The averaging principle for fully coupled two time-scale stochastic systems

Jinghai SHAO

Center for Applied Mathematics, Tianjin University

@17th Workshop on Markov Processes and Related Topics

Nov. 26, 2022

This talk is based on the works:

- Yonghua, Mao, Shao, *Averaging and large deviation principles for two time-scale regime-switching processes*, preprint 2022
- Shao, *On the application of ergodic condition to averaging principle for multiscale stochastic systems*, preprint 2022

Contents

Consider the SDEs:

$$\mathrm{d}X_t^{\varepsilon,\alpha} = b(X_t^{\varepsilon,\alpha}, Y_t^{\varepsilon,\alpha})\mathrm{d}t + \sqrt{\varepsilon}\sigma(X_t^{\varepsilon,\alpha}, Y_t^{\varepsilon,\alpha})\mathrm{d}B_t, \qquad \text{(e-1)}$$

with initial condition $X_0^{\varepsilon,\alpha} = x_0 \in \mathbb{R}^d, Y_0^{\varepsilon,\alpha} = i$. $(Y_t^{\varepsilon,\alpha})_{t\geq 0}$ is a purely jumping process over $\mathcal{S} = \{1, 2, \ldots, N\}$, $N \leq \infty$, satisfying

$$\mathbb{P}(Y_{t+\delta}^{\varepsilon,\alpha} = j | Y_t^{\varepsilon,\alpha} = i, X_t^{\varepsilon,\alpha} = x) = \begin{cases} \frac{1}{\alpha} q_{ij}(x)\delta + o(\delta), & i \neq j, \\ 1 + \frac{1}{\alpha} q_{ii}(x)\delta + o(\delta), & i = j. \end{cases} \qquad \text{(e-2)}$$

- $\varepsilon, \alpha > 0$, Assume always $\alpha = \alpha(\varepsilon) \to 0$ as $\varepsilon \to 0$.
- $(X_t^{\varepsilon,\alpha})$ is *slow process*, $(Y_t^{\varepsilon,\alpha})$ is *fast process*.
- $X_t^{\varepsilon,\alpha}$ and $Y_t^{\varepsilon,\alpha}$ are fully coupled. Namely, $b(x,y)$ $\sigma(x,y)$, $q_{ij}(x)$.

OUR MAIN CONCERNS:

1. The dependence of Q-matrix $(q_{ij}(x))$ on x
2. The state space $\mathcal{S}$ is infinitely countable
3. The ratio ε/α as $\varepsilon, \alpha \to 0$

Key feature of fully coupled slow-fast system

1. The fast component reaches its equilibrium state at much shorter time, but its equilibrium depends on the state of the slow component.
2. The slow component evolves approximately as a system by averaging its coefficients over the local stationary distributions of the fast component.
3. Such approximations yield a significant model simplification, which is justified mathematically by establishing Averaging Principle

Parts of related works

♣ Both slow process and fast process are diffusions

- R. Khasminskii, 1968
- R. Khasminskii, G. Yin, *J. Differential Eqs.* 2005
- R. Liptser, *PTRF*, 1996.
- A. Veretennikov, *Ann. Probab.* 1999
- A. Puhalskii, *Ann. Probab.* 2016
- W. Liu, Rockner, X. Sun, Y. Xie, *JDE* 2020
- J. Bao, Q. Song, G. Yin, C. Yuan, *SAA*, 2017

Parts of related works

♣ One is diffusion process, another is Markov chain

1. Q. He, G. Yin, *Asymptotic Analysis*, 2014
2. R. Kraaij, M. Schlottke, *A large deviation principle for Markovian slow-fast systems*, arXiv, 2021
3. R. Kumar, L. Popovic, *Large deviations for multi-scale jump-diffusion processes*, Stoch. Proc. Appl. 2017
4. A. Faggionato, D. Gabrielli, M. Crivellari, *Markov Process. Related Fields* 2010
5. A. Budhiraja, P. Dupuis, A. Ganguly, *Electron. J. Probab.* 2018

Ⓒ Ref. 4 and Ref. 5 studied "fully coupled" systems with jumping over a finite state space.

LDP on such a slow-fast system

- R. Kumar, L. Popovic, 2016, SPA
  - General two time-scale jump diffusions
  - Critical assumption: Comparison Principle
  - Nonlinear semigroup method: cf. Jin Feng and Kurtz (2006)
- Budhiraja, Dupuis, Ganguly, 2018, Electron. J. Probab.
  - weak convergence method
  - $\mathcal{S}$ is a finite state space, i.e. $N < \infty$, $\alpha = \varepsilon$
  - establish the averaging principle based on Freidlin-Wentzell (1979): When $\mathcal{S}$ is finite, invariant measure $(\pi_i^x)_{i\in\mathcal{S}}$ for $(q_{ij}(x))$ is given as a ratio of polynomials of transition probabilities, and so $x \mapsto \pi_i^x$ is Lipschitz continuous.

Parts of related works

- Beznidenhout, 1987, Ann. Probab. $\alpha = 1$, $q_{ij}(x)$ independent of x, *Markovian switching*
- A. Eizenberg, M. Freidlin, 1993, Ann. Probab.
  - X_t^ε in a bounded domain, Y_t^α in a finite state space
  - Additive noise, $\alpha \equiv 1$.
- M. Freidlin, Lee, 1996, Probab. Theory Relat. Fields
  - multiplicative noise, $\alpha = \varepsilon$
  - Study the limiting behavior of reaction-diffusion system:

$$\begin{cases} L_i^\varepsilon u^\varepsilon(x,i) + \frac{1}{\varepsilon}\sum_{j=1}^N q_{ij}(x)(u^\varepsilon(x,j) - u^\varepsilon(x,i)) = 0, \ x \in G, \text{bounded domain} \\ u^\varepsilon(x,i)\big|_{\partial G} = g(x,i), \quad i = 1, \dots, N. \end{cases}$$

Contents

Conditions on the coefficients

For the system $(X_t^{\varepsilon,\alpha}, Y_t^{\varepsilon,\alpha})$ given in (e-1), (e-2),

(A1) There exist constants K_1, $K_2 > 0$ such that

$$|b(x,i)-b(y,i)| + \|\sigma(x,i)-\sigma(y,i)\| \leq K_1|x-y|,$$
$$|b(x,i)| + \|\sigma(x,i)\| \leq K_2, \qquad x,y\in\mathbb{R}^d,\ i\in\mathcal{S}.$$

(A2) $\forall\, x\in\mathbb{R}^d$, $(q_{ij}(x))_{i,j\in\mathcal{S}}$ is conservative, irreducible, and

$$\sup_x \sup_{i\in\mathcal{S}} q_i(x) < \infty.$$

(A3) There exists a constant K_3 such that

$$|q_{ij}(x)-q_{ij}(y)| \leq K_3|x-y|, \qquad x,y\in\mathbb{R}^d,\ i,j\in\mathcal{S}.$$

♠ Our challenge in establishing Averaging Principle:

the continuity of $x \mapsto \pi^x$ in $\|\cdot\|_{\mathrm{var}}$.

- The Markov chain P_t is called *ergodic* if

$$\lim_{t\to\infty} \|P_t(i,\cdot) - \pi\|_{\mathrm{var}} = 0, \quad i \in \mathcal{S};$$

- is called *exponentially ergodic* if $\exists\, \lambda, C_i > 0$

$$\|P_t(i,\cdot) - \pi\|_{\mathrm{var}} \le C_i \mathrm{e}^{-\lambda t}, \quad t > 0, i \in \mathcal{S};$$

- is called *strongly ergodic* if

$$\lim_{t\to\infty} \sup_{i\in\mathcal{S}} \|P_t(i,\cdot) - \pi\|_{\mathrm{var}} = 0.$$

Rem. If Markov chain is strongly ergodic, its convergence rate must be

$$\sup_{i\in\mathcal{S}} \|P_t(i,\cdot) - \pi\|_{\mathrm{var}} \le C \mathrm{e}^{-\lambda t}, \quad t > 0.$$

Let P_t^x be the semigroup associated with $(q_{ij}(x))$, and π^x its invariant probability measure.

(A4) Suppose that $\exists\, c_1, \lambda_1 > 0$ such that

$$\sup_{i\in\mathcal{S}} \|P_t^x(i,\cdot) - \pi^x\|_{\mathrm{var}} \le c_1 \mathrm{e}^{-\lambda_1 t}, \quad t > 0, x \in \mathbb{R}^d.$$

Proposition 1. (Strongly ergodic case)

Assume (A2), (A3) and (A4) hold. Then, the functional $x \mapsto \pi^x$ from $\mathbb{R}^d$ to $\mathscr{P}(\mathcal{S})$ is Lipschitz continuous, i.e.

$$\|\pi^x - \pi^y\|_{\mathrm{var}} \le C_\pi |x - y|, \quad x, y \in \mathbb{R}^d,$$

for a constant $C_\pi = 2c_1 K_3/\lambda_1$.

Example 1: For each $x \in (0,1)$, let $(Y_t^x)_{t\geq 0}$ be a birth-death process on $\mathcal{S} = \{1, 2, \ldots\}$ with birth rate $q_{ii+1}(x) = b_i(x) = x$ for $i \geq 1$ and death rate $q_{ii-1}(x) = a_i(x) = 1$ for $i \geq 2$. It is clear that $q_{ij}(x)$ is Lipschitz continuous in x for all $i, j \in \mathcal{S}$. Then,

(i) for each $x \in (0,1)$, the birth-death Markov chain $(Y_t^x)_{t\geq 0}$ is exponentially ergodic and not strongly ergodic, satisfying

$$\|P_t^x(i,\cdot) - \pi^x\|_{\mathrm{var}} \leq C_i(x)\mathrm{e}^{-(1-\sqrt{x})^2 t}, \quad t > 0,\, i \in \mathcal{S},$$

for some $C_i(x) > 0$ depending on $i \in \mathcal{S}$ and $x \in (0,1)$.

(ii) Its invariant probability measure $\pi^x = (\pi_i^x)_{i\geq 1}$ is given by

$$\pi_i^x = (1-x)x^{i-1}, \quad i \geq 1,$$

and satisfies

$$\sup_{x\neq y} \frac{\|\pi^x - \pi^y\|_{\mathrm{var}}}{|x-y|^\beta} = \infty.$$

This means that $x \mapsto \pi^x$ is **not** Hölder continuous with any exponent $\beta \in (0,1]$.

Conditions on the coefficients

(A5) Assume there exist a positive function $\theta : \mathcal{S} \to (0, \infty)$, a decreasing function $\eta : [0, \infty) \to [0, 2]$ satisfying $\int_0^\infty \eta_s \mathrm{d}s < \infty$ such that

$$\|P_t^x(i, \cdot) - \pi^x\|_{\mathrm{var}} \le \theta(i)\eta_t, \quad t \ge 0,\ x \in \mathbb{R}^d,\ i \in \mathcal{S}.$$

Proposition 2. (Weak ergodicity case)

Assume the conditions (A2), (A3) and (A5) hold, then $x \mapsto \pi^x$ is $1/2$-Hölder continuous, i.e.

$$\|\pi^x - \pi^y\|_{\mathrm{var}} \le K_4 \sqrt{|x - y|}, \qquad x, y \in \mathbb{R}^d,$$

where $K_4 = \sqrt{K_3 (\inf_{i \in \mathcal{S}} \theta(i)) \int_0^\infty \eta_s \mathrm{d}s}$.

Recall that $(X_t^{\varepsilon,\alpha})$ satisfies the SDE

$$\mathrm{d}X_t^{\varepsilon,\alpha} = b(X_t^{\varepsilon,\alpha}, Y_t^{\varepsilon,\alpha})\mathrm{d}t + \sqrt{\varepsilon}\sigma(X_t^{\varepsilon,\alpha}, Y_t^{\varepsilon,\alpha})\mathrm{d}B_t, \quad X_0^{\varepsilon,\alpha} = x_0, Y_0^{\varepsilon,\alpha} = i.$$

The limiting system of $X_t^{\varepsilon,\alpha}$ is of the form

$$\mathrm{d}\bar{X}_t = \bar{b}(\bar{X}_t)\mathrm{d}t, \quad \bar{X}_0 = x_0, \tag{e-3}$$

where

$$\bar{b}(x) = \sum_{i\in\mathcal{S}} b(x,i)\pi_i^x,$$

and (π_i^x) is the invariant probability associated with $(q_{ij}(x))$.

- The continuity of $x \mapsto \pi^x$ impacts the continuity of $\bar{b}(x) = \sum_i b(x,i)\pi_i^x$.
- The existence and uniqueness of limit system as $\varepsilon, \alpha \to 0$.
- Does the ratio ε/α as $\varepsilon, \alpha \to 0$ impact the limit system?

Theorem 3 (Averaging Principle:strongly ergodic)

Assume (A1)-(A4) hold, then

$$\lim_{(\varepsilon,\alpha)\to 0} \mathbb{E}|X_t^{\varepsilon,\alpha} - \bar{X}_t| = 0, \quad t > 0.$$

- According to this theorem, under conditions (A1)-(A4), especially (A4), the L^1-convergence of $X_t^{\varepsilon,\alpha}$ to $\bar{X}_t$ does not depend on the ratio of ε/α as $\varepsilon, \alpha \to 0$.

Theorem 4 (Averaging Principle: weakly ergodic)

Assume that (A1)-(A3) and (A5) hold. In addition, suppose that there exist constants $c_2 > 0$, $c_3 < \infty$ such that the function $\theta(\cdot)$ given in (A5) also satisfies

$$Q(x)\theta(i) = \sum_{j\in\mathcal{S}} q_{ij}(x)\theta(j) \le -c_2\theta(i) + c_3, \quad x \in \mathbb{R}^d,\ i \in \mathcal{S}.$$

Then

(i) for each $T > 0$, the set of distributions of $\{(X_t^{\varepsilon,\alpha})_{t\in[0,T]}; \varepsilon, \alpha \in (0,1)\}$ in $\mathcal{C}([0,T];\mathbb{R}^d)$ is tight, and any convergent subsequence of $(X_t^{\varepsilon,\alpha})_{t\in[0,T]}$ shall converges weakly to a solution $(\bar{X}_t)_{t\in[0,T]}$ of ODE (e-3).

(ii) If ODE (e-3) admits a unique solution, then $(X_t^{\varepsilon,\alpha})_{t\in[0,T]}$ converges weakly to the unique solution $(\bar{X}_t)_{t\in[0,T]}$.

An interesting example

Let us recall an example presented in G. Yin, Q. Zhang, *Continuous-time Markov chains and applications: a singular perturbation approach*, 1998.

Example (Yin-Zhang, Example 7.3)

Let $(\Lambda_t^\alpha)_{t\in[0,T]}$ be a continuous time Markov chain on the state space $\mathcal{S} = \{1,2\}$ with transition rate

$$\frac{1}{\alpha}\begin{pmatrix} -\lambda & \lambda \\ \mu & -\mu \end{pmatrix},$$

for some $\lambda, \mu > 0$. Then, for each $T > 0$ the collection of distributions of $(\Lambda_t^\alpha)_{t\in[0,T]}$ for $\alpha \in (0,1)$ is not tight.

Contents

Consider the following two time-scale stochastic systems:

$$\begin{cases} \mathrm{d}X_t^\varepsilon = b(X_t^\varepsilon, Y_t^\varepsilon)\mathrm{d}t + \sigma(X_t^\varepsilon, Y_t^\varepsilon)\mathrm{d}W_t, & X_0^\varepsilon = x_0, \\ \mathrm{d}Y_t^\varepsilon = \frac{1}{\varepsilon} f(X_t^\varepsilon, Y_t^\varepsilon)\mathrm{d}t + \frac{1}{\sqrt{\varepsilon}} g(X_t^\varepsilon, Y_t^\varepsilon)\mathrm{d}B_t, & Y_0^\varepsilon = y_0, \end{cases}$$

where (W_t) and (B_t) are d-dimensional mutually independent Wiener processes, $b(x,y) \in \mathbb{R}^d$ and $f(x,y) \in \mathbb{R}^d$ are drifts, $\sigma(x,y) \in \mathbb{R}^{d\times d}$ and $g(x,y) \in \mathbb{R}^{d\times d}$ are diffusion coefficients.

- For each fixed x, let $(Y_t^{x,y})$ be the solution to SDE:

$$\mathrm{d}Y_t^{x,y} = f(x, Y_t^{x,y})\mathrm{d}t + g(x, Y_t^{x,y})\mathrm{d}B_t, \quad Y_0^{x,y} = y.$$

- Denote the semigroup by P_t^x, the invariant probability measure by π^x.
- Let $\bar{b}(x) = \int b(x,y)\pi^x(\mathrm{d}y)$, $\bar{\sigma}(x) = \int \sigma(x,y)\pi^x(\mathrm{d}y)$,

$$\mathrm{d}\bar{X}_t = \bar{b}(\bar{X}_t)\mathrm{d}t + \bar{\sigma}(\bar{X}_t)\mathrm{d}W_t. \tag{e-4}$$

- The averaging principle suggests that often (X_t^ε) converges to $(\bar{X}_t)$ in certain sense.

In the fully coupled case, Veretennikov (1991) provided a general condition on the fast component:

- There exist functions $\bar{b}$, $\bar{\sigma}$ and $K(T)$ such that $\lim_{T\to\infty} K(T) = 0$, and for all $t \geq 0$, $T > 0$, $x, y \in \mathbb{R}^d$,

$$\Big|\frac{1}{T}\mathbb{E}\Big[\int_t^{t+T} b(x, Y_s^{x,y})\mathrm{d}s\Big] - \bar{b}(x)\Big| \leq K(T)(1 + |x|^2 + |y|^2),$$
$$\Big|\frac{1}{T}\mathbb{E}\Big[\int_t^{t+T} \sigma(x, Y_s^{x,y})\mathrm{d}s\Big] - \bar{\sigma}(x)\Big| \leq K(T)(1 + |x|^2 + |y|^2),$$

- $\bar{\sigma}(x)$ is nondegenerate, continuous in x.
- Aim: wellposedness of the limit system: $\mathrm{d}\bar{X}_t = \bar{b}(\bar{X}_t)\mathrm{d}t + \bar{\sigma}(\bar{X}_t)\mathrm{d}W_t$.
- The crucial point: the continuity of $x \mapsto \pi^x$.

An example

Let (X_t^ε) and (Y_t^ε) be stochastic processes respectively on $[0,1]$ and on $[0,\infty)$ with reflection boundary satisfying

$$\begin{cases} \mathrm{d}X_t^\varepsilon = Y_t^\varepsilon \mathrm{d}t + Y_t^\varepsilon \mathrm{d}W_t, & X_0^\varepsilon = x_0 \in (0,1), \\ \mathrm{d}Y_t^\varepsilon = \frac{1}{\varepsilon}\tilde{f}(X_t^\varepsilon, Y_t^\varepsilon)\mathrm{d}t + \frac{1}{\sqrt{\varepsilon}}\mathrm{d}B_t, & Y_0^\varepsilon = y_0 \in (0,\infty), \end{cases}$$

where

$$\tilde{f}(x,y) = \frac{-x^3\mathrm{e}^{-xy} - (1-x)\mathrm{e}^{-y}}{x^2\mathrm{e}^{-xy} + (1-x)\mathrm{e}^{-y}}, \quad x \in [0,1],\ y \in [0,\infty).$$

For this example, $b(x,y) = \sigma(x,y) = y$ are Lipschitz continuous.

An example

For each $x \in [0,1]$, the invariant probab. measure π^x associated with

$$\mathrm{d}Y_t^{x,y} = \tilde{f}(x, Y_t^{x,y})\mathrm{d}t + \mathrm{d}B_t, \quad Y_0^{x,y} = y \in (0,\infty),$$

is given by

$$\pi^x(\mathrm{d}y) = \big(x^2 \mathrm{e}^{-xy} + (1-x)\mathrm{e}^{-y}\big)\mathrm{d}y.$$

Then

$$\bar{b}(x) := \int_0^\infty b(x,y)\pi^x(\mathrm{d}y) = \begin{cases} 2-x, & \text{if } x \in (0,1], \\ 1, & \text{if } x = 0. \end{cases}$$

- $\bar{b}(x)$ is not continuous at $x = 0$.
- Check directly that for each $x \in [0,1]$ $(Y_t^{x,y})$ is exponentially ergodic but not strongly ergodic.

Consider the parabolic equation:

$$\partial_t u(t,y) = \mathscr{L}^x u(t,y), \quad t>0,\ y\in\mathbb{R}^d, \tag{e-5}$$

with $u(0,y)=h(y)$, $y\in\mathbb{R}^d$ and $h\in C_b(\mathbb{R}^d)$, where $\mathscr{L}^x$ is given by

$$\mathscr{L}^x v(y) = \sum_{k=1}^d f_k(x,y)\frac{\partial v(y)}{\partial y_k} + \frac{1}{2}\sum_{k,l=1}^d G_{kl}(x,y)\frac{\partial^2 v(y)}{\partial y_k \partial y_l}, \quad v\in C^2(\mathbb{R}^d),$$

and $(G_{kl}(x,y)) = (gg^*)(x,y)$.
Let $\Theta^x(t,z;s,y)$, $0\le s<t$, denotes fundamental solution to (e-5).

Proposition 5.

Assume that

1. Strong ergodicity: $\sup_y \|P_t^x(y,\cdot) - \pi^x\|_{\mathrm{var}} \le \kappa_1 \mathrm{e}^{-\lambda_1 t}, \quad t > 0.$
2. $\exists\, \kappa_2 > 0$ $\big(f(x_1, y) - f(x_2, y)\big) \cdot z \le \kappa_2 |x_1 - x_2||z|, \forall, x_1, x_2, y, z \in \mathbb{R}^d.$
3. $\exists\, c_1, c_2 > 0,$

$$|\nabla_z \Theta^x(t, z; s, y)| \le \frac{c_1}{(t-s)^{(d+1)/2}} \mathrm{e}^{-c_2 \frac{|y-z|^2}{t-s}}.$$

4. $g(x, y) = g(y)$ depends only on the fast component.

Then $C > 0$ such that

$$\|\pi^{x_1} - \pi^{x_2}\|_{\mathrm{var}} \le C|x_1 - x_2|^{2/3}.$$

Proposition 6.

Assume that

1. Strong ergodicity: $\sup_y \|P_t^x(y,\cdot)-\pi^x\|_{\mathrm{var}} \le \kappa_1 \mathrm{e}^{-\lambda_1 t}, \quad t>0.$
2. $\exists\, \kappa_3>0,$

$$(f(x_1,y_1)-f(x_2,y_2))\cdot(y_1-y_2)+\|g(x_1,y_1)-g(x_2,y_2)\|^2 \\ \le \kappa_3(|x_1-x_2|^2+|y_1-y_2|^2).$$

Then for $x_1,x_2\in\mathbb{R}^d$

$$\mathbb{W}_{bL}(\pi^{x_1},\pi^{x_2})\le C|x_1-x_2|^{\frac{\lambda_1}{\lambda_1+\kappa_3}}.$$

For $\mu,\nu\in\mathscr{P}(\mathbb{R}^d)$,

$$\mathbb{W}_{bL}(\mu,\nu):=\sup\big\{|\mu(h)-\nu(h)|;|h|\le 1,|h|_{\mathrm{Lip}}\le 1\big\}.$$

Theorem 7. (Averaging principle)

The conditions of Proposition 6 for the fast component are valid. For the slow component, assume that

1. $\exists\, \kappa_4 > 0$ such that

$$|b(x_1, y_1) - b(x_2, y_2)|^2 + \|\sigma(x_1, y_1) - \sigma(x_2, y_2)\|^2 \\ \leq \kappa_4\big(|x_1 - x_2|^2 + |y_1 - y_2|^2\big).$$

2. $\exists\, \kappa_5 > 0$ such that $|b(x, y)| + \|\sigma(x, y)\| \leq \kappa_5(1 + |x|)$.
3. $\inf_{x,y\in\mathbb{R}^d} \inf_{\xi\in\mathbb{R}^d, |\xi|=1} \xi^*(\sigma\sigma^*)(x, y)\xi > 0$.

Then $(X_t^\varepsilon)_{t\in[0,T]}$ converges weakly in $\mathcal{C}([0, T]; \mathbb{R}^d)$ as $\varepsilon \to 0$ to the process $(\bar{X}_t)_{t\in[0,T]}$.

Thank You For Your Attention !

EMAIL:shaojh@tju.edu.cn