Random Colorings of a Cayley Tree

Graham R. Brightwell *
Department of Mathematics
London School of Economics
Houghton St.
London WC2A 2AE England

Peter Winkler†
Bell Labs 2C-365
Lucent Technologies
700 Mountain Ave.
Murray Hill, NJ 07974 U.S.A.

January 9, 2000

Abstract

Probability measures on the space of proper colorings of a Cayley tree (that is, an infinite regular connected graph with no cycles) are of interest not only in combinatorics but also in statistical physics, as states of the antiferromagnetic Potts model at zero temperature, on the "Bethe lattice".

We concentrate on a particularly nice class of such measures which remain invariant under parity-preserving automorphisms of the tree. Making use of a correspondence with branching random walks on certain bipartite graphs, we determine when more than one such measure exists. The case of "uniform" measures is particularly interesting, and as it turns out, plays a special role.

Some of the results herein are deducible from previous work of the authors and by members of the statistical physics community, but many are new. We hope that this work will serve as a helpful glimpse into the rapidly expanding intersection of combinatorics and statistical physics.

1 Introduction

1.1 Uniform Proper Colorings

The *Cayley tree* T_r is uniquely defined by being connected, $r+1$-regular and cycle-free; thus T_1 is the two-way infinite path and T_2 the complete binary tree. A (proper) q-coloring of T_r can be thought of as a homomorphism φ from T_r to the complete graph K_q on q nodes, which we label $\{1, 2, \ldots, q\}$. (We call the nodes of T_r "sites" in accordance with physics tradition and to distinguish them from the nodes of the target graph K_q.)

What would it mean to choose such a coloring φ *uniformly* at random? If we were instead coloring a finite graph, there would be no difficulty in selecting φ with equal probability from the finitely many possibilities. Even on T_r, if the coloring were not constrained to be proper, uniformity would be easily achievable by choosing the color of each site uniformly and independently.

In the case at hand, however, we are reduced to agreeing on what properties a "uniform" probability measure on the space of colorings should have, and then asking whether there is a unique measure with these properties.

*Author's net address: g.r.brightwell@lse.ac.uk
†Author's net address: pw@lucent.com

We will begin by listing every property we can think of that makes sense for the infinite tree T_r and holds for the uniform distribution on colorings of any finite graph. Letting μ be a measure on proper colorings of T_r by q colors, we demand:

1. μ is a *Gibbs measure.* This means that for any finite subtree U of T_r, and any proper coloring ψ of U, the probability that φ agrees with ψ on U, given that $\varphi = \psi$ on the boundary of U, is the same for φ drawn from μ as it is for φ drawn uniformly from the set of all proper colorings of U. (The boundary ∂U of U is the set of those points in U which are adjacent to some point of $T_r \setminus U$.) For much more information about Gibbs measures, in far more general settings, the reader is referred to the excellent texts by Baxter and Georgii [1, 9].

2. μ is *simple.* This means that if we draw φ from μ, conditioned on any particular site u being colored by any fixed color j, then the restrictions of φ to each connected component of $T_r \setminus \{u\}$ are independent.

3. μ is *invariant.* This means that for any automorphism θ of T_r, $\mu \circ \theta = \mu$; in other words, if A is any measurable set of colorings, then $\mu(\{\varphi \circ \theta : \ \varphi \in A\}) = \mu(A)$.

In statistical physics, the Cayley tree T_r is frequently called a "Bethe lattice". Coloring by q colors with symmetric interactions produces the "q-state Potts model"; the model is "antiferromagnetic" if adjacent sites are discouraged (and "at zero temperature", forbidden) from having the same color. The Bethe lattice is not a statistical physicist's first choice as a setting for theorems on model behavior—a lattice such as a cubic grid, in some Euclidean space, is preferred—but it has yielded many useful insights. Although variational methods used on more "amenable" lattices do not always work on the Bethe lattice (see e.g. [6]), the absence of cycles permits direct methods for analysis such as those we employ below.

1.2 Constraint Graphs and Activities

We now generalize the setting in two ways. First, let us replace the target graph K_q by an arbitrary, finite "constraint graph" H. We can think of H as representing a generalized set of coloring constraints, namely that adjacent sites of T_r may receive colors i and j if and only if i and j are adjacent in H; if H has a loop at node i then color i may be used on both adjacent sites. Such a coloring will be called a *coloring by H* when confusion with ordinary proper coloring is threatened. The use of H is intended to model physical systems with hard constraints.

There are constraint graphs other than K_q of natural combinatorial interest, for instance the graph J consisting of nodes 0 and 1 with an edge connecting the two nodes and a loop at node 0. In any J-coloring of T_r the sites colored "1" form an independent set, and conversely. We are thus talking now about random independent sets instead of random proper colorings, and there is again a name for this model in statistical physics: the (discrete) *hard-core lattice gas model*, or "hard-core model" for short. The intuition is that the sites colored "1" are occupied by a gas molecule, and forces between molecules prevent occupation of two neighboring sites. (An elegant and accessible treatment of the discrete hard-core model can be found in [2].)

The second generalization replaces the uniform distribution with a "multiplicative" distribution, defined by positive reals $\lambda_1, \ldots, \lambda_n$ associated with the nodes of H. The *activity* λ_i of node i reflects the relative preference given to color i; in particular, if the graph G to be colored by H is finite, the probability of a particular coloring φ is proportional to

$$\prod_{u \in G} \lambda_{\varphi(u)} \ .$$

This implies in particular that if the colors of the neighbors of a site u are fixed in such a way that colors $j_1, \ldots, j_k$ are permissible at u, then the conditional probability that u is colored j_1 is $\lambda_{j_1}/(\lambda_{j_1} + \cdots + \lambda_{j_k})$. This condition then applies also when G is infinite, as a special case of the Gibbs condition.

To see that the multiplicative measures really are a natural extension of the uniform, consider for example the constraint graph which consists of a path on three nodes with a loop at the center node. Coloring uniformly by this H is equivalent to selecting an independent set and then coloring the nodes of that set by two colors; the distribution thus induced on uncolored independent sets is the same as that given by coloring by J, with activities satisfying $\lambda_1/\lambda_0 = 2$.

For the hard-core model, activity can be thought of as the result of applying pressure (positive or negative) to the gas, and is often called "fugacity".

For the Potts model, the activity of any single color is controlled by an "external field". The effect of the field, however, is understood by physicists to become infinite when the temperature reaches 0; thus non-uniform activity sets fall outside their framework at zero temperature. The result is that many of the questions we attack, motivated by our combinatorial point of view, are not explored e.g. in the analysis of the antiferromagnetic Potts model on the Bethe lattice found in Peruggi et al. [14, 15, 16].

A physicist might even argue that there is nothing special at all about the uniform distribution, for example in the hard-core model. However, if the constraint graph is symmetric then the uniform distribution (where all the activities are equal) preserves this symmetry; and in fact we will see that in critical cases, the uniform set of activities stands apart from all others.

1.3 Semi-invariance

In much more general settings than ours, a result of Dobrushin [7] states that there is always at least one Gibbs measure for a given "specification". The problem of determining when there is more than one such Gibbs measure is a central problem of statistical mechanics, often called the *DLR problem* after Dobrushin, Lanford and Ruelle (see e.g. [17]); when more than one Gibbs measure exists, there is said (by some) to be a *phase transition*.

The DLR problem for uniformly random colorings of T_r can be put entirely in finite combinatorial terms. Let T_r^n be an n-level r-branching tree, and choose a proper q-coloring of T_r^n uniformly at random. Let p_n be the maximum probability that the root is colored "red" given the colors of the leaves; then $p_n \to 1/q$ if and only if there is a unique uniform Gibbs measure for q-colorings of T_r. Roughly speaking, if there is a phase transition, then there is "long-range order" meaning that the color of a site can be influenced by the colors of sites arbitrarily far away. We give more details in Section 3.4.

We are particularly interested in Gibbs measures satisfying the nice conditions given in Section 1.1, but we would like sometimes to relax one of our conditions slightly as follows. A measure is said to be *semi-invariant* if it is invariant under *even* automorphisms of T_r, i.e. those automorphisms which preserve the parity of a site. (For specificity, we choose a root w of T_r which is deemed to be of parity 0 mod 2, the parity of all other points being equivalent mod 2 to their distance to w.)

The hard-core model on T_r has been extensively studied (see e.g. [10, 9, 4, 3]); for any r and any $\lambda := \lambda_1/\lambda_0$, there is a unique simple invariant Gibbs measure. However, for $\lambda > r^r/(r-1)^{r+1}$ there are two additional simple semi-invariant Gibbs measures, one with most of its independent set on even sites, the other mostly on odd.

We will see that the situation for proper colorings is similar in two respects: there is always a unique simple invariant Gibbs measure, and sometimes multiple simple semi-invariant Gibbs measures. For instance, in the case $q = r+1$ (q-coloring the q-regular tree), it will turn out that there is a unique simple semi-invariant Gibbs measure *only* when $\lambda_1 = \lambda_2 = \cdots = \lambda_q$.

Our results come from a very useful correspondence between invariant (and semi-invariant) simple Gibbs measures, and node-weighted branching random walks on a constraint graph.

2 Gibbs Measures

2.1 Homomorphisms

As noted T_r will always stand for the $(r+1)$-regular Cayley tree and H for a finite constraint graph, possibly with loops. If G is any finite or infinite graph (the "board") we denote the space of homomorphisms from G to H by $\mathrm{Hom}(G,H)$. Adjacency between sites of G or nodes of H is denoted by $\sim$, so that $\varphi \in \mathrm{Hom}(G,H)$ and $u \sim v$ in G imply $\varphi(u) \sim \varphi(v)$ in H.

Assume that a positive real activity λ_i is associated with each node i of H. If the board G is finite, we denote the resulting multiplicative probability distribution on $\mathrm{Hom}(G,H)$ by m_G; thus

$$m_G(\{\psi\}) := \frac{\prod_{u\in G} \lambda_{\psi(u)}}{\sum_{\varphi\in\mathrm{Hom}(G,H)} \prod_{u\in G} \lambda_{\varphi(u)}} .$$

When the board is infinite, there is a bit more work to be done in determining when a measure on $\mathrm{Hom}(G,H)$ meets the "specifications" given by the set of activities on H. To begin with, a finite set U of sites of G will be called a *patch* and we will deliberately confuse U with the subgraph of G induced by U. As before, the *boundary* ∂U of U will be the set of sites in U which are adjacent to at least one site of $G\backslash U$.

If U is a patch and $\varphi \in \mathrm{Hom}(G,H)$, we denote by $\varphi\restriction U$ the restriction of φ to U; thus $\varphi\restriction U \in \mathrm{Hom}(U,H)$. If A is an event of the form

$$A = \{\varphi \in \mathrm{Hom}(G,H) : \ \varphi\restriction U \in F\}$$

for some patch U and some $F \subset \mathrm{Hom}(U,H)$, then we call A a "patch event". We equip $\mathrm{Hom}(G,H)$ with the σ-field (denoted by $\mathcal{F}$) generated by the patch events, and consider henceforth only measures μ on $\langle \mathrm{Hom}(G,H), \mathcal{F}\rangle$ such that $\mu(\mathrm{Hom}(G,H)) = 1$.

Let $U_1 \subset U_2 \subset \cdots$ be a nested sequence of patches whose union is G. We can define a compact topology on the space $\mathcal{M}$ of all probability measures on $\mathrm{Hom}(G,H)$ via the metric ρ defined as follows:

$$\rho(\mu,\nu) = \sum_{i=1}^{\infty} 2^{-i} \|\mu\restriction U_i, \nu\restriction U_i\|$$

where $\mu\restriction U$ is the measure induced by μ on $\mathrm{Hom}(U,H)$, and $\|\cdot\|$ is the total variation metric between probability distributions (a.k.a. half the ℓ^1 distance). Of course this topology is not sensitive to the choice of the U_i's, but it is useful to fix them anyway.

A measure μ on $\mathrm{Hom}(G,H)$ is said to be a *Gibbs measure* if for any patch $U \subset G$, and almost every $\psi \in \mathrm{Hom}(G,H)$,

$$\Pr_{\mu}\left(\varphi\restriction U = \psi\restriction U \,\middle|\, \varphi\restriction((G\setminus U)\cup\partial U) = \psi\restriction((G\setminus U)\cup\partial U)\right) = \Pr_{m_U}\left(\varphi\restriction U = \psi\restriction U \,\middle|\, \varphi\restriction\partial U = \psi\restriction\partial U\right) .$$

It suffices to verify the Gibbs condition for the U_i's in our nested sequence; e.g. in the case $G = T_r$, for U_i equal to the set of sites at distance at most i from the root.

It may be possible for a measure to satisfy the Gibbs condition in a trivial and somewhat unsatisfactory way. For example, suppose we are q-coloring T_r (with root w) for some $q \le r+1$,

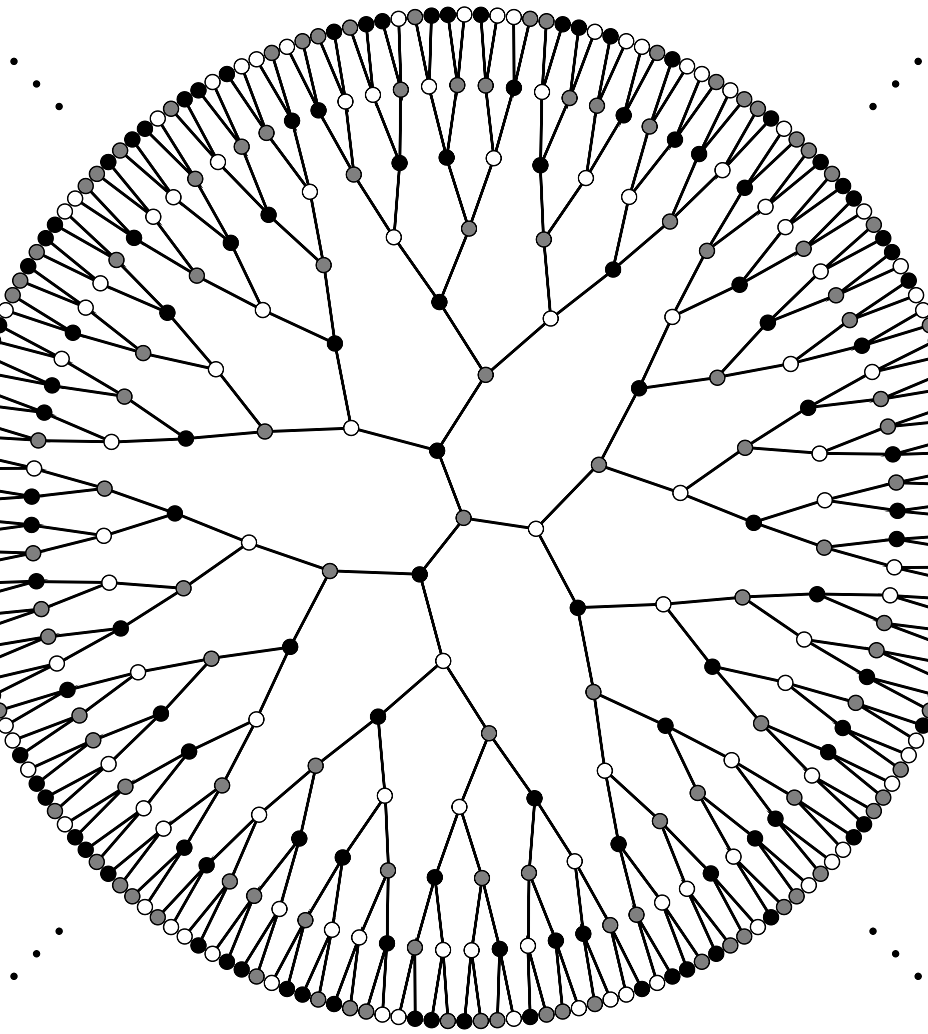

Figure 1: A frozen 3-coloring of the complete binary tree

and let ψ be any fixed coloring in which the children of every node exhibit all colors other than the color of the parent. Let μ be the measure which assigns probability 1 to ψ. Then the colors of ψ on any ∂U_i force the colors on ∂U_{i-1} and so on all the way to the root, thus matching m_U regardless of the activities. Thus μ is a ("frozen") Gibbs measure, and is also vacuously simple—but not invariant or semi-invariant. A frozen state of $\mathrm{Hom}(T_2, K_3)$ is illustrated in Fig. 1; for more about frozen Gibbs measures the reader is referred to [5]. (Note that frozen Gibbs measures can only occur when adjacent sites are absolutely forbidden to have the same color; that is, appropriately, at zero temperature. This accounts for their absence from most of the statistical physics literature.)

Normally the activities affect the definition of Gibbs measure via their tacit role in m_U. The remaining properties of measures that we use—simplicity, invariance and semi-invariance—are defined in the previous section for the case $G = T_r$ (which is all we shall need) and do not depend on the activities.

2.2 Simple Invariant Gibbs Measures

Simple invariant Gibbs measures on $\mathrm{Hom}(T_r, H)$ for arbitrary H were studied in [4], and the results can be applied directly to the case $H = K_q$. We need only this half of the main theorem of [4]:

Theorem 2.1. *Let H be a connected graph which satisfies the following two conditions:*

(a) *Every looped node of H is adjacent to all other nodes of H;*

(b) *With its loops deleted, H is a complete multipartite graph.*

Then for every set of activities on H, and every $r \geq 1$, there is a unique simple invariant Gibbs measure on the space $\mathrm{Hom}(T_r, H)$.

Since the graphs K_q (as well as J) satisfy the two conditions, we conclude that for the scope of this paper, we always have a unique simple invariant Gibbs measure.

The proof of the main theorem in [4] relied strongly on the fact that simple invariant Gibbs measures on $\mathrm{Hom}(T_r, H)$ are relatively tangible objects, obtainable by branching random walks on H. The measures are obtained as follows.

Let H be arbitrary but connected, and give each node i of H a positive real weight w_i. Let the probability distribution π on the nodes of H be defined by letting π_i be proportional to $w_i \sum_{j \sim i} w_j$ (π is then the stationary distribution for a random walk on H weighted by the w_i). We can (randomly) determine a coloring φ of T_r by H as follows:

First, drop an amoeba onto H which lands on node i with probability π_i; this i becomes the color of the root w of T_r.

Second, the amoeba divides into $r+1$ baby amoebas and each takes an independent random step to a neighbor j of i, with probabilities proportional to w_j. These j's become the colors of w's $r+1$ children.

Next, each baby amoeba divides into r new amoebas, and all take independent random steps to neighbors, always in accordance with the weights. These grand-babies determine the colors of the $r(r+1)$ grandchildren of w.

The process continues, each new amoeba splitting r ways at the next step, and every site of T_r is thus colored. The following theorem from [4] is critical, though quite straightforward to prove.

Theorem 2.2. *The above process produces a simple invariant Gibbs measure on* $\mathrm{Hom}(T_r, H)$ *for activities given by:*

$$\lambda_i = \frac{w_i}{\left(\sum_{j \sim i} w_j\right)^r} \ .$$

Moreover, every simple invariant Gibbs measure on $\mathrm{Hom}(T_r, H)$ *can be obtained in this way.*

Results similar to, but not quite the same as, Theorem 2.2 had earlier been proved by Peruggi [13] and Zachary [20].

The map which transforms the weights into activities is surjective and, in the cases considered in this paper, injective as well. We will see that things get much more interesting when the invariance condition is relaxed to semi-invariance.

2.3 Constraint Graphs and Their Doubles

A graph H is said to be *ergodic* if it is connected and not bipartite (note that a graph containing a loop cannot be bipartite). Thus the graphs K_q for $q > 2$, and the graph J above associated with the hard-core model, are ergodic.

Given an ergodic graph H on nodes $1, 2, \ldots, q$ we form its bipartite "double", denoted $2H$, as follows: the nodes of $2H$ are $\{1, 2 \ldots, q\} \cup \{-1, -2, \ldots, -q\}$ with an edge between i and j just when $i \sim -j$ or $-i \sim j$ in H. Note that $2H$ is loopless; a loop at node i in H becomes the edge $\{-i, i\}$ in $2H$.

A coloring ψ of T_r by $2H$ induces a coloring $|\psi|$ by H via $|\psi|(v) = |\psi(v)|$. In the reverse direction, a coloring φ of T_r by H may be transformed to a coloring $\bar{\varphi}$ by $2H$, by putting $\bar{\varphi}(v) = \varphi(v)$ for even sites $v \in T_r$ and $\bar{\varphi}(v) = -\varphi(v)$ for odd v.

Let $\lambda = (\lambda_1, \ldots, \lambda_n)$ be a set of activities for H and suppose that μ is a simple invariant Gibbs measure on $\mathrm{Hom}(T_r, H)$ corresponding to λ. From μ we can obtain a simple invariant Gibbs measure $\bar{\mu}$ on $\mathrm{Hom}(T_r, 2H)$ by selecting φ from μ, and flipping a fair coin to decide between $\bar{\varphi}$ (as defined above) and $-\bar{\varphi}$. Obviously $\bar{\mu}$ yields the activity set $\bar{\lambda}$ on $2H$ given by $\bar{\lambda}_i = \lambda_{|i|}$. Furthermore, the weights on H which produce μ extend to $2H$ by $w_{-i} = w_i$.

Conversely, suppose ν is a simple invariant Gibbs measure on $\mathrm{Hom}(T_r, 2H)$ whose activity set satisfies $\lambda_{-i} = \lambda_i$ for each i. Then the measure $|\nu|$, obtained by choosing ψ from ν and taking its absolute value, is certainly an invariant Gibbs measure on $\mathrm{Hom}(T_r, H)$ for $\lambda \restriction \{1, \ldots, q\}$, but is it simple?

In fact, if the weights on $2H$ which produce ν do not satisfy $w_{-i} = cw_i$, then $|\nu|$ will fail to be simple. To see this, observe that if the weights are not proportional then there are nodes $i \sim j$ of H such that $p_{-i,-j} \neq p_{i,j}$ in the random walk on $2H$. Suppose that $|\psi|$ is conditioned on the color of the root w of T_r being fixed at i, and let x and y be distinct neighbors of w. Set $\alpha = \Pr(\psi(w) = i \mid |\psi(w)| = i)$. Then

$$\Pr(|\psi|(x) = j) = (1 - \alpha)p_{-i,-j} + \alpha p_{i,j}$$

but

$$\Pr(|\psi|(x) = j \wedge |\psi|(y) = j) = (1 - \alpha)p_{-i,-j}^2 + \alpha p_{i,j}^2 > \Pr(|\psi|(x) = j)^2$$

so the colors of x and y are not independent given $|\psi|(w)$.

However, we can recover simplicity at the expense of one bit worth of symmetry. Let ν^+ be ν conditioned on $\psi(u) > 0$, and define ν^- similarly. Then $|\nu^+|$ and $|\nu^-|$ are essentially the same as ν^+ and ν^-, respectively, and all are simple; but these measures are only semi-invariant.

On the other hand, suppose μ is a simple semi-invariant Gibbs measure on $\mathrm{Hom}(T_r, H)$. Let θ be a parity-reversing automorphism of T_r and define $\mu' := \mu \circ \theta$, so that $\frac{1}{2}\mu + \frac{1}{2}\mu'$ is fully invariant (but generally no longer simple). However, $\nu := \frac{1}{2}\bar{\mu} + \frac{1}{2}(-\overline{\mu'})$ is a simple *and* invariant Gibbs measure on $\mathrm{Hom}(T_r, 2H)$, thus given by a node-weighted random walk on $2H$. We can recover μ as ν_+, hence:

Theorem 2.3. *Every simple semi-invariant Gibbs measure on* $\mathrm{Hom}(T_r, H)$ *is obtainable from a node-weighted branching random walk on* $2H$*, with its initial state drawn from the stationary distribution on positive nodes of* $2H$.

Suppose, instead of beginning with a measure, we start by weighting the nodes of $2H$ and creating a Gibbs measure as in Theorem 2.3. Suppose the activities of the measure are $\{\lambda_i : i = \pm 1, \ldots, \pm q\}$. By identifying color $-i$ with i for each $i > 0$, we create a measure on H-colorings, but this will not be a Gibbs measure unless it happens that $\lambda_{-1}, \ldots, \lambda_{-q}$ are proportional to $\lambda_1, \ldots, \lambda_q$.

We could assure this easily enough by making the weights proportional as well, e.g. by $w_{-i} = w_i$; then the resulting measure on $\mathrm{Hom}(G, H)$ could have been obtained directly by applying these weights to H, and is thus a fully invariant simple Gibbs measure. To get new, semi-invariant Gibbs measures on $\mathrm{Hom}(G, H)$, we must somehow devise weights for $2H$ such that $w_{-i} \not\propto w_i$ yet $\lambda_{-i} \propto \lambda_i$.

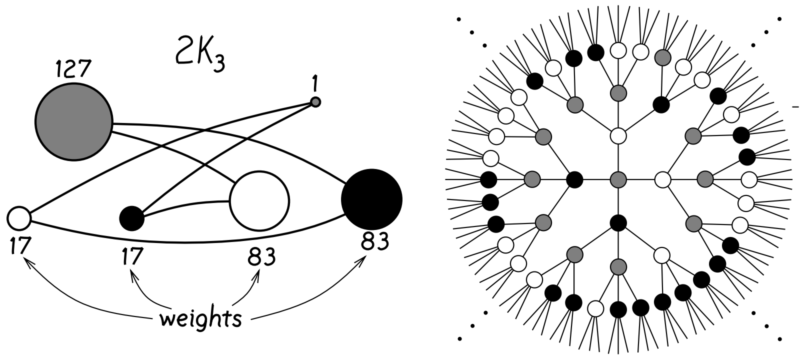

Figure 2: An asymmetric phase for 3-coloring a 3-branching tree

Restated with slightly different notation, simple semi-invariant Gibbs measures are in 1–1 correspondence with solutions to the "fundamental equations"

$$\lambda_i = \frac{u_i}{\left(\sum_{j\sim i} v_j\right)^r} = \frac{v_i}{\left(\sum_{j\sim i} u_j\right)^r}$$

for $i = 1, \ldots, q$. Such a solution will be invariant if $u_i = v_i$ for each i.

Figure 2 illustrates a semi-invariant, but not invariant, simple Gibbs measure for uniform 3-colorings of T_3. Approximate weights of the nodes of $2H = 2K_3$ are given along with part of a sample coloring drawn from this measure. Additional measures may be obtained by permuting the colors or by making all the weights equal (invariant case).

3 Results

3.1 Semi-invariant Measures: General r and q

We are now in a position to begin answering the key question of this paper: for which r and q, and which sets of activities, is there a unique simple semi-invariant Gibbs measure on $\mathrm{Hom}(T_r, K_q)$?

Let us dispose of the less interesting cases first. If $q = 2$ there are only two possible colorings of T_r, and every measure on $\mathrm{Hom}(T_r, K_2)$ is a convex combination of the two; all of these are simple semi-invariant Gibbs measures but only the $\frac{1}{2}$–$\frac{1}{2}$ mix is fully invariant.

The Cayley tree T_1 is a doubly-infinite path, known for not exhibiting phase transitions even in much more general circumstances than ours. For all $q > 2$ and all sets of activities for K_q, the only Gibbs measure of any kind on $\mathrm{Hom}(T_1, K_q)$ is the one given by a (non-branching) stationary node-weighted random walk on K_q. The weights are uniquely determined by the activities (see [4] for a direct proof).

Henceforth we will assume that $r \geq 2$ and $q \geq 3$. When $q < r+1$, we will see that all choices of activities including the uniform set yield multiple simple semi-invariant Gibbs measures.

When $q > r+1$, there is only one simple semi-invariant Gibbs measure for the uniform set of activities, but multiple simple semi-invariant Gibbs measures for some other choices of activities.

The critical case is at $q = r+1$, that is, when the number of colors is equal to the degree of the Cayley tree. Here we will show that there are multiple simple semi-invariant Gibbs measures for all activities *except* the uniform case, where there is just one.

If we consider all possible Gibbs measures, we find frozen measures (such as the ones mentioned above) and others whenever $q \leq r+1$. We believe that when $q > r+1$ and the activities are equal, the unique simple semi-invariant Gibbs measure is in fact the only Gibbs measure of any kind; we are able to prove this, however, only for $q > cr$, with fixed $c > 1$.

3.2 Multiple Simple Semi-invariant Gibbs Measures

Our next task is to show the existence of multiple simple semi-invariant Gibbs measures for certain values of q and r and certain ranges of the activities.

Theorem 3.1. *There are multiple simple semi-invariant Gibbs measures for* $\mathrm{Hom}(T_r, K_q)$ *in the following cases:*

- $q < r+1$, *for all sets of activities;*
- $q \leq r+1$, *for all non-uniform sets of activities;*
- *any* $q, r > 1$, *for some non-uniform sets of activities.*

The uniform case with $q < r+1$ is contained in the work of Peruggi, di Liberto and Monroy [16], where they find phase transitions of several types at zero temperature and zero external field. The other cases are in line with their work but, as we mentioned earlier, the case of hard constraints and finite variation from uniform activities does not have an exact counterpart in the physicists' setting.

Proof. We proceed indirectly, by showing that, in certain circumstances, the solution to a certain optimization problem is an "asymmetric" solution to the fundamental equations, and so corresponds to a simple Gibbs measure that is semi-invariant but not invariant. There will also always be a simple invariant Gibbs measure, and of course another simple semi-invariant Gibbs emasure can be obtained by switching the parity. The quantities involved in our optimization problem do not seem to be equivalent to any of the usual physical parameters such as pressure or mean free energy.

Given any non-negative real vectors $\lambda = (\lambda_1, \ldots, \lambda_q)$, $\mathbf{u} = (u_1, \ldots, u_q)$ and $\mathbf{v} = (v_1, \ldots, v_q)$, set

$$G(\mathbf{u}, \mathbf{v}) = \sum_{j \neq k} u_j v_k \ ; \quad H_\lambda(\mathbf{u}) = \sum_{j=1}^{q} \frac{u_j^{1+1/r}}{\lambda_j^{1/r}} \ .$$

For any fixed $q \geq 3$, $r \geq 2$ and vector $\lambda > 0$, consider the problem of maximizing $G(\mathbf{u}, \mathbf{v})$ over non-negative $\mathbf{u}$, $\mathbf{v}$, subject to the constraints $H_\lambda(\mathbf{u}) \leq 1$, $H_\lambda(\mathbf{v}) \leq 1$.

Note that the constraints define a compact region of $\mathbb{R}^{2q}$, and that G is continuous over this region, so the maximum is attained.

Suppose that some variable u_j or v_j is zero at the maximum. Then there are some k, l distinct from j with $u_j = 0$, $v_k > 0$, and $u_l > 0$ (or similarly with the roles of $\mathbf{u}$ and $\mathbf{v}$ interchanged). Then $\partial G / \partial u_j \geq v_k > 0$, $\partial H_\lambda(\mathbf{u}) / \partial u_j = 0$, while $\partial H_\lambda(\mathbf{u}) / \partial u_l > 0$ at the supposed maximum. Hence it is possible, by increasing u_j and decreasing u_l, to increase G while holding $H_\lambda(\mathbf{u})$ fixed—a contradiction.

Therefore all the variables are strictly positive at the maximum, and clearly we have equality in the constraints $H_\lambda(\mathbf{u})$, $H_\lambda(\mathbf{v}) \leq 1$. This means that the maximum will be a stationary point of the Lagrangian

$$L(\mathbf{u}, \mathbf{v}, \mu_1, \mu_2) = G(\mathbf{u}, \mathbf{v}) - \mu_1 \left(H_\lambda(\mathbf{u}) - 1 \right) - \mu_2 \left(H_\lambda(\mathbf{v}) - 1 \right) \ .$$

Setting the partial derivatives of L equal to 0 gives us:

$$\begin{aligned} \textstyle\sum_{k\neq j} v_k &= \mu_1 \tfrac{r+1}{r} \left(\tfrac{u_j}{\lambda_j}\right)^{1/r} \qquad (j=1,\ldots,q)\ , \\ \textstyle\sum_{k\neq j} u_k &= \mu_2 \tfrac{r+1}{r} \left(\tfrac{v_j}{\lambda_j}\right)^{1/r} \qquad (j=1,\ldots,q)\ , \end{aligned}$$

together with the constraints $H_\lambda(\mathbf{u}) = H_\lambda(\mathbf{v}) = 1$. Solutions to this system of equations are in 1–1 correspondence with solutions to the fundamental equations, and therefore with simple semi-invariant Gibbs measures for the set of activities λ.

We know that there is one invariant solution, and therefore a stationary point with $u_j = v_j$ for all j and $\mu_1 = \mu_2 = \mu$. Let w_j be the common value of u_j and v_j at this point. If this is the only solution to the fundamental equations, then this must be the solution to our optimization problem.

Suppose that $w_1 \geq w_2 \geq \cdots \geq w_q$ (this is equivalent to $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_q$). Suppose also that either $q < r+1$ or $q = r+1$ and the λ_j (hence the w_j) are not all equal. Observe then that, in either case, $\sum_{j\neq 1} w_j \leq (q-1)w_2 \leq rw_2$, and $\sum_{j\neq 2} w_j \leq (q-1)w_1 \leq rw_1$. Furthermore, equality is not possible in the latter inequality, and we can choose $\gamma > 0$ such that

$$2 > \frac{(1+\gamma)}{r} \left(\frac{\sum_{j\neq 2} w_j}{w_1} + \frac{\sum_{j\neq 1} w_j}{w_2} \right).$$

In the case $q > r+1$, we can find weights w_j, and $\gamma > 0$, satisfying the above inequality. Such weights define a set of activities, and it is this set for which we claim that there are multiple simple semi-invariant Gibbs measures.

Now choose a small $\delta > 0$, set

$$\delta_1 = \delta \left(\frac{w_2}{\lambda_2}\right)^{1/r}, \quad \delta_2 = \delta \left(\frac{w_1}{\lambda_1}\right)^{1/r},$$

and consider the effect of the following small changes to the variables:

$$\begin{aligned} u_1 &= w_1 + \delta_1 - \tfrac{\delta_1^2(1+\gamma)}{2rw_1}\ ; \\ u_2 &= w_2 - \delta_2 - \tfrac{\delta_2^2(1+\gamma)}{2rw_2}\ ; \\ v_1 &= w_1 - \delta_1 - \tfrac{\delta_1^2(1+\gamma)}{2rw_1}\ ; \\ v_2 &= w_2 + \delta_2 - \tfrac{\delta_2^2(1+\gamma)}{2rw_2}\ . \end{aligned}$$

For $j > 2$, we keep $u_j = v_j = w_j$.

We have that

$$\begin{aligned} H_\lambda(\mathbf{u}) &= H_\lambda(\mathbf{w}) + \tfrac{w_1^{1+1/r}}{\lambda_1^{1/r}} \left(\left(\tfrac{u_1}{w_1}\right)^{1+1/r} - 1 \right) + \tfrac{w_2^{1+1/r}}{\lambda_2^{1/r}} \left(\left(\tfrac{u_2}{w_2}\right)^{1+1/r} - 1 \right) \\ &= H_\lambda(\mathbf{w}) + \tfrac{w_1^{1+1/r}}{\lambda_1^{1/r}} \left(\tfrac{\delta_1}{w_1} - \left(\tfrac{\delta_1}{w_1}\right)^2 \tfrac{\gamma(r+1)}{2r^2} + O\left(\tfrac{\delta_1}{w_1}\right)^3 \right) \\ &\quad - \tfrac{w_2^{1+1/r}}{\lambda_2^{1/r}} \left(\tfrac{\delta_2}{w_2} + \left(\tfrac{\delta_2}{w_2}\right)^2 \tfrac{\gamma(r+1)}{2r^2} + O\left(\tfrac{\delta_2}{w_2}\right)^3 \right). \end{aligned}$$

By choice of δ_1 and δ_2, the first order terms cancel. Thus, for δ sufficiently small, we have $H_\lambda(\mathbf{u}) \leq H_\lambda(\mathbf{w}) = 1$. Similarly $H_\lambda(\mathbf{v}) \leq H_\lambda(\mathbf{w}) = 1$ for small δ, so our new choice is feasible.

Now let us determine the effect on $G(\mathbf{u},\mathbf{v})$. To simplify matters, define $x_j = (w_j/\lambda_j)^{1/r}$ for each j, and $\nu = \mu(1+1/r)$, so that $\sum_{k\neq j} w_k = \nu x_j$ for $j = 1,2$.

Then we have

$$\begin{aligned}
G(\mathbf{u},\mathbf{v}) - G(\mathbf{w},\mathbf{w}) &= \nu\big((u_1-w_1)x_1 + (u_2-w_2)x_2 + (v_1-w_1)x_1 + (v_2-w_2)x_2\big) \\
&\quad +(u_1-w_1)(v_2-w_2) + (u_2-w_2)(v_1-w_1) \\
&= \nu\delta x_1 x_2 \Big[\Big(1 - \tfrac{\delta x_2(1+\gamma)}{2rw_1}\Big) + \Big(-1 - \tfrac{\delta x_1(1+\gamma)}{2rw_2}\Big) + \Big(-1 - \tfrac{\delta x_2(1+\gamma)}{2rw_1}\Big) \\
&\quad + \Big(1 - \tfrac{\delta x_1(1+\gamma)}{2rw_2}\Big)\Big] + 2\delta^2 x_1 x_2 + O(\delta^3) \\
&= \delta^2 x_1 x_2 \Big(2 - \tfrac{\nu(1+\gamma)}{r}\Big(\tfrac{x_2}{w_1} + \tfrac{x_1}{w_2}\Big)\Big) + O(\delta^3) \\
&= \delta^2 x_1 x_2 \Big(2 - \tfrac{1+\gamma}{r}\Big(\tfrac{\sum_{k\neq 2} w_k}{w_1} + \tfrac{\sum_{k\neq 1} w_k}{w_2}\Big)\Big) + O(\delta^3)\ .
\end{aligned}$$

By choice of γ, this is positive for δ sufficiently small.

In conclusion, making the indicated change to $\mathbf{u}$ and $\mathbf{v}$ keeps the solution feasible while increasing the objective function, thus proving that the choice $\mathbf{w}$ is not a maximum. ☐

3.3 Unique Simple Semi-invariant Gibbs Measures

We saw in the previous section that for $q < r+1$, there are multiple simple semi-invariant Gibbs measures for all activities, including uniform. Here we show that in the uniform case, this threshold is exact.

Theorem 3.2. *For $q \geq r+1$, there is only one simple semi-invariant Gibbs measure on* $\mathrm{Hom}(T_r, K_q)$ *with uniform activities.*

Proof. Suppose that there are weights $(u_1, u_2, \ldots, u_q, v_1, v_2, \ldots, v_q)$ satisfying the fundamental equations with $\lambda_1 = \cdots = \lambda_q = 1$, so that

$$u_j = \Big(\sum_{k\neq j} v_k\Big)^r; \quad v_j = \Big(\sum_{k\neq j} u_k\Big)^r \quad (j = 1,\ldots,q)\ .$$

Suppose that $u_1 \geq u_2 \geq \cdots \geq u_q$, which implies that $v_1 \leq v_2 \leq \cdots \leq v_q$. Set

$$\alpha = u_1/u_q; \quad \beta = \frac{u_1 + u_2 + \cdots + u_{q-1}}{(q-1)u_q}\ ,$$

so that $\alpha \geq \beta \geq 1$.

We know that all the weights are equal at the unique invariant solution, so what we are claiming is that there is no solution to the equations above with $\beta > 1$.

We have

$$\alpha^{1/r} = \left(\frac{u_1}{u_q}\right)^{1/r} = 1 + \frac{v_q - v_1}{\sum_{k=1}^{q-1} v_k}\ .$$

Note that

$$\sum_{k=1}^{q-1} v_k = \sum_{k=1}^{q-1} \big(u_q + (q-2)\beta u_q + (\beta u_q - u_k)\big)^r \geq (q-1)\left(u_q + (q-2)\beta u_q\right)^r\ ,$$

since the terms $(\beta u_q - u_k)$ sum to zero. Hence, using also that $v_1 = ((q-1)\beta + 1 - \alpha)^r u_q^r$, $v_q = ((q-1)\beta)^r u_q^r$, we see that

$$\begin{aligned} \alpha^{1/r} &\leq 1 + \frac{\big((q-1)\beta\big)^r - \big((q-1)\beta+1-\alpha\big)^r}{(q-1)\big(1+(q-2)\beta\big)^r} \\ &= 1 + \frac{1-\big(1-\frac{\alpha-1}{(q-1)\beta}\big)^r}{(q-1)\big(1-\frac{\beta-1}{(q-1)\beta}\big)^r} . \end{aligned}$$

The right hand side here is decreasing in q, so for $q \geq r+1$ we have

$$0 \leq F_r(\alpha,\beta) = -\alpha^{1/r} + 1 + \frac{1-\left(1-\frac{\alpha-1}{r\beta}\right)^r}{r\left(1-\frac{\beta-1}{r\beta}\right)^r} .$$

We next claim that $F_r(\alpha,\beta)$ is decreasing in α. To see this, we start by observing that

$$\begin{aligned} \frac{\partial F_r}{\partial \alpha} &= \frac{1}{r}\left(-\alpha^{-(r-1)/r} + \frac{1}{\beta}\frac{(1-(\alpha-1)/r\beta)^{r-1}}{(1-(\beta-1)/r\beta)^r}\right) \\ &= \alpha^{-(r-1)/r}\left(-1 + \frac{G(\alpha)^{(r-1)/r}}{G(\beta)}\right), \end{aligned}$$

where the function $G(x) = x(1-(x-1)/r\beta)^r$ is decreasing for $x \geq \beta$, and satisfies $G(\beta) \geq 1$. Thus

$$G(\beta) \geq G(\beta)^{(r-1)/r} \geq G(\alpha)^{(r-1)/r} ,$$

and so indeed F_r is decreasing in α. Therefore we have

$$0 \leq F_r(\alpha,\beta) \leq F_r(\beta,\beta) = H_r(\beta) = -\beta^{1/r} + 1 + \frac{1}{r}\left(\left(1 - \frac{\beta-1}{r\beta}\right)^{-r} - 1\right).$$

Now we claim that $H_r(\beta)$ is decreasing in β. Indeed, its derivative, multiplied by r, is

$$\frac{-1}{\beta^{(r-1)/r}} + \frac{\beta^{r-1}}{\left(1+\frac{r-1}{r}(\beta-1)\right)^{r+1}} .$$

This is negative since $1 + \frac{r-1}{r}(\beta-1) \geq (1+(\beta-1))^{(r-1)/r} = \beta^{(r-1)/r}$, and also we have strict inequality if $\beta > 1$, so $H_r(\beta) \leq H_r(1) = 0$, with strict inequality, and therefore a contradiction, unless $\beta = 1$.

In conclusion, in the case we are considering, there is no solution to the fundamental equations unless all the weights are equal, which is the required result. □

3.4 Uniqueness of Gibbs Measures

In this subsection we will show that when q is sufficiently large relative to r, there is just one Gibbs measure for $\mathrm{Hom}(T_r, K_q)$ with uniform activities. Salas and Sokal [18], and apparently also Kotecký as cited in [9], were able to show that with any board of maximum degree $r+1$, there is a unique Gibbs measure in the uniform case as long as $q > 2r+2$. We can do somewhat better in our special case.

Our approach is to show that there is, indeed, no long-range order; that is, the restriction of a random coloring φ to far-away sites has diminishing effect on the values of φ on a fixed patch of T_r.

Given r probability vectors $\mathbf{p}_1 = (p_{11}, \dots, p_{1q})$, $\dots$, $\mathbf{p}_r = (p_{r1}, \dots, p_{rq})$ in $[0,1]^q$, each with entries summing to 1, we form their *Potts product* $\mathbf{p}_1 * \cdots * \mathbf{p}_r$ as the probability vector proportional to $(p'_{11}p'_{21} \cdots p'_{r1}, \dots, p'_{1q}p'_{2q} \cdots p'_{rq})$, where $p'_{ij} = 1 - p_{ij}$, for $i = 1, \dots, r$, and $j = 1, \dots, q$.

Given q and r, and a positive integer n, consider a finite portion T_r^n of T_r, defined from T_r by deleting one of the $r+1$ branches from the root, and taking for T_r^n all sites at distance at most n from the root. Now, for every assignment of a probability vector $\mathbf{p}(u)$ to each leaf u of T_r^n, we recursively define a probability vector $\mathbf{p}(x)$ at each internal site x of T_r^n as the Potts product $\mathbf{p}(x_1) * \cdots * \mathbf{p}(x_r)$, where $x_1, \dots, x_r$ are the children of x. In particular, any assignment of a probability vector to each leaf gives rise to a probability vector $\mathbf{p}(w)$ at the root w.

Suppose that each probability vector assigned to a leaf is a 0–1 vector, to be thought of as specifying the color of that leaf, and $\mathbf{p}(w) = (p_1, \dots, p_q)$. We claim then that the number of proper q-colorings of T_r^n, such that the leaf colors are as specified and the root has color j, is proportional to p_j. This can be seen easily by working in from the leaves.

We would like to find conditions (on q and r) guaranteeing that, as $n \to \infty$, the root vector $\mathbf{p}(w)$ tends to the uniform vector $(1/q, \dots, 1/q)$, regardless of the leaf vectors. As we shall see, this implies that there is a unique Gibbs measure in this case.

To get some feel for this, consider the case $q = r+1$, take a frozen q-coloring φ of the tree T_r, and delete a branch from the root corresponding to a repeated color among the children of the root. Now the colors on sites u at distance n from the root produce 0–1 probability vectors $\mathbf{p}(u)$. For any internal vertex x, we claim that the derived probability vector $\mathbf{p}(x)$ is also the 0–1 vector corresponding to $\varphi(x)$, which we may take to be the last color q. Indeed, note that the r children of x are each given one of the colors 1, 2, $\dots$, $q-1$, so we may assume for an induction that their assigned vectors are $(1, 0, \dots, 0)$, $(0, 1, 0, \dots, 0)$, $\dots$, $(0, \dots, 0, 1, 0)$. The Potts product of these vectors is $(0, \dots, 0, 1)$, which indeed corresponds to the assignment of color q to x. Hence the root vector $\mathbf{p}(w)$ does not necessarily converge to uniform as $n \to \infty$ in the case $q = r+1$.

Define the *eccentricity* $E(\mathbf{x})$ of a non-negative vector $\mathbf{x} = (x_1, \dots, x_q)$ to be

$$\sum_{j \neq k} (x_j - x_k)^2 \Big/ \left(\sum_{j=1}^{q} x_j \right)^2 .$$

Note that the denominator is 1 for a probability vector, while the eccentricity is invariant under scaling: $E(t\mathbf{x}) = E(\mathbf{x})$ for all scalars t and vectors $\mathbf{x}$.

Let $M(q, r; n)$ be the maximum eccentricity of $\mathbf{p}(w)$, over all assignments of probability vectors $\mathbf{p}(u)$ to the leaves u of T_r^n.

Theorem 3.3. *Suppose that q and r are such that $M(q, r; n) \to 0$ as $n \to \infty$. Then there is a unique Gibbs measure (namely, the simple invariant Gibbs measure generated by the node-weighted branching random walk) for* $\mathrm{Hom}(T_r, K_q)$ *with uniform activities.*

Proof. Let μ be any Gibbs measure; we aim to show that μ is equal to the measure ν defined from the r-branching random walk on K_q, with all nodes weighted equally.

To show that the two measures are equal, take any patch U, any proper coloring φ of U, and any $\varepsilon > 0$. We need to show that $|\Pr_\mu(\psi = \varphi) - \Pr_\nu(\psi = \varphi)| < \varepsilon$. We may assume (by taking a larger patch if necessary) that U is connected (so it is a subtree of T_r), and that all its elements have degree 1 or $r+1$. Note that $\Pr_\nu(\psi = \varphi) = 1/q(q-1)^{|U|-1}$, since all of the $q(q-1)^{|U|-1}$ proper q-colorings of U are equally likely under ν.

Suppose that U has ℓ leaves. Now choose n sufficiently large (to be specified later), and let V denote the set of all sites of T_r at distance at most n from our patch U. Note that $(V \setminus U) \cup \partial U$ consists of ℓ disjoint copies of T_r^k, one rooted at each leaf of U.

We generate a random sample from the Gibbs measure μ according to the following procedure: first take a random sample ψ from μ, then uncolor all sites in $V \setminus \partial V$. Now choose, uniformly at random, a proper coloring of $V \setminus \partial V$ consistent with the coloring of ∂V. By the Gibbs property, this is a valid way of constructing a sample coloring ψ'.

Now take any coloring φ of U, and observe that the probability that $\psi' \restriction U = \varphi$ is equal to the sum, over all colorings ρ of ∂V, of $\Pr_\mu(\psi \restriction \partial V = \rho) \Pr(\psi' \restriction U = \varphi \mid \psi \restriction \partial V = \rho)$. The second probability here refers only to the recoloring phase, and so is proportional to the number of proper colorings of $V \setminus (U \cup \partial V)$ consistent with both the coloring ρ of ∂V and the coloring $\varphi \restriction \partial U$ of the leaves of U. This is, by the earlier discussion, proportional to the product, over all the leaves u of U, of the $\varphi(u)$-entry of the vector $\mathbf{p}(u)$ generated from the 0–1 probability vectors set (by ρ) at the leaves of the copy of T_r^n rooted at u. This vector $\mathbf{p}(u)$ has eccentricity at most $M = M(q, r; n)$, by definition, and so each entry is within $\sqrt{M}$ of $1/q$. Hence the probability $\Pr(\psi' \restriction U = \varphi \mid \psi \restriction \partial V = \rho)$ is proportional to a number lying between $(1/q - \sqrt{M})^\ell$ and $(1/q + \sqrt{M})^\ell$, for every φ and ρ. Therefore

$$\left(\frac{1/q - \sqrt{M}}{1/q + \sqrt{M}}\right)^\ell \frac{1}{q(q-1)^{|U|-1}} \leq \Pr(\psi' \restriction U = \varphi) \leq \left(\frac{1/q + \sqrt{M}}{1/q - \sqrt{M}}\right)^\ell \frac{1}{q(q-1)^{|U|-1}} ,$$

so we can ensure that $\Pr(\psi' \restriction U = \varphi)$ is within ε of $1/q(q-1)^{|U|-1}$ by choosing n to make M small enough. ☐

To prove that $M(q, r; n) \to 0$, the obvious approach is to show that, if $\mathbf{p}_1, \ldots, \mathbf{p}_r$ are any probability vectors with eccentricities at most m, then the eccentricity of $\mathbf{p}_1 * \cdots * \mathbf{p}_r$ is at most tm for some $t < 1$. It turns out to be worthwhile to refine this approach somewhat.

We start with a lemma dealing with the normalization involved in the definition of the Potts product.

Lemma 3.4. *Suppose* $\mathbf{p}_1 = (p_{11}, \ldots, p_{1q})$, $\ldots$, $\mathbf{p}_r = (p_{r1}, \ldots, p_{rq})$ *are probability vectors, all of whose entries are at most* U. *Then*

$$\sum_{j=1}^{q} p'_{1j} \cdots p'_{rj} \geq \max\left(q - r, q(1-U)^{r/qU}\right).$$

and so in particular every entry of $\mathbf{p}_1 * \cdots * \mathbf{p}_r$ *is at most* $1/(q-r)$.

Proof. Consider the problem of minimizing $\sum_{j=1}^q p'_{1j} \cdots p'_{rj}$, subject to the conditions that all the p'_{ij} are (to start with) between 0 and 1, and $\sum_{j=1}^q p'_{ij} = q - 1$ for all $i = 1, \ldots, r$. For any fixed assignment of values to all the variables except $p'_{i1}, \ldots, p'_{iq}$ (for some fixed i), it is clear that the minimum is obtained by setting $q - 1$ of these variables to 1, and the other (that value p'_{ij} whose multiplier $\prod_{h \neq i} p'_{hj}$ is maximum) to 0. Hence, at the optimum, exactly r of the qr variables p'_{ij} are set to 0, making r of the terms $p'_{1j} \cdots p'_{rj}$ equal to 0, and the other $q - r$ terms equal to 1. Therefore

$$\sum_{j=1}^{q} p'_{1j} \cdots p'_{rj} \geq q - r .$$

To see the other bound, we start by using the arithmetic-geometric mean inequality to find that

$$\sum_{j=1}^{q} \prod_{i=1}^{r} p'_{ij} \geq q \left(\prod_{j=1}^{q} \prod_{i=1}^{r} p'_{ij}\right)^{1/q} .$$

Now each of the qr terms p'_{ij} can be written as $1-\lambda_{ij}U$, where $0 \leq \lambda_{ij} \leq 1$. Since the sum of all the qr terms p'_{ij} is $r(q-1)$, the sum of all the λ_{ij} is equal to r/U. Now since

$$p'_{ij} = 1 - \lambda_{ij}U \geq (1-U)^{\lambda_{ij}} \ ,$$

we have

$$\sum_{j=1}^{q} \prod_{i=1}^{r} p'_{ij} \geq q \left((1-U)^{r/U} \right)^{1/q} ,$$

as desired.

The final statement follows because an entry of $\mathbf{p}_1 * \cdots * \mathbf{p}_r$ is obtained by dividing a number at most 1 by a normalizing factor of at least $q-r$. $\square$

Note: We can also see informally that the entries of $\mathbf{p}_1 * \cdots * \mathbf{p}_r$ are at least $1/(q-r)$, by interpreting p_{ij} as the probability that the ith child of a site u has color j; since any possible coloring of the children permits at least $q-r$ colors at u, the probability of any one color at u is at most $1/(q-r)$.

Thus, if we take any probability vectors at the leaves of our T_r^n, the entries of their Potts products, one layer in, are all at most $U = 1/(q-r)$. This in turn implies a lower bound on the entries two layers in, namely $L = (1-1/(q-r))^r/(q-1)$, since $q-1$ is an easy upper bound on the normalization constant. Better bounds could be found by iterating this procedure.

To prove that $M(q,r;n) \to 0$ as $n \to \infty$, it thus suffices to show that, if $\mathbf{p}_1 = (p_{11}, \cdots, p_{1q})$, $\ldots$, $\mathbf{p}_r = (p_{r1}, \ldots, p_{rq})$ are probability vectors with eccentricity at most m and entries between L and U, then $E(\mathbf{p}_1 * \cdots * \mathbf{p}_r) \leq tm$ for some $t < 1$.

Accordingly, take such vectors $\mathbf{p}_1, \ldots, \mathbf{p}_r$, and consider

$$E(\mathbf{p}_1 * \cdots * \mathbf{p}_r) = \frac{\sum_{k \neq j} \left(p'_{1j} \cdots p'_{rj} - p'_{1k} \ldots p'_{rk} \right)^2}{\left(\sum_{j=1}^{q} p'_{1j} \cdots p'_{rj} \right)^2} .$$

We need an upper bound on the numerator of this expression, which we can then combine with the lower bound on the denominator given by Lemma 3.4.

By telescoping, we see that

$$p'_{1j} \cdots p'_{rj} - p'_{1k} \cdots p'_{rk} = \sum_{i=1}^{r} (p'_{ij} - p'_{ik}) \prod_{h<i} p'_{hj} \prod_{h>i} p'_{hk} .$$

Since all the p'_{ij} are at most $1-L$, while $p'_{ij} - p'_{ik} = p_{ik} - p_{ij}$, we find that

$$\left| p'_{1j} \cdots p'_{rj} - p'_{1k} \cdots p'_{rk} \right| \leq (1-L)^{r-1} \sum_{i=1}^{r} |p_{ik} - p_{ij}| \ .$$

Thus we have

$$\begin{aligned} \sum_{j \neq k} \left(p'_{1j} \cdots p'_{rj} - p'_{1k} \cdots p'_{rk} \right)^2 &\leq (1-L)^{2(r-1)} \sum_{j \neq k} \left(\sum_{i=1}^{r} |p_{ik} - p_{ij}| \right)^2 \\ &\leq (1-L)^{2(r-1)} r \sum_{i=1}^{r} \sum_{j \neq k} (p_{ik} - p_{ij})^2 \\ &\leq r^2 (1-L)^{2(r-1)} m \ . \end{aligned}$$

Combining this bound with that from Lemma 3.4, we see that

$$E(\mathbf{p}_1 * \cdots * \mathbf{p}_r) \leq \min \left\{ \left(\frac{r(1-L)^{r-1}}{q-r} \right)^2 , \left(\frac{r(1-L)^{r-1}}{q(1-U)^{r/qU}} \right)^2 \right\} \cdot m \ .$$

Hence if q and r are such that either of the two terms is less than 1, for the values of L and U given earlier, then there is a unique Gibbs measure for $\mathrm{Hom}(T_r, K_q)$, with uniform activities.

In particular, if $q \geq 2r$, then $r(1-L)^{r-1}/(q-r) < 1$ and we have uniqueness. Also, if r and q are both large, with $q = (d+1)r$, then $U = 1/dr$, $L \approx e^{-1/d}/(d+1)r$, and

$$\frac{r(1-L)^{r-1}}{q(1-U)^{r/qU}} \approx \frac{1}{d+1} \exp\left(\frac{1-e^{-1/d}}{d+1}\right).$$

This quantity is less than 1 for $d > 0.6296$. Therefore we have the following result.

Theorem 3.5. *Suppose that $q \geq 2r$, or that $q \geq 1.6296\, r$ and r is sufficiently large. Then there is a unique Gibbs measure for $\mathrm{Hom}(T_r, K_q)$, with uniform activities.*

The reader will have no difficulty in improving the constant 1.6296 in the above theorem. Our own calculations give also a more specific bound, namely that we have a unique uniform Gibbs measure when $r \geq 5$ and $q > (\frac{5}{3})r$. However, what we believe is that there is a unique Gibbs measure even for $q = r+2$, and this method does not seem able to give any result of this type. Even the case $q = 5$, $r = 3$ remains a challenge, although it is not impossible that the method used here can be improved to cover this case.

4 Remarks and Conjectures

We have pretty much settled the question of when there is more than one simple semi-invariant Gibbs measure, but have left some important open questions concerning general Gibbs measures.

We have shown that, when $q \geq r+1$, there is only one simple semi-invariant Gibbs measure in the uniform case, but multiple such measures for some non-uniform activity sets. We suspect that when $q > r+1$, there is only one Gibbs measure of any kind for the uniform set of activities. We have been able to prove this for $q = 4$ and $r = 2$, and generally when $q \geq 2r$ and, for large r, when $q \geq 1.6296\, r$.

For the critical case $q = r+1$, all non-uniform sets of activities produce multiple simple semi-invariant Gibbs measures. There are other Gibbs measures, however, regardless of the activity set, and it appears to be a daunting challenge to classify them—even in the ground case $q = 3$, $r = 2$.

When $q < r+1$, we know that there are multiple simple semi-invariant Gibbs measures for every set of activities (plus frozen Gibbs measures and more).

The state of affairs is roughly illustrated in Figure 3.

There are many related issues which we have not considered in this work. For example, the question of whether our simple semi-invariant Gibbs measures are *extremal* (that is, not mixes of other Gibbs measures) is complex, and linked to the issue of whether information about the starting state of a branching random walk is available from states reached at a much later time (see e.g. [5, 8, 11, 12]). (We can show that for $q \geq r+1$, the simple invariant Gibbs measure for the uniform set of activities is indeed extremal.)

The DLR problem (uniqueness of Gibbs measures) is itself closely allied with the problem of whether "heat bath" dynamics, which make local changes to colorings of a large finite board, constitute rapidly mixing Markov chains. For example, Vigoda [19] has shown that when the number of colors is at least 11/6 times the maximum degree of the board, rapid mixing ensues; it is thus arguably noteworthy that our constant 1.6296 undercuts 11/6.

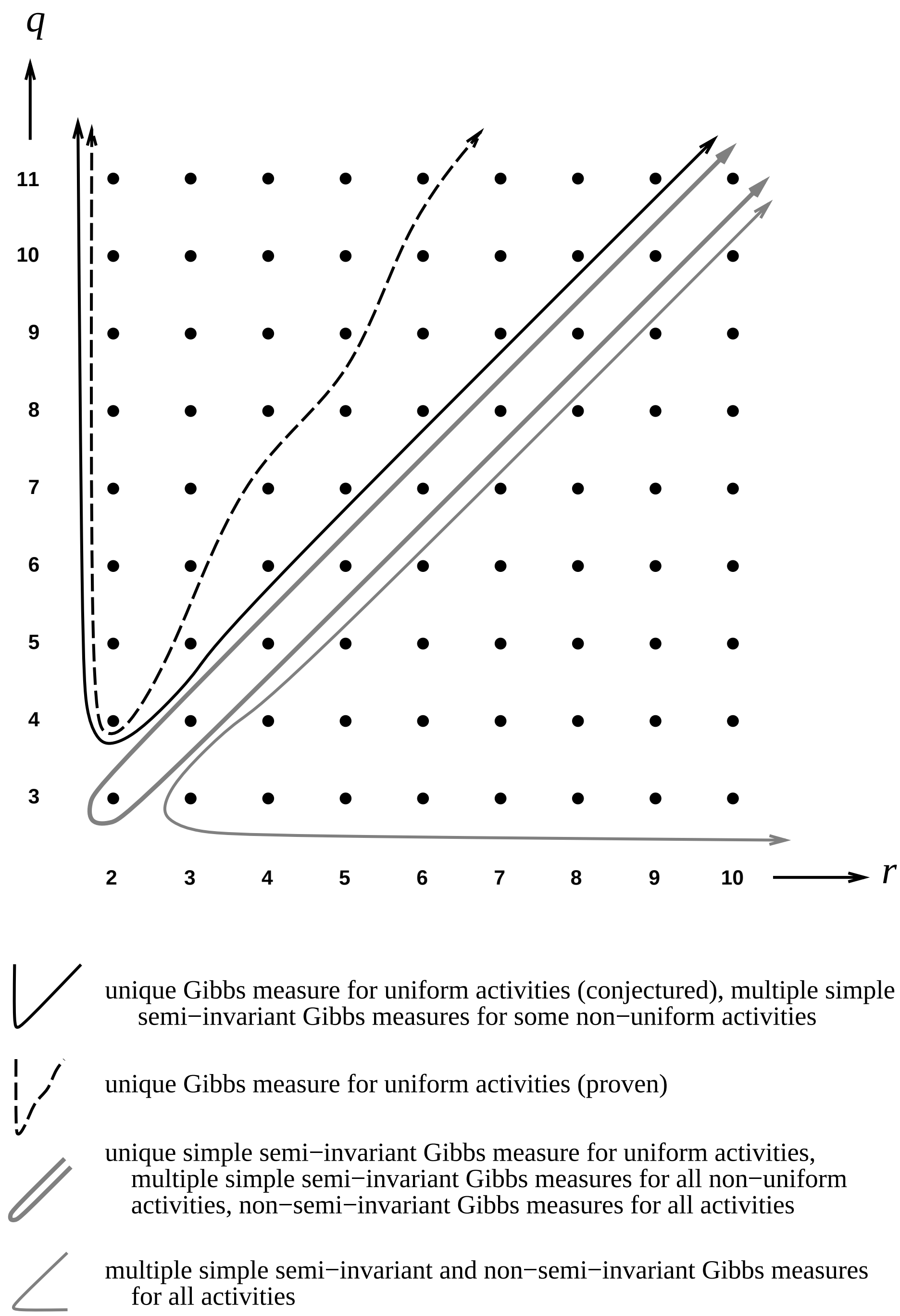

Figure 3: Simple semi-invariant, and general, Gibbs measures for q-coloring the r-branching Cayley tree

References

[1] R.J. Baxter, *Exactly Solved Models in Statistical Mechanics*, Academic Press, London (1982).

[2] J. van den Berg and J.E. Steif, Percolation and the hard-core lattice gas model, *Stochastic Proc. and their Appls.* **49** (1994), 179–197.

[3] G.R. Brightwell, O. Häggström and P. Winkler, Nonmonotonic behavior in hard-core and Widom–Rowlinson models, *J. Stat. Physics* **94** (1999), 415–435.

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

[5] G.R. Brightwell and P. Winkler, Gibbs measures and dismantlable graphs, *J. Comb. Theory (Series B)*, to appear (2000).

[6] R.M. Burton, C.-E. Pfister and J.E. Steif, The variational principle for Gibbs states fails on trees, *Markov Process. Related Fields* **1** #3 (1995), 387–406.

[7] R.L. Dobrushin, The description of a random field by means of conditional probabilities and conditions of its regularity, *Theory Probab. Appl.* **13** #2 (1968), 197–224.

[8] W. Evans, C. Kenyon, Y. Peres and L.J. Schulman, Broadcasting on trees and the Ising model, *Ann. Appl. Probab.*, to appear (2000).

[9] H.-O. Georgii, *Gibbs Measures and Phase Transitions*, de Gruyter, Berlin (1988).

[10] F.P. Kelly, Stochastic models of computer communication systems, *J. R. Statist. Soc. B* **47** #3 (1985), 379–395.

[11] E. Mossel, Recursive reconstruction on periodic trees, *Random Structures & Algorithms* **13** #1 (1998), 81–97.

[12] E. Mossel, Reconstruction on trees: Beating the second eigenvalue, preprint (1999).

[13] F. Peruggi, Probability measures and Hamiltonian models on Bethe lattices. I. Properties and construction of MRT probability measures, J. Math. Phys. **25** (1984), 3303–3315.

[14] F. Peruggi, F. di Liberto and G. Monroy, The Potts model on Bethe lattices: I. General results, *J. Phys. A: Math. Gen.* **16** (1983), 811–828.

[15] F. Peruggi, F. di Liberto and G. Monroy, Critical behaviour in three-state Potts antiferromagnets on a Bethe lattice, *Physica* **131A** (1985), 300–310.

[16] F. Peruggi, F. di Liberto and G. Monroy, Phase diagrams of the q-state Potts model on Bethe lattices, *Physica* **141A** (1987), 151–186.

[17] B. Prum and J.C. Fort, *Stochastic Processes on a Lattice and Gibbs Measures*, Kluwer, Dordrecht (1991).

[18] J. Salas and A.D. Sokal, Absence of phase transition for antiferromagnetic Potts models via the Dobrushin uniqueness theorem, *J. Stat. Phys.* **86** (1997), 551—579.

[19] E. Vigoda, Improved bounds for sampling colorings, *Proc. 40th Symp. on Foundations of Comp. Sci.*, I.E.E.E. Computer Society, Los Alamitos CA (1999), 51–59.

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