12.2)).

We say that two polymers γ_1, γ_2 are *compatible*, denoted $\gamma_1 \sim \gamma_2$, if $d_G(\gamma_1, \gamma_2) > 2$ (i.e. $\gamma_1 \cup \gamma_2$ is not G^2-connected).

We let $\mathcal{P}$ denote the set of all polymers in G and let Ω denote the family of all sets of mutually compatible polymers from $\mathcal{P}$.

Definition 3.2. For each dominant pattern $(A,B) \in \mathcal{D}_\lambda(H)$, the (A,B) *polymer model* is the polymer model with polymer set $\mathcal{P}$, compatibility relation $\sim$ and weight function $w_{A,B}$.

The partition function of the (A,B) polymer model is

$$\Xi_{A,B} = \sum_{\Gamma \in \Omega} \prod_{\gamma \in \Gamma} w_{A,B}(\gamma). \tag{3.3}$$

We thus have $|\mathcal{D}_\lambda(H)|$ distinct polymer models determined by the choice of weight function $w_{A,B}$ where $(A,B) \in \mathcal{D}_\lambda(H)$ (note that the set of polymers, $\mathcal{P}$, is the same for each polymer model).

A useful property of the weight function is that it is multiplicative over G^2-connected components of a set S.

Lemma 3.3. *Let $(A,B) \in \mathcal{D}_\lambda(H)$. For $S \subseteq V$ with G^2-connected components $\gamma_1, \dots, \gamma_k$ we have*

$$w_{A,B}(S) = \prod_{i=1}^{k} w_{A,B}(\gamma_i)\,.$$

Proof. Since the γ_i are the G^2-connected components of S, the sets γ_i^+ are mutually disjoint. Moreover there is a one-one correspondence between elements $f \in \hat{\chi}_{A,B}(S)$ and tuples $(f_1, \dots, f_k) \in \hat{\chi}_{A,B}(\gamma_1) \times \ldots \times \hat{\chi}_{A,B}(\gamma_k)$. Using the expression (3.2) for the weight function we have

$$\begin{aligned} w_{A,B}(S) &= \frac{\sum_{f \in \hat{\chi}_{A,B}(S)} \prod_{v \in S^+} \lambda_{f(v)}}{\lambda_A^{|S^+ \cap \mathcal{O}|} \lambda_B^{|S^+ \cap \mathcal{E}|}} = \prod_{i=1}^{k} \frac{\sum_{f \in \hat{\chi}_{A,B}(\gamma_i)} \prod_{v \in \gamma_i^+} \lambda_{f(v)}}{\lambda_A^{|\gamma_i^+ \cap \mathcal{O}|} \lambda_B^{|\gamma_i^+ \cap \mathcal{E}|}} \\ &= \prod_{i=1}^{k} w_{A,B}(\gamma_i)\,. \quad \square \end{aligned}$$

The partition function $\Xi_{A,B}$ from (3.3) is the normalising constant of a probability distribution $\nu_{A,B}$ on Ω defined by

$$\nu_{A,B}(\Gamma) = \frac{\prod_{\gamma \in \Gamma} w_{A,B}(\gamma)}{\Xi_{A,B}}\,. \tag{3.4}$$

These measures allow us to build a new probability measure $\hat{\mu}_{H,\lambda}$ on $\mathrm{Hom}(G,H)$.

Definition 3.4. For $f \in \mathrm{Hom}(G,H)$, let $\hat{\mu}_{H,\lambda}(f)$ denote the probability that f is selected by the following four-step process:

1. Choose $(A,B) \in \mathcal{D}_\lambda(H)$ with probability proportional to $\Xi_{A,B}$.
2. Choose a random polymer configuration $\Gamma \in \Omega$ from the distribution $\nu_{A,B}$.
3. Letting $S = \bigcup_{\gamma \in \Gamma} \gamma$, select a colouring $f \in \hat{\chi}_{A,B}(S)$ with probability proportional to $\prod_{v \in S^+} \lambda_{f(v)}$.
4. Independently assign each $v \in \mathcal{O} \backslash S^+$ the colour $i \in A$ with probability λ_i/λ_A and each $v \in \mathcal{E} \backslash S^+$ the colour $j \in B$ with probability λ_j/λ_B.

We note that this definition is equivalent to Definition 1.1 from the introduction where we have expanded the last step into two steps. We do this to emphasise the independence in Step 4 which will be useful for later calculations.

We end this section by recording a couple of the basic properties of the measure $\hat{\mu}_{H,\lambda}$.

Definition 3.5. Let $(A,B) \in \mathcal{D}_\lambda(H)$. We say that a colouring $f \in \mathrm{Hom}(G,H)$ is *captured* by (A,B) if each of the G^2-connected components of $S(f) \coloneqq (f^{-1}(A^c) \cap \mathcal{O}) \cup (f^{-1}(B^c) \cap \mathcal{E})$ is a polymer.

In the following lemma, we let $\mathbf{f}$ denote a random colouring chosen according to $\hat{\mu}_{H,\lambda}$. Let $\mathbf{D}$ denote the random pattern selected at Step 1 in the definition of $\hat{\mu}_{H,\lambda}$ (Definition 3.4). Moreover, we let

$$\tilde{Z}_G^H(\lambda) \coloneqq \eta_\lambda(H)^{m^n/2} \sum_{(A,B)\in\mathcal{D}_\lambda(H)} \Xi_{A,B}\,. \tag{3.5}$$

Lemma 3.6. *Let $f \in \mathrm{Hom}(G,H)$.*

(i) If f is captured by precisely p polymer models then

$$\mathbb{P}(\mathbf{f}=f) = p\frac{\prod_{v\in V}\lambda_{f(v)}}{\tilde{Z}_G^H(\lambda)}\,.$$

(ii) For $(A,B)\in\mathcal{D}_\lambda(H)$,

$$\mathbb{P}(\mathbf{f}=f \mid \mathbf{D}=(A,B)) = \frac{\prod_{v\in V}\lambda_{f(v)}}{\eta_\lambda(H)^{m^n/2}\cdot\Xi_{A,B}} \cdot \mathbb{1}_{\{f \text{ captured by } (A,B)\}}\,.$$

Proof. We begin with the second claim. Fix $f\in\mathrm{Hom}(G,H)$, $(A,B)\in\mathcal{D}_\lambda(H)$ and let $S=S(f)$ as in Definition 3.5. Let the G^2-connected components of S be $\gamma_1,\dots,\gamma_k$ and let $\Gamma=\{\gamma_1,\dots,\gamma_k\}$. If f is captured by the (A,B) polymer model then

$$\begin{aligned}\mathbb{P}(\mathbf{f}=f \mid \mathbf{D}=(A,B)) &= \frac{\prod_{\gamma\in\Gamma} w_{A,B}(\gamma)}{\Xi_{A,B}} \cdot \frac{\prod_{v\in V}\lambda_{f(v)}}{\sum_{h\in\chi_{A,B}(S)}\prod_{v\in V}\lambda_{h(v)}} \\ &= \frac{\prod_{v\in V}\lambda_{f(v)}}{\eta_\lambda(H)^{m^n/2}\cdot\Xi_{A,B}}\end{aligned}$$

where for the final equality we used (3.2) and Lemma 3.3. If f is not captured by the (A,B) polymer model then one of the G^2-connected components of S is not a polymer and so the probability that Γ is selected at Step 2 in Definition 3.4 is 0. Therefore $\mathbb{P}(\mathbf{f}=f \mid \mathbf{D}=(A,B))=0$. The first claim follows by noting that

$$\mathbb{P}(\mathbf{D}=(A,B)) = \frac{\Xi_{A,B}}{\sum_{(C,D)\in\mathcal{D}_\lambda(H)}\Xi_{C,D}}$$

for any dominant pattern (A,B). □

4. Overview of the proof

In this section we give an overview of the proof of our main theorem, Theorem 1.2. We also show how the proof of Theorem 1.2 ties in with the proof of Theorem 1.4 which provides detailed asymptotic approximations to the partition function $Z_G^H(\lambda)$.

Theorem 1.2 asserts that $\hat{\mu}_{H,\lambda}$ is very close to the original spin measure $\mu_{H,\lambda}$ (defined in (1.1)) in total variation distance. On the way to proving Theorem 1.2 we will also show that the linear combination of partition functions $\Xi_{A,B}$ defined at (3.5) serves as a very good approximation to the spin model partition function $Z_G^H(\lambda)$ (defined in (1.2)).

Lemma 4.1. *We have*

$$1 - e^{-\Omega(m^n/n^2)} \leq \frac{\tilde{Z}_G^H(\lambda)}{Z_G^H(\lambda)} \leq 1 + e^{-\Omega(m^n/n^2)}\,.$$

In order to prove Lemma 4.1, we will show that almost every element of $\mathrm{Hom}(G,H)$ is captured (see Definition 3.5) by precisely one polymer model. This is also key for the proof of Theorem 1.2.

Let $\mathrm{Hom}_0(G,H)$, $\mathrm{Hom}_1(G,H)$, $\mathrm{Hom}_2(G,H)$ denote the sets of all colourings which are captured by 0, precisely one, and ≥ 2 dominant patterns respectively.

Lemma 4.2. *There exists $\zeta = \zeta(H,\lambda) < \eta_\lambda(H)$ so that*

$$Z_0 := \sum_{f \in \mathrm{Hom}_0(G,H)} \prod_{v \in V} \lambda_{f(v)} \leq \zeta^{m^n/2}\,.$$

Lemma 4.3.

$$Z_2 := \sum_{f \in \mathrm{Hom}_2(G,H)} \prod_{v \in V} \lambda_{f(v)} \leq e^{-\Omega(m^n/n^2)} \tilde{Z}_G^H(\lambda)\,.$$

We will prove Lemmas 4.2 and 4.3 in Section 12. We now show how both Lemma 4.1 and Theorem 1.2 follow from these two lemmas.

Proof of Lemma 4.1 assuming Lemmas 4.2 and 4.3. By Lemma 3.6 (i) we have

$$\tilde{Z}_G^H(\lambda) = \sum_{f \in \mathrm{Hom}(G,H)} p_f \prod_{v \in V} \lambda_{f(v)}\,,$$

where p_f denotes the number of polymer models that capture f. Comparing this to

$$Z_G^H(\lambda) = \sum_{f \in \mathrm{Hom}(G,H)} \prod_{v \in V} \lambda_{f(v)}$$

we obtain

$$\tilde{Z}_G^H(\lambda) + Z_0 - 4^q Z_2 \leq Z_G^H(\lambda) \leq \tilde{Z}_G^H(\lambda) + Z_0 - Z_2\,,$$

where 4^q is a bound on the number of dominant patterns. The lemma follows from Lemmas 4.2 and 4.3. □

Proof of Theorem 1.2 assuming Lemmas 4.2 and 4.3. Let $\mu, \hat{\mu}$ denote $\mu_{H,\lambda}, \hat{\mu}_{H,\lambda}$ respectively. Recalling a formula for the total variation distance between discrete probability measures we have

$$\|\mu - \hat{\mu}\|_{TV} = \sum_{\substack{f \in \mathrm{Hom}(G,H):\\ \hat{\mu}(f) > \mu(f)}} \hat{\mu}(f) - \mu(f)\,. \tag{4.1}$$

We consider the contribution to the above sum from elements of $\mathrm{Hom}_0(G,H)$, $\mathrm{Hom}_1(G,H)$ and $\mathrm{Hom}_2(G,H)$ separately. If $f \in \mathrm{Hom}_0(G,H)$ then $\hat{\mu}(f) = 0$ by Lemma 3.6 (i) and so elements of $\mathrm{Hom}_0(G,H)$ do not contribute to the sum in (4.1). Since $\tilde{Z}_G^H(\lambda)$ is within a factor $1 + e^{-\Omega(m^n/n^2)}$ of $Z_G^H(\lambda)$ by Lemma 4.1 (which holds by our assumption that Lemmas 4.2 and 4.3 hold) we have,

$$\begin{aligned}\sum_{\substack{f \in \mathrm{Hom}_1(G,H):\\ \hat{\mu}(f) > \mu(f)}} \hat{\mu}(f) - \mu(f) &= \sum_{\substack{f \in \mathrm{Hom}_1(G,H):\\ \hat{\mu}(f) > \mu(f)}} \left(\frac{\prod_{v \in V} \lambda_{f(v)}}{\tilde{Z}_G^H(\lambda)} - \frac{\prod_{v \in V} \lambda_{f(v)}}{Z_G^H(\lambda)} \right) \\ &\leq e^{-\Omega(m^n/n^2)} \end{aligned} \tag{4.2}$$

where for the first equality we used Lemma 3.6 (i). To bound the contribution from elements of $\mathrm{Hom}_2(G,H)$, we use Lemma 4.3 to obtain

$$\begin{aligned}\sum_{\substack{f \in \mathrm{Hom}_2(G,H):\\ \hat{\mu}(f) > \mu(f)}} \hat{\mu}(f) - \mu(f) &\leq \sum_{\substack{f \in \mathrm{Hom}_2(G,H):\\ \hat{\mu}(f) > \mu(f)}} \left(4^q \frac{\prod_{v \in V} \lambda_{f(v)}}{\tilde{Z}_G^H(\lambda)} - \frac{\prod_{v \in V} \lambda_{f(v)}}{Z_G^H(\lambda)} \right) \\ &\leq e^{-\Omega(m^n/n^2)} \end{aligned} \tag{4.3}$$

where again we use Lemma 3.6 (i) for the first inequality and the fact that elements of $\mathrm{Hom}_2(G,H)$ are captured by $\leq 4^q$ dominant patterns. The result follows by summing the bounds (4.2) and (4.3) to bound the sum in (4.1). □

In order to prove Theorem 1.2 it remains to prove Lemmas 4.2 and 4.3. A crucial step toward the proof of these lemmas is to verify that the Kotecký-Preiss condition (2.1) holds for each of the polymer models introduced in Definition 3.2.

Lemma 4.4. *There exist functions $f : \mathcal{P} \to [0,\infty)$ and $g : \mathcal{P} \to [0,\infty)$ such that for each $(A,B) \in \mathcal{D}_\lambda(H)$ and all polymers $\gamma \in \mathcal{P}$ we have*

$$\sum_{\gamma':d(\gamma',\gamma)\leq 2} w_{A,B}(\gamma')e^{f(\gamma')+g(\gamma')} \leq f(\gamma)\,. \tag{4.4}$$

In particular the cluster expansion of $\ln \Xi_{A,B}$ *converges absolutely.*

We will prove Lemma 4.4 in Section 10. A key consequence of Lemma 4.4 is that it provides strong tail bounds for the cluster expansion. In order to state our tail bound, we introduce a family of parameters that will play an important role throughout this paper. For a dominant pattern $(A,B) \in \mathcal{D}_\lambda(H)$ let

$$\delta_{A,B} := \max\left\{ \max_{v\in B^c} \frac{\lambda_{N(v)\cap A}}{\lambda_A}, \max_{u\in A^c} \frac{\lambda_{N(u)\cap B}}{\lambda_B} \right\}. \tag{4.5}$$

In Section 10 we prove the following.

Lemma 4.5. *For* $(A,B) \in \mathcal{D}_\lambda(H)$ *and* $k \in \mathbb{N}$*,*

$$\sum_{j=k}^{\infty} |L_{A,B}(j)| = O\left(m^n d^{2(k-1)} \delta_{A,B}^{dk}\right).$$

We note that $\delta_{A,B} < 1$ for each dominant pattern (A,B) and so Lemma 4.5 does indeed provide an effective tail bound.

Lemma 4.6. $\delta_{A,B} < 1$ *for each* $(A,B) \in \mathcal{D}_\lambda(H)$.

Proof. Let $(A,B) \in \mathcal{D}_\lambda(H)$. For any $u \in A^c$ there must exist $w \in B$ such that $u \nsim w$ else $A \cup \{u\} \sim B$ contradicts the fact that (A,B) is a dominant pattern. In other words, $N(u) \cap B \neq B$, so $\lambda_{N(u)\cap B} < \lambda_B$ for all $u \in A^c$. Similarly $\lambda_{N(v)\cap A} < \lambda_A$ for all $v \in B^c$. □

To end this section we show how the following, refined version of Theorem 1.4 follows from the above results.

Theorem 4.7. *For* m *an even integer,* $k \in \mathbb{N}$ *and* (H,λ) *a weighted graph, we have*

$$Z_G^H(\lambda) = \eta_\lambda(H)^{\frac{m^n}{2}} \sum_{(A,B)\in\mathcal{D}_\lambda(H)} \exp\left\{ \sum_{j=1}^{k-1} L_{A,B}(j) + \varepsilon_{k,A,B} \right\}$$

where $\varepsilon_{k,A,B} = O(m^n n^{2(k-1)} \delta_{A,B}^{dk})$.

Proof of Theorem 4.7 assuming Lemmas 4.2, 4.3and 4.5. Let $k \in \mathbb{N}$. By Lemma 4.5

$$\ln \Xi_{A,B} = \sum_{j=1}^{\infty} L_{A,B}(j) = \sum_{j=1}^{k-1} L_{A,B}(j) + \varepsilon_{k,A,B} \tag{4.6}$$

where $\varepsilon_{k,A,B} = O(m^n d^{2(k-1)} \delta_{A,B}^{dk})$. The result follows from Lemma 4.1 (which holds by our assumption that Lemmas 4.2 and 4.3 hold). □

In the next section we introduce the entropy tools that play a central role in the proofs of Lemmas 4.2 and 4.4. Pioneered by Kahn and Lawrenz [37], entropy methods have seen a huge range of applications in the study of graph homomorphisms and beyond (see for example [14,15,27,35,36,38,48,50]). The proof of Lemma 4.2 is an adaptation of a delicate entropy argument due to Engbers and Galvin [14]. Entropy arguments will also play a central role in the proof of Lemma 4.4. Roughly speaking, to prove Lemma 4.4 we need to balance two quantities: (*i*) the number of polymers of a given size and (*ii*) the total weight (or 'cost') of polymers of that size. In actuality, rather than fixing the size, a more complicated constraint will be used but we defer the details for now. A careful combination of entropy tools, the graph container method, isoperimetric and algebraic properties of the torus will ultimately allow us to balance (*i*) and (*ii*) and verify the Kotecký-Preiss condition.

Starting with the work of Korshunov and Sapozhenko [39], the graph container method has been widely applied in the study of spin models on the torus and integer lattice (see for example [17,20,22,23,21,24–26,28,38,39,44,45,48,50]). As noted in the introduction, the synthesis of container and entropy methods first appears in the work of Peled and Spinka [50] and later appears in the work of Kahn and Park [38] (with our treatment following the latter more closely).

5. Entropy tools

In this section we gather some tools based on entropy. In what follows all random variables will be discrete. Recall that we use log to denote the base 2 logarithm and that the *binary entropy function* is the map $H : [0,1] \to \mathbb{R}$ given by

$$H(p) = -p \log p - (1-p) \log(1-p)\,,$$

where we interpret $0 \log 0$ as 0. The *entropy* of a random variable X is

$$H(X) = \sum_x -\mathbb{P}(X = x) \log \mathbb{P}(X = x)\,.$$

The inequality that makes entropy a useful tool for counting is

$$H(X) \leq \log |\mathrm{range}(X)|\,, \tag{5.1}$$

and we have equality if and only if X has the uniform distribution. For two random variables X and Y, the *conditional entropy* of X given Y is

$$H(X|Y) = -\sum_y \mathbb{P}(Y = y) \sum_x \mathbb{P}(X = x \mid Y = y) \log \mathbb{P}(X = x \mid Y = y)\,.$$

We collect a few useful properties of entropy in the form of a lemma.

Lemma 5.1.

1. $H(X,Y) = H(X) + H(Y|X)$,
2. $H(X_1,\dots,X_n|Y) \le \sum_i H(X_i|Y)$,
3. *if Z is determined by Y, then* $H(X|Y) \le H(X|Z)$.

For a random vector $X = (X_1,\dots,X_k)$ and $S \subseteq [k]$, set $X_S = (X_i : i \in S)$. We also use the following version of Shearer's Lemma.

Lemma 5.2. *If $X = (X_1,\dots,X_k)$ is a random vector and $\varphi : 2^{[k]} \to \mathbb{R}_{\ge 0}$ satisfies*

$$\sum_{S \ni i} \varphi(S) \ge 1 \text{ for each } i \in [k]\,,$$

then for any partial order $\prec$ on $[k]$,

$$H(X) \le \sum_{S \subseteq [k]} \varphi(S) H(X_S | (X_i : i \prec S))\,,$$

where $i \prec S$ means $i \prec s$ for all $s \in S$.

Throughout the paper we will also use the following standard estimate on binomial coefficients in terms of the binary entropy function.

$$\binom{x}{\le \beta x} := \sum_{i=0}^{\lfloor \beta x \rfloor} \binom{x}{i} \le 2^{H(\beta)x} \quad \text{where } x \in \mathbb{N},\ \beta \in [0,1/2]\,.$$

We refer the reader to [11] for more background on information theory.

6. Isoperimetry in the torus

In this short section we collect some isoperimetric properties of the torus. These results will be key for bounding the weights of polymers.

Our first lemma gives a lower bound on the bipartite expansion factor of the torus. The following proof is an adaptation of an argument due to Christofides, Ellis and the second author [10, Theorem 3].

Lemma 6.1. *If $X \subseteq \mathcal{E}, \mathcal{O}$ with $|X| = \beta m^n/2$, where $0 < \beta \le 1$ then*

$$|N(X)| \ge |X| \left(1 + 4\sqrt{2} \cdot \frac{(1-\beta)}{m^2\sqrt{n}}\right)\,.$$

Proof. We proceed by induction on n and note that the case $n = 1$ is simple to verify (note that if $m = 2$ then $\beta > 0$ implies $\beta = 1$). We let $\mathcal{E}_n, \mathcal{O}_n$ denote the even/odd side of the bipartition of the graph $\mathbb{Z}_m^n$ (we make the dependence on n explicit to clarify the induction step). For a subset $Y \subseteq \mathbb{Z}_m^n$ and $i \in \mathbb{Z}_m$ we let

$$Y_i := \left\{(y_1, \dots, y_{n-1}) \in \mathbb{Z}_m^{n-1} : (y_1, \dots, y_{n-1}, i) \in Y\right\}.$$

We note that since $X \subseteq \mathcal{O}_n$ or $X \subseteq \mathcal{E}_n$, we have $X_i \subseteq \mathcal{O}_{n-1}$ or $X \subseteq \mathcal{E}_{n-1}$ for each i. Letting $|X_i| = \beta_i m^{n-1}/2$, we may appeal to the inductive hypothesis to conclude that

$$|N(X_i)| \geq |X_i| \left(1 + 4\sqrt{2} \cdot \frac{(1-\beta_i)}{m^2\sqrt{n-1}}\right) \tag{6.1}$$

for each $i \in \mathbb{Z}_m$. Now, we have

$$|N(X)| = \sum_{i \in \mathbb{Z}_m} |N(X)_i| = \sum_{i \in \mathbb{Z}_m} |N(X_i) \cup X_{i+1} \cup X_{i-1}|. \tag{6.2}$$

Let $\delta = 4\sqrt{2}\beta(1-\beta)/(m\sqrt{n})$ and suppose that $|\beta_i - \beta_{i+1}| \geq \delta$ for some $i \in \mathbb{Z}_m$. By the above inequality and the fact that $|N(X_i) \cup X_{i+1} \cup X_{i-1}| \geq \max\{|X_{i-1}|, |X_i|, |X_{i+1}|\}$ we would then have

$$|N(X)| \geq \delta m^{n-1}/2 + \sum_{i \in \mathbb{Z}_m} |X_i| = \delta m^{n-1}/2 + |X| = |X| \left(1 + 4\sqrt{2} \cdot \frac{(1-\beta)}{m^2\sqrt{n}}\right).$$

We may therefore assume that $|\beta_i - \beta_{i+1}| < \delta$ for each $i \in \mathbb{Z}_m$. Since $\beta = \sum_i \beta_i/m$ we have $\beta_i \in (\beta - m\delta/4, \beta + m\delta/4)$ for each i. By (6.1) and (6.2) we have

$$|N(X)| \geq |X| + \frac{m^{n-1}}{2} \sum_{i \in \mathbb{Z}_m} 4\sqrt{2} \cdot \frac{\beta_i(1-\beta_i)}{m^2\sqrt{n-1}}.$$

Letting $g(\beta) = \beta(1-\beta)$, we would therefore be done if we could show that

$$\frac{1}{m} \sum_{i \in \mathbb{Z}_m} g(\beta_i) \geq g(\beta)\sqrt{1 - 1/n}.$$

Since g is concave, the minimum of the sum $\sum_{i \in \mathbb{Z}_m} g(\beta_i)$ under the constraints $\beta = \sum_i \beta_i/m$ and $\beta_i \in [\beta - m\delta/4, \beta + m\delta/4]$ for all i is achieved when half of the β_i take value $\beta - m\delta/4$ and the other half take value $\beta + m\delta/4$. Therefore

$$\begin{aligned} \frac{1}{mg(\beta)} \sum_{i \in \mathbb{Z}_m} g(\beta_i) &\geq \frac{g(\beta - m\delta/4) + g(\beta + m\delta/4)}{2g(\beta)} \\ &= 1 - \frac{m^2\delta^2}{16\beta(1-\beta)} \end{aligned}$$

$$\begin{aligned} &= 1 - \frac{2\beta(1-\beta)}{n} \\ &\geq 1 - 1/(2n) \\ &\geq \sqrt{1-1/n} \end{aligned}$$

where we have used the inequality $\sqrt{1-x} \leq 1 - x/2$ for $x \in [0,1]$. □

We remark that Riordan [51] exhibited an ordering of the vertices of $\mathbb{Z}_m^n$ whose initial segments have minimal vertex boundary for their size. We suspect that initial segments of Riordan's ordering restricted to $\mathcal{E}$ give the correct lower bound for Lemma 6.1. However, the approximate statement of Lemma 6.1 will suffice for our purposes.

We will also make use of the fact that sets of polynomial size have much more substantial expansion. The following is a consequence of [51, Theorem 1.1].

Lemma 6.2. *There exists $C > 0$ such that if $X \subseteq V$, $|X| = O(n^4)$ then $|N(X)| \geq C|X|$, and if $|X| \leq n/2$, then $|N(X)| \geq n|X| - 2|X|(|X|-1)$.*

7. Approximate polymers

In this section we introduce some tools belonging to the graph container method introduced by Korshunov and Sapozhenko [39,52]. The results of this section are a minor modification of known results in this area. In particular, we adapt the arguments from the excellent exposition of Galvin [20].

Recall from the introduction to Section 5 that our goal is to balance the number of polymers satisfying a certain constraint with the total weight of such polymers.

Sapozhenko's approximation tools will allow us to replace polymers by 'approximate polymers' which are far less numerous, but retain enough information so that we can effectively bound the total weight of polymers with a given approximation.

We write 2^V to denote the power set of V. For a subset $Y \subseteq V$ and $v \in V$, we let $d_Y(v) := |N(v) \cap Y|$. Recall that G is a d-regular graph where $d = (1 + \mathbb{1}_{m>2})n$.

Let $\psi = \psi(d) > 0$. A *ψ-approximating pair* for a subset $X \subseteq V$ is a pair $(F, S) \in 2^V \times 2^V$ satisfying

$$F \subseteq N(X),\ S \supseteq X\,, \tag{7.1}$$

$$d_{V\setminus F}(u) \leq \psi \quad \forall u \in S \tag{7.2}$$

and

$$d_S(v) \leq \psi \quad \forall v \in V\setminus F\,. \tag{7.3}$$

We think of ψ as small compared to d and so conditions (7.2) and (7.3) say that the graph between S and $V\setminus F$ is sparse. On the other hand the graph between X

and $V\backslash N(X)$ is empty. We therefore view S as an approximation of X and F as an approximation of $N(X)$.

For any $g \in \mathbb{N}$, let

$$\mathcal{H}(g) = \{X \subseteq V : X \text{ is } G^2\text{-connected}, |N(X)| = g\}. \tag{7.4}$$

Our goal in this section is to prove the following.

Lemma 7.1. *Let $\psi \leq d/2$. There is a family $\mathcal{U} = \mathcal{U}(g) \subseteq 2^V \times 2^V$ with*

$$|\mathcal{U}| \leq 2^{O(g(\log d)/\psi + d)}$$

such that every $X \in \mathcal{H}(g)$ has a ψ-approximating pair in $\mathcal{U}$.

The key point here is that the bound on the size of $\mathcal{U}(g)$ is much smaller than the size of $\mathcal{H}(g)$.

We say that a set $Y \subseteq V$ *covers* a set $X \subseteq V$ if each $x \in X$ has a neighbour in Y, and each $y \in Y$ has a neighbour in X (note that this relation is symmetric despite the asymmetric terminology).

Lemma 7.2. *Let $X \subseteq V$, then there exists a set $Y \subseteq V$ such that Y covers X and $|Y| \leq 4|N(X)|/d$.*

Proof. Let $\ell = 2^k$ be the largest power of 2 less than or equal to n. We label the first ℓ coordinates of $V = \{0, \dots, m-1\}^n$ with elements of the group $\mathbb{Z}_2^k$, that is we fix a bijection $\phi : [\ell] \to \mathbb{Z}_2^k$. We then define the function $\Phi : V \to \mathbb{Z}_2^k$ given by

$$\Phi(x) = \sum_{i \in [\ell]} x_i \cdot \phi(i).$$

Observe that for each $i \in [\ell]$, the set $N(X) \cap \Phi^{-1}(\phi(i))$ covers X and one such set must have size at most $|N(X)|/\ell \leq 2|N(X)|/n$. The result follows by recalling that $d = (1 + \mathbb{1}_{m>2})n$. □

The function Φ in the above proof was inspired by a construction of Linial, Meshulam and Tarsi [43] used to bound the chromatic number of the graph on vertex set $\{0,1\}^n$ where $x \sim y$ if and only if x and y are at Hamming distance 2. A similar construction will reappear in the next section in a crucial role (see Lemma 8.1).

For the remainder of this section, it will be useful to generalise our notion of a G^2-connected set: for $k \in \mathbb{N}$, say that a subset $S \subseteq V(G)$ is *G^k-connected* if the graph $G^k[S]$ is connected (here G^k denotes the kth power of the graph G).

The following well-known lemma is a useful enumeration tool when considering G^k-connected sets.

Lemma 7.3 *([7], Lemma 2.1). In a graph of maximum degree at most Δ, the number of connected, induced subgraphs of order t containing a fixed vertex v is at most $(e\Delta)^{t-1}$.*

Lemma 7.4. *For all $g \in \mathbb{N}$, there exists a family $\mathcal{V} = \mathcal{V}(g) \subseteq 2^V$ with*

$$|\mathcal{V}| \leq 2^{O(g(\log d)/d+d)}$$

such that for any $X \in \mathcal{H}(g)$, $\mathcal{V}$ contains a set which covers X.

Proof. Given $X \in \mathcal{H}(g)$, Lemma 7.2 gives us a set $F \subseteq N(X)$ which covers X where $|F| \leq 4g/d$. Moreover, since every vertex of F is adjacent to a vertex of X and vice versa, we see that F must be G^4-connected. We may therefore take $\mathcal{V}$ to be the set of all G^4-connected subsets of V of size at most $4g/d$. Note that G^4 has maximum degree $\leq d^4$ and so the number of G^4-connected sets of some size t, through a given vertex in G is, by Lemma 7.3, at most $(ed^4)^{t-1}$. The number of G^4-connected subsets of V of size at most $4g/d$ is therefore at most

$$m^d \sum_{t=1}^{4g/d} (ed^4)^{t-1} = 2^{O(g(\log d)/d+d)}. \quad \square$$

In order to prove Lemma 7.1, it will suffice to prove the following.

Lemma 7.5. *Let $\psi \leq d/2$. For each $Y \subseteq V$ there is a family $\mathcal{W} = \mathcal{W}(Y, g) \subseteq 2^V \times 2^V$ with*

$$|\mathcal{W}| \leq 2^{O(g(\log d)/\psi)}$$

such that any $X \subseteq V$ covered by Y with $|N(X)| = g$ has a ψ-approximating pair in $\mathcal{W}$.

Indeed, letting

$$\mathcal{U}(g) = \bigcup_{Y \in \mathcal{V}(g)} \mathcal{W}(Y, g)$$

where $\mathcal{V}(g)$ and $\mathcal{W}(Y, g)$ are the families from Lemmas 7.4 and 7.5, it is clear that $\mathcal{U}(g)$ has the properties claimed in Lemma 7.1. It remains to prove Lemma 7.5.

We give an algorithm which, given a set $X \subseteq V$ and $Y \subseteq N(X)$, produces a ψ-approximating pair for X.

Algorithm 1:

Input: $(X, Y) \in 2^V \times 2^V$ such that $Y \subseteq N(X)$.

Step 1: Set $F' = Y$. If $\{u \in X : d_{N(X)\setminus F'}(u) \geq \psi\} \neq \emptyset$, pick u in this set and update F' by $F' \longleftarrow F' \cup N(u)$. Repeat until $\{u \in X : d_{N(X)\setminus F'}(u) \geq \psi\} = \emptyset$.

Step 2: Set $S' = \{u \in V : d_{V\setminus F'}(u) \leq \psi\}$. If $\{w \in V\setminus N(X) : d_{S'}(w) \geq \psi\} \neq \emptyset$, pick w in this set and update S' by $S' \longleftarrow S'\setminus N(w)$. Repeat until $\{w \in V\setminus N(X) : d_{S'}(w) \geq \psi\} = \emptyset$.

Output: $(S, F) = (S', F' \cup \{w \in V : d_{S'}(w) \geq \psi\})$.

Lemma 7.6. *The output of Algorithm 1 is a ψ-approximating pair for X.*

Proof. To see that $F \subseteq N(X)$ first note that at termination, $F' \subseteq N(X)$ since $Y \subseteq N(X)$, and the vertices added to F' in Step 1 are all in $N(X)$. Moreover $\{w \in V : d_{S'}(w) \geq \psi\} \subseteq N(X)$ else Step 2 would not have terminated. To see that $S \supseteq X$, note that initially $S' \supseteq X$ else Step 1 would not have terminated and Step 2 deletes from S' only neighbours of $V\setminus N(X)$.

To verify (7.2), note that at the start of Step 2, $d_{V\setminus F'}(u) \leq \psi$ for all $u \in S'$ by definition and $F' \subseteq F$ and $S \subseteq S'$. Condition (7.3) is immediate from the definition of F. □

We now prove Lemma 7.5 (and hence also Lemma 7.1).

Proof of Lemma 7.5. Let $Y \subseteq V$. We will show that as the input (X, Y) to Algorithm 1 runs over X such that Y covers X and $|N(X)| = g$, we get at most $2^{O(g(\log d)/\psi)}$ distinct outputs. We are then done by taking $\mathcal{W}$ to be the set of these outputs.

Suppose then that Y covers X where $|N(X)| = g$. Since Y covers X we have $X \subseteq N(Y)$ and $Y \subseteq N(X)$. In particular $|Y| \leq g$. Consider Algorithm 1 with input (X, Y). The output of Step 1 of the algorithm is determined by the set of u's whose neighbourhoods are added to F' in Step 1, and the set of w's whose neighbourhoods are removed from S' in Step 2.

Each iteration in Step 1 reduces the size of $N(X)\setminus F'$ by at least ψ, so there are at most g/ψ iterations. The u's in Step 1 are all drawn from X and hence $N(Y)$, a set of size at most dg (since $|Y| \leq g$). The number of possible outputs for Step 1 is therefore at most

$$\binom{dg}{\leq g/\psi} = 2^{O(g(\log d)/\psi)}.$$

We perform a similar analysis on Step 2. At the start of Step 2, $d_{N(X)}(u) \geq d - \psi$ for each $u \in S'$ so that

$$|S'|(d-\psi) \leq \sum_{u\in S'} d_{N(X)}(u) = \sum_{u\in N(X)} d_{S'}(u) \leq dg$$

and so initially $|S'| \leq 2g$ (since $\psi \leq d/2$). Each iteration in Step 2 reduces the size of S' by at least ψ, so there are at most $2g/\psi$ iterations. Each iteration draws a vertex w from $N(S')$, a set whose elements are all at distance at most 4 from Y (indeed, $S' \subseteq N(F')$

by definition, $F' \subseteq N(X)$ and $X \subseteq N(Y)$). We thus have $|N(S')| \leq d^4 g$ and so as in Step 1, the total number of outputs for Step 2 is $2^{O(g(\log d)/\psi)}$. □

8. Grouping polymers and decomposing the torus

In this section we build on the results of the previous one in order to enumerate polymers satisfying certain constraints. One of these constraints corresponds to a measure of 'roughness' of the polymer which we refer to as the *neighbourhood distribution* (see (8.4) below). Again, as outlined at the beginning of Section 5, our current goal is to verify the Kotecký-Preiss condition (2.1) (for suitably chosen functions f and g) for each of our polymer models (Lemma 4.4). A key component in our enumeration steps will be the entropy tools introduced in Section 5. We will also rely on an algebraically constructed vertex partition of $\mathbb{Z}_m^n$ which enjoys good 'covering properties' (see Lemma 8.1). We note that this is one of the key places where we take advantage of the specific structure of the torus.

A quirk of our algebraic construction is that the results of this section go through more easily if the dimension of the torus is assumed to be a power of 2. In the general case, we must decompose the torus into a collection of subtori each of dimension n' where n' is divisible by a large power of 2. The reason for this will become clearer after our covering property is stated formally. First we define our decomposition of G into subtori.

Choose $k \in \mathbb{N}$ such that

$$\ell := 2^k = \Theta(\sqrt{d} \log^4 d)\,. \tag{8.1}$$

The precise value of ℓ will become relevant in the next section (see the proof of Lemma 9.1). Let $0 \leq p \leq \ell - 1$ be such that $p \equiv n \pmod{\ell}$. We split the torus $G = \mathbb{Z}_m^n$ into m^p subtori in the following way. For $x \in \{0, \dots, m-1\}^p$ let

$$T_x := G\left[\{v = (v_1, \dots, v_n) \in V : (v_{n-p+1}, \dots, v_n) = x\}\right]\,.$$

We note that $T_x \cong \mathbb{Z}_m^{n-p}$ for each $x \in \{0, \dots, m-1\}^p$. We let G' denote the graph that is the disjoint union of all the T_x and note that $V(G') = V(G) = V$. For $v \in V$, let $N'(v)$ denote the neighbourhood of v in G' and let $d'(v)$ denote the degree of v in G'. We allow for similar modifications of degree notation, for example if $X \subseteq V$ we let $d'_X(v)$ denote the size of the set $N'(v) \cap X$. We note that T_x is a d'-regular graph where

$$d - 2\ell \leq d' = (1 + \mathbb{1}_{m>2})(n-p) \leq d\,. \tag{8.2}$$

We may now introduce the covering property alluded to above.

Lemma 8.1. *There exists a vertex partition $V = V_1 \cup \ldots \cup V_\ell$ with the following property. For each $i \in [\ell]$ and $v \in V$,*

$$d'_{V_i}(v) = d'/\ell\,.$$

Proof. Fix a bijection

$$\phi : [\ell] \to \mathbb{Z}_2^k\,.$$

For $t \in \mathbb{N}$, let $\overline{t}$ denote the element of $[\ell]$ for which $t \equiv \overline{t} \pmod{\ell}$. We consider the function $\Phi : V \to \mathbb{Z}_2^k$ given by

$$\Phi(v) = \sum_{t\in[n]} v_t \cdot \phi\left(\overline{t}\right)\,.$$

Let V_i denote the set of $v \in V$ such that $\Phi(v) = \phi(i)$ and note that $V_1 \cup \ldots \cup V_\ell$ forms a partition of V. For $t \in [n]$, let $e_t = (0, \ldots, 1, \ldots, 0) \in V$ be the standard basis vector in direction t. For $v \in V$ we have

$$N'(v) = \{v \pm e_1, \ldots, v \pm e_{n-p}\}$$

(where the addition is componentwise modulo m). Since for $t \in [n]$

$$\Phi(v \pm e_t) = \Phi(v) + \phi(\overline{t})\,,$$

we have for $s \in \{1, \ldots, (n-p)/\ell\}$

$$\{\Phi(v \pm e_t) : (s-1)\ell + 1 \le t \le s\ell\} = \mathbb{Z}_2^k\,.$$

For each $i \in [\ell]$, we therefore have $(n-p)/\ell$ distinct values of $t \in [n-p]$ for which $\Phi(v \pm e_t) = \phi(i)$ (or equivalently $v \pm e_t \in V_i$). It follows that for each $i \in [\ell]$, $d'_{V_i}(v) = (1 + \mathbb{1}_{m>2})(n-p)/\ell = d'/\ell$. □

The fact each vertex $v \in V$ has identical degree (with respect to G') into each set V_i will be crucial. By contrast, if we were to consider degrees with respect to G in the above lemma, we would find that $d_{V_i}(v)$ takes one of two values. This seemingly minor discrepancy introduces an imperfect covering system in an application of Shearer's lemma in the proof of the main enumeration result of this section (Lemma 8.3). The resulting error term produces a bound that is too weak for our purposes.

Before stating our polymer enumeration result we require some preliminaries. We define

$$\mathcal{N}(g) = \{X \subseteq V : |N(X)| = g\}\,,$$

and note that this differs from the definition of $\mathcal{H}(g)$ (see (7.4)) in that the elements of $\mathcal{N}(g)$ need not be G^2-connected.

For the remainder of this section we fix

$$0 < \psi < d/3\,. \tag{8.3}$$

For a pair $(F, S) \in 2^V \times 2^V$, we let

$$\mathcal{N}(g, F, S) := \{X \in \mathcal{N}(g) : \ (F, S) \text{ is a } \psi\text{-approximating pair for } X\ \}\ .$$

For a set $X \subseteq V$, we define the *neighbourhood distribution* of X to be the *multiset*

$$D_X := \{d'_X(v) : v \in N(X)\}\,. \tag{8.4}$$

We emphasise that in the definition of D_X, v ranges over the set $N(X)$ rather than $N'(X)$ and so some elements of the multiset may be 0. It turns out that the weight of a subset $X \subseteq V$ in our polymer models (as defined in (3.1)) is closely related to its neighbourhood distribution and so it will be useful to group sets according to their distribution.

Our goal is to enumerate the number of polymers $\gamma \in \mathcal{N}(g)$ with a fixed ψ-approximating pair (F, S) and neighbourhood distribution D (Lemma (8.3) below). Fixing the degree distribution is a rather delicate constraint and the covering lemma (Lemma 8.1) is crucial for dealing with this. Roughly speaking, our strategy is to specify a set $X \in \mathcal{N}(g)$ with a given neighbourhood distribution as follows: first specify $N(X) \cap V_i$ (with V_i as in Lemma 8.1) for some i and then specify X by selecting subsets of $N'(u)$ for each $u \in N(X) \cap V_i$. By Lemma 8.1, each vertex will be specified d'/ℓ times which we account for with an application of Shearer's Lemma. We choose the subsets of $N'(u)$ to have sizes compatible with the fixed neighbourhood distribution. We do this for each i and then average. We gain a critical entropy saving from the fact that on average the sets $N(X) \cap V_i$ are smaller than the full neighbourhood $N(X)$ by a factor of ℓ.

Let D be a multiset of size g and let $(F, S) \in 2^V \times 2^V$. We define

$$\mathcal{N}(D, F, S) := \{X \in \mathcal{N}(g, F, S) : D_X = D\}\ .$$

Let us also define

$$x = x_{F,S,D} : F \to \mathbb{N} \tag{8.5}$$

to be a function that maximises the quantity

$$\prod_{v \in F} \binom{d'_S(v)}{x(v)}$$

subject to the constraint $\{x(v) : v \in F\} \subseteq D$ (multiset inclusion). Note that we use the convention $\binom{0}{0} = 1$.

In the following we will make use of the classical theorem of Hardy and Ramanujan [29] on integer partitions. An *integer partition* of a natural number k is a multiset of natural numbers whose elements sum to k. We let $p(k)$ denote the number of integer partitions of k.

Theorem 8.2 *(Hardy and Ramanujan [29]).*

$$p(k) \sim \frac{1}{4k\sqrt{3}} e^{\pi\sqrt{2k/3}} \quad \text{as } k \to \infty\,.$$

We will not need the precise asymptotics of the above theorem, the bound $p(k) = e^{O(\sqrt{k})}$ will suffice for our purposes. The paper of Hardy and Ramanujan [29] in fact begins with an elementary proof of this weaker bound.

We may now state our main enumeration lemma.

Lemma 8.3. *For $(F,S) \in 2^V \times 2^V$ and multiset D of size g we have*

$$|\mathcal{N}(D,F,S)| \leq 2^{O(g(\log d)/\ell + \sqrt{dg\ell}) + (g-|F|)\psi/d'} \prod_{v\in F} \binom{d'_S(v)}{x(v)}^{1/d'},$$

where $x = x_{F,S,D}$ is defined in (8.5).

Proof. Let $V = V_1 \cup \ldots \cup V_\ell$ be the partition of V from Lemma 8.1. With this partition in mind, we refine our view of the set $\mathcal{N} := \mathcal{N}(D,F,S)$ in the following way. Let us fix an ordered ℓ-tuple $\Pi = (\Pi_1, \ldots, \Pi_\ell)$, where each Π_i is a multiset of elements in $[d']$ and $D = \bigcup_i \Pi_i$ (multiset union). For each $i \in [\ell]$, we let $\mathcal{N}(\Pi_i)$ denote the elements $X \in \mathcal{N}$ for which

$$\Pi_i = \{d'_X(v) : v \in V_i \cap N(X)\}\,,$$

and we let

$$\mathcal{N}(\Pi) := \bigcap_{i=1}^{\ell} \mathcal{N}(\Pi_i)\,.$$

We now bound the number of tuples Π for which $\mathcal{N}(\Pi)$ is non-empty. Observe that by Lemma 8.1, for $X \in \mathcal{N}$ we have

$$\sum_{v\in V_i} d'_X(v) = \sum_{v\in X} d'_{V_i}(v) = d'|X|/\ell \leq dg/\ell\,.$$

For $\mathcal{N}(\Pi)$ to be non-empty, the non-zero elements of Π_i must therefore be an integer partition of some integer $k_i \leq dg/\ell$. The number of choices for each Π_i is therefore at

most $2^{O(\sqrt{dg/\ell})}$ by Theorem 8.2 (note that there are $< g$ choices for the multiplicity of 0 in Π_i) and so the number of choices for $\Pi = (\Pi_1, \dots, \Pi_\ell)$ is at most

$$2^{O(\sqrt{dg\ell})}\,.$$

We may therefore bound the size of $\mathcal{N}$ by

$$|\mathcal{N}| \le 2^{O(\sqrt{dg\ell})} \max_{\Pi} |\mathcal{N}(\Pi)|\,. \tag{8.6}$$

Given $\Pi = (\Pi_1, \dots, \Pi_\ell)$, we bound the size of $\mathcal{N}(\Pi)$ via the following inequality which follows from the definition of $\mathcal{N}(\Pi)$

$$|\mathcal{N}(\Pi)| \le \min_i |\mathcal{N}(\Pi_i)| \le \left(\prod_{i=1}^{\ell} |\mathcal{N}(\Pi_i)|\right)^{1/\ell}\,. \tag{8.7}$$

Fix $i \in [\ell]$ and let $X \in \mathcal{N}(\Pi_i)$.

Recall that (F, S) is a ψ-approximating pair for X and so by condition (7.2) in the definition of an approximating pair

$$(d-\psi)|S| \le \sum_{v\in S} d_F(v) = \sum_{v\in F} d_S(v) \le dg\,.$$

For the final inequality we used that $|F| \le |N(X)| = g$ by condition (7.1). Since $\psi < d/3$ by assumption (8.3) we have

$$|S| \le 2g \tag{8.8}$$

and so

$$|N(S)| \le 2dg\,. \tag{8.9}$$

We have $N(X) \subseteq N(S)$ by (7.1). There are therefore at most $\binom{2dg}{|\Pi_i|}$ choices for $V_i \cap N(X)$. Let us fix such a choice and denote it by N_i. Fix a function

$$\pi : N_i \to \{0, \dots, d'\}\,,$$

such that $\{\pi(v) : v \in N_i\} = \Pi_i$. Crudely there are at most $d^{|N_i|} = d^{|\Pi_i|}$ choices for the function π. We now bound the number of choices for $X \in \mathcal{N}(\Pi_i)$ such that $V_i \cap N(X) = N_i$ and

$$d'_X(v) = \pi(v) \text{ for each } v \in N_i\,.$$

Recall that by Lemma 8.1, for every $v \in V$, we have $d'_{V_i}(v) = d'/\ell$. Let

$$S' = \{v \in S : d'_{N_i}(v) = d'/\ell\}$$

and note that $X \subseteq S'$ as $N_i = V_i \cap N(X)$. Let Ψ denote the family of subsets $X \subseteq S'$ such that

$$d'_X(v) = \pi(v) \text{ for each } v \in N_i.$$

We will estimate $|\Psi|$ by bounding the entropy $H(X)$, where X is a uniformly random element of Ψ. For each $v \in N_i$, let $X_v = X \cap N'(v)$. Note that by definition of S', each vertex $w \in S'$ is contained in d'/ℓ of the sets $\{N'(v) : v \in N_i\}$. Let $F_i := V_i \cap F$. It follows by (5.1) and Shearer's Lemma (Lemma 5.2), that

$$\begin{aligned}
\log|\Psi| = H(X) &\le \frac{\ell}{d'}\sum_{v\in N_i} H(X_v)\\
&\le \frac{\ell}{d'}\sum_{v\in N_i}\log\binom{d'_S(v)}{\pi(v)}\\
&\le \frac{\ell}{d'}\left(\sum_{v\in F_i}\log\binom{d'_S(v)}{\pi(v)} + |N_i\backslash F_i|\psi\right)\\
&\le \frac{\ell}{d'}\left(\sum_{v\in F_i}\log\binom{d'_S(v)}{x_i(v)} + |N_i\backslash F_i|\psi\right)
\end{aligned}$$

where $x_i := x_{F_i,S,\Pi_i}$ (see (8.5)). For the penultimate inequality we used $d'_S(v) \le \psi$ for each $v \in N_i\backslash F_i$ by the definition of an approximating pair.

Putting everything together we have

$$|\mathcal{N}(\Pi_i)| \le \binom{2dg}{|\Pi_i|} d^{|\Pi_i|}\left[2^{(|\Pi_i|-|F_i|)\psi}\prod_{v\in F_i}\binom{d'_S(v)}{x_i(v)}\right]^{\ell/d'}.$$

By (8.7) we then have

$$\begin{aligned}
|\mathcal{N}(\Pi)| &\le \left[\prod_{i=1}^{\ell}\binom{2dg}{|\Pi_i|} d^{|\Pi_i|}2^{(|\Pi_i|-|F_i|)\psi\ell/d'}\prod_{v\in F_i}\binom{d'_S(v)}{x_i(v)}^{\ell/d'}\right]^{1/\ell}\\
&\le \binom{2dg}{g/\ell} d^{g/\ell}2^{(g-|F|)\psi/d'}\prod_{v\in F}\binom{d'_S(v)}{x(v)}^{1/d'}\\
&= 2^{O(g(\log d)/\ell)+(g-|F|)\psi/d'}\prod_{v\in F}\binom{d'_S(v)}{x(v)}^{1/d'}.
\end{aligned}$$

The lemma follows from (8.6). □

We remark that the important factor in the above lemma is the product of binomial coefficients. It will be useful to keep the following bound in mind:

$$\prod_{v\in F}\binom{d'_S(v)}{x(v)}^{1/d'} \le \prod_{t\in D}\binom{d'}{t}^{1/d'} \le \exp_2\left\{\sum_{t\in D} H(t/d')\right\}. \tag{8.10}$$

The right hand side should be compared with the entropy terms that arise in our bound for the total weight of polymers with a given neighbourhood distribution (Lemma 9.2 to appear in the next section).

One could now bound the number of polymers $\gamma \in \mathcal{H}(g)$ with a given neighbourhood distribution simply by summing the bounds of Lemma 8.3 over all approximating pairs in $\mathcal{U}(g)$ (as defined in Lemma 7.1). However, it turns out to be too costly to specify the entire neighbourhood distribution of a polymer and we have to settle for specifying only part of it.

Definition 8.4. Let $\gamma \in \mathcal{P}$ be such that $|N(\gamma\cap\mathcal{E})| \ge |N(\gamma\cap\mathcal{O})|$, and suppose (F,S) is a ψ-approximating pair for γ. We call the pair

$$(\gamma\cap\mathcal{E}, S\cap\mathcal{O})$$

a *fingerprint* of the polymer γ. We make the same definition with the roles of $\mathcal{E},\mathcal{O}$ reversed.

Note that multiple polymers may have the same fingerprint. Our goal is to bound the number of possible fingerprints $R=(\rho,T)$ such that R is the fingerprint of a polymer γ with $|N(\gamma)| = g$ and the neighbourhood distribution of ρ is D. Call the set of all such fingerprints $\mathcal{R}(g,D)$. Given $(F,S)\in 2^V\times 2^V$, let $\mathcal{R}(g,D,F,S)$ denote the set of $R\in\mathcal{R}(g,D)$ such that R is the fingerprint for some $\gamma\in\mathcal{H}(g)$ with ψ-approximating pair (F,S). Finally let

$$\mathcal{R}_{\mathcal{E}}(g,D) = \{R\in\mathcal{R}(g,D) : R=(\rho,T) \text{ where } \rho\subseteq\mathcal{E}\},$$

and

$$\mathcal{R}_{\mathcal{E}}(g,D,F,S) = \{R\in\mathcal{R}(g,D,F,S) : R=(\rho,T) \text{ where } \rho\subseteq\mathcal{E}\},$$

and define $\mathcal{R}_{\mathcal{O}}(g,D)$ and $\mathcal{R}_{\mathcal{O}}(g,D,F,S)$ similarly.

Lemma 8.5. *For all g, $(F,S)\in 2^V\times 2^V$ and multisets D,*

$$|\mathcal{R}_{\mathcal{E}}(g,D,F,S)| \le 2^{O(g(\log d)/\ell+\sqrt{dg\ell})+(|D|-|F\cap\mathcal{O}|)\psi/d'} \prod_{v\in F\cap\mathcal{O}}\binom{d'_{S\cap\mathcal{E}}(v)}{x(v)}^{1/d'},$$

where $x = x_{F\cap\mathcal{O},S\cap\mathcal{E},D}$ (as defined in (8.5)). The same bound holds with $\mathcal{O},\mathcal{E}$ swapped.

Proof. Let $R = (\rho, T) \in \mathcal{R}_{\mathcal{E}}(g, D, F, S)$, so that $(F \cap \mathcal{O}, S \cap \mathcal{E})$ is a ψ-approximating pair for the set ρ. Note that by the definition of a fingerprint we also have $|D| = |N(\rho)| \leq g$. By Lemma 8.3, there are therefore at most

$$2^{O(g(\log d)/\ell + \sqrt{dg\ell}) + (|D| - |F\cap\mathcal{O}|)\psi/d'} \prod_{v \in F\cap\mathcal{O}} \binom{d'_{S\cap\mathcal{E}}(v)}{x(v)}^{1/d'},$$

choices for ρ. The final claim of the lemma follows by symmetry. □

9. Bounding weights with entropy

In this section we establish a bound on the total weight of polymers in the set $\mathcal{H}(g)$ (as defined in (7.4)). This will be a crucial step toward the proof of Lemma 4.4 (verifying the Kotecký-Preiss condition). We complete the proof of Lemma 4.4 in the next section.

Recall from Definition 3.1 that for a polymer γ, $|N(\gamma \cap \mathcal{O})|, |N(\gamma \cap \mathcal{E})| < (1-\alpha)m^n/2$ where $\alpha = \alpha(H, \lambda)$ will be specified later (see (12.2)).

Lemma 9.1. *There exists $\xi > 0$ (depending on m and (H, λ)) such that for $d^4 \leq g \leq (1-\alpha)m^n/2$, we have*

$$\sum_{\gamma \in \mathcal{H}(g)} w(\gamma) \leq e^{-\xi g/(\sqrt{d}\log^2 d)}.$$

There are two broad steps in the proof of Lemma 9.1. In the first step, we bound the total weight of all polymers with a given *fingerprint*. In the second step, we sum over different classes of fingerprints and appeal to Lemma 8.5 (a bound on the number of fingerprints satisfying a certain set of constraints). To carry out the first step, we require a careful entropy argument inspired by Kahn and Park [38].

Given a fingerprint R, let $\mathcal{P}(R)$ denote the set of all polymers $\gamma \in \mathcal{P}$ with fingerprint R. Henceforth we let q denote the number of vertices in H and identify $V(H)$ with $\{1, \ldots, q\}$.

For the next lemma we introduce the following parameter of the weighted graph (H, λ):

$$r = r(H, \lambda) := \frac{\min_i \lambda_i}{\lambda_1 + \ldots + \lambda_q}. \tag{9.1}$$

For the remainder of this section we set

$$\psi = r^2 d/3$$

and note $0 < \psi < d/3$ so that the definition is consistent with assumption (8.3). We note also that

$$\psi < r^2 d'/2 \tag{9.2}$$

by (8.1) and (8.2).

In the following lemma, we bound polymer weights $w(\gamma)$ via the inclusion $\mathrm{Hom}(G,H) \subseteq \mathrm{Hom}(G',H)$ (recall the definition of G' from the start of the previous section). Although this may seem wasteful, it will help streamline the comparison with the polymer count of Lemma 8.3.

Lemma 9.2. *Suppose $R=(\rho,T)$ is a fingerprint of a polymer in $\mathcal{H}(g)$. Then*

$$\sum_{\gamma\in\mathcal{P}(R)} w(\gamma) \le \exp_2\left\{ -\sum_{t\in D} H\left(t/d'\right) - \frac{r|D_\psi|}{5} + \frac{3qg}{d'} \right\}.$$

where D is the neighbourhood distribution of ρ and D_ψ is the multiset of elements in D which are $\le \psi$.

For the proof the following notation will be useful. For $c \subseteq V(H)$, let

$$n(c) \coloneqq |\{v \in V(H) : v \sim c\}|.$$

For a graph F let

$$\eta(F) \coloneqq \max\{|A||B| : A,B \subseteq V(F), A \sim B\}, \tag{9.3}$$

that is $\eta(F)=\eta_i(F)$ (as defined in (1.3)) where $i : V(F) \to \mathbb{R}$ takes only the value 1 (we say that F is unweighted in this case).

Proof of Lemma 9.2. Suppose $\gamma \in \mathcal{P}(R)$. Recall from (3.1) that

$$w(\gamma) = \frac{\sum_{f\in\chi_{A,B}(\gamma)} \prod_{v\in V} \lambda_{f(v)}}{\eta_\lambda(H)^{m^n/2}},$$

where $\chi_{A,B}(\gamma)$ is the set of colourings $f \in \mathrm{Hom}(G,H)$ such that f disagrees with (A,B) at each $v \in \gamma$ and agrees at each $v \in V\backslash\gamma$. Let $\chi(\rho,T)$ denote the set of colourings f such that f disagrees with (A,B) at each $v \in \rho$ and agrees at each $v \in V\backslash(\rho\cup T)$. Note that we do not specify whether f agrees or disagrees with (A,B) on the vertices of T. Suppose without loss of generality that $\rho \subseteq \mathcal{E}$. Since $\gamma\cap\mathcal{E}=\rho$ and $\gamma\cap\mathcal{O}\subseteq T$ for each $\gamma\in\mathcal{P}(R)$ we have

$$\sum_{\gamma\in\mathcal{P}(R)} w(\gamma) \le \frac{\sum_{f\in\chi(\rho,T)} \prod_{v\in V} \lambda_{f(v)}}{\eta_\lambda(H)^{m^n/2}}. \tag{9.4}$$

Our strategy is to use a standard 'blowup trick' that allows us bound the right hand side of (9.4) by counting homomorphisms into a blowup of H. Having reduced to a

counting problem, we may bring entropy tools to bear. However, the blowup trick only works if the weight function λ takes only *rational* values. We deal with general λ via a limiting argument.

Claim 9.3. *If $\lambda : V(H) \to \mathbb{R}_{>0}$ takes only rational values, then*

$$\frac{\sum_{f\in\chi(\rho,T)}\prod_{v\in V}\lambda_{f(v)}}{\eta_\lambda(H)^{m^n/2}} \leq \exp_2\left\{-\sum_{t\in D} H\left(t/d'\right) - \frac{r|D_\psi|}{5} + \frac{3qg}{d'}\right\}.$$

Proof. First we introduce the blowup trick then we introduce the entropy arguments. Since λ takes only rational values, letting $V(H) = [q]$ we may write,

$$\lambda(i) = \frac{a_i}{b}$$

for each $i \in [q]$, where a_i, b are positive integers. For the sake of uniqueness, we assume b is chosen as small as possible.

Define the *blowup* $H[\lambda] = H(a_1, \ldots, a_q)$, to be the graph on vertex set $W_1 \cup \ldots \cup W_q$ where the W_i are disjoint, $|W_i| = a_i$ for all i and $x \sim y$ if and only if $x \in W_i$, $y \in W_j$ for some pair $\{i, j\} \in E(H)$. For a vertex $v \in V(H[\lambda])$, we let $\overline{v}$ denote the unique $i \in [q]$ for which $v \in W_i$. We extend this notation to subsets $Y \subseteq H[\lambda]$, letting $\overline{Y} \coloneqq \{\overline{y} : y \in Y\}$.

We let $A[\lambda] = \bigcup_{i\in A} W_i$ and define $B[\lambda]$ similarly. Let $\chi(\rho, T, \lambda)$ denote the set of $f \in \mathrm{Hom}(G, H[\lambda])$ such that f disagrees with $(A[\lambda], B[\lambda])$ at each $v \in \rho$ and agrees at each $v \in V\backslash(\rho \cup T)$. We have

$$|\chi(\rho, T, \lambda)| = \sum_{f\in\chi(\rho,T)}\prod_{v\in V} a_{f(v)}$$

and also (recalling the definition (9.3))

$$\eta(H[\lambda]) = b^2\eta_\lambda(H),$$

so that

$$\frac{|\chi(\rho, T, \lambda)|}{\eta(H[\lambda])^{m^n/2}} = \frac{\sum_{f\in\chi(\rho,T)}\prod_{v\in V}\lambda_{f(v)}}{\eta_\lambda(H)^{m^n/2}}.$$

We have therefore related the quantity we want to bound to the size of the set $\chi(\rho, T, \lambda)$. We bound $|\chi(\rho, T, \lambda)|$ using an entropy argument.

Suppose that $f \in \chi(\rho, T, \lambda)$ is chosen uniformly at random. For ease of notation we write N_u instead of $N'(u)$ for a vertex $u \in V$ in the following. Recall that for a colouring f and set $X \subseteq V$, we write f_X for the restriction of f to X.

By Lemma 5.1 we may decompose the entropy of f as

$$H(f) = H(f_\mathcal{E}) + H(f_\mathcal{O}|f_\mathcal{E}). \tag{9.5}$$

We bound these terms separately. Let $U = N'(\rho) \cup T \subseteq \mathcal{O}$. With an application of Shearer's lemma (Lemma 5.2) in mind we define the function $\varphi : 2^{\mathcal{E}} \to \mathbb{R}_{\geq 0}$ by

$$\varphi(S) = \begin{cases} 1/d' & \text{if } S = N_u \text{ for some } u \in U, \\ 1 - d'_U(u)/d' & \text{if } S = \{u\} \text{ for some } u \in \mathcal{E}, \\ 0 & \text{otherwise.} \end{cases}$$

Observe that $\sum_{S \ni u} \varphi(S) = 1$ for each $u \in \mathcal{E}$ and so we may apply Lemma 5.2 to obtain

$$H(f_{\mathcal{E}}) \leq \sum_{u \in U} \frac{1}{d'} H(f_{N_u}) + \sum_{u \in \mathcal{E}} \left(1 - \frac{d'_U(u)}{d'}\right) H(f_u)\,. \tag{9.6}$$

Turning to the second term of (9.5), we have by Lemma 5.1,

$$H(f_{\mathcal{O}}|f_{\mathcal{E}}) \leq \sum_{u \in U} H(f_u|\overline{f(N_u)}) + \sum_{u \in \mathcal{O}\setminus U} H(f_u)\,. \tag{9.7}$$

For a vertex $u \in V$, let

$$I(u) = \frac{1}{d'} H(f_{N_u}) + H(f_u|\overline{f(N_u)}) = \frac{1}{d'}\left[H(f_{N_u}|\overline{f(N_u)}) + H(\overline{f(N_u)})\right] + H(f_u|\overline{f(N_u)})\,.$$

Then by (9.5), (9.6) and (9.7) we may write

$$H(f) \leq \sum_{u \in U} I(u) + \sum_{u \in \mathcal{E}} \left(1 - \frac{d'_U(u)}{d'}\right) H(f_u) + \sum_{u \in \mathcal{O}\setminus U} H(f_u). \tag{9.8}$$

Again we bound each term individually. For a subset $c \subseteq V(H)$, let $a_c \coloneqq \sum_{i \in c} a_i$. For the second sum we note that $f(u) \in A[\lambda]$ for each $u \in \mathcal{E}\setminus\rho$, $d'_U(u) = d'$ for each $u \in \rho$ and $|A[\lambda]| = a_A$ so that

$$\sum_{u \in \mathcal{E}} \left(1 - \frac{d'_U(u)}{d'}\right) H(f_u) \leq \sum_{u \in \mathcal{E}\setminus\rho} \left(1 - \frac{d'_U(u)}{d'}\right) \log a_A = (m^n/2 - |U|) \log a_A\,.$$

For the third sum in (9.8) we simply use $f(u) \in B[\lambda]$ for each $u \in \mathcal{E}\setminus U$ so that

$$\sum_{u \in \mathcal{O}\setminus U} H(f_u) \leq (m^n/2 - |U|) \log a_B.$$

It remains to bound the first sum in (9.8). For any $u \in V$ we may bound $I(u)$ as follows:

$$I(u) \le \max_{c \subseteq V(H)} \left\{ \frac{1}{d'} H(f_{N_u} | \overline{f(N_u)} = c) + H(f_u | \overline{f(N_u)} = c) \right\} + \frac{1}{d'} H(\overline{f(N_u)}) \quad (9.9)$$

$$\le \max_{c \subseteq [q]} \left\{ \log a_c + \log a_{n(c)} \right\} + \frac{q}{d'}$$

$$\le \log \eta(H[\lambda]) + \frac{q}{d'} . \quad (9.10)$$

For the second inequality we used (5.1), noting that there are at most $a_c^{d'}$ possible colourings of N_u given that $\overline{f(N_u)} = c$, and that there are at most 2^q choices for c.

For $u \in N'(\rho)$ we will improve on the bound (9.10) by considering vertices in $N_u \cap \rho$, on which f disagrees with the pattern $(A[\lambda], B[\lambda])$. Let d_u denote $d'_\rho(u)$. Since $f \in \chi(\rho, T, \lambda)$ we know that $f(N_u \cap \rho) \subseteq (A[\lambda])^c$ and $f(N_u \backslash \rho) \subseteq A[\lambda]$. By (9.9), we then have

$$I(u) \le \max_{c \subseteq V(H)} \left\{ \frac{d_u}{d'} \log a_{c \cap A^c} + \left(1 - \frac{d_u}{d'}\right) \log a_{c \cap A} + \log a_{n(c)} \right\} + \frac{q}{d'} . \quad (9.11)$$

Holding c fixed and considering $a_{c \cap A^c}$, $a_{c \cap A}$ as continuous variables which sum to a_c, the above expression is maximised when $a_{c \cap A^c} = \frac{d_u}{d'} a_c$ and $a_{c \cap A} = \left(1 - \frac{d_u}{d'}\right) a_c$ and so

$$I(u) \le -H(d_u/d') + \max_{c \subseteq V(H)} \{ \log a_c + \log a_{n(c)} \} + \frac{q}{d'}$$

$$\le -H(d_u/d') + \log \eta(H) + \frac{q}{d'} . \quad (9.12)$$

The appearance of the $-H(d_u/d')$ terms here is key. These 'entropy penalties' will exactly balance the binomial terms in our polymer count, Lemma 8.3 (see also (8.10)).

Recalling the definition of r (9.1), we have

$$r = r(H, \lambda) = \frac{\min_i a_i}{a_1 + \ldots + a_q} .$$

If $d_u \le \psi$, we can do better in (9.11) by considering $a_{c \cap A^c}$, $a_{c \cap A}$ as integer variables which sum to a_c. In this case, since $\psi < rd'$ by (9.2), the expression in (9.11) is maximised when $a_{c \cap A^c} = \min_i a_i$ and $a_{c \cap A} = a_c - \min_i a_i$. We then have

$$I(u) \le \max_{c \subseteq V(H)} \{ \log(a_c - \min_i a_i) + \log a_{n(c)} \} + q/d'$$

$$\le \max_{c \subseteq V(H)} \{ \log a_c + \log a_{n(c)} + \log(1 - r) \} + q/d'$$

$$\le \log(1 - r) + \log \eta(H[\lambda]) + q/d'$$

$$\le -H(d_u/d') - r/5 + \log \eta(H[\lambda]) + q/d' . \quad (9.13)$$

For the final inequality we used $d_u/d' \le \psi/d' \le r^2/2$ and $H(x) \le -\log(1 - r) - r/5$ for $0 \le x \le r^2/2$.

By (9.10), (9.12) and (9.13),

$$\sum_{u\in U} I(u) \le |U|\left(\log \eta(H[\lambda]) + \frac{q}{d'}\right) - \sum_{t\in D} H(t/d') - r|D_\psi|/5\,.$$

Combining our bounds on the three sums in (9.8), we have

$$\log|\chi(\rho,T,\lambda)| = H(f) \le \frac{m^n}{2}\log\eta(H[\lambda]) - \sum_{t\in D} H\left(\frac{t}{d'}\right) - \frac{r|D_\psi|}{5} + \frac{q|U|}{d'}\,.$$

Recall that R is a fingerprint of a polymer $\gamma \in \mathcal{H}(g)$ and so $|N'(\rho)| \le g$. Moreover, since $T \subseteq S$ where S is a part of a ψ-approximating pair (F,S) of γ we have $|T| \le 2g$ by (8.8) and so $|U| \le 3g$. The claim follows. □

It remains to deal with the case where λ takes irrational values. We proceed by a limiting argument. Choose a sequence (λ^k) where $\lambda^k : V(H) \to \mathbb{Q}_{>0}$ and $\lambda^k \to \lambda$ pointwise. Note that, as $k \to \infty$,

$$\eta_{\lambda^k}(H) \to \eta_\lambda(H) \tag{9.14}$$

and (recalling (9.1))

$$r(H,\lambda^k) \to r(H,\lambda)\,. \tag{9.15}$$

With Claim 9.3 applied to $H[\lambda^k]$ and letting $\psi_k := r^2(H,\lambda^k)/3$ we obtain

$$\frac{\sum_{f\in\chi(\rho,T)}\prod_{v\in V}\lambda^k(f(v))}{\eta_{\lambda^k}(H)^{m^n/2}} \le \exp_2\left\{-\sum_{t\in D} H\left(t/d'\right) - \frac{r_k|D_{\psi_k}|}{5} + \frac{3qg}{d'}\right\}.$$

Taking the limit $k\to\infty$, using (9.14) and (9.15), and returning to (9.4) we obtain

$$\sum_{\gamma\in\mathcal{P}(R)} w(\gamma) \le \exp_2\left\{-\sum_{t\in D} H\left(t/d'\right) - \frac{r|D_\psi|}{5} + \frac{3qg}{d'}\right\}. \quad \square$$

By summing over possible fingerprints, we can bound the total weight of polymers in $\mathcal{H}(g)$.

Proof of Lemma 9.1. Let Δ denote the collection of multisets D for which $\mathcal{R}(g,D) \ne \emptyset$ (defined in the discussion following Definition 8.4). For $D \in \Delta$, the non-zero elements of D must be an integer partition of some multiple of d' which is $\le gd'$ and so by Theorem 8.2

$$|\Delta| \le 2^{O(\sqrt{gd})}\,. \tag{9.16}$$

Recall from Lemma 9.2 that for $D \in \Delta$, D_ψ is the multiset of elements in D that are $\leq \psi$.

Claim 9.4. *If $(F,S) \in 2^V \times 2^V$ is such that $\mathcal{R}_\mathcal{E}(g,D,F,S) \neq \emptyset$ then*

$$|D| - |F \cap \mathcal{O}| \leq |D_\psi|\,.$$

The same statement holds with $\mathcal{O}, \mathcal{E}$ swapped.

Proof. Let $(\rho, T) \in \mathcal{R}_\mathcal{E}(g,D,F,S)$ so that $(F \cap \mathcal{O}, S \cap \mathcal{E})$ is an approximating pair for ρ. By (7.1) and (7.3) in the definition of a ψ-approximating pair, we have $d_\rho(v) \leq d_{S\cap\mathcal{E}}(v) \leq \psi$ for each $v \in \mathcal{O}\backslash F$. Since D is the neighbourhood distribution of ρ, we thus have at least $|N(\rho)\backslash(F\cap\mathcal{O})| = |D| - |F\cap\mathcal{O}|$ elements of D which are $\leq \psi$. The claim holds with $\mathcal{O}, \mathcal{E}$ swapped by symmetry. □

Recall from (8.1) that we chose $k \in \mathbb{N}$ such that

$$\ell = 2^k = \Theta(\sqrt{d}\log^4 d) \tag{9.17}$$

and that $d - 2\ell \leq d' \leq d$ from (8.2).

Let $\Delta_1 \subseteq \Delta$ be the collection of multisets $D \in \Delta$ such that $|D_\psi| > g(\log d)^2/\ell$. Let $\Delta_0 = \Delta\backslash\Delta_1$. Suppose that $D \in \Delta_1$. Then

$$\begin{aligned}
&\sum_{R\in\mathcal{R}(g,D)} \sum_{\gamma\in\mathcal{P}(R)} w(\gamma)\\
&\leq 2 \sum_{(F,S)\in\mathcal{U}(g)} \sum_{R\in\mathcal{R}_\mathcal{E}(g,D,F,S)} \sum_{\gamma\in\mathcal{P}(R)} w(\gamma)\\
&\leq 2 \sum_{(F,S)\in\mathcal{U}(g)} \sum_{R\in\mathcal{R}_\mathcal{E}(g,D,F,S)} \exp_2\left\{-\sum_{t\in D} H(t/d') - r|D_\psi|/5 + 3qg/d'\right\}\\
&\leq 2^{O(g(\log d)/\ell)+|D_\psi|\psi/d'} \prod_{t\in D}\binom{d'}{t}^{1/d'} \exp_2\left\{-\sum_{t\in D} H(t/d') - r|D_\psi|/5\right\}\\
&\leq 2^{-\Omega(g(\log d)^2/\ell)}\,.
\end{aligned} \tag{9.18}$$

For the first inequality we use Corollary 7.1 where the factor of 2 accounts for the sum over $\mathcal{R}_\mathcal{O}(g,D,F,S)$ by symmetry. For the second inequality we use Lemma 9.2. For the third inequality we used Corollary 7.1, Lemma 8.5 and Claim 9.4. For the final inequality we use $r \leq 1/2$ in collecting the D_ψ terms.

Let us assume now that $D \in \Delta_0$. It will be useful to refine the notion of approximation for polymers $\gamma \in \mathcal{H}(g)$ with a fingerprint in $\mathcal{R}_\mathcal{E}(g,D)$ by taking advantage of the fact that any approximating pair for γ must specify almost all of $N(\gamma\cap\mathcal{E})$. Indeed suppose

that $\gamma \in \mathcal{H}(g)$ such that $|N(\gamma \cap \mathcal{E})| \geq |N(\gamma \cap \mathcal{O})|$. We call $(F, S) \in 2^V \times 2^V$ a *super-approximating pair* for γ if (F, S) is an approximating pair for γ and moreover

$$N(\gamma \cap \mathcal{E}) = F \cap \mathcal{O} \text{ and } N(S \cap \mathcal{E}) = F \cap \mathcal{O}\,. \tag{9.19}$$

Claim 9.5. *There exists a family $\mathcal{U}^* = \mathcal{U}^*(g) \subseteq 2^V \times 2^V$ such that*

$$|\mathcal{U}^*| \leq 2^{O(g(\log d)^3/\ell)} \tag{9.20}$$

with the following property: for each $D \in \Delta_0$ and $\gamma \in \mathcal{H}(g)$ with fingerprint in $\mathcal{R}_{\mathcal{E}}(g, D)$, there exists $(F, S) \in \mathcal{U}^*$ such that (F, S) is a super-approximating pair for γ.

We first show how our lemma follows from Claim 9.5.

Let $D \in \Delta_0$ and suppose that $(F, S) \in \mathcal{U}^*$ is such that $\mathcal{R}_{\mathcal{E}}(g, D, F, S) \neq \emptyset$. Recall that polymers are defined to be G^2-connected subsets $\gamma \subseteq V$ such that $|N(\gamma \cap \mathcal{E})|, |N(\gamma \cap \mathcal{O})| < (1-\alpha)m^n/2$ and so $|F \cap \mathcal{O}| < (1-\alpha)m^n/2$ by (9.19). Since $N(S \cap \mathcal{E}) = F \cap \mathcal{O}$ (again by (9.19)), we have by Lemma 6.1,

$$|S \cap \mathcal{E}| = \left(1 - \Omega(1/\sqrt{n})\right) |F \cap \mathcal{O}|\,. \tag{9.21}$$

Noting that $|F \cap \mathcal{O}| \geq g/2$, by (9.19) and the definition of a fingerprint (Definition 8.4), we thus have

$$e_{G'}(F \cap \mathcal{O}, (S \cap \mathcal{E})^c) = d'(|F \cap \mathcal{O}| - |S \cap \mathcal{E}|) = \Omega(g\sqrt{d})\,. \tag{9.22}$$

We pause here to highlight the fact that we used vertex isoperimetry in G to obtain (9.21) and then degree regularity in G' to deduce (9.22), an edge isoperimetric inequality in G'. We now use this edge expansion to more effectively bound the product in Lemma 8.5. We remark that edge expansion in G alone would not suffice for this purpose since boundary edges in G might be missing in G'.

Returning to the proof, note that since $|D| = |F \cap \mathcal{O}|$ by (9.19), we have by Lemma 8.5,

$$|\mathcal{R}_{\mathcal{E}}(g, D, F, S)| \leq 2^{O(g(\log d)/\ell)} \prod_{v \in F \cap \mathcal{O}} \binom{d'_{S \cap \mathcal{E}}(v)}{x(v)}^{1/d'}, \tag{9.23}$$

where $x = x_{F \cap \mathcal{O}, S \cap \mathcal{E}, D}$ (as defined in (8.5)).

Recall that $D \in \Delta_0$ so that all but at most $g(\log d)^2/\ell$ elements of D are greater that ψ. Suppose that $x(v) \geq \psi$, then for $x(v) \leq x \leq d'$ we have

$$\binom{x-1}{x(v)} \Big/ \binom{x}{x(v)} = 1 - x(v)/x \leq 1 - \psi/d'.$$

Recalling the value of ℓ from (9.17) we have

$$\prod_{v\in F\cap\mathcal{O}}\binom{d'_{S\cap\mathcal{E}}(v)}{x(v)} \le (1-\psi/d')^{e_{G'}(F\cap\mathcal{O},(S\cap\mathcal{E})^c)-gd(\log d)^2/\ell}\prod_{t\in D}\binom{d'}{t}$$
$$= (1-\psi/d')^{\Omega(g\sqrt{d})}\prod_{t\in D}\binom{d'}{t}.$$

Returning to (9.23) we therefore have

$$|\mathcal{R}_{\mathcal{E}}(g,D,F,S)| \le 2^{-\Omega(g/\sqrt{d})}\prod_{t\in D}\binom{d'}{t}^{1/d'}.$$

Thus by Lemma 9.2

$$\sum_{R\in\mathcal{R}_{\mathcal{E}}(g,D,F,S)}\sum_{\gamma\in\mathcal{P}(R)} w(\gamma) \le 2^{-\Omega(g/\sqrt{d})}\prod_{t\in D}\binom{d'}{t}^{1/d'}\exp_2\left\{-\sum_{t\in D}H(t/d')+3qg/d'\right\}$$
$$\le 2^{-\Omega(g/\sqrt{d})}. \tag{9.24}$$

By Claim 9.5 we have

$$\sum_{\gamma\in\mathcal{H}(g)} w(\gamma) \le 2\sum_{D\in\Delta_1}\sum_{R\in\mathcal{R}_{\mathcal{E}}(g,D)}\sum_{\gamma\in\mathcal{P}(R)} w(\gamma) + 2\sum_{D\in\Delta_0}\sum_{(F,S)\in\mathcal{U}^*}\sum_{R\in\mathcal{R}_{\mathcal{E}}(g,D,F,S)}\sum_{\gamma\in\mathcal{P}(R)} w(\gamma).$$

The factors of 2 account for the sums over $R_{\mathcal{O}}(g,D)$ by symmetry. By the bounds (9.16), (9.18), (9.20) and (9.24) we have that the right hand side is at most $2^{-\Omega(g/(\sqrt{d}\log^2 d))}$ (recall that $g \ge d^4$ by assumption).

It remains to verify Claim 9.5.

Proof of Claim 9.5. Let $(F,S)\in\mathcal{U}(g)$ where $\mathcal{U}(g)$ is the set from Corollary 7.1 and let

$$\mathcal{V}_{F,S} := \left\{(F\cap\mathcal{O})\cup Q : Q\in\binom{N(S\cap\mathcal{E})}{\le g(\log d)^2/\ell}\right\} \subseteq 2^{\mathcal{O}}.$$

For $X\subseteq V$ let $B(X) := \{v\in V : N(v)\subseteq X\}$. Define

$$\mathcal{U}^* := \bigcup_{(F,S)\in\mathcal{U}(g)}\bigcup_{X\in\mathcal{V}_{F,S}}((F\cap\mathcal{E})\cup X, (S\cap\mathcal{O})\cup B(X)).$$

We now show that $\mathcal{U}^*$ has the desired properties. Let $D\in\Delta_0$ and let $\gamma\in\mathcal{H}(g)$ such that $|N(\gamma\cap\mathcal{E})|\ge|N(\gamma\cap\mathcal{O})|$. Let $\rho := \gamma\cap\mathcal{E}$ have neighbourhood distribution D. By Corollary 7.1 there exists $(F,S)\in\mathcal{U}(g)$ such that (F,S) is an approximating pair for γ. In particular by (7.1), $F\cap\mathcal{O}\subseteq N(\rho)\subseteq N(S)$. By Claim 9.4, $|N(\rho)\setminus(F\cap\mathcal{O})|\le g(\log d)^2/\ell$ and so $N(\rho)\in\mathcal{V}_{F,S}$. It follows that the pair

$$(F',S') = ((F\cap\mathcal{E})\cup N(\rho), (S\cap\mathcal{O})\cup B(N(\rho)))$$

belongs to $\mathcal{U}^*$. It is easy to verify that (F', S') is a super-approximating pair for γ. It remains to bound on the size of $\mathcal{U}^*$.

For $(F, S) \in \mathcal{U}(g)$ we have $|N(S)| \leq 2dg$ by (8.9) and so $|\mathcal{V}_{F,S}| \leq \binom{2dg}{\leq g(\log d)^2/\ell} = 2^{O(g(\log d)^3/\ell)}$. By Corollary 7.1 we thus have

$$|\mathcal{U}^*| \leq 2^{O(g(\log d)/\psi + g(\log d)^3/\ell)} = 2^{O(g(\log d)^3/\ell)} \quad \square \quad \square$$

10. Verifying the Kotecký-Preiss condition

In this section we prove Lemma 4.4. That is, we verify the Kotecký-Preiss condition (2.1) for each of our polymer models with suitable functions f and g. We then use this lemma to establish tail bounds on the cluster expansions of the log partition functions $\ln \Xi_{A,B}$. In particular we prove a more detailed version Lemma 4.5.

Proof of Lemma 4.4. With ξ as in Lemma 9.1 and r as defined in (9.1), let

$$f(\gamma) = |\gamma|/d \text{ and } \quad g(\gamma) = \begin{cases} r|N(\gamma)|/3 & \text{if } |N(\gamma)| \leq d^4 \\ \xi|N(\gamma)|/(4\sqrt{d}\log^2 d) & \text{otherwise .} \end{cases} \tag{10.1}$$

Let us fix a dominant pattern $(A, B) \in \mathcal{D}_\lambda(H)$. We denote $w_{A,B}$ simply by w. We want to show that the condition (4.4) holds with the above choice of f and g. That is, we want to show that

$$\sum_{\gamma' : d(\gamma', \gamma) \leq 2} w(\gamma') e^{f(\gamma') + g(\gamma')} \leq f(\gamma) \tag{10.2}$$

for all polymers $\gamma \in \mathcal{P}$.

We will in fact show that for each $v \in V(G)$

$$\sum_{\gamma' : \gamma' \ni v} w(\gamma') \cdot e^{f(\gamma') + g(\gamma')} \leq \frac{1}{d^3} = \frac{f(\gamma)}{d^2|\gamma|}, \tag{10.3}$$

for all $\gamma \in \mathcal{P}$. By summing this inequality over all v at distance at most 2 from γ in G (noting that there are $\leq d^2|\gamma|$ such vertices) we establish (10.2).

We split the sum appearing in (10.3) according to the size of $|N(\gamma')|$. First we consider those γ' for which $|N(\gamma')| \geq d^4$. By Lemma 9.1 we have

$$\begin{aligned} \sum_{g=d^4}^{(1-\alpha)m^n/2} \sum_{\gamma \in \mathcal{H}(g)} w(\gamma) \cdot e^{f(\gamma)+g(\gamma)} &\leq \sum_{g=d^4}^{(1-\alpha)m^n/2} e^{\xi g/(2\sqrt{d}\log^2 d)} \sum_{\gamma \in \mathcal{H}(g)} w(\gamma) \\ &\leq \sum_{g=d^4}^{\infty} e^{-\xi g/(2\sqrt{d}\log^2 d)} \end{aligned}$$

$$\leq \frac{1}{2d^3}\,. \tag{10.4}$$

for d sufficiently large.

Next we consider the contribution to the sum in (10.3) coming from polymers $\gamma \in \mathcal{H}(g)$ where $g \leq d^4$. To this end, we need the following bound on the weight of a polymer. Recall that for $S \subseteq V$ we let $\partial(S) = N(S)\backslash S$ and $S^+ = N(S) \cup S$.

Claim 10.1. *For $(A,B) \in \mathcal{D}_\lambda(H)$ and $\gamma \in \mathcal{P}$ we have*

$$w(\gamma) \leq (1-r)^{|\partial\gamma|} r^{-|\gamma|}\,.$$

Proof. By (3.2) we have

$$w(\gamma) = \frac{\sum_{f \in \hat{\chi}_{A,B}(\gamma)} \prod_{v \in \gamma^+} \lambda_{f(v)}}{\lambda_A^{|\gamma^+ \cap \mathcal{O}|} \lambda_B^{|\gamma^+ \cap \mathcal{E}|}}\,,$$

where $\hat{\chi}_{A,B}(\gamma)$ denotes the set of possible restrictions f_{γ^+} where $f \in \chi_{A,B}(\gamma)$.

Given any $v \in A^c$ there must exist a $w \in B$ such that $v \nsim w$ else $A \cup \{v\} \sim B$ contradicting the fact that (A,B) is a dominant pattern. Similarly for each $v \in B^c$ there is a $w \in A$ such that $v \nsim w$. It follows that

$$\sum_{f \in \hat{\chi}_{A,B}(\gamma)} \prod_{v \in \gamma^+} \lambda_{f(v)} \leq \lambda_{A^c}^{|\gamma \cap \mathcal{O}|} \left(\lambda_B - \min_{i \in B} \lambda_i\right)^{|\partial\gamma \cap \mathcal{E}|} \lambda_{B^c}^{|\gamma \cap \mathcal{E}|} \left(\lambda_A - \min_{i \in A} \lambda_i\right)^{|\partial\gamma \cap \mathcal{O}|}\,,$$

and so

$$w(\gamma) \leq \left(1 - \frac{\min_{i \in A} \lambda_i}{\lambda_A}\right)^{|\partial\gamma \cap \mathcal{O}|} \left(1 - \frac{\min_{i \in B} \lambda_i}{\lambda_B}\right)^{|\partial\gamma \cap \mathcal{E}|} \left(\frac{\lambda_{A^c}}{\lambda_A}\right)^{|\gamma \cap \mathcal{O}|} \left(\frac{\lambda_{B^c}}{\lambda_B}\right)^{|\gamma \cap \mathcal{E}|}\,. \tag{10.5}$$

The result follows by comparing the above quotients to r. □

If $|N(\gamma)| = g \leq d^4$ then it follows from Lemma 6.2 that $|\gamma| \leq g/(Cd)$. We then have by Claim 10.1

$$w(\gamma) \leq (1-r)^{g/2}\, r^{-g/(Cd)}\,,$$

for d sufficiently large.

Note that the graph G^2 has maximum degree at most d^2 and so by Lemma 7.3 the number of G^2-connected sets of size t containing a fixed vertex v is at most $(ed^2)^{t-1}$. Note also if $\gamma \neq \emptyset$ then $|N(\gamma)| \geq d$. It follows that

$$\sum_{g=d}^{d^4} \sum_{\substack{\gamma \in \mathcal{H}(g): \\ \gamma \ni v}} w(\gamma) \cdot e^{f(\gamma)+g(\gamma)} \leq \sum_{g=d}^{d^4} (e^2 d^2/r)^{g/(Cd)} \, (1-r)^{g/2} \, e^{gr/3} \leq \frac{1}{2d^3} \tag{10.6}$$

for d sufficiently large. By summing (10.4) and (10.6), we obtain (10.3) and hence also (10.2). This concludes the proof. □

We end this section by using Lemma 4.4 to establish strong tail bounds on the cluster expansion of $\ln \Xi_{A,B}$.

Recall that for a dominant pattern $(A,B) \in \mathcal{D}_\lambda(H)$

$$\delta_{A,B} := \max \left\{ \max_{v \in B^c} \frac{\lambda_{N(v)\cap A}}{\lambda_A}, \max_{u \in A^c} \frac{\lambda_{N(u)\cap B}}{\lambda_B} \right\},$$

and recall that $\delta_{A,B} < 1$ (Lemma 4.6). Recall also that for $k \geq 1$ we define

$$L_{A,B}(k) = \sum_{\substack{\Gamma \in \mathcal{C}: \\ \|\Gamma\| = k}} w_{A,B}(\Gamma).$$

The following lemma extends Lemma 4.5.

Lemma 10.2. *Let $(A,B) \in \mathcal{D}_\lambda(H)$ and let $\delta = \delta_{A,B}$.*

$$L_{A,B}(1) = \Theta(m^n \delta^d).$$

Moreover for $t \geq 1$ fixed

$$\sum_{k=t}^{\infty} |L_{A,B}(k)| = O(m^n d^{2(t-1)} \delta^{dt}),$$

and for $s \geq 0$ fixed

$$\sum_{k=1}^{\infty} k^s |L_{A,B}(k)| = (1+o(1)) L_{A,B}(1) \ \text{ as } n \to \infty. \tag{10.7}$$

In particular,

$$\ln \Xi_{A,B} = \sum_{j=1}^{\infty} L_{A,B}(j) = \Theta(m^n \delta^d). \tag{10.8}$$

Proof. Let us fix $(A,B) \in \mathcal{D}_\lambda(H)$. For ease of notation we let $L_k = L_{A,B}(k)$ and $w = w_{A,B}$.

The first claim follows by recalling the formula

$$L_1 = \left[\frac{1}{\lambda_A \lambda_B^d} \sum_{v \in A^c} \lambda_v \lambda_{N(v)\cap B}^d + \frac{1}{\lambda_B \lambda_A^d} \sum_{v \in B^c} \lambda_v \lambda_{N(v)\cap A}^d \right] \frac{m^n}{2}$$

and using the definition of δ (4.5).

Suppose that $k \in \mathbb{N}$ is fixed. We now bound L_k from above. Let Γ be a cluster with $\|\Gamma\| = k$. Since $V(\Gamma) := \bigcup_{S\in\Gamma} S$ is a G^2-connected set of size at most k, there are $O(m^n d^{2(k-1)})$ possibilities for $V(\Gamma)$ by Lemma 7.3. Given a set $X \subseteq V(Q_n)$ of size at most k, there are at most a constant number of clusters Γ of size k such that $V(\Gamma) = X$. The number of clusters of size k is therefore $O(m^n d^{2(k-1)})$.

Suppose now γ is a polymer with $|\gamma \cap \mathcal{O}| = i$ and $|\gamma \cap \mathcal{E}| = j$ where $i + j \leq k$. Using the fact that $|\partial(\gamma \cap \mathcal{O})| \geq di - O(1)$ and $|\partial(\gamma \cap \mathcal{E})| \geq dj - O(1)$ by Lemma 6.2 and (3.2), we have

$$\begin{aligned} w(\gamma) &= O\left(\frac{1}{\lambda_A^{dj} \lambda_B^{di}} \sum_{\substack{v_1,\dots,v_i \in A^c \\ u_1,\dots,u_j \in B^c}} \lambda_{N(u_1)\cap A}^d \cdots \lambda_{N(u_j)\cap A}^d \cdot \lambda_{N(v_1)\cap B}^d \cdots \lambda_{N(v_i)\cap B}^d \right) \\ &= O\left(\delta^{d(i+j)}\right), \end{aligned}$$

(this is a refinement of Claim 10.1 in the case where the size of the polymer is constant). Thus if Γ is a cluster of size k then $w(\Gamma) = O\left(\delta^{dk}\right)$ (note that k is fixed so that the Ursell function $\phi(\Gamma)$ is bounded by a constant). We therefore have

$$L_k = O(m^n d^{2(k-1)} \delta^{dk}). \tag{10.9}$$

We now show how the lemma follows from the following claim.

Claim 10.3. *For $s \geq 0$, $t \geq 1$ fixed, there exists a constant K such that for n sufficiently large,*

$$\sum_{k>K} k^s L_k \leq \delta^{td}.$$

Combining Claim 10.3 with (10.9) (for each $k \leq K$), we have

$$\sum_{k=1}^{\infty} k^s L_k = L_1 + \sum_{k=2}^{K} O\left(m^n d^{2(k-1)} \delta^{dk}\right) + O(\delta^{td}) = L_1 + O(m^n d^2 \delta^{2d}),$$

and also

$$\sum_{k=t}^{\infty} L_k = \sum_{k=t}^{K} O(m^n d^{2(k-1)} \delta^{dk}) + O(\delta^{td}) = O(m^n d^{2(t-1)} \delta^{dt}).$$

The lemma follows. It remains to establish Claim 10.3.

Proof of Claim 10.3. Having established the Kotecký-Preiss condition (Lemma 4.4) for our polymer model (with f, g as in (10.1)), we may apply Theorem 2.1. In particular, applying (2.2) where $\gamma = \{v\}$ is a single vertex polymer we have

$$\sum_{\substack{\Gamma \in \mathcal{C} \\ \Gamma \nsim \{v\}}} |w_{A,B}(\Gamma)|\, e^{g(\Gamma)} \leq 1/d\,. \tag{10.10}$$

Recall that we write $\Gamma \nsim \{v\}$ if there exists $\gamma \in \Gamma$ so that $\{v\} \nsim \gamma$. For each $\Gamma \in \mathcal{C}$, $\Gamma \nsim \{v\}$ for some $v \in V$ and so by summing (10.10) over all $v \in V$ we obtain

$$\sum_{\Gamma \in \mathcal{C}} |w_{A,B}(\Gamma)| e^{g(\Gamma)} \leq m^n/d\,,$$

(recall that $g(\Gamma) \coloneqq \sum_{\gamma \in \Gamma} g(\gamma)$). By (10.1), the definition of the function $g(\cdot)$, we have for any cluster Γ and n sufficiently large,

$$e^{g(\Gamma)/2} \geq \|\Gamma\|^s\,. \tag{10.11}$$

Moreover, there is a constant $K = K(m, H, \lambda)$ such that if Γ is a cluster of size $> K$, then for n large

$$e^{g(\Gamma)/2} \geq m^n \cdot \delta^{-td}\,. \tag{10.12}$$

Thus

$$\sum_{k > K} k^s L_k \leq \sum_{\substack{\Gamma \in \mathcal{C} \\ \|\Gamma\| > K}} |w(\Gamma)| e^{g(\Gamma)/2} \leq \delta^{td}\,,$$

where for the first inequality we used (10.11) and for the second inequality we used (10.10) and (10.12). □ □

11. Large deviations for polymer configurations

In this section we use the Kotecký-Preiss condition (Lemma 4.4) established in the previous section to prove a large deviation result for the total size of all polymers in a random sample from the polymer measure $\nu_{A,B}$ (as defined in (3.4)). As a corollary, we show that a typical sample from $\hat{\mu}_{H,\lambda}$ has very few defect vertices and is 'highly-balanced' in a sense that we will make precise later in the section.

In the next section we use these results to prove Lemmas 4.2 and 4.3 and hence also Theorem 1.2. The results of this section will also play a key role in Sections 13 and 18.

The proof of the following theorem and the proofs in Section 15, exploit a connection between the cluster expansion and cumulant generating functions. Recently Cannon and Perkins [9] made novel use of this connection to prove correlation decay results for the hard-core model on unbalanced bipartite graphs. Our treatment parallels that of the first author and Perkins [34] who prove a large deviation result and central limit theorems for the hard-core model on Q_n.

Suppose X is a random variable whose moment generating function $\mathbb{E}e^{tX}$ is defined for t in a neighbourhood of 0. We will use the *cumulant generating function* of X, defined as

$$h_t(X) = \ln \mathbb{E}e^{tX}\,, \tag{11.1}$$

that is, the logarithm of the moment generating function.

Theorem 11.1. *Let $(A,B) \in \mathcal{D}_\lambda(H)$ and let $\delta = \delta_{A,B}$ (as defined in (4.5)). Let $\mathbf{\Gamma}$ be a random configuration drawn from the distribution $\nu_{A,B}$. There exist constants $C, c > 0$, depending on (H,λ) and m, such that if $t \geq C \cdot d\delta^d$, then*

$$\mathbb{P}(\|\mathbf{\Gamma}\| \geq tm^n) \leq e^{-c\cdot tm^n/n}\,.$$

Proof. Let w denote $w_{A,B}$ and let Ξ denote $\Xi_{A,B}$. We introduce an auxiliary polymer model with modified polymer weights:

$$\tilde{w}(S) = w(S)e^{|S|/d}\,.$$

Let $\tilde{\Xi}$ be the associated polymer model partition function. Recall that Ω denotes the collection of all sets of mutually compatible polymers. We have

$$\frac{\tilde{\Xi}}{\Xi} = \frac{1}{\Xi}\sum_{\Gamma\in\Omega}\prod_{\gamma\in\Gamma}\tilde{w}(\gamma) = \frac{1}{\Xi}\sum_{\Gamma\in\Omega} e^{\|\Gamma\|/d}\prod_{\gamma\in\Gamma} w(\gamma) = \mathbb{E}e^{\|\mathbf{\Gamma}\|/d}$$

where $\mathbf{\Gamma}$ is a random polymer configuration drawn according to $\nu_{A,B}$ (the unmodified polymer model). In other words

$$h_{1/d}(\|\mathbf{\Gamma}\|) = \ln\tilde{\Xi} - \ln\Xi\,. \tag{11.2}$$

We claim that the cluster expansion of $\ln\tilde{\Xi}$ converges absolutely. Indeed let $v \in V$ be a fixed vertex. The sum in the Kotecký-Preiss condition (10.3) for the polymer model with modified weights $\tilde{w}(\cdot)$ with f and g as in (10.1) is

$$\sum_{\gamma\ni v}|\tilde{w}(\gamma)|e^{f(\gamma)+g(\gamma)} = \sum_{\gamma\ni v}|w(\gamma)|e^{f(\gamma)+g(\gamma)+|\gamma|/d}\,.$$

Returning to inequalities (10.4) and (10.6), we observe that they both equally hold with $f(\gamma)$ replaced by $f(\gamma)+|\gamma|/d$. We conclude that

$$\sum_{\gamma \ni v} |\tilde{w}(\gamma)| e^{f(\gamma)+g(\gamma)} \leq 1/d^3. \tag{11.3}$$

In other words we have verified the Kotecký-Preiss condition (10.3) for the modified polymer model and so the cluster expansion of $\ln \tilde{\Xi}$ converges absolutely. Using (11.2) and the bound $\Xi \geq 1$, then applying Lemma 10.2 (noting that the proofs remain valid with the modified weights), gives

$$\begin{aligned} h_{1/d}(\|\mathbf{\Gamma}\|) &\leq \ln \tilde{\Xi} \\ &\leq \sum_{\Gamma \in \mathcal{C}} |\tilde{w}(\Gamma)| \\ &= O(m^n \delta^d)\,. \end{aligned} \tag{11.4}$$

By Markov's inequality we have,

$$\begin{aligned} \mathbb{P}[\|\mathbf{\Gamma}\| > tm^n] &= \mathbb{P}\left[e^{\|\mathbf{\Gamma}\|/d} > e^{tm^n/d}\right] \\ &\leq e^{-tm^n/d} \mathbb{E} e^{|\mathbf{\Gamma}|/d}\,. \end{aligned}$$

Applying the bound (11.4) gives

$$\mathbb{P}[\|\mathbf{\Gamma}\| > tm^n] \leq \exp\left[-\frac{tm^n}{d} + O(m^n \delta^d)\right]\,.$$

The result follows. □

For $s > 0$, and weighted graph (H, λ), we say that a colouring $f \in \mathrm{Hom}(G, H)$ is (s, λ)-*balanced* with respect to (A, B) if for each $k \in A$, the proportion of vertices of $\mathcal{E}$ coloured k is within s of λ_k/λ_A and for each $\ell \in B$, the proportion of vertices of $\mathcal{O}$ coloured ℓ is within s of λ_ℓ/λ_B.

The following lemma shows that with high probability, a colouring sampled from the measure $\hat{\mu}_{H,\lambda}$ (as defined in Definition 3.4) is well-balanced with respect to some dominant pattern.

Recall that we call the set of vertices at which a colouring $f \in \mathrm{Hom}(\mathbb{Z}_m^n, H)$ differs from its closest dominant colouring (breaking ties arbitrarily if necessary) the *defect vertices* of f.

Lemma 11.2. *Let* $\mathbf{f}$ *denote a random element of* $\mathrm{Hom}(G, H)$ *sampled according to* $\hat{\mu}_{H,\lambda}$*, let* $\mathbf{D}$ *denote the random dominant pattern selected at Step 1 in the definition of* $\hat{\mu}_{H,\lambda}$ *(Definition 3.4) and let* $\mathbf{\Gamma}$ *be the random polymer configuration selected at Step 2. Let* X *denote the number of defect vertices of* $\mathbf{f}$*. The following hold:*

(i)

$$\mathbb{P}(X = \|\mathbf{\Gamma}\|) \geq 1 - e^{-\Omega(m^n/n^2)}\,.$$

(ii) If $(A,B) \in \mathcal{D}_\lambda(H)$ *and* $Cd\delta_{A,B}^d \leq t \leq 1/(4d)$*, with* C *as in Theorem 11.1, then*

$$\mathbb{P}(X \geq tm^n \mid \mathbf{D} = (A,B)) \leq e^{-\Omega(tm^n/n)}\,.$$

(iii) If $(A,B) \in \mathcal{D}_\lambda(H)$ *and* $s \geq 10Cd^2\delta_{A,B}^d$*, then*

$$\mathbb{P}(\mathbf{f} \text{ is } (s,\lambda)\text{-balanced wrt } (A,B) \mid \mathbf{D} = (A,B)) \geq 1 - e^{-\Omega(s^2m^n)} - e^{-\Omega(sm^n/n^2)}\,.$$

Proof. We begin by proving (i) and (ii) in conjunction. Let $(A,B) \in \mathcal{D}_\lambda(H)$ and $Cd\delta_{A,B}^d \leq t \leq 1/(4d)$. By Theorem 11.1, we have

$$\mathbb{P}(\|\mathbf{\Gamma}\| \leq tm^n \mid \mathbf{D} = (A,B)) \geq 1 - e^{-\Omega(tm^n/n)}\,. \tag{11.5}$$

Let $\mathbf{\Gamma}^+$ denote $\bigcup_{\gamma\in\mathbf{\Gamma}} \gamma^+$ and let $\|\mathbf{\Gamma}^+\| \coloneqq \sum_{\gamma\in\mathbf{\Gamma}} |\gamma^+|$.

For $k \in A$, let Z_k denote the number of vertices in $\mathcal{O}\backslash\mathbf{\Gamma}^+$ that receive colour k conditioned on the event $\mathbf{D} = (A,B)$. Since the vertices of $V\backslash\mathbf{\Gamma}^+$ are coloured at Step 4 in the definition of $\hat{\mu}_{H,\lambda}$ we see that

$$Z_k \sim \mathrm{Bin}(|\mathcal{O}\backslash\mathbf{\Gamma}^+|, \lambda_k/\lambda_A)\,.$$

For $k \in B$ we define Z_k similarly and note that $Z_k \sim \mathrm{Bin}(|\mathcal{E}\backslash\mathbf{\Gamma}^+|, \lambda_\ell/\lambda_B)$. Since $t \leq 1/(4d)$ and $\|\mathbf{\Gamma}^+\| \leq (d+1)\|\mathbf{\Gamma}\|$, if $\|\mathbf{\Gamma}\| \leq tm^n$, then $|\mathcal{O}\backslash\mathbf{\Gamma}^+|, |\mathcal{E}\backslash\mathbf{\Gamma}^+| \geq m^n/6$.

For $k \in A$, let Y_k denote the number of vertices of $\mathcal{O}$ that receive colour k in the colouring $\mathbf{f}$ conditioned on the event $\mathbf{D} = (A,B)$. Define Y_k similarly for $k \in B$. Note that for $k \in A \cup B$,

$$0 \leq Y_k - Z_k \leq \|\mathbf{\Gamma}^+\|\,. \tag{11.6}$$

By Chernoff's bound (applied to the random variables Z_k) and a union bound we then have

$$\mathbb{P}\left(\min_{k\in A\cup B} Y_k = \Omega(m^n) \;\middle|\; \|\mathbf{\Gamma}\| \leq tm^n, \mathbf{D} = (A,B)\right) \geq 1 - e^{-\Omega(m^n)}\,. \tag{11.7}$$

Suppose that the events $\mathbf{D} = (A,B)$, $\min_{k\in A\cup B} Y_k = \Omega(m^n)$, and $\|\mathbf{\Gamma}\| \leq tm^n$ all hold. If $(C,D) \in \mathcal{D}_\lambda(H)$ is a dominant pattern distinct from (A,B), then either $A\backslash C \neq \emptyset$ or $B\backslash D \neq \emptyset$ and so, since $\min_{k\in A\cup B} Y_k = \Omega(m^n)$, $\mathbf{f}$ must disagree with (C,D) on $\Omega(m^n)$ vertices. On the other hand, since $\mathbf{D} = (A,B)$, $\mathbf{f}$ disagrees with (A,B) on $\|\mathbf{\Gamma}\| \leq tm^n$ vertices. We conclude that (A,B) must be the closest dominant pattern to $\mathbf{f}$ and $X =$

$\|\mathbf{\Gamma}\|$. Conclusion (ii) therefore follows from (11.5) and (11.7) and (i) follows by taking $t = 1/(4d)$ in (11.5) and (11.7) for each $(A, B) \in \mathcal{D}_\lambda(H)$.

We now turn to (iii). By Theorem 11.1 again we have that if $s \geq 10C \cdot d^2\delta^d$, then

$$\mathbb{P}\left(\|\mathbf{\Gamma}^+\| \leq s\frac{|\mathcal{O}|}{4} \;\middle|\; \mathbf{D} = (A, B)\right) \geq 1 - e^{-\Omega(sm^n/n^2)}. \tag{11.8}$$

For $k \in A$, by (11.6) and Chernoff's bound (applied to Z_k), we also have that

$$\mathbb{P}\left(\left|\frac{Y_k}{|\mathcal{O}|} - \frac{\lambda_k}{\lambda_A}\right| \leq s \;\middle|\; \|\mathbf{\Gamma}^+\| \leq s\frac{|\mathcal{O}|}{4}, \mathbf{D} = (A, B)\right) \geq 1 - e^{-\Omega(s^2m^n)}. \tag{11.9}$$

The analogous statement holds for $k \in B$. Conclusion (iii) follows from (11.8), (11.9) and a union bound. □

12. Capturing by polymer models

In this section we prove Lemmas 4.2 and 4.3, which together show that almost all elements of $\mathrm{Hom}(G, H)$ are captured by precisely one polymer model.

In Section 4, we showed how Lemmas 4.2 and 4.3 imply both Lemma 4.1 and our main result Theorem 1.2. This will therefore conclude the proof of these results. We also showed how Lemmas 4.2, 4.3 and 4.5 imply Theorem 4.7 which provides asymptotic formulae for the partition function $Z_G^H(\lambda)$. Since we established Lemma 4.5 by proving Lemma 10.2, this will also conclude the proof of Theorem 4.7 (a refinement of Theorem 1.4).

We begin by recalling some terminology from Sections 3 and 4. Recall from Definition 3.1 that a polymer is a G^2-connected set γ such that $|N(\gamma \cap \mathcal{O})|, |N(\gamma \cap \mathcal{E})| < (1 - \alpha)m^n/2$ where $\alpha = \alpha(H, \lambda)$ will be specified later in the section (see (12.2)). We say a colouring $f \in \mathrm{Hom}(G, H)$ is captured by $(A, B) \in \mathcal{D}_\lambda(H)$ if each of the G^2-connected components of $(f^{-1}(A^c) \cap \mathcal{O}) \cup (f^{-1}(B^c) \cap \mathcal{E})$ is a polymer (cf. Definition 3.5). $\mathrm{Hom}_0(G, H)$, $\mathrm{Hom}_1(G, H)$, $\mathrm{Hom}_2(G, H)$ denote the sets of all colourings which are captured by 0, precisely one, and ≥ 2 dominant patterns respectively.

We first prove Lemma 4.3 which we restate for convenience. Recall the definition of $\tilde{Z}_G^H(\lambda)$ from (3.5).

Lemma 12.1.

$$Z_2 := \sum_{f \in \mathrm{Hom}_2(G,H)} \prod_{v \in V} \lambda_{f(v)} \leq e^{-\Omega(m^n/n^2)} \tilde{Z}_G^H(\lambda).$$

Proof. As before, let $\mathbf{f}$ denote a random element of $\mathrm{Hom}(H, G)$ selected according to $\hat{\mu}_{H,\lambda}$. Let $\mathbf{D}$ denote the random dominant pattern selected at Step 1 in the definition of $\hat{\mu}_{H,\lambda}$ (Definition 3.4).

Let us fix two distinct dominant patterns (A, B) and (C, D) (we may assume two dominant patterns exist else the conclusion of the lemma holds trivially). Let $\mathcal{F} \subseteq$

$\mathrm{Hom}(H,G)$ denote the subset of colourings that are captured by both (A,B) and (C,D). There exists a constant $s>0$ (depending only on (H,λ)) such that a colouring cannot be (s,λ)-balanced with respect to both (A,B) and (C,D). We may therefore partition $\mathcal{F}$ as $\mathcal{F}=\mathcal{F}_{A,B}\cup\mathcal{F}_{C,D}$ where $\mathcal{F}_{A,B}$ consists of those elements of $\mathcal{F}$ that are *not* (s,λ)-balanced with respect to (A,B) (and similarly for $\mathcal{F}_{C,D}$).

It follows from Lemma 3.6 (ii) and Lemma 11.2 (iii) that

$$\mathbb{P}(\mathbf{f}\in\mathcal{F}_{A,B}|\mathbf{D}=(A,B))=\frac{\sum_{f\in\mathcal{F}_{A,B}}\prod_{v\in V}\lambda_{f(v)}}{\eta_\lambda(H)^{m^n/2}\cdot\Xi_{A,B}}\leq e^{-\Omega(m^n/n^2)}\,.$$

Since there are only a constant number of pairs of dominant patterns we have

$$\sum_{f\in\mathrm{Hom}_2(G,H)}\prod_{v\in V}\lambda_{f(v)}\leq e^{-\Omega(m^n/n^2)}\eta_\lambda(H)^{m^n/2}\sum_{(A,B)\in\mathcal{D}_\lambda(H)}\Xi_{A,B}\,.\quad\square$$

We now prove Lemma 4.2 which is a more cumbersome task. To help in this endeavour we adapt an artful entropy argument of Engbers and Galvin [14]. We note that the proof of Lemma 4.2 does not rely on the convergence of the cluster expansion, and may therefore be considered as independent of Sections 6–11. Again we restate the lemma.

Lemma 12.2. *There exists $\zeta=\zeta(H,\lambda)<\eta_\lambda(H)$ so that*

$$Z_0:=\sum_{f\in\mathrm{Hom}_0(G,H)}\prod_{v\in V}\lambda_{f(v)}\leq\zeta^{m^n/2}\,.$$

Before giving the proof we require some preliminaries and notation. Let

$$V^*=\{x\in V:x_n=0,x\in\mathcal{E}\}\,.$$

For $v\in V$, and $0\leq i\leq m-1$, let

$$v^i:=v+(0,\dots,0,i)$$

and let

$$C(v):=\{v^0,\dots,v^{m-1}\}\,.$$

For $v\in V$, set

$$M_v:=N(v)\backslash\{v^1,v^{m-1}\}$$

and

$$M_{C(v)}:=\bigcup_{u\in C(v)}M_u\,.$$

We note that the subgraph of $\mathbb{Z}_m^n$ induced by $M_{C(v)}$ is a disjoint union of $2n-2$ cycles of length m (when $m \geq 4$) or $n-1$ disjoint edges (when $m = 2$).

To carry out a judicious application of Shearer's Lemma, Engbers and Galvin construct a partial order on V with certain desirable properties. We record one of these properties here in the form of a lemma.

Lemma 12.3 *(Engbers and Galvin [14]). There exists a partial order $\prec$ on V such that for all $v \in V^*$*

$$M_{C(v)} \subseteq \{u : u \prec C(v)\}.$$

We are now in a position to prove our lemma.

Proof of Lemma 12.2. We proceed with an entropy argument and blowup trick similar to that used in the proof of Lemma 9.2. First for each $i \in [q]$, let us fix a sequence of rationals $(\lambda_{i,k})_{k\in\mathbb{N}}$ such that $\lambda_{i,k} \to \lambda_i$. For each k we may write

$$(\lambda_{1,k}, \dots, \lambda_{q,k}) = (a_{1,k}/b_k, \dots, a_{q,k}/b_k)\,,$$

where $a_{i,k}, b_k \in \mathbb{N}$. Let $H_k := H(a_{1,k}, \dots, a_{q,k})$, that is the graph on vertex set $U_{1,k} \cup \dots \cup U_{q,k}$ where the $U_{i,k}$ are disjoint, $|U_{i,k}| = a_{i,k}$ for all i and $x \sim y$ if and only if $x \in U_{i,k}$, $y \in U_{j,k}$ for some pair $\{i,j\} \in E(H)$. For a vertex $v \in V(H_k)$, we let $\overline{v}$ denote the unique $i \in [q]$ for which $v \in U_{i,k}$. For $Y \subseteq V(H_k)$, we let $\overline{Y} := \{\overline{y} : y \in Y\}$. For convenience let us set $H_0 := H$.

For a vertex $v \in V^*$ and colouring $f \in \mathrm{Hom}(G, H_k)$, let

$$P_v(f) := (\overline{f(M_{v^0})}, \dots, \overline{f(M_{v^{m-1}})})\,.$$

We write $(A_0, \dots, A_{m-1})$ to indicate tuples of sets $A_i \subseteq V(H)$. Write $\mathrm{alt}(A,B)$ to denote the alternating tuple $(A, B, \dots, A, B)$. We say that $\mathcal{A} = (A_0, \dots, A_{m-1})$ is *ideal* if $\mathcal{A} = \mathrm{alt}(A,B)$ for some $(A,B) \in \mathcal{D}_\lambda(H)$. For $f \in \mathrm{Hom}(G, H_k)$, we call a vertex $u \in V^*$ *ideal* (wrt f) if $P_v(f)$ is ideal. For $X \subseteq V^*$, let $\mathrm{Hom}_X(G, H_k)$ denote the set of all $f \in \mathrm{Hom}(G, H_k)$ for which the set of ideal vertices of V^* is precisely X. Note that

$$\sum_{f \in \mathrm{Hom}_X(G,H)} \prod_{v \in V} \lambda_{f(v),k} = |\mathrm{Hom}_X(G, H_k)|/b_k^{m^n}\,. \tag{12.1}$$

For $\beta \in (0,1]$, let

$$\mathrm{Hom}^\beta(G, H_k) := \bigcup_{X \in \binom{V^*}{\leq \beta |V^*|}} \mathrm{Hom}_X(G, H_k)\,,$$

that is $\mathrm{Hom}^\beta(G, H_k)$ is the set of all $f \in \mathrm{Hom}(G, H_k)$ for which $\leq \beta |V^*|$ vertices of V^* are ideal.

Establishing the following claim will complete the proof.

Claim 12.4. *There exists $\zeta = \zeta(H,\lambda) < \eta_\lambda(H)$ and $\beta = \beta(H,\lambda) > 0$ such that for n sufficiently large*

$$\sum_{f\in \mathrm{Hom}^\beta(G,H)} \prod_{v\in V} \lambda_{f(v)} \le \zeta^{m^n/2}.$$

Indeed if $f \in \mathrm{Hom}(G,H)$ has $> \beta|V^*|$ ideal vertices in V^*, then by pigeonholing we can find a set $V^{**} \subseteq V^*$ such that $|V^{**}| \ge 4^{-q}\beta|V^*|$ and $P_v(f) = \mathrm{alt}(A,B)$ for all $v \in V^{**}$ for some fixed $(A,B) \in \mathcal{D}_\lambda(H)$. Letting $U = \bigcup_{v\in V^{**}} C(v)$, we have $|U\cap\mathcal{O}| = |U\cap\mathcal{E}| = m|V^{**}|/2 \ge 2^{-2q-1}\beta m^n/2$. Moreover f_{U^+} agrees with the pattern (A,B). We set

$$\alpha = \alpha(H,\lambda) \coloneqq 2^{-2q-1}\beta \tag{12.2}$$

and declare this to be the parameter appearing in Definition 3.1 of a polymer (a polymer is a G^2-connected subset $\gamma \subseteq V$ such that $|N(\gamma\cap\mathcal{E})|, |N(\gamma\cap\mathcal{O}))| < (1-\alpha)m^n/2$).

As in Definition 3.5, let $S(f) = (f^{-1}(A^c)\cap\mathcal{O}) \cup (f^{-1}(B^c)\cap\mathcal{E})$, the set of vertices at which f disagrees with (A,B). Since f_{U^+} agrees with the pattern (A,B), we have $S(f)\cap U^+ = \emptyset$ and so $S(f)^+\cap U = \emptyset$. By the choice of α, both $U\cap\mathcal{O}$ and $U\cap\mathcal{E}$ have size at least $\alpha m^n/2$ and so each of the G^2-connected components of $S(f)$ is a polymer i.e. f is captured by (A,B). We conclude that $\mathrm{Hom}_0(G,H) \subseteq \mathrm{Hom}^\beta(G,H)$ and the lemma follows. It remains to prove Claim 12.4.

Proof of Claim 12.4. Fix $X \in \binom{V^*}{\le\beta|V^*|}$ and choose $f \in \mathrm{Hom}_X(G,H_k)$ uniformly at random. Following [14] we will upper bound $H(f)$ by Shearer's Lemma (Lemma 5.2). We take as our covering family $\{M_{C(v)} : v \in V^*\}$ together with $(1+\mathbb{1}_{\{m>2\}})(n-1)$ copies of $C(v)$ for each $v \in V^*$. Each vertex of G is covered $(1+\mathbb{1}_{\{m>2\}})(n-1)$ times and so by Lemmas 5.2 and 12.3 we have

$$H(f) \le \sum_{v\in V^*} H(f_{C(v)}|f_{M_{C(v)}}) + \left(\frac{1+\mathbb{1}_{\{m=2\}}}{2n-2}\right)\sum_{v\in V^*} H(f_{M_{C(v)}}). \tag{12.3}$$

For a tuple $\mathcal{A} = (A_0,\dots,A_{m-1})$ where each $A_i \subseteq V(H)$, we let $z_k(\mathcal{A})$ denote the number of ways of choosing $(x_0,\dots,x_{m-1})$ with $x_i \in U_{y_i,k}$ where $y_i \in A_i$ for each i and $y_0 \sim \dots \sim y_{m-1} \sim y_0$. Let $z(\mathcal{A})$ denote the sum of $\prod_{i=0}^{m-1}\lambda_{x_i}$ over all tuples $(x_0,\dots,x_{m-1})$ with $x_i \in A_i$ for each i and with $x_0 \sim \dots \sim x_{m-1} \sim x_0$. Observe that as $k\to\infty$

$$z_k(\mathcal{A})/b_k^m \to z(\mathcal{A}) \tag{12.4}$$

for any tuple $\mathcal{A}$.

We let $n(\mathcal{A})$ denote the tuple $(n(A_0), \dots, n(A_{m-1}))$ where we recall that for $A \subseteq V(H)$, $n(A)$ denotes the common neighbourhood of A in H. We bound the terms in the first sum of (12.3) as follows:

$$\begin{aligned} H(f_{C(v)}|f_{M_{C(v)}}) &\leq H(f_{C(v)}|P_v(f)) \\ &\leq \sum_{\mathcal{A}} \mathbb{P}[P_v(f) = \mathcal{A}] H(f_{C(v)}|P_v(f) = \mathcal{A}) \\ &\leq \sum_{\mathcal{A}} \mathbb{P}[P_v(f) = \mathcal{A}] \log z_k(n(\mathcal{A}))\,, \end{aligned} \tag{12.5}$$

where the sum is over all tuples $\mathcal{A} = (A_0, \dots, A_{m-1})$ where each $A_i \subseteq V(H)$. The final inequality follows from the fact that $C(v)$ induces a cycle of length m in G and so if $P_v(f) = \mathcal{A}$ we must have for $i \in \{0, \dots, m-1\}$, $f(v^i) \in U_{y_i,k}$ for some $y_i \in n(A_i)$ where the y_i form a cycle in H.

To bound the terms in the second sum of (12.3) we write

$$\begin{aligned} H(f_{M_{C(v)}}) &= H(f_{M_{C(v)}}|P_v(f)) + H(P_v(f)) \\ &\leq \left(\frac{2n-2}{1+\mathbb{1}_{\{m=2\}}}\right) \sum_{\mathcal{A}} \mathbb{P}[P_v(f) = \mathcal{A}] \log z_k(\mathcal{A}) + qm\,. \end{aligned} \tag{12.6}$$

The inequality follows from the fact that $M_{C(v)}$ induces $2n-2$ cycles (when $m \geq 4$) or $n-1$ disjoint edges (when $m = 2$) and that there are at most 2^{qm} possible values that $P_v(f)$ can take. Summing the bounds (12.5) and (12.6) gives

$$H(f_{C(v)}|f_{M_{C(v)}}) + \left(\frac{1+\mathbb{1}_{\{m=2\}}}{2n-2}\right) H(f_{M_{C(v)}}) \leq \sum_{\mathcal{A}} \mathbb{P}[P_v(f) = \mathcal{A}] \log(z_k(\mathcal{A}) z_k(n(\mathcal{A})) + \frac{qm}{n-1}\,.$$

Summing this inequality over $v \in V^*$, recalling that the vertices of $X \subseteq V^*$ are ideal and the vertices of $V^* \backslash X$ are not, and using (12.3) we obtain

$$\begin{aligned} &H(f) \\ &\leq |X| \max_{\mathcal{A}} \log(z_k(\mathcal{A}) z_k(n(\mathcal{A})) + (|V^*| - |X|) \max_{\substack{\mathcal{A} \\ \text{not ideal}}} \log(z_k(\mathcal{A}) z_k(n(\mathcal{A}))) + \frac{qm}{n-1}|V^*| \\ &\leq \frac{m^{n-1}}{2} \left[\beta \max_{\mathcal{A}} \log(z_k(\mathcal{A}) z_k(n(\mathcal{A})) + (1-\beta) \max_{\substack{\mathcal{A} \\ \text{not ideal}}} \log(z_k(\mathcal{A}) z_k(n(\mathcal{A}))) + \frac{qm}{n-1}\right]. \end{aligned}$$

Subtracting $m^n \log b_k$ from both sides of the above inequality, taking the limit $k \to \infty$ and using (12.1) and (12.4) gives

$$\log\left(\sum_{f\in \mathrm{Hom}_X(G,H)}\prod_{v\in V}\lambda_{f(v)}\right) \tag{12.7}$$
$$\leq \frac{m^{n-1}}{2}\left[\beta\max_{\mathcal{A}}\log(z(\mathcal{A})z(n(\mathcal{A}))+(1-\beta)\max_{\substack{\mathcal{A}\\ \text{not ideal}}}\log(z(\mathcal{A})z(n(\mathcal{A}))+\frac{qm}{n-1}\right].$$

Now, for $\mathcal{A}=(A_0,\dots,A_{m-1})$ we have the inequality

$$z(\mathcal{A})\leq\prod_{i=0}^{m-1}\lambda_{A_i}$$

and so

$$z(\mathcal{A})z(n(\mathcal{A}))\leq\prod_{i=0}^{m-1}\lambda_{A_i}\lambda_{n(A_i)}\,. \tag{12.8}$$

Observe that for $A\subseteq V(H)$ we have $\lambda_A\lambda_{n(A)}\leq\eta_\lambda(H)$ with equality if and only if $(A,n(A))$ is a dominant pattern. Returning to (12.8) we see that

$$\max_{\mathcal{A}} z(\mathcal{A})z(n(\mathcal{A}))\leq\eta_\lambda(H)^m\,, \tag{12.9}$$

with equality only if $(A_i,n(A_i))\in\mathcal{D}_\lambda(H)$ for each i. We will now show that if $\mathcal{A}$ is not ideal then

$$z(\mathcal{A})z(n(\mathcal{A}))<\eta_\lambda(H)^m\,. \tag{12.10}$$

By (12.9) we may assume that $\mathcal{A}=(A_0,\dots,A_{m-1})$ where $(A_i,n(A_i))\in\mathcal{D}_\lambda(H)$ for all i. Observe that if $(A,B)\in\mathcal{D}_\lambda(H)$ then

$$n(B)=A \text{ and } n(A)=B\,. \tag{12.11}$$

Indeed it is clear that $A\subseteq n(B)$ and the inclusion cannot be strict else $(n(B),B)$ is a pattern with $\lambda_A\lambda_B<\lambda_{n(B)}\lambda_B$ contradicting the fact that (A,B) is dominant. The equality $n(A)=B$ follows similarly.

Since $\mathcal{A}$ is not ideal, we may assume without loss of generality that (A_0,A_1) is not a dominant pattern and so $A_1\neq n(A_0)$. For $A,B\subseteq V(H)$, let $p(A,B)$ be the set of pairs $(a,b)\in A\times B$ such that $a\nsim b$. We have

$$z(\mathcal{A})\leq\left(\lambda_{A_0}\lambda_{A_1}-\sum_{\{x,y\}\in p(A_0,A_1)}\lambda_x\lambda_y\right)\prod_{i=2}^{m-1}\lambda_{A_i}$$

and

$$z(n(\mathcal{A})) \leq \left(\lambda_{n(A_0)}\lambda_{n(A_1)} - \sum_{\{x,y\}\in p(n(A_0),n(A_1))} \lambda_x\lambda_y \right) \prod_{i=2}^{m-1} \lambda_{n(A_i)}\,.$$

If both $p(A_0, A_1)$ and $p(n(A_0), n(A_1))$ are empty then $A_0 \sim A_1$ and $n(A_0) \sim n(A_1)$ hence $A_1 \subseteq n(A_0)$ and $n(A_0) \subseteq n(n(A_1))$. Since $(A_1, n(A_1))$ is a dominant pattern we have $A_1 = n(n(A_1))$ by (12.11) and so $n(A_0) = A_1$ contrary to assumption. We conclude that one of $p(A_0, A_1)$ and $p(n(A_0), n(A_1))$ is non-empty and so $z(\mathcal{A})z(n(\mathcal{A})) < \eta_\lambda(H)^m$ which proves (12.10). Letting

$$\eta' := \left(\max_{\substack{\mathcal{A} \\ \text{not ideal}}} z(\mathcal{A})z(n(\mathcal{A})) \right)^{1/m} < \eta_\lambda(H)$$

and returning to (12.7) we have

$$\log\left(\sum_{f\in \mathrm{Hom}_X(G,H)} \prod_{v\in V} \lambda_{f(v)} \right) \leq \frac{m^n}{2}\left[\beta \log \eta_\lambda(H) + (1-\beta)\log\eta' + \frac{q}{n-1} \right]. \quad (12.12)$$

It follows from (12.12) that

$$\log\left(\sum_{f\in \mathrm{Hom}^\beta(G,H)} \prod_{v\in V} \lambda_{f(v)} \right) \leq \frac{m^n}{2}\left[\beta \log \eta_\lambda(H) + (1-\beta)\log\eta' + \frac{H(\beta)}{m} + \frac{q}{n-1} \right].$$

The claim follows by choosing $\beta = \beta(H, \lambda)$ sufficiently small and n sufficiently large. □ □

13. A structure theorem for Hom($\mathbb{Z}_m^n, H$)

In this section we show how the results proved thus far can be used to prove Theorem 1.3 and resolve conjectures of Engbers and Galvin [14, Conjectures 6.1, 6.2, 6.3].

Recall that for $s > 0$ and a weighted graph (H, λ), we say that an H-colouring of $\mathbb{Z}_m^n$ is (s, λ)-*balanced* with respect to (A, B) if for each $k \in A$, the proportion of vertices of $\mathcal{E}$ coloured k is within s of λ_k/λ_A and for each $\ell \in B$, the proportion of vertices of $\mathcal{O}$ coloured ℓ is within s of λ_ℓ/λ_B.

Theorem 1.3 was inspired by the following theorem of Engbers and Galvin.

Theorem 13.1 *(Engbers and Galvin [14]). Fix an even integer $m \geq 2$ and a weighted graph (H, λ) where λ takes only rational values. There is a partition of* $\mathrm{Hom}(\mathbb{Z}_m^n, H)$ *into $|\mathcal{D}_\lambda(H)| + 1$ classes*

$$\mathrm{Hom}(\mathbb{Z}_m^n, H) = F(0) \cup \bigcup_{(A,B)\in\mathcal{D}_\lambda(H)} F(A, B)$$

with the following properties (we write μ for $\mu_{H,\lambda}$).

1. $\mu(F(0)) \leq e^{-\Omega(n)}$.
2. *For $(A,B) \in \mathcal{D}_\lambda(H)$, each $f \in F(A,B)$ is $(e^{-\Omega(n)},\lambda)$-balanced w.r.t. (A,B).*
3. *If $A \neq B$ is such that $(A,B),(B,A) \in \mathcal{D}_\lambda(H)$ then*
$$\mu(F(A,B)) = \mu(F(B,A))\left(1 \pm e^{-\Omega(n)}\right).$$
4. *If $(A,B),(\tilde{A},\tilde{B}) \in \mathcal{D}_\lambda(H)$ are such that $\varphi(A) = \tilde{A}$ and $\varphi(B) = \tilde{B}$ for some weight preserving automorphism φ of (H,λ), then*
$$\mu(F(A,B)) = \mu(F(\tilde{A},\tilde{B}))\left(1 \pm e^{-\Omega(n)}\right).$$
5. *For each $(A,B) \in \mathcal{D}_\lambda(H)$, $x \in \mathcal{O}$, $y \in \mathcal{E}$, $k \in A$ and $\ell \in B$,*
$$\mathbb{P}_\mu(f(x) = k | f \in F(A,B)) = \frac{\lambda_k}{\lambda_A}\left(1 \pm e^{-\Omega(n)}\right)$$
and
$$\mathbb{P}_\mu(f(y) = \ell | f \in F(A,B)) = \frac{\lambda_\ell}{\lambda_B}\left(1 \pm e^{-\Omega(n)}\right).$$

Engbers and Galvin note that Theorem 13.1 does not make a general statement about the measures of the sets $F(A,B)$, however they conjecture an explicit formula for the asymptotics of $\ln \mu(F(A,B))$ [14, Conjectures 6.1, 6.2]. They go on to conjecture [14, Conjecture 6.3] that there is a decomposition of $\mathrm{Hom}(\mathbb{Z}_m^n, H)$ satisfying the conclusions of Theorem 13.1 such that $\mu_{H,\lambda}(F(0)) \leq e^{-m^n/p(n)}$ for some polynomial $p(n)$ whose degree depends only on (H,λ) and m. We remark that although Theorem 13.1 includes the restriction that λ takes rational values, both of the aforementioned conjectures are made for an arbitrary weighted graph (H,λ). The following theorem resolves these conjectures in a strong form.

We have the following strengthening of Theorem 1.3.

Theorem 13.2. *Fix a weighted graph (H,λ) and $m \geq 2$ even. There exists $\xi \in (0,1)$ such that if $\xi^d \leq s \leq 1$ then there is a partition of $\mathrm{Hom}(\mathbb{Z}_m^n, H)$ into $|\mathcal{D}_\lambda(H)| + 1$ classes*
$$\mathrm{Hom}(\mathbb{Z}_m^n, H) = F(0) \cup \bigcup_{(A,B)\in\mathcal{D}_\lambda(H)} F(A,B)$$
with the following properties. With $\mu = \mu_{H,\lambda}$ and $t := \min\{s^2, s/n^2\}$,

1. $\mu(F(0)) \leq e^{-\Omega(tm^n)}$.
2. *For $(A,B) \in \mathcal{D}_\lambda(H)$, each $f \in F(A,B)$ is (s,λ)-balanced with respect to (A,B).*

3. For each $(A,B) \in \mathcal{D}_\lambda(H)$,

$$\mu(F(A,B)) = \frac{1}{Z_G^H(\lambda)} \eta_\lambda(H)^{\frac{m^n}{2}} \Xi_{A,B} \left(1 \pm e^{-\Omega(tm^n)}\right) .$$

4. For each $(A,B) \in \mathcal{D}_\lambda(H)$, $x \in \mathcal{O}$, $y \in \mathcal{E}$, $k \in A$ *and* $\ell \in B$,

$$\mathbb{P}_\mu(f(x) = k | f \in F(A,B)) = \frac{\lambda_k}{\lambda_A} \left(1 \pm e^{-\Omega(n)}\right)$$

and

$$\mathbb{P}_\mu(f(y) = \ell | f \in F(A,B)) = \frac{\lambda_\ell}{\lambda_B} \left(1 \pm e^{-\Omega(n)}\right) .$$

Proof. Let $\mu, \hat{\mu}$ denote $\mu_{H,\lambda}, \hat{\mu}_{H,\lambda}$ respectively. Recall that $\mathrm{Hom}_0(G,H)$, $\mathrm{Hom}_1(G,H)$, $\mathrm{Hom}_2(G,H)$ denote the sets of all colourings which are captured (cf. Definition 3.5) by 0, precisely one, and ≥ 2 dominant patterns respectively. Let $E(0) = \mathrm{Hom}_0(G,H) \cup \mathrm{Hom}_2(G,H)$ and note that by Lemmas 12.1 and 12.2,

$$\hat{\mu}(E(0)) \leq e^{-\Omega(m^n/n^2)} . \tag{13.1}$$

We now consider the partition

$$\mathrm{Hom}_1(G,H) = \bigcup_{(A,B)\in\mathcal{D}_\lambda(H)} E(A,B) ,$$

where $E(A,B)$ denotes the set of colourings $f \in \mathrm{Hom}_1(G,H)$ which are captured by the (A,B) polymer model.

As in Lemma 12.1, we let

$$Z_2 = \sum_{f\in\mathrm{Hom}_2(G,H)} \prod_{v\in V} \lambda_{f(v)} .$$

Fix $(A,B) \in \mathcal{D}_\lambda(H)$. Since each element of $E(A,B)$ is captured by the (A,B) polymer model and no others we have by Lemma 3.6 (i) that

$$\frac{1}{\tilde{Z}_G^H(\lambda)} \left(\eta_\lambda(H)^{m^n/2} \Xi_{A,B} - 4^q Z_2\right) \leq \hat{\mu}(E(A,B)) \leq \frac{1}{\tilde{Z}_G^H(\lambda)} \eta_\lambda(H)^{m^n/2} \Xi_{A,B} . \tag{13.2}$$

Here 4^q is used as a crude upper bound for $|\mathcal{D}_\lambda(H)|$.

Let

$$\rho := \max_{(A,B)\in\mathcal{D}_\lambda(H)} \delta_{A,B} ,$$

with $\delta_{A,B}$ as defined in (4.5). Note that $\rho < 1$ since $\delta_{A,B} < 1$ for each (A, B) (Lemma 4.6). Pick

$$\max\{m^{-1/4}, \rho^{1/2}\} < \xi < 1\,. \tag{13.3}$$

Suppose that $s \geq \xi^d$. Let us fix $(A, B) \in \mathcal{D}_\lambda(H)$ and let $F(A, B)$ denote the set of elements of $E(A, B)$ that are (s, λ)-balanced with respect to (A, B). Let $t := \min\{s^2, s/n^2\}$. By Lemma 11.2 (iii) we have

$$\hat{\mu}(E(A, B)\backslash F(A, B)) \leq e^{-\Omega(tm^n)}\,. \tag{13.4}$$

Let $F(0) = E(0) \cup \bigcup_{(A,B)\in\mathcal{D}_\lambda(H)}(E(A, B)\backslash F(A, B))$, and consider the partition

$$\mathrm{Hom}(G, H) = F(0) \cup \bigcup_{(A,B)\in\mathcal{D}_\lambda(H)} F(A, B)\,.$$

By (13.1), (13.4) and Theorem 1.2, the above partition satisfies conclusions 1 and 2 of our theorem. It remains to verify conclusions 3 and 4. By (13.2), (13.4) and Lemma 12.1,

$$\hat{\mu}(F(A, B)) = \left(1 \pm e^{-\Omega(tm^n)}\right) \frac{1}{\tilde{Z}_G^H(\lambda)} \eta_\lambda(H)^{m^n/2} \Xi_{A,B}\,. \tag{13.5}$$

By Theorem 1.2, the same holds for $\mu(F(A, B))$ and so, applying Lemma 4.1, we obtain conclusion 3. Before we turn to conclusion 4, we use (13.5) to prove the following simple lower bound on $\hat{\mu}(F(A, B))$.

Claim 13.3.

$$\hat{\mu}(F(A, B)) \geq e^{-O(m^n\rho^d+q)}\,. \tag{13.6}$$

Proof of Claim 13.3. By Lemma 10.2,

$$\ln \Xi_{A,B} = O(m^n\delta_{A,B}^d)\,.$$

Using the crude bound $|\mathcal{D}_\lambda(H)| \leq 4^q$ we have

$$\tilde{Z}_G^H(\lambda) := \eta_\lambda(H)^{m^n/2} \sum_{(A,B)\in\mathcal{D}_\lambda(H)} \Xi_{A,B} \leq 4^q \eta_\lambda(H)^{m^n/2} e^{O(m^n\rho^d)}\,.$$

The claim follows from (13.5), noting that $\Xi_{A,B} \geq 1$ and $tm^n \to \infty$ by the definition of t. □

We now turn to conclusion 4. As before, let $\mathbf{D}$ denote the random pattern selected at Step 1 in the definition of $\hat{\mu}$ (Definition 3.4). First we bound the probability

$$\mathbb{P}_{\hat{\mu}}(f(x) = k \mid \mathbf{D} = (A, B)) .$$

Let $\mathbf{\Gamma}$ be the random polymer configuration selected at Step 2 in the definition of $\hat{\mu}$ conditioned on the event $\mathbf{D} = (A, B)$. In particular $\mathbf{\Gamma}$ has distribution $\nu_{A,B}$. Let $\mathbf{\Gamma}^+$ denote the union $\bigcup_{\gamma \in \mathbf{\Gamma}} \gamma^+$. By symmetry, $\mathbb{P}_{\hat{\mu}}(v \in \mathbf{\Gamma}^+)$ is the same for each $v \in \mathcal{O}$. Let us denote this probability by $p_{\mathcal{O}}$ and define $p_{\mathcal{E}}$ similarly. Then by Lemma 10.2 and (15.1) we have

$$(p_{\mathcal{O}} + p_{\mathcal{E}})m^n/2 = \mathbb{E}\left|\mathbf{\Gamma}^+\right| \leq (d+1)\mathbb{E}\left|\mathbf{\Gamma}\right| = O(dm^n \delta_{A,B}^d) ,$$

and so

$$p_{\mathcal{O}} + p_{\mathcal{E}} = e^{-\Omega(n)} .$$

It follows that

$$\begin{aligned} \mathbb{P}_{\hat{\mu}}\left(f(x) = k \mid \mathbf{D} = (A, B)\right) &= \mathbb{P}_{\hat{\mu}}\left(f(x) = k \mid \mathbf{D} = (A, B), x \notin \mathbf{\Gamma}^+\right) + e^{-\Omega(n)} \\ &= \lambda_k/\lambda_A + e^{-\Omega(n)} , \end{aligned} \tag{13.7}$$

where for the final equality we used that if $x \notin \mathbf{\Gamma}^+$, then x is coloured at Step 4 in the definition of $\hat{\mu}$.

By (13.1) and (13.4) the symmetric difference of the events $\{\mathbf{D} = (A, B)\}$ and $\{f \in F(A, B)\}$ has $\hat{\mu}$-measure $e^{-\Omega(tm^n)}$. By Claim 13.3,

$$\mathbb{P}_{\hat{\mu}}(f \in F(A, B)) \geq 4^{-q-1} e^{-O(m^n \rho^d)} .$$

By (13.3) we have $tm^n \to \infty$ and also $t/\rho^d \to \infty$. It then follows that

$$\mathbb{P}_{\hat{\mu}}\left(f(x) = k \mid f \in F(A, B)\right) = \mathbb{P}_{\hat{\mu}}\left(f(x) = k \mid \mathbf{D} = (A, B)\right)\left(1 \pm e^{-\Omega(tm^n)}\right)$$

and so by (13.7)

$$\mathbb{P}_{\hat{\mu}}\left(f(x) = k \mid f \in F(A, B)\right) = \lambda_k/\lambda_A + e^{-\Omega(n)} .$$

Finally using the fact that $\|\mu - \hat{\mu}\|_{TV} \leq e^{-\Omega(m^n/n^2)}$ (Theorem 1.2), conclusion 4 follows. □

We note that if $(A, B), (\tilde{A}, \tilde{B}) \in \mathcal{D}_\lambda(H)$ are such that $\varphi(A) = \tilde{A}$ and $\varphi(B) = \tilde{B}$ for some weight preserving automorphism φ of (H, λ), then $\Xi_{A,B} = \Xi_{\tilde{A},\tilde{B}}$. Similarly if $(A, B) \in \mathcal{D}_\lambda(H)$ where $A \neq B$ then $\Xi_{A,B} = \Xi_{B,A}$. Conclusion 3 of Theorem 13.2 therefore strengthens conclusions 3 and 4 of Theorem 13.1 significantly. Moreover by the cluster expansion $\ln \Xi_{A,B} = \sum_{j=1}^{\infty} L_{A,B}(j)$ and Lemma 10.2 we can obtain detailed

explicit asymptotic formulae for each $\mu(F(A,B))$ in Theorem 13.2 (see Section 16 for an algorithm to compute the terms $L_{A,B}(j)$).

As mentioned in the introduction, one may wonder to what extent the measures of the classes $F(A,B)$ in Theorem 13.2 can differ. We note that it is easy to construct a weighted graph (H,λ) with two dominant patterns $(A,B),(A',B')$ such that $\delta_{A',B'} < \delta_{A,B}$ (as defined at (4.5)). In particular, by (10.8) and Theorem 13.2,

$$\ln(\mu(F(A',B')/\mu(F(A,B)) \sim \ln(\Xi_{A',B'}/\Xi_{A,B}) \sim -m^n\delta_{A,B}^d\,.$$

We end this section by showing that the tradeoff in Theorem 13.2 between the size of $\mu(F(0))$ and the degree to which the elements of $F(A,B)$ are balanced (see conclusions 1 and 2) is essentially optimal.

Proposition 13.4. *Let ξ be as in Theorem 13.2, $\xi^d \le s = o(1)$, and let*

$$\mathrm{Hom}(\mathbb{Z}_m^n,H) = F(0) \cup \bigcup_{(A,B)\in\mathcal{D}_\lambda(H)} F(A,B)$$

be a partition of $\mathrm{Hom}(\mathbb{Z}_m^n,H)$ such that for $(A,B)\in\mathcal{D}_\lambda(H)$, each $f\in F(A,B)$ is (s,λ)-balanced with respect to (A,B). Then

$$\mu(F(0)) = e^{-O(s^2m^n)}\,.$$

Proof sketch. Fix $(A,B)\in\mathcal{D}_\lambda(H)$, and let F denote the set of $f\in\mathrm{Hom}(\mathbb{Z}_m^n,H)$ such that $f(\mathcal{O})\subseteq A$, $f(\mathcal{E})\subseteq B$ and f is *not* (s,λ)-balanced with respect to (A,B). Since $s=o(1)$, each $f\in F$ is not (s,λ)-balanced with respect to any dominant pattern (in fact we only require s to be less than a sufficiently small constant). Thus $F\subseteq F(0)$.

Consider the random experiment where we independently assign each $v\in\mathcal{O}$ the colour $k\in A$ with probability λ_k/λ_A and each $v\in\mathcal{E}$ the colour $\ell\in B$ with probability λ_ℓ/λ_B. By tightness of the Chernoff bound,

$$\frac{\sum_{f\in F}\prod_{v\in V}\lambda_{f(v)}}{\eta_\lambda(H)^{m^n/2}} = e^{-O(s^2m^n)}. \tag{13.8}$$

As before, let $\rho=\max_{(A,B)\in\mathcal{D}_\lambda(H)}\delta_{A,B}$. By Theorem 4.7 we then have

$$\mu(F(0)) \ge \mu(F) = \frac{\sum_{f\in F}\prod_{v\in V}\lambda_{f(v)}}{Z_G^H(\lambda)} = \frac{\sum_{f\in F}\prod_{v\in V}\lambda_{f(v)}}{\eta_\lambda(H)^{m^n/2}}\cdot e^{-O(m^n\rho^d+q)} = e^{-O(s^2m^n)},$$

where for the final equality we used (13.8) and $\xi > \max\{m^{-1/4},\rho^{1/2}\}$ (see (13.3)). □

14. Torpid mixing via conductance

In this section we show how Theorem 1.11 can be deduced from our decomposition result Theorem 13.2. We will make use of a well-known conductance argument, using a form of the argument presented in [13].

Let $\mathcal{M}$ be an ergodic (connected and aperiodic) Markov chain on a finite state space Ω, with transition probabilities $P(\omega,\omega')$, $\omega,\omega' \in \Omega$ and stationary distribution π.

Let $X \subseteq \Omega$ and $Y \subseteq \Omega\backslash X$ satisfy $\pi(X) \leq 1/2$ and $P(\omega,\omega') = 0$ for all $\omega \in X, \omega' \in \Omega\backslash(X \cup Y)$. Then from [13, Claim 2.3] we have

$$\tau_{\mathcal{M}} \geq \frac{\pi(X)}{8\pi(Y)}\,. \tag{14.1}$$

We will apply this bound when $\pi(Y)$ is small. Intuitively, Y acts as a bottleneck: for the chain to leave the set X it must pass through Y. If $\pi(Y)$ is small, the chain is unlikely to pass through Y, which drives up the mixing time.

Proof of Theorem 1.11. We use the conductance argument outlined above. Let $\mu = \mu_{H,\lambda}$ and let

$$\beta := \min_{(A,B)\in\mathcal{D}_\lambda(H)} \min\left\{\frac{\min_{k\in A}\lambda_k}{\lambda_A}, \frac{\min_{\ell\in B}\lambda_\ell}{\lambda_B}\right\}.$$

Consider the decomposition from Theorem 13.2 with $s = \beta/(2q)$

$$\mathrm{Hom}(\mathbb{Z}_m^n, H) = F(0) \cup \bigcup_{(A,B)\in\mathcal{D}_\lambda(H)} F(A,B)\,.$$

Suppose that (H,λ) is non-bipartite and non-trivial and so has at least two dominant patterns. Suppose also that (A,B) is such that $\mu(F(A,B))$ is minimal and so in particular $\mu(F(A,B)) \leq 1/2$. Let $X = F(A,B)$ and $Y = F(0)$. By (13.6) and Theorem 1.2 we also have

$$\mu(X) \geq e^{-O(m^n\rho^d+q)}\,, \tag{14.2}$$

where $\rho = \max_{(A,B)\in\mathcal{D}_\lambda(H)} \delta_{A,B} < 1$.

Suppose now that (C,D) is a dominant pattern distinct from (A,B) and note that either $A\backslash C \neq \emptyset$ or $B\backslash D \neq \emptyset$. Without loss of generality suppose $A\backslash C \neq \emptyset$. Let $f_1 \in F(A,B)$, $f_2 \in F(C,D)$. By conclusion 2 of Theorem 13.2, f_1 is (s,λ)-balanced with respect to (A,B) and f_2 is (s,λ)-balanced with respect to (C,D). Letting $k \in A\backslash C$ we then have

$$|f_1^{-1}(k)\backslash f_2^{-1}(k)| \geq (\beta - s - |C|s)\frac{m^n}{2} \geq \frac{\beta m^n}{4}$$

and so f_1 and f_2 differ in at least $\beta m^n/4$ vertices.

Now let $\mathcal{M}$ be a $(\beta/5)$-local ergodic Markov chain on $\mathrm{Hom}(\mathbb{Z}_m^n, H)$ with transition probabilities $P(f_1, f_2)$, $f_1, f_2 \in \mathrm{Hom}(\mathbb{Z}_m^n, H)$ and stationary distribution μ. By the above, $P(f_1, f_2) = 0$ for all $f_1 \in X$, $f_2 \in (X \cup Y)^c$ (i.e. for the chain to leave X it must pass through Y). By conclusion 1 of Theorem 13.2, the conductance bound (14.1) and (14.2), we therefore have

$$\tau_{\mathcal{M}} \geq \frac{\mu(X)}{8\mu(Y)} = e^{\Omega(m^n/n^2)}.$$

The proof of the second statement of the theorem, dealing with bipartite (H, λ), goes through with trivial modifications and so we omit it. □

We end this section by remarking that the assumption of ergodicity in Theorem 1.11 is a non-trivial restriction. We illustrate this with the example where $H = K_4$, i.e. 4-colourings of $\mathbb{Z}_m^n$. Let $\mathcal{H}$ denote the graph on vertex set $\mathrm{Hom}(\mathbb{Z}_m^n, K_4)$ where two vertices f_1, f_2 are adjacent if and only $f_1(v) \neq f_2(v)$ for precisely one vertex $v \in V(\mathbb{Z}_m^n)$. It is well-known that $\mathcal{H}$ has isolated vertices for any $m \geq 2$ (these are referred to as *frozen colourings*: updating the colour of any single vertex results in a non-proper colouring). One can use Theorem 1.2 to show that the component structure of $\mathcal{H}$ undergoes a phase transition at $m = 256$ in the following sense: If $m < 256$ is even then $\mathcal{H}$ has a giant connected component occupying $1 - o(1)$ of the vertices whereas if $m > 256$ the largest component of $\mathcal{H}$ occupies an exponentially small fraction of the total number of vertices. If $m = 256$, then the largest component in $\mathcal{H}$ is polynomially small. This transition is driven by the appearance of copies of the three dimensional hypercube Q_3 which is 4-coloured in such a way that no single vertex update is valid (i.e. a frozen colouring of Q_3). We omit the details for brevity.

15. The defect distribution

In this section we prove our central limit theorem for the distribution of the number of vertices in a sample from the defect distributions $\nu_{A,B}$ (Theorem 1.7).

We begin with some preliminaries on cumulants of random variables. Recall from (15.1), the cumulant generating function of a random variable X, $h_t(X) = \ln \mathbb{E}e^{tX}$. The ℓth *cumulant* of X is defined by taking derivatives of $h_t(X)$ and evaluating at 0:

$$\kappa_\ell(X) = \left.\frac{\partial^\ell h_t(X)}{\partial t^\ell}\right|_{t=0}.$$

In fact the cumulants of X are related to the moments of X by a non-linear change of basis (see e.g. [41]). In particular, $\kappa_1(X) = \mathbb{E}X$ and $\kappa_2(X) = \mathrm{var}(X)$. Moreover, if a random variable X has a distribution determined by its moments, and if for a sequence of random variables X_n we have $\lim_{n\to\infty} \kappa_\ell(X_n) = \kappa_\ell(X)$ for all $\ell \geq 1$, then $X_n \xrightarrow{d} X$

(recall that $\xrightarrow{d}$ denotes convergence in distribution). We will use this in conjunction with the following fact.

Fact 1. If X has a Poisson distribution with mean m, then $\kappa_r(X) = m$ for all r. If X has a standard normal distribution (mean 0, variance 1) then $\kappa_1(X) = 0$, $\kappa_2(X) = 1$, and $\kappa_\ell(X) = 0$ for all $\ell \geq 3$.

We will also use the following basic property of cumulants.

Fact 2. If X is a random variable, $a, b \in \mathbb{R}$ and $\ell \geq 2$, then $\kappa_\ell(aX + b) = a^\ell \kappa_\ell(X)$.

We are now in a position to prove Theorem 1.7.

Proof of Theorem 1.7. For brevity let us denote the weight function $w_{A,B}$ (see (3.1)) by w and denote the partition function $\Xi_{A,B}$ (see (3.3)) by Ξ. For $t \geq 0$, let us define the modified weight function

$$w_t(\gamma) = w(\gamma)e^{t|\gamma|}$$

and let Ξ_t denote the accompanying modified partition function. We then have

$$h_t(\|\mathbf{\Gamma}\|) = \log \Xi_t - \log \Xi \,.$$

As was shown in the proof of Theorem 11.1, the cluster expansion of $\log \Xi_t$ converges absolutely for $t \leq 1/d$ (see inequality (11.3)).

$$\log \Xi_t = \sum_{\Gamma \in \mathcal{C}} w_t(\Gamma) = \sum_{\Gamma \in \mathcal{C}} w(\Gamma) e^{t\|\mathbf{\Gamma}\|} \,.$$

It follows that for fixed $s \in \mathbb{N}$

$$\begin{aligned} \kappa_s(\|\mathbf{\Gamma}\|) &= \left. \frac{\partial^s \log \Xi_t}{\partial t^s} \right|_{t=0} \qquad (15.1) \\ &= \sum_{\Gamma \in \mathcal{C}} \|\mathbf{\Gamma}\|^s w(\Gamma) \\ &= \sum_{k=1}^{\infty} k^s L_{A,B}(k) \\ &= (1 + o(1)) L_{A,B}(1) \,, \qquad (15.2) \end{aligned}$$

where for the last inequality we used (10.7) from Lemma 10.2.

Let $L_1 = L_{A,B}(1)$. If $L_1 \to 0$, then $\kappa_\ell(\|\mathbf{\Gamma}\|) \to 0$ for each ℓ so that $\|\mathbf{\Gamma}\| = 0$ whp. If $L_1 \to \rho > 0$ then $\kappa_\ell(\|\mathbf{\Gamma}\|) \to \rho$ for each ℓ so that $\|\mathbf{\Gamma}\| \xrightarrow{d} \mathrm{Pois}(\rho)$. Finally suppose

that $L_1 \to \infty$ and let Y denote $(\|\mathbf{\Gamma}\| - \mathbb{E}\|\mathbf{\Gamma}\|)/\sqrt{\mathrm{var}\|\mathbf{\Gamma}\|}$. We have $\kappa_1(Y) = \mathbb{E}Y = 0$. By Fact 2 and (15.2) we have, for $\ell \geq 2$,

$$\kappa_\ell(Y) = \frac{\kappa_\ell(\|\mathbf{\Gamma}\|)}{(\mathrm{var}\|\mathbf{\Gamma}\|)^{\ell/2}} = \frac{\kappa_\ell(\|\mathbf{\Gamma}\|)}{\kappa_2(\|\mathbf{\Gamma}\|)^{\ell/2}} = (1+o(1))L_{A,B}(1)^{1-\ell/2}$$

so that $\kappa_2(Y) \to 1$ and $\kappa_\ell(Y) \to 0$ for $\ell \geq 3$. We conclude that $Y \xrightarrow{d} N(0,1)$ by Fact 1. □

Proof of Corollary 1.8. We specialise to the q-colouring model so that $(H,\lambda) = (K_q, i)$ where $i \equiv 1$. We denote $\hat{\mu}_{H,\lambda}, \mu_{H,\lambda}$ by $\hat{\mu}, \mu$ respectively. Let $\hat{X}$ denote the number of defect vertices in a sample from $\hat{\mu}$. Note that since $\|\mu - \hat{\mu}\| \leq e^{-\Omega(m^n/n^2)}$ by Theorem 1.2, it suffices to establish the statement of the corollary with $\hat{\mu}$ and $\hat{X}$ in place of μ and X.

Let $\mathbf{\Gamma}$ denote the random polymer configuration selected at Step 2 in the definition of $\hat{\mu}$. By Lemma 11.2 (i) we have that $\hat{X} = \|\mathbf{\Gamma}\|$ with probability $\geq 1 - e^{-\Omega(m^n/n^2)}$. The result follows from Theorem 1.7 and the symmetry of the dominant patterns in (K_q, i). □

16. Computing terms of the cluster expansion

In this section we give a general description of the terms in the cluster expansion of $\Xi_{A,B}$ and provide an algorithm for computing these terms as an explicit function of n. Combined with Theorem 1.4, this provides, for fixed (H,λ) and $m \geq 2$ even, a finite time algorithm to compute an asymptotic formula (correct up to a multiplicative factor $(1+o(1))$) for the partition function $Z_G^H(\lambda)$ (such as those presented in Corollary 1.6).

For the following lemma we recall that for $(A,B) \in \mathcal{D}_\lambda(H)$

$$L_{A,B}(k) = \sum_{\substack{\Gamma \in \mathcal{C}: \\ \|\Gamma\| = k}} w_{A,B}(\Gamma)$$

and

$$\delta_{A,B} = \max\left\{ \max_{v \in B^c} \frac{\lambda_{N(v)\cap A}}{\lambda_A}, \max_{u \in A^c} \frac{\lambda_{N(u)\cap B}}{\lambda_B} \right\}. \tag{16.1}$$

Lemma 16.1. *Fix a weighted graph (H,λ), a dominant pattern $(A,B) \in \mathcal{D}_\lambda(H)$, $m \geq 2$ even and $k \in \mathbb{N}$. We have*

$$L_{A,B}(k) = m^n \sum_{i \in I} p_i(n) \cdot \alpha_i^n,$$

where I is an index set of size at most $e^{O(k \log k)}$ and for each $i \in I$, $\alpha_i \in [0, \delta_{A,B}^k)$ depends only on (H,λ) and p_i is a polynomial of degree at most $2(k-1)$ whose coefficients

depend only on (H, λ). Moreover the α_i and the coefficients of the polynomials p_i can all be computed in time $e^{O(k \log k)}$.

Proof. For $v \in V$ we say $i \in [n]$ is an *active coordinate* of v if $v_i \neq 0$. For $S \subseteq V$ we define the active coordinates of S to be the union of the sets of active coordinates of the elements of S. For a cluster Γ, we define the active coordinates of Γ to be the active coordinates of the set $V(\Gamma) = \bigcup_{S \in \Gamma} S$.

For $j, k, a \in \mathbb{N}$, let $\mathcal{G}_{j,k,a}$ denote the set of all clusters Γ containing $\mathbf{0}$ with $\|\Gamma\| = k$ and $|V(\Gamma)| = j$ and whose set of active coordinates is precisely $[a]$. Note that if $S \subseteq V$ is a G^2-connected set of size j containing $\mathbf{0}$ then S has at most $2(j-1)$ active coordinates. Therefore, if $\mathcal{G}_{j,k,a} \neq \emptyset$ we must have $a \leq 2(j-1)$.

By symmetry of coordinates and vertex transitivity of $\mathbb{Z}_m^n$ we have

$$L_{A,B}(k) = m^n \sum_{j=1}^{k} \frac{1}{j} \sum_{a=1}^{2(j-1)} \binom{n}{a} \sum_{\Gamma \in \mathcal{G}_{j,k,a}} w_{A,B}(\Gamma)\,. \tag{16.2}$$

We first show how to efficiently generate the sets $\mathcal{G}_{j,k,a}$, then we show how to efficiently compute the weights $w_{A,B}(\Gamma)$.

Claim 16.2. *For $k \in \mathbb{N}$, $j \in [k]$, and $a \in [2k]$ the set $\mathcal{G}_{j,k,a}$ has size $e^{O(k \log k)}$ and can be generated in time $e^{O(k \log k)}$.*

Proof of Claim 16.2. First we construct the list $\mathcal{L}_j$ of all G^2-connected subsets of V of size j which contain $\mathbf{0} = (0, \dots, 0)$ and whose set of active coordinates are a subset of $[2k]$. We do so iteratively. Let $1 \leq t < j$ and suppose we have constructed the list $\mathcal{L}_t$. For each $S \in \mathcal{L}_t$ and $v \in S$, we run through the list of vertices w at distance ≤ 2 from v such that the active coordinates of w are a subset of $[2k]$, and we add $S \cup \{w\}$ to the list $\mathcal{L}_{t+1}$ if $w \notin S$. Note that there are at most $4\binom{2k}{2} + 2k \leq 8k^2$ choices for w. This procedure generates the whole list $\mathcal{L}_{t+1}$ and shows that $|\mathcal{L}_{t+1}| \leq 8tk^2|\mathcal{L}_t|$ and so $|\mathcal{L}_j| \leq j!(8k^2)^j = e^{O(k \log k)}$.

Let $\mathcal{L}_j^a$ denote the subset of $\mathcal{L}_j$ consisting of those sets whose active coordinates are precisely $[a]$. Note that we can generate the list $\mathcal{L}_j^a$ in time $e^{O(k \log k)}$ by checking the elements of $\mathcal{L}_j$ one by one.

We now generate the list $\mathcal{G}_{j,k,a}$. To do so we run through each $S \in \mathcal{L}_j^a$ and create the list of clusters Γ with $\|\Gamma\| = k$ and $V(\Gamma) = S$. We claim that this can be done in time $e^{O(k \log k)}$. Recall that a cluster Γ with $\|\Gamma\| = k$ is an ordered set of polymers $\Gamma = (\gamma_1, \dots, \gamma_\ell)$ such that $\sum_{i=1}^{\ell} |\gamma_i| = k$. Let us fix $S \in \mathcal{L}_j^a$. Since there are at most 2^k ordered integer partitions of k, it suffices to show that for a fixed such partition $(j_1, \dots, j_\ell)$ (so that $\sum_i j_i = k$) we may find, in time $e^{O(k \log k)}$, all clusters $(\gamma_1, \dots, \gamma_\ell)$ for which $|\gamma_i| = j_i$ for all i, and $\bigcup_i \gamma_i = S$. To do this we can simply check each element of $\binom{S}{j_1} \times \dots \times \binom{S}{j_\ell}$ (a set of size at most $e^{O(k \log k)}$) to see if it constitutes a legitimate cluster. $\square$

We now turn to computing the weights $w_{A,B}(\Gamma)$ in (16.2). For $S \subseteq V$, let $\tilde{\chi}_{A,B}(S)$ denote the set of all colourings of $G[S]$ which disagree with (A, B) at every vertex of S.

Claim 16.3. *If $S \subseteq V$ is a G^2-connected set of size at most k whose active coordinates lie in $[2k]$, then we can write*

$$w_{A,B}(S) = \sum_{f \in \tilde{\chi}_{A,B}(S)} \beta_f \cdot \alpha_f^d$$

where $0 \leq \alpha_f < \delta_{A,B}^{|S|}$, $\beta_f \geq 0$ and both can be computed in time polynomial in k.

Note that if $|S| \leq k$, then $|\tilde{\chi}_{A,B}(S)| = e^{O(k)}$ and we can list the elements of $\tilde{\chi}_{A,B}(S)$ by brute force in $e^{O(k)}$ time.

We recall that for a cluster Γ, $w_{A,B}(\Gamma) := \phi(I_\gamma) \prod_{\gamma \in \Gamma} w_{A,B}(\gamma)$ where $\phi(I_\gamma)$ is the Ursell function of the graph I_γ (as defined in (1.8)). The Ursell function $\phi(I_\gamma)$ is an evaluation of the Tutte polynomial of I_γ and therefore can be computed in time $e^{O(k)}$ by an algorithm of Björklund, Husfeldt, Kaski, and Koivisto [5, Theorem 1].

The lemma therefore follows from (16.2) and Claims 16.2 and 16.3. It remains to prove Claim 16.3.

Proof of Claim 16.3. Recall from (3.2) that

$$w_{A,B}(S) = \frac{\sum_{f \in \hat{\chi}_{A,B}(S)} \prod_{v \in S^+} \lambda_{f(v)}}{\lambda_A^{|S^+ \cap \mathcal{O}|} \lambda_B^{|S^+ \cap \mathcal{E}|}}, \tag{16.3}$$

where $\hat{\chi}_{A,B}(S)$ is the set of all colourings of $G[S^+]$ which disagree with (A, B) precisely on S.

For $f \in \tilde{\chi}_{A,B}(S)$ and $v \in \partial S \cap \mathcal{O}$, let $A(f, v)$ denote the subset of A available to v given the colouring f of $G[S]$. More precisely $A(f, v) = A \cap \bigcap_{u \in N_G(v) \cap S} N_H(f(u))$. For $v \in \partial S \cap \mathcal{E}$, define $B(f, v)$ similarly. Then

$$\sum_{f \in \hat{\chi}_{A,B}(S)} \prod_{v \in S^+} \lambda_{f(v)} = \sum_{f \in \tilde{\chi}_{A,B}(S)} \prod_{v \in S} \lambda_{f(v)} \prod_{u \in \partial S \cap \mathcal{O}} \lambda_{A(f,u)} \prod_{w \in \partial S \cap \mathcal{E}} \lambda_{B(f,w)}.$$

Let $\bar{S}$ be the set of vertices in V which are either contained in S or have ≥ 2 neighbours in S. By grouping elements of $\partial S \backslash \bar{S}$ according to their unique neighbour in S we may rewrite the right hand side as

$$\sum_{f \in \tilde{\chi}_{A,B}(S)} \prod_{v \in S} \lambda_{f(v)} \prod_{u \in \bar{S} \cap \partial S \cap \mathcal{O}} \lambda_{A(f,u)} \prod_{w \in \bar{S} \cap \partial S \cap \mathcal{E}} \lambda_{B(f,w)} \prod_{x \in S \cap \mathcal{E}} \lambda_{A \cap N(f(x))}^{d - d_{\bar{S}}(x)} \prod_{y \in S \cap \mathcal{O}} \lambda_{B \cap N(f(y))}^{d - d_{\bar{S}}(y)}.$$

We now consider the denominator in (16.3). Note that

$$|S^+ \cap \mathcal{O}| = |\bar{S} \cap \mathcal{O}| + \sum_{v \in S \cap \mathcal{E}} (d - d_{\bar{S}}(v)),$$

and similarly for $|S^+ \cap \mathcal{E}|$. Thus

$$w_{A,B}(S) = \sum_{f \in \tilde{\chi}_{A,B}(S)} \left[\frac{1}{\lambda_A^{|\bar{S} \cap \mathcal{O}|} \lambda_B^{|\bar{S} \cap \mathcal{E}|}} \prod_{v \in S} \lambda_{f(v)} \prod_{u \in \bar{S} \cap \partial S \cap \mathcal{O}} \lambda_{A(f,u)} \prod_{w \in \bar{S} \cap \partial S \cap \mathcal{E}} \lambda_{B(f,w)} \right.$$
$$\left. \prod_{x \in S \cap \mathcal{E}} \left(\frac{\lambda_{A \cap N(f(x))}}{\lambda_A} \right)^{d - d_{\bar{S}}(x)} \prod_{y \in S \cap \mathcal{O}} \left(\frac{\lambda_{B \cap N(f(y))}}{\lambda_B} \right)^{d - d_{\bar{S}}(y)} \right].$$

For each $f \in \tilde{\chi}_{A,B}(S)$ the corresponding term in the above sum takes the form $\beta_f \cdot \alpha_f^d$ where $\alpha_f < \delta_{A,B}^{|S|}$ by (16.1), the definition of $\delta_{A,B}$. We now show that the set $\bar{S}$ can be generated in time polynomial in k and that $|\bar{S}| \leq k^2$ so that α_f, β_f can be computed in time polynomial in k.

Since $\mathbb{Z}_m^n$ has maximum codegree 2, by considering the joint neighbourhood of each pair of elements in S, we have $|\bar{S}| \leq k + 2\binom{k}{2} = k^2$. Now let $u, v \in S$ be distinct. Since the active coordinates of S are contained in $[2k]$, each element of $N(u) \cap N(v)$ must differ from u and v in a coordinate in $[2k]$. We may therefore generate the set $N(u) \cap N(v)$, and hence also $\bar{S}$, in time polynomial in k. □ □

17. Counting proper q-colourings

In the previous section we gave a general description of the terms of the cluster expansion of the log partition functions $\ln \Xi_{A,B}$ along with an algorithm for how to compute these terms. In this section we specialise to the case of the uniform measure on q-colourings and present a more detailed picture of the cluster expansion. In particular, we prove a strengthening of Theorem 1.5 and prove Corollary 1.6, thus resolving [38, Conjecture 5.2].

Throughout this section we let $(H, \lambda) = (K_q, i)$ where $i \equiv 1$. We let $V(K_q) = [q]$. The dominant patterns of (K_q, i) are the pairs (A, B) where $A \cup B = [q]$ and $\{|A|, |B|\} = \{\lceil q/2 \rceil, \lfloor q/2 \rfloor\}$.

17.1. Typical polymers

Define the *type* of a G^2-connected set $S \subseteq V$ to be the isomorphism class of the induced subgraph $G^2[S]$. For any fixed type T, we denote by n_T the number of G^2-connected sets of type T. Let us call a set $S \subseteq V$ *typical* if it is G^2-connected, its type is a tree, S is a subset of one of the partition classes $\mathcal{E}, \mathcal{O}$ and for every edge $\{u, v\}$ in $G^2[S]$, the codegree of u and v in G is 2 (we note that this last condition is redundant in the case where $m = 2$). The reason for considering such sets is that almost all G^2-connected sets of a fixed size are typical and the weights of these sets take a particularly simple form (see Lemma 17.2). For a tree T, let n'_T denote the number of typical sets of type T.

Lemma 17.1. *Let $t \in \mathbb{N}$ and let T be a type with t vertices. If T is a tree then*

$$n'_T = (c_T + o(1))m^n n^{2t-2} \text{ and } n_T - n'_T = O(m^n n^{2t-3})$$

where $c_T = \left(\frac{1+3\mathbb{1}_{m>2}}{2}\right)^{t-1} |\mathrm{Aut}(T)|^{-1}$. If T is not a tree then

$$n_T = O(m^n n^{2t-3}).$$

Proof. Since T is a connected graph we may fix an ordering $(x_1, \dots, x_t)$ of the vertices of T so that $T_i := T[\{x_1, \dots, x_i\}]$ is connected for all $i \in [t]$. We let d_i denote the degree of the vertex x_i in the graph T_i.

Let h_i denote the number of injective graph homomorphisms from T_i to G^2. We note that $h_1 = m^n$ and

$$n_T = h_t/|\mathrm{Aut}(T)|. \tag{17.1}$$

We will construct an injective graph homomorphism $\varphi : T \to G^2$ recursively as follows. Suppose that we have constructed an injective graph homomorphism $\varphi_i : T_i \to G^2$ for some $i \leq t-1$. We will extend φ_i to an injective graph homomorphism $\varphi_{i+1} : T_{i+1} \to G^2$. Let E_i denote the set of possible choices for $\varphi_{i+1}(x_{i+1})$. We consider two cases.

If $d_{i+1} > 1$, then $\varphi_{i+1}(x_{i+1})$ must lie in the joint neighbourhood of $\varphi_i(x)$ and $\varphi_i(y)$ for some $x, y \in V(T_i)$. For any pair of vertices $u, v \in V$ their codegree in G^2 is at most $4n$ and so

$$|E_i| \leq 4n. \tag{17.2}$$

Suppose now that $d_{i+1} = 1$. We note that $u \in E_i$ if and only if u is adjacent to $\varphi_i(x_i)$ and non-adjacent to $\varphi_i(x_j)$ for $j < i$ in G^2. Again using the fact that the maximum codegree in G^2 is at most $4n$ we have

$$|E_i| = (1 + 3\mathbb{1}_{m>2})\binom{n}{2} + O(n). \tag{17.3}$$

If T is not a tree then $d_{i+1} > 1$ for some $i \leq t-1$. It follows by (17.2) and (17.3) that $h_t = O(m^n n^{2(t-1)-1}) = O(m^n n^{2t-3})$. The bound $n_T = O(m^n n^{2t-3})$ follows from (17.1).

Suppose now that T is a tree so that $d_{i+1} = 1$ for all $i \leq t-1$. Let h'_i denote the number of injective graph homomorphisms from T_i to G^2 whose image is typical. Let $E'_i \subseteq E_i$ be the set of possible choices for $\varphi_{i+1}(x_{i+1})$ so that $\varphi_{i+1}(x_i), \varphi_{i+1}(x_{i+1})$ have codegree 2 in G (in particular x_i, x_{i+1} both lie in $\mathcal{E}$ or $\mathcal{O}$). Given $u \in E_i$, we have $u \notin E'_i$ if and only if either u is adjacent to $\varphi_i(x_i)$ in G or the codegree of u and $\varphi_i(x_i)$ is 1 (the latter is only possible when $m > 2$). It follows that

$$|E_i \backslash E'_i| \leq 4n. \tag{17.4}$$

By (17.3) and (17.4), we have

$$h'_t = (1+o(1))(1+3\mathbb{1}_{m>2})^{t-1}2^{-(t-1)}m^n n^{2(t-1)}$$

and $h_t - h'_t = O(m^n n^{2t-3})$. The result follows by (17.1) and the analogous identity $n'_T = h'_t/|\mathrm{Aut}(T)|$. □

Lemma 17.2. *Fix $k \in \mathbb{N}$ and let γ be a polymer of size k. We have*

$$w_{A,B}(\gamma) = O\left(\left(1 - \frac{1}{\lceil q/2 \rceil}\right)^{dk}\right).$$

Moreover if γ is typical of type T, then if $\gamma \subseteq \mathcal{O}$

$$w_{A,B}(\gamma) = \frac{|B|^k}{|A|^k}\left(1 - \frac{1}{|B|}\right)^{dk-k+1}$$

and if $\gamma \subseteq \mathcal{E}$

$$w_{A,B}(\gamma) = \frac{|A|^k}{|B|^k}\left(1 - \frac{1}{|A|}\right)^{dk-k+1}.$$

Proof. We let w denote $w_{A,B}$. By (10.5), noting that in this setting $\lambda_i = 1$ for all i and $\{\lambda_A, \lambda_B\} = \{\lceil q/2 \rceil, \lfloor q/2 \rfloor\}$,

$$w(\gamma) = O\left(\left(1 - \frac{1}{\lceil q/2 \rceil}\right)^{|\partial\gamma|}\right).$$

Since the size of γ is k, a constant, we have by Lemma 6.2 that $|\partial\gamma| = dk - O(1)$ and so

$$w(\gamma) = O\left(\left(1 - \frac{1}{\lceil q/2 \rceil}\right)^{dk}\right).$$

Suppose now that γ is typical and of type T where T is a tree on k vertices. We may abuse notation slightly and let T denote the graph $G^2[\gamma]$ (as well as the isomorphism class of this graph). Suppose that $\gamma \subseteq \mathcal{O}$ (the proof where $\gamma \subseteq \mathcal{E}$ is identical). Let $a = |A|$ and let $b = |B|$. Since γ is typical we have $|N(\gamma)| = dk - 2(k-1)$.

By (16.3)

$$w_{A,B}(\gamma) = \frac{|\hat{\chi}_{A,B}(\gamma)|}{a^{|\gamma|}b^{|N(\gamma)|}} = \frac{|\hat{\chi}_{A,B}(\gamma)|}{a^k b^{dk-2(k-1)}} \tag{17.5}$$

where $\hat{\chi}_{A,B}(\gamma)$ is the set of all colourings of $G[\gamma^+]$ which disagree with (A,B) precisely on γ. Since $\gamma \subseteq \mathcal{O}$, $\hat{\chi}_{A,B}(\gamma)$ is the set of proper colourings $f : \gamma^+ \to B$. Fix a map

$c : \gamma \to B$ and suppose that we want to extend it to an element of $\hat{\chi}_{A,B}(\gamma)$. For every vertex $v \in N(\gamma)$ we have $b-1$ choices for its colour except if $v \in N(u) \cap N(w)$ for some $u, w \in \gamma$ where $c(u) \neq c(v)$. For such a vertex we have $b-2$ choices for its colour. We therefore have

$$|\hat{\chi}_{A,B}(\gamma)| = \sum_{c:\gamma\to B} (b-1)^{dk-2(k-1)} \left(\frac{b-2}{b-1}\right)^{\sum_{\{u,v\}\in E(T)} \mathbb{1}_{c(u)\neq c(v)}}$$
$$= (b-1)^{dk-3(k-1)}(b-2)^{k-1} \sum_{c:\gamma\to B} \left(\frac{b-1}{b-2}\right)^{\sum_{\{u,v\}\in E(T)} \mathbb{1}_{c(u)=c(v)}} .$$

Though not essential for the proof, we note that this final sum is an evaluation of the b-state Potts model partition function of the graph T. It is well known that the Potts model partition function is the same for all trees T on k vertices and indeed a simple induction on k reveals that

$$\sum_{c:\gamma\to B} x^{\sum_{\{u,v\}\in E(T)} \mathbb{1}_{c(u)=c(v)}} \equiv b(x+b-1)^{k-1} .$$

We therefore have

$$|\hat{\chi}_{A,B}(\gamma)| = b(b-1)^{dk-k+1} .$$

The result now follows from (17.5). □

17.2. Terms of the cluster expansion

In this section we show how Lemmas 17.1 and 17.2 can be used to provide a detailed picture of the cluster expansion. Recall that we are in the specialised setting $(H, \lambda) = (K_q, i)$ where $\lambda \equiv 1$, so that the expression $L_{A,B}(k)$ (as defined in (1.10)) is the same for all dominant patterns (A, B) by symmetry. We therefore denote $L_{A,B}(k)$ simply by L_k.

Let $\mathcal{T}_k$ denote the set of trees on k vertices.

Lemma 17.3. *For $k \in \mathbb{N}$ fixed, as $n \to \infty$*

$$L_k = (1+o(1))c_k \left(1 - \frac{1}{\lceil q/2 \rceil}\right)^{dk} m^n n^{2k-2}$$

where

$$c_k = \frac{(1+\mathbb{1}_{q \text{ even}})(1+3\mathbb{1}_{m>2})^{k-1}}{2^k(\lceil q/2 \rceil - 1)^{k-1}} \frac{\lceil q/2 \rceil^k}{\lfloor q/2 \rfloor^k} \sum_{T\in\mathcal{T}_k} \frac{1}{|\mathrm{Aut}(T)|} .$$

Proof. Let Γ be a cluster of size k. By Lemma 17.2

$$w(\Gamma) = \phi(I_\Gamma) \prod_{\gamma \in \Gamma} w(\gamma) = O\left(\left(1 - \frac{1}{\lceil q/2 \rceil}\right)^{dk}\right) \tag{17.6}$$

where $\phi(I_\Gamma)$ is the Ursell function of the incompatibility graph of Γ as defined in (1.8).

We say a cluster Γ is of type T if the set $V(\Gamma) \coloneqq \bigcup_{\gamma \in \Gamma} \gamma$ is of type T. Similarly we say Γ is typical if $V(\Gamma)$ is typical.

Note that given a set $X \subseteq V$ of size at most k, there are at most a constant (dependent on k) number of clusters Γ of size k such that $V(\Gamma) = X$. The number of clusters of size k and type T is therefore $O(n_T)$ (recall that n_T is the number of G^2-connected sets of type T).

For $T \in \mathcal{T}_k$, let $\mathcal{C}_k(T)$ denote the set of all typical clusters Γ of type T with $\|\Gamma\| = k$. Then by Lemma 17.1 and (17.6)

$$L_k = \sum_{\Gamma \in \mathcal{C}_k} w(\Gamma) = \sum_{T \in \mathcal{T}_k} \sum_{\Gamma \in \mathcal{C}_k(T)} w(\Gamma) + O\left(\left(1 - \frac{1}{\lceil q/2 \rceil}\right)^{dk} m^n n^{2k-3}\right). \tag{17.7}$$

Let us now fix $T \in \mathcal{T}_k$. Since T is a tree, for $1 \leq \ell \leq k$, deleting $\ell - 1$ edges from T results in a graph with precisely ℓ connected components. Therefore the number of vertex partitions $V(T) = V_1 \cup \ldots \cup V_\ell$ for which $T[V_i]$ is connected for each i is $\binom{k-1}{\ell-1}$. Thus if $X \subseteq V$ is of type T, the number of clusters $\Gamma \in \mathcal{C}_k(T)$ consisting of ℓ polymers with $V(\Gamma) = X$ is $\ell! \binom{k-1}{\ell-1}$ (recall that a cluster is an ordered multiset of polymers). Moreover for any such cluster Γ, the incompatibility graph I_Γ is a tree on ℓ vertices and so has Ursell function $\phi(I_\Gamma) = (-1)^{\ell-1}/\ell!$. If $X = V(\Gamma)$ is typical with $X \subseteq \mathcal{E}$, then by Lemma 17.2, letting $a = |A|, b = |B|$,

$$w(\Gamma) = \frac{1}{\ell!}(-1)^{\ell-1} \frac{a^k}{b^k} \left(1 - \frac{1}{a}\right)^{dk-k+\ell};$$

we obtain the same formula with a, b swapped if $X \subseteq \mathcal{O}$. Recalling that n'_T denotes the number of typical sets of type T, we then have

$$\begin{aligned}
\sum_{\Gamma \in \mathcal{C}_k(T)} w(\Gamma) &= \frac{n'_T}{2} \sum_{\ell=1}^{k} \binom{k-1}{\ell-1} (-1)^{\ell-1} \frac{a^k}{b^k} \left(1 - \frac{1}{a}\right)^{dk-k+\ell} \\
&\quad + \frac{n'_T}{2} \sum_{\ell=1}^{k} \binom{k-1}{\ell-1} (-1)^{\ell-1} \frac{b^k}{a^k} \left(1 - \frac{1}{b}\right)^{dk-k+\ell} \\
&= \frac{n'_T}{2} \left[\frac{a^k}{b^k (a-1)^{k-1}} \left(1 - \frac{1}{a}\right)^{dk} + \frac{b^k}{a^k (b-1)^{k-1}} \left(1 - \frac{1}{b}\right)^{dk} \right]
\end{aligned}$$

$$= (1+o(1))(1+\mathbb{1}_{q \text{ even}})\frac{n_T'}{2}\frac{\lceil q/2\rceil^k}{\lfloor q/2\rfloor^k(\lceil q/2\rceil-1)^{k-1}}\left(1-\frac{1}{\lceil q/2\rceil}\right)^{dk}.$$

The result follows from (17.7) and Lemma 17.1. □

The above lemma shows that each term L_k is positive for large n. Theorem 1.4 then shows that in order to compute an expression for $c_q(\mathbb{Z}_m^n)$ that is correct up to a multiplicative $(1+o(1))$ factor, one necessarily has to compute $L_1,\dots,L_{k-1}$ where k is the least integer for which $L_k = o(1)$. By Lemma 17.3, this is the least integer k such that $m(1-1/\lceil q/2\rceil)^{(1+\mathbb{1}_{m>2})k} < 1$.

17.3. The hypercube

To end this section we specialise to the case where $m=2$ and fully determine L_1, L_2 in order to arrive at the expressions for $c_q(Q_n)$ stated in Corollary 1.6. The calculations required to compute the expressions for $i(\mathbb{Z}_m^n)$ in Corollary 1.6 are similar and so we omit them.

The following is a strengthening of Theorem 1.5.

Theorem 17.4. *For $q \geq 4$*

$$c_q(Q_n) = (1+\mathbb{1}_{\{q \text{ odd}\}})\binom{q}{\lfloor q/2\rfloor}\left\lfloor\frac{q}{2}\right\rfloor^{2^{n-1}}\left\lceil\frac{q}{2}\right\rceil^{2^{n-1}}\cdot\exp\{L_1+L_2+\varepsilon\},$$

where

$$L_1 = \frac{\lceil q/2\rceil}{2\lfloor q/2\rfloor}\left(2-\frac{2}{\lceil q/2\rceil}\right)^n + \frac{\lfloor q/2\rfloor}{2\lceil q/2\rceil}\left(2-\frac{2}{\lfloor q/2\rfloor}\right)^n,$$

$$\begin{aligned} L_2 = 2^n\Bigg[&\frac{q-1}{2(\lfloor q/2\rfloor-1)(\lceil q/2\rceil-1)}\cdot n\left(1-\frac{1}{\lfloor q/2\rfloor}\right)^n\left(1-\frac{1}{\lceil q/2\rceil}\right)^n \\ &+\frac{\lfloor q/2\rfloor^2}{8\lceil q/2\rceil^2(\lfloor q/2\rfloor-1)^3}(n^2-n-2(\lfloor q/2\rfloor-1)^3)\left(1-\frac{1}{\lfloor q/2\rfloor}\right)^{2n} \\ &+\frac{\lceil q/2\rceil^2}{8\lfloor q/2\rfloor^2(\lceil q/2\rceil-1)^3}(n^2-n-2(\lceil q/2\rceil-1)^3)\left(1-\frac{1}{\lceil q/2\rceil}\right)^{2n}\Bigg], \end{aligned}$$

and

$$\varepsilon = O\left(\left(1-\frac{1}{\lceil q/2\rceil}\right)^{3n}2^n n^4\right).$$

Proof. The result will follow from Theorem 1.4 once we verify the expressions for L_1, L_2.

Let (A,B) be a dominant pattern. For ease of notation we let $a = |A| = \lfloor q/2\rfloor$ and $b = |B| = \lceil q/2\rceil$.

Polymers There are two types of polymer of size 1: a single vertex in $\mathcal{O}$ and a single vertex in $\mathcal{E}$. There are 2^{n-1} of the first type and each has weight $\frac{b}{a}(1-\frac{1}{b})^n$. There are 2^{n-1} of the second type and each has weight $\frac{a}{b}(1-\frac{1}{a})^n$.

There are three types of polymer of size 2:

(i) two vertices in $\mathcal{O}$ at distance 2 in Q_n,
(ii) two vertices in $\mathcal{E}$ at distance 2, and
(iii) an edge of Q_n.

There are $2^{n-3}n(n-1)$ of type (i) and each has weight

$$\frac{b(b-1)}{a^2}\cdot\frac{(b-2)^2(b-1)^{2n-4}}{b^{2n-2}}+\frac{b}{a^2}\cdot\frac{(b-1)^{2n-2}}{b^{2n-2}}=\frac{1}{a^2}(b^2-3b+3)(1-\tfrac{1}{b})^{2n-3}.$$

The first term in the sum on the LHS accounts for colourings where the two vertices of the polymer receive distinct colours in B and the second term accounts for colourings where the two vertices receive the same colour. Similarly, there are $2^{n-3}n(n-1)$ of type (ii) and each has weight $\frac{1}{b^2}(a^2-3a+3)(1-\frac{1}{a})^{2n-3}$. There are $n2^{n-1}$ of type (iii) and each has weight $(1-\frac{1}{a})^{n-1}(1-\frac{1}{b})^{n-1}$.

Clusters There are two cluster types of size 1, each consisting of single polymer of size 1, with Ursell function 1 and count and weight given above. Thus

$$L_1=\frac{a}{2b}\left(2-\frac{2}{a}\right)^n+\frac{b}{2a}\left(2-\frac{2}{b}\right)^n.$$

There are four types of clusters of size 2:

1. An ordered pair of incompatible polymers of size 1 both from $\mathcal{O}$,
2. An ordered pair of incompatible polymers of size 1 both from $\mathcal{E}$,
3. An ordered pair of incompatible polymers of size 1, one from $\mathcal{O}$ and one from $\mathcal{E}$,
4. One polymer of type (i), (ii) or (iii) above.

There are $2^{n-1}+2^{n-2}n(n-1)$ of type 1, with Ursell function $-1/2$ and weight $\frac{b^2}{a^2}(1-\frac{1}{b})^{2n}$. There are $2^{n-1}+2^{n-2}n(n-1)$ of type 2, with Ursell function $-1/2$ and weight $\frac{a^2}{b^2}(1-\frac{1}{a})^{2n}$. There are $n2^n$ of type 3, with Ursell function $-1/2$ and weight $(1-\frac{1}{a})^n(1-\frac{1}{b})^n$. The rest have Ursell function 1 with counts and weights given above.

All together this gives:

$$L_2=2^n\Bigg[\frac{a+b-1}{2(a-1)(b-1)}\cdot n\left(1-\frac{1}{a}\right)^n\left(1-\frac{1}{b}\right)^n + \frac{a^2}{8b^2(a-1)^3}(n^2-n-2(a-1)^3)\left(1-\frac{1}{a}\right)^{2n}$$

$$+\frac{b^2}{8a^2(b-1)^3}(n^2-n-2(b-1)^3)\left(1-\frac{1}{b}\right)^{2n}\Bigg]. \quad \square$$

The expressions for $c_q(Q_n)$ in Corollary 1.6 follow once we observe that $\varepsilon_n = o(1)$ for $q \in \{5,6,7,8\}$.

18. k-Bounded functions and height functions

In this section we apply our detailed understanding of the set $\mathrm{Hom}(Q_n, H)$ (for appropriately chosen H) in order to prove Theorem 1.9 and then Theorem 1.10.

18.1. k-Bounded functions

Recall from the introduction that a function $f : 2^{[n]} \to \mathbb{N}$ is *k-bounded* if $f(\emptyset) = 0$ and

$$0 \le f(A \cup \{x\}) - f(A) \le k$$

for all $A \subseteq [n]$ and $x \in [n] \backslash A$. We let $\mathcal{B}_k(n)$ denote the set of all k-bounded functions on $2^{[n]}$. Recall also that

$$\mathcal{F}(n) = \{f : V(Q_n) \to \mathbb{Z} : f(\mathbf{0}) = 0 \text{ and } u \sim v \implies |f(u) - f(v)| = 1\}.$$

As observed by Mossel (see [36]), there is a bijection from $\mathcal{B}_1(n)$ (rank functions) to the set $\mathcal{F}(n)$. The next lemma generalises this observation, giving a bijection from $\mathcal{B}_k(n)$ to an appropriate class of functions $f : V(Q_n) \to \mathbb{Z}$.

Given a subset $\mathcal{S} \subseteq \mathbb{N}$, let

$$\mathrm{Lip}(Q_n; \mathcal{S}) \coloneqq \{f : V(Q_n) \to \mathbb{Z} : f(\mathbf{0}) = 0 \text{ and } u \sim v \implies |f(u) - f(v)| \in \mathcal{S}\}.$$

In particular $\mathcal{F}(n) = \mathrm{Lip}(Q_n; \{1\})$.

Lemma 18.1. *There exists a bijection*

$$\varphi : \mathcal{B}_k(n) \to \mathrm{Lip}(Q_n; \mathcal{S}_k),$$

where

$$\mathcal{S}_k = \{0, \dots, k\} \cap (k + 2\mathbb{Z}).$$

Proof. Throughout we identify $2^{[n]}$ with $V(Q_n) = \{0,1\}^n$ in the usual way. For $v \in V(Q_n)$, we let $|v| \coloneqq \sum_i v_i$. Given $f \in \mathcal{B}_k(n)$, let $\varphi(f)$ denote the map $V(Q_n) \to \mathbb{Z}$ given by

$$v \mapsto 2f(v) - k|v|\,.$$

We claim that $\varphi(f) \in \mathrm{Lip}(Q_n; \mathcal{S}_k)$. Clearly $\varphi(f)(\mathbf{0}) = 0$ since $f(\mathbf{0}) = 0$. Now suppose $u \sim v$ in Q_n and without loss of generality let $|u| = |v| + 1$. Then

$$|\varphi(f)(u) - \varphi(f)(v)| = |2f(u) - 2f(v) - k| \in \mathcal{S}_k$$

since $f \in \mathcal{B}_k(n)$ and so $0 \leq f(u) - f(v) \leq k$.

For $g \in \mathrm{Lip}(Q_n; \mathcal{S}_k)$, let $\varphi'(g)$ denote the map $V(Q_n) \to \mathbb{Q}$ given by

$$v \mapsto (g(v) + k|v|)/2\,.$$

It suffices to show that $\varphi'(g) \in \mathcal{B}_k(n)$ since then clearly φ, φ' are inverse to each other. Note first that $\varphi'(g)(\mathbf{0}) = 0$ since $g(\mathbf{0}) = 0$. If k is even then $g(v)$ is even for all $v \in V(Q_n)$. If k is odd then $g(v)$ and $|v|$ have the same parity for all $v \in V(Q_n)$. In either case the image of $\varphi'(g)$ is a subset of $\mathbb{Z}$. Suppose now that $u \sim v$ in Q_n where $|u| = |v| + 1$. Then

$$\varphi'(g)(u) - \varphi'(g)(v) = \frac{g(u) - g(v) + k}{2} \in \{0, \ldots, k\}$$

since $g \in \mathrm{Lip}(Q_n; \mathcal{S}_k)$. This completes the proof. □

We now show that for any finite set $\mathcal{S} \subseteq \mathbb{N}$, there is a bijection between the sets $\mathrm{Lip}(Q_n; \mathcal{S})$ and a subset of $\mathrm{Hom}(Q_n, H)$ for an appropriately chosen graph H. To this end we define the following class of Cayley graphs. Given $N \in \mathbb{N}$ and $\mathcal{S} \subseteq \mathbb{Z}_N$, let $C(N; \mathcal{S})$ denote the graph on vertex set $\mathbb{Z}_N$ where $u \sim v$ if and only if $u - v = \pm x$ for some $x \in \mathcal{S}$. Note that $C(N; \mathcal{S})$ has loops if $0 \in \mathcal{S}$. Let

$$\mathrm{Hom}_{\mathbf{0}}(Q_n, C(N; \mathcal{S})) \coloneqq \{f \in \mathrm{Hom}(G, C(N; \mathcal{S})) : f(\mathbf{0}) = 0\}.$$

Suppose now $f : V(Q_n) \to \mathbb{Z}$. Let $\mathrm{Mod}_N(f)$ denote the map $V(Q_n) \to \{0, \ldots, N-1\}$ where $\mathrm{Mod}_N(f)(v) \equiv f(v) \pmod N$. The following lemma is an adaptation of [16, Proposition 2.1].

Lemma 18.2. *Fix a finite set $\mathcal{S} \subseteq \mathbb{N} \cup \{0\}$. For $N \geq 4 \cdot \max \mathcal{S} + 1$*

$$\mathrm{Mod}_N : \mathrm{Lip}(Q_n; \mathcal{S}) \to \mathrm{Hom}_{\mathbf{0}}(Q_n, C(N; \mathcal{S}))$$

is a bijection.

Proof. We first show that Mod_N is injective. Indeed suppose that $f, g \in \mathrm{Lip}(Q_n; \mathcal{S})$ where $f \neq g$. Since $f(\mathbf{0}) = g(\mathbf{0}) = 0$, there must exist $\{u, v\} \in E(Q_n)$ such that $f(u) = g(u) =: x$ whereas $f(v) \neq g(v)$. Let $s \coloneqq \max \mathcal{S}$. Since $f, g \in \mathrm{Lip}(Q_n; \mathcal{S})$, we have $f(v), g(v) \in [x - s, x + s]$. Since $N > 2s$, it follows that $\mathrm{Mod}_N(f)(v) \neq \mathrm{Mod}_N(g)(v)$.

It remains establish surjectivity. For $v \in V(Q_n) = \{0,1\}^n$, we let $|v| \coloneqq \sum_i v_i$. We define a spanning tree T of Q_n, rooted at $\mathbf{0}$, as follows. If $|v| > 0$, with j the first coordinate for which $v_j \neq 0$, let $v^* \coloneqq (0,\dots,0,v_j-1,v_{j+1},\dots,v_n)$ be the parent of v in T.

Suppose now that $g \in \mathrm{Hom}_{\mathbf{0}}(Q_n, C(N;\mathcal{S}))$. We will construct $f \in \mathrm{Lip}(Q_n;\mathcal{S})$ such that $g = \mathrm{Mod}_N(f)$ recursively as follows. Set $f(\mathbf{0}) = 0$. Suppose now that $|v| > 0$ and that we have defined $f(w)$ for all w such that $|w| < |v|$. We then let $f(v)$ be the unique integer z such that $|f(v^*) - z| \in \mathcal{S}$ and $z \equiv g(v) \pmod N$. It remains to check that f is indeed an element of $\mathrm{Lip}(Q_n;\mathcal{S})$. We need to check that

$$|f(u) - f(v)| \in \mathcal{S} \text{ for all } \{u,v\} \in E(Q_n)\,. \tag{$\dagger$}$$

Note that by construction ($\dagger$) holds for all pairs $\{v, v^*\}$. We proceed by induction on $|u| + |v|$. If $|u| + |v| = 1$ then $u = v^*$ or vice versa and so ($\dagger$) holds. Suppose now that $\{u,v\} \in E(Q_n)$ with $|u| + |v| > 1$.

If $\min\{i : u_i \neq 0\} > \min\{i : v_i \neq 0\}$ then $u = v^*$ and so ($\dagger$) holds. If $\min\{i : u_i \neq 0\} = \min\{i : v_i \neq 0\}$, then $u \sim u^* \sim v^* \sim v$. By the induction hypothesis and the construction of f, $|f(u) - f(u^*)|, |f(u^*) - f(v^*)|, |f(v) - f(v^*)| \in \mathcal{S}$ and so $|f(u) - f(v)| \leq 3s$. Moreover $f(u) - f(v) \equiv g(u) - g(v) \equiv \pm x \pmod N$ for some $x \in \mathcal{S}$. We deduce that $|f(u) - f(v)| \in \mathcal{S}$ since $N > 4s$. □

Proof of Theorem 1.9. By Lemmas 18.1 and 18.2,

$$|\mathcal{B}_k(n)| = |\mathrm{Hom}_{\mathbf{0}}(Q_n, C(N;\mathcal{S}_k))|$$

where $\mathcal{S}_k = \{0,\dots,k\} \cap (k + 2\mathbb{Z})$ and $N \coloneqq 4k+1$. By vertex transitivity of $C(N;\mathcal{S}_k)$ we also have

$$|\mathcal{B}_k(n)| = |\mathrm{Hom}_{\mathbf{0}}(Q_n, C(N;\mathcal{S}_k))| = \frac{1}{N}|\mathrm{Hom}(Q_n, C(N;\mathcal{S}_k))|\,. \tag{18.1}$$

We now use Theorem 1.4 to obtain accurate asymptotics for $|\mathrm{Hom}(Q_n, C(N;\mathcal{S}_k))|$ (and therefore also $|\mathcal{B}_k(n)|$). Let $H = C(N;\mathcal{S}_k)$. We begin by identifying the dominant patterns of H. First let us establish some terminology. We call a set $\{y_1,\dots,y_t\} \subseteq V(H) = \mathbb{Z}_N$, an *interval* if the elements of the set can be relabelled such that $y_{i+1} = y_i + 1$ for all $1 \leq i \leq t-1$. Similarly, we call the set $\{y_1,\dots,y_t\}$, a *skip interval* if the elements of the set can be renamed such that $y_{i+1} = y_i + 2$ for all $1 \leq i \leq t-1$.

Suppose now that (A,B) is a pattern in H and write $B = \{x_1,\dots,x_\ell\} \subseteq \mathbb{Z}_N$. Since $a \sim b$ in H for all $a \in A, b \in B$, we have

$$A \subseteq (x_1 \pm \mathcal{S}_k) \cap \ldots \cap (x_\ell \pm \mathcal{S}_k) \subseteq \mathbb{Z}_N\,,$$

where $x_i \pm \mathcal{S}_k \coloneqq x_i + (-\mathcal{S}_k \cup \mathcal{S}_k)$. Now, if $\ell \leq k+2$,

$$|(x_1 \pm \mathcal{S}_k) \cap \ldots \cap (x_\ell \pm \mathcal{S}_k)| \leq |-\mathcal{S}_k \cup \mathcal{S}_k| - (\ell - 1) = k + 2 - \ell$$

with equality if and only if B is a skip interval. Thus,

$$|A||B| \leq \ell(k+2-\ell) \leq \left\lfloor \frac{k}{2} + 1 \right\rfloor \left\lceil \frac{k}{2} + 1 \right\rceil$$

with equality if and only if A, B are both skip intervals with $\{|A|, |B|\} = \left\{ \left\lfloor \frac{k}{2} + 1 \right\rfloor, \left\lceil \frac{k}{2} + 1 \right\rceil \right\}$. If k is even and $B = \{x_1, \ldots, x_{k/2+1}\}$, then

$$A = (x_1 \pm \mathcal{S}_k) \cap \ldots \cap (x_{k/2+1} \pm \mathcal{S}_k) = B\,.$$

If k is odd and $B = \{x_1, \ldots, x_{(k+3)/2}\}$ where $x_{i+1} \equiv x_i + 2$ for $1 \leq i \leq (k+1)/2$, then

$$A = (x_1 \pm \mathcal{S}_k) \cap \ldots \cap (x_{(k+3)/2} \pm \mathcal{S}_k) = \{x_1 + 1, x_2 + 1, \ldots, x_{(k+1)/2} + 1\}\,.$$

If k is even, we therefore have N dominant patterns (one for each choice of skip interval with $k/2+1$ elements). If k is odd we have $2N$ dominant patterns: we choose a skip interval B of length $(k+3)/2$, then we have dominant patterns (A, B) and (B, A).

Given a dominant pattern (A, B), let us calculate $L_{A,B}(1)$ (as defined in (1.10)). Again we split into cases depending on the parity of k. If k is even, by symmetry we may assume that $A = B = \{0, 2, \ldots, k\}$. Calculating $L_{A,B}(1)$ (e.g. by using the explicit expression (1.11)) we have

$$L_{A,B}(1) = 2^{n-1} \left[\frac{2}{(k/2+1)^{n+1}} \sum_{v \in A^c} |N(v) \cap B|^n \right]\,.$$

If $v = k+2t$, where $1 \leq t \leq k/2$ then $N(v) \cap B = \{2t, 2t+2, \ldots, k\}$ and so $|N(v) \cap B| = k/2 - t + 1$. Similarly if $v = -2t$ then $|N(v) \cap B| = k/2 - t + 1$. For all other $v \in A^c$, $N(v) \cap B = \emptyset$. We therefore have

$$L_{A,B}(1) = \frac{2^{n+2}}{(k+2)} \sum_{t=1}^{k/2} \left(\frac{2t}{k+2} \right)^n = \frac{2^{n+2}}{(k+2)} \left(\frac{k}{k+2} \right)^n + O\left(2^n \left(\frac{k-2}{k+2} \right)^n \right)$$

and $\delta_{A,B} = k/(k+2)$ (with $\delta_{A,B}$ as defined in (4.5)). By Lemma 10.2 we therefore have

$$\sum_{j=2}^{\infty} L_{A,B}(j) = O\left(n^2 2^n \left(\frac{k}{k+2} \right)^{2n} \right)\,.$$

By Theorem 4.7 we then have

$$|\mathcal{B}_k(n)| = \left(\frac{k}{2} + 1 \right)^{2^n} \exp\left\{ \frac{2^{n+2}}{k+2} \left(\frac{k}{k+2} \right)^n + O\left(n^2 2^n \left(\frac{k}{k+2} \right)^{2n} \right) \right\}$$

(note we have N dominant patterns which cancels the factor of $1/N$ in (18.1)).

If k is odd, by symmetry we may assume that $A = \{1, 3, \dots, k\}, B = \{0, 2, \dots, k+1\}$. Then

$$L_{A,B}(1) = 2^{n-1}\left[\frac{1}{|A||B|^n}\sum_{v\in A^c}|N(v)\cap B|^n + \frac{1}{|B||A|^n}\sum_{v\in B^c}|N(v)\cap A|^n\right].$$

If $v = k+2t$, where $1 \le t \le (k+1)/2$ then $N(v)\cap B = \{2t, 2t+2, \dots, k+1\}$ and so $|N(v)\cap B| = (k+1)/2 - t + 1$. Similarly if $v = 1-2t$ then $|N(v)\cap B| = (k+1)/2 - t + 1$. For all other $v \in A^c$, $N(v)\cap B = \emptyset$.

If $v = k+2t+1$, where $1 \le t \le (k-1)/2$ then $N(v)\cap A = \{2t+1, 2t+3, \dots, k\}$ and so $|N(v)\cap A| = (k-1)/2 - t + 1$. Similarly if $v = -2t$ then $|N(v)\cap A| = (k-1)/2 - t + 1$. For all other $v \in B^c$, $N(v)\cap A = \emptyset$. Thus $\delta_{A,B} = (k+1)/(k+3)$ and

$$\begin{aligned} L_{A,B}(1) &= \frac{2^{n+1}}{k+1}\sum_{t=1}^{(k+1)/2}\left(\frac{2t}{k+3}\right)^n + \frac{2^{n+1}}{k+3}\sum_{t=1}^{(k-1)/2}\left(\frac{2t}{k+1}\right)^n \\ &= \frac{2^{n+1}}{k+1}\left(\frac{k+1}{k+3}\right)^n + O\left(2^n\left(\frac{k-1}{k+1}\right)^n\right). \end{aligned}$$

By Lemma 10.2

$$\sum_{j=2}^{\infty} L_{A,B}(j) = O\left(n^2 2^n\left(\frac{k+1}{k+3}\right)^{2n}\right)$$

and so if $k \ge 3$, by Theorem 4.7,

$$|\mathcal{B}_k(n)| = 2\left(\frac{k+1}{2}\right)^{2^{n-1}}\left(\frac{k+3}{2}\right)^{2^{n-1}}\exp\left\{\frac{2^{n+1}}{k+1}\left(\frac{k+1}{k+3}\right)^n + O\left(2^n\left(\frac{k-1}{k+1}\right)^n\right)\right\}.$$

We have a leading factor of 2 in the above since there are $2N$ dominant patterns and a factor of $1/N$ in (18.1).

If $k = 1$ then the error term is dominated by $L_{A,B}(2)$ and so

$$|\mathcal{B}_1(n)| = \left(1 + O\left(\frac{n^2}{2^n}\right)\right)2e2^{2^{n-1}}.$$

For $k \ge 2$, we may summarise our results as follows:

$$|\mathcal{B}_k(n)| = (1 + \mathbb{1}_{k \text{ odd}})\left(\left\lfloor\frac{k}{2}+1\right\rfloor\left\lceil\frac{k}{2}+1\right\rceil\right)^{2^{n-1}}\exp\left\{\frac{(1+\mathbb{1}_{k \text{ even}})}{\lfloor k/2+1\rfloor}\left(\frac{2\lceil k/2\rceil}{\lceil k/2+1\rceil}\right)^n + \varepsilon\right\}$$

where

$$\varepsilon = \begin{cases} O\left(n^2 2^n \left(\frac{k}{k+2}\right)^{2n}\right) & k \text{ even} \\ O\left(2^n \left(\frac{k-1}{k+1}\right)^{n}\right) & k \text{ odd}. \end{cases} \quad \square$$

18.2. Height functions

As a further application of our large deviation result (Theorem 11.1), we now show how it can be used to prove Theorem 1.10. Recall that our goal is to estimate the probability that a uniformly chosen height function $f \in \mathcal{F}(n)$ takes $> tn$ values (for t fixed and n large). We let $R(f)$ denote the size of the range of f.

The study of height functions (on Q_n and more general graphs) and their concentration properties was initiated by Benjamini, Häggström and Mossel [3]. In [3] the authors conjecture that for any $t > 0$, if f is chosen uniformly at random from $\mathcal{F}(n)$, then $\mathbb{P}(R(f) > tn)$ converges to 0 as $n \to \infty$. Kahn [36] resolved this conjecture in a strong form showing that there is in fact a constant b such that $\mathbb{P}(R(f) > b) = e^{-\Omega(n)}$. Later, Galvin [20] strengthened Kahn's result further still showing that one can in fact take $b = 5$. We note that this lies in stark contrast to a subsequent result of Benjamini, Yadin, and Yehudayoff [4] who show that if the degree of G is 'sub-logarithmic', then the range of a random height function on G is super-constant.

Recently, Peled [44] vastly generalised the work of Galvin [20] obtaining results for a general class of tori including $\mathbb{Z}_m^n$ (where m is allowed to be large with respect to n) and provides strong bounds on $\mathbb{P}(R(f) \geq k)$ for arbitrary k. Since the upper bound on $\mathbb{P}(R(f) \geq tn)$ in Theorem 1.10 is a special case of Peled's result [44, Theorem 2.1], we sketch this part of the proof.

First we note that by Lemma 18.2, the map Mod_5 is a bijection $\mathcal{F}(n) \to \mathrm{Hom}_{\mathbf{0}}(Q_n, C_5)$ where C_5 denotes a cycle of length 5. A closer look at the proof of Lemma 18.2 reveals that Mod_3 provides a bijection $\mathcal{F}(n) \to \mathrm{Hom}_{\mathbf{0}}(Q_n, K_3)$ (this bijection is attributed to Randall in [20]). In the following it is slightly more convenient to work with this latter bijection.

Sketch proof of Theorem 1.10. Let $f \in \mathcal{F}(n)$. Note that since $f(u), f(v)$ have opposite parity for all $\{u, v\} \in E(Q_n)$ and $f(\mathbf{0}) = 0$ we have $f(\mathcal{E}) \subseteq 2\mathbb{Z}$ and $f(\mathcal{O}) \subseteq 1 + 2\mathbb{Z}$.

Let $t > 0$ and suppose that $R(f) \geq tn$. As remarked above, $\mathrm{Mod}_3(f) \in \mathrm{Hom}_{\mathbf{0}}(Q_n, K_3)$. Let (A, B) be the dominant pattern of (K_3, i) (where $i \equiv 1$) that agrees the most with $\mathrm{Mod}_3(f)$ (breaking ties arbitrarily if necessary). We say an integer x is *good* if either x is even and $x \pmod 3 \in A$ or x is odd and $x \pmod 3 \in B$. We say an integer is *bad* otherwise and note that if $f(v)$ is a bad integer, then $\mathrm{Mod}_3(f)(v)$ disagrees with (A, B) at v.

Let p be the midpoint of the range of f. Since good and bad integers occur in alternating intervals of length 3, we may choose $\{b, b+1\} \subseteq [p-3, p+3]$ such that b and $b+1$ are both bad. Let

$$V_{\text{small}} \coloneqq \{v \in V(Q_n) : f(v) \leq b\} \text{ and } V_{\text{large}} \coloneqq \{v \in V(Q_n) : f(v) \geq b+1\}$$

and note that $V_{\text{small}}, V_{\text{large}}$ partition $V(G)$. Choosing $v, w \in V(Q_n)$ so that $f(v) - f(w)$ is maximal, we have $B_v(tn/2-4) \subseteq V_{\text{large}}$ and $B_w(tn/2-4) \subseteq V_{\text{small}}$. Suppose that $|V_{\text{small}}| \leq |V_{\text{large}}|$ (we argue similarly if $|V_{\text{small}}| > |V_{\text{large}}|$) so that

$$|B_w(tn/2-4)| \leq |V_{\text{small}}| \leq 2^{n-1}.$$

By Harper's isoperimetric inequality [30, Theorem 1],

$$|\partial V_{\text{small}}| \geq |\partial B_w(tn/2-4)| = \binom{n}{tn/2-3}.$$

Since $f(u) = b+1$ for all $u \in \partial V_{\text{small}}$, and $b+1$ is bad we have that $\text{Mod}_3(f) \in \text{Hom}_{\mathbf{0}}(Q_n, K_3)$ disagrees with (A,B) at each vertex of $\partial V_{\text{small}}$.

By Lemma 11.2 (ii) and Theorem 1.2, the probability that a uniformly chosen element of $\text{Hom}(Q_n, K_3)$ disagrees with its closest dominant pattern at $\geq \binom{n}{tn/2-3}$ vertices is at most

$$\leq \exp\left\{-\Omega\left(\frac{1}{n}\binom{n}{tn/2-3}\right)\right\}.$$

The same is therefore also true of $\text{Hom}_{\mathbf{0}}(Q_n, K_3)$. Since $\text{Hom}_{\mathbf{0}}(Q_n, K_3)$ and $\mathcal{F}(n)$ are in bijection, it follows that if f is a uniformly chosen element of $\mathcal{F}(n)$, then

$$\mathbb{P}(R(f) \geq tn) \leq \exp\left\{-\Omega\left(\frac{1}{n}\binom{n}{tn/2}\right)\right\} = \exp\left\{-2^{H(t/2)n(1+o(1))}\right\}.$$

We now give a lower bound for $\mathbb{P}(R(f) \geq tn)$. We construct a set of height functions $f \in \text{Hom}(Q_n, \mathbb{Z})$ with $R(f) \geq tn$ as follows. Let τ denote the least integer $\geq tn-1$ with the same parity as n such that $\lfloor \tau/2 \rfloor$ is even. Let $\mathbf{1} = (1, \dots, 1) \in V(Q_n)$. Let $\mathcal{F}$ denote the family of $f : V(Q_n) \to \mathbb{Z}$ satisfying the following constraints:

1. $f(v) = d(v, \mathbf{0})$ for all $v \in B_{\mathbf{0}}(\lfloor \tau/2 \rfloor)$,
2. $f(v) = \tau - d(v, \mathbf{1})$ for all $v \in B_{\mathbf{1}}(\lceil \tau/2 \rceil)$,
3. $f(v) = \lfloor \tau/2 \rfloor$ for all $v \in \mathcal{E} \setminus (B_{\mathbf{0}}(\lfloor \tau/2 \rfloor) \cup B_{\mathbf{1}}(\lceil \tau/2 \rceil))$,
4. $f(v) \in \{\lfloor \tau/2 \rfloor - 1, \lfloor \tau/2 \rfloor + 1\}$ for all $v \in \mathcal{O} \setminus (B_{\mathbf{0}}(\lfloor \tau/2 \rfloor) \cup B_{\mathbf{1}}(\lceil \tau/2 \rceil))$.

It is clear that $\mathcal{F} \subseteq \text{Hom}(Q_n, \mathbb{Z})$ and $R(f) = \tau + 1 \geq tn$ for all $f \in \mathcal{F}$. Moreover,

$$|\mathcal{F}| \geq \exp_2\left\{2^{n-1} - 2|B_{\mathbf{0}}(\lceil \tau/2 \rceil)|\right\} \geq \exp_2\left\{2^{n-1} - O\left(2^{H(t/2)n}\right)\right\}.$$

Since $|\text{Hom}(Q_n, \mathbb{Z})| \sim 2e2^{2^{n-1}}$ by Theorem 1.4 and Lemma 18.2 (or alternatively [20, Corollary 1.5]), we have

$$\mathbb{P}(R(f) \geq tn) \geq \exp_2\left\{-O\left(2^{H(t/2)n}\right)\right\}. \quad \square$$

19. Concluding remarks

We conclude by mentioning some open problems and directions for future research.

n fixed, m large.

- The techniques of this paper were inspired by tools used to analyse spin models on the integer lattice $\mathbb{Z}^n$. In this context, it is common to study the torus $\mathbb{Z}_m^n$ (viewed as the integer lattice with 'periodic boundary conditions') where n is fixed and $m \to \infty$. The main obstacle in adapting the proofs of our paper to this setting is that the isoperimetric properties of the torus $\mathbb{Z}_m^n$ become too weak in the limit $m \to \infty$. One way to overcome this obstacle, inspired by [32], would be to replace our polymer models with *contour models* from Pirogov-Sinai theory (see also [18] for an excellent introduction to this topic). Roughly speaking, our polymer models encode the regions on which a colouring disagrees with a fixed dominant pattern whereas contour models can be used to encode the *boundary* between regions agreeing with distinct patterns. The advantage is that contour models encode configurations more efficiently, using fewer vertices to describe deviations from a dominant pattern. This type of approach was recently applied with great success by Peled and Spinka [48,50,49], to show that a general class of spin systems exhibit long-range order on $\mathbb{Z}^n$ (see the end of this section).
  Despite the obstacles posed by isoperimetry, we believe that the intuition provided by our polymer models is still a useful guide. In particular, in the regime where n is a fixed large constant and m is large with respect to n, we expect a typical configuration from $\mu_{H,\lambda}$ to globally agree with a single dominant pattern with regions of disagreeing vertices appearing in G^2-connected components of size at most $C \log m$ (where C may depend on n and (H, λ)).
  As a starting point, we offer the following conjecture, inspired by Theorem 4.7. As usual we let G denote $\mathbb{Z}_m^n$.

  Conjecture 19.1. *For a fixed weighted graph (H, λ)*

$$\lim_{m\to\infty} \frac{1}{m^n} \ln Z_G^H(\lambda) = \frac{1}{2} \ln \eta_\lambda(H) + (1 + o_n(1))\varepsilon(H, \lambda, n)\,, \tag{19.1}$$

  where the limit is taken over even m and

$$\varepsilon(H, \lambda, n) := \max_{(A,B)\in\mathcal{D}_\lambda(H)} \left(\frac{1}{2\lambda_A \lambda_B^{2n}} \sum_{v\in A^c} \lambda_v \lambda_{N(v)\cap B}^{2n} + \frac{1}{2\lambda_B \lambda_A^{2n}} \sum_{v\in B^c} \lambda_v \lambda_{N(v)\cap A}^{2n} \right).$$

  We note that the expression in parentheses in the above conjecture is $L_{A,B}(1)/m^n$. The RHS of (19.1) is easily seen to be a lower bound. Specialising to the case of approximating $c_q(G)$, the number of proper q-colourings of G, the above conjecture predicts that for q even

$$\lim_{m\to\infty} \frac{1}{m^n} \ln c_q(\mathbb{Z}_m^n) = \ln(q/2) + (1+o_n(1))(1-2/q)^{2n}\,, \tag{19.2}$$

where again we take the limit over even m. We obtain an analogous prediction for q odd. The current best upper bound is due to Peled and Spinka [50] whose results imply that there exists $c>0$, such that for q fixed and n large, the LHS of (19.2) is at most $\frac{1}{2}\ln\left(\left\lfloor\frac{q}{2}\right\rfloor\left\lceil\frac{q}{2}\right\rceil\right) + e^{-\frac{cn}{q^3(q+\log n)}}$.

- In the regime where the dimension n is fixed and m is large, various analogues of our torpid mixing result Theorem 1.11 have been proved throughout the literature for specific choices of (H,λ): Galvin [22] established such a result in the case of the hard-core model; Galvin and Randall [26] for the 3-colouring model; and Borgs, Chayes, Dyer and Tetali [8] for (H,λ) with carefully chosen λ. It would be interesting to study to what extent the intuition of Theorem 1.11, that non-trivial weighted graphs give rise to torpid mixing, holds true in this regime.

Odd sidelength. Throughout this paper we assumed that the sidelength m of the torus is even. It is natural to wonder what happens in the case where m is odd. One starting point would be to investigate the asymptotics of $Z_G^H(\lambda)$ for fixed odd m, and $n\to\infty$. To our knowledge, even the asymptotics of $i(\mathbb{Z}_3^n)$ are unknown.

General spin systems. The spin systems considered in this paper fit into a more general framework where one allows for soft interactions between spins. For a set of spins $[q]$ and functions $\lambda:[q]\to\mathbb{R}_{\geq 0}$ and $\delta:[q]\times[q]\to\mathbb{R}_{\geq 0}$ where $\delta_{i,j}=\delta_{j,i}$ for all $i,j\in[q]$, one can define the probability of a configuration $f:V(G)\to[q]$ to be proportional to

$$\prod_{v\in V(G)} \lambda_{f(v)} \prod_{\{u,v\}\in E(G)} \delta_{f(u),f(v)}\,.$$

This well-studied class of probability distributions notably includes the Ising model and Potts model.

The results of Peled and Spinka [48,49] mentioned above, show that a general class of such spin systems exhibit long-range order in $\mathbb{Z}^n$. Soon after the first appearance of our paper, Peled and Spinka posted [49] (a companion to [48]) detailing this remarkable result. It would be very interesting to explore the extent to which the results of our paper enjoy generalisations in this setting.

Acknowledgments

We thank the anonymous referees for their careful reading of the paper and helpful comments.

References

[1] C.A. Athanasiadis, Algebraic combinatorics of graph spectra, subspace arrangements and Tutte polynomials, PhD thesis, Massachusetts Institute of Technology, 1996.

[2] J. Balogh, R.I. Garcia, L. Li, Independent sets in the middle two layers of boolean lattice, J. Comb. Theory, Ser. A 178 (2021) 105341.
[3] I. Benjamini, O. Häggström, E. Mossel, On random graph homomorphisms into $\mathbb{Z}$, J. Comb. Theory, Ser. B 78 (1) (2000) 86–114.
[4] I. Benjamini, A. Yadin, A. Yehudayoff, Random graph-homomorphisms and logarithmic degree, Electron. J. Probab. 12 (2006) 12.
[5] A. Björklund, T. Husfeldt, P. Kaski, M. Koivisto, Computing the Tutte polynomial in vertex-exponential time, in: Proceedings of the Forty-ninth Annual Symposium on Foundations of Computer Science, FOCS 2008, IEEE, 2008, pp. 677–686.
[6] C. Borgs, J. Chayes, T. Helmuth, W. Perkins, P. Tetali, Efficient sampling and counting algorithms for the Potts model on $\mathbb{Z}^d$ at all temperatures, in: Proceedings of the 52nd Annual ACM SIGACT Symposium on Theory of Computing, 2020, pp. 738–751.
[7] C. Borgs, J. Chayes, J. Kahn, L. Lovász, Left and right convergence of graphs with bounded degree, Random Struct. Algorithms 42 (1) (2013) 1–28.
[8] C. Borgs, J.T. Chayes, M. Dyer, P. Tetali, On the sampling problem for H-colorings on the hypercubic lattice, in: Graphs, Morphisms, and Statistical Physics: DIMACS Workshop Graphs, Morphisms and Statistical Physics, March 19-21, 2001, DIMACS Center, vol. 63, American Mathematical Soc., 2004, p. 13.
[9] S. Cannon, W. Perkins, Counting independent sets in unbalanced bipartite graphs, in: Proceedings of the Fourteenth Annual ACM-SIAM Symposium on Discrete Algorithms, SIAM, 2020, pp. 1456–1466.
[10] D. Christofides, D. Ellis, P. Keevash, An approximate vertex-isoperimetric inequality for r-sets, Electron. J. Comb. 20 (4) (2013) P15.
[11] T.M. Cover, J.A. Thomas, Elements of information theory, John Wiley & Sons, 2012.
[12] R. Dobrushin, Estimates of semi-invariants for the Ising model at low temperatures, Transl. Am. Math. Soc. 2 (177) (1996) 59–82.
[13] M. Dyer, A. Frieze, M. Jerrum, On counting independent sets in sparse graphs, SIAM J. Comput. 31 (5) (2002) 1527–1541.
[14] J. Engbers, D. Galvin, H-coloring tori, J. Comb. Theory, Ser. B 102 (5) (2012) 1110–1133.
[15] J. Engbers, D. Galvin, H-colouring bipartite graphs, J. Comb. Theory, Ser. B 102 (3) (2012) 726–742.
[16] O.N. Feldheim, R. Peled, Rigidity of 3-colorings of the discrete torus, Ann. Inst. Henri Poincaré, Probab. Stat. 54 (2018) 952–994.
[17] O.N. Feldheim, Y. Spinka, Long-range order in the 3-state antiferromagnetic Potts model in high dimensions, J. Eur. Math. Soc. 21 (5) (2019) 1509–1570.
[18] S. Friedli, Y. Velenik, Statistical mechanics of lattice systems: a concrete mathematical introduction, Cambridge University Press, 2017.
[19] A. Galanis, L.A. Goldberg, J. Stewart, Fast algorithms for general spin systems on bipartite expanders, ACM Trans. Comput. Theory 13 (4) (2021) 1–18.
[20] D. Galvin, On homomorphisms from the Hamming cube to $\mathbb{Z}$, Isr. J. Math. 138 (1) (2003) 189–213.
[21] D. Galvin, Sampling 3-colourings of regular bipartite graphs, Electron. J. Probab. 12 (2007) 481–497.
[22] D. Galvin, Sampling independent sets in the discrete torus, Random Struct. Algorithms 33 (3) (2008) 356–376.
[23] D. Galvin, A threshold phenomenon for random independent sets in the discrete hypercube, Comb. Probab. Comput. 20 (1) (2011) 27–51.
[24] D. Galvin, J. Kahn, On phase transition in the hard-core model on $\mathbb{Z}^d$, Comb. Probab. Comput. 13 (2) (2004) 137–164.
[25] D. Galvin, J. Kahn, D. Randall, G.B. Sorkin, Phase coexistence and torpid mixing in the 3-coloring model on $\mathbb{Z}^d$, SIAM J. Discrete Math. 29 (3) (2015) 1223–1244.
[26] D. Galvin, D. Randall, Torpid mixing of local Markov chains on 3-colorings of the discrete torus, in: Proceedings of the eighteenth annual ACM-SIAM symposium on Discrete algorithms, Society for Industrial and Applied Mathematics, 2007, pp. 376–384.
[27] D. Galvin, P. Tetali, On weighted graph homomorphisms, in: DIMACS Series in Discrete Mathematics and Theoretical Computer Science, vol. 63, 2004, pp. 97–104.
[28] D. Galvin, P. Tetali, Slow mixing of Glauber dynamics for the hard-core model on regular bipartite graphs, Random Struct. Algorithms 28 (4) (2006) 427–443.
[29] G.H. Hardy, S. Ramanujan, Asymptotic formulæ in combinatory analysis, Proc. Lond. Math. Soc. 2 (1) (1918) 75–115.
[30] L.H. Harper, Optimal numberings and isoperimetric problems on graphs, J. Comb. Theory 1 (3) (1966) 385–393.

[31] T. Helmuth, M. Jenssen, W. Perkins, Finite-size scaling, phase coexistence, and algorithms for the random cluster model on random graphs, Ann. Inst. Henri Poincaré, Probab. Stat. 59 (2023) 817–848.
[32] T. Helmuth, W. Perkins, G. Regts, Algorithmic Pirogov–Sinai theory, Probab. Theory Relat. Fields (2019) 1–45.
[33] M. Jenssen, P. Keevash, W. Perkins, Algorithms for #BIS-hard problems on expander graphs, SIAM J. Comput. 49 (4) (2020) 681–710.
[34] M. Jenssen, W. Perkins, Independent sets in the hypercube revisited, J. Lond. Math. Soc. (2020).
[35] J. Kahn, An entropy approach to the hard-core model on bipartite graphs, Comb. Probab. Comput. 10 (3) (2001) 219–237.
[36] J. Kahn, Range of cube-indexed random walk, Isr. J. Math. 124 (1) (2001) 189–201.
[37] J. Kahn, A. Lawrenz, Generalized rank functions and an entropy argument, J. Comb. Theory, Ser. A 87 (2) (1999) 398–403.
[38] J. Kahn, J. Park, The number of 4-colorings of the Hamming cube, Isr. J. Math. (2020) 1–21.
[39] A. Korshunov, A. Sapozhenko, The number of binary codes with distance 2, Probl. Kibern. 40 (1983) 111–130.
[40] R. Kotecký, D. Preiss, Cluster expansion for abstract polymer models, Commun. Math. Phys. 103 (3) (1986) 491–498.
[41] V. Leonov, A.N. Shiryaev, On a method of calculation of semi-invariants, Theory Probab. Appl. 4 (3) (1959) 319–329.
[42] C. Liao, J. Lin, P. Lu, Z. Mao, Counting independent sets and colorings on random regular bipartite graphs, arXiv preprint, arXiv:1903.07531, 2019.
[43] N. Linial, R. Meshulam, M. Tarsi, Matroidal bijections between graphs, J. Comb. Theory, Ser. B 45 (1) (1988) 31–44.
[44] R. Peled, High-dimensional Lipschitz functions are typically flat, Ann. Probab. 45 (3) (2017) 1351–1447.
[45] R. Peled, W. Samotij, Odd cutsets and the hard-core model on $\mathbb{Z}^d$, Ann. Inst. Henri Poincaré, Probab. Stat. 50 (2014) 975–998.
[46] R. Peled, W. Samotij, A. Yehudayoff, Grounded Lipschitz functions on trees are typically flat, Electron. Commun. Probab. 18 (2013).
[47] R. Peled, W. Samotij, A. Yehudayoff, Lipschitz functions on expanders are typically flat, Comb. Probab. Comput. 22 (4) (2013) 566–591.
[48] R. Peled, Y. Spinka, A condition for long-range order in discrete spin systems with application to the antiferromagnetic Potts model, arXiv preprint, arXiv:1712.03699, 2017.
[49] R. Peled, Y. Spinka, Long-range order in discrete spin systems, arXiv preprint, arXiv:2010.03177, 2020.
[50] R. Peled, Y. Spinka, Rigidity of proper colorings of $\mathbb{Z}^d$, Invent. Math. (2022) 1–84.
[51] O. Riordan, An ordering on the even discrete torus, SIAM J. Discrete Math. 11 (1) (1998) 110–127.
[52] A. Sapozhenko, On the number of connected subsets with given cardinality of the boundary in bipartite graphs, Metody Diskret. Analiz. 45 (1987) 42–70.
[53] A.D. Scott, A.D. Sokal, The repulsive lattice gas, the independent-set polynomial, and the Lovász local lemma, J. Stat. Phys. 118 (5–6) (2005) 1151–1261.
[54] A. Sinclair, M. Jerrum, Approximate counting, uniform generation and rapidly mixing Markov chains, Inf. Comput. 82 (1) (1989) 93–133.