Parameter estimation for a discrete sampling of an integrated Ornstein–Uhlenbeck Process.

Arnaud Gloter
Université de Marne la Vallée
Equipe d'Analyse et de Mathématiques appliquées
5, boulevard Descartes
Champs-sur-Marne
77454 MARNE-LA-VALLEE CEDEX 2
e-mail : gloter@math.univ-mlv.fr

July 2000
Published in Statistics, Vol 35 (2001), p.225-243

Abstract

We study the estimation of parameters $\theta = (\mu, \sigma^2)$ for a diffusion $dX_t = a(X_t, \sigma^2)dB_t + b(X_t, \mu)dt$, when we observe a discretization with step Δ of the integral $I_t = \int_0^t X_s ds$. To keep computations tractable we focus on the case of an Ornstein–Uhlenbeck process, but our results provide information on how to deal with other processes. We study an efficient estimator $\hat{\theta}_n$ based on the Gaussian property of the process $(\int_{i\Delta}^{(i+1)\Delta} X_s ds)_{i \geq 0}$, and we give an estimator $\tilde{\theta}_n$ based on Ryden's idea of maximum likelihood split data. We compare these different estimators: first we give some numerical results, then we give a theoretical explanation for these results.

AMS 1991 *subject classification.* 62F12, 62M09, 62M10.
key words: diffusion processes, discrete time observation, parametric inference, minimum contrast, Whittle approximation.

1 Introduction

Parameter estimation for a diffusion process $(X_t)_{t \geq 0}$, which solves $dX_t = a(X_t, \sigma^2)dB_t + b(X_t, \mu)dt$ is now classical. It has been treated under many different assumptions for the observation of the sample path: (X_t) may be continuously observed throughout a time interval $[0, T]$ (see Kutoyants (1981)); or only a discretization may be observed, the sampling interval Δ being fixed or tending to zero as the number of observations tends to infinity (see Bibby and Sørensen (1995), Dacunha–Castelle and Florens-Zmirou (1986), Genon–Catalot and Jacod (1993), Kessler (1997)).

In this paper we suppose that we don't observe the process (X_t) itself but a discrete sampling of the integrated process $I_t = \int_0^t X_s ds$.

Integrals of diffusion processes have recently been considered in the field of finance in relation with the stochastic volatility models (see e.g. Ghysels *et al.* (1996) for a survey of these models). Data may be obtained from option prices and their associated implied volatilities (see e.g., Pastorello *et al* (1994), Taylor and Xu (1994, 1995)).

The process $(I_t)_{t \geq 0}$ is discretely observed with a regular sampling interval Δ. For a general diffusion X, the exact distribution of a n–sample $(I_{i\Delta},\ i \leq n)$ is not explicit. Therefore, we consider one of the few models for which computations are possible in order to try and compare different inference methods in view of further generalizations. In this paper, we study the case where (X_t) is a strictly stationary Ornstein–Uhlenbeck process:

$$dX_t = \mu X_t dt + \sigma dB_t.$$

The unknown parameter $\theta = (\mu, \sigma^2)$ is to be estimated from the observation of $(I_{i\Delta}, i \leq n)$ which is equivalent to the observation of the increments $(J_i = I_{(i+1)\Delta} - I_{i\Delta},\ i \leq n-1)$.

We first investigate the probabilistic properties of the process $(J_i, i \in \mathbb{N})$ (Section 2). It is a Gaussian ARMA(1,1) process with exponentially decaying α–mixing coefficient.

In Section 3, we study efficient estimators of θ. Although the exact distribution of $(J_i, i \leq n)$ is explicit, the likelihood function is hardly tractable. So we study the Whittle estimator of θ which is known to be asymptotically equivalent to the maximum likelihood estimator (see Dzhaparidze and Yaglom (1983)). It turns out that the Whittle contrast is explicit (Theorem 3.1). In particular, when μ is known, the Whittle estimator $\hat{\sigma^2}_n$ is explicit and its asymptotic variance is equal to $2\sigma^4$.

In Section 4, noting that $(J_i, i \in \mathbb{N})$ is a special case of Hidden Markov Chain, we use Ryden's approach to build other estimators namely the maximum likelihood split data estimator (MLSDE) (see Ryden (1994)). For m integer, the MLSDE $\tilde{\theta}_n^{(m)}$ is based on the maximization of the likelihood of a n-sample distributed as $(J_0, J_1, \dots, J_{m-1})$. Thus $\tilde{\theta}_n^{(m)}$ uses nm datas. We prove that it is consistent, asymptotically Gaussian and give an expression for its asymptotic covariance matrix.

Section 5 deals with the comparison of asymptotic covariances of the previous estimators. We give numerical values and a theoretical comparison. The conclusions are the following. For large values of m, the MLSDE and the Whittle estimator have a close asymptotic behaviour. For small Δ, the results are more surprising. For the drift parameter, the asymptotic variances of estimators are similar. For the diffusion coefficient parameter, they are quite different. The asymptotic variance of $\hat{\sigma^2}_n$ does not depend on Δ, whereas the asymptotic variance of $\tilde{\sigma^2}_n^{(m)}$ is of order $\frac{1}{\Delta}$ (Theorem 5.1).

In Section 6, we discuss possibilities of extensions of the two methods to more general diffusion processes.

2 Model and assumptions

In this section, we describe the probabilistic properties of the observed process.

Let $(C = \mathcal{C}(\mathbb{R}_+, \mathbb{R}), \mathcal{C}, (\mathcal{C}_t)_{t\geq 0}, (X_t, t \geq 0), \mathbb{P}_\theta)$ be the canonical probability space associated with the observation of a strictly stationary Ornstein-Uhlenbeck process with parameter $\theta = (\mu, \sigma^2)$: for $t \geq 0$, X_t is defined on C by $X_t(w) = w_t$, $\mathcal{C}_t = \sigma(X_s,\ 0 \leq s \leq t)$, $\mathcal{C} = \sigma(X_t,\ t \geq 0)$; $\theta = (\mu, \sigma^2)$ is a two-dimensional parameter: $\theta_1 = \mu$ is negative, and $\theta_2 = \sigma^2$ positive; $\mathbb{P}_\theta$ is the probability on $(C, \mathcal{C})$, such that, under $\mathbb{P}_\theta$, there exists a standard Brownian motion $(B_t^\theta, t \geq 0)$, adapted to $(\mathcal{C}_t)_{t\geq 0}$ and such that the canonical process $(X_t)_{t\geq 0}$ is solution of:

$$dX_t = \mu X_t dt + \sigma dB_t^\theta \tag{1}$$

with X_0 centered, Gaussian, with variance $\frac{\sigma^2}{2|\mu|}$ and independent of B^θ.

Under $\mathbb{P}_\theta$, the process $(X_t, t \geq 0)$ is a stationary Ornstein-Uhlenbeck process. Solving (1), we obtain, for all $t \geq 0$ and $h \geq 0$:

$$X_{t+h} = e^{\mu h} X_t + e^{\mu(t+h)} \sigma \int_t^{t+h} e^{-\mu s} dB_s^\theta \tag{2}$$

For Δ a positive real and $i \in \mathbb{N}$, let

$$J_i = \int_{i\Delta}^{(i+1)\Delta} X_s ds \quad \text{for } i \in \mathbb{N} \tag{3}$$

The process $(J_i)_{i\geq 0}$ is not Markov, but we can link J_i and J_{i+1} by a relation of ARMA(1,1) type.

Proposition 2.1. *Under $\mathbb{P}_\theta$, for all $i \geq 0$,*

$$J_{i+1} - e^{\mu\Delta} J_i = \frac{\sigma}{\mu} \int_{i\Delta}^{(i+1)\Delta} (e^{\mu\Delta} - e^{\mu((i+1)\Delta - s)}) dB_s^\theta + \frac{\sigma}{\mu} \int_{(i+1)\Delta}^{(i+2)\Delta} (e^{\mu((i+2)\Delta - s)} - 1) dB_s^\theta$$

Hence for all $i \geq 1$, $J_{i+1} - e^{\mu\Delta} J_i$ is independent of $(J_0, \ldots, J_{i-1})$

Proof. See the appendix. □

We define the following expressions

$$r(0,\mu) = \frac{1}{\mu^2} \left(\Delta + \frac{1 - e^{\mu\Delta}}{\mu} \right) \tag{4}$$

$$r(k,\mu) = -\frac{1}{2\mu^3} e^{\mu|k|\Delta} e^{-\mu\Delta} \left(e^{\mu\Delta} - 1 \right)^2 \quad \text{for} \quad k \neq 0 \tag{5}$$

$$A_0(\mu) = \frac{1}{\mu^2} \left(\Delta + \frac{1 - e^{2\mu\Delta}}{\mu} + \Delta e^{2\mu\Delta} \right) \tag{6}$$

$$A_1(\mu) = \frac{1}{2\mu^2} \left(\frac{e^{2\mu\Delta} - 1}{\mu} - 2e^{\mu\Delta}\Delta \right) \tag{7}$$

$$B_0(\mu) = 1 + e^{2\mu\Delta}, \quad B_1(\mu) = -e^{\mu\Delta} \tag{8}$$

Proposition 2.2. *The process $(J_i)_{i\in\mathbb{N}}$ is strictly stationary and Gaussian with for $i, j \geq 0$,*

$$E(J_i) = 0, \quad Var(J_i) = r(0,\mu)\sigma^2, \quad Cov(J_i, J_j) = r(i-j,\mu)\sigma^2 \quad \textit{for} \quad i \neq j.$$

Its spectral density $f(\lambda, \theta)$ has the explicit form:

$$f(\lambda, \theta) = \sigma^2 \frac{A_0(\mu) + 2A_1(\mu)\cos\lambda}{B_0(\mu) + 2B_1(\mu)\cos\lambda} \tag{9}$$

with

$$B_0(\mu) - 2B_1(\mu) > 0 \ , \ A_1(\mu) > 0 \ , \ A_0(\mu) - 2A_1(\mu) > 0. \tag{10}$$

Proof. See the appendix □

An immediate consequence of (10) is that $\inf_{\lambda\in\mathbb{R}} f(\lambda, \theta) > 0$. Due to the form of its spectral density, the process $(J_i)_{i\in\mathbb{N}}$ has an ARMA(1,1) representation (see e.g. Brockwell and Davis (1991)). Moreover, we can study its α–mixing coefficient.

Proposition 2.3. *Let $\alpha_J(k)$ denote the α–mixing coefficient of $(J_i)_{i\in\mathbb{N}}$ (see e.g. Doukhan (1994) chap.1). We have*

$$\alpha_J(k) \leq e^{\mu(k+1)\Delta}.$$

Hence, the process is ergodic.

Proof. Let Q_θ denote the stationary distribution of $(X_t)_{t\geq 0}$, i.e. $Q_\theta = \mathcal{N}\left(0, \frac{\sigma^2}{2|\mu|}\right)$. The infinitesimal generator of the Ornstein-Uhlenbeck process considered as an operator on $L^2(Q_\theta)$ is self-adjoint and has discrete spectrum equal to $\{n\mu, n \in \mathbb{N}\}$ (see Karlin and Taylor (1981) p.332). Thus, using Proposition 1 p.112 in Doukhan (1994), we know that the α-mixing coefficient $\alpha_X(t)$ of $(X_t)_{t\geq 0}$ satisfies:

$$\alpha_X(t) = \alpha\left(\sigma(X_0), \sigma(X_t)\right) \leq e^{\mu t}$$

(The first equality above is valid because $(X_t)_{t\geq 0}$ is a strictly stationary Markov process). Since J_i is $\sigma(X_s, i\Delta \leq s \leq (i+1)\Delta)$ measurable, $\alpha_J(k-1) \leq e^{\mu k\Delta}$. Now, $\mu < 0$ implies $\lim_{k\to\infty} \alpha_J(k) = 0$, which gives the ergodicity. □

3 An efficient estimator

The likelihood function of the $(J_0, \ldots, J_{n-1})$ is explicitly known but its exact formula is difficult to compute. Therefore, instead of the exact likelihood, we shall use its Whittle approximation which provides efficient estimators (see e.g., Dzhaparidze and Yaglom (1983)). This approximation is also studied in Dacunha–Castelle and Duflo (1986) (chapter 3) and called the Whittle contrast.

3.1 The Whittle contrast

Recall the definition of the periodogram $I_n(\lambda)$ for $n \in \mathbb{N}$ and $\lambda \in \mathbb{R}$:

$$I_n(\lambda) = \frac{1}{n}\left|\sum_{p=0}^{n-1} J_p e^{-ip\lambda}\right|^2 \tag{11}$$

The Whittle contrast is given by:

$$U_n(\theta) = \frac{1}{2\pi}\int_{-\pi}^{\pi}\left[\ln f(\lambda,\theta) + \frac{I_n(\lambda)}{f(\lambda,\theta)}\right] d\lambda \tag{12}$$

Let $\hat{\theta}_n = \operatorname{arginf}_{\theta\in\Theta} U_n(\theta)$. This estimator is called the Whittle estimator.

We can actually compute the Whittle contrast explicitly.

Theorem 3.1. *We have*

$$U_n(\theta) = \ln\left(\sigma^2 \frac{A_1(\mu)}{-\xi(\mu)}\right) + \frac{1}{\sigma^2 n}\sum_{k,l=0}^{n-1} J_k J_l c(k-l,\mu) \tag{13}$$

where (see (6) – (8)*)*

$$c(0,\mu) = \left(\frac{B_1(\mu)\xi^2(\mu) + B_0(\mu)\xi(\mu) + B_1(\mu)}{\xi(\mu)\sqrt{A_0^2(\mu) - 4A_1^2(\mu)}} + \frac{B_1(\mu)}{A_1(\mu)}\right) \tag{14}$$

$$c(k,\mu) = \xi(\mu)^{|k|-1}\left(\frac{B_1(\mu)\xi^2(\mu) + B_0(\mu)\xi(\mu) + B_1(\mu)}{\sqrt{A_0^2(\mu) - 4A_1^2(\mu)}}\right) \text{ if } |k| \neq 0 \tag{15}$$

$$\xi(\mu) = \frac{-A_0(\mu) + \sqrt{A_0^2(\mu) - 4A_1^2(\mu)}}{2A_1(\mu)} \tag{16}$$

(Note that, by (10)*,* $\xi(\mu)$ *is well defined and is negative.)*

Proof. For the sake of simplicity, we omit μ in all expressions depending on μ only. Using (11) and (12) we have to prove that

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} \ln f(\lambda,\theta)d\lambda = \ln(\sigma^2\frac{A_1}{-\xi}) \tag{17}$$

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} f(\lambda,\theta)^{-1}d\lambda = \left(\frac{B_1\xi^2 + B_0\xi + B_1}{\xi\sqrt{A_0^2 - 4A_1^2}} + \frac{B_1}{A_1}\right)\frac{1}{\sigma^2} \tag{18}$$

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} e^{ik\lambda} f(\lambda,\theta)^{-1}d\lambda = \xi^{|k|-1}\left(\frac{B_1\xi^2 + B_0\xi + B_1}{\sqrt{A_0^2 - 4A_1^2}}\right)\frac{1}{\sigma^2} \text{ if } |k| \neq 0 \tag{19}$$

First, we state a useful equality (see e.g. Theorem 15.18 p.307 in Rudin (1987)):

$$\text{For } |x| \leq 1, \quad \int_{-\pi}^{\pi} \ln(1 + x^2 - 2x\cos\lambda)d\lambda = 0 \tag{20}$$

Let us prove (17). Using (9),(8) we have

$$\ln f(\lambda,\theta) = \ln(\sigma^2\frac{A_1}{-\xi}) + \ln(-\frac{A_0\xi}{A_1} - 2\xi\cos\lambda) - \ln(1 + e^{2\mu\Delta} - 2e^{\mu\Delta}\cos\lambda)$$

But using (16), we have $A_1\xi^2 + A_0\xi + A_1 = 0$, and so

$$\ln f(\lambda,\theta) = \ln(\sigma^2\frac{A_1}{-\xi}) + \ln(1 + \xi^2 - 2\xi\cos\lambda) - \ln(1 + e^{2\mu\Delta} - 2e^{\mu\Delta}\cos\lambda)$$

Since, by (16), $|\xi| \leq 1$ and $e^{\mu\Delta} \leq 1$, we can apply (20), and this gives (17).

Let us prove (18). We compute:

$$\begin{aligned}\frac{1}{2\pi}\int_{-\pi}^{\pi} f(\lambda,\theta)^{-1}d\lambda &= \frac{1}{2\pi\sigma^2}\int_{-\pi}^{\pi}\left(\frac{B_0 + 2B_1\cos\lambda}{A_0 + 2A_1\cos\lambda}\right)d\lambda \\ &= \frac{1}{\sigma^2 2i\pi}\int_{\mathbb{U}} \frac{B_0 + B_1(z + z^{-1})}{(A_0 + A_1(z + z^{-1}))z}dz, \text{ where } \mathbb{U} \text{ is the unit circle} \\ &= \frac{1}{\sigma^2 2i\pi}\int_{\mathbb{U}} \frac{B_1 z^2 + B_0 z + B_1}{(A_1 z^2 + A_0 z + B_1)z}dz\end{aligned}$$

Using the residue Theorem we have (with $G(z) = \frac{B_1 z^2 + B_0 z + B_1}{(A_1 z^2 + A_0 z + B_1) z}$ for the sake of simplicity)

$$c(0) = \sum_{\alpha \text{ pole of } G, \ |\alpha|<1} res(G, \alpha)$$

The pole zero has residue $\frac{B_1}{A_1}$, and ξ is the only other pole with residue $\frac{B_1 \xi^2 + B_0 \xi + B_1}{\xi \sqrt{A_0^2 - 4A_1^2}}$. This gives (18).

To get (19), we compute the expression of $\frac{1}{2\pi} \int_{-\pi}^{\pi} e^{ik\lambda} f(\lambda, \theta)^{-1} d\lambda$, in the same way as for (18). □

3.2 Properties of the Whittle estimator

To study the Whittle estimator, we assume that $\Theta = [\underline{\mu}, \overline{\mu}] \times [\underline{\sigma^2}, \overline{\sigma^2}]$ with $\underline{\mu} < \overline{\mu} < 0$, $0 < \underline{\sigma^2} < \overline{\sigma^2}$. The assumption of compacity for Θ is used in Dacunha–Castelle and Duflo (1986) in order to simplify the proof of consistency of minimum contrast estimators (see Dacunha–Castelle and Duflo (1986), Theorem 3.2.8). We denote by $\theta_0 = (\mu_0, \sigma_0^2)$ the true value of the parameter and assume that $\theta_0 \in \overset{\circ}{\Theta}$.

Let $I(\theta)$ be the 2×2 matrix defined for $\theta \in \overset{\circ}{\Theta}$, by

$$\text{for } i, j \in \{1, 2\} \quad I(\theta)_{i,j} = \frac{1}{4\pi} \int_{-\pi}^{\pi} \frac{\partial}{\partial \theta_i} \ln f(\lambda, \theta) \frac{\partial}{\partial \theta_j} \ln f(\lambda, \theta) d\lambda \tag{21}$$

Proposition 3.2. 1) *For all $\theta \in \overset{\circ}{\Theta}$, the matrix $I(\theta)$ is non singular.*

2) $\sqrt{n}(\hat{\theta}_n - \theta_0) \xrightarrow{n \to \infty} \mathcal{N}(0, I^{-1}(\theta_0))$ *in distribution under* $\mathbb{P}_{\theta_0}$.

Proof. Provided $I(\theta_0)$ is non singular, 2) is easily obtained by a classical proof (see Dacunha–Castelle and Duflo (1986), Dzhaparidze and Yaglom (1983)).

Let us prove 1). By noticing, using (9), that $f(\lambda, \theta) = \sigma^2 g(\lambda, \mu)$, we obtain:

$$\frac{\partial}{\partial \sigma^2} \ln(f(\lambda, \theta)) = \frac{1}{\sigma^2}$$

Hence the matrix $I(\theta)$ is:

$$I(\theta) = \begin{bmatrix} \frac{1}{4\pi} \int_{-\pi}^{\pi} \left(\frac{\partial}{\partial \mu} \ln g(\lambda, \mu) \right)^2 d\lambda & \frac{1}{4\pi\sigma^2} \int_{-\pi}^{\pi} \left(\frac{\partial}{\partial \mu} \ln g(\lambda, \mu) \right) d\lambda \\ \frac{1}{4\pi\sigma^2} \int_{-\pi}^{\pi} \left(\frac{\partial}{\partial \mu} \ln g(\lambda, \mu) \right) d\lambda & \frac{1}{2\sigma^4} \end{bmatrix} \tag{22}$$

Suppose that $I(\theta)$ is singular, then $\det I(\theta) = 0$. Because of (22) we have $\int_{-\pi}^{\pi} \left(\frac{\partial}{\partial \mu} \ln g(\lambda, \mu) \right)^2 d\lambda = \left(\int_{-\pi}^{\pi} \left(\frac{\partial}{\partial \mu} \ln g(\lambda, \mu) \right) d\lambda \right)^2$; but equality in the Cauchy-Schwarz inequality implies that $\frac{\partial}{\partial \mu} \ln g(\lambda, \mu)$ is independent of λ. We deduce that

$$\frac{\partial}{\partial \mu} \left(\frac{\partial}{\partial \lambda} \ln g(\lambda, \mu) \right) = \frac{\partial}{\partial \lambda} \left(\frac{\partial}{\partial \mu} \ln g(\lambda, \mu) \right) = 0$$

Derivating (9), we find

$$\left(\frac{\partial}{\partial\lambda}\ln g(\lambda,\mu)\right)_{|\lambda=\frac{\pi}{2}} = \frac{-2A_1}{A_0} - \frac{-2B_1}{B_0}$$

where this expression should not depend on μ. Replacing (6) and (7) above yields the fact that

$$-\frac{\frac{e^{2\mu\Delta}-1}{\mu} - 2e^{\mu\Delta}\Delta}{\Delta + \frac{1-e^{2\mu\Delta}}{\mu} + \Delta e^{2\mu\Delta}} - \frac{2e^{\mu\Delta}}{1+e^{2\mu\Delta}} \text{ should be independent of } \mu$$

By letting $\mu \to -\infty$ we find it is equal to zero. Hence, for all μ,

$$-\left(e^{2\mu\Delta} - 1 - 2\mu\Delta e^{\mu\Delta}\right)\left(1+e^{2\mu\Delta}\right) - \left(\mu\Delta + 1 - e^{\mu\Delta} + \mu\Delta e^{2\mu\Delta}\right)2e^{\mu\Delta} = 0$$

This is absurd. So $I^{-1}(\theta)$ exists. □

Remark 3.3. *If μ_0 is known, then the Whittle estimator $\hat{\sigma^2}_n$ of σ_0^2 is given by*

$$\hat{\sigma^2}_n = \frac{1}{n}\sum_{k,l=0}^{n-1} J_k J_l c(k-l,\mu_0)$$

and satisfies $\sqrt{n}(\hat{\sigma^2}_n - \sigma_0^2) \xrightarrow{n\to\infty} \mathcal{N}(0, 2\sigma_0^4)$. This is the same asymptotic distribution as the MLE of σ^2 based on the observation of $(X_{i\Delta})_{i\leq n}$.

4 Maximum likelihood split data estimator

4.1 Introduction and notations

The process $(J_i)_{i\in\mathbb{N}}$ is not Markov, but is a deterministic function of the 2-dimensional Markov chain $(J_i, X_{(i+1)\Delta})_{i\in\mathbb{N}}$. This is a special case of Hidden Markov Model, therefore we use Ryden's idea (Ryden (1994)). We split the observation into groups of fixed size, consider these groups as independent and then maximize the resulting likelihood. The resulting estimator is called the maximum likelihood split data estimator (MLSDE). In this section, we prove the consistency and asymptotic normality of the MLSDE.

Let m be an integer, $m \geq 1$. Define for $i = 0, 1, \ldots, n-1$,

$$K_i^{(m)} = (J_{im}, J_{im+1}, \ldots, J_{im+m-1})^* \tag{23}$$

If x is a vector or a matrix, we denote by x^* its transpose. Since m is fixed throughout this section, we shall set,

$$K_i^{(m)} = K_i \tag{24}$$

The process $(K_i, i \in \mathbb{N})$ is ergodic and its α-mixing coefficient satisfies $\alpha_K(k) \leq \alpha_J((k+1)m)$. By $\inf_{\lambda \in \mathbb{R}} f(\lambda, \theta) > 0$, the covariance matrix of K_0,

$$\left(\sigma^2 M_{i,j}^{(m)}(\mu)\right)_{0 \leq i,j \leq m-1} = \sigma^2 \left(r(i-j,\mu)\right)_{0 \leq i,j \leq m-1}, \tag{25}$$

is invertible.

Let $P_\theta^{(m)}$ be the distribution of K_0, $p^{(m)}(.,\theta)$ its density under $\mathbb{P}_\theta$, and:

$$l^{(m)}(.,\theta) = \ln p^{(m)}(.,\theta) \tag{26}$$

$$U_n^{(m)}(\theta) = \frac{1}{n} \sum_{i=0}^{n-1} l^{(m)}(K_i,\theta), \quad \tilde{\theta}_n^{(m)} = \underset{\theta \in \Theta}{\operatorname{argmax}}\, U_n^{(m)}(\theta)$$

For $i, j \in \{1, 2\}$,

$$I_{i,j}^{(m)}(\theta) = E_\theta \left[\frac{\partial}{\partial \theta_i} l^{(m)}(K_0,\theta) \frac{\partial}{\partial \theta_j} l^{(m)}(K_0,\theta) \right] \tag{27}$$

For $i, j \in \{1, 2\}$, $k \geq 1$,

$$\gamma_{i,j}^{(m)}(k,\theta) = E_\theta \left[\frac{\partial}{\partial \theta_i} l^{(m)}(K_0,\theta) \frac{\partial}{\partial \theta_j} l^{(m)}(K_k,\theta) \right] \tag{28}$$

4.2 Asymptotic behaviour of the maximum likelihood split data estimator

Before stating results for $\tilde{\theta}_n^{(m)}$ we need two preliminary propositions. The first one is the identifiability assumption.

Proposition 4.1. *If $m \geq 2$, then $P_\theta^{(m)} = P_{\theta'}^{(m)}$ if and only if $\theta = \theta'$.*

Proof. Assume that $\theta = (\mu, \sigma^2)$, $\theta' = (\mu', {\sigma^2}')$ and $P_\theta^{(m)} = P_{\theta'}^{(m)}$. Since $P_\theta^{(m)}$ is a m-dimensional Gaussian law and $m \geq 2$, $P_\theta^{(m)} = P_{\theta'}^{(m)}$ implies the equality between the variance of J_0 and the covariance of (J_0, J_1) under $\mathbb{P}_\theta$ and $\mathbb{P}_{\theta'}$. By Proposition 2.2 and (4), (5):

$$\frac{\sigma^2}{\mu^2}\left(\Delta + \frac{1 - e^{\mu\Delta}}{\mu}\right) = \frac{{\sigma^2}'}{{\mu'}^2}\left(\Delta + \frac{1 - e^{\mu'\Delta}}{\mu'}\right) \tag{29}$$

$$\frac{\sigma^2}{2\mu^3}\left(1 - e^{\mu\Delta}\right)^2 = \frac{{\sigma^2}'}{2{\mu'}^3}\left(1 - e^{\mu'\Delta}\right)^2$$

It follows by a simple calculation that $\mu = \mu'$ and $\sigma^2 = {\sigma^2}'$. □

Remark 4.2. *If $m = 1$, then only one parameter may be identified.*

Now, for the asymptotic normality, we need the following result.

Proposition 4.3. *For $m \geq 2$ and $\theta \in \overset{\circ}{\Theta}$, $I^{(m)}(\theta)$ is non singular.*

Proof. Assume that $I^{(m)}(\theta)$ is singular, then $\det I^{(m)}(\theta)=0$. By (27),

$$E_\theta\left[\left(\frac{\partial}{\partial\mu}l^{(m)}(K_0,\theta)\right)^2\right]E_\theta\left[\left(\frac{\partial}{\partial\sigma^2}l^{(m)}(K_0,\theta)\right)^2\right]$$
$$=\left(E_\theta\left[\frac{\partial}{\partial\mu}l^{(m)}(K_0,\theta)\frac{\partial}{\partial\sigma^2}l^{(m)}(K_0,\theta)\right]\right)^2$$

This equality in the Cauchy-Schwarz inequality implies that there exits a constant $c(\theta)$ such that (recall (26)):

$$\frac{\partial}{\partial\mu}\ln p^{(m)}(x,\theta)=c(\theta)\frac{\partial}{\partial\sigma^2}\ln p^{(m)}(x,\theta)\quad\forall x\ \in\mathbb{R}^m \tag{30}$$

Using the fact that the covariance matrix of K_0 is $\sigma^2M^{(m)}(\mu)$, we get:

$$\ln p^{(m)}(x,\theta)=-\frac{1}{2}\left[m2\pi+m\ln\sigma^2+\ln\det M^{(m)}(\mu)+\frac{1}{\sigma^2}x^*\left(M^{(m)}(\mu)\right)^{-1}x\right]$$

$$\frac{\partial}{\partial\sigma^2}\ln p^{(m)}(x,\theta)=-\frac{1}{2}\left[\frac{m}{\sigma^2}-\frac{1}{\sigma^4}x^*(M^{(m)}(\mu))^{-1}x\right] \tag{31}$$

$$\frac{\partial}{\partial\mu}\ln p^{(m)}(x,\theta)=-\frac{1}{2}\left[\frac{\partial}{\partial\mu}(\ln\det M^{(m)}(\mu))+\frac{1}{\sigma^2}x^*\frac{\partial}{\partial\mu}\left((M^{(m)}(\mu))^{-1}\right)x\right]$$

So, by (30), we must have $\frac{\partial}{\partial\mu}\left((M^{(m)}(\mu))^{-1}\right)=-\frac{c(\theta)}{\sigma^2}(M^{(m)}(\mu))^{-1}$; since $M^{(m)}(\mu)$ does not depend on σ^2, the same is true for $-\frac{c(\theta)}{\sigma^2}$. Set $\tilde{c}(\mu)=-\frac{c(\theta)}{\sigma^2}$, then we have:

$$\frac{\partial}{\partial\mu}\left((M^{(m)}(\mu))^{-1}\right)=\tilde{c}(\mu)(M^{(m)}(\mu))^{-1}$$

We can solve this equation: $(M^{(m)}(\mu))^{-1}=(M^{(m)}(\mu_0))^{-1}\exp\left(\int_{\mu_0}^{\mu}\tilde{c}(s)ds\right)$. Hence,

$$M^{(m)}(\mu)=M^{(m)}(\mu_0)\exp\left(-\int_{\mu_0}^{\mu}\tilde{c}(s)ds\right)$$

But this implies that $M^{(m)}_{0,0}(\mu)$ and $M^{(m)}_{0,1}(\mu)$ have the same asymptotic behaviour as $\mu\to-\infty$ ($\sim$ constant $\exp\left(-\int_{\mu_0}^{\mu}\tilde{c}(s)ds\right)$). And this is absurd: $M^{(m)}_{0,0}(\mu)$ is of order μ^{-2}, and $M^{(m)}_{0,1}(\mu)$ is of order μ^{-3} (using (4) and (5)). □

We can now prove that $\tilde{\theta}_n^{(m)}$ is asymptotically normal.

Theorem 4.4. *Assume $m\geq 2$ then*

$$H^{(m)}_{i,j}(\theta)=\sum_{k=1}^{\infty}\gamma^{(m)}_{i,j}(k,\theta),\quad\Gamma^{(m)}_{i,j}(\theta)=I^{(m)}_{i,j}(\theta)+2H^{(m)}_{i,j}(\theta) \tag{32}$$

are well defined (for $i,j\in\{1,2\}$ and $\theta\in\overset{\circ}{\Theta}$) and

$$\sqrt{n}(\tilde{\theta}_n^{(m)}-\theta_0)\xrightarrow{n\to\infty}\mathcal{N}\left(0,I^{(m)}(\theta_0)^{-1}\Gamma^{(m)}(\theta_0)I^{(m)}(\theta_0)^{-1}\right)$$
$$=\mathcal{N}\left(0,I^{(m)}(\theta_0)^{-1}+2I^{(m)}(\theta_0)^{-1}H^{(m)}(\theta_0)I^{(m)}(\theta_0)^{-1}\right)$$

in distribution under $\mathbb{P}_{\theta_0}$.

Proof. Consistency is obtained by adapting Ryden's proof (Ryden (1994)) to this model. We only consider the asymptotic normality.

Denote
$\nabla l^{(m)}(x,\theta) = \left(\frac{\partial}{\partial\mu} l^{(m)}(x,\theta), \frac{\partial}{\partial\sigma^2} l^{(m)}(x,\theta)\right)$ and $\nabla^2 l^{(m)}(x,\theta) = \left[\frac{\partial^2}{\partial\theta_i\theta_j} l^{(m)}(x,\theta)\right]_{i,j\in\{1,2\}}$.
By using standard arguments it is enough to prove

1) $\frac{1}{n}\sum_{i=0}^{n-1} \nabla^2 l^{(m)}(K_i,\theta_0) \xrightarrow{n\to\infty} -I^{(m)}(\theta_0)$ $\mathbb{P}_{\theta_0} a.s.$

2) $\frac{1}{\sqrt{n}}\sum_{i=0}^{n-1} \nabla l^{(m)}(K_i,\theta_0) \xrightarrow{n\to\infty} \mathcal{N}(0,\Gamma^{(m)}(\theta_0))$, in law under $\mathbb{P}_{\theta_0}$ and $\Gamma^{(m)}(\theta_0)$ is well defined.

3) For i,j,k in $\{1,2\}$, $\sup_{n\in\mathbb{N},\theta\in\Theta} \frac{1}{n}\sum_{i=0}^{n-1} \left|\frac{\partial^3}{\partial\theta_i\theta_j\theta_k} l^{(m)}(K_i,\theta)\right|$ is bounded in $\mathbb{P}_{\theta_0}$ probability.

Point 1) is obtained by the ergodicity of $(K_i,\ i\geq 0)$, since $E_{\theta_0}(\nabla^2 l^{(m)}(K_0,\theta_0)) = -I^{(m)}(\theta_0)$.

To obtain 2) by Theorem 1 p.46 in Doukhan (1994), it is enough to show that $H^{(m)}(\theta)$ is well defined . (Because the other assumptions of the theorem are easy to check: $E[|K_0|^{2+\delta}] < \infty$ and $\sum_{k=0}^{\infty} \alpha_K(k)^{\frac{\delta}{2+\delta}} < \infty$, for some $\delta > 0$, since $\alpha_K(k-1) \leq \alpha_J(km)$.)

Applying the first covariance inequality given in Doukhan (1994), Theorem 3 p.9 we get, for all $\theta \in \overset{\circ}{\Theta}$, (see (28))

$$\begin{aligned}
\gamma_{i,j}^{(m)}(k,\theta) &\leq 8\alpha_J^{\frac{1}{2}}(mk)\left\{E_\theta\left[\frac{\partial}{\partial\theta_i} l^{(m)}(K_0,\theta)\right]^4 E_\theta\left[\frac{\partial}{\partial\theta_j} l^{(m)}(K_k,\theta)\right]^4\right\}^{\frac{1}{4}} \\
&= 8\alpha_J^{\frac{1}{2}}(mk)\left\{E_\theta\left[\frac{\partial}{\partial\theta_i} l^{(m)}(K_0,\theta)\right]^4 E_\theta\left[\frac{\partial}{\partial\theta_j} l^{(m)}(K_0,\theta)\right]^4\right\}^{\frac{1}{4}} .
\end{aligned}$$

Using Proposition 2.3, we see that $\sum_{k=0}^{\infty} \alpha_J^{\frac{1}{2}}(mk) \leq \infty$. Hence, $H^{(m)}(\theta)$ is well defined.

Finally, for i,j,k in $\{1,2\}$ and $n > 0$,

$$\sup_{\theta\in\Theta} \frac{1}{n}\sum_{i=0}^{n-1}\left|\frac{\partial^3}{\partial\theta_i\theta_j\theta_k} l^{(m)}(K_i,\theta)\right| \leq \frac{1}{n}\sum_{i=0}^{n-1} \sup_{\theta\in\Theta}\left|\frac{\partial^3}{\partial\theta_i\theta_j\theta_k} l^{(m)}(K_i,\theta)\right|,$$

which converges when, $n\to\infty$, to $E_{\theta_0}\left[\sup_{\theta\in\Theta}\left|\frac{\partial^3}{\partial\theta_i\theta_j\theta_k} l^{(m)}(\theta)\right|\right]$.

So we have the result. □

Remark 4.5. *If μ_0 is known, using* (31) *we get:*

$$\tilde{\sigma^2}_n^{(m)} = \frac{1}{mn}\sum_{k=0}^{n-1} K_k^*(M^{(m)}(\mu_0))^{-1} K_k$$

5 Comparison of the theoretical asymptotic variances

In this section our aim is to compare the efficient estimator $\hat{\theta}_n$ with the MLSD estimators $\tilde{\theta}_n^{(m)}$ by means of their asymptotic theoretical variances for different values of Δ and m.

Clearly, when m increases, the MLSD estimator must behave better: indeed, the asymptotic covariance matrices of $\sqrt{nm}(\hat{\theta}_{mn} - \theta_0)$ and $\sqrt{nm}(\tilde{\theta}_n^{(m)} - \theta_0)$ tend to be similar as m becomes large (see Table 2).

When Δ varies, for fixed m, the results are more surprising. The asymptotic variances of $\hat{\mu}_n$ and $\tilde{\mu}_n^{(m)}$ are very similar: Table 1 shows that this variance is high for small Δ. This is consistent with the usual results of drift estimation for diffusions based on the discrete observation of the diffusion itself (Dacunha-Castelle and Florens-Zmirou (1986)), where the asymptotic variance is shown to be of order $O(\frac{1}{\Delta})$.

On the contrary, the asymptotic variances of $\hat{\sigma^2}_n$ and $\tilde{\sigma^2}_n^{(m)}$ behave differently as shown in Table 2. When Δ is small, the variance of the MLSD estimator is high. Indeed, the numerical results are confirmed by the theorical result of Theorem 5.1.

Table 1: We assume that $\sigma_0 = 1$ is known, and $\mu_0 = -1$, the figures take account of the fact that $\tilde{\mu}_n^{(2)}$ uses $2n$ datas.

Theoretical asymptotic variances of the estimator [1]				
	$\Delta = 2$	$\Delta = 1$	$\Delta = 0.1$	$\Delta = 0.01$
$Var\ \sqrt{n}(\hat{\mu}_n - \mu_0)$	1.1	2.0	20.0	200.0
$Var\ \sqrt{2n}(\tilde{\mu}_n^{(2)} - \mu_0)$	1.1	2.0	20.0	200.0

Table 2: We assume that $\mu_0 = -1$ is known, and $\sigma_0 = 1$, the figures take account of the fact that $\tilde{\sigma^2}_n^{(m)}$ uses mn datas.

Theoretical asymptotic variances of the estimator [1]				
	$\Delta = 2$	$\Delta = 1$	$\Delta = 0.1$	$\Delta = 0.01$
$Var\ \sqrt{n}(\hat{\sigma^2}_n - \sigma^2)$	2	2	2	2
$Var\ \sqrt{2n}(\tilde{\sigma^2}_n^{(2)} - \sigma_0^2)$	2.2	2.7	7.4	52.4
$Var\ \sqrt{4n}(\tilde{\sigma^2}_n^{(4)} - \sigma_0^2)$	2.0	2.3	3.6	14.8
$Var\ \sqrt{8n}(\tilde{\sigma^2}_n^{(8)} - \sigma_0^2)$	2.1	2.2	2.6	5.3

The rest of the section is now devoted to the theoretical study of matrices $I^{(m)}(\theta)$, $H^{(m)}(\theta)$, and $\Gamma^{(m)}(\theta)$. As $m \to \infty$, it is possible to prove that $I^{(m)}(\theta)^{-1}\Gamma^{(m)}(\theta)I^{(m)}(\theta)^{-1} \sim_{m\to\infty} m^{-1}I(\theta)$ (a detailed proof is available upon request). The latter property is consistent with the numerical results of Table 2.

For $m = 2$, as $\Delta \to 0$, the following theorem precises the difference between both types of estimator.

Theorem 5.1. *In the case $m = 2$, we may precise the following expressions for $I^{(2)}(\theta)$ and $\Gamma^{(2)}(\theta)$*

$$I^{(2)}_{2,2}(\theta) = \frac{1}{\sigma^4}, \quad \Gamma^{(2)}_{2,2}(\theta) \sim_{\Delta\to 0} \frac{1}{4\sigma^4\,|\mu|\,\Delta}. \tag{33}$$

Hence, if μ_0 is known, the asymptotic variance of $\tilde{\sigma^2}^{(2)}_n$ is equivalent when $\Delta \to 0$ to $\frac{\sigma_0^4}{4|\mu_0|\Delta}$.

Proof. $I^{(2)}_{2,2}(\theta)$ is the Fisher information for the parameter σ^2 of a Gaussian vector with covariance matrix $\sigma^2 M^{(2)}(\mu)$ of size 2×2. So we know that $I^{(2)}_{2,2}(\sigma^2) = \frac{1}{\sigma^4}$.

To prove the second part of (33), let us calculate the sum $H^{(2)}_{2,2}(\theta)$. We write the following diagonalization of $M^{(2)}(\mu)$ (recall that, by (25), $M^{(2)}(\mu) = \begin{bmatrix} r(0,\mu) & r(1,\mu) \\ r(1,\mu) & r(0,\mu) \end{bmatrix}$):

$$M^{(2)}(\mu) = V^{(2)^*}D^{(2)}(\mu)V^{(2)}$$

$$V^{(2)} = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}, \quad D^{(2)}(\mu) = \begin{bmatrix} r(0,\mu)+r(1,\mu) & 0 \\ 0 & r(0,\mu)-r(1,\mu) \end{bmatrix}$$

Denote $L^{(2)}_k = V^{(2)}K_k$, for $k = 0,\dots,n-1$ and $N^{(2)}_i, \quad i = 2k,\ 2k+1$ the components of $L^{(2)}_k$:

$$L^{(2)}_0 = \begin{bmatrix} N^{(2)}_0 \\ N^{(2)}_1 \end{bmatrix}, \ L^{(2)}_1 = \begin{bmatrix} N^{(2)}_2 \\ N^{(2)}_3 \end{bmatrix}, \ \dots, \ L^{(2)}_{n-1} = \begin{bmatrix} N^{(2)}_{2(n-1)} \\ N^{(2)}_{2n-1} \end{bmatrix}$$

Now, formula (31) writes:

$$\frac{\partial}{\partial\sigma^2}\ln p^{(2)}(K_k,\theta) = \frac{1}{2\sigma^4}\left[\left(\frac{{N^{(2)}_{2k}}^2}{r(0,\mu)+r(1,\mu)} - \sigma^2\right) + \left(\frac{{N^{(2)}_{2k+1}}^2}{r(0,\mu)-r(1,\mu)} - \sigma^2\right)\right]$$

[1] We have calculated $var\hat{\mu}_n$ numerically with the formulae (9) and (21) the integral being classically approximated.

We have calculated $var\tilde{\mu}^{(2)}_n$ and $var\tilde{\sigma^2}^{(m)}_n$ numerically with the formulae (27), (28), (32) the expectations beeing calculated by using the special Gaussian form of the considered r.v. We numerically diagonalize their covariance matrices to get computation on independent variables.

To compute $\gamma^{(2)}_{2,2}(k,\theta)$ (see (28)), we use the fact that if (Z,Z') is a Gaussian vector with law $\mathcal{N}\left(0,\begin{bmatrix} a & b \\ b & a\end{bmatrix}\right)$, $E[(Z^2-a)(Z'^2-a)] = 2b^2$, and that $Cov_\theta(N_i^{(2)},N_{2k+j}^{(2)}) = e^{2(k-1)\mu\Delta}Cov_\theta(N_i^{(2)},N_j^{(2)})$. we obtain:

$$\gamma^{(2)}_{2,2}(k,\theta) = \frac{e^{4(k-1)\mu\Delta}}{2\sigma^8}S(\Delta) \quad ,$$

where

$$S(\Delta) = \frac{Cov_\theta(N_0^{(2)},N_2^{(2)})^2}{(r(0,\mu)+r(1,\mu))^2} + \frac{Cov_\theta(N_0^{(2)},N_3^{(2)})^2}{r(0,\mu)^2-r(1,\mu)^2} + \frac{Cov_\theta(N_1^{(2)},N_2^{(2)})^2}{r(0,\mu)^2-r(1,\mu)^2} + \frac{Cov_\theta(N_1^{(2)},N_3^{(2)})^2}{(r(0,\mu)-r(1,\mu))^2}$$

Hence,

$$\sum_{k=1}^{\infty}\gamma^{(2)}_{2,2}(k,\theta) = \frac{1}{2\sigma^8(1-e^{4\mu\Delta})}S(\Delta)$$

So, we have to give the limit of $S(\Delta)$ when $\Delta \to 0$. Using $N_0^{(2)} = \frac{1}{\sqrt{2}}(J_0 + J_1), N_1^{(2)} = \frac{1}{\sqrt{2}}(-J_0+J_1), N_2^{(2)} = \frac{1}{\sqrt{2}}(J_2+J_3), N_3^{(2)} = \frac{1}{\sqrt{2}}(J_2-J_3)$, and Proposition 2.2, we obtain:

$$S(\Delta) = \frac{\sigma^4}{4}\frac{(2r(2,\mu)+r(1,\mu)+r(3,\mu))^2}{(r(0,\mu)+r(1,\mu)^2} + \frac{\sigma^4}{4}\frac{(r(1,\mu)-r(3,\mu))^2}{r(0,\mu)^2-r(1,\mu)^2} + \frac{\sigma^4}{4}\frac{(r(3,\mu)-r(1,\mu))^2}{r(0,\mu)^2-r(1,\mu)^2} + \frac{\sigma^4}{4}\frac{(2r(2,\mu)-r(1,\mu)-r(3,\mu))^2}{(r(0,\mu)+r(1,\mu))^2}$$

And we easily deduce that $S(\Delta) \xrightarrow{\Delta\to 0} \sigma^4$, by using the following straightforward equalities (by (4)–(5)):

$$r(0,\mu) = -\frac{\Delta^2}{2\mu} - \frac{\Delta^3}{6} + o(\Delta^3), \quad r(1,\mu) = -\frac{\Delta^2}{2\mu} - \frac{\Delta^3}{2} + o(\Delta^3)$$
$$r(2,\mu) = -\frac{\Delta^2}{2\mu} - \Delta^3 + o(\Delta^3), \quad r(3,\mu) = -\frac{\Delta^2}{2\mu} - \frac{3\Delta^3}{2} + o(\Delta^3).$$

□

6 Conclusions and possible extensions

Let us now draw some conclusions on the two methods in view of possible extensions. Suppose we want to estimate unknown parameters of an ergodic one-dimensional diffusion (X_t) from the observation of the sample $J_i = \int_{i\Delta}^{(i+1)\Delta} X_s ds$, $0 \le i \le n-1$. First note that the exact distribution of a m-tuple $(J_i, i \le m-1)$ is hardly tractable for large m (hence for the whole sample $m=n$). So, actually, we started with the idea of using Ryden's method for small values of m in the general case. In fact, our result enlight the fact that the method will not be appropriate at least for estimating the diffusion coefficient parameters.

Moreover, even for small values of m ($m = 1, 2$) Ryden's likelihood will not be easily computable.

On the contrary, the Whittle contrast seems more suitable for generalization since it relies only on the covariance structure of the J_i's and there are several ergodic diffusions for which these covariances are explicit and simple. Further work is in progess in this direction.

7 Appendix

7.1 Proof of the proposition 2.1

We integrate (2), (with $h = \Delta$), between $i\Delta$ and $(i+1)\Delta$:

$$J_{i+1} - e^{\mu\Delta} J_i = \sigma \int_{i\Delta}^{(i+1)\Delta} e^{\mu(t+\Delta)} \int_t^{t+\Delta} e^{-\mu s} dB_s^\theta dt$$

Using the Fubini Theorem, we get:

$$\begin{aligned} J_{i+1} - e^{\mu\Delta} J_i = \sigma \int_{i\Delta}^{(i+1)\Delta} dB_s^\theta \left(e^{-\mu s} \int_{i\Delta}^{s} e^{\mu(t+\Delta)} dt \right) \\ + \sigma \int_{(i+1)\Delta}^{(i+2)\Delta} dB_s^\theta \left(e^{-\mu s} \int_{s-i\Delta}^{(i+1)\Delta} e^{\mu(t+\Delta)} dt \right) \end{aligned}$$

This gives the results □

7.2 Proof of proposition 2.2

Since $(X_s)_{s\geq 0}$ is a strictly stationary Gaussian process, so is the process $(J_i)_{i\in\mathbb{N}}$. Because the expectation of X_s is zero, by the Fubini Theorem $E[J_i] = 0$, for all i.

Let us calculate the covariance function of $(J_i)_{i\in\mathbb{N}}$. Elementary computations show that covariance function of $(X_s)_{s\geq 0}$ is given by: $Cov(X_s, X_{s'}) = \frac{\sigma^2}{-2\mu} e^{\mu|s'-s|}$.

So, with some computations, for $0 \leq i \leq j$,

$$E(J_i J_j) = \int_{i\Delta}^{(i+1)\Delta} \int_{j\Delta}^{(j+1)\Delta} E[X_s X_{s'}] ds ds' = \sigma^2 r(j-i, \mu)$$

The spectral density $f(\lambda, \theta)$ is given by:

$$f(\lambda, \theta) = \sigma^2 r(0, \mu) + \sum_{k=1}^{\infty} \sigma^2 (e^{i\lambda k} r(k, \mu) + e^{-i\lambda k} r(k, \mu)) = \sigma^2 \frac{A_0 + 2A_1 \cos\lambda}{B_0 + 2B_1 \cos\lambda}$$

The inequalities of (10) are obtained by elementary computations.

References

- Bibby B.M. and Sørensen S. (1995). Martingale estimations functions for discretely observed diffusion processes. *Bernoulli*, **1** p.17–39.

- Brockwell P.J. and Davis R.A. (1991). *Time Series: Theory and Methods.* Springer-Verlag.

- Dacunha-Castelle D. and Duflo M. (1986). *Probability and statistics Volume 2.* Springer–Verlag.

- Dacunha-Castelle D. and Florens-Zmirou D. (1986). Estimation of the coefficients of a diffusion from discrete obsrevations. *Stochastics*, **19**.

- Doukhan P. (1994). *Mixing. Properties and examples.* Springer-Verlag.

- Dzhaparidze K.O. and Yaglom A.M. (1983). Spectrum parameter estimation in time series analysis. In *Developments in Statistics* (P.R. Krishnaiah, ed.) **4** p. 1–181 Academic, New York.

- Genon-Catalot V. and Jacod J. (1993). On the estimation of the diffusion coefficient for multidimensional diffusion processes. *Annales de l'I.H.P.,Probabilité-Statistiques*, **29** p.119–151.

- Ghysels E., Harvey A. and Renault E. (1996), Stochastic volatility. *Handbook of statistics*, **14** p.119–192.

- Karlin S. and Taylor H.M. (1981). *A second course in stochastic processes.* Academic Press, New–York.

- Kessler M. (1997). Estimation of an ergodic diffusion from discrete observations. *Scandinavian Journal of Statistics No.* **24***(2)* p.211–229.

- Kutoyants Yu. A. (1981). *Parameter Estimation for Stochastic Process.* Heldermann, Berlin.

- Pastorello S., Renaud E. and Touzi N. (1994). Statistical inference for random variance option pricing. *Southern European Economics Discussion Series*, D.P.136.

- Rudin W. (1987). *Real and complex analysis, third edition.* Mc.Graw Hill.

- Ryden T. (1994). Consistent and asymptotically normal parameter estimates for hidden markov models. *The annals of statistics*, **22**(4) p.1884–1895.

- Taylor S.J. and Xu X. (1994). The magnitude of implied volatility smiles: theory and empirical evidence for exchange rates. *Rev. Futures Markets*, **13**, p.355–380.
- Taylor S.J. and Xu X. (1995). The incremental volatility information in one million foreign exchange quotations. *Preprint, Conference in high frequency data in finance, Zürich.*