EURANDOM PREPRINT SERIES

2019-016

December 30, 2019

Glauber Dynamics on the Erdős-Rényi Random Graph

F. den Hollander, O. Jovanovski
ISSN 1389-2355

GLAUBER DYNAMICS ON THE ERDŐS-RÉNYI RANDOM GRAPH

F. DEN HOLLANDER AND O. JOVANOVSKI

Abstract. We investigate the effect of disorder on the Curie-Weiss model with Glauber dynamics. In particular, we study metastability for spin-flip dynamics on the Erdős-Rényi random graph $\mathrm{ER}_n(p)$ with n vertices and with edge retention probability $p \in (0,1)$. Each vertex carries an Ising spin that can take the values -1 or $+1$. Single spins interact with an external magnetic field $h \in (0,\infty)$, while pairs of spins at vertices connected by an edge interact with each other with ferromagnetic interaction strength $1/n$. Spins flip according to a Metropolis dynamics at inverse temperature β. The standard Curie-Weiss model corresponds to the case $p = 1$, because $\mathrm{ER}_n(1) = K_n$ is the complete graph on n vertices. For $\beta > \beta_c$ and $h \in (0, p\chi(\beta p))$ the system exhibits *metastable behaviour* in the limit as $n \to \infty$, where $\beta_c = 1/p$ is the *critical inverse temperature* and χ is a certain *threshold function* satisfying $\lim_{\lambda\to\infty} \chi(\lambda) = 1$ and $\lim_{\lambda\downarrow 1} \chi(\lambda) = 0$. We compute the average crossover time from the *metastable set* (with magnetization corresponding to the 'minus-phase') to the *stable set* (with magnetization corresponding to the 'plus-phase'). We show that the average crossover time grows exponentially fast with n, with an exponent that is the same as for the Curie-Weiss model with external magnetic field h and with ferromagnetic interaction strength p/n. We show that the correction term to the exponential asymptotics is a multiplicative error term that is *at most polynomial* in n. For the complete graph K_n the correction term is known to be a multiplicative constant. Thus, apparently, $\mathrm{ER}_n(p)$ is so homogeneous for large n that the effect of the fluctuations in the disorder is small, in the sense that the metastable behaviour is controlled by the average of the disorder. Our model is the first example of a metastable dynamics on a random graph where the correction term is estimated to high precision.

Contents

Date: December 23, 2019.
2010 *Mathematics Subject Classification.* 60C05; 60K35; 60K37; 82C27.
Key words and phrases. Erdős-Rényi random graph, Glauber spin-flip dynamics, metastability, crossover time.
The research in this paper was supported by NWO Gravitation Grant 024.002.003-NETWORKS.

1. Introduction and main results

In Section 1.1 we provide some background on metastability. In Section 1.2 we define our model: spin-flip dynamics on the Erdős-Rényi random graph $\mathrm{ER}_n(p)$. In Section 1.3 we identify the metastable pair for the dynamics, corresponding to the 'minus-phase' and the 'plus-phase', respectively. In Section 1.4 we recall the definition of spin-flip dynamics on the complete graph K_n, which serves as a comparison object, and recall what is known about the average metastable crossover time for spin-flip dynamics on K_n (Theorem 1.3 below). In Section 1.5 we transfer the sharp asymptotics for K_n to a somewhat rougher asymptotics for $\mathrm{ER}_n(p)$ (Theorem 1.4 below). In Section 1.6 we close by placing our results in the proper context and giving an outline of the rest of the paper.

1.1. **Background.** Interacting particle systems, evolving according to a *Metropolis dynamics* associated with an energy functional called the *Hamiltonian*, may end up being trapped for a long time near a state that is a local minimum but not a global minimum. The deepest local minima are called *metastable states*, the global minimum is called the *stable state*. The transition from a metastable state to the stable state marks the relaxation of the system to equilibrium. To describe this relaxation, it is of interest to compute the crossover time and to identify the set of critical configurations the system has to visit in order to achieve the transition. The critical configurations represent the saddle points in the energy landscape: the set of mini-max configurations that must be hit by any path that achieves the crossover.

Metastability for interacting particle systems on *lattices* has been studied intensively in the past three decades. Various different approaches have been proposed, which are summarised in the monographs by Olivieri and Vares [13], Bovier and den Hollander [4]. Recently, there has been interest in metastability for interacting particle systems on *random graphs*, which is much more challenging because the crossover time typically depends in a delicate manner on the realisation of the graph.

In the present paper we are interested in metastability for spin-flip dynamics on the *Erdős-Rényi random graph*. Our main result is an estimate of the average crossover time from the 'minus-phase' to the 'plus-phase' when the spins feel an external magnetic field at the vertices in the graph as well as a ferromagnetic interaction along the edges in the graph. Our paper is part of a larger enterprise in which the goal is to understand metastability on graphs. Jovanovski [12] analysed the case of the *hypercube*, Dommers [6] the case of the *random regular graph*, Dommers, den Hollander, Jovanovski and Nardi [9] the case of the *configuration model*, and den Hollander and Jovanovski [11] the case of the *hierarchical lattice*. Each case requires carrying out a detailed combinatorial analysis that is model-specific, even though the metastable behaviour is ultimately universal. For lattices like the hypercube and the hierarchical lattice a full identification of the relevant quantities is possible, while for random graphs like the random regular graph and the configuration model so far only the communication height is well understood, while the set of critical configurations and the prefactor remain somewhat elusive.

The equilibrium behaviour of the Ising model on random graphs is well understood. See e.g. Dommers, Giardinà and van der Hofstad [7], [8].

1.2. **Spin-flip dynamics on** $\mathrm{ER}_n(p)$**.** Let $\mathrm{ER}_n(p) = (V, E)$ be a realisation of the Erdős-Rényi random graph on $|V| = n \in \mathbb{N}$ vertices with edge retention probability $p \in (0, 1)$, i.e., each edge in the complete graph K_n is present with probability p and absent with probability $1-p$, independently of other edges (see Fig. 1). We write $\mathbb{P}_{\mathrm{ER}_n(p)}$ to denote the law of $\mathrm{ER}_n(p)$. For typical properties of $\mathrm{ER}_n(p)$, see van der Hofstad [10, Chapters 4–5].

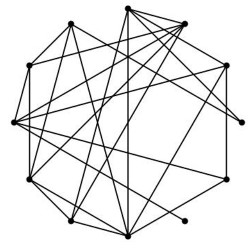

FIGURE 1. A realization of the Erdős-Rényi random graph with $n = 12$ and $p = \frac{1}{3}$.

Each vertex carries an Ising spin that can take the values -1 or $+1$. Let $S_n = \{-1,+1\}^V$ denote the set of spin configurations on V, and let H_n be the *Hamiltonian* on S_n defined by

$$H_n(\sigma) = -\frac{1}{n}\sum_{(v,w)\in E}\sigma(v)\sigma(w) - h\sum_{v\in V}\sigma(v), \quad \sigma\in S_n. \tag{1.1}$$

In other words, single spins interact with an *external magnetic field* $h \in (0,\infty)$, while pairs of neighbouring spins interact with each other with a *ferromagnetic coupling strength* $1/n$.

Let $\boxminus = \{-1\}^V$ and $\boxplus = \{+1\}^V$ denote the configuration where all spins are -1, respectively, $+1$. Since

$$H_n(\boxminus) = -\frac{|E|}{n} + hn, \tag{1.2}$$

we have the geometric representation

$$H_n(\sigma) = H_n(\boxminus) + \frac{2}{n}|\partial_E\sigma| - 2h|\sigma|, \quad \sigma\in S_n, \tag{1.3}$$

where

$$\partial_E\sigma = \{(v,w)\in E\colon \sigma(v) = -\sigma(w) = +1\} \tag{1.4}$$

is the *edge-boundary* of σ and

$$|\sigma| = \{v\in \mathrm{ER}_n(p)\colon\ \sigma(v) = +1\} \tag{1.5}$$

is the *vertex-volume* of σ.

In the present paper we consider a spin-flip dynamics on S_n commonly referred to as *Glauber dynamics*, defined as the continuous-time Markov process with transition rates

$$r(\sigma,\sigma) = \begin{cases} e^{-\beta[H_n(\sigma')-H_n(\sigma)]_+}, & \text{if } \|\sigma-\sigma'\| = 2,\\ 0, & \text{if } \|\sigma-\sigma'\| > 2,\end{cases} \quad \sigma,\sigma'\in S_n, \tag{1.6}$$

where $\|\cdot\|$ is the ℓ_1-norm on S_n. This dynamics has as *reversible* stationary distribution the Gibbs measure

$$\mu_n(\sigma) = \frac{1}{Z_n}e^{-\beta H_n(\sigma)}, \quad \sigma\in S_n, \tag{1.7}$$

where $\beta\in(0,\infty)$ is the *inverse temperature* and Z_n is the normalizing partition sum. We write

$$\{\xi_t\}_{t\geq 0} \tag{1.8}$$

to denote the path of the random dynamics and $\mathbb{P}_\xi$ to denote its law given $\xi_0 = \xi$. For $\chi\subset S_n$, we write

$$\tau_\chi = \inf\{t\geq 0\colon\ \xi_t\in\chi,\, \xi_{t-}\in\chi\}. \tag{1.9}$$

to denote the *first hitting/return time* of χ.

We define the *magnetization* of σ by

$$m(\sigma) = \frac{1}{n}\sum_{v\in V}\sigma(v), \tag{1.10}$$

and observe the relation

$$m(\sigma) = \frac{2|\sigma|}{n} - 1, \quad \sigma \in S_n. \tag{1.11}$$

We will frequently switch between working with volume and working with magnetization. Equation (1.11) ensures that these are in one-to-one correspondence. Accordingly, we will frequently look at the dynamics from the perspective of the induced *volume process* and *magnetization process*,

$$\{|\xi_t|\}_{t\geq 0}, \quad \{m(\xi_t)\}_{t\geq 0}, \tag{1.12}$$

which are *not* Markov.

1.3. **Metastable pair.** For fixed n, the Hamiltonian in (1.1) achieves a *global minimum* at $\boxplus$ and a *local minimum* at $\boxminus$. In fact, $\boxminus$ is the deepest local minimum not equal to $\boxplus$ (at least for h small enough). However, in the limit as $n \to \infty$, these do *not* form a metastable pair of configurations because *entropy* comes into play.

Definition 1.1 **(Metastable regime).** The parameters β, h are said to be in the *metastable regime* when

$$\beta \in (1/p, \infty), \qquad h \in \big(0, p\chi(\beta p)\big) \tag{1.13}$$

with (see Fig. 2)

$$\chi(\lambda) = \sqrt{1-\tfrac{1}{\lambda}} - \tfrac{1}{2\lambda}\log\left[\lambda\left(1+\sqrt{1-\tfrac{1}{\lambda}}\right)^2\right], \qquad \lambda \geq 1. \tag{1.14}$$

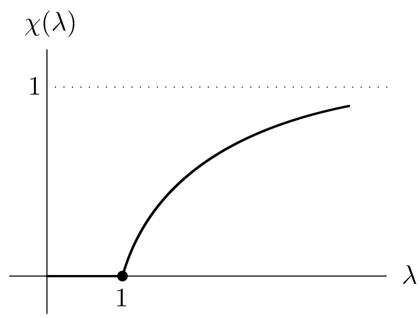

FIGURE 2. Plot of $\lambda \mapsto \chi(\lambda)$.

We have $\lim_{\lambda\to\infty}\chi(\lambda) = 1$ and $\lim_{\lambda\downarrow 1}\chi(\lambda) = 0$ (with slope $\frac{1}{2}$). Hence, for $\beta \to \infty$ any $h \in (0,p)$ is metastable, while for $\beta \downarrow 1/p$ or $p \downarrow 0$ no $h \in (0,\infty)$ is metastable. The latter explains why we do not consider the non-dense Erdős-Rényi random graph with $p = p_n \downarrow 0$ as $n \to \infty$. ■

The threshold $\beta_c = 1/p$ is the *critical temperature*: the static model has a phase transition at $h = 0$ when $\beta > \beta_c$ and no phase transition when $\beta \leq \beta_c$ (see e.g. Dommers, Giardinà and van der Hofstad [8]).

To define the proper metastable pair of configurations, we need the following definitions. Let

$$\Gamma_n = \{-1, -1+\tfrac{2}{n}, \ldots, 1-\tfrac{2}{n}, 1\}, \quad I_n(a) = -\frac{1}{n}\log\binom{n}{\frac{1+a}{2}n}, \quad J_n(a) = 2\beta(pa+h) - 2I_n'(a). \tag{1.15}$$

Define

$$\begin{aligned} \mathbf{m}_n &= \min\{a \in \Gamma_n : \ J_n(a) \leq 0\}, \\ \mathbf{t}_n &= \min\{a \in \Gamma_n : \ a > \mathbf{m}_n, \, J_n(a) \geq 0\}, \\ \mathbf{s}_n &= \min\{a \in \Gamma_n : \ a > \mathbf{t}_n, \, J_n(a) \leq 0\}. \end{aligned} \tag{1.16}$$

The numbers in the left-hand side of (1.16) play the role of magnetizations. Further define

$$\mathbf{M}_n = \frac{n}{2}(\mathbf{m}_n + 1), \quad \mathbf{T}_n = \frac{n}{2}(\mathbf{t}_n + 1), \quad \mathbf{S}_n = \frac{n}{2}(\mathbf{s}_n + 1), \tag{1.17}$$

which are the volumes corresponding to (1.16), and

$$A_k = \{\sigma \in S_n\colon\ |\sigma| = k\}, \qquad k \in \{0,1,\dots,n-1,n\}, \tag{1.18}$$

the set of configurations with volume k. Define

$$R_n(a) = -\frac{1}{2}pa^2 - ha + \frac{1}{\beta}I_n(a) \tag{1.19}$$

and note that

$$R_n'(a) = -pa - h + \frac{1}{\beta}I_n'(a) = -\frac{1}{2\beta}J_n(a). \tag{1.20}$$

The motivation behind the definitions in (1.15), (1.16) and (1.19) will become clear in Section 2. Via Stirling's formula it follows that

$$J_n(a) = 2\beta(pa+h) + \log\left(\frac{1-a+\frac{1}{n}}{1+a+\frac{1}{n}}\right) + O(n^{-2}), \quad a \in \Gamma_n. \tag{1.21}$$

We will see that, in the limit as $n \to \infty$ when (β, h) is in the metastable regime defined by (1.13), the numbers in (1.16) are well-defined: $A_{\mathbf{M}_n}$ is the *metastable set*, $A_{\mathbf{S}_n}$ is the *stable set*, $A_{\mathbf{T}_n}$ is the *top set*, i.e., the set of saddle points that lie in between $A_{\mathbf{M}_n}$ and $A_{\mathbf{S}_n}$. Our key object of interest will be the *crossover time* from $A_{\mathbf{M}_n}$ to $A_{\mathbf{S}_n}$ via $A_{\mathbf{T}_n}$.

Note that

$$\Gamma_n \to [-1,1], \quad I_n(a) \to I(a), \quad J_n(a) \to J_{p,\beta,h}(a), \qquad n \to \infty, \tag{1.22}$$

with

$$J_{p,\beta,h}(a) = 2\beta(pa+h) + \log\left(\frac{1-a}{1+a}\right) \tag{1.23}$$

and

$$I(a) = \frac{1-a}{2}\log\left(\frac{1-a}{2}\right) + \frac{1+a}{2}\log\left(\frac{1+a}{2}\right). \tag{1.24}$$

Accordingly,

$$\mathbf{m}_n \to \mathbf{m}, \quad \mathbf{t}_n \to \mathbf{t}, \quad \mathbf{s}_n \to \mathbf{s}, \qquad n \to \infty, \tag{1.25}$$

with $\mathbf{m}$, $\mathbf{t}$, $\mathbf{s}$ the three successive zeroes of J (see Fig. 4 and recall (1.16)). Define

$$R_{p,\beta,h}(a) = -\frac{1}{2}pa^2 - ha + \frac{1}{\beta}I(a) \tag{1.26}$$

Note that $R_{p,\beta,h}(a)$ plays the role of *free energy*: $-\frac{1}{2}pa^2 - ha$ and $\frac{1}{\beta}I(a)$ represent the energy, respectively, entropy at magnetisation a. Note that $I(a)$ equals the relative entropy of the probability measure $\frac{1}{2}(1+a)\delta_{+1} + \frac{1}{2}(1-a)\delta_{-1}$ with respect to the counting measure $\delta_{+1}+\delta_{-1}$. Also note that

$$R_{p,\beta,h}'(a) = -pa - h + \frac{1}{\beta}I'(a) = -\frac{1}{2\beta}J_{p,\beta,h}(a). \tag{1.27}$$

Remark 1.2. As shown in Corollary 3.6 below, if $h \in (p,\infty)$, then (1.6) leads to *non-metastable* behaviour where the dynamics 'drifts' through a sequence of configurations with volume growing from $\mathbf{M}$ to $\mathbf{S}$ within time $O(1)$. ∎

1.4. **Metastability on K_n.** Let K_n be the complete graph on n vertices (see Fig. 3). Spin-flip dynamics on K_n, commonly referred to as *Glauber dynamics for the Curie-Weiss model*, is defined as in Section 1.2 but with the *Curie-Weiss Hamiltonian*

$$H_n(\sigma) = -\frac{1}{2n}\sum_{1\le i,j\le n}\sigma(i)\sigma(j) - h\sum_{1\le i\le n}\sigma(i), \quad \sigma \in S_n. \tag{1.28}$$

This is the special case of (1.1) when $p=1$, except for the diagonal term $-\frac{1}{2n}\sum_{1\le i\le n}\sigma(i)\sigma(i) = -\frac{1}{2}$, which shifts H_n by a constant and has no effect on the dynamics. The advantage of (1.28) is that we may write

$$H_n(\sigma) = n\big[-\tfrac{1}{2}m(\sigma)^2 - hm(\sigma)\big], \tag{1.29}$$

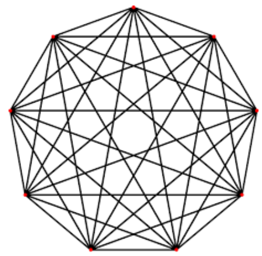

FIGURE 3. The complete graph with $n = 9$.

which shows that the energy is a function of the magnetization only, i.e., the Curie-Weiss model is a *mean-field* model. Clearly, this property fails on $\mathrm{ER}_n(p)$.

For the Curie-Weiss model it is known that there is a *critical inverse temperature* $\beta_c = 1$ such that, for $\beta > \beta_c$, h small enough and in the limit as $n \to \infty$, the stationary distribution μ_n given by (1.7) and (1.28) has two phases: the 'minus-phase', where the majority of the spins are -1, and the 'plus-phase', where the majority of the spins are $+1$. These two phases are the *metastable state*, respectively, the *stable state* for the dynamics. In the limit as $n \to \infty$, the dynamics of the magnetization introduced in (1.12) (which is Markov) converges to a Brownian motion on $[-1, +1]$ in the *double-well potential* $a \mapsto R_{1,\beta,h}(a)$ (see Fig. 4).

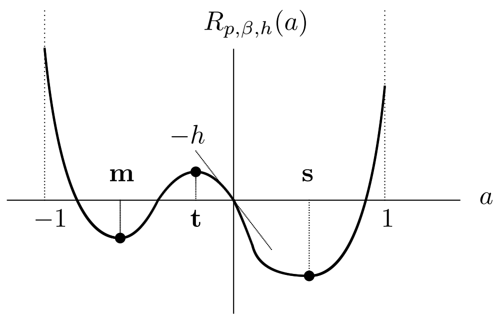

FIGURE 4. Plot of $R_{p,\beta,h}(a)$ as a function of the magnetization a. The metastable set $A_{\mathbf{M}}$ has magnetization $\mathbf{m} < 0$, the stable set $A_{\mathbf{S}}$ has magnetization $\mathbf{s} > 0$, the top set has magnetization $\mathbf{t} < 0$. Note that $R_{p,\beta,h}(-1) = -\frac{1}{2}p+h$, $R_{p,\beta,h}(0) = -\beta^{-1}\log 2$, $R_{p,\beta,h}(+1) = -\frac{1}{2}p - h$ and $R'_{p,\beta,h}(-1) = -\infty$, $R'_{p,\beta,h}(0) = -h$, $R'_{p,\beta,h}(+1) = \infty$.

The following theorem can be found in Bovier and den Hollander [4, Chapter 13]. For $p = 1$, the metastable regime in (1.13) becomes

$$\beta \in (1, \infty), \qquad h \in \big(0, \chi(\beta)\big). \tag{1.30}$$

Theorem 1.3 (**Average crossover time on K_n**). *Subject to* (1.30)*, as $n \to \infty$, uniformly in $\xi \in A_{\mathbf{M}_n}$,*

$$\mathbb{E}_\xi\left[\tau_{A_{\mathbf{S}_n}}\right] = [1 + o_n(1)]\,\frac{\pi}{1-\mathbf{t}}\sqrt{\frac{1-\mathbf{t}^2}{1-\mathbf{m}^2}}\,\frac{1}{\beta\sqrt{R''_{1,\beta,h}(\mathbf{m})[-R''_{1,\beta,h}(\mathbf{t})]}}\,e^{\beta n[R_{1,\beta,h}(\mathbf{t})-R_{1,\beta,h}(\mathbf{m})]}. \tag{1.31}$$

Fig. 4 illustrates the setting: the average crossover time from $A_{\mathbf{M}_n}$ to $A_{\mathbf{S}_n}$ depends on the energy barrier $R_{1,\beta,h}(\mathbf{t}) - R_{1,\beta,h}(\mathbf{m})$ and on the curvature of $R_{1,\beta,h}$ at $\mathbf{m}$ and $\mathbf{t}$. Note that $\mathbf{m}, \mathbf{s}, \mathbf{t}$ in Fig. 4 are the limits as $n \to \infty$ of $\mathbf{m}_n, \mathbf{s}_n, \mathbf{t}_n$ defined in (1.16) for $p = 1$.

1.5. **Metastability on $\mathrm{ER}_n(p)$.** Unlike for the spin-flip dynamics on K_n, the induced processes defined in (1.12) are *not Markovian*. This is due to the random geometry of $\mathrm{ER}_n(p)$. However, we will see that they are *almost Markovian*, a fact that we will exploit by comparing the dynamics on $\mathrm{ER}_n(p)$ with that on K_n, but with a ferromagnetic coupling strength p/n rather than $1/n$ and with an external magnetic field that is a *small perturbation* of h.

As shown in Lemma 2.4 below, in the metastable regime the function $a \mapsto R_p(a)$ has a double-well structure just like in Fig. 4, so that the metastable state $A_{\mathbf{M}}$ and the stable state $A_{\mathbf{S}}$ are separated by an *energy barrier* represented by $A_{\mathbf{T}}$.

We are finally in a position to state our main theorem. This is stated under a slightly more restrictive condition than (1.13), namely,

$$\frac{h}{p} < \chi^*(\beta p), \qquad \chi^*(\lambda) = 1 - \frac{1}{2\lambda} - \frac{\log(4\lambda - 1)}{2\lambda}. \tag{1.32}$$

Theorem 1.4 (**Average crossover time on** $\mathrm{ER}_n(p)$)**.** *Subject to* (1.32)*, with* $\mathbb{P}_{\mathrm{ER}_n(p)}$*-probability tending to* 1 *as* $n \to \infty$*, uniformly in* $\xi \in A_{\mathbf{M}_n}$*,*

$$\mathbb{E}_\xi \left[\tau_{A_{\mathbf{S}_n}}\right] = n^{E_n}\, e^{\beta n [R_{p,\beta,h}(\mathbf{t}) - R_{p,\beta,h}(\mathbf{m})]} \tag{1.33}$$

where the random exponent E_n *satisfies*

$$\lim_{n\to\infty} \mathbb{P}_{\mathrm{ER}_n(p)}\big(|E_n| \le \beta(\mathbf{t}-\mathbf{m})\tfrac{11}{6}\big) = 1. \tag{1.34}$$

Thus, apart from a polynomial error term, the average crossover time is the same as on the complete graph with ferromagnetic interaction strength p/n instead of $1/n$.

1.6. **Discussion and outline.** We discuss the significance of our main theorem.

1. Theorem 1.4 provides an estimate on the average crossover time from $A_{\mathbf{M}_n}$ to $A_{\mathbf{S}_n}$ on $\mathrm{ER}_n(p)$ (recall Fig. 4). The estimate is *uniform* in the starting configuration. The exponential term in the estimate is the *same* as on K_n, but with a ferromagnetic interaction strength p/n rather than $1/n$. The multiplicative error term is *at most polynomial* in n. Such an error term is *not* present on K_n, for which the prefactor is known to be a constant up to a multiplicative factor $1 + o(1)$ (as shown in Theorem 1.3). The randomness of $\mathrm{ER}_n(p)$ manifests itself through a more complicated prefactor, which we do not know how to identify. What is interesting is that, apparently, $\mathrm{ER}_n(p)$ is so homogeneous for large n that the prefactor is at most polynomial. We expect the prefactor to be *random* under the law $\mathbb{P}_{\mathrm{ER}_n(p)}$.

2. It is known that on K_n the crossover time divided by it average has an exponential distribution in the limit as $n \to \infty$, as is typical for metastable behaviour. The same is true on $\mathrm{ER}_n(p)$. A proof of this fact can be obtained in a straightforward manner from the comparison properties underlying the proof of Theorem 1.4. These comparison properties, which are based on *coupling* of trajectories, also allow us to identify the *typical set of trajectories* followed by the spin-flip dynamics prior to the crossover. We will not spell out the details.

3. The proof of Theorem 1.4 is based on estimates of transition probabilities and transition times between pairs of configurations with different volume, in combination with a *coupling argument.* Thus we are following the *path-wise* approach to metastabuilty (see [4] for background). Careful estimates are needed because on $\mathrm{ER}_n(p)$ the processes introduced in (1.12) are *not* Markovian, unlike on K_n.

4. Bovier, Marello and Pulvirenti [5] use capacity estimates and concentration of measure estimates to show that the prefactors form a *tight* family of random variables under the law $\mathbb{P}_{\mathrm{ER}_n(p)}$ as $n \to \infty$, which constitutes a considerable sharpening of (1.33). The result is valid for $\beta > \beta_c$ and h small enough. The starting configuration is not arbitrary, but is drawn according to the *last-exit-biased distribution* for the transition from $A_{\mathbf{M}_n}$ to $A_{\mathbf{S}_n}$, as is common in the *potential-theoretic* approach to metastability. The exponential limit law is therefore not immediate.

5. Another interesting model is where *the randomness sits in the vertices rather than in the edges*, namely, Glauber spin-flip dynamics with Hamiltonian

$$H_n(\sigma) = -\frac{1}{n} \sum_{1\le i,j\le n} \sigma(i)\sigma(j) - \sum_{1\le i\le n} h_i \sigma(i), \tag{1.35}$$

where h_i, $1 \le i \le n$, are i.i.d. random variables drawn from a common probability distribution ν on $\mathbb{R}$. The metastable behaviour of this model was analysed in Bovier *et al.* [3] (discrete ν) and Bianchi *et al.* [1] (continuous ν). In particular, the prefactor was computed up to a multiplicative factor $1+o(1)$, and turns out to be rather involved (see [4, Chapters 14–15]). Our model is even harder because the interaction between the spins runs along the edges of $\mathrm{ER}_n(p)$, which has an *intricate spatial structure.* Consequently, the so-called *lumping technnique* (employed in [3] and [1] to monitor the magnetization on the level sets of the magnetic field) can no longer be used. For the dynamics under (1.35) the exponential law was proved in Bianchi *et al.* [2].

Outline. The remainder of the paper is organized as follows. In Section 2 we define the perturbed spin-flip dynamics on K_n, in Definition 2.1 below, and state two comparison lemmas, in Lemmas 2.2–2.3 below, that are needed to approximate the spin-flip dynamics on $\mathrm{ER}_n(p)$. We also explain, in Lemma 2.4 below, why Definition 1.1 identifies the metastable regime. In Section 3 we collect a few basic facts about the geometry of $\mathrm{ER}_n(p)$ and the spin-flip dynamics on $\mathrm{ER}_n(p)$. In Section 4 we derive rough capacity estimates for the spin-flip dynamics on $\mathrm{ER}_n(p)$. In Section 5 we derive refined capacity estimates and use these to prove Lemma 2.2. In Section 6 we compare hitting times and prove Lemma 2.3. In Section 7 we show that two copies of the spin-flip dynamics starting near the metastable state can be coupled in a short time. In Section 8, finally, we use Lemmas 2.2–2.3 to prove Theorem 1.4.

2. Preparations

In Section 2.1 we define the perturbed spin-flip dynamics on K_n that will be used as comparison object. In Section 2.2 we state two lemmas that quantify the comparison (Lemmas 2.2–2.3 below). In Section 2.3 we do a rough metastability analysis of the perturbed model. In Section 2.4 we show that $R_{p,\beta,h}$ has a double-well structure if and only if (β, h) is in the metastable regime, in the sense of Definition 1.1 (Lemma 2.4 below).

Define

$$J_n^*(a) = 2\beta\left(p\left(a+\tfrac{2}{n}\right)+h\right)+\log\left(\frac{1-a}{1+a+\frac{2}{n}}\right), \quad a \in \Gamma_n. \tag{2.1}$$

We see from (1.21) that $J_n(a) = J_n^*(a) + O(n^{-2})$ when $\beta p = \frac{1}{1-a^2}$. This will be useful for the analysis of the 'free energy landscape'.

2.1. Perturbed Curie-Weiss.

We will compare the dynamics on $\mathrm{ER}_n(p)$ with that on K_n, but with a ferromagnetic coupling strength p/n rather than $1/n$, and with an external magnetic field field that is a *small perturbation* of h.

Definition 2.1 (Perturbed Curie-Weiss).
(1) Let

$$H_n^u(\sigma) = -\tfrac{p}{2n}\sum_{1\le i,j\le n}\sigma(i)\sigma(j) - h_n^u\sum_{1\le i\le n}\sigma(i), \quad \sigma\in S_n, \tag{2.2}$$

$$H_n^l(\sigma) = -\tfrac{p}{2n}\sum_{1\le i,j\le n}\sigma(i)\sigma(j) - h_n^l\sum_{1\le i\le n}\sigma(i), \quad \sigma\in S_n, \tag{2.3}$$

be the Hamiltonians on S_n corresponding to the Curie-Weiss model on n vertices with ferromagnetic coupling strength p/n, and with external magnetic fields h_n^u and h_n^l given by

$$h_n^u = h + \tfrac{(1+\epsilon)\log(n^{11/6})}{n}, \qquad h_n^l = h - \tfrac{(1+\epsilon)\log(n^{11/6})}{n}, \tag{2.4}$$

where $\epsilon > 0$ is arbitrary. The indices u and l stand for upper and lower, and the choice of exponent $\frac{11}{6}$ will become clear in Section 4.
(2) The equilibrium measures on S_n corresponding to (2.2) and (2.3) are denoted by μ_n^u and μ_n^l, respectively (recall (1.7)).
(3) The Glauber dynamics based on (2.2) and (2.3) are denoted by

$$\{\xi_t^u\}_{t\ge0}, \qquad \{\xi_t^l\}_{t\ge0}, \tag{2.5}$$

respectively.
(4) The analogues of (1.16) and (1.17) are denoted by $\mathbf{m}_n^u, \mathbf{t}_n^u, \mathbf{s}_n^u$, $\mathbf{M}_n^u, \mathbf{T}_n^u, \mathbf{S}_n^u$ and $\mathbf{m}_n^l, \mathbf{t}_n^l, \mathbf{s}_n^l$, $\mathbf{M}_n^l, \mathbf{T}_n^l, \mathbf{S}_n^l$, respectively. ■

In what follows we will *suppress the n-dependence from most of the notation.* Almost all of the analysis in Sections 2–8 pertains to the dynamics on $\mathrm{ER}_n(p)$.

2.2. **Comparison lemmas.** We define the shorthand notations

$$\tau_{\mathbf{m}} = \tau_{A_{\mathbf{m}}}, \qquad \tau_{\mathbf{s}} = \tau_{A_{\mathbf{s}}}, \qquad \tau_{\mathbf{t}} = \tau_{A_{\mathbf{t}}}. \tag{2.6}$$

Recall that $\mathbb{P}_{\mathrm{ER}_n(p)}$ denotes the law of $\mathrm{ER}_n(p)$. We will see that, for a typical realization of the random graph $\mathrm{ER}_n(p)$, when $\{\xi_t\}_{t\geq 0}$ starts from a configuration slightly to the left of the *top*, it takes less time to reach the stable set than $\{\xi^l_t\}_{t\geq 0}$ and more time than $\{\xi^u_t\}_{t\geq 0}$.

Our first comparison lemma establishes a degree of uniformity of the average crossover time to the stable set as a function of the initial configuration inside the basin of the metastable set. Its proof will be given in Section 5.

Lemma 2.2 (Uniformity of average crossover times). *For any $\epsilon > 0$ and any initial configurations ξ_0, $\tilde{\xi}_0$ with $|\xi_0|, |\tilde{\xi}_0| \leq (1-\epsilon)\,\mathbf{t}$,*

$$\mathbb{E}_{\xi_0}[\tau_{\mathbf{s}}]/\mathbb{E}_{\tilde{\xi}_0}[\tau_{\mathbf{s}}] = 1 + o_n(1). \tag{2.7}$$

Our second comparison lemma establishes a degree of uniformity of the hitting probabilities of upward volume level sets as a function of the initial configuration. Its proof will be given in Section 6.

Lemma 2.3 (Uniformity of hitting probabilities of volume level sets). *For any $0 \leq k < m$ with $m \in (\mathbf{m}, \mathbf{s})$ and any configurations $\xi \in A_k$ and $\sigma, \sigma' \in A_m$,*

$$\mathbb{P}_{\xi}[\tau_\sigma < \tau_{\sigma'}] = \tfrac{1}{2} + O\left(\tfrac{1}{\sqrt{n}}\right). \tag{2.8}$$

2.3. **Metastability for perturbed Curie-Weiss.** Recall that $\{\xi^u_t\}_{t\geq 0}$ and $\{\xi^l_t\}_{t\geq 0}$ denote the Glauber dynamics for the Curie-Weiss model driven by (2.2) and (2.3), respectively. An important feature is that their magnetization processes

$$\begin{aligned} \{\theta^u_t\}_{t\geq 0} &= \{m(\xi^l_{\tau^l_s})\}_{t\geq 0} \\ \{\theta^l_t\}_{t\geq 0} &= \{m(\xi^u_{\tau^u_s})\}_{t\geq 0} \end{aligned} \tag{2.9}$$

are continuous-time Markov processes themselves (see e.g. Bovier and den Hollander [4, Chaper 13]). The state space of these two processes is $\Gamma_n = \left\{-1, -1+\frac{2}{n}, \dots, 1-\frac{2}{n}, 1\right\}$, and the transition rates are

$$q^u(a,a') = \begin{cases} \frac{n}{2}(1-a)\, e^{-\beta[p(-2a-\frac{2}{n})-2h^u]_+}, & \text{if } a' = a+\frac{2}{n}, \\ \frac{n}{2}(1+a)\, e^{-\beta[p(2a+\frac{2}{n})+2h^u]_+}, & \text{if } a' = a-\frac{2}{n}, \\ 0, & \text{otherwise}, \end{cases} \tag{2.10}$$

$$q^l(a,a') = \begin{cases} \frac{n}{2}(1-a)\, e^{-\beta[p(-2a-\frac{2}{n})-2h^l]_+}, & \text{if } a' = a+\frac{2}{n}, \\ \frac{n}{2}(1+a)\, e^{-\beta[p(2a+\frac{2}{n})+2h^l]_+}, & \text{if } a' = a-\frac{2}{n}, \\ 0, & \text{otherwise}, \end{cases} \tag{2.11}$$

respectively. The processes in (2.9) are reversible with respect to the Gibbs measures

$$\nu^u(a) = \frac{1}{z^u} e^{\beta n(\frac{1}{2}pa^2 + h^u a)} \binom{n}{\frac{1+a}{2}n}, \quad a \in \Gamma_n, \tag{2.12}$$

$$\nu^l(a) = \frac{1}{z^l} e^{\beta n(\frac{1}{2}pa^2 + h^l a)} \binom{n}{\frac{1+a}{2}n}, \quad a \in \Gamma_n, \tag{2.13}$$

respectively.

Define

$$\Psi^u(a) = -\tfrac{1}{2}pa^2 - h^u a, \quad a \in \Gamma_n, \tag{2.14}$$

$$\Psi^l(a) = -\tfrac{1}{2}pa^2 - h^l a, \quad a \in \Gamma_n. \tag{2.15}$$

Note that (2.10) and (2.12) can be written as

$$\begin{aligned} q^u\left(a, a+\tfrac{2}{n}\right) &= \tfrac{n}{2}(1-a)\, e^{-\beta n[\Psi^u(a+\frac{2}{n})-\Psi^u(a)]_+}, \\ \nu^u(a) &= \frac{1}{z^u} e^{-\beta n \Psi^u(a)} \binom{n}{\frac{n}{2}(1+a)}, \end{aligned} \tag{2.16}$$

and similar formulas hold for (2.11) and (2.13). The properties of the function $\nu^u\colon \Gamma_n \to [0,1]$ can be analysed by looking at the ratio of adjacent values:

$$\frac{\nu^u\left(a+\frac{2}{n}\right)}{\nu^u\left(a\right)} = \exp\left(2\beta\left(p\left(a+\tfrac{2}{n}\right)+h^u\right)+\log\left(\tfrac{1-a}{1+a+\frac{2}{n}}\right)\right), \tag{2.17}$$

which suggests that 'local free energy wells' in ν^u can be found by looking at where the sign of

$$2\beta\left(p\left(a+\tfrac{2}{n}\right)+h^u\right)+\log\left(\tfrac{1-a}{1+a+\frac{2}{n}}\right) \tag{2.18}$$

changes from negative to positive. To that end note that, in the limit $n\to\infty$, the second term is positive for $a<0$, tends to ∞ as $a\to-1$, is negative for $a\geq 0$, tends to $-\infty$ as $a\to 1$, and tends to 0 as $a\to 0$. The first term is linear in a, and for appropriate choices of p, β and h^u (see Definition 1.1) is negative near $a=-1$ and becomes positive at some value $a<0$. This implies that, for appropriate choices of p, β and h^u, the sum of the two terms in (2.18) can change sign $+\to-\to+$ on the interval $[-1,0]$, and can change sign $+\to-$ on $[0,1]$. Assuming that our choice of p, β and h^u corresponds to this change-of-signs sequence, we define $\mathbf{m}^u$, $\mathbf{t}^u$ and $\mathbf{s}^u$ as in (1.16) with h replaced by h^u. This observation makes it clear that the sets in the right-hand side of (1.16) indeed are non-empty.

The interval $[\mathbf{m}^u,\mathbf{t}^u]$ poses a barrier for the process $\{\theta^u_t\}_{t\geq 0}$ due to a negative drift, which delays the initiation of the convergence to equilibrium while the process passes through the interval $[\mathbf{t}^u,\mathbf{s}^u]$. The same is true for the process $\{\xi^u_t\}_{t\geq 0}$. Similar observations hold for $\{\theta^l_t\}_{t\geq 0}$ and $\{\xi^l_t\}_{t\geq 0}$. Recall Fig. 4.

2.4. Double-well structure.

Lemma 2.4 **(Metastable regime).** *The potential $R_{p,\beta,h}$ defined in* (1.26) *has a double-well structure if and only if $\beta p>1$ and $0<h<p\chi(\beta p)$, with χ defined in* (1.14).

Proof. In order for $R_{p,\beta,h}$ to have a double-well structure, the measure ν^u must have two distinct maxima on the interval $(-1,1)$. From (1.22), (1.27) and (2.17) it follows that

$$J_{p,\beta,h}(a) = 2\lambda\left(a+\tfrac{h}{p}\right)+\log\left(\tfrac{1-a}{1+a}\right), \qquad \lambda=\beta p, \tag{2.19}$$

must have one local minimum and two zeroes in $(-1,1)$. Since

$$J'_{p,\beta,h}(a) = 2\left(\lambda-\tfrac{1}{1-a^2}\right), \qquad a\in[-1,1], \tag{2.20}$$

it must therefore be that $\lambda>1$. The local minimum is attained when

$$\lambda = \tfrac{1}{1-a^2}, \tag{2.21}$$

i.e., when $a=a_\lambda=-\sqrt{1-\frac{1}{\lambda}}$ (a_λ must be negative because it lies in $(\mathbf{m},\mathbf{t})$; recall Fig. 4). Since

$$0 > J_{p,\beta,h}(a_\lambda) = 2\lambda\left(a_\lambda+\tfrac{h}{p}\right)+\log\left(\tfrac{1-a_\lambda}{1+a_\lambda}\right), \tag{2.22}$$

it must therefore be that

$$\tfrac{h}{p} < \chi(\lambda) \tag{2.23}$$

with $\chi(\lambda)$ given by (1.14). □

3. Basic facts

In this section we collect a few facts that will be needed in Section 4 to derive capacity estimates for the dynamics on $\mathrm{ER}_n(p)$. In Section 3.1 we derive a large deviation bound for the degree of typical vertices $\mathrm{ER}_n(p)$ (Lemma 3.2 below). In Section 3.2 we do the same for the edge-boundary of typical configurations (Lemma 3.3 below). In Section 3.3 we derive upper and lower bounds for the jump rates of the volume process (Lemmas 3.4–3.5 and Corollary 3.6 below), and show that the return times to the metastable set *conditional* on not hitting the top set are small (Lemma 3.7). In Section 3.4 we use the various bounds to show that the probability for the volume process to grow by $n^{1/3}$ is almost uniform in the starting configuration (Lemma 3.8 below).

Definition 3.1 (Notation). For a vertex $v \in V$, we will write $v \in \sigma$ to mean $\sigma(v) = +1$ and $v \notin \sigma$ to mean $\sigma(v) = -1$. Similarly, we will denote by $\overline{\sigma}$ the configuration obtained from σ by flipping the spin at every vertex, i.e., $\sigma(v) = +1$ if and only if $\overline{\sigma}(v) = -1$. For two configurations σ, σ' we will say that $\sigma \subseteq \sigma'$ if and only if $v \in \sigma \implies v \in \sigma'$. By $\sigma \cup \sigma'$ we denote the configuration satisfying $v \in \sigma \cup \sigma'$ if and only if $v \in \sigma$ or $v \in \sigma'$. A similar definition applies to $\sigma \cap \sigma'$. We will also write $\sigma \sim \sigma'$ when there is a $v \in V$ such that $\sigma = \sigma' \cup \{v\}$ or $\sigma' = \sigma \cup \{v\}$. We will say that σ and σ' are neighbours. We write $\deg(v)$ for the degree of $v \in V$. ■

3.1. **Concentration bounds for** $\mathrm{ER}_n(p)$**.** Recall that $\mathbb{P}_{\mathrm{ER}_n(p)}$ denotes the law $\mathrm{ER}_n(p)$.

Lemma 3.2 (Concentration of degrees and energies). *With $\mathbb{P}_{\mathrm{ER}_n(p)}$-probability tending to 1 as $n \to \infty$ the following is true. For any $\varrho\colon \mathbb{N} \to \mathbb{R}_+$ satisfying*

$$\lim_{n\to\infty} \varrho(n) = \infty \tag{3.1}$$

and any $c > \sqrt{\frac{1}{8}\log 2}$,

$$pn - \varrho(n)\sqrt{n\log n} < \deg(v) < pn + \varrho(n)\sqrt{n\log n} \qquad \forall\, v \in V, \tag{3.2}$$

$$\begin{aligned} \tfrac{1}{n}\left(2p|\xi|(n-|\xi|) - cn^{3/2}\right) - 2h|\xi| &\le H_n(\xi) - H_n(\boxminus) \\ &\le \tfrac{1}{n}\left(2p|\xi|(n-|\xi|) + cn^{3/2}\right) - 2h|\xi| \qquad \forall\, \xi \in S_n. \end{aligned} \tag{3.3}$$

Proof. These bounds are immediate from Hoeffding's inequality and a union bound. □

3.2. **Edge boundaries of** $\mathrm{ER}_n(p)$**.** We partition the configuration space as

$$S_n = \bigcup_{k=0}^{n} A_k, \tag{3.4}$$

where A_k is defined in (1.18). For $0 \le k \le n$ and $-pk\,(n-k) \le i \le (1-p)\,k\,(n-k)$, define

$$\phi_i^k = \left|\{\sigma \in A_k\colon\; |\partial_E \sigma| = pk\,(n-k) + i\}\right|, \tag{3.5}$$

i.e., ϕ_i^k counts the configurations σ with volume k whose edge-boundary size $|\partial_E\sigma|$ deviates by i from its mean, which is equal to $pk\,(n-k)$. For $0 \le k \le n$, let $\mathbb{P}_k$ denote the uniform distribution on A_k.

Lemma 3.3 (Upper bound on edge-boundary sizes). *With $\mathbb{P}_{\mathrm{ER}_n(p)}$-probability tending to 1 as $n \to \infty$ the following are true. For $-pk(n-k) \le j \le (1-p)k(n-k)$ and $\varrho\colon \mathbb{N} \to \mathbb{R}_+$ satisfying* (3.1),

$$\mathbb{P}_k\left[\phi_j^k \ge \varrho\,(n)\binom{n}{k} p^{pk(n-k)+j}(1-p)^{(1-p)k(n-k)-j}\binom{k(n-k)}{pk(n-k)+j}\right] \le \frac{1}{\varrho(n)} \tag{3.6}$$

and

$$\begin{aligned} \mathbb{P}_k\left[\sum_{j\ge i}\phi_j^k \ge \varrho\,(n)\binom{n}{k} e^{-\frac{2i^2}{k(n-k)}}\right] &\le \frac{1}{\varrho\,(n)}, \\ \mathbb{P}_k\left[\sum_{j\le -i}\phi_j^k \ge \varrho\,(n)\binom{n}{k} e^{-\frac{2i^2}{k(n-k)}}\right] &\le \frac{1}{\varrho\,(n)}. \end{aligned} \tag{3.7}$$

Proof. Write $\simeq$ to denote equality in distribution. Note that if $\sigma \simeq \mathbb{P}_k$, then $|\partial_E\sigma| \simeq \mathrm{Bin}\,(k\,(n-k)\,,p)$, and hence

$$\mathbb{P}_k\left[|\partial_E\sigma| = i\right] = p^i\,(1-p)^{k(n-k)-i}\binom{k\,(n-k)}{i}. \tag{3.8}$$

In particular,

$$\mathbb{E}_k\left[\phi_j^k\right] = \mathbb{E}_k\left[\sum_{\sigma\in A_k}\mathbb{1}_{\{|\partial_E\sigma| = pk(n-k)+j\}}\right] = \binom{n}{k}p^{pk(n-k)+j}(1-p)^{(1-p)k(n-k)-j}\binom{k\,(n-k)}{pk(n-k)+j}. \tag{3.9}$$

Hence, by Markov's inequality, the claim in (3.6) follows. Moreover,

$$\mathbb{E}_k\left[\sum_{j\geq i}\phi_j^k\right]=\mathbb{E}_k\left[\sum_{\sigma\in A_k}\mathbb{1}_{\{|\partial_E\sigma|\geq pk(n-k)+i\}}\right]\leq\binom{n}{k}e^{-2\frac{i^2}{k(n-k)}},\tag{3.10}$$

where we again use Hoeffding's inequality. Hence, by Markov's inequality, we get the first line in (3.7). The proof of the second line is similar. □

3.3. Jump rates for the volume process. The following lemma establishes bounds on the rate at which configurations in A_k jump forward to A_{k+1} and backward to A_{k-1}. In Appendix A we will sharpen the error in the prefactors in (3.11)–(3.12) from $2n^{2/3}$ to $O(1)$ and the error in the exponents in (3.11)–(3.12) from $3n^{-1/3}$ to $O(n^{-1/2})$. The formulas in (3.14) and (3.15) show that for small and large magnetization the rate forward, respectively, backward are maximal.

Lemma 3.4 (Bounds on forward jump rates). *With $\mathbb{P}_{\mathrm{ER}_n(p)}$-probability tending to 1 as $n\to\infty$ the following are true.*
(a) For $2n^{1/3}\leq k\leq n-2n^{1/3}$,

$$\begin{aligned}&\left(n-k-2n^{2/3}\right)e^{-2\beta[\vartheta_k+3n^{-1/3}]_+}\\&\quad\leq\sum_{\xi\in A_{k+1}}r(\sigma,\xi)\leq\left(n-k-2n^{2/3}\right)e^{-2\beta[\vartheta_k-3n^{-1/3}]_+}+2n^{2/3},\qquad\sigma\in A_k,\end{aligned}\tag{3.11}$$

and

$$\begin{aligned}&\left(k-2n^{2/3}\right)e^{-2\beta[-\vartheta_k+3n^{-1/3}]_+}\\&\quad\leq\sum_{\xi\in A_{k-1}}r(\sigma,\xi)\leq\left(k-2n^{2/3}\right)e^{-2\beta[-\vartheta_k-3n^{-1/3}]_+}+2n^{2/3},\qquad\sigma\in A_k,\end{aligned}\tag{3.12}$$

where

$$\vartheta_k=p\left(1-\tfrac{2k}{n}\right)-h.\tag{3.13}$$

(b) For $n-\frac{n}{3}(p+h)\leq k<n$,

$$\sum_{\xi\in A_{k+1}}r(\sigma,\xi)=n-k,\qquad\sigma\in A_k.\tag{3.14}$$

(c) For $0<k\leq\frac{n}{3}(p-h)$,

$$\sum_{\xi\in A_{k-1}}r(\sigma,\xi)=k,\qquad\sigma\in A_k.\tag{3.15}$$

Proof. The proof is via probabilistic counting.

(a) Write $\mathbb{P}$ for the law under which $\sigma\in S_n$ is a uniformly random configuration and $v\in\overline{\sigma}$ is a uniformly random vertex. By Hoeffding's inequality, the probability that v has more than $p|\sigma|+n^{2/3}$ neighbours in σ (i.e., $w\in V$ such that $(v,w)\in E$ and $\sigma(w)=+1$) is bounded by

$$\mathbb{P}\left[|E(v,\sigma)|\geq p\,|\overline{\sigma}|+n^{2/3}\right]\leq e^{-2n^{1/3}}.\tag{3.16}$$

where

$$E(v,\sigma)=\{w\in\sigma:\ (v,w)\in E\}.\tag{3.17}$$

Define the event

$$R^+(\sigma)=\left\{\exists\,\zeta\subseteq\overline{\sigma},\,\zeta\in A_{2n^{2/3}}:\ |E(v,\sigma)|\geq p\,|\sigma|+n^{2/3}\ \forall\,v\in\zeta,\right\},\tag{3.18}$$

i.e., the configuration $\overline{\sigma}$ has at least $2n^{2/3}$ vertices like v, each with at least $p|\sigma|+n^{2/3}$ neighbours in σ. Then, for $0\leq k\leq n-2n^{2/3}$,

$$\mathbb{P}\left[R^+(\sigma)\right]\leq\binom{|\overline{\sigma}|}{2n^{2/3}}\left(e^{-2n^{1/3}}\right)^{2n^{2/3}}\leq 2^ne^{-4n}.\tag{3.19}$$

Hence the probability that some configuration $\sigma \in S_n$ satisfies condition $R^+(\sigma)$ is bounded by

$$\mathbb{P}\left[\bigcup_{\sigma\in S_n} R^+(\sigma)\right] \le 4^n e^{-4n} \le e^{-2n}. \tag{3.20}$$

Thus, with $\mathbb{P}_{\mathrm{ER}_n(p)}$-probability tending to 1 as $n\to\infty$ there are no configurations $\sigma\in S_n$ satisfying condition $R^+(\sigma)$. The same holds for the event

$$R^-(\sigma) = \left\{\exists\,\zeta\subseteq\overline{\sigma},\,\zeta\in A_{2n^{2/3}}\colon\ |E(v,\sigma)| \le p\,|\sigma| - n^{2/3}\ \forall\, v\in\zeta\right\}, \tag{3.21}$$

for which

$$\mathbb{P}\left[\bigcup_{\sigma\in S_n} R^-(\sigma)\right] \le e^{-2n}. \tag{3.22}$$

Now let $\sigma\in A_k$, and observe that σ has $n-k$ neighbours in A_{k+1} and k neighbours in A_{k-1}. But if $\xi=\sigma\cup\{v\}\in A_{k+1}$, then by (1.3),

$$\begin{aligned} H_n(\xi)-H_n(\sigma) &= \tfrac{2}{n}\Big(|E(v,\overline{\sigma})| - |E(v,\sigma)|\Big) - 2h \\ &= \tfrac{2}{n}\big(\deg(v) - 2\,|E(v,\sigma)|\big) - 2h \\ &\le \tfrac{2}{n}\big(pn + n^{1/2}\log n - 2\,|E(v,\sigma)|\big) - 2h, \end{aligned} \tag{3.23}$$

where the last inequality uses (3.2) with $\varrho(n)=\log n$. Similarly,

$$H_n(\xi)-H_n(\sigma) \ge \tfrac{2}{n}\Big(pn - n^{1/2}\log n - 2\,|E(v,\sigma)|\Big) - 2h. \tag{3.24}$$

The events $R^+(\sigma)$ in (3.18) and $R^-(\sigma)$ in (3.21) guarantee that for any configuration σ at most $2n^{2/3}$ vertices in the configuration $\overline{\sigma}$ can have more than $n^{2/3}$ neighbours in σ. In other words, the configuration σ has at most $2n^{2/3}$ neighbouring configurations in A_{k+1} that differ in energy by more than $6n^{-1/3}-2h$. Since on the complement of $R^+(\sigma)$ with $\sigma\in A_k$ we have $|\{w\in\sigma\colon (v,w)\in E\}| \le 2pk+2n^{1/3}$ (because $n^{1/2}\log n\le n^{2/3}$ for n large enough), from (3.20) and (3.22) we get that, with $\mathbb{P}_{\mathrm{ER}_n(p)}$-probability at least $1-e^{-2n}$,

$$\begin{aligned} &\left|\left\{\xi\in A_{k+1}\colon\ \xi\sim\sigma,\ H_n(\xi)-H_n(\sigma)\ge \tfrac{2}{n}\Big(pn-2pk+3n^{2/3}\Big)-2h\right\}\right| \le 2n^{2/3}, \\ &\left|\left\{\xi\in A_{k+1}\colon\ \xi\sim\sigma,\ H_n(\xi)-H_n(\sigma)\le \tfrac{2}{n}\Big(pn-2pk-3n^{2/3}\Big)-2h\right\}\right| \le 2n^{2/3}, \end{aligned} \tag{3.25}$$

and hence, by (1.6), the rate at which the Markov chain starting at $\sigma\in A_k$ jumps to A_{k+1} satisfies

$$\sum_{\xi\in A_{k+1}} r(\sigma,\xi) \ge \big(n-k-2n^{2/3}\big)e^{-2\beta[\theta_k+3n^{-1/3}]_+}, \tag{3.26}$$

$$\sum_{\xi\in A_{k+1}} r(\sigma,\xi) \le \big(n-k-2n^{2/3}\big)e^{-2\beta[\theta_k-3n^{-1/3}]_+} + 2n^{2/3}. \tag{3.27}$$

Here the term $n-k-2n^{2/3}$ comes from exclusion of the at most $2n^{2/3}$ neighbours in configurations that differ from σ in energy by more than $6n^{-1/3}-2h$. Similarly, with $\mathbb{P}_{\mathrm{ER}_n(p)}$-probability at least $1-e^{-2n}$,

$$\begin{aligned} &\left|\left\{\xi\in A_{k-1}\colon\ \xi\sim\sigma,\ H_n(\xi)-H_n(\sigma)\ge \tfrac{2}{n}\Big(-pn+2pk+3n^{2/3}\Big)+2h\right\}\right| \le 2n^{2/3}, \\ &\left|\left\{\xi\in A_{k-1}\colon\ \xi\sim\sigma,\ H_n(\xi)-H_n(\sigma)\le \tfrac{2}{n}\Big(-pn+2pk-3n^{2/3}\Big)+2h\right\}\right| \le 2n^{2/3}, \end{aligned} \tag{3.28}$$

and hence, by (1.6), the rate at which the Markov chain starting at $\sigma\in A_k$ jumps to A_{k-1} satisfies

$$\sum_{\xi\in A_{k-1}} r(\sigma,\xi) \le \big(k-2n^{2/3}\big)e^{-2\beta[-\theta_k-3n^{-1/3}]_+} + 2n^{2/3}, \tag{3.29}$$

$$\sum_{\xi\in A_{k-1}} r(\sigma,\xi) \ge \big(k-2n^{2/3}\big)e^{-2\beta[-\theta_k+3n^{-1/3}]_+}. \tag{3.30}$$

This proves (3.11) and (3.12).

(b) To get (3.14), note that for $\xi = \sigma \cup \{v\}$ with $v \notin \sigma$,

$$H_n(\xi) - H_n(\sigma) = \tfrac{2}{n}\Big(|E(v,\overline{\sigma})| - |E(v,\sigma)| \Big) - 2h = \tfrac{2}{n}\Big(2|E(v,\sigma)| - \deg(v) \Big) - 2h \leq 2\Big(2(n-k) - p + n^{-1/2}\log n - h \Big) \tag{3.31}$$

for n large enough, which is ≤ 0 when $n - k \leq \frac{n}{3}(p+h)$, so that $r(\sigma,\xi) = 1$ by (1.6).

(c) To get (3.15), note that for $\xi = \sigma \setminus \{v\}$ with $v \notin \sigma$,

$$H_n(\xi) - H_n(\sigma) = \tfrac{2}{n}\Big(|E(v,\sigma)| - |E(v,\overline{\sigma})| \Big) + 2h = \tfrac{2}{n}\Big(2|E(v,\sigma)| - \deg(v) \Big) + 2h \leq 2\Big(2k - p + n^{-1/2}\log n + h \Big) \tag{3.32}$$

for n large enough, which is ≤ 0 when $k \leq \frac{n}{3}(p-h)$, so that $r(\sigma,\xi) = 1$ by (1.6). □

The following lemma is technical and merely serves to show that near $A_{\mathbf{M}}$ transitions involving a flip from -1 to $+1$ typically occur at rate 1.

Lemma 3.5 (Attraction towards the metastable state). *Suppose that $|\xi| = [1 + o_n(1)]\,\mathbf{M}$. Then $r(\xi,\xi^v) = 1$ for all but $O(n^{2/3})$ many $v \in \xi$.*

Proof. We want to show that

$$H_n(\xi^v) < H_n(\xi) \tag{3.33}$$

for all but $O(n^{2/3})$ many $v \in \xi$. Note that by (3.21) and (3.22) there are at most $2n^{2/3}$ many $v \in \xi$ such that $|E(v,\overline{\xi})| \leq p(n - |\xi|) - n^{2/3}$, and at most $2n^{2/3}$ many $v \in \xi$ such that $|E(v,\xi)| \geq p|\xi| + n^{2/3}$. Hence, by (1.3), for all but at most $4n^{2/3}$ many $v \in \xi$ we have that

$$\begin{aligned} H_n(\xi^v) &= H_n(\xi) + \tfrac{2}{n}\left(|E(v,\xi)| - \left|E\left(v,\overline{\xi}\right)\right|\right) + 2h \\ &= H_n(\xi) + \tfrac{2p}{n}(2|\xi| - n) + 2h + o_n(1) \\ &= H_n(\xi) + \tfrac{2p}{n}(2\mathbf{M} - n) + 2h + o_n(1) \\ &= H_n(\xi) + 2p\mathbf{m} + 2h + o_n(1), \end{aligned} \tag{3.34}$$

where we use (1.17). From the definition of $\mathbf{m}$ in (1.16) it follows that $2p\mathbf{m} + 2h + o_n(1) < 0$, where we recall from the discussion near the end of Section 2.3 that $\mathbf{m} < 0$ and hence $\log(\frac{1-\mathbf{m}}{1+\mathbf{m}}) > 0$. Hence (3.33) follows. □

We can now prove the claim made in Remark 1.2, namely, there is no metastable behaviour outside the regime in (1.13). Recall the definition of $\mathbf{S}_n$ in (1.17), which requires the function J in (1.23) to have two zeroes. If it has only one zero, then denote that zero by a' and define $\mathbf{S}_n = \frac{n}{2}(a'+1)$. Let $A_{\mathbf{S}_n + O(n^{2/3})}$ be the union of all A_k with $|k - \mathbf{S}_n| = O(n^{2/3})$.

Corollary 3.6 (Non-metastable regime). *Suppose that $\beta \in (1/p,\infty)$ and $h \in (p,\infty)$. Then $\{\xi_t\}_{t\geq 0}$ has a drift towards $A_{\mathbf{S}_n + O(n^{2/3})}$. Consequently, $\mathbb{E}_{\xi_0}[\tau_{\mathbf{S}}] = O(1)$ for any initial configuration $\xi_0 \in S_n$.*

Proof. If $\beta \in (1/p,\infty)$ and $h \in (p,\infty)$, then the function $a \mapsto J_{p,\beta,h}(a) = 2\beta(pa+h) + \log(\frac{1-a}{1+a})$ has a unique root in the interval $(0,1)$. Indeed, $J_{p,\beta,h}(a) > 0$ for $a \in [-1,0]$, $J'_{p,\beta,h}(0) = 2(\beta p - 1) > 0$, while $a \mapsto \log(\frac{1-a}{1+a})$ is concave and tends to $-\infty$ as $a \uparrow 1$. We claim that the process $\{\xi_t\}_{t\geq 0}$ drifts towards that root, i.e., if we denote the root by a', then the process drifts towards the set $A_{\frac{n}{2}(a'+1)}$, which by convention we identify with A_{S_n}. Note that if $h \in (p,\infty)$, then $\vartheta_k = p(1 - \frac{2k}{n}) - h < 0$ for all $0 \leq k \leq n$ and so, by Lemma 3.4,

$$\begin{aligned} \sum_{\xi \in A_{k+1}} r(\sigma,\xi) &\geq n - k - 2n^{2/3}, \\ \sum_{\xi \in A_{k-1}} r(\sigma,\xi) &\leq \left(k - 2n^{2/3}\right)e^{-2\beta[-\vartheta_k - 3n^{-1/3}]} + 2n^{2/3}. \end{aligned} \tag{3.35}$$

Thus, for $k \leq \frac{n}{2} - 4n^{2/3}$, $\sum_{\xi \in A_{k+1}} r(\sigma,\xi) > \sum_{\xi \in A_{k-1}} r(\sigma,\xi)$. Similarly, for $k \geq \frac{n}{2} + 4n^{2/3}$, the opposite inequality holds. Therefore there is a drift towards $A_{\mathbf{S}_n + O(n^{2/3})}$. □

We close this section with a lemma stating that the average return time to $A_{\mathbf{M}_n}$ conditional on not hitting $A_{\mathbf{T}_n}$ is of order 1 and has an exponential tail. This will be needed to control the time between successive attempts to go from $A_{\mathbf{M}_n}$ to $A_{\mathbf{T}_n}$, until the dynamics crosses $A_{\mathbf{T}_n}$ and moves to $A_{\mathbf{S}_n}$ (recall Fig. 4).

Lemma 3.7 (**Conditional return time to the metastable set**). *There exist a $C > 0$ such that, with $\mathbb{P}_{\mathrm{ER}_n(p)}$-probability tending to 1 as $n \to \infty$, uniformly in $\xi \in A_{\mathbf{M}_n}$,*

$$\mathbb{P}_\xi \left[\tau_{A_{\mathbf{M}_n}} \geq k \mid \tau_{A_{\mathbf{M}_n}} < \tau_{A_{\mathbf{T}_n}} \right] \leq e^{-Ck}, \qquad \forall k. \tag{3.36}$$

Proof. The proof is given in Appendix A. □

3.4. **Uniformity in the starting configuration.** The following lemma shows that the probability of the event $\{\tau_{A_{k+o(n^{1/3})}} < \tau_{A_k}\}$ is almost uniform as a function of the starting configuration in A_k.

Lemma 3.8 (**Uniformity of hitting probability of volume level sets**).
With $\mathbb{P}_{\mathrm{ER}_n(p)}$-probability tending to 1 as $n \to \infty$, the following is true. For every $0 \leq k < m \leq n$,

$$\frac{\max_{\sigma \in A_k} \mathbb{P}_\sigma \left[\tau_{A_m} < \tau_{A_k} \right]}{\min_{\sigma \in A_k} \mathbb{P}_\sigma \left[\tau_{A_m} < \tau_{A_k} \right]} \leq \left[1 + o_n(1) \right] e^{K(m-k)n^{-1/3}} \tag{3.37}$$

with $K = K(\beta, h, p) \in (0, \infty)$.

Proof. The proof proceeds by estimating the probability of trajectories from A_k to A_m. Observe that

$$\begin{aligned} e^{-2\beta[\vartheta_k + 3n^{-1/3}]_+} &\geq e^{-2\beta[\vartheta_k]_+} \left(1 - 6\beta n^{-1/3} \right), \quad \forall n, \\ e^{-2\beta[\vartheta_k - 3n^{-1/3}]_+} &\leq e^{-2\beta[\vartheta_k]_+} \left(1 + 7\beta n^{-1/3} \right), \quad n \text{ large enough}, \end{aligned} \tag{3.38}$$

and that similar estimates hold for $e^{-2\beta[-\vartheta_k + 3n^{-1/3}]_+}$ and $e^{-2\beta[-\vartheta_k - 3n^{-1/3}]_+}$. We will bound the ratio in the left-hand side of (3.37) by looking at two random processes on $\{0, \ldots, n\}$, one of which bounds $\max_{\sigma \in A_k} \mathbb{P}_\sigma [\tau_{A_m} < \tau_{A_k}]$ from above and the other of which bounds $\min_{\sigma \in A_k} \mathbb{P}_\sigma [\tau_{A_m} < \tau_{A_k}]$ from below. The proof comes in 3 Steps.

1. We begin with the following observation. Suppose that $\{X_t^+\}_{t \geq 0}$ and $\{X_t^-\}_{t \geq 0}$ are two continuous-time Markov chains taking unit steps in $\{0, \ldots, n\}$ at rates $r^-(k, k \pm 1)$ and $r^+(k, k \pm 1)$, respectively. Furthermore, suppose that for every $0 \leq k \leq n-1$,

$$r^-(k, k+1) \leq \min_{\sigma \in A_k} \sum_{\xi \in A_{k+1}} r(\sigma, \xi) \leq \max_{\sigma \in A_k} \sum_{\xi \in A_{k+1}} r(\sigma, \xi) \leq r^+(k, k+1), \tag{3.39}$$

and for every $1 \leq k \leq n$,

$$r^-(k, k-1) \geq \max_{\sigma \in A_k} \sum_{\xi \in A_{k-1}} r(\sigma, \xi) \geq \min_{\sigma \in A_k} \sum_{\xi \in A_{k-1}} r(\sigma, \xi) \geq r^+(k, k-1). \tag{3.40}$$

Then

$$\frac{\max_{\sigma \in A_k} \mathbb{P}_\sigma \left[\tau_{A_m} < \tau_{A_k} \right]}{\min_{\sigma \in A_k} \mathbb{P}_\sigma \left[\tau_{A_m} < \tau_{A_k} \right]} \leq \frac{\mathbb{P}_k^{X^+} \left[\tau_m < \tau_k \right]}{\mathbb{P}_k^{X^-} \left[\tau_m < \tau_k \right]}. \tag{3.41}$$

Indeed, from (3.39) and (3.40) it follows that we can couple the three Markov chains $\{X_t^+\}_{t \geq 0}$, $\{X_t^-\}_{t \geq 0}$ and $\{\xi_t\}_{t \geq 0}$ in such a way that, for any $0 \leq k \leq n$ and any $\sigma_0 \in A_k$, if $X_0^- = X_0^+ = |\sigma_0| = k$, then

$$X_t^- \leq |\sigma_t| \leq X_t^+, \qquad t \geq 0. \tag{3.42}$$

This immediately guarantees that, for any $0 \leq k \leq m \leq n$,

$$\mathbb{P}_k^{X^-} \left[\tau_m < \tau_k \right] \leq \min_{\sigma \in A_k} \mathbb{P}_\sigma \left[\tau_{A_m} < \tau_{A_k} \right] \leq \max_{\sigma \in A_k} \mathbb{P}_\sigma \left[\tau_{A_m} < \tau_{A_k} \right] \leq \mathbb{P}_k^{X^+} \left[\tau_m < \tau_k \right], \tag{3.43}$$

which proves the claim in (3.41). To get (3.39) and (3.40), we pick $r^-(i,j)$ and $r^+(i,j)$ such that

$$r^-(i,j) = \begin{cases} \left(n-i-(2+6\beta)\,n^{2/3}\right)e^{-2\beta[\vartheta_i]_+} + (2+6\beta)\,n^{2/3}\mathbb{1}_{\left\{i\geq n\left(1-\frac{1}{3}(p+h)\right)\right\}}, & j=i+1,\\ \min\left\{i,\ \left(i+(-2+7\beta)\,n^{2/3}\right)e^{-2\beta[-\vartheta_i]_+}+2n^{2/3}\right\}, & j=i-1,\\ 0, & \text{otherwise,}\end{cases} \tag{3.44}$$

and

$$r^+(i,j) = \begin{cases} \min\left\{n-i,\ \left(n-i+(-2+7\beta)\,n^{2/3}\right)e^{-2\beta[\vartheta_i]_+}+2n^{2/3}\right\}, & j=i+1,\\ \left(i-(2+6\beta)\,n^{2/3}\right)e^{-2\beta[-\vartheta_i]_+}, & j=i-1,\\ 0, & \text{otherwise,}\end{cases} \tag{3.45}$$

and note that, by Lemma 3.4 and (3.38), (3.39)–(3.40) are indeed satisfied.

2. We continue from (3.41). Our task is to estimate the right-hand side of (3.41). Let $\mathcal{G}$ be the set of all unit-step paths from k to m that only hit m after their final step:

$$\mathcal{G} = \bigcup_{M\in\mathbb{N}} \left\{\{\gamma_i\}_{i=0}^{M-1} :\ \gamma_0=k,\ \gamma_M=m,\ \gamma_i\in\{0,\dots,m-1\} \text{ and } |\gamma_{i+1}-\gamma_i|=1 \text{ for } 0\leq i<M\right\}. \tag{3.46}$$

We will show that

$$\frac{\mathbb{P}_k^{X^+}\left[X_t^+ \text{ follows trajectory } \gamma\right]}{\mathbb{P}_k^{X^-}\left[X_t^- \text{ follows trajectory } \gamma\right]} \leq \exp\left(\left[24\beta+4e^{2\beta(p+h+1)}\right](m-k)\,n^{-1/3}\right) \quad \forall\,\gamma\in\mathcal{G}, \tag{3.47}$$

which will settle the claim. (Note that the paths realising $\{\tau_m<\tau_k\}$ form a subset of $\mathcal{G}$.) To that end, let $\gamma^\star\in\mathcal{G}$ be the path $\gamma^\star=\{k,k+1,\dots,m\}$. We claim that

$$\sup_{\gamma\in\mathcal{G}} \frac{\mathbb{P}_k^{X^+}\left[X_t^+ \text{ follows trajectory } \gamma\right]}{\mathbb{P}_k^{X^-}\left[X_t^- \text{ follows trajectory } \gamma\right]} \leq \frac{\mathbb{P}_k^{X^+}\left[X_t^+ \text{ follows trajectory } \gamma^\star\right]}{\mathbb{P}_k^{X^-}\left[X_t^- \text{ follows trajectory } \gamma^\star\right]}. \tag{3.48}$$

Indeed, if $\gamma=(\gamma_1,\dots,\gamma_M)\in\mathcal{G}$, then by the Markov property we have that

$$\mathbb{P}_k^{X^+}\left[X_t^+ \text{ follows trajectory } \gamma\right] = \prod_{i=0}^{M-1}\mathbb{P}_{\gamma_i}^{X^+}\left[\tau_{\gamma_{i+1}}<\tau_{\gamma_i}\right], \tag{3.49}$$

with a similar expression for $\mathbb{P}_k^{X^-}[X_t^- \text{ follows trajectory } \gamma]$. Therefore, noting that $\gamma_i-1=2\gamma_i-\gamma_{i+1}$ when $\gamma_{i+1}=\gamma_i+1$ and $\gamma_i+1=2\gamma_i-\gamma_{i+1}$ when $\gamma_{i+1}=\gamma_i-1$, we have

$$\begin{aligned}\frac{\mathbb{P}_k^{X^+}\left[X_t^+ \text{ follows trajectory } \gamma\right]}{\mathbb{P}_k^{X^-}\left[X_t^- \text{ follows trajectory } \gamma\right]} &= \prod_{i=0}^{M-1}\frac{\mathbb{P}_{\gamma_i}^{X^+}\left[\tau_{\gamma_{i+1}}<\tau_{\gamma_i}\right]}{\mathbb{P}_{\gamma_i}^{X^-}\left[\tau_{\gamma_{i+1}}<\tau_{\gamma_i}\right]}\\ &= \prod_{i=1}^{M}\left(\frac{r^+(\gamma_i,\gamma_{i+1})}{r^+(\gamma_i,\gamma_{i+1})+r^+(\gamma_i,2\gamma_i-\gamma_{i+1})}\right)\left(\frac{r^-(\gamma_i,\gamma_{i+1})}{r^-(\gamma_i,\gamma_{i+1})+r^-(\gamma_i,2\gamma_i-\gamma_{i+1})}\right)^{-1}.\end{aligned} \tag{3.50}$$

Since, whenever $\gamma_{i+1}=\gamma_i-1$,

$$\begin{aligned}\frac{r^-(\gamma_i,\gamma_{i+1})}{r^-(\gamma_i,\gamma_{i+1})+r^-(\gamma_i,2\gamma_i-\gamma_{i+1})} &= \frac{r^-(\gamma_i,\gamma_i-1)}{r^-(\gamma_i,\gamma_i-1)+r^-(\gamma_i,\gamma_i+1)}\\ &\geq \frac{r^+(\gamma_i,\gamma_i-1)}{r^+(\gamma_i,\gamma_i-1)+r^+(\gamma_i,\gamma_i+1)}\\ &= \frac{r^+(\gamma_i,\gamma_{i+1})}{r^+(\gamma_i,\gamma_{i+1})+r^+(\gamma_i,2\gamma_i-\gamma_{i+1})},\end{aligned} \tag{3.51}$$

we get

$$\prod_{i=0}^{M-1}\left(\frac{r^{+}(\gamma_i,\gamma_{i+1})}{r^{+}(\gamma_i,\gamma_{i+1})+r^{+}(\gamma_i,2\gamma_i-\gamma_{i+1})}\right)\left(\frac{r^{-}(\gamma_i,\gamma_{i+1})}{r^{-}(\gamma_i,\gamma_{i+1})+r^{-}(\gamma_i,2\gamma_i-\gamma_{i+1})}\right)^{-1} \tag{3.52}$$
$$\leq \prod_{i=k}^{m-1}\left(\frac{r^{+}(i,i+1)}{r^{+}(i,i+1)+r^{+}(i,i-1)}\right)\left(\frac{r^{-}(i,i+1)}{r^{-}(i,i+1)+r^{-}(i,i-1)}\right)^{-1}$$
$$=\frac{\mathbb{P}_k^{X^+}\left[X_t^+ \text{ follows trajectory } \gamma^\star\right]}{\mathbb{P}_k^{X^-}\left[X_t^- \text{ follows trajectory } \gamma^\star\right]}.$$

This proves the claim in (3.48).

3. Next, consider the ratio

$$\frac{r^{-}(i,i+1)+r^{-}(i,i-1)}{r^{+}(i,i+1)+r^{+}(i,i-1)}=\frac{A}{B} \tag{3.53}$$

with

$$\begin{aligned} A &= \left(n-i-(2+6\beta)\,n^{2/3}\right)e^{-2\beta[\vartheta_i]_+}+(2+6\beta)\,n^{2/3}\mathbb{1}_{\left\{i\geq n\left(1-\frac{1}{3}(p+h)\right)\right\}} \\ &\quad+\left(i+(-2+7\beta)\,n^{2/3}\right)e^{-2\beta[-\vartheta_i]_+}+2n^{2/3}, \\ B &= \left(n-i+(-2+7\beta)\,n^{2/3}\right)e^{-2\beta[\vartheta_i]_+}+2n^{2/3} \\ &\quad+\left(i-(2+6\beta)\,n^{2/3}\right)e^{-2\beta[-\vartheta_i]_+}, \end{aligned} \tag{3.54}$$

and the ratio

$$\frac{r^{+}(i,i+1)}{r^{-}(i,i+1)}=\frac{C}{D} \tag{3.55}$$

with

$$\begin{aligned} C &= \left(n-i+(-2+7\beta)\,n^{2/3}\right)e^{-2\beta[\vartheta_i]_+}+2n^{2/3}, \\ D &= \left(n-i-(2+6\beta)\,n^{2/3}\right)e^{-2\beta[\vartheta_i]_+}+(2+6\beta)\,n^{2/3}\mathbb{1}_{\left\{i\geq n\left(1-\frac{1}{3}(p+h)\right)\right\}}. \end{aligned} \tag{3.56}$$

Note from (3.53) that for $\vartheta_i \geq 0$ (i.e., $i \leq \frac{n}{2}(1-p^{-1}h)$, in which case also $i < n(1-\frac{1}{3}(p+h)))$,

$$\frac{r^{-}(i,i+1)+r^{-}(i,i-1)}{r^{+}(i,i+1)+r^{+}(i,i-1)} \leq 1+\frac{13\beta e^{2\beta(p-h)}}{n^{1/3}}, \tag{3.57}$$

and from (3.55) it follows that in this case

$$\frac{r^{+}(i,i+1)}{r^{-}(i,i+1)} \quad\leq\quad 1+\frac{3\,(3+13\beta)\,e^{2\beta(p-h)}}{n^{1/3}\,(p+h)}. \tag{3.58}$$

Similarly, for $\vartheta_i < 0$ we have that

$$\frac{r^{-}(i,i+1)+r^{-}(i,i-1)}{r^{+}(i,i+1)+r^{+}(i,i-1)} \quad\leq\quad 1+\frac{2e^{2\beta(p+h)}}{n^{1/3}} \tag{3.59}$$

and

$$\frac{r^{+}(i,i+1)}{r^{-}(i,i+1)} \quad\leq\quad 1+\frac{6\,(2+6\beta)}{n^{1/3}\,(p+h)}. \tag{3.60}$$

Combining (3.57)–(3.60), we get that, for all $1 \leq i \leq n-1$,

$$\frac{r^{-}(i,i+1)+r^{-}(i,i-1)}{r^{+}(i,i+1)+r^{+}(i,i-1)} \times \frac{r^{+}(i,i+1)}{r^{-}(i,i+1)} \leq 1+Kn^{-1/3}, \tag{3.61}$$

where

$$K=\max\left\{e^{2\beta(p-h)}\left(\tfrac{9+39\beta}{p+h}+13\beta\right),\ 2e^{2\beta(p+h)}+\tfrac{12+36\beta}{p+h}\right\}. \tag{3.62}$$

Therefore

$$\frac{\mathbb{P}_k^{X^+}\left[X_t^+ \text{ follows trajectory } \gamma^\star\right]}{\mathbb{P}_k^{X^-}\left[X_t^- \text{ follows trajectory } \gamma^\star\right]} \leq \prod_{i=k}^{m-1}\left(1+\tfrac{K}{n^{1/3}}\right) \leq e^{Kn^{-1/3}(m-k)}. \tag{3.63}$$

□

4. Capacity bounds

The goal of this section is to derive various capacity bounds that will be needed to prove Theorem 1.4 in Sections 7–8. In Section 4.1 we derive capacity bounds for the processes $\{\xi_t^l\}_{t\geq 0}$ and $\{\xi_t^u\}_{t\geq 0}$ on K_n introduced in (2.9) (Lemma 4.1 below). In Section 4.2 we do the same for the process $\{\xi_t\}_{t\geq 0}$ on $\mathrm{ER}_n(p)$ (Lemma 4.2 below). In Section 4.3 we use the bounds to rank-order the mean return times to $A_{\mathbf{M}^l}$, $A_{\mathbf{M}}$ and $A_{\mathbf{M}^u}$, respectively (Lemma 4.3 below). This ordering will be needed in the construction of a coupling in Section 7.

Define the Dirichlet form for $\{\xi_t\}_{t\geq 0}$ by

$$\mathscr{E}(f,f) = \tfrac{1}{2}\sum_{\sigma,\sigma'\in S_n}\mu(\sigma)\,r(\sigma,\sigma')\left[f(\sigma)-f(\sigma')\right]^2, \qquad f\colon\, S_n \to [0,1], \tag{4.1}$$

and for $\{\theta_t^u\}_{t\geq 0}$ and $\{\theta_t^l\}_{t\geq 0}$ by

$$\begin{aligned} \mathscr{E}^u(f,f) &= \tfrac{1}{2}\sum_{a,a'\in\Gamma_n}\nu^u(a)\,q^u(a,a')\left[f(a)-f(a')\right]^2, \\ \mathscr{E}^l(f,f) &= \tfrac{1}{2}\sum_{a,a'\in\Gamma_n}\nu^l(a)\,q^l(a,a')\left[f(a)-f(a')\right]^2, \qquad f\colon\,\Gamma_n\to[0,1]. \end{aligned} \tag{4.2}$$

For $A,B\subseteq S_n$, define the capacity between A and B for $\{\xi_t\}_{t\geq 0}$ by

$$\mathrm{cap}(A,B) = \min_{f\in Q(A,B)} \mathscr{E}(f,f), \tag{4.3}$$

where

$$Q(A,B) = \left\{f\colon\, S_n\to[0,1],\ f_{|A}\equiv 1,\ f_{|B}\equiv 0\right\}, \tag{4.4}$$

and similarly for $\mathrm{cap}^u(A,B)$ and $\mathrm{cap}^l(A,B)$.

4.1. **Capacity bounds on K_n.** First we derive capacity bounds for $\{\xi_t^l\}_{t\geq 0}$ and $\{\xi_t^u\}_{t\geq 0}$ on K_n. A useful reformulation of (4.3) is given by

$$\mathrm{cap}(A,B) = \sum_{\sigma\in A}\sum_{\sigma'\in S_n}\mu(\sigma)\,r(\sigma,\sigma')\,\mathbb{P}_\sigma(\tau_B<\tau_A). \tag{4.5}$$

Lemma 4.1 (**Capacity bounds for $\{\xi_t^u\}_{t\geq 0}$ and $\{\xi_t^l\}_{t\geq 0}$**)**.** *For $a,b\in[\mathbf{m}^u,\mathbf{s}^u]$ with $a<b$,*

$$\frac{\left(1-b+\frac{2}{n}\right)}{2n\,(b-a)^2} \leq \frac{\mathrm{cap}^u(a,b)}{C^\star(b)} \leq \frac{n\,(1-b)}{2} \tag{4.6}$$

with

$$C^\star(b) = \frac{1}{z^u}e^{-\beta n\Psi^u(b)}\binom{n}{\frac{n}{2}\left(1+\min\left(b,\mathbf{t}^u\right)\right)}. \tag{4.7}$$

For $a,b\in\left[\mathbf{m}^l,\mathbf{s}^l\right]$ with $a<b$, analogous bounds hold for $\mathrm{cap}^l(a,b)$.

Proof. We will prove the upper and lower bounds only for $\mathrm{cap}^u(a,b)$, the proof for $\mathrm{cap}^l(a,b)$ being identical. Note from the definition in (4.3) that

$$\begin{aligned} &\mathrm{cap}^u(a,b) \\ &= \min_{f\in Q(a,b)}\sum_{i=0}^{\frac{n}{2}(b-a)-1}\nu^u\left(a+\tfrac{2i}{n}\right)q^u\left(a+\tfrac{2i}{n},a+\tfrac{2(i+1)}{n}\right)\left[f\left(a+\tfrac{2i}{n}\right)-f\left(a+\tfrac{2(i+1)}{n}\right)\right]^2, \end{aligned} \tag{4.8}$$

where it is easy to see that the set $Q(a,b)$ in (4.4) may be reduced to

$$Q(a,b) = \Big\{f\colon\Gamma_n\to[0,1],\ f(x)=1 \text{ for } x\leq a,\ f(x)=0 \text{ for } x\geq b\Big\}. \tag{4.9}$$

Note that for every $f \in Q(a,b)$ there is some $0 \le i \le \frac{n}{2}(b-a)-1$ such that

$$\left| f\left(a+\tfrac{2i}{n}\right) - f\left(a+\tfrac{2(i+1)}{n}\right)\right| \ge \left(\tfrac{n}{2}(b-a)\right)^{-1}. \tag{4.10}$$

Also note that, by (2.16),

$$\begin{aligned}\nu^u\left(a+\tfrac{2i}{n}\right) q^u\left(a+\tfrac{2i}{n}, a+\tfrac{2(i+1)}{n}\right)\\ = \frac{1}{z^u}\frac{n}{2}\left(1-a-\tfrac{2i}{n}\right) e^{-\beta n \max\left\{\Psi^u\left(a+\frac{2i}{n}\right),\Psi^u\left(a+\frac{2(i+1)}{n}\right)\right\}}\binom{n}{\frac{n}{2}(1+a)+i},\end{aligned} \tag{4.11}$$

and that, for any $\delta \in \mathbb{R}$,

$$\Psi^u(a+\delta) - \Psi^u(a) = -\delta\left(pa + h^u + \tfrac{p}{2}\delta\right), \tag{4.12}$$

so that

$$\max\left\{\Psi^u\left(a+\tfrac{2i}{n}\right), \Psi^u\left(a+\tfrac{2(i+1)}{n}\right)\right\} \le \tfrac{2}{n}\left(pa+h+\tfrac{p}{n}\right) + \Psi^u\left(a+\tfrac{2i}{n}\right). \tag{4.13}$$

Combining, (4.9)–(4.13) with $\delta = \frac{2}{n}$, we get

$$\begin{aligned}\mathrm{cap}^u(a,b)\\ &\ge \min_{0\le i\le \frac{n}{2}(b-a)-1} \frac{2\left(1-b+\frac{2}{n}\right)e^{-2\beta(p+h^u)}}{nz^u(b-a)^2} e^{-\beta n \Psi^u\left(a+\frac{2i}{n}\right)}\binom{n}{\frac{n}{2}(1+a)+i}\\ &= \frac{2\left(1-b+\frac{2}{n}\right)e^{-2\beta(p+h^u+\frac{p}{n})}}{nz^u(b-a)^2} e^{-\beta n\Psi^u(\min(b,\mathbf{t}^u))}\binom{n}{\frac{n}{2}(1+\min(b,\mathbf{t}^u))},\end{aligned} \tag{4.14}$$

where we use (4.12) that, by the definition of $\mathbf{m}^u$, $\mathbf{t}^u$, $\mathbf{s}^u$, for $a,b \in [\mathbf{m}^u, \mathbf{s}^u]$ with $a<b$, the function $i \mapsto e^{-\beta n\Psi^u\left(a+\frac{2i}{n}\right)}\binom{n}{\frac{n}{2}(1+a)+i}$ is decreasing on $[\mathbf{m}^u,\mathbf{t}^u]$ and increasing on $[\mathbf{t}^u,\mathbf{s}^u]$. This settles the lower bound in (4.6).

Arguments similar to the ones above give

$$\begin{aligned}\mathrm{cap}^u(a,b) &\le \nu^u\left(\min(b,\mathbf{t}^u)-\tfrac{2}{n}\right) q^u\left(\min(b,\mathbf{t}^u)-\tfrac{2}{n}, \min(b,\mathbf{t}^u)\right)\\ &\le \frac{n(1-b)e^{2\beta(p+h^u)}}{2z^u} e^{-\beta n \Psi^u(\min(b,\mathbf{t}^u))}\binom{n}{\frac{n}{2}(1+\min(b,\mathbf{t}^u))},\end{aligned} \tag{4.15}$$

where for the first equality we use the test function $f \equiv 1$ on $\left[-1, \min(b,\mathbf{t}^u)-\frac{2}{n}\right]$ and $f\equiv 0$ on $[\min(b,\mathbf{t}^u),1]$ in (4.9). □

4.2. **Capacity bounds on** $\mathrm{ER}_n(p)$**.** Next we derive capacity bounds for $\{\xi_t\}_{t\ge 0}$ on $\mathrm{ER}_n(p)$. The proof is analogous to what was done in Lemma 4.1 for $\{\theta^u_t\}_{t\ge0}$ and $\{\theta^l_t\}_{t\ge0}$ on K_n.

Define the set of direct paths between $A \subseteq S_n$ and $B\subseteq S_n$ by

$$\mathcal{L}_{A,B} = \left\{\gamma = (\gamma_0,\dots,\gamma_{|\gamma|}\colon A \to B : |\gamma_{i+1}| = |\gamma_i| + 1 \text{ for all } \gamma_i \in \gamma\right\}, \tag{4.16}$$

which may be empty. Abbreviate $\theta_k = p(1-\frac{k}{n}) - h$.

Lemma 4.2 (**Capacity bounds for** $\{\xi_t\}_{t\ge0}$**).** *With* $\mathbb{P}_{\mathrm{ER}_n(p)}$*-probability tending to* 1 *as* $n\to\infty$ *the following is true. For every* $0 \le k < k' \le n$ *and every* $\varrho\colon \mathbb{N}\to\mathbb{R}_+$ *satisfying* $\lim_{n\to\infty}\varrho(n) = \infty$,

$$\begin{aligned}\mathrm{cap}(A_k, A_{k'}) &\le \frac{1}{Z} e^{-\beta H_n(\boxminus)} O\big(\varrho(n) n^{11/6}\big)\binom{n}{k_m} e^{-\beta\, 2k_m\theta_{k_m}},\\ \mathrm{cap}(A_k, A_{k'}) &\ge \frac{1}{Z} e^{-\beta H_n(\boxminus)}\,\Omega\left(n^{-1} e^{-\left(\beta+\frac{1}{\sqrt{3}}\right)\sqrt{\log n}}\right)\binom{n}{k_m} e^{-\beta\, 2k_m\theta_{k_m}},\end{aligned} \tag{4.17}$$

where

$$k_m = \mathrm{argmin}_{k\le j\le k'} \binom{n}{j} e^{-\beta\, 2j\theta_j}. \tag{4.18}$$

Proof. Recall from (4.1) and (4.3) that

$$\operatorname{cap}(A_k, A_{k'}) = \min_{f\in Q} \mathscr{E}(f,f) = \min_{f\in Q} \tfrac{1}{2} \sum_{\sigma,\xi\in S_n} \mu(\sigma)\, r(\sigma,\xi)\, [f(\sigma) - f(\xi)]^2, \tag{4.19}$$

where

$$Q(A_k, A_{k'}) = \left\{ f\colon\ S_n \to [0,1] :\ f|_{A_k} \equiv 1,\ f|_{A_{k'}} \equiv 0 \right\}. \tag{4.20}$$

The proof comes in 3 Steps.

1. We first prove the upper bound in (4.17). Let $B = \bigcup_{j=k}^{k_m-1} A_j$, and note that, by (1.7),

$$\begin{aligned}
\operatorname{cap}(A_k, A_{k'}) &\le \tfrac{1}{2} \sum_{\sigma,\xi\in S_n} \mu(\sigma)\, r(\sigma,\xi)\, [\mathbb{1}_B(\sigma) - \mathbb{1}_B(\xi)]^2 = \sum_{\sigma\in A_{k_m-1}} \sum_{\xi\in A_{k_m}} \mu(\sigma)\, r(\sigma,\xi) \\
&= \frac{1}{Z} \sum_{\sigma\in A_{k_m-1}} \sum_{\xi\in A_{k_m},\xi\sim\sigma} e^{-\beta\max\{H_n(\xi),H_n(\sigma)\}} \\
&= \frac{1}{Z} \left(\sum_{\sigma\in A_{k_m-1}} \sum_{\substack{\xi\in A_{k_m},\xi\sim\sigma \\ H_n(\sigma)\ge H_n(\xi)}} e^{-\beta H_n(\sigma)} + \sum_{\sigma\in A_{k_m-1}} \sum_{\substack{\xi\in A_{k_m},\xi\sim\sigma \\ H_n(\sigma)< H_n(\xi)}} e^{-\beta H_n(\xi)} \right) \\
&\le \frac{1}{Z} \max\{k_m, n-k_m\} \left(\sum_{\sigma\in A_{k_m-1}} e^{-\beta H_n(\sigma)} + \sum_{\xi\in A_{k_m}} e^{-\beta H_n(\xi)} \right). \qquad (4.21)
\end{aligned}$$

Recall from (3.5) that ϕ_i^k denotes the cardinality of the set of all $\sigma \in A_k$ with $|\partial_E \sigma| = pk(n-k) + i$. Note from (1.3) that for any $\xi \in A_{k_m}$ such that $|\partial_E \xi| = pk_m(n-k_m) + i$,

$$e^{-\beta H_n(\xi)} = e^{-\beta H_n(\boxminus)} e^{-\beta\,(2k_m\theta_{k_m} + \frac{2i}{n})}. \tag{4.22}$$

There are $\binom{n}{k_m}$ terms in the sum, and therefore we get

$$\begin{aligned}
\sum_{\xi\in A_{k_m}} e^{-\beta H_n(\xi)} &= e^{-\beta H_n(\boxminus)} \sum_{i=-pk_m(n-k_m)}^{(1-p)k_m(n-k_m)} \phi_i^{k_m} e^{-\beta\left(2k_m\theta_{k_m} + \frac{2i}{n}\right)} \\
&= e^{-\beta H_n(\boxminus)} \left(\sum_{i<-Y} \phi_i^{k_m} e^{-\beta\left(2k_m\theta_{k_m} + \frac{2i}{n}\right)} + \sum_{i\ge -Y} e^{-\beta\left(2k_m\theta_{k_m} + \frac{2i}{n}\right)} \right) \qquad (4.23) \\
&\le e^{-\beta H_n(\boxminus)} \left(\binom{n}{k_m} e^{-\beta(2k_m\theta_{k_m} - \frac{2Y}{n})} + \sum_{i<-Y} \phi_i^{k_m} e^{-\beta\left(2k_m\theta_{k_m} + \frac{2i}{n}\right)} \right)
\end{aligned}$$

with $Y = \sqrt{\log(\varrho(n)^2 n^{5/6}) k_m(n-k_m)}$. The choice of Y will become clear shortly. The summand in the right-hand side can be bounded as follows. By (3.3) in Lemma 3.2, the sum over $i < -Y$ can be restricted to $-cn^{3/2} \le i < -Y$, since with high probability no configuration has a boundary size that deviates by more than $cn^{3/2}$ from the mean. But, using Lemma 3.3, we can also bound from above the number of configurations that deviate by at most Y from the mean, i.e., we can bound $\phi_i^{k_m}$ for $-cn^{3/2} \le i < -Y$. Taking a union bound over $0 \le k \le n$ and $-cn^{3/2} \le i < -Y$, we get

$$\mathbb{P}\left[\bigcup_{k=0}^{n} \bigcup_{i=-cn^{3/2}}^{-Y} \left\{ \phi_i^{k_m} \ge \varrho(n)\, n^{5/2} \binom{n}{k_m} e^{-\frac{2i^2}{k_m(n-k_m)}} \right\} \right] \le \frac{1}{\varrho(n)}. \tag{4.24}$$

Thus, with $\mathbb{P}_{\mathrm{ER}_n(p)}$-probability at least $\geq 1 - \frac{1}{\varrho(n)}$,

$$\begin{aligned}\sum_{i<-Y} \phi_i^{k_m} e^{-\beta\left(2k_m\,\theta_{k_m} + \frac{2i}{n}\right)} &\leq \sum_{i>Y} \varrho(n)\, n^{5/2} \binom{n}{k_m} e^{-2i\left(\frac{i}{k_m(n-k_m)} - \frac{\beta}{n}\right)} e^{-\beta\, 2k_m\,\theta_{k_m}} \\ &\leq \varrho(n)\, n^{5/2} \binom{n}{k_m} e^{-\beta\, 2k_m\,\theta_{k_m}} e^{-2\log(\varrho(n) n^{5/6})} \\ &\leq \frac{n^{5/6}}{\varrho(n)} \binom{n}{k_m} e^{-\beta\, 2k_m\,\theta_{k_m}},\end{aligned} \tag{4.25}$$

where we use that, for $i > Y$ and n sufficiently large,

$$\tfrac{i}{k_m(n-k_m)} - \tfrac{\beta}{n} \geq \sqrt{\tfrac{\log(\varrho(n)^2 n^{5/6})}{k_m(n-k_m)}} - \tfrac{\beta}{n} \geq \sqrt{\tfrac{\log(\varrho(n) n^{5/6})}{k_m(n-k_m)}}. \tag{4.26}$$

The above inequality also clarifies our choice of Y. Substituting this into (4.24), we see that

$$\begin{aligned}\sum_{\xi\in A_{k_m}} e^{-\beta H_n(\xi)} &\leq [1+o_n(1)]\, e^{-\beta H_n(\boxminus)} e^{\frac{2\beta Y}{n}} e^{-\beta\, 2k_m\,\theta_{k_m}} \binom{n}{k_m} \\ &= O\!\left(\varrho(n) n^{5/6}\right) e^{-\beta H_n(\boxminus)} e^{-\beta\, 2k_m\theta_{k_m}} \binom{n}{k_m}.\end{aligned} \tag{4.27}$$

A similar bound holds for $\sum_{\xi\in A_{k_m-1}} e^{-\beta H_n(\xi)}$. A union bound over $1 \leq k_m \leq n$ increases the exponent $\frac{5}{6}$ to $\frac{11}{6}$. Together with (4.21), this proves the upper bound in (4.17).

2. We next derive a combinatorial bound that will be used later for the proof of the lower bound in (4.17). Note that if $f \in Q(A_k, A_{k'})$ and $\gamma \in \mathcal{L}_{A_k,A_{k'}}$ (recall (4.16)), then there must be some $1 \leq i \leq k'-k$ such that

$$|f(\gamma_i) - f(\gamma_{i+1})| \geq (k'-k)^{-1}. \tag{4.28}$$

A simple counting argument shows that

$$\left|\mathcal{L}_{A_k,A_{k'}}\right| = \binom{n}{k} \frac{(n-k)!}{(n-k')!}, \tag{4.29}$$

since for each $\sigma \in A_k$ there are $(n-k)\times(n-k-1)\times\cdots\times(n-k'+1)$ paths in $\mathcal{L}_{A_k,A_{k'}}$ from σ to $A_{k'}$. Let

$$b_i = \left|\left\{(\sigma,\xi) \in A_{k+i-1}\times A_{k+i}\colon\ |f(\sigma) - f(\xi)| \geq (k'-k)^{-1},\ \sigma\sim\xi\right\}\right|, \quad 1 \leq i \leq k'-k. \tag{4.30}$$

We claim that

$$\exists\, 1 \leq i_\blacklozenge \leq k'-k\colon \qquad b_{i_\blacklozenge} \geq \frac{k}{k'-k}\binom{n}{k+i_\blacklozenge}. \tag{4.31}$$

Indeed, the number of paths in $\mathcal{L}_{A_k,A_{k'}}$ that pass through $\sigma \in A_{k+i_\blacklozenge-1}$ followed by a move to $\xi \in A_{k+i_\blacklozenge}$ equals

$$z_{i_\blacklozenge} = \frac{(k+i_\blacklozenge-1)!}{k!} \times \frac{(n-k-i_\blacklozenge)!}{(n-k')!}, \tag{4.32}$$

where the first term in the product counts the number of paths from $\sigma \in A_{k+i_\blacklozenge+1}$ to A_k, while the second term counts the number of paths from $\xi \in A_{k+i_\blacklozenge}$ to $A_{k'}$. Thus, if (4.31) fails, then

$$\sum_{i=1}^{k'-k} b_i z_i < \frac{k}{k'-k} \sum_{i=1}^{k-k'} \frac{1}{k+i} \frac{n!}{k!\,(n-k')!} \leq \frac{n!}{k!\,(n-k')!} = \left|\mathcal{L}_{A_k,A_{k'}}\right|, \tag{4.33}$$

which in turn implies that (4.28) does not hold for some $\gamma \in \mathcal{L}_{A_k,A_{k'}}$ (use that $b_i z_i$ counts the paths that satisfy condition (4.28)), which is a contradiction. Hence the claim in (4.31) holds.

3. In this part we prove the lower bound in (4.17). By Lemma 3.3 we have that, with $\mathbb{P}_{\mathrm{ER}_n(p)}$-probability at least $1-\frac{1}{\varrho(n)}$, for any $Y \geq 0$,

$$\sum_{j \geq \sqrt{Y(k+i_\blacklozenge)(n-k-i_\blacklozenge)}} \phi_j^{k+i_\blacklozenge} \leq \varrho(n) \binom{n}{k+i_\blacklozenge} e^{-2Y}. \tag{4.34}$$

Picking $Y=\log(\varrho(n)\,k^{-1}2n^{3/2})$, we get that

$$\sum_{j \geq \sqrt{Y(k+i_\blacklozenge)(n-k-i_\blacklozenge)}} \phi_j^{k+i_\blacklozenge} \leq \frac{1}{4}\frac{k^2\varrho(n)}{\varrho(n)^2 n^3}\binom{n}{k+i_\blacklozenge} \leq \frac{1}{2}\frac{k}{n\,(k'-k)}\binom{n}{k+i_\blacklozenge}, \tag{4.35}$$

and so at least half of the configurations contributing to $b_{i_\blacklozenge}$ have an edge-boundary of size at most

$$p\,(k+i_\blacklozenge)\,(n-k-i_\blacklozenge)+\sqrt{Y\,(k+i_\blacklozenge)\,(n-k-i_\blacklozenge)}. \tag{4.36}$$

If $\xi \in A_{k+i_\blacklozenge}$ is such a configuration, then by Lemma 3.2 the same is true for any $\sigma \sim \xi$ (i.e., configurations differing at only one vertex), since

$$|\partial_E \sigma| \leq |\partial_E \xi| + \max_{v \in \sigma \triangle \xi} \deg(v) \leq |\partial_E \xi| + pn + o\left(\rho(n)\sqrt{n \log n}\right). \tag{4.37}$$

This implies

$$\begin{aligned} \mathscr{E}(f,f) &= \tfrac{1}{2}\sum_{\sigma,\xi \in S_n} \mu(\sigma)\, r(\sigma,\xi)\,[f(\sigma)-f(\xi)]^2 \\ &\geq \frac{1}{Zn^{2/3}} \sum_{\xi \in A_{k+i_\blacklozenge}} \sum_{\sigma \in A_{k+i_\blacklozenge -1}} e^{-\beta \max\{H_n(\sigma),H_n(\xi)\}} \\ &\geq e^{-\beta H_n(\boxminus)} \frac{k}{2Zn^2}\binom{n}{k+i_\blacklozenge} \exp\left(-\beta\left(2(k+i_\blacklozenge)\theta_{k+i_\blacklozenge} + 2\frac{\sqrt{Y(k+i_\blacklozenge)(n-k-i_\blacklozenge)}}{n}\right)\right). \end{aligned} \tag{4.38}$$

Therefore

$$\begin{aligned} \mathscr{E}(f,f) &\geq e^{-\beta H_n(\boxminus)} e^{-\beta\sqrt{Y}} \min_{1 \leq i \leq k'-k} \frac{k}{2Zn^2}\binom{n}{k+i} e^{-\beta(2(k+i)\theta_{k+i})} \\ &= e^{-\beta H_n(\boxminus)} e^{-\beta\sqrt{Y}} \frac{k}{2Zn^2}\binom{n}{k_m} e^{-\beta(2(k+i)\theta_{k-m})}. \end{aligned} \tag{4.39}$$

Since (4.39) is true for any $f \in Q(A_k, A_{k'})$, the lower bound in (4.17) follows, with k_m defined in (4.18). □

4.3. **Hitting probabilities on** $\mathrm{ER}_n(p)$**.** Let $\mu_{A_\mathbf{M}}$ be the equilibrium distribution μ conditioned to the set $A_\mathbf{M}$. Write $\mathbb{P}^l$ and $\mathbb{P}^u$ denote the laws of the processes $\{\xi_t^l\}_{t\geq 0}$ and $\{\xi_t^u\}_{t\geq 0}$, respectively.

Lemma 4.3 (Rank ordering of hitting probabilities). *With $\mathbb{P}_{\mathrm{ER}_n(p)}$-probability tending to 1 as $n \to \infty$,*

$$\max_{\xi \in A_{\mathbf{M}^l}} \mathbb{P}_\xi^l\left[\tau_{\mathbf{S}^l} < \tau_{\mathbf{M}^l}\right] \leq \mathbb{P}_{\mu_{A_\mathbf{M}}}\left[\tau_\mathbf{S} < \tau_\mathbf{M}\right] \leq \min_{\sigma \in A_{\mathbf{M}^u}} \mathbb{P}_\sigma^u\left[\tau_{\mathbf{S}^u} < \tau_{\mathbf{M}^u}\right]. \tag{4.40}$$

Proof. The proof comes in 3 Steps.

1. Recall from (1.10) that the magnetization of $\sigma \in A_k$ is $m(\sigma) = 2\frac{k}{n}-1$. We first observe that the maximum and the minimum in (4.40) are redundant, because by symmetry

$$\begin{aligned} \max_{\xi \in A_k} \mathbb{P}_\xi^l\left[\tau_{A_{k'}} < \tau_{A_k}\right] &= \min_{\xi \in A_k} \mathbb{P}_\xi^l\left[\tau_{A_{k'}} < \tau_{A_k}\right], \\ \min_{\xi \in A_k} \mathbb{P}_\xi^u\left[\tau_{A_{k'}} < \tau_{A_k}\right] &= \max_{\xi \in A_k} \mathbb{P}_\xi^u\left[\tau_{A_{k'}} < \tau_{A_k}\right]. \end{aligned} \tag{4.41}$$

Recall that $\{\xi_t^l\}_{t\geq 0}$ is the Markov process on S_n governed by the Hamiltonian H_n^l in (2.3), and that the associated magnetization process $\{\theta_t^l\}_{t\geq 0} = \{m(\xi_t^l)\}_{t\geq 0}$ is a Markov process on the set Γ_n in (1.15) with transition rates given by q^l in (2.11). Denoting by $\hat{\mathbb{P}}^l$ the law of $\{\theta_t^l\}_{t\geq 0}$, we get from (4.5) that for any $0 \leq k \leq k' < n$, and with $a = \frac{2k}{n} - 1$ and $b = \frac{2k'}{n} - 1$,

$$cap^l\left(a,b\right) = \sum_{u\in\Gamma_n} \nu^l\left(a\right) q^l\left(a,u\right) \hat{\mathbb{P}}_a^l\left[\tau_b < \tau_a\right] = \nu^l\left(a\right)\left[q^l\left(a,a+\tfrac{2}{n}\right) + q^l\left(a,a-\tfrac{2}{n}\right)\right]\hat{\mathbb{P}}_a^l\left[\tau_b<\tau_a\right], \tag{4.42}$$

and therefore

$$\max_{\xi\in A_k} \mathbb{P}_\xi^l\left[\tau_{A_{k'}} < \tau_{A_k}\right] = \hat{\mathbb{P}}_a^l\left[\tau_b<\tau_a\right] = \left[\nu^l\left(a\right)\left(q^l\left(a,a+\tfrac{2}{n}\right)+q^l\left(a,a-\tfrac{2}{n}\right)\right)\right]^{-1}\mathrm{cap}^l\left(a,b\right). \tag{4.43}$$

By (2.16), using the abbreviations

$$\Psi_1 = \max\left\{\Psi^l\left(a\right), \Psi^l\left(a+\tfrac{2}{n}\right)\right\}, \qquad \Psi_2 = \max\left\{\Psi^l\left(a\right), \Psi^l\left(a-\tfrac{2}{n}\right)\right\}, \tag{4.44}$$

we have, with the help of (4.12),

$$\begin{aligned}\nu^l\left(a\right)\left(q^l\left(a,a+\tfrac{2}{n}\right)+q^l\left(a,a-\tfrac{2}{n}\right)\right) &= \frac{1}{z^l}\frac{n}{2}\binom{n}{\frac{n}{2}\left(1+a\right)}\left(\left(1-a\right)e^{-\beta n\Psi_1} + \left(1+a\right)e^{-\beta n \Psi_2}\right) \\ &\geq \frac{1}{z^l} n e^{-2\beta\left(p|a|+h^l+\frac{p}{n}\right)} e^{-\beta n \Psi^l(a)}\binom{n}{\frac{n}{2}\left(1+a\right)}.\end{aligned} \tag{4.45}$$

From Lemma 4.1 we have that

$$\mathrm{cap}^l\left(a,b\right) \leq \tfrac{n(1-a)}{2z^l} e^{-\beta n \Psi^l(b)}\binom{n}{\frac{n}{2}\left(1+b\right)}. \tag{4.46}$$

Putting (4.43), (4.45) and (4.46) together, we get

$$\max_{\xi\in A_k}\mathbb{P}_\xi^l\left[\tau_{A_{k'}}<\tau_{A_k}\right] \leq \tfrac{(1-a)}{2} e^{2\beta\left(p|a|+h^l+\frac{p}{n}\right)} e^{-\beta n\left[\Psi^l(b)-\Psi^l(a)\right]}\binom{n}{\frac{n}{2}\left(1+b\right)}\binom{n}{\frac{n}{2}\left(1+a\right)}^{-1}. \tag{4.47}$$

Similarly, denoting by $\hat{\mathbb{P}}^u$ the law of $\{\theta_t^u\}_{t\geq 0}$, we have

$$\begin{aligned}\min_{\xi\in A_k}\mathbb{P}_\xi^u\left[\tau_{A_{k'}}<\tau_{A_k}\right] &= \hat{\mathbb{P}}_a^u\left[\tau_b<\tau_a\right] \\ &= \left[\nu^u\left(a\right)\left(q^u\left(a,a+\tfrac{2}{n}\right)+q^u\left(a,a-\tfrac{2}{n}\right)\right)\right]^{-1}\mathrm{cap}^u\left(a,b\right),\end{aligned} \tag{4.48}$$

where

$$\left[\nu^u\left(a\right)\left(q^u\left(a,a+\tfrac{2}{n}\right)+q^u\left(a,a-\tfrac{2}{n}\right)\right)\right]^{-1} \geq \left[\frac{n}{z^u} e^{2\beta\left(p|a|+h^l+\frac{p}{n}\right)} e^{-\beta n\Psi^u(a)}\binom{n}{\frac{n}{2}\left(1+a\right)}\right]^{-1}, \tag{4.49}$$

and, Lemma 4.1,

$$\mathrm{cap}^u\left(a,b\right) \geq \frac{1}{2nz^u} e^{-\beta n\Psi^u(b)}\binom{n}{\frac{n}{2}\left(1+b\right)}. \tag{4.50}$$

Putting (4.48)–(4.50) together, we get

$$\min_{\xi\in A_k}\mathbb{P}_\xi^u\left[\tau_{A_{k'}}<\tau_{A_k}\right] \geq \frac{1}{n} e^{-\beta n\left[\Psi^l(b)-\Psi^l(a)\right]}\binom{n}{\frac{n}{2}\left(1+b\right)}\binom{n}{\frac{n}{2}\left(1+a\right)}^{-1}. \tag{4.51}$$

2. Recall from (4.5) that

$$\mathrm{cap}\left(A_k, A_{k'}\right) = \sum_{\sigma\in A_k}\sum_{\xi\in S_n}\mu(\sigma)r(\sigma,\xi)\mathbb{P}_\sigma\left[\tau_{A_{k'}}<\tau_{A_k}\right]. \tag{4.52}$$

Split

$$\begin{aligned}\sum_{\sigma\in A_k}\sum_{\xi\in S_n}\mu(\sigma)r(\sigma,\xi) &= \sum_{\sigma\in A_k}\sum_{\xi\in A_{k+1}}\mu(\sigma)r(\sigma,\xi) + \sum_{\sigma\in A_k}\sum_{\xi\in A_{k-1}}\mu(\sigma)r(\sigma,\xi) \\ &= \frac{1}{Z}\sum_{\sigma\in A_k}\sum_{\xi\in A_{k+1}} e^{-\beta\max\{H_n(\sigma),H_n(\xi)\}} + \frac{1}{Z}\sum_{\xi\in A_k}\sum_{\xi' A_{k-1}} e^{-\beta\max\{H_n(\sigma),H_n(\xi)\}}.\end{aligned} \tag{4.53}$$

By Lemma 3.3 and a reasoning similar to that leading to (4.38),

$$
\begin{aligned}
\sum_{\xi\in A_k} e^{-\beta H_n(\xi)} &= e^{-\beta H_n(\boxminus)} \sum_{i=-pk(n-k)}^{(1-p)k(n-k)} \phi_i^k e^{-\beta\left(2k\theta_k+2\frac{i}{n}\right)} \\
&\geq \tfrac{1}{2}\binom{n}{k} e^{-\beta H_n(\boxminus)} e^{-\beta\left(2k\theta_k+2\frac{\sqrt{Yk(n-k)}}{n}\right)} \\
&\geq \tfrac{1}{2}\binom{n}{k} e^{-\beta H_n(\boxminus)} e^{-\beta\sqrt{\log(\sqrt{2\varrho(n)})}} e^{-\beta\, 2k\theta_k}
\end{aligned}
\tag{4.54}
$$

with $Y=\log(\sqrt{2\varrho(n)})$. Indeed, by (3.7) fewer than $\frac{1}{2}\binom{n}{k}$ configurations in A_k have an edge-boundary of size $\geq pk\,(n-k)+\sqrt{k\,(n-k)\,Y}$. Moreover, if $\xi\sim\xi'$, then, by Lemma 3.2,

$$
e^{\beta[H_n(\xi')-H_n(\xi)]} \leq [1+o(1)]\, e^{\beta(p+h)}, \tag{4.55}
$$

and since we may absorb this constant inside the error term $\varrho\,(n)$, we get that

$$
\sum_{\sigma\in A_k}\sum_{\xi\in A_{k+1}} e^{-\beta\max\{H_n(\sigma),H_n(\xi)\}} \geq e^{-\beta H_n(\boxminus)}\tfrac{1}{2}\,(n-k)\binom{n}{k} e^{-\frac{\beta}{2}\sqrt{\log\sqrt{2\varrho(n)}}} e^{-\beta\,2k\,\theta_k}, \tag{4.56}
$$

$$
\sum_{\sigma\in A_k}\sum_{\xi\in A_{k-1}} e^{-\beta\max\{H_n(\sigma),H_n(\xi)\}} \geq e^{-\beta H_n(\boxminus)}\tfrac{1}{2}k\binom{n}{k} e^{-\frac{\beta}{2}\sqrt{\log\sqrt{2\varrho(n)}}} e^{-\beta\,2k\,\theta_k}, \tag{4.57}
$$

and hence

$$
\sum_{\sigma\in A_k}\sum_{\xi\in S_n} \mu(\sigma) r(\sigma,\xi) \geq e^{-\beta H_n(\boxminus)}\frac{1}{2Z}\binom{n}{k} e^{-\frac{\beta}{2}\sqrt{\log\sqrt{2\varrho(n)}}} e^{-\beta\,2k\,\theta_k}. \tag{4.58}
$$

3. Similar bounds can be derived for $\mathbb{P}_{\mu_{A_k}}\left[\tau_{A_{k'}}<\tau_{A_k}\right]$. Indeed, note that, by Lemma (3.5), $r\,(\sigma,\xi)=1$ for all $\sigma\in A_k$ and all but $O(n^{2/3})$ many configurations $\xi\in S_n$. Therefore

$$
\begin{aligned}
\mathrm{cap}(A_k,A_{k'}) &= n\,[1+o(1)]\sum_{\sigma\in A_k}\mu(\sigma)\mathbb{P}_\sigma\left[\tau_{A_{k'}}<\tau_{A_k}\right] \\
&= n\,[1+o(1)]\,\mu(A_k)\,\mathbb{P}_{\mu_{A_k}}\left[\tau_{A_{k'}}<\tau_{A_k}\right]
\end{aligned}
\tag{4.59}
$$

and hence

$$
\mathbb{P}_{\mu_{A_k}}\left[\tau_{A_{k'}}<\tau_{A_k}\right] = [1+o(1)]\,\frac{\mathrm{cap}(A_k,A_{k'})}{n\mu(A_k)}. \tag{4.60}
$$

Note that

$$
\mu(A_k)=\frac{1}{Z}\sum_{\sigma\in A_k} e^{-\beta H_n(\sigma)}, \tag{4.61}
$$

and we have already produced bounds for a sum like (4.61) in Lemma 4.2. Referring to (4.28), we see that

$$
\mathbb{P}_{\mu_{A_k}}\left[\tau_{A_{k'}}<\tau_{A_k}\right] \leq [1+o(1)]\,\frac{\mathrm{cap}(A_k,A_{k'})}{\frac{1}{Z}\,e^{-(\beta+\frac{1}{\sqrt{3}})\sqrt{\log n}}\binom{n}{k}e^{-\beta 2k\theta_k}}, \tag{4.62}
$$

$$
\mathbb{P}_{\mu_{A_k}}\left[\tau_{A_{k'}}<\tau_{A_k}\right] \geq [1+o(1)]\,\frac{\mathrm{cap}(A_k,A_{k'}}{\frac{1}{Z}\,n^{17/6}e^{-(\beta+\frac{1}{\sqrt{3}})\sqrt{\log n}}\binom{n}{k}e^{-\beta 2k\theta_k}}. \tag{4.63}
$$

Finally, we note that if we let $\Delta_h=h-h^u$, then

$$
\frac{e^{-\beta n[\Psi^u(\mathbf{s}^u)-\Psi^u(\mathbf{m}^u)]}}{e^{-\beta n[\Psi(\mathbf{s})-\Psi(\mathbf{m})]}} = e^{\beta n C_{\beta,h,p}\Delta_h}, \tag{4.64}
$$

where $C_{\beta,h,p}$ is a constant that depends on the parameters β, p and h. A similar expression follows for the ratio

$$
\binom{n}{\frac{n}{2}(1+\mathbf{s}^u)}\binom{n}{\frac{n}{2}(1+\mathbf{m}^u)}^{-1}\left[\binom{n}{\frac{n}{2}(1+\mathbf{s})}\binom{n}{\frac{n}{2}(1+\mathbf{m}}^{-1}\right]^{-1}. \tag{4.65}
$$

From this the statement of the lemma follows. ☐

5. Invariance under initial states and refined capacity estimates

In this section we use Lemma 4.3 to control the time it takes $\{m(\xi_t)\}_{t\geq 0}$ to cross the interval $[\mathbf{t}^u,\mathbf{s}^u]\cap[\mathbf{t}^l,\mathbf{s}^l]$, which will be a good indicator of the time it takes $\{\xi_t\}_{t\geq 0}$ to reach the basin of the stable state $\mathbf{s}$. In particular, our aim is to control this time by comparing it with the time it takes $\{\theta^u_t\}_{t\geq 0}$ and $\{\theta^l_t\}_{t\geq 0}$ defined in (2.9) to do the same for $\mathbf{s}^u$ and $\mathbf{s}^{(l}$. In Section 5.1 we derive bounds on the probability of certain rare events for the dynamics on $\mathrm{ER}_n(p)$ (Lemmas 5.1–5.6 below). In Section 5.2 we use these bounds to prove Lemma 2.2.

5.1. **Estimates for rare events.** The aim of this section is to show that, for $\{\xi_t\}_{t\geq 0}$ starting in $A_{\mathbf{T}}$, the exact starting configuration is irrelevant for the metastable crossover time because it is very large. We will do this by showing that "local mixing" takes place long before the crossover occurs. More precisely, we will show that if $\xi,\tilde{\xi}_0$ is any two initial configurations in $A_{\mathbf{T}}$, then there is a coupling such that the trajectory $t\mapsto\xi_t$ intersects the trajectory $t\mapsto\tilde{\xi}_t$ well before either strays too far from $A_{\mathbf{T}}$. The idea of the proof is to show that there is a small but sufficiently large overlap between the distributions of ξ_t and $\tilde{\xi}_t$ once every spin at every vertex has had a chance to update (the meaning of this will become clear in the sequel). By treating this probability as the chance of success for a geometric random variable, it follows that, after a large number of trials, with high probability the trajectories intersect.

We begin by deriving upper and lower bounds on the number of jumps $N_\xi\left(t\right)$ taken by the process $\{\xi_t\}_{t\geq 0}$ up to time t. By Lemma 3.4, the jump rate from any $\sigma\in S_n$ is bounded by

$$n\,e^{-2\beta(p+h)}\leq\sum_{\sigma'\in S_n}r\left(\sigma,\sigma'\right)\leq n. \tag{5.1}$$

Hence $N_\xi\left(t\right)$ can be stochastically bounded from above by a Poisson random variable with parameter tn, and from below by a Poisson random variable with parameter $tne^{(-2\beta(p+h)}$. It therefore follows that, for any $M\geq 0$,

$$\begin{aligned}&\mathbb{P}\left[N_\xi\left(t\right)\geq M\right]\leq\chi_M(nt),\\&\mathbb{P}\left[N_\xi\left(t\right)< M\right]\leq 1-\chi_M\big(nt\,e^{-2\beta(p+h)}\big),\end{aligned} \tag{5.2}$$

where we abbreviate $\chi_M(u)=e^{-u}\sum_{k\geq M}u^k/k!$, $u\in\mathbb{R}$, $M\in\mathbb{N}$.

5.1.1. *Localisation.* The purpose of the next lemma is to show that the probability of $\{\xi_t\}_{t\geq 0}$ straying too far from $A_{\mathbf{M}}$ within its first $n^2\log n$ jumps is very small. The seemingly arbitrary choice of $n^2\log n$ is in fact related to the Coupon Collector's problem (see (5.29) below).

Lemma 5.1 (Localisation). *Let $\xi_0\in A_{\mathbf{M}}$, $T=\inf\left\{t\geq 0\colon N_\xi\left(t\right)\geq n^2\log n\right\}$, and let $C_1\in\mathbb{R}$ be a sufficiently large constant, possibly dependent on p and h (but not on n). Then*

$$\mathbb{P}_{\xi_0}\left[\xi_t\in A_{\mathbf{M}+C_1n^{5/6}}\text{ for some }0\leq t\leq T\right]\leq e^{-n^{2/3}}. \tag{5.3}$$

Proof. The idea of the proof is to show that $\{\xi_t\}_{t\geq 0}$ returns many times to $A_{\mathbf{M}}$ before reaching $A_{\mathbf{M}+C_1n^{5/6}}$. The proof come sin 3 Steps.

1. We begin by showing that $T\leq n^2\log n$ with probability $\geq 1-e^{-n^3}$. Indeed, by the second line of (5.2),

$$\begin{aligned}&\mathbb{P}\left[T>n^2\log n\right]\\&\quad=\mathbb{P}\left[N_\xi\left(n^2\log n\right)<n^2\log n\right]\\&\quad\leq 1-\chi_{n^2\log n}\Big(\left(n^3\log n\right)e^{-2\beta(p+h)}\Big)\\&\quad\leq\sum_{k=0}^{n^2\log n}\exp\left(-(n^3\log n)\,e^{-2\beta(p+h)}+k\log\left(\tfrac{e\,n^3\log n}{k}\right)\right)\\&\quad\leq(n^2\log n)\exp\left(-(n^3\log n)\,e^{-2\beta(p+h)}+n^{5/2}\right)\\&\quad\leq e^{-n^3},\end{aligned} \tag{5.4}$$

where for the second inequality we use that $k! \geq (\frac{k}{e})^k$, $k \in \mathbb{N}$, and for the third inequality that, for n sufficiently large,

$$k \log\left(\tfrac{en^3 \log n}{k}\right) \leq (n^2 \log n) \log\left(e\, n^3 \log n\right) \leq n^{5/2}. \tag{5.5}$$

Next, observe that

$$\begin{aligned} &\mathbb{P}_{\xi_0}\left[\xi_t \in A_{\mathbf{M}+C_1 n^{5/6}} \text{ for some } 0 \leq t \leq T\right] \\ &= \mathbb{P}_{\xi_0}\left[\xi_t \in A_{\mathbf{M}+C_1 n^{5/6}} \text{ for some } 0 \leq t \leq T,\, T \leq n^2 \log n\right] \\ &\quad + \mathbb{P}_{\xi_0}\left[\xi_t \in A_{\mathbf{M}+C_1 n^{5/6}} \text{ for some } 0 \leq t \leq T,\, T > n^2 \log n\right] \\ &\leq (n^2 \log n) \max_{\sigma \in A_{\mathbf{M}}} \mathbb{P}_\sigma\left[\tau_{A_{\mathbf{M}+C_1 n^{5/6}}} < \tau_{A_{\mathbf{M}}}\right] + e^{-n^3}. \end{aligned} \tag{5.6}$$

Here, the inequality follows from (5.4) and the observation that the event $\xi_t \in A_{\mathbf{M}+C_1 n^{5/6}}$ for some $0 \leq t \leq T$ with $T \leq n^2 \log n$ is contained in the event that $A_{\mathbf{M}+C_1 n^{5/6}}$ is visited before the $(n^2 \log n)$-th return to $A_{\mathbf{M}}$. From Lemma 4.3 and (4.47) it follows that

$$\max_{\sigma \in A_{\mathbf{M}}} \mathbb{P}_\sigma\left[\tau_{A_{\mathbf{m}+C_1 n^{5/6}}} < \tau_{A_{\mathbf{M}}}\right] \leq \tfrac{(1-a)}{2} e^{2\beta\left(p|a|+h^l+\frac{p}{n}\right)} e^{-\beta n[\Psi(b)-\Psi(a)]} \binom{n}{\frac{n}{2}(1+b)} \binom{n}{\frac{n}{2}(1+a)}^{-1} \tag{5.7}$$

with $a = \mathbf{m}/n$ and $b = \left(\mathbf{m} + C_1 n^{5/6}\right)/n$.

2. Our assumption on the parameters β, p and h is that $2\beta(p(a+\frac{2}{n})+h)+\log(\frac{1-a}{1+a+\frac{2}{n}})$ is negative in two disjoint regions. Recall that the first region lies between $a_1 = \frac{2\mathbf{m}}{n} - 1$ and $a_2 = \frac{2\chi_2}{n} - 1$, where

$$\begin{aligned} \mathbf{m} &= \min_{a \in \Gamma_n} \left\{2\beta\left(p\left(a+\tfrac{2}{n}\right)+h\right) + \log\left(\tfrac{1-a}{1+a+\frac{2}{n}}\right) \leq 0\right\}, \\ \chi_2 &= \min_{a \in \Gamma_n} \left\{2\beta\left(p\left(a+\tfrac{2}{n}\right)+h\right) + \log\left(\tfrac{1-a}{1+a+\frac{2}{n}}\right) \leq 0\right\}. \end{aligned} \tag{5.8}$$

This, in particular, implies that the derivative of $2\beta(p(a+\frac{2}{n})+h)+\log(\frac{1-a}{1+a+\frac{2}{n}})$ at $a = a_1$ is

$$2\beta p - \tfrac{1}{1-a_1} - \tfrac{1}{1+a_1} = -\delta_1 < 0 \tag{5.9}$$

for some $\delta_1 > 0$. Recall that $\Psi(a) = -\frac{p}{2}a^2 - ha$, so that $\Psi(b)-\Psi(a) = (a-b)(\frac{p}{2}(a+b)+h)$, which gives

$$\begin{aligned} &e^{-\beta n[\Psi(b)-\Psi(a)]} \binom{n}{\frac{n}{2}(1+b)} \binom{n}{\frac{n}{2}(1+a)}^{-1} \\ &= \exp\left(\beta n\,(b-a)\,(pa+h) + \beta n\,(b-a)^2\, \tfrac{p}{2} + \tfrac{n}{2}\log\left(\tfrac{(1+a)^{(1+a)}(1-a)^{(1-a)}}{(1+b)^{(1+b)}(1-b)^{(1-b)}}\right) + O(\log n)\right), \end{aligned} \tag{5.10}$$

where we use Stirling's approximation in the last line. Since $b = a + C_1 n^{-1/6}$, we have

$$\text{r.h.s. } (5.10) = \exp\left(\beta C_1 n^{5/6}\,(pa+h) + \tfrac{p}{2}\beta C_1^2 n^{2/3} + \tfrac{n}{2}\log F\right) \tag{5.11}$$

with

$$F = (1-U_n(a))^{1+a}\,(1+V_n(a))^{1-a}\,(W_n(a))^{C_1 n^{-1/6}}, \tag{5.12}$$

where

$$U_n(a) = \tfrac{C_1 n^{-1/6}}{1+a+C_1 n^{-1/6}}, \qquad V_n(a) = \tfrac{C_1 n^{-1/6}}{1-a-C_1 n^{-1/6}}. \tag{5.13}$$

From the Taylor series expansion of $\log(1+x)$ for $0 \leq |x| < 1$, we obtain

$$\begin{aligned} \tfrac{n}{2}(1+a)\log(1-U_n(a)) &\leq \tfrac{n}{2}(1+a)\left(-U_n(a) - \tfrac{1}{2}(U_n(a))^2\right), \\ \tfrac{n}{2}(1-a)\log(1+V_n(a)) &\leq \tfrac{n}{2}(1-a)\left(V_n(a) - \tfrac{1}{2}(V_n(a))^2 + O(n^{-1/2})\right), \end{aligned} \tag{5.14}$$

and

$$\begin{aligned} \tfrac{1}{2}C_1 n^{5/6}\log\left(\frac{U_n(a)}{V_n(a)}\right) &= \tfrac{1}{2}C_1 n^{5/6}\log\left(\tfrac{1-a}{1+a}\,\tfrac{1-a-C_1 n^{-1/6}}{1-a}\,\tfrac{1+a}{1+a+C_1 n^{-1/6}}\right) \\ &\leq \tfrac{1}{2}C_1 n^{5/6}\left(\log\left(\tfrac{1-a}{1+a}\right) - \tfrac{C_1 n^{-1/6}}{1-a} - U_n(a) - O(n^{-2/3})\right). \end{aligned} \tag{5.15}$$

By the definition of $\mathbf{m}$, we have

$$C_1 n^{5/6}\left(\beta\,(pa+h)+\log\left(\tfrac{1-a}{1+a}\right)\right)\leq 0. \tag{5.16}$$

Hence we get

$$\begin{aligned}&\beta C_1 n^{5/6}\,(pa+h)+\tfrac{p}{2}\beta n^{2/3}C_1^2+\tfrac{n}{2}\log F\\ &\leq \tfrac{p}{2}\beta n^{2/3}C_1^2-\tfrac{\frac{1}{2}C_1(1+a)n^{5/6}}{1+a+C_1n^{-1/6}}+\tfrac{\frac{1}{2}C_1(1-a)n^{5/6}}{1-a-C_1n^{-1/6}}-\tfrac{1}{2}C_1^2n^{2/3}\,G\end{aligned} \tag{5.17}$$

with

$$G=\tfrac{1}{1-a}+\tfrac{1}{1+a+C_1n^{-1/6}}+\left(\tfrac{1-a}{2}\right)\left(\tfrac{1}{1-a-C_1n^{-1/6}}\right)^2+\left(\tfrac{1+a}{2}\right)\left(\tfrac{1}{1+a+C_1n^{-1/6}}\right)^2. \tag{5.18}$$

Hence

$$\begin{aligned}\text{r.h.s. (5.17)}\;&\leq\;\tfrac{p}{2}\beta n^{2/3}C_1^2+\tfrac{1}{2}C_1^2n^{2/3}\left(\tfrac{1}{1-a-cn^{-1/6}}+\tfrac{1}{1+a+C_1n^{-1/6}}\right)-\tfrac{1}{2}C_1^2n^{1/6}\,G\\ &\leq\; n^{2/3}\tfrac{1}{2}C_1^2\left(p\beta-\tfrac{1}{2}\tfrac{1}{1-a-C_1n^{-1/6}}-\tfrac{1}{2}\tfrac{1}{1+a+C_1n^{-1/6}}+O(n^{-1/6})\right)\\ &=\; n^{2/3}\tfrac{1}{2}C_1^2\left(p\beta-\tfrac{1}{2}\tfrac{1}{1-a}-\tfrac{1}{2}\tfrac{1}{1+a}+O(n^{-1/3})\right)\leq-\tfrac{1}{4}C_1^2\delta_1n^{2/3}.\end{aligned} \tag{5.19}$$

3. Combine (5.7), (5.10) and (5.17), and pick C_1 large enough, to get the claim in (5.3). □

5.1.2. *Update times.* The following two lemmas give useful bounds for the coupling scheme. The symbol $\simeq$ stands for equality in distribution.

Lemma 5.2 **(Total variation between exponential distributions).**
Let $X\simeq\mathrm{Exp}\,(\lambda)$ and $Y\simeq\mathrm{Exp}\,(\lambda+\delta)$. Then the total variation distance between the distributions of X and Y is bounded by

$$d_{TV}\,(X,Y)\leq\frac{2\delta}{\lambda+\delta}. \tag{5.20}$$

Proof. Elementary. □

Lemma 5.3 **(Update times).** *Let T^{ξ}_{update} be the first time $\{\xi_t\}_{t\geq0}$ has experienced an update at every site:*

$$T^{\xi}_{update}=\inf\{t\geq0\colon\ \forall\,v\in V\,\exists\,0\leq s\leq t\colon\ \xi_s\,(v)=-\xi_0\,(v)\}\,. \tag{5.21}$$

Then, for any $y>0$,

$$\mathbb{P}\left[T^{\xi}_{update}\geq y\right]\leq\frac{\exp\left(-\lambda y+\log n\right)}{1-\exp\left(-\lambda y\right)},\qquad \lambda=e^{-\beta(2p+h)}. \tag{5.22}$$

Proof. Recall that for $\sigma\in S_n$ and $v\in V$, σ^v denotes the configuration satisfying $\sigma^v\,(w)=\sigma\,(w)$ for $w\neq v$, and $\sigma^v\,(v)=-\sigma\,(v)$. From (1.3) and (1.6) it follows that

$$r\,(\sigma,\sigma^v)\geq\lambda, \tag{5.23}$$

and so T^{ξ}_{update} is dominated by the maximum of n i.i.d. $\mathrm{Exp}\,(\lambda)$ random variables. Therefore

$$\begin{aligned}\mathbb{P}\left[T^{\xi}_{update}\leq y\right]\;&\geq\;\left(1-e^{-\lambda y}\right)^n=\exp\left(n\log(1-e^{-\lambda y})\right)\\ &\geq\;\exp\left(-\tfrac{ne^{-\lambda y}}{1-e^{-\lambda y}}\right)\geq1-\tfrac{ne^{-\lambda y}}{1-e^{-\lambda y}},\end{aligned} \tag{5.24}$$

which proves the claim. □

5.1.3. *Returns.* The next lemma establishes a lower bound on the number of returns to $A_{\mathbf{M}}$ before reaching $A_{\mathbf{S}}$. Let $g_{\xi_0}(A_{\mathbf{M}}, A_{\mathbf{S}})$ denote the number of jumps that $\{\xi_t\}_{t\geq 0}$ makes into the set $A_{\mathbf{M}}$ before reaching $A_{\mathbf{S}}$. More precisely, let $\{s_i\}_{i\in\mathbb{N}_0}$ denote the jump times of the process $\{\xi_t\}_{t\geq 0}$, i.e., $s_0 = 0$ and

$$s_i = \inf\left\{s > s_{i-1} \colon\ \xi_s \neq \xi_{s_{i-1}}\right\}, \tag{5.25}$$

and define for the process ξ_t commencing at ξ_0

$$g_{\xi_0}(A_{\mathbf{M}}, A_{\mathbf{S}}) = \left|\{i \in \mathbb{N}_0 \colon\ \xi_{s_i} \in A_{\mathbf{M}}, \xi_s \notin A_{\mathbf{S}}\ \forall\, s \leq s_i\}\right|. \tag{5.26}$$

Lemma 5.4 (Bound on number of returns). *For any $\xi_0 \in A_{\mathbf{M}}$ and any $\delta > 0$,*

$$\mathbb{P}_{\xi_0}\left[g_{\xi_0}(A_{\mathbf{M}}, A_{\mathbf{S}}) < e^{([R_p(t)-R_p(m)]-\delta)n}\right] \leq e^{-\delta n + Cn^{2/3}} \tag{5.27}$$

for some constant C that does not depend on n.

Proof. Let Y be a geometric random variable with probability of success given by $e^{-(R_p(t)-R_p(m))n+Cn^{2/3}}$. Then, by Lemma 4.3, every time the process $\{\xi_t\}_{t\geq 0}$ starts over from A_M, it has a probability less than $\mathbb{P}^u_\xi[\tau_{S^u} < \tau_{M^u}]$ of making it to A_S. Using the bounds from that lemma, it follows that Y is stochastically dominated by $g_{\xi_0}(A_{\mathbf{M}}, A_{\mathbf{T}})$. Hence

$$\mathbb{P}\left[Y \leq e^{([R_p(t)-R_p(m)]-\delta)n}\right] \leq e^{([R_p(t)-R_p(m)]-\delta)n}\, e^{-[R_p(t)-R_p(m)]n+Cn^{2/3}} \leq e^{-\delta n + Cn^{2/3}}. \tag{5.28}$$

□

Lemma 5.5 (Coupon collector). *The time S it takes a coupon collector to collect all of n coupons satisfies*

$$\mathbb{P}\left[S \geq n^2 \log n\right] \leq ne^{-n\log n}. \tag{5.29}$$

Proof. Define

$$s_v = \max_{t\leq S}\{\text{vertex } v \text{ is updated}\}, \tag{5.30}$$

and let

$$p_{in} = \mathbb{P}\left[\mathbf{m} - 10n^{2/3} \leq |\xi_{x_v}| \leq \mathbf{m} + 10n^{2/3}\right]. \tag{5.31}$$

Then

$$\frac{\mathbb{P}\left[\xi_S = \sigma\right]}{\mathbb{P}[\tilde{\xi}_{\tilde{S}} = \sigma]} \geq p_{in}^{2n}\left(1 - \frac{c}{n^{1/3}}\right)^n \geq e^{-Cn^{2/3}}. \tag{5.32}$$

□

Recall from (5.25) that s_i is the i^{th} configuration visited by $\{\xi_t\}_{t\geq 0}$. Let $\xi_0 = x$ and $\tilde{\xi}_0 = y$. For $x, y \in A_{\mathbf{M}}$, let

$$R = \min\{t \geq 0 \colon\ \text{in } \xi_t \text{ every vertex has been refreshed by time } t\}, \tag{5.33}$$

and similarly for $\tilde{R}$. Then, for any $\sigma \in S_n$,

$$\frac{\mathbb{P}[\xi_S = \sigma]}{\mathbb{P}[\tilde{\xi}_{\tilde{S}} = \sigma]} \geq e^{-Cn^{2/3}}. \tag{5.34}$$

5.1.4. *Coupling estimate.* We close with the following observation.

Lemma 5.6 (Coupling estimate). *With $\mathbb{P}_{\mathrm{ER}_n(p)}$-probability tending to 1 as $n \to \infty$ there is a coupling such that, for some $0 \leq s, t \leq T_0$ and any $\xi_0, \tilde{\xi}_0$ and $\delta > 0$,*

$$\mathbb{P}\left[\xi_s \neq \tilde{\xi}_t\right] \leq e^{-\Gamma^\star n + \delta n}. \tag{5.35}$$

Proof. Immediate from Lemmas 5.1, 5.4 and 5.5. □

5.2. **Uniform hitting time.** In this section we prove Lemma 2.2.

Proof. Consider two copies of the process, $\{\xi_t\}_{t\geq 0}$ and $\{\tilde{\xi}_t\}_{t\geq 0}$. Let $T_0 = e^{(\Gamma^\star-\delta)n}$, and for $A\subseteq S$ and $s\geq 0$ define

$$\tau^{\xi}_{s,A} = \inf\{t\geq s\colon \xi_t\in A\} \tag{5.36}$$

and similarly for $\tau^{\tilde{\xi}}_{s,A}$. We know from Lemma 5.6 that

$$\max_{\sigma\in A_{\mathbf{M}}} \mathbb{E}_\sigma\left[\tau_{A_{\mathbf{M}}}\right] \leq T_0. \tag{5.37}$$

Let $\{\xi_t,\tilde{\xi}_t\}_{t\geq 0}$ be the coupling of the two processes described above, and note that

$$\mathbb{E}_{\xi_0}\left[\tau_{A_{\mathbf{T}}}\right] = \hat{\mathbb{E}}_{(\xi_0,\tilde{\xi}_0)}[\tau^{\xi}_{0,A_{\mathbf{T}}}] = \hat{\mathbb{E}}_{(\xi_0,\tilde{\xi}_0)}\left[\tau^{\xi}_{0,A_{\mathbf{T}}}\mathbb{1}_{\{\xi_{T_0}=\tilde{\xi}_{T_0}\}}\right] + \hat{\mathbb{E}}_{(\xi_0,\tilde{\xi}_0)}\left[\tau^{\xi}_{0,A_{\mathbf{T}}}\mathbb{1}_{\{\xi_{T_0}\neq\tilde{\xi}_{T_0}\}}\right], \tag{5.38}$$

where $\hat{\mathbb{E}}$ denotes expectation with respect to the law of the joint process. Note that

$$\tau^{\xi}_{0,A_{\mathbf{T}}}\mathbb{1}_{\{\xi_{T_0}=\tilde{\xi}_{T_0}\}} \leq \tau^{\tilde{\xi}}_{0,A_{\mathbf{T}}}\mathbb{1}_{\left\{\xi_{T_0}=\tilde{\xi}_{T_0},\,\tau^{\tilde{\xi}}_{0,A_{\mathbf{T}}}\geq T_0\right\}} + \tau^{\xi}_{0,A_{\mathbf{T}}}\mathbb{1}_{\left\{\xi_{T_0}=\tilde{\xi}_{T_0},\,\tau^{\tilde{\xi}}_{0,A_{\mathbf{T}}}<T_0\right\}}, \tag{5.39}$$

and that

$$\begin{aligned}
&\tau^{\xi}_{0,A_{\mathbf{T}}}\mathbb{1}_{\left\{\xi_{T_0}=\tilde{\xi}_{T_0},\,\tau^{\tilde{\xi}}_{0,A_{\mathbf{T}}}<T_0\right\}}\\
&\quad= \tau^{\xi}_{0,A_{\mathbf{T}}}\mathbb{1}_{\left\{\xi_{T_0}=\tilde{\xi}_{T_0},\,\tau^{\tilde{\xi}}_{0,A_{\mathbf{T}}}<T_0,\,|\tilde{\xi}_{T_0}|<\mathbf{T}\right\}} + \tau^{\xi}_{0,A_{\mathbf{T}}}\mathbb{1}_{\left\{\xi_{T_0}=\tilde{\xi}_{T_0},\,\tau^{\tilde{\xi}}_{0,A_{\mathbf{T}}}<T_0,\,|\tilde{\xi}_{T_0}|\geq\mathbf{T}\right\}}\\
&\quad\leq \tau^{\xi}_{0,A_{\mathbf{T}}}\mathbb{1}_{\left\{\xi_{T_0}=\tilde{\xi}_{T_0},\,\tau^{\tilde{\xi}}_{0,A_{\mathbf{T}}}<T_0,\,|\tilde{\xi}_{T_0}|<\mathbf{T}\right\}} + T_0.
\end{aligned} \tag{5.40}$$

Also note from the definition of the coupling that, for any $\sigma\in S$ and any $A\subseteq S$, $\hat{\mathbb{E}}_{(\sigma,\sigma)}[\tau^{\xi}_{A}] = \mathbb{E}_\sigma[\tau_A]$ because the two trajectories merge when they start from the same site. Hence

$$\begin{aligned}
&\hat{\mathbb{E}}_{(\xi_0,\tilde{\xi}_0)}\left[\tau^{\xi}_{0,A_{\mathbf{T}}}\mathbb{1}_{\left\{\xi_{T_0}=\tilde{\xi}_{T_0},\,\tau^{\tilde{\xi}}_{0,A_{\mathbf{T}}}<T_0,\,|\tilde{\xi}_{T_0}|<\mathbf{T}\right\}}\right]\\
&\quad= \sum_{\sigma\in\bigcup_{i<\chi_1}A_i} \hat{\mathbb{E}}_{(\xi_0,\tilde{\xi}_0)}\left[\tau^{\xi}_{0,A_{\mathbf{T}}}\mathbb{1}_{\left\{\xi_{T_0}=\tilde{\xi}_{T_0}=\sigma,\,\tau^{\tilde{\xi}}_{0,A_{\mathbf{T}}}<T_0\right\}}\right]\\
&\quad\leq \sum_{\sigma\in\bigcup_{i<\mathbf{t}}A_i} \hat{\mathbb{E}}_{(\sigma,\sigma)}\left[\tau^{\xi}_{T_0,A_{\mathbf{T}}}\right]\hat{\mathbb{P}}_{(\xi_0,\tilde{\xi}_0)}\left[\mathbb{1}_{\left\{\xi_{T_0}=\tilde{\xi}_{T_0}=\sigma,\,\tau_{0,A_{\mathbf{T}}}<T_0\right\}}\right]\\
&\quad= \sum_{\sigma\in\bigcup_{i<\chi_1}A_i}\left(\hat{\mathbb{E}}_{(\sigma,\sigma)}\left[\tau^{\xi}_{A_{\mathbf{T}}}\right]+T_0\right)\hat{\mathbb{P}}_{(\xi_0,\tilde{\xi}_0)}\left[\xi_{T_0}=\tilde{\xi}_{T_0}=\sigma,\tau^{\tilde{\xi}}_{A_{\mathbf{T}}}<T_0\right]\\
&\quad\leq \left(T_0+\max_{\sigma\in\bigcup_{i<\chi_1}A_i}\mathbb{E}_\sigma\left[\tau^{\sigma}_{A_{\mathbf{T}}}\right]\right)\mathbb{P}_{\tilde{\xi}_0}\left[\tau_{A_{\mathbf{T}}}<T_0\right],
\end{aligned} \tag{5.41}$$

where we use the Markov property. Similarly, observe that

$$\begin{aligned}
\tau^{\xi}_{0,A_{\mathbf{T}}}\mathbb{1}_{\{\xi_{T_0}\neq\tilde{\xi}_{T_0}\}} &= \tau^{\xi}_{0,A_{\mathbf{T}}}\mathbb{1}_{\left\{\xi_{T_0}\neq\tilde{\xi}_{T_0},\,\tau^{\xi}_{0,A_{\mathbf{T}}}\leq T_0\right\}} + \tau^{\xi}_{0,A_{\mathbf{T}}}\mathbb{1}_{\left\{\xi_{T_0}\neq\tilde{\xi}_{T_0},\,\tau^{\xi}_{0,A_{\mathbf{T}}}>T_0\right\}}\\
&\leq T_0 + \tau^{\xi_{T_0}}_{0,A_{\mathbf{T}}}\mathbb{1}_{\left\{\xi_{T_0}\neq\tilde{\xi}_{T_0}\,\tau^{\xi}_{0,A_{\mathbf{T}}}>T_0\right\}},
\end{aligned} \tag{5.42}$$

and

$$\begin{aligned}
&\hat{\mathbb{E}}_{(\xi_0,\tilde{\xi}_0)}\left[\tau^{\xi}_{T_0,A_{\mathbf{T}}}\mathbb{1}_{\left\{\xi_{T_0}\neq\tilde{\xi}_{T_0},\,\tau^{\xi}_{0,A_{\mathbf{T}}}>T_0\right\}}\right]\\
&\quad= \sum_{\sigma\colon|\sigma|<A_{\mathbf{T}}}\sum_{\sigma'\neq\sigma}\hat{\mathbb{E}}_{(\xi_0,\tilde{\xi}_0)}\left[\tau^{\xi}_{T_0,A_{\mathbf{T}}}\mathbb{1}_{\left\{\xi_{T_0}=\sigma,\,\tilde{\xi}_{T_0}=\sigma',\,\tau^{\xi}_{0,A_{\mathbf{T}}}>T_0\right\}}\right]\\
&\quad= \sum_{\sigma\colon|\sigma|<\mathbf{t}}\sum_{\sigma'\neq\sigma}\hat{\mathbb{E}}_{(\sigma,\sigma')}\left[T_0+\tau^{\xi}_{0,A_{\mathbf{T}}}\right]\hat{\mathbb{E}}_{(\xi_0,\tilde{\xi}_0)}\left[\mathbb{1}_{\left\{\xi_{T_0}=\sigma,\,\tilde{\xi}_{T_0}=\sigma',\,\tau^{\xi}_{0,A_{\mathbf{T}}}>T_0\right\}}\right]\\
&\quad\leq \max_{\sigma\in\bigcup_{i<\mathbf{t}}A_i}\left(T_0+\mathbb{E}_\sigma\left[\tau^{\xi}_{0,A_{\mathbf{T}}}\right]\right)\hat{\mathbb{P}}_{(\xi_0,\tilde{\xi}_0)}\left[\xi_{T_0}\neq\tilde{\xi_{T_0}}\right].
\end{aligned} \tag{5.43}$$

Thus, (5.38) becomes

$$\mathbb{E}_{\xi_0}[\tau_{A_{\mathbf{T}}}] \le 2T_0 + \mathbb{E}_{\tilde{\xi}_0}[\tau_{A_{\mathbf{T}}}] + \Big(\mathbb{P}_{\tilde{\xi}_0}[\tau_{A_{\mathbf{T}}} < T_0] + \hat{\mathbb{P}}_{(\xi_0,\tilde{\xi}_0)}[\xi_{T_0} \neq \tilde{\xi}_{T_0}]\Big)\Big(T_0 + \max_{\sigma \in \bigcup_{i<\mathbf{t}} A_i} \mathbb{E}_\sigma[\tau_{A_{\mathbf{T}}}]\Big). \tag{5.44}$$

To bound $\mathbb{P}_{\tilde{\xi}_0}[\tau^{\tilde{\xi}}_{A_{\mathbf{T}}} < T_0]$, we limit the number of steps that $\{\tilde{\xi}_t\}_{t\ge 0}$ can take until time T_0. From (5.2) and Stirling's approximation we have that

$$\begin{aligned} \mathbb{P}\Big[N_{\tilde{\xi}}(T_0) \ge 3nT_0\Big] &\le \sum_{k=0}^{\infty} e^{nT_0+k}\left(\frac{nT_0}{3nT_0+k}\right)^{3nT_0+k} \\ &\le e^{nT_0}\left(\tfrac{1}{3}\right)^{3nT_0}\sum_{k=0}^{\infty} e^{k}\left(\tfrac{1}{3}\right)^{k} \le 11\,(0.91)^{3ne^{\frac{1}{2}\Gamma^\star}}. \end{aligned} \tag{5.45}$$

It therefore follows that with high probability $\{\tilde{\xi}_t\}_{t\ge0}$ does not make more than $3nT_0$ steps until time T_0. Hence

$$\mathbb{P}_{\tilde{\xi}_0}\big[\tau^{\tilde{\xi}}_{A_{\mathbf{T}}} < T_0\big] \le \mathbb{P}_{\tilde{\xi}_0}\Big[\tau^{\tilde{\xi}}_{A_{\mathbf{T}}} < T_0, N_{\tilde{\xi}}(T_0) < 3nT_0\Big] + 11\,(0.91)^{3ne^{\frac{1}{2}\Gamma^\star}}. \tag{5.46}$$

Finally, note that the event $\{\tau^{\tilde{\xi}}_{A_{\mathbf{T}}} < T_0, N_{\tilde{\xi}}(T_0) < 3nT_0\}$ implies that $\{\tilde{\xi}_t\}_{t\ge0}$ makes fewer than $3nT_0$ returns to the set $A_{\mathbf{M}}$ before reaching $A_{\mathbf{T}}$. The probability of this event is bounded from above by $3nT_0e^{\Gamma^\star n + Cn^{2/3}} = 3ne^{\frac{1}{2}\Gamma^\star n + Cn^{2/3}}$, and hence

$$\mathbb{P}_{\tilde{\xi}_0}\big[\tau^{\tilde{\xi}}_{A_{\mathbf{T}}} < T_0\big] \le 4n\, e^{\frac{1}{2}\Gamma^\star n + Cn^{2/3}}. \tag{5.47}$$

Finally, from (5.35) and (5.47) we obtain

$$\mathbb{E}_{\xi_0}\big[\tau^{\xi}_{A_{\mathbf{T}}}\big] = \mathbb{E}\big[\tau^{\tilde{\xi}}_{A_{\mathbf{T}}}\big]\,[1+o(1)], \tag{5.48}$$

which settles the claim. □

6. Hitting time comparisons

In Section 6.1 we estimate ratios of hitting probabilities via ratios of capacities (Lemma 6.1 below). In Section 6.2 we show that the dynamics has a downward drift before it reaches the energetic barrier (Lemma 6.2 below). In Section 6.3 we prove Lemma 2.3 via a further uniformity lemma (Lemma 6.3 below).

6.1. Comparison of hitting probabilities.

For $x \in A_k$, $1 \le k \le n$ and $m \ge k$, define

$$B_m(x) = \{y \in A_m\colon\ x \subseteq y\}. \tag{6.1}$$

Lemma 6.1 (**Ratios of hitting probabilities**).
With $\mathbb{P}_{\mathrm{ER}_n(p)}$-probability tending to 1 as $n \to \infty$ the following is true. For any $\sigma_1, \sigma_2 \in A_k$, $1 \le k \le n$ (possibly $\sigma_1 = \sigma_2$) and any $a \in B_m(\sigma_1)$, $b \in B_m(\sigma_2)$, $m \ge k$,

$$\frac{\mathbb{P}_{\sigma_1}[\tau_a < \tau_{\sigma_1}]}{\mathbb{P}_{\sigma_2}[\tau_b < \tau_{\sigma_2}]} = \frac{\mathrm{cap}(\sigma_1, a)}{\mathrm{cap}(\sigma_2, b)} = 1 + O\left(e^{-\sqrt{n}}\right), \tag{6.2}$$

$$\frac{\mathbb{P}_a[\tau_{\sigma_1} < \tau_a]}{\mathbb{P}_b[\tau_{\sigma_2} < \tau_b]} = \frac{\mathrm{cap}(a, \sigma_1)}{\mathrm{cap}(b, \sigma_1)} = 1 + O\left(e^{-\sqrt{n}}\right). \tag{6.3}$$

Proof. We will only prove (6.2), the proof of (6.3) being similar.

The first equality in (6.2) is immediate from (4.5). To prove the second equality, recall from (4.1) and (4.3) that

$$\mathrm{cap}(\sigma_1, a) = \min_{f \in Q_{(\sigma_1,a)}} \mathcal{E}(f,f), \qquad \mathcal{E}(f,f) = \tfrac{1}{2}\sum_{\sigma,\sigma' \in S_n} \mu(\sigma)\, r(\sigma,\sigma')\,[f(\sigma) - f(\sigma')]^2, \tag{6.4}$$

where

$$Q_{(\sigma_1,a)} = \{f\colon\ S_n \to [0,1],\ f(\sigma_1) = 1,\ f(a) = 0\}. \tag{6.5}$$

A similar definition holds for cap (σ_2, b) and $Q_{(\sigma_2,b)}$. Let $\tilde{\psi}$ be any bijection on the vertex set V that satisfies

$$
\begin{aligned}
\tilde{\psi}(v) &= v \qquad \forall\, v \in \{\bar{b}\cap\bar{a}\} \cup \{\sigma_1\cap\sigma_2\} \cup \{b\cap a\cap\overline{\sigma_1}\cap\overline{\sigma_2}\}, \\
\tilde{\psi}(\{\sigma_1\cap b\cap\overline{\sigma_2}\}) &= \{\sigma_2\cap a\cap\overline{\sigma_1}\}, \\
\tilde{\psi}\left(\{\sigma_1\cap\bar{b}\}\right) &= \{\sigma_2\cap\bar{a}\}, \\
\tilde{\psi}\left(\{a\cap\overline{\sigma_1}\cap\bar{b}\}\right) &= \{b\cap\overline{\sigma_2}\cap\bar{a}\}.
\end{aligned} \tag{6.6}
$$

It is easy to check that such a bijection exists, which requires verifying that the image sets on the right-hand side have the same cardinality as the argument on the left-hand side. Note that in the special case $\sigma_1 = \sigma_2$ the bijection maps a vertex v with $a(v) = +1 = -b(v)$ to some vertex w with $a(v) = -1 = -b(v)$, and leaves the vertices on which a and b agree fixed.

Let $\psi\colon\ S_n \to S_n$ be the bijection that maps $\xi \in S_n$ to the configuration $\psi(\xi)$ defined by $(\psi(\xi))(v) = \tilde{\psi}(v)$. Observe that $\psi(\sigma_1) = \sigma_2$ and $\psi(a) = b$. Thus, for any $f \in Q_{(\sigma_1,a)}$, the function $\tilde{f} = f\circ\psi$ is in $Q_{(\sigma_2,b)}$. Furthermore, it is clear that $|\xi| = |\psi(\xi)|$, and that ψ is a graph isomorphism on S_n, i.e., $\xi \sim \xi'$ if and only if $\psi(\xi) \sim \psi(\xi')$. Hence

$$
\begin{aligned}
&\mu(\xi)\, r(\xi,\xi')\left[\tilde{f}(\xi) - \tilde{f}(\xi')\right]^2 = \mu(\xi)\, r(\xi,\xi')\left[f(\psi(\xi)) - f(\psi(\xi'))\right]^2 \\
&\le \left(\max_{\xi,\xi'\in A_{|\xi|}} \frac{\mu(\xi)\, r(\xi,\xi')}{\mu(\psi(\xi))\, r(\psi(\xi),\psi(\xi'))}\right) \mu(\psi(\xi))\, r(\psi(\xi),\psi(\xi'))\left[f(\psi(\xi)) - f(\psi(\xi'))\right]^2 \\
&\le \left(\frac{\min_{\xi\in A_{|\xi|}} e^{-\beta H_n(\xi)}}{\max_{\xi\in A_{|\xi|}} e^{-\beta H_n(\xi)}}\right) \mu(\psi(\xi))\, r(\psi(\xi),\psi(\xi'))\left[f(\psi(\xi)) - f(\psi(\xi'))\right]^2 \\
&\le \left(1 + O\big(e^{-\sqrt{n}}\big)\right)\mu(\psi(\xi))\, r(\psi(\xi),\psi(\xi'))\left[f(\psi(\xi)) - f(\psi(\xi'))\right]^2,
\end{aligned} \tag{6.7}
$$

where the last inequality uses Lemma 3.2. It follows that

$$
\mathscr{E}(\tilde{f},\tilde{f}) \le \left[1 + O\big(e^{-\sqrt{n}}\big)\right] \mathscr{E}(f,f), \tag{6.8}
$$

and likewise

$$
\mathscr{E}(f,f) \le \left[1 + O\big(e^{-\sqrt{n}}\big)\right] \mathscr{E}(\tilde{f},\tilde{f}). \tag{6.9}
$$

Since this is true for any $f \in Q_{(\sigma_1,a)}$, the second inequality in (6.2) follows via (6.4). □

6.2. **Downward drift.** Define (recall (3.13))

$$
\mathbf{t} = \min\left\{1 \le i \le n\colon\ 2\beta\left(-\theta_i^l + \tfrac{2}{n}\right) + \log\left(\tfrac{n-i}{i+1}\right) \le 0\right\}, \tag{6.10}
$$

$$
\chi_2 = \min\left\{\mathbf{t} \le i \le n\colon\ 2\beta\left(-\theta_i^l + \tfrac{2}{n}\right) + \log\left(\tfrac{n-i}{i+1}\right) > 0\right\}, \tag{6.11}
$$

and

$$
\tilde{\chi} = \tfrac{1}{2}(\mathbf{t} + \chi_2). \tag{6.12}
$$

Note that $\frac{\mathbf{t}}{n} < \frac{\chi_2}{n} < \frac{1}{2}$. For any $x, a, b \in S_n$ (see [4, Lemma 8.4]),

$$
\begin{aligned}
\mathbb{P}_x[\tau_a < \tau_b] &= \sum_{i\in\mathbb{N}_0}\left(\mathbb{P}_x\left[\tau_x < \tau_{\{a,b\}}\right]\right)^i \mathbb{P}_x\left[\tau_a < \tau_{\{b,x\}}\right] \\
&= \frac{\mathbb{P}_x\left[\tau_a < \tau_{\{b,x\}}\right]}{1 - \mathbb{P}_x\left[\tau_x < \tau_{\{a,b\}}\right]} = \frac{\mathbb{P}_x\left[\tau_a < \tau_{\{b,x\}}\right]}{\mathbb{P}_x\left[\tau_{\{a,b\}} < \tau_x\right]} \qquad (6.13) \\
&\le \frac{\mathbb{P}_x[\tau_a < \tau_x]}{\mathbb{P}_x[\tau_b < \tau_x]}. \qquad (6.14)
\end{aligned}
$$

We will make use of inequality (6.14), as well as the following more precise bound on (6.13):

$$
\frac{\mathbb{P}_x\left[\tau_a < \tau_{\{b,x\}}\right]}{\mathbb{P}_x\left[\tau_{\{a,b\}} < \tau_x\right]} \le \frac{\mathbb{P}_x[\tau_a < \tau_x]}{\mathbb{P}_x[\tau_a < \tau_x] + \mathbb{P}_x[\tau_b < \tau_x] - \mathbb{P}_x[\tau_a < \tau_x, \tau_b < \tau_x]}. \tag{6.15}
$$

Lemma 6.2 **(Downward drift).** *With $\mathbb{P}_{\mathrm{ER}_n(p)}$-probability tending to 1 as $n \to \infty$ the following is true. For any $a, b \in A_{\tilde{\chi}}$ (with possibly $a = b$) and $x \in A^{\mathbf{t}}$,*

$$\mathbb{P}_a\left[\tau_x < \tau_b\right] = 1 + O\left(\tfrac{1}{\sqrt{n}}\right). \tag{6.16}$$

Proof. The proof comes in 3 Steps.

1. Intuitively, (6.16) holds because the magnetization has a downward drift for all $\sigma \in \bigcup_{i=\mathbf{t}}^{\chi_2} A_i$. In particular, this implies that if $\{\xi_t\}_{t\geq 0}$ starts at $\xi_0 = a$, then $X_t = |\xi_t|$ has a negative drift whenever $\xi_t \in \cup_{i=\mathbf{t}}^{\chi_2} A_i$. To see why, note that if $\xi \in A_i$ for $\tilde{\chi} - n^{1/2} \leq i \leq \tilde{\chi} + n^{1/2}$, then by Lemma 3.4 we have, for any $\epsilon > 0$,

$$\begin{aligned}
\mathbb{P}_\xi&\left[\tau_{A_{i-1}} < \tau_{A_{i+1}}\right] \qquad (6.17)\\
&= \tfrac{\sum_{\xi'\in A_{i-1}} r(\xi,\xi')}{\sum_{\xi'\in A_{i-1}} r(\xi,\xi')+\sum_{\xi'\in A_{i+1}} r(\xi,\xi')}\\
&\geq \tfrac{(i-2n^{2/3})e^{-2\beta[-\vartheta_i+3n^{-1/3}]_+}}{(i-2n^{2/3})\,e^{-2\beta[-\vartheta_i+3n^{-1/3}]_+}+(n-i-2n^{2/3})\,e^{-2\beta[\vartheta_i-3n^{-1/3}]_+}+2n^{2/3}}\\
&\geq \tfrac{1}{1+(\frac{n}{i}-1)\,e^{-2\beta\vartheta_i}+\epsilon}, \qquad (6.18)
\end{aligned}$$

where $\vartheta_i = p(1 - \frac{2i}{n} - h$. Since, by (6.10) and (6.11), $\log\left(\frac{n}{i} - 1\right) - 2\beta\vartheta_i < 0$, we have that $\mathbb{P}_\xi[\tau_{A_{i-1}} < \tau_{A_{i+1}}] \geq \tilde{p}$ for some $\tilde{p} > \frac{1}{2}$. Let t_j denote the time of the j^{th} jump of $\{\xi_t\}_{t\geq 0}$. Then for $j \leq \sqrt{n}$ (and hence $|X_{t_j} - X_0| \leq \sqrt{n}$) it follows from Hoeffding's inequality that if $\xi_0 \in A_{\tilde{\chi}}$ (and hence $X_0 = \tilde{\chi}$), then for any $\epsilon > 0$,

$$\mathbb{P}_{\xi_0}\left[X_{t_{\sqrt{n}}} \geq m - \sqrt{n}\,(2\tilde{p} - 1 - \epsilon)\right] \leq e^{-\epsilon\sqrt{n}}. \tag{6.19}$$

Thus, with high probability, after $\sqrt{n}$ steps ξ_t is in $D_{\tilde{\chi}} = \cup_{i\geq \tilde{\chi}-\sqrt{n}(2\tilde{p}-1-\epsilon)} A_i$. We check that, in the first $\sqrt{n}$ steps, with high probability ξ_t does not visit (or return to) b. Indeed, note that for any neighbour $\xi' \sim b$, from [???] we have that

$$\max_{\xi'} \mathbb{P}_{\xi'}\left[\xi'_{t_1} = b\right] \leq \frac{1}{n}\, e^{2\beta(p+h+1)}, \tag{6.20}$$

and so by a union bound we get

$$\mathbb{P}_a\left[\bigcup_{j=1}^{\sqrt{n}} \xi_{t_j} = b\right] \leq \frac{1}{\sqrt{n}}\, e^{2\beta(p+h+1)}. \tag{6.21}$$

Hence we conclude that, with probability at least $1 - O(\frac{1}{\sqrt{n}})$, $\{\xi_t\}_{t\geq 0}$ reaches the set $D_{\tilde{\chi}}$ before hitting (or returning to) b.

2. We next show that, after reaching $D_{\tilde{\chi}}$, with high probability the process $\{\xi_t\}_{t\geq 0}$ reaches $A^{\mathbf{t}}$ before returning to $A_{\tilde{\chi}}$, and hence also before reaching b. We will split the journey from $A_{\tilde{\chi}}$ to $A^{\mathbf{t}}$ into two parts: (1) from $A_{\tilde{\chi}}$ to $A^{\mathbf{t}+\tilde{\epsilon}n}$ for some small $\tilde{\epsilon} > 0$; (2) from $A^{\mathbf{t}+\tilde{\epsilon}n}$ to $A_{\tilde{\chi}}$. Indeed, as in (6.18), there is some $\tilde{p} > \frac{1}{2}$ such that, for all $\mathbf{t} + \tilde{\epsilon}n \leq i \leq \tilde{\chi}$ and all $\xi \in A_i$,

$$\mathbb{P}_\xi\left[\tau_{A_{i-1}} < \tau_{A_{i+1}}\right] \geq \tilde{p}. \tag{6.22}$$

It follows from the Gambler's ruin problem that for the first part, for any $0 < \alpha \leq 2\tilde{p} - 1 - \epsilon$ and any $\xi' \in A_{m-\alpha\sqrt{n}}$,

$$\mathbb{P}_{\xi'}\left[\tau_{A_{k+\tilde{\epsilon}n}} < \tau_{A_m}\right] \geq \frac{1 - r^{\alpha\sqrt{n}}}{1 - r^{m-k-\tilde{\epsilon}n}} \geq 1 - r^{\alpha\sqrt{n}}, \qquad r = \frac{1-\tilde{p}}{\tilde{p}}. \tag{6.23}$$

For the second part, from $A^{\mathbf{t}+\tilde{\epsilon}n}$ to $A^{\mathbf{t}}$, we have

$$\mathbb{P}_\xi\left[\tau_{A_{i-1}} < \tau_{A_{i+1}}\right] \geq \tfrac{1}{2}, \tag{6.24}$$

and here we can use the following argument. A simple random walk, starting at $\mathbf{t} + \tilde{\epsilon}n$ and restricted to $[\mathbf{t}, \mathbf{t} + 2\tilde{\epsilon}n]$, reaches $\mathbf{t}$ within $\leq n^3$ steps with probability $\geq 1 - \frac{C_1}{n}$ for some $C_1 > 0$. But $\{X_t\}_{t\geq 0} = \{|\xi_t|\}_{t\geq 0}$ is stochastically dominated by simple random walk, and so the same bounds on the number of steps it takes to reach $\mathbf{t}$ apply to it as well, conditional on $\{X_t\}_{t\geq 0}$ not reaching $\mathbf{t} + 2\tilde{\epsilon}n$. But the latter event occurs with high probability because, with probability at leat $1 - O(e^{-C_2 n})$ for some $C_2 > 0$, $\{X_t\}_{t\geq 0}$ does not reach

$2\tilde{\epsilon}n$ before it has made at least $e^{C_3 n}$ steps, for some $C_3 > 0$. This proves that with high probability $\{\xi_t\}_{t\geq 0}$, starting at some $a \in A_{\tilde{\chi}}$, visits $A^{\mathbf{t}}$ before reaching the state b.

3. Having established the above, we next show that, for any $x \in A^{\mathbf{t}}$, with high probability the process starting in $A^{\mathbf{t}}$ visits x before it visits $b \in A_{\tilde{\chi}}$. Let $y \in A^{\mathbf{t}}$ and note that, by (6.14),

$$\mathbb{P}_y\left[\tau_b < \tau_x\right] \leq \frac{\mathbb{P}_y\left[\tau_b < \tau_y\right]}{\mathbb{P}_y\left[\tau_x < \tau_y\right]} = \frac{\operatorname{cap}(y,b)}{\operatorname{cap}(y,x)}. \tag{6.25}$$

From the definition of the capacity, using the test function $f \in Q_{(y,b)}$ given by $f(b) = 0$ and $f(\sigma) = 1$ for all $\sigma \in S_n$, $\sigma \neq b$, we get

$$\operatorname{cap}(y,b) \leq e^{-\beta H_n(\boxminus)} \sum_{\zeta \in S_n} \mu(b)\, r(b,\zeta) \leq \frac{1}{Z} e^{-2\beta\left(\left(p\tilde{\chi}\left(1-\frac{\tilde{\chi}}{n}\right)-cn^{1/2}\right)-h\tilde{\chi}\right)}, \tag{6.26}$$

where the second inequality in (6.26) follows from (3.3). Now let $\gamma\colon\, y \to x$ be a path of shortest distance between y and x, such that $\gamma_i \in A^{\mathbf{t}} \cup A^{\mathbf{t}-1}$ for all $1 \leq i \leq |\gamma|$, where we observe also that $|\gamma| = |y \triangle x| \leq n$. For any $f \in Q_{(y,x)}$ we have $f(\gamma_1) = f(y) = 1$ and $f(\gamma_{|y\triangle x|}) = f(x) = 0$, and so it must be that $|f(\gamma_{i_\blacklozenge}) - f(\gamma_{i_\blacklozenge+1})| \geq \frac{1}{n}$ for some $1 \leq i_\blacklozenge \leq |y\triangle x| - 1$. This implies that

$$\begin{aligned}
\operatorname{cap}(y,x) &\geq \frac{1}{n^2}\mu\left(\gamma_{i_\blacklozenge}\right) r\left(\gamma_{i_\blacklozenge}, \gamma_{i_\blacklozenge+1}\right) \\
&\geq \frac{1}{n^2}\frac{1}{Z} e^{-\beta H_n(\boxminus)}, e^{-2\beta\left(\left(p\mathbf{t}\left(1-\frac{\mathbf{t}}{n}\right)+cn^{1/2}\right)-h\mathbf{t}\right)},
\end{aligned} \tag{6.27}$$

where again the second inequality follows from (3.3). From (6.25) and the observation that the map $i \mapsto ipi\left(1-\frac{i}{n}\right) - hi$ is increasing on $[\mathbf{t}, \chi_2]$, we conclude that

$$\mathbb{P}_y\left[\tau_b < \tau_x\right] \leq n^2 \exp\left(-2\beta\left(p\tilde{\chi}\left(1-\tfrac{\tilde{\chi}}{n}\right) - p\mathbf{t}\left(1-\tfrac{\mathbf{t}}{n}\right) - 2cn^{1/2} - h\left(\tilde{\chi}-\mathbf{t}\right)\right)\right) = O\big(e^{-Cn}\big)$$

for some $C > 0$. $\square$

6.3. **Uniform hitting probability.** In this section we prove Lemma 2.3, which is implied by the following lemma.

Lemma 6.3 (Uniformity estimate). *With $\mathbb{P}_{\mathrm{ER}_n(p)}$-probability tending to 1 as $n \to \infty$ the following is true. For any $x \in A^{\mathbf{t}}$ and $a, b \in A_{\tilde{\chi}}$,*

$$\mathbb{P}_x\left[\tau_a < \tau_b\right] = \tfrac{1}{2} + O\left(\tfrac{1}{\sqrt{n}}\right). \tag{6.28}$$

Proof. The proof comes in 3 Steps.

1. We first prove the result for the special case $a, b \in B_{\tilde{\chi}}(x)$. Recall from (6.13) and (6.15) that

$$\mathbb{P}_x\left[\tau_a < \tau_b\right] \leq \frac{\mathbb{P}_x\left[\tau_a < \tau_x\right]}{\mathbb{P}_x\left[\tau_a < \tau_x\right] + \mathbb{P}_x\left[\tau_b < \tau_x\right] - \mathbb{P}_x\left[\tau_a < \tau_x, \tau_b < \tau_x\right]}, \tag{6.29}$$

while from Lemma 6.1 we have that

$$\frac{\mathbb{P}_x\left[\tau_a < \tau_x\right]}{\mathbb{P}_x\left[\tau_b < \tau_x\right]} = \frac{\operatorname{cap}(x,a)}{\operatorname{cap}(x,b)} = 1 + O\big(e^{-\sqrt{n}}\big). \tag{6.30}$$

Thus, (6.29) becomes

$$\mathbb{P}_x\left[\tau_a < \tau_b\right] \leq \frac{\mathbb{P}_x\left[\tau_b < \tau_x\right]}{\left(2 + O(e^{-\sqrt{n}})\right)\mathbb{P}_x\left[\tau_b < \tau_x\right] - \mathbb{P}_x\left[\tau_a < \tau_x, \tau_b < \tau_x\right]}. \tag{6.31}$$

Hence it suffices to show that $\mathbb{P}_x\left[\tau_a < \tau_x, \tau_b < \tau_x\right]$ is much smaller than $\mathbb{P}_x\left[\tau_b < \tau_x\right]$. To that end we observe that

$$\begin{aligned}
\mathbb{P}_x\left[\tau_a < \tau_x, \tau_b < \tau_x\right] &= \mathbb{P}_x\left[\tau_a < \tau_b < \tau_x\right] + \mathbb{P}_x\left[\tau_b < \tau_a < \tau_x\right] \\
&\leq \mathbb{P}_x\left[\tau_a < \tau_x\right]\mathbb{P}_a\left[\tau_b < \tau_x\right] + \mathbb{P}_x\left[\tau_b < \tau_x\right]\mathbb{P}_b\left[\tau_a < \tau_x\right] \\
&= \mathbb{P}_x\left[\tau_a < \tau_x\right] O\left(\tfrac{1}{\sqrt{n}}\right) + \mathbb{P}_x\left[\tau_b < \tau_x\right] O\left(\tfrac{1}{\sqrt{n}}\right) \\
&= \mathbb{P}_x\left[\tau_a < \tau_x\right] O\left(\tfrac{1}{\sqrt{n}}\right),
\end{aligned} \tag{6.32}$$

where the second equality follows from Lemma 6.2 and the third equality from (6.30). Hence (6.31) becomes

$$\mathbb{P}_x[\tau_a < \tau_b] \leq \frac{\mathbb{P}_x[\tau_a < \tau_x]}{\left(2 + O\left(\frac{1}{\sqrt{n}}\right)\right)\mathbb{P}_x[\tau_b < \tau_x]} = \tfrac{1}{2} + o(1). \tag{6.33}$$

Since the same inequality holds also for $\mathbb{P}_x[\tau_b < \tau_a] = 1 - \mathbb{P}_x[\tau_a < \tau_b]$, the claim follows for the case $a, b \in B_{\tilde{\chi}}(x)$.

2. We next proceed with the general case: $a, b \in A_{\tilde{\chi}}$. Let $y \in A^{\mathbf{t}}$ be any configuration such that $|y \cap a| = |y \cap b|$. For instance, if $|a \cap b| \geq \mathbf{t}$, then we can take y to be any subset of $a \cap b$ that is of size $\mathbf{t}$ (note that this is equivalent to $a, b \in B_{\tilde{\chi}}(y)$). If $|\overline{a \cup b}| \geq \mathbf{t}$, then set y to be some subset of $\overline{a \cup b}$ of size $\mathbf{t}$. Otherwise we can take y to be any selection of $\mathbf{t}$ vertices from the sets $a \cap b$, $\overline{a \cup b}$ and pairs of vertices $(v_1, v_2) \in \{a \backslash b\} \times \{b \backslash a\}$. We claim that

$$\frac{\mathbb{P}_y[\tau_b < \tau_y]}{\mathbb{P}_y[\tau_a < \tau_x]} = 1 + O\left(e^{-\sqrt{n}}\right), \tag{6.34}$$

and similarly,

$$\frac{\mathbb{P}_b[\tau_y < \tau_b]}{\mathbb{P}_a[\tau_y < \tau_a]} = 1 + O\left(e^{-\sqrt{n}}\right). \tag{6.35}$$

In other words, the result of Lemma 6.1 extends to the general case, with x replaced by y. The proof is analogous to that of Lemma 6.1. We will first construct a bijection of V, then a bijection of S_n, and then use this bijection as was done before.

3. Let $\tilde{\psi}$ be any bijection on the vertex set V that satisfies

$$\begin{aligned} \tilde{\psi}(\{y\}) &= \{y\}, \\ \tilde{\psi}(\{y \cap a\}) &= \{y \cap b\}, \\ \tilde{\psi}(\{\overline{y} \cap a\}) &= \{\overline{y} \cap b\}, \end{aligned} \tag{6.36}$$

and let $\psi \colon S_n \to S_n$ be the bijection that maps $\sigma \in S_n$ to the configuration $\psi(\sigma)$ defined by $\psi(\sigma)(v) = \tilde{\psi}(v)$. Clearly, ψ is an isomorphism on S_n preserving the relation $\sim$, $\psi(y) = y$, $\psi(a) = b$, and $|\psi(\xi)| = |\xi|$ for all $\xi \in S_n$. Therefore (6.34) and (6.35) follow by an argument identical to (6.7), (6.8) and (6.9) in the proof of Lemma 6.1. Note that by an argument identical to (6.28),

$$\mathbb{P}_x\left[\tau_y < \tau_{\{a,b\}}\right] = 1 - O\left(e^{-Cn}\right), \tag{6.37}$$

and by an argument identical to (6.32),

$$\mathbb{P}_y[\tau_a < \tau_y, \tau_b < \tau_y] \leq \mathbb{P}_y[\tau_a < \tau_y]\, O\left(\tfrac{1}{\sqrt{n}}\right). \tag{6.38}$$

Hence, from (6.13), (6.15), (6.34), (6.37), (6.38) and the Markov property, we have that

$$\begin{aligned} \mathbb{P}_x[\tau_a < \tau_b] &= \mathbb{P}_x\left[\tau_a < \tau_b \middle| \tau_y < \tau_{\{a,b\}}\right] + O\left(e^{-Cn}\right) \\ &= \mathbb{P}_y[\tau_a < \tau_b] + O\left(te^{-Cn}\right) \\ &\leq \frac{\mathbb{P}_y[\tau_a < \tau_y]}{\mathbb{P}_y[\tau_b < \tau_y] + \mathbb{P}_y[\tau_b < \tau_y] - \mathbb{P}_y[\tau_a < \tau_x, \tau_b < \tau_x]} + O\left(e^{-Cn}\right) \\ &= \tfrac{1}{2} + O\left(\tfrac{1}{\sqrt{n}}\right). \end{aligned} \tag{6.39}$$

By symmetry, we also have

$$\mathbb{P}_x[\tau_b < \tau_a] \leq \tfrac{1}{2} + O\left(\tfrac{1}{\sqrt{n}}\right). \tag{6.40}$$

□

7. A coupling scheme

In this section we define a coupling of $(\xi_t)_{t\geq 0}$ and $\{\tilde{\xi}_t\}_{t\geq 0}$ with arbitrary starting configurations in $A_{\mathbf{M}}$. The coupling is divided into a short-term scheme, defined in Section 7.1 and analysed in Lemma 7.1 below, followed by a long-term scheme, defined in Section 7.2 and analysed Corollary 7.3 below. The goal of the coupling is to keep the process $\{m(\xi_t)\}_{t\geq 0}$ bounded by $\{\theta_t^u\}_{t\geq 0}$ from above and bounded by $\{\theta_t^l\}_{t\geq 0}$ from below (the precise meaning will become clear in the sequel).

7.1. Short-term scheme.

Lemma 7.1 (Short-term coupling). *With $\mathbb{P}_{\mathrm{ER}_n(p)}$-probability tending to 1 as $n \to \infty$, there is a coupling $\{\xi_t, \tilde{\xi}_t\}_{t\geq 0}$ of $\{\xi_t\}_{t\geq 0}$ and $\{\tilde{\xi}_t\}_{t\geq 0}$ such that*

$$\mathbb{P}\big[\xi_{2n} \neq \tilde{\xi}_{2n}\big] \leq e^{n^{-2/3}} \tag{7.1}$$

for any initial states $\xi_0 \in A_{\mathbf{M}}$ and $\tilde{\xi}_0 \in A_{\mathbf{M}}$.

Proof. The main idea behind the proof is as follows. Define

$$W_1^t = \{v \in V\colon\ \xi_t(v) = -\tilde{\xi}_t(v)\} = \xi_t \Delta \tilde{\xi}_t, \tag{7.2}$$

i.e., the symmetric difference between the two configurations ξ_t and $\tilde{\xi}_t$, and

$$W_2^t = \{v \in V\colon\ \xi_t(v) = \tilde{\xi}_t(v)\} = V \setminus W_1^t. \tag{7.3}$$

The coupling we are about to define will result in the set W_1^t shrinking at a higher rate than the set W_2^t, which will imply that W_1^t contracts to the empty set. The proof comes in 8 Steps.

1. We begin with bounds on the relevant transition rates that will be required in the proof. Recall from Lemma 3.4 (in particular, (3.20) and (3.22)) that with $\mathbb{P}_{\mathrm{ER}_n(p)}$-probability at least $1 - e^{-2n}$ there are at most $2n^{2/3}$ vertices $v \in \overline{\xi_t}$ (i.e., $\xi_t(v) = -1$) such that $|E(v,\xi_t)| = |\{w \in \xi_t\colon\ (v,w) \in E\}| \geq p\,|\xi_t| + n^{2/3}$, and similarly at most $2n^{2/3}$ vertices $v \in \overline{\xi_t}$ such that $|E(v,\xi_t)| \leq p\,|\xi_t| - n^{2/3}$. Analogous bounds are true for $\tilde{\xi}_t$, $t \geq 0$. Denote the set of *bad* vertices for ξ_t by

$$B_t = \big\{v \in \overline{\xi_t}\colon\ \big||E(v,\xi_t)| - p|\xi_t|\big| \geq n^{2/3}\big\}, \tag{7.4}$$

and the set of bad vertices for $\tilde{\xi}_t$ by $\tilde{B}_t$. Let $\hat{B}_t = B_t \cup \tilde{B}_t$. Recall that ξ_t^v denotes the configuration obtained from ξ_t by flipping the sign at vertex $v \in V$. If $v \notin \hat{B}_t$, then from (1.3) and Lemma 3.2 it follows that, for $v \notin \xi_t$,

$$\begin{aligned} H_n(\xi_t^v) - H_n(\xi_t) &= \tfrac{2}{n}\left(|\partial_E \xi_t^v| - |\partial_E \xi_t|\right) - 2h \\ &= \tfrac{2}{n}\left(\deg(v) - 2\,|E(v,\xi_t)|\right) - 2h \\ &\leq \tfrac{2}{n}\left(pn + n^{1/2}\log n - 2p\,|\xi_t| + 2n^{2/3}\right) - 2h, \end{aligned} \tag{7.5}$$

and similarly, for $v \in \xi_t$,

$$H_n(\xi_t^v) - H_n(\xi_t) \leq \tfrac{2}{n}\left(pn + n^{1/2}\log n - 2p\,(n - |\xi_t|) + 2n^{2/3}\right) + 2h. \tag{7.6}$$

Again, by (1.3) and Lemma 3.2, we have similar lower bounds, namely, if $v \notin \hat{B}_t$, then, for $v \notin \xi_t$,

$$H_n(\xi_t^v) - H_n(\xi_t) \geq \tfrac{2}{n}\left(pn - n^{1/2}\log n - 2p\,|\xi_t| - 2n^{2/3}\right) - 2h, \tag{7.7}$$

and, for $v \in \xi_t$,

$$H_n(\xi_t^v) - H_n(\xi_t) \geq \tfrac{2}{n}\left(pn - n^{1/2}\log n - 2p\,(n - |\xi_t|) - 2n^{2/3}\right) + 2h. \tag{7.8}$$

Identical bounds hold for $H_n(\tilde{\xi}_t^v) - H_n(\tilde{\xi}_t)$. Therefore, if $v \notin \hat{B}_t$, and if either $v \in \xi_t \cap \tilde{\xi}_t$ or $v \notin \xi_t \cup \tilde{\xi}_t$, then

$$\begin{aligned}
\left| r(\xi_t, \xi_t^v) - r(\tilde{\xi}_t, \tilde{\xi}_t^v) \right| &= \left| e^{-\beta[H_n(\xi_t^v) - H_n(\xi_t)]_+} - e^{-\beta[H_n(\tilde{\xi}_t^v) - H_n(\tilde{\xi}_t)]_+} \right| \\
&= e^{-\beta[H_n(\xi_t^v) - H_n(\xi_t)]_+} \left| 1 - e^{\beta\left([H_n(\xi_t^v) - H_n(\xi_t)]_+ - [H_n(\tilde{\xi}_t^v) - H_n(\tilde{\xi}_t)]_+\right)} \right| \\
&\leq [1 + o_n(1)]\, e^{-\beta[H_n(\xi_t^v) - H_n(\xi_t)]_+} \left(e^{8\beta n^{-1/3} + \frac{4p}{n}\left(|\xi_t| - |\tilde{\xi}_t|\right)} - 1 \right) \qquad (7.9) \\
&\leq [1 + o_n(1)] \left(e^{8\beta n^{-1/3} + \frac{4p}{n}\left(|\xi_t| - |\tilde{\xi}_t|\right)} - 1 \right) \\
&\leq [1 + o_n(1)] \left(8\beta n^{-1/3} + \tfrac{4p}{n}\left(|\xi_t| - |\tilde{\xi}_t|\right) \right).
\end{aligned}$$

2. Having established the above bounds on the transition rates, we give an explicit construction of the coupling $\{\xi_t, \tilde{\xi}_t\}_{t \geq 0}$.

Definition 7.2.

(I) We first define the coupling for time $t = 0$. For $t > 0$ this coupling will be *renewed* after each renewal of $\{\xi_t, \tilde{\xi}_t\}_{t \geq 0}$, i.e., whenever either of the two processes jumps to a new state. To that end, for every $v \in W_2^0$ (i.e., $\xi_0(v) = \tilde{\xi}_0(v)$), couple the exponential random variables $e_0^v \simeq \mathrm{Exp}(r(\xi_0, \xi_0^v))$ and $\tilde{e}_0^v \simeq \mathrm{Exp}(r(\tilde{\xi}_0, \tilde{\xi}_0^v))$ associated with the transitions $\xi_0 \to \xi_0^v$ and $\tilde{\xi}_0 \to \tilde{\xi}_0^v$ according to the following scheme:

(1) Choose a point
$$(x, y) \in \{(x', y') : \ 0 \leq x' < \infty,\ 0 \leq y' \leq r(\xi_0, \xi_0^v)\, e^{-r(\xi_0, \xi_0^v)x'}\}$$
uniformly and set $e_0^v = x$. Note that, indeed, this gives $e_0^v \simeq \mathrm{Exp}(r(\xi_0, \xi_0^v))$.

(2) If the value y from step 1 satisfies $y \leq r(\tilde{\xi}_0, \tilde{\xi}_0^v) \exp(-r(\tilde{\xi}_0, \tilde{\xi}_0^v)x)$, then set $\tilde{e}_0^v = e_0^v = x$. Else, choose
$$(x*, y*) \in \left\{(x', y') : 0 \leq x' < \infty,\ r(\xi_0, \xi_0^v)\, e^{-r(\xi_0, \xi_0^v)x'} < y' \leq r(\tilde{\xi}_0, \tilde{\xi}_0^v) e^{-r(\tilde{\xi}_0, \tilde{\xi}_0^v)x'}\right\}$$
uniformly and independently from the sampling in step 1, and set $\tilde{e}_0^v = x*$. Note that this too gives $e_0^v \simeq \mathrm{Exp}(r(\tilde{\xi}_0, \tilde{\xi}_0^v))$.

(II) For every $v \in W_1^0$, sample the random variables $e_0^v \simeq \mathrm{Exp}(r(\xi_0, \xi_0^v))$ and $\tilde{e}_0^v \simeq \mathrm{Exp}(r(\tilde{\xi}_0, \tilde{\xi}_0^v))$ associated with the transitions $\xi_0 \to \xi_0^v$ and $\tilde{\xi}_0 \to \tilde{\xi}_0^v$ independently. At time $t = 0$, we use the above rules to define the *jump* times associated with any vertex $v \in V$. Recall that W_2^0 is the set of vertices where the two configurations agree in sign. The aim of the coupling defined above is to preserve that agreement. Following every renewal, we re-sample all transition times anew (i.e., we choose new copies of the exponential variables as was done above). We proceed in this way until the first of the following two events happens: either $\xi_t = \tilde{\xi}_t$, or $n \log n$ transitions have been made by either one of the two processes.

3. Note that the purpose of limiting the number of jumps to $n \log n$ is to permit us to employ Lemma 5.1, which in turn we use to maintain control on the two processes being similar in volume. Further down we will also show that, with high probability, in time $2n$ no more than $n \log n$ transitions occur. By (7.9) and Lemma 5.2, if $v \notin \hat{B}_t$, then

$$\mathbb{P}\left[e_t^v \neq e_t^v\right] \leq \frac{2(8\beta n^{-1/3} + \frac{4p}{n}(|\xi_t| - |\tilde{\xi}_t|))}{e^{-2\beta(p+h)}}. \qquad (7.10)$$

On the other hand, if $v \in \hat{B}_t$ and we let $z = \frac{2\beta(p+h)}{1 - e^{-2\beta(p+h)}}$, then

$$\begin{aligned}
\mathbb{P}\left[e_t^v \neq e_t^v\right] &= d_{TV}\left(e_t^v, e_t^v\right) \leq e^{-2\beta(p+h)} \int_0^z dx\, \exp\left(-x e^{-2\beta(p+h)}\right) \\
&= 1 - \exp\left(-\frac{2\beta\,(p+h)\, e^{-2\beta(p+h)}}{1 - e^{-2\beta(p+h)}} \right). \qquad (7.11)
\end{aligned}$$

Observe that, for $v \in W_1^t$, with $\mathbb{P}_{\mathrm{ER}_n(p)}$-high probability

$$\sum_{v \in W_1^t} \Big[r(\xi_t, \xi_t^v) + r(\tilde{\xi}_t, \tilde{\xi}_t^v) \Big] \geq [1 + o_n(1)] \left| W_1^t \right| . \tag{7.12}$$

Indeed, by the concentration inequalities of Lemma 3.2 and the bound in Lemma 5.1, it follows that $|\xi_t|$ and $|\tilde{\xi}_t|$ are of similar magnitude:

$$\mathbb{P}\big[||\xi_t| - |\tilde{\xi}_t|| \geq n^{5/6} \big] \leq e^{-n^{2/3}}. \tag{7.13}$$

Therefore, with $\mathbb{P}_{\mathrm{ER}_n(p)}$-high probability, for all but $O(n^{2/3})$ such v,

$$H(\xi_t) - H(\xi_t^v) = [1 + o_n(1)] \left[H\big(\tilde{\xi}_t^v\big) - H\big(\tilde{\xi}_t\big) \right], \tag{7.14}$$

from which (7.12) follows. The rate at which the set W_2^t shrinks is equal to the rate at which it loses $v \in W_2^t$ such that also $v \notin \hat{B}_t$, plus the rate at which it loses $v \in W_2^t$ such that $v \in \hat{B}_t$. From (7.9) it follows that the former is bounded from above by $|W_2^t|(8\beta n^{-1/3} + \frac{4p}{n}(|\xi_t| - |\tilde{\xi}_t|))$, while by (3.20) the latter is bounded by $4n^{2/3}$. Therefore, defining the stopping time

$$\upsilon_i = \inf \left\{ t \colon |W_1^t| = i \right\}, \tag{7.15}$$

we have that

$$\mathbb{P}_{(\xi_t, \tilde{\xi}_t)} \big[\upsilon_{|W_1^t| - 1} < \upsilon_{|W_1^t| + 1} \big] \geq \frac{|W_1^t|}{|W_2^t| [8\beta n^{-1/3} + \frac{4p}{n}(|\xi_t| - |\tilde{\xi}_t|)] + 4n^{2/3} + |W_1^t|}. \tag{7.16}$$

From Lemma 5.1 we know that (with probability $\geq 1 - e^{-n^{2/3}}$) neither $|\xi_t|$ nor $|\tilde{\xi}_t|$ will stray beyond $\mathbf{M} + Cn^{5/6}$ and $\mathbf{M} - Cn^{5/6}$ within $n^2 \log n$ steps. Thus,

$$\Big| |\xi_t| - |\tilde{\xi}_t| \Big| \leq Cn^{5/6}. \tag{7.17}$$

Hence, for $|W_1^t| \geq n^{6/7}$ we have that (7.16) is equal to $1 - o_n(1)$.

4. Next suppose that $|W_1^t| < n^{6/7}$. To bound the rate at which the set W_2^t shrinks, we argue as follows. The rate at which a matching vertex v becomes non-matching equals

$$|r(\xi_t, \xi_t^v) - r(\tilde{\xi}_t, \tilde{\xi}_t^v)|. \tag{7.18}$$

Let

$$\begin{aligned} B_1 &= -2h + \tfrac{2}{n} \left(\deg(v) - 2|E(v, \xi_t)| \right), \\ B_2 &= -2h + \tfrac{2}{n} \Big(\deg(v) - 2|E(v, \tilde{\xi}_t)| \Big), \\ B_3 &= -h + \tfrac{1}{n} \Big(\deg(v) - 2|E(v, \xi_t \cap \tilde{\xi}_t)| \Big). \end{aligned} \tag{7.19}$$

For $v \notin \xi_t \cup \tilde{\xi}_t$, we can estimate

$$\begin{aligned} |r(\xi_t, \xi_t^v) - r(\tilde{\xi}_t, \tilde{\xi}_t^v)| &= \Big| e^{-\beta [B_1]_+} - e^{-\beta [B_2]_+} \Big| \\ &\leq e^{-2\beta [B_3]_+} \Big| e^{-\frac{4\beta}{n} |E(v, \xi_t \backslash \tilde{\xi}_t)|} - e^{-\frac{4\beta}{n} |E(v, \tilde{\xi}_t \backslash \xi_t)|} \Big| \\ &\leq e^{-2\beta [B_3]_+} \tfrac{4\beta}{n} \Big| |E(v, \xi_t \backslash \tilde{\xi}_t)| - |E(v, \tilde{\xi}_t \backslash \xi_t)| \Big| \\ &\leq [1 + o_n(1)]\, e^{-2\beta [-p\mathbf{m} - h]_+} \tfrac{4\beta}{n} \left| E\left(v, W_1^t\right) \right|, \end{aligned} \tag{7.20}$$

where we note that $W_1^t = \xi_t \backslash \tilde{\xi}_t \cup \tilde{\xi}_t \backslash \xi_t$ and use that, by Lemma 3.2 and the bound $|\xi_t \backslash \tilde{\xi}_t| \leq |W_1^t| \leq n^{-6/7}$,

$$\begin{aligned} &\frac{1}{n} \Big(\deg(v) - 2|E(v, \xi_t \cap \tilde{\xi}_t)| \Big) = [1 + o_n(1)]\, p \left(1 - \tfrac{2|\xi_t \cap \tilde{\xi}_t|}{n} \right) \\ &= [1 + o_n(1)]\, p \left(1 - \tfrac{2|\xi_t|}{n} \right) = [1 + o_n(1)]\, p \left(1 - \tfrac{2\mathbf{M}}{n} \right) = -[1 + o_n(1)]\, p\mathbf{m}. \end{aligned} \tag{7.21}$$

Since $|E(v, W_1^t)| \leq [1 + o_n(1)] p |W_1^t|$ and $|V| = n$, we have that

$$\sum_{v \notin \xi_t \cup \tilde{\xi}_t} |r(\xi_t, \xi_t^v) - r(\tilde{\xi}_t, \tilde{\xi}_t^v)| \leq [1 + o_n(1)]\, e^{-2\beta [-p\mathbf{m} - h]_+}\, 4\beta p\, |W_1^t|. \tag{7.22}$$

For $v \in \xi_t \cap \tilde{\xi}_t$, on the other hand, Lemma 3.5 gives that

$$r(\xi_t, \xi_t^v) = r(\tilde{\xi}_t, \tilde{\xi}_t^v) = 1 \tag{7.23}$$

for all but $O(n^{2/3})$ many such v. If v is such that $r(\xi_t, \xi_t^v) \neq r(\tilde{\xi}_t, \tilde{\xi}_t^v)$, then a computation identical to the one leading to (7.22) gives that

$$\sum_{v \in \xi_t \cap \tilde{\xi}_t} |r\,(\xi_t, \xi_t^v) - r(\tilde{\xi}_t, \tilde{\xi}_t^v)| = O(n^{-1/6})\big|W_1^t\big|. \tag{7.24}$$

Combining (7.22) and (7.24), we obtain

$$\sum_{v \in W_t^2} |r\,(\xi_t, \xi_t^v) - r(\tilde{\xi}_t, \tilde{\xi}_t^v)| \leq [1 + o_n(1)]\, e^{-2\beta[-p\mathbf{m}-h]_+}\, 4\beta p\,\big|W_1^t\big|, \tag{7.25}$$

which bounds the rate at which W_2^t shrinks.

5. To bound the rate at which the set W_1^t shrinks, we argue as follows. The rate at which a non-matching vertex v becomes a matching equals

$$r(\xi_t, \xi_t^v) + r(\tilde{\xi}_t, \tilde{\xi}_t^v). \tag{7.26}$$

Note that, for every $v \in W_1^t$,

$$H\,(\xi_t^v) - H\,(\xi_t) = -\,[1 + o_n(1)]\,\big[H(\tilde{\xi}_t^v) - H(\tilde{\xi}_t)\big], \tag{7.27}$$

since, up to an arithmetic correction of magnitude $|W_1^t| = O(n^{6/7})$, v has the same number of neighbours in ξ_t as in $\tilde{\xi}_t$. Hence it follows that

$$\sum_{v \in W_1^t} \Big[r(\xi_t, \xi_t^v) + r(\tilde{\xi}_t, \tilde{\xi}_t^v)\Big] = [1 + o_n(1)]\,\big(e^{-2\beta[-p\mathbf{m}-h]_+} + e^{-2\beta[p\mathbf{m}+h]_+}\big)|W_1^t|, \tag{7.28}$$

which bounds the rate at which W_1^t shrinks.

6. Combining (7.25) and (7.28), and noting that $p\mathbf{m} + h < 0$, we see that $|W_1^t|$ is contracting when

$$[1 + o_n(1)]\,\Big(e^{2\beta(p\mathbf{m}+h)} + 1\Big)\,\big|W_1^t\big| > [1 + o_n(1)]\,\Big(e^{2\beta(p\mathbf{m}+h)}\,4\beta p\Big)\,\big|W_1^t\big|. \tag{7.29}$$

For this in turn it suffices that

$$e^{2\beta(p\mathbf{m}+h)} + 1 > e^{2\beta(p\mathbf{m}+h)} 4\beta p. \tag{7.30}$$

7. Note from the definition of $\mathbf{m}$ in (1.16) that, up to a correction factor of $1 + o_n(1)$, $\mathbf{m}$ solves the equation $J(\mathbf{m}) = 0$ with

$$J_{p,\beta,h}(a) = 2\lambda\left(a + \tfrac{h}{p}\right) + \log\left(\tfrac{1-a}{1+a}\right), \qquad \lambda = \beta p, \tag{7.31}$$

i.e.,

$$\tfrac{1+\mathbf{m}}{1-\mathbf{m}} = e^{2\lambda\left(\mathbf{m}+\frac{h}{p}\right)}. \tag{7.32}$$

Hence (7.30) amounts to $\mathbf{m} + \frac{h}{p} < -\frac{\log(4\lambda-1)}{2\lambda}$. The latter in turn amounts to $J(-\frac{\log(4\lambda-1)}{2\lambda} - \frac{h}{p}) < 0$ (recall Lemma 2.4), i.e.,

$$4\lambda - 1 > \frac{1 + \frac{\log(4\lambda-1)}{2\lambda} + \frac{h}{p}}{1 - \frac{\log(4\lambda-1)}{2\lambda} - \frac{h}{p}}. \tag{7.33}$$

Abbreviate $\omega = 1 - \frac{\log(4\lambda-1)}{2\lambda} - \frac{h}{p}$. Then (7.33) holds when $\omega > \frac{1}{2\lambda}$. Hence we require that

$$\frac{h}{p} < \chi^*(\lambda), \qquad \chi^*(\lambda) = 1 - \frac{1}{2\lambda} - \frac{\log(4\lambda - 1)}{2\lambda}. \tag{7.34}$$

This is the constraint in (1.32).

8. To conclude, we summarise the implication of the contraction of the process $|W_1^t|$. The probability in (7.16) is equal to $1 - O_n(n^{\frac{5}{6}-\frac{6}{7}})$ for $|W_1^t| > n^{6/7}$, and is strictly larger than $\frac{1}{2}$ for $|W_1^t| \leq n^{6/7}$. Furthermore, from (7.12) we know that the rate at which W_1^t shrinks is ≥ 1. This allows us to ensure that sufficiently many steps are made by time $2n$ to allow W_1^t to contract to the empty set. In particular, the number steps taken by W_1^t up to time $2n$ is bounded from below by a Poisson point process $N(t)$ with unit rate, for which we have

$$\mathbb{P}\left[N(2n) \le \tfrac{3n}{2}\right] \le 2n \frac{(2n)^{\left(\frac{3n}{2}\right)} e^{-2n}}{\left(\frac{3n}{2}\right)!} \le 2n \left(\tfrac{4n}{3}\right)^{\left(\frac{3n}{2}\right)} e^{-\frac{n}{2}} \le 1.07^{\left(-\frac{n}{2}\right)}. \tag{7.35}$$

In other words, with probability exponentially close to 1, we have that at least $3n/2$ jumps are made in time $2n$. To bound the probability that W_1^t has not converged to the empty set, note that this probability decreases in the number of transitions made by W_1^t. Therefore, without loss of generality, we may assume that $\frac{3n}{2}$ transitions were made, and that we start with $|W_1^0| = n$. We claim that, with high probability, in time $2n$, W_1^t takes at most $\frac{100n}{\log n}$ increasing steps (i.e., $i \to i+1$) in the interval $[n^{5/6}, n]$. Indeed, note that the probability of the latter occurring is less than

$$2^M O\big(n^{-1/42}\big)^{\frac{100n}{\log n}} = O\big(e^{-n}\big). \tag{7.36}$$

It follows that at least $\frac{n}{2}[1+o_n(1)]$ steps are taken in the interval $[0, n^{5/6}]$. But then, using (7.16), we have that the probability of an increasing step is at most $\frac{1}{2} - \epsilon$ for some $\epsilon > 0$, and therefore the probability of that event is at most

$$2^{\frac{n}{2}[1+o_n(1)]} \left(\tfrac{1}{2}+\epsilon\right)^{\frac{n}{4}[1+o_n(1)]} \left(\tfrac{1}{2}-\epsilon\right)^{\frac{n}{4}[1+o_n(1)]} = 4^{\frac{n}{4}[1+o_n(1)]} \left(\tfrac{1}{4}-\epsilon^2\right)^{\frac{n}{4}[1+o_n(1)]} = \left(1-4\epsilon^2\right)^{\frac{n}{4}[1+o_n(1)]}. \tag{7.37}$$

Finally, observing that in the entire proof so far, the largest probability for any of the bounds not to hold is $O(e^{n^{-2/3}})$ (see (7.13) and the paragraph following (7.16)), we get

$$\mathbb{P}\left[|W_1^t| > 0\right] \le O\big(e^{-n^{2/3}}\big) \tag{7.38}$$

and so the claim of the lemma follows. □

7.2. **Long-term scheme.**

Corollary 7.3 (Long-term coupling). *With $\mathbb{P}_{\mathrm{ER}_n(p)}$-probability tending to 1 as $n \to \infty$, there is a coupling of $\{\xi_t\}_{t\ge0}$ and $\{\tilde{\xi}_{\tilde{t}}\}_{t\ge0}$, and there are times t and $\tilde{t}$ with $\max(t,\tilde{t}) < e^{n\Gamma^\star - \delta n}$, such that*

$$\mathbb{P}\big[\xi_t \ne \tilde{\xi}_{\tilde{t}}\big] \le e^{-n\delta + O(n^{2/3})}. \tag{7.39}$$

Proof. Define $\{\tilde{\mathbf{s}}_i\}_{i\ge0}$ in an analogous manner for $\{\tilde{\xi}_t\}_{t\ge0}$. Then we can define a coupling of $\{\xi_t\}_{t\ge0}$ and $\{\tilde{\xi}_t\}_{t\ge0}$ as follows. For $i \ge 0$ and $0 \le s \le 2n$, couple $\xi_{\mathbf{s_i}+s}$ and $\tilde{\xi}_{\tilde{\mathbf{s}}_\mathbf{i}+s}$ as described in Lemma 7.1. For times $t \in (\mathbf{s_i}+2n, \mathbf{s_{i+1}})$ and $\tilde{t} \in (\tilde{\mathbf{s}}_\mathbf{i}+2n, \tilde{\mathbf{s}}_\mathbf{i+1})$, let $\{\xi_t\}_{t\ge0}$ and $\{\tilde{\xi}_{\tilde{t}}\}_{t\ge0}$ run independently of each other. Terminate this coupling at the first time t such that $t = \mathbf{s_i} + s$ for some $s \le 2n$ and $\xi_t = \tilde{\xi}_{\tilde{t}}$ with $\tilde{t} = \tilde{\mathbf{s}}_\mathbf{i} + s$, from which point onward we simply let $\xi_t = \tilde{\xi}_{\tilde{t}}$. It is easy to see that the coupling above is an attempt at repeating the coupling scheme of Lemma 7.1 until the paths of the two processes have crossed. To avoid having to wait until both processes are in $A_\mathbf{M}$ at the same time, the coupling defines a joint distribution of ξ_t and $\tilde{\xi}_{\tilde{t}}$.

Note that, by Lemma 5.4, with probability of at least $1 - e^{-\delta n + O(n^{2/3})}$, $\{\xi_t\}_{t\ge0}$ will visit $A_\mathbf{M}$ at least $e^{\Gamma^\star n - \delta n}$ times before reaching $A_\mathbf{S}$, for any $\delta > 0$. The same statement is true for $\{\tilde{\xi}_{\tilde{t}}\}_{\tilde{t}\ge0}$. Assuming that the aforementioned event holds for both ξ_t and $\tilde{\xi}_{\tilde{t}}$, the probability that the coupling does not succeed (i.e., the two trajectories do not intersect as described earlier) is at most

$$\left[O\big(e^{n^{-2/3}}\big)\right]^{e^{(\Gamma^\star-\delta)n}}. \tag{7.40}$$

Therefore, the probability that the coupling does not succeed before either of $\{\xi_t\}_{t\ge0}$ or $\{\tilde{\xi}_{\tilde{t}}\}_{\tilde{t}\ge0}$ reaches $A_\mathbf{S}$ is at most $e^{-\delta n + O(n^{2/3})}$. □

8. Proof of the main metastability theorem

In the section we prove Theorem 1.4.

Proof. The key is to show that for any $\xi_0^u \in A_{\mathbf{M}^u}$, $\xi_0 \in A_{\mathbf{M}}$ and $\xi_0^l \in A_{\mathbf{M}^l}$,

$$\mathbb{E}_{\xi_0^u}\left[\tau_{\mathbf{S}^u}\right] \leq \mathbb{E}_{\xi_0}\left[\tau_{\mathbf{S}}\right] \leq \mathbb{E}_{\xi_0^l}\left[\tau_{\mathbf{S}^l}\right]. \tag{8.1}$$

Note that, by Lemma 2.2, $\mathbb{E}_{\xi_0}[\tau_{\mathbf{S}}]$ is the same for all $\xi_0 \in A_{\mathbf{M}}$ up to a multiplicative factor of $1+o_n(1)$. Therefore it is suffices to find *some* convenient $\xi \in A_{\mathbf{M}}$ for which we can prove the aforementioned theorem.

1. Our proof follows four steps:

(1) Starting from the initial distribution $\mu_{A_{\mathbf{M}}}$ on the set $A_{\mathbf{M}}$, the trajectory segment taken by ξ_t from ξ_0 to ξ_τ with $\tau = \min(\tau_{\mathbf{M}}, \tau_{\mathbf{S}})$, can be coupled to the analogous trajectory segments taken by ξ_t^l and ξ_t^u, starting in $A_{\mathbf{M}^l}$ and $A_{\mathbf{M}^u}$, respectively, and this coupling can be done in such a way that the following two conditions hold:
   (a) If ξ_t reaches $A_{\mathbf{S}}$ before returning to $A_{\mathbf{M}}$ (i.e., $\tau_{\mathbf{S}} < \tau_{\mathbf{M}}$), then ξ_t^u reaches $A_{\mathbf{S}^u}$ before returning to $A_{\mathbf{M}^u}$.
   (b) If ξ_t returns to $A_{\mathbf{M}}$ before reaching $A_{\mathbf{S}}$ (i.e., $\tau_{\mathbf{M}} < \tau_{\mathbf{S}}$), then ξ_t^l returns to $A_{\mathbf{M}^l}$ before reaching $A_{\mathbf{S}^l}$.

(2) We show that if ξ_t has initial distribution $\mu_{A_{\mathbf{M}}}$ and $\tau_{\mathbf{M}} < \tau_{\mathbf{S}}$, then upon returning to $A_{\mathbf{M}}$ the distribution of ξ_t is once again given by $\mu_{A_{\mathbf{M}}}$. This implies that the argument in Step (1) can be applied repeatedly, and that the number of returns ξ_t makes to $A_{\mathbf{M}}$ before reaching $A_{\mathbf{S}}$ is bounded from below by the number of returns ξ_t^u makes to $A_{\mathbf{M}^u}$ before reaching $A_{\mathbf{S}^u}$, and is bounded from above by the number of returns ξ_t^l makes to $A_{\mathbf{M}^l}$ before reaching $A_{\mathbf{S}^l}$.

(3) Using Lemma 3.7, we bound the time between unsuccessful excursions, i.e., the expected time it takes for ξ_t, when starting from $\mu_{A_{\mathbf{M}}}$, to return to $A_{\mathbf{M}}$, given that $\tau_{\mathbf{M}} < \tau_{\mathbf{S}}$. This bound is used together with the outcome of Step (2) to obtain the bound

$$\mathbb{E}^u_{\mu_{A_{\mathbf{M}^u}}}\left[\tau^u_{\mathbf{S}}\right] \leq \mathbb{E}_{\mu_{A_{\mathbf{M}}}}\left[\tau_{\mathbf{S}}\right] \leq \mathbb{E}^l_{\mu_{A_{\mathbf{M}^l}}}\left[\tau^l_{\mathbf{S}}\right]. \tag{8.2}$$

Here, the fact that the conditional average return time is bounded by some large constant rather than 1 does not affect the sandwich in (8.2), because the errors coming from the perturbation of the magnetic field in the Curie-Weiss model are polynomial in n (see below).

(4) We complete the proof by using Lemma 2.2 and showing that, for any distribution μ_0 restricted to $A_{\mathbf{M}}$,

$$\mathbb{E}_{\mu_0}\left[\tau_{\mathbf{S}}\right] = [1+o_n(1)]\, \mathbb{E}_{\mu_{A_{\mathbf{M}}}}\left[\tau_{\mathbf{S}}\right]. \tag{8.3}$$

2. Before we turn to the proof of these steps, we explain how the bound on the exponent in the prefactor of Theorem 1.4 comes about. Return to (2.4). The magnetic field h is perturbed to $h \pm (1+\epsilon)\log(n^{11/6})/n$. We need to show how this affects the formulas for the average crossover time in the Curie-Weiss model. For this we use the computations carried out in [4, Chapter 13]. According to [4, Eq. (13.2.4)] we have, for any $\xi \in A_{\mathbf{M}_n}$ and any $\epsilon > 0$,

$$\mathbb{E}_\xi\left[\tau_{A_{\mathbf{S}_n}}\right] = [1+o_n(1)]\, \frac{2}{1-\mathbf{t}}\, e^{\beta n[R_n(\mathbf{t})-R_n(\mathbf{m})]}\, \frac{1}{n}\, \mathbf{S}_n \tag{8.4}$$

with

$$\mathbf{S}_n = \sum_{\substack{a,a' \in \Gamma_n \\ |a-\mathbf{t}|<\epsilon,\, |a'-\mathbf{m}|<\epsilon}} e^{\beta n[R_n(\mathbf{a})-R_n(\mathbf{t})]-\beta n[R_n(a')-R_n(\mathbf{m})]}, \tag{8.5}$$

where R_n is the free energy defined by $R_n' = -J_n/2\beta$ (recall (1.20)). (Here we suppress the dependence on β, h and note that (8.4) carries an extra factor $\frac{1}{n}$ because [4, Chapter 13] considers a discrete-time dynamics where at every unit of time a single spin is drawn uniformly at random and is flipped with a probability that is given by the right-hand side of (1.6).) According to [4, Eq. (13.2.5)–(13.2.6)] we have

$$I_n(a) - I(a) = [1+o_n(1)]\frac{1}{2n}\log\left(\tfrac{1}{2}\pi n(1-a^2)\right), \qquad a \in [-1,1], \tag{8.6}$$

so that

$$e^{\beta n[R_n(a)-R(a)]} = [1+o_n(1)]\sqrt{\tfrac{1}{2}\pi n(1-a^2)}, \qquad a \in [-1,1]. \tag{8.7}$$

where R is the limiting free energy defined by $R' = -J/2\beta$ (recall (1.27)). Inserting (8.7) into (8.4), we get

$$\mathbb{E}_\xi\left[\tau_{A_{\mathbf{S}_n}}\right] = [1+o_n(1)]\,\frac{2}{1-\mathbf{t}}\sqrt{\frac{1-\mathbf{t}^2}{1-\mathbf{m}^2}}\,e^{\beta n[R(\mathbf{t})-R(\mathbf{m})]}\,\frac{1}{n}\,\mathbf{S}_n^* \tag{8.8}$$

with

$$\mathbf{S}_n^* = \sum_{\substack{a,a'\in\Gamma_n\\ |a-\mathbf{t}|<\epsilon,\,|a'-\mathbf{m}|<\epsilon}} e^{\beta n[R(\mathbf{a})-R(\mathbf{t})]-\beta n[R(a')-R(\mathbf{m})]}, \tag{8.9}$$

Finally, according to [4, Eq. (13.2.9)–(13.2.11)] we have, with the help of a Gaussian approximation,

$$\lim_{n\to\infty}\frac{1}{n}\,\mathbf{S}_n^* = \frac{\pi}{2\beta\sqrt{[R''(\mathbf{m})[-R''(\mathbf{t})]}}. \tag{8.10}$$

Putting together (8.8) and (8.10), we see how Theorem 1.3 arises as the correct formula for the Curie-Weiss model.

3. The above computations are for β, h fixed and $p=1$. We need to investigate what changes when $p\in(0,1)$, β is fixed, but h depends on n:

$$h_n = h \pm (1+\epsilon)\,\frac{\log(n^{11/6})}{n}. \tag{8.11}$$

We write R_n^n to denote R_n when h is replaced by h_n. In the argument in [4] leading up to (8.4), the approximation only enters through the prefactor. But since $h_n\to h$ as $n\to\infty$, the perturbation affects the prefactor only by a factor $1+o_n(1)$. Since h plays no role in (8.6) and $R_n^n(a)-R^n(a) = \frac{1}{\beta}[I_n(a)-I(a)]$ (recall (1.19) and (1.26)), we get (8.8) with exponent $\beta n[R^n(\mathbf{t})-R^n(\mathbf{m})]$ and (8.9) with exponent $\beta n[R^n(\mathbf{a})-R^n(\mathbf{t})]-\beta n[R^n(a')-R^n(\mathbf{m})]$. The latter affects the Gaussian approximation behind (8.10) only by a factor $1+o_n(1)$. However, the former leads to an error term in the exponent, compared to the Curie Weiss model, that equals

$$\begin{aligned}\beta n[R^n(\mathbf{t})-R^n(\mathbf{m})]-\beta n[R(\mathbf{t})-R(\mathbf{m})] &= \beta n\int_{\mathbf{m}}^{\mathbf{t}} da\,[(R^n)'(a)-R'(a)]\\ = \beta n\int_{\mathbf{m}}^{\mathbf{t}} da\,[-(h_n-h)] &= \beta(\mathbf{t}-\mathbf{m})\,n(h-h_n) = \mp\beta(\mathbf{t}-\mathbf{m})\,(1+\epsilon)\log(n^{11/6}).\end{aligned} \tag{8.12}$$

The exponential of this equals $n^{\mp\beta(\mathbf{t}-\mathbf{m})\,(1+\epsilon)(11/6)}$, which proves Theorem 1.4 with the bound in (1.34) because ϵ is arbitrary.

Proof of Step (1): This step is a direct application of Lemma 4.3.

Proof of Step (2): Let $\xi_0 \overset{d}{=} \mu_{A_\mathbf{M}}$, and recall that $\tau_\mathbf{M}$ is the first return time of ξ_t to $A_\mathbf{M}$ once the initial state ξ_0 has been left. We want to show that $\xi_{\tau_\mathbf{M}} \overset{d}{=} \mu_{A_\mathbf{M}}$ or, in other words, that $\mathbb{P}_{\mu_{A_\mathbf{M}}}[\xi_{\tau_\mathbf{M}}=\sigma] = \mu_{A_\mathbf{M}}(\sigma)$ for any $\sigma\in A_\mathbf{M}$. To facilitate the argument, we begin by defining the set of all finite permissible trajectories $\mathscr{T}$, i.e.,

$$\mathscr{T} = \bigcup_{N\in\mathbb{N}}\left\{\gamma=\{\gamma_i\}_{i=0}^N \in S_n^N:\ ||\gamma_i|-|\gamma_{i+1}|| = 1\ \forall\, 0\le i\le N-1\right\}. \tag{8.13}$$

Let $\gamma\in\mathscr{T}$ be any finite trajectory beginning at $\gamma_0\in A_\mathbf{M}$, ending at $\gamma_{|\gamma|-1}=\sigma\in A_\mathbf{M}$, and satisfying $\gamma_i\notin A_\mathbf{M}$ for $0<i<|\gamma|-1$. Then the probability that ξ_t follows the trajectory γ is given by

$$\begin{aligned}\mathbb{P}[\xi_t \text{ follows } \gamma] &= \mu_{A_\mathbf{M}}(\gamma_0)\,P(\gamma_0,\gamma_1)\times\cdots\times P\left(\gamma_{|\gamma|-2},\sigma\right)\\ &= \frac{1}{\mu(A_\mathbf{M})}\mu(\gamma_0)\,P(\gamma_0,\gamma_1)\times\cdots\times P\left(\gamma_{|\gamma|-2},\sigma\right)\\ &= \frac{1}{\mu(A_\mathbf{M})}\mu(\sigma)\,P\left(\sigma,\gamma_{|\gamma|-2}\right)\times\cdots\times P(\gamma_1,\gamma_0)\\ &= \mu_{A_\mathbf{M}}(\sigma)\,P\left(\sigma,\gamma_{|\gamma|-2}\right)\times\cdots\times P(\gamma_1,\gamma_0),\end{aligned} \tag{8.14}$$

where the third line follows from reversibility. Thus, if we let $\mathscr{T}(\sigma)$ be the set of all trajectories in $\mathscr{T}$ that begin in $A_{\mathbf{M}}$, end at σ, and do not visit $A_{\mathbf{M}}$ in between, then we get

$$\begin{aligned}\mathbb{P}_{\mu_{A_{\mathbf{M}}}}\left[\xi_{\tau_{\mathbf{M}}}=\sigma\right] &= \sum_{\gamma\in\mathscr{T}(\sigma)} \mu_{A_{\mathbf{M}}}(\sigma)\, P\left(\sigma,\gamma_{|\gamma|-2}\right)\times\cdots\times P\left(\gamma_1,\gamma_0\right)\\ &= \mu_{A_{\mathbf{M}}}(\sigma)\,\mathbb{P}_\sigma\left[\tau_{\mathbf{M}}<\infty\right]\\ &= \mu_{A_{\mathbf{M}}}(\sigma),\end{aligned} \tag{8.15}$$

where we use recurrence and the law of total probability, since the trajectories in $\mathscr{T}(\sigma)$, when reversed, give all possible trajectories that start at $\sigma\in A_{\mathbf{M}}$ and return to $A_{\mathbf{M}}$ in a finite number of steps. This shows that if ξ_t has initial distribution $\mu_{A_{\mathbf{M}}}$, then it also has this distribution upon every return to $A_{\mathbf{M}}$.

We can now define a segment-wise coupling of the trajectory taken by ξ_t with the trajectories taken by ξ_t^u and ξ_t^l. First, we define the subsets of trajectories that start and end in particular regions of the state space: (i) $\mathscr{T}_{\sigma,L,K}$ is the set of trajectories that start at a particular configuration σ and end in A_K without ever visiting A_K or A_L in between, for some $K<L$; (ii) $\mathscr{T}_{\sigma,L,L}$ is the set of trajectories that start at some σ and end in A_L without ever visiting A_K or A_L in between; (iii) $\mathscr{T}_{\sigma,L}$ is the union of the two aforementioned sets. In explicit form,

$$\begin{aligned}\mathscr{T}_{\sigma,L,K} &= \left\{\gamma\in\mathscr{T}\colon \gamma_0=\sigma, \gamma_{|\gamma|-1}\in A_K, K<|\gamma_j|<L\ \forall\, 0<j<|\gamma|-1\right\},\\ \mathscr{T}_{\sigma,L,L} &= \left\{\gamma\in\mathscr{T}\colon \gamma_0=\sigma, \gamma_{|\gamma|-1}\in A_L, K<|\gamma_j|<L\ \forall\, 0<j<|\gamma|-1\right\},\\ \mathscr{T}_{\sigma,L} &= \mathscr{T}_{\sigma,L,K}\cup\mathscr{T}_{\sigma,L,L}.\end{aligned} \tag{8.16}$$

By Step (1), for any $\xi_0^l\in A_{\mathbf{M}^l}$ and $\xi_0^u\in A_{\mathbf{M}^u}$,

$$\mathbb{P}^l_{\xi_0^l}\left[\mathscr{T}_{\xi_0^l,\mathbf{S}^l,\mathbf{S}^l}\right]\le\mathbb{P}_{\xi_0}\left[\mathscr{T}_{\xi_0,\mathbf{S},\mathbf{S}}\right]\le\mathbb{P}^u_{\xi_0^u}\left[\mathscr{T}_{\xi_0^u,\mathbf{S}^u,\mathbf{S}^u}\right]. \tag{8.17}$$

It is clear that the two probabilities at either end of (8.17) are independent of the starting points ξ_0^l and ξ_0^u. By the argument given above, if for the probability in the middle $\xi_0\stackrel{d}{=}\mu_{A_{\mathbf{M}}}$, then each subsequent return to $A_{\mathbf{M}}$ also has this distribution. For this reason, we may define a coupling of the trajectories as follows.

Sample a trajectory segment γ^l from $\mathscr{T}_{\xi_0^l,\mathbf{S}^l}$ for the process ξ_t^l. If γ^l happens to be in $\mathscr{T}_{\xi_0^l,\mathbf{S}^l,\mathbf{S}^l}$, then by (8.17) we may sample a trajectory segment γ from $\mathscr{T}_{\xi_0,\mathbf{S},\mathbf{S}}$ for the process ξ_t, and a trajectory segment γ^u from $\mathscr{T}_{\xi_0^u,\mathbf{S}^u,\mathbf{S}^u}$ for the process ξ^u. Otherwise, $\gamma^l\in\mathscr{T}_{\xi_0^l,\mathbf{S}^l,\mathbf{M}^l}$, and we independently take $\gamma\in\mathscr{T}_{\xi_0,\mathbf{S},\mathbf{S}}$ with probability $\mathbb{P}_{\xi_0}[\mathscr{T}_{\xi_0,\mathbf{S},\mathbf{S}}]-\mathbb{P}^l_{\xi_0^l}[\mathscr{T}_{\xi_0^l,\mathbf{S}^l,\mathbf{S}^l}]$, and $\gamma\in\mathscr{T}_{\xi_0,\mathbf{S},\mathbf{M}}$ otherwise. If $\gamma\in\mathscr{T}_{\xi_0,\mathbf{S},\mathbf{S}}$, then sample γ^u from $\mathscr{T}_{\xi_0^u,\mathbf{S}^u,\mathbf{S}^u}$. Otherwise $\gamma\in\mathscr{T}_{\xi_0,\mathbf{S},\mathbf{M}}$, and so take independently $\gamma^u\in\mathscr{T}_{\xi_0^u,\mathbf{S}^u,\mathbf{S}^u}$ with probability $\mathbb{P}^u_{\xi_0^u}[\mathscr{T}_{\xi_0^u,\mathbf{S}^u,\mathbf{S}^u}]-\mathbb{P}_{\xi_0}[\mathscr{T}_{\xi_0,\mathbf{S},\mathbf{S}}]$, and $\gamma^u\in\mathscr{T}_{\xi_0^u,\mathbf{S}^u,\mathbf{M}^u}$ with the remaining probability. We glue together the sampling of segments leaving and returning to $A_{\mathbf{M}^l}/A_{\mathbf{M}}/A_{\mathbf{M}^u}$ with the next sampling of such segments. This results in trajectories for ξ^u, ξ, and ξ^l that reach $A_{\mathbf{S}^u}/A_{\mathbf{S}}/A_{\mathbf{S}^l}$, in that particular order.

Proof of Step (3) and Step (4): These two steps are immediate from Lemma 2.2 and Lemma 3.7.

□

Appendix A. Conditional average return time for inhomogeneous random walk

In this appendix we prove Lemma 3.7. In Appendices A.1–A.2 we compute the harmonic function and the conditional average return time for an arbitrary nearest-neighbour random walk on a finite interval. In Appendix A.3 we use these computations to prove the lemma.

A.1. **Harmonic function.** Consider a *nearest-neighbour* random walk on the set $\{0,\dots,N\}$ with strictly positive transition probabilities $p(x,x+1)$ and $p(x,x-1)$, $0<x<N$, and with 0 and N acting as absorbing boundaries. Let τ_0 and τ_N denote the first hitting times of 0 and N. The *harmonic function* is defined as

$$h_N(x)=\mathbb{P}_x(\tau_N<\tau_0),\quad 0\le x\le N, \tag{A.1}$$

where $\mathbb{P}_x$ is the law of the random walk starting from x. This is the unique solution of the recursion relation

$$h_N(x)=p(x,x+1)h_N(x+1)+p(x,x-1)h_N(x-1),\quad 0<x<N, \tag{A.2}$$

with boundary conditions

$$h_N(0) = 0, \qquad h_N(N) = 1. \tag{A.3}$$

Since $p(x, x+1) + p(x, x-1) = 1$, the recursion can be written as

$$p(x, x+1)[h_N(x+1) - h_N(x)] = p(x, x-1)[h_N(x) - h_N(x-1)]. \tag{A.4}$$

Define $\Delta h_N(x) = h_N(x+1) - h_N(x)$, $0 \leq x < N$. Iteration gives

$$\Delta h_N(x) = \pi[1, x]\, \Delta h_N(0), \quad 0 \leq x < N, \tag{A.5}$$

where we define

$$\pi(I) = \prod_{z \in I} \frac{p(z, z-1)}{p(z, z+1)}, \quad I \subseteq \{1, \dots, N-1\}, \tag{A.6}$$

with the convention that the empty product equals 1. Since $h_N(0) = 0$, we have

$$h_N(x) = \sum_{z=0}^{x-1} \Delta h_N(z) = \left(\sum_{z=0}^{x-1} \pi[1, z] \right) \Delta h_N(0), \qquad 0 < x \leq N. \tag{A.7}$$

Put $C = \Delta h_N(0)$, and abbreviate

$$R(x) = \sum_{z=0}^{x-1} \pi[1, z], \quad 0 \leq x \leq N. \tag{A.8}$$

Since $h_N(N) = 1$, we have $C = 1/R(N)$. Therefore we arrive at

$$h_N(x) = \frac{R(x)}{R(N)}, \quad 0 \leq x \leq N. \tag{A.9}$$

Remark A.1. For simple random walk we have $p(x, x \pm 1) = \frac{1}{2}$, hence $\pi[1, x] = 1$ and $R(x) = x$, and so

$$h_N(x) = \frac{x}{N}, \quad 0 \leq x \leq N, \tag{A.10}$$

which is the standard gambler's ruin formula.

A.2. **Conditional average hitting time.** We are interested in the quantity

$$e_N(x) = \mathbb{E}_x(\tau_N \mid \tau_N < \tau_0), \qquad 0 < x \leq N. \tag{A.11}$$

The conditioning amounts to taking the *Doob transformed* random walk, which has transition probabilities

$$q(x, x \pm 1) = p(x, x \pm 1)\, \frac{h_N(x \pm 1)}{h_N(x)}. \tag{A.12}$$

We have the recursion relation

$$e_N(x) = 1 + q(x, x+1) e_N(x+1) + q(x, x-1) e_N(x-1), \quad 0 < x < N, \tag{A.13}$$

with boundary conditions

$$e_N(N) = 0, \quad e_N(1) = 1 + e_N(2). \tag{A.14}$$

Putting $f_N(x) = h_N(x) e_N(x)$, we get the recursion

$$f_N(x) = h_N(x) + p(x, x+1) f_N(x+1) + p(x, x-1) f_N(x-1), \quad 0 < x < N, \tag{A.15}$$

which can be rewritten as

$$p(x, x+1)[f_N(x+1) - f_N(x)] = p(x, x-1)[f_N(x) - f_N(x-1)] - h_N(x). \tag{A.16}$$

Define $\Delta f_N(x) = f_N(x+1) - f_N(x)$, $0 < x < N$. Iteration gives

$$\Delta f_N(x) = \pi(1, x]\, \Delta f_N(1) - \chi(1, x], \quad 0 < x < N, \tag{A.17}$$

with

$$\chi(1, x] = \sum_{y=2}^{x} \pi(y, x]\, \frac{h_N(y)}{p(y, y+1)}, \quad 0 < x < N, \tag{A.18}$$

Since $f_N(N) = 0$, we have

$$f_N(x) = -\sum_{z=x}^{N-1} \Delta f_N(z) = \sum_{z=x}^{N-1} \chi(1,z] - \left(\sum_{z=x}^{N-1} \pi(1,z]\right) \Delta f_N(1), \quad 0 < x < N, \tag{A.19}$$

or

$$e_N(x) = \frac{1}{h_N(x)} \sum_{z=x}^{N-1} \chi(1,z] - \frac{1}{h_N(x)} \left(\sum_{z=x}^{N-1} \pi(1,z]\right) \Delta f_N(1), \quad 0 < x < N. \tag{A.20}$$

Put $C = \Delta f_N(1)$, and abbreviate

$$A(x) = \sum_{z=x}^{N-1} \pi(1,z], \quad B(x) = \sum_{z=x}^{N-1} \chi(1,z], \quad 0 < x \le N. \tag{A.21}$$

Then

$$e_N(x) = \frac{1}{h_N(x)} \big[B(x) - CA(x)\big]. \tag{A.22}$$

Since $e_N(1) = 1 + e_N(2)$, we have

$$C = \frac{[h_N(2)B(1) - h_N(1)B(2)] - h_N(1)h_N(2)}{h_N(2)A(1) - h_N(1)A(2)}. \tag{A.23}$$

Abbreviate

$$\bar{R}(x) = \sum_{z=0}^{x-1} \pi(1,z], \quad \bar{S}(x) = \sum_{z=0}^{x-1} \chi(1,z], \quad 0 < x \le N. \tag{A.24}$$

Then

$$A(x) = \bar{R}(N) - \bar{R}(x), \quad B(x) = \bar{S}(N) - \bar{S}(x), \quad 0 < x < N. \tag{A.25}$$

Note that $h_N(x) = R(x)/R(N) = \bar{R}(x)/\bar{R}(N)$, because $\pi[1,z] = \pi(1)\pi(1,z]$ and a common factor $\pi(1)$ drops out. Note further that $\bar{R}(1) = 1$, $\bar{R}(2) = 2$, while $\bar{S}(1) = \bar{S}(2) = 0$. Therefore

$$C = \frac{\bar{S}(N)}{\bar{R}(N)} - \frac{2}{\bar{R}(N)^2}. \tag{A.26}$$

Therefore we arrive at

$$e_N(x) = \bar{S}(N) - \frac{\bar{R}(N)}{\bar{R}(x)}\, \bar{S}(x) + \frac{2}{\bar{R}(x)} - \frac{2}{\bar{R}(N)}, \quad 0 < x \le N. \tag{A.27}$$

Abbreviating

$$\bar{T}(x) = \bar{S}(x)\bar{R}(N) = \sum_{z=0}^{x-1} \sum_{y=2}^{z} \frac{\pi(y,z]}{p(y,y+1)}\, \bar{R}(y), \quad \bar{U}(x) = \frac{\bar{T}(x) - 2}{\bar{R}(x)}, \tag{A.28}$$

we can write

$$e_N(x) = \bar{U}(N) - \bar{U}(x), \quad 0 < x \le N. \tag{A.29}$$

Remark A.2. For simple random walk we have $p(x, x\pm1) = \frac{1}{2}$, $\pi(y,z] = 1$, $\bar{R}(x) = x$, $\bar{S}(x) = \frac{1}{3N}(x^3 - 7x + 6)$ and $\bar{U}(x) = \frac{1}{3}(x^2 - 7)$, and so

$$e_N(x) = \tfrac{1}{3}\,(N^2 - x^2), \qquad 0 < x \le N. \tag{A.30}$$

This is to be compared with the *unconditional* average hitting time $\mathbb{E}_x(\tau) = x(N-x)$, $0 \le x \le N$, where $\tau = \tau_0 \wedge \tau_N$ is the first hitting time of $\{0, N\}$.

A.3. **Application to spin-flip dynamics.** We will use the formulas in (A.6), (A.24) and (A.28)–(A.29) to obtain an upper bound on the conditional return time to the metastable state. This bound will be sharp enough to prove Lemma 3.7. We first do the computation for the complete graph (Curie-Weiss model). Afterwards we turn to the Erdős-Rényi Random Graph (our spin-flip dynamics).

A.3.1. *Complete graph.* We monitor the magnetization of the *continuous-time* Curie-Weiss model by looking at the magnetization at the times of the spin-flips. This gives a *discrete-time* random walk on the set Γ_n defined in (1.15). This set consists of $n+1$ sites. We first consider the excursions to the *left* of $\mathbf{m}_n$ (recall (1.16)). After that we consider the excursions to the *right*.

1. For the Curie-Weiss model we have (use the formulas in Lemma 3.4 without the error terms)

$$\sigma \in A_k\colon \qquad \sum_{\xi\in A_{k+1}} r(\sigma,\xi) = (n-k)\, e^{-2\beta[\theta_k]_+}, \qquad \sum_{\xi\in A_{k-1}} r(\sigma,\xi) = k\, e^{-2\beta[-\theta_k]_+}, \tag{A.31}$$

where $\theta_k = p(1-\frac{2k}{n}) - h$. Hence, the quotient of the rate to move downwards, respectively, upwards in magnetization equals

$$Q(k) = \frac{\sum_{\xi\in A_{k-1}} r(\sigma,\xi)}{\sum_{\xi\in A_{k+1}} r(\sigma,\xi)} = \frac{k}{n-k}\, e^{2\beta([\theta_k]_+ - [-\theta_k]_+)}. \tag{A.32}$$

It is convenient to change variables by writing $k = \frac{n}{2}(a_k+1)$, so that $\theta_k = -pa_k - h$. The metastable state corresponds to $k = \mathbf{M}_n = \frac{n}{2}(\mathbf{m}_n+1)$, i.e., $a_k = \mathbf{m}_n$. We know from (1.16)–(1.15) that $\mathbf{m}_n$ is the smallest solution of the equation $J_n(\mathbf{m}_n) = 0$ (rounded off by $1/n$ to fall in Γ_n). Hence $\mathbf{m}_n = \mathbf{m} + O(1/n)$ with $\mathbf{m}$ the smallest solution of the equation $J_{p,\beta,h}(\mathbf{m}) = 0$, satisfying $\frac{1-\mathbf{m}}{1+\mathbf{m}} = e^{-2\beta(p\mathbf{m}+h)}$ (recall (1.23)). Hence we can write (for ease of notation we henceforth ignore the error $O(1/n)$)

$$Q(k) = \frac{F(\mathbf{m}_n)}{F(a_k)}, \qquad F(a) = \frac{1-a}{1+a}\, e^{2\beta p a}. \tag{A.33}$$

Here, we use that $[\theta_k]_+ - [-\theta_k]_+ = \theta_k$, which holds because $0 = R'_{p,\beta,h}(\mathbf{m}) = -p\mathbf{m} - h + \beta^{-1} I'(\mathbf{m})$ with $I'(\mathbf{m}) < 0$ because $\mathbf{m} < 0$ (recall (1.27)), so that $-p\mathbf{m}_n - h > 0$ for n large enough, which implies that also $-pa - h > 0$ for all $a < \mathbf{m}_n$ for n large enough. We next note that (recall (1.27) and (2.20))

$$\frac{d}{da} \log\left[\frac{F(\mathbf{m}_n)}{F(a)}\right] = -2\left(\beta p - \frac{1}{1-a^2}\right) = -J'_{p,\beta,h}(a) = 2\beta R''_{p,\beta,h}(a) \geq \delta \quad \text{for some } \delta > 0, \tag{A.34}$$

where the inequality comes from the fact that $a \mapsto R_{p,\beta,h}(a)$ has a positive curvature that is bounded away from zero on $[-1,\mathbf{m}]$ (recall Figure 4).

2. We view the excursions to the left of $\mathbf{m}_n$ as starting from site N in the set $\{0,\dots,N\}$ with $N = \mathbf{M}_n = \frac{n}{2}(\mathbf{m}_n+1)$. From (A.28)–(A.29), we get

$$\begin{aligned} e_N(x) &= \sum_{z=0}^{N-1}\sum_{y=2}^{z} \frac{\pi(y,z]}{p(y,y+1)}\frac{\bar R(y)}{\bar R(N)} - \sum_{z=0}^{x-1}\sum_{y=2}^{z} \frac{\pi(y,z]}{p(y,y+1)}\frac{\bar R(y)}{\bar R(x)} + \frac{2}{\bar R(N)\bar R(x)}[\bar R(N) - \bar R(x)] \\ &\leq \sum_{z=x}^{N-1}\sum_{y=2}^{z} \frac{\pi(y,z]}{p(y,y+1)}\frac{\bar R(y)}{\bar R(N)} + \frac{2}{\bar R(x)} \\ &\leq 2\sum_{z=1}^{N-1}\sum_{y=2}^{z} \pi(y,z] + 2. \end{aligned} \tag{A.35}$$

Here, we use that $p(y,y+1) \geq \frac12$ and $1 = \bar R(0) \leq \bar R(y) \leq \bar R(N)$ for all $0 < y < N$ (recall (A.24) and note that $x \mapsto \bar R(x)$ is non-decreasing). The bound is independent of x. Using the estimate

$$Q(x) = \frac{p(x,x-1)}{p(x,x+1)} \leq e^{-\epsilon(N-x)/N}, \quad 0 < x < N, \quad \text{for some } \epsilon = \epsilon(\delta) > 0, \tag{A.36}$$

which comes from (A.34), we can estimate

$$\pi(y,z] \leq \prod_{x=y+1}^{z} e^{-\epsilon(N-x)/N} = \exp\left[-\epsilon \sum_{x=y+1}^{z} (N-x)/N\right], \qquad 0 \leq y \leq z < N, \tag{A.37}$$

from which it follows that

$$\sum_{z=1}^{N-1}\sum_{y=2}^{z} \pi(y,z] = O(N/\epsilon), \qquad N \to \infty. \tag{A.38}$$

Thus we arrive at

$$e_N(x) = O(N), \qquad N \to \infty, \quad \text{uniformly in } 0 < x < N. \tag{A.39}$$

To turn (A.39) into a tail estimate, we use the Chebyshev inequality: (A.39) implies that every N time units there is a probability at least c to hit N, for some $c > 0$ and uniformly in $0 < x < N$. Hence

$$\mathbb{P}_x(\tau_N \geq kN \mid \tau_N < \tau_0) \leq (1-c)^k \qquad \forall\, k \in \mathbb{N}_0. \tag{A.40}$$

3. For excursions to the right of $\mathbf{m}_n$ the argument is similar. Now $N = \mathbf{T}_n - \mathbf{M}_n = \frac{n}{2}(\mathbf{t}_n - \mathbf{m}_n)$ (recall (1.17)), and the role of 0 and N is interchanged. Both near 0 and near N the drift towards $\mathbf{M}_n$ vanishes *linearly* (because of the non-zero curvature). If we condition the random walk not to hit N, then the average hitting time of 0 starting from x is again $O(N)$, uniformly in x.

4. Returning from the discrete-time random walk to the continuous-time Curie-Weiss model, we note that order n spin-flips occur per unit of time. Since $N \asymp n$ as $n \to \infty$, (A.40) and its analogue for excursions to the right give that, uniformly in $\xi \in A_{\mathbf{M}_n}$,

$$\mathbb{P}_\xi\left[\tau_{A_{\mathbf{M}_n}} \geq k \mid \tau_{A_{\mathbf{M}_n}} < \tau_{A_{\mathbf{T}_n}}\right] \leq e^{-Ck} \qquad \forall\, k \in \mathbb{N}_0. \tag{A.41}$$

for some $C > 0$, which is the bound in (3.36).

A.3.2. *Erdős-Rényi Random Graph.* We next argue that the above argument can be extended to our spin-flip dynamics after taking into account that the rates to move downwards and upwards in magnetization are *perturbed by small errors*. In what follows we will write $p^{\mathrm{CW}}(x, x \pm 1)$ for the transition probabilities in the Curie-Weiss model and $p^{\mathrm{ER}}(x, x \pm 1)$ for the transition probabilities that serve as *uniform upper and lower bounds* for the transition probabilities in our spin-flip model. Recall that the latter actually depend on the configuration and not just on the magnetization, but Lemma 3.4 provides us with uniform bounds that allow us to *sandwich* the magnetization between the magnetizations of *two perturbed Curie-Weiss models*.

1. Suppose that

$$\frac{p^{\mathrm{ER}}(x, x-1)}{p^{\mathrm{ER}}(x, x+1)} = \frac{p^{\mathrm{CW}}(x, x-1)}{p^{\mathrm{CW}}(x, x+1)}[1 + O(N^{-1/2})]. \tag{A.42}$$

Then there exists a $C > 0$ large enough such that

$$\pi^{\mathrm{ER}}(y, z] \leq C\pi^{\mathrm{CW}}(y, z], \qquad 0 \leq y \leq z < N. \tag{A.43}$$

Indeed, as long as $z - y \leq C_1 N^{1/2}$ we have the bound in (A.43) (with C depending on C_1). On the other hand, if $z - y > C_1 N^{1/2}$ with C_1 large enough, then *the drift of the Curie-Weiss model sets in and overrules the error*: recall from (A.36) that the drift at distance $N^{1/2}$ from N is of order $N^{1/2}/N = N^{-1/2}$. It follows from (A.43) that (A.38)–(A.40) carry over, with suitably adapted constants, and hence so does (A.41).

2. To prove (A.42), we must show that (A.32) holds up to a multiplicative error $1 + O(n^{-1/2})$. In the argument that follows we assume that k is such that $\theta_k \geq \delta$ for some fixed $\delta > 0$. We comment later on how to extend the argument to other k values. Recall that $\theta_k = -pa_k - h$ and that $\theta_k \geq \delta > 0$ for all $a_k \in [-1, \mathbf{m}]$ for n large enough.

3. Let $\sigma \in A_k$ and $\sigma^v \in A_{k-1}$, where σ^v is obtained from σ by flipping the sign at vertex $v \in \sigma$ from $+1$ to -1. Write the transition rate from σ to σ^v as

$$\begin{aligned} r(\sigma, \sigma^v) &= \exp\left(-\beta\left[2p(\tfrac{2k}{n} - 1) + 2h + \tfrac{2}{n}\big(\epsilon(\sigma, v) - \epsilon(\overline{\sigma}, v)\big)\right]_+\right) \\ &= \exp\left(-2\beta\left[-\theta_k + \tfrac{1}{n}\big(\epsilon(\sigma, v) - \epsilon(\overline{\sigma}, v)\big)\right]_+\right). \end{aligned} \tag{A.44}$$

Here, $2p(\frac{2k}{n} - 1) = \frac{2}{n}p[k - (n-k)]$ equals $\frac{2}{n}$ times the average under $\mathbb{P}_{\mathrm{ER}_n(p)}$ of $E(\sigma, v) - E(\overline{\sigma}, v)$, with $E(\sigma, v)$ the number of edges between the support of σ and v and $E(\overline{\sigma}, v)$ the number of edges between the support of $\overline{\sigma}$ and v (recall the notation in Definition 3.1), and $\epsilon(\sigma, v) - \epsilon(\overline{\sigma}, v)$ is an *error term* that arises from deviations of this average. Since $-\theta_k \leq -\delta$, the error terms are *not seen except* when they represent a large deviation of size at least δn. A union bound over all the vertices and all the configurations, in combination with Hoeffding's inequality, guarantees that, with $\mathbb{P}_{\mathrm{ER}_n(p)}$-probability tending to 1 as $n \to \infty$, for any σ

there are at most $(\log 2)/2\delta^2 = O(1)$ many vertices that can lead to a large deviation of size at least δn. Since $r(\sigma,\sigma^v) \leq 1$, we obtain

$$\sum_{v\in\sigma} r(\sigma,\sigma^v) = O(1) + [n-k-O(1)]\, e^{-2\beta[-\theta_k]_+}. \tag{A.45}$$

This is a refinement of (3.11).

4. Similarly, let $\sigma \in A_k$ and $\sigma^v \in A_{k+1}$, where σ^v is obtained from σ by flipping the sign at vertex $v \notin \sigma$ from -1 to $+1$. Write the transition rate from σ to σ^v as

$$\begin{aligned} r(\sigma,\sigma^v) &= \exp\Big(-\beta\left[2p(1-\tfrac{2k}{n}) - 2h + \tfrac{2}{n}\big(\epsilon(\overline{\sigma},v) - \epsilon(\sigma,v)\big)\right]_+\Big) \\ &= \exp\Big(-2\beta\left[\theta_k + \tfrac{1}{n}\big(\epsilon(\overline{\sigma},v) - \epsilon(\sigma,v)\big)\right]_+\Big). \end{aligned} \tag{A.46}$$

We *cannot remove* $[\cdot]_+$ when the error terms represent a large deviation of order δn. By the same argument as above, this happens for all but $(\log 2)/2\delta^2 = O(1)$ many vertices v. For all other vertices, we *can remove* $[\cdot]_+$ and write

$$r(\sigma,\sigma^v) = e^{-2\beta\theta_k}\, \exp\big(\tfrac{1}{n}\big(\epsilon(\overline{\sigma},v) - \epsilon(\sigma,v)\big)\big). \tag{A.47}$$

Next, we sum over v and use the inequality, valid for δ small enough,

$$e^{-(1+\delta)\frac{1}{M}|\sum_{i=1}^M a_i|} \leq \frac{1}{M}\sum_{i=1}^M e^{a_i} \leq e^{(1+\delta)\frac{1}{M}|\sum_{i=1}^M a_i|} \qquad \forall\, 0 \leq |a_i| \leq \delta,\ 1 \leq i \leq M. \tag{A.48}$$

This gives

$$\sum_{v\notin\sigma} r(\sigma,\sigma^v) = O(1) + [k-O(1)]\, e^{-2\beta\theta_k}\, e^{O(|S_n|)}, \qquad S_n = \frac{1}{[k-O(1)]}\frac{1}{n}\sum_{v\notin\sigma}\big(\epsilon(\overline{\sigma},v) - \epsilon(\sigma,v)\big). \tag{A.49}$$

We know from Lemma 3.2 that, with $\mathbb{P}_{\mathrm{ER}_n(p)}$-probability tending to 1 as $n\to\infty$,

$$|S_n| \leq \frac{cn^{3/2}}{[k-O(1)]n} \qquad \forall\, c > \sqrt{\tfrac{1}{8}\log 2}. \tag{A.50}$$

Since we may take $k \geq \frac{n}{3}(p-h)$ (recall (3.15)), we obtain

$$\sum_{v\notin\sigma} r(\sigma,\sigma^v) = O(1) + [k-O(1)]\, e^{-2\beta\theta_k}\, e^{O(n^{-1/2})}. \tag{A.51}$$

This is a refinement of (3.12).

5. The same argument works when we assume that k is such that $\theta_k \leq -\delta$ for some fixed $\delta > 0$: simply reverse the arguments in Steps 3 and 4. It therefore remains to explain what happens when $\theta_k \approx 0$, i.e., $a_k \approx -\frac{h}{p}$. We then see from (1.27) that $R'_{p,\beta,h}(a_k) \approx \beta^{-1}I'(a_k) < 0$, so that a_k lies in the interval $[\mathbf{t},0]$, which is *beyond* the top state (recall Fig. 4).

References

[1] A. Bianchi, A. Bovier and D. Ioffe, Sharp asymptotics for metastability in the random field Curie-Weiss model, Electron. J. Probab. 14 (2009) 1541–1603.

[2] A. Bianchi, A. Bovier and D. Ioffe, Pointwise estimates and exponential laws in metastable systems via coupling methods, Ann. Probab. 40 (2012) 339–379.

[3] A. Bovier, M. Eckhoff, V. Gayrard and M. Klein, Metastability in stochastic dynamics of disordered mean-field models, Probab. Theory Relat. Fields 119 (2001) 99–161.

[4] A. Bovier and F. den Hollander, *Metastability – A Potential-Theoretic Approach*, Grundlehren der mathematischen Wissenschaften 351, Springer, 2015.

[5] A. Bovier, S. Marello and E. Pulvirenti, Metastability for the dilute Curie-Weiss model with Glauber dynamics, preprint 23 December 2019.

[6] S. Dommers, Metastability of the Ising model on random regular graphs at zero temperature, Probab. Theory Relat. Fields 167 (2017) 305–324.

[7] S. Dommers, C. Giardinà and R. van der Hofstad, Ising models on power-law random graphs, J. Stat. Phys. 141 (2010) 638–660.

[8] S. Dommers, C. Giardinà and R. van der Hofstad, Ising critical exponents on random trees and graphs, Commun. Math. Phys. 328 (2014) 355–395.

[9] S. Dommers, F. den Hollander, O. Jovanovski and F.R. Nardi, Metastability for Glauber dynamics on random graphs, Ann. Appl. Probab. 27 (2017) 2130–2158.
[10] R. van der Hofstad, *Random Graphs and Complex Networks, Volume I*, Cambridge University Press, 2017.
[11] F. den Hollander and O. Jovanovski, Metastability on the hierarchical lattice, J. Phys. A: Math. Theor. 50 (2017) 305001.
[12] O. Jovanovski, Metastability for the Ising model on the hypercube, J. Stat. Phys. 167 (2017) 135–159.
[13] E. Olivieri and M.E. Vares, *Large Deviations and Metastability*, Encyclopedia of Mathematics and its Applications 100, Cambridge University Press, Cambridge, 2005.

Mathematical Institute, Leiden University, P.O. Box 9512, 2300 RA Leiden, The Netherlands.
E-mail address: denholla@math.leidenuniv.nl

Bank of Montreal, 100 King Street West, Toronto, ON M5X 1A9 Canada.
E-mail address: oliver.jovanovski@bmo.com