TAIWANESE JOURNAL OF MATHEMATICS
Vol. 15, No. 4, pp. 1783-1805, August 2011
This paper is available online at http://www.tjm.nsysu.edu.tw/

# MEAN-FIELD MODELS INVOLVING CONTINUOUS-STATE-DEPENDENT RANDOM SWITCHING: NONNEGATIVITY CONSTRAINTS, MOMENT BOUNDS, AND TWO-TIME-SCALE LIMITS

G. Yin[1], Guangliang Zhao[1] and Fubao Xi[2]

**Abstract.** This work concerns mean-field models, which are formulated using stochastic differential equations. Different from the existing formulations, a random switching process is added. The switching process can be used to describe the random environment and other stochastic factors that cannot be explained in the usual diffusion models. The added switching component makes the formulation more realistic, but it adds difficulty in analyzing the underlying processes. Several properties of the mean-field models are provided including regularity, nonnegativity, finite moments, and continuity. In addition, the paper addresses the issue when the switching takes place an order of magnitude faster than that of the continuous state. It derives a limit that is an average with respect to the invariant measure of the switching process.

## 1. Introduction

This work concerns mean-field models, which are many-body systems with interactions. The origin of the problem stems from statistical mechanics. As is well known, many-body problems are notoriously difficult due to the many bodies involved and the interactions among them. To overcome the difficulties, one of the main ideas is to replace all interactions to any one body with an average or effective interaction, which reduces any multi-body problem to an effective one-body problem. Although the motivation is mainly from statistical mechanics, such models have also enjoyed recent applications in, for examples, graphical models in artificial

Received April 12, 2010, accepted April 12, 2010.
Communicated by Jen-Chih Yao.
2000 *Mathematics Subject Classification*: 60J27, 60J60, 93D30.
*Key words and phrases*: Mean-field model, Random switching, Regularity, Moment, Continuity, Ergodicity, Two-time scale.
[1]The research of this author was supported in part by the National Science Foundation under DMS-0907753, and in part by the Air Force Office of Scientific Research under FA9550-10-1-0210.
[2]The research of this author was supported in part by the Basic Research.

intelligence. It should be mentioned that in financial engineering, a somewhat related idea is termed mean-reversion models, in which one uses a dynamic model that has a force pushing the system moving towards its "equilibrium" in a suitable way.

Concerning the mean-field models, intuitively, if the particles exhibit many interactions in the original system, the mean field will be more accurate for such a system. The usefulness, the potential impact on many practical situations, and the challenges arise have attracted much attention in recent years. The work [1] presents a detailed study on cooperative behavior of such systems, and the subsequent work [2] delineates the law of large numbers and central limit theorem for jump mean fields; see also related work [4, 9] and references therein. From a probabilistic view point, it is important to have an in-depth understanding of basic properties of such systems.

Owing to the progress in technology, more complicated systems are encountered in applications. In response to such challenges, much effort has been devoted to modeling and analysis for sophisticated systems. One of the ideas is to bring regime switching into the formulation, so as to deal with the coexistence of continuous dynamics and discrete events. For example, the underlying dynamic systems may be influenced by or subject to not only the usual dynamics represented by differential equations, but also movements that exhibit jump or switching behavior. These discrete events are used to depict random environment or other uncertainty. Recently, there are growing interests on formulating complex systems by use of regime-switching processes, which largely enriched the applicability of the dynamic models; see [14] and many references therein.

Continuing our effort in the study of regime-switching diffusions, this work focuses on investigating properties of regime-switching mean-field models. This paper is a continuation of our recent work [11], in which regularity, Feller continuity, strong Feller continuity, and exponential ergodicity are obtained. In the previous work, for example, in [1] as well as in [11], each component of the system is allowed to take values in $\mathbb{R}$. That is, any of the $r$ bodies is allowed to take negative values. However, in statistical physics, typically, these many bodies are only allowed to be nonnegative. Thus it will be more realistic to consider a formulation with nonnegativity constraint. In this paper, we take nonnegativity constraint into consideration, which puts further challenges to the analysis. If $\mathbb{R}$ is used, to ensure the system is non-explosive, it suffices to verify the regularity. Under the nonnegativity constraint, it is necessary to show that each component of the system remains to be nonnegative or to be confined to the first quadrant only. This in turn, requires more careful analysis and special attention. In addition, we are interested in getting several moment bounds. With such bounds at our hands, we can proceed to obtain sample continuity as well as further asymptotic behavior. Furthermore, when the switching process is varying an order of magnitude faster than the continuous state, certain average takes place. We show that the continuous state process has a limit, which is an average with respect to the quasi-stationary measure of the fast varying switching process (more precise definition will be given in the subsequent section). This limit can be obtained by means of a martingale problem formulation.

The rest of the paper is arranged as follows. The precise formulation of the mean-field model is provided next. Section 3 concerns properties of solutions of the stochastic differential equations for the mean-field models. Regularity together with the existence of the solution in the quarter plane is provided. Also given in that section are moment properties, existence of moment generating functions, and sample path continuity. Section 4 continues with positive recurrence. Section 5 proceeds with the study of large-time asymptotic properties. Section 6 examines asymptotic properties of a mean-field model, in which the switching process is fast varying. When the switching takes place an order of magnitude faster than that of the continuous state, we derive a limit that is an average with respect to the invariant measure of the switching process. Finally, the paper is concluded with some further remarks.

## 2. Formulation

Suppose that $\alpha(\cdot)$ is a randomly switching process taking values in $\mathcal{M} := \{1, \dots, m\}$, and that $\gamma(\cdot)$, $\beta(\cdot) : \mathcal{M} \mapsto \mathbb{R}_+$, where $\mathbb{R}_+ := \{z \in \mathbb{R} : z > 0\}$. Consider an $r$-body mean-field model with switching described by the following system of stochastic differential equations. For $i = 1, 2, \dots, r$,

$$\begin{aligned} dX_i(t) &= \big[\gamma(\alpha(t))X_i(t) - X_i^3(t) - \beta(\alpha(t))(X_i(t) - \overline{X}(t))\big]dt \\ &\quad + \sigma_{ii}(X(t), \alpha(t))dW_i(t), \end{aligned} \tag{2.1}$$

where $W_i(\cdot)$ is a one-dimensional standard Brownian motion,

$$\overline{X}(t) = \frac{1}{r}\sum_{i=1}^{r} X_i(t),\ \ X(t) = \big(X_1(t), X_2(t), \dots, X_r(t)\big)',$$

$x'$ denotes the transpose of $x$. For $\alpha \in \mathcal{M} = \{1, \dots, m\}$, the transition rules of $\alpha(t)$ are specified by

$$\mathbf{P}\{\alpha(t+\Delta) = k | \alpha(t) = \alpha, X(t) = x\} = q_{\alpha k}(x)\Delta + o(\Delta) \ \text{ if } \ k \neq \alpha, \tag{2.2}$$

where $\Delta \downarrow 0$ and $\sum_{k\in\mathcal{M}} q_{\alpha k}(x) = 0$ for each $\alpha \in \mathcal{M}$.

We will use the following assumptions, which are conditions on the coefficients of (2.1) and the transition rate (2.2). Note that (A1) allows the diffusion to grow at the order of $4 - \delta$ for some $\delta > 0$, whereas (A2) allows similar growth rate and requires also $\sigma_{ii}(0) = 0$. Condition (A1) is sufficient to ensure the existence and uniqueness of the solution of the switching stochastic differential equation, and condition (A2) enables us to obtain further properties such as nonnegativity etc. More details will be seen in the subsequent sections.

$(A1)$ The $Q(\cdot)$ is bounded and continuous. For each $\alpha \in \mathcal{M}$ and $x \in \mathbb{R}^r$,

(a) $q_{\alpha k}(x) > 0$ for $k \neq \alpha$ and $\sigma_{ii}(x,\alpha) > 0$ for each $1 \leq i \leq r$;

(b) $\sigma_{ii}(x,\alpha)$ and $q_{\alpha k}(x)$ are locally Lipschitz with respect to $x$;

(c) $\sigma_{ii}(x,\alpha)$ is infinitely differentiable in $x$;

(d) there exist constants $K_0 > 0$ and $\delta > 0$ such that

$$\sum_{i=1}^{r} \sigma_{ii}^2(x,\alpha) \leq K_0(|x|^{4-\delta} + 1). \tag{2.3}$$

$(A2)$ Assume (A1) but with (2.3) replaced by

$$\sum_{i=1}^{r} \sigma_{ii}^2(x,\alpha) \leq K_0|x|^{4-\delta}. \tag{2.4}$$

To proceed, it is convenient to use a vector notation. For $(x,\alpha) \in \mathbb{R}^r \times \mathcal{M}$, set

$$b(x,\alpha) = \begin{pmatrix} b_1(x,\alpha) \\ b_2(x,\alpha) \\ \vdots \\ b_r(x,\alpha) \end{pmatrix} = \begin{pmatrix} \gamma(\alpha)x_1 - x_1^3 - \beta(\alpha)(x_1 - \overline{x}) \\ \gamma(\alpha)x_2 - x_2^3 - \beta(\alpha)(x_2 - \overline{x}) \\ \vdots \\ \gamma(\alpha)x_r - x_r^3 - \beta(\alpha)(x_r - \overline{x}) \end{pmatrix} \in \mathbb{R}^r, \tag{2.5}$$

and $\sigma(x,\alpha) = \text{diag}\{\sigma_{ii}(x,\alpha)\} \in \mathbb{R}^r \times \mathbb{R}^r$, where $\overline{x} := \sum_{j=1}^r x_j/r$. Then stochastic differential equation (2.1) can be rewritten as

$$dX(t) = b(X(t),\alpha(t))dt + \sigma(X(t),\alpha(t))dW(t). \tag{2.6}$$

For a function $f(\cdot,\cdot) : \mathbb{R}^r \times \mathcal{M} \mapsto \mathbb{R}$ such that $f(\cdot,\alpha)$ is twice continuously differentiable with respect to the variable $x$ for each $\alpha \in \mathcal{M}$, the operator associated with the switching diffusion is given by

$$\begin{aligned} \mathcal{L}f(x,\alpha) = {} & \frac{1}{2}\sum_{i=1}^{r} \sigma_{ii}^2(x,\alpha)\frac{\partial^2 f(x,\alpha)}{\partial x_i^2} + \sum_{i=1}^{r} b_i(x,\alpha)\frac{\partial f(x,\alpha)}{\partial x_i} \\ & + \sum_{k\in\mathcal{M}} q_{\alpha k}(x)\big(f(x,k) - f(x,\alpha)\big). \end{aligned} \tag{2.7}$$

## 3. Properties of Solutions

### 3.1. Regularity and Existence of Solutions

We begin by stating an existence and uniqueness of solution for the system of differential equations of interest. Its proof can be found in [11, Theorem 3.3].

**Lemma 3.1.** *Assume condition* (A1). *Then for each initial condition* $(X(0), \alpha(0)) = (X_0, \alpha)$ *with* $\alpha \in \mathcal{M} = \{1, \dots, m\}$, *there exists a unique solution* $(X(t), \alpha(t))$ *to* (2.6) *and* (2.2) *for* $t \geq 0$.

**Remark 3.2.** To proceed, we explore the regularity and nonnegativity of solutions to (2.6) and (2.2). To get the regularity only, one can use a Liapunov function $V(x, \alpha) = |x|$. Then it can be verified that $\mathcal{L}V(x, \alpha) \leq cV(x, \alpha)$ for some $c > 0$. However, to show that the process will remain in the first quadrant, more complex Liapunov function is needed as can be seen in the proof to follow.

**Theorem 3.3.** *Assume* (A2) *and* $X_0 \in \mathbb{R}_+^r := \{(x_1, \dots, x_r) : \ x_i > 0, i = 1, \dots, r\}$. *Then the solution to* (2.6) *will remain in* $\mathbb{R}_+^r$ *almost surely. That is,* $X(t) \in \mathbb{R}_+^r$ *a.s. for any* $t \geq 0$.

*Proof.* Consider (2.1). Using an argument of [5] for diffusions, assumption (A2) indicates that the coefficients of the stochastic differential equation (2.6) are locally Lipschitz and "locally" linear growth; see [14, Chapter 2]. Therefore, there is an explosion time $\rho_e$ such that for all $t \in [0, \rho_e)$, there exists a local solution for (2.6). Let $k_0 > 0$ be sufficiently large such that $X_i(0) \in ((1/k_0), k_0)$ for each $i = 1, \dots, r$. For each $k \geq k_0$, we define

$$\tau_k := \inf \left\{ t \in [0, \rho_e) : \ X_i(t) \notin \left( \frac{1}{k}, k \right) \ \text{ for some } \ i = 1, 2, \dots, r \right\}. \tag{3.1}$$

The sequence $\tau_k, k = 1, 2, \dots$ is monotonically increasing. Set $\tau_\infty := \lim_{k \to \infty} \tau_k$. Then $\tau_\infty \leq \rho_e$.

We are in a position to prove $\tau_\infty = \infty$ a.s. Suppose that this were not true. Then there would exist a $T > 0$ and $\varepsilon > 0$ such that $\mathbf{P}\{\tau_\infty < T\} > \varepsilon$. Thus, there is a $k_1$ such that $P\{\tau_k < T\} > \varepsilon$ for all $k \geq k_1$. Denote

$$S(x) = \sum_{i=1}^{r} x_i, \tag{3.2}$$

and define a Liapunov function

$$V(x, \alpha) = \sum_{i=1}^{r} x_i - \log S(x) \quad \text{ where } \ (x, \alpha) \in \mathbb{R}_+^r \times \mathcal{M}.$$

It is easily seen that

$$\frac{\partial}{\partial x_i} \log S(x) = \frac{1}{S(x)}, \quad \frac{\partial^2}{\partial x_i^2} \log S(x) = -\frac{1}{S^2(x)}.$$

Direct calculation leads to

$$
\begin{aligned}
\mathcal{L}V(x,\alpha) &= \sum_{i=1}^{r}[\gamma(\alpha)x_i - x_i^3 - \beta(\alpha)(x_i-\overline{x})] \\
&\quad + \frac{1}{S(x)}\sum_{i=1}^{r}[-\gamma(\alpha)x_i + x_i^3 + \beta(\alpha)(x_i-\overline{x})] \\
&\quad + \frac{1}{2}\frac{1}{S^2(x)}\sum_{i=1}^{r}\sigma_{ii}^2(x,\alpha) \quad \text{for each} \quad \alpha\in\mathcal{M}.
\end{aligned} \tag{3.3}
$$

Since $x_i > 0$ for each $i$, using the familiar inequality

$$
\sum_{i=1}^{r} x_i^p \le \Big(\sum_{i=1}^{r} x_i\Big)^p, \quad p>1, \tag{3.4}
$$

$$
\begin{aligned}
&-\frac{1}{S(x)}\sum_{i=1}^{r}\gamma(\alpha)x_i = -\gamma(\alpha), \\
&\frac{1}{S(x)}\sum_{i=1}^{r}\beta(\alpha)(x_i-\overline{x}) = 0, \\
&\frac{1}{S(x)}\sum_{i=1}^{r}x_i^3 \le \frac{1}{S(x)}\Big(\sum_{i=1}^{r}x_i\Big)^3 = S^2(x), \\
&\frac{1}{S(x)}\sum_{i=1}^{r}\sigma_{ii}(x,\alpha) \le \frac{1}{S^2(x)}\sum_{i=1}^{r}x_i^{4-\delta} \le S^{2-\delta}(x).
\end{aligned} \tag{3.5}
$$

Using (3.5) in (3.3), detailed estimates lead to that when $x_i > 0$ is large, the value of $\mathcal{L}V(x,\alpha)$ is dominated by $-x_i^3$; when $x_i>0$ is small, the value of $\mathcal{L}V(x,\alpha)$ is dominated by a constant by using the bound of $\sigma(x,\alpha)$ in assumption (2.4). Thus, in any event,

$$
\mathcal{L}V(x,\alpha) \le K \quad \text{where} \quad K>0 \quad \text{is independent of} \quad k. \tag{3.6}
$$

By virtue of the definitions of $\tau_k$ and $V(x,\alpha)$,

$$
V(X(\tau_k),\alpha(\tau_k)) \ge (k - r\log k)\wedge\Big(\frac{1}{k}+\log k\Big).
$$

By means of Dynkin's formula,

$$
\begin{aligned}
&\mathbf{E}V(X(T\wedge\tau_k),\alpha(T\wedge\tau_k)) - V(X(0),\alpha(0)) \\
&= \mathbf{E}\int_0^{\tau_k\wedge T}\mathcal{L}V(X(s),\alpha(s))ds \\
&\le KT.
\end{aligned}
$$

By rearrangement,

$$
\begin{aligned}
KT + V(X(0),\alpha(0)) &\ge \mathbf{E}V(X(\tau_k\wedge T),\alpha(\tau_k\wedge T)) \\
&\ge \mathbf{E}V(X(\tau_k),\alpha(\tau_k))I_{\{\tau_k<T\}}
\end{aligned}
$$

$$
\begin{aligned}
&\geq (k - r\log k)) \wedge \Big(\frac{1}{k} + \log k\Big)\mathbf{P}(\tau_k < T) \\
&\geq \Big[(k - r\log k) \wedge \Big(\frac{1}{k} + \log k\Big)\Big]\,\varepsilon \\
&\to \infty \quad \text{as } k \to \infty.
\end{aligned}
$$

This is a contradiction. As a result, $\lim_{k\to\infty} \tau_k = \infty$ a.s. and hence the explosion time $\rho_e = \infty$ a.s. ■

### 3.2. Moment Properties

In this section, we consider moment properties of the process $X(t)$. We show that the moment generating function

$$M(z) = \mathbf{E}\exp(z'X(t)), \quad \text{for any} \quad t \geq 0,\ z' = (z_1, \dots, z_r) \in \mathbb{R}^r \quad \text{with} \quad z_i \in \mathbb{R}$$

exists. To proceed, we first obtain a finite moment result.

**Lemma 3.4.** *Under the conditions of Theorem* 3.3*, for any* $p \geq 2$*,*

$$\sup_{t\geq 0} \mathbf{E}[\sum_{i=1}^{r} X_i^p(t)] \leq K < \infty.$$

*Proof.* For any $(x,\alpha) \in \mathbb{R}_+^r \times \mathcal{M}$, consider $V(x,\alpha) = S^p(x)$ with $S(x)$ defined in (3.2). Using the stopping time $\tau_k$ defined in (3.1), we have

$$
\begin{aligned}
\mathcal{L}V(x,\alpha) &= pS^{p-1}(x)\sum_{i=1}^{r}[\gamma(\alpha)x_i - x_i^3 - \beta(\alpha)(x_i - \overline{x})] \\
&\quad + \frac{1}{2}p(p-1)S^{p-2}(x)\sum_{i=1}^{r}\sigma_{ii}^2(x,\alpha) \\
&= p\gamma(\alpha)S^p(x) - pS^{p-1}(x)\sum_{i=1}^{r} x_i^3 + \frac{1}{2}p(p-1)S^{p-2}\sum_{i=1}^{r}\sigma_{ii}^2(x,\alpha).
\end{aligned}
$$

By virtue of Dynkin's formula,

$$
\begin{aligned}
&\mathbf{E}[e^{t\wedge\tau_k}S^p(X(t\wedge\tau_k)) - S^p(X(0)) \\
&\quad = \mathbf{E}\int_0^{t\wedge\tau_k} e^s[V(X(s),\alpha(s)) + \mathcal{L}V(X(s),\alpha(s))]ds \\
&\quad \leq \mathbf{E}\int_0^{t\wedge\tau_k} e^s\Big([1+p\gamma(\alpha)]S^p(x) + \frac{1}{2}p(p-1)S^{p-2}\Big(\sum_{i=1}^{r} x_i\Big)^2\Big)ds \qquad (3.7)\\
&\quad \leq \mathbf{E}\int_0^{t\wedge\tau_k} e^s K ds \\
&\quad \leq K(e^t - 1).
\end{aligned}
$$

Since $\lim_{k\to\infty} \tau_k = \infty$ a.s., letting $k \to \infty$, we obtain

$$\mathbf{E}[e^t S^p(X(t))] - S^p(X(0)) \le K(e^t - 1)$$

so

$$\mathbf{E}S^p(X(t)) \le e^{-t} S^p(X(0)) + K(1 - e^{-t}) \le K < \infty.$$

Using (3.4) and taking $\sup_{t\ge 0}$, the desired result then follows. ■

By virtue of Theorem 3.3, we can show $\mathbf{E}X_i(t) \le K < \infty$. This together with Lemma 3.4 yields that for any positive integer $l$, $\mathbf{E}X_i^l(t) \le K < \infty$ for each $i = 1, \dots, r$. Thus for any given $z \in \mathbb{R}^r$, we have $\mathbf{E}z_i^l X_i^l(t) \le K < \infty$ for each $i = 1, \dots, r$. As a result,

$$\sum_{l=0}^{\infty} \sum_{i=1}^{r} \frac{z_i^l \mathbf{E}X_i^l(t)}{l!} \quad \text{converges absolutely and uniformly.}$$

The existence of the moment generating function then follows. We summarize this into the following proposition.

**Proposition 3.5.** *The moment generating function $M(z) = \mathbf{E}\exp(z'X(t))$ exists for any $z \in \mathbb{R}^r$ and $t \ge 0$.*

### 3.3. Continuity

This section establishes the sample path continuity. In fact, the desired result is obtained by means of an auxiliary bound. To proceed, we first establish the following lemma.

**Lemma 3.6.** *Under the conditions of Theorem* 3.3, *for any positive integer $\kappa$, $0 < T < \infty$, and any $0 \le t, s \le T$, there is a positive constant $K$ such that*

$$\mathbf{E}|X(t) - X(s)|^{2\kappa} \le K|t-s|^{\kappa}. \tag{3.8}$$

*Proof.* It suffices to examine each component $X_i(\cdot)$. It is easily seen that for any positive integer $\kappa$, $0 < T < \infty$, and any $0 \le t, s \le T$,

$$X_i(t) - X_i(s) = \int_s^t b_i(X(u), \alpha(u))du + \int_s^t \sigma_{ii}(X(u), \alpha(u))dW_i(u),$$

and as a result

$$\begin{aligned}&\Big|X_i(t) - X_i(s)\Big|^{2\kappa}\\ &\le 2^{2\kappa-1}\left[\left|\int_s^t b_i(X(u), \alpha(u))du\right|^{2\kappa} + \left|\int_s^t \sigma_{ii}(X(u), \alpha(u))dW_i(u)\right|^{2\kappa}\right].\end{aligned}$$

An application of the Hölder inequality leads to

$$\begin{aligned}
&\mathbf{E}\left|\int_s^t b_i(X(u),\alpha(u))du\right|^{2\kappa}\\
&\leq \left(\int_s^t du\right)^{2\kappa-1}\int_s^t \mathbf{E}|b_i(X(u),\alpha(u))|^{2\kappa}du\\
&\leq K(t-s)^{2\kappa}.
\end{aligned}\tag{3.9}$$

The last line above follows from the moment estimate in Lemma 3.4.

Next, we estimate the diffusion term. By using [7, Lemma 4.12, p. 131], we have

$$\begin{aligned}
&\mathbf{E}\left|\int_s^t \sigma_{ii}(X(u),\alpha(u))dW_i(u)\right|^{2\kappa}\\
&\leq [\kappa(2\kappa-1)^{\kappa}(t-s)^{\kappa-1}]\int_s^t \mathbf{E}|\sigma_{ii}(X(u),\alpha(u))|^{2\kappa}du\\
&\leq K(t-s)^{\kappa}.
\end{aligned}\tag{3.10}$$

Combining the estimates in (3.9) and (3.10), the desired moment estimate follows. ∎

**Theorem 3.7.** *Under the conditions of Theorem* 3.3*, the process* $X(\cdot)$*, which is the solution of* (2.1)*, has continuous sample paths almost surely.*

*Proof.* It suffices to examine each component. The assertion is a direct consequence of Lemma 3.6 and the well-known Kolmogorov continuity criterion [10, Theorem 2, p. 3]. ∎

## 4. Positive Recurrence

Recall that an $\mathbb{R}^r$-valued Markov process $\xi(t)$ satisfying $\xi(0)=x$ (denoted by $\xi^x(t)$ when we want to emphasize the initial data $x$-dependence) is recurrent with respect to some nonempty bounded open set $G\subset\mathbb{R}^r$ if $\mathbf{P}\{\tau^x<\infty\}=1$ for any $x\notin G$, where $x=\xi(0)$ and $\tau^x$ is the hitting time of $G$ for $\xi^x(t)$ (i.e., the first time that the process $\xi^x(t)$ enters the set $G$, or $\tau^x:=\inf\{t\geq 0:\xi^x(t)\in G\}$). The process $\xi(t)$ is said to be positive recurrent with respect to $G$ if $\mathbf{E}\tau^x<\infty$ for any $x\notin G$.

**Theorem 4.1.** *The solution is positive recurrent with respect to the domain*

$$G_\rho:=\{x\in\mathbb{R}^r_+:\ 0<x_i<\rho,\ i=1,2,\dots,r\},$$

*where* $\rho$ *is a positive number to be specified.*

*Proof.* By virtue of [14], it suffices for each $\alpha \in \mathcal{M}$ to find a nonnegative Liapunov function $V(\cdot, \alpha)$ defined on $(\mathbb{R}^r_+ - \overline{G}_\rho) \times \mathcal{M}$ such that $V(\cdot, \alpha)$ is twice continuously differentiable with respect to $x$ and satisfies $\mathcal{L}V(x, \alpha) \le -1$ for all $(x, \alpha) \in (\mathbb{R}^r_+ - \overline{G}_\rho) \times \mathcal{M}$, where $\overline{G}_\rho$ denotes the closure of $G_\rho$. We consider the nonnegative function

$$V(x, \alpha) = \sum_{i=1}^r (\log x_i)^2, \qquad (x, \alpha) \in (\mathbb{R}^r_+ - \overline{G}_\rho) \times \mathcal{M}.$$

Note that for getting the positive recurrence, it suffices to work with a Liapunov function that is defined in the exterior of the bounded set $G_\rho$. Note also that when $x_i > e$, $\log x_i > 1$. Thus, in view of the fact $\log x_i / x_i \le 1$,

$$\begin{aligned} \sum_{i=1}^r 2\beta(\alpha) \frac{\log x_i}{x_i} \sum_{j=1}^r \frac{x_j}{r} &\le \frac{2\beta(\alpha)}{r} \sum_{i=1}^r \sum_{j=1}^r x_j \\ &= 2\beta(\alpha) \sum_{i=1}^r \log x_i \frac{x_i}{\log x_i} \\ &\le \sum_{i=1}^r 2\beta(\alpha) x_i \log x_i \end{aligned} \tag{4.1}$$

when $\rho$ is large enough. Note also that for $x_i > e$, $(1 - \log x_i)/x_i^2 < 0$. As a result,

$$\sum_{i=1}^r \sigma_{ii}^2(x, \alpha) \frac{1 - \log x_i}{x_i^2} < 0. \tag{4.2}$$

Consequently, using (4.1) and (4.2),

$$\begin{aligned} \mathcal{L}V(x, \alpha) &= \sum_{i=1}^r 2 \log x_i (\gamma(\alpha) - x_i^2 - \beta(\alpha)) + \sum_{i=1}^r 2\beta(\alpha) \frac{\log x_i}{x_i} \sum_{j=1}^r \frac{x_j}{r} \\ &\quad + \sum_{i=1}^r \sigma_{ii}^2(x, \alpha) \frac{1 - \log x_i}{x_i^2} \\ &< \sum_{i=1}^r 2 \log x_i (\gamma(\alpha) - x_i^2 - \beta(\alpha) + \beta(\alpha) x_i) \\ &= \sum_{i=1}^r 2 \log x_i [-(x_i - \frac{\beta(\alpha)}{2})^2 + (\frac{\beta(\alpha)}{2})^2 - \beta(\alpha) + \gamma(\alpha)] \end{aligned} \tag{4.3}$$

Thus we can find $\rho$ large enough such that for all $(x, \alpha) \in (\mathbb{R}^r_+ - \overline{G}_\rho) \times \mathcal{M}$, $\mathcal{L}V(x, \alpha) < -1$. Therefore $X(t)$ is positive recurrent with respect to the domain $G_\rho$.

Since the process is positive recurrent with respect to $G_\rho$, there is an invariant measure there ([15] and also [14, Chapter 4]). Furthermore, there is an invariant density for the joint process $(X(t), \alpha(t))$, denoted by $\{\mu(x,i) : i \in \mathcal{M}\}$ such that

$$\mathcal{L}^* \mu(x,i) = 0, \quad \sum_{i \in \mathcal{M}} \int_{\mathbb{R}^r} \mu(x,i) dx = 1,$$

where $\mathcal{L}^*$ is the adjoint of the operator $\mathcal{L}$.

## 5. Further Asymptotic Bounds

In this section, we derive further asymptotic bounds in the sense of almost sure estimates. The result reveals long-time behavior and stability.

**Theorem 5.1.** *The solution $X(t)$ satisfies*

$$\limsup_{T \to \infty} \frac{\log (\overline{X}(T))}{\log T} \le K \quad \textit{a.s.}$$

*for some $K > 0$.*

*Proof.* Choose

$$V(t,x,\alpha) = e^t \log (\overline{x}) \quad \text{for} \quad (t,x,\alpha) \in [0,\infty) \times \mathbb{R}^r_+ \times \mathcal{M},$$

where $\overline{x} = \sum_{i=1}^r x_i / r$. It is readily seen that

$$\frac{\partial V}{\partial t} = e^t \log(\overline{x}), \quad \frac{\partial V}{\partial x_i} = \frac{1}{r\overline{x}} e^t, \quad \text{and} \quad \frac{\partial^2 V}{\partial x_i^2} = -\frac{1}{r^2 \overline{x}^2} e^t.$$

Then by virtue of Itô's formula,

$$\begin{aligned}
&e^t \log (\overline{X}(t)) - \log (\overline{X}(0)) \\
&= \int_0^t e^s \log (\overline{X}(s)) ds \\
&\quad + \int_0^t e^s \Big\{ \sum_{i=1}^r \frac{1}{r\overline{X}(s)} [\gamma(\alpha(s)) X_i(s) - X_i^3(s) - \beta(\alpha(s))(X_i(s) - \overline{X}(s))] \\
&\quad + \sum_{i=1}^r \frac{1}{2} \sigma_{ii}^2 (X(s), \alpha(s)) \frac{-1}{r^2 \overline{X}^2(s)} \Big\} ds \\
&\quad + \int_0^t e^s \sum_{i=1}^r \frac{\sigma_{ii}(X(s), \alpha(s))}{r\overline{X}(s)} dw_i(s).
\end{aligned} \tag{5.1}$$

Denote

$$M_i(t) = \int_0^t e^s \frac{1}{r\overline{X}(s)} \sigma_{ii}(X(s), \alpha(s)) dw_i(s),$$

whose quadratic variation is

$$\langle M_i, M_i\rangle(t) = \int_0^t e^{2s}\frac{\sigma_{ii}^2(X(s),\alpha(s))}{r^2\overline{X}^2(s)}ds.$$

By the familiar exponential martingale inequality (e.g., [8, p. 49]), for any positive constants $T$, $\delta$, and $\eta$, we have

$$\mathbf{P}\{\sup_{0\le t\le T}[M_i(t) - \frac{\delta}{2}\langle M_i, M_i\rangle(t)] > \eta\} \le e^{-\delta\eta}.$$

Choose

$$T = k\zeta,\ \ \delta = \varepsilon e^{-k\zeta},\ \ \text{and}\ \ \eta = \frac{\theta e^{k\zeta}\log k}{\varepsilon},$$

where $k \in \mathbb{N}$, $0 < \varepsilon < 1$, $\theta > 1$, and $\zeta > 0$. Then similarly to that of [16], we can show that

$$M_i(t) \le \frac{\varepsilon e^{-k\zeta}}{2}\langle M_i, M_i\rangle(t) + \frac{\theta e^{k\theta}\log k}{\varepsilon},\ \text{for all } 0 \le t \le k\gamma.$$

$$\begin{aligned} & e^t\log(\overline{X}(t)) - \log(\overline{X}(0)) \\ & \le \int_0^t e^s\{\log\overline{X}(s) + \sum_{i=1}^r\frac{1}{r\overline{X}(s)}[\gamma(\alpha(s))X_i(s) \\ & -X_i^3(s) - \beta(\alpha(s))(X_i(s) - \overline{X}(s))] \\ & + \sum_{i=1}^r\frac{\sigma_{ii}^2}{2}(X(s),\alpha(s))\frac{-1}{r^2\overline{X}^2(s)} \\ & + \sum_{i=1}^r\frac{\varepsilon e^{-k\zeta}e^s\sigma_{ii}^2(X(s),\alpha(s))}{2}\frac{1}{r^2\overline{X}^2(s)}\}ds + \frac{r\theta e^{k\zeta}\log k}{\varepsilon}. \end{aligned} \tag{5.2}$$

Since $\varepsilon e^{-k\zeta}e^s < 1$,

$$\sum_{i=1}^r\frac{\sigma_{ii}^2(X(s),\alpha)}{2}\frac{-1}{r^2\overline{x}^2} + \frac{\varepsilon e^{-k\zeta}e^s\sigma_{ii}^2(x,\alpha)}{2}\frac{1}{r^2\overline{x}^2} = \sum_{i=1}^r\frac{\sigma_{ii}^2(x,\alpha)}{2r^2\overline{x}^2}(-1+\varepsilon e^{-k\zeta}e^s) < 0.$$

In addition, note that

$$\frac{1}{r\overline{x}}\sum_{i=1}^r[\gamma(\alpha)x_i - \beta(\alpha)(x_i - \overline{x})] = \gamma(\alpha).$$

Using the Hölder inequality,

$$\Big(\sum_{i=1}^r x_i\Big)^3 \le \Big(\sum_{i=1}^r 1^{3/2}\Big)^2\sum_{i=1}^r x_i^3 \le r^2\sum_{i=1}^r x_i^3.$$

As a result,

$$\log \overline{x} - \frac{1}{r\overline{x}} \sum_{i=1}^{r} x_i^3 \le \log \overline{x} - \frac{1}{r^3 \overline{x}} \Big( \sum_{i=1}^{r} x_i \Big)^3 = \log \overline{x} - \overline{x}^2 < 0.$$

Using the above estimates, we obtain

$$\log \overline{x} + \frac{1}{r\overline{x}} \sum_{i=1}^{r} [\gamma(\alpha) x_i - x_i^3 - \beta(\alpha)(x_i - \overline{x})] \le \gamma(\alpha). \tag{5.3}$$

Therefore, using (5.3) and (5.2), we obtain

$$\begin{aligned} e^t \log(\overline{X}(t)) - \log(\overline{X}(0)) &\le \int_0^t e^s C ds + \frac{r\theta e^{k\zeta} \log k}{\varepsilon} \\ &= C(e^t - 1) + \frac{r\theta e^{k\zeta} \log k}{\varepsilon}, \end{aligned} \tag{5.4}$$

where $C$ is a positive constant. Thus for $(k-1)\zeta \le t \le k\zeta$, similar to the development in [16], by sending $\zeta \downarrow 0$, $\varepsilon \uparrow 1$, and $\theta \downarrow 1$, we have

$$\limsup_{T \to \infty} \frac{\log \overline{X}(T)}{\log T} \le K \quad \text{a.s.}$$

The result is proved. ■

As a direct consequence, we obtain the following corollary.

**Corollary 5.2.** *Under the conditions of Theorem* 5.1*,*

$$\limsup_{T \to \infty} \frac{\log \overline{X}(T)}{T} \le 0 \quad \textit{a.s.,}$$

*and*

$$\limsup_{T \to \infty} \frac{\log |X(T)|}{T} \le 0 \quad \textit{a.s.}$$

## 6. A Two-time-scale Limit

This section is concerned with a class of mean field processes, in which the random switching process changes an order of magnitude faster than the continuous state (or the switching process jump change much more frequently). The basic premise is that there are inherent two-time scales. Our interest focuses on the limit behavior of the resulting process. Suppose that $\varepsilon > 0$ is a small parameter and the system of mean field equations is given by

$$\begin{aligned} dX_i^\varepsilon(t) = &\big[\gamma(\alpha(t)) X_i^\varepsilon(t) - (X_i^\varepsilon(t))^3 - \beta(\alpha^\varepsilon(t))(X_i^\varepsilon(t) - \overline{X}^\varepsilon(t))\big] dt \\ &+ \sigma_{ii}(X^\varepsilon(t), \alpha^\varepsilon(t)) dW_i(t), \end{aligned} \tag{6.1}$$

or using the definition of $b(x, \alpha)$,

$$dX^{\varepsilon}(t) = b(X^{\varepsilon}(t), \alpha^{\varepsilon}(t))dt + \sigma(X^{\varepsilon}(t), \alpha^{\varepsilon}(t))dW(t), \tag{6.2}$$

where $\alpha^{\varepsilon}(t)$ is a fast-varying process whose generator is $Q(x)/\varepsilon$ when $X^{\varepsilon}(t) = x$. Compared with our previous work on two-time-scale Markov processes [12], where time-inhomogeneous Markov chains are treated, the new contribution is featured in the $x$-dependence of the switching process. In this paper, the switching process itself is non-Markov. To overcome the difficulty, we sub-divide the interval into small part, and use careful approximation techniques to resolve the $x$-dependency issue. Recall that $Q(x)$ is bounded and continuous.

($A3$) For each $x \in \mathbb{R}^r$, $Q(x)$ is weakly irreducible. That is, the system of equations

$$\nu(x)Q(x) = 0, \quad \nu(x)\mathbb{1} = 1$$

has a unique solution where $\mathbb{1} := (1, \ldots, 1)' \in \mathbb{R}^{m\times 1}$ is a vector with all component being 1, where $\nu(x) = (\nu_1(x), \ldots, \nu_m(x))$ with $\nu_i(x) \geq 0$ for each $i \in \mathcal{M}$.

As can be seen, $\alpha^{\varepsilon}(\cdot)$ is subject to fast variation, whereas $X^{\varepsilon}(\cdot)$ changes relatively slowly compared with $\alpha^{\varepsilon}(\cdot)$. Although it is subject to rapid variations, the $\alpha^{\varepsilon}(\cdot)$ does not go to $\infty$; it is essentially a noise process having an invariant measure. As $\varepsilon \to 0$, the noise is averaged out, and the slow component of the evolution $X^{\varepsilon}(\cdot)$ converges weakly to $X(\cdot)$ that is an average with respect to the invariant measure $\nu(x)$ given in (A3); see also the idea of averaging in systems with singularly perturbed diffusions in [6] and references therein. Let us consider the process $\{X^{\varepsilon}(\cdot)\}$, and work with $t \in [0, T]$ for some $T > 0$.

**Lemma 6.1.** *Assume both* (A1) *and* (A3). *Then*

$$\sup_{t\in[0,T]} \mathbf{E}|X^{\varepsilon}(t)|^2 < \infty.$$

This is barely a restatement of the moment bounds in the previous sections. Using similar idea as in Lemma 3.4, we can also show that $\mathbf{E}|X^{\varepsilon}(t)|^p < \infty$.

**Lemma 6.2.** *Assume both* (A1) *and* (A3). *Then* $\{X^{\varepsilon}(\cdot)\}$ *is tight in* $D([0, T] : \mathbb{R}^r)$, *the space of functions that are right continuous with left limits endowed with the Skorohod topology.*

*Proof.* For any $\Delta > 0$, and $t, s > 0$ satisfying $s \leq \Delta$, using Lemma 6.1 and the same technique as in Lemma 3.6,

$$\mathbf{E}_t|X^{\varepsilon}(t + s) - X^{\varepsilon}(t)|^2 \leq O(s) \leq O(\Delta),$$

where $\mathbf{E}_t$ denotes the conditioning on the $\sigma$-algebra generated by $\{X^\varepsilon(u), \alpha^\varepsilon(u) : u \le t\}$. Taking $\limsup\limits_{\varepsilon\to 0}$ followed by $\lim\limits_{\Delta\to 0}$, we obtain

$$\lim_{\Delta\to 0}\limsup_{\varepsilon\to 0}\mathbf{E}\mathbf{E}_t|X^\varepsilon(t+s)-X^\varepsilon(t)|^2 = 0.$$

Thus, by virtue of the well-known tightness criterion (for example, see [13, Lemma 14.12, p. 320]), $\{X^\varepsilon(\cdot)\}$ is tight. ■

Since $\{X^\varepsilon(\cdot)\}$ is tight, we can extract weakly convergent subsequences by the well-known Prohorov's theorem. Select such a subsequence and for notational simplicity, still denote the subsequence indexed by $\varepsilon$ with limit $X(\cdot)$. By virtue of the Skorohod representation, there is an augmented probability space on which there is a sequence $\widetilde{X}^\varepsilon(\cdot)$ defined on it having the same distribution as $X^\varepsilon(\cdot)$ such that $\widetilde{X}^\varepsilon(\cdot)$ converges to $\widetilde{X}(\cdot)$ in the sense of w.p.1, where $\widetilde{X}(\cdot)$ have the same distribution as that of $X(\cdot)$. With a slight abuse of notation without changing notation, we still denote this sequence by $\{X^\varepsilon(\cdot)\}$ such that $X^\varepsilon(\cdot)\to X(\cdot)$ w.p.1. Using the argument as in Theorem 3.7, $X(\cdot)$ has continuous sample paths w.p.1. We proceed to characterize the limit process.

To proceed, for an arbitrary $N$ satisfying $0 < N < \infty$, we can confine ourselves with $B_N = \{x : |x| \le N\}$, the ball with radius $N$, and work with a truncated process $X^{\varepsilon,N}(\cdot)$, known as $N$-truncation [13, p.321]. We then obtain the limit of $X^{\varepsilon,N}(\cdot)$. Finally by letting $N\to\infty$ and using a piecing together argument, we prove the convergence of $X^\varepsilon(\cdot)$. However, For notational simplicity and without or loss of generality, we can assume that $X^\varepsilon(\cdot)$ is bounded in what follows. We shall show that the limit $X(\cdot)$ is a solution of the mean field equation

$$dX(t) = \overline{b}(X(t))dt + \overline{\sigma}(X(t))dW(t), \tag{6.3}$$

where

$$\begin{aligned}
\overline{b}(x) &= \sum_{i\in\mathcal{M}}\nu_i(x)b(x,i),\\
\overline{a}(x) &= \sum_{i\in\mathcal{M}}\nu_i(x)a(x,i),\\
\overline{a}(x) &= \overline{\sigma}(x)\overline{\sigma}'(x), \quad a(x,i)=\sigma(x,i)\sigma'(x,i),\\
\nu(x) &= (\nu_1(x),\ldots,\nu_m(x))\in\mathbb{R}^{1\times m}.
\end{aligned}\tag{6.4}$$

Equivalently, $X(\cdot)$ is a solution of the martingale problem with operator $\overline{\mathcal{L}}$ defined by

$$\overline{\mathcal{L}}f(x) = \overline{b}'(x)\nabla f(x) + \mathrm{tr}[\overline{a}(x)Hf(x)], \tag{6.5}$$

for any $f(\cdot) \in C^2(\mathbb{R}^r)$, where $\nabla f(x)$ and $Hf(x)$ are the usual gradient and Hessian matrix of $f(x)$, respectively.

**Theorem 6.3.** *Under the conditions of Lemma* 6.2*, the process* $X^{\varepsilon}(\cdot)$ *converges weakly to* $X(\cdot)$*, which is the solution of the martingale problem with operator* $\overline{\mathcal{L}}$ *given by* (6.5) *or* $X(\cdot)$ *is a solution of the limit mean field equation given by* (6.3).

*Proof.* Since we have already established the tightness of the process $\{X^{\varepsilon}(\cdot)\}$, what remains to be done is to characterize the limit process $X(\cdot)$. To show that $X(\cdot)$ is a solution of the martingale problem with operator $\overline{\mathcal{L}}$, pick out any $F(\cdot) \in C_0^2(\mathbb{R}^r)$ ($C^2$ function with compact support). We need only show that

$$F(X(t)) - F(X(0)) - \int_0^t \overline{\mathcal{L}} F(X(u))du \ \text{ is a martingale.} \tag{6.6}$$

To verify (6.6), it suffices to show that for any bounded and continuous function $h(\cdot)$, any positive integer $\ell$, any $t, s > 0$, and any $t_l \le t$ with $l \le \ell$,

$$\mathbf{E}h(X(t_l) : l \le \ell)\Big[F(X(t+s)) - F(X(t)) - \int_t^{t+s} \overline{\mathcal{L}} F(X(u))du\Big] = 0. \tag{6.7}$$

To verify (6.7), we begin with the process indexed by $\varepsilon$, namely $\{X^{\varepsilon}(\cdot)\}$. Because $F(\cdot)$ is independent of $i \in \mathcal{M}$,

$$\sum_{j\in\mathcal{M}} q_{ij}^{\varepsilon}(x)F(x) = 0 \ \text{ for each } \ i \in \mathcal{M} \ \text{ and each } \ x. \tag{6.8}$$

Since the joint process $(X^{\varepsilon}(\cdot), \alpha^{\varepsilon}(\cdot))$ is Markov,

$$F(X^{\varepsilon}(t+s)) - F(X^{\varepsilon}(t)) - \int_t^{t+s} \mathcal{L}F(X^{\varepsilon}(u))du$$

is a martingale, and as a result,

$$\mathbf{E}h(X^{\varepsilon}(t_l) : l \le \ell)\Big[F(X^{\varepsilon}(t+s)) - F(X^{\varepsilon}(t)) - \int_t^{t+s} \mathcal{L}F(X^{\varepsilon}(u))du\Big] = 0,$$

where $\mathcal{L}$ is the operator defined in (2.7) with $Q(x)$ replaced by $Q(x)/\varepsilon$. That is, in view of (6.8),

$$\mathcal{L}F(x) = \frac{1}{2}\mathrm{tr}[a(x,i)HF(x)] + b'(x,i)\nabla F(x), \ i \in \mathcal{M}.$$

Note that since $F(\cdot)$ is independent of $i \in \mathcal{M}$, the term involving $Q(x)/\varepsilon$ disappears. Note also that $\mathcal{L}$ depends on $\varepsilon$ and should have been written as $\mathcal{L}^{\varepsilon}$, but for notational simplicity, we suppress the $\varepsilon$-dependence. By the weak convergence of $X^{\varepsilon}(\cdot)$ to $X(\cdot)$ and the Skorohod representation, we have

$$\begin{aligned} &\mathbf{E}h(X^{\varepsilon}(t_l) : l \le \ell)\Big[F(X^{\varepsilon}(t+s)) - F(X^{\varepsilon}(t))\Big] \\ &\quad \to \mathbf{E}h(X(t_l) : l \le \ell)\Big[F(X(t+s)) - F(X(t))\Big] \ \text{ as } \ \varepsilon \to 0. \end{aligned} \tag{6.9}$$

On the other hand,

$$
\begin{aligned}
&\mathbf{E}h(X^\varepsilon(t_l): l \le \ell)\Big[\int_t^{t+s} \mathcal{L}F(X^\varepsilon(u))du\Big]\\
&\quad = \mathbf{E}h(X^\varepsilon(t_l): l \le \ell)\Big[\int_t^{t+s} [b'(X^\varepsilon(u),\alpha^\varepsilon(u))\nabla F(X^\varepsilon(u))\\
&\qquad +\frac{1}{2}\mathrm{tr}[a(X^\varepsilon(u),\alpha^\varepsilon(u))HF(X^\varepsilon(u)]du\Big]
\end{aligned}
$$

First, consider the drift term, we obtain

$$
\begin{aligned}
&\mathbf{E}h(X^\varepsilon(t_l): l \le \ell)\Big[\int_t^{t+s} b'(X^\varepsilon(u),\alpha^\varepsilon(u))\nabla F(X^\varepsilon(u))du\Big]\\
&= \mathbf{E}h(X^\varepsilon(t_l): l \le \ell)\Big[\int_t^{t+s} \sum_{i\in\mathcal{M}} b'(X^\varepsilon(u),i)\nabla F(X^\varepsilon(u))I_{\{\alpha^\varepsilon(u)=i\}}\Big]\\
&= \mathbf{E}h(X^\varepsilon(t_l): l \le \ell)\\
&\quad\Big[\int_t^{t+s} \sum_{i\in\mathcal{M}} b'(X^\varepsilon(u),i)\nabla F(X^\varepsilon(u))[I_{\{\alpha^\varepsilon(u)=i\}}-\nu_i(X^\varepsilon(u))]\Big]\\
&\quad +\mathbf{E}h(X^\varepsilon(t_l): l \le \ell)\Big[\int_t^{t+s} \sum_{i\in\mathcal{M}} b'(X^\varepsilon(u),i)\nabla F(X^\varepsilon(u))\nu_i(X^\varepsilon(u))\Big].
\end{aligned}
\tag{6.10}
$$

By virtue of the weak convergence of $X^\varepsilon(\cdot)$ to $X(\cdot)$ and the Skorohod representation (without changing notation by our convention), it can be shown that for the last term in (6.10),

$$
\begin{aligned}
&\mathbf{E}h(X^\varepsilon(t_l): l\le\ell)\Big[\int_t^{t+s} \sum_{i\in\mathcal{M}} b'(X^\varepsilon(u),i)\nabla F(X^\varepsilon(u))\nu_i(X^\varepsilon(u))\Big]\\
&\to \mathbf{E}h(X(t_l): l\le\ell)\Big[\int_t^{t+s} \sum_{i\in\mathcal{M}} b'(X(u),i)\nabla F(X(u))\nu_i(X(u))\Big] \text{ as } \ \varepsilon\to 0\\
&= \mathbf{E}h(X(t_l): l\le\ell)\Big[\int_t^{t+s} \overline{b}'(X(u))\nabla F(X(u))du\Big].
\end{aligned}
\tag{6.11}
$$

As for the next to the last term in (6.10), we partition the interval $[t, t+s]$ as follows. For any $0<\Delta<1$, let $t=t_0<t_1<t_2<\cdots<t_{l_\varepsilon}\le t+s$ such that $t_k = k\varepsilon^{1-\Delta}$. Note that $l_\varepsilon = \lfloor s/\varepsilon^{1-\Delta}\rfloor = O(1/\varepsilon^{1-\Delta})$. Without loss of generality and for notational simplicity, we will assume the $t_{l_\varepsilon}$ coincides with $t+s$. Then we can rewrite

$$
\begin{aligned}
&\int_t^{t+s} \sum_{i\in\mathcal{M}} b'(X^\varepsilon(u),i)\nabla F(X^\varepsilon(u))[I_{\{\alpha^\varepsilon(u)=i\}}-\nu_i(X^\varepsilon(u))]\\
&= \sum_{k=0}^{l_\varepsilon-1}\int_{t_k}^{t_{k+1}} \sum_{i\in\mathcal{M}} b'(X^\varepsilon(u),i)\nabla F(X^\varepsilon(u))[I_{\{\alpha^\varepsilon(u)=i\}}-\nu_i(X^\varepsilon(u))]du.
\end{aligned}
$$

By the continuity of $b(\cdot, i)$ and the smoothness of $F(\cdot)$, we obtain

$$\begin{aligned}
&\lim_{\varepsilon\to 0} \mathbf{E} h(X^\varepsilon(t_l) : l \le \ell) \\
&\quad \Big[\sum_{i\in\mathcal{M}} \int_t^{t+s} b'(X^\varepsilon(u), i)\nabla F(X^\varepsilon(u))[I_{\{\alpha^\varepsilon(u)=i\}} - \nu_i(X^\varepsilon(u))]du\Big] \\
&= \lim_{\varepsilon\to 0} \mathbf{E} h(X^\varepsilon(t_l) : l \le \ell) \\
&\quad \Big[\sum_{k=0}^{l_\varepsilon-1}\sum_{i\in\mathcal{M}} \int_{t_k}^{t_{k+1}} b'(X^\varepsilon(t_k), i)\nabla F(X^\varepsilon(t_k))[I_{\{\alpha^\varepsilon(u)=i\}} - \nu_i(X^\varepsilon(u))]du\Big].
\end{aligned}$$

By the choice of $t_l$, we can rewrite the last line above as

$$\begin{aligned}
&\mathbf{E} h(X^\varepsilon(t_l) : l \le \ell) \\
&\quad \Big[\sum_{k=0}^{l_\varepsilon-1}\sum_{i\in\mathcal{M}} \int_{t_k}^{t_{k+1}} b'(X^\varepsilon(t_k), i)\nabla F(X^\varepsilon(t_k))[I_{\{\alpha^\varepsilon(u)=i\}} - \nu_i(X^\varepsilon(u))]du\Big] \\
&= \mathbf{E} h(X^\varepsilon(t_l) : l \le \ell) \\
&\quad \Big[\sum_{k=0}^{l_\varepsilon-1}\sum_{i\in\mathcal{M}} b'(X^\varepsilon(t_k), i)\nabla F(X^\varepsilon(t_k))\mathbf{E}_{t_k} \int_{t_k}^{t_{k+1}} [I_{\{\alpha^\varepsilon(u)=i\}} - \nu_i(X^\varepsilon(u))]du\Big].
\end{aligned}$$

By virtue of the Cauchy-Schwarz inequality,

$$\begin{aligned}
&\mathbf{E}\Big| h(X^\varepsilon(t_l) : l \le \ell)\Big[\sum_{k=0}^{l_\varepsilon-1}\sum_{i\in\mathcal{M}} \\
&\qquad b'(X^\varepsilon(t_k), i)\nabla F(X^\varepsilon(t_k))\mathbf{E}_{t_k} \int_{t_k}^{t_{k+1}} [I_{\{\alpha^\varepsilon(u)=i\}} - \nu_i(X^\varepsilon(u))]du\Big]\Big| \\
&\le \sum_{k=0}^{l_\varepsilon-1}\sum_{i\in\mathcal{M}} \mathbf{E}\Big| h(X^\varepsilon(t_l) : l \le \ell) \\
&\qquad b'(X^\varepsilon(t_k), i)\nabla F(X^\varepsilon(t_k))\mathbf{E}_{t_k} \int_{t_k}^{t_{k+1}} [I_{\{\alpha^\varepsilon(u)=i\}} - \nu_i(X^\varepsilon(u))]du\Big| \\
&\le K \sum_{k=0}^{l_\varepsilon-1}\sum_{i\in\mathcal{M}} (1 + \mathbf{E}^{1/2}|X^\varepsilon(t_k)|^2)\mathbf{E}^{1/2} \\
&\qquad \Big|\mathbf{E}_{t_k} \int_{t_k}^{t_{k+1}} [I_{\{\alpha^\varepsilon(u)=i\}} - \nu_i(X^\varepsilon(u))]du\Big|^2.
\end{aligned} \tag{6.12}$$

**Lemma 6.4.** *Assume the conditions of Theorem* 6.3 *are fulfilled. For* $t \in [0, T]$, *and each fixed $x$ in a bounded subset of* $\mathbb{R}^r$*, consider the generator* $Q(x)/\varepsilon$*. Then*

$$\left|\exp\Big(\frac{Q(x)t}{\varepsilon}\Big) - \mathbb{1}\nu(x)\right| \le K\exp\Big(-\frac{\kappa_0 t}{\varepsilon}\Big), \tag{6.13}$$

*for some* $K > 0$ *and* $\kappa_0 > 0$*, where* $\mathbb{1} = (1, 1, \ldots, 1)' \in \mathbb{R}^m$*.*

*Proof of Lemma* **6.4**. For each $x$ in a bounded subset of $\mathbb{R}^r$, consider the switching process with a generator $Q(x)/\varepsilon$. Then the results in [12, Lemma A.2, p. 300] are applicable. In fact, $\exp(Q(x)t/\varepsilon)$ is the associated transition matrix. The weak irreducibility of $Q(x)$ implies that $\exp(Q(x)t/\varepsilon)$ converges to a matrix with identical rows, namely, $\mathbb{1}\nu(x)$. Moreover, the convergence takes place exponentially fast. Thus (6.13) holds. Since the set $x$ living in is bounded, $K$ and $\kappa_0$ can be chosen to be independent of $x$. ■

We now examine the last term in (6.12). We have by the continuity of $\nu(x)$,

$$\mathbf{E}_{t_k}\int_{t_k}^{t_{k+1}}[\nu_i(X^\varepsilon(t_k)) - \nu_i(X^\varepsilon(u))]du = o(t_{k+1} - t_k) = o(\varepsilon^{1-\Delta}).$$

Thus

$$\begin{aligned}&\sum_{k=0}^{l_\varepsilon - 1}\sum_{i\in\mathcal{M}}(1 + \mathbf{E}^{1/2}|X^\varepsilon(t_k)|^2)\mathbf{E}^{1/2}\\ &\qquad\left|\mathbf{E}_{t_k}\int_{t_k}^{t_{k+1}}[\nu_i(X^\varepsilon(t_k)) - \nu_i(X^\varepsilon(u))]du\right|^2 \\ &\le K\sum_{k=0}^{l_\varepsilon - 1} o(\varepsilon^{1-\Delta}) \le K l_\varepsilon o(\varepsilon^{1-\Delta}) = o(1) \to 0 \quad\text{as}\quad \varepsilon \to 0.\end{aligned} \tag{6.14}$$

Next, consider

$$\mathbf{E}_{t_k}\int_{t_k}^{t_{k+1}}[I_{\{\alpha^\varepsilon(u)=i\}} - \nu_i(X^\varepsilon(t_k))]du.$$

For $u \in [t_k, t_{k+1}]$, to consider $\mathbf{E}[I_{\{\alpha^\varepsilon(u)=i\}} - \nu_i(X^\varepsilon(t_k))]$, first let us examine the associated transition matrix

$$P^\varepsilon(u, t_k) = (p^\varepsilon_{ij}(u, t_k)) = (\mathbf{P}(\alpha^\varepsilon(u) = j|\alpha^\varepsilon(t_k) = i, X^\varepsilon(t_k)).$$

It satisfies the forward equation

$$\begin{aligned}\frac{d}{du}P^\varepsilon(u, t_k) &= P^\varepsilon(u, t_k)\frac{Q(X^\varepsilon(u))}{\varepsilon}\\ &= P^\varepsilon(u, t_k)\frac{Q(X^\varepsilon(t_k))}{\varepsilon} + P^\varepsilon(u, t_k)\frac{Q(X^\varepsilon(u)) - Q(X^\varepsilon(t_k))}{\varepsilon},\\ P^\varepsilon(t_k, t_k) &= I.\end{aligned}$$

The solution of the above matrix differential equation is given by

$$
\begin{aligned}
P^\varepsilon(u,t_k) &= \exp\Big(\frac{Q(X^\varepsilon(t_k))(u-t_k)}{\varepsilon}\Big) \\
&\quad + \int_{t_k}^u P^\varepsilon(s,t_k)\frac{Q(X^\varepsilon(s))-Q(X^\varepsilon(t_k))}{\varepsilon} \\
&\quad \exp\Big(\frac{Q(X^\varepsilon(t_k))(s-t_k)}{\varepsilon}\Big)ds.
\end{aligned} \tag{6.15}
$$

Note that $Q(X^\varepsilon(s))1\!\!1 = Q(X^\varepsilon(t_k))1\!\!1 = 0$ and $\nu(x)Q(x)=0$ for each $x$. It then follows from (6.15) that

$$
\begin{aligned}
&P^\varepsilon(u,t_k) - 1\!\!1\nu(X^\varepsilon(u)) \\
&= \Big(\exp\Big(\frac{Q(X^\varepsilon(t_k))(u-t_k)}{\varepsilon}\Big) - 1\!\!1\nu(X^\varepsilon(t_k))\Big) \\
&\quad + \int_{t_k}^u \Big\{[P^\varepsilon(s,t_k) - 1\!\!1\nu(X^\varepsilon(t_k))]\frac{Q(X^\varepsilon(s))-Q(X^\varepsilon(t_k))}{\varepsilon} \\
&\quad + 1\!\!1[\nu(X^\varepsilon(t_k)) - \nu(X^\varepsilon(s))]\frac{Q(X^\varepsilon(s))}{\varepsilon}\Big\} \\
&\quad \times\Big(\exp\Big(\frac{Q(X^\varepsilon(t_k))(s-t_k)}{\varepsilon}\Big) - 1\!\!1\nu(X^\varepsilon(t_k))\Big)ds.
\end{aligned} \tag{6.16}
$$

Note that using Lemma 6.4, for some for some $0<\kappa_1<\kappa_0$,

$$
\begin{aligned}
&\Big| \int_{t_k}^u 1\!\!1[\nu(X^\varepsilon(t_k)) - \nu(X^\varepsilon(s))]\frac{Q(X^\varepsilon(s))}{\varepsilon} \\
&\Big(\exp\Big(\frac{Q(X^\varepsilon(t_k))(s-t_k)}{\varepsilon}\Big) - 1\!\!1\nu(X^\varepsilon(t_k))\Big)ds\Big| \\
&\le K\int_{t_k}^u \frac{1}{\varepsilon}g_1^\varepsilon(s)\exp\Big(-\frac{\kappa_0(s-t_k)}{\varepsilon}\Big)ds \\
&\le K\int_{t_k}^u g_1^\varepsilon(s)\exp\Big(-\frac{\kappa_1(s-t_k)}{\varepsilon}\Big)ds \\
&\le o(\varepsilon).
\end{aligned}
$$

In the above $g_1^\varepsilon(s)$ is a continuous function satisfying $g_1^\varepsilon(s)\to 0$ as $\varepsilon\to 0$.

Thus, using Lemma 6.4, for some $\kappa_0>0$ and $0<\kappa_2<\kappa_0$,

$$
\begin{aligned}
&|P^\varepsilon(u,t_k) - 1\!\!1\nu(X^\varepsilon(u))| \\
&\le \exp\Big(-\frac{\kappa_0(u-t_k)}{\varepsilon}\Big) + o(\varepsilon)
\end{aligned}
$$

$$
\begin{aligned}
&+K\int_{t_k}^{u}|P^\varepsilon(s,t_k)-\mathbb{1}\nu(X^\varepsilon(t_k))|g^\varepsilon(s)\frac{1}{\varepsilon}\exp\Big(\frac{-\kappa_0(s-t_k)}{\varepsilon}\Big)ds\\
&\le \exp\Big(-\frac{\kappa_0(u-t_k)}{\varepsilon}\Big)+o(\varepsilon)\\
&\quad +K\int_{t_k}^{u}|P^\varepsilon(s,t_k)-\mathbb{1}\nu(X^\varepsilon(t_k))|g^\varepsilon(s)\exp\Big(\frac{-\kappa_2(s-t_k)}{\varepsilon}\Big)ds,
\end{aligned}
$$

where $g^\varepsilon(s)\to 0$ as $\varepsilon\to 0$. An application of Gronwall's inequality [3, p. 36, Lemma 6.2] then yields that

$$
|P^\varepsilon(u,t_k)-\mathbb{1}\nu(X^\varepsilon(u))|\le o(\varepsilon). \tag{6.17}
$$

It then follows that

$$
\begin{aligned}
&\left|\mathbf{E}_{t_k}\int_{t_k}^{t_{k+1}}[I_{\{\alpha^\varepsilon(u)=i\}}-\nu_i(X^\varepsilon(u))]du\right|\\
&\quad\le K\int_{t_k}^{t_{k+1}}o(\varepsilon)du\\
&\quad=O(\varepsilon^{2-\Delta}).
\end{aligned} \tag{6.18}
$$

Using (6.18) in the last term in (6.12), we obtain

$$
\begin{aligned}
&\sum_{k=0}^{l_\varepsilon-1}\sum_{i\in\mathcal{M}}(1+\mathbf{E}^{1/2}|X^\varepsilon(t_k)|^2)\mathbf{E}^{1/2}\\
&\left|\mathbf{E}_{t_k}\int_{t_k}^{t_{k+1}}[I_{\{\alpha^\varepsilon(u)=i\}}-\nu_i(X^\varepsilon(u))]du\right|^2\\
&\quad\le K\sum_{k=0}^{l_\varepsilon-1}\sum_{i\in\mathcal{M}}(1+\mathbf{E}^{1/2}|X^\varepsilon(t_k)|^2)\mathbf{E}^{1/2}\left|\int_{t_k}^{t_{k+1}}o(\varepsilon)du\right|^2\\
&\quad\le Kl_\varepsilon O(\varepsilon^{2-\Delta})=O(\varepsilon^{1-\Delta})\to 0\ \text{ as }\ \varepsilon\to 0.
\end{aligned} \tag{6.19}
$$

By virtue of the weak convergence of $X^\varepsilon(\cdot)$ to $X(\cdot)$, the Skorohod representation, (6.10), (6.11), and the estimates up to now, we obtain

$$
\begin{aligned}
&\mathbf{E}h(X^\varepsilon(t_l):l\le\ell)\Big[\int_t^{t+s}b'(X^\varepsilon(u),\alpha^\varepsilon(u))\nabla F(X^\varepsilon(u))du\Big]\\
&\quad=\mathbf{E}h(X(t_l):l\le\ell)\Big[\int_t^{t+s}\overline{b}'(X(u))\nabla F(X(u))du\Big].
\end{aligned} \tag{6.20}
$$

Likewise, we can show

$$
\begin{aligned}
&\mathbf{E}h(X^\varepsilon(t_l):l\le\ell)\Big[\int_t^{t+s}\mathrm{tr}[a(X^\varepsilon(u),\alpha^\varepsilon(u))HF(X^\varepsilon(u))]du\Big]\\
&\quad=\mathbf{E}h(X(t_l):l\le\ell)\Big[\int_t^{t+s}\mathrm{tr}[a(X(u))HF(X(u))]du\Big].
\end{aligned} \tag{6.21}
$$

Combining the estimates obtained thus far, we obtain that $X(\cdot)$ is a solution of the martingale problem with operator $\overline{\mathcal{L}}$. The theorem is thus proved. ■

## 7. Concluding Remarks

In this paper, we have established several properties for a class of mean-field models with random switching. The random switching is continuous-state dependent. One difficulties considered here is that each of the particle is required to be nonnegative. Our results include moment estimates, regularity, continuity, and certain tightness. Furthermore, we also examine the asymptotic behavior when the switching process is subject to fast variation. In the future study, it will be interesting to examine the equivalent or mean field behavior when the number of particles or bodies becomes large. In addition, the study of behavior of phase transitions will also be a worthwhile undertaking.

## References

1. D. A. Dawson, Critical dynamics and fluctuations for a mean-field model of cooperative behavior, *J. Statist. Phys.*, **31** (1983), 29-85.
2. D. A. Dawson and X. Zheng, Law of large numbers and central limit theorem for unbounded jump mean-field models, *Adv. Appl. Math.*, **12** (1991), 293-326.
3. J. K. Hale, *Ordinary Differential Equations*, 2nd ed., R.E. Krieger, Malabar, FL, 1980.
4. M. Hitsuda and I. Mitoma, Tightness problem and stochastic evolution equation arising from fluctuation phenomena for interacting diffusions, *J. Multivariate Anal.*, **19** (1986), 311-328.
5. R. Z. Khasminskii, *Stochastic Stability of Differential Equations*, Stijhoff and Noordhoff, Alphen, 1980.
6. R. Khasminskii and G. Yin, On averaging principles: An asymptotic expansion approach, *SIAM J. Math. Anal.*, **35** (2004), 1534-1560.
7. R. S. Liptser and A. N. Shiryayev, *Statistics of Random Processes I*, Springer-Verlag, New York, NY, 2001.
8. X. Mao, *Stochastic Differential Equations and Applications*, 2nd Ed., Horwood, Chichester, UK, 2008.
9. T. Shiga and H. Tanaka, Central limit theorem for a system of Markovian particles with mean field interactions, *Z. Wahr. Verw. Geb.*, **69** (1985), 439-459.
10. A. V. Skorohod, *Studies in the Theory of Random Processes*, Dover, New York, 1982.
11. F. Xi and G. Yin, Asymptotic properties of a mean-field model with a continuous-state-dependent switching process, *J. Appl. Probab.*, **46** (2009), 221-243.

12. G. Yin and Q. Zhang, *Continuous-time Markov Chains and Applications: A Singular Perturbations Approach*, Springer-Verlag, New York, NY, 1998.

13. G. Yin and Q. Zhang, *Discrete-time Markov Chains: Two-time-scale Methods and Applications*, Springer, New York, NY, 2005.

14. G. Yin and C. Zhu, *Hybrid Switching Diffusions: Properties and Applications*, Springer, New York, 2010.

15. C. Zhu and G. Yin, Asymtotic properties of hybrid diffusion systems, *SIAM J. Control Optim.*, **46** (2007), 1155-1179.

16. C. Zhu and G. Yin, On competitive Lotka-Volterra model in random environments, *J. Math. Anal. Appl.*, **357** (2009), 154-170.

G. Yin
Department of Mathematics
Wayne State University
Detroit, MI 48202
U.S.A.
E-mail: gyin@math.wayne.edu

Guangliang Zhao
Department of Mathematics
Wayne State University
Detroit, MI 48202
U.S.A.
E-mail: ea4628@wayne.edu

Fubao Xi
Department of Mathematics
Beijing Institute of Technology
Beijing 100081
P. R. China
E-mail: xifb@bit.edu.cn