

# POINTWISE ESTIMATES AND EXPONENTIAL LAWS IN METASTABLE SYSTEMS VIA COUPLING METHODS<sup>1,2</sup>

BY ALESSANDRA BIANCHI, ANTON BOVIER AND DMITRY IOFFE

*Weierstrass-Institute for Applied Analysis and Stochastics, Institut für Angewandte Mathematik, Rheinische Friedrich–Wilhelms-Universität and Faculty of Industrial Engineering and Management, Technion*

We show how coupling techniques can be used in some metastable systems to prove that mean metastable exit times are almost constant as functions of the starting microscopic configuration within a “meta-stable set.” In the example of the Random Field Curie Weiss model, we show that these ideas can also be used to prove asymptotic exponentiality of normalized metastable escape times.

## 1. Introduction.

1.1. *The problem.* Metastable systems are characterized by the fact that the state space can be decomposed into several disjoint subsets, with the property that transition times between subspaces are long compared to characteristic mixing times within each subspace. The mathematically rigorous analysis of Markov processes exhibiting metastable behavior was first developed in the *large deviation theory* of Freidlin and Wentzell [8, 14]. This approach yields logarithmic asymptotics of transition times and other quantities of interest. Over the last decade, a *potential theoretic approach* [2, 5] to metastability was developed that in many instances yields more precise asymptotics, and in particular the exact prefactors of exponential terms.

In this work we study metastability for a class of stochastic Ising models. The main objective is to extend the potential theoretical approach for deriving asymptotics of transition times for processes starting from individual microscopic configurations, and, subsequently, for studying exponential scaling laws for these transition times.

So far the existing methods work well in the following situations:

(1) The process is strongly recurrent in the sense that it visits an individual atom of the state space in each metastable state many times with overwhelming probability *before* a metastable transition happens. This situation occurs, for ex-

---

Received March 2010; revised September 2010.

<sup>1</sup>Supported in part through a grant by the German–Israeli Foundation (GIF).

<sup>2</sup>Supported in part by the DFG in the SFB 611 and the Hausdorff Center for Mathematics. *MSC2010 subject classifications.* Primary 82C44, 60K35; secondary 60G70.

*Key words and phrases.* Disordered system, random field Curie–Weiss model, Glauber dynamics, metastability, potential theory, coupling, exponential law.

ample, in Markov chains with finite state space, and on discrete state space, such as  $\mathbb{Z}^d$ , in the presence of a confining potential.

(2) In models where strong symmetries allow the analysis of the dynamics through a *lumped chain* that satisfies the requirements of (1). This situation occurs, for example, in mean field models such as the Curie–Weiss model [7] and the Curie–Weiss model with random magnetic fields that take only finitely many values [4, 12].

(3) In situations where the process returns often to small neighborhoods,  $\mathcal{O}_\varepsilon(x)$  of points,  $x$ , in a metastable state where the oscillations of harmonic functions on these neighborhoods can be made arbitrarily small. This is the case in finite and some infinite-dimensional diffusion processes [6, 10].

One would expect that situation (3) also arises in a wide variety of stochastic Ising models or stochastic particle systems exhibiting metastable behavior. *Proving* the respective regularity properties of microscopic harmonic functions appears, however, to be a difficult issue in general.

The purpose of the present paper is to develop an approach to this problem via coupling techniques that allow to cover at least some interesting situations.

A key idea of the potential theoretic approach is to express quantities of physical interest in terms of *capacities* and to use *variational principles* to compute the latter. A fundamental identity used systematically in this approach is a representation formula for the Green’s function,  $g_B(x, y)$ , with Dirichlet conditions in a set  $B$ , that reads (in the context of arbitrary discrete state space)

$$(1.1) \quad g_B(x, y) = \mu(y) \frac{h_{x,B}(y)}{\text{cap}(x, B)},$$

where  $B$  is a subset of the configuration space,  $h_{x,B}(y) = h_{\{x\},B}$  and  $h_{A,B}$  is the equilibrium potential, that is,

$$(1.2) \quad h_{A,B}(y) = \begin{cases} 1, & \text{if } y \in A, \\ 0, & \text{if } y \in B, \\ \mathbb{P}_y(\tau_A < \tau_B), & \text{otherwise.} \end{cases}$$

We use

$$\tau_C = \min\{t > 0 : x(t) \in C\}$$

for the first hitting times of sets  $C$ , and  $\text{cap}(A, B)$  is the capacity between the sets  $A$  and  $B$ ;  $\text{cap}(x, B) = \text{cap}(\{x\}, B)$ .

Equation (1.1) immediately leads to a formula for the mean hitting time  $\mathbb{E}_x \tau_B$  of  $B$ , for the process starting in  $x$ . However, the resulting expression for  $\mathbb{E}_x \tau_B$  is useful as long as the ratio appearing in (1.1) is under control and is not seriously of the form  $0/0$ .

To be more precise, it may happen that  $h_{x,B}(y) = f(A)h_{A,B}(y)$  and  $\text{cap}(x, B) = f(A) \text{cap}(A, B)$ , for “macroscopic” sets  $A \ni x$ . Then

$$(1.3) \quad \frac{h_{x,B}(y)}{\text{cap}(x, B)} = \frac{h_{A,B}(y)}{\text{cap}(A, B)},$$

but except in cases where (1.3) is manifest by some symmetry, it will be very hard to establish such relations by a direct pointwise estimation of numerator and denominator in (1.1).

Examples where this problem occurs are diffusion processes in  $d > 1$ , Glauber dynamics in the case of finite temperature, etc. In such cases, a useful version can be extracted by averaging equation (1.1) with respect to  $x$  after multiplying both sides by  $\text{cap}(x, B)$  over as suitable neighborhood  $A \equiv A_x$ . This yields the formula

$$(1.4) \quad \mathbb{E}_{\nu_A} \tau_B = \frac{1}{\text{cap}(A, B)} \sum_y h_{A,B}(y) \mu(y),$$

where  $\nu_A$  is a specific probability distribution on  $A$ . Actually (1.4) can be derived without a recourse to (1.1): if  $P$  is the transition kernel of a reversible Markov chain  $x(t)$ , then the equilibrium potential  $h_{A,B}$  is harmonic outside  $A \cup B$ ;  $(I - P)h_{A,B} = Lh_{A,B} = 0$ . Thus,

$$(1.5) \quad h_{A,B}(y) = \sum_{x \in A} g_B(y, x) Lh_{A,B}(x)$$

for all  $y \notin B$ . By reversibility  $\mu(y)g_B(y, x) = \mu(x)g_B(x, y)$ , and it follows that

$$(1.6) \quad \sum_{y \notin A \cup B} \mu(y)h_{A,B}(y) = \sum_{x \in A} \mu(x)Lh_{A,B}(x)\mathbb{E}_x \tau_B,$$

which is (1.4) with  $\nu_A(x) = \mu(x)Lh_{A,B}(x)/\text{cap}(A, B)$ .

The point is that the right-hand side of (1.4) can be evaluated in many cases of interest when formula (1.1) suffers from the problem discussed above. This has been demonstrated recently in two examples, the Glauber dynamics of the random field Curie–Weiss model at finite temperature [1], and the Kawasaki dynamics in the zero temperature limit on volumes that diverge exponentially with the inverse temperature [3].

An obvious question is whether the mean hitting time of  $B$  really depends on the specific initial distribution  $\nu_A$  or whether, for all  $z \in A$ ,  $\mathbb{E}_z \tau_B$  is equal to  $\mathbb{E}_{\nu_A} \tau_B$  up to a small error. This question, and related one concerning other functions of initial conditions is of much further reaching importance. In particular, it is relevant for proving the asymptotic exponentiality of the transition time using approximate renewal arguments. Let us mention that the same issue also arises in the case of diffusion equations in the Wentzell–Freidlin regime. Here, Martinelli and Scoppola [11], Martinelli, Olivieri and Scoppola [10] showed that solutions of the stochastic differential equation starting at two different points in a neighborhood of a stable equilibrium and driven by the same noise are converging exponentially fast to each other with probability tending to one. From this, they deduced regularity of exit probabilities  $\mathbb{P}_x[\tau_B > t\mathbb{E}\tau_B]$  as functions of  $x$  and hence exponentiality of  $\tau_B$  and asymptotic independence of  $\mathbb{E}_x \tau_B$  of the starting point  $x \in A$ . Such a strong contraction property is, however, not available in stochastic Ising models on the level of microscopic paths.

In the present paper, we will develop a method that allows us to obtain similar results, at least in some cases, with an alternative and, weaker input. It is based on coupling techniques and allows us to turn the following simple heuristic argument into a rigorous proof: the Markov chain should mix quickly before it leaves a substantial neighborhood of the starting point  $x$ ; since the mixing time is short compared to the hitting time  $\tau_B$ , the mean of  $\tau_B$  should be the same for all starting configuration in  $A$ . Moreover, the chain will return many times to  $A$  before reaching  $B$ ; by rapid mixing, the return times will be essentially i.i.d., hence the number of returns will be geometric, and the scaled hitting time will be exponential.

To demonstrate the usefulness of this approach, our key example will be the Random Field Curie–Weiss model with continuous distribution of the random fields. In that sense, the present result is also a completion of our previous paper [1]. Technically, the coupling construction we employ is based on [9] and still contains model dependent elements. However, the basic ideas are more general and will be of relevance for the treatment of a wider range of metastable systems.

The remainder of this paper is organized as follows. In the next subsection, we describe a general setting of Markov chains to which our method applies. In Section 1.3, we state our two main theorems. In Section 2, we recall the definition of Glauber dynamics for the random field Curie–Weiss model and recall the main result from [1]. In Section 3, we recall the coupling constructed by Levin, Luczak and Peres for the standard Curie–Weiss model and show how this can be modified to be useful in the random field model. We then prove Theorem 1.1. In Section 4, we show how to prove the asymptotic exponentiality of the transition times and give the proof of Theorem 1.2.

1.2. *Setting.* In this subsection, we describe a general setting in which our methods can be applied.

In the sequel,  $N$  will be a large parameter. We consider (families of) Markov processes,  $\sigma(t)$ , with finite state space,  $\mathcal{S}_N \equiv \{-1, 1\}^N$ , and transition probabilities  $p_N$  that are reversible w.r.t. a (Gibbs) measure,  $\mu_N$ . Transition probabilities  $p_N$  always have the following structure: at each step, a site  $x \in \Lambda$  is chosen with uniform probability  $1/N$ . Then the spin at  $x$  is set to  $\pm 1$  with probabilities  $p_x^\pm(\sigma)$ ;  $p_x^+(\sigma) + p_x^-(\sigma) \equiv 1$ . In the sequel, we shall assume that there exists  $\alpha \in [1/2, 1)$  such that

$$(1.7) \quad \max_{x, \sigma, \pm} p_x^\pm(\sigma) \leq \alpha.$$

A key hypothesis is the existence of a family of “good” mesoscopic approximations of our processes. By this, we mean the following: there is a sequence of disjoint partitions,  $\{\Lambda_1^n, \dots, \Lambda_{k_n}^n\}$ , of  $\Lambda \equiv \{1, \dots, N\}$ , and a family of maps,  $\underline{m}^{(n)} : \mathcal{S}_N \rightarrow \Gamma_n \subset \mathbb{R}^n$ , given by

$$(1.8) \quad m_i^n(\sigma) = \frac{1}{N} \sum_{x \in \Lambda_i^n} \sigma_x.$$

We will always think of these partitions as nested, that is,  $\{\Lambda_1^{n+1}, \dots, \Lambda_{k_{n+1}}^{n+1}\}$  is a refinement of  $\{\Lambda_1^n, \dots, \Lambda_{k_n}^n\}$ . On the other hand, to lighten the notation, we will mostly drop the superscript and identify  $k_n = n$ , and refer to the generic partition  $\Lambda_1, \dots, \Lambda_n$ . It will be convenient to introduce the notation

$$\mathcal{S}^n[\underline{m}] \equiv (\underline{m}^n)^{-1}(\underline{m}) = \{\sigma : \underline{m}^n(\sigma) = \underline{m}\}$$

for the set-valued inverse images of  $\underline{m}^n$ . We think of the maps  $\underline{m}^n$  as some block averages of our “microscopic” variables  $\sigma_i$  over blocks of decreasing (in  $n$ ) “mesoscopic” sizes.

As is well known, the image process,  $\underline{m}^n(\sigma(t))$ , is in general not Markovian. However, there is a canonical Markov process,  $\underline{m}^n(t)$ , with state space  $\Gamma_n$  and reversible measure  $\mathcal{Q}_n \equiv \mu_N \circ (\underline{m}^n)^{-1}$ , that is a “best” approximation of  $\underline{m}^n(\sigma(t))$ , in the sense that if  $\underline{m}^n(\sigma(t))$  is Markov, then  $\underline{m}^n(t) = \underline{m}^n(\sigma(t))$  (in law). For all  $\underline{m}, \underline{m}' \in \Gamma_n$ , the transition probabilities of this chain are given by

$$(1.9) \quad r_N(\underline{m}, \underline{m}') \equiv \frac{1}{\mathcal{Q}_n(\underline{m})} \sum_{\substack{\sigma \in \mathcal{S}^n[\underline{m}] \\ \sigma' \in \mathcal{S}^n[\underline{m}']}} \mu_N(\sigma) p_N(\sigma, \sigma').$$

In the models, we consider here the following two assumptions are satisfied:

(A.1) The sequence of chains  $\underline{m}^n(t)$  approximates  $\underline{m}^n(\sigma(t))$  in the strong sense that there exists  $\varepsilon(n) \downarrow 0$ , as  $n \uparrow \infty$ , such that for any  $\underline{m}, \underline{m}' \in \Gamma_n$ ,

$$(1.10) \quad \max_{\substack{\sigma \in \mathcal{S}^n[\underline{m}], \sigma' \in \mathcal{S}^n[\underline{m}'] \\ r_N(\underline{m}, \underline{m}') > 0}} \left| \frac{p_N(\sigma, \sigma') |\mathcal{S}^n[\underline{m}']|}{r_N(\underline{m}, \underline{m}')} - 1 \right| \leq \varepsilon(n).$$

(A.2) The microscopic flip rates satisfy: if  $\underline{m}^n(\sigma) = \underline{m}^n(\eta)$  and  $\sigma_x = \eta_x$ , then  $p_x^\pm(\sigma) = p_x^\pm(\eta)$ .

Note that our assumption (A.1) is much stronger than the maybe more natural looking

$$\max_{\sigma \in \mathcal{S}^n[\underline{m}]} \left| \frac{\sum_{\sigma' \in \mathcal{S}^n[\underline{m}']} p(\sigma, \sigma')}{r_N(\underline{m}, \underline{m}')} - 1 \right| \leq \varepsilon(n).$$

Finally, we need to place us in a “metastable” situation. Specifically, we will assume that there exist two disjoint sets  $A = \{\sigma \in \mathcal{S}_N : \underline{m}^{n_0}(\sigma) \in \mathbf{A}\}$  and  $B = \{\sigma \in \mathcal{S}_N : \underline{m}^{n_0}(\sigma) \in \mathbf{B}\}$ , for some  $n_0$  and sets  $\mathbf{A}, \mathbf{B} \subseteq \Gamma_{n_0}$ , a constant  $C > 0$  and a sequence  $a_n < \infty$ , such that, for all  $n \geq n_0$  and for all  $\sigma, \eta \in A$ ,

$$(1.11) \quad \mathbb{P}_\sigma[\tau_B < \tau_{\underline{m}^n(\eta)}] \leq a_n e^{-CN},$$

where, with a little abuse of notation, we denote by  $\tau_{\underline{m}^n(\eta)}$  the first hitting time of the set  $\mathcal{S}^n[\underline{m}^n(\eta)]$ .

1.3. *Main results.* In the setting outlined above, we will prove the following theorem.

**THEOREM 1.1.** *Consider a Markov process as described above, and let  $A, B$  be such that (1.11) holds. Then*

$$(1.12) \quad \max_{\sigma, \eta \in A} \left| \frac{\mathbb{E}_\sigma \tau_B}{\mathbb{E}_\eta \tau_B} - 1 \right| \leq e^{-CN/2}.$$

**REMARK.** Assumptions (A1) and (A2) are formulated in the context in which we will prove our results. The restriction to the state space  $\{-1, 1\}^N$  is mainly done because we need to construct an explicit coupling. It is rather straightforward to generalize everything to the case of Potts spins ( $\mathcal{S}_N \equiv \{1, \dots, q\}^N$ ) and maps  $\underline{m}^n$  whose components are permutation invariant functions of the spin variables on  $\Lambda_t^n$ .

The claim of Theorem 1.1 is trivial whenever  $\mathbb{P}_\sigma(\tau_\eta < \tau_B)$  is exponentially close to one, as  $N \uparrow \infty$ . However, in the context of stochastic Ising models it is reasonable to expect that, for fixed  $\sigma, \eta \in A$ ,  $\mathbb{P}_\sigma(\tau_\eta < \tau_B)$  is exponentially small. That is, despite the fact that a chain starting at  $\sigma$  spends an exponentially large amount of time in  $A$ , this time is not long enough for visiting more than a small fraction of the exponentially large number of microscopic points in  $A$ . An alternative approach is to try to construct a coupling between  $\sigma$  and  $\eta$  chains. In the case of the Curie–Weiss model (without random fields), a useful coupling algorithm was suggested in the recent paper [9]. This algorithm ensures that:

- (a) If  $\underline{m}^n(\sigma_s) = \underline{m}^n(\eta_s)$ , then  $\underline{m}^n(\sigma_t) = \underline{m}^n(\eta_t)$  for all  $t \geq s$ .
- (b) The Hamming distance between  $\sigma_t$  and  $\eta_t$  is nonincreasing in time.

In a way, this is reminiscent of the stochastic stability results of [11]. It is straightforward to adjust the construction of [9] to the general context we consider here. But both (a) and (b) above would be lost, and it is not clear that such a coupling would work globally.

Instead, our strategy is to use (1.11) and to keep trying to couple the  $\sigma$ -chain with a typical  $\eta$ -chain each time when  $\sigma_t$  enters  $\mathcal{S}^n(\underline{m}(\eta))$ . In the sequel, we call this the *basic coupling attempt*. Clearly, in view of a possible biased sampling, basic coupling attempts should be designed with care, which explains the relatively complicated construction in Section 3.2. It is based on [9], but we need to enlarge the probability space in order to achieve sufficient independence between decision making and properties of the eventually chosen  $\eta$ -path. In particular, the fact that  $\sigma$ -chain and  $\eta$ -chain meet will not automatically imply coupling.

A second and related problem that tends to arise in the situation that we are interested in is the breakdown of strict renewal properties. This is a well-known issue in the theory of continuous space Markov processes where methods such

as Nummelin splitting [13] were devised to prove ergodic theorem for the Harris recurrent chains. Here we would like to use renewal arguments, for example, to prove asymptotic exponentiality of the law of  $\tau_B$ . We will show that again coupling arguments can be used to solve such problems.

As an example, we will prove the following theorem.

**THEOREM 1.2.** *In the random field Curie–Weiss model, for  $A$  and  $B$  chosen to satisfy the hypothesis of Theorem 1.1,*

$$(1.13) \quad \mathbb{P}_\sigma(\tau_B / \mathbb{E}_\sigma \tau_B > t) \rightarrow e^{-t} \quad \text{as } N \uparrow \infty$$

for all  $\sigma \in A$  and for all  $t \in \mathbb{R}_+$ .

Theorem 1.2 is proven in Section 3. The basic idea is to use our iterative coupling procedure for deriving a renewal-type equation for the Laplace transform of  $\tau_B$ .

**2. The random field Curie–Weiss model.** The results of this paper are motivated by the study of the Glauber dynamics of the random field Curie–Weiss model (RFCW). We will show that the assumptions of the two theorems above can be verified in that model. In this section, we briefly recall results for this model obtained recently in [1] and prove an elementary local recurrence estimate.

*2.1. The model and equilibrium properties.* In the RFCW model, the state space is  $\mathcal{S}_N \equiv \{-1, 1\}^N$ , the Gibbs measure is given by

$$(2.1) \quad \mu_N(\sigma) = Z_N^{-1} \exp(-\beta H_N(\sigma)),$$

and the *random Hamiltonian*,  $H_N$ , is defined as

$$(2.2) \quad H_N(\sigma) \equiv -\frac{N}{2} \left( \frac{1}{N} \sum_{i \in \Lambda} \sigma_i \right)^2 - \sum_{i \in \Lambda} h_i \sigma_i,$$

where  $\Lambda \equiv \{1, \dots, N\}$  and  $h_i, i \in \Lambda$ , are i.i.d. random variables on some probability space  $(\Omega, \mathcal{F}, \mathbb{P}_h)$ .

The total magnetization

$$(2.3) \quad m_N(\sigma) \equiv \frac{1}{N} \sum_{i \in \Lambda} \sigma_i$$

is an effective order parameter of the model, and the sets of configurations where the magnetization takes particular values play the rôle of *metastable states*. More specifically, we introduce the law of  $m_N$  through

$$(2.4) \quad \mathcal{Q}_{\beta, N} \equiv \mu_{\beta, N} \circ m_N^{-1}$$

on the set of possible values  $\Gamma_N \equiv \{-1, -1 + 2/N, \dots, 1\}$ .  $\mathcal{Q}_{\beta,N}$  satisfies a large deviation property, in particular

$$(2.5) \quad Z_{\beta,N} \mathcal{Q}_{\beta,N}(m) = \sqrt{\frac{2I_N''(m)}{N\pi}} \exp(-N\beta F_{\beta,N}(m))(1 + o(1))$$

with  $I_N$  being the Legendre transform of

$$(2.6) \quad t \mapsto \frac{1}{N} \sum_{i \in \Lambda} \log \cosh(t + \beta h_i),$$

and with an explicit form for the rate function (“free energy”),  $F_{\beta,N}$ . The metastable states correspond to multiple local minima of  $F_{\beta,N}$ , whenever they exist.

A crucial feature of the model is that we can introduce a family of *mesoscopic* variables in such a way that the dynamics on these mesoscopic variables is well approximated by a Markov process. Let us briefly describe these mesoscopic variables.

**2.2. Coarse graining.** Let  $I$  denote the support of the distribution of the random fields. Let  $I_\ell$ , with  $\ell \in \{1, \dots, n\}$ , be a partition of  $I$  such that, for some  $C < \infty$  and for all  $\ell$ ,  $|I_\ell| \leq C/n \equiv \varepsilon$ .

Each realization of the random field  $\{h_i\}_{i \in \mathbb{N}}$  induces a random partition of the set  $\Lambda \equiv \{1, \dots, N\}$  into subsets

$$(2.7) \quad \Lambda_k \equiv \{i \in \Lambda : h_i \in I_k\}.$$

We may introduce  $n$  order parameters

$$(2.8) \quad m_k(\sigma) \equiv \frac{1}{N} \sum_{i \in \Lambda_k} \sigma_i.$$

We denote by  $\underline{m}$  the  $n$ -dimensional vector  $(m_1, \dots, m_n)$ .  $\underline{m}$  takes values in the set

$$(2.9) \quad \Gamma_N^n \equiv \prod_{k=1}^n \left\{ -\rho_{N,k}, -\rho_{N,k} + \frac{2}{N}, \dots, \rho_{N,k} - \frac{2}{N}, \rho_{N,k} \right\},$$

where

$$(2.10) \quad \rho_k \equiv \rho_{N,k} \equiv \frac{|\Lambda_k|}{N}.$$

Note that the random variables  $\rho_k$  concentrate exponentially (in  $N$ ) around their mean value  $\mathbb{E}_h \rho_{N,k} = \mathbb{P}_h[h_i \in I_k] \equiv p_k$ . The Hamiltonian can be written as

$$(2.11) \quad H_N(\sigma) = -NE(\underline{m}(\sigma)) + \sum_{\ell=1}^n \sum_{i \in \Lambda_\ell} \sigma_i \tilde{h}_i,$$



where  $E : \mathbb{R}^n \rightarrow \mathbb{R}$  is the function

$$(2.12) \quad E(\underline{x}) \equiv \frac{1}{2} \left( \sum_{k=1}^n x_k \right)^2 + \sum_{k=1}^n \bar{h}_k x_k$$

with

$$(2.13) \quad \bar{h}_\ell \equiv \frac{1}{|\Lambda_\ell|} \sum_{i \in \Lambda_\ell} h_i \quad \text{and} \quad \tilde{h}_i \equiv h_i - \bar{h}_\ell.$$

The equilibrium distribution of the variables  $\underline{m}(\sigma)$  is given by

$$(2.14) \quad \begin{aligned} \mathcal{Q}_{\beta,N}(\underline{x}) &\equiv \mu_{\beta,N}(\underline{m}(\sigma) = \underline{x}) \\ &= \frac{1}{Z_N} e^{\beta N E(\underline{x})} \mathbb{E}_\sigma \mathbb{1}_{\{\underline{m}(\sigma) = \underline{x}\}} e^{\sum_{\ell=1}^n \sum_{i \in \Lambda_\ell} \sigma_i (h_i - \bar{h}_\ell)}. \end{aligned}$$

For a mesoscopic subset,  $\mathbf{A} \subseteq \Gamma_N^n$ , we define its microscopic counterpart,  $A$ , as

$$(2.15) \quad A = \mathcal{S}^n[\mathbf{A}] = \{\sigma \in \mathcal{S}_N : \underline{m}(\sigma) \in \mathbf{A}\}.$$

Note that, as in the one-dimensional case, we can express the right-hand side of (2.14) as

$$(2.16) \quad Z_{\beta,N} \mathcal{Q}_{\beta,N}(\underline{x}) = \prod_{\ell=1}^n \sqrt{\frac{(I''_{N,\ell}(x_\ell/\rho_\ell)/\rho_\ell)}{N\pi/2}} \exp(-N\beta F_{\beta,N}(\underline{x}))(1 + o(1))$$

with an explicit expression for the function  $F_{\beta,N}$ ,

$$(2.17) \quad F_{\beta,N}(\underline{x}) \equiv -\frac{1}{2} \left( \sum_{\ell=1}^n x_\ell \right)^2 - \sum_{\ell=1}^n x_\ell \bar{h}_\ell + \frac{1}{\beta} \sum_{\ell=1}^n \rho_\ell I_{N,\ell}(x_\ell/\rho_\ell).$$

The key point of the construction above is that it places the RFCW model in the context described in Section 1.2. Namely, defining the mesoscopic rates,  $r_N(\underline{m}, \underline{m}')$ , in (1.9) for the functions  $\underline{m}$  defined in (2.8), one can easily verify that the estimates (1.10) hold, as was exploited in [1]. In the next subsection, we will show that the recurrence hypothesis (1.11) also holds in this model.

In [1], we proved the following.

**THEOREM 2.1.** *Assume that  $\beta$  and the distribution of the magnetic field are such that there exist more than one local minimum of  $F_{\beta,N}$ . Let  $m^*$  be a local minimum of  $F_{\beta,N}$ ,  $M \equiv M(m^*)$  be the set of minima of  $F_{\beta,N}$  such that  $F_{\beta,N}(m) < F_{\beta,N}(m^*)$ , and  $z^*$  be the minimax between  $m$  and  $M$ , that is, the lower of the highest maxima separating  $m$  from  $M$  to the left, respectively, right. Then,  $\mathbb{P}_h$ -almost surely,*

$$(2.18) \quad \begin{aligned} \mathbb{E}_{\nu_{S[m^*], S[M]}} \tau_{S[M]} &= C(\beta, m^*, M) N \exp(\beta N [F_{\beta,N}(z^*) - F_{\beta,N}(m^*)]) \\ &\quad \times (1 + o(1)), \end{aligned}$$

where  $C(\beta, m^*, M)$  is a constant that is computed explicitly in [1].

Here the initial measure,  $\nu_{S[m^*], S[M]}$ , is the so-called *last exit biased distribution* on the set  $S[m^*] \equiv \{\sigma \in \mathcal{S}_N : m_N(\sigma) = m^*\}$ , given by the formula

$$(2.19) \quad \nu_{A,B}(\sigma) = \frac{\mu_{\beta,N}(\sigma) \mathbb{P}_\sigma[\tau_B < \tau_A]}{\sum_{\sigma \in A} \mu_{\beta,N}(\sigma) \mathbb{P}_\sigma[\tau_B < \tau_A]}.$$

Although the theorem is stated in [1] for the starting measure in a set defined with respect to the one-dimensional order parameter, the estimates given there immediately imply that the same formulas hold replacing  $m^*$  with a local minimum,  $\underline{m}^*$ , in the  $n$ -dimensional order parameter space.

Theorem 2.1 implies that the estimate (2.18) holds for  $\mathbb{E}_\sigma \tau_{S[M]}$  for any  $\sigma$  in a neighborhood of  $S[\underline{m}^*]$ , for  $n$  large enough.

**2.3. Local recurrence.** Before starting the proof of (1.12), let us verify that the hypothesis (1.11) holds for the RFCW model. Specifically, let us define the metastable set  $\mathbf{A}_\delta \subset \Gamma_n$  as the ball, with respect to the Hamming distance, of fixed radius  $\delta N$ ,  $\delta > 0$ , centered on a local minimum  $\underline{m}^*$  of  $F_{\beta,N}$ . Let  $A_\delta \subset \mathcal{S}_N$  be the corresponding microscopic metastable set and denote by  $\tau_{\underline{m}}$  the first hitting time of the set  $\mathcal{S}^n[\underline{m}]$ . With this notation, we have the following lemma.

LEMMA 2.2. *There exist  $\delta > 0$  and  $c_1 > 0$  such that, for all  $n$  large enough,  $\sigma, \sigma' \in A_\delta$ ,*

$$(2.20) \quad \mathbb{P}_\sigma[\tau_B < \tau_{\underline{m}(\sigma')}] \leq e^{-c_1 N}.$$

PROOF. We first notice that if  $\sigma' \in \mathcal{S}^n[\underline{m}^*]$ , then the assertion of the lemma holds for all  $n$  sufficiently large with a constant  $c_0$  independent of  $n$ , as has been proven in [1] (see Proposition 6.12).

Moreover, for all  $\sigma, \sigma' \in A_\delta$ ,

$$(2.21) \quad \mathbb{P}_\sigma[\tau_{\underline{m}(\sigma')} < \tau_{\underline{m}^*}] \geq e^{-c\delta N}$$

for some positive constant  $c$ . To see this, notice that, due to the property of  $\mathbf{A}_\delta$ , one can find a mesoscopic path from  $\underline{m}(\sigma)$  to  $\underline{m}(\sigma')$  with length at most  $\delta N$ . Implementing the argument that is used in the proof of Lemma 6.11 of [1], one gets (2.21).

To prove (2.20), we use a renewal argument. Let us consider a configuration  $\sigma \in \mathcal{S}^n[\underline{m}^*]$  and a generic  $\sigma' \in A_\delta$ , and set  $\underline{m} \equiv \underline{m}(\sigma')$ . Then

$$(2.22) \quad \begin{aligned} \mathbb{P}_\sigma(\tau_B < \tau_{\underline{m}}) &\leq \mathbb{P}_\sigma(\tau_B < \tau_{\underline{m}} \wedge \tau_{\underline{m}^*}) + \mathbb{P}_\sigma(\tau_{\underline{m}^*} < \tau_B < \tau_{\underline{m}}) \\ &\leq \mathbb{P}_\sigma(\tau_B < \tau_{\underline{m}^*}) + \max_{\eta \in \mathcal{S}^n[\underline{m}^*]} \mathbb{P}_\eta(\tau_B < \tau_{\underline{m}}) \mathbb{P}_\sigma(\tau_{\underline{m}^*} < \tau_{\underline{m}}) \\ &\leq e^{-c_0 N} + \max_{\eta \in \mathcal{S}^n[\underline{m}^*]} \mathbb{P}_\eta(\tau_B < \tau_{\underline{m}}) (1 - e^{-c\delta N}), \end{aligned}$$

where in the second line we used the Markov property, and in the last line we used the inequality (2.20) and (2.21). Taking the maximum over  $\sigma \in \mathcal{S}^n[\underline{m}^*]$  on

both sides of (2.22) and rearranging the summation, we get the inequality (2.20) for  $\sigma \in \mathcal{S}^n[\underline{m}^*]$ , with a constant  $c_1 = c_0 - c\delta$  which is strictly positive for small enough  $\delta$ .

Now let us consider the general case when  $\sigma, \sigma' \in A_\delta$  and set again  $\underline{m} \equiv \underline{m}(\sigma')$ . As before, we have

$$\begin{aligned}
 \mathbb{P}_\sigma(\tau_B < \tau_{\underline{m}}) &\leq \mathbb{P}_\sigma(\tau_B < \tau_{\underline{m}} \wedge \tau_{\underline{m}^*}) + \mathbb{P}_\sigma(\tau_{\underline{m}^*} < \tau_B < \tau_{\underline{m}}) \\
 (2.23) \qquad &\leq \mathbb{P}_\sigma(\tau_B < \tau_{\underline{m}^*}) + \max_{\eta \in \mathcal{S}^n[\underline{m}^*]} \mathbb{P}_\eta(\tau_B < \tau_{\underline{m}}) \mathbb{P}_\sigma(\tau_{\underline{m}^*} < \tau_B) \\
 &\leq e^{-c_0 N} + e^{-c_1 N} = e^{-c_1 N} (1 + o(1)),
 \end{aligned}$$

where in the third line we used the fact that the inequality (2.20) was already established for  $\eta \in \mathcal{S}^n[\underline{m}^*]$ . This concludes the proof of the lemma.  $\square$

Lemma 2.2 shows that, for all  $n$  large enough, the RFCW model parameterized by the variables  $\underline{m} \in \Gamma_N^n$  satisfies the hypothesis (1.11) with  $A = A_\delta$  as defined above.

### 3. Construction of the coupling.

3.1. *The coupling by Levin, Luczak and Peres.* Recall that we consider a partition,  $\{\Lambda_1, \dots, \Lambda_n\}$ , of  $\Lambda \equiv \{1, \dots, N\}$  and let  $\underline{m} = (m_1(\sigma), \dots, m_n(\sigma))$  be the vector of partial magnetizations as defined in (2.8).

We begin by explaining a coupling that was used by Levin, Luczak and Peres [9] in the usual Curie–Weiss model. In that case, the transition rates have the following properties: whenever  $x, y$  and  $\sigma, \eta$  are such that:

- (i)  $\underline{m}(\sigma) = \underline{m}(\eta)$ , and
- (ii)  $\sigma_x = \eta_y$ ,

then

$$(3.1) \qquad p_N(\sigma, \sigma^{(x)}) = p_N(\eta, \eta^{(y)}).$$

We continue to employ the notation  $p_x^\pm(\sigma)$ ,

$$(3.2) \qquad p_x^{-\sigma_x}(\sigma) \equiv N p_N(\sigma, \sigma^{(x)}) \quad \text{and} \quad p_x^+(\sigma) + p_x^-(\sigma) = 1,$$

where, as usual,  $\sigma^{(x)}$  is the configuration obtained from  $\sigma$  by setting  $\sigma_x^x = -\sigma_x$  and leaving all other components of  $\sigma$  unchanged.

The coupling of Levin, Luczak and Peres is constructed as follows. Let  $\sigma$  and  $\eta$  be two initial conditions such that  $\underline{m}(\sigma) = \underline{m}(\eta)$ . Let  $I_t, t = 0, 1, 2, \dots$ , be a family of independent random variables that are uniformly distributed on  $\Lambda$ . Assume that at time  $t$ ,  $\underline{m}(\sigma(t)) = \underline{m}(\eta(t))$  and do the following:

- (O1) Draw the random variable  $I_t$ ;

- (O2) Set  $\eta_{I_t}(t + 1) = \pm 1$  with probabilities  $p_{I_t}^{\pm}(\eta(t))$ , respectively, and set  $\eta_x(t + 1) = \eta_x(t)$  for all  $x \neq I_t$ ;
- (A) Then do the following:
- (i) If  $\sigma_{I_t}(t) = \eta_{I_t}(t)$ , then set:
    - \*  $\sigma_{I_t}(t + 1) = \eta_{I_t}(t + 1)$ ;
    - \*  $\sigma_x(t + 1) = \sigma_x(t)$ , for all  $x \neq I_t$ .
  - (ii) If  $\sigma_{I_t}(t) \neq \eta_{I_t}(t)$ , then let  $\Lambda_k$  be the element of the partition such that  $I_t \in \Lambda_k$  and choose  $y$  uniformly at random on the set  $\{z \in \Lambda_k : \sigma_z(t) \neq \eta_z(t) \neq \eta_{I_t}(t)\}$ . Note that this set is not empty, since  $\underline{m}(\sigma(t)) = \underline{m}(\eta(t))$  and  $\sigma(t)$  and  $\eta(t)$  differ in one site of  $\Lambda_k$ . Then set:
    - \*  $\sigma_y(t + 1) = \eta_{I_t}(t + 1)$ ;
    - \*  $\sigma_x(t + 1) = \sigma_x(t)$ , for all  $x \neq y$ .

Note that this construction has the virtue that  $\underline{m}(\sigma(t + 1)) = \underline{m}(\eta(t + 1))$ , so that the assumption inherent in the construction is always verified, if it is verified at time zero.

Moreover, if  $\sigma(t) = \eta(t)$  for some  $t$ , then  $\sigma(t + s) = \eta(t + s)$ , for all  $s \geq 0$ . Finally, one easily checks that the marginal distributions of  $\sigma(t)$  and  $\eta(t)$  coincide and are given by the law of the original dynamics. This latter fact depends crucially on the fact that the flip rates do not depend on which site in a given subset  $\Lambda_i$  the spin is flipped, provided they are flipped in the same direction.

*3.2. Coupling attempt in the general case.* In the general case, we consider here, including the RFCW, (3.1) does not hold unless  $x = y$ . All we assume is (A.2) and (1.10). The problem is now that the probabilities to update the  $\sigma$ -chain in a chosen point  $y$  are typically not the same as those of the  $\eta$ -chain in the original point  $I_t$ . However, by (1.10), these probabilities are still close to each other, in the sense that there exists  $\nu = \nu(n)$  with  $\nu \downarrow 0$  as  $n \uparrow \infty$ , for example,  $\nu(n) = 3\epsilon(n)$ , such that for any  $k$ , for any  $x, y \in \Lambda_k$  and for any  $\sigma$  and  $\eta$  with  $\underline{m}(\sigma) = \underline{m}(\eta)$  and  $\sigma_x = \eta_y$ ,

$$(3.3) \quad \frac{p_x^{\pm}(\eta)}{p_y^{\pm}(\sigma)} \leq 1 + \nu.$$

Thus, in order to maintain the correct marginal distribution for the processes, we have to change the updating rules in such a way that the  $\sigma$ -chain will sometimes not maintain the same magnetization as the  $\eta$ -chain, which implies that the coupling cannot be continued.

The basic strategy to overcome this difficulty is to use iterated coupling attempts. We shall decompose the  $\sigma$ -path on  $[0, \tau_B^{\sigma})$  into cycles and during each cycle we shall attempt to couple it with an independent copy of the  $\eta$ -chain. In the case of success, both chains will run together until  $\tau_B$ . Such procedure necessarily involves a sampling of  $\eta$ -paths. In order to control its bias, it will be important to separate the path properties of  $\eta$ -chains with which we try to couple from the probability of whether a subsequent coupling attempt is successful or not. This will be achieved by constructing a coupling on an extended probability space.

*Basic coupling attempt.* There are two parameters  $c_2 > 0$  and  $\kappa < \infty$  whose values will be quantified in the sequel.

Let  $\eta$  and  $\sigma$  satisfy  $\underline{m}(\eta) = \underline{m}(\sigma)$ . We shall try to couple a  $\sigma$ -path with an  $\eta$ -path during the first  $N^\kappa$ -steps of their life. Let  $M = c_2 N$  and let  $V_i, i = 1, \dots, M$ , be a family of i.i.d. Bernoulli random variables with

$$(3.4) \quad \mathbb{P}[V_i = 1] = 1 - \mathbb{P}[V_i = 0] = 1 - \nu(n).$$

We now describe how the coupling construction is adjusted using the random variables  $V_i, i = 1, \dots, M$ . Let  $\underline{m}(\eta(0)) = \underline{m}(\sigma(0))$ .

As before, let  $I_t, t = 0, 1, 2, \dots, N^\kappa$ , be a family of independent random variables that are uniformly distributed on  $\Lambda$ . Let  $\mathcal{M}_0 = 0$  and  $\chi_0 = 0, \eta(0) = \eta, \sigma(0) = \sigma$ . At time  $t \geq 1$ , do the following:

- (O1) Draw the random variable  $I_t$ ;
- (O2) Set  $\eta_{I_t}(t + 1) = \pm 1$  with probability  $p_{I_t}^\pm(\eta_{I_t}(t))$  and set  $\eta_x(t + 1) = \eta_x(t)$  for all  $x \neq I_t$ ;
- (A) If at time  $1 \leq t \leq N^\kappa, \chi_t = 0$  and  $\mathcal{M}_t < M$ , then do the following:
  - (i) If  $\sigma_{I_t}(t) = \eta_{I_t}(t)$ , then set:
    - \*  $\sigma_{I_t}(t + 1) = \eta_{I_t}(t + 1)$ ;
    - \*  $\sigma_x(t + 1) = \sigma_x(t)$ , for all  $x \neq I_t$ ;
    - \*  $\mathcal{M}_{t+1} = \mathcal{M}_t$ .
  - (ii) If  $\sigma_{I_t}(t) \neq \eta_{I_t}(t)$ , let  $\Lambda_k$  be the element of the partition such that  $I_t \in \Lambda_k$  and, as before, choose  $y$  uniformly at random on the set  $\{z \in \Lambda_k : \sigma_z(t) \neq \eta_z(t) \neq \eta_{I_t}(t)\}$ . Then set:
    - \*  $\sigma_y(t + 1) = \eta_{I_t}(t + 1)$  with probability

$$(3.5) \quad \begin{cases} 1, & \text{if } V_{\mathcal{M}_t} = 1, \\ \frac{p_{I_t}^\pm(\eta(t)) \wedge p_y^\pm(\sigma(t)) - (1 - \nu)p_{I_t}^\pm(\eta(t))}{\nu p_{I_t}^\pm(\eta(t))}, & \text{if } V_{\mathcal{M}_t} = 0, \end{cases}$$

and  $\sigma_y(t + 1) = -\eta_{I_t}(t + 1)$  with probability

$$(3.6) \quad \begin{cases} 0, & \text{if } V_{\mathcal{M}_t} = 1, \\ \frac{p_{I_t}^\pm(\eta(t)) - p_{I_t}^\pm(\eta(t)) \wedge p_y^\pm(\sigma(t))}{\nu p_{I_t}^\pm(\eta(t))}, & \text{if } V_{\mathcal{M}_t} = 0; \end{cases}$$

- \*  $\sigma_x(t + 1) = \sigma_x(t)$ , for all  $x \neq y$ ;
- \* If  $V_{\mathcal{M}_t} = 0$ , then set  $\chi_s = 1$  for  $s = t + 1, \dots, N^\kappa$ , otherwise set  $\chi_{t+1} = \chi_t$ ;
- \* Set  $\mathcal{M}_{t+1} = \mathcal{M}_t + 1$ ;

- (B) If at time  $t$ , either  $\chi_t = 1$  or  $\mathcal{M}_t = M$ , then update  $\sigma$  independently of  $\eta$ , that is:

- (i) draw  $I'_t$  independently with the same law as  $I_t$ , and

- (ii) set  $\sigma_{I'_t}(t + 1) = \pm 1$  with probability  $p_{I'_t}^\pm(\sigma(t))$ , and  $\sigma_x(t + 1) = \sigma_x(t)$ , for all  $x \neq I'_t$ .

The process  $\mathcal{M}_t$  is a counter that increases by one each time a new coin  $V_i$  is used by the coupling. The value  $\chi_t = 1$  of the variable  $\chi_t$  indicates that a zero coin was used by time  $t$ .

The following lemma collects the basic properties of the process constructed above.

LEMMA 3.1. *Let  $\tilde{\mathbb{P}}$  denote the joint distribution of the processes  $\sigma, \eta, V$  defined above. Then the above is a good coupling in the sense that the marginal distributions of both  $\eta(t), t \leq N^k$ , and  $\sigma(t), t \leq N^k$ , under the law  $\tilde{\mathbb{P}}$  are  $\mathbb{P}_{\sigma(0)}$  and  $\mathbb{P}_{\eta(0)}$ , respectively.*

PROOF. The assertion is obvious for the process  $\eta(t)$ . It is also clear for the  $\sigma(t)$  process if updates are done according to case B. Therefore, we only need to check that it holds for process  $\sigma(t)$  at such times  $t \leq N^k$  when  $\chi_t$  is still 0 and  $\mathcal{M}_t$  is still less than  $M = c_2N$ . In other words, we have to compute

$$(3.7) \quad \tilde{\mathbb{P}}[\sigma(t + 1) = \sigma_x^+(t) | \sigma(t); \chi_t = 0; \mathcal{M}_t < c_2N],$$

where  $\sigma_x^+ \triangleq (\sigma_1, \dots, \sigma_{x-1}, +1, \dots, \sigma_N)$ . First, it is clear that, given that  $I_t = x$  and  $\sigma_{I_t}(t) = \eta_{I_t}(t)$ , we get the desired result, that is,

$$(3.8) \quad \begin{aligned} \tilde{\mathbb{P}}[\sigma(t + 1) = \sigma_x^+(t) | I_t = x; \sigma_x(t) = \eta_x(t); \sigma(t); \chi_t = 0; \mathcal{M}_t < c_2N] \\ = p_x^+(\eta(t)) = p_x^+(\sigma(t)). \end{aligned}$$

In the case  $I_t = y \neq x$ , we get a contribution to (3.7) only if:

- (i)  $x, y$  are in the same set  $\Lambda_i$ ,
- (ii)  $\sigma_y(t) \neq \eta_y(t)$ ,
- (iii)  $\sigma_x(t) \neq \eta_x(t)$ , and
- (iv)  $\sigma_x(t) = \eta_y(t)$ .

If these conditions are satisfied, the probability to flip  $\sigma_x$  to +1 is

$$(3.9) \quad \begin{aligned} (1 - \nu)p_y^+(\eta(t)) + \nu p_x^+(\eta(t)) \frac{p_y^+(\eta(t)) \wedge p_x^+(\sigma(t)) - (1 - \nu)p_y^+(\eta(t))}{\nu p_y^+(\eta(t))} \\ + \nu p_y^- \frac{p_y^-(\eta(t)) - p_y^-(\eta(t)) \wedge p_x^-(\sigma(t))}{\nu p_y^-(\eta(t))} \\ = p_y^+(\eta(t)) \wedge p_x^+(\sigma(t)) + p_y^-(\eta(t)) - p_y^-(\eta(t)) \wedge p_x^-(\sigma(t)) \\ = p_x^+(\sigma(t)). \end{aligned}$$

The last line is easily verified by distinguishing cases. It follows that the probability in (3.7) is equal to  $N^{-1} p_x^+(\sigma(t))$ , as desired. This proves the lemma.  $\square$

The construction above tries to merge the processes  $\eta(t)$  and  $\sigma(t)$  only as long as  $\chi_t = 0$  and  $\mathcal{M}_t < M$ . Note that if for some  $t < N^k$  both these conditions still hold and, in addition,  $\eta(t) = \sigma(t)$ , then the two dynamics automatically stay together until  $N^k$  and, indeed, stay coupled forever. Naively, one would want to classify such situation as “successful coupling.” However, this would involve an implicit sampling of  $\eta$ -trajectories which may lead to distortion of their statistical properties. For example, it is not clear whether the correct value of  $\mathbb{E}_\eta \tau_B$  would survive such a procedure.

In order to circumvent this obstacle, we use a more restrictive definition of what a “successful coupling” should be. Namely, we say that our basic coupling attempt is successful if the following two *independent* events  $\mathcal{A}$  and  $\mathcal{B}$  simultaneously happen on the enlarged probability space:

- (1) The event

$$(3.10) \quad \mathcal{A} \equiv \left\{ \bigvee_{i=1}^M V_i = 1 \right\}$$

is the event that all  $M$  random variables  $V_i$  should be equal to 1.

- (2) The event  $\mathcal{B}$  depends only on the random variables  $\eta(t), t \leq N^k$ . To define it, we introduce two stopping times,  $S$  and  $\mathcal{N}$ . Let

$$(3.11) \quad S_x = \inf\{t : \eta_x(t + 1) = -\eta_x(0)\}$$

and set  $S \equiv \max_{1 \leq x \leq N} S_x$ . Clearly,  $S_x$  is the first time the spin at site  $x$  has been flipped and  $S$  is the first time all coordinates of  $\eta$  have been flipped.  $\mathcal{N}$  is defined as

$$(3.12) \quad \mathcal{N} \equiv \sum_{t=0}^S \sum_{x=1}^N \mathbb{1}_{\{I_t=x\}} \mathbb{1}_{\{t \leq S_x\}},$$

which is the total number of flipping attempts until time  $S$ . Finally,

$$(3.13) \quad \mathcal{B} \equiv \{\tau_B^\eta \geq N^k\} \cap \{S < N^k\} \cap \{\mathcal{N} \leq M\}.$$

The important observation is the following.

LEMMA 3.2. *On  $\mathcal{A} \cap \mathcal{B}$ , the coupling is successful in the sense that*

$$(3.14) \quad \mathcal{A} \cap \mathcal{B} \subset \{\eta(N^k) = \sigma(N^k)\}.$$

PROOF. On the event  $\mathcal{B} \cap \mathcal{A}$ , by time  $N^k$   $\eta(t)$  has not reached  $B$ , all spins have been flipped once, and each flip that involved a site where  $\eta(t) \neq \sigma(t)$  was done when the coin  $V_i$  took the value  $+1$ . Therefore, on each first flip the corresponding  $\eta$  and  $\sigma$  spins became aligned, hence  $\eta(N^k) = \sigma(N^k)$ .  $\square$

REMARK. Note that the inclusion (3.14) is in general strict. The rationale for the introduction of the events  $\mathcal{A}$  and  $\mathcal{B}$  is that the unlikely event  $\mathcal{A}$  does not affect the  $\eta$ -chain at all and that the (likely) event  $\mathcal{B}$  does not distort the hitting times of the  $\eta$ -chain in the sense that  $\mathbb{E}_\eta(\tau_B \mathbb{1}_{\mathcal{B}}) \geq \mathbb{E}_\eta \tau_B (1 - e^{-cN})$ . This will be part of the content of Lemma 3.3 which we formulate and prove below.

3.3. *Construction of a cycle and cycle decomposition of  $\sigma$ -paths.* We have seen that  $\mathcal{B} \cap \mathcal{A}$  indicates that our coupling is successful and  $\eta(t)$  and  $\sigma(t)$  arrive together in  $B$ . However, the probability of  $\mathcal{A} \cap \mathcal{B}$  is very small, essentially due to the fact that the probability of  $\mathcal{A}$  is small, namely  $\mathbb{P}(\mathcal{A}) = (1 - \nu)^M$ . What will be essential is that the probability of  $\mathcal{B}$  is otherwise close to one, and therefore the  $\eta$ -paths (which are independent of the  $V_i$ ) will be affected very little by the occurrence of  $\mathcal{A} \cap \mathcal{B}$ .

We then have to decide what to do on  $(\mathcal{A} \cap \mathcal{B})^c$  at time  $N^\kappa$ . Define the stopping time

$$(3.15) \quad \Delta = \min\{t > N^\kappa : \sigma(t) \in \mathcal{S}^n[\underline{m}(\eta)]\}.$$

If

$$(3.16) \quad \mathcal{D} = \{\Delta < \tau_B^\sigma\}$$

happens, then we initiate a new basic coupling attempt at time  $\Delta$  for a new, independent copy of the  $\eta$ -chain and a chain starting from  $\sigma(\Delta)$ . Otherwise, on the event  $\mathcal{D}^c \cap (\mathcal{A} \cap \mathcal{B})^c$ , the process stops and coupling has not occurred.

The cycle decomposition of  $\sigma[0, \tau_B^\sigma]$  is based on a collection  $\{\eta^\ell[0, \tau_B^{\ell, \eta}]\}$  of independent copies of  $\eta$ -chains and on a collection  $\{\underline{V}^\ell = (V_0^\ell, \dots, V_{N^\kappa}^\ell)\}$  of i.i.d. stacks of coins. The events  $\{\mathcal{A}^\ell, \mathcal{B}^\ell\}$  are well defined and independent. The events  $\{\mathcal{D}^\ell\}$  are defined iteratively as follows: the event  $\mathcal{D}^0$  is simply the above event  $\mathcal{D}$  defined with respect the coupling attempt based on  $\{\eta^0, \underline{V}^0\}$ . If  $\mathcal{D}^0$  occurs, we denote by  $\theta_0 = \Delta_0$  the random time at which the first cycle ends. Assume now that  $\bigcap_0^{k-1} \mathcal{D}^\ell \cap \bigcap(\mathcal{A}^\ell \cap \mathcal{B}^\ell)^c$  happened and that the  $(k - 1)$ st cycle was finished at a random time  $\theta_{k-1}$  and at some random point  $\sigma(\theta_{k-1}) \in \mathcal{S}^n[\underline{m}(\eta)]$ . Let us initiate a new basic coupling attempt using a new independent copy  $\{\eta^k, \underline{V}^k\}$  for a chain starting at  $\eta$  and a chain starting from  $\sigma(\theta_{k-1})$ . The event  $\mathcal{D}^k$ , and accordingly the cycle length  $\Delta_k$ , are then defined appropriately. If  $\mathcal{D}^k$  happens and  $\bigcap(\mathcal{A}^\ell \cap \mathcal{B}^\ell)^c$  then  $\theta_k \equiv \theta_{k-1} + \Delta_k$ ,  $\sigma(\theta_k) \in \mathcal{S}^n[\underline{m}(\eta)]$  is well defined as well, and the iterative procedure goes on.

In light of the above definitions, the enlarged probability space  $\tilde{\Omega}$  has the following disjoint decomposition:

$$(3.17) \quad \mathbb{1} = \sum_{k=0}^{\infty} \mathbb{1}_{\mathcal{A}^k} \mathbb{1}_{\mathcal{B}^k} \prod_{\ell=0}^{k-1} (1 - \mathbb{1}_{\mathcal{A}^\ell} \mathbb{1}_{\mathcal{B}^\ell}) \mathbb{1}_{\mathcal{D}^\ell} + \sum_{k=0}^{\infty} (1 - \mathbb{1}_{\mathcal{A}^k} \mathbb{1}_{\mathcal{B}^k}) (1 - \mathbb{1}_{\mathcal{D}^k}) \prod_{\ell=0}^{k-1} (1 - \mathbb{1}_{\mathcal{A}^\ell} \mathbb{1}_{\mathcal{B}^\ell}) \mathbb{1}_{\mathcal{D}^\ell}.$$



As a consequence, we arrive at the following decomposition of the hitting time  $\tau_B^\sigma$  in terms of the (independent) hitting times  $\{\tau_B^{k,\eta}\}$ :

$$\begin{aligned}
 (3.18) \quad \tau_B^\sigma &= \sum_{k=0}^\infty \left\{ \prod_{\ell=0}^{k-1} \mathbb{1}_{\mathcal{D}^\ell} (1 - \mathbb{1}_{\mathcal{A}^\ell} \mathbb{1}_{\mathcal{B}^\ell}) \right\} (\theta_{k-1} + \tau_B^{k,\eta}) \mathbb{1}_{\mathcal{A}^k} \mathbb{1}_{\mathcal{B}^k} \\
 &+ \tau_B^\sigma \sum_{k=0}^\infty \left\{ \prod_{\ell=0}^{k-1} \mathbb{1}_{\mathcal{D}^\ell} (1 - \mathbb{1}_{\mathcal{A}^\ell} \mathbb{1}_{\mathcal{B}^\ell}) \right\} (1 - \mathbb{1}_{\mathcal{D}^k}) (1 - \mathbb{1}_{\mathcal{A}^k} \mathbb{1}_{\mathcal{B}^k}).
 \end{aligned}$$

(In both formulas above, we use the convention that products with a negative number of terms are equal to 1 and set  $\theta_{-1} \equiv 0$ .) Note that the first terms in (3.17) and (3.18) correspond to the cases when the iterative coupling eventually succeeds, whereas the second term corresponds to the case when it eventually fails.

3.4. Upper bounds on probabilities and proof of Theorem 1.1.

LEMMA 3.3. *The following estimates hold uniformly in  $\sigma, \eta \in A$ :*

(i) *There is a constant  $c > 0$ , independent of  $n$ , such that, for  $N$  large enough,*

$$(3.19) \quad \mathbb{P}(\mathcal{B}^c) \leq e^{-cN}$$

and

$$(3.20) \quad \mathbb{E}_\eta(\tau_B \mathbb{1}_{\mathcal{B}}) \geq \mathbb{E}_\eta \tau_B (1 - e^{-cN}).$$

(ii) *If  $N$  is large enough,*

$$(3.21) \quad \mathbb{P}_\sigma(\mathcal{D}) \geq 1 - e^{-cN}.$$

PROOF. Item (ii) follows from Lemma 2.2 with, for example,  $c = c_1/2$ . To prove item (i), we write  $\mathcal{B}^c = \{\tau_B^\eta \leq N^\kappa\} \cup \{S \geq N^\kappa\} \cup \{\mathcal{N} > M\}$ . Thus,

$$(3.22) \quad \mathbb{1}_{\mathcal{B}^c} = \mathbb{1}_{\{\tau_B^\eta \leq N^\kappa\}} + \mathbb{1}_{\{\tau_B^\eta > N^\kappa\}} \mathbb{1}_{\{S \geq N^\kappa\}} + \mathbb{1}_{\{\tau_B^\eta > N^\kappa\}} \mathbb{1}_{\{S < N^\kappa\}} \mathbb{1}_{\{\mathcal{N} > M\}}.$$

Inserting this into (3.19) and (3.20), there are three terms to bound. The first term is easy:

$$(3.23) \quad \mathbb{P}_\eta(\tau_B \leq N^\kappa) \leq N^\kappa \max_{\sigma': \underline{m}(\sigma') = \underline{m}} \mathbb{P}_{\sigma'}(\tau_B < \tau_{\underline{m}}) \leq N^\kappa e^{-c_1 N}.$$

The first inequality used the fact that in order to reach  $B$ , the process has to make one final excursion to  $B$  without return to the starting set  $\underline{m}$ , and that there are at most  $N^\kappa$  attempts to do so. The last inequality uses (2.20). The corresponding term for (3.20) is

$$(3.24) \quad \mathbb{E}_\eta(\tau_B \mathbb{1}_{\{\tau_B^\eta < N^\kappa\}}) \leq N^{2\kappa} e^{-c_1 N}.$$

The second term is also easy: first,

$$(3.25) \quad \mathbb{P}_\eta(\{\tau_B > N^\kappa\} \cap \{S \geq N^\kappa\}) \leq \mathbb{P}_\eta(S \geq N^\kappa)$$

and

$$(3.26) \quad \begin{aligned} \mathbb{E}_\eta(\tau_B \mathbb{1}_{\{\tau_B > N^\kappa\}} \mathbb{1}_{\{S \geq N^\kappa\}}) &\leq \sum_{\sigma'} (N^\kappa + \mathbb{E}_{\sigma'} \tau_B) \mathbb{P}_\eta(\eta_{N^\kappa} = \sigma'; S \geq N^\kappa) \\ &\leq (N^\kappa + \max_{\sigma'} \mathbb{E}_{\sigma'} \tau_B) \mathbb{P}_\eta(S \geq N^\kappa). \end{aligned}$$

Using the formula (1.4) with  $A = \{\sigma'\}$ , and bounding the corresponding capacity  $\text{cap}(\sigma', B) \geq e^{-c_3 N}$  from below in the crudest way (e.g., retaining a single one-dimensional path from  $\sigma'$  to  $B$ ; see [2]), one gets that

$$(3.27) \quad N^\kappa + \max_{\sigma'} \mathbb{E}_{\sigma'} \tau_B \leq e^{2c_3 N},$$

where  $c_3$  does not depend on  $n$ .

Next, we show that if  $\kappa > 2$  the probability  $\mathbb{P}_\eta(S \geq N^\kappa)$  is super-exponentially small. Indeed, since at each step the probability to flip each particular spin is bounded from below by  $(1 - \alpha)/N$ ,

$$(3.28) \quad \mathbb{P}_\eta(S \geq N^\kappa) \leq N \left(1 - \frac{1 - \alpha}{N}\right)^{N^\kappa} \leq e^{-c_4 N^{\kappa-1}}.$$

Finally, even the third term is easy:

$$(3.29) \quad \mathbb{E}_\eta(\mathbb{1}_{\{\tau_B > N^\kappa\}} \mathbb{1}_{\{S < N^\kappa\}} \mathbb{1}_{\{\mathcal{N} > M\}}) \leq \mathbb{P}_\eta(\mathcal{N} > M)$$

and, as in (3.26),

$$(3.30) \quad \begin{aligned} \mathbb{E}_\eta(\tau_B \mathbb{1}_{\{\tau_B > N^\kappa\}} \mathbb{1}_{\{S < N^\kappa\}} \mathbb{1}_{\{\mathcal{N} > M\}}) &\leq (N^\kappa + \max_{\sigma'} \mathbb{E}_{\sigma'} \tau_B) \mathbb{P}_\eta(\mathcal{N} > M) \\ &\leq e^{2c_3 N} \mathbb{P}_\eta(\mathcal{N} > M). \end{aligned}$$

It remains to bound  $\mathbb{P}_\eta(\mathcal{N} > M)$ . In order to do this, we split the time interval  $[0, S]$  into epochs

$$(3.31) \quad [0, S] = [0, S_{i_1}] \cup (S_{i_1}, S_{i_2}] \cup \dots \cup (S_{i_{N-1}}, S],$$

where  $\underline{i} = \{i_1, \dots, i_N\}$  is a permutation of  $\{1, \dots, N\}$  which is fixed by the order in which spins are flipped for the first time,

$$(3.32) \quad S_{i_1} < S_{i_2} < \dots < S_{i_N} = S.$$

Fix a particular permutation  $\underline{i}$  and let  $\mathcal{E}[\underline{i}]$  be the event that (3.32) happens. Let us first derive a lower bound on  $\mathbb{P}_\eta(\mathcal{E}[\underline{i}])$ . It is convenient to decompose  $\mathcal{E}[\underline{i}] = \bigcap_{k=0}^{N-1} \mathcal{E}_k[\underline{i}]$ , where

$$(3.33) \quad \begin{aligned} \mathcal{E}_0[\underline{i}] &= \{\text{No spin was flipped on } [0, S_{i_1} - 1]\} \\ &\cap \{\text{Spin } i_1 \text{ was flipped on } S_{i_1} - t \text{ step}\} \end{aligned}$$

and

$$(3.34) \quad \mathcal{E}_k[\underline{i}] = \{\text{No spin was flipped for the first time during } [S_{i_k}, S_{i_{k+1}} - 1)\} \\ \cap \{\text{Spin } i_{k+1} \text{ was flipped on } S_{i_{k+1}} - t \text{ step}\}.$$

Let  $\mathcal{N}_k$  be the number of times previously unflipped spins were attempted to flip during the interval  $(S_{i_k}, S_{i_{k+1}}]$ . Clearly  $\mathcal{N} = \sum_{k=0}^{N-1} \mathcal{N}_k$ .

In view of (1.7),

$$(3.35) \quad \mathbb{P}_\eta(\mathcal{E}_0[\underline{i}]; \mathcal{N}_0 = \ell_0) \leq \frac{\alpha^{\ell_0}}{N}.$$

To give an upper bound on the probability of the events  $\{\mathcal{E}_k[\underline{i}]; \mathcal{N}_k = \ell_k\}$ , for  $k > 0$ , we distinguish between two types of trials, which happen during the intervals  $(S_{i_k}, S_{i_{k+1}}]$ . First, one might choose yet unflipped spins from  $\{i_{k+1}, \dots, i_N\}$  but then fail to flip them. On the event  $\{\mathcal{N}_k = \ell_k\}$  this happens exactly  $\ell_k - 1$  times. Second, one might choose already flipped spins from the set  $\{i_1, \dots, i_k\}$ . The probability of the latter is  $k/N$ , whereas, according to (1.7), a uniform upper bound for the probability of the former option is  $\alpha(N - k)/N$ . Thus, if  $\mathcal{G}_k$  is the  $\sigma$ -field generated by  $\eta_{[0, S_k]}$ , then

$$(3.36) \quad \mathbb{P}_\eta(\mathcal{E}_k; \mathcal{N}_k = \ell_k | \mathcal{G}_k) \leq \frac{\alpha}{N} \left( \frac{\alpha(N - k)}{N} \right)^{\ell_k - 1} \left( \sum_{j=0}^{\infty} \left( \frac{k}{N} \right)^j \right)^{\ell_k} \\ = \frac{\alpha^{\ell_k}}{N - k}.$$

Therefore,

$$(3.37) \quad \mathbb{E}_\eta \left( \prod_{\ell=0}^k \mathbb{1}_{\mathcal{E}_\ell} \mathbb{1}_{\{\mathcal{N}_k = \ell_k\}} | \mathcal{G}_k \right) \leq \frac{\alpha^{\ell_k}}{N - k} \prod_{\ell=0}^{k-1} \mathbb{1}_{\mathcal{E}_\ell[\underline{i}]}$$

and, consequently,

$$(3.38) \quad \mathbb{P}_\eta(\mathcal{E}[\underline{i}]; \mathcal{N}_0 = \ell_0; \dots; \mathcal{N}_{N-1} = \ell_{N-1}) \leq \frac{1}{N!} \alpha^{\sum \ell_k}.$$

As a result, we get that

$$(3.39) \quad \mathbb{P}_\eta(\mathcal{N} > M) \leq \sum_{L > M} \alpha^L \binom{N + L}{L} \leq \sum_{L > M} e^{-\ln(1/\alpha)L} e^{N(\ln(1 + L/N) + 1)}.$$

For  $M \equiv c_2 N$  and providing that  $c_2$  is large enough, we finally obtain that

$$(3.40) \quad \mathbb{P}_\eta(\mathcal{N} > M) \leq e^{-c_5 N}$$

for a constant,  $c_5$ , increasing linearly with  $c_2$ . Putting all estimates together concludes the proof of the lemma.  $\square$

Notice that if  $\mathcal{A} \cap \mathcal{B}$  happens, then  $\underline{m}(\eta_t) \equiv \underline{m}(\sigma_t)$ . In particular,  $\mathcal{A} \cap \mathcal{B} \subset \{\tau_B^\sigma \geq N^k\}$  and hence  $\tau_B^\sigma = \tau_B^\eta$  on  $\mathcal{A} \cap \mathcal{B}$ .

Let us go back to the cycle decomposition (3.18). Using  $\tilde{\mathbb{E}}$  for the expectation on the enlarged probability space,

$$(3.41) \quad \mathbb{E}\tau_B^\sigma \geq \sum_{k=0}^\infty \tilde{\mathbb{E}} \left\{ \prod_{\ell=0}^{k-1} \mathbb{1}_{\mathcal{D}^\ell} (1 - \mathbb{1}_{\mathcal{A}^\ell} \mathbb{1}_{\mathcal{B}^\ell}) \right\} \tau_B^{k,\eta} \mathbb{1}_{\mathcal{A}^k} \mathbb{1}_{\mathcal{B}^k}.$$

Let  $\mathcal{F}_{\theta_k}$  be the  $\sigma$ -algebra generated by all the events and trajectories  $\mathcal{A}^\ell, \mathcal{B}^\ell, \mathcal{D}^\ell, \eta^\ell$  and  $\sigma(\theta_{\ell-1}, \theta_\ell), \ell \leq k$ . In view of the independence of the copies  $\{\eta^\ell, \underline{V}^\ell\}$ ,

$$(3.42) \quad \tilde{\mathbb{E}}(\tau_B^\eta \mathbb{1}_{\mathcal{A}^k} \mathbb{1}_{\mathcal{B}^k} | \mathcal{F}_{\theta_{k-1}}) = \mathbb{P}(\mathcal{A}) \mathbb{E}_\eta(\tau_B \mathbb{1}_{\mathcal{B}}) = (1 - \nu)^M \mathbb{E}_\eta(\tau_B \mathbb{1}_{\mathcal{B}}).$$

On the other hand,

$$(3.43) \quad \begin{aligned} \tilde{\mathbb{E}}(\mathbb{1}_{\mathcal{D}^\ell} (1 - \mathbb{1}_{\mathcal{A}^\ell} \mathbb{1}_{\mathcal{B}^\ell}) | \mathcal{F}_{\theta_{\ell-1}}) &\geq \tilde{\mathbb{E}}(1 - \mathbb{1}_{\mathcal{A}^\ell} \mathbb{1}_{\mathcal{B}^\ell}) - \max_{\sigma' \in \mathcal{S}^n[\underline{m}]} (1 - \mathbb{P}_{\sigma'}(\mathcal{D})) \\ &\geq 1 - (1 - \nu)^M - e^{-cN}. \end{aligned}$$

Altogether (recall that  $M = c_2 N$ ),

$$(3.44) \quad \begin{aligned} \mathbb{E}_\sigma \tau_B &\geq \mathbb{E}_\eta(\tau_B \mathbb{1}_{\mathcal{B}}) (1 - \nu)^{c_2 N} \sum_{k=0}^\infty (1 - (1 - \nu)^{c_2 N} - e^{-cN})^k \\ &\geq \mathbb{E}_\eta \tau_B \frac{1 - e^{-cN}}{1 + (1 - \nu)^{-c_2 N} e^{-cN}}, \end{aligned}$$

which tends to  $\mathbb{E}_\eta \tau_B$  if  $\nu < c/c_2$ . This concludes the proof of Theorem 1.1.

3.5. *Extension to the case  $\underline{m}(\sigma) \neq \underline{m}(\eta)$ .* Very little has to be changed if we replace the condition that we start in a configuration  $\sigma$  that has the same mesoscopic magnetization as  $\eta$ , but for which (1.11) still holds. In that case, we cannot start the coupling in the first cycle, so we simply have to wait until time  $\Delta_0$  (provided  $\mathcal{D}^0$  occurs, i.e.,  $\sigma_t$  does not hit  $B$  before that time). This means that we replace (3.18) by

$$(3.45) \quad \begin{aligned} \tau_B^\sigma &= \sum_{k=1}^\infty \left\{ \mathbb{1}_{\mathcal{D}^0} \prod_{\ell=1}^{k-1} \mathbb{1}_{\mathcal{D}^\ell} (1 - \mathbb{1}_{\mathcal{A}^\ell} \mathbb{1}_{\mathcal{B}^\ell}) \right\} (\theta_{k-1} + \tau_B^{k,\eta}) \mathbb{1}_{\mathcal{A}^k} \mathbb{1}_{\mathcal{B}^k} + \tau_B^\sigma \mathbb{1}_{(\mathcal{D}^0)^c} \\ &\quad + \tau_B^\sigma \sum_{k=1}^\infty \left\{ \mathbb{1}_{\mathcal{D}^0} \prod_{\ell=1}^{k-1} \mathbb{1}_{\mathcal{D}^\ell} (1 - \mathbb{1}_{\mathcal{A}^\ell} \mathbb{1}_{\mathcal{B}^\ell}) \right\} (1 - \mathbb{1}_{\mathcal{D}^k}) (1 - \mathbb{1}_{\mathcal{A}^k} \mathbb{1}_{\mathcal{B}^k}). \end{aligned}$$

We then proceed exactly as before to get

$$(3.46) \quad \begin{aligned} \mathbb{E}_\sigma \tau_B &\geq \mathbb{E}_\eta(\tau_B \mathbb{1}_{\mathcal{B}}) (1 - \nu)^{c_2 N} \sum_{k=0}^\infty (1 - e^{-c_2 N}) (1 - (1 - \nu)^{c_1 N} - e^{-c_2 N})^k \\ &\geq \mathbb{E}_\eta \tau_B \frac{(1 - e^{-cN})(1 - e^{-c_2 N})}{1 + (1 - \nu)^{-c_1 N} e^{-c_2 N}}, \end{aligned}$$

which is virtually equivalent to the previous case.

3.6. *The Laplace transform.* Next, we show that the same coupling can also be used to show that the Laplace transform of  $\tau_B$  depends very little on the initial conditions within a set  $A$ . Set  $T \equiv \mathbb{E}_{\nu_A} \tau_B$ .

PROPOSITION 3.4. *If  $A, B$  satisfy the hypothesis of Theorem 1.1, then, for every configurations  $\sigma, \eta \in A$  and  $\lambda \geq 0$ ,*

$$(3.47) \quad R_\sigma(\lambda) \equiv \mathbb{E}_\sigma(e^{-(\lambda/T)\tau_B}) = \mathbb{E}_\eta(e^{-(\lambda/T)\tau_B})(1 + o(1)).$$

The proof of Proposition 3.4 involves some estimates and computations that we collect in the following lemmas.

LEMMA 3.5. *There exists a constant,  $c > 0$ , independent of  $n$ , such that, for any  $\eta \in A$ ,*

$$(3.48) \quad \mathbb{E}_\eta(\mathbb{1}_B e^{-(\lambda/T)\tau_B}) \geq \mathbb{E}_\eta(e^{-(\lambda/T)\tau_B})(1 - e^{-cN}).$$

PROOF. The proof is similar to that of (3.20) and uses some of the estimates given there. The aim is to prove that

$$(3.49) \quad \mathbb{E}_\eta(\mathbb{1}_{B^c} e^{-(\lambda/T)\tau_B}) \leq \mathbb{E}_\eta(e^{-(\lambda/T)\tau_B})e^{-cN}.$$

By Jensen’s inequality, for every  $\eta \in A$ ,

$$(3.50) \quad \mathbb{E}_\eta(e^{-(\lambda/T)\tau_B}) \geq e^{-(\lambda/T)\mathbb{E}_\eta \tau_B} = e^{-\lambda}(1 + o(1)),$$

where the second line follows from the pointwise estimate on  $\mathbb{E}_\eta \tau_B$  that was proven in the previous subsections. To prove (3.49), it is enough to notice that, by Lemma 3.3,

$$(3.51) \quad \mathbb{E}_\eta(\mathbb{1}_{B^c} e^{-(\lambda/T)\tau_B}) \leq e^{-cN}. \quad \square$$

PROOF OF PROPOSITION 3.4. For simplicity, we consider the case when  $\underline{m}(\sigma) = \underline{m}(\eta) \equiv \underline{m}$ . Analogously to (3.18), we obtain

$$(3.52) \quad \begin{aligned} & \mathbb{E}_\sigma(e^{-(\lambda/T)\tau_B}) \\ &= \tilde{\mathbb{E}}\left(\sum_{k=0}^{\infty} e^{-(\lambda/T)(\theta_{k-1} + \tau_B^{k,\eta})} \mathbb{1}_{A^k} \mathbb{1}_{B^k} \prod_{\ell=0}^{k-1} \mathbb{1}_{D^\ell} (1 - \mathbb{1}_{A^\ell} \mathbb{1}_{B^\ell})\right) \\ &+ \tilde{\mathbb{E}}\left(e^{-(\lambda/T)\tau_B^\sigma} \sum_{k=0}^{\infty} (1 - \mathbb{1}_{D^k})(1 - \mathbb{1}_{A^k} \mathbb{1}_{B^k}) \prod_{\ell=0}^{k-1} \mathbb{1}_{D^\ell} (1 - \mathbb{1}_{A^\ell} \mathbb{1}_{B^\ell})\right) \\ &\leq \sum_{k=0}^{\infty} \tilde{\mathbb{E}}\left(e^{-(\lambda/T)\tau_B^{k,\eta}} \mathbb{1}_{A^k} \mathbb{1}_{B^k} \prod_{\ell=0}^{k-1} \mathbb{1}_{D^\ell} (1 - \mathbb{1}_{A^\ell} \mathbb{1}_{B^\ell})\right) \\ &+ \sum_{k=0}^{\infty} \tilde{\mathbb{E}}\left((1 - \mathbb{1}_{D^k})(1 - \mathbb{1}_{A^k} \mathbb{1}_{B^k}) \prod_{\ell=0}^{k-1} \mathbb{1}_{D^\ell} (1 - \mathbb{1}_{A^\ell} \mathbb{1}_{B^\ell})\right). \end{aligned}$$

Now, for every  $k, \ell \geq 0$ , as in (3.42),

$$(3.53) \quad \tilde{\mathbb{E}}(\mathbb{1}_{\mathcal{A}^k} \mathbb{1}_{\mathcal{B}^\ell} e^{-(\lambda/T)\tau_B^{k,\eta}} | \mathcal{F}_{\theta_{k-1}}) \leq (1 - \nu)^M \mathbb{E}_\eta(e^{-(\lambda/T)\tau_B}).$$

Moreover, as in (3.43),

$$(3.54) \quad \begin{aligned} \tilde{\mathbb{E}}(\mathbb{1}_{\mathcal{D}^{(\ell)}} (1 - \mathbb{1}_{\mathcal{A}^{(\ell)}} \mathbb{1}_{\mathcal{B}^\ell}) | \mathcal{F}_{\theta_{\ell-1}}) &\leq \tilde{\mathbb{E}}(1 - \mathbb{1}_{\mathcal{A}^{(\ell)}} \mathbb{1}_{\mathcal{B}^\ell} | \mathcal{F}_{\theta_{\ell-1}}) \\ &= 1 - \mathbb{P}(\mathcal{A})\mathbb{P}(\mathcal{B}) \\ &\leq 1 - (1 - \nu)^M (1 - e^{-cN}). \end{aligned}$$

This last estimate, together with (3.21) of Lemma 3.3, shows that the term in the last line of (3.52) is smaller than

$$(3.55) \quad \sum_{k=0}^{\infty} e^{-cN} (1 - (1 - \nu)^M (1 - e^{-cN}))^k \leq 2e^{-N(c-c_2\nu)}.$$

Combining these estimates, we arrive at

$$(3.56) \quad \begin{aligned} \mathbb{E}_\sigma(e^{-(\lambda/T)\tau_B}) &\leq \mathbb{E}_\eta(e^{-(\lambda/T)\tau_B}) (1 - \nu)^{c_2N} \sum_{k=0}^{\infty} (1 - (1 - \nu)^{c_2N} (1 - e^{-cN}))^k \\ &\leq \mathbb{E}_\eta(e^{-(\lambda/T)\tau_B}) (1 - e^{-cN}) + 2e^{-N(c-c_2\nu)} \\ &= \mathbb{E}_\eta(e^{-(\lambda/T)\tau_B}) (1 + 3e^{-N(c-c_2\nu)}), \end{aligned}$$

which tends to  $\mathbb{E}_\eta(e^{-(\lambda/T)\tau_B})$  if  $\nu < c/c_2$ .  $\square$

**4. Renewal and the exponential distribution for the RFCW.** We will use the results of Section 2 and the notation introduced therein. In particular, for each  $n$  fixed, we set  $A = \mathcal{S}^n[\underline{m}^*]$  and  $A_\delta = \mathcal{S}^n[\mathbf{A}_\delta]$ , where  $\mathbf{A}_\delta$  is the mesoscopic  $\delta$ -neighborhood of  $\underline{m}^*$ . In the sequel we choose  $n$  appropriately large and  $\delta$  appropriately small.

In the case of the RFCW model, we prove the convergence of the law of the normalized metastable time,  $\tau_B$ , to an exponential distribution, via convergence of the Laplace transform,  $R_\sigma(\lambda)$ , defined in (3.47). The proof of the latter is based on renewal arguments.

**4.1. Renewal equations.** By Proposition 3.4, instead of studying the process starting in a given point,  $\sigma$ , for which no exact renewal equation will hold, it is enough to study the process starting on a suitable measure on  $A$ , for which such a relation will be shown to hold. For  $\lambda \geq 0$ , let  $\rho_\lambda$  denote the probability measure on  $A$  that satisfies the equation

$$(4.1) \quad \sum_{\sigma \in A} \rho_\lambda(\sigma) \mathbb{E}_\sigma(e^{-(\lambda/T)\tau_A} \mathbb{1}_{\tau_A < \tau_B} \mathbb{1}_{\sigma(\tau_A) = \sigma'}) = C(\lambda) \rho_\lambda(\sigma')$$

for all  $\sigma' \in A$ , where

$$(4.2) \quad C(\lambda) = \mathbb{E}_{\rho_\lambda} (e^{-(\lambda/T)\tau_A} \mathbb{1}_{\tau_A < \tau_B}).$$

Existence and uniqueness of such a measure follow in a standard way from the Perron–Frobenius theorem.

The usefulness of this definition comes from the fact that the Laplace transform of  $\tau_B$  started in this measure satisfies an exact renewal equation.

LEMMA 4.1. *Let  $R_{\rho_\lambda}(\lambda) = \sum_{\sigma} \rho_\lambda(\sigma) R_\sigma(\lambda)$ . Then*

$$(4.3) \quad R_{\rho_\lambda}(\lambda) = \frac{\mathbb{E}_{\rho_\lambda} (e^{-(\lambda/T)\tau_B} \mathbb{1}_{\tau_B < \tau_A})}{1 - \mathbb{E}_{\rho_\lambda} (e^{-(\lambda/T)\tau_A} \mathbb{1}_{\tau_A < \tau_B})}.$$

PROOF. Using that  $1 = \mathbb{1}_{\tau_B < \tau_A} + \mathbb{1}_{\tau_A < \tau_B}$  and the strong Markov property, we see that

$$(4.4) \quad \begin{aligned} R_{\rho_\lambda}(\lambda) &= \mathbb{E}_{\rho_\lambda} (e^{-(\lambda/T)\tau_B} \mathbb{1}_{\tau_B < \tau_A}) \\ &+ \sum_{\sigma' \in A} \mathbb{E}_{\rho_\lambda} (e^{-(\lambda/T)\tau_A} \mathbb{1}_{\tau_A < \tau_B} \mathbb{1}_{\sigma(\tau_A) = \sigma'}) \mathbb{E}_{\sigma'} e^{-(\lambda/T)\tau_B} \\ &= \mathbb{E}_{\rho_\lambda} (e^{-(\lambda/T)\tau_B} \mathbb{1}_{\tau_B < \tau_A}) + \sum_{\sigma' \in A} C(\lambda) \rho_\lambda(\sigma') \mathbb{E}_{\sigma'} e^{-(\lambda/T)\tau_B}. \end{aligned}$$

Equation (4.3) is now immediate.  $\square$

4.2. *Convergence.* As a result of the representation (4.3), Theorem 1.2 will follow from (3.47) once we prove the following lemma.

LEMMA 4.2. *With the notation from Lemma 4.1, for any  $\lambda \geq 0$ ,*

$$(4.5) \quad \lim_{N \rightarrow \infty} \frac{\mathbb{E}_{\rho_\lambda} (e^{-(\lambda/T)\tau_B} \mathbb{1}_{\tau_B < \tau_A})}{1 - \mathbb{E}_{\rho_\lambda} (e^{-(\lambda/T)\tau_A} \mathbb{1}_{\tau_A < \tau_B})} = \frac{1}{1 + \lambda}.$$

PROOF. The proof of this lemma comprises seven steps.

STEP 1. Define  $T_\lambda = \mathbb{E}_{\rho_\lambda}$ . We claim:

LEMMA 4.3. *There exists  $c_6 > 0$ , such that, for any  $\lambda \geq 0$  fixed,*

$$(4.6) \quad T_\lambda = \frac{\mathbb{E}_{\rho_\lambda} \tau_{A \cup B}}{\mathbb{P}_{\rho_\lambda}(\tau_B < \tau_A)} (1 + o(e^{-c_6 N})).$$

Indeed,

$$(4.7) \quad \begin{aligned} T_\lambda &= \mathbb{E}_{\rho_\lambda} (\tau_B \mathbb{1}_{\{\tau_B < \tau_A\}}) + \mathbb{E}_{\rho_\lambda} (\tau_B \mathbb{1}_{\{\tau_A < \tau_B\}}) \\ &= \mathbb{E}_{\rho_\lambda} \tau_{A \cup B} + \mathbb{E}_{\rho_\lambda} (\mathbb{1}_{\{\tau_A < \tau_B\}} \mathbb{E}_{\sigma(\tau_A)} \tau_B) \\ &= \mathbb{E}_{\rho_\lambda} \tau_{A \cup B} + T_\lambda \mathbb{P}_{\rho_\lambda}(\tau_A < \tau_B) + \mathbb{E}_{\rho_\lambda} (\mathbb{1}_{\{\tau_A < \tau_B\}} (\mathbb{E}_{\sigma(\tau_A)} \tau_B - T_\lambda)). \end{aligned}$$

However, by the invariance of  $\rho_\lambda$ ,

$$(4.8) \quad \mathbb{E}_{\rho_\lambda}(\mathbb{1}_{\{\tau_A < \tau_B\}} e^{-(\lambda/T)\tau_A} (\mathbb{E}_{\sigma(\tau_A)} \tau_B - T_\lambda)) = 0.$$

It follows that the absolute value of the last term in (4.7) is bounded above as

$$(4.9) \quad \begin{aligned} & \mathbb{E}_{\rho_\lambda}((1 - e^{-(\lambda/T)\tau_A}) \mathbb{1}_{\{\tau_A < \tau_B\}}) \max_{\sigma \in A} |\mathbb{E}_\sigma \tau_B - T_\lambda| \\ & \leq \lambda \max_{\sigma \in A} \left| \frac{\mathbb{E}_\sigma \tau_B - T_\lambda}{T} \right| \mathbb{E}_{\rho_\lambda}(\tau_A \mathbb{1}_{\{\tau_A < \tau_B\}}) \\ & = o(e^{-cN/2}) \mathbb{E}_{\rho_\lambda}(\tau_A \mathbb{1}_{\{\tau_A < \tau_B\}}), \end{aligned}$$

where we used (1.12) in the last step. This implies the claim of the lemma.

STEP 2. Control of  $\rho_\lambda$ -measure.

LEMMA 4.4. *There exists  $c_7 < \infty$ , such that for any  $n$  [and hence  $\epsilon = \epsilon(n)$ ] fixed,*

$$(4.10) \quad \max_{\sigma \in A} \frac{\rho_\lambda(\sigma)}{\mu(\sigma)/\mu(A)} \leq e^{c_7 \epsilon N}$$

as soon as  $N$  is large enough.

PROOF. In order to prove (4.10), first of all, note that by reversibility

$$(4.11) \quad \sum_{\sigma' \in A} \mu(\sigma') \mathbb{P}_{\sigma'}(\tau_A^r < \tau_B; \sigma(\tau_A^r) = \sigma) = \mu(\sigma) \mathbb{P}_\sigma(\tau_A^r < \tau_B),$$

where  $\tau_A^r$  is the  $r$ th hitting time of  $A$ . Assume now that we are able to prove that there exists  $r$  and  $M$  such that

$$(4.12) \quad \mathbb{P}_\eta(\tau_A^r < \tau_B; \sigma(\tau_A^r) = \sigma) \leq (1 - \epsilon)^{-M} \mathbb{P}_{\sigma'}(\tau_A^r < \tau_B; \sigma(\tau_A^r) = \sigma),$$

uniformly in  $\eta, \sigma, \sigma' \in A$ . In view of (4.1), this would imply

$$(4.13) \quad \begin{aligned} \rho_\lambda(\sigma) & \leq \frac{1}{C(\lambda)^r} \sum_{\eta} \rho_\lambda(\eta) \mathbb{P}_\eta(\tau_A^r < \tau_B; \sigma(\tau_A^r) = \sigma) \\ & \leq \frac{(1 - \epsilon)^{-M}}{C(\lambda)^r} \mathbb{P}_{\sigma'}(\tau_A^r < \tau_B; \sigma(\tau_A^r) = \sigma). \end{aligned}$$

Multiplying both sides above by  $\mu(\sigma')$  and applying (4.11), we conclude that (4.12) implies that

$$(4.14) \quad \rho(\sigma) \leq \frac{(1 - \epsilon)^{-M}}{C(\lambda)^r} \frac{\mu(\sigma)}{\mu(A)} \mathbb{P}_\sigma(\tau_A^r < \tau_B),$$

uniformly in  $\sigma \in A$ . The target (4.10), therefore, will be a consequence of the following two claims: there exists  $c > 0$ , such that, independently of the coarse graining parameter  $n$ ,

$$(4.15) \quad C(\lambda) \geq 1 - e^{-cN}$$



as soon as  $N$  is sufficiently large. Furthermore, for sufficiently large  $c_2$  and  $\kappa$ , (4.12) holds with  $M = c_2N$  and  $r = N^\kappa$ .

We first show that (4.15) holds. By the uniform bound (2.21) and Jensen’s inequality, it follows that

$$\begin{aligned}
 C(\lambda) &\geq (1 - e^{-cN}) \sum_{\sigma} \rho_{\lambda}(\sigma) \mathbb{E}_{\sigma}(e^{-(\lambda/T)\tau_A} | \tau_A < \tau_B) \\
 (4.16) \quad &\geq (1 - e^{-cN}) \exp\left\{-\frac{\lambda}{T} \sum_{\sigma} \rho_{\lambda}(\sigma) \mathbb{E}_{\sigma}(\tau_A | \tau_A < \tau_B)\right\} \\
 &\geq (1 - e^{-cN}) \exp\left\{-\frac{\lambda \mathbb{E}_{\rho_{\lambda}}(\tau_A \mathbb{1}_{\{\tau_A < \tau_B\}})}{T(1 - e^{-cN})}\right\}.
 \end{aligned}$$

By (4.6),

$$(4.17) \quad \mathbb{E}_{\rho_{\lambda}}(\tau_A \mathbb{1}_{\{\tau_A < \tau_B\}}) \leq \frac{T \mathbb{P}_{\rho_{\lambda}}(\tau_B < \tau_A)}{1 + o(e^{-c_8N})},$$

and (4.15) follows by (2.21) and (1.12).

Next, we show that (4.12) holds. There exists  $c_8 < \infty$  such that

$$(4.18) \quad \mathbb{P}_{\eta}(\tau_A^r < \tau_B; \sigma_{\tau_A^r} = \sigma) \geq e^{-c_8N} \mathbb{P}_{\eta}(\tau_A^r < \tau_B),$$

uniformly in  $\sigma, \eta \in A$ . This is a rough estimate: by the Markov property,

$$(4.19) \quad \mathbb{P}_{\eta}(\tau_A^r < \tau_B; \sigma_{\tau_A^r} = \sigma) \geq \mathbb{P}_{\eta}(\tau_A^r < \tau_B) \min_{\eta' \in A} \mathbb{P}_{\eta'}(\tau_A < \tau_B; \sigma_{\tau_A} = \sigma).$$

Let  $\eta' \in A$  and let the Hamming distance between  $\sigma$  and  $\eta'$  be  $K$ . Then we can reach  $\sigma$  from  $\eta'$  by flipping exactly  $K$  spins; since this can be done in  $K!$  orders, and each flip has probability at least  $((1 - \alpha)/N)$  by (1.7), we see that

$$(4.20) \quad \mathbb{P}_{\eta'}(\tau_A < \tau_B; \sigma_{\tau_A} = \sigma) \geq \frac{K!}{N^K} (1 - \alpha)^K,$$

and (4.18) follows.

Next, let  $\eta \in A$  and consider a dynamics starting from  $\eta$ . We shall try to couple it with a dynamics starting from  $\sigma'$  using just one *basic coupling attempt*. Employing the same notation as in Section 3.2, we know [see (3.28) and (3.40)] that for  $\kappa > 2$  and  $M = c_2N$ ,

$$(4.21) \quad \mathbb{P}_{\eta}(\mathcal{S} > N^{\kappa}, \mathcal{N} > M) \leq e^{-c_9N},$$

where  $c_9$  grows linearly with  $c_2$ . In the sequel, we choose  $c_2$  so large that  $c_9$  becomes larger than the constant  $c_8$  in (4.18).

Let us redefine the event  $\mathcal{B}$  in (3.13) as  $\mathcal{B} = \{\mathcal{S} \leq N^{\kappa}\} \cap \{\mathcal{N} \leq M\}$ . The coins  $V_1, \dots, V_M$  and the event  $\mathcal{A} = \{\bigvee_{i=1}^M V_i = 1\}$  remain the same. Consider the enlarged probability space  $(\tilde{\Omega}, \tilde{\mathbb{P}})$  which corresponds to a single basic coupling attempt to couple a dynamics  $\sigma(t)$  from  $\sigma'$  to the dynamics  $\eta(t)$  which starts at  $\eta$ .

The coupling is successful if and only if the event  $\mathcal{A} \cap \mathcal{B}$ , which depends on at most  $N^\kappa$  steps, happens. Therefore,

$$\begin{aligned}
 & \mathbb{P}_{\sigma'}(\tau_A^r < \tau_B; \sigma(\tau_A^r) = \sigma) \\
 (4.22) \quad & \geq \tilde{\mathbb{P}}(N^\kappa \leq \tau_A^r < \tau_B; \eta(\tau_A^r) = \sigma; \mathcal{A}; \mathcal{B}) \\
 & = \mathbb{P}_\eta(N^\kappa \leq \tau_A^r < \tau_B; \eta(\tau_A^r) = \sigma; \mathcal{B})\tilde{\mathbb{P}}(\mathcal{A}) \\
 & = \mathbb{P}_\eta(N^\kappa \leq \tau_A^r < \tau_B; \eta(\tau_A^r) = \sigma; \mathcal{B})(1 - \epsilon)^M.
 \end{aligned}$$

Now, let us choose  $r = N^\kappa$ . In particular, the constraint  $N^\kappa \leq \tau_A^r$  becomes redundant. By (2.21) and in view of (4.18) and our choice of  $M$  which leads to a large  $c_9$  in (4.21), there exists  $c_{10} > 0$  such that

$$\begin{aligned}
 & \mathbb{P}_\eta(\tau_A^r < \tau_B; \eta(\tau_A^r) = \sigma; \mathcal{B}) \\
 (4.23) \quad & \geq \mathbb{P}_\eta(\tau_A^r < \tau_B; \eta(\tau_A^r) = \sigma) - \mathbb{P}_\eta(\mathcal{B}^c) \\
 & \geq \mathbb{P}_\eta(\tau_A^r < \tau_B; \eta(\tau_A^r) = \sigma)(1 - e^{-c_{10}N}).
 \end{aligned}$$

Equation (4.12) follows.  $\square$

STEP 3. The following crucial bound, to which we refer to a *uphill lemma*, will be proven in the next subsection.

LEMMA 4.5. *There exists  $c_{11} > 0$  such that*

$$(4.24) \quad \mathbb{E}_{\rho_\lambda}(\tau_B \mathbb{1}_{\tau_B < \tau_A}) \leq e^{-c_{11}N} \mathbb{E}_{\rho_\lambda}(\tau_A \mathbb{1}_{\tau_A < \tau_B}).$$

REMARK. Intuitively, the bound (4.24) should follow from the decomposition

$$\begin{aligned}
 (4.25) \quad \mathbb{E}_\sigma \tau_{A \cup B} & = \mathbb{P}_\sigma(\tau_A < \tau_B) \mathbb{E}_\sigma(\tau_A | \tau_A < \tau_B) \\
 & \quad + \mathbb{P}_\sigma(\tau_B < \tau_A) \mathbb{E}_\sigma(\tau_B | \tau_B < \tau_A)
 \end{aligned}$$

since the first probability on the right-hand side is close to one, the second is exponentially small, and the two conditional expectations should be of the same order. It seems, however, remarkably difficult to establish such a result uniformly in the starting point  $\sigma \in A$ , for the same reasons why the pointwise control of mean exit times is difficult.

We shall proceed with the proof assuming that (4.24) holds.

STEP 4. In view of (4.6), a look at (4.24) reveals that the conditional expectation

$$(4.26) \quad \frac{\mathbb{E}_{\rho_\lambda}(\tau_B | \tau_B < \tau_A)}{T} = o(e^{-c_{11}N}).$$

Using that, for  $x \geq 0$ ,  $1 \geq e^{-x} \geq 1 - x$ , it follows that the numerator in (4.3) satisfies

$$(4.27) \quad \mathbb{E}_{\rho_\lambda}(e^{-(\lambda/T)\tau_B} \mathbb{1}_{\tau_B < \tau_A}) = \mathbb{P}_{\rho_\lambda}(\tau_B < \tau_A)(1 + o(e^{-c_{11}N})).$$

STEP 5. Let us turn now to the denominator in (4.3). We rewrite it as

$$(4.28) \quad \mathbb{P}_{\rho_\lambda}(\tau_B < \tau_A) \left( 1 + \lambda \frac{\mathbb{E}_{\rho_\lambda}((1 - e^{-(\lambda/T)\tau_A})\mathbb{1}_{\{\tau_A < \tau_B\}})}{\lambda \mathbb{P}_{\rho_\lambda}(\tau_B < \tau_A)} \right).$$

Using (4.6) for  $1/\mathbb{P}_{\rho_\lambda}(\tau_B < \tau_A)$ , we are left with the computation of

$$(4.29) \quad \frac{T}{\lambda \mathbb{E}_{\rho_\lambda}(\tau_A \mathbb{1}_{\{\tau_A < \tau_B\}})} \mathbb{E}_{\rho_\lambda}((1 - e^{-(\lambda/T)\tau_A})\mathbb{1}_{\{\tau_A < \tau_B\}}).$$

Since,

$$(4.30) \quad \mathbb{E}_{\rho_\lambda}((1 - e^{-(\lambda/T)\tau_A})\mathbb{1}_{\{\tau_A < \tau_B\}}) = \frac{\lambda}{T} \int_0^1 \mathbb{E}_{\rho_\lambda}(e^{-(s\lambda/T)\tau_A} \tau_A \mathbb{1}_{\{\tau_A < \tau_B\}}) \, ds,$$

we deduce that the expression in (4.29) belongs to the interval

$$(4.31) \quad \left[ \frac{\mathbb{E}_{\rho_\lambda}(e^{-(\lambda/T)\tau_A} \tau_A \mathbb{1}_{\{\tau_A < \tau_B\}})}{\mathbb{E}_{\rho_\lambda}(\tau_A \mathbb{1}_{\{\tau_A < \tau_B\}})}, 1 \right].$$

The target (4.5) follows once we show that

$$(4.32) \quad \lim_{N \rightarrow \infty} \frac{\mathbb{E}_{\rho_\lambda}(e^{-(\lambda/T)\tau_A} \tau_A \mathbb{1}_{\{\tau_A < \tau_B\}})}{\mathbb{E}_{\rho_\lambda}(\tau_A \mathbb{1}_{\{\tau_A < \tau_B\}})} = 1.$$

It is clear that (4.32) follows as soon as we check that there exists a sequence  $\alpha_N \downarrow 0$ , such that

$$(4.33) \quad \lim_{N \rightarrow \infty} \frac{\mathbb{E}_{\rho_\lambda}(\tau_A \mathbb{1}_{\{\tau_A < \tau_B\}} \mathbb{1}_{\{\tau_A < \alpha_N T\}})}{\mathbb{E}_{\rho_\lambda}(\tau_A \mathbb{1}_{\{\tau_A < \tau_B\}})} = 1.$$

This will be our next goal.

Let  $B_\delta = S_N \setminus A_\delta$ . Our proof of (4.33) is based on the following decomposition:

$$(4.34) \quad \begin{aligned} \mathbb{E}_{\rho_\lambda}(\tau_A \mathbb{1}_{\{\tau_A < \tau_B\}} \mathbb{1}_{\{\tau_A > \alpha_N T\}}) &\leq \mathbb{E}_{\rho_\lambda}(\tau_A \mathbb{1}_{\{\tau_A < \tau_{B_\delta}\}} \mathbb{1}_{\{\tau_A > \alpha_N T\}}) \\ &\quad + \mathbb{E}_{\rho_\lambda}(\tau_A \mathbb{1}_{\{\tau_{B_\delta} < \tau_A < \tau_B\}}) \\ &\equiv \text{I}_\delta + \text{II}_\delta. \end{aligned}$$

The logic behind this decomposition should be transparent: the conditional (on  $\tau_A < \tau_B$ ) landscape should have the global mesoscopic minima at  $A$ . The term  $\text{I}_\delta$  is a local one and should be small, since the dynamics cannot spend too much time inside a local well  $A_\delta$  without hitting  $A$ . On the other hand, the term  $\text{II}_\delta$  should be small because of the price paid for the uphill run toward  $B_\delta$  before hitting  $A$ . We claim that there exists  $\alpha_N \downarrow 0$  and  $c > 0$  such that

$$(4.35) \quad \max\{\text{I}_\delta, \text{II}_\delta\} \leq e^{-cN} \mathbb{E}_{\rho_\lambda}(\tau_A \mathbb{1}_{\{\tau_A < \tau_B\}}).$$

Evidently, (4.33) is a consequence of (4.35).

STEP 6. Bound on  $I_\delta$ . The term  $I_\delta$  is bounded above as

$$(4.36) \quad I_\delta \leq \max_{\sigma \in A} \mathbb{E}_\sigma (\tau_{A \cup B_\delta} \mathbb{1}_{\{\tau_{A \cup B_\delta} > \alpha_N T\}}).$$

The right-hand side of (4.36) depends on the dynamics in a  $\delta$ -neighborhood of a nondegenerate local minimum  $A = \mathcal{S}^n[\underline{m}^*]$ . We try to formalize an intuitive idea that such dynamics mixes up on time scale much shorter than  $T$  and cannot afford spending  $\alpha_N T$  units of time without hitting  $A \cup B_\delta$ . This is a somewhat coarse estimate. Let us start with estimating hitting times from equilibrium measure over mesoscopic slots:

LEMMA 4.6. *Let  $A_\delta$  and  $B_\delta$  be as defined above. Then there exists  $c(\delta)$ , satisfying  $c(\delta) \downarrow 0$ , as  $\delta \downarrow 0$ , such that, for all  $\underline{m}' \in \mathbf{A}_\delta \setminus \underline{m}^*$ ,*

$$(4.37) \quad \mathbb{E}_{\nu_{\underline{m}'}} \tau_{A \cup B_\delta} \leq e^{c(\delta)N},$$

where  $\nu_{\underline{m}'}$  is the probability measure on  $\mathcal{S}^n[\underline{m}']$ , which we referred to in (1.4).

PROOF. By formula (1.4), we have that

$$(4.38) \quad \begin{aligned} \mathbb{E}_{\nu_{\underline{m}'}} \tau_{A \cup B_\delta} &= \frac{1}{\text{cap}(\underline{m}', A \cup B_\delta)} \sum_{\sigma \in A_\delta \setminus A} \mu_{\beta, N}(\sigma) h_{\mathcal{S}^n[\underline{m}'], S[A \cup B_\delta]}(\sigma) \\ &\leq \frac{1}{\text{cap}(\underline{m}', A)} \sum_{\sigma \in A_\delta \setminus A} \mu_{\beta, N}(\sigma) = \frac{\mu_{\beta, N}(A \setminus A_\delta)}{\text{cap}(\underline{m}', A)}. \end{aligned}$$

Note that we used here only the crudest possible estimate on the harmonic function  $h_{\mathcal{S}^n[\underline{m}'], A \cup B_\delta}(\sigma)$ , but the results of [1] do not give us anything much better. It remains to bound the capacity  $\text{cap}(\underline{m}', A)$  from below. However, this is relatively easy using the methods explained in Section 5 of [1], to which we refer for further details. One gets that

$$(4.39) \quad \text{cap}(\underline{m}', A) \geq e^{-c\delta\epsilon N} \mu_{\beta, N}(\underline{m}'). \quad \square$$

As a consequence we obtain the following lemma.

LEMMA 4.7. *Let  $A_\delta$  and  $B_\delta$  be as defined above. Then there exists  $c(\delta)$  satisfying  $c(\delta) \downarrow 0$  as  $\delta \downarrow 0$ , such that, for all  $\eta \in A_\delta \setminus A$ ,*

$$(4.40) \quad \mathbb{P}_\eta (\tau_{A \cup B_\delta} \leq 2e^{c(\delta)N}) \geq \frac{(1 - \nu(n))^M}{3},$$

where  $1 - \nu(n)$  is the probability (3.4) of a successful single coin-flip and  $M = c_2 N$  is the number of coins.

PROOF. As the formulation of the lemma suggests, we use the basic coupling as described in the preceding section: let  $\underline{m}' \in \mathbf{A}_\delta$  and  $\eta, \sigma \in \mathcal{S}^n[\underline{m}']$ . Define the event  $\mathcal{B}$  as in (3.13). In fact, since we are interested in  $\tau_{A \cup B_\delta}$ , the first constraint in (3.13) becomes redundant and we can redefine  $\mathcal{B}$  simply as

$$(4.41) \quad \mathcal{B} = \{S < N^\kappa\} \cap \{\mathcal{N} < M\}.$$

Then, performing our basic coupling attempt we infer that, for any  $\eta, \sigma \in \mathcal{S}[\underline{m}']$ ,

$$(4.42) \quad \mathbb{P}_\eta(\tau_{A \cup B_\delta} \leq 2e^{c(\delta)N}) \geq (1 - \nu(n))^M \mathbb{P}_\sigma(\tau_{A \cup B_\delta} \leq e^{c(\delta)N}; \mathcal{B}).$$

By Lemma 4.6 and Chebyshev’s inequality

$$(4.43) \quad \mathbb{P}_{\nu_{\underline{m}'}}(\tau_{A \cup B_\delta} \leq 2e^{c(\delta)N}) \geq \frac{1}{2}$$

and, in view of the bound (3.19), (4.40) follows.  $\square$

Let us go back to (4.36). By Lemma 4.7,

$$(4.44) \quad \max_{\sigma \in A} \mathbb{P}_\sigma(\tau_{A \cup B_\delta} > k2e^{c(\delta)N}) \leq \left(1 - \frac{(1 - \nu(n))^M}{3}\right)^k.$$

Therefore, as follows by a straightforward application of the tail formula,

$$(4.45) \quad I_\delta \leq e^{-c_{12}N}$$

as soon as

$$(4.46) \quad \alpha_N T > 3c_{13}N e^{(c(\delta) + \nu(n))N}.$$

Since  $T \sim e^{CN}$  with  $C > 0$  being, of course, independent of our choice of  $\delta$  and  $n$ , it is always possible to tune the parameters  $\delta, n$  and  $\alpha_N \downarrow 0$  in such a way that (4.46) holds.

STEP 7. Bound on  $\Pi_\delta$ . Note that

$$(4.47) \quad \begin{aligned} \mathbb{E}_{\rho_\lambda}(\tau_A \mathbb{1}_{\{\tau_{B_\delta} < \tau_A < \tau_B\}}) &= \mathbb{E}_{\rho_\lambda}(\tau_{B_\delta} \mathbb{1}_{\{\tau_{B_\delta} < \tau_A\}}) \\ &\quad + \mathbb{E}_{\rho_\lambda}(\mathbb{1}_{\{\tau_{B_\delta} < \tau_A\}} \mathbb{E}_{\sigma(\tau_{B_\delta})}(\tau_A \mathbb{1}_{\{\tau_A < \tau_B\}})). \end{aligned}$$

By the *Uphill lemma* [see (4.24) above] the first term in (4.47) is negligible with respect to  $\mathbb{E}_{\rho_\lambda} \tau_A \mathbb{1}_{\{\tau_A < \tau_{B_\delta}\}}$ . Therefore, the bulk of the remaining work is to find an appropriate upper bound on the second term in (4.47).

By the *Downhill lemma* [see (4.61) below] we would be in good shape if we would have the original reversible measure  $\mu$  instead of the  $\rho_\lambda$  eigen-measure defined in (4.1). Namely, as it is explained in the end of Section 4.3, (4.61) implies that, independently of  $n$ , there exists  $c_\delta > 0$  such that

$$(4.48) \quad \frac{1}{\mu(A)} \sum_{\sigma \in A} \mu(\sigma) \mathbb{E}_\sigma(\mathbb{1}_{\{\tau_{B_\delta} < \tau_A\}} \mathbb{E}_{\sigma(\tau_{B_\delta})}(\tau_A \mathbb{1}_{\{\tau_A < \tau_B\}})) \leq e^{-c_\delta N}$$

as soon as  $N$  is large enough. However, since we have already established in (4.10) that  $\rho_\lambda$  is, up to arbitrary small exponential corrections, controlled by  $\mu$ , it follows that the second term in (4.47) is exponentially small and hence also negligible with respect to  $\mathbb{E}_{\rho_\lambda}(\tau_A \mathbb{1}_{\{\tau_A < \tau_{B_\delta}\}})$ .

The proof of Lemma 4.2 is now complete.  $\square$

4.3. *Uphill and Downhill lemmas.* In this subsection, we shall prove (4.24) and (4.48).

PROOF OF LEMMA 4.5. Instead of proving (4.24) directly, we will first show the (more natural) estimate

$$(4.49) \quad \sum_{\sigma \in A} \mu(\sigma) \mathbb{E}_\sigma(\tau_B \mathbb{1}_{\tau_B < \tau_A}) \leq e^{-cN} \mu(A)$$

for some  $c > 0$ . To do so, we use the fact that

$$(4.50) \quad \mathbb{E}_\sigma \tau_{A \cup B} = \mathbb{E}_\sigma(\tau_A \mathbb{1}_{\tau_A < \tau_B}) + \mathbb{E}_\sigma(\tau_B \mathbb{1}_{\tau_B < \tau_A}).$$

Define the function

$$(4.51) \quad w_{A,B}(\sigma) \equiv \begin{cases} \mathbb{E}_\sigma(\tau_A \mathbb{1}_{\tau_A < \tau_B}), & \text{if } \sigma \notin A \cup B, \\ 0, & \text{else,} \end{cases}$$

$w_{A,B}$  solves the Dirichlet problem

$$(4.52) \quad Lw_{A,B}(\sigma) = h_{A,B}(\sigma), \quad \sigma \notin A \cup B,$$

$$(4.53) \quad w_{A,B}(\sigma) = 0, \quad \sigma \in A \cup B,$$

where  $L \equiv 1 - P$ . Notice that, for  $\sigma \in A$ ,

$$(4.54) \quad \mathbb{E}_\sigma(\tau_A \mathbb{1}_{\tau_A < \tau_B}) = \mathbb{P}_\sigma(\tau_A < \tau_B) - Lw_{A,B}(\sigma).$$

Next, using reversibility,

$$(4.55) \quad \sum_{\sigma} \mu(\sigma) h_{A,B}(\sigma) Lw_{A,B}(\sigma) = \sum_{\sigma} \mu(\sigma) Lh_{A,B}(\sigma) w_{A,B}(\sigma).$$

By the properties of the functions  $h_{A,B}$  and  $w_{A,B}$ , this equation reduces to

$$(4.56) \quad - \sum_{\sigma \in A} \mu(\sigma) Lw_{A,B}(\sigma) = \sum_{\sigma \notin A \cup B} \mu(\sigma) h_{A,B}(\sigma)^2.$$

Hence,

$$(4.57) \quad \begin{aligned} \sum_{\sigma \in A} \mu(\sigma) \mathbb{E}_\sigma(\tau_A \mathbb{1}_{\tau_A < \tau_B}) &= \sum_{\sigma \in A} \mu(\sigma) \mathbb{P}_\sigma(\tau_A < \tau_B) \\ &+ \sum_{\sigma \notin A \cup B} \mu(\sigma) h_{A,B}(\sigma)^2. \end{aligned}$$

Using a completely similar procedure, one shows that

$$(4.58) \quad \sum_{\sigma \in A} \mu(\sigma) \mathbb{E}_\sigma(\tau_{A \cup B}) = \sum_{\sigma \in A} \mu(\sigma) + \sum_{\sigma \notin A \cup B} \mu(\sigma) h_{A,B}(\sigma).$$

Therefore, taking into account (4.50),

$$(4.59) \quad \begin{aligned} \sum_{\sigma \in A} \mu(\sigma) \mathbb{E}_\sigma(\tau_B \mathbb{1}_{\tau_B < \tau_A}) &= \sum_{\sigma \in A} \mu(\sigma) \mathbb{P}_\sigma(\tau_B < \tau_A) \\ &+ \sum_{\sigma \notin A \cup B} \mu(\sigma) h_{A,B}(\sigma) h_{B,A}(\sigma). \end{aligned}$$

The first term on the right-hand side is exponentially small compared to  $\mu(A)$  by Lemma 2.2. The same holds true for the second term, by the same estimates that were used in the proof of Lemmas 6.1 and 6.2 in [1]. Thus (4.49) holds. By Lemma 4.4, it follows that for a slightly smaller constant  $c'$ ,  $\mathbb{E}_{\rho_\lambda}(\tau_B \mathbb{1}_{\tau_B < \tau_A}) \leq e^{-c'N}$ . Finally,  $\mathbb{E}_{\rho_\lambda} \tau_{A \cup B} \geq 1$  and so  $\mathbb{E}_{\rho_\lambda}(\tau_A \mathbb{1}_{\tau_A < \tau_B}) \geq 1 - e^{-c'N}$ , and we can deduce (4.24). This concludes the proof of the lemma.  $\square$

The microscopic harmonic function  $h(\sigma) \equiv \mathbb{P}(\tau_A < \tau_B)$  gives rise to the so-called  $h$ -transformed chain with transition probabilities

$$(4.60) \quad p_N^h(\sigma, \sigma') = h(\sigma)^{-1} p_N(\sigma, \sigma') h(\sigma').$$

This  $h$ -transformed chain lives on  $\{\sigma : h(\sigma) > 0\}$  and it is reversible with respect to  $\mu^h \equiv h^2 \mu$ . The following *Downhill lemma* holds.

LEMMA 4.8. *With the notation introduced before,*

$$(4.61) \quad \begin{aligned} \sum_{\sigma \in A} \mu(\sigma) \mathbb{E}_\sigma(\mathbb{1}_{\{\tau_{B_\delta} < \tau_A\}} \mathbb{E}_{\sigma \tau_{B_\delta}}(\tau_A \mathbb{1}_{\{\tau_A < \tau_B\}})) \\ \leq \sum_{\sigma' \in A^\delta \setminus A} \mu^h(\sigma') \mathbb{P}_{\sigma'}^h(\tau_{B_\delta} < \tau_A). \end{aligned}$$

PROOF. By reversibility,

$$(4.62) \quad \mu(\sigma) \mathbb{E}_\sigma(\mathbb{1}_{\{\tau_{B_\delta} < \tau_A\}} \mathbb{1}_{\{\sigma(\tau_{B_\delta}) = \eta\}}) = \mu(\eta) \mathbb{E}_\eta(\mathbb{1}_{\{\tau_A < \tau_{B_\delta}\}} \mathbb{1}_{\{\sigma(\tau_A) = \sigma\}}).$$

Hence,

$$(4.63) \quad \begin{aligned} \sum_{\sigma \in A} \mu(\sigma) \mathbb{E}_\sigma(\mathbb{1}_{\{\tau_{B_\delta} < \tau_A\}} \mathbb{E}_{\sigma(\tau_{B_\delta})}(\tau_A \mathbb{1}_{\{\tau_A < \tau_B\}})) \\ = \sum_{\eta \in B_\delta} \mu(\eta) \mathbb{P}_\eta(\tau_A < \tau_{B_\delta}) \mathbb{E}_\eta(\tau_A \mathbb{1}_{\{\tau_A < \tau_B\}}). \end{aligned}$$

Since the only nonzero contribution to the latter sum comes from  $\eta$  in the exterior boundary of  $A_\delta$ , we can bound it from above in terms of the  $h$ -transformed quantities as

$$(4.64) \quad \sum_{\eta \in B_\delta} \mu^h(\eta) \mathbb{P}_\eta^h(\tau_A < \tau_{B_\delta}) \mathbb{E}_\eta^h \tau_A.$$

Applying the representation formula (1.4) for hitting times for the  $h$ -transformed dynamics, we can represent the above sum as

$$(4.65) \quad \sum_{\sigma' \in A_\delta \setminus A} \mu^h(\sigma') \mathbb{P}_{\sigma'}^h(\tau_{B_\delta} < \tau_A),$$

and (4.61) follows.  $\square$

Let us go back to (4.48). Using an estimate completely analogous to Lemma 2.2, one sees that

$$(4.66) \quad \begin{aligned} \sum_{\sigma' \in A_\delta \setminus A} \mu^h(\sigma') \mathbb{P}_{\sigma'}^h(\tau_{B_\delta} < \tau_A) &\leq \sum_{\sigma' \in A_\delta \setminus A} \mu(\sigma') h(\sigma') \mathbb{P}_{\sigma'}(\tau_{B_\delta} < \tau_A) \\ &\leq \mu(A_\delta \setminus A) e^{-c_\delta N} \end{aligned}$$

for some  $c_\delta > 0$ . This allows us to deduce (4.48) from (4.61).

**Acknowledgments.** We thank Malwina Luczak for stimulating discussions on [9] and coupling methods in general. We are grateful to Martin Slowik for pointing out an error in a previous version and for suggesting the proof of Lemma 4.5.

The kind hospitality of the Technion, Haifa, the Weierstrass-Institute for Applied Analysis and Stochastics, and the Institute of Applied Mathematics at Bonn University is gratefully acknowledged.

## REFERENCES

- [1] BIANCHI, A., BOVIER, A. and IOFFE, D. (2009). Sharp asymptotics for metastability in the random field Curie–Weiss model. *Electron. J. Probab.* **14** 1541–1603. [MR2525104](#)
- [2] BOVIER, A. (2009). Metastability. In *Methods of Contemporary Mathematical Statistical Physics. Lecture Notes in Math.* **1970** 177–221. Springer, Berlin. [MR2581606](#)
- [3] BOVIER, A., DEN HOLLANDER, F. and SPITONI, C. (2010). Homogeneous nucleation for Glauber and Kawasaki dynamics in large volumes and low temperature. *Ann. Probab.* **38** 661–713.
- [4] BOVIER, A., ECKHOFF, M., GAYRARD, V. and KLEIN, M. (2001). Metastability in stochastic dynamics of disordered mean-field models. *Probab. Theory Related Fields* **119** 99–161. [MR1813041](#)
- [5] BOVIER, A., ECKHOFF, M., GAYRARD, V. and KLEIN, M. (2002). Metastability and low lying spectra in reversible Markov chains. *Comm. Math. Phys.* **228** 219–255. [MR1911735](#)
- [6] BOVIER, A., ECKHOFF, M., GAYRARD, V. and KLEIN, M. (2004). Metastability in reversible diffusion processes. I. Sharp asymptotics for capacities and exit times. *J. Eur. Math. Soc. (JEMS)* **6** 399–424. [MR2094397](#)



- [7] CASSANDRO, M., GALVES, A., OLIVIERI, E. and VARES, M. E. (1984). Metastable behavior of stochastic dynamics: A pathwise approach. *J. Stat. Phys.* **35** 603–634. [MR0749840](#)
- [8] FREIDLIN, M. I. and WENTZELL, A. D. (1984). *Random Perturbations of Dynamical Systems. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]* **260**. Springer, New York. [MR0722136](#)
- [9] LEVIN, D. A., LUCZAK, M. J. and PERES, Y. (2010). Glauber dynamics for the mean-field Ising model: Cut-off, critical power law, and metastability. *Probab. Theory Related Fields* **146** 223–265. [MR2550363](#)
- [10] MARTINELLI, F., OLIVIERI, E. and SCOPPOLA, E. (1989). Small random perturbations of finite- and infinite-dimensional dynamical systems: Unpredictability of exit times. *J. Stat. Phys.* **55** 477–504. [MR1003525](#)
- [11] MARTINELLI, F. and SCOPPOLA, E. (1988). Small random perturbations of dynamical systems: Exponential loss of memory of the initial condition. *Comm. Math. Phys.* **120** 25–69. [MR0972542](#)
- [12] MATHIEU, P. and PICCO, P. (1998). Metastability and convergence to equilibrium for the random field Curie–Weiss model. *J. Stat. Phys.* **91** 679–732. [MR1632726](#)
- [13] NUMMELIN, E. (1978). A splitting technique for Harris recurrent Markov chains. *Z. Wahrsch. Verw. Gebiete* **43** 309–318. [MR0501353](#)
- [14] OLIVIERI, E. and VARES, M. E. (2005). *Large Deviations and Metastability. Encyclopedia of Mathematics and Its Applications* **100**. Cambridge Univ. Press, Cambridge. [MR2123364](#)

A. BIANCHI  
 WEIERSTRASS-INSTITUTE FOR APPLIED ANALYSIS  
 AND STOCHASTICS  
 MOHRENSTRASSE 39  
 10117 BERLIN  
 GERMANY  
 E-MAIL: [alessandra.bianchi7@unibo.it](mailto:alessandra.bianchi7@unibo.it)

A. BOVIER  
 INSTITUT FÜR ANGEWANDTE MATHEMATIK  
 RHEINISCHE FRIEDRICH–WILHELMS-UNIVERSITÄT  
 ENDENICHER ALLEE 60  
 53115 BONN  
 GERMANY  
 E-MAIL: [bovier@uni-bonn.de](mailto:bovier@uni-bonn.de)

D. IOFFE  
 FACULTY OF INDUSTRIAL ENGINEERING  
 AND MANAGEMENT  
 TECHNION—ISRAEL INSTITUTE OF TECHNOLOGY  
 HAIFA 32000  
 ISRAEL  
 E-MAIL: [ieioffe@technion.ac.il](mailto:ieioffe@technion.ac.il)