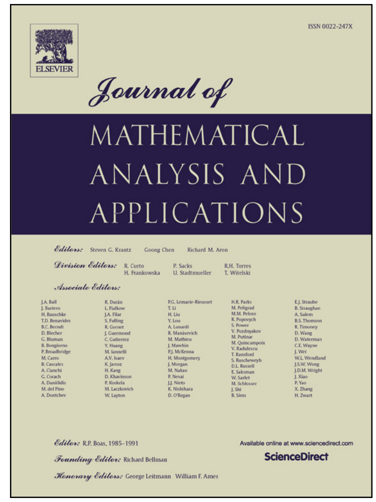

Variational Principles of Hitting Times for Non-reversible Markov Chains

Lu-Jing Huang, Yong-Hua Mao

PII: S0022-247X(18)30700-5
DOI: https://doi.org/10.1016/j.jmaa.2018.08.036
Reference: YJMAA 22493

To appear in: *Journal of Mathematical Analysis and Applications*

Received date: 4 December 2017

Variational Principles of Hitting Times for Non-reversible Markov Chains

Lu-Jing Huang, Yong-Hua Mao*

School of Mathematical Sciences, Beijing Normal University,

Laboratory of Mathematics and Complex Systems, Ministry of Education

Beijing 100875, People's Republic of China

Abstract

We give some new kinds of variational formulas for the first hitting time of non-reversible Markov chain on countable state space. Some comparison theorems are obtained for the non-reversible Markov chain and its corresponding reversible one. As an application, we prove a stronger version of a conjecture in [1, Chapter 9, Conjecture 22].

Keywords and phrases: Variational formula, Markov chain, non-reversible, first hitting time, fundamental matrix.

AMS 2000 Subject classification: 60J10, 60J27

Running head Variational principles of hitting times

1 Introduction

Hitting times play an important role in the theory for Markov processes. Especially for Markov chains, the concept of hitting time is the start point to study recurrence, various ergodicity etc.. Refer to [14] for discrete time Markov chains and to [2] for continuous time Markov chains. In recent years, hitting times are used to derive the convergence rate for a Markov process toward its limiting distribution, cf. [3, 18]. In this paper, we give some new kinds of variational formulas for the first hitting time of a non-reversible Markov chain on a countable state space. The existing variational formulas were mainly for the symmetric Markov processes, see [12] for symmetric diffusion processes, [1] for finite symmetric Markov chains. Very recently, in [8] we give a variational formula for finite asymmetric Markov chains, via that of the capacity for asymmetric Markov chains in [6]. As pointed out in [10, Section 3], hitting times are adopted to be an important criterion to show advantages of non-reversible Markov chains, and there are no general results in this direction.

Let V be a countable state space and $Q = (q_{ij} : i, j \in V)$ be an irreducible, totally stable and conservative Q-matrix. That is, for $i \neq j$, there exist distinct $i_1, ..., i_m \in V$

*Corresponding author: maoyh@bnu.edu.cn

such that $q_{ii_1}q_{i_1i_2}...q_{i_mj} > 0$, and

$$q_i := -q_{ii} < \infty; \quad \sum_{j\neq i} q_{ij} = q_i, \quad i \in V.$$

We also assume Q determines a unique Q-process $X = \{X_t : t \geq 0\}$ and the process admits the stationary distribution π: $\pi > 0$, $\sum_{i\in V} \pi_i = 1$ and $\sum_{i\in V} \pi_i q_{ij} = 0$ for $j \in V$.

Denote by $\mathscr{K}$ the space of functions $f : V \to \mathbb{R}$ with finite support $\mathscr{S}(f) = \{i : f(i) \neq 0\}$. For $f,\ g \in \mathscr{K}$, define

$$Qf(i) = \sum_{j\in V} q_{ij}f_j = \sum_{j\neq i} q_{ij}(f_j - f_i),\ i \in V.$$

Let $L^2(\pi)$ be the space of functions on V which are square-integrable with respect to π, with inner product $\langle\cdot,\cdot\rangle$ defined by

$$\langle f, g\rangle = \sum_{i\in V} \pi_i f_i g_i,$$

and write $\pi(f) = \sum_{i\in V} \pi_i f_i$ for $f \in L^2(\pi)$. For convenience, we denote the generator of chain X acting on domain $\mathscr{D}(Q)$ of $L^2(\pi)$ by Q, and for $f \in L^2(\pi), g \in \mathscr{D}(Q)$, define

$$D_\lambda(f, g) = \langle f, -Qg\rangle + \lambda\pi(fg), \quad \lambda \geq 0,$$

with the natural convention $D := D_0$. Note that in general $D(f, g) \neq D(g, f)$. For any subset A of V, let

$$\tau_A = \inf\{t \geq 0 : X_t \in A\}$$

denote the first hitting time to A of chain X or Q, and

$$\tau_A^+ = \inf\{t > 0 : X_t \in A \text{ and there exists } s \in (0, t) \text{ such that } X_s \neq X_0\}$$

denote the first return time to A. In particular, denote by τ_i and τ_i^+ when A is a singleton $\{i\}$. We also need the notations

$$\mathscr{F}_A := \{f \in \mathscr{K} :\ f|_A = 0 \text{ and } \pi(f) = 1\}, \quad \mathscr{G}_A := \{g \in \mathscr{K} :\ g|_A = 0 \text{ and } \pi(g) = 0\}$$

and $\mathscr{G}_f := \{g :\ \mathscr{S}(g) \subseteq \mathscr{S}(f) \text{ and } \pi(g) = 0\}$ for any $f \in \mathscr{K}$.

Now, we can state our main results on the variational formulas of the first hitting time for non-reversible Markov chains.

Theorem 1.1. *With the notations defined above,*

(1) for any non-trivial subset $A \subseteq V$ and $\lambda > 0$,

$$\frac{\lambda}{1 - \mathbb{E}_\pi[\exp(-\lambda\tau_A)]} = \inf_{f\in\mathscr{F}_A} \sup_{g\in\mathscr{G}_f} D_\lambda(f - g, f + g); \tag{1.1}$$

(2) for any non-trivial subset $A \subseteq V$,

$$1/\mathbb{E}_\pi[\tau_A] = \inf_{f\in\mathscr{F}_A} \sup_{g\in\mathscr{G}_f} D(f - g, f + g). \tag{1.2}$$

In particular, if the generator Q is bounded on $L^2(\pi)$, we can replace $\mathscr{G}_f$ in (1.1) by $\mathscr{G}_A$. If moreover there exists $i \in V$ such that $\mathbb{E}_i[(\tau_i^+)^3] < \infty$, then $\mathscr{G}_f$ in (1.2) also can be replaced by $\mathscr{G}_A$.

Remark 1.2. (1) It should be mentioned here that in an unpublished paper in 1994, Doyle [5] gave a deep insight into the Dirichlet principle for non-reversible Markov chains, even though his proofs are rather conceptual. Indeed, we benefit many ideas from [5].

(2) A result similar to Theorem 1.1 can also be established for discrete time Markov chains, which will be shown as a corollary in Section 5.

(3) When Q is reversible with respect to its stationary distribution π ($\pi_i q_{ij} = \pi_j q_{ji}$: $i, j \in V$), the supremum in (1.2) is attained at $g = 0$. In this case, the variational formula reduces to:

$$1/\mathbb{E}_\pi[\tau_A] = \inf_{f \in \mathscr{F}_A} D(f, f),$$

which was proved in [1, Chapter 3, Proposition 41] for finite V. Similarly, when Q is reversible with respect to π, we obtain a new kind of variational formula:

$$\frac{\lambda}{1 - \mathbb{E}_\pi[\exp(-\lambda\tau_A)]} = \inf_{f \in \mathscr{F}_A} D_\lambda(f, f), \quad \lambda > 0. \tag{1.3}$$

The method to prove Theorem 1.1 is to solve a pair of Poisson equations: one is for Q, another is for Q^*, the dual of Q. Usually $Q^* = (q_{ij}^* : i, j \in V)$ is also called the "time-reversal" of Q, given by

$$q_{ij}^* = \pi_j q_{ji} / \pi_i, \quad i, j \in V.$$

We remark that Q^* is also induced an operator on $L^2(\pi)$ with domain $\mathscr{D}(Q^*)$, which is adjoint with Q. Similarly, we can define, for $f \in \mathscr{D}(Q^*), g \in L^2(\pi)$,

$$D_\lambda^*(f, g) = \langle -Q^* f, g \rangle + \lambda\pi(fg), \quad \lambda \geq 0.$$

It is clear that $D^*(f, g) = D(f, g)$ for $f, g \in \mathscr{D}(Q) \cap \mathscr{D}(Q^*)$, especially for $f, g \in \mathscr{K}$.

In Section 2, we will first prove a variational formula from a pair of general Poisson equations for Q and Q^* when Q, Q^* are bounded, and then we use an approximation argument to derive the variational formulas for general unbounded Q. At first glance, Theorem 1.1 has nothing to do with Q^*, but it really depends heavily on Q^*. This is the main difference between the non-reversible Markov chains and the reversible ones. Roughly speaking, we have to use an "inf-sup" form in the non-reversible case rather than an "inf" form in the reversible case.

Next, we give another type of variational formulas for the first hitting time, which seems difficult to be derived directly from Theorem 1.1. We can see that an "inf" form appears, but both Q and Q^* are involved in a rather complicated form. For this, define the reversible Q-matrix $\overline{Q} = (\overline{q}_{ij} : i, j \in V)$ by

$$\overline{q}_{ij} = \frac{1}{2}(q_{ij} + q_{ij}^*).$$

Theorem 1.3. *With the notations defined in Theorem 1.1,*

(1) for any non-trivial subset $A \subseteq V$ and $\lambda > 0$,

$$\frac{\lambda}{1 - \mathbb{E}_\pi[\exp(-\lambda\tau_A)]} = \inf_{f \in \mathscr{F}_A} \sup_{\mathscr{S}(g) \subseteq \mathscr{S}(f)} \{2D_\lambda(f, g) - D_\lambda(g, g)\}; \tag{1.4}$$

(2) for any non-trivial subset $A \subseteq V$,

$$1/\mathbb{E}_\pi[\tau_A] = \inf_{f\in\mathscr{F}_A} \sup_{\mathscr{S}(g)\subseteq\mathscr{S}(f)} \{2D(f,g) - D(g,g)\}. \tag{1.5}$$

The new variational formulas in Theorem 1.1 or Theorem 1.3 enable ones to compare the hitting times between a Q-process and its corresponding reversible Q-process $\overline{Q}$. For this, for any subset A of V, denote by τ_A^*, $\overline{\tau}_A$ the first hitting time to A of chains Q^* and $\overline{Q}$, respectively. Let $T_0 = \sum_{j\in V} \pi_j \mathbb{E}_i[\tau_j]$ be the average hitting time of chain Q. It is well known that T_0 is independent of the state i, see Corollary 13 in [1, Chapter 2]. And denote by $\overline{T}_0$ the average hitting time of chain $\overline{Q}$. As a corollary of Theorem 1.1, we compare the first hitting time between chains Q and $\overline{Q}$ in following.

Corollary 1.4. *For any subset A of V, there exists the Laplace transform order between τ_A and $\overline{\tau}_A$. Namely,*

$$\mathbb{E}_\pi[\exp(-\lambda\overline{\tau}_A)] \le \mathbb{E}_\pi[\exp(-\lambda\tau_A)], \quad \lambda > 0.$$

In particular, $\mathbb{E}_\pi[\overline{\tau}_A] \ge \mathbb{E}_\pi[\tau_A]$ and $\overline{T}_0 \ge T_0$.

As we did in [8], we can get a non-reversible Q-matrix by adding a vorticity matrix to a reversible Q-matrix in Section 3.2 below. And we will present some comparison theorems for them and give a proof of Corollary 1.4. Moreover, in [8], we gave an affirmative answer to a conjecture in [1, Chapter 9, Conjecture 22], by using the new variational formulas for discrete time Markov chains in Section 5, we can obtain a stronger version for this conjecture. For this, we assume that V is finite and P is an irreducible probability transition matrix on V with the stationary distribution π. Define

$$z_{ij} = \sum_{n=0}^{\infty} [p_{ij}^{(n)} - \pi_j]$$

as the fundamental matrix of P. In fact, the fundamental matrix Z can be view as the inverse of operator $I - P$ on the linear space of functions $f : V \to \mathbb{R}$ satisfying $\pi(f) = 0$. In [1], Aldous and Fill conjectured that

$$\mathrm{trace}[Z^2(P^* - P)] \ge 0. \tag{1.6}$$

They also proved that (1.6) implies that the average hitting time for a Markov chain is smaller than that of its reversibilizations, see Corollary 24 in [1, Chapter 9].

Corollary 1.5. *Assume that X is an irreducible Markov chain on finite state space V with probability transition matrix P and stationary distribution π. Then $[Z(P^* - P)Z]_{ii} \ge 0$ for $i \in V$. Consequently, trace$[Z^2(P^* - P)] \ge 0$.*

The rest of the paper is organized as follows. In Section 2 and 4, we prove the main results, Theorem 1.1 and 1.3, and in Section 3, we present several comparison theorems. In Section 5, discrete time Markov chains are considered, and Corollary 1.5 is proven.

2 Proof of Theorem 1.1

Fix a non-trivial subset A of V and constants $\lambda,\ c \geq 0$. Consider the Poisson equation for chain Q:

$$\begin{cases} (\lambda - Q)x(i) = c, & i \in A^c; \\ x_i = 0, & i \in A. \end{cases} \tag{2.1}$$

If its solution exists, then we denote it by φ. Similarly, write φ^* be the solution of Poisson equation (2.1) for chain Q^* if it exists. We can obtain a variational formula for this pair of Poisson equations, from which we derive the proof of Theorem 1.1.

Theorem 2.1. *Assume that the generator Q is bounded. Suppose that for given a non-trivial subset $A \subseteq V$ and $\lambda,\ c \geq 0$, Poisson equations (2.1) for chains Q and Q^* have solutions $\varphi,\ \varphi^*$, respectively, such that $\varphi,\ \varphi^* \in L^2(\pi)$ and $\pi(\varphi) = \pi(\varphi^*) = 1$. Then*

$$D_\lambda(\varphi^*, \varphi) = \inf_{f \in \mathscr{F}_A} \sup_{g \in \mathscr{G}_A} D_\lambda(f - g, f + g).$$

Proof. Since $\mathscr{K}$ is dense in $L^2(\pi)$, it is sufficient to prove that

$$D_\lambda(\varphi^*, \varphi) = \inf_{f \in \mathscr{F}'_A} \sup_{g \in \mathscr{G}'_A} D_\lambda(f - g, f + g), \tag{2.2}$$

where $\mathscr{F}'_A = \{f \in L^2(\pi) : f|_A = 0 \text{ and } \pi(f) = 1\}$ and $\mathscr{G}'_A = \{g \in L^2(\pi) : g|_A = 0 \text{ and } \pi(g) = 0\}$.

For this, we need two inequalities:

$$D_\lambda(\varphi^* - g, \varphi + g) \leq D_\lambda(\varphi^*, \varphi); \tag{2.3}$$

$$D_\lambda(\varphi^* + g, \varphi + g) \geq D_\lambda(\varphi^*, \varphi), \tag{2.4}$$

for all $g \in \mathscr{G}'_A$. Indeed, from the definitions of φ and φ^* it follows that for $g \in \mathscr{G}'_A$

$$D_\lambda(g, \varphi) = \sum_{i \in A^c} \pi_i g_i (\lambda - Q)\varphi(i) = c\pi(g) = 0;$$

$$D_\lambda(\varphi^*, g) = \langle (\lambda - Q)^* \varphi^*, g \rangle = c\pi(g) = 0.$$

Thus, combine them with the fact that $D_\lambda(g, g) \geq 0$ for all $g \in \mathscr{G}'_A$, one finds

$$D_\lambda(\varphi^* - g, \varphi + g) = D_\lambda(\varphi^*, \varphi) - D_\lambda(g, g) \leq D_\lambda(\varphi^*, \varphi)$$

and

$$D_\lambda(\varphi^* + g, \varphi + g) = D_\lambda(\varphi^*, \varphi) + D_\lambda(g, g) \geq D_\lambda(\varphi^*, \varphi).$$

Now we return to prove (2.2). Write $\overline{\varphi} = (\varphi + \varphi^*)/2$ and $\widehat{\varphi} = (\varphi - \varphi^*)/2$. It is easy to check that $\overline{\varphi} \in \mathscr{F}'_A$ and $\widehat{\varphi} \in \mathscr{G}'_A$. Note that $Q\phi^*$ and $Q^*\phi$ are in $L^2(\pi)$ since Q (Q^*) is bounded.

For $g \in \mathscr{G}'_A$, define $g_1 = g - \widehat{\varphi}$, and then $g_1 \in \mathscr{G}'_A$ is clear from the boundary conditions of $\widehat{\varphi}$ and g. So if we take $f = \overline{\varphi}$ in variational formula (2.2), then (2.3) implies that

$$\begin{aligned} D_\lambda(\overline{\varphi} - g, \overline{\varphi} + g) &= D_\lambda(\overline{\varphi} - \widehat{\varphi} - g_1, \overline{\varphi} + \widehat{\varphi} + g_1) \\ &= D_\lambda(\varphi^* - g_1, \varphi + g_1) \\ &\leq D_\lambda(\varphi^*, \varphi). \end{aligned}$$

Thus,

$$D_\lambda(\varphi^*, \varphi) \geq \inf_{f \in \mathscr{F}'_A} \sup_{g \in \mathscr{G}'_A} D_\lambda(f-g, f+g). \tag{2.5}$$

Conversely, for $f \in \mathscr{F}'_A$, define $f_1 = f - \overline{\varphi}$ and then $f_1 \in \mathscr{G}'_A$ similarly. By setting the test function $g = \widehat{\varphi}$ in (2.2), from (2.4) it follows that

$$D_\lambda(f - \widehat{\varphi}, f + \widehat{\varphi}) = D_\lambda(\varphi^* + f_1, \varphi + f_1) \geq D_\lambda(\varphi^*, \varphi).$$

So we obtain that

$$D_\lambda(\varphi^*, \varphi) \leq \inf_{f \in \mathscr{F}'_A} \sup_{g \in \mathscr{G}'_A} D_\lambda(f-g, f+g). \tag{2.6}$$

Combining (2.5) and (2.6) gives the desired result. □

To prove Theorem 1.1, we specify the general Poisson equation (2.1) for the first hitting time.

Lemma 2.2. *(1) For any non-trivial subset A and $\lambda > 0$, let*

$$\varphi_i = \frac{1 - \mathbb{E}_i[\exp(-\lambda\tau_A)]}{1 - \mathbb{E}_\pi[\exp(-\lambda\tau_A)]}, \quad i \in V.$$

Then $\varphi = (\varphi_i : i \in V)$ is a solution of the Poisson equation:

$$\begin{cases} (\lambda - Q)x(i) = \lambda/(1 - \mathbb{E}_\pi[\exp(-\lambda\tau_A)]), & i \in A^c; \\ x_i = 0, & i \in A. \end{cases} \tag{2.7}$$

(2) For any non-trivial subset A, let $\psi_i = \mathbb{E}_i[\tau_A]/\mathbb{E}_\pi[\tau_A]$, $i \in V$. Then $\psi = (\psi_i : i \in V)$ is a solution of the Poisson equation:

$$\begin{cases} -Qx(i) = 1/\mathbb{E}_\pi[\tau_A], & i \in A^c; \\ x_i = 0, & i \in A. \end{cases} \tag{2.8}$$

Proof. From [7, Chapter 9], we know that $\tilde{\varphi} := (\mathbb{E}_i[\exp(-\lambda\tau_A)] : i \in V)$ is the minimal non-negative solution of the equation

$$\begin{cases} x_i = \sum_{j \neq i} \frac{q_{ij}}{\lambda + q_i} x_j, & i \in A^c; \\ x_i = 1, & i \in A. \end{cases}$$

Multiply both sides of the first equality by $\lambda + q_i$ and rearrange the terms, we can see that $\tilde{\varphi}$ is a solution of the equation

$$\begin{cases} (\lambda - Q)x(i) = 0, & i \in A^c; \\ x_i = 1, & i \in A. \end{cases}$$

Since $Q1 = 0$, it is clear that $1 - \tilde{\varphi}$ is a solution of the equation

$$\begin{cases} (\lambda - Q)x(i) = \lambda, & i \in A^c; \\ x_i = 0, & i \in A. \end{cases} \tag{2.9}$$

Thus we obtain (1) by multiplying by $1/(1 - \mathbb{E}_\pi[\exp(-\lambda\tau_A)])$ both sides of equation (2.9). By using a similar transform, we can check that (2) holds. □

We also need following lemma, which says that τ_A and τ_A^* are identically distributed when started from the stationary distribution π.

Lemma 2.3. *For any subset A of V, the first hitting times τ_A and τ_A^* have the same Laplace transform under $\mathbb{P}_\pi$. Namely,*

$$\mathbb{E}_\pi[\exp(-\lambda\tau_A)] = \mathbb{E}_\pi[\exp(-\lambda\tau_A^*)], \quad \lambda > 0.$$

In particular, $\mathbb{E}_\pi[\tau_A] = \mathbb{E}_\pi[\tau_A^*]$.

Proof. Fix $\lambda > 0$. It is easy to check, as in the proof of Lemma 2.2, that $\varphi^* = ((1 - \mathbb{E}_i[\exp(-\lambda\tau_A^*)])/(1 - \mathbb{E}_\pi[\exp(-\lambda\tau_A^*)]) : i \in V)$ is a solution of Poisson equation (2.7) for chain Q^*. So we have

$$D_\lambda(\varphi^*, \varphi) = \langle \varphi^*, (\lambda - Q)\varphi\rangle = \sum_{i \in A^c} \pi_i \varphi_i^*(\lambda - Q)\varphi(i) = \frac{\lambda}{1 - \mathbb{E}_\pi[\exp(-\lambda\tau_A)]}.$$

But note that on the left we have

$$D_\lambda(\varphi^*, \varphi) = \langle (\lambda - Q)^*\varphi^*, \varphi\rangle = \sum_{i \in A^c} \pi_i \varphi_i(\lambda - Q)^*\varphi^*(i) = \frac{\lambda}{1 - \mathbb{E}_\pi[\exp(-\lambda\tau_A^*)]}.$$

Thus the proof be complete by combining above two equalities. □

Now we are ready to prove Theorem 1.1.

Proof of Theorem 1.1. Fix a non-trivial subset A and $\lambda > 0$.

(a) We assume firstly that Q is bounded. Since φ which defined as at Lemma 2.2 is square-integrable, by taking $c = \lambda/(1 - \mathbb{E}_\pi[\exp(-\lambda\tau_A)])$ on Poisson equation (2.1), we obtain

$$\frac{\lambda}{1 - \mathbb{E}_\pi[\exp(-\lambda\tau_A)]} = \inf_{f \in \mathscr{F}_A} \sup_{g \in \mathscr{G}_A} D_\lambda(f - g, f + g) \tag{2.10}$$

from Theorem 2.1 and Lemma 2.2-2.3.

If moreover $\mathbb{E}_i[(\tau_i^+)^3] < \infty$ for some $i \in V$. Take and fix $a \in A$, we have $\mathbb{E}_a[(\tau_a^+)^3] < \infty$ by irreducibility. For $\psi_i^{'} = \mathbb{E}_i[\tau_a]$,

$$\sum_{i \in V} \pi_i \psi_i^{'2} \le \sum_{i \in V} \pi_i \mathbb{E}_i[\tau_a^2] = \mathbb{E}_\pi[(\tau_a^+)^2] - \pi_a \mathbb{E}_a[(\tau_a^+)^2].$$

But [4, Lemma 3.4] shows that

$$\mathbb{E}_\pi[(\tau_a^+)^2] < \infty \Leftrightarrow \mathbb{E}_a[(\tau_a^+)^3] < \infty,$$

so $\psi^{'} \in L^2(\pi)$. Furthermore, $\psi \in L^2(\pi)$ follows immediately by $\psi \le \psi^{'}$. Similarly $\psi^* \in L^2(\pi)$. Now by taking $\lambda = 0$ and $c = 1/\mathbb{E}_\pi[\tau_A]$ in Poisson equation (2.1), from Theorem 2.1 and Lemma 2.2-2.3, we get

$$1/\mathbb{E}_\pi[\tau_A] = \inf_{f \in \mathscr{F}_A} \sup_{g \in \mathscr{G}_A} D(f - g, f + g).$$

(b) To prove (1) and (2) for general (unbounded) Q, we use an approximation argument. But since the proofs of them are quite similar, we shall be brief and only prove (1) here.

When A is a non-trivial subset with finite A^c, we will define a collapsed chain on $V_a = A^c \cup \{a\}$ $(a \notin V)$. Let $Q^A = (q_{ij}^A :\ i, j \in V_a)$ be defined by

$$q_{ij}^A = q_{ij}, \quad i, j \in A^c; \qquad q_{ia}^A = \sum_{j \in A} q_{ij}, \quad i \in A^c;$$
$$q_{ai}^A = \frac{1}{\pi(A)} \sum_{j \in A} \pi_j q_{ji}, \quad i \in A^c; \qquad q_{aa}^A = -\sum_{i \in A^c} q_{ai}.$$

The collapsed chain admits stationary distribution π^A given by $\pi_i^A = \pi_i$ for $i \in A^c$ and $\pi_a^A = \pi(A)$. Then τ_A has the same distribution as that of τ_a^A the hitting time to the singleton $\{a\}$ of Q^A-process, as indicated in [1, Section 7.2, Chapter 1]. Since Q^A is a bounded operator and φ, ψ defined in Lemma 2.2 are square-integrable and so are φ^*, ψ^*. As proved above in (a), variational formula (2.10) holds for τ_a^A, and then for τ_A.

(c) For $f \in \mathscr{F}_A$, we set $B^c := \mathscr{S}(f)$. So B^c is finite and $A \subseteq B$. Applying (2.10) to subset B, we can see that

$$\begin{aligned}\frac{\lambda}{1 - \mathbb{E}_\pi[\exp(-\lambda \tau_B)]} &= \inf_{h \in \mathscr{F}_B} \sup_{g \in \mathscr{G}_B} D_\lambda(h - g, h + g) \\ &\le \sup_{g \in \mathscr{G}_f} D_\lambda(f - g, f + g).\end{aligned}$$

And the fact $\tau_B \le \tau_A$ gives

$$\frac{\lambda}{1 - \mathbb{E}_\pi[\exp(-\lambda \tau_A)]} \le \sup_{g \in \mathscr{G}_f} D_\lambda(f - g, f + g). \tag{2.11}$$

Conversely, without loss of generality, let $V = \{1, 2, ...\}$ and set $E_n := \{n, n+1, ...\}$ and $A_n := E_n \cup A$. Then A_n^c is finite. Applying (2.10) to subset A_n again gives that

$$\begin{aligned}\frac{\lambda}{1 - \mathbb{E}_\pi[\exp(-\lambda \tau_{A_n})]} &\ge \inf_{f \in \mathscr{F}_{A_n}} \sup_{g \in \mathscr{G}_f} D_\lambda(f - g, f + g) \\ &\ge \inf_{f \in \mathscr{F}_A} \sup_{g \in \mathscr{G}_f} D_\lambda(f - g, f + g).\end{aligned}$$

And we get

$$\frac{\lambda}{1 - \mathbb{E}_\pi[\exp(-\lambda \tau_A)]} \ge \inf_{f \in \mathscr{F}_A} \sup_{g \in \mathscr{G}_f} D_\lambda(f - g, f + g), \tag{2.12}$$

by passing to the limit as $n \to \infty$. Thus combining (2.11) and (2.12) gives the desired result. $\square$

3 Comparison theorems

In this section, we will present some comparison theorems for the first hitting times. There are the Monotonicity law and comparison theorems between reversible and non-reversible Markov chains. We also give a proof of Corollary 1.4 in this section.

3.1 Monotonicity law

Recall that Q is an irreducible and regular Q-matrix, which determines uniquely a Q-process. In order to distinguish hitting times between different chains, we denote $\tau_{\cdot}(Q)$ as the first hitting time of chain Q in this subsection. For two Q-matrices Q and K, we say that $K \le Q$, if $k_{ij} \le q_{ij}$ for any $i \neq j$.

Theorem 3.1. *Assume that K, Q be irreducible and regular Q-matrices on V with same stationary distribution π. If $K \le Q$ and K is reversible with respect to π, then for any subset A,*

$$\mathbb{E}_\pi[\exp(-\lambda\tau_A(K))] \le \mathbb{E}_\pi[\exp(-\lambda\tau_A(Q))], \quad \lambda > 0. \tag{3.1}$$

In particular, $\mathbb{E}_\pi[\tau_A(K)] \ge \mathbb{E}_\pi[\tau_A(Q)]$.

Proof. Fix a subset A and $\lambda > 0$. If $A = \emptyset$ or V, it is easy to see that inequality (3.1) holds. Now assume that A is a non-trivial subset. Since for f with finite support,

$$\langle f, (\lambda - Q)f\rangle = \frac{1}{2}\sum_{i,j}\pi_i(\lambda + q_{ij})(f_j - f_i)^2,$$

the assumption $K \le Q$ implies

$$\langle f, (\lambda - Q)f\rangle \ge \langle f, (\lambda - K)f\rangle.$$

By Theorem 1.1 and variational formula (1.3) for the reversible case, and taking $g = 0$ in the following supremum,

$$\begin{aligned}
\frac{\lambda}{1 - \mathbb{E}_\pi[\exp(-\lambda\tau_A(Q))]} &= \inf_{f\in\mathscr{F}_A}\sup_{g\in\mathscr{G}_f}\langle f - g, (\lambda - Q)(f + g)\rangle \\
&\ge \inf_{f\in\mathscr{F}_A}\langle f, (\lambda - Q)f\rangle \\
&\ge \inf_{f\in\mathscr{F}_A}\langle f, (\lambda - K)f\rangle \\
&= \frac{\lambda}{1 - \mathbb{E}_\pi[\exp(-\lambda\tau_A(K))]}.
\end{aligned}$$

□

Remark 3.2. (1) Let ξ and η be two non-negative random variables. We say that ξ is smaller than η in the Laplace transform order(denoted by $\xi \le_{Lt} \eta$) if

$$\mathbb{E}[\exp(-\lambda\xi)] \ge \mathbb{E}[\exp(-\lambda\eta)], \quad \text{for all } \lambda > 0.$$

According to Theorem 3.1, we can see that $\tau_A(Q) \le_{Lt} \tau_A(K)$ for any subset A. Furthermore, More properties of the Laplace transform order can be found in [16].

(2) Correspondingly, Monotonicity law for the asymptotic variance and the capacity can be found in [5, 15, 17] and references therein.

3.2 Comparison theorems between reversible and non-reversible Markov chains

As another important application of Theorem 1.1, we construct non-reversible Markov chains by adding anti-symmetric perturbations to reversible Markov chains and present some properties of their first hitting times.

Let us recall some notations. Consider an irreducible Markov chain $X = \{X_t : t \geq 0\}$ on V with generator $K = (k_{ij} : i, j \in V)$ and stationary distribution π. Assume that X is reversible with respect to π. Recall that matrix Γ is a vorticity matrix if it satisfies $\Gamma 1 = 0$ and $\Gamma^T = -\Gamma$, where Γ^T is the transpose of matrix Γ. Assume that there exists a nonzero vorticity matrix Γ, such that

$$\Gamma_{ij} > -\pi_i k_{ij}, \ i \neq j \in V. \tag{3.2}$$

Define

$$Q_\Gamma = K + \mathrm{diag}(\pi)^{-1}\Gamma,$$

where $\mathrm{diag}(\pi)$ is defined by the diagonal matrix for vector π and $\mathrm{diag}(\pi)^{-1}$ is its inverse. Then Q_Γ is also irreducible and it has the same stationary distribution π. Assume that Y is the Markov chain that is determined by Q_Γ. Denote by $\tau_\cdot$ and $\tau_\cdot(\Gamma)$ the first hitting times of chains X and Y, respectively.

The following theorem shows that the first hitting times of chains X and Y exist the Laplace transform order.

Theorem 3.3. *Let Markov chains X and Y be defined as above. Then for any subset $A \subseteq V$,*

$$\tau_A(\Gamma) \leq_{Lt} \tau_A \quad \text{under } \mathbb{P}_\pi. \tag{3.3}$$

In particular, $\mathbb{E}_\pi[\tau_A] \geq \mathbb{E}_\pi[\tau_A(\Gamma)]$.

Proof. When $A = \emptyset$ or V, it is obvious that $\mathbb{E}_\pi[\exp(-\lambda\tau_A)] = \mathbb{E}_\pi[\exp(-\lambda\tau_A(\Gamma))]$. We just need to prove that the Laplace transform order (3.3) is also true for non-trivial subsets. Note that $\langle f, (\lambda - Q_\Gamma)f\rangle = \langle f, (\lambda - K)f\rangle$ for $f \in \mathscr{K}$. From Theorem 1.1 and (1.3) it follows that, for $\lambda > 0$,

$$\begin{aligned}
\frac{\lambda}{1 - \mathbb{E}_\pi[\exp(-\lambda\tau_A(\Gamma))]} &= \inf_{f \in \mathscr{F}_A} \sup_{g \in \mathscr{G}_f} \langle f - g, (\lambda - Q_\Gamma)(f + g)\rangle \\
&\geq \inf_{f \in \mathscr{F}_A} \langle f, (\lambda - Q_\Gamma)f\rangle \\
&= \inf_{f \in \mathscr{F}_A} \langle f, (\lambda - K)f\rangle \\
&= \frac{\lambda}{1 - \mathbb{E}_\pi[\exp(-\lambda\tau_A)]}.
\end{aligned}$$

□

Remark 3.4. Denote the average hitting time of chains X and Y, by T_0 and $T_0(\Gamma)$, respectively. Theorem 3.3 implies that $T_0 \geq T_0(\Gamma)$. In fact, the average hitting time has closed connection with strong ergodicity and asymptotic variance, see e.g. [9, 13].

According to the properties of the Laplace transform order, we present an interesting result for the first hitting time as follows.

Corollary 3.5. *Suppose that $f : [0, +\infty) \to \mathbb{R}$ is a completely monotone function, i.e., its derivatives $f^{(n)}$ exist and satisfy $(-1)^n f^{(n)} \geq 0, n = 0, 1, 2, \ldots.$ Then*

$$\mathbb{E}_\pi[f(\tau_A(\Gamma))] \geq \mathbb{E}_\pi[f(\tau_A)]$$

provided the expectations exist. Especially, if we take $f(x) = (1+x)^{-n}$, then

$$\mathbb{E}_\pi[(1+\tau_A(\Gamma))^{-n}] \geq \mathbb{E}_\pi[(1+\tau_A)^{-n}], \quad n \geq 1.$$

Proof. Combine Theorem 5.A.7 in [16] and Theorem 3.3, we can finish the proof easily. □

Finally, we give a proof of Corollary 1.4.

Proof of Corollary 1.4. Define $\Gamma = \mathrm{diag}(\pi)(Q - Q^*)/2$. It is easy to check that Γ is a vorticity matrix and

$$Q = \overline{Q} + \mathrm{diag}(\pi)^{-1}\Gamma.$$

Then the assertion follows from Theorem 3.3. □

3.3 Parameterization

Based on Section 3.2, we will introduce a parameter to control the anti-symmetric perturbations and investigate the properties of the hitting times that viewed as functions of the parameter.

Let X be defined as at Section 3.2. Assume that there exists a nonzero vorticity matrix Γ satisfying (3.2). Define a family of Q-matrices on V by

$$Q_\alpha = K + \alpha \mathrm{diag}(\pi)^{-1}\Gamma, \quad -1 \leq \alpha \leq 1. \tag{3.4}$$

Obviously, all of them have the same stationary distribution π. For any subset $A \subseteq V$, define the first hitting time to A of chain Q_α by $\tau_A(\alpha)$ and the average hitting time by $T_0(\alpha)$. The following result shows that as functions of variable α, the hitting times of Q_α have monotone and symmetry properties.

Theorem 3.6. *For the reversible Markov chain X, let Q_α be defined as at (3.4) with nonzero vorticity matrix Γ satisfying (3.2). For any subset $A \subseteq V$ and $\lambda > 0$, denote by $R(\alpha)$ either $\mathbb{E}_\pi[\tau_A(\alpha)]$ or $T_0(\alpha)$. Then*

(1) $\mathbb{E}_\pi[\exp(-\lambda\tau_A(\alpha))] = \mathbb{E}_\pi[\exp(-\lambda\tau_A(-\alpha))]$ and $R(-\alpha) = R(\alpha)$ for any $\alpha \in [-1, 1]$;

(2) $\mathbb{E}_\pi[\exp(-\lambda\tau_A(\alpha))]$ is non-increasing for $\alpha \in [-1, 0]$ and $R(\alpha)$ is non-decreasing for $\alpha \in [-1, 0]$.

Proof. (1) As we know that the dual matrix $(\lambda - Q_\alpha)^* = \lambda - Q_{-\alpha}$, so we get the first assertion by Lemma 2.3.

(2) The second assertion is obvious when $A = \emptyset$ or V. In following, assume that A is a non-trivial subset. It suffices to prove that for any $\alpha_1, \alpha_2 \in [-1, 0]$ with $\alpha_1 < \alpha_2$ and $f \in \mathscr{F}_A$,

$$\sup_{g \in \mathscr{G}_f} \langle f - g, (\lambda - Q_{\alpha_1})(f+g) \rangle \geq \sup_{g \in \mathscr{G}_f} \langle f - g, (\lambda - Q_{\alpha_2})(f+g) \rangle, \tag{3.5}$$

which implies that $\mathbb{E}_\pi[\exp(-\lambda\tau_A(\alpha))]$ is non-increasing for $\alpha \in [-1, 0]$ by variational formula (1.1).

Now we prove (3.5). Define $N = \mathrm{diag}(\pi)^{-1}\Gamma$. Since N is anti-symmetric with respect to π, we have

$$\begin{aligned}&\langle f-g,(\lambda-Q_\alpha)(f+g)\rangle\\ =&\langle f-g,(\lambda-K)(f+g)\rangle-\alpha\langle f-g,N(f+g)\rangle\\ =&\langle f-g,(\lambda-K)(f+g)\rangle+2\alpha\langle Nf,g\rangle.\end{aligned}\tag{3.6}$$

And also

$$\langle f-(-g),(\lambda-Q_\alpha)(f+(-g))\rangle=\langle f-g,(\lambda-K)(f+g)\rangle-2\alpha\langle Nf,g\rangle$$

by symmetry of K. Thus,

$$\langle f-g,(\lambda-Q_\alpha)(f+g)\rangle<\langle f-(-g),(\lambda-Q_\alpha)(f+(-g))\rangle$$

for $\alpha\in[-1,0]$ and function $g\in\mathscr{G}_f$ with $\langle Nf,g\rangle>0$. This means that the supremum in (3.5) can not be attained by those g such that $\langle Nf,g\rangle>0$, so it is sufficient to consider the function $g\in\mathscr{G}_f$ such that $\langle Nf,g\rangle\le 0$. Apply (3.6) to those g to get that

$$\langle f-g,(\lambda-Q_{\alpha_1})(f+g)\rangle\ge\langle f-g,(\lambda-Q_{\alpha_2})(f+g)\rangle,$$

for $-1\le\alpha_1\le\alpha_2\le 0$. Thus (3.5) holds.

Finally, we can also prove the monotone properties of $R(\alpha)$ in a similar way via variational formula (1.2). □

Remark 3.7. In fact, we proved that the commute time have monotone and symmetry properties, and so as the average hitting time on discrete time Markov chains in [8].

4 Proof of Theorem 1.3

In this section, we will firstly prove a new version of variational formulas (1.4)-(1.5) for the subset A with finite A^c, and then use an approximation argument to complete the proof of Theorem 1.3. But for the reason that the approximation is very similar to that in the last part of the proof of Theorem 1.1, we choose to omit it after we prove the following result for A with finite A^c.

We introduce some notations firstly. Fix a non-trivial subset A with finite A^c. In what follows, denote matrix $R=(r_{ij}:i,j\in V)$ restricted on A^c by R^a, which is defined by

$$r^a_{ij}=r_{ij},\quad i,j\in A^c,$$

and define π^a the measure on A^c by $\pi^a_i=\pi_i$ for $i\in A^c$.

Recall that $\overline{Q}$ is symmetric with respect to π, so $\overline{Q}^a$ is symmetric with respect to π^a . For $\lambda>0$, $(\lambda-\overline{Q}^a)^{-1}$ is the resolution for $\overline{Q}^a$-process $(\overline{p}^a_{ij}(t):i,j\in A^c)$, that is,

$$(\lambda-\overline{Q}^a)^{-1}=\int_0^\infty e^{-\lambda t}\overline{p}^a_{ij}(t)dt,$$

so $(\lambda-\overline{Q}^a)^{-1}$ is well-defined and self-adjoint on $L^2(A^c,\pi^a)$. Let

$$\mathscr{A}_\lambda=(\lambda-Q^a)(\lambda-\overline{Q}^a)^{-1}(\lambda-Q^a)^*.$$

Then $\mathscr{A}_\lambda$ is a self-adjoint operator on $L^2(A^c,\pi^a)$. Since Q is irreducible, we can see that Q^a must be not conservative. Thus $\mathscr{A}:=(-Q^a)(-\overline{Q}^a)^{-1}(-Q^a)^*$ is also well defined, and self-adjoint on $L^2(A^c,\pi^a)$.

Proposition 4.1. *With the notations in Theorem 1.1,*

(1) for non-trivial subset A with finite A^c and $\lambda > 0$,

$$\frac{\lambda}{1 - \mathbb{E}_\pi[\exp(-\lambda\tau_A)]} = \inf_{\pi^a(\tilde f)=1} \langle \tilde f, \mathscr{A}_\lambda \tilde f\rangle_{\pi^a}$$
$$= \inf_{f\in\mathscr{F}_A} \sup_{g|_A=0} \{2D_\lambda(f,g) - D_\lambda(g,g)\};$$

(2) for non-trivial subset A with finite A^c,

$$1/\mathbb{E}_\pi[\tau_A] = \inf_{\pi^a(\tilde f)=1} \langle \tilde f, \mathscr{A} \tilde f\rangle_{\pi^a}$$
$$= \inf_{f\in\mathscr{F}_A} \sup_{g|_A=0} \{2D(f,g) - D(g,g)\}.$$

Remark 4.2. Gaudillière and Landim [6] gave a variational formula of capacity for non-reversible Markov chains, from which we borrow some ideas for our proof here.

The following lemma is a key to prove Proposition 4.1.

Lemma 4.3. *Fix a non-trivial subset A with finite A^c.*

(1) For $\lambda > 0$, let φ, φ^* be the solutions of Poisson equations (2.7) *for chains Q and Q^*, respectively. Then $\overline{\varphi} := (\varphi+\varphi^*)/2$ on A^c is the unique solution of the Poisson equation*

$$\mathscr{A}_\lambda x(i) = \frac{\lambda}{1 - \mathbb{E}_\pi[\exp(-\lambda\tau_A)]}, \quad i \in A^c. \tag{4.1}$$

(2) Let ψ, ψ^* be the solutions of Poisson equations (2.8) *for chains Q and Q^*, respectively. Then $\overline{\psi} = (\psi+\psi^*)/2$ on A^c is the unique solution of the Poisson equation*

$$\mathscr{A} x(i) = 1/E_\pi[\tau_A], \quad i \in A^c.$$

Proof. We only prove (1) here while the proof of (2) is similar. It follows from Lemma 2.3 that for all $i \in A^c$,

$$(\lambda - Q^a)^*\varphi^*(i) = (\lambda - Q^a)\varphi(i) = \frac{\lambda}{1 - \mathbb{E}_\pi[\exp(-\lambda\tau_A)]}.$$

So it implies that

$$\begin{aligned}(\lambda - Q^a)^*\overline{\varphi}(i) &= \frac{1}{2}[(\lambda - Q^a)^*\varphi(i) + (\lambda - Q^a)^*\varphi^*(i)] \\ &= \frac{1}{2}[(\lambda - Q^a)^*\varphi(i) + (\lambda - Q^a)\varphi(i)] \\ &= (\lambda - \overline{Q}^a)\varphi(i),\end{aligned}$$

and thus

$$\mathscr{A}_\lambda\overline{\varphi}(i) = (\lambda - Q^a)\varphi(i) = \frac{\lambda}{1 - \mathbb{E}_\pi[\exp(-\lambda\tau_A)]}, \quad i \in A^c.$$

□

Proof of Proposition 4.1. We just give the proof of (1) here, since that of (2) is similar. Fix a non-trivial subset A with finite A^c and $\lambda > 0$.

From Lemma 4.3 and the fact $\pi^a(\overline{\varphi}) = 1$ it follows that

$$\langle \overline{\varphi}, \mathscr{A}_\lambda \overline{\varphi} \rangle_{\pi^a} = \frac{\lambda}{1 - \mathbb{E}_\pi[\exp(-\lambda \tau_A)]}. \tag{4.2}$$

For any $\tilde{f} : A^c \to \mathbb{R}$ with $\pi^a(\tilde{f}) = 1$, define $\tilde{f}_1 = \tilde{f} - \overline{\varphi}$ and then $\pi^a(\tilde{f}_1) = 0$. Using Lemma 4.3 again gives

$$\langle \tilde{f}_1, \mathscr{A}_\lambda \overline{\varphi} \rangle_{\pi^a} = \frac{\lambda \pi^a(\tilde{f}_1)}{1 - \mathbb{E}_\pi[\exp(-\lambda \tau_A)]} = 0,$$

so that

$$\langle \tilde{f}, \mathscr{A}_\lambda \tilde{f} \rangle_{\pi^a} = \langle \overline{\varphi}, \mathscr{A}_\lambda \overline{\varphi} \rangle_{\pi^a} + \langle \tilde{f}_1, \mathscr{A}_\lambda \tilde{f}_1 \rangle_{\pi^a} \geq \langle \overline{\varphi}, \mathscr{A}_\lambda \overline{\varphi} \rangle_{\pi^a},$$

since $\mathscr{A}_\lambda$ is positive definite on $L^2(A^c, \pi^a)$. Therefore,

$$\frac{\lambda}{1 - \mathbb{E}_\pi[\exp(-\lambda \tau_A)]} = \inf_{\pi^a(\tilde{f})=1} \langle \tilde{f}, \mathscr{A}_\lambda \tilde{f} \rangle_{\pi^a}.$$

Notice that $(\lambda - \overline{Q}^a)^{-1}$ is a bounded positive definite operator on $L^2(A^c, \pi^a)$, [11, Theorem 3.1.2] provides that

$$\langle \tilde{f}, (\lambda - \overline{Q}^a)^{-1} \tilde{f} \rangle_{\pi^a} = \sup_{\tilde{g}} \{ 2 \langle \tilde{f}, \tilde{g} \rangle_{\pi^a} - \langle \tilde{g}, (\lambda - \overline{Q}^a) \tilde{g} \rangle_{\pi^a} \}.$$

For any $\tilde{f}, \tilde{g}$ on A^c with $\pi^a(\tilde{f}) = 1$, we can extend $\tilde{f}, \tilde{g}$ to V by letting $f|_{A^c} = \tilde{f}$, $g|_{A^c} = \tilde{g}$ and f, g vanish on A. Thus $\pi(f) = \pi^a(\tilde{f}) = 1$, and

$$\langle (\lambda - Q^a)^* \tilde{f}, \tilde{g} \rangle_{\pi^a} = D_\lambda(f, g), \quad \langle \tilde{g}, (\lambda - Q^a) \tilde{g} \rangle_{\pi^a} = D_\lambda(g, g),$$

so that

$$\langle \tilde{f}, \mathscr{A}_\lambda \tilde{f} \rangle_{\pi^a} = \sup_{g|A=0} \{ 2 D_\lambda(f, g) - D_\lambda(g, g) \}.$$

This completes the proof. $\square$

Next, we will give an interesting application of Theorem 1.3. Recall the notations given as at Section 3.2. In Theorem 3.3, we prove that

$$\mathbb{E}_\pi[\exp(-\lambda \tau_A(\Gamma))] \geq \mathbb{E}_\pi[\exp(-\lambda \tau_A)], \quad \text{for all } A \subseteq V.$$

What we are interested in is the inverse problem: whether there exists a subset A such that the above strict inequality holds. This answer is yes as follows.

Theorem 4.4. *Let $Q_\Gamma = K + \mathrm{diag}(\pi)^{-1}\Gamma$ for some vorticity matrices which satisfy* (3.2). *Then $Q_\Gamma = K$ if and only if one of the following conditions fulfilled:*

(1) there exists $\lambda > 0$ such that $\mathbb{E}_\pi[\exp(-\lambda \tau_A(\Gamma))] = \mathbb{E}_\pi[\exp(-\lambda \tau_A)]$ for any $A \subseteq V$;

(2) $\mathbb{E}_\pi[\tau_A] = \mathbb{E}_\pi[\tau_A(\Gamma)]$ for any $A \subseteq V$.

Proof. It is obvious that the necessity is true, and for the sufficiency, since the proofs under (1) and (2) are similar, we only prove the assertion by assuming that there exists $\lambda > 0$ such that

$$\mathbb{E}_\pi[\exp(-\lambda\tau_A(\Gamma))] = \mathbb{E}_\pi[\exp(-\lambda\tau_A)], \quad A \subseteq V. \tag{4.3}$$

For convenience, denote $N = \mathrm{diag}(\pi)^{-1}\Gamma$. We will prove that $N = 0$ so that $Q_\Gamma = K$. Obviously, the diagonal of N is dull since Γ is anti-symmetric. We are going to prove that

$$n_{ij} = 0 \quad \text{for } i \neq j. \tag{4.4}$$

Before doing that, we need do some preparations. Fix a non-trivial subset A with finite A^c and denote by $\overline{\varphi}_\Gamma$ the solution of Poisson equation (4.1) for chain Q_Γ. By equality (4.2) in the proof of Proposition 4.1,

$$\begin{aligned}\frac{\lambda}{1-\mathbb{E}_\pi[\exp(-\lambda\tau_A(\Gamma))]} &= \langle\overline{\varphi}_\Gamma, (\lambda - Q_\Gamma^a)(\lambda - K^a)^{-1}(\lambda - Q_\Gamma^a)^*\overline{\varphi}_\Gamma\rangle_{\pi^a} \\ &= \langle\overline{\varphi}_\Gamma, (\lambda - K^a)\overline{\varphi}_\Gamma\rangle_{\pi^a} + \langle\overline{\varphi}_\Gamma, N^{a*}(\lambda - K^a)^{-1}N^a\overline{\varphi}_\Gamma\rangle_{\pi^a},\end{aligned} \tag{4.5}$$

where Q_Γ^a, K^a and π^a are defined as above. For the chain K, since it is reversible with respect to π, we use (1.3) to its first hitting time and obtain

$$\begin{aligned}\frac{\lambda}{1-\mathbb{E}_\pi[\exp(-\lambda\tau_A)]} &= \inf_{\pi^a(f)=1}\langle f, (\lambda - K^a)f\rangle_{\pi^a} \\ &\leq \langle\overline{\varphi}_\Gamma, (\lambda - K^a)\overline{\varphi}_\Gamma\rangle_{\pi^a}.\end{aligned} \tag{4.6}$$

But according to hypothesis (4.3) and (4.5)-(4.6), it must holds

$$\langle\overline{\varphi}_\Gamma, N^{a*}(\lambda - K^a)^{-1}N^a\overline{\varphi}_\Gamma\rangle_{\pi^a} = \langle N^a\overline{\varphi}_\Gamma, (\lambda - K^a)^{-1}N^a\overline{\varphi}_\Gamma\rangle_{\pi^a} = 0.$$

Thus

$$N^a\overline{\varphi}_\Gamma = 0, \tag{4.7}$$

since $(\lambda - K^a)^{-1}$ is positive definite on $L^2(A^c, \pi^a)$.

Now we prove (4.4) step by step. Without loss of generality, let $V = \{1, 2, ...\}$. Take $A = \{3, 4, ...\}$, from the definition of $\overline{\varphi}_\Gamma$ it follows that

$$\overline{\varphi}_\Gamma(1),\ \overline{\varphi}_\Gamma(2) > 0 \text{ and } \overline{\varphi}_\Gamma(k) = 0,\ k \geq 3. \tag{4.8}$$

Combine (4.7)-(4.8) with the fact $n_{11} = n_{22} = 0$, we obtain that $n_{12}^a = n_{21}^a = 0$, i.e., $n_{12} = n_{21} = 0$. A similar argument shows that $n_{13} = n_{23} = n_{31} = n_{32} = 0$ when we take $A = \{4, 5, ...\}$. Thus inductively $n_{ij} = 0$ for any $i \neq j$. □

5 Discrete time Markov chains

In this section, we consider discrete time Markov chains. Recall that V is a countable state space, on which $P = (p_{ij} : i, j \in V)$ is the probability transition matrix of an irreducible discrete time Markov chain $X = \{X_n : n \geq 0\}$. Assume that the chain X admits the unique stationary distribution $\pi = (\pi_i : i \in V)$:

$$\sum_{i\in V}\pi_i p_{ij} = \pi_j, \quad j \in V.$$

Let $\tau_\cdot(\tau_\cdot^+)$ denote the first hitting(return) time of X. Similar to the Dirichlet form D_λ in continuous time case, we denote a new Dirichlet form $E_\lambda(\lambda \geq 0)$ for chain X by

$$E_\lambda(f,g) = \langle f, (I - \mathrm{e}^{-\lambda}P)g\rangle, \text{ for } f,g \in L^2(\pi),$$

with the natural convention $E := E_0$.

A result analogous to Theorem 1.1 and 1.3 holds for discrete time Markov chains, we write down it in following corollary.

Corollary 5.1. *For the discrete time Markov chain X,*

(1) for any non-trivial subset $A \subseteq V$ and $\lambda > 0$,

$$\begin{aligned}\frac{1-\mathrm{e}^{-\lambda}}{1-\mathbb{E}_\pi[\exp(-\lambda\tau_A)]} &= \inf_{f\in\mathscr{F}_A}\sup_{g\in\mathscr{G}_A} E_\lambda(f-g,f+g)\\ &= \inf_{f\in\mathscr{F}_A}\sup_{\mathscr{S}(g)\subseteq\mathscr{S}(f)}\{2E_\lambda(f,g)-E_\lambda(g,g)\};\end{aligned} \tag{5.1}$$

(2) for any non-trivial subset $A \subseteq V$,

$$\begin{aligned}1/\mathbb{E}_\pi[\tau_A] &= \inf_{f\in\mathscr{F}_A}\sup_{g\in\mathscr{G}_f} E(f-g,f+g)\\ &= \inf_{f\in\mathscr{F}_A}\sup_{\mathscr{S}(g)\subseteq\mathscr{S}(f)}\{2E(f,g)-E(g,g)\}.\end{aligned} \tag{5.2}$$

In particular, if there exists $i \in V$ such that $\mathbb{E}_i[(\tau_i^+)^3] < \infty$, then $\mathscr{G}_f$ in (5.2) can be replaced by $\mathscr{G}_A$.

Proof. For the assertion (2), we just replace Q by $I-P$ in the proofs of variational formulas (1.2) and (1.5), we can complete its proof easily. Now, fix $\lambda > 0$ and a non-trivial subset A. As presented in [7, Chapter 6], $(\mathbb{E}_i[\exp(-\lambda\tau_A)] : i \in V)$ is the minimal non-negative solution of the equation

$$\begin{cases} x_i = \mathrm{e}^{-\lambda}\sum_{j\in V} p_{ij}x_j, & i \in A^c;\\ x_i = 1, & i \in A.\end{cases}$$

So by some simple transforms, we obtain that

$$\phi_i = \frac{1-\mathbb{E}_i[\exp(-\lambda\tau_A)]}{1-\mathbb{E}_\pi[\exp(-\lambda\tau_A)]},\ i \in V$$

is a solution of Poisson equation

$$\begin{cases} (I-\mathrm{e}^{-\lambda}P)x(i) = \frac{1-\mathrm{e}^{-\lambda}}{1-\mathbb{E}_\pi[\exp(-\lambda\tau_A)]}, & i \in A^c;\\ x_i = 0, & i \in A.\end{cases} \tag{5.3}$$

Now using the arguments like the ones for Poisson equation (2.7) to (5.3) gives the desired result. □

To prove Corollary 1.5, we need a monotone property of hitting times. For this, denote by $P^*, \overline{P}$ the dual and reversible parts of P with respect to π, respectively. Define

$$P_\alpha = (1-\alpha)P + \alpha P^*, \quad 0 \le \alpha \le 1. \tag{5.4}$$

Then P_α is an irreducible probability transition matrix and $P_{1/2} = \overline{P}$. They all have the same stationary distribution π. Denote the first hitting time and the average hitting time of chain P_α by $\tau_\cdot(\alpha)$, $T_0(\alpha)$, and then the monotone and symmetry properties in Theorem 3.6 can be extended to discrete time Markov chains by the new variational formulas (5.1)-(5.2).

Corollary 5.2. *Let chains P_α be defined as at* (5.4)*. For any subset $A \subseteq V$ and $\lambda > 0$, denote by $R(\alpha)$ either $\mathbb{E}_\pi[\tau_A(\alpha)]$ or $T_0(\alpha)$. Then*

(*a*) $\mathbb{E}_\pi[\exp(-\lambda\tau_A(\alpha))] = \mathbb{E}_\pi[\exp(-\lambda\tau_A(1-\alpha))]$ *and $R(\alpha) = R(1-\alpha)$ for $\alpha \in [0,1]$.*

(*b*) $\mathbb{E}_\pi[\exp(-\lambda\tau_A(\alpha))]$ *is non-increasing for $\alpha \in [0,1/2]$ and $R(\alpha)$ is non-decreasing for $\alpha \in [0,1/2]$.*

Now, we can give a proof of Corollary 1.5 as follows.

Proof of Corollary 1.5. Assume that the state space V is finite and fix $i \in V$. As shown in the proof of Corollary 24 in [1, Chapter 9], when $\alpha \in [0,1/2)$,

$$\frac{d}{d\alpha}\mathbb{E}_j[\tau_i(\alpha)] = \frac{1}{\pi_i}\sum_k [(z_\alpha)_{ik} - (z_\alpha)_{jk}] \sum_l (p^*_{kl} - p_{kl})(z_\alpha)_{li}, \quad j \in V.$$

Average over j and use that fact $\sum_j \pi_j (z_\alpha)_{jk} = 0$ to derive

$$\begin{aligned}
\frac{d}{d\alpha}\mathbb{E}_\pi[\tau_i(\alpha)] &= \frac{1}{\pi_i}\sum_k\sum_l (z_\alpha)_{ik}(p^*_{kl} - p_{kl})(z_\alpha)_{li} \\
&= \frac{1}{\pi_i}[Z_\alpha(P^* - P)Z_\alpha]_{ii} \\
&= \frac{1}{(1-2\alpha)\pi_i}[Z_\alpha(P^*_\alpha - P_\alpha)Z_\alpha]_{ii},
\end{aligned}$$

where Z_α is the fundamental matrix of P_α and $\tau_i(\alpha)$ is the first hitting time to state i of chain P_α. Notice that Corollary 5.2 gives that $\mathbb{E}_\pi[\tau_i(\alpha)]$ is non-decreasing on $[0,1/2]$. So $\frac{d}{d\alpha}\mathbb{E}_\pi[\tau_i(\alpha)] \ge 0$, i.e., $[Z_\alpha(P^*_\alpha - P_\alpha)Z_\alpha]_{ii} \ge 0$ for $\alpha \in [0,1/2)$. In particular, $[Z(P^* - P)Z]_{ii} \ge 0$ for $i \in V$ and then the proof be complete by $\mathrm{trace}[Z^2(P^* - P)] = \mathrm{trace}[Z(P^* - P)Z]$. $\square$

Acknowledgement Research supported in part by NSFC (No 11501531, 11571043, 11771047).

References

[1] D.-J. Aldous and J.-A. Fill. *Reversible Markov chains and random walks on graphs.* 2002. URL www.berkeley.edu/users/aldous/book.html.

[2] W.-J. Anderson. *Continuous-time Markov chains.* Springer-Verlag, New York, 1991.

[3] M.-F. Chen. *From Markov chains to non-equilibrium particle systems*. World Scientific, Singapore, 1992.

[4] R. Cogburn. A uniform theory for sums of Markov chain transition probabilities. *The Annals of Probability*, 3(2):191–214, 1975.

[5] P.-G. Doyle. Energy for Markov chains. 1994. http://www.math.dartmouth.edu/doyle.

[6] A. Gaudillière and C. Landim. A Dirichlet principle for non reversible Markov chains and some recurrence theorems. *Probability Theory and Related Fields*, 158:55–89, 2014.

[7] Z.-T. Hou and Q.-F. Guo. *Time-homogeneous Markov processes with countable state space(in Chinese)*. Beijing Sci. Press and Springer-Verlag, 1978. Beijing Sci. Press, English translation(1988).

[8] L.-J. Huang and Y.-H. Mao. On some hitting times for nonreversible Markov chains. *Journal of Applied Probability*, 54(2):627–637, 2017.

[9] C.-R. Hwang. Accelerating Monte Carlo Markov processes. *Cosmos*, 1(1):87–94, 2005.

[10] J. Bierkens. Non-reversible Metropolis-Hastings. *Statistics and Computing*, 26(6):1213–1228, 2016.

[11] S. Kesavan. *Topics in functional analysis and applications*. Wiley, New York, 1989.

[12] K.-K.-J. Kinateder and P. McDonald. Variational principles for average exit time moments for diffusions in Euclidean space. *Proceedings of the American Mathematical Society*, 127:2767–2772, 1999.

[13] Y.-H. Mao. The eigentime identity for continuous-time ergodic Markov chains. *Journal of Applied Probability*, 41(4):1071–1080, 2004.

[14] S.-P. Meyn and R.-L. Tweedie. *Markov chains and stochastic stability, 2nd edition*. Cambridge University Press, New York, 2009.

[15] P.-H. Peskun. Optimum Monte-Carlo sampling using Markov chains. *Biometrika*, 60(3):607–612, 1973.

[16] M. Shaked and J.-G. Shanthikumar. *Stochastic orders*. Springer Series in Statistics, 2007.

[17] L. Tierney. A note on Metropolis-Hastings kernels for general state spaces. *The Annals of Applied Probability*, 8(1):1–9, 1998.

[18] F.-Y. Wang. *Functional inequalities, Markov semigroups and spectral theory*. Science Press, Beijing, 2005.