Stability of regime-switching diffusion processes under perturbation of transition rate matrices

Chenggui Yuan

Department of Mathematics
Swansea University
Swansea, SA1 8EN
(Joint work with Jinghai Shao)

Outline

- A regime-switching diffusion process (RSDP), is a diffusion process in random environments characterized by a Markov chain.
- The state vector of a RSDP is a pair $(X(t), \Lambda(t))$, where $\{X(t)\}_{t\geq 0}$ satisfies a stochastic differential equation (SDE)

$$dX(t) = b(X(t), \Lambda(t))dt + \sigma(X(t), \Lambda(t))dW_t, \quad t > 0, \quad (1.1)$$

with the initial data $X_0 = x \in \mathbb{R}^n, \Lambda_0 = i \in \mathbb{S}$, and $\{\Lambda(t)\}_{t\geq 0}$ denotes a continuous-time Markov chain with the state space $\mathbb{S} := \{1, 2 \cdots, N\}$, $1 \leq N \leq \infty$, and the transition rules specified by

$$\mathbb{P}(\Lambda(t+\Delta) = j | \Lambda(t) = i) = \begin{cases} q_{ij}\Delta + o(\Delta), & i \neq j, \\ 1 + q_{ii}\Delta + o(\Delta), & i = j. \end{cases} \quad (1.2)$$

- A regime-switching diffusion process (RSDP), is a diffusion process in random environments characterized by a Markov chain.
- The state vector of a RSDP is a pair $(X(t), \Lambda(t))$, where $\{X(t)\}_{t\geq 0}$ satisfies a stochastic differential equation (SDE)

$$dX(t) = b(X(t), \Lambda(t))dt + \sigma(X(t), \Lambda(t))dW_t, \quad t > 0, \quad (1.1)$$

  with the initial data $X_0 = x \in \mathbb{R}^n, \Lambda_0 = i \in \mathbb{S}$, and $\{\Lambda(t)\}_{t\geq 0}$ denotes a continuous-time Markov chain with the state space $\mathbb{S} := \{1, 2 \cdots, N\}$, $1 \leq N \leq \infty$, and the transition rules specified by

$$\mathbb{P}(\Lambda(t+\Delta) = j | \Lambda(t) = i) = \begin{cases} q_{ij}\Delta + o(\Delta), & i \neq j, \\ 1 + q_{ii}\Delta + o(\Delta), & i = j. \end{cases} \quad (1.2)$$

Motivations (Cont.)

- RSDPs have considerable applications in e.g. control problems, storage modeling, neutral activity, biology and mathematical finance (see e.g. the monographs by Mao-Yuan (2006), and Yin-Zhu (2010)). The dynamical behavior of RSDPs may be markedly different from diffusion processes without regime switchings, see e.g. Pinsky -Scheutzow (1992), Mao-Yuan (2006).

Mao, X. and Yuan, C., Stochastic Differential Equations with Markovian Switching, Imperial College Press, 2006

It is interesting to have a look of the following two equations

$$dx(t) = x(t)dt + 2x(t)dW(t) \tag{1.3}$$

and

$$dx(t) = 2x(t) + x(t)dW(t) \tag{1.4}$$

switching from one to the other according to the movement of the Markov chain $\Lambda(t)$. We observe that Eq. (1.3) is almost surely exponentially stable since the Lyapunov exponent is $\lambda_1 = -1$ while Eq. (1.4) is almost surely exponentially unstable since the Lyapunov exponent is $\lambda_2 = 1.5$.

Let $\Lambda(t)$ be a right-continuous Markov chain taking values in $S = \{1, 2\}$ with the generator

$$\Gamma = (\gamma_{ij})_{2\times 2} = \begin{pmatrix} -1 & 1 \\ \gamma & -\gamma \end{pmatrix}.$$

Of course $W(t)$ and $\Lambda(t)$ are assumed to be independent. Consider a one-dimensional linear SDEwMS

$$dx(t) = a(\Lambda(t))x(t)dt + b(\Lambda(t))x(t)dW(t) \tag{1.5}$$

on $t \geq 0$, where

$$a(1) = 1, \quad a(2) = 2, \quad b(1) = 2, \quad b(2) = 1.$$

However, as the result of Markovian switching, the overall behaviour, i.e. Eq. (1.5) will be exponentially stable if $\gamma > 1.5$ but exponentially unstable if $\gamma < 1.5$ while the Lyapunov exponent of the solution is 0 when $\gamma = 1.5$.

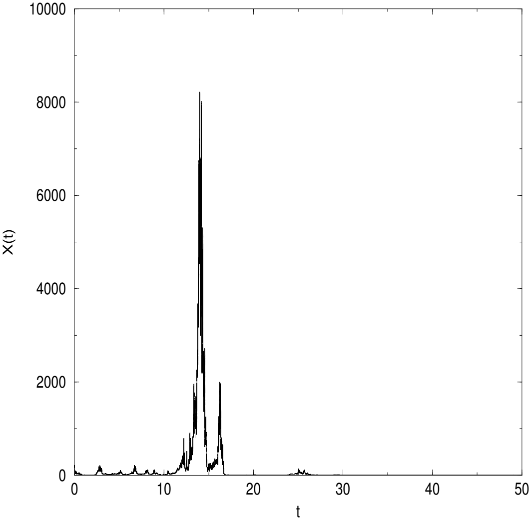

Figure: The graph of numerical solution when $\gamma = 2$.

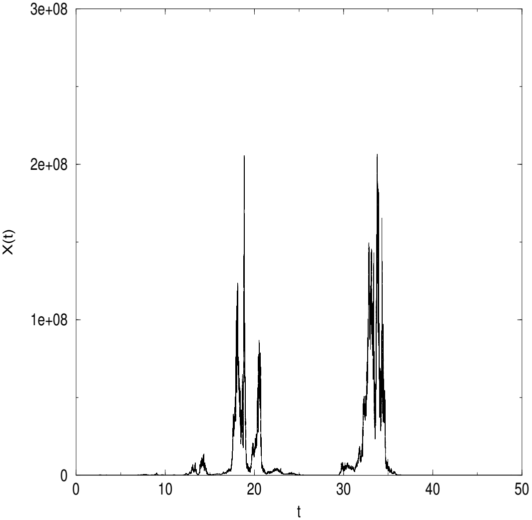

Figure: The graph of numerical solution when $\gamma = 1.5$.

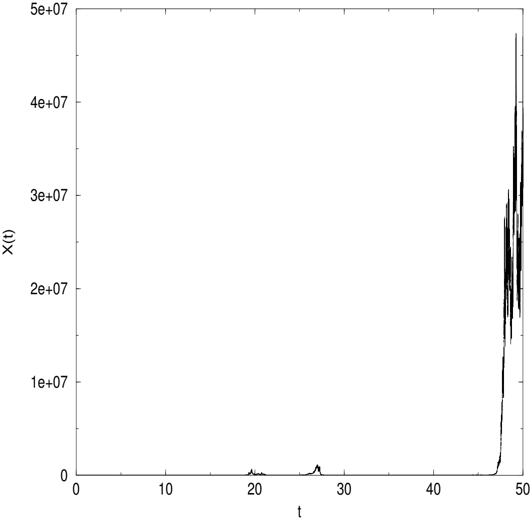

Figure: The graph of numerical solution when $\gamma = 0.5$.

Motivations (Cont.)

- So far, the works on RSDPs have included ergodicity (Cloez-Hairer (2013), Shao (2014)) stability in distribution (Mao-Yuan (03), Xi-Yin (2010)), recurrence and transience (Pinsky-Scheutzow (1992),invariant densities (Bakhtin et al. (2014)) and so forth
- For the counterpart associated with Euler-Maruyama (EM) algorithms, we refer to Yuan-Mao (2005), and Yin-Zhu (2010), where RSDPs therein enjoy finite state space.

Motivations (Cont.)

- So far, the works on RSDPs have included ergodicity (Cloez-Hairer (2013), Shao (2014)) stability in distribution (Mao-Yuan (03), Xi-Yin (2010)), recurrence and transience (Pinsky-Scheutzow (1992),invariant densities (Bakhtin et al. (2014)) and so forth
- For the counterpart associated with Euler-Maruyama (EM) algorithms, we refer to Yuan-Mao (2005), and Yin-Zhu (2010), where RSDPs therein enjoy finite state space.

Shao, J., Yuan, C., Stability of regime-switching processes under perturbation of transition rate matrices. Nonlinear Analysis: Hybrid Systems 33 (2019), 211-226.

In this work we are concerned with the stability of the process (X_t) under perturbation of the transition rate matrix of $(\Lambda(t))$. From the application point of view, there are mainly two types of perturbations of Q.

First type of perturbation: The size of Q is fixed, however, each entry q_{ij} of Q may have small perturbation. Namely, there is another transition rate matrix $\widetilde{Q} = (\tilde{q}_{ij})_{i,j\in\mathbb{S}}$, and each entry $\tilde{q}_{ij}$ acts as an estimator of the element q_{ij} of Q. Without loss of generality, assume that $\widetilde{Q}$ is conservative and totally stable, then a unique transition function $\widetilde{P}_t,\ t \geq 0$ is determined (cf. e.g. [3, Corollary 3.12]). Let $(\tilde{\Lambda}(t))$ be a continuous-time Markov chain starting from $i_0 = \Lambda_0$ corresponding to $\widetilde{Q}$. Then the distribution of $\tilde{\Lambda}_t$ is fixed, so, a new dynamical system $(\tilde{X}(t))$ is induced from the process $(\tilde{\Lambda}(t))$, i.e.

$$d\tilde{X}(t) = b(\tilde{X}(t), \tilde{\Lambda}(t))dt + \sigma(\tilde{X}(t), \tilde{\Lambda}(t))dW(t), \quad \tilde{X}_0 = x_0 \in \mathbb{R}^d,\ \tilde{\Lambda}(0) = i_0 \in \mathbb{S}. \tag{1.6}$$

Under some suitable conditions of the coefficients $b(\cdot,\cdot)$ and $\sigma(\cdot,\cdot)$, SDEs (1.1) and (1.6) admit a unique solution. Therefore, the distribution $\mathcal{L}(X(t))$ of $X(t)$ (resp. $\mathcal{L}(\tilde{X}(t))$ of $\tilde{X}(t)$) is determined in some sense by the transition rate matrix Q (resp. $\widetilde{Q}$). The following basic and important question therefore arise:

- Can we use the difference between Q and $\widetilde{Q}$ to charactrize the difference between the distributions of $X(t)$ and $\tilde{X}(t)$?

Second type of perturbation: The size of Q can be changed. In applications, when facing the graphs drawn from experimental data, it is hard sometimes to determine the number of the regimes for the regime-switching processes. For example, if there are actually three regimes, the process stays for a very short period of time at one of them. From this kind of experimental data, it is very likely that a regime-switching model with only two regimes are detected. What is the impact caused by this incorrect choice of the number of states for the regime-switching processes?

Precisely, let $\widehat{Q}$ be a conservative transition rate matrix on $E := \mathbb{S}\backslash\{1,\dots,m\}$ with $m < N$, which determines uniquely the semigroup $\hat{P}_t = e^{t\hat{Q}},\ t \geq 0$ on E. Let $(\hat{\Lambda}_t)$ be a continuous-time Markov chain on E corresponding to $(\hat{P}_t)$ or equivalently $\hat{Q}$. Using the same coefficients $b(\cdot,\cdot)$, $\sigma(\cdot,\cdot)$ as those of SDE (1.1), and considering the new dynamical system $(\hat{X}_t)$ corresponding to $(\hat{\Lambda}_t)$ defined by:

$$d\hat{X}_t = b(\hat{X}_t, \hat{\Lambda}_t)dt + \sigma(\hat{X}_t, \hat{\Lambda}_t)dW_t, \quad \hat{X}_0 = x_0 \in \mathbb{R}^d,\ \hat{\Lambda}_0 = i_1 \in E. \quad (1.7)$$

Under suitable conditions of b and σ, the solutions of (1.1) and (1.7) are uniquely determined. This means that given $\hat{Q}$ on E, the distribution of $\hat{X}_t$ is then determined. Denote $\mathcal{L}(X_t)$ and $\mathcal{L}(\hat{X}_t)$ the distributions of X_t and $\hat{X}_t$ respectively. We aim to measure the Wasserstein distance $W_2(\mathcal{L}(X_t), \mathcal{L}(\hat{X}_t))$ via the difference between the transition rate matrices $Q = (q_{ij})_{i,j\in\mathbb{S}}$ and $\hat{Q} = (\hat{q}_{ij})_{i,j\in E}$. To achieve this, rewrite Q in the following form:

$$Q = \begin{pmatrix} Q_0 & A \\ B & Q_1 \end{pmatrix}, \quad (1.8)$$

where $Q_0 \in \mathbb{R}^{m\times m}$, $A \in \mathbb{R}^{m\times(N-m)}$, $B \in \mathbb{R}^{(N-m)\times m}$, and $Q_1 \in \mathbb{R}^{(N-m)\times(N-m)}$.

To analyze the impact of the regularity of the coefficients in SDE (1.1), we will consider separately two situations: SDEs with regular coefficients and SDEs with irregular coefficients. Let us first consider the situation that the coefficients of (1.1) are regular. Assume the coefficients $b:\mathbb{R}^d\times\mathbb{S}\to\mathbb{R}^d$ and $\sigma:\mathbb{R}^d\times\mathbb{S}\to\mathbb{R}^{d\times d}$ satisfy:

(H1) For each $i\in\mathbb{S}$ there exists a constant κ_i such that

$$2\langle x-y, b(x,i)-b(y,i)\rangle+2\|\sigma(x,i)-\sigma(y,i)\|_{\mathrm{HS}}^2\le\kappa_i|x-y|^2,\quad x,\ y\in\mathbb{R}^d.$$

(H2) There exists a constant K such that

$$|b(x,i)|^2\le K(1+|x|^2),\quad \|\sigma(x,i)\|_{\mathrm{HS}}^2\le K(1+|x|^2),\quad x\in\mathbb{R}^d,\ i\in\mathbb{S}.$$

In this case, we shall use the Wasserstein distance $W_2(\cdot,\cdot)$ to measure the difference between the distributions of $X(t)$ and $\tilde{X}(t)$, which is defined by

$$W_2(\nu_1,\nu_2)^2=\inf_{\Pi\in\mathcal{C}(\nu_1,\nu_2)}\Big\{\int_{\mathbb{R}^d\times\mathbb{R}^d}|x-y|^2\Pi(dx,dy)\Big\}, \tag{2.1}$$

where $\mathcal{C}(\nu_1,\nu_2)$ denotes the set of all probability measures on $\mathbb{R}^d\times\mathbb{R}^d$ with marginals ν_1 and ν_2.

To analyze the impact of the regularity of the coefficients in SDE (1.1), we will consider separately two situations: SDEs with regular coefficients and SDEs with irregular coefficients. Let us first consider the situation that the coefficients of (1.1) are regular. Assume the coefficients $b:\mathbb{R}^d\times\mathbb{S}\to\mathbb{R}^d$ and $\sigma:\mathbb{R}^d\times\mathbb{S}\to\mathbb{R}^{d\times d}$ satisfy:

(H1) For each $i\in\mathbb{S}$ there exists a constant κ_i such that

$$2\langle x-y, b(x,i)-b(y,i)\rangle+2\|\sigma(x,i)-\sigma(y,i)\|_{\mathrm{HS}}^2\le\kappa_i|x-y|^2,\quad x,\ y\in\mathbb{R}^d.$$

(H2) There exists a constant K such that

$$|b(x,i)|^2\le K(1+|x|^2),\quad \|\sigma(x,i)\|_{\mathrm{HS}}^2\le K(1+|x|^2),\quad x\in\mathbb{R}^d,\ i\in\mathbb{S}.$$

In this case, we shall use the Wasserstein distance $W_2(\cdot,\cdot)$ to measure the difference between the distributions of $X(t)$ and $\tilde X(t)$, which is defined by

$$W_2(\nu_1,\nu_2)^2=\inf_{\Pi\in\mathcal{C}(\nu_1,\nu_2)}\Big\{\int_{\mathbb{R}^d\times\mathbb{R}^d}|x-y|^2\Pi(dx,dy)\Big\},\qquad (2.1)$$

where $\mathcal{C}(\nu_1,\nu_2)$ denotes the set of all probability measures on $\mathbb{R}^d\times\mathbb{R}^d$ with marginals ν_1 and ν_2.

For an irreducible transition rate matrix Q on $\mathbb{S}$, its corresponding transition probability measure $P_t(i,\cdot)$ must be ergodic. Denote $\pi = (\pi_i)$ the invariant probability measure of Q. Define τ to be the largest positive constant such that

$$\sup_{i\in\mathbb{S}} \|P_t(i,\cdot) - \pi\|_{\mathrm{var}} = O(\varepsilon^{-\tau t}), \quad t > 0, \tag{2.2}$$

where $\|\mu - \nu\|_{\mathrm{var}}$ stands for the total variation distance between two probability measures μ and ν, i.e. $\|\mu - \nu\|_{\mathrm{var}} = 2\sup\{|\mu(A) - \nu(A)|; A \in \mathscr{B}(S)\}$. Additionally, for $p > 0$, let

$$Q_p = Q + p\,\mathrm{diag}(\kappa_0, \kappa_1, \ldots, \kappa_N),$$

and

$$\eta_p = -\max\left\{\mathrm{Re}(\gamma);\ \gamma \in \mathrm{spec}(Q_p)\right\}, \tag{2.3}$$

where $\mathrm{spec}(Q_p)$ denotes the spectrum of the operator Q_p.

Theorem

Let (X_t, Λ_t) *and* $(\widetilde{X}_t, \tilde{\Lambda}_t)$ *be the solutions of* (1.1) *and* (1.6) *respectively. Assume* (H1) *and* (H2) *hold. Then*

$$W_2(\mathcal{L}(X_t), \mathcal{L}(\widetilde{X}_t))^2 \leq \left(4\varepsilon^{-1}+8\right) K C_2(p)^{\frac{1}{p}} \left(N^2 t^2 \|Q - \widetilde{Q}\|_{\ell_1}\right)^{\frac{1}{q}} \Psi(t, \varepsilon, \eta_p, K, p), \tag{2.4}$$

where $p > 1$, $q = p/(p-1)$, ε *and* $C_2(p)$ *are positive constants,* ℓ_1*-norm is the maximum absolute row sum norm,* η_p *is defined by* (2.3), *and*

$$\Psi(t, \varepsilon, \eta_p, K, p) = \left(\int_0^t \left[1 + (|x_0|^2 + 2Ks)e^{(2K+1)s}\right]^p e^{-(\eta_p - \varepsilon p)(t-s)} ds\right)^{\frac{1}{p}}. \tag{2.5}$$

Theorem

If assume further that

$$|b(x,i)|^2 \le K, \quad \|\sigma(x,i)\|_{\mathrm{HS}}^2 \le K, \quad x \in \mathbb{R}^d,\ i \in \mathbb{S}, \tag{2.6}$$

then we have a simple estimate:

$$\begin{aligned} &W_2(\mathcal{L}(X_t), \mathcal{L}(\widetilde{X}_t))^2 \\ &\le (4\varepsilon^{-1}+8) K C_2(p)^{\frac{1}{p}} \big(N^2 t^2 \|Q - \widetilde{Q}\|_{\ell_1}\big)^{\frac{1}{q}} \Big(\frac{1 - e^{-(\eta_p - \varepsilon p)t}}{\eta_p - \varepsilon p}\Big)^{\frac{1}{p}}. \end{aligned} \tag{2.7}$$

Theorem

Let (X_t, Λ_t) *and* $(\hat{X}_t, \hat{\Lambda}_t)$ *be the solutions of* (1.1) *and* (1.7) *respectively. Assume* (H1) *and* (H2) *hold. Then*

$$W_2(\mathcal{L}(X_t), \mathcal{L}(\hat{X}_t))^2 \le (4\varepsilon^{-1}+8)KC_2(p)^{\frac{1}{p}}(Nt)^{\frac{2}{q}}\left(\|B\|_{\ell_1} + \|Q_1 - \widehat{Q}\|_{\ell_1}\right)^{\frac{1}{q}}\Psi(t,\varepsilon,\eta_p,K,p), \tag{2.8}$$

where $p > 1$, $q = p/(p-1)$, ε *and* $C_2(p)$ *are positive constants,* η_p *is defined by* (2.3), *and* $\Psi(t,\varepsilon,\eta_p,K,p)$ *is given by* (2.5). *Assume further that* b *and* σ *satisfy* (2.6), *then*

$$W_2(\mathcal{L}(X_t), \mathcal{L}(\hat{X}_t))^2 \le (4\varepsilon^{-1}+8)KC_2(p)^{\frac{1}{p}}(Nt)^{\frac{2}{q}}\left(\|B\|_{\ell_1} + \|Q_1 - \widehat{Q}\|_{\ell_1}\right)^{\frac{1}{q}}\left(\frac{1-\varepsilon^{-(\eta_p-\varepsilon p)t}}{\eta_p-\varepsilon p}\right)^{\frac{1}{p}}. \tag{2.9}$$

Next, we consider the stability of the dynamical system (X_t) under the perturbation of the transition rate matrix when the coefficients of the studied SDE are irregular. Precisely, let

$$dX_t = b(X_t, \Lambda_t)dt + \sigma(X_t)dW_t, \quad X_0 = x_0 \in \mathbb{R}^d,\ \Lambda_0 = i_0 \in \mathbb{S}, \quad (2.10)$$

where $\sigma : \mathbb{R}^d \to \mathbb{R}^{d\times d}$ is still Lipschitz continuous, but b only satisfies some integrability condition. Here, (Λ_t) is also a continuous time Markov chain with a conservative and irreducible transition rate matrix $Q = (q_{ij})_{i,j\in\mathbb{S}})$. (Λ_t) is assumed to be independent of (W_t). An interesting example (see F.Y. Wang) is

$$b(x, i) = \beta_i \Big\{ \sum_{k=1}^{\infty} \log\Big(1 + \frac{1}{|x-k|^2}\Big)\Big\}^{\frac{1}{2}} - x, \qquad (2.11)$$

where $\beta : \mathbb{S} \to \mathbb{R}_+$. This drift b is rather singular, whereas we can show that (X_t) is still stable in a suitable sense w.r.t. the perturbation of Q even in this situation.

Similar to (1.1) and (1.7), we consider the processes $(\widetilde{X}_t)$ and $(\hat{X}_t)$ corresponding to the perturbations $\widetilde{Q} = (\tilde{q}_{ij})_{i,j\in\mathbb{S}}$ and $\hat{Q} = (\hat{q}_{ij})_{i,j\in E}$. Namely,

$$d\widetilde{X}_t = b(\widetilde{X}_t, \tilde{\Lambda}_t)dt + \sigma(\widetilde{X}_t)dW_t,\ \widetilde{X}_0 = x_0,\ \tilde{\Lambda}_0 = i_0, \tag{2.12}$$

where $(\tilde{\Lambda}_t)$ is associated with $\widetilde{Q}$ and is independent of (W_t).

$$d\hat{X}_t = b(\hat{X}_t, \hat{\Lambda}_t)dt + \sigma(\hat{X}_t)dW_t,\ \hat{X}_0 = x_0,\ \hat{\Lambda}_0 = i_1 \in E, \tag{2.13}$$

where $(\hat{\Lambda}_t)$ is associated with $\hat{Q}$ on the state space E and is independent of (W_t). We shall measure the difference between the distribution $\mathcal{L}(X_t)$ and $\mathcal{L}(\widetilde{X}_t)$ by the Fortet-Mourier distance (also called bounded Lipschitz distance):

$$W_{bL}(\mu, \nu) = \sup\Big\{ \int_{\mathbb{R}^d} \phi\, d\mu - \int_{\mathbb{R}^d} \phi\, d\nu;\ \|\phi\|_{\mathrm{Lip}} + \|\phi\|_\infty \le 1 \Big\} \tag{2.14}$$

for two probability measures $\mu,\ \nu$ on $\mathbb{R}^d$, $\|\phi\|_{\mathrm{Lip}} := \sup_{x,y,\in\mathbb{R}^d, x\neq y} \frac{|\phi(x)-\phi(y)|}{|x-y|}$.

To provide a suitable integrability condition on the drift b, we need to introduce an auxiliary function V and its associated probability measure μ_0. Let $V \in C^2(\mathbb{R}^d)$, define

$$Z_0(x) = -\sum_{i,j=1}^{d} \big(a_{ij}(x)\partial_j V(x)\big)e_i, \tag{2.15}$$

where $(a_{ij}(x)) = \sigma(x)\sigma^*(x)$, σ^* denotes the transpose of σ, $\{e_i\}_{i=1}^d$ is the canonical orthonormal basis of $\mathbb{R}^d$ and ∂_j is the directional derivative along e_j. Let

$$\mu_0(dx) = e^{-V(x)}dx. \tag{2.16}$$

Assume that V satisfies:

(A) there exists a $K_0 > 0$ such that $|Z_0(x) - Z_0(y)| \leq K_0|x - y|$ for all $x, y \in \mathbb{R}^d$, and $\mu_0(\mathbb{R}^d) = 1$.

Let

$$Z(x,i) = b(x,i) - Z_0(x), \quad x \in \mathbb{R}^d,\ i \in \mathbb{S}. \tag{2.17}$$

Theorem

Suppose that condition (A) holds for the function $V \in C^2(\mathbb{R}^d)$. Let $T > 0$ be fixed. Assume that there exists a constant $\eta > 2Td$ such that

$$\max_{i\in\mathbb{S}} \mu_0\left(e^{\eta|\sigma^{-1}(\cdot)Z(\cdot,i)|^2}\right) < \infty. \tag{2.18}$$

Then

$$W_{bL}(\mathcal{L}(X_t), \mathcal{L}(\widetilde{X}_t)) \leq C \max\left\{\|Q-\widetilde{Q}\|_{\ell_1}^{\frac{1}{2q_0}}, \|Q-\widetilde{Q}\|_{\ell_1}^{\frac{1}{2q_0\gamma}}\right\}, \quad t \in [0,T], \tag{2.19}$$

for some constant C depending on $T, x_0, \tau_1, K_0, \gamma, p_0$ and $\max_{i\in\mathbb{S}} \mu_0\left(e^{\eta|\sigma^{-1}(\cdot)Z(\cdot,i)|^2}\right)$, where $p_0 > 1$ is a constant satisfying $2p_0^2Td < \eta$, $q_0 = p_0/(p_0-1)$ and $\gamma > 1$ is a constant.

Theorem

Suppose that condition (A) holds for the function $V \in C^2(\mathbb{R}^d)$. Let $T > 0$ be fixed. Assume that there exists a constant $\eta > 2Td$ such that (2.18) still holds. The representation (1.8) holds. Then

$$W_{bL}(\mathcal{L}(X_t), \mathcal{L}(\hat{X}_t)) \\ \leq C \max\left\{ \left(\|B\|_{\ell_1} + \|Q_1 - \widehat{Q}\|_{\ell_1}\right)^{\frac{1}{2q_0}}, \left(\|B\|_{\ell_1} + \|Q_1 - \widehat{Q}\|_{\ell_1}\right)^{\frac{1}{2q_0\gamma}} \right\}, \quad t \in [0, T], \tag{2.20}$$

for some constant C depending on $N, T, x_0, \tau_1, K_0, \gamma, p_0$ and $\max_{i \in \S} \mu_0\left(\varepsilon^{\eta|\sigma^{-1}(\cdot)Z(\cdot,i)|^2}\right)$, where $p_0 > 1$ is a constant satisfying $2p_0^2 Td < \eta$, $q_0 = p_0/(p_0 - 1)$ and $\gamma > 1$ is a constant.

Proofs of Main Results

Consider the following SDEs:

$$dX_t = b(X_t, \Lambda_t)dt + \sigma(X_t, \Lambda_t)dW_t, \quad X_0 = x_0,\ \Lambda_0 = i_0, \quad (2.21)$$

$$d\widetilde{X}_t = b(\widetilde{X}_t, \tilde{\Lambda}_t)dt + \sigma(\widetilde{X}_t, \tilde{\Lambda}_t)dW_t, \quad \widetilde{X}_0 = x_0,\ \tilde{\Lambda}_0 = i_0. \quad (2.22)$$

Here (Λ_t) and $(\tilde{\Lambda}_t)$ are continuous-time Markov chains on $\mathbb{S} = \{1, \dots, N\}$ with transition rate matrices $Q = (q_{ij})_{i,j\in\mathbb{S}}$ and $\widetilde{Q} = (\tilde{q}_{ij})_{i,j\in\mathbb{S}}$ respectively.

For the regime-switching diffusions (X_t, Λ_t) and $(\widetilde{X}_t, \tilde{\Lambda}_t)$ with Markovian switching, as usual we assume (Λ_t) and $(\tilde{\Lambda}_t)$ are independent of the Brownian motion (W_t). To be precise, we introduce the following probability space $(\Omega, \mathscr{F}, \mathbb{P})$ used throughout this work.

Let

$$\Omega_1 = \{\omega \big| \omega : [0, \infty) \to \mathbb{R}^d \text{ continuous}, \ \omega_0 = 0\},$$

which is endowed with the local uniform convergence topology and the Wiener measure $\mathbb{P}_1$ so that its coordinate process $W(t, \omega) = \omega(t)$, $t \geq 0$, is a d-dimensional Brownian motion.
Put

$$\Omega_2 = \{\omega \big| \omega : [0, \infty) \to \mathbb{S} \text{ right continuous with left limit}\},$$

endowed with the Skorokhod topology and a probability measure $\mathbb{P}_2$. The Markov chains (Λ_t) and $(\tilde{\Lambda}_t)$ are all constructed in the space $(\Omega_2, \mathscr{B}(\Omega_2), \mathbb{P}_2)$. Set

$$(\Omega, \mathscr{F}, \mathbb{P}) = (\Omega_1 \times \Omega_2, \mathscr{B}(\Omega_1) \times \mathscr{B}(\Omega_2), \mathbb{P}_1 \times \mathbb{P}_2).$$

Thus under $\mathbb{P} = \mathbb{P}_1 \times \mathbb{P}_2$, (Λ_t), $(\tilde{\Lambda}_t)$ are independent of the Brownian motion (W_t). Denote by $\mathbb{E}_{\mathbb{P}_1}$ taking the expectation with respect to the probability measure $\mathbb{P}_1$, and similarly $\mathbb{E}_{\mathbb{P}_2}$.

Lemma

Let (X_t, Λ_t), $(\widetilde{X}_t, \tilde{\Lambda}_t)$ *be the solution of* (2.21) *and* (2.22) *respectively and* $X_0 = \widetilde{X}_0 = x_0 \in \mathbb{R}^d$. *Assume (H2) holds. Then, for* $\mathbb{P}_2$*-almost surely* $\omega_2 \in \Omega_2$,

$$\begin{aligned} \mathbb{E}_{\mathbb{P}_1}[|X_t|^2](\omega_2) &\le (|x_0|^2 + 2Kt)\varepsilon^{(2K+1)t}, \\ \mathbb{E}_{\mathbb{P}_1}[|\widetilde{X}_t|^2](\omega_2) &\le (|x_0|^2 + 2Kt)\varepsilon^{(2K+1)t}, \quad t > 0. \end{aligned} \tag{2.23}$$

This can be proved by using the Itô formula, and taking the expectation w.r.t. $\mathbb{P}_1$.

We also need the following lemma.
Next, we construct a coupling process $(\Lambda_t, \tilde{\Lambda}_t)$ such that (Λ_t) and $(\tilde{\Lambda}_t)$ are continuous-time Markov chains with transition rate matrix Q and $\widetilde{Q}$ respectively.

Lemma

It holds that

$$\int_0^t \mathbb{P}(\Lambda_s \neq \tilde{\Lambda}_s) ds \leq N^2 t^2 \|Q - \widetilde{Q}\|_{\ell_1}. \tag{2.24}$$

Proof of the Theorem

For simplicity of notation, let $Z_t = X_t - \widetilde{X}_t$. Then, due to (H1) and (H2), Itô's formula yields that

$$
\begin{aligned}
d|Z_t|^2 &= \big\{2\langle Z_t, b(X_t,\Lambda_t) - b(\widetilde{X}_t,\tilde{\Lambda}_t)\rangle + \|\sigma(X_t,\Lambda_t) - \sigma(\widetilde{X}_t,\tilde{\Lambda}_t)\|_{\mathrm{HS}}^2\big\}dt + dM_t \\
&\le \big\{\kappa_{\Lambda_t}|Z_t|^2 + 2\langle Z_t, b(\widetilde{X}_t,\Lambda_t) - b(\widetilde{X}_t,\tilde{\Lambda}_t)\rangle + 2\|\sigma(\widetilde{X}_t,\Lambda_t) - \sigma(\widetilde{X}_t,\tilde{\Lambda}_t)\|_{\mathrm{HS}}^2\big\}dt + dM_t \\
&\le \big\{(\kappa_{\Lambda_t}+\varepsilon)|Z_t|^2 + \frac{1}{\varepsilon}\big(|b(\widetilde{X}_t,\Lambda_t)| + |b(\widetilde{X}_t,\tilde{\Lambda}_t)|\big)^2\mathbf{1}_{\{\Lambda_t\neq\tilde{\Lambda}_t\}} \\
&\qquad + 4\big(\|\sigma(\widetilde{X}_t,\Lambda_t)\|_{\mathrm{HS}}^2 + \|\sigma(\widetilde{X}_t,\tilde{\Lambda}_t)\|_{\mathrm{HS}}^2\big)\mathbf{1}_{\{\Lambda_t\neq\tilde{\Lambda}_t\}}\big\}dt + dM_t \\
&\le \big\{(\kappa_{\Lambda_t}+\varepsilon)|Z_t|^2 + \frac{2K}{\varepsilon}(1+|\widetilde{X}_t|^2)\mathbf{1}_{\{\Lambda_t\neq\tilde{\Lambda}_t\}} + 8K(1+|\widetilde{X}_t|^2)\mathbf{1}_{\{\Lambda_t\neq\tilde{\Lambda}_t\}}\big\}dt + dM_t
\end{aligned}
$$

for any $\varepsilon > 0$, where $M_t = \int_0^t 2\langle Z_s, (\sigma(X_s,\Lambda_s) - \sigma(\widetilde{X}_s,\tilde{\Lambda}_s))dW_s\rangle$ for $t \ge 0$ is a martingale.

Taking the expectation w.r.t. $\mathbb{P}_1$ on both sides of the previous inequality, we get

$$d\,\mathbb{E}_{\mathbb{P}_1}[|Z_t|^2](\omega_2) \leq \big(4\varepsilon^{-1}+8\big)K\mathbb{E}_{\mathbb{P}_1}\big[1+|\widetilde{X}_t|^2\big](\omega_2)\mathbf{1}_{\{\Lambda_t\neq\tilde{\Lambda}_t\}}(\omega_2)dt + (\kappa_{\Lambda_t}+\varepsilon)(\omega_2)\mathbb{E}_{\mathbb{P}_1}[|Z_t|^2](\omega_2)dt. \tag{2.25}$$

Using the Gronwall inequality and Lemma, we obtain that

$$\mathbb{E}_{\mathbb{P}_1}[|Z_t|^2](\omega_2) \leq \big(4\varepsilon^{-1}+8\big)K\int_0^t\Big(1+(|x_0|^2+2Ks)e^{(2K+1)s}\Big) \times \mathbf{1}_{\{\Lambda_s\neq\tilde{\Lambda}_s\}}\, e^{\int_s^t(\kappa_{\Lambda_r}+\varepsilon)(\omega_2)dr}ds.$$

Taking the expectation w.r.t. $\mathbb{P}_2$ and using Hölder's inequality, we get

$$\begin{aligned}\mathbb{E}|Z_t|^2 \leq \int_0^t &\Big\{(4\varepsilon^{-1}+8)K\big[1+(|x_0|^2+2Ks)e^{(2K+1)s}\big] \\ &\cdot \big(\mathbb{E}\mathbf{1}_{\{\Lambda_s\neq\tilde{\Lambda}_s\}}(\omega_2)\big)^{\frac{1}{q}}\big(\mathbb{E}e^{p\int_s^t(\kappa_{\Lambda_r}+\varepsilon)(\omega_2)dr}\big)^{\frac{1}{p}}\Big\}ds\end{aligned} \tag{2.26}$$

for $p,\ q>1$ with $1/p+1/q=1$.

In order to estimate the term $\mathbb{E}\, e^{q\int_0^t(\kappa_{\Lambda_s}+1)ds}$, we need the following notation. Let

$$Q_p = Q + p\mathrm{diag}(\kappa_0, \kappa_1, \ldots, \kappa_N),$$

and

$$\eta_p = -\max\left\{\mathrm{Re}(\gamma);\ \gamma \in \mathrm{spec}(Q_p)\right\},$$

where $\mathrm{diag}(\kappa_0, \kappa_1, \ldots, \kappa_N)$ denotes the diagonal matrix generated by the vector $(\kappa_0, \kappa_1, \ldots, \kappa_N)$, $\mathrm{spec}(Q_p)$ denotes the spectrum of the operator Q_p. According to [1, Proposition 4.1], for any $p > 0$, there exist two positive constants $C_1(p)$ and $C_2(p)$ such that

$$C_1(p)e^{-\eta_p t} \le \mathbb{E}\, e^{p\int_0^t \kappa_{\Lambda_s} ds} \le C_2(p)e^{-\eta_p t}, \quad t > 0. \tag{2.27}$$

To estimate the term $\int_0^t \mathbb{E}\mathbf{1}_{\{\Lambda_s \neq \tilde{\Lambda}_s\}} ds$ is in previous lemma. Consequently, substituting the estimates (2.27) and (2.24) into (2.26), we arrive at

$$\mathbb{E}[|Z_t|^2] \leq \big(4\varepsilon^{-1}+8\big) K C_2(p)^{\frac{1}{p}} \Big(N^2 t^2 \|Q - \widetilde{Q}\|_{\ell_1} \Big)^{\frac{1}{q}} \cdot \Big(\int_0^t \big[1 + (|x_0|^2 + 2Ks) e^{(2K+1)s}\big]^p e^{-(\eta_p - \varepsilon p)(t-s)} ds \Big)^{\frac{1}{p}}. \tag{2.28}$$

Note that the solutions of (2.21) and (2.22) exist uniquely. Then the distribution of $(X_t, \widetilde{X}_t)$ on $\mathbb{R}^d \times \mathbb{R}^d$ is a coupling of $\mathcal{L}(X_t)$ and $\mathcal{L}(\widetilde{X}_t)$. By the definition of the Wasserstein distance, it follows

$$\begin{aligned} W_2(\mathcal{L}(X_t), \mathcal{L}(\widetilde{X}_t))^2 &\leq \mathbb{E}[|X_t - \widetilde{X}_t|^2] \\ &\leq \big(4\varepsilon^{-1}+8\big)KC_2(p)^{\frac{1}{p}} N^{\frac{2}{q}} t^{\frac{2}{q}} \|Q - \widetilde{Q}\|_{\ell_1}^{\frac{1}{q}} \\ &\cdot \Big(\int_0^t \big[1 + (|x_0|^2 + 2Ks)\varepsilon^{(2K+1)s}\big]^p e^{-(\eta_p - \varepsilon p)(t-s)} ds\Big)^{\frac{1}{p}}, \end{aligned}$$

which is the desired estimate (2.4).

- J. Bardet, H. Guerin, F. Malrieu, Long time behavior of diffusions with Markov switching, ALEA Lat. Am. J. Probab. Math. Stat., 7 (2010), 151-170.
- A.Y. Mitrophanov, The spectral gap and perturbation bounds for reversible continuous-time Markov chains, J. Appl. Prob. 41 (2004), 1219-1222.
- M.-F. Chen, From Markov chains to non-equilibrium particle systems, 2nd ed. Singapore: World Scientific, 2004.
- B. Cloez, and M. Hairer, Exponential ergodicity for Markov processes with random switching, Bernoulli, 21 (2015), 505–536.

- J. Bardet, H. Guerin, F. Malrieu, Long time behavior of diffusions with Markov switching, ALEA Lat. Am. J. Probab. Math. Stat., 7 (2010), 151-170.
- A.Y. Mitrophanov, The spectral gap and perturbation bounds for reversible continuous-time Markov chains, J. Appl. Prob. 41 (2004), 1219-1222.
- M.-F. Chen, From Markov chains to non-equilibrium particle systems, 2nd ed. Singapore: World Scientific, 2004.
- B. Cloez, and M. Hairer, Exponential ergodicity for Markov processes with random switching, Bernoulli, 21 (2015), 505–536.

- J. Bardet, H. Guerin, F. Malrieu, Long time behavior of diffusions with Markov switching, ALEA Lat. Am. J. Probab. Math. Stat., 7 (2010), 151-170.
- A.Y. Mitrophanov, The spectral gap and perturbation bounds for reversible continuous-time Markov chains, J. Appl. Prob. 41 (2004), 1219-1222.
- M.-F. Chen, From Markov chains to non-equilibrium particle systems, 2nd ed. Singapore: World Scientific, 2004.
- B. Cloez, and M. Hairer, Exponential ergodicity for Markov processes with random switching, Bernoulli, 21 (2015), 505–536.

J. Bardet, H. Guerin, F. Malrieu, Long time behavior of diffusions with Markov switching, ALEA Lat. Am. J. Probab. Math. Stat., 7 (2010), 151-170.

A.Y. Mitrophanov, The spectral gap and perturbation bounds for reversible continuous-time Markov chains, J. Appl. Prob. 41 (2004), 1219-1222.

M.-F. Chen, From Markov chains to non-equilibrium particle systems, 2nd ed. Singapore: World Scientific, 2004.

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

- M. Pinsky, and R. Pinsky, Transience recurrence and central limit theorem behavior for diffusions in random temporal environments, Ann. Probab. 21 (1993), 433–452.
- J. Shao, Ergodicity of one-dimensional regime-switching diffusion processes, Science China Math., 57 (2014), 2407–2414.
- J. Shao, Criteria for transience and recurrence of regime-switching diffusion processes, Electron. J. Probab., 20 (2015), 1–15.
- J. Shao, and F.B. Xi, Stability and recurrence of regime-switching diffusion processes, SIAM J. Control Optim., 52 (2014), 3496–3516.
- F. Y. Wang, Estimates for Invariant Probability Measures of Degenerate SPDEs with Singular and Path-Dependent Drifts, PTRF (2017).

- M. Pinsky, and R. Pinsky, Transience recurrence and central limit theorem behavior for diffusions in random temporal environments, Ann. Probab. 21 (1993), 433–452.
- J. Shao, Ergodicity of one-dimensional regime-switching diffusion processes, Science China Math., 57 (2014), 2407–2414.
- J. Shao, Criteria for transience and recurrence of regime-switching diffusion processes, Electron. J. Probab., 20 (2015), 1–15.
- J. Shao, and F.B. Xi, Stability and recurrence of regime-switching diffusion processes, SIAM J. Control Optim., 52 (2014), 3496–3516.
- F. Y. Wang, Estimates for Invariant Probability Measures of Degenerate SPDEs with Singular and Path-Dependent Drifts, PTRF (2017).

- M. Pinsky, and R. Pinsky, Transience recurrence and central limit theorem behavior for diffusions in random temporal environments, Ann. Probab. 21 (1993), 433–452.
- J. Shao, Ergodicity of one-dimensional regime-switching diffusion processes, Science China Math., 57 (2014), 2407–2414.
- J. Shao, Criteria for transience and recurrence of regime-switching diffusion processes, Electron. J. Probab., 20 (2015), 1–15.
- J. Shao, and F.B. Xi, Stability and recurrence of regime-switching diffusion processes, SIAM J. Control Optim., 52 (2014), 3496–3516.
- F. Y. Wang, Estimates for Invariant Probability Measures of Degenerate SPDEs with Singular and Path-Dependent Drifts, PTRF (2017).

M. Pinsky, and R. Pinsky, Transience recurrence and central limit theorem behavior for diffusions in random temporal environments, Ann. Probab. 21 (1993), 433–452.

J. Shao, Ergodicity of one-dimensional regime-switching diffusion processes, Science China Math., 57 (2014), 2407–2414.

J. Shao, Criteria for transience and recurrence of regime-switching diffusion processes, Electron. J. Probab., 20 (2015), 1–15.

J. Shao, and F.B. Xi, Stability and recurrence of regime-switching diffusion processes, SIAM J. Control Optim., 52 (2014), 3496–3516.

F. Y. Wang, Estimates for Invariant Probability Measures of Degenerate SPDEs with Singular and Path-Dependent Drifts, PTRF (2017).

M. Pinsky, and R. Pinsky, Transience recurrence and central limit theorem behavior for diffusions in random temporal environments, Ann. Probab. 21 (1993), 433–452.

J. Shao, Ergodicity of one-dimensional regime-switching diffusion processes, Science China Math., 57 (2014), 2407–2414.

J. Shao, Criteria for transience and recurrence of regime-switching diffusion processes, Electron. J. Probab., 20 (2015), 1–15.

J. Shao, and F.B. Xi, Stability and recurrence of regime-switching diffusion processes, SIAM J. Control Optim., 52 (2014), 3496–3516.

F. Y. Wang, Estimates for Invariant Probability Measures of Degenerate SPDEs with Singular and Path-Dependent Drifts, PTRF (2017).