Independent sets in the hypercube revisited

Jenssen, Matthew; Perkins, Will

INDEPENDENT SETS IN THE HYPERCUBE REVISITED

MATTHEW JENSSEN AND WILL PERKINS

ABSTRACT. We revisit Sapozhenko's classic proof on the asymptotics of the number of independent sets in the discrete hypercube $\{0, 1\}^d$ and Galvin's follow-up work on weighted independent sets. We combine Sapozhenko's graph container methods with the cluster expansion and abstract polymer models, two tools from statistical physics, to obtain considerably sharper asymptotics and detailed probabilistic information about the typical structure of (weighted) independent sets in the hypercube. These results refine those of Korshunov and Sapozhenko and Galvin, and answer several questions of Galvin.

1. INTRODUCTION

Let Q_d denote the discrete hypercube of dimension d: the graph with vertex set $\{0, 1\}^d$ with edges between vectors that differ in exactly one coordinate. An independent set in a graph G is a set of vertices that induce no edges. Let $i(G)$ denote the number of independent sets of G.

Korshunov and Sapozhenko proved the following result on the number of independent sets of the hypercube.

Theorem 1 (Korshunov and Sapozhenko [17])**.**

$$i(Q_d) = (1 + o(1)) \cdot 2\sqrt{e} \cdot 2^{2^{d-1}}$$

as $d \to \infty$.

A beautiful and influential proof of Theorem 1 was later given by Sapozhenko in [22]. See [8] for an exposition of this proof.

One of our main results in this paper will be to reinterpret Sapozhenko's proof in terms of the *cluster expansion* from statistical physics. This allows us to compute additional terms in the asymptotic expansion of $i(Q_d)$ among other things. For instance, we can compute the asymptotics to the third order in 2^{-d}.

Theorem 2.

$$i(Q_d) = 2\sqrt{e} \cdot 2^{2^{d-1}} \left(1 + \frac{3d^2 - 3d - 2}{8 \cdot 2^d} + \frac{243d^4 - 646d^3 - 33d^2 + 436d + 76}{384 \cdot 2^{2d}} + O\left(d^6 \cdot 2^{-3d} \right) \right)$$

as $d \to \infty$.

More generally, we give a formula and an algorithm for computing the asymptotics to arbitrary order in 2^{-d}.

Theorem 1 (along with Sapozhenko's techniques) provided the first glimpse of a rich landscape of phenomena concerning independent sets in Q_d. To describe the phenomena we take the perspective of statistical physics. The independence polynomial of the hypercube is

$$Z(\lambda) = \sum_{I \in \mathcal{I}(Q_d)} \lambda^{|I|},$$

where $\mathcal{I}(Q_d)$ is the set of all independent sets of Q_d. In particular, $Z(1) = i(Q_d)$. The independence polynomial is the partition function of the *hard-core model* from statistical physics: a probability distribution on independent sets weighted by the fugacity parameter λ. This distribution is defined by

$$\mu(I) = \frac{\lambda^{|I|}}{Z(\lambda)}.$$

The hard-core model (or hard-core lattice gas) is a simple model of a gas, and in statistical physics it is most commonly studied on the integer lattice $\mathbb{Z}^d$. As is common in the literature, we will refer to vertices contained in an independent set drawn from the hard-core model as 'occupied'.

Let $\mathcal{E} \subset V(Q_d)$ be the set of 'even' vertices of the hypercube whose coordinates sum to an even number and let $\mathcal{O} \subset V(Q_d)$ be the 'odd' vertices whose coordinates sum to an odd number. We note that Q_d is a bipartite graph with bipartition $(\mathcal{E}, \mathcal{O})$. Kahn [15] showed that for constant λ, typical independent sets drawn from μ contain either mostly even vertices or mostly odd vertices, and thus the hard-core model on Q_d exhibits a kind of 'phase coexistence' in the language of statistical physics.

By generalizing Sapozhenko's techniques, Galvin [7] was able to describe the typical structure of independent sets drawn from μ in greater detail and for a wider range of parameters λ. We need two definitions to describe these results.

Definition 3. *For an independent set $I \in \mathcal{I}(Q_d)$, we say $\mathcal{E}$ is the minority side of the bipartition if $|\mathcal{E} \cap I| < |\mathcal{O} \cap I|$ and the majority side otherwise. If $\mathcal{E}$ is the minority side, then $\mathcal{O}$ is the majority side and vice versa.*

Definition 4. *A set $S \subseteq \mathcal{E}$ (or $\mathcal{O}$), is 2-linked if the subgraph of Q_d induced by the vertex set $S \cup N(S)$ is connected; in other words, S is connected in the graph Q_d^2 (the square of the graph Q_d).*

Galvin showed that for the hard-core model on Q_d at fugacity $\lambda = 1 + s/d$ with s constant, the number of occupied vertices on the minority side is asymptotically distributed as a Poisson random variable with mean $e^{-s/2}/2$ and with high probability (whp) all 2-linked components of occupied vertices on the minority side are of size 1 [7, Theorem 1.4]. He conjectured that there is in fact a series of thresholds at which 2-linked components of size t emerge in a Poisson fashion and asked as an open problem for the distribution of occupied 2-linked components of size t on the minority side for all t.

Here we prove his conjecture and answer his question in a strong form. We show that the emergence of a 2-linked occupied component of size t on the minority side has a sharp threshold at $\lambda_t = 2^{1/t} - 1$ and we identify precisely the scaling window about

this threshold. We also essentially determine the asymptotic joint distribution of the number of such components (see Theorem 6).

Theorem 5. *For $t \geq 1$ fixed, let*

$$\lambda_t(d) = 2^{1/t} - 1 + \frac{2^{1+1/t}(t-1)\log d}{td} + \frac{s(d)}{d} \, .$$

Then for the hard-core model on Q_d at fugacity $\lambda_t(d)$:

- *if $s(d) \to \infty$ as $d \to \infty$ then whp there are no 2-linked occupied components of size t on the minority side;*
- *if $s(d) \to -\infty$ then whp there are 2-linked occupied components of size t on the minority side;*
- *if $s(d)$ tends to a constant s then the distribution of the number of 2-linked occupied components of size t on the minority side converges to a Poisson distribution with mean*

$$e^{-st2^{-1/t}}2^{2-2/t-t}(2^{1/t}-1)^t \sum |Aut\ (T)|^{-1}$$

  where the sum is over all trees T on t vertices and $Aut(T)$ denotes the automorphism group of the tree T.

In fact we prove much more detailed probabilistic results. Define the *defect type* of a 2-linked component S of $\mathcal{E}$ or $\mathcal{O}$ to be the isomorphism class of the induced subgraph $Q_d^2[S]$. In particular there is a unique defect type of size 1 (an isolated vertex), a unique defect type of size 2 (two vertices at distance 2 in Q_d), but two defect types of size 3: 3 vertices whose distance-2 graph forms a clique and 3 vertices whose distance-2 graph forms a path. For a given defect type T, let X_T be the random variable that counts the number of 2-linked occupied components of type T on the minority side in the hard-core model on Q_d. Let $m_T = \mathbb{E}X_T$ and $\sigma_T^2 = \text{var}(X_T)$ (for the asymptotics of m_T and σ_T^2 see Lemma 19 and Corollary 21 below).

We determine the limiting distribution of the number of each type of defect and show that the number of defects of different types are asymptotically independent.

Theorem 6. *There is a constant $C_0 > 0$ such that if $\lambda \geq C_0 \log d/d^{1/3}$ and T is a defect type then the following holds. If T and λ are such that $m_T \to \rho$ as $d \to \infty$ for some constant $\rho > 0$, then*

$$X_T \Rightarrow Pois(\rho) \, ,$$

where '$\Rightarrow$' denotes convergence in distribution. If T and λ are such that $m_T \to \infty$ as $d \to \infty$, then

$$\tilde{X}_T = \frac{X_T - m_T}{\sigma_T} \Rightarrow N(0,1) \, .$$

Moreover, suppose we have two finite sets of defect types $\mathcal{T}_1, \mathcal{T}_2$ so that for each $T \in \mathcal{T}_1$, there exists $\rho_T > 0$ so that $m_T \to \rho_T$, and for each $T \in \mathcal{T}_2$, $m_T \to \infty$. Then the collection of random variables $\{X_T\}_{T \in \mathcal{T}_1} \cup \{\tilde{X}_T\}_{T \in \mathcal{T}_2}$ converges in distribution to a collection of independent Poisson and standard normal random variables.

We remark that the condition that $\lambda \geq C_0 \log d / d^{1/3}$ is a technical requirement of a container lemma due to Galvin which is a key ingredient in our proofs (see Lemma 11 below). We expect that Theorem 6 in fact extends to the range $\lambda > (1 + \Omega(1)) \log d / d$.

There is a close connection between computing accurate estimates of the partition function and deriving probabilistic information about the hard-core model. As a key step in proving his probabilistic results, Galvin gave a significant generalization of Theorem 1 to counting weighted independent sets in the hypercube; that is, computing the asymptotics of $Z(\lambda)$ for general λ.

Theorem 7 (Galvin [7]). *For $\lambda \geq \sqrt{2} - 1 + \frac{(\sqrt{2} + \Omega(1)) \log d}{d}$,*

$$
(1) \qquad Z(\lambda) = (2 + o(1)) \cdot \exp\left[\frac{\lambda}{2} \left(\frac{2}{1 + \lambda} \right)^d \right] \cdot (1 + \lambda)^{2^{d-1}} \, .
$$

Moreover, there is a constant $C_0 > 0$ so that for $\lambda \geq C_0 \log d / d^{1/3}$,

$$
(2) \qquad Z(\lambda) = 2(1 + \lambda)^{2^{d-1}} \cdot \exp\left(\frac{\lambda}{2} \left(\frac{2}{1 + \lambda} \right)^d (1 + o(1)) \right) \, .
$$

The formula (1) generalises Theorem 1, and determines the asymptotics of $Z(\lambda)$ for $\lambda > \sqrt{2} - 1$, while the formula (2) finds the asymptotics of $\log Z(\lambda)$ for $\lambda = \Omega(\log d / d^{1/3})$.

Our techniques based on the cluster expansion will allow us to sharpen Theorem 7 considerably: we find a formula that can be used not only to determine the asymptotics of $Z(\lambda)$ for all constant λ but also to give an expansion of $\log Z(\lambda)$ to arbitrary order in 2^{-d}.

To write the formula we need some notation that comes from *polymer models* in the statistical mechanics of lattice systems [18]. Before we introduce these notions formally, we describe some of the intuition underlying the proof of Theorem 2 and the results to come. An immediate lower bound on $Z(\lambda)$ of $2(1 + \lambda)^{2^{d-1}} - 1$ comes by considering the contribution from independent sets which lie entirely in one side of the bipartition of Q_d. We call the collection of independent sets which lie entirely in $\mathcal{E}$ (or $\mathcal{O}$) the even (odd) *ground state*. Taking $\lambda = 1$, for example, there is a constant factor gap between this trivial lower bound $i(Q_d) \geq 2 \cdot 2^{2^{d-1}} - 1$ and the correct asymptotics of Theorem 1. Therefore a constant proportion of independent sets do not belong to a ground state. However, almost all independent sets are very close to a ground state independent set. Thus it is natural to describe independent sets in terms of their deviations from a ground state: given a subset $X \subseteq \mathcal{E}$, let $p(X)$ denote the probability that an independent set I chosen according to μ satisfies $I \cap \mathcal{E} = X$. When X is small, we think of it as a deviation from the odd ground state and note that the relative 'cost' of such a deviation is

$$
(3) \qquad \frac{p(X)}{p(\emptyset)} = \frac{\lambda^{|X|}}{(1 + \lambda)^{|N(X)|}} \, .
$$

We denote this cost, or *weight*, of a deviation X by $w(X)$. Crucially, the weight $w(X)$ factorises over the 2-linked components of X, and so we define an *even polymer* to be any 2-linked subset of $\mathcal{E}$, and define its weight by (3). We define odd polymers similarly.

The language of polymer models allows us to relate the partition function $Z(\lambda)$ to the partition function of a (multivariate) hard-core model on an auxiliary graph whose

vertices are polymers and each polymer S has its associated weight $w(S)$ as its fugacity. A key feature of this transformation is that while at large λ an independent set drawn from $\mu_{Q_d,\lambda}$ is typically very structured, the corresponding deviations on the minority side are typically unstructured and behave almost independently. Using the cluster expansion, we are then able to extract almost complete probabilistic information from our model. In particular it allows us to precisely quantify the contribution to $Z(\lambda)$ from small deviations, and allows us to compute $\log Z(\lambda)$ to essentially arbitrary accuracy.

The cluster expansion is a powerful and classical tool in the rigorous study of statistical mechanics. In our context, it is the multivariate Taylor expansion of the logarithm of the partition function of our auxiliary hard-core model. Studying this infinite series naturally leads to the question of convergence. Verifying the convergence of the cluster expansion amounts to showing that the number of polymers of a given weight is not too large. This is where the container method of Sapozhenko comes in. In fact, all of the ingredients needed to show that this polymer model has a convergent cluster expansion are already present in Sapozhenko's work and Galvin's extensions. In some sense Sapozhenko rediscovered the concept of a polymer model and computed the smallest order terms of the cluster expansion by hand. Certainly the intuition behind the specific polymer model is clear in his work.

We now venture to make some of the above mentioned notions more concrete. Recall that an even/odd polymer is a 2-linked subset of $\mathcal{E}/\mathcal{O}$ respectively. The size of a polymer S, $|S|$, is the number of vertices in S. Since Q_d exhibits symmetry between $\mathcal{E}$ and $\mathcal{O}$ we will restrict our attention to even polymers. We say two even polymers S_1, S_2 are *compatible* if $d_G(S_1, S_2) > 2$; that is if $S_1 \cup S_2$ is not 2-linked. Otherwise S_1 and S_2 are incompatible (and note that each polymer is incompatible with itself). For a tuple Γ of even polymers, the *incompatibility graph*, $H(\Gamma)$, is the graph with vertex set Γ and an edge between any two incompatible polymers. An even *cluster* Γ is an ordered tuple of even polymers so that $H(\Gamma)$ is connected. The *size* of a cluster Γ is $\|\Gamma\| = \sum_{S \in \Gamma} |S|$. Let $\mathcal{C}$ be the set of all even clusters and $\mathcal{C}_k$ the set of all even clusters of size k.

Recall that for a polymer S, we define its weight to be $w(S) = \lambda^{|S|}(1+\lambda)^{-|N(S)|}$. For a cluster Γ we define

$$w(\Gamma) = \phi(H(\Gamma)) \prod_{S \in \Gamma} w(S),$$

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

$$(4) \qquad \phi(H) = \frac{1}{|V(H)|!} \sum_{\substack{A \subseteq E(H) \\ \text{spanning, connected}}} (-1)^{|A|}.$$

Finally for $k \geq 1$ we define

$$L_k = \sum_{\Gamma \in \mathcal{C}_k} w(\Gamma).$$

Note that by symmetry L_k would be identical if we had considered odd polymers and odd clusters instead.

We can now state our formula for $Z(\lambda)$.

Theorem 8. *Suppose $\lambda \geq C_0 \log d / d^{1/3}$ and λ is bounded as $d \to \infty$. Then for all fixed $k \geq 1$,*

$$Z(\lambda) = 2(1+\lambda)^{2^{d-1}} \cdot \exp\left(\sum_{j=1}^{k} L_j + \varepsilon_k\right)$$

where for each fixed k, $|\varepsilon_k| = O\left(\frac{2^d \lambda^{k+1} d^{2k}}{(1+\lambda)^{d(k+1)}}\right)$ as $d \to \infty$. Moreover, L_k can be computed in time $e^{O(k \log k)}$.

In fact, it is not essential that λ remain bounded as $d \to \infty$: a similar formula holds for all values of λ with an addition of $\exp(-2^d/d^4)$ to ε_k, but for simplicity here we focus on the more interesting cases when λ is bounded or tends to 0.

As a quick check, note that at $\lambda = 1$, $L_1 = 1/2$ since there are 2^{d-1} polymers of size 1 and each has weight 2^{-d}. Moreover, $\varepsilon_2 = O(d^4 2^{-2d}) = o(1)$ and so Theorem 8 implies that $i(Q_d) = 2 \cdot 2^{2^{d-1}} e^{1/2+o(1)}$, recovering Theorem 1.

More generally, Theorem 8 extends Theorem 7. For instance, we can give a closed-form formula for the asymptotics of $Z(\lambda)$ for any constant λ.

Corollary 9. *For any fixed $t \geq 1$ and for $\lambda \geq 2^{1/t} - 1 + \frac{2^{1+1/t}(t-1)\log d}{td} + \frac{\omega(1)}{d}$,*

$$(5) \qquad Z(\lambda) = (2 + o(1))2^{2^{d-1}} \exp\left(\sum_{j=1}^{t-1} L_j\right).$$

For example, if $\lambda \geq 2^{1/3} - 1 + \frac{2^{7/3} \log d}{3d} + \frac{\omega(1)}{d}$, then

$$(6) \quad Z(\lambda) = (2 + o(1)) \cdot \exp\left[\frac{\lambda}{2}\left(\frac{2}{1+\lambda}\right)^d \left(1 + \frac{(2\lambda^2 + \lambda^3)d(d-1) - 2}{4(1+\lambda)^d}\right)\right] (1+\lambda)^{2^{d-1}}.$$

We find it rather remarkable how well the two tools from statistical physics, polymer models and the cluster expansion, work with the graph container method, and we expect many further applications of this combination of methods. See [21] for a survey of the graph container method. In forthcoming work, Keevash and the first author [13] apply this combination of methods to resolve conjectures of Galvin and Engbers [5] and Kahn and Park [16] on the number of q-colourings of Q_d. As a future research direction, we ask whether these statistical physics tools can be used in conjunction with the method of hypergraph containers [1, 23] to derive finer asymptotics and probabilistic information in some of the many extremal combinatorics problems in which hypergraph containers have been deployed.

The paper is organised as follows: We introduce abstract polymer models and the cluster expansion in Section 2, and then specialise to the hypercube and prove Theorem 8 in Section 3. We prove the probabilistic results of Theorems 5 and 6 in Section 4. We explicitly compute L_1, L_2, and L_3 and prove Theorem 2 in Section 5.

Related work. As Galvin remarked in [7], only a few properties of the hypercube Q_d are needed in deriving Theorem 7; the same is true for Theorems 6 and 8. The essential properties are that the graph be bipartite and that some isoperimetric estimates hold (of the form of Lemma 12 below). In fact, using an approach to approximate counting based

on the cluster expansion [12, 14], one could obtain efficient algorithms to approximate the partition function $Z_G(\lambda)$ and to sample from the hard-core model for a class of graphs with these properties. The polymer models used in [14, 20, 3] to sample from the hard-core model on random regular bipartite graphs are very similar to the ones used here. For a similar class of bipartite graphs Galvin and Tetali [10] showed that the Glauber dynamics Markov chain for sampling from the hard-core model exhibits slow mixing; that proof is also based on extending the ideas of Sapozhenko.

2. Polymer models and the cluster expansion

Here we introduce the main tools we will use, abstract polymer models [11, 18] and the cluster expansion, both tools from statistical physics that have been used extensively to study phase diagrams of lattice spin models. 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 $w(S)$ for each polymer S as well as a symmetric and reflexive incompatibility relation between polymers. We write $S \nsim S'$ if polymers S and S' 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

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

where the contribution from the empty set is 1.

A *cluster* is an ordered tuple of polymers whose incompatibility graph $H(\Gamma)$ is connected. Let $\mathcal{C}$ be the set of all clusters. The *cluster expansion* is the formal power series in the weights $w(S)$

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

where

$$w(\Gamma) = \phi(H(\Gamma)) \prod_{S \in \Gamma} w(S)\,,$$

and $\phi(H)$ is the Ursell function as defined in (4). In fact the cluster expansion is simply the multivariate Taylor series for $\log \Xi(\mathcal{P})$ in the variables $w(S)$, as observed by Dobrushin [4]. See also Scott and Sokal [24] 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 10 ([18]). *Let $f : \mathcal{P} \to [0, \infty)$ and $g : \mathcal{P} \to [0, \infty)$ be two functions. Suppose that for all polymers $S \in \mathcal{P}$,*

$$(7) \qquad \sum_{S' \nsim S} |w(S')| e^{f(S') + g(S')} \leq f(S)\,,$$

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

$$(8) \qquad \sum_{\substack{\Gamma \in \mathcal{C} \\ \Gamma \nsim S}} |w(\Gamma)| \, e^{g(\Gamma)} \leq f(S) \,.$$

We remark that one could simply take $g \equiv 0$ in (7) in order to establish convergence of the cluster expansion. However, allowing g to take non-zero values (thus strengthening (7)) allows us to give strong tail bounds on the cluster expansion via (8). 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. Polymers in the hypercube

We now return to our specific setting with polymers derived from the hard-core model on Q_d. These polymers will essentially be the same as those defined in Section 1. Here we will study the cluster expansion of this polymer model in depth.

3.1. **Preliminaries.** We begin with some notation and lemmas from [7].

For a set $A \subseteq \mathcal{E}$ (and analogously for $A \subseteq \mathcal{O}$), let $|A|$ denote the number of vertices of A, $N(A)$ be the set of neighbours of A, and $[A] = \{v \in \mathcal{E} : N(v) \subseteq N(A)\}$ the bipartite closure of A. Clearly $|[A]| \geq |A|$. Let

$$\mathcal{G}(a,b) = \{A \subseteq \mathcal{E} : A \text{ 2-linked}, |[A]| = a, |N(A)| = b\}.$$

The following lemma of Galvin is based on the graph container method of Sapozhenko [22]. This is a key technical ingredient in [22, 7] and in the results of this paper.

Lemma 11 ([7]). *There exist constants $C_0, C_1 > 0$, so that for all $\lambda \geq C_0 \log d / d^{1/3}$, all $a \leq 2^{d-2}$,*

$$\sum_{A \in \mathcal{G}(a,b)} \frac{\lambda^{|A|}}{(1+\lambda)^b} \leq 2^d \exp\left(-\frac{C_1(b-a) \log d}{d^{2/3}} \right) \,.$$

In what follows, we will always assume that $\lambda \geq C_0 \log d / d^{1/3}$ to allow us to apply Lemma 11.

We will also use the following isoperimetric estimates, which come from [6, 17] but can also be found in [7].

Lemma 12. *Suppose $S \subseteq \mathcal{E}$ (or $S \subseteq \mathcal{O}$). Then*

  (1) If $|S| \leq d/10$, then $|N(S)| \geq d|S| - 2|S|^2$.
  (2) If $|S| \leq d^4$, then $|N(S)| \geq d|S|/10$.
  (3) If $|S| \leq 2^{d-2}$, then $|N(S)| \geq \left(1 + \frac{1}{2\sqrt{d}}\right)|S|$.

We also make use of the following, from, e.g. [9].

Lemma 13. *The number of 2-linked subsets $S \subseteq \mathcal{E}$ of size t which contain a given vertex v is at most $(ed^2)^{t-1}$.*

3.2. **The defect polymer model.** We begin by fixing a side of the bipartition which we call the *defect side*. Let us suppose this side is $\mathcal{E}$ (the case where $\mathcal{O}$ is the defect side will be identical).

We define a polymer to be a 2-linked subset S of the defect side in Q_d so that $|[S]| \leq 2^{d-2}$. Let $\mathcal{P}$ be the set of all such polymers (we will make use of a subscript, as in $\mathcal{P}_\mathcal{E}$ or $\mathcal{P}_\mathcal{O}$, if we want to indicate which is the defect side). Two polymers S, S' are compatible if $S \cup S'$ is not 2-linked. Let Ω be the set of all pairwise compatible sets of polymers from $\mathcal{P}$. The weight functions are defined as

$$w(S) = \frac{\lambda^{|S|}}{(1+\lambda)^{|N(S)|}} \ .$$

Let $\Xi = \Xi(\mathcal{P})$ denote the resulting polymer model partition function (and note that by symmetry Ξ is the same regardless of the defect side).

The partition function Ξ is the normalizing constant of a probability distribution ν on Ω defined by

$$\nu(\Gamma) = \frac{\prod_{S \in \Gamma} w(S)}{\Xi(\mathcal{P})} \ .$$

Using ν we can define a probability measure $\hat{\mu}$ on $\mathcal{I}(Q_d)$ as follows:

(1) With probability $1/2$ choose $\mathcal{D} = \mathcal{E}$ or $\mathcal{D} = \mathcal{O}$ to be the defect side.
(2) Choose a polymer configuration $\Gamma \in \Omega_\mathcal{D}$ from ν and assign all vertices of $\cup_{S \in \Gamma} S$ to be occupied on the defect side $\mathcal{D}$.
(3) For each vertex v on the non-defect side that is not blocked by an occupied vertex on the defect side, include v in the independent set independently with probability $\frac{\lambda}{1+\lambda}$.

The resulting distribution $\hat{\mu}$ is not exactly the hard-core model μ on Q_d, but we will show that the two distributions are very close in total variation distance. Moreover, we will show that a scaling of the partition function Ξ is a very good approximation of the hard-core partition function $Z(\lambda)$. Note that the defect side need not be the minority side: in step 3 we may choose no vertices to be occupied opposite the defect side. Nevertheless, we will show below that with very high probability the defect side is in fact the minority side of an independent set sampled according to $\hat{\mu}$ (Lemma 17 below).

Lemma 14. *We have*

(9) $$\left| \log Z(\lambda) - \log \left[2(1+\lambda)^{2^{d-1}} \Xi \right] \right| = O\left(\exp(-2^d/d^4) \right) \ .$$

Moreover,

$$\|\hat{\mu} - \mu\|_{TV} = O\left(\exp(-2^d/d^4) \right) \ .$$

We will prove Lemma 14 after showing that the polymer model satisfies the Kotecký–Preiss condition. Lemma 14 allows us to work with Ξ and ν to prove Theorems 6 and 8. In particular, to prove Theorem 8 we will approximate Ξ by truncating the cluster expansion for $\log \Xi$ and exponentiating. To prove Theorem 6 we will prove the probabilistic statements for polymer configurations sampled from ν and then use

Lemmas 14 and 17 to transfer these results to results about the minority side of an independent set drawn from μ.

We define the truncated cluster expansion of $\log \Xi$ as

$$T_k = \sum_{\substack{\Gamma \in \mathcal{C}: \\ \|\Gamma\| < k}} w(\Gamma) \,.$$

We now show that condition (7) holds for the defect polymer model with appropriate choices of functions $f(\cdot)$ and $g(\cdot)$, and thus T_k gives a good approximation to $\log \Xi$.

Lemma 15. *For integers $d, k \geq 1$, let*

$$\gamma(d,k) = \begin{cases} \log(1+\lambda)(dk - 3k^2) - 7k \log d & \text{if } k \leq \frac{d}{10} \\ \frac{d \log(1+\lambda)k}{20} & \text{if } \frac{d}{10} < k \leq d^4 \\ \frac{k}{d^{3/2}} & \text{if } k > d^4 \,. \end{cases}$$

Then for d sufficiently large

$$(10) \qquad \sum_{\substack{\Gamma \in \mathcal{C} \\ \|\Gamma\| \geq k}} |w(\Gamma)| \leq d^{-3/2} 2^{d-1} e^{-\gamma(d,k)} \,.$$

In particular, for k fixed and d sufficiently large,

$$(11) \qquad |T_k - \log \Xi| \leq d^{7k - 3/2} 2^d (1+\lambda)^{-dk + 3k^2} \,.$$

Proof. Let $g : \mathcal{P} \to [0, \infty)$ be defined by $g(S) = \gamma(d, |S|)$ and define $f : \mathcal{P} \to [0, \infty)$ by $f(S) = |S|/d^{3/2}$. We will show that the Kotecký–Preiss condition (7) holds. That is, for every $S \in \mathcal{P}$,

$$(12) \qquad \sum_{S' \not\sim S} w(S') e^{d^{-3/2}|S'| + g(S')} \leq |S|/d^{3/2} \,.$$

To prove this we will show that for all $v \in \mathcal{E}$,

$$\sum_{S \ni v} w(S) e^{d^{-3/2}|S| + g(S)} \leq \frac{1}{d^{7/2}} \,,$$

and this will suffice since $S' \not\sim S$ if and only if $S' \ni v$ for some $v \in N^2(S)$ and $|N^2(S)| \leq d^2|S|$. We will break up the sum according to the different cases of $\gamma(d, k)$.

First we sum over S with $|S| \leq \frac{d}{10}$. We use the fact that for such S, $|N(S)| \geq d|S| - 2|S|^2$ by Lemma 12, and that there are at most $\exp(3k \log d)$ 2-linked sets S of size k that contain a fixed vertex v by Lemma 13.

$$\begin{aligned} \sum_{\substack{S \ni v \\ |S| \leq \frac{d}{10}}} w(S) e^{f(S) + g(S)} &\leq \sum_{k=1}^{d/10} e^{3k \log d} \frac{\lambda^k}{(1+\lambda)^{dk - 2k^2}} e^{kd^{-3/2} + \log(1+\lambda)(dk - 3k^2) - 7k \log d} \\ &\leq \sum_{k \geq 1} \exp\left(3k \log d + k \log \lambda + kd^{-3/2} - k^2 \log(1+\lambda) - 7k \log d\right) \\ &\leq \sum_{k \geq 1} \exp\left(-4k \log d + kd^{-3/2}\right) \end{aligned}$$

which is at most $\frac{1}{3d^{7/2}}$ for d large enough.

We next sum over S with $\frac{d}{10} < |S| \leq d^4$. We use the fact that for such S, $|N(S)| \geq d|S|/10$ by Lemma 12.

$$\sum_{\substack{S \ni v \\ d/10 < |S| \leq d^4}} w(S)e^{f(S)+g(S)} \leq \sum_{k=d/10}^{d^4} e^{3k \log d} \frac{\lambda^k}{(1+\lambda)^{dk/10}} e^{kd^{-3/2}+dk\log(1+\lambda)/20}$$

$$= \sum_{k=d/10}^{d^4} \exp\left(k \left(d^{-3/2} + \log \lambda + 3\log d - \frac{d\log(1+\lambda)}{20} \right) \right),$$

and so if $\lambda \geq C_0 \log d/d$ and d is large enough, then this sum is at most $\frac{1}{3d^{7/2}}$.

Now turning to S with $d^4 < |S| \leq 2^{d-2}$, we have that $|N(S)| \geq |S|(1 + 1/(2\sqrt{d}))$, and so

$$\sum_{\substack{S \ni v \\ d^4 < |S| \leq 2^{d-2}}} w(S)e^{f(S)+g(S)} = \sum_{\substack{d^4 < a \leq 2^{d-2} \\ (1+1/(2\sqrt{d}))a \leq b \leq 2^{d-1}}} \sum_{\substack{S \ni v \\ |[S]|=a,|N(S)|=b}} \frac{\lambda^{|S|}}{(1+\lambda)^b} e^{2|S|d^{-3/2}}$$

$$\leq \sum_{\substack{d^4 < a \leq 2^{d-2} \\ (1+1/(2\sqrt{d}))a \leq b \leq 2^{d-1}}} e^{2ad^{-3/2}} \sum_{\substack{S \ni v \\ |[S]|=a,|N(S)|=b}} \frac{\lambda^{|S|}}{(1+\lambda)^b}$$

$$\leq \sum_{\substack{a > d^4 \\ b \geq (1+1/(2\sqrt{d}))a}} e^{2ad^{-3/2}} d \exp\left(-\frac{C_1(b-a)\log d}{d^{2/3}} \right),$$

where the last inequality comes from applying Lemma 11. In the sum, we have $(b-a) \geq a/(2\sqrt{d})$ and $a > d^4$, and so

$$\frac{2a}{d^{3/2}} + \log d - \frac{C_1(b-a)\log d}{d^{2/3}} \leq -ad^{-7/6}$$

for large enough d, and so

$$\sum_{\substack{S \ni v \\ d^4 < |S| \leq 2^{d-2}}} w(S)e^{f(S)+g(S)} \leq \sum_{\substack{d^4 < a \leq 2^{d-2} \\ (1+1/(2\sqrt{d}))a \leq b \leq 2^{d-1}}} \exp(-ad^{-7/6})$$

$$\leq 2^d \sum_{a > d^4} \exp(-ad^{-7/6})$$

$$\leq \frac{1}{3d^{7/2}}$$

for d large enough. Putting the three bounds together gives (12).

To prove the lemma we now apply Theorem 10, applying (8) for the polymer S containing the single vertex v to obtain:

$$\sum_{\Gamma \in \mathcal{C}, \Gamma \nsim v} |w(\Gamma)|e^{g(\Gamma)} \leq d^{-3/2}.$$

Summing over all v gives

$$\text{(13)} \qquad \sum_{\Gamma \in \mathcal{C}} |w(\Gamma)| e^{g(\Gamma)} \leq 2^{d-1} d^{-3/2} \,.$$

Since $\gamma(d,k)/k$ is non-increasing in k, we have, recalling that $g(\Gamma) = \sum_{S \in \Gamma} g(S)$ and $\|\Gamma\| = \sum_{S \in \Gamma} |S|$,

$$g(\Gamma) = \sum_{S \in \Gamma} \gamma(d, |S|) \geq \sum_{S \in \Gamma} \frac{|S|}{\|\Gamma\|} \gamma(d, \|\Gamma\|) = \gamma(d, \|\Gamma\|) \,.$$

Keeping only terms in inequality (13) corresponding to clusters of size at least k, we have

$$\sum_{\substack{\Gamma \in \mathcal{C} \\ \|\Gamma\| \geq k}} |w(\Gamma)| \leq d^{-3/2} 2^{d-1} e^{-\gamma(d,k)}$$

as desired. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

The Kotecký–Preiss condition also allows us to prove a simple large deviation result for the total size of all polymers in a random polymer configuration drawn from ν.

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 make extensive use of the *cumulant generating function* of X, defined as

$$h_t(X) = \log \mathbb{E}e^{tX} \,,$$

that is, the logarithm of the moment generating function.

Lemma 16. *Let $\mathbf{\Gamma}$ be a random configuration drawn from the distribution ν. Then with probability at least $1 - O\left(\exp(-2^d/d^4)\right)$, we have*

$$\|\mathbf{\Gamma}\| \leq 2^d/d^2 \,,$$

where $\|\mathbf{\Gamma}\| = \sum_{S \in \mathbf{\Gamma}} |S|$.

Proof. We introduce an auxiliary polymer model with modified polymer weights:

$$\tilde{w}(S) = w(S) e^{|S| d^{-3/2}} \,.$$

Let $\tilde{\Xi}$ be the associated polymer model partition function. Then $\log \tilde{\Xi} - \log \Xi = h_t(\|\mathbf{\Gamma}\|)$ at $t = d^{-3/2}$ where $\mathbf{\Gamma}$ is a random polymer configuration from the original polymer model.

In the proof of Lemma 15, all of the estimates hold if we were to replace $f(S) = |S|/d^{3/2}$ by $\tilde{f}(S) = 2|S|/d^{3/2}$. Therefore the proof shows that the Kotecký–Preiss condition holds for the polymer weights $\tilde{w}(S)$, and the functions $f(S), g(S)$ as above. Applying (8) and summing over all polymers of size 1 gives

$$\begin{aligned} \log \tilde{\Xi} &\leq \sum_{\Gamma \in \mathcal{C}} |\tilde{w}(\Gamma)| \\ &\leq 2^{d-1} d^{-3/2} e^{-\gamma(d,1)} \\ &\leq 2^{d-1} d^{11/2} (1+\lambda)^{3-d} \,. \end{aligned}$$

Then since $\Xi \geq 1$, we have

$$h_{d^{-3/2}}(\|\mathbf{\Gamma}\|) \leq \log \tilde{\Xi}$$
$$\leq 2^{d-1} d^{11/2} (1+\lambda)^{3-d} .$$

By Markov's inequality we have

$$\Pr[\|\mathbf{\Gamma}\| > 2^d/d^2] \leq e^{-t 2^d/d^2} \mathbb{E} e^{t\|\mathbf{\Gamma}\|} ,$$

and setting $t = d^{-3/2}$ gives

$$\Pr[\|\mathbf{\Gamma}\| > 2^d/d^2] \leq \exp\left[-\frac{2^d}{d^{7/2}} + \frac{d^{11/2} 2^{d-1}}{(1+\lambda)^{d-3}} \right]$$
$$\leq \exp(-2^d/d^4)$$

for large enough d since for $\lambda \geq C_0 \log d / d^{1/3}$, $(1+\lambda)^d$ grows faster than any fixed polynomial in d. $\qquad\square$

This large deviation bound allows us to show that with very high probability over an independent set drawn from $\hat{\mu}$, the defect side is the minority side.

Lemma 17. *With probability at least $1 - O\left(\exp(-2^d/d^4)\right)$ over the random independent set $\mathbf{I}$ drawn from $\hat{\mu}$, the minority side of the bipartition is the defect side.*

Proof. Let $\mathcal{D}$ and $\mathcal{M}$ denote the defect and minority side respectively selected under $\hat{\mu}$. By Lemma 16 we have

$$\Pr[\mathcal{M} \neq \mathcal{D}] \leq \Pr[\mathcal{M} \neq \mathcal{D} \mid \|\mathbf{\Gamma}\| \leq 2^d/d^2] + O\left(\exp(-2^d/d^4)\right) .$$

Let us fix an element Γ in the sample space of $\mathbf{\Gamma}$ with $\|\Gamma\| \leq 2^d/d^2$ and let $V(\Gamma) := \bigcup_{S \in \Gamma} S$. Let $\mathbf{X}$ denote the size of the intersection of $\mathbf{I}$ with the non-defect side conditioned on the event that $\mathbf{\Gamma} = \Gamma$. Note that $\mathbf{X}$ has a $\mathrm{Bin}(M, \lambda/(1+\lambda))$ distribution where $M = 2^{d-1} - |N(V(\Gamma))| \geq (1 - 2/d) 2^{d-1}$. We then have

$$\Pr[\mathcal{M} \neq \mathcal{D} \mid \mathbf{\Gamma} = \Gamma] \leq \Pr[\mathbf{X} \leq \|\Gamma\|]$$
$$\leq \Pr[\mathbf{X} \leq 2^d/d^2]$$
$$\leq \Pr[\mathbf{X} \leq \mathbb{E}\mathbf{X}/2]$$
$$\leq e^{-\mathbb{E}\mathbf{X}/8}$$
$$\leq \exp(-2^d/d^4) ,$$

for d sufficiently large. For the penultimate inequality we applied the Chernoff bound. The result follows. $\qquad\square$

Now we can prove Lemma 14.

Proof of Lemma 14. We say an independent set I is captured by the odd polymer model if every 2-linked component S of $\mathcal{O} \cap I$ has $|[S]| \leq 2^{d-2}$ and captured by the even polymer model if every 2-linked component S of $\mathcal{E} \cap I$ has $|[S]| \leq 2^{d-2}$. If we view $2(1+\lambda)^{2^{d-1}} \Xi$ as the sum of $(1+\lambda)^{2^{d-1}} \Xi$ for Ξ representing the odd polymer model and $(1+\lambda)^{2^{d-1}} \Xi$ for Ξ representing the even polymer model, then each I that is captured by the odd

polymer model contributes $\lambda^{|I|}$ to the first summand and each I that is captured by the even polymer model contributes $\lambda^{|I|}$ to the second summand.

Observe first that every $I \in \mathcal{I}(Q_d)$ is captured by either the odd or the even polymer model. Indeed suppose not, then there exists $I \in \mathcal{I}(Q_d)$ which contains a set $S \subseteq \mathcal{O}$ with $|[S]| > 2^{d-2}$ and a set $S' \subseteq \mathcal{E}$ with $|[S']| > 2^{d-2}$. It follows that $|N(S)| = |N([S])| > 2^{d-2}$ (since, for example, Q_d contains a perfect matching). However then $N(S) \cap [S'] \neq \emptyset$ and so $S \cap N(S') = S \cap N([S']) \neq \emptyset$, contradicting the fact that I is an independent set.

It remains to bound the contribution to $2(1+\lambda)^{2^{d-1}} \Xi$ from independent sets that are counted twice. That is, bound $\sum_{I \in B} \lambda^{|I|}$ where B denotes the collection of independent sets that are captured by both the odd and even polymer models. However, any such independent set can be selected by $\hat{\mu}$ conditioned on the event that $\mathcal{M} \neq \mathcal{D}$ (using the notation of Lemma 17). Letting $\mathbf{I}$ denote the independent set selected by $\hat{\mu}$ we have by Lemma 17 that

$$(14) \qquad \Pr[\mathbf{I} \in B \wedge \mathcal{M} \neq \mathcal{D}] = \frac{\sum_{I \in B} \lambda^{|I|}}{2(1+\lambda)^{2^{d-1}} \Xi} = O\left(\exp(-2^d/d^4)\right).$$

All together this gives the inequalities

$$\left(1 - O\left(\exp(-2^d/d^4)\right)\right) 2(1+\lambda)^{2^{d-1}} \Xi \leq Z(\lambda) \leq 2(1+\lambda)^{2^{d-1}} \Xi,$$

and so

$$\log[2(1+\lambda)^{2^{d-1}} \Xi] - O(\exp(-2^d/d^4)) \leq \log Z(\lambda) \leq \log[2(1+\lambda)^{2^{d-1}} \Xi],$$

which gives (9). Recall one formula for the total variation distance between discrete probability measures:

$$\|\mu - \hat{\mu}\|_{TV} = \sum_{I : \hat{\mu}(I) > \mu(I)} \hat{\mu}(I) - \mu(I).$$

The total variation distance bound is then immediate from (14) as the only independent sets that have higher probability under $\hat{\mu}$ than μ are those that are counted twice. $\square$

Now we can prove Theorem 8.

Proof of Theorem 8. First we prove the estimate $|L_r| = O\left(\frac{2^d \lambda^r d^{2(r-1)}}{(1+\lambda)^{dr-r^2}}\right)$ for r fixed. Let Γ be a cluster with $\|\Gamma\| = r$. Since $V(\Gamma) := \bigcup_{S \in \Gamma} S$ is a 2-linked set of size at most r, there are $O(2^d d^{2(r-1)})$ possibilities for $V(\Gamma)$ by Lemma 13. Given a set $X \subseteq V(Q_d)$ of size at most r, there are at most a constant number of clusters Γ of size r such that $V(\Gamma) = X$. It follows that the number of clusters of size r is $O(2^d d^{2(r-1)})$. By Lemma 12, the weight of any cluster of size r is $O(\lambda^r/(1+\lambda)^{dr-r^2})$ (note that the Ursell function of a cluster of size r is simply a constant). The claimed estimate on $|L_r|$ follows.

Let $k \geq 1$ be fixed. By (11) we have that

$$\varepsilon'_k := |T_{k+2} - \log \Xi| \leq d^{7(k+2)-3/2} 2^d (1+\lambda)^{-d(k+2)+3(k+2)^2}$$

for d sufficiently large and so

$$\Xi = \exp\left\{\sum_{j=1}^{k+1} L_j + \varepsilon'_k\right\}.$$

It follows from Lemma 14 that

$$Z(\lambda) = 2(1+\lambda)^{2^{d-1}} \exp\left\{ \sum_{j=1}^{k} L_j + L_{k+1} + \varepsilon_k' + O(\exp(-2^d/d^4)) \right\}$$

$$= 2(1+\lambda)^{2^{d-1}} \exp\left\{ \sum_{j=1}^{k} L_j + \varepsilon_k \right\},$$

where $|\varepsilon_k| = O\left(\frac{2^d \lambda^{k+1} d^{2k}}{(1+\lambda)^{d(k+1)}} \right)$ (it is here we use that λ is bounded as $d \to \infty$).

Finally we show that L_k can be computed in time $e^{O(k \log k)}$. Let X be the family of all 2-linked subsets of $\mathcal{E}$ of size at most k which contain the vertex $\underline{0} = (0, \ldots, 0)$. Given $S \in X$, we call a coordinate i *active* for S if $x_i = 1$ for some $x \in S$. We note that every $S \in X$ has at most $2k$ active coordinates. For $A \subseteq [d]$, we let X_A denote the set of elements in X whose set of active coordinates is precisely A.

For $m \in [k]$, we will construct the list $\mathcal{L}_m$ of all the elements $S \in X$ with $|S| = m$ and whose set of active coordinates are a subset of $[2k]$. We do so iteratively. Suppose we have constructed the list $\mathcal{L}_m$. For a vertex $v \in V(Q_d)$, and $\{i, j\} \subseteq [d]$, let v_{ij} denote the vertex of Q_d obtained by flipping the ith and jth coordinate of v. For each pair $\{i, j\} \subseteq [2k]$, $S \in \mathcal{L}_m$ and $v \in S$, add $S \cup \{v_{ij}\}$ to the list $\mathcal{L}_{m+1}$ if $v_{ij} \notin S$. This procedure generates the whole list $\mathcal{L}_{m+1}$ and shows that $|\mathcal{L}_{m+1}| \leq m\binom{2k}{2}|\mathcal{L}_m|$ and so $|\mathcal{L}_k| \leq k!\binom{2k}{2}^k = e^{O(k \log k)}$. For $m \in [k]$ and $a \in [2k]$, let $\mathcal{L}_m^a$ denote the subset of $\mathcal{L}_m$ consisting of those sets whose active coordinates are precisely $[a]$. Note that we can generate the list $\mathcal{L}_m^a$ in time $e^{O(k \log k)}$ by checking the elements of $\mathcal{L}_m$ one by one.

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 fixed $a \in [2k]$ and $m \in [k]$, we generate the list $\mathcal{G}_{m,k,a}$ of all clusters of size k containing $\underline{0}$ with active coordinates $[a]$ and $|V(\Gamma)| = m$. To do this we run through each $S \in \mathcal{L}_m^a$ and create the list of clusters Γ of size k with $V(\Gamma) = S$. We claim that this can be done in time $e^{O(k \log k)}$. Recall that a cluster of size k is an ordered set of polymers $(\gamma_1, \ldots, \gamma_\ell)$ such that $\sum_{i=1}^{\ell} |\gamma_i| = k$. Let us fix $S \in \mathcal{L}_m^a$. Since there are at most 2^k ordered integer partitions of k, it suffices to show that for a fixed such partition $(m_1, \ldots, m_\ell)$ (so that $\sum_i m_i = k$) we may find, in time $e^{O(k \log k)}$, all clusters $(\gamma_1, \ldots, \gamma_\ell)$ for which $|\gamma_i| = m_i$ for all i and $\bigcup_i \gamma_i = S$. To do this we can simply check each element of $\binom{S}{m_1} \times \ldots \times \binom{S}{m_\ell}$ (a set of size at most $e^{O(k \log k)}$) to see if it constitutes a legitimate cluster.

By symmetry of coordinates and vertex transitivity of Q_d we have

$$L_k = 2^{d-1} \sum_{j=1}^{k} \frac{1}{j} \sum_{a=1}^{2k} \binom{d}{a} \sum_{\Gamma \in \mathcal{G}_{j,k,a}} w(\Gamma).$$

Finally we note that by using an algorithm of Björklund, Husfeldt, Kaski, and Koivisto [2, Theorem 1], we may calculate the Ursell function of a cluster $\Gamma \in \mathcal{G}_{j,k,a}$ in time $e^{O(k)}$. Moreover for a set $S \in \mathcal{L}_j^a$ where $j \in [k]$, we can calculate $|N(S)|$ in time $O(k^2)$. We can therefore calculate $w(\Gamma)$ in time $e^{O(k)}$. $\qquad\square$

4. Probabilistic properties via the cluster expansion

Here we use the cluster expansion to prove Theorems 5 and 6 and Corollary 9. Using Lemmas 14 and 17 we see that up to $O(\exp(-2^d/d^4))$ total variation error, we may replace the minority side of an independent set drawn from μ with the defect side of an independent set drawn from $\hat{\mu}$; or in other words, a polymer configuration drawn from ν. Thus in this section we will let X_T denote the (random) number of polymers of type T in a random polymer configuration $\boldsymbol{\Gamma}$ drawn from ν, and prove the conclusions of Theorems 6 and 5 for these random variables. We will also assume throughout this section that $C_0 \log d/d^{1/3} \leq \lambda \leq 2$. Theorem 6 is vacuous if $\lambda > 2$ since $m_T \to 0$ for all types T in that case; the formula (5) in Corollary 9 holds for $\lambda > 2$ by Theorem 7.

We begin with some preliminaries on cumulants of random variables. Recall the cumulant generating function of a random variable X, $h_t(X) = \log \mathbb{E}e^{tX}$. The kth *cumulant* of X is defined by taking derivatives of $h_t(X)$ and evaluating at 0:

$$\kappa_k(X) = \left.\frac{\partial^k h_t(X)}{\partial t^k}\right|_{t=0}.$$

In fact the cumulants of X are related to the moments of X by a non-linear change of basis (see e.g. [19]). In particular, $\kappa_1(X) = \mathbb{E}X$ and $\kappa_2(X) = \text{var}(X)$. 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 $\lim_{n\to\infty} \kappa_k(X_n) = \kappa_k(X)$ for all $k \geq 1$, then X_n converges to X in distribution (denoted $X_n \Rightarrow X$). We will use this fact in conjunction with the following fact.

Fact 18. *If X has a Poisson distribution with mean m, then $\kappa_k(X) = m$ for all k. If X has a standard normal distribution (mean 0, variance 1) then $\kappa_1(X) = 0$, $\kappa_2(X) = 1$, and $\kappa_k(X) = 0$ for all $k \geq 3$.*

We also need a few preliminaries about defect types. First, for fixed t the number of defect types of size t is bounded independent of d. Let $\tau(S)$ denote the type of a polymer S. The weight of a polymer S is determined by $\tau(S)$, since $|N(S)|$ is determined by the number of edges of S in the graph $Q_d^2[S]$. Let w_T denote $w(S)$ for S of type T. Using Lemma 12, we have the simple bounds

$$(15) \qquad \frac{\lambda^t}{(1+\lambda)^{dt}} \leq w_T \leq \frac{\lambda^t}{(1+\lambda)^{dt-2t^2}}$$

for a type T of size t and d large enough. Note that for any fixed $k \geq 1$ and any type T, we have $d^k w_T \to 0$ as $d \to \infty$; that is, each polymer weight decays super-polynomially fast in d. We denote by $n_T = n_T(d)$ the number of polymers of type T.

Lemma 19. *Let T be a defect type of a fixed size t. Then*

$$(16) \qquad \frac{2^{d-1}}{t} \leq n_T \leq \frac{2^{d-1}}{t}(ed^2)^{(t-1)}.$$

Moreover, if T is a tree defect type then

$$n_T = (c_T + o(1))2^d d^{2t-2}$$

where $c_T = 2^{-t}|Aut(T)|^{-1}$ and if T is not a tree then

$$n_T = O(2^d d^{2t-3}).$$

Proof. By the vertex transitivity of Q_d, every vertex of $\mathcal{E}$ (or $\mathcal{O}$) is contained in the same number of polymers of type T. Let us denote this number by $n_{T,v}$ and note that $n_T = 2^{d-1}n_{T,v}/t$. The lower bound in (16) follows from the fact that if there exists a polymer with type T, then certainly $n_{T,v} \geq 1$. The upper bound follows from the fact that every vertex of Q_d is contained in at most $(ed^2)^{(t-1)}$ 2-linked sets of size t by Lemma 13.

Since T is a connected graph we may fix an ordering $(x_1, \ldots, x_t)$ of the vertices of T so that $T_i := T[\{x_1, \ldots, x_i\}]$ is connected for all $i \in [t]$. We let d_i denote the degree of the vertex x_i in the graph T_i.

We will construct an injective graph homomorphism $\varphi : T \to Q_d^2[\mathcal{E}]$ recursively as follows. Suppose that we have constructed an injective graph homomorphism $\varphi_i : T_i \to Q_d^2[\mathcal{E}]$ for some $i \leq t-1$ and let m_i denote the number of such homomorphisms. We now extend φ_i to an injective graph homomorphism $\varphi_{i+1} : T_{i+1} \to Q_d^2[\mathcal{E}]$. 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 $u, v \in \mathcal{E}$ their codegree in $Q_d^2[\mathcal{E}]$ is at most $2(d-2)$ and so there are at most $2(d-2)$ choices for $\varphi_{i+1}(x_{i+1})$ whence

(17) $$m_{i+1} \leq 2(d-2)m_i .$$

Suppose now that $d_{i+1} = 1$ and let R_i denote the set of possible choices for $\varphi_{i+1}(x_{i+1})$. We note that $u \in R_i$ if and only if u is adjacent to $\varphi_i(x_i)$ and non-adjacent to $\varphi_i(x_j)$ for $j < i$ in $Q_d^2[\mathcal{E}]$. Again using the fact that the maximum codegree in $Q_d^2[\mathcal{E}]$ is $2(d-2)$ it follows that $\binom{d}{2} - 2(d-2) \leq |R_i| \leq \binom{d}{2}$. We then have that

(18) $$\left(\binom{d}{2} - 2(d-2) \right) m_i \leq m_{i+1} \leq \binom{d}{2} m_i .$$

If T is not a tree then $d_{i+1} > 1$ for some $i \leq t-1$. It follows by (17) and the upper bound of (18) that $m_t = O(2^d d^{2(t-1)-1}) = O(2^d d^{2t-3})$. The bound $n_T = O(2^d d^{2t-3})$ follows from the fact that $n_T = m_t/|\mathrm{Aut}(T)|$ where $\mathrm{Aut}(T)$ denotes the automorphism group of the graph T (recall that t is a constant).

If T is a tree then $d_{i+1} = 1$ for all $i \leq t-1$ and so by (18) $m_t = (1+o(1))2^{d-1}d^{2(t-1)}2^{-(t-1)}$. The result follows. $\square$

Now fix a defect type T and let X_T be the number of polymers of type T in $\boldsymbol{\Gamma}$. We introduce modified polymer weights $\tilde{w}$, given by

$$\tilde{w}(S) = w(S)e^{t\mathbf{1}_{\tau(S)=T}} .$$

Let $\tilde{\Xi}$ be the corresponding polymer model partition function. Then we have

$$\mathbb{E}e^{tX_T} = \sum_{\Gamma} \nu(\Gamma)e^{t\sum_{S\in\Gamma}\mathbf{1}_{\tau(S)=T}}$$

$$= \frac{1}{\Xi}\sum_{\Gamma}\prod_{S\in\Gamma}w(S)e^{t\mathbf{1}_{\tau(S)=T}}$$

$$= \frac{\tilde{\Xi}}{\Xi},$$

and so

$$\kappa_k(X_T) = \frac{\partial^k}{\partial t^k} \log \frac{\tilde{\Xi}}{\Xi} \bigg|_{t=0}$$

$$= \frac{\partial^k \log \tilde{\Xi}}{\partial t^k} \bigg|_{t=0}.$$

If the cluster expansion for $\log \tilde{\Xi}$ converges absolutely, we can write

$$\kappa_k(X_T) = \frac{\partial^k}{\partial t^k} \sum_{\Gamma \in \mathcal{C}} \tilde{w}(\Gamma) \bigg|_{t=0}$$

(19)
$$= \sum_{\Gamma \in \mathcal{C}} w(\Gamma) Y_T(\Gamma)^k,$$

where $Y_T(\Gamma) = \sum_{S \in \Gamma} \mathbf{1}_{\tau(S)=T}$, the number of polymers of type T in the cluster Γ.

The following lemma gives bounds on cluster weights using the Kotecký–Preiss condition. Theorem 6 will then follow in a series of corollaries.

Lemma 20. *Consider a fixed defect type T, and let $k \geq 1$ be a fixed integer. Then*

(20)
$$\sum_{\Gamma \in \mathcal{C}} w(\Gamma) Y_T(\Gamma)^k = (1 + o(1)) n_T w_T$$

as $d \to \infty$.

Moreover if $\{T_1, \ldots, T_\ell\}$ is a fixed set of distinct defect types, and $k_1, \ldots k_\ell$ are fixed positive integers, then

(21)
$$\sum_{\Gamma \in \mathcal{C}} |w(\Gamma)| \prod_{i=1}^{\ell} Y_{T_i}(\Gamma)^{k_i} = O\left(d^{7\ell} 2^d \prod_{i=1}^{\ell} w_{T_i}\right).$$

Proof. In the sum in (20), if we consider only clusters made up of a single polymer of type T then we get a contribution of exactly $n_T w_T$, and so it remains to show that the contribution of all other terms is $o(n_T w_T)$. Let t denote the number of vertices in a graph of type T. We first consider the contribution to the sum (20) from clusters Γ with $Y_T(\Gamma) = 1$ and $\|\Gamma\| > t$. By (10), we may bound this contribution by

$$\sum_{\substack{\Gamma \in \mathcal{C} \\ \|\Gamma\| \geq t+1}} |w(\Gamma)| \leq d^{-3/2} 2^{d-1} e^{-\gamma(d,t+1)} = d^{11/2} 2^{d-1} (1+\lambda)^{-d(t+1)+3(t+1)^2} = o(n_T w_T),$$

since from (15) and (16)

$$n_T w_T \geq \frac{2^{d-1} \lambda^t}{t(1+\lambda)^{dt}}.$$

Consider now the contribution to the sum (20) from clusters Γ with $Y_T(\Gamma) = y > 1$. For such a cluster, recalling that $g(S) = \gamma(d, |S|)$ for a polymer S and $g(\Gamma) = \sum_{S \in \Gamma} g(S)$, we have

$$g(\Gamma) \geq y[\log(1+\lambda)(dt - 3t^2) - 7\log d].$$

Thus, using (13) we may bound this contribution by

$$2^{d-1}d^{-3/2}e^{-y[\log(1+\lambda)(dt-3t^2)-7\log d]}y^k = y^k d^{7y-3/2}2^{d-1}(1+\lambda)^{-dyt+3yt^2} \leq 2^{d-1}(1+\lambda)^{-3dyt/4},$$

where the above inequality holds for d large enough (independent of y). The result follows since

$$\sum_{y=2}^{\infty} 2^{d-1}(1+\lambda)^{-3ydt/4} \leq 2^d(1+\lambda)^{-3dt/2} = o(n_T w_T).$$

Next we turn to (21). Consider a cluster Γ with $Y_{T_1}(\Gamma) = y_1, \ldots, Y_{T_\ell}(\Gamma) = y_\ell$, where $y_1, \ldots, y_\ell \geq 1$. Then we have

$$g(\Gamma) \geq \sum_{j=1}^{\ell} y_j[\log(1+\lambda)(dt_j - 3t_j^2) - 7\log d]$$

where t_j is the size of a polymer of type T_j. By (13) and (15), the contribution of such clusters to the sum in (21) is therefore at most

$$\frac{2^{d-1}}{d^{3/2}}\prod_{j=1}^{\ell} y_j^{k_j} d^{7y_j}(1+\lambda)^{-dt_j y_j + 3t_j^2 y_j} = O\left(2^d d^{7\ell}\prod_{j=1}^{\ell} w_{T_j}\right)\prod_{j=1}^{\ell} y_j^{k_j} d^{7(y_j-1)}(1+\lambda)^{-dt_j(y_j-1)+3t_j^2 y_j}.$$

Finally, let $K = \max\{k_1, \ldots k_\ell, t_1, \ldots t_\ell\}$, so that summing over all positive integer vectors $\vec{y} = (y_1, \ldots, y_\ell)$, we have

$$\sum_{\vec{y}}\prod_{j=1}^{\ell} y_j^{k_j} d^{7(y_j-1)}(1+\lambda)^{-dt_j(y_j-1)+3t_j^2 y_j} \leq \sum_{\vec{y}}\prod_{j=1}^{\ell} y_j^K d^{7(y_j-1)}(1+\lambda)^{-d(y_j-1)+3K^2 y_j}$$

$$\leq \sum_{s=0}^{\infty}\sum_{\substack{\vec{y}: \\ \sum(y_j-1)=s}} d^{7s}(1+\lambda)^{-ds}\prod_j y_j^K(1+\lambda)^{3K^2 y_j}$$

$$= (1+\lambda)^{3K^2\ell}\sum_{s=0}^{\infty}\sum_{\substack{\vec{y}: \\ \sum(y_j-1)=s}} d^{7s}(1+\lambda)^{-(d-3K^2)s}\prod_j y_j^K$$

$$\leq (1+\lambda)^{3K^2\ell}\sum_{s=0}^{\infty}\sum_{\substack{\vec{y}: \\ \sum(y_j-1)=s}} (s+\ell)^K d^{7s}(1+\lambda)^{-(d-3K^2)s}$$

$$\leq (1+\lambda)^{3K^2\ell}\sum_{s=0}^{\infty} s^\ell(s+\ell)^K d^{7s}(1+\lambda)^{-(d-3K^2)s}$$

$$= O(1).$$

Putting these estimates together yields (21). □

An immediate corollary of Lemma 20 gives the asymptotics of m_T, σ_T^2 for a given type T.

Corollary 21. *Let T be a defect type. Then*

$$m_T = (1 + o(1))n_T w_T$$

and

$$\sigma_T^2 = (1 + o(1))n_T w_T \,.$$

Proof. These formulae follow from (19) and (20) by taking $k = 1$ and $k = 2$ respectively.
□

We can also use Lemma 20 to prove Poisson convergence.

Corollary 22. *Suppose for a given type T and fugacity λ we have $m_T \to \rho > 0$ as $d \to \infty$. Then $X_T \Rightarrow Pois(\rho)$.*

Proof. Using Fact 18, it is enough to show that $\kappa_k(X_T) \to \rho$ for all $k \geq 1$. By (19) and our assumption we have

$$m_T = \sum_{\Gamma \in \mathcal{C}} w(\Gamma) Y_T(\Gamma) = \rho + o(1),$$

and therefore using (19) again,

$$\sum_{\Gamma \in \mathcal{C}} w(\Gamma) Y_T(\Gamma)^k = \rho + o(1)$$

for all $k \geq 1$.
□

In a similar fashion, we obtain asymptotic normality if $m_T \to \infty$.

Corollary 23. *Fix a type T. If λ is such that $m_T \to \infty$ as $d \to \infty$, then $\tilde{X}_T = (X_T - m_T)/\sigma_T \Rightarrow N(0, 1)$.*

Proof. By Fact 18, it suffices to show that $\kappa_1(\tilde{X}_T) \to 0$, $\kappa_2(\tilde{X}_T) \to 1$, and $\kappa_k(\tilde{X}_T) \to 0$ for all $k \geq 3$. By the definition of $\tilde{X}_T$, we have $\kappa_1(\tilde{X}_T) = 0$ and $\kappa_2(\tilde{X}_T) = 1$. By translation invariance and scaling of higher cumulants, for $k \geq 3$ we have

$$\kappa_k(\tilde{X}_T) = \frac{1}{\sigma_T^k} \kappa_k(X_T)$$

$$= \frac{1}{\sigma_T^k} \sum_{\Gamma \in \mathcal{C}} w(\Gamma) Y_T(\Gamma)^k$$

by (19). By Lemmas 20 and 21 we have $\sum_{\Gamma \in \mathcal{C}} w(\Gamma) Y_T(\Gamma)^k = (1 + o(1))\sigma_T^2$, and so for $k \geq 3$,

$$\kappa_k(\tilde{X}_T) = O\left(\sigma_T^{2-k}\right) \to 0$$

as $d \to \infty$ since our assumption on m_T implies $\sigma_T \to \infty$.
□

To study the joint distribution of the counts of different defect types, it is convenient to work with the *joint cumulants* of a collection of random variables. Given a set of random variables $(X_1, \ldots, X_\ell)$ and non-negative integers $k_1, \ldots, k_\ell$, we define the joint cumulant

$$\kappa\left(X_1^{(k_1)}, \ldots, X_\ell^{(k_\ell)}\right) = \frac{\partial^{\sum_i k_i}}{\prod_i \partial t_i^{k_i}} \log \mathbb{E} e^{\sum_{i=1}^{\ell} t_i X_i}\bigg|_{t_1, \ldots, t_\ell = 0} \,.$$

In particular, with this notation

$$\kappa_k(X) = \kappa(X^{(k)}).$$

We will use the fact that the joint cumulants of independent random variables vanish; that is, if $\ell \geq 2$, $X_1, \ldots, X_\ell$ are independent random variables, and $k_1, \ldots, k_\ell$ are positive integers, then

(22) $$\kappa\left(X_1^{(k_1)}, \ldots, X_\ell^{(k_\ell)}\right) = 0.$$

Generalizing formula (19) to collections of random variables, we can express the joint cumulants of defect type counts via a modified cluster expansion. Let $\{T_1, \ldots, T_\ell\}$ be a set of distinct defect types and let $k_1, \ldots, k_\ell$ be non-negative integers. Then

$$\kappa\left(X_{T_1}^{(k_1)}, \ldots, X_{T_\ell}^{(k_\ell)}\right) = \sum_{\Gamma \in \mathcal{C}} w(\Gamma) \prod_{i=1}^{\ell} Y_{T_i}(\Gamma)^{k_i}.$$

Corollary 24. *Consider two fixed sets $\mathcal{T}_1$ and $\mathcal{T}_2$ of distinct defect types so that for each $T \in \mathcal{T}_1$, $m_T \to \rho_T$ for some $\rho_T > 0$, and for each $T \in \mathcal{T}_2$, $m_T \to \infty$ as $d \to \infty$. Then the collection of random variables $\{X_T\}_{T \in \mathcal{T}_1} \cup \{\tilde{X}_T\}_{T \in \mathcal{T}_2}$ converges in distribution to a collection of independent Poisson and standard normal random variables.*

Proof. We will use the fact that the distribution of a collection of Poisson and normal random variables is determined by its joint moments, or equivalently, by its joint cumulants. Here, working with cumulants instead of moments will simplify calculations considerably. From Corollaries 22 and 23 we know that the cumulants of each of the individual random variables in the collection converge to the corresponding cumulants of the corresponding Poisson or normal random variable. Therefore it is enough to show convergence of the joint cumulants involving at least two of the random variables, and from (22), we must show that these converge to 0. In particular, for $T_1, \ldots, T_j \in \mathcal{T}_1$, and $T_{j+1}, \ldots, T_\ell \in \mathcal{T}_2$, we will show

(23) $$\kappa(X_{T_1}^{(k_1)}, \ldots, X_{T_j}^{(k_j)}, \tilde{X}_{T_{j+1}}^{(k_{j+1})}, \tilde{X}_{T_\ell}^{(k_\ell)}) \to 0$$

as $d \to \infty$ as long as least two of the k_i's are positive. Since $\sigma_T^2 \to \rho_T > 0$ for $T \in \mathcal{T}_1$, it will suffice to show (23) when we center and normalise all of the random variables, that is, for $T_1, \ldots, T_\ell \in \mathcal{T}_1 \cup \mathcal{T}_2$,

$$\kappa(\tilde{X}_{T_1}^{(k_1)}, \ldots, \tilde{X}_{T_\ell}^{(k_\ell)}) \to 0$$

as long as at least two of the k_i's are positive. WLOG we can assume that $\ell \geq 2$, $k_i \geq 1$ for all i, and that $w_{T_1} \geq w_{T_2} \geq \cdots \geq w_{T_\ell}$. By scaling and translation invariance, we have

$$\kappa(\tilde{X}_{T_1}^{(k_1)}, \ldots, \tilde{X}_{T_\ell}^{(k_\ell)}) = \prod_{i=1}^{\ell} \frac{1}{\sigma_{T_i}^{k_i}} \kappa(X_{T_1}^{(k_1)}, \ldots, X_{T_\ell}^{(k_\ell)})$$

$$= \prod_{i=1}^{\ell} \frac{1}{\sigma_{T_i}^{k_i}} \sum_{\Gamma} w(\Gamma) \prod_{i=1}^{\ell} Y_{T_i}(\Gamma)^{k_i}.$$

Then using (21) from Lemma 20 we have

$$\left| \kappa(\tilde{X}_{T_1}^{(k_1)}, \dots, \tilde{X}_{T_\ell}^{(k_\ell)}) \right| = O\left(\prod_{i=1}^{\ell} \frac{1}{\sigma_{T_i}^{k_i}} \cdot d^{7\ell} \cdot 2^d \prod_{i=1}^{\ell} w_{T_i} \right) .$$

First suppose that $k_1 \geq 2$. Then since $\sigma_{T_i} = \Omega(1)$ for all i and $2^d w_{T_1} = O(\sigma_{T_1}^2)$, we have

$$\left| \kappa(\tilde{X}_{T_1}^{(k_1)}, \dots, \tilde{X}_{T_\ell}^{(k_\ell)}) \right| = O\left(d^{7\ell} \sigma_{T_1}^{2-k_1} \prod_{i=2}^{\ell} w_{T_i} \sigma_{T_i}^{-k_i} \right)$$

$$= O\left(d^{7\ell} w_{T_2} \right)$$

$$= o(1)$$

since w_T tends to 0 faster than any fixed polynomial in d for any type T. On the other hand if we have $k_1 = 1$, then

$$\left| \kappa(\tilde{X}_{T_1}^{(k_1)}, \dots, \tilde{X}_{T_\ell}^{(k_\ell)}) \right| = O\left(d^{7\ell} w_{T_1}^{1/2} w_{T_2}^{1/2} \sigma_{T_2}^{1-k_2} \prod_{i=3}^{\ell} w_{T_i} \sigma_{T_i}^{-k_i} \right)$$

$$= O\left(d^{7\ell} w_{T_1}^{1/2} \right)$$

$$= o(1) .$$

$\square$

Theorem 6 follows from Corollaries 22, 23, and 24. We now prove Theorem 5.

Proof of Theorem 5. We can assume in what follows that $\lambda \leq 2$, since if $\lambda > 2$ whp there are no occupied vertices on the minority side (Theorem 1.2 of [7]).

First we show that if $\lambda = 2^{1/t} - 1 + \frac{2^{1+1/t}(t-1)\log d}{td} + \frac{\omega(1)}{d}$, then whp there are no 2-linked components of size t on the defect side.

Let T the type of a polymer of size t. We have the bounds

$$\frac{\lambda^t}{(1+\lambda)^{td}} \leq w_T \leq \frac{\lambda^t}{(1+\lambda)^{td-2t^2}} ,$$

where the upper bound uses Lemma 12, and so $w_T = \Theta(\lambda^t(1+\lambda)^{-dt}) = \Theta((1+\lambda)^{-dt})$ for this range of λ. By Lemma 13, $n_T = O(2^d d^{2(t-1)})$, and so by Corollary 21,

$$m_T = O\left(\frac{2^d d^{2(t-1)}}{(1+\lambda)^{dt}} \right) .$$

Now plugging in $\lambda = 2^{1/t} - 1 + \frac{2^{1+1/t}(t-1)\log d}{td} + \frac{s}{d}$ for some s, we have

$$m_T = O\left(\frac{2^d d^{2(t-1)}}{\left(2^{1/t} + \frac{2^{1+1/t}(t-1)\log d}{td} + \frac{s}{d} \right)^{dt}} \right)$$

$$= O\left(\frac{d^{2(t-1)}}{\left(1 + \frac{2(t-1)\log d}{td} + \frac{s 2^{-1/t}}{d} \right)^{dt}} \right)$$

$$= O\left(e^{-st2^{-1/t}}\right),$$

and so as $s \to \infty$, $m_T \to 0$. This is true for any type T of size t, and since there are a constant number of such types, Markov's inequality shows that whp there are no polymers of size t in $\mathbf{\Gamma}$ if $\lambda = 2^{1/t} - 1 + \frac{2^{1+1/t}(t-1)\log d}{td} + \frac{\omega(1)}{d}$.

Now suppose $\lambda = 2^{1/t} - 1 + \frac{2^{1+1/t}(t-1)\log d}{td} - \frac{\omega(1)}{d}$. Consider a type T where T is isomorphic to a tree on t vertices. In this case we have $n_T = \Theta(2^d d^{2(t-1)})$ by Lemma 19, and so for $\lambda = 2^{1/t} - 1 + \frac{2^{1+1/t}(t-1)\log d}{td} + \frac{s}{d}$ the previous calculation gives

$$m_T = \Omega\left(e^{-st2^{-1/t}}\right).$$

In particular if $s \to -\infty$ as $d \to \infty$ then $m_T \to \infty$. By Corollary 21, $\sigma_T^2 \sim m_T$, and so by the second-moment method (Paley-Zygmund inequality), $X_T \geq 1$ whp.

To prove the second part of Theorem 5, suppose that $\lambda = 2^{1/t} - 1 + \frac{2^{1+1/t}(t-1)\log d}{td} + \frac{s(d)}{d}$ where $s(d)$ converges to a constant s as $d \to \infty$. Then for any type T of size t that is not a tree, by Lemma 19 we have $m_T = o(1)$ as $d \to \infty$, and since there is a constant number of such types, we know that whp there are no non-tree 2-linked components of size t on the minority side. Let $T_1, \ldots, T_\ell$ be the defect types of size t that are trees. The proof of Lemma 19 shows that in fact every tree on t vertices is a defect type. Note that for each i we have that $w_{T_i} = \lambda^t(1+\lambda)^{-dt+2(t-1)}$. Then by Lemma 19 and Corollary 21 we have that $m_{T_i} = (c_{T_i} + o(1))(\lambda^t(1+\lambda)^{-dt+2(t-1)}2^d d^{2(t-1)})$ for each i, and so by a similar calculation as above we have that $m_{T_i} \to \rho_i$ as $d \to \infty$ where

$$\rho_i = \frac{1}{2^t|\mathrm{Aut}(T_i)|}e^{-st2^{-1/t}}(2^{1/t} - 1)^t 2^{2(1-1/t)}.$$

By Corollary 24, the collection of random variables $X_{T_1}, \ldots, X_{T_\ell}$ converges to a collection of independent Poisson random variables with mean $\rho_1, \ldots, \rho_\ell$, and therefore their sum is distributed as Poisson with mean $\sum_{i=1}^{\ell} \rho_i$, completing the proof of Theorem 5. Calculating this mean explicitly amounts to calculating $|\mathrm{Aut}(T)|$ for every tree T on t vertices, a task whose running time depends only on t (a constant). □

The proof of Corollary 9 involves a similar calculation.

Proof of Corollary 9. We may again assume that $\lambda \leq 2$ since for larger λ, $Z(\lambda) = (2 + o(1))(1+\lambda)^{2^{d-1}}$ by Theorem 7. Now fix $t \geq 1$ and take $\lambda = 2^{1/t} - 1 + \frac{2^{1+1/t}(t-1)\log d}{td} + \frac{\omega(1)}{d}$. We then can apply Theorem 8 with $k = t$ to obtain

$$Z(\lambda) = 2(1+\lambda)^{2^{d-1}} \cdot \exp\left(\sum_{j=1}^{t} L_j + \varepsilon_t\right).$$

But by the same calculation as above in the proof of Theorem 5 we have

$$|L_t| = O\left(\frac{2^d d^{2(t-1)}}{(1+\lambda)^{dt}}\right)$$
$$= o(1),$$

and

$$|\varepsilon_t| = O\left(\frac{2^d d^{2t}}{(1+\lambda)^{d(t+1)}}\right) = o(1),$$

and so

$$Z(\lambda) = (2 + o(1))(1+\lambda)^{2^{d-1}} \exp\left(\sum_{j=1}^{t-1} L_j\right).$$

The example formula (6) follows from the computation of L_1, L_2 given below in Section 5.   □

5. Computation of the cluster weights

Here we compute L_1, L_2, L_3 explicitly to use in Theorem 2 and Corollary 9.

Proposition 25. *We have*

$$L_1 = \frac{2^d \lambda}{(1+\lambda)^d} \cdot \frac{1}{2}$$

$$L_2 = \frac{2^d \lambda^2}{(1+\lambda)^{2d}} \cdot \frac{(2\lambda + \lambda^2)d(d-1) - 2}{8}$$

$$L_3 = \frac{2^d \lambda^3}{48(1+\lambda)^{3d}}\big[8 + 2(8\lambda - 2\lambda^2 + 4\lambda^3 + 11\lambda^4 + 4\lambda^5)d + 3(-4\lambda + 12\lambda^2 + 4\lambda^3 - 9\lambda^4 - 4\lambda^5)d^2$$

$$+ 2(-2\lambda - 22\lambda^2 - 16\lambda^3 + \lambda^4 + 2\lambda^5)d^3 + 3(4\lambda^2 + 4\lambda^3 + \lambda^4)d^4\big]$$

At $\lambda = 1$, this is

$$L_1 = \frac{1}{2}$$

$$L_2 = 2^{-d} \cdot \frac{3d^2 - 3d - 2}{8}$$

$$L_3 = 2^{-2d} \cdot \frac{27d^4 - 74d^3 - 3d^2 + 50d + 8}{48}.$$

Polymers. There is a single type of polymer of size 1. There are 2^{d-1} of these, and each has weight $\lambda(1+\lambda)^{-d}$.

There is a single type of polymer of size 2. There are $2^{d-3}d(d-1)$ of these and each has weight $\lambda^2(1+\lambda)^{-2d+2}$.

There are two types of polymers of size 3: those that form a clique in the distance 2 graph and those that form a path on 3 vertices. There are $2^{d-2}d(d-1)(d-2)/3$ of the first type and each has weight $\lambda^3(1+\lambda)^{-3d+5}$; there are $2^{d-4}d(d-1)(d-2)(d-3)$ of the second type and each has weight $\lambda^3(1+\lambda)^{-3d+4}$.

Clusters. There is a single cluster type of size 1, each consisting of single polymer of size 1, with Ursell function 1. Thus

$$L_1 = \frac{2^d \lambda}{(1+\lambda)^d} \cdot \frac{1}{2}$$

There are two types of clusters of size 2: an ordered pair of incompatible polymers of size 1, of which there are $2^{d-1} + 2^{d-2}d(d-1)$, with Ursell function $-1/2$ and weight $\lambda^2(1+\lambda)^{-2d}$, and one polymer of size 2 with Ursell function 1 and count and weight given above.

All together this gives:

$$L_2 = -\frac{1}{2}\left(2^{d-1} + 2^{d-2}d(d-1)\right)\lambda^2(1+\lambda)^{-2d} + 2^{d-3}d(d-1)\lambda^2(1+\lambda)^{-2d+2}$$

$$= \frac{2^d\lambda^2}{(1+\lambda)^{2d}} \cdot \frac{(2\lambda + \lambda^2)d(d-1) - 2}{8}.$$

At $\lambda = 1$ this is

$$2^{-d} \cdot \frac{3d^2 - 3d - 2}{8}.$$

There are five types of clusters of size 3:
(1) One polymer of size 3, first type: $2^{d-2}d(d-1)(d-2)/3$ of weight $\lambda^3(1+\lambda)^{-3d+5}$, Ursell function 1.
(2) One polymer of size 3, second type: $2^{d-4}d(d-1)(d-2)(d-3)$ of weight $\lambda^3(1+\lambda)^{-3d+4}$, Ursell function 1.
(3) Three polymers of size 1, incompatibility graph is a triangle: $2^{d-1} + 3 \cdot 2^{d-2}d(d-1) + 2^{d-1}d(d-1)(d-2)$ of weight $\lambda^3(1+\lambda)^{-3d}$, Ursell function $1/3$.
(4) Three polymers of size 1, incompatibility graph is a path on 3 vertices: $3 \cdot 2^{d-3}d(d-1)(d-2)(d-3)$ of weight $\lambda^3(1+\lambda)^{-3d}$, Ursell function $1/6$.
(5) One polymer of size 2, one of size 1: $2^{d-2}d(d-1)[d(d-1) - 2(d-2)]$ of weight $\lambda^3(1+\lambda)^{-3d+2}$, Ursell function $-1/2$.

All together this gives:

$$L_3 = \frac{2^d\lambda^3}{48(1+\lambda)^{3d}}\Big[8 + 2(8\lambda - 2\lambda^2 + 4\lambda^3 + 11\lambda^4 + 4\lambda^5)d + 3(-4\lambda + 12\lambda^2 + 4\lambda^3 - 9\lambda^4 - 4\lambda^5)d^2$$

$$+ 2(-2\lambda - 22\lambda^2 - 16\lambda^3 + \lambda^4 + 2\lambda^5)d^3 + 3(4\lambda^2 + 4\lambda^3 + \lambda^4)d^4\Big]$$

At $\lambda = 1$ this is

$$2^{-2d} \cdot \frac{27d^4 - 74d^3 - 3d^2 + 50d + 8}{48}$$

Proof of Theorem 2. Theorem 8 tells us that

$$i(Q_d) = 2 \cdot 2^{2^{d-1}} \cdot \exp\left(L_1 + L_2 + L_3 + O(L_4)\right)$$

$$= 2\sqrt{e} \cdot 2^{2^{d-1}} \cdot \exp\left(L_2 + L_3 + O(L_4)\right)$$

since $L_1 = 1/2$. If we write $L_k = a_{k-1}2^{-(k-1)d}$, then we have

$$i(Q_d) = 2\sqrt{e} \cdot 2^{2^{d-1}} \cdot \exp\left(a_1 2^{-d} + a_2 2^{-2d} + O(a_3 2^{-3d})\right).$$

Since the Taylor series for $\exp(a_1 x + a_2 x^2 + a_3 x^3)$ around $x = 0$ is

$$1 + a_1 x + \left(\frac{a_1^2}{2} + a_2 \right) x^2 + O\left(\left(a_1^3 + a_1 a_2 + a_3 \right) x^3 \right) ,$$

we have

$$i(Q_d) = 2\sqrt{e} \cdot 2^{2^{d-1}} \left(1 + L_2 + \frac{L_2^2}{2} + L_3 + O(L_2^3 + L_2 L_3 + L_4) \right)$$

$$= 2\sqrt{e} \cdot 2^{2^{d-1}} \left(1 + \frac{3d^2 - 3d - 2}{8 \cdot 2^d} + \frac{243d^4 - 646d^3 - 33d^2 + 436d + 76}{384 \cdot 2^{2d}} + O\left(d^6 \cdot 2^{-3d} \right) \right)$$

which gives Theorem 2. □

Acknowledgements

References

[1] J. Balogh, R. Morris, and W. Samotij. Independent sets in hypergraphs. *Journal of the American Mathematical Society*, 28(3):669–709, 2015.

[2] A. Björklund, T. Husfeldt, P. Kaski, and 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, pages 677–686. IEEE, 2008.

[3] S. Cannon and W. Perkins. Counting independent sets in unbalanced bipartite graphs. In *Proceedings of the Fourteenth Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 1456–1466. SIAM, 2020.

[4] R. Dobrushin. Estimates of semi-invariants for the Ising model at low temperatures. *Translations of the American Mathematical Society-Series 2*, 177:59–82, 1996.

[5] J. Engbers and D. Galvin. H-coloring tori. *Journal of Combinatorial Theory, Series B*, 102(5):1110–1133, 2012.

[6] D. Galvin. On homomorphisms from the Hamming cube to $\mathbb{Z}$. *Israel Journal of Mathematics*, 138(1):189–213, 2003.

[7] D. Galvin. A threshold phenomenon for random independent sets in the discrete hypercube. *Combinatorics, probability and computing*, 20(1):27–51, 2011.

[8] D. Galvin. Independent sets in the discrete hypercube. *arXiv preprint arXiv:1901.01991*, 2019.

[9] D. Galvin and J. Kahn. On phase transition in the hard-core model on $\mathbb{Z}^d$. *Combinatorics, Probability and Computing*, 13(2):137–164, 2004.

[10] D. Galvin and P. Tetali. Slow mixing of Glauber dynamics for the hard-core model on regular bipartite graphs. *Random Structures & Algorithms*, 28(4):427–443, 2006.

[11] C. Gruber and H. Kunz. General properties of polymer systems. *Communications in Mathematical Physics*, 22(2):133–161, 1971.

[12] T. Helmuth, W. Perkins, and G. Regts. Algorithmic Pirogov-Sinai theory. *Probability Theory and Related Fields*, to appear. Extended abstract at STOC 2019.

[13] M. Jenssen and P. Keevash. Homomorphisms from the torus. *In preparation*.

[14] M. Jenssen, P. Keevash, and W. Perkins. Algorithms for #BIS-hard problems on expander graphs. In *Proceedings of the Thirtieth Annual ACM-SIAM Symposium on Discrete Algorithms (SODA 2019)*, pages 2235–2247. SIAM, 2019. Full version at http://arxiv.org/abs/1807.04804v2.

[15] J. Kahn. An entropy approach to the hard-core model on bipartite graphs. *Combinatorics, Probability and Computing*, 10(3):219–237, 2001.

[16] J. Kahn and J. Park. The number of 4-colorings of the Hamming cube. *Israel Journal of Mathematics*, to appear.

[17] A. Korshunov and A. Sapozhenko. The number of binary codes with distance 2. *Problemy Kibernet*, 40:111–130, 1983.

[18] R. Kotecký and D. Preiss. Cluster expansion for abstract polymer models. *Communications in Mathematical Physics*, 103(3):491–498, 1986.

[19] V. Leonov and A. N. Shiryaev. On a method of calculation of semi-invariants. *Theory of Probability & its Applications*, 4(3):319–329, 1959.

[20] C. Liao, J. Lin, P. Lu, and Z. Mao. Counting independent sets and colorings on random regular bipartite graphs. In *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques (APPROX/RANDOM 2019)*. Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik, 2019.

[21] W. Samotij. Counting independent sets in graphs. *European Journal of Combinatorics*, 48:5–18, 2015.

[22] A. Sapozhenko. On the number of connected subsets with given cardinality of the boundary in bipartite graphs. *Metody Diskret Analiz*, 45:42–70, 1987.

[23] D. Saxton and A. Thomason. Hypergraph containers. *Inventiones mathematicae*, 201(3):925–992, 2015.

[24] A. D. Scott and A. D. Sokal. The repulsive lattice gas, the independent-set polynomial, and the Lovász local lemma. *Journal of Statistical Physics*, 118(5-6):1151–1261, 2005.

School of Mathematics, University of Birmingham, Birmingham, UK
E-mail address: m.jenssen@bham.ac.uk.

Department of Mathematics, Statistics, and Computer Science, University of Illinois at Chicago, 851 S. Morgan, Chicago, IL
E-mail address: math@willperkins.org.