

# ASYMPTOTIC PROPERTIES OF A MEAN-FIELD MODEL WITH A CONTINUOUS-STATE-DEPENDENT SWITCHING PROCESS

FUBAO XI,\* *Beijing Institute of Technology*

G. YIN,\*\* *Wayne State University*

## Abstract

This work is concerned with a class of mean-field models given by a switching diffusion with a continuous-state-dependent switching process. Focusing on asymptotic properties, the regularity or nonexplosiveness, Feller continuity, and strong Feller continuity are established by means of introducing certain auxiliary processes and by making use of the truncations. Based on these results, exponential ergodicity is obtained under the Foster–Lyapunov drift conditions. By virtue of the coupling methods, the strong ergodicity or uniform ergodicity in the sense of convergence in the variation norm is established for the mean-field model with a Markovian switching process. Besides this, several examples are presented for demonstration and illustration.

*Keywords:* Mean-field model; continuous-state-dependent switching; Markovian switching; Feller continuity; strong Feller continuity; exponential ergodicity; strong ergodicity

2000 Mathematics Subject Classification: Primary 60J60

Secondary 60J27

## 1. Introduction

Originating from statistical mechanics, mean-field models are concerned with many-body systems with interactions. To overcome the difficulty of interactions due to the many bodies, one of the main ideas is to replace all interactions to any one body with an average or effective interaction. This reduces any multi-body problem to an effective one-body problem. Although their main motivation and development are in statistical mechanics, such models have also enjoyed recent applications in, for example, graphical models in artificial intelligence. It should be mentioned that in financial engineering, a somewhat related formulation is termed a mean-reversion model, in which the formulation uses a dynamic model that has a force pushing the system towards its ‘equilibrium’ in a suitable way.

Intuitively, if the field or particle exhibits many interactions in the original system, the mean field will be more accurate for such a system. The usefulness of such systems, the potential impact on many practical situations, and the challenges they present to both physicists and mathematicians have attracted much attention in recent years. Dawson [6] presented a detailed study on the cooperative behavior of such systems, and, subsequently, in [7] delineated the law

---

Received 24 November 2008; revision received 6 February 2009.

\* Postal address: Department of Mathematics, Beijing Institute of Technology, Beijing, 100081, P. R. China.

Email address: xifb@bit.edu.cn

Supported in part by the National Natural Science Foundation of China under grant number 10671037.

\*\* Postal address: Department of Mathematics, Wayne State University, Detroit, MI 48202, USA.

Email address: gyin@math.wayne.edu

Supported in part by the National Science Foundation under grant DMS-0603287 and in part by the National Security Agency under grant MSPF-068-029.

of large numbers and central limit theorem for jump mean fields; see also the related works of Hitsuda and Mitoma [9] and Shiga and Tanaka [23], and the references therein. From a probabilistic point of view, one of the many important problems is that of gaining an in depth understanding of the ergodicity of such systems.

Owing to the rapid progress in technology, much more complicated systems are encountered in applications. In response to such challenges, much effort has been devoted to the modeling and analysis of more 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 not only be influenced by or subject to the usual dynamics represented by differential equations, but also contain movements that exhibit jump or switching behavior. These discrete events are used to depict random environments or other uncertainties. One of the methods of formulating complex systems is to use a regime-switching formulation, which largely enriched the applicability of the dynamic models. Continuing our efforts in the study of regime-switching diffusions, this work is devoted to the study of ergodicity of regime-switching mean-field models.

For motivational purposes, let us begin with two examples. When the switching component  $Z(t)$ , to be defined precisely in Section 2, is missing, these examples are used to study the quenching problem in chemical systems; see Section 13 of [10].

**Example 1.1.** For simplicity, let us start with a one-body dynamic system given by

$$dX(t) = [\alpha(Z(t))X(t) - X^3(t)] dt + \sigma(X(t), Z(t)) dW(t), \quad (1.1)$$

where  $\alpha(\cdot)$  and  $\sigma(\cdot)$  are appropriate functions to be defined precisely in Section 2,  $W(t)$  is a standard one-dimensional Wiener process, and  $Z(t)$  is a switching process depending on the dynamics of  $X(t)$ . When  $Z(t)$  is missing, the model defined by (1.1) is called the Schlögl model with Gaussian white noise perturbation and it is an appropriate model for studying the quenching problem (see Section 13.1 of [10]). This equation is also used to describe the movement of a single anharmonic oscillator subject to Gaussian white noise perturbation (see [6]).

**Example 1.2.** Consider a two-body problem with the dynamics given by

$$\begin{aligned} dX_1(t) &= [\alpha(Z(t))X_1(t) - X_1^3(t) - \beta(Z(t))(X_1(t) - X_2(t))] dt \\ &\quad + \sigma_{11}(X(t), Z(t)) dW_1(t), \\ dX_2(t) &= [\alpha(Z(t))X_2(t) - X_2^3(t) - \beta(Z(t))(X_2(t) - X_1(t))] dt \\ &\quad + \sigma_{22}(X(t), Z(t)) dW_2(t), \end{aligned} \quad (1.2)$$

where  $\beta(\cdot)$ ,  $\sigma_{11}(\cdot)$ , and  $\sigma_{22}(\cdot)$  are appropriate functions to be defined precisely in Section 2, and  $W(t)$  is a standard two-dimensional Wiener process. In statistical physics, the model defined by (1.2) without  $Z(t)$  is known as the double-vessel diffusion Schlögl model with Gaussian white noise perturbation and was used to study the quenching problem for multidimensional or nonuniform systems in Section 13.3 of [10]. Moreover, adopting this modeling method also enables us to avoid using stochastic partial differential equations, which are usually very complex; see Section 13.3 of [10].

In this work we consider a class of mean-field models with randomly varying switching. By concentrating on the two-component Markov process  $(X(t), Z(t))$ , we obtain the Feller continuity and the strong Feller continuity by virtue of the truncation methods. Moreover, we investigate the ergodic properties of diffusion processes with continuous-state-dependent

switching. We prove the  $\Psi$ -exponential ergodicity for  $(X(t), Z(t))$  under the Foster–Lyapunov drift condition. When the switching process  $Z(t)$  is independent of the component  $X(t)$  (i.e. the case of Markovian switching), we establish the strong ergodicity or uniform ergodicity for  $(X(t), Z(t))$  by the coupling methods.

The rest of the paper is arranged as follows. The precise formulation of the problem is presented in Section 2. Section 3 is devoted to the regularity and Feller continuity. Section 4 proceeds with the strong Feller property, and in Section 5 we exploit exponential ergodicity. In Section 6 we establish strong ergodicity.

### 2. Formulation

For a vector or a matrix  $A$ , we use  $A^\top$  to denote its transpose. Let  $K$  be a positive integer, let  $(\Omega, \mathcal{F}, \mathbb{P})$  be a probability space, let  $\{\mathcal{F}_t\}$  be an increasing family of sub- $\sigma$ -algebras of  $\mathcal{F}$ , and let  $W(t) = (W_1(t), W_2(t), \dots, W_K(t))^\top$  be an  $\mathcal{F}_t$ -adapted, standard,  $K$ -dimensional Wiener process. Suppose that  $Z(t)$  is a right-continuous random process taking values in a finite state space  $N = \{1, 2, \dots, n_0\}$ . Consider a  $K$ -body mean-field model with switching described by the following system of Itô stochastic differential equations. For  $i = 1, 2, \dots, K$ ,

$$dX_i(t) = [\alpha(Z(t))X_i(t) - X_i^3(t) - \beta(Z(t))(X_i(t) - \bar{X}(t))]dt + \sigma_{ii}(X(t), Z(t))dW_i(t), \tag{2.1}$$

where  $\bar{X}(t) = K^{-1} \sum_{j=1}^K X_j(t)$ ,  $X(t) = (X_1(t), X_2(t), \dots, X_K(t))^\top$ , and  $\alpha(k) > 0$  and  $\beta(k) > 0$  for  $k \in N$ . Moreover, the transition rules of  $Z(t)$  are specified by

$$\mathbb{P}(Z(t + \Delta) = l \mid Z(t) = k, X(t) = x) = \begin{cases} q_{kl}(x)\Delta + o(\Delta) & \text{if } k \neq l, \\ 1 + q_{kk}(x)\Delta + o(\Delta) & \text{if } k = l, \end{cases} \tag{2.2}$$

which hold uniformly in  $\mathbb{R}^K$  as  $\Delta \downarrow 0$ . For  $x = (x_1, x_2, \dots, x_K)^\top \in \mathbb{R}^K$  and  $\sigma = (\sigma_{ij}) \in \mathbb{R}^K \times \mathbb{R}^K$ , we use the usual Euclidean length and trace norm:

$$|x| = \sqrt{\sum_{i=1}^K x_i^2}, \quad |\sigma| = \sqrt{\text{tr}(\sigma\sigma^\top)} = \sqrt{\sum_{i,j=1}^K \sigma_{ij}^2}.$$

Let  $N$  be endowed with the discrete topology, and let  $\mathcal{B}(\mathbb{R}^K \times N)$  be the product  $\sigma$ -algebra on  $\mathbb{R}^K \times N$ . We introduce a metric  $\lambda(\cdot, \cdot)$  on  $\mathbb{R}^K \times N$  as follows:

$$\lambda((x, m), (y, n)) = \rho(x, y) + d(m, n),$$

where

$$\rho(x, y) = |x - y|, \quad d(m, n) = \begin{cases} 0, & m = n, \\ 1, & m \neq n. \end{cases}$$

Therefore,  $(\mathbb{R}^K \times N, \lambda(\cdot, \cdot), \mathcal{B}(\mathbb{R}^K \times N))$  is a locally compact and separable metric space. For the existence and uniqueness of the solution  $(X(t), Z(t))$  satisfying system (2.1) and (2.2), we make the following assumption.

**Assumption 2.1.** For all  $1 \leq i \leq K$  and  $k \neq l \in N$ , the functions  $\sigma_{ii}(x, k) > 0$  in (2.1) and  $q_{kl}(x) > 0$  in (2.2) satisfy the local Lipschitz condition with respect to  $x$ . Moreover, for

all  $1 \leq i \leq K$  and  $k \in N$ , each  $\sigma_{ii}(x, k) > 0$  is infinitely differentiable in  $x$  and there exist constants  $H > 0$  and  $\delta \in (0, 4]$  such that

$$\sum_{i=1}^K \sigma_{ii}^2(x, k) \leq H(|x|^{4-\delta} + 1). \quad (2.3)$$

Analogously to the proof given in Section 6 of [28], under Assumption 2.1, adopting the method of proof of Theorem 4.1 in Chapter III of [13], or Theorem 2.1 of [22] together with the use of Lemma 3 in Chapter II of [24], we can also prove that the stochastic differential equation (2.1) has a unique solution  $X(t)$  on  $\mathbb{R}^K$  even though the drift coefficients in (2.1) satisfy neither the Lipschitz condition nor the linear growth condition (see Section 3). Throughout the rest of this paper, as a standing hypothesis, we assume that Assumption 2.1 holds. Hence, system (2.1) and (2.2) has a unique solution  $(X(t), Z(t))$ . Moreover, the process  $(X(t), Z(t))$  is a Markov process with state space  $\mathbb{R}^K \times N$ . In what follows we often denote the process  $(X(t), Z(t))$  determined by system (2.1) and (2.2) with the initial condition  $(X(0), Z(0)) = (x, k)$  by  $(X^{x,k}(t), Z^{x,k}(t))$  if the emphasis on the initial data is needed.

The term  $-\beta(k)(X_i(t) - \bar{X}(t))$  on the right-hand side of (2.1) (when  $Z(t) = k \in N$ ) creates a tendency for each component  $X_i(t)$  to move toward the arithmetic mean of the ensemble  $X(t) = (X_1(t), X_2(t), \dots, X_K(t))^T$ . Thus, the model defined by (2.1) has a mean-field-like interaction in which each component  $X_i(t)$  interacts with every other component  $X_j(t)$  ( $j \neq i$ ) and then provides a simple model of a cooperative interaction. Here, the discrete component  $Z(t)$  is the switching process and it represents the cooperative interaction of  $X_i(t)$  ( $1 \leq i \leq K$ ) evolving in a random environment which is otherwise not representable by the traditional differential equations. Note also that the switching process depends on the continuous component  $X(t)$ . Therefore, our model defined by  $(X(t), Z(t))$  using (2.1) can be called a mean-field model with continuous-state-dependent switching. The Markov process  $(X(t), Z(t))$  is a continuous-state-dependent, regime-switching (or hybrid) diffusion. When the functions  $q_{kl}(x)$  in (2.2) are independent of  $x$  and the switching process  $Z(t)$  is a Markov chain, itself being independent of  $W(t)$ , the above continuous-state-dependent switching diffusion reduces to the Markovian switching diffusion. As explained in [27], [28], and [32], the notion of continuous-state-dependent switching is a nontrivial generalization of the notion of Markovian switching.

When the switching component  $Z(t)$  is missing, the model defined by (2.1) reduces to a classical mean-field model. For more detailed interpretation about this type of mean-field model without continuous-state-dependent switching, an interested reader is referred to [6] and the references therein. In fact, the mean-field method has many applications and is often considered to be an approximation to describe the complicated original models in statistical physics. Moreover, various mean-field models have been well investigated by many authors; see [6], [7], [9], [23], and the references therein. However, these investigations mainly dealt with limit theorems such as laws of large numbers and central limit theorems. In this paper we focus on ergodic properties of the mean-field model with continuous-state-dependent switching defined by  $(X(t), Z(t))$ .

In practice, stochastic systems with Markovian switching are often called Markovian jump systems and are found to be useful in many applications; see [29, Chapters 3, 8–10], [30], and the references therein. They have successfully been used to model flexible manufacturing systems, production planning, power systems, and economic systems, among others. Meanwhile, such systems could also be regarded as a class of stochastic systems in the Markovian random

environment. Recently, stability of stochastic systems with Markovian switching has received a great deal of attention. Stochastic stability, moment stability, robust stability, exponential stability, stability in distribution, and stability in the total variation norm have been studied extensively by many authors; see [1], [14], [16], [25], [26], [29, Chapters 3, 8–10], [30], [31], and [32], among others. For example, stability in distribution of switching diffusion processes with Markovian switching was studied in [1] and [31], and stability in the total variation was discussed in [25] and [26]. Since convergence in total variation leads to weak convergence, stability in total variation implies that of stability in distribution. To obtain the results of convergence in total variation, the coupling methods in [25, Section 3] and [26, Section 3] were used. As witnessed in many fields, coupling methods are effective and powerful. For background on coupling methods and their applications, the reader is referred to [2, Chapter 2], [3, Chapter 5], [4], and the references therein.

### 3. Nonexplosiveness and Feller continuity

In this section we prove the nonexplosiveness and Feller continuity of  $(X(t), Z(t))$ . To do so, let us make some preparations first. For  $(x, k) \in \mathbb{R}^K \times N$ , set

$$b(x, k) = \begin{pmatrix} b_1(x, k) \\ b_2(x, k) \\ \vdots \\ b_K(x, k) \end{pmatrix} = \begin{pmatrix} \alpha(k)x_1 - x_1^3 - \beta(k)(x_1 - \bar{x}) \\ \alpha(k)x_2 - x_2^3 - \beta(k)(x_2 - \bar{x}) \\ \vdots \\ \alpha(k)x_K - x_K^3 - \beta(k)(x_K - \bar{x}) \end{pmatrix} \in \mathbb{R}^K \tag{3.1}$$

and  $\sigma(x, k) = \text{diag}\{\sigma_{ii}(x, k)\} \in \mathbb{R}^K \times \mathbb{R}^K$ , where  $\bar{x} = K^{-1} \sum_{j=1}^K x_j$ . Then the stochastic differential equation (2.1) can be rewritten as

$$dX(t) = b(X(t), Z(t)) dt + \sigma(X(t), Z(t)) dW(t). \tag{3.2}$$

Let  $C^2(\mathbb{R}^K; \mathbb{R}_+)$  denote the family of all nonnegative functions on  $\mathbb{R}^K$  that are twice continuously differentiable, where  $\mathbb{R}_+ = [0, \infty)$ . As in [27] and [28], for each  $k \in N$  and any function  $V(\cdot, k) \in C^2(\mathbb{R}^K; \mathbb{R}_+)$ , we define an operator  $L$  as follows:

$$LV(x, k) = \frac{1}{2} \sum_{i=1}^K \sigma_{ii}^2(x, k) \frac{\partial^2}{\partial x_i^2} V(x, k) + \sum_{i=1}^K b_i(x, k) \frac{\partial}{\partial x_i} V(x, k) + \sum_{l \in N} q_{kl}(x)(V(x, l) - V(x, k)). \tag{3.3}$$

Next, for convenience, we recall the definitions of the Feller continuity and the strong Feller continuity. As in [28], denote the family of transition probabilities of the Markov process  $(X(t), Z(t))$  by  $\{P(t, (x, k), A) : t \geq 0, (x, k) \in \mathbb{R}^K \times N, A \in \mathcal{B}(\mathbb{R}^K \times N)\}$ . Moreover, for any given  $t > 0$ , set

$$P(t) := \{P(t, (x, k), A) : (x, k) \in \mathbb{R}^K \times N, A \in \mathcal{B}(\mathbb{R}^K \times N)\}, \tag{3.4}$$

which is the transition probability kernel of the Markov process  $(X(t), Z(t))$ . For any given  $t > 0$ , if, for any bounded and continuous function  $f(x, k)$  on  $\mathbb{R}^K \times N$ ,

$$P_t f(x, k) := \sum_{l \in N} \int_{\mathbb{R}^K} f(y, l) P(t, (x, k), dy \times \{l\}) = E f(X^{x,k}(t), Z^{x,k}(t))$$

is still bounded and continuous with respect to  $(x, k)$ , then we say that the transition probability kernel  $P(t)$  is Feller continuous; if, for any bounded measurable function  $f(x, k)$  on  $\mathbb{R}^K \times N$ ,  $P_t f(x, k)$  is bounded and continuous with respect to  $(x, k)$ , then we say that the transition probability kernel  $P(t)$  is strong Feller continuous. Moreover, if the Markov process  $(X(t), Z(t))$  has a Feller or strong Feller continuous transition probability kernel, we sometimes say that  $(X(t), Z(t))$  has the Feller or strong Feller property, respectively.

**Remark 3.1.** It follows from Proposition 6.1.1 of [20] that, for any  $t > 0$ , the transition probability kernel  $P(t)$  is Feller continuous if and only if  $P(t, (x, k), O)$  is lower semicontinuous with respect to  $(x, k)$  for any open set  $O \in \mathcal{B}(\mathbb{R}^K \times N)$ ; the transition probability kernel  $P(t)$  is strong Feller continuous if and only if  $P(t, (x, k), A)$  is lower semicontinuous with respect to  $(x, k)$  for any set  $A \in \mathcal{B}(\mathbb{R}^K \times N)$ .

**Remark 3.2.** The  $b(\cdot, k)$  given in (3.1) and Assumption 2.1 imply that  $b(\cdot, k)$ ,  $\sigma(\cdot, k)$ , and  $q_{kl}(\cdot)$  are locally bounded and locally Lipschitz. That is, for any  $M > 0$ , there exists a constant  $H_M > 0$  such that

$$|b(x, k)| \leq H_M, \quad |\sigma(x, k)| \leq H_M, \quad q_{kl}(x) \leq H_M, \\ |b(x, k) - b(y, k)| + |\sigma(x, k) - \sigma(y, k)| \leq H_M|x - y|,$$

and

$$|q_{kl}(x) - q_{kl}(y)| \leq H_M|x - y|,$$

for all  $x, y \in U(M)$  and  $k \neq l \in N$ , where  $U(M) := \{x \in \mathbb{R}^K : |x| \leq M\}$ . Moreover, setting  $V(x, k) = |x|^2 + 1$ , under Assumption 2.1, it is easy to verify that

$$LV(x, k) \leq -V(x, k) + \beta \quad \text{for } (x, k) \in \mathbb{R}^K \times N, \tag{3.5}$$

where  $\beta > 0$  is a constant. In fact, for all  $(x, k) \in \mathbb{R}^K \times N$  and  $1 \leq i \leq K$ , it is obvious that  $\partial V(x, k)/\partial x_i = 2x_i$  and  $\partial^2 V(x, k)/\partial x_i^2 = 2$ . Substituting these into (3.3), we have

$$LV(x, k) = \sum_{i=1}^K \sigma_{ii}^2(x, k) + 2(\alpha(k) - \beta(k)) \sum_{i=1}^K x_i^2 - 2 \sum_{i=1}^K x_i^4 + 2\beta(k)\bar{x} \sum_{i=1}^K x_i. \tag{3.6}$$

Note that

$$|x|^4 \leq K \sum_{i=1}^K x_i^4 \quad \text{and} \quad \bar{x} \sum_{i=1}^K x_i \leq |x|^2.$$

Hence, it follows from (3.6) that

$$LV(x, k) \leq \sum_{i=1}^K \sigma_{ii}^2(x, k) + 2\alpha(k)|x|^2 - \frac{2}{K}|x|^4 \\ = -V(x, k) \left[ \frac{2|x|^4}{K(|x|^2 + 1)} - \frac{2\alpha(k)|x|^2}{|x|^2 + 1} - \frac{\sum_{i=1}^K \sigma_{ii}^2(x, k)}{|x|^2 + 1} \right] \\ = -V(x, k)\tilde{F}(x, k), \tag{3.7}$$

where  $\tilde{F}(x, k)$  denotes the function in the square brackets. Meanwhile, it follows from (2.3) that there exists a compact set  $C_0 \subset \mathbb{R}^K$  such that

$$\tilde{F}(x, k) \mathbf{1}_{C_0^c \times N}(x, k) \geq \mathbf{1}_{C_0^c \times N}(x, k),$$

where  $\mathbf{1}_{C_0^c \times N}$  is the indicator function of the set  $C_0^c \times N$ . Combining this fact with (3.7), we obtain (3.5).

In what follows we will prove the nonexplosiveness and Feller continuity for  $(X(t), Z(t))$  determined by system (2.2) and (3.2) by making use of a truncation argument. We denote by  $\mathbb{P}^{(x,k)}$  the distribution of  $(X(t), Z(t))$  starting from  $(x, k)$ , and we denote by  $\mathbb{E}^{(x,k)}$  the corresponding expectation.

**Theorem 3.1.** *Under Assumption 2.1, the system given by (2.2) and (3.2) has a unique nonexplosive solution  $(X(t), Z(t))$ , and it is a strong Markov process and has the Feller property.*

*Proof.* Noting Remark 3.2, under Assumption 2.1, we can construct sequences of functions  $\{b^{(n)}(x, k)\}$ ,  $\{\sigma^{(n)}(x, k)\}$ , and  $\{q_{kl}^{(n)}(x)\}$  such that, for all  $x \in O_n := \{x \in \mathbb{R}^K : |x| < n\}$  and  $k \neq l \in N$ ,

$$b^{(n)}(x, k) = b(x, k), \quad \sigma^{(n)}(x, k) = \sigma(x, k), \quad q_{kl}^{(n)}(x) = q_{kl}(x),$$

and, for each positive integer  $n$ , the newly constructed functions are bounded and continuously differentiable in  $x$ , and satisfy the Lipschitz condition globally. Then we have a unique strong Markov process  $(X^{(n)}(t), Z^{(n)}(t))$  corresponding to the functions constructed above for system (2.2) and (3.2). It is clear that  $(X(t), Z(t)) = (X^{(n)}(t), Z^{(n)}(t))$  whenever  $t < \gamma_n$ , where  $\gamma_n := \inf\{t \geq 0 : (X(t), Z(t)) \in O_n^c \times N\}$ .

Denote by  $\gamma$  the exit time (i.e. explosion time) of the Markov process  $(X(t), Z(t))$ , defined as  $\gamma = \lim_{n \rightarrow \infty} \gamma_n$ . To obtain the desired nonexplosiveness, we need only prove that  $\mathbb{P}^{(x,k)}(\gamma = \infty) = 1$ . To this end, take  $V(x, k) = |x|^2 + 1$ , as in Remark 3.2. It is easy to see from (3.5) that, for some positive constant  $\alpha_0$ ,

$$LV(x, k) \leq \alpha_0 V(x, k) \quad \text{for } (x, k) \in \mathbb{R}^K \times N, \tag{3.8}$$

$$\inf\{V(x, k) : |x| \geq M, k \in N\} \rightarrow \infty \quad \text{as } M \rightarrow \infty. \tag{3.9}$$

Next, we define a function

$$W(x, t, k) = V(x, k) \exp(-\alpha_0 t).$$

From (3.3) and (3.8),  $[(\partial/\partial t) + L]W(x, t, k) \leq 0$ . Therefore, by virtue of Dynkin’s formula (cf. Lemma 3 in Chapter II of [24]), for  $\gamma_n(t) := \min\{\gamma_n, t\}$ , we obtain

$$\begin{aligned} & \mathbb{E}^{(x,k)}(V(X(\gamma_n(t)), Z(\gamma_n(t))) \exp(-\alpha_0 \gamma_n(t))) - V(x, k) \\ &= \mathbb{E}^{(x,k)} \int_0^{\gamma_n(t)} \left( \frac{\partial}{\partial s} + L \right) W(X(s), s, Z(s)) \, ds \\ &\leq 0. \end{aligned}$$

Noting the inequality  $\gamma_n(t) \leq t$  and the nonnegativity of  $V(x, k)$ , we arrive at

$$\mathbb{E}^{(x,k)} V(X(\gamma_n(t)), Z(\gamma_n(t))) \leq V(x, k) \exp(\alpha_0 t).$$

From this we derive the estimate

$$\mathbb{P}^{(x,k)}(\gamma_n \leq t) \leq (\inf\{V(x, k) : |x| \geq n\})^{-1} V(x, k) \exp(\alpha_0 t). \tag{3.10}$$

Letting  $n \rightarrow \infty$  and using (3.9), we conclude that  $\mathbb{P}^{(x,k)}(\gamma = \infty) = 1$ .

Finally, we prove the Feller continuity. By virtue of Theorem 3.2 of [27], we need only verify Assumptions 1.1 and 1.2 of [27]. Remark 3.2, (3.8), and (3.9) imply Assumptions 1.1 and 1.2 of [27]. The proof of the theorem is thus concluded.

### 4. Strong Feller continuity

In this section we prove the strong Feller continuity for  $(X(t), Z(t))$ . For this purpose, we introduce a family of diffusion processes. For each  $k \in N$ , let the single diffusion process  $X^{(k)}(t)$  satisfy the following stochastic differential equation in  $\mathbb{R}^K$ :

$$dX^{(k)}(t) = b(X^{(k)}(t), k) dt + \sigma(X^{(k)}(t), k) dW(t). \tag{4.1}$$

In the sequel, we use  $\{P^{(k)}(t, x, A) : t \geq 0, x \in \mathbb{R}^K, A \in \mathcal{B}(\mathbb{R}^K)\}$  to denote a family of transition probabilities of the diffusion process  $X^{(k)}(t)$ . Moreover, for any given  $t > 0$ , set

$$P^{(k)}(t) := \{P^{(k)}(t, x, A) : x \in \mathbb{R}^K, A \in \mathcal{B}(\mathbb{R}^K)\},$$

which is the transition probability kernel of the Markov process  $X^{(k)}(t)$ . Correspondingly, we denote by  $P^{(k)(x)}$  the distribution of  $X^{(k)}(t)$  starting from  $x$ . Hereafter, we denote the Lebesgue measure on  $\mathbb{R}^K$  by  $m(\cdot)$ .

**Theorem 4.1.** *For any given  $k \in N$ , (4.1) has a unique nonexplosive solution  $X^{(k)}(t)$ . For each  $t \geq 0$  and each  $x \in \mathbb{R}^K$ , the transition probability  $P^{(k)}(t, x, \cdot)$  is absolutely continuous with respect to the Lebesgue measure  $m(\cdot)$ , and, for each  $t > 0$ , the transition probability kernel  $P^{(k)}(t)$  is strongly Feller continuous. Moreover, for each  $t > 0$ , the support of  $X^{(k)}(t)$  is equal to  $\mathbb{R}^K$ .*

*Proof.* For an arbitrary  $k \in N$ , the generator  $L^{(k)}$  of (4.1) is given by

$$L^{(k)} = \frac{1}{2} \sum_{i=1}^K \sigma_{ii}^2(x, k) \frac{\partial^2}{\partial x_i^2} + \sum_{i=1}^K b_i(x, k) \frac{\partial}{\partial x_i}.$$

Set  $V^{(k)}(x) = |x|^2 + 1$  for  $x \in \mathbb{R}^K$ . Similarly to (3.5), we have

$$L^{(k)} V^{(k)}(x) \leq -V^{(k)}(x) + \beta \quad \text{for } x \in \mathbb{R}^K, \tag{4.2}$$

where  $\beta > 0$  is a constant as in Remark 3.2. Next, analogously to the proof of Theorem 3.1, we can construct sequences of functions  $\{b^{(k)(n)}(x)\}$  and  $\{\sigma^{(k)(n)}(x)\}$  such that, for all  $x \in O_n := \{x \in \mathbb{R}^K : |x| < n\}$ ,

$$b^{(k)(n)}(x) = b(x, k), \quad \sigma^{(k)(n)}(x) = \sigma(x, k),$$

and, for each positive integer  $n$ ,

$$b^{(k)(n)}(x) = (b_1^{(k)(n)}(x), \dots, b_K^{(k)(n)}(x))^T : \mathbb{R}^K \rightarrow \mathbb{R}^K$$

and

$$\sigma^{(k)(n)}(x) = \text{diag}\{\sigma_{ii}^{(k)(n)}(x)\} : \mathbb{R}^K \rightarrow \mathbb{R}^K \times \mathbb{R}^K$$

are bounded smooth functions with bounded derivatives, and also that  $\sigma_{ii}^{(k)(n)}(x) > 0$  for all  $1 \leq i \leq K$  and  $x \in \mathbb{R}^K$ . For each positive integer  $n$ , we have a unique strong Markov process  $X^{(k)(n)}(t)$  corresponding to the functions constructed above via (4.1). Moreover, it is clear that  $X^{(k)}(t) = X^{(k)(n)}(t)$  whenever  $t < \xi_n^{(k)}$ , where  $\xi_n^{(k)} = \inf\{t \geq 0 : X^{(k)}(t) \in O_n^c\}$ . Using (4.2), similarly to the proof of Theorem 3.1 (see (3.8)–(3.10)), we can prove that, for each  $t > 0$ ,

$$\lim_{n \rightarrow \infty} P^{(k)(x)}(\xi_n^{(k)} \leq t) = 0 \quad \text{for } x \in \mathbb{R}^K, \tag{4.3}$$

which implies the desired nonexplosiveness.



Now we begin to prove the strong Feller continuity. For each positive integer  $n$ , let the  $C^\infty$  vector fields  $A_0^{(k)}, A_1^{(k)}, \dots, A_K^{(k)}$  on  $\mathbb{R}^K$  be defined as follows:

$$A_0^{(k)(n)}(x) = \sum_{j=1}^K \left( b_j^{(k)(n)}(x) - \frac{1}{2} \sigma_{jj}^{(k)(n)}(x) \frac{\partial}{\partial x_j} \sigma_{jj}^{(k)(n)}(x) \right) \frac{\partial}{\partial x_j},$$

$$A_i^{(k)(n)}(x) = \sigma_{ii}^{(k)(n)}(x) \frac{\partial}{\partial x_i}, \quad i = 1, 2, \dots, K.$$

Since  $\sigma_{ii}^{(k)(n)}(x) > 0$ , we have

$$\dim \text{Lie}\{A_l^{(k)}(x) : 1 \leq l \leq K\} = K \quad \text{for all } x \in \mathbb{R}^K, \tag{4.4}$$

and so

$$\dim \text{Lie}\{A_l^{(k)}(x) : 0 \leq l \leq K\} = K \quad \text{for all } x \in \mathbb{R}^K. \tag{4.5}$$

From (4.5), the hypoelliptic condition is satisfied (see [11] and [15] for the details). Therefore,  $X^{(k)(n)}(t)$  is a strong Feller process with a transition density  $p^{(k)(n)}(t, x, y)$  with respect to the Lebesgue measure  $m(\cdot)$ . Thus, for each  $x \in \mathbb{R}^K$  and each  $t > 0$ , for each set  $A \in \mathcal{B}(\mathbb{R}^K)$  with  $m(A) = 0$ ,

$$P^{(k)(x)}(X^{(k)}(t) \in A, \xi_n^{(k)} \geq t) = 0, \tag{4.6}$$

as long as  $n$  is sufficiently large such that  $n > |x|$ . Meanwhile, we also have

$$P^{(k)(x)}(X^{(k)}(t) \in A) \leq P^{(k)(x)}(X^{(k)}(t) \in A, \xi_n^{(k)} \geq t) + P^{(k)(x)}(\xi_n^{(k)} \leq t).$$

This together with (4.3) and (4.6) shows that the transition probability  $P^{(k)}(t, x, \cdot)$  is absolutely continuous with respect to the Lebesgue measure  $m(\cdot)$ . In view of Theorem 3.1, the transition probability kernel  $P^{(k)}(t)$  is Feller continuous. For a set  $A \in \mathcal{B}(\mathbb{R}^K)$ , let  $A^\circ$  and  $\partial A$  denote the interior and the boundary of  $A$ , respectively. It is easy to see that  $m(\partial A) = 0$  for all  $A \in \mathcal{B}(\mathbb{R}^K)$ . Combining all these facts together with Remark 3.1, we obtain, for every  $x \in \mathbb{R}^K$  and every  $t > 0$ ,

$$\liminf_{y \rightarrow x} P^{(k)}(t, y, A) \geq \liminf_{y \rightarrow x} P^{(k)}(t, y, A^\circ) \geq P^{(k)}(t, x, A^\circ) = P^{(k)}(t, x, A),$$

which implies that  $P^{(k)}(t)$  is strong Feller continuous.

Finally, for each positive integer  $n$ , from (4.4) we also find that the support of  $X^{(k)(n)}(t)$  is equal to  $\mathbb{R}^K$  for all  $t > 0$  (see Section 8 in Chapter VI of [12]). Combining this fact with (4.3), for each  $x \in \mathbb{R}^K$  and each  $t > 0$ , we conclude that, for each set  $A \in \mathcal{B}(\mathbb{R}^K)$  with  $m(A) > 0$ ,

$$\begin{aligned} P^{(k)(x)}(X^{(k)}(t) \in A) &\geq P^{(k)(x)}(X^{(k)}(t) \in A, \xi_n^{(k)} \geq t) \\ &= P^{(k)(x)}(\xi_n^{(k)} \geq t) P^{(k)(x)}(X^{(k)}(t) \in A \mid \xi_n^{(k)} \geq t) \\ &\geq \frac{1}{2} P^{(k)(x)}(X^{(k)}(t) \in A \mid \xi_n^{(k)} \geq t) \\ &> 0 \end{aligned}$$

for sufficiently large  $n$ . This implies that, for each  $t > 0$ , the support of  $X^{(k)}(t)$  is also equal to  $\mathbb{R}^K$ .

**Remark 4.1.** By virtue of the hypoellipticity we have proved the strong Feller continuity for the solution  $X^{(k)}(t)$  of (4.1). The main point is that the drift and diffusion coefficients are not bounded, neither do they satisfy the linear growth condition. As a remedy, we have used the truncation argument.

Note that  $(X(t), Z(t))$  could be viewed as the result of the family of diffusion processes  $X^{(k)}(t), k \in N$ , switching from one to another according to the movement of process  $Z(t)$  that also depends on process  $X(t)$ . To prove the strong Feller continuity for  $(X(t), Z(t))$ , we still need to give a result on the number of jumps of  $Z(t)$  on the interval  $[0, t]$  for any given  $t > 0$ . To this end, we make the following interim assumption. Note that this assumption facilitates our proof and it is used as a bridge in the process of deriving the desired result. Eventually, we will remove it to obtain the desired result.

**Assumption 4.1.** *There exists a positive constant  $\tilde{H} > 0$  such that*

$$\sup\{q_{kl}(x) : x \in \mathbb{R}^K, k \neq l \in N\} \leq \tilde{H} < +\infty.$$

Let us introduce another auxiliary process  $\phi(t)$  as follows. Let  $\phi(t)$  be a right-continuous and time-homogeneous Markov chain with finite state space  $N$  such that

$$P(\phi(t + \Delta) = l \mid \phi(t) = k) = \begin{cases} \tilde{H}\Delta + o(\Delta) & \text{if } k \neq l, \\ 1 - (n_0 - 1)\tilde{H}\Delta + o(\Delta) & \text{if } k = l, \end{cases}$$

as  $\Delta \downarrow 0$ , where  $n_0$  is the number of elements in  $N$ .

Intuitively, Assumption 4.1 prevents  $Z(t)$  from jumping more quickly than  $\phi(t)$ . As in [28], we let  $\{\zeta_\ell\}$  be the sequence of Markov times defined by

$$\zeta_0 = 0, \quad \zeta_\ell = \inf\{s > \zeta_{\ell-1} : Z(t) \neq Z(\zeta_{\ell-1})\},$$

and set  $\zeta = \lim_{\ell \rightarrow \infty} \zeta_\ell$ . Then  $\zeta$  is the first instant prior to which the second component of the strong Markov process  $(X(t), Z(t))$  has infinitely many jumps. Next, by the sequence of Markov times  $\{\zeta_\ell\}$  we define  $J(t) = \max\{\ell : \zeta_\ell \leq t\}$ . Then  $J(t)$  is the number of jumps of the second component of the strong Markov process  $(X(t), Z(t))$  prior to  $t$ . Analogously to the proof of Proposition 3.4 of [28], by virtue of the auxiliary process  $\phi(t)$  we can prove the following lemma; see Section 3 of [28] for a detailed proof.

**Lemma 4.1.** *Suppose that Assumption 4.1 holds. For every  $(x, k) \in \mathbb{R}^K \times N$ , we have*

$$P^{(x,k)}(\zeta = +\infty) = 1, \quad \text{and, consequently,} \quad P^{(x,k)}(J(t) < +\infty) = 1 \tag{4.7}$$

for every  $t > 0$ .

**Proposition 4.1.** *Under Assumption 2.1, for each  $(x, k) \in \mathbb{R}^K \times N$ , each  $t > 0$ , and each set  $D \in \mathcal{B}(\mathbb{R}^K \times N)$ , we have*

$$P^{(x,k)}((X(t), Z(t)) \in D) = \sum_{\ell=0}^{\infty} P^{(x,k)}((X(t), Z(t)) \in D, J(t) = \ell). \tag{4.8}$$

*Proof.* As in the proof of Theorem 3.1, noting Remark 3.2, under Assumption 2.1 only (without Assumption 4.1), we can construct a sequence of functions  $\{q_{kl}^{(n)}(x)\}$  such that, for all  $x \in O_n = \{x \in \mathbb{R}^K : |x| < n\}$  and  $k \neq l \in N, q_{kl}^{(n)}(x) = q_{kl}(x)$ , and, for each positive integer  $n$ , the newly constructed functions also satisfy Assumption 4.1. For each integer  $n \geq 1$ , set  $\zeta_n := \inf\{t \geq 0 : (X(t), Z(t)) \in O_n^c \times N\}$ . Using (3.5), similar to the proof of Theorem 3.1 (see (3.8)–(3.10)), we can prove that, for every  $(x, k) \in \mathbb{R}^K \times N$  and every  $t > 0$ ,  $\lim_{n \rightarrow \infty} P^{(x,k)}(\zeta_n \leq t) = 0$ . Thus, we have

$$\{J(t) < +\infty\} = \bigcap_{n=1}^{\infty} \{J(\zeta_n \wedge t) < +\infty\} \tag{4.9}$$

almost everywhere with respect to  $P^{(x,k)}$ . On the other hand, for any given  $(x, k) \in \mathbb{R}^K \times N$ , by virtue of the second equality in (4.7), for each sufficiently large integer  $n$  satisfying  $n > |x|$  and each  $t > 0$ ,  $P^{(x,k)}(J(\zeta_n \wedge t) < +\infty) = 1$ . Combining this fact with (4.9) and also noting that

$$\{J(\zeta_n \wedge t) < +\infty\} \supset \{J(\zeta_{n+1} \wedge t) < +\infty\},$$

we find that, for every  $(x, k) \in \mathbb{R}^K \times N$  and every  $t > 0$ ,

$$P^{(x,k)}(J(t) < +\infty) = \lim_{n \rightarrow \infty} P^{(x,k)}(J(\zeta_n \wedge t) < +\infty) = 1.$$

Then (4.8) follows.

For the subsequent uses, we introduce a reference measure. Hereafter, we let the reference measure  $\mu(\cdot)$  be the product measure on  $\mathbb{R}^K \times N$  (the Lebesgue measure  $m(\cdot)$  on  $\mathbb{R}^K$  and the counting measure on  $N$ ). Hereafter, for a given set  $D \in \mathcal{B}(\mathbb{R}^K \times N)$  and each  $k \in N$ , let

$$D_k = \{x \in \mathbb{R}^K : (x, k) \in D\}$$

be the section of  $D$  at  $k$ . Then, the product measure theorem (cf. Theorem 1 in Section 6.3 of [5]) gives us the following useful lemma.

**Lemma 4.2.** *For any set  $D \in \mathcal{B}(\mathbb{R}^K \times N)$  with  $\mu(D) = 0$ , we have  $m(D_k) = 0$  for all  $k \in N$ , where  $m(\cdot)$  is just the Lebesgue measure on  $\mathbb{R}^K$ .*

**Proposition 4.2.** *We have the following two assertions.*

(i) *For each  $(x, k) \in \mathbb{R}^K \times N$ , each  $t > 0$ , and each integer  $\ell \geq 0$ ,*

$$P^{(x,k)}((X(t), Z(t)) \in D, J(t) = \ell) = 0$$

*for all sets  $D \in \mathcal{B}(\mathbb{R}^K \times N)$  satisfying  $\mu(D) = 0$ .*

(ii) *For every  $(x, k) \in \mathbb{R}^K \times N$  and every  $t > 0$ ,*

$$P^{(x,k)}((X(t), Z(t)) \in D) = 0$$

*for all sets  $D \in \mathcal{B}(\mathbb{R}^K \times N)$  satisfying  $\mu(D) = 0$ .*

*Proof.* Note that, for each  $k \in N$ , the transition probability  $P^{(k)}(t, x, \cdot)$  of the diffusion process  $X^{(k)}(t)$  determined in (4.1) is absolutely continuous with respect to the Lebesgue measure  $m(\cdot)$  that was proved in Theorem 4.1. Also, note that the first component of  $(X(t), Z(t))$  coincides with the diffusion process  $X^{(k)}(t)$  on the interval  $[\zeta_0, \zeta_1]$  when  $Z(\zeta_0) = k$ . To proceed, we prove Proposition 4.2(i) by mathematical induction on  $\ell \geq 0$ . When  $\ell = 0$ , from the two facts noted above and Lemma 4.2, we obtain

$$\begin{aligned} &P^{(x,k)}((X(t), Z(t)) \in D, J(t) = 0) \\ &= P(X(t) \in D_k, Z(t) = k, \zeta_0 \leq t < \zeta_1 \mid X(0) = x, Z(0) = k) \\ &\leq P(X^{(k)}(t) \in D_k \mid X^{(k)}(0) = x) \\ &= 0 \end{aligned}$$

for any set  $D \in \mathcal{B}(\mathbb{R}^K \times N)$  satisfying  $\mu(D) = 0$ . Now, we suppose that assertion (i) holds for  $J(t) = \ell$ . By virtue of the strong Markov property and the induction hypothesis, we conclude that

$$\begin{aligned} & \mathbb{P}^{(x,k)}((X(t), Z(t)) \in D, J(t) = \ell + 1) \\ &= \mathbb{P}^{(x,k)}((X(t), Z(t)) \in D, \zeta_{\ell+1} \leq t < \zeta_{\ell+2}) \\ &= \mathbb{E}^{(x,k)}(\mathbb{P}^{(X(\zeta_1), Z(\zeta_1))}((X(t), Z(t)) \in D, \zeta_\ell \leq t < \zeta_{\ell+1})) \\ &= 0 \end{aligned}$$

for any set  $D \in \mathcal{B}(\mathbb{R}^K \times N)$  satisfying  $\mu(D) = 0$ . This proves assertion (i) of Proposition 4.2, while assertion (ii) follows from assertion (i) and (4.8) in Proposition 4.1.

**Theorem 4.2.** *For each  $(x, k) \in \mathbb{R}^K \times N$  and each  $t > 0$ , the transition probability  $\mathbb{P}(t, (x, k), \cdot)$  of  $(X(t), Z(t))$  is absolutely continuous with respect to  $\mu(\cdot)$ . In addition, for any given  $t > 0$ , the transition probability kernel  $\mathbb{P}(t)$  defined in (3.4) is strong Feller continuous. Furthermore, for each  $t > 0$ , the support of  $(X(t), Z(t))$  is equal to  $\mathbb{R}^K \times N$ .*

*Proof.* The first assertion follows from Proposition 4.2. We proceed to prove the second assertion. For a set  $A \in \mathcal{B}(\mathbb{R}^K \times N)$ , let  $A^\circ$  and  $\partial A$  denote the interior and the boundary of  $A$ , respectively. It follows from the definition of the reference measure  $\mu(\cdot)$  that  $\mu(\partial A) = 0$  for all  $A \in \mathcal{B}(\mathbb{R}^K \times N)$ . Therefore, by virtue of the absolute continuity stated in the first assertion and the Feller continuity proved in Theorem 3.1, and recalling Remark 3.1, we conclude that, for every  $(x, k) \in \mathbb{R}^K \times N$  and every  $t > 0$ ,

$$\begin{aligned} \liminf_{y \rightarrow x} \mathbb{P}(t, (y, k), A) &\geq \liminf_{y \rightarrow x} \mathbb{P}(t, (y, k), A^\circ) \\ &\geq \mathbb{P}(t, (x, k), A^\circ) \\ &= \mathbb{P}(t, (x, k), A), \end{aligned}$$

which implies the second assertion. Finally, let us prove the third assertion. Note that, for any given  $k \in N$ , the first component of  $(X(t), Z(t))$  coincides with the diffusion process  $X^{(k)}(t)$  determined by (4.1) on the interval  $[\zeta_0, \zeta_1)$  when  $Z(\zeta_0) = k$  and the support of  $X^{(k)}(t)$  is equal to  $\mathbb{R}^K$  for all  $t > 0$  (see Theorem 4.1). Therefore, by virtue of the fact that  $0 < q_{kl}(x) < +\infty$  and by its continuity with respect to  $x$  for any  $k \neq l \in N$  in Assumption 2.1, the third assertion holds. The proof is completed.

### 5. Exponential ergodicity

In this section we investigate the exponential ergodicity for the strong Markov process  $(X(t), Z(t))$ . As in [28], we first consider a form of stability called the boundedness in probability on average for  $(X(t), Z(t))$  (see Section 3 of [21] for more details about the corresponding concept). The process  $(X(t), Z(t))$  is called bounded in probability on average if, for each  $(x, k) \in \mathbb{R}^K \times N$  and each  $\varepsilon > 0$ , there exists a compact subset  $C \subset \mathbb{R}^K$  such that

$$\liminf_{t \rightarrow \infty} \frac{1}{t} \int_0^t \mathbb{P}(t, (x, k), C \times N) \, ds \geq 1 - \varepsilon.$$

We introduce a Foster–Lyapunov drift condition. For some  $\alpha, \beta > 0$ ,  $f(x, k) \geq 1$ , a compact set  $C \subset \mathbb{R}^K$ , and some  $V(x, k) \geq 0$  that is twice continuously differentiable in  $x$  on  $\mathbb{R}^K \times N$ ,

$$LV(x, k) \leq -\alpha f(x, k) + \beta \mathbf{1}_{C \times N}(x, k), \quad (x, k) \in \mathbb{R}^d \times N, \tag{5.1}$$

where  $L$  is the operator defined in (3.3) and  $\mathbf{1}_{C \times N}$  is the indicator function of the set  $C \times N$ .

**Theorem 5.1.** *The strong Markov process  $(X(t), Z(t))$  is bounded in probability on average.*

*Proof.* By virtue of Theorem 4.2,  $(X(t), Z(t))$  is a strong Feller process and so it is a so-called  $T$ -process (see [21] for the detailed definition of  $T$ -processes). From Theorem 3.1,  $(X(t), Z(t))$  is nonexplosive. Thus, (3.5) implies (5.1). Therefore, the desired result follows from Theorem 4.7 of [22].

We proceed to investigate the exponential ergodicity for the strong Markov process  $(X(t), Z(t))$ . As in [22], for any positive function  $\Psi(x, k) \geq 1$  defined on  $\mathbb{R}^K \times N$  and any signed measure  $\nu(\cdot)$  defined on  $\mathcal{B}(\mathbb{R}^K \times N)$ , we write

$$\|\nu\|_\Psi = \sup\{|\nu(\Phi)| : \text{all measurable } \Phi(x, k) \text{ satisfying } |\Phi| \leq \Psi\},$$

where

$$\nu(\Phi) = \sum_{k \in N} \int_{\mathbb{R}^K} \Phi(x, k) \nu(dx, k).$$

Note that the total variation norm  $\|\nu\|_{\text{var}}$  is just  $\|\nu\|_\Psi$  in the case  $\Psi \equiv 1$ .

For a function  $1 \leq \Psi < \infty$  on  $\mathbb{R}^K \times N$ , the Markov process  $(X(t), Z(t))$  is said to be  $\Psi$ -exponentially ergodic if there exist a probability measure  $\pi(\cdot)$ , a positive constant  $\theta < 1$ , and a finite-valued function  $\Theta(x, k)$  such that

$$\|\mathbb{P}(t, (x, k), \cdot) - \pi(\cdot)\|_\Psi \leq \Theta(x, k)\theta^t$$

for all  $t \geq 0$  and all  $(x, k) \in \mathbb{R}^K \times N$ . Moreover, a nonnegative function  $V(x, k)$  defined on  $\mathbb{R}^K \times N$  is called a norm-like function if  $V(x, k) \rightarrow \infty$  as  $|x| \rightarrow \infty$  for all  $k \in N$  (also known as radially unbounded in the literature of ordinary differential equations). For some  $\alpha, \beta > 0$  and a norm-like function  $V(x, k)$  that is twice continuously differentiable in  $x$ , the Foster–Lyapunov drift condition requires that

$$LV(x, k) \leq -\alpha V(x, k) + \beta, \quad (x, k) \in \mathbb{R}^K \times N. \tag{5.2}$$

**Theorem 5.2.** *The strong Markov process  $(X(t), Z(t))$  is  $\Psi$ -exponentially ergodic with*

$$\Psi(x, k) = |x|^2 + 1 \quad \text{and} \quad \Theta(x, k) = B(|x|^2 + 1),$$

where  $B$  is a finite constant.

*Proof.* By the proof of Theorem 5.1,  $(X(t), Z(t))$  is a  $T$ -process. Since the support of  $(X(t), Z(t))$  is equal to  $\mathbb{R}^K \times N$  for all  $t > 0$  (see Theorem 4.2),  $(X(t), Z(t))$  is irreducible with respect to the reference measure  $\mu(\cdot)$  defined in the previous section. Consequently,  $(X(t), Z(t))$  is a  $\mu$ -irreducible  $T$ -process. Hence, for any given constant  $h > 0$ , the  $h$ -skeleton chain  $(X(nh), Z(nh))_{n \geq 0}$  is an irreducible  $T$ -chain. By virtue of Theorem 3.1,  $(X(t), Z(t))$  is nonexplosive. In view of Theorem 3.2 of [19], we find that all compact sets of the state space  $\mathbb{R}^K \times N$  are petite for the  $h$ -skeleton chain  $(X(nh), Z(nh))_{n \geq 0}$  (see [19] for detailed definitions of  $T$ -chains and petite sets). Clearly, (3.5) implies (5.2) with  $V(x, k) = |x|^2 + 1$ . Then, this implies (5.2) with  $V(x, k) = |x|^2$  too. Consequently, applying Theorem 6.1 of [22] to the strong Markov process  $(X(t), Z(t))$ , we finally obtain the desired result. This completes the proof.

**Remark 5.1.** Because  $(X(t), Z(t))$  is  $\Psi$ -exponentially ergodic,  $(X(t), Z(t))$  is automatically positive Harris recurrent (see [8] and [21]). Moreover, since  $\Psi$ -exponential ergodicity carries over to the seemingly stronger context of  $\Psi$ -uniform ergodicity,  $(X(t), Z(t))$  is  $\Psi$ -uniformly ergodic; that is, for a positive constant  $\theta < 1$ ,

$$\sup \left\{ \frac{\|P(t, (x, k), \cdot) - \pi(\cdot)\|_{\Psi}}{\Psi(x, k)} : (x, k) \in \mathbb{R}^K \times N \right\} \leq B\theta^t \quad \text{for all } t \geq 0, \quad (5.3)$$

where  $\Psi(x, k) = |x|^2 + 1$  and  $B$  is a finite constant. In fact, the right-hand side of (5.3) can be defined as an operator norm and it describes the  $\Psi$ -uniform ergodicity which is a generalization of the classical uniform ergodicity (or strong ergodicity). However, the denominator  $\Psi(x, k)$  is a radially unbounded function. Nevertheless, in the next section we will consider classical uniform ergodicity (i.e. when  $\Psi(x, k) \equiv 1$ ) for  $(X(t), Z(t))$  when it is in a Markovian switching case.

**Remark 5.2.** Note that Theorems 3.1, 4.2, 5.1, and 5.2 can be applied to Example 1.1 and Example 1.2.

### 6. Strong ergodicity

For Markov processes, the strong ergodicity or uniform ergodicity in the sense of convergence in the variation norm is the strongest one; see [3, Chapter 4] and [20, Chapter 16]. Strong ergodicity for some Markov processes, especially for one-dimensional diffusion processes and birth–death processes, was studied in [17] and [18] using coupling methods. From the previous section, the Markov process  $(X(t), Z(t))$  that we are interested in admits a unique invariant probability measure  $\pi(\cdot)$ . In this section we will use the coupling methods to obtain the strong ergodicity for Markovian switching diffusion  $(X(t), Z(t))$ . To do so, we make an assumption.

**Assumption 6.1.** For all  $k \neq l \in N$ , the functions  $q_{kl}(x) \equiv q_{kl} > 0$  are independent of  $x$  and the second component  $Z(t)$  determined by (2.2) is a Markov chain independent of  $W(t)$ . Moreover, the functions  $\sigma_{11}(x, k) = \sigma_{22}(x, k) = \dots = \sigma_{KK}(x, k)$  in (2.1) satisfy  $\kappa \leq \sigma_{11}(x, k) \leq \kappa^{-1}$  with some constant  $\kappa \in (0, 1]$  for all  $(x, k) \in \mathbb{R}^K \times N$ .

Under Assumption 6.1,  $(X(t), Z(t))$  described by (3.2) (or (2.1) and (2.2)) is a mean-field model with Markovian switching, which is a special case of the general diffusion processes with continuous-state-dependent switching. In what follows we will study the strong ergodicity for this special case by making use of the coupling methods. More precisely, we will construct a coupling of  $(X(t), Z(t))$  and itself, and then estimate the coupling time to prove the desired strong ergodicity. To proceed, let us cite the following definition of strong ergodicity.

**Definition 6.1.** ([18].) The Markov process  $(X(t), Z(t))$  is said to be strongly ergodic in variational norm (in short, strongly ergodic) if there exists an  $\varepsilon > 0$  such that

$$\sup\{\|P(t, (x, k), \cdot) - \pi(\cdot)\|_{\text{var}} : (x, k) \in \mathbb{R}^K \times N\} = O(e^{-\varepsilon t}) \quad \text{as } t \rightarrow \infty.$$

For  $\gamma \geq 2$ , define

$$\hat{\alpha}(\gamma) = \sup\{\varepsilon \geq 0 : \sup\{\|P(t, (x, k), \cdot) - \pi(\cdot)\|_{\text{var}} : (x, k) \in \mathbb{R}^K \times N\} \leq \gamma \exp(-\varepsilon t) \text{ for all } t \geq 0\}.$$

Furthermore, let  $\hat{\alpha} := \hat{\alpha}(\infty) = \lim_{\gamma \rightarrow \infty} \hat{\alpha}(\gamma)$ .

We begin to construct the coupling of  $(X(t), Z(t))$ . We first construct the classical coupling of the discrete component  $Z(t)$  as follows. Let  $(Z(t), Z'(t))$  be the Markov chain with phase space  $N \times N$  and coupling operator (see [3] for details)

$$\begin{aligned} \Omega f(k_1, k_2) = & \mathbf{1}_{\Delta^c}(k_1, k_2) \left( \sum_{l_1 \in N} q_{k_1 l_1} (f(l_1, k_2) - f(k_1, k_2)) \right. \\ & \left. + \sum_{l_2 \in N} q_{k_2 l_2} (f(k_1, l_2) - f(k_1, k_2)) \right) \\ & + \mathbf{1}_{\Delta}(k_1, k_2) \sum_{l_1 \in N} q_{k_1 l_1} (f(l_1, l_1) - f(k_1, k_2)) \quad \text{for } k_1, k_2 \in N, \end{aligned} \tag{6.1}$$

where  $f$  is a bounded function on  $N \times N$  and  $\Delta := \{(k_1, k_2) \in N \times N : k_1 = k_2\}$ . Set

$$S = \inf\{t \geq 0 : Z(t) = Z'(t)\}.$$

Clearly,  $Z(t)$  and  $Z'(t)$  move independently until the moment  $S$  and they move together from the moment  $S$  onwards. From [3], we can construct many kinds of couplings of  $Z(t)$ . However, in what follows we use only  $(Z(t), Z'(t))$ , the classical coupling constructed above.

**Lemma 6.1.** *Suppose that Assumption 6.1 holds. Set*

$$\varepsilon_0 := \min \left\{ \min \left\{ (1 - e^{-q_k}) \frac{q_{kl}}{q_k} : l \in N \setminus \{k\} \right\}, e^{-q_k} : k \in N \right\}, \tag{6.2}$$

where  $q_k = \sum_{l \neq k} q_{kl}$ . Then, for any  $k, l \in N$ , we have

$$P^{(k,l)}(S > t) \leq (1 - \varepsilon_0)^{\lfloor t \rfloor}, \quad t > 0, \tag{6.3}$$

$$E^{(k,l)}(S) \leq \varepsilon_0^{-1} < +\infty, \tag{6.4}$$

where  $P^{(k,l)}$  and  $E^{(k,l)}$  denote the probability and expectation for the coupling  $(Z(t), Z'(t))$  starting from  $(k, l)$ , respectively, and  $\lfloor t \rfloor$  denotes the integral part of  $t$ .

*Proof.* Let  $\tau_1 = \inf\{t \geq 0 : Z(t) \neq Z(0)\}$  be the first jump time of the Markov chain  $Z(t)$ . Then, we have

$$\begin{aligned} P(Z(\tau_1) = l, \tau_1 < 1 \mid Z(0) = k) &= (1 - e^{-q_k}) \frac{q_{kl}}{q_k}, \quad l \neq k, \\ P(\tau_1 \geq 1 \mid Z(0) = k) &= e^{-q_k}. \end{aligned}$$

Thus, the classical coupling  $(Z(t), Z'(t))$  constructed by (6.1) yields, for any  $k_1 \neq k_2 \in N$ ,

$$\begin{aligned} P^{(k_1, k_2)}(S \leq 1) &\geq \sum_{l \neq k_1, k_2} (1 - e^{-q_{k_1}}) \frac{q_{k_1 l}}{q_{k_1}} (1 - e^{-q_{k_2}}) \frac{q_{k_2 l}}{q_{k_2}} \\ &\quad + (1 - e^{-q_{k_1}}) \frac{q_{k_1 k_2}}{q_{k_1}} e^{-q_{k_2}} + (1 - e^{-q_{k_2}}) \frac{q_{k_2 k_1}}{q_{k_2}} e^{-q_{k_1}} \\ &\geq \varepsilon_0 \sum_{l \neq k_1, k_2} (1 - e^{-q_{k_1}}) \frac{q_{k_1 l}}{q_{k_1}} + \varepsilon_0 (1 - e^{-q_{k_1}}) \frac{q_{k_1 k_2}}{q_{k_1}} + \varepsilon_0 e^{-q_{k_1}} \\ &= \varepsilon_0. \end{aligned}$$

Meanwhile, for  $k_1 = k_2$ , we have  $\mathbf{P}^{(k_1, k_2)}(S \leq 1) = 1 \geq \varepsilon_0$ . Combining these two facts, we arrive at

$$\mathbf{P}^{(k, l)}(S > 1) \leq (1 - \varepsilon_0) \quad \text{for all } k, l \in N.$$

From this and the Markov property, it is easy to prove (6.3) by the induction method. In fact, suppose that  $\mathbf{P}^{(k, l)}(S > \lfloor t \rfloor - 1) \leq (1 - \varepsilon_0)^{\lfloor t \rfloor - 1}$ . Then

$$\begin{aligned} \mathbf{P}^{(k, l)}(S > t) &\leq \mathbf{E}^{(k, l)}(\mathbf{1}_{(S > \lfloor t \rfloor)}) \\ &= \mathbf{E}^{(k, l)}(\mathbf{1}_{(S > 1)} \mathbf{E}^{(X(1), Z(1))}(\mathbf{1}_{(S > \lfloor t \rfloor - 1)})) \\ &\leq (1 - \varepsilon_0)^{\lfloor t \rfloor - 1} \mathbf{P}^{(k, l)}(S > 1) \\ &\leq (1 - \varepsilon_0)^{\lfloor t \rfloor}. \end{aligned}$$

Finally, from (6.3) we obtain

$$\mathbf{E}^{(k, l)}(S) \leq \sum_{n=1}^{\infty} n \mathbf{P}^{(k, l)}((n - 1) \leq S < n) = \sum_{n=1}^{\infty} \mathbf{P}^{(k, l)}(S \geq (n - 1)) \leq \frac{1}{\varepsilon_0},$$

which is just (6.4). This completes the proof.

Combining Lemma 4.1 of [18] and (6.4), we obtain the following proposition.

**Proposition 6.1.** *Suppose that Assumption 6.1 holds. For any nonnegative integer  $n$ , we have*

$$\sup\{\mathbf{E}^{(k, l)}(S^n) : k, l \in N\} \leq \frac{n!}{\varepsilon_0^n}, \tag{6.5}$$

where  $\varepsilon_0$  is defined in (6.2).

Now, for all  $x, y \in \mathbb{R}^K$  and  $k \in N$ , set the  $K \times K$  matrix

$$c(x, y, k) = \sigma_{11}(x, k)\sigma_{11}(y, k) \left( \frac{I - 2(x - y)(x - y)^\top}{|x - y|^2} \right),$$

where  $I$  denotes the  $K \times K$  identity matrix. Next, set the  $2K \times 2K$  matrix

$$\tilde{a}(x, y, k) = \begin{pmatrix} \sigma_{11}^2(x, k)I & c(x, y, k) \\ c(x, y, k)^\top & \sigma_{11}^2(y, k)I \end{pmatrix}.$$

Moreover, recalling that  $(Z(t), Z'(t))$  is the classical coupling constructed above, set

$$\hat{a}(t, x, Z(t), y, Z'(t)) = \begin{pmatrix} \sigma_{11}^2(x, Z(t))I & \hat{c}(t, x, Z(t), y, Z'(t)) \\ \hat{c}(t, x, Z(t), y, Z'(t))^\top & \sigma_{11}^2(y, Z'(t))I \end{pmatrix},$$

where  $\hat{c}(t, x, Z(t), y, Z'(t)) = \mathbf{1}_{[S, \infty)}(t)c(x, y, Z(t))$ . Thus, we can construct a coupling of  $(X(t), Z(t))$  and itself by letting  $(X(t), Y(t))$  satisfy the following stochastic differential equation in  $\mathbb{R}^{2K}$ :

$$d \begin{pmatrix} X(t) \\ Y(t) \end{pmatrix} = \begin{pmatrix} b(X(t), Z(t)) \\ b(Y(t), Z'(t)) \end{pmatrix} dt + \hat{\sigma}(t, X(t), Z(t), Y(t), Z'(t)) dB(t), \tag{6.6}$$

where

$$\hat{\sigma}(t, x, Z(t), y, Z'(t))\hat{\sigma}(t, x, Z(t), y, Z'(t))^\top = \hat{a}(t, x, Z(t), y, Z'(t))$$



and  $B(t)$  is a Brownian motion independent of  $(Z(t), Z'(t))$  in  $\mathbb{R}^{2K}$ . On the interval  $[0, S)$ ,  $(X(t), Y(t))$  given by (6.6) is the independent coupling of  $X(t)$  and  $Y(t)$ , whereas on the interval  $[S, \infty)$ ,  $(X(t), Y(t))$  given by (6.6) is the coupling by reflection of  $X(t)$  and  $Y(t)$  (see [4] for detailed definitions of various couplings). Hence, we have constructed a coupling  $(X(t), Z(t), Y(t), Z'(t))$ . In what follows we let  $P^{(x,k,y,l)}$  denote the distribution of the coupling  $(X(t), Z(t), Y(t), Z'(t))$  starting from  $(x, k, y, l)$ , and let  $E^{(x,k,y,l)}$  denote the corresponding expectation. Set

$$T = \inf\{t \geq 0: X(t) = Y(t), Z(t) = Z'(t)\}.$$

Then  $T$  is the coupling time of  $(X(t), Z(t))$  and  $(Y(t), Z'(t))$ .

To estimate the coupling time  $T$ , as was done in [4], let us set

$$\begin{aligned} A(x, y, k) &= \sigma_{11}^2(x, k)I + \sigma_{11}^2(y, k)I - 2c(x, y, k), \\ \bar{A}(x, y, k) &= \frac{\langle x - y, A(x, y, k)(x - y) \rangle}{|x - y|^2}, \quad x \neq y, \\ \hat{B}(x, y, k) &= \langle x - y, b(x, k) - b(y, k) \rangle, \end{aligned}$$

for  $x, y \in \mathbb{R}^K$  and  $k \in N$ , where  $\langle \cdot, \cdot \rangle$  denotes the inner product in  $\mathbb{R}^K$ . With some straightforward calculations, we have the following estimates on the above functions under Assumption 6.1:

$$\text{tr } A(x, y, k) = \bar{A}(x, y, k) = 4\sigma_{11}(x, k)\sigma_{11}(y, k), \tag{6.7}$$

$$\frac{1}{4}\kappa^2 \leq \frac{1}{\bar{A}(x, y, k)} \leq \frac{1}{4}\kappa^{-2}. \tag{6.8}$$

Furthermore, it is readily seen that

$$\begin{aligned} \hat{B}(x, y, k) &= (\alpha(k) - \beta(k)) \sum_{i=1}^K (x_i - y_i)^2 - \sum_{i=1}^K (x_i - y_i)^2 (x_i^2 + x_i y_i + y_i^2) \\ &\quad + \frac{\beta(k)}{K} \sum_{i=1}^K \sum_{j=1}^K (x_i - y_i)(x_j - y_j). \end{aligned}$$

Noting that

$$\begin{aligned} x_i^2 + x_i y_i + y_i^2 &\geq \frac{1}{4}(x_i - y_i)^2, \\ (x_i - y_i)(x_j - y_j) &\leq \frac{1}{2}[(x_i - y_i)^2 + (x_j - y_j)^2], \end{aligned}$$

we have

$$\hat{B}(x, y, k) \leq \alpha(k) \sum_{i=1}^K (x_i - y_i)^2 - \frac{1}{4} \sum_{i=1}^K (x_i - y_i)^4.$$

It can be shown that

$$\sum_{i=1}^K (x_i - y_i)^4 \geq \frac{1}{K} \left( \sum_{i=1}^K (x_i - y_i)^2 \right)^2.$$

Combining the last two inequalities, we obtain

$$\begin{aligned} \hat{B}(x, y, k) &\leq \alpha(k) \sum_{i=1}^K (x_i - y_i)^2 - \frac{1}{4K} \left( \sum_{i=1}^K (x_i - y_i)^2 \right)^2 \\ &= \alpha(k)|x - y|^2 - \frac{1}{4K}|x - y|^4. \end{aligned} \tag{6.9}$$

From (6.7), (6.8), and (6.9), we obtain

$$\frac{\text{tr}A(x, y, k) - \bar{A}(x, y, k) + 2\hat{B}(x, y, k)}{\bar{A}(x, y, k)} \leq \frac{1}{2}\kappa^{-2}\alpha(k)|x - y|^2 - \frac{1}{8K}\kappa^2|x - y|^4,$$

which gives us, for every  $r \in \mathbb{R}_+$ ,

$$\begin{aligned} & \sup \left\{ \frac{\text{tr}A(x, y, k) - \bar{A}(x, y, k) + 2\hat{B}(x, y, k)}{\bar{A}(x, y, k)} : |x - y| = r, k \in N \right\} \\ & \leq \frac{1}{2}\kappa^{-2}\hat{\alpha}r^2 - \frac{1}{8K}\kappa^2r^4 \\ & = \gamma(r), \end{aligned} \tag{6.10}$$

where  $\hat{\alpha} = \max\{\alpha(k) : k \in N\}$ , a positive constant, and  $\gamma(r)$  denotes the function on the second line of (6.10). Moreover, we define two functions on  $\mathbb{R}_+$  as follows:

$$C(r) = \exp\left(\int_1^r \frac{\gamma(u)}{u} du\right), \tag{6.11}$$

$$\eta(r) := 4\kappa^2 \leq \inf\{\bar{A}(x, y, k) : |x - y| = r, k \in N\}. \tag{6.12}$$

**Lemma 6.2.** *Suppose that Assumption 6.1 holds. We have*

$$\sup\{E^{(x,k,y,k)}(T) : x, y \in \mathbb{R}^K, k \in N\} < \infty. \tag{6.13}$$

*Proof.* Recall that  $S \equiv 0$  if  $Z(0) = Z'(0)$ . In this case, for all  $x, y \in \mathbb{R}^K$  and  $t \geq 0$ , the diffusion matrix  $\hat{a}(t, x, Z(t), y, Z'(t))$ , defined above (6.6), is identical to the matrix  $\tilde{a}(x, y, Z(t))$ . Analogously to the proofs of Theorems 4.2 and 5.1 of [4], for positive integers  $\ell$  and  $n$ , we set

$$\begin{aligned} S_\ell &= \inf\{t \geq 0 : |X(t) - Y(t)| > \ell\}, & \ell > 1, \\ T_n &= \inf\left\{t \geq 0 : |X(t) - Y(t)| < \frac{1}{n}\right\}, & n > 1, \\ T_{n,\ell} &= T_n \wedge S_\ell, \\ F_{n,\ell}(r) &= -\int_{1/n}^r C(s)^{-1} ds \int_s^\ell \frac{C(u)}{\eta(u)} du, & \frac{1}{n} \leq r \leq \ell, n, \ell > 1. \end{aligned}$$

It is easy to verify that

$$\begin{aligned} -\infty < F_{n,\ell}(r) \leq 0, & \quad F'_{n,\ell}(r) \leq 0, \\ F''_{n,\ell}(r) + \frac{F'_{n,\ell}(r)\gamma(r)}{r} &= \frac{1}{\eta(r)}. \end{aligned} \tag{6.14}$$

Using Equation (2.8) of [4] and noting (6.10), (6.12), and (6.14), we can verify that

$$2\tilde{L}(k)F_{n,\ell}(\rho(x, y)) \geq 1 \tag{6.15}$$

for all  $x, y \in \mathbb{R}^K$  with  $1/n \leq \rho(x, y) \leq \ell$  and all  $k \in N$ , where  $\tilde{L}(k)$  is the coupling differential operator defined by the above  $\tilde{a}(x, y, k)$  and  $\tilde{b}(x, y, k) := (b(x, k), b(y, k))^\top$ , as in [4].

Set  $r = \rho(x, y)$ . Using Dynkin’s formula (cf. Lemma 3 in Chapter II of [24]), from (6.15) we deduce that

$$\begin{aligned} & \mathbb{E}^{(x,k,y,k)} F_{n,\ell}(\rho(X(t \wedge T_{n,\ell}), Y(t \wedge T_{n,\ell}))) - F_{n,\ell}(r) \\ &= \mathbb{E}^{(x,k,y,k)} \int_0^{t \wedge T_{n,\ell}} \tilde{L}(Z(u)) F_{n,\ell}(\rho(X(u), Y(u))) du \\ &\geq \frac{1}{2} \mathbb{E}^{(x,k,y,k)}(t \wedge T_{n,\ell}), \end{aligned}$$

so that

$$\mathbb{E}^{(x,k,y,k)}(t \wedge T_{n,\ell}) \leq -2F_{n,\ell}(r).$$

Letting  $t \uparrow \infty$ , we arrive at

$$\mathbb{E}^{(x,k,y,k)}(T_{n,\ell}) \leq -2F_{n,\ell}(r). \tag{6.16}$$

Next, we define a function on  $\mathbb{R}_+$  as follows:

$$F(r) := \lim_{n \rightarrow \infty} \lim_{\ell \rightarrow \infty} F_{n,\ell}(r) = - \int_0^r C(s)^{-1} ds \int_s^\infty \frac{C(u)}{\eta(u)} du. \tag{6.17}$$

Letting  $\ell \uparrow \infty$  and then  $n \uparrow \infty$  on the both sides of (6.16), we obtain

$$\mathbb{E}^{(x,k,y,k)}(T) \leq -2F(r). \tag{6.18}$$

To prove (6.13), it suffices to show that

$$F(\infty) := - \int_0^\infty C(s)^{-1} ds \int_s^\infty \frac{C(u)}{\eta(u)} du > -\infty. \tag{6.19}$$

Using (6.10), (6.11) and (6.12), we obtain

$$C(s) = \exp\left(\frac{1}{4}\kappa^{-2}\hat{\alpha}s^2 - \frac{1}{4}\kappa^{-2}\hat{\alpha} - \frac{1}{32K}\kappa^2s^4 + \frac{1}{32K}\kappa^2\right).$$

Hence, for the integrand in (6.19), we have

$$\begin{aligned} & C(s)^{-1} \int_s^\infty \frac{C(u)}{\eta(u)} du \\ &= \frac{1}{4}\kappa^{-2} \exp\left(\frac{1}{32K}\kappa^2s^4 - \frac{1}{4}\kappa^{-2}\hat{\alpha}s^2\right) \int_s^\infty \exp\left(\frac{1}{4}\kappa^{-2}\hat{\alpha}u^2 - \frac{1}{32K}\kappa^2u^4\right) du \\ &= f(s), \end{aligned} \tag{6.20}$$

where  $f(s)$  denotes the function on the second line of (6.20). Define an auxiliary function

$$g(s) := \frac{2K\kappa^{-4}}{s^3}, \quad \text{and note that } \int_1^\infty g(s) ds < \infty. \tag{6.21}$$

By virtue of l’Hôpital’s rule and with some lengthy but elementary calculations, we can verify that  $\lim_{s \rightarrow \infty} f(s)/g(s) = 1$ . Combining this with (6.20) and (6.21), we obtain

$$\int_1^\infty C(s)^{-1} ds \int_s^\infty \frac{C(u)}{\eta(u)} du < \infty. \tag{6.22}$$

Meanwhile, from (6.20) we have

$$\int_0^1 C(s)^{-1} ds \int_s^\infty \frac{C(u)}{\eta(u)} du < \infty. \tag{6.23}$$

Then (6.22) and (6.23) imply (6.19), so (6.13) is proved. This completes the proof.

**Proposition 6.2.** *Suppose that Assumption 6.1 holds. For any nonnegative integer  $n$ , we have*

$$\sup\{E^{(x,k,y,k)}(T^n) : x, y \in \mathbb{R}^K, k \in N\} \leq n!(-2F(\infty))^n, \tag{6.24}$$

where  $F(\infty)$  is defined in (6.19).

*Proof.* In view of (6.15), for the function  $F(r)$  defined in (6.17) and the coupling differential operator  $\tilde{L}(k)$  defined by the  $\tilde{a}(x, y, k)$  and  $\tilde{b}(x, y, k)$  as above, we have

$$\tilde{L}(k)(-2F(\rho(x, y))) \leq -1, \quad x \neq y \in \mathbb{R}^K, k \in N.$$

Therefore, from Theorem 5.18 and Equation (5.35) of [2], (6.24) holds.

**Theorem 6.1.** *Suppose that Assumption 6.1 holds. We have*

$$\sup\{E^{(x,k,y,l)}(T) : x, y \in \mathbb{R}^K, k, l \in N\} \leq \varepsilon_0^{-1} - 2F(\infty) < \infty, \tag{6.25}$$

where  $\varepsilon_0$  and  $F(\infty)$  are defined in (6.2) and (6.19), respectively. Furthermore, for any nonnegative integer  $n$ , we have

$$\sup\{E^{(x,k,y,l)}(T^n) : x, y \in \mathbb{R}^K, k, l \in N\} \leq n! \left( \varepsilon_0^{-1} - 2F(\infty) \right)^n. \tag{6.26}$$

Moreover, the Markov process  $(X(t), Z(t))$  is strongly ergodic with

$$\hat{\alpha} \geq [\varepsilon_0^{-1} - 2F(\infty)]^{-1} > 0, \tag{6.27}$$

where  $\hat{\alpha}$  is the rate constant defined in Definition 6.1.

*Proof.* With a slight abuse of notation, we denote by  $\theta$  the usual shift operator in what follows. Hence, we obtain

$$E^{(x,k,y,l)}(T) = E^{(x,k,y,l)}(S) + E^{(x,k,y,l)}(\theta_S T). \tag{6.28}$$

Clearly, using (6.4), for the first term on the right-hand side of (6.28), we have

$$E^{(x,k,y,l)}(S) = E^{(k,l)}(S) \leq \varepsilon_0^{-1}. \tag{6.29}$$

On the other hand, by virtue of the strong Markov property, and (6.18) and (6.19), for the second term on the right-hand side of (6.28), we have

$$E^{(x,k,y,l)}(\theta_S T) \leq E^{(x,k,y,l)}(E^{(X(S), Z(S), Y(S), Z'(S))}(T)) \leq -2F(\infty). \tag{6.30}$$

Substituting (6.29) and (6.30) into (6.28) yields (6.25).

Next, we prove (6.26). Using the strong Markov property, we obtain

$$\begin{aligned} E^{(x,k,y,l)}(T^n) &= E^{(x,k,y,l)}((S + \theta_S T)^n) \\ &= \sum_{m=0}^n \frac{n!}{m!(n-m)!} E^{(x,k,y,l)}(S^m (\theta_S T)^{n-m}) \\ &= \sum_{m=0}^n \frac{n!}{m!(n-m)!} E^{(x,k,y,l)}(S^m E^{(X(S),Z(S),Y(S),Z'(S))}(T^{n-m})). \end{aligned} \tag{6.31}$$

Substituting estimates (6.24) and (6.5) into (6.31) yields

$$\begin{aligned} E^{(x,k,y,l)}(T^n) &\leq n! \sum_{m=0}^n \frac{(-2F(\infty))^{n-m}}{\varepsilon_0^m} \\ &\leq n! \sum_{m=0}^n \frac{n!}{m!(n-m)!} \frac{(-2F(\infty))^{n-m}}{\varepsilon_0^m}, \end{aligned}$$

which implies (6.26).

Finally, we prove the strong ergodicity. Setting  $M := \varepsilon_0^{-1} - 2F(\infty)$ , from (6.26) we have, for any  $\lambda < M^{-1}$ ,

$$E^{(x,k,y,l)} e^{\lambda T} = \sum_{n=0}^{\infty} \frac{\lambda^n}{n!} E^{(x,k,y,l)}(T^n) \leq \sum_{n=0}^{\infty} \lambda^n M^n = (1 - \lambda M)^{-1} < \infty. \tag{6.32}$$

Analogously to the proof of Theorem 4.2 of [18], by virtue of Theorem 2.1 of [18] and (6.32), we prove that the Markov process  $(X(t), Z(t))$  is strongly ergodic and that  $\hat{\alpha}(\gamma) \geq M^{-1}(1 - 2/\gamma)$  for  $\gamma > 2$ ; hence, (6.27) holds.

To conclude this section, we provide an example for which Theorem 6.1 can be applied. To do so, we return to the model introduced in Example 1.2.

**Example 6.1.** Take  $K = 2$  and  $N = \{1, 2\}$ , and consider the model introduced in Example 1.2. Let  $\alpha(1) = 1, \alpha(2) = 2, \beta(1) = 2, \beta(2) = 1$ ,

$$Q := (q_{kl}) = \begin{pmatrix} -1 & 1 \\ 2 & -2 \end{pmatrix}, \quad \text{and} \quad \sigma_{11}(x, k) = \sigma_{22}(x, k) \equiv 1$$

for all  $(x, k) \in \mathbb{R}^2 \times \{1, 2\}$ . Now, by (6.2), we first obtain  $\varepsilon_0 = e^{-2}$ . Next,  $\hat{\alpha} = 2$  and  $\kappa = 1$ , and then we can take

$$\gamma(r) = r^2 - \frac{r^4}{16} \quad \text{and} \quad \eta(r) \equiv 4.$$

Hence, we can take

$$C(r) = \exp\left(\int_1^r \frac{\gamma(u)}{u} du\right) = \exp\left(\frac{r^2}{2} - \frac{1}{2} - \frac{r^4}{64} + \frac{1}{64}\right).$$

Therefore, we obtain

$$F(\infty) = -\frac{1}{4} \int_0^\infty \exp\left(\frac{s^4}{64} - \frac{s^2}{2}\right) ds \int_s^\infty \exp\left(\frac{u^2}{2} - \frac{u^4}{64}\right) du > -\infty. \tag{6.33}$$

In summary, by virtue of Theorem 6.1, we conclude that  $(X(t), Z(t))$  defined by the above model is strongly ergodic with

$$\hat{\alpha} \geq [e^2 - 2F(\infty)]^{-1} > 0,$$

where  $F(\infty)$  is calculated using (6.33).

### Acknowledgements

The authors thank the anonymous referee and the Editor for careful readings of this paper, and for their comments and suggestions.

### References

- [1] BASAK, G. K., BISI, A. AND GHOSH, M. K. (1996). Stability of a random diffusion with linear drift. *J. Math. Anal. Appl.* **202**, 604–622.
- [2] CHEN, M.-F. (2004). *Eigenvalues, Inequalities, and Ergodic Theory*. Springer, London.
- [3] CHEN, M.-F. (2004). *From Markov Chains to Non-Equilibrium Particle Systems*, 2nd edn. World Scientific, River Edge, NJ.
- [4] CHEN, M.-F. AND LI, S.-F. (1989). Coupling methods for multidimensional diffusion processes. *Ann. Prob.* **17**, 151–177.
- [5] CHOW, Y. S. AND TEICHER, H. (1978). *Probability Theory*. Springer, New York.
- [6] DAWSON, D. A. (1983). Critical dynamics and fluctuations for a mean-field model of cooperative behavior. *J. Statist. Phys.* **31**, 29–85.
- [7] DAWSON, D. A. AND ZHENG, X. (1991). Law of large numbers and central limit theorem for unbounded jump mean-field models. *Adv. Appl. Math.* **12**, 293–326.
- [8] DOWN, D., MEYN, S. P. AND TWEEDIE, R. L. (1995). Exponential and uniform ergodicity of Markov processes. *Ann. Prob.* **23**, 1671–1691.
- [9] HITSUDA, M. AND MITOMA, I. (1986). Tightness problem and stochastic evolution equation arising from fluctuation phenomena for interacting diffusions. *J. Multivariate Anal.* **19**, 311–328.
- [10] HU, G. (1994). *Stochastic Forces and Nonlinear Systems*. Shanghai Scientific and Technological Education Publishing House, Shanghai (in Chinese).
- [11] ICHIHARA, K. AND KUNITA, H. (1974). A classification of the second order degenerate elliptic operators and its probabilistic characterization. *Z. Wahrscheinlichkeitsth.* **30**, 235–254. (Correction: **39** (1977), 81–84.)
- [12] IKEDA, N. AND WATANABE, S. (1981). *Stochastic Differential Equations and Diffusion Processes*. North-Holland, Amsterdam.
- [13] KHAS'MINSKIĪ, R. Z. (1980). *Stochastic Stability of Differential Equations*. Stijhoff and Noordhoff, Alphen.
- [14] KHAS'MINSKIĪ, R. Z., ZHU, C. AND YIN, G. (2007). Stability of regime-switching diffusions. *Stoch. Process. Appl.* **117**, 1037–1051.
- [15] KLIEMANN, W. (1987). Recurrence and invariant measures for degenerate diffusions. *Ann. Prob.* **15**, 690–707.
- [16] MAO, X. (1999). Stability of stochastic differential equations with Markovian switching. *Stoch. Process. Appl.* **79**, 45–67.
- [17] MAO, Y.-H. (2002). Strong ergodicity for Markov processes by coupling methods. *J. Appl. Prob.* **39**, 839–852.
- [18] MAO, Y.-H. (2006). Convergence rates in strong ergodicity for Markov processes. *Stoch. Process. Appl.* **116**, 1964–1976.
- [19] MEYN, S. P. AND TWEEDIE, R. L. (1992). Stability of Markovian processes. I. Criteria for discrete-time chains. *Adv. Appl. Prob.* **24**, 542–574.
- [20] MEYN, S. P. AND TWEEDIE, R. L. (1993). *Markov Chains and Stochastic Stability*. Springer, London.
- [21] MEYN, S. P. AND TWEEDIE, R. L. (1993). Stability of Markovian processes. II. Continuous-time processes and sampled chains. *Adv. Appl. Prob.* **25**, 487–517.
- [22] MEYN, S. P. AND TWEEDIE, R. L. (1993). Stability of Markovian processes. III. Foster–Lyapunov criteria for continuous-time processes. *Adv. Appl. Prob.* **25**, 518–548.
- [23] SHIGA, T. AND TANAKA, H. (1985). Central limit theorem for a system of Markovian particles with mean field interactions. *Z. Wahrscheinlichkeitsth.* **69**, 439–459.
- [24] SKOROKHOD, A. V. (1989). *Asymptotic Methods in the Theory of Stochastic Differential Equations*. American Mathematical Society, Providence, RI.
- [25] XI, F.-B. (2002). Stability for a random evolution equation with Gaussian perturbation. *J. Math. Anal. Appl.* **272**, 458–472.

- [26] XI, F.-B. (2004). Stability of a random diffusion with nonlinear drift. *Statist. Prob. Lett.* **68**, 273–286.
- [27] XI, F.-B. (2008). Feller property and exponential ergodicity of diffusion processes with state-dependent switching. *Sci. China Ser. A* **51**, 329–342.
- [28] XI, F.-B. AND ZHAO, L.-Q. (2006). On the stability of diffusion processes with state-dependent switching. *Sci. China Ser. A* **49**, 1258–1274.
- [29] YIN, G. G. AND ZHANG, Q. (1998). *Continuous-Time Markov Chains and Applications* (Appl. Math. **37**). Springer, New York.
- [30] YUAN, C. AND LYGEROS, J. (2005). On the exponential stability of switching diffusion processes. *IEEE Trans. Automatic Control* **50**, 1422–1426.
- [31] YUAN, C. AND MAO, X. (2003). Asymptotic stability in distribution of stochastic differential equations with Markovian switching. *Stoch. Process. Appl.* **103**, 277–291.
- [32] ZHU, C. AND YIN, G. (2007). Asymptotic properties of hybrid diffusion systems. *SIAM J. Control Optimization* **46**, 1155–1179.