Cutoff phenomenon for the simple exclusion process on the complete graph

Hubert Lacoin and Rémi Leblond

CEREMADE Université Paris Dauphine,
Place du Marchal De Lattre De Tassigny,
75775 Paris CEDEX 16 - France
E-mail address: `lacoin@ceremade.dauphine.fr`

Ecole Polytechnique,
Département de Mathématiques Appliquées,
91128 Palaiseau, France
E-mail address: `remi.leblond@polytechnique.edu`

Abstract. We study the time that the simple exclusion process on the complete graph needs to reach equilibrium in terms of total variation distance. For the graph with n vertices and $1 \ll k < n/2$ particles, we show that the mixing time is of order $\frac{1}{2} n \log \min(k, \sqrt{n})$, and that around this time, for any ε, the total variation distance drops from $1 - \varepsilon$ to ε in a time window whose width is of order n (i.e. in a much shorter time). Our proof is purely probabilistic and self-contained.

1. Introduction

Let $G = (V, E)$ be a finite connected graph and $1 \leqslant k \leqslant |V| - 1$ an integer. We define a *configuration* as an element of $\eta \in \{0, 1\}^V$ with k ones and $|V| - k$ zeros (ones can be considered as particles moving on the graph). The simple exclusion process on the graph G with k particles, can be described as follows: One starts with a given configuration and at each time step, one chooses an edge e uniformly at random in E and one interchanges the contents (zero or one) of the two vertices adjacent to e. This Markov chain is reversible and has the uniform measure over all configurations as equilibrium measure.

In this paper we study the rate of convergence to equilibrium of the exclusion process on the complete graph with n vertices, $k(n)$ particles, and k going to infinity with n. For obvious symmetry reasons, one can restrict the problem to the case $k \leqslant n/2$ without any loss of generality. For the sake of clarity we give a formal definition of the exclusion process

The configuration space is

$$\Omega(n, k(n)) := \Big\{ \eta \in \{0,1\}^{[1,n] \cap \mathbb{N}} \mid \sum_{x \in [1,n] \cap \mathbb{N}} \eta(x) = k(n) \Big\}.$$

We consider the discrete-time Markov chain on $\Omega(n, k(n))$ that at each time step independently selects two vertices uniformly at random in $\{1, \ldots, n\}$ and interchanges their contents (emptiness or particle). Note that with probability $1/n$ the two chosen vertices are the same: in that case nothing happens. We give now the transition kernel of this process: the symmetric group $\mathcal{S}_n$ acts transitively on $\Omega(n, k(n))$ in a natural way. If $\sigma \in \mathcal{S}_n$ then

$$(\sigma.\eta)(x) := \eta(\sigma^{-1}(x)),$$

and for $\eta, \eta' \in \Omega(n, k(n))$, $\eta \neq \eta'$, the transition rates are given by

$$P(\eta, \eta') = \begin{cases} \frac{2}{n^2} & \text{if } \eta' = \tau.\eta \text{ for some transposition } \tau \\ 0 & \text{if it is not the case} \end{cases}$$

One can check that this implies $P(\eta, \eta) \geqslant 1/2$.

This Markov chain is reversible, aperiodic and its equilibrium measure is the uniform measure on $\Omega(n, k(n))$ that we denote by π. Given an initial configuration $\xi \in \Omega(n, k(n))$, one writes $\mathbb{P}^\xi$ (and $\mathbb{E}^\xi$ denotes the associated expectation) for the law of the Markov chain $(\eta_t)_{t \geqslant 0}$ started from the configuration $\eta_0 = \xi$ and μ_t^ξ for the marginal distribution of $\mathbb{P}^\xi$ at time t ($t \in \mathbb{N}$).

We study the convergence to equilibrium of this chain. Distance to equilibrium is given by the following quantity

$$d^{(n)}(t) := \max_{\xi \in \Omega(n, k(n))} \| \mu_t^\xi - \pi \|.$$

where $\| \cdot \|$ denotes the total variation distance: for two measures on $\Omega(n, k(n))$

$$\| \mu - \pi \| := \frac{1}{2} \sum_{\eta \in \Omega(n, k(n))} |\mu(\eta) - \pi(\eta)|.$$

The main result of this paper is a sharp estimate of the time needed to reach equilibrium.

Theorem 1.1. *If* $\lim_{n \to \infty} k(n)/\sqrt{n} = \infty$, *then for every* $\varepsilon > 0$ *there exists* $\beta > 0$ *such that for all* n

$$d^{(n)} \left(\frac{1}{4} n \log n + \beta n \right) \leqslant \varepsilon$$

$$d^{(n)} \left(\frac{1}{4} n \log n - \beta n \right) \geqslant 1 - \varepsilon.$$

If $\lim_{n \to \infty} k(n)/\sqrt{n} = 0$, *and* $\lim_{n \to \infty} k(n) = \infty$, *then for every* $\varepsilon > 0$ *there exists* $\beta > 0$ *such that for all* n

$$d^{(n)} \left(\frac{1}{2} n \log k(n) + \beta n \right) \leqslant \varepsilon$$

$$d^{(n)} \left(\frac{1}{2} n \log k(n) - \beta n \right) \geqslant 1 - \varepsilon.$$

If $k(n)/\sqrt{n} \to l \in (0, \infty)$, *then* (1.1) *and* (1.1) *hold.*

The function $t \mapsto d^{(n)}(t)$ is non-increasing. Thus for any ε one can set

$$T_{\mathrm{mix}}^{(n)}(\varepsilon) := \inf\{t \mid d^{(n)}(t) \leqslant \varepsilon\} = \sup\{t \mid d^{(n)}(t) > \varepsilon\}.$$

From the above theorem one has that for any $\varepsilon > 0$

$$T_{\mathrm{mix}}^{(n)}(\varepsilon) - T_{\mathrm{mix}}^{(n)}(1 - \varepsilon) = O_\varepsilon(n) = o_\varepsilon(T_{\mathrm{mix}}),$$

where O_ε and o_ε underline dependence in ε. In words: the time in which the distance to equilibrium drops from close to one to close to zero is at most of order n and much smaller than T_{mix}. This phenomenon is known as *cutoff*. It was first identified by Diaconis and Shahshahani (1981) for the random walk on the symmetric group generated by transpositions (see also Berestycki et al., 2010 for a recent extension of this result with a probabilistic proof), and was given its name in the celebrated paper of Aldous and Diaconis (1986) where it is shown that cutoff occurs for top-to-random card shuffle. The bound that is obtained for $T_{\mathrm{mix}}^{(n)}(\varepsilon) - T_{\mathrm{mix}}^{(n)}(1 - \varepsilon)$ (in the present case, $O(n)$) is often called the cutoff window . Our result is optimal in the sense that $O(n)$ is the best window one can obtain: it will be shown in the proof that

$$\lim_{\varepsilon \to 0} \liminf_{n \to \infty} \frac{T_{\mathrm{mix}}^{(n)}(\varepsilon) - T_{\mathrm{mix}}^{(n)}(1 - \varepsilon)}{n} = \infty.$$

The simple exclusion process on the complete graph maps to another problem: the Bernouilli Laplace Diffusion Process. In Diaconis and Shahshahani (1987), Diaconis and Shashahani studied this model and, using purely algebraic methods, proved cutoff in the case $k(n) = n/2$. Their method should be extendable to some other values of k (e.g. using the same method Donnelly et al., 1994 extended the result to the case where $k = \alpha n$ for some $\alpha \in (0, 1)$), but it clearly fails to give the right result when $k(n) \ll \sqrt{n}$ (e.g. the upper-bound given in Theorem 2 fails to be sharp in that particular case). We also underline that the methods we present here are purely probabilistic.

The simple exclusion process on the complete graph can be seen as a projection of the random walk on the symmetric group generated by transposition, and therefore the mixing time for simple exclusion is always smaller than the mixing time for random transposition. What our result underlines is that while the spectral gaps (this a general result that holds for every graph, see Caputo et al., 2010) for the two processes are the same, the mixing time differ: the mixing time for the random transposition model is $n/2 \log n(1+o(1))$, whereas the mixing time for the exclusion process is at most $n/4 \log n(1 + o(1))$. This result is specific to the complete graph: for the exclusion process on the segment or on the circle, mixing time for exclusion process with a density of particle and interchange process are expected to coincide Wilson (2004).

Let us also compare the mixing time of the simple exclusion process with the mixing time of the simple exclusion process with k labeled particles: the space of configurations is

$$\Omega'(n, k(n)) := \left\{ \eta \in \{0, 1, \ldots, k\}^{[1,n] \cap \mathbb{N}} \mid \forall i \in \{1, \ldots, k\} \; \exists! x, \eta(x) = i \right\}.$$

The rules for the evolution are the same: at each time step, one chooses two vertices at random and interchanges their contents. As there is no risk of confusion we use for the labeled process the same notation as that for the exclusion process. The equilibrium measure π for this process is the uniform measure over $\Omega'(n, k(n))$.

What we can show is that if $k \ll \sqrt{n}$ then the mixing times of labeled and unlabeled exclusion process coincide but that they differ as soon as $k \gg \sqrt{n}$. We suspect that for every value of k and every ε one has

$$T'_{\mathrm{mix}}(\varepsilon) := \frac{n}{2}\log k + O(n).$$

We prove in the following that

Theorem 1.2. *For the exclusion process on the complete graph with n vertices and k labeled particles, for every ε, there exists $\beta > 0$ such that for every k and n*

$$d^{(n)}\left(\frac{1}{2}n\log k - \beta n\right) \geqslant 1 - \varepsilon.$$

Moreover if $\lim_{n\to\infty} k(n)/\sqrt{n} = 0$ then for every ε, there exist $\beta > 0$ such that for all n

$$d^{(n)}\left(\frac{1}{2}n\log k(n) + \beta n\right) \leqslant \varepsilon.$$

The mixing time of the simple exclusion process has been studied for some other graphs than the complete graph. However, to our knowledge, cutoff has not been proved for any other graph. We refer to Wilson (2004) for a study of the mixing time of the simple exclusion process on the segment $\{0, \ldots, n\}$ (the edges of the graph are the $(k, k+1)$, $k \in [0, n-1]$), Morris (2006) for simple exclusion on the d-dimensional torus, and Oliveira (2011) for a recent general study of the exclusion process mixing time.

The sequel of the paper is organized as follows

- In Section 2, we reduce the study of the the unlabeled exclusion process to the study of a birth of death chain, which is a first step towards the proof of Theorem 1.1.
- In Section 3, we prove Theorem 1.1 in the case of small k.
- In Section 4, we prove Theorem 1.2.
- In Section 5, we prove Theorem 1.1 in the case of large k.

2. Reduction to the study of a birth and death chain

Our Markov chain is a lazy simple random walk on a transitive graph. Therefore, by transitivity, the distance $\|\mu_t^\xi - \pi\|$ does not depend on the initial configuration ξ. We can set η_0 to be

$$(\eta_0)(x) := \begin{cases} 1 & \text{if } x \in [1, k(n)] \\ 0 & \text{if not} \end{cases}$$

and we simply write $\mathbb{P}$ - and $\mathbb{E}$ for the associated expectation - (resp. μ_t) for the law of $(\eta_t)_{t \geqslant 0}$ (resp. η_t) starting from this configuration.

We now claim that for every t, μ_t is invariant under permutations of the coordinates in $\{1, \ldots, k(n)\}$ and in $\{k(n) + 1, \ldots, n\}$. This is obviously true for $t = 0$, and this remains true for $t > 0$ as the dynamic itself is invariant under these permutations. Therefore, if one sets

$$W(\eta) := \sum_{x=1}^{k(n)} \eta(x),$$
$$W_t := W(\eta_t).$$

then for any m and t such that $\mu(W_t = m) > 0$, $\mu_t(\cdot|W = m)$ is the uniform measure over all the configuration η such that $W(\eta) = m$.

Let $\bar{\mu}_t$ and $\bar{\pi}$ be the law of W under μ_t and π respectively. The preceding remarks imply that

$$d(t) = \|\bar{\mu}_t - \bar{\pi}_t\|$$

One can check that the evolution of W_t is Markovian. Our problem is now confined to the study of the mixing time of this new Markov chain, which is what is called a *birth and death chain* on $\{0, \ldots, k(n)\}$.

We write $\bar{\mu}_t^i$ for the law of W_t starting from $W_0 = i$. As $(W_t)_{t \geqslant 0}$ is a projection of the Markov chain $(\eta_t)_{\geqslant 0}$, one has

$$\|\bar{\mu}_t - \pi\| = \max_{\xi \in \Omega(n,k(n))} \|\mu_t^\xi - \pi\| \geqslant \max_{i \in \{1,\ldots,k\}} \|\bar{\mu}_t^i - \bar{\pi}\| \geqslant \|\bar{\mu}_t - \bar{\pi}\|,$$

that is to say that for any value of t, W_t is farther from the equilibrium measure if $W_0 = k$.

Remark 2.1. At this point of our analysis, one can already show that there is cutoff for our process. Indeed, the cutoff phenomenon for general birth and death chain has been studied in Ding et al. (2010), in which the authors prove that $T_{\mathrm{rel}} = o(T_{\mathrm{mix}}(1/4))$ is a necessary and sufficient for having cutoff (where T_{rel}, the relaxation time is by definition the inverse of the spectral gap). This condition can be checked rather easily in our case. However, one cannot get the location of the cutoff, nor the correct order for the size of the window by using only this general result.

We use the notation $\bar{P}$ to denote the transition probability of $(W_t)_{t \geqslant 0}$. For the sake of clarity, we often omit the dependence in n in the notation. We have

$$\bar{P}(i, i+1) = \frac{2(k-i)^2}{n^2},$$
$$\bar{P}(i, i-1) = \frac{2i(n-2k+i)}{n^2},$$
$$\bar{P}(i, i) = \frac{n^2 - 2[(k-i)^2 + i(n-2k+i)]}{n^2}.$$

We end this section with a first simple Lemma giving the expectation of W_t. It will constantly be used in the sequel.

Lemma 2.2. *One has, for any value of k, for any t*

$$\mathbb{E}^\xi[W_t] = \left(W(\xi) - \frac{k^2}{n}\right)\left(1 - \frac{2}{n}\right)^t + \frac{k^2}{n}$$

Proof: Using the jump rates we compute the expected value of W_{t+1} given W_t. One has

$$\mathbb{E}^\xi[W_{t+1}|W_t] = W_t + \bar{P}(W_t, W_t+1) - \bar{P}(W_t, W_t-1) = \frac{k^2}{n} + \left(W_t - \frac{k^2}{n}\right)\left(1 - \frac{2}{n}\right).$$

Taking the expectation on both sides, and making a trivial induction, one gets the desired result. $\qquad\square$

3. The case $k(n) \ll n^{1/2}$

In this section we prove the main theorem with the assumptions that

$$\lim_{n\to\infty} \frac{k(n)}{\sqrt{n}} = 0 \text{ and } \lim_{n\to\infty} k(n) = \infty.$$

3.1. *Upper bound on T_{mix}.* Under the assumption that $\lim_{n\to\infty} \frac{k(n)}{\sqrt{n}} = 0$, $\bar\pi(W = 0) = 1 - o(1)$. Indeed, with this condition, the expectation of W at equilibrium is

$$\bar\pi(W) = \frac{k^2(n)}{n} = o(1).$$

Let us choose $\gamma > 0$ and set $t_\gamma = \frac{n}{2} \log k(n) + \gamma n$ (is has to be thought as the integer part, but we omit this in the notation to keep things simpler; at any rate it would not change the proof). One has

$$\bar\mu_{t_\gamma}(W) = k \left(1 - \frac{2}{n}\right)^{\frac{n}{2}\log k + \gamma n} + o(1)$$

Altogether we get that

$$\|\bar\mu_{t_\gamma} - \bar\pi\| \leqslant e^{-2\gamma} + o(1).$$

And therefore

$$T_{\mathrm{mix}}(\varepsilon) \leqslant \frac{n}{2} \left(\log k(n) - \log \varepsilon + o(1)\right).$$

$\square$

3.2. *Lower bound on T_{mix}.* To get the other bound, we make the following consideration: the equilibrium measure $\bar\pi$ is concentrated on the event $\{W = 0\}$. Therefore, on the original exclusion process, every particle has to be moved at least once in order to be significantly close to equilibrium. To formalize this properly, we present an alternative construction of the simple exclusion process.

Let $(X_t, Y_t)_{t \geqslant 1}$ be a sequence of i.i.d. random variables distributed uniformly on $\{1, \ldots, n\}^2$ (we include this process in the probability law $\mathbb{P}$). Under $\mathbb{E}^\xi$, we start from $\eta_0 = \xi$ and we build η_t from η_{t-1} by interchanging the content of sites X_t and Y_t if $X_t \neq Y_t$:

$$\eta_t(x) := \begin{cases} \eta_{t-1}(x) & \text{if } x \notin \{X_t, Y_t\}, \\ \eta_t(Y_t) & \text{if } x = X_t, \\ \eta_t(X_t) & \text{if } x = Y_t. \end{cases}$$

We define τ as the time were all the sites in $\{0, \ldots, k(n)\}$ have been selected at least once by the process (X, Y)

$$\tau := \inf \left\{ t \geqslant 0 \mid \bigcup_{s=1}^{t} \{X_s, Y_s\} \supset \{1, \ldots, k(n)\} \right\}.$$

Notice that if $t < \tau$, then $W_t \neq 0$, so that

$$\|\bar\mu_t - \delta_{W=0}\| \geqslant \mathbb{P}\left[\tau > t\right]$$

Therefore one has

$$\|\bar\mu_t - \bar\pi\| \geqslant \mathbb{P}\left[\tau > t\right] - o(1)$$

Estimating the time τ boils down to the so-called coupon collector problem (see Levin et al., 2009, Section 2.2 in particular). Set $X'_{2s-1} := X_s$ and $X'_{2s} = Y_s$. Then X'_s is an i.i.d. sequence One has the following equality in law

$$\tau = \lceil \tau'/2 \rceil = \left\lceil \frac{1}{2} \sum_{i=1}^{k} \mathcal{E}_i \right\rceil .$$

where

$$\tau' := \inf \left\{ t \geqslant 0 \,|\{X'_s, s = 1, \ldots, t\} \supset \{1, \ldots, k(n)\} \right\}.$$

and the $\mathcal{E}_i$ are defined by

$$\sum_{j=1}^{i} \mathcal{E}_j := \inf \left\{ t \geqslant 0 \,\big|\#(\{X'_s | s = 1, \ldots, t\} \cap [1, k(n)]) = i \right\}.$$

It is not difficult to check that $(\mathcal{E}_i)_{i \in [1,k]}$ are independent geometric variables of mean $(\frac{n}{k-i+1})_{i \in [1,k]}$. From this, one gets the following moment estimates: for some constant C

$$\mathbb{E}\left[\sum_{i=1}^{k} \mathcal{E}_i \right] \geqslant n \log k - Cn,$$

$$\mathrm{Var}_{\mathbb{P}}\left[\sum_{i=1}^{k} \mathcal{E}_i \right] = \sum_{i=1}^{k} \frac{1 - (i/n)}{(i/n)^2} \leqslant Cn^2.$$

Therefore, if one chooses $t_\gamma := \frac{1}{2} n \log k - \gamma n$ (suppose that this in an integer), one has by classical second moment inequality:

$$\mathbb{P}\left[\sum_{i=1}^{k} \mathcal{E}_i \leqslant 2t_\gamma \right] \leqslant \frac{C}{(2\gamma - C)^2}$$

And therefore from (3.2) and (3.2)

$$d(t_\gamma) \geqslant 1 - \frac{C}{(2\gamma - C)^2} - o(1),$$

and hence, for any $\varepsilon > 0$

$$T_{\mathrm{mix}}(1 - \varepsilon) \geqslant \frac{n}{2} \left(\log k(n) - \sqrt{C/\varepsilon} - C + o(1) \right).$$

(One could get a tighter bound with $\log \varepsilon$ instead of $-\varepsilon^{-1/2}$ by using exponential moments instead of second moment). $\square$

4. Bounds for the labeled process

The methods of the previous section can be applied for the proof of Theorem 1.2. For the lower-bound we remark that at equilibrium, for any $i \in \{1, \ldots, k\}$, $\pi(\eta(i) = i) = \frac{1}{n}$. Thus in every case the expected number of fixed points is less than one: $\pi(\#\{i | \eta(i) = i\}) \leqslant 1$, and for any integer K

$$\pi(\#\{i | \eta(i) = i\} \geqslant K) \leqslant 1/K.$$

One constructs the exclusion process from $(X_t, Y_t)_{t \geqslant 0}$ as in the previous section. We define τ as the first time at which all of the k first sites have been selected.

$$\tau := \inf \left\{ t \geqslant 0 \mid \#((\bigcup_{s=1}^{t} \{X_s, Y_s\}) \cap \{1, \ldots, k(n)\}) \geqslant k - K \right\}.$$

If one starts the process from

$$(\eta_0)(x) := \left\{ \begin{array}{ll} x & \text{if } x \in [1, k(n)] \\ 0 & \text{if not} \end{array} \right.$$

One has

$$\mu_t^{\eta_0}(\#\{i | \eta(i) = i\} \geqslant K) \geqslant \mathbf{P}(t > \tau)$$

and therefore

$$\|\mu_t^{\eta_0} - \pi\| \geqslant \mathbf{P}(t > \tau) - \frac{1}{K}.$$

Taking the same definition for $\mathcal{E}_i$ as in the previous section one has

$$\tau = \lceil \sum_{i=1}^{k-K} \mathcal{E}_i \rceil,$$

and $\mathcal{E}_i$ are independent geometric variables of respective mean $\left(\frac{n}{k-i-n} \right)$. One has:

$$\mathbb{E} \left[\sum_{i=1}^{k-K} \mathcal{E}_i \right] \geqslant n \log k - n \log K - Cn,$$

$$\text{Var}_{\mathbb{P}} \left[\sum_{i=1}^{k} \mathcal{E}_i \right] = \sum_{i=K}^{k} \frac{1 - (i/n)}{(i/n)^2} \leqslant Cn^2/K,$$

and therefore, using Chebychev inequality one gets:

$$P(\tau \leqslant \frac{n}{2} \log k - n \log K) \leqslant \frac{C}{K(\log K - C)^2}.$$

Overall, this gives that for K sufficiently large and for $t = \frac{n}{2} \log k - n \log K$:

$$\|\mu_t^{\eta_0} - \pi\| \geqslant 1 - \frac{2}{K}.$$

which gives

$$T_{\text{mix}}(1 - \varepsilon) \geqslant \frac{n}{2} \log k + n \log \varepsilon / 2.$$

For the upper bound, we assume that $\lim_{n \to \infty} k(n)/\sqrt{n} = 0$. We notice that as for the unlabeled process, the distance to equilibrium is the same for every starting position by symmetry.

$$d(t) = \|\mu_t^{\eta_0} - \pi\|.$$

Let W be the number of particles lying on the vertices $\{1, \ldots, k\}$. $(W(\eta) := \sum_{x=1}^{k} \mathbf{1}_{\eta(x) \neq 0})$. Now notice that for every $t \geqslant k$ one has

$$\mu_t^{\eta_0}(\cdot | W = 0) = \pi(\cdot | W = 0),$$

as the initial condition and the dynamics are invariant under permutation of $\{k(n) + 1, \ldots, n\}$. Therefore the same analysis as in section 3.1 gives

$$T_{\text{mix}}(\varepsilon) \leqslant \frac{n}{2} \left(\log k(n) - \log \varepsilon + o(1) \right).$$

5. The cases $k(n) \gg \sqrt{n}$ and $k(n) \approx \sqrt{n}$

In this section we prove the main theorem with the assumption that either $\lim_{n\to\infty} k(n)/n = \infty$ or $\lim_{n\to\infty} k(n)/n = l > 0$. The latter case is a bit more complicated than the other as the distribution of W at equilibrium is asymptotically non-degenerate. One makes use of second moment arguments for the lower bound of the mixing time, and a diffusion argument to get an upper bound.

5.1. *Lower bound on the mixing time.* In this section we work with the weaker assumption $\liminf_{n\to\infty} k(n)/\sqrt{n} > 0$. To get the right bound on the mixing time one uses a second moment method. The first and essential step is to compute a tight estimate of $\bar{\mu}_t [W^2]$. We start by writing an explicit formula for the second moment of W_t.

Lemma 5.1.

$$\mathbb{E}[W_t^2] = \left(1 - \frac{4}{n} + \frac{4}{n^2}\right)^t k^2$$

$$+ \sum_{s=0}^{t-1} \left[\left(\frac{4k^2}{n^2} - \frac{8k}{n^2} + \frac{2i}{n}\right)\mathbb{E}[W_s] + \frac{2k^2}{n}\right]\left(1 - \frac{4}{n} + \frac{4}{n^2}\right)^{t-1-s}.$$

Proof: One simply uses the transition of the Markov chain to get a recurrence relation

$$\mathbb{E}[W_{t+1}^2|W_t] = W_t^2 + 2W_t(\bar{P}(W_t, W_t + 1) - \bar{P}(W_t, W_t - 1))$$
$$+ \bar{P}(W_t, W_t + 1) + \bar{P}(W_t, W_t - 1).$$

By taking the expectation on both sides, one gets the following recursive relation

$$\mathbb{E}(W_{t+1}^2) = \left(1 - \frac{4}{n} + \frac{4}{n^2}\right)\mathbb{E}[W_t^2] + \left(\frac{4k^2}{n^2} - \frac{8k}{n^2} + \frac{2i}{n}\right)\mathbb{E}[W_t] + \frac{k^2}{n}.$$

which after induction gives the expected result. $\qquad\square$

Then, using the formula above, we get a clean bound on the variance of W_t.

Lemma 5.2. *There exists a constant C such that for any K, n large enough (depending on K), and $t = \frac{n}{4}\log n - \frac{\gamma n}{2}$, $\gamma \in [-K, K]$*

$$\mathrm{Var}_{\mathbb{P}}(W_t^2) \leqslant \frac{Ck^2}{n} + \frac{e^\gamma k}{\sqrt{n}}.$$

Proof: Suppose that $t = \frac{n}{4}(\log n - 2\gamma)$ for some $\gamma \in [-K, K]$. All the $O(\cdot)$ are uniform in $\gamma \in [-K, K]$ when n is large enough. We start by giving an estimate of the expectation squared

$$[\mathbb{E}(W_t)]^2 = \left[\frac{k^2}{n} + \left(k - \frac{k^2}{n}\right)\left(1 - \frac{1}{2n}\right)^t\right]^2$$

$$= \left[\frac{k^2}{n} + \left(k - \frac{k^2}{n}\right)\frac{e^\gamma}{\sqrt{n}}\left(1 + O\left(\frac{\log n}{n}\right)\right)\right]^2$$

$$= \frac{k^4}{n^2} + \frac{2k^2}{n^{3/2}}\left(k - \frac{k^2}{n}\right)e^\gamma + n^{-1}\left(k - \frac{k^2}{n}\right)^2 e^{2\gamma} + O(k^2/n).$$

Estimating $\mathbb{E}(W_t^2)$ is a bit more tricky. For practical reasons we divide it in three terms

$$\mathbb{E}(W_t^2) = \left(1 - \frac{4}{n} + \frac{4}{n^2}\right)^t k^2 + \frac{2k^2}{n^2} \sum_{s=0}^{t-1} \left(1 - \frac{4}{n} + \frac{4}{n^2}\right)^{t-1-s}$$

$$+ \sum_{s=0}^{t-1} \left(\frac{4k^2}{n^2} - \frac{8k}{n^2} + \frac{2}{n}\right) \mathbb{E}(W_s) \left(1 - \frac{4}{n} + \frac{4}{n^2}\right)^{t-1-s}$$

The first one gives

$$\left(1 - \frac{4}{n} + \frac{4}{n^2}\right)^t k^2 = \frac{k^2 e^{2\gamma}}{n} + O(k^2/n).$$

The second:

$$\frac{2k^2}{n^2} \sum_{s=0}^{t-1} \left(1 - \frac{4}{n} + \frac{4}{n^2}\right)^{t-1-s} = O(k^2/n).$$

We divide the third term it into two contributions

$$\sum_{s=0}^{t-1} \left(\frac{4k^2}{n^2} - \frac{8k}{n^2} + \frac{2}{n}\right) \mathbb{E}(W_s) \left(1 - \frac{4}{n} + \frac{4}{n^2}\right)^{t-1-s}$$

$$= \left(\frac{4k^2}{n^2} - \frac{8k}{n^2} + \frac{2}{n}\right) \left(k - \frac{k^2}{n}\right) \left(1 - \frac{4}{n} + \frac{4}{n^2}\right)^{t-1} \sum_{s=0}^{t-1} \left(\frac{1 - \frac{2}{n}}{1 - \frac{4}{n} + \frac{4}{n^2}}\right)^s$$

$$+ \left(\frac{4k^2}{n^2} - \frac{8k}{n^2} + \frac{2}{n}\right) \frac{k^2}{n} \sum_{s=0}^{t-1} \left(1 - \frac{4}{n} + \frac{4}{n^2}\right)^s.$$

The first contribution can be estimated as follows

$$\left(\frac{4k^2}{n^2} + \frac{2}{n} + O(k/n^2)\right) \left(k - \frac{k^2}{n}\right) \frac{e^{2\gamma}}{n}.$$

$$\cdot \left(1 + O\left(\frac{\log n}{n}\right)\right) \frac{\sqrt{n} e^{-\gamma} - 1 + O(n^{-1/2} \log n)}{\frac{2}{n} + O(n^{-2})}$$

$$= \frac{2k^2}{n^{3/2}} \left(k - \frac{k^2}{n}\right) e^{\gamma} - \frac{2k^2}{n^2} \left(k - \frac{k^2}{n}\right) e^{2\gamma} + \frac{e^{\gamma} k}{\sqrt{n}} + O(k^2/n).$$

The second is equal to

$$\left(\frac{4k^2}{n^2} + O(1/n)\right) \frac{k^2}{4} \left(1 - \frac{e^{2\gamma}}{n} + O\left(\frac{1}{n}\right)\right) = \frac{k^4}{n^2} - e^{2\gamma} \frac{k^4}{n^3} + O(k^2/n).$$

Summing everything up gives the expected result.   □

Now we use the bounds that we have on the second and first moment to bound the mixing time. Set $t = \frac{n}{4}(\log n - 2\gamma)$. Let (W_1, W_2) be a maximal coupling between $\bar{\mu}_t$ and $\bar{\pi}$, and let $\bar{\nu}_t$ be its law (such that $\|\bar{\mu}_t - \bar{\pi}\| = \bar{\nu}_t(W_1 \neq W_2)$).

One has

$$\bar{\nu}_t(W_1 - W_2)^2 = \bar{\nu}_t\{W_1 \neq W_2\} \bar{\nu}_t\left((W_1 - W_2)^2 \mid W_1 \neq W_2\right)$$

$$\geqslant \bar{\nu}_t\{W_1 \neq W_2\} \left[\bar{\nu}_t\left((W_1 - W_2) \mid W_1 \neq W_2\right)\right]^2 = \frac{(\nu_t(W_1 - W_2))^2}{\|\bar{\mu}_t - \bar{\pi}\|}.$$

And hence
$$\|\bar{\mu}_t - \bar{\pi}\| \geqslant \frac{[\nu_t(W_1 - W_2)]^2}{\bar{\nu}_t(W_1 - W_2)^2} = \frac{1}{1 + \frac{\text{Var}_{\bar{\nu}_t}(W_1 - W_2)}{[\nu_t(W_1 - W_2)]^2}}.$$

It is then easy to compute the moments. The expectation is given by
$$\bar{\nu}_t(W_1 - W_2) = \bar{\mu}_t(W) - \bar{\pi}(W) = e^{\gamma} \frac{1}{\sqrt{n}} \left(k - \frac{k^2}{n}\right)(1 + o(1)).$$

For the variance, first notice that:
$$\text{Var}_{\bar{\pi}}(W) = \sum_{i,j=1}^{k} \bar{\pi}(\eta(i)\eta(j)) - \frac{k^4}{n^2} = k\bar{\pi}(\eta(1)) + k(k-1)\bar{\pi}(\eta(1)\eta(2)) - \frac{k^4}{n^2} \leqslant \frac{k^2}{n}.$$

And therefore
$$\text{Var}\bar{\nu}_t(W_1 - W_2)^2 \leqslant 2[\text{Var}_{\bar{\mu}_t}(W) + \text{Var}_{\bar{\pi}}(W)] \leqslant C' \frac{k^2}{n} + \frac{e^{\gamma}k}{\sqrt{n}}.$$

Hence one gets
$$\|\bar{\mu}_t - \bar{\pi}\| \geqslant \left(1 + \frac{C'\frac{k^2}{n} + \frac{e^{\gamma}k}{\sqrt{n}}}{\frac{e^{2\gamma}}{n}\left(k - \frac{k^2}{n}\right)^2(1 + o(1))}\right)^{-1}$$
$$\geqslant \left(1 + 4C'e^{-2\gamma} + \frac{4\sqrt{n}}{k}e^{-\gamma} + o(1)\right)^{-1}.$$

where in the last line one used $k \leqslant \frac{n}{2}$. Using the assumption that $\liminf k(n)/\sqrt{n} > 0$, one obtains that for n large enough
$$\|\bar{\mu}_t - \bar{\pi}\| \geqslant \left(1 + C'' \max(e^{-2\gamma}, e^{-\gamma}) + o(1)\right)^{-1}$$

Therefore
$$T_{\text{mix}}(\varepsilon) \geqslant \begin{cases} \frac{n}{4}\log n - \frac{n}{4}\log\left(\frac{C''}{\varepsilon^{-1}-1}\right) & \text{if } \varepsilon \geqslant (1 + C'')^{-1} \\ \frac{n}{4}\log n + \frac{n}{2}\log\left(\frac{\varepsilon^{-1}-1}{C''}\right) & \text{if } \varepsilon \leqslant (1 + C'')^{-1} \end{cases}$$

5.2. *Upper bound on* T_{mix}. To give an upper bound on the mixing time, we bound the following quantity
$$\bar{d}(t) = \max_{x,y \in [0,k]^2} \|\bar{\mu}_t^x - \bar{\mu}_t^y\| \geqslant d(t).$$

We define a coupling of two replicas of the Markov chain W starting from different states as follows: let $(W_t^{(1)}, W_t^{(2)})$ be the Markov chain on $\{1, \ldots, k\}^2$ given by the following transition
$$\mathbf{P}((i,j), (i \pm 1, j)) := \bar{P}(i, i \pm 1), \text{ if } i \neq j,$$
$$\mathbf{P}((i,j), (i, j \pm 1)) := \bar{P}(i, j \pm 1), \text{ if } i \neq j,$$
$$\mathbf{P}((i,i), (i \pm 1, i \pm 1)) := \bar{P}((i, i \pm 1),$$
$$\mathbf{P}((i,j), (i, j)) := \bar{P}(i, i) + \bar{P}(j, j) - 1,$$
$$\mathbf{P}((i,i), (i, i)) := \bar{P}(i, i),$$

where all the other transitions have zero probability. One can check that the coefficients are positive and that this indeed defines a stochastic matrix. This coupling has the property that once $W^{(1)}$ and $W^{(2)}$ merge, they stay together. Moreover,

before the merging, at most one of the two coordinates changes at each time step, and therefore the sign of $W^{(1)} - W^{(2)}$ is constant in time (*i.e.* $W^{(1)}$ and $W^{(2)}$ cannot cross without merging). With our choice for initial condition, in the sequel it is always non-negative.

If one denotes by $\mathbf{P}^{x,y}$ the law of $(W_t^{(1)}, W_t^{(2)})$ with initial condition (x, y), $x \geqslant y$ and defines

$$\tau := \inf\{t > 0 \mid W_t^{(1)} = W_t^{(2)}\},$$

then

$$\|\bar{\mu}_t^x - \bar{\mu}_t^y\| \leqslant \mathbf{P}^{x,y}[\tau > t].$$

Set $D_t := W_t^{(1)} - W_t^{(2)} \geqslant 0$. According to Lemma 2.2

$$\mathbf{E}^{x,y}[D_t] = (x - y)\left(1 - \frac{2}{n}\right)^t.$$

A first moment analysis is enough to treat the case $\lim_{n\to\infty} k(n)/\sqrt{n} = l$. Indeed for any x and y

$$\mathbf{P}^{x,y}[\tau > t] = \mathbf{P}^{x,y}[D_t \geqslant 1] \leqslant \mathbf{E}^{x,y}[D_t] \leqslant l\sqrt{n}(1 + o(1))\left(1 - \frac{2}{n}\right)^t.$$

Setting $t_\beta = \frac{1}{4}n\log n + \beta n$ one gets

$$\mathbf{P}^{x,y}[\tau > t_\beta] = le^{-2\beta}(1 + o(1)).$$

and therefore

$$T_{\mathrm{mix}}(\varepsilon) \leqslant \frac{1}{4}n\log n - \frac{n}{2}\log(\varepsilon/l).$$

The rest of the section is therefore devoted to the case $\lim_{n\to\infty} k(n)/\sqrt{n} = \infty$. Given that $W_t^{(1)} > W_t^{(2)}$, D_t has the following transition probabilities

$$\mathbf{P}(D_{t+1} = D_t - 1 \mid W_t^{(1)}, W_t^{(2)}) = \mathbf{P}(W_t^{(1)}, W_t^{(1)} - 1) + \mathbf{P}(W_t^{(2)}, W_t^{(2)} + 1)$$

$$= 2\frac{W_t^{(1)}(n - 2k + W_t^{(1)}) + (k - W_t^{(2)})^2}{n^2},$$

$$\mathbf{P}(D_{t+1} = D_t + 1 \mid W_t^{(1)}, W_t^{(2)}) = \mathbf{P}(W_t^{(1)}, W_t^{(1)} + 1) + \mathbf{P}(W_t^{(2)}, W_t^{(2)} - 1)$$

$$= 2\frac{(k - W_t^{(1)})^2 + W_t^{(2)}(n - 2k + W_t^{(2)})}{n^2}.$$

One can check that the evolution of D is not Markovian (it depends on the values of $W_t^{(1)}$ and $W_t^{(2)}$ and not only on D_t). This makes the analysis of τ difficult.

We now sketch the method we use to tackle this problem:

- First, we use a first moment method to show that after a time $t_0 := \frac{n}{4}\log n$, with probability close to one $W_{t_0}^{(1)} - W_{t_0}^{(2)}$ is of order $k/\sqrt{n}$.
- Then, we do a sequence of stochastic comparisons to state that starting from $W_{t_0}^{(1)} - W_{t_0}^{(2)} \leqslant k/\sqrt{n}$, τ is stochastically dominated by (t_0 plus) the hitting time of zero for a simple symmetric random walk on $\mathbb{Z}$ with jump rate k^2/n^2 .
- Finally, we use a reflection argument to show that the typical time for hitting zero when starting from $k/\sqrt{n}$ for such a walk is of order n. Altogether this gives that τ is smaller than $t_0 + Kn$ with probability close to one if K is sufficiently large.

The idea of combining a first moment method and diffusion in order to evaluate mixing times has already been used to compute the mixing time of the mean field Ising model in Levin et al. (2010), the method was then refined in Ding et al. (2009). The interest of the method here lies in the particular manner the coupling has to be constructed.

Set $t_\alpha = \frac{n}{4} \log n + \alpha n$. One uses Markov property to get the following bound on τ.

$$
\mathbf{P}^{x,y}(\tau > t_{\alpha+1}) \leqslant \mathbf{E}^{x,y}\left[\mathbf{P}^{W_{t_\alpha}^{(1)}, W_{t_\alpha}^{(2)}}(\tau > n)\right]
$$

$$
\leqslant \mathbf{P}^{x,y}[W_{t_\alpha}^{(1)} - W_{t_\alpha}^{(2)} \leqslant M] + \max_{x' \geqslant y', x'-y' \leqslant M} \mathbf{P}^{x',y'}[\tau > n]
$$

$$
\leqslant \frac{\mathbf{E}^{x,y}[W_{t_\alpha}^{(1)} - W_{t_\alpha}^{(2)}]}{M} + \max_{x' \geqslant y', x'-y' \leqslant M} \mathbf{P}^{x',y'}[\tau > n].
$$

We apply it for $M = \frac{ke^{-\alpha}}{\sqrt{n}}$. As we have

$$
\mathbf{E}^{x,y}[W_{t_\alpha}^{(1)} - W_{t_\alpha}^{(2)}] = (x-y)\left(1 - \frac{2}{n}\right)^{t_\alpha} \leqslant \frac{k}{\sqrt{n}} e^{-2\alpha},
$$

this gives

$$
\bar{d}(t_{\alpha+1}) \leqslant e^{-\alpha}(1 + o(1)) + \max_{x' \geqslant y', x'-y' \leqslant \frac{ke^{-\alpha}}{\sqrt{n}}} \mathbf{P}^{x',y'}[\tau > n],
$$

and all that remains to do is estimating the second term. The next step is to show that

$$
\mathbf{P}^{x',y'}[\tau > n] \leqslant Q^{x'-y'}[\tau' > n]
$$

where τ' is the first hitting time of zero for a nearest neighbor symmetric random walk on $\mathbb{Z}$ with "jump rate" n/k^2, starting from $x' - y'$ (law $Q^{x'-y'}$). The result is rather intuitive, as $D_t = W_t^{(1)} - W_t^{(2)}$ has a drift towards zero and the probability of jumping is bounded from below by k^2/n^2. However, stochastic comparisons have to be made with some care in order to prove the result rigorously. We construct a coupling explicitly.

We define J_i the sequence of moving time for $(W^{(1)}, W^{(2)})$ i.e. $J_0 = 0$ and for $i \geqslant 0$

$$
J_{i+1} := \inf\{t \geqslant J_i \mid (W_t^{(1)}, W_t^{(2)}) \neq (W_{J_i}^{(1)}, W_{J_i}^{(2)})\}.
$$

We remark that $(W_{J_i}^{(1)}, W_{J_i}^{(2)})_{i \geqslant 0} = (\bar{W}_i^{(1)}, \bar{W}_i^{(2)})_{i \geqslant 0}$ is itself a Markov chain (something similar to the skeleton of a continuous time/discrete space Markov chain), and that conditionally to $(W_{J_i}^{(1)}, W_{J_i}^{(2)})_{i \geqslant 0}$, $(J_{i+1} - J_i)_{i \geqslant 0}$, it is a sequence of geometric variables of mean $(2 - \bar{P}(W_{J_i}^{(1)}, W_{J_i}^{(1)}) - \bar{P}(W_{J_i}^{(2)}, W_{J_i}^{(2)}))^{-1}$ if $W_{J_i}^{(1)} \neq W_{J_i}^{(2)}$, and of mean $(1 - \bar{P}(W_{J_i}^{(1)}, W_{J_i}^{(1)}))^{-1}$ if the processes have merged.

Let $(U_i)_{i \geqslant 0}$ and $(U_i')_{i \geqslant 0}$ be two independent sequences of i.i.d. random variables. One constructs the process $(W_t^{(1)}, W_t^{(2)})_{t \geqslant 0}$ starting from (x,y) deterministically from the sequences $(U_i)_{i \geqslant 0}$ and $(U_i')_{i \geqslant 0}$ as follows (for the sake of simplicity, we do not give details of how the construction is done after the merging as we do not use it):

- First, one constructs recursively $(\bar{W}_i^{(1)}, \bar{W}_i^{(2)})_{i \geqslant 0}$, $(\bar{W}_0^{(1)}, \bar{W}_0^{(2)}) = (x, y)$ and

$$(\bar{W}_{i+1}^{(1)}, \bar{W}_{i+1}^{(2)}) := \begin{cases} (\bar{W}_i^{(1)} + 1, \bar{W}_i^{(2)}) & \text{if } U'_{i+1} \in [0, a(\bar{W}_i^{(1)}, W_i^{(2)})] \\ (\bar{W}_i^{(1)}, \bar{W}_i^{(2)} - 1) & \text{if } U'_{i+1} \in (a(\bar{W}_i^{(1)}, W_i^{(2)}), b(\bar{W}_i^{(1)}, W_i^{(2)}))] \\ (\bar{W}_i^{(1)} - 1, \bar{W}_i^{(2)}) & \text{if } U'_{i+1} \in (b(\bar{W}_i^{(1)}, W_i^{(2)}), c(\bar{W}_i^{(1)}, W_i^{(2)}))] \\ (\bar{W}_i^{(1)}, \bar{W}_i^{(2)} + 1) & \text{if } U'_{i+1} \in (c(\bar{W}_i^{(1)}, W_i^{(2)}), 1] \end{cases}$$

  where $a(i, j)$, $b(i, j)$, $c(i, j)$ are chosen such that the chain has the right transition probabilities. It is important to notice that $b(i, j) \leqslant 1/2$ for all i, j.
- Then, given $(\bar{W}_i^{(1)}, \bar{W}_i^{(2)})_{i \geqslant 0}$ we construct the sequence of moving times: $J_0 = 0$ and

$$J_{i+1} - J_i = m \quad \text{if} \quad U_{i+1} \in \left[1 - (1 - q(\bar{W}_i^{(1)}, \bar{W}_i^{(2)}))^{m-1}, 1 - (1 - q(\bar{W}_{i-1}^{(1)}, \bar{W}_{i-1}^{(2)})^m \right),$$

  where $q(i, j) := (2 - \bar{P}(i, i) - \bar{P}(j, j)) \geqslant k^2/n^2$ is the inverse of the mean jumping time from (i, j).

Using the same variables $(U_i)_{i \geqslant 0}$ and $(U'_i)_{i \geqslant 0}$, one constructs a simple random walk on $\mathbb{Z}$.

- First, one defines $\bar{X}_i$ as

$$\bar{X}_i = (x - y) + \sum_{j=1}^{i} \mathbf{1}_{\{U'_i \leqslant 1/2\}} - \mathbf{1}_{\{U'_i > 1/2\}}.$$

- Then, one defines H_i by $H_0 = 0$ and

$$H_{i+1} - H_i = m \text{ if } U_{i+1} \in \left[1 - (1 - k^2/n^2,)^{m-1}, 1 - (1 - k^2/n^2)^m \right).$$

- Finally, one sets

$$X_t = \bar{X}_i \text{ if } t \in [H_i, H_{i+1}).$$

From this construction one has that $H_i \geqslant J_i$ and $\bar{X}_i \geqslant \bar{W}_i^{(1)} - \bar{W}_i^{(2)}$ for all i. Therefore

$$\tau' := \inf\{t \mid X_t = 0\} = H_{\inf\{i \mid \bar{X}_i = 0\}} \geqslant J_{\inf\{i \mid \bar{W}_i^{(1)} = \bar{W}_i^{(2)}\}} = \tau.$$

By construction $(X_t)_{t \geqslant 0}$ is a random walk with transition probability $p(x, x \pm 1) = k^2/(2n^2)$ and $p(x, x) = 1 - k^2/n^2$, and therefore we proved (5.2). We now finish the proof of the main theorem. From (5.2) and (5.2)

$$\bar{d}(t_{\alpha+1}) \leqslant e^{-\alpha}(1 + o(1)) + Q^{\lceil \frac{ke^{-\alpha}}{\sqrt{n}} \rceil}[\tau' > n],$$

where Q^m is the law of $(X_t)_{t \geqslant 0}$ starting from m. Then from Proposition 5.3

$$Q^{\lceil \frac{ke^{-\alpha}}{\sqrt{n}} \rceil}[\tau' > n] \leqslant e^{-\alpha}(1 + o(1))$$

and therefore

$$T_{\mathrm{mix}}(\varepsilon) \leqslant \frac{n}{4} \log n + n \log(\varepsilon/2) + o(n).$$

Using the same technique one can get that there exists a function $c(\varepsilon)$ that goes to infinity when ε goes to 0 such that

$$T_{\mathrm{mix}}(1 - \varepsilon) \leqslant \frac{n}{4} \log n - c(\varepsilon)n + o(n).$$

(this, together with the lower-bound part shows that the window of size $O(n)$ is an optimal result (1)).

5.3. *Diffusion bounds.* One is left with proving the approximation we used for the law of τ'. Let n be a fixed integer. Let $(X_t)_{t \geqslant 0}$ be a nearest-neighbor random walk on $\mathbb{Z}$ with transitions $p(x, x \pm 1) = q(n)/2$, $p(x, x) = 1 - q(n)$. We start our random walk from $\alpha s(n)$ where $s(n)$ and $q(n)$ satisfy:

$$s(n)^2 = q(n)n.$$

We denote Q^x the probability associated to this random walk starting from x and $Q = Q^0$. We want to estimate τ', the hitting time of zero for this random walk.

Proposition 5.3. *One has for any given positive α and β*

$$\lim_{n \to \infty} Q^{\alpha s(n)} [\tau' > \beta n] = \frac{1}{\sqrt{2\pi}} \int_{\left[-\frac{\alpha}{\sqrt{\beta}} , \frac{\alpha}{\sqrt{\beta}}\right]} e^{-\frac{s^2}{2}} \, ds \leqslant \frac{\alpha}{\sqrt{\beta}}.$$

In fact the proof can almost be reduced to proving the following lemma.

Lemma 5.4. *For any positive integer m and n,*

$$Q^m [\tau' > n] = Q [X_n \in [-m+1, m]]$$

Proof: We have

$$Q^m [\tau > n] = \sum_{j > 0} Q^m (\forall i \in [1, n-1] \ X_i > 0, X_n = j).$$

Moreover,

$$Q^m (\forall i \in [1, n-1] \ X_i > 0, X_n = j)$$
$$= Q^m (X_n = j) - Q^m (\exists i \in [1, n-1] \ X_i = 0, \ X_n = j)$$
$$= Q^m (X_n = j) - Q^m [X_n = -j] = Q [X_n = j - m] - Q [X_n = j + m].$$

The second inequality just comes from the application of the reflection principle. Therefore

$$\sum_{j > 0} Q^m (\forall i \in [1, n-1] X_i > 0, X_n = j) = Q (X_n \in [-m+1, m]).$$

$\square$

With the previous lemma, all one has to do is prove the convergence of $X_{\beta n}/\sqrt{\beta n q(n)}$ to a Gaussian variable. We do it by computing the Fourrier transform. For any fixed K we have that uniformly for all $|t| \leqslant K$

$$Q \left[e^{it \frac{X_{\beta n}}{\sqrt{\beta n q(n)}}} \right] = \left[1 - q(n) \left(1 - \cos \frac{t}{\sqrt{\beta n q(n)}} \right) \right]^{\beta n}$$

$$= \left[1 - q(n) \frac{t^2}{2\beta n q(n)} (1 + o(1)) \right]^{\beta n} = e^{-\frac{t^2}{2}} (1 + o(1)).$$

Therefore

$$\lim_{n\to\infty} Q^{\alpha s(n)}\left[\tau' > \beta n\right] = \lim_{n\to\infty} Q\left[\frac{X_{\beta n}}{\sqrt{nqn}} \in \frac{[-\alpha s(n) + 1, \alpha s(n)]}{\sqrt{\beta nq(n)}}\right]$$

$$= P\left[\mathcal{N} \in \left[-\frac{\alpha}{\sqrt{\beta}} \; ; \; \frac{\alpha}{\sqrt{\beta}}\right]\right].$$

where $\mathcal{N}$ (with law denoted by P) is a standard normal variable.

Acknowledgment:

References

D. Aldous and P. Diaconis. Shuffling cards and stopping times. *Amer. Math. Monthly* **93** (5), 333–348 (1986). MR841111.

N. Berestycki, O. Schramm and O. Zeitouni. Mixing times for random k-cycles and coalescence-fragmentation chains. *ArXiv Mathematics e-prints* (2010). http://arxiv.org/abs/1001.1894.

P. Caputo, T. M. Liggett and T. Richthammer. Proof of Aldous' spectral gap conjecture. *J. Amer. Math. Soc.* **23** (3), 831–851 (2010). MR2629990.

P. Diaconis and M. Shahshahani. Generating a random permutation with random transpositions. *Z. Wahrsch. Verw. Gebiete* **57** (2), 159–179 (1981). MR626813.

P. Diaconis and M. Shahshahani. Time to reach stationarity in the Bernoulli-Laplace diffusion model. *SIAM J. Math. Anal.* **18** (1), 208–218 (1987). MR871832.

J. Ding, E. Lubetzky and Y. Peres. The mixing time evolution of Glauber dynamics for the mean-field Ising model. *Comm. Math. Phys.* **289** (2), 725–764 (2009). MR2506768.

J. Ding, E. Lubetzky and Y. Peres. Total variation cutoff in birth-and-death chains. *Probab. Theory Related Fields* **146** (1-2), 61–85 (2010). MR2550359.

P. Donnelly, P. Lloyd and A. Sudbury. Approach to stationarity of the Bernoulli-Laplace diffusion model. *Adv. in Appl. Probab.* **26** (3), 715–727 (1994). MR1285456.

D. A. Levin, M. J. Luczak and Y. Peres. Glauber dynamics for the mean-field Ising model: cut-off, critical power law, and metastability. *Probab. Theory Related Fields* **146** (1-2), 223–265 (2010). MR2550363.

D. A. Levin, Y. Peres and E. L. Wilmer. *Markov chains and mixing times.* American Mathematical Society, Providence, RI (2009). ISBN 978-0-8218-4739-8. With a chapter by James G. Propp and David B. Wilson. MR2466937.

B. Morris. The mixing time for simple exclusion. *Ann. Appl. Probab.* **16** (2), 615–635 (2006). MR2244427.

R. I. Oliveira. Mixing of the symmetric exclusion processes in terms of the corresponding single-particle random walk. *ArXiv Mathematics e-prints* (2011). http://arxiv.org/abs/1007.2669.

D. B. Wilson. Mixing times of Lozenge tiling and card shuffling Markov chains. *Ann. Appl. Probab.* **14** (1), 274–325 (2004). MR2023023.