Reconstruction/Non-reconstruction Thresholds for Colourings of General Galton-Watson Trees.

Charilaos Efthymiou
Goethe University, Mathematics Institute, Frankfurt 60325, Germany
efthymiou@gmail.com

Abstract

The broadcasting models on trees arise in many contexts such as discrete mathematics, biology, information theory, statistical physics and computer science. Here we consider the $k$-colouring model. A basic question here is whether the root’s assignment affects the distribution of the colourings at the vertices at distance $h$ from the root. This is the so-called reconstruction/non-reconstruction problem. For the case where the underlying tree is $d$-ary and $d$ is sufficiently large, it is well known that $d/\ln d$ is a threshold function. That is, for $k = (1 + \epsilon)d/\ln d$ we have non-reconstruction while for $k = (1 - \epsilon)d/\ln d$ we have reconstruction.

In this work, we consider the case where the underlying tree is a random one. In particular, our focus is on the well-known Galton-Watson trees. This model arises naturally in many contexts, e.g. the theory of spin-glasses and its applications on random Constraint Satisfaction Problems (rCSP). The aforementioned study focuses on Galton-Watson trees with offspring distribution binomial with parameters $n$ and $d/n$, i.e. $B(n,d/n)$. However, here we consider a broader version of the problem, as we assume general offspring distribution for the tree.

Our aim is to relate the corresponding bounds of $k$ for (non)reconstruction with certain concentration properties of the offspring distribution. Interestingly enough, our analysis allows to decouple the distributions of the trees and the colourings. E.g. when non-reconstruction is a typical phenomenon of the tree instances, then we provide the structural properties which imply the corresponding spatial mixing property.

For the case of Galton-Watson trees with offspring $B(n,d/n)$ we verify that indeed $d/\ln d$ is a threshold function for the problem. A very interesting consequence of our analysis is that for offspring distribution with expected offspring $d$, we get the threshold function $d/\ln d$ under weaker concentration conditions compared to $B(n,d/n)$.

1 Introduction

The broadcasting models on trees and the closely related reconstruction problem were originally studied in statistical physics. Since then, they have found applications in other areas including biology (in phylogenetic reconstruction [8] [25]), communication theory (in the study of noisy computation [13]). The models on trees appear in computer science, too. They arise in the study of random Constraint Satisfaction Problems (rCSP) such as random $k$-SAT, random graph colouring e.t.c. Very impressively, the models on trees seem to capture some of the most fundamental properties of the corresponding models on random (hyper)graphs. [17].

The most basic problem in the study of broadcasting models is to determine the reconstruction/non-reconstruction threshold. I.e. whether the configuration of the root affects the distribution of the configuration of distant vertices. The transition from non-reconstruction to reconstruction can be achieved by adjusting appropriately the parameters of the model. Typically, this transition exhibits a threshold behaviour. So far, the main focus of the study was to determine the precise location of this threshold for various models when the underlying graph is a fixed (infinite) tree. In this work we extend the study of
the phenomenon to the cases where the tree is not fixed but it is chosen according to some probability distribution. In particular, we consider Galton-Watson trees (GW-trees) with some general offspring distribution with expectation $d$.

One could remark that in our setting we have to deal with two levels of randomness. One is w.r.t. the tree instances and the second is w.r.t. the random $k$-colourings of the tree. Our analysis manages to “decouple” these two different kinds of randomness. In our study of (non)reconstruction on GW trees we search for the structural properties of the underlying tree which somehow play prominent role in the (non)reconstruction. That is, we not only get the reconstruction/non-reconstruction bounds, but we get the structural properties of the tree which implies the corresponding spatial mixing, too.

The main technical challenge we have to deal with is the so-called “effect of high degrees”. That is, we expect to have vertices in the tree which are of degree much higher than the expected offspring. Their involvement in the analysis in non-trivial. The deviation from the expected degree is so large that expressing the (non)reconstruction bounds in terms of maximum degree leads to highly suboptimal results. This challenge also appears in related problems in random graphs $G(n, d/n)$, e.g. sampling colourings \cite{11, 10}. In our analysis, we should accommodate groups of high degree vertices of the tree and show that they have a negligible effect on the spatial mixing.

The hypothesis is that in GW-trees with $B(n, d/n)$ offspring distribution we have a reconstruction/non-reconstruction threshold that coincides with that of a $d$-ary tree i.e. the threshold function is $d/\ln d$ (we extrapolate when $d$ is non-integer). The intuition behind this hypothesis relies on the fact that for sufficiently large $d > 0$ the offspring distribution is concentrated about its expectation $d$. Our aim is to make the intuitive base of this relation rigorous. As a matter of fact, we do not restrict our study to the binomial case only. We consider a general offspring distribution.

When the offspring distribution is sufficiently concentrated about its expectation $d$, then we show that the transition from non-reconstruction to reconstruction exhibits a threshold behaviour at the critical point $d/\ln d$. Interestingly, our analysis implies that we get a threshold behaviour under much weaker concentration conditions than what we have in $B(n, d/n)$. When the concentration is not sufficiently high to provide thresholds, we still get bounds for reconstruction and non-reconstruction that depend on the tails of the offspring distribution.

2 Definitions and Results

So as to give a definition of the model and the reconstruction problem it is more convenient to describe them first in a setting where the underlying tree is fixed. That is, we consider a fixed tree $T$ of height $h$ and then we will add randomness on the tree instances. W.l.o.g. the reader my well assume that the tree is $\Delta$-ary for some integer $\Delta > 0$.

The broadcasting models on a tree $T$ are models in which information is sent from the root over the edges to the leaves. We assume that the edges represent noisy channels. For some finite set of spins $\Sigma = \{1, 2, \ldots, k\}$, a configuration on $T$ is an element in $\Sigma^T$, i.e. it is an assignment of spins to the vertices of $T$. The spin of the root $r$ is chosen according to some initial distribution over $\Sigma$. The information propagates along the edges of the tree as follows: There is a $k \times k$ stochastic matrix $M$ such that if the vertex $v$ is assigned spin $i$, then its child $u$ is assigned spin $j$ with probability $M_{i,j}$.

Our focus is on the $k$-colouring model (also known as $k$-state Potts model at zero temperature), i.e. the matrix $M$ is as follows:

$$M_{i,j} = \begin{cases} \frac{1}{k-1} & \text{for } i \neq j \\ 0 & \text{otherwise.} \end{cases}$$

Broadcasting models give rise to Gibbs measures on trees. E.g. for the colouring model, assuming that the broadcasting process over $T$ starts with the root $r(T)$ coloured $i$, then the $k$-colouring we get after the processes has finished is a random $k$-colouring of $T$ conditional that $r(T)$ is coloured $i$. 

$2$
We let $L_h(T)$ denote the set of vertices at distance $h$ from the root $r(T)$. Also, we let $\mu^i$ denote the uniform distribution over the $k$-colourings of $T$ conditional that $r(T)$ is assigned colour $i$. Reconstructibility is defined as follows:

**Definition 1** For any $i, j \in [k]$ let $||\mu^i - \mu^j||_{L_h}$ denote the total variation distance of the projections of $\mu^i$ and $\mu^j$ on $L_h$. We say that a model is reconstructible on a tree $T$ if there exists $i, j \in [k]$ for which

$$\lim_{h \to \infty} ||\mu^i - \mu^j||_{L_h(T)} > 0.$$ 

When the above limit is zero for every $i, j$, then we say that the model has non-reconstruction.

(Non)Reconstructibility expresses how information decays along the tree. As a matter of fact, non-reconstruction is equivalent to the mutual information between the colouring of root $r(T)$ and that of $L_h$ going to zero as $h$ grows (see [24]). Non-reconstruction is equivalent to the Gibbs measure being extremal [15]. For a finite tree $T$, non-reconstruction implies that typical colourings of the vertices at level $h$ of the tree have a vanishing effect on the distribution of the colouring of $r(T)$, as $h$ grows.

The reconstruction problem extends naturally to the cases where the underlying tree is not fixed but it is chosen from a distribution. Here we consider the well-known Galton Watson trees (GW-trees) with a general offspring distribution. That is,

**Definition 2** Let $\xi$ be a distribution on the non negative integers with expectation $d_\xi$. We let $T_\xi$ denote a Galton-Watson tree with offspring distribution $\xi$. Also, for any integer $h > 0$ we let $T_\xi^h$ denote the restriction of $T_\xi$ to its first $h$ levels.

For convenience we consider $\xi$ as a stochastic vector, i.e. $\xi_i$ is the probability for a random variable distributed as in $\xi$ to be equal to $i$, for some integer $i \geq 0$. Also, w.l.o.g. we assume that there is a set of parameters for $\xi$ which can be adjusted so as to have sufficiently large $d_\xi > 0$. This kind of parametrization makes it possible to study reconstruction/non-reconstruction threshold for the random colourings of $T_\xi^h$.

For $T_\xi^h$, the reconstruction/non-reconstruction from Definition 1 extends as follows:

**Definition 3** We say that a model is reconstructible on $T_\xi^h$ if there exists $i, j \in [k]$ for which

$$\lim_{h \to \infty} \mathbb{E}||\mu^i - \mu^j||_{L_h} > 0,$$

where the expectation is w.r.t. the instances of the tree. When the above limit is zero for every $i, j$, then we say that the model has non-reconstruction.

It is possible that the height of the random tree does not reach height $h$. In this case the total variation distance in Definition 1 is zero. Usually the transition from non-reconstruction to reconstruction exhibits a threshold behaviour. We investigate this behaviour for GW-trees, too.

**Definition 4** For the $k$-colouring model on $T_\xi^h$ the transition from reconstruction to non-reconstruction exhibits a threshold behaviour if the following holds: There is a threshold function $K : \mathbb{R}^+ \to \mathbb{R}^+$ such that for any $\alpha > 0$ and sufficiently large $d_\xi$, we have non-reconstruction when $k \geq (1 + \alpha)K(d_\xi)$, while we have reconstruction when $k \leq (1 - \alpha)K(d_\xi)$.

One of the main results of this work is the following theorem.

---

1Gibbs distribution

2In other words, $T_\xi^h$ is the induced subtree of $T_\xi$ which contains all the vertices within graph distance $h$ from the root.
Theorem 1  Let some fixed $\gamma > 0$. Consider $T^h_\xi$ such that $d_\xi \geq d(\gamma)$. Also, assume that for any random variable $Z$ distributed as in $\xi$ the following is true for every $x \geq (1 + \gamma)d_\xi$

$$\Pr[Z \geq x] \leq x^{-c} \quad \text{and} \quad \Pr[Z \leq (1 - \gamma)d_\xi] \leq (d_\xi)^{-c},$$

(1)

for sufficiently large $c > 0$, which is independent of $d_\xi$ and $\gamma$. Then, $d_\xi / \ln d_\xi$ is a threshold function for reconstruction/non-reconstruction transition.

The above theorem follows as a corollary of the results in Section 4. In particular it follows from the more general (also more technical) Theorem 2. Theorem 2 is more general in that it covers non-threshold cases, too.

Observe that the concentration conditions in Theorem 1 are weaker than what we can get for $\mathcal{B}(n,d/n)$. I.e. using standard Chernoff bounds (e.g. [27]) it is direct to show that the quantities of interest in (1) for $\mathcal{B}(n,d/n)$ are of smaller order of magnitude than polynomial. That is, the binomial distribution exhibits stronger concentration properties than what we actually need for reconstruction/non-reconstruction threshold. From Theorem 1 and standard tail bounds for $\mathcal{B}(n,d/n)$ we get the following Corollary 1

Consider the $k$-colouring model on $T_\xi$ where $\xi$ is the distribution $\mathcal{B}(n,d/n)$. Then $d_\xi / \ln d$ is the non-reconstruction/reconstruction threshold function.

3 High Level Description

In this section, we provide an high level description of how do we derive upper and lower bound for reconstruction and non-reconstruction, respectively. Consider an instance of $T^h_\xi$. We assume that the expectation of $\xi$, i.e. $d_\xi$, is sufficiently large, thus, the tree with positive probability will reach to height $h$.

Remark 1. We will use the term “random colouring of a vertex set $\Lambda$” to indicate the following colouring of the vertices in $\Lambda$: Take a random colouring of the tree and keep only the colouring of the vertices in $\Lambda$. Also, when we refer to “typical colourings of vertex set $\Lambda$”, we imply that they are typical w.r.t. the aforementioned distribution.

Depending on the tails of $\xi$ we get appropriate quantities $\Delta_+$ and $\Delta_-$ such that $\Delta_- < d_\xi < \Delta_+$. Given these two quantities we show that we have non-reconstruction for $k \geq (1 + \alpha)\Delta_+/\ln \Delta_+$ and we have reconstruction for $k \leq (1 - \alpha)\Delta_-/\ln \Delta_-$, where $\alpha > 0$ is fixed.

For the non-reconstruction our approach has two steps. The first step is to show that there is a sufficiently large and well spread set of vertices in $T^h_\xi$ such that a typical $k$-colouring of the vertices at level $h$ does not bias their colour assignments too much when $k \geq (1 + \alpha)\Delta_+ / \ln \Delta_+$. A vertex is biased if the boundary forces it to prefer a specific, relatively small, set of colours. E.g. one vertex could be the root of a subtree of $T^h_\xi$ whose all vertices are of degrees much larger than $\Delta_+$. Then, for such
vertex a typical colouring of the vertices at level $h$ of $T^h$ is biasing. These vertices are the “bad” part of the tree. The rest is the good part of the tree. The vertices of the good part are called mixing.

To be more specific about the structural properties, the requirements are as follows: For every path from the root of $T^h$ to the vertices at level $h$ we should have a sufficiently large fraction of mixing vertices. Additionally, the number of vertices at level $h$ should not deviate significantly from their expectation. For any, arbitrary, instance of $T^h$ which has these properties, we show that non-reconstruction holds. The non-reconstruction bound follows by showing that $T^h$ has these structural properties with probability that tends to 1 as $h \to \infty$.

The second step is to show non-reconstruction given that the underlying tree has the aforementioned structural properties. For showing this, we use the idea, introduced in [4], which shows non-reconstruction by upper bounding appropriately the second moment of a quantity called “magnetization of the root”. This approach underlies the derivation of non-reconstruction bounds for various models on fixed trees e.g. [2, 29, 3, 4]. Our approach builds on the very elegant combinatorial approach from [2], which uses the notion of unbiasing boundary to bound the second moment of the magnetization of the root. That is, the setting we develop allows to use the idea of unbiasing boundaries and derive appropriate upper bounds for the second moment of the magnetization.

Before proceeding, let us be more detailed about the mixing vertices. So as to decide whether a vertex is mixing, we should check the subtree that includes all its decadents. Roughly speaking, a non leaf vertex $v$ is mixing if the following (inductively defined) criterion is satisfied: Its degree is at most $\Delta_+$. If the vertex $v$ has children, then at most $o(\Delta_+)$ of them are non-mixing vertices. The rest should be mixing. We consider leaves to be mixing vertices.

In Figure 1 the “good subtree” contains exactly the mixing vertices. The “bad subtrees” contains the rest. The above restriction implies that a vertex in the good part can have only a limited number of neighbours in the bad subtrees. The root of $T^h$ is not necessarily a mixing vertex. The value of $\Delta_+$ is taken such that the typical instances of $T^h$ has the desired structural properties with probability tending to 1 as the height $h$ tends to infinity.

Consider, now, the reconstruction bound, which is an easier case. The reconstruction bound is a well known when the offspring distribution is $B(n, d/n)$, e.g. [22, 28]. Our approach deviates from both [22, 28] in that we focus on the structural properties of the underlying tree which imply reconstruction. That is, we propose a set of structural properties that $T^h$ has with sufficiently large probability and then we show that for any tree with such properties reconstruction holds. Our approach applies to general offspring distribution.

Let us be more specific. For any tree $T$ of height $h$ and a given $k$, a sufficient condition to have reconstruction is the following one: With probability bounded away from zero for any $h$, a random colouring at the vertices at level $h$ “freezes” the colour assignment of a great proportion of the tree, which includes the root. That is, the colour assignments of the vertices specify a unique colour assignment for the root of $T$. For a $\Delta_+$-ary tree and $k \leq (1 - \alpha)(\Delta_+)/\ln \Delta_-$, it has been shown in [28, 26] that a random $k$-colouring of the vertices at level $h$ freezes the colour assignment at the root with this large probability.

Somehow, the above implies that if $T^h$ has a $(\Delta_+)$-ary subtree of height $h$ with the same root as $T^h$, then a random $k$-colouring of the vertices at level $h$ freezes the colouring of the root with probability bounded away from zero when $k \leq (1 - \alpha)\Delta_+ / \ln \Delta_-$. Depending on $\xi$, we choose $\Delta_-$ such that $T^h$ has such a subtree with probability bounded away from zero for any $h$. Then reconstruction follows.

Figure 2 illustrates exactly the idea. In the “freezable subtree” every non-leaf vertex is of degree at least $(\Delta_-)$. The effect from the rest of the tree is not important as the existence of the freezable part guarantees that the colouring of the root is going to freeze with positive probability over the random colourings of the vertices at level $h$.

---

3The quantity $o(\Delta_+)$ should be of smaller order than linear in $\Delta_+$. 

5
We would like to remark the following: A non leaf vertex in the freezable subtree does not only have degree at least \( \Delta \), but a sufficiently large number of its children should belong to the freezable subtree too. Each leaf is defined to be a freezable subtree.

3.1 Motivation from Statistical Physics

Even though we study a more general problem, the case of GW-trees with \( B(n, d/n) \) offspring distribution is of particular interest in Computer Science, Statistical Physics and many other areas. This interest originates from the study of random Constraint Satisfaction Problems, e.g. random \( k \)-SAT, colouring or independent sets of a random graph \( G(n, d/n) \) etc.

The ingenious, however, mathematically non-rigorous Cavity Method \[21, 17\] makes very impressive predictions about some of the most fundamental properties of rCSP. One of the most interesting parts of these predictions involve the Gibbs distribution, with a focus on its spatial mixing properties. The Cavity Method predicts that some of the most basic properties of the Gibbs distribution of the models on random (hyper)graphs can be captured by the Gibbs distributions of the corresponding models on aforementioned GW-trees. This assumption is studied in \[14, 23\].

For the \( k \)-colouring problem on a random graph \( G(n, d/n) \) the Cavity Method predicts that there is the critical value \( k_0 = k_0(d) \) such that the following holds: For \( k = (1 + \epsilon)k_0 \) we have non-reconstruction, while for \( k = (1 - \epsilon)k_0 \) we have reconstruction. The asymptotic value\(^4\) for \( k_0 \) is \( k_0 \sim d/\ln d \). It also predicts the same critical value for the reconstruction/non-reconstruction transition for the \( k \)-colourings of the GW-trees with offspring \( B(n, d/n) \). The part of the prediction which is related to GW-trees follows as a corollary from Theorem \[1\] (i.e. Corollary \[1\]). Using the result \[23\], we could argue that Corollary \[1\] implies non-reconstruction/reconstruction transition for the \( k \)-colourings of \( G(n, d/n) \) at the critical point \( k_0 \sim d/\ln d \).

Generally, it is reasonable to expect that the case of trees is much easier to handle, than that of the graphs. Still, the study the spatial mixing on GW-trees is still very challenging task, as we discussed above, due to the fluctuating degrees of the vertices. The effect of high degrees become even more dramatic if someone considers non-symmetric models like \( k \)-SAT or independent sets.

The non-reconstruction phenomenon in rCSP seems to play a vital role in algorithmic problems. In particular, it has been related to the efficiency of local algorithms which search for satisfying solutions. That is, when we have non-reconstruction, usually there is an efficient (simple) local algorithm which finds solutions e.g. \[5, 16\]. On the other hand, in the reconstruction regime there is no efficient algorithm which finds solutions. For this reason, the transition from non-reconstruction to reconstruction has been given the name “algorithmic barrier”, in \[1\]. We should mention that there is no rigorous result of computational hardness related to this transition.

Another, bold, prediction from statistical physics is that in the non-reconstruction regime it is possible not only to find solutions of rCSP efficiently but to (approximately) count them efficiently, too. That is, the spatial mixing properties we have in the non-reconstruction regime allows efficient (approximate) sampling from the Gibbs distribution. This area has a lot of open problems as the state of the art of the approximate sampling algorithm has not gone beyond the Gibbs uniqueness region, e.g. \[10, 11, 12\]. For colouring, the Gibbs uniqueness threshold is by factor \( \ln d \) off from the reconstruction/non-reconstruction threshold.

4 Upper and Lower Bounds

We start our analysis by focusing on the upper and lower bounds for the number of colours \( k \) for reconstruction and non-reconstruction, respectively, for some GW-tree \( T^h_{t\xi} \). These bounds are expressed,
respectively, in terms of the quantities $\Delta_-$ and $\Delta_+$. Both of them depend on the statistics of the offspring distribution $\xi$, i.e., they depend on the lower and the upper tail of $\xi$, respectively. In particular the we have the following:

**Definition 5** Consider a distribution $\xi$ over the non negative integers with expectation $d_\xi$. Given some fixed $\delta \in (0, 1/10)$, we let $\Delta_+ = \Delta_+(\delta) > d_\xi$ be the minimum integer such that the following holds: There is sufficiently large $\beta \geq 4$ independent of $d_\xi$ and $q \in [0, 3/4)$ such that

\[
q \geq \sum_{i > \Delta_+} \xi_i + \Pr \left[ B(\Delta_+, q) \geq (\Delta_+)^\delta \right]
\]

and

\[
\sum_{t > \Delta_+} t \cdot \xi_t \leq \exp (-2\beta \ln d_\xi), \quad \Pr \left[ B(\Delta_+, q) > (\Delta_+)^\delta \right] \leq \exp (-2\beta \ln d_\xi).
\]

Observe that once we have $\xi$ and $\delta$ we can specify $\Delta_+$. To get an intuition about the above conditions consider first $[2]$. Choosing appropriately $\delta$, the condition $[2]$ somehow implies that, regardless of the height of the tree, the probability of the root to be non-mixing is always upper bounded by $q$. Observe that $q$ does not depend on $h$. We will define (later) that the root of $T^h_\xi$ is non-mixing if either of the following two events holds. The degree of the root is greater than $\Delta_+$. The degree is at most $\Delta_+$ but there are more than $\Delta_+^\delta$ children of the root which are non-mixing roots in their own subtrees. If $q$ is an upper bound for the root to be non-mixing, then the r.h.s. of the inequality in $[2]$ is an upper bound of the probability of either of these two events occur. Additionally, the conditions $[3]$ guarantees that every path from the root to the vertices at level $h$ of the tree $T^h_\xi$ has sufficiently many mixing vertices.

**Definition 6** Let $\xi$ be a distribution over the non negative integers. Given some $\delta \in (0, 1/10)$, we let $\Delta_- = \Delta_-(\delta) < d_\xi$ be the minimum integer such that the following holds: There is $g \in [0, 3/4)$ such that

\[
g \geq \sum_{i < \Delta_-} \xi_i + \sum_{i \geq \Delta_-} \xi_i \Pr \left[ B(i, 1 - g) < (\Delta_-) - (\Delta_-)^\delta \right].
\]

The intuition for $[4]$ is very similar to that for $[2]$. Given $\xi$ and $\delta$ we specify $\Delta_-$. The quantity $g$ is an upper bound for the probability that the root of $T^h_\xi$ is non-freezable, i.e. it is the root of a freezable subtree of $T^h_\xi$. The quantity $g$ does not depend on the height of the tree. The root of $T^h_\xi$ is non-freezable if either of the following (sufficient) conditions hold. The root is of degree less than $\Delta_-$. The root is of degree at least $\Delta_-$ while there are at least $(\Delta_-)^\delta$ children of the root such that each of them is non-freezable in its own subtree. If $g$ is an upper bound for the root to be non-freezable, then the r.h.s. of $[4]$ is an upper bound of the probability that either of the two conditions hold.

The following theorem is the main technical result of our work.

**Theorem 2** Let some fixed $\alpha > 0$. Consider an instance of $T^h_\xi$ such that $d_\xi$ is sufficiently large. Consider also some $\delta \in (\alpha/2, 1/10)$, which specifies both $\Delta_-$ and $\Delta_+$. For $\mu$, the Gibbs distribution over the $k$-colourings of $T^h_\xi$ the following is true:

**non-reconstruction:** For $k = (1 + \alpha)\Delta_+/\ln \Delta_+$ and any $i, j \in [k]$ it holds that

\[
\mathbb{E}||\mu^i - \mu^j||_{L_h} \leq 8k (2\Delta_+)^{-0.456h}.
\]

**reconstruction:** For $k = (1 - \alpha)\Delta_-/\ln \Delta_-$ there are $i, j \in [k]$ such that

\[
\mathbb{E}||\mu^i - \mu^j||_{L_h} \geq \left(1 - \frac{2}{\log k}\right)(1 - g),
\]
where \( g \) is from Definition 2. Both of the expectations above are taken w.r.t. the tree instances.

The proof of Theorem 2 appears in two sections. In Section 3 we present the proof for the non-reconstruction part. In Section 4 we present the proof for the reconstruction part.

Given Theorem 2, it is elementary to show that Theorem 1 holds. I.e. given the conditions in 1, we have to show that \( \Delta_- \) and \( \Delta_+ \) are sufficiently close to each other. The derivations are simple and they are presented in full detail in Section 14.

In the rest of this section we present some notation which is going to be very useful for the proofs of the aforementioned results.

**Notation.** For any tree \( T \) we let \( r(T) \) or \( r_T \) denote its root. Let \( L_h(T) \) denote the set of vertices at graph distance \( h \) from \( r(T) \). For every vertex \( v \in T \), we define \( T_v \) the subtree of \( T \) as follows: Delete the edge between \( v \) and its parent in \( T \). Then \( T_v \) is the connected component that contains \( v \). We use the convention that \( r(T_v) = v \).

We use capital letter of the Latin alphabet to indicate random variables which are colourings of the tree \( T \), e.g. \( X, Y, \) etc. We use small letter of the greek alphabet to indicate fixed colourings, e.g. \( \sigma, \tau, \) etc. We use the notation \( \sigma_\Lambda \) or \( X(\Lambda) \) do indicate that the vertices in \( \Lambda \) have a colour assignment specified by the colouring \( \sigma \) or \( X \), respectively.

Given a tree \( T \), we let \( \mu \) denote the Gibbs distribution for its \( k \)-colourings. Usually we consider \( \mu \) under certain boundary conditions, i.e. given some \( \Lambda \subset T \), and some \( k \)-colouring of \( T, \sigma \), we need to consider the Gibbs distribution where the vertices in \( \Lambda \) have fixed colouring \( \sigma_\Lambda \). For this case we denote the Gibbs distribution \( \mu^{\sigma_\Lambda} \). For \( \Xi \subseteq T \) we let \( \mu_{\Xi} \) denote the marginal of the Gibbs distribution for the vertices in \( \Xi \). We denote marginals over the vertex set \( \Xi \) of a Gibbs distribution with boundary \( \sigma_\Lambda \) in the natural way, i.e. \( \mu^{\sigma_\Lambda}_{\Xi} \).

## 5 Proof of Theorem 2 - Non Reconstruction

First, consider a fixed tree \( T \) of height \( h \) and we let \( L = L_h(T) \). From (5) we have that

\[
||\mu^T - \mu||_{rT} \leq k \sum_{\sigma(L) \in [k]^L} \mu_L(\sigma_L) \cdot ||\mu^{\sigma(L)} - \mu||_{rT}
\]

Furthermore, from the definition of the total variation distance we have that

\[
\sum_{\sigma(L) \in [k]^L} \mu_L(\sigma_L) \cdot ||\mu^{\sigma(L)} - \mu||_{rT} = \frac{1}{2} \sum_{\sigma(L) \in [k]^L} \mu_L(\sigma_L) \cdot \sum_{c \in [k]} \left| \mu^{\sigma(L)}_{r(T)}(c) - 1/k \right|
\]

\[
= \frac{1}{2} \sum_{c \in [k]} \sum_{\sigma(L) \in [k]^L} \mu_L(\sigma_L) \cdot \left| \mu^{\sigma(L)}_{r(T)}(c) - 1/k \right|.
\]

The quantity \( \left| \mu^{\sigma(L)}_{r(T)}(c) - 1/k \right| \), is usually refereed as magnetization of the root \( r(T) \) (5). The inner sum is the average magnetization at the root, w.r.t. boundaries at the set \( L \). We bound this average magnetization by using the following standard result.

**Proposition 1** Consider a fixed tree \( T \) of height \( h \) and some integer \( k > 0 \). For every \( c \in [k] \) the following is true: Let \( X \) be a random \( k \)-colouring of \( T \) conditional that the root is coloured \( c \). It holds that

\[
\sum_{\sigma(L) \in [k]^L} \mu_L(\sigma(L)) \cdot \left| \mu^{\sigma(L)}_{r(T)}(c) - 1/k \right| \leq \sqrt{\frac{1}{k} \left| \mu^{\sigma(L)}_{r(T)}(c) - 1/k \right|^2},
\]

where \( \mu^{\sigma(L)}_{r(T)}(c) = \frac{1}{|L|} \sum_{l \in L} \sigma_L(l) \).
where \( Z \) is random colouring of \( T \) conditional that \( r_T \) is coloured \( q \), while \( q \) is the colour that maximizes the r.h.s. of (7).

Our proof of Proposition 1, which is very similar to the proof of Lemma 1 in [4], appears in Section 12. The quantity in the square root in (7) is closely related to the expected of the square of the magnetization of the root. Essentially we are going to show that the expectation of the square of the magnetization of the root tends to zero with \( h \). This idea is used in others previous works e.g. [3, 4, 2, 29]. We set our analysis in such a way that the ideas from [2] can be adopted in our context so as to show non-reconstruction.

The reader should observe that the quantity on the r.h.s. of (7) is a deterministic one, i.e. it depends only on the underlying tree \( T, c \) and \( k \). We let

\[
G_{c,k}(T) = \left| \mu^X(\cdot) - \mu^{Z} L(\cdot) \right|_{r_T}. \tag{7}
\]

Since we impose randomness on the instances of the trees, clearly, \( G_{c,k}(T_h) \) is a random variable. It suffices to show that for any \( c \in [k] \) the expectation \( E[G_{c,k}(T_h)] \) tends to zero with \( h \) sufficiently fast.

**Definition 7 (Mixing Root)** For an instance of \( T_h \) the root is mixing if the following holds: For \( h = 0 \), then \( r(T) \) is always mixing. For \( h > 0 \), \( r(T) \) is mixing if and only if \( \text{deg}(r_T) < \Delta_+ \) and there are at most \((\Delta_+)^\delta \) many vertices \( v \) children of \( r(T) \) such that \( \tilde{T}_v \) does not have a mixing root.

**Remark.** In Definition, the quantity \( q \) is an upper bound for the probability that the root of \( T_h \) is not mixing.

**Definition 8** We let \( A_{h,\zeta} \) denote the set of trees \( T \) of height \( h \) such that the following holds: Any path \( P \) of length \( h \) from \( r(T) \) to \( L_h(T) \) contains at least \((1 - \zeta)h \) vertices \( v \) such that \( \tilde{T}_v \) has a mixing root.

Before presenting the following result, we need to remind the reader that in Definition\(^5\) given \( \xi \) and \( \delta \), for \( \Delta_+ \) among others the following inequality should hold

\[
\sum_{t \geq \Delta_+} t\xi_t < \exp \left( -2\beta \ln d_\xi \right)
\]

where \( \beta \geq 4 \). In the following proposition we need to use the quantity \( \beta \) from the above condition.

**Proposition 2** Assume that the distribution \( \xi, \delta, \Delta_+ \) are as defined in the statement of Theorem 2. Let \( \zeta \in (0, 1) \) and \( \theta = \theta(\zeta) > 1 \) be such that \((1 - \zeta)\theta < 1 \) and \( \beta(1 - \theta) < -1 \). Then, for every \( h \geq 1 \) it holds that

\[
\Pr[T_h \in A_{h,\zeta}] \geq 1 - \exp \left[ -\left(1 - \theta(1 - \zeta)\right)C \cdot h \right],
\]

where \( C = \beta \ln d_\xi \).

The proof of Proposition 2 appears in Section 11.

**Theorem 3** Let \( \xi, \Delta_+ \) and \( \alpha \) be as in the statement of Theorem 2. For \( k = (1 + \alpha)\Delta_+/\ln \Delta_+ \), it holds that

\[
E \left[ G \left( T_h \right) \mid T_h \in A_{h,\zeta} \right] \leq \frac{4(2\Delta_+)^{-0.9(3/4-\zeta)h}}{\Pr[T_h \in A_{h,\zeta}]}
\]

\(^5\)The square root comes from the standard application of Cauchy-Schwarz inequality \( \sqrt{E[X]} \leq \sqrt{E[X^2]} \).

9
The proof of Theorem 3 appears in Section 6.

By applying Proposition 2 we set \( \zeta = 1/4 \) and \( \theta = 1.3 \). Then we get that

\[
\Pr[\mathcal{T}_\xi^h \notin \mathcal{A}_{h,\zeta}] \leq d_\zeta^{.1h}.
\]  
(8)

Also, from Theorem 3 we get that

\[
\mathbb{E} \left[ \mathcal{G}(\mathcal{T}_\xi^h) \middle| \mathcal{T}_\xi^h \in \mathcal{A}_{h,\zeta} \right] \leq 8(2\Delta_+)^{-0.45h}
\]  
(9)

Since we always have \( 0 \leq \mathcal{G}(T) \leq 1 \) we get that

\[
\mathbb{E} \left[ \mathcal{G}(\mathcal{T}_\xi^h) \right] \leq \mathbb{E} \left[ \mathcal{G}(\mathcal{T}_\xi^h) \middle| \mathcal{T}_\xi^h \in \mathcal{A}_{h,1/4} \right] + \Pr \left[ \mathcal{T}_\xi^h \notin \mathcal{A}_{h,1/4} \right] \\
\leq 16(2\Delta_+)^{-0.45h}.
\]

where the last inequality follows from (8) and (9). The theorem follows.

6 Proof of Theorem 3

Consider first the quantity \( \mathcal{G}_{c,k}(T) \), for some fixed tree \( T \). Then, it holds that

\[
\mathcal{G}_{c,k}(T) = \left\| \mu^{X_L}(\cdot) - \mu^{Z^q_L}(\cdot) \right\|_{r_T}.
\]  
(10)

An important remark from Proposition 1 is that it allows to use any kind of correlation between the \( X, Z^q \). For this reason we assume that \( (X, Z^q) \) is distributed as in \( \nu_{c,q}^T \). We are going to specify this distribution soon. First we get the following result.

**Proposition 3** Let a fixed \( \alpha > 0 \) and \( \delta, \gamma \leq \min\{\alpha/2, 1/10\} \). Then for \( k = (1 + \alpha)\Delta/\ln \Delta \), it hold that

\[
\mathbb{E} \left[ \mathcal{G}_{c,k}(\mathcal{T}_\xi^h) \middle| \mathcal{T}_\xi^h \in \mathcal{A}_{h,\zeta} \right] \leq \frac{1}{\Pr \left[ \mathcal{T}_\xi^h \in \mathcal{A}_{h,\zeta} \right]} \left( 2 \exp \left( -\frac{1}{8}(\Delta_+)^{h/4-1/2+\frac{1}{8}+\frac{\alpha}{1+\alpha}} \right) \cdot \mathbb{E} \left[ L_h \left( \mathcal{T}_\xi^h \right) \right] + 2(\Delta_+)^{-\gamma(3/4-\zeta)h} \cdot \mathbb{E}[H(X_L, Z^q_L)] \right)
\]

On the r.h.s., the rightmost expectation term is w.r.t. both the joint distribution of \( X, Z^q \) and the distribution over the tree \( \mathcal{T}_\xi^h \). The rest expectations are w.r.t. the distributions over trees only, i.e. \( \mathcal{T}_\xi^h \).

The proof of Proposition 3 appears in Section 7.

It is direct to see that

\[
\mathbb{E} \left[ L_h \left( \mathcal{T}_\xi^h \right) \right] = (d_\zeta)^h.
\]  
(11)

The theorem will follow by bounding appropriately the quantity \( \mathbb{E}[H(X_L, Z^q_L)] \). For this we need to specify a coupling between the random variables \( X \) and \( Z^q \) which minimizes their expected Hamming distance. Observe that the expected hamming distance is both w.r.t. the coupling and the randomness of the trees.

The coupling \( X \) and \( Z^q \) is inductively defined and it is as follows: We colour the vertices from the root down to the leaves. For a vertex \( v \) whose father \( w \) is such that \( X(w) = Z^q(w) \) we couple \( X(v) \) and \( Z^q(v) \) identically, i.e. \( X(v) = Z^q(v) \). On the other hand, when \( X(w) \neq Z^q(w) \) we set \( X(v) = Z^q(v) \)
unless \(X(v) = Z^q(w)\) then we set \(Z^q(v) = X(w)\). In this coupling, the following holds: Let \(w\) be a vertex in the tree and let \(u\) one child of \(w\). It holds that

\[
\Pr[X(u) \neq Z^q(u) | X(w) \neq Z^q(w)] = \frac{1}{k}.
\]

In a tree \(T\), the expected number of children per vertex is \(d\). Then, it is elementary to show that for a disagreeing vertex, the expected number of disagreeing children is \(d/k \leq \ln \Delta + 1 + \alpha\), since \(\Delta > d\). From standard arguments about branching processes we get the following:

\[
\mathbb{E}[H(X_L, Y_L)] \leq \left(\frac{\ln \Delta}{1 + \alpha}\right)^h.
\]

Observe that the above expectation is w.r.t. both tree instances and random colourings. The theorem follows be combining the above inequality and \((11)\) with in Proposition 3.

### 7 Proof of Proposition 3

The previous setting allows to use ideas based on the notion of biasing-unbiasing boundary (introduced in [2]) to prove Proposition 3. Many of the results in this proof are extensions of results from [2] set to fit in our new context.

**Definition 9 (Biasing Boundary)** Consider \(k = (1 + \alpha)\frac{\Delta^+}{\ln \Delta^+}\) where \(\alpha > 0\) a fixed positive. Let \(\gamma = \min\{\alpha/2, 1/10\}\). For a fixed tree \(T\) of height \(h\), we let \(\sigma, X\) be a fixed \(k\)-colouring and a random \(k\)-colouring of \(T\), respectively. For \(L = L_h(T)\), we say that \(\sigma\) biases \(r(T)\) if there is \(c \in [k]\) such that

\[
\Pr[X_{r(T)} = c | X_L = \sigma_L] \geq (\Delta^+)^{-\gamma}.
\]

Also, we let \(\mathcal{U}(T)\) denote the set of all boundary conditions which are not biasing.

The following useful lemma provides a sufficient condition for a boundary not to be biasing.

**Lemma 1** Let \(k = (1 + \alpha)\Delta^+/\ln \Delta^+\), for fixed \(\alpha > 0\) and let \(H\) be a tree of height \(t > 0\) which has mixing root. Let \(v_1, \ldots, v_s\) be the vertices children of the root of \(H\) and we let the set \(S \subseteq \{\tilde{H}_{v_1}, \tilde{H}_{v_2}, \ldots, \tilde{H}_{v_s}\}\) contain only the subtrees whose roots are mixing. A proper \(k\)-colouring of \(T\) \(\sigma\) does not bias the root if the following holds: There are at most \(\Delta^\delta\) many subtrees \(\tilde{H}_{v_i} \in S\) such that \(\sigma(L_{t-1}(\tilde{H}_{v_i}))\) biases the root \(r(\tilde{H}_{v_i})\). Then \(\sigma(L_{t}(H))\) does not bias \(r(H)\).

The proof of Lemma 1 appears in Section 10.1.

**Definition 10** We let \(T\) be a tree of height \(h\) and \(L = L_h(T)\). For every vertex \(w \in L\) we denote the set of boundaries \(\mathcal{U}_w \subseteq [k]^L\) as follows: Let \(\mathcal{P}\) denote the path that connects \(r_T\) and \(w\) and we let

\[
\mathcal{M} = \left\{v \in \mathcal{P} : \text{dist}(r_T, v) \leq \frac{3}{4}h, \tilde{T}_v \text{ has mixing root}\right\}.
\]

Then \(\mathcal{U}_w\) contains the boundaries on \(L\) which does not bias any of the subtrees \(\tilde{T}_v\) where \(v \in \mathcal{M}\), that is

\[
\mathcal{U}_w = \left\{\sigma_L \subseteq \mathcal{U}(T) : \forall v \in \mathcal{M} \sigma_L \cap \tilde{T}_v \in \mathcal{U}(\tilde{T}_v)\right\}.
\]
Proposition 4 Let some fixed tree $T \in \mathcal{A}_{h,\xi}$ and let $L = L_h(T)$. Consider $\sigma, \tau$ to be two $k$-colourings of $T$ such that $H(\sigma_L, \tau_L) = 1$. Assume that both $\sigma_L, \tau_L \in \mathcal{U}_w$, where $w \in L$ and $\sigma(w) \neq \tau(w)$. Then it holds that
\[
||\mu^\sigma_L - \mu^\tau_L||_{r(T)} \leq \Delta^*_\xi,h = (2\Delta_+^{-\gamma})(3/4 - \zeta)h.
\]

The proof of Proposition 4 appears in Section 8.

Proposition 5 Let some fixed tree $T \in \mathcal{A}_{h,\xi}$. Let $X$ be a $k$-random colouring of $T$. For $k = (1 + \alpha)\Delta_+ / \ln \Delta_+$, where $\alpha > 0$ is fixed, and any $w \in L_h(T)$ it holds that
\[
\Pr [X_L \notin \mathcal{U}_w] \leq 2\exp \left( -\frac{1}{8}(\Delta_+^{\frac{\alpha}{2} - \zeta}) \right).
\]

The proof of Proposition 5 appears in Section 9.

Proof of Proposition 3: First, consider some fixed tree $T \in \mathcal{A}_{h,\xi}$ and we let $L = L_h(T)$. Consider, also, two colourings of the leaves $\sigma(L)$ and $\tau(L)$. We let $m$ be the Hamming distance between $\sigma(L)$ and $\tau(L)$, i.e. $m = H(\sigma_L, \tau_L)$. Let $v_1, \ldots, v_m$ be the vertices in $L$ for which $\sigma_L$ and $\tau_L$ disagree. Consider the sequence of boundary conditions $Z_0, \ldots, Z_m \in [k]^h$ such that $\sigma_L = Z_1$, $\tau_L = Z_m$, while the rest of the members are as follows: For $i \leq m$, we get $Z_i$ from $Z_{i-1}$ by substituting the assignment of $v_i$ from $\sigma(v_i)$ to “free”. Also, for $i \geq m$ we get $Z_{i+1}$ from $Z_i$ by substituting $Z(v_i-m)$ from “free” to $\tau(v_{i-m})$. It is direct that $H(Z_i, Z_{i+1}) = 1$.

It holds that
\[
||\mu^\sigma_L - \mu^\tau_L||_{r(T)} \leq \sum_{i=0}^{2m-1} ||\mu^{Z_i} - \mu^{Z_{i+1}}||_{r(T)}.
\] (12)

Also, it is not hard to see that for every $w \in L$ the following is true: if $\sigma_L \in \mathcal{U}_w$, then $Z_i \in \mathcal{U}_w$ for every $i = 1, \ldots, m$. Similarly, if $\tau_L \in \mathcal{U}_w$, then $Z_i \in \mathcal{U}_w$ for every $i = m, \ldots, 2m$.

Let the event $\mathcal{U}_\xi$ be that “$\sigma_L \notin \mathcal{U}_v \cup \tau_L \notin \mathcal{U}_v$”. Then it holds that
\[
||\mu^{Z_i} - \mu^{Z_{i+1}}||_{r(T)} \leq \mathbb{I}_{\mathcal{U}_v} + \left(1 - \mathbb{I}_{\mathcal{U}_v}\right) \Delta^*_\xi,h,
\] (13)

where $\Delta^*_\xi,h$ is defined in the statement of Proposition 4. In words, the above inequality states the following: if at least one of the $\sigma_L, \tau_L$ are not in $\mathcal{U}_v$, then the l.h.s. of (13) is at most 1. On the other hand, if both $\sigma_L, \tau_L \in \mathcal{U}_v$, then the total variation distance on the l.h.s. can be upper bounded using Proposition 4.

Plugging (13) into (12) we have that
\[
||\mu^\sigma_L - \mu^\tau_L||_{r(T)} \leq 2 \cdot \sum_{w \in L_h(T)} \mathbb{I}_{\sigma(w) \neq \tau(w)} \cdot \left[\mathbb{I}_{\mathcal{U}_v} + \left(1 - \mathbb{I}_{\mathcal{U}_v}\right) \Delta^*_\xi,h\right].
\] (14)

Now, we consider the quantity $G_{c,k}(T)$, i.e. $G_{c,k}(T) = ||\mu^{X_L} - \mu^{Z_L}||_{r(T)}$. For bounding $G_{c,k}(T)$ we are going to use (14). That is

4Free boundary on some set A, is the lack of vertices with fixed colour assignment.
\[ G_{c,k}(T) = ||\mu^{L} - \mu^{Z^q_{L}}||_{r(T)} \]
\[ \leq \sum_{\sigma_L, \tau_L \in [k]^L} \Pr \left[ X_L = \sigma_L, Z^q_{L} = \tau_L \right] \cdot ||\mu^{\sigma_L} - \mu^{\tau_L}||_{r(T)} \]
\[ \leq 2 \cdot \sum_{\sigma_L, \tau_L \in [k]^L} \Pr \left[ X_L = \sigma_L, Z^q_{L} = \tau_L \right] \cdot \sum_{v \in L_h(T)} \left( 1 - I \left( \sigma_v \neq r_v \right) \cdot \left( I \left( \sigma_v = r_v \right) + \left( 1 - I \left( \sigma_v = r_v \right) \right) \Delta_{z,h}^* \right) \right) \quad \text{[from (13)]} \]

Due to symmetry it holds that \( \Pr \left[ X(L) \notin \mathcal{U}_v \right] = \Pr \left[ Z^q(L) \notin \mathcal{U}_v \right] \). Using this observation and a union bound, the above inequality implies that
\[ G_{c,k}(T) \leq 4 \sum_{v \in L} \Pr \left[ X(L) \notin \mathcal{U}_w \right] + \Delta_{z,h}^* \cdot \sum_{v \in L} \Pr \left[ X(v) \neq Z^q(v) \right] \]
\[ \leq 2 \exp \left( -\frac{1}{8} \left( \Delta_f + \frac{h/4 - 1}{2} \cdot \frac{\delta - 1}{8 + \frac{1}{1 + 0}} \right) \cdot |L_h(T)| \right) + 2 \Delta_{z,h}^* \cdot \mathbb{E}_{\nu_{c,q}}[H(X_L, Z^q_{L})] \]

where in the last inequality we used Proposition 5 to bound \( \Pr \left[ X(L) \notin \mathcal{U}_w \right] \cdot \mathbb{E}_{\nu_{c,q}}[H(X(L), Z^q(L))] \) is the expected Hamming distance between \( X_L \) and \( Z^q_{L} \) and depends only on the joint distribution of \( X, Z^q \), which is denoted as \( \nu_{c,q} \).

The proposition follows by averaging over \( T_h^* \), conditional that we have a tree in \( A_{h,\xi} \), that is
\[ \mathbb{E} \left[ G_{c,k} \left( T_h^* \right) \left| T_h^* \in A_{h,\xi} \right. \right] \leq \frac{1}{\Pr \left[ T_h^* \in A_{h,\xi} \right]} \left( 2 \exp \left( -\frac{1}{8} \left( \Delta_f + \frac{h/4 - 1}{2} \cdot \frac{\delta - 1}{8 + \frac{1}{1 + 0}} \right) \right) \cdot \mathbb{E} \left[ |L_h(T_h^*)| \right] + 2(2\Delta_{z,h}^*)^{(3/4 - \epsilon)h} \cdot \mathbb{E}[H(X_L, Z^q_{L})] \right) \]

The rightmost expectation term is w.r.t. both \( \nu_{c,q} \) and the distribution of random trees \( T_h^* \). In the above derivations we used the following, easy to derive inequality.
\[ \mathbb{E} \left[ f \left( T_h^* \right) \left| T_h^* \in A_{h,\xi} \right. \right] \leq \frac{\mathbb{E} \left[ f \left( T_h^* \right) \right]}{\Pr \left[ T_h^* \in A_{h,\xi} \right]} \]

where \( f \) is any non-negative functions on instances of the distribution \( T_h^* \). The proposition follows. \[ \square \]

### 8 Proof of Proposition 4

**Proof:** For this we use coupling. The same ideas was used in [9][10]. The idea works as follows. Observe that we have disagreement only on exactly one vertex. Also, observe that the underlying structure of the graph is tree. So as to bound \( ||\mu^{\sigma_L} - \mu^{\tau_L}||_{r(T)} \) we take two \( k \)-colourings of \( T, X \) and \( Y \) distributed as in \( \mu^{\sigma_L}, \mu^{\tau_L} \) respectively. We are going to couple \( X, Y \). Then clearly it holds that
\[ ||\mu^{\sigma_L} - \mu^{\tau_L}||_{r(T)} \leq \Pr[X(r_T) \neq Y(r_T)]. \quad (15) \]
The coupling of the two random variables is done in a step-wise fashion moving away from the disagreeing vertex \( w \). In particular what is of our interest is the vertices on the path \( P \) that connects \( w \) with \( r_T \), i.e. \( P = v_0, v_1, \ldots v_h \) where \( v_0 = w \) and \( v_h = r_T \).

We couple \( X, Y \) by considering the pairs \((X(v_i), Y(v_i))\), for \( i = 1, \ldots, h \). Observe that if for some \( j \) we have that \( X(v_j) = Y(v_j) \), then we can couple the remaining vertices identically, i.e. for every \( i > j \) we have \( X(v_i) = Y(v_i) \). Clearly this holds due to the fact that we deal with a tree and once we have \( X(v_j) = Y(v_j) \) there is no alternative path for the disagreement to propagate to the pairs \( X(v_i), Y(v_i) \) where \( i > j \).

On the other hand, consider the case that \( X(v_i) \neq Y(v_j) \), for some \( 0 \leq j < h \). We need to bound the probability that \( X(v_{j+1}) \neq Y(v_{j+1}) \) in the coupling. For this we consider two cases, depending on whether the tree \( \tilde{T}_{v_{j+1}} \) has a mixing root or not. Then we have the following:

\[
\Pr[X(v_{j+1}) \neq Y(v_{j+1})] \leq \begin{cases} 
(2\Delta^\gamma) & \text{if } \tilde{T}_{v_{j+1}} \text{ has mixing root} \\
1 & \text{otherwise}
\end{cases}
\]

Observe that once we show that indeed the above bounds holds, then it is a matter of very simple inductive argument to see that the proposition holds.

Thus, it remains to show the bound above. In particular, it suffices to show the bound regarding the case where the \( \tilde{T}_{v_{j+1}} \) has mixing root, the other one is trivial. For this case assume that \( X(v_j) = c, Y(v_j) = q \) for two different \( c, q \in [k] \). In this situation we have disagreement between \( X(v_{j+1}), Y(v_{j+1}) \) if either \( X(v_{j+1}) = q \) or \( Y(v_{j+1}) = c \) or both. Otherwise we have \( X(v_{j+1}) = Y(v_{j+1}) \). Then it becomes apparent that

\[
\Pr[X(v_{j+1}) \neq Y(v_{j+1}) | X(v_j) = c, Y(v_j) = q] \leq \max \{\Pr[X(v_{j+1}) = q | X(v_j) = c], \Pr[Y(v_{j+1}) = c | Y(v_j) = q]\}
\]

The the result follows almost directly. W.l.o.g. consider the term \( \Pr[X(v_{j+1}) = q | X(v_j) = c] \). Clearly there is a \( c' \in [k] \) such that

\[
\Pr[X(v_{j+1}) = q | X(v_j) = c] \leq \Pr[X(v_{j+1}) = q | X(v_j) = c, X(v_{j+2}) = c'].
\]

The above holds because \( \Pr[X(v_{j+1}) = q | X(v_j) = c] \) can be written as a convex combination of boundaries on \( v_{j+2} \).

The subtree we deal with has mixing root and the boundary biases at most \( \Delta^\delta + 2 \) neighbours. It is a matter of elementary calculations to verify that \( \Pr[X(v_{j+1}) = q | X(v_j) = c, X(v_{j+2}) = c'] \leq 2\Delta^\gamma \). The proposition follows.

\[\square\]

9 Proof of Proposition 5

So as to show Proposition 5 we use the following result.

**Proposition 6** Let \( k = (1 + \alpha)\Delta/\ln \Delta \), for fixed \( \alpha > 0 \). Consider some tree \( H \), of height \( t > 0 \), which has mixing root. For \( Z \), a random \( k \)-colouring of \( H \), the following is true

\[
\Pr[Z_{L_h(H)} \notin \mathcal{U}(H)] \leq \exp\left(-\frac{1}{8}(\Delta^\delta + 2) + \frac{\alpha}{4\Delta^\gamma}\right),
\]

(16)

where \( \delta > 0 \) comes from Definition 7.
The proof of Proposition 6 appears in Section 10.

**Proof of Proposition 6:** The proposition follows by using Proposition 8 and a simple union bound. In particular, let $L = L_h(T)$. Also, let $P$ denote the path that connects $r_T$ and $w \in L_h(T)$ while

$$M = \left\{ v \in P : \text{dist}(r_T, v) \leq \frac{3}{4} h, \ T_v \text{ has mixing root} \right\}.$$  

Clearly, $X_L \notin U_w$ if for some vertex $u \in M$, it holds that $X(L \cap \tilde{T}_u) \notin U(\tilde{T}_u)$, i.e., the boundary $X(L \cap \tilde{T}_u)$ biases the root of the subtree $\tilde{T}_u$. That is,

$$\Pr[X(L_h(T)) \notin U_P] = \Pr \left[ \bigcup_{u \in M} X_{L \cap \tilde{T}_u} \notin U(\tilde{T}_u) \right] \leq \sum_{u \in M} \Pr \left[ X_{L \cap \tilde{T}_u} \notin U(\tilde{T}_u) \right] \quad \text{[union bound]}$$

$$\leq \sum_{t=(1/4)h}^h \exp \left( -\frac{1}{8} (\Delta_+)^{t-\frac{\alpha}{2}} - \frac{7}{8} \frac{\Delta_+}{t+\alpha} \right) \leq 2 \exp \left( -\frac{1}{8} (\Delta_+)^{\frac{h/4-1-\delta+\frac{7}{8} \delta}{t+\alpha}} \right),$$

in the last line, above, we used Proposition 6. The proposition follows. \qed

**10 Proof of Proposition 6**

Since we assumed that the tree $H$ has a mixing root, it holds that $\text{deg}(r_H) = s \leq \Delta$. We let $v_1, v_2, \ldots, v_s$ denote the children of $r_H$. We remind the reader that the set $S \subseteq \{ \tilde{H}_{v_1}, \tilde{H}_{v_2}, \ldots, \tilde{H}_{v_s} \}$ contain only the subtrees whose roots are mixing.

So as to prove Proposition 6 we need the following result.

**Lemma 2** Let $X$ be a random $k$-colouring $H$. For $L_i = L_{h-1}(\tilde{H}_v)$, let $B_i$ denote the event that in $H_v$, the boundary $X(L_i)$ does not bias $r(H_v)$. For any $\Gamma \subseteq \{ 1, \ldots, s \}$ it holds that

$$\Pr[\bigcap_{i \in \Gamma} B_i] = \prod_{i \in \Gamma} \Pr[B_i] = (\Pr[B_i])^{|\Gamma|}.$$  

The proof of this lemma is straightforward. The interested reader may find one in [2], i.e., this follows from Lemma 3.5 and Lemma 3.6 there. Thus, we omit the proof.

**Proof of Proposition 6:** The proof is by induction on $t \geq 1$. Consider the case where $t = 1$. Then, $H$ is one level tree whose root is of degree at most $\Delta$. Let $Y$ denote the number of colours that do not appear in $X(L_1)$. It holds that

$$\Pr[X_{L_1(H)} \notin U(H)] \leq \Pr[Y \leq \Delta_+] \quad \text{(17)}.$$  

Observe that $\Pr[Y \leq \Delta]$ is an increasing function of the degree of $r(H)$. That is, the larger the degree of $r(H)$ the more colours are expected to be used to colour the leaves of $H$. For this reason, we are going to upper bound the r.h.s. of (17) by assuming that $\text{deg}(r_H) = \Delta$. It holds that

$$E[Y] = (k-1) \left( 1 - \frac{1}{k-1} \right)^{\Delta_+}$$

$$\geq (k-1) \exp \left( -\frac{\Delta_+}{k-2} \right) \quad \text{[as $1 - x \geq e^{-x}$ for $0 < x < 1/5$]}$$

$$\geq (k-1) \exp \left( -\left( 1 - \frac{\alpha}{1+\alpha} \right) \ln \Delta_+ - \frac{\ln \Delta_+}{k-2} \right) \geq (\Delta_+)^{\frac{7}{8} \frac{\alpha}{1+\alpha}}. \quad \text{(18)}$$
Viewing the $k - 1$ colour which are available for the leaves of $H$ as bins and each leaf of $H$ as a ball which is thrown to a random bin, $Y$ corresponds to the number of empty bins. It is a standard result that we can apply Chernoff bounds for bounding the tails of $Y$, e.g. see [27]. Then we get that

$$
\Pr[Y < (\Delta_+)^\gamma] \leq \Pr[Y \leq \frac{1}{2} \mathbb{E}[Y]] \leq \exp\left(-\frac{\mathbb{E}[Y]}{8}\right) \leq \exp\left(-\frac{(\Delta_+)^7}{8} \frac{\alpha}{\gamma}\right), \quad \text{[since } \gamma \leq \min\{\alpha/2, 1/10\}\text{]}
$$

where in the last inequality we use (18). We have proved the basis of our induction.

Assume, now, that (16) is true for every tree of height $t - 1$ which has mixing root. It suffices to show that (16) is true for a tree $H$ of height $t$ with a mixing root. For such a tree $H$ let $L = L_t(H)$. Consider also a random $k$-colouring $X$ for this tree. Let $Z$, denote the number of subtrees in $\mathbb{S}$ which are biased under the random colouring $X$, i.e. the number of trees $H_{v_i} \in \mathbb{S}$ such that $X(L \cap H_{v_i})$ is biasing for $r(H_{v_i})$. From Lemma [1] we have the following

$$
\Pr[X_L \notin U(H)] \leq \Pr\left[Z > \Delta_+^\delta\right]. \quad (19)
$$

Let

$$
\varrho = \max_{H_v \in \mathbb{S}} \left\{ \Pr[X(L \cap H_v) \notin U(H_v)] \right\},
$$

where for the subtree $H_v$, the set $U(H_v)$ contains all the boundary conditions (at level $t - 1$) $H_v$ which do not bias the root of $r(H_v)$. From Lemma [2] we conclude that $Z$ is dominated by $B(\Delta_+,\varrho)$, i.e. the binomial distribution with parameters $\Delta_+$ and $\varrho$. Observe that due to our assumptions it holds that $\Delta_+^\delta \gg \Delta_+ \cdot \varrho$. We have that

$$
\Pr\left[Z > \Delta_+^\delta\right] \leq \sum_{j=\Delta_+^\delta}^{\Delta_+} \binom{\Delta_+}{j} \varrho^j (1 - \varrho)^{\Delta_+ - j} \leq \Delta_+ \left(\Delta_+^\delta \varrho\right)^{\Delta_+} \varrho^{\Delta_+ - \Delta_+^\delta} (1 - \varrho)^{\Delta_+ - \Delta_+^\delta}
$$

$$
\leq \frac{\Delta_+}{(\Delta_+^\delta/e)} \left(\Delta_+ \varrho\right)^{\Delta_+^\delta} \left[\text{as } \binom{n}{k} \leq (ne/i)^i \right]
$$

$$
\leq (\Delta_+ \varrho)^{\Delta_+^\delta} \left[\text{as } \frac{\Delta_+}{(\Delta_+^\delta/e)} < 1 \right]
$$

$$
\leq \left(\Delta_+ \exp\left(-\frac{1}{8} \Delta_+^7 \frac{\alpha}{\gamma} \frac{\alpha}{\gamma}\right)\right)^{\Delta_+^\delta} \left[\text{by the induction hypothesis}\right]
$$

$$
\leq \left(\exp\left(-\frac{1}{8} \Delta_+^7 \frac{\alpha}{\gamma} \frac{\alpha}{\gamma}\right)\right)^{\Delta_+^\delta} \leq \exp\left(-\frac{1}{8} \Delta_+^7 \frac{\alpha}{\gamma} \frac{\alpha}{\gamma}\right). \quad (20)
$$

The proposition follows by plugging (20) into (19). \hfill \Box

10.1 Proof of Lemma [1]

Assume that $\text{deg}(r_H) = s$ for some integer $s$. Clearly $s \leq \Delta_+$ since we assume that $H$ has a mixing root. Let $L = L_t(H)$. Let $v_1, \ldots, v_s$ denote the children of the root. Let $L_i$ denote the vertices at level $t - 1$ of the subtree $H_{v_i}$. That is $L_i = L \cap H_{v_i}$.

Let $X$ be a random $k$-colouring of $H$ such that $X_L = \sigma_L$ also, for $i = 1, \ldots, s$, let $X_i$ denote a random colouring of $H_{v_i}$ such that $X_i(L_i) = \sigma_{L_i}$. A simple recursive argument yields the following relation: For any $c \in [k]$ it holds that

$$
\Pr[X(r_H) = c] = \frac{\prod_{i=1}^s \Pr[X_i(v_i) \neq c]}{\sum_{c' \in [k]} \prod_{i=1}^s \Pr[X_i(v_i) \neq c']} \leq \frac{1}{\sum_{c' \in [k]} \prod_{i=1}^s \Pr[X_i(v_i) \neq c']}. \quad (21)
$$
We are going to show that \( r(H) \) is not biased by \( \sigma_{L_h} \) by showing that the denominator in \(|21|\) is sufficiently small.

Let \( B \subset [k] \) denote the set of colours \( c \) for which there is some \( i \) such that \( \Pr[X_i(v_i) = c] \geq \Delta_+^{\gamma} \). Clearly, only \( \Delta_+^{\gamma} \) many colours can have increased bias at the root of \( H \), since \( \sum_{c \in [k]} \Pr[X_i(v_i) = c] = 1 \).

Also, by our assumptions, there are at most \( \Delta_+^4 \) trees \( \tilde{H}_v \) whose root is biased but the boundary biases the colour assignment of the root. Furthermore, there are \( \Delta_+^4 \) trees \( \tilde{H}_v \) with non-mixing roots. For the trees with the non-mixing root we assume the root is biased by the boundary conditions. That is, there can be at most \( 2\Delta_+^{4} \) trees \( \tilde{H}_v \) whose roots are biased.

Clearly, all the above imply that \( |B| \leq 2\Delta_+^{\gamma+\delta} \). Let \( U = [k] \setminus B \). Then we can rewrite \(|21|\) as follows:

\[
\Pr[X(r_H) = c] \leq \left( \sum_{c' \in U} \prod_{i=1}^s (1 - \Pr[X_i(v_i) = c']) \right)^{-1}
\leq \left( \sum_{c' \in U} \prod_{i=1}^s \exp \left( - \frac{\Pr[X_i = c']}{1 - \Pr[X_i = c']} \right) \right)^{-1}
\leq \left( |U| \sum_{c' \in U} \frac{1}{|U|} \prod_{i=1}^s \exp \left( - \frac{\Pr[X_i(v_i) = c']}{1 - \Pr[X_i(v_i) = c']} \right) \right)^{-1}
\leq \left( \left| U \right| \prod_{c' \in U} \exp \left( - \frac{1}{|U|} \sum_{i=1}^s \frac{\Pr[X_i(v_i) = c']}{1 - \Pr[X_i(v_i) = c']} \right) \right)^{-1}
\leq \left( \left| U \right| \prod_{i=1}^s \frac{\Pr[X_i(v_i) \in U]}{1 - \Delta_+^{\gamma}} \right)^{-1}
\leq \left( \left| U \right| \exp \left( - \frac{1}{1 - \Delta_+^{\gamma}} \frac{s}{|U|} \right) \right)^{-1}
\]

[by the arithmetic-geometric mean relation]

\[\Pr[X(v_i) = c] < \Delta_+^{\gamma} \text{ for } c \in U\]

It is straightforward to show that \( |U| \geq \left[ 1 - \Delta_+^{\gamma+\delta-1} \right] \geq (1 + \frac{\alpha}{10}) \Delta_+^{\frac{\Delta_+^{\gamma}}{\ln \Delta_+}}, \) since \( \gamma + \delta < 1 \). Also it holds that \( \frac{1}{1 - \Delta_+^{\gamma}} \frac{s}{|U|} \leq \frac{\ln \Delta_+}{1 + 4\alpha^2/\gamma} \), since \( s \leq \Delta_+ \). Thus, we get that

\[\Pr[X = c] \leq \frac{1}{(1 + \alpha/2) \Delta_+^{\frac{\Delta_+^{\gamma}}{\ln \Delta_+} \Delta_+^{\frac{\Delta_+^{\gamma}}{1 + 4\alpha^2/\gamma}}} \leq \Delta_+^{-\frac{2\alpha}{1 + 4\alpha^2/\gamma}} < \Delta_+^{-\gamma},\]

as \( \gamma \leq \min\{\alpha/2, 1/10\} \). The lemma follows.

### 11 Proof of Proposition 2

For \( i = (1-h)h \) we let \( Q_{h,i} = \Pr[T_h^h \notin A_{h,i}] \). Also, we let \( Q_{h,i} = \Pr[T_h^h \notin A_{h,i}] \) deg(r(T_h^h)) = t

Using a simple union bound we get the following: For \( t \leq (\Delta_+)^\delta \) it holds that

\[Q_{h,i} \leq t \cdot Q_{h-1,i-1} \tag{22}\]

For \( (\Delta_+)^\delta \leq t \leq (\Delta_+) \), we use the following lemma
Lemma 3 For \((\Delta_+)^{\delta} < t \leq \Delta_+\), it holds that
\[
Q_{h,i}^{t} \leq 2t \left( Q_{h-1,i-1} + Q_{h-1,i} \cdot \Pr \left[ B(\Delta_+, q) \geq (\Delta_+)^{\delta} \right] \right).
\]
The proof of Lemma 3 appears in Section 11.1

Finally, using a simple union bound we get that for \(t > \Delta_+\) it holds that
\[
Q_{h,i}^{t} \leq t \cdot Q_{h-1,i}.
\]
We use the bounds for \(Q_{h,i}^{t}\) from above, to bound \(Q_{h,i}\) by working as follows:
\[
Q_{h,i} = \sum_{t=0}^{n} Q_{h,i}^{t} \xi_t
= Q_{h-1,i-1} \cdot \sum_{t=0}^{\Delta_+^{\delta}} t \cdot \xi_t + 2Q_{h-1,i-1} \cdot \sum_{t=(\Delta_+)^{\delta}+1}^{\Delta_+} t \cdot \xi_t + 2Q_{h-1,i} \cdot \Pr \left[ B(\Delta_+, q) \geq (\Delta_+)^{\delta} \right] \cdot \sum_{t=(\Delta_+)^{\delta}+1}^{\Delta_+} t \cdot \xi_t + Q_{h-1,i} \cdot \sum_{t=(\Delta_+)^{\delta}+1}^{\Delta_+} t \cdot \xi_t
\leq 2Q_{h-1,i-1} \sum_{t=0}^{\Delta_+^{\delta}} t \cdot \xi_t + Q_{h-1,i} \left( 2 \Pr \left[ B(\Delta_+, q) \geq (\Delta_+)^{\delta} \right] \sum_{t=(\Delta_+)^{\delta}+1}^{\Delta_+} t \cdot \xi_t + \sum_{t=(\Delta_+)^{\delta}+1}^{\Delta_+} t \cdot \xi_t \right)
\leq 2d \cdot Q_{h-1,i-1} + Q_{h-1,i} \left( 2d \cdot \Pr \left[ B(\Delta_+, q) \geq (\Delta_+)^{\delta} \right] + \sum_{t=(\Delta_+)^{\delta}+1}^{\Delta_+} t \cdot \xi_t \right).
\]

The following lemma uses (24) to derive an upper bound on \(Q_{h,i}\).

Lemma 4 Let \(\lambda \in (0, 1)\) and \(\theta' > 1\) be a fixed numbers such that \(\lambda \theta' < 1\) and \(\beta(1 - \theta') < -1\). Then for \(Q_{h,i}\) that satisfy (24) where \(i\) is such that \(i = \lambda h\), it holds that
\[
Q_{i} \leq \exp \left[ -(1 - \lambda \theta') C \cdot h \right].
\]
The quantities \(\beta, C\) are defined in the statement of Proposition 2.

The proof of Lemma 4 appears in Section 11.2

The proposition follows by using the above lemma and set \(\lambda = (1 - \zeta)\) and \(\theta' = \theta\).

11.1 Proof of Lemma 3

We let \(Q_{h,i}^{M} = \Pr \left[ T_{\xi, h}^{\lambda} \notin A_{h, \zeta} \right] \cdot \Pr \left[ r \left( T_{\xi, h}^{\lambda} \right) \text{ is mixing} \right]\) and \(Q_{h,i}^{N} = \Pr \left[ T_{\xi, h}^{\lambda} \notin A_{h, \zeta} \right] \cdot \Pr \left[ T_{\xi, h}^{\lambda} \text{ is not mixing} \right]\).

We remind the reader that it holds that \((\Delta_+)^{\delta} < t \leq \Delta_+\). We denote with \(q_h\) the probability for a \(T_{\xi, h}^{\lambda}\) to not have a mixing root. It is direct to show that for any integer \(h > 0\) it holds that \(q_h \leq q\), where \(q\) is from Definition 5.

Let \(Q\) denote the number of vertices \(u\) the children of \(r \left( T_{\xi}^{h} \right)\) such that \(T_{\xi, u}\) does not have a mixing root. It is direct that \(T_{\xi, u}\) is distributed as \(T_{\xi}^{h-1}\). Conditioning that the degree of \(r \left( T_{\xi}^{h} \right)\) is \(t\), \(Q\) is binomially distributed with parameters \(t\) and \(q_{h-1}\), i.e. \(B(t, q_{h-1})\). It holds that
\[
Q_{h,i}^{t} \leq \sum_{j=0}^{(\Delta_+)^{\delta}} \binom{t}{j} q_{h-1}^{j} (1 - q_{h-1})^{t-j} \left( (t-j) Q_{h-1,i-1}^{M} + j Q_{h-1,i-1}^{N} \right)
+ \sum_{j=(\Delta_+)^{\delta}+1}^{t} \binom{t}{j} q_{h-1}^{j} (1 - q_{h-1})^{t-j} \left( (t-j) Q_{h-1,i}^{M} + j Q_{h-1,i}^{N} \right).
\]
Using the standard equality that \((t - j)\binom{t}{j} = t^{t-1} - j\), we get that

\[
Q_{h,i}^t \leq t(1 - q)Q_{h-1,i-1}^{t-1} \sum_{j=0}^{(\Delta_+)^t} \binom{t-1}{j} q_h^j (1 - q_{h-1})^{t-1-j} + t q Q_{h-1,i-1}^{t-1} \sum_{j=1}^{(\Delta_+)^t} \binom{t-1}{j-1} q_{h-1}^{j-1} (1 - q_{h-1})^{t-1-j} + t(1 - q_h)Q_{h-1,i}^t \sum_{j=(\Delta_+)^t+1}^{t-1} \binom{t-1}{j} q_h^j (1 - q_{h-1})^{t-1-j} + t q_{h-1} Q_{h-1,i}^t \sum_{j=(\Delta_+)^t+1}^{t} \binom{t-1}{j-1} q_{h-1}^{j-1} (1 - q_{h-1})^{t-1-j},
\]

It is not hard to see that for any \(h, i\) it holds that \(q_h Q_{h,i}^t \leq Q_{h,i}^t\) and \((1 - q_h)Q_{h,i}^t \leq Q_{h,i}^t\). Using these two inequalities we get that

\[
Q_{h,i}^t \leq tQ_{h-1,i-1} \left[ \Pr[B(t - 1, q_{h-1}) \leq (\Delta_+)^t] + \Pr[B(t - 1, q_{h-1}) \leq (\Delta_+)^t - 1] \right] + tQ_{h-1,i} \left[ \Pr[B(t - 1, q_{h-1}) \geq (\Delta_+)^t + 1] + \Pr[B(t - 1, q_{h-1}) \geq (\Delta_+)^t] \right] \leq 2tQ_{h-1,i-1} + 2tQ_{h-1,i} \Pr[B(t - 1, q_{h-1}) \geq (\Delta_+)^t].
\]

Note that that \(\Pr[B(t - 1, q_{h-1}) \geq (\Delta_+)^t]\) is increasing with \(t\). That is, for \(t \leq \Delta_+\) it holds that

\[
\Pr[B(t - 1, q_{h-1}) \geq (\Delta_+)^t] \leq \Pr[B(\Delta_+, q_{h-1}) \geq (\Delta_+)^t] \leq \Pr[B(\Delta_+, q) \geq (\Delta_+)^t],
\]

where the last inequality follows from the fact that \(B(\Delta_+, q_{h-1})\) is stochastically dominated by \(B(\Delta_+, q)\), since, \(q_{h-1} \leq q\), for any \(h\). The lemma follows.

### 11.2 Proof of Lemma 4

We are going to use induction to prove the lemma. First we are going to show that if (25) is true for some \(h > 1\) then it is also true for \(h + 1\). Let \(\lambda = \frac{1}{h}, \lambda^- = \frac{1}{h - 1}\) and \(\lambda^+ = \frac{1}{h + 1}\). We rewrite (24) in terms of \(\lambda, \lambda^+\) and \(\lambda^-\) as follows:

\[
Q_{\{h, \lambda h\}} \leq 2d \cdot Q_{\{h-1, \lambda^-(h-1)\}} + Q_{\{h-1, \lambda^+(h-1)\}} \left[ 2d \Pr[B(\Delta_+, q) \geq (\Delta_+)^t] + \sum_{t \geq (\Delta_+)^t+1} t \cdot \xi_t \right] . \tag{26}
\]

Using the induction hypothesis and noting that \(\lambda^- = \lambda - \frac{1 - \lambda}{h - 1}\) we have that

\[
Q_{\{h-1, \lambda^-(h-1)\}} \leq \exp \left[ -(1 - \theta \cdot \lambda^-) C(h - 1) \right] \leq \exp \left[ \frac{-1 - \theta'}{(1 - \theta') \left( \lambda - \frac{1 - \lambda}{h - 1} \right)} C(h - 1) \right] \leq \exp \left[ -(1 - \theta' \cdot \lambda^-) C(h - 1) \right] \cdot \exp \left[ -(1 - \theta') \cdot C(h - 1) \right] \leq \exp \left[ -\theta' \cdot \lambda^- C(h - 1) \right] \cdot \exp \left[ (1 - \theta') C \right].
\]
As far as \( Q_{h-1,i} \) is regarded, we use the fact that \( \lambda^+ = \lambda + \frac{\lambda}{h-1} \) and we get that
\[
Q_{\{h-1, \lambda^+: (h-1)\}} \leq \exp \left[-(1 - \theta' \cdot \lambda^+) \cdot C(h-1) \right] \\
\leq \exp \left[- \left(1 - \theta' \cdot \lambda - \frac{\theta' \lambda}{h-1} \right) C(h-1) \right] \\
\leq \exp \left[- (1 - \theta' \cdot \lambda) C(h-1) \cdot \exp[\theta' \lambda C] \right] \\
\leq \exp \left[- (1 - \theta' \cdot \lambda) C \right] \exp[C].
\]

Substituting the bounds for \( Q_{\{h-1,i:-1\}}, Q_{\{h-1,i\}} \) above into (26) we get that
\[
Q_{\{h, \lambda h\}} \leq \exp \left[- (1 - \theta' \cdot \lambda) Ch \right] \times \\
\times \left(2d \cdot \exp \left[(1 - \theta') C \right] + \exp(C) \left(2d Pr \left[ \mathcal{B}(\Delta_+, q) \geq (\Delta_+)^\beta \right] + \sum_{t \geq (\Delta_+)+1} t \cdot \xi_t \right) \right).
\]

Due to our assumptions it is direct that it holds that
\[
2d \cdot \exp \left[(1 - \theta') C \right] = 2d^{1+\beta(1-\theta')} \leq \frac{1}{5},
\]
i.e. we have assumed that \( \beta(1 - \theta') < -1 \). Also due to our assumptions it is direct that the following is true
\[
\exp(C) \left(2d \Pr \left[ \mathcal{B}(\Delta_+, q) \geq (\Delta_+)^\beta \right] + \sum_{t \geq (\Delta_+)+1} t \cdot \xi_t \right) \leq \frac{2}{5}.
\]

Using the two bounds above (27) writes as follows:
\[
Q_{\{h, \lambda h\}} \leq \exp \left[- (1 - \theta' \cdot \lambda) Ch \right].
\]

It remains to show the base of induction, i.e to study the case \( h = 1 \). For any fixed \( \lambda \in (0, 1) \) and \( h = 1 \) it holds that
\[
Q_{\{h, \lambda h\}} \leq Pr[\text{deg}(r(T)) \geq \Delta_+] = \sum_{t \geq \Delta_+} \xi_t \leq \exp[-2C].
\]

The lemma follows.

12 \ Proof of Proposition

Given some \( \sigma = [k]^L \), we let the variable \( Y = Y(\sigma_L) \) be such that \( Y = \mu_L^r(c) - 1/k \). Let the colouring of the root \( \tau_r = c \). By definition, we have that
\[
\mathbb{E}_{\mu^r} [Y] = \sum_{\sigma_L \in [k]^L} \mu_L^r(\sigma_L) Y(\sigma_L) \\
= \sum_{\sigma_L \in [k]^L} \mu_L^r(\sigma_L)(\mu^\sigma_L(c) - 1/k) = \mu^X(L)(c) - 1/k.
\]

Also, we have that
\[
\mathbb{E}_{\mu^r} [Y] = \sum_{\sigma_L \in [k]^L} \frac{\mu_L^r(\sigma_L)}{\mu_L(\sigma_L)} \left(\mu^\sigma_L(c) - 1/k\right) \cdot \mu_L(\sigma_L) \\
= \sum_{\sigma_L \in [k]^L} \frac{\mu^\sigma_L(c)}{\mu^r(c)} \left(\mu^\sigma_L(c) - 1/k\right) \cdot \mu_L(\sigma_L).
\]
That is, in order to compute the expectation above we calculate the Randon-Nikodym derivative. The derivation in the second line is just an application of Bayes’ rule. Letting \( \frac{\mu_L^X(c)}{\mu^L(c)} = r(\sigma_L) \) and noting that \( \mu^r(c) = 1/k \), it is elementary to verify that

\[
 k \cdot Y(\sigma_L) + 1 = r(\sigma_L).
\]

Using the above equality we get that

\[
 E_{\mu^r} [Y] = k \sum_{\sigma_L \in [k]^L} (\mu_L^\sigma(c) - 1/k)^2 \mu(\sigma_L) + \sum_{\sigma_L \in [k]^L} (\mu_L^r(c) - 1/k) \mu(\sigma_L). \tag{27}
\]

It is direct to show that \( \sum_{\sigma_L \in [k]^L} (\mu_L^\sigma(c) - 1/k) \mu(\sigma_L) = 0 \). Thus, we get that

\[
 E_{\mu^r} [Y] = E[Y^2] = \mu^X_L(c) - 1/k. \tag{28}
\]

where the second expectation is w.r.t. unconditional Gibbs distribution. Observe that \( E_{\mu^r} [Y] \geq 0 \).

Using the above equality and Cauchy-Schwarz inequality we get the following:

\[
 \sum_{\sigma(L) \in [k]^L} \mu_L(\sigma_L) \cdot |\mu_L^\sigma_r(T)(c) - 1/k| \leq \sqrt{\sum_{\sigma(L) \in [k]^L} \mu_L(\sigma_L) \cdot |\mu_L^\sigma_r(T)(c) - 1/k|^2} \leq \frac{1}{k} \left| \mu_L^X_r(c) - 1/k \right|. \tag{29}
\]

Observe that in (29) the quantity inside the absolute value is always non-negative (e.g. from (28)). Also, it holds that

\[
 \left| \mu_L^X_r(c) - 1/k \right| \leq ||\mu^X_L(\cdot) - \mu(\cdot)||_r = ||\mu^X_L(\cdot) - \mu^Z_L(\cdot)||_r. \tag{30}
\]

where \( Z \) is a random \( k \)-colouring of \( T \). The equality, above, holds since the distributions \( \mu^r \) and \( \mu^Z_L \) are identical. For every \( q \in [k] \) let \( Z^q \) denote a random colouring of \( T \) conditional that \( r(T) \) is coloured \( q \). By the definition of total variation distance we get the following:

\[
 ||\mu^X_L(\cdot) - \mu^Z_L(\cdot)||_r = \frac{1}{2} \sum_{c' \in [k]} \left| \mu_L^X_r(c') - \mu_L^Z(r(c')) \right| \leq \frac{1}{2} \sum_{c' \in [k]} \left| \mu_L^X(c') - 1/k \sum_{q \in [k]} \mu_L^Z(r(c')) \right| \leq \frac{1}{k} \sum_{q \in [k]} \frac{1}{2} \sum_{c' \in [k]} |\mu_L^X(c') - \mu_L^Z(r(c'))| \leq \frac{1}{k} \sum_{q \in [k]} \left| \mu_L^X(\cdot) - \mu_L^Z(\cdot) \right|. \tag{31}
\]

Since the r.h.s. of (31) is a convex combination, it follows that

\[
 ||\mu^X_L(\cdot) - \mu^Z_L(\cdot)||_r \leq \max_{q \in [k]} \left\{ \left| \mu^X_L(\cdot) - \mu^Z_L(\cdot) \right| \right\}.
\]

The proposition follows by combining the above inequality, (30) and (29).
13 Proof of Theorem 2 - Reconstruction

Consider the following.

Definition 11 (Freezable Root) For a tree $T$ of height $h$, its root is freezable if the following holds: If $h = 1$, then $r(T)$ is of degree is at least $\Delta_-$. For $h > 1$, $r(T)$ is freezable if and only if $\deg(r_T) \geq \Delta_-$ and there are at least $\Pr\{\delta(v) \geq \delta\}$ many vertices $v$ children of $r(T)$ such that $\bar{T}_v$ has a freezable root.

Definition 12 (Freezing Boundary) Consider a tree $T$ of height $h$ and let $L = L_h(T)$. Then a boundary condition $\sigma_L$ freezes the colouring $r_T$ if the following holds: There exists $c \in [k]$ such that $\mu^*_{\sigma_L} (c) = 1$.

That is, a freezing boundary condition forces a unique colouring assignment at the root $T$.

Let $F_h$ denote the set of trees of height $h$ which have freezable root. Since the total variation distance is always non-negative, it holds that

$$\mathbb{E}[[\mu^1 - \mu^2]_L] \geq \Pr[T^h_ξ \in F_h] \cdot \mathbb{E}[|\mu^1 - \mu^2|_L T^h_ξ \in F_h]$$ (32)

The proof is going to be done in two steps. First we are going to show that with probability bounded away from zero a random boundary at the vertices of level $h$ of the tree is freezing. This would imply that the expected total variation distance, above, is positive due to the standard relation

$$\sum_{\sigma_L \in [k]} \mu_L(\sigma_L) ||\mu^*_{\sigma_L}(\cdot) - \mu(\cdot)||_{r_T} \leq ||\mu^1 - \mu^2||_L$$ (33)

Then the proposition, will follow.

Lemma 5 It holds that $\Pr[T^h_ξ \in F_h] \geq 1 - g$, where $g$ is from Definition 6.

Remark 1 We should observe that the quantity $g$ does not depend on $h$, the height of the tree. I.e. the probability bound holds for any $h > 0$.

Proof of Lemma 5 We are going to use induction to show that $\Pr[T^h_ξ \notin F_h] < g$. Let $\deg_r$ denote the degree of the root $T^h_ξ$. For $h = 1$, we use Definition 11, i.e.

$$\Pr[T^h_ξ \notin F_h] = \Pr[\deg_r < \Delta_-] = \sum_{i < \Delta_-} \xi_i \leq g,$$

where the last inequality follows from the definition of the quantity $g$, i.e. from Definition 6. Assume now that $\gamma = \Pr[T^{h-1}_ξ \notin F_{h-1}] \leq g$ is true for some $h > 1$. We are going to show that it is also true that $\Pr[T^h_ξ \notin F_h] \leq g$. Let the $\mathcal{Y}_r$ denote the event that $r_T$ has less than $(\Delta_-) - (\Delta_-)^\delta$ children which $v$ such that $\bar{T}_v$ does not have a freezable root. It holds that

$$\Pr[T^h_ξ \notin F_h] \leq \Pr[\deg_r < \Delta_-] + \Pr[\deg_r \geq \Delta_-] \Pr[\mathcal{Y}_r | \deg_r \geq \Delta_-]$$

$$\leq \sum_{\xi_i < \Delta_-} \xi_i + \sum_{\xi_i \geq \Delta_-} \Pr[\mathcal{Y}_r, \deg_r = i]$$

$$\leq \sum_{\xi_i < \Delta_-} \xi_i + \sum_{\xi_i \geq \Delta_-} \xi_i \Pr[B(i, 1 - \gamma) < (\Delta_-) - (\Delta_-)^\delta]$$

$$\leq \sum_{\xi_i < \Delta_-} \xi_i + \sum_{\xi_i \geq \Delta_-} \xi_i \Pr[B(i, 1 - g) < (\Delta_-) - (\Delta_-)^\delta] \leq g \quad \text{by Definition 6.}$$
Lemma 6  For any fixed $\alpha > 0$ and $k = (1 + \alpha) \Delta_\text{L} / \ln \Delta_\text{L}$ it holds that

$$\mathbb{E} \left[ ||\mu_i - \mu_j||_{L_h} \mid T^h \in \mathcal{F}_h \right] \geq \left( 1 - \frac{2}{\log k} \right).$$

Proof: The lemma will follow by assuming any instance of the trees in $\mathcal{F}_h$, i.e. we consider a fixed tree $T \in \mathcal{F}_h$. We let $\mathbb{P}$ denote the set of these vertices $v$ children of $r(T)$ such that $T_v$ has a freezable root. Since we have assumed that $T \in \mathcal{F}_h$ it holds that $|\mathbb{P}| \geq \Delta_\text{L} - (\Delta_\text{L})^\delta$.

Take a random colouring of $T$. W.l.o.g. assume that the root is coloured with colour $c$. This means that each of the children of the root has a colour which is distributed uniformly at random in $[k] \setminus \{c\}$ and each of the colour assignments is independent of the other. So as the colour assignment of the root to be frozen, it suffices to have the following: For every colour $q \in [k] \setminus \{c\}$ there should be at least one child in $\mathbb{F}$ which is assigned $q$ and its colouring is frozen. Clearly, examining only the children of the $r(T)$ which are in $\mathbb{F}$ will yield a lower bound for the probability that we have a frozen colouring at $r(T)$. Let $P_h$ denote the probability that the root of $T$ is frozen. For the Gibbs distribution of the tree $T$ then it holds that

$$||\mu_i - \mu_j||_{L_h} \geq P_h.$$ 

Also, since the tree $T$ is chosen arbitrarily $P_h$ is a lower bound for the expectation of the total variation distance. The lemma follows by bounding appropriately $P_h$.

At this point, we can derive the bound by working, essentially, as in [26, 28, 29]. We present the steps for bounding $P_h$, just for the sake of completeness.

Letting $w_q$ denote the number of occurrences of the colour $q$ between the vertices in $\mathbb{F}$ we have that

$$P_h = \mathbb{E} \left[ \prod_{q \in [k] \setminus \{c\}} \left( 1 - (1 - P_{h-1})^w_q \right) \right],$$

where the expectation is w.r.t. the random variables $w_q$. Clearly the variables $w_q$ for different $q$ follow the multinomial distribution. E.g. the should sum to $|\mathbb{F}|$. Clearly the random variables are correlated with each other.

Consider a set of $k - 1$ independent random variables $\tilde{w}_q$ for every $q \in [k] \setminus \{c\}$. Each $\tilde{w}_q$ follows a Poisson distribution with parameter $D = \frac{|\mathbb{F}|}{k-1} \left( 1 - \frac{1}{\log k} \right)$.

It is elementary to show that conditional that $\sum_{q \in [k] \setminus \{c\}} \tilde{w}_q \leq |\mathbb{F}|$ there is a coupling of $(w_1, \ldots, w_{k-1})$ and $(\tilde{w}_1, \ldots, \tilde{w}_{k-1})$ such that for every $q$ it holds that $w_q \geq \tilde{w}_q$, (e.g. see Lemma 4 in [29]). Then clearly we get that

$$P_h \geq \mathbb{E} \left[ \prod_{q \in [k] \setminus \{c\}} \left( 1 - (1 - P_{h-1})^{\tilde{w}_q} \right) \right] - \Pr \left[ \sum_{q \in [k] \setminus \{c\}} \tilde{w}_q > |\mathbb{F}| \right] \geq \prod_{q \in [k] \setminus \{c\}} \mathbb{E} \left[ \left( 1 - (1 - P_{h-1})^{\tilde{w}_q} \right) \right] - \Pr \left[ \sum_{q \in [k] \setminus \{c\}} \tilde{w}_q > |\mathbb{F}| \right] \geq (1 - \exp(P_{h-1} D))^{k-1} - \Pr \left[ \sum_{q \in [k] \setminus \{c\}} \tilde{w}_q > |\mathbb{F}| \right],$$

in the second inequality we use the fact that $\tilde{w}_q$s are independent with each other. It holds that $\sum_{q \in [k] \setminus \{c\}} \tilde{w}_q$ is distributed as in Po($|\mathbb{F}| (1 - 1/\log k)$). Thus, it holds that $s = \Pr \left[ \sum_{q \in [k] \setminus \{c\}} \tilde{w}_q > |\mathbb{F}| \right] \leq 1/k^2$. 

The lemma follows.
Let \( f(x) = (1 - \exp(xD))^k - 1 \). Then it is direct to verify that \( f(1 - \frac{1}{\log k}) > 1 - \frac{1}{\log k} \). Since \( P_0 = 1 \) and \( f(x) \) is increasing function we get that \( P_h > 1 - \frac{1}{\log k} \), for any \( h \geq 0 \).

\[ \square \]

14 Proof of Theorem

We will show the theorem by using Theorem\(^2\)

Consider some fixed \( \alpha > 0 \). Let \( k_1 = (1 + \alpha)d_\xi/\ln d_\xi \) and \( k_2 = (1 - \alpha)d_\xi/\ln d_\xi \). Clearly, we can find appropriate \( \gamma_1 = \gamma_1(\alpha) > 0 \) bounded above and below by a fixed number (independent of \( d_\xi \)) such that \( \Delta_+ = (1 + \gamma_1)d_\xi \) and \( k_1 = (1 + \alpha/2)\Delta_+ / \ln \Delta_+ \). Similarly, we can find appropriate \( \gamma_2 = \gamma_2(\alpha) > 0 \) bounded above and below by a fixed number (independent of \( d_\xi \)) such that \( \Delta_- = (1 - \gamma_2)d_\xi \) and \( k = (1 - \alpha/2)\Delta_- / \ln \Delta_- \).

The theorem follows by showing that both \( \Delta_+ \) and \( \Delta_- \) (as chosen above) satisfy the restrictions from Theorem\(^2\)(and those that are implied from Definitions\(^5\)\(6\)). For this, we use the assumptions about \( \xi \) in \(\text{[1]}\).

Consider first the conditions related to \( \Delta_+ \). For this we need to use the assumption from \(\text{[1]}\) that for any \( x \geq (1 + \gamma_1)d_\xi \) it holds that

\[ \Pr[\mathcal{Z} \geq x] \leq x^{-c}, \]

for sufficiently large \( c > 0 \). According to Theorem\(^2\), we should choose \( \delta \in (\alpha/4, 1/10) \). For any such \( \delta \) and \( \Delta_+ \) as defined above observe that condition \(\text{[2]}\) is trivially satisfied for \( g = 2 \Pr[\mathcal{Z} > (1 + \gamma_1)d_\xi] \leq 2d_\xi^{-c} \), where \( c > 0 \) is sufficiently large.

The leftmost conditions in \(\text{[3]}\) is also satisfied for sufficiently large \( c > 0 \). I.e. it holds that

\[ \sum_{t > (1 + \gamma_1)d_\xi} t \cdot \xi_t \leq \sum_{t > (1 + \gamma_1)d_\xi} t \cdot \Pr[\mathcal{Z} \geq t] \leq 2[(1 + \gamma_1)d_\xi]^{-(c-1)}. \]

The second condition in \(\text{[3]}\) is trivially satisfied for the specific value of \( g \) we chose.

Consider now the conditions related to \( \Delta_- \). We are going to show that \( g = 2 \Pr[\mathcal{Z} < (1 - \gamma_2)d_\xi] \) satisfies \(\text{[4]}\) if \( \Delta_- = (1 - \gamma_2)d_\xi \) and we choose any \( \delta \in (\alpha/4, 1/10) \). In particular, we have to upper bound appropriately the rightmost sum in \(\text{[4]}\).

To see why \( g = 2 \Pr[\mathcal{Z} < (1 - \gamma_2)d_\xi] \) satisfies \(\text{[4]}\) observe that \( g \leq 2d_\xi^{-c} \) (this is from \(\text{[1]}\)). It holds that \( g \cdot \Delta_- < d_\xi^{-c/2} \ll (\Delta_-)^{-1+\delta} \). This implies that for any \( i \geq \Delta_- \) we have that

\[ \Pr[B(i, 1 - g) < (\Delta_-) - (\Delta_-)^\delta] \leq \Pr[B(\Delta_-, 1 - g) < (\Delta_-) - (\Delta_-)^\delta], \]

as \( \Delta_- \) is bounded from above and \( g > 0 \) and the larger \( i \) gets the larger the deviation gets. Thus, it holds that

\[ \sum_{i \geq \Delta_-} \xi_i \Pr[B(i, 1 - g) < (\Delta_-) - (\Delta_-)^\delta] \leq \Pr[B(\Delta_-, 1 - g) < (\Delta_-) - (\Delta_-)^\delta] \sum_{i \geq \Delta_-} \xi_i \]

\[ \leq \Pr[B(\Delta_-, 1 - g) < (\Delta_-) - (\Delta_-)^\delta] \]

\[ = \Pr[B(\Delta_-, g) > (\Delta_-)^\delta] \leq \exp(-\Delta^\delta). \]

The inequality in the second line follows from the fact that \( \sum_{i \geq \Delta_-} \xi_i \leq 1 \). The last inequality follows from a direct application of Chernoff bounds, i.e. Corollary 2.4 in \(\text{[13]}\). The theorem follows.
Acknowledgement. I would like to thank Amin Coja-Oghlan for the discussions on the problem and for his comments and suggestions for improving the content of this work.

References

[1] D. Achlioptas, A. Coja-Oghlan. Algorithmic Barriers from Phase Transitions. In Proc. of 49th Annual IEEE Symposium on Foundations of Computer Science, FOCS 2008.

[2] N. Bhatnagar, J. Vera, E. Vigoda and D. Weitz. Reconstruction for coloring on Trees. SIAM J. Discrete Math. 25 (2), (2011), pp 809-826.

[3] N. Bhatnagar, A. Sly, and P. Tetali. Reconstruction Threshold for the Hardcore Model. In Proc. of the 14th International Conference on Randomization and Computation (RANDOM), 6302 LNCS, pp 434-447, 2010.

[4] C. Borgs, J. Chayes, E. Mossel, and S. Roch. The kesten-stigum reconstruction bound is tight for roughly symmetric binary channels. In Proc. of the 47th Annual IEEE Symposium on Foundations of Computer Science (FOCS06), pp 518530, 2006.

[5] J. T. Chayes, L. Chayes, J. Sethna and D. J. Thoules. A mean field spin glass with short-range interactions. Comm. Math.. Phys., 106(1), pp41-89, 1986.

[6] Amin Coja-Oghlan. A better algorithm for random k-SAT. SIAM Journal on Computing 39:2823-2864.

[7] A. Coja-Oghlan C. Efthymiou. On independent sets in random graphs. To appear in the journal Random Structures and Algorithms. Also, in proc. of 22nd Symposium on Discrete Algorithms (SODA’11), pp 136-144, 2011.

[8] C. Daskalakis, E. Mossel and S. Roch. Optimal Phylogenetic Reconstruction. In proc. of the 38th Annual ACM Symposium on Theory of Computing (STOC), 159-168, 2006.

[9] M. Dyer, A. Flaxman, A. M. Frieze and E. Vigoda. Random colouring sparse random graphs with fewer colours than the maximum degree. Random Struct. and Algorithms 29, pp. 450-465, 2006.

[10] C. Efthymiou, MCMC sampling colourings and independent sets of G(n, d/n) near uniqueness threshold. In proc of SODA’14, pp 305-316, 2014.

[11] C. Efthymiou A simple algorithm for sampling colouring of G(n, d/n) using (2 + ε)d colours. In proc of SODA’12, pp 272-280, 2012.

[12] C. Efthymiou Switching Colouring of G(n, d/n) for Sampling up to Gibbs Uniqueness Threshold. Technical Report, April 2014.

[13] W. Evan, C. Kenyon, Y. Peres and L. J. Schulman. Broadcasting on trees, the Ising model. The Annals of Applied Probability, 2000, Vol. 10 No. 2, , 410-433

[14] A. Gerschenfeld and A. Montanari. Reconstruction for models on random graphs. In Proc. of 48th Annual IEEE Symposium on Foundations of Computer Science, FOCS 2007.

[15] H. O. Georgii. Gibbs Measures and Phase Transitions, de Gruyter Stud. Math. 9, de Gruyter, Berlin, 1988

[16] G. R. Grimmett and C. J. H. McDiarmid. On colouring random graphs. Math. Proc. of the Camb. Phil. Soc. 77 (02), pp 313-32, 1975.
[17] F. Krzakala, A. Montanari, F. Ricci-Tersenghi, G. Semerjianc, L. Zdeborova. *Gibbs states and the set of solutions of random constraint satisfaction problems*. In Proc. National Academy of Sciences 104 (2007) pp 10318-10323.

[18] S. Janson, T. Łuczak and A. Ruciński. *Random Graphs*. John Wiley & Sons, Inc. 2000

[19] F. Martinelli, A. Sinclair and D. Weitz. *Glauber dynamics on trees: Boundary conditions and mixing time*. Comm. Mathm. Phys., 250 (2004), pp 301-334.

[20] F. Martinelli, A. Sinclair and D. Weitz. *Fast mixing for independent sets, colorings and other models on trees*. Random Structures Algorithms, 31, (2007), pp 134-172.

[21] M. Mézard, G. Parisi and R. Zecchina. *Analytic and Algorithmic Solution of Random Satisfiability Problems*. Science 297 no. 5582 pp. 812-815, 2002.

[22] M. Molloy. *The freezing threshold for k-colourings of a random graph*. In proc. of the 44th ACM Symposium on Theory of Computing (STOC’12), pp 921-930, 2012.

[23] A. Montanari, R. Restrepo and P. Tetali. *Reconstruction and Clustering in Random Constraint Satisfaction Problems*, SIAM Journal on Discrete Mathematics, 2011

[24] E. Mossel. *Reconstruction on Trees: Beating the second eignevalue*. The Annals of Applied Probability 2001, Vol. 11, No. 1 285-300.

[25] E. Mossel. *Phase transitions in phylogeny*. Trans. Amer. Math. Soc., 356(6):2379-2404, 2004

[26] E. Mossel and Y. Peres. *Information flow on trees*. The Annals of Applied Probability 2003, Vol. 13, No. 3, 817-844.

[27] R. Motwani & P. Raghavan. *Randomized Algorithms*. Cambridge University Press, 1995.

[28] G. Semerjian. *On the freezing of variables in random constraint satisfaction problems*. J. Stat. Phys, 2008, Vol. 130, No 2, 251-293.

[29] A. Sly. *Reconstruction of Random Colourings*, Commun. Math. Phys., 188 (2009), pp 943-961.