RAZUMIKHIN-TYPE THEOREMS ON MOMENT EXPONENTIAL STABILITY OF FUNCTIONAL DIFFERENTIAL EQUATIONS INVOLVING TWO-TIME-SCALE MARKOVIAN SWITCHING

Fuke Wu
School of Mathematics and Statistics
Huazhong University of Science and Technology, Wuhan, Hubei 430074, China

George Yin
Department of Mathematics
Wayne State University, Detroit, MI 48202, USA

Le Yi Wang
Department of Electrical and Computer Engineering
Wayne State University, MI 48202, USA

Abstract. This work develops moment exponential stability of functional differential equations (FDEs) with Markovian switching, in which a two-time-scale (real time $t$ and fast time $t/\varepsilon$ with a small parameter $\varepsilon > 0$) continuous-time and finite-state Markov chain is used to represent the switching process. The essence is that there is a time-scale separation, which is motivated by the consideration of networked control systems and manufacturing systems. Under suitable conditions, we establish a Razumikhin-type theorem on the $p$th moment exponential $\varepsilon$-stability for the small parameter $\varepsilon$. By virtue of the Razumikhin-type theorem, we further deduce mean-square exponential stability results for delay differential equations (DDEs) and ordinary differential equations (ODEs) with two-time-scale Markovian switching. These stability results show that the overall system may be stabilized by the Markov switching even when some of the underlying subsystems are unstable. It is noted that in the presence of the Markovian switching, the stationary distribution of the fast changing part of the Markov chain plays an important role. Explicit conditions for the mean-square exponential stability of linear equations are derived and illustrative examples are provided to demonstrate our results.

1. Introduction. Recently, Markovian switching systems driven by continuous-time Markovian chains have drawn growing attention in biological sciences [1,22,35], financial engineering [3, 5–7, 31, 39], communications and manufacturing [32, 38], among others. Such systems have been used to model many practical scenarios in which abrupt changes are experienced in the structure and parameters caused by phenomena such as component failures, communication link interruption, packet loss, and power line contingency.

2010 Mathematics Subject Classification. Primary: 34K50, 60J28, 34F05, 93E15.

Key words and phrases. Functional differential equation, Razumikhin-type theorem, two-time-scale Markov chain, stationary distribution, moment exponential stability.

The research of F. Wu was supported in part by the National Natural Science Foundations of China (No. 11422110 and No. 61473125). The research of G. Yin and L.Y. Wang was supported in part by the Army Research Office under grant W911NF-15-1-0218.
In many applications, stability is one of the major concerns and a great challenge. The stochastic stability analysis was initiated more than fifty years ago [10,12,18,19] and has been substantially developed ever since. Comprehensive stability analysis, methodology exploration, algorithm development, and numerical solutions for random systems have emerged more recently [13,25,36,37]. Because time delays are often unavoidable in practical systems, significant efforts have also been devoted to delay differential systems and more generally functional differential systems, which are exemplified by chemical processes, biological systems, and communication systems, see for example [4,11,17,28]). It is well observed that delays have detrimental impact on stability and performance. Their presence adds substantial difficulties in stability analysis, mainly due to the fact that these systems are described by delay differential equations (DDEs) or functional differential equations (FDEs); see [9,15,16]. When Markovian switching systems with time delays are considered, the stability analysis becomes even more challenging due to the coexistence and interactions of time delays and abrupt changes in system structures and parameters.

Treating systems with delays, Razumikhin methods have been widely used for analyzing stability; see Haddock et al. [8] and Hale [9]. In essence, a Razumikhin-type theorem states that if the derivative of a Liapunov function along all system trajectories is negative whenever the present value of the function dominates that of the past over the delay interval, then the Liapunov function along the system trajectories will converge to zero. In recent years, this theorem has been extensively applied to analyze input-to-state stability (ISS) and input-to-output stability (IOS) of time-delay systems and functional differential systems, as well as to design feedback stabilizers [11,23,28]. In a series of papers, Mao and his colleagues established the Razumikhin-type theorem on moment exponential stability for stochastic DDEs and stochastic FDEs with Markovian switching and Brownian motion [25–27].

When the Markovian switching evolves in a fast pace, certain “averaging” effect takes place [2,14,30,31,40]. Using such an averaging idea, Wu, Yin, and Wang [29] established the Razumikhin-type theorem on the moment exponential stability of systems with random delays modeled by a two-time-scale Markov Chain in which the driving Markov chain has both fast and slow motions and involves strong and weak interactions. It was shown that the “averaged” delay with respect to the stationary distribution of the fast motion plays an important role in asymptotic stability. In particular, the overall system may be stabilized by the Markov switching, even if some of the underlying subsystems are unstable. In [29], the Markovian switching systems with random delays were treated, in which a Markov chain is used to model the random delay. This paper substantially extends our previous results to general switching functional systems in which the Markov chain is embedded in a nonlinear and more general structure. To illustrate, consider a switching FDE of the form

\[ \dot{x}(t) = f(x_t, t, r(t)), \quad t \geq 0, \] (1.1)

where

\[ x_t = x_t(\theta) := \{ x(t + \theta) : -\tau \leq \theta \leq 0 \}, \] (1.2)

where \( r(t) \) is a continuous-time Markov chain with state space \( \mathbb{S} = \{1, 2, \ldots, m\} \), and \( f : C([-\tau, 0]; \mathbb{R}^n) \times \mathbb{R}_+ \times \mathbb{S} \to \mathbb{R}^n \) is Borel measurable. This paper establishes a Razumikhin-type theorem of moment exponential stability for this class of switching systems. The main distinctive features from the existing literature (cf. [25–27]) is
that the fast-switching motion of the Markov chain yields an “averaging” effect with respect to its stationary measure and potentially produces a stabilizing factor.

The investigation on stability of stochastic delay equations goes back to Kushner [18]. Systems perturbed or driven by wide-band noises, possibly non-Markovian, were treated extensively in [20]; and later studied comprehensively for their related control and filtering problems [21]. This paper focuses on stability of functional differential diffusion systems under two-time-scale Markovian switching. Based on this formulation and employing certain technical results on two-time-scale Markovian switching is derived using the Razumikhin-type theorem. Finally, Section 6 presents illustrative examples to demonstrate our results together with some additional remarks.

The rest of the paper is arranged as follows. In the next section, we introduce notation and main assumptions. Section 3 establishes the Razumikhin-type theorem of the $p$th moment exponential $\varepsilon$-stability on FDEs with Markovian switching (1.1). A mean-square exponential $\varepsilon$-stability criterion is presented among others. Applying the Razumikhin-type theorem, Section 4 establishes criteria for the $p$th moment exponential stability of DDEs with two-time-scale Markovian switching. In Section 5, the moment exponential stability of ODEs with two-time-scale Markovian switching is derived using the Razumikhin-type theorem. Finally, Section 6 presents illustrative examples to demonstrate our results together with some additional remarks.

2. FDEs with Markovian switching: A two-time-scale formulation. Unless otherwise specified, we use the following notation throughout the paper. Let $| \cdot |$ be the Euclidean norm in $\mathbb{R}^n$. If $A$ is a vector or matrix, its transpose is denoted by $A'$. If $A$ is a matrix, denote its trace norm by $|A| = \sqrt{\text{tr}(A' A)}$. Let $\mathbb{R}_+ = [0, \infty)$. Denote by $C([-\tau, 0]; \mathbb{R}^n)$ the family of continuous functions from $[-\tau, 0]$ to $\mathbb{R}^n$, which forms a Banach space with the norm $\|\xi\| = \sup_{-\tau \leq \theta \leq 0} |\xi(\theta)|$. $\lambda_{\min}(A)$ and $\lambda_{\max}(A)$ denote the smallest and largest eigenvalues of $A$, respectively. If $a, b \in \mathbb{R}$, let $a \wedge b = \min\{a, b\}$ and $a \vee b = \max\{a, b\}$. Define $a^+ = a \vee 0$ and $a^- = a \wedge 0$. Let $\mathcal{M}$ denote all probability measures on $[-\tau, 0]$. In this paper, these probability measures on $\mathcal{M}$ can be replaced by any right-continuous nondecreasing functions on $[-\tau, 0]$. Let $I_A$ represent the indicator function of the set $A$.

Let $(\Omega, \mathcal{F}, P)$ be a complete probability space with a filtration $\{\mathcal{F}_t\}_{t \geq 0}$ satisfying the usual conditions, that is, it is right continuous and increasing while $\mathcal{F}_0$ contains all $\mathbb{P}$-null sets. Denoted by $C^b_{\mathcal{F}_0}([-\tau, 0]; \mathbb{R}^n)$ the family of all bounded $\mathcal{F}_t$-measurable, $C([-\tau, 0]; \mathbb{R}^n)$-valued random variables. Let $p > 0$ and $L^p_{\mathcal{F}_t}([-\tau, 0]; \mathbb{R}^n)$ be the family of $\mathcal{F}_t$-measurable stochastic processes $\varphi = \{\varphi(\theta) : -\tau \leq \theta \leq 0\}$ such that $\|\varphi\|^p := \sup_{-\tau \leq \theta \leq 0} E|\varphi(\theta)|^p < \infty$. Let $\tilde{r}(t)$, $t \geq 0$ be an continuous-time Markov chain defined on the probability space $(\Omega, \mathcal{F}, P)$ taking values in a finite state space $\mathbb{S} = \{1, 2, \ldots, m\}$ with generator $\tilde{Q} = (\tilde{q}_{ij}) \in \mathbb{R}^{m \times m}$. Recall that the generator $\tilde{Q}$ is weakly irreducible [33], if the system of equations

\[
\begin{cases}
\tilde{\nu} \tilde{Q} = 0 \\
\sum_{i=1}^m \tilde{\nu}_i = 1
\end{cases}
\]  

(2.1)

has a unique solution satisfying $\tilde{\nu}_i \geq 0$ for each $i = 1, \ldots, m$. The solution $\tilde{\nu} = (\tilde{\nu}_1, \tilde{\nu}_2, \ldots, \tilde{\nu}_m)$ is termed quasi-stationary distribution. In what follows, we use $O(y)$ to denote the function of $y$ satisfying $\sup_y |O(y)|/|y| < \infty$. 


In our setup, the Markov chain \( r(t) \) has both fast and slow motions with strong and weak interactions. To reflect the fast and slow motions of the Markov chain, we introduce a small parameter \( \varepsilon > 0 \) and rewrite the Markov chain \( r(t) \) as \( r^\varepsilon(t) \) and the generator \( Q \) as \( Q^\varepsilon \). Thus, the Markov chain displays two-time scales (a usual running time \( t \) and a stretched (fast) time \( t/\varepsilon \)). In the analysis to follow, we choose \( \varepsilon > 0 \) to be simply a small constant. The essence is that there is a time-scale separation motivated by the consideration of networked control systems. Suppose that the generator of the Markov chain is given by

\[
Q^\varepsilon = \frac{Q}{\varepsilon} + Q_0,
\]

where both \( Q \) and \( Q_0 \) are generators of suitable continuous-time Markov chains. Throughout the paper, we assume that \( Q \) is weakly irreducible in the sense defined in (2.1). Denote the quasi-stationary distribution associated with \( Q \) by \( \nu = (\nu_1, \ldots, \nu_m) \). As demonstrated in [33], using asymptotic expansions, one can show that the asymptotic properties of the probability vector \( p^\varepsilon(t) = (P(r^\varepsilon(t) = 1), \ldots, P(r^\varepsilon(t) = m)) \) and the transition probability matrix \( P^\varepsilon(t) = (p^\varepsilon_{ij}(t)) = (P(r^\varepsilon(t + s) = j| r^\varepsilon(s) = i)) \in \mathbb{R}^{m \times m} \) is dominated by \( Q/\varepsilon \). Applying the asymptotic extensions of the probability vector \( p^\varepsilon(t) \) (see [33]) yields

\[
|p^\varepsilon(t) - \nu| \leq O(\varepsilon + \exp\left(-\frac{\kappa t}{\varepsilon}\right)),
\]

where \( \kappa > 0 \) determined by the nonzero eigenvalues of \( Q \).

We are interested in the asymptotic stability of the associated switching FDEs (1.1). To highlight the effect of the time-scale separation, let us rewrite system (1.1) as

\[
\dot{x}^\varepsilon(t) = f(x^\varepsilon_t, t, r^\varepsilon(t)).
\]

In this paper, assume that the initial data \( \xi \in C^1_{\bar{\nu}_{k_0}}([-\tau, 0]; \mathbb{R}^n) \) satisfying the Lipschitz condition, namely, there exists constant \( K \) such that \( |\xi(\theta_1) - \xi(\theta_2)| \leq K|\theta_1 - \theta_2| \) for all \( \theta_1, \theta_2 \in [-\tau, 0] \). This system may be seen as a switching system among the following \( m \) functional differential subsystems

\[
\dot{x}(t) = f(x_t, t, i), \quad i = 1, 2, \ldots, m
\]

with the random switching governed by the Markov chain \( r^\varepsilon(t) \). The switching functional differential system (2.4) may therefore be rewritten as

\[
\dot{x}^\varepsilon(t) = \sum_{i=1}^m f(x^\varepsilon_t, t, i)I_{r^\varepsilon(t) = i}.
\]

Let \( C^{1,1}(\mathbb{R}^n \times \mathbb{R}_+ \times \mathbb{S}; \mathbb{R}_+) \) denote the family of all nonnegative functions \( V \) on \( \mathbb{R}^n \times \mathbb{R}_+ \times \mathbb{S} \) which are continuously differentiable in \( x \in \mathbb{R}^n \) and \( t \in \mathbb{R}_+ \). For a given \( V \in C^{1,1}(\mathbb{R}^n \times \mathbb{R}_+ \times \mathbb{S}; \mathbb{R}_+) \), define an operator \( \mathcal{L} \) from \( C([-\tau, 0]; \mathbb{R}^n) \times \mathbb{R}_+ \times \mathbb{S} \) to \( \mathbb{R} \) by

\[
\mathcal{L}(\varphi, t, i)V(\varphi(0), t, i) = V_t(\varphi(0), t, i) + \nabla V(\varphi(0), t, i)f(\varphi, t, i) + \sum_{j=1}^m q_{ij}V(\varphi(0), t, j)
\]

for any \( \varphi \in C([-\tau, 0]; \mathbb{R}^n) \), where

\[
\nabla V(x, t, i) = \left( \frac{\partial V(x, t, i)}{\partial x_1}, \frac{\partial V(x, t, i)}{\partial x_2}, \ldots, \frac{\partial V(x, t, i)}{\partial x_n} \right).
\]
Remark 2.1. Let us emphasize that $V$ is a function defined on $\mathbb{R}^n \times \mathbb{R}_+ \times \mathbb{S}$, but the operator $\mathcal{L}$ is a functional defined on $C([-\tau, 0]; \mathbb{R}^n) \times \mathbb{R}_+ \times \mathbb{S}$. Note that $\varphi \in C([-\tau, 0]; \mathbb{R}^n)$ is from the coefficient of $f$, not from $V$. Since $\varphi(0)$ is only one point of $\varphi$, for notational convenience, when $\mathcal{L}$ is applied to $V$ we write $\mathcal{L}(\varphi, t, i)V(\varphi(0), t, i)$ as $\mathcal{L}V(\varphi, t, i)$ in what follows. Thus, $\mathcal{L}V$ is only a notation; $V$ on the other hand depends only on $(\varphi(0), t, i)$.

Let $x^\varepsilon(t)$ be the solution of (2.4). According to the chain rule for differentiation,

$$\dot{V}(x^\varepsilon(t), t, r^\varepsilon(t)) = \mathcal{L}V(x^\varepsilon_t(t), t, r^\varepsilon(t)) = \sum_{i=1}^{m} \mathcal{L}V(x^\varepsilon_t, t, i)I_{i(r^\varepsilon(t)=i)}. \quad (2.8)$$

Let us impose the linear growth condition and the local Lipschitz condition on $f$:

**Assumption 2.2** (The linear growth condition). For any $\varphi \in C([-\tau, 0]; \mathbb{R}^n)$, there exist constants $K_i$ such that

$$|f(\varphi, t, i)| \leq K_i \|\varphi\|. \quad (2.9)$$

**Assumption 2.3** (The local Lipschitz condition). For each $j = 1, 2, \ldots$, there exist constant $K_{ij}$ dependent on $j$ and a probability measure $\mu \in \mathcal{M}$ such that for every $i \in \mathbb{S}$,

$$|f(\varphi, t, i) - f(\psi, t, i)| \leq K_{ij} \int_{-\tau}^{0} |\varphi(\theta) - \psi(\theta)| d\mu(\theta) \quad (2.10)$$

for any $\varphi, \psi \in C([-\tau, 0]; \mathbb{R}^n)$ with $\|\varphi\| \vee \|\psi\| \leq j$.

According to the classical result (for example [9,27]), Assumptions 2.2 and 2.3 can guarantee that there exists a unique solution for switching FDE (2.4). Assumption 2.2 implies that $f(0, t, i) = 0$ for all $i \in \mathbb{S}$. This implies that the switching FDE (2.4) admits a trivial solution (or zero solution). Let us now present the definition of the moment exponential $\varepsilon$-stability.

**Definition 2.4.** The trivial solution of the switching FDE (2.4), or simply (2.4) is said to be $p$th moment exponentially $\varepsilon$-stable if there exist positive constants $C, \gamma$ and $\beta$ such that for sufficiently small $\varepsilon$ and any $t \geq 0$,

$$\mathbb{E}|x^\varepsilon(t)|^p \leq Ce^{-\gamma t} + O(\varepsilon^\beta).$$

3. Moment exponential stability. By the Razumikhin-type technique, this section examines the $p$th moment exponential stability for switching FDE with the two-time-scale Markovian switching (2.4).

**Theorem 3.1.** Let $\lambda, p, c_1, c_2$ be all positive numbers and $q > 1$. Under Assumptions 2.2 and 2.3, if there exists a function $V \in C^{1,1}(\mathbb{R}^n \times \mathbb{R}_+ \times \mathbb{S}; \mathbb{R}_+)$ such that the following conditions hold:

(i) for all $(x, t, i) \in \mathbb{R}^n \times [-\tau, \infty) \times \mathbb{S}$,

$$c_1|x|^p \leq V(x, t, i) \leq c_2|x|^p; \quad (3.1)$$

(ii) for all $t \geq 0$ and each $j = 1, 2, \ldots$ and all $i \in \mathbb{S}$, there exist positive constants $H_{ij}$ such that

$$|V(x_1, t, i) - V(x_2, t, i)| \vee |V_i(x_1, t, i) - V_i(x_2, t, i)| \vee |\nabla V(x_1, t, i) - \nabla V(x_2, t, i)| \leq H_{ij}|x_1 - x_2| \quad (3.2)$$

for $|x_1| \vee |x_2| \leq j;
(iii) for all $t \geq 0$,
\[
\sum_{i=1}^{m} \nu_i \mathbb{E}[\mathcal{L}V(\varphi, t, i) + \lambda EV(\varphi(0), t, i)] \leq 0 \tag{3.3}
\]
provided $\varphi = \{\varphi(\theta); -\tau \leq \theta \leq 0\} \in L^p_{\mathcal{F}_t}([-\tau, 0]; \mathbb{R}^n)$ satisfying
\[
\mathbb{E}\left[\min_{i \in \mathcal{S}} V(\varphi(\theta), t + \theta, i)\right] < q \mathbb{E}\left[\max_{i \in \mathcal{S}} V(\varphi(0), t, i)\right] \text{ on } -\tau \leq \theta \leq 0,
\]
then for all initial data $\xi \in C^b_{\mathcal{F}_0}([-\tau, 0]; \mathbb{R}^n)$ and the small parameter $\varepsilon > 0$ and any $\varsigma \in (0, 1/2)$,
\[
\mathbb{E}|x^\varepsilon(t)|^p \leq \frac{C_2}{C_1} \|\xi\|_E^p e^{-\gamma t} + O(\varepsilon^{1-\varsigma}), \tag{3.4}
\]
where
\[
\gamma = \min\left\{\lambda, \frac{\log q}{\tau}\right\}. \tag{3.5}
\]

In other words, switching FDE (2.4) is $p$th moment exponentially $\varepsilon$-stable.

**Remark 3.2.** In \cite{26} or \cite{27}, the stochastic Razumikhin-type theorem required that condition (i) be satisfied and condition (iii) be replaced by (iii’) for all $t \geq 0$,
\[
\mathbb{E}\left[\max_{i \in \mathcal{S}} \mathcal{L}V(\varphi, t, i)\right] + \lambda \mathbb{E}\left[\max_{i \in \mathcal{S}} V(\varphi(0), t, i)\right] \leq 0 \tag{3.6}
\]
provided $\varphi = \{\varphi(\theta); -\tau \leq \theta \leq 0\} \in L^p_{\mathcal{F}_t}([-\tau, 0]; \mathbb{R}^n)$ satisfying
\[
\mathbb{E}\left[\min_{i \in \mathcal{S}} V(\varphi(\theta), t + \theta, i)\right] < q \mathbb{E}\left[\max_{i \in \mathcal{S}} V(\varphi(0), t, i)\right] \text{ on } -\tau \leq \theta \leq 0.
\]
Comparing with (iii’), condition (iii) shows that the asymptotic properties of the Markov chain, namely its stationary distribution, plays an important role to determine stability of the equation, which shows the “average” idea. In fact, (3.6) implies that for all $i \in \mathcal{S}$,
\[
\mathbb{E}[\mathcal{L}V(\varphi, t, i)] + \lambda EV(\varphi(0), t, i) \leq 0,
\]
but (3.3) only requires that the above inequality holds in the sense of “average” with respect to the stationary distribution of the driving Markov chain. This implies that even if some subsystems do not satisfy Razumikhin Theorem, the switching system may be stable, namely, the Markov chain becomes a stabilizing factor.

**Remark 3.3.** Comparing with the result in \cite{9} or \cite{27}, Theorem 3.1 requires that the additional condition (ii) is satisfied. In general, if $V \in C^{2,1}(\mathbb{R}^n \times \mathbb{R}_+ \times \mathcal{S}; \mathbb{R}_+)$, namely $V$ is twice continuously differentiable in $x \in \mathbb{R}^n$, this condition holds.

Let us first present the boundedness for the solution of Eq. (2.4).

**Lemma 3.4.** Under Assumption 2.2, for any initial data $\xi \in C^b_{\mathcal{F}_0}([-\tau, 0]; \mathbb{R})$ and $T > 0$, there exists a constant $N_T$ dependent on $T$ such that
\[
\sup_{-\tau \leq t \leq T} |x^\varepsilon(t)| \leq N_T.
\]

**Proof.** We rewrite (2.4) as
\[
x^\varepsilon(t) = x^\varepsilon(0) + \int_0^t f(x^\varepsilon_s, s, r^\varepsilon(s))ds,
\]
where $x^\varepsilon(t) = (x^\varepsilon_1(t), \ldots, x^\varepsilon_m(t))$. By the boundedness of $f$, if $\varepsilon > 0$ is sufficiently small, then for all $t \geq 0$,
\[
|f(x^\varepsilon(t), s, r^\varepsilon(s))| \leq C(1 + |x^\varepsilon(t)|^p + |s|^p + |r^\varepsilon(s)|^p),
\]
where $C$ is a constant. Therefore, by the boundedness of $x^\varepsilon(0)$ and the fact that $x^\varepsilon(t)$ is a solution of (2.4), we have
\[
\sup_{-\tau \leq t \leq T} |x^\varepsilon(t)| \leq N_T.
\]
which, together with the linear growth condition, implies that

\[ |x^e(t)| \leq |\xi(0)| + \int_0^t |f(x^e_s, s, \tau^e(s))|ds \]

\[ \leq |\xi(0)| + K_{\max} \int_0^t \|x^e_s\|ds \]

\[ \leq |\xi(0)| + K_{\max} \int_0^t \sup_{-\tau \leq \theta \leq 0} |x^e(s + \theta)|ds, \]

where \( K_{\max} = \max_{i\in S} \{K_i\} \). For any \( t_1 \in (0, T] \), we have

\[ \sup_{-\tau \leq t \leq t_1} |x^e(t)| \leq 2\|\xi\| + K_{\max} \int_0^{t_1} \sup_{-\tau \leq \theta \leq 0} |x^e(s + \theta)|ds. \]

Applying the Gronwall inequality yields

\[ \sup_{-\tau \leq t \leq t_1} |x^e(t)| \leq 2\|\xi\|e^{K_{\max}t_1} \leq 2\|\xi\|e^{K_{\max}T}. \]

Choosing \( t_1 = T \) gives the desired assertion. \( \square \)

Let \( x^e(t) \) be the solution of the switching functional system (2.4). For \( i \in S \) and any \( t \in [0, T] \), define a sequence of weighted occupation measures \( Z^e_i(t) \) as follows:

\[ Z^e_i(t) = \int_0^t L(x^e_s, s, i)[I_{\{\tau^e(s)=i\}} - \nu_i]ds. \] (3.7)

Let us establish the following lemma on convergence of \( Z^e_i(t) \):

**Lemma 3.5.** Let Assumptions 2.2 and 2.3 hold and condition (ii) in Theorem 3.1 be satisfied. For any \( \zeta \in (0, 1) \) sufficiently small,

\[ \sup_{0 \leq t \leq T} \mathbb{E}[Z^e_i(t)]^2 = O(\varepsilon^{1-\zeta}). \] (3.8)

This proof can be found in Appendix A.

For \( V \in C^{1,1}(\mathbb{R}^n \times \mathbb{R}_+ \times S; \mathbb{R}_+) \), let us define another sequence of weighted occupation measures \( X^e_i(t) \) as

\[ X^e_i(t) = \int_0^t V(x^e(t), t, i)[I_{\{\tau^e(s)=i\}} - \nu_i]ds. \] (3.9)

Similar to Lemma 3.5, we can establish the following lemma on convergence of \( X^e_i(t) \):

**Lemma 3.6.** Let Assumptions 2.2 and 2.3 hold. Then there exists \( \zeta \in (0, 1) \) such that

\[ \sup_{0 \leq t \leq T} \mathbb{E}[X^e_i(t)]^2 = O(\varepsilon^{1-\zeta}). \] (3.10)

This proof is similar to Lemma 3.5, so it is omitted.
Proof of Theorem 3.1. If we can prove that for sufficiently small $\varepsilon$,
\[
e^{-\gamma t}[\mathbb{E}[x^{\varepsilon}(t)]^{p} - O(\varepsilon^{\frac{1}{2}-\gamma})] \leq \frac{c_{2}}{c_{1}}\|\xi\|_{E}^{p} =: \kappa
\]
for any $\bar{\gamma} \in (0, \gamma)$, the proof is complete. By (3.1), it follows from
\[
\hat{W}^{\varepsilon}(t) := e^{\bar{\gamma}t}[\mathbb{E}[x^{\varepsilon}(t), t, r^{\varepsilon}(t)] - c_{1}O(\varepsilon^{\frac{1}{2}-\gamma})] \leq \kappa_{c_{1}}.
\tag{3.11}
\]
When $t \in [-\tau, 0]$,
\[
\hat{W}^{\varepsilon}(t) \leq \mathbb{E}[x^{\varepsilon}(t), t, r^{\varepsilon}(t)] - c_{1}e^{\bar{\gamma}t}O(\varepsilon^{\frac{1}{2}-\gamma}) \leq \mathbb{E}(\xi(t), t, r^{\varepsilon}(t)) \leq c_{2}\|\xi\|_{E}^{p} = \kappa_{c_{1}}.
\]
Then for sufficiently small $\varepsilon$, we claim (3.11) holds for all $t \geq 0$. Otherwise, for any given $\varepsilon > 0$ sufficiently small, by the continuity of $\hat{W}^{\varepsilon}(t)$, there exists the smallest $\rho^{\varepsilon} \in [0, \infty)$ such that for all $t \in [-\tau, \rho^{\varepsilon}]$, $\hat{W}^{\varepsilon}(t) < \kappa_{c_{1}}$ and $\hat{W}^{\varepsilon}(\rho^{\varepsilon}) = \kappa_{c_{1}}$ as well as $\hat{W}^{\varepsilon}(\rho^{\varepsilon} + \delta) > \hat{W}^{\varepsilon}(\rho^{\varepsilon})$ for all sufficiently small $\delta$. Then for any $t \in [\rho^{\varepsilon} - \tau, \rho^{\varepsilon}]$,
\[
\mathbb{E}[x^{\varepsilon}(t), t, r^{\varepsilon}(t)] - c_{1}O(\varepsilon^{\frac{1}{2}-\gamma}) = e^{-\bar{\gamma}t}\hat{W}^{\varepsilon}(t)
\leq e^{-\bar{\gamma}t}[\mathbb{E}(\xi(t), t, r^{\varepsilon}(t)) - c_{1}O(\varepsilon^{\frac{1}{2}-\gamma})]
\leq e^{-\bar{\gamma}t}\mathbb{E}[x^{\varepsilon}(\rho^{\varepsilon}), \rho^{\varepsilon}, r^{\varepsilon}(\rho^{\varepsilon})] - c_{1}e^{\bar{\gamma}(\rho^{\varepsilon} - t)}O(\varepsilon^{\frac{1}{2}-\gamma}),
\]
which implies that
\[
\mathbb{E}[x^{\varepsilon}(t), t, r^{\varepsilon}(t)] \leq e^{-\bar{\gamma}t}\mathbb{E}(x^{\varepsilon}(\rho^{\varepsilon}), \rho^{\varepsilon}, r^{\varepsilon}(\rho^{\varepsilon}))
\tag{3.12}
\]
since $e^{\bar{\gamma}(\rho^{\varepsilon} - t)} \geq 1$. If $\mathbb{E}(x^{\varepsilon}(\rho^{\varepsilon}), \rho^{\varepsilon}, r^{\varepsilon}(\rho^{\varepsilon})) = 0$, then $\mathbb{E}[x^{\varepsilon}(t), t, r^{\varepsilon}(t)] = 0$ for all $t \in [\rho^{\varepsilon} - \tau, \rho^{\varepsilon}]$. By condition (3.1),
\[
\mathbb{E}[x^{\varepsilon}(t)]^{p} \leq \frac{1}{c_{1}}E\left[\min_{i \in \mathcal{S}} V(x^{\varepsilon}(t), t, i)\right]
\leq \frac{1}{c_{1}}\mathbb{E}(x^{\varepsilon}(t), t, r^{\varepsilon}(t)) = 0,
\]
which implies that $x^{\varepsilon}(t) = 0$ a.s. for all $t \in [\rho^{\varepsilon} - \tau, \rho^{\varepsilon}]$. By the existence and uniqueness of the solution, $x^{\varepsilon}(t) = 0$ a.s. for all $t \geq 0$. This contradicts the definition of $\rho^{\varepsilon}$. Therefore, we have
\[
\mathbb{E}(x^{\varepsilon}(\rho^{\varepsilon}), \rho^{\varepsilon}, r^{\varepsilon}(\rho^{\varepsilon})) > 0.
\]
By $\bar{\gamma} < (\log q)/\tau$, (3.12) implies that for all $t \in [\rho^{\varepsilon} - \tau, \rho^{\varepsilon}]$,
\[
\mathbb{E}[x^{\varepsilon}(t), t, r^{\varepsilon}(t)] < q\mathbb{E}(x^{\varepsilon}(\rho^{\varepsilon}), \rho^{\varepsilon}, r^{\varepsilon}(\rho^{\varepsilon}))
\]
which is equivalent to
\[
\mathbb{E}(x^{\varepsilon}(\rho^{\varepsilon} + \theta), \rho^{\varepsilon} + \theta, r^{\varepsilon}(\rho^{\varepsilon} + \theta)) < q\mathbb{E}(x^{\varepsilon}(\rho^{\varepsilon}), \rho^{\varepsilon}, r^{\varepsilon}(\rho^{\varepsilon}))
\]
for any $\theta \in [-\tau, 0]$. Consequently, for any $\theta \in [-\tau, 0]$,
\[
\mathbb{E}\left[\min_{i \in \mathcal{S}} V(x^{\varepsilon}(\rho^{\varepsilon} + \theta), \rho^{\varepsilon} + \theta, i)\right]
\leq \mathbb{E}(x^{\varepsilon}(\rho^{\varepsilon} + \theta), \rho^{\varepsilon} + \theta, r^{\varepsilon}(\rho^{\varepsilon} + \theta))
< q\mathbb{E}(x^{\varepsilon}(\rho^{\varepsilon}), \rho^{\varepsilon}, r^{\varepsilon}(\rho^{\varepsilon}))
\leq q\mathbb{E}\left[\max_{i \in \mathcal{S}} V(x^{\varepsilon}(\rho^{\varepsilon}), \rho^{\varepsilon}, i)\right].
\]
Applying condition (3.3) yields
\[
\sum_{i=1}^{m} \nu_i \left[ E \mathcal{LV}(x_{\rho^+}^\varepsilon, \rho^+, i) + \lambda E V(x(\rho^+), \rho^+, i) \right] \leq 0.
\]
Noting that \( \tilde{\gamma} < \gamma \leq \lambda \), we have
\[
\sum_{i=1}^{m} \nu_i [ E \mathcal{LV}(x_{\rho^+}^\varepsilon, \rho^+, i) + \tilde{\gamma} E V(x(\rho^+), \rho^+, i) ] < 0,
\] (3.13)
which implies that for sufficiently small \( \delta > 0 \), for any \( t \in [\rho^+, \rho^+ + \delta] \),
\[
\sum_{i=1}^{m} \nu_i [ E \mathcal{LV}(x_t^\varepsilon, t, i) + \tilde{\gamma} E V(x(t), t, i) ] < 0.
\] (3.14)

It follows from this that
\[
\tilde{W}^\varepsilon(\rho^+ + \delta) - \tilde{W}^\varepsilon(\rho^+) = \left[ e^{\tilde{\gamma}(\rho^+ + \delta)} E V(x^\varepsilon(\rho^+ + \delta), \rho^+ + \delta, x^\varepsilon(\rho^+ + \delta)) - e^{\tilde{\gamma}\rho^+} E V(x^\varepsilon(\rho^+), \rho^+, x^\varepsilon(\rho^+)) \right]
- c_1 e^{\tilde{\gamma}\rho^+} (e^{\tilde{\gamma} \delta} - 1) O(\varepsilon^{\frac{1}{2} - \varsigma})
\]
\[
= E \int_{\rho^+}^{\rho^+ + \delta} e^{\tilde{\gamma} t} [ E \mathcal{LV}(x_t^\varepsilon, t, r^\varepsilon(t)) + \tilde{\gamma} V(x^\varepsilon(t), t, r^\varepsilon(t))] dt - c_1 e^{\tilde{\gamma}\rho^+} (e^{\tilde{\gamma} \delta} - 1) O(\varepsilon^{\frac{1}{2} - \varsigma})
\]
\[
= \int_{\rho^+}^{\rho^+ + \delta} e^{\tilde{\gamma} t} \sum_{i=1}^{m} \nu_i [ E \mathcal{LV}(x_t^\varepsilon, t, i) + \tilde{\gamma} E V(x^\varepsilon(t), t, i) ] dt
\]
\[
+ E \int_{\rho^+}^{\rho^+ + \delta} e^{\tilde{\gamma} t} \sum_{i=1}^{m} \mathcal{LV}(x_t^\varepsilon, t, i)(I_{\{ r^\varepsilon(t) = i \}} - \nu_i) dt
\]
\[
+ \tilde{\gamma} E \int_{\rho^+}^{\rho^+ + \delta} e^{\tilde{\gamma} t} \sum_{i=1}^{m} V(x^\varepsilon(t), t, i)(I_{\{ r^\varepsilon(t) = i \}} - \nu_i) dt - c_1 e^{\tilde{\gamma}\rho^+} (e^{\tilde{\gamma} \delta} - 1) O(\varepsilon^{\frac{1}{2} - \varsigma})
\]
\[
< \sum_{i=1}^{m} E \int_{\rho^+}^{\rho^+ + \delta} e^{\tilde{\gamma} t} \mathcal{LV}(x_t^\varepsilon, t, i)(I_{\{ r^\varepsilon(t) = i \}} - \nu_i) dt
\]
\[
+ \tilde{\gamma} \sum_{i=1}^{m} E \int_{\rho^+}^{\rho^+ + \delta} e^{\tilde{\gamma} t} V(x^\varepsilon(t), t, i)(I_{\{ r^\varepsilon(t) = i \}} - \nu_i) dt - c_1 e^{\tilde{\gamma}\rho^+} (e^{\tilde{\gamma} \delta} - 1) O(\varepsilon^{\frac{1}{2} - \varsigma}).
\] (3.16)

Letting \( \varsigma = \varsigma/2 \) and applying Lemmas 3.5 and 3.6 lead to
\[
E \int_{\rho^+}^{\rho^+ + \delta} e^{\tilde{\gamma} t} \mathcal{LV}(x_t^\varepsilon, t, i)(I_{\{ r^\varepsilon(t) = i \}} - \nu_i) dt = O(\varepsilon^{\frac{1}{2} - \varsigma}),
\] (3.17)
and
\[
E \int_{\rho^+}^{\rho^+ + \delta} e^{\tilde{\gamma} t} V(x^\varepsilon(t), t, i)(I_{\{ r^\varepsilon(t) = i \}} - \nu_i) dt = O(\varepsilon^{\frac{1}{2} - \varsigma}).
\] (3.18)

Substituting (3.17) and (3.18) into (3.15) yields
\[
\tilde{W}^\varepsilon(\rho^+ + \delta) - \tilde{W}^\varepsilon(\rho^+) < O(\varepsilon^{\frac{1}{2} - \varsigma}).
\]
Noting that \( \varepsilon \) is sufficiently small, we therefore have
\[
\tilde{W}^\varepsilon(\rho^+ + \delta) - \tilde{W}^\varepsilon(\rho^+) \leq 0,
\] (3.19)
which contradicts the definition of $\rho^\varepsilon$. Thus, (3.11) holds for all $t \geq 0$. This completes this proof.

Clearly, in Theorem 3.1, condition (3.3) plays a crucial rule to determine the moment stability. Nevertheless, it is not convenient to check this condition because it is not related to the coefficient of $f$ explicitly. In what follows, we impose some conditions on $f$ to guarantee Theorem 3.1. These conditions show the roles of coefficient $f$ as well as the Markov switching $r^\varepsilon(t)$ in the moment stability of switching FDE (2.4) for sufficiently small $\varepsilon > 0$.

**Corollary 3.7.** Under Assumptions 2.2 and 2.3, assume that there exist constants $\zeta_i, \sigma_i$ and a probability measure $\eta \in \mathcal{M}$ such that for any $\varphi \in C([-\tau, 0]; \mathbb{R}^n)$ and $t \geq 0$,

$$
\varphi'(0)f(\varphi, t, i) \leq -\zeta_i|\varphi(0)|^2 + \sigma_i \int_{-\tau}^0 |\varphi(\theta)|^2 d\eta(\theta). \tag{3.20}
$$

If $\bar{\zeta} > \bar{\sigma} \vee 0$, where $\bar{x} = \sum_{i=1}^m \nu_i x_i$ for $x_i = \zeta_i, \sigma_i$, then for any $\xi \in (0, 1/2)$,

$$
E|x^\varepsilon(t)|^2 \leq \|\xi\|^2 e^{-\gamma_1 t} + O(\varepsilon^{2-\xi}),
$$

where

$$
\gamma_1 = \max_{q>1} \left\{ 2(\bar{\zeta} - q\bar{\sigma} \vee 0) \wedge \frac{\log q}{\tau} \right\}.
$$

That is, the trivial solution of the system (2.4) is mean-square exponentially $\varepsilon$-stable.

**Proof.** Choose $V(x) = |x|^2$. It is readily seen that conditions (i) and (ii) of Theorem 3.1 hold. By the definition of $LV$ and (3.20),

$$
LV(\varphi, t, i) \leq -2\zeta_i|\varphi(t)|^2 + 2\sigma_i \int_{-\tau}^0 |\varphi(\theta)|^2 d\eta(\theta), \tag{3.21}
$$

which shows that

$$
\sum_{i=1}^n \nu_i ELV(\varphi, t, i) \leq -2\bar{\zeta}E|\varphi(0)|^2 + 2(2\bar{\sigma} \vee 0) \left( \sup_{-\tau \leq \theta \leq 0} E|\varphi(\theta)|^2 \right). \tag{3.22}
$$

Recalling that $\bar{\zeta} > \bar{\sigma} \vee 0$, we can choose $q > 1$ such that $\bar{\zeta} > q\bar{\sigma} \vee 0$. Hence, if $E|\varphi(\theta)|^2 < qE|\varphi(0)|^2$ for $\theta \in [-\tau, 0]$, which implies that

$$
E[\min_{i \in S} V(\varphi(\theta), t + \theta, i)] < qE[\max_{i \in S} V(\varphi(0), t, i)]
$$
on $-\tau \leq \theta \leq 0$ since $V$ is independent of $t$ and $r^\varepsilon(t)$, (3.22) implies that

$$
\sum_{i=1}^m \nu_i ELV(\varphi, t, i) \leq -2(\bar{\zeta} - q\bar{\sigma} \vee 0)E|\varphi(0)|^2.
$$

This shows that condition (iii) holds. Applying Theorem 3.1 gives the desired result.

**Remark 3.8.** The conditions in Theorem 3.1 are given in terms of the function $V$, so they are not easily verifiable. Condition (3.20) in Corollary 3.20 is given in terms of the function $f$. As a result, stability can be checked by using $f$. Comparing these two results, Theorem 3.1 is a general one, whereas Corollary 3.7 considers a special case.
4. DDEs with Markovian switching. As a class of special cases of FDEs with Markovian switching, the Markovian switching DDEs have broad applications. By the moment stability theorem established above on switching FDEs, this section examines the moment exponential stability of switching DDEs with multiple delays. Let us consider the DDE with Markovian switching of the form

\[ \dot{x}(t) = F(x(\tau_1), \ldots, x(\tau_k), t, r(t)) \quad (4.1) \]

on \( t \geq 0 \) with initial data \( x_0 = \xi \in C_{\mathscr{P}}^{b}([0,0]; \mathbb{R}^n) \) satisfying the Lipschitz condition, where \( \tau = \max\{\tau_1, \ldots, \tau_k\} \), \( F : \mathbb{R}^n \times \mathbb{R}^{n \times k} \times \mathbb{R}_+ \times \mathcal{S} \to \mathbb{R}^n \), is a Borel measurable function. Define \( \mu_0, \mu_1, \ldots, \mu_k \) as the Dirac measures in \( 0, \tau_1, \tau_2, \ldots, \tau_k \).

Let us choose

\[ d\mu = \frac{1}{k}(d\mu_1 + d\mu_2 + \cdots + d\mu_k). \]

It is clear that \( \mu \) is a probability measure. For (4.1), Assumption 2.3 may be specialized as follows.

**Assumption 4.1.** For each \( j = 1, 2, \ldots \), there exists constant \( K_{ij} \) dependent on \( j \) such that \( t \geq 0 \) and \( i \in \mathcal{S} \), there exist constants \( K_i \) such that

\[ |F(x_0, \ldots, x_k, t, i) - F(y_0, \ldots, y_k, t, i)| \leq K_{ij} \sum_{j=0}^{k} |x_j - y_j| \quad (4.2) \]

for any \( x_0, \ldots, x_k, y_0, \ldots, y_k \in \mathbb{R}^n \) with \( |x_0| |x_2| \cdots |x_k| |y_0| |y_2| \cdots |y_k| \leq j \).

The linear growth condition may be rewritten in the following way.

**Assumption 4.2.** For any \( x_0, \ldots, x_k \in \mathbb{R}^n \), there exist constants \( K_i \) such that

\[ F(x_0, \ldots, x_k, t, i) \leq K_i (|x_0| + \cdots |x_k|). \quad (4.3) \]

This implies that \( F(0, \ldots, 0, t, i) = 0 \), so there exists a trivial solution for Eq. (4.1). Moreover, the operator (2.7) becomes

\[ L\mathcal{V}(x_0, \ldots, x_k, t, i) = V_t(x_0, t, i) + \nabla \mathcal{V}(x_0, t, i)F(x_0, \ldots, x_k, t, i) + \sum_{j=1}^{m} q_{ij} V(x_0, t, j) \quad (4.4) \]

for any \( x_0, \ldots, x_k \in \mathbb{R}^n \), \( t \geq 0 \) and \( i \in \mathcal{S} \).

Under Assumptions 4.1 and 4.2, Eq. (4.1) has a unique solution and the results in Lemma 3.4 hold. Then by using Theorem 3.1, let us establish the moment exponential stability.

**Theorem 4.3.** Let \( \lambda, p, c_1, c_2 \) be all positive numbers and \( q > 1 \). Under Assumptions 4.1 and 4.2, if there exists a function \( \mathcal{V} \in C^{1,1}(\mathbb{R}^n \times \mathbb{R}_+ \times \mathcal{S}; \mathbb{R}_+) \) such that

(i) and (ii) in Theorem 3.1 holds and the following condition is satisfied:
Proof. For any \( q \) where
\[
\sum_{i=1}^{m} \nu_i \mathbb{E}[\mathcal{L}V(x_0, \ldots, x_k, t, i) + \lambda V(x_0, t, i)] \leq 0
\]
(4.5)
provided \( x_0, \ldots, x_k \in \mathbb{R}^n \) satisfying for all \( \tau_j \), \( j = 1, \ldots, k \),
\[
\mathbb{E}\left[ \min_{i \in \mathbb{S}} V(x_j, t - \tau_j, i) \right] < q \mathbb{E}\left[ \max_{i \in \mathbb{S}} V(x_0, t, i) \right],
\]
then for all initial data \( \xi \in C^0_{\mathbb{F}_T}([-\tau, 0]; \mathbb{R}^n) \) and small parameter \( \varepsilon > 0 \), Eq. (4.1) holds the property
\[
\mathbb{E}[x^\varepsilon(t)]^p \leq \frac{C_2}{c_1} \|\xi\|_p^p e^{-\gamma t} + O(\varepsilon^{1/2}),
\]
(4.6)
where
\[
\gamma = \min \left\{ \lambda, \frac{\log q}{\tau} \right\},
\]
namely, Eq. (4.1) is \( p \)th moment exponentially \( \varepsilon \)-stable.

Proof. For \( \varphi \in C([-\tau, 0]; \mathbb{R}^n) \) and \( t \geq 0 \), define \( f(\varphi, t, i) = F(\varphi(0), \varphi(-\tau_1), \ldots, \varphi(-\tau_k), t, i) \). Then (4.1) becomes (2.4). It is clear that condition (iii) implies that condition (iii) of Theorem 3.1 holds. Applying Theorem 3.1 gives the desired result.

By this theorem, we may establish the following corollary.

Corollary 4.4. Let \( \lambda_0, \lambda_1, \ldots, \lambda_k, p, c_1, c_2 \) be all positive numbers with
\[
\lambda_0 > \sum_{j=1}^{k} \lambda_j.
\]
(4.7)

Under Assumptions 4.1 and 4.2, if there exists a function \( V \in C^{1,1}(\mathbb{R}^n \times [0, T] \times \mathbb{S}; \mathbb{R}_+) \) such that conditions (i) and (ii) of Theorem 4.3 are satisfied and moreover, for any \( x_0, \ldots, x_k \in \mathbb{R}^n \) and \( t \geq 0 \),
\[
\sum_{i=1}^{m} \nu_i \mathcal{L}V(x_0, \ldots, x_k, t, i) \leq -\lambda_0 \left( \max_{i \in \mathbb{S}} V(x_0, t, i) \right) + \sum_{j=1}^{k} \lambda_j \left( \min_{i \in \mathbb{S}} V(x_j, t, i) \right),
\]
(4.8)
then, for the solution of (4.1),
\[
\mathbb{E}[x^\varepsilon(t)]^p \leq \frac{C_2}{c_1} \|\xi\|_p^p e^{-\left(\lambda_0 - q \sum_{j=1}^{k} \lambda_j\right)t} + O(\varepsilon^{1/2}),
\]
(4.9)
where \( q \in (1, \frac{\lambda_0}{\sum_{j=1}^{k} \lambda_j}) \) is the unique root of \( \lambda_0 - q \sum_{j=1}^{k} \lambda_j = (\log q)/\tau \).

Proof. For any \( t \geq 0 \) and \( x_0, \ldots, x_k \in \mathbb{R}^n \) satisfying
\[
\mathbb{E}\left[ \min_{i \in \mathbb{S}} V(x_j, t - \tau_j, i) \right] < q \mathbb{E}\left[ \max_{i \in \mathbb{S}} V(x_0, t, i) \right], \quad \text{for } j = 1, \ldots, k,
\]
it follows from (4.8) that
\[
\sum_{i=1}^{m} \nu_i \mathbb{E}[\mathcal{L}V(x_0, \ldots, x_k, t, i)]
\]
\[
\leq -\lambda_0 \mathbb{E}\left( \max_{i \in \mathbb{S}} V(x_0, t, i) \right) + \sum_{j=1}^{k} \lambda_j \mathbb{E}\left( \min_{i \in \mathbb{S}} V(x_j, t - \tau_j, i) \right)
\]
If there exist positive constants \( \beta \) and \( \zeta \) such that \[ 2 \bar{\zeta} = \sum_{i=1}^{m} \nu_i \beta_i \zeta_i, \quad \bar{K}_\beta = \sum_{i=1}^{m} \nu_i \beta_i K_i \quad \text{and} \quad \bar{q}_\beta = \sum_{i=1}^{m} \sum_{j=1}^{m} \nu_i q_{ij} \beta_j. \] Then we may present the following corollary, which may illustrate the “averaging” idea on the stability of stochastic switching equations with multiple delays.

**Corollary 4.5.** Let Assumptions 4.1 and 4.2 hold. Assume that there exist constant \( \zeta_i \) such that

\[ x'F(x,0,\ldots,0,t,i) \leq -\zeta_i |x|^2 \quad \text{for all} \ x \in \mathbb{R}^n \quad \text{and} \quad t \geq 0. \]  

If there exist positive constants \( \beta_i, \ i = 1,\ldots,m \) such that

\[ \frac{2 \bar{\zeta} - \bar{K}_\beta - \bar{q}_\beta}{\max_{i \in S} \{ \beta_i \}} > \frac{2k \bar{K}_\beta}{\min_{i \in S} \{ \beta_i \}}, \]  

then for any initial data \( \xi \in C([-\tau,0]; \mathbb{R}^n) \) and sufficiently small \( \varepsilon > 0, \)

\[ E|x^\varepsilon(t)|^2 \leq \| \xi \|^2 e^{-(\bar{\lambda}_0-q\lambda)t} + O(\varepsilon^{1-\varepsilon}), \]  

where \( q \in (1,\bar{\lambda}_0/\hat{\lambda}_0) \) is the unique root of the equation \( \bar{\lambda}_0 - q\lambda = \log q/\tau, \) where

\[ \bar{\lambda}_0 = \frac{2 \bar{\zeta} - \bar{K}_\beta - \bar{q}_\beta}{\max_{i \in S} \{ \beta_i \}}, \quad \hat{\lambda} = \frac{2k \bar{K}_\beta}{\min_{i \in S} \{ \beta_i \}}. \]

That is, \( x(t) \), the solution of (4.1) is mean square exponentially \( \varepsilon \)-stable.

**Proof.** Choose \( V(x,i) = \beta_i |x|^2 \). By (4.4),

\[
\mathcal{L}V(x_0,\ldots,x_k,t,i) = 2 \beta_i x_0' F(x_0,\ldots,x_k,t,i) + \sum_{j=1}^{m} q_{ij} \beta_j |x_0|^2
\]

\[
= 2 \beta_i x_0' F(x_0,0,\ldots,0,t,i) + \sum_{j=1}^{m} q_{ij} \beta_j |x_0|^2
\]

\[
+ 2 \beta_i x_0' [ F(x_0,x_1,\ldots,x_k,t,i) - F(x_0,0,\ldots,0,t,i) ].
\]

This, together with (4.2), (4.3), and (4.10) gives that for any \( \eta_1 > 0, \)

\[
\mathcal{L}V(x^\varepsilon(t),x^\varepsilon(t-\tau_1),\ldots,x^\varepsilon(t-\tau_k),t,i)
\]
Choosing $\eta_1 = K_1$, we therefore have

$$
\sum_{i=1}^{m} \nu_i \mathcal{L}V(x^\varepsilon(t), x^\varepsilon(t - \tau_1), \ldots, x^\varepsilon(t - \tau_k), t, i)
\leq -\frac{m}{\nu_1} \left[ \beta_i (2 \zeta_i - K_i) - \beta_i q_{ij} \right] |x^\varepsilon(t)|^2 + 2 \sum_{i=1}^{m} \nu_i \beta_i \sum_{j=1}^{k} |x^\varepsilon(t - \tau_j)|^2
\leq -(2 \zeta_i - \bar{K}_\beta - \bar{q}_\beta) |x^\varepsilon(t)|^2 + 2 \bar{K}_\beta \sum_{j=1}^{k} |x^\varepsilon(t - \tau_j)|^2.
$$

By the definition of $V(x, i)$,

$$
\max_{i \in \mathcal{S}} \{V(x, i)\} = \max_{i \in \mathcal{S}} \{ \beta_i \} |x|^2, \quad \min_{i \in \mathcal{S}} \{V(x, i)\} = \min_{i \in \mathcal{S}} \{ \beta_i \} |x|^2.
$$

Thus, we have

$$
\sum_{i=1}^{m} \nu_i \mathcal{L}V(x^\varepsilon(t), x^\varepsilon(t - \tau_1), \ldots, x^\varepsilon(t - \tau_k), t, i)
\leq -\frac{2 \zeta_i - \bar{K}_\beta - \bar{q}_\beta}{\max_{i \in \mathcal{S}} \{ \beta_i \}} \max_{i \in \mathcal{S}} \{V(x^\varepsilon(t), i)\} + \frac{2 \bar{K}_\beta}{\min_{i \in \mathcal{S}} \{ \beta_i \}} \sum_{j=1}^{k} \min_{i \in \mathcal{S}} \{V(x^\varepsilon(t - \tau_j), i)\},
$$

By condition (4.11), applying Corollary 4.4 yields the desired result. □

**Remark 4.6.** In this corollary, condition (4.11) is very important. This condition is expressed as the average value with respect to the stationary distribution of the fast-varying part of Markov chain $r^\varepsilon(t)$ (with weight $\beta_i$). This implies that both the switching Lyapunov function and the stationary distribution will have crucial effect on the stability result.

5. **ODEs with Markovian switching.** Letting $\tau = 0$, Eq. (2.4) or (4.1) becomes the ordinary differential equation (ODE) with Markovian switching of the form

$$
\dot{x}^\varepsilon(t) = f(x^\varepsilon(t), t, r^\varepsilon(t)), t \geq 0,
$$

where $f$ is a function from $\mathbb{R}^n \times \mathbb{R}^+ \times \mathcal{S}$ to $\mathbb{R}^n$, and the initial value $\xi$ becomes a bounded $\mathcal{F}_0$-measurable $\mathbb{R}^n$-valued random variable. Letting $\mu \in \mathcal{M}$ be the Dirac measure in the origin, the local Lipschitz assumption (2.10) becomes

$$
|f(x, t, i) - f(y, t, i)| \leq K_{ij} |x - y|
$$

(5.2)
for $x, y \in \mathbb{R}^n$ with $|x| \lor |y| < j$, where $K_{ij}$ is a constant dependent on $j$. Correspondingly, (2.7) becomes a function from $\mathbb{R}^n \times \mathbb{R}_+ \times S$ to $\mathbb{R}$ of the form

$$LV(x, t, i) = V_t(x, t, i) + \nabla V(x, t, i)f(x, t, i) + \sum_{j=1}^{m} q_{ij} V(x, t, j).$$

(5.3)

The following result follows directly from Theorem 3.1 or Theorem 4.3.

**Theorem 5.1.** Let $\lambda, p, c_1, c_2$ be all positive numbers and $q > 1$. Under condition (5.2), if there exists a function $V \in C^{1,1}(\mathbb{R}^n \times \mathbb{R}_+ \times S; \mathbb{R}_+)$ such that conditions (i) and (ii) in Theorem 3.1 hold and the following condition is satisfied:

(iii) for $t \geq 0$ and all $x \in \mathbb{R}^n$,

$$\sum_{i=1}^{m} \nu_i [LV(x, t, i) + \lambda V(x, t, i)] \leq 0,$$

then for all bounded initial value $\xi \in \mathbb{R}^n$, (5.1) is $p$th moment exponentially $\varepsilon$-stable and the $p$th moment Liapunov exponent is not greater than $-\lambda$.

**Remark 5.2.** In general, for $\lambda > 0$, $LV(x, t, i) + \lambda V(x, t, i) \leq 0$ for some $i \in S$ may guarantee that the $i$th subsystem is stable and $LV(x, t, i) - \lambda V(x, t, i) > 0$ implies that the $i$th system is unstable. condition (iii) in this theorem shows that, even if some subsystems are unstable under the condition $LV(x, t, i) - \lambda V(x, t, i) > 0$, the switching system may still be stable if condition (iii) holds. That is, the Markov chain $r^\varepsilon(t)$ may serve as a stabilizing factor for sufficiently small $\varepsilon > 0$.

To illustrate, let us consider the $n$-dimensional linear differential equation with Markovian switching

$$\dot{x}^\varepsilon(t) = A(r^\varepsilon(t))x^\varepsilon(t),$$

(5.4)

where $A : S \to \mathbb{R}^{n \times n}$. This equation is a random switching system among the following $m$ subsystems

$$\dot{y}(t) = A(i)y(t), \quad i = 1, 2, \ldots, m.$$

Let $A_i = A(i)$. It is well-known that the $i$th subsystem is exponentially stable if $A_i + A_i'$ is negative definite and it is unstable if $A_i + A_i'$ is positive definite. Choose $V(x) = |x|^2$. Then we have

$$LV(x, t, i) = x'(A_i + A_i')x,$$

which implies that

$$\sum_{i=1}^{m} \nu_i LV(x, t, i) \leq \tilde{\lambda}|x|^2 = \tilde{\lambda}V(x),$$

where $\tilde{\lambda} = \sum_{i=1}^{m} \nu_i \lambda_{\max}(A_i + A_i')$. If

$$\tilde{\lambda} = \sum_{i=1}^{m} \nu_i \lambda_{\max}(A_i + A_i') < 0,$$

(5.5)

then condition (iii) of Theorem 5.1 is satisfied. Applying Theorem 5.1 gives the following corollary.

**Corollary 5.3.** If condition (5.5) holds, the linear equation (5.4) is mean-square exponentially $\varepsilon$-stable with the Liapunov exponent less than $-\tilde{\lambda}$. 

Note that $\lambda_{\text{max}}(A_i + A_i') < 0$ shows that $A_i + A_i'$ is negative definite. By Corollary 5.3, even if $A_i + A_i'$ is positive definite for some $i = 1, 2, \ldots, m$, the switching system (5.4) may still be mean-square exponentially stable if condition (5.5) holds. This shows that if even not all subsystems are stable, it is possible that the switched system may be stable. Namely, the Markov chain may be a stabilizing factor for small parameter $\varepsilon$.

6. Examples and remarks. The model for the Markov chain given in (2.2) can be generalized to

$$Q^\varepsilon = \frac{Q}{\varepsilon} + \frac{Q_0}{\varepsilon^{\gamma_0}}, \quad 0 \leq \gamma_0 < 1,$$

(6.1)

where both $Q$ and $Q_0$ are generators of suitable continuous-time Markov chains. Here, $Q/\varepsilon$ represents the fast-varying part and $Q_0/\varepsilon^{\gamma_0}$ represents the slow-changing part. All the results obtained before hold true in this system. To proceed, we present three examples next, in which we use (6.1) with $\gamma_0 = 0$.

6.1. Illustrative examples.

Example 6.1. Let $r^\varepsilon(t)$ be a continuous-time Markov chain generated $Q_\varepsilon = Q/\varepsilon + Q_0$ with

$$Q = \begin{bmatrix} -0.2 & 0.2 \\ 0.8 & -0.8 \end{bmatrix} \quad \text{and} \quad Q_0 = \begin{bmatrix} -0.2 & 0.2 \\ 0.1 & -0.1 \end{bmatrix},$$

(6.2)

and its state space $S = \{1, 2\}$. The stationary distribution associated with $Q$ is given by $\nu = (0.8, 0.2)$. Consider the 2-dimensional SDE with Markovian switching

$$\dot{x}^\varepsilon(t) = A(r(t))x^\varepsilon(t),$$

(6.3)

where $A(i) : \mathbb{R}^2 \to \mathbb{R}^{2\times2}$ for $i = 1, 2$ with

$$A_1 = A(1) = \begin{pmatrix} -3 \\ 2\sqrt{2} - 1 \end{pmatrix}, \quad \text{and} \quad A_2 = A(2) = \begin{pmatrix} 1 & 2 \\ -2 & 1 \end{pmatrix}.$$

The system (6.3) may be seen as a switching system between the two subsystems

$$\dot{x}_1(t) = A_1x_1(t)$$

(6.4)

and

$$\dot{x}_2(t) = A_2x_2(t)$$

(6.5)

according to the Markov chain $r^\varepsilon(t)$. Since $\lambda_{\text{max}}(A_1 + A_1') = -2 < 0$, namely, $A_1 + A_1'$ is negative definite, the subsystem (6.4) is exponentially stable with the Liapunov exponent less than $-2$. Note that $\lambda_{\text{min}}(A_2 + A_2') = 2 > 0$, so $A_2 + A_2'$ is positive definite. This shows that the subsystem (6.5) is unstable.

Note that $\lambda_{\text{max}}(A_2 + A_2') = 2$. Therefore, we have

$$\lambda = 0.8\lambda_{\text{max}}(A_1 + A_1') + 0.2\lambda_{\text{max}}(A_2 + A_2') = 0.8 \times (-2) + 0.2 \times 2 = -1.2. \hspace{1cm} (6.6)$$

Applying Corollary 5.3 gives that the switching system (6.3) is mean-square exponentially $\varepsilon$-stable with the Liapunov exponent less than $-1.2$ for small parameter $\varepsilon > 0$ although the subsystem (6.5) is not stable.

Example 6.2. Consider a scalar linear Volterra delay-integro-differential equations with the Markovian Switching of the form

$$\dot{x}^\varepsilon(t) = \alpha(r^\varepsilon(t))x^\varepsilon(t) + \beta(r^\varepsilon(t)) \int_{-1}^{0} x^\varepsilon(t + \theta)d\theta$$

(6.7)
with the initial data \( \xi \in C([-1,0]; \mathbb{R}) \), where \( r^\varepsilon(t) \) is the same as Example 6.1, \( \alpha_1 = \alpha(1) = 1, \sigma_2 = \alpha(2) = -2 \), \( \beta_1 = \beta(1) = 1, \beta_2 = \beta(2) = 0.5 \). System (6.7) may be considered as a switching system between the two subsystems

\[
\dot{x}_1(t) = x_1(t) + \int_{-1}^{0} x_1(t + \theta) d\theta 
\]

and

\[
\dot{x}_2(t) = -2x_2(t) + 0.5 \int_{-1}^{0} x_2(t + \theta) d\theta 
\]

according to the Markov chain \( r^\varepsilon(t) \). We know that (6.8) is not stable while (6.9) is. We shall show that the switching system (6.7) is mean-square exponentially stable.

Define

\[
f(\varphi, t, 1) = \varphi(0) + \int_{-1}^{0} \varphi(\theta) d\mu(\theta), \quad f(\varphi, t, 2) = -2\varphi(0) + 0.5 \int_{-1}^{0} \varphi(\theta) d\mu(\theta),
\]

where \( \mu \) is the uniform distribution on \([-1,0]\), then the switching system (6.7) becomes (2.4). It is easy to test that

\[
\varphi(0)f(\varphi, t, 1) = \varphi^2(0) + \varphi(0) \int_{-1}^{0} \varphi(\theta) d\mu(\theta) \leq 1.5\varphi^2(0) + 0.5 \int_{-1}^{0} \varphi^2(\theta) d\mu(\theta),
\]

\[
\varphi(0)f(\varphi, t, 2) = -2\varphi^2(0) + 0.5\varphi(0) \int_{-1}^{0} \varphi(\theta) d\mu(\theta) \leq -1.75\varphi^2(0) + 0.25 \int_{-1}^{0} \varphi^2(\theta) d\mu(\theta),
\]

which implies that condition (3.20) holds with \( \zeta_1 = -1.5, \zeta_2 = 1.75 \) and \( \sigma_1 = 0.5, \sigma_2 = 0.25 \). The simple computation gives \( \tilde{\sigma} = 1.5 \times 0.2 + 1.75 \times 0.8 = 1.1 > 0.5 \times 0.2 + 0.25 \times 0.8 = 0.3 \). Applying Corollary 3.7 gives that switching equation (6.7) is mean-square exponentially \( \varepsilon \)-stable with the Liapunov exponent less than \( \max_{q \geq 1}\{ (2.2 - 0.6q) \wedge \log q \} \). If we choose \( q > 1 \) to be the unique solution of \( 2.2 - 0.6q = \log q \), then \( q = 2.2875 \) and the stable exponent is less than \(-0.8275\). This shows that for small parameter \( \varepsilon > 0 \), switching equation (6.7) is mean-square exponentially \( \varepsilon \)-stable although the subsystem (6.9) is not stable.

### 6.2. Remarks

This paper has focused on stability of functional differential equations with Markov switching. The Razumikhin-type techniques state that if the derivative of a Liapunov function along system trajectories is negative whenever the current value of the function dominates its recent history of the delay interval, then the Liapunov function along the trajectories will converge to zero. In this work, we have established the Razumikhin-type theorem on exponential \( \varepsilon \)-stability in the sense of the \( p \)th moment for functional and delay differential systems with two-time-scale Markov chain as the switching factor. This result implies that the Markov chain may serve as a stochastic stabilizing factor. In this stochastic stabilizing process, the rapid-switching part plays an important role.

### Appendix A: Proof of Lemma 3.5

For any \( \zeta \in (0, 1) \) sufficiently and \( t \in (0, T) \), let \( N = \lfloor t/\varepsilon^{1-\zeta} \rfloor \), where \( \lfloor a \rfloor \) denotes the integer part of \( a \). Use a partition of \([0, t]\) given by

\[
[0, t] = [t_0, t_1] \cup [t_1, t_2] \cup [t_2, t_3] \cup \cdots \cup [t_N, t_{N+1}],
\]
where \( t_k = \varepsilon^{1-\zeta} k \) for \( k = 0, 1, \ldots, N \) and \( t_{N+1} = t \). Define a piecewise functional:

\[
L_t = \begin{cases} \mathcal{L}V(x_0^t, 0, i), & \text{if } 0 \leq t < t_2; \\ \mathcal{L}V(x_{t_{k-1}}^t, t_{k-1}, i), & \text{if } t_k \leq t < t_{k+1}; \\ \mathcal{L}V(x_{t_{N-1}}^t, t_{N-1}, i), & \text{if } t = t_{N+1}. 
\end{cases}
\]

In view of the elementary inequality \((a+b)^2 \leq 2a^2 + 2b^2\),

\[
\mathbb{E}|Z(t)|^2 \leq 2\mathbb{E}\left( \int_t^0 |\mathcal{L}V(x^s, t, i) - L_s| |I_{t_{\tau}}(s) = i) - \nu_i| ds \right)^2
+ 2\mathbb{E}\left( \int_t^0 L_s |I_{t_{\tau}}(s) = i) - \nu_i| ds \right). 
\] (A.1)

The Cauchy-Schwarz inequality yields that for any \( t \in [0, T]\),

\[
\mathbb{E}\left( \int_0^t |\mathcal{L}V(x^s, s, i) - L_s| I_{t_{\tau}}(s) = i) - \nu_i| ds \right)^2 \leq T \int_0^t \mathbb{E}(\mathcal{L}V(x^s, s, i) - L_s)^2 ds.
\]

By the definition of \( L_s \), for \( t \in [0, t_2] \),

\[
\int_0^t \mathbb{E}(\mathcal{L}V(x^s, s, i) - L_s)^2 ds = \int_0^t \mathbb{E}(\mathcal{L}V(x^s, s, i) - \mathcal{L}V(x_0^s, 0, i))^2 ds.
\]

According to Lemma 3.4, there exists an integer \( j \) such that \( |x^s(t)| \leq N_T \leq j \) for all \( t \in [-\tau, T] \). This, together with Assumption 2.3 and condition (ii) in Theorem 3.1, implies that there exists constant \( H_{ij} \) such that

\[
|\mathcal{L}V(x^s, s, i) - \mathcal{L}V(x_0^s, 0, i)|^2 \leq H_{ij} \left[ |x^s(s) - \xi(0)|^2 + \int_{-\tau}^0 |x^s(s + \theta) - \xi(\theta)|^2 d\mu(\theta) \right].
\] (A.2)

We therefore have

\[
\int_0^t \mathbb{E}(\mathcal{L}V(x^s, s, i) - \mathcal{L}V(x_0^s, 0, i))^2 ds
\leq H_{ij} \left[ \int_0^t \mathbb{E}|x^s(s) - \xi(0)|^2 ds + \int_{-\tau}^0 \mathbb{E}|x^s(s + \theta) - \xi(\theta)|^2 d\mu(\theta) ds \right]. (A.3)
\]

According to the linear growth condition and Lemma 3.4,

\[
\int_0^t \mathbb{E}|x^s(s) - \xi(0)|^2 ds = \int_0^t \mathbb{E} \left| \int_0^s f(x^u, u, \tau(u)) du \right|^2 ds
\leq \int_0^t \mathbb{E} K_{\text{max}} \int_0^s \|x^u_0\|^2 ds
\leq j^2 K_{\text{max}}^2 \int_0^t s^2 ds
\leq \frac{1}{3} j^2 K_{\text{max}}^2 t^2
= O(t^{3-3\zeta}),
\] (A.4)
where \( K_{\text{max}} \) is defined in Lemma 3.4. It is obvious that
\[
\int_0^t \int_{-\tau}^0 \mathbb{E}|x^\varepsilon(s + \theta) - \xi(\theta)|^2 d\mu(\theta) d\tau \leq K \int_{-\tau}^0 s^2 d\mu(\theta)
\]
where \( D_1 = \{(s, \theta) : s \in [0, t], \theta \in [-\tau, 0], s + \theta < 0\} \) and \( D_2 = \{(s, \theta) : s \in [0, t], \theta \in [-\tau, 0], s + \theta \geq 0\} \). Recall that the initial data \( \xi \in C^0([\gamma, 0]; \mathbb{R}) \) satisfies the Lipschitz condition. There exists constant \( K \) such that
\[
\int_{D_1} \mathbb{E}(\xi(s + \theta) - \xi(\theta))^2 d\mu(\theta) ds \leq K^2 \int_{D_1} s^2 d\mu(\theta) ds \leq K^2 \int_0^t s^2 ds \int_{-\tau}^0 d\mu(\theta) \leq \frac{K^2 \tau^3}{3} = O(\varepsilon^{3-3\zeta}). \tag{A.6}
\]
Noting that \( s + \theta \geq 0 \) for \( (s, \theta) \in D_2 \), we have
\[
\int_{D_2} \mathbb{E}(\xi(s + \theta) - \xi(\theta))^2 d\mu(\theta) ds \leq 2 \int_{D_2} \mathbb{E}(x^\varepsilon(s + \theta) - x^\varepsilon(0))^2 d\mu(\theta) ds + 2 \int_{D_2} \mathbb{E}(\xi(0) - \xi(\theta))^2 d\mu(\theta) ds. \tag{A.7}
\]
Since \( 0 \leq -\theta \leq s \) for \( (s, \theta) \in D_2 \), similar to the computation of (A.6),
\[
\int_{D_2} \mathbb{E}(\xi(0) - \xi(\theta))^2 d\mu(\theta) ds = O(\varepsilon^{3-3\zeta}). \tag{A.8}
\]
Noting that \( 0 \leq s + \theta \leq s \) for \( (s, \theta) \in D_2 \), by the linear growth condition and Lemma 3.4, we have
\[
\int_{D_2} \mathbb{E}(x^\varepsilon(s + \theta) - x^\varepsilon(0))^2 d\mu(\theta) ds = \int_{D_2} \mathbb{E} \left| \int_0^{s+\theta} \dot{x}^\varepsilon(\zeta) d\zeta \right|^2 d\mu(\theta) ds = \int_{D_2} \mathbb{E} \left| \int_0^{s+\theta} f(x^\varepsilon(\zeta), r(\zeta)) d\zeta \right|^2 d\mu(\theta) ds \leq K_{\text{max}}^2 \int_{D_2} \mathbb{E} \left| \int_0^{s+\theta} \|x^\varepsilon\| d\zeta \right|^2 d\mu(\theta) ds \leq j^2 K_{\text{max}}^2 \int_{D_2} (s + \theta)^2 d\mu(\theta) ds \leq j^2 K_{\text{max}}^2 \int_0^s s^2 ds \int_{-\tau}^0 d\mu(\theta) \leq \frac{1}{3} j^2 K_{\text{max}}^2 t^3 = O(\varepsilon^{3-3\zeta}). \tag{A.9}
\]
Substituting (A.8) and (A.9) into (A.7) followed by substituting (A.6) and (A.7) into (A.5) yield
\[
\int_0^t \int_{-\tau}^0 \mathbb{E}|x^\varepsilon(s + \theta) - \xi(\theta)|^2 d\mu(\theta) d\tau = O(\varepsilon^{3-3\zeta}).
\]
This, together with (A.4) implies that for \( t \in [0, t_2] \),
\[
\int_0^t \mathbb{E}[\mathcal{L}V(x_s^x, s, i) - \mathcal{L}V(x_0^x, 0, i)]^2 ds = O(\varepsilon^{3-3\zeta}). \tag{A.10}
\]
By Lemma 3.4 and Assumption 2.3, we have
\[
\int_0^t \mathbb{E}[\mathcal{L}V(x_s^x, s, i) - L_s]^2 ds
= \sum_{k=2}^{N} \int_{t_k}^{t_{k+1}} \mathbb{E}[f(x_s^x, s, i) - f(x_{t_{k-1}}^x, t_{k-1}, i)]^2 ds + O(\varepsilon^{3-3\zeta})
\leq K_{ij}^2 \sum_{k=2}^{N} \int_{t_k}^{t_{k+1}} \int_{-\tau}^{0} \mathbb{E}[x^\varepsilon(s + \theta) - x^\varepsilon(t_{k-1} + \theta)]^2 d\mu(\theta) ds + O(\varepsilon^{3-3\zeta}). \tag{A.11}
\]
Now let us estimate \( \int_{t_k}^{t_{k+1}} \int_{-\tau}^{0} \mathbb{E}[x^\varepsilon(s + \theta) - x^\varepsilon(t_{k-1} + \theta)]^2 d\mu(\theta) ds \) for \( k \geq 2 \). If \( t_{k-1} \geq \tau \), similar to the computation of (A.9),
\[
\int_{t_k}^{t_{k+1}} \int_{-\tau}^{0} \mathbb{E}[x^\varepsilon(s + \theta) - x^\varepsilon(t_{k-1} + \theta)]^2 d\mu(\theta) ds
= \int_{t_k}^{t_{k+1}} \int_{-\tau}^{0} \mathbb{E}\left[ \int_{t_{k-1} + \theta}^{t_k + \theta} x^\varepsilon(\zeta) d\zeta \right]^2 d\mu(\theta) ds
= \int_{t_k}^{t_{k+1}} \int_{-\tau}^{0} \mathbb{E}\left[ \int_{t_{k-1} + \theta}^{t_k + \theta} f(x_s^x, s, r(\zeta)) d\zeta \right]^2 d\mu(\theta) ds
\leq j^2 K_{max}^2 \int_{t_k}^{t_{k+1}} (s - t_{k-1})^2 ds
\leq j^2 K_{max}^2 (t_{k+1} - t_{k-1})^3
= O(\varepsilon^{3-3\zeta}). \tag{A.12}
\]
If \( t_{k-1} < \tau \),
\[
\int_{t_k}^{t_{k+1}} \int_{-\tau}^{0} \mathbb{E}[x^\varepsilon(s + \theta) - x^\varepsilon(t_{k-1} + \theta)]^2 d\mu(\theta) ds
= \int_{t_k}^{t_{k+1}} \int_{-\tau}^{t_{k-1}} \mathbb{E}[x^\varepsilon(s + \theta) - \xi(t_{k-1} + \theta)]^2 d\mu(\theta) ds
+ \int_{t_k}^{t_{k+1}} \int_{-\tau}^{0} \mathbb{E}[x^\varepsilon(s + \theta) - x^\varepsilon(t_{k-1} + \theta)]^2 d\mu(\theta) ds.
\]
Similar to the computation of (A.12),
\[
\int_{t_k}^{t_{k+1}} \int_{-\tau}^{0} \mathbb{E}[x^\varepsilon(s + \theta) - x^\varepsilon(t_{k-1} + \theta)]^2 d\mu(\theta) ds = O(\varepsilon^{3-3\zeta}). \tag{A.13}
\]
Define \( D_3 = \{(s, \theta) : s \in [t_k, t_{k+1}], \theta \in [-\tau, -t_{k-1}], s + \theta \geq 0\} \) and \( D_4 = \{(s, \theta) : s \in [t_k, t_{k+1}], \theta \in [-\tau, -t_{k-1}], s + \theta < 0\} \). Then
\[
\int_{t_k}^{t_{k+1}} \int_{-\tau}^{-t_{k-1}} \mathbb{E}[x^\varepsilon(s + \theta) - \xi(t_{k-1} + \theta)]^2 d\mu(\theta) ds
\]

\[ \int_{D_3} \mathbb{E}|x^\varepsilon(s + \theta) - \varepsilon(t_{k-1} + \theta)|^2 d\mu(\theta) ds + \int_{D_4} \mathbb{E}|\xi(s + \theta) - \xi(t_{k-1} + \theta)|^2 d\mu(\theta) ds. \]

Similar to the computation of (A.6),
\[ \int_{D_3} \mathbb{E}|\xi(s + \theta) - \xi(t_{k-1} + \theta)|^2 d\mu(\theta) ds = O(\varepsilon^{3-3\zeta}) \tag{A.14} \]
and similar to (A.7),
\[ \int_{D_4} \mathbb{E}|x^\varepsilon(s + \theta) - x^\varepsilon(t_{k-1} + \theta)|^2 d\mu(\theta) ds = O(\varepsilon^{3-3\zeta}). \tag{A.15} \]

(A.13), (A.14), and (A.15) show that if \( t_{k-1} < \tau \),
\[ \int_{t_k}^{t_{k+1}} \int_{-\tau}^{0} \mathbb{E}|x^\varepsilon(s + \theta) - x^\varepsilon(t_{k-1} + \theta)|^2 d\mu(\theta) ds = O(\varepsilon^{3-3\zeta}). \]

This, together with (A.12), gives
\[ \int_{t_k}^{t_{k+1}} \int_{-\tau}^{0} \mathbb{E}|\xi(s + \theta) - \xi(t_{k-1} + \theta)|^2 d\mu(\theta) ds = O(\varepsilon^{3-3\zeta}). \]

Substituting the above expression into (A.11) gives
\[ \int_{0}^{t} \mathbb{E}[\mathcal{L}V(x^\varepsilon, s, i) - L_s]^2 ds = \sum_{k=0}^{N} O(\varepsilon^{3-3\zeta}) = O(\varepsilon^{2-2\zeta}). \tag{A.16} \]

Let us now estimate the second term of (A.1). Denote
\[ \eta^\varepsilon(t) = \mathbb{E}\left[ \int_{0}^{t} L_s[I_{\{r^\varepsilon(s) = i\}} - \nu_i] ds \right]^2. \]

Note that for all \( t \in [0, T] \), \( L_t \) is a bounded functional. Then the derivative of \( \eta^\varepsilon(t) \) is given by
\[ \frac{d\eta^\varepsilon(t)}{dt} = 2\mathbb{E} \int_{0}^{t} L_s L_t (I_{\{r^\varepsilon(t) = i\}} - \nu_i) (I_{\{r^\varepsilon(s) = i\}} - \nu_i) ds. \]

Since \( L_t L_s (I_{\{r^\varepsilon(t) = i\}} - \nu_i) (I_{\{r^\varepsilon(s) = i\}} - \nu_i) \) is bounded for any \( s, t \in [0, T] \), there must exist a constant \( K_1 \) such that
\[ \mathbb{E} \int_{0}^{t} L_s L_t (I_{\{r^\varepsilon(t) = i\}} - \nu_i) (I_{\{r^\varepsilon(s) = i\}} - \nu_i) ds = K_1 t \leq K_1 t_2 = O(\varepsilon^{1-\zeta}) \]
for any \( t \in [0, t_2] \). If \( t \in [t_k, t_{k+1}] \) for \( k \geq 2 \), then using the same argument,
\[ \mathbb{E} \int_{t_k}^{t} L_s L_t (I_{\{r^\varepsilon(t) = i\}} - \nu_i) (I_{\{r^\varepsilon(s) = i\}} - \nu_i) ds = O(\varepsilon^{1-\zeta}). \]

Hence we have
\[ \frac{d\eta^\varepsilon(t)}{dt} = 2 \int_{0}^{t_{k-1}} \mathbb{E}[L_s L_t |I_{\{r^\varepsilon(t) = i\}} - \nu_i] |I_{\{r^\varepsilon(s) = i\}} - \nu_i|] ds + O(\varepsilon^{1-\zeta}). \tag{A.17} \]

Define \( \mathcal{F}^\varepsilon_t = \sigma\{x^\varepsilon(s), 0 \leq s \leq t\} \). When \( t \in [t_k, t_{k+1}) \), according the definition of \( L_t \), \( L_t \) is \( \mathcal{F}^\varepsilon_{t-1} \)-measurable. For any \( s_{k-1} < s \leq t_{k-1} < t_k \leq t < t_{k+1} \), using the
asymptotic expansion of the probability vector of \( r^\varepsilon(t) \) (see [33, Lemma 5.1, p81]), there exists a constant \( \kappa \) such that

\[
\mathbb{E}(L_s L_t [I_{\{r^\varepsilon(s) = i\}} - \nu_i][I_{\{r^\varepsilon(t) = i\}} - \nu_i]) \\
= \mathbb{E}(L_s L_t [I_{\{r^\varepsilon(s) = i\}} - \nu_i]\mathbb{E}([I_{\{r^\varepsilon(t) = i\}} - \nu_i]|\mathcal{F}^\varepsilon_{t_{k-1}}]) \\
= \mathbb{E}[L_s L_t [I_{\{r^\varepsilon(s) = i\}} - \nu_i]O \left( \varepsilon + e^{\frac{(T - tk - 1)}{\varepsilon}} \right)] \\
= \mathbb{E}[L_s L_t [I_{\{r^\varepsilon(s) = i\}} - \nu_i]O \left( \varepsilon + e^{\frac{(T - tk - 1)}{\varepsilon}} \right)] \\
= \mathbb{E}[L_s L_t [I_{\{r^\varepsilon(s) = i\}} - \nu_i]O(\varepsilon) \\
= O(\varepsilon).
\]

This together with (A.17) gives

\[
\frac{d\hat{\eta}^\varepsilon(t)}{dt} = O(\varepsilon^{1-\zeta}), \quad (A.18)
\]

which holds uniformly on \([0, T]\). Equation (A.18) together with \( \hat{\eta}^\varepsilon(0) = 0 \) yields

\[
\sup_{0 \leq t \leq T} \eta^\varepsilon(t) = \sup_{0 \leq t \leq T} \int_0^t \left( \frac{d\hat{\eta}^\varepsilon(s)}{ds} \right) ds = O(\varepsilon^{1-\zeta}). \quad (A.19)
\]

Combining (A.19) with (A.16) gives \( \sup_{0 \leq t \leq T} \mathbb{E}[Z^\varepsilon(t)]^2 = O(\varepsilon^{1-\zeta}) \), as required.

REFERENCES

[1] M. Abundo and L. Caramellino, Some remarks on a Markov chain modelling cooperative biological systems, *Open Systems & Information Dynamics*, 3 (1995), 325–343.

[2] G. Badowski and G. Yin, Stability of hybrid dynamic systems containing singularly perturbed random processes, *IEEE Transactions on Automatic Control*, 47 (2002), 2021–2032.

[3] G. Barone-Adesi and R. Whaley, Efficient analytic approximation of American option values, *Journal of Finance*, 42 (1987), 301–320.

[4] M. Barrio, K. Burrage, A. Leier and T. Tian, Oscillatory regulation of Hes1: Discrete stochastic delay modelling and simulation, *PLoS Computational Biology*, 2 (2006), 1017–1039.

[5] A. Bensoussan, Z. Yan and G. Yin, Threshold-type policies for real options using regime-switching models, *SIAM J. Financial Math.*, 3 (2012), 667–689.

[6] G. B. Di Masi, Y. M. Kabanov and W. J. Runggaldier, Mean variance hedging of options on stocks with Markov volatility, *Theory of Probability and Applications*, 39 (1994), 172–182.

[7] P. Chen, H. Yang and G. Yin, Markowitz’s mean-variance asset-liability management with regime switching: A continuous-time model, *Insurance: Mathematics and Economics*, 43 (2008), 456–465.

[8] J. R. Haddock, T. Krisztin, J. Terjéki and J. H. Wu, An invariance principle of Lyapunov–Razumikhin type for neutral functional differential equations, *Journal of Differential Equations*, 107 (1994), 395–417.

[9] J. K. Hale and S. M. V. Lunel, *Introduction to Functional Differential Equations*, Springer-Verlag, New York, 1993.

[10] I. I. Kac and N. N. Krasovskii, On the stability of systems with random parameters, *J. Appl. Math. Mech.*, 24 (1960), 1225–1246.

[11] I. Karafyllis, P. Pepe and Z. P. Jiang, Input-to-output stability for systems described by retarded functional differential equations, *European Journal of Control*, 14 (2008), 539–555.

[12] R. Z. Hasminskii, *Stochastic Stability of Differential Equations*, Sijithoff and Noordhoff, Alphen a/d Rijn, 1980.

[13] R. Z. Hasminskii, C. Zhu and G. Yin, Stability of regime-switching diffusions, *Stochastic Processes and Their Applications*, 117 (2007), 1037–1051.

[14] R. Z. Hasminskii and G. Yin, On averaging principle: An asymptotic expansion approach, *SIAM Journal on Mathematical Analysis*, 35 (2004), 1534–1560.
[15] V. B. Kolmanovskii and V. R. Nosov, *Stability of Functional Differential Equations*, Academic Press, New York, 1986.
[16] V. B. Kolmanovskii and A. Myshkis, *Applied Theory of Functional Differential Equations*, Dordrecht, The Netherlands: Kluwer, 1992.
[17] Y. Kuang, *Delay Differential Equations with Applications in Population Dynamics*, Academic Press, Boston, 1993.
[18] H. J. Kushner, *On the stability of processes defined by stochastic difference-differential equations*, *Journal of Differential Equations*, 4 (1968), 424–443.
[19] H. J. Kushner, *Stability and existence of diffusions with discontinuous or rapidly growing drift terms*, *Journal of Differential Equations*, 11 (1972), 156–168.
[20] H. J. Kushner, *Approximation and Weak Convergence Methods for Random Processes, with Applications to Stochastic Systems Theory*, MIT Press, Cambridge, MA, 1984.
[21] H. J. Kushner, *Weak Convergence Methods and Singularly Perturbed Stochastic Control and Filtering Problems*, Birkhäuser, Boston, MA, 1990.
[22] Q. Luo and X. Mao, *Stochastic population dynamics under regime switching II*, *Journal of Mathematical Analysis and Applications*, 355 (2009), 577–593.
[23] M. Jankovic, Control Lyapunov-Razumikhin functions and robust stabilization of time delay systems, *IEEE Transactions on Automatic Control*, 46 (2001), 1048–1060.
[24] X. Mao, *Stochastic Differential Equations and Applications*, Horwood, Chichester, 1997.
[25] X. Mao, *Stability of stochastic differential equations with Markovian switching*, Stochastic Processes and their Applications, 79 (1999), 45–67.
[26] X. Mao, Stochastic functional differential equations with Markovian switching, *Functional Differential Equations*, 6 (1999), 375–396.
[27] X. Mao and C. Yuan, *Stochastic Differential Equations with Markovian Switching*, London, Imperial College Press, 2006.
[28] A. R. Teel, Connections between Razumikhin-type theorems and the ISS nonlinear small gain theorem, *IEEE Transactions on Automatic Control*, 43 (1998), 960–964.
[29] F. Wu, G. Yin and L. Y. Wang, Moment exponential stability of random delay systems with two-time-scale Markovian switching, *Nonlinear Analysis Series B: Real World Applications*, 13 (2012), 2476–2490.
[30] F. Wu, G. Yin and L. Y. Wang, Stability of a Pure Random Delay System with Two-Time-Scale Markovian Switching, *Journal of Differential Equations*, 253 (2012), 878–905.
[31] G. Yin, Asymptotic expansions of option price under regime-switching diffusions with a fast-varying switching process, *Asymptotic Analysis*, 65 (2009), 203–222.
[32] G. Yin, S. Kan, L. Y. Wang and C. Xu, Identification of systems with regime switching and unmodeled dynamics, *IEEE Transaction on Automatical Control*, 54 (2009), 34–47.
[33] G. Yin and Q. Zhang, *Continuous-time Markov Chains and Applications: A Singular Perturbation Approach*, Springer-Verlag, New York, 1998.
[34] G. Yin, G. Zhao and F. Wu, Regularization and stabilization of randomly switching dynamic systems, *SIAM J. Appl. Math.*, 72 (2012), 1361–1382.
[35] G. Yin and C. Zhu, *Hybrid Switching Diffusions: Properties and Applications*, Springer-Verlag, New York, 2010.
[36] C. Yuan and X. Mao, Asymptotic stability in distribution of stochastic differential equations with Markovian switching, *Stochastic Processes and their Applications*, 103 (2003), 277–291.
[37] C. Yuan and G. Yin, Stability of hybrid stochastic delay systems whose discrete components have a large state space: A two-time-scale approach, *Journal of Mathematical Analysis and Applications*, 368 (2010), 103–119.
[38] Q. Zhang, Hybrid filtering for linear systems with non-Gaussian disturbances, *IEEE Transaction on Automatical Control*, 45 (2000), 50–61.
[39] X. Y. Zhou and G. Yin, Markowitz’s mean-variance portfolio selection with regime switching: A continuous-time model, *SIAM Journal on Control and Optimization*, 42 (2003), 1466–1482.
[40] C. Zhu, G. Yin and Q. S. Song, Stability of random-switching systems of differential equations, *Quarterly of Applied Mathematics*, 67 (2009), 201–220.

Received March 2014; revised June 2014.

E-mail address: wufuke@mail.hust.edu.cn
E-mail address: gyin@math.wayne.edu
E-mail address: lywang@wayne.edu