Gibbs properties of the fuzzy Potts model on trees and in mean field

Olle Häggström*† and Christof Külske‡

November 21, 2003

Abstract

We study Gibbs properties of the fuzzy Potts model in the mean field case (i.e. on a complete graph) and on trees. For the mean field case, a complete characterization of the set of temperatures for which non-Gibbsianness happens is given. The results for trees are somewhat less explicit, but we do show for general trees that non-Gibbsianness of the fuzzy Potts model happens exactly for those temperatures where the underlying Potts model has multiple Gibbs measures.

Key words: Potts model, fuzzy Potts model, Gibbs measures, non-Gibbsian measures, trees, mean-field models.
AMS 2000 subject classification: 82B20, 82B26.

1 Introduction

It used to be taken for granted that simple transformations of Gibbs measures are themselves Gibbsian. A few counterexamples were found in the 70's and 80's [14, 29], but these were usually referred to as being somehow exceptional or pathological. In the seminal paper from 1993 by van Enter, Fernández and Sokal [8], further examples were found, and a systematic study of Gibbsianness vs. non-Gibbsianness of large classes of transformed or projected versions of Gibbs systems began; see [7, 26, 5, 12, 19, 11, 10, 20, 6, 18] for some of the subsequent work in this area.

In particular, Gibbs properties of the so-called fuzzy Potts model were studied in Maes and Vande Velde [26] and Häggström [18]. Like almost all work in the study of Gibbsianness vs. non-Gibbsianness, these papers focused on the case where the underlying lattice is $\mathbb{Z}^d$. There are a few exceptions: Häggström [16] and Le Ny [23, 24] studied these issues for certain models living on trees, and Külske [21] considered analogous questions for models living on a complete graph, also known as the Curie–Weiss or mean field case. In this paper, we shall continue in these directions by studying Gibbs properties of the fuzzy Potts model on trees and in the mean field setup.

The fuzzy Potts model arises, loosely speaking, from the standard q-state Potts model by looking at it with a pair of glasses that prevents from distinguishing some of

*Research supported by the Swedish Research Council

†Mathematical Statistics, Chalmers University of Technology, 412 96 Göteborg, SWEDEN, olleh@math.chalmers.se, http://www.math.chalmers.se/~olleh/

‡Weierstrass-Institut für Angewandte Analysis und Stochastik, Mohrenstr. 39, D-10117 Berlin, GERMANY, kuelske@wias-berlin.de, http://www.wias-berlin.de/people/kuelske/

the spin values; see Section 2 for precise definitions. This makes the fuzzy Potts model one of the most basic examples of a hidden Markov random field [22], and it has also turned out to be useful in the study of percolation-theoretic properties of the underlying Potts model [3, 17]. Maes and Vande Velde [26] speculated that Gibbsianness of the fuzzy Potts model on $\mathbb{Z}^d$ might hold precisely in the Gibbs uniqueness regime (i.e., above the critical temperature) of the underlying Potts model, but this was shown in [18] not to be the case: non-Gibbsianness of the fuzzy Potts model happens also for some parameter values where the underlying Potts model has a unique Gibbs measure. In the following result, which is our main result for trees, we see that the desired equivalence between on one hand Gibbsianness of the fuzzy Potts model and on the other hand Gibbs uniqueness of the underlying Potts model does hold when $\mathbb{Z}^d$ is replaced by a tree.

Theorem 1.1 *Consider the q-state Potts model on a tree Γ at inverse temperature β, and let s and $r_1, \dots, r_s$ be positive integers with $1 < s < q$ and $\sum_{i=1}^{s} r_i = q$. The set $\mathcal{G}$ of Gibbs measures for this Potts model contains an element whose corresponding fuzzy Potts measure with spin partition $(r_1, \dots, r_s)$ is non-Gibbsian, if and only if $|\mathcal{G}| > 1$.*

Here $|\cdot|$ denotes cardinality, while the remaining notation and terminology will be explained in later sections. Our proof of the non-Gibbsianness part of Theorem 1.1 will proceed by showing that conditional probabilities in certain fuzzy Potts measures lack the continuity properties needed for Gibbsianness. We shall even see that the set of discontinuities has positive measure, so that not only Gibbsianness but also so-called *almost sure Gibbsianness* fails.

We move on to the mean-field fuzzy Potts model, which lives on a complete graph on N vertices and for which we consider asymptotics as $N \to \infty$. Here the situation is quite different. Before we state our main result a few general remarks are in order. First of all one has to be careful to find the right way of asking for "Gibbssianness" vs. "non-Gibbsianness" for mean-field models. It must be asked in an appropriate sense if we want to see non-trivial behavior that reflects the lattice-phenomenon in a natural way. We remind the reader that the Gibbs measures of simple mean-field models usually converge weakly to (linear combinations of) product measures. A (non-trivial) linear combination of product measures is non-Gibbsian and has each spin configuration as a point of discontinuity when we are looking at it in the *product topology* [9]. So the problem of finding non-Gibbsianness in mean-field models would always have a trivial (negative) answer as soon as there is a phase transition, and a trivial (positive) answer as soon as there is no phase transition, independently of the model. We stress that this is a very different phenomenon than the one happening for the fuzzy Potts model on the tree described above. However, if we don't want to stop at this point but want to see something meaningful we must proceed differently. As it was argued in [21] *non-Gibbsianness for mean-field models* should be understood as *discontinuity of conditional probabilities as a function of the conditioning*, but the notion of continuity must *not* be taken with respect to product topology. More precisely, we need to perform the following limiting procedure.

1. Take the conditioning of the conditional probabilities of the finite volume Gibbs-measures while staying in finite volume. Due to permutation invariance, these conditional probabilities are automatically volume-dependent *functions of the empirical average* over all the spins in the conditioning.

2. Show that the large volume-limit for these functions exists, and look at their continuity properties.

When these limiting conditional probabilities are discontinuous, we have found an analogue of "non-Gibbsian" behavior in the mean-field model. When they are continuous, the mean-field model behaves in a "Gibbsian" way. In the case of non-Gibbsian behavior we can carry the analogue between mean-field and lattice to the notion of "almost sure Gibbsianness" (that is familiar on the lattice). For the mean-field model we look at the size of the set of the discontinuity points in the large volume-limit, with respect to the limiting measure on the empirical distribution. If the discontinuity points get measure zero, we have found the mean-field analogue of "almost sure Gibbsian" behavior.

An analysis of this sort was carried out in [21] for the decimation transformation of the Ising ferromagnet, and examples of joint measures in random systems including the random field Ising model. For the models we were looking at we saw a surprising analogy between mean-field and lattice results.

We are now ready to state our main result for the mean-field version of the fuzzy Potts model in short form. Precise definitions and more details will be given in 5.

Theorem 1.2 *Consider the q-state mean-field Potts model at inverse temperature β, and let s and $r_1, \dots, r_s$ be positive integers with $1 < s < q$ and $\sum_{i=1}^{s} r_i = q$. Consider the limiting conditional probabilities of the corresponding fuzzy Potts model with spin partition $(r_1, \dots, r_s)$.*

(i) *Suppose that $r_i \leq 2$ for all $i = 1, \dots, s$. Then the limiting conditional probabilities are continuous functions of the empirical mean of the conditioning, for all $\beta \geq 0$.*

Assume that $r_i \geq 3$ for some i and put $r_* := \min\{r \geq 3, r = r_i \text{ for some } i = 1, \dots, s\}$. Denote by $\beta_c(r)$ the inverse critical temperature of the r-state Potts model. Then the following holds.

(ii) *The limiting conditional probabilities are continuous for all $\beta < \beta_c(r_*)$.*

(iii) *The limiting conditional probabilities are discontinuous for all $\beta \geq \beta_c(r_*)$.*

(iv) *The set of discontinuity points has zero measure in the infinite volume limit in all cases.*

Thus, we have a rather complete picture for the limiting behavior of the model on complete graphs. Note that from (iii) follows in particular that there is an interesting range of temperatures $\beta_c(r_*) \leq \beta < \beta_c(q)$ when the underlying Potts model shows no phase transition but the fuzzy model is non-Gibbsian. (It is well-known that $\beta_c(q)$ is increasing with q.) As mentioned above, the existence of such a region was shown on the lattice in [18]; in the present mean-field model the lower endpoint of the interval is moreover proved to be $\beta_c(r_*)$ (which is only a conjecture on the lattice). For such a non-Gibbsianness to occur in mean-field we need however that there is at least one fuzzy class containing three or more spin-values. This is due to the fact that the discontinuity of the limiting conditional probabilities is related to a first order transition within one fuzzy class, and such a transition exists if and only if there are at least three spin values.

Controlling the size of the set of discontinuities is a more subtle task, but we manage in Theorem 1.2 (iv) to provide the complete answer in the mean-field case: almost sure Gibbsianness holds regardless of the choice of parameter values.

The rest of the paper is organized as follows. In Section 2 we define the models. In Section 3 we briefly explain why, in the case of the fuzzy Potts model on trees and on $\mathbb{Z}^d$, Gibbsianness is the same thing as so-called quasilocality. Our main results for trees are stated and proved in Section 4, whereas those in the mean field setup are treated in Section 5. We mention that Section 5 can be read independently of Sections 3 and 4.

2 The models

In this section we give the definitions (following [18]) of the Potts model and the fuzzy Potts model, first on finite graphs, and then on infinite graphs. The results in Section 2.3 concerning infinite-volume limits of the Potts model date back to Aizenman et al. [1]; see also [13] for a detailed account of these results.

2.1 Potts in finite volume

For a positive integer q, the q-state **Potts model** on a finite graph $G = (V, E)$ is a random assignment of $\{1, \dots, q\}$-valued spins to the vertices of G. The Gibbs measure $\pi^G_{q,\beta}$ for the q-state Potts model on G at inverse temperature $\beta \geq 0$, is the probability measure $\pi^G_{q,\beta}$ on $\{1, \dots, q\}^V$ which to each $\xi \in \{1, \dots, q\}^V$ assigns probability

$$\pi^G_{q,\beta}(\xi) = \frac{1}{Z^G_{q,\beta}} \exp\left(2\beta \sum_{\langle x,y\rangle \in E} I_{\{\xi(x)=\xi(y)\}}\right). \tag{1}$$

Here $\langle x, y\rangle$ denotes the edge connecting $x, y \in V$, I_A is the indicator function of the event A, and $Z^G_{q,\beta}$ is a normalizing constant.

2.2 Fuzzy Potts in finite volume

Next, let s and $r_1, \dots, r_s$ be positive integers such that $\sum_{i=1}^s r_i = q$. The **fuzzy Potts model** on G with these parameters arises by taking the q-state Potts model on G, and then identifying the first r_1 Potts states with a single fuzzy spin value 1, the next r_2 of the states with fuzzy spin value 2, and so on. A more precise definition is as follows. Fix q, β and $(r_1, \dots, r_s)$ as above. Let X be a $\{1, \dots, q\}^V$-valued random object distributed according to the Gibbs measure $\pi^G_{q,\beta}$. Then take Y to be the $\{1 \dots, s\}^V$-valued random object obtained from X by setting

$$Y(x) = \begin{cases} 1 & \text{if } X(x) \in \{1, \dots, r_1\} \\ 2 & \text{if } X(x) \in \{r_1 + 1, \dots, r_1 + r_2\} \\ \vdots & \quad\vdots \\ s & \text{if } X(x) \in \{q - r_s + 1, \dots, q\} \end{cases} \tag{2}$$

for each $x \in V$. We write $\mu^G_{q,\beta,(r_1,\dots,r_s)}$ for the probability measure on $\{1, \dots, s\}^V$ which describes the distribution of Y, and call it the fuzzy Potts measure with parameters q, β, and $(r_1, \dots, r_s)$. We call $(r_1, \dots, r_s)$ the **spin partition** for this fuzzy Potts model.

Of course, $\mu^G_{q,\beta,(r_1,\dots,r_s)}$ is uninteresting for $s = 1$, whereas for $s = q$ it just reproduces the ordinary Potts model. We therefore require that $1 < s < q$, and consequently that $q \geq 3$.

2.3 Potts in infinite volume

Now let $G = (V, E)$ be infinite and locally finite. For $W \subset V$, we define its boundary ∂W as

$$\partial W = \{x \in V \setminus W : \exists y \in W \text{ such that } \langle x, y \rangle \in E\}\,.$$

A probability measure π on $\{1, \ldots, q\}^V$ is said to be a Gibbs measure for the q-state Potts model on G at inverse temperature β, if it admits conditional probabilities such that for all finite $W \subset V$, all $\xi \in \{1, \ldots, q\}^W$ and all $\eta \in \{1, \ldots, q\}^{V \setminus W}$ we have

$$\begin{aligned}
&\pi(X(W) = \xi \,|\, X(V \setminus W) = \eta) \\
&\quad = \frac{1}{Z^{W,\eta}_{q,\beta}} \exp\left(2\beta\left(\sum_{\substack{\langle x,y\rangle \in E\\ x,y\in W}} I_{\{\xi(x)=\xi(y)\}} + \sum_{\substack{\langle x,y\rangle \in E\\ x\in W, y\in \partial W}} I_{\{\xi(x)=\eta(y)\}}\right)\right)
\end{aligned} \tag{3}$$

where the normalizing constant $Z^{W,\eta}_{q,\beta}$ depends on η but not on ξ. Note that the corresponding relation holds in the finite graph case where π is defined by (1).

The basic examples of Gibbs measures for the Potts model are constructed as follows. Let $\Lambda = \{\Lambda_n\}_{n=1}^{\infty}$ denote a sequence of subsets of V, which is an exhaustion of V in the sense that (i) each Λ_n is finite, (ii) $\Lambda_1 \subset \Lambda_2 \subset \cdots$, and (iii) $\bigcup_{n=1}^{\infty} \Lambda_n = V$. Let G_n denote the graph whose vertex set is $\Lambda_n \cup \partial\Lambda_n$, and whose edge set consists of pairs of vertices in $\Lambda_n \cup \partial\Lambda_n$ at distance 1 from each other. It is well-known that the Gibbs measures $\pi^{G_n}_{q,\beta}$ converge to a probability measure on $\{1, \ldots, q\}^V$ which is a Gibbs measure for the Potts model on G with the given parameters. Convergence takes place in the sense that probabilities of cylinder sets converge. The limiting probability measure on $\{1, \ldots, q\}^V$ is denoted $\pi^{G,0}_{q,\beta}$, and is called the Gibbs measure (for the Potts model on G with the given parameters) with **free boundary condition**. Other Gibbs measures are those with so-called **spin i boundary condition**, denoted $\pi^{G,i}_{q,\beta}$, for $i = 1, \ldots q$. These are obtained by conditioning $\pi^{G_n}_{q,\beta}$ on taking spin value i all over $\partial\Lambda_n$ and then taking limits as $n \to \infty$. The existence of these limits, and the fact that each of the measures $\pi^{G,0}_{q,\beta}, \ldots, \pi^{G,q}_{q,\beta}$ is independent of the particular choice of exhaustion $\{\Lambda_n\}_{n=1}^{\infty}$, follows from the work of Aizenman et al. [1].

The Gibbs measures $\pi^{G,0}_{q,\beta}, \pi^{G,1}_{q,\beta}, \ldots \pi^{G,q}_{q,\beta}$ may or may not coincide depending on G and on the parameter values. It is a fundamental result from [1] that the occurrence of more than one distinct Gibbs measure is (for fixed G and q) increasing in β. Hence, there exists a critical value $\beta_c = \beta_c(G, q) \in (0, \infty)$, such that for $\beta < \beta_c$, there is only one Gibbs measure (so that in particular $\pi^{G,0}_{q,\beta} = \cdots = \pi^{G,q}_{q,\beta}$), whereas for $\beta > \beta_c$, there are multiple Gibbs measures (and moreover the measures $\pi^{G,0}_{q,\beta}, \ldots, \pi^{G,q}_{q,\beta}$ are all different). The critical value may be ∞ if the graph is "too small" or 0 if the graph is "too large" (requiring unbounded degree and more than that) but in many interesting cases there is a nontrivial critical value $\beta_c \in (0, \infty)$, such as for cubic lattices in $d \geq 2$ dimensions and regular trees of degree at least 3. Yet another important result from [1] is that nonuniqueness of Gibbs measures is equivalent to having

$$\pi^{G,1}_{q,\beta}(\text{spin 1 at } x) > \frac{1}{q} \tag{4}$$

for some $x \in V$, and that if G is connected, then this is in turn equivalent to having (4) for *every* $x \in V$. (For symmetry reasons, we have

$$\pi^{G,0}_{q,\beta}(\text{spin 1 at } x) = \frac{1}{q} \tag{5}$$

for every $x \in V$. Whenever we are in the uniqueness regime of the parameter space, we then of course have (5) with $\pi^{G,0}_{q,\beta}$ replaced by any of the other Gibbs measures $\pi^{G,i}_{q,\beta}$.)

2.4 Fuzzy Potts in infinite volume

Given the Gibbs measures $\pi^{G,0}_{q,\beta}, \pi^{G,1}_{q,\beta}, \ldots, \pi^{G,q}_{q,\beta}$, we define fuzzy Potts measures as in the case of finite graphs. More precisely, for q, β, and $(r_1, \ldots, r_s)$ as above, and $i \in \{0, \ldots, q\}$, we define the fuzzy Potts measure $\mu^{G,i}_{q,\beta,(r_1,\ldots,r_s)}$ to be the distribution of the $\{1, \ldots, s\}^V$-valued random object Y obtained by first picking $X \in \{1, \ldots, q\}^V$ according to the Gibbs measure $\pi^{G,i}_{q,\beta}$, and then constructing Y from X as in (2).

3 Gibbsianness and quasilocality

When S is a finite set, $G = (V, E)$ is an infinite locally finite graph, and μ is a probability measure on S^V, it is well known (see, e.g., [8, Thm. 2.12]) that μ is Gibbsian if and only if it satisfies the properties of quasilocality and uniform nonnullness. The latter property means that μ admits conditional probabilities such that

$$\min_{s \in S} \inf_{\eta \in S^{V \setminus \{x\}}} \mu(X(x) = s \mid X(V \setminus \{x\}) = \eta) > 0$$

for each $x \in V$. Uniformly nonnullness holds in the Potts model, and it is easy to see that this property is inherited by the fuzzy Potts model; see [18, Lem. 4.5]. Hence, the problem of determining whether the fuzzy Potts model with given parameter values is Gibbsian is reduced to that of whether it is quasilocal. Quasilocality is defined as follows, where (as in Section 2) $\Lambda = \{\Lambda_n\}_{n=1}^{\infty}$ is an exhaustion of V (the definition does not depend on the particular choice of Λ).

Definition 3.1 *Let S be a finite set and let $G = (V, E)$ be an infinite locally finite graph. A probability measure μ on S^V is said to be* **quasilocal** *if it admits conditional probabilities such that for all finite $W \subset V$ and all $\xi \in S^W$ we have*

$$\lim_{n\to\infty} \sup_{\substack{\eta,\eta' \in S^{V\setminus W} \\ \eta(\Lambda_n \setminus W) = \eta'(\Lambda_n \setminus W)}} \left| \mu(X(W) = \xi \mid X(V \setminus W) = \eta) - \mu(X(W) = \xi \mid X(V \setminus W) = \eta') \right| = 0 \,. \tag{6}$$

Because of the asserted equivalence between Gibbsianness and quasilocality for the fuzzy Potts model, we shall in the following focus entirely on quasilocality. In Section 4 on trees, this means studying the property in Definition 3.1 verbatim, whereas in Section 5 we need to adapt the definition of quasilocality somewhat (following [21]), as hinted in Section 1.

4 The fuzzy Potts model on trees

4.1 Trees: definitions

A tree Γ is a connected graph without cycles. In addition to these properties, we assume that Γ is locally finite, and we denote its vertex set and edge set by V_Γ and E_Γ, respectively. Pick an arbitrary vertex in $\rho \in V_\Gamma$ and call it the **root** of Γ. For $x, y \in V_\Gamma$, let $\mathrm{dist}(x, y)$ denote the graph-theoretic distance between x and y in Γ. If x and y share an edge and $\mathrm{dist}(y, \rho) = \mathrm{dist}(x, \rho) + 1$, then we call y a **child** of x, and x is the **parent** of y. More generally, if x is on the unique self-avoiding path from ρ to y, then y is called a **descendant** of x, and x is an **ancestor** of y. Each vertex x except for the root has exactly one parent, denoted $\mathrm{parent}(x)$ while the number of children may vary. If two vertices x and y have the same parent, then we call them **siblings**.

An important example is when, for some $d \geq 2$, the root has $d+1$ children and all others have d children; this is referred to as the regular tree with degree d. See, e.g., [28] for a variety of other interesting examples of trees.

For $n = 0, 1, \ldots$, let $\Gamma_n = (V_{\Gamma_n}, E_{\Gamma_n})$ be the subgraph (subtree) of Γ given by

$$V_{\Gamma_n} = \{x \in V_\Gamma : \mathrm{dist}(x, \rho) \leq n\}$$

and

$$E_{\Gamma_n} = \{e \in E_\Gamma : \text{both endpoints of } e \text{ are in } V_{\Gamma_n}\},$$

and note that $\{V_{\Gamma_n}\}_{n=1}^{\infty}$ is an exhaustion of V_Γ. For $x \in \Gamma$, let $\Gamma_{(x)}$ denote the induced subtree of Γ whose vertex set consists of x and all its descendants. In other words, $\Gamma_{(x)} = (V_{\Gamma_{(x)}}, E_{\Gamma_{(x)}})$ with

$$V_{\Gamma_{(x)}} = \{y \in V_\Gamma : x \text{ is an ancestor of } y\}$$

and

$$E_{\Gamma_{(x)}} = \{e \in E_\Gamma : \text{both endpoints of } e \text{ are in } V_{\Gamma_{(x)}}\}.$$

Finally, for $x \in \Gamma$ and $n \geq \mathrm{dist}(x, \rho)$, define the subtree $\Gamma_{(x,n)} = (V_{\Gamma_{(x,n)}}, E_{\Gamma_{(x,n)}})$ by setting

$$V_{\Gamma_{(x,n)}} = V_{\Gamma_{(x)}} \cap V_{\Gamma_n}$$

and

$$E_{\Gamma_{(x,n)}} = E_{\Gamma_{(x)}} \cap E_{\Gamma_n}.$$

4.2 Proofs

The key results for proving Theorem 1.1 are the following two propositions.

Proposition 4.1 *Let Γ be a tree, and fix the parameter values q, β, s and $(r_1, \ldots r_s)$ with $1 < s < q$ for the Potts model and the fuzzy Potts model on Γ. Then the fuzzy Potts measure $\mu^{G,0}_{q,\beta,(r_1,\ldots,r_s)}$ corresponding to the Gibbs measure with free boundary condition, is quasilocal.*

Proposition 4.2 *Let Γ be a tree, and fix the parameter values q, β, s and $(r_1, \ldots r_s)$ with $1 < s < q$ and $r_1 > 1$ for the Potts model and the fuzzy Potts model on Γ. Suppose that $\pi^{G,1}_{q,\beta} \neq \pi^{G,0}_{q,\beta}$. Then $\mu^{G,1}_{q,\beta,(r_1,\ldots,r_s)}$ is nonquasilocal.*

Proof of Theorem 1.1 from Propositions 4.1 and 4.2: Since $s < q$, we must have $r_i > 1$ for some $i \in \{1, \ldots, s\}$, and there is no loss of generality in assuming that $r_1 > 1$. If $|\mathcal{G}| > 1$, then $\pi^{G,1}_{q,\beta} \neq \pi^{G,0}_{q,\beta}$ due to (4) and (5). Hence, using Proposition 4.2, $\mu^{G,1}_{q,\beta,(r_1,\ldots,r_s)}$ is nonquasilocal and therefore non-Gibbsian, and the 'if' part of the theorem is established. For the 'only if' part, note that if $|\mathcal{G}| = 1$, then $\mathcal{G} = \{\pi^{G,0}_{q,\beta}\}$, so that $\mu^{G,0}_{q,\beta,(r_1,\ldots,r_s)}$ is the only fuzzy Potts measure, which by Proposition 4.1 is quasilocal and therefore Gibbsian. $\square$

It remains to prove Propositions 4.1 and 4.2. To this end, we need to introduce the notion of a tree-indexed Markov chain on Γ, and its relation to Gibbs measures for the Potts model on Γ. This relation is well-known for regular trees (see for instance [30, 31, 2]), while the extension to general trees seems to be less well-studied.

Let $(x_0, x_1, \ldots)$ be an enumeration of V_Γ such that the root ρ comes first ($x_0 = \rho$), then all vertices in $V_{\Gamma_1} \setminus \{\rho\}$, then all vertices in $V_{\Gamma_2} \setminus V_{\Gamma_1}$, and so on. Fix q, let ν be a probability measure on $\{1, \ldots, q\}$ (which will play the role of an initial distribution), and let $P = (P_{ij})_{i,j \in \{1,\ldots,q\}}$ be a transition matrix. Let X be the $\{1, \ldots, q\}^{V_\Gamma}$-valued random spin configuration obtained as follows. First pick $X(x_0) \in \{1, \ldots, q\}$ according to ν. Then, inductively, once $X(x_0), \ldots, X(x_n)$ have been determined, pick $X(x_{n+1}) \in \{1, \ldots, q\}$ with distribution $(P_{i1}, \ldots, P_{iq})$ where $i = X(\text{parent}(x_{n+1}))$. For obvious reasons, X is called a tree-indexed Markov chain on Γ.

There is sometimes reason to consider inhomogeneous tree-indexed Markov chains, where the transition matrix P is allowed to depend on where in the tree we are: for every $x \in V_\Gamma \setminus \{\rho\}$, we then have a transition matrix $P^x = (P^x_{ij})_{i,j \in \{1,\ldots,q\}}$, and X is generated as above with $X(x)$ chosen according to the distribution $(P^x_{i1}, \ldots, P^x_{iq})$ where $i = X(\text{parent}(x))$.

It is readily checked that a (possibly inhomogeneous) tree-indexed Markov chain X is also a Markov random field on Γ, meaning that for any finite $W \subset V_\Gamma$, the conditional distribution of $X(W)$ given $X(V_\Gamma \setminus W)$ depends on $X(V_\Gamma \setminus W)$ only via $X(\partial W)$. Hence the supremum in (6) becomes 0 for all n large enough so that Λ_n contains $W \cap \partial W$, so that we have the following lemma.

Lemma 4.3 *The distribution of any homogeneous or inhomogeneous tree-indexed Markov chain on Γ is quasilocal.*

Fix $\beta \geq 0$, and consider the tree-indexed Markov chain given by $\nu = (\frac{1}{q}, \ldots, \frac{1}{q})$ and transition matrix $P = (P_{ij})_{i,j \in \{1,\ldots,q\}}$ given by

$$P_{ij} = \begin{cases} \frac{e^{2\beta}}{e^{2\beta}+q-1} & \text{if } i = j \\ \frac{1}{e^{2\beta}+q-1} & \text{otherwise.} \end{cases} \tag{7}$$

Let $X \in \{1, \ldots, q\}^{V_\Gamma}$ be given by this particular tree-indexed Markov chain. By directly checking (1), we see that $X(\Lambda_n)$ has distribution $\pi^{\Gamma_n}_{q,\beta}$. By taking limits as $n \to \infty$ and considering the construction of $\pi^{G,0}_{q,\beta}$ in Section 2.3, we see that X is distributed according to the Gibbs measure $\pi^{\Gamma,0}_{q,\beta}$ for the Potts model on Γ with free boundary condition.

Proof of Proposition 4.1: Construct $X \in \{1, \ldots, q\}^{V_\Gamma}$ sequentially as above, with $\nu = (\frac{1}{q}, \ldots, \frac{1}{q})$ and P given by (7), and let $Y \in \{1, \ldots, r\}^{V_\Gamma}$ from X as in (2). Then the conditional distribution of $Y(x_{n+1})$ given $X(x_0), \ldots, X(x_n)$ such that $X(\text{parent}(x_{n+1})) = i$ and $Y(\text{parent}(x_{n+1})) = k$), is given by

$$\mathbb{P}(Y(x_{n+1}) = l \mid \cdots) = \begin{cases} \frac{e^{2\beta}+r_k-1}{e^{2\beta}+q-1} & \text{if } l = k \\ \frac{r_k}{e^{2\beta}+q-1} & \text{otherwise,} \end{cases} \tag{8}$$

which follows by summing over the possible values of $X(x_{n+1})$. Note that the right-hand side of (8) depends on $X(x_0), \ldots, X(x_n)$ only through $Y(\text{parent}(x_{n+1}))$. It follows that Y is a tree-indexed Markov chain with state space $\{1, \ldots, s\}$, initial distribution

$(\frac{r_1}{q}, \ldots, \frac{r_s}{q})$ and transition matrix $P = (P_{kl})_{k,l \in \{1,\ldots,s\}}$ given by

$$P_{kl} = \begin{cases} \frac{e^{2\beta}+r_l-1}{e^{2\beta}+q-1} & \text{if } l = k \\ \frac{r_l}{e^{2\beta}+q-1} & \text{otherwise.} \end{cases} \tag{9}$$

Quasilocality of Y now follows from Lemma 4.3. $\square$

For the proof of Proposition 4.2, we need to consider the tree-indexed Markov chain on Γ corresponding to the Gibbs measure $\pi^{\Gamma,1}_{q,\beta}$ with the "all 1" boundary condition. This is a bit more complicated than the case of $\pi^{\Gamma,0}_{q,\beta}$ due to the lack of full symmetry among the spin values.

For $x \in V_\Gamma$, consider the Gibbs measure $\pi^{\Gamma_{(x)},1}_{q,\beta}$, and in particular the probability $\pi^{\Gamma_{(x)},1}_{q,\beta}(\text{spin 1 at } x)$, which we denote by a_x. (Note that a_x is in general distinct from $\pi^{\Gamma,1}_{q,\beta}(\text{spin 1 at } x)$, because it fails to take into account, e.g., the possible influence from parent(x) on x.) For symmetry reasons, the $\pi^{\Gamma_{(x)},1}_{q,\beta}$-distribution of the spin at x is

$$\left(a_x, \frac{1-a_x}{q-1}, \frac{1-a_x}{q-1}, \ldots, \frac{1-a_x}{q-1}\right).$$

Also define

$$b_x = \frac{a_x}{(1-a_x)/(q-1)} = \frac{\pi^{\Gamma_{(x)},1}_{q,\beta}(\text{spin 1 at } x)}{\pi^{\Gamma_{(x)},1}_{q,\beta}(\text{spin 2 at } x)}. \tag{10}$$

The constants $\{b_x\}_{x \in V_\Gamma}$ satisfy the following recursion.

Lemma 4.4 *Suppose $x \in V_\Gamma$ is a vertex with k children $y_1, \ldots, y_k$. We then have*

$$b_x = \frac{\prod_{i=1}^k (e^{2\beta} b_{y_i} + q - 1)}{\prod_{i=1}^k (e^{2\beta} + b_{y_i} + q - 2)}. \tag{11}$$

Proof: For n large enough so that $x \in V_{\Lambda_n}$, define, as a finite-volume analogue of (10),

$$b_{x,n} = \frac{\pi^{\Gamma_{(x,n)},1}_{q,\beta}(\text{spin 1 at } x)}{\pi^{\Gamma_{(x,n)},1}_{q,\beta}(\text{spin 2 at } x)},$$

where $\pi^{\Gamma_{(x,n)},1}_{q,\beta}$ is the finite-volume Gibbs measure for $\Gamma_{(x,n)}$ with spin 1 boundary condition on those vertices sitting furthest away from x in $\Gamma_{(x,n)}$, i.e., those at distance n from ρ in Γ. By the construction of Gibbs measures in Section 2.3, we have

$$\lim_{n\to\infty} b_{x,n} = b_x. \tag{12}$$

Imagine now the modified graph $\Gamma^*_{(x,n)}$ obtained from $\Gamma_{(x,n)}$ by removing all edges incident to x. In other words, $\Gamma^*_{(x,n)}$ is a disconnected graph with an isolated vertex x together with k connected components isomorphic to $\Gamma_{(y_1,n)}, \ldots, \Gamma_{(y_k,n)}$. When picking $X \in \{1,\ldots,q\}^{V_{\Gamma^*_{(x,n)}}}$ according to $\pi^{\Gamma^*_{(x,n)},1}_{q,\beta}$, the spin configurations on different connected components obviously become independent. In particular, if we only consider the spins $(X(x), X(y_1), \ldots, X(y_k))$, then we can note that these spins become independent, with $X(x)$ having distribution $(\frac{1}{q}, \ldots, \frac{1}{q})$, and $X(y_i)$ having distribution $(\frac{b_{y_i,n}}{b_{y_i,n}+q-1}, \frac{1}{b_{y_i,n}+q-1}, \ldots, \frac{1}{b_{y_i,n}+q-1})$.

If we now reinsert the edges between x and $y_1, \ldots, y_k$, thus recovering $\Gamma_{(x,n)}$, then the $\pi_{q,\beta}^{\Gamma_{(x,n)},1}$-distribution of $(X(x), X(y_1), \ldots, X(y_k))$ becomes the same as the corresponding $\pi_{q,\beta}^{\Gamma^*_{(x,n)},1}$-distribution above except that each configuration $\xi \in \{1, \ldots, q\}^{\{x,y_1,\ldots,y_k\}}$ is reweighted by a factor $\exp(2\beta \sum_{i=1}^k I_{\{\xi(y_i)=\xi(x)\}})$. Hence

$$\pi_{q,\beta}^{\Gamma_{(x,n)},1}((X(x), X(y_1), \ldots, X(y_k)) = \xi) = \frac{1}{Z}\prod_{i=1}^{k}(e^{2\beta I_{\{\xi(y_i)=\xi(x)\}}}\, b_{y_i,n}^{I_{\{\xi(y_i)=1\}}})$$

for some normalizing constant Z. By integrating out $X(y_1), \ldots, X(y_k)$, we get

$$b_{x,n} = \frac{\prod_{i=1}^k (e^{2\beta} b_{y_i,n} + q - 1)}{\prod_{i=1}^k (e^{2\beta} + b_{y_i,n} + q - 2)}\,.$$

Sending $n \to \infty$ in this expression, and using (12) $k+1$ times (substituting x with itself and with $y_1, \ldots, y_k$), we obtain (11), as desired. $\square$

Note that the above proof yields that given $X(x) = 1$, the spins $X(y_1), \ldots, X(y_k)$ become conditionally independent, with $X(y_i)$ having distribution

$$\left(\frac{b_{y_i}e^{2\beta}}{b_{y_i}e^{2\beta}+q-1}, \frac{1}{b_{y_i}e^{2\beta}+q-1}, \ldots, \frac{1}{b_{y_i}e^{2\beta}+q-1}\right)\,.$$

Likewise, for $l \neq 1$, conditioning on $X(x) = l$ makes $X(y_1), \ldots, X(y_k)$ conditionally independent with $X(y_i)$ taking value 1 with probability $\frac{b_{y_i}}{b_{y_i}+e^{2\beta}+q-2}$, value l with probability $\frac{e^{2\beta}}{b_{y_i}+e^{2\beta}+q-2}$, and other values with probabilities $\frac{1}{b_{y_i}+e^{2\beta}+q-2}$.

By iterating the above argument, we arrive at the following tree-indexed Markov chain description of the Gibbs measure $\pi_{q,\beta}^{\Gamma,1}$.

Lemma 4.5 *Suppose that the random spin configuration $X \in \{1, \ldots, q\}^{V_\Gamma}$ is obtained as an inhomogeneous tree-indexed Markov chain with initial distribution*

$$\nu = \left(\frac{b_\rho}{b_\rho+q-1}, \frac{1}{b_\rho+q-1} \ldots, \frac{1}{b_\rho+q-1}\right)$$

and transition matrices $P^x = (P^x_{ij})_{i,j\in\{1,\ldots,q\}}$ given by

$$P^x_{ij} = \begin{cases} \frac{b_x e^{2\beta}}{b_x e^{2\beta}+q-1} & \text{if } i=j=1 \\ \frac{1}{b_x e^{2\beta}+q-1} & \text{if } i=1,\ j\neq 1 \\ \frac{b_x}{b_x+e^{2\beta}+q-2} & \text{if } i\neq 1, j=1 \\ \frac{e^{2\beta}}{b_x+e^{2\beta}+q-2} & \text{if } i=j\neq 1 \\ \frac{1}{b_x+e^{2\beta}+q-2} & \text{otherwise.} \end{cases}$$

Then X has distribution $\pi_{q,\beta}^{\Gamma,1}$.

A crucial difference now compared to the Gibbs measure $\pi_{q,\beta}^{\Gamma,0}$ with free boundary condition is that if any $b_x \neq 1$, then there is not enough state-symmetry in the tree-indexed Markov chain in Lemma 4.5 to make the corresponding fuzzy Potts model a tree-indexed Markov chain. This will soon become clear.

A key lemma for proving nonquasilocality in the fuzzy Potts model is the following.

Lemma 4.6 *If $\pi_{q,\beta}^{\Gamma,1} \neq \pi_{q,\beta}^{\Gamma,0}$, then there exist two siblings $y_1, y_2 \in V_\Gamma$ such that $b_{y_i} > 1$ for both $i = 1$ and $i = 2$.*

Proof: It follows from the assumption $\pi_{q,\beta}^{\Gamma,1} \neq \pi_{q,\beta}^{\Gamma,0}$ using (4) that $a_\rho > \frac{1}{q}$, so that

$$b_\rho > 1\,. \tag{13}$$

Furthermore, (4) and (5) imply that $a_x \geq \frac{1}{q}$ for all $x \in V_\Gamma$, whence $b_x \geq 1$ for all $x \in V_\Gamma$. Note also that 1 is a fixed point of the recursion (11), in the sense that if all children $y_1, \ldots, y_k$ satisfy $b_{y_i} = 1$, then $b_x = 1$.

Hence, ρ must have at least one child x with $b_x > 1$. By iterating this argument we see that for any n, it must have at least one descendant x at distance n such that $b_x > 1$. Fix n and such a vertex x with $b_x > 1$ at distance n from ρ. Write $(z_0, z_1, \ldots, z_n)$ for the vertices on the self-avoiding path from x to ρ (so that in particular $z_0 = x$ and $z_n = \rho$). Next, note that the recursion (11) has the property that if one of the children y_i has $b_{y_1} > 1$, then $b_x > 0$ as well. Since $b_{z_0} > 1$ it follows that $b_{z_i} > 1$ for $i = 1, \ldots, n$.

Suppose now for contradiction that the assertion of the lemma is false, i.e., that there are no two siblings $y_1, y_2 \in V_\Gamma$ for which $b_{y_1} > 1$ and $b_{y_2} > 1$. Then none of the vertices $z_0, \ldots, z_{n-1}$ has a sibling y with $b_y > 1$. The recursion (11) along the path $(z_0, z_1, \ldots, z_n)$ then turns into a simple one-dimensional dynamical system on the space $[1, \infty)$ given by $b_{z_{i+1}} = f(b_{z_i})$ where

$$f(b) = \frac{e^{2\beta} b + q - 1}{e^{2\beta} + b + q - 2}\,.$$

This dynamical system is contractive with a unique fixed point at $b = 1$, so that – if we just keep iterating beyond the n'th iteration – for any initial value $b_{z_0} \in [1, \infty)$ we obtain

$$\lim_{n \to \infty} b_{z_n} = 1\,. \tag{14}$$

Since f is increasing and bounded by $e^{2\beta}$, we get that b_{z_1} is bounded by $e^{2\beta}$ and, therefore, that the convergence in (14) is in fact uniform in the initial value b_{z_0}. Thus we can, for any $\varepsilon > 0$, find an n which guarantees that $b_{z_n} < 1 + \varepsilon$. Thus, $b_\rho < 1 + \varepsilon$ for any $\varepsilon > 0$, whence $b_\rho = 1$. But this contradicts (13), so the proof is complete. $\Box$

Proof of Proposition 4.2: By Lemma 4.6, Γ has at least one vertex which has (at least) two children y_1 and y_2 that both have $b_{y_i} > 1$. The choice of root ρ for the tree does not influence the Gibbs measure $\pi_{q,\beta}^{\Gamma,1}$, and therefore we may assume that ρ has two such children y_1 and y_2. We shall for simplicity first prove the proposition under the assumption that

$$\rho \text{ has no other children,} \tag{15}$$

and in the end show how to remove this assumption.

We shall have a look at the conditional distribution of the fuzzy spin $Y(\rho)$ at the root, given that its neighbors (i.e., its children) take value

$$Y(y_1) = Y(y_2) = 1\,. \tag{16}$$

By summing over all $X \in \{1, \ldots, q\}^{\{\rho, y_1, y_2\}}$ such that (16) holds, and using Lemma 4.5, we obtain

$$\begin{aligned}
&\frac{\mathbb{P}(Y(\rho) = 1 \,|\, Y(y_1) = Y(y_2) = 1)}{\mathbb{P}(Y(\rho) \neq 1 \,|\, Y(y_1) = Y(y_2) = 1)} \qquad\qquad (17)\\
&= \frac{\frac{b_\rho}{b_\rho + q - 1} \prod_{i=1}^2 \frac{b_{y_i} e^{2\beta} + r_1 - 1}{b_{y_i} e^{2\beta} + q - 1} + \frac{r_1 - 1}{b_\rho + q - 1} \prod_{i=1}^2 \frac{b_{y_i} + e^{2\beta} + r_1 - 2}{b_{y_i} + e^{2\beta} + q - 2}}{\frac{q - r_1}{b_\rho + q - 1} \prod_{i=1}^2 \frac{b_{y_i} + r_1 - 1}{b_{y_i} + e^{2\beta} + q - 2}}\\
&= \frac{\prod_{i=1}^2 (b_{y_i} e^{2\beta} + r_1 - 1) + (r_1 - 1) \prod_{i=1}^2 (b_{y_i} + e^{2\beta} + r_1 - 2)}{(q - r_1) \prod_{i=1}^2 (b_{y_i} + r_1 - 1)}
\end{aligned}$$

where in the last line we have used (11) to express b_ρ in terms of the b_{y_i}'s.

Now pick an n, and consider conditioning further on some $\eta_n \in \{1, \ldots, s\}^{V_{\Gamma_{n+1}} \setminus \{\rho\}}$ such that $\eta_n(y_1) = \eta_n(y_2) = 1$. The conditional probability $\mathbb{P}(Y(\rho) = 1 \,|\, Y(y_1) = Y(y_2) = 1)$ is a convex combination of terms $\mathbb{P}(Y(\rho) = 1 \,|\, Y(V_{\Gamma_{n+1}} \setminus \{\rho\}) = \eta_n)$ for such η_n's. We can therefore find a particular $\eta_n \in \{1, \ldots, s\}^{\Lambda_{n+1} \setminus \{\rho\}}$ such that

$$\begin{aligned}
&\frac{\mathbb{P}(Y(\rho) = 1 \,|\, Y(V_{\Gamma_{n+1}} \setminus \{\rho\}) = \eta_n)}{\mathbb{P}(Y(\rho) \neq 1 \,|\, Y(V_{\Gamma_{n+1}} \setminus \{\rho\}) = \eta_n)}\\
&\geq \frac{\prod_{i=1}^2 (b_{y_i} e^{2\beta} + r_1 - 1) + (r_1 - 1) \prod_{i=1}^2 (b_{y_i} + e^{2\beta} + r_1 - 2)}{(q - r_1) \prod_{i=1}^2 (b_{y_i} + r_1 - 1)}\,. \qquad (18)
\end{aligned}$$

Fix such an η_n. Next, construct another configuration $\eta_n' \in \{1, \ldots, s\}^{V_{\Gamma_{n+1}} \setminus \{\rho\}}$ by taking

$$\eta_n'(x) = \begin{cases} \eta_n(x) & \text{for } x \in V_{\Gamma_n} \setminus \{\rho\} \\ (\eta_n(\mathrm{parent}(x)) + 1) \bmod s & \text{for } x \in V_{\Gamma_{n+1}} \setminus V_{\Gamma_n}\,. \end{cases}$$

The crucial aspects of this choice of η_n' is that (a) $\eta_n = \eta_n'$ on V_{Γ_n} and (b) each x in the remotest layer $V_{\Gamma_{n+1}} \setminus V_{\Gamma_n}$ of Γ_{n+1} has a fuzzy spin value which is different from its parent. It is readily checked that property (b) implies that the conditional distribution of $Y(V_{\Gamma_{n-1}})$ given $Y(V_{\Gamma_{n+1}} \setminus V_{\Gamma_{n-1}}) \eta_n'(V_{\Gamma_{n+1}} \setminus V_{\Gamma_{n-1}})$ becomes the same as if the underlying Gibbs measure had been not $\pi_{q,\beta}^{\Gamma,1}$ but rather the finite-volume Gibbs measure $\pi_{q,\beta}^{\Gamma_{n+1}}$ (cf. [18, Lem. 9.2]). Hence the conditional distribution of $Y(\rho)$ given that $Y(V_{\Gamma_{n+1}} \setminus \{\rho\}) = \eta_n')$ can be calculated from the tree-indexed Markov chain corresponding to free boundary condition, i.e., the one defined in (9). We get

$$\frac{\mathbb{P}(Y(\rho) = 1 \,|\, Y(V_{\Gamma_{n+1}} \setminus \{\rho\}) = \eta_n')}{\mathbb{P}(Y(\rho) \neq 1 \,|\, Y(V_{\Gamma_{n+1}} \setminus \{\rho\}) = \eta_n')} = \frac{(e^{2\beta} + r_1 - 1)^2}{(q - r_1) r_1} \tag{19}$$

Note that the right-hand sides of (18) and (19) do not depend on n. We now make the following crucial claim.

> **Claim:** If $b_{y_1} > 1$ and $b_{y_2} > 1$, then the right-hand side of (18) is strictly greater than the right-hand side of (19).

To prove the claim, define

$$a = \frac{b_{y_1} b_{y_2} + r_1 - 1}{(b_{y_1} + r_1 - 1)(b_{y_2} + r_1 - 1)}$$

and note that a can be rewritten as

$$\begin{aligned} a &= \frac{b_{y_1} b_{y_2} + r_1 - 1}{(b_{y_1} + r_1 - 1)(b_{y_2} + r_1 - 1)} \\ &= \frac{1}{r_1} \frac{r_1}{(b_{y_1} + r_1 - 1)} + \frac{b_{y_2}}{(b_{y_2} + r_1 - 1)} \frac{(b_{y_1} - 1)}{(b_{y_1} + r_1 - 1)} . \end{aligned}$$

Assuming that $b_{y_1} > 1$ and $b_{y_2} > 1$, we get that $\frac{b_{y_2}}{b_{y_2}+r_1-1} > \frac{1}{r_1}$ and that $\frac{b_{y_1}-1}{b_{y_1}+r_1-1} > 0$, whence

$$\begin{aligned} a &> \frac{1}{r_1} \frac{r_1}{(b_{y_1} + r_1 - 1)} + \frac{1}{r_1} \frac{(b_{y_1} - 1)}{(b_{y_1} + r_1 - 1)} \\ &= \frac{1}{r_1} . \end{aligned} \tag{20}$$

Next, an elementary but tedious calculation shows that the right-hand side of (18) can be rewritten as

$$\frac{a(e^{4\beta} + r_1 - 1) + (1 - a)(2e^{2\beta} + r_1 - 2)}{q - r_1} . \tag{21}$$

Analogously, the right-hand side of (19) can be rewritten as

$$\frac{\frac{1}{r_1}(e^{4\beta} + r_1 - 1) + (1 - \frac{1}{r_1})(2e^{2\beta} + r_1 - 2)}{q - r_1} . \tag{22}$$

Now, using (20) and the observation that

$$e^{4\beta} + r_1 - 1 > 2e^{2\beta} + r_1 - 2 \, ,$$

we get that the expression in (21) is strictly greater than that in (22), and the claim is proved.

Hence the difference between the left-hand sides of (18) and (19) is bounded away from 0 uniformly in n. The denominators of the left-hand sides are bounded away from 0 uniformly in n due to uniform nonnullness of the fuzzy Potts model (see Section 3). Hence

$$\mathbb{P}(Y(\rho) = 1 \,|\, Y(V_{\Gamma_{n+1}} \setminus \{\rho\}) = \eta_n) - \mathbb{P}(Y(\rho) = 1 \,|\, Y(V_{\Gamma_{n+1}} \setminus \{\rho\}) = \eta'_n)$$

is bounded away from 0 uniformly on n. By plugging in these η_n and η'_n in (6), we get, since $\eta_n = \eta'_n$ on V_{Γ_n}, that quasilocality of Y fails. This proves the proposition modulo the assumption (15).

It remains to remove the assumption (15). To do this, suppose that ρ has $k - 2$ additional children $y_3, \ldots, y_k$. We can then extend the configurations η and η' that we condition on above, to $y_3, \ldots, y_k$ and their descendants, as follows. We insist that η and η' that they take value 1 at $y_3, \ldots, y_k$, and that they take some value other than 1 at all children of $y_3, \ldots, y_k$ (they may otherwise be arbitrary on the further descendants of

$y_3, \ldots, y_k)$. Easy modifications of the calculations above show that (18) and (19) hold as before, with the modification that both right-hand sides are multiplied by

$$\left(\frac{e^{2\beta} + r_1 - 1}{r_1}\right)^{k-2}.$$

Since this factor is the same in (18) and (19), the rest of the proof goes through as before. □

Remark: Since the event conditioned on in (17) has positive measure, it is easy to extract from the above proof that the set of discontinuities of the conditional probability $\mathbb{P}(Y(\rho) = 1 | Y(V_\Gamma \setminus \{\rho\}) = \eta)$ as a function of ρ, has positive measure under $\mu^{G,1}_{q,\beta,(r_1,\ldots,r_s)}$. Hence, so-called almost sure quasilocality and almost sure Gibbsianness fails in general for the fuzzy Potts model on trees, in contrast to the $\mathbb{Z}^d$ case (see Maes and Vande Velde [26]) and the mean-field case (Theorem 1.2 (iv)). This contrast between the fuzzy Potts model on $\mathbb{Z}^d$ and on trees is analogous to the corresponding almost sure Gibbsianness issue for the random-cluster model; see [16].

4.3 Discussion

What concrete information can we extract from Theorem 1.1? Let $\beta_c = \beta_c(\Gamma, q)$ denote, as in Section 2.3, the critical value for the q-state Potts model on the tree Γ. For $q \geq 3$, we then have from Theorem 1.1 that $\beta < \beta_c$ implies that any corresponding fuzzy Potts measure is Gibbsian, while $\beta > \beta_c$ yields existence of corresponding fuzzy Potts measures that are non-Gibbsian.

It remains to specify the critical value $\beta_c(\Gamma, q)$. If we know the critical value $p_c(\Gamma, q)$ of the corresponding random-cluster model, then we can calculate $\beta_c = -\frac{1}{2}\log(1 - p_c)$ (see, e.g., [13]). For the case when Γ is a regular tree, the critical value $p_c(\Gamma, q)$ can be characterized in terms of the solutions of a certain algebraic equation given in [15, p. 235].

For general trees the situation is more complicated. For a variety of stochastic models on trees, critical values can be calculated in terms of a natural quantity known as the branching number of the tree, denoted $\mathrm{br}(\Gamma)$; see for instance [28]. Lyons [25] calculated $\beta_c(\Gamma, q)$ in terms of $\mathrm{br}(\Gamma)$ for the case $q = 2$. In contrast, and perhaps somewhat surprisingly, the critical values $\beta_c(\Gamma, q)$ for larger q do *not* admit a characterization in terms of $\mathrm{br}(\Gamma)$; this was shown by Pemantle and Steif [27]. Bounds for $\beta_c(\Gamma, q)$ that only depend on $\mathrm{br}(\Gamma)$ and on q can, however, be obtained using the standard comparison techniques for the random-cluster model reviewed in [13].

5 The fuzzy Potts model on complete graphs

In this section we treat the case of complete graphs. We start with precise definitions of the model and a detailed explanation of the limiting process for the conditional probabilities that was sketched in the introduction. The proofs are essentially self-contained but use some standard knowledge (whose main reference is Ellis and Wang [4]) on the infinite volume limit of the empirical distribution of the order parameter in the mean-field Potts model.

5.1 Mean-field Potts in finite volume N

For a positive integer q, the Gibbs measure $\pi^N_{q,\beta}$ for the q-state Potts model on the complete graph with N vertices at inverse temperature $\beta \geq 0$, is the probability measure on $\{1,\dots,q\}^N$ which to each $\xi \in \{1,\dots,q\}^N$ assigns probability

$$\pi^N_{q,\beta}(\xi) = \frac{1}{Z^N_{q,\beta}} \exp\left(\frac{\beta}{N} \sum_{1\leq x\neq y\leq N} I_{\{\xi(x)=\xi(y)\}}\right) . \tag{23}$$

Here $Z^N_{q,\beta}$ is the normalizing constant. Note that this definition slightly deviates from the definition (1) by the factor $1/N$ appearing in the exponential. Such a convention is appropriate because, clearly, the interaction must be chosen depending on the size of the graph in a mean-field model. This definition of the finite volume Gibbs-measures is standard in the literature; see e.g. [4].

5.2 Mean-field fuzzy Potts in finite volume N

The mean-field fuzzy Potts measure in finite volume N is then defined in the same way as it is defined on every graph. To be explicit, fix q, β and the spin-partition $(r_1,\dots,r_s)$ as above. Let X be the $\{1,\dots,q\}^N$-valued random object distributed according to the mean field finite volume Gibbs measure $\pi^N_{q,\beta}$. Take Y to be the $\{1\dots,s\}^N$-valued random object obtained from X by the site-wise application of the spin-partitioning as in (2). Then $\mu^N_{q,\beta,(r_1,\dots,r_s)}$ is the probability measure on $\{1,\dots,s\}^N$ which describes the distribution of Y.

5.3 Gibbsianness vs. non-Gibbsianness for mean-field models: continuity vs. discontinuity of limiting conditional probabilities

We start with some general remarks about mean-field models to explain the appropriate analogue of non-Gibbsianness in more detail than we did in the introduction. To begin with, the following lemma makes explicit that we can always describe the single-site conditional probabilities of the finite volume Gibbs measures of a mean-field model in terms of a *single-site kernel* from the empirical distribution vector of the conditioning to the single-site state space. It is the infinite volume limit of this kernel that shall then be considered in the analysis of the model.

So, suppose that S is a finite set (local spin space) and for any N we are given an exchangeable (that is permutation-invariant) measure μ^N on S^N. This permutation invariance is certainly true for the mean-field Potts model. Moreover it carries over trivially to the fuzzy Potts model. This is clear since the distribution of the latter is simply obtained by an application of the same map to the spin variable at each site.

In a general context, denote by $\mathcal{P} = \{(p_i)_{i\in S}, 0 \leq p_i \leq 1, \sum_{i\in S} p_i = 1\}$ the space of probability vectors on the set S. We use the obvious short notation $x^c = \{1,\dots,N\}\setminus\{x\}$.

Lemma 5.1 *For each N there is a probability kernel $Q^N : S \times \mathcal{P} \to [0,1]$ from $\mathcal{P}$ to the single-site state space S such that the single-site conditional expectations at any site x can be written in the form*

$$\mu^N\big(X(x) = i \big| X(x^c) = \eta\big) = Q^N\big(i\big|(n_j)_{j\in S}\big) . \tag{24}$$

Here $n_j = \frac{1}{N-1}\#\big(1 \le y \le N, y \ne x, \eta(x) = j\big)$ *is the fraction of sites for which the spin-values of the conditioning are in the state* $j \in S$.

Proof: By exchangeability it is clear that the right hand side of (24) depends on the sets $\big\{1 \le y \le N, y \ne x, \eta(y) = j\big\}$, for all $j \in S$, only through their size. Equivalently we may express this dependence in terms of the empirical distribution $(n_j)_{j\in S}$. $\square$

In turn, the knowledge of the kernel Q^N uniquely determines the measure μ^N. This is clear since the knowledge of all one-site conditional probabilities of finitely many random variables uniquely determines the joint distribution. So we may as well consider the Q^Ns as the basic objects and regard them as the starting point of the definition of a mean field model. This is of course only meaningful if the Q^Ns are related to each other in a meaningful way.

Let us turn now to the concrete case of the mean-field Potts model to point out two very simple observations that shall serve as a motivation of our further investigation. In this case we have directly from the definition (23) the explicit formula

$$Q^N_{q,\beta}\big(i\big|(n_j)_{1\le j\le q}\big) = \frac{\exp\big(\beta(1-\frac{1}{N})n_i\big)}{\sum_{j=1}^q \exp\big(\beta(1-\frac{1}{N})n_j\big)}\,. \tag{25}$$

We note the following.

(i) $Q^N_{q,\beta}$ converges to $Q^\infty_{q,\beta} = \frac{\exp(\beta n_i)}{\sum_{j=1}^q \exp(\beta n_j)}$ when N tends to infinity. Indeed, the trivial $1/N$-factor appearing in (25) could of course even be removed by a harmless redefinition of the model that would lead to the same infinite volume behavior of the Gibbs measures, making all $Q^N_{q,\beta}$ identical.

(ii) The limiting kernel $Q^\infty_{q,\beta}$ is a *continuous function* of the probability vector $(n_j)_{1\le j\le q}$, as a function on $\mathbb{R}^q$.

The existence of the infinite volume limit (i) is a minimal ingredient for the definition of a mean-field model. Assuming this we can talk about limiting or "infinite volume" conditional probabilities. Then, *continuous dependence of the limiting conditional probability* as it is stated in (ii) is the obvious analogue to the continuous dependence of the conditional expectation of a lattice model on the conditioning with respect to product topology.

So, properties (i) and (ii) are the analogues of a proper Gibbsian structure for mean-field models. "Non-Gibbsianness" may then manifest itself by the failure of (ii) at certain points of discontinuity. The reader may find a number of examples of this in [21]. After these introductory remarks we will show in the following that discontinuities in fact occur for the mean-field fuzzy Potts model, for certain values of the parameters, and discuss them in detail.

5.4 Conditional probabilities for fuzzy Potts in finite volume

Let us use the following notation for the single-site probability kernel that describes the conditional probabilities of the fuzzy model.

$$\mu^N_{q,\beta,(r_1,\dots,r_s)}\Big(Y(x) = k\big|Y(x^c) = \eta\Big) =: Q^N_{q,\beta,(r_1,\dots,r_s)}\big(k\big|(n_l)_{1\le l\le s}\big)\,. \tag{26}$$

where $n_l = \frac{1}{N-1}\#\big(1 \le y \le N, y \ne x, \eta(x) = l\big)$, for $l = 1, \dots, s$ is the empirical distribution of fuzzy spin-values in the conditioning.

Now, it is not difficult to derive an explicit expression in terms of expectations with respect to ordinary mean-field Potts measures, having the number of states given by the sizes of the classes r_l. Clearly, the infinite volume analysis relies on this result.

Proposition 5.2 *For each finite N we have the representation*

$$Q^N_{q,\beta,(r_1,\dots,r_s)}\big(k\big|(n_l)_{1\le l\le s}\big) = \frac{r_k\, A(\beta_k, r_k, N_k)}{\sum_{l=1}^s r_l\, A(\beta_l, r_l, N_l)} \tag{27}$$

where

$$A(\tilde\beta, r, M) \equiv \pi^M_{r,\tilde\beta}\Big(\exp\Big(\frac{\tilde\beta}{M}\sum_{x=1}^M I_{X(x)=1}\Big)\Big)\,,$$

$$N_k = (N-1)n_k\,,$$

and

$$\beta_k = \frac{\beta N_k}{N} = \beta\Big(1 - \frac{1}{N}\Big)n_k\,.$$

Remark: In particular we have $A(\tilde\beta, r = 1, N) = e^{\tilde\beta}$. From this we see immediately that the case of the original Potts model is recovered by setting all r_l equal to one.

Proof of Proposition 5.2: To compute the left hand side of (27) we may choose $x = 1$ and write

$$\begin{aligned}&\mu^N_{q,\beta,(r_1,\dots,r_s)}\Big(Y(1) = k\Big|Y([2,N]) = \eta([2,N])\Big)\\ &= \tfrac{1}{Norm.\,(\eta([2,N]))}\textstyle\sum_{\xi(1)\mapsto k}\sum_{\xi([2,N])\mapsto\eta([2,N])}\pi^N_{q,\beta}\Big(\xi(1),\xi([2,N])\Big)\,.\end{aligned}$$

Here we are summing over Potts configurations ξ that are mapped to the fuzzy Potts configuration (k,η) by means of the definition of the fuzzy model given in (2). The normalization has to be chosen such that summing over $k = 1, \dots, s$ yields one, for each fixed $\eta([2,N])$. The partition function appearing in the Gibbs-average on the right hand side only gives a constant that can be absorbed in the normalization, and so we need only consider

$$\begin{aligned}&\textstyle\sum_{\xi(1)\mapsto k}\sum_{\xi([2,N])\mapsto\eta([2,N])}\exp\Big(\frac{\beta}{N}\sum_{1\le x\ne y\le N} I_{\{\xi(x)=\xi(y)\}}\Big)\\ &\quad= \textstyle\sum_{\xi(1)\mapsto k}\sum_{\xi([2,N])\mapsto\eta([2,N])}\exp\Big(\frac{\beta}{N}\sum_{2\le y\le N} I_{\{\xi(1)=\xi(y)\}}\Big)\\ &\qquad\times\exp\Big(\frac{\beta}{N}\sum_{2\le x\ne y\le N} I_{\{\xi(x)=\xi(y)\}}\Big)\,.\end{aligned}$$

For fixed $\eta([2,N])$ we denote $\Lambda_l := \#\big\{x \in \{2,\dots,N\} : \eta(x) = l\big\}$. Then the sum in the last exponential decomposes over these sets, and we can rewrite the right hand side of the last equation in the form

$$\begin{aligned}&\textstyle\sum_{\xi(1)\mapsto k}\sum_{\xi([2,N])\mapsto\eta([2,N])}\exp\Big(\frac{\beta_k}{N_k}\sum_{z\in\Lambda_k} I_{\{\xi(z)=\xi(1)\}}\Big)\\ &\quad\times\textstyle\prod_{l=1}^s\exp\Big(\frac{\beta_l}{N_l}\sum_{x<y,x,y\in\Lambda_l} I_{\{\xi(x)=\xi(y)\}}\Big)\,.\end{aligned}$$

Next we divide the last line by the product of partition functions which is obtained by omitting the first exponential and the first sum. This only yields another $\eta([2,N])$-dependent constant. Using cancellations for the terms with $l \neq k$ we see in this way that

$$\mu^N_{q,\beta,(r_1,\dots,r_s)}\Big(Y(1)=k\Big|Y([2,N])=\eta([2,N])\Big)$$
$$= \frac{1}{Norm.\,(\eta([2,N]))}\sum_{\xi(1)\mapsto k}\pi^{N_k}_{k,\beta_k}\Big(\exp\Big(\frac{\beta_k}{N_k}\sum_{z=1}^{N_k}I_{\{X(z)=1\}}\Big)\Big),$$

which concludes the proof. □

5.5 Continuity vs. discontinuity of limiting conditional probabilities for fuzzy Potts

In this subsection we will derive an explicit formula for the limiting conditional probabilities of the fuzzy model. From this parts (i), (ii), (iii) of Theorem 1.2 follow.

We can build on well-known results about the limiting behavior of the empirical distribution of the mean-field Potts model. The main point is that it exhibits a first-order phase transition at a finite inverse critical temperature $\beta_c(q)$, for all $q \geq 3$. For the special case $q = 2$ (Ising model) there is only a second order phase transition. The following pieces of information about the mean-field Potts model can be found in [4, Thms 2.1. and 2.3]. To get a first impression of the model, the reader is advised to begin by focusing on the case $\beta \neq \beta_c(q)$, i.e. off the critical temperature.

Theorem 5.3 (Ellis, Wang) *Assume that $q \geq 3$, and define $\beta_c(q) := \frac{2(q-1)}{q-2}\log(q-1)$. Then we have the weak limit*

$$\lim_{N\uparrow\infty}\pi^N_{q,\beta}\Big(\frac{1}{N}\sum_{x=1}^{N}(I_{\{X(x)=1\}},\dots,I_{\{X(x)=q\}})\in\cdot\Big) \tag{28}$$
$$= \begin{cases} \delta_{\frac{1}{q}(1,1,\dots,1)}, & \text{if } \beta<\beta_c(q)\\ \frac{1}{q}\sum_{\nu=1}^{q}\delta_{u(\beta,q)\,e_\nu+\frac{1-u(\beta,q)}{q}(1,1,\dots,1)}, & \text{if } \beta>\beta_c(q)\\ \lambda_0(q)\delta_{\frac{1}{q}(1,1,\dots,1)}+\frac{1-\lambda_0(q)}{q}\sum_{\nu=1}^{q}\delta_{u(\beta_c(q),q)\,e_\nu+\frac{1-u(\beta_c(q),q)}{q}(1,1,\dots,1)} & \text{if } \beta=\beta_c(q)\,, \end{cases}$$

where e_i is the unit vector in the i'th coordinate direction of $\mathbb{R}^q$.

The quantity $u(\beta,q)$ is well defined for $\beta \geq \beta_c(q)$. It is the largest solution of the mean field equation

$$u = \frac{1-e^{-\beta u}}{1+(q-1)e^{-\beta u}} \tag{29}$$

and obeys the following properties: It is strictly increasing in β, and we have $u(q,\beta_c(q)) = \frac{q-2}{q-1}$. The constant appearing at the critical point obeys the strict inequality $0 < \lambda_0(q) < 1$.

Some comments are in order: Obviously, $u(\beta,q)$ plays the role of an order parameter. Now, for $\beta > \beta_c(q)$ the system is in a symmetric linear combination of ν-like states. The limiting empirical distribution becomes the equidistribution on the possible spin values for $\beta < \beta_c(q)$. It jumps at the critical point for $q \geq 3$. At the critical point itself there is a non-trivial linear combination between both types of measures.

To feel comfortable with the mean-field equation (29) the reader may note that it is obtained from the equations $n_i = \frac{\exp(\beta n_i)}{\sum_{j=1}^q \exp(\beta n_j)}$ for $i = 1, \ldots, q$ with the following ansatz: Denote by i the index with the largest n_j. Assume that n_j is independent of j, for $j \neq i$, and put $u = n_i - n_j$ for some $j \neq i$.

Let us mention that the results of Theorem 5.3 can be obtained by a Gaussian transformation and saddle point estimates on the resulting integrals (all of which is omitted here). At the critical point a little care is needed: To obtain the proper value of the constant $\lambda_0(q)$ a Gaussian approximation around the minima and estimates showing positive curvature are needed.

The well-known case of the mean field Ising model $q = 2$ can be recovered from the theorem by taking the formal limit $q \downarrow 2$ in the explicit formula for $\beta_c(q)$ and noting that $u(q, \beta_c(q)) = 0$. So (28) describes a second order transition in that case.

The following explicit formula for the limiting conditional probabilities of the fuzzy model now follows easily from our finite volume representation of the conditional probabilities given in Proposition 5.2 and the known limiting statement of Theorem 5.3.

Theorem 5.4 *We have*

$$\lim_{N\uparrow\infty} Q^N_{q,\beta,(r_1,\ldots,r_s)}\left(k \middle| (n_l)_{1\leq l\leq s}\right) = \frac{C(\beta n_k, r_k)}{\sum_{l=1}^s C(\beta n_l, r_l)}$$

whenever $n_k \neq \beta_c(r_k)/\beta$ for all k with $r_k \geq 3$. Here

$$C(\tilde\beta, r) = \exp\Big(\frac{\tilde\beta}{r}\Big) \times \begin{cases} r, & \text{if } \tilde\beta < \beta_c(r) \\ \exp\Big(\frac{\tilde\beta(r-1)u(\tilde\beta,r)}{r}\Big) + (r-1)\exp\Big(-\frac{\tilde\beta\, u(\tilde\beta,r)}{r}\Big), & \text{if } \tilde\beta > \beta_c(r)\,. \end{cases}$$

Proof of Theorem 5.4: Let $\tilde\beta \neq \beta_c(q)$. By Theorem 5.3 we have $\lim_{M\uparrow\infty} rA(\tilde\beta, r, M) = C(\tilde\beta, r)$. $\square$

Remark: Obviously this gives the right answer for $\beta = 0$ or in the case of the original Potts model (letting all r_l be equal to one). We see however that the limiting form of the conditional expectations has a nontrivial form in general. This expression has jumps for $n_l = \beta_c(r_l)/\beta$ whenever $r_l \geq 3$. (For matters of simplicity we state the result only outside these critical values.) Indeed, for $r \geq 2$ we have

$$C(\beta_c(r) \mp 0, r) = (r-1)^{\frac{2(r-1)}{r(r-2)}} \times \begin{cases} r \\ r(r-1)^{\frac{r-2}{r}} \end{cases}$$

which jumps for $r \geq 3$. (For $r = 2$ this expression has to be interpreted as the limit of the right hand side with $r \downarrow 2$.)

The reader should notice the following: First of all we have shown the pointwise existence of the limit

$$(n_l)_{1\leq l\leq s} \mapsto \lim_{N\uparrow\infty} Q^N_{q,\beta,(r_1,\ldots,r_s)}\left(k \middle| (n_l)_{1\leq l\leq s}\right).$$

The notion of "continuity of limiting conditional probabilities" that was introduced in Theorem 1.2 has the precise meaning of continuity of the right hand side as a function on the closed set $\mathcal{P}$ of s-dimensional probability vectors with respect to the ordinary Euclidean topology. From the explicit limiting formula given in the theorem and the well-known knowledge of the jumps of the order parameter the proof of the first three parts of our main theorem 1.2 is now immediate.

Proof of Theorem 1.2 (i),(ii),(iii): The points of discontinuity are precisely given by the values $n_k = \frac{\beta_c(r_k)}{\beta}$ for those k with $r_k \geq 3$ for which $\frac{\beta_c(r_k)}{\beta} < 1$. So (i) is immediate. To see (ii) and (iii) we use that $\beta_c(r)$ is an increasing function of r. $\square$

5.6 Typicality of continuity points – "almost sure Gibbsianness"

What can be said about the measure of the discontinuity points? We will answer this question now and prove the remaining part (iv) of Theorem 1.2. To start with, from Theorem 5.3 follows trivially by "contraction" that the typical values of the order parameter in the fuzzy model are as follows. (Recall that e_l is the unit vector in the l'th coordinate direction of $\mathbb{R}^s$.)

Corollary 5.5 *We have*

$$\lim_{N\uparrow\infty} \mu^N_{q,\beta,(r_1,\dots,r_s)}\Bigl(\frac{1}{N}\sum_{x=1}^{N}(I_{\{Y(x)=1\}},\dots,I_{\{Y(x)=s\}})\in\cdot\Bigr)$$
$$= \begin{cases} \delta_{\frac{1}{q}(r_1,r_2,\dots,r_s)} & \text{if } \beta<\beta_c(q)\\ \lambda_0(q)\delta_{\frac{1}{q}(r_1,r_2,\dots,r_s)} + \sum_{l=1}^{s}\frac{(1-\lambda_0(q))r_l}{q}\delta_{u(\beta,q)e_l+\frac{1-u(\beta,q)}{q}(r_1,r_2,\dots,r_s)} & \text{if } \beta=\beta_c(q)\\ \sum_{l=1}^{s}\frac{r_l}{q}\delta_{u(\beta,q)e_l+\frac{1-u(\beta,q)}{q}(r_1,r_2,\dots,r_s)} & \text{if } \beta>\beta_c(q)\,. \end{cases}$$

In other words, the values for the fuzzy densities n_l that occur with non-zero probability are: The values r_l/q in the high-temperature regime (including the critical point) and the two values

$$n^+(\beta,q,r_l) \equiv u(q,\beta) + \frac{1-u(q,\beta)}{q}r_l$$

and

$$n^-(\beta,q,r_l) \equiv \frac{1-u(q,\beta)}{q}r_l \quad \Bigl(\leq n_l^+(\beta,q,r_l)\Bigr)$$

in the low temperature regime (including the critical point).

Now, the non-trivial question is: Can it happen that these values coincide with the points of discontinuity of the limiting conditional probability, for certain choices of the parameter?

The following proposition tells us that this can never be the case, and so the points of discontinuity are always atypical. This immediately proves (iv) of Theorem 1.2. As we will see the proof of the proposition is elementary but slightly tricky; it makes use of specific properties of the solution of the mean-field equation. In that sense it is the most difficult part of our analysis of the mean field fuzzy Potts model.

Proposition 5.6 *Assume that $q > r \geq 2$.*

(i) *For the high-temperature range $\beta \leq \beta_c(q)$ we have*

$$\frac{r}{q} < \frac{\beta_c(r)}{\beta}\,.$$

(ii) *For the low-temperature range $\beta \geq \beta_c(q)$ we have that*

$$n^-(\beta, q, r) < \frac{\beta_c(r)}{\beta} < n^+(\beta, q, r)\,.$$

Remark: (i) says that that the typical density of each fuzzy class is too small to create a first order transition. The left inequality of (ii) says that the typical density of a fuzzy class not containing the predominant spin-value of the underlying Potts model is always too small to create a first order transition. The corresponding conditional Potts model is always in a high-temperature state. The right inequality of (ii) says that the typical density of the fuzzy class that contains the predominant spin-value of the underlying Potts model is always too big to create a first order transition. The corresponding conditional Potts model is always in a low-temperature state.

Proof: The claim (i) follows from that fact that $\frac{r}{q} < \frac{\beta_c(r)}{\beta_c(q)}$ for all $q > r$. This in turn is implied by the fact that $\frac{\beta_c(q)}{q}$ is decreasing in q. It is obvious that this holds for large enough q, by the explicit expression for $\beta_c(q)$. It is elementary to verify that it holds in fact for any $q \geq 2$.

Next we prove (ii). We show first the right inequality which is equivalent to

$$u(q, \beta) > \frac{q}{q-r}\frac{\beta_c(r)}{\beta} - \frac{r}{q-r}\,.$$

By Theorem 5.3 the order parameter $u(q, \beta)$ is an increasing function in β. The right hand side is decreasing in β. So it suffices to prove the inequality for $\beta = \beta_c(q)$. Using $u(q, \beta_c(q)) = \frac{q-2}{q-1}$ this can be put equivalently as

$$\beta_c(r) < \beta_c(q)\left(1 - \frac{q-r}{q(q-1)}\right)\,. \tag{30}$$

We will use now the elementary property that

$$\beta_c(q) < q, \quad \text{for all real } q > 2\,. \tag{31}$$

This implies also that $\beta_c(q)$ is concave because

$$\beta_c''(q) = \frac{-2q(q-2) + 4(q-1)\log(q-1)}{(q-2)^3(q-1)}$$

and the denominator is negative, by the last inequality.

In order to show (30) we note, by concavity that

$$\beta_c(r) \leq \beta_c(q) + \beta_c'(q)(r-q) \tag{32}$$

and show that the right hand side of (32) is strictly bounded from above by the right hand side of (30). But the latter statement is equivalent to

$$\beta_c'(q) > \beta_c(q)\frac{1}{q(q-1)}\,.$$

Computing the logarithmic derivative $\frac{\beta_c'(q)}{\beta_c(q)}$ we see that this is equivalent to

$$\frac{1}{q-1} - \frac{1}{q-2} + \frac{1}{(q-1)\log(q-1)} > \frac{1}{q(q-1)}\,.$$

This inequality in turn reduces after trivial computation to the statement (31) and this concludes the proof of the right inequality of (i).

Let us come to the proof of the left inequality of (ii). The claim says $\frac{1-u(q,\beta)}{q} r < \frac{\beta_c(r)}{\beta}$. Using the mean-field equation we may write

$$1 - u(q, \beta) = \frac{q}{e^{+\beta u(q,\beta)} + q - 1} \,.$$

So the claim is equivalent to

$$\beta \frac{r}{\beta_c(r)} < e^{+\beta u(q,\beta)} + q - 1 \,.$$

Now, the left hand side is increasing as a function of r, for $r \geq 2$. So the claim follows from

$$\beta \frac{q-1}{\beta_c(q-1)} < e^{+\beta u(q,\beta)} + q - 1$$

for all $q \geq 3$. Next we use again that the order parameter $u(q, \beta)$ is an increasing function of β. Thus the last inequality follows if we can show

$$\beta_c(q) \frac{q-1}{\beta_c(q-1)} < e^{+\beta_c(q) u(q,\beta_c(q))} + q - 1 \,.$$

We have $e^{+\beta_c(q)u(q,\beta_c(q)+0)} = (q-1)^2$ from the explicit expressions and so the last inequality is equivalent to

$$\frac{\beta_c(q)}{\beta_c(q-1)} < q \,.$$

It is elementary to verify from the explicit expression for $\beta_c(q)$ that this actually holds for all $q \geq 3$. $\square$

References

[1] Aizenman, M., Chayes, J.T., Chayes, L. and Newman, C.M. (1988) Discontinuity of the magnetization in one-dimensional $1/|x-y|^2$ Ising and Potts models, *J. Statist. Phys.* **50**, 1–40.

[2] Brightwell, G.R. and Winkler, P. (1999) Graph homomorphisms and phase transitions, *J. Combin. Theory B* **77**, 221–262.

[3] Chayes, L. (1996) Percolation and ferromagnetism on $\mathbf{Z}^2$: the q-state Potts cases, *Stoch. Proc. Appl.* **65**, 209–216.

[4] Ellis, R.S. and Wang, K.W. (1989) Limit theorems for the empirical vector of the Curie–Weiss–Potts model, *Stoch. Proc. Appl.* **35**, 59–79.

[5] van Enter, A.C.D. (1996) On the possible failure of the Gibbs property for measures on lattice systems, *Markov Proc. Relat. Fields* **2**, 209–224.

[6] van Enter, A.C.D., Fernández, R., den Hollander, F. and Redig, F. (2002) Possible loss and recovery of Gibbsianness during the stochastic evolution of Gibbs measures, *Comm. Math. Phys.* **226**, 101–130.

[7] van Enter, A.C.D., Fernández, R. and Kotecký, R. (1995) Pathological behavior of renormalization-group maps at high fields and above the transition temperature, *J. Statist. Phys.* **79**, 969–992.

[8] van Enter, A.C.D., Fernández, R. and Sokal, A.D. (1993) Regularity properties of position-space renormalization group transformations: Scope and limitations of Gibbsian theory, *J. Statist. Phys.* **72**, 879–1167.

[9] van Enter, A.C.D. and Lőrinczi, J. (1996) Robustness of the non-Gibbsian property: some examples, *J. Phys. A* **29**, 2465–2473.

[10] van Enter, A.C.D., Maes, C., Schonmann, R.H. and Shlosman, S. (2000) The Griffiths singularity random field, *On Dobrushin's way. From probability to statistical mechanics* (R. Minlos, Yu. Suhov and S. Shlosman, eds) pp 59–70, AMS.

[11] van Enter, A.C.D., Maes, C. and Shlosman, S. (2000) Dobrushin's program on Gibbsianity restoration: weakly Gibbs and almost Gibbs random fields, *On Dobrushin's way. From probability to statistical mechanics* (R. Minlos, Yu. Suhov and S. Shlosman, eds), pp 51–58, AMS.

[12] Fernández, R. and Pfister, C.-E. (1997) Global specifications and nonquasilocality of projections of Gibbs measures, *Ann. Probab.* **25**, 1284–1315.

[13] Georgii, H.-O., Häggström, O. and Maes, C. (2001) The random geometry of equilibrium phases, in *Phase Transitions and Critical Phenomena, vol 18* (C. Domb and J.L. Lebowitz, eds), pp 1–142, Academic Press, London.

[14] Griffiths, R.B. and Pearce, P.A. (1978) Position-space renormalization-group transformations: Some proofs and some problems, *Phys. Rev. Lett.* **41**, 917–920.

[15] Häggström, O. (1996) The random-cluster model on a homogeneous tree, *Probab. Th. Relat. Fields* **104**, 231–253.

[16] Häggström, O. (1996) Almost sure quasilocality fails for the random-cluster model on a tree, *J. Statist. Phys.* **84**, 1351–1361.

[17] Häggström, O. (1999) Positive correlations in the fuzzy Potts model, *Ann. Appl. Probab.* **9**, 1149–1159.

[18] Häggström, O. (2003) Is the fuzzy Potts model Gibbsian?, *Ann. Inst. H. Poincaré, Probab. Statist.* **39**, 891–917.

[19] Külske, C. (1999) (Non-) Gibbsianness and phase transitions in random lattice spin models, *Markov Proc. Relat. Fields* **5**, 357–383.

[20] Külske, C. (2001) Weakly Gibbsian representations for joint measures of quenched lattice spin models, *Probab. Th. Relat. Fields* **119**, 1–30.

[21] Külske, C. (2003) Analogues of non-Gibbsianness in joint measures of disordered mean field models, *J. Statist. Phys.* **112**, 1101–1130.

[22] Künsch, H., Geman, S. and Kehagias, A. (1995) Hidden Markov random fields, *Ann. Appl. Probab.* **5**, 577–602.

[23] Le Ny, A. (2000) Fractal failure of quasilocality for a majority rule transformation on a tree, *Lett. Math. Phys.* **54**, 11–24.

[24] Le Ny, A. (2002) Almost and fractal quasilocality for a majority rule on a tree, preprint.

[25] Lyons, R. (1989) The Ising model and percolation on trees and tree-like graphs, *Comm. Math. Phys* **125**, 337–353.

[26] Maes, C. and Vande Velde, K. (1995) The fuzzy Potts model, *J. Phys. A* **28**, 4261–4271.

[27] Pemantle, R. and Steif, J. (1999) Robust phase transitions for Heisenberg and other models on general trees, *Ann. Probab.* **27**, 876–912.

[28] Peres, Y. (1999) Probability on trees: an introductory climb, *Lectures on probability theory and statistics (Saint-Flour, 1997)*, pp 193–280.

[29] Schonmann, R.H. (1989) Projections of Gibbs measures may be non-Gibbsian, *Comm. Math. Phys.* **124**, 1–7.

[30] Spitzer, F. (1975) Markov random fields on an infinite tree, *Ann. Probab.* **3**, 387–398.

[31] Zachary, S. (1983) Countable state space Markov random fields and Markov chains on trees, *Ann. Probab.* **11**, 894–903.