A note on disagreement percolation

Olle Häggström*

February 2, 2001

Abstract

We construct a coupling of two distinct Gibbs measures for Markov random fields with the same specifications, such that the existence of an infinite path of disagreements between the two configurations has probability 0. This shows that the independence assumption in the disagreement percolation method for proving Gibbsian uniqueness, cannot be dropped without being replaced by other conditions. A similar counterexample is given for couplings of Markov chains.

1 Introduction

There are many interesting and useful connections between percolation theory on one hand, and Gibbs systems and Markov random fields on the other. One example is the random-cluster model, which puts bond percolation and Potts models in a common parameterized family, and gives percolation-theoretic proofs of phase transition in the latter [8, 1, 10]; another is the equivalence between spin percolation and Gibbs state multiplicity in Ising and Potts models on the square lattice [7, 6]. We refer to [9] for a general introduction to such connections.

A particularly striking and intuitive example is van den Berg's [2] so-called disagreement percolation technique for proving uniqueness of Gibbs measures in Markov random fields. The main ingredient of this technique is the following result from [2] (for precise definitions of all concepts involved, see Section 2).

Theorem 1.1 *Let $G = (V, E)$ be an infinite locally finite graph, and let S be a finite set. Let μ_1 and μ_2 be two Gibbs measures for the same specification of a Markov random field on G taking values in S. Pick two configurations $X_1, X_2 \in S^V$ independently with respective distributions μ_1 and μ_2. If*

$$\mathbf{P}(G \textit{ contains an infinite path of disagreements between } X_1 \textit{ and } X_2) = 0\,, \qquad (1)$$

then $\mu_1 = \mu_2$.

(By an infinite path, we mean an infinite *self-avoiding* path in G.) For applications of this result, see e.g. [2, 5, 3]. The intuition behind Theorem 1.1 is the following. If the event in (1) has probability 0, then every finite subset of V is, with probability 1, surrounded by some finite random set W on which X_1 and X_2 agree (take the same values). In general, there are (for a given pair of configuration (X_1, X_2)) many choices of such a surrounding set, but if W is chosen with sufficient care, then the conditional

*Research supported by the Swedish Natural Science Research Council

distributions of X_1 and X_2 "inside" W, given their (common) values on W, are the same. The distributions of X_1 and X_2 on any finite subset of V are therefore the same, and it follows that $\mu_1 = \mu_2$.

An extension of the above result to certain dependent couplings of μ_1 and μ_2 is given by van den Berg and Maes [4] (a coupling of μ_1 and μ_2 is a joint construction on the same probability space of two random objects with respective distributions μ_1 and μ_2). The above intuition is so strong that one might be tempted to believe that the conclusion $\mu_1 = \mu_2$ might follow from (1), for *any* coupling of μ_1 and μ_2. The following result (which is the main one of this note) says that this is *not* the case.

Theorem 1.2 *There exist a countably infinite, locally finite, graph $G = (V, E)$, a finite set S, two Gibbs measures μ_1 and μ_2 on S^V for the same Markov random field specification, and a pair of random configurations $X_1, X_2 \in S^V$ with respective distributions μ_1 and μ_2, such that* (1) *holds, while on the other hand $\mu_1 \neq \mu_2$.*

The existence of such a coupling was conjectured by Steif [15] in the early days of disagreement percolation (perhaps 1992).

In Section 3, we will (after some preliminaries in Section 2) examine a specific example which establishes Theorem 1.2. The example is an Ising model on a graph which is obtained by local modifications of the square lattice $\mathbf{Z}^2$. Finally, in Section 4, we will give a similar, and very simple, counterexample in the more classical theory of couplings of Markov chains [13, 16]. Section 4 can be read independently of Sections 2 and 3, and some readers will perhaps benefit from reading Section 4 first, before going into (the somewhat more technical) Sections 2 and 3.

2 Preliminaries

In this section, we recall some well-known definitions and results, needed to prove our main result in Section 3. Some general references for the results quoted here, are [10] and [9].

2.1 (Quasi-)transitivity and invariant couplings

A **graph automorphism** of a graph $G = (V, E)$ is a bijective mapping $\gamma : V \to V$ which preserves adjacency, i.e. for any $x, y \in V$ we have that $\gamma(x)$ and $\gamma(y)$ share an edge in E if and only if x and y do. The graph G is said to be **transitive** if for any $x, y \in V$, there exists a graph automorphism of G mapping x on y. The graph is said to be **quasi-transitive** if V can be partitioned into finitely many sets $\{V_1, \ldots, V_k\}$ such that for each $i \in \{1, \ldots, k\}$ and every $x, y \in V_i$, there is a graph automorphism mapping x on y. Intuitively, quasi-transitivity means that the graph has only finitely many "types" of vertices.

Let S be a finite set. An S^V-valued random element X is said to be **invariant** if for any graph automorphism γ, any positive integer n, any $x_1, \ldots, x_n \in V$, and any $s_1, \ldots, s_n \in S$, we have

$$\mathbf{P}(X(x_1) = s_1, \ldots, X(x_n) = s_n) = \mathbf{P}(X(\gamma(x_1)) = s_1, \ldots, X(\gamma(x_n)) = s_n) \,.$$

Similarly, a coupling of two invariant S^V-valued random elements X and Y is said to be invariant if for all γ, n, $x_1, \ldots, x_n$ as above, and all $s_1, \ldots, s_{2n} \in S$, we have

$$\mathbf{P}(X(x_1) = s_1, \ldots, X(x_n) = s_n, Y(x_1) = s_{n+1}, \ldots, Y(x_n) = s_{2n})$$

$$= \quad \mathbf{P}(X(\gamma(x_1)) = s_1, \ldots, X(\gamma(x_n)) = s_n, Y(\gamma(x_1)) = s_{n+1}, \ldots, Y(\gamma(x_n)) = s_{2n}) .$$

Intuitively, this means for (quasi-)transitive graphs that the joint behavior of X and Y is the same "everywhere in the graph".

2.2 Stochastic domination

Now suppose that S is a subset of $\mathbf{R}$. It is then natural to introduce the following partial order $\preceq$ on S^V: for $\xi, \eta \in S^V$, we have $\xi \preceq \eta$ if and only if $\xi(x) \leq \eta(x)$ for all $x \in V$.

A function $f : S^V \to \mathbf{R}$ is said to be **increasing** if $f(\xi) \leq f(\eta)$ whenever $\xi \preceq \eta$. For two S^V-valued random objects X and Y with respective distributions μ and ν, we say that X is **stochastically dominated** by Y, denoted $X \preceq_d Y$, if for all bounded increasing functions $f : S^V \to \mathbf{R}$ we have

$$\int_{S^V} f(\xi)\, d\mu(\xi) \leq \int_{S^V} f(\xi)\, d\nu(\xi) .$$

By Strassen's theorem (see, e.g., [13]), we have that $X \preceq_d Y$ if and only if there exists a coupling of X and Y such that

$$\mathbf{P}(X \preceq Y) = 1 .$$

Such a coupling of X and Y is said to be **monotone**.

2.3 Markov random fields

We take $x \sim y$ to mean that the two vertices x and y share an edge in G. For any subset W of V, we define the **boundary** ∂W of W as

$$\partial W = \{x \in V \setminus W : \exists y \in W \text{ such that } x \sim y\} .$$

Also, for $X \in S^V$ and $W \subset V$, we write $X(W)$ as shorthand notation for $\{X(x)\}_{x \in W}$.

A probability measure μ on S^V, and the corresponding S^V-valued random element X, are said to be a **Markov random field** if μ admits conditional probabilities such that for all finite vertex sets $W \subset V$, all $\xi \in S^W$, and all $\eta \in S^{V \setminus W}$, we have

$$\mu(X(W) = \xi \,|\, X(V \setminus W) = \eta) = \mu(X(W) = \xi \,|\, X(\partial W) = \eta(\partial W)) .$$

In this case, we also call μ a **Gibbs measure** for the Markov random field.

Two probability measures μ and ν on S^V, and the corresponding S^V-valued random elements X and Y, are said to be **Markov random fields with the same specifications** if, for all W, ξ and η as above, we have

$$\mu(X(W) = \xi \,|\, X(\partial W) = \eta(\partial W)) = \nu(Y(W) = \xi \,|\, Y(\partial W) = \eta(\partial W)) .$$

2.4 The Ising model

Perhaps the most famous example of a Markov random field is the Ising model. For $\beta \geq 0$, a probability measure μ on $\{-1, 1\}^V$ is said to be a **Gibbs measure for the**

Ising model on G at inverse temperature β, if μ is a Markov random field such that for all W, ξ and η as above, we have

$$\begin{aligned} &\mu(X(W) = \xi \,|\, X(\partial W) = \eta(\partial W)) \\ &= \frac{1}{Z} \exp\left(-2\beta \left(\sum_{\substack{x,y \in W \\ x \sim y}} I_{\{\xi(x) \neq \xi(y)\}} + \sum_{\substack{x \in W,\, y \in \partial W \\ x \sim y}} I_{\{\xi(x) \neq \eta(y)\}} \right) \right) \end{aligned} \qquad (2)$$

where Z is a normalizing constant which is allowed to depend on W and η, but not on ξ. For any G and β, one can construct a particular Gibbs measure μ^+ as follows. First let $W_1 \subset W_2 \subset \cdots$ be an increasing sequence of finite subsets of V, converging to V in the sense that each $x \in V$ is in all but at most finitely many of the W_i's. Then, for each n, let $\mu_{n,+}$ be the distribution of the random element $X \in \{-1,1\}^V$ obtained by letting $X(V \setminus W_n) \equiv 1$, and picking $X(W_n)$ according to (2) with $\eta \equiv 1$. The measures $\mu_{n,+}$ are known to converge, as $n \to \infty$, to a limiting probability measure μ_+, called the **plus measure** for the Ising model on G at inverse temperature β. The **minus measure** μ_- is defined analogously. Both μ_+ and μ_- turn out to be Gibbs measures (for the Ising model with the given β).

It is well-known that any Gibbs measure μ for the Ising model on G at inverse temperature β is sandwiched between μ_+ and μ_- in the sense that

$$\mu_- \preceq_d \mu \preceq_d \mu_+ \,.$$

Hence, uniqueness of Gibbs measures for the given G and β is equivalent to having $\mu_- = \mu_+$. Another standard result is the following, where, for any finite $W \subset V$ and any boundary condition $\xi \in \{-1,1\}^{\partial W}$, we let $\mu_{W,\xi}$ denote the probability distribution on $\{-1,1\}^W$ given in (2).

Lemma 2.1 *For any W as above, and any $\xi, \xi' \in \{-1,1\}^{\partial W}$ such that $\xi \preceq \xi'$, we have*

$$\mu_{W,\xi} \preceq_d \mu_{W,\xi'} \,.$$

Next, for the case where G is the square lattice (i.e., G is the graph with vertex set $\mathbf{Z}^2$, and edge set consisting of all pairs of vertices at Euclidean distance 1 from each other), it is known that $\mu_+ = \mu_-$ if and only if $\beta \leq \beta_c$, where $\beta_c = \frac{1}{2}\log(1+\sqrt{2})$ is the reciprocal of the so-called Onsager critical temperature.

Note also that for the case where G is finite, the above definition automatically gives a unique Gibbs measure for the Ising model at a given β.

3 Proof of the main result

The following result clearly implies Theorem 1.2.

Theorem 3.1 *There exists a countably infinite, locally finite, quasi-transitive graph $G = (V, E)$ and a $\beta > 0$ such that the Ising model on G at inverse temperature β has the following properties:*

(i) $\mu_- \neq \mu_+$

(ii) *There exists a coupling of μ_- and μ_+ such that with probability* 1, *G contains no infinite path of disagreements between the two configurations.*

Moreover, the coupling in (ii) *can be taken to be invariant and monotone.*

It is natural to ask whether the corresponding statement is true if we restrict to transitive graphs. I expect this to be the case (see Jonasson [12] for a very recent result in this direction), and would even go as far as saying that an example ought to exist even for $G = \mathbf{Z}^2$. Finding such an example would probably require techniques other than those used in the present paper.

The rest of this section is devoted to the construction of an example which proves Theorem 3.1.

We first describe the graph on which the Ising model will live. For positive integers n and k, and any graph H, we define **the (n, k)-decoration of H** to be the graph obtained by replacing each edge of H by n parallel paths, each of length k. Our choice of graph will be the $(23, 3)$-decoration of the square lattice, and we write $\mathbf{Z}^2_{(23,3)}$ for this graph. See Figure 1.

In the graph $\mathbf{Z}^2_{(23,3)}$, we will use the term **bridge** for a pair of vertices not in $\mathbf{Z}^2$ sitting on one of the paths of length 3 linking two vertices in $\mathbf{Z}^2$. Note that each vertex in $\mathbf{Z}^2_{(23,3)}$ is either in $\mathbf{Z}^2$, or in a (uniquely specified) bridge.

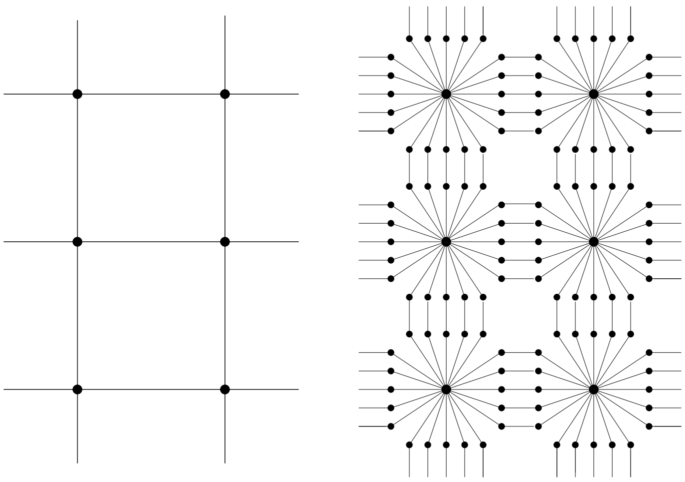

Figure 1: A part of the square lattice, and the corresponding part of its $(n, 3)$-decoration with $n = 5$. Each edge of the square lattice is replaced by n parallel paths of length 3. In our main example, we will take $n = 23$ rather than $n = 5$.

Next, we fix the inverse temperature at $\beta = \frac{1}{4}\log(3)$ (this is precisely the value we need to make the calculations in equations (3), (4) and (5) work out neatly). As a preparation for our analysis of the Ising model on $\mathbf{Z}^2_{(23,3)}$, we first study it on $(n,3)$-decorations of some other graphs. What now follows is just an exercise in well-known series-parallel laws for the Ising model. The most basic example is to take H to be the graph on two vertices x and y, and a single edge connecting x to y. The $(1,3)$-decoration $H_{(1,3)}$ of H, is simply a path of length 3. A direct calculation (summing over all possible configurations of -1's and $+1$'s on $H_{(1,3)}$) shows that if we take the Ising model on $H_{(1,3)}$ at $\beta = \frac{1}{4}\log(3)$, then the probability that the endpoints x and y of $H_{(1,3)}$ take the same spin value, satisfies

$$\frac{\mathbf{P}(x \text{ and } y \text{ get the same spin})}{\mathbf{P}(x \text{ and } y \text{ get different spins})} = \frac{1+\frac{1}{3}+\frac{1}{3}+\frac{1}{3}}{\frac{1}{\sqrt{3}}+\frac{1}{\sqrt{3}}+\frac{1}{\sqrt{3}}+\frac{1}{3\sqrt{3}}} = \frac{3\sqrt{3}}{5}.$$

A similar summation for the Ising model on $H_{(n,3)}$ at the same inverse temperature, gives

$$\frac{\mathbf{P}(x \text{ and } y \text{ get the same spin})}{\mathbf{P}(x \text{ and } y \text{ get different spins})} = \frac{(1+\frac{1}{3}+\frac{1}{3}+\frac{1}{3})^n}{(\frac{1}{\sqrt{3}}+\frac{1}{\sqrt{3}}+\frac{1}{\sqrt{3}}+\frac{1}{3\sqrt{3}})^n} = \left(\frac{3\sqrt{3}}{5}\right)^n.$$

But this is exactly the same ratio as we would get if we took the Ising model on H at inverse temperature $\frac{n}{2}\log(\frac{3\sqrt{3}}{5})$. This, together with the ± 1 symmetry of the Ising model, means that if we take the Ising model on $H_{(n,3)}$ at $\beta = \frac{1}{4}\log(3)$, and then just look at the spins at x and y, what we see is the Ising model on H at inverse temperature $\frac{n}{2}\log(\frac{3\sqrt{3}}{5})$. The following more general statement is an easy consequence.

Lemma 3.2 *Let $H = (V_H, E_H)$ be any finite graph, and let $H_{(n,3)} = (V_{H_{(n,3)}}, E_{H_{(n,3)}})$ be its $(n,3)$-decoration. If we now pick $X \in \{-1,1\}^{V_{H_{(n,3)}}}$ according to the Ising model on $H_{(n,3)}$ with $\beta = \frac{1}{4}\log(3)$, then $X(V_H)$ is distributed according to the Ising model on H with $\beta = \frac{n}{2}\log(\frac{3\sqrt{3}}{5})$.*

A standard limiting argument then implies the following, where we write $V_{\mathbf{Z}^2_{(23,3)}}$ for the vertex set of the graph $\mathbf{Z}^2_{(23,3)}$.

Lemma 3.3 *Pick the $V_{\mathbf{Z}^2_{(23,3)}}$-valued random element X (resp. Y) according to the minus (resp. plus) measure for the Ising model on $\mathbf{Z}^2_{(23,3)}$ at inverse temperature $\beta = \frac{1}{4}\log(3)$. Then $X(\mathbf{Z}^2)$ (resp. $Y(\mathbf{Z}^2)$) is distributed according to the minus (resp. plus) measure for the Ising model on the square lattice at inverse temperature $\beta^* = \frac{23}{2}\log(\frac{3\sqrt{3}}{5})$.*

The alert reader has probably been wondering: why 23? The answer is that with this choice, we get

$$\beta^* = \frac{23}{2}\log\left(\frac{3\sqrt{3}}{5}\right) = 0.1921\ldots > 0.1913\ldots = \frac{1}{2}\log(1+\sqrt{2})$$

which is the Onsager critical value. Lemma 3.3 therefore implies that for the Ising model on $\mathbf{Z}^2_{(23,3)}$ at $\beta = \frac{1}{4}\log(3)$, we have $\mu_+ \neq \mu_-$.

(It is worth stressing at this point that our method does not really rely on knowing the Onsager critical value. It is enough to know the much more elementary fact that the critical value for the Ising model on $\mathbf{Z}^2$ is finite, although we then have to replace

$n = 23$ by some unspecified "n large enough". See also the remark at the end of this section.)

In order to prove Theorem 3.1, it only remains for us to find a monotone and invariant coupling of X and Y (with distributions as in Lemma 3.3), such that with probability 1 there is no infinite path of vertices in $\mathbf{Z}^2_{(23,3)}$ on which X and Y take opposite values.

Consider first one of the spin configurations, say X, and note that if we condition on $X(\mathbf{Z}^2)$, then the spin values on different bridges (including those bridges that go between the same pair of vertices) are conditionally independent. This is immediate from the Markov random field property of the Ising model.

Hence, a correct coupling of X and Y can be obtained as follows. First pick $X(\mathbf{Z}^2)$ and $Y(\mathbf{Z}^2)$ according to some coupling with the correct first and second marginals. Then, for each bridge B *independently*, pick $X(B)$ and $Y(B)$ according to some coupling whose first and second marginals are given by (2) with boundary conditions provided by $X(\mathbf{Z}^2)$ and $Y(\mathbf{Z}^2)$.

We go on to make this more specific. It is well-known that for the Ising model on the square lattice, there exists a coupling of the minus and plus measures which is monotone and invariant; for concreteness, we may take the coupling based on Glauber dynamics and coupling-from-the-past provided in [11, Section 5]. So pick $X(\mathbf{Z}^2)$ and $Y(\mathbf{Z}^2)$ according to such a coupling for the Ising model on the square lattice at inverse temperature $\beta^* = \frac{23}{2}\log(\frac{3\sqrt{3}}{5})$. Then, for each bridge $B = \{z_1, z_2\}$ independently, pick $X(B)$ and $Y(B)$ according to the following rules. With v_1 denoting the vertex in $\mathbf{Z}^2$ adjacent to z_1, and v_2 denoting the vertex in $\mathbf{Z}^2$ adjacent to z_2, we have the following nine cases to take care of:

$$
\begin{array}{lll}
\text{(i)} & (X(v_1), X(v_2)) = (-1,-1) & (Y(v_1), Y(v_2)) = (-1,-1) \\
\text{(ii)} & (X(v_1), X(v_2)) = (-1,-1) & (Y(v_1), Y(v_2)) = (-1,+1) \\
\text{(iii)} & (X(v_1), X(v_2)) = (-1,-1) & (Y(v_1), Y(v_2)) = (+1,-1) \\
\text{(iv)} & (X(v_1), X(v_2)) = (-1,-1) & (Y(v_1), Y(v_2)) = (+1,+1) \\
\text{(v)} & (X(v_1), X(v_2)) = (-1,+1) & (Y(v_1), Y(v_2)) = (-1,+1) \\
\text{(vi)} & (X(v_1), X(v_2)) = (-1,+1) & (Y(v_1), Y(v_2)) = (+1,+1) \\
\text{(vii)} & (X(v_1), X(v_2)) = (+1,-1) & (Y(v_1), Y(v_2)) = (+1,-1) \\
\text{(viii)} & (X(v_1), X(v_2)) = (+1,-1) & (Y(v_1), Y(v_2)) = (+1,+1) \\
\text{(ix)} & (X(v_1), X(v_2)) = (+1,+1) & (Y(v_1), Y(v_2)) = (+1,+1)\,.
\end{array}
$$

(All other combinations of $(X(v_1), X(v_2))$ and $(Y(v_1), Y(v_2))$ are impossible since $X(\mathbf{Z}^2)$ and $Y(\mathbf{Z}^2)$ were chosen according to a monotone coupling.) In the cases (i), (v), (vii) and (ix) where the boundary conditions in X and Y are identical, we simply take $X(B)$ and $Y(B)$ to be identical. For each of the cases (ii) and (vi), we prescribe some monotone coupling (whose existence is guaranteed by Lemma 2.1), and use the same coupling for all bridges (independently) with the given pair of boundary conditions. The cases (iii) and (viii) are simply the "mirror images" of (ii) and (vi), and are treated in the same way.

This leaves the case (iv), for which we need to make the coupling more explicit: A direct calculation gives (thanks to our choice of $\beta = \frac{1}{4}\log(3)$)

$$
\begin{aligned}
&\mathbf{P}((X(z_1), X(z_2)) = (i,j) \,|\, (X(v_1), X(v_2)) = (-1,-1)) \\
&\quad = \left\{ \begin{array}{ll} \frac{1}{2} & \text{for } (i,j) = (-1,-1) \\ \frac{1}{6} & \text{for } (i,j) = (-1,+1),\ (+1,-1), \text{ or } (+1,+1), \end{array} \right.
\end{aligned} \tag{3}
$$

and, similarly,

$$\begin{aligned}\mathbf{P}((Y(z_1),Y(z_2)) = (i,j) \mid (Y(v_1),Y(v_2)) = (+1,+1)) \\ = \begin{cases} \frac{1}{2} & \text{for } (i,j) = (+1,+1) \\ \frac{1}{6} & \text{for } (i,j) = (-1,-1),\ (-1,+1), \text{ or } (+1,-1). \end{cases}\end{aligned} \tag{4}$$

These conditional distributions can be coupled by setting

$$((X(z_1),X(z_2)),(Y(z_1),Y(z_2))) = \begin{cases} ((-1,-1),(-1,+1)) & \text{with probability } \frac{1}{6} \\ ((-1,-1),(+1,-1)) & \text{with probability } \frac{1}{6} \\ ((-1,-1),(-1,-1)) & \text{with probability } \frac{1}{6} \\ ((-1,+1),(+1,+1)) & \text{with probability } \frac{1}{6} \\ ((+1,-1),(+1,+1)) & \text{with probability } \frac{1}{6} \\ ((+1,+1),(+1,+1)) & \text{with probability } \frac{1}{6}. \end{cases} \tag{5}$$

A key feature of this coupling, besides monotonicity, is that in each of the six cases in (5), we have that X and Y are equal in at least one of the vertices z_1 and z_2.

This is the coupling for which we claim that the properties listed in Theorem 3.1 hold. We have already noted that the distributions μ_- and μ_+ of X and Y are different. Furthermore, monotonicity and invariance of the coupling are immediate from the construction. It remains to show that, with probability 1, $\mathbf{Z}^2_{(23,3)}$ does not contain an infinite path of disagreements between X and Y. Such an path would have to have every third vertex lying in $\mathbf{Z}^2$, and for each $v \in \mathbf{Z}^2$ on this path we would have to have $X(v) = -1$ and $Y(v) = +1$, because the coupling is monotone. Suppose for contradiction that $v \in \mathbf{Z}^2$ is on such a path. Then there has to be another vertex $v' \in \mathbf{Z}^2$ three steps away from v on this path, such that $X(v') = -1$ and $Y(v') = +1$. But the coupling in (5) prevents any bridge between v and v' from having disagreements between X and Y at both of its vertices. Hence, an infinite path of disagreements between X and Y in $\mathbf{Z}^2_{(23,3)}$ cannot exist. The proof of Theorem 3.1 is therefore complete.

Remark. Our counterexample can be modified in various ways. For instance, the choice of the square lattice as the graph to decorate can be replaced by any graph G with a finite critical value β_c for nonuniqueness of Gibbs measures in the Ising model; we then just have to use $(n,3)$-decorations rather than $(23,3)$-decorations, where n is taken to be large enough so that $\frac{n}{2}\log(\frac{3\sqrt{3}}{5}) > \beta_c$.

4 A Markov chain coupling example

One of the main questions in Markov theory is whether or not the asymptotic behavior of a chain depends on its initial state. Fix a transition matrix M for a discrete time Markov chain on a finite or countable state space A. Let $X = \{X(t)\}_{t=0,1,2\ldots}$ and $Y = \{Y(t)\}_{t=0,1,2\ldots}$ be two Markov chain with this transition matrix, starting in two different states $X(0) = x$ and $Y(0) = y$. Write $\mu_{x,t}$ (resp. $\mu_{y,t}$) for the distribution of $X(t)$ (resp. $Y(t)$). A common way of making the above question precise is to ask whether or not

$$\lim_{t\to\infty} \mathrm{d}_{\mathrm{TV}}(\mu_{x,t},\mu_{y,t}) = 0 \tag{6}$$

where d_{TV} is the total variation distance, defined by

$$\mathrm{d}_{\mathrm{TV}}(\mu_x(t),\mu_y(t)) = \sup_{E\subset A} |\mu_{x,t}(E) - \mu_{y,t}(E)|\,.$$

A standard method for bounding $d_{TV}(\mu_{x,t}, \mu_{y,t})$ is to use the a coupling argument of the following kind (see, e.g., [13] or [16]). Suppose that we have a coupling of the chains X and Y with the property that if $X(\tau) = Y(\tau)$ for some τ, then the chains stay together, i.e., $X(t) = Y(t)$ for all $t \geq \tau$. Define the random time $T = \inf\{t : X(t) = Y(t)\}$. The total variation distance between $\mu_{x,t}$ and $\mu_{y,t}$ can then be bounded as follows:

$$\begin{aligned} d_{TV}(\mu_{x,t}, \mu_{y,t}) &= \sup_{E \subset A} |\mu_{x,t}(E) - \mu_{y,t}(E)| \\ &= \sup_{E \subset A} |\mathbf{P}(X(t) \in E) - \mathbf{P}(Y(t) \in E)| \\ &\leq \mathbf{P}(X(t) \neq Y(t)) \\ &= \mathbf{P}(T > t)\,. \end{aligned}$$

In particular, (6) holds if it can be established that $\mathbf{P}(T < \infty) = 1$, i.e., that the two chains eventually meet almost surely.

In most standard constructions of couplings, there is no problem in assuring that $X(t) = Y(t)$ for all $t \geq T$. It may therefore be tempting to think that $\mathbf{P}(T > t)$ is an upper bound on $d_{TV}(\mu_{x,t}, \mu_{y,t})$ for *any* coupling of the chains X and Y (i.e., that we do not need the assumption that $X(t) = Y(t)$ for all $t \geq T$). But this is false!

Proposition 4.1 *There exists a finite state Markov chain such that for two of its states x and y we have*

$$\lim_{t \to \infty} d_{TV}(\mu_{x,t}, \mu_{y,t}) > 0$$

while on the other hand there exists a coupling of the chains $X = \{X(t)\}_{t=0,1,2\ldots}$ and $Y = \{Y(t)\}_{t=0,1,2\ldots}$, starting at $X(0) = x$ and $Y(0) = y$, with the property that their first meeting time $T = \inf\{t : X(t) = Y(t)\}$ is finite with probability 1.

What we will now do is to construct an example which establishes Proposition 4.1. Related examples, together with conditions under which the conclusion $d_{TV}(\mu_{x,t}, \mu_{y,t}) \leq \mathbf{P}(T > t)$ is warranted, are discussed by Rosenthal [14].

Consider a Markov chain with state space $\{1, 2, \ldots, 6\}$ and transition probabilities given by Figure 2, where $a \in [0, 1]$. Observe that 5 and 6 are absorbing states, so that the chain fails to be irreducible. Let the chains $X = \{X(t)\}_{t=0,1,2\ldots}$ and $Y = \{Y(t)\}_{t=0,1,2\ldots}$ start in states $X(0) = 1$ and $Y(0) = 2$. We thus write $\mu_{1,t}$ and $\mu_{2,t}$ for the distributions of $X(t)$ and $Y(t)$. Note that $X(2) = X(3) = X(4) = \cdots$ with probability 1, and similarly for Y. A direct calculation shows that

$$\begin{aligned} \lim_{t \to \infty} d_{TV}(\mu_{1,t}, \mu_{2,t}) &= d_{TV}(\mu_{1,2}, \mu_{2,2}) \\ &= \mathbf{P}(X(2) = 5) - \mathbf{P}(Y(2) = 5) \\ &= (1 - 2a(1-a)) - 2a(1-a) \;=\; 1 - 4a(1-a) \end{aligned}$$

which is nonzero unless $a = \frac{1}{2}$.

In contrast, there exists, for any $a \in [1 - \frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}]$, a coupling of X and Y with the property that with probability 1, the chains take the same value either at time 1 or at time 2, so that $\mathbf{P}(T < \infty) = 1$. We will restrict our calculation to the borderline case $a = 1 - \frac{\sqrt{2}}{2}$, for which the coupling becomes particularly clean, and leave the general case as an exercise to the reader.

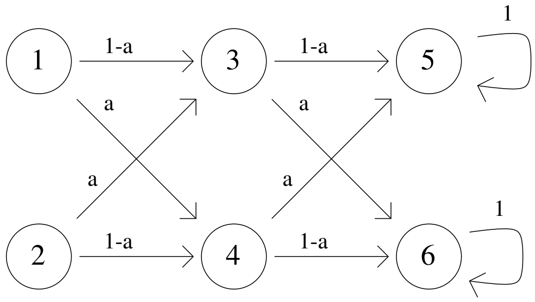

Figure 2: Transition graph for the Markov chain in our example. In the concrete calculation, we will take $a = \frac{\sqrt{2}}{2}$.

With $a = 1 - \frac{\sqrt{2}}{2}$, we get

$$(X(1), X(2)) = \begin{cases} (3,5) & \text{with probability } \frac{1}{2} \\ (3,6) & \text{with probability } \frac{\sqrt{2}-1}{2} \\ (4,5) & \text{with probability } \frac{3-2\sqrt{2}}{2} \\ (4,6) & \text{with probability } \frac{\sqrt{2}-1}{2}, \end{cases}$$

and, similarly,

$$(Y(1), Y(2)) = \begin{cases} (3,5) & \text{with probability } \frac{\sqrt{2}-1}{2} \\ (3,6) & \text{with probability } \frac{3-2\sqrt{2}}{2} \\ (4,5) & \text{with probability } \frac{\sqrt{2}-1}{2} \\ (4,6) & \text{with probability } \frac{1}{2}. \end{cases}$$

We can therefore couple X and Y by setting

$$((X(1), X(2)), (Y(1), Y(2)) = \begin{cases} ((3,5),(3,5)) & \text{with probability } \frac{\sqrt{2}-1}{2} \\ ((3,5),(3,6)) & \text{with probability } \frac{3-2\sqrt{2}}{2} \\ ((3,5),(4,5)) & \text{with probability } \frac{\sqrt{2}-1}{2} \\ ((3,6),(4,6)) & \text{with probability } \frac{\sqrt{2}-1}{2} \\ ((4,5),(4,6)) & \text{with probability } \frac{3-2\sqrt{2}}{2} \\ ((4,6),(4,6)) & \text{with probability } \frac{\sqrt{2}-1}{2}. \end{cases} \tag{7}$$

Note that in each of the six cases in (7), we have either $X(1) = Y(1)$, or $X(2) = Y(2)$. We therefore have $\mathbf{P}(T < \infty) = 1$, as desired. (Note also the similarlity with the coupling (5) in the disagreement percolation example.)

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. Stat. Phys.* **50**, 1–40.

[2] van den Berg, J. (1993) A uniqueness condition for Gibbs measures, with application to the 2–dimensional Ising antiferromagnet, *Commun. Math. Phys.* **152**, 161–166.

[3] van den Berg, J. (1999) On the absence of phase transition in the monomer-dimer model, *Perplexing Problems in Probability: Festschrift in Honor of Harry Kesten* (eds M. Bramson and R. Durrett), pp 185–195, Birkhäuser, Boston.

[4] van den Berg, J. and Maes, C. (1994) Disagreement percolation in the study of Markov fields, *Ann. Probab.* **22**, 749–763.

[5] van den Berg, J. and Steif, J. (1994) Percolation and the hard core lattice gas model, *Stoch. Proc. Appl.* **49**, 179–197.

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

[7] Coniglio, A., Nappi, C.R., Peruggi, F. and Russo, L. (1976) Percolation and phase transitions in the Ising model, *Commun. Math. Phys.* **51**, 315–323.

[8] Fortuin, C.M. and Kasteleyn, P.W. (1972) On the random-cluster model. I. Introduction and relation to other models, *Physica* **57**, 536–564.

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

[10] Häggström, O. (1998) Random-cluster representations in the study of phase transitions, *Markov Proc. Rel. Fields* **4**, 275–321.

[11] Häggström, O., Schonmann, R. and Steif, J. (2000) The Ising model on diluted graphs and strong amenability, *Ann. Probab.* **28**, 1111–1137.

[12] Jonasson, J. (2000) On disagreement percolation and maximality of the critical value for iid percolation, preprint, `http://www.math.chalmers.se/~jonasson/recent.html`

[13] Lindvall, T. (1992) *Lectures on the Coupling Method*, Wiley, New York.

[14] Rosenthal, J.S. (1997) Faithful couplings of Markov chains: now equals forever, *Adv. Appl. Math.* **18**, 372–381.

[15] Steif, J., personal communication.

[16] Thorisson, H. (2000) *Coupling, Stationarity, and Regeneration*, Springer, New York.

Dept of Mathematics
Chalmers University of Technology and Göteborg University
412 96 Göteborg
Sweden
`olleh@math.chalmers.se`
`http://www.math.chalmers.se/~olleh/`