



HAL
open science

Stationary distributions for stochastic differential equations with random effects and statistical applications.

Valentine Genon-Catalot, Catherine Larédo

► **To cite this version:**

Valentine Genon-Catalot, Catherine Larédo. Stationary distributions for stochastic differential equations with random effects and statistical applications.. 2013. hal-00807258v1

HAL Id: hal-00807258

<https://hal.science/hal-00807258v1>

Preprint submitted on 3 Apr 2013 (v1), last revised 7 Dec 2015 (v3)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Stationary distributions for stochastic differential equations with random effects and statistical applications.

V. Genon-Catalot¹, C. Larédo²

¹*Laboratoire MAP5, Université Paris Descartes, U.F.R. de Mathématiques et Informatique, CNRS UMR 8145, PRES Sorbonne Paris Cité, 45, rue des Saints-Pères, 75270 Paris Cedex 06, France.
e-mail: Valentine.Genon-Catalot@parisdescartes.fr.*

²*(Corresponding author) Laboratoire MIA, I.N.R.A. and LPMA, Université Denis Diderot, CNRS-UMR 7599
Laboratoire MIA, I.N.R.A. Domaine de Vilvert, 78350, Jouy-en-Josas, France.
e-mail: catherine.laredo@jouy.inra.fr*

Abstract

Let $(X(t), t \geq 0)$ be defined by a stochastic differential equation including a random effect ϕ in the drift and diffusion coefficients. We characterize the stationary distributions of the joint process $((\phi, X(t)), t \geq 0)$ which are non unique and prove limit theorems and central limit theorems for functionals of the sample path $(X(t), t \in [0, T])$ as T tends to infinity. This allows to build several estimators of the random variable ϕ which are consistent and asymptotically mixed normal with rate \sqrt{T} . Examples are given fulfilling the assumptions of the limit theorems. Parametric estimation of the distribution of the random effect from N *i.i.d.* processes $(X_j(t), t \in [0, T]), j = 1, \dots, N$ is considered. Parametric estimators are built and proved to be \sqrt{N} -consistent and asymptotically Gaussian as both N and $T = T(N)$ tend to infinity with $T(N)/N$ tending to infinity.

April 2, 2013

AMS 2000 subject classification: Primary: 62M05, 62F12 ; secondary: 60J60, 60F05.

Keywords and phrases: Asymptotic mixed normality, consistency, limit theorems, mixed-effects models, stochastic differential equations.

Running title: Stationary distributions for mixed effects SDEs.

1 Introduction

Mixed-effects or random effects models are of common use in many fields of applications and especially in biomedical research. In such models, the distribution of an observation X is given in two stages. First, there is an unobserved random variable ϕ (the random effect) with distribution $\nu(d\phi)$ on some space Φ . Second, given that $\phi = \varphi$, X has distribution $P_{\varphi,\beta}(dx)$ which depends on φ and possibly on some other fixed parameter β . The resulting distribution is thus a mixture $P_{\nu,\beta}(dx) = \int \nu(d\phi)P_{\phi,\beta}(dx)$. The interest of random effects models is that they take into account two sources of variability. Usually, X represents the behaviour of an individual, ϕ describes the individual specificity and given that $\phi = \varphi$, the model describes a general evolution for all individuals with individual value φ (see *e.g.* Beal and Sheiner, 1982, Wolfinger, 1993, Davidian and Giltinan, 1995, Pinheiro and Bates, 2000, Kuhn and Lavielle, 2004, Nie and Yang, 2005, Nie, 2006, 2007).

In particular, in stochastic differential equations (SDEs) with random effects, $X = (X(t), t \geq 0)$ is a continuous-time stochastic process with dynamics given by:

$$dX(t) = b(X(t), \phi, \beta)dt + \sigma(X(t), \phi, \beta) dW(t), \quad X(0) = \eta, \quad (1)$$

where W is a Brownian motion, (ϕ, η) is a random variable independent of W . SDEs with random effects have been introduced to generalize classical ordinary differential equations models (Ditlevsen and de Gaetano, 2005, Overgaard *et al.*, 2005, Donnet and Samson, 2008) and also to model neuronal data (Picchini *et al.*, 2010).

We focus here on one-dimensional SDEs with just random effects, *i.e.* $(X(t))$ is real-valued, W is a standard Wiener process, $b(x, \varphi, \beta) = b(x, \varphi)$, $\sigma(x, \varphi, \beta) = \sigma(x, \varphi)$ are real-valued functions defined on $\mathbb{R} \times \Phi$ and (ϕ, η) takes values in $\Phi \times \mathbb{R}$.

Recently, the estimation of the unknown distribution $\nu(d\phi)$ of ϕ whether parametric or nonparametric based on the observation of N *i.i.d.* processes $(X_j(t), t \in [0, T]), j = 1, \dots, N$ distributed as (1) has been the subject of several contributions. Parametric estimation is investigated in Ditlevsen and de Gaetano, 2005 for the specific model of mixed effects Brownian motion with drift and in Donnet, S. and Samson, A. (2008). In Delattre *et al.*, 2012, the MLE for random effects SDEs is studied more generally for fixed T and N tending to infinity. In particular, for $\phi \in \mathbb{R}$ having a Gaussian distribution with unknown mean μ and unknown variance ω^2 , and with $b(x, \varphi) = \varphi b(x)$, $\sigma(x, \varphi) = \sigma(x)$ with $b(\cdot), \sigma(\cdot)$ known functions, an explicit formula for the exact likelihood is obtained and a complete asymptotic study of the exact MLE of (μ, ω^2) is given. Approximations of the likelihood are also proposed for general mixed SDEs in Picchini *et al.*, 2010, Picchini and Ditlevsen, 2011.

For ϕ real-valued and $\nu(d\phi) = f(\varphi)d\varphi$, nonparametric estimation of the density f is investigated in Comte *et al.*, 2012, under the asymptotic framework that both N and $T = T(N)$ tend to infinity in such a way that $T(N)/N$ tends to infinity. Only two specific models are studied: $b(x, \varphi) = \varphi b(x)$, $b(x, \varphi) = \varphi + b(x)$, $\sigma(x, \varphi) = \sigma(x)$.

In this paper, we first study properties of model (1) for one process X . We prove that the joint process $((\phi, X(t)), t \geq 0)$ is Markov and characterize its stationary distributions which are non unique. Then, we prove limit theorems as T tends to infinity, for functionals of the form $T^{-1} \int_0^T g(X(s))ds$ and $T^{-1} \int_0^T h(X(s))dW(s)$ with g, h given functions. Afterwards, we study the estimation of the random variable ϕ from the observation of one trajectory $(X(t), t \in [0, T])$. Our limit theorems allow us to build several kind of estimators $\hat{\phi}_T$ of ϕ which are consistent and such that $\sqrt{T}(\hat{\phi}_T - \phi)$ converges stably in distribution to a mixed normal distribution. We also build estimators satisfying $\mathbb{E}|\tilde{\phi}_T - \phi|^\gamma \leq CT^{-\gamma/2}$ under appropriate assumptions on the model and the distribution of ϕ .

Finally, we assume that we observe N *i.i.d.* processes $(X_j(t), t \in [0, T], j = 1, \dots, N)$ such that

$$dX_j(t) = b(X_j(t), \phi_j)dt + \sigma(X_j(t), \phi_j) dW_j(t), \quad X_j(0) = \eta_j, j = 1, \dots, N, \quad (2)$$

where (W_1, \dots, W_N) are N independent Wiener processes, $(\phi_1, \eta_1) \dots, (\phi_N, \eta_N)$ are N *i.i.d.* random variables taking values in $\mathbb{R}^d \times \mathbb{R}$, $((\phi_1, \eta_1) \dots, (\phi_N, \eta_N))$ and (W_1, \dots, W_N) are independent. We set a parametric assumption on the common distribution of the ϕ_j 's, *i.e.* $\nu(d\varphi) = f(\theta, \varphi)d\alpha(\varphi)$ with unknown parameters θ . We deduce from the previous sections estimators of θ which are \sqrt{N} -consistent and asymptotically Gaussian under the asymptotic framework that N and $T = T(N)$ tend to infinity with $T(N)/N$ tending also to infinity.

Section 2 is devoted to the model properties of (1) and the limit theorems. Classical examples are given and proved to fulfill the assumptions of the theorems. In Section 3, we focus on the estimation of the random effect ϕ . In Section 4, we consider the estimation of a multidimensional random effect $\phi = (\phi^1, \dots, \phi^d)' \in \mathbb{R}^d$ when $b(x, \varphi) = \sum_{j=1}^d \varphi^j b_j(x)$, $\sigma(x, \varphi) = \sigma(x)$. In Section 5, we consider the parametric estimation of the distribution of the random effects from N *i.i.d.* processes $(X_j(t), t \in [0, T], j = 1, \dots, N)$. Section 6 contains some concluding remarks. Proofs are gathered in Section 7.

2 Model properties.

Consider a real valued stochastic process $(X(t), t \geq 0)$, with dynamics ruled by (1). The Wiener process W and the r.v. (ϕ, η) are defined on a common probability space $(\Omega, \mathcal{F}, \mathbb{P})$ endowed with a filtration $(\mathcal{F}_t, t \geq 0)$ satisfying the usual conditions. The couple (ϕ, η) is assumed to be \mathcal{F}_0 -measurable and W is a $(\mathcal{F}_t, t \geq 0)$ -Brownian motion. We introduce below assumptions ensuring that the process (1) is well defined.

2.1 Existence and uniqueness of strong solutions

It is convenient to look at (1) as a two dimensional process:

$$\begin{aligned} d\phi(t) &= 0, \quad \phi(0) = \phi, \\ dX(t) &= b(X(t), \phi(t))dt + \sigma(X(t), \phi(t))dW(t), \quad X(0) = \eta. \end{aligned}$$

(H1) The real valued functions $(x, \varphi) \rightarrow b(x, \varphi)$ and $(x, \varphi) \rightarrow \sigma(x, \varphi)$ are continuous on $\mathbb{R} \times \mathbb{R}^d$ and such that the above two-dimensional system admits a unique strong solution adapted to the filtration (\mathcal{F}_t) for all initial condition (ϕ, η) .

Under (H1), the process $(\phi, X(t))$ is strong Markov and there exists a measurable functional F

$$(\varphi, x, w.) \in (\mathbb{R}^d \times \mathbb{R} \times C(\mathbb{R}^+, \mathbb{R})) \rightarrow (\varphi, F(\varphi, x, w.)) \in \mathbb{R}^d \times C(\mathbb{R}^+, \mathbb{R}) \quad (3)$$

such that

$$(\phi, X(\cdot)) = (\phi, F(\phi, \eta, W(\cdot))). \quad (4)$$

Standard regularity assumptions ensure the existence and uniqueness of a strong solution for the system. For instance, this holds if either (A) or (B) below are fulfilled (see *e.g.* Comte *et al.*, 2012):

(A) The functions $(x, \varphi) \rightarrow b(x, \varphi)$ and $(x, \varphi) \rightarrow \sigma(x, \varphi)$ are C^1 on $\mathbb{R} \times \mathbb{R}^d$, and such that:
 $\exists K > 0, \forall (x, \varphi) \in \mathbb{R} \times \mathbb{R}^d, \quad |b(x, \varphi)| + |\sigma(x, \varphi)| \leq K(1 + |x| + |\varphi|).$

(B) The functions $(x, \varphi) \rightarrow b(x, \varphi)$ and $(x, \varphi) \rightarrow \sigma(x, \varphi)$ are C^1 on $\mathbb{R} \times \mathbb{R}^d$, and such that:

$$|b'_x(x, \varphi)| + |\sigma'_x(x, \varphi)| \leq L(\varphi), \quad |b'_\varphi(x, \varphi)| + |\sigma'_\varphi(x, \varphi)| \leq L(\varphi)(1 + |x|),$$

with $\varphi \rightarrow L(\varphi)$ continuous.

Under (H1), for all φ and x , the stochastic differential equation

$$dX^{\varphi, x}(t) = b(X^{\varphi, x}(t), \varphi)dt + \sigma(X^{\varphi, x}(t), \varphi)dW(t), \quad X^{\varphi, x}(0) = x \quad (5)$$

admits the unique strong solution process $(X^{\varphi, x}(t), t \geq 0) = (F_t(\varphi, x, W(\cdot)), t \geq 0)$ (see (3)-(4)). By the Markov property of the joint process $(\phi, X(\cdot))$, given $\phi = \varphi, X(0) = x$, the conditional distribution of X is identical to the distribution of $X^{\varphi, x}(\cdot)$. Analogously, denote by μ_φ the conditional distribution of η given $\phi = \varphi$, then the conditional distribution of X given $\phi = \varphi$ is identical to the distribution of the process

$$dX^\varphi(t) = b(X^\varphi(t), \varphi)dt + \sigma(X^\varphi(t), \varphi)dW(t), \quad X^\varphi(0) \sim \mu_\varphi \quad (6)$$

with $X^\varphi(0)$ independent of W (see Lemma 1 below).

2.2 Stationary distributions

We assume (H1) and introduce additional assumptions for stationary distributions.

(H2)-(i) There exists a Borel subset Φ of \mathbb{R}^d and an interval $(\ell, r) \subset \mathbb{R}$ such that $\forall \varphi \in \Phi, \forall x \in (\ell, r), \sigma^2(x, \varphi) > 0$, the scale density of (6)

$$s_\varphi(x) = \exp\left(-2 \int_{x_0}^x \frac{b(u, \varphi)}{\sigma^2(u, \varphi)} du\right), \quad x_0 \in (\ell, r)$$

satisfies $\int_\ell^r s_\varphi(x) dx = \int^r s_\varphi(x) dx = +\infty$ and the speed density $m_\varphi(x)$ of (6) satisfies $M(\varphi) := \int_\ell^r m_\varphi(x) dx < +\infty$.

(H2)-(ii) The function $(x, \varphi) \rightarrow b(x, \varphi)$ is C^1 on $(\ell, r) \times \mathbb{R}^d$ and $(x, \varphi) \rightarrow \sigma(x, \varphi)$ is C^2 on $(\ell, r) \times \mathbb{R}^d$.

(H2)-(iii) The random variable ϕ takes values in Φ . We denote its distribution by $\nu(d\varphi)$.

Let

$$\pi_\varphi(x) = 1_{(\ell, r)}(x) m_\varphi(x) / M(\varphi). \quad (7)$$

This is the unique invariant density of model (6). Denote by $p_t^\varphi(x, y)$ the transition density of model (6). Assumptions (H1)-(H2) imply that

$$(x, y, \varphi) \rightarrow p_t^\varphi(x, y) \quad \text{and} \quad (x, \varphi) \rightarrow \pi_\varphi(x) \quad (8)$$

are measurable. Moreover, for $g : (\ell, r) \rightarrow \mathbb{R}$ a positive Borel function,

$$\varphi \rightarrow \pi_\varphi(g) := \int_\ell^r g(x) \pi_\varphi(x) dx$$

is measurable on Φ . The random variables $\pi_\phi(g)$ are thus well defined.

Proposition 1. *Assume (H1)-(H2). Then, the distribution on $\Phi \times (\ell, r)$ given by*

$$\pi(d\varphi, dx) = \nu(d\varphi) \times \pi_\varphi(x) dx$$

is a stationary distribution for the Markov process $(\phi, X(t))$ given by (1).

Thus, if (ϕ, η) has distribution π , the process $(\phi, X(t))$ is strictly stationary with marginal distribution π . It is worth stressing that the Markov process $(\phi, X(t))$ may admit infinitely many invariant distributions. Hence, it is not ergodic. Note that the marginal distribution of $X(0)$ when $(\phi, X(0))$ has distribution π is a mixture distribution with density given by:

$$p(x) = \int_{\Phi} \nu(d\varphi) \pi_\varphi(x).$$

Example 1. Ornstein-Uhlenbeck process with random effect.

1. Consider $(X(t))$ given by: $dX(t) = -\phi X(t)dt + dW(t)$. For ϕ with distribution $\nu(d\phi)$ supported in $(0, \infty)$, the distribution on $(0, \infty) \times \mathbb{R}$ defined by $\pi(d\phi, dx) = \nu(d\phi) \otimes \mathcal{N}(0, (2\phi)^{-1})(dx)$ is a stationary distribution for $(\phi, X(t))$. The marginal distribution of X has density:

$$p(x) = \frac{1}{\sqrt{\pi}} \int_0^{+\infty} \phi^{1/2} e^{-\phi x^2} d\nu(\phi)$$

For instance, if ϕ is Gamma $\Gamma(a, \lambda)$, we get:

$$p(x) = \frac{\lambda^a \Gamma(a + (1/2))}{\sqrt{\pi} \Gamma(a) (\lambda + x^2)^{a + (1/2)}}.$$

2. We can also consider the simpler case $b(x, \phi) = \phi - x, \sigma(x) = 1$. Here, $\Phi = \mathbb{R}$, $(\ell, r) = \mathbb{R}$, $\pi_\phi = \mathcal{N}(\phi, 1/2)$. For any distribution $\nu(d\phi)$ on \mathbb{R} , $\nu(d\phi) \otimes \mathcal{N}(\phi, 1/2)(dx)$ is a stationary distribution for $(\phi, X(t))$.
3. A model with random effects in the drift and diffusion coefficient is also possible: $dX(t) = -AX(t)dt + \Sigma dW(t)$ with $\phi = (A, \Sigma) \in (0, \infty) \times (0, +\infty)$. If (A, Σ) has distribution $\nu(da, d\sigma)$, the invariant distribution of $(\phi, X(t))$ is $\pi(da, d\sigma, dx) = \nu(da, d\sigma) \otimes \mathcal{N}(0, \sigma^2/(2a))(dx)$.

2.3 Limit theorems.

We consider the process $(\phi, X(t))$ given by (1) where the initial variable (ϕ, η) has distribution $\nu(d\phi) \otimes d\mu_\phi(x)$ with ν a distribution on Φ and μ_ϕ a distribution on (ℓ, r) .

Lemma 1. *Given that $\phi = \varphi$, the conditional distribution of the process $(X(t))$ given by (1) where the initial variable (ϕ, η) has distribution $\nu(d\phi) \otimes d\mu_\phi(x)$ with μ_ϕ a distribution on (ℓ, r) is identical to the distribution of X^φ given by*

$$dX^\varphi(t) = b(X^\varphi(t), \varphi)dt + \sigma(X^\varphi(t), \varphi)dW(t), \quad X^\varphi(0) \sim \mu_\varphi \quad (9)$$

with $X^\varphi(0)$ independent of W .

Theorem 1 and 2 use Lemma 1 and a conditioning device to deduce properties of the process $(X(t))$ from those of $(X^\varphi(t))$.

Theorem 1. (1) *Let $g : (\ell, r) \rightarrow \mathbb{R}$ be a Borel function such that $\pi_\varphi(|g|) < +\infty$ for all $\varphi \in \Phi$. Then, almost surely, as T tends to infinity,*

$$\frac{1}{T} \int_0^T g(X(s))ds \rightarrow \pi_\phi(g)$$

(2) Let $g, h : (\ell, r) \rightarrow \mathbb{R}$ be Borel functions such that $\pi_\varphi(|g|) < +\infty$ and $\pi_\varphi(h^2) < +\infty$ for all $\varphi \in \Phi$. Then, as T tends to infinity,

$$\left(\frac{1}{\sqrt{T}} \int_0^T h(X(s)) dW(s), \frac{1}{T} \int_0^T g(X(s)) ds, \phi \right) \rightarrow \left(\varepsilon (\pi_\phi(h^2))^{1/2}, \pi_\phi(g), \phi \right)$$

in distribution, where ε is a standard Gaussian variable, independent of ϕ . Consequently, $\frac{1}{\sqrt{T}} \int_0^T h(X(s)) dW(s)$ converges stably in distribution to a mixed normal law $MN(0, \pi_\phi(h^2))$.

Remark 1. • In Theorem 1, it is possible to weaken the assumptions: in point (1), if $\pi_\varphi(|g|) < +\infty$ for all $\varphi \in \tilde{\Phi}$, a Borel subset of Φ , then,

$$\mathbb{P}\left(\frac{1}{T} \int_0^T g(X(s)) ds \rightarrow \pi_\phi(g), \phi \in \tilde{\Phi}\right) = \mathbb{P}(\phi \in \tilde{\Phi}).$$

• Analogously, in point (2), if $\pi_\varphi(|g|) < +\infty$ and $\pi_\varphi(h^2) < +\infty$ for all φ belonging to a subset $\tilde{\Phi}$ of Φ , the convergence in distribution holds on the set $(\phi \in \tilde{\Phi})$.

Theorem 2. Let $g : (\ell, r) \rightarrow \mathbb{R}$ be a Borel function such that $\pi_\varphi(|g|) < +\infty$ for all $\varphi \in \Phi$. Let

$$V_\varphi(g) = 4M(\varphi) \int_\ell^r s_\varphi(x) \left(\int_\ell^x (g(u) - \pi_\varphi(g)) d\pi_\varphi(u) \right)^2 dx, \quad (10)$$

and assume that $V_\varphi(g) < +\infty$ for all $\varphi \in \Phi$. Then, the following convergence in distribution holds, as T tends to infinity:

$$\left(\sqrt{T} \left(\frac{1}{T} \int_0^T g(X(s)) ds - \pi_\phi(g) \right), \phi \right) \rightarrow \left(\varepsilon V_\phi^{1/2}(g), \phi \right).$$

where ε is a standard Gaussian variable, independent of ϕ . Thus, $\sqrt{T} \left(\frac{1}{T} \int_0^T g(X(s)) ds - \pi_\phi(g) \right)$ converges stably in law to the mixed normal law $MN(0, V_\phi(g))$.

Remark 2. The same remark as Remark 1 holds. If we only have $V_\varphi(g) < +\infty$ on a subset of Φ , the convergence in distribution in Theorem 2 only holds on this subset.

Other expressions of (10) are available and simpler. To get these, we recall some known facts (see e.g. Genon-Catalot *et al.*, 2000). Assume (H1)-(H2). We denote again X^φ the process given by

$$dX^\varphi(t) = b(X^\varphi(t), \varphi) dt + \sigma(X^\varphi(t), \varphi) dW(t), \quad X^\varphi(0) \sim \pi_\varphi \quad (11)$$

with $X^\varphi(0)$ independent of W . For all $\varphi \in \Phi$, X^φ is stationary and ergodic. Let \mathcal{L}_φ denote its infinitesimal generator on $L^2(\pi_\varphi)$ with domain \mathcal{D}_φ and range \mathcal{R}_φ (see (19)). We have:

$$\begin{aligned} \mathcal{D}_\varphi &= \{F \in L^2(\pi_\varphi), F' \text{ absolutely continuous}, \mathcal{L}_\varphi F \in L^2(\pi_\varphi), \\ &\quad \lim F'(x)/s_\varphi(x) = 0, \text{ as } x \uparrow r \text{ and } \downarrow \ell\} \end{aligned}$$

For $F \in \mathcal{D}_\varphi$, the following formula holds:

$$-2 \langle \mathcal{L}_\varphi F, F \rangle_\varphi = \int_\ell^r (F' \sigma(\cdot, \varphi))^2 \pi_\varphi(x) dx,$$

where $\langle \cdot, \cdot \rangle_\varphi$ denotes the scalar product of $L^2(\pi_\varphi)$. For G such that $\pi_\varphi G = 0$, and for any F solving $\mathcal{L}_\varphi F = -G$, we have:

$$-2 \langle \mathcal{L}_\varphi F, F \rangle_\varphi = 2 \langle G, F \rangle_\varphi = 4M(\varphi) \int_\ell^r s_\varphi(x) \left(\int_\ell^x G(u) \pi_\varphi(u) du \right)^2 dx.$$

The property that $G \in \mathcal{R}_\varphi$ is equivalent to the fact that the above integral is finite. Let

$$\lambda(\varphi) = \inf \left\{ \frac{-\langle \mathcal{L}_\varphi F, F \rangle_\varphi}{\langle F, F \rangle_\varphi} \right\}, F \in \mathcal{D}_\varphi, \pi_\varphi F = 0.$$

The range \mathcal{R}_φ of the infinitesimal generator \mathcal{L}_φ is identical to the set $\{G \in L^2(\pi_\varphi), \pi_\varphi G = 0\}$ if and only if $\lambda(\varphi) > 0$. A necessary and sufficient condition for $\lambda(\varphi) > 0$ is given in Genon-Catalot *et al.*, 2000.

We can now link these properties with the result of Theorem 2. Let g be such that $\pi_\varphi g^2 < +\infty$ for all $\varphi \in \Phi$. The function $G_\varphi = g - \pi_\varphi g$ satisfies $\pi_\varphi G_\varphi = 0$. If $\lambda(\varphi) > 0$, the equation $\mathcal{L}_\varphi F_\varphi = -G_\varphi$ admits a solution $F_\varphi \in \mathcal{D}_\varphi$. Then, we know that quantity

$$V_\varphi(g) = -2 \langle \mathcal{L}_\varphi F_\varphi, F_\varphi \rangle_\varphi = \int_\ell^r (F_\varphi' \sigma(\cdot, \varphi))^2 \pi_\varphi(x) dx$$

is finite.

Example 1 (continued). Ornstein-Uhlenbeck process with random effect.

1. We consider $b(x, \varphi) = -\varphi x, \sigma(x) = 1$. We apply Theorem 1 to $h(x) = x, g(x) = x^2$. We have $\Phi = (0, \infty), (\ell, r) = \mathbb{R}, \pi_\varphi = \mathcal{N}(0, (2\varphi)^{-1}), \pi_\varphi g = (2\varphi)^{-1}$. Let $\nu(d\varphi)$ be supported in Φ . As T tends to infinity,

$$\frac{1}{T} \int_0^T X^2(s) ds \rightarrow (2\phi)^{-1} \quad \text{a.s. and}$$

$$\left(\frac{1}{\sqrt{T}} \int_0^T X(s) dW(s), \frac{1}{T} \int_0^T X^2(s) ds, \phi \right) \rightarrow_{\mathcal{D}} \left(\varepsilon (2\phi)^{-1/2}, (2\phi)^{-1}, \phi \right) \quad (12)$$

where ε is a standard Gaussian independent of ϕ . For applying Theorem 2 to g , we have to compute $V_\phi(g)$. The process X^φ corresponding to a fixed value φ with initial variable having distribution $\pi_\varphi = \mathcal{N}(0, (2\varphi)^{-1})$ is strictly stationary, ergodic, ρ -mixing with $\lambda(\varphi) = -\varphi$. Its infinitesimal generator operating on $L^2(\pi_\varphi)$ admits a sequence

of eigenvalues (the sequence $-n\varphi$, $n \in \mathbb{N}$) and associated eigenfunctions which are Hermite polynomials. In particular,

$$F_\varphi(x) = \frac{1}{2\varphi}\left(x^2 - \frac{1}{2\varphi}\right)$$

satisfies $\mathcal{L}_\varphi F_\varphi = -(x^2 - \frac{1}{2\varphi}) = -(g(x) - \pi_\varphi g)$. Actually, F_φ is an eigenfunction associated with the eigenvalue -2φ . We have $F'_\varphi(x) = x/\varphi$, $V_\varphi(g) = \int_{\mathbb{R}} (x^2/\varphi^2) d\pi_\varphi(x) = 1/(2\varphi^3)$. Consequently,

$$\sqrt{T}\left(\frac{1}{T} \int_0^T X^2(s) ds - 1/(2\phi)\right) \rightarrow_{\mathcal{D}} \varepsilon/\sqrt{2\phi^3}$$

with ε a standard Gaussian variable, independent of ϕ .

2. In the case $b(x, \varphi) = \varphi - x$, $\sigma(x) = 1$, $\Phi = \mathbb{R}$, $(\ell, r) = \mathbb{R}$, $\pi_\varphi = \mathcal{N}(\varphi, (1/2))$. Consider $g(x) = x$. As T tends to infinity,

$$\frac{1}{T} \int_0^T X(s) ds \rightarrow \pi_\phi g = \phi \quad \text{a.s. .}$$

The function $g_\varphi(x) = x - \varphi$ satisfies $\mathcal{L}_\varphi g_\varphi = -g_\varphi$. Thus, $V_\varphi(g) = 1$ and in this case,

$$\sqrt{T}\left(\frac{1}{T} \int_0^T X(s) ds - \phi\right) \rightarrow_{\mathcal{D}} \mathcal{N}(0, 1).$$

3. Consider the case of two random effects $\phi = (A, \Sigma)$, $b(x, \varphi) = -ax$, $\sigma(x, \varphi) = \sigma$, $\Phi = (0, +\infty)^2$, $\pi_\varphi = \mathcal{N}(0, \sigma^2/(2a))$, $h(x) = x$, $g(x) = x^2$. As T tends to infinity,

$$\frac{1}{T} \int_0^T X^2(s) ds \rightarrow \Sigma^2(2A)^{-1} \quad \text{a.s. and}$$

$$\left(\frac{1}{\sqrt{T}} \int_0^T X(s) dW(s), \frac{1}{T} \int_0^T X^2(s) ds, \phi\right) \rightarrow_{\mathcal{D}} \left(\varepsilon \Sigma(2A)^{-1/2}, \Sigma^2(2A)^{-1}, \phi\right) \quad (13)$$

where ε is a standard Gaussian independent of $\phi = (A, \Sigma)$. Analogously, we can compute $V_\phi(g)$. We find

$$F_\phi(x) = \frac{1}{2A}\left(x^2 - \frac{\Sigma^2}{2A}\right), \quad V_\phi(g) = \frac{\Sigma^4}{2A^3},$$

and

$$\sqrt{T}\left(\frac{1}{T} \int_0^T X^2(s) ds - \Sigma^2(2A)^{-1}\right) \rightarrow_{\mathcal{D}} \varepsilon \Sigma^2(2A^3)^{-1/2}.$$

3 Estimation of the random effect.

3.1 General drift.

We consider in this section, the estimation of the random variable ϕ based on the observation of one trajectory $(X(t), t \in [0, T])$. We assume that the random effect is not present in the diffusion coefficient, *i.e.* $\sigma(x, \varphi) = \sigma(x)$, that (H1)-(H2) hold and that (ϕ, η) has distribution $\nu(d\varphi) \otimes d\mu_\varphi(x)$ on $\Phi \times (\ell, r)$, where μ_φ is a distribution on (ℓ, r) (possibly deterministic). Under Assumptions (H1)-(H2), the process (18) is positive recurrent on (ℓ, r) for all $\varphi \in \Phi$. As ϕ takes values in Φ , $\mathbb{P}(\forall s \geq 0, \sigma(X(s)) > 0) = 1$. Therefore, the function

$$L_T(\psi) = \exp \int_0^T \frac{b(X(s), \psi)}{\sigma^2(X(s))} dX(s) - (1/2) \int_0^T \frac{b^2(X(s), \psi)}{\sigma^2(X(s))} ds$$

is well-defined for all $\psi \in \Phi$. Thus, we can introduce:

$$\hat{\phi}_T = \hat{\phi}_T(X(s), s \leq T) = \text{Argsup}_\psi L_T(\psi). \quad (14)$$

The functional $\hat{\varphi}_T = \hat{\phi}_T(X^{\varphi, x}(s), s \leq T)$ is the exact maximum likelihood of the true value φ based on the observation $(X^{\varphi, x}(s), s \leq T)$. Consider the additional assumptions:

(H3) The function $(x, \varphi) \rightarrow b(x, \varphi)$ is twice continuously differentiable on $(\ell, r) \times \mathbb{R}^d$.

(H4) The matrix $I(\varphi) = (I_{jk}(\varphi))$ with

$$I_{jk}(\varphi) = \int_{(\ell, r)} \frac{(\partial b / \partial \varphi_j)(x, \varphi)(\partial b / \partial \varphi_k)(x, \varphi)}{\sigma^2(x)} \pi_\varphi(x) dx$$

is invertible for $\varphi \in \Phi$.

Proposition 2. *Assume that the maximum likelihood estimator $\hat{\varphi}_T$ based on the observation $(X^{\varphi, x}(s), s \leq T)$ with fixed (unknown) φ is consistent and that $\sqrt{T}(\hat{\varphi}_T - \varphi)$ converges in distribution to $\mathcal{N}(0, I^{-1}(\varphi))$ for all $\varphi \in \Phi$. Then, $\hat{\phi}_T$ converges in probability to ϕ and $\sqrt{T}(\hat{\phi}_T - \phi)$ converges in distribution to the mixed normal law $MN(0, I^{-1}(\phi))$.*

Standard assumptions ensuring the consistency and the asymptotic normality for the estimator $\hat{\varphi}_T$ may be found *e.g.* in Kutoyants (2004). We do not detail these assumptions here. Note that, in the previous reference, the convergence of moments of $\sqrt{T}(\hat{\varphi}_T - \varphi)$ is obtained. But, we cannot deduce the convergence of moments of $\sqrt{T}(\hat{\phi}_T - \phi)$ since these random variables are not bounded. This is why we study below models where estimators are explicit and we can directly study their moments.

3.2 Linear random effect in the drift.

In this paragraph, let us assume that $b(\cdot, \varphi) = \varphi b(\cdot)$ with $\varphi \in \mathbb{R}$, $\sigma(x, \varphi) = \sigma(x)$ and for all $T > 0$

$$\int_0^T \frac{b^2(X(s))}{\sigma^2(X(s))} ds < +\infty.$$

To estimate the random variable ϕ from the observation of $(X(t), t \leq T)$, we consider the explicit formula for (14):

$$\hat{\phi}_T = \frac{\int_0^T \frac{b(X(s))}{\sigma^2(X(s))} dX(s)}{\int_0^T \frac{b^2(X(s))}{\sigma^2(X(s))} ds}. \quad (15)$$

We can prove directly the following result.

Proposition 3. *Assume (H1)-(H2) and that $\pi_\varphi(b^2/\sigma^2) < +\infty$ for all $\varphi \in \Phi$. Then, as T tends to infinity, $\hat{\phi}_T$ converges a.s. to ϕ and*

$$\sqrt{T}(\hat{\phi}_T - \phi) \rightarrow_{\mathcal{D}} \varepsilon / \sqrt{\pi_\phi(b^2/\sigma^2)}$$

where ε is a standard Gaussian variable independent of ϕ (stable convergence in distribution).

Example 1 (continued). Ornstein-Uhlenbeck process with random effect.

1. For $b(x) = -x$ ($b(x, \varphi) = -\varphi x$), $\sigma(x) = 1$ and ϕ supported in $(0, +\infty)$, the distribution π_φ is $\mathcal{N}(0, 1/(2\varphi))$. We have $\pi_\phi(b^2) = (2\phi)^{-1}$. The estimator $\hat{\phi}_T = -\int_0^T X(s)dX(s) / \int_0^T X^2(s)ds$ converges a.s. to ϕ and, using (12), $\sqrt{T}(\hat{\phi}_T - \phi)$ converges in distribution to $\varepsilon\sqrt{2\phi}$.
3. It is worth noting that we can consider the model with two random effects $\phi = (A, \Sigma)$ seen above. The estimator $\hat{A}_T = -\int_0^T X(s)dX(s) / \int_0^T X^2(s)ds$ converges a.s. to A and, by (13), $\sqrt{T}(\hat{A}_T - A)$ converges in distribution to $\varepsilon\sqrt{2A}$.

3.3 Convergence of moments.

For further use, we need estimators $\tilde{\phi}_T$ of the random variable ϕ such that $\mathbb{E}|\tilde{\phi}_T - \phi|^{2\gamma} \leq C T^{-\gamma}$ for some $\gamma \geq 1$. Let us consider, under Assumptions (H1)-(H2), that the process $(X(t))$ given by (1) has initial variable (ϕ, η) with distribution $\nu(d\varphi) \otimes d\pi_\varphi(x)$ with π_φ defined in (7). Thus, $(\phi, X(t))$ is Markov and strictly stationary. Given $\phi = \varphi$, the distribution of X is equal to the distribution of X^φ given in (11). We first state a preliminary result.

Proposition 4. *Assume (H1)-(H2) and that the process $(X(t))$ given by (1) has initial variable (ϕ, η) with distribution $\nu(d\varphi) \otimes d\pi_\varphi(x)$ with π_φ defined in (7). Let g such that*

$\pi_\varphi g^2 < +\infty$ for all $\varphi \in \Phi$. Let F_φ satisfy $-\mathcal{L}_\varphi F_\varphi = g - \pi_\varphi g$ for all $\varphi \in \Phi$ (see (19)), i.e.

$$F_\varphi(x) = F_\varphi^g(x) = -2M(\varphi) \int_\ell^x s_\varphi(u) \left(\int_\ell^u (g(v) - \pi_\varphi g) dv \right) du.$$

Let $K(\phi)$ be a non negative measurable function of ϕ . For $\gamma \geq 1$, assume that

$$\mathbb{E} \left[K(\phi) \left((F_\phi(\eta))^{2\gamma} + (F'_\phi(\eta)\sigma(\eta, \phi))^{2\gamma} \right) \right] < +\infty. \quad (16)$$

Then,

$$\begin{aligned} & \mathbb{E} \left[K(\phi) \left(\sqrt{T} \left(\frac{1}{T} \int_0^T g(X(s)) ds - \pi_\phi(g) \right) \right)^{2\gamma} \right] \\ & \leq c_\gamma \left(\frac{1}{T^\gamma} \mathbb{E}(K(\phi)(F_\phi(\eta))^{2\gamma}) + \mathbb{E}(K(\phi)(F'_\phi(\eta)\sigma(\eta, \phi))^{2\gamma}) \right), \end{aligned}$$

where c_γ is a numerical constant.

Example 1. (continued [2.]) Ornstein-Uhlenbeck process with random effect.

Consider $dX(t) = (\phi - X(t))dt + dW(t)$, $X(0) = \eta$, where ϕ takes values in $\Phi = \mathbb{R}$, $\pi_\varphi = \mathcal{N}(\varphi, 1/2)$. With $g(x) = x$, $\pi_\varphi g^2 = \varphi^2 + (1/2)$, $\pi_\varphi g = \varphi$, $g_\varphi(x) = x - \varphi$ satisfies $\mathcal{L}_\varphi g_\varphi = -g_\varphi$, i.e. $F_\varphi^g = g_\varphi$, $g'_\varphi = 1$, $\pi_\varphi (g_\varphi)^{2\gamma} = C_{2\gamma}(1/2^\gamma)$ where $C_{2\gamma}$ is the moment of order 2γ of a standard Gaussian variable. Consequently

$$\mathbb{E} \left(\sqrt{T} \left(\frac{1}{T} \int_0^T X(s) ds - \phi \right) \right)^{2\gamma} \leq c_\gamma (T^{-\gamma} C_{2\gamma}(1/2^\gamma) + 1).$$

We study now the moments of the estimator $\hat{\phi}_T$ in the case of a linear effect in the drift (see Section 3.2). To deal with the denominator, we use a truncated version of $\hat{\phi}_T$ (see (15)) and define:

$$\tilde{\phi}_T = \hat{\phi}_T \mathbf{1}_{\left(\frac{V_T}{T} \geq \frac{k}{\sqrt{T}}\right)} \quad \text{where} \quad V_T = \int_0^T h^2(X(s)) ds, \quad h(x) = \frac{b(x)}{\sigma(x)} \quad (17)$$

and k is a constant. Assuming that $\pi_\varphi h^4 < +\infty$, denote by $H_\varphi = F_\varphi^{h^2}$ the function defined in Proposition 4, i.e. solution of $-\mathcal{L}_\varphi H_\varphi = h^2 - \pi_\varphi h^2$. Now, we can state:

Proposition 5. Assume (H1)-(H2) and that the initial variable (ϕ, η) of (1) with $b(x, \varphi) = \varphi b(x)$, $\sigma(x, \varphi) = \sigma(x)$ has distribution $\nu(d\varphi) \otimes d\pi_\varphi(x)$. Let $\gamma \geq 1$. Assume that for all $\varphi \in \Phi$, $\pi_\varphi h^4 < +\infty$ and

$$\mathbb{E} \left[\frac{\phi^{2\gamma}}{(\pi_\phi(h^2))^{2\gamma}} \right] < +\infty, \quad \mathbb{E} \left[\frac{\pi_\phi(h^{4\gamma})}{(\pi_\phi(h^2))^{4\gamma}} \right] < +\infty,$$

$$\mathbb{E}(H_\phi(\eta))^{4\gamma} + \mathbb{E}(H'_\phi(\eta) \sigma(\eta, \phi))^{4\gamma} < +\infty, \quad \mathbb{E} \left(\left(\frac{\phi^{2\gamma}}{(\pi_\phi(h^2))^{2\gamma}} (H_\phi(\eta))^{2\gamma} + \mathbb{E}(H'_\phi(\eta) \sigma(\eta, \phi))^{2\gamma} \right) \right).$$

Then,

$$\mathbb{E}(\tilde{\phi}_T - \phi)^{2\gamma} \leq C T^{-\gamma}.$$

From the above proof, we can deduce the following corollary.

Corollary 1. Assume that $\mathbb{E}(\pi_\phi h^2)^{-2} < +\infty$ and that

$$\mathbb{E} \left(\frac{1}{(\pi_\phi(h^2))^2} (H_\phi(\eta))^2 + \mathbb{E}(H'_\phi(\eta) \sigma(\eta, \phi))^2 \right) < +\infty.$$

Then, the estimators $\hat{\phi}_T$ and $\tilde{\phi}_T$ are asymptotically equivalent, i.e. $\sqrt{T}(\hat{\phi}_T - \tilde{\phi}_T) = o_P(1)$.

Example 1 (continued [1.]) .

Let us look at the conditions of Proposition 5 in the case of the Ornstein-Uhlenbeck process $dX(t) = -\phi X(t)dt + dW(t)$ where $\phi > 0$ and $\pi_\phi = \mathcal{N}(0, (2\phi)^{-1})$. We have $h(x) = x$, $\pi_\phi(h^2) = (2\phi)^{-1}$, $H_\phi(x) = (2\phi)^{-1}(x^2 - (2\phi)^{-1})$, thus,

$$(\phi/\pi_\phi(h^2)) \propto \phi^2, \quad (\pi_\phi(h^{4\gamma})/(\pi_\phi(h^2))^{4\gamma}) \propto \phi^{2\gamma}, \quad \mathbb{E}(H_\phi(\eta))^{4\gamma} \leq C\mathbb{E}\phi^{-8\gamma}.$$

The other conditions bring no other constraints. To sum up, Proposition 5 holds if $\mathbb{E}\phi^{-8\gamma} < +\infty$ and $\mathbb{E}\phi^{4\gamma} < +\infty$. For instance, for $\gamma = 1$, and ϕ with distribution Gamma $G(a, \lambda)$, we must take $a > 8$. If $\gamma = 1$ and ϕ is inverse Gamma $\text{Inv}G(a, \lambda)$, we must impose $a > 4$. If ϕ has compact support $[a, b]$ with $0 < a < b$, all conditions are fulfilled.

3.4 Examples.

Up to now, we have illustrated our results on the Ornstein-Uhlenbeck model. In this section, we consider other classical examples and show that they are covered by the above theory.

Example 2 . Bilinear model.

Let $b(x, \varphi) = \varphi - x$, $\sigma(x) = x$. We have $\Phi = (0, +\infty)$, $(\ell, r) = (0, +\infty)$, $s_\varphi(x) = x^2 \exp 2\varphi/x$. The invariant density $\pi_\varphi(x) = 4\varphi^3 x^{-4} \exp(-2\varphi/x) 1_{(0, +\infty)}(x)$ is the Inverse Gamma $(3, 2\varphi)$ and $\Phi = (0, +\infty)$. As $\int x \pi_\varphi(x) dx = \varphi$, we have:

$$\frac{1}{T} \int_0^T X(s) ds \rightarrow \phi \quad a.s..$$

As $\int x^2 \pi_\varphi(x) dx = 2\varphi^2 < +\infty$, $g_\varphi(x) = x - \varphi \in L^2(\pi_\varphi)$ and satisfies $-\mathcal{L}_\varphi g_\varphi = g_\varphi$. Thus, with $g(x) = x$, $V_\varphi(g) = \int x^2 \pi_\varphi(x) dx = 2\varphi^2$. Hence,

$$\sqrt{T} \left(\frac{1}{T} \int_0^T X(s) ds - \phi \right) \rightarrow_{\mathcal{D}} \sqrt{2\varepsilon} \phi.$$

The function $g_\varphi(x)$ satisfies for $1 \leq \gamma < 3/2$, $\pi_\varphi(g_\varphi^{2\gamma}) \leq C\varphi^{2\gamma}(1+\Gamma(3-2\gamma))$. The assumptions of Proposition 4 hold for $1 \leq \gamma < 3/2$ and $\mathbb{E}\phi^{2\gamma} < +\infty$.

The estimator (14) is given by:

$$\hat{\phi}_T = \frac{\int_0^T (\frac{dX(s)}{X^2(s)} + \frac{ds}{X(s)})}{\int_0^T \frac{ds}{X^2(s)}}.$$

As $\int_0^{+\infty} x^{-2}\pi_\phi(x)dx = 3\phi^{-2}$, the above estimator is consistent and $\sqrt{T}(\hat{\phi}_T - \phi)$ converges in distribution as T goes to infinity to the mixed normal distribution $\mathcal{MN}(0, \phi^2/3)$.

Example 3 . C.I.R. model.

Let $b(x, \varphi) = \varphi - x$, $\sigma(x) = \sqrt{x}$, $(\ell, r) = (0, +\infty)$. We have $s_\varphi(x) = x^{-2\varphi} \exp 2x$. The invariant density $\pi_\varphi(x) = (2^{2\varphi}/\Gamma(2\varphi))x^{2\varphi-1} \exp(-2x)1_{(0,+\infty)}(x)$ is the Gamma $(2\varphi, 2)$ and $\Phi = [1/2, +\infty)$. We also have $\int x \pi_\varphi(x) dx = \varphi$. The function $g_\varphi(x) = x - \varphi$ satisfies $-\mathcal{L}_\varphi g_\varphi = g_\varphi$ and $V_\varphi(g) = \varphi$. Therefore, $(1/T) \int_0^T X(s)ds \rightarrow \phi$ *a.s.* and

$$\sqrt{T}(\frac{1}{T} \int_0^T X(s)ds - \phi) \rightarrow_{\mathcal{D}} \varepsilon \sqrt{\phi}.$$

We find that, for $2\gamma \in \mathbb{N}$,

$$\pi_\varphi(g_\varphi^{2\gamma}) \leq C(\frac{\Gamma(2\varphi + 2\gamma)}{\Gamma(2\varphi)} + \varphi^{2\gamma}) \leq C'\varphi^{2\gamma}.$$

The conditions of Proposition 4 are satisfied for all integer $\gamma \geq 1$ such that $\mathbb{E}\phi^{2\gamma} < +\infty$.

The estimator (14) is given by:

$$\hat{\phi}_T = \frac{\int_0^T (\frac{dX(s)}{X(s)} + ds)}{\int_0^T \frac{ds}{X(s)}}.$$

Here $I(\varphi) = \int_0^{+\infty} x^{-1}\pi_\varphi(x)dx < +\infty$ for $\varphi \in \Phi_0 = (1/2, +\infty)$. In this case, $I(\varphi) = (\varphi - (1/2))^{-1}$. We must assume that ϕ takes values in Φ_0 . Then, the estimator $\hat{\phi}_T$ is consistent and $\sqrt{T}(\hat{\phi}_T - \phi)$ converges in distribution as T goes to infinity to the mixed normal distribution $\mathcal{MN}(0, \phi - (1/2))$.

Example 4. Hyperbolic diffusion with random effect.

Consider $(X(t))$ given by: $dX(t) = \phi X(t)dt + \sqrt{1 + X^2(t)}dW(t)$. For the process with fixed φ , $s_\varphi(x) = (1 + x^2)^{-\varphi}$, $m_\varphi(x) = (1 + x^2)^{\varphi-1}$, condition (H2) is fulfilled for $(\ell, r) = \mathbb{R}$, $\Phi = (-\infty, 1/2)$. The constant $M(\varphi)$ is not explicit except for negative integer values. However, explicit formulae for some moments of π_φ are available.

Consider the function $h(x) = x/\sqrt{1+x^2}$. Consider the integrals $I(a) = \int dt/(1+t^2)^a < +\infty$ for $a > 1/2$ and $J(a) = \int t^2 dt/(1+t^2)^a < +\infty$ for $a > 3/2$. Using integration by parts, we have, for $a > 3/2$:

$$J(a) = \frac{1}{2(a-1)}I(a-1), \quad I(a) = \frac{2a-3}{2(a-1)}I(a-1).$$

For the invariant density π_φ , we have $M(\varphi) = I(1-\varphi)$. Thus, for $\varphi < 1/2$,

$$\pi_\varphi h^2 = \frac{1}{2(1-\varphi)}.$$

Assume that ϕ has distribution supported in $\Phi := (-\infty, 1/2)$. Then,

$$\frac{1}{T} \int_0^T \frac{X^2(s)}{1+X^2(s)} ds \rightarrow \frac{1}{2(1-\phi)} \quad \text{a.s. .}$$

For the associated central limit theorem, we have to study the ρ -mixing property of the process X^φ . In this model, the spectrum of the infinitesimal generator \mathcal{L}_φ is not discrete. However, we can check the sufficient conditions for ρ -mixing (*i.e.* spectral gap) given in Proposition 2.8 of Genon-Catalot *et al.*, 2000. First, we have $\lim_{|x| \rightarrow \infty} \sqrt{1+x^2} m_\varphi(x) = 0$. We compute the function $\gamma_\varphi(x) = (\sqrt{1+x^2})' - 2\varphi x/\sqrt{1+x^2}$. As $x \rightarrow \pm\infty$, $\gamma_\varphi(x) \rightarrow \pm(1-2\varphi)$. As these limits exist and have finite inverse, the process is ρ -mixing. Consequently, for any function G_φ such that $\pi_\varphi G_\varphi = 0$, the equation $\mathcal{L}_\varphi F_\varphi = -G_\varphi$ admits a solution $F_\varphi \in \mathcal{D}_\varphi$. For $G_\varphi = h^2 - \frac{1}{2(1-\varphi)}$, we thus have $V_\varphi(h^2) < +\infty$ for all $\varphi > 1/2$. The expression of $V_\varphi(h^2)$ is not explicit.

Note that, in this model, the discrete spectrum is included in $(-\tau(\varphi), 0]$ with $\tau(\varphi) = (1-2\varphi)^2/8$. We have:

$$\sqrt{T} \left(\frac{1}{T} \int_0^T \frac{X^2(s)}{1+X^2(s)} ds - \frac{1}{2(1-\phi)} \right) \rightarrow_{\mathcal{D}} \varepsilon (V_\phi(h^2))^{1/2}.$$

The estimator

$$\hat{\phi}_T = \int_0^T \frac{X(s)}{1+X^2(s)} dX(s) / \int_0^T \frac{X^2(s)}{1+X^2(s)} ds$$

converges *a.s.* to ϕ and $\sqrt{T}(\hat{\phi}_T - \phi)$ converges in distribution to $\varepsilon \sqrt{2(1-\phi)}$.

4 Estimation of a multidimensional linear random effect in the drift.

Let us assume that $b(\cdot, \varphi) = \sum_{j=1}^d \varphi^j b_j(\cdot)$ with $\varphi = (\varphi^j, j = 1, \dots, d) \in \mathbb{R}^d$, b_j, σ Lipschitz. Assume also that

$$\int_0^T \frac{b_j^2(X(s))}{\sigma^2(X(s))} ds < +\infty$$

for all $T > 0$ and all $j = 1, \dots, d$. Consider the vector

$$U_T = \left(\int_0^T \frac{b_1(X(s))}{\sigma^2(X(s))} dX(s) \dots \int_0^T \frac{b_d(X(s))}{\sigma^2(X(s))} dX(s) \right)'$$

and the matrix $\mathcal{I}_T = (\mathcal{I}_T^{k,\ell})_{1 \leq k, \ell \leq d}$ with

$$\mathcal{I}_T^{k,\ell} = \int_0^T \frac{b_k(X(s))b_\ell(X(s))}{\sigma^2(X(s))} ds.$$

Consider the estimator $\hat{\phi}_T$ of ϕ defined by the equation

$$\mathcal{I}_T \hat{\phi}_T = U_T.$$

Assume that $\pi_\varphi(b_k b_\ell / \sigma^2) < +\infty$ for all $\varphi \in \tilde{\Phi} \subset \Phi$, let

$$\mathcal{I}(\phi) = (\pi_\phi(b_k b_\ell / \sigma^2)).$$

Then,

$$\mathbb{P}(\mathcal{I}_T/T \rightarrow \mathcal{I}(\phi), \phi \in \tilde{\Phi}) = \mathbb{P}(\phi \in \tilde{\Phi}).$$

Denote by Φ_0 the subset of $\tilde{\Phi}$ such that $\mathcal{I}(\varphi)$ invertible for all $\varphi \in \Phi_0$. On Φ_0 , $\sqrt{T}(\hat{\phi}_T - \phi)$ converges in distribution to the centered mixed normal distribution with covariance matrix $\mathcal{I}(\phi)^{-1}$ ($MN(0, \mathcal{I}(\phi)^{-1})$).

Example 1 (continued [3]). Ornstein-Uhlenbeck process with two random effects.

Consider the stochastic differential equation $dX(t) = (-AX(t) + B)dt + dW(t)$. Let us assume that $\phi = (A, B)$ has distribution $\nu(da, db)$ with support $\Phi = (0, \infty) \times \mathbb{R}$ and that given $A = a, B = b$, $X(0)$ has distribution $\mathcal{N}(b/a, 1/2a)$. We have

$$\mathcal{I}_T = \begin{pmatrix} \int_0^T X^2(s) ds & -\int_0^T X(s) ds \\ -\int_0^T X(s) ds & T \end{pmatrix},$$

$$U_T = \left(-\int_0^T X(s) dX(s), X(T) - X(0) \right)'$$

$$\mathcal{I}(\phi) = \begin{pmatrix} 1/2A & B/A \\ B/A & 1 \end{pmatrix},$$

We can compute $\det \mathcal{I}(\varphi) = 1/2a - (b/a)^2$. The set Φ_0 is therefore $\Phi_0 = \{\varphi = (a, b) \in \Phi, a > 2b^2\}$. The estimator $\mathbb{P}(\hat{\phi}_T \rightarrow \phi, \phi \in \Phi_0) = \mathbb{P}(\phi \in \Phi_0)$. On the set $\phi \in \Phi_0$, $\sqrt{T}(\hat{\phi}_T - \phi)$ is asymptotically $MN(0, \mathcal{I}(\phi)^{-1})$.

5 Estimation of the distribution of the random effect from *i.i.d.* trajectories.

In this section, we assume that we observe N *i.i.d.* sample paths $(X_j(t), t \leq T)$, $j = 1, \dots, N$ and want to estimate the distribution of ϕ_j . In the asymptotic framework $N \rightarrow +\infty$ with fixed T , it is possible to estimate consistently this distribution. For specific diffusion models and specific distributions of the ϕ_j 's, explicit maximum likelihood estimators can be studied (see *e.g.* Delattre *et al.*, 2012 and the references therein). For general models, explicit computations are not possible. Numerical results are available or theoretical and heavy conditions have been proposed (see references given in the introduction). Here, we adopt a different point of view and consider a double asymptotic framework $N \rightarrow +\infty$ and $T = T(N) \rightarrow +\infty$. This is the approach of Comte *et al.*, 2012. Indeed, it allows to simplify the problem and obtain explicit estimators with a two-step procedure. From each sample path $(X_j(t), t \in [0, T])$, we build an estimator $\tilde{\phi}_{j,T}$ which is an approximation of ϕ_j for large T . Then, the N *i.i.d.* estimators $(\tilde{\phi}_{j,T}, j = 1, \dots, N)$ can be used to estimate the common distribution of the ϕ_j 's.

In what follows, we consider a parametric model for the distribution of ϕ_j , *i.e.* we set $\nu(d\varphi) = \nu_\theta(d\varphi) = f(\theta, \varphi)d\alpha(\varphi)$ for some dominating measure α on Φ and $\theta \in \Theta \subset \mathbb{R}^p$. As usual, we denote by θ_0 the true value of the parameter.

5.1 Moment method.

Let $F : \Phi \rightarrow \mathbb{R}$ be a measurable function such that $\forall \theta \in \Theta$, $\mathbb{E}_\theta F^2(\phi_j) < +\infty$, and set $m_\theta(F) = \int_\Phi F(\varphi)f(\theta, \varphi)d\alpha(\varphi)$.

Proposition 6. *Let $\tilde{\phi}_{j,T}$ be an estimator of ϕ_j built from the trajectory $(X_j(t), t \in [0, T])$ satisfying, for some constant $C(\theta_0)$,*

$$\mathbb{E}_{\theta_0} |F(\tilde{\phi}_{j,T}) - F(\phi_j)| \leq \frac{C(\theta_0)}{\sqrt{T}}.$$

If both N and T tend to infinity, $N^{-1} \sum_{j=1}^N F(\tilde{\phi}_{j,T})$ tends to $m_{\theta_0}(F)$ in \mathbb{P}_{θ_0} -probability. If moreover, $T = T(N)$ satisfies $T(N)/N \rightarrow +\infty$ as $N \rightarrow +\infty$, then,

$$\sqrt{N} \left(N^{-1} \sum_{j=1}^N F(\tilde{\phi}_{j,T}) - m_{\theta_0}(F) \right) \rightarrow_{\mathcal{D}(\mathbb{P}_{\theta_0})} \mathcal{N}(0, \sigma_{\theta_0}^2)$$

where $\sigma_{\theta_0}^2 = m_{\theta_0}(F^2) - m_{\theta_0}^2(F)$.

The proof is elementary using the assumption and the standard large law of numbers and central limit theorem for $N^{-1} \sum_{j=1}^N F(\phi_j)$.

5.2 Gaussian random effects.

Consider a model where $\Phi = \mathbb{R}$, $\phi_j \sim \mathcal{N}(\mu, \omega^2)$ and let $\theta = (\mu, \omega^2)$.

Proposition 7. *Let $\tilde{\phi}_{j,T}$ be an estimator of ϕ_j built from the trajectory $(X_j(t), t \in [0, T])$ satisfying, for some constant $C(\theta_0)$, $\mathbb{E}_{\theta_0}(\tilde{\phi}_{j,T} - \phi_j)^2 \leq \frac{C(\theta_0)}{T}$. Set*

$$\tilde{\mu}_{N,T} = N^{-1} \sum_{j=1}^N \tilde{\phi}_{j,T}, \quad \tilde{\omega}_{N,T}^2 = N^{-1} \sum_{j=1}^N (\tilde{\phi}_{j,T} - \tilde{\mu}_{N,T})^2.$$

Then, as N, T go to infinity, $(\tilde{\mu}_{N,T}, \tilde{\omega}_{N,T}^2)$ is a consistent estimator of $\theta_0 = (\mu_0, \omega_0^2)$. If moreover, $T = T(N)$ tends to infinity with $T(N)/N$ tending to infinity, then $\sqrt{N}(\tilde{\mu}_{N,T} - \hat{\mu}_N)$ and $\sqrt{N}(\tilde{\omega}_{N,T}^2 - \hat{\omega}_N^2)$ tend to 0 in \mathbb{P}_{θ_0} -probability, where $\hat{\mu}_N = N^{-1} \sum_{j=1}^N \phi_j$ and $\hat{\omega}_N^2 = N^{-1} \sum_{j=1}^N (\phi_j - \hat{\mu}_N)^2$. Thus, the estimator $(\tilde{\mu}_{N,T}, \tilde{\omega}_{N,T}^2)$ is asymptotically equivalent to the exact MLE based on $(\phi_j, j = 1, \dots, N)$.

The proof is elementary.

Example 1 (continued [2].) Ornstein-Uhlenbeck process.

Assume that $dX_j(t) = (\phi_j - X_j(t))dt + dW_j(t)$, $X_j(0) = \eta_j$ where $\phi_j \sim \mathcal{N}(\mu, \omega^2)$ and given that $\phi_j = \varphi$, $X_j(0) \sim \mathcal{N}(\varphi, 1/2)$. Then, as seen above, $\tilde{\phi}_{j,T} = T^{-1} \int_0^T X_j(s)ds$ satisfies the condition of the proposition. The parameter $\theta_0 = (\mu_0, \omega_0^2)$ can be estimated via $(\tilde{\mu}_{N,T}, \tilde{\omega}_{N,T}^2)$.

5.3 Contrast method.

More generally, assume that ϕ_j has the true distribution $f(\theta_0, \varphi)d\alpha(\varphi)$ on Φ and set

$$u(\theta, \varphi) = -\log f(\theta, \varphi) 1_{\varphi \in \Phi}, \quad U_N(\theta) = \frac{1}{N} \sum_{j=1}^N u(\theta, \tilde{\phi}_{j,T}),$$

where $\tilde{\phi}_{j,T}$ is an estimator of ϕ_j built from $(X_j(t), t \leq T)$. Let

$$\tilde{\theta}_{N,T} = \text{Arg inf}_{\theta \in \Theta} U_N(\theta)$$

be an associated minimum contrast estimator. Then, by an appropriate choice of the $\tilde{\phi}_{j,T}$'s, the minimum contrast estimator will be asymptotically equivalent to the exact MLE based on the ϕ_j 's. The choice of the estimators $\tilde{\phi}_{j,T}$ is linked with the model for $(X_j(t))$ and the function $u(\theta, \varphi)$.

6 Concluding remarks

In this paper, we provide sufficient conditions for the existence of stationary distributions for a process $(X(t))$ defined by a SDE with a random effect ϕ . There is no unicity of the stationary distributions.

Generally, authors are interested in estimating the unknown distribution of the random effect. Here, we propose several estimators of the random variable ϕ itself based on functionals of the continuous sample path $(X(t), t \leq T)$. The estimators are proved to be consistent and asymptotically mixed normal as T goes to infinity.

In practice, only discrete time observations are available. Then, as it is standard, we can use the one-step discretizations of integrals in formulae to obtain the estimators of ϕ .

To estimate the common distribution of the ϕ_j 's when N *i.i.d.* processes $(X_j(t), t \leq T), j = 1, \dots, N)$ are observed, one may proceed as follows. From each sample path $(X_j(t), t \leq T)$ we build an estimator $\tilde{\phi}_{j,T}$ which appears as a noisy observation of ϕ_j for large T . We simply replace ϕ_j by $\tilde{\phi}_{j,T}$ to obtain estimators. The method relies on a double asymptotic framework where both N and T go to infinity.

7 Appendix: proofs

Proof of Proposition 1

Assume that $(\phi, X(0))$ has distribution π on $\Phi \times (\ell, r)$. Let $X^{\varphi,x}$ denote the process given by

$$dX^{\varphi,x}(t) = b(X^{\varphi,x}(t), \varphi)dt + \sigma(X^{\varphi,x}(t), \varphi)dW(t), \quad X^{\varphi,x}(0) = x. \quad (18)$$

For $H : \mathbb{R} \rightarrow \mathbb{R}^+, G : \mathbb{R}^d \rightarrow \mathbb{R}^+$ two Borel functions, by the Markov property of $(X(t), \phi)$, we have:

$$\mathbb{E}(H(X(t))G(\phi)|X(0) = x, \phi = \varphi) = \mathbb{E}(H(X^{\varphi,x}(t)))G(\varphi).$$

Using $p_t^\varphi(x, y)$, the transition density of (18), we have:

$$\mathbb{E}(H(X^{\varphi,x}(t))) = \int_{(\ell,r)} H(y)p_t^\varphi(x, y)dy.$$

As $\pi_\varphi(x)$ is the invariant density of (18), we get:

$$\begin{aligned} \int_{(\ell,r)} \pi_\varphi(x) dx \mathbb{E}(H(X^{\varphi,x}(t))) &= \int_{(\ell,r)} H(y)dy \int_{(\ell,r)} dx \pi_\varphi(x) p_t^\varphi(x, y) \\ &= \int_{(\ell,r)} H(y)\pi_\varphi(y)dy. \end{aligned}$$

Therefore,

$$\begin{aligned}\mathbb{E}(H(X(t))G(\phi)) &= \\ \int_{\Phi \times (\ell, r)} \nu(d\varphi) \pi_\varphi(x) dx \mathbb{E}(H(X^{\varphi, x}(t)))G(\varphi) &= \int_{\Phi \times (\ell, r)} G(\varphi)H(y)\nu(d\varphi)\pi_\varphi(y) dy \\ &= \mathbb{E}(H(X(0))G(\phi)).\end{aligned}$$

□

Proof of Lemma 1

We compare the finite-dimensional distributions. For $n \geq 1$, $0 \leq t_1 \leq t_2 \leq \dots \leq t_n$, $H: \mathbb{R}^n \rightarrow \mathbb{R}^+$ and $k: \mathbb{R}^d \rightarrow \mathbb{R}^+$, we have,

$$\begin{aligned}\mathbb{E}(k(\phi)H(X(t_1), \dots, X(t_n))) &= \int_{\Phi} d\nu(\varphi)k(\varphi) \mathbb{E}(H(X(t_1), \dots, X(t_n))|\phi = \varphi) \\ &= \int_{\Phi \times (\ell, r)^{n+1}} d\nu(\varphi) \times d\mu_\varphi(x_0)k(\varphi)H(x_1, \dots, x_n) \prod_{i=1}^n p_{t_i - t_{i-1}}^\varphi(x_{i-1}, x_i) dx_1 \dots, dx_n.\end{aligned}$$

Thus, the conditional density of $(X(t_1), \dots, X(t_n))$ given $\phi = \varphi$ is equal to:

$$\int_{(\ell, r)} d\mu_\varphi(x_0) \prod_{i=1}^n p_{t_i - t_{i-1}}^\varphi(x_{i-1}, x_i).$$

This is exactly the density of $(X^\varphi(t_1), \dots, X^\varphi(t_n))$ given by (9). □

Proof of Theorem 1

Point (1). The process X^φ defined in (9) satisfies the ergodic theorem for all $\varphi \in \Phi$:

$$\mathbb{P}\left(\frac{1}{T} \int_0^T g(X^\varphi(s)) ds \xrightarrow{T \rightarrow +\infty} \pi_\varphi(g)\right) = 1.$$

By Lemma 1,

$$\mathbb{P}\left(\frac{1}{T} \int_0^T g(X(s)) ds \rightarrow \pi_\phi(g) | \phi = \varphi\right) = \mathbb{P}\left(\frac{1}{T} \int_0^T g(X^\varphi(s)) ds \rightarrow \pi_\varphi(g)\right).$$

Thus,

$$\mathbb{P}\left(\frac{1}{T} \int_0^T g(X(s)) ds \rightarrow \pi_\phi(g)\right) = 1.$$

Point (2). The conditional distribution of $(X = F(\phi, \eta; W), W)$ given $\phi = \varphi$ is equal to the distribution of $(X^\varphi = (F(\varphi, X^\varphi(0); W), W)$ with $X^\varphi(0)$ independent of W and with distribution μ_φ .

As $\frac{1}{T} \int_0^T h^2(X^\varphi(s)) ds \rightarrow \pi_\varphi(h^2)$ a.s., the central limit theorem for stochastic integrals (see e.g. Kutoyants, 1986) implies:

$$\frac{1}{\sqrt{T}} \int_0^T h(X^\varphi(s)) dW(s) \xrightarrow{T \rightarrow +\infty} \mathcal{N}(0, \pi_\varphi(h^2)), \quad \text{in distribution.}$$

We also have $(\frac{1}{T} \int_0^T g(X^\varphi(s))ds \rightarrow \pi_\varphi(g)$ a.s. as T tends to infinity. Thus, the following convergence in distribution holds:

$$\left(\frac{1}{\sqrt{T}} \int_0^T h(X^\varphi(s))dW(s), \frac{1}{T} \int_0^T g(X^\varphi(s))ds\right) \rightarrow (\mathcal{N}(0, \pi_\varphi(h^2)), \pi_\varphi(g)).$$

Set

$$Z_T = \frac{1}{\sqrt{T}} \int_0^T h(X(s))dW(s), \quad V_T = \frac{1}{T} \int_0^T g(X(s))ds.$$

For $u, v, w \in \mathbb{R}$,

$$\mathbb{E}(\exp(iuZ_T + ivV_T)|\phi = \varphi) \exp(iw\varphi) \rightarrow \exp\left(-\frac{u^2}{2}\pi_\varphi(h^2) + iv\pi_\varphi(g) + iw\varphi\right).$$

This implies

$$\mathbb{E}(\exp(iuZ_T + ivV_T + iw\phi)) \rightarrow \mathbb{E}\left[\exp\left(-\frac{u^2}{2}\pi_\phi(h^2) + iv\pi_\phi(g) + iw\phi\right)\right].$$

For ε a random variable independent of ϕ with distribution $\mathcal{N}(0, 1)$, we have:

$$\mathbb{E}\left[\exp\left(iu\varepsilon(\pi_\phi(h^2))^{1/2} + iv\pi_\phi(g) + iw\phi\right)\right] = \mathbb{E}\left[\exp\left(-\frac{u^2}{2}\pi_\phi(h^2) + iv\pi_\phi(g) + iw\phi\right)\right].$$

□

Proof of Theorem 2

We consider for fixed φ the process X^φ given by (9) and introduce its infinitesimal generator \mathcal{L}_φ defined on $C^2((\ell, r))$ -functions by:

$$\mathcal{L}_\varphi F(x) = \frac{\sigma^2(x, \varphi)}{2} F''(x) + b(x, \varphi) F'(x) = \frac{1}{2} \frac{1}{m_\varphi} \left(\frac{F'}{s_\varphi}\right)'. \quad (19)$$

By the Ito formula, for F a C^2 -function on (ℓ, r) , we have:

$$F(X^\varphi(T)) = F(X^\varphi(0)) + \int_0^T \mathcal{L}_\varphi F(X^\varphi(s))ds + \int_0^T F'(X^\varphi(s))\sigma(X^\varphi(s), \varphi)dW(s).$$

Therefore,

$$-\frac{1}{\sqrt{T}} \int_0^T \mathcal{L}_\varphi F(X^\varphi(s))ds = \frac{1}{\sqrt{T}} \int_0^T F'(X^\varphi(s))\sigma(X^\varphi(s), \varphi)dW(s) + R_T,$$

with

$$R_T = \frac{1}{\sqrt{T}}(F(X^\varphi(0)) - F(X^\varphi(T))).$$

The ergodic properties of X^φ imply that $X^\varphi(T)$ converges in distribution to the stationary distribution $\pi_\varphi(x)dx$ as T tends to infinity. Hence, R_T tends to 0 in probability. We can conclude that, if

$$\pi_\varphi((F'\sigma(\cdot, \varphi))^2) < +\infty,$$

the following convergence in distribution holds:

$$-\frac{1}{\sqrt{T}} \int_0^T \mathcal{L}_\varphi F(X^\varphi(s)) ds \rightarrow \mathcal{N}(0, \pi_\varphi((F' \sigma(\cdot, \varphi))^2)).$$

Now, setting $g_\varphi(x) = g(x) - \pi_\varphi(g)$, we search F_φ solving $\mathcal{L}_\varphi F_\varphi = -g_\varphi$. Using (19) yields:

$$F'_\varphi(x) = -2s_\varphi(x) \int_\ell^x g_\varphi(u) m_\varphi(u) du. \quad (20)$$

A simple computation shows that $V_\varphi(g) = \pi_\varphi((F' \sigma(\cdot, \varphi))^2)$. By the same conditioning device, we obtain:

$$\mathbb{E}(\exp(iu \frac{1}{\sqrt{T}} \int_0^T g_\varphi(X(s)) ds + iv\phi)) \rightarrow \mathbb{E}(\exp(-\frac{u^2}{2} V_\varphi(g) + iv\phi)),$$

which achieves the proof. \square

Proof of Proposition 2

As above, we use the conditioning device:

$$\mathbb{P}(|\hat{\phi}_T - \phi| > h | \phi = \varphi) = \mathbb{P}(|\hat{\phi}_T - \varphi| > h).$$

And, analogously for the convergence in distribution. \square

Proof of Proposition 3

We have

$$\hat{\phi}_T - \phi = \frac{M_T}{\langle M \rangle_T},$$

with $M_T = \int_0^T \frac{b(X(s))}{\sigma(X(s))} dW(s)$ and $\langle M \rangle_T = \int_0^T \frac{b^2(X(s))}{\sigma^2(X(s))} ds$. First, by theorem 1, $\langle M \rangle_T / T$ converges *a.s.* to $\pi_\phi(b^2/\sigma^2) > 0$. Hence, $\langle M \rangle_{+\infty} = +\infty$ *a.s.* and $M_T / \langle M \rangle_T$ converges *a.s.* to 0. Second, $(M_T/\sqrt{T}, \langle M \rangle_T / T)$ converges in distribution to $(\varepsilon \sqrt{\pi_\phi(b^2/\sigma^2)}, \pi_\phi(b^2/\sigma^2))$. This implies the result. \square

Proof of Proposition 4

We start with the process X^φ given in (11) and set $G_\varphi = g - \pi_\varphi g$. Using that $-\mathcal{L}_\varphi F_\varphi = G_\varphi$ and the Ito formula, we get:

$$\int_0^T G_\varphi(X^\varphi(s)) ds = \int_0^T F'_\varphi(X^\varphi(s)) \sigma(X^\varphi(s), \varphi) dW(s) + F_\varphi(X^\varphi(0)) - F_\varphi(X^\varphi(T)).$$

The process X^φ is stationary with marginal distribution π_φ . Therefore, using the Hölder and the Burkholder-Davis-Gundy inequalities yields:

$$\mathbb{E} \left(\int_0^T G_\varphi(X^\varphi(s)) ds \right)^{2\gamma} \leq c_\gamma \left(T^\gamma \pi_\varphi((F'_\varphi \sigma(\cdot, \varphi))^{2\gamma}) + \pi_\varphi((F_\varphi)^{2\gamma}) \right).$$

Now,

$$\mathbb{E}\left(\left(\int_0^T G_\phi(X(s))ds\right)^{2\gamma} \mid \phi = \varphi\right) = \mathbb{E}\left(\left(\int_0^T G_\varphi(X^\varphi(s))ds\right)^{2\gamma}\right).$$

Noting that

$$\mathbb{E}K(\phi)(\pi_\phi((F'_\phi\sigma(\cdot, \phi))^{2\gamma})) = \mathbb{E}K(\phi)(F'_\phi(\eta)\sigma(\eta, \phi))^{2\gamma},$$

and

$$\mathbb{E}K(\phi)\pi_\phi((F_\phi)^{2\gamma}) = \mathbb{E}K(\phi)F_\phi(\eta)^{2\gamma},$$

gives the result. \square

Proof of Proposition 5

We have: $\sqrt{T}(\tilde{\phi}_T - \phi) = A_1 + A_2$, with

$$A_1 = -\sqrt{T}\phi \mathbf{1}_{\left(\frac{V_T}{T} < \frac{k}{\sqrt{T}}\right)}, \quad A_2 = \frac{M_T/\sqrt{T}}{V_T/T} \mathbf{1}_{\left(\frac{V_T}{T} \geq \frac{k}{\sqrt{T}}\right)}, \quad (21)$$

where $M_T = \int_0^T h(X(s))dW(s)$ and $h(x) = \frac{b(x)}{\sigma(x)}$. We first study the term A_2 . Note that:

$$\mathbb{E}\left[\frac{V_T}{T} \mid \phi\right] = \frac{1}{T} \int_0^T \mathbb{E}[h^2(X(s)) \mid \phi] ds,$$

and $\mathbb{E}[h^2(X(s)) \mid \phi = \varphi] = \mathbb{E}[h^2(X^\varphi(s))] = \pi_\varphi(h^2)$. Thus,

$$L = \pi_\phi(h^2) = \mathbb{E}\left(\frac{V_T}{T} \mid \phi\right) = \mathbb{E}h^2(X(0)) \mid \phi.$$

We have: $A_2 = A'_2 + A''_2$ with

$$A'_2 = \frac{M_T}{\sqrt{T}} \left(\frac{1}{V_T/T} - \frac{1}{L} \right) \mathbf{1}_{\left(\frac{V_T}{T} \geq \frac{k}{\sqrt{T}}\right)}, \quad A''_2 = \frac{M_T}{L\sqrt{T}} \mathbf{1}_{\left(\frac{V_T}{T} \geq \frac{k}{\sqrt{T}}\right)}. \quad (22)$$

Using that L is \mathcal{F}_0 -measurable, $\langle M \rangle_T = V_T$, the Burkholder-Davis-Gundy and the Hölder inequalities, we get:

$$\begin{aligned} \mathbb{E}(A''_2)^{2\gamma} &\leq C_{2\gamma} \mathbb{E} \frac{V_T^\gamma}{L^{2\gamma} T^\gamma} \\ &\leq C_{2\gamma} \mathbb{E} \frac{1}{L^{2\gamma} T} \int_0^T h^{2\gamma}(X(s)) ds \\ &= C_{2\gamma} \mathbb{E} \frac{\pi_\phi(h^{2\gamma})}{L^{2\gamma}} = C_{2\gamma} \mathbb{E} \frac{\pi_\phi(h^{2\gamma})}{(\pi_\phi(h^2))^{2\gamma}}, \end{aligned}$$

where $C_{2\gamma}$ is the constant of the B-D-G inequality. For $\gamma = 1$, we have:

$$\mathbb{E}(A''_2)^2 = \mathbb{E} \frac{1}{L} = \mathbb{E} \frac{1}{\pi_\phi(h^2)}.$$

We thus need $\mathbb{E}X < +\infty$ with $X = \pi_\phi(h^{2\gamma})/(\pi_\phi(h^2))^{2\gamma}$. We look at A'_2 :

$$\mathbb{E}(A'_2)^{2\gamma} \leq \frac{1}{k^{2\gamma}} \left(\mathbb{E} \left(\frac{M_T}{L\sqrt{T}} \right)^{4\gamma} \mathbb{E} \left(\sqrt{T}(L - V_T/T) \right)^{4\gamma} \right)^{1/2}.$$

We have:

$$\mathbb{E} \left(\frac{M_T}{L\sqrt{T}} \right)^{4\gamma} \leq C_{4\gamma} \mathbb{E} \frac{1}{L^{4\gamma}} \left(\frac{V_T}{T} \right)^{2\gamma} \leq C_{4\gamma} \mathbb{E} \frac{\pi_\phi(h^{4\gamma})}{(\pi_\phi(h^2))^{4\gamma}}.$$

Therefore, we impose $\mathbb{E}Y < +\infty$ with $Y = \pi_\phi(h^{4\gamma})/(\pi_\phi(h^2))^{4\gamma}$. As $X^2 \leq Y$, the condition $\mathbb{E}Y < +\infty$ implies $\mathbb{E}X < +\infty$.

Then, we apply Proposition 4 with the function $H_\phi = F_\phi^{h^2}$:

$$\mathbb{E} \left(\sqrt{T}(L - V_T/T) \right)^{4\gamma} \leq C \left(\mathbb{E}(H_\phi(\eta))^{4\gamma} + \mathbb{E}(H'_\phi(\eta) \sigma(\eta, \phi))^{4\gamma} \right).$$

There remains the term A_1 . We have $A_1^{2\gamma} \leq T^\gamma \phi^{2\gamma} \mathbf{1}_{(\frac{V_T}{T} < \frac{k}{\sqrt{T}})}$. Therefore, setting $\ell = \pi_\phi h^2$, we have, for all $p \geq 1$,

$$\begin{aligned} \mathbb{P}\left(\frac{V_T}{T} < \frac{k}{\sqrt{T}} \mid \phi = \varphi\right) &\leq \mathbb{P}\left(|\ell - \frac{V_T}{T}| > \ell - \frac{k}{\sqrt{T}} \mid \phi = \varphi\right) \mathbf{1}_{(\ell > \frac{2k}{\sqrt{T}})} + \mathbf{1}_{(\ell \leq \frac{2k}{\sqrt{T}})} \\ &\leq \mathbb{P}\left(|\ell - \frac{V_T}{T}| > \frac{\ell}{2} \mid \phi = \varphi\right) + \mathbf{1}_{(\ell^{-1} \geq \frac{\sqrt{T}}{2k})} \\ &\leq \left(\frac{2}{\ell}\right)^{2p} \mathbb{E}\left(\ell - \frac{V_T}{T}\right)^{2p} \mid \phi = \varphi + \left(\frac{2k}{\sqrt{T}}\right)^{2p} \ell^{-2p}. \end{aligned}$$

Thus,

$$\mathbb{E}(A_1^{2\gamma}) \leq T^{\gamma-p} \mathbb{E} \left[\phi^{2\gamma} \left(\frac{2}{L}\right)^{2p} \left(\sqrt{T}(L - \frac{V_T}{T}) \right)^{2p} \right] + T^{\gamma-p} (2k)^{2p} \mathbb{E}(\phi^{2\gamma} L^{-2p}).$$

We need $p \geq \gamma$ and that the left-hand side be finite. Taking $p = \gamma$, we apply Proposition 4 with $K(\phi) = (\phi/L)^{2\gamma}$ and get the result with the condition stated. Note that, for any function $F(\varphi, x)$, the following relation holds:

$$\mathbb{E}F(\phi, \eta) = \mathbb{E}\pi_\phi F(\phi, \cdot).$$

Note also that, taking $p > \gamma$ yields that $\mathbb{E}(A_1^{2\gamma})$ tends to 0 as T tends to infinity provided that the left-hand side expectations are finite. But this yields a stronger moment constraint. \square

Proof of Corollary 1

We have $\sqrt{T}(\hat{\phi}_T - \tilde{\phi}_T) = \sqrt{T}\hat{\phi}_T \mathbf{1}_{(\frac{V_T}{T} < \frac{k}{\sqrt{T}})}$. As $\hat{\phi}_T$ converges to ϕ , it is enough to prove that $\sqrt{T}\mathbb{P}(\frac{V_T}{T} < \frac{k}{\sqrt{T}})$ tends to 0. We use the same inequality as above with $p = 1$ and get:

$$\sqrt{T}\mathbb{P}\left(\frac{V_T}{T} < \frac{k}{\sqrt{T}} \mid \phi = \varphi\right) \leq \frac{1}{\sqrt{T}} \left(\frac{2}{\ell}\right)^2 \mathbb{E}\left((\sqrt{T}(\ell - \frac{V_T}{T}))^2 \mid \phi = \varphi\right) + \frac{2k}{\sqrt{T}} \ell^{-2}. \quad (23)$$

We apply Proposition 4 with $K(\phi) = (\pi_\phi h^2)^{-2}$. This ends the proof. \square

References

- [1] Comte F., Genon-Catalot V. and Samson A. (2012). Nonparametric estimation for stochastic differential equations with random effects. Preprint MAP5 2012-35. To appear in *Stoch. Proc. and Appl.*.
- [2] Davidian, M. and Giltinan, D.(1995). *Non linear models to repeated measruement data*. Chapman and Hall.
- [3] Ditlevsen, S. and De Gaetano, A. (2005). Mixed effects in stochastic differential equation models. *Statistical Journal* **3**, 137-153.
- [4] Donnet, S. and Samson, A. (2008). Parametric inference for mixed models defined by stochastic differential equations. *ESAIM P & S* **12**, 196-218.
- [5] Kuhn, E. and Lavielle, M. (2004). Coupling an approximation version of em with an mcmc procedure. *ESAIM P & S* **8**, 115-131.
- [6] Kuotoyants, Yu. A. (2004). *Statistical inference for ergodic diffusion processes*. Springer Series in Statistics. Springer-Verlag, London.
- [7] Nie, L. (2006). Strong consistency of the MLE in generalized linear and nonlinear mixed-effects models. *Metrika* **63**, 123-143.
- [8] Nie, L. (2007). Convergence rate of the MLE in generalized linear and nonlinear mixed-effects models: theory and applications. *Journal of Statistical Planning and Inference* **137**, 1787-1804.
- [9] Nie, L. & Yang, M. (2005). Strong consistency of the MLE in nonlinear mixed-effects models with large cluster size. *Sankhya: The Indian Journal of Statistics* **67**, 736-763.
- [10] Overgaard, R.V., Jonsson, N., Tornoe, W. and Madsen, H.(2005). Non-linear Mixed-effects models with stochastic differential equations: implementation of an estimation algorithm. *Journal of Pharmacokinetics and Pharmacodynamics* **32**, 85-107.
- [11] Picchini, U., De Gaetano, A. and Ditlevsen, S. (2010). Stochastic differential mixed-effects models. *Scand. J. Statist.* **37**, 67-90.
- [12] Picchini, U. and Ditlevsen, S. (2011). Practicle estimation of high dimensional stochastic differential mixed-effects models. *Computational Statistics & Data Analysis* **55**, 1426-1444.
- [13] Pinheiro, J.C. and Bates, D.M. (2000). *Mixed-Effects Models in S and S-PLUS*. New York, Springer.
- [14] Wolfinger, R. (1993). Laplace's approximation for nonlinear mixed models. *Biometrika* **80**, 791-795.