Available online at www.sciencedirect.com
ScienceDirect
Stochastic Processes and their Applications 116 (2006) 1660–1675
stochastic processes and their applications
www.elsevier.com/locate/spa

Simulation of conditioned diffusion and application to parameter estimation

Bernard Delyon, Ying Hu*

IRMAR, Université Rennes 1, Campus de Beaulieu, 35042 Rennes Cedex, France

Received 16 March 2005; received in revised form 10 January 2006; accepted 11 April 2006
Available online 2 May 2006

Abstract

In this paper, we propose some algorithms for the simulation of the distribution of certain diffusions conditioned on a terminal point. We prove that the conditional distribution is absolutely continuous with respect to the distribution of another diffusion which is easy for simulation, and the formula for the density is given explicitly. An example of parameter estimation for a Duffing–Van der Pol oscillator is given as an application.

Keywords: Simulation; Conditioned diffusion; Monte Carlo Markov Chain

1. Introduction

The aim of this paper is to propose algorithms for the simulation of the distribution of a diffusion

$$\mathrm{d}x_t = b(t, x_t)\mathrm{d}t + \sigma(t, x_t)\mathrm{d}w_t, \quad x_0 = u, \quad 0 \le t \le T,$$

conditioned on $x_T = v$, where b and σ are given functions with appropriate dimensions, and w is a standard Brownian motion.

From the point of view of application, this allows us to do posterior sampling when the diffusion is observed at instants $\{t_1, \ldots, t_n\} \subset [0, T]$.

Let us recall that in the usual conditioning (see, e.g., [7]), the distribution of the diffusion x conditioned on $x_T = v$ is the same as that of another diffusion y satisfying

$$\mathrm{d}y_t = \tilde{b}(t, y_t)\mathrm{d}t + \sigma(t, y_t)\mathrm{d}w_t, \quad y_0 = u, \quad 0 \le t \le T,$$

* Corresponding author. Tel.: +33 223235820; fax: +33 223236790.
E-mail addresses: bernard.delyon@univ-rennes1.fr (B. Delyon), ying.hu@univ-rennes1.fr (Y. Hu).

0304-4149/$ - see front matter
doi:10.1016/j.spa.2006.04.004

where

$$\tilde{b}(t,x) = b(t,x) + [\sigma\sigma^*](t,x)\nabla_x(\log p(t,x;T,v)),$$

and $p(s,u;t,z)$ is the density of $x_t^{s,u}$. However, this is not suitable for simulations because in general, one does not know the transition density p.

We will prove that, in certain cases, the conditional distribution of the diffusion is absolutely continuous with respect to the distribution of another diffusion which is easy for simulation, and we give the explicit formula for the density. This leads to an efficient simulation algorithm.

Two different cases will be considered:

1. The matrix $\sigma(t,x)$ depends only on t, and b has the form $b(t,x) = b_0(t) + A(t)x + \sigma(t)b_1(t,x)$.
2. The matrix $\sigma(t,x)$ is uniformly invertible.

We apply also our simulation algorithm to a problem of parameter estimation for a Duffing–Van Der Pol oscillator. We give our simulation for illustration.

This paper is organized as follows: In Section 2, we recall a Girsanov theorem for unbounded drift which is essential for our simulation algorithm. In Section 3, we consider Case 1, and in Section 4, we consider Case 2. The last section is devoted to the application.

2. A Girsanov theorem for unbounded drifts

This section is devoted to give a slightly generalized Girsanov theorem which will be used in the next section. We call a measurable function $F(t,x)$ from $\mathbb{R}_+ \times \mathbb{R}^d$ to $\mathbb{R}^n$ locally Lipschitz with respect to x, if for any $R > 0$, there exists a constant $C_R > 0$, such that, for any $(t,x,y) \in \mathbb{R}_+ \times \mathbb{R}^d \times \mathbb{R}^d$ with $|x| \le R$, $|y| \le R$,

$$|F(t,x) - F(t,y)| \le C_R|x-y|.$$

And on the metric space $C([0,T];\mathbb{R}^m)$, we define the filtration $\{\mathcal{F}_t\}_t$ to be the natural filtration of the coordinate process.

Theorem 1. *Let $b(t,x)$, $h(t,x)$, $\sigma(t,x)$ be measurable functions from $\mathbb{R}_+ \times \mathbb{R}^d$ to $\mathbb{R}^d$, $\mathbb{R}^m$, and $\mathbb{R}^{d\times m}$ which are locally Lipschitz with respect to x; consider the following stochastic differential equations:*

$$\mathrm{d}x_t = b(t,x_t)\mathrm{d}t + \sigma(t,x_t)\mathrm{d}w_t, \tag{1}$$

$$\mathrm{d}y_t = (b(t,y_t) + \sigma(t,y_t)h(t,y_t))\mathrm{d}t + \sigma(t,y_t)\mathrm{d}w_t, \quad y_0 = x_0, \tag{2}$$

on the finite interval $[0,T]$. We assume the existence of strong solution for each equation. We assume in addition that h is bounded on compact sets. Then the Girsanov formula holds: for any non-negative Borel function $f(x,w)$ defined on $C([0,T];\mathbb{R}^d) \times C([0,T];\mathbb{R}^m)$, one has

$$E_y[f(y,w^h)] = E_x[f(x,w)\mathrm{e}^{\int_0^T h^*(t,x_t)\mathrm{d}w_t - \frac{1}{2}\int_0^T |h(t,x_t)|^2\mathrm{d}t}], \tag{3}$$

$$E_x[f(x,w)] = E_y[f(y,w^h)\mathrm{e}^{-\int_0^T h^*(t,y_t)\mathrm{d}w_t - \frac{1}{2}\int_0^T |h(t,y_t)|^2\mathrm{d}t}], \tag{4}$$

where $w_t^h = w_t + \int_0^t h(s,y_s)\mathrm{d}s$, and h^* stands for the transpose of h.

Proof. We assume first that the positive supermartingale

$$M_t = \exp\left\{\int_0^t h^*(s, x_s)\mathrm{d}w_s - \frac{1}{2}\int_0^t |h(s, x_s)|^2 \mathrm{d}s\right\}$$

is a martingale under P_x which will be proved later. In this case $\tilde{w}_t = w_t - \int_0^t h(s, x_s)\mathrm{d}s$ is a Brownian motion under $M_T P_x$, leading to a solution $(x, \tilde{w})$ of (2):

$$\mathrm{d}x_t = b(t, x_t)\mathrm{d}t + \sigma(t, x_t)h(t, x_t)\mathrm{d}t + \sigma(t, x_t)\mathrm{d}\tilde{w}_t.$$

As $b(t, x)$, $h(t, x)$, $\sigma(t, x)$ are locally Lipschitz with respect to x, pathwise uniqueness holds for (1) and (2). The standard Girsanov theorem implies that (3) holds.

We prove now that M_t is a martingale. For any $R > 0$, consider the stopping time

$$\tau_R = \inf\{t \geq 0 : |x_t| \geq R\} \wedge T.$$

Taking into consideration that h is locally bounded, we have, according to the Girsanov theorem for bounded drift:

$$P_{y|\mathcal{F}_{\tau_R}} = M_{\tau_R} P_{x|\mathcal{F}_{\tau_R}}.$$

Hence

$$E_x[M_T] \geq E_x[1_{\tau_R = T} M_T] = E_x[1_{\tau_R = T} M_{\tau_R}] = P_y[\tau_R = T]$$

which converges to 1 as $R \to \infty$. It implies that $E_x[M_T] = 1$, and M is a martingale.

Finally, (4) follows in the same way. □

3. Case when σ is independent of x

We assume here that x_t has the specific form

$$\mathrm{d}x_t = (\sigma_t h(t, x_t) + A_t x_t + b_t)\mathrm{d}t + \sigma_t \mathrm{d}w_t, \quad x_0 = u, \tag{5}$$

where σ_t and A_t are time dependent deterministic matrices and $h(t, x)$, b_t are vector valued with appropriate dimension.

For example the two-dimensional process (x, y) which satisfies the following SDE:

$$\mathrm{d}x_t = y_t \mathrm{d}t \tag{6}$$
$$\mathrm{d}y_t = b(t, x_t, y_t)\mathrm{d}t + \sigma \mathrm{d}w_t \tag{7}$$

and which is the noisy version of $\ddot{x}_t = b(t, x_t, \dot{x}_t)$, see, e.g. [1].

We shall prove the following result:

Theorem 2. *Assume that A_t, b_t and σ_t are bounded measurable functions of t with values in $\mathbb{R}^{d\times d}$, $\mathbb{R}^d$ and $\mathbb{R}^{d\times m}$, respectively. Assume also that $h(t, x)$ is locally Lipschitz with respect to x uniformly with respect to t with values in $\mathbb{R}^m$, and locally bounded; and the SDE (5) has a strong solution. Moreover, we assume that σ admits a measurable left inverse almost everywhere,*[1] *denoted by σ^+; and that h, A, b and σ^+ are left continuous with respect to t. Then,*

[1] This requires essentially that $\sigma^*\sigma$ is almost everywhere >0.

(i) *the covariance matrix* R_{st} *of the Gaussian process* ξ_t *corresponding to* (5) *with* $h = 0$ *is given by:*

$$R_{st} = P_s \int_0^{\min(s,t)} P_u^{-1}\sigma_u\sigma_u^* P_u^{-*}\mathrm{d}u\ P_t^*,$$

where

$$\frac{\mathrm{d}P_t}{\mathrm{d}t} = A_t P_t, \quad P_0 = Id,$$

and $P_u^{-*} = (P_u^{-1})^*$;

(ii) *the distribution of the process*

$$p_t = \xi_t - R_{tT}R_{TT}^+(\xi_T - v) \tag{8}$$

is the same as the distribution of ξ *conditioned on* $\xi_T = v$ *(*M^+ *stands for the left pseudo-inverse*[2] *of* M*). For any non-negative measurable function* f,

$$E[f(x)|x_0 = u, x_T = v] = CE[f(p)\mathrm{e}^{\int_0^T h^*(t,p_t)(\sigma_t^+\mathrm{d}p_t - \sigma_t^+(A_t p_t + b_t)\mathrm{d}t) - \frac{1}{2}\int_0^T \|h(t,p_t)\|^2\mathrm{d}t}], \tag{9}$$

where C *is a constant depending on* u, v *and* T.

Proof. (i) The formula for R_{st} is classic and comes from $\xi_t = P_t\int_0^t P_u^{-1}(b_u\mathrm{d}u + \sigma_u\mathrm{d}w_u) + P_t\xi_0$, see e.g. [6].

(ii) Let us first recall that if (Y, Z) is a Gaussian vector, the distribution of Y conditioned on $Z = z_0$ coincides with the distribution of another Gaussian vector $Y - R_{YZ}R_{ZZ}^+(Z - z_0)$, where R_{ZZ}^+ is the left pseudo-inverse of R_{ZZ}; its covariance is $R_{YY} - R_{YZ}R_{ZZ}^+R_{ZY}$. Taking Y as the vector $(\xi_{t_1}, \ldots, \xi_{t_k})$, and $Z = \xi_T$, we observe that, defining the process p by (8), $(p_{t_1}, \ldots, p_{t_k})$ has the same distribution as that of $(\xi_{t_1}, \ldots, \xi_{t_k})$ conditioned on $\xi_T = v$. And the covariance of p_t is $C_{st} = R_{st} - R_{sT}R_{TT}^+R_{Tt}$.

Denote by p_t^v the process (8); in particular for any non-negative measurable function $\varphi(\cdot)$, $E[\varphi(\xi)] = \int E[\varphi(p^v)]\mu_T(\mathrm{d}v)$ where μ_T is the distribution of ξ_T. For any non-negative measurable functions f and g,

$$\begin{aligned} E[f(x)g(x_T)] &= E[f(\xi)g(\xi_T)\mathrm{e}^{\int_0^T h^*(t,\xi_t)\mathrm{d}w_t - \frac{1}{2}\int_0^T \|h(t,\xi_t)\|^2\mathrm{d}t}] \\ &= E[f(\xi)g(\xi_T)\mathrm{e}^{\int_0^T h^*(t,\xi_t)(\sigma_t^+\mathrm{d}\xi_t - \sigma_t^+(A_t\xi_t + b_t)\mathrm{d}t) - \frac{1}{2}\int_0^T \|h(t,\xi_t)\|^2\mathrm{d}t}]. \end{aligned} \tag{10}$$

Given a sequence of partitions $(\Delta_n)_{n\ge1}$ of $[0, T]$:

$$\Delta_n = \{t_0^n < t_1^n < \cdots < t_{k_n}^n = T\}$$

with $|\Delta_n| = \max_{0\le i\le k_n-1}(t_{i+1}^n - t_i^n) \to 0$, and a continuous stochastic process X, we define:

$$S_n(X) = \sum_{i=0}^{k_n-1} h^*(t_i^n, X_{t_i^n})\sigma_{t_i^n}^+(X_{t_{i+1}^n} - X_{t_i^n}).$$

Then

$$E[|S_n(\xi) - S_m(\xi)| \wedge 1] = \int_{\mathbb{R}^d} E[|S_n(p^v) - S_m(p^v)| \wedge 1]\mu_T(\mathrm{d}v),$$

[2] $M^+ = (M^*M)^{-1}M^*$ and the symmetric matrix is inverted by diagonalisation with $1/0 = 0$.

which implies that $S_n(p^v)$ converges in probability $P \otimes \mu_T$. Hence, we can define $\int_0^T h^*(t, p_t^v)\sigma_t^+ \mathrm{d}p_t^v$ as the limit (in probability $P \otimes \mu_T$) of the sequence $S_n(p^v)$. Obviously, this limit is independent of the sequence of partitions $(\Delta_n)_n$ which satisfies $|\Delta_n| \to 0$.

Finally, defining the continuous function $\Theta_N(x) = N \wedge x$, $x \geq 0$, we have

$$\begin{aligned} &E[\Theta_N(f(\xi)g(\xi_T)\mathrm{e}^{S_n(\xi)-\int_0^T h^*(t,\xi_t)\sigma_t^+(A_t\xi_t+b_t)\mathrm{d}t-\frac{1}{2}\int_0^T \|h(t,\xi_t)\|^2\mathrm{d}t})] \\ &= \int_{\mathbb{R}^d} E[\Theta_N(f(p^v)\mathrm{e}^{S_n(p^v)-\int_0^T h^*(t,p_t^v)\sigma_t^+(A_t p_t^v+b_t)\mathrm{d}t-\frac{1}{2}\int_0^T \|h(t,p_t^v)\|^2\mathrm{d}t} g(v))]\mu_T(\mathrm{d}v). \end{aligned}$$

Taking the limit first in n and then in N, and returning to (10), we deduce:

$$\begin{aligned} E[f(x)g(x_T)] = \int_{\mathbb{R}^d} &E[f(p^v)\mathrm{e}^{\int_0^T h^*(t,p_t^v)(\sigma_t^+\mathrm{d}p_t^v-\sigma_t^+(A_t p_t^v+b_t)\mathrm{d}t)-\frac{1}{2}\int_0^T \|h(t,p_t^v)\|^2\mathrm{d}t}] \\ &\times g(v)\mu_T(\mathrm{d}v) \end{aligned}$$

which implies (9) and C is the value of the density of μ_T with respect to the distribution of x_T at v. □

As the Brownian bridge, we have:

Proposition 3. *Let us assume that* $M_t = \int_t^T P_u^{-1}\sigma_u\sigma_u^* P_u^{-*}\mathrm{d}u$ *is positive definite for any* $t \in [0, T)$. *Then the distribution of the process* p *is the same as that of* q *which is the solution to the following linear SDE*

$$\mathrm{d}q_t = A_t q_t \mathrm{d}t + b_t \mathrm{d}t + \sigma_t\sigma_t^* P_t^{-*} M_t^{-1}(P_t^{-1}(E[\xi_t] - q_t) - P_T^{-1}(E[\xi_T] - v))\mathrm{d}t + \sigma_t \mathrm{d}w_t, \tag{11}$$

with $q_0 = u$.

Proof. The matrix $Q_t = P_t M_t$ is the solution to $\dot{Q}_t = (A_t - \sigma_t\sigma_t^* P_t^{-*} M_t^{-1} P_t^{-1})Q_t$, implying that the covariance of q_t can be rewritten as follows: for $s < t$,

$$\begin{aligned} Q_s \int_0^s Q_u^{-1}\sigma_u\sigma_u^* Q_u^{-*}\mathrm{d}u\ Q_t^* &= Q_s \int_0^s M_u^{-1} P_u^{-1}\sigma_u\sigma_u^* P_u^{-*} M_u^{-1}\mathrm{d}u\ Q_t^* \\ &= Q_s\left(M_s^{-1} - M_0^{-1}\right)Q_t^* \\ &= P_s M_s (M_s^{-1} - M_0^{-1}) M_t P_t^* \\ &= P_s (M_0 - M_s)(Id - M_0^{-1}(M_0 - M_t))P_t^* \\ &= C_{st}. \end{aligned}$$

On the other hand, from (8), the expectation $\bar{p}_t$ of the process p_t satisfies

$$\frac{\mathrm{d}}{\mathrm{d}t}\bar{p}_t - A_t\bar{p}_t - b_t = -\sigma_t\sigma_t^* P_t^{-*} P_T^* R_{TT}^{-1}(E[\xi_T] - v).$$

Elementary algebra shows $P_T^* R_{TT}^{-1} = -Q_t^{-1}(R_{tT} R_{TT}^{-1} - P_t P_T^{-1})$, hence

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}\bar{p}_t - A_t\bar{p}_t - b_t &= \sigma_t\sigma_t^* P_t^{-*} Q_t^{-1}(R_{tT} R_{TT}^{-1} - P_t P_T^{-1})(E[\xi_T] - v) \\ &= \sigma_t\sigma_t^* P_t^{-*} Q_t^{-1}(E[\xi_t] - \bar{p}_t - P_t P_T^{-1}(E[\xi_T] - v)) \end{aligned}$$

which is the equation satisfied by $E[q_t]$. The conclusion follows by noting that both p and q are Gaussian processes. □

Remark. M_t is positive definite for any $t \in [0, T)$ if and only if the pair of functions (A, σ) is controllable on $[t, T]$ for any $t \in [0, T)$. See, e.g. [6] for some discussions.

Example. Consider the two-dimensional stochastic differential equation defined by (6) and (7), where $\sigma \neq 0$. Let us assume that b is locally Lipschitz with respect to (x, y), and this equation admits a strong solution (the strong solution exists if there exists a Lyapunov function, see, e.g. [1]). Then we have:

$$\begin{aligned} &E[f(x, y)|(x_0, y_0) = u, (x_T, y_T) = v] \\ &\quad = CE\left[f(p, q)\mathrm{e}^{\sigma^{-2}\int_0^T b(t,p_t,q_t)\mathrm{d}q_t - \frac{1}{2\sigma^2}\int_0^T b(t,p_t,q_t)^2\mathrm{d}t}\right] \end{aligned} \tag{12}$$

where (p, q) is the following bridge starting from $(p_0, q_0) = u$:

$$\begin{pmatrix} p_t \\ q_t \end{pmatrix} = \begin{pmatrix} z_t \\ \dot{z}_t \end{pmatrix} - \frac{t}{T^3}\begin{pmatrix} t(3T - 2t) & -tT(T - t) \\ 6(T - t) & T(3t - 2T) \end{pmatrix}\begin{pmatrix} z_T - v_1 \\ \dot{z}_T - v_2 \end{pmatrix}, \tag{13}$$

$$\text{with} \quad z_t = u_1 + tu_2 + \sigma\int_0^t w_s\mathrm{d}s, \quad \dot{z}_t = u_2 + \sigma w_t;$$

or (p, q) can be chosen as:

$$\begin{aligned} \mathrm{d}p_t &= q_t\mathrm{d}t, \\ \mathrm{d}q_t &= \left(-6\frac{p_t - v_1}{(T - t)^2} - 2\frac{2q_t + v_2}{T - t}\right)\mathrm{d}t + \sigma\mathrm{d}w_t. \end{aligned} \tag{14}$$

4. σ invertible, general b

4.1. Bounded drift

Let us consider the following SDEs:

$$\mathrm{d}x_t = b(t, x_t)\mathrm{d}t + \sigma(t, x_t)\mathrm{d}w_t, \quad x_0 = u, \tag{15}$$

$$\mathrm{d}y_t = b(t, y_t)\mathrm{d}t - \frac{y_t - v}{T - t}\mathrm{d}t + \sigma(t, y_t)\mathrm{d}w_t, \quad y_0 = u. \tag{16}$$

Remark. If $b = 0$, and $\sigma = Id$, then x is a Brownian motion. It is well known (see, e.g. [6]) that the law of the Brownian motion x conditioned on $x_T = v$ is the same as that of the Brownian bridge y satisfying the following SDE:

$$\mathrm{d}y_t = -\frac{y_t - v}{T - t}\mathrm{d}t + \mathrm{d}w_t, \quad y_0 = u.$$

The form of SDE (16) is inspired by the above SDE in order to fit the simplest case: the Brownian bridge case.

The objective of this section is to prove that the distribution of x (solution of (15)) conditioned on $x_T = v$ is absolutely continuous with respect to y (solution of (16)) with an explicit density. We shall assume some regularity conditions on b and σ here.

Assumption 4.1. The functions $b(t, x)$ and $\sigma(t, x)$ are $C^{1,2}$ with values in $\mathbb{R}^d$ and $\mathbb{R}^{d\times d}$ respectively; and the functions b, σ, together with their derivatives, are bounded. Moreover, σ is invertible with a bounded inverse.

Let $x^{s,u}$ be the solution of (15) starting at $s \in [0, T]$. Under Assumption 4.1, x is a strong Markov process with positive transition density. For $(s, u) \in [0, T]$, we denote $p(s, u; t, z)$ to be the density of $x_t^{s,u}$. Then there exist constants $m, \lambda, M, \Lambda > 0$, such that the density function $p(s, u; t, z)$ satisfies Aronson's estimation [2]: for $t > s$,

$$m(t-s)^{-\frac{d}{2}} \mathrm{e}^{-\frac{\lambda|z-u|^2}{t-s}} \le p(s, u; t, z) \le M(t-s)^{-\frac{d}{2}} \mathrm{e}^{-\frac{\Lambda|z-u|^2}{t-s}}.$$

We first study SDE (16).

Lemma 4. *Let Assumption 4.1 hold. Then the SDE (16) admits a unique solution on* $[0, T)$. *Moreover,* $\lim_{t\to T} y_t = v$, *a.s. and* $|y_t - v|^2 \le C(T-t)\log\log[(T-t)^{-1} + e]$, *a.s., where* C *is a positive random variable.*

Proof. The fact that the SDE (16) admits a unique solution on $[0, T)$ is classic. Applying Itô's formula to $\frac{y_t - v}{T-t}$, we deduce easily the following:

$$\frac{y_t - v}{T-t} = \frac{u-v}{T} + \int_0^t (T-s)^{-1} b(s, y_s)\mathrm{d}s + \int_0^t (T-s)^{-1}\sigma(s, y_s)\mathrm{d}w_s.$$

For each i, $\{(\int_0^t (T-s)^{-1}\sigma(s, y_s)\mathrm{d}w_s)_i, t \ge 0\} = \{\sum_{j=1}^d \int_0^t (T-s)^{-1}\sigma_{ij}(s, y_s)\mathrm{d}w_s^j, t \ge 0\}$ is a continuous local martingale, and its quadratic variation process $\tau_t = \int_0^t \sum_{j=1}^d (T-s)^{-2}\sigma_{ij}^2(s, y_s)\mathrm{d}s$ satisfies $\tau_t \to \infty$ as $t \to T$, and $\tau_t \le \frac{c}{T-t}$ for a constant $c > 0$. Applying Dambis–Dubins–Schwarz's theorem, for each i, there exists a standard one-dimensional Brownian motion B^i, such that

$$\left(\int_0^t (T-s)^{-1}\sigma(s, y_s)\mathrm{d}w_s\right)_i = B^i(\tau_t), \qquad t \ge 0.$$

Taking into consideration of the law of the iterated logarithm for the Brownian motion B^i, the conclusion follows easily. □

Now we can state the main theorem of this section.

Theorem 5. *Let Assumption 4.1 hold. Then*

$$E[f(x)|x_T = v] = CE\left[f(y)\exp\left\{-\int_0^T \frac{2\tilde{y}_t^* A_t(y_t)b_t(y_t)\mathrm{d}t + \tilde{y}_t^*(\mathrm{d}A_t(y_t))\tilde{y}_t + \sum_{ij}\mathrm{d}\langle A_t^{ij}(y_t), \tilde{y}_t^i\tilde{y}_t^j\rangle}{2(T-t)}\right\}\right] \tag{17}$$

where $A(t, y) = (\sigma(t, y)^*)^{-1}\sigma(t, y)^{-1}$, $\tilde{y}_t = y_t - v$, *and* $\langle\cdot,\cdot\rangle$ *is the quadratic variation of semimartingales.*

Remark. From Lemma 4, the integral in (17) is well defined.

Proof. Let $f(x)$ be an $\mathcal{F}_t$-measurable non-negative function, $t < T$, then

$$E[f(y)] = E\left[f(x)\exp\left\{-\int_0^t \frac{(\sigma_s^{-1}(x_s)(x_s-v))^*}{T-s}\mathrm{d}w_s - \frac{1}{2}\int_0^t \left\|\frac{\sigma_s^{-1}(x_s)(x_s-v)}{T-s}\right\|^2 \mathrm{d}s\right\}\right]. \tag{18}$$

On the other hand, Itô's formula gives:

$$\mathrm{d}\,\frac{\|\sigma^{-1}(t,x_t)(x_t-v)\|^2}{T-t} = 2\frac{(x_t-v)^*A(t,x_t)\mathrm{d}x_t}{T-t} + \frac{\|\sigma^{-1}(t,x_t)(x_t-v)\|^2}{(T-t)^2}\mathrm{d}t + \frac{\mathrm{d}\cdot\mathrm{d}t}{T-t} + \frac{(x_t-v)^*(\mathrm{d}A(t,x_t))(x_t-v)}{T-t} + \frac{\sum_{ij}\mathrm{d}\langle A^{ij}(t,x_t),(x_t^i-v^i)(x_t^j-v^j)\rangle}{T-t}.$$

Combining the above equation with (18), we deduce that,

$$E[f(y)] = CC_tE\left[f(x)\exp\left\{-\frac{\|\sigma_t^{-1}(x_t)(x_t-v)\|^2}{2(T-t)} + \int_0^t \frac{(x_s-v)^*A_s(x_s)b_s(x_s)}{T-s}\mathrm{d}s + \frac{1}{2}\int_0^t \frac{(x_s-v)^*(\mathrm{d}A_s(x_s))(x_s-v)}{T-s} + \frac{\sum_{ij}\mathrm{d}\langle A_s^{ij}(x_s),(x_s^i-v^i)(x_s^j-v^j)\rangle}{T-s}\right\}\right],$$

where $C > 0$ is a constant, and $C_t = (T-t)^{-\frac{d}{2}}$.

Or equivalently,

$$E[f(y)\varphi_t] = CC_tE\left[f(x)\exp\left\{-\frac{\|\sigma(t,x_t)^{-1}(x_t-v)\|^2}{2(T-t)}\right\}\right], \tag{19}$$

where

$$\varphi_t = \exp\left\{-\int_0^t \frac{\tilde{y}_s^*A_s(y_s)b_s(y_s)}{T-s}\mathrm{d}s - \frac{1}{2}\int_0^t \frac{\tilde{y}_s^*(\mathrm{d}A_s(y_s))\tilde{y}_s}{T-s} + \frac{\sum_{ij}\mathrm{d}\langle A_s^{ij}(y_s),\tilde{y}_s^i\tilde{y}_s^j\rangle}{T-s}\right\}. \tag{20}$$

Note that $\{\varphi_t, t\in[0,T]\}$ is a well defined continuous process, thanks to Lemma 4.

Putting $f = 1$ in (19), we deduce then:

$$\frac{\mathbb{E}[f(y)\varphi_t]}{\mathbb{E}[\varphi_t]} = \frac{E\left[f(x)\exp\left\{-\frac{\|\sigma(t,x_t)^{-1}(x_t-v)\|^2}{2(T-t)}\right\}\right]}{E\left[\exp\left\{-\frac{\|\sigma(t,x_t)^{-1}(x_t-v)\|^2}{2(T-t)}\right\}\right]}. \tag{21}$$

Assuming that $f(x)$ takes the form $f(x) = g(x_{t_1},\ldots,x_{t_N})$, $0 < t_1 < t_2 < \cdots < t_N < T$, $g \in C_b(\mathbb{R}^{Nd})$, and letting $t \to T$, from the Lemmas 7 and 8 in the Appendix, we get:

$$\frac{\mathbb{E}[f(y)\varphi_T]}{\mathbb{E}[\varphi_T]} = \mathbb{E}[f(x)|x_T = v].$$

This completes the proof of the theorem. □

Remark. For practical implementation, it is useful to note that the second and third terms of the integral in (17) are the limit of $\sum \tilde{y}^*_{t_k}(A(t_k, y_{t_k}) - A(t_{k-1}, y_{t_{k-1}}))\tilde{y}_{t_k}\frac{1}{2(T-t_k)}$.

4.2. Unbounded drift

Let us now consider the following SDE:

$$\mathrm{d}x_t = b(t, x_t)\mathrm{d}t + \sigma(t, x_t)\mathrm{d}w_t, \quad x_0 = u, \tag{22}$$

where the drift b can be unbounded. We assume instead

Assumption 4.2. The function $\sigma(t, x)$ is $C^{1,2}$ with values in $\mathbb{R}^{d\times d}$; the function σ together with its derivatives are bounded; and σ is invertible with a bounded inverse. The function b is locally Lipschitz with respect to x and is locally bounded. Moreover, the SDE (22) admits a strong solution.

Combining the Theorems 1 and 5, we are able to prove the following

Theorem 6. *Let Assumption 4.2 hold, and y be the solution of*

$$\mathrm{d}y_t = -\frac{y_t - v}{T-t}\mathrm{d}t + \sigma(t, y_t)\mathrm{d}w_t, \quad y_0 = u. \tag{23}$$

Then,

$$E[f(x)|x_T = v] = CE\left[f(y)\exp\left\{-\int_0^T \frac{\tilde{y}^*_t(\mathrm{d}A(t,y_t))\tilde{y}_t + \sum_{ij}\mathrm{d}\langle A^{ij}(t,y_t), \tilde{y}^i_t\tilde{y}^j_t\rangle}{2(T-t)}\right.\right.$$
$$\left.\left. + \int_0^T (b^*A)(t, y_t)\mathrm{d}y_t - \frac{1}{2}\int_0^T |\sigma^{-1}b|^2(t, y_t)\mathrm{d}t\right\}\right],$$

where $A(t, y) = \sigma(t, y)^{-*}\sigma(t, y)^{-1}$, $\tilde{y}_t = y_t - v$, and $\langle\cdot,\cdot\rangle$ is the quadratic variation of semimartingales.

Proof. Let $\bar{x}$ be the solution of:

$$\mathrm{d}\bar{x}_t = \sigma(t, \bar{x}_t)\mathrm{d}w_t, \quad \bar{x}_0 = u. \tag{24}$$

Then, from Theorem 1, for non-negative measurable functions f and g,

$$\begin{aligned}\mathbb{E}[f(x)g(x_T)] &= \mathbb{E}[f(\bar{x})g(\bar{x}_T)\mathrm{e}^{\int_0^T (b^*A)(t,\bar{x}_t)\mathrm{d}\bar{x}_t - \frac{1}{2}\int_0^T |\sigma^{-1}b|^2(t,\bar{x}_t)\mathrm{d}t}]\\ &= \int_{\mathbb{R}^d} E[f(\bar{x})\mathrm{e}^{\int_0^T (b^*A)(t,\bar{x}_t)\mathrm{d}\bar{x}_t - \frac{1}{2}\int_0^T |\sigma^{-1}b|^2(t,\bar{x}_t)\mathrm{d}t}|\bar{x}_T = v]g(v)\mathrm{d}v.\end{aligned}$$

It remains to apply Theorem 5. □

Remark. If the drift b is bounded, both formulas in Theorems 5 and 6 are available. Unfortunately, it is difficult to compare the efficiency of simulation when applying these two formulas.

5. Example: Parameter estimation for a Duffing–Van der Pol oscillator

Consider the Van der Pol process: [1]

$$\mathrm{d}x_t = y_t \mathrm{d}t$$
$$\mathrm{d}y_t = -y_t \mathrm{d}t + (\theta_* - x_t^2)x_t \mathrm{d}t + \sigma \mathrm{d}w_t = b(x_t, y_t)\mathrm{d}t + \sigma \mathrm{d}w_t, \quad \sigma \neq 0.$$

It is easy to check that the function $V(x, y) = \frac{1}{2}x^4 + y^2 + 1$ is a Lyapunov function for this system. Hence, a strong solution exists (see, e.g. [5]) and Theorem 2 can be applied.

Assume that the process is observed at instants $\{t_1, \ldots, t_n\} \subset [0, T]$; the observation set is thus $Z = \{(x_{t_1}, y_{t_1}), \ldots, (x_{t_n}, y_{t_n})\}$. The statistical problem here is the estimation of θ_*. The maximum likelihood estimate is:

$$\hat{\theta} = \arg\max_{\theta} P_\theta(Z),$$

where $P_\theta(Z)$ stands for the density of the distribution of Z in $\mathbb{R}^{2n}$. There is no simple form for this function; on the other hand the density $p_\theta(x, y)$ of the law of (x, y) with respect to the law of the process corresponding to the equation with $b = 0$ (Girsanov formula (12)) is easily derived:

$$\begin{aligned}\sigma^2 \log p_\theta(x, y) &= \int_0^T b(x_t, y_t)\mathrm{d}y_t - \frac{1}{2}\int_0^T b(x_t, y_t)^2 \mathrm{d}t \\ &= \int_0^T (-y_t + (\theta - x_t^2)x_t)\mathrm{d}y_t - \frac{1}{2}\int_0^T (-y_t + (\theta - x_t^2)x_t)^2 \mathrm{d}t.\end{aligned}$$

As a consequence, the maximum likelihood estimation of θ_* where the whole trajectory is observed is easy. A standard approach for the computation of $\hat{\theta}$ is to use the EM algorithm (see, e.g. [3]) where the full trajectory is considered as missing data. Let us explain informally in a few words how it is derived: One can show that

$$\hat{\theta} = \arg\max_{\theta} E_{\hat{\theta}}[\log p_\theta(x, y)|Z]. \tag{25}$$

Indeed, for any θ,

$$\begin{aligned}E_{\hat{\theta}}[\log p_\theta(x, y)|Z] - E_{\hat{\theta}}[\log p_{\hat{\theta}}(x, y)|Z] &= E_{\hat{\theta}}[\log(p_\theta(x, y)/p_{\hat{\theta}}(x, y))|Z] \\ &\leq \log E_{\hat{\theta}}[p_\theta(x, y)/p_{\hat{\theta}}(x, y)|Z] \\ &= \log P_\theta(Z)/P_{\hat{\theta}}(Z),\end{aligned} \tag{26}$$

since the "conditional expectation of the density is the density of the marginal". The EM algorithm consists in the computation of the solution to (25) as the limit of the sequence

$$\hat{\theta}_{j+1} = \arg\max_{\theta} E_{\hat{\theta}_j}[\log p_\theta(x, y)|Z].$$

The proof of the convergence of this sequence to $\hat{\theta}$ is beyond the scope of this paper. In our situation these expectations cannot be directly calculated but can be estimated through Monte Carlo simulations, and we get the new sequence

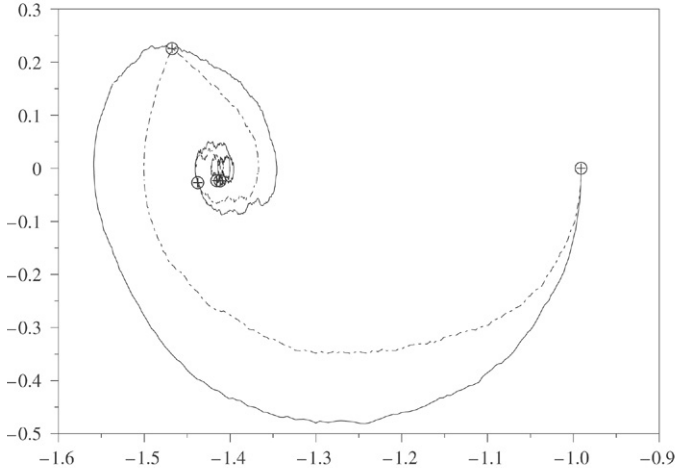

Fig. 1. $\sigma = 0.02, n = 5, \hat{\theta} = 1.98$.

$$\check{\theta}_{j+1} = \arg\max_{\theta} \sum_{k=1}^{K} \log p_\theta(x^k, y^k),$$

where (x^k, y^k) are independent realizations of the distribution of (x, y) conditioned on Z under $\check{\theta}_j$. Finally, according to the results of Section 3, we obtain the algorithm

$$\theta_{j+1} = \arg\max_{\theta} \sum_{k=1}^{K} \log(p_\theta(p^k, q^k))\, p_{\theta_j}(p^k, q^k), \tag{27}$$

where (p^k, q^k) are simulated according to (13) or (14). The maximization in (27) is immediate:

$$\theta_{j+1} = \frac{\langle \int_0^T p_t \mathrm{d}q_t + \int_0^T p_t (q_t + p_t^3) \mathrm{d}t \rangle_j}{\langle \int_0^T p_t^2 \mathrm{d}t \rangle_j}, \tag{28}$$

where $\langle . \rangle_j$ stands for the weighted mean over the K samples with weights $p_{\theta_j}(p^k, q^k)$.

Remark. Instead of the EM algorithm, we could apply the Bayesian approach (see, e.g. [8]) which would lead to an analogous MCMC method.

We present now the results of three typical estimation runs with different parameters (we do not pretend here to make any precise study of convergence). For these three experiments, we chose

$$x_0 = (-1, 0), \quad T = 10, \quad t_i = \frac{i-1}{n-1} T, \quad h = 0.01, \quad K = 100, \quad J = 10, \quad \theta_0 = 4,$$
$$\theta = 2.$$

h is the step in the Euler scheme used for calculating (28). We performed the estimation of θ with two different values of σ or n; the true trajectory is represented in plain style, the observation points are $\oplus$ and the conditionally simulated process is represented with dashed style (see Figs. 1–3).

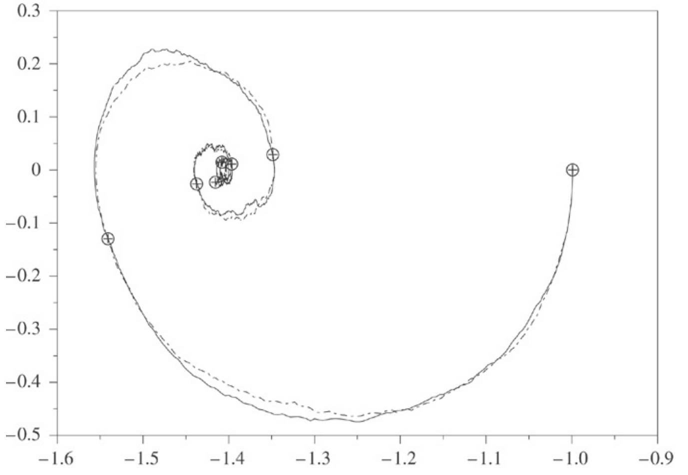

Fig. 2. $\sigma = 0.02, n = 7, \hat{\theta} = 1.993$.

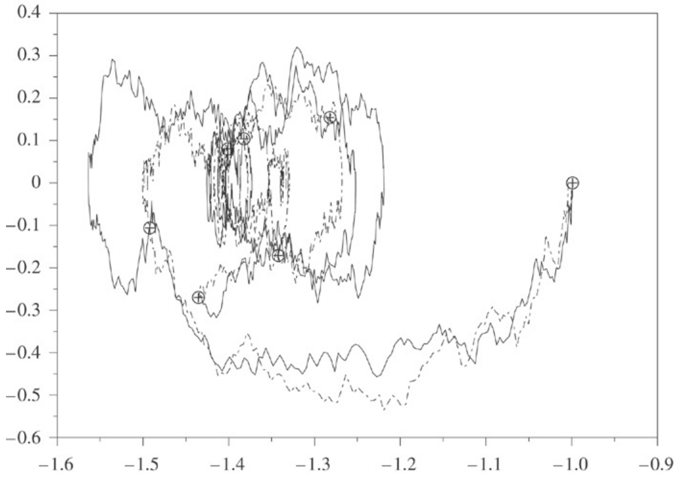

Fig. 3. $\sigma = 0.2, n = 7, \hat{\theta} = 1.9$.

Acknowledgment

The authors would like to thank an anonymous referee for his helpful comments.

Appendix

Lemma 7. *Let* $0 < t_1 < t_2 < \cdots < t_N < T$, *and* $g \in C_b(\mathbb{R}^{Nd})$. *Then, putting*

$$\psi_t = \exp\left\{-\frac{\|\sigma(t, x_t)^{-1}(x_t - v)\|^2}{2(T - t)}\right\},$$

$$\lim_{t\to T} \frac{\mathbb{E}[g(x_{t_1}, x_{t_2}, \ldots, x_{t_N})\psi_t]}{\mathbb{E}[\psi_t]} = \mathbb{E}[g(x_{t_1}, x_{t_2}, \ldots, x_{t_N})|x_T = v]. \tag{29}$$

Proof. For any $t \in (t_N, T)$,

$$\frac{\mathbb{E}[g(x_{t_1}, x_{t_2}, \ldots, x_{t_N})\psi_t]}{\mathbb{E}[\psi_t]} = \frac{\int_{\mathbb{R}^d} \Phi_g(t,z) \exp\left\{-\frac{\|\sigma(t,z)^{-1}(z-v)\|^2}{2(T-t)}\right\} \mathrm{d}z}{\int_{\mathbb{R}^d} \Phi_1(t,z) \exp\left\{-\frac{\|\sigma(t,z)^{-1}(z-v)\|^2}{2(T-t)}\right\} \mathrm{d}z},$$

where

$$\Phi_g(t,z) = \int_{\mathbb{R}^{Nd}} g(z_1, \ldots, z_N) p(0, u; t_1, z_1) \cdots p(t_N, z_N; t, z) \mathrm{d}z_1 \cdots \mathrm{d}z_N,$$

which is continuous thanks to Aronson's estimation. Evidently, $\Phi_1(t,z) = p(0, u; t, z)$.

Moreover, applying a simple change of variable $z = v + (T-t)^{\frac{1}{2}} z'$,

$$\begin{aligned}
&(T-t)^{-\frac{d}{2}} \int_{\mathbb{R}^d} \Phi_g(t,z) \exp\left\{-\frac{\|\sigma(t,z)^{-1}(z-v)\|^2}{2(T-t)}\right\} \mathrm{d}z \\
&= \int_{\mathbb{R}^d} \Phi_g(t, v + (T-t)^{\frac{1}{2}} z') \exp\left\{-\frac{\|\sigma(t, v + (T-t)^{\frac{1}{2}} z')^{-1} z'\|^2}{2}\right\} \mathrm{d}z' \\
&\to \Phi_g(T, v) \int_{\mathbb{R}^d} \exp\left\{-\frac{\|\sigma(T,v)^{-1} z'\|^2}{2}\right\} \mathrm{d}z'.
\end{aligned}$$

Hence,

$$\lim_{t\to T} \frac{\mathbb{E}[g(x_{t_1}, x_{t_2}, \ldots, x_{t_N})\psi_t]}{\mathbb{E}[\psi_t]} = \frac{\Phi_g(T,v)}{\Phi_1(T,v)},$$

from which we deduce (29) by the Bayes formula, since

$$\Phi_g(T,v) = \int_{\mathbb{R}^{Nd}} g(z_1, \ldots, z_N) q(z_1, \ldots, z_N, v) \mathrm{d}z_1 \cdots \mathrm{d}z_N,$$

where q is the density of $(x_{t_1}, \ldots, x_{t_N}, x_T)$. □

Lemma 8.

$$\lim_{t\to T} E[|\varphi_t - \varphi_T|] = 0.$$

We need the following two propositions to prove this lemma.

Proposition 9. (i) *There exist two constants $c_1 > 0$, $c_2 > 0$, such that*

$$c_1 \le C_t E[\psi_t] \le c_2, \forall t \in [0, T),$$

where

$$C_t = (T-t)^{-\frac{d}{2}}.$$

(ii) *There exists a constant $c_3 > 0$, such that*

$$E[\varphi_t] \le c_3, \quad \forall t \in [0, T).$$

Proof. (i) We note that

$$C_t\mathbb{E}[\psi_t] = (T-t)^{-\frac{d}{2}} \int \exp\left\{-\frac{\|\sigma(t,z)^{-1}(z-v)\|^2}{2(T-t)}\right\} p(0,u;t,z)\mathrm{d}z.$$

We get easily the conclusion taking into consideration Aronson's estimation after a change of variable $z = v + (T-t)^{\frac{1}{2}}z'$.

(ii) It follows from (19) and (i). □

Proposition 10. *For any $\varepsilon > 0$, there exists an adapted bounded process α_t such that*

$$\mathrm{d}C_t\psi_t = \mathrm{d}M_t + \alpha_t(C_t\psi_t)^{1-\varepsilon}(T-t)^{-h}\mathrm{d}t, \quad h = \frac{\varepsilon d + 1}{2}$$

where $(M_t)_{0\le t<T}$ is a martingale.

Proof. Set $\tilde{x}_t = x_t - v$, $p_t = \|\sigma^{-1}(t,x_t)\tilde{x}_t\|$, and $A_t = \sigma^{-*}(t,x_t)\sigma^{-1}(t,x_t)$. We have

$$\begin{aligned}
\mathrm{d}\frac{p_t^2}{T-t} &= 2\frac{\tilde{x}_t^* A_t \mathrm{d}x_t}{T-t} + \frac{p_t^2}{(T-t)^2}\mathrm{d}t + \frac{\mathrm{d}}{T-t}\mathrm{d}t + \frac{\tilde{x}_t^*(\mathrm{d}A_t)\tilde{x}_t}{T-t} + \frac{1}{T-t}\sum_{i,j}\mathrm{d}\langle A_t^{ij}, \tilde{x}_t^i\tilde{x}_t^j\rangle \\
&= 2\frac{\tilde{x}_t^*\sigma(t,x_t)^{-*}\mathrm{d}w_t}{T-t} + \frac{p_t^2}{(T-t)^2}\mathrm{d}t + \frac{\mathrm{d}}{T-t}\mathrm{d}t + r_t\frac{p_t^2+p_t}{T-t}\mathrm{d}t + \frac{p_t^2}{T-t}r_t'\mathrm{d}w_t,
\end{aligned}$$

where r_t and r_t' are two adapted bounded processes. Hence we get:

$$\begin{aligned}
\mathrm{d}C_t\psi_t &= \frac{\mathrm{d}}{2(T-t)}C_t\psi_t\mathrm{d}t - \frac{1}{2}C_t\psi_t\mathrm{d}\left(\frac{p_t^2}{T-t}\right) + \frac{1}{8}C_t\psi_t\mathrm{d}\left\langle\frac{p_t^2}{T-t}\right\rangle \\
&= \mathrm{d}M_t + C_t\psi_t r_t''\left(\frac{p_t^2+p_t}{T-t} + \frac{p_t^4+p_t^3}{(T-t)^2}\right)\mathrm{d}t,
\end{aligned}$$

where r_t'' is an adapted bounded process. For any $\varepsilon > 0$, $\mathrm{e}^{-\varepsilon\frac{x^2}{2}}|x|^k$, $k = 1, 2, 3, 4$, are all bounded functions, then there exists a constant $c_\varepsilon > 0$ such that

$$\psi_t^\varepsilon\left(\frac{p_t^2+p_t}{T-t} + \frac{p_t^4+p_t^3}{(T-t)^2}\right) \le \frac{c_\varepsilon}{\sqrt{T-t}}.$$

Hence,

$$\mathrm{d}C_t\psi_t = \mathrm{d}M_t + (C_t\psi_t)^{1-\varepsilon}(T-t)^{-h}r_t'''c_\varepsilon\mathrm{d}t,$$

where r_t''' is still an adapted bounded process. □

Let us now return to the proof of Lemma 8.

Proof. First, from Fatou's lemma and Proposition 9,

$$E[\varphi_T] \le \liminf_{t\to T} E[\varphi_t] \le c_3.$$

We choose $t_0 \in (0,T)$ which is close enough to T, and A large enough, and put

$$\sigma = \inf\left\{t_0 < t < T, C_t\psi_t \le \frac{1}{A}\right\} = \inf\left\{t_0 < t < T, p_t^2 \ge 2(T-t)\log\frac{A}{(T-t)^{\frac{\mathrm{d}}{2}}}\right\}.$$

Under the distribution of $x_.$, $\sigma < T$ a.s. However under the distribution of $y_.$, $\lim_{A\to+\infty} \sigma = T$, a.s., taking into consideration of Lemma 4. We have, from (21),

$$\frac{E[\varphi_t 1_{\sigma<t}]}{E[\varphi_t]} = \frac{E[\psi_t 1_{\sigma<t}]}{E[\psi_t]} \leq \frac{1}{c_1} E[C_t \psi_t 1_{\sigma<t}].$$

On the other hand, from Proposition 10 with a fixed $\varepsilon \in (0, 1/d)$,

$$\mathrm{d}C_t\psi_t = \mathrm{d}M_t + \alpha_t (C_t\psi_t)^{1-\varepsilon}(T-t)^{-h}\mathrm{d}t,$$

i.e.,

$$C_t\psi_t = C_\sigma\psi_\sigma + M_t - M_\sigma + \int_\sigma^t \alpha_s (C_s\psi_s)^{1-\varepsilon}(T-s)^{-h}\mathrm{d}s.$$

Hence,

$$E[C_t\psi_t 1_{\sigma<t}] \leq A^{-1} + \bar\alpha \int_{t_0}^t E[C_s\psi_s 1_{\sigma<s}]^{1-\varepsilon}(T-s)^{-h}\mathrm{d}s, \quad \text{with } \bar\alpha = \sup_t \|\alpha_t\|_\infty.$$

Therefore, $E[C_t\psi_t 1_{\sigma<t}]$ is bounded by u_t, which is the solution of the following differential equation,

$$\mathrm{d}u_t = \bar\alpha u_t^{1-\varepsilon}(T-t)^{-h}\mathrm{d}t, \quad u_{t_0} = A^{-1};$$

and this equation has an explicit solution:

$$u_t = \left\{ \frac{\varepsilon\bar\alpha}{1-h}[(T-t_0)^{1-h} - (T-t)^{1-h}] + A^{-\varepsilon} \right\}^{1/\varepsilon} \leq \left\{ c_0(T-t_0)^{1-h} + A^{-\varepsilon} \right\}^{1/\varepsilon},$$

where $c_0 > 0$ is a constant. We get finally,

$$\begin{aligned} \frac{E[\varphi_t 1_{t\leq\sigma}]}{E[\varphi_t]} &= 1 - \frac{E[\varphi_t 1_{\sigma<t}]}{E[\varphi_t]} \\ &\geq 1 - \frac{1}{c_1}(c_0(T-t_0)^{1-h} + A^{-\varepsilon})^{1/\varepsilon}. \end{aligned}$$

We note that $\{\varphi_t 1_{t\leq\sigma}\}_t$ is a uniformly integrable family due to Novikov's lemma, since we have

$$1_{t\leq\sigma}\varphi_t \leq C \exp\left\{ \int_0^{t\wedge\sigma} \frac{|y_s - v|^2}{T-s} v_s \mathrm{d}w_s - \frac{1}{2}\int_0^{t\wedge\sigma} \frac{|y_s-v|^4}{|T-s|^2}|v_s|^2 \mathrm{d}s \right\},$$

where for fixed A, C is a positive constant and v_t is an adapted bounded process.

Taking the $\liminf_{t\to T}$, we get,

$$\frac{E[\varphi_T 1_{\sigma=T}]}{\limsup\limits_{t\to T} E[\varphi_t]} \geq 1 - \frac{1}{c_1}(c_0(T-t_0)^{1-h} + A^{-\varepsilon})^{1/\varepsilon}.$$

Since $1_{\sigma=T}$ converges to one a.s. as $A\to\infty$, we get

$$\frac{E[\varphi_T]}{\limsup\limits_{t\to T} E[\varphi_t]} \geq 1 - \frac{1}{c_1}(c_0(T-t_0)^{1-h})^{1/\varepsilon}.$$

It remains to let $t_0 \to T$ to get:

$$\limsup_{t\to T} E[\varphi_t] \leq E[\varphi_T].$$

Hence,

$$\lim_{t\to T} E[\varphi_t] = E[\varphi_T],$$

and we finish the proof by Scheffé's lemma (see, e.g. [4]). $\square$

References

[1] L. Arnold, Random Dynamical Systems, Springer, Berlin, 1998.
[2] D.G. Aronson, Bounds for the fundamental solution of a parabolic equation, Bull. Amer. Math. Soc. 73 (1967) 890–896.
[3] G. Casella, C.P. Robert, Monte Carlo Statistical Methods, Springer, New York, 1999.
[4] C. Dellacherie, P.A. Meyer, Probabilités et Potentiels. Chapitre I à IV, Hermann, Paris, 1975.
[5] R.Z. Has'minskii, Stochastic Stability of Differential Equations, Sijthoff and Noordhoff, 1980.
[6] I. Karatzas, S. Shreve, Brownian Motion and Stochastic Calculus, 2nd ed., Springer, New York, 1991.
[7] T.J. Lyons, W.A. Zheng, On conditional diffusion processes, Proc. Roy. Soc. Edinburgh Sect. A 115 (1990) 243–255.
[8] G.O. Roberts, O. Stramer, On inference for partially observed nonlinear diffusion models using the Metropolis–Hastings algorithms, Biometrika 88 (2001) 603–621.