The cutoff phenomenon for Ehrenfest chains

Guan-Yu Chen*, Yang-Jen Fang, Yuan-Chung Sheu

Department of Applied Mathematics, National Chiao Tung University, Hsinchu 300, Taiwan

Abstract

We consider families of Ehrenfest chains and provide a simple criterion on the L^p-cutoff and the L^p-precutoff with specified initial states for $1 \le p < \infty$. For the family with an L^p-cutoff, a cutoff time is described and a possible window is given. For the family without an L^p-precutoff, the exact order of the L^p-mixing time is determined. The result is consistent with the well-known conjecture on cutoffs of Markov chains proposed by Peres in 2004, which says that a cutoff exists if and only if the multiplication of the spectral gap and the mixing time tends to infinity.

1. Introduction

Consider a time-homogeneous Markov chain on a finite set Ω with one-step transition matrix K. Let $K^t(x, \cdot)$ denote the probability distribution of the chain at time t started at x. It is well-known that if K is ergodic (irreducible and aperiodic), then

$$\lim_{t \to \infty} K^t(x, y) = \pi(y) \quad \forall x, y \in \Omega,$$

* Corresponding author. Tel.: +1 886 3 5712121x56421; fax: +1 886 3 5724679.
E-mail addresses: gychen@math.nctu.edu.tw (G.-Y. Chen), imyjfan@hotmail.com (Y.-J. Fang), sheu@math.nctu.edu.tw (Y.-C. Sheu).

where π is the stationary distribution of K on Ω. Denote by k_x^t the relative density of $K^t(x, \cdot)$ with respect to π, that is, $k_x^t(y) = K^t(x, y)/\pi(y)$. For $1 \le p < \infty$, define the L^p-distance by

$$D_p(x, t) = \|k_x^t - 1\|_{L^p(\pi)} = \left(\sum_{y \in \Omega} |k_x^t(y) - 1|^p \pi(y) \right)^{1/p}.$$

For $p = \infty$, the L^∞-distance is set to be $D_\infty(x, t) = \max_y |k_x^t(y) - 1|$. In the case $p = 1$, this is exactly twice of the total variation distance between $K^t(x, \cdot)$ and π, which is defined by

$$D_{\text{TV}}(x, t) = \|K^t(x, \cdot) - \pi\|_{\text{TV}} = \max_{A \subset \Omega} \{K^t(x, A) - \pi(A)\}. \tag{1.1}$$

In the case $p = 2$, it is the so-called chi-square distance. For any $\epsilon > 0$ and $1 \le p \le \infty$, define the L^p-mixing time by

$$T_p(x, \epsilon) = \min\{t \ge 0 : D_p(x, t) \le \epsilon\}.$$

Consider a family of finite ergodic Markov chains (Ω_n, K_n, π_n) with specified initial states x_n. For $1 \le p \le \infty$, the family is said to present an L^p-cutoff with cutoff time t_n if

$$\lim_{n \to \infty} D_{n,p}(x_n, (1 + a)t_n) = \begin{cases} 0 & \text{if } a > 0 \\ M_p & \text{if } -1 < a < 0 \end{cases},$$

where $D_{n,p}$ denotes the L^p-distance for the nth Markov chain, $M_1 = 2$ and $M_p = \infty$ for $p \in (1, \infty]$. In total variation and separation, the cutoff is defined in the same spirit and has the replacement of M_p with 1. The concept of cutoffs was introduced by Aldous and Diaconis in [1–3] to capture the fact that many ergodic Markov chains converge abruptly to their stationary distributions. We refer the reader to [7,8,13,15,16] for details and further discussions on variant examples.

In this paper, we treat the Ehrenfest chains, a classical example introduced by Paul Ehrenfest to remark the second law of thermodynamics. In detail, let $\Omega_n = \{0, 1, \ldots, n\}$ and K_n be the Markov kernel of the Ehrenfest chain on Ω_n given by

$$K_n(i, i + 1) = 1 - \frac{i}{n}, \qquad K_n(i + 1, i) = \frac{i + 1}{n}, \quad \forall 0 \le i \le n - 1. \tag{1.2}$$

Clearly, the unbiased binomial distribution, $\pi_n(i) = \binom{n}{i} 2^{-n}$, is the stationary distribution of K_n and the pair (K_n, π_n) is reversible, i.e. $\pi_n(i)K_n(i, j) = \pi_n(j)K_n(j, i)$ for all $i, j \in \Omega_n$. By lifting the chain to a random walk on the hypercube, one may use the group representation of $(\mathbb{Z}_2)^n$ to identify the eigenvalues and eigenvectors of K_n. See Lemma B.1.

The aim of this paper is to provide a necessary and sufficient condition on the L^p-cutoff of Ehrenfest chains with $1 \le p < \infty$ and describe the L^p-cutoff time if any. The following is our main result achieved in Theorems 3.1 and 4.1.

Theorem 1.1. *Let K_n be defined in* (1.2) *and set $K_n' = (I + nK_n)/(n + 1)$, $\pi_n(i) = \binom{n}{i} 2^{-n}$. For $p \in [1, \infty)$, the following are equivalent.*

(1) *The family $\{(\Omega_n, K_n', \pi_n) : n = 1, 2, \ldots\}$ with starting states $(x_n)_{n=1}^\infty$ has an L^p-cutoff.*
(2) *$|n - 2x_n|/\sqrt{n} \to \infty$ as $n \to \infty$.*

Moreover, if (3) *holds, then, as $n \to \infty$,*

$$T_{n,p}(x_n, \epsilon) = \frac{n}{2} \log \frac{|n - 2x_n|}{\sqrt{n}} + O(n), \quad \forall \epsilon > 0, \ p \in (1, \infty),$$

where $O(n)$ denotes a function of order less than or equal to n. For $p = 1$, the above identity remains true with $\epsilon \in (0, 2)$.

For the total variation, [8, Theorem 6.5] provides a sufficient condition on cutoffs, while Theorem 3.1 proves that such a condition is necessary. For the L^p-cutoff with $1 < p < \infty$, Theorem [8, Theorem 6.5] gives the L^2 case, while Theorem 4.1 gives the L^p case. It is worthwhile to remark that if there is a cutoff, then the main term of the mixing time is the same for all $1 \le p < \infty$. In fact, the equivalence in Theorem 1.1 is also valid for the precutoff, a concept with a more general sense on the rapid convergence of Markov chains, which will be introduced in the next section. The L^p-precutoff in Theorem 1.1 is determined by Theorems 3.1 and 4.1.

The remaining of this article is organized in the following way. In Section 2, we recall various notions of cutoffs in [7]. In Section 3, the total variation mixing of Ehrenfest chains is discussed and a path comparison technique is introduced in the proof of the main theorem. In Section 4, we treat the L^p-cutoff with a precise estimation on the L^p-norm of eigenfunctions. In Section 5, we put some remarks summarizing from the proof and address a connection of Theorem 1.1 to well-studied results. In Appendix A, we derive the limiting distribution of the hitting probability for the simple random walk on $\mathbb{Z}$. This is not only of interests by itself but also plays an important role in proving the total variation cutoff. The other essential techniques are relegated to the Appendix B.

2. Cutoffs

Throughout the remaining of this paper, we let (Ω, K, π, μ) denote a time-homogeneous irreducible Markov chain on Ω with one-step transition matrix K, stationary distribution π and initial distribution μ. Write (Ω, H_t, π, μ) as the continuous time Markov chain associated with (Ω, K, π, μ) if $H_t = e^{-t(I-K)}$, the semigroup associated with K. If the chain starts at state x, we write (Ω, K, π, x) and (Ω, H_t, π, x) instead. For any two sequences of positive numbers, say t_n, s_n, the notation $s_n = O(t_n)$ means that there are $N > 0$ and $C > 0$ such that $s_n \le C t_n$ for all $n \ge N$. If both $s_n = O(t_n)$ and $t_n = O(s_n)$ hold, we simply write $t_n \asymp s_n$. If $t_n/s_n \to 1$ as $n \to \infty$, write $t_n \sim s_n$ for short.

First, we recall the definition of cutoff in [7] and write it in L^p-distance.

Definition 2.1. Let $\mathcal{F} = \{(\Omega_n, K_n, \pi_n, \mu_n) : n = 1, 2, \ldots\}$ be a family of irreducible finite Markov chains. For $p \in (1, \infty]$, the family $\mathcal{F}$ is said to present:

(1) An L^p-precutoff if there is a sequence $t_n > 0$ and constants $0 < A < B$ such that

$$\lim_{n \to \infty} D_{n,p}(\mu_n, B_n) = 0, \qquad \liminf_{n \to \infty} D_{n,p}(\mu_n, A_n) > 0,$$

where $B_n = \inf\{j \ge 0 : j > Bt_n\}$ and $A_n = \sup\{j \ge 0 : j < At_n\}$.

(2) An L^p-cutoff if there is a sequence $t_n > 0$ such that, for all $\epsilon \in (0, 1)$,

$$\lim_{n \to \infty} D_{n,p}(\mu_n, \overline{k_n}(\epsilon)) = 0, \qquad \lim_{n \to \infty} D_{n,p}(\mu_n, \underline{k_n}(-\epsilon)) = \infty,$$

where $\overline{k_n}(\epsilon) = \inf\{j \ge 0 : j > (1+\epsilon)t_n\}$ and $\underline{k_n}(\epsilon) = \sup\{j \ge 0 : j < (1+\epsilon)t_n\}$.

(3) A $(t_n, b_n)L^p$-cutoff if $t_n > 0$, $b_n > 0$, $b_n = o(\overline{t_n})$ and

$$\lim_{c \to \infty} \overline{F_p}(c) = 0, \qquad \lim_{c \to -\infty} \underline{F_p}(c) = \infty,$$

where

$$\overline{F_p}(c) = \limsup_{n \to \infty} D_{n,p}(\mu_n, \overline{k}(n, c)), \qquad \underline{F_p}(c) = \liminf_{n \to \infty} D_{n,p}(\mu_n, \underline{k}(n, c)),$$

and $\overline{k}(n, c) = \inf\{j \ge 0 : j > t_n + cb_n\}$ and $\underline{k}(n, c) = \sup\{j \ge 0 : j < t_n + cb_n\}$.

The definition of cutoffs for $p = 1$ is the same except the replacement of the limit ∞ with 2. In (2) and (3), t_n is called an L^p-cutoff time and b_n is called a window with respect to t_n.

Remark 2.1. If t_n tends to infinity, it makes no difference to replace A_n with $\lfloor At_n \rfloor$ or $\lceil At_n \rceil$, and so for the replacements of $B_n, \overline{k_n}(\epsilon), \underline{k_n}(\epsilon), \overline{k}(n, c)$, and $\underline{k}(n, c)$. In the continuous time case, the definition of cutoffs follows in the same way with $A_n = At_n, B_n = Bt_n, \overline{k_n}(\epsilon) = \underline{k_n}(\epsilon) = (1 + \epsilon)t_n$ and $\overline{k}(n, c) = \underline{k}(n, c) = t_n + cb_n$.

For any Markov chain, how fast of the convergence to the stationary distribution can also be captured by the following simple concept.

Definition 2.2. Let (Ω, K, π, μ) be an irreducible finite Markov chain and $p \in [1, \infty]$. For $\epsilon > 0$, the ϵ-L^p-mixing time (or briefly the L^p-mixing time) is defined to be

$$T_p(\mu, \epsilon) := \inf\{t \geq 0 : D_p(\mu, t) \leq \epsilon\},$$

where the right side is set to be infinity if the infimum is taken on an empty set. If (Ω, H_t, π, μ) is the continuous time chain associated with K, write the L^p-mixing time as

$$T_p^c(\mu, \epsilon) := \inf\{t \geq 0 : D_p^c(\mu, t) \leq \epsilon\},$$

where $D_p^c(\mu, t)$ is the L^p-distance between μH_t and π.

The concept of cutoff can also be described using the notion of mixing time. For instance, assuming $T_{n,p}(\epsilon) \to \infty$ for some $\epsilon > 0$, a family of irreducible Markov chains has an L^p-cutoff if and only if

$$\lim_{n \to \infty} T_{n,p}(\mu_n, \epsilon)/T_{n,p}(\mu_n, \delta) = 1, \quad \forall \epsilon, \delta \in (0, M_p),$$

where $M_p = \infty$ if $p > 1$ and $M_1 = 2$. See [7, Propositions 2.3 and 2.4] for further details and relationships.

We end this section by introducing the following lemma and corollary, which will be used in proving the main results.

Lemma 2.1. *Let $\mathcal{F} = \{(\Omega_n, K_n, \pi_n, \mu_n) : n = 1, 2, \ldots\}$ be a family of irreducible and aperiodic Markov chains and $p \in [1, \infty]$. Suppose that there is $\epsilon > 0$ and $a_n \to \infty$ such that $T_{n,p}(\mu_n, \epsilon) \asymp a_n$ and $T_{n,p}(\mu_n, \delta) = O(a_n)$ for all $0 < \delta < \epsilon$. Then, the following are equivalent.*

(1) *$\mathcal{F}$ has no L^p-precutoff.*
(2) *For all $c > 0$,*

$$\limsup_{n \to \infty} D_{n,p}(\mu_n, \lfloor ca_n \rfloor) > 0.$$

(3) *As $\delta \to 0$,*

$$\limsup_{n \to \infty} \frac{T_{n,p}(\mu_n, \delta)}{a_n} \to \infty.$$

Proof. See Appendix B. $\square$

The following corollary comes immediate from the above lemma.

Corollary 2.2. *As in the setting of Lemma* 2.1, *the following are equivalent.*

(1) *No subfamily of $\mathcal{F}$ has an L^p-precutoff.*
(2) *For all $c > 0$,*

$$\liminf_{n\to\infty} D_{n,p}(\mu_n, \lfloor ca_n \rfloor) > 0.$$

(3) *As $\delta \to 0$,*

$$\liminf_{n\to\infty} \frac{T_{n,p}(\mu_n, \delta)}{a_n} \to \infty.$$

Lemma 2.1 and Corollary 2.2 also hold in the continuous time case without the assumptions $T_{n,p}(\mu_n, \epsilon) \to \infty$ and $a_n \to \infty$. It makes no difference to replace $\lfloor ca_n \rfloor$ with $\lceil ca_n \rceil$ Lemma 2.1(2) and Corollary 2.2(2).

3. The total variation cutoff of Ehrenfest chains

This section is dedicated to the total variation cutoff of Ehrenfest chains. For $n \geq 1$, let K_n be the transition matrix in (1.2), K_n' be the modification of K_n in Theorem 1.1 and $H_{n,t} = e^{-t(I-K_n)}$ be the semigroup associated with K_n. Referring to the setting of (1.1), let $D_{n,\mathrm{TV}}(x_n, t)$, $D_{n,\mathrm{TV}}^c(x_n, t)$ be respectively the total variation distances between $(K_n')^t$, $H_{n,t}$ and π_n with initial state x_n, and let $T_{n,\mathrm{TV}}(x_n, \epsilon)$, $T_{n,\mathrm{TV}}^c(x_n, \epsilon)$ be the corresponding mixing times. For $p \in [1, \infty]$, let $D_{n,p}$, $D_{n,p}^c$ and $T_{n,p}$, $T_{n,p}^c$ be the L^p-distances and the L^p-mixing time in the discrete and continuous time cases.

Theorem 3.1. *Consider the families $\mathcal{F} = \{(\Omega_n, K_n', \pi_n, x_n) : n = 1, 2, \ldots\}$ and $\mathcal{F}_c = \{(\Omega_n, H_{n,t}, \pi_n, x_n) : n = 1, 2, \ldots\}$. The following are equivalent.*

(1) *$\mathcal{F}$ (resp, $\mathcal{F}_c$) has a total variation precutoff.*
(2) *$\mathcal{F}$ (resp, $\mathcal{F}_c$) has a total variation cutoff.*
(3) *$|n - 2x_n|/\sqrt{n} \to \infty$.*

Furthermore, if (3) holds, then both $\mathcal{F}$ and $\mathcal{F}_c$ have a (t_n, n) total variation cutoff with

$$t_n = \frac{n}{2} \log \frac{|n - 2x_n|}{\sqrt{n}}.$$

Remark 3.1. The window size n is optimal in the sense that, if $\mathcal{F}$ or $\mathcal{F}_c$ has a (t_n, b_n) total variation cutoff, then $n = O(b_n)$. See [7] for details.

Proof of Theorem 3.1. (3) $\Rightarrow$ (2) and the (t_n, n) total variation cutoff under (3) was given by [8]. (2) $\Rightarrow$ (1) follows from the definition. For (1) $\Rightarrow$ (3), we assume (3) fails and prove $\mathcal{F}$ and $\mathcal{F}_c$ have no total variation precutoff. It suffices to show that, if $|x_n - n/2|/\sqrt{n}$ is bounded, then no subfamily of $\mathcal{F}$ and $\mathcal{F}_c$ has a total variation precutoff. The proof consists of three steps.

Step 1: *Bounding the total variation from above.* Note that the total variation distance is bounded above by the chi-square distance. That is,

$$2D_{n,\mathrm{TV}}(x, t) \leq D_{n,2}(x, t), \qquad 2D_{n,\mathrm{TV}}^c(x, t) \leq D_{n,2}^c(x, t).$$

Using the reversibility of K_n and Lemma B.1, the L^2-distance can be expressed as follows.

$$[D_{n,2}(x,t)]^2 = \sum_{i=1}^{n} |\psi_{n,i}(x)|^2 \left(1 - \frac{2i}{n+1}\right)^{2t}$$

$$\leq 2 \sum_{i=1}^{\lfloor n/2 \rfloor} |\psi_{n,i}(x)|^2 e^{-4it/(n+1)} + e^{-4tn/(n+1)},$$

where $\psi_{n,i}$ is the function defined in (B.1) and the inequality applies the identity $\psi_{n,n-i}(x) = (-1)^x \psi_{n,i}(x)$ for all $x, i \in \{0, 1, \ldots, n\}$. It is worthwhile to note that the summation in the last line is also an upper bound for the continuous time case since

$$[D_{n,2}^c(x,t)]^2 = \sum_{i=1}^{n} |\psi_{n,i}(x)|^2 e^{-4it/n}$$

$$\leq 2 \sum_{i=1}^{\lfloor n/2 \rfloor} |\psi_{n,i}(x)|^2 e^{-4it/(n+1)} + e^{-4nt/(n+1)}.$$

Recall [8, Eq. (6.6)] as follows.

$$\psi_{n,i+1}(x) = \frac{n - 2x}{\sqrt{n}} A_{n,i} \psi_{n,i}(x) - B_{n,i} \psi_{n,i-1}(x), \tag{3.1}$$

where

$$A_{n,i} = \sqrt{\frac{n}{(i+1)(n-i)}}, \qquad B_{n,i} = \sqrt{\frac{i(n-i+1)}{(i+1)(n-i)}}.$$

Obviously, for $n \geq 2$ and $1 \leq i < n$, $A_{n,i} \leq 1$ and $B_{n,i} \leq 2$. By setting $r = 2 + \sup_n\{|n - 2x_n|/\sqrt{n}\} < \infty$, we obtain

$$|\psi_{n,i+1}(x_n)| \leq (r-2)|\psi_{n,i}(x_n)| + 2|\psi_{n,i-1}(x_n)|, \quad \forall 1 \leq i < n.$$

Along with the following boundary condition,

$$|\psi_{n,0}(x_n)| = 1, \qquad |\psi_{n,1}(x_n)| = |n - 2x_n|/\sqrt{n} \leq (r-2),$$

the above inequality yields

$$|\psi_{n,i}(x_n)| \leq r^i, \quad \forall 0 \leq i \leq n.$$

Putting this back to the computation of the L^2-distance, one has

$$\max\{D_{n,\mathrm{TV}}(x_n, N(n+1)), D_{n,\mathrm{TV}}^c(x_n, N(n+1))\}$$

$$\leq \frac{1}{2} \left(2 \sum_{i=1}^{\lfloor n/2 \rfloor} r^{2i} e^{-4iN} + e^{-4nN}\right)^{1/2} < \left(\frac{1}{2} \sum_{i=1}^{\infty} r^{2i} e^{-4iN}\right)^{1/2},$$

where N is any positive integer. Using the last inequality in the above, we obtain

$$\max\{T_{n,\mathrm{TV}}(x_n, \epsilon), T_{n,\mathrm{TV}}^c(x_n, \epsilon)\} \leq \left\lceil \frac{1}{2} \log \frac{r}{\epsilon} \right\rceil (n+1), \quad \forall \epsilon \in (0,1), n \geq 2.$$

Step 2: *Bounding the total variation from below: Discrete time case.* Note that the evolution of the chain under K_n' is very similar to that under K_n. Intuitively, the set of even integers less than

$(n + 1)$, say A_n, is an appropriate testing set for the total variation due to the periodicity of K_n. In fact, $2 \cdot \mathbf{1}_{A_n} - 1 = \psi_{n,n}$ and

$$
\begin{aligned}
D_{n,\mathrm{TV}}(x_n, t) &\geq |(K'_n)^t(x_n, A_n) - \pi_n(A_n)| \\
&= \frac{1}{2} |(K'_n)^t(x_n, \psi_{n,n})| \geq \frac{1}{2} e^{-4t/(n+1)},
\end{aligned}
\tag{3.2}
$$

for $n \geq 3$. This implies, for $0 < \epsilon \leq 1/(2e^4)$ and $n \geq 3$,

$$
T_{n,\mathrm{TV}}(x_n, \epsilon) \geq \left\lfloor \frac{1}{4} \log \frac{1}{2\epsilon} \right\rfloor (n + 1).
$$

Note that such a lower bound is independent of the initial state x_n.

Along with the upper bound in Step 1, we obtain $T_{n,\mathrm{TV}}(x_n, 1/(2e^4)) \asymp n$ and $T_{n,\mathrm{TV}}(x_n, \epsilon) = O_\epsilon(n)$ for all $\epsilon < 1/(2e^4)$. Using the last inequality of (3.2), it is easy to see that, for any $c \geq 1$ and $n \geq 1$,

$$
D_{n,\mathrm{TV}}(x_n, \lfloor cn \rfloor) \geq D_{n,\mathrm{TV}}(x_n, \lfloor 2c \rfloor (n + 1)) \geq \frac{1}{2} e^{-4\lfloor 2c \rfloor}.
$$

By Corollary 2.2, no subfamily of $\mathcal{F}$ has a total variation precutoff.

Step 3: *Bounding the total variation from below: Continuous time case.* Again, we suppose $|n - 2x_n|/\sqrt{n}$ is bounded. It has been developed in Step 1 that $T^c_{n,\mathrm{TV}}(x_n, \epsilon) = O_\epsilon(n)$ for all $\epsilon \in (0, 1)$. By Corollary 2.2, it suffices to show that

$$
\liminf_{n \to \infty} D^c_{n,\mathrm{TV}}(x_n, cn) > 0, \quad \forall c > 0.
\tag{3.3}
$$

The trick used in Step 2 does not work for the continuous time case. Our policy is as follows. First, we compare the original discrete time Ehrenfest chain K_n with the simple random walk on $\mathbb{Z}$. The comparison will generate a lower bound on the total variation distance related to the first passage time in Appendix A. This will lead to (3.3).

Observe that, for any $A \subset \Omega_n$ and $t \geq 0$,

$$
D^c_{n,\mathrm{TV}}(x_n, t) \geq H_{n,t}(x_n, A) - \pi_n(A) = \sum_{i=0}^{\infty} \left(e^{-t} \frac{t^i}{i!} \right) K^i_n(x_n, A) - \pi_n(A).
\tag{3.4}
$$

By the symmetry of K_n and the boundedness of $|x_n - n/2|/\sqrt{n}$, it loses no generality to assume that $n/4 \leq x_n \leq n/2$ for all $n \geq 1$. Concerning the subfamily of $\mathcal{F}_c$, it suffices to deal with the following subcases.

$$
(n/2 - x_n)/\sqrt{n} \to a \in [0, \infty), \quad \text{as } n \to \infty.
\tag{3.5}
$$

The next proposition is helpful in the selection of the testing set A.

Proposition 3.2. *Let K_n be the transition matrix on Ω_n defined by (1.2). Suppose μ_n is a probability concentrated on $A = \{0, 1, \ldots, \lceil n/2 \rceil\}$. Then, $\mu_n K^t_n(A) \geq 1/2$ for all $t \geq 0$.*

See Appendix B for a proof of this proposition. Now, let $A = \{0, 1, \ldots, \lceil n/2 \rceil\}$. Clearly, $\pi_n(A) \leq 1/2 + \pi_n(\lceil n/2 \rceil)$ and $\pi_n(\lceil n/2 \rceil) \sim (\pi n/2)^{-1/2}$. Let T be the first time the chain K_n hits state $\lfloor n/2 \rfloor$. By the strong Markov property, we have

$$
K^i_n(x_n, A) = \sum_{j=0}^{i} K^{i-j}_n(\lfloor n/2 \rfloor, A) \mathbb{P}_{x_n}(T = j) + \mathbb{P}_{x_n}(T > i) \geq \frac{1}{2} + \frac{1}{2} \mathbb{P}_{x_n}(T > i).
$$

Putting this back to (3.4) yields, for $m \geq 0$,

$$D^c_{n,\mathrm{TV}}(x_n, t) \geq \frac{1}{2} \left(e^{-t} \sum_{i=0}^{m} \frac{t^i}{i!} \right) \mathbb{P}_{x_n}(T > m) - \pi_n(\lceil n/2 \rceil). \qquad (3.6)$$

Next, we use Theorem A.1 to bound $\mathbb{P}_{x_n}(T > m)$ from below. Consider the simple random walk on $\mathbb{Z}$. For $m \geq 1$, $k \geq 1$ and $i \in \mathbb{Z}$, let $\mathcal{P}(m, k, i)$ be the set containing paths of length m starting from 0, ending at i and staying in $\{0, \pm 1, \pm 2, \ldots, \pm(k-1)\}$ up to time m and write

$$x + \mathcal{P}(m, k, i) = \{(x + w_0, x + w_1, \ldots, x + w_m) : (w_0, w_1, \ldots, w_m) \in \mathcal{P}(m, k, i)\}.$$

Clearly,

$$\mathbb{P}_{x_n}(T > m) \geq \sum_{i=0}^{\lfloor n/2 \rfloor - x_n - 1} \mathbb{P}_{x_n}(x_n + \mathcal{P}(m, \lfloor n/2 \rfloor - x_n, i))$$

Let $\mathbb{P}'$ be the probability where the simple random walk on $\mathbb{Z}$ starting from the origin sits. For any path $w = (w_0, w_1, \ldots, w_m) \in \mathcal{P}(m, k, i)$ with $|i| < k$, one may partition the edges $\{(w_j, w_{j+1}) : 0 \leq k < m\}$ into two subsets, say $B_1(w)$ and $B_2(w)$, where $B_1(w) = \{(j, j+1) : 0 \leq j < i\}$ for $i > 0$, $B_1(w) = \{(j, j-1) : 0 \geq j > i\}$ for $i < 0$, and $B_2(w)$ is a union of pairs in the form $\{(j, j+1), (j+1, j)\}$ with $-k < j < k-1$. Note that, for $2x_n - n/2 \leq j \leq n/2$,

$$1 - \frac{j}{n} \geq \frac{j}{n} \geq \frac{1}{2} \left(\frac{4x_n}{n} - 1 \right) = \frac{1}{2} \left(1 - \frac{2(n - 2x_n)}{n} \right)$$

and

$$\left(1 - \frac{j}{n} \right) \frac{j+1}{n} \geq \frac{1}{4} \left[1 - \left(\frac{n - 2j}{n} \right)^2 \right] \geq \frac{1}{4} \left[1 - 4 \left(\frac{n - 2x_n}{n} \right)^2 \right].$$

This leads to $\mathbb{P}_{x_n}(w) \geq c_n(m)\mathbb{P}'(w)$ for all $w \in x_n + \mathcal{P}(m, \lfloor n/2 \rfloor - x_n, i - x_n)$ and $2x_n - n/2 \leq i \leq n/2$ with

$$c_n(m) = \left[1 - 4 \left(\frac{n - 2x_n}{n} \right)^2 \right]^{m/2} \left(1 - \frac{2(n - 2x_n)}{n} \right)^{n/2 - x_n}.$$

Letting N be any positive integer and $m = Nn$, we obtain

$$\mathbb{P}_{x_n}(T > Nn) \geq c_n(Nn)\mathbb{P}'(T_{\lfloor n/2 \rfloor - x_n} > Nn),$$

where T_l is the first passage time to $\{\pm l\}$ of the simple random walk on $\mathbb{Z}$. See (A.1) for details. Putting this back to (3.6) yields

$$D^c_{n,\mathrm{TV}}(x_n, t) \geq \frac{1}{2} \left(e^{-t} \sum_{i=0}^{Nn} \frac{t^i}{i!} \right) c_n(Nn)\mathbb{P}'(T_{\lfloor n/2 \rfloor - x_n} > Nn) - \pi_n(\lfloor n/2 \rfloor).$$

By Theorem A.1 and Lemma B.4, if $a > 0$ in the setting of (3.5), then

$$\liminf_{n \to \infty} D^c_{n,\mathrm{TV}}(x_n, cn) \geq \frac{\alpha_N}{2} e^{-(8N+4)a^2}, \quad \forall N > c,$$

where $\alpha_N = \frac{4}{\pi} \sum_{k=0}^{\infty} \frac{(-1)^k}{2k+1} e^{-N(2k+1)^2 \pi^2 / (8a^2)} > 0$. Hence, for $a > 0$, no subfamily of $\mathcal{F}_c$ has a total variation precutoff.

In the end, we deal with the subcase $a = 0$. First, we write

$$K_n^t(x, y)/\pi_n(y) - 1 = \sum_{i=1}^n \psi_{n,i}(x)\psi_{n,i}(y)\beta_{n,i}^t.$$

See [15, Lemma 1.3.3] for a proof. Applying this identity to the case $(K_n')^t$ and $H_{n,t}$ gives

$$\frac{(K_n')^t(x, y)}{\pi_n(y)} - 1 = \sum_{i=1}^n \psi_{n,i}(x)\psi_{n,i}(y)\left(\frac{1 + n\beta_{n,i}}{n + 1}\right)^t \tag{3.7}$$

and

$$\frac{H_{n,t}(x, y)}{\pi_n(y)} - 1 = \sum_{i=1}^n \psi_{n,i}(x)\psi_{n,i}(y)e^{-t(1-\beta_{n,i})} \tag{3.8}$$

For $n \geq 1$, set

$$H_{n,t}(x_n, y)/\pi_n(y) - 1 = f_n(t, y) + g_n(t, y),$$

where

$$f_n(t, y) = \psi_{n,2}(x_n)e^{-t(1-\beta_{n,2})}\psi_{n,2}(y)$$

and

$$g_n(t, y) = \sum_{i=1,i\neq 2}^n \psi_{n,i}(x_n)e^{-t(1-\beta_{n,i})}\psi_{n,i}(y).$$

By Jensen's inequality, one can see that

$$2D_{n,\mathrm{TV}}^c(x_n, t) = \|f_n(t, \cdot) + g_n(t, \cdot)\|_{L^1(\pi_n)} \geq \|f_n(t, \cdot)\|_{L^1(\pi_n)} - \|g_n(t, \cdot)\|_{L^2(\pi_n)}.$$

It remains to prove that, for all $c > 0$,

$$\liminf_{n\to\infty}\left[\|f_n(cn, \cdot)\|_{L^1(\pi_n)} - \|g_n(cn, \cdot)\|_{L^2(\pi_n)}\right] > 0.$$

Note that

$$\|g_n(t, \cdot)\|_{L^2(\pi_n)} = \left(\frac{(n - 2x_n)^2}{n}e^{-4t/n} + \sum_{i=3}^n |\psi_{n,i}(x_n)|^2 e^{-4it/n}\right)^{1/2}.$$

Recall that, in Step 1, if $r = 2 + \sup_n \{|n - 2x_n|/\sqrt{n}\} < \infty$, then $|\psi_{n,i}(x_n)| \leq r^i$ for all $0 \leq i \leq n$. Putting this back to the $L^2(\pi_n)$-norm of $g_n(t, \cdot)$ gives

$$\|g_n(cn, \cdot)\|_{L^2(\pi_n)} \leq \left(\frac{(n - 2x_n)^2}{n}e^{-4c} + \frac{(r^2e^{-4c})^3}{1 - r^2e^{-4c}}\right)^{1/2},$$

provided $r < e^{2c}$. Also, it is an easy exercise to compute $|\psi_{n,2}(x_n)| \sim 1/\sqrt{2}$ and

$$\|\psi_{n,2}\|_{L^1(\pi_n)} \geq \frac{1}{2}\pi_n\left(\{x : |x - n/2| < \sqrt{n}/4\}\right) \sim \frac{1}{\sqrt{2\pi}}\int_0^{1/2} e^{-u^2/2}du \geq \frac{1}{12}.$$

Under the assumption $(n/2 - x_n)/\sqrt{n} \to 0$, if $r < e^{2c}$, then

$$\liminf_{n\to\infty} \left[\|f_n(cn, \cdot)\|_{L^1(\pi_n)} - \|g_n(cn, \cdot)\|_{L^2(\pi_n)} \right]$$

$$\geq \frac{1}{12\sqrt{2}} e^{-4c} - \frac{r^3}{\sqrt{1 - r^2 e^{-4c}}} e^{-6c} = e^{-4c} \left(\frac{1}{12\sqrt{2}} - \frac{r^3}{\sqrt{1 - r^2 e^{-4c}}} e^{-2c} \right).$$

The last term is positive for large c and this implies

$$\liminf_{n\to\infty} D^c_{n,\mathrm{TV}}(x_n, cn) > 0, \quad \forall c > 0.$$

Hence, in the case $a = 0$, no subfamily of $\mathcal{F}_c$ has a total variation precutoff. $\quad\square$

4. The L^p-cutoff of Ehrenfest chains

This section is contributed to the development of the L^p-cutoff of Ehrenfest chains with $p \in (1, \infty)$. The main theorem states as follows.

Theorem 4.1. *Let $\mathcal{F}$ and $\mathcal{F}_c$ be the families in Theorem 3.1. For $p \in (1, \infty)$, the following are equivalent.*

(1) *$\mathcal{F}$ (resp. $\mathcal{F}_c$) has an L^p-precutoff.*
(2) *$\mathcal{F}$ (resp. $\mathcal{F}_c$) has an L^p-cutoff.*
(3) *$|x_n - n/2|/\sqrt{n} \to \infty$.*

Moreover, if (3) holds, then both $\mathcal{F}$ and $\mathcal{F}_c$ have a $(t_n, n)L^p$-cutoff with

$$t_n = \frac{n}{2} \log \frac{|n - 2x_n|}{\sqrt{n}}.$$

Proof. In this proof, the $L^p(\pi)$-norm of f is written as $\|f\|_p$. Obviously, (2) $\Rightarrow$ (1) comes immediate from Definition 2.1 for all $1 < p < \infty$. For (3) $\Rightarrow$ (2) and the $(t_n, n)L^p$-cutoff, we set

$$\overline{F_p}(a) = \limsup_{n\to\infty} D_{n,p}(x_n, t_n + an), \qquad \underline{F_p}(a) = \liminf_{n\to\infty} D_{n,p}(x_n, t_n + an)$$

and

$$\overline{G_p}(a) = \limsup_{n\to\infty} D^c_{n,p}(x_n, t_n + an), \qquad \underline{G_p}(a) = \liminf_{n\to\infty} D^c_{n,p}(x_n, t_n + an).$$

The case $p = 2$ is given in [8]. By the monotonicity (in p) of L^p-norm, it remains to show that, for $1 < p < 2$,

$$\lim_{a\to-\infty} \min\{\underline{F_p}(a), \underline{G_p}(a)\} = \infty,$$

and, for $2 < p < \infty$,

$$\lim_{a\to\infty} \max\{\overline{F_p}(a), \overline{G_p}(a)\} = 0.$$

Case 1: $(1 < p < 2)$ Set $q = (1 - 1/p)^{-1}$. As a consequence of the central limit theorem, we have

$$\|\psi_{n,1}\|_q = \left(\sum_{x=0}^{n} \left(\frac{|n - 2x|}{\sqrt{n}} \right)^q \pi_n(x) \right)^{1/q} \to C_q := [\mathbb{E}(|X|^q)]^{1/q} < \infty,$$

where X is a standard normal random variable and $\mathbb{E}$ denotes the expectation. By (3.7) and (3.8), this implies

$$\underline{F_p}(a) \geq \liminf_{n \to \infty} \frac{|\langle (K_n')^{t_n + an}(x_n, \cdot)/\pi_n - 1, \psi_{n,1} \rangle_{\pi_n}|}{\|\psi_{n,1}\|_q} = e^{-2a}/C_q$$

and

$$\underline{G_p}(a) \geq \liminf_{n \to \infty} \frac{|\langle H_{n,t_n + an}(x_n, \cdot)/\pi_n - 1, \psi_{n,1} \rangle_{\pi_n}|}{\|\psi_{n,1}\|_q} = e^{-2a}/C_q.$$

This proves the desired $(t_n, n)L^p$-cutoff.

Case 2: $(2 < p < \infty)$ Using the fact $\psi_{n,n-i}(x) = (-1)^x \psi_{n,i}(x)$, the right sides of (3.7) and (3.8) yields

$$D_{n,p}(x_n, t) \leq 2 \sum_{i=1}^{\lceil n/2 \rceil} |\psi_{n,i}(x_n)| \|\psi_{n,i}\|_p \left(1 - \frac{2i}{n+1} \right)^t + \left(1 - \frac{2}{n+1} \right)^t \leq 2d_p(n, t)$$

and

$$D_{n,p}^c(x_n, t) \leq 2 \sum_{i=1}^{\lceil n/2 \rceil} |\psi_{n,i}(x_n)| \|\psi_{n,i}\|_p e^{-2it/n} + e^{-2t} \leq 2d_p(n, t),$$

where

$$d_p(n, t) = \sum_{i=1}^{\lceil n/2 \rceil} |\psi_{n,i}(x_n)| \|\psi_{n,i}\|_p e^{-2it/(n+1)} + e^{-2t/(n+1)}.$$

It remains to compute $\|\psi_{n,i}\|_p$. By (3.1), it is easy to see

$$|\psi_{n,i+1}(x)| \leq \left(\sqrt{\frac{2}{i+1}} \times \frac{|n - 2x|}{\sqrt{n}} \right) |\psi_{n,i}(x)| + |\psi_{n,i-1}(x)|, \quad \forall i \leq n/2.$$

With the initial conditions, $\psi_{n,0} \equiv 1$ and $\psi_{n,1}(x) = (n - 2x)/\sqrt{n}$, one can prove inductively

$$|\psi_{n,i}(x)| \leq \sqrt{\frac{2^i}{i!}} \prod_{j=1}^{i} \left(|\psi_{n,1}(x)| + \sqrt{\frac{j}{2}} \right), \quad \forall x \in \Omega_n, \ i \leq n/2. \tag{4.1}$$

For convenience, we write $i! = \alpha_i i^{i+1/2} e^{-i}$. By Stirling's formula, we may choose $\beta > 1$ such that

$$i^{i+1/2} e^{-i}/\beta \leq i! \leq \beta i^{i+1/2} e^{-i}, \quad \forall i \geq 1. \tag{4.2}$$

In the above setting, (4.1) gives

$$|\psi_{n,i}(x)| \leq (2e)^{i/2} i^{-1/4} \beta^{1/2} \left(|\psi_{n,1}(x)| i^{-1/2} + 1 \right)^i, \tag{4.3}$$

which implies

$$\|\psi_{n,i}\|_p^p \leq (2e)^{pi/2} i^{-p/4} \beta^{p/2} \pi_n \left[\left(|\psi_{n,1}| i^{-1/2} + 1 \right)^{pi} \right]$$

$$\leq (2e)^{pi/2} i^{-p/4} \beta^{p/2} 2^{pi} \left[i^{-pi/2} \pi_n \left(|\psi_{n,1}|^{pi} \right) + 1 \right],$$

where the last inequality uses the fact $(s + t)^r \leq 2^{r-1}(s^r + t^r)$ for any $s > 0, t > 0$ and $r \geq 1$. By Lemma B.5, one may choose $C > 1$ such that

$$\pi_n(|\psi_{n,1}|^{pi}) \leq C4^{pi} \Gamma\left(\frac{pi+1}{2}\right) \leq 5\beta C4^{pi} pi[(pi)/(2e)]^{pi/2}, \qquad (4.4)$$

where $\Gamma(\cdot)$ is the Gamma function. To see the last inequality, recall that $\Gamma(t + 1) = t\Gamma(t)$. Observe that $\sup_{1 \leq \alpha \leq 2} \Gamma(\alpha) \leq 2$ and $e^{3/2} \leq 5$. By (4.2), we obtain

$$\Gamma\left(\frac{pi+1}{2}\right) = \frac{pi-1}{2}\Gamma\left(\frac{pi-1}{2}\right) \leq pi\left(\left\lceil\frac{pi-3}{2}\right\rceil!\right)$$
$$\leq \beta pi(\lceil(pi-3)/2\rceil)^{\lceil(pi-3)/2\rceil+1/2}e^{-\lceil(pi-3)/2\rceil}$$
$$\leq 5\beta pi(pi/2e)^{pi/2}.$$

Plugging (4.4) back to the upper bound for $\|\psi_{n,i}\|_p^p$, one has

$$\|\psi_{n,i}\|_p \leq (2e)^{i/2}i^{-1/4}\beta^{1/2}2^i\left[5\beta C4^{pi}pi(p/(2e))^{pi/2}+1\right]^{1/p} \leq 10\beta Ci^{1/4}(8p)^i$$

and, applying (4.3) with $x = x_n$, this leads to

$$d_p(n, t) \leq 10\beta^2 C \sum_{i=1}^{\lceil n/2 \rceil} (20p)^i \left(|\psi_{n,1}(x_n)|+1\right)^i e^{-2it/(n+1)} + e^{-2t/(n+1)}, \qquad (4.5)$$

where $8\sqrt{2e} < 20$ is used.

Finally, let $a > 1$. It is obvious that, for n large enough,

$$t_n = \frac{n}{2}\log|\psi_{n,1}(x_n)|, \qquad t_n + an \geq \frac{n+1}{2}\log|\psi_{n,1}(x_n)| + \frac{n+1}{2}(a-1).$$

By (4.5), this implies

$$d_p(n, t_n + an) \leq 10\beta^2 C \sum_{i=1}^{\lceil n/2 \rceil} \left(\frac{20p}{e^{a-1}} \times \frac{|\psi_{n,1}(x_n)|+1}{|\psi_{n,1}(x_n)|}\right)^i + \exp\{-|\psi_{n,1}(x_n)|\}.$$

Letting $n \to \infty$ yields that, for $e^{a-1} > 20p$,

$$\max\{\overline{F_p}(a), \overline{G_p}(a)\} \leq 20\beta^2 C \sum_{i=1}^{\infty} i(20pe^{1-a})^i = \frac{400\beta^2 Cpe^{1-a}}{1 - 20pe^{1-a}}.$$

This proves the desired cutoff.

For (1) $\Rightarrow$ (3), we assume that $|x_n - n/2|/\sqrt{n}$ is bounded and prove that no subfamily of $\mathcal{F}$ and $\mathcal{F}_c$ has an L^p-precutoff. Set $M = \sup_{n \geq 1}\{|2x_n - n|/\sqrt{n}\} + 1$. By (4.5), if $p > 2$ and $e^a \geq 20Mp$, then

$$\max\{D_{n,p}(x_n, \lceil an \rceil), D_{n,p}^c(x_n, an)\} \leq \frac{400M\beta^2 Cpe^{-a}}{1 - 20Mpe^{-a}} + 2e^{-a}.$$

On one hand, the right side converges to 0 as a tends to infinity. This implies, for all $\epsilon > 0$ and $p < \infty$,

$$T_{n,p}(x_n, \epsilon) = O_\epsilon(n), \qquad T_{n,p}^c(x_n, \epsilon) = O_\epsilon(n).$$

On the other hand, one may conclude from the proof of Theorem 3.1 (Steps 2–3) that

$$\liminf_{n \to \infty} \min\{D_{n,\mathrm{TV}}(x_n, an), D^c_{n,\mathrm{TV}}(x_n, an)\} > 0, \quad \forall a > 0.$$

By Corollary 2.2, no subfamily of $\mathcal{F}$ and $\mathcal{F}_c$ has an L^p-precutoff for $1 < p < \infty$. □

5. Some remarks

In this section, we make some remarks summarizing from the content of the previous sections and establish a connection with known results. First, it is worthwhile to remark from the proofs of Theorems 3.1 and 4.1 that if $|n - 2x_n|/\sqrt{n}$ is bounded, then the L^p-mixing time of Ehrenfest chains is of order n in both discrete and continuous time case for $1 \leq p < \infty$, though the families do not present any cutoff at all. This implies that (3) of Theorems 3.1 and 4.1 is equivalent to $\lambda_n T_{n,p}(x_n, \epsilon) \to \infty$ and $\lambda^c_n T^c_{n,p}(x_n, \epsilon) \to \infty$ for $1 \leq p < \infty$, where λ_n, λ^c_n are respectively the spectral gaps of discrete and continuous time chains. This is consistent with the conjecture proposed by Peres during the ARCC workshop held by AIM in Palo Alto, December 2004.

For the L^∞-cutoff, the equivalence in Theorem 4.1 might fail. Assume that n is even, $x_n = n/2$ and consider the continuous time case. Recall the separation distance as follows.

$$D_{n,\mathrm{sep}}(x, t) = \max_y \left\{ 1 - \frac{(K'_n)^t(x, y)}{\pi_n(y)} \right\}, \qquad D^c_{n,\mathrm{sep}}(x, t) = \max_y \left\{ 1 - \frac{H_{n,t}(x, y)}{\pi_n(y)} \right\}.$$

It is easy to see that the separation distance is closely related to the L^∞-distance. For $n \geq 1$, let L_n be the Markov kernel on $\{0, 1, \ldots, n/2\}$ given by

$$L_n(i, i) = 0, \quad \forall 0 \leq i \leq n/2, \qquad L_n(i, i + 1) = 1 - \frac{i}{n}, \quad \forall 0 \leq i < n/2,$$

and

$$L_n(i + 1, i) = \frac{i + 1}{n}, \quad \forall 0 \leq i < n/2 - 1, \qquad L_n(n/2, n/2 - 1) = 1.$$

Obviously, the stationary distribution of L_n is given by $\widetilde{\pi}_n(i) = 2^{1-n} \binom{n}{i}$ for $i < n/2$ and $\widetilde{\pi}_n(n/2) = 2^{-n} \binom{n}{n/2}$. Set $\widetilde{D}^c_{n,\mathrm{sep}}(x, t)$ be the separation distances between $e^{-t(I-L_n)}(x, \cdot)$ and $\widetilde{\pi}_n$. Then, $D^c_{n,\mathrm{sep}}(n/2, t) = \widetilde{D}^c_{n,\mathrm{sep}}(n/2, t)$. Observe that $I - L_n$ has eigenvalues $4i/n$ with $0 \leq i \leq n/2$. Clearly, the spectral gap of L_n is $\lambda_n = 4/n$ and

$$t_n = \sum_{i=1}^{n/2} \frac{n}{4i} = \frac{n \log n}{4} + O(n).$$

By[10, Theorems 5.1 and 6.1], the family $\mathcal{F}_c$ has a $\left(\frac{1}{4}n \log n, n\right)$ separation cutoff, but has no L^p-precutoff for $1 \leq p < \infty$ according to Theorems 3.1 and 4.1.

Finally, we get our results involved in several well-studied results. For continuous time Markov chains with countable state spaces, Martínez and Ycart [14] provide an equivalence to the cutoff of families with nominated sequence of initial states x_n. According to their framework, a special state, called 0, is selected and the hitting time T_0 to 0 is considered. In [14, Theorem 4.1], they showed that the family of continuous time Markov chains with starting state x_n has a total variation cutoff with cutoff time $\mathbb{E}_{x_n}(T_0)$ if and only if T_0 has a concentration phenomenon as the distance between x_n and 0 tends to infinity. Back to the setting of Ehrenfest chain, if one

selects the typical state $\lfloor n/2 \rfloor$ as 0 and lets $|x_n - n/2|$ tends to infinity, then the equivalence of the total variation cutoff should be achieved through a precise estimation of the average hitting time to $\lfloor n/2 \rfloor$ starting from x_n. A formula on the expected hitting times of birth and death chains is available in [4, Eq. (4.20)] and the comparison of the mixing time and the expected hitting time would be of interest. Such a comparison was made by Lachaud in [12] for the Ornstein–Uhlenbeck process, which can be approximated by Ehrenfest chains.

In [4], Barrera et al. observed that, assuming a drift toward a typical state is associated with an energy function, if the energy well is sufficiently steep with sufficiently smooth walls, descents are abrupt and the hitting time to the typical state starting from n has a concentration phenomenon and, hence, a cutoff exists. This was proved in [4] to be applicable for Ehrenfest chains. If $|x_n - n/2|$ is bounded, the result in Theorem 1.1 can be related to the fact that the energy well is not steep enough between $\lfloor n/2 \rfloor$ and x_n so that no cutoff exists. For discrete time Ehrenfest chains, Bertoncini showed in [5, Chapter 7] that the family starting from the boundary points has a total variation cutoff if there is a concentration phenomenon for the time to access a neighborhood of $\lfloor n/2 \rfloor$ within a distance of order less than $\sqrt{n}$. In some sense, this means that if the initial state is not too close to $\lfloor n/2 \rfloor$, then there is a cutoff, which is a converse of Theorem 1.1.

Acknowledgments

Appendix A. The first passage time of simple random walks on $\mathbb{Z}$

This section is contributed to the estimation of the hitting probability for the simple random walk on integers. A simple random walk is a discrete time Markov chain $(X_n)_{n=0}^{\infty}$ with transition matrix

$$K(i, i+1) = K(i, i-1) = 1/2, \quad \forall i \in \mathbb{Z}.$$

For $m \geq 1$, let T_m be the first passage time to the set $\{\pm m\}$, i.e.

$$T_m = \inf\{n \geq 0 : X_n = m \text{ or } X_n = -m\}. \tag{A.1}$$

For the continuous time case, let $N(t)$ be a Poisson process with parameter 1 and independent of X_n and set $Y_t = X_{N(t)}$. Clearly, Y_t is a realization of the semigroup $H_t = e^{-t(I-K)}$ associated with K and the corresponding first passage time to $\{\pm m\}$ is denoted by

$$\widetilde{T}_m = \inf\{t \geq 0 : Y_t = m \text{ or } Y_t = -m\}. \tag{A.2}$$

The following is the main theorem in this section.

Theorem A.1. *Let $T_m, \widetilde{T}_m$ be the random times defined in (A.1) and (A.2) and $\mathbb{P}_0$ be the conditional probability given the initial state is 0. Then, for $b > 0$,*

$$\lim_{m \to \infty} \mathbb{P}_0(T_m > bm^2) = \lim_{m \to \infty} \mathbb{P}_0(\widetilde{T}_m > bm^2) = \frac{4}{\pi} \sum_{k=0}^{\infty} \frac{(-1)^k}{2k+1} e^{-b(2k+1)^2 \pi^2/8}.$$

Theorem A.1 is a consequence of the convergence of the renormalized processes $(X_{\lfloor m^2 t \rfloor}/m)_{t\geq 0}$ and $(Y_{m^2 t/m})_{t\geq 0}$ toward the Brownian motion on $\mathbb{R}$ in the weak topology with respect to the uniform norm over compact time intervals. The computation of the Laplace transform of the hitting times is a standard approach and the inversion of the Laplace transform leads to a special function whose expansion is given in Theorem A.1. Here, we consider another approach by computing the Laplace transform of the hitting times for $(X_n)_{n\geq 0}$, $(Y_t)_{t\geq 0}$ and determining the limiting distribution directly from a detailed analysis of the sequence of transforms.

The following proposition is useful in characterizing the distribution of the first passage time.

Proposition A.2 (*[6, Section 2]*)**.** *Let K be the transition matrix of an irreducible birth-and-death chain on $\{0, 1, \ldots\}$. For $m \geq 1$, let τ_m and $\widetilde{\tau}_m$ be respectively the first passage times to state m associated with the discrete time and continuous time chains. Let $\lambda_1, \ldots, \lambda_m$ be the eigenvalues of the submatrix of $I - K$ indexed by $\{0, 1, \ldots, m - 1\}$. Then, $\lambda_i \in (0, 2)$ for $1 \leq i \leq m$, $\lambda_i \neq \lambda_j$ for $i \neq j$, and*

$$\mathbb{P}_0(\tau_m > k) = \sum_{i=1}^{m} \left(\prod_{j:j\neq i} \frac{\lambda_j}{\lambda_j - \lambda_i} \right) (1 - \lambda_i)^k \tag{A.3}$$

and

$$\mathbb{P}_0(\widetilde{\tau}_m > t) = \sum_{i=1}^{m} \left(\prod_{j:j\neq i} \frac{\lambda_j}{\lambda_j - \lambda_i} \right) e^{-t\lambda_i}. \tag{A.4}$$

Remark A.1. The right side of (A.4) is exactly $\mathbb{P}(\widetilde{T} > t)$, where $\widetilde{T}$ is a sum of m independent exponential random variables with parameters $\lambda_1, \ldots, \lambda_m$. Assuming $\lambda_i \in (0, 1)$ for all $1 \leq i \leq m$, the right side of (A.3) is equal to $\mathbb{P}(T > k)$, where T is a sum of independent geometric random variables with success probabilities $\lambda_1, \ldots, \lambda_m$.

To prove Theorem A.1, we need the following lemmas of which proofs are deferred to the end of this section.

Lemma A.3. *For $n \geq 1$, let $X_{n,1}, \ldots, X_{n,n}$ be independent exponential random variables with parameters $\lambda_{n,1}, \ldots, \lambda_{n,n}$, where $\lambda_{n,m} = 1 - \cos \frac{(2m-1)\pi}{2n}$. Set $S_n = X_{n,1} + \cdots + X_{n,n}$. Then, S_n/n^2 converges in distribution to a positive continuous random variable with density*

$$f(t) = \frac{\pi}{2} \sum_{m=0}^{\infty} (-1)^m (2m + 1) e^{-t(2m+1)^2\pi^2/8}, \quad \forall t > 0.$$

In particular,

$$\lim_{n\to\infty} \mathbb{P}(S_n/n^2 > t) = \frac{4}{\pi} \sum_{m=0}^{\infty} \frac{(-1)^m}{2m + 1} e^{-t(2m+1)^2\pi^2/8}.$$

Lemma A.4. *For $n \geq 1$, let $X_{n,1}, \ldots, X_{n,n}$ be independent geometric random variables with parameters $\lambda_{n,1}, \ldots, \lambda_{n,n}$, where $\lambda_{n,m} = \frac{1}{2}\left[1 - \cos \frac{(2m-1)\pi}{2n}\right]$. Set $S_n = X_{n,1} + \cdots + X_{n,n}$.*

Then, S_n/n^2 converges in distribution to a positive continuous random variable with density

$$f(t) = \frac{\pi}{4} \sum_{m=0}^{\infty} (-1)^m (2m+1) e^{-t(2m+1)^2 \pi^2/16}, \quad \forall t > 0.$$

In particular,

$$\lim_{n \to \infty} \mathbb{P}(S_n/n^2 > t) = \frac{4}{\pi} \sum_{m=0}^{\infty} \frac{(-1)^m}{2m+1} e^{-t(2m+1)^2 \pi^2/16}.$$

Proof of Theorem A.1. Back to the setting of the simple random walk. Observe that

$$\mathbb{P}_0(T_m > k) = \mathbb{P}_0(|X_i| < m, \ \forall i \le k), \qquad \mathbb{P}_0(\widetilde{T}_m > t) = \mathbb{P}_0(|Y_s| < m, \ \forall s \le t).$$

By the symmetry of the walk starting from 0, one may collapse states $\pm i$ to achieve

$$\mathbb{P}_0(T_m > k) = \mathbb{P}'_0(\tau_m > k), \qquad \mathbb{P}_0(\widetilde{T}_m > t) = \mathbb{P}'_0(\widetilde{\tau}_m > t),$$

where $\mathbb{P}'_0$ is the probability for the birth-and-death chain on $\{0, 1, \ldots\}$ with initial state 0 and transition matrix K' given by

$$K'(0, 1) = 1, \qquad K'(i, i-1) = K'(i, i+1) = 1/2, \quad \forall i \ge 1.$$

Here, τ_m and $\widetilde{\tau}_m$ are the first passage times to state m associated with the discrete time and continuous time chains driven by K'. Applying the method introduced in [11, Section XIV.5], the eigenvalues and eigenvectors for the submatrix of $I - K'$ indexed by $0, 1, \ldots, m-1$ are

$$\lambda_{m,i} = 1 - \cos \frac{(2i-1)\pi}{2m}, \qquad \phi_{m,i}(j) = \cos \frac{(2i-1)j\pi}{2m},$$

for $1 \le i \le m$ and $0 \le j \le m-1$.

We first treat the continuous time case. Let $S_{m,1}, \ldots, S_{m,m}$ be independent exponential random variables with parameters $\lambda_{m,1}, \ldots, \lambda_{m,m}$. Set $S_m = S_{m,1} + \cdots + S_{m,m}$. By Proposition A.2, $\mathbb{P}_0(\widetilde{T}_m > bm^2) = \mathbb{P}(S_m > bm^2)$. As a consequence of Lemma A.3, letting m tend to infinity yields the desired identity in continuous time case.

For the discrete time case, the periodicity of K', which is of period 2, implies $\lambda_{m,i} > 1$ for some i. Consider the lazy walk with transition matrix $\frac{1}{2}(I + K')$. It is clear that the eigenvalues of the submatrix of $I - \frac{1}{2}(I + K')$ indexed by $\{0, \ldots, m-1\}$ are $\lambda_{m,1}/2, \ldots, \lambda_{m,m}/2$, which are contained in $(0, 1)$. To relate the discrete time case and the transition matrix $\frac{1}{2}(I + K')$, let $(X'_n)_{n=0}^{\infty}$ be the birth-and-death chain with transition matrix K' and define $Z_n = X'_{2n}/2$. Obviously,

$$\mathbb{P}'_0(Z_{n+1} = 1 | Z_n = 0) = \mathbb{P}'_0(X'_{2n+2} = 2 | X'_{2n} = 0) = 1/2.$$

For $i > 0$,

$$\mathbb{P}'_0(Z_{n+1} = i + 1 | Z_n = i) = \mathbb{P}'_0(X'_{2n+2} = 2i + 2 | X'_{2n} = 2i) = 1/4$$

and

$$\mathbb{P}'_0(Z_{n+1} = i - 1 | Z_n = i) = \mathbb{P}'_0(X'_{2n+2} = 2i - 2 | X'_{2n} = 2i) = 1/4$$

and, for $i \ge 0$,

$$\mathbb{P}'_0(Z_{n+1} = i | Z_n = i) = \mathbb{P}'_0(X'_{2n+2} = 2i | X'_{2n} = 2i) = 1/2.$$

This implies that given $X_0' = 0$, or equivalently $Z_0 = 0$, $(Z_n)_{n=0}^{\infty}$ is a Markov chain on $\{0, 1, \ldots\}$ with initial state 0 and transition matrix $\frac{1}{2}(I + K')$. Furthermore, by the periodicity of K', if m is even and positive, then

$$\mathbb{P}_0'(\tau_m > k) = \mathbb{P}_0'(X_i' < m, \ \forall i \le k) = \mathbb{P}_0'(Z_i < m/2, \ \forall i \le \lfloor k/2 \rfloor).$$

If m is odd and $m > 1$, then

$$\mathbb{P}_0'(\tau_m > k) = \mathbb{P}_1'(X_i' < m, \ \forall i \le k - 1) = \mathbb{P}_0'(Z_i < (m-1)/2, \ \forall i \le \lfloor (k-1)/2 \rfloor),$$

where the last equality uses the fact that, given $X_0' = 1$, the chain $(X_{2n}' - 1)_{n=1}^{\infty}$ has the same distribution as $(Z_n)_{n=1}^{\infty}$ with $Z_0 = 0$. Let τ_m' be the first passage time to m of the chain $(Z_n)_{n=0}^{\infty}$. Putting all above together yields, for $m > 1$,

$$\mathbb{P}_0(T_m > k) = \mathbb{P}_0'(\tau_m > k) \begin{cases} \ge \mathbb{P}_0'(\tau_{\lfloor m/2 \rfloor}' > k/2) \\ \le \mathbb{P}_0'(\tau_{\lfloor m/2 \rfloor}' > k/2 - 1) \end{cases}.$$

For $m \ge 1$, let $S_{m,1}', \ldots, S_{m,m}'$ be independent geometric random variables with success probabilities $\lambda_{m,1}/2, \ldots, \lambda_{m,m}/2$ and set $S_m' = S_{m,1}' + \cdots + S_{m,m}'$. By Proposition A.2, for any positive integer k,

$$\mathbb{P}_0(T_m > k) \begin{cases} \ge \mathbb{P}(S_{\lfloor m/2 \rfloor}' > k/2) \\ \le \mathbb{P}(S_{\lfloor m/2 \rfloor}' > k/2 - 1) \end{cases}.$$

By Lemma A.4, replacing k with $\lfloor bm^2 \rfloor$ yields the desired identity. $\quad\square$

Proof of Lemma A.3. We first show that S_n/n^2 converges in distribution. Consider the characteristic function of S_n/n^2, $\phi_n(t) = \mathbb{E}(e^{itS_n/n^2})$, where $i = \sqrt{-1}$. Clearly, $\phi_n(t) = \prod_{m=1}^{n}(1 - itn^{-2}\lambda_{n,m}^{-1})^{-1}$. Write $\phi_n = g_n h_n$, where

$$g_n(t) = \prod_{1 \le m \le \sqrt{n}} \frac{1}{1 - itn^{-2}\lambda_{n,m}^{-1}}, \qquad h_n(t) = \prod_{\sqrt{n} < m \le n} \frac{1}{1 - itn^{-2}\lambda_{n,m}^{-1}}.$$

Let $\theta_m = (2m - 1)^2 \pi^2/8$ and set

$$\psi_n(t) = \prod_{1 \le m \le \sqrt{n}} \frac{1}{1 - it\theta_m^{-1}}.$$

For $w_1, \ldots, w_k, z_1, \ldots, z_k \in \mathbb{C}$, it is easy to see that

$$\left| \prod_{i=1}^{k} w_i - \prod_{i=1}^{k} z_i \right| \le \sum_{i=1}^{k} |w_i - z_i| \left(\prod_{j=1}^{i-1} |z_j| \right) \left(\prod_{j=i+1}^{k} |w_j| \right).$$

Using the above inequality, we obtain

$$|h_n(t) - 1| \le t \sum_{\sqrt{n} < m \le n} \frac{1}{n^2 \lambda_{n,m}} \le 4t \sum_{\sqrt{n} < m \le n} \theta_m^{-1}$$

and

$$|g_n(t) - \psi_n(t)| \le t \sum_{1 \le m \le \sqrt{n}} |\theta_m^{-1} - n^{-2}\lambda_{n,m}^{-1}| \le t \sum_{1 \le m \le \sqrt{n}} \frac{1}{6n^2 - \theta_m},$$

where the last inequalities use the following fact

$$\max\{s^2/2 - s^4/24, s^2/8\} \leq 1 - \cos s \leq s^2/2, \quad \forall 0 \leq s \leq \pi. \tag{A.5}$$

Note that $\theta_m \leq 6n$ for $1 \leq m \leq \sqrt{n}$. This leads to

$$|g_n(t) - \psi_n(t)| \leq \frac{t}{6\sqrt{n}(n-1)} \leq tn^{-3/2}, \quad \forall n > 1.$$

Putting all above together yields

$$
\begin{aligned}
|\phi_n(t) - \psi_n(t)| &\leq |g_n(t)[h_n(t) - 1]| + |g_n(t) - \psi_n(t)| \\
&\leq |h_n(t) - 1| + |g_n(t) - \psi_n(t)| \to 0,
\end{aligned}
$$

as $n \to \infty$, for all $t \in \mathbb{R}$. It is clear that ψ_n converges pointwise to $\psi(t) = \prod_{m=1}^{\infty}(1 - it\theta_m^{-1})^{-1}$ and ψ is exactly a sum of independent exponential random variables with parameters $(\theta_m)_{m=1}^{\infty}$, say X. Thus, S_n/n^2 converges in distribution to X.

To see the distribution of X, let $(X_m)_{m \geq 1}$ be independent exponential random variables with parameters $(\theta_m)_{m \geq 1}$ and $Y_m = X_1 + \cdots + X_m$. Obviously, Y_m converges in distribution to X. Note that, for $t > 0$,

$$\mathbb{P}(Y_m > t) = \sum_{k=1}^{m} c_{m,k} e^{-\theta_k t},$$

where

$$c_{m,k} = \prod_{j=1, j \neq k}^{m} \frac{\theta_j}{\theta_j - \theta_k} = \frac{4(-1)^{k-1}}{2k-1} \left[2^{-2m} \binom{2m}{m} \right]^2 \frac{(m!)^2}{(m-k)!(m+k-1)!}.$$

It is easy too see that $|c_{m,k}| > |c_{m,k+1}|$ for $1 \leq k < m$ and $\theta_k < \theta_{k+1}$ for all k. This implies that $|c_{m,k}|e^{-\theta_k t}$ is decreasing in k. As $c_{m,k}e^{-\theta_k t}$ is an alternating sequence, one has

$$\left| \mathbb{P}(Y_m > t) - \sum_{k=1}^{j} c_{m,k} e^{-\theta_k t} \right| \leq |c_{m,j}|, \quad \forall j \leq m.$$

By Stirling's formula,

$$\lim_{m \to \infty} c_{m,k} = (-1)^{k-1} \frac{4}{(2k-1)\pi}, \quad \forall k \geq 1.$$

Letting $m \to \infty$ and then $j \to \infty$ yields

$$\mathbb{P}(X > t) = \frac{4}{\pi} \sum_{k=0}^{\infty} \frac{(-1)^k}{2k+1} e^{-t(2k+1)^2\pi^2/8}, \quad \forall t \geq 0.$$

The desired density function of X is then obtained by the uniform convergence of the right side and its derivative. $\square$

Proof of Lemma A.4. Set $\phi_n(t) = \mathbb{E}(e^{it S_n/n^2})$ and $\psi_n(t) = \prod_{m=1}^{n}(1 - itn^{-2}\lambda_{n,m}^{-1})^{-1}$ with $i = \sqrt{-1}$. As a consequence of the proof for Lemma A.3, we have

$$\lim_{n \to \infty} \psi_n(t) = \prod_{m=1}^{\infty} \frac{1}{1 - i(2t)\theta_m^{-1}}, \quad \forall t \in \mathbb{R},$$

where $\theta_m = (2m - 1)^2 \pi^2 / 8$. Thus, it remains to show that $\phi_n(t)/\psi_n(t) \to 1$ as $n \to \infty$ for all $t \in \mathbb{R}$. Note that

$$\phi_n(t) = \prod_{m=1}^{n} \frac{1}{1 - \lambda_{n,m}^{-1}(1 - e^{-it/n^2})}.$$

Fix $t \in \mathbb{R}$ and write $(1 - e^{-it/n^2})n^2 = a_n + ib_n$ with $a_n, b_n \in \mathbb{R}$. By (A.5), one has, for $n \geq \max\{|2t|, 1\}$,

$$0 \leq a_n \leq \frac{t^2}{2n^2} \leq \frac{1}{8}, \qquad |b_n| = |\sin(t/n^2)|n^2 \leq |t|,$$

and

$$n^2 \lambda_{n,m} \geq \frac{n^2}{2} \times \frac{[(2m - 1)\pi/2n]^2}{8} \geq \frac{(2m - 1)^2}{8}.$$

This implies

$$|1 - \lambda_{n,m}^{-1}(1 - e^{-it/n^2})| \leq |1 - n^{-2}\lambda_{n,m}^{-1}it|, \quad \forall 1 \leq m \leq n,$$

which leads to

$$|\psi_n(t)/\phi_n(t) - 1| \leq \sum_{m=1}^{n} \left| \frac{n^{-2}\lambda_{n,m}^{-1}[n^2(e^{-it/n^2} - 1) + it]}{1 - n^{-2}\lambda_{n,m}^{-1}it} \prod_{k=1}^{m-1} \frac{1 - \lambda_{n,m}^{-1}(1 - e^{-it/n^2})}{1 - n^{-2}\lambda_{n,m}^{-1}it} \right|$$

$$\leq 8|n^2(e^{-it/n^2} - 1) + it| \sum_{m=1}^{\infty} (2m - 1)^{-2} \to 0,$$

when $n \to \infty$, as desired. $\quad\square$

Appendix B. Techniques and proofs

Lemma B.1 (*[9, Theorem 1]*). *The matrix defined in* (1.2) *has eigenvalues*

$$\beta_{n,i} = 1 - \frac{2i}{n} \quad 0 \leq i \leq n,$$

with $L^2(\pi_n)$-normalized eigenvectors

$$\psi_{n,i}(x) = \binom{n}{i}^{-1/2} \sum_{k=0}^{i} (-1)^k \binom{x}{k} \binom{n - x}{i - k}, \quad 0 \leq i, x \leq n. \tag{B.1}$$

Proof of Lemma 2.1. (2) $\Leftrightarrow$ (3) is obvious from the definition of the L^p-mixing time. By the monotonicity of the L^p-distance, the converse statements for (1) and (2) are exactly

(1)′ $\mathcal{F}$ has an L^p-precutoff.
(2)′ There is $C > 0$ such that

$$\lim_{n \to \infty} D_{n,p}(\mu_n, \lfloor Ca_n \rfloor) = 0.$$

We prove the equivalence of (1) and (2) by showing $(1)' \Leftrightarrow (2)'$ instead. First, assume that $\mathcal{F}$ has an L^p-precutoff and, according to Remark 2.1, let $t_n > 0$ and $0 < A < B$ be constants such that

$$\liminf_{n \to \infty} D_{n,p}(\mu_n, \lfloor At_n \rfloor) = \epsilon_0 > 0, \qquad \lim_{n \to \infty} D_{n,p}(\mu_n, \lfloor Bt_n \rfloor) = 0.$$

Let $\delta < \min\{\epsilon, \epsilon_0\}$ and choose $N > 0, C_1 > 0$ such that

$$D_{n,p}(\mu_n, \lfloor At_n \rfloor) > \delta > D_{n,p}(\mu_n, \lfloor Bt_n \rfloor), \qquad T_{n,p}(\mu_n, \delta) \leq C_1 a_n, \quad \forall n \geq N.$$

The former implies $At_n \leq T_{n,p}(\mu_n, \delta) \leq Bt_n$ and, then,

$$Bt_n \leq \frac{BT_{n,p}(\mu_n, \delta)}{A} \leq \frac{BC_1}{A} a_n.$$

This yields

$$\limsup_{n \to \infty} D_{n,p}(\mu_n, \lfloor BC_1 a_n / A \rfloor) \leq \limsup_{n \to \infty} D_{n,p}(\mu_n, \lfloor Bt_n \rfloor) = 0.$$

Second, assume $(2)'$ and choose $C_2 > 0$ such that $T_{n,p}(\mu_n, \epsilon) \geq C_2 a_n$ and $a_n \geq 2/C_2$. Then, for $n \geq 1$,

$$D_{n,p}(\mu_n, \lfloor C_2 a_n/2 \rfloor) \geq D_{n,p}(\mu_n, \lfloor C_2 a_n - 1 \rfloor) \geq D_{n,p}(\mu_n, T_{n,p}(\mu_n, \epsilon) - 1) > \epsilon > 0.$$

This proves the L^p-precutoff. $\square$

We consider Proposition 3.2 in a more general setting.

Lemma B.2. *Let K be the transition matrix of a periodic birth-and-death chain on $\Omega = \{0, 1, \ldots, m\}$ with birth rate p_i and death rate $q_i = 1 - p_i$. That is,*

$$K(i, i + 1) = p_i, \qquad K(i, i - 1) = q_i = 1 - p_i, \quad \forall 0 \leq i \leq m,$$

with the convention $p_m = q_0 = 0$. Let $l = \lfloor m/2 \rfloor$ and μ be a probability on Ω. Suppose that, for any $i \geq 0$,

$$\mu(l - 2i) \geq \mu(l + 2i + 2) \geq \mu(l - 2i - 2), \quad p_{l+2i} \geq q_{l-2i} \geq p_{l+2i+2}, \qquad (B.2)$$

and

$$p_{l+2i} + q_{l+2i+2} \geq p_{l-2i-2} + q_{l-2i} \geq p_{l+2i+2} + q_{l+2i+4}. \qquad (B.3)$$

Then, for all $i \geq 0$,

$$\mu K(l + 2i + 1) \geq \mu K(l - 2i - 1) \geq \mu K(l + 2i + 3).$$

Proof. By the periodicity of K,

$$\mu K(j) = \mu(j - 1)p_{j-1} + \mu(j + 1)q_{j+1}, \quad \forall 0 \leq j \leq m,$$

where

$$\mu(-1) = \mu(m + 1) = p_{-1} = q_{m+1} = 0. \qquad (B.4)$$

It is easy to check that both (B.2) and (B.3) hold under the extension in (B.4). If $i \leq (l - 1)/2$, then $l + 2i + 1 \leq 2l \leq m$ and

$$\mu K(l + 2i + 1) - \mu K(l - 2i - 1)$$
$$= [\mu(l + 2i)p_{l+2i} + \mu(l + 2i + 2)q_{l+2i+2}]$$

$$- [\mu(l - 2i)q_{l-2i} + \mu(l - 2i - 2)p_{l-2i-2}]$$
$$\geq \mu(l - 2i)(p_{l+2i} - q_{l-2i}) + \mu(l + 2i + 2)(q_{l+2i+2} - p_{l-2i-2})$$
$$\geq \mu(l + 2i + 2)(p_{l+2i} - q_{l-2i} + q_{l+2i+2} - p_{l-2i-2}) \geq 0.$$

If $l + 2i + 3 \leq m$, then $l - 2i - 1 \geq 2l + 2 - m \geq 1$ and

$$\mu K(l - 2i - 1) - \mu K(l + 2i + 3)$$
$$= [\mu(l - 2i)q_{l-2i} + \mu(l - 2i - 2)p_{i-2i-2}]$$
$$\quad - [\mu(l + 2i + 2)p_{l+2i+2} + \mu(l + 2i + 4)q_{l+2i+4}]$$
$$\geq \mu(l + 2i + 2)(q_{l-2i} - p_{l+2i+2}) + \mu(l - 2i - 2)(p_{i-2i-2} - q_{l+2i+4})$$
$$\geq \mu(l - 2i - 2)(q_{l-2i} - p_{l+2i+2} + p_{i-2i-2} - q_{l+2i+4}) \geq 0.$$

This finishes the proof. □

Remark B.1. Lemma B.2 also holds for the case that m is even and $l = m/2 - 1$. The proof goes similarly and is omitted.

The following is a simple corollary of Lemma B.2.

Corollary B.3. *Let K be the transition matrix on $\Omega = \{0, 1, \ldots, m\}$ given by*

$$K(i, i + 1) = p_i, \qquad K(i, i - 1) = q_i = 1 - p_i, \quad \forall 0 \leq i \leq m,$$

where $p_m = q_0 = 0$, and let μ be a probability on Ω. Suppose that

$$p_i = q_{m-i}, \qquad p_i \geq p_{i+1}, \quad \forall i \leq m/2,$$

and

$$p_i + q_{i+2} \leq p_{i+1} + q_{i+3}, \quad \forall 0 \leq i \leq \lfloor m/2 \rfloor - 2.$$

(1) *If $m = 2l$ and*

$$\mu(l + 2i) \geq \mu(l - 2i - 2) \geq \mu(l + 2i + 2), \quad \forall i \geq 0,$$

then, for all $i \geq 0$ and $t \in \{0, 1, 2, \ldots\}$,

$$\mu K^{2t+1}(l - 2i - 1) \geq \mu K^{2t+1}(l + 2i + 1) \geq \mu K^{2t+1}(l - 2i - 3)$$

and

$$\mu K^{2t}(l + 2i) \geq \mu K^{2t}(l - 2i - 2) \geq \mu K^{2t}(l + 2i + 2).$$

(2) *If $m = 2l$ and*

$$\mu(l - 2i - 1) \geq \mu(l - 2i + 1) \geq \mu(l - 2i - 3), \quad \forall i \geq 0,$$

then, for all $i \geq 0$ and $t \in \{0, 1, 2, \ldots\}$,

$$\mu K^{2t+1}(l + 2i) \geq \mu K^{2t+1}(l - 2i - 2) \geq \mu K^{2t+1}(l + 2i + 2).$$

and

$$\mu K^{2t}(l - 2i - 1) \geq \mu K^{2t}(l + 2i + 1) \geq \mu K^{2t}(l - 2i - 3).$$

(3) If $m = 2l + 1$ and

$$\mu(l - 2i) \geq \mu(l + 2i + 2) \geq \mu(l - 2i - 2), \quad \forall i \geq 0,$$

then, for all $i \geq 0$ and $t \in \{0, 1, 2, \ldots\}$,

$$\mu K^{2t+1}(l + 2i + 1) \geq \mu K^{2t+1}(l - 2i - 1) \geq \mu K^{2t+1}(l + 2i + 3)$$

and

$$\mu K^{2t}(l - 2i) \geq \mu K^{2t}(l + 2i + 2) \geq \mu K^{2t}(l - 2i - 2).$$

Proof of Proposition 3.2. For the birth-and-death chain in Proposition 3.2, it is obvious that $p_i = 1 - i/n$ and $q_i = i/n$. This implies

$$p_i = q_{n-i}, \quad p_i > p_{i+1}, \qquad p_i + q_{i+2} = 1 + \frac{2}{n}, \quad \forall i \geq 0.$$

Applying Corollary B.3 with $K = K_n$ and $\mu = \delta_{\lceil n/2 \rceil}$, the Dirac mass on $\lceil n/2 \rceil$, yields

$$K_n^t(\lceil n/2 \rceil, A) \geq 1/2, \quad \forall t \geq 0.$$

For the general case with $\mu_n(A) \geq 1/2$, let $(X_t)_{t=0}^{\infty}$ be a Markov chain with transition matrix K_n and let T be the first passage time to state $\lceil n/2 \rceil$, i.e., $T = \min\{t \geq 0 : X_t = \lceil n/2 \rceil\}$. By the irreducibility of K_n, $\mathbb{P}_{\mu_n}(T < \infty) = 1$. Using the strong Markov property, we obtain, for $t \geq 0$,

$$\mu_n K_n^t(A) = \sum_{i=0}^{t} \mathbb{P}_{\mu_n}(X_t \in A, \ T = i) + \mathbb{P}_{\mu_n}(X_t \in A, \ T > t)$$

$$= \sum_{i=0}^{t} \mathbb{P}(X_{t-i} \in A | X_0 = \lceil n/2 \rceil) \mathbb{P}_{\mu_n}(T = i) + \mathbb{P}_{\mu_n}(T > t)$$

$$\geq \frac{1}{2} \mathbb{P}_{\mu_n}(T \leq t) + \mathbb{P}_{\mu_n}(T > t) \geq 1/2. \quad \square$$

Lemma B.4 ([7, Lemma A.1]). *For $n > 0$, let $a_n \in \mathbb{R}^+$, $b_n \in \mathbb{Z}^+$, $c_n = \frac{b_n - a_n}{\sqrt{a_n}}$ and $d_n = e^{-a_n} \sum_{i=0}^{b_n} \frac{a_n^i}{i!}$. Assume that $a_n + b_n \to \infty$. Then*

$$\limsup_{n \to \infty} d_n = \Phi\left(\limsup_{n \to \infty} c_n\right), \qquad \liminf_{n \to \infty} d_n = \Phi\left(\liminf_{n \to \infty} c_n\right), \tag{B.5}$$

where $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-t^2/2} dt$.

In particular, if c_n converges(the limit can be $+\infty$ and $-\infty$), then $\lim_{n \to \infty} d_n = \Phi(\lim_{n \to \infty} c_n)$.

Lemma B.5. *For $n \geq 1$, let ξ_n be a binomial random variable with parameters $(n, 1/2)$. Then, there is a universal constant $C > 0$ such that*

$$E\left(\left|\frac{n - 2\xi_n}{\sqrt{n}}\right|^{\theta}\right) \leq C 4^{\theta} \Gamma\left(\frac{\theta + 1}{2}\right), \quad \forall \theta > 0, \ n \geq 1,$$

where Γ is the Gamma function.

Proof. Set $\Omega_n = \{0, 1, \ldots, n\}$ and $\pi_n(x) = \binom{n}{x} 2^{-n}$. According to the definition of ξ_n, $\mathbb{P}(\xi_n = x) = \pi_n(x)$ for $x \in \Omega_n$. For $0 \le j < \sqrt{n}$, set

$$E_{n,j} = \{x \in \Omega_n : |n - 2x|/\sqrt{n} \in (j, j+1]\}, \qquad y_{n,j} = \max\{x \in E_{n,j} : x \le n/2\}.$$

Clearly, $\left[n - (j+1)\sqrt{n}\right]/2 \le y_{n,j} < \left(n - j\sqrt{n}\right)/2$ and

$$E\left(\left|\frac{n - 2\xi_n}{\sqrt{n}}\right|^\theta\right) \le \sum_{j=0}^{\lfloor \sqrt{n} \rfloor} (j+1)^\theta \pi_n(E_{n,j}). \tag{B.6}$$

Using (4.2), we obtain, for $y_{n,j} \ne 0$,

$$\pi_n(E_{n,j}) = 2^{-n} \sum_{x \in E_{n,j}} \frac{n!}{x!(n-x)!} \le 2^{1-n} \lceil \sqrt{n} \rceil \frac{n!}{y_{n,j}!(n - y_{n,j})!}$$

$$\le 2^{2-n} \sqrt{n} \beta^3 \frac{n^{n+1/2}}{y_{n,j}^{y_{n,j}+1/2}(n - y_{n,j})^{n - y_{n,j}+1/2}} = 8\beta^3/z_{n,j},$$

where

$$z_{n,j} = \left[1 - \left(1 - \frac{2y_{n,j}}{n}\right)^2\right]^{(n+1)/2} \left[\frac{1 + (1 - 2y_{n,j}/n)}{1 - (1 - 2y_{n,j}/n)}\right]^{n/2 - y_{n,j}}.$$

Since $t \mapsto (1-t)^{1/t}$ is strictly decreasing on $(0, 1)$, it is easy to see that

$$1 - \left(1 - \frac{2y_{n,j}}{n}\right)^2 \ge \left[1 - \left(1 - \frac{2y_{n,j}}{n}\right)\right]^{1 - 2y_{n,j}/n}.$$

As $y_{n,j} \le n/2$, this implies

$$z_{n,j} \ge \frac{2y_{n,j}}{n} \left[1 + \left(1 - \frac{2y_{n,j}}{n}\right)\right]^{n/2 - y_{n,j}}.$$

In the case $y_{n,j} \ge n/6$, one may use the inequality, $\log(1 + t) \ge t/2$ for $t \in [0, 1]$, to get

$$z_{n,j} \ge \frac{1}{3} \exp\left\{\frac{n}{4}\left(1 - \frac{2y_{n,j}}{n}\right)^2\right\} \ge \frac{1}{3} e^{j^2/4}.$$

In the case $1 \le y_{n,j} \le n/6$, it is clear that

$$z_{n,j} \ge \frac{2}{n}\left(\frac{5}{3}\right)^{n/3} \ge \frac{2}{n} e^{n/6} \ge \frac{2}{n} e^{n/24} e^{j^2/8},$$

where the last inequality applies the fact $j < \sqrt{n}$. Putting both cases together, we may choose a universal constant $C > 1$ such that

$$z_{n,j} \ge \frac{e^{j^2/8}}{C}, \quad \forall 0 \le j \le \sqrt{n}, \qquad y_{n,j} \ne 0, \quad n \ge 1.$$

Back to the computation of $\pi_n(E_{n,j})$, this gives

$$\pi_n(E_{n,j}) \le 8C\beta^3 e^{-j^2/8}, \quad \forall 0 \le j \le \sqrt{n}, \qquad y_{n,j} \ne 0, \quad n \ge 1.$$

In fact, the above inequality also holds for $y_{n,j} = 0$ (which must imply $j = \lfloor \sqrt{n} \rfloor$) because, in such a case, $\pi_n(E_{n,j}) = 2^{1-n} \leq 2e^{-(\log 2)j^2} \leq 2e^{-j^2/8}$. Back to (B.6), we achieve

$$E\left(\left|\frac{n - 2\xi_n}{\sqrt{n}}\right|^{\theta}\right) \leq 32C\beta^3 4^{\theta} \Gamma\left(\frac{\theta + 1}{2}\right). \quad \square$$

References

[1] D. Aldous, Random walks on finite groups and rapidly mixing Markov chains, in: Seminar on Probability, XVII, in: Lecture Notes in Math., vol. 986, Springer, Berlin, 1983, pp. 243–297. MR 770418.

[2] D. Aldous, P. Diaconis, Shuffling cards and stopping times, Amer. Math. Monthly 93 (1986) 333–348. MR 0841111.

[3] D. Aldous, P. Diaconis, Strong uniform times and finite random walks, Acta Appl. Math. 8 (1987) 69–97. MR 0876954.

[4] J. Barrera, O. Bertoncini, R. Fernndez, Abrupt convergence and escape behavior for birth and death chains, J. Stat. Phys. 137 (4) (2009) 595–623. MR 2565098.

[5] O. Bertoncini, Convergence abrupte et métastabilité. Ph.D. Thesis, Université de Rouen, 2007.

[6] M. Brown, Y.-S. Shao, Identifying coefficients in the spectral representation for first passage time distributions, Probab. Engrg. Inform. Sci. 1 (1987) 69–74.

[7] G.-Y. Chen, L. Saloff-Coste, The cutoff phenomenon for ergodic markov processes, Electron. J. Probab. 13 (2008) 26–78. MR 2375599.

[8] G.-Y. Chen, L. Saloff-Coste, The L^2-cutoff for reversible Markov processes, J. Funct. Anal. 258 (2010) 2246–2315. MR 2584746.

[9] P. Diaconis, P. Hanlon, Eigen analysis for some examples of the Metropolis algorithm, in: Hypergeometric Functions on Fomains of Positivity, Jack Polynomials, and Applications, Vol. 138, Amer. Math. Soc, Providence, RI, 1992, pp. 99–117. MR 1199122.

[10] P. Diaconis, L. Saloff-Coste, Separation cut-offs for birth and death chains, Ann. Appl. Probab. 16 (2006) 2098–2122. MR 2288715.

[11] W. Feller, An Introduction to Probability Theory and Its Applications. I, third ed., John Wiley & Sons Inc., New York, 1968, MR 0228020.

[12] B. Lachaud, Cut-off and hitting times of a sample of Ornstein–Uhlenbeck processes and its average, J. Appl. Probab. 42 (4) (2005) 1069–1080. MR 2203823.

[13] D. Levin, Y. Peres, E. Wilmer, Markov Chains and Mixing Times, American Mathematical Society, Providence, RI, 2009, MR 2466937.

[14] S. Martnez, B. Ycart, Decay rates and cutoff for convergence and hitting times of Markov chains with countably infinite state space, Adv. Appl. Probab. 33 (1) (2001) 188–205. MR 1825322.

[15] L. Saloff-Coste, Lectures on finite Markov chains, in: Lectures on Probability Theory and Statistics, in: Lecture Notes in Math., vol. 1665, Springer, Berlin, 1997, pp. 301–413. MR 1490046.

[16] L. Saloff-Coste, Random walks on finite groups, in: Probability on Discrete Structures, in: Encyclopaedia Math. Sci., vol. 110, Springer, Berlin, 2004, pp. 263–346. MR 2023654.