ALEA, Lat. Am. J. Probab. Math. Stat. 10 (1), 293–321 (2013)

On the mixing time and spectral gap for birth and death chains

Guan-Yu Chen[1] **and Laurent Saloff-Coste**[2]

Department of Applied Mathematics, National Chiao Tung University
Hsinchu 300, Taiwan
E-mail address: gychen@math.nctu.edu.tw
URL: http://jupiter.math.nctu.edu.tw/~gychen/index2.htm

Malott Hall, Department of Mathematics, Cornell University
Ithaca, NY 14853-4201
E-mail address: lsc@math.cornell.edu
URL: http://www.math.cornell.edu/~lsc/lau.html

Abstract. For birth and death chains, we derive bounds on the spectral gap and mixing time in terms of birth and death rates. Together with the results of Ding et al. (2010), this provides a criterion for the existence of a cutoff in terms of the birth and death rates. A variety of illustrative examples are treated.

1. Introduction

Let Ω be a countable set and (Ω, K, π) be an irreducible Markov chain on Ω with transition matrix K and stationary distribution π. Let I be the identity matrix indexed by Ω and

$$H_t = e^{-t(I-K)} = \sum_{i=0}^{\infty} e^{-t} t^i K^i / i!$$

be the associated semigroup which describes the corresponding natural continuous time process on Ω. For $\delta \in (0, 1)$, set

$$K_\delta = \delta I + (1 - \delta)K. \tag{1.1}$$

Clearly, K_δ is similar to K but with an additional holding probability depending of δ. We call K_δ the δ-lazy walk or δ-lazy chain of K. It is well-known that if K is irreducible with stationary distribution π, then

$$\lim_{m\to\infty} K_\delta^m(x, y) = \lim_{t\to\infty} H_t(x, y) = \pi(y), \quad \forall x, y \in \Omega,\ \delta \in (0, 1).$$

Received by the editors October 22, 2012; accepted April 9, 2013.
2010 *Mathematics Subject Classification.* 60J10,60J27.
Key words and phrases. Birth and death chains, Cutoff phenomenon.
[1]Partially supported by NSC grant NSC100-2115-M-009-003-MY2.
[2]Partially supported by NSF grant DMS-1004771.

In this paper, we consider convergence in total variation. The total variation between two probabilities μ, ν on Ω is defined by $\|\mu-\nu\|_{\mathrm{TV}} = \sup\{\mu(A)-\nu(A)|A \subset \Omega\}$. For any irreducible K with stationary distribution π, the (maximum) total variation distance is defined by

$$d_{\mathrm{TV}}(m) = \sup_{x\in\Omega} \|K^m(x,\cdot) - \pi\|_{\mathrm{TV}}, \tag{1.2}$$

and the corresponding mixing time is given by

$$T_{\mathrm{TV}}(\epsilon) = \inf\{m \geq 0 | d_{\mathrm{TV}}(m) \leq \epsilon\}, \quad \forall \epsilon \in (0,1). \tag{1.3}$$

We write $d^{(c)}_{\mathrm{TV}}, T^{(c)}_{\mathrm{TV}}$ for the total variation distance and mixing time for the continuous semigroup and $d^{(\delta)}_{\mathrm{TV}}, T^{(\delta)}_{\mathrm{TV}}$ for the δ-lazy walk.

A birth and death chain on $\{0,1,...,n\}$ with birth rate p_i, death rate q_i and holding rate r_i is a Markov chain with transition matrix K given by

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

where $p_i + q_i + r_i = 1$ and $p_n = q_0 = 0$. It is obvious that K is irreducible if and only if $p_i q_{i+1} > 0$ for $0 \leq i < n$. Under the assumption of irreducibility, the unique stationary distribution π of K is given by $\pi(i) = c(p_0 \cdots p_{i-1})/(q_1 \cdots q_i)$, where c is a positive constant such that $\sum_{i=0}^n \pi(i) = 1$. The following theorem provides a bound on the mixing time using the birth and death rates and is treated in Theorems 3.1 and 3.9.

Theorem 1.1. *Let K be an irreducible birth and death chain on $\{0,1,...,n\}$ with birth, death and holding rates p_i, q_i, r_i. Let i_0 be a state satisfying $\pi([0,i_0]) \geq 1/2$ and $\pi([i_0,n]) \geq 1/2$, where $\pi(A) = \sum_{i\in A} \pi(i)$, and set*

$$t = \max\left\{\sum_{k=0}^{i_0-1} \frac{\pi([0,k])}{\pi(k)p_k}, \sum_{k=i_0+1}^{n} \frac{\pi([k,n])}{\pi(k)q_k}\right\}.$$

Then, for any $\delta \in [1/2,1)$,

$$\min\left\{T^{(c)}_{\mathrm{TV}}(1/10), T^{(\delta)}_{\mathrm{TV}}(1/20)\right\} \geq \frac{t}{6},$$

and

$$\max\left\{T^{(c)}_{\mathrm{TV}}(\epsilon), T^{(\delta)}_{\mathrm{TV}}(\epsilon)\right\} \leq \frac{18t}{\epsilon^2}, \quad \forall \epsilon \in (0,1).$$

Ding et al. (2010) derive a similar upper bound. Note that if $(X_m)_{m=0}^\infty$ is a Markov chain on Ω_n with transition matrix K and $\tau_i := \min\{m \geq 0 | X_m = i\}$, then $t = \max\{\mathbb{E}_0\tau_{i_0}, \mathbb{E}_n\tau_{i_0}\}$, where $\mathbb{E}_i$ denotes the conditional expectation given $X_0 = i$. See Lemma 3.3 for details.

A sharp transition phenomenon, known as cutoff, was observed by Aldous and Diaconis in early 1980s. See e.g. Diaconis (1996); Chen and Saloff-Coste (2008) for an introduction and a general review of cutoffs. In total variation, a family of irreducible Markov chains $(\Omega_n, K_n, \pi_n)_{n=1}^\infty$ is said to present a cutoff if

$$\lim_{n\to\infty} \frac{T_{n,\mathrm{TV}}(\epsilon)}{T_{n,\mathrm{TV}}(\eta)} = 1, \quad \forall 0 < \epsilon < \eta < 1. \tag{1.4}$$

The family is said to present a (t_n, b_n) cutoff if $b_n = o(t_n)$ and

$$|T_{n,\mathrm{TV}}(\epsilon) - t_n| = O(b_n), \quad \forall 0 < \epsilon < 1.$$

The cutoff for the associated continuous semigroups is defined in a similar way. Given a family $\mathcal{F}$ of irreducible Markov chains, we write $\mathcal{F}_c$ and $\mathcal{F}_\delta$ for the families of corresponding continuous time chain and δ-lazy discrete time chains.

Let $\mathcal{F} = \{(\Omega_n, K_n, \pi_n)|n = 1, 2, ...\}$ be a family of birth and death chains, where $\Omega_n = \{0, 1, ..., n\}$ and K_n has birth rate $p_{n,i}$, death rate $q_{n,i}$ and holding rate $r_{n,i}$. Suppose that K_n is irreducible with stationary distribution π_n. For the family $\{(\Omega_n, K_n, \pi_n)|n = 1, 2, ...\}$, Ding et al. (2010) showed that, in the discrete time case and assuming $\inf_{i,n} r_{n,i} > 0$, the cutoff in total variation exists if and only if the product of the total variation mixing time and the spectral gap, i.e. the smallest non-zero eigenvalue of $I - K$, tends to infinity. There is also a similar version for the continuous time case. In Chen and Saloff-Coste (2012a), we use the results of Diaconis and Saloff-Coste (2006); Ding et al. (2010) to provide another criterion on the cutoff using the eigenvalues of K_n. In both cases, the spectral gap is needed to determine if there is a cutoff. The following theorem provides a bound on the spectral gap using the birth and death rates.

Theorem 1.2. *Consider an irreducible birth and death chain K on $\{0, 1, ..., n\}$ with birth, death and holding rates, p_i, q_i, r_i. Let π and λ be the stationary distribution and spectral gap of K and set*

$$\ell = \max\left\{\max_{j:j<i_0} \sum_{k=j}^{i_0-1} \frac{\pi([0,j])}{\pi(k)p_k}, \max_{j:j>i_0} \sum_{k=i_0+1}^{j} \frac{\pi([j,n])}{\pi(k)q_k}\right\},$$

where i_0 is a state such that $\pi([0, i_0]) \geq 1/2$ and $\pi([i_0, n]) \geq 1/2$. Then,

$$\frac{1}{4\ell} \leq \lambda \leq \frac{2}{\ell}.$$

The above theorem is motivated by Miclo (1999), where the author considers the spectral gap of birth and death chains on $\mathbb{Z}$. We refer the reader to Miclo (1999) and the references therein for more information. Note that if t, ℓ are the constants in Theorem 1.1-1.2, then $t \geq \ell$. Based on the results in Ding et al. (2010), we obtain a theorem regarding cutoffs for birth and death chains.

Theorem 1.3. *Consider a family of irreducible birth and death chains*

$$\mathcal{F} = \{(\Omega_n, K_n, \pi_n)|n = 1, 2, ...\},$$

where $\Omega_n = \{0, 1, ..., n\}$ and K_n has birth, death and holding rates, $p_{n,i}, q_{n,i}, r_{n,i}$. For $n \geq 1$, let $i_n \in \{0, ..., n\}$ be a state satisfying $\pi_n([0, i_n]) \geq 1/2$ and $\pi_n([i_n + 1, n]) \geq 1/2$ and set

$$t_n = \max\left\{\sum_{k=0}^{i_n-1} \frac{\pi_n([0,k])}{\pi_n(k)p_{n,k}}, \sum_{k=i_n+1}^{n} \frac{\pi_n([k,n])}{\pi_n(k)q_{n,k}}\right\}.$$

and

$$\ell_n = \max\left\{\max_{j:j<i_n} \sum_{k=j}^{i_n-1} \frac{\pi_n([0,j])}{\pi_n(k)p_{n,k}}, \max_{j:j>i_n} \sum_{k=i_n+1}^{j} \frac{\pi_n([j,n])}{\pi_n(k)q_{n,k}}\right\},$$

Then, for any $\epsilon \in (0, 1/2)$ and $\delta \in (0, 1)$, there is a constant $C = C(\epsilon, \delta) > 1$ such that

$$C^{-1}t_n \leq \min\{T^{(c)}_{n,\mathrm{TV}}(\epsilon), T^{(\delta)}_{n,\mathrm{TV}}(\epsilon)\} \leq \max\{T^{(c)}_{n,\mathrm{TV}}(\epsilon), T^{(\delta)}_{n,\mathrm{TV}}(\epsilon)\} \leq Ct_n,$$

for n large enough. Moreover, the following are equivalent.

(1) $\mathcal{F}_c$ *has a total variation cutoff.*
(2) *For* $\delta \in (0,1)$, $\mathcal{F}_\delta$ *has a total variation cutoff.*
(3) $t_n \ell_n \to \infty$.

The above theorem is immediate from Theorems 1.1, 1.2, 2.4 and 2.5. The selection of i_n can be relaxed. See Theorem 3.10 for a precise statement. By the results in Chen and Saloff-Coste (2012a), Theorem 1.3 also holds when t_n is replaced by the following constant

$$s_n = \frac{1}{\lambda_{n,1}} + \cdots + \frac{1}{\lambda_{n,n}},$$

where $\lambda_{n,1},...,\lambda_{n,n}$ are nonzero eigenvalues of $I - K_n$. Furthermore, Theorem 1.3 also holds in separation with $\delta \in [1/2, 1)$. We will use Theorem 1.3 to study the cutoff of several examples including the following theorem which concerns random walks with bottlenecks. It is a special case of Theorem 4.12.

Theorem 1.4. *For* $n \geq 1$, *let* $\Omega_n = \{0,1,...,n\}$, $\pi_n \equiv 1/(n+1)$ *and* K_n *be an irreducible birth and death chain on* Ω_n *satisfying*

$$K_n(i-1,i) = K_n(i,i-1) = \begin{cases} 1/2 & \text{for } i \notin \{x_{n,1},...,x_{n,k_n}\} \\ \epsilon_n & \text{for } i = x_{n,j},\ 1 \leq j \leq k_n \end{cases},$$

where $0 \leq k_n \leq n$, $\epsilon_n \in (0,1/2]$, $x_{n,1},...,x_{n,k_n} \in \Omega_n$ *are distinct and the holding rate at* i *is adjusted accordingly. Set* $t_n = n^2 + a_n/\epsilon_n$, *where*

$$a_n = \sum_{i=1}^{k_n} \min\{x_{n,i}, n+1-x_{n,i}\},$$

and set

$$b_n = \max_{j:j\leq n/2} \{(j+1) \times |\{1 \leq i \leq k_n : j < x_{n,i} \leq n-j\}|\}.$$

Then, for any $\epsilon \in (0,1/2)$ *and* $\delta \in (0,1)$, *there is* $C = C(\epsilon,\delta) > 1$ *such that*

$$C^{-1} t_n \leq \min\{T^{(c)}_{n,\mathrm{TV}}(\epsilon), T^{(\delta)}_{n,\mathrm{TV}}(\epsilon)\} \leq \max\{T^{(c)}_{n,\mathrm{TV}}(\epsilon), T^{(\delta)}_{n,\mathrm{TV}}(\epsilon)\} \leq C t_n,$$

for n *large enough.*

Moreover, the following are equivalent.

(1) $\mathcal{F}_c$ *has a total variation cutoff.*
(2) *For* $\delta \in (0,1)$, $\mathcal{F}_\delta$ *has a total variation cutoff.*
(3) $a_n/(n^2\epsilon_n) \to \infty$ *and* $a_n/b_n \to \infty$.

The remaining of this article is organized as follows. In Section 2, the concepts of cutoffs and mixing times and fundamental results are reviewed. In Section 3, we give a proof for Theorems 1.1 and 1.2. For illustration, we consider several nontrivial examples in Section 4, where the mixing time and cutoff are determined. Note that the assumption regarding birth and death rates in Sections 3 and 4 can be relaxed using the comparison technique in Diaconis and Saloff-Coste (1993a,b).

2. Backgrounds

Throughout this paper, for any two sequences s_n, t_n of positive numbers, we write $s_n = O(t_n)$ if there are $C > 0, N > 0$ such that $|s_n| \leq C|t_n|$ for $n \geq N$. If $s_n = O(t_n)$ and $t_n = O(s_n)$, we write $s_n \asymp t_n$. If $t_n/s_n \to 1$ as $n \to \infty$, we write $t_n \sim s_n$.

2.1. *Cutoffs and mixing time.* Consider the following definitions.

Definition 2.1. Referring to the notation in (1.2), a family $\mathcal{F} = \{(\Omega_n, K_n, \pi_n)|n = 1, 2, ...\}$ is said to present a total variation

(1) precutoff if there is a sequence t_n and $B > A > 0$ such that

$$\lim_{n\to\infty} d_{n,\mathrm{TV}}(\lceil Bt_n\rceil) = 0, \quad \liminf_{n\to\infty} d_{n,\mathrm{TV}}(\lfloor At_n\rfloor) > 0.$$

(2) cutoff if there is a sequence t_n such that, for all $\epsilon > 0$,

$$\lim_{n\to\infty} d_{n,\mathrm{TV}}(\lceil(1+\epsilon)t_n\rceil) = 0, \quad \lim_{n\to\infty} d_{n,\mathrm{TV}}(\lfloor(1-\epsilon)t_n\rfloor) = 1.$$

In definition 2.1(2), t_n is called a cutoff time. The definition of a cutoff for continuous semigroups is similar with $\lceil\cdot\rceil$ and $\lfloor\cdot\rfloor$ deleted.

Remark 2.2. In Definition 2.1, if $t_n \to \infty$ (or equivalently $T_{n,\mathrm{TV}}(\epsilon) \to \infty$ for some $\epsilon \in (0, 1)$), then the cutoff is consistent with (1.4). This is also true for cutoffs in continuous semigroups without the assumption $t_n \to \infty$. See Chen (2006); Chen and Saloff-Coste (2008) for further discussions on cutoffs.

It is well-known that the mixing time can be bounded below by the reciprocal of the spectral gap up to a multiple constant. We cite the bound in Chen and Saloff-Coste (2012a) as follows.

Lemma 2.3. *Let K be an irreducible transition matrix on a finite set Ω with stationary distribution π. For $\delta \in (0, 1)$, let K_δ be the δ-lazy walk given by (1.1). Suppose (π, K) is reversible, that is, $\pi(x)K(x, y) = \pi(y)K(y, x)$ for all $x, y \in \Omega$ and let λ be the smallest non-zero eigenvalue of $I - K$. Then, for $\epsilon \in (0, 1/2)$,*

$$T^{(c)}_{\mathrm{TV}}(\epsilon) \ge \frac{-\log(2\epsilon)}{\lambda}, \quad T^{(\delta)}_{\mathrm{TV}}(\epsilon) \ge \left\lfloor \frac{-\log(2\epsilon)}{2\max\{1-\delta, \log(2/\delta)\}\lambda} \right\rfloor,$$

where the second inequality requires $|\Omega| \ge 2/\delta$.

2.2. *Cutoffs for birth and death chains.* Consider a family of irreducible birth and death chains

$$\mathcal{F} = \{(\Omega_n, K_n, \pi_n)|n = 1, 2, ...\},$$

where $\Omega_n = \{0, 1, ..., n\}$ and K_n has birth rate $p_{n,i}$, death rate $q_{n,i}$ and holding rate $r_{n,i}$. We write $\mathcal{F}_c, \mathcal{F}_\delta$ as families of the corresponding continuous time chains and δ-lazy discrete time chains in $\mathcal{F}$. A criterion on total variation cutoffs for families of birth and death chains was introduced in Ding et al. (2010), which say that, for $\delta \in (0, 1)$, $\mathcal{F}_c, \mathcal{F}_\delta$ have total variation cutoffs if and only if the product of the mixing time and the spectral gap tends to infinity. As the total variation distance is comparable with the separation distance, the authors of Ding et al. (2010) identify cutoffs in total variation and separation, where a criterion on separation cutoffs was proposed in Diaconis and Saloff-Coste (2006). In the recent work Chen and Saloff-Coste (2012a), the cutoffs for $\mathcal{F}_c$ and $\mathcal{F}_\delta$ are proved to be equivalent and this leads to the following theorems.

Theorem 2.4. *Chen and Saloff-Coste (2012a, Section 4) Let $\mathcal{F} = \{(\Omega_n, K_n, \pi_n)| n = 1, 2, ...\}$ be a family of irreducible birth and death chain with $\Omega_n = \{0, 1, ..., n\}$. For $n \ge 1$, let $\lambda_{n,1}, ..., \lambda_{n,n}$ be nonzero eigenvalues of $I - K_n$ and set*

$$\lambda_n = \min_{1\le i\le n} \lambda_{n,i}, \quad s_n = \frac{1}{\lambda_{n,1}} + \cdots + \frac{1}{\lambda_{n,n}}.$$

Then, the following are equivalent.

(1) $\mathcal{F}_c$ *has a total variation cutoff.*
(2) $\mathcal{F}_\delta$ *has a total variation cutoff.*
(3) $\mathcal{F}_c$ *has a total variation precutoff.*
(4) $\mathcal{F}_\delta$ *has a total variation precutoff.*
(5) $T^{(c)}_{n,\mathrm{TV}}(\epsilon)\lambda_n \to \infty$ *for some* $\epsilon \in (0,1)$.
(6) $T^{(\delta)}_{n,\mathrm{TV}}(\epsilon)\lambda_n \to \infty$ *for some* $\epsilon \in (0,1)$.
(7) $s_n\lambda_n \to \infty$.

Theorem 2.5. Chen and Saloff-Coste (2012a, *Section 4) Referring to Theorem* 2.4, *it holds true that, for* $\epsilon, \eta \in (0,1/2)$ *and* $\delta \in (0,1)$,

$$T^{(c)}_{n,\mathrm{TV}}(\epsilon) \asymp T^{(\delta)}_{n,\mathrm{TV}}(\eta).$$

Further, if there is $\epsilon_0 \in (0,1/2)$ *such that* $T^{(c)}_{n,\mathrm{TV}}(\epsilon_0)\lambda_n$ *or* $T^{(\delta)}_{n,\mathrm{TV}}(\epsilon_0)\lambda_n$ *is bounded, then, for any* $\epsilon \in (0,1/2)$ *and* $\delta \in (0,1)$,

$$T^{(c)}_{n,\mathrm{TV}}(\epsilon) \asymp T^{(\delta)}_{n,\mathrm{TV}}(\epsilon) \asymp \lambda_n^{-1}.$$

2.3. *A remark on the precutoff.* Note that if there is no cutoff in total variation, the approximation in Theorem 2.5 may fail for $\epsilon \in (1/2,1)$. This means that, for $0<\epsilon<1/2<\eta<1$, the orders of $T^{(c)}_{n,\mathrm{TV}}(\epsilon)$ and $T^{(c)}_{n,\mathrm{TV}}(\eta)$ can be different. Consider the following example. For $n \ge 3$, let $\Omega_n = \{0,1,...,n\}$, $M_n = \lfloor n/2 \rfloor$ and

$$\begin{cases} K_n(i,i+1) = K_n(i+1,i) = 1/2 \quad \text{for } 0 \le i < n,\ i \ne M_n \\ K_n(M_n,M_n+1) = K_n(M_n+1,M_n) = \epsilon_n \\ K_n(0,0) = K_n(n,n) = 1/2 \\ K_n(M_n,M_n) = K_n(M_n+1,M_n+1) = 1/2-\epsilon_n \end{cases}, \tag{2.1}$$

with $\epsilon_n \le 1/2$. Assume that $\epsilon_n = o(n^{-2})$. By Theorem 1.4, we have

$$T^{(c)}_{n,\mathrm{TV}}(\epsilon) \asymp T^{(\delta)}_{n,\mathrm{TV}}(\epsilon) \asymp n/\epsilon_n, \quad \forall \epsilon \in (0,1/2),\ \delta \in (0,1).$$

Next, we consider the δ-lazy discrete time case with $\delta = 1/2$. Let $K_{n,1/2} = (I+K_n)/2$ and K'_n be the 1/2-lazy simple random walk on $\{0,1,...,M_n\}$, that is,

$$\begin{cases} K'_n(i,i+1) = K'_n(i+1,i) = 1/4, \quad \forall 0 \le i < M_n \\ K'_n(i,i) = 1/2, \quad \forall 0 < i < M_n \\ K'_n(0,0) = K'_n(M_n,M_n) = 3/4 \end{cases}.$$

For $n \ge 3$, set

$$c_n = \min_{0\le i,j\le M_n} \frac{K^{m_n}_{n,1/2}(i,j)}{(K'_n)^{m_n}(i,j)}, \quad C_n = \max_{0\le i,j\le M_n} \frac{K^{m_n}_{n,1/2}(i,j)}{(K'_n)^{m_n}(i,j)}.$$

Proposition 2.6. *If* $m_n \asymp n^2$, *then*

$$c_n \to 1, \quad C_n \to 1, \quad \text{as } n \to \infty.$$

Proof: For $\ell \ge 1$, let $(i_0,i_1,...,i_\ell)$ be a path in $\{0,1,...,M_n\}$. Note that

$$\prod_{k=1}^{\ell} K_{n,1/2}(i_{k-1},i_k) \ge \left(\frac{3/4-\epsilon_n/2}{3/4}\right)^{\ell} \prod_{k=1}^{\ell} K'_n(i_{k-1},i_k).$$

This implies $c_n \geq (1-2\epsilon_n/3)^{m_n} \sim 1$ as $n \to \infty$. To see an upper bound of C_n, one may use Lemma 4.4 in Ding et al. (2010) to conclude that, for $0 \leq i \leq n$ and $\ell \geq 0$,

$$\begin{cases} K_{n,1/2}^{\ell}(i,j) \geq K_{n,1/2}^{\ell}(i,j-1) & \forall 1 \leq j \leq i \\ K_{n,1/2}^{\ell}(i,j) \geq K_{n,1/2}^{\ell}(i,j+1) & \forall i \leq j < n \end{cases},$$

and, for $0 \leq i \leq M_n$ and $\ell \geq 0$,

$$\begin{cases} (K_n')^{\ell}(i,j) \geq (K_n')^{\ell}(i,j-1) & \forall 1 \leq j \leq i \\ (K_n')^{\ell}(i,j) \geq (K_n')^{\ell}(i,j+1) & \forall i \leq j < M_n \end{cases}.$$

By the induction, the above observation implies that, for any probabilities μ, ν on $\{0, ..., n\}, \{0, ..., M_n\}$ satisfying $\mu(i) = \nu(i)$ for $0 \leq i \leq M_n$,

$$\mu K_{n,1/2}^{\ell}(j) \leq \nu (K_n')^{\ell}(j), \quad \forall 0 \leq j \leq M_n,\ \ell \geq 0.$$

This yields $C_n \leq 1$ for all $n \geq 3$. □

For $\epsilon \in (0,1)$, let $T'_{n,\mathrm{TV}}(\epsilon)$ be the total variation mixing time for K_n'. It is well-known that, for $\epsilon \in (0,1)$, $T'_{n,\mathrm{TV}}(\epsilon) \asymp n^2$. Let $d_{n,\mathrm{TV}}^{(1/2)}, d'_{n,\mathrm{TV}}$ be the total variation distance for $K_{n,1/2}, K_n'$. As a consequence of the above discussion, we obtain, for $\epsilon \in (0,1)$,

$$\limsup_{n\to\infty} d_{n,\mathrm{TV}}^{(1/2)}(T'_{n,\mathrm{TV}}(\epsilon)) \leq \frac{1}{2}\left(1 + \limsup_{n\to\infty} d'_{n,\mathrm{TV}}(T'_{n,\mathrm{TV}}(\epsilon))\right) \leq \frac{1+\epsilon}{2}.$$

Thus, for $\epsilon \in (1/2, 1)$, $T_{n,\mathrm{TV}}^{(1/2)}(\epsilon) = O(n^2)$. Note that, for $m_n = o(n^2)$,

$$\lim_{n\to\infty} \sum_{i \leq an} K_{n,1/2}^{m_n}(0,i) = 1, \quad \forall a > 0.$$

This yields $n^2 = O(T_{n,\mathrm{TV}}^{(1/2)}(\epsilon))$ for $\epsilon > 0$. The above discussion is also valid for the continuous time case and any δ-lazy discrete time case. We summarizes the results in the following theorem.

Theorem 2.7. *Let $\mathcal{F} = \{(\Omega_n, K_n, \pi_n) | n = 1, 2, ...\}$ be the family of birth and death chains in (2.1) and $\delta \in (0,1)$. Suppose that $\epsilon_n = o(n^{-2})$. Then, there is no total variation cutoff for $\mathcal{F}_c$ and $\mathcal{F}_\delta$. Furthermore, for $\epsilon \in (0, 1/2)$,*

$$T_{n,\mathrm{TV}}^{(c)}(\epsilon) \asymp T_{n,\mathrm{TV}}^{(\delta)}(\epsilon) \asymp n/\epsilon_n,$$

and, for $\epsilon \in (1/2, 1)$,

$$T_{n,\mathrm{TV}}^{(c)}(\epsilon) \asymp T_{n,\mathrm{TV}}^{(\delta)}(\epsilon) \asymp n^2.$$

Remark 2.8. Figure 2.1 displays the total variaton distances of the birth and death chains on $\{1, 2, ..., 100\}$ with transition matrices K_1 and K_2 given by

$$\begin{cases} K_1(i,i) = 1/2, & \text{for } i \notin \{1, 50, 51, 100\} \\ K_1(i,i+1) = K_1(i+1,i) = 1/4, & \text{for } i < 50 \text{ or } i > 51 \\ K_1(i,i) = 3/4 & \text{for } k \in \{1, 100\} \\ K_1(i,i+1) = K_1(i+1,i) = 10^{-3} & \text{for } k = 50 \\ K_1(i,i) = K_1(i,i) = 3/4 - 10^{-3} & \text{for } i \in \{50, 51\} \\ K_1(i,j) = 0 & \text{otherwise} \end{cases}$$

and

$$\begin{cases} K_2(i,i+1)=K_2(i+1,i)=10^{-2} & \text{for } i=25 \\ K_2(i,i)=3/4-10^{-2} & \text{for } i\in\{25,26\} \\ K_2(i,j)=K_1(i,j) & \text{otherwise} \end{cases} .$$

Note that each curve has only one sharp transition for $d_{\rm TV}(t)\le 1/2$. This is consistent with Theorem 1.3. These examples show that multiple sharp transitions may occur for $d_{\rm TV}(t)>1/2$. Note also that the flat part of the curves occupy very large time regions. For instance, the left most curve stays near the value $1/2$ for t between 10^3 and 10^6.

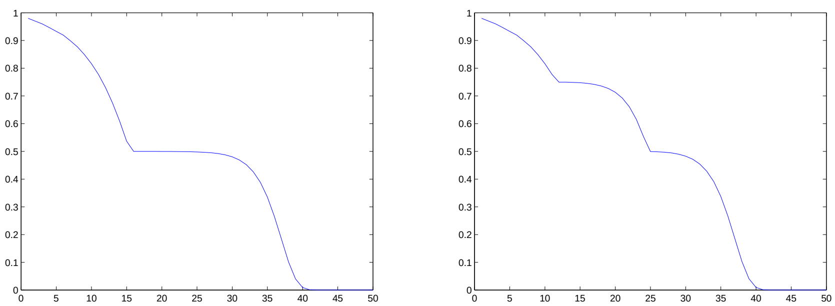

FIGURE 2.1. The curves display the total variation distance of the chains in Remark 2.8, where the left most curve is for K_1 and the right most curve is for K_2. The curve consists of the points $(m, d_{\rm TV}(100^{\lfloor 0.1\times m\rfloor}))$ with $m=1,2,...,50$. The right most point of each curve corresponds to $d_{\rm TV}(t)$ with $t=10^{10}$.

3. Bounds for mixing time and spectral gap

This section is dedicated to proving Theorems 1.1 and 1.2. In the first two subsections, we treat respectively the upper and lower bounds of the total variation mixing time. This leads to Theorem 1.1. In the third subsection, we provide a relaxation of the choice of i_n in Theorem 1.3. In the last subsection, we introduce a bound on the spectral gap which includes Theorem 1.2.

3.1. *An upper bound of the mixing time.* Let (Ω,K,π) be an irreducible birth and death chain, where $\Omega=\{0,1,...,n\}$ and K has birth rate p_i, death rate q_i and holding rate r_i. Let $(X_m)_{m=0}^\infty$ be a realization of the discrete time chain. Obviously, if N_t is a Poisson process with parameter 1 and independent of $(X_m)_{m=0}^\infty$, then $(X_{N_t})_{t\ge 0}$ is a realization of the continuous time chain. For $\delta\in[0,1)$, if $(B_m^{(\delta)})_{m=1}^\infty$ is a sequence of independent Bernoulli$(1-\delta)$ trials which are independent of $(X_m)_{m=0}^\infty$, then $Y_m^{(\delta)}=X_{B_1^{(\delta)}+\cdots+B_m^{(\delta)}}$ is a realization of the δ-lazy chain. For $0\le i\le n$, we define the first passage time to i by

$$\widetilde{\tau}_i := \inf\{t\ge 0|X_{N_t}=i\},\quad \tau_i^{(\delta)} := \min\{m\ge 0|Y_m=i\}, \tag{3.1}$$

and simply put $\tau_i := \tau_i^{(0)} = \min\{m \geq 0 | X_m = i\}$. Briefly, we write $\mathbb{P}_i(\cdot)$ for $\mathbb{P}(\cdot|X_0 = i)$ and write $\mathbb{E}_i$, Var_i as the expectation and variance under $\mathbb{P}_i$. The main result of this subsection is as follows.

Theorem 3.1 (Upper bound)**.** *Let (Ω, K, π) be an irreducible birth and death chain with $\Omega = \{0, 1, ..., n\}$. Let $\tau_i = \tau_i^{(0)}$ be the first passage time to i defined in (3.1). For $\epsilon \in (0, 1)$ and $\delta \in [1/2, 1)$,*

$$\max\left\{T_{\mathrm{TV}}^{(c)}(\epsilon), (1-\delta)T_{\mathrm{TV}}^{(\delta)}(\epsilon)\right\} \leq \frac{9(\mathbb{E}_0\tau_{i_0} + \mathbb{E}_n\tau_{i_0})}{\epsilon^2}, \tag{3.2}$$

where $i_0 \in \{0, ..., n\}$ satisfies $\pi([0, i_0 - 1]) \leq 1/2$ and $\pi([i_0 + 1, n]) \leq 1/2$.

Remark 3.2. Chen and Saloff-Coste (2012a) obtain a slightly improved upper bound similar to (3.2), which says that

$$\max\left\{T_{\mathrm{TV}}^{(c)}(\epsilon), (1-\delta)T_{\mathrm{TV}}^{(\delta)}(\epsilon)\right\} \leq \frac{(\sqrt{\epsilon} + \sqrt{1-\epsilon})(\mathbb{E}_0\tau_{i_0} + \mathbb{E}_n\tau_{i_0})}{\sqrt{\epsilon}}.$$

Comparing with (3.2), the above inequality has an improved dependence on ϵ.

To understand the right side of (3.2), we introduce the following lemma.

Lemma 3.3. *Referring to the setting in (3.1), it holds true that, for $i < j$, $\mathbb{E}_i(\tau_j^{(\delta)}) = \mathbb{E}_i(\tau_j)/(1-\delta)$ and $\mathbb{E}_i(\tau_j) = \mathbb{E}_i(\widetilde{\tau}_j) = \sum_{k=i}^{j-1} \pi([0, k])/(p_k\pi(k))$.*

Proof: The proof is based on the strong Markov property. See Barrera et al. (2009, Proposition 2) for a reference on the discrete time case, whereas the continuous time case is an immediate result of the fact $\{\widetilde{\tau}_i > t\} = \{\tau_i > N_t\}$. □

Remark 3.4. By Theorem 3.1 and Lemma 3.3, the total variation mixing time for the continuous time and the δ-lazy, with $\delta \geq 1/2$, discrete time birth and death chain on $\{0, 1, ..., n\}$ are bounded above by the following term up to a multiple constant.

$$\sum_{k=0}^{i_0-1} \frac{\pi([0, k])}{p_k\pi(k)} + \sum_{k=i_0+1}^{n} \frac{\pi([k, n])}{q_k\pi(k)},$$

where $i_0 \in \{0, ..., n\}$ satisfies $\pi([0, i_0 - 1]) \leq 1/2$ and $\pi([i_0 + 1, n]) \leq 1/2$.

Remark 3.5. In Theorem 3.1, i_0 is unique if $\pi([0, i]) \neq 1/2$ for all $0 \leq i \leq n$. If $\pi([0, j]) = 1/2$, then i_0 can be j or $j + 1$, but the right side of (3.2) is the same in either case using Lemma 3.3.

Remark 3.6. Let K be an irreducible birth and death chain with birth, death and holding rates p_i, q_i, r_i and stationary distribution π. Let λ be the spectral gap of K. As a consequence of Lemma 2.3 and theorem 3.1, we obtain, for $\epsilon \in (0, 1/2)$,

$$\lambda \geq \frac{\epsilon^2 \log(1/(2\epsilon))}{9} \left(\sum_{k=0}^{i_0-1} \frac{\pi([0, k])}{p_k\pi(k)} + \sum_{k=i_0+1}^{n} \frac{\pi([k, n])}{q_k\pi(k)}\right)^{-1},$$

where i_0 is such that $\pi([0, i_0 - 1]) \leq 1/2$ and $\pi([i_0 + 1, n]) \leq 1/2$. The maximum of $\epsilon^2 \log(1/(2\epsilon))$ on $(0, 1/2)$ is attained at $\epsilon = 1/(2\sqrt{e})$ and equal to $1/(8e)$. A similar lower bound of the spectral gap is also derived in Chen and Saloff-Coste (2012b) with improved constant.

As a simple application of Lemma 3.3, we have

Corollary 3.7. *Referring to Lemma 3.3, for $i \le j$,*

$$\mathbb{E}_i\tau_j \le \left(\frac{1}{\pi([j,n])} - 1\right)\mathbb{E}_n\tau_i.$$

Proof: By Lemma 3.3, one has

$$\mathbb{E}_i\tau_j = \sum_{k=i}^{j-1}\frac{\pi([0,k])}{p_k\pi(k)}, \quad \mathbb{E}_n(\tau_i) = \sum_{k=i}^{n-1}\frac{\pi([k+1,n])}{q_{k+1}\pi(k+1)} = \sum_{k=i}^{n-1}\frac{\pi([k+1,n])}{p_k\pi(k)}.$$

The inequality is then given by the fact $\pi([0,k])/\pi([k+1,n]) = 1/\pi([k+1,n]) - 1 \le 1/\pi([j,n]) - 1$ for $k < j$. □

The following proposition is the main technique used to prove Theorem 3.1.

Proposition 3.8. *Referring to the setting in (3.1), it holds true that, for $j < k$,*

$$d_{\mathrm{TV}}^{(c)}(i,t) \le \mathbb{P}_i(\max\{\widetilde{\tau}_j, \widetilde{\tau}_k\} > t) + 1 - \pi([j,k]),$$

and

$$d_{\mathrm{TV}}^{(1/2)}(i,t) \le \mathbb{P}_i(\max\{\tau_j^{(1/2)}, \tau_k^{(1/2)}\} > t) + 1 - \pi([j,k]),$$

In particular,

$$d_{\mathrm{TV}}^{(c)}(t) \le \frac{\mathbb{E}_0\widetilde{\tau}_k + \mathbb{E}_n\widetilde{\tau}_j}{t} + 1 - \pi([j,k])$$

and

$$d_{\mathrm{TV}}^{(1/2)}(t) \le \frac{2(\mathbb{E}_0\tau_k^{(1/2)} + \mathbb{E}_n\tau_j^{(1/2)})}{t} + 1 - \pi([j,k]).$$

In the above proposition, the discrete time case is discussed in Lemma 2.3 in Ding et al. (2010). Our method to prove this proposition is to construct a no-crossing coupling. We give the proof of the continuous time case for completeness and refer to Ding et al. (2010) for the discrete time case, where a heuristic idea on the construction of no-crossing coupling is proposed.

Proof of Proposition 3.8: Let $(Y_t)_{t\ge 0}$ be another process corresponding to H_t with $Y_0 \overset{d}{=} \pi$. Set $T := \inf\{t \ge 0 | X_t = Y_t\}$ and $Z_t := Y_t\mathbf{1}_{\{t\le T\}} + X_t\mathbf{1}_{\{t>T\}}$. Clearly, $(X_t, Z_t)_{t\ge 0}$ is a coupling for the semigroup H_t and must be no-crossing according to the continuous time setting. Note that $T = \inf\{t \ge 0 | X_t = Z_t\}$ is the coupling time of X_t and Z_t. The classical coupling statement implies that

$$d_{\mathrm{TV}}^{(c)}(i,t) \le \mathbb{P}_i(T > t). \tag{3.3}$$

See e.g. Aldous (1983) for a reference. Note that $X_{\tau_j} = j$, $X_{\tau_k} = k$ and

$$\mathbb{P}_i(X_{\widetilde{\tau}_j} \le Y_{\widetilde{\tau}_j}) = \pi([j,n]), \quad \mathbb{P}_i(X_{\widetilde{\tau}_k} \ge Y_{\widetilde{\tau}_k}) = \pi([0,k]).$$

As X_t, Y_t can not cross each other without coalescing in advance, this implies

$$\begin{aligned}\mathbb{P}_i(T \le \max\{\widetilde{\tau}_j, \widetilde{\tau}_k\}) &\ge \mathbb{P}_i(\min\{\widetilde{\tau}_j, \widetilde{\tau}_k\} \le T \le \max\{\widetilde{\tau}_j, \widetilde{\tau}_k\}) \\ &\ge \mathbb{P}_i(X_{\widetilde{\tau}_j} \le Y_{\widetilde{\tau}_j}, X_{\widetilde{\tau}_k} \ge Y_{\widetilde{\tau}_k}) \ge \pi([j,k]).\end{aligned}$$

Putting this back to (3.3) gives the desired result.

For the last part, note that if $i \le j$, then $\widetilde{\tau}_j < \widetilde{\tau}_k$ and, by Markov's inequality, this implies

$$\mathbb{P}_i(\max\{\widetilde{\tau}_j, \widetilde{\tau}_k\} > t) \le \mathbb{P}_0(\widetilde{\tau}_k > t) \le \mathbb{E}_0\widetilde{\tau}_k/t.$$

Similarly, for $i \geq k$, one can show that

$$\mathbb{P}_i(\max\{\widetilde{\tau}_j, \widetilde{\tau}_k\} > t) \leq \mathbb{P}_n(\widetilde{\tau}_j > t) \leq \mathbb{E}_n\widetilde{\tau}_j/t.$$

For $j < i < k$, we have

$$\mathbb{P}_i(\max\{\widetilde{\tau}_j, \widetilde{\tau}_k\} > t) \leq \mathbb{P}_i(\widetilde{\tau}_j > t) + \mathbb{P}_i(\widetilde{\tau}_k > t) \leq \frac{\mathbb{E}_n\widetilde{\tau}_j + \mathbb{E}_0\widetilde{\tau}_k}{t}.$$

□

Proof of Theorem 3.1: Set $j_\epsilon = \min\{i \geq 0 | \pi([0,i]) \geq \epsilon/3\}$ and $k_\epsilon = \min\{i \geq 0 | \pi([0,i]) \geq 1-\epsilon/3\}$. By Proposition 3.8 and Lemma 3.3, the choice of $j = j_\epsilon$ and $k = k_\epsilon$ implies that

$$T_{\mathrm{TV}}^{(c)}(\epsilon) \leq \frac{3(\mathbb{E}_0\tau_{k_\epsilon} + \mathbb{E}_n\tau_{j_\epsilon})}{\epsilon}.$$

By Corollary 3.7, one has

$$\mathbb{E}_0\tau_{k_\epsilon} = \mathbb{E}_0\tau_{i_0} + \mathbb{E}_{i_0}\tau_{k_\epsilon} \leq \mathbb{E}_0\tau_{i_0} + \left(\frac{3}{\epsilon} - 1\right)\mathbb{E}_n\tau_{i_0}$$

and

$$\mathbb{E}_n\tau_{j_\epsilon} = \mathbb{E}_n\tau_{i_0} + \mathbb{E}_{i_0}\tau_{j_\epsilon} \leq \mathbb{E}_n\tau_{i_0} + \left(\frac{3}{\epsilon} - 1\right)\mathbb{E}_0\tau_{i_0}.$$

Adding up both terms gives the upper bound in continuous time case. The proof for the (1/2)-lazy discrete time case is similar and, by Proposition 3.8, we obtain $T_{\mathrm{TV}}^{(1/2)}(\epsilon) \leq 18(\mathbb{E}_0\tau_{i_0} + \mathbb{E}_n\tau_{i_0})/\epsilon^2$. For $\delta \in (1/2, 1)$, note that $K_\delta = (K_{2\delta-1})_{1/2}$. Since the birth and death rates of $K_{2\delta-1}$ are $2(1-\delta)p_i$ and $2(1-\delta)q_i$, the above result and Lemma 3.3 lead to $T_{\mathrm{TV}}^{(\delta)}(\epsilon) \leq 9(\mathbb{E}_0\tau_{i_0} + \mathbb{E}_n\tau_{i_0})/((1-\delta)\epsilon^2)$. □

3.2. *A lower bound of the mixing time.* The goal of this subsection is to establish a lower bound on the total variation mixing time for birth and death chains. Recall the notations in the previous subsection. Let $(X_m)_{m=0}^\infty$ be an irreducible birth and death chain with transition matrix K and stationary distribution π. Let N_t be a Poisson process of parameter 1 that is independent of X_m. For $0 \leq i \leq n$, let $\tau_i = \min\{m \geq 0 | X_m = i\}$ and $\widetilde{\tau}_i = \inf\{t \geq 0 | X_{N_t} = i\}$. Then, the total variation mixing time satisfies

$$d_{\mathrm{TV}}(0,t) \geq K^t(0,[0,i-1]) - \pi([0,i-1]) \geq \mathbb{P}_0(\tau_i > t) - \pi([0,i-1]) \tag{3.4}$$

and

$$d_{\mathrm{TV}}^{(c)}(0,t) \geq H_t(0,[0,i-1]) - \pi([0,i-1]) \geq \mathbb{P}_0(\widetilde{\tau}_i > t) - \pi([0,i-1]). \tag{3.5}$$

Brown and Shao discuss the distribution of $\widetilde{\tau}_i$ in Brown and Shao (1987), of which proof also works for the discrete time case. In detail, if $-1 < \beta_1 < \cdots < \beta_i < 1$ are the eigenvalues of the submatrix of K indexed by $\{0, ..., i-1\}$ and $\lambda_j = 1 - \beta_j$, then

$$\mathbb{P}_0(\tau_i > t) = \sum_{j=1}^{i}\left(\prod_{k \neq j}\frac{\lambda_k}{\lambda_k - \lambda_j}\right)(1-\lambda_j)^t \tag{3.6}$$

and

$$\mathbb{P}_0(\widetilde{\tau}_i > t) = \sum_{j=1}^{i}\left(\prod_{k \neq j}\frac{\lambda_k}{\lambda_k - \lambda_j}\right)e^{-t\lambda_j}. \tag{3.7}$$

Note that, under $\mathbb{P}_0$, $\widetilde{\tau}_i$ is the sum of independent exponential random variables with parameters $\lambda_1, ..., \lambda_i$. If $\beta_1 > 0$, then τ is the sum of independent geometric random variables with parameters $\lambda_1, ..., \lambda_i$. In discrete time case, the requirement $\beta_1 > 0$ holds automatically for the δ-lazy chain with $\delta \geq 1/2$. The above formula leads to the following theorem.

Theorem 3.9 (Lower bound)**.** *Let K be the transition matrix of an irreducible birth and death chain on $\{0, 1, ..., n\}$. Let $\tau_i = \tau_i^{(0)}$ be the first passage time to i defined in (3.1). For $\delta \in [1/2, 1)$,*

$$\min\{T_{\mathrm{TV}}^{(c)}(1/10), 2(1-\delta)T_{\mathrm{TV}}^{(\delta)}(1/20)\} \geq \frac{\max\{\mathbb{E}_0\tau_{i_0}, \mathbb{E}_n\tau_{i_0}\}}{6},$$

where $i_0 \in \{0, ..., n\}$ satisfies $\pi([0, i_0 - 1]) \leq 1/2$ and $\pi([i_0 + 1, n]) \leq 1/2$.

Proof of Theorem 3.9: First, we consider the continuous time case. Let $\lambda_1, ..., \lambda_i$ be eigenvalues of the submatrix of $I - K$ indexed by $0, ..., i-1$ and $\widetilde{\tau}_{i,1}, ..., \widetilde{\tau}_{i,i}$ be independent exponential random variables with parameters $\lambda_1, ..., \lambda_i$. By (3.7), $\widetilde{\tau}_i$ and $\widetilde{\tau}_{i,1} + \cdots + \widetilde{\tau}_{i,i}$ are identically distributed under $\mathbb{P}_0$ and, by (3.5), this implies

$$d_{\mathrm{TV}}^{(c)}(0, t) \geq \mathbb{P}(\widetilde{\tau}_{i,1} + \cdots + \widetilde{\tau}_{i,i} > t) - \pi([0, i-1]).$$

It is easy to see that

$$\mathbb{E}_0\widetilde{\tau}_i = \frac{1}{\lambda_1} + \cdots + \frac{1}{\lambda_i}, \quad \mathrm{Var}_0(\widetilde{\tau}_i) = \frac{1}{\lambda_1^2} + \cdots + \frac{1}{\lambda_i^2}.$$

Let $a \in (0, 1)$ and consider the following two cases. If $1/\lambda_j > a\mathbb{E}_0\widetilde{\tau}_i$ for some $1 \leq j \leq i$, then

$$\mathbb{P}_0(\widetilde{\tau}_i > t) \geq \mathbb{P}(\widetilde{\tau}_{i,j} > t) > e^{-t/(a\mathbb{E}_0\widetilde{\tau}_i)}.$$

If $1/\lambda_j \leq a\mathbb{E}_0\widetilde{\tau}_i$ for all $1 \leq j \leq i$, then $\mathrm{Var}_0(\widetilde{\tau}_i) \leq a(\mathbb{E}_0\widetilde{\tau}_i)^2$ and, by the one-sided Chebyshev inequality, we have

$$\mathbb{P}_0(\widetilde{\tau}_i > t) \geq \frac{(t - \mathbb{E}_0\widetilde{\tau}_i)^2}{\mathrm{Var}_0(\widetilde{\tau}_i) + (t - \mathbb{E}_0\widetilde{\tau}_i)^2} \geq \frac{(t - \mathbb{E}_0\widetilde{\tau}_i)^2}{a(\mathbb{E}_0\widetilde{\tau}_i)^2 + (t - \mathbb{E}_0\widetilde{\tau}_i)^2} = \frac{(1-b)^2}{a + (1-b)^2},$$

for $t = b\mathbb{E}_0\widetilde{\tau}_i$ with $b \in (0, 1)$. Combining both cases and setting $i = i_0$ in (3.5) yields that, for $a, b \in (0, 1)$,

$$d_{\mathrm{TV}}^{(c)}(0, b\mathbb{E}_0\widetilde{\tau}_{i_0}) \geq \min\left\{e^{-b/a}, \frac{(1-b)^2}{a + (1-b)^2}\right\} - \frac{1}{2}. \tag{3.8}$$

Putting $a = 1/3$ and $b = 1/6$ gives $T_{\mathrm{TV}}^{(c)}(0, 1/10) \geq \mathbb{E}_0\widetilde{\tau}_{i_0}/6$.

For the discrete time case, note that the eigenvalues of the submatrix of $I - K_{1/2} = \frac{1}{2}(I - K)$ indexed by $0, ..., i-1$ are $\lambda_1/2, ..., \lambda_i/2$. Let $\tau_{i,1}, ..., \tau_{i,i}$ be independent geometric random variables with success probabilities $\lambda_1/2, ..., \lambda_i/2$. Replacing K with $K_{1/2}$ in (3.4), we obtain

$$d_{\mathrm{TV}}^{(1/2)}(0, t) \geq \mathbb{P}_0(\tau_{i,1} + \cdots + \tau_{i,i} > t) - \pi([0, i-1]).$$

Note that, under $\mathbb{P}_0$, $\tau_i^{(1/2)}$ has the same distribution as $\tau_{i,1} + \cdots + \tau_{i,i}$ and this implies

$$\mathbb{E}_0\tau_i^{(1/2)} = \frac{2}{\lambda_1} + \cdots + \frac{2}{\lambda_i}, \quad \mathrm{Var}_0(\tau_i^{(1/2)}) = \sum_{j=1}^{i} \frac{4(1 - \lambda_j/2)}{\lambda_j^2} \leq \sum_{j=1}^{i} \frac{4}{\lambda_j^2}.$$

Using the same analysis as before, one may derive, for $1/\mathbb{E}_0\tau_i^{(1/2)} < a < 1$ and $t < \mathbb{E}_0\tau_i^{(1/2)}$,

$$\mathbb{P}_0(\tau_i^{(1/2)} > t) \geq \min\left\{\left(1 - \frac{1}{a\mathbb{E}_0\tau_i^{(1/2)}}\right)^t, \frac{\left(t - \mathbb{E}_0\tau_i^{(1/2)}\right)^2}{a\left(\mathbb{E}_0\tau_i^{(1/2)}\right)^2 + \left(t - \mathbb{E}_0\tau_i^{(1/2)}\right)^2}\right\}.$$

By Lemma 3.3, $\mathbb{E}_0\tau_i^{(1/2)} \geq 2i$. Obviously, if $i_0 = 0$, then $T_{\mathrm{TV}}^{(1/2)}(0, 1/20) \geq 0 = \mathbb{E}_0\tau_{i_0}^{(1/2)}$. For $i_0 \geq 1$, $\mathbb{E}_0\tau_{i_0}^{(1/2)} \geq 2$ and the setting, $a = 2/3$ and $t = \left\lfloor \mathbb{E}_0\tau_{i_0}^{(1/2)}/12 \right\rfloor$, implies

$$d_{\mathrm{TV}}^{(1/2)}\left(0, \left\lfloor \mathbb{E}_0\tau_{i_0}^{(1/2)}/12 \right\rfloor\right) \geq \min\left\{2^{-1/3}, \frac{(11/12)^2}{2/3 + (11/12)^2}\right\} - \frac{1}{2} > \frac{1}{20},$$

where the first inequality use the fact that $s\log(1 - 3/(2s))$ is increasing on $[2, \infty)$. Hence, we have $T_{\mathrm{TV}}^{(1/2)}(0, 1/20) \geq \mathbb{E}_0\tau_{i_0}^{(1/2)}/12 = \mathbb{E}_0\tau_{i_0}/6$. For $\delta > 1/2$, the combination of the above result and the observation $K_\delta = (K_{2\delta-1})_{1/2}$ implies that $T_{\mathrm{TV}}^{(\delta)}(0, 1/20) \geq \mathbb{E}_0\tau_{i_0}/(12(1-\delta))$.

The analysis from the other end point gives the other lower bound. This finishes the proof. □

3.3. *Relaxation of the median condition.* In some cases, it is not easy to determine the value of i_n in Theorem 1.3. Let t_n be the constants in Theorem 3.1. For $c \in (0, 1)$, let $i_n(c) \in \{0, ..., n\}$ be the state such that $\pi_n([0, i_n(c) - 1]) \leq c$, $\pi_n([i_n(c) + 1, n]) \leq 1 - c$ and let $t_n(c)$ be the following constant

$$t_n(c) = \sum_{k=0}^{i_n(c)-1} \frac{\pi_n([0, k])}{\pi_n(k)p_{n,k}} + \sum_{k=i_n(c)+1}^{n} \frac{\pi_n([k, n])}{\pi_n(k)q_{n,k}}.$$

Assume that $c \geq 1/2$. In this case, if i_n is the smallest median, then $i_n \leq i_n(c)$ and

$$\sum_{k=i_n}^{i_n(c)-1} \frac{\pi([0, k])}{\pi_n(k)p_{n,k}} = \sum_{k=i_n+1}^{i_n(c)} \frac{\pi_n([0, k-1])}{\pi_n(k)q_{n,k}}.$$

Note that, for $i_n < k \leq i_n(c)$,

$$\frac{1}{2} \leq \pi_n([0, i_n]) \leq \frac{\pi_n([0, k-1])}{\pi_n([k, n])} \leq \frac{1}{\pi_n([i_n(c), n])} \leq \frac{1}{1-c}.$$

This implies $t_n/2 \leq t_n(c) \leq t_n/(1-c)$. Similarly, for $c \leq 1/2$, one can show that $t_n/2 \leq t_n(c) \leq t_n/c$. Combining both cases gives

$$t_n/2 \leq t_n(c) \leq t_n/\min\{c, 1-c\}. \tag{3.9}$$

As a consequence of the above discussion, we obtain the following theorem.

Theorem 3.10. *Referring to Theorem 1.3. For $n \geq 1$, let $j_n \in \{0, 1, ..., n\}$ and set*

$$t_n' = \max\left\{\sum_{k=0}^{j_n-1} \frac{\pi_n([0, k])}{\pi_n(k)p_{n,k}}, \sum_{k=j_n+1}^{n} \frac{\pi_n([k, n])}{\pi_n(k)q_{n,k}}\right\}.$$

Suppose that

$$0 < \liminf_{n\to\infty} \pi_n([0, j_n]) \leq \limsup_{n\to\infty} \pi_n([0, j_n]) < 1.$$

Then, Theorem 1.3 remains true if t_n is replaced by t'_n.

Proof: The proof comes immediately from (3.9) with $c = \pi_n([0, j_n])$. □

We use this observation to bound the cutoff time in the following theorem.

Theorem 3.11. *Referring to Theorem 1.3. Suppose that $\mathcal{F}_c$ has a total variation cutoff. Then, for any $\epsilon \in (0,1)$,*

$$\frac{2\log 2}{5} \leq \liminf_{n\to\infty} \frac{T^{(c)}_{n,\mathrm{TV}}(\epsilon)}{t_n} \leq \limsup_{n\to\infty} \frac{T^{(c)}_{n,\mathrm{TV}}(\epsilon)}{t_n} \leq 2$$

Proof of Theorem 3.11: The upper bound is given by Remark 3.2 and the fact, $\max\{s,t\} \geq (s+t)/2$, whereas the lower bound is obtained by applying $a = 2/5$ and $b = a\log(2/(1+2\epsilon))$ in (3.8) with $\epsilon \to 0$. □

3.4. *Bounding the spectral gap.* This subsection is devoted to poviding bounds on the specral gap for birth and death chains. As the graph associated with a birth and death chain is a path, weighted Hardy's inequality can be used to bound the spectral gap. We refer to the Appendix for a detailed discussion of the following results. See Theorems A.1-A.5.

Theorem 3.12. *Consider an irreducible birth and death chain on $\{0,...,n\}$ with birth, death and holding rates p_i, q_i, r_i and stationary distribution π. Let λ be the spectral gap and set, for $0 \leq i \leq n$,*

$$C(i) = \max\left\{\max_{j:j<i}\sum_{k=j}^{i-1}\frac{\pi([0,j])}{\pi(k)p_k}, \max_{j:j>i}\sum_{k=i+1}^{j}\frac{\pi([j,n])}{\pi(k)q_k}\right\}.$$

Then, for $0 \leq m \leq n$,

$$\frac{1}{4C(m)} \leq \lambda \leq \frac{1}{\min\{\pi([0,m]),\pi([m,n])\}C(m)}.$$

In particular, if M is a median of π, that is, $\pi([0,M]) \geq 1/2$ and $\pi([M,n]) \geq 1/2$, then

$$\frac{1}{4C(M)} \leq \lambda \leq \frac{2}{C(M)}.$$

Theorem 3.13. *Consider an irreducible birth and death chain on $\{0,...,n\}$ with birth, death and holding rates p_i, q_i, r_i and stationary distribution π. Let λ be the spectral gap and set $N = \lceil n/2 \rceil$. Suppose that $p_i = q_{n-i}$ for $0 \leq i \leq n$. Then,*

$$\frac{1}{4C} \leq \lambda \leq \frac{1}{C},$$

where

$$C = \max_{0\leq i\leq N-1}\left\{\pi([0,i])\sum_{j=i}^{N-1}\frac{1}{\pi(j)p_j}\right\} \quad \textit{if } n \textit{ is even},$$

and

$$C = \max_{0\leq i\leq N-1}\left\{\pi([0,i])\left(\sum_{j=i}^{N-2}\frac{1}{\pi(j)p_j} + \frac{1}{2\pi(N-1)p_{N-1}}\right)\right\} \quad \textit{if } n \textit{ is odd}.$$

Remark 3.14. In Saloff-Coste (1999), the author also obtained bounds similar to Theorem 3.13 for the case $\pi(i) \geq \pi(i+1)$ with $0 \leq i < n/2$ using the path technique. For more information on path techniques, see Diaconis and Saloff-Coste (1993a,b); Diaconis and Stroock (1991) and the references therein.

4. Examples

In this section, we will apply the theory developed in the previous section to examples of special interest. First, we give a criterion on the cutoff using the birth and death rates.

Theorem 4.1 (Cutoffs from birth and death rates)**.** *Let $\mathcal{F} = \{(\Omega_n, K_n, \pi_n)|n = 1, 2, ...\}$ be a family of irreducible birth and death chains on $\Omega_n = \{0, 1, ..., n\}$ with birth rate, $p_{n,i}$, death rate $q_{n,i}$ and holding rate $r_{n,i}$. Let λ_n be the spectral gap of K_n. For $n \geq 1$, let $j_n \in \{0, ..., n\}$ and set*

$$t_n = \max\left\{\sum_{k=0}^{j_n-1} \frac{\pi_n([0,k])}{\pi_n(k)p_{n,k}}, \sum_{k=j_n+1}^{n} \frac{\pi_n([k,n])}{\pi_n(k)q_{n,k}}\right\}$$

and

$$\ell_n = \max\left\{\max_{j:j<j_n} \sum_{k=j}^{j_n-1} \frac{\pi_n([0,j])}{\pi_n(k)p_{n,k}}, \max_{j:j>j_n} \sum_{k=j_n+1}^{j} \frac{\pi_n([j,n])}{\pi_n(k)q_{n,k}}\right\}.$$

Suppose that

$$0 < \liminf_{n\to\infty} \pi_n([0, j_n]) \leq \limsup_{n\to\infty} \pi_n([0, j_n]) < 1.$$

Then, for $\epsilon \in (0, 1/2)$ and $\delta \in (0, 1)$,

$$\lambda_n \asymp 1/\ell_n, \quad T^{(c)}_{n,\mathrm{TV}}(\epsilon) \asymp t_n \asymp T^{(\delta)}_{n,\mathrm{TV}}(\epsilon).$$

Furthermore, the following are equivalent.

1. *$\mathcal{F}_c$ has a cutoff in total variation.*
2. *For $\delta \in (0, 1)$, $\mathcal{F}_\delta$ has a cutoff in total variation.*
3. *$\mathcal{F}_c$ has precutoff in total variation.*
4. *For $\delta \in (0, 1)$, $\mathcal{F}_\delta$ has a precutoff in total variation.*
5. *$t_n/\ell_n \to \infty$.*

The above theorem is obvious from Theorems 2.4, 3.10 and 3.12. We use two classical examples, simple random walks and Ehrenfest chains, to illustrate how to apply Theorem 4.1 to determine the total variation cutoff and mixing times.

Example 4.2 (Simple random walks on finite paths). For $n \geq 1$, the simple random walk on $\{0, ..., n\}$ is a birth and death chain with $p_{n,i} = q_{n,i+1} = 1/2$ for $0 \leq i < n$ and $r_{n,0} = r_{n,n} = 1/2$. It is clear that K_n is irreducible and aperiodic with uniform stationary distribution. Let t_n, ℓ_n be the constants in Theorem 4.1. It is an easy exercise to show that $\ell_n \asymp n^2 \asymp t_n$. By Theorem 4.1, neither $\mathcal{F}_c$ nor $\mathcal{F}_\delta$ has total variation precutoff, but $T^{(c)}_{n,\mathrm{TV}}(\epsilon) \asymp n^2 \asymp T^{(\delta)}_{n,\mathrm{TV}}(\epsilon)$ for $\epsilon \in (0, 1/2)$ and $\delta \in (0, 1)$. In fact, one may use a hitting time statement to prove that the mixing time has order at least n^2, when $\epsilon \in [1/2, 1)$. This implies that the above approximation of mixing time holds for $\epsilon \in (0, 1)$.

Example 4.3 (Ehrenfest chains). Consider the Ehrenfest chain on $\{0,...,n\}$, which is a birth and death chain with rates $p_{n,i} = 1 - i/n$ and $q_{n,i} = i/n$. It is obvious that K_n is irreducible and periodic with stationary distribution $\pi_n(i) = 2^{-n}\binom{n}{i}$. An application of the representation theory shows that, for $0 \le i \le n$, $2i/n$ is an eigenvalue of $I - K_n$. Let λ_n, s_n be the constants in Theorem 2.4. Clearly, $\lambda_n = 2/n$ and $s_n \asymp n\log n$ and, by Theorem 2.4, both $\mathcal{F}_c$ and $\mathcal{F}_\delta$ have a total variation cutoff. Note that, as a simple corollary, one obtains the non-trivial estimates

$$\sum_{i=0}^{\lceil \frac{n}{2}\rceil - 1} \frac{\binom{n}{0} + \cdots + \binom{n}{i}}{\binom{n}{i}} \asymp n\log n, \quad \max_{0\le i<n/2} \sum_{j=0}^{i}\binom{n}{j} \times \sum_{j=i}^{\lceil \frac{n}{2}\rceil - 1}\binom{n}{i}^{-1} \asymp n.$$

For a detailed computation on the total variation and the L^2-distance, see e.g. Diaconis (1988).

In the next subsections, we consider birth and death chains of special types.

4.1. *Chains with valley stationary distributions.* In this subsection, we consider birth and death chains with valley stationary distribution. For $n \ge 1$, let $\Omega_n = \{0, 1, ..., n\}$ and K_n be an irreducible birth and death chain on Ω_n with birth, death and holding rates, $p_{n,i}, q_{n,i}, r_{n,i}$. Suppose that there is $j_n \in \Omega_n$ such that

$$p_{n,i} \le q_{n,i+1},\ \forall i < j_n, \quad p_{n,i} \ge q_{n,i+1},\ \forall i \ge j_n. \tag{4.1}$$

Obviously, the stationary distribution π_n of K_n satisfies $\pi_n(i) \ge \pi_n(i+1)$ for $i < j_n$ and $\pi_n(i) \le \pi_n(i+1)$ for $i \ge j_n$.

Let t_n, ℓ_n be the constants in Theorem 4.1 and write

$$\ell_n = \max\left\{ \max_{j:j<j_n} \sum_{k=j+1}^{j_n} \frac{\pi_n([0,j])}{\pi_n(k)q_{n,k}}, \max_{j:j>j_n} \sum_{k=j_n}^{j-1} \frac{\pi_n([j,n])}{\pi_n(k)p_{n,k}} \right\}.$$

Set

$$M_L = \max_{0<i\le j_n} q_{n,i},\ m_L = \min_{0<i\le j_n} q_{n,i},\ M_R = \max_{j_n\le i<n} p_{n,i},\ m_R = \min_{j_n\le i<n} p_{n,i}.$$

Clearly,

$$\ell_n \le \max\left\{ \frac{\pi_n([0,j_n])}{m_L} \sum_{i=0}^{j_n} \frac{1}{\pi_n(i)}, \frac{\pi_n([j_n,n])}{m_R} \sum_{i=j_n}^{n} \frac{1}{\pi_n(i)} \right\}.$$

Let j_n' be such that $\pi_n([0,j_n']) \ge \pi_n([0,j_n])/2$ and $\pi_n([j_n',j_n]) \ge \pi_n([0,j_n])/2$. Note that if $j_n \ge 1$, then $j_n \ge \max\{2j_n', j_n'+1\}$. By (4.1), this implies

$$\sum_{k=j_n'+1}^{j_n} \frac{\pi_n([0,j_n'])}{\pi_n(k)} \ge \frac{\pi_n([0,j_n])}{4} \sum_{k=j_n'}^{j_n} \frac{1}{\pi_n(k)} \ge \frac{\pi_n([0,j_n])}{8} \sum_{k=0}^{j_n} \frac{1}{\pi_n(k)}.$$

One can derive a similar inequality from the other end point and this yields

$$\ell_n \ge \frac{1}{8}\min\left\{ \frac{\pi_n([0,j_n])}{M_L} \sum_{i=0}^{j_n} \frac{1}{\pi_n(i)}, \frac{\pi_n([j_n,n])}{M_R} \sum_{i=j_n}^{n} \frac{1}{\pi_n(i)} \right\}.$$

For t_n, note that

$$\frac{\pi_n([0,j_n-1])}{2} \sum_{k=0}^{j_n-1} \frac{1}{\pi_n(k)} \le \sum_{k=0}^{j_n-1} \frac{\pi_n([0,k])}{\pi_n(k)} \le \pi_n([0,j_n-1]) \sum_{k=0}^{j_n-1} \frac{1}{\pi_n(k)}$$

and

$$\frac{\pi_n([j_n+1,n])}{2}\sum_{k=j_n+1}^{n}\frac{1}{\pi_n(k)}\le\sum_{k=j_n+1}^{n}\frac{\pi_n([k,n])}{\pi_n(k)}\le\pi_n([j_n+1,n])\sum_{k=j_n+1}^{n}\frac{1}{\pi_n(k)}$$

This implies

$$t_n\le\max\left\{\frac{\pi_n([0,j_n])}{m_L}\sum_{i=0}^{j_n}\frac{1}{\pi_n(i)},\frac{\pi_n([j_n,n])}{m_R}\sum_{i=j_n}^{n}\frac{1}{\pi_n(i)}\right\}$$

and

$$t_n\ge\frac{1}{8}\max\left\{\frac{\pi_n([0,j_n])}{M_L}\sum_{i=0}^{j_n}\frac{1}{\pi_n(i)},\frac{\pi_n([j_n,n])}{M_R}\sum_{i=j_n}^{n}\frac{1}{\pi_n(i)}\right\}$$

The following theorem is an immediate consequence of the above discussion and Theorem 4.1.

Theorem 4.4. *Let $\mathcal{F}=\{(\Omega_n,K_n,\pi_n)|n=1,2,...\}$ be a family of birth and death chains satisfying (4.1). Assume that $\pi_n([0,j_n])\asymp\pi_n([j_n,n])$ and*

$$\max_{0<i\le j_n}q_{n,i}\asymp\min_{0<i\le j_n}q_{n,i},\quad\max_{j_n\le i<n}p_{n,i}\asymp\min_{j_n\le i<n}p_{n,i}.$$

Then, there is no cutoff for $\mathcal{F}_c,\mathcal{F}_\delta$ and, for $\epsilon\in(0,1/2)$ and $\delta\in(0,1)$,

$$T^{(c)}_{n,\mathrm{TV}}(\epsilon)\asymp T^{(\delta)}_{n,\mathrm{TV}}(\epsilon)\asymp\frac{1}{\lambda_n}\asymp\max\left\{\frac{1}{q_{n,j_n}}\sum_{i=0}^{j_n}\frac{1}{\pi_n(i)},\frac{1}{p_{n,j_n}}\sum_{i=j_n}^{n}\frac{1}{\pi_n(i)}\right\}.$$

For an illustration of the above theorem, we consider the following Markov chains. For $n\ge1$, let $\Omega_n=\{0,1,...,n\}$, π_n be a non-uniform probability distribution on Ω_n satisfying (4.1) and M_n be a transition matrix given by

$$M_n(i,j)=\begin{cases}1/2 & \text{for } j=i-1, i\le j_n,\\ 1/2 & \text{for } j=i+1, i\ge j_n,\\ \pi_n(i+1)/(2\pi_n(i)) & \text{for } j=i+1, i<j_n,\\ \pi_n(i-1)/(2\pi_n(i)) & \text{for } j=i-1, i>j_n,\\ 1/2-\pi_n(i+1)/(2\pi_n(i)) & \text{for } j=i<j_n,\\ 1/2-\pi_n(i-1)/(2\pi_n(i)) & \text{for } j=i>j_n.\end{cases}\tag{4.2}$$

Note that M_n is the Metropolis chain for π_n associated to the simple random walk on Ω_n. For more information on the Metropolis chain, see Diaconis and Saloff-Coste (1998) and the references therein. The next theorem is a corollary of Theorem 4.4.

Theorem 4.5. *Let $\mathcal{F}=\{(\Omega_n,M_n,\pi_n)|n=1,2,..\}$ be the family of Metropolis chains satisfying (4.1)-(4.2). Suppose $\pi_n([0,j_n])\asymp\pi_n([j_n,n])$. Then, neither $\mathcal{F}_c$ nor $\mathcal{F}_\delta$ has a total variation precutoff but, for $\epsilon\in(0,1/2)$ and $\delta\in(0,1)$,*

$$T^{(c)}_{n,\mathrm{TV}}(\epsilon)\asymp\sum_{i=0}^{n}\frac{1}{\pi_n(i)}\asymp T^{(\delta)}_{n,\mathrm{TV}}(\epsilon).$$

Example 4.6. Let $a>0$ and $\check{\pi}_{n,a},\hat{\pi}_{n,a}$ be probability measures on $\{0,\pm1,...,\pm n\}$ given by

$$\check{\pi}_{n,a}(i)=\check{c}_{n,a}(|i|+1)^a,\quad\hat{\pi}_{n,a}(i)=\hat{c}_{n,a}(n-|i|+1)^a,\tag{4.3}$$

where $\check{c}_{n,a}, \hat{c}_{n,a}$ are normalizing constants. Let $\check{\mathcal{F}}, \hat{\mathcal{F}}$ be families of the Metropolis chains for $\check{\pi}_{n,a}, \hat{\pi}_{n,a}$ associated to the simple random walks on $\{0, \pm 1, ..., \pm n\}$, that is,

$$\check{M}_{n,a}(i,j) = \check{M}_{n,a}(-i,-j), \quad \hat{M}_{n,a}(i,j) = \hat{M}_{n,a}(-i,-j)$$

and

$$\check{M}_{n,a}(i,j) = \begin{cases} \frac{1}{2} & \text{if } j = i+1, i \in [0, n-1] \\ \frac{i^a}{2(i+1)^a} & \text{if } j = i-1, i \in [1, n] \\ \frac{(i+1)^a - i^a}{2(i+1)^a} & \text{if } j = i, i \notin \{0, n\} \\ 1 - \frac{n^a}{2(n+1)^a} & \text{if } i = j = n \end{cases}$$

and

$$\hat{M}_{n,a}(i,j) = \begin{cases} \frac{1}{2} & \text{if } j = i-1, i \in [1, n] \\ \frac{(n-i)^a}{2(n-i+1)^a} & \text{if } j = i+1, i \in [0, n-1] \\ \frac{(n-i+1)^a - (n-i)^a}{2(n-i+1)^a} & \text{if } j = i \neq 0 \\ 1 - \frac{n^a}{(n+1)^a} & \text{if } i = j = 0 \end{cases}.$$

Let $\check{\lambda}_{n,a}, \hat{\lambda}_{n,a}$ and $\check{T}_{n,a}, \hat{T}_{n,a}$ be the spectral gaps and total variation mixing times of $\check{M}_{n,a}, \hat{M}_{n,a}$. It has been proved in Chen and Saloff-Coste (2012b); Saloff-Coste (1999) that there is $C > 1$ such that, for all $a > 0$ and $n \geq 1$,

$$\frac{1}{C\check{\lambda}_{n,a}} \asymp n^a \left(\left(1 + \frac{1}{n} \right)^a + \frac{n}{1+a} \right) (1 + v(n,a)) \leq \frac{C}{\check{\lambda}_{n,a}}$$

and

$$\frac{1}{C\hat{\lambda}_{n,a}} \leq \frac{(n+a)^2}{(1+a)^2} \leq \frac{C}{\hat{\lambda}_{n,a}},$$

where $v(n,1) = \log n$ and $v(n,a) = (n^{1-a} - 1)/(1-a)$ for $a \neq 1$. By Theorem 4.4, $\check{\mathcal{F}}_c$ and $\check{\mathcal{F}}_\delta$ have no cutoff in total variation but, for fixed $a > 0$, $\epsilon \in (0, 1/2)$ and $\delta \in (0, 1)$,

$$\check{T}^{(c)}_{n,a}(\epsilon) \asymp \check{T}^{(\delta)}_{n,a}(\epsilon) \asymp \begin{cases} n^2 & \text{if } a \in (0,1) \\ n^2 \log n & \text{if } a = 1 \\ n^{1+a} & \text{if } a \in (1, \infty) \end{cases}.$$

The above result in continuous time case is also obtained in Saloff-Coste (1999).

To see the cutoff for $\hat{\mathcal{F}}$, let

$$t_n = \sum_{k=0}^{n-1} \frac{\hat{\pi}_{n,a}([-n, -n+k])}{\hat{\pi}_{n,a}(-n+k)} = \sum_{k=1}^{n} k^{-a} \sum_{j=1}^{k} j^a.$$

By Theorems 3.1-3.9, we have

$$\frac{2t_n}{3} \leq \hat{T}^{(c)}_{n,a}(1/10) \leq 3600 t_n.$$

Note that, for $k \geq 1$ and $a > 0$,

$$\frac{k^a(k+a)}{2(1+a)} \leq \sum_{j=1}^{k} j^a \leq \frac{2k^a(k+a)}{1+a}.$$

This implies

$$\frac{n(n+a)}{6(1+a)} \leq \hat{T}^{(c)}_{n,a}(1/10) \leq \frac{14400 n(n+a)}{1+a}.$$

We collect the above results in the following theorem.

Theorem 4.7. For $n \geq 1$, let $a_n > 0$ and $\check{\pi}_{n,a_n}, \hat{\pi}_{n,a_n}$ be probability measures given by (4.3). Let $\check{\mathcal{F}}, \hat{\mathcal{F}}$ be the families of Metropolis chains for $\check{\pi}_{n,a_n}, \hat{\pi}_{n,a_n}$ as above with total variation mixing time $\check{T}_{n,\mathrm{TV}}, \hat{T}_{n,\mathrm{TV}}$. Then, for $\epsilon \in (0, 1/2)$ and $\delta \in (0, 1)$,

$$\hat{T}^{(c)}_{n,\mathrm{TV}}(\epsilon) \asymp \hat{T}^{(\delta)}_{n,\mathrm{TV}}(\epsilon) \asymp \frac{n(n + a_n)}{1 + a_n}$$

and

$$\check{T}^{(c)}_{n,\mathrm{TV}}(\epsilon) \asymp \check{T}^{(\delta)}_{n,\mathrm{TV}}(\epsilon) \asymp n^{a_n}\left(\left(1 + \frac{1}{n}\right)^{a_n} + \frac{n}{1 + a_n}\right)(1 + v(n, a_n)),$$

where $v(n, 1) = \log n$ and $v(n, a) = (n^{1-a} - 1)/(1 - a)$ for $a \neq 1$.

Moreover, neither $\check{\mathcal{F}}_c$ nor $\check{\mathcal{F}}_\delta$ has a total variation cutoff. Also, $\hat{\mathcal{F}}_c$ and $\hat{\mathcal{F}}_\delta$ have a total variation cutoff if and only if $a_n \to \infty$.

4.2. *Chains with monotonic stationary distributions.* In this subsection, we consider birth and death chains with monotonic stationary distributions. For $n \geq 1$, let $\Omega_n = \{0, 1, ..., n\}$ and K_n be a birth and death chain on Ω_n with birth, death and holding rates, $p_{n,i}, q_{n,i}, r_{n,i}$. Suppose that

$$p_{n,i} \geq q_{n,i+1}, \quad \forall 0 \leq i < n. \tag{4.4}$$

If K_n is irreducible, then the stationary distribution π_n satisfying $\pi_n(i) \leq \pi_n(i+1)$ for $0 \leq i < n$. Let $j_n \in \Omega_n$ and t_n, ℓ_n be the constants in Theorem 4.1. Assume that $\pi_n([0, j_n]) \asymp \pi_n([j_n, n])$ and

$$\max_{0 \leq i < j_n} p_{n,i} \asymp \min_{0 \leq i < j_n} p_{n,i}, \quad \max_{j_n \leq i < n} p_{n,i} \asymp \min_{j_n \leq i < n} p_{n,i}. \tag{4.5}$$

Using a discussion similar to that in front of Theorem 4.4, one can show that

$$t_n \asymp \max\left\{\frac{1}{p_{n,1}} \sum_{k=0}^{j_n - 1} \frac{\pi_n([0, k])}{\pi_n(k)}, \frac{1}{p_{n,j_n}} \sum_{k=j_n}^{n} \frac{1}{\pi_n(k)}\right\}$$

and

$$\ell_n \asymp \max\left\{\frac{1}{p_{n,1}} \max_{0 \leq j < j_n} \sum_{k=j}^{j_n - 1} \frac{\pi_n([0, j])}{\pi_n(k)}, \frac{1}{p_{n,j_n}} \sum_{k=j_n}^{n} \frac{1}{\pi_n(k)}\right\}.$$

This leads to the following theorem.

Theorem 4.8. *Let $\mathcal{F} = \{(\Omega_n, K_n, \pi_n) | n = 1, 2, ...\}$ be a family of irreducible birth and death chains with $\Omega_n = \{0, 1, ..., n\}$ and birth, death and holding rates $p_{n,i}, q_{n,i}, r_{n,i}$. Let $\lambda_n, T_{n,\mathrm{TV}}$ be the spectral gap and total variation mixing time of K_n and set*

$$u_n = \sum_{k=0}^{j_n - 1} \frac{\pi_n([0, k])}{\pi_n(k)}, \quad v_n = \max_{0 \leq j < j_n} \sum_{k=j}^{j_n - 1} \frac{\pi_n([0, j])}{\pi_n(k)}, \quad w_n = \sum_{k=j_n}^{n} \frac{1}{\pi_n(k)}.$$

Assume that $\pi_n([0, j_n]) \asymp \pi_n([j_n, n])$ and (4.5) holds. Then, for $\epsilon \in (0, 1/2)$ and $\delta \in (0, 1)$,

$$\lambda_n^{-1} \asymp \max\left\{\frac{v_n}{p_{n,1}}, \frac{w_n}{p_{n,j_n}}\right\}, \quad T^{(c)}_{n,\mathrm{TV}}(\epsilon) \asymp T^{(\delta)}_{n,\mathrm{TV}}(\epsilon) \asymp \max\left\{\frac{u_n}{p_{n,1}}, \frac{w_n}{p_{n,j_n}}\right\}.$$

Moreover, $\mathcal{F}_c$ and $\mathcal{F}_\delta$ have a total variation cutoff if and only if

$$u_n/v_n \to \infty, \quad (u_n p_{n,j_n})/(w_n p_{n,1}) \to \infty.$$

For $n \geq 1$, let f_n be a non-decreasing function on $[0,n]$ and set $F_n(x) = \int_0^x f_n(t)dt$ and $G_n(x,m) = \int_x^m 1/f_n(t)dt$. Note that if there is $C > 1$ such that

$$C^{-1} f_n(i)\pi_n(0) \leq \pi_n(i) \leq C f_n(i)\pi_n(0), \quad \forall 0 \leq i \leq n,\, n \geq 1,$$

then

$$\frac{1}{2C^2}\left(\frac{F_n(k)}{f_n(k)} + 1\right) \leq \frac{\pi_n([0,k])}{\pi_n(k)} \leq C^2\left(\frac{F_n(k)}{f_n(k)} + 1\right)$$

and

$$\frac{1}{2C}\left(G_n(j,j_n) + \frac{1}{f_n(j)}\right) \leq \pi_n(0)\sum_{k=j}^{j_n-1}\frac{1}{\pi_n(k)} \leq C\left(G_n(j,j_n) + \frac{1}{f_n(j)}\right).$$

This implies

$$\pi_n([0,j])\sum_{k=j}^{j_n-1}\frac{1}{\pi_n(k)} \leq C^2\left(G_n(j,j_n) + \frac{1}{f_n(j)}\right)(F_n(j) + f_n(j))$$

and

$$\pi_n([0,j])\sum_{k=j}^{j_n-1}\frac{1}{\pi_n(k)} \geq \frac{1}{4C^2}\left(G_n(j,j_n) + \frac{1}{f_n(j)}\right)(F_n(j) + f_n(j)).$$

Let u_n, v_n, w_n be the constants in Theorem 4.8 and assume that

$$\min_{0\leq i<n} p_{n,i} \asymp \max_{0\leq i<n} p_{n,i} \asymp 1.$$

Consider the following cases.

Case 1: $f_n(x) = \exp\{\alpha_n x^{\beta_n}\}$ with $\inf_n \alpha_n > 0$ and $\inf_n \beta_n \geq 1$. In this case, $F_n(x) = O(f_n(x))$ and $G_n(x,m) = O(1/f_n(x))$ for $1 \leq x < m$. By setting $j_n = n$, we obtain

$$\pi_n([0,j_n]) \asymp \pi_n([j_n,n]), \quad u_n \asymp n, \quad v_n \asymp w_n \asymp 1.$$

By Theorem 4.8, $\lambda_n \asymp 1$ and, for $\epsilon \in (0,1/2)$ and $\delta \in (0,1)$,

$$T^{(c)}_{n,\mathrm{TV}}(\epsilon) \asymp T^{(\delta)}_{n,\mathrm{TV}}(\epsilon) \asymp n.$$

There is a total variation cutoff for $\mathcal{F}_c$ or $\mathcal{F}_\delta$.

Case 2: $f_n(x) = \exp\{\alpha_n x^{\beta_n}\}$ with $0 < \inf_n \alpha_n \leq \sup_n \alpha_n < \infty$ and $0 < \inf_n \beta_n \leq \sup_n \beta_n < 1$. Note that, for $\alpha \in \mathbb{R}$ and $\beta \in (0,1)$,

$$\frac{d}{dx}\left(x^{1-\beta}e^{\alpha x^\beta}\right) = \left(\alpha\beta + (1-\beta)x^{-\beta}\right)e^{\alpha x^\beta}.$$

This implies that, uniformly for $n/2 \leq x$ and $1 + x \leq m \leq n$,

$$F_n(x) \asymp x^{1-\beta_n} f_n(x), \quad G_n(x,m) \asymp \left(\frac{x^{1-\beta_n}}{f_n(x)} - \frac{m^{1-\beta_n}}{f_n(m)}\right).$$

Letting $j_n = \lfloor n - n^{1-\beta_n} \rfloor$ yields

$$\pi_n([0,j_n]) \asymp \pi_n([j_n,n]), \quad u_n \asymp n^{2-\beta_n}, \quad v_n \asymp n^{2-2\beta_n} \asymp w_n.$$

By Theorem 4.8, $\mathcal{F}_c$ and $\mathcal{F}_\delta$ have a total variation cutoff and

$$\lambda_n \asymp n^{2\beta_n - 2}, \quad T^{(c)}_{n,\mathrm{TV}}(\epsilon) \asymp T^{(\delta)}_{n,\mathrm{TV}}(\epsilon) \asymp n^{2-\beta_n}, \quad \forall \epsilon \in (0,1/2),\, \delta \in (0,1).$$

Case 3: $f_n(x) = \exp\{\alpha_n[\log(x+1)]^{\beta_n}\}$ with $0 < \inf_n \alpha_n \le \sup_n \alpha_n < \infty$ and $1 < \inf_n \beta_n \le \sup_n \beta_n < \infty$. Note that, for $\alpha \in \mathbb{R}$ and $\beta > 1$,

$$\frac{d}{dx}\left(\frac{(x+1)e^{\alpha[\log(x+1)]^\beta}}{[\log(x+1)]^{\beta-1}}\right) = \left(\alpha\beta + \frac{1-(\beta-1)/\log(x+1)}{[\log(x+1)]^{\beta-1}}\right)e^{\alpha[\log(x+1)]^\beta}.$$

This implies that, uniformly for $n/2 \le x < m \le n$,

$$F_n(x) \asymp \frac{(x+1)}{[\log(x+1)]^{\beta_n-1}}e^{\alpha_n[\log(x+1)]^{\beta_n}}$$

and

$$G_n(x,m) \asymp \left(\frac{(x+1)e^{-\alpha_n[\log(x+1)]^{\beta_n}}}{[\log(x+1)]^{\beta_n-1}} - \frac{(m+1)e^{-\alpha_n[\log(m+1)]^{\beta_n}}}{[\log(m+1)]^{\beta_n-1}}\right).$$

Set $j_n = n[1-(\log n)^{1-\beta_n}]$. The above computation leads to

$$\pi_n([0,j_n]) \asymp \pi_n([j_n,n]), \quad u_n \asymp n^2(\log n)^{1-\beta_n}, \quad v_n \asymp n^2(\log n)^{2-2\beta_n} \asymp w_n.$$

By Theorem 4.8, both $\mathcal{F}_c$ and $\mathcal{F}_\delta$ have a total variation cutoff and, for $\epsilon \in (0,1/2)$ and $\delta \in (0,1)$,

$$\lambda_n \asymp n^{-2}(\log n)^{2\beta_n-2}, \quad T^{(c)}_{n,\mathrm{TV}}(\epsilon) \asymp n^2(\log n)^{1-\beta_n} \asymp T^{(\delta)}_{n,\mathrm{TV}}(\epsilon).$$

Case 4: $f_n(x) = \exp\{\alpha_n[\log(x+1)]^{\beta_n}\}$ with $\sup_n \alpha_n < \infty$ and $\sup_n \beta_n \le 1$. Note that, as a consequence of the mean values theorem, one may choose, for each $0 < a < 1$, a constant $b \in (a,1)$ such that

$$\begin{aligned} 1 < \frac{f_n(n)}{f_n(an)} &= \exp\left\{\alpha_n\left[(\log(n+1))^{\beta_n} - (\log(an+1))^{\beta_n}\right]\right\} \\ &= \exp\left\{\alpha_n\beta_n(1-a)n\frac{(\log(bn+1))^{\beta_n-1}}{bn+1}\right\} \\ &\le \exp\left\{\frac{1-a}{a}\sup_n \alpha_n\right\} < \infty. \end{aligned}$$

This implies that, for $a \in (0,1)$, one may choose a constant $A > 1$ (depending on a) such that

$$\frac{1}{An} \le \pi_n(x) \le \frac{A}{n}, \quad \forall x \ge an,\ n \ge 1.$$

Choosing $j_n = \lfloor n/2 \rfloor$ yields $\pi_n([0,j_n]) \asymp \pi_n([j_n,n])$ and $u_n \asymp v_n \asymp w_n \asymp n^2$. By Theorem 4.8, there is no total variation cutoff for $\mathcal{F}_c$ or $\mathcal{F}_\delta$ and

$$T^{(c)}_{n,\mathrm{TV}}(\epsilon) \asymp T^{(\delta)}_{n,\mathrm{TV}}(\epsilon) \asymp \lambda_n^{-1} \asymp n^2, \quad \forall \epsilon \in (0,1/2),\ \delta \in (0,1).$$

4.3. *Chains with symmetric stationary distributions.* This subsection is dedicated to the study of birth and death chains with symmetric stationary distributions. Let K be an irreducible birth and death chain on $\{0,...,n\}$ with stationary distribution π. Note that π is symmetric at $n/2$, that is, $\pi(n-i) = \pi(i)$ for $0 \le i \le n/2$, if and only if

$$p_i p_{n-i-1} = q_{i+1}q_{n-i}, \quad \forall 0 \le i \le n/2.$$

By the symmetry of π, we will fix $j_n = \lfloor n/2 \rfloor$ when applying Theorem 4.1.

Consider a family of irreducible birth and death chains, $\mathcal{F} = \{(\Omega_n, K_n, \pi_n)|n = 1,2,...\}$ with $\Omega_n = \{0,1,...,n\}$. Let $p_{n,i}, q_{n,i}, r_{n,i}$ be respectively the birth, death and holding rates of K_n and t_n, ℓ_n be constants in Theorem 4.1. Assume that π_n is symmetric at $n/2$. Continuously using the fact $(a+b)/2 \le \max\{a,b\} \le a+b$ for $a \ge 0, b \ge 0$, we obtain

$$t_n \asymp \sum_{k:k\le n/2} \frac{\pi_n([0,k])}{\pi_n(k)\min\{p_{n,k}, q_{n,n-k}\}}$$

and

$$\ell_n \asymp \max_{j:j\le n/2} \sum_{k:j\le k\le n/2} \frac{\pi_n([0,j])}{\pi_n(k)\min\{p_{n,k}, q_{n,n-k}\}}.$$

Theorem 4.1 can be rewritten as follows.

Theorem 4.9. *Let $\mathcal{F} = \{(\Omega_n, K_n, \pi_n)|n = 1,2,...\}$ be a family of irreducible birth and death chains with $\Omega_n = \{0,1,...,n\}$. Let λ_n and $p_{n,i}, q_{n,i}, r_{n,i}$ be the spectral gap and the birth, death and holding rates of K_n. Assume that*

$$p_{n,i}p_{n,n-i-1} = q_{n,i+1}q_{n,n-i}, \quad \forall 0 \le i \le n/2.$$

Then, for $\epsilon \in (0,1/2)$ and $\delta \in (0,1)$,

$$\lambda_n \asymp 1/\ell_n, \quad T^{(c)}_{n,\mathrm{TV}}(\epsilon) \asymp T^{(\delta)}_{n,\mathrm{TV}}(\epsilon) \asymp t_n,$$

where

$$t_n = \sum_{k:k\le n/2} \frac{\pi_n([0,k])}{\pi_n(k)\min\{p_{n,k}, q_{n,n-k}\}}$$

and

$$\ell_n = \max_{j:j\le n/2} \left\{ \pi_n([0,j]) \sum_{k:j\le k\le n/2} \frac{1}{\pi_n(k)\min\{p_{n,k}, q_{n,n-k}\}} \right\}.$$

Moreover, the following are equivalent.

(1) *$\mathcal{F}_c$ has a cutoff in total variation.*
(2) *For $\delta \in (0,1)$, $\mathcal{F}_\delta$ has a cutoff in total variation.*
(3) *$\mathcal{F}_c$ has a precutoff in total variation.*
(4) *For $\delta \in (0,1)$, $\mathcal{F}_\delta$ has a precutoff in total variation.*
(5) *$t_n/\ell_n \to \infty$.*

The next theorem considers a perturbation of birth and death chains which has the same stationary distribution as the original chains. The new chains keep the order of mixing time and spectral gap unchanged.

Theorem 4.10. *Consider the family in Theorem 4.9 and assume that*

$$p_{n,i}p_{n,n-i-1} = q_{n,i+1}q_{n,n-i}, \quad \forall 0 \le i \le n/2.$$

For $n \ge 1$, let $A_n \subset \{0,...,n-1\}$, $c_{n,i} \in [0,1]$ for $i \in A_n$ and $\widetilde{K}_n$ be a birth and death chain on Ω_n with birth and death rates, $\widetilde{p}_{n,i}, \widetilde{q}_{n,i}$, satisfying

$$\begin{cases} \widetilde{p}_{n,i} = c_{n,i}p_{n,i} + (1-c_{n,i})\min\{p_{n,i}, q_{n,n-i}\} & \text{for } i \in A_n, \\ \widetilde{q}_{n,i+1} = q_{n,i+1}\widetilde{p}_{n,i}/p_{n,i} & \text{for } i \in A_n, \\ \widetilde{p}_{n,i} = p_{n,i}, \quad \widetilde{q}_{n,i+1} = q_{n,i+1} & \text{for } i \notin A_n. \end{cases}$$

Let $\lambda_n, \widetilde{\lambda}_n$ and $T_{n,\mathrm{TV}}(\epsilon), \widetilde{T}_{n,\mathrm{TV}}(\epsilon)$ be the spectral gaps and total variation mixing times of $K_n, \widetilde{K}_n$. Then, given $\epsilon \in (0,1/2)$ and $\delta \in (0,1)$,

$$\widetilde{\lambda}_n \asymp \lambda_n, \quad \widetilde{T}^{(c)}_{n,\mathrm{TV}}(\epsilon) \asymp T^{(c)}_{n,\mathrm{TV}}(\epsilon) \asymp \widetilde{T}^{(\delta)}_{n,\mathrm{TV}}(\epsilon) \asymp T^{(\delta)}_{n,\mathrm{TV}}(\epsilon),$$

where the approximation is uniform on the choice of $A_n, c_{n,i}$.

Proof: The approximation of the spectral gap and the total variation mixing time is immediate from Theorem 4.9, whereas the uniformity of the approximation is given by Theorems 3.1, 3.9 and 3.12. □

Example 4.11. For $n \ge 1$, let K_n be a birth and death chain on $\{0, 1, ..., 2n\}$ given by

$$K_n(i, i+1) = K_n(i+1, i) = \begin{cases} 1/2 & \text{for even } i \\ 1/(2n) & \text{for odd } i \end{cases}.$$

By Theorem 4.10, the mixing time and spectral gap of K_n are comparable with those of $\widetilde{K}_n$, where $\widetilde{K}_n(i, i+1) = \widetilde{K}_n(i+1, i) = 1/(2n)$ for $0 \le i < 2n$. Let $\mathcal{F}$ be the family consisting of K_n. By Theorem 4.9, neither $\mathcal{F}_c$ nor $\mathcal{F}_\delta$ has a total variation precutoff and $T^{(c)}_{n,\mathrm{TV}}(\epsilon) \asymp T^{(\delta)}_{n,\mathrm{TV}}(\epsilon) \asymp \lambda_n^{-1} \asymp n^3$ for all $\epsilon \in (0, 1/2)$ and $\delta \in (0, 1)$, which is nontrivial.

Next, we consider simple random walks on finite paths with bottlenecks. For $n \ge 1$, let $k_n \le n$ and $x_{n,1}, ..., x_{n,k_n}$ be positive integers satisfying $1 \le x_{n,i} < x_{n,i+1} \le n$ for $i = 1, ..., k_n - 1$. Let K_n be the birth and death chain on $\{0, 1, ..., n\}$ of which birth, death and holding rates are given by

$$p_{n,i-1} = q_{n,i} = \begin{cases} 1/2 & \text{for } i \notin \{x_{n,1}, ..., x_{n,k_n}\} \\ \epsilon_{n,j} & \text{for } i = x_{n,j},\ 1 \le j \le k_n \end{cases}, \tag{4.6}$$

where $\epsilon_{n,j} \in (0, 1/2]$ for $1 \le j \le k_n$. Clearly, K_n is irreducible and the stationary distribution, say π_n, is uniform on $\{0, 1, ..., n\}$. The following theorem is immediate from Theorems 4.9.

Theorem 4.12. *Let* $\mathcal{F}$ *be a family of birth and death chains given by* (4.6) *and* λ_n *be the spectral gap of* K_n. *For* $n \ge 1$, *set*

$$t_n = n^2 + \sum_{i=1}^{k_n} \frac{\min\{x_{n,i}, n+1-x_{n,i}\}}{\epsilon_{n,i}}$$

and

$$\ell_n = n^2 + \max_{j: j \le n/2} \left\{ \sum_{i: |x_{n,i} - n/2| \le j} \frac{n/2 + 1 - j}{\epsilon_{n,i}} \right\}.$$

Then, for all $\epsilon \in (0, 1/2)$ *and* $\delta \in (0, 1)$,

$$T^{(c)}_{n,\mathrm{TV}}(\epsilon) \asymp T^{(\delta)}_{n,\mathrm{TV}}(\epsilon) \asymp t_n, \quad \lambda_n \asymp 1/\ell_n.$$

Furthermore, the following are equivalent.

(1) $\mathcal{F}_c$ *has a cutoff in total variation.*
(2) *For* $\delta \in (0, 1)$, $\mathcal{F}_\delta$ *has a cutoff in total variation.*
(3) $\mathcal{F}_c$ *has precutoff in total variation.*
(4) *For* $\delta \in (0, 1)$, $\mathcal{F}_\delta$ *has a precutoff in total variation.*
(5) $t_n/\ell_n \to \infty$.

Remark 4.13. Let t_n, ℓ_n be the constants in Theorem 4.12. Then,

$$t_n \asymp n^2 + \sum_{j \in L_n} \frac{x_{n,j}}{\epsilon_{n,j}} + \sum_{j \in R_n} \frac{n+1-x_{n,j}}{\epsilon_{n,j}}$$

and

$$\ell_n \asymp n^2 + \max_{i \in L_n} \sum_{j \in L_n : j \geq i} \frac{x_{n,i}}{\epsilon_{n,j}} + \max_{i \in R_n} \sum_{j \in R_n : j \leq i} \frac{n+1-x_{n,i}}{\epsilon_{n,j}}.$$

where $L_n = \{i : x_{n,i} \leq n/2\}$ and $R_n = \{i : x_{n,i} > n/2\}$.

Theorem 1.4 considers a special case of Theorem 4.12 with $\epsilon_{n,i} = \epsilon_n$ for $1 \leq i \leq k_n$. It is clear from Theorem 1.4 that if k_n is bounded, then no cutoff exists for $\mathcal{F}_c$ or $\mathcal{F}_\delta$. The following example shows a case of cutoffs for the family in Theorem 1.4.

Example 4.14. Let $\mathcal{F}$ be the family in Theorem 1.4, with $k_n = \lfloor n^{1/3} \rfloor - 1$ and

$$x_{n,i} = \left\lfloor \frac{n^{5/6}}{n^{1/3} - i} \right\rfloor, \quad \forall 1 \leq i \leq k_n.$$

Clearly, for n large enough, $x_{n,i} \neq x_{n,j}$ when $i \neq j$. Let a_n, b_n be the constant in Theorem 1.4. It is not hard to show that

$$a_n \asymp n^{5/6} \log n, \quad b_n \asymp n^{5/6}.$$

By Theorem 1.4, $\mathcal{F}_c$ and $\mathcal{F}_\delta$, with $\delta \in (0,1)$, have a total variation cutoff if and only if $\epsilon_n = o(n^{-7/6} \log n)$. Furtheromre, if $\epsilon_n = o(n^{-7/6} \log n)$, then

$$T^{(c)}_{n,\mathrm{TV}}(\epsilon) \asymp \frac{n^{5/6} \log n}{\epsilon_n} \asymp T^{(\delta)}_{n,\mathrm{TV}}(\epsilon), \quad \forall \epsilon, \delta \in (0,1).$$

The following two theorems treat special cases of Theorem 4.12.

Theorem 4.15. *Let $\mathcal{F}$ be a family of birth and death chains satisfying* (4.6)*. Let N be a positive constant. Suppose, for $n \geq 1$, there are constants $J^{(n)}_1, ..., J^{(n)}_N$ and a partition of $\{1, ..., k_n\}$, say $I^{(n)}_1, ..., I^{(n)}_N$, such that, for $1 \leq k \leq N$,*

$$\max_{i \in I^{(n)}_k} \{x_{n,i} \wedge (n+1-x_{n,i})\} \asymp \min_{i \in I^{(n)}_k} \{x_{n,i} \wedge (n+1-x_{n,i})\} \asymp J^{(n)}_k,$$

where $a \wedge b = \min\{a, b\}$. Then, neither $\mathcal{F}_c$ nor $\mathcal{F}_\delta$ has a total variation cutoff. Moreover,

$$T^{(c)}_{n,\mathrm{TV}}(\epsilon) \asymp T^{(\delta)}_{n,\mathrm{TV}}(\epsilon) \asymp \lambda_n^{-1} \asymp t_n, \quad \forall \epsilon \in (0,1/2), \, \delta \in (0,1)$$

where

$$t_n = n^2 + \max_{1 \leq k \leq N} \left\{ J^{(n)}_k \sum_{l \in I^{(n)}_k} \frac{1}{\epsilon_{n,l}} \right\}.$$

The next theorem gives an example that no total variation cutoff exists for $\mathcal{F}_c, \mathcal{F}_\delta$ even when the constant N in Theorem 4.15 tends to infinity.

Theorem 4.16. *Let $\mathcal{F}$ be a family of birth and death chains satisfying* (4.6)*. Suppose that $\min_j \epsilon_{n,j} \asymp \max_j \epsilon_{n,j}$ and $x_{n,i} = \lfloor in/k_n \rfloor$ with $k_n \leq n/2$, then neither $\mathcal{F}_c$ nor $\mathcal{F}_\delta$ has a total variation cutoff, but*

$$T^{(c)}_{n,\mathrm{TV}}(\epsilon) \asymp T^{(\delta)}_{n,\mathrm{TV}}(\epsilon) \asymp \lambda_n^{-1} \asymp \max\{n^2, nk_n/\epsilon_{n,1}\}, \quad \forall \epsilon \in (0,1/2), \, \delta \in (0,1).$$

Remark 4.17. Note that the assumption regarding the birth and death rates in this section can be relaxed using the comparison technique in Diaconis and Saloff-Coste (1993a,b).

Appendix A. Spectral gaps of finite paths

This section is devoted to finding the correct order of spectral gaps of finite paths. Let $G = (V, E)$ be the undirected finite graph with vertex set $V = \{0, 1, 2, ...n\}$ and edge set $E = \{\{i, i+1\} : i = 0, 1, ..., n-1\}$. Given two positive measures π, ν on V, E with $\pi(V) = 1$, the Dirichlet form and variance associated with ν and π are defined by

$$\mathcal{E}_\nu(f, g) := \sum_{i=1}^{n-1} [f(i) - f(i+1)][g(i) - g(i+1)]\nu(i, i+1)$$

and

$$\mathrm{Var}_\pi(f) := \pi(f^2) - \pi(f)^2,$$

where f, g are functions on V. The spectral gap of G with respect to π, ν is defined as

$$\lambda^G_{\pi,\nu} := \min \left\{ \frac{\mathcal{E}_\nu(f, f)}{\mathrm{Var}_\pi(f)} \middle| f \text{ is non-constant} \right\}.$$

To bound the spectral gap, we need the following notations. Let $C_+(i)$ and $C_-(i)$ be constants defined by

$$C_+(i) = \max_{j:j>i} \sum_{k=i+1}^{j} \frac{\pi([j, n])}{\nu(k-1, k)}, \quad C_-(i) = \max_{j:j<i} \sum_{k=j}^{i-1} \frac{\pi([0, j])}{\nu(k, k+1)}, \tag{A.1}$$

where $\max \emptyset := 0$.

Theorem A.1. *Let $G = (V, E)$ be a path on $\{0, 1, ..., n\}$ and π, ν be positive measures on V, E with $\pi(V) = 1$. Referring to (A.1), set $C(m) = \max\{C_+(m), C_-(m)\}$. Then, for $0 \le m \le n$,*

$$\frac{1}{4C(m)} \le \lambda^G_{\pi,\nu} \le \frac{1}{\min\{\pi([0, m]), \pi([m, n])\} C(m)}.$$

In particular, if M is a median of π, that is, $\pi([0, M]) \ge 1/2$ and $\pi([M, n]) \ge 1/2$, then

$$\frac{1}{4C(M)} \le \lambda^G_{\pi,\nu} \le \frac{2}{C(M)}.$$

Remark A.2. Referring to the setting in Theorem A.1, the authors of Chen and Saloff-Coste (2012b) obtained $\lambda^G_{\pi,\nu} \ge 1/C'$, where

$$C' = \min_{0 \le j \le n} \max \left\{ \sum_{k=0}^{j-1} \frac{\pi([0, k])}{\nu(k, k+1)}, \sum_{k=j+1}^{n} \frac{\pi([k, n])}{\nu(k-1, k)} \right\}.$$

Theorem A.1 indicates that $1/C(M)$ is always of the same order as the spectral gap and provides an estimate that can be significantly better than $1/C'$.

The proof of Theorem A.1 is based on the following proposition, which is related to weighted Hardy's inequality on $\{1, ..., n\}$.

Proposition A.3. *Fix $n \ge 1$. Let μ, π be positive measures on $\{1, ..., n\}$ and A be the smallest constant such that*

$$\sum_{i=1}^{n} \left(\sum_{j=1}^{i} g(j) \right)^2 \pi(i) \le A \sum_{i=1}^{n} g^2(i)\mu(i), \quad \forall g \ne \mathbf{0}. \tag{A.2}$$

Then, $B \le A \le 4B$*, where*

$$B = \max_{1 \le i \le n} \left\{ \pi([i,n]) \sum_{j=1}^{i} \frac{1}{\mu(j)} \right\}.$$

Remark A.4. Miclo (1999) discussed the infinity case $\{1,2,...\}$ using the method in Muckenhoupt (1972), which was introduced by Muckenhoupt to study the continuous case $[0,\infty)$. For more information on the weighted Hardy inequality, see Miclo (1999) and the references therein.

Proof of Theorem A.1: We first consider the lower bound of $\lambda^G_{\pi,\nu}$. Let f be any function defined on V and set $f_+ = [f - f(m)]\mathbf{1}_{\{m,...,n\}}$ and $f_- = [f - f(m)]\mathbf{1}_{\{0,...,m\}}$. Then,

$$\frac{\mathcal{E}_\nu(f,f)}{\mathrm{Var}_\pi(f)} \ge \frac{\mathcal{E}_\nu(f,f)}{\pi(f - f(m))^2} = \frac{\mathcal{E}_\nu(f_+,f_+) + \mathcal{E}_\nu(f_-,f_-)}{\pi(f_+^2) + \pi(f_-^2)} \tag{A.3}$$

Set $g(j) = f(m+j) - f(m+j-1)$ for $1 \le j \le n-m$ and $h(i) = f(m-i) - f(m-i+1)$ for $1 \le i \le m$. Note that

$$\mathcal{E}_\nu(f_+,f_+) = \sum_{j=1}^{n-m} g^2(j)\nu(m+j-1,m+j),\ \pi(f_+^2) = \sum_{j=1}^{n-m}\left(\sum_{k=1}^{j} g(k)\right)^2 \pi(m+j),$$

and

$$\mathcal{E}_\nu(f_-,f_-) = \sum_{i=1}^{m} h^2(i)\nu(m-i,m-i+1),\ \pi(f_-^2) = \sum_{j=1}^{m}\left(\sum_{k=1}^{j} h(k)\right)^2 \pi(m-j).$$

By Proposition A.3, the above computation implies that

$$\frac{\mathcal{E}_\nu(f_+,f_+)}{\pi(f_+^2)} \ge \frac{1}{4C_+(m)}, \quad \frac{\mathcal{E}_\nu(f_-,f_-)}{\pi(f_-^2)} \ge \frac{1}{4C_-(m)}.$$

Putting this back to (A.3) gives the desired lower bound.

For the upper bound, we first consider the case $C = C_+(m)$. By Proposition A.3, $C_+(m) \le A$, where A is the smallest constant A such that, for any function ϕ defined on $\{1,2,...,n-m+1\}$,

$$\sum_{j=1}^{n-m}\left(\sum_{k=1}^{j}\phi(k)\right)^2 \pi(m+j) \le A \sum_{j=1}^{n-m} \phi^2(j)\nu(m+j-1,m+j).$$

Let ϕ be a minimizer for A, which must exist, and define ψ by setting

$$\psi(i) = \begin{cases} \phi(1) + \cdots + \phi(i-m) & \text{for } m < i \le n \\ 0 & \text{for } 0 \le i \le m \end{cases}.$$

Clearly, $1/C_+(m) \ge 1/A = \mathcal{E}_\nu(\psi,\psi)/\pi(\psi^2)$. Without loss of generality, we may assume further that ϕ is nonnegative. Note that $\pi(\{\psi = 0\}) \ge \pi([0,m])$. By the Cauchy-Schwartz inequality, this implies $\pi(\psi)^2 \le \pi(\{\psi > 0\})\pi(\psi^2) \le \pi([m+1,n])\pi(\psi^2)$ and, then, $\mathrm{Var}_\pi(\psi) \ge \pi([0,m])\pi(\psi^2)$. This leads to $1/C = 1/C_+(m) \ge \pi([0,m])\lambda^G_{\pi,\nu}$. Similarly, if $C = C_-(m)$, one can prove that $1/C \ge \pi([m,n])\lambda^G_{\pi,\nu}$. This yields the upper bound of the spectral gap. □

Proof of Proposition A.3: The proofs of Theorem A.1 and Proposition A.3 are very similar to those in Miclo (1999). Note that A is attained at functions of the same sign and we assume that g is non-negative. As A is attainable, the minimizer g for A satisfies the following Euler-Lagrange equations.

$$Ag(i)\mu(i) = \sum_{j=i}^{n}(g(1)+\cdots+g(j))\pi(j), \quad \forall 1 \le i \le n. \tag{A.4}$$

This is equivalent to the following system of equations.

$$A[g(i)\mu(i) - g(i+1)\mu(i+1)] = (g(1)+\cdots+g(i))\pi(i), \quad \forall 1 \le i \le n,$$

with the convention that $\mu(n+1) := 0$. Inductively, one can show that $g > 0$. Summing up (A.4) over $\{1,...,\ell\}$ yields

$$\begin{aligned} A\sum_{i=1}^{\ell} g(i) &= \sum_{i=1}^{\ell}\frac{1}{\mu(i)}\sum_{j=i}^{n}(g(1)+\cdots+g(j))\pi(j) \\ &\ge \sum_{i=1}^{\ell}\sum_{j=\ell}^{n}\frac{(g(1)+\cdots+g(j))\pi(j)}{\mu(i)} \\ &\ge \left(\sum_{i=1}^{\ell} g(i)\right)\left(\sum_{i=1}^{\ell}\frac{1}{\mu(i)}\right)\pi([\ell,n]). \end{aligned}$$

This leads to $A \ge B$.

To see the upper bound, we use Miclo's method in Miclo (1999). Set $N(j) = \sum_{i=1}^{j} 1/\mu(i)$. By the Cauchy inequality, the left side of (A.2) is bounded above by

$$\sum_{i=1}^{n}\pi(i)\sum_{j=1}^{i} g^2(j)\mu(j)N^{1/2}(j)\sum_{l=1}^{i}\frac{1}{\mu(l)N^{1/2}(l)}.$$

Note that, for $s>0, t>0$, $t^{1/2}-s^{1/2} \ge (t-s)/(2t^{1/2})$. This implies $2(N^{1/2}(l) - N^{1/2}(l-1)) \ge 1/(\mu(l)N^{1/2}(l))$ with the convention that $N(0) := 0$. Consequently, we have

$$\sum_{l=1}^{i}\frac{1}{\mu(l)N^{1/2}(l)} \le 2N^{1/2}(i) \le \left(\frac{4B}{\pi([i,n])}\right)^{1/2},$$

and, thus,

$$\begin{aligned} \sum_{i=1}^{n}\left(\sum_{j=1}^{i} g(j)\right)^2\pi(i) &\le \sqrt{4B}\sum_{i=1}^{n}\frac{\pi(i)}{\pi([i,n])^{1/2}}\sum_{j=1}^{i} g^2(j)\mu(j)N^{1/2}(j) \\ &\le \sqrt{4B}\sum_{j=1}^{n} g^2(j)\mu(j)N^{1/2}(j)\sum_{i=j}^{n}\frac{\pi(i)}{\pi([i,n])^{1/2}}. \end{aligned}$$

Again, the inequality for s,t implies

$$\sum_{i=j}^{n}\frac{\pi(i)}{\pi([i,n])^{1/2}} \le 2\pi([j,n])^{1/2} \le \frac{\sqrt{4B}}{N^{1/2}(j)}.$$

This gives the desired upper bound. □

Next, we consider a special case. Let π, ν are measures on $V = \{0, 1, ..., n\}, E = \{\{i, i+1\}|0 \le i < n\}$ with $\pi(V) = 1$. Suppose

$$\pi(i) = \pi(n-i), \quad \nu(i, i+1) = \nu(n-i-1, n-i), \quad \forall 0 \le i \le n/2. \tag{A.5}$$

By the symmetry of π and ν, if ψ is a minimizer for $\lambda^G_{\pi,\nu}$ with $\pi(\psi) = 0$, then ψ is either symmetric or anti-symmetric at $n/2$. The former is set aside because ψ is known to be monotonic and this leads to the case $\psi(n-i) = -\psi(i)$ for $0 \le i \le n/2$. If n is even with $n = 2k$, then $\psi(k) = 0$ and this implies

$$\lambda^G_{\pi,\nu} = \inf \left\{ \frac{\sum_{i=1}^k (f(i) - f(i-1))^2 \nu(i-1, i)}{\sum_{i=0}^{k-1} f^2(i)\pi(i)} \middle| f(k) = 0, f \neq \mathbf{0} \right\}.$$

Equivalently, if one sets $g(i) = f(k-i) - f(k-i+1)$ and $\mu(i) = \nu(k-i, k-i+1)$ for $1 \le i \le k$, then $1/\lambda^G_{\pi,\nu}$ is the smallest constant A such that

$$\sum_{i=1}^k \left(\sum_{j=1}^i g(j) \right)^2 \pi(k-i) \le A \sum_{i=1}^k g^2(i)\mu(i), \quad \forall g \neq \mathbf{0}. \tag{A.6}$$

Similarly, if n is odd with $n = 2k-1$, one has

$$\lambda^G_{\pi,\nu} = \min \left\{ \frac{\sum_{i=1}^{k-1} (f(i) - f(i-1))^2 \nu(i-1, i) + 2f^2(k-1)\nu(k-1, k)}{\sum_{i=0}^{k-1} f^2(i)\pi(i)} \middle| f \neq \mathbf{0} \right\},$$

and this leads to (A.6) with $g(1) = f(k-1)$, $\mu(1) = 2\nu(k-1, k)$ and, for $2 \le i \le k$, $g(i) = f(k-i) - f(k-i+1)$ and $\mu(i) = \nu(k-i, k-i+1)$. A direct application of Proposition A.3 implies the following theorem.

Theorem A.5. *Let $G = (V, E)$ be the graph with $V = \{0, 1, ..., n\}$, $E = \{\{i, i+1\}|i = 0 \le i < n\}$ and let π, ν be positive measures on V, E satisfying $\pi(V) = 1$ and* (A.5)*. Set $N = \lceil n/2 \rceil$. Then, $1/(4C) \le \lambda^G_{\pi,\nu} \le 1/C$, where*

$$C = \max_{0 \le i < N} \left\{ \pi([0, i]) \sum_{j=i}^{N-1} \frac{1}{\nu(j, j+1)} \right\} \quad \textit{if } n \textit{ is even,}$$

and

$$C = \max_{0 \le i < N} \left\{ \pi([0, i]) \left(\sum_{j=i}^{N-2} \frac{1}{\nu(j, j+1)} + \frac{1}{2\nu(N-1, N)} \right) \right\} \quad \textit{if } n \textit{ is odd.}$$

Remark A.6. The symmetry of π, ν in Theorems A.5 can be relaxed using the comparison technique.

Acknowledgements

The authors would like to thank an anonymous referee for pointing out several mistakes in a preliminary version.

References

D. Aldous. Random walks on finite groups and rapidly mixing Markov chains. In *Seminar on probability, XVII*, volume 986 of *Lecture Notes in Math.*, pages 243–297. Springer, Berlin (1983). MR770418.

J. Barrera, O. Bertoncini and R. Fernández. Abrupt convergence and escape behavior for birth and death chains. *J. Stat. Phys.* **137** (4), 595–623 (2009). MR2565098.

M. Brown and Y.-S. Shao. Identifying coefficients in the spectral representation for first passage time distributions. *Probab. Engrg. Inform. Sci.* **1**, 69–74 (1987). DOI 10.1017/S0269964800000309.

G.-Y. Chen. *The cutoff phenomenon for finite Markov chains*. Ph.D. thesis, Cornell University (2006).

G.-Y. Chen and L. Saloff-Coste. The cutoff phenomenon for ergodic Markov processes. *Electron. J. Probab.* **13**, no. 3, 26–78 (2008). MR2375599.

G.-Y. Chen and L. Saloff-Coste. Comparison of cutoffs between lazy walks and markovian semigroups (2012a). In preparation.

G.-Y. Chen and L. Saloff-Coste. Spectral computations for birth and death chains (2012b). In preparation.

P. Diaconis. *Group representations in probability and statistics*. Institute of Mathematical Statistics Lecture Notes—Monograph Series, 11. Institute of Mathematical Statistics, Hayward, CA (1988). ISBN 0-940600-14-5. MR964069.

P. Diaconis. The cutoff phenomenon in finite Markov chains. *Proc. Nat. Acad. Sci. U.S.A.* **93** (4), 1659–1664 (1996). MR1374011.

P. Diaconis and L. Saloff-Coste. Comparison techniques for random walk on finite groups. *Ann. Probab.* **21** (4), 2131–2156 (1993a). MR1245303.

P. Diaconis and L. Saloff-Coste. Comparison theorems for reversible Markov chains. *Ann. Appl. Probab.* **3** (3), 696–730 (1993b). MR1233621.

P. Diaconis and L. Saloff-Coste. What do we know about the Metropolis algorithm? *J. Comput. System Sci.* **57** (1), 20–36 (1998). 27th Annual ACM Symposium on the Theory of Computing (STOC'95) (Las Vegas, NV). MR1649805.

P. Diaconis and L. Saloff-Coste. Separation cut-offs for birth and death chains. *Ann. Appl. Probab.* **16** (4), 2098–2122 (2006). MR2288715.

P. Diaconis and D. Stroock. Geometric bounds for eigenvalues of Markov chains. *Ann. Appl. Probab.* **1** (1), 36–61 (1991). MR1097463.

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.

L. Miclo. An example of application of discrete Hardy's inequalities. *Markov Process. Related Fields* **5** (3), 319–330 (1999). MR1710983.

B. Muckenhoupt. Hardy's inequality with weights. *Studia Math.* **44**, 31–38 (1972). Collection of articles honoring the completion by Antoni Zygmund of 50 years of scientific activity, I. MR0311856.

L. Saloff-Coste. Simple examples of the use of Nash inequalities for finite Markov chains. In *Stochastic geometry (Toulouse, 1996)*, volume 80 of *Monogr. Statist. Appl. Probab.*, pages 365–400. Chapman & Hall/CRC, Boca Raton, FL (1999). MR1673142.