Available online at www.sciencedirect.com

ScienceDirect

Stochastic Processes and their Applications 125 (2015) 3623–3635

stochastic processes and their applications

www.elsevier.com/locate/spa

Convergence of switching diffusions

Sören Christensen[a,b,*], Albrecht Irle[a]

[a] Christian-Albrechts-Universität, Mathematisches Seminar, Ludewig-Meyn-Str. 4, 24098 Kiel, Germany

[b] University of Hamburg, Department of Mathematics, SPST, Bundesstraße 55, 20146 Hamburg, Germany

Received 10 March 2014; received in revised form 26 November 2014; accepted 30 March 2015
Available online 7 April 2015

Abstract

This paper studies the asymptotic behavior of processes with switching. More precisely, the stability under fast switching for diffusion processes and discrete state space Markovian processes is considered. The proofs are based on semimartingale techniques, so that no Markovian assumption for the modulating process is needed.

© 2015 Elsevier B.V. All rights reserved.

MSC: 93E03; 60H20; 60H30

Keywords: Processes with switching; Switching diffusions; Fast switching; Asymptotic stability; Semimartingales

1. Introduction

Stochastic processes with dynamics depending on a further source of randomness have been of interest as well for theoretical reasons as from the point of view of application. Such processes are called processes with switching and usually switching involves an additional Markovian source of randomness with a finite number of states. For diffusion processes $(X_t)_{t\in[0,\infty)}$ given by a stochastic differential equation, the dynamics then additionally depend on a modulating Markovian process $(Y_t)_{t\in[0,\infty)}$ with finite state space. Mao and Yuan [8] give an extensive treatment of this subject. If $(X_t)_{t\in[0,\infty)}$ is itself a Markovian process with a discrete state space then the intensity matrix will depend on the modulating process.

* Corresponding author at: Christian-Albrechts-Universität, Mathematisches Seminar, Ludewig-Meyn-Str. 4, 24098 Kiel, Germany.

E-mail address: soren.christensen@gu.se (S. Christensen).

http://dx.doi.org/10.1016/j.spa.2015.03.010
0304-4149/© 2015 Elsevier B.V. All rights reserved.

In this paper we are interested in the stability under fast switching. This concerns the behavior of the switching process when the modulating process depends on an additional parameter $\epsilon > 0$ which lets it fluctuate more and more rapidly when ϵ tends to 0, so we have processes $(X_t^\epsilon)_{t\in[0,\infty)}$ and $(Y_t^\epsilon)_{t\in[0,\infty)}$. The question of interest concerns the asymptotic stability, i.e. convergence in a distributional sense, of $(X_t^\epsilon)_{t\in[0,\infty)}$ as ϵ tends to 0 which is by no means obvious as the processes $(Y_t^\epsilon)_{t\in[0,\infty)}$ fluctuate more and more rapidly.

In fast Markovian switching we look at processes $(Y_t^\epsilon)_{t\in[0,\infty)}$ with intensity matrix $\frac{1}{\epsilon}G$ for a given intensity matrix G; in pathwise terms we would look at $(Y_{t/\epsilon})_{t\in[0,\infty)}$. In [11] and [9] the asymptotic stability in the stochastic differential equation setting was shown, and in [4,3] this stability was derived for conditionally Poisson processes. For the proofs the assumption of Markovian switching was essential and the technical details can be seen as complicated and technically involved. Note that the processes $(X_t^\epsilon)_{t\in[0,\infty)}$ themselves are not Markovian so that the usual machinery for showing distributional convergence of Markov processes cannot be applied directly and has to be adapted.

This note stems from the observation that the processes $(X_t^\epsilon)_{t\in[0,\infty)}$ are semimartingales, so that we may show stability using the convergence theorem for families of semimartingales. As demonstrated here, this can indeed be done. Section 2 is devoted to the case of diffusion processes which is technically more involved relying on some uniform estimates for switching diffusions; Section 3 treats discrete state space Markovian processes where the proofs are simpler. The Appendix contains the proof of the analytical Lemma 2.4 which is essential for obtaining the main results.

An advantage of the semimartingale approach is that the Markovian assumption for the modulating process is no longer needed, only an assumption of ergodicity. Furthermore, the proofs turn out to be less complicated than using an approach based Markov theory. As switching processes have various applications in financial market modeling, see e.g. [5] or [2], where the modulating process may correspond to macroeconomic influences, this generalization might be of interest in this field. In particular, the findings discussed in [2] are based on the results presented in this paper.

2. Diffusion processes

We consider càdlàg processes $(X_t)_{t\in[0,\infty)}$, $(Y_t)_{t\in[0,\infty)}$, where $X_t : \Omega \to I$ for some interval $I \subseteq \mathbb{R}$, $Y_t : \Omega \to \mathcal{Y}$ for some suitable space $\mathcal{Y}$. Assume that for

$$b : I \times \mathcal{Y} \to \mathbb{R}, \qquad \sigma : I \times \mathcal{Y} \to \mathbb{R}$$

the process $(X_t)_{t\in[0,\infty)}$ fulfils the stochastic differential equation

$$dX_t = b(X_t, Y_t)dt + \sigma(X_t, Y_t)dW_t,$$

where $(W_t)_{t\in[0,\infty)}$ is a Wiener process independent of $(Y_t)_{t\in[0,\infty)}$. Note that the dynamics of the process X depend on the modulating process Y, so that we may also write $X = X^Y$.

Lemma 2.1. *Assume that there exist C_1, C_2 such that*

$$\max\{|b(x, y)|, |\sigma(x, y)|\} \le C_1 + C_2|x| \quad \textit{for all } x \in I, y \in \mathcal{Y}.$$

Let $q \ge 1$, $E|X_0|^q < \infty$. Then for all $T > 0$ there exists some constant C_3, only depending on q, C_1, C_2, T and $E|X_0|^q$, such that

$$\sup_{t\in[0,T]} E|X_t|^q \le C_3.$$

Proof. This follows immediately from [7, Lemma 2 and Corollary 6 in Section 2.5]. □

It is important to note that this estimate holds uniformly in all processes $(Y_t)_{t\in[0,\infty)}$ taking values in $\mathcal{Y}$, and this will also be explicitly stated in the following result:

Lemma 2.2. *Under the assumptions of Lemma 2.1 let $E|X_0| < \infty$. Then for any $\delta > 0$ there exists some $K > 0$ such that*

$$P\left(\sup_{t\le T}|X_t| \ge K\right) \le \delta$$

uniformly in all $\mathcal{Y}$-valued modulating processes $(Y_t)_{t\in[0,\infty)}$ for $X = X^Y$.

Proof. Use $3K$ instead of K. Then we see that

$$\begin{aligned} P&\left(\sup_{t\le T}\left|X_0 + \int_0^t b(X_s, Y_s)ds + \int_0^t \sigma(X_s, Y_s)dW_s\right| \ge 3K\right) \\ &\le P(|X_0| \ge K) + P\left(\sup_{t\le T}\left|\int_0^t b(X_s, Y_s)ds\right| \ge K\right) \\ &\quad + P\left(\sup_{t\le T}\left|\int_0^t \sigma(X_s, Y_s)dW_s\right| \ge K\right). \end{aligned}$$

The first term is trivial, so we start by looking at the second term. Clearly

$$\begin{aligned} \sup_{t\le T}\left|\int_0^t b(X_s, Y_s)ds\right| &\le \int_0^T |b(X_s, Y_s)|ds \le \int_0^T (C_1 + C_s|X_s|)\, ds \\ &= C_1T + \int_0^T C_2|X_s|ds. \end{aligned}$$

Now, using Lemma 2.1,

$$E\int_0^T C_2|X_s|ds = \int_0^T C_2E|X_s|ds \le C_2C_3T,$$

so that the expectation of the second term is bounded by $(C_1 + C_2C_3)T$, and the second term is bounded by Markov's inequality by

$$P\left(\sup_{t\le T}\left|\int_0^t b(X_s, Y_s)ds\right| \ge K\right) \le \frac{(C_1 + C_2C_3)T}{K}.$$

Now we consider the third term: The stochastic process $\left(\int_0^t \sigma(X_s, Y_s)dW_s\right)_{t\ge 0}$ is, for another Wiener process $(W'_t)_{t\ge 0}$, equal to $(W'_{\beta(t)})_{t\ge 0}$, where $\beta(t) = \int_0^t \sigma^2(X_s, Y_s)ds$. Hence, for any $\gamma > 0$

$$\begin{aligned} P\left(\sup_{t\le T}\left|\int_0^t \sigma(X_s, Y_s)dW_s\right| \ge K\right) &= P\left(\sup_{t\le T}|W'_{\beta(t)}| \ge K\right) = P\left(\sup_{t\le \beta(T)}|W'_t| \ge K\right) \\ &\le P\left(\beta(T) \ge \gamma\right) + P\left(\sup_{t\le\gamma}|W'_t| \ge K\right). \end{aligned}$$

Note that

$$E\beta(T) \le E\int_0^T (C_1 + C_2|X_s|)^2 ds = \int_0^T E(C_1 + C_2|X_s|)^2 ds \le C_4$$

by Lemma 2.1 for some C_4 only depending on C_1, C_2 and T. So firstly choose γ with $C_4/\gamma = \delta/4$, hence

$$P\left(\beta(T) \ge \gamma\right) \le \frac{\delta}{4}.$$

Now choose K with

$$P\left(\sup_{t\le\gamma} |W'_t| \ge K\right) \le \frac{\delta}{4}, \qquad P(|X_0| \ge K) \le \frac{\delta}{4}, \qquad \frac{(C_1 + C_2C_3)T}{K} \le \frac{\delta}{4}.$$

Altogether, we obtain

$$P\left(\sup_{t\le T} |X_t| \ge 3K\right) \le \delta. \quad \square$$

For fixed $\rho, T > 0$ define $\tau_0 = 0$ and

$$\tau_i = \inf\{s \ge \tau_{i-1} : |X_s - X_{\tau_{i-1}}| = \rho\} \wedge T, \qquad n_T = \sup\{k : \tau_k < T\}.$$

It is rather obvious that for a fixed process Y it holds that $P(\tau_1 = 0) = 0$ and $P(n_T < \infty) = 1$. In the following lemma, we show a version of this observation, that holds uniformly in all modulating processes $(Y_t)_{t\in[0,\infty)}$ taking values in $\mathcal{Y}$.

Lemma 2.3. *In addition to the assumptions of Lemma 2.1, let us assume that*

$$\sup_{|x|\le K, y\in\mathcal{Y}} (|b(x,y)| + \sigma^2(x,y)) < \infty \quad \textit{for all } K > 0. \tag{1}$$

Then for any $\rho, T, \delta > 0$ there exist $K', \rho' > 0$, such that

$$P(\tau_1 \ge \rho') \ge 1 - \delta, \qquad P(n_T \le K') \ge 1 - \delta$$

for all processes $(Y_t)_{t\in[0,\infty)}$ taking values in $\mathcal{Y}$.

Proof. (a) We start by looking at τ_1. As in the proof of Lemma 2.2 and using the same notation we have for $\rho' < T$

$$\begin{aligned}
P(\tau_1 \le \rho') &= P\left(\sup_{t\le\rho'} |X_t - X_0| \ge \rho\right) \\
&\le \frac{(C_1 + C_2C_3)\rho'}{\rho/2} + P(\beta(\rho') \ge \gamma) + P\left(\sup_{t\le\gamma} |W'_t| \ge \frac{\rho}{2}\right) \\
&\le C'_1\frac{\rho'}{\rho} + \frac{E\beta(\rho')}{\gamma} + P\left(\sup_{t\le\gamma} |W'_t| \ge \frac{\rho}{2}\right) \\
&\le C'_1\frac{\rho'}{\rho} + C'_2\frac{\rho'}{\gamma} + P\left(\sup_{t\le\gamma} |W'_t| \ge \frac{\rho}{2}\right).
\end{aligned}$$

Firstly we choose γ such that $P\left(\sup_{t\le\gamma}|W'_t|\ge\rho/2\right)\le\delta/3$. Then we choose ρ' such that $C'_1\frac{\rho'}{\rho}\le\delta/3,\ C'_2\frac{\rho'}{\gamma}\le\delta/3$, which gives the first estimate of the assertion.

(b) Let $\delta>0$. Choose K according to Lemma 2.2 such that

$$P\left(\sup_{t\le T}|X_t|\ge K\right)\le\frac{\delta}{2}.$$

We set

$$C=\sup_{|x|\le K,y\in\mathcal{Y}}|b(x,y)|,\qquad D=\sup_{|x|\le K,y\in\mathcal{Y}}\sigma^2(x,y).$$

The following estimates are always considered on $A=\{\sup_{t\le T}|X_t|\le K\}$. Note that for any $\tau,s\ge 0$

$$\left|\int_\tau^{\tau+s}b(X_s,Y_s)ds+\int_\tau^{\tau+s}\sigma(X_s,Y_s)dW_s\right|\ge\rho$$

implies

$$\int_\tau^{\tau+s}|b(X_s,Y_s)|ds\ge\frac{\rho}{2}\quad\text{or}\quad\left|\int_\tau^{\tau+s}\sigma(X_s,Y_s)dW_s\right|\ge\frac{\rho}{2}.$$

Hence on A

$$\int_\tau^{\tau+s}Cds\ge\frac{\rho}{2}\quad\text{or}\quad\left|\int_\tau^{\tau+s}\sigma(X_s,Y_s)dW_s\right|\ge\frac{\rho}{2}$$

so that

$$\tau_i-\tau_{i-1}\ge\min\left\{\frac{\rho}{2C},\ \inf\left\{s:\left|\int_{\tau_{i-1}}^{\tau_{i-1}+s}\sigma(X_s,Y_s)dW_s\right|\ge\frac{\rho}{2}\right\}\right\}.$$

Now we may argue in the following way. If $\tau_k<T$ then there exist k disjoint stochastic intervals $[\tau_{i-1},\tau_i)=J_i\subseteq[0,T)$ such that

$$\text{the length of } J_i \text{ is } \ge\frac{\rho}{2C}\quad\text{or}\ \sup_{s\in J_i}\left|\int_{\tau_{i-1}}^{s}\sigma(X_s,Y_s)dW_s\right|\ge\rho/2.$$

The number m of intervals J_i with length $\ge\frac{\rho}{2C}$ must fulfil $m\frac{\rho}{2C}<T$ so that there are at least

$$k-\frac{T2C}{\rho}\ \text{intervals } J_i\quad\text{with}\ \sup_{s\in J_i}\left|\int_{\tau_{i-1}}^{s}\sigma(X_s,Y_s)dW_s\right|\ge\frac{\rho}{2}.$$

To obtain a bound independent of the particular process $(Y_t)_t$ we transfer this to the process $(W'_t)_t$. Looking at the intervals $[\beta(\tau_{i-1}),\beta(\tau_i))=J'_i$ these are disjoint intervals $\subseteq[0,\beta(T))$ and for at least $k-\frac{T2C}{\rho}$ of them we have

$$\sup_{t\in J'_i}|W'_t-W'_{\beta(\tau_{i-1})}|\ge\frac{\rho}{2}.$$

Note that on A

$$\beta(T)=\int_0^T\sigma^2(X_t,Y_t)dt\le DT.$$

So it follows that $\tau_k<T$ implies the existence of at least $k-\frac{T2C}{\rho}$ such disjoint intervals $J'_i\subseteq[0,DT]$.

For a formal statement define the random variable

$$Z_{\rho',T'} = \sup\{k : \text{ There exist } k \text{ disjoint intervals } \subseteq [0,T') \\ \text{with } \sup_{a_i\le t\le b_i} |W'_t - W'_{a_i}| \ge \rho'\}.$$

By path continuity we see that

$$P(Z_{\rho',T'} = \infty) = 0, \quad \text{hence } P(Z_{\rho',T'} \ge k) \to 0 \text{ as } k \to \infty.$$

The foregoing reasoning implies

$$P(n_T \ge k) = P(\tau_k < T) \le P(A^c) + P\left(Z_{\rho/2,DT} \ge k - \frac{T2C}{\rho}\right).$$

So we only have to choose K' such that $P(Z_{\rho/2,DT} \ge K' - \frac{T2C}{\rho}) \le \delta/2$ to obtain the second estimate. Note that K' is independent of the particular process $(Y_t)_{t\in[0,\infty)}$. $\square$

Before coming to the main result of this section, we provide an analytical lemma which is essential for the following.

Lemma 2.4. *Let $f : \mathbb{R} \to \mathbb{R}$ be measurable such that*

$$\frac{1}{T}\int_0^T f(x)dx \to 0 \quad \textit{as } T \to \infty, \qquad \sup_T \frac{1}{T}\int_0^T |f(x)|dx < \infty.$$

Then for any continuous $h : [0,1] \to \mathbb{R}$

$$\frac{1}{T}\int_0^T h\left(\frac{x}{T}\right) f(x)dx \to 0 \quad \textit{as } T \to \infty.$$

The proof can be found in the Appendix.

Let us now fix some process $(Y_t)_{t\in[0,\infty)}$ with state space $\mathcal{Y}$. From now on we assume that $\mathcal{Y}$ is a finite set, which implies that condition (1) holds true. We assume that $(Y_t)_{t\in[0,\infty)}$ is ergodic in the sense that

$$\frac{1}{t}\int_0^t (1_{\{Y_s=y\}} - \pi(y))ds \to 0 \quad \text{a.s., as } t \to \infty$$

for some probability distribution $\pi(y)$, $y \in \mathcal{Y}$. We increase the speed of the process by looking at processes $(Y^\epsilon_t)_{t\in[0,\infty)}$, $\epsilon > 0$, having the same distribution as $(Y_{t/\epsilon})_{t\in[0,\infty)}$. The first processes have to be adapted to a filtration for which $(W_t)_{t\in[0,\infty)}$ is a Wiener process, and we let $(X^\epsilon_t)_{t\in[0,\infty)}$ be the solution of the corresponding stochastic differential equation

$$dX^\epsilon_t = b(X^\epsilon_t, Y^\epsilon_t)dt + \sigma(X^\epsilon_t, Y^\epsilon_t)dW_t,$$

with starting value independent of ϵ. In our proof we will work with $Y^\epsilon_t = Y_{t/\epsilon}$. For these processes to live on a common filtration with the Wiener process $(W_t)_{t\in[0,\infty)}$ we assume that

$$(W_t)_{t\in[0,\infty)} \text{ and } (Y_t)_{t\in[0,\infty)} \text{ are independent}$$

and then we may use $(Y_{t/\epsilon})_{t\in[0,\infty)}$ for $(Y^\epsilon_t)_{t\in[0,\infty)}$. This is no restriction in generality when compared with Skorokhod [11], Sarafyan and Skorokhod [9]. When $(Y_t)_{t\in[0,\infty)}$ is a Markov process with discrete state space living on the same filtration as $(W_t)_{t\in[0,\infty)}$ then these two processes are necessarily independent. This is shown in [10] for a Poisson process and can be generalized to general Markov processes with discrete state space; see e.g. [1].

Theorem 2.5. *Let $B : I \times \mathcal{Y} \to \mathbb{R}$ be such that $B(\cdot, y)$ is continuous for all $y \in \mathcal{Y}$. Then for all $T > 0$, $y \in \mathcal{Y}$ it holds that*

$$\sup_{0 \le r \le T} \left| \int_0^r B(X_t^\epsilon, y) \left(1_{\{Y_t^\epsilon = y\}} - \pi(y)\right) dt \right| \to 0 \quad \textit{in probability as } \epsilon \to 0.$$

Proof. Fix $T > 0$, $y \in \mathcal{Y}$. Let $\eta > 0$. We want to show that

$$P\left(\sup_{r \le T} \left| \int_0^r B(X_t^\epsilon, y) \left(1_{\{Y_t^\epsilon = y\}} - \pi(y)\right) dt \right| > \eta \right) \to 0 \quad \text{as } \epsilon \to 0.$$

Let $\delta > 0$. Choose K according to Lemma 2.2 such that

$$P\left(\sup_{t \le T} |X_t^\epsilon| \ge K \right) \le \delta \quad \text{for all } \epsilon > 0.$$

Let

$$A_\epsilon = \left\{ \sup_{t \le T} |X_t^\epsilon| < K \right\}.$$

The following estimates are always considered on A_ϵ. Let

$$C = \sup\{|B(x, y)| : |x| \le K\} < \infty;$$

thus for $r_0 = \eta / C$

$$\sup_{r \le r_0} \int_0^r \left| B(X_t^\epsilon, y) \left(1_{\{Y_t^\epsilon = y\}} - \pi(y)\right) \right| dt \le \eta.$$

By a change of variable, we obtain

$$\begin{aligned} \int_0^r B(X_t^\epsilon, y)(1_{\{Y_t^\epsilon = y\}} - \pi(y))dt &= \int_0^r B(X_t^\epsilon, y)(1_{\{Y_{t/\epsilon} = y\}} - \pi(y))dt \\ &= \epsilon \int_0^{r/\epsilon} B(X_{\epsilon s}^\epsilon, y)(1_{\{Y_s = y\}} - \pi(y))ds. \end{aligned}$$

Note that the integral

$$\frac{\epsilon}{r} \int_0^{r/\epsilon} B(X_{rs\epsilon/r}^\epsilon, y)(1_{\{Y_s = y\}} - \pi(y))ds$$

has exactly the form considered in the proof of Lemma 2.4 (see Appendix) with $t = r/\epsilon$, $g = B(\cdot, y)$, $h(s) = X_{rs}^\epsilon$, and $f(s) = 1_{\{Y_s = y\}} - \pi(y)$. On A_ϵ one has

$$|X_{rs}^\epsilon| \le K \quad \text{for all } s \le 1.$$

Let $\delta > 0$. Since $B(\cdot, y)$ is uniformly continuous on $[-K, K]$ we may choose $\rho > 0$ such that

$$|B(x, y) - B(x', y)| \le \frac{\eta}{2T} \quad \text{for } |x - x'| \le \rho, \; x, x' \in [-K, K].$$

Next, let $\sigma_0^{\epsilon,r} = 0$,

$$\sigma_i^{\epsilon,r} = \inf\{s \ge \sigma_{i-1}^{\epsilon,r} : |X_{rs}^\epsilon - X_{\sigma_{i-1}^{\epsilon,r}}^\epsilon| = \rho\} \wedge 1, \qquad n^{\epsilon,r} = \sup\{i : \sigma_i^{\epsilon,r} < 1\}.$$

The estimate in the proof of Lemma 2.4 with $C_1 = 1, C_2 = C$ yields

$$\left|\epsilon \int_0^{r/\epsilon} B(X^\epsilon_{rs\epsilon/r}, y)(1_{\{Y_s=y\}} - \pi(y))ds\right| \\ \le r\frac{\eta}{2} + 2rC \sum_{i=1}^{n^{\epsilon,r}+1} \left|\epsilon \int_0^{\sigma_i^{\epsilon,r} r/\epsilon} (1_{\{Y_s=y\}} - \pi(y))dy\right|.$$

Setting $\tau_0^\epsilon = 0$,

$$\tau_i^\epsilon = \inf\{s \ge \tau_{i-1}^\epsilon : |X_s^\epsilon - X^\epsilon_{\tau^\epsilon_{i-1}}| = \rho\} \wedge T, \qquad n^\epsilon = \sup\{k : \tau_k^\epsilon < T\},$$

it follows that $r\sigma_i^{\epsilon,r} = \tau_i^\epsilon \wedge r,\ \ n^{\epsilon,r} \le n^\epsilon$ and

$$\left|\epsilon \int_0^{r/\epsilon} B(X^\epsilon_{\epsilon s}, y)(1_{\{Y_s=y\}} - \pi(y))ds\right| \le \frac{\eta}{2} + 2C \sum_{i=1}^{n^{\epsilon,r}+1} \epsilon \int_0^{\tau_i^\epsilon \wedge r/\epsilon} (1_{\{Y_s=y\}} - \pi(y))ds.$$

Next, according to Lemma 2.3, we may choose $K', \rho' \le r_0$ such that

$$P(n^\epsilon \le K') \ge 1-\delta, \qquad P(\tau_1^\epsilon \ge \rho') \ge 1-\delta \quad \text{for all } \epsilon > 0,$$

hence also

$$P(\tau_1^\epsilon \wedge r \ge \rho') \ge 1-\delta \quad \text{for all } \epsilon > 0,\ r \ge r_0.$$

Due to the ergodicity assumption

$$\frac{1}{t}\int_0^t (1_{\{Y_s=y\}} - \pi(y))ds \to 0 \quad \text{a.s., as } t \to \infty.$$

So we may choose $\epsilon_0 > 0$ such that for

$$D_\epsilon = \left\{\sup_{t \ge \rho'} \left|\epsilon \int_0^{t/\epsilon} (1_{\{Y_s=y\}} - \pi(y))ds\right| \le \frac{\eta}{2}\frac{1}{2(K'+1)C}\right\}$$

we have $P(D_\epsilon) \ge 1-\delta$ for all $\epsilon \le \epsilon_0$.

Altogether, we obtain on $A_\epsilon \cap \{n^\epsilon \le K'\} \cap \{\tau_1^\epsilon \ge \rho'\} \cap D_\epsilon$ that for all $r_0 \le r \le T,\ 0 < \epsilon \le \epsilon_0$

$$\left|\epsilon \int_0^{r/\epsilon} B(X^\epsilon_{\epsilon s}, y)(1_{\{Y_s^\epsilon=y\}} - \pi(y))ds\right| \le \frac{\eta}{2} + 2C(n^\epsilon+1)\frac{\eta}{2}\frac{1}{2(K'+1)C} \le \eta,$$

hence

$$\sup_{0<r\le T}\left|\int_0^r B(X_s^\epsilon, y)(1_{\{Y_s^\epsilon=y\}} - \pi(y))ds\right| \le \eta,$$

the case $r \le r_0$ being obvious as remarked in the beginning of the proof. It follows that

$$P\left(\sup_{0<r\le T}\left|\int_0^r B(X_s^\epsilon, y)(1_{\{Y_s^\epsilon=y\}} - \pi(y))ds\right| > \eta\right) \\ \le P(A_\epsilon^c) + P(n^\epsilon \ge K') + P(\tau_1^\epsilon \le \rho') + P(D_\epsilon^c) \le 4\delta,$$

for all $0 < \epsilon \le \epsilon_0$. $\square$

Now, the previous result can be utilized to prove the convergence in distribution for fast switching diffusions as follows.

Define

$$\hat{b} : I \to \mathbb{R}, \qquad \hat{b}(x) = \sum_{y \in \mathcal{Y}} b(x, y)\pi(y),$$

$$\hat{\sigma} : I \to \mathbb{R}, \qquad \hat{\sigma}(x) = \sum_{y \in \mathcal{Y}} \sigma(x, y)\pi(y),$$

and $(\hat{X}_t)_{t\in[0,\infty)}$ as the solution of the corresponding stochastic differential equation

$$d\hat{X}_t = \hat{b}(\hat{X}_t)dt + \hat{\sigma}(\hat{X}_t)dW_t.$$

Theorem 2.6.

$$(X^{\epsilon_t})_{t\in[0,\infty)} \to (\hat{X}_t)_{t\in[0,\infty)} \quad \textit{in distribution as } \epsilon \to 0.$$

Proof. The infinitesimal characteristics are

$$b(X_t^\epsilon, Y_t^\epsilon),\ \sigma(X_t^\epsilon, Y_t^\epsilon) \text{ for the semimartingale } (X^{\epsilon_t})_{t\in[0,\infty)},$$

and

$$\hat{b}(\hat{X}_t),\ \hat{\sigma}(\hat{X}_t) \text{ for the semimartingale } (\hat{X}_t)_{t\in[0,\infty)}.$$

Theorem 2.5 shows that for all $T > 0$

$$\sup_{0\le r\le T} \int_0^r \left(b(X_t^\epsilon, Y_t^\epsilon) - \hat{b}(X_t^\epsilon)\right) dt \to 0 \quad \text{in probability as } \epsilon \to 0,$$

and similarly for $\sigma, \hat{\sigma}$. This implies the assertion by the semimartingale convergence theorem; see [6, Theorem 3.21, Chapter IX]. □

3. Discrete state processes

In the case of a discrete state space we start with a càdlàg process $(Y_t)_{t\in[0,\infty)}$ with a finite state space $\mathcal{Y}$, assumed a subset of $\mathbb{R}$ without loss of generality, and discrete jump times $\gamma_0 = 0 < \gamma_1 < \gamma_2 < \cdots$. Here, the term discrete jump times means that these times are strictly increasing and the process Y is constant on $[\gamma_{i+1}, \gamma_i)$ for all i. Furthermore, let $I \subseteq \mathbb{R}$ be countable, and for each $y \in \mathcal{Y}$ let $q(\cdot, \cdot|y)$ be an intensity matrix. Conditionally, we generate the switching process in the following way: In zero, we start a continuous time Markov chain $\tilde{X}^0$ with starting state x_0 and intensity matrix $q(\cdot, \cdot|Y_0)$. In the first jump time γ_1, we start a new chain $\tilde{X}^1$ with starting point $\tilde{X}^0_{\gamma_1}$ and intensity matrix $q(\cdot, \cdot|Y_{\gamma_1})$. In the same way we define $\tilde{X}^{i+1}$, starting in $\tilde{X}^i_{\gamma_i}$. We define the switching process $X = X^Y$ by setting $X_t = \tilde{X}^i_{t-\gamma_i}$ for $\gamma_i \le t < \gamma_{i+1}$.

Note that when using diffusion processes $\tilde{X}^i$ with coefficients depending on y instead of Markov chains in the previous construction, (ignoring some technical issues) the process X is a switching diffusion as considered in Section 2. In this section we consider the jump counterpart for the results obtained there.

We make the assumption that for some finite set $J \subseteq \mathbb{R} \setminus \{0\}$

$$q(i, i + j|y) = 0 \quad \text{for all } i \in I,\ y \in \mathcal{Y},\ j \notin J,$$

so $i + J$ is the set of states which can be reached from i. Thus $q(i|y) := \sum_{i'\ne i} q(i, i'|y) < \infty$ for all $y \in \mathcal{Y}$, and we furthermore assume that

$$q := \sup_{i,y} q(i|y) < \infty.$$

If the state space I of the process X is finite, these assumptions are of course fulfilled, and they imply that there is no explosion in finite time.

Define the jumps time of $(X_t)_{t\geq 0}$ by $\tau_0 = 0$ and

$$\tau_i = \inf\{t \geq \tau_{i-1} : X_t \neq X_{\tau_{i-1}}\}.$$

The following lemma is analog to Lemma 2.3 and gives again an estimate, which is uniform in all modulating processes Y. The proof now turns out to be much easier due to the discrete nature of situation.

Lemma 3.1. *Let $T > 0$, $n_T = \sup\{k : \tau_k < T\}$. Let $\delta > 0$. Then there exist $\rho' > 0$, $K' \in \mathbb{N}$ such that*

$$P(\tau_1 \geq \rho') \geq 1 - \delta, \qquad P(n_T < K') \geq 1 - \delta$$

for all processes $(Y_t)_{t\in[0,\infty)}$ taking values in $\mathcal{Y}$.

Proof. Denote the counting process of $(X_t)_{t\in[0,\infty)}$ by $(N_t)_{t\in[0,\infty)}$, and by $(N_t^i)_{t\in[0,\infty)}$ for $(X_t^i)_{t\in[0,\infty)}$, i.e. N_t (resp. N_t^i) denotes the number of jumps of X (resp. X^i) before time t. Let the random index j_t be given by $\gamma_{j_t} \leq t < \gamma_{j_t+1}$. Then we consider

$$\{\tau_1 > t\} = \{N_t = 0\} = \{N^0_{\gamma_1} = 0\} \cap \{N^1_{\gamma_2-\gamma_1} = 0\} \cap \cdots \cap \{N^{j_t}_{t-\gamma_{j_t}} = 0\}.$$

It follows by conditioning that

$$P(N_t = 0) = E\left(e^{-q(X_0|Y_0)\gamma_1} e^{-q(X_{\gamma_1}|Y_{\gamma_1})(\gamma_2-\gamma_1)} \ldots e^{-q(X_{\gamma_{j_t}}|Y_{\gamma_{j_t}})(t-\gamma_{j_t})}\right) \geq e^{-qt}.$$

Now, let $(N_t^*)_{t\in[0,\infty)}$ denote a Poisson process with intensity q and corresponding jump times τ_j^*, $j \geq 0$. Then the previous arguments show that $P(\tau_1 > t) \geq P(\tau_1^* > t)$. Together with the Markov property, the same arguments can be used for τ_j, τ_j^*, $j > 1$, so that $P(\tau_j > t) \geq P(\tau_j^* > t)$. Hence

$$P(\tau_1 \geq \rho) \geq P(\tau_1^* \geq \rho) \to 1 \quad \text{as } \rho \to 0,$$
$$P(n_T \leq K) = P(\tau_{k+1} > T) \geq P(\tau_{K+1}^* > T) \to 1 \quad \text{as } K \to \infty. \qquad \square$$

Lemma 3.2. *Assume the situation of Lemma 3.1. For any $\epsilon > 0$ there exists K such that*

$$P\left(\sup_{t\leq T} |X_t| \geq K\right) \leq \epsilon$$

uniformly in all $\mathcal{Y}$-valued modulating processes $(Y_t)_{t\in[0,\infty)}$ for $X = X^Y$.

Proof. Using Markov's inequality we have

$$P\left(\sup_{t\leq T} |X_t| \geq K\right) \leq E(\sup_{t\leq T} |X_t|)/K \leq (|x_0| + E(n_T + 1) \sup_{j\in J} |j|)/K.$$

Therefore, it is enough to show that $E(n_T)$ is bounded uniformly in Y. But this holds by the proof of the preceding Lemma since En_T is not larger than the expected number of jumps of the Poisson process $(N_t^*)_{t\in[0,\infty)}$ in $[0, T]$, which is well-known to be finite. $\square$

As in Section 2 we fix a process $(Y_t)_{t\in[0,\infty)}$ with the property

$$\frac{1}{t}\int_0^t (1_{\{Y_s=y\}} - \pi(y))ds \to 0 \quad \text{a.s., as } t \to \infty.$$

Furthermore, we take $(Y_t^\epsilon)_{t\in[0,\infty)} = (Y_{t/\epsilon})_{t\in[0,\infty)}$ with corresponding $(X_t^\epsilon)_{t\in[0,\infty)}$.

Theorem 3.3. *Let $Q : I \times \mathcal{Y} \to \mathbb{R}$ be such that $\sup_{|i|\le K} |Q(i, y)| < \infty$ for all $y \in \mathcal{Y}$, $K > 0$. For all $y \in \mathcal{Y}$, $T > 0$ it holds that*

$$\sup_{0\le r\le T} \left| \int_0^r Q(X_t^\epsilon, y)\left(1_{\{Y_t^\epsilon=y\}} - \pi(y)\right) dt \right| \to 0 \quad \textit{in probability as } \epsilon \to 0.$$

Proof. Fix y, T. For $\delta > 0$ choose K according to Lemma 3.2 such that

$$P\left(\sup_{t\le T} |X_t^\epsilon| \ge K\right) \le \delta \quad \text{for all } \epsilon > 0.$$

Write

$$A_\epsilon = \{\sup_{t\le T} |X_t^\epsilon| \le K\}, \qquad C = \sup\{|Q(i, y)| : |i| \le K\} < \infty.$$

Then for $r_0 = \eta/C$ on A_ϵ

$$\sup_{r\le r_0} \int_0^r |Q(X_t^\epsilon, y)(1_{\{Y_t=y\}} - \pi(y))|dt \le \eta.$$

We may proceed with a simplified version of the proof of Theorem 2.5 without the use of Lemma 2.4. It holds that

$$\int_0^r Q(X_t^\epsilon, y)(1_{\{Y_t^\epsilon=y\}} - \pi(y))dt = \epsilon \int_0^{r/\epsilon} Q(X_{rs\epsilon/r}^\epsilon, y)(1_{\{Y_s=y\}} - \pi(y))ds.$$

Define $\sigma_i^{\epsilon,r}$ as in the proof of Theorem 2.5 replacing $= \rho$ by > 0 to obtain the jump times, with corresponding $n^{\epsilon,r}$. Then, on A_ϵ

$$\begin{aligned}
&\left|\epsilon \int_0^{r/\epsilon} Q(X_{rs\epsilon/r}^\epsilon, y)(1_{\{Y_s^\epsilon=y\}} - \pi(y))ds\right| \\
&\le \left|\epsilon \sum_{i=1}^{n^{\epsilon,r}+1} Q(X_{r\sigma_i^{\epsilon,r}}^\epsilon, y) \int_{\sigma_{i-1}^{\epsilon,r}r/\epsilon}^{\sigma_i^{\epsilon,r}r/\epsilon} (1_{\{Y_s=y\}} - \pi(y))ds\right| \\
&\le 2C(n^{\epsilon,r} + 1) \left|\epsilon \int_0^{\sigma_{n^{\epsilon,r}+1}^{\epsilon,r}r/\epsilon} (1_{\{Y_s=y\}} - \pi(y))ds\right|.
\end{aligned}$$

Using Lemma 3.1, the proof is concluded as in Theorem 2.5. □

Define the intensity matrix

$$\hat{q}(\cdot, \cdot) = \sum_{y\in\mathcal{Y}} q(\cdot, \cdot|y)\pi(y)$$

with corresponding Markov process $(\hat{X}_t)_{t\in[0,\infty)}$.

Theorem 3.4.

$$(X_t^\epsilon)_{t\in[0,\infty)} \to (\hat{X}_t)_{t\in[0,\infty)} \quad \textit{in distribution as } \epsilon \to 0.$$

Proof. The infinitesimal jump characteristics are given by

$$q(X_t^\epsilon, X_t^\epsilon + j|Y_t^\epsilon) \text{ for the semimartingale } (X^{\epsilon_t})_{t\in[0,\infty)},$$

and

$$\hat{q}(\hat{X}_t, \hat{X}_t + j) \text{ for the semimartingale } (\hat{X}_t)_{t\in[0,\infty)}.$$

Theorem 2.5 shows that for all $T > 0,\ j \in J$

$$\sup_{0\le r\le T} \int_0^r \left(q(X_t^\epsilon, X_t^\epsilon + j|Y_t^\epsilon) - \hat{q}(X_t^\epsilon, X_t^\epsilon + j)\right) dt \to 0 \quad \text{in probability as } \epsilon \to 0$$

using $Q(i, y) = q(i, i+j|y)$. Again this implies the assertion by the semimartingale convergence theorem; see [6, Theorem 3.21, Chapter IX]. □

Acknowledgment

We thank the anonymous referee for carefully reading an earlier version of this manuscript and for thoughtful and valuable comments that helped to improve the presentation of this article.

Appendix. Proof of Lemma 2.4

Proof of Lemma 2.4. Note that in the proof we shall provide a more precise inequality which will be used in proving Theorem 2.5. Due to this reason we shall use a further continuous mapping $g : \mathbb{R} \to \mathbb{R}$ and write $g(h(x/T))$ instead of $h(x/T)$. For the assertion of this analytical lemma, g is just the identity. Let $[\alpha, \beta] = h([0, 1])$. Let $\delta > 0$. Since g is uniformly continuous on $[\alpha, \beta]$ there exists $\rho > 0$ such that

$$|g(y) - g(y')| \le \frac{\delta}{C_1} \quad \text{for } |y - y'| \le \rho,\ y, y' \in [\alpha, \beta],$$

where $C_1 = \sup_T \frac{1}{T} \int_0^T |f(x)|dx$; also let $C_2 = \sup_{y\in[\alpha,\beta]} |g(y)|$. Set

$$s_0 = 0, \qquad s_i = \inf\{s \ge s_{i-1} : |h(s) - h(s_{i-1})| = \rho\} \wedge 1,$$

furthermore $n = \sup\{i : s_i < 1\}$, where, due to the continuity of h, n is finite. Now

$$\begin{aligned}
&\left|\frac{1}{t}\int_0^t g\left(h\left(x/t\right)\right) f(x)dx\right| \\
&\quad\le \left|\frac{1}{t}\sum_{i=1}^{n+1}\int_{ts_{i-1}}^{ts_i} \left(g\left(h\left(x/t\right)\right) - g\left(h\left(s_i\right)\right)\right) f(x)dx\right| + \left|\frac{1}{t}\sum_{i=1}^{n+1}\int_{ts_{i-1}}^{ts_i} g\left(h\left(s_i\right)\right) f(x)dx\right| \\
&\quad\le \frac{1}{t}\frac{\delta}{C_1}\sum_{i=1}^{n+1}\int_{ts_{i-1}}^{ts_i} |f(x)|dx + \frac{1}{t}\sum_{i=1}^{n+1}\left|g\left(h\left(s_i\right)\right)\right|\left|\int_{ts_{i-1}}^{ts_i} f(x)dx\right| \\
&\quad\le \frac{\delta}{C_1}\frac{1}{t}\int_0^t |f(x)|dx + \frac{1}{t}\sum_{i=1}^{n+1} C_2\left|\int_0^{ts_i} f(x)dx - \int_0^{ts_{i-1}} f(x)dx\right| \\
&\quad\le \delta + 2C_2\sum_{i=1}^{n+1}\left|\frac{1}{t}\int_0^{ts_i} f(x)dx\right|.
\end{aligned}$$

Now choose t_0 such that

$$\sup_{s \geq s_1} \left| \frac{1}{t} \int_0^{ts} f(x)dx \right| \leq \frac{\delta}{2C_2(n+1)} \quad \text{for all } t \geq t_0,$$

thus

$$\left| \frac{1}{t} \int_0^{t} g\left(h\left(x/t\right)\right) f(x)dx \right| \leq \delta + 2C_2(n+1) \frac{\delta}{2C_2(n+1)} = 2\delta. \quad \square$$

References

[1] S. Christensen, A note on the independence of Wiener processes and Markov processes, Working Paper, University of Kiel, 2014.
[2] S. Christensen, A. Irle, J. Kauschke, Agent Based Models With Varying Switching Rates and Fast Information Changes, Preprint, 2014.
[3] G.B. DiMasi, Yu.M. Kabanov, The strong convergence of two-scale stochastic systems and singular perturbations of filtering equations, J. Math. Systems Estim. Control 3 (1993) 207–224.
[4] G.B. DiMasi, Yu.M. Kabanov, A first order approximation for the convergence of distributions of the Cox processes with fast Markov switches, Stoch. Stoch. Rep. 54 (1995) 211–219.
[5] A. Irle, J. Kauschke, T. Lux, M. Milakovic, Switching rates and the asymptotic behaviour of herding models, Adv. Complex Syst. 14 (2011) 359–376.
[6] J. Jacod, A.N. Shiryaev, Limit Theorems for Stochastic Processes, second ed., in: Grundlehren der Mathematischen Wissenschaften (Fundamental Principles of Mathematical Sciences), vol. 288, Springer-Verlag, Berlin, 2003.
[7] N.V. Krylov, Controlled Diffusion Processes, Springer-Verlag, Berlin, 1980.
[8] X. Mao, C. Yuan, Stochastic Differential Equations with Markovian Switching, Imperial College Press, London, 2006.
[9] V.V. Sarafyan, A.V. Skorokhod, Dynamical systems with fast switchings, Teor. Veroyatn. Primen. 32 (1987) 658–669 (in Russian).
[10] S. Shreve, Stochastic Calculus for Finance II, Springer, New York, 2004.
[11] A.V. Skorokhod, Asymptotic Methods in the Theory of Stochastic Differential Equations, in: Translations of Mathematical Monographs, vol. 78, American Mathematical Society, Providence, RI, 1989, Translated from the Russian by H.H. McFaden.