J. Appl. Probab. **61,** 540–557 (2024)
*doi:*10.1017/jpr.2023.54

COMPARISON THEOREM AND STABILITY UNDER PERTURBATION OF TRANSITION RATE MATRICES FOR REGIME-SWITCHING PROCESSES

JINGHAI SHAO,* *Tianjin University*

Abstract

A comparison theorem for state-dependent regime-switching diffusion processes is established, which enables us to pathwise-control the evolution of the state-dependent switching component simply by Markov chains. Moreover, a sharp estimate on the stability of Markovian regime-switching processes under the perturbation of transition rate matrices is provided. Our approach is based on elaborate constructions of switching processes in the spirit of Skorokhod's representation theorem varying according to the problem being dealt with. In particular, this method can cope with switching processes in an infinite state space and not necessarily of birth–death type. As an application, some known results on the ergodicity and stability of state-dependent regime-switching processes can be improved.

Keywords: Comparison theorem; regime-switching diffusions; ergodicity; perturbation theory

2020 Mathematics Subject Classification: Primary 60J27
Secondary 60J60; 60K37

1. Introduction

Stochastic processes with regime-switching have been extensively studied in many research fields due to their characterization of random changes of the environment between different regimes; see, for instance, [4, 8, 9, 13, 17, 18, 19, 22, 28, 31] and references therein. In particular, when the switching of the environment depends on the state of the dynamic system studied, usually called a state-dependent regime-switching process (RSP), the properties of such a system become much more complicated due to their intensive interaction. The monograph [28] introduced various properties of state-dependent RSPs, telling us that it is a very challenging task to provide easily verifiable conditions to justify the ergodicity and stability of state-dependent RSPs.

In view of the relatively abundant results on state-independent RSPs, it is natural to simplify a state-dependent RSP into a state-independent one. So, we need to establish some kind of comparison theorem for state-dependent RSPs. Such an idea has been used in many works. For instance, [4] used this idea to study the exponential ergodicity in the Wasserstein distance for birth–death-type state-dependent RSPs based on an application of the weak Harris's theorem. The state-dependent switching process and the constructed state-independent process

Received 29 November 2022; accepted 24 June 2023.
* Postal address: Center for Applied Mathematics, Tianjin University, Tianjin 300072, China. Email: shaojh@tju.edu.cn

constitute a coupling process. However, the constructed coupling process in [4] is no longer a Markovian process and needs modification in applications (cf. [4, Remark 3.10]). Majda and Tong used the same method to study the exponential ergodicity in the setting of piecewise deterministic processes with regime-switching and applied their results to tropical stochastic lattice models. In [20], the author constructed a Markovian coupling for birth–death-type state-dependent RSPs; this was extended in [22] to a general case under the condition of the existence of an order-preserving coupling for Markov chains (cf. [22, Lemma 2.7]). This result is not constructive, and the verification of [22, Assumption 2.6] is not easy in applications.

Accordingly, our first purpose is to establish a comparison theorem for more general state-dependent RSPs, in particular to get rid of the restriction of birth–death-type switching and to be applicable to switching processes in an infinitely countable state space. Our comparison theorem is in the pathwise sense. As a consequence, the corresponding results in [4, 12, 20, 22] can be generalized after certain necessary modifications.

For state-dependent RSPs, study of the Feller property and the smooth dependence of initial values is of great interest. It is quite different to that of Markovian RSPs and diffusion processes, as noted in [17], [27], [28, Chapter 2], and references therein. For instance, in [17, Theorem 3.1], the continuous differentiability of the continuous component of an RSP with respect to the initial value in L^p with $p \in (0, 1)$ was proved under suitable regular conditions. Moreover, restricted to a bounded set, certain gradient estimates associated with the continuous component of an RSP can be derived from [17, Theorem 4.1]. Recently, [24] studied the continuous dependence of initial values for state-dependent RSPs, which are solutions to certain stochastic functional differential equations with regime-switching. As shown in [24], the key point is the estimate for the quantity

$$\Theta(t, \Lambda, \tilde{\Lambda}) := \frac{1}{t}\int_0^t \mathbb{P}\big(\Lambda_s \neq \tilde{\Lambda}_s\big)\, \mathrm{d}s, \qquad t > 0. \tag{1.1}$$

Moreover, it was also shown in [23] that the quantity $\Theta(t, \Lambda, \tilde{\Lambda})$ plays a crucial role in study of the stability of RSPs under perturbation of the Q-matrix. However, the estimates in [23, 24] do not work for switching processes in an infinite state space, which needs to be generalized from the viewpoint of applications. Also, the estimation of $\Theta(t, \Lambda, \tilde{\Lambda})$ develops the classical perturbation theory for Markov chains (cf. [14, 15, 16, 29]). More details on this are given in Section 3.

Our improvements in the previous two concerned problems are all based on a new observation on Skorokhod's representation theorem for jumping processes. The approach of using Skorokhod's representation theorem to study RSPs has been widely used in the literature; see, e.g., [8, 17, 18, 23, 28], etc. The basic idea of Skorokhod's representation theorem is to represent the switching process in terms of an integral with respect to a Poisson random measure based on a sequence of constructed intervals on the real line. The length of each interval is determined by $(q_{ij}(x))$. However, the impact of the construction and sort order of the intervals on the jumping processes obtained has been neglected in all the previous works. In this work, we show that it is necessary to carry out different constructions of the intervals to solve different problems. This is illustrated through establishing the comparison theorem and studying the stability problem under perturbation of the Q-matrix.

This work is organized as follows. In Section 2, we first provide the construction of the coupling process, then establish the comparison theorem for state-dependent RSPs. As an illustrative example, we use this comparison theorem to study the stability of state-dependent RSPs. In Section 3, we also first provide the construction of the coupling process $(\Lambda_t, \tilde{\Lambda}_t)$ in (1.1), the key point of which is the construction of intervals used in Skorokhod's representation theorem. Then we provide an estimate of (1.1) which improves the one given in [23].

2. Comparison theorem for state-dependent RSPs

Let $\mathcal{S} = \{1, 2, \ldots, N\}$ with $2 \le N \le \infty$. So, $\mathcal{S}$ is allowed to be an infinitely countable state space by taking $N = \infty$. Consider

$$\mathrm{d}X_t = b(X_t, \Lambda_t)\,\mathrm{d}t + \sigma(X_t, \Lambda_t)\,\mathrm{d}B_t, \tag{2.1}$$

where $b : \mathbb{R}^d \times \mathcal{S} \to \mathbb{R}^d$, $\sigma : \mathbb{R}^d \times \mathcal{S} \to \mathbb{R}^{d\times d}$ satisfying suitable conditions, and (B_t) is a d-dimensional Brownian motion. Here, (Λ_t) is a jumping process on $\mathcal{S}$ satisfying

$$\mathbb{P}(\Lambda_{t+\delta} = j \mid \Lambda_t = i, X_t = x) = \begin{cases} q_{ij}(x)\delta + o(\delta) & \text{if } i \ne j, \\ 1 + q_{ii}(x)\delta + o(\delta) & \text{if } i = j, \end{cases} \tag{2.2}$$

provided $\delta > 0$ is sufficiently small. When $(q_{ij}(x))$ does not depend on x, (X_t, Λ_t) is called a state-independent RSP or a Markovian RSP. Meanwhile, when $(q_{ij}(x))$ does depend on x, (X_t, Λ_t) is called a state-dependent RSP, which is used to model the phenomenon that the dynamic system (X_t) can impact the change rate of the random environment in applications. There is an intensive interaction between (X_t) and (Λ_t) for state-dependent RSPs, and hence it is quite desirable to develop a stochastic comparison theorem to simplify such a system. Since (X_t) can be simplified, if necessary, by using a classical comparison theorem for diffusion processes (cf. [10, 11]) or for Lévy processes (cf. [26]) performed separately in each fixed regime i, the key point is to control the jumping component (Λ_t) whose transition rates vary with (X_t).

Theorem 2.1. *Assume that the stochastic differential equations (SDEs) (2.1) and (2.2) admit a solution (X_t, Λ_t) for any initial value $(X_0, \Lambda_0) = (x, i) \in \mathbb{R}^d \times \mathcal{S}$. Suppose that $(q_{ij}(x))$ is conservative for every $x \in \mathbb{R}^d$ and satisfies*

- *(Q1)* $K_0 := \sup_{x\in\mathbb{R}^d} \sup_{i\in\mathcal{S}} |q_{ii}(x)| < \infty$;
- *(Q2) for all $i \in \mathcal{S}$, there is a $c_i \in \mathbb{N}$ such that $q_{ij}(x) = 0$ for all $j \in \mathcal{S}$ with $|j - i| > c_i$ and all $x \in \mathbb{R}^d$.*

Define

$$q^*_{ij} = \begin{cases} \sup\limits_{x\in\mathbb{R}^d} \max\limits_{j<\ell\le i} q_{\ell j}(x), & j < i, \\ \inf\limits_{x\in\mathbb{R}^d} \min\limits_{\ell\le i} q_{\ell j}(x), & j > i, \\ -\sum_{i\ne j} q^*_{ij}, & j = i, \end{cases} \quad \text{and} \quad \bar{q}_{ij} = \begin{cases} \inf\limits_{x\in\mathbb{R}^d} \min\limits_{j<\ell\le i} q_{\ell j}(x), & j < i, \\ \sup\limits_{x\in\mathbb{R}^d} \max\limits_{\ell\le i} q_{\ell j}(x), & j > i, \\ -\sum_{j\ne i} \bar{q}_{ij}, & j = i. \end{cases}$$

Then there exist two continuous-time Markov chains (Λ^*_t) and $(\bar{\Lambda}_t)$ on $\mathcal{S}$ with transition rate matrices (q^*_{ij}) and $(\bar{q}_{ij})$ respectively such that

$$\mathbb{P}\big(\Lambda^*_t \le \Lambda_t \le \bar{\Lambda}_t \text{ for all } t \ge 0\big) = 1.$$

https://doi.org/10.1017/jpr.2023.54 Published online by Cambridge University Press

Remark 2.1.

(i) For birth–death-type $(q_{ij}(x))$, i.e. $q_{ij}=0$ for $|i-j|\geq 2$, we have

$$q^*_{i(i-1)}=\sup_{x\in\mathbb{R}^d} q_{i(i-1)}(x), \qquad q^*_{i(i+1)}=\inf_{x\in\mathbb{R}^d} q_{i(i+1)}(x),$$
$$\bar{q}_{i(i-1)}=\inf_{x\in\mathbb{R}^d} q_{i(i-1)}(x), \qquad \bar{q}_{i(i+1)}=\sup_{x\in\mathbb{R}^d} q_{i(i+1)}(x).$$

This coincides with the Markov chain constructed in [20].

(ii) There are many works that investigate the existence of a solution to (2.1) and (2.2); see [28] under Lipschitzian conditions, [18] under non-Lipschitzian conditions, and [31] under integrable conditions.

(iii) Assumption (Q1) ensures that there exists a unique Markov chain $(\Lambda^*_t)_{t\geq 0}$ $\big((\bar{\Lambda}_t)_{t\geq 0}\big)$ associated with (q^*_{ij}) $\big((\bar{q}_{ij})\big)$; see, e.g., [3, Corollary 2.24].

As mentioned in the introduction, this comparison theorem can help us to generalize the corresponding results in [4, 12, 20, 22]. More precisely, we can remove Assumption 3.1 of birth–death-type restriction in [4] and generalize [4, Theorems 3.3, 3.4] there. Also, the exponential convergence results, [12, Theorems 2.1, 2.3], for the two stochastic lattice models for moist tropical convection in climate science studied can be generalized by considering more general jumping processes besides the birth–death processes used in [12].

As an illustrative example, we apply Theorem 2.1 to investigate the stability of state-dependent RSPs. The stability of stochastic processes with regime switching has been studied in many works. We refer the reader to [13, 28] for surveys on this topic, and also to [21] for some recent results on the stability of state-dependent RSPs based on M-matrix theory, the Perron–Frobenius theorem, and the Friedholm alternative.

Theorem 2.2. *Let $\big(X^{x,i}_t, \Lambda^{x,i}_t\big)$ be the solution to (2.1) and (2.2) with initial value (x,i). Assume that the conditions of Theorem 2.1 hold. Suppose that there exist a function $\rho\in C^2(\mathbb{R}^d)$, constants $\beta_i\in\mathbb{R}$ for $i\in\mathcal{S}$, and constants $p, \tilde{c}>0$ such that*

$$\mathscr{L}^{(i)}\rho(x)\leq\beta_i\rho(x), \qquad \rho(x)\geq\tilde{c}|x|^p, \qquad x\in\mathbb{R}^d,\ i\in\mathcal{S}, \tag{2.3}$$

where

$$\mathscr{L}^{(i)}\rho(x)=\sum_{k=1}^d b_k(x,i)\partial_k\rho(x)+\frac{1}{2}\sum_{k,l=1}^d a_{kl}(x,i)\partial_k\partial_l\rho(x), \qquad a_{kl}(x,i)=\sum_{j=1}^d\sigma_{kj}(x,i)\sigma_{lj}(x,i).$$

Through reordering $\mathcal{S}$, without loss of the generality we may assume that $(\beta_i)_{i\in\mathcal{S}}$ is nondecreasing. Let $(\bar{q}_{ij})$ be defined as in Theorem 2.1. Assume that $(\bar{q}_{ij})$ is irreducible, and admits a unique invariant probability measure $\bar{\mu}$ such that

$$\sum_{i\in\mathcal{S}}\bar{\mu}_i\beta_i<0. \tag{2.4}$$

Then there exists $p'\in(0,p]$ such that $\lim_{t\to\infty}\mathbb{E}\big|X^{x,i}_t\big|^{p'}=0$ for $x\in\mathbb{R}^d$, $i\in\mathcal{S}$.

https://doi.org/10.1017/jpr.2023.54 Published online by Cambridge University Press

Remark 2.2. In Theorem 2.2, via condition (2.3), we characterize the stability property of the process (X_t) at each fixed state $i \in \mathcal{S}$ by a constant $\beta_i \in \mathbb{R}$, which is measured under a common Lyapunov type function ρ. Then, with the help of the auxiliary Markov chain $(\bar{q}_{ij})$, the average condition (2.4) yields the stability of (X_t).

Next, we provide an example to illustrate the application of Theorem 2.2.

Example 2.1. Consider $\mathrm{d}X_t = b_{\Lambda_t} X_t\,\mathrm{d}t + \sigma_{\Lambda_t} X_t\,\mathrm{d}B_t$, where (Λ_t) is a jumping process on $\mathcal{S} = \{1, 2, 3\}$ with the state-dependent transition rate matrix $(q_{ij}(x))_{i,j\in\mathcal{S}}$ given by $q_{ij}(x) = 1 + |i - j|(x^2 \wedge 1)$, $i, j \in \mathcal{S}$, $x \in \mathbb{R}$.

Take $\rho(x) = x^2$ in (2.3) to yield $\beta_i = 2b_i + \sigma_i^2$, $i \in \mathcal{S}$. By virtue of Theorem 2.1, direct calculation yields

$$(\bar{q}_{ij})_{i,j\in\mathcal{S}} = \begin{pmatrix} -5 & 2 & 3 \\ 1 & -4 & 3 \\ 1 & 1 & -2 \end{pmatrix},$$

and its unique invariant probability measure is $(\bar{\mu}_1, \bar{\mu}_2, \bar{\mu}_3) = \left(\frac{1}{6}, \frac{7}{30}, \frac{3}{5}\right)$. Then, by Theorem 2.2, if

$$\frac{1}{6}\left(2b_1 + \sigma_1^2\right) + \frac{7}{30}\left(2b_2 + \sigma_2^2\right) + \frac{3}{5}\left(2b_3 + \sigma_3^2\right) < 0,$$

the process (X_t) is stable in the sense that $\lim_{t\to\infty} \mathbb{E}\left[\left|X_t^{x,i}\right|^p\right] = 0$, $x \in \mathbb{R}$, $i \in \mathcal{S}$, for some $p \in (0, 2]$.

Example 2.2. Consider the regime-switching diffusion process $\mathrm{d}X_t = b_{\Lambda_t} X_t\,\mathrm{d}t + \sigma_{\Lambda_t} X_t\,\mathrm{d}B_t$, and (Λ_t) a jumping process on $\mathcal{S} = \{1, 2, \ldots\}$ with state-dependent transition rate matrix $(q_{ij}(x))$ given by $q_{12}(x) = 1 + \sin x$, $q_{1k}(x) = 0$ if $k \geq 3$, $q_{11}(x) = -1 - \sin x$, and, for $n \geq 2$, $q_{n(n+1)}(x) = 1 + \sin x$, $q_{n1}(x) = 3 - \sin x$, $q_{nk}(x) = 0$ if $k \notin \{1, n + 1\}$, and $q_{nn}(x) = -4$. According to the definition of $(\bar{q}_{ij})$ in Theorem 2.1, direct calculation yields $\bar{q}_{12} = 2$, $\bar{q}_{1k} = 0$ if $k \geq 3$, $\bar{q}_{11} = -2$, and, for $n \geq 2$, $\bar{q}_{n1} = 2$, $\bar{q}_{n(n+1)} = 2$, $\bar{q}_{nk} = 0$ if $k \notin \{1, n, n + 1\}$, $\bar{q}_{nn} = -4$. The invariant probability measure $\bar{\mu} = (\bar{\mu}_n)_{n\in\mathcal{S}}$ for $(\bar{q}_{ij})$ is given by $\bar{\mu}_n = 1/2^n$, $n \geq 1$. Take $\rho(x) = x^2$, then $\rho \in C^2(\mathbb{R})$ and $\mathscr{L}^{(i)}\rho(x) \leq \beta_i \rho(x)$ with $\beta_i = 2b_i + \sigma_i^2$, $i \in \mathcal{S}$. According to Theorem 2.2, if

$$\sum_{n=1}^{\infty} \frac{1}{2^n}\left(2b_n + \sigma_n^2\right) < 0,$$

the process is stable in the sense that $\lim_{t\to\infty} \mathbb{E}\left[\left|X_t^{x,i}\right|^p\right] = 0$, $x \in \mathbb{R}$, $i \in \mathcal{S}$, for some $p \in (0, 2]$.

2.1. Construction of the coupling process

In the spirit of Skorokhod [25] in the study of processes with rapid switching, [8] presented a representation of a Markov chain in terms of a Poisson random measure and studied the stability of RSPs. This kind of representation theorem has been widely applied in the study of various properties of RSPs.

Before giving out our precise representation theorem for establishing the comparison theorem, let us recall the construction in [8] for comparison. When $\mathcal{S} = \{1, 2, \ldots, N\}$ is a finite

https://doi.org/10.1017/jpr.2023.54 Published online by Cambridge University Press

state space, let $\Delta_{ij}(x)$ be consecutive, left-closed, right-open intervals on $[0, \infty)$, each having length $q_{ij}(x)$. More precisely,

$$\Delta_{12}(x)=[0, q_{12}(x)),\ \ \Delta_{13}(x)=[q_{12}(x), q_{12}(x)+q_{13}(x)),\ \ \ldots,\ \ \Delta_{1N}(x)=\big[\textstyle\sum_{j\neq 1} q_{1j}(x), q_1(x)\big),$$
$$\Delta_{21}(x)=[q_1(x), q_1(x)+q_{21}(x)),\ \ \ldots,$$
$$\vdots$$

Define a function $h:\mathbb{R}^d\times\mathcal{S}\times\mathbb{R}\to\mathbb{R}$ by $h(x, i, z)=\sum_{j\in\mathcal{S},j\neq i}(j-i)\mathbf{1}_{\Delta_{ij}(x)}(z)$. Then, (Λ_t) is a jumping process satisfying (2.2) as a solution to the SDE $\mathrm{d}\Lambda_t=\int_{[0,\infty)}h(X_t, \Lambda_{t-}, z)\,\mathcal{N}(\mathrm{d}t, \mathrm{d}z)$, where (X_t) satisfies (2.1), and $\mathcal{N}(\mathrm{d}t, \mathrm{d}z)$ is a Poisson random measure with intensity $\mathrm{d}t\times\mathrm{d}z$.

We can now introduce our construction of the coupling process in three steps. We assume that conditions (Q1) and (Q2) hold in this section.

Step 1: Construction of intervals. For every fixed $x\in\mathbb{R}^d$ and $i, j\in\mathcal{S}$, we define the intervals $\Gamma_{ij}(x)$, Γ^*_{ij}, and $\bar\Gamma_{ij}$ in the following way. Starting from 0, the intervals $\Gamma_{ij}(x)$ for $j<i$ are defined on the positive half-line, while $\Gamma_{ij}(x)$ for $j>i$ are defined on the negative half-line. More precisely,

$$\begin{aligned}\Gamma_{i1}(x)&=[0, q_{i1}(x)),\\ \Gamma_{i2}(x)&=[q_{i1}(x), q_{i1}(x)+q_{i2}(x)),\\ &\ \ \vdots\\ \Gamma_{i(i-1)}(x)&=\Big[\textstyle\sum_{j=1}^{i-2}q_{ij}(x), \sum_{j=1}^{i-1}q_{ij}(x)\Big),\end{aligned}\tag{2.5}$$

and

$$\begin{aligned}\Gamma_{i(i+c_i)}(x)&=[-q_{i(i+c_i)}(x), 0),\\ \Gamma_{i(i+c_i-1)}(x)&=[-q_{i(i+c_i-1)}(x)-q_{i(i+c_i)}(x), -q_{i(i+c_i)}(x)),\\ &\ \ \vdots\\ \Gamma_{i(i+1)}(x)&=\Big[-\textstyle\sum_{j=i+1}^{i+c_i}q_{ij}(x), -\sum_{j=i+2}^{i+c_i}q_{ij}(x)\Big),\end{aligned}\tag{2.6}$$

where c_i is given in (Q2). Analogously, by replacing $q_{ij}(x)$ in (2.5) and (2.6) with $\bar q_{ij}$ and q^*_{ij} respectively, we can define the intervals $\bar\Gamma_{ij}$ and Γ^*_{ij}. Here and in what follows, we put $\Gamma_{ij}(x)=\emptyset$ if $q_{ij}(x)=0$ and $\Gamma_{ii}(x)=\emptyset$ for the convenience of notation. This convention also applies to the intervals $\bar\Gamma_{ij}$ and Γ^*_{ij}.

The sort order of the intervals $\Gamma_{ij}(x)$, $\bar\Gamma_{ij}$, and Γ^*_{ij} according to $j>i$ or $j<i$ will play an important role in the argument below. The assumption on the existence of c_i is used here such that on the negative half-line, the first interval starting from 0 is associated with the state $j\in\mathcal{S}$ satisfying $j=\max\{k\in\mathcal{S}; k>i, q_{ik}(x)>0\}$. The common starting point 0 of $\Gamma_{ij}(x)$, $\bar\Gamma_{ij}$, and Γ^*_{ij} for different $i\in\mathcal{S}$ also plays an important role in our construction of the order-preservation coupling process.

Step 2: Explicit construction of Poisson random measure. Here we use the method of [11, Chapter I, p. 44] to present a concrete construction of the Poisson random measure. Denote by $\mathbf{m}(\mathrm{d}x)$ the Lebesgue measure over $\mathbb{R}$. Let ξ_k, $k=1, 2, \ldots$, be random variables taking values in $[-K_0, K_0]$ with $\mathbb{P}(\xi_k\in\mathrm{d}x)=\mathbf{m}(\mathrm{d}x)/2K_0$, and τ_k, $k=1, 2, \ldots$, be non-negative random variables such that $\mathbb{P}(\tau_k>t)=\mathrm{e}^{-2tK_0}$, $t\geq 0$. Assume that all ξ_k and τ_k, $k\geq 1$, are mutually independent. Let $\zeta_n=\tau_1+\tau_2+\cdots+\tau_n$, $n=1, 2, \ldots$, and $\zeta_0=0$, $\mathcal{D}_{\mathbf{p}}=\bigcup_{n\geq 1}\{\zeta_n\}$, and

https://doi.org/10.1017/jpr.2023.54 Published online by Cambridge University Press

$\mathbf{p}(t)=\sum_{0\le s<t}\Delta\mathbf{p}(s)$, with $\Delta\mathbf{p}(s)=0$ for $s\notin\mathcal{D}_{\mathbf{p}}$, and $\Delta\mathbf{p}(\zeta_n)=\xi_n$, $n\ge 1$, where $\Delta\mathbf{p}(s):=\mathbf{p}(s)-\mathbf{p}(s-)$. The finiteness of K_0 means that $\lim_{n\to\infty}\zeta_n=\infty$ almost surely (a.s.); that is, during each finite time period, there exists a finite number of jumps for $(\mathbf{p}(t))$. Let

$$\mathcal{N}_{\mathbf{p}}([0,t]\times A)=\#\{s\in\mathcal{D}_{\mathbf{p}};0\le s\le t,\ \Delta\mathbf{p}(s)\in A\},\qquad t>0,\ A\in\mathscr{B}(\mathbb{R}).$$

As a consequence, $\mathbf{p}(t)$ and $\mathcal{N}_{\mathbf{p}}(\mathrm{d}t,\mathrm{d}z)$ are respectively a Poisson point process and a Poisson random measure with intensity measure $\mathrm{d}t\,\mathbf{m}(\mathrm{d}x)$. It is always assumed that $\mathbf{p}(t)$ is independent of the Brownian motion (B_t) in (2.1).

Step 3: Construction of coupling processes. Define three functions ϑ, ϑ^*, and $\bar{\vartheta}$ as follows:

$$\begin{aligned}\vartheta(x,i,z)&=\sum_{j\in\mathcal{S},j\neq i}(j-i)\mathbf{1}_{\Gamma_{ij}(x)}(z),\\ \vartheta^*(i,z)&=\sum_{j\in\mathcal{S},j\neq i}(j-i)\mathbf{1}_{\Gamma^*_{ij}}(z),\\ \bar{\vartheta}(i,z)&=\sum_{j\in\mathcal{S},j\neq i}(j-i)\mathbf{1}_{\bar{\Gamma}_{ij}}(z).\end{aligned}$$

Then, consider the following SDEs:

$$\mathrm{d}\Lambda_t=\int_{\mathbb{R}}\vartheta(X_t,\Lambda_{t-},z)\,\mathcal{N}_{\mathbf{p}}(\mathrm{d}t,\mathrm{d}z),\tag{2.7}$$

$$\mathrm{d}\bar{\Lambda}_t=\int_{\mathbb{R}}\bar{\vartheta}(\bar{\Lambda}_{t-},z)\,\mathcal{N}_{\mathbf{p}}(\mathrm{d}t,\mathrm{d}z),\tag{2.8}$$

$$\mathrm{d}\Lambda^*_t=\int_{\mathbb{R}}\vartheta^*\big(\Lambda^*_{t-},z\big)\,\mathcal{N}_{\mathbf{p}}(\mathrm{d}t,\mathrm{d}z),\tag{2.9}$$

and $\Lambda_0=\bar{\Lambda}_0=\Lambda^*_0=i_0\in\mathcal{S}$. Here, recall that (X_t) satisfies (2.1).

The fact that the solution to (2.7) satisfies (2.2) can be checked directly using the property that $\mathbb{P}(\mathcal{N}_{\mathbf{p}}((0,\delta]\times A)\ge 2)=o(\delta)$, $\delta>0$. Next, we verify that $\big(\bar{\Lambda}_t\big)$ and $\big(\Lambda^*_t\big)$ given by (2.8) and (2.9) are the jumping processes associated with $\big(\bar{q}_{ij}\big)$ and $\big(q^*_{ij}\big)$ respectively. According to Itô's formula, for any bounded measurable function F on $\mathcal{S}$,

$$\begin{aligned}\mathbb{E}F\big(\bar{\Lambda}_t\big)&=F\big(\bar{\Lambda}_0\big)+\mathbb{E}\int_0^t\int_{\mathbb{R}}\Big(F\big(\bar{\Lambda}_{s-}+\bar{\vartheta}\big(\bar{\Lambda}_{s-},z\big)\big)-F\big(\bar{\Lambda}_{s-}\big)\Big)\,\mathcal{N}_{\mathbf{p}}(\mathrm{d}s,\mathrm{d}z)\\ &=F\big(\bar{\Lambda}_0\big)+\mathbb{E}\int_0^t\int_{\mathbb{R}}\sum_{j\in\mathcal{S}}\big(F(j)-F\big(\bar{\Lambda}_{s-}\big)\big)\mathbf{1}_{\bar{\Gamma}_{\bar{\Lambda}_{s-}j}}(z)\,\mathrm{d}s\,\mathbf{m}(\mathrm{d}z)\\ &=F\big(\bar{\Lambda}_0\big)+\mathbb{E}\int_0^t\sum_{j\in\mathcal{S}}\bar{q}_{\bar{\Lambda}_{s-}j}\big(F(j)-F\big(\bar{\Lambda}_{s-}\big)\big)\,\mathrm{d}s.\end{aligned}$$

Writing $\bar{P}_tF(i_0)=\mathbb{E}F\big(\bar{\Lambda}_t\big)$ with $\bar{\Lambda}_0=i_0$, we obtain from the above integral equation that $\bar{P}_tF(i_0)=F(i_0)+\int_0^t\bar{P}_s(\bar{Q}F)\,\mathrm{d}s$, where $\bar{Q}F(i)=\sum_{j\in\mathcal{S},j\neq i}\bar{q}_{ij}(F(j)-F(i))$, and the corresponding differential form is

$$\frac{\mathrm{d}\bar{P}_tF}{\mathrm{d}t}=\bar{P}_t\bar{Q}F.\tag{2.10}$$

https://doi.org/10.1017/jpr.2023.54 Published online by Cambridge University Press

Due to (Q1), the Kolmogorov forward equation (2.10) admits a unique solution, and hence $(\bar{\Lambda}_t)$ is a continuous-time Markov chain with transition rate matrix $\bar{Q} = (\bar{q}_{ij})$ (cf. [3, Corollary 2.24]). The corresponding conclusion for (Λ_t^*) can be proved by the same method.

Consequently, through the previous three steps, we have completed the construction of the desired Markov chains (Λ_t^*) and $(\bar{\Lambda}_t)$ used in Theorem 2.1.

Remark 2.3. Our constructed coupling process $(X_t, \Lambda_t, \Lambda_t^*, \bar{\Lambda}_t)$ in terms of a common Poisson random measure presents a good order relation, and is not restricted to be of birth–death type. Let us compare it with the coupling constructed in [4] (presented in the argument of [4, Lemma 3.9]). In [4], the constructed Markov chain (L_t) and the original jumping process (Λ_t) will move independently of each other until the time when they meet. During their meeting time, the coupling is designed such that $\Lambda_t \geq L_t$. After the meeting time, the two processes Λ_t and L_t locate at different states and then they move independently once again until the next meeting time. However, the restriction of jumping in a birth–death type could ensure that $\Lambda_t \geq L_t$ after the meeting time.

2.2. Proofs of the comparison theorem and its application

Let us first present the proof of the comparison theorem.

Proof of Theorem 2.1. We only provide the proof of $\Lambda_t \leq \bar{\Lambda}_t$ for all $t \geq 0$ almost surely; the corresponding solution for Λ_t and Λ_t^* can be proved in the same way.

According to the representation of (2.7) and (2.8), the processes (Λ_t) and $(\bar{\Lambda}_t)$ have no jumps outside $\mathcal{D}_{\mathbf{p}}$. So, we only need to prove that

$$\mathbb{P}\big(\Lambda_t \leq \bar{\Lambda}_t,\ t \in \mathcal{D}_{\mathbf{p}}\big) = 1. \tag{2.11}$$

By the definition of $\bar{q}_{ij}$, for $j > i$, $\bar{q}_{ij} \geq q_{lj}(x)$ for all $1 \leq l \leq i$ and all $x \in \mathbb{R}^d$; for $j < i$, $\bar{q}_{ij} \leq q_{lj}(x)$ for all $j < l \leq i$ and all $x \in \mathbb{R}^d$. By the sort order of the intervals $\Gamma_{ij}(x)$ and $\bar{\Gamma}_{ij}$, we have, for $i < k \in \mathcal{S}$,

$$\bigcup_{r \geq m} \Gamma_{ir}(x) \subset \bigcup_{r \geq m} \bar{\Gamma}_{kr} \qquad \text{for all } m > k \text{ and all } x \in \mathbb{R}^d, \tag{2.12}$$

$$\bigcup_{r \leq m} \Gamma_{ir}(x) \supset \bigcup_{r \leq m} \bar{\Gamma}_{kr} \qquad \text{for all } m < i \text{ and all } x \in \mathbb{R}^d. \tag{2.13}$$

Assuming $i = \Lambda_{\zeta_{n-1}} \leq \bar{\Lambda}_{\zeta_{n-1}} = k$ for some $n \geq 1$, we are going to show that $\Lambda_{\zeta_n} \leq \bar{\Lambda}_{\zeta_n}$, whose proof is divided into four cases.

For case (i), $\Lambda_{\zeta_n} = m \geq k$, by (2.7) and the construction of the Poisson random measure $\mathcal{N}_{\mathbf{p}}(\mathrm{d}t, \mathrm{d}z)$, we get $\xi_n \in \Gamma_{im}(X_{\zeta_n})$. By (2.12), this yields that $\xi_n \in \bigcup_{r \geq m} \bar{\Gamma}_{kr}$. Together with (2.8), this means that $\bar{\Lambda}_{\zeta_n}$ must jump into the set $\{l \in \mathcal{S}; l \geq m\}$. Whence, $\Lambda_{\zeta_n} \leq \bar{\Lambda}_{\zeta_n}$.

For case (ii), $\Lambda_{\zeta_n} = m$ with $i < m < k$, (2.7) implies that $\xi_n \in \Gamma_{im}(X_{\zeta_n})$, and hence $\xi_n < 0$. So, $\xi_n \notin \bigcup_{j \leq k} \bar{\Gamma}_{kj} \subset [0, \infty)$, which means that $\bar{\Lambda}_{\zeta_n}$ cannot jump into the set $\{j \in \mathcal{S}; j \leq k\}$, and hence $\Lambda_{\zeta_n} \leq \bar{\Lambda}_{\zeta_n}$. But, if $\Lambda_{\zeta_n} = m$ with $m \leq i$ and $\bar{\Lambda}_{\zeta_n} \leq i$, this situation is studied in case (iii).

For case (iii), $\Lambda_{\zeta_n} = m$ with $m \leq i$ and $\bar{\Lambda}_{\zeta_n} > i$, it is obvious that $\Lambda_{\zeta_n} \leq \bar{\Lambda}_{\zeta_n}$. If $\bar{\Lambda}_{\zeta_n} = m' \leq i$, (2.8) and (2.13) yield $\xi_n \in \bar{\Gamma}_{km'} \subset \bigcup_{r \leq m'} \Gamma_{ir}(X_{\zeta_n})$. Whence, Λ_{ζ_n} jumps into $\{j \in \mathcal{S}; j \leq m'\}$, and hence $\Lambda_{\zeta_n} \leq \bar{\Lambda}_{\zeta_n}$ still holds.

For case (iv), if $\bar{\Lambda}_{\zeta_n} = m$ with $i < m < k$, then $\xi_n \in \bar{\Gamma}_{km}$, $\xi_n > 0$, and $\xi_n \notin \bigcup_{j > i} \Gamma_{ij}(X_{\zeta_n})$. So, $\Lambda_{\zeta_n} \leq i < m = \bar{\Lambda}_{\zeta_n}$.

https://doi.org/10.1017/jpr.2023.54 Published online by Cambridge University Press

Consequently, if $\Lambda_{\zeta_{n-1}} \le \bar{\Lambda}_{\zeta_{n-1}}$, we have $\Lambda_{\zeta_n} \le \bar{\Lambda}_{\zeta_n}$ for $n \ge 1$. By induction on n, we prove that (2.11) holds, and finally $\mathbb{P}(\Lambda_t \le \bar{\Lambda}_t,\ t \ge 0) = 1$. The proof of Theorem 2.1 is complete. □

Proof of Theorem 2.2. For each $i \in \mathcal{S}$, denote by $\big(X_t^{(i)}\big)$ the solution to the SDE

$$\mathrm{d}X_t^{(i)} = b\big(X_t^{(i)}, i\big)\,\mathrm{d}t + \sigma\big(X_t^{(i)}, i\big)\,\mathrm{d}B_t, \qquad X_0^{(i)} = x,$$

and by $\big(P_t^{(i)}\big)$ its associated semigroup. According to Itô's formula and Gronwall's inequality, it follows from (2.3) that $P_t^{(i)}\rho(x) = \mathbb{E}\rho\big(X_t^{(i)}\big) \le \mathrm{e}^{\beta_i t}\rho(x)$. Let ζ_n, $\mathcal{N}_{\mathbf{p}}(\mathrm{d}t, \mathrm{d}z)$, and $\big(\bar{\Lambda}_t\big)$ be defined as at the beginning of this section. Let $\mathbb{E}^{\mathcal{N}_{\mathbf{p}}}[\cdot] = \mathbb{E}[\,\cdot \mid \mathscr{F}^{\mathcal{N}_{\mathbf{p}}}]$ be the conditional expectation with respect to the σ-algebra $\mathscr{F}^{\mathcal{N}_{\mathbf{p}}} = \sigma\{\mathbf{p}(s); s \ge 0\}$. The mutual independence of $\mathcal{N}_{\mathbf{p}}$ and (B_t) yields

$$\mathbb{E}^{\mathcal{N}_{\mathbf{p}}}\big[\rho\big(X_{\zeta_n}\big)\big] \le \mathbb{E}^{\mathcal{N}_{\mathbf{p}}}\big[\rho\big(X_{\zeta_{n-1}}\big)\big] + \mathbb{E}^{\mathcal{N}_{\mathbf{p}}}\bigg[\int_{\zeta_{n-1}}^{\zeta_n} \beta_{\Lambda_{\zeta_{n-1}}}\rho(X_s)\,\mathrm{d}s\bigg].$$

By Theorem 2.1, $\beta_{\Lambda_s} \le \beta_{\bar{\Lambda}_s}$ a.s. Furthermore, since $(\bar{\Lambda}_s)$ depends only on $\mathcal{N}_{\mathbf{p}}$, we have

$$\mathbb{E}^{\mathcal{N}_{\mathbf{p}}}\big[\rho\big(X_{\zeta_n}\big)\big] \le \mathbb{E}^{\mathcal{N}_{\mathbf{p}}}\big[\rho\big(X_{\zeta_{n-1}}\big)\big] + \beta_{\bar{\Lambda}_{\zeta_{n-1}}}\mathbb{E}^{\mathcal{N}_{\mathbf{p}}}\bigg[\int_{\zeta_{n-1}}^{\zeta_n} \rho(X_s)\,\mathrm{d}s\bigg].$$

Then, as $\bar{\Lambda}_s \equiv \bar{\Lambda}_{\zeta_{n-1}}$ for $s \in [\zeta_{n-1}, \zeta_n)$ by (2.8),

$$\mathbb{E}^{\mathcal{N}_{\mathbf{p}}}\big[\rho\big(X_{\zeta_n}\big)\big] \le \exp\bigg\{\int_{\zeta_{n-1}}^{\zeta_n} \beta_{\bar{\Lambda}_s}\,\mathrm{d}s\bigg\}\mathbb{E}^{\mathcal{N}_{\mathbf{p}}}\big[\rho\big(X_{\zeta_{n-1}}\big)\big].$$

Setting $m_t := \sup\{n \in \mathbb{N}; \zeta_n \le t\}$, deducing recursively, we obtain

$$\begin{aligned}\mathbb{E}^{\mathcal{N}_{\mathbf{p}}}[\rho(X_t)] &\le \exp\bigg\{\int_{\zeta_{m_t}}^{t} \beta_{\bar{\Lambda}_s}\,\mathrm{d}s\bigg\}\mathbb{E}^{\mathcal{N}_{\mathbf{p}}}\big[\rho\big(X_{\zeta_{m_t}}\big)\big] \\ &\le \exp\bigg\{\int_{\zeta_{m_t}}^{t} \beta_{\bar{\Lambda}_s}\,\mathrm{d}s\bigg\}\exp\bigg\{\int_{\zeta_{m_t-1}}^{\zeta_{m_t}} \beta_{\bar{\Lambda}_s}\,\mathrm{d}s\bigg\}\mathbb{E}^{\mathcal{N}_{\mathbf{p}}}\big[\rho\big(X_{\zeta_{m_t-1}}\big)\big] \\ &\le \cdots \le \exp\bigg\{\int_0^t \beta_{\bar{\Lambda}_s}\,\mathrm{d}s\bigg\}\rho(x).\end{aligned}$$

Consequently, $\mathbb{E}[\rho(X_t)] \le \mathbb{E}\big[\exp\big\{\int_0^t \beta_{\bar{\Lambda}_s}\,\mathrm{d}s\big\}\big]\rho(x)$, and further, by Hölder's inequality, for $p' \in (0, 1]$,

$$\mathbb{E}[\rho^{p'}(X_t)] \le \mathbb{E}\bigg[\exp\bigg\{\int_0^t p'\beta_{\bar{\Lambda}_s}\,\mathrm{d}s\bigg\}\bigg]\rho^{p'}(x). \tag{2.14}$$

According to [1, Propositions 4.1, 4.2], (2.4) yields that there exist $p' \in (0, 1]$, $C, \eta > 0$ such that $\mathbb{E}\big[\exp\big\{\int_0^t p'\beta_{\bar{\Lambda}_s}\,\mathrm{d}s\big\}\big] \le C\mathrm{e}^{-\eta t}$. Combining this with (2.3) and (2.14), the desired conclusion follows immediately. □

https://doi.org/10.1017/jpr.2023.54 Published online by Cambridge University Press

3. Perturbation of continuous-time Markov chain

Given two transition rate matrices $Q=(q_{ij})_{i,j\in\mathcal{S}}$ and $\widetilde{Q}=(\tilde{q}_{ij})_{i,j\in\mathcal{S}}$ on $\mathcal{S}=\{1,2,\ldots,N\}$, $2\le N\le\infty$, there are two continuous-time Markov chains (Λ_t) and $(\tilde{\Lambda}_t)$ associated with Q and $\widetilde{Q}$ respectively with $\Lambda_0=\tilde{\Lambda}_0$; the purpose of this section is to estimate the quantity

$$\frac{1}{t}\int_0^t \mathbb{P}\big(\Lambda_s\neq\tilde{\Lambda}_s\big)\,\mathrm{d}s,\qquad t>0, \tag{3.1}$$

in terms of the difference between Q and $\widetilde{Q}$. The quantity in (3.1) plays an important role in the study of regime-switching processes. For instance, in [23], it is the key point to show the stability of the process (X_t) under the perturbation of Q. In [20], it is the basis for proving the Euler–Maruyama approximation of state-dependent regime-switching processes. Also, as mentioned in the introduction, it was used in [24] to study the smooth dependence of initial values for state-dependent RSPs. In this section, we improve the estimate of the quantity in (3.1), and apply it to develop the perturbation theory associated with regime-switching processes.

In the classical perturbation theory of Markov chains, there has been a lot of research on an upper estimate of the total variation distance $\|P_t-\widetilde{P}_t\|_{\mathrm{var}}$ between the semigroups P_t and $\widetilde{P}_t$ associated with (Λ_t) and $(\tilde{\Lambda}_t)$ respectively; see, e.g., [5, 14–16] and references therein. For instance, [14, 15] showed that

$$\|P_t-\widetilde{P}_t\|_{\mathrm{var}}\le\frac{\mathrm{e}\tau_1}{\mathrm{e}-1}\|Q-\widetilde{Q}\|_{\ell_1}, \tag{3.2}$$

where $\beta(t)=\frac{1}{2}\max_{i,j\in\mathcal{S}}\|(e_i-e_j)\exp(tQ)\|_{\ell_1}$, $\tau_1=\inf\{t>0;\,\beta(t)\le\mathrm{e}^{-1}\}$. Recall that the ℓ_1-norm of a matrix $A=(a_{ij})_{i,j\in\mathcal{S}}$ is defined by $\|A\|_{\ell_1}=\sup_{i\in\mathcal{S}}\sum_{j\in\mathcal{S}}|a_{ij}|$, and the total variation distance between any two probability measures μ and ν on $\mathcal{S}$ is defined by $\|\mu-\nu\|_{\mathrm{var}}=\sup_{|h|\le1}\big|\sum_{i\in\mathcal{S}}h_i(\mu_i-\nu_i)\big|$. However, the estimate in (3.2) and the method of establishing it is not applicable to (3.1). Moreover, notice that $\|P_t-\widetilde{P}_t\|_{\mathrm{var}}=2\inf\mathbb{P}\big(\xi\neq\tilde{\xi}\big)\le2\mathbb{P}(\Lambda_t\neq\tilde{\Lambda}_t)$, where the infimum is over all couplings $(\xi,\tilde{\xi})$ with marginal distributions P_t and $\widetilde{P}_t$ respectively. The method in [5, 14, 15] cannot be extended to deal with (3.1) because the total variation norm plays an essential role in establishing (3.2). In [23], we provided an estimate of (3.1) through constructing a coupling process $(\Lambda_t,\tilde{\Lambda}_t)$ using Skorokhod's representation, where the intervals were constructed as in [8]. Consequently, [23] can only cope with jumping processes in a finite state space and the estimate obtained is not satisfactory, especially for large t.

In this section we use the following assumptions:

(H1) $K_0:=\sup\{q_i,\tilde{q}_j;\,i,j\in\mathcal{S}\}<\infty$.

(H2) There exists a $c_0\in\mathbb{N}$ such that $q_{ij}=\tilde{q}_{ij}=0$ for all $i,j\in\mathcal{S}$ with $|j-i|>c_0$.

Theorem 3.1. *Assume that (H1) and (H2) hold. Then there exist processes (Λ_t) and $(\tilde{\Lambda}_t)$ such that, for all $t>0$,*

$$\frac{1}{t}\int_0^t\mathbb{P}\big(\Lambda_s\neq\tilde{\Lambda}_s\big)\,\mathrm{d}s\le1-\frac{1}{t\|Q-\widetilde{Q}\|_{\ell_1}}\Big(1-\mathrm{e}^{-\|Q-\widetilde{Q}\|_{\ell_1}t}\Big), \tag{3.3}$$

which implies that

$$\frac{1}{t}\int_0^t\mathbb{P}\big(\Lambda_s\neq\tilde{\Lambda}_s\big)\,\mathrm{d}s\le\min\left\{\frac{1}{2}\|Q-\widetilde{Q}\|_{\ell_1}t,\,1\right\}. \tag{3.4}$$

https://doi.org/10.1017/jpr.2023.54 Published online by Cambridge University Press

Also, for all $t > 0$,

$$\frac{1}{t}\int_0^t \mathbb{P}\big(\Lambda_s \neq \tilde{\Lambda}_s\big)\,\mathrm{d}s \geq \frac{\inf_{i\in\mathcal{S}} \sum_{j\neq i} |q_{ij} - \tilde{q}_{ij}|}{M + \|Q - \widetilde{Q}\|_{\ell_1}} \left(1 - \frac{1}{\big(M + \|Q - \widetilde{Q}\|_{\ell_1}\big)t}\Big(1 - \mathrm{e}^{-\big(M + \|Q-\widetilde{Q}\|_{\ell_1}\big)t}\Big)\right),$$

where $M = 4c_0K_0$.

Remark 3.1. In [23, Lemma 2.2], given two transition rate matrices Q and $\widetilde{Q}$ on a finite state space $\mathcal{S} = \{1, 2, \ldots, N\}$, a coupling process $(\Lambda_t, \tilde{\Lambda}_t)$ was constructed that satisfies

$$\frac{1}{t}\int_0^t \mathbb{P}\big(\Lambda_s \neq \tilde{\Lambda}_s\big)\,\mathrm{d}s \leq N^2 t \|Q - \widetilde{Q}\|_{\ell_1}. \tag{3.5}$$

There is an important drawback in (3.5): the appearance of N^2 on the right-hand side, which restricts the application of this result to Markov chains on an infinite state space. This drawback has been removed in Theorem 3.1, and, further, a lower estimate of $\frac{1}{t}\int_0^t \mathbb{P}\big(\Lambda_s \neq \tilde{\Lambda}_s\big)\,\mathrm{d}s$ is provided in the current work.

3.1. Construction of the coupling process

In this part, we introduce the coupling process $(\Lambda_t, \tilde{\Lambda}_t)$ on $\mathcal{S}\times\mathcal{S}$ that will be used in the proof of Theorem 3.1. Similarly to Section 2, it is also constructed using the Skorokhod representation theorem. However, there are several subtle differences to fit the current purpose.

Step 1: Construction of intervals Due to (H1) and (H2), we let

$$\Gamma_{1k} = [(k-2)K_0, (k-2)K_0 + q_{1k}), \qquad \widetilde{\Gamma}_{1k} = [(k-2)K_0, (k-2)K_0 + \tilde{q}_{1k})$$

for $2 \leq k \leq c_0 + 1$, and $U_1 = [0, c_0K_0)$. By (H1), $q_{1k} \leq K_0$ and $\tilde{q}_{1k} \leq K_0$, and hence $\Gamma_{1k} \bigcap \Gamma_{1j} = \emptyset$ and $\widetilde{\Gamma}_{1k} \bigcap \widetilde{\Gamma}_{1j} = \emptyset$, $k \neq j$. Moreover, $\Gamma_{1k} \subset U_1$ and $\widetilde{\Gamma}_{1k} \subset U_1$ for all $2 \leq k \leq c_0 + 1$. For $n \geq 2$, define

$$\Gamma_{nk} = \begin{cases} \big[2(n-1)c_0K_0 - (n-k)K_0,\ 2(n-1)c_0K_0 - (n-k)K_0 + q_{nk}\big), & n - c_0 \leq k < n, \\ \big[2(n-1)c_0K_0 + (k-n-1)K_0,\ 2(n-1)c_0K_0 + (k-n-1)K_0 + q_{nk}\big), & n + c_0 \geq k > n. \end{cases}$$

Define, similarly, $\widetilde{\Gamma}_{nk}$ by replacing q_{nk} above with $\tilde{q}_{nk}$. Let $U_n = \big[(2n-3)c_0K_0, (2n-1)c_0K_0\big)$, $n \geq 2$. So $\Gamma_{nk} \subset U_n$ and $\widetilde{\Gamma}_{nk} \subset U_n$ for all $k, n \in \mathcal{S}$ with $|k - n| \leq c_0$.

Compared with the intervals constructed in [8], the starting point of each interval Γ_{nk} constructed above does not depend on any other intervals Γ_{ij}, $i, j \in \mathcal{S}$ with $i \neq n$, $j \neq k$. We construct Γ_{nk} in this way in order to remove the term N^2 that appears in (3.5).

Step 2: Construction of Poisson random measure Denote by $\mathbf{m}(\mathrm{d}x)$ the Lebesgue measure over the real line $\mathbb{R}$. Let $\xi_i^{(k)}$, $k, i = 1, 2, \ldots$, be U_k-valued random variables with $\mathbb{P}\big(\xi_i^{(k)} \in \mathrm{d}x\big) = \mathbf{m}(\mathrm{d}x)/\mathbf{m}(U_k)$, and $\tau_i^{(k)}$, $k, i = 1, 2, \ldots$, be non-negative random variables such that $\mathbb{P}\big(\tau_i^{(k)} > t\big) = \exp[-t\mathbf{m}(U_k)]$, $t \geq 0$. Assume all $\xi_i^{(k)}$, $\tau_i^{(k)}$ to be mutually independent. Write $\zeta_n^{(k)} = \tau_1^{(k)} + \cdots + \tau_n^{(k)}$ for $k, n \geq 1$ and $\zeta_0^{(k)} = 0$ for $k \geq 1$. We define $\mathcal{D}_{\tilde{\mathbf{p}}} = \bigcup_{k\geq 1}\bigcup_{n\geq 0}\big\{\zeta_n^{(k)}\big\}$ and $\tilde{\mathbf{p}}(t) = \sum_{0\leq s<t} \Delta\tilde{\mathbf{p}}(s)$, $\Delta\tilde{\mathbf{p}}(s) = 0$ for $s \notin \mathcal{D}_{\tilde{\mathbf{p}}}$, and $\Delta\tilde{\mathbf{p}}\big(\zeta_n^{(k)}\big) = \xi_n^{(k)}$, $k, n = 1, 2, \ldots$ Let

$$\mathcal{N}_{\tilde{\mathbf{p}}}([0, t] \times A) = \#\big\{s \in \mathcal{D}_{\tilde{\mathbf{p}}};\, 0 < s \leq t,\ \tilde{\mathbf{p}}(s) \in A\big\}, \qquad t > 0,\ A \in \mathscr{B}([0, \infty)).$$

As a consequence, we get a Poisson random process $(\tilde{\mathbf{p}}(t))$ and its associated Poisson random measure $\mathcal{N}_{\tilde{\mathbf{p}}}(\mathrm{d}t, \mathrm{d}x)$ with intensity measure $\mathrm{d}t\,\mathbf{m}(\mathrm{d}x)$.

The construction of $(\tilde{\mathbf{p}}(t))$ is more complicated than that of $(\mathbf{p}(t))$ in Section 2, which is caused by the fact that the union of Γ_{nk} for $n, k = 1, 2, \ldots$ may be unbounded.

Step 3: Construction of coupling processes Define two functions ϑ, $\tilde{\vartheta}$ associated with Γ_{ij} and $\widetilde{\Gamma}_{ij}$, $i, j \in \mathcal{S}$, by

$$\vartheta(i, z) = \sum_{j\in\mathcal{S}, j\neq i} (j-i)\mathbf{1}_{\Gamma_{ij}}(z), \qquad \tilde{\vartheta}(i, z) = \sum_{j\in\mathcal{S}, j\neq i} (j-i)\mathbf{1}_{\widetilde{\Gamma}_{ij}}(z). \tag{3.6}$$

Then, the desired coupling process $(\Lambda_t, \tilde{\Lambda}_t)$ is given as the solution of the following SDEs:

$$\mathrm{d}\Lambda_t = \int_{[0,\infty)} \vartheta(\Lambda_{t-}, z)\,\mathcal{N}_{\tilde{\mathbf{p}}}(\mathrm{d}t, \mathrm{d}z), \qquad \Lambda_0 = i_0 \in \mathcal{S}, \tag{3.7}$$

$$\mathrm{d}\tilde{\Lambda}_t = \int_{[0,\infty)} \tilde{\vartheta}(\tilde{\Lambda}_{t-}, z)\,\mathcal{N}_{\tilde{\mathbf{p}}}(\mathrm{d}t, \mathrm{d}z), \qquad \tilde{\Lambda}_0 = i_0 \in \mathcal{S}. \tag{3.8}$$

Remark 3.2. The coupling process $(\Lambda_t, \tilde{\Lambda}_t)$ constructed above is different from the basic coupling of (Λ_t) and $\big(\tilde{\Lambda}_t\big)$ (see [3, p. 11]). Indeed, the transition rate matrix $\bar{Q} = (\bar{q}_{(ij)(k\ell)})_{i,j,k,\ell\in\mathcal{S}}$ of $(\Lambda_t, \tilde{\Lambda}_t)$ is given as follows: for $i, j, k, r \in \mathcal{S}$, which are different from each other, $\bar{q}_{(ij)(kr)} = 0$, $\bar{q}_{(ij)(ik)} = \tilde{q}_{jk}$, $\bar{q}_{(ij)(kj)} = q_{ik}$, $\bar{q}_{(ii)(jj)} = q_{ij} \wedge \tilde{q}_{ij}$, $\bar{q}_{(ii)(ji)} = (q_{ij} - \tilde{q}_{ij}) \vee 0$, $\bar{q}_{(ii)(ij)} = (\tilde{q}_{ij} - q_{ij}) \vee 0$, and $\bar{q}_{(ii)(ii)} = -\sum_{(\ell,\ell')\neq(i,i)} \bar{q}_{(ii)(\ell\ell')}$.

The transition rate matrix of the basic coupling is given by $q_{(ij)(kj)} = (q_{ik} - \tilde{q}_{jk}) \vee 0$, $q_{(ij)(ik)} = (\tilde{q}_{jk} - q_{ik}) \vee 0$, and $q_{(ij)(kk)} = q_{ik} \wedge \tilde{q}_{jk}$, for all $i, j, k \in \mathcal{S}$, and i,j not necessarily different.

By the previous construction of Γ_{ij}, $\widetilde{\Gamma}_{ij}$, and $\mathcal{N}_p(\mathrm{d}t, \mathrm{d}x)$, the following properties hold:

(i) The processes (Λ_t) and $\big(\tilde{\Lambda}_t\big)$ can jump only when the Poisson process $(\tilde{\mathbf{p}}(t))$ jumps.

(ii) If $\Lambda_t = \tilde{\Lambda}_t = k$, for $\delta > 0$, $\Lambda_{t+\delta} \neq \tilde{\Lambda}_{t+\delta}$ may happen only when $\zeta_n^{(k)} \in [t, t+\delta)$ and $\xi_n^{(k)} \in \Gamma_{kj}\Delta\widetilde{\Gamma}_{kj}$ for some $n \geq 1$, where $A\Delta B := (A\backslash B) \cup (B\backslash A)$ for Borel sets A, B.

3.2. Proof of Theorem 3.1 and its application

Proof of Theorem 3.1. To get the upper estimate, for $\delta > 0$, let

$$\alpha(\delta) = \sup_{i\in\mathcal{S}} \mathbb{P}(\Lambda_\delta \neq \tilde{\Lambda}_\delta \mid \Lambda_0 = \tilde{\Lambda}_0 = i), \qquad \beta(\delta) = 1 - \alpha(\delta).$$

Now we use the representations in (3.7) and (3.8) of (Λ_t) and $\big(\tilde{\Lambda}_t\big)$ to estimate $\alpha(\delta)$. Noting that $\Lambda_0 = \tilde{\Lambda}_0 = i_0$ for some $i_0 \in \mathcal{S}$, by (3.6), (3.7), and (3.8) we have

$$\begin{aligned}\mathbb{P}\big(\Lambda_\delta \neq \tilde{\Lambda}_\delta\big) &= \mathbb{P}\big(\Lambda_\delta \neq \tilde{\Lambda}_\delta \mid \Lambda_0 = \tilde{\Lambda}_0\big) = \mathbb{P}\big(\Lambda_\delta \neq \tilde{\Lambda}_\delta, \mathcal{N}_{\tilde{\mathbf{p}}}([0, \delta] \times U_{i_0}) \geq 1\big) \\ &= \mathbb{P}\big(\Lambda_\delta \neq \tilde{\Lambda}_\delta, \mathcal{N}_{\tilde{\mathbf{p}}}\big([0, \delta] \times U_{i_0}\big) = 1\big) + \mathbb{P}\big(\Lambda_\delta \neq \tilde{\Lambda}_\delta, \mathcal{N}_{\tilde{\mathbf{p}}}([0, \delta] \times U_{i_0}) \geq 2\big).\end{aligned} \tag{3.9}$$

Since $H := \mathbf{m}(U_{i_0}) = 2c_0K_0 < \infty$, there exists a $C > 0$ independent of the choice of $i_0 \in \mathcal{S}$ such that

$$\mathbb{P}\big(\mathcal{N}_{\tilde{\mathbf{p}}}([0, \delta] \times U_{i_0}) \geq 2\big) = 1 - \mathrm{e}^{-H\delta} - H\delta\mathrm{e}^{-H\delta} \leq C\delta^2.$$

Moreover,

$$\begin{aligned}\mathbb{P}\big(\Lambda_\delta \neq \tilde{\Lambda}_\delta, \mathcal{N}_{\tilde{\mathbf{p}}}([0,\delta]\times U_{i_0})=1\big) &= \int_0^\delta \mathbb{P}\Big(\Lambda_\delta \neq \tilde{\Lambda}_\delta,\, \tau_1^{(i_0)} \in \mathrm{d}s,\, \tau_2^{(i_0)} > \delta - s\Big)\\ &= \int_0^\delta \mathbb{P}\Big(\xi_1^{(i_0)} \in \cup_{j\in\mathcal{S}}\big(\Gamma_{i_0 j}\Delta\widetilde{\Gamma}_{i_0 j}\big),\, \tau_1^{(i_0)} \in \mathrm{d}s,\, \tau_2^{(i_0)} > \delta - s\Big)\\ &= \delta \mathrm{e}^{-H\delta} \sum_{j\in\mathcal{S}, j\neq i_0} |q_{i_0 j} - \tilde{q}_{i_0 j}|\\ &\le \delta \mathrm{e}^{-H\delta}\big\|Q-\widetilde{Q}\big\|_{\ell_1} \le \delta\big\|Q-\widetilde{Q}\big\|_{\ell_1}. \end{aligned} \tag{3.10}$$

In the light of (3.9) and (3.10), we get

$$\mathbb{P}\big(\Lambda_\delta \neq \tilde{\Lambda}_\delta\big) = \mathbb{P}(\Lambda_\delta \neq \tilde{\Lambda}_\delta \mid \Lambda_0 = \tilde{\Lambda}_0 = i_0) \le \delta\big\|Q-\widetilde{Q}\big\|_{\ell_1} + C\delta^2, \tag{3.11}$$

and hence

$$\alpha(\delta) \le \delta\big\|Q-\widetilde{Q}\big\|_{\ell_1} + C\delta^2, \qquad \beta(\delta) = 1-\alpha(\delta) \ge 1 - \delta\big\|Q-\widetilde{Q}\big\|_{\ell_1} - C\delta^2. \tag{3.12}$$

Next, let us consider the time 2δ.

$$\begin{aligned}\mathbb{P}\big(\Lambda_{2\delta} \neq \tilde{\Lambda}_{2\delta}\big) &= \mathbb{P}\big(\Lambda_{2\delta} \neq \tilde{\Lambda}_{2\delta},\, \Lambda_\delta = \tilde{\Lambda}_\delta\big) + \mathbb{P}\big(\Lambda_{2\delta} \neq \tilde{\Lambda}_{2\delta},\, \Lambda_\delta \neq \tilde{\Lambda}_\delta\big)\\ &= \mathbb{P}\big(\Lambda_{2\delta} \neq \tilde{\Lambda}_{2\delta} \mid \Lambda_\delta = \tilde{\Lambda}_\delta\big)\mathbb{P}\big(\Lambda_\delta = \tilde{\Lambda}_\delta\big) + \mathbb{P}\big(\Lambda_{2\delta} \neq \tilde{\Lambda}_{2\delta} \mid \Lambda_\delta \neq \tilde{\Lambda}_\delta\big)\mathbb{P}\big(\Lambda_\delta \neq \tilde{\Lambda}_\delta\big)\\ &\le \alpha(\delta)\mathbb{P}\big(\Lambda_\delta = \tilde{\Lambda}_\delta\big) + \mathbb{P}\big(\Lambda_\delta \neq \tilde{\Lambda}_\delta\big)\\ &= \alpha(\delta) + (1-\alpha(\delta))\mathbb{P}\big(\Lambda_\delta \neq \tilde{\Lambda}_\delta\big)\\ &\le \alpha(\delta)(1+\beta(\delta)),\end{aligned}$$

where we have used the time homogeneity of Markov chain $(\Lambda_t, \tilde{\Lambda}_t)$ to get the estimate $\mathbb{P}\big(\Lambda_{2\delta} \neq \tilde{\Lambda}_{2\delta} \mid \Lambda_\delta = \tilde{\Lambda}_\delta\big) \le \alpha(\delta)$. Analogously, using the time homogeneity of (Λ_t) and $(\tilde{\Lambda}_t)$,

$$\begin{aligned}\mathbb{P}\big(\Lambda_{3\delta} \neq \tilde{\Lambda}_{3\delta}\big) &\le \mathbb{P}\big(\Lambda_{3\delta} \neq \tilde{\Lambda}_{3\delta} \mid \Lambda_{2\delta} = \tilde{\Lambda}_{2\delta}\big)\mathbb{P}\big(\Lambda_{2\delta} = \tilde{\Lambda}_{2\delta}\big) + \mathbb{P}\big(\Lambda_{2\delta} \neq \tilde{\Lambda}_{2\delta}\big)\\ &\le \alpha(\delta) + \beta(\delta)\mathbb{P}\big(\Lambda_{2\delta} \neq \tilde{\Lambda}_{2\delta}\big)\\ &\le \alpha(\delta)\big(1+\beta(\delta)+\beta(\delta)^2\big) = 1-\beta(\delta)^3.\end{aligned}$$

Deducing inductively, we get

$$\mathbb{P}\big(\Lambda_{k\delta} \neq \tilde{\Lambda}_{k\delta}\big) \le \alpha(\delta)\sum_{m=1}^{k}\beta(\delta)^{m-1} = 1-\beta(\delta)^k \qquad \text{for } k \ge 4.$$

https://doi.org/10.1017/jpr.2023.54 Published online by Cambridge University Press

Therefore, for $t>0$, setting $K(t)=[t/\delta]$, $t_k=k\delta$ for $k\le K$, and $t_{K+1}=t$, by (3.12),

$$\begin{aligned}\int_0^t \mathbb{P}\big(\Lambda_s\neq\tilde\Lambda_s\big)\mathrm{d}s &\le \sum_{k=0}^{K(t)}\int_{t_k}^{t_{k+1}}\big(\mathbb{P}\big(\Lambda_s\neq\tilde\Lambda_s,\,\Lambda_{k\delta}=\tilde\Lambda_{k\delta}\big)+\mathbb{P}\big(\Lambda_s\neq\tilde\Lambda_s,\,\Lambda_{k\delta}\neq\tilde\Lambda_{k\delta}\big)\big)\,\mathrm{d}s\\ &\le \sum_{k=0}^{K(t)}\int_{t_k}^{t_{k+1}}\mathbb{P}\big(\Lambda_s\neq\tilde\Lambda_s\mid\Lambda_{k\delta}\neq\tilde\Lambda_{k\delta}\big)\,\mathrm{d}s+\sum_{k=0}^{K(t)}\int_{t_k}^{t_{k+1}}\mathbb{P}\big(\Lambda_{k\delta}\neq\tilde\Lambda_{k\delta}\big)\,\mathrm{d}s\\ &\le \alpha(\delta)t+\delta\left(K(t)+1-\sum_{k=0}^{K(t)}\beta(\delta)^k\right)\\ &\le \alpha(\delta)t+\delta(K(t)+1)-\delta\sum_{k=0}^{K(t)}\big(1-\delta\|Q-\widetilde Q\|_{\ell_1}-C\delta^2\big)^k\\ &\le \alpha(\delta)t+\delta(K(t)+1)-\frac{1-\big(1-\delta\|Q-\widetilde Q\|_{\ell_1}-C\delta^2\big)^{K(t)+1}}{\|Q-\widetilde Q\|_{\ell_1}+C\delta}.\end{aligned}$$

Letting $\delta\downarrow 0$ and then dividing both sides by t, we get the upper estimate

$$\frac1t\int_0^t\mathbb{P}\big(\Lambda_s\neq\tilde\Lambda_s\big)\,\mathrm{d}s\le 1-\frac{1}{t\|Q-\widetilde Q\|_{\ell_1}}\Big(1-\mathrm{e}^{-\|Q-\widetilde Q\|_{\ell_1}t}\Big)\le 1.$$

Using the inequality $\mathrm{e}^{-x}\le 1-x+\frac12x^2$ for $x\ge 0$, we further get

$$\frac1t\int_0^t\mathbb{P}\big(\Lambda_s\neq\tilde\Lambda_s\big)\,\mathrm{d}s\le\min\left\{\frac12\|Q-\widetilde Q\|_{\ell_1}t,\,1\right\}.$$

Therefore, the upper estimates (3.3) and (3.4) have been proved.

For the lower estimate, We will estimate the difference by induction. Due to (3.9),

$$\begin{aligned}\mathbb{P}\big(\Lambda_\delta\neq\tilde\Lambda_\delta\big)&=\mathbb{P}\big(\Lambda_\delta\neq\tilde\Lambda_\delta\mid\Lambda_0=\tilde\Lambda_0\big)\\ &\ge\mathbb{P}\big(\Lambda_\delta\neq\tilde\Lambda_\delta,\,N_{\tilde{\mathbf p}}\big([0,\delta]\times U_{i_0}\big)=1\big)\\ &=\delta\mathrm{e}^{-H\delta}\sum_{j\neq i}|q_{i_0j}-\tilde q_{i_0j}|\ge\delta\mathrm{e}^{-H\delta}\inf_{i\in\mathcal S}\sum_{j\neq i}|q_{ij}-\tilde q_{ij}|\;=:\tilde\alpha(\delta).\end{aligned}$$

Then,

$$\begin{aligned}\mathbb{P}\big(\Lambda_{2\delta}\neq\tilde\Lambda_{2\delta}\big)&=\mathbb{P}\big(\Lambda_{2\delta}\neq\tilde\Lambda_{2\delta},\,\Lambda_\delta=\tilde\Lambda_\delta\big)+\mathbb{P}\big(\Lambda_{2\delta}\neq\tilde\Lambda_{2\delta},\,\Lambda_\delta\neq\tilde\Lambda_\delta\big)\\ &=\mathbb{P}\big(\Lambda_{2\delta}\neq\tilde\Lambda_{2\delta}\mid\Lambda_\delta=\tilde\Lambda_\delta\big)\big(1-\mathbb{P}\big(\Lambda_\delta\neq\tilde\Lambda_\delta\big)\big)\\ &\quad+\mathbb{P}\big(\Lambda_\delta\neq\tilde\Lambda_\delta\big)-\mathbb{P}\big(\Lambda_{2\delta}=\tilde\Lambda_{2\delta}\mid\Lambda_\delta\neq\tilde\Lambda_\delta\big)\mathbb{P}\big(\Lambda_\delta\neq\tilde\Lambda_\delta\big)\\ &\ge\mathbb{P}\big(\Lambda_{2\delta}\neq\Lambda_{2\delta}\mid\Lambda_\delta=\tilde\Lambda_\delta\big)\big(1-\mathbb{P}\big(\Lambda_\delta\neq\tilde\Lambda_\delta\big)\big)+\mathbb{P}\big(\Lambda_\delta\neq\tilde\Lambda_\delta\big)\\ &\quad-\mathbb{P}\Big(\text{there exist jumps for }(\tilde{\mathbf p}(t))\text{ during }[\delta,2\delta)\text{ localizing in }U_{\Lambda_\delta}\cup U_{\tilde\Lambda_\delta}\Big)\\ &\mathbb{P}\big(\Lambda_\delta\neq\tilde\Lambda_\delta\big)\\ &\ge\mathbb{P}\big(\Lambda_{2\delta}\neq\Lambda_{2\delta}\mid\Lambda_\delta=\tilde\Lambda_\delta\big)\big(1-\mathbb{P}\big(\Lambda_\delta\neq\tilde\Lambda_\delta\big)\big)\\ &\quad+\mathbb{P}\big(\Lambda_\delta\neq\tilde\Lambda_\delta\big)-\big(1-\mathrm{e}^{-M\delta}\big)\mathbb{P}\big(\Lambda_\delta\neq\tilde\Lambda_\delta\big)\\ &\ge\tilde\alpha(\delta)(1-\alpha(\delta))+\mathrm{e}^{-M\delta}\tilde\alpha(\delta)\\ &\ge\tilde\alpha(\delta)+\big(\mathrm{e}^{-M\delta}-\delta\|Q-\widetilde Q\|_{\ell_1}-C\delta^2\big)\tilde\alpha(\delta),\end{aligned}$$

https://doi.org/10.1017/jpr.2023.54 Published online by Cambridge University Press

where we have used (3.11) and $\mathbb{P}\big(N_{\tilde{\mathbf{p}}}\big([\delta, 2\delta]\times\big(U_{\Lambda_\delta}\cup U_{\tilde{\Lambda}_\delta}\big)\big)\geq 1\big)\leq 1-\mathrm{e}^{-M\delta}$, since $\mathbf{m}(U_i\cup U_j)\leq M=4c_0K_0$ for all $i, j\in\mathcal{S}$. Setting $\gamma(\delta)=\mathrm{e}^{-M\delta}-\delta\big\|Q-\widetilde{Q}\big\|_{\ell_1}-C\delta^2$, which is positive when δ is small enough, we rewrite the previous estimate in the form $\mathbb{P}\big(\Lambda_{2\delta}\neq\tilde{\Lambda}_{2\delta}\big)\geq\tilde{\alpha}(\delta)+\gamma(\delta)\tilde{\alpha}(\delta)$. Repeating this procedure, we obtain

$$\mathbb{P}(\Lambda_{k\delta}\neq\tilde{\Lambda}_{k\delta})\geq\tilde{\alpha}(\delta)\sum_{m=1}^{k}\gamma(\delta)^{m-1},\qquad k\geq 3.$$

Then, from this, for $K\in\mathbb{N}$,

$$\begin{aligned}\int_0^{K\delta}\mathbb{P}\big(\Lambda_s\neq\tilde{\Lambda}_s\big)\,\mathrm{d}s&=\sum_{k=0}^{K}\int_{k\delta}^{(k+1)\delta}\big(\mathbb{P}\big(\Lambda_s\neq\tilde{\Lambda}_s,\,\Lambda_{k\delta}\neq\tilde{\Lambda}_{k\delta}\big)+\mathbb{P}\big(\Lambda_s\neq\tilde{\Lambda}_s,\,\Lambda_{k\delta}=\tilde{\Lambda}_{k\delta}\big)\big)\,\mathrm{d}s\\&\geq\sum_{k=0}^{K}\int_{k\delta}^{(k+1)\delta}\big(\mathbb{P}\big(\Lambda_{k\delta}\neq\tilde{\Lambda}_{k\delta}\big)-\mathbb{P}\big(\Lambda_{k\delta}\neq\tilde{\Lambda}_{k\delta},\,\Lambda_s=\tilde{\Lambda}_s\big)\big)\,\mathrm{d}s\\&\geq\delta\sum_{k=1}^{K}\mathbb{P}\big(\Lambda_{k\delta}\neq\tilde{\Lambda}_{k\delta}\big)-\delta\big(1-\mathrm{e}^{-M\delta}\big)\sum_{k=1}^{K}\mathbb{P}\big(\Lambda_{k\delta}\neq\tilde{\Lambda}_{k\delta}\big)\\&\geq\delta\mathrm{e}^{-M\delta}\sum_{k=1}^{K}\tilde{\alpha}(\delta)\frac{1-\gamma(\delta)^k}{1-\gamma(\delta)}\\&=\mathrm{e}^{-M\delta}\frac{\delta\tilde{\alpha}(\delta)}{1-\gamma(\delta)}\left(K-\frac{\gamma(\delta)\big(1-\gamma(\delta)^K\big)}{1-\gamma(\delta)}\right).\end{aligned}$$

Since

$$\lim_{\delta\downarrow 0}\frac{\tilde{\alpha}(\delta)}{1-\gamma(\delta)}=\frac{\inf_{i\in\mathcal{S}}\sum_{j\neq i}|q_{ij}-\tilde{q}_{ij}|}{M+\big\|Q-\widetilde{Q}\big\|_{\ell_1}},\qquad\lim_{\delta\downarrow 0}\gamma(\delta)^{\frac{t}{\delta}}=\mathrm{e}^{-\left(M+\left\|Q-\widetilde{Q}\right\|_{\ell_1}\right)t},$$

by taking $K=[t/\delta]$ in the previous estimation and letting δ tend downward to 0, we finally get

$$\int_0^t\mathbb{P}\big(\Lambda_s\neq\tilde{\Lambda}_s\big)\,\mathrm{d}s\geq\frac{\inf_{i\in\mathcal{S}}\sum_{j\neq i}|q_{ij}-\tilde{q}_{ij}|}{M+\big\|Q-\widetilde{Q}\big\|_{\ell_1}}\left(t-\frac{1}{M+\big\|Q-\widetilde{Q}\big\|_{\ell_1}}\left(1-\mathrm{e}^{-\left(M+\left\|Q-\widetilde{Q}\right\|_{\ell_1}\right)t}\right)\right),$$

which is the desired lower estimate. The proof is complete. □

Next, we consider the application of Theorem 3.1. First, let us consider its application to perturbation theory on the invariant probability measures of continuous-time Markov chains on infinite state spaces. There have been many works on perturbation theory of Markov chains, such as [16, 29] and references therein. We refer readers to the recent review paper [30] for more discussions on this topic.

Let P_t and $\widetilde{P}_t$ denote the semigroups with respect to the transition rate matrices Q and $\widetilde{Q}$ respectively. Assume that there exist invariant probability measures $\pi=(\pi_i)_{i\in\mathcal{S}}$ and $\tilde{\pi}=(\tilde{\pi}_i)_{i\in\mathcal{S}}$ associated respectively with P_t and $\widetilde{P}_t$, i.e. $\pi P_t=\pi$, $\tilde{\pi}\widetilde{P}_t=\tilde{\pi}$.

https://doi.org/10.1017/jpr.2023.54 Published online by Cambridge University Press

Corollary 3.1. *Assume (H1) and (H2) hold. Suppose that there exists a function* $\eta : [0, \infty) \to [0, 2]$ *satisfying* $\int_0^\infty \eta_s \, \mathrm{d}s < \infty$ *such that, for some* $i_0 \in \mathcal{S}$, $\|P_t(i_0, \cdot) - \pi\|_{\mathrm{var}} \le \eta_t$ *and* $\|\widetilde{P}_t(i_0, \cdot) - \tilde{\pi}\|_{\mathrm{var}} \le \eta_t$ *for* $t \ge 0$*. Then,*

$$\|\pi - \tilde{\pi}\|_{\mathrm{var}} \le 2\sqrt{2}\bigg(\int_0^\infty \eta_s \, \mathrm{d}s\bigg)^{1/2} \sqrt{\|Q - \widetilde{Q}\|_{\ell_1}}.$$

Proof. For any bounded function h on $\mathcal{S}$ with $|h|_\infty := \sup_{i \in \mathcal{S}} |h_i| \le 1$,

$$\begin{aligned}
|\pi(h) - \tilde{\pi}(h)| &= \bigg|\sum_{i\in\mathcal{S}} \pi_i h_i - \sum_{i\in\mathcal{S}} \tilde{\pi}_i h_i\bigg| \\
&\le \bigg|\pi(h) - \frac{1}{t}\int_0^t P_s h(i_0)\,\mathrm{d}s\bigg| + \bigg|\tilde{\pi}(h) - \frac{1}{t}\int_0^t \widetilde{P}_s h(i_0)\,\mathrm{d}s\bigg| \\
&\quad + \bigg|\frac{1}{t}\int_0^t P_s h(i_0) - \widetilde{P}_s h(i_0)\,\mathrm{d}s\bigg| \\
&\le \frac{1}{t}\int_0^t |P_s h(i_0) - \pi(h)|\,\mathrm{d}s + \frac{1}{t}\int_0^t |\widetilde{P}_s h(i_0) - \tilde{\pi}(h)|\,\mathrm{d}s + \|Q - \widetilde{Q}\|_{\ell_1} t \\
&\le \frac{2}{t}\int_0^\infty \eta_s\,\mathrm{d}s + \|Q - \widetilde{Q}\|_{\ell_1} t \qquad \text{for all } t > 0,
\end{aligned}$$

where we have used (3.4) in Theorem 3.1. By taking $t = \Big(2\int_0^\infty \eta_s\,\mathrm{d}s/\|Q - \widetilde{Q}\|_{\ell_1}\Big)^{1/2}$, we arrive at

$$\|\pi - \tilde{\pi}\|_{\mathrm{var}} = \sup_{|h|\le 1} |\pi(h) - \tilde{\pi}(h)| \le 2\sqrt{2}\bigg(\int_0^\infty \eta_s\,\mathrm{d}s\bigg)^{\frac{1}{2}} \sqrt{\|Q - \widetilde{Q}\|_{\ell_1}},$$

which is the desired conclusion. □

Remark 3.3. When $\mathcal{S}$ is a finite state space, the stability of π in terms of the perturbation of Q has been studied in [6, 7]. [7] proved it through expressing π as a polynomial of transition probabilities. This result was applied in [2] to establish the averaging principle for multiple-timescale systems. [6] proved a similar result by the Perron–Frobenius theorem to express π in terms of a non-zero right eigenvector of the Q-matrix with eigenvalue 0.

Second, we can apply Theorem 3.1 to improve all the main results [23, Theorems 1.1–1.4]. Here we only state the improvement of [23, Theorem 1.1] to save space.

Consider the regime-switching system (X_t, Λ_t) satisfying

$$\mathrm{d}X_t = b(X_t, \Lambda_t)\,\mathrm{d}t + \sigma(X_t, \Lambda_t)\,\mathrm{d}B_t, \qquad X_0 = x_0 \in \mathbb{R}^d,\ \Lambda_0 = i_0 \in \mathcal{S},$$

with (Λ_t) a Markov chain on $\mathcal{S} = \{1, 2, \ldots, N\}$, $2 \le N \le \infty$. In realistic applications, we can sometimes only get an estimation $\widetilde{Q}$ of the original transition rate matrix Q of (Λ_t). $\widetilde{Q}$ determines another Markov chain $\big(\tilde{\Lambda}_t\big)$. Correspondingly, the studied system (X_t) turns into $(\widetilde{X}_t)$ satisfying

$$\mathrm{d}\widetilde{X}_t = b\big(\widetilde{X}_t, \tilde{\Lambda}_t\big)\,\mathrm{d}t + \sigma\big(\widetilde{X}_t, \tilde{\Lambda}_t\big)\,\mathrm{d}B_t, \qquad \widetilde{X}_0 = x_0,\ \tilde{\Lambda}_0 = i_0.$$

It is necessary to measure the difference between X_t and $\widetilde{X}_t$ caused by the difference between Q and $\widetilde{Q}$. We shall characterize the difference between X_t and $\widetilde{X}_t$ via the Wasserstein distance between their distributions.

For any two probability measures μ, ν on $\mathbb{R}^d$, define the L_2-Wasserstein distance between them by

$$W_2(\mu, \nu)^2 = \inf_{\pi \in \mathscr{C}(\mu,\nu)} \left\{ \int_{\mathbb{R}^d \times \mathbb{R}^d} |x-y|^2 \, \pi(\mathrm{d}x, \mathrm{d}y) \right\},$$

where $\mathscr{C}(\mu, \nu)$ denotes the set of all the couplings of μ, ν on $\mathbb{R}^d \times \mathbb{R}^d$.

Corollary 3.2. *Assume that (Q1), (Q2), (H1), and (H2) hold. Denote by $\mathcal{L}(X_t)$ and $\mathcal{L}(\widetilde{X}_t)$ the distributions of X_t and $\widetilde{X}_t$ respectively. Then*

$$W_2\big(\mathcal{L}(X_t), \mathcal{L}(\widetilde{X}_t)\big)^2 \le C(p, t)\Big(\big\|Q - \widetilde{Q}\big\|_{\ell_1}\Big)^{(p-1)/p}, \qquad t > 0,$$

where $p > 1$ and C(p,t) is a positive constant depending only on p and t.

Proof. This result can be proved along the lines of [23, Theorem 1.1] by replacing the upper bound of [23, Lemma 2.2] with the upper bound given in Theorem 3.1. The constant $C(p,t)$ can be explicitly expressed as in [23]. □

Funding information

This work is supported in part by National Key R&D Program of China (No. 2022YFA1000033) and NNSFs of China (No. 12271397, 11831014).

Competing interests

There were no competing interests to declare which arose during the preparation or publication process of this article.

References

[1] Bardet, J., Guerin, H. and Malrieu, F. (2010). Long time behavior of diffusions with Markov switching. *ALEA Lat. Am. J. Prob. Math. Stat.* **7**, 151–170.

[2] Budhiraja, A., Dupuis, P. and Ganguly, A. (2018). Large deviations for small noise diffusions in a fast Markovian environment. *Electron. J. Prob.* **23**, 1–33.

[3] Chen, M. (2004). *From Markov Chains to Non-Equilibrium Particle Systems*, 2nd edn. World Scientific, Singapore.

[4] Cloez, B. and Hairer, M. (2015). Exponential ergodicity for Markov processes with random switching. *Bernoulli* **21**, 505–536.

[5] Daleckii, J. and Krein, M. (1974). *Stability of Solutions of Differential Equations in Banach Space*. American Mathematical Society, Providence, RI.

[6] Faggionato, A., Gabrielli, D. and Crivellari, M. (2010). Averaging and large deviation principles for fully-coupled piecewise deterministic Markov processes and applications to molecular motors. *Markov Process. Relat. Fields* **16**, 497–548.

[7] Freidlin, M. and Wentzell, A. (1998). *Random Perturbations of Dynamical Systems* (Grundlehren der Mathematischen Wissenschaften 260), 2nd ed,. Springer, New York.

[8] Ghosh, M., Arapostathis, A. and Marcus, S. (1997). Ergodic control of switching diffusions. *SIAM J. Control Optim.* **35**, 1952–1988.

[9] Guo, X. and Zhang, Q. (2002). Closed-form solutions for perpetual American put options with regime-switching. *SIAM J. Appl. Math.* **64**, 2034–2049.

[10] Ikeda, N. and Watanabe, S. (1977). A comparison theorem for solutions of stochastic differential equations and its applications. *Osaka J. Math.* **14**, 619–633.

[11] Ikeda, N. and Watanabe, S. (1989). *Stochastic Differential Equations and Diffusion Processes*, 2nd ed,. North-Holland, Amsterdam.

[12] Majda, A. and Tong, X. (2016). Geometric ergodicity for piecewise contracting processes with applications for tropical stochastic lattice models. *Commun. Pure Appl. Math.* **69**, 1110–1153.

https://doi.org/10.1017/jpr.2023.54 Published online by Cambridge University Press

[13] MAO, X. AND YUAN, C. (2006). *Stochastic Differential Equations with Markovian Switching*. Imperial College Press, London.
[14] MITROPHANOV, A. (2003). Stability and exponential convergence of continuous-time Markov chains. *J. Appl. Prob.* **40**, 970–979.
[15] MITROPHANOV, A. (2004). The spectral gap and perturbation bounds for reversible continuous-time Markov chains. *J. Appl. Prob.* **41**, 1219–1222.
[16] MITROPHANOV, A. (2005). Sensitivity and convergence of uniformly ergodic Markov chains. *J. Appl. Prob.* **42**, 1003–1014.
[17] NGUYEN, D., YIN, G. AND ZHU, C. (2017). Certain properties related to well posedness of switching diffusions. *Stoch. Process. Appl.* **127**, 3135–3158.
[18] SHAO, J. (2015). Strong solutions and strong Feller properties for regime-switching diffusion processes in an infinite state space. *SIAM J. Control Optim.* **53**, 2462–2479.
[19] SHAO, J. (2015). Criteria for transience and recurrence of regime-switching diffusion processes. *Electron. J. Prob.* **20**, 1–15.
[20] SHAO, J. (2018). Invariant measures and Euler–Maruyma's approximations of state-dependent regime-switching diffusions. *SIAM J. Control Optim.* **56**, 3215–3238.
[21] SHAO, J. AND XI, F. (2014). Stability and recurrence of regime-switching diffusion processes. *SIAM J. Control Optim.* **52**, 3496–3516.
[22] SHAO, J. AND XI, F. (2019). Stabilization of regime-switching processes by feedback control based on discrete time observations II: state-dependent case. *SIAM J. Control Optim.* **57**, 1413–1439.
[23] SHAO, J. AND YUAN, C. (2019). Stability of regime-switching processes under perturbation of transition rate matrices. *Nonlinear Anal. Hybrid Syst.* **33**, 211–226.
[24] SHAO, J. AND ZHAO, K. (2021). Continuous dependence for stochastic functional differential equations with state-dependent regime-switching on initial values. *Acta Math. Sin. English Ser.* **37**, 389–407.
[25] SKOROKHOD, A. (1989). *Asymptotic Methods in the Theory of Stochastic Differential Equations*. American Mathematical Society, Providence, RI.
[26] WANG, J. (2013). Stochastic comparison for Lévy-type processes. *J. Theoret. Prob.* **26**, 997–1019.
[27] XI, F. (2008). Feller property and exponential ergodicity of diffusion processes with state-dependent switching. *Sci. China A* **51**, 329–342.
[28] YIN, G. AND ZHU, C. (2010). *Hybrid Switching Diffusions: Properties and Applications* (Stoch. Model. Appl. Prob. **63**). Springer, New York.
[29] ZEIFMAN, A. AND ISAACSON, D. (1994). On strong ergodicity for nonhomogeneous continuous-time Markov chains. *Stoch. Process. Appl.* **50**, 263–273.
[30] ZEIFMAN, A., KOROLEV, V. AND SATIN, Y. (2020). Two approaches to the construction of perturbation bounds for continuous time Markov chains. *Mathematics* **8**, 253.
[31] ZHANG, S. (2019). On invariant probability measures of regime-switching diffusion processes with singular drifts. *J. Math. Anal. Appl.* **478**, 655–688.