RANDOM CLUSTER DYNAMICS FOR THE ISING MODEL IS RAPIDLY MIXING

BY HENG GUO[1,2,*] AND MARK JERRUM[1,†]

University of Edinburgh * *and Queen Mary, University of London* [†]

We show that the mixing time of Glauber (single edge update) dynamics for the random cluster model at $q = 2$ on an *arbitrary* n-vertex graph is bounded by a polynomial in n. As a consequence, the Swendsen–Wang algorithm for the ferromagnetic Ising model at any temperature also has a polynomial mixing time bound.

1. Introduction. The Ising model is perhaps the best known model in statistical physics, and it has also been widely studied from an algorithmic perspective. An instance of the model is an undirected graph G, together with a parameter $\beta > 0$. A *configuration* of the model is an assignment $\sigma \in \{0, 1\}^V$ of "spins" to the vertices of G. The *weight* $w(\sigma)$ of configuration σ is $\beta^{m(\sigma)}$ where $m(\sigma)$ is the number of monochromatic edges [edges $\{i, j\}$ with $\sigma(i) = \sigma(j)$] in G. It is of importance to compute the partition function of the system, which is the sum of weights $w(\sigma)$ over all configurations $\sigma \in \{0, 1\}^V$.

If $\beta < 1$, then the system is antiferromagnetic, and the partition function is computationally hard, even to approximate. However, in the ferromagnetic case ($\beta > 1$), Jerrum and Sinclair (1993) gave a *fully polynomial-time randomized approximation scheme* (FPRAS) for the partition function, which is efficient and achieves any specified relative error. A direct approach using Markov chain Monte Carlo (MCMC) on the spin configurations described above fails, as the spin model exhibits multiple phases for sufficiently large β on, say, a complete graph. On the other hand, the "even subgraphs" model has the same partition function as that of the Ising model, and it does form the basis for a successful application of MCMC, as was shown by Jerrum and Sinclair (1993). (See Sections 2 and 3 for details of the various models referred to in this Introduction.)

There is a third model which is equivalent to the Ising model in the sense of having the same partition function up to an easily computable factor, namely the random cluster model introduced by Fortuin and Kasteleyn (1972). Similar to the

even subgraphs model, the configurations of the random cluster model are subsets of the edge set of G. However, the random cluster model is more tightly related to the Ising model; in fact, a random Ising configuration can be obtained by colouring the connected components (clusters) of a random cluster configuration independently and uniformly at random by 0 and 1. Although we already have a polynomial-time algorithm for estimating the partition function of the Ising model, it is natural to wonder about the mixing time of the Gibbs sampler for random cluster configurations, which makes single edge-flip moves with Metropolis rejection probabilities. Indeed, it is conceivable that this dynamics mixes faster than the standard dynamics for the even subgraphs model.

Another reason for focusing on the random cluster model is that it extends the other two models in the following sense. There is a generalisation of the Ising model to $q \geq 2$ spins, known as the q-state Potts model, of which the Ising model is the special case $q = 2$. Although the even subgraphs and spin formulations are defined only for integer q, the random cluster model makes sense for arbitrary positive real q. Thus, by studying the dynamics of the random cluster model at $q = 2$, we may gain insight into the complexity of computing the partition function of the random cluster model at other values of q, particularly (for reasons that will be explained presently) in the range $0 \leq q < 2$. Stated in other terms, we would hope to gain information about the complexity of approximating the Tutte polynomial $T(G; x, y)$ in the region $0 \leq (x - 1)(y - 1) < 2$, and $x, y \geq 1$, about which nothing is currently known except for the point $x = y = 1$ and the (trivial) hyperbola $(x - 1)(y - 1) = 1$ [Goldberg and Jerrum (2008, 2014)].

In this paper, we prove that the Gibbs sampler (single edge-flip dynamics) for the random cluster model on an *arbitrary* graph mixes in time polynomial in $n = |V(G)|$, the number of vertices of G. (See Theorem 2.) One main tool is the well-known canonical paths technique for bounding mixing times via a parameter known as congestion, in the form presented by Sinclair (1992); see also [Diaconis and Stroock (1991)]. Another tool is a coupling between random cluster and even subgraph configurations discovered by Grimmett and Janson (2009). The existence of this coupling invites us to bound the congestion of the edge-flip dynamics on random cluster configurations in terms of the known bounds on congestion for the edge-flip dynamics on (augmented) even subgraph configurations, established by Jerrum and Sinclair (1993). Unfortunately, this translation between the models cannot be handled by existing comparison techniques [Diaconis and Saloff-Coste (1993), Dyer et al. (2006)], and an extension of comparison methods to the current situation is a contribution of the paper, and one that may find applications elsewhere.

Swendsen and Wang (1987) proposed a Markov chain that is widely considered to be an efficient method for sampling random cluster configurations (and Ising spin configurations) in practice. (Refer to Section 2.2 for a description of this Markov chain.) Prior to this work, the study of the Swendsen–Wang algorithm and related cluster dynamics was focused on special graphs, such as the complete

graph (the "mean-field" situation) or the two-dimensional lattice $\mathbb{Z}^2$. For complete graphs, the mixing time is very well understood for all $q \geq 1$ [Blanca and Sinclair (2015), Galanis, Štefankovič and Vigoda (2015), Long et al. (2014)]. For $\mathbb{Z}^2$, for all $q \geq 1$, the dynamics is fast mixing at all temperatures other than the critical one [Blanca and Sinclair (2016), Ullrich (2013, 2014a)]. When $q = 2$, a polynomial upper bound was known in the critical case on $\mathbb{Z}^2$ [Lubetzky and Sly (2012), Ullrich (2013, 2014a)], whereas on $\mathbb{Z}^3$ this was unknown. Recently, exponential mixing time lower bounds were established at the critical temperature on $\mathbb{Z}^2$ when $q > 4$, and for other q, the mixing time is at most subexponential [Duminil-Copin et al. (2016), Gheissari and Lubetzky (2016)], improving upon previous slow mixing results [Borgs et al. (1999)].

On the other hand, little is known for the Swendsen–Wang algorithm on arbitrary graphs. Ullrich (2014b) has shown that the relaxation time of the Swendsen–Wang dynamics is always no more than that of the edge-flip dynamics, so our result provides the first polynomial upper bound on the mixing time of the Swendsen–Wang algorithm for the ferromagnetic Ising model on *arbitrary* graphs (see Theorem 9). This confirms a conjecture raised by Sokal in the 1990s. However, the exponent in the bound we derive here is likely to be well above the true answer. Indeed, Peres has made and circulated [Peres (2017)] the following conjecture.

CONJECTURE 1 (Peres). *The mixing time of the Swendsen–Wang algorithm for the ferromagnetic Ising model is $O(n^{1/4})$ on an arbitrary graph.*

The conjectured bound $O(n^{1/4})$ comes from the mixing time at the critical temperature in a complete graph [Long et al. (2014)], which is believed to be the worst case. In contrast, the bound we show is $O(n^4 m^3)$ (see Theorem 9). Hopefully, the result presented here may be the first step on the road towards settling this conjecture.

Since the random cluster model is defined for all positive real q, it is natural to speculate on the mixing time of the Glauber dynamics when $q \neq 2$. For $q > 2$, the mixing time cannot be polynomial in general, owing to a first-order phase transition of the model on the complete graph identified by Bollobás, Grimmett and Janson (1996). This phase transition is a barrier to rapid mixing when $q > 2$, as shown by Gore and Jerrum (1999) when q is an integer, and by Blanca and Sinclair (2015) for general $q > 2$. In fact, there is no polynomial-time algorithm of any sort for evaluating the partition function of the random cluster model on general graphs when $q > 2$, unless there is an FPRAS for counting independent sets in a bipartite graph [Goldberg and Jerrum (2012)]. In contrast, in the range $0 \leq q \leq 2$ there is no known barrier to rapid mixing, and there is cause to be optimistic, particularly in the range $1 < q < 2$, in which the random cluster model is monotonic.

2. Definitions and the main results. For a graph $G = (V, E)$, we will use $n = |V|$ and $m = |E|$ throughout the paper. The ferromagnetic Ising model on a graph $G = (V, E)$ with parameter $\beta > 1$ is defined by the following: for any $\sigma \in \{0, 1\}^V$, the probability of being in configuration σ is

$$(1) \qquad \pi(\sigma) = \frac{\beta^{m(\sigma)}}{Z_{\mathrm{Ising}}(\beta)},$$

where $m(\sigma)$ is the number of mono-chromatic edges in σ, and its normalizing factor, the so-called partition function, is defined as

$$Z_{\mathrm{Ising}}(\beta) = \sum_{\sigma \in \{0,1\}^V} \beta^{m(\sigma)}.$$

The random cluster model with parameters (p, q) is defined on subsets of edges $S \subseteq E$ such that

$$(2) \qquad \pi_{\mathrm{RC}}(S) \propto p^{|S|}(1 - p)^{|E \setminus S|} q^{\kappa(S)},$$

where $\kappa(S)$ is the number of connected components in the subgraph (V, S). The partition function is

$$Z_{\mathrm{RC}}(p, q) = \sum_{S \subseteq E} p^{|S|}(1 - p)^{|E \setminus S|} q^{\kappa(S)}.$$

Denote this measure by $\pi_{\mathrm{RC};p,q}(\cdot)$ or simply $\pi_{\mathrm{RC}}(\cdot)$ when there is no confusion. We use Ω throughout this article to denote the state space of random cluster models, namely $\{0, 1\}^E$. It is well known that, for $q = 2$ and $p = 1 - \frac{1}{\beta}$, the random cluster model is equivalent to the Ising model in the sense that their partition functions are equal up to some easily computable factor [see (10)]. The random cluster model was introduced by Fortuin and Kasteleyn (1972), who also showed the equivalence between partition functions of the random cluster and the Ising model. Edwards and Sokal (1988) further elucidated the connection: for general integer $q > 0$, there is a coupling between the random cluster and the Potts model with q spins (where Ising is the special case of $q = 2$).

The (lazy) single bond flip dynamics P_{RC} is defined as follows based on the Metropolis filter:

$$(3) \quad P_{\mathrm{RC}}(x, y) = \begin{cases} \dfrac{1}{2m} \min\left\{1, \dfrac{\pi_{\mathrm{RC}}(y)}{\pi_{\mathrm{RC}}(x)}\right\} & \text{if } |x \oplus y| = 1; \\[2mm] 1 - \dfrac{1}{2m} \sum_{e \in E} \min\left\{1, \dfrac{\pi_{\mathrm{RC}}(x \oplus \{e\})}{\pi_{\mathrm{RC}}(x)}\right\} & \text{if } x = y; \\[2mm] 0 & \text{otherwise}, \end{cases}$$

where $x, y \in \Omega$. It is not hard to see, for example, by checking the detailed balance condition, that $\pi_{\mathrm{RC}}(\cdot)$ is the stationary distribution of P_{RC}. Note that the Markov chain is *lazy*, that is, it remains at its current state with probability at least $\frac{1}{2}$. This

eliminates the possibility of the transition matrix P having negative eigenvalues, and simplifies the analysis later.

For a Markov chain with transition matrix P and stationary distribution π, we are interested in its *mixing time*, that is, how fast it converges to the stationary distribution, defined as follows:

$$(4) \qquad \tau_\varepsilon(P) := \min\Big\{t : \max_{x \in \Omega} \| P^t(x, \cdot) - \pi \| \le \varepsilon \Big\},$$

where $\| \cdot \|$ is the total variation distance, namely

$$\| \pi - \pi' \| = \frac{1}{2} \sum_{x \in \Omega} |\pi(x) - \pi'(x)|.$$

Our main result is a mixing time upper bound of the single bond flip dynamics P_{RC} that is polynomial in the number of vertices.

THEOREM 2.   *For the random cluster model with parameters $0 < p < 1$ and $q = 2$, we have that*

$$\tau_\varepsilon(P_{\mathrm{RC}}) \le 8n^4 m^2 \big(m \ln(1 - p)^{-1} + \ln \varepsilon^{-1}\big).$$

2.1. *Preliminaries on Markov chains.*   If a Markov chain is lazy, then the second-largest eigenvalue (in absolute value) of the transition matrix P is just the second eigenvalue, denoted by λ_2. The *relaxation time* is defined as

$$(5) \qquad\qquad\qquad T_{\mathrm{rel}}(P) := \frac{1}{1 - \lambda_2}.$$

The relaxation time is closely related to the mixing time, as shown by the following proposition [see, e.g., Diaconis and Stroock (1991), Proposition 3].

PROPOSITION 3.   *For a lazy, ergodic, reversible Markov chain P and any initial state $x_0 \in \Omega$,*

$$\tau_\varepsilon(P) \le T_{\mathrm{rel}}\big(\ln \pi(x_0)^{-1} + \ln \varepsilon^{-1}\big).$$

Our goal is to bound $\tau_\varepsilon(P_{\mathrm{RC}})$. We can choose the initial state to be the empty set of edges, which has weight $\pi(\varnothing) = \frac{(1-p)^{|E|} 2^{|V|}}{Z_{\mathrm{RC}}}$. Also for $\beta = \frac{1}{1-p}$ we have $Z_{\mathrm{RC}}(p, 2) = \beta^{-|E|} Z_{\mathrm{Ising}}(\beta) \le 2^{|V|}$ and, therefore, $\pi(\varnothing) \ge (1 - p)^{|E|}$. Hence, $\ln \pi(x_0)^{-1} \le m \ln(1 - p)^{-1}$.

Canonical paths are a useful technique to bound the relaxation time of Markov chains, introduced by Jerrum and Sinclair (1989), Sinclair and Jerrum (1989). Let $\Gamma = \{\gamma_{xy} : x, y \in \Omega\}$ be a collection of paths, where $\gamma_{xy} = \{z_0, \ldots, z_\ell\}$ is a "canonical" path from $x = z_0$ to $y = z_\ell$ of length ℓ where each step (z_i, z_{i+1}) is a valid

transition of the Markov chain, namely $P(z_i, z_{i+1}) > 0$. The *congestion* $\varrho(\Gamma)$ associated with these paths is

$$(6) \qquad \varrho(\Gamma) := \max_{(z,z') \in \Omega^2, P(z,z') > 0} \frac{L}{\pi(z) P(z,z')} \sum_{\substack{x,y \in \Omega^2 \\ \gamma_{xy} \ni (z,z')}} \pi(x)\pi(y),$$

where $L = L(\Gamma)$ denotes the maximum length of paths in Γ.

A more general technique is provided by the flow formulation for congestion. A (valid) flow Γ is a collection of paths, where each path $\gamma \in \Gamma$ is assigned a weight $wt(\gamma)$, such that

$$(7) \qquad \sum_{\gamma \text{ is from } x \text{ to } y} wt(\gamma) = \pi(x)\pi(y).$$

The congestion of Γ is defined as

$$(8) \qquad \varrho(\Gamma) := \max_{(z,z') \in \Omega^2, P(z,z') > 0} \frac{L}{\pi(z) P(z,z')} \sum_{\gamma \in \Gamma, (z,z') \in \gamma} wt(\gamma).$$

The canonical paths are just a flow where for each pair (x, y) there is only one path with positive weight.

Sinclair (1992) showed that the relaxation time can be bounded by the congestion of any flow Γ.

PROPOSITION 4. *For a lazy, ergodic, reversible Markov chain P and any flow Γ,*

$$T_{\text{rel}}(P) \leq \varrho(\Gamma).$$

The main task is to design a good flow Γ_{RC} so that $\varrho(\Gamma_{\text{RC}})$ is bounded by a polynomial in n.

THEOREM 5. *There is a collection Γ_{RC} of paths for the random cluster dynamics such that its congestion $\varrho(\Gamma_{\text{RC}}) \leq 8m^2 n^4$.*

Theorem 5 will be proved in Section 4. In particular, it implies a bound on the relaxation time of P_{RC}.

COROLLARY 6. *For the Markov chain, P_{RC}, $T_{\text{rel}}(P_{\text{RC}}) \leq 8m^2 n^4$.*

Theorem 2 follows from Proposition 3 and Corollary 6.

2.2. *The Swendsen–Wang algorithm.* Swendsen and Wang (1987) proposed the following algorithm to sample Ising configurations with parameter β. For an Ising configuration $\sigma = \sigma_t \in \{0, 1\}^V$ at time t:

- Let M be the set of monochromatic edges under σ, that is, $(u, v) \in M$ if $\sigma(u) = \sigma(v)$.
- For each edge $e \in M$, delete it with probability β^{-1}. Let M' denote the set of monochromatic edges that were not deleted.
- In the subgraph (V, M'), for each connected component, choose uniformly at random and independently from $\{0, 1\}$, and assign the chosen spin to all vertices in that component.

The resulting spin configuration is σ_{t+1} at time $t + 1$.

Ullrich (2014b) showed that the relaxation time of the Swendsen–Wang algorithm is no larger than that of the single bond flip dynamics. In fact, the result of [Ullrich (2014b)] holds for any integer $q > 0$, but here we only need it for $q = 2$.

PROPOSITION 7 [Ullrich (2014b), Theorem 5]. *Let P_{SW} be the transition matrix of the Swendsen–Wang algorithm to sample Ising configurations with a parameter $\beta > 1$. Let P_{RC} be the transition matrix of the corresponding single bond flip dynamics for the random cluster model with $p = 1 - \beta^{-1} \in (0, 1)$ and $q = 2$. Then for any graph G, $T_{rel}(P_{SW}) \le T_{rel}(P_{RC})$.*

Combining Corollary 6 and Proposition 7, we have the following.

COROLLARY 8. *For the Swendsen–Wang algorithm, $T_{rel}(P_{SW}) \le 8m^2n^4$.*

Again, we use Corollary 8 together with Proposition 3, implying a polynomial mixing time upper bound for the Swendsen–Wang algorithm.

THEOREM 9. *Let P_{SW} be the transition matrix of the Swendsen–Wang algorithm to sample Ising configurations with a parameter $\beta > 1$. We have that*

$$\tau_\varepsilon(P_{SW}) \le 8n^4 m^2 \big(m \ln \beta + \ln \varepsilon^{-1}\big).$$

3. Random even subgraphs. There is yet another formalism of the Ising model, that is, the so-called "high-temperature expansion" or even subgraphs model. We still pick a subset of edges $S \subseteq E$ but with the further restriction that every vertex in the induced subgraph (V, S) has even degree. Denote by $\Omega_{even}(G)$ the state space of all such even subgraphs of G. We usually simply write Ω_{even} when there is no confusion. In this even subgraphs model, we want to sample from Ω_{even} with parameter $p \le 1/2$, so that edges are more inclined to be "out" than "in". That is, for any $S \in \Omega_{even}$,

$$(9) \qquad\qquad \pi(S) \propto p^{|S|}(1 - p)^{|E \setminus S|}$$

and

$$Z_{\text{even}}(p) = \sum_{S \in \Omega_{\text{even}}} p^{|S|}(1-p)^{|E \setminus S|}.$$

Distributions (1), (2) and (9) have in fact the same partition function, up to certain scaling factors:

$$(10) \qquad Z_{\text{Ising}}(\beta) = \beta^{|E|} Z_{\text{RC}}\left(1 - \frac{1}{\beta}, 2\right) = 2^{|V|} \beta^{|E|} Z_{\text{even}}\left(\frac{1}{2}\left(1 - \frac{1}{\beta}\right)\right).$$

The first equivalence was discovered by Fortuin and Kasteleyn (1972); [see also [Grimmett (2006)]]. The second one is also a classical result, known as early as in [van der Waerden (1941)]. More detailed explanations can be found in Appendix A.

Grimmett and Janson (2009) discovered the following coupling between even subgraphs and random cluster configurations.

THEOREM 10 [Grimmett and Janson (2009), Theorem 3.5]. *Take a random even subgraph S from distribution* (9) *with parameter $p \le 1/2$, and add each edge $e \notin S$ independently with probability $\frac{p}{1-p}$ to get a random subgraph R. Then R is a random cluster configuration, that is, it satisfies* (2) *with parameters* $(2p, 2)$.

For completeness, we give a proof of Theorem 10.

PROOF. The number of even subgraphs of a (not necessarily simple) graph $G = (V, E)$ is well known to be

$$(11) \qquad |\Omega_{\text{even}}(G)| = 2^{|E|-|V|+\kappa(G)},$$

where $\kappa(G)$ is the number of connected components of G.

For each $r \subseteq E$,

$$\Pr(R = r) \propto \sum_{s \subseteq r, s \text{ even}} \left(\frac{p}{1-p}\right)^{|s|} \left(\frac{p}{1-p}\right)^{|r \setminus s|} \left(\frac{1-2p}{1-p}\right)^{|E \setminus r|}$$

$$\propto p^{|r|}(1-2p)^{|E \setminus r|} N(r),$$

where $N(r)$ is the number of even subgraphs of (V, r). By (11), $N(r) = 2^{|r|-|V|+\kappa(r)}$. Hence,

$$\Pr(R = r) \propto (2p)^{|r|}(1-2p)^{|E \setminus r|} 2^{\kappa(r)}. \qquad \square$$

However, it is not clear how to sample from Ω_{even} with edge weights directly in an efficient way, partly because of the rigid structure of the all even requirement. On the other hand, Jerrum and Sinclair (1993) designed a Markov chain to do so by moving among all subgraphs, but with each odd degree vertex incurring a penalty.

Note that the Jerrum–Sinclair Markov chain together with the Grimmett–Janson coupling (Theorem 10) yields an efficient sampler for random cluster models and Ising configurations. It is more straightforward and efficient than the one given by Randall and Wilson (1999), which also uses the Jerrum–Sinclair chain.

An alternative (but similar) Markov chain is to move between even subgraphs and near-even subgraphs, for which we allow exactly two odd degree vertices (or "holes"). This is the so-called "worm" process, introduced by Prokof'ev and Svistunov (2001).

Let Ω_k be the collection of subgraphs where k many vertices have odd degrees. Then $\Omega_0 = \Omega_{\text{even}}$ and the state space Ω_{worm} of the "worm" process is $\Omega_{\text{worm}} := \Omega_0 \cup \Omega_2$. For each pair of vertices (u, v) such that $u \neq v$, denote by $\Omega(u, v)$ the set of subgraphs of G in which u and v have odd degrees and all other vertices are even. Then

$$\Omega_2 = \bigcup_{u,v \in V} \Omega(u, v).$$

For a subset of edges $S \subseteq E$, let $w_p(S) := p^{|S|}(1 - p)^{|E \setminus S|}$. We give a penalty of n^{-2} to each near-even subgraph:

$$(12) \qquad w_{\text{worm}}(S) := \begin{cases} w_p(S) & \text{if } S \in \Omega_0; \\ n^{-2} w_p(S) & \text{if } S \in \Omega_2; \\ 0 & \text{otherwise.} \end{cases}$$

The "worm" measure is defined as the following:

$$(13) \qquad \pi_{\text{worm}}(S) := \begin{cases} \dfrac{w_{\text{worm}}(S)}{Z_{\text{worm}}(p)} & \text{if } S \in \Omega_{\text{worm}}; \\ 0 & \text{otherwise,} \end{cases}$$

where $Z_{\text{worm}}(p) = \sum_{S \in \Omega_{\text{worm}}} w_{\text{worm}}(S)$.

The winding idea of [Jerrum and Sinclair (1993)] provides a way to design canonical paths between states in Ω_{worm} with low congestion. We will not need to analyze it in full detail for the worm process. Instead, we only care about paths from one even subgraph to another.

THEOREM 11. *There is a collection of paths*

$$\Gamma_{\text{worm}} = \{\gamma_{xy} \mid x, y \in \Omega_0\}$$

equipped with a weight function $wt(\cdot)$ *such that the following hold*:

1. *For any path* $\gamma \in \Gamma_{\text{worm}}$ *and any state* $w \in \gamma$, $w \in \Omega_{\text{worm}}$;
2. $wt(\gamma_{xy}) = \pi_{\text{even}}(x)\pi_{\text{even}}(y)$;
3. *Each state* w *appears at most once in* γ *and* $L(\Gamma_{\text{worm}}) \leq m$;

4. *for any transition (w, w') where $w' = w \oplus \{e\}$,*

$$\sum_{\gamma \in \Gamma_{\text{worm}} \text{ and } \gamma \ni (w, w')} wt(\gamma) \leq n^4 \pi_{\text{worm}}(w).$$

Moreover, in the special case $w' = w \cup \{e\}$ for some $e \notin w$, we have the additional bound

$$\sum_{\gamma \in \Gamma_{\text{worm}} \text{ and } \gamma \ni (w, w')} wt(\gamma) \leq n^4 \pi_{\text{worm}}(w) \frac{p}{1 - p}.$$

Note that Γ_{worm} is not a complete collection of canonical paths for $\pi_{\text{worm}}(\cdot)$. The proof of Theorem 11 is an adaptation of [Jerrum and Sinclair (1993)] and is given in Appendix B. Note that Collevecchio et al. (2016) give an analysis of a complete set of canonical paths for the worm process, but their result does not quite fit our situation.

Since paths in Γ_{worm} go through Ω_{worm} instead of Ω_{even}, we need to extend Theorem 10 to Ω_{worm}. It will no longer be exact.

Take a random subgraph S from π_{worm} (13) with parameter $p \leq 1/2$. Again we add each edge $e \notin S$ independently with probability $\frac{p}{1-p}$ to get R. Denote by $\hat{\pi}(\cdot)$ the law of R.

LEMMA 12. *For any $R \subseteq E$,*

$$\frac{\hat{\pi}(R)}{\pi_{\text{RC};2p,2}(R)} \leq \frac{3}{2}.$$

PROOF. Similar to the proof of Theorem 10, it is not hard to see that

$$\hat{\pi}(R) \propto p^{|R|}(1 - 2p)^{|E \setminus R|}(N(R) + n^{-2}N'(R)),$$

where $N(R)$, as before, is the number of even subgraphs of (V, R), and $N'(R)$ is the number of subgraphs of R that belong to Ω_2. Note that for each near-even subgraph there is a penalty of n^{-2} for its weight [see (12)]. We use (11) to count the number of even subgraphs of R, which is $2^{|R|-|V|+\kappa(R)}$.

Let $\Omega_R(u, v)$ be the set of near-even subgraphs of R with holes u and v. If u, v are in different connected components of (V, R), then there is no possible such subgraph and $|\Omega_R(u, v)| = 0$. Otherwise, u, v are in the same component of (V, R), and we can add an extra edge (u, v) to R to get a graph R'. Applying (11) to R', we get that

$$N(R') = 2^{|R|+1-|V|+\kappa(R)} = N(R) + |\Omega_R(u, v)|.$$

The second equality is because each even subgraph of R' either uses the new edge (u, v) or not. If it does not use (u, v), then it is an even subgraph of R. Otherwise,

it is [after removing the edge (u, v)] a near-even subgraph of R with holes u and v. Hence,

$$\left|\Omega_R(u, v)\right| = 2^{|R|-|V|+\kappa(R)},$$

as $N(R) = 2^{|R|-|V|+\kappa(R)}$.

Let $c(R)$ be the number of pairs of vertices from every component of (V, R). That is,

$$(14) \qquad c(R) := \sum_{i=1}^{\kappa(R)} \binom{n_i}{2},$$

where n_i is the size of the ith component of (V, R) with the convention that $\binom{1}{2} = 0$. Then we have that

$$N'(R) = 2^{|R|-|V|+\kappa(R)} c(R),$$

and

$$\widehat{\pi}(R) \propto (2p)^{|R|}(1 - 2p)^{|E \setminus R|} 2^{\kappa(R)} \left(1 + \frac{c(R)}{n^2}\right).$$

The lemma follows by noticing that $0 \le c(R) \le \frac{n(n-1)}{2}$.   □

4. Lifting canonical paths. Let $p \le 1/2$ be the parameter of the even subgraph and the worm measure. Let Γ_{worm} be the collection of paths as in Theorem 11. We will use Lemma 12 to lift Γ_{worm} to a flow Γ_{RC} for P_{RC}, the single edge-flip Markov chain for the random cluster model with parameter $2p$. This flow Γ_{RC} will be the one used in the proof of Theorem 5.

We first construct a flow Γ'_{RC} from Γ_{worm}. Let $\gamma = \{w_0, w_1, \ldots, w_\ell\}$ be a path in Γ_{worm} where $w_0, w_\ell \in \Omega_0$, and $\ell \le L(\Gamma_{\text{worm}})$. We lift γ to a flow (random path) as follows. First, we add each edge $e \notin w_0$ with probability $p' = \frac{p}{1-p}$ independently as in Lemma 12, to obtain the starting state Z_0 of the path. In other words, letting

$$\delta(w, z) := (p')^{|z \setminus w|}(1 - p')^{|E \setminus z|},$$

for subsets of edges $w \subseteq z \subseteq E$, we draw a superset Z_0 of w_0 such that $\Pr(Z_0 = z) = \delta(w_0, z)$ for any $z \supseteq w_0$. Note that

$$\pi_{\text{RC}}(z) = \sum_{w \subseteq z, w \in \Omega_0} \pi_{\text{even}}(w)\delta(w, z)$$

by Theorem 10, and

$$\widehat{\pi}(z) = \sum_{w \subseteq z, w \in \Omega_{\text{worm}}} \pi_{\text{worm}}(w)\delta(w, z)$$

by definition.

We construct $Z_1, \ldots, Z_\ell$ inductively. Given Z_{k-1} for $1 \le k \le \ell$, we construct Z_k by mimicking the transition from w_{k-1} to w_k while ensuring that

$$\Pr_\gamma(Z_k = z) = \delta(w_k, z),$$

for any $z \supseteq w_k$ at the same time. Here, the subscript γ emphasises that probabilities are with respect to a fixed path γ. By the induction hypothesis, $\Pr_\gamma(Z_{k-1} = z) = \delta(w_{k-1}, z)$ for any $z \supseteq w_{k-1}$. For Z_k, there are two cases:

- If $w_k = w_{k-1} \cup \{e\}$ for some edge $e \notin w_{k-1}$, then let $Z_k = Z_{k-1} \cup \{e\}$. We have that

$$\begin{aligned}
\Pr_\gamma(Z_k = z) &= \Pr_\gamma(Z_{k-1} = z) + \Pr_\gamma(Z_{k-1} = z \setminus \{e\}) \\
&= \delta(w_{k-1}, z) + \delta(w_{k-1}, z \setminus \{e\}) \\
&= \delta(w_k, z) p' + \delta(w_k, z)(1 - p') = \delta(w_k, z),
\end{aligned}$$

  for any $z \supseteq w_k$.
- If $w_k = w_{k-1} \setminus \{e\}$ for some edge $e \in w_{k-1}$, then let $Z_k = Z_{k-1}$ with probability p' and $Z_k = Z_{k-1} \setminus \{e\}$ with probability $1 - p'$. For any $z \supseteq w_k$ such that $e \in z$,

$$\Pr_\gamma(Z_k = z) = \Pr_\gamma(Z_{k-1} = z) p' = \delta(w_{k-1}, z) p' = \delta(w_k, z),$$

  and for any $z \supseteq w_k$ such that $e \notin z$,

$$\begin{aligned}
\Pr_\gamma(Z_k = z) &= \Pr_\gamma(Z_{k-1} = z \cup \{e\})(1 - p') \\
&= \delta(w_{k-1}, z \cup \{e\})(1 - p') = \delta(w_k, z).
\end{aligned}$$

Given γ, the lifted path $Z = \{Z_0, Z_1, \ldots, Z_\ell\}$ is constructed as above. A particular flow path $\zeta = \{z_0, z_1, \ldots, z_\ell\}$ in the random cluster world may be lifted from multiple paths from Γ_{worm}, and its weight is assigned to be the aggregation:

$$wt(\zeta) = \sum_{\gamma \in \Gamma_{\text{worm}}} wt(\gamma) \Pr_\gamma(Z = \zeta).$$

This completes the construction of Γ'_{RC}.

However, Γ'_{RC} is not a valid flow for $\pi_{\text{RC}}(\cdot)$. Recall that the valid flow should satisfy (7). An equivalent view of (7) is that if we draw a random path according to the weight function $wt(\cdot)$, the initial and final states should be *independently* distributed according to $\pi_{\text{RC}}(\cdot)$. Under this view, the problem with Γ'_{RC} is that Z_0 and Z_ℓ are correlated, even though the marginal distribution of each is $\pi_{\text{RC}}(\cdot)$.

We resolve this issue next by constructing Γ_{RC}. Given $\gamma \in \Gamma_{\text{worm}}$ with length ℓ, we construct $Z_0, \ldots, Z_\ell$ the same as in Γ'_{RC}. To repair the distribution of Z_ℓ, we append further transitions to re-randomize edges that are not in w_ℓ. More precisely, let $\{e_1, e_2, \ldots, e_k\}$ be the edges that are not in w_ℓ where $k = |E \setminus w_\ell|$. Given $Z_{\ell+i-1}$ for $1 \le i \le k$, let $Z_{\ell+i} = Z_{\ell+i-1} \setminus \{e_i\}$ with probability $1 - p'$ and

$Z_{\ell+i} = Z_{\ell+i-1} \cup \{e_i\}$ with probability p'. As in Γ'_{RC}, for a random cluster path $\zeta = \{z_0, z_1, \ldots, z_{\ell+k}\}$, its weight is defined to be

$$wt(\zeta) = \sum_{\gamma \in \Gamma_{\text{worm}}} wt(\gamma) \Pr_{\gamma}(Z = \zeta).$$

This completes the construction of Γ_{RC}. The longest path in Γ_{RC} has length at most $L(\Gamma_{\text{worm}}) + m$, that is, $L(\Gamma_{RC}) \leq L(\Gamma_{\text{worm}}) + m \leq 2m$.

Fix a path $\gamma = \{w_0, w_1, \ldots, w_\ell\}$. For any $0 \leq i \leq \ell$ and $z \supseteq w_i$, we have $\Pr_{\gamma}(Z_i = z) = \delta(w_i, z)$, because of the construction of Γ'_{RC}. Moreover, for any $1 \leq i \leq |E \setminus w_\ell|$ and $z \supseteq w_\ell$, we have $\Pr_{\gamma}(Z_{\ell+i} = z) = \delta(w_\ell, z)$. This can be shown by inductively going through the construction above. The re-randomization does not change the marginal distribution but removes the correlation between Z_0 and $Z_{\ell'}$, where $\ell' = \ell + |E \setminus w_\ell|$ (conditional on γ).

LEMMA 13.   *The flow Γ_{RC} is valid for $\pi_{RC}(\cdot)$, namely it satisfies* (7).

PROOF.   We verify (7) as follows:

$$\sum_{\zeta \text{ is from } x \text{ to } y} wt(\zeta)$$

$$= \sum_{\substack{w \subseteq x, w' \subseteq y \\ w, w' \in \Omega_0}} \sum_{\gamma \text{ is from } w \text{ to } w'} wt(\gamma) \Pr_{\gamma}(Z_0 = x, Z_{\ell'} = y)$$

$$= \sum_{\substack{w \subseteq x, w' \subseteq y \\ w, w' \in \Omega_0}} \sum_{\gamma \text{ is from } w \text{ to } w'} wt(\gamma) \Pr_{\gamma}(Z_0 = x) \Pr_{\gamma}(Z_{\ell'} = y)$$

$$= \sum_{\substack{w \subseteq x, w' \subseteq y \\ w, w' \in \Omega_0}} \sum_{\gamma \text{ is from } w \text{ to } w'} wt(\gamma) \delta(w, x) \delta(w', y)$$

$$= \sum_{\substack{w \subseteq x, w' \subseteq y \\ w, w' \in \Omega_0}} \pi_{\text{even}}(w) \pi_{\text{even}}(w') \delta(w, x) \delta(w', y)$$

$$= \left(\sum_{w \subseteq x, w \in \Omega_0} \pi_{\text{even}}(w) \delta(w, x) \right) \left(\sum_{w' \subseteq y, w' \in \Omega_0} \pi_{\text{even}}(w') \delta(w', y) \right)$$

$$= \pi_{RC}(x) \pi_{RC}(y),$$

where in the last step we use Theorem 10.   □

LEMMA 14.   *Let $2p \leq 1$ be the parameter for the random cluster model*:

1. *For a transition (z, z') where $z' = z \cup \{e\}$ for some $e \notin z$,*

$$\sum_{\zeta \in \Gamma_{RC}, \zeta \ni (z,z')} wt(\zeta) \leq \frac{p}{1-p} \cdot 2n^4 \pi_{RC}(z).$$

2. *For a transition (z, z') where $z' = z \setminus \{e\}$ for some $e \in z$,*

$$\sum_{\zeta \in \Gamma_{RC}, \zeta \ni (z,z')} wt(\zeta) \leq \frac{1-2p}{1-p} \cdot 2n^4 \pi_{RC}(z).$$

3. *For a transition (z, z),*

$$\sum_{\zeta \in \Gamma_{RC}, \zeta \ni (z,z)} wt(\zeta) \leq 2mn^4 \pi_{RC}(z).$$

PROOF. Fix γ, let Z be a random path lifted from γ and ℓ be the length of γ. Thus the path is $\gamma = (w_0, \ldots, w_\ell)$ and, in particular, the final state of the path is w_ℓ. For a state $w \in \gamma$, let $i(\gamma, w)$ be index of w in γ and $k(w, e)$ be the index of e in the set $E \setminus w$, following the enumeration mentioned previously. Any w only appears once in $\gamma \in \Gamma_{worm}$ and hence $i(\gamma, w)$ is well defined.

We want to bound the traffic in Γ_{RC} that goes through (z, z'). Let $p' = \frac{p}{1-p}$. Depending on z', we have three cases:

1. First, assume that $z' = z \cup \{e\}$ where $e \notin z$. The traffic may be from Γ'_{RC} transitions or from the part we append at the end of each Γ'_{RC} path. Hence we have the following bound:

$$\sum_{\zeta \in \Gamma_{RC}, \zeta \ni (z,z')} wt(\zeta)$$

$$= \sum_{w \subseteq z} \left(\sum_{\gamma \ni (w, w \cup \{e\})} wt(\gamma) \Pr_\gamma \left(Z_{i(\gamma,w)} = z, Z_{i(\gamma,w)+1} = z' \right) \right.$$

$$\left. + \sum_{\gamma = (w_1, \ldots, w_\ell), w_\ell = w} wt(\gamma) \Pr_\gamma \left(Z_{\ell + k(w,e)-1} = z, Z_{\ell + k(w,e)} = z' \right) \right)$$

$$= \sum_{w \subseteq z} \left(\sum_{\gamma \ni (w, w \cup \{e\})} wt(\gamma) \Pr_\gamma (Z_{i(\gamma,w)} = z) \right.$$

$$\left. + \sum_{\gamma, w_\ell = w} wt(\gamma) \Pr_\gamma (Z_{\ell + k(w,e)-1} = z) p' \right)$$

$$= \sum_{w \subseteq z} \delta(w, z) \left(\sum_{\gamma \ni (w, w \cup \{e\})} wt(\gamma) + \sum_{\gamma, w_\ell = w} wt(\gamma) p' \right).$$

Hence by Theorem 11,

$$\sum_{\zeta \in \Gamma_{RC}, \zeta \ni (z,z')} wt(\zeta)$$

$$= \sum_{w \subseteq z} \delta(w, z)\left(\sum_{\gamma \ni (w, w \cup \{e\})} wt(\gamma) + \sum_{\gamma, w_\ell = w} wt(\gamma) p' \right)$$

$$\leq \sum_{w \subseteq z} \delta(w, z)\left(n^4 \pi_{\text{worm}}(w)\frac{p}{1-p} + \pi_{\text{even}}(w)p' \right)$$

$$= p'n^4 \sum_{w \subseteq z} \delta(w, z)\pi_{\text{worm}}(w) + p' \sum_{w \subseteq z, w \in \Omega_0} \delta(w, z)\pi_{\text{even}}(w)$$

$$= p'n^4 \hat{\pi}(z) + p'\pi_{RC}(z)$$

$$\leq 2p'n^4 \pi_{RC}(z),$$

where we use Lemma 12 in the last line. Also note that $\pi_{\text{even}}(w) = 0$ if $w \notin \Omega_0$.

2. Next, assume that $z' = z \setminus \{e\}$ where $e \in z$. Similar to the previous case, we have that

$$\sum_{\zeta \in \Gamma_{RC}, \zeta \ni (z,z')} wt(\zeta)$$

$$= \sum_{w \subseteq z, w \ni e} \sum_{\gamma \ni (w, w \setminus \{e\})} wt(\gamma) \Pr_\gamma\big(Z_{i(\gamma,w)} = z, Z_{i(\gamma,w)+1} = z'\big)$$

$$+ \sum_{w \subseteq z, w \not\ni e} \sum_{\gamma, w_\ell = w} wt(\gamma) \Pr_\gamma\big(Z_{\ell+k(w,e)-1} = z, Z_{\ell+k(w,e)} = z'\big)$$

$$= \sum_{w \subseteq z, w \ni e} \sum_{\gamma \ni (w, w \setminus \{e\})} wt(\gamma) \Pr_\gamma\big(Z_{i(\gamma,w)} = z\big)(1 - p')$$

$$+ \sum_{w \subseteq z, w \not\ni e} \sum_{\gamma, w_\ell = w} wt(\gamma) \Pr_\gamma\big(Z_{\ell+k(w,e)-1} = z\big)(1 - p')$$

$$= \sum_{w \subseteq z, w \ni e} (1 - p')\delta(w, z) \sum_{\gamma \ni (w, w \setminus \{e\})} wt(\gamma)$$

$$+ \sum_{w \subseteq z, w \not\ni e} (1 - p')\delta(w, z) \sum_{\gamma, w_\ell = w} wt(\gamma).$$

Again we use Theorem 11 and Lemma 12:

$$\sum_{\zeta \in \Gamma_{RC}, \zeta \ni (z,z')} wt(\zeta)$$

$$\leq \sum_{w \subseteq z} \delta(w, z)(1 - p')\big(n^4 \pi_{\text{worm}}(w) + \pi_{\text{even}}(w)\big)$$

$$\leq (1 - p')n^4 \sum_{w \subseteq z} \delta(w, z)\pi_{\text{worm}}(w)$$

$$+ (1 - p') \sum_{w \subseteq z, w \in \Omega_0} \delta(w, z)\pi_{\text{even}}(w)$$

$$= (1 - p')n^4 \hat{\pi}(z) + (1 - p')\pi_{\text{RC}}(z)$$

$$\leq 2(1 - p')n^4 \pi_{\text{RC}}(z).$$

3. At last, we handle the case that $z = z'$. Then we have the following bound:

$$\sum_{\zeta \in \Gamma_{\text{RC}}, \zeta \ni (z,z)} wt(\zeta)$$

$$= \sum_{w \subseteq z} \left(\sum_{\gamma \ni w} wt(\gamma) \Pr_\gamma (Z_{i(\gamma,w)} = z, Z_{i(\gamma,w)+1} = z) \right.$$

$$\left. + \sum_{\gamma, w_\ell = w} wt(\gamma) \sum_{i=1}^{|E \backslash w|} \Pr_\gamma (Z_{\ell(\gamma)+i-1} = z, Z_{\ell(\gamma)+i} = z) \right)$$

$$\leq \sum_{w \subseteq z} \left(\sum_{\gamma \ni w} wt(\gamma) \Pr_\gamma (Z_{i(\gamma,w)} = z) + \sum_{\gamma, w_\ell = w} wt(\gamma)\delta(w, z)|E \backslash w| \right)$$

$$= \sum_{w \subseteq z} \delta(w, z) \left(\sum_{e \in z} \sum_{\gamma \ni (w, w \oplus \{e\})} wt(\gamma) + |E \backslash w| \sum_{\gamma, w_\ell = w} wt(\gamma) \right).$$

By Theorem 11 and Lemma 12,

$$\sum_{\zeta \in \Gamma_{\text{RC}}, \zeta \ni (z,z)} wt(\zeta) \leq m \sum_{w \subseteq z} \delta(w, z)\left(n^4 \pi_{\text{worm}}(w) + \pi_{\text{even}}(w)\right)$$

$$\leq 2mn^4 \pi_{\text{RC}}(z). \qquad \square$$

Note that in Lemma 14 we analyzed self-loop transitions (z, z) as well. We may remove self-loop transitions from the flow Γ_{RC} without increasing the congestion. However, doing so would make one transition in Γ_{RC} correspond to potentially more than one steps before the lifting, making the analysis more difficult.

Now we are ready to prove Theorem 5.

PROOF OF THEOREM 5. The flow Γ_{RC} is constructed at the beginning of this section and is verified in Lemma 13. We analyze its congestion in the following. There are three cases depending on the transition (z, z'). Note that the parameter of the random cluster is $2p$, where $p < 1/2$ is the parameter for the even subgraph model.

For any transition (z, z') where $z' = z \cup \{e\}$ for some $e \notin z$,

$$\frac{L(\Gamma_{RC})}{\pi_{RC}(z) P_{RC}(z, z')} \sum_{\zeta \in \Gamma_{RC}, \zeta \ni (z, z')} wt(\zeta)$$

$$\leq \frac{L(\Gamma_{RC})}{\pi_{RC}(z) P_{RC}(z, z')} \cdot \frac{p}{1 - p} \cdot 2n^4 \pi_{RC}(z)$$

$$\leq 2mn^4 \cdot \frac{p}{1 - p} \cdot \frac{2m}{\min\{1, \frac{2p}{2(1-2p)}\}}$$

$$\leq 4m^2 n^4,$$

where we use Lemma 14 in the first inequality and $p \leq 1/2$ in the last.

Similarly, for a transition (z, z') where $z' = z \setminus \{e\}$ for some $e \in z$,

$$\frac{L(\Gamma_{RC})}{\pi_{RC}(z) P_{RC}(z, z')} \sum_{\zeta \in \Gamma_{RC}, \zeta \ni (z, z')} wt(\zeta)$$

$$\leq \frac{L(\Gamma_{RC})}{\pi_{RC}(z) P_{RC}(z, z')} \cdot \frac{1 - 2p}{1 - p} \cdot 2n^4 \pi_{RC}(z)$$

$$\leq 2mn^4 \cdot \frac{1 - 2p}{1 - p} \cdot \frac{2m}{\min\{1, \frac{1-2p}{2p}\}}$$

$$\leq 8m^2 n^4,$$

where we use Lemma 14 in the first inequality and $p \leq 1/2$ in the last.

For any transition (z, z') where $z' = z$, since the chain is lazy, $P_{RC}(z, z') \geq 1/2$ and

$$\frac{L(\Gamma_{RC})}{\pi_{RC}(z) P_{RC}(z, z')} \sum_{\zeta \in \Gamma_{RC}, \zeta \ni (z, z')} wt(\zeta) \leq \frac{L(\Gamma_{RC})}{\pi_{RC}(z) P_{RC}(z, z')} \cdot 2mn^4 \pi_{RC}(z)$$

$$\leq 4m^2 n^4,$$

where we use Lemma 14 in the first line. $\quad \square$

APPENDIX A: EQUIVALENCE OF THE THREE MODELS

The equivalence between the Ising model and the random cluster model with $q = 2$ can be found, for example, in [Grimmett (2006)]. An alternative explanation is as follows. In the Ising model, instead of assigning vertices 0 or 1, we assign "equal" or "independent" to edges. Each "equal" edge has a weight of $\beta - 1$, and the "independent" edge has weight 1. This does not change the partition function of the Ising model, since for each edge, if the two endpoints are equal, the weight

is $\beta - 1 + 1 = \beta$, whereas if the two endpoints are different, the weight is 1. For a subset $S \subseteq E$ of edges assigned "equal", each component of S has two possible assignments. Therefore, the weight of S is $(\beta - 1)^{|S|} 2^{\kappa(S)}$. After rescaling by $\beta^{|E|}$, this matches the random cluster formulation (2) with $p = 1 - \frac{1}{\beta}$ and $q = 2$. This gives the first equality of (10).

The equivalence between the Ising model and even subgraphs model can be explained via a holographic transformation by the Hadamard matrix $H = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$.[3] This view will be useful in the next section. In the Ising model, vertices have functions EQUALITY on their adjacent d many half-edges, which after the transformation becomes EVEN function, defined as follows:

$$\text{EVEN}(x_1, \ldots, x_d) = \begin{cases} 2 & \text{if } \bigoplus_i x_i = 0; \\ 0 & \text{otherwise.} \end{cases}$$

On the edges, the function (on the two half-edges) is

$$\text{ISING}(x_1, x_2) = \begin{cases} \beta & \text{if } x_1 = x_2; \\ 1 & \text{otherwise,} \end{cases}$$

whereas after the transformation it is a weighted equality function:

$$\text{WEQ}(x_1, x_2) = \begin{cases} \dfrac{\beta + 1}{2} & \text{if } x_1 = x_2 = 0; \\ \dfrac{\beta - 1}{2} & \text{if } x_1 = x_2 = 1; \\ 0 & \text{otherwise.} \end{cases}$$

Therefore, for a subset S of edges (both half-edges are 1), its weight is

$$wt(S) = \begin{cases} 2^{|V|} \left(\dfrac{\beta - 1}{2} \right)^{|S|} \left(\dfrac{\beta + 1}{2} \right)^{|E \setminus S|} & \text{if } S \in \Omega_{\text{even}}; \\ 0 & \text{otherwise.} \end{cases}$$

The requirement of $S \in \Omega_{\text{even}}$ arises because each vertex requires even degree, and when all degree constraints are satisfied, the vertices contribute $2^{|V|}$ in total. We may rewrite the weight of $S \in \Omega_{\text{even}}$:

$$2^{|V|} \left(\frac{\beta - 1}{2} \right)^{|S|} \left(\frac{\beta + 1}{2} \right)^{|E \setminus S|} = 2^{|V|} \beta^{|E|} \left(\frac{1}{2} \left(1 - \frac{1}{\beta} \right) \right)^{|S|} \left(\frac{1}{2} \left(1 + \frac{1}{\beta} \right) \right)^{|E \setminus S|}.$$

Hence setting $p = \frac{1}{2}(1 - \frac{1}{\beta})$ matches (9) and taking out an appropriate scaling factor yields the second equality of (10).

---

[3] For a treatment of holographic transformations, see, for example, [Cai, Guo and Williams (2016)].

APPENDIX B: CONGESTION OF THE WORM PROCESS

Throughout this section, fix $p \leq 1/2$. Recall that Ω_k is the collection of subgraphs where k many vertices have odd degrees. Then $\Omega_0 = \Omega_{\text{even}}$, and $\Omega_0 \cup \Omega_2 = \Omega_{\text{worm}}$. Define

$$Z_k := \sum_{S \in \Omega_k} w_p(S),$$

where $w_p(S) = p^{|S|}(1-p)^{|E \setminus S|}$. Then $Z_0 = Z_{\text{even}}(p)$ and $Z_{\text{worm}}(p) = Z_0 + n^{-2}Z_2$.

LEMMA 15. $Z_2 \leq \binom{n}{2} Z_0$.

PROOF.   We adopt the holographic transformation view of the even subgraphs model. A vertex that only allows odd degrees is equivalent to the following function:

$$\text{ODD}(x_1, \ldots, x_d) = \begin{cases} 2 & \text{if } \bigoplus_i x_i = 1; \\ 0 & \text{otherwise.} \end{cases}$$

Transforming back to the Ising model, this vertex is still an EQUALITY on all adjacent half-edges, but with a weight of -1 when all half-edges are assigned 1. Hence for every $u, v \in V$,

$$(15) \qquad\qquad Z_{u,v} := \sum_{S \in \Omega(u,v)} w_p(S) \leq Z_0,$$

because the left-hand side can be transformed to the original Ising model with u and v having weights -1. Summing over all possible pairs of vertices in (15) yields $Z_2 \leq \binom{n}{2} Z_0$.   □

In particular, Lemma 15 implies that $Z_{\text{worm}} = Z_0 + n^{-2}Z_2 \leq Z_0 + n^{-2}\binom{n}{2}Z_0 \leq 3Z_0/2$. Now we are ready to prove Theorem 11.

PROOF OF THEOREM 11.   Let I and F be two configurations in Ω_0, denoting the initial and final states. Then $I \oplus F \in \Omega_0$. The canonical path from I to F will be identical to those in [Jerrum and Sinclair (1993)]. Fix an arbitrary ordering of all cycles in G. For each cycle, we designate a starting vertex and a direction around the cycle. Hence each cycle is an ordered tuple of edges. Since $I \oplus F$ is an even subgraph, we can cover $I \oplus F$ by a collection of edge-disjoint cycles. Let $\{C_1, \ldots, C_r\}$ be the first such in our ordering. Let $e_1, \ldots, e_k$ be the edges of $\{C_1, \ldots, C_r\}$ taken in order (first order the edges according to the cycle they occur in, and then by their position within the cycle, counting from the start vertex). The canonical path γ from I to F is defined to be $Z_0 = I$, $Z_i = Z_{i-1} \oplus e_i$, and $Z_k = F$. Intuitively, the canonical path unwinds C_i one by one from $i = 1$ to $i = r$. Clearly, $L = L(\Gamma_{\text{worm}}) \leq m$ as it can use every edge at most once.

This path is always in $\Omega_0 \cup \Omega_2$ because if we start to unwind a cycle, then the current state is an even subgraph. If we are unwinding a path, then we always flip an edge that is adjacent to an odd degree vertex.

For any transition (w, w') where $w' = w \oplus e$ for some edge $e \in E$, we use a combinatorial encoding as in [Jerrum and Sinclair (1993)] for all paths passing through (w, w'). For any two configurations $I, F \in \Omega_0$, let $\varphi(I, F) = I \oplus F \oplus w$. We claim that $\varphi : \Omega_0^2 \to \Omega_0 \cup \Omega_2$ is an injection. This is because given (w, w') and $U = \varphi(I, F)$, we can recover the unique (I, F). First, since $w \oplus U = I \oplus F$, all edges *not* in $w \oplus U$ have the same state in both I and F, and their states are the same as those in w. Then for edges in $w \oplus U$, due to the construction of the canonical path, there is a unique ordering among those edges, including $e = w \oplus w'$. For any edge before e, its status in w has been changed to that in F, and its status in U is still the same as that in I. For any edge after e (including e itself), its status in w is still the same as that in I, and in U is the same as in F.

Recall that $w_p(S) = p^{|S|}(1 - p)^{|E \setminus S|}$ for any subset of edges $S \subset E$. Since $I \oplus F = w \oplus U$ and $I \cap F = w \cap U$, we have that

$$w_p(I)w_p(F) = w_p(w)w_p(U).$$

Therefore,

$$
\sum_{\gamma \ni (w,w')} wt(\gamma) = \sum_{\substack{I,F \in \Omega_0^2 \\ \gamma_{IF} \ni (w,w')}} \pi_{\text{even}}(I)\pi_{\text{even}}(F) = \sum_{\substack{I,F \in \Omega_0^2 \\ \gamma_{IF} \ni (w,w')}} \frac{w_p(I)w_p(F)}{Z_0^2}
$$

$$
= \sum_{\substack{I,F \in \Omega_0^2 \\ \gamma_{IF} \ni (w,w')}} \frac{w_p(w)w_p(\varphi(I, F))}{Z_0^2}
$$

$$
\leq w_p(w) \sum_{U \in \Omega_0 \cup \Omega_2} \frac{w_p(U)}{Z_0^2}
$$

$$
= \frac{Z_0 + Z_2}{Z_0^2} \cdot w_p(w).
$$

By the definition of π_{worm} (12) and (13), $\pi_{\text{worm}}(w) = \frac{w_{\text{worm}}(w)}{Z_{\text{worm}}} \geq \frac{w_p(w)}{n^2 Z_{\text{worm}}}$. This implies that

$$
\sum_{\gamma \ni (w,w')} wt(\gamma) \leq \frac{Z_0 + Z_2}{Z_0} \cdot \frac{Z_{\text{worm}}}{Z_0} \cdot n^2 \pi_{\text{worm}}(w)
$$

(by Lemma 15)
$$
\leq \left(1 + \binom{n}{2}\right)\left(1 + \binom{n}{2}n^{-2}\right)n^2 \pi_{\text{worm}}(w)
$$

$$
\leq n^4 \pi_{\text{worm}}(w).
$$

For the last claim of the theorem, let $w' = w \cup \{e\}$ for some $e \notin w$. We can do the same combinatorial encoding for w'. That is, let $U' = \varphi'(I, F) = I \oplus F \oplus w'$. It is easy to verify as above that φ' is an injection. Then as above,

$$
\sum_{\gamma \ni (w, w')} wt(\gamma) \le \frac{Z_0 + Z_2}{Z_0^2} \cdot w_p(w')
$$

$$
= \frac{Z_0 + Z_2}{Z_0^2} \cdot w_p(w) \cdot \frac{p}{1 - p}
$$

$$
\le n^4 \pi_{\mathrm{worm}}(w) \frac{p}{1 - p}. \qquad \square
$$

REFERENCES

BLANCA, A. and SINCLAIR, A. (2015). Dynamics for the mean-field random-cluster model. In *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques. LIPIcs. Leibniz Int. Proc. Inform.* **40** 528–543. Schloss Dagstuhl. Leibniz-Zent. Inform., Wadern. MR3441983

BLANCA, A. and SINCLAIR, A. (2016). Random-cluster dynamics in $\mathbb{Z}^2$. *Probab. Theory Relat. Fields.* DOI: 10.1007/s00440-016-0725-1.

BOLLOBÁS, B., GRIMMETT, G. and JANSON, S. (1996). The random-cluster model on the complete graph. *Probab. Theory Related Fields* **104** 283–317. MR1376340

BORGS, C., CHAYES, J. T., FRIEZE, A., KIM, J. H., TETALI, P., VIGODA, E. and VU, V. H. (1999). Torpid mixing of some Monte Carlo Markov chain algorithms in statistical physics. In 40*th Annual Symposium on Foundations of Computer Science* (*New York*, 1999) 218–229. IEEE Computer Soc., Los Alamitos, CA. MR1917562

CAI, J.-Y., GUO, H. and WILLIAMS, T. (2016). A complete dichotomy rises from the capture of vanishing signatures. *SIAM J. Comput.* **45** 1671–1728. MR3543175

COLLEVECCHIO, A., GARONI, T. M., HYNDMAN, T. and TOKAREV, D. (2016). The worm process for the Ising model is rapidly mixing. *J. Stat. Phys.* **164** 1082–1102. MR3534485

DIACONIS, P. and SALOFF-COSTE, L. (1993). Comparison theorems for reversible Markov chains. *Ann. Appl. Probab.* **3** 696–730. MR1233621

DIACONIS, P. and STROOCK, D. (1991). Geometric bounds for eigenvalues of Markov chains. *Ann. Appl. Probab.* **1** 36–61. MR1097463

DUMINIL-COPIN, H., GAGNEBIN, M., HAREL, M., MANOLESCU, I. and TASSION, V. (2016). Discontinuity of the phase transition for the planar random-cluster and Potts models with $q > 4$. Available at arXiv:abs/1611.09877.

DYER, M., GOLDBERG, L. A., JERRUM, M. and MARTIN, R. (2006). Markov chain comparison. *Probab. Surv.* **3** 89–111. MR2216963

EDWARDS, R. G. and SOKAL, A. D. (1988). Generalization of the Fortuin–Kasteleyn–Swendsen–Wang representation and Monte Carlo algorithm. *Phys. Rev. D* **38** 2009–2012. MR0965465

FORTUIN, C. M. and KASTELEYN, P. W. (1972). On the random-cluster model. I. Introduction and relation to other models. *Physica* **57** 536–564. MR0359655

GALANIS, A., ŠTEFANKOVIČ, D. and VIGODA, E. (2015). Swendsen–Wang algorithm on the mean-field Potts model. In *Approximation, Randomization, and Combinatorial Optimization.*

Algorithms and Techniques. LIPIcs. Leibniz Int. Proc. Inform. **40** 815–828. Schloss Dagstuhl. Leibniz-Zent. Inform., Wadern. MR3441999

GHEISSARI, R. and LUBETZKY, E. (2016). Mixing times of critical 2D Potts models. Available at arXiv:abs/1607.02182.

GOLDBERG, L. A. and JERRUM, M. (2008). Inapproximability of the Tutte polynomial. *Inform. and Comput.* **206** 908–929. MR2433765

GOLDBERG, L. A. and JERRUM, M. (2012). Approximating the partition function of the ferromagnetic Potts model. *J. ACM* **59** Art. 25, 31. MR2995824

GOLDBERG, L. A. and JERRUM, M. (2014). The complexity of computing the sign of the Tutte polynomial. *SIAM J. Comput.* **43** 1921–1952. MR3291539

GORE, V. K. and JERRUM, M. R. (1999). The Swendsen–Wang process does not always mix rapidly. *J. Stat. Phys.* **97** 67–86. MR1733467

GRIMMETT, G. (2006). *The Random-Cluster Model. Grundlehren der Mathematischen Wissenschaften* [*Fundamental Principles of Mathematical Sciences*] **333**. Springer, Berlin. MR2243761

GRIMMETT, G. and JANSON, S. (2009). Random even graphs. *Electron. J. Combin.* **16** Research Paper, 46, 19. MR2491648

JERRUM, M. and SINCLAIR, A. (1989). Approximating the permanent. *SIAM J. Comput.* **18** 1149–1178. MR1025467

JERRUM, M. and SINCLAIR, A. (1993). Polynomial-time approximation algorithms for the Ising model. *SIAM J. Comput.* **22** 1087–1116. MR1237164

LONG, Y., NACHMIAS, A., NING, W. and PERES, Y. (2014). A power law of order 1/4 for critical mean field Swendsen–Wang dynamics. *Mem. Am. Math. Soc.* **232**.

LUBETZKY, E. and SLY, A. (2012). Critical Ising on the square lattice mixes in polynomial time. *Comm. Math. Phys.* **313** 815–836.

PERES, Y. (2017). Personal communication.

PROKOF'EV, N. and SVISTUNOV, B. (2001). Worm algorithms for classical statistical models. *Phys. Rev. Lett.* **87** 160601.

RANDALL, D. and WILSON, D. (1999). Sampling spin configurations of an Ising system. In *SODA* 959–960.

SINCLAIR, A. (1992). Improved bounds for mixing rates of Markov chains and multicommodity flow. *Combin. Probab. Comput.* **1** 351–370.

SINCLAIR, A. and JERRUM, M. (1989). Approximate counting, uniform generation and rapidly mixing Markov chains. *Inform. and Comput.* **82** 93–133. MR1003059

SWENDSEN, R. and WANG, J.-S. (1987). Nonuniversal critical dynamics in Monte Carlo simulations. *Phys. Rev. Lett.* **58** 86–88.

ULLRICH, M. (2013). Comparison of Swendsen–Wang and heat-bath dynamics. *Random Structures Algorithms* **42** 520–535.

ULLRICH, M. (2014a). Rapid mixing of Swendsen–Wang dynamics in two dimensions. *Dissertationes Math.* (*Rozprawy Mat.*) **502**.

ULLRICH, M. (2014b). Swendsen–Wang is faster than single-bond dynamics. *SIAM J. Discrete Math.* **28** 37–48.

VAN DER WAERDEN, B. L. (1941). Die lange Reichweite der regelmäßigen Atomanordnung in Mischkristallen. *Z. Phys.* **118** 473–488.

SCHOOL OF INFORMATICS
UNIVERSITY OF EDINBURGH
INFORMATICS FORUM
EDINBURGH EH8 9AB
UNITED KINGDOM
E-MAIL: hguo@inf.ed.ac.uk

SCHOOL OF MATHEMATICAL SCIENCES
QUEEN MARY, UNIVERSITY OF LONDON
MILE END ROAD
LONDON E1 4NS
UNITED KINGDOM
E-MAIL: m.jerrum@qmul.ac.uk