Markov chain mixting time, card shuffling and spin systems dynamics

Weiyang Ning

A dissertation submitted in partial fulfillment of the
requirements for the degree of

Doctor of Philosophy

University of Washington

2013

Reading Committee:

Christopher Hoffman, Chair

Zhen-Qing Chen

Krzysztof Burdzy

Peter Hoff

Program Authorized to Offer Degree:
Mathematics

University of Washington

Abstract

Markov chain mixting time, card shuffling and spin systems dynamics

Weiyang Ning

Chair of the Supervisory Committee:
Professor Christopher Hoffman
Mathematics

The mixing time of a Markov chain describes how fast the Markov chain converges to its stationary distribution. In this thesis, we survey some of the knowledge and main tools available in this field by looking at examples. We focus on various models of card shuffling (random walk on the permutation group S_n) and the Swendsen-Wang dynamics of the mean-field Ising Model. We show that the Card-Cyclic to Random shuffle has mixing time of order $\Theta(n \log n)$ (joint work of Ben Morris and Yuval Peres). We also determine the order of the mixing time of the mean field Swendsen-Wang dynamics at all temperatures. In particular, at criticality, it mixes at time $\Theta(n^{\frac{1}{4}})$ (joint work of Yun Long, Asaf Nachmias, and Yuval Peres).

TABLE OF CONTENTS

ACKNOWLEDGMENTS

First of all, I would like to thank Yuval Peres, for leading me into the world of probability, and his constant support. I also want to thank Asaf Nachmias and Ben Morris, my main collaborators, for generously sharing their ideas during our delightful collaboration. I am also indebted to Chris Hoffman, for carefully reading my thesis and providing numerous valuable suggestions.

Chapter 1

INTRODUCTION

A finite irreducible and aperiodic Markov chain has an unique stationary distribution π and the distribution of the chain at time t will converge to π as $t \to \infty$. Given two probability distributions μ and ν on the state space Ω, define the total variation distance by

$$\|\mu - \nu\|_{\mathrm{TV}} = \max_{A \subset \Omega} |\mu(A) - \nu(A)| = \frac{1}{2}\|\mu - \nu\|_1.$$

Estimating the maximal total variation distance (over all starting states x) between $\mathbb{P}^t(x, \cdot)$ and π is among our primary objectives. It is therefore convenient to define

$$d(t) := \max_{x \in \Omega} \|\mathbb{P}^t(x, \cdot) - \pi\|_{TV}.$$

One can show that $d(t)$ is non-increasing in t. The mixing time is defined by

$$t_{mix}(\epsilon) := \min\{t : d(t) \leq \epsilon\}$$

and

$$t_{mix} := t_{mix}(1/4).$$

There are various methods to bound $t_{mix}(\epsilon)$ from above and below. We will explore some of the available tools as we go through examples. The main examples covered in this thesis are various card shuffling Markov chains and the Swendsen-Wang dynamics of spin systems. They are covered in chapter 2 and 3 respectively.

Chapter 2

RANDOM WALK ON THE PERMUTATION GROUP

The card shuffling Markov chains are examples of random walks on the permutation group S_n. A probability distribution μ on the permutation group describes a mechanism of shuffling cards: apply permutation σ to the deck of n cards with probability $\mu(\sigma)$. Repeatedly shuffling the deck using this mechanism is equivalent to running the random walk on S_n with increment distribution μ. If the support of μ generates all of S_n, the corresponding chain is irreducible. If $\mu(id) > 0$, the chain is aperiodic. As examples of random walk on groups, all card shuffling chains have uniform stationary distributions.

Example I: Top to Random Insertion. Let σ_t be the shuffle of taking the top card from the deck and inserting it to an uniformly position, which is known as the Top-to-Random Insertion. Let τ_{top} be the time one move after the first occasion when the original bottom card has been moved to the top of the deck. One can prove by induction that τ_{top} is independent of $\sigma_{\tau_{top}}$ and $\sigma_{\tau_{top}}$ is uniform over S_n. A **strong stationary time** for a Markov chain X_t is a randomized stopping time τ such that X_τ has distribution π, and X_τ, τ are independent. Thus τ_{top} is a strong stationary time of the Top to Random Insertion.

Lemma 2.0.1. *If τ is a strong stationary time, then*

$$d(t) \leq \max_x \mathbb{P}^x(\tau > t).$$

Observe that the distribution of τ_{top} is the same as the classical coupon collector's time. Applying the lemma, one has

$$d(n \log n + \alpha n) \leq e^{-\alpha}.$$

This gives an upper bound of the mixing time.

For a lower bound, Aldous and Diaconis [1] observed that the event

$$A_j = \{\text{the original bottom } j \text{ cards are in their original relative order}\}$$

illustrates the great difference between the uniform distribution and the distribution after $n \log n - \alpha n$ shuffles starting from the identity.

Proposition 2.0.2. *For any $\epsilon > 0$, there exists a constant α_0 such that $\alpha > \alpha_0$ implies that for all sufficiently large n,*

$$d_n(n \log n - \alpha n) \geq 1 - \epsilon.$$

In particular, there is a constant α_1 such that for all sufficiently large n,

$$t_{mix} \geq n \log n - \alpha_1 n.$$

Notice that for the Top-to-Random Insertion, combining the results in the upper and lower bound, one has

$$\lim_{\alpha\to\infty}\limsup_{n\to\infty} d(n\log n+\alpha n)=0,$$

and

$$\lim_{\alpha\to-\infty}\limsup_{n\to\infty} d(n\log n+\alpha n)=1.$$

This is an example of the cut-off phenomenon which states that the total variation distance $d(t)$ drops abruptly from near 1 to near 0 in a time period which is of lower order of the mixing time. See [23] Chapter 18.

Example II: Random Transposition. Another example of random walks on S_n that cut-off has been established is the Random Transposition. In this shuffle one picks two uniform random cards (with probability $\frac{1}{n}$ the two cards are the same) and transpose them. For a lower bound of mixing time, we look at the number of cards that have not been picked yet by the time t. Notice that these cards are the fix points of the permutation if starting from identity. Since two random cards are picked each time, by the coupon collector problem, one need $\frac{1}{2}n\log n$ shuffles to get almost all cards picked. Thus it is reasonable to believe for $t<\frac{1}{2}n\log n-\alpha n$, there are still too many fix points of the permutation.

Proposition 2.0.3. *Let $0<\epsilon<1$. For the Random Transposition chain,*

$$t_{mix}(\epsilon)\geq\frac{n-1}{2}\log\big(\frac{1-\epsilon}{6}n\big).$$

Diaconis and Shahshahani [12] proved a matching upper bound. Their argument uses representation theory. Roughly speaking, the L_2 distance between the distribution of the permutation after t shuffles and the uniform distribution is bounded from above by an expression involving all irreducible representation of S_n, which can be characterized by the Young diagrams. See [10] for details. The cut-off of the Random Transposition was also proved by Berestycki, Schramm and Zeitouni using purely probabilistic argument in [8]. They showed that a random walk on S_n whose step distribution is uniform on k-cycles has the mixing time of $(1/k)n\log n$, with threshold of width linear in n. The Random Transposition is a special case of this result. Their proof relies on a delicate coupling of Schramm [38].

An important general question about the mixing times of card shuffling problems is what happens when we move from the world of random updates, where at each step a card is chosen at random and updated, to systematic scan, when the updates are done in a more deterministic fashion. On the one hand, systematic scan is less random, so one might expect that the mixing time is larger. On the other hand, systematic scan can update n cards in n steps, whereas with random updates $n\log n$ steps are required by the coupon collector problem. Mironov [25], Saloff-Coste and Zuniga [40], and Mossel, Peres and Sinclair [28] analyzed the Cyclic-to-Random shuffle, which is a systematic scan version of the Random Transposition:

Example III: Cyclic-to-Random shuffle. At step t the card in position $t \bmod n$ is swapped with a randomly chosen card. They found that the mixing time for this chain is still on the order of $n \log n$. The upper bound in [28] follows from a strong stationary time argument and in [40], the comparison technique is used. For the lower bound of this model, in [28] the authors generalized an idea of Wilson which first appeared in [42]. The approach in [42] uses an eigenvector Φ of the Markov chain. One has

$$\mathbb{E}(\Phi(X_{t+1})|X_t) = \lambda\Phi(X_t).$$

To obtain a good lower bound, one finds a $\lambda < 1$ but $\lambda \approx 1$. Since $\lambda < 1$, in stationarity $\mathbb{E}\Phi(X) = 0$. But since $\lambda \approx 1$, it takes a long time before $\mathbb{E}\Phi(X_t) \approx 0$. If furthermore the eigenvector is smooth in the sense that $\mathbb{E}\big((\Phi(X_{t+1}) - \Phi(X_t))^2|X_t\big)$ is never large, then we can bound the variance of $\Phi(X_t)$, showing that it is with high probability confined to a small interval around its expected value. Provided that $\mathbb{E}\Phi(X_t)$ is large enough, we can reliably distinguish $\Phi(X_t)$ from $\Phi(X)$ in stationarity, which implies that the Markov chain has not yet mixed by the time t.

Example IV: Random to Random Insertion. The Random to Random Insertion is defined as follows. At each step a card is chosen uniformly at random and then moved to a uniform random position. It was shown in [13], [41] and [37] that the mixing time of this shuffle is on the order of $n \log n$. The upper bound follows from the comparison technique introduced by Diaconis and Saloff-Coste. The lower bound is obtained via analyzing the longest increasing subsequence of the permutation. Notice that the cards that have yet to be moved form a such subsequence. For the random permutation, the longest increasing subsequence is typically of length $\Theta(\sqrt{n})$. However, for $t < (\frac{1}{2} - o(1))n \log n$, the number of cards that have yet to be moved is of higher order by the coupon collector argument. Subag [37] further refined this argument and improved the lower bound to $(\frac{3}{4} - o(1))n \log n$.

2.1 Card-Cyclic to Random shuffle

Pinsky [34] invented the following model, called the Card-Cyclic to Random shuffle, which is a systematic scan version of the Random to Random Insertion: at time t move the card with the label $t \bmod n$ to a uniform random position. It is not obvious that the mixing time is greater than n: after n steps each card has been randomized, so one might expect the whole deck to be close to uniform. However, Pinsky showed that the mixing time is indeed greater than n, as the total variation distance at this time converges to 1 as n goes to infinity. We show that in fact the mixing time is on the order of $n \log n$. This is joint work with Ben Morris and Yuval Peres [27].

Remark. One can study another systematic scan version of the Random to Random Insertion where one cycles through the cards by position rather than label. Consider the shuffle where at time t, the card in position $t \bmod n$ is removed and inserted in a uniform random position. Call this the Position-Cyclic-to-Random Insertion. For this shuffle the coupon collector problem implies a lower bound for the mixing time of order $n \log n$ by the following argument (Ross Pinsky, personal communication): Note that the time-reversal of this chain is the shuffle which at time t picks a uniform card and inserts it to location t mod n. Thus, if we start with the cards in increasing order then the cards that have never been chosen for reinsertion by time t form an increasing subsequence of the permutation at time t. Since the longest increasing subsequence of a uniform permutation is $O(\sqrt{n})$ with high probability, the mixing time must be at least of order $n \log n$, since this is the number of steps required to ensure that the number of "uncollected coupons" in the coupon collector problem is $O(\sqrt{n})$. A matching upper-bound of $O(n \log n)$ follows from the work of Saloff-Coste and Zuniga. See [40, Theorem 4.8].

2.2 Lower bound of the Card-Cyclic to Random shuffle

To prove a lower bound, we show that for some small constant c after $cn \log n$ steps of shuffles, there exists a certain set of positions that tend to be occupied by a certain set of cards. This involves some analysis of an interesting ordinary differential equation. Recall that in the Card-Cyclic-to-Random shuffle, at time t we remove the card with label t mod n and then reinsert it into a uniform random location.

Define a round to be n consecutive such shuffles. Note that the Markov chain that performs a round of the Card-Cyclic-to-Random shuffle at each step is time-homogeneous with a doubly-stochastic transition matrix, irreducible and aperiodic, hence converges to the uniform stationary distribution. It follows that the Card-Cyclic-to-Random shuffle converges to uniform as well. Our main results show that the mixing time is on the order of $\log n$ rounds.

Theorem 2.2.1. *There exists c_0 such that for any $c < c_0$ and $0 < \epsilon < 1$, when n is sufficiently large, we have*

$$t_{\text{mix}}(\epsilon) \geq cn \log n.$$

Here $c_0 = \frac{1}{2+2a}$ where a is the smallest positive solution of equations $b = e^a \sin b$ and $a = e^a \cos b - 1$. Numerically $c_0 = 0.161875162....$

2.2.1 The barrier

The key idea for the lower bound is to imagine a barrier between two parts of the deck, that moves along with the cards as the shuffling is performed. If a card is inserted into the gap that the barrier occupies, we use the convention that the card is inserted on the same side of the barrier as it was in the previous step. We illustrate this with the following example. Suppose there is a deck of 8 cards with a barrier between cards 3 and 5. In the next step, card 7 is inserted between cards 3 and 5.

$$\begin{array}{ccccccccc} 2 & 1 & 3 & | & 5 & 4 & 6 & 8 & 7 \\ 2 & 1 & 3 & | & 7 & 5 & 4 & 6 & 8 \end{array}$$

Let $\{\sigma_t\}_{t=0}^{\infty}$ be a Card-Cyclic-to-Random shuffle. We think of $\sigma_t(i)$ as the position of card i at time t, where the positions range from 1 at the left to n at the right. Define the position of the barrier as the position of the card immediately to its left, and throughout the present chapter, let B_t be the position of the barrier at time t. Use the convention that $B_t = 0$ if at time t the barrier is to the left of all cards. We will call the pair process (σ_t, B_t) the *auxiliary process.*

Note that at any time $t > n$, every card has been reinserted exactly once in the previous n steps. Furthermore, if a card is reinserted to the left of the barrier then it stays there until it is reinserted again. Hence

$$B_t = \sum_{i=1}^{n} \mathbf{1}(\text{the card moved at time } t-i \text{ is inserted to the left of barrier}). \tag{2.1}$$

Since the conditional probability that the card at time t is inserted to the left of the barrier, given B_t, is $\frac{1}{n}B_t$, taking expectations in (2.1) gives

$$\mathbb{E}(B_t) = \frac{1}{n}\sum_{i=1}^{n} \mathbb{E}(B_{t-i}). \tag{2.2}$$

Define $g(t) = \mathbb{E}(\frac{1}{n}B_t)$. Then g satisfies the following *moving average* condition:

$$g(t) = \frac{1}{n}\sum_{i=1}^{n} g(t-i), \tag{2.3}$$

for $t > n$. We shall approximate $g(t)$ by $f(t/n)$, where $f : \mathbf{R} \to [0,1]$ is a continuous function satisfying 2.5. Our first lemma gives an example of such a function.

Lemma 2.2.2. *There exists $a > 0$ and $b > 2\pi$ such that $f(x) = \frac{1}{2} + \frac{1}{2}e^{-ax}\sin(bx)$ satisfies*

$$f'(x) = f(x) - f(x-1). \tag{2.4}$$

Moreover,

$$f(x) = \int_{x-1}^{x} f(s)ds, \tag{2.5}$$

for all x.

Proof. Since properties (2.4) and (2.5) are preserved under shifting and scaling, it is enough to show that they apply to $h(x) = e^{-ax}\sin(bx)$, for suitable a and b.

First, we show that for suitable choice of a and b we have $h'(x) = h(x) - h(x-1)$. By the product rule,

$$h'(x) = -ae^{-ax}\sin(bx) + be^{-ax}\cos(bx), \tag{2.6}$$

and a calculation shows that

$$h(x) - h(x-1) = (1 - e^{a}\cos b)e^{-ax}\sin(bx) + (e^{a}\sin b)e^{-ax}\cos(bx). \tag{2.7}$$

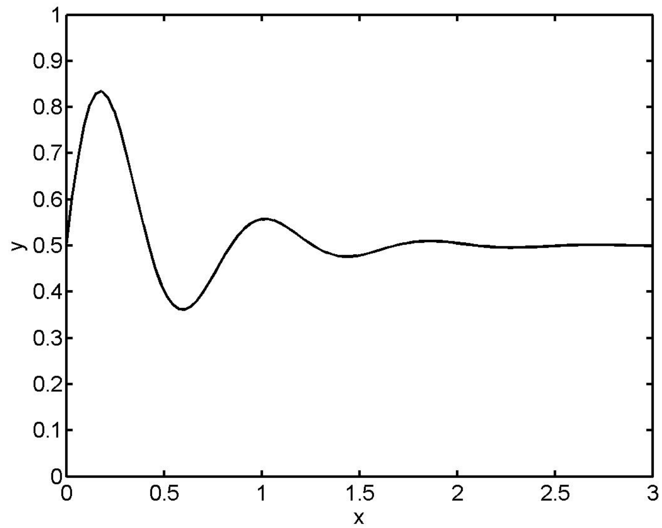

Figure 2.1: Graph of f(s)

The quantities (2.6) and (2.7) are equal if $b = e^a \sin b$ and $-a = 1 - e^a \cos b$. Solving for a in the first equation gives

$$a = \log \frac{b}{\sin b},$$

and substituting this into the second one gives

$$\log \frac{\sin b}{b} = 1 - \frac{b \cos b}{\sin b}.$$

By the intermediate value theorem, this equation has a solution with b in the interval $[2\pi + \frac{\pi}{4}, 2\pi + \frac{\pi}{2}]$, since when $b = 2\pi + \frac{\pi}{4}$ the right-hand side is smaller than the left-hand side, but when $b = 2\pi + \frac{\pi}{2}$ the right-hand side is larger. Furthermore, since $\sin b < b$ when $b > 0$, we have $a = \log \frac{b}{\sin b} > 0$. (Numerical approximation gives the solution as $b = 7.4615...$ and $a = 2.0888....$)

Next we claim that since $h'(x) = h(x) - h(x-1)$, we must have $h(x) = \int_{x-1}^{x} h(s)ds$. To see this, define $\hat{q}(x) = \int_{x-1}^{x} h(s)ds$ and note that $\hat{q}'(x) = h'(x)$. This implies that $h(x) - \hat{q}(x) = C$ for a constant C. But since $a > 0$, we have $h(x) \to 0$ as $x \to \infty$. Consequently $\hat{q}(x) \to 0$ as $x \to \infty$ by the definition of $\hat{q}$, and so $C = 0$. $\square$

Recall that $g(t) = \mathbb{E}(\frac{1}{n} B_t)$, where B_t is the position of the barrier at time t. A key part of our proof will be to show that g closely follows the continuous function f of Lemma 2.2.2. However, in order for this to be the case we must start with a permutation chosen from a certain probability distribution. It is most convenient to describe this starting permutation as being generated in the first n time steps, which we call the *startup round.* In the startup round, we begin with only a barrier. At time t, for $1 \leq t \leq n$, we put card t to the left of the barrier with probability $f(\frac{t}{n})$. The location among the already existing cards in the left (right) side of the barrier is arbitrary. We must modify the definition of g to handle the startup round. Define $g : \{1, 2, \dots\} \to \mathbf{R}$ by

$$g(t) = \begin{cases} f(\frac{t}{n}) & \text{if } 1 \le t \le n; \\ \mathbb{E}(\frac{1}{n}B_t) & \text{otherwise.} \end{cases}$$

Thus g satisfies the moving average condition, and, because of the insertion probabilities used in the startup round, g matches f for the first n steps. (That is, $g(\cdot) = f(\frac{\cdot}{n})$ on $\{1, \cdots, n\}$.) As we show below, this is enough to ensure that g is well-approximated by f for a number of rounds on the order of $\log n$.

Lemma 2.2.3. *There exists a constant $C > 0$ such that*

$$\left|g(t) - f(\tfrac{t}{n})\right| \le \tfrac{C}{2n} e^{2(t+1)/n},$$

for all $t > 0$.

Proof. First, note that if $t > n$ then

$$\begin{aligned} g(t+1) - g(t) &= \tfrac{1}{n} \sum_{i=1}^{n} g(t+1-i) - \tfrac{1}{n} \sum_{i=1}^{n} g(t-i) \\ &= \tfrac{1}{n} \Big(g(t) - g(t-n) \Big). \end{aligned}$$

Rearranging terms gives

$$g(t+1) = (1 + \tfrac{1}{n}) g(t) - \tfrac{1}{n} g(t-n). \tag{2.8}$$

Recall that $f(x) = \frac{1}{2} + \frac{1}{2} e^{-ax} \sin(bx)$ and $a > 0$. Some calculus shows that the second derivative of f is uniformly bounded on $[0, \infty)$. Hence

$$\begin{aligned} f(\tfrac{t+1}{n}) - f(\tfrac{t}{n}) &= \tfrac{1}{n} f'(\tfrac{t}{n}) + \mathrm{O}(\tfrac{1}{n^2}) \\ &= \tfrac{1}{n} \Big(f(\tfrac{t}{n}) - f(\tfrac{t}{n} - 1) \Big) + \mathrm{O}(\tfrac{1}{n^2}), \end{aligned}$$

where the first line follows from Taylor's theorem and the second line follows from Lemma 2.2.2. Rearranging terms gives

$$f(\tfrac{t+1}{n}) = (1 + \tfrac{1}{n}) f(\tfrac{t}{n}) - \tfrac{1}{n} f(\tfrac{t-n}{n}) + \mathrm{O}(\tfrac{1}{n^2}). \tag{2.9}$$

Combining (2.8) and (2.9) and using the triangle inequality gives

$$\left|g(t+1) - f(\tfrac{t+1}{n})\right| \le (1 + \tfrac{1}{n}) \left|g(t) - f(\tfrac{t}{n})\right| + \tfrac{1}{n} \left|g(t-n) - f(\tfrac{t-n}{n})\right| + \tfrac{C}{n^2}, \tag{2.10}$$

for a universal constant C. We claim that for all t we have

$$\left|g(t) - f(\tfrac{t}{n})\right| \le \tfrac{C}{n^2} \sum_{i=0}^{t} \left(1 + \tfrac{2}{n}\right)^i. \tag{2.11}$$

We prove this by induction. For the base case, note that $g(t) = f(\frac{t}{n})$ for $t = 1, \dots, n$. Now if we suppose that (2.11) holds for $1, \dots, t$, then the two absolute values on the right-hand side of (2.10) can be bounded by $\frac{C}{n^2} \sum_{i=0}^{t} (1 + \frac{2}{n})^i$. Hence

$$\begin{aligned} \left| g(t+1) - f(\tfrac{t+1}{n}) \right| &\leq (1 + \tfrac{2}{n}) \Big[\tfrac{C}{n^2} \sum_{i=0}^{t} (1 + \tfrac{2}{n})^i \Big] + \tfrac{C}{n^2} \\ &= \tfrac{C}{n^2} \sum_{i=0}^{t+1} (1 + \tfrac{2}{n})^i, \end{aligned}$$

which verifies (2.11) for $t + 1$. To finish the proof of the lemma, note that

$$\begin{aligned} \tfrac{C}{n^2} \sum_{i=0}^{t} (1 + \tfrac{2}{n})^i &= \tfrac{C}{n^2} \frac{(1 + \frac{2}{n})^{t+1} - 1}{\frac{2}{n}} \\ &\leq \tfrac{C}{2n} e^{2(t+1)/n}. \end{aligned}$$

□

2.2.2 Deviation estimates

In the previous subsection we proved that the expected barrier location is well-approximated by a continuous function. In the present subsection we show that the barrier stays reasonably close to its expectation with high probability when the number of rounds is on the order of $\log n$.

Define a *configuration* as a pair (σ, b), where σ is a permutation and b is a barrier location. (Thus the state space of the auxiliary process is the set of all configurations.) We define the *insertion distance* between two configurations as the minimum number of cards we would need to remove and re-insert to get from one configuration to the other. For example the insertion distance between the two configurations below is 2. (Move cards 4 and 7.)

$$\begin{array}{cccccccccc} 2 & 1 & 4 & 3 & | & 5 & 6 & 8 & 7 \\ 2 & 1 & 3 & 7 & | & 5 & 4 & 6 & 8 \end{array}$$

Lemma 2.2.4. *Let (σ_t^1, B_t^1) and (σ_t^2, B_t^2) be auxiliary processes, and define $\hat{\sigma}_t^i = (\sigma_t^i, B_t^i)$ for $i = 1, 2$. Let d be the insertion distance between $\hat{\sigma}_0^1$ and $\hat{\sigma}_0^2$. Then*

$$|\mathbb{E}B_t^1 - \mathbb{E}B_t^2| \leq d \left(1 + \tfrac{1}{n}\right)^t.$$

Proof. There is a natural coupling of σ_t^1 and σ_t^2 that we call *label coupling.* In label coupling, at time t we choose a label X uniformly at random. If $X = t \bmod n$, then we move card $t \bmod n$ to the leftmost position in both processes. Otherwise, we insert card $t \bmod n$ to the right of the card with label X in both processes.

Suppose that $A = \{a_1, \dots, a_d\}$ is a minimal set of cards that can be moved to get from $\hat{\sigma}_0^1$ to $\hat{\sigma}_0^2$. Note that under the label coupling, only in the case when we move a card not in

A can the insertion distance be increased. In such moves, if the card is put to the right of a card in A, the insertion distance increases by 1 and otherwise it stays the same. Thus the expected insertion distance after one step is at most

$$(d+1)\frac{d}{n} + d\,\frac{n-d}{n} = d\left(1+\frac{1}{n}\right).$$

Iterating this argument shows that the expected insertion distance after t steps is at most $d\left(1+\frac{1}{n}\right)^t$. The lemma follows from this, since the barrier can move by at most one position with each re-insertion. □

We are now ready to state the main lemma of this subsection.

Lemma 2.2.5. *Let (σ_t, B_t) be an auxiliary process. Fix $c > 0$ and suppose T satisfies $n < T \le cn \log n$. Then for any $x > 0$ we have*

$$\mathbb{P}(\left|\tfrac{1}{n}B_T - g(T)\right| \ge x) \le 2\exp\left(-x^2 n^{1-2c}\right).$$

Proof. Fix T with $n < T \le cn \log n$. Since $g(T) = \frac{1}{n}\mathbb{E}(B_T)$, it is enough to show that for any $x > 0$ we have

$$\mathbb{P}(|B_T - \mathbb{E}(B_T)| \ge x) \le 2\exp\left(-x^2 n^{-(1+2c)}\right).$$

Let $\mathcal{F}_t$ be the sigma-field generated by the process up to time t, and consider the Doob martingale

$$M_t := \mathbb{E}(B_T \,|\, \mathcal{F}_t).$$

Applying Lemma 2.2.4 to the case of two configurations that differ by one insertion gives

$$|M_t - M_{t-1}| \le \left(1+\tfrac{1}{n}\right)^{T-t},$$

for t with $1 \le t \le T$. Thus the Azuma-Hoeffding bound gives

$$\begin{aligned} \mathbb{P}(|B_T - \mathbb{E}(B_T)| \ge x) &= \mathbb{P}(|M_T - \mathbb{E}(M_T)| \ge x) \\ &\le 2\exp\left(\frac{-x^2}{2\sum_{t=1}^T b_t^2}\right), \end{aligned} \tag{2.12}$$

where $b_t = \left(1+\frac{1}{n}\right)^{T-t}$. Let $r = \left(1+\frac{1}{n}\right)^2$. The sum in (2.12) can be written as

$$\begin{aligned} \sum_{i=0}^{T-1} r^i &= \frac{r^T - 1}{r-1} \\ &\le \tfrac{n}{2} r^T, \end{aligned} \tag{2.13}$$

since $r - 1 = \frac{2}{n} + \frac{1}{n^2} \ge \frac{2}{n}$. Since $T < cn \log n$, the quantity (2.13) is at most

$$\tfrac{n}{2}\left(1+\tfrac{1}{n}\right)^{2cn\log n} \le \tfrac{1}{2} n^{1+2c}.$$

Substituting this into (2.12) yields the lemma. □

2.2.3 Proof of the Lower bound

Recall that $f(s) = \frac{1}{2} + \frac{1}{2}e^{-as}\sin(bs)$, for $a = 2.0888...$ and $b = 7.4615....$ The rough idea for the lower bound is as follows. Note that if c is sufficiently small and $s < c\log n$, then the fluctuation of $f(s)$ between s and $s+1$ is of higher order than $n^{-1/2}$. Thus in the corresponding round of the Card-Cyclic-to-Random shuffle, there will be an interval of cards where the probability of inserting to the left of the barrier is detectably high. Before we give the proof, we recall Hoeffding's bounds in [17].

Theorem 2.2.6. *Let $X_1, \ldots, X_k$ be samples from a population of 0's and 1's, and let $p = \mathbb{E}(X_1)$ be the proportion of 1's in the population. Then for $\alpha > 0$,*

$$\mathbb{P}\left(\sum_{i=1}^{k} X_i - kp \geq \alpha\right) \leq e^{-2\alpha^2/k}. \tag{2.14}$$

The bound (2.14) applies whether the sampling is done with or without replacement.

Proof. Proof of Theorem 2.2.1 Let $c > 0$ be small enough so that

$$c < \frac{1}{2+2a}. \tag{2.15}$$

Fix T with $n < T < cn\log n$ and let $x = T/n$. Suppose that $\sin(bx) \leq 0$. The case $\sin(bx) > 0$ is similar. Since $b > 2\pi$, there exist x_1, x_2 with $x-1 < x_1 < x_2 < x$, such that

$$bx_1 = 2\pi k + \pi/4, \quad \text{and}$$

$$bx_2 = 2\pi k + 3\pi/4,$$

for an integer k. Note that for $s \in [x_1, x_2]$ we have

$$\begin{aligned} f(s) &\geq \tfrac{1}{2} + \beta e^{-as} \\ &\geq \tfrac{1}{2} + \beta n^{-ac}, \end{aligned} \tag{2.16}$$

where $\beta = \frac{1}{2}\sin(\pi/4)$. The second inequality holds because $x \leq c\log n$.

Let A be the event that $|\frac{1}{n}B_t - f(t/n)| \leq \frac{\beta}{4}n^{-ac}$ for all t with $T-n < t \leq T$. Note that since $T < cn\log n$, substituting T into the upper bound of Lemma 2.2.3 implies that if $t \leq T$ then $|g(t) - f(t/n)| < Bn^{2c-1}$, for a constant $B > 0$. Since $2c-1 < -ac$ by (2.15), for sufficiently large n we have

$$Bn^{2c-1} < \tfrac{\beta}{8}n^{-ac},$$

and hence $|g(t) - f(t/n)| < \frac{\beta}{8}n^{-ac}$ for $t \leq T$. Hence

$$\begin{aligned} \mathbb{P}(A^c) &\leq \mathbb{P}\big(\big|\tfrac{1}{n}B_t - g(t)\big| > \tfrac{\beta}{8}n^{-ac} \text{ for some } t \text{ with } T-n < t \leq T\big) \\ &\leq 2n\exp\Big(-\Big[\tfrac{\beta}{8}n^{-ac}\Big]^2 n^{1-2c}\Big) \\ &= 2n\exp\Big(-\tfrac{\beta^2}{64}n^{1-2c(a+1)}\Big), \end{aligned} \tag{2.17}$$

where the second inequality follows from Lemma 2.2.5 and a union bound. Since $1-2c(a+1)>0$ by (2.15), the quantity (2.17), and hence $\mathbb{P}(A^c)$, converges to 0 as $n\to\infty$.

Let $I=\{t \bmod n : nx_1<t<nx_2\}$ and $m=|I|$. Since $x_2-x_1=\pi/2b$, there is a constant $\lambda>0$ such that $m\ge\lambda n$ for sufficiently large n. Let N be the number of cards in I (that is, cards whose label is in I) placed to the left of the barrier between times nx_1 and nx_2. Then N is also the number of cards from I to the left of the barrier at time T. By (2.16), on the event A the insertion probabilities $\frac{B_t}{n}$ are bounded below by $\frac{1}{2}+\frac{3\beta}{4}n^{-ac}$ for t with $nx_1<t<nx_2$. Hence the conditional distribution of N given A stochastically dominates the Binomial$(m,\frac{1}{2}+\frac{3\beta}{4}n^{-ac})$ distribution. Thus Hoeffding's bounds give

$$\begin{aligned}\mathbb{P}\left(N<\tfrac{m}{2}+\tfrac{\beta}{2}mn^{-ac}\,|\,A\right) &\le \exp\left(\frac{-2\left(\frac{\beta}{4}mn^{-ac}\right)^2}{m}\right)\\ &\le \exp\left(-\tfrac{\beta^2}{8}\lambda n^{1-2ac}\right), \qquad (2.18)\end{aligned}$$

where the second line follows from the fact that $m\ge\lambda n$. Since $1-2ac>0$ by (2.15), the quantity (2.18) converges to 0 as $n\to\infty$.

Now let Y be the number of cards in I having position less than $\frac{n}{2}+\frac{\beta}{4}n^{1-ac}$ at time T. Since $f\left(\frac{T}{n}\right)\le\frac{1}{2}$, we have $B_T\le\frac{n}{2}+\frac{\beta}{4}n^{1-ac}$ on the event A, and hence

$$\mathbb{P}(Y\le\tfrac{m}{2}+\tfrac{\beta}{2}mn^{-ac}) \le \mathbb{P}(N\le\tfrac{m}{2}+\tfrac{\beta}{2}mn^{-ac}\,|\,A)+\mathbb{P}(A^c), \qquad (2.19)$$

which converges to 0 as $n\to\infty$.

To complete the proof, let Y_u be the number of cards in I whose position is less than $\frac{n}{2}+\frac{\beta}{4}n^{1-ac}$ in a uniform random permutation.

Hoeffding's bounds imply that

$$\begin{aligned}\mathbb{P}(Y_u>\tfrac{m}{2}+\tfrac{\beta}{2}mn^{-ac}) &\le \exp\left(\frac{-2\left(\frac{\beta}{4}mn^{-ac}\right)^2}{m}\right)\\ &\le \exp\left(-\frac{\beta^2}{8}\lambda n^{1-2ac}\right), \qquad (2.20)\end{aligned}$$

for sufficiently large n. Since $1-2ac>0$, the quantity (2.20) converges to 0 as $n\to\infty$. Combining this with (2.19), we conclude that $t_{\mathrm{mix}}(\epsilon)\ge cn\log n$ for large enough n. □

2.3 Upper bound of the Card-Cyclic to Random shuffle

To analyze upper bound, we prove the following theorem.

Theorem 2.3.1. *For any $\epsilon>0$ and $n\ge4$, we have*

$$t_{\mathrm{mix}}(\epsilon)\le C(n\log n-2n\log\epsilon),$$

where

$$C = \frac{1}{\log 2 - \log(e-1)} = 6.58664655...$$

We use the path coupling method of Bubley and Dyer [6]. Suppose the state space Ω of a Markov Chain is a metric space with given distance ρ, the transportation metric between two distributions on Ω is defined by

$$\rho_K(\mu,\nu) := \inf\{\mathbb{E}(\rho(X,Y)) : (X,Y) \text{ is a coupling of } \mu \text{ an } \nu\}.$$

The following theorem is from [6]. See also [23] Chapter 14.

Theorem 2.3.2. *Suppose that there exists $\alpha > 0$ such that for each edge $\{x,y\}$ in G there exists a coupling (X_1,Y_1) of the distributions $\mathbb{P}(x,\cdot)$ and $\mathbb{P}(y,\cdot)$ such that*

$$\mathbb{E}_{x,y}\rho(X_1,Y_1) \leq \rho(x,y)e^{-\alpha}.$$

Then

$$t_{mix}(\epsilon) \leq \frac{-\log\epsilon + \log(diam(G))}{\alpha}.$$

The analysis of Card-Cyclic to Random shuffle is based on the transition of one round of moves, which is n consecutive shuffles. The underlying graph is S_n with an edge existing between two permutations that differ by an adjacent transposition. The path metric on G is defined by

$$\rho(x,y) = \min\{\text{length of } \eta : \eta \text{ is a path from } x \text{ to } y\}.$$

Define

$$\text{diam}(G) = \sup_{x,y} \rho(x,y).$$

For a permutation x, define σ_t^x to be the Card-Cyclic-to-Random shuffle starting at x. Our mixing time upper bound follows from the following lemma.

Lemma 2.3.3. *If permutations x and y differ by an adjacent transposition and $n \geq 4$, there is a coupling of σ_n^x and σ_n^y such that*

$$\mathbb{E}\rho(\sigma_n^x, \sigma_n^y) \leq e^{-\alpha},$$

where $\alpha = 2(\log 2 - \log(e-1))$.

Proof. There is another natural coupling of two Card-Cyclic-to-Random processes besides label coupling; we call this second coupling *position coupling.* In position coupling, the card is inserted into the same locations in both processes. Now assume that for some $i < j$, the permutation x can be obtained from y by transposing the cards with label i and j, as shown below. In the diagram, the kth X in the top row represents the same card as the kth X in the bottom row.

$$\begin{array}{lcccccccc} x: & X & X & X & i & j & X & X & X \\ y: & X & X & X & j & i & X & X & X \end{array}$$

The coupling strategy is divided into 3 stages, corresponding to t in $\{1, \cdots, i-1\}$, $\{i, \cdots, j-1\}$, and $\{j, \cdots, n\}$ respectively.

Stage 1: moving cards $1, \ldots, i-1$. In this stage use position coupling. As is shown by diagram 1 below, at the end of this stage we still have two permutations that differ only by a transposition of i and j. However, there may have been some cards inserted between cards i and j; we represent these cards with a's.

$$\begin{array}{rcccccccc} \sigma_{i-1}^x : & X & X & i & a & a & j & X & X \\ \sigma_{i-1}^y : & X & X & j & a & a & i & X & X \end{array}$$

diagram 1

Stage 2: moving cards $i, \ldots, j-1$. In this stage we use label coupling. At the end of this stage, some cards might have been inserted into the group of a's. We denote such cards with α''s. In addition, some cards might have been inserted between card j and the first X to the right of the card j. We represent them with b's. Diagram 2 shows a typical pair of permutations after stage 2.

$$\begin{array}{rccccccccccc} \sigma_{j-1}^x : & X & X & a & \alpha' & a & \alpha' & j & b & b & X & X \\ \sigma_{j-1}^y : & X & X & j & b & b & a & \alpha' & a & \alpha' & X & X \end{array}$$

diagram 2

Stage 3: moving cards $j, \ldots, n$. Here we use label coupling again. Cards inserted into the group of a's and α''s are represented with a_{**}'s, and cards inserted into the group of b's are represented with β's. See diagram 3 below. Notice that the a's, α''s and a_{**}'s maintain the same relative order in σ_n^x and as in σ^n, and similarly for the b's and β''s.

$$\begin{array}{rcccccccccccccc} \sigma_n^x : & X & X & a & \alpha' & a_{**} & a & a_{**} & \alpha' & b & \beta' & b & X & X \\ \sigma_n^y : & X & X & b & \beta' & b & a & \alpha' & a_{**} & a & a_{**} & \alpha' & X & X \end{array}$$

diagram 3

For $t \leq n$, let A_t be the number of a's, α''s and a_{**}'s, and let B_t be the number of b's and β''s, after card t has been moved. Note that

$$\rho(\sigma_n^x, \sigma_n^y) \leq A_n B_n.$$

Thus we are left to estimate $\mathbb{E}(A_n B_n)$.

Initially we have $A_0 = B_0 = 0$. Recall that in the first stage we use position coupling. For $t \leq i-1$ we have $B_t = 0$ and A_t satisfies

$$\mathbb{P}(A_{t+1} = A_t | A_t) = \frac{n - A_t - 1}{n},$$

and

$$\mathbb{P}(A_{t+1} = A_t + 1 | A_t) = \frac{A_t + 1}{n}.$$

This implies

$$\mathbb{E}(A_{t+1}+1|A_t) = (A_t+1)\left(1+\frac{1}{n}\right). \tag{2.21}$$

Hence

$$\mathbb{E}A_{i-1} = \left(1+\frac{1}{n}\right)^{i-1} - 1. \tag{2.22}$$

Recall that we use label coupling in the second stage. For $i \le t \le j-1$, we have the following transition rule:

$$\mathbb{P}(A_{t+1} = A_t, B_{t+1} = B_t|A_t, B_t) = \frac{n - A_t - B_t - 1}{n},$$

and

$$\mathbb{P}(A_{t+1} = A_t + 1, B_{t+1} = B_t|A_t, B_t) = \frac{A_t}{n},$$

and

$$\mathbb{P}(A_{t+1} = A_t, B_{t+1} = B_t + 1|A_t, B_t) = \frac{B_t + 1}{n}.$$

This implies

$$\mathbb{E}(A_{t+1}(B_{t+1}+1)|A_t, B_t) = A_t(B_t+1)\left(1+\frac{2}{n}\right).$$

Recall that $B_t = 0$ for all $t \le i-1$. Thus we have

$$\mathbb{E}A_{j-1}(B_{j-1}+1) = \mathbb{E}A_{i-1}\left(1+\frac{2}{n}\right)^{j-i}. \tag{2.23}$$

Note that for t with $i \le t < j$ we have

$$\mathbb{E}(A_{t+1}|A_t) = A_t\left(1+\frac{1}{n}\right). \tag{2.24}$$

Thus $\mathbb{E}A_{j-1} = \mathbb{E}A_{i-1}(1+\frac{1}{n})^{j-i}$. Combining this with (2.23) and (2.22) gives

$$\mathbb{E}A_{j-1}B_{j-1} = \left(\left(1+\frac{1}{n}\right)^{i-1} - 1\right)\left(\left(1+\frac{2}{n}\right)^{j-i} - \left(1+\frac{1}{n}\right)^{j-i}\right). \tag{2.25}$$

For $j \le t \le n$ we have the following transition probabilities:

$$\mathbb{P}(A_{t+1} = A_t, B_{t+1} = B_t|A_t, B_t) = \frac{n - A_t - B_t}{n};$$

$$\mathbb{P}(A_{t+1} = A_t + 1, B_{t+1} = B_t|A_t, B_t) = \frac{A_t}{n};$$

$$\mathbb{P}(A_{t+1} = A_t, B_{t+1} = B_t + 1|A_t, B_t) = \frac{B_t}{n}.$$

This implies

$$\mathbb{E}(A_{t+1}B_{t+1}|A_t, B_t) = A_tB_t\left(1+\frac{2}{n}\right).$$

Using (2.25), we obtain

$$\mathbb{E}A_nB_n = \left(\left(1+\frac{1}{n}\right)^{i-1} - 1\right)\left[\left(1+\frac{2}{n}\right)^{j-i} - \left(1+\frac{1}{n}\right)^{j-i}\right]\left(1+\frac{2}{n}\right)^{n-j+1}.$$

Since $1+\frac{2}{n} \leq \left(1+\frac{1}{n}\right)^2$, the expression in square brackets is at most $\left(1+\frac{1}{n}\right)^{j-i}\left(\left(1+\frac{1}{n}\right)^{i-j} - 1\right)$. Thus if we define β and γ so that $i = \beta n$ and $j = \gamma n$, calculation yields that

$$\mathbb{E}A_nB_n \leq (e^{\beta} - 1)e^{\gamma-\beta}(e^{\gamma-\beta} - 1)e^{2(1-\gamma)},$$

if $0 \leq \beta \leq \log 2$, and

$$\mathbb{E}A_nB_n \leq (e^{\beta} - 1)e^{\gamma-\beta}(e^{\gamma-\beta} - 1)e^{2(1-\gamma)}\left(1+\frac{2}{n}\right),$$

if $\log 2 < \beta \leq 1$. The former expression is maximized, for γ and β with $0 \leq \beta \leq \gamma \leq 1$, by $\left(\frac{e-1}{2}\right)^2$. The maximum occurs when $\gamma = 1$ and $\beta = \log\frac{2e}{e+1}$. Notice that $\log\frac{2e}{e+1} < \log 2$. Therefore, if $\alpha = 2(\log 2 - \log(e-1))$, then

$$\mathbb{E}(A_nB_n) \leq e^{-\alpha},$$

for all $0 \leq \beta \leq \gamma \leq 1$ and $n \geq 4$. which completes the proof. $\square$

Proof. Proof of Theorem 2.3.1 We apply Theorem 3.2.1 to a round of the Card-Cyclic-to-Random shuffle. Since the diameter of S_n with respect to adjacent transpositions is $\frac{n(n-1)}{2} < n^2$, substituting the α of Lemma 2.3.3 into Theorem 3.2.1 gives

$$t_{\text{mix}}(\epsilon) \leq \frac{1}{\log 2 - \log(e-1)}(\log n - \frac{1}{2}\log\epsilon).$$

$\square$

Remark. Theorem 2.2.1 and Theorem 2.3.1 together establish that the Card-Cyclic-to-Random shuffle has a pre-cutoff in total variation distance. It is an interesting open problem to determine if cutoff occurs in this shuffle. For reference on cutoff and pre-cutoff phenomenon, see [23, Chapter 18].

Chapter 3

SPIN SYSTEMS AND THE SWENDSEN-WANG DYNAMICS

3.1 Introduction

Given a graph $G = (V, E)$, a spin system is a probability distribution on $\Omega = \{1, -1\}^V$. For a configuration σ, the value $\sigma(v)$ is called the spin at v. The physical interpretation is that magnets, each having one of the two possible orientations represented by $+1$ and -1, are placed on the vertices of the graph; a configuration specifies the orientations of these magnets. The (nearest-neighbor) Ising Model is the most widely studied spin system. In this system, the energy of a configuration σ is defined to be

$$H(\sigma) = -\sum_{v,w \in V, v \sim w} \sigma(v)\sigma(w).$$

The energy increases with the number of pairs whose spins disagree.

The Gibbs distribution (without the external field) corresponding to the energy H is the probability distribution μ on Ω defined by

$$\mu(\sigma) = \frac{1}{Z(\beta)} e^{-\beta H(\sigma)}.$$

The partition function $Z(\beta)$ is the normalizing constant such that μ is a probability distribution. The parameter β determines the importance of the energy function. The case $\beta > 0$ (resp. $\beta < 0$) is called ferromagnetic (resp. anti-ferromagnetic) Ising model.

The Potts model is defined in similar fashion except that we assign q possible colors to each vertex where $q \geq 2$ is an integer. The state space is $\Omega = \{1, 2, 3, \cdots, q\}^V$ where a color is an element of $\{1, 2, 3, \cdots, q\}$. The energy function is

$$H(\sigma) = -\sum_{v,w \in V, v \sim w} 1\{\sigma(v) = \sigma(w)\}.$$

The Ising model is the case when $q = 2$.

Introduced in 1963, the Glauber dynamics are commonly used to sample from the Potts (Ising) model on graphs. It updates the configuration by picking a vertex w uniformly at random from V and then generating a new configuration according to μ conditioned on the set of configuration agreeing with σ on vertices different from w. At low temperatures, the mixing time of the Glauber dynamics becomes very large in some cases (sometimes exponential in the size of the graph), making it computationally hard to sample from the equilibrium measure. Global Markov chains, which allow moves such as cluster flipping, yield much faster mixing sometimes and those are the algorithms of choice when practitioners actually sample. The Swendsen-Wang (SW) algorithm and its variations are frequently used in practice.

Let

$$E(\sigma) = \{\{u, v\} \in E : \sigma(u) = \sigma(v)\}.$$

The Swendsen-Wang algorithm performs the following two steps.

1. Given a Potts configuration, delete each edge in $E(\sigma_t)$ independently with probability $1 - p = e^{-\beta}$. This gives some open edges in E, which induce some clusters.

2. For each cluster obtained in the previous step, assign a random color from $\{1, 2, 3, \cdots, q\}$ to it. This gives a new configuration σ_{t+1}.

It is easy to show using Fortuin and Kasteleyn's Random Cluster model [15] (see Edwards and Sokal [14] for this derivation) that the Potts model measure is invariant under the SW dynamics. Moreover, the SW dynamics is clearly an aperiodic and irreducible Markov chain. Hence from any starting configuration σ_0, the law of σ_t obtained after t updates, converges in distribution to the stationary Potts model.

It worths to mention that the Swendsen-Wang dynamics do not mix rapidly in all cases. Gore and Jerrum [16] discovered that for any $q > 2$, on the complete graph K_n there are temperatures where the SW dynamics has mixing time of order at least $\exp(\Omega(\sqrt{n}))$. Borgs, Chayes, Frieze, Kim, Tetali, Vigoda and Vu [5] proved a similar lower bound on the mixing time of the SW algorithm at the critical temperatures, on the d-dimensional lattice torus for any $d \geq 2$ and q sufficiently large.

One natural question is how does the SW algorithm perform when $q = 2$, i.e., the Ising model. The first positive result in this direction is due to Cooper, Dyer, Frieze and Rue [9]. They proved that the SW algorithm on the complete graph on n vertices has mixing time at most $O(\sqrt{n})$ for all non-critical temperatures. We improve their result and show that the mixing time of the SW algorithm on the complete graph of n vertices is of order $O(\log n)$ at the non-critical temperatures (in fact, we will show it is of order $\Theta(1)$ in high temperatures) and is of order $\Theta(n^{1/4})$ at the critical temperature. Thus, the chain is rapidly mixing at all temperatures and critical slowing down indeed occurs, but on a much smaller scale. Heuristic arguments for the exponent $1/4$ at criticality were found earlier by physicists, see [39] and [35]. This is joint work with Yun Long, Asaf Nachmias and Yuval Peres [21].

Theorem 3.1.1. *Consider the SW dynamics on the complete graph K_n of n vertices, with percolation parameter $p = \frac{c}{n}$, where c is a constant independent of n. Then,*

(i) If $c > 2$ then $T_{\mathrm{mix}} = \Theta(\log n)$.

(ii) If $c = 2$ then $T_{\mathrm{mix}} = \Theta(n^{1/4})$.

(iii) If $c < 2$ then $T_{\text{mix}} = \Theta(1)$.

It is instructive to compare these results with the mixing time of the Glauber dynamics for the critical Ising model on K_n. In [22], the authors show that this mixing time is $\Theta(n^{3/2})$. Since in the SW dynamics we update n vertices in each step, the number of vertex updates up to the mixing time is $\Theta(n^{5/4})$, that is, it performs faster by $n^{1/4}$ than the Glauber dynamics.

For the Potts model on any tree with n vertices, we prove the following theorem.

Theorem 3.1.2. *The mixing time of the Swendsen-Wang process for the q-state ferromagnetic Potts model on any tree with n vertices is $O(\log n)$.*

The main bulk of the research were devoted to study how to prove the upper bound as it stated in Theorem 3.1.1. Coupling is the main technique applied here. A coupling of two distribution μ and ν is two random variables X and Y defined on a common space such that the marginal distribution of X and Y are μ and ν respectively.

Any coupling of Markov chains with transition matrix P can be modified so that the two chains stay together at all times after their first simultaneous visits to a single state. More precisely, if $X_s = Y_s$, then $X_t = Y_t$ for all $t \geq s$. For such modified coupling, we have the following theorem.

Theorem 3.1.3. *Let (X_t, Y_t) be a coupling satisfying the previous requirement for which $X_0 = x$ and $Y_0 = y$. Let τ be the first time that the two chains meet. Then*

$$\|P^t(x,\cdot) - P^t(y,\cdot)\|_{TV} \leq \mathbb{P}(\tau > t).$$

Further,

Lemma 3.1.4. *Assume that for every two states $x, y \in \Omega$ there exists a coupling $\{X_t, Y_t\}$ with $X_0 = x$, $Y_0 = y$ such that $X_t = Y_t$ for some $t \leq L$ with a constant probability $\epsilon > 0$. Then, $T_{\text{mix}}(X, 1-\epsilon) \leq L$.*

3.1.1 Random graph estimates

Due to the percolative nature of the dynamics we require several estimates about percolation on the complete graph. Recall that $G(m, p)$ is the random graph obtained from the complete graph on m vertices by retaining each edge with probability p and erasing it with probability $1-p$, independently of all other edges. There is rich literature about this model (see [18] and the references within). In particular, there is an interesting phase transition when $p = \frac{1}{m}$. In this work we required several estimates on the size of the largest connected component $\mathcal{C}_1$ for various values of p. In the following we highlight some of these estimates which are interesting for the random graph community.

Fix $\theta > 1$. It is a well known result of Pittel [32] that for $G(m,p)$ with $p = \frac{\theta}{m}$ we have that $\frac{|\mathcal{C}_1|-\beta m}{\sqrt{m}}$ converges in probability to a normal distribution, where β is the unique positive solution of the equation

$$1 - e^{-\theta x} = x.$$

This does not imply, however, the moderate deviation bound on $|\mathcal{C}_1 - \beta m|$

$$\mathbb{P}(\big||\mathcal{C}_1| - \beta m\big| > A\sqrt{m}) \le Ce^{-cA^2}\,, \tag{3.1}$$

for any $A > 0$, which we prove here, see Lemma 3.3.4.

The study of the random graph "inside" the phase transition was initiated by Bollobás [4] where it is shown that if $p = \frac{1+\epsilon(m)}{m}$ when $\epsilon(m)$ is a positive sequence satisfying $\epsilon(m) \ge m^{-1/3}\log m$, then with high probability $|\mathcal{C}_1| = (2+o(1))\epsilon m$. The logarithmic corrections were removed by Luczak [24] and this statement holds whenever $\epsilon(m) \gg m^{-1/3}$. A stronger result was recently proved by Pittel and Wormald [33]. They show that in this regime of p, the distribution of $\frac{|\mathcal{C}_1|-2\epsilon m}{\sqrt{m/\epsilon}}$ converges to a normal distribution (this is a corollary of Theorem 6 of [33], but in fact the authors prove much more than this statement).

Surprisingly however, the above results do not give good estimates on $\mathbb{E}|\mathcal{C}_1|$ and on moderate and large deviations of $|\mathcal{C}_1| - 2\epsilon m$. These are crucial in our analysis of the Swendsen-Wang chain since these determine the moments of the increments of the chain. In Section 3.3 we prove several such estimates, for instance

$$\mathbb{E}|\mathcal{C}_1| = 2\epsilon m + O(\epsilon^{-2} + \epsilon^2 m).$$

See the more accurate inequality in Theorem 3.3.8. Another interesting estimate is a bound on the deviation of $|\mathcal{C}_1|$,

$$\mathbf{P}\Big(\Big||\mathcal{C}_1| - 2\epsilon m\Big| > A\sqrt{\frac{m}{\epsilon}}\Big) \le Ce^{-cA^2}\,,$$

for any A satisfying $0 \le A \le \sqrt{\epsilon^3 m}$. See Theorem 3.3.9.

Combining these new results of the Erdos-Renyi random graph, one can deduce that the magnetization chain X_t, which is the sum of all spins, is a random walk with some certain increment distribution. Using some classical ideas of the overshoot of random walks, a coupling of two copies of magnetization chains is provided. A rough guide of the whole proof is given in the next section.

3.2 Outline of the proof of Theorem 3.1.1

Due to the length of the proof of Theorem 3.1.1, we provide here a "road-map" of whole argument. The reader is advised to follow this outline to get the general idea of the proof and go to the main contents for further details whenever needed. Let $\{\sigma_t\}_{t=0}^{\infty}$ be the SW

Markov chain. Consider the chain X_t defined by

$$X_t = \Big| \sum_v \sigma_t(v) \Big| . \tag{3.2}$$

Since the underlying graph is complete, $\{X_t\}$ is a Markov chain with state space $\{0, \ldots, n\}$. Given X_0, the random variable X_1 is determined by two independently drawn random graphs $G(\frac{n+X_0}{2}, \frac{c}{n})$ and $G(\frac{n-X_0}{2}, \frac{c}{n})$. If we denote by $\{\mathcal{C}_j^+\}_{j\geq 1}$ and $\{\mathcal{C}_j^-\}_{j\geq 1}$ the connected components of the corresponding two random graphs, then X_1 is distributed as

$$\Big| \sum_{j\geq 1} \epsilon_j |\mathcal{C}_j^+| + \sum_{j\geq 1} \epsilon_j' |\mathcal{C}_j^-| \Big| , \tag{3.3}$$

where $\{\epsilon_j\}$ and $\{\epsilon_j'\}$ are i.i.d. random variables taking 1 with probability 1/2 and -1 otherwise. This is the reason that the moments and large deviation estimates of random graph component sizes are useful in our approach.

Frequently, to obtain upper bounds on the mixing time of the SW chain we will obtain a bound on the mixing time of the chain X_t and then use the following lemma, which appears in [9] and is based on the path coupling idea of Bubley and Dyer [6].

Lemma 3.2.1. *Suppose $\{\sigma_t\}$ and $\{\sigma_t'\}$ are two SW chains such that $X_0 = X_0'$. There exists a coupling of the two chains such that with probability at least $\frac{1}{2}$ the two chains meet after $O(\log n)$ steps.*

3.2.1 Outline of the Proof of Theorem 3.1.1 (i)

By Lemmas 3.1.4 and 3.2.1 it suffices to show that we can couple two copies of the magnetization chain X_t and X_t' such that they meet in $O(\log n)$ with probability $\Omega(1)$ which is uniform over all initial values x_0 and x_0'. It turns out that the stationary distribution is concentrated in a window of length $\sqrt{n}$ around $\gamma_0 n$ for some $\gamma_0 = \gamma_0(c) \in [0, 1]$. In fact, the one step evolution of X_t essentially contracts the second moment of $|X_t - \gamma_0 n|$. That is, we have

$$\mathbb{E}(X_1 - \gamma_0 n)^2 \leq \delta(x_0 n - \gamma_0 n)^2 + Bn , \tag{3.4}$$

for some constants $\delta \in (0, 1)$ and large B. See Theorem 3.4.2. It follows quickly that there exists an interval of values $I = [\gamma_0 n - A\sqrt{n}, \gamma_0 n + A\sqrt{n}]$ for some large constant A such that for any initial value x_0 we have that $X_t \in I$ with probability $\Omega(1)$ whenever $t = \Theta(\log n)$.

Once the two chains are both in the interval I one can show that they can be coupled to meet in the next step with probability $\Omega(1)$. This is the content of Theorem 3.4.5. The main idea of that argument is that the random graph $G(n, \frac{c}{n})$ has $\Theta(n)$ isolated vertices with high probability and that the difference of the sums of the spins of the two chains *before* we assign spins to the isolated vertices is $O(\sqrt{n})$. Thus, we one can couple the two chains to correct the $O(\sqrt{n})$ error by assignment of those isolated vertices. This follows from the classical local central limit theorem for the simple random walk.

3.2.2 Outline of the Proof of Theorem 3.1.1 (iii)

Since we need to prove an $O(1)$ upper bound we cannot use Lemma 3.2.1 here. However, the study of X_t's evolution will still be useful. As in the supercritical case, the stationary measure is concentrated in a window of width $\Theta(\sqrt{n})$, but this time around 0 and mixing occurs much faster. We will show, as before, that we have contraction, that is,

$$\mathbb{E}\big(X_1^2 \,|\, X_0\big) \le \delta X_0^2 + Bn \tag{3.5}$$

for some constants $\delta \in (0,1)$ and large B and for all $x_0 \in [0,1]$. Moreover, if $0 \le X_0 \le \frac{1}{c} - \frac{1}{2}$, we have

$$\mathbb{E} X_1^2 \le Bn\,, \tag{3.6}$$

see Theorem 3.5.2. The first inequality implies that X_t will be in the window $[0, (\frac{1}{c} - \frac{1}{2})n]$ in $O(1)$ steps with probability $\Omega(1)$. The second inequality implies that from this window X_t jumps into $[0, A\sqrt{n}]$ with high probability in just one more step. This gives the mixing time upper bound on the chain X_t.

To go further and obtain a mixing time of the SW chain one needs to consider the following two-dimensional chain. For a starting configuration σ_0, let G_1 denote the vertices with positive spin and G_2 be its complement. Let (Y_t, Z_t) be a two-dimensional Markov chain, where Y_t records the number of vertices with positive spin in G_1 and Z_t records the number vertices with positive spin in G_2. By symmetry, the probability of the SW chain of being at σ at time t is the same for all σ which have the same two-dimensional chain value. Consequently, the total variation distance of σ_t from stationarity is the same as the total variation distance of the two-dimensional chain from its stationary distribution. By Lemma 3.1.4, it suffices to provide a coupling of such two-dimensional chains so that they meet in $O(1)$ steps with probability $\Omega(1)$. By our previous argument, $Y_t + Z_t$ will be in the window $[\frac{n}{2} - A\sqrt{n}, \frac{n}{2} + A\sqrt{n}]$ within $O(1)$ steps. One can show that once inside this window, such a coupling does exist. See Proposition 3.5.3. The idea is similar to the proof of part (i) by considering the isolated vertices in the two random graphs.

3.2.3 Outline of the Proof of Theorem 3.1.1 (ii)

We again use Lemma 3.2.1 and Lemma 3.1.4 to bound the mixing time of the magnetization chain. However, to simplify our calculations, we will consider a slight modification to the magnetization chain X_t. Instead of choosing a random spin for each component after the percolation step, we assign a positive spin to the *largest* component and random spins for all other components. Let X'_t be the sum of spins at time t (notice that we do *not* take absolute values here), that is,

$$X'_{t+1} \stackrel{d}{=} \max\{|\mathcal{C}_1^+(t)|, |\mathcal{C}_1^-(t)|\} + \epsilon \min\{|\mathcal{C}_1^+(t)|, |\mathcal{C}_1^-(t)|\} + \sum_{j\ge 2} \epsilon_j |\mathcal{C}_j^+(t)| + \sum_{j \ge 2} \epsilon'_j |\mathcal{C}_j^-(t)|\,, \tag{3.7}$$

where as usual $\epsilon, \{\epsilon_j\}$ and $\{\epsilon'_j\}$ are independent mean zero $\pi_{(m)}$ signs. This chain has state space $\{-n, \ldots, n\}$ and its absolute value has the same distribution as our original chain X_t. As a consequence, any upper bound on the mixing time of X'_t implies the same upper bound on the mixing time of X_t. For convenience, we will now denote this modified chain by X_t. Let X_t and Y_t be two such chains such that X_t start from an arbitrary location X_0 and Y_t starts from the stationary distribution. We will show that we can couple X_t and Y_t so that they meet in $O(n^{1/4})$ steps with probability $\Omega(1)$. It will become evident that it suffices to restrict the attention to $X_0 \in [0, n]$. We will divide this into two subcases:

(i) $X_0 \in [n^{3/4}, n]$,

(ii) $X_0 \in [0, n^{3/4}]$,

and consider them separately, let us begin with case (i). In this case the coupling strategy is as follows.

Consider the first *crossing time* of X_t and Y_t, that is, the first time t such that $\text{sign}(X_t - Y_t) \neq \text{sign}(X_{t-1} - Y_{t-1})$. We will show that this is likely to occur only when the two chains take values $\Theta(n^{3/4})$ and, more importantly, the distance between the chains one step before the crossing time is of order $n^{5/8}$. This is the content of Theorem 3.6.5. The fact that one time step before the crossing time is not a stopping time is problematic and requires an *overshoot* estimate stating that the two chains are not likely to cross each other from distance larger than $O(n^{5/8})$. For random walks, these kind of estimates are classical (see for instance [20]). The key estimate here is Theorem 3.6.8.

Next we show that when the chains take values $\Theta(n^{3/4})$ they satisfying a local central limit theorem in scale $n^{5/8}$. In particular, the chain has probability $\Omega(n^{-5/8})$ to move to any point x in an interval of size $\Theta(n^{5/8})$ around the starting point. We use the standard characteristic function technique to show this, see Lemma 3.6.19. Now we are ready to conclude the proof in this case since we know that a step before the crossing time the chains have already been at distance $O(n^{5/8})$ from each other, so the local CLT provides a way to couple them in a few additional steps after the crossing time. See the proof of Theorem 3.6.1.

Let us consider now case (ii) in which $X_0 \in [0, n^{3/4}]$. To handle this case define $I = [-An^{2/3}, An^{2/3}]$ and proceed in two steps.

1. With high probability X_t will visit the interval I by time $O(n^{1/4})$. This is proved in Theorem 3.6.24 and is based on the fact that the drift of the chain $|X_t|$ in this regime is approximately $-n^{1/2}$ (this is a small negative drift).

2. Once the chain is inside I, it will be pushed above $\Omega(n^{3/4})$ within $O(n^{1/4})$ steps. See Theorem 3.6.23.

Thus, with these two claims we see that in at most $O(n^{1/4})$ steps the chain is pushed into the $n^{3/4}$ regime and we may use the theorems of case (i) to conclude. Let us briefly expand on the proofs of (1) and (2).

The proof of (1) relies on the fact that the chain has a negative drift of magnitude $\Omega(n^{1/2})$ as long as $X_t \notin I$. This follows rather easily from the random graph estimate Theorem 3.3.15. Note, however, that Theorem 3.3.15 estimates the expected size of the cluster discovered in time $\delta\epsilon m$ in the exploration process for some small $\delta > 0$ and *not* of the largest cluster $\mathcal{C}_1$. We denote the former cluster by $\mathcal{C}_{\delta\epsilon m}$ and remark that it has high probability of being the largest. However, we were unable to prove the estimate of Theorem 3.3.15 for $\mathcal{C}_1$ but only for $\mathcal{C}_{\delta\epsilon m}$. This is the reason we need to consider yet another slight modification of the magnetization chain: instead of giving a plus sign to $\mathcal{C}_1$ and drawing random signs for the rest of the clusters, we give the plus sign to $\mathcal{C}_{\delta\epsilon m}$ and the rest receive random signs. From this point on the proof of (1) is rather straightforward.

For the proof of (2) one has to show that when X_t is in I, even though the drift is negative there is still enough noise to eventually push X_t to the $n^{3/4}$ regime. We were unable to pursue this strategy since it involves very delicate random graph estimates we were unable to obtain. Instead we use the following coupling idea. Since the stationary distribution normalized by $n^{3/4}$ has a weak limit with positive density at 0, the expected number of visits to I by the stationary chain before time T is $\Theta\big(\frac{n^{2/3}}{n^{3/4}}T\big)$. In Lemma 3.6.25 we show that when $T = \Theta(n^{1/4})$ the actual number of visits to I is positive with high probability. Next, to show that X_t is pushed upwards we start a stationary chain Z_t and wait until it enters I. We then couple X_t and Z_t such that they meet inside I and from that point they stay together. The only technical issue with this strategy is how to perform the coupling of Z_t and X_t inside I. This will follow, as before, from a uniform lower bound stating that for any $x, x_0 \in I$ we have

$$\mathbb{P}(X_1 = x \mid X_0 = x_0) \geq cn^{-2/3} .$$

This estimate is done inside the scaling window of the random graph phase transition and so the proofs are different from the previous ones and require some combinatorial estimates. See Lemmas 3.6.26 and Lemma 3.6.27.

3.3 Random graph estimates

In this section we prove some facts about random graphs which will be used in the proof. These lemmas might also be of sperate interests in random graph theory. Recall that $G(m, p)$ is obtained from the complete graph on m vertices by retaining independently each edge with probability p and deleting it with probability $1 - p$. We denote by $\mathcal{C}_j$ the j-th largest component of $G(m, p)$.

3.3.1 The exploration process

We recall an exploration process, due to Karp and Martin-Löf (see [19] and [26]), in which vertices will be either *active, explored* or *neutral.* After the completion of step $t \in \{0, 1, \ldots, m\}$ we will have precisely t explored vertices and the number of the active and neutral vertices is denoted by A_t and N_t respectively. Fix an ordering of the vertices $\{v_1, \ldots, v_m\}$. In step $t = 0$ of the process, we declare vertex v_1 active and all other vertices neutral. Thus $A_0 = 1$ and $N_0 = m - 1$. In step $t \in \{1, \ldots, m\}$, if $A_{t-1} > 0$, then let w_t be the first active vertex; if $A_{t-1} = 0$, let w_t be the first neutral vertex. Denote by η_t the number of neutral neighbors of w_t in $G(m, p)$, and change the status of these vertices to active. Then, set w_t itself explored.

Denote by $\mathcal{F}_t$ the σ-algebra generated by $\{\eta_1, \ldots, \eta_t\}$. Observe that given $\mathcal{F}_{t-1}$ the random variable η_t is distributed as $\mathrm{Bin}(N_{t-1} - \mathbf{1}_{\{A_{t-1}=0\}}, p)$ and we have the recursions

$$N_t = N_{t-1} - \eta_t - \mathbf{1}_{\{A_{t-1}=0\}}\,, \qquad t \leq m\,, \tag{3.8}$$

and

$$A_t = \begin{cases} A_{t-1} + \eta_t - 1, & A_{t-1} > 0 \\ \eta_t, & A_{t-1} = 0\,, \qquad t \leq m\,. \end{cases} \tag{3.9}$$

As every vertex is either neutral, active or explored,

$$N_t = m - t - A_t\,, \qquad t \leq m\,. \tag{3.10}$$

At each time $j \leq m$ in which $A_j = 0$, we have finished exploring a connected component. Hence the random variable Z_t defined by

$$Z_t = \sum_{j=1}^{t-1} \mathbf{1}_{\{A_j=0\}}\,,$$

counts the number of components completely explored by the process before time t. Define the process $\{Y_t\}$ by $Y_0 = 1$ and

$$Y_t = Y_{t-1} + \eta_t - 1\,.$$

By (3.9) we have that $Y_t = A_t - Z_t$, i.e. Y_t counts the number of active vertices at step t minus the number of components completely explored before step t.

Lemma 3.3.1. *For any t we have*

$$Y_t \stackrel{d}{\leq} \mathrm{Bin}(m - 1, 1 - (1 - p)^t) + 1 - t, \tag{3.11}$$

and

$$Y_t \stackrel{d}{\geq} \mathrm{Bin}(m - t - 1, 1 - (1 - p)^t) + 1 - t. \tag{3.12}$$

Proof. For each vertex v at each step of the process we examine precisely one of its edges emanating from it unless the vertex is active or explored at this step. Thus, all the vertices for which the process discovered an open edge emanating from them between time 1 and t are active, except for at least t of them which are explored. The probability of a vertex having no open edges explored from it between time 1 and t is precisely $(1-p)^t$. This shows (3.11).

The reason this bound is not precise is that it is possible that a neutral vertex turns to be active because there were no more active vertices at this step of the exploration process. This, however, can only happen at most t times between time 1 and t and this gives the lower bound (3.12). $\square$

At each step we marked as explored precisely one vertex. Hence, the component of v_1 has size $\min\{t \geq 1 : A_t = 0\}$. Moreover, let $t_1 < t_2 \ldots$ be the times at which $A_{t_j} = 0$; then $(t_1, t_2 - t_1, t_3 - t_2, \ldots)$ are the sizes of the components. Observe that $Z_t = Z_{t_j} + 1$ for all $t \in \{t_j + 1, \ldots, t_{j+1}\}$. Thus $Y_{t_{j+1}} = Y_{t_j} - 1$ and if $t \in \{t_j + 1, \ldots, t_{j+1} - 1\}$ then $A_t > 0$, and thus $Y_{t_{j+1}} < Y_t$. By induction we conclude that $A_t = 0$ if and only if $Y_t < Y_s$ for all $s < t$. In other words $A_t = 0$ if and only if $\{Y_t\}$ has hit a new record minimum at time t. By induction we also observe that $Y_{t_j} = -(j-1)$ and that for $t \in \{t_j + 1, \ldots t_{j+1}\}$ we have $Z_t = j$. Also, by our previous discussion for $t \in \{t_j + 1, \ldots t_{j+1}\}$ we have $\min_{s \leq t-1} Y_s = Y_{t_j} = -(j-1)$, hence by induction we deduce that $Z_t = -\min_{s \leq t-1} Y_s + 1$. Consequently,

$$A_t = Y_t - \min_{s \leq t-1} Y_s + 1 \,. \tag{3.13}$$

Lemma 3.3.2. *For all $p \leq \frac{2}{m}$ there exists a constant $c > 0$ such that for any integer $t > 0$,*

$$\mathbf{P}\Big(N_t \leq m - 5t\Big) \leq e^{-ct} \,.$$

Where we recall N_t is the number of neutral points in exploration process at time t.

The proof of Lemma 3.3.2 can be found in Lemma 3 of [29].

3.3.2 Random graph lemmas for non-critical cases

Let $G(m,p)$ be the random graph where $p = \frac{\theta}{m}$.

Lemma 3.3.3. *Suppose $\theta < 1$ is a constant. Then we have*

$$\mathbb{E}(\sum_{j\geq 1} |\mathcal{C}_j|^2) \leq \frac{\mathcal{M}}{1-\theta}. \tag{3.14}$$

Proof. Observe that $\sum_{j\geq 1} |\mathcal{C}_j|^2 = \sum_v |\mathcal{C}(v)|$ since in the right hand side each component $\mathcal{C}(v)$ is counted precisely $|\mathcal{C}(v)|$ times. Hence

$$\mathbb{E}(\sum_{j\geq 1} |\mathcal{C}_j|^2) = \mathbb{E} \sum_v |\mathcal{C}(v)| = mE|\mathcal{C}(v)|. \tag{3.15}$$

In the exploration process, we can couple Y_t with a process W_t with i.i.d. increment of $\mathrm{Bin}(m, \theta/m) - 1$ and $W_0 = 1$ such that $W_t \geq Y_t$. Thus, the hitting time of 0 for Y_t which equals to $|\mathcal{C}(v)|$ is bounded from above by the hitting time of 0 for W_t. For W_t, we have $\mathbb{E}\tau = 1/(1-\theta)$ by Wald's Lemma. This concludes the proof. $\square$

For $\theta > 1$ let $\beta = \beta(\theta)$ be the unique positive solution of the equation

$$1 - e^{-\theta x} = x. \tag{3.16}$$

In [32] it was proved that $\frac{|\mathcal{C}_1| - \beta m}{\sqrt{m}}$ converges in distribution to a normal distribution. We were unable to deduce from that result moderate deviation estimates, and we provide them in the following lemma.

Lemma 3.3.4. *There exists constants $c = c(\theta) > 0$ and universal constant C such that for any $A > 0$ we have*

$$\mathbb{P}\big(\big||\mathcal{C}_1| - \beta m\big| \geq A\sqrt{m}\big) \leq Ce^{-cA^2}. \tag{3.17}$$

Proof. Assume $A \leq \sqrt{m}$ otherwise this probability is 0. Let $\xi = \xi(\theta) > 0$ be a large constant that we will determine later. We will show that for some $c > 0$

$$\mathbb{P}(Y_{\beta m + A\sqrt{m}} \geq -cA\sqrt{m}) \leq e^{-cA^2}, \tag{3.18}$$

and that

$$\mathbb{P}\Big(\bigcup_{cA\sqrt{m} \leq t \leq \beta m - \xi A\sqrt{m}} Y_t < 0\Big) \leq Ce^{-cA^2}. \tag{3.19}$$

If these two events do not occur, then there exists a component of size in $[\beta m - (\xi + c)A\sqrt{m}, \beta m + A\sqrt{m}]$. The remaining graph is a subcritical random graph and it is a classical result that the probability that it contains a component of size $\Theta(m)$ decays exponentially in m, and this will conclude the proof.

The proof of (3.18) is based on the stochastic upper bound of Y_t in (3.11). Plugging in $t = \beta m + A\sqrt{m}$ and using the fact that $1 - x \geq e^{-x-x^2}$ for small enough x we get

$$\begin{aligned}
\mathbb{P}(Y_{\beta m + A\sqrt{m}} \geq -cA\sqrt{m}) &\leq \mathbb{P}\Big(\mathrm{Bin}(m, 1 - (1 - \frac{\theta}{m})^{\beta m + A\sqrt{m}}) \geq \beta m + (1-c)A\sqrt{m}\Big) \\
&\leq \mathbb{P}\Big(\mathrm{Bin}(m, 1 - e^{-\theta\beta - A\theta m^{-1/2} - A\theta^2 m^{-3/2}}) \geq \beta m + (1-c)A\sqrt{m}\Big).
\end{aligned}$$

A quick calculation using the fact that $1 - e^{-\theta\beta} = \beta$ gives that the expected value of this binomial random variable is at most

$$\beta m + A\theta e^{-\theta\beta}\sqrt{m} + O(1).$$

Since $\theta e^{-\theta\beta} < 1$ it follows that we can choose c so small so that this expectation is less than $\beta m + (1-2c)A\sqrt{m}$, and then Azuma-Hoeffding inequality (see for instance Theorem 7.2.1 of [2]) gives that

$$\mathbb{P}(Y_{\beta m + cA\sqrt{m}} \geq -A\sqrt{m}) \leq e^{-cA^2}.$$

We now turn to prove (3.19). For this we will divide $[cA\sqrt{m}, \beta m - \xi A\sqrt{m}]$ into two subintervals $[cA\sqrt{m}, \delta\beta m]$ and $[\delta\beta m, \beta m - \xi A\sqrt{m}]$ where $\delta > 0$ is a small constant that will be chosen later. For convenience write $\alpha = \frac{t}{m}$. For any $t \in [cA\sqrt{m}, \delta\beta m]$ we have by (3.12) and the fact that $1 - x \le e^{-x}$ for all $x \ge 0$ that

$$\begin{aligned}\mathbb{P}(Y_t < 0) &\le \mathbb{P}\Big(\mathrm{Bin}\big((1-\alpha)m, 1 - (1 - \frac{\theta}{m})^{\alpha m}\big) \le \alpha m\Big) \\ &\le \mathbb{P}\Big(\mathrm{Bin}\big((1-\alpha)m, 1 - e^{-\theta\alpha}\big) \le \alpha m\Big).\end{aligned}$$

Since $1 - e^{-x} \ge x - x^2$ for all $x \ge 0$ we deduce that the expectation of the last binomial is at least

$$(1-\alpha)(\theta\alpha - \theta^2\alpha^2) > \alpha,$$

since $\theta > 1$ when $\alpha = t/m \le \delta$ and $\delta = \delta(\theta) > 0$ is chosen small enough. By a standard large deviation estimate (see for instance, Corollary A.1.14 of [2]), we have that

$$\mathbb{P}(Y_t < 0) \le e^{-c\alpha m},$$

for some $c = c(\theta)$ and all $t \in [cA\sqrt{m}, \delta\beta m]$. It follows from the union bound that

$$\mathbb{P}\Big(\bigcup_{cA\sqrt{m} \le t \le \delta\beta m} Y_t < 0\Big) = O(e^{-cA\sqrt{m}}). \tag{3.20}$$

For the interval $[\delta\beta m, \beta m - \xi A\sqrt{m}]$ we will use the process $\widetilde{Y}_t$ which approximates Y_t introduced by Bollobas and Riordan [7]. We write

$$D_t = \mathbb{E}(\eta_t - 1 | F_{t-1}),$$

and define

$$\Delta_t = \eta_t - 1 - D_t.$$

Let $y_t = m - t - m(1-p)^t$ and define the approximation process by

$$\widetilde{Y}_t = y_t + \sum_{i=1}^{t} (1-p)^{t-i}\Delta_i. \tag{3.21}$$

In [7] Lemma 3 it is proved that for any $p > 0$ and any $1 \le t \le m$ we have

$$|Y_t - \widetilde{Y}_t| \le ptZ_t. \tag{3.22}$$

Put $\tau = \min\{t \ge \delta\beta m, A_t = 0\}$. We have

$$\begin{aligned}\mathbb{P}(\tau < \beta m - \xi A\sqrt{m}) &\le \mathbb{P}(|Y_\tau - \widetilde{Y}_\tau| \ge \theta A\sqrt{m}) \\ &+ \mathbb{P}(|Y_\tau - \widetilde{Y}_\tau| < \theta A\sqrt{m}, \tau < \beta m - \xi A\sqrt{m}).\end{aligned} \tag{3.23}$$

By (3.22) the first term has the upper bound

$$\mathbb{P}(|Y_\tau - \widetilde{Y}_\tau| \ge \theta A\sqrt{m}) \le \mathbb{P}(Z_\tau \ge A\sqrt{m}) = O(e^{-cA\sqrt{m}}),$$

since $Z_\tau \geq A\sqrt{m}$ implies that there exists at least one time t in $[A\sqrt{m}, \delta\beta m]$ such that $Y_t < 0$. The bound follows immediately from our estimate in (3.20).

To bound the second term of (3.23) observe that on $[\delta\beta m, \beta m - \xi A\sqrt{m}]$, the minimum of y_t is attained at the right end of the interval with value $(1-\theta e^{-\theta\beta})\xi A\sqrt{m}(1+o(1))$. Thus if we choose $\xi = \xi(\theta)$ large enough such that $(1-\theta e^{-\theta\beta})\xi > \theta$ and write $c = (1-\theta e^{-\theta\beta})\xi - \theta$, we have

$$\mathbb{P}(|Y_\tau - \widetilde{Y}_\tau| < \theta A\sqrt{m}, \tau < \beta m - \xi A\sqrt{m}) \leq \mathbb{P}\Big(\sum_{i=1}^{\tau}(1-p)^{\tau-i}\Delta_i < -cA\sqrt{m}\Big), \tag{3.24}$$

since $Y_\tau \leq 0$ by definition. Notice that τ is at most m, thus it suffices to bound from above

$$\mathbb{P}\Big(\max_{1\leq t\leq m}(-\sum_{i=1}^{t}(1-p)^{t-i}\Delta_i) > cA\sqrt{m}\Big).$$

Notice that $(1-p)^t = \Theta(1)$, hence it is equivalent to bound

$$\mathbb{P}\Big(\max_{1\leq t\leq m}(-\sum_{i=1}^{t}(1-p)^{-i}\Delta_i) > cA\sqrt{m}\Big).$$

Let $a > 0$ be a small number (eventually we will take $a = \Theta(m^{-1/2})$). Direct computation and the fact that $1+x > e^{x-x^2}$ for negative x when $|x|$ is small enough and some Taylor expansion yield

$$\begin{aligned}\mathbb{E}(e^{-a(1-p)^{-i}\Delta_i}|F_{i-1}) &= (1+p(e^{-a(1-p)^{-i}}-1))^{N_{i-1}-\mathbf{1}_{\{A_{i-1}=0\}}}e^{a(1-p)^{-i}p(N_{i-1}-\mathbf{1}_{\{A_{i-1}=0\}})}\\ &\geq e^{\frac{a^2p}{3}(N_{i-1}-\mathbf{1}_{\{A_{i-1}=0\}})(1-p)^{-2i}} \geq 1,\end{aligned} \tag{3.25}$$

when a is small enough. Thus we conclude $e^{-a\sum_{i=1}^{t}(1-p)^i\Delta_i}$ is a submartingale. By Doob's maximal inequality (see [11]) we have

$$\mathbb{E}\big(\max_{1\leq t\leq m} e^{-a\sum_{i=1}^{t}(1-p)^i\Delta_i}\big)^2 \leq 4\mathbb{E}e^{-2a\sum_{i=1}^{m}(1-p)^i\Delta_i}.$$

On the other hand, the fact that $1+x\leq e^x$ for all x yields

$$\mathbb{E}(e^{-a(1-p)^{-i}\Delta_i}|F_{i-1}) \leq e^{(\frac{a^2}{2}+O(a^3))p(N_{i-1}-\mathbf{1}_{\{A_{i-1}=0\}})(1-p)^{-2i}}. \tag{3.26}$$

Since $N_t \leq m$ we get

$$\mathbb{E}e^{-2a\sum_{i=1}^{m}(1-p)^i\Delta_i} \leq 4e^{Cma^2},$$

where $C = C(\theta)$. By Markov's inequality, we have

$$\mathbb{P}\big(\max_{1\leq t\leq m}(-\sum_{i=1}^{t}(1-p)^{-i}\Delta_i) > cA\sqrt{m}\big) \leq e^{Cma^2-2acA\sqrt{m}}.$$

Choosing $a = \frac{cA\sqrt{m}}{Cm}$ to minimize the right hand side, we conclude

$$\mathbb{P}\big(\max_{1\leq t\leq m}(-\sum_{i=1}^{t}(1-p)^{-i}\Delta_i) > \theta A\sqrt{m}\big) \leq 4e^{-cA^2},$$

for some $c = c(\theta)$ which is a continuous function of θ, and this concludes the proof. □

Corollary 3.3.5. *Suppose $\theta \in [a,b]$ where $a > 1$. Then, there exists a constant $C = C(a,b)$ such that for $G(m, \frac{\theta}{m})$ we have*

$$\Big|\mathbb{E}|\mathcal{C}_1| - \beta(\theta)m\Big| \leq C\sqrt{m}, \tag{3.27}$$

Proof. It follows immediately by integrating Lemma 3.3.4. □

Corollary 3.3.6. *Suppose $\theta \in [a,b]$ where $a > 1$. There exists a constant $C = C(a,b)$ such that for $G(m, \frac{\theta}{m})$ we have*

$$\mathbb{E}(\sum_{j\geq 1} |\mathcal{C}_j|^2) \leq (\mathbb{E}|\mathcal{C}_1|)^2 + Cm. \tag{3.28}$$

Proof. Notice that

$$\mathbb{E}\sum_{j\geq 1} |\mathcal{C}_j|^2 - (\mathbb{E}|\mathcal{C}_1|)^2 = \Big(\mathbb{E}|\mathcal{C}_1|^2 - (\mathbb{E}|\mathcal{C}_1|)^2\Big) + \mathbb{E}\sum_{j\geq 2} |\mathcal{C}_j|^2. \tag{3.29}$$

We have that

$$\mathbb{E}|\mathcal{C}_1|^2 - (\mathbb{E}|\mathcal{C}_1|)^2 = \mathbb{E}(|\mathcal{C}_1| - \mathbb{E}|\mathcal{C}_1|)^2 \leq \mathbb{E}(|\mathcal{C}_1| - \beta m)^2.$$

By integrating Lemma 3.3.4 we get

$$\mathbb{E}|\mathcal{C}_1|^2 - (\mathbb{E}|\mathcal{C}_1|)^2 \leq Cm.$$

For supercritical random graph $G(m, \frac{\theta}{m})$, it is a classical result that $|\mathcal{C}_1| \in ((\beta - \epsilon)m, (\beta + \epsilon)m)$ with probability at least $1 - e^{-c_\epsilon m}$ for fixed ϵ. Conditioned on this and the vertex set of $\mathcal{C}_1$, the other components are distributed as $G(m - |\mathcal{C}_1|, \frac{\theta}{m})$ (which is subcritical) restricted to the event that it does not contain any component larger than $|\mathcal{C}_1|$. This event happens with probability at most e^{-cm}. Thus we obtain

$$\mathbb{E}(\sum_{j\geq 2} |\mathcal{C}_j|^2) \leq (1 + o(1))\frac{m}{1-(1-\beta)\theta},$$

□

Lemma 3.3.7. *Let $M = \sum_{v\in V} \mathbf{1}_{\{v \text{ is isolated}\}}$ be the number of isolated vertices in $G(m, \theta/m)$ where $\theta > 0$ is a constant. There exists a constant $C > 0$ such that*

$$\mathbb{P}(M \geq Cm) = 1 - O(\frac{1}{m}).$$

Proof. We have

$$\mathbb{E}M = \sum_{v\in V} \mathbb{P}(v \text{ is isolated.}) = \mathcal{M}(1 - \frac{\theta}{\mathcal{M}})^{\mathcal{M}-1},$$

and

$$\begin{aligned} \mathbb{E}M^2 &= \sum_{v\in V} \mathbb{P}(v \text{ is isolated.}) + \sum_{\text{v,w}\in\text{V}} \mathbb{P}(\text{v, w are both isolated.}) \\ &= \mathcal{M}(1 - \frac{\theta}{\mathcal{M}})^{\mathcal{M}-1} + \mathcal{M}(\mathcal{M}-1)(1 - \frac{\theta}{\mathcal{M}})^{2\mathcal{M}-3}. \end{aligned}$$

Thus, we obtain

$$\mathbb{E}(M-\mathbb{E}M)^2 = O(m).$$

By Markov's inequality,

$$\mathbb{P}(M \le \tfrac{1}{2}\mathbb{E}M) \le \mathbb{P}\Big((M-\mathbb{E}M)^2 \ge \frac{1}{4}(\mathbb{E}M)^2\Big) \le \frac{\mathbb{E}(M-\mathbb{E}M)^2}{\frac{1}{4}(\mathbb{E}M)^2} = O(\frac{1}{\mathcal{M}}).$$

Since $\mathbb{E}M = \Theta(m)$, we finished the proof. □

3.3.3 Random graph lemmas for the near-critical case

In [33], Pittel and Wormald study the near-critical random graph $G(m,p)$ where $p = \frac{1+\epsilon}{m}$ with $\epsilon = o(1)$ but $\epsilon^3 m \to \infty$. A direct corollary of Theorem 6 of [33] shows that in this regime $\frac{|\mathcal{C}_1|-2\epsilon m}{\sqrt{m/\epsilon}}$ converges in distribution to a normal random variable (see also [7] for a recent simple proof of this fact), and that a local central limit theorem holds. Unfortunately, one cannot deduce from that precise bounds on the average size of $|\mathcal{C}_1|$ and moderate deviations estimates on $|\mathcal{C}_1| - 2\epsilon m$. The following two theorems give these estimates.

Theorem 3.3.8. *Consider $G(m,p)$ with $p = \frac{1+\epsilon}{m}$ where $\epsilon = o(1)$ and there exists a large constant $A > 0$ such that $\epsilon^3 m \ge A \log m$. Then we have that*

$$\mathbb{E}|\mathcal{C}_1| \le 2\epsilon m - \frac{8}{3}\epsilon^2 m + O(\epsilon^3 m),$$

and there exists a constant $C > 0$ such that

$$\mathbb{E}|\mathcal{C}_1| \ge 2\epsilon m - C(\epsilon^{-2} + \epsilon^2 m)\,.$$

Theorem 3.3.9. *Consider $G(m,p)$ with $p = \frac{1+\epsilon}{m}$ where $\epsilon^3 m \ge 1$. Then there exists some $c > 0$ such that*

$$\mathbf{P}\Big(\Big||\mathcal{C}_1| - 2\epsilon m\Big| > A\sqrt{\frac{m}{\epsilon}}\Big) = O\big(e^{-cA^2}\big),$$

for any A satisfying $2 \le A \le \sqrt{\epsilon^3 m}/10$.

Corollary 3.3.10. *Consider $G(m,p)$ with $p = \frac{1+\epsilon}{m}$ where $\epsilon^3 m \ge 1$, then*

$$\mathbb{E}\Big||\mathcal{C}_1| - 2\epsilon m\Big|^k \le C\Big(\frac{m}{\epsilon}\Big)^{k/2}.$$

Theorem 3.3.11. *For any large constant A and small $\delta > 0$ there exists a constant $q_1(A,\delta) > 0$ such that the following hold. Consider $G(m,p)$ with $p = \frac{1+\epsilon}{m}$ where $\epsilon \in [A^{-1}m^{-1/4}, Am^{-1/4}]$, then*

$$\mathbb{P}\big(|\mathcal{C}_1| \in [2\epsilon m - \delta m^{5/8},\, 2\epsilon m + \delta m^{5/8}]\big) \ge q_1 > 0\,.$$

Theorem 3.3.11 is a direct corollary of Theorem 6 of [33] which provides a central limit theorem for the giant component. Next we provide some moment estimates of component sizes in the subcritical and supercritical regime.

Theorem 3.3.12. *Consider $G(m,p)$ with $p = \frac{1-\epsilon}{m}$ and $\epsilon^3 m \geq 1$. Then we have*

(i) $\mathbb{E}\sum_{j\geq 1}|\mathcal{C}_j|^k = O(m\epsilon^{-2k+3})$ *for any fixed $k \geq 2$,*

(ii) $\mathbb{E}\sum_{i,j}|\mathcal{C}_i|^2|\mathcal{C}_j|^2 = O(m^2\epsilon^{-2})$,

(iii) $\mathbb{E}\sum_{j\geq 1}|\mathcal{C}_j|^2 \geq cm\epsilon^{-1}$.

Theorem 3.3.13. *Consider $G(m,p)$ with $p = \frac{1+\epsilon}{m}$ with $\epsilon^3 m \geq 1$. Then we have*

(i) $\mathbb{E}|\mathcal{C}(v)|^k = O(\epsilon^{k+1}m^k)$, *for any fixed $k \geq 2$.*

(ii) $\mathbb{E}\sum_{j\geq 2}|\mathcal{C}_j|^k = O(m\epsilon^{-2k+3})$,

(iii) $\mathbb{E}\sum_{i,j\geq 2}|\mathcal{C}_i|^2|\mathcal{C}_j|^2 = O(m^2\epsilon^{-2})$.

Theorem 3.3.14. *Consider $G(m,p)$ with $p = \frac{1-\epsilon}{m}$ where $\epsilon \in [A^{-1}m^{-1/4}, Am^{-1/4}]$. Then, for any small positive constant δ, there exist $K = K(A,\delta)$ and $q_2 = q_2(A,\delta)$ such that*

$$\mathbb{P}\Big(\sum_{|\mathcal{C}_j|\leq\delta\sqrt{m}}|\mathcal{C}_j|^2 \geq Km^{5/4}\Big) \geq q_2 > 0.$$

In the following theorem we derive estimates on the expected cluster size valid as long as $\epsilon^3 m \geq 1$. We believe these estimates should hold for $\mathcal{C}_1$ but we were not able to prove that. The difficulty rises because when $\epsilon^3 m$ is large but does not grow at least logarithmically, it is hard to rule out the possibility that $\mathcal{C}_1$ is discovered after time $\delta\epsilon m$ for some fixed $\delta > 0$. Luckily, for the main proof it suffices to have these estimate for $\mathcal{C}_{\delta\epsilon m}$, that component discovered at time $\delta\epsilon m$, rather than $\mathcal{C}_1$. This becomes evident in the proof of Theorem 3.6.24.

Theorem 3.3.15. *Consider $G(m,p)$ with $p = \frac{1+\epsilon}{m}$ and assume $\epsilon^3 m \geq 1$. For some fixed $\delta > 0$ let $\mathcal{C}_{\delta\epsilon m}$ be the component which is discovered by the exploration process at time $\delta\epsilon m$ (in other words, the length of the excursion of Y_t containing the time $\delta\epsilon m$). Then there is some small value of $\delta > 0$ such that*

(i) $\mathbb{E}|\mathcal{C}_{\delta\epsilon m}| \leq 2\epsilon m - c\epsilon^{-2}$.

(ii) $\mathbb{E}\sum_{\mathcal{C}_j\neq\mathcal{C}_{\delta\epsilon m}}|\mathcal{C}_j|^k \leq Cm\epsilon^{-2k+3}\,, \quad \text{for} \quad k = 2, 4\,,$

where C and c are positive universal constants.

Before proceeding to the proofs of the theorems stated in this section, we first require some preparations about processeses with i.i.d. increments.

Processes with i.i.d. increments

Fix some small $\epsilon > 0$ and let $p = \frac{1+\epsilon}{m}$ for some integer $m > 1$. Let $\{\beta_j\}$ be a sequence of random variables distributed as $\mathrm{Bin}(m, p)$. Let $\{W_t\}_{t\geq 0}$ be a process defined by

$$W_0 = 1, \qquad W_t = W_{t-1} + \beta_t - 1\,.$$

Let τ be the hitting time of 0, i.e.

$$\tau = \min_t \{W_t = 0\}\,.$$

Lemma 3.3.16. *We have*

$$\mathbf{P}(\tau = \infty) = 2\epsilon - \frac{8}{3}\epsilon^2 + O(\epsilon^3)\,, \tag{3.30}$$

and there exists constant $C, c > 0$ such that for all $T \geq \epsilon^{-2}$ we have

$$\mathbf{P}(T \leq \tau < \infty) \leq C\Big(\epsilon^{-2} T^{-3/2} e^{-\frac{(\epsilon^2 - c\epsilon^3)T}{2}}\Big)\,. \tag{3.31}$$

We say that t_0 is a *record minimum* of $\{W_t\}$ if $W_t > W_{t_0}$ for all $t < t_0$.

Lemma 3.3.17. *Denote by Z^w the number of record minima of W_t. Then*

$$\mathbb{E}Z^w = \frac{\epsilon^{-1}}{2} + O(1)\,, \quad \text{and} \quad \mathbb{E}(Z^w)^2 = O(\epsilon^{-2})\,.$$

Lemma 3.3.18. *Denote by γ the random variable*

$$\gamma = \max\{t \,:\, t \text{ is a record minimum of } W_t\}\,.$$

Then we have

$$\mathbb{E}\gamma = O(\epsilon^{-2})\,.$$

For the subcritical case we have the following.

Lemma 3.3.19. *Assume $\epsilon < 0$ in the previous setting. There exists constant $C_1, C_2, c_1, c_2 > 0$ such that for all $T \geq \epsilon^{-2}$ we have*

$$\mathbf{P}(\tau \geq T) \leq C_1\Big(\epsilon^{-2} T^{-3/2} e^{-\frac{(\epsilon^2 - c_1\epsilon^3)T}{2}}\Big)\,,$$

and

$$\mathbf{P}(\tau \geq T) \geq c_2\Big(\epsilon^{-2} T^{-3/2} e^{-\frac{(\epsilon^2 + C_2\epsilon^3)T}{2}}\Big)\,.$$

Furthermore, for any fixed $k \geq 1$

$$\mathbb{E}\tau^k = O(\epsilon^{-2k+1})\,.$$

The proof of Lemma 3.3.19 can be found in [29] Lemma 4.

For the proof of Lemma 3.3.16 we will use the following proposition due to Spitzer (see [36]).

Proposition 3.3.20. *Let $a_0, \ldots, a_{k-1} \in \mathbb{Z}$ satisfy $\sum_{i=0}^{k-1} a_i = -1$. Then there is precisely one $j \in \{0, \ldots, k-1\}$ such that for all $r \in \{0, \ldots, k-2\}$*

$$\sum_{i=0}^{r} a_{(j+i) \bmod k} \geq 0 \,.$$

Proof of Lemma 3.3.16. Let β be a random variable distributed as $\mathrm{Bin}(m,p)$ and let $f(s) = \mathbb{E}s^{\beta}$. It is a classical fact (see [3]) that $1 - \mathbf{P}(\tau = \infty)$ is the unique fixed point of $f(s)$ in $(0,1)$. For $s \in (0,1)$ we have

$$\mathbb{E}s^{\beta} = \Big[1 - p(1-s)\Big]^m = 1-(1+\epsilon)(1-s) + \frac{(1+\epsilon)^2(1-s)^2}{2} - \frac{(1+\epsilon)^3(1-s)^3}{6} + O\Big((1-s)^4\Big)\,,$$

since $(1-x)^m = 1 - mx + \frac{m^2x^2}{2} - \frac{m^3x^3}{6} + O(m^4x^4)$. Write $q = 1-s$ and put $\mathbb{E}s^{\beta} = s$. We get that

$$1 - (1+\epsilon)q + \frac{(1+2\epsilon)q^2}{2} - \frac{q^3}{6} + O(q^4) + O(\epsilon q^3) + O(\epsilon^2 q^2) = 1 - q\,.$$

Solving this gives that that $q = 2\epsilon - \frac{8}{3}\epsilon^2 + O(\epsilon^3)$, as required.

We now turn to proving (3.31). By Proposition 3.3.20, $\mathbf{P}(\tau = t) = \frac{1}{t}\mathbf{P}(W_t = 0)$. Since $\sum_{j=1}^{t} \beta_j$ is distributed as a $\mathrm{Bin}(mt, p)$ random variable we have

$$\mathbf{P}(W_t = 0) = \binom{mt}{t-1} p^{t-1}(1-p)^{mt-(t-1)}\,.$$

Replacing $t-1$ with t in the above formula only changes it by a multiplicative constant which is always between $1/2$ and 2. A straightforward computation using Stirling's approximation gives

$$\mathbf{P}(W_t = 0) = \Theta\Big\{t^{-1/2}(1+\epsilon)^t\Big(1 + \frac{1}{m-1}\Big)^{t(m-1)}\Big(1 - \frac{1+\epsilon}{m}\Big)^{t(m-1)}\Big\}\,. \tag{3.32}$$

Denote $x = (1+\epsilon)\Big(1 + \frac{1}{m-1}\Big)^{m-1}\Big(1 - \frac{1+\epsilon}{m}\Big)^{m-1}$, then

$$\mathbf{P}(\tau \geq T) = \sum_{t \geq T} \mathbf{P}(\tau = t) = \sum_{t \geq T} \frac{1}{t}\mathbf{P}(W_t = 0) = \Theta\Big(\sum_{t \geq T} t^{-3/2}x^t\Big)\,.$$

This sum can be bounded above by

$$T^{-3/2}\sum_{t \geq T} x^t = T^{-3/2}\frac{x^T}{1-x}\,.$$

Observe that as $m \to \infty$ we have that x tends to $(1+\epsilon)e^{-\epsilon}$. By expanding $e^{-\epsilon}$ we find that

$$x = (1+\epsilon)(1-\epsilon+\frac{\epsilon^2}{2}) + \Theta(\epsilon^3) = 1 - \frac{\epsilon^2}{2} + \Theta(\epsilon^3)\,.$$

Using this and the previous bounds on $\mathbf{P}(\tau = \infty)$ we conclude the proof of (3.31). □

Proof of Lemma 3.3.17. This follows immediately since Z^w is a geometric random variable with success probability $p = P(\tau = \infty) = 2\epsilon - \frac{8}{3}\epsilon^2 + O(\epsilon^3)$ by (3.30) of Lemma 3.3.16. □

Proof of Lemma 3.3.18. At each record minimum the process has probability $\Theta(\epsilon)$ of never going below its current location by (3.30) of Lemma 3.3.16. It is a classical fact that the expected size of each excursion between record minimum, on the event that it is finite, is $O(\epsilon^{-1})$. Thus, by Wald's Lemma

$$\mathbb{E}(\gamma) \le C\epsilon^{-1}\mathbb{E}Z^w = O(\epsilon^{-2}).$$

□

Exploration process estimates

In this section we study the process Y_t defined in Section 3.3.1 and provide some useful estimates.

Lemma 3.3.21. *For $p = \frac{1+\epsilon}{m}$ we have*

$$\mathbf{P}\Big(Y_t \ge -45\epsilon^2 m \ \text{ for all } 1 \le t \le 3\epsilon m\Big) \ge 1 - 5e^{-48\epsilon^3 m}\,.$$

Proof. Denote by γ the stopping time

$$\gamma = \min\{t \,:\, N_t \le m - 15\epsilon m\}\,,$$

and consider the process $\{W_t\}$ which has i.i.d. increments distributed as $\text{Bin}(m-15\epsilon m, p)-1$ and $W_0 = 1$. Then we can couple the processes $\{Y_t\}$ and $\{W_t\}$ such that $Y_{t\wedge\gamma} \ge W_{t\wedge\gamma}$ and hence on the event $\gamma > 3\epsilon m$ we have

$$\min_{t \le 3\epsilon m} Y_t \ge \min_{t \le 3\epsilon m} W_t\,. \tag{3.33}$$

Note that the expectation of the increment of W_t is $-15\epsilon - 15\epsilon^2$, thus for any positive $\alpha > 0$ the process $-\alpha W_t$ is a submartingale whence $\exp(-\alpha W_t)$ is a submartingale as well. We put $\alpha = 8\epsilon$ and applying Doob's maximal L^2 inequality (see [11]) yields that

$$\mathbb{E}\Big[\max_{t\le 3\epsilon m} e^{-16\epsilon W_t}\Big] \le 4\mathbb{E}\Big[e^{-16\epsilon W_{3\epsilon m}}\Big]\,.$$

Since $W_{3\epsilon m}$ is distributed as $\mathrm{Bin}(3\epsilon(1-15\epsilon)m^2,p)-3\epsilon m+1$ we obtain by direct computation that

$$\mathbb{E}\Big[\max_{t\le 3\epsilon m} e^{-16\epsilon W_t}\Big] \le 4e^{672\epsilon^3 m}\,.$$

Markov's inequality implies that

$$\mathbf{P}\Big(\exists\, t\le 3\epsilon m \text{ with } W_t \le -45\epsilon^2 m\Big) \le \mathbf{P}\Big(\max_{t\le 3\epsilon m} e^{-16\epsilon W_t} \ge e^{720\epsilon^3 m}\Big) \le 4e^{-48\epsilon^3 m}\,.$$

Note that if there exists $t\le 3\epsilon m$ with $Y_t \le -45\epsilon^2 m$ then by (3.33) either $\gamma\le 3\epsilon m$ or there exists $t\le 3\epsilon m$ such that $W_t\le -45\epsilon^2 m$. Lemma 3.3.2 shows that $\mathbf{P}(\gamma\le 3\epsilon m)\le e^{-c\epsilon m} = o(e^{-48\epsilon^3 m})$ and this concludes the proof of the lemma. $\square$

We now use the estimates of the previous lemma to amplify Lemma 3.3.2.

Lemma 3.3.22. *For $p=\frac{1+\epsilon}{m}$ there exists some fixed $c>0$ such that*

$$\mathbf{P}\Big(\exists t\le 3\epsilon m \text{ with } N_t \le m-t-50\epsilon^2 m\Big) \le 9e^{-c\epsilon^3 m}\,.$$

Proof. Let α_i be independent random variables distributed as $\mathrm{Bin}(m,p)$ and we couple such that $\eta_i\le\alpha_i$ for all i. By (3.8) and the fact that Z_t is non-decreasing we have that for $t\le 3\epsilon m$

$$N_t \ge m-1-\sum_{i=1}^{t}\alpha_i - Z_{3\epsilon m}\,. \tag{3.34}$$

Observe that if for some positive k we have $Y_t\ge -k$ for all $t\le T$ then $Z_T\le k$. Thus, Lemma 3.3.21 together with the fact that $\{Z_t\}$ is increasing implies that

$$\mathbf{P}\Big(Z_{3\epsilon m}\ge 45\epsilon^2 m\Big)\le 5e^{-48\epsilon^3 m}\,.$$

We have that $\sum_{i=1}^{t}\alpha_i$ is distributed as $\mathrm{Bin}(mt,p)$ and has mean $t+\epsilon t$. The same argument using Doob's maximal inequality, as in the proof of Lemma 3.3.21, gives that

$$\mathbf{P}\Big(\exists t\le 3\epsilon m \text{ with } \sum_{i=1}^{t}\alpha_i \ge t+4\epsilon^2 m\Big)\le 4e^{-c\epsilon^3 m}\,,$$

for some fixed $c>0$. The assertion of the lemma follows by putting the last two inequalities into (3.34). $\square$

Lemma 3.3.23. *Assume that $p=\frac{1+\epsilon}{m}$ and that $\epsilon^3 m\ge 1$. Then there exist a constant $c>0$ such that for any a satisfying $1\le a\le\sqrt{\epsilon^3 m}$ we have*

$$\mathbf{P}\Big(Y_t>0 \;\text{ for all } a\sqrt{m/\epsilon}\le t\le 2\epsilon m - a\sqrt{m/\epsilon}\Big)\ge 1-2e^{-ca^2}\,.$$

Proof. Denote by γ the stopping time

$$\gamma = \min\{t \,:\, N_t < m - t - 50\epsilon^2 m\}\,.$$

Lemma 3.3.22 states that

$$\mathbf{P}(\gamma \le 3\epsilon m) \le 9e^{-c\epsilon^3 m}\,,$$

for some constant $c > 0$. Let $\{W_t\}$ be a process with independent increments distributed as $\mathrm{Bin}(m - t - 50\epsilon^2 m, p) - 1$ (note that the increments are not identically distributed) and $W_0 = 1$. As usual we can couple such that $Y_{t\wedge\gamma} \ge W_{t\wedge\gamma}$ for all t. Hence, if $\gamma \ge 2\epsilon m$ and there exists $t \le 2\epsilon m$ with $Y_t \le 0$ then it must be that $W_t \le 0$. We conclude that it suffices to show the assertion of the lemma to the process $\{W_t\}$ and this is our next goal.

For any $\alpha > 0$ we have

$$\mathbb{E}\Big[e^{-\alpha(W_t - W_{t-1})} \mid W_{t-1}\Big] = e^{\alpha}[1 - p(1 - e^{-\alpha})]^{m-t-50\epsilon^2 m}\,.$$

We use $1 - x \le e^{-x}$ with $x = p(1 - e^{-\alpha})$ and $1 - e^{-\alpha} \ge \alpha - \alpha^2$ for α small enough (we will eventually take $\alpha = O(\epsilon)$) to get

$$\mathbb{E}\Big[e^{-\alpha(W_t - W_{t-1})} \mid W_{t-1}\Big] \le e^{\alpha^2(1+\epsilon) - \alpha(\epsilon - \frac{t}{m}(1+\epsilon) - 50\epsilon^2(1+\epsilon))}\,. \tag{3.35}$$

Thus, we learn that the process

$$e^{-\alpha W_t} e^{-(1+\epsilon)\alpha^2 t - (1+\epsilon)\alpha\frac{t^2}{2m} + \epsilon\alpha t(1 - 50\epsilon(1+\epsilon))}\,,$$

is a supermartinagle. Write

$$f(t) = t\big[-(1+\epsilon)\alpha^2 + \epsilon\alpha(1 - 50\epsilon(1+\epsilon))\big] - t^2\frac{(1+\epsilon)\alpha}{2m}\,.$$

We apply the optional stopping theorem on the stopping time $\tau = \min\{t \ge \sqrt{m/\epsilon} : W_t = 0\}$ and get that

$$\mathbb{E}e^{f(\tau)} \le 1\,.$$

Direct calculation gives that when we put $\alpha = \frac{1}{3}\epsilon$ the function f attains its minimum on the interval $[a\sqrt{m/\epsilon}, \epsilon m]$ at $\tau = a\sqrt{m/\epsilon}$ for any $a \in [1, \sqrt{\epsilon^3 m}/3]$. Hence

$$\mathbf{P}\big(a\sqrt{m/\epsilon} \le \tau \le \epsilon m\big) \le \mathbf{P}\big(e^{f(\tau)} \ge e^{f(a\sqrt{m/\epsilon})}\big)\,.$$

An immediate calculation shows that $f(a\sqrt{m/\epsilon}) \ge ca\sqrt{m\epsilon^3}$ and we learn by Markov's inequality that

$$\mathbf{P}\big(a\sqrt{m/\epsilon} \le \tau \le \epsilon m\big) \le e^{-ca\sqrt{m\epsilon^3}} \le e^{-ca^2}\,, \tag{3.36}$$

since $a \le \sqrt{m\epsilon^3}$.

We are left to estimate $\mathbf{P}(\epsilon m \leq \tau \leq 2\epsilon m - a\sqrt{m/\epsilon})$. To that aim we define a new process $\{X_t\}_{t\geq 0}$ by $X_t = W_{\epsilon m+t}$. By (3.35), for positive α we have that

$$\mathbb{E}\Big[e^{-\alpha(X_t - X_{t-1})} \mid X_{t-1}\Big] \leq e^{\alpha^2(1+\epsilon)-\alpha(\epsilon - \frac{t+\epsilon m}{m}(1+\epsilon)-50\epsilon^2(1+\epsilon))} .$$

This together with a straight forward computation yields that the process

$$e^{-\alpha X_t} e^{-\alpha^2 t(1+\epsilon)-\alpha\left(\frac{(1+\epsilon)t^2}{2m}+55\epsilon^2 t\right)} ,$$

is a supermartingale. Write τ for the stopping time

$$\tau = \min\{t \geq 0 \,:\, X_t = 0\} .$$

Optional stopping yields that

$$\mathbb{E}\Big[e^{-\alpha^2\tau(1+\epsilon)-\alpha\left(\frac{\tau^2}{2m}+55\epsilon^2(\tau\wedge 4\epsilon m)\right)}\Big] \leq \mathbb{E}\Big[e^{-\alpha X_0}\Big] \leq e^{\alpha^2\epsilon m(1+\epsilon)-\alpha\left(\frac{\epsilon^2 m}{2}-55\epsilon^3 m\right)} , \tag{3.37}$$

where the last inequality is an immediate calculation with (3.35) and the fact that $X_0 = W_{\epsilon m}$. Observe that the exponent on the left hand side of the previous display is

$$f(\tau) = -\alpha^2\tau(1+\epsilon) - \alpha(\tau^2/2m + 55\epsilon^2\tau) ,$$

which is a non-increasing function of τ on $[0,\infty)$. Hence, for any $a \in [1, \sqrt{\epsilon^3 m}]$ we get that

$$\mathbf{P}\big(\tau \leq \epsilon m - a\sqrt{m/\epsilon}\big) \leq \mathbf{P}\Big(e^{f(\tau)} \geq e^{f(\epsilon m - a\sqrt{m/\epsilon})}\Big) . \tag{3.38}$$

We have that

$$f(\epsilon m - a\sqrt{m/\epsilon}) \geq -2\alpha^2\epsilon m - \alpha\Big(\frac{\epsilon^2 m}{2} - a\sqrt{\epsilon m} + \frac{1}{2}a^2\epsilon^{-1} + 55\epsilon^3 m\Big) .$$

We use Markov inequality and (3.37) to get

$$\begin{aligned}\mathbf{P}\Big(e^{f(\tau)} \geq e^{f(\epsilon m - a\sqrt{m/\epsilon})}\Big) &\leq e^{4\alpha^2\epsilon m - \alpha\left(a\sqrt{\epsilon m} - \frac{1}{2}a^2\epsilon^{-1} - 110\epsilon^3 m\right)} \\ &\leq e^{4\alpha^2\epsilon m - c\alpha a\sqrt{\epsilon m}} ,\end{aligned}$$

where in the last inequality we used our assumption on a and ϵ. We choose $\alpha \approx a(\epsilon m)^{-1/2}$ that minimizes the last expression. This yields

$$\mathbf{P}\Big(e^{f(\tau)} \geq e^{f(\epsilon m - a\sqrt{m/\epsilon})}\Big) \leq e^{-ca^2} .$$

We put this into (3.38), which together with (3.36) yields the assertion of the lemma. $\square$

Lemma 3.3.24. *Assume that* $p = \frac{1+\epsilon}{m}$. *Write* $\tau = \min\{t : Y_t = 0\}$, *then for any small* $\alpha > 0$

$$\mathbb{E}\big[e^{\alpha Y_{\epsilon^{-2}}} \mid \tau \geq \epsilon^{-2}\big] \leq C e^{2\alpha\epsilon^{-1} + \alpha^2\epsilon^{-2}} .$$

Proof. We have that $\mathbf{P}(\tau \geq \epsilon^{-2}) \geq c\epsilon$. To see this we perform the usual argument of bounding Y_t below by a process of independent increments (until a stopping time, using Lemma 3.3.2) and using Lemma 3.3.16. This has been done in this section several times so we omit the details. Thus, it suffices to bound from above $\mathbb{E}e^{\alpha Y_{\epsilon^{-2}}}\mathbf{1}_{\{\tau\geq\epsilon^{-2}\}}$. Since we can bound Y_t by a process W_t which has i.i.d. $\mathrm{Bin}(m,p)-1$ increments, it suffices to bound the same expectation for W_t. Write $\gamma = \min\{t : W_t = 0 \text{ or } W_t \geq \epsilon^{-1}\}$. We have

$$\mathbb{E}e^{\alpha W_{\epsilon^{-2}}}\mathbf{1}_{\{\tau\geq\epsilon^{-2}\}} \leq \mathbb{E}e^{\alpha W_{\epsilon^{-2}}}\mathbf{1}_{\{\tau\geq\epsilon^{-2},\gamma\geq\epsilon^{-2}\}} + \mathbb{E}e^{\alpha W_{\epsilon^{-2}}}\mathbf{1}_{\{\tau\geq\epsilon^{-2},\gamma<\epsilon^{-2}\}}\,.$$

For the first term on the right hand side we note that on $\gamma \geq \epsilon^{-2}$ we have that $W_{\epsilon^{-2}} \leq \epsilon^{-1}$, so

$$\mathbb{E}e^{\alpha W_{\epsilon^{-2}}}\mathbf{1}_{\{\tau\geq\epsilon^{-2},\gamma\geq\epsilon^{-2}\}} \leq C\epsilon e^{\alpha\epsilon^{-1}}\,.$$

For the second term we condition on $\{\tau \geq \epsilon^{-2}, \gamma < \epsilon^{-2}\}$ (which implies $W_\gamma \geq \epsilon^{-1}$ and $\gamma < \epsilon^2$) to get that

$$\mathbb{E}e^{\alpha W_{\epsilon^{-2}}}\mathbf{1}_{\{\tau\geq\epsilon^{-2},\gamma<\epsilon^{-2}\}} \leq \mathbf{P}(W_\gamma \geq \epsilon^{-1})\mathbb{E}[e^{\alpha W_\gamma}e^{\alpha(W_{\epsilon^{-2}}-W_\gamma)} \mid W_\gamma \geq \epsilon^{-1}, \gamma < \epsilon^{-2}]\,. \tag{3.39}$$

We have that $\mathbf{P}(W_\gamma \geq \epsilon^{-1}) = O(\epsilon)$ by Lemma 7 of [31]. We condition in addition on W_γ and γ and pull out the $e^{\alpha W_\gamma}$ factor. By the strong Markov property we have that conditioned on all these, the random variable $W_{\epsilon^{-2}} - W_\gamma$ is distributed as the sum of $\epsilon^{-2} - \gamma$ i.i.d. copies of $\mathrm{Bin}(m,p) - 1$ random variables. Thus,

$$\mathbb{E}[e^{\alpha(W_{\epsilon^{-2}}-W_\gamma)} \mid W_\gamma, \gamma < \epsilon^2] \leq e^{-\alpha\epsilon^{-2}}[1 + p(e^\alpha - 1)]^{m\epsilon^{-2}}\,.$$

Furthermore, Lemma 5 of [30] states that conditioned on $W_\gamma \geq \epsilon^{-1}$ and $\gamma < \epsilon^2$ the distribution of $W_\gamma - \epsilon^{-1}$ is bounded above by $\mathrm{Bin}(m,p)$, whence

$$\mathbb{E}[e^{\alpha W_\gamma} \mid W_\gamma \geq \epsilon^{-1}, \gamma < \epsilon^{-2}] \leq e^{\alpha\epsilon^{-1}}[1 + p(e^\alpha - 1)]^m\,.$$

Putting this back into (3.39) gives

$$\mathbb{E}e^{\alpha W_{\epsilon^{-2}}}\mathbf{1}_{\{\tau\geq\epsilon^{-2},\gamma<\epsilon^{-2}\}} \leq C\epsilon e^{-\alpha(\epsilon^{-2}-\epsilon^{-1})}[1 + p(e^\alpha - 1)]^{m(\epsilon^{-2}+1)}\,.$$

Putting all these together we get

$$\begin{aligned}\mathbb{E}\big[e^{\alpha Y_{\epsilon^{-2}}} \mid \tau \geq \epsilon^{-2}\big] &\leq Ce^{\alpha\epsilon^{-1}} + Ce^{-\alpha(\epsilon^{-2}-\epsilon^{-1})}[1 + p(e^\alpha - 1)]^{m(\epsilon^{-2}+1)} \\ &\leq Ce^{\alpha\epsilon^{-1}} + Ce^{-\alpha(\epsilon^{-2}-\epsilon^{-1})}e^{(1+\epsilon)(\alpha+\alpha^2)(\epsilon^{-2}+1)}\,,\end{aligned}$$

The lemma follows now by an immediate calculation. □

Lemma 3.3.25. *Let $p = \frac{1+\epsilon}{m}$ and assume $\epsilon^3 m \geq 1$. Then for any $\ell > 0$, we have*

$$\mathbf{P}(|\mathcal{C}(v)| \geq 2\epsilon m + \ell) \leq C\epsilon e^{\frac{-c\ell^2(2\epsilon m+\ell)}{m^2}}\,.$$

Proof. We assume that $\ell \geq 2\sqrt{m/\epsilon}$ since otherwise the exponential is of constant order and the assertion of the lemma follows simply from Lemma 3.3.16. Recall that $|\mathcal{C}(v)|$ is distributed as the first hitting time τ of Y_t at 0. We put $T = 2\epsilon m + \ell$ and condition on $Y_{\epsilon^{-2}}$ and on $\tau \geq \epsilon^{-2}$. That is,

$$\mathbf{P}(\tau \geq 2\epsilon m + \ell) = \mathbf{P}(\tau \geq \epsilon^{-2})\mathbb{E}\big[\mathbf{P}(\tau \geq T \mid Y_{\epsilon^{-2}}, \tau \geq \epsilon^{-2})\big]\,. \tag{3.40}$$

Since Y_t is bounded above by a process with increments distributed as $\mathrm{Bin}(m,p)-1$, we learn by Lemma 3.3.16 that $\mathbf{P}(\tau \geq \epsilon^{-2}) = O(\epsilon)$. The second term will give us the exponential in the assertion of the Lemma simply because Y_T has small probability of being positive at this time. Indeed, since the increments of Y_t are stochastically bounded above by $\mathrm{Bin}(m-t,p)-1$ we have that for any small $\alpha > 0$

$$\mathbb{E}\Big[e^{\alpha(Y_t - Y_{t-1})} \mid Y_{t-1}\Big] \leq e^{-\alpha}[1 + p(e^{\alpha}-1)]^{m-t} \leq e^{-\alpha + (1+\epsilon)(\alpha+\alpha^2)(1-t/m)}\,,$$

since $e^{\alpha} - 1 \leq \alpha + \alpha^2$ for small enough α. Summing this over t ranging from ϵ^{-2} to T gives

$$\begin{aligned}
\mathbb{E}\big[e^{\alpha Y_T} \mid Y_{\epsilon^{-2}}, \tau \geq \epsilon^{-2}\big] &\leq e^{-\alpha(T-\epsilon^{-2}) + (1+\epsilon)(\alpha+\alpha^2)\left(T - \epsilon^{-2} - \frac{T^2-\epsilon^{-4}}{2m}\right)} e^{\alpha Y_{\epsilon^{-2}}} \\
&\leq e^{\alpha^2 T(1+\epsilon) - \alpha\left[\frac{T^2-\epsilon^{-4}}{2m} - \epsilon T\right]} e^{\alpha Y_{\epsilon^{-2}}}\,.
\end{aligned}$$

Hence,

$$\begin{aligned}
\mathbb{E}\big[e^{\alpha Y_T} \mid \tau \geq \epsilon^{-2}\big] &\leq \mathbb{E}[e^{\alpha Y_{\epsilon^{-2}}} \mid \tau \geq \epsilon^{-2}] e^{\alpha^2 T(1+\epsilon) - \alpha\left[\frac{T^2-\epsilon^{-4}}{2m} - \epsilon T\right]} \\
&\leq C e^{\alpha^2 (T+\epsilon^{-2})(1+\epsilon) - \alpha\left[\frac{T^2-\epsilon^{-4}}{2m} - \epsilon T - 2\epsilon^{-1}\right]}\,,
\end{aligned}$$

where the last inequality is due to Lemma 3.3.24. Hence, by Markov's inequality this is also an upper bound on $\mathbf{P}(Y_T \geq 0 \mid \tau \geq \epsilon^{-2})$ which is what we aim to estimate. We now choose α

$$\alpha = \frac{\frac{T^2-\epsilon^{-4}}{2m} - \epsilon T - 2\epsilon^{-1}}{2(T+\epsilon^{-2})}\,,$$

which is positive and of order ℓ/m since $\ell \geq 2\sqrt{m/\epsilon}$ and minimizes the above expectation. We get that

$$\mathbf{P}(Y_T \geq 0 \mid \tau \geq \epsilon^{-2}) \leq C e^{-\frac{cT(T-2\epsilon m)^2}{m^2}}\,,$$

for some $c > 0$ by a straightforward calculation, concluding our proof. $\square$

Proof of near-critical random graph theorems.

We are now ready to prove the Theorems stated in Section 3.3.3.

Proof of Theorem 3.3.8. We begin by proving the upper bound on $\mathbb{E}|\mathcal{C}_1|$. For any positive integer ℓ define by X_ℓ the random variable

$$X_\ell = \left|\left\{v \,:\, |\mathcal{C}(v)| \geq \ell\right\}\right|.$$

Observe that if $|\mathcal{C}_1| \geq \ell$, then we must have that $|X_\ell| \geq |\mathcal{C}_1|$. Thus for any positive integer ℓ we have

$$\mathbb{E}|\mathcal{C}_1| \leq \ell\,\mathbf{P}(|\mathcal{C}_1| < \ell) + \mathbb{E}X_\ell\,. \tag{3.41}$$

We take $\ell = \frac{1}{20}\epsilon m$ and since Lemma 3.3.23 implies that $\mathbf{P}(|\mathcal{C}_1| \leq \ell) \leq Ce^{-c\epsilon^3 m}$ and $\epsilon^3 m \geq A\log m$ we have that the first term on the right hand side of (3.41) is $o(1)$. We now turn to bound the second term on the right hand side of (3.41). Since $\mathbb{E}X_\ell = m\mathbf{P}(|\mathcal{C}(v)| \geq \ell)$ it suffices to bound from above $\mathbf{P}(|\mathcal{C}(v)| \geq \ell)$. Recall that $|\mathcal{C}(v)|$ is the hitting time of the process $\{Y_t\}$ at 0. Let $\{W_t\}$ be a process with independent increments distributed as $\mathrm{Bin}(m,p)-1$ and $W_0 = 1$, as in Lemma 3.3.16. Let $\tau = \min_t\{W_t = 0\}$ be the hitting time of W_t at 0, then it is clear that we can couple W_t and Y_t such that $|\mathcal{C}(v)| \leq \tau$. Thus

$$\mathbf{P}(|\mathcal{C}(v)| \geq \ell) \leq \mathbf{P}(\tau \geq \ell) = \mathbf{P}(\tau = \infty) + \mathbf{P}(\ell \leq \tau < \infty)\,.$$

We now apply Lemma 3.3.16 with $T = \ell = \frac{1}{20}\epsilon m$ and get by the previous display that

$$\begin{aligned}
\mathbf{P}(|\mathcal{C}(v)| \geq \epsilon m/20) &\leq 2\epsilon - \frac{8}{3}\epsilon^2 + O(\epsilon^3) + C_1\epsilon^{-7/2}m^{-3/2}e^{-\epsilon^3 m/4} \\
&= 2\epsilon - \frac{8}{3}\epsilon^2 + O(\epsilon^3)\,,
\end{aligned}$$

as long as $\epsilon^3 m \geq A\log m$ for large enough A. We conclude that

$$\mathbb{E}X_\ell \leq 2\epsilon m - \frac{8}{3}\epsilon^2 m + O(\epsilon^3 m)\,,$$

which together with (3.41) concludes the proof of the upper bound on $\mathbb{E}|\mathcal{C}_1|$.

We turn to the proof of the lower bound on $\mathbb{E}|\mathcal{C}_1|$. Recall that at each record minimum of the process $\{Y_t\}$ we are starting the exploration of a new component. Write

$$\gamma = \max\left\{t \leq \epsilon m \,:\, Y_t \text{ is at a record minimum}\right\},$$

and

$$\tau = \min\{t \geq 0 : Y_{\epsilon m + t} < 0\}\,.$$

Then we have that

$$|\mathcal{C}_1| \geq \epsilon m - \gamma + \tau\,. \tag{3.42}$$

Thus, in order to complete the proof we will provide an upper bound on $\mathbb{E}\gamma$ and a lower bound on $\mathbb{E}\tau$. Let $\{W_t\}$ be a process defined as in Lemma 3.3.18 with i.i.d. increments distributed as $\mathrm{Bin}(m(1-\epsilon/2),p)-1$. Define the stopping time β by

$$\beta = \min\{t \,:\, N_t \leq m(1-\epsilon/2)\}\,,$$

then it is clear we can couple $\{Y_{t\wedge\beta}\}$ with $\{W_{t\wedge\beta}\}$ such that the increments of the first are larger than of the latter process. This guarantees that every record minimum of the first process is also a record minimum of the second, and thus if we put

$$\gamma^w = \max\Big\{t \,:\, W_t \text{ is at a record minimum}\Big\}\,,$$

then we can couple such that

$$\gamma \mathbf{1}_{\{\text{no record minima at times } [\epsilon m/10, \epsilon m]\}} \leq \gamma^w + \epsilon m \mathbf{1}_{\{\beta \leq \epsilon m/10\}}\,.$$

Lemma 3.3.23 shows that the probability that there is a record minimum at some time between $\epsilon m/10$ and ϵm decays faster than m^{-2} provided that $\epsilon^3 m \geq A \log m$ for A large enough. Hence, taking expectations on both sides and using Lemma 3.3.18 and Lemma 3.3.2 gives that $\mathbb{E}\gamma = O(\epsilon^{-2})$.

We now turn to give a lower bound on $\mathbb{E}\tau$. We begin by estimating $\mathbb{E}\tau^2$. As before, define the process $\{X_t\}_{t\geq 0}$ by $X_t = Y_{\epsilon m + t}$ and note that $X_t - X_{t-1} = \eta_{\epsilon m + t}$. For any t such that $N_{t+\epsilon m} \geq m - (t + \epsilon m) - 50\epsilon^2 m$ we have

$$\mathbb{E}[X_{t+1} - X_t \mid \mathcal{F}_t] \geq -\frac{t(1+\epsilon)}{m} - 55\epsilon^2\,.$$

Thus the process $\{X_{t\wedge T} + \frac{(t\wedge T)^2(1+\epsilon)}{2m} + 55\epsilon^2 (t \wedge T)\}$ is a submartingale, where T is defined as

$$T = \min\{t - \epsilon m : t \geq \epsilon m, N_t \leq m - t - 50\epsilon^2 m\}\,.$$

Optional stopping yields that

$$\mathbb{E}(\tau \wedge T)^2 \geq \frac{2m}{1+\epsilon}(\mathbb{E}X_0 - \mathbb{E}X_{\tau\wedge T}) - \frac{110\epsilon^2 m}{1+\epsilon}\mathbb{E}[\tau \wedge T]\,. \tag{3.43}$$

By Lemma 3.3.23, we have

$$\mathbb{P}(\tau < \epsilon m - a\sqrt{m/\epsilon}) \leq e^{-ca^2}.$$

Also by lemma 3.3.23, one can deduce

$$\begin{aligned}
\mathbb{P}(\tau > \epsilon m + a\sqrt{m/\epsilon}) &\leq \mathbb{P}(\tau > \epsilon m + a\sqrt{m/\epsilon}, Y_t > 0 \text{ for } \ t \in [\frac{a}{2}\sqrt{m/\epsilon}, \epsilon m]) + e^{-ca^2} \\
&\leq \mathbb{P}(|\mathcal{C}_1| > 2\epsilon m + \frac{a}{2}\sqrt{m/\epsilon}) + e^{-ca^2} \\
&\leq \mathbb{P}(X_{2\epsilon m a\sqrt{m/\epsilon}/2} > 2\epsilon m) + e^{-ca^2},
\end{aligned}$$

where $X_{2\epsilon m a\sqrt{m/\epsilon}/2}$ is the number of vertices v such that $|\mathcal{C}_v| \geq 2\epsilon m + \frac{a}{2}\sqrt{m/\epsilon}$ as defined in the beginning of the proof. By Lemma 3.3.25, we have

$$\mathbb{E}X_{2\epsilon m a\sqrt{m/\epsilon}/2} \leq C m\epsilon e^{-ca^2}\,.$$

Plugging this into the previous inequality and using Markov's inequality shows that $\mathbb{E}[\tau \wedge T] = O(\epsilon m)$ and

$$\mathbb{P}(|\tau - \epsilon m| > a\sqrt{m/\epsilon}) \leq Ce^{-ca^2}. \tag{3.44}$$

Lemma 3.3.22 shows that $\mathbf{P}(T \leq 2\epsilon m) \leq m^{-2}$, and so $\mathbb{E}X_{\tau\wedge T} = o(1)$ and $\mathbb{E}(\tau \wedge T)^2 = \mathbb{E}\tau^2 + o(1)$. We get that

$$\mathbb{E}\tau^2 \geq 2m\mathbb{E}Y_{\epsilon m} - o(1)\,.$$

We bound from below $\mathbb{E}Y_{\epsilon m}$ using the approximating process $\widetilde{Y}_t$ defined in (3.21). We have that $\mathbb{E}\widetilde{Y}_{\epsilon m} = \epsilon^2 m^2 + O(\epsilon^3 m)$ and using (3.22) and Lemma 3.3.26 we deduce the same estimate for $\mathbb{E}Y_{\epsilon m}$. This yields that

$$\mathbb{E}\tau^2 \geq \epsilon^2 m^2 - C\epsilon^3 m^2\,,$$

for some $C > 0$. Inequality (3.44) gives that for some $C > 0$ we have

$$\mathrm{Var}(\tau) \leq \mathbb{E}\Big[(\tau - \epsilon m)^2\Big] \leq \frac{Cm}{\epsilon}\,.$$

We conclude

$$\mathbb{E}\tau = \sqrt{\mathbb{E}\tau^2 - \mathrm{Var}(\tau)} \geq \epsilon m\sqrt{1 - C\epsilon - \frac{C}{\epsilon^3 m}} \geq \epsilon m - C\epsilon^2 m - C\epsilon^{-2}\,,$$

since $\sqrt{1-x} \geq 1 - x$ for $x \in (0,1)$. Using this and our estimate on $\mathbb{E}\gamma$ in (3.42) finishes the proof. $\square$

Proof of Theorem 3.3.9. Since component sizes are excursions' length above past minima and $Y_0 = 1$, Lemma 3.3.23 immediately yields the bound

$$\mathbf{P}\Big(|\mathcal{C}_1| \leq 2\epsilon m - A\sqrt{\frac{m}{\epsilon}}\Big) \leq e^{-cA^2}\,, \tag{3.45}$$

valid for any A satisfying $1 \leq A \leq \sqrt{\epsilon^3 m}$. For the upper bound we use Lemma 3.3.25 stating that

$$\mathbf{P}(|\mathcal{C}(v)| \geq 2\epsilon m + A\sqrt{m/\epsilon}) = O(\epsilon e^{-cA^2})\,.$$

Write $X = |\{v : |\mathcal{C}(v)| \geq 2\epsilon m + A\sqrt{m/\epsilon}\}|$ so that $\mathbb{E}X = O(\epsilon m e^{-cA^2})$. As usual we have

$$\mathbf{P}(|\mathcal{C}_1| \geq 2\epsilon m + A\sqrt{m/\epsilon}) \leq \mathbf{P}(X \geq 2\epsilon m) = O(e^{-cA^2})\,,$$

by Markov's inequality, concluding the proof. $\square$

Proof of Corollary 3.3.10. Part (i) of the corollary follows immediately from Theorem 3.3.9 by integration. Indeed,

$$\mathbb{E}\Big[\Big||\mathcal{C}_1| - 2\epsilon m\Big|^k\Big] = \sum_{\ell} \ell^{k-1}\mathbb{P}(|\mathcal{C}_1 - 2\epsilon m| > \ell)$$

$$\leq \sum_{\ell=1}^{\sqrt{\frac{m}{\epsilon}}} \ell^{k-1} + C \sum_{\ell=\sqrt{\frac{m}{\epsilon}}}^{\epsilon m} \ell^{k-1} e^{\frac{-c\ell^2\epsilon}{m}} + \sum_{\ell\geq \epsilon m} \ell^{k-1} e^{\frac{-c\ell^3}{m^2}} ,$$

where we bounded the second sum on the right hand side using Theorem 3.3.9 and the last sum using Lemma 3.3.25 (which is valid for all $\ell > 0$ and not limited by $\ell \leq \sqrt{\epsilon^3 m}$) and the usual Markov inequality on the variable $X = |\{v : |\mathcal{C}(v)| \geq 2\epsilon m + \ell\}|$. A quick calculation now shows each term is of order at most $(m/\epsilon)^{k/2}$, concluding our proof. □

Proof of Theorem 3.3.12. We begin by proving (i). As before, $|\mathcal{C}(v)|$ is stochastically dominated by the random variable τ defined in Lemma 3.3.19. This Lemma gives that for any fixed $k \geq 1$

$$\mathbb{E}|\mathcal{C}(v)|^k = O(\epsilon^{-2k+1}) .$$

Number the vertices of $G(m, p)$ arbitrarily $v_1, \ldots, v_m$ and observe that

$$\sum_{j\geq 1} |\mathcal{C}_j|^k = \sum_{i=1}^{m} |\mathcal{C}(v_i)|^{k-1} ,$$

because each component $\mathcal{C}_j$ is counted in the sum in the right hand size precisely $|\mathcal{C}_j|$ times. By symmetry we learn that

$$\mathbb{E}\sum_{j\geq 1} |\mathcal{C}_j|^k = m\mathbb{E}|\mathcal{C}(v)|^{k-1} = O(m\epsilon^{-2k+3}) ,$$

finishing the first assertion of the theorem.

We proceed to prove (ii). Recall that $\sum_j |\mathcal{C}_j|^2 = \sum_v |\mathcal{C}(v)|$. Thus,

$$\mathbb{E}\Big(\sum_j |\mathcal{C}_j|^2\Big)^2 = \mathbb{E}\sum_{v,w} |\mathcal{C}(v)||\mathcal{C}(w)|1_{\{\mathcal{C}(v)=\mathcal{C}(w)\}} + \mathbb{E}\sum_{v,w} |\mathcal{C}(v)||\mathcal{C}(w)|1_{\{\mathcal{C}(v)\neq\mathcal{C}(w)\}} .$$

The first term on the right hand side is $\sum_v \mathbb{E}|\mathcal{C}(v)|^3$ which equals $\mathbb{E}\sum_j |\mathcal{C}_j|^4$ and is upper bounded by $O(m\epsilon^{-5})$ by part (i) of the theorem. This bound is $O(m^2\epsilon^{-2})$ since $\epsilon^3 m \geq 1$. For the second term we note that we can write $\mathbb{E}\sum_{v,w} |\mathcal{C}(v)||\mathcal{C}(w)|1_{\{\mathcal{C}(v)\neq\mathcal{C}(w)\}}$ as

$$\mathbb{E}\sum_{w} |\mathcal{C}(w)| \sum_{v} |\mathcal{C}(v)|1_{\{v\notin\mathcal{C}(w)\}} = \mathbb{E}\sum_{w} |\mathcal{C}(w)| \sum_{\{v\notin\mathcal{C}(w)\}} |\mathcal{C}(v)| .$$

Conditioned on $\mathcal{C}(w)$ the distribution of the rest of the graph is also subcritical random graph with ϵ' bigger than ϵ. Thus the estimate of part (i) of the theorem (together with the fact that $\sum_v |C(v)| = \sum_j |\mathcal{C}_j|^2$) can be applied and we may bound

$$\mathbb{E}\sum_{v,w} |\mathcal{C}(v)||\mathcal{C}(w)|1_{\{\mathcal{C}(v)\neq\mathcal{C}(w)\}} \leq Cm\epsilon^{-1}\mathbb{E}\sum_{w} |\mathcal{C}(w)| = O(m^2\epsilon^{-2}) ,$$

which finishes the proof of (ii).

To prove part (iii) of the theorem, let W_t be a process with i.i.d. increment distributed as $\mathrm{Bin}(m - 5\epsilon^{-2}, \frac{1-\epsilon}{m}) - 1$ and $W_0 = 1$. Let

$$\tau = \min\{t : N_t < m - 5\epsilon^{-2}\}.$$

As usual we can couple such that $Y_{t\wedge\tau} \geq W_{t\wedge\tau}$. Let $\gamma = \min\{t : W_t \leq 0\}$. For any T We have

$$\begin{aligned} \mathbb{P}(\gamma \geq T) &= \mathbb{P}(\gamma \geq T, \tau \leq T) + \mathbb{P}(\gamma \geq T, \tau > T) \\ &\leq \mathbb{P}(\tau \leq T) + \mathbb{P}(|\mathcal{C}(v)| \geq T), \end{aligned}$$

which implies

$$\mathbb{P}(|\mathcal{C}(v)| \geq T) \geq \mathbb{P}(\gamma \geq T) - \mathbb{P}(\tau \leq T)\,. \tag{3.46}$$

Put $T = \epsilon^{-2}$ we have by Lemma 3.3.2 that

$$\mathbb{P}(\tau \leq T) \leq e^{-c\epsilon^{-2}}\,.$$

Furthermore, Lemma 3.3.19 shows that

$$\mathbb{P}(\gamma \geq \epsilon^{-2}) \geq c\epsilon\,,$$

for some constant $c > 0$. Thus, by (3.46) we get that

$$\mathbb{P}(|\mathcal{C}(v)| \geq \epsilon^{-2}) \geq c\epsilon - e^{-c\epsilon^{-2}}\,,$$

which implies $\mathbb{E}|\mathcal{C}(v)| \geq c\epsilon^{-1}$ and concludes the proof. $\square$

Proof of Theorem 3.3.13. The proof of (i) is a calculation using Lemma 3.3.25. We have

$$\mathbb{E}|\mathcal{C}(v)|^k = \sum_{\ell=1}^{\epsilon^{-2}} \ell^{k-1}\mathbf{P}(|\mathcal{C}(v)| \geq \ell) + \sum_{\ell=\epsilon^{-2}}^{10\epsilon m} \ell^{k-1}\mathbf{P}(|\mathcal{C}(v)| \geq \ell) + \sum_{\ell=10\epsilon m}^{m} \ell^{k-1}\mathbf{P}(|\mathcal{C}(v)| \geq k)\,.$$

For the first sum we use the estimate $\mathbf{P}(|\mathcal{C}(v)| \geq l) \leq O(\epsilon + \ell^{-1/2})$ appearing in the proof of Proposition 1 of [31]. We get

$$\sum_{\ell=1}^{\epsilon^{-2}} \ell^{k-1}\mathbf{P}(|\mathcal{C}(v)| \geq \ell) \leq C\sum_{\ell=1}^{\epsilon^{-2}} \ell^{k-1}(\epsilon + \ell^{-1/2}) = O(\epsilon^{-2k+1})\,.$$

For the second sum, since Y_t is bounded above by a process with i.i.d. increments $\mathrm{Bin}(m,p)$-1, each term is of order ϵ by Lemma 3.3.16. This gives the main contribution of $O(\epsilon^{k+1}m^k)$. Lastly, the third sum we bound using Lemma 3.3.25 to get

$$\sum_{\ell=10\epsilon m}^{m} \ell^{k-1}\mathbf{P}(|\mathcal{C}(v)| \geq k) \leq C\epsilon \sum_{\ell=10\epsilon m}^{m} \ell^{k-1}e^{-cm^{-2}\ell^3}\,.$$

Since $\epsilon^3 m \geq 1$ we may bound the sum above by summing from $m^{2/3}$ to m. A straightforward calculation then gives that

$$\sum_{\ell=10\epsilon m}^{m} \ell^{k-1}\mathbf{P}(|\mathcal{C}(v)| \geq k) \leq C\epsilon(\epsilon m)^{k-1}m^{2/3} = O(\epsilon^{k+1}m^k)\,,$$

which finishes the proof of (i). We proceed to prove (ii). We have that

$$\mathbb{E}\sum_{j\geq 2}|\mathcal{C}_j|^k = \mathbb{E}\sum_{j\geq 2}|\mathcal{C}_j|^k\mathbf{1}_{\{|\mathcal{C}_1|<1.5\epsilon m\}} + \mathbb{E}\sum_{j\geq 2}|\mathcal{C}_j|^k\mathbf{1}_{\{|\mathcal{C}_1|\geq 1.5\epsilon m\}}\,. \tag{3.47}$$

For the first term of (3.47) we apply FKG inequality to get

$$\mathbb{E}\sum_{j\geq 2}|\mathcal{C}_j|^k\mathbf{1}_{\{|\mathcal{C}_1|<1.5\epsilon m\}} \leq \mathbb{E}\sum_{j\geq 1}|\mathcal{C}_j|^k\mathbf{1}_{\{|\mathcal{C}_1|<1.5\epsilon m\}} \leq \mathbb{P}(|\mathcal{C}_1| < 1.5\epsilon m)\mathbb{E}\sum_{j\geq 1}|\mathcal{C}_j|^k.$$

By Theorem 3.3.9, we have

$$\mathbb{P}(|\mathcal{C}_1| < 1.5\epsilon m) \leq Ce^{-c\epsilon^3 m}\,,$$

and so

$$\mathbb{E}\sum_{j\geq 2}|\mathcal{C}_j|^k\mathbf{1}_{\{|\mathcal{C}_1|<1.5\epsilon m\}} \leq Ce^{-c\epsilon^3 m}m\mathbb{E}|\mathcal{C}(v)|^{k-1}\,.$$

By part (i) of the theorem this is at most $C\epsilon^k m^k e^{-c\epsilon^3 m}$ which is $O(m\epsilon^{-2k+3})$ since $\epsilon \geq m^{-1/3}$. This shows the required bound for the first term of (3.47).

To take care of the second term of (3.47) we condition on $\mathcal{C}_1$ and note that the graph remaining is distributed as $G(m-|\mathcal{C}_1|, p)$ conditioned on the event of not having a component larger than $|\mathcal{C}_1|$. But since $|\mathcal{C}_1| \geq 1.5\epsilon m$ this random graph is in the subcritical regime, and the probability of having such a component is smaller than $1/2$ (in fact, it is exponentially small). The required estimate follows by part (i) of Theorem 3.3.12. This finishes the proof of (ii).

The proof of (iii) goes in similar lines of (ii). We have

$$\mathbb{E}(\sum_{j\geq 2}|\mathcal{C}_j|^2)^2 = \mathbb{E}(\sum_{j\geq 2}|\mathcal{C}_j|^2)^2\mathbf{1}_{\{|\mathcal{C}_1|<1.5\epsilon m\}} + \mathbb{E}(\sum_{j\geq 2}|\mathcal{C}_j|^2)^2\mathbf{1}_{\{|\mathcal{C}_1|\geq 1.5\epsilon m\}}\,.$$

As in the proof of (ii), to control the first term we use FKG inequality, extract $\mathbf{P}(|\mathcal{C}_1| < 1.5\epsilon m)$ and bound the rest by $\mathbb{E}(\sum_{j\geq 1}|\mathcal{C}_j|^2)^2$ (instead of $j \geq 2$). The analysis performed in the proof of part (ii) of Theorem 3.3.12 shows that $\mathbb{E}(\sum_{j\geq 1}|\mathcal{C}_j|^2)^2$ is controlled by $(\mathbb{E}\sum_{j\geq 1}|\mathcal{C}_j|^2)^2$. We get that

$$\mathbb{E}(\sum_{j\geq 2}|\mathcal{C}_j|^2)^2\mathbf{1}_{\{|\mathcal{C}_1|<1.5\epsilon m\}} \leq Ce^{-c\epsilon^3 m}m^4\epsilon^4 = O(m^2\epsilon^{-2})\,.$$

To control the second term, as in the proof of (ii), we condition on $\mathcal{C}_1$ and use part (ii) of Theorem 3.3.12 to estimate the remaining subcritical graph. This is done identically to part

(ii) and we omit the details. $\square$

Proof of Theorem 3.3.14. We will use a second moment method. First we show that

$$\mathbb{E} \sum_{|\mathcal{C}_j| \le \delta\sqrt{m}} |\mathcal{C}_j|^2 \ge cm^{5/4} \,,$$

for some $c = c(\delta) > 0$. Indeed, we have

$$\mathbb{E}|\mathcal{C}(v)|\mathbf{1}_{\{|\mathcal{C}(v)|\le\delta\sqrt{m}\}} \ge \mathbb{E}|\mathcal{C}(v)|\mathbf{1}_{\{\frac{\delta}{2}\sqrt{m}\le|\mathcal{C}(v)|\le\delta\sqrt{m}\}} \ge \frac{\delta}{2}\sqrt{m}\mathbb{P}(\frac{\delta}{2}\sqrt{m} \le |\mathcal{C}(v)| \le \delta\sqrt{m}). \tag{3.48}$$

We proceed further by restricting to the case that $\mathcal{C}_v$ is tree. Indeed, we have

$$\begin{aligned} \mathbb{P}(\frac{\delta}{2}\sqrt{m} \le |\mathcal{C}(v)| \le \delta\sqrt{m}) &\ge \sum_{k=\delta/2\sqrt{m}}^{\delta\sqrt{m}} \mathbb{P}(|\mathcal{C}(v)| = k, \mathcal{C}(v) \text{ is a tree}) \\ &= \sum_{k=\delta/2\sqrt{m}}^{\delta\sqrt{m}} \binom{m-1}{k-1} k^{k-2} p^{k-1} (1-p)^{k(m-k)+\binom{k}{2}-(k-1)}. \end{aligned}$$

A quick calculation using Stirling's formula gives that for all such k, each summand is of order $\Theta(m^{-4/3})$ and so the probability is of order at least $m^{-1/4}$ and the expectation in (3.48) is of order at least $m^{1/4}$. This gives the first moment estimate since

$$\mathbb{E} \sum_{|\mathcal{C}_j| \le \delta\sqrt{m}} |\mathcal{C}_j|^2 = \mathbb{E} \sum_{v:|\mathcal{C}(v)|\le\delta\sqrt{m}} |C(v)| \,.$$

We continue with the second moment estimate. By Theorem 3.3.12 the second moment satisfies

$$\mathbb{E}\big[\sum_j |\mathcal{C}_j|^2\big]^2 = O(m^{5/2}) \,,$$

and so the assertion of the Theorem follows by the inequality (see [11])

$$\mathbf{P}(V > a) \ge \frac{(\mathbb{E}V - a)^2}{\mathbb{E}V^2} \,,$$

valid for any non-negative random variable V and $a < \mathbb{E}V$. $\square$

Now we turn to the proof of Theorem 3.3.15. Recall that Z_t counts the number of record minima of $\{Y_s\}$ before time t.

Lemma 3.3.26. *For any fixed $\delta \in (0, 1/10)$, there exists an universal constant $C > 0$ such that as long as $\epsilon^3 m \ge 1$ we have*

$$\mathbb{E}Z_{\delta\epsilon m} \le \frac{1}{2[(1-5\delta-5\delta\epsilon)\epsilon]} + O(1) \,,$$

and

$$\mathbb{E}Z^2_{\delta\epsilon m} = O(\epsilon^{-2}) \,.$$

Proof. Define the stopping time τ by

$$\tau = \min\{t : N_t \leq m(1 - 5\delta\epsilon)\}\,,$$

and $\{W_t\}$ to be the process with increments distributed as $\mathrm{Bin}(m(1-5\delta\epsilon), p)$ and $W_0 = 1$. As usual we can couple such that $Y_{t\wedge\tau} \geq W_{t\wedge\tau}$ and that the increments of the first process are always larger than of the second. This guarantees that the number of record minimum of $Y_{t\wedge\tau}$ is bounded from above by the record minimum of $W_{t\wedge\tau}$. Denote by Z^w the number of record minima of the process $\{W_t\}$, then by the above discussion we have

$$\mathbb{E}Z_{\delta\epsilon m} \leq \delta\epsilon m \mathbb{E}\mathbf{1}_{\{\tau < \delta\epsilon m\}} + \mathbb{E}Z^w\,.$$

The order of the first term can be arbitrarily small since $\mathbb{P}(\tau < \delta\epsilon m)$ is exponentially small in ϵm by Lemma 3.3.2. Lemma 3.3.17 bound the second term by the required amount. This concludes the bound on $\mathbb{E}Z_{\delta\epsilon m}$. For the second moment estimate, note that by the same argument, we have

$$\mathbb{E}Z_{\delta\epsilon m}^2 \leq \delta^2\epsilon^2 m^2 \mathbb{E}\mathbf{1}_{\{\tau < \delta\epsilon m\}} + \mathbb{E}(Z^w)^2\,,$$

and the exponential decay of $\mathbb{P}(\tau < \delta\epsilon m)$ and Lemma 3.3.17 concludes the proof. □

Lemma 3.3.27. *For any fixed $\delta \in (0, 1/10)$ denote by τ_δ the stopping time*

$$\tau_\delta = \min_{t \geq \delta\epsilon m} \Big\{ t \text{ is a record minimum of } Y_t \Big\} - \delta\epsilon m\,.$$

Then

$$\mathbb{E}\tau_\delta \leq (2-\delta)\epsilon m - \frac{1}{4\epsilon^2}\,.$$

Proof. Define the process $\{X_t\}$ by $X_t = Y_{\delta\epsilon m + t}$ so that

$$\tau_\delta = \min\{t \geq 0 \,:\, X_t = -Z_{\delta\epsilon m}\}\,.$$

Let $\{W_t\}$ be a process defined by $W_0 = X_0$ and with independent increments distributed as $\mathrm{Bin}(m - t - \delta\epsilon m, p) - 1$ and let τ denote the stopping time $\min\limits_t \{W_t = -Z_{\delta\epsilon m}\}$. As usual, X_t can be stochastically bounded above by W_t and hence $\mathbb{E}\tau_\delta \leq \mathbb{E}\tau$ and we are left to estimate $\mathbb{E}\tau$. We have

$$\mathbb{E}[W_t - W_{t-1} \mid \mathcal{F}_{t-1}] = (1-\delta)\epsilon - \frac{t(1+\epsilon)}{m} - \delta\epsilon^2\,. \tag{3.49}$$

Put

$$\begin{aligned} f(t) &= \frac{t^2}{2m} - (1-\delta)\epsilon t - (\delta - \delta^2/2)\epsilon^2 m - \delta\epsilon^2 t + \frac{t(1+\epsilon) + \epsilon t^2}{2m} \\ &= \frac{[t - (2-\delta)\epsilon m]^2}{2m} + \epsilon[t - (2-\delta)\epsilon m] - \delta\epsilon^2 t + \frac{t(1+\epsilon)+\epsilon t^2}{2m}\,, \end{aligned}$$

then by (3.49) we deduce that $M_t = W_t + f(t)$ is a martingale. A direct calculation with (3.11) gives that

$$\mathbb{E}W_0 = \mathbb{E}Y_{\delta\epsilon m} \le -\delta\epsilon m + \delta\epsilon^2 m - \delta^2\epsilon^2 m/2 + O(\epsilon^3 m)\,,$$

and so we deduce that $\mathbb{E}M_0 \le C\epsilon^3 m$. Furthermore, we have that $\mathbb{E}\tau = O(\epsilon m)$ since after time $2\epsilon m$ the process becomes subcritical with drift $-\epsilon$. Put $\bar\tau = \tau - (2-\delta)\epsilon m$, then by the above and optional stopping if follows that

$$\frac{\mathbb{E}\bar\tau^2}{2m} + \epsilon\mathbb{E}\bar\tau - \mathbb{E}Z_{\delta\epsilon m} \le C\epsilon^3 m\,.$$

This and Lemma 3.3.26 gives that

$$\mathbb{E}\bar\tau \le \frac{1}{2[(1-5\delta-5\delta\epsilon)\epsilon]} - \frac{\mathbb{E}\bar\tau^2}{2\epsilon m} + O(\epsilon^2 m)\,. \tag{3.50}$$

Next, we wish to derive a lower bound on $\mathbb{E}\bar\tau^2$. Put $T = \delta m$, then for $t \le T$ we have that

$$\mathbb{E}\Big[(M_t - M_{t-1})^2\Big] \ge 1-\delta\,,$$

hence the process

$$M_{t\wedge T}^2 - (1-\delta)(t\wedge T)\,,$$

is a submartingale and optional stopping gives

$$(1-\delta)\mathbb{E}[\tau\wedge T] \le \mathbb{E}M_{\tau\wedge T}^2\,. \tag{3.51}$$

We now bound $\mathbb{E}M_{\tau\wedge T}^2$ from above. We have

$$W_{\tau\wedge T} = -Z_{\delta\epsilon m}\mathbf{1}_{\{\tau\le T\}} + W_T\mathbf{1}_{\{\tau>T\}}\,.$$

Thus,

$$\mathbb{E}W_{\tau\wedge T}^2 \le \mathbb{E}Z_{\delta\epsilon m}^2 + O(m^2)\mathbf{P}(\tau>T)\,.$$

Since after time $\delta m/2$ the process is subcritical with constant negative drift we have that $\mathbf{P}(\tau > T)$ decays exponentially in m. Lemma 3.3.17 now yields that $\mathbb{E}W_{\tau\wedge T}^2 = \mathbb{E}Z_{\delta\epsilon m}^2 + o(1) = O(\epsilon^{-2})$. Next we estimate $\mathbb{E}f^2(\tau\wedge T)$. Write $\mu = (2-\delta)\epsilon m$ and simplify $f(t)$ to get

$$\begin{aligned} f(t) &= \frac{(t-\mu)^2(1-\epsilon)}{2m} + (t-\mu)\Big[\epsilon - \delta\epsilon^2 + \frac{1+\epsilon}{2m} - \frac{\mu\epsilon}{m}\Big] - \delta\epsilon^2\mu + \frac{1+\epsilon}{2m}\mu - \frac{\epsilon}{2m}\mu^2 \\ &= \frac{(t-\mu)^2(1-\epsilon)}{2m} + (t-\mu)\Big[\epsilon + O(\epsilon^2)\Big] + O(\epsilon^3 m)\,. \end{aligned}$$

Hence

$$f^2(t) = \frac{(t-\mu)^4(1-\epsilon)^2}{4m^2} + \frac{(t-\mu)^3(\epsilon+O(\epsilon^2))}{m} + (t-\mu)^2\epsilon^2(1+O(\epsilon)) + (t-\mu)O(\epsilon^4 m).$$

Lemmas 3.3.23 and 3.3.25 imply that $\mathbb{E}\bar{\tau}^k$ is of order $(m/\epsilon)^{k/2}$ and hence the third term on the right hand side is dominant so,

$$\mathbb{E}f^2(\tau \wedge T) = (1+o(1))\epsilon^2 \mathbb{E}\Big[(\tau \wedge T - \mu)^2\Big] . \tag{3.52}$$

We also use Cauchy-Schwartz to estimate

$$\Big|\mathbb{E}W_{\tau\wedge T} f(\tau \wedge T)\Big| \le \sqrt{\mathbb{E}W^2_{\tau\wedge T}}\sqrt{\mathbb{E}f^2(\tau\wedge T)} = O(\sqrt{m/\epsilon}) = o(\mathbb{E}f^2(\tau\wedge T)) ,$$

since $\mathbb{E}W^2_{\tau\wedge T} = O(\epsilon^{-2})$ and $\sqrt{m/\epsilon} = o(\epsilon m)$. We put this and (3.52) into (3.51) and get that

$$(1+o(1))\epsilon^2\mathbb{E}\Big[(\tau\wedge T-\mu)^2\Big] \ge (1-\delta)\mu - (1-\delta)\mathbb{E}[\tau\wedge T - \mu] = (1+o(1))(1-\delta)\mu ,$$

since $\mathbb{E}\bar{\tau} = O(\epsilon^{-1}m)$ and $\mathbf{P}(\tau > T)$ decays exponentially in m. We learn that

$$\mathbb{E}\bar{\tau}^2 \ge (1-o(1))(1-\delta)(2-\delta)\frac{m}{\epsilon} .$$

Putting this into (3.50) gives that if $\delta > 0$ is chosen small enough (but fixed) and m is large enough

$$\mathbb{E}\bar{\tau} \le -\frac{1}{4\epsilon^2} ,$$

concluding the proof of the lemma. $\square$

Proof of Theorem 3.3.15. Part (i) follows immediately from Lemma 3.3.27 since $|\mathcal{C}_{\delta\epsilon m}| \le \delta\epsilon m + \tau_\delta$. To prove (ii) we proceed as in the proof of Lemma 3.3.10 and write

$$\mathbb{E}\sum_{\mathcal{C}_j\neq\mathcal{C}_{\delta\epsilon m}} |\mathcal{C}_j|^k = \mathbb{E}\sum_{\mathcal{C}_j\neq\mathcal{C}_{\delta\epsilon m}} |\mathcal{C}_j|^k \mathbf{1}_{\{\mathcal{C}_{\delta\epsilon m}\le 1.5\epsilon m\}} + \mathbb{E}\sum_{j\ge 2}|\mathcal{C}_j|^k\mathbf{1}_{\{\mathcal{C}_{\delta\epsilon m}\ge 1.5\epsilon m\}} .$$

Lemma 3.3.23 shows that $\mathbf{P}(\mathcal{C}_{\delta\epsilon m} \le 1.5\epsilon m) \le Ce^{-c\epsilon^3 m}$ and so FKG inequality gives

$$\mathbb{E}\sum_{\mathcal{C}_j\neq\mathcal{C}_{\delta\epsilon m}} |\mathcal{C}_j|^k \mathbf{1}_{\{\mathcal{C}_{\delta\epsilon m}\le 1.5\epsilon m\}} \le Ce^{-c\epsilon^3 m}\mathbb{E}\sum_j |\mathcal{C}_j|^k = O(m\epsilon^{-2k+3}) ,$$

by part (i) of Lemma 3.3.13. The second term is handled as in the proof of Lemma 3.3.10 by conditioning on $\mathcal{C}_{\delta\epsilon m}$ and using Lemma 3.3.13 for the remaining subcritical graph. $\square$

3.4 *Supercritical case*

In this section we show that the mixing time of the Swendsen-Wang chain is $\Theta(\log n)$ in the supercritical case $c > 2$. This is part (i) of Theorem 3.1.1. Let $\{X_t\}_{t\ge 0}$ be the one dimensional chain defined in (3.2) and write $x_0 = X_0/n$. For $x > \frac{2}{c} - 1$ (so that $\frac{c(1+x)}{2} > 1$), define

$$\Phi(x) = \beta\Big(\frac{c(1+x)}{2}\Big)\frac{1+x}{2}, \tag{3.53}$$

where $\beta(\cdot)$ is defined in (3.16). Since $\beta : \mathbf{R}^+ \to \mathbf{R}$ we have that $\Phi : [-1,\infty] \to \mathbf{R}$. We begin with some preparations for the proof.

Lemma 3.4.1. *For $c > 2$, there exists an unique fixed point $\gamma_0 \in (1 - \frac{2}{c}, 1)$ of $\Phi(x)$. Furthermore, we have*

$$\frac{1}{2} < \Phi'(x) < 1 \qquad \text{for } x > 2c^{-1} - 1\,, \tag{3.54}$$

and there exists a constant $\delta \in (0,1)$ such that for every $x \in [0,1] \setminus \{\gamma_0\}$ we have

$$\tfrac{1}{2} < \frac{\Phi(x) - \gamma_0}{x - \gamma_0} \le \delta. \tag{3.55}$$

Theorem 3.4.2. *There exist constants $\delta \in (0,1)$ and $B > 0$ such that*

$$\mathbb{E}(X_1 - \gamma_0 n)^2 \le \delta(X_0 - \gamma_0 n)^2 + Bn\,. \tag{3.56}$$

Proposition 3.4.3. *We have*

$$\mathbb{E}\Big(X_1 - \Phi(x_0)n \ \mid X_0 \in [\gamma_0 n, n]\Big)^2 = O(n)\,.$$

Proposition 3.4.4. *If X is distributed as the stationary distribution of the magnetization Swendsen-Wang chain, then*

$$\mathbb{E}(X - \gamma_0 n)^2 = O(n).$$

Theorem 3.4.5. *Suppose X_0, Y_0 are two magnetization Swendsen-Wang chains such that $X_0, Y_0 \in [\gamma_0 n - A\sqrt{n}, \gamma_0 n + A\sqrt{n}]$ where A is a constant, we can couple X_1 and Y_1 such that $X_1 = Y_1$ with probability $\Omega(1)$ (which may depend on A).*

Proof of part (i) of Theorem 3.1.1: Rearranging Theorem 3.4.2 and taking expectations gives

$$\mathbb{E}(X_{t+1} - \gamma_0 n)^2 - \frac{B}{1-\delta}n \le \delta\Big[\mathbb{E}(X_t - \gamma_0 n)^2 - \frac{B}{1-\delta}n\Big]$$

for all t. We apply this inductively and get

$$\mathbb{E}(X_{C\log n} - \gamma_0 n)^2 - \frac{B}{1-\delta}n \le \delta^{C\log n}\Big[\mathbb{E}(X_0 - \gamma_0 n)^2 - \frac{B}{1-\delta}n\Big]\,.$$

Hence, when $C = C(\delta)$ is large enough we get that

$$\mathbb{E}(X_{C\log n} - \gamma_0 n)^2 = O(n)\,,$$

and so Markov's inequality gives

$$\mathbb{P}(|X_{C\log n} - \gamma_0 n| \le A\sqrt{n}) \ge \frac{3}{4} \tag{3.57}$$

for some large constant A. Let X'_t be a magnetization SW chain starting at stationarity. By Theorem 3.4.4 and Markov's inequality we have

$$\mathbb{P}(|X'_{C\log n} - \gamma_0 n| \le A\sqrt{n}) \ge \frac{3}{4} \tag{3.58}$$

for some large constant A. Now, to couple X_t and X'_t we first run them independently until time $C\log n$. By (3.57) and (3.58), we have that $X_{C\log n}, X'_{C\log n} \in [\gamma_0 n - A\sqrt{n}, \gamma_0 n + A\sqrt{n}]$ with probability at least 1/2. By Theorem 3.4.5, we can couple $X_{C\log n+1}$ and $X'_{C\log n+1}$ such that $X_{C\log n+1} = X'_{C\log n+1}$ with probability $\Omega(1)$. Then by Lemma 3.2.1, we have that $\{\sigma_t\}$ and $\{\sigma'_t\}$ can be coupled such that $\sigma_t = \sigma'_t$ in $O(\log n)$ steps with probability $\Omega(1)$. The upper bound of mixing time follows from Lemma 3.1.4.

For the lower bound, we will show that if $X_0 = n$, then

$$\|X_{\alpha\log n} - X_\pi\|_{TV} \geq 1/4$$

for some small constant $\alpha > 0$, where X_π is the stationary distribution of the magnetization Swendsen-Wang chain. By (3.55), we have that

$$\begin{aligned}\mathbb{P}\Big(X_{t+1} - \gamma_0 n \leq \frac{1}{4}(X_t - \gamma_0 n)\Big) &\leq \mathbb{P}\Big(X_{t+1} - \gamma_0 n \leq \frac{1}{2}(\Phi(X_t/n)n - \gamma_0 n)\Big) \\ &= \mathbb{P}\Big(X_{t+1} - \Phi(X_t/n)n \leq \frac{1}{2}(\gamma_0 n - \Phi(X_t/n)n)\Big).\end{aligned}$$

When $X_t \geq \gamma_0 n$ we have that $\Phi(X_t/n)n \geq \gamma_0 n$ by (3.55), hence Proposition 3.4.3 and Markov's inequality imply that

$$\mathbb{P}\Big(X_{t+1} - \gamma_0 n \leq \frac{1}{4}(X_t - \gamma_0 n) \mid X_t \geq \gamma_0 n\Big) \leq \frac{O(n)}{(\Phi(X_t/n)n - \gamma_0 n)^2}. \tag{3.59}$$

Furthermore, if $X_t - \gamma_0 n \geq n^{\frac{3}{4}}$, then $\Phi(X_t/n)n - \gamma_0 n \geq n^{\frac{3}{4}}/2$ by (3.55). Plugging this into (3.59) gives

$$\mathbb{P}\Big(X_{t+1} - \gamma_0 n \geq \frac{1}{4}(X_t - \gamma_0 n)\big|X_t - \gamma_0 n \geq n^{\frac{3}{4}}\Big) \geq 1 - O(n^{-\frac{1}{2}}). \tag{3.60}$$

Starting from $X_0 = n$, by applying (3.60) iteratively we have

$$\mathbb{P}\big(X_{\alpha\log n} - \gamma_0 n \geq n^{\frac{3}{4}}\big) \geq (1 - O(n^{-\frac{1}{2}}))^{\alpha\log n} = 1 - o(1), \tag{3.61}$$

when $\alpha > 0$ is small enough constant. On the other hand, by Proposition 3.4.4 and the Markov's inequality, we have $\mathbb{P}\Big(|X_\pi - \gamma_0 n| \geq A\sqrt{n}\Big) \leq \frac{1}{4}$ for some constant A. Putting the two inequalities together, we get

$$\|X_{\alpha\log_4 n} - X_\pi\|_{TV} \geq \frac{3}{4} - o(1) \geq \frac{1}{4}, \tag{3.62}$$

which gives a lower bound on the mixing time of magnetization SW chain X_t. This concludes the proof since any lower bound of the mixing time of X_t implies the same lower bound of mixing time of σ_t. $\square$

Proof of Lemma 3.4.1: By the definition of $\beta(\cdot)$ in equation (3.16), we know $\Phi(x)$ is the positive solution of

$$1 - e^{-c\Phi(x)} = \frac{2\Phi(x)}{x+1} \tag{3.63}$$

for all $x > \frac{2}{c} - 1$. Taking derivative of both sides yields

$$ce^{-c\Phi}\Phi' = \frac{2(x+1)\Phi' - 2\Phi}{(x+1)^2}.$$

By plugging in $x + 1 = \frac{2\Phi}{1-e^{-c\Phi}}$ we get

$$\Phi' = \frac{1 - 2e^{-c\Phi} + e^{-2c\Phi}}{2(1 - e^{-c\Phi} - c\Phi e^{-c\Phi})}. \tag{3.64}$$

By (3.64), we have that $\frac{1}{2} < \Phi'$ if and only if $e^{-c\Phi} > 1 - c\Phi$ which is true for all $c\Phi > 0$. We also have that

$$\Phi' < 1 \Longleftrightarrow c\Phi < \sinh(c\Phi) \tag{3.65}$$

which holds for all $c\Phi > 0$.

Since $c > 2$ (which implies $\frac{2}{c} - 1 < 0$), we have that $\Phi' < 1$ for all $x \in [0, 1]$. Since Φ' is continuous, we have a constant $\delta_1 \in (0, 1)$ such that $\Phi'(x) < \delta_1$ for all $x \in [0, 1]$. Note that $\Phi(0) = \frac{1}{2}\beta(\frac{c}{2}) > 0$, $\Phi(1) = \beta(c) < 1$ and Φ is strictly increasing in $[0, 1]$, we have by Rolle's theorem that there exists an unique point $\gamma_0 \in (0, 1)$ such that $\Phi(\gamma_0) = \gamma_0$. By plugging in $x = 1 - 2/c$ into (3.53) and the definition of $\beta(\cdot)$, we get

$$\begin{aligned} \Phi(1 - 2/c) > 1 - 2/c \quad &\Leftrightarrow \quad \beta(c-1) > 1 - \frac{1}{c-1} \\ &\Leftrightarrow \quad e^{-(c-2)} < \frac{1}{c-1} \end{aligned}$$

which is always true for $c > 2$. It follows immediately that $\gamma_0 > 1 - \frac{2}{c}$. □

Recall that given $X_0 = x_0 n$, we have that X_1 is distributed as in (3.3). To prove Theorem 3.4.2 we first state a useful lemma.

Lemma 3.4.6. *Let $c > 2$.*

(i) There exists a non-negative function $h(\cdot)$ with $h(\epsilon) \to 0$ as $\epsilon \to 0$ such that if $|x_0 - (1 - \frac{2}{c})| \le \epsilon$, then $\mathbb{E}\Big(\sum_{j\ge1} |\mathcal{C}_j^-|^2\Big) \le h(\epsilon)n^2$.

(ii) For any fixed $\epsilon > 0$, if $x_0 \in [0, (1 - \frac{2}{c} - \epsilon)]$, then $\mathbb{E}\Big(\sum_{j\ge1} |\mathcal{C}_j^-|^2\Big) \le (\Phi^2(-x_0) + o(1))n^2$.

(iii) For any fixed $\epsilon > 0$, if $x_0 \in [1 - \frac{2}{c} + \epsilon, 1]$, then $\mathbb{E}\Big(\sum_{j\ge1} |\mathcal{C}_j^-|^2\Big) \le O(n)$.

Proof of Theorem 3.4.2: By (3.3), we have

$$\mathbb{E}X_1^2 = \mathbb{E}\Big(\sum_{j\ge1} |\mathcal{C}_j^+|^2\Big) + \mathbb{E}\Big(\sum_{j\ge1} |\mathcal{C}_j^-|^2\Big) \tag{3.66}$$

and

$$\begin{aligned}\mathbb{E}X_1 &= \mathbb{E}\Big|\sum_{j\geq 1}\epsilon_j|\mathcal{C}_j^+| + \sum_{j\geq 1}\epsilon_j'|\mathcal{C}_j^-|\Big| = \mathbb{E}\Big|\epsilon_1\big(\sum_{j\geq 1}\epsilon_j|\mathcal{C}_j^+| + \sum_{j\geq 1}\epsilon_j'|\mathcal{C}_j^-|\big)\Big| \\ &\geq \mathbb{E}\Big(\epsilon_1\big(\sum_{j\geq 1}\epsilon_j|\mathcal{C}_j^+| + \sum_{j\geq 1}\epsilon_j'|\mathcal{C}_j^-|\big)\Big) \\ &= \mathbb{E}\Big[|\mathcal{C}_1^+| + \epsilon_1\Big(\sum_{j\geq 2}\epsilon_j|\mathcal{C}_j^+| + \sum_{j\geq 1}\epsilon_j'|\mathcal{C}_j^-|\Big)\Big] = \mathbb{E}|\mathcal{C}_1^+|. \end{aligned} \tag{3.67}$$

Combining (3.66) and (3.67), we get

$$\mathbb{E}\Big(X_1 - \gamma_0 n\Big)^2 \leq \mathbb{E}\Big(\sum_{j\geq 1}|\mathcal{C}_j^+|^2\Big) + \mathbb{E}\Big(\sum_{j\geq 1}|\mathcal{C}_j^-|^2\Big) - 2\gamma_0 n\cdot\mathbb{E}|\mathcal{C}_1^+| + \gamma_0^2 n^2. \tag{3.68}$$

The random graph $G(\frac{1+x_0}{2}n, \frac{c}{n})$ is supercritical with $\theta = \frac{1+x_0}{2}n\frac{c}{n} \geq \frac{c}{2} > 1$. By Corollary 3.3.6 we have

$$\mathbb{E}\Big(\sum_{j\geq 1}|\mathcal{C}_j^+|^2\Big) \leq \Big(\mathbb{E}|\mathcal{C}_1^+|\Big)^2 + O(n). \tag{3.69}$$

Plugging (3.69) into (3.68), we get

$$\mathbb{E}\Big(X_1 - \gamma_0 n\Big)^2 \leq \Big(\mathbb{E}|\mathcal{C}_1^+| - \gamma_0 n\Big)^2 + \mathbb{E}\Big(\sum_{j\geq 1}|\mathcal{C}_j^-|^2\Big) + O(n). \tag{3.70}$$

By Corollary 3.3.5, we have $\Big|\mathbb{E}|\mathcal{C}_1^+| - \Phi(x_0)n\Big| \leq O(\sqrt{n})$. Thus,

$$\begin{aligned}\Big(\mathbb{E}|\mathcal{C}_1^+| - \gamma_0 n\Big)^2 &\leq \Big|\mathbb{E}|\mathcal{C}_1^+| - \Phi(x_0)n\Big|^2 + \Big|\Phi(x_0)n - \gamma_0 n\Big|^2 \\ &+ 2\Big|\mathbb{E}|\mathcal{C}_1^+| - \Phi(x_0)n\Big|\Big|\Phi(x_0)n - \gamma_0 n\Big| \\ &\leq \Big|\Phi(x_0)n - \gamma_0 n\Big|^2 + O(\sqrt{n})\Big|\Phi(x_0)n - \gamma_0 n\Big| + O(n). \end{aligned} \tag{3.71}$$

Applying Lemma 3.4.1 gives that

$$\Big(\mathbb{E}|\mathcal{C}_1^+| - \gamma_0 n\Big)^2 \leq \delta_1^2|x_0 - \gamma_0|^2 n^2 + |x_0 - \gamma_0|O(n^{3/2}) + O(n). \tag{3.72}$$

If $|x_0 - \gamma_0| = O(n^{-\frac{1}{2}})$, then $|x_0 - \gamma_0|n^{3/2} = O(n)$. If $|x_0 - \gamma_0|n^{\frac{1}{2}} \to \infty$, we have $|x_0 - \gamma_0|O(n^{3/2}) = o(|x_0 - \gamma_0|^2 n^2)$. Plugging these back into (3.72), we get

$$\mathbb{E}(X_1 - \gamma_0 n)^2 \leq (\delta_1^2 + o(1))|x_0 - \gamma_0|^2 n^2 + O(n) + \mathbb{E}\Big(\sum_{j\geq 1}|\mathcal{C}_j^-|^2\Big). \tag{3.73}$$

To estimate $\mathbb{E}\Big(\sum_{j\geq 1}|\mathcal{C}_j^-|^2\Big)$, choose a small constant ϵ such that $\delta_1^2 + h(\epsilon) < 1$ where $h(\cdot)$ is defined in part (i) of Lemma 3.4.6. If $\Big|x_0 - (1 - \frac{2}{c})\Big| < \epsilon$, we have that

$$\mathbb{E}(X_1 - \gamma_0 n)^2 \leq (\delta_1^2 + h(\epsilon))|x_0 - \gamma_0|^2 n^2 + O(n) \tag{3.74}$$

by plugging part (i) of Lemma 3.4.6 into (3.73).

If $x_0 \in [0, (1 - \frac{2}{c} - \epsilon)]$, we have that $|\Phi(x_0) - \gamma_0|$ is uniformly bounded from below by Lemma 3.4.1. As a result, we have that $\Big(\mathbb{E}|\mathcal{C}_1^+| - \gamma_0 n\Big)^2 \leq (\Phi(x_0) - \gamma_0)^2 n^2 + O(n)$ in (3.71). Plugging this and part (ii) of Lemma 3.4.6 into (3.71) gives

$$\mathbb{E}(X_1 - \gamma_0 n)^2 \leq ((\Phi(x_0) - \gamma_0)^2 + \Phi^2(-x_0) + o(1))n^2 + O(n). \tag{3.75}$$

By Lemma 3.4.1 and Rolle's Theorem, we have

$$\frac{\Phi(x_0) - \Phi(-x_0)}{2x_0} \geq \frac{1}{2},$$

which leads to $\Phi^2(-x_0) \leq \Big(\Phi(x_0) - x_0\Big)^2$. This gives

$$\frac{\Big(\Phi(x_0) - \gamma_0\Big)^2 + \Phi^2(-x_0)}{(x_0 - \gamma_0)^2} \leq \frac{\Big(\Phi(x_0) - \gamma_0\Big)^2 + \Big(\Phi(x_0) - x_0\Big)^2}{(x_0 - \gamma_0)^2} < 1, \tag{3.76}$$

since $x_0 < \Phi(x_0) < \gamma_0$. The left hand side of (3.76) is smaller than 1 for all $x_0 \in [0, (1 - \frac{2}{c} - \beta)]$, so it is smaller than some constant $\delta_2 < 1$ uniformly. Plugging this into (3.75), we get

$$\mathbb{E}(X_1 - \gamma_0 n)^2 \leq \delta_2 (x_0 - \gamma_0)^2 n^2 + O(n). \tag{3.77}$$

If $x_0 \in [1 - \frac{2}{c} + \epsilon, 1]$, we plug (iii) of Lemma 3.4.6 into (3.73) and obtain

$$\mathbb{E}(X_1 - \gamma_0 n)^2 \leq (\delta_1^2 + o(1))|x_0 - \gamma_0|^2 n^2 + O(n). \tag{3.78}$$

Combining (3.74),(3.77) and (3.78) concludes our proof. □

Proof of Lemma 3.4.6. We begin with case (ii). In this regime, the random graph $G(\frac{1-x_0}{2}n, \frac{c}{n})$ is supercritical with $\theta > 1 + \frac{c\epsilon}{2}$. In the same way we obtained (3.69) we also have

$$\mathbb{E}\Big(\sum_{j\geq 1} |\mathcal{C}_j^-|^2\Big) \leq \Big(\mathbb{E}|\mathcal{C}_1^-|\Big)^2 + O(n). \tag{3.79}$$

By Corollary 3.3.5 we have that $|\mathbb{E}|\mathcal{C}_1^-| - \Phi(-x_0)n| \leq O(\sqrt{n})$ showing that

$$\begin{aligned}
\mathbb{E}\Big(\sum_{j\geq 1} |\mathcal{C}_j^-|^2\Big) &\leq \Big(\Phi(-x_0)n + O(\sqrt{n})\Big)^2 + O(n) \\
&\leq \Phi^2(-x_0)n^2 + O(n) + \Phi(-x_0)O(n^{3/2}) \\
&= \big(\Phi^2(-x_0) + o(1)\big)n^2,
\end{aligned} \tag{3.80}$$

since $|\Phi(x_0)|$ is uniformly bounded from below, as required.

We now prove case (i). Note that we have

$$\mathbb{E}\Big(\sum_{j\geq 1} |\mathcal{C}_j^-|^2\Big) = \Big(\frac{1 - x_0}{2}n\Big)\mathbb{E}|C_v|$$

as in (3.15). Since $\mathbb{E}|C_v|$ is decreasing in x_0, we have that $\mathbb{E}\Big(\sum_{j\geq 1}|\mathcal{C}_j^-|^2\Big)$ reaches its maximum at $x_0 = 1-\frac{2}{c}-\epsilon$. Plugging in this value into (3.80) gives

$$\mathbb{E}\Big(\sum_{j\geq 1}|\mathcal{C}_j^-|^2\Big) \leq \big(\beta^2(1+\frac{\epsilon c}{2})+o(1)\big)n^2 \tag{3.81}$$

Note that $\beta(x)\to 0$ as $x\to 1$, so we can take $h(x)=\beta^2(1+\frac{cx}{2})+o(1)$.

To prove case (iii) note that $\frac{c(1-x_0)}{2}\leq 1-\frac{\epsilon c}{2}$, so the random graph $G(\frac{1-x_0}{2}n,\frac{c}{n})$ is subcritical in this regime with θ bounded from above away from 1. Applying Lemma 3.3.3, we get

$$\mathbb{E}\Big[\sum_{j\geq 1}|\mathcal{C}_j^-|^2\Big] = O(n). \tag{3.82}$$

□

Proof of Proposition 3.4.3: Note that (3.70) is valid for all $\gamma_0\in[0,1]$ and in particular for $\Phi(x_0)$. Thus,

$$\mathbb{E}\Big(X_1-\Phi(x_0)n\Big)^2 \leq \Big(\mathbb{E}|\mathcal{C}_1^+|-\Phi(x_0)n\Big)^2 + \mathbb{E}\Big(\sum_{j\geq 1}|\mathcal{C}_j^-|^2\Big)+O(n). \tag{3.83}$$

Recall that $x_0\geq\gamma_0>1-\frac{2}{c}$. By Corollary 3.3.5 we have that $\big(\mathbb{E}|\mathcal{C}_1^+|-\Phi(x_0)n\big)^2=O(n)$. The random graph of $G(\frac{1-x_0}{2}n,\frac{c}{n})$ is in regime (iii) of Lemma 3.4.6. Plugging (3.82) into (3.83), we get

$$\mathbb{E}\Big(X_1-\Phi(x_0)n\Big)^2 = O(n)\,, \tag{3.84}$$

as required. □

Proof of Proposition 3.4.4: If X_0 follows the stationary distribution of the magnetization SW chain, so does X_1. Taking expectation of both sides of (3.56) gives

$$\mathbb{E}(X_1-\gamma_0 n)^2 \leq \delta\mathbb{E}(X_0-\gamma_0 n)^2+Bn\,,$$

as required. □

To prove Theorem 3.4.5 we need the following lemma.

Lemma 3.4.7. *Let Y and Z be two random variables distributed as the sum of n independent random $\pi_{(m)}$ signs. Then for any fixed constant a, there exists a constant $\kappa(a)\in(0,1]$ such that for any $-a\sqrt{n}\leq y\leq a\sqrt{n}$, we can couple Y and Z such that $Y-y=Z$ with probability at least κ.*

Proof. Direct corollary of the local central limit theorem of simple random walk. □

Proof of Theorem 3.4.5. To couple X_1 and Y_1, we first apply the percolation step of the Swendsen-Wang dynamics in both chains independently. By Lemma 3.3.7, with probability

$1-O(\frac{1}{n})$, the number of isolated points after percolation is bigger than $\frac{1}{3e^c}n$ in both chains. Conditioned on this, we assign each component a $\pi_{(m)}$ spin using the following procedure.

First assign the spins of components independently in descending order of their size until there are $\frac{1}{3e^c}n$ components left. Note the remaining components are all isolated vertices. Denote by $\bar{X}_1$ and $\bar{Y}_1$ as the absolute value of the sum of spins at this time respectively.

Note that (3.67) and (3.68) are still valid if we replace X_1 by $\bar{X}_1$. Consequently, Theorem 3.4.2 is also valid if replacing X_1 by $\bar{X}_1$. Hence, since $|X_0 - \gamma_0 n| \le A\sqrt{n}$ we have

$$\mathbb{E}(\bar{X}_1 - \gamma_0 n)^2 = O(n)\,.$$

By Markov's inequality, there exists a constant A_1 such that

$$\mathbb{P}\Big(|\bar{X}_1 - \gamma_0 n| \ge A_1\sqrt{n}\Big) \le \frac{1}{4}\,, \tag{3.85}$$

and similarly

$$\mathbb{P}\Big(|\bar{Y}_1 - \gamma_0 n| \ge A_1\sqrt{n}\Big) \le \frac{1}{4}\,. \tag{3.86}$$

Consider the event

$$\mathcal{A} := \{|\bar{X}_1 - \gamma_0 n| < A_1\sqrt{n}\} \cap \{|\bar{Y}_1 - \gamma_0 n| < A_1\sqrt{n}\} \cap \{\text{There are at least } \tfrac{n}{3e^c} \text{ isolated vertices}\}\,.$$

By (3.85) and (3.86) we have that $\mathbb{P}(\mathcal{A}) \ge \frac{1}{4}$.

Conditioned on $\mathcal{A}$, we have $|\bar{X}_1 - \bar{Y}_1| \le 2A_1\sqrt{n}$. Denote by $\hat{X}_1$ and $\hat{Y}_1$ the sum of spins of the rest of the components (all of them being isolated vertices) of the two chains respectively. Note $\hat{X}_1$ and $\hat{Y}_1$ are i.i.d. sums of $\pi_{(m)}$ spins. By Lemma 3.4.7 we can couple $\hat{X}_1$ and $\hat{Y}_1$ so that $\hat{X}_1 + \bar{X}_1 = \hat{Y}_1 + \bar{Y}_1$ with probability $\Omega(1)$. Finally, notice that $X_1 \overset{(d)}{=} |\bar{X}_1 + \hat{X}_1|$ and $Y_1 \overset{(d)}{=} |\bar{Y}_1 + \hat{Y}_1|$, concluding the proof. $\square$

3.5 Subcritical case

In this section, we prove that in the subcritical case $c < 2$, the mixing time of the Swendsen-Wang chain is $\Theta(1)$. This is part (iii) of Theorem 3.1.1.

Lemma 3.5.1. *For $c \in (1,2)$ there exists a constant $\delta \in (0,1)$ such that for all $x \in [\frac{2}{c}-1, 1]$, we have*

$$\frac{\Phi(x)}{x} \le \delta \tag{3.87}$$

where $\Phi(\cdot)$ is defined in (3.53).

Theorem 3.5.2. *There exist two constants $\delta \in (0,1)$ and $B > 0$ such that*

$$\mathbb{E}\big(X_1^2 \mid X_0\big) \le \delta X_0^2 + Bn\,. \tag{3.88}$$

Moreover, if $0 \le x_0 \le \frac{1}{c} - \frac{1}{2}$, we have

$$\mathbb{E}X_1^2 \le Bn. \tag{3.89}$$

To get the constant upper bound of mixing time we need to consider the following two-dimensional chain. Let G_1 be a fixed subset of the vertices and G_2 its complement. Let (Y_t, Z_t) be a two-dimensional Markov chain, where Y_t record the number of vertices with positive spin in G_1 and Z_t record the number vertices with positive spin in G_2.

Proposition 3.5.3. *Let* (Y_t, Z_t) *and* $(\widetilde{Y}_t, \widetilde{Z}_t)$ *be two two-dimensional chains as defined above. Suppose* $Y_0 + Z_0$ *and* $\widetilde{Y}_0 + \widetilde{Z}_0$ *lie in the window* $I = [\frac{n}{2} - A\sqrt{n}, \frac{n}{2} + A\sqrt{n}]$ *where* A *is a constant. Then we can couple* (Y_1, Z_1) *and* $(\widetilde{Y}_1, \widetilde{Z}_1)$ *such that* $(Y_1, Z_1) = (\widetilde{Y}_1, \widetilde{Z}_1)$ *with probability* $\Omega(1)$ *(which may depend on* A*).*

Proof of part (iii) of Theorem 3.1.1: For any starting configuration σ, let G_1 be the vertices with positive spin and G_2 be its complement. Let X_t be the magnetization chain and (Y_t, Z_t) be the two-dimensional chain as described above. As usual $\mathbb{P}$ and π are the transition matrix and the stationary distribution of the Swendsen-Wang chain, respectively, and let $\widetilde{\mathbb{P}}$ and $\widetilde{\pi}$ be the corresponding transition matrix and stationary distribution of (Y_t, Z_t), respectively. By symmetry, configurations with same two-dimensional chain value have same distributions for any t. Consequently

$$||\sigma\mathbb{P}^t - \pi||_{TV} = ||(|G_1|, 0)\widetilde{\mathbb{P}}^t, \widetilde{\pi}||_{TV} \,. \tag{3.90}$$

Thus, by Lemma 3.1.4 it suffices to couple the chains (Y_t, Z_t) and $(\widetilde{Y}_t, \widetilde{Z}_t)$ such that they meet with probability $\Omega(1)$ in time $t = \Theta(1)$. By Lemma 3.5.2, we have

$$\mathbb{E}(X_{t+1}^2) - \frac{B}{1-\delta}n \le \delta\Big[\mathbb{E}(X_t^2) - \frac{B}{1-\delta}n\Big].$$

Applying this inductively we get

$$\mathbb{E}(X_t^2) - \frac{B}{1-\delta}n \le \delta^t \mathbb{E}(X_0^2) \le \delta^t n^2.$$

For $t \ge 2\log_\delta \frac{1}{8}(\frac{1}{c} - \frac{1}{2})$ and large n, we have

$$\mathbb{E}(X_t^2) \le \frac{1}{4}(\frac{1}{c} - \frac{1}{2})^2 n^2.$$

For such t Markov's inequality gives

$$\mathbb{P}\Big(X_t \ge (\frac{1}{c} - \frac{1}{2})n\Big) \le \frac{1}{4}. \tag{3.91}$$

By Theorem 3.5.2 and Markov's inequality, if $X_t \in [0, (\frac{1}{c} - \frac{1}{2})n]$, then $X_{t+1} \in [0, A\sqrt{n}]$ with probability at least $1/2$ for some large constant A. Combining this and (3.91), we have that after constant number of steps, the chain X_t will jump into the window $I = [0, A\sqrt{n}]$ with probability $\Omega(1)$.

For any two Swendsen-Wang chains σ and $\tilde{\sigma}$, Let X_t and $\widetilde{X}_t$ be the corresponding magnetization chains. Running the two Swenden-Wang dynamics independently first, by the argument above, we have that X_t and $\widetilde{X_t}$ both jump into $[0, A\sqrt{n}]$ after constant steps with probability $\Omega(1)$. By Proposition 3.5.3, we can couple the two two-dimensional chains so that $(Y_t, Z_t) = (\widetilde{Y}_t, \widetilde{Z}_t)$ with probability $\Omega(1)$, which concludes the whole proof. $\square$

Proof of Lemma 3.5.1: Note Φ is differentiable on $[\frac{2}{c}-1, 1]$. Recalling (3.65), we have $\Phi' < 1$ for all $x > \frac{2}{c} - 1$. By Rolle's Theorem, we have $\Phi(x) - 0 \le x - (\frac{2}{c} - 1)$ for all $x > \frac{2}{c} - 1$. So

$$\frac{\Phi(x)}{x} \le 1 - \frac{\frac{2}{c}-1}{x} \le 1 - (\frac{2}{c} - 1)$$

for all $x \in [\frac{c}{2} - 1, 1]$. $\square$

Proof of Theorem 3.5.2: We use the fact that (3.66) is still valid. The random graph $G(\frac{1-x_0}{2}n, \frac{c}{n})$ is subcritical with $\theta = (\frac{1-x_0}{2}n)\frac{c}{n} = \frac{c}{2}$. By Lemma 3.3.3, we have

$$\mathbb{E}\Big(\sum_{j\ge 1} |\mathcal{C}_j^-|^2\Big) = O(n). \tag{3.92}$$

If $c < 1$, the random graph $G(\frac{1+x_0}{2}n, \frac{c}{n})$ is subcritical with $\theta = (\frac{1+x_0}{2}n)\frac{c}{n} \le c < 1$. By Lemma 3.3.3, we have

$$\mathbb{E}\Big(\sum_{j\ge 1} |\mathcal{C}_j^+|^2\Big) = O(n). \tag{3.93}$$

If $c \ge 1$, then let $\epsilon > 0$ be a small constant that we will determine later and consider the following three cases.

(i) $0 \le x_0 \le \frac{2}{c} - 1 - \epsilon$. In this case, the random graph $G(\frac{1+x_0}{2}n, \frac{c}{n})$ is subcritical with $\theta \le 1 - \frac{\epsilon c}{2}$. By Lemma 3.3.3,

$$\mathbb{E}\Big(\sum_{j\ge 1} |\mathcal{C}_j^+|^2\Big) = O(n). \tag{3.94}$$

(ii) $\frac{2}{c} - 1 + \epsilon \le x_0 \le 1$(in case $c > 1$). In this case, the random graph $G(\frac{1+x_0}{2}n, \frac{c}{n})$ is supercritical with $\theta \ge 1 + \frac{\epsilon c}{2}$. By Corollary 3.3.6, we have

$$\begin{aligned}
\mathbb{E}\Big(\sum_{j\ge 1} |\mathcal{C}_j^+|^2\Big) &\le \big(\mathbb{E}|\mathcal{C}_1^+|\big)^2 + O(n) \\
&= \Big(\Phi(x_0)n\Big)^2 + \Big(\mathbb{E}|\mathcal{C}_1^+| - \Phi(x_0)n\Big)\Big(\mathbb{E}|\mathcal{C}_1^+| + \Phi(x_0)n\Big) + O(n).
\end{aligned}$$

By Corollary 3.3.5, we have $\Big|\mathbb{E}|\mathcal{C}_1^+| - \Phi(x_0)n\Big| = O(\sqrt{n})$. By Lemma 3.5.1, we have $\Phi(x_0)n \le \delta x_0 n$. So we have

$$\mathbb{E}\Big(\sum_{j\ge 1} |\mathcal{C}_j^+|^2\Big) \le \delta^2 x_0^2 n^2 + O(n^{3/2}) \le (\delta^2 + o(1))x_0^2 n^2.$$

(iii) $\frac{2}{c} - 1 - \epsilon \leq x_0 \leq \frac{2}{c} - 1 + \epsilon$ (or $1 - \epsilon \leq x_0 \leq 1$ in case $c = 1$). Recall that $\mathbb{E}\Big(\sum_{j\geq 1} |\mathcal{C}_j^+|^2\Big) = \frac{1+x_0}{2} n\mathbb{E}|C_v|$. So $\mathbb{E}\Big(\sum_{j\geq 1} |\mathcal{C}_j^+|^2\Big)$ reaches its maximum at $x_0 = \frac{2}{c} - 1 + \epsilon$ for $1 < c < 2$ or $x_0 = 1$ for $c = 1$. In the former case, by the estimate in case (ii), we get

$$\mathbb{E}\Big(\sum_{j\geq 1} |\mathcal{C}_j^+|^2\Big) \leq (\delta^2 + o(1))(\frac{2}{c} - 1 + \epsilon)^2 n^2.$$

Now we choose ϵ to be small enough such that $\delta^2 \Big(\frac{2/c-1+\epsilon}{2/c-1-\epsilon}\Big)^2 < 1$, then we choose a constant δ_1 such that $\delta^2 \Big(\frac{2/c-1+\epsilon}{2/c-1-\epsilon}\Big)^2 < \delta_1 < 1$. Then We have

$$\mathbb{E}\Big(\sum_{j\geq 1} |\mathcal{C}_j^+|^2\Big) \leq \delta x_0^2 n^2.$$

In the latter case, by Theorem 1 of [30], we have that

$$\mathbb{E}\Big(\sum_{j\geq 1} |\mathcal{C}_j^+|^2\Big) = o(n^2).$$

The Lemma follows from combining case (i), (ii) and (iii). □

Proof of Proposition 3.5.3: Suppose without lost of generality that $|G_2| \leq |G_1|$. Since $Y_0 + Z_0 \in I$, the random graphs $G(Y_0 + Z_0, \frac{c}{n})$ and $G(n - (Y_0 + Z_0), \frac{c}{n})$ are both subcritical for large n. The same is true for the chain $(\widetilde{Y}_t, \widetilde{Z}_t)$. In the first chain after the percolation step, denote by $\{(\frac{t}{n} \wedge 1)_j\}_{j\geq 1}$ and $\{\mathcal{B}_j\}_{j\geq 1}$ the components with vertices completely in G_1 and G_2 respectively. Note that there are also components that have vertices in both G_1 and G_2. Denote such components by $\{\mathcal{C}_j\}_{j\geq 1}$. In the second chain, we denote by $\widetilde{(\frac{t}{n} \wedge 1)_j}$,$\widetilde{\mathcal{B}_j}$ and $\widetilde{\mathcal{C}_j}$ to be these components. Lemma 3.3.7 implies that for some $c > 0$ with probability $\Omega(1)$ we have that the number of isolated vertices in $\{A_j\}$ is at least $c|G_1|$ and at least $c|G_2|$ for $\{\widetilde{A_j}\}$. Denote this event by $(\frac{t}{n} \wedge 1)$.

Furthermore, by Lemma 3.3.3 we have

$$\mathbb{E}(\sum_{j\geq 1} |(\frac{t}{n} \wedge 1)_j|^2 + \sum_{j\geq 1} |\mathcal{C}_j \cap G_1|^2) = O(|G_1|)\,, \tag{3.95}$$

$$\mathbb{E}(\sum_{j\geq 1} |\mathcal{B}_j|^2 + \sum_{j\geq 1} |\mathcal{C}_j \cap G_2|^2) = O(|G_2|)\,, \tag{3.96}$$

$$\mathbb{E}(\sum_{j\geq 1} |\tilde{A}_j|^2 + \sum_{j\geq 1} |\tilde{\mathcal{C}}_j \cap G_1|^2) = O(|G_1|)\,, \tag{3.97}$$

$$\mathbb{E}(\sum_{j\geq 1} |\tilde{B}_j|^2 + \sum_{j\geq 1} |\tilde{\mathcal{C}}_j \cap G_2|^2) = O(|G_2|). \tag{3.98}$$

Now, we first assign spins to all components except the isolated vertices in $\{A_j\}$ and $\{\widetilde{A_j}\}$ independently in both chains. Let M_1, N_1 be the sum of spins in G_1 and G_2 respectively in first chain before assigning the rest of the spins, and similarly $\widetilde{M_1}$, $\widetilde{N_1}$ be the same for the second chain at this time. By (3.95),(3.96),(3.97),(3.98) and Markov's inequality that we have

$$\{\mathcal{A}, |M_1 - \widetilde{M_1}| = O(\sqrt{|G_1|}), |N_1 - \widetilde{N_1}| = O(\sqrt{|G_2|})\}$$

occurs with probability $\Omega(1)$. Then by Lemma 3.4.7, we can couple the sum of spins in both G_1 and G_2 so that they are the same in both chains with probability $\Omega(1)$. This gives the required coupling of (Y_1, Z_1) and $(\widetilde{Y_1}, \widetilde{Z_1})$. □

3.6 Critical Case

In this section, we prove that the mixing time for the Swendsen-Wang dynamics in the critical case $c = 2$ is of order $n^{1/4}$. This is part (ii) of Theorem 3.1.1.

Let X_t and Y_t be two magnetization chains such that X_t starts from an arbitrary location and Y_t starts from the stationary distribution. To prove an upper bound of order $n^{1/4}$ to the mixing time we show that we can couple X_t and Y_t so that they meet in time $O(n^{1/4})$ with probability $\Omega(1)$. For a high level view of this coupling strategy we refer the reader to Section 3.2.3.

Consider the following slight modification to the magnetization chain X_t. Instead of choosing a random spin for each component after the percolation step, we assign a positive spin to the largest component and random spins for all other components. Let X'_t be the sum of spins at time t (notice that we do *not* take absolute values here), that is,

$$X'_{t+1} \overset{d}{=} \max\{|\mathcal{C}_1^+(t)|, |\mathcal{C}_1^-(t)|\} + \epsilon \min\{|\mathcal{C}_1^+(t)|, |\mathcal{C}_1^-(t)|\} + \sum_{j\geq 2} \epsilon_j |\mathcal{C}_j^+(t)| + \sum_{j \geq 2} \epsilon'_j |\mathcal{C}_j^-(t)| \,, \quad (3.99)$$

where as usual $\epsilon, \{\epsilon_j\}$ and $\{\epsilon'_j\}$ are independent mean zero $\pi_{(m)}$ signs. This chain has state space $[-n, n]$ and its absolute value is distributed as our original chain. As a consequence, any upper bound on the mixing time of the modified chain implies the same upper bound on the original chain.

The bulk of this section is devoted to the proof of the upper bound on the mixing time (the corresponding lower bound is much easier to prove and this is done in subsection 3.6.3). To ease the notation, in this section we will refer to this modified chain by X_t and Y_t. The only exception to this in this section is Theorem 3.6.24 where another modification to the chain was required for the proof.

The upper bound asserted in part (ii) of Theorem 3.1.1 will follow immediately by the following two theorems. Though their statement is almost identical, the difference in the starting point X_0 give rise to completely different proof methods so we chose to specify them as two separate theorems for convenience.

Theorem 3.6.1. *Let X_t and Y_t be two SW magnetization chains such that $X_0 \geq n^{3/4}$ and $Y_0 \overset{d}{=} \pi$. Then we can couple X_t and Y_t so that they meet each other within $O(n^{1/4})$ steps with probability $\Omega(1)$.*

Theorem 3.6.2. *Let X_t and Y_t be two SW magnetization chains such that $0 \leq X_0 \leq n^{3/4}$ and $Y_0 \stackrel{d}{=} \pi$. Then we can couple X_t and Y_t so that they meet each other within $O(n^{1/4})$ steps with probability $\Omega(1)$.*

Proof of the upper bound of part (ii) of Theorem 3.1.1: Theorem 3.6.1 and Theorem 3.6.2 give that for any $X_0 \geq 0$ we can couple X_t and Y_t so that they meet within $O(n^{1/4})$ steps. If $X_0 < 0$, then by (3.99) and symmetry we have that

$$P(X_1 \geq 0) \geq \frac{1}{2},$$

so we may apply Theorem 3.6.1 and Theorem 3.6.2 again. This shows that the mixing time of X_t is bounded above by $O(n^{1/4})$. Note that $|X_t|$ and the original magnetization chain has the same distribution. Now Lemma 3.2.1 gives the required upper bound and concludes the proof.. □

3.6.1 Starting at the $[n^{3/4}, n]$ regime: Proof of Theorem 3.6.1

Theorem 3.6.3. [Crossing and overshoot] *Let X_t and Y_t be two SW magnetization chains with $X_0 \geq n^{3/4}$ and $Y_0 \stackrel{d}{=} \pi$. Put*

$$T = \min\left\{t : X_t, Y_t \in [A^{-1}n^{3/4}, An^{3/4}] \text{ and } |X_t - Y_t| \leq hn^{5/8}\right\},$$

for some constant $h > 0$ and large constant A. Then we can choose positive constants h, q, K depending only on A such that

$$\mathbb{P}(T \leq Kn^{1/4}) \geq q\,.$$

Theorem 3.6.4. [Local CLT] *For any constants $A > 1$ and $h > 0$, there exist constants $\delta = \delta(A, h) > 0$ and $k = k(A, h) \in \mathbb{N}$ such that for any $x_0 \in [A^{-1}n^{3/4}, An^{3/4}]$ and any $x \in n + 2\mathbb{Z}$ with $|x - x_0| \leq hn^{5/8}$, we have*

$$\mathbf{P}(X_k = x | X_0 = x_0) \geq \delta n^{-5/8}.$$

Proof of Theorem 3.6.1: By Theorem 3.6.3, the event $T \leq Kn^{1/4}$ occurs with probability at least q. By Theorem 3.6.4 and the strong Markov Property we learn that there exist $\delta > 0$ and $k \in \mathbb{N}$ such that for any $x \in n + 2\mathbb{Z}$ with $|x - X_T| \leq hn^{5/8}$ and $|x - Y_T| \leq hn^{5/8}$, we have

$$\mathbf{P}(X_{T+k} = x \mid T \leq Kn^{1/4}) \geq \delta n^{-5/8}\,,$$

and

$$\mathbf{P}(Y_{T+k} = x \,| \tau \leq Kn^{1/4}) \geq \delta n^{-5/8}.$$

Thus, for any such x we can couple X_t and Y_t so that $X_{T+k} = Y_{T+k} = x$ with probability at least $\delta n^{-5/8}$. We have at least $\frac{hn^{5/8}}{2}$ such x's so in this coupling we have that $X_{T+k} = Y_{T+k}$ with probability at least $h\delta/2$. Lemma 3.1.4 concludes the proof. □

Crossing and overshoot: Proof of Theorem 3.6.3

For any two magnetization chains X_t and Y_t, define $J_t = X_t - Y_t$. Let τ be the first time the two chains cross each other, i.e.

$$\tau := \min\{t : \mathrm{sign} J_t \neq \mathrm{sign} J_0\}. \tag{3.100}$$

The following theorem implies Theorem 3.6.3 immediately.

Theorem 3.6.5. *Let X_t and Y_t be two independent magnetization SW chain with $X_0 \geq n^{3/4}$ and $Y_0 \stackrel{d}{=} \pi$. There exists positive constants δ, K, A and h such that*

$$\mathbb{P}\Big(\tau \leq Kn^{1/4}\,;\; X_{\tau-1}, Y_{\tau-1} \in [A^{-1}n^{3/4},\, An^{3/4}]\,; J_{\tau-1} \leq hn^{5/8}\Big) \geq \delta\,.$$

To prove Theorem 3.6.5 we will use the following results.

Theorem 3.6.6. *The stationary distribution π of the modified magnetization chain satisfies*

$$\lim_{n\to\infty} \pi[a_1 n^{3/4},\, a_2 n^{3/4}] = \frac{1}{Z}\int_{a_1}^{a_2} \exp(-\frac{1}{12}x^4)dx\,,$$

for any constants $a_2 \geq a_1 \geq 0$ where $Z = \int_0^\infty exp(-\frac{1}{12}x^4)dx$ is the normalizing constant.

Lemma 3.6.7. *For any constant $A > 0$ there exists N such that for all $n \geq N$ we have that if $X_0 \in [A^{-1}n^{3/4}, An^{3/4}]$, then the following hold:*

(i). $-Cn^{1/2} \leq \mathbb{E}X_1 - X_0 \leq 0$.

(ii). $\mathbb{E}|X_1 - x_0|^k \leq Cn^{5k/8}$ *for* $k = 2, 3, 4$.

(iii). $\mathbb{E}\sum_{j\geq 1} |\mathcal{C}_j^-|^2 \geq cn^{5/4}$.

where $C = C(A)$ and $c = c(A)$ are constants.

Theorem 3.6.8. *Let X_t and Y_t be two independent magnetization chains with $X_0, Y_0 \in [b_1 n^{3/4}, b_2 n^{3/4}]$ for constants $b_2 > b_1 > 0$. Put $h = \frac{x_0 - y_0}{n^{5/8}}$ and suppose that $h > 0$ and that $h = o(n^{1/8})$. Let τ be the crossing time of X_t and Y_t defined in (3.100). Then there exist positive constants M and δ which only depend on b_1 and b_2 such that*

$$\mathbb{P}(\tau \leq Mh^2) \geq \delta\,.$$

Lemma 3.6.9. *Let X_t be a magnetization SW chain and $I = [a_1 n^{3/4},\, a_2 n^{3/4}]$ where $a_2 > a_1 > 0$ are two constants. Let $h \in (0, a_1)$ and $\xi \in [0, a_1/4]$ be two constants. Then for any $b \in I$, we have*

$$\mathbb{P}\big(\mathrm{sign}(X_1 - b) \neq \mathrm{sign}(X_0 - b) \;\big|\; X_0 > -\xi n^{3/4}\,, |X_0 - b| \geq hn^{3/4}\big) \leq Dn^{-1/3},$$

where $D = D(a_1, a_2, h, \xi)$ is a constant.

Theorem 3.6.10. *For any fixed constants $b_2 > b_1 > 0$, $q < 1$ and $K > 0$, there exists a constant $B = B(b_1, b_2, q, K)$ such that for every $X_0 \in [b_1 n^{3/4},\, b_2 n^{3/4}]$, we have*

$$\mathbb{P}\Big(X_t \le B n^{3/4} \text{ for all } t \in [0, K n^{1/4}]\Big) \ge q.$$

Theorem 3.6.11. *Let X_t be a magnetization SW chain with $X_0 > a n^{3/4}$ where $a > 0$ is a constant. Define $\tau_a = \min\{t : X_t \le a n^{3/4}\}$. Then for any positive constant $b > 0$ we have*

$$\mathbb{P}(\tau_a > b n^{1/4}) \le \sqrt{\frac{6}{ab}}\,. \tag{3.101}$$

We begin by showing how these results imply the main theorem of this subsection.

Proof of Theorem 3.6.5: Let a_1, K and C be three positive constants to be selected later. Define

$$\tau_1 := \min\{t : X_t < a_1 n^{3/4}\}\,,$$

and define $\mathcal{A}$ to be the event that

1. $Y_0 \in [\frac{a}{2} n^{3/4}, a n^{3/4}]$ and
2. $\tau_1 \le K n^{1/4}$ and
3. $Y_{\tau_1} \ge a_1 n^{3/4}$ and
4. $Y_t \le C n^{3/4}$ for all $t \le K n^{1/4}$.

First we determine constants a_1, δ, K and C so that $\mathbb{P}(\mathcal{A}) \ge \delta > 0$. By Theorem 3.6.6, there exists a constant $q > 0$ such that

$$\mathbb{P}\Big(Y_0 \in [\frac{n^{3/4}}{2},\, n^{3/4}]\Big) \ge q.$$

By Theorem 3.6.6 again, we can choose $a_1 > 0$ such that

$$\mathbb{P}\Big(Y_0 \in [-n,\, a_1 n^{3/4}]\Big) \le \frac{q}{2}.$$

Since X_t and Y_t are independent we have that $Y_{\tau_1} \overset{d}{=} \pi_n$. Thus

$$\mathbb{P}\big(Y_0 \in [\frac{n^{3/4}}{2}, n^{3/4}], Y_{\tau_1} > a_1 n^{3/4}\big) \ge \frac{q}{2}.$$

By Lemma 3.6.11 there exists a constant $K = K(a_1, q)$ such that

$$\mathbb{P}\big(Y_0 \in [\frac{n^{3/4}}{2}, n^{3/4}], Y_{\tau_1} > a_1 n^{3/4}, \tau_1 \le K n^{1/4}\big) \ge \frac{q}{4}. \tag{3.102}$$

By Lemma 3.6.10, there is a constant $C = C(K, q)$ such that

$$\mathbb{P}\Big(Y_t \le Cn^{3/4} \text{ for all } t \le Kn^{1/4}\Big) \ge 1 - \frac{q}{8}. \tag{3.103}$$

Combining (3.102) and (3.103) shows that $\mathbb{P}(\mathcal{A}) \ge \frac{q}{8}$. Note that if $\mathcal{A}$ occurs, then

$$\tau \le \tau_1 \le Kn^{1/4}. \tag{3.104}$$

Next we show $\big\{J_{\tau-1} \le \frac{a_1}{2}n^{3/4}\big\} \cap (\frac{t}{n} \wedge 1)$ has positive probability. We do this by proving $J_{\tau-1} \le \frac{a_1}{2}n^{3/4}$ occurs with high probability on $\mathcal{A}$. Note that $\big\{J_{\tau-1} \le \frac{a_1}{2}n^{3/4}\big\} \cap (\frac{t}{n} \wedge 1)$ implies $X_{\tau-1}, Y_{\tau-1} \in [A^{-1}n^{3/4}, An^{3/4}]$ for some large constant A.

If $\{J_{\tau-1} > \frac{a_1}{2}n^{3/4}\} \cap (\frac{t}{n} \wedge 1)$ occurs, then there exists some $t \le Kn^{1/4}$ such that $J_t > \frac{a_1}{2}n^{3/4}$ and $J_{t+1} < 0$. This implies that there is a point $y \in [\frac{a_1}{2}n^{3/4}, (C + \frac{a_1}{2})n^{3/4}]$ with $|X_t - y| \ge \frac{a_1}{4}n^{3/4}$ and $|Y_t - y| \ge \frac{a_1}{4}n^{3/4}$ and at least one of X_{t+1} and Y_{t+1} crosses y. Suppose first that $Y_{\tau-1} \ge -\xi n^{3/4}$ where ξ is a small positive constant. Then Lemma 3.6.9 and the union bound give that

$$\mathbb{P}\Big(\mathcal{A}, J_{\tau-1} > \frac{a_1}{2}n^{3/4}, Y_{\tau-1} \ge -\xi n^{3/4}\Big) \le Dn^{-1/3}Kn^{1/4} = o(1). \tag{3.105}$$

Next suppose that $Y_{\tau-1} < -\xi n^{3/4}$. Then there is a $t \in [0, Kn^{1/4}]$ such that $Y_t \le -\xi n^{3/4}$. By (3.99), for any starting location, we have

$$\mathbb{P}(X_1 < -\xi n^{3/4}) \le \mathbb{P}\Big(\epsilon \min\{|\mathcal{C}_1^+(t)|, |\mathcal{C}_1^-(t)|\} + \sum_{j \ge 2} \epsilon_j |\mathcal{C}_j^+(t)| + \sum_{j \ge 2} \epsilon_j' |\mathcal{C}_j^-(t)| < -\xi n^{3/4}\Big).$$

By Theorem 3.3.13 we have that

$$\mathbb{E}\Big(\epsilon \min\{|\mathcal{C}_1^+(t)|, |\mathcal{C}_1^-(t)|\} + \sum_{j \ge 2} \epsilon_j |\mathcal{C}_j^+(t)| + \sum_{j \ge 2} \epsilon_j' |\mathcal{C}_j^-(t)|\Big)^4 = O(n^{8/3}), \tag{3.106}$$

so Markov's inequality gives that

$$\mathbb{P}(X_1 < -\xi n^{3/4}) = O(n^{-1/3}). \tag{3.107}$$

The union bound implies now that

$$\mathbb{P}\Big(\mathcal{A}, J_{\tau-1} > \frac{a_1}{2}n^{3/4}, Y_{\tau-1} < -\xi n^{3/4}\Big) \le O(n^{-1/3})Kn^{1/4} = o(1),$$

and so together with (3.105) we conclude that $\{\mathcal{A}, X_{\tau-1}, Y_{\tau-1} \in [A^{-1}n^{3/4}, An^{3/4}]\}$ occurs with probability $\Omega(1)$ for some constant A. We denote this event by $\mathcal{B}$.

It remains to prove that $\{J_{\tau-1} \le hn^{5/8}\} \cap \mathcal{B}$ occurs with probability $\Omega(1)$ for some constant $h > 0$. Suppose first $J_{\tau-1} > n^{23/32}$. Notice that $\{\mathcal{B}, J_{\tau-1} > n^{23/32}\}$ implies there is a $t < Kn^{1/4}$ such that $X_t, Y_t \in [A^{-1}n^{3/4}, An^{3/4}]$, $J_t > n^{23/32}$ and $J_{t+1} < 0$. This implies at least one of X_t and Y_t has to make a huge jump of order at least $n^{23/32}$. By part (ii) of Lemma 3.6.7 with $k = 4$, Markov's inequality and the union bound we have

$$\mathbb{P}(J_{\tau-1} > n^{23/32}, \mathcal{B}) \leq O(n^{4(5/8-23/32)}n^{1/4}) = o(1). \tag{3.108}$$

To handle the case $J_{\tau-1} < n^{23/32}$, let

$$W_k = [2^k n^{5/8}, 2^{k+1}n^{5/8}],$$

and consider the probability $\mathbb{P}(J_{\tau-1} \in W_k, \mathcal{B})$. Let T_m be the first time that $J_t \in W_k$ for m-th time and

$$(\frac{t}{n} \wedge 1)_m = \bigcap_{1\leq m'\leq m} \{X_{T_{m'}}, Y_{T_{m'}} \in [A^{-1}n^{3/4}, A^{3/4}]\}.$$

Note that $(\frac{t}{n} \wedge 1)_m \in \mathcal{F}_{T_m}$. We have

$$\mathbb{P}(J_{\tau-1} \in W_k, \mathcal{B}) \leq \sum_{m=1}^{\infty} \mathbb{P}(T_m \leq \tau - 1\,, J_{T_m+1} < 0\,, \mathcal{B}). \tag{3.109}$$

Notice that $\{T_m \leq \tau - 1, J_{T_m+1} < 0, \mathcal{B}\}$ implies that for all $m' \leq m$, we have $X_{T_{m'}} > a_1 n^{3/4}$, $Y_{T_{m'}} < Cn^{3/4}$ and $|X_{T_{m'}} - Y_{T_{m'}}| \leq 2^{k+1}n^{5/8}$. This in particular implies that $X_{T_{m'}}, Y_{T_{m'}} \in [A^{-1}n^{3/4}, A^{3/4}]$. Hence $\{T_m \leq \tau - 1\,, J_{T_m+1} < 0\,, \mathcal{B}\}$ implies $\{T_m \leq \tau - 1\,, J_{T_m+1} < 0\,, (\frac{t}{n} \wedge 1)_m\}$. Also, by part (ii) of Lemma 3.6.7 and Markov's inequality, we have

$$\mathbb{P}(|X_{t+1} - X_t| \geq 2^{k-1}n^{5/8} \,|\, X_t = \Theta(n^{3/4})) \leq \frac{C}{2^{4k}}.$$

The same inequality holds for Y_t by the same reason. Thus

$$\mathbb{P}\Big(|J_{t+1} - J_t| \geq 2^k n^{5/8} \Big| X_t, Y_t = \Theta(n^{3/4})\Big) \leq \frac{C}{2^{4k}}.$$

We now use the strong Markov property on the stopping time T_m and plug the above estimate in (3.109) to get that

$$\begin{aligned}
\mathbb{P}(J_{\tau-1} \in W_k,\, \mathcal{B}) &\leq \sum_{m=1}^{\infty} \mathbb{P}(T_m \leq \tau - 1, J_{T_m+1} < 0, (\frac{t}{n} \wedge 1)_m) \\
&\leq \sum_{m=1}^{\infty} \frac{C}{2^{4k}} \mathbb{P}(T_m \leq \tau - 1,\, (\frac{t}{n} \wedge 1)_m)\,.
\end{aligned} \tag{3.110}$$

If $\{T_m \leq \tau - 1\,, (\frac{t}{n} \wedge 1)_m\}$ occurs, then for any $l \leq m$, we have $T_{m-l} \leq \tau - 1$ and $X_{T_{m-l}}, Y_{T_{m-l}} \in [A^{-1}n^{3/4}, An^{3/4}]$ and most importantly, the chains do not cross between time T_{m-l} and T_m, which is at least l steps. Now, let M and r be the constants from Lemma 3.6.8 and put $l = M2^{2k}$. The strong Markov property on the stopping time $T_{m-M2^{2k}}$ and Lemma 3.6.8 gives that

$$\mathbb{P}(T_m \leq \tau - 1,\, (\frac{t}{n} \wedge 1)_m) \quad \leq \quad (1-r)\mathbb{P}(T_{m-M2^{2k}} \leq \tau - 1,\, (\frac{t}{n} \wedge 1)_{m-M2^{2k}})\,.$$

Applying this recursively gives that

$$\mathbb{P}(T_m \leq \tau - 1, \, (\frac{t}{n} \wedge 1)_m) \leq (1-r)^{[\frac{m}{M2^{2k}}]}.$$

Plugging this into (3.110), we get

$$\mathbb{P}(J_{\tau-1} \in W_k, \, \mathcal{B}) \leq \frac{C}{2^{4k}} \sum_{m=1}^{\infty} (1-r)^{[\frac{m}{M2^{2k}}]} = \frac{C}{2^{2k}}.$$

Combining this and (3.108) we have for large enough k_0, $\{J_{\tau-1} \leq 2^{k_0} n^{5/8}, \mathcal{B}\}$ occurs with probability $\Omega(1)$, which concludes the proof of the theorem. $\square$

We now proceed with proving the statements we have used so far in the proof of Theorem 3.6.5. To prove Theorem 3.6.6 we will use the following small lemmas.

Lemma 3.6.12 (Simon and Griffiths (1973))**.** *Denote by S_n the sum of spins for Ising model on the complete graph. If the inverse temperature $\beta = \frac{1}{n}$, then there exists a random variable X with density proportional to* $\exp(-\frac{1}{12}x^4)$ *such that*

$$\frac{S_n}{n^{3/4}} \xrightarrow{d} X,$$

as $n \to \infty$.

Corollary 3.6.13. *Consider Ising model on the complete graph with inverse temperature $\beta = \frac{1}{n} + O(\frac{1}{n^2})$. For any fixed constants $a_2 \geq a_1 \geq 0$, we have*

$$\lim_{n\to\infty} \mathbb{P}(|S_n| \in [a_1 n^{3/4}, a_2 n^{3/4}]) = \frac{1}{A} \int_{a_1}^{a_2} exp(-\frac{1}{12}x^4) dx,$$

where $A = \int_0^\infty exp(-\frac{1}{12}x^4)dx$ is the normalizing constant.

Proof of Corollary 3.6.13: By Lemma 3.6.12 we have that the conclusion of the corollary holds for $\beta_1 = \frac{1}{n}$. Thus it suffices to prove that for any configuration σ in which $|S_n(\sigma)| \in [a_1 n^{3/4}, a_2 n^{3/4}]$ we have

$$\mathbf{P}_\beta(\sigma) = (1+o(1))\mathbf{P}_{\beta_1}(\sigma)\,.$$

Observe that on complete graph we have that

$$\sum_{u,v,u\neq v} \sigma(u)\sigma(v) = \frac{S_n^2 - n}{2}.$$

Thus, for any σ with $S_n(\sigma) \in [a_1 n^{3/4}, a_2 n^{3/4}]$, we have

$$\frac{\mathbb{P}_\beta(\sigma)}{\mathbb{P}_{\beta_1}(\sigma)} = \frac{e^{\beta(\frac{S_n^2-n}{2})}/Z(\beta)}{e^{\beta_1(\frac{S_n^2-n}{2})}/Z(\beta_1)} = (1+o(1))\frac{Z(\beta_1)}{Z(\beta)}\,, \tag{3.111}$$

so it is enough to show $Z(\beta) = (1 + o(1))Z(\beta_1)$. Indeed, Lemma 3.6.12 implies that

$$\mathbb{P}_{\beta_1}(|S_n| \geq n^{7/8}) = o(1)\,,$$

but for any configuration σ with $|S_n(\sigma)| \leq n^{7/8}$ we have that

$$e^{\beta(\frac{S_n^2-n}{2})} = (1 + o(1))e^{\beta_1(\frac{S_n^2-n}{2})}\,,$$

and the assertion follows. □

Lemma 3.6.14. *Let π_n be the stationary distribution of the modified magnetization SW chain X_t, then we have*

$$\lim_{n\to\infty} \pi_n[-\infty, 0] = 0.$$

Proof of Lemma 3.6.14: Recall that SW dynamics with parameter p has stationary distribution of Ising model with $p = 1 - e^{-2\beta}$. Plugging in $p = \frac{2}{n}$ we get $\beta = \frac{1}{n} + O(\frac{1}{n^2})$. By Corollary 3.6.13, for any $\epsilon > 0$, there exists constant b_1 and b_2 such that $0 < b_1 < b_2$ and

$$\pi_n\big([b_1 n^{3/4}, b_2 n^{3/4}] \cup [-b_2 n^{3/4}, -b_1 n^{3/4}]\big) > 1 - \epsilon.$$

By definition of stationarity, for any set S we have

$$\sum_{y\in[-n,n]} \pi_n(y)\mathbb{P}(y, S) = \pi_n(S). \tag{3.112}$$

Put $S = [-n, 0]$ and denote $\pi_n[-n, 0]$ by δ_n. For any X_0 we have

$$\mathbb{P}(X_1 \leq 0) \leq 1/2\,,$$

by symmetry. For $X_0 \in [b_1 n^{3/4}, b_2 n^{3/4}]$ Lemma 3.6.9 gives that

$$P(X_1 < 0) \leq Dn^{-1/3}\,.$$

Plugging these into (3.112), we have

$$\delta_n = \sum_{y\in[-n,n]} \pi_n(y)\mathbb{P}(y, S) \leq \frac{1}{2}\delta_n + \epsilon + Dn^{-1/3},$$

which gives

$$\delta_n \leq 2(\epsilon + Dn^{-1/3})\,,$$

concluding the proof. □

Proof of Theorem 3.6.6: Directly follows from Corollary 3.6.13 and Lemma 3.6.14. □

The following is an easy estimate which use frequently to show that the main contribution from the first term of (3.99) comes from the $|\mathcal{C}_1^+|$ element rather than the $|\mathcal{C}_1^-|$ element.

Proposition 3.6.15. *If $X_0 \geq Cn^{2/3}\log^2 n$ for some large constant C, then*

$$\mathbf{P}(|\mathcal{C}_1^-| \geq |\mathcal{C}_1^+|) \leq O(e^{-c\log^2 n})\,.$$

Proof. By our condition on X_0 we have that $|\mathcal{C}_1^+|$ is distributed as the size of the largest component in a supercritical random graph $G(m,p)$ with $m = \frac{n+X_0}{2}$ and $p = \frac{1+\epsilon}{m}$ with $\epsilon = X_0/n = \Omega(n^{-1/3}\log^2 n)$. Theorem 3.3.9 gives that

$$\mathbf{P}(|\mathcal{C}_1^+| \geq cn^{2/3}\log^2 n) \geq 1 - Ce^{-c\log^2 n}\,,$$

for some small $c > 0$. On the other hand $|\mathcal{C}_1^-|$ is distributed as a subcritical random graph. Theorem 1 of [30] gives that

$$\mathbf{P}(|\mathcal{C}_1^-| \geq cn^{2/3}\log^2 n) \leq Ce^{-c\log^2 n}\,,$$

which finishes the proof. □

Lemma 3.6.16. *If $X_t \geq Cn^{2/3}\log n$ for some large constant C, then*

$$\mathbb{E}[X_{t+1} \mid X_t] \leq X_t\Big(1 - \frac{X_t}{6n}\Big). \tag{3.113}$$

Proof. By (3.99) we have $\mathbb{E}[X_{t+1} \mid X_t] = \mathbb{E}\big[|\max\{|\mathcal{C}_1^+|, |\mathcal{C}_1^-|\}|\big|X_t\big]$, hence Proposition 3.6.15 gives that

$$\mathbb{E}[X_{t+1} \mid X_t] = \mathbb{E}|\mathcal{C}_1^+| + O(e^{-c\log^2 n})\,.$$

Thus, Theorem 3.3.8 yields that

$$\mathbb{E}[X_{t+1} \mid X_t] \leq 2\frac{X_t}{n}\frac{n+X_t}{2} - \frac{7}{3}\frac{X_t^2}{n^2}\frac{n+X_t}{2} + O(e^{-c\log^2 n}) \leq X_t\Big(1 - \frac{X_t}{6n}\Big)\,,$$

when n is large enough. □

Proof of Lemma 3.6.7: As in the previous proof we have

$$\mathbb{E}X_1 = \mathbb{E}|\mathcal{C}_1^+| + O(e^{-c\log^2 n})\,.$$

Since $\epsilon = \frac{x_0}{n} = \Theta(n^{-1/4})$ Theorem 3.3.8 gives that

$$\begin{aligned}
\mathbb{E}|\mathcal{C}_1^+| &= 2\frac{x_0}{n}\frac{n+x_0}{2} - \frac{8}{3}\Big(\frac{x_0}{n}\Big)^2\frac{n+x_0}{2} + O\Big(\Big(\frac{x_0}{n}\Big)^3\frac{n+x_0}{2}\Big) \\
&= x_0 - \frac{x_0^2}{3n} + O\Big(\frac{x_0^3}{n^2}\Big) = x_0 - \frac{x_0^2}{3n} + O(n^{1/4})\,,
\end{aligned}$$

which gives part (i) of the lemma since $x_0 \in [A^{-1}n^{3/4}, An^{3/4}]$. We now prove part (ii). For $k = 2, 3, 4$, by (3.99) and Jensen's inequality we have that

$$\begin{aligned}
\mathbb{E}|X_1 - x_0|^k &= \mathbb{E}\Big||\mathcal{C}_1^+| - x_0 + \sum_{j\geq 2}\epsilon_j|\mathcal{C}_j^+| + \sum_{j\geq 1}\epsilon_j'|\mathcal{C}_j^-|\Big|^k + \Theta(e^{-cn^{1/8}}) \\
&\leq 2^{k-1}\Big(\mathbb{E}\Big||\mathcal{C}_1^+| - x_0\Big|^k + \mathbb{E}\Big|\sum_{j\geq 2}\epsilon_j|\mathcal{C}_j^+| + \sum_{j\geq 1}\epsilon_j'|\mathcal{C}_j^-|\Big|^k\Big). \qquad (3.114)
\end{aligned}$$

Theorem 3.3.10 now gives that

$$\begin{aligned}\mathbb{E}\Big||\mathcal{C}_1^+| - 2\frac{x_0}{n}\frac{n+x_0}{2}\Big|^k &\le C\Big(\frac{n+x_0}{2}\Big/\frac{x_0}{n}\Big)^{k/2} \\ &\le C\Big(\frac{n^2}{x_0}\Big)^{k/2} \le O(n^{5k/8}).\end{aligned}$$

Another application of Jensen's inequality gives that

$$\begin{aligned}\mathbb{E}\big||\mathcal{C}_1^+| - x_0\big|^k &= \mathbb{E}\big|\big(|\mathcal{C}_1^+| - 2\frac{x_0}{n}\frac{n+x_0}{2}\big) + \big(2\frac{x_0}{n}\frac{n+x_0}{2} - x_0\big)\big|^k \\ &\le 2^{k-1}\Big(O(n^{5k/8}) + \Big(\frac{x_0^2}{n}\Big)^k\Big) \le O(n^{5k/8}). \qquad (3.115)\end{aligned}$$

To bound the rest of (3.114), notice that by Holder's inequality, we only need to consider the case $k = 4$. We have

$$\begin{aligned}\mathbb{E}\Big|\sum_{j\ge2}\epsilon_j|\mathcal{C}_j^+| + \sum_{j\ge1}\epsilon_j'|\mathcal{C}_j^-|\Big|^4 &\le \sum_{j\ge2}\mathbb{E}|\mathcal{C}_j^+|^4 + \sum_{j\ge1}\mathbb{E}|\mathcal{C}_j^-|^4 + \Big(\sum_{j\ge2}\mathbb{E}|\mathcal{C}_j^+|^2\Big)\Big(\sum_{j\ge1}\mathbb{E}|\mathcal{C}_j^-|^2\Big) \\ &+ \mathbb{E}\sum_{i,j\ge2,i\ne j}|\mathcal{C}_i^+|^2|\mathcal{C}_j^+|^2 + \mathbb{E}\sum_{i,j\ge1,i\ne j}|\mathcal{C}_i^-|^2|\mathcal{C}_j^-|^2\,.\end{aligned}$$

By Theorem 3.3.10 we have

$$\sum_{j\ge2}\mathbb{E}|\mathcal{C}_j^+|^2 \le C_2 n\Big(\frac{x_0}{n}\Big)^{-1} = O(n^{5/4})$$

and

$$\sum_{j\ge2}\mathbb{E}|\mathcal{C}_j^+|^4 \le C_4 n\Big(\frac{x_0}{n}\Big)^{-5} = O(n^{9/4}).$$

By Theorem 3.3.12, we have

$$\sum_{j\ge1}\mathbb{E}|\mathcal{C}_j^-|^2 \le C_2 n\Big(\frac{x_0}{n}\Big)^{-1} = O(n^{5/4})$$

and

$$\sum_{j\ge1}\mathbb{E}|\mathcal{C}_j^-|^4 \le C_4 n\Big(\frac{x_0}{n}\Big)^{-5} = O(n^{9/4}).$$

These together with Theorem 3.3.13 to handle the cross terms finishes the proof of part (ii) of the lemma. Part (iii) follows immediately by Theorem 3.3.12,

$$\mathbb{E}\sum_{j\ge1}|\mathcal{C}_j^-|^2 \ge c_2\frac{n-x_0}{2}\frac{n}{x_0} \ge c_2\frac{n}{4}A^{-1}n^{1/4} \ge cn^{5/4}\,.$$

□

Lemma 3.6.17. *Let X be a real valued random variable with $\mathbb{E}X = 0$ and $\mathbb{E}X^2 \ge h^2$ and $\mathbb{E}X^4 \le bh^4$ where $b \ge 1$. Then for any $\rho \in [0,1]$ we have*

$$\mathbb{P}(X \le -\rho h) \ge \frac{(1-\rho^2)^2}{2b}\,.$$

Proof of Lemma 3.6.17: By Cauchy-Schwartz

$$\mathbb{E}[X^2\mathbf{1}_{\{X^2\geq\rho^2h^2\}}] \leq \sqrt{\mathbb{E}X^4\mathbb{E}\mathbf{1}_{\{X^2\geq\rho^2h^2\}}} \leq \sqrt{bh^4\mathbb{P}(X^2 \geq \rho^2h^2)}\,.$$

Hence,

$$h^2 \leq \mathbb{E}X^2 \leq \rho^2h^2 + \mathbb{E}[X^2\mathbf{1}_{\{X^2\geq\rho^2h^2\}}] \leq \rho^2h^2 + \sqrt{bh^4\mathbb{P}(X^2 \geq \rho^2h^2)}\,.$$

We conclude that

$$\mathbb{P}(|X| \geq \rho h) \geq \frac{(1-\rho^2)^2}{b}\,,$$

and the assertion follows by symmetry since $\mathbb{P}(X \leq -\rho h) = \mathbb{P}(-X \leq -\rho h)$. □

The following will be used in the proof of Theorem 3.6.8.

Theorem 3.6.18. *Let X_t be a magnetization chain with $X_0 \in [b_1n^{3/4}, b_2n^{3/4}]$ where $b_2 > b_1 > 0$ are two constants. Let τ_1 be the first time that $X_t \notin [\frac{b_1}{2}n^{3/4},\ (b_2+\frac{b_1}{2})n^{3/4}]$. Then there exists a constant $C = C(b_1, b_2) > 0$ such that for all constant $\delta > 0$ we have*

$$\mathbb{P}(\tau_1 \leq \delta n^{1/4}) \leq C\delta^2.$$

Proof of Theorem 3.6.18 Denote by I the interval $[\frac{b_1}{2}n^{3/4},\ (b_2+\frac{b_1}{2})n^{3/4}]$. Part (ii) of Lemma 3.6.7 gives

$$\mathbb{E}\Big[(X_{(t+1)\wedge\tau_1} - X_{t\wedge\tau_1})^k\Big|\mathcal{F}_t\Big] \leq Cn^{5k/8} \tag{3.116}$$

for $k = 2, 3, 4$. Define

$$Z := X_{(t+1)\wedge\tau_1} - X_{t\wedge\tau_1} - (\mathbb{E}X_{(t+1)\wedge\tau_1} - \mathbb{E}X_{t\wedge\tau_1}).$$

Note that $\left|\mathbb{E}X_{(t+1)\wedge\tau_1} - \mathbb{E}X_{t\wedge\tau_1}\right| \leq Cn^{1/2}$ by part (i) of Lemma 3.6.7, hence

$$\mathbb{E}\Big[Z^k\Big|\mathcal{F}_t\Big] \leq Cn^{\frac{5k}{8}} \tag{3.117}$$

for $k = 2, 3, 4$. Also, for $k = 1$, part (i) of Lemma 3.6.7 gives that

$$\mathbb{E}[Z|\mathcal{F}_t] \leq C\sqrt{n}. \tag{3.118}$$

Denote

$$f(t) = \Big(\mathbb{E}[X_{t\wedge\tau_1} - \mathbb{E}X_{t\wedge\tau_1}]^4\Big)^{1/2}.$$

Note that

$$f(t+1)^2 = \mathbb{E}[X_{(t+1)\wedge\tau} - \mathbb{E}X_{(t+1)\wedge\tau_1}]^4 = \mathbb{E}\Big[(X_{t\wedge\tau_1} - \mathbb{E}X_{t\wedge\tau_1}) + Z\Big]^4. \tag{3.119}$$

For $k = 1, 2, 3, 4$, we have

$$\begin{aligned}
\mathbb{E}\Big[\Big(X_{t\wedge\tau_1} - \mathbb{E}X_{t\wedge\tau_1}\Big)^{4-k}Z^k\Big] &= \mathbb{E}\Big(\mathbb{E}\Big[\Big(X_{t\wedge\tau_1} - \mathbb{E}X_{t\wedge\tau_1}\Big)^{4-k}Z^k\Big|\mathcal{F}_t\Big]\Big) \\
&= \mathbb{E}\Big[\Big(X_{t\wedge\tau_1} - \mathbb{E}X_{t\wedge\tau_1}\Big)^{4-k}\mathbb{E}[Z^k|\mathcal{F}_t]\Big].
\end{aligned}$$

Hölder's inequality implies that

$$\mathbb{E}\Big[\Big(X_{t\wedge\tau_1}-\mathbb{E}X_{t\wedge\tau_1}\Big)^{4-k}Z^k\Big]\le Cn^{\frac{5k}{8}}f(t)^{\frac{4-k}{2}}, \tag{3.120}$$

for $k=2,3,4$ and by (3.118)

$$\mathbb{E}\Big[\Big(X_{t\wedge\tau_1}-\mathbb{E}X_{t\wedge\tau_1}\Big)^{3}Z\Big]\le C\sqrt{n}f(t)^{\frac{3}{2}}. \tag{3.121}$$

Expanding the right hand side of (3.119) and plugging (3.120) and (3.121) into it, we get

$$f(t+1)^2\le f(t)^2+C\sqrt{n}f(t)^{3/2}+Cn^{5/4}f(t)+Cn^{15/8}f(t)^{1/2}+Cn^{5/2}. \tag{3.122}$$

Comparing the right hand side of (3.122) with

$$\Big(f(t)+Cn^{1/2}f(t)^{1/2}+Cn^{5/4}\Big)^2, \tag{3.123}$$

we find that the first, second, third and fifth term of (3.122) is dominated by expanding (3.123). For the forth term, if $f(t)=O(n^{5/4})$, then it is dominated by $(Cn^{5/4})^2$. Otherwise it is dominated by $Cn^{5/4}f(t)$. The conclusion is that

$$f(t+1)^2\le\Big(f(t)+Cn^{1/2}f(t)^{1/2}+Cn^{5/4}\Big)^2. \tag{3.124}$$

Thus, if $f(t)=O(n^{3/2})$, then we have

$$f(t+1)\le f(t)+Cn^{5/4}. \tag{3.125}$$

Since $f(0)=0$, by iterating (3.125) we get that $f(t)\le Ctn^{5/4}$ for all $t\le\delta n^{1/4}$ where $\delta>0$ is a constant. Put $t=\delta n^{1/4}$. Markov's inequality gives that

$$\mathbb{P}(|X_{\delta n^{1/4}\wedge\tau_1}-\mathbb{E}X_{\delta n^{1/4}\wedge\tau_1}|\ge\frac{b_1}{4}n^{3/4})\le\frac{(C\delta)^2}{(\frac{b_1}{4})^4}. \tag{3.126}$$

By part (i) of Lemma 3.6.7 we have that

$$|\mathbb{E}X_{\delta n^{1/4}\wedge\tau_1}-X_0|\le C\delta n^{3/4}.$$

Thus, for small enough δ we have

$$\mathbb{P}\Big(\Big|X_{\delta n^{1/4}\wedge\tau_1}-x_0\Big|\le\frac{b_1}{2}n^{3/4}\Big)\ge 1-C\delta^2, \tag{3.127}$$

which means that X_t has not jumped out of the window I within $\delta n^{1/4}$ steps with probability at least $1-C\delta^2$. □

Proof of Theorem 3.6.8: Recall that $J_t=X_t-Y_t$. Let M be a large constant that will be chosen later. Assume without loss of generality that $J_0\ge 0$, we will prove that J_{Mh^2}

is negative with probability $\Omega(1)$, which implies the theorem. Denote by I the interval $[\frac{b_1}{2}n^{3/4}, (b_2+\frac{b_1}{2})n^{3/4}]$ and define

$$\tau_1 = \min\{t : X_t \notin I \text{ or } Y_t \notin I\}\,.$$

We will prove our claim by precisely estimating the first, second and forth moment of $J_{Mh^2\wedge\tau_1}$ and then apply Lemma 3.6.17 to $J_{Mh^2\wedge\tau_1} - \mathbb{E}J_{Mh^2\wedge\tau_1}$. We start with first moment estimate. By part (i) of Lemma 3.6.7 and the optional stopping theorem we get

$$\mathbb{E}X_{t\wedge\tau_1} - C\sqrt{n} \le \mathbb{E}X_{(t+1)\wedge\tau_1} \le \mathbb{E}X_{t\wedge\tau_1} \tag{3.128}$$

for some constant $C = C(b_1, b_2) > 0$. Applying (3.128) recursively gives that

$$X_0 - CMh^2\sqrt{n} \le \mathbb{E}X_{Mh^2\wedge\tau_1} \le X_0. \tag{3.129}$$

The same formula holds for Y_t, hence

$$\mathbb{E}J_{Mh^2\wedge\tau_1} \le hn^{5/8} + CMh^2n^{1/2}. \tag{3.130}$$

We proceed with the second moment estimate. Notice that if $X_0 \in I$, we have that

$$\begin{aligned}\mathbb{E}(X_1 - \mathbb{E}(X_1|\mathcal{F}_0))^2 \;&=\; \mathbb{E}\Big[\max\{|\mathcal{C}_1^+|, |\mathcal{C}_1^-|\} - \mathbb{E}\max\{|\mathcal{C}_1^+|, |\mathcal{C}_1^-|\}\\ &+\; \epsilon\min\{|\mathcal{C}_1^+|, |\mathcal{C}_1^-|\} + \sum_{j\ge2}\epsilon_j|\mathcal{C}_j^+| + \sum_{j\ge2}\epsilon_j'|\mathcal{C}_j^-|\Big]^2\end{aligned}$$

by (3.99). In the $n^{3/4}$ regime, we have $\mathbb{P}(|C_1^-| \ge |C_1^+|) = O(e^{-c\log^2 n})$ by Proposition 3.6.15, hence

$$\mathbb{E}(X_1 - \mathbb{E}(X_1|\mathcal{F}_0))^2 \ge (1 - Ce^{-c\log^2 n})\mathbb{E}\sum_{j\ge1}|C_j^-|^2 \ge c_1 n^{5/4}\,,$$

by part (iii) of Lemma (3.6.7). Also, by part (ii) of Lemma 3.6.7 we have that

$$\mathbb{E}(X_1 - \mathbb{E}(X_1|\mathcal{F}_0))^2 \le Cn^{5/4}.$$

Now let

$$A_t = \sum_{i=0}^{t-1}(X_{(i+1)\wedge\tau_1} - X_{i\wedge\tau_1}) - \mathbb{E}(X_{(i+1)\wedge\tau_1} - X_{i\wedge\tau_1}|\mathcal{F}_i)$$

and

$$B_t = X_0 - \mathbb{E}X_{t\wedge\tau_1} + \sum_{i=0}^{t-1}\mathbb{E}(X_{(i+1)\wedge\tau_1} - X_{i\wedge\tau_1}|\mathcal{F}_i).$$

Then it is easy to verify that $A_t + B_t = X_{t\wedge\tau_1} - \mathbb{E}X_{t\wedge\tau_1}$. Moreover, since the martingale increments are orthogonal we have that

$$\mathbb{E}A_t^2 = \sum_{i=0}^{t-1}\mathbb{E}(X_{(i+1)\wedge\tau_1} - \mathbb{E}(X_{(i+1)\wedge\tau_1}|\mathcal{F}_i))^2.$$

Since $h = o(n^{1/8})$ Theorem 3.6.18 gives that

$$\mathbb{P}(\tau_1 \le Mh^2) = o(1)\,.$$

This implies that

$$cMh^2n^{5/4} \le \mathbb{E}A^2_{Mh^2} \le CMh^2n^{5/4}.$$

By (3.129) and part (i) of Lemma 3.6.7, we get that $|B_t| \le Ctn^{1/2}$. This gives

$$\mathbb{E}B_t^2 \le Ct^2n.$$

Cauchy-Schwarz inequality gives

$$\mathbb{E}|A_{Mh^2}B_{Mh^2}| \le CMh^2n^{9/8}.$$

Thus, we have

$$\mathbf{Var}X_{Mh^2\wedge\tau_1} = \mathbb{E}A^2_{Mh^2} + \mathbb{E}B^2_{Mh^2} + 2\mathbb{E}A_{Mh^2}B_{Mh^2} \ge (c - o(1))Mh^2n^{5/4}.$$

The same estimates hold for Y_t. Since X_t and Y_t are independent we have

$$\mathbf{Var}J_{Mh^2\wedge\tau_1} \ge c_1Mh^2n^{5/4}. \tag{3.131}$$

For the fourth moment estimate, by (3.125) we have

$$\mathbb{E}[X_{Mh^2\wedge\tau_1} - \mathbb{E}X_{Mh^2\wedge\tau_1}]^4 \le (Mh^2Cn^{5/4})^2$$

and

$$\mathbb{E}[Y_{Mh^2\wedge\tau_1} - \mathbb{E}Y_{Mh^2\wedge\tau_1}]^4 \le (Mh^2Cn^{5/4})^2.$$

By the Jensen's inequality, we get

$$\mathbb{E}[J_{Mh^2\wedge\tau_1} - \mathbb{E}J_{Mh^2\wedge\tau_1}]^4 \le 16(Mh^2Cn^{5/4})^2. \tag{3.132}$$

Putting (3.131) and (3.132) together, taking $\rho = \frac{1}{\sqrt{Mc_1}}$ and using Lemma 3.6.17, we get

$$\mathbb{P}(J_{Mh^2\wedge\tau_1} - \mathbb{E}J_{Mh^2\wedge\tau_1} \le -hn^{5/8}) \ge \delta,$$

where $\delta > 0$ is a constant. Combining this with (3.130), we get

$$\mathbb{P}\Big(J_{Mh^2\wedge\tau_1} \le 0\Big) \ge \delta. \tag{3.133}$$

Here we choose M so that $Mc_1 \ge 2$, concluding the proof. $\square$

Proof of Lemma 3.6.9: If $X_0 \le b - hn^{3/4}$, then assume first $X_0 \in [\xi n^{3/4}, b - hn^{3/4}]$. In this regime, by part (ii) of Lemma 3.6.7 with $k = 4$ and Markov's inequality we have

$$\mathbb{P}\big(\text{sign}(X_1 - b) \ne \text{sign}(X_0 - b)\big) \le \frac{Cn^{5/2}}{h^4n^3} = O(n^{-1/2}).$$

Assume now $X_0 \in [-\xi n^{3/4}, \xi n^{3/4}]$. Recall the distribution of X_1 in (3.99) and that $b \geq a_1 n^{3/4}$ and $\xi \leq a_1/4$. If $X_1 \leq a_1 n^{3/4}$, then either

$$\max\{|\mathcal{C}_1^+|, |\mathcal{C}_1^-|\} > 2\xi n^{3/4},$$

or

$$\epsilon \min\{|\mathcal{C}_1^+(t)|, |\mathcal{C}_1^-(t)|\} + \sum_{j\geq 2} \epsilon_j |\mathcal{C}_j^+(t)| + \sum_{j\geq 2} \epsilon_j' |\mathcal{C}_j^-(t)| \geq \frac{a_1}{2} n^{3/4}.$$

By Theorem 3.3.9 and monotonicity of $|\mathcal{C}_1|$, we have

$$\mathbb{P}(\max\{|\mathcal{C}_1^+|, |\mathcal{C}_1^-|\} > 2\xi n^{3/4}) \leq Ce^{-cn^{1/8}}.$$

By Theorem 3.3.13 and Markov's inequality, we have

$$\mathbb{P}(\epsilon \min\{|\mathcal{C}_1^+(t)|, |\mathcal{C}_1^-(t)|\} + \sum_{j\geq 2} \epsilon_j |\mathcal{C}_j^+(t)| + \sum_{j\geq 2} \epsilon_j' |\mathcal{C}_j^-(t)| \geq \frac{a_1}{2} n^{3/4}) = O(n^{-1/3}).$$

Thus, we have

$$\mathbb{P}(X_1 \geq b) = O(n^{-1/3}).$$

If $X_0 \geq b + hn^{3/4}$, then assume first $X_0 \in [b + h^{3/4}, Bn^{3/4}]$ for some large constant B. By part (ii) of Lemma 3.6.7 with $k = 4$ and Markov's inequality, we have

$$\mathbb{P}(X_1 \leq b) \leq \frac{Cn^{5/2}}{h^4 n^3} = O(n^{-1/2}).$$

Assume $X_0 \geq Bn^{3/4}$. If $X_1 \leq bn^{3/4}$, then either

$$\max\{|\mathcal{C}_1^+|, |\mathcal{C}_1^-|\} \leq \frac{B}{2} n^{3/4},$$

or

$$\epsilon \min\{|\mathcal{C}_1^+(t)|, |\mathcal{C}_1^-(t)|\} + \sum_{j\geq 2} \epsilon_j |\mathcal{C}_j^+(t)| + \sum_{j\geq 2} \epsilon_j' |\mathcal{C}_j^-(t)| \leq -(\frac{B}{2} - a_2) n^{3/4}.$$

By Theorem 3.3.9, we have

$$\mathbb{P}(\max\{|\mathcal{C}_1^+|, |\mathcal{C}_1^-|\} \leq \frac{B}{2} n^{3/4}) \leq Ce^{-cn^{1/8}}.$$

By Theorem 3.3.13 and Markov's inequality we have

$$\mathbb{P}(\epsilon \min\{|\mathcal{C}_1^+(t)|, |\mathcal{C}_1^-(t)|\} + \sum_{j\geq 2} \epsilon_j |\mathcal{C}_j^+(t)| + \sum_{j\geq 2} \epsilon_j' |\mathcal{C}_j^-(t)| \leq O(n^{-1/3}).$$

Thus, we have

$$\mathbb{P}(X_1 \leq b) = O(n^{-1/3})\,.$$

□

Proof of Theorem 3.6.10: Denote by I the interval $[b_1n^{3/4}, b_2n^{3/4}]$ and let B be a large constant to be chosen later. For any $X_0 \in I$, define

$$A_{t,X_0} = \mathbb{P}(X_t \ \text{ exceeds } Bn^{3/4} \text{ within } t \text{ steps} \mid X_0).$$

Let

$$A_t = \max_{X_0 \in I} A_{t,X_0}.$$

Then A_t is increasing in t. Let

$$\tau = \min\{t : X_t \notin [\frac{b_1}{2}n^{3/4}, Bn^{3/4}]\}\,.$$

Then $X_{t\wedge\tau}$ is a supermartingale by part (i) of Lemma 3.6.7. Thus we have

$$\mathbb{E}X_{Kn^{1/4}\wedge\tau} \leq b_2n^{3/4}. \tag{3.134}$$

For simplicity denote $g(B) = \max_{X_0\in I}\mathbb{P}(X_{Kn^{1/4}\wedge\tau} \geq Bn^{3/4} \mid X_0)$. We get from the above estimate and (3.134) that $g(B) \to 0$ as $B \to \infty$. For all $X_0 \in I$ and $t \leq Kn^{1/4}$, we have

$$A_{t,X_0} \leq g(B, X_0) + \mathbb{P}\Big(X_{Kn^{1/4}\wedge\tau} \leq \frac{b_1}{2}n^{3/4}, X_t \text{ exceeds } Bn^{3/4} \text{ before } \ t\Big). \tag{3.135}$$

Denote

$$(\frac{t}{n}\wedge 1) = \big\{X_{Kn^{1/4}\wedge\tau} \leq \frac{b_1}{2}n^{3/4}, X_t \text{ exceeds } Bn^{3/4} \text{ before } \ t\big\}.$$

Let τ_1 be the exit time of $[\frac{b_1}{2}n^{3/4}, (b_2+\frac{b_1}{2})n^{3/4}]$. By Theorem 3.6.18, we have

$$\mathbb{P}(\tau_1 > \delta n^{1/4}) \geq 1 - C\delta^2,$$

for any sufficiently small constant $\delta > 0$. On the event $\{\tau_1 > \delta n^{1/4}\}$, there are three cases:

(i) $X_{\delta n^{1/4}} \in [\frac{b_1}{2}n^{3/4}, b_1n^{3/4}]$.

(ii) $X_{\delta n^{1/4}} \in [b_1n^{3/4}, b_2n^{3/4}]$.

(iii) $X_{\delta n^{1/4}} \in [b_2n^{3/4}, (b_2+\frac{b_1}{2})n^{3/4}]$.

For case (ii), by the Markov property at time $\delta n^{1/4}$, we have that

$$\mathbb{P}((\frac{t}{n}\wedge 1) \,\big|\, \tau_1 > \delta n^{1/4}, X_{\delta n^{1/4}} \in [b_1n^{3/4}, b_2n^{3/4}]) \leq A_{t-\delta n^{1/4}}.$$

For case (i), define

$$T = \min\{t > \delta n^{1/4} : X_t \in [b_1n^{3/4}, b_2n^{3/4}]\}\,.$$

By monotonicity of A_t and the strong Markov property on T we have

$$\mathbb{P}\big((\frac{t}{n}\wedge 1) \,\big|\, \tau_1 > \delta n^{1/4}, X_{\delta n^{1/4}} \in [\frac{b_1}{2}n^{3/4}, b_1n^{3/4}], T < t\big) \leq A_{t-\delta n^{1/4}}.$$

The event $\{(\frac{t}{n} \wedge 1), \tau_1 > \delta n^{1/4}, X_{\delta n^{1/4}} \in [\frac{b_1}{2} n^{3/4}, b_1 n^{3/4}], T \geq t\}$ implies that there exists $t \leq K n^{1/4}$ such that $X_t < \frac{b_1}{2} n^{3/4}$ and $X_{t+1} > b_2 n^{3/4}$. By Lemma 3.6.9 and the union bound, this happens with probability at most

$$D n^{-\frac{1}{3}} K n^{1/4} = O(n^{-1/12}).$$

For case (iii), the event $\big\{(\frac{t}{n} \wedge 1), \tau_1 > \delta n^{1/4}, X_{\delta n^{1/4}} \in [b_2 n^{3/4}, (b_2 + \frac{b_1}{2}) n^{3/4}]\big\}$ implies that X_t first goes below $\frac{b_1}{2} n^{3/4}$ and then goes above $B n^{3/4}$. Let

$$T' = \min\{t : t > \tau, X_t \in I\}.$$

By monotonicity of A_t and the strong Markov property on T', we obtain

$$\mathbb{P}\big(\mathcal{A} \,\big|\, \tau_1 > \delta n^{1/4}, X_{\delta n^{1/4}} \in [b_2 n^{3/4}, (b_2 + \frac{b_1}{2}) n^{3/4}], T' < t\big) \leq A_{t - \delta n^{1/4}}.$$

By similar argument in case (ii), we have

$$\mathbb{P}\big(\mathcal{A}, \tau_1 > \delta n^{1/4}, X_{\delta n^{1/4}} \in [b_2 n^{3/4}, (b_2 + \frac{b_1}{2}) n^{3/4}], T' \geq t\big) = O(n^{-1/12}).$$

Summing up the above estimates, we obtain

$$\mathbb{P}((\frac{t}{n} \wedge 1), \tau_1 > \delta n^{1/4}) \leq A_{t - \delta n^{1/4}} + O(n^{-1/12}). \tag{3.136}$$

On the event $\{(\frac{t}{n} \wedge 1), \tau_1 \leq \delta n^{1/4}\}$, which happens with probability at most $C\delta^2$, there are two cases to consider:

(i) $X_{\tau_1} < \frac{b_1}{2} n^{3/4}$,

(ii) $X_{\tau_1} > (b_2 + \frac{b_1}{2}) n^{3/4}$.

In case (i), let

$$T_1 = \min\{t : t > \tau_1, X_t \in I\}.$$

By monotonicity of A_t and the strong Markov property on T_1, we have

$$\mathbb{P}((\frac{t}{n} \wedge 1) \,|\, \tau_1 \leq \delta n^{1/4}, X_{\tau_1} < \frac{b_1}{2} n^{3/4}, T_1 < t) \leq A_t.$$

A similar argument as before gives us

$$\mathbb{P}((\frac{t}{n} \wedge 1), \tau_1 \leq \delta n^{1/4}, X_{\tau_1} < \frac{b_1}{2} n^{3/4}, T_1 \geq t) = O(n^{-1/12}).$$

In case (ii), let

$$T_2 = \min\{t : t > \tau, X_t \in I\}.$$

Similar arguments gives

$$\mathbb{P}((\frac{t}{n} \wedge 1) | \tau_1 \leq \delta n^{1/4}, X_{\tau_1} < \frac{b_1}{2} n^{3/4}, T_2 < t) \leq A_t,$$

and

$$\mathbb{P}((\frac{t}{n}\wedge 1), \tau_1 \le \delta n^{1/4}, X_{\tau_1} < \frac{b_1}{2}n^{3/4}, T_2 \ge t) = O(n^{-1/12}).$$

Summing over these estimate, we obtain

$$\mathbb{P}((\frac{t}{n}\wedge 1), \tau_1 \le \delta n^{1/4}) \le C\delta^2(A_t + O(n^{-1/12})). \tag{3.137}$$

Plugging (3.136) and (3.137) into (3.135), we get

$$A_{t,X_0} \le g(B) + (A_{t-\delta n^{1/4}} + O(n^{-1/12})) + C\delta^2(A_t + O(n^{-1/12})).$$

Maximizing over X_0 and rearranging give

$$A_t \le \frac{1}{1-C\delta^2}(A_{t-\delta n^{1/4}} + g(B) + O(n^{-1/12})).$$

Telescoping gives

$$A_{Kn^{1/4}} \le \frac{1}{1-C\delta^2}^{\left\lceil \frac{K}{\delta}\right\rceil}\Big(C\delta^2 + g(B) + O(n^{-1/12})\Big).$$

Since $\frac{1}{1-C\delta^2}^{\left\lceil \frac{K}{\delta}\right\rceil}$ converges as δ goes to 0, we conclude that we can choose $\delta > 0$ small enough and B so large to make $A_{Kn^{1/4}}$ arbitrarily small, as required. □

Proof of Theorem 3.6.11: Notice that

$$\begin{aligned}\mathbb{E}\Big[X_{(t+1)\wedge\tau_a}\Big|\mathcal{F}_t\Big] &= \mathbb{E}\Big[X_{t+1}\mathbf{1}_{\{\tau_a\ge t+1\}} + X_{\tau_a}\mathbf{1}_{\{\tau_a\le t\}}\Big|\mathcal{F}_t\Big] \\ &= \mathbb{E}[X_{t+1}|\mathcal{F}_t]\mathbf{1}_{\{\tau_a\ge t+1\}} + X_{\tau_a}\mathbf{1}_{\{\tau_a\le t\}}.\end{aligned}$$

By Lemma 3.6.16, we have

$$\mathbb{E}\Big[X_{(t+1)\wedge\tau_a}\Big|\mathcal{F}_t\Big] \le X_t\Big(1-\frac{X_t}{6n}\Big)\mathbf{1}_{\{\tau_a\ge t+1\}} + X_{\tau_a}\mathbf{1}_{\{\tau_a\le t\}} = X_{t\wedge\tau_a} - \frac{X_t^2}{6n}\mathbf{1}_{\{\tau_a\ge t+1\}}. \tag{3.138}$$

Taking expectations on both sides of (3.138), we get

$$\mathbb{E}X_{(t+1)\wedge\tau_a} \le \mathbb{E}X_{t\wedge\tau_a} - \frac{1}{6n}\mathbb{E}X_t^2\mathbf{1}_{\{\tau_a\ge t+1\}}. \tag{3.139}$$

Note that

$$\mathbb{E}\Big(X_t^2\mathbf{1}_{\{\tau_a\ge t+1\}}\Big) \ge a^2n^{3/2}\mathbb{P}(\tau_a\ge t+1),$$

and

$$a^2n^{3/2} \ge \frac{\mathbb{E}\Big(X_{\tau_a}^2\mathbf{1}_{\{\tau_a\le t\}}\Big)}{\mathbb{P}(\tau_a\le t)}.$$

Hence we have

$$\mathbb{E}\Big(X_t^2\mathbf{1}_{\{\tau_a\ge t+1\}}\Big) \ge \frac{\mathbb{E}\Big(X_{\tau_a}^2\mathbf{1}_{\{\tau_a\le t\}}\Big)}{\mathbb{P}(\tau_a\le t)}\mathbb{P}(\tau_a\ge t+1),$$

which implies

$$\mathbb{E}\Big(X_{\tau_a}^2 \mathbf{1}_{\{\tau_a \le t\}}\Big) \le \frac{\mathbb{P}(\tau_a \le t)}{\mathbb{P}(\tau_a \ge t+1)} \mathbb{E}\Big(X_t^2 \mathbf{1}_{\{\tau_a \ge t+1\}}\Big).$$

Adding $\mathbb{E}\Big(X_t^2 \mathbf{1}_{\{\tau_a \ge t+1\}}\Big)$ to both sides, we obtain

$$\frac{\mathbb{E}\Big[X_t^2 \mathbf{1}_{\{\tau_a \ge t+1\}}\Big]}{\mathbb{P}(\tau_a \ge t+1)} \ge \mathbb{E}X_{t\wedge\tau_a}^2 \ge (\mathbb{E}X_{t\wedge\tau_a})^2. \tag{3.140}$$

Plugging into (3.139), we get

$$\mathbb{E}X_{(t+1)\wedge\tau_a} \le \mathbb{E}X_{t\wedge\tau_a} - \frac{1}{6n}\mathbb{P}(\tau_a \ge t+1)(\mathbb{E}X_{t\wedge\tau_a})^2. \tag{3.141}$$

Note that $\mathbb{E}X_{(t+1)\wedge\tau_a} > 0$. Taking the inverse of (3.141) leads to

$$\frac{1}{\mathbb{E}X_{(t+1)\wedge\tau_a}} \ge \frac{1}{\mathbb{E}X_{t\wedge\tau_a}} + \frac{1}{6n}\mathbb{P}(\tau_a \ge t+1).$$

Summing t from 0 to $\lceil bn^{1/4}\rceil - 1$, we get

$$\frac{1}{\mathbb{E}X_{\lceil bn^{1/4}\rceil\wedge\tau_a}} \ge \frac{1}{6n}\sum_{t=0}^{\lceil bn^{1/4}\rceil-1} \mathbb{P}(\tau_a \ge t+1) \ge \frac{1}{6n}\mathbb{P}(\tau_a \ge bn^{1/4})bn^{1/4}. \tag{3.142}$$

On the other hand, for any $x \in [0,n]$, observe that $X_{bn^{1/4}\wedge\tau_a} \le -x$ implies there exists $t \le bn^{1/4}$ such that $X_t > an^{3/4}$ and $X_{t+1} < -x$. This implies either

$$\max\{|\mathcal{C}_1^+|, |\mathcal{C}_1^-|\} \le \frac{a}{2}n^{3/4},$$

or

$$\epsilon \min\{|\mathcal{C}_1^+(t)|, |\mathcal{C}_1^-(t)|\} + \sum_{j\ge 2}\epsilon_j|\mathcal{C}_j^+(t)| + \sum_{j\ge 2}\epsilon_j'|\mathcal{C}_j^-(t)| \le -x - \frac{a}{2}n^{3/4}.$$

By Theorem 3.3.9, we have $\mathbb{P}(\max\{|\mathcal{C}_1^+|, |\mathcal{C}_1^-|\} \le \frac{a}{2}n^{3/4}) = O(e^{-cn^{1/8}})$. By Theorem 3.3.13 and Markov's inequality, we have

$$\mathbb{P}(\epsilon \min\{|\mathcal{C}_1^+(t)|, |\mathcal{C}_1^-(t)|\} + \sum_{j\ge 2}\epsilon_j|\mathcal{C}_j^+(t)| + \sum_{j\ge 2}\epsilon_j'|\mathcal{C}_j^-(t)| \le -x - \frac{a}{2}n^{3/4})) \le \frac{Cn^{8/3}}{(x+\frac{a}{2}n^{3/4})^4}.$$

Hence by union bound we obtain

$$\mathbb{P}(X_{bn^{1/4}\wedge\tau_a} \le -x) \le \frac{Cn^{8/3}}{(x+\frac{a}{2}n^{3/4})^4}bn^{1/4}.$$

By a direct computation we obtain

$$\mathbb{E}(|X_{bn^{1/4}\wedge\tau_a}|\mathbf{1}_{\{X_{bn^{1/4}\wedge\tau_a}\le 0\}}) \le \sum_{x=0}^{n} bn^{1/4}\frac{Cn^{8/3}}{(x+\frac{a}{2}n^{3/4})^4} = O(n^{2/3}).$$

Thus we get

$$\mathbb{E}X_{\lceil bn^{1/4}\rceil\wedge\tau_a} \geq \mathbb{P}(\tau_a > \lceil bn^{1/4}\rceil)an^{3/4} - O(n^{2/3}).$$

Multiplying this and (3.142) we get

$$1 \geq \frac{1}{6n}bn^{1/4}an^{3/4}\Big[\mathbb{P}(\tau_a > \lceil bn^{1/4}\rceil)\Big]^2,$$

which gives (3.101). $\square$

Coupling inside the scaling window: Proof of Theorem 3.6.4

Lemma 3.6.19. *For any fixed constant $A > 1$ there exist positive constants $q = q(A), \beta = \beta(A)$, such that if $X_0 \in [A^{-1}n^{3/4}, An^{3/4}]$, then*

$$\mathbf{P}(X_1 = x|X_0) \geq qn^{-5/8}$$

for any $x \in n + 2\mathbb{Z}$ and $|x - X_0| \leq \beta n^{5/8}$.

Proof of Theorem 3.6.4 We will use induction to prove that for any $\ell > 0$ and any $x \in n + 2\mathbb{Z}$ such that $|X_0 - x| \leq \beta(1/2 + \ell/2)n^{5/8}$, we have

$$\mathbf{P}(X_\ell = x|X_0) \geq q^\ell\Big(\frac{\beta}{2}\Big)^{\ell-1}n^{-5/8}. \tag{3.143}$$

This implies Theorem 3.6.4 immediately.

We prove this assertion by induction on ℓ. Lemma 3.6.19 implies (3.143) is true for $\ell = 1$. Suppose now (3.143) holds for ℓ and we prove for $\ell + 1$. If $x \in n + 2\mathbb{Z}$ and $|x - X_0| \leq \beta\Big(1/2 + (\ell+1)/2\Big)n^{5/8}$, then the number of y such that $y \in n + 2\mathbb{Z}$ and $|y - x| \leq \beta n^{5/8}$ and $|y - X_0| \leq \beta(1/2 + \ell/2)n^{5/8}$ is at least $\frac{\beta}{2}n^{5/8}$. Thus, we have

$$\begin{aligned}\mathbb{P}\Big(|X_\ell - x| \leq \beta n^{5/8}\Big) &= \sum_{|y-x|\leq\beta n^{5/8}} \mathbb{P}(X_\ell = y) \\ &\geq q^\ell\Big(\frac{\beta}{2}\Big)^{\ell-1}n^{-5/8}\frac{\beta}{2}n^{5/8} = q^\ell\Big(\frac{\beta}{2}\Big)^\ell,\end{aligned} \tag{3.144}$$

where we used the induction hypothesis. Since $|x - X_0| \leq \beta\Big(1/2 + (\ell+1)/2\Big)n^{5/8}$, we get

$$\mathbb{P}\Big(X_{\ell+1} = x\,\Big|\,|X_\ell - x| \leq \beta n^{5/8}\Big) \geq qn^{-5/8}$$

by Lemma 3.6.19. Together with (3.144) we get (3.143) for $\ell + 1$, concluding the proof. $\square$

Recall that conditioned on the cluster sizes, X_1 is a summation of independent but not identically distributed random variables. The following is a local central limit theorem for such sums, tailored to our particular needs, and is used to prove Lemma 3.6.19. We have not found in the literature a statement general enough to be valid in our setting. The proof is the standard proof of the local CLT using characteristic function.

Lemma 3.6.20. *Suppose K_n are positive integers such that $K_n \geq qn$ for some constant $q > 0$ and $a_1, a_2, \cdots, a_{K_n}$ are positive integers such that $a_j = 1$ for $1 \leq j \leq qn$ and $a_j \leq \sqrt{\frac{qn}{2}}$ for all j. Let $b(n) = \sum_{j=1}^{K_n} a_j$ and $c(n) = \sqrt{\sum_{j=1}^{K_n} a_j^2 / n^{5/4}}$. Assume that there are two positive constants δ and C such that $\delta < c(n) < C$ for all n. Let $X_n = \sum_{j=1}^{K_n} \epsilon_j a_j$ where $\{\epsilon_j\}$ is independent random $\pi_{(m)}$ signs. Then for any $x \in b(n) + 2\mathbf{Z}$ and large enough n, we have*

$$\mathbb{P}\Big(X_n = x\Big) \geq \frac{\sqrt{2}}{\sqrt{\pi} c(n) n^{5/8}} \Big(e^{-\frac{x^2}{2}} - 1/2\sqrt{2}\Big). \tag{3.145}$$

Proof of Lemma 3.6.19: We need to show that with probability $\Omega(1)$ the percolation configuration fits the setting of Lemma 3.6.20. Define $(\frac{t}{n} \wedge 1)_1$ and $(\frac{t}{n} \wedge 1)_2$ as the following events:

$$(\frac{t}{n} \wedge 1)_1 = \{|\mathcal{C}_1^+| \in \big[X_0 - \frac{c}{4} n^{5/8}, X_0 + \frac{c}{4} n^{5/8}\big], \sum_{j \geq 2} |\mathcal{C}_j^+|^2 \leq D n^{5/4}, \sum_{j \geq 2} |\mathcal{C}_j^+|^3 \leq D n^{7/4}\},$$

$$(\frac{t}{n} \wedge 1)_2 = \{\sum_{j \geq 1} |\mathcal{C}_j^-|^2 \leq D n^{5/4}, \sum_{j \geq 1} |\mathcal{C}_j^-|^3 \leq D n^{7/4}, \sum_{|\mathcal{C}_j^-| \leq \frac{\sqrt{n}}{6}} |\mathcal{C}_j^-|^2 \geq c^2 n^{5/4}, \big|\{j : |\mathcal{C}_j^-| = 1\}\big| \geq \frac{n}{18}\}$$

where D and c are constants to be selected later. First we prove that $(\frac{t}{n} \wedge 1)_1$ and $(\frac{t}{n} \wedge 1)_2$ both happen with probability $\Omega(1)$. To bound from below the probability of $(\frac{t}{n} \wedge 1)_2$, take $\delta = \frac{1}{3\sqrt{2}}$ in Theorem 3.3.14. We get

$$\mathbb{P}\Big(\sum_{|\mathcal{C}_j^-| \leq \frac{1}{6}\sqrt{n}} |\mathcal{C}_j^-|^2 \geq c n^{5/4}\Big) \geq q = q(A) > 0, \tag{3.146}$$

for some $c = c(A) > 0$. By Theorem 3.3.12, for $k = 2, 3$ we have

$$\mathbb{E} \sum_{j \geq 1} |\mathcal{C}_j^-|^k \leq C n (A^{-1} n^{-1/4})^{-2k+3}.$$

Thus, for

$$D \geq \frac{4CA^3}{q}, \tag{3.147}$$

we have by Markov's inequality that

$$\mathbb{P}\Big(\sum_{j \geq 1} |\mathcal{C}_j^-|^2 \geq D n^{5/4}\Big) \leq \frac{q}{4} \tag{3.148}$$

and

$$\mathbb{P}\Big(\sum_{j \geq 1} |\mathcal{C}_j^-|^3 \geq D n^{7/4}\Big) \leq \frac{q}{4}. \tag{3.149}$$

By Lemma 3.3.7, we have

$$\mathbb{P}\Big(\big|\{j : |\mathcal{C}_j^-| = 1\}\big| \geq \frac{n}{18}\Big) \geq 1 - C/n \geq 1 - \frac{q}{4}. \tag{3.150}$$

Putting (3.146), (3.148), (3.149) and (3.150) together, we get

$$\mathbb{P}((\frac{t}{n} \wedge 1)_2) \geq \frac{q}{4}.$$

To bound from below the probability of $(\frac{t}{n} \wedge 1)_1$, we apply Theorem 3.3.11 to get that

$$\mathbb{P}\Big(\Big||\mathcal{C}_1^+| - x_0(1 + \frac{x_0}{n})\Big| \leq \frac{c}{4}(\frac{n + x_0}{2})^{5/8}\Big) \geq q = q(A) > 0.$$

Since $\frac{x_0^2}{n} = o(n^{5/8})$ and $\frac{n+x_0}{2} \leq \frac{3}{4}n$, we get

$$\mathbb{P}\Big(|\mathcal{C}_1^+| \in \Big[x_0 - \frac{c}{4}n^{5/8},\, x_0 + \frac{c}{4}n^{5/8}\Big]\Big) \geq q. \tag{3.151}$$

By Theorem 3.3.13, for $k = 2, 3$ we have

$$\mathbb{E}\sum_{j \geq 2} |\mathcal{C}_j^+|^k \leq C_k n(A^{-1}n^{-1/4})^{-2k+3}.$$

Again, when D satisfies 3.147 we get by Markov's inequality that

$$\mathbb{P}\Big(\sum_{j \geq 2} |\mathcal{C}_j^+|^2 \geq Dn^{5/4}\Big) \leq \frac{q}{4} \tag{3.152}$$

and

$$\mathbb{P}\Big(\sum_{j \geq 2} |\mathcal{C}_j^+|^3 \geq Dn^{7/4}\Big) \leq \frac{q}{4}. \tag{3.153}$$

By (3.151), (3.152) and (3.153), we have

$$\mathbb{P}((\frac{t}{n} \wedge 1)_1) \geq \frac{q}{2}. \tag{3.154}$$

Since $(\frac{t}{n} \wedge 1)_1$ and $(\frac{t}{n} \wedge 1)_2$ are independent, we get

$$\mathbb{P}((\frac{t}{n} \wedge 1)_1 \cap (\frac{t}{n} \wedge 1)_2) \geq \frac{q^2}{8},$$

providing D satisfies (3.147). By Proposition 3.6.15 we have $\mathbb{P}(|C_1^-| \geq |C_1^+|) = O(e^{-c\log^2 n})$. Hence the event

$$(\frac{t}{n} \wedge 1) = \{(\frac{t}{n} \wedge 1)_1, (\frac{t}{n} \wedge 1)_2, |C_1^-| < |C_1^+|)\}\,,$$

occurs with probability $\Omega(1)$.

Next we prove that for every $x \in n + 2\mathbb{Z}$ and $|x - X_0| \leq \frac{c}{2}n^{5/8}$, there exist a constant $\delta > 0$ such that

$$\mathbb{P}(X_1 = x\,|(\frac{t}{n} \wedge 1)) \geq \delta n^{-5/8}\,, \tag{3.155}$$

which will conclude the proof. Denote

$$M_1 = |\mathcal{C}_1^+| + \sum_{j \geq 2} \epsilon_j |\mathcal{C}_j^+| + \sum_{|\mathcal{C}_j| > \sqrt{n}/6} \epsilon_j' |\mathcal{C}_j^-|$$

and

$$M_2 = \sum_{|\mathcal{C}_j| \le \sqrt{n}/6} \epsilon'_j |\mathcal{C}_j^-|.$$

Note that M_1 and M_2 are independent conditioned on $(\frac{t}{n} \wedge 1)$. We will first prove that there exist a constant $\alpha > 0$ such that

$$\mathbb{P}\Big(|M_1 - X_0| \le \frac{c}{2} n^{5/8} \big| (\frac{t}{n} \wedge 1)\Big) \ge \alpha. \tag{3.156}$$

On $(\frac{t}{n} \wedge 1)$ we have that

$$\sum_{j \ge 2} |\mathcal{C}_j^+|^2 + \sum_{|\mathcal{C}_j| > \sqrt{n}/6} |\mathcal{C}_j^-|^2 \le 2Dn^{5/4} \tag{3.157}$$

and

$$\sum_{j \ge 2} |\mathcal{C}_j^+|^3 + \sum_{|\mathcal{C}_j| > \sqrt{n}/6} |\mathcal{C}_j^-|^3 \le 2Dn^{7/4}. \tag{3.158}$$

If $\sum_{j \ge 2} |\mathcal{C}_j^+|^2 + \sum_{|\mathcal{C}_j| > \sqrt{n}/6} |\mathcal{C}_j^-|^2 \le \frac{c^2}{32} n^{5/4}$, by Markov's inequality, we have

$$\mathbb{P}\Big(\Big| \sum_{j \ge 2} \epsilon_j |\mathcal{C}_j^+| + \sum_{|\mathcal{C}_j| > \sqrt{n}/6} \epsilon'_j |\mathcal{C}_j^-| \Big| \le \frac{c}{4} n^{5/8}\Big) \ge 1/2. \tag{3.159}$$

Otherwise

$$\frac{|\mathcal{C}_j|}{\big(\sum_{j \ge 2} |\mathcal{C}_j^+|^2 + \sum_{|\mathcal{C}_j| > \sqrt{n}/6} |\mathcal{C}_j^-|^2\big)^{1/2}} > \epsilon$$

implies

$$|\mathcal{C}_j| \ge \frac{\epsilon c^2}{32} n^{5/8},$$

and since $\{|\mathcal{C}_j|\}$ also satisfy (3.158), we learn that the Lindeberg condition is satisfied. By Lindeberg-Feller theorem (see [10], (4.5)), we have

$$\mathbb{P}\Big(\Big| \sum_{j \ge 2} \epsilon_j |\mathcal{C}_j^+| + \sum_{|\mathcal{C}_j| > \sqrt{n}/6} \epsilon'_j |\mathcal{C}_j^-| \Big| \le \frac{c}{4} n^{5/8}\Big) \ge \alpha > 0. \tag{3.160}$$

Combining this and (3.159) yields (3.156).

To estimate M_2 let

$$b = \sum_{|\mathcal{C}_j| \le \sqrt{n}/6} |\mathcal{C}_j^-| \quad \text{and} \quad a = n^{-9/8} b^{1/2}.$$

By Lemma 3.6.20, for every $x \in b + 2\mathbb{Z}$, we have

$$\mathbb{P}\big(M_2 = x \big| (\frac{t}{n} \wedge 1)\big) \ge \frac{\sqrt{2}}{\sqrt{\pi} a n^{5/8}} \Big(e^{-\frac{x^2}{2a^2 n^{5/4}}} - 1/2\Big).$$

For all x such that $|x| \le cn^{5/8}$, we have

$$\frac{\sqrt{2}}{\sqrt{\pi}an^{5/8}}\Big(e^{-\frac{x^2}{2a^2n^{5/4}}} - 1/2\Big) \ge \frac{\sqrt{2}}{\sqrt{\pi}Dn^{5/8}}\Big(e^{-1/2} - 1/2\Big) \ge \delta n^{-5/8}$$

where δ is a constant. So for every $x \in b + 2\mathbb{Z}$ and $|x| \le cn^{5/8}$, we have

$$\mathbb{P}\big(M_2 = x\big|(\frac{t}{n} \wedge 1)\big) \ge \delta n^{-5/8}. \tag{3.161}$$

By (3.156) and (3.161), for every $x \in n + 2\mathbb{Z}$ with $|x - x_0| \le \frac{c_1}{2}n^{5/8}$, we have

$$\begin{aligned} & \mathbb{P}\big(M_1 + M_2 = x\big|(\frac{t}{n} \wedge 1)\big) \\ \ge\ & \mathbb{P}\Big(|M_1 - x_0| \le \frac{c_1}{2}n^{5/8}\,,\ M_2 = (x - x_0) - (M_1 - x_0)\big|(\frac{t}{n} \wedge 1)\Big) \\ \ge\ & \alpha\delta n^{-5/8}. \end{aligned}$$

This proves (3.155), which concludes the whole proof. □

To prove Lemma 3.6.20 we need the following two small assertions. The first is Exercise 3.2 of [10].

Lemma 3.6.21. *If $\mathbb{P}(X \in b + h\mathbf{Z}) = \mathbf{1}$, where b is a complex number and $h > 0$ is a real number. Then for any $x \in b + h\mathbf{Z}$, we have*

$$\mathbb{P}(X = x) = \frac{h}{2\pi}\int_{-\pi/h}^{\pi/h} e^{-itx}\Phi(t)\partial t,$$

where $\Phi(t)$ is the characteristic function of X.

Lemma 3.6.22. *For any x in $\mathbb{R}$, let $m(x)$ be the integer that is closest to x (if $x - \frac{1}{2}$ is an integer, then we put $m(x) = x - \frac{1}{2}$). Then for any x*

$$|\cos x| \le \exp\Big(-\frac{(x - m(\frac{x}{\pi})\pi)^2}{2}\Big).$$

Proof. Since $m(\frac{x}{\pi})\pi \in \{k\pi\}_{k\in\mathbf{Z}}$, we have $|\cos x| = |\cos(x - m(\frac{x}{\pi})\pi)|$. Also, we have $-\frac{\pi}{2} \le x - m(\frac{x}{\pi})\pi \le \frac{\pi}{2}$. Since $\cos x \le e^{-\frac{x^2}{2}}$ for all $x \in [-\frac{\pi}{2}, \frac{\pi}{2}]$ we have that

$$|\cos x| = \cos\Big(x - m(\frac{x}{\pi})\pi\Big) \le \exp\Big(-\frac{(x - m(\frac{x}{\pi})\pi)^2}{2}\Big).$$

□

Proof of Lemma 3.6.20: For simplicity we will abbreviate $c(n)$ by c. Let

$$d_j = \frac{a_j}{cn^{5/8}}.$$

Then we have $\sum_{j=1}^{K_n} d_j^2 = 1$ and $\frac{X_n}{cn^{5/8}} = \sum_{j=1}^{K_n} \epsilon_j d_j$. Since $a_j = O(n^{1/2})$, we have that $d_j = O(n^{-1/8})$. Thus it satisfies Lindeberg condition (see [10]). Consequently, we have that

$$\frac{X_n}{cn^{5/8}} \xrightarrow{d} N(0,1).$$

Denote the characteristic function of $\frac{X_n}{cn^{5/8}}$ by $\Phi_n(t)$. A straightforward computation gives that

$$\Phi_n(t) = \Big(\cos\frac{t}{cn^{5/8}}\Big)^{qn} \prod_{j=qn+1}^{K_n} \cos(td_j)\,, \tag{3.162}$$

and we have $\Phi_n(t) \to e^{-\frac{t^2}{2}}$ for all fixed $t \in \mathbb{R}$. Taking $h = \frac{2}{cn^{5/8}}$ in Lemma 3.6.21, for $x \in b(n) + 2\mathbf{Z}$ we have

$$\mathbb{P}\Big(X_n = x\Big) = \frac{1}{\pi cn^{5/8}} \int_{-\frac{\pi}{2}cn^{5/8}}^{\frac{\pi}{2}cn^{5/8}} e^{-itx}\Phi_n(t)\partial t. \tag{3.163}$$

Let M be a large constant to be selected later. Note that $\Phi_n(t)$ is an even function so

$$\begin{aligned}
\int_{-\frac{\pi}{2}cn^{5/8}}^{\frac{\pi}{2}cn^{5/8}} e^{-itx}\Phi_n(t)\partial t &= \int_{-M}^{M} e^{-itx}\Phi_n(t)\partial t + 2\int_{M}^{\frac{\pi}{2}cn^{5/8}} e^{-itx}\Phi_n(t)\partial t \\
&\geq \int_{-M}^{M} e^{-itx}\Phi_n(t)\partial t - 2\int_{M}^{\frac{\pi}{2}cn^{5/8}} |\Phi_n(t)|\partial t.
\end{aligned} \tag{3.164}$$

We will first bound from above the second term of (3.164). Let $m_j(t) = m\Big(\frac{td_j}{\pi}\Big)\frac{\pi}{d_j}$, i.e., $m_j(t)$ is the element in $\Big\{k\frac{\pi}{d_j}\Big\}_{k\in\mathbf{Z}}$ that is closest to t. Note that by Lemma 3.6.22, we have

$$\cos(td_j) \leq \exp\Big\{-\Big[td_j - m(\frac{td_j}{\pi})\pi\Big]^2\Big/2\Big\} = \exp\Big\{-d_j^2\frac{(t-m_j(t))^2}{2}\Big\}.$$

For large enough n, we have $\frac{1}{c^2n^{5/4}} \geq \frac{1}{2}\frac{1}{c^2n^{5/4}-qn}$. Thus, we get

$$\Big|\cos(td_j)\Big| \leq \exp\Big\{-\frac{a_j^2}{c^2n^{5/4}-qn}\cdot\frac{(t-m_j(t))^2}{4}\Big\}.$$

Since $\sum_{j=qn+1}^{K_n} \frac{(a_j)^2}{c^2n^{5/4}-qn} = 1$ and e^{-x} is a convex function, we have by Jensen's inequality that

$$\begin{aligned}
\prod_{j=qn+1}^{K_n} \Big|\cos(td_j)\Big| &\leq \exp\Big\{-\sum_{j=qn+1}^{K_n} \frac{a_j^2}{c^2n^{5/4}-qn}\frac{(t-m_j(t))^2}{4}\Big\} \\
&\leq \sum_{j=qn+1}^{K_n} \frac{a_j^2}{c^2n^{5/4}-qn}\exp\Big(-\frac{(t-m_j(t))^2}{4}\Big).
\end{aligned} \tag{3.165}$$

Recall that $|t| \leq \frac{\pi}{2}cn^{5/8}$ and $|\cos(x)| \leq e^{-\frac{x^2}{2}}$ for $x \in [-\frac{\pi}{2}, \frac{\pi}{2}]$, whence

$$\Big|\cos\frac{t}{cn^{5/8}}\Big|^{qn} \leq \exp\Big(-\frac{qt^2}{2c^2n^{1/4}}\Big). \tag{3.166}$$

Plugging (3.165) and (3.166) into (3.162), we get

$$|\Phi_n(t)| \leq \sum_{j=qn+1}^{K_n} \frac{a_j^2}{c^2 n^{5/4} - qn} \exp\Big(- \frac{(t-m_j(t))^2}{4} - \frac{qt^2}{2c^2 n^{1/4}} \Big).$$

Hence, we have

$$\int_M^{\frac{\pi}{2}cn^{5/8}} |\Phi_n(t)| \partial t \leq \sum_{j=qn+1}^{K_n} \frac{a_j^2}{c^2 n^{5/4} - qn} \int_M^{\infty} \exp\Big(- \frac{(t-m_j(t))^2}{4} - \frac{qt^2}{2c^2 n^{1/4}} \Big) \partial t. \quad (3.167)$$

We will divide the integral into two parts such that the first part converges to 0 as M goes to infinity and the second part is bounded by a constant. Recall that $m_j(t) = 0$ for $t \in [-\frac{\pi}{2d_j}, \frac{\pi}{2d_j}]$, so for any $j \in [qn+1, K_n]$, we have

$$\begin{aligned} & \int_M^{\infty} \exp\Big(- \frac{(t-m_j(t))^2}{4} - \frac{qt^2}{2c^2 n^{1/4}} \Big) \partial t \\ = & \int_M^{\frac{\pi}{2d_j}} \exp\Big(- \frac{t^2}{4} - \frac{qt^2}{2c^2 n^{1/4}} \Big) \partial t \\ + & \sum_{\ell=1}^{\infty} \int_{\frac{\pi}{2d_j}(2\ell-1)}^{\frac{\pi}{2d_j}(2\ell+1)} \exp\Big(- \frac{(t-m_j(t))^2}{4} - \frac{qt^2}{2c^2 n^{1/4}} \Big) \partial t \end{aligned} \quad (3.168)$$

The first term of the right hand side of (3.168) is bounded by $\int_M^\infty e^{-\frac{t^2}{4}} \partial t$. For the second term, note that for $t \geq \frac{\pi}{2d_j}(2\ell-1)$, we have

$$\exp\Big(- \frac{qt^2}{2c^2 n^{1/4}} \Big) \leq \exp\Big(- \frac{q\pi^2 n (2\ell-1)^2}{8a_j^2} \Big)$$

and

$$\begin{aligned} \int_y^{y+\frac{\pi}{d_j}} e^{-\frac{1}{4}(t-m_j(t))^2} \partial t &= \int_{-\frac{\pi}{2d_j}}^{\frac{\pi}{2d_j}} e^{-\frac{1}{4}(t-m_j(t))^2} \partial t \\ &= \int_{-\frac{\pi}{2d_j}}^{\frac{\pi}{2d_j}} e^{-\frac{t^2}{4}} \partial t \leq 2\sqrt{\pi}. \end{aligned} \quad (3.169)$$

for any y, since $\frac{1}{4}(t-m_j(t))^2$ is a periodic function. Thus, we get

$$\int_M^{\infty} \exp\Big(- \frac{(t-m_j(t))^2}{4} - \frac{qt^2}{2c^2 n^{1/4}} \Big) \partial t \leq \int_M^\infty e^{-\frac{t^2}{4}} \partial t + \sum_{\ell=1}^{\infty} 2\sqrt{\pi} \exp\Big(- \frac{q\pi^2 n(2\ell-1)^2}{8a_j^2} \Big). \quad (3.170)$$

Recall that $a_j \leq \sqrt{qn/2}$, hence

$$\begin{aligned} \sum_{\ell=1}^{\infty} \exp\Big(- \frac{q\pi^2 n(2\ell-1)^2}{8a_j^2} \Big) &\leq \sum_{\ell=1}^{\infty} \exp\Big(- \frac{\pi^2 (2\ell-1)^2}{4} \Big) \\ &\leq \frac{e^{-\pi^2/4}}{1-e^{-\pi^2/2}} \leq \frac{1}{8}. \end{aligned}$$

Plugging into (3.170), we get

$$\int_M^\infty \exp\Big(-\frac{(t-m_j(t))^2}{4}-\frac{qt^2}{2c^2n^{1/4}}\Big)\partial t \le 2\sqrt{\pi}\Big(1-\Phi(\frac{M}{\sqrt{2}})\Big)+\frac{\sqrt{\pi}}{4},$$

where $\Phi(\cdot)$ is the distribution function of $N(0,1)$. Plugging back into (3.167), we get

$$\int_M^{\frac{\pi}{2}cn^{5/8}} |\Phi_n(t)|\partial t \le 2\sqrt{\pi}\Big(1-\Phi(\frac{M}{\sqrt{2}})\Big)+\frac{\sqrt{\pi}}{4}. \tag{3.171}$$

Now we go back to the first term of the right hand side of (3.164). Recall that $\Phi_n(t)$ converge to $e^{-t^2/2}$ for all t. We have the following estimate:

$$\begin{aligned}
& \int_{-M}^{M} e^{-itx}\Phi_n(t)\partial t \\
= \; & \int_{-\infty}^{\infty} e^{-itx}e^{-t^2/2}\partial t - \Big(\int_{-\infty}^{-M}+\int_M^\infty\Big)e^{-itx}e^{-t^2/2}\partial t + \int_{-M}^{M} e^{-itx}(\Phi_n(t)-e^{-t^2/2})\partial t \\
\ge \; & \sqrt{2\pi}e^{-x^2/2} - 2\int_M^\infty e^{-t^2/2}\partial t - \int_{-M}^{M}\big|\Phi_n(t)-e^{-t^2/2}\big|\partial t.
\end{aligned} \tag{3.172}$$

Note that the second term of the left most side of (3.172) converges to 0 as $M\to\infty$ and for fixed M we have $\int_{-M}^{M}|\Phi_n(t)-e^{-\frac{t^2}{2}}|\partial t \to 0$ by the Dominated Convergence Theorem. Plugging these and (3.171) into (3.164), we get

$$\liminf_{n\to\infty}\int_{-\frac{\pi}{2}cn^{5/8}}^{\frac{\pi}{2}cn^{5/8}} e^{-itx}\Phi_n(t)\partial t \ge \sqrt{2\pi}e^{-x^2/2}-\frac{\sqrt{\pi}}{2},$$

which concludes the whole proof. □

3.6.2 Starting at the $[0, n^{3/4}]$ regime: Proof of Theorem 3.6.2

Theorem 3.6.23. *Let $I=[-An^{2/3}, An^{2/3}]$ where A is a fixed large constant.Then there exist positive constants K, a, q such that*

$$\mathbb{P}\Big(\tau_a \le Kn^{1/4}\,|\,X_0\in I\Big)\ge q \tag{3.173}$$

where $\tau_a=\inf\{t\ge 0: X_t\ge an^{3/4}\}$.

Theorem 3.6.24. *For a constant A put $I=[-An^{2/3},\, An^{2/3}]$ and $\tau=\inf\{t\ge 0: X_t\in I\}$. Then there exist constant $c>0$ such that for sufficiently large A, we have*

$$\mathbb{P}(\tau>t)\le\frac{2|X_0|}{ct\sqrt{n}}.$$

Proof of Theorem 3.6.2: Let A, c be constants such that the assertion of Theorem 3.6.24 holds and write $I = [-An^{2/3}, An^{2/3}]$. Since $X_0 \leq n^{3/4}$, by Lemma 3.6.24 with $t = \frac{4n^{1/4}}{c}$, we have

$$\mathbb{P}(\tau \leq \frac{4n^{1/4}}{c}) > \frac{1}{2}$$

where $\tau = \min\{t : X_t \in I\}$. By Theorem 3.6.23 and the strong Markov property, we get that X_t exceeds $an^{3/4}$ within $(K + \frac{4}{c})n^{1/4}$ steps with probability at least $\frac{q}{2}$. By Theorem 3.6.1, we can couple X_t and with the stationary chain Y_t within $O(n^{1/4})$ steps such that they meet each other with probability $\Omega(1)$. Applying Lemma 3.1.4 concludes the proof. □

We now proceed to the proof of Theorem 3.6.23. We begin with some lemmas.

Lemma 3.6.25. *Let A be a large constant and put $I = [-An^{2/3}, An^{2/3}]$. For any $q \in (0, 1)$, there exist a state $Z = Z(q) \in I$ and constants $a = a(q) > 0$, $K = K(q) > 0$ such that*

$$\mathbb{P}\Big(\tau_a \leq Kn^{1/4} | X_0 = Z\Big) \geq q \tag{3.174}$$

where $\tau_a = \inf\{t \geq 0 : X_t \geq an^{3/4}\}$.

Proof. Let Y_t be a SW chain with $Y_0 \overset{d}{=} \pi_n$. By Theorem 3.6.6, there exists a constant B such that

$$\pi_n([B^{-1}n^{3/4},\, Bn^{3/4}]) \geq 1 - \frac{1-q}{4}. \tag{3.175}$$

We will prove the lemma for $a = B^{-1}$ and $K = \frac{6B}{c}$ where c is the constant in Theorem 3.6.24. Write $J = [B^{-1}n^{3/4},\, Bn^{3/4}]$, then by (3.175) we have

$$\mathbb{P}\Big(Y_0 \in J \text{ and } Y_{Kn^{1/4}} \in J\Big) \geq 1 - \frac{1-q}{2}. \tag{3.176}$$

Put $\tau = \inf\{t \geq 0 : Y_t \in I\}$. We have

$$\mathbb{P}(\tau \leq Kn^{1/4} | Y_0 \in J) \geq 1 - \frac{2Bn^{3/4}}{c\sqrt{n}Kn^{1/4}} = \frac{2}{3}$$

by Theorem 3.6.24. Thus we have

$$\mathbb{P}\Big(Y_0 \in J,\, \tau \leq Kn^{1/4}\Big) \geq \Big(1 - \frac{1-q}{4}\Big)\frac{2}{3} \geq \frac{1}{2}.$$

Let

$$\delta = \max_{W \in I} \mathbb{P}(\tau_a \leq Kn^{1/4} | X_0 = W).$$

By the strong Markov property

$$\begin{aligned} &\mathbb{P}\Big(Y_0 \in J,\, \tau \leq Kn^{1/4},\, Y_{Kn^{1/4}} \leq an^{3/4}\Big) \\ = \;&\mathbb{P}\Big(Y_0 \in J,\, \tau \leq Kn^{1/4}\Big)\mathbb{P}\big(Y_{Kn^{1/4}} \leq an^{3/4} | \tau, Y_\tau\big) \geq \frac{1-\delta}{2}\,, \end{aligned}$$

since $\tau_a > Kn^{1/4}$ implies that $Y_{Kn^{1/4}} \leq an^{3/4}$. We deduce that

$$\mathbb{P}\Big(Y_0 \in J,\, Y_{Kn^{1/4}} \notin J\Big) \geq \frac{1-\delta}{2}. \tag{3.177}$$

Combining (3.176) and (3.177) we get that $\delta \geq q$, concluding our proof. $\square$

Lemma 3.6.26. *Consider the random graph $G(\frac{n+X_0}{2}, \frac{2}{n})$ where $X_0 \in [-An^{2/3}, An^{2/3}]$ for some large constant A. Then the intersection of the following events occurs with probability at least $\delta = \delta(A) > 0$:*

- $|\mathcal{C}_1| + |\mathcal{C}_2| \in [4An^{2/3}, 8An^{2/3}]$, *and* $|\mathcal{C}_2| > \frac{4A}{3}n^{2/3}$,
- $\mathcal{C}_1$ *and* $\mathcal{C}_2$ *are trees, and*
- $\sum_{j\geq 3} |\mathcal{C}_j|^2 \leq n^{4/3}$.

Proof. Let $(\frac{t}{n} \wedge 1)$ be the event

- $|\mathcal{C}_1| + |\mathcal{C}_2| \in [4An^{2/3}, 8An^{2/3}], |\mathcal{C}_2| > \frac{4A}{3}n^{2/3}$,
- $\mathcal{C}_1$ and $\mathcal{C}_2$ are trees.

By Theorem 5.20 of [18], we have $\mathbb{P}((\frac{t}{n} \wedge 1)) \geq \delta = \delta(A) > 0$. Conditioned on $(\frac{t}{n} \wedge 1)$ and on $\mathcal{C}_1$ and $\mathcal{C}_2$ the remaining graph, $\{\mathcal{C}_j\}_{j\geq 3}$, is distributed as $G(\frac{n+X_0}{2} - |\mathcal{C}_1| - |\mathcal{C}_2|, \frac{2}{n})$ conditioned to the event that it does not have components larger than $|\mathcal{C}_2|$. By Theorem 7 of [30] the complement of this event has probability decaying exponentially in A. Let $\{\mathcal{C}'_j\}$ be the component size in the unconditioned space $G(\frac{n+X_0}{2} - |\mathcal{C}_1| - |\mathcal{C}_2|, \frac{2}{n})$. We have

$$\mathbb{E}\sum_{j\geq 1} |\mathcal{C}'_j|^2 = \big(\frac{n+X_0}{2} - |\mathcal{C}_1| - |\mathcal{C}_2|\big)\mathbb{E}|\mathcal{C}(v)|\,.$$

Since $X_0 \leq An^{2/3}$ and $|\mathcal{C}_1| + |\mathcal{C}_2| \geq 4An^{2/3}$, Theorem 7 of [30] gives that

$$\mathbb{E}|\mathcal{C}(v)| \leq O(e^{-cA})n^{1/3}\,,$$

and so

$$\mathbb{E}\sum_{j\geq 1} |\mathcal{C}'_j|^2 \leq O(e^{-cA})n^{4/3}\,.$$

The lemma now follows since in the conditioned space, the event we condition on has probability exponentially close to 1. $\square$

Lemma 3.6.27. *Let $I = [-An^{2/3}, An^{2/3}]$ for some large A. There exist a constant $c = c(A) > 0$ such that*

$$\mathbf{P}(X_1 = x \mid X_0 \in I) \geq cn^{-2/3} \tag{3.178}$$

for any $x \in n + 2\mathbf{Z}$ with $x \in I$ and $x > 0$.

Proof. Write $(\frac{t}{n} \wedge 1)$ for the event of the assertion of Lemma 3.6.26 in $\{\mathcal{C}_j^+\}$, so that $\mathbb{P}((\frac{t}{n} \wedge 1)) \geq \delta(A) > 0$. In $G(\frac{n-|X_0|}{2}, \frac{2}{n})$ we have by Theorem 3.3.13 that

$$\mathbb{E}\sum_{j\geq 1} |\mathcal{C}_j^-|^2 \leq Dn^{4/3}$$

where $D = D(A)$ is a constant. We have by Markov's inequality that

$$\mathbb{P}\Big(\big|\sum_{j\geq 3} \epsilon_j |\mathcal{C}_j^+|\big| \leq Dn^{2/3} \big|(\frac{t}{n} \wedge 1)\Big) \geq 1 - 1/D^2\,,$$

and

$$\mathbb{P}\big(\big|\sum_{j\geq 1} \epsilon_j |\mathcal{C}_j^-|\big| \leq Dn^{2/3}\big) \geq 1 - 1/D^2\,,$$

and these two events are independent. Thus, the following event which we denote by $\mathcal{B}$ happens with probability $\Omega(1)$.

- $|\mathcal{C}_1^+| + |\mathcal{C}_2^+| \in [4An^{2/3}, 8An^{2/3}], |\mathcal{C}_2^+| > \frac{4A}{3}n^{2/3}$,
- $\mathcal{C}_1^+$ and $\mathcal{C}_2^+$ are trees,
- $\big|\sum_{j\geq 3} \epsilon_j |\mathcal{C}_j^+| + \sum_{j\geq 1} \epsilon_j |\mathcal{C}_j^-|\big| \leq 2Dn^{2/3}$.

Note that if a negative spin is assigned to $\mathcal{C}_2^+$ then

$$X_1 = |\mathcal{C}_1^+| - |\mathcal{C}_2^+| + \sum_{j\geq 3} \epsilon_j |\mathcal{C}_j^+| + \sum_{j\geq 1} \epsilon_j |\mathcal{C}_j^-|.$$

Thus

$$\mathbb{P}(X_1 = x) \geq \frac{1}{2}\mathbb{P}\big(|\mathcal{C}_1^+| - |\mathcal{C}_2^+| + \sum_{j\geq 3} \epsilon_j |\mathcal{C}_j^+| + \sum_{j\geq 1} \epsilon_j |\mathcal{C}_j^-| = x\big).$$

So we only need to show that for any $x \in [-An^{2/3}, An^{2/3}]$ we have

$$\mathbb{P}(|\mathcal{C}_1^+| - |\mathcal{C}_2^+| = x \mid \mathcal{B}) \geq cn^{-2/3}, \tag{3.179}$$

for some constant $c = c(A) > 0$. For any $m \in [4An^{2/3}, 8An^{2/3}]$ let $l = \frac{m+x}{2}$. By Cayley's formula we have that

$$\begin{aligned}
&\mathbb{P}(|\mathcal{C}_1^+| - |\mathcal{C}_2^+| = x \,\big|\, |\mathcal{C}_1^+| + |\mathcal{C}_2^+| = m, \mathcal{C}_1 \cup \mathcal{C}_2, \mathcal{B}) \\
= \;& \mathbb{P}(|\mathcal{C}_1^+| = l, |\mathcal{C}_2^+| = m - l \,\big|\, |\mathcal{C}_1^+| + |\mathcal{C}_2^+| = m, \mathcal{C}_1 \cup \mathcal{C}_2, \mathcal{B}) \\
= \;& \frac{\binom{m}{l} l^{l-2} (m-l)^{(m-l)-2}}{\sum_{k=\frac{4A}{3}n^{2/3}}^{m/2} \binom{m}{k} k^{k-2} (m-k)^{(m-k)-2}}.
\end{aligned} \tag{3.180}$$

Let

$$a(k) = \binom{m}{k} k^{k-2}(m-k)^{(m-k)-2}.$$

By Stirling's formula, there are two constants c and C such that for large enough n and any $k_1, k_2 \in [\frac{4A}{3}n^{2/3}, \frac{m}{2}]$, we have

$$c \leq \frac{a(k_1)}{a(k_2)} \leq C.$$

This implies

$$\frac{\binom{m}{k} k^{k-2}(m-k)^{(m-k)-2}}{\sum_{k=\frac{4A}{3}n^{2/3}}^{m/2} \binom{m}{k} k^{k-2}(m-k)^{(m-k)-2}} \geq cn^{-2/3}$$

which proves (3.179). □

Proof of Theorem 3.6.23: Let A be large and $q \in (0,1)$ will be chosen later very close to 1. Let Z be the site and $K > 0$ the number satisfying the assertion of Lemma 3.6.25. Let $\{\widetilde{X}_t\}$ be an independent SW chain starting at Z and $\widetilde{\tau}_a$ is as in Lemma 3.6.25. Then we have

$$\mathbb{P}\Big(\widetilde{\tau}_a \geq Kn^{1/4} \mid \widetilde{X}_0 = Z\Big) \leq 1 - q\,. \tag{3.181}$$

Let $c > 0$ be the constant from Lemma 3.6.27. This lemma implies that we can couple X_t and $\widetilde{X}_t$ such that $X_1 = \widetilde{X}_1$ with probability at least c. From that point we can couple such that the two processes stay together with probability 1.

$$\mathbb{P}(X_t = \widetilde{X}_t \text{ for } t \geq 1) \geq c\,. \tag{3.182}$$

Thus, we have

$$\mathbb{P}(\tau_a \leq Kn^{1/4}) \geq \mathbb{P}(X_t = \widetilde{X}_t \text{ for } t \geq 1 \text{ and } \widetilde{\tau}_a \leq Kn^{1/4}) \geq c - (1-q)\,,$$

so we choose $q \geq 1 - c/2$ and conclude the proof. □

To prove Theorem 3.6.24 we consider yet another modification of the SW dynamics $\{X'_t\}$. For any X'_0, in the supercritical random graph $G(\frac{n+|X_0|'}{2}, \frac{2}{n})$, let $\mathcal{C}_{\delta\epsilon n}$ be the component discovered by the exploration process at time $\delta\epsilon n$ where $\epsilon = \frac{X'_0}{n}$ and δ is a small constant (see Lemma 3.3.15). We assign positive spin to this component and random spins to all other components in $G(\frac{n+|X_0|'}{2}, \frac{2}{n})$ and all components in $G(\frac{n-|X_0|'}{2}, \frac{2}{n})$. Let X'_1 be the sum of spins after this assigning process.

The reason we require this change is that we were not able to obtain the bounds of Theorem 3.3.15 for $\mathcal{C}_1$, but only for $\mathcal{C}_{\delta\epsilon n}$ which is very likely to be $\mathcal{C}_1$. This will become evident in the proof. We first state a key lemma and then use it to prove Theorem 3.6.24.

Lemma 3.6.28. *For any constant A put $I = [-An^{2/3}, An^{2/3}]$. Then there exists a constant $c > 0$ such that for sufficiently large A we have*

$$\mathbb{E}\Big(|X'_1|\mathbf{1}_{\{X'_1 \notin I\}} + X'_1\mathbf{1}_{\{X'_1 \in I\}} \mid |X'_0| > An^{2/3}\Big) \leq |X'_0| - c\sqrt{n}\,. \tag{3.183}$$

Proof of Theorem 3.6.24: Notice that of $|X_0| = |X_0'|$ then $|X_1| \stackrel{d}{=} |X_1'|$, and so $|X_t| \stackrel{d}{=} |X_t'|$ for all $t \geq 1$. Thus, we only need to prove the assertion of the Theorem for $\{X_t'\}$. For simplicity of notation we write X_t for X_t'. Assume that $|X_0| > An^{2/3}$ otherwise the assertion is trivial. We begin by noticing that

$$\mathbb{E}\Big(|X_{t+1}|\mathbf{1}_{\{\tau>t+1\}} + X_\tau \mathbf{1}_{\{\tau\leq t+1\}}\Big|\mathcal{F}_t\Big)\mathbf{1}_{\{\tau\leq t\}} = \mathbb{E}(X_\tau \mathbf{1}_{\{\tau\leq t\}}|\mathcal{F}_t) = X_\tau \mathbf{1}_{\{\tau\leq t\}}\,. \tag{3.184}$$

By Lemma 3.6.28 we have

$$\begin{aligned}\mathbb{E}\Big(|X_{t+1}|\mathbf{1}_{\{\tau>t+1\}} + X_\tau \mathbf{1}_{\{\tau\leq t+1\}}\Big|\mathcal{F}_t\Big)\mathbf{1}_{\{\tau\geq t+1\}} &= \mathbb{E}\Big(|X_{t+1}|\mathbf{1}_{\{\tau>t+1\}} + X_\tau \mathbf{1}_{\{\tau= t+1\}}\Big|\mathcal{F}_t\Big) \\ &\leq |X_t|\mathbf{1}_{\{\tau>t\}} - c\sqrt{n}\mathbf{1}_{\{\tau>t\}}. \end{aligned} \tag{3.185}$$

Thus, We have that $\{X_t\}$ satisfies the following inequality:

$$\mathbb{E}\Big(|X_{t+1}|\mathbf{1}_{\{\tau>t+1\}} + X_\tau \mathbf{1}_{\{\tau\leq t+1\}}\Big|\mathcal{F}_t\Big) \leq |X_t|\mathbf{1}_{\{\tau>t\}} + X_\tau\mathbf{1}_{\{\tau\leq t\}} - c\sqrt{n}\mathbf{1}_{\{\tau>t\}}. \tag{3.186}$$

Taking expectations of both sides of (3.186), we get

$$\mathbb{E}\Big(|X_{t+1}|\mathbf{1}_{\{\tau>t+1\}} + X_\tau \mathbf{1}_{\{\tau\leq t+1\}}\Big) \leq \mathbb{E}\Big(|X_t|\mathbf{1}_{\{\tau>t\}} + X_\tau\mathbf{1}_{\{\tau\leq t\}}\Big) - c\sqrt{n}\mathbb{P}(\tau > t)\,.$$

Summing over t from 0 to $k-1$, we get

$$\begin{aligned}\mathbb{E}\Big(|X_k|\mathbf{1}_{\{\tau>k\}} + X_\tau \mathbf{1}_{\{\tau\leq k\}}\Big) &\leq |X_0| - \sum_{t=0}^{k-1} c\sqrt{n}\mathbb{P}(\tau > t) \\ &\leq |X_0| - kc\sqrt{n}\mathbb{P}(\tau > k). \end{aligned} \tag{3.187}$$

We also have

$$\mathbb{E}\Big(|X_k|\mathbf{1}_{\{\tau>k\}} + X_\tau \mathbf{1}_{\{\tau\leq k\}}\Big) \geq \mathbb{E}\Big(X_\tau \mathbf{1}_{\{\tau\leq k\}}\Big) \geq -An^{2/3} \geq -|X_0|.$$

Combining this with (3.187), we have

$$kc\sqrt{n}\mathbb{P}(\tau > k) \leq 2|X_0|$$

which implies the required result. □

Lemma 3.6.29. *Let X be a random variable. Then for any $b < 0$ and positive integer k, we have*

$$\mathbb{E}\Big(|X|\mathbf{1}_{(X\leq b)}\Big) \leq \frac{\mathbb{E}|X|^k}{|b|^{k-1}}.$$

Proof of Lemma 3.6.29: We have

$$\mathbb{E}|X|^k \geq \int_{-\infty}^{b} (-x)^k dF(x) \geq |b|^{k-1}\int_{-\infty}^{b}(-x)dF(x) = |b|^{k-1}\mathbb{E}|X|\mathbf{1}_{(X\leq b)}.$$

$\square$

Proof of Lemma 3.6.28: For simplicity write again X_t for X'_t. Notice that

$$|X_1|\mathbf{1}_{\{X_1\notin I\}} + X_1\mathbf{1}_{\{X_1\in I\}} = X_1 + 2|X_1|\mathbf{1}_{\{X_1<-An^{2/3}\}}\,.$$

We first bound $\mathbb{E}X_1$ from above. Recall that in our modified chain we have

$$\mathbb{E}X_1 = \mathbb{E}|\mathcal{C}_{\delta\epsilon n}|\,.$$

By part (i) of Theorem 3.3.15, for sufficiently large A and $|X_0| \geq An^{2/3}$, we have

$$\mathbb{E}|\mathcal{C}_{\delta\epsilon n}| \leq 2\frac{X_0}{n}\frac{n+X_0}{2} - c\Big(\frac{X_0}{n}\Big)^{-2} = X_0 + \frac{X_0^2}{n} - \frac{cn^2}{X_0^2}.$$

If $X_0 \leq \sqrt[4]{\frac{c}{2}}n^{3/4}$, then we have $\frac{X_0^2}{n} \leq \frac{cn^2}{2X_0^2}$. In this case we have

$$\mathbb{E}X_1 = \mathbb{E}|\mathcal{C}_{\delta\epsilon n}| \leq X_0 - \frac{cn^2}{2X_0^2}. \tag{3.188}$$

If $X_0 > \sqrt[4]{\frac{c}{2}}n^{3/4}$, then by Lemma 3.6.16, we have

$$\mathbb{E}X_1 \leq X_0 - \frac{X_0^2}{6n}. \tag{3.189}$$

Next we bound $\mathbb{E}|X_1|\mathbf{1}_{\{X_1<-An^{2/3}\}}$ from above. Let

$$M = \sum_{\mathcal{C}_j^+\neq\mathcal{C}_{\delta\epsilon n}^+} \epsilon_j|\mathcal{C}_j^+| + \sum_{j\geq 1}\epsilon'_j|\mathcal{C}_j^-|.$$

Then

$$X_1 = |\mathcal{C}_{\delta\epsilon n}| + M.$$

Since $|\mathcal{C}_{\delta\epsilon n}| > 0$, if $X_1 < -An^{2/3}$, then $M < -An^{2/3}$ and $|X_1| \leq -M$. Thus,

$$\mathbb{E}\Big(|X_1|\mathbf{1}_{\{X_1<-An^{2/3}\}}\Big) \leq \mathbb{E}\Big((-M)\mathbf{1}_{\{X_1<-An^{2/3}\}}\Big) \leq \mathbb{E}\Big((-M)\mathbf{1}_{\{M<-An^{2/3}\}}\Big).$$

Lemma 3.6.29 with $k=4$ gives

$$\mathbb{E}\Big((-M)\mathbf{1}_{\{M<-An^{2/3}\}}\Big) \leq \frac{\mathbb{E}M^4}{(An^{2/3})^3}. \tag{3.190}$$

We also have

$$\begin{aligned}\mathbb{E}M^4 \quad &\leq \quad \sum_{\mathcal{C}_j^+\neq\mathcal{C}_{\delta\epsilon n}^+}\mathbb{E}|\mathcal{C}_j^+|^4 + \sum_{j\geq1}\mathbb{E}|\mathcal{C}_j^-|^4 + 6\Big[\sum_{\mathcal{C}_j^+\neq\mathcal{C}_{\delta\epsilon n}^+}\mathbb{E}|\mathcal{C}_j^+|^2\Big]\Big[\sum_{j\geq1}\mathbb{E}|\mathcal{C}_j^-|^2\Big] \\ &+ \quad 6\Big[\sum_{\mathcal{C}_j^+,\mathcal{C}_i^+\neq\mathcal{C}_{\delta\epsilon n}^+,i\neq j}\mathbb{E}|\mathcal{C}_i^+|^2|\mathcal{C}_j^+|^2 + \sum_{i,j\geq1,i\neq j}\mathbb{E}|\mathcal{C}_i^-|^2|\mathcal{C}_j^-|^2\Big].\end{aligned} \tag{3.191}$$

By (ii) of Theorem 3.3.15 and Lemma 3.3.12, we have

$$\mathbb{E}M^4 = O\Big(\frac{n^6}{X_0^5}\Big) + O\Big(\frac{n^4}{X_0^2}\Big) = O\Big(\frac{n^4}{X_0^2}\Big) \tag{3.192}$$

since $|X_0| \geq An^{2/3}$. Plugging into (3.190), we have

$$\mathbb{E}\Big(|X_1|\mathbf{1}_{\{X_1<-An^{2/3}\}}\Big) = O\big(\frac{n^2}{A^3X_0^2}\big)\,. \tag{3.193}$$

If $X_0 \leq \sqrt[4]{\frac{c}{2}}n^{3/4}$, then combining (3.193) with (3.188) for large enough A, we get

$$\mathbb{E}X_1 + 2\mathbb{E}|X_1|\mathbf{1}_{\{X_1<-An^{2/3}\}} \leq X_0 - \frac{cn^2}{4X_0^2} \leq X_0 - c\sqrt{n}.$$

If $X_0 > \sqrt[4]{\frac{c}{2}}n^{3/4}$, then combining (3.193) with (3.189) for large enough A, we get

$$\mathbb{E}X_1 + 2\mathbb{E}|X_1|\mathbf{1}_{\{X_1<-An^{2/3}\}} \leq X_0 - \frac{X_0^2}{6n} + O(A^{-3}\frac{n^2}{X_0^2}) \leq X_0 - c\sqrt{n}.$$

Combining these two cases finishes the proof. □

3.6.3 The lower bound on the mixing time

Recall that in this section X_t is the original magnetization chain we defined in (3.3).

Proof of the lower bound of part (ii) of Theorem 3.1.1: Suppose X'_t is a modified magnetization chain and π' is its stationary distribution. By Theorem 3.6.6, we can choose an interval $[a_1n^{3/4},\, a_2n^{3/4}]$ with $0 < a_1 < a_2$ such that

$$\pi'(a_1n^{3/4}, a_2n^{3/4}) > \frac{3}{4}. \tag{3.194}$$

Suppose $X'_0 = 3a_2n^{3/4}$. By Theorem 3.6.18, there exists a constant k such that

$$\mathbb{P}(\tau > kn^{1/4}) \geq \frac{1}{2}$$

where τ is the first time that X_t exit $[a_2n^{3/4}, 4a_2n^{3/4}]$. This implies

$$\mathbb{P}\Big(X'_{kn^{1/4}} \geq a_2n^{3/4}\Big) \geq \frac{1}{2}\,.$$

By Theorem 3.6.6, we have $\pi'(-a_2n^{3/4}, -a_1n^{3/4})$ converges to 0 as n goes to infinity. Also, by (3.107), we have $\mathbb{P}(X'_{kn^{1/4}} \in [-a_2n^{3/4}, -a_1n^{3/4}]) = O(n^{-1/12})$. Combining these, we get that

$$\pi'[(a_1n^{3/4}, a_2n^{3/4}), (-a_2n^{3/4}, -a_1n^{3/4})] - \mathbb{P}(X'_{kn^{1/4}} \in [(a_1n^{3/4}, a_2n^{3/4}), (-a_2n^{3/4}, -a_1n^{3/4})]) > \frac{1}{4}$$

for large enough n. Recall that $X_t \overset{d}{=} |X'_t|$, so this is equivalent to

$$\pi(a_1n^{3/4}, a_2n^{3/4}) - \mathbb{P}(X_{kn^{1/4}} \in (a_1n^{3/4}, a_2n^{3/4})) > \frac{1}{4},$$

i.e.,

$$\Big\|X_{kn^{1/4}} - \pi\Big\|_{TV} > \frac{1}{4}. \tag{3.195}$$

□

3.7 Fast mixing of the Swendsen-Wang process on trees

In this section we provide an upper bound estimate of the mixing time of the Swendsen-Wang process on any tree with n vertices. We will prove in a more general setting for the Swendsen-Wang process for the q-state ferromagnetic Potts model. Recall that Ising model is the case $q = 2$.

For any given graph $G = (V, E)$, consider the set $S = \{0, 1\}^{|E|}$ of all edge configuration $\eta : E \to \{0, 1\}$. We consider the following Markov chain σ_t on S. At each step, we first color each component independently and uniformly from the q colors. Then we add all edges that connect vertices with the same color. Finally, delete each existing edge with probability $(1-p)$ to get a new state in S. It is easy to see that this process the dual of the Swendsen-Wang process for the q-state ferromagnetic Potts model on vertices configurations and the stationary distribution of σ_t is the random cluster model. For any two Swendsen-Wang chains, if we can couple the corresponding edge models so that they are the same(i.e., they have same clusters) at some time, we therefore couple the original Swendsen-Wang process at the same time. Consequently, any upper bound of the mixing time of this edge model implies the same upper bound on ferromagnetic Potts model.

There is an exploration process on trees to present σ_t. Notice that on trees each edge with state 0 connects two separate components. For any given $\eta \in S$, we color each components independently and uniformly from the q colors, starting from the root. We add edges connects vertices with the same color. Notice that this procedure is equivalent to setting every edge originally has configuration 0 with configuration 1 with probability $\frac{1}{q}$ and maintain configuration 0 otherwise. Thus, the process σ_t can be described as follows: First change each edge of 0 to 1 with probability $\frac{1}{q}$ and stay 0 otherwise, independently for each of them. Then, change each edge of 1 ,including those who have changed from 0 to 1 in the previous step, to 0 with probability $1 - p$, and stay 1 otherwise, independently for each of them. Each bit evolves independently as a Markov chain on $\{0, 1\}$, with transition matrix

$$\mathbb{P} = \begin{pmatrix} 1 - \frac{p}{Q} & \frac{p}{Q} \\ 1 - p & p \end{pmatrix}. \tag{3.196}$$

Proof of Theorem 3.1.2: The transition matrix (3.196) gives that we can couple every single edge with probability at least $1 - p + \frac{p}{q} \geq \frac{1}{q}$. Using the path coupling method of Bubley and Dyer (see Theorem 14.6 and Corollary 14.7 of [23]), we have

$$T_{\text{mix}} \leq \frac{\log n + \log 4}{-\log p(1 - \frac{1}{q})}.$$

$\square$

BIBLIOGRAPHY

[1] Aldous D. and Diaconis P. (1986), Shuffling cards and stopping times, Amer. Math. Monthly 93, no. 5, 333-348.

[2] Alon N. and Spencer J., The Probabilistic Method, third edition, Wiley, New York.

[3] Athreya K. B. and Ney P. E. (1972), Branching processes. Dover Publications, Inc., Mineola, NY.

[4] Bollobás B. (1984), The evolution of random graphs, *Trans. Amer. Math. Soc.*, **286**, 257–274.

[5] Borgs C., Chayes J. T., Frieze A., Kim J. H., Tetali P., Vigoda E. and Vu V. (1999), Torpid mixing of some MCMC algorithms in statistical physics, Proceedings of the 40th IEEE Symposium on Foundations of Computer Science (FOCS), 218-229.

[6] Bubley R. and Dyer M. (1997), Path Coupling: A technique for proving rapid mixing in Markov Chains, Proceedings of the 38th Annual Symposium on Foundation of Computer Science, pp. 223-231.

[7] Bollobás B. and Riordan O. (2011), Asymptotic normality of the size of the giant component via a random walk, preprint. Available at `http://arxiv.org/abs/1010.4595`

[8] Berestycki N., Schramm O., Zeitouni O. (2011), Mixing times for random k-cycles and coalescence-fragmentation chains, Annals of Probability, Vol. 39, No. 5, 1815-1843.

[9] Cooper C., Dyer M. E., Frieze A. M. and Rue R. (2000), *Mixing Properties of the Swendsen-Wang Process on the Complete Graph and Narrow Grids.* Journal of Mathematical Physics. 41: 1499–1527.

[10] Diaconis P., Group Representations in Probability and Statistics. IMS Lecture Notes - Monograph Series, 11 Institute of Mathematical Statistics, Hayward Ca.

[11] Durrett R. (1996), Probability: Theory and Examples. *Duxbury Press*, 2nd edition.

[12] Diaconis P. and Shahshahani M. (1981), Generating a random permutation with random transpositions, Z. Wahr. verw. Gebiete, 57 159-179.

[13] Diaconis P. and Saloff-Coste L. (1993), Comparison techniques for random walks on finite groups, Ann. Probab. 21 2131-2156.

[14] Edwards R. G. and Sokal A. D. (1988), *Generalizations of the Fortuin-Kasteleyn-Swendsen-Wang representation and Monte Carlo algorithm.* Physical Review D 38: 2009-2012.

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

[16] Gore V. and Jerrum M. R. (1996), *The Swendsen-Wang process does not always mix rapidly.* Proceedings of the 29th Annual ACM Symposium on Theory of Computing, (1997) 674-681.

[17] Hoeffding W. (1963), Probability inequalities for sums of bounded random variables. Journal of the American Statistical Association 58 (301): 13-30.

[18] Janson S., Luczak T. and Rucinski A. (2000), Random Graphs, Wiley, New York.

[19] Karp, R. M. (1990), The transitive closure of a random digraph. *Random Structures Algorithm* **1**, 73-93

[20] Lawler G. and Limic V. (2009) Random Walk: A Modern Introduction, Cambridge Studies in Advanced Mathematics (No. 123).

[21] Long Y., Nachmias A., Ning W., and Peres Y. (2012), A power law of order 1/4 for critical mean-field Swendsen-Wang dynamics, to appear in Memoir of American Mathematics Society.

[22] Levin D., Luczak M. and Peres Y (2010), Glauber dynamics for the mean-field Ising model: cut-off, critical power law, and metastability, *Probability Theory and Related Fields*, **146**, 223-265.

[23] Levin D., Peres Y. and Wilmer E. (2009), Markov Chains and mixing time, American Mathematical Society, Providence, RI. With a chapter by James G. Propp and David B. Wilson.

[24] Luczak T. (1990), Component behavior near the critical point of the random graph process, *Random Structures Algorithms*, **1**, 287-310.

[25] Mironov I. (2002), (Not So) Random Shuffles of RC4, Proceedings of CRYPTO, 304C319.

[26] Martin-Löf, A. (1986), Symmetric sampling procedures, general epidemic processes and their threshold limit theorems. *J. Appl. Probab.* **23**, 265-282

[27] Morris B., Ning W., and Peres Y. (2013), Mixting time of the Card-Cyclic to Random shuffle. Submitted.

[28] Mossel E., Peres Y. and Sinclair A. (2004), Shuffling by semi-random transpositions, Proceedings of the 45th Annual IEEE Symposium on Foundations of Computer Science(FOCS'04) October 17-19, 2004, Rome, Italy, 572-581, IEEE.

[29] Nachmias A. and Peres Y. (2007), Component sizes of the random graph outside the scaling window, *Latin American Journal of Probability and Mathematical Statistics (ALEA)*, **3**, 133-142.

[30] Nachmias A. and Peres Y. (2010), The critical random graph, with martingales, *Israel Journal of Math*, **176**, 29-43.

[31] Nachmias A and Peres Y. (2008), Critical percolation on random regular graphs, *Random structures and algorithms*, to appear.

[32] Pittel B. (1990), On Tree Census and the Giant Component in Sparse Random Graphs, Random Struct. Algorithms 1(3): 311-342.

[33] Pittel B. and Wormald N. (2005),Counting connected graphs inside-out, *J. Combinatorial Theory, Series B* **93**, 127-172.

[34] Pinsky R. (2011), Probabilistic and Combinatorial Aspects of the Card-Cyclic to Random Shuffle, preprint.

[35] Persky N., Ben-Av R., Kanter I. and Domany E. (1996), Mean-field behavior of cluster dynamics, *Phys. Rev. E* **54**, 2351-2358.

[36] Spitzer, F. (1956), A combinatorial lemma and its application to probability theory. Amer. Math. Soc. 82, 323-339.

[37] Subag E. (2011), A Lower Bound for the Mixing Time of the Random-to-Random Insertions Shuffle, preprint.

[38] Schramm O. (2005), Compositions of random transpositions, Israel Journal of Mathematics, vol 147, 221-244.

[39] Ray T., Tamayo P. and Klein W. (1989), Mean-field study of the Swendsen-Wang dynamics. *Phys. Rev. A* **39**, 5949-5953.

[40] Saloff-Coste L. and Zuniga J. (2007), Convergence of some time inhomogeneous Markov chains via spectral techniques. Stochastic Processes and their Applications 117, 961-979.

[41] Uyemura-Reyes J. (2002), Random Walk, semi-direct products, and card shuffling, Ph.D. Thesis, Stanford University.

[42] Wilson D. (2004), Mixing times of lozenge tiling and card shuffling Markov chains, The Annals of Applied Probability, 14(1),274-325.