Decay of correlations for the hardcore model on the d-regular random graph

Nayantara Bhatnagar[*] Allan Sly[†] Prasad Tetali [‡]

Abstract

A key insight from statistical physics about spin systems on random graphs is the central role played by Gibbs measures on trees. We determine the local weak limit of the hardcore model on random regular graphs upto a density for the largest independent set that is bounded by and goes asymptotically to the condensation threshold. We show that the hardcore measure converges in probability locally in a strong sense to the free boundary condition Gibbs measure on the tree. As a consequence we show that the reconstruction threshold on the random graph, indicative of the onset of point to set spatial correlations, is equal to the reconstruction threshold on the d-regular tree for which we determine precise asymptotics. We expect that our methods will generalize to a wide range of spin systems for which the second moment method holds.

1 Introduction

In this paper we consider the hardcore model on random d-regular graphs and study its local spatial mixing properties. We determine the location of a phase transition where the model undergoes a spatial mixing transition after which the spin at a typical vertex becomes dependent over long distances. Theory from statistical physics relates this transition to the clustering or shattering threshold. These transitions appear to be related to the absence of efficient algorithms finding large independent sets, although no hardness results are known for random satisfiability problems. Currently, no algorithms are known to find independent sets of size $\frac{(1+\epsilon)\log d}{d}n$ in a random d-regular graph on n vertices, which coincides with the spatial mixing threshold. In contrast the

maximal independent set is of size $\frac{(2-o_d(1))\log d}{d}n$ [16]. In this work, we show that the reconstruction or extremality threshold on the infinite d-regular tree determines the onset of long distance point to set spatial correlations in the random d-regular graph. We prove an asymptotic lower bound on the reconstruction threshold which matches the known upper bound in the first two terms of the asymptotic series. Together, these results determine the asymptotic location of the threshold for the random d-regular graph for the onset of point to set correlations over long distances.

In a finite graph $G = (V, E)$, an independent set is a subset of the vertices containing no adjacent vertices. Denote the set of independent sets as $I(G)$. We will view an independent set as a *spin configuration* σ, taking values in $\{0,1\}^V$ with σ_v denoting the spin at the vertex v. We say a vertex or *site* v is *occupied* in σ if $\sigma_v = 1$. The *hardcore model* (or *hardcore measure*) is the probability measure over the set of independent sets $\sigma \in I(G)$ given by

$$\mathbb{P}(\sigma) = \frac{1}{Z}\lambda^{\sum_{v \in V}\sigma_v}\mathbb{1}_{\sigma \in I(G)}. \tag{1.1}$$

The parameter $\lambda > 0$ is known as the *fugacity* and controls the typical size of an independent set with larger values of λ putting more of the weight of the distribution on larger independent sets. As usual, Z is a normalizing constant called the partition function. The definition of Gibbs measures and the hardcore model in particular can be extended to infinite graphs by way of the Dobrushin-Lanford-Ruelle condition which essentially says that for every finite set A, the probability of a configuration on A is given by the Gibbs distribution given by a random boundary generated by the measure outside of A. Such a measure is called a Gibbs measure, and it may not be unique (see e.g. [18] for more details).

On the infinite d-regular tree T_d, there is a unique Gibbs measure for the hardcore model if and only if $\lambda \leq \frac{(d-1)^{d-1}}{(d-2)^d}$. However, for every λ, there exists a translation invariant Gibbs measure given by a Markov model on the tree which we denote by $\mathbb{P}_{T_d}$ (henceforth, we refer to this as "the translation invariant measure" on T_d). We denote the density of $\mathbb{P}_{T_d}$, that is, the probability that a site is occupied, by $\alpha = \alpha(\lambda, d)$ which satisfies the relation

$$\lambda = \frac{\alpha}{1 - 2\alpha}\left(\frac{1-\alpha}{1-2\alpha}\right)^{d-1}. \tag{1.2}$$

Since $\alpha = \alpha(\lambda, d)$ is a strictly monotone increasing function of λ we will use both parameters interchangeably to specify the model depending on the context. The density of the largest independent set of a d-regular random graph is asymptotically $(2\log d - (2 + o_d(1))\log\log d)/d$ [16]. The results we present hold very close to this threshold, up to

$$\alpha < \alpha_c(d) := \frac{(2\log d - (3 + o_d(1))\log\log d)}{d}.$$

We take λ_c to be the corresponding value of λ. The bulk of this paper is devoted to establishing that the hardcore measure on the random d-regular graph is well approximated locally by the measure $\mathbb{P}_{T_d}$ when $\lambda < \lambda_c$. We prove that the measure converges in a strong notion of local weak convergence described in Section 1.2.

Theorem 1. *Let G_n be the random d-regular graph on n vertices. Then for large enough d, the hardcore measure on G_n with fugacity $\lambda < \lambda_c$ converges in probability locally to the measure $\mathbb{P}_{T_d}$.*

Our methods provide a general framework for proving convergence in probability locally which we expect will apply to various other Gibbs measures on random graphs such as colorings or NAE-SAT. Having established Theorem 1, it is natural to consider properties of the measure $\mathbb{P}_{T_d}$. The set of Gibbs measures is convex and so we may

ask whether $\mathbb{P}_{T_d}$ is extremal, that is, it is not a convex combination of other Gibbs measures. Extremality is equivalent to a notion of point to set correlation on trees called the *reconstruction problem* (for a survey, see [30]).

To formalize the definition of the problem, we will make use of a description of $\mathbb{P}_{T_d}$ as a Markov model on the tree generated as follows. First the spin at the root is chosen to be occupied with probability α and unoccupied with probability $1 - \alpha$, where α is chosen as in (1.2). The spins of the remaining vertices of the graph are generated from their parents' spins by taking one step of the Markov transition matrix

$$M = \begin{pmatrix} p_{11} & p_{10} \\ p_{01} & p_{00} \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ \frac{\alpha}{1-\alpha} & \frac{1-2\alpha}{1-\alpha} \end{pmatrix},$$

where p_{ij} denotes the probability of the spin at a vertex being j given that the spin of the parent is state i. Since $(\alpha, 1 - \alpha)$ is reversible with respect to M this gives a translation invariant measure on T_d which corresponds to the measure $\mathbb{P}_{T_d}$ with fugacity λ.

Let $\sigma(L)$ denote the spins of the vertices $L = L(\ell)$ at depth ℓ in the tree, as generated by the Markov model described above. The reconstruction problem on the tree asks if we can recover information on σ_x, the spin of the root x from the spins $\sigma(L)$ as $\ell \to \infty$. Formally, we say that the model (T_d, M) has *non-reconstruction* if

$$\lim_{\ell \to \infty} \mathbb{P}_{T_d}(\sigma_x = 1 | \sigma(L)) \to \alpha(\lambda, d) \tag{1.3}$$

in probability as $\ell \to \infty$, and otherwise, the model has *reconstruction*. Non-reconstruction is equivalent to extremality of the Gibbs measure or that the tail σ-algebra of the Gibbs measure is trivial [30].

Information theoretically, non-reconstruction corresponds to fast decay of correlations between the spin at the root and the spins of far away vertices [30]. Proposition 12 of [29] implies that there exists a critical fugacity λ_R (or, equivalently, a critical density α_R) such that reconstruction holds for the hardcore model with fugacity $\lambda > \lambda_R$ and non-reconstruction holds for $\lambda < \lambda_R$. The reconstruction problem on the tree was originally studied as a problem in statistical physics but has since found many applications including in computational phylogenetic reconstruction [12], the study of the geometry of the space of random constraint satisfaction problems (CSP's) [1, 22] and the mixing time of Markov chains [4, 9, 25, 33, 38].

Here we establish tight bounds on the reconstruction threshold for the hardcore model on the d-regular tree[1]. The upper bound was shown by Brightwell and Winkler [10], and our contribution is the lower bound.

Theorem 2. *For large enough d, the reconstruction threshold for $\mathbb{P}_{T_d}$ on the d-regular tree satisfies*
$$\frac{(\log 2 - o_d(1)) \log^2 d}{2 \log \log d} \leq \lambda_R \leq (e + o_d(1)) \log^2 d.$$

Prior to our work, Martin [24] had shown that $\lambda_R > e - 1$. Restating Theorem 2 in terms of α we have that the critical density for reconstruction satisfies

$$\frac{1}{d} \left(\log d + \log \log d - \log \log \log d - \log 2 + \log \log 2 - o_d(1) \right)$$
$$\leq \alpha_R \leq \frac{1}{d} \left(\log d + \log \log d + 1 + o_d(1) \right) \tag{1.4}$$

leaving only an additive $(\log \log \log d)/d$ gap between the bounds. The form of our bound in equation (1.4) is strikingly similar to the bound for the q-coloring model [35] which

---

[1]This result previously appeared in extended abstract form in [5].

states that reconstruction (resp. non-reconstruction) holds when the degree d is at least (resp. at most) $q(\log q + \log \log q + O(1))$.

The next theorem, combined with Theorem 2 gives a precise picture of the local spatial mixing properties of the hardcore model on the random d-regular graph. In [19] a natural extension of the reconstruction problem was introduced for graphs. Let $\{G_n\}$ be a family of random graphs whose size n goes to infinity, and let σ be distributed according to the hardcore model with fugacity λ. We will use $\sigma(S)$ to denote the configuration on a subset of vertices S and σ_v to denote the spin at a vertex v. The model has non-reconstruction if for a uniformly chosen $u \in V(G_n)$,

$$\lim_{\ell \to \infty} \limsup_n \mathbb{E}\left|\mathbb{P}\left(\sigma_u = 1 | \sigma(\partial B_u(\ell)), u\right) - \alpha(\lambda, d)\right| = 0 \tag{1.5}$$

where $B_u(\ell)$ denotes the vertices within distance ℓ of u (and by abuse of notation, the induced subgraph), $\partial B_u(\ell)$ denotes the boundary of $B_u(\ell)$ and $\alpha(\lambda, d)$ is the density given by (1.2). Here, the expectation is over the choice of the random vertex u and the configuration on the boundary $\partial B_u(\ell)$.

Theorem 3. *Let $\lambda < \lambda_c$ and let $\alpha(\lambda, d)$ be the density given by (1.2). Let G_n be the random d-regular graph on n vertices and let u be a uniformly random vertex in $V(G_n)$. Then, for large enough d,*

$$\mathbb{P}_{T_d}(\sigma_x = 1 | \sigma(L)) \xrightarrow{\mathbb{P}} \alpha(\lambda, d) \ \text{ as } \ell \to \infty$$
$$\Leftrightarrow$$
$$\lim_{\ell \to \infty} \limsup_n \mathbb{E}\left|\mathbb{P}\left(\sigma_u = 1 | \sigma(\partial B_u(\ell)), u\right) - \alpha(\lambda, d)\right| = 0.$$

That is, the random d-regular graph has non-reconstruction if and only if (T_d, M) has non-reconstruction.

1.1 Related work

A significant body of work has been devoted to the reconstruction problem on the d-regular tree by probabilists, computer scientists and physicists for a number of different spin configuration models. The earliest such result is the Kesten-Stigum bound [21] which states that for a Markov model defined on the tree, reconstruction holds whenever $\theta^2(d-1) > 1$, where θ is the second largest eigenvalue of the corresponding Markov matrix. This bound was shown to be tight in the case of the Ising model [7, 14] where it was shown that non-reconstruction holds when $\theta^2(d-1) \leq 1$. Similar results were derived for the Ising model with small external field [9] and the 3-state Potts model [34] which constitute the only models for which exact thresholds are known. On the other hand, for the hardcore model $\theta^2(d-1) = (1 + o_d(1))\frac{1}{d}\log^2 d$ and thus at least when d is large, the Kesten-Stigum bound is known not to be tight [10].

In both the coloring model and the hardcore model the reconstruction threshold is far from the Kesten-Stigum bound for large d. In the coloring model close to optimal bounds on the reconstruction threshold [6, 35] were obtained by first showing that, when n is small, the information on the root is sufficiently small. Then a quantitative version of [20] establishes that the information on the root converges to 0 exponentially quickly. In this work, we show that the hardcore model behaves similarly.

1.1.1 Replica symmetry breaking and finding large independent sets

The reconstruction problem plays a deep role in the geometry of the space of solutions of random CSPs. While for problems with few constraints the space of solutions is

connected and finding solutions is generally easy, as the number of constraints increases the space may break into exponentially many small clusters. Physicists, using powerful but non-rigorous "replica symmetry breaking" heuristics, predicted that the clustering phase transition exactly coincides with the reconstruction region on the associated tree model [23, 22]. This picture was rigorously established (up to first order terms) for the coloring and satisfiability problems [1, 28]. When solutions are far apart, local search algorithms will in general fail. Indeed for both the coloring and SAT models, no algorithm is known to find solutions in the clustered phase. It has been conjectured to be computationally intractable beyond this phase transition [1].

Previous results [19, 28] have related the reconstruction problem on the Poisson tree with constant expected degree with reconstruction in sparse random graph ensembles. These results established a "replica" condition saying that the empirical distribution of pairs of spins at a vertex from two independent configurations are from a product measure. This does not apply in the case of the hardcore model since the degree of a vertex affects its probability of being in the independent set. At the same time, for the d-regular random graph the methods of [28] do not seem to be directly applicable and we approach the problem instead using the theory of local weak convergence of Gibbs measures.

The associated CSP for the hardcore model corresponds to finding large independent sets in random d-regular graphs. For the Erdös-Rényi random graph of average degree d or the random regular graph of degree d it was shown that there exist independent sets of density $\frac{(2-o_d(1))\log d}{d}$ [16, 15]. On the other hand the best known algorithm finds independent sets only of density $\frac{(1+o_d(1))\log d}{d}$ [39], which is equal to α_R asymptotically as $d \to \infty$ [5], and is roughly half the density of the largest independent set. It has been shown that local algorithms will fail to find or sample from dense independent sets beyond this threshold [17, 11]. It was shown that asymptotically close to the reconstruction threshold independent sets exhibit the same clustering phenomena in random regular graphs [32, 17] and Erdős-Renyi random graphs [11] as colorings and SAT [1, 22].

On the other hand, in the statistical physics community, it was thought until recently that the hardcore model behaves differently from the coloring or SAT [3]. Statistical physicists had shown that as the density of occupied sites increases, the hardcore model undergoes a continuous phase transition, rather than a discontinuous transition as the mathematical work suggests. In [3] the authors address this discrepancy and use the cavity method to predict that for graphs of small degree ($d < 16$), the phase transition is indeed continuous while for $d > 16$, the model should have a discontinuous phase transition at a location agreeing with the theoretical work showing clustering and reconstruction [11, 5]. By determining the reconstruction threshold on random regular graphs we provide further evidence supporting the connection of the reconstruction threshold with the computational hardness of finding large independent sets in random graphs.

Just below the satisfiability threshold many CSPs including the hardcore model are believed to undergo an additional phase transition, different from the clustering transition, called condensation [22, 26, 3]. Beyond the condensation transition the second moment method fails and the distribution places most of its weight on a constant number of clusters. For k-colorings on the Erdös-Rényi random graph, the cavity method predicts a precise location for the condensation transition [26] in terms of a distributional fixed point problem. For a sufficiently large number of colors, Bapst et al. [2] have proved that the condensation transition occurs exactly at this value. As far as we know, there is no precise conjecture about the location of the condensation transition in the case of the hardcore model although Barbier et al. [3] use another heuristic known as

the "1RSB approach" to conjecture that the transition exists and give its asymptotic location. After the condensation transition it is believed that the hardcore measure no longer converges locally to $\mathbb{P}_{T_d}$ explaining the necessity of an upper bound on λ in our theorems.

1.2 Local weak convergence

There are a number of natural notions of local weak convergence of Gibbs measures and we introduce these now, following the notation used in [27]. Let $\mathcal{T}_d$ denote the space of hardcore Gibbs measures on T_d endowed with the topology of weak convergence and let $\mathcal{M}_d$ to be the space of probability measures over $\mathcal{T}_d$. For a sequence of graphs G_n we denote a hardcore measure by μ_n while ν denotes a hardcore measure on T_d. The notation $T_d(\ell)$ will denote the restriction of the tree T_d to a ball of radius ℓ around the root (and by abuse of notation, we also use it to denote the set of vertices of the restriction). The shorthand ν^ℓ denotes the restriction of the corresponding measure to a ball of radius ℓ around the root. For a measure on Gibbs measures $\mathfrak{m} \in \mathcal{M}_d$, we let $\mathfrak{m}^\ell$ denote the measure on the space of measures on $\{0,1\}^{T_d(\ell)}$ induced by such projections.

Definition 1.1. *Consider a sequence of graphs-Gibbs measure pairs $\{(G_n, \mu_n)\}_{n \in \mathbb{N}}$ and for $v \in V(G_n)$, let $\mathbb{P}_n^\ell(v)$ denote the law of the pair $(B_v(\ell), \sigma(B_v(\ell)))$ when σ is drawn with distribution μ_n. Let U_n denote the uniform measure over a random vertex $u \in V(G_n)$. Let $\mathbb{P}_n^\ell = \mathbb{E}_{U_n}(\mathbb{P}_n^\ell(u))$ denote the average of $\mathbb{P}_n^\ell(u)$. Let $\delta_{T_d(\ell)}$ denote the Dirac measure on graphs which is 1 on $T_d(\ell)$.*

  A. *The first mode of convergence concerns picking a random vertex u and a random local configuration in the neighbourhood of u. Formally, for $\bar\nu \in \mathcal{T}_d$ we say that* $\{\mu_n\}_{n \in \mathbb{N}}$ *converges locally on average to $\bar\nu$ if for any ℓ,*

  $$\lim_{n \to \infty} d_{\mathrm{TV}}\left(\mathbb{P}_n^\ell, \delta_{T_d(\ell)} \times \bar\nu^\ell\right) = 0. \tag{1.6}$$

  B. *A stronger form of convergence involves picking a random vertex u and the associated random local measure $\mathbb{P}_n^\ell(u)$ and asking if this distribution of distributions converges. Formally, we say that* the local distributions of $\{\mu_n\}_{n \in \mathbb{N}}$ converge locally to $\mathfrak{m} \in \mathcal{M}_d$ *if it holds that the law of $\mathbb{P}_n^\ell(u)$ converges weakly to $\delta_{T_d(\ell)} \times \mathfrak{m}^\ell$ for all ℓ.*

  C. *If $\mathfrak{m}$ is a point mass on $\bar\nu \in \mathcal{T}_d$ and if the local distributions of $\{\mu_n\}_{n \in \mathbb{N}}$ converge locally to $\mathfrak{m}$ then we say that* $\{\mu_n\}_{n \in \mathbb{N}}$ *converges in probability locally to $\bar\nu$. Equivalently convergence in probability locally to $\bar\nu$ says that for any ℓ and any $\varepsilon > 0$ it holds that*

  $$\lim_{n \to \infty} U_n\left(d_{\mathrm{TV}}(\mathbb{P}_n^\ell(u), \delta_{T_d(\ell)} \times \nu^\ell) > \varepsilon\right) = 0. \tag{1.7}$$

Remark 1.2. As noted in [27], $C \Rightarrow B \Rightarrow A$ while in [37] it is noted that if the measures ν are *extremal* Gibbs measures then the three notions of convergence A, B and C are equivalent.

At a high level, convergence locally on average to ν means that after averaging the local distribution of configurations over all the vertices, the random configuration converges weakly to ν while convergence in probability locally to ν means that the local distribution at almost every vertex is close to ν eventually. As noted above the former is a weaker condition and is in fact much simpler to prove. One can apply the second moment method for the hardcore model on the random d-regular graph for a large range of λ to relate the hardcore measure to its planted version where one first chooses a random independent set and then constructs a uniformly chosen graph compatible with

the set. By exploring the graph in the planted measure by progressively revealing its edges one can show convergence locally on average to the measure $\mathbb{P}_{T_d}$ and via the second moment method this can be extended to the original hardcore distribution. This argument does not imply the stronger local convergence in probability and indeed, if one assumes the picture developed in statistical physics, in the condensation phase one expects local convergence of type A but not convergence of type B or C.

In order to investigate the reconstruction problem it is necessary to work with local convergence in probability. Much of the work of the paper involves showing how the second moment method can be used to imply this stronger notion of convergence. Thus, our proof shows that for the hardcore model, up to the fugacity for which the second moment method holds, the notions A and C of local convergence of measures are equivalent. Our methods are quite general and should apply to a broad range of CSPs and Gibbs measures on graphs. Roughly speaking, one would need to show a corresponding bound on the second moment of the partition function and concavity of the log-partition function. One would also need to show that the partition function changes by a bounded amount when an edge is added and as such, our method should be applicable to non-zero temperature models.

1.3 Outline of the proof

We begin by establishing a lower bound on the reconstruction threshold for the d-regular tree in Section 2, proving Theorem 2. We show that when α is bounded by the lower bound in (1.4) then even for a tree of depth 3 there is already significant loss of information of the spin at the root. In particular we show that if the spin of the root was 1 then the typical posterior probability that the spin of the root is 1 given the spins at level 3 will be less than $\frac{1}{2}$. The result is completed by linearizing the tree posterior probability recursion similarly to [9, 34]. In this part of the proof we closely follow the analysis of [9] who analyzed the reconstruction problem for the Ising model with small external field. We do not require the full strength of their analysis, however, as in our case we are far from the Kesten-Stigum bound. We show that a quantity referred to as the *magnetization* decays exponentially fast to 0. The magnetization provides a bound on the posterior probabilities and this completes the result.

The $\log \log d$ term in our bound on λ_R in Theorem 2 is explained as the first point at which there is significant decay of information at level 3 on the tree. In particular the analysis in Proposition 2.3 part $c)$ is essentially tight. It may be possible to get improved bounds by considering higher depth trees although the description of the posterior distribution necessarily becomes more complex. A sharper analysis of this sort was done in [35] for the coloring model although the method there made crucial use of the symmetry of the states.

The bulk of the paper concerns proving local weak convergence to $\mathbb{P}_{T_d}$ for the hardcore model on the random d-regular graph and this is shown in Theorem 5.4 in Section 5. Our main tool is a new approach to the use of the second moment method. We select say, $n^{\frac{3}{5}}$ randomly chosen vertices in the d-regular random graph, and consider a "punctured" graph with the local neighborhoods of these vertices removed. The punctured graph is used to study the partition function of the original graph conditional on the configuration of the boundaries of these neighborhoods. The second moment method in combination with Azuma's inequality implies that the partition function conditioned on a boundary configuration is within a multiplicative factor of $\exp(O(n^{\frac{1}{2}+\varepsilon}))$ of the expected partition function. We prove convergence in probability locally by showing that it is extremely unlikely that a constant fraction of the $n^{\frac{3}{5}}$ randomly chosen vertices have a local measure which is far from the translation invariant measure on the tree. Indeed, we show that this would entail the existence of a set of configurations

on the set of boundary vertices which has a constant probability under the hardcore measure but expected probability of only $\exp(-cn^{\frac{3}{5}})$. In Proposition 5.1 we show that this is precluded by the second moment method.

One strength of our approach is that it does not require the detailed calculations of the small graph conditioning method. In many spin systems, including the one studied here, the ratio of the second moment of the partition function to the square of the first moment tends to a value greater than 1 and so the second moment method cannot be used to estimate the partition function with probability tending to 1. In this case, small graph conditioning can be used to give estimates on the partition function [40].

The first and second moments of the hardcore partition function for a d-regular random graph are derived in Section 3 while the calculations for the punctured random graph appear in Section 4. The remaining proof involves establishing the requisite bound on the second moment itself. This involves determining the maximum of a function which corresponds to the expected log number of pairs of independent sets in a random regular graph with a given overlap between them. In Proposition 3.2, which is proved in Section 6, we consider the scaled log-partition function, determine its maximum and show that it decays quadratically near its maximum. This is a key fact used in relating the first and second moments of the partition functions of the random graph in Section 4.

Acknowledgements

2   Upper bound on the reconstruction threshold on the tree

In this section we present the proof of Theorem 2. We start by noting that for any finite restriction of T_d to its first ℓ levels, we can use the Markov matrix M as before to generate an independent set σ from the hardcore measure by setting the spin of the root to be occupied with probability α and then applying the matrix as before to generate the spins at the children recursively until we reach the leaves of the tree.

We define the following quantities which are related to the transition probabilities of the Markov matrix M. Let

$$\pi_{01} = \frac{1-\alpha}{\alpha}, \qquad \Delta := \pi_{01} - 1 = \frac{1-2\alpha}{\alpha}, \text{ and}$$
$$\theta := p_{00} - p_{10} = p_{11} - p_{01} = -\frac{\alpha}{1-\alpha}.$$

As mentioned in the introduction, θ, the second eigenvalue of M, plays a particularly important role in the reconstruction problem.

For ease of notation, we will establish non-reconstruction for the model $(\widetilde{T}_d, M)$ where $\widetilde{T}_d$ is the d-ary tree (where each vertex has d children) rather than on the d-regular tree. It is not difficult to modify the recursion we will obtain for the d-ary tree to a recursion for the $(d+1)$-regular tree, showing that non-reconstruction also holds in that case. Finally, we can show that non-reconstruction on the d-regular tree is equivalent to non-reconstruction on the $(d+1)$-regular tree once we note that in equation (1.4) we have that $\alpha_R(d+1) - \alpha_R(d) = o_d(1)$ so the difference can be absorbed in the error term. We will use T to denote a finite tree (not necessarily $(d+1)$-regular) whose root will be denoted by x. We let $\mathbb{P}_T$ denote the measure on independent set of T obtained using the Markov model on the tree where the root is chosen to be occupied with probability α and the spin of each vertex is chosen according to an independent channel M given the spin of its parent. Let $\mathbb{P}_T^1, \mathbb{E}_T^1$ (and resp. $\mathbb{P}_T^0, \mathbb{E}_T^0$ and $\mathbb{P}_T, \mathbb{E}_T$) denote the probabilities and

expectations with respect to the measure on the leaves of T obtained by conditioning on the root x to be 1 (resp. 0, and stationary). Let $L = L(\ell)$ denote the set of vertices of T at depth ℓ and let $\sigma(L)$ denote the configuration on level ℓ. We will write $\mathbb{P}_T(\cdot|\sigma(\ell) = A)$ to denote the measure conditioned on the leaves being in state $A \in \{0,1\}^L$.

As in [9] we analyze the *weighted magnetization of the root* of T which is a function of the random configuration the vertices at distance ℓ from the root and defined as follows:

$$
\begin{aligned}
X = X(\ell) &:= (1-\alpha)^{-1}[(1-\alpha)\mathbb{P}_T(\sigma_x = 1|\sigma(L)) - \alpha\mathbb{P}_T(\sigma_x = 0|\sigma(L))] \\
&= \frac{1}{\pi_{01}}\left(\frac{\mathbb{P}_T(\sigma_x = 1|\sigma(L))}{\alpha} - 1\right).
\end{aligned}
\tag{2.1}
$$

Notice that since $\mathbb{E}_T(\mathbb{P}_T(\sigma_x = 1|\sigma(L))) = \mathbb{P}_T(\sigma_x = 1) = \alpha$, by (2.1), we have that $\mathbb{E}_T(X) = 0$. Also, from the first line of (2.1), it can be verified that $X \leq 1$ since $\mathbb{P}_T(\sigma_x = 1|\sigma(L) = A) \leq 1$ for any A. We will also make use of the following second moments of the magnetization.

$$
\overline{X} = \overline{X}(\ell) := \mathbb{E}_T(X^2), \quad \overline{X}_1 = \overline{X}_1(\ell) := \mathbb{E}_T^1(X^2), \quad \overline{X}_0 = \overline{X}_0(\ell) := \mathbb{E}_T^0(X^2)
\tag{2.2}
$$

The following equivalent definition of non-reconstruction is well known and follows from the definition in (1.3) using (2.1).

Proposition 2.1. *Non-reconstruction for the model $(\widetilde{T}_d, M)$ is equivalent to*

$$
\lim_{\ell \to \infty} \overline{X} = 0.
$$

In the remainder of the proof we derive bounds for $\overline{X}$. We begin by showing that already for a 3 level tree, $\overline{X}$ becomes small. Then we establish a recurrence along the lines of [9] that shows that once $\overline{X}$ is sufficiently small, it must converge to 0. As this part of the derivation follows the calculation in [9] we will adopt their notation in places. Non-reconstruction is then a consequence of Proposition 2.1. In the next lemma we determine some basic properties of X.

Lemma 2.2. *For any finite tree T, the following relations hold:*

a) $\mathbb{E}_T(X) = \alpha\mathbb{E}_T^1(X) + (1-\alpha)\mathbb{E}_T^0(X)$.

b) $\overline{X} = \alpha\overline{X}_1 + (1-\alpha)\overline{X}_0$.

c) $\mathbb{E}_T^1(X) = \pi_{01}\overline{X}$ and $\mathbb{E}_T^0(X) = -\overline{X}$.

Proof. Note that for any random variable $Y = Y(A)$ which depends only on the states at the leaves, we have $E_T(Y) = \alpha\mathbb{E}_T^1(Y) + (1-\alpha)\mathbb{E}_T^0(Y)$. Parts $a)$ and $b)$ therefore follow since X is a random variable that is a function of the states at the leaves. For part $c)$ we proceed as follows. The first and last equalities below follow from (2.1).

$$
\begin{aligned}
\mathbb{E}_T^1(X) &= \pi_{01}^{-1}\sum_A \mathbb{P}_T(\sigma(L) = A|\sigma_x = 1)\left(\frac{\mathbb{P}_T(\sigma_x = 1|\sigma(L) = A)}{\alpha} - 1\right) \\
&= \pi_{01}^{-1}\sum_A \mathbb{P}_T(\sigma(L) = A)\frac{\mathbb{P}_T(\sigma_x = 1|\sigma(L) = A)}{\alpha}\left(\frac{\mathbb{P}_T(\sigma_x = 1|\sigma(L) = A)}{\alpha} - 1\right) \\
&= \pi_{01}^{-1}\left(\frac{\mathbb{E}_T((\mathbb{P}_T(\sigma_x = 1|\sigma(L)))^2)}{\alpha^2} - 1\right) \\
&= \pi_{01}\mathbb{E}_T(X^2)
\end{aligned}
$$

The second part of $c)$ follows by combining this with $a)$ and the fact that $\mathbb{E}_T(X) = 0$. $\square$

The following proposition estimates typical posterior probabilities which we will use to bound $\overline{X}$. Let $T(\ell)$ denote the tree which is the restriction of $\widetilde{T}_d$ to its first ℓ levels. For a finite tree T, let T^i denote the subtree rooted at the child of the root u_i and let L_i denote the restriction of L to the subtree T_i.

Proposition 2.3. *For a finite d-ary tree T we have that*

a) For any configuration at the leaves $A = (A_1, \cdots, A_d)$,

$$\mathbb{P}_T(\sigma_x = 0 | \sigma(L) = A) = \left(1 + \lambda \prod_i \mathbb{P}_{T^i}(\sigma_{u_i} = 0 | \sigma_{L_i} = A_i)\right)^{-1}.$$

b) Let $\mathcal{G}$ be the set of leaf configurations

$$\mathcal{G} = \left\{ \sigma(L) \mid \mathbb{P}_T(\sigma_x = 0 | \sigma(L)) = \frac{1}{2}\left(1 + \frac{1}{1 + 2\lambda}\right)\right\}.$$

Then

$$\frac{\mathbb{P}_T^0(\sigma(L) \in \mathcal{G})}{\mathbb{P}_T^1(\sigma(L) \in \mathcal{G})} = \frac{\alpha}{1 - \alpha}\frac{1 + \lambda}{\lambda}.$$

c) Let $\beta > \log 2 - \log\log 2$ and $\alpha = \frac{1}{d}\left(\log d + \log\log d - \log\log\log d - \beta\right)$. Then in the 3-level d-ary tree $T(3)$ we have that

$$\mathbb{E}_{T(3)}^1(\mathbb{P}(\sigma_x = 1 | \sigma(L))) \leq \frac{1}{2}.$$

Proof. Part $a)$ is a consequence of standard tree recursions for Markov models established using Bayes rule. For part $b)$ first note that

$$
\begin{aligned}
\mathbb{P}_T(\sigma_x = 1 \mid \sigma(L) \in \mathcal{G}) &= 1 - \mathbb{P}_T(\sigma_x = 0 \mid \sigma(L) \in \mathcal{G}) \\
&= \frac{1}{2}\left(1 - \frac{1}{1 + 2\lambda}\right).
\end{aligned}
\tag{2.3}
$$

Now,

$$
\begin{aligned}
\mathbb{P}_T^0(\sigma(L) \in \mathcal{G}) &= \frac{\mathbb{P}_T(\sigma_x = 0 \mid \sigma(L) \in \mathcal{G})\mathbb{P}_T(\sigma(L) \in \mathcal{G})}{1 - \alpha} \\
&= \frac{\alpha}{1 - \alpha}\frac{1 + \lambda}{\lambda}\left(\frac{\mathbb{P}_T(\sigma_x = 1 \mid \sigma(L) \in \mathcal{G})\mathbb{P}_T(\sigma(L) \in \mathcal{G})}{\alpha}\right) \\
&= \frac{\alpha}{1 - \alpha}\frac{1 + \lambda}{\lambda}\mathbb{P}_T^1(\sigma(L) \in \mathcal{G})
\end{aligned}
$$

where the first and third equations follow by the definition of conditional probability and the second follows from (2.3) and the definition of $\mathcal{G}$ which establishes $b)$.

For part $c)$, we start by calculating the probability of certain posterior probabilities for trees of small depth. With our assumption on α we have that

$$\lambda = \frac{\alpha}{1 - 2\alpha}\left(1 + \frac{\alpha}{1 - 2\alpha}\right)^d = \frac{(1 + o_d(1))e^{-\beta}\log^2 d}{\log\log d}.
\tag{2.4}$$

Since $\sigma(L) = \{0\}^L$ under $\mathbb{P}_{T(1)}^1$, by part $a)$ we have that

$$\mathbb{P}_{T(1)}^1(\sigma_x = 0 | \sigma(L)) = \frac{1}{1 + \lambda} \ w.p. \ 1.$$

Also,

$$\mathbb{P}_{T(1)}(\forall\, i,\ u_i = 0 | \sigma_x = 0) = \left(\frac{1 - 2\alpha}{1 - \alpha}\right)^d.$$

Using the two equations above, we have that

$$\mathbb{P}^0_{T(1)}(\sigma_x = 0 | \sigma(L)) = \begin{cases} 1 & w.p.\ 1 - \left(\frac{1-2\alpha}{1-\alpha}\right)^d \\ \frac{1}{1+\lambda} & w.p.\ \left(\frac{1-2\alpha}{1-\alpha}\right)^d. \end{cases}$$

The first case above corresponds to leaf configurations of the tree $T(1)$ where at least one of the leaves is 1, while the second case corresponds to the configurations where all the leaves are 0. Next, applying part $a)$ to a tree of depth 2, we have

$$\mathbb{P}^1_{T(2)}(\sigma_x = 0 | \sigma(L)) = \frac{1}{1 + \lambda \prod_i \mathbb{P}^0_{T(1)}(\sigma_{u_i} = 0 | \sigma(L_i))}.$$

Using this expression we write this conditional probability based on the leaf configurations of the subtrees of the root of depth 1 below.

$$\mathbb{P}^1_{T(2)}(\sigma_x = 0 | \sigma(L)) = \begin{cases} \frac{1}{1+\lambda} & w.p.\ \left(1 - \left(\frac{1-2\alpha}{1-\alpha}\right)^d\right)^d \\ \frac{1}{2}\left(1 + \frac{1}{1+2\lambda}\right) & w.p.\ \left(1 - \left(\frac{1-2\alpha}{1-\alpha}\right)^d\right)^{d-1}\left(\frac{1-2\alpha}{1-\alpha}\right)^d d & (2.5) \\ > \frac{1}{2}\left(1 + \frac{1}{1+2\lambda}\right) & o.w. \end{cases}$$

The first case above corresponds to the situation when each subtree of the root of depth 1 has a leaf configuration where at least one of the leaves is 1. The second case is when one of the d subtrees has a leaf configuration where all leaves are 0, while the remaining subtrees have leaf configurations where at least one leaf is 1. The third case corresponds to the remaining possibilities.

By part $b)$ and (2.5), after substituting the expressions for λ from (2.4), we have that

$$\begin{aligned}
\mathbb{P}^0_{T(2)}(\sigma(L) \in \mathcal{G}) &= \frac{\alpha}{1-\alpha}\frac{1+\lambda}{\lambda}\mathbb{P}^1_{T(2)}[\sigma(L) \in \mathcal{G}] \\
&= \frac{\alpha(1+\lambda)}{\lambda(1-\alpha)}\left(1 - \left(\frac{1-2\alpha}{1-\alpha}\right)^d\right)^{d-1}\left(\frac{1-2\alpha}{1-\alpha}\right)^d d \\
&\geq (1 - o_d(1))\frac{e^\beta \log\log d}{d}. \quad (2.6)
\end{aligned}$$

We can now calculate the values of $P^1_{T(3)}(\sigma_x = 0 | \sigma(L))$ as follows. By part $a)$

$$\mathbb{P}^1_{T(3)}(\sigma_x = 0 | \sigma(L)) = \frac{1}{1 + \lambda \prod_i \mathbb{P}^0_{T(2)}(\sigma_{u_i} = 0 | \sigma(L_i))}.$$

Denote

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

Thus, p is the probability that if we started with $\sigma_x = 0$ in $T(2)$, the configuration at the leaves is from $\mathcal{G}$. If we start with $\sigma_x = 1$ in $T(3)$, the number subtrees of the root with leaf configurations in $\mathcal{G}$ is distributed binomially and will be about dp. By Chernoff bounds, and the bound on p from (2.6),

$$\mathbb{P}\left(Bin(d, p) < e^\beta \log\log d - 2\sqrt{e^\beta \log\log d}\right) < \frac{1}{3}.$$

Finally, by the definition of $\mathcal{G}$,

$$\mathbb{P}^0_{T(2)}(\sigma_{u_i} = 0 | \sigma(L) \in \mathcal{G}) = \frac{1}{2}\left(1 + \frac{1}{1 + 2\lambda}\right)$$

and hence, by (2.4)

$$
\begin{aligned}
\mathbb{E}^1_{T(3)}(\mathbb{P}(\sigma_x = 1 | \sigma(L))) &= \mathbb{E}^1_{T(3)}(1 - \mathbb{P}(\sigma_x = 0 | \sigma(L))) \\
&\leq \left(1 - \frac{1}{1 + \lambda(2(1 - o_d(1)))^{-(e^\beta \log\log d - 2\sqrt{e^\beta \log\log d})}}\right)\frac{2}{3} + \frac{1}{3}.
\end{aligned}
$$

By taking d large enough above, for β as in the assumptions of the claim, the term in the bracket above can be made arbitrarily small and thus we conclude that

$$\mathbb{E}^1_{T(3)}(\mathbb{P}(\sigma_x = 1 | \sigma(L))) \leq \frac{1}{2}. \qquad \square$$

Lemma 2.4. *Let $\beta > \log 2 - \log\log 2$ and $\alpha = \frac{1}{d}\left(\log d + \log\log d - \log\log\log d - \beta\right)$. For d large enough,*

$$\overline{X}(3) \leq \frac{\alpha}{2}.$$

Proof. Combining (2.1), part $c)$ of Lemma 2.2, and part $c)$ of Proposition 2.3,

$$
\begin{aligned}
\overline{X}(3) &= \frac{1}{\pi_{01}^2}\left(\frac{\mathbb{E}^1_{T_3}(\mathbb{P}(\sigma_x = 1 \mid \sigma(L)))}{\alpha} - 1\right) \\
&\leq \frac{1}{\pi_{01}^2}\left(\frac{1}{2\alpha} - 1\right) \\
&\leq \frac{\alpha}{2}. \qquad \square
\end{aligned}
$$

Next, we present a recursion for $\overline{X}$ and complete the proof of the main result of this section. The development of the recursion follows the steps in [9] closely. In particular, Lemmas 2.5, 2.6, and 2.7 are from [9], so we omit some of the calculations. We refer the reader to [9] for the complete details of the proofs.

Magnetization of a child

Let T be a finite tree with root x. Let x' be a child of x and let T' be the subtree of T rooted at x' (see Figure 1). Let A' be the restriction of A to the leaves L' of T'. Let $X' = X(A')$ denote the magnetization at x', the root of T'. That is

$$X' := \frac{1}{\pi_{01}}\left(\frac{\mathbb{P}_T(\sigma_{x'} = 1 \mid \sigma(L'))}{\alpha} - 1\right)$$

when σ is the independent set from the hardcore measure obtained by the Markov process on T. Note that X' has the same distribution as the weighted magnetization of the root of a tree of the same depth as T'. We will abuse notation and use $\mathbb{E}_{T'}(X'), \mathbb{E}_{T'}((X')^2)$ etc. to mean the moments of the weighted magnetization of the root of a tree of the same depth as T'.

Lemma 2.5. *We have*

a) $\mathbb{E}^1_T(X') = \theta\mathbb{E}^1_{T'}(X')$ *and* $\mathbb{E}^0_T(X') = \theta\mathbb{E}^0_{T'}(X')$.

b) $\mathbb{E}^1_T((X')^2) = (1 - \theta)\mathbb{E}_{T'}((X')^2) + \theta\mathbb{E}^1_{T'}((X')^2)$.
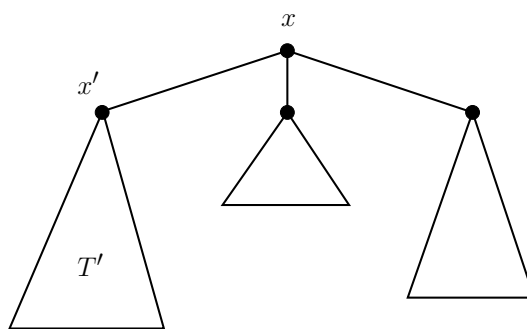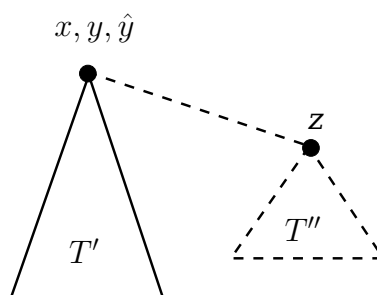
Figure 1: The tree T.

Figure 2: The tree T obtained by merging T' and T''. The dashed subtree is $\hat{T}$.

c) $\mathbb{E}_T^0((X')^2) = (1-\theta)\mathbb{E}_{T'}((X')^2) + \theta\mathbb{E}_{T'}^0((X')^2)$.

The proof follows from the first part of Lemma 2.2 and the Markov property when we condition on x.

In [9], the authors derive moment recursions for the magnetization for the Ising model by defining an *Add-Merge* operation for trees. In the case of the hardcore model, the moment recursions can be derived in exactly the same way and we describe the sequence of intermediate results without full calculations, referring the reader to [9] for the details. The idea is to write the effect on the magnetization of the root of a tree when an edge is added to the root and another tree is attached to this edge by its root. Referring to Figure 2, let T' (resp. T'') be a finite tree rooted at y (resp. z) with the channel on all edges being given by M, leaf states A' (resp A'') and weighted magnetization at the root Y (resp. Z). Add an edge $(\hat{y}, z)$ to T'' to obtain a new tree $\hat{T}$. Then merge $\hat{T}$ with T' by identifying $y = \hat{y}$ to obtain a new tree T. To avoid ambiguities, denote by x the root of T and X the magnetization of the root of T. We let $A = (A', A'')$ be the leaf state of T. Let $\hat{Y}$ be the magnetization of the root of $\hat{T}$.

Note: In the above construction, the vertex y is a vertex "at the same level" as x, and not a child of x as it was in Lemma 2.5.

Lemma 2.6. *With the notation above, $\hat{Y} = \theta Z$.*

The proof follows by applying Bayes rule, the Markov property and Lemma 2.2. These facts also imply the following result.

Lemma 2.7. *For any tree $\hat{T}$,*

$$X = \frac{Y + \hat{Y} + \Delta Y\hat{Y}}{1 + \pi_{01}Y\hat{Y}}$$

where $\Delta := \pi_{01} - 1 = \frac{1-2\alpha}{\alpha}$.

With these lemmas in hand we can derive a recursive upper bound on the second moments. We will use the expansion

$$\frac{1}{1+r} = 1 - r + r^2 \frac{1}{1+r}.$$

Taking $r = \pi_{01} Y \hat{Y}$, by Lemma 2.7 we have

$$
\begin{aligned}
X &= (Y + \hat{Y} + \Delta Y \hat{Y}) \left(1 - \pi_{01} Y \hat{Y} + (\pi_{01} Y \hat{Y})^2 \frac{1}{1 + \pi_{01} Y \hat{Y}} \right) \\
&= Y + \hat{Y} + \Delta Y \hat{Y} - \pi_{01} Y \hat{Y} \left(Y + \hat{Y} + \Delta Y \hat{Y} \right) + (\pi_{01})^2 (Y \hat{Y})^2 X \\
&\leq Y + \hat{Y} + \Delta Y \hat{Y} - \pi_{01} Y \hat{Y} \left(Y + \hat{Y} + \Delta Y \hat{Y} \right) + (\pi_{01})^2 (Y \hat{Y})^2 \qquad (2.7)
\end{aligned}
$$

where the last inequality follows since $X \leq 1$ with probability 1.

Let $\rho' = \overline{Y}_1 / \overline{Y}$ and $\rho'' = \overline{Z}_1 / \overline{Z}$. The second moments $\overline{Y}, \overline{Z}, \overline{Y}_1$ and $\overline{Z}_1$ are defined analogously to (2.2) for the variables Y and Z according to the appropriate measures over the trees T' and T''.

The following relations can be derived by applying respectively: Lemma 2.2 (part c)) for the first two equalities, the definition of $\overline{Y}_1$, Lemma 2.2 (part c)) in combination with Lemma 2.6, and finally, Lemma 2.5.

$$
\begin{aligned}
\mathbb{E}_T^1(X) &= \pi_{01} \overline{X}, \quad \mathbb{E}_T^1(Y) = \pi_{01} \overline{Y}, \quad \mathbb{E}_T^1(Y^2) = \overline{Y} \rho' \\
\mathbb{E}_T^1(\hat{Y}) &= \pi_{01} \theta^2 \overline{Z}, \quad \mathbb{E}_T^1(\hat{Y}^2) = \theta^2 \overline{Z}((1 - \theta) + \theta \rho'').
\end{aligned} \qquad (2.8)
$$

Applying $(\pi_{01})^{-1} E_T^1(\cdot)$ to both sides of (2.7), and using the relations (2.8), we obtain the following inequality. Below, we have used the fact that when we condition on the root x being 1, the magnetizations Y and $\hat{Y}$ are independent.

$$
\begin{aligned}
\overline{X} &\leq \overline{Y} + \theta^2 \overline{Z} + \theta^2 \Delta \pi_{01} \overline{Y} \cdot \overline{Z} - \pi_{01} \theta^2 \overline{Y} \cdot \overline{Z} \rho' - \pi_{01} \theta^2 \overline{Y} \cdot \overline{Z}((1 - \theta) + \theta \rho'') \\
&\quad - \Delta \theta^2 \overline{Y} \cdot \overline{Z} \rho'((1 - \theta) + \theta \rho'') + \pi_{01} \theta^2 \overline{Y} \cdot \overline{Z} \rho'((1 - \theta) + \theta \rho'') \\
&= \overline{Y} + \theta^2 \overline{Z} - \pi_{01} \theta^2 \overline{Y} \overline{Z} (\mathcal{A} - \Delta \mathcal{B})
\end{aligned}
$$

where

$$
\begin{aligned}
\mathcal{A} &= \rho' + (1 - \rho')((1 - \theta) + \theta \rho''), \\
\text{and} \quad \mathcal{B} &= 1 - (\pi_{01})^{-1} \rho'((1 - \theta) + \theta \rho'') = 1 - \frac{\alpha}{1 - 2\alpha} \rho'((1 - \theta) + \theta \rho'').
\end{aligned}
$$

If $\mathcal{A} - \Delta \mathcal{B} \geq 0$, this would already give a sufficiently good recursion to show that $\overline{X}(\ell)$ goes to 0, so we will assume is negative and try to get a good (negative) lower bound. First note that by their definition $\rho', \rho'' \geq 0$. Further since $\overline{Y} = \alpha \overline{Y}_1 + (1 - \alpha) \overline{Y}_0$,

$$\rho' \leq \frac{1}{\alpha}.$$

Similarly,

$$\rho'' \leq \frac{1}{\alpha}.$$

Since $\mathbb{E}_T^1(\hat{Y}^2)$ and $\overline{Z} \geq 0$, it follows from (2.8) that $(1 - \theta) + \theta \rho'' \geq 0$. Together with the fact that $\rho' \geq 0$, this implies that $\mathcal{B} \leq 1$.

Since $\mathcal{A}$ is multi-linear in (ρ', ρ''), to minimize it, it is sufficient to consider the extreme cases. When $\rho' = 0$, $\mathcal{A}$ is minimized at the upper bound of ρ'' and hence

$$\mathcal{A} \geq 1 - \pi_{01} \frac{\alpha}{1 - \alpha} = 0.$$

When $\rho' = \frac{1}{\alpha}$,

$$\mathcal{A} = \frac{1}{\alpha} + \left(1 - \frac{1}{\alpha}\right)[1 - \theta(1 - \rho'')] \geq 0.$$

Hence, we have

$$\overline{X} \leq \overline{Y} + \theta^2 \overline{Z} + \frac{1 - 2\alpha}{1 - \alpha} \overline{Y} \cdot \overline{Z}.$$

We can apply this bound to the d-ary tree $k - 1$ times in succession for the first $k - 1$ children of the root. Applying Lemma 2.6 and simplifying one can verify that

$$\overline{X} \leq \frac{1 - \alpha}{1 - 2\alpha} \theta^2 \left(\left(1 + \overline{Z} \frac{1 - 2\alpha}{1 - \alpha}\right)^k - 1\right).$$

Applying the following inequality

$$|(1 + x)^k - 1| \leq e^{|x|k} - 1 = \int_0^{|x|k} e^s \, ds \leq e^{|x|k} k |x|$$

implies the recursion below.

Proposition 2.8. *If for some ℓ, $\overline{X}(\ell) \leq \frac{\alpha}{2}$, we have that*

$$\overline{X}(\ell + 1) \leq \theta^2 \left(\frac{1 - \alpha}{1 - 2\alpha}\right)^2 e^{\frac{1}{2}\alpha d} d \overline{X}(\ell).$$

Thus if $\left(\frac{\alpha}{1 - 2\alpha}\right)^2 e^{\frac{1}{2}\alpha d} d < 1$ then it follows from the recursion that

$$\lim_{\ell \to \infty} \overline{X}(\ell) = 0. \tag{2.9}$$

When $\alpha = \frac{1}{d}\left(\log d + \log\log d - \log\log\log d - \beta\right)$ and $\beta > \log 2 - \log\log 2$, by Lemma 2.4, for d large enough, $\overline{X}(3) \leq \frac{\alpha}{2}$. Hence by equation (2.9) we have that $\overline{X}(\ell) \to 0$ and so by Proposition 2.1 we have non-reconstruction. Since reconstruction is monotone in λ and hence in α it follows that we have non-reconstruction for $\alpha \leq \alpha_R$ for large enough d. This completes the proof of Theorem 2.

3 Partition function of the hardcore model for random d-regular graph

In this section, we derive expressions for the first and second moments of the hardcore partition function for the d-regular random graph. The calculations are along the lines of those in [31] and we adopt their notation here. We will work with the configuration model for random graphs, described below, in order to simplify the calculations.

3.1 Configuration model

Let $\mathcal{H}(n, d)$ denote the set of all d-regular (multi)graphs on n vertices and $\mathcal{G}(n, d)$ the subset of d-regular simple graphs. The analysis of the properties of a random graph in $\mathcal{G}(n, d)$ can often be simplified by making use of the *configuration model*, introduced by Bollóbas [8]. Fix d and n such that dn is even. Define a d-regular multigraph on n vertices via the configuration model as follows. Begin by replacing each vertex with d distinct copies and then generate a uniformly random pairing of the dn distinct points. Finally, collapse the d copies corresponding to each vertex back into one vertex, obtaining a uniformly random multigraph in $\mathcal{H}(n, d)$. Let S be the event that the multigraph obtained

is simple. Clearly, on the event S, the graph obtained is uniformly distributed over $\mathcal{G}(n,d)$. Moreover, for fixed d,

$$\mathbb{P}(S) = (1 + o_n(1)) \exp\left(\frac{1 - d^2}{4}\right). \tag{3.1}$$

Since the probability in (3.1) is uniformly bounded away from 0, any event that holds asymptotically with high probability for $\mathcal{H}(n,d)$ also holds asymptotically with high probability when we condition on S, i.e. for $\mathcal{G}(n,d)$. In what follows, by "d-regular random graph", we will mean the multigraph generated by the configuration model, unless mentioned otherwise.

One useful property of the configuration model that we will make use of repeatedly is that the pairings of the dn distinct points may be revealed sequentially. That is, given a vertex v, we may reveal the pairings of its d copies one by one so that the distribution of pairings over the remaining unmatched points remains uniform.

In the sequel, we will use $f(n) = \widetilde{\Theta}(g(n))$ to mean equality of the functions up to polynomial factors in n. We will assume throughout that quantities of the form $an, \alpha n, \gamma n, \varepsilon n$ are integers. We use "with high probability" to mean with probability going to 1 as $n \to \infty$. We will use σ_u to denote the restriction of an independent set σ of the graph to a vertex u. Finally, the restriction of σ to a subset of vertices S will be denoted by $\sigma(S)$.

3.2 The first moment of the partition function

In this section, we calculate the first moment of the partition function for the hardcore model on the d-regular random graph. For an independent set $\sigma \in I(G)$, let $|\sigma|$ denote the number of vertices in I. For fugacity λ, the partition function is given by

$$Z_G = Z_G(\lambda) = \sum_{\sigma \in I(G)} \lambda^{|\sigma|}.$$

Let $0 \le \alpha \le 1/2$ and let $Z_{G,\alpha} = Z_{G,\alpha}(\lambda)$ be the contribution to the partition function from independent sets of size αn, i.e.

$$Z_{G,\alpha} := \sum_{\sigma \in I(G): |\sigma| = \alpha n} \lambda^{\alpha n}, \qquad Z_G = \sum_{\alpha} Z_{G,\alpha}.$$

The following approximation will be useful in simplifying the probabilities obtained in the sequel. Let $a > 0$ be a constant. Then, by Stirling's approximation,

$$\prod_{j=1}^{an} j = \exp\left(n \int_0^a \log(x)dx + an\log(n) + O(\log n)\right). \tag{3.2}$$

Let

$$H(x) = -x\log(x) - (1-x)\log(1-x).$$

Let $G \sim \mathcal{H}(n,d)$. Fix $\lambda > 0$ and $0 \le \alpha \le \frac{1}{2}$. The first moment of $Z_{G,\alpha}$ is given by

$$\mathbb{E}(Z_{G,\alpha}) = \binom{n}{\alpha n} \lambda^{\alpha n} \prod_{i=0}^{\alpha nd - 1} \frac{(1-\alpha)nd - i}{nd - 1 - 2i} = \widetilde{\Theta}(1) \exp\left(n\Phi(\alpha)\right) \tag{3.3}$$

where

$$\Phi(\alpha) = \Phi(\alpha, \lambda) = H(\alpha) + \alpha \log(\lambda) + d\left((1-\alpha)\log(1-\alpha) - \left(\frac{1-2\alpha}{2}\right)\log(1-2\alpha)\right). \tag{3.4}$$

The first equality in (3.3) follows by calculating the probability in the configuration model that a given subset of αn vertices is an independent set, i.e. that the vertices in the subset are not matched to vertices in the subset itself. The second equality follows by (3.2).

For $\lambda > 0$, it can be verified that the maximum of Φ is achieved at $\alpha^* = \alpha^*(\lambda, d)$, which is the solution to the equation

$$\lambda \frac{1-\alpha}{\alpha} \left(\frac{1-2\alpha}{1-\alpha} \right)^d = 1 \tag{3.5}$$

which is obtained by differentiating Φ. Solving, we obtain that

$$\alpha^*(\lambda, d) = (1 + \varepsilon(\lambda, d)) \frac{\log(\lambda d / \log(\lambda d))}{d}$$

where $\varepsilon(\lambda, d)$ tends to 0 as d tends to ∞. Notice that the relation (3.5) between α, λ and d is equivalent to (1.2).

3.3 The second moment of the partition function

To estimate the second moment $\mathbb{E}((Z_{G,\alpha})^2)$, we consider the contributions from pairs of independent sets σ', σ'' each of size αn. We divide this according to the size of the overlap $|\sigma' \cap \sigma''| = \gamma n$ and according to the number εn of edges of the graph which go from each of σ', σ'' to the complement $V \setminus (\sigma' \cup \sigma'')$. Call this contribution $Z_{G,\alpha,\gamma,\varepsilon}^{(2)}$. That is, for $(\alpha, \gamma, \varepsilon)$ in the region

$$\mathcal{R} = \left\{ (\alpha, \gamma, \varepsilon) : 0 \leq \alpha, \gamma, \varepsilon \leq \frac{1}{2}, \ \alpha - \gamma - \varepsilon \geq 0, \ 1 - 2\alpha - 2\varepsilon \geq 0 \right\}, \tag{3.6}$$

we define

$$Z_{G,\alpha,\gamma,\varepsilon}^{(2)} := \lambda^{2\alpha n} \Big| \Big\{ \sigma', \sigma'' \in I(G) : |\sigma'| = |\sigma''| = \alpha n, |\sigma' \cap \sigma''|$$
$$= \gamma n, |E_G(\sigma', V \setminus (\sigma' \cup \sigma''))| = |E_G(\sigma'', V \setminus (\sigma' \cup \sigma''))| = \varepsilon n \Big\} \Big|$$

and

$$\mathbb{E}((Z_{G,\alpha})^2) = \sum_{\gamma,\varepsilon} \mathbb{E}\left(Z_{G,\alpha,\gamma,\varepsilon}^{(2)} \right).$$

Calculating the probability in the configuration model that a pair of subsets of vertices σ' and σ'' as above are both independent sets, we have that for $G \sim \mathcal{H}(n, d)$,

$$\mathbb{E}\left(Z_{G,\alpha,\gamma,\varepsilon}^{(2)} \right) = \lambda^{2\alpha n} \binom{n}{\alpha n} \binom{\alpha n}{\gamma n} \binom{(1-\alpha)n}{(\alpha - \gamma)n} \left(\frac{\prod_{i=0}^{\gamma nd-1} (1 - 2\alpha + \gamma)dn - i}{\prod_{i=0}^{\gamma nd-1} dn - 1 - 2i} \right) \times$$

$$\times \left(\frac{\prod_{i=0}^{\varepsilon dn-1} ((1-2\alpha)dn - i) \cdot \prod_{i=0}^{(\alpha-\gamma-\varepsilon)dn-1} (\alpha - \gamma)dn - i}{\prod_{i=0}^{(\alpha-\gamma)dn-1} ((1 - 2\gamma)dn - 1 - 2i)} \cdot \frac{\prod_{i=0}^{\varepsilon dn-1} (1 - 2\alpha - \varepsilon)dn - i}{\prod_{i=0}^{\varepsilon dn-1} (1 - 2\alpha)dn - 1 - 2i} \right). \tag{3.7}$$

The following function arises in the approximation of the expression in (3.7)

$$f(\alpha, \gamma, \varepsilon) = 2\alpha \log(\lambda) + H(\alpha) + H_1(\gamma, \alpha) + H_1(\alpha - \gamma, 1 - \alpha) + d\Psi_2(\alpha, \gamma, \varepsilon)$$

where

$$H_1(x, y) = -x(\log(x) - \log(y)) + (x - y)(\log(y - x) - \log(y))$$

and

$$\Psi_2(\alpha, \gamma, \varepsilon) = H_1(\varepsilon, \alpha - \gamma) + \int_0^\gamma \log(1 - 2\alpha + \gamma - x)dx - \int_0^\gamma \log(1 - 2x)dx$$

$$+ \int_0^\varepsilon \log(1 - 2\alpha - x)dx + \int_0^{\alpha - \gamma - \varepsilon} \log(\alpha - \gamma - x)dx - \int_0^{\alpha - \gamma} \log(1 - 2\gamma - 2x)dx$$

$$+ \int_0^\varepsilon \log(1 - 2\alpha - \varepsilon - x)dx - \int_0^\varepsilon \log(1 - 2\alpha - 2x)dx.$$

In particular, in Section 6 we will show that the logarithm of $\mathbb{E}(Z_{G,\alpha,\gamma,\varepsilon}^{(2)})$ scaled by n is well approximated by f, and for $\lambda < \lambda_c$, f decays quadratically around its maximum.

Proposition 3.1. *Let $G \sim \mathcal{H}(n, d)$ and $0 \leq \alpha \leq \frac{1}{2}$. Then,*

$$\mathbb{E}\left(Z_{G,\alpha,\gamma,\varepsilon}^{(2)}\right) = \exp\left(nf(\alpha, \gamma, \varepsilon) + O(\log(n))\right).$$

For any α, define $\hat{\gamma} = \hat{\gamma}(\alpha) := \alpha^2$ and $\hat{\varepsilon} = \hat{\varepsilon}(\alpha) := \alpha(1 - 2\alpha)$. We will also use the shorthand $\gamma^* := \hat{\gamma}(\alpha^*)$ and $\varepsilon^* := \hat{\varepsilon}(\alpha^*)$.

Proposition 3.2. *Let $\lambda < \lambda_c$, and $(\alpha, \gamma, \varepsilon) \in \mathcal{R}$. Then, the function $f(\alpha, \gamma, \varepsilon)$ attains its maximum at $(\alpha^*, \gamma^*, \varepsilon^*)$ and is strictly concave at this point. In particular for some $C = C(d, \lambda)$,*

$$f(\alpha^*, \gamma^*, \varepsilon^*) - f(\alpha, \gamma, \varepsilon) \geq C(|\alpha - \alpha^*|^2 + |\gamma - \gamma^*|^2 + |\varepsilon - \varepsilon^*|^2).$$

Finally, the following second moment-type bound is also proved in Section 6.

Proposition 3.3. *Let $G \sim \mathcal{H}(n, d)$ and $\lambda < \lambda_c$. Then,*

$$\mathbb{E}((Z_G)^2) = \widetilde{\Theta}(1)(\mathbb{E}(Z_G))^2.$$

4 Partition function of a punctured random graph

In this section we study the effect on the hardcore measure of a d-regular random graph of conditioning on the spins of a small number of the vertices. In order to do this, we analyze the partition function of a punctured graph $\tilde{G}$ obtained from a d-regular random graph G by deleting a small fraction of vertices and their neighborhoods. Define the following quantities with respect to a graph $G = (V, E)$. Let $d(u, v) = d_G(u, v)$ denote the distance between two vertices $u, v \in V$. For a vertex $u \in V$ and integer r, the r-*neighborhood of u*, denoted $B_r(u)$ and its (vertex) boundary are defined as

$$B_r(u) := \{v \in V : d(u, v) \leq r\}, \ \partial B_r(u) := B_r(u) \setminus B_{r-1}(u).$$

Lemma 4.1. *Let $G = (V, E) \sim \mathcal{H}(n, d)$. Let $S \subset V$ be a random subset of vertices with $|S| = n^{3/5}$ and let r be some large constant. Then, the expected number of $u \in S$ such that the neighborhood $B_r(u)$ contains another vertex in S is $O(n^{1/5})$ and the probability of a neighbourhood with 3 vertices in S is $O(n^{-1/5})$. Furthermore, with high probability, for all $u \in S$ the neighbourhood $B_r(u)$ is a tree.*

Proof. The first claim follows by linearity of expectation. By the independence of the choice of S and the graph G, we can first choose the graph and then choose which vertices are in S. Choose the edges of the graph, fix a random vertex u and consider the neighborhood $B_r(u)$. It can be calculated that the chance that another vertex in S is in

$B_r(u)$ is $O(n^{-2/5})$ (where the constant may depend on d as well as r). Summing over the $n^{3/5}$ vertices of S, we obtain the claimed bound.

The probability that there is a neighborhood with 3 vertices of S can be bounded by the union bound. We make use of the independence of S and G as before to choose the random graph first. The probability that a fixed neighborhood around a vertex u has 2 other vertices of S can be calculated to be bounded by $O(n^{-4/5})$.

Finally, the number of cycles of length at most $2r$ has constant expected value and thus with high probability no neighborhood of a vertex in S contains such a cycle as n tends to ∞. $\qquad\square$

Let $G = (V, E) \sim \mathcal{H}(n, d)$ and fix a large constant r. Let $S \subset V$ be a uniformly chosen set of vertices with $|S| = n^{3/5}$. Let $\tilde{G} = (\tilde{V}, \tilde{E})$ be the graph obtained by deleting from V the set of vertices $\cup_{v \in S} B_{r-1}(v)$ and any edges adjacent to these vertices. Define

$$B := \tilde{G} \bigcap \left(\bigcup_{u \in S} \partial B_r(u) \right). \tag{4.1}$$

Let

$$S' = \{s \in S \ : \ \forall s' \in S \setminus \{s\}, B_r(s) \cap B_r(s') = \varnothing, B_r(s) \text{ is a tree}\} \tag{4.2}$$

Let $k = |S'|$. Let $s_1, \ldots, s_k$ be an arbitrary ordering of the elements of S' and for $1 \le i \le k$ define $W_i := \partial B_r(s_i)$. Define $W_{k+1} := B \setminus \cup_{i=1}^{k} W_i$.

Corollary 4.2. *The vertices of B have degree $d-1$ or $d-2$ in $\tilde{G}$ with high probability. With high probability, the number of vertices in $\tilde{G}$ of degree $d-2$ is $O(n^{\frac{1}{5}})$ and the number of vertices of degree $d-1$ is $n^{\frac{3}{5}}(1 - o_n(1))d(d-1)^r$. The size of S', $k = (1 - o_n(1))n^{\frac{3}{5}}$ with high probability.*

Proof. Suppose a vertex $v \in B$ is in $\partial B_r(u_1) \cap \partial B_r(u_2)$ for some u_1, u_2. We know that with high probability it is not in any third $\partial B_r(u_3)$, otherwise there are 3 centers close together contradicting Lemma 4.1. Further, by Lemma 4.1, with high probability in G the neighborhood around every vertex of S is a tree. Thus, there are at most two vertices adjacent to v in $\cup_{u \in S} B_{r-1}(u)$ and its degree in $\tilde{G}$ is at least $d - 2$. In the other case, $v \in \partial B_r(u)$ for a unique vertex $u \in S$ and hence its degree in $\tilde{G}$ is $d - 1$. The bounds on the numbers of these vertices follow by Lemma 4.1 and applying the second moment method. The bound on the size of k then follows immediately from Lemma 4.1. $\qquad\square$

In what follows we will sometimes work in the conditional space of the configuration model where G is such that the high probability conclusions of Lemma 4.1 and Corollary 4.2 hold for $\tilde{G}$. Since the configuration model allows us to expose edges and maintain the uniform distribution over pairings of the unmatched pairs, under the conditioning, $\tilde{G}$ is a graph chosen according to the configuration model where the degrees of the vertices are modified appropriately, and we denote this set of graphs by $\hat{\mathcal{H}}(n, d)$. We use $\hat{\mathbb{P}}$ and $\hat{\mathbb{E}}$ to denote the corresponding conditional expectation and probabilities.

4.1 The first moment of the partition function of $\tilde{G}$

Let B be the subset of vertices defined in (4.1) and let $\sigma \in \{0, 1\}^B$. Define $Z_{G,\sigma}$ to be the partition function over independent sets of G whose restriction to B is σ, i.e.,

$$Z_{G,\sigma} := \sum_{\omega \in I(G): \omega(B) = \sigma} \lambda^{|\omega|}.$$

Similarly, define

$$Z_{\tilde{G},\sigma} := \sum_{\omega \in I(\tilde{G}):\omega(B)=\sigma} \lambda^{|\omega|}.$$

In this section, we will show that in expectation, for the boundary B as defined in (4.1) and any $\sigma \in \{0,1\}^B$, $Z_{\tilde{G},\sigma}$ is essentially proportional to a product measure on B. Let $m = |V(\tilde{G}) \setminus B|$. Define $Z_{\tilde{G},\alpha,\sigma}$ to be the partition function for independent sets of $\tilde{G}$ whose restriction to B is $\sigma \in \{0,1\}^B$ and for which α fraction of the vertices $V(\tilde{G}) \setminus B$ are in the independent set

$$Z_{\tilde{G},\alpha,\sigma} := \sum_{\omega \in I(\tilde{G}):\omega(B)=\sigma, \sum_{v \in V \setminus B} \omega_v = \alpha m} \lambda^{|\omega|}.$$

Fix an independent set ω of $\tilde{G}$ whose restriction to B is σ such that $\sum_{v \in V(\tilde{G}) \setminus B} \omega_v = \alpha m$. Let $L = |\sigma|$ and let L_i be the number of vertices in σ of degree $d - i$ for $i = 1, 2$. Let M_i denote the number of vertices of B of degree $d - i$ for $i = 1, 2$. We can calculate the expectation of the partition function as before using the exploration process in the configuration model. Let $N_1 = (d-1)L_1 + (d-2)L_2 + d\alpha m$ be the number of half-edges adjacent to a vertex in the independent set. Let $N_T = (d-1)M_1 + (d-2)M_2 + dm$ be the total number of half-edges overall. Calculating the probability that the pairing of the half edges does not pair vertices which are in the independent set, we have

$$\hat{\mathbb{E}}\left(Z_{\tilde{G},\alpha,\sigma}\right) = \lambda^{L_1+L_2+\alpha m}\binom{m}{\alpha m}\frac{\prod_{i=0}^{N_1-1}(N_T - N_1 - i)}{\prod_{i=0}^{N_1-1}(N_T - 1 - 2i)}. \tag{4.3}$$

In what follows, let $G \sim \mathcal{H}(n,d)$ and let $\tilde{G}$ be defined as above. Recall that α^* is given by the solution to (3.5). Let σ_0 denote the empty independent set on B.

Proposition 4.3. *Fix $\lambda > 0$.*

(1) For all $\sigma \in \{0,1\}^B$, and $0 \leq \alpha \leq \frac{1}{2}$,

$$\frac{\hat{\mathbb{E}}\left(Z_{\tilde{G},\alpha,\sigma}\right)}{\hat{\mathbb{E}}\left(Z_{\tilde{G},\alpha,\sigma_0}\right)} = \exp\left(O(n^{\frac{1}{5}})\right)\left(\lambda\left(\frac{1-2\alpha}{1-\alpha}\right)^{d-1}\right)^{|\sigma|}.$$

(2) Let α be such that $|\alpha - \alpha^*| < Cn^{-\frac{2}{5}}$ for a constant $C > 0$. Then,

$$\frac{\hat{\mathbb{E}}\left(Z_{\tilde{G},\alpha,\sigma}\right)}{\hat{\mathbb{E}}\left(Z_{\tilde{G},\alpha,\sigma_0}\right)} = \exp\left(O(n^{\frac{1}{5}})\right)\chi(\sigma)$$

where

$$\chi(\sigma) = \left(\lambda\left(\frac{1-2\alpha^*}{1-\alpha^*}\right)^{d-1}\right)^{|\sigma|}.$$

Proof. We compare the formula (4.3) for σ and σ_0. Let N'_1, N'_T be the corresponding quantities for σ_0 as defined before. Note that for $i = 1, 2$, $L'_i = 0$ and $N'_T = N_T$.

Comparing the numerators and denominators of the fraction in (4.3) we obtain that

$$
\frac{\hat{\mathbb{E}}(Z_{\tilde{G},\alpha,\sigma})}{\hat{\mathbb{E}}(Z_{\tilde{G},\alpha,\sigma_0})} = \lambda^{L_1+L_2} \frac{\displaystyle\prod_{i=0}^{N_1-N_1'-1}(N_T - 2N_1' - 2i)}{\displaystyle\prod_{i=0}^{N_1-N_1'-1}(N_T - N_1' - i)} = \lambda^{L_1+L_2} \left(\frac{1-2\alpha}{1-\alpha}\right)^{N_1-N_1'} \times
$$

$$
\times \frac{\displaystyle\prod_{i=0}^{N_1-N_1'-1}\left(1 + \frac{1}{(1-2\alpha)dm}\left((d-1)M_1 + (d-2)M_2 - 2i\right)\right)}{\displaystyle\prod_{i=0}^{N_1-N_1'-1}\left(1 + \frac{1}{(1-\alpha)dm}\left((d-1)M_1 + (d-2)M_2 - i\right)\right)}
$$

$$
= \lambda^{L_1+L_2}\left(\frac{1-2\alpha}{1-\alpha}\right)^{N_1-N_1'}\exp\left(O\left(\frac{(N_1 - N_1')n^{\frac{3}{5}}}{dm}\right)\right),
$$

where the last line follows since $M_1 = O(n^{\frac{3}{5}})$ and $M_2 = O(n^{\frac{1}{5}})$. Since $N_1 - N_1' = (d-1)L_1 + (d-2)L_2 = O(n^{\frac{3}{5}})$ and $m = n(1 - o_n(1))$, we obtain that

$$
\frac{\hat{\mathbb{E}}\left(Z_{\tilde{G},\alpha,\sigma}\right)}{\hat{\mathbb{E}}\left(Z_{\tilde{G},\alpha,\sigma_0}\right)} = \lambda^{L_1+L_2}\exp\left(O(n^{\frac{1}{5}})\right)\left(\frac{1-2\alpha}{1-\alpha}\right)^{(d-1)L_1+(d-2)L_2}
$$

$$
= \exp\left(O(n^{\frac{1}{5}})\right)\left(\lambda\left(\frac{1-2\alpha}{1-\alpha}\right)^{d-1}\right)^{|\sigma|}\left(\frac{1-2\alpha}{1-\alpha}\right)^{-L_2} = \exp\left(O(n^{\frac{1}{5}})\right)\left(\lambda\left(\frac{1-2\alpha}{1-\alpha}\right)^{d-1}\right)^{|\sigma|}.
$$

The last bound follows since $L_2 = O(n^{\frac{1}{5}})$ by the assumed conditioning, giving part (1) of the proposition. Finally, by the assumption that $|\alpha - \alpha^*| \leq Cn^{-\frac{2}{5}}$ the last expression above can be bounded by

$$
= \exp\left(O(n^{\frac{1}{5}})\right)\left(\lambda\left(\frac{1-2\alpha^*}{1-\alpha^*}\right)^{d-1}\right)^{|\sigma|},
$$

completing part (2) of the proposition. $\qquad\square$

Proposition 4.4. *For all $\sigma \in \{0,1\}^B$, and for large enough constant $C = C(\lambda, d)$,*

$$
\hat{\mathbb{E}}(Z_{\tilde{G},\sigma}) = (1 - o(1)) \sum_{\alpha:|\alpha-\alpha^*|\leq Cn^{-\frac{2}{5}}} \hat{\mathbb{E}}(Z_{\tilde{G},\alpha,\sigma}).
$$

Corollary 4.5. *For any $\sigma \in \{0,1\}^B$,*

$$
\hat{\mathbb{E}}(Z_{\tilde{G},\sigma}) = \exp\left(O(n^{\frac{1}{5}})\right)\chi(\sigma_B)\hat{\mathbb{E}}(Z_{\tilde{G},\sigma_0}).
$$

Proof. The claim follows by putting together part (2) of Proposition 4.3 and Proposition 4.4. In particular, taking C to be large enough as in Proposition 4.4,

$$
\hat{\mathbb{E}}(Z_{\tilde{G},\sigma}) = (1 - o_n(1)) \sum_{\alpha:|\alpha-\alpha^*|\leq Cn^{-\frac{2}{5}}} \hat{\mathbb{E}}(Z_{\tilde{G},\alpha,\sigma})
$$

$$
= \exp\left(O(n^{\frac{1}{5}})\right)\chi(\sigma) \sum_{\alpha:|\alpha-\alpha^*|\leq Cn^{-\frac{2}{5}}} \hat{\mathbb{E}}(Z_{\tilde{G},\sigma_0})
$$

$$
= \exp\left(O(n^{\frac{1}{5}})\right)\chi(\sigma)\hat{\mathbb{E}}(Z_{\tilde{G},\sigma_0}). \qquad\square
$$

To prove Proposition 4.4, we need a few intermediate results. Let $G_m \sim \mathcal{H}(m, d)$ where $m = |V(\tilde{G}) \setminus B|$, as defined above. Define the partition functions $Z_{G_m, \alpha}$ and $Z^{(2)}_{G_m, \alpha, \gamma, \varepsilon}$ as $Z_{G, \alpha}$ and $Z^{(2)}_{G, \alpha, \gamma, \varepsilon}$ were defined respectively, with $G = G_m$.

Lemma 4.6. *For any $0 \leq \alpha \leq \frac{1}{2}$,*

$$\frac{\hat{\mathbb{E}}(Z_{\tilde{G}, \alpha, \sigma_0})}{\hat{\mathbb{E}}(Z_{\tilde{G}, \alpha^*, \sigma_0})} = \frac{\mathbb{E}(Z_{G_m, \alpha})}{\mathbb{E}(Z_{G_m, \alpha^*})} \exp\left(O(n^{\frac{3}{5}}|\alpha - \alpha^*|)\right).$$

Proof. Let N_1 and N_T be defined as in (4.3). Define N_1^* and N_T^* analogously for α^*. Note that for the configuration σ_0, $L_i = 0$ for $i = 1, 2$ and $N_T = N_T^*$. Comparing the expressions in (4.3) and (3.3), we have that

$$\frac{\hat{\mathbb{E}}(Z_{\tilde{G}, \alpha, \sigma_0})}{\hat{\mathbb{E}}(Z_{\tilde{G}, \alpha^*, \sigma_0})} = \frac{\mathbb{E}(Z_{G_m, \alpha})}{\mathbb{E}(Z_{G_m, \alpha^*})} \times \frac{\displaystyle\prod_{i=0}^{d\alpha m - 1} \frac{dm - 2i - 1}{dm + (d-1)M_1 + (d-2)M_2 - 2i - 1}}{\displaystyle\prod_{i=0}^{d\alpha^* m - 1} \frac{dm - 2i - 1}{dm + (d-1)M_1 + (d-2)M_2 - 2i - 1}} \times$$

$$\times \frac{\displaystyle\prod_{i=0}^{d\alpha m - 1} \frac{(1-\alpha)dm + (d-1)M_1 + (d-2)M_2 - i}{(1-\alpha)dm - i}}{\displaystyle\prod_{i=0}^{d\alpha^* m - 1} \frac{(1-\alpha^*)dm + (d-1)M_1 + (d-2)M_2 - i}{(1-\alpha^*)dm - i}}$$

$$= \frac{\mathbb{E}(Z_{G_m, \alpha})}{\mathbb{E}(Z_{G_m, \alpha^*})} \exp\left(O(n^{\frac{3}{5}}|\alpha - \alpha^*|)\right)$$

where the last line follows by the bounds on m and M_i for $i = 1, 2$. $\qquad\square$

Lemma 4.7. *For $0 \leq \alpha \leq \frac{1}{2}$, there is a constant $C = C(\lambda, d) > 0$ such that*

$$\frac{\mathbb{E}(Z_{G_m, \alpha})}{\mathbb{E}(Z_{G_m, \alpha^*})} = \tilde{\Theta}(1) \exp\left(-Cn|\alpha - \alpha^*|^2\right).$$

Proof. Recall the expression (3.4) for $\Phi(\alpha)$. Writing the Taylor expansion for $\Phi(\alpha)$ around α^* and noting that $\Phi'(\alpha^*) = 0$, we have

$$\Phi(\alpha) - \Phi(\alpha^*) = \frac{\partial \Phi(\alpha^*)}{\partial \alpha}|\alpha - \alpha^*| + \frac{1}{2}\frac{\partial^2 \Phi(\alpha^*)}{\partial \alpha^2}|\alpha - \alpha^*|^2 + o\left(|\alpha - \alpha^*|^2\right)$$

$$= -\left(\frac{1}{\alpha^*(1 - \alpha^*)} + \frac{d}{(1 - \alpha^*)(1 - 2\alpha^*)}\right)|\alpha - \alpha^*|^2 + o\left(|\alpha - \alpha^*|^2\right)$$

and therefore

$$\frac{\mathbb{E}(Z_{G_m, \alpha})}{\mathbb{E}(Z_{G_m, \alpha^*})} = \tilde{\Theta}(1) \exp\left(m\left(\Phi(\alpha) - \Phi(\alpha^*)\right)\right) = \tilde{\Theta}(1) \exp\left(-Cn|\alpha - \alpha^*|^2\right). \qquad\square$$

Proof of Proposition 4.4. Combining part (1) of Proposition 4.3, Lemma 4.6 and Lemma 4.7, we have that for any $0 \leq \alpha \leq \frac{1}{2}$,

$$\frac{\hat{\mathbb{E}}(Z_{\tilde{G}, \alpha, \sigma})}{\hat{\mathbb{E}}(Z_{\tilde{G}, \alpha^*, \sigma_B})} = \frac{\hat{\mathbb{E}}(Z_{\tilde{G}, \alpha, \sigma})}{\hat{\mathbb{E}}(Z_{\tilde{G}, \alpha, \sigma_0})} \cdot \frac{\hat{\mathbb{E}}(Z_{\tilde{G}, \alpha, \sigma_0})}{\hat{\mathbb{E}}(Z_{\tilde{G}, \alpha^*, \sigma_0})} \cdot \frac{\hat{\mathbb{E}}(Z_{\tilde{G}, \alpha^*, \sigma_0})}{\hat{\mathbb{E}}(Z_{\tilde{G}, \alpha^*, \sigma_B})}$$

$$= \exp\left(O\left(n^{\frac{1}{5}} - Cn|\alpha - \alpha^*|^2 + n^{\frac{3}{5}}|\alpha - \alpha^*|\right)\right).$$

When $|\alpha - \alpha^*| > Cn^{-\frac{2}{5}}$, for a large enough constant $C = C(\lambda, d)$, the second term in the parenthesis above dominates and right hand side can be made arbitrarily small. Therefore

$$\hat{\mathbb{E}}(Z_{\tilde{G},\sigma}) = \sum_{\alpha:|\alpha-\alpha^*|>Cn^{-2/5}} \hat{\mathbb{E}}(Z_{\tilde{G},\alpha,\sigma}) + \sum_{\alpha:|\alpha-\alpha^*|\leq Cn^{-2/5}} \hat{\mathbb{E}}(Z_{\tilde{G},\alpha,\sigma})$$

$$= (1 - o_n(1)) \sum_{\alpha:|\alpha-\alpha^*|\leq Cn^{-2/5}} \hat{\mathbb{E}}(Z_{\tilde{G},\alpha,\sigma}). \qquad \square$$

4.2 The second moment of the partition function of $\tilde{G}$

As before, we divide the second moment $\mathbb{E}((Z_{\tilde{G},\alpha,\sigma})^2)$ into the contribution from pairs of independent sets σ' and σ'' of $\tilde{G}$ whose restriction to B is σ, $\sum_{v\in V\setminus B}\sigma'_v = \sum_{v\in V\setminus B}\sigma''_v = \alpha m$ and $|(\sigma'\cap\sigma'')\setminus B| = \gamma m$. We can further divide according the number εdm of half-edges which are paired from each of σ' and σ'' to $V(\tilde{G})\setminus(\sigma'\cup\sigma'')$. Denote this contribution by $Z^{(2)}_{\tilde{G},\alpha,\gamma,\varepsilon,\sigma}$. Thus, we can write

$$\mathbb{E}\left((Z_{\tilde{G},\alpha,\sigma})^2\right) = \sum_{\gamma,\varepsilon}\mathbb{E}\left(Z^{(2)}_{\tilde{G},\alpha,\gamma,\varepsilon,\sigma}\right).$$

As before, let L denote $|\sigma|$ with L_i denoting the numbers of vertices of B in the independent set of degrees $d-i$ for $i = 1, 2$. Define M_i as before and let $K_i = M_i - L_i$. Calculating the probability in the configuration model that a pair of subsets σ' and σ'' as above are independent sets we obtain

$$\hat{\mathbb{E}}\left(Z^{(2)}_{\tilde{G},\alpha,\gamma,\varepsilon,\sigma}\right) = \lambda^{2\alpha m+2(L_1+L_2)}\binom{m}{\alpha m}\binom{\alpha m}{\gamma m}\binom{(1-\alpha)m}{(\alpha-\gamma)m}\times$$

$$\times\left(\frac{\displaystyle\prod_{i=0}^{\gamma md+(d-1)L_1+(d-2)L_2-1}(1-2\alpha+\gamma)dm+(d-1)K_1+(d-2)K_2-i}{\displaystyle\prod_{i=0}^{\gamma md+(d-1)L_1+(d-2)L_2-1}dm+(d-1)M_1+(d-2)M_2-1-2i}\right)\times$$

$$\times\left(\frac{\displaystyle\prod_{i=0}^{\varepsilon dm-1}(1-2\alpha)dm+(d-1)K_1+(d-2)K_2-i\cdot\prod_{i=0}^{(\alpha-\gamma-\varepsilon)dm-1}(\alpha-\gamma)dm-i}{\displaystyle\prod_{i=0}^{(\alpha-\gamma)dm-1}(1-2\gamma)dm+(d-1)(M_1-2L_1)+(d-2)(M_2-2L_2)-1-2i}\times\right.$$

$$\left.\times\frac{\displaystyle\prod_{i=0}^{\varepsilon dm-1}(1-2\alpha-\varepsilon)dm+(d-1)K_1+(d-2)K_2-i}{\displaystyle\prod_{i=0}^{\varepsilon dm-1}(1-2\alpha)dm+(d-1)(M_1-2L_1)+(d-2)(M_2-2L_2)-1-2i}\right). \qquad (4.4)$$

In the main result of this section, we show in Proposition 4.15, that the second moment $\hat{\mathbb{E}}\left((Z_{G,\sigma})^2\right)$ is roughly the square of the first moment $\hat{\mathbb{E}}\left(Z_{G,\sigma}\right)$ by an analysis similar to that in [31, Theorem 6.11] and [36, Lemma 3.5]. We begin with a series of intermediate results that are needed for the proof.

Proposition 4.8. *Let* $(\alpha, \gamma, \varepsilon) \in \mathcal{R}$. *Then, for any* $\sigma \in \{0,1\}^B$,

$$\hat{\mathbb{E}}\left(Z^{(2)}_{\tilde{G},\alpha,\gamma,\varepsilon,\sigma}\right) = \exp\left(O\left(n^{\frac{1}{5}}\right)\right)\hat{\mathbb{E}}\left(Z^{(2)}_{\tilde{G},\alpha^*,\gamma^*,\varepsilon^*,\sigma}\right).$$

To prove Proposition 4.8, we need a few intermediate lemmas.

Lemma 4.9. *Let* $(\alpha, \gamma, \varepsilon) \in \mathcal{R}$. *Then, for some* $C = C(\lambda, d) > 0$

$$\frac{\hat{\mathbb{E}}(Z^{(2)}_{\tilde{G},\alpha,\gamma,\varepsilon,\sigma_0})}{\hat{\mathbb{E}}(Z^{(2)}_{\tilde{G},\alpha^*,\gamma^*,\varepsilon^*,\sigma_0})}$$
$$\leq \exp\left(-Cn\left(|\alpha-\alpha^*|^2 + |\gamma-\gamma^*|^2 + |\epsilon-\epsilon^*|^2\right) + O\left(n^{\frac{3}{5}}\left(|\alpha-\alpha^*| + |\gamma-\gamma^*| + |\varepsilon-\varepsilon^*|\right)\right)\right).$$

Proof. For the configuration σ_0, for $i = 1, 2$, $L_i = 0$ and $K_i = M_i$. Comparing the expressions (3.7) and (4.4), we obtain that for $G_m \sim \mathcal{H}(m, d)$,

$$\frac{\hat{\mathbb{E}}(Z^{(2)}_{\tilde{G},\alpha,\gamma,\varepsilon,\sigma_0})}{\hat{\mathbb{E}}(Z^{(2)}_{\tilde{G},\alpha^*,\gamma^*,\varepsilon^*,\sigma_0})} = \frac{\mathbb{E}(Z^{(2)}_{G_m,\alpha,\gamma,\varepsilon})}{\mathbb{E}(Z^{(2)}_{G_m,\alpha^*,\gamma^*,\varepsilon^*})} \times$$

$$\times \frac{\displaystyle\prod_{i=0}^{\gamma md-1} \frac{(1-2\alpha+\gamma)dm + (d-1)M_1 + (d-2)M_2 - i}{(1-2\alpha+\gamma)dm - i}}{\displaystyle\prod_{i=0}^{\gamma^* md-1} \frac{(1-2\alpha^*+\gamma^*)dm + (d-1)M_1 + (d-2)M_2 - i}{(1-2\alpha^*+\gamma^*)dm - i}} \times$$

$$\times \frac{\displaystyle\prod_{i=0}^{\gamma^* md-1} \frac{dm + (d-1)M_1 + (d-2)M_2 - 1 - 2i}{dm - 1 - 2i}}{\displaystyle\prod_{i=0}^{\gamma md-1} \frac{dm + (d-1)M_1 + (d-2)M_2 - 1 - 2i}{dm - 1 - 2i}} \times$$

$$\times \frac{\displaystyle\prod_{i=0}^{\varepsilon md-1} \frac{(1-2\alpha)dm + (d-1)M_1 + (d-2)M_2 - i}{(1-2\alpha+\gamma)dm - i}}{\displaystyle\prod_{i=0}^{\varepsilon^* md-1} \frac{(1-2\alpha^*)dm + (d-1)M_1 + (d-2)M_2 - i}{(1-2\alpha^*)dm - i}} \times$$

$$\times \frac{\displaystyle\prod_{i=0}^{(\alpha^*-\gamma^*)md-1} \frac{(1-2\gamma^*)dm + (d-1)M_1 + (d-2)M_2 - 1 - 2i}{(1-2\gamma^*)dm - 1 - 2i}}{\displaystyle\prod_{i=0}^{(\alpha-\gamma)md-1} \frac{(1-2\gamma)dm + (d-1)M_1 + (d-2)M_2 - 1 - 2i}{(1-2\gamma)dm - 1 - 2i}} \times$$

$$\times \frac{\displaystyle\prod_{i=0}^{\varepsilon md-1} \frac{(1-2\alpha-\varepsilon)dm + (d-1)M_1 + (d-2)M_2 - i}{(1-2\alpha-\varepsilon)dm - i}}{\displaystyle\prod_{i=0}^{\varepsilon^* md-1} \frac{(1-2\alpha^*-\varepsilon^*)dm + (d-1)M_1 + (d-2)M_2 - i}{(1-2\alpha^*-\varepsilon^*)dm - i}} \times$$

$$\times \frac{\displaystyle\prod_{i=0}^{\varepsilon^* md-1} \frac{(1-2\alpha^*)dm + (d-1)M_1 + (d-2)M_2 - 1 - 2i}{(1-2\alpha^*)dm - 1 - 2i}}{\displaystyle\prod_{i=0}^{\varepsilon md-1} \frac{(1-2\alpha)dm + (d-1)M_1 + (d-2)M_2 - 1 - 2i}{(1-2\alpha)dm - 1 - 2i}}$$

$$= \exp\left(m\left(f(\alpha,\gamma,\varepsilon) - f(\alpha^*,\gamma^*,\varepsilon^*)\right) + O(\log(m))\right) \times$$
$$\times \exp\left(O\left(n^{\frac{3}{5}}\left(|\alpha-\alpha^*| + |\gamma-\gamma^*| + |\varepsilon-\varepsilon^*|\right)\right)\right).$$

The final equality follows by Proposition 3.1 and the bounds on the sizes of m, M_1 and M_2 by the assumed conditioning. By Proposition 3.2, we have that for some constant C, possibly depending on d and λ,

$$f(\alpha, \gamma, \varepsilon) - f(\alpha^*, \gamma^*, \varepsilon^*) \leq -C\left(|\alpha - \alpha^*|^2 + |\gamma - \gamma^*|^2 + |\varepsilon - \varepsilon^*|^2\right).$$

Therefore, we obtain that

$$\frac{\hat{\mathbb{E}}(Z^{(2)}_{\check{G},\alpha,\gamma,\varepsilon,\sigma_0})}{\hat{\mathbb{E}}(Z^{(2)}_{\check{G},\alpha^*,\gamma^*,\varepsilon^*,\sigma_0})}$$
$$\leq \exp\left(-Cn\left(|\alpha - \alpha^*|^2 + |\gamma - \gamma^*|^2 + |\epsilon - \epsilon^*|^2\right) + O\left(n^{\frac{3}{5}}\left(|\alpha - \alpha^*| + |\gamma - \gamma^*| + |\varepsilon - \varepsilon^*|\right)\right)\right). \ \square$$

Lemma 4.10. Let $(\alpha, \gamma, \varepsilon) \in \mathcal{R}$. For any $\sigma \in \{0,1\}^B$,

$$\frac{\hat{\mathbb{E}}(Z^{(2)}_{\check{G},\alpha,\gamma,\varepsilon,\sigma})}{\hat{\mathbb{E}}(Z^{(2)}_{\check{G},\alpha,\gamma,\varepsilon,\sigma_0})} = (\chi(\sigma))^2 \exp\left(O\left(|\alpha - \alpha^*| + |\gamma - \gamma^*| + |\varepsilon - \varepsilon^*|\right)n^{\frac{3}{5}} + n^{\frac{1}{5}}\right)\right).$$

Proof. Using the expression (4.4) for each of $\mathbb{E}(Z^{(2)}_{\check{G},\alpha,\gamma,\varepsilon,\sigma_0})$ and $\mathbb{E}(Z^{(2)}_{\check{G},\alpha,\gamma,\varepsilon,\sigma})$ and taking ratios of the numerators and denominators separately for each of the products, we have

$$\frac{\hat{\mathbb{E}}(Z^{(2)}_{\check{G},\alpha,\gamma,\varepsilon,\sigma})}{\hat{\mathbb{E}}(Z^{(2)}_{\check{G},\alpha,\gamma,\varepsilon,\sigma_0})} = \lambda^{2|\sigma|} \exp\left(O(n^{\frac{1}{5}})\right) \times$$

$$\times \left(\frac{(1-2\alpha)^2}{(1-2\alpha+\gamma)(1-2\gamma)^{1/2}} \frac{(1-2\alpha-\varepsilon)}{(1-2\alpha)} \frac{(1-2\gamma)^{1/2}}{(1-2\alpha)^{1/2}} \times\right.$$
$$\left. \times \frac{(1-2\alpha-2\varepsilon)}{(1-2\alpha-\varepsilon)} \frac{(1-2\alpha)^{1/2}}{(1-2\alpha-2\varepsilon)^{1/2}}\right)^{(d-1)L_1+(d-2)L_2}$$

$$= \lambda^{2|\sigma|} \exp\left(O(n^{\frac{1}{5}})\right) \cdot \left(\frac{(1-2\alpha)(1-2\alpha-2\varepsilon)^{1/2}}{(1-2\alpha+\gamma)}\right)^{(d-1)L_1+(d-2)L_2}$$

$$= \lambda^{2|\sigma|} \exp\left(O\left(n^{\frac{1}{5}} + n^{\frac{3}{5}}\left(|\alpha - \alpha^*| - |\gamma - \gamma^*| + |\varepsilon - \varepsilon^*|\right)\right)\right) \times$$

$$\times \left(\frac{(1-2\alpha^*)(1-2\alpha^*-2\varepsilon^*)^{1/2}}{(1-2\alpha^*+\gamma^*)}\right)^{(d-1)L_1+(d-2)L_2}.$$

Since $\gamma^* = (\alpha^*)^2$, $\varepsilon^* = \alpha^*(1-2\alpha^*)$ and $|\sigma| = L_1 + L_2$, the last line gives that

$$\frac{\hat{\mathbb{E}}(Z^{(2)}_{\check{G},\alpha,\gamma,\varepsilon,\sigma})}{\hat{\mathbb{E}}(Z^{(2)}_{\check{G},\alpha,\gamma,\varepsilon,\sigma_0})} = \left(\lambda\left(\frac{1-2\alpha^*}{1-\alpha^*}\right)^{d-1}\right)^{2|\sigma|} \exp\left(O\left(|\alpha - \alpha^*| + |\gamma - \gamma^*| + |\varepsilon - \varepsilon^*|\right)n^{\frac{3}{5}} + n^{\frac{1}{5}}\right)$$

and the lemma follows. $\square$

Putting these results together we now prove Proposition 4.8.

Proof of Proposition 4.8. Let $\sigma \in \{0,1\}^B$ and write

$$\frac{\hat{\mathbb{E}}\left(Z^{(2)}_{\check{G},\alpha,\gamma,\varepsilon,\sigma}\right)}{\hat{\mathbb{E}}\left(Z^{(2)}_{\check{G},\alpha^*,\gamma^*,\varepsilon^*,\sigma}\right)} = \frac{\hat{\mathbb{E}}\left(Z^{(2)}_{\check{G},\alpha,\gamma,\varepsilon,\sigma}\right)}{\hat{\mathbb{E}}\left(Z^{(2)}_{\check{G},\alpha,\gamma,\varepsilon,\sigma_0}\right)} \frac{\hat{\mathbb{E}}\left(Z^{(2)}_{\check{G},\alpha,\gamma,\varepsilon,\sigma_0}\right)}{\hat{\mathbb{E}}\left(Z^{(2)}_{\check{G},\alpha^*,\gamma^*,\varepsilon^*,\sigma_0}\right)} \frac{\hat{\mathbb{E}}\left(Z^{(2)}_{\check{G},\alpha^*,\gamma^*,\varepsilon^*,\sigma_0}\right)}{\hat{\mathbb{E}}\left(Z^{(2)}_{\check{G},\alpha^*,\gamma^*,\varepsilon^*,\sigma}\right)}.$$

Applying Lemmas 4.9 and 4.10 to the terms in the product above, we have that

$$\frac{\hat{\mathbb{E}}\left(Z^{(2)}_{\tilde{G},\alpha,\gamma,\varepsilon,\sigma}\right)}{\hat{\mathbb{E}}\left(Z^{(2)}_{\tilde{G},\alpha^*,\gamma^*,\varepsilon^*,\sigma}\right)} \leq \exp\left(O\left(n^{\frac{3}{5}}\left(|\alpha-\alpha^*|+|\gamma-\gamma^*|+|\varepsilon-\varepsilon^*|\right)+n^{\frac{1}{5}}\right)\right.$$
$$\left. - Cn\left(|\alpha-\alpha^*|^2+|\gamma-\gamma^*|^2+|\varepsilon-\varepsilon^*|^2\right)\right). \qquad (4.5)$$

For a constant C, define the set

$$\mathcal{R}_C = \left\{(\alpha,\gamma,\varepsilon)\in\mathcal{R} \text{ s.t. } |\alpha-\alpha^*|,|\gamma-\gamma^*|,|\varepsilon-\varepsilon^*|\leq Cn^{-\frac{2}{5}}\right\}.$$

Note that for some sufficiently large C, if $(\alpha,\gamma,\varepsilon)\in\mathcal{R}_C$, the the right-hand side of (4.5) can be bounded by $\exp\left(O(n^{\frac{1}{5}})\right)$. On the other hand, if $(\alpha,\gamma,\varepsilon)\notin\mathcal{R}_C$ for any constant C, then the right-hand side of (4.5) can be made arbitraily small and therefore, for any $(\alpha,\gamma,\varepsilon)\in\mathcal{R}$ and σ

$$\hat{\mathbb{E}}\left(Z^{(2)}_{\tilde{G},\alpha,\gamma,\varepsilon,\sigma}\right) = \exp\left(O(n^{\frac{1}{5}})\right)\hat{\mathbb{E}}\left(Z^{(2)}_{\tilde{G},\alpha^*,\gamma^*,\varepsilon^*,\sigma}\right). \qquad \square$$

Proposition 4.11. *For any $\sigma\in\{0,1\}^B$,*

$$\hat{\mathbb{E}}\left(Z^2_{\tilde{G},\sigma}\right) = \exp\left(O(n^{\frac{1}{5}})\right)(\chi(\sigma))^2\,\hat{\mathbb{E}}\left(Z^2_{\tilde{G},\sigma_0}\right).$$

Proof. Applying the Cauchy-Schwarz inequality, Proposition 4.8 and Lemma 4.10, we have

$$\hat{\mathbb{E}}\left(Z^2_{\tilde{G},\sigma}\right) = \hat{\mathbb{E}}\left(\left(\sum_\alpha Z_{\tilde{G},\alpha,\sigma}\right)^2\right) = \tilde{\Theta}(1)\sum_\alpha\hat{\mathbb{E}}\left(Z^2_{\tilde{G},\alpha,\sigma}\right) = \tilde{\Theta}(1)\sum_{\alpha,\gamma,\varepsilon}\hat{\mathbb{E}}\left(Z^{(2)}_{\tilde{G},\alpha,\gamma,\varepsilon,\sigma}\right) =$$

$$= \exp\left(O(n^{\frac{1}{5}})\right)\hat{\mathbb{E}}\left(Z^{(2)}_{\tilde{G},\alpha^*,\gamma^*,\varepsilon^*,\sigma}\right) = \exp\left(O(n^{\frac{1}{5}})\right)(\chi(\sigma))^2\,\hat{\mathbb{E}}\left(Z^{(2)}_{\tilde{G},\alpha^*,\gamma^*,\varepsilon^*,\sigma_0}\right) =$$

$$= \exp\left(O(n^{\frac{1}{5}})\right)(\chi(\sigma))^2\sum_{\alpha,\gamma,\varepsilon}\hat{\mathbb{E}}\left(Z^{(2)}_{\tilde{G},\alpha,\gamma,\varepsilon,\sigma_0}\right) =$$

$$= \exp\left(O(n^{\frac{1}{5}})\right)(\chi(\sigma))^2\sum_\alpha\hat{\mathbb{E}}\left(Z^2_{\tilde{G},\alpha,\sigma_0}\right) =$$

$$= \exp\left(O(n^{\frac{1}{5}})\right)(\chi(\sigma))^2\,\hat{\mathbb{E}}\left(\left(\sum_\alpha Z_{\tilde{G},\alpha,\sigma_0}\right)^2\right) =$$

$$= \exp\left(O(n^{\frac{1}{5}})\right)(\chi(\sigma))^2\,\hat{\mathbb{E}}\left(Z^2_{\tilde{G},\sigma_0}\right). \qquad \square$$

The next step is to show a bound on the second moment of $Z_{\tilde{G},\sigma_0}$ by the square of the first moment, and we begin with the following intermediate result.

Lemma 4.12.

$$\hat{\mathbb{E}}\left(Z^{(2)}_{\tilde{G},\alpha^*,\gamma^*,\varepsilon^*,\sigma_0}\right) = \exp\left(O(n^{\frac{1}{5}})\right)\left(\hat{\mathbb{E}}(Z_{\tilde{G},\alpha^*,\sigma_0})\right)^2.$$

Proof. As before, we note that for the configuration σ_0, for $i=1,2$, L_i-0 and $K_i=M_i$.

Comparing the expressions (3.7) and (4.4) we obtain that

$$\frac{\hat{\mathbb{E}}\left(Z^{(2)}_{\tilde{G},\alpha^*,\gamma^*,\varepsilon^*,\sigma_0}\right)}{\mathbb{E}\left(Z^{(2)}_{G_m,\alpha^*,\gamma^*,\varepsilon^*}\right)} = \exp\left(O(n^{\frac{1}{5}})\right) \times$$

$$\times \left(\frac{(1-2\alpha^*+\gamma^*)(1-2\gamma^*)^{\frac{1}{2}}}{(1-2\alpha^*)} \frac{(1-2\alpha^*)(1-2\alpha^*)^{\frac{1}{2}}}{(1-2\alpha^*-\varepsilon^*)(1-2\gamma^*)^{\frac{1}{2}}} \times \right.$$

$$\left. \times \frac{(1-2\alpha^*)(1-2\alpha^*-\varepsilon^*)^{\frac{1}{2}}}{(1-2\alpha^*-2\varepsilon^*)(1-2\alpha^*)^{\frac{1}{2}}} \right)^{(d-1)K_1+(d-2)K_2}$$

$$= \exp\left(O(n^{\frac{1}{5}})\right) \left(\frac{(1-\alpha^*)^2}{1-2\alpha^*} \right)^{(d-1)K_1+(d-2)K_2} \quad (4.6)$$

where the last equality follows by canceling terms and using the fact that $\gamma^* = (\alpha^*)^2$ and $\varepsilon^* = \alpha^*(1-2\alpha^*)$. Similarly, comparing (3.3) and (4.3), we obtain that

$$\frac{\hat{\mathbb{E}}(Z_{\tilde{G},\alpha^*,\sigma_0})}{\mathbb{E}(Z_{G_m,\alpha^*})} = \exp\left(O(n^{\frac{1}{5}})\right) \left(\frac{1-\alpha^*}{(1-2\alpha^*)^{\frac{1}{2}}} \right)^{(d-1)K_1+(d-2)K_2}. \quad (4.7)$$

Combining Proposition 3.1, (3.3), and Lemma 6.2, we have that

$$\frac{\hat{\mathbb{E}}\left(Z^{(2)}_{G_m,\alpha^*,\gamma^*,\varepsilon^*}\right)}{\left(\hat{\mathbb{E}}(Z_{G_m,\alpha^*})\right)^2} = \exp\left(O(n^{\frac{1}{5}})\right) \frac{\exp\left(mf(\alpha^*,\gamma^*,\varepsilon^*)+O(\log m)\right)}{\exp\left(2n\Phi(\alpha^*)\right)} = \exp\left(O(n^{\frac{1}{5}})\right). \quad (4.8)$$

Finally, putting together (4.6), (4.7) and (4.8) proves the lemma. $\qquad\square$

Proposition 4.13.

$$\hat{\mathbb{E}}\left(Z^2_{\tilde{G},\sigma_0}\right) = \exp\left(O(n^{\frac{1}{5}})\right) \left(\hat{\mathbb{E}}\left(Z_{\tilde{G},\sigma_0}\right)\right)^2.$$

Proof. Applying the Cauchy-Schwartz inequality, we have

$$\hat{\mathbb{E}}\left(Z^2_{\tilde{G},\sigma_0}\right) = \hat{\mathbb{E}}\left(\left(\sum_\alpha Z_{\tilde{G},\alpha,\sigma_0}\right)^2\right) \le \tilde{\Theta}(1) \sum_\alpha \hat{\mathbb{E}}\left(Z^2_{\tilde{G},\alpha,\sigma_0}\right) = \tilde{\Theta}(1) \sum_{\alpha,\gamma,\varepsilon} \hat{\mathbb{E}}\left(Z^{(2)}_{\tilde{G},\alpha,\gamma,\varepsilon,\sigma_0}\right) =$$

$$= \exp\left(O(n^{\frac{1}{5}})\right) \hat{\mathbb{E}}\left(Z^{(2)}_{\tilde{G},\alpha^*,\gamma^*,\varepsilon^*,\sigma_0}\right) = \exp\left(O(n^{\frac{1}{5}})\right) \left(\hat{\mathbb{E}}\left(Z_{\tilde{G},\alpha^*,\sigma_0}\right)\right)^2 =$$

$$= \exp\left(O(n^{\frac{1}{5}})\right) \left(\hat{\mathbb{E}}\left(Z_{\tilde{G},\sigma_0}\right)\right)^2$$

where the third equality is by Proposition 4.8 and the fourth equality is by Lemma 4.12. $\qquad\square$

Define $Z_{G,\alpha,\sigma}$ to be the partition function over independent sets of G whose restriction to B is σ and for which α fraction of the vertices in $V(\tilde{G}) \setminus B$ are in the independent set. That is,

$$Z_{G,\alpha,\sigma} := \sum_{\omega \in I(G):\omega(B)=\sigma,\sum_{v \in V(\tilde{G}) \setminus B} \omega_v = \alpha m} \lambda^{|\omega|}.$$

Lemma 4.14. *For any $\sigma \in \{0,1\}^B$ and $0 \le \alpha < \frac{1}{2}$, the partition functions for G and $\tilde{G}$ can be related by*

$$Z_{G,\sigma} = \kappa(\sigma)Z_{\tilde{G},\sigma}$$

and

$$Z_{G,\alpha,\sigma} = \kappa(\sigma)Z_{\tilde{G},\alpha,\sigma}$$

where $\kappa(\sigma)$ is a constant depending only on the configuration σ and has a product structure

$$\kappa(\sigma) = \prod_{i=1}^{k+1} \kappa_i(\sigma(W_i)).$$

and $\kappa_i = \kappa_j$ for $1 \le i, j \le k$.

Proof. By the Markov property. □

Putting these results together, we obtain the following.

Proposition 4.15. *For any $\sigma \in \{0,1\}^B$,*

$$\hat{\mathbb{E}}((Z_{G,\sigma})^2) = \exp(O(n^{\frac{1}{5}}))(\hat{\mathbb{E}}(Z_{G,\sigma}))^2.$$

Proof.

$$
\begin{aligned}
\hat{\mathbb{E}}((Z_{G,\sigma})^2) &= \hat{\mathbb{E}}((Z_{\tilde{G},\sigma})^2(\kappa(\sigma))^2) && \text{(by Lemma 4.14)}\\
&= \exp\left(O(n^{\frac{1}{5}})\right)\hat{\mathbb{E}}((Z_{\tilde{G},\sigma_0})^2(\chi(\sigma)\kappa(\sigma))^2) && \text{(by Proposition 4.11)}\\
&= \exp\left(O(n^{\frac{1}{5}})\right)(\hat{\mathbb{E}}(Z_{\tilde{G},\sigma_0}\chi(\sigma)\kappa(\sigma)))^2 && \text{(by Proposition 4.13)}\\
&= \exp\left(O(n^{\frac{1}{5}})\right)(\hat{\mathbb{E}}(Z_{G,\sigma}))^2. && \text{(by Corollary 4.5 and Lemma 4.14)}
\end{aligned}
$$

□

5 Local weak convergence to the free measure on the tree

The first result in this section shows that there does not exist a "bad set" of neighborhoods with large stationary probability where the partition function is much larger than the expected partition function. In the sequel, we will show that if on a constant fraction of the $n^{\frac{3}{5}}$ random neighborhoods, the local measure is far from the free measure on the tree, then such a bad set exists with high probability.

Proposition 5.1. *Let $c > 0$ and suppose that $G \sim \hat{\mathcal{H}}(n,d)$. The probability that there exists a set of independent set configurations $\mathcal{B} \subset \{0,1\}^B$ such that*

$$\mathcal{E}_1(\mathcal{B}) := \left\{ \sum_{\sigma \in \mathcal{B}} Z_{G,\sigma} > \exp\left(cn^{\frac{3}{5}}\right) \sum_{\sigma \in \mathcal{B}} \hat{\mathbb{E}}\left(Z_{G,\sigma}\right) \right\} \tag{5.1}$$

and

$$\mathcal{E}_2(\mathcal{B}) := \left\{ \sum_{\sigma \in \mathcal{B}} Z_{G,\sigma} > \exp\left(-n^{\frac{4}{7}}\right) \hat{\mathbb{E}}\left(Z_G\right) \right\} \tag{5.2}$$

is at most $\exp\left(-n^{\frac{3}{5}}(c - o_n(1))\right)$.

The choice of $\frac{4}{7}$ could be replaced with any constant less than $\frac{3}{5}$ and greater than $\frac{1}{2}$.

Proof. In what follows, let B be a fixed set vertices, rather than a random set as defined in (4.1). Since the claim holds for each fixed B, it holds for random B. Suppose that there is a set $\mathcal{B}$ of boundary configurations such that $\mathcal{E}_1(\mathcal{B})$ and $\mathcal{E}_2(\mathcal{B})$ hold. Define the set of configurations

$$\mathcal{D} := \left\{ \sigma \in \mathcal{B} : Z_{G,\sigma} > \frac{1}{3}\exp\left(cn^{\frac{3}{5}}\right)\hat{\mathbb{E}}\left(Z_{G,\sigma}\right) \right\}.$$

Suppose it was the case that

$$\sum_{\sigma \in \mathcal{B} \setminus \mathcal{D}} Z_{G,\sigma} > \frac{1}{2} \sum_{\sigma \in \mathcal{B}} Z_{G,\sigma}.$$

Then,

$$\frac{1}{3} \exp\left(cn^{\frac{3}{5}}\right) \sum_{\sigma \in \mathcal{B} \setminus \mathcal{D}} \hat{\mathbb{E}}(Z_{G,\sigma}) > \sum_{\sigma \in \mathcal{B} \setminus \mathcal{D}} Z_{G,\sigma} > \frac{1}{2} \sum_{\sigma \in \mathcal{B}} Z_{G,\sigma}.$$

This contradicts (5.1) (that $\mathcal{E}_1(\mathcal{B})$ holds) and thus we may assume that

$$\sum_{\sigma \in \mathcal{D}} Z_{G,\sigma} > \frac{1}{2} \sum_{\sigma \in \mathcal{B}} Z_{G,\sigma}.$$

Therefore, by (5.2) (that $\mathcal{E}_2(\mathcal{B})$ holds) we have

$$\sum_{\sigma \in \mathcal{D}} Z_{G,\sigma} > \frac{1}{2} \sum_{\sigma \in \mathcal{B}} Z_{G,\sigma} > \frac{1}{2} \exp\left(-n^{\frac{4}{7}}\right) \sum_{\sigma \in \{0,1\}^B} \hat{\mathbb{E}}\left(Z_{G,\sigma}\right). \qquad (5.3)$$

By (5.3), and Markov's inequality,

$$\hat{\mathbb{P}}\left(\exists \mathcal{B} \text{ s.t. } \mathcal{E}_1(\mathcal{B}) \cap \mathcal{E}_2(\mathcal{B})\right) \leq \hat{\mathbb{P}}\left(\sum_{\sigma \in \mathcal{D}} Z_{G,\sigma} > \frac{1}{2} \exp\left(-n^{\frac{4}{7}}\right) \sum_{\sigma \in \{0,1\}^B} \hat{\mathbb{E}}\left(Z_\sigma\right)\right) \leq$$

$$\leq \frac{2\hat{\mathbb{E}}\left(\sum_{\sigma \in \mathcal{D}} Z_{G,\sigma}\right)}{\exp\left(-n^{\frac{4}{7}}\right) \sum_{\sigma \in \{0,1\}^B} \hat{\mathbb{E}}\left(Z_\sigma\right)}. \qquad (5.4)$$

By the definition of $\mathcal{D}$ and Proposition 4.15 we have

$$\hat{\mathbb{E}}\left(\sum_{\sigma \in \mathcal{D}} Z_{G,\sigma}\right) < 3 \exp\left(-cn^{\frac{3}{5}}\right) \hat{\mathbb{E}}\left(\sum_{\sigma \in \mathcal{D}} Z_{G,\sigma} \frac{Z_{G,\sigma}}{\hat{\mathbb{E}}\left(Z_{G,\sigma}\right)}\right) \leq$$

$$\leq 3 \exp\left(-cn^{\frac{3}{5}}\right) \hat{\mathbb{E}}\left(\sum_{\sigma \in \{0,1\}^B} Z_{G,\sigma} \frac{Z_{G,\sigma}}{\hat{\mathbb{E}}\left(Z_{G,\sigma}\right)}\right) \leq$$

$$\leq \exp\left(-cn^{\frac{3}{5}}\right) \exp\left(O(n^{\frac{1}{5}})\right) \sum_{\sigma \in \{0,1\}^B} \hat{\mathbb{E}}\left(Z_{G,\sigma}\right). \qquad (5.5)$$

Putting together (5.4) and (5.5), we obtain that

$$\hat{\mathbb{P}}\left(\exists \mathcal{B} \text{ s.t. } \mathcal{E}_1(\mathcal{B}) \cap \mathcal{E}_2(\mathcal{B})\right) \leq \exp\left(-n^{\frac{3}{5}}(c - o(1))\right),$$

which completes the proof of Proposition 5.1. $\qquad \square$

Recall the definition of the set of vertices S' from (4.2) and recall that for each $s_i \in S'$, W_i is the set of vertices on the boundary $\partial B_r(s_i)$ and $|S'| = k$. Fix $1 \leq i \leq |S'|$. Let $T_1, \ldots, T_{|W_i|}$ be $(d-1)$-ary trees and let $\tilde{T}$ be their union. Let $\mathbb{P}_{\tilde{T}}$ be the product of the free measures on the trees. Let us identify the roots $u^{(i)} = \{u_1, \ldots, u_{|W_i|}\}$ of these trees with the vertices of W_i. Let T be the tree obtained by joining to $\tilde{T}$ a d-ary tree of depth r whose leaves are identified with the u_i.

Lemma 5.2. *Define the distribution ν on $\{0,1\}^B$ by*

$$\nu(A) := \frac{\displaystyle\sum_{\sigma \in A} \chi(\sigma)\kappa(\sigma)}{\displaystyle\sum_{\sigma \in \{0,1\}^B} \chi(\sigma)\kappa(\sigma)}.$$

If $\sigma \sim \nu$, then for each $1 \leq i \leq k$, the $\sigma(W_i)$ are independent and

$$\nu(\sigma(W_i) \in \cdot) = \mathbb{P}_T(\sigma(u^{(i)}) \in \cdot) = \mathbb{P}_{T_d}(B_r(x) \in \cdot). \tag{5.6}$$

Proof. By relating the occupation probability of the root for the free measure for the d-regular tree and the occupation probability of the root of the free measure on the $(d-1)$-ary tree, it can be verified that

$$\mathbb{P}_{\tilde{T}}(\sigma(W_i) = \omega) \propto \chi(\omega).$$

By the Markov property,

$$\mathbb{P}_T(\sigma(W_i) = \omega) \propto \chi(\omega)\kappa(\omega).$$

Therefore

$$\nu(\sigma(W_i) = \omega) = \mathbb{P}_T(\sigma(u^{(i)}) = \omega) \tag{5.7}$$

and (5.6) follows. $\qquad\square$

Let $\mathbb{P}_{G_n}$ denote the hardcore measure on the random d-regular graph of size n. Recall that we have local weak convergence to the free measure if for all all r, with high probability over a uniformly chosen vertex u, for $\varepsilon > 0$, as $n \to \infty$,

$$\mathbb{P}\left(d_{tv}(\mathbb{P}_{G_n}(\sigma(B_r(u)) \in \cdot), \mathbb{P}_{T_d}(\sigma(B_r(x)) \in \cdot) > \varepsilon)\right) \to 0. \tag{5.8}$$

The following lemma will be used in the next result.

Lemma 5.3. *Let I be an index set, $(X_i)_{i \in I}$ random variables and $(a_i)_{i \in I}$ constants. Suppose that for each i,*

$$\mathbb{P}(X_i < a_i) \leq \varepsilon$$

for some $\varepsilon > 0$. Then,

$$\mathbb{P}\left(\sum_{i \in I} X_i < \frac{1}{2}\sum_{i \in I} a_i\right) \leq 2\varepsilon.$$

Proof. We show first that

$$\sum_i X_i < \frac{1}{2}\sum_i a_i \quad \Rightarrow \quad \sum_i a_i \mathbb{1}_{X_i > a_i} < \frac{1}{2}\sum_i a_i \quad \Leftrightarrow \quad \sum_i a_i \mathbb{1}_{X_i \leq a_i} \geq \frac{1}{2}\sum_i a_i.$$

The equivalence above is immediate and the first implication can be seen as follows:

$$\frac{1}{2}\sum_i a_i > \sum_i X_i \geq \sum_i X_i \mathbb{1}_{X_i > a_i} \geq \sum_i a_i \mathbb{1}_{X_i > a_i}.$$

Applying Markov's inequality, we have

$$\mathbb{P}\left(\sum_i X_i < \frac{1}{2}\sum_i a_i\right) \leq \mathbb{P}\left(\sum_i a_i \mathbb{1}_{X_i \leq a_i} \geq \frac{1}{2}\sum_i a_i\right) \leq 2\varepsilon. \qquad\square$$

We will show the following result which in turn implies (5.8), since with high probability $k = (1 - o_n(1))n^{3/5}$.

Theorem 5.4. *Let $G \sim \mathcal{H}(n, d)$. Let ω be an independent set drawn according to the hardcore measure on G. For any $\varepsilon > 0$,*

$$\frac{|\{1 \le i \le k \ : \ d_{tv}(\mathbb{P}_{G_n}(\omega(W_i)) \in \cdot), \mathbb{P}_{T_d}(B_r(x) \in \cdot)) > \varepsilon\}|}{k} \xrightarrow{\mathbb{P}} 0, \tag{5.9}$$

as $n \to \infty$.

Proof. Let $\mathcal{E}$ be the event that the left hand side of (5.9) is at least $\delta > 0$. If $\mathcal{E}$ occurs, then by the definition of total variation distance, there exists a set J of indices of size δk and $A_i \in \{0, 1\}^{W_i}$ for $i \in J$ such that

$$\mathbb{P}_G\left(\omega(W_i) \in A_i\right) - \mathbb{P}_T(\omega(u^{(i)}) \in A_i) > \varepsilon \quad \forall i \in J.$$

This implies

$$\mathbb{E}\left(\sum_{i \in J} \mathbb{1}(\omega(W_i) \in A_i)\right) - \sum_{i \in J} \mathbb{P}_T(\omega(u^{(i)}) \in A_i) \ge \varepsilon \delta k. \tag{5.10}$$

Using the fact that $\sum_{i \in J} \mathbb{1}(\omega(W_i) \in A_i) \le \delta k$, we obtain that

$$\mathbb{E}\left(\sum_{i \in J} \mathbb{1}(\omega(W_i) \in A_i)\right) - \sum_{i \in J} \mathbb{P}_T(\omega(u^{(i)}) \in A_i) \le$$
$$\le \frac{\varepsilon \delta n^{\frac{3}{5}}}{2} + \delta k \mathbb{P}\left(\sum_{i \in J} \mathbb{1}(\omega(W_i) \in A_i) > \frac{\varepsilon \delta n^{\frac{3}{5}}}{2} + \sum_{i \in J} \mathbb{P}_T(\omega(u^{(i)}) \in A_i)\right). \tag{5.11}$$

Combining (5.10) and (5.11) and using the fact that $k \ge (1 - o(1))n^{\frac{3}{5}}$, we get

$$\mathbb{P}\left(\sum_{i \in J} \mathbb{1}(\omega(W_i) \in A_i) > \frac{\varepsilon \delta n^{\frac{3}{5}}}{2} + \sum_{i \in J} \mathbb{P}_T(\omega(u^{(i)}) \in A_i)\right) \ge \frac{\varepsilon}{3}. \tag{5.12}$$

Define the set of configurations

$$\mathcal{B} := \left\{\sigma \in \{0, 1\}^{\cup_i W_i} \ \text{s.t.} \ \sum_{i \in J} \mathbb{1}(\sigma \in A_i) > \frac{\varepsilon \delta n^{\frac{3}{5}}}{2} + \sum_{i \in J} \mathbb{P}_T(\sigma \in A_i)\right\}.$$

By (5.12), on the event $\mathcal{E}$

$$\sum_{\sigma \in \mathcal{B}} Z_{G,\sigma} \ge \frac{\varepsilon}{3} Z_G. \tag{5.13}$$

In particular, (5.13) holds when $G \sim \hat{\mathcal{H}}(n, d)$. By Proposition 4.15, for any $\sigma \in \{0, 1\}^B$ and $G \sim \hat{\mathcal{H}}(n, d)$,

$$\frac{\hat{\mathbb{E}}(Z_{G,\sigma}^2)}{(\hat{\mathbb{E}}(Z_{G,\sigma}))^2} \le \exp\left(O(n^{\frac{1}{5}})\right),$$

so that by the Paley-Zygmund inequality,

$$\hat{\mathbb{P}}\left(Z_{G,\sigma} > \frac{1}{2}\hat{\mathbb{E}}(Z_{G,\sigma})\right) \ge \exp\left(-O(n^{\frac{1}{5}})\right).$$

Indeed, by Markov's inequality, it follows that for any $\kappa > 0$

$$\hat{\mathbb{P}}\left(Z_{G,\sigma} \in \left[\frac{\hat{\mathbb{E}}(Z_{G,\sigma})}{2}, e^{n^\kappa}\hat{\mathbb{E}}(Z_{G,\sigma})\right]\right) \geq \exp\left(-O(n^{\frac{1}{5}})\right). \tag{5.14}$$

On the other hand, by Azuma's inequality, for all $\kappa > 0$,

$$\hat{\mathbb{P}}\left(|\log Z_{G,\sigma} - \hat{\mathbb{E}}(\log Z_{G,\sigma})| \geq n^{\frac{1}{2}+\kappa}\right) \leq \exp\left(-\frac{n^{1+2\kappa}}{2nd\lambda}\right) \leq \exp(-n^{2\kappa}). \tag{5.15}$$

Together, (5.14) and (5.15) imply that the interval

$$[\log(\hat{\mathbb{E}}(Z_{G,\sigma}) - \log 2, n^\kappa + \log\hat{\mathbb{E}}(Z_{G,\sigma})] \cap [\hat{\mathbb{E}}(\log Z_{G,\sigma}) - n^{1/2+\kappa}, \hat{\mathbb{E}}(\log Z_{G,\sigma}) + n^{1/2+\kappa}]$$

is non-empty and hence

$$|\hat{\mathbb{E}}(\log Z_{G,\sigma}) - \log(\hat{\mathbb{E}}(Z_{G,\sigma}))| \leq 2n^{\frac{1}{2}+\kappa}.$$

Plugging this into (5.15), we obtain that for each fixed σ,

$$\hat{\mathbb{P}}\left(Z_{G,\sigma} \geq \exp\left(-O(n^{\frac{1}{2}+\kappa})\right)\hat{\mathbb{E}}(Z_{G,\sigma})\right) \geq 1 - \exp(-n^{2\kappa}).$$

By Lemma 5.3,

$$\hat{\mathbb{P}}\left(\sum_{\sigma \in \mathcal{B}} Z_{G,\sigma} \geq \exp\left(-O(n^{\frac{1}{2}+\kappa})\right)\sum_{\sigma \in \mathcal{B}}\hat{\mathbb{E}}(Z_{G,\sigma})\right) \geq 1 - 2\exp(-n^{2\kappa})$$

and combining with (5.13), we have that

$$\hat{\mathbb{P}}\left(\sum_{\sigma \in \mathcal{B}} Z_{G,\sigma} \geq \exp\left(-O(n^{\frac{1}{2}+\kappa})\right)\hat{\mathbb{E}}(Z_G)\right) \geq 1 - 2\exp(-n^{2\kappa}).$$

That is, we have shown that the event $\mathcal{E}_2(\mathcal{B})$, as defined in Proposition 5.1 holds with high probability on the event $\mathcal{E}$. By Corollary 4.5 and Lemma 4.14 for each $\sigma \in \mathcal{B}$, we have

$$\hat{\mathbb{E}}(Z_{G,\sigma}) = \exp\left(O(n^{\frac{1}{5}})\right)\chi(\sigma)\kappa(\sigma)\hat{\mathbb{E}}(Z_{G,\sigma_0}).$$

Summing this over all possible σ, we have

$$\hat{\mathbb{E}}(Z_G) = \sum_{\sigma \in \{0,1\}^B} \exp\left(O(n^{\frac{1}{5}})\right)\chi(\sigma)\kappa(\sigma)\hat{\mathbb{E}}(Z_{G,\sigma_0}),$$

and comparing these two equalities, we obtain that

$$\sum_{\sigma \in \mathcal{B}}\hat{\mathbb{E}}(Z_{G,\sigma}) = \exp\left(O(n^{\frac{1}{5}})\right)\frac{\displaystyle\sum_{\sigma \in \mathcal{B}}\chi(\sigma)\kappa(\sigma)}{\displaystyle\sum_{\sigma \in \{0,1\}^B}\chi(\sigma)\kappa(\sigma)}\hat{\mathbb{E}}(Z_G) = \exp\left(O(n^{\frac{1}{5}})\right)\nu(\mathcal{B})\hat{\mathbb{E}}(Z_G). \tag{5.16}$$

By Lemma 5.2 and Azuma-Hoeffding,

$$\nu(\mathcal{B}) = \nu\left(\sum_{i \in J}\mathbb{1}(\sigma \in A_i) > \frac{\varepsilon\delta n^{\frac{3}{5}}}{2} + \sum_{i \in J}\mathbb{P}_T(\sigma \in A_i)\right)$$

$$= \nu\left(\sum_{i \in J}\mathbb{1}(\sigma \in A_i) - \nu\left(\sum_{i \in J}\mathbb{1}(\sigma \in A_i)\right) > \frac{\varepsilon\delta n^{\frac{3}{5}}}{2}\right) \leq \exp\left(-cn^{\frac{3}{5}}\right).$$

Combining the above bound with (5.16), we obtain that

$$\sum_{\sigma \in \mathcal{B}} \hat{\mathbb{E}}(Z_{G,\sigma}) \le \exp\left(-cn^{\frac{3}{5}}\right) \hat{\mathbb{E}}(Z_G). \tag{5.17}$$

Using (5.17), we have that if $\mathcal{E}$ and $\mathcal{E}_2(\mathcal{B})$ hold, then

$$\sum_{\sigma \in \mathcal{B}} Z_{G,\sigma} > \exp\left(-O(n^{\frac{1}{2}+\kappa})\right) \hat{\mathbb{E}}(Z_G) \ge \exp\left(cn^{\frac{3}{5}}\right) \hat{\mathbb{E}}\left(\sum_{\sigma \in \mathcal{B}} Z_{G,\sigma}\right),$$

and therefore, the event $\mathcal{E}_1(\mathcal{B})$ as defined in Proposition 5.1 holds. Therefore, on the event $\{\mathcal{E} \cap \mathcal{E}_2(\mathcal{B})\}$, $\mathcal{E}_1(\mathcal{B})$ holds with probability at least $1 - 2\exp(-n^{2\kappa})$.

To summarise, we have shown that on the event $\mathcal{E}$, the event $\{\mathbb{E}_1(\mathcal{B}) \cap \mathcal{E}_2(\mathcal{B})\}$ holds with high probability. Applying Proposition 5.1,

$$\exp(-cn^{\frac{3}{5}}) \ge \hat{\mathbb{P}}(\mathcal{E}_1(\mathcal{B}) \cap \mathcal{E}_2(\mathcal{B})) \ge \hat{\mathbb{P}}(\mathcal{E}_1(\mathcal{B}) \cap \mathcal{E}_2(\mathcal{B})|\mathcal{E})\hat{\mathbb{P}}(\mathcal{E}) \ge (1 - o(1))\hat{\mathbb{P}}(\mathcal{E}).$$

Thus, $\hat{\mathbb{P}}(\mathcal{E}) \to 0$ as $n \to \infty$. Since $G \sim \hat{\mathcal{H}}(n,d)$ with high probability, $\mathbb{P}(\mathcal{E}) \to 0$ as $n \to \infty$ and the claim follows. $\qquad \square$

Theorem 5.4 implies (5.8), which establishes local weak convergence to the free measure, proving Theorem 1. Given local weak convergence, the equivalence of the reconstruction thresholds and hence Theorem 3, follow.

6 Technical lemmas about the partition function

In this section we show that the second moment $\mathbb{E}(Z_G^2)$ is close to the the square of the first moment $\mathbb{E}(Z_G)$ and satisfies a quadratic decay property. Let

$$H_1(x,y) = -x(\log(x) - \log(y)) + (x-y)(\log(y-x) - \log(y)).$$

Recall that

$$\mathbb{E}\left(Z_{G,\alpha,\gamma,\varepsilon}^{(2)}\right) = \lambda^{2\alpha n}\binom{n}{\alpha n}\binom{\alpha n}{\gamma n}\binom{(1-\alpha)n}{(\alpha-\gamma)n}\left(\frac{\prod_{i=0}^{\gamma nd-1}(1-2\alpha+\gamma)dn-i}{\prod_{i=0}^{\gamma nd-1}dn-1-2i}\right) \times$$

$$\times \left(\frac{\prod_{i=0}^{\varepsilon dn-1}(1-2\alpha)dn-i \cdot \prod_{i=0}^{(\alpha-\gamma-\varepsilon)dn-1}(\alpha-\gamma)dn-i}{\prod_{i=0}^{(\alpha-\gamma)dn-1}(1-2\gamma)dn-1-2i} \cdot \frac{\prod_{i=0}^{\varepsilon dn-1}(1-2\alpha-\varepsilon)dn-i}{\prod_{i=0}^{\varepsilon dn-1}(1-2\alpha)dn-1-2i}\right). \tag{6.1}$$

We also recall that the following function arises naturally in the estimation of the second moment:

$$f(\alpha,\gamma,\varepsilon) = 2\alpha \log(\lambda) + H(\alpha) + H_1(\gamma,\alpha) + H_1(\alpha-\gamma,1-\alpha) + d\Psi_2(\alpha,\gamma,\varepsilon),$$

where

$$\Psi_2(\alpha,\gamma,\varepsilon) = H_1(\varepsilon,\alpha-\gamma) + \int_0^\gamma \log(1-2\alpha+\gamma-x)dx - \int_0^\gamma \log(1-2x)dx$$

$$+ \int_0^\varepsilon \log(1-2\alpha-x)dx + \int_0^{\alpha-\gamma-\varepsilon} \log(\alpha-\gamma-x)dx - \int_0^{\alpha-\gamma} \log(1-2\gamma-2x)dx$$

$$+ \int_0^\varepsilon \log(1-2\alpha-\varepsilon-x)dx - \int_0^\varepsilon \log(1-2\alpha-2x)dx. \tag{6.2}$$

Using (3.2) to compare terms in (6.1) and (6.2) proves Proposition 3.1 showing that

$$\mathbb{E}\left(Z_{\alpha,\gamma,\varepsilon}^{(2)}\right) = \exp(nf(\alpha,\gamma,\varepsilon) + O(\log n)).$$

Thus the second moment depends on the behavior of the function f. We will show a series of technical lemmas showing that f attains its maximum at $(\alpha^*, \gamma^*, \varepsilon^*)$ and decays quadratically around this point. Define

$$\alpha_c := \frac{(2 - \delta_d)\log(d)}{d},$$

where $\delta_d = C\frac{\log\log d + 1}{\log d} \to 0$ as $d \to \infty$ and $C > 3$.

Lemma 6.1. *For each fixed α, γ in the region $\mathcal{R}$, the function f has a local maximum at*

$$\overline{\varepsilon} = \overline{\varepsilon}(\alpha,\gamma) = \frac{1}{2}\left(1 - 2\gamma - \sqrt{(1 - 2\alpha)^2 + 4(\alpha - \gamma)^2}\right).$$

Proof. Differentiating (6.2), we have that the derivative of f is given by

$$\frac{\partial f}{\partial \varepsilon} = d\log\left(\frac{(\alpha - \gamma - \varepsilon)(1 - 2\alpha - \varepsilon)(1 - 2\alpha - 2\varepsilon)^2}{\varepsilon^2(1 - 2\alpha - \varepsilon)(1 - 2\alpha - 2\varepsilon)}\right).$$

Since the hardcore model is a permissive model, we may assume that the local maxima of f are in the interior of $\mathcal{R}$ (see e.g. [13, Proposition 3.2]). Solving for ε by setting $\frac{\partial f}{\partial \varepsilon} = 0$, gives that the unique solution in the interior of $\mathcal{R}$ is $\varepsilon = \overline{\varepsilon}$. Further, we check that the second derivative

$$\frac{\partial^2 f}{\partial \varepsilon^2} = -d\left(\frac{1}{\alpha - \gamma - \varepsilon} + \frac{2}{1 - 2\alpha - 2\varepsilon} + \frac{2}{\varepsilon}\right) < 0,$$

and hence $\overline{\varepsilon}$ is a maximum. $\qquad\square$

We will also use the following technical lemma.

Lemma 6.2. *For any $0 \le \alpha \le \frac{1}{2}$,*

$$2\Phi(\alpha) = f(\alpha, \hat{\gamma}, \hat{\varepsilon}).$$

Proof. Substituting $\hat{\gamma} = \alpha^2$ and $\hat{\varepsilon} = \alpha(1 - 2\alpha)$ in the expression for $f(\alpha, \gamma, \varepsilon)$, and simplifying, we have

$$2\Phi(\alpha) - f(\alpha, \hat{\gamma}, \hat{\varepsilon}) = 2d\left((1 - \alpha)\log(1 - \alpha) - \frac{1 - 2\alpha}{2}\log(1 - 2\alpha) - 2\right) - d\Psi_2(\alpha, \hat{\gamma}, \hat{\varepsilon})$$

$$+ H(\alpha) + H_1(\alpha^2, \alpha) + H_1(\alpha(1 - \alpha), 1 - \alpha) = 0. \qquad\square$$

Define the function

$$g(\alpha, \gamma) \equiv f(\alpha, \gamma, \overline{\varepsilon}(\alpha, \gamma))$$

and consider its extremal values for fixed α in the region where $\alpha, \gamma \ge 0$ and $\alpha - \gamma \ge 0$. The derivative of g is given by

$$\frac{\partial g}{\partial \gamma}(\alpha, \gamma) = \frac{\partial f}{\partial \gamma}(\alpha, \gamma, \overline{\varepsilon}(\alpha, \gamma)), \tag{6.3}$$

while its second derivative is

$$\frac{\partial^2 g}{\partial \gamma^2}(\cdot, \cdot) = \frac{\partial f}{\partial^2 \gamma}(\cdot, \overline{\varepsilon}) + \frac{\partial \overline{\varepsilon}}{\partial \gamma}\frac{\partial f}{\partial \gamma \partial \varepsilon}(\cdot, \cdot, \overline{\varepsilon}). \tag{6.4}$$

We establish the behavior of the function f near its maximum by showing several facts about g.

Proof of Proposition 3.2. By Lemma 6.1, for fixed α and γ, f has a maximum at $\bar{\varepsilon}(\alpha, \gamma)$. Thus, it remains to find the maximum of $g(\alpha, \gamma)$. In what follows we will show that for fixed α, g is maximized at (α, α^2). Noting that $\bar{\varepsilon}(\alpha, \hat{\gamma}) = \alpha(1 - 2\alpha)$, it follows that f is maximized at $(\alpha^*, \gamma^*, \varepsilon^*)$ since by Lemma 6.2, $f(\alpha, \alpha^2, \alpha(1 - 2\alpha)) = 2\Phi(\alpha)$, and by (3.5), Φ is maximized at α^*.

We can verify that (α, α^2) is a stationary point of g by computing the first derivative. In what follows, we establish that for fixed $\alpha < \alpha_c$, (α, α^2) is a global maximizer of $g(\alpha, \gamma)$ by considering several possible ranges for γ. This will consist of two main steps. The first is to show that (α, α^2) is a maximum and the second is to show that the function g is larger at (α, α^2) than at any other possible maximum. Let $\gamma_i := c_i \frac{\alpha}{\log(\alpha^{-1})}$ for $i = 1, 2, 3$ with the constants c_i to be set later.

Computing the derivatives using equations (6.3) and (6.4) gives that

$$\frac{\partial^2 g}{\partial \gamma^2} = -\left(\frac{2}{\alpha - \gamma} + \frac{1}{\gamma} + \frac{1}{1 - 2\alpha + \gamma} \right)$$
$$+ d \left(\frac{1}{1 - 2\alpha + \gamma} + \frac{2}{\alpha - \gamma} - \frac{2(\alpha - \gamma)}{(\alpha - \gamma - \bar{\varepsilon})\sqrt{(1 - 2\alpha)^2 + 4(\alpha - \gamma)^2}} \right). \tag{6.5}$$
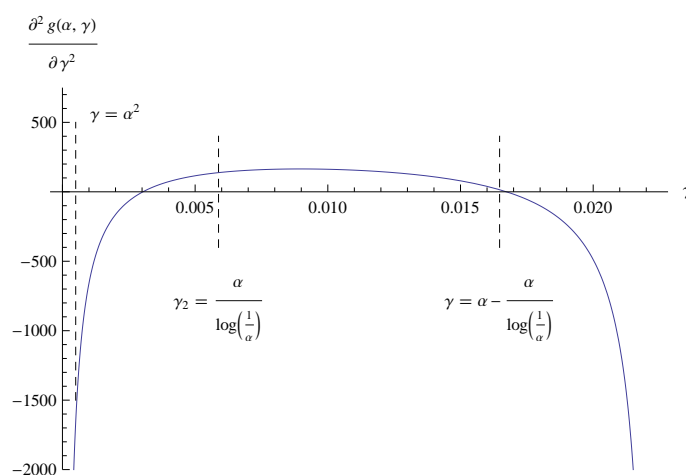

Figure 3: Second derivative of $g(\alpha, \gamma)$ with respect to γ.

1) Let d be sufficiently large so that by the assumption that $\alpha < \alpha_c$, we have that $d < 3\alpha^{-1} \log(\alpha^{-1})$. Let d also be large enough so that for a small constant c_1 to be chosen later, $\alpha^2 < c_1 \alpha / \log(\alpha^{-1})$. We show that there is a constant c_1 so that for $\gamma \in [0, \gamma_1]$, $\frac{\partial^2 g}{\partial \gamma^2} < 0$ (see Figure 3) and hence the stationary point (α, α^2) of g is a maximum (see Figures 4 and 5).

Note that the first term of the second derivative (6.5) is negative and has magnitude at least $1/\gamma \geq \alpha^{-1} \log(\alpha^{-1})/c_1$ for this range of γ. The term on the second line of (6.5) is positive, but we will argue that its magnitude is $O(\alpha^{-1} \log(\alpha^{-1}))$. The claim will follow by taking the constant c_1 to be small enough.

The terms in the second bracket of (6.5) can be bounded by $O(1)$ as can be seen below from the series expansion in Mathematica (note that this calculation does not
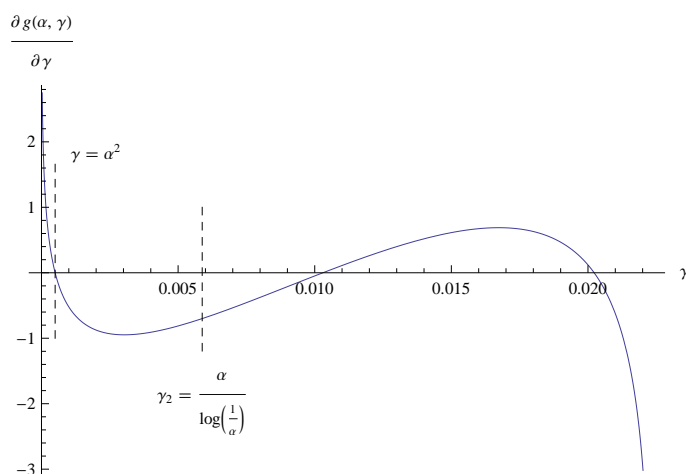
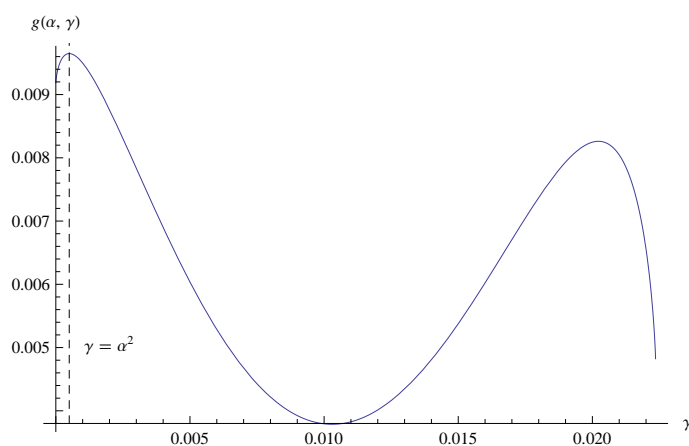Figure 4: First derivative of $g(\alpha, \gamma)$ with respect to γ.

Figure 5: The function $g(\alpha, \gamma)$.

depend on the size of γ or on c_1).

$$\frac{1}{1 - 2\alpha + \gamma} + \frac{2}{\alpha - \gamma} - \frac{2(\alpha - \gamma)}{\sqrt{(1 - 2\alpha)^2 + 4(\alpha - \gamma)^2}\left(\alpha - \gamma - \frac{\left(1 - 2\gamma - \sqrt{(1-2\alpha)^2 + 4(\alpha - \gamma)^2}\right)}{2}\right)}$$
$$= \left(1 + 4\alpha + O(\alpha^2)\right) + \left(-3 - 12\alpha + O(\alpha^2)\right)\gamma + O(\gamma^2). \tag{6.6}$$

Therefore

$$\frac{\partial^2 g}{\partial \gamma^2} < -\frac{1}{\gamma} + Cd < -\frac{1}{c_1}\alpha^{-1}\log(\alpha^{-1}) + C\alpha^{-1}\log(\alpha^{-1}),$$

where above the particular value of C may change in each appearance. Hence by choosing c_1 to be sufficiently small, the claim follows.

We now divide the analysis showing that $\gamma = \alpha^2$ corresponds to a global maximum into two cases based on the size of α.

2) When α is small enough, we show that g has no stationary point for $\gamma \in (\alpha^2, \alpha]$. Suppose that $\alpha < \varepsilon \log(d)/d$ for ε sufficiently small. Then, for some ε', $d < \varepsilon'\alpha^{-1}\log(\alpha^{-1})$. Expanding the terms in the second bracket of (6.5) as in (6.6), and recalling that

$\partial g(\alpha, \alpha^2)/\partial \gamma = 0$, we have that for $\gamma > \alpha^2$,

$$\frac{\partial g}{\partial \gamma}(\alpha, \gamma) < \int_{\alpha^2}^{\gamma} (-\frac{1}{s} + Cd)ds < \log\left(\frac{\alpha^2}{\gamma}\right) + C\varepsilon'\alpha^{-1}\log(\alpha^{-1})(\gamma - \alpha^2). \tag{6.7}$$

We claim that the bound on the right hand side of (6.7) is maximized at the end points of the interval $[\alpha^2, \alpha]$. Indeed, differentiating the bound with respect to γ, the only stationary point in the interval is at $\gamma = \alpha/(C\varepsilon'\log(\alpha^{-1}))$. Furthermore, the second derivative of the bound is positive so that the stationary point can only be a minimum.

At $\gamma = \alpha^2$, the bound evaluates to 0. Evaluating (6.7) at $\gamma = \alpha$, we obtain that

$$\frac{\partial g}{\partial \gamma}(\alpha, \alpha) < \log(\alpha) + C\varepsilon'\alpha^{-1}\log(\alpha^{-1})(\alpha - \alpha^2) \leq \log(\alpha)(1 - C\varepsilon') < 0$$

since ε' can be made arbitrarily small by our choice of ε. It follows that $\partial g(\alpha, \gamma)/\partial \gamma < 0$ in $(\alpha^2, \alpha]$ and the interval does not contain any stationary points.

When $\alpha \geq \varepsilon \log(d)/d$, there may be a second stationary point near (α, α) and the values of g at the stationary points must be compared to show that (α, α^2) is the global maximum. This will be done in the points that follow.

3) We show that there is a constant $c_2 > c_1$ so that for $\gamma \in (\alpha^2, \gamma_2]$, $\frac{\partial g}{\partial \gamma} < 0$. Thus, for this range of γ there are no stationary points of g (see Figure 4). Integrating the second derivative in this range we obtain that

$$\frac{\partial g}{\partial \gamma}(\alpha, \gamma_2) = \int_{\alpha^2}^{\gamma_1} \frac{\partial^2 g(\alpha, s)}{\partial s^2} ds + \int_{\gamma_1}^{\gamma_2} \frac{\partial^2 g(\alpha, s)}{\partial s^2} ds \leq -\int_{\alpha^2}^{\gamma_1} \frac{1}{2s} ds + \int_{\gamma_1}^{\gamma_2} Cd\,ds$$

$$= \frac{1}{2}\log\left(\frac{\alpha^2}{\gamma_1}\right) + Cd(\gamma_2 - \gamma_1) \leq \frac{1}{2}\log\left(\frac{\alpha\log(\alpha^{-1})}{c_1}\right) + C(c_2 - c_1).$$

The upper bounds in the first line above follow by the arguments similar to those of 1) and 2) above. In particular, c_1 can be chosen small enough so that the first bound follows. The last inequality follows since $\alpha < \alpha_c$ which implies $d < 3\alpha^{-1}\log(\alpha^{-1})$. As $d \to \infty$, $\alpha \to 0$ and therefore for large enough d, the first derivative will be negative as claimed.

4) There are constants c_2, c_3 such that $c_2 > c_1$ and for $\gamma \in [\gamma_2, \alpha - \gamma_3]$, $\frac{\partial^2 g}{\partial \gamma^2} > 0$ (see Figure 3). This implies that g does not have a maximum in this range.

For this range of γ, the first term of $\frac{\partial^2 g}{\partial \gamma^2}$ in (6.5) can be bounded as

$$-\left(\frac{2}{\alpha - \gamma} + \frac{1}{\gamma} + \frac{1}{1 - 2\alpha + \gamma}\right) \geq -\left(\frac{2}{c_3} + \frac{1}{c_2}\right)\alpha^{-1}\log(\alpha^{-1}). \tag{6.8}$$

The second term can be bounded below as follows. We use the Taylor series expansion

$$\sqrt{x + y} = \sqrt{x} + \sum_{i=1}^{\infty} x^{\frac{1}{2} - i} y^i \left(\prod_{k=1}^{i}\left(\frac{3}{2} - k\right)\frac{1}{i!}\right)$$

to expand as follows with $x = (1 - 2\alpha)^2$ and $y = 4(\alpha - \gamma)^2$:

$$\bar{\varepsilon} = (\alpha - \gamma) - \frac{1}{2}\sum_{i=1}^{\infty}(1 - 2\alpha)^{1-2i}(2(\alpha - \gamma))^{2i}\left(\prod_{k=1}^{i}\left(\frac{3}{2} - k\right)\frac{1}{i!}\right).$$

Rearranging, we have

$$\alpha - \gamma - \bar{\varepsilon} = \frac{1}{2} \sum_{i=1}^{\infty} \frac{(2(\alpha - \gamma))^{2i}}{(1 - 2\alpha)^{2i-1}} \left(\prod_{k=1}^{i} \left(\frac{3}{2} - k \right) \frac{1}{i!} \right).$$

Similarly,

$$\sqrt{(1 - 2\alpha)^2 + 4(\alpha - \gamma)^2} = \sum_{i=0}^{\infty} \frac{(2(\alpha - \gamma))^{2i}}{(1 - 2\alpha)^{2i-1}} \left(\prod_{k=1}^{i} \left(\frac{3}{2} - k \right) \frac{1}{i!} \right).$$

Therefore, the second term of (6.5) can be bounded as follows

$$d \left(\frac{1}{1 - 2\alpha + \gamma} + \frac{2}{\alpha - \gamma} - \frac{2(\alpha - \gamma)}{(\alpha - \gamma - \bar{\varepsilon})\sqrt{(1 - 2\alpha)^2 + 4(\alpha - \gamma)^2}} \right)$$

$$= d \left(1 + \frac{2}{\alpha - \gamma} \left(1 - \frac{1}{\left(\frac{1}{1-2\alpha} - \frac{(\alpha-\gamma)^2}{(1-2\alpha)^3} + \cdots \right) \left((1 - 2\alpha) + \frac{2(\alpha-\gamma)^2}{1-2\alpha} - \frac{2(\alpha-\gamma)^4}{(1-2\alpha)^3} + \cdots \right)} \right) \right)$$

$$= d \left(1 + \frac{2(\alpha - \gamma)}{(1 - 2\alpha)^2} \frac{1 - 4\frac{(\alpha-\gamma)^2}{(1-2\alpha)^2} \cdots}{1 + \frac{(\alpha-\gamma)^2}{(1-2\alpha)^2} - 4\frac{(\alpha-\gamma)^4}{(1-2\alpha)^4} \cdots} \right)$$

$$\geq d \left(1 + \frac{2(\alpha - \gamma)}{(1 - 2\alpha)^2} \left(1 - 5\frac{(\alpha - \gamma)^2}{(1 - 2\alpha)^2} \right) \right) \geq d. \tag{6.9}$$

Combining (6.8) and (6.9) we see that to prove the claim it is enough to choose constants c_2 and c_3 to satisfy $\left(\frac{2}{c_3} + \frac{1}{c_2} \right) \alpha^{-1} \log(\alpha^{-1}) < d$. Recall that we are now in the case that $d > \varepsilon' \alpha^{-1} \log(\alpha^{-1})$. Thus, the claim follows by choosing $c_2 > c_1$ and c_2 and c_3 large enough.

5) Lastly, we show that for $\gamma \in [\alpha - \gamma_3, \alpha]$, the maximum obtained has a smaller value than the maximum at $\gamma = \alpha^2$ (see Figure 5). By (6.5) we have

$$\frac{\partial g^2}{\partial \gamma^2} \geq -\frac{2}{\alpha - \gamma} \left(1 + \frac{\alpha - \gamma}{\gamma} \right) \geq \frac{-2 \left(1 + \frac{c_3}{\log(\alpha^{-1}) - c_3} \right)}{\alpha - \gamma}.$$

Suppose that γ_4 is any maximum of g in this interval so that $\partial g(\alpha, \gamma)/\partial \gamma = 0$. We can bound the first derivative as follows.

$$\frac{\partial g(\alpha, \gamma)}{\partial \gamma} = \int_{\gamma_4}^{\gamma} \frac{\partial^2 g(\alpha, s)}{\partial s^2} ds \geq -2 \left(1 + \frac{c_3}{\log(\alpha^{-1}) - c_3} \right) \int_{\gamma_4}^{\gamma} \frac{1}{\alpha - s} ds$$

$$= -2 \left(1 + \frac{c_3}{\log(\alpha^{-1}) - c_3} \right) \log \left(\frac{\alpha - \gamma_4}{\alpha - \gamma} \right).$$

Now, using the above bound, we integrate to obtain

$$g(\alpha, \alpha) - g(\alpha, \gamma_4) = \int_{\gamma_4}^{\alpha} \frac{\partial g(\alpha, s)}{\partial s} ds \geq -2 \left(1 + \frac{c_3}{\log(\alpha^{-1}) - c_3} \right) (\alpha - \gamma_4)$$

$$> -2(1 + \varepsilon_d)(\alpha - \gamma_4)$$

$$\Rightarrow \quad g(\alpha, \gamma_4) < g(\alpha, \alpha) + 2(1 + \varepsilon_d)(\alpha - \gamma_4),$$

where $\varepsilon_d = 1/(\log d - \log \log d)$. We would like to show that $g(\alpha, \alpha^2) > g(\alpha, \gamma_4)$. Thus it suffices to show that $g(\alpha, \alpha^2) - g(\alpha, \alpha) > 2(1 + \varepsilon_d)(\alpha - \gamma_4)$. Roughly, $g(\alpha, \alpha)$ should

behave like Φ, the first moment, while $g(\alpha, \alpha^2)$ is 2Φ. Recall that $g(\alpha, \alpha) = f(\alpha, \alpha, 0)$ since $\bar{\varepsilon}(\alpha, \alpha) = 0$. Comparing (3.3) and (3.7), we have

$$\exp\left(nf(\alpha, \alpha, 0)\right) = \lambda^{\alpha n} e^{n\Phi(\alpha)}$$
$$\Rightarrow g(\alpha, \alpha) = f(\alpha, \alpha, 0) = \alpha \log(\lambda) + \Phi(\alpha).$$

Also,

$$\Phi(\alpha) = \alpha \log(\lambda) + \frac{1}{n} \log(|\{\sigma : |\sigma| = \alpha n\}|).$$

Therefore,

$$g(\alpha, \alpha^2) - g(\alpha, \alpha) = 2\Phi(\alpha) - \alpha \log(\lambda) - \Phi(\alpha)$$
$$= \frac{1}{n} \log(|\{\sigma : |\sigma| = \alpha n\}|)$$
$$= H(\alpha) + d\left((1-\alpha)\log(1-\alpha) + (\alpha - \frac{1}{2})\log(1-2\alpha)\right). \quad (6.10)$$

Using Mathematica to expand the terms of the expression in (6.10), we obtain for $\alpha < \alpha_c$,

$$H(\alpha) + d\left((1-\alpha)\log(1-\alpha) + (\alpha - \frac{1}{2})\log(1-2\alpha)\right)$$
$$= (1 - \log(\alpha))\alpha - \frac{d+1}{2}\alpha^2 - \left(\frac{d}{2} + \frac{1}{6}\right)\alpha^3 + O(\alpha^4)$$
$$> 2(1 + \varepsilon_d)\alpha > 2(1 + \varepsilon_d)(\alpha - \gamma_4),$$

for d sufficiently large.

The first part of Proposition 3.2 follows by the five facts above. Next we show that $f(\alpha, \gamma, \varepsilon)$ decays quadratically near $\alpha^*, \gamma^*, \varepsilon^*$. We start by calculating the Hessian matrix for f. The partial second mixed derivatives of $f(\alpha, \gamma, \varepsilon)$ evaluated at $\hat{\gamma} = \alpha^2$ and $\hat{\varepsilon} = \alpha(1 - 2\alpha)$ are given by:

$$\frac{\partial^2 f}{\partial \alpha^2} = \frac{\alpha^2(6 - 21d) + 2\alpha^4(-4 + d) - d + 16\alpha^3 d + \alpha(-2 + 8d)}{(1 - 2\alpha)^2(-1 + \alpha)^2 \alpha^2}$$
$$\frac{\partial^2 f}{\partial \alpha \partial \gamma} = \frac{\alpha(2 - 4d) + d + \alpha^2 d}{(-1 + \alpha)^2 \alpha^2}$$
$$\frac{\partial^2 f}{\partial \alpha \partial \varepsilon} = \frac{\left(1 - 4\alpha + 2\alpha^2\right) d}{(1 - 2\alpha)^2 \alpha^2}$$
$$\frac{\partial^2 f}{\partial \gamma^2} = \frac{-1 + \left(-1 + 4\alpha - 2\alpha^2\right) d}{(-1 + \alpha)^2 \alpha^2}$$
$$\frac{\partial^2 f}{\partial \gamma \partial \varepsilon} = -\frac{d}{\alpha^2}$$
$$\frac{\partial^2 f}{\partial \varepsilon^2} = -\frac{\left(1 - 2\alpha + 2\alpha^2\right) d}{(1 - 2\alpha)^2 \alpha^2}.$$

Using Mathematica to calculate the characteristic polynomial of the Hessian matrix, we obtain

$$\frac{1}{(1-\alpha)^3 \alpha^5 (1-2\alpha)^3}\left((2d - 8\alpha d + 12\alpha^2 d - 8\alpha^3 d + 2\alpha d^2 - 6\alpha^2 d^2 + 8\alpha^3 d^2 - 2\alpha^3 d^3)\right.$$

$$
\begin{aligned}
&+ (2\alpha^2 - 12\alpha^3 + 24\alpha^4 - 16\alpha^5 + 2\alpha d - 14\alpha^2 d + 50\alpha^3 d - 102\alpha^4 d + 120\alpha^5 d - 92\alpha^6 d \\
&\qquad + 40\alpha^7 d + 2\alpha^3 d^2 - 10\alpha^4 d^2 + 10\alpha^5 d^2 + 16\alpha^6 d^2 - 20\alpha^7 d^2)x \\
&+ (\alpha^3 - 5\alpha^4 + 6\alpha^5 + 2\alpha^6 + 4\alpha^7 - 24\alpha^8 + 16\alpha^9 + 3\alpha^3 d - 29\alpha^4 d + 116\alpha^5 d - 236\alpha^6 d \\
&\qquad + 246\alpha^7 d - 116\alpha^8 d + 16\alpha^9 d)x^2 \\
&+ (\alpha^5 - 9\alpha^6 + 33\alpha^7 - 63\alpha^8 + 66\alpha^9 - 36\alpha^{10} + 8\alpha^{11})x^3 \Big) .
\end{aligned}
$$

Recall that $\alpha = O(\log d/d)$. Thus by taking d sufficiently large, it is enough to consider only the leading terms in the coefficients of the characteristic polynomial. Recall also that $\alpha \le 1/2$. First, note that the value of the polynomial at $x = 0$ which is given by the constant coefficient is positive. Secondly, each of the other coefficients is also positive, so that for all $x > 0$, the derivative of the characteristic polynomial is positive. Hence the polynomial itself has only real negative roots and the Hessian is negative definite. Thus, f is strictly concave at $\alpha^*, \gamma^*, \varepsilon^*$ and must decay quadratically around its maximum. Hence we have that for any $(\alpha, \gamma, \varepsilon)$,

$$
f(\alpha^*, \gamma^*, \varepsilon^*) - f(\alpha, \gamma, \varepsilon) \ge C(|\alpha - \alpha^*|^2 + |\gamma - \gamma^*|^2 + |\varepsilon - \varepsilon^*|^2). \qquad \square
$$

We conclude by establishing Proposition 3.3 which says that second moment of the partition function can be bounded by the square of the first moment up to a polynomial term.

Proof of Proposition 3.3. Applying the Cauchy-Schwartz inequality we have

$$
\begin{aligned}
\mathbb{E}((Z_G)^2) &= \sum_{\alpha, \alpha'} \mathbb{E}\left(Z_{G,\alpha} Z_{G,\alpha'}\right) = \tilde{\Theta}(1) \sum_{\alpha} \mathbb{E}((Z_{G,\alpha})^2) \\
&= \tilde{\Theta}(1) \sum_{\alpha, \gamma, \epsilon} \mathbb{E} Z^{(2)}_{G,\alpha,\gamma,\varepsilon} \le \exp(n f(\alpha^*, \gamma^*, \epsilon^*) + O(\log n)) \\
&= \exp(2\Phi(\alpha^*)n + O(\log n)) = \tilde{\Theta}(1)(\mathbb{E} Z_G)^2 , \qquad \square
\end{aligned}
$$

where the inequality is by Proposition 3.1 while the penultimate equality is by Lemma (6.2).

References

[1] D. Achlioptas and A. Coja-Oghlan. Algorithmic barriers from phase transitions. In *Proceedings of IEEE Foundations of Computer Science*, pages 793–802, 2008.

[2] V. Bapst, A. Coja-Oghlan, S. Hetterich, F. Rassmann, and D. Vilenchik. The condensation phase transition in random graph coloring. Preprint, 2014. `http://arxiv.org/abs/1404.5513`. MR-3319009

[3] J. Barbier, F. Krzakala, L. Zdeborová, and P. Zhang. The hard-core model on random graphs revisited. *Journal of Physics: Conference Series*, 473:012–021, IOP Publishing, 2013.

[4] N. Berger, C. Kenyon, E. Mossel, and Y. Peres. Glauber dynamics on trees and hyperbolic graphs. *Probability Theory and Related Fields*, 131:311–340, 2005. MR-2123248

[5] N. Bhatnagar, A. Sly, and P. Tetali. Reconstruction threshold for the hardcore model. In *Proceedings of the 14th International Conference on Randomization and Computation (RANDOM)*, volume 6302 of *Lecture Notes in Computer Science*, pages 434–447. Springer, 2010. MR-2755854

[6] N. Bhatnagar, J. Vera, E. Vigoda, and D. Weitz. Reconstruction for colorings on trees. *SIAM Journal on Discrete Mathematics, Special Issue for Message Passing Algorithms*, 25(2):809–826, 2011. MR-2817532

[7] P. Bleher, J. Ruiz, and V. Zagrebnov. On the purity of the limiting gibbs state for the ising model on the bethe lattice. *Journal of Statistical Physics*, 79(1):473–482, 1995. MR-1325591

[8] B. Bollobás. A probabilistic proof of an asymptotic formula for the number of labeled random graphs. *European Journal of Combinatorics*, 1:311–316, 1980. MR-0595929

[9] C. Borgs, J. Chayes, E. Mossel, and S. Roch. The kesten-stigum reconstruction bound is tight for roughly symmetric binary channels. In *Proceedings of the 47th IEEE Symposium on Foundations of Computer Science*, pages 518–530, 2006.

[10] G. Brightwell and P. Winkler. A second threshold for the hard-core model on a Bethe lattice. *Random Structures and Algorithms*, 24:303–314, 2004. MR-2068871

[11] A. Coja-Oghlan and C. Efthymiou. On independent sets in random graphs. In *Proceedings of the Twenty-second Annual ACM-SIAM Symposium on Discrete Algorithms*, SODA '11, pages 136–144, 2011. MR-2857116

[12] C. Daskalakis, E. Mossel, and S. Roch. Optimal phylogenetic reconstruction. In *Proceedings of the ACM Symposium on Theory of Computating*, pages 159–168, 2006. MR-2277141

[13] A. Dembo, A. Montanari, and S. Sun. Factor models on locally tree-like graphs. *Annals of Probability*, 41, 2013. MR-3161472

[14] W. Evans, C. Kenyon, Y. Peres, and L. J. Schulman. Broadcasting on trees and the Ising model. *The Annals of Applied Probability*, 10(2):410–433, 2000. MR-1768240

[15] A. Frieze. On the independence number of random graphs. *Discrete Mathematics*, 81(2):171–175, 1990. MR-1054975

[16] Alan M Frieze and T Łuczak. On the independence and chromatic numbers of random regular graphs. *Journal of Combinatorial Theory, Series B*, 54:123–132, 1992. MR-1142268

[17] D. Gamarnik and M. Sudan. Limits of local algorithms over sparse random graphs. In *Proceedings of the 5th Conference on Innovations in Theoretical Computer Science*, ITCS '14, pages 369–376, 2014. MR-3359490

[18] H.-O. Georgii. *Gibbs measures and phase transitions.* de Gruyter, Berlin, 1988. MR-0956646

[19] A. Gerschenfeld and A. Montanari. Reconstruction for models on random graphs. In *Proceedings of the IEEE Symposium on Foundations of Computer Science*, pages 194–204, 2007.

[20] S. Janson and E. Mossel. Robust reconstruction on trees is determined by the second eigenvalue. *Annals of Probability*, 32(3B):2630–2649, 2003. MR-2078553

[21] H. Kesten and B.P. Stigum. Additional limit theorems for indecomposable multidimensional Galton-Watson processes. *Annals of Mathematical Statistics*, 37:1463–1481, 1966. MR-0200979

[22] F. Krzakala, A. Montanari, F. Ricci-Tersenghi, G. Semerjian, and L. Zdeborová. Gibbs states and the set of solutions of random constraint satisfaction problems. *Proceedings of the National Academy of Sciences*, 104:10318–10323, 2007. MR-2317690

[23] Mézard M. and A. Montanari. Reconstruction on trees and spin glass transition. *Journal of Statistical Physics*, 124:1317–1350, 2006. MR-2266446

[24] J. Martin. Reconstruction thresholds on regular trees. In *DMTCS volume Proc. Discrete Random Walks*, pages 191–204, 2003. MR-2042387

[25] F. Martinelli, A. Sinclair, and D. Weitz. Fast mixing for independent sets, colorings, and other models on trees. In *Proceedings of ACM-SIAM SODA*, pages 449–458, 2004. MR-2291084

[26] Marc Mezard and Andrea Montanari. *Information, physics, and computation.* Oxford University Press, 2009. MR-2518205

[27] A. Montanari, E. Mossel, and A. Sly. The weak limit of ising models on locally tree-like graphs. *Probability Theory and Related Fields*, 152:31–51, 2012. MR-2875752

[28] A. Montanari, R. Restrepo, and P. Tetali. Reconstruction and clustering in random constraint satisfaction problems. *SIAM Journal on Discrete Mathematics*, 25(2):771–808, 2011. MR-2823097

[29] E. Mossel. Reconstruction on trees: Beating the second eigenvalue. *Annals of Applied Probabability*, 11:285–300, 2001. MR-1825467

[30] E. Mossel. Survey: Information flow on trees graphs, morphisms and statistical physics. *DIMACS Series in Discrete Mathematics and Theoretical Computer Science*, pages 155–170, 2004. MR-2056226

[31] E. Mossel, D. Weitz, and N. Wormald. On the hardness of sampling independent sets beyond the tree threshold. *Probability Theory and Related Fields*, 143:401–439, 2009. MR-2475668

[32] M. Rahman and B. Virag. Local algorithms for independent sets are half-optimal. Preprint, 2014. `http://arxiv.org/abs/1402.0485`.

[33] R. Restrepo, D. Stefankovic, J.C. Vera, E. Vigoda, and L. Yang. Phase transition for glauber dynamics for independent sets on regular trees. In *Proceedings of the ACM-SIAM Symposium on Discrete Algorithms*, pages 945–956, 2011. MR-2857176

[34] A. Sly. Reconstruction for the Potts model. In *Proceedings of the Symposium for Theory of Computing*, pages 581–590, 2009. MR-2780104

[35] A. Sly. Reconstruction of random colourings. *Communications in Mathematical Physics*, 288:943–961, 2009. MR-2504861

[36] A. Sly. Computational transition at the uniqueness threshold. In *IEEE Symposium on Foundations of Computer Science*, pages 287–296, 2010. MR-3025202

[37] A. Sly and N. Sun. The computational hardness of counting in two-spin models on d-regular graphs. In *Proceedings of the 2012 IEEE 53rd Annual Symposium on Foundations of Computer Science*, FOCS '12, pages 361–369, 2012. MR-3186623

[38] P. Tetali, J. Vera, E. Vigoda, and L. Yang. Phase transition for the mixing time of the glauber dynamics for coloring regular trees. In *Proceedings of the Twenty-First Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 1646–1656, 2010. MR-2809774

[39] N. Wormald. Differential equations for random processes and random graphs. *Annals of Applied Probability*, 5:1217–1235, 1995. MR-1384372

[40] N. Wormald. Models of random regular graphs. *London Mathematical Society Lecture Note Series*, pages 239–298, 1999. MR-1725006