

DIFFUSIONS WITH MEASUREMENT ERRORS. II. OPTIMAL ESTIMATORS

ARNAUD GLOTER¹ AND JEAN JACOD²

Abstract. We consider a diffusion process X which is observed at times i/n for $i = 0, 1, \dots, n$, each observation being subject to a measurement error. All errors are independent and centered Gaussian with known variance ρ_n . There is an unknown parameter to estimate within the diffusion coefficient. In this second paper we construct estimators which are asymptotically optimal when the process X is a Gaussian martingale, and we conjecture that they are also optimal in the general case.

Mathematics Subject Classification. 60J60, 62F12, 62M05.

Received February 2, 2001. Revised October 24, 2001.

1. INTRODUCTION AND RESULTS

1) In this paper we pursue the investigation of statistical inference for diffusion processes observed on a regular grid on the time interval $[0, 1]$, when each measurement is blurred by an error. A general introduction may be found in [3], as well as the proof of the LAN property for Gaussian martingales. Here we consider rather general 1-dimensional diffusions and exhibit estimators which “are likely” to be asymptotically optimal, although we are still unable to fully justify this claim by mathematical proofs. However in the special case of Gaussian martingales these estimators are indeed asymptotically optimal.

2) The setting is as follows. We consider an SDE of the form:

$$dX_t = b_t dt + \sqrt{c(\theta, t, X_t)} dW_t, \quad \mathcal{L}(X_0) = \eta, \quad t \in [0, 1]. \quad (1.1)$$

Here W is a standard Brownian motion, and η is an arbitrary initial law on \mathbb{R} , and b is (non-anticipative) drift term which may depend on the path of X or W . The parameter takes its values in a bounded closed interval Θ of \mathbb{R} . The basic assumptions on the coefficients are (with a dot meaning the derivative w.r.t. θ):

Hypothesis (HS): $(\theta, t, x) \rightsquigarrow c(\theta, t, x)$ is a function from $\Theta \times [0, 1] \times \mathbb{R} \mapsto (0, \infty)$ which is twice continuously differentiable in θ and once continuously differentiable in t , and c , \dot{c} and \ddot{c} are twice continuously differentiable in x and continuous in t ; the process $b = b_t(\omega)$ is optional and locally bounded (locally in time).

Keywords and phrases: Statistics of diffusions, measurement errors, LAN property.

¹ G.R.A.P.E., UMR 5113 du CNRS, Université Montesquieu (Bordeaux), Avenue Léon Duguit, 33608 Pessac, France; e-mail: gloter@montesquieu.u-bordeaux.fr

² Laboratoire de Probabilités et Modèles Aléatoires, UMR 7599 du CNRS, Université Paris 6, 4 place Jussieu, 75252 Paris, France; e-mail: jj@ccr.jussieu.fr

Hypothesis (HS) implies first that (1.1) admits a unique weak solution, denoted by P_θ , and second that this solution is equivalent to the weak solution of the same equation with b identically 0. Our second assumption is an identifiability assumption:

Hypothesis (HI): (i) For $\zeta \neq \theta$ we have $P_\theta(c(\zeta, t, X_t) = c(\theta, t, X_t) \forall t \in [0, 1]) = 0$.
 (ii) We have $P_\theta(\dot{c}(\theta, t, X_t) = 0 \forall t \in [0, 1]) = 0$.

At stage n , we are given an i.i.d. sequence of $\mathcal{N}(0, 1)$ variables (U_i) , independent of W , and we observe the variables

$$Y_i^n = X_{i/n} + \sqrt{\rho_n} U_i, \quad i = 0, \dots, n, \tag{1.2}$$

where ρ_n is a *known* positive number for each n .

The rate of convergence u_n will be the same as in [3], namely we consider the following three cases:

$$\left. \begin{array}{ll} \text{Case 1} & n\rho_n \rightarrow u = 0: \quad \text{take } u_n = 1/\sqrt{n} \\ \text{Case 2} & n\rho_n \rightarrow u \in (0, \infty): \quad \text{take } u_n = 1/\sqrt{n} \\ \text{Case 3} & n\rho_n \rightarrow u = \infty, \quad \sup_n \rho_n < \infty: \quad \text{take } u_n = (\rho_n/n)^{1/4} \end{array} \right\}. \tag{1.3}$$

Set also

$$\iota_u(x, y) = \begin{cases} \frac{y^2}{2x^2} & \text{if } u = 0 \\ \frac{y^2(2 + x/u)}{2\sqrt{u}x^{3/2}(4 + x/u)^{3/2}} & \text{if } 0 < u < \infty \\ \frac{y^2}{8x^{3/2}} & \text{if } u = \infty \end{cases} \tag{1.4}$$

$$I(\theta) = \int_0^1 \iota_u(c(\theta, s, X_s), \dot{c}(\theta, s, X_s)) ds. \tag{1.5}$$

3) In order to completely state our main result, we need to define all variables U_i (up to infinity) at once, and this is most conveniently done in a way which is coherent with the way our observation scheme is related with time. So we may and will assume, without loss of generality, that we have another Brownian motion W' , independent of W (hence of X), and such that at stage n we have

$$U_i = U_i^n = \sqrt{n} \left(W'_{(i+1)/n} - W'_{i/n} \right). \tag{1.6}$$

Below we state our result in the form of “stable convergence in law”, as introduced by Renyi [7], and for which we refer for example to [4]. Let us simply recall that a sequence of real-valued variables V_n on $(\Omega, \mathcal{F}, P_\theta)$ stably converges in law to a limit V defined on an extension $(\bar{\Omega}, \bar{\mathcal{F}}, \bar{P}_\theta)$ of the original space if we have $E_\theta(\Phi f(V_n)) \rightarrow \bar{E}_\theta(\Phi f(V))$ for any bounded \mathcal{F} -measurable variable Φ and any continuous bounded function f on \mathbb{R} .

Theorem 1.1. *Assume (HS), (HI), and also that we are in one of the three cases of (1.3) and that the probability space accommodates the two independent Brownian motions W and W' , with (1.6). Then we can construct estimators $\hat{\theta}_n$ which converge to θ in P_θ -probability, and if further θ is in the interior of Θ the sequence $\frac{1}{u_n}(\hat{\theta}_n - \theta)$*

converges stably in law under P_θ to a limit which, conditionally on the process (X, W, W') , is centered Gaussian with variance $1/I(\theta)$ (recall (1.5)).

The estimators $\hat{\theta}_n$ will be constructed below. Since $I(\theta)$ does not depend on W' , the limit in law above is also Gaussian with variance $1/I(\theta)$ conditionally on (X, W) : this is the only relevant property from a statistical point of view; but getting the above conditionally on (X, W, W') is for free!

A perhaps more usual way of stating the result would be to say that $\frac{1}{u_n}(\hat{\theta}_n - \theta)$ converges in law under P_θ to a limit U which is a mixture of Gaussian laws, and more precisely which satisfies $P(U \leq x) = \int \Phi_u(x)\mu(du)$ where Φ_u is the distribution function of $\mathcal{N}(0, u)$ and μ is the law (under P_θ) of the variable $1/I(\theta)$. But the stable convergence is a necessary step for getting this, and it gives more insight since it relates the asymptotic behaviour of the estimators with the behaviour of the underlying process X itself.

Remarks. 1) In the Gaussian martingale case, that is when $b_t = 0$ and $c(\theta, t, x) = c(\theta, t)$ does not depend on x , in view of Theorem 2.1 of [3] we see that these estimators are asymptotically optimal (observe that (HS) and (HI) imply the assumption (HI $_\theta$) of this theorem).

2) In view of the form of the $I(\theta)$, which as said before is the Fisher information in the Gaussian case, it seems quite likely that our estimators are asymptotically optimal and that the LAMN property with conditional Fisher information $I(\theta)$ holds. However this is not proved yet.

3) In Case 1, we know (see *e.g.* Dohnal [1], or [2]) that with no measurement errors we have the LAMN property with conditional Fisher information $I(\theta)$ above (corresponding to $u = 0$): so in this case the estimators given below are indeed asymptotically optimal.

4) In Case 1 again, it is also interesting to see what happens if we use the optimal estimators for the non-noisy case derived *e.g.* in [2] and plug in the noisy observations. We have done that in a very simple case, where $b_t = 0$ and $c(\theta, s, x) = \theta$ (so X is a Brownian motion with unknown variance): it turns out that the optimal estimators in the non-noisy case remain optimal in the noisy case if $n^{3/2}\rho_n \rightarrow 0$, they still converge with the same rate $(1/\sqrt{n})$ but to a non-centered normal variable if $n^{3/2}\rho_n \rightarrow v \in (0, \infty)$, and they do not even converge with the correct rate if $n^{3/2}\rho_n \rightarrow \infty$.

5) The fact that the errors are Gaussian can probably be dispensed with, but our method of proof explicitly uses this fact.

The paper is organized as follows: in Section 2 we construct $\hat{\theta}_n$. In Section 3 we strengthen the assumption (HS). Section 4 is devoted to some preliminary estimates, additional lemmas are given in Sections 5 and 6, and the proof itself is carried out in Section 7, together with a brief analysis of the example hinted at in Remark 4 above.

2. CONSTRUCTION OF ESTIMATORS

Our estimator $\hat{\theta}_n$ will be a minimum contrast estimator, for a contrast Δ_n which is an approximation of minus the loglikelihood, since the real likelihoods are unavailable here. This contrast function may look difficult to understand, so we give more details than strictly needed. The idea is as follows:

First, we divide the observations into l_n blocks of size k_n , with $l_n k_n$ being as close as possible to n , and forget the last observation in each block. k_n should be small enough, so that the increments $X_{\frac{i}{n}} - X_{\frac{i-1}{n}}$ are very close to being independent and Gaussian within each block, conditionally on the value of X at the left hand point of the block, but for technical reasons it should not be too small either. A “good” choice is as follows: The sequence u_n being given by (1.3), then k_n is an integer such that

$$\frac{nu_n^2}{k_n} \rightarrow 0, \quad \frac{k_n}{nu_n} \rightarrow 0. \tag{2.1}$$

This implies $k_n \rightarrow \infty$ and $k_n/n \rightarrow 0$, and this is achieved by taking for example

$$k_n = \begin{cases} [n^{1/4}] & \text{in Cases 1 and 2} \\ [n^{5/8} \rho_n^{3/8}] & \text{in Case 3.} \end{cases} \tag{2.2}$$

We also set $l_n = [n/k_n]$, and we take n large enough to have $k_n \geq 2$ and $l_n \geq 2$. The times at which the various blocks begin are

$$s_m^n = \frac{k_n(m-1)}{n}, \quad m = 1, \dots, l_n. \tag{2.3}$$

Next, we introduce the $(k_n - 1) \times (k_n - 1)$ -matrix D^n whose entries are

$$D_{i,j}^n = \begin{cases} 2\rho_n & \text{if } i = j \\ -\rho_n & \text{if } |i - j| = 1 \\ 0 & \text{otherwise.} \end{cases} \tag{2.4}$$

The eigenvalues of D^n are

$$\lambda_i^n = 2\rho_n \left(1 - \cos \frac{i\pi}{k_n} \right), \quad i = 1, \dots, k_n - 1, \tag{2.5}$$

and we write $D^n = P^n \Lambda^n P^{n,*}$ where Λ^n is diagonal with entries given above and P^n is orthogonal.

Next we consider the differences between successive observations, renumbering them within each block of size k_n . This leads to set for $m = 1, \dots, l_n$ and $i = 0, 1, \dots, k_n$:

$$\left. \begin{aligned} Y_i^{n,m} &= Y_{k_n(m-1)+i}^n, & U_i^{n,m} &= U_{k_n(m-1)+i}^n, \\ X_i^{n,m} &= X_{\frac{k_n(m-1)+i}{n}}, & W_i^{n,m} &= W_{\frac{k_n(m-1)+i}{n}}, \\ R_i^{n,m} &= Y_i^{n,m} - Y_{i-1}^{n,m} = X_i^{n,m} - X_{i-1}^{n,m} + \sqrt{\rho_n}(U_i^{n,m} - U_{i-1}^{n,m}) & \text{if } i \geq 1 \end{aligned} \right\}. \tag{2.6}$$

Then, conditionally on the value $X_{s_m^n}$, and with $a = \sqrt{c}$, the increments $X_i^{n,m} - X_{i-1}^{n,m}$ for $i = 1, \dots, k_n$ are (approximately) equal under P_θ to $a(\theta, s_m^n, X_{s_m^n})(W_i^{n,m} - W_{i-1}^{n,m})$, hence are i.i.d. with law $\mathcal{N}(0, c(\theta, s_m^n, X_{s_m^n})/n)$, at least if the drift b_t in (1.1) vanishes. Then the vector $R^{n,m} = (R_i^{n,m})_{1 \leq i \leq k_n-1}$ is (approximately) centered Gaussian with covariance matrix $D^n + \frac{c(\theta, s_m^n, X_{s_m^n})}{n} I_{k_n-1}$, where I_d denotes the $d \times d$ -identity matrix.

Since $X_{s_m^n}$ is unobservable, we replace it by the following approximation:

$$\widehat{Y}_m^n = \frac{1}{k_n - 1} \sum_{i=1}^{k_n-1} Y_i^{n,m-1}, \tag{2.7}$$

so under P_θ and conditionally on the value \widehat{Y}_m^n , the vector $R^{n,m}$ is (approximately) centered Gaussian with covariance matrix $C^{n,m}(\theta) = D^n + \frac{c(\theta, s_m^n, \widehat{Y}_m^n)}{n} I_{k_n-1}$.

Now, recall from [3] that the eigenvalues of $C^{n,m}(\zeta)$ are $\rho_n \Phi_j^{n,m}(\zeta)$, where

$$\Phi_j^{n,m}(\zeta) = 2 \left(1 - \cos \frac{j\pi}{k_n} \right) + \frac{c(\zeta, s_m^n, \widehat{Y}_m^n)}{n\rho_n}, \quad j = 1, \dots, k_n - 1, \tag{2.8}$$

and that the same matrix P^n allows to diagonalise $C^{n,m}(\zeta)$ for all values of ζ , as well as D^n . So if we set

$$F_j^{n,m} = \sum_{i=1}^{k_n-1} P_{ij}^n R_i^{n,m}, \quad j = 1, \dots, k_n - 1, \quad m = 1, \dots, l_n, \tag{2.9}$$

the variables $(F_j^{n,m})_{1 \leq j \leq k_n-1}$ are, under P_ζ and conditionally on \widehat{Y}_m^n , approximatively independent, centered Gaussian with variances $\Phi_j^{n,m}(\zeta)$. So a natural contrast function is

$$\Delta_n(\zeta) = \sum_{m=2}^{l_n} \sum_{j=1}^{k_n-1} \left(\frac{(F_j^{n,m})^2}{\rho_n \Phi_j^{n,m}(\zeta)} + \log \Phi_j^{n,m}(\zeta) \right), \tag{2.10}$$

that is $-2 \log Z_n(\zeta)$ where $Z_n(\zeta)$ is the likelihood which we would obtain if the Gaussian approximation made above were exact (we drop the first block $m = 1$ for technical reasons). Then $\widehat{\theta}_n$ is a point achieving the minimum of $\Delta_n(\cdot)$ over Θ (such a point exists since Θ is compact and Δ_n is continuous). Note that $\Delta_n(\zeta)$ can actually be computed from the observations.

3. REDUCTION OF HYPOTHESES

Our aim here is to show that we can replace (HS) in Theorem 1.1 by the following assumption:

Hypothesis (HS’): We have (HS) and $c, \dot{c}, \ddot{c}, c'_t$ (partial derivative in time), c'_x and c''_{xx} (partial derivatives in space) are all bounded, and $\varepsilon = \inf_{\theta,t,x} c(\theta, t, x) > 0$. Further $b_t = 0$ identically.

Note that the function $a = \sqrt{c}$ also satisfies the same hypotheses (with $\sqrt{\varepsilon}$ instead of ε).

Proposition 3.1. *If the results of Theorem 1.1 hold under (HS’) and (HI), they also hold under (HS) and (HI).*

Proof. Let us call (HS’’) the same assumption than (HS’), except that the process b is locally bounded but not necessarily 0. Call (A) the set of results in Theorem 1.1. Denote by Q_θ the probability measure under which the process X solves (1.1) with $b_t = 0$.

1) First, assume (HS’’) and (HI). By (HS’’), the variable $\int_0^1 \frac{b_s^2}{c(\theta, s, X_s)} ds$ is everywhere finite, hence Girsanov’s theorem yields that Q_θ and P_θ are equivalent on the σ -field \mathcal{F}_1 generated by the processes W and W' up to time 1. This yields first that (HI) is satisfied if P_θ is replaced by Q_θ . Second, since stable convergence of \mathcal{F}_1 -measurable variables (like $\frac{1}{u_n}(\widehat{\theta}_n - \theta)$) is preserved by any given (fixed) equivalent change of measure, if (A) holds for Q_θ it also holds for P_θ . But since Q_θ satisfies (HS’) and (HI), our assumption yields that (A) holds for Q_θ , henceforth for P_θ as well.

2) Next we suppose (HS) and (HI). For any $q \in \mathbb{N}$ we can find a coefficient c_q satisfying the assumption (HS’’) and such that: first $c(\zeta, t, x) = c_q(\zeta, t, x)$ for all $\zeta \in \Theta$ and $t \in [0, 1]$ and x with $|x| \leq q$; and second that $c_q(\zeta, t, x) \neq c_q(\theta, t, x)$ and $\dot{c}_q(\theta, t, x) \neq 0$ for all x with $|x| > q$ and $\zeta \neq \theta$ and $t \in [0, 1]$. If $T_q = \inf(t : |X_t| \geq q)$, we get that on the set $\{T_q \geq 1\}$ the process X equals on $[0, 1]$ the process X^q which solves equation (1.1) with c_q instead of c (and the same Brownian motion W).

On the one hand, the estimators $\widehat{\theta}_n^q$ associated with X^q have (A), because for X^q we have (HS’’) and (HI); on the other hand, $\widehat{\theta}_n^q = \widehat{\theta}_n$ on the set $\{T_q \geq 1\}$, and $P_\theta(T_q \geq 1) \rightarrow 1$ as $q \rightarrow 1$: the fact that (A) holds for $\widehat{\theta}_n$ is then trivial. \square

4. PRELIMINARY ESTIMATES

We denote by θ the “true” value of the parameter, and we assume (HS’), (HI), and that $n\rho_n \rightarrow u \in [0, \infty]$. Then our equation has a unique strong solution and the Wiener process W is also adapted to the filtration

generated by X . We also assume that at stage n , the variables U_i are given by (1.6) with another Brownian motion W' independent of W and X_0 (hence of X). We write $P = P_\theta$.

Set $\sigma_n = \frac{k_n}{nu_n^2}$ and $\sigma'_n = \sigma_n \log(e + n\rho_n)$, so that $\sigma_n \leq \sigma'_n$. In view of the definitions of u_n and l_n and (2.1) and of the fact that $u_n \log(n\rho_n) \rightarrow 0$ in Case 3, we get

$$k_n \rightarrow \infty, \quad l_n \rightarrow \infty, \quad \sigma_n \rightarrow \infty, \quad \frac{k_n l_n}{n} \rightarrow 1, \quad l_n u_n^2 \rightarrow 0, \quad \sigma'_n u_n \rightarrow 0. \tag{4.1}$$

In the sequel “universal” constants, which may in fact depend on the coefficients of equation (1.1) and change from line to line, are all denoted by C ; any dependency of further parameters will explicitly be stated, like C_ε for example.

4.1. Some integrals and applications

Exactly as in [3] we set for every $a > 0, x \in \mathbb{R}$: $\phi(x, a) = 2(1 - \cos(x)) + a$. We first observe that

$$0 < \frac{\phi(x, a)}{\phi(x, b)} \leq \begin{cases} 1 & \text{if } a \leq b \\ \frac{a}{b} & \text{if } a > b. \end{cases} \tag{4.2}$$

Next we consider the integrals (where $a > 0, b > 0$; I_2 was already considered in [3]):

$$I(a, b) = \int_0^\pi \log \frac{\phi(x, a)}{\phi(x, b)} dx, \quad I_p(a) = \int_0^\pi \frac{1}{\phi(x, a)^p} dx \quad \text{for } p \in (0, \infty).$$

A simple calculation shows that

$$I(a, b) = 2\pi \log \frac{\sqrt{a} + \sqrt{4+a}}{\sqrt{b} + \sqrt{4+b}}, \quad I_1(a) = \frac{\pi}{\sqrt{a(4+a)}}, \quad I_2(a) = \frac{\pi(2+a)}{a^{3/2}(4+a)^{3/2}}. \tag{4.3}$$

We will also need the following quantities, for $k = 2, 3, \dots$ and $p > 0$:

$$J(a, b, k) = \frac{\pi}{k} \sum_{i=1}^{k-1} \log \frac{\phi(i\pi/k, a)}{\phi(i\pi/k, b)}, \quad J_p(a, k) = \frac{\pi}{k} \sum_{i=1}^{k-1} \frac{1}{\phi(\frac{i\pi}{k}, a)^p}. \tag{4.4}$$

The next estimates are then elementary:

$$|J(a, b, k) - I(a, b)| \leq \frac{\pi}{k} \left| \log \frac{a}{b} \right|, \quad I_p(a) - \frac{\pi}{ka^p} \leq J_p(a, k) \leq I_p(a). \tag{4.5}$$

The first estimate below is trivial, and the second one obtains by observing that $1 - \cos x \geq Ax^2$ for $0 \leq x \leq \pi$ for some $A > 0$ and by making the change of variable $x = y\sqrt{a}$ in the integral defining $I_p(a)$:

$$I_p(a) \leq \frac{C_p}{a^p}, \quad I_p(a) \leq \begin{cases} \frac{C_p}{a^{p-1/2}} & \text{if } p > 1/2 \\ C(1 + |\log a|) & \text{if } p = 1/2. \end{cases} \tag{4.6}$$

As a first application of the above, we give a very useful inequality. With ε being the minimum of the function c (recall (HS')), we set

$$\eta_{n,j} = \frac{1}{n\rho_n \phi(j\pi/k_n, \varepsilon/n\rho_n)}. \tag{4.7}$$

We obviously have

$$0 < \eta_{n,j} \leq \frac{1}{\varepsilon}. \tag{4.8}$$

Furthermore, equations (4.4) and (4.5) yield for $p \geq 1/2$:

$$\sum_{j=1}^{k_n-1} (\eta_{n,j})^p = \frac{k_n}{\pi} \frac{1}{(n\rho_n)^p} J_p \left(\frac{\varepsilon}{n\rho_n}, k_n \right) \leq \frac{k_n}{\pi} \frac{1}{(n\rho_n)^p} I_p \left(\frac{\varepsilon}{n\rho_n} \right).$$

This is smaller than $C_p k_n$ by the first estimate in (4.6), and if further $n\rho_n \geq 1$ it is also smaller than $C_p k_n / \sqrt{n\rho_n}$ if $p > 1/2$ and than $C k_n (1 + \log(n\rho_n)) / \sqrt{n\rho_n}$ if $p = 1/2$. So

$$\sum_{j=1}^{k_n-1} (\eta_{n,j})^p \leq \begin{cases} C_p \sigma_n & \text{if } p > 1/2 \\ C \sigma'_n & \text{if } p = 1/2. \end{cases} \tag{4.9}$$

4.2. More notation

Recall that $a = \sqrt{c}$. Set for $j = 1, \dots, k_n - 1$ (compare with (2.8) and (2.9)):

$$\widehat{X}_m^n = X_0^{n,m}, \quad \Psi_j^{n,m}(\zeta) = 2 \left(1 - \cos \frac{j\pi}{k_n} \right) + \frac{c(\zeta, s_m^n, \widehat{X}_m^n)}{n\rho_n} \tag{4.10}$$

$$S_i^{n,m} = a(\theta, s_m^n, \widehat{X}_m^n)(W_i^{n,m} - W_{i-1}^{n,m}) + \sqrt{\rho_n}(U_i^{n,m} - U_{i-1}^{n,m}), \quad G_j^{n,m} = \sum_{i=1}^{k_n-1} P_{ij}^n S_i^{n,m}. \tag{4.11}$$

Then, analogous to (2.10), we set

$$\Gamma_n(\zeta) = \sum_{m=2}^{l_n} \sum_{j=1}^{k_n-1} \left(\frac{(G_j^{n,m})^2}{\rho_n \Psi_j^{n,m}(\zeta)} + \log \Psi_j^{n,m}(\zeta) \right). \tag{4.12}$$

Most of the rest of the paper is devoted to controlling the difference $\Delta_n - \Gamma_n$ and to study the convergence of Δ_n and its first two derivatives in ζ . For this, we first observe that Δ_n and Γ_n are twice differentiable, with the first two derivatives given by

$$\dot{\Delta}_n(\zeta) = \sum_{m=2}^{l_n} \sum_{j=1}^{k_n-1} \left(\frac{-(F_j^{n,m})^2}{n\rho_n^2 \Phi_j^{n,m}(\zeta)^2} + \frac{1}{n\rho_n \Phi_j^{n,m}(\zeta)} \right) \dot{c}(\zeta, s_m^n, \widehat{Y}_m^n), \tag{4.13}$$

$$\dot{\Gamma}_n(\zeta) = \sum_{m=2}^{l_n} \sum_{j=1}^{k_n-1} \left(\frac{-(G_j^{n,m})^2}{n\rho_n^2 \Psi_j^{n,m}(\zeta)^2} + \frac{1}{n\rho_n \Psi_j^{n,m}(\zeta)} \right) \dot{c}(\zeta, s_m^n, \widehat{X}_m^n), \tag{4.14}$$

$$\begin{aligned} \ddot{\Delta}_n(\zeta) = & \sum_{m=2}^{l_n} \sum_{j=1}^{k_n-1} \left\{ \left(\frac{2(F_j^{n,m})^2}{n^2 \rho_n^3 \Phi_j^{n,m}(\zeta)^3} - \frac{1}{n^2 \rho_n^2 \Phi_j^{n,m}(\zeta)^2} \right) \dot{c}(\zeta, s_m^n, \widehat{Y}_m^n)^2 \right. \\ & \left. + \left(\frac{-(F_j^{n,m})^2}{n\rho_n^2 \Phi_j^{n,m}(\zeta)^2} + \frac{1}{n\rho_n \Phi_j^{n,m}(\zeta)} \right) \ddot{c}(\zeta, s_m^n, \widehat{Y}_m^n) \right\}, \end{aligned} \tag{4.15}$$

$$\begin{aligned} \ddot{\Gamma}_n(\zeta) &= \sum_{m=2}^{l_n} \sum_{j=1}^{k_n-1} \left\{ \left(\frac{2(G_j^{n,m})^2}{n^2 \rho_n^3 \Psi_j^{n,m}(\zeta)^3} - \frac{1}{n^2 \rho_n^2 \Psi_j^{n,m}(\zeta)^2} \right) \dot{c}(\zeta, s_m^n, \widehat{X}_m^n)^2 \right. \\ &\quad \left. + \left(\frac{-(G_j^{n,m})^2}{n \rho_n^2 \Psi_j^{n,m}(\zeta)^2} + \frac{1}{n \rho_n \Psi_j^{n,m}(\zeta)} \right) \ddot{c}(\zeta, s_m^n, \widehat{X}_m^n) \right\}. \end{aligned} \tag{4.16}$$

4.3. Inequalities

Now we derive a number of more or less classical inequalities. The first one, utterly well known, is

$$E \left(\sup_{0 \leq u \leq s} |X_{t+u} - X_t|^p | \mathcal{F}_t \right) \leq C s^{p/2} \tag{4.17}$$

(here (\mathcal{F}_t) is the canonical filtration of the pair (W, W')).

Next, let us introduce some σ -fields:

$$\begin{aligned} \mathcal{G}_j^{n,m} &= \sigma \left(X_0; W_s : s \leq s_m^n + \frac{j}{n}; W'_s : s \leq s_m^n + \frac{j+1}{n} \right), \\ \mathcal{G}_m^n &= \sigma(X_0, W_s, W'_s : s \leq s_m^n). \end{aligned}$$

Observe that

$$\mathcal{G}_{k_n-1}^{n,m-1} \subset \mathcal{G}_m^n \subset \mathcal{G}_{k_n}^{n,m-1} = \mathcal{G}_0^{n,m}$$

and also that all variables in (2.6) are $\mathcal{G}_i^{n,m}$ -measurable. Further, \widehat{Y}_m^n and \widehat{X}_m^n are \mathcal{G}_m^n -measurable, while $F_j^{n,m}$ and $G_j^{n,m}$ are \mathcal{G}_{m+1}^n -measurable. It readily follows, since X is a martingale and W' is a Brownian motion independent of X and W , that

$$E(\widehat{Y}_m^n | \mathcal{G}_{m-1}^n) = E(\widehat{X}_m^n | \mathcal{G}_{m-1}^n) = \widehat{X}_{m-1}^n, \tag{4.18}$$

$$E(F_j^{n,m} | \mathcal{G}_m^n) = 0. \tag{4.19}$$

Using (4.17) and the fact that $\sum_{i=1}^{k_n-1} U_i^{n,m}$ is centered normal with variance $k_n - 1$ and independent of X and \mathcal{G}_m^n , we get (observing also that $k_n^2/n\rho_n \rightarrow \infty$ in all cases by (2.1)):

$$E(|\widehat{Y}_m^n - \widehat{X}_m^n|^p | \mathcal{G}_m^n) \leq C_p \left(\left(\frac{k_n}{n} \right)^{p/2} + \left(\frac{\rho_n}{k_n} \right)^{p/2} \right) \leq C_p \left(\frac{k_n}{n} \right)^{p/2}. \tag{4.20}$$

Our next observation is about the law of the vector $(G_j^{n,m})_{1 \leq j \leq k_n-1}$, conditional on \mathcal{G}_m^n : first the random vector $(S_j^{n,m})_{1 \leq i \leq k_n-1}$ is clearly centered Gaussian with covariance matrix $D^n + (c(\theta, s_m^n, \widehat{X}_m^n)/n)I_{k_n-1}$ (conditional on \mathcal{G}_m^n); second, by [3] this matrix is orthogonalized by the same orthogonal matrix P^n than D^n , and its eigenvalues are $\rho_n \Psi_j^{n,m}(\theta)$ (see (2.5, 4.10)). Then we readily get from (4.11) that

$$\text{conditional on } \mathcal{G}_m^n, \text{ the } \left(\frac{G_j^{n,m}}{\sqrt{\rho_n \Psi_j^{n,m}(\theta)}} \right)_{1 \leq j \leq k_n-1} \text{ are i.i.d. } \mathcal{N}(0, 1), \tag{4.21}$$

and it clearly follows that (here and below $p \in [2, \infty)$):

$$E(G_j^{n,m} | \mathcal{G}_m^n) = 0, \quad E(|G_j^{n,m}|^p | \mathcal{G}_m^n) = C_p \rho_n^{p/2} \Psi_j^{n,m}(\theta)^{p/2}. \tag{4.22}$$

Finally we wish to compare $G_j^{n,m}$ and $F_j^{n,m}$, and this is done in the following lemma:

Lemma 4.1. *We have the following inequalities (with $1 \leq j \leq k_n - 1$ and $p \in [2, \infty)$):*

$$E(|F_j^{n,m} - G_j^{n,m}|^p | \mathcal{G}_m^n) \leq C_p \frac{k_n^{p/2}}{n^p}, \tag{4.23}$$

$$|E((F_j^{n,m})^2 - (G_j^{n,m})^2 | \mathcal{G}_m^n)| \leq C \frac{k_n}{n^2}, \tag{4.24}$$

$$E(|(F_j^{n,m})^2 - (G_j^{n,m})^2|^2 | \mathcal{G}_m^n) \leq C \frac{k_n}{n^2} \left(\frac{k_n}{n^2} + \rho_n \Psi_j^{n,m}(\theta) \right). \tag{4.25}$$

Proof. 1) In all the proof we fix j, m and n , and we set $u_i = s_m^n + i/n$ and $c_i = P_{i,j}^{n,m}$ and $V_i = R_i^{n,m} - S_i^{n,m}$. So $F_j^{n,m} - G_j^{n,m} = \sum_{i=1}^{k_n-1} c_i V_i$, while $V_i = \int_{u_{i-1}}^{u_i} (a(\theta, s, X_s) - a(\theta, s_m^n, \widehat{X}_m^n)) dW_s$.

2) Observing that V_i is $\mathcal{G}_i^{n,m}$ -mesurable and that $E(V_i | \mathcal{G}_{i-1}^{n,m}) = 0$, we can use Rosenthal's inequality to get

$$E(|F_j^{n,m} - G_j^{n,m}|^p | \mathcal{G}_m^n) \leq C_p E \left(\left(\sum_{i=1}^{k_n-1} E(c_i^2 V_i^2 | \mathcal{G}_{i-1}^{n,m}) \right)^{p/2} \middle| \mathcal{G}_m^n \right) + C_p \sum_{i=1}^{k_n-1} E(|c_i V_i|^p | \mathcal{G}_m^n).$$

We have $E(V_i^2 | \mathcal{G}_{i-1}^{n,m}) = E(\int_{u_{i-1}}^{u_i} (a(\theta, s, X_s) - a(\theta, s_m^n, \widehat{X}_m^n))^2 ds | \mathcal{G}_{i-1}^{n,m})$; hence the boundedness of a'_t and a'_x (see (HS')) yields

$$E(c_i^2 V_i^2 | \mathcal{G}_{i-1}^{n,m}) \leq C c_i^2 \frac{1}{n} \left(\left(\frac{k_n}{n} \right)^2 + |X_{i-1}^{n,m} - \widehat{X}_m^n|^2 \right).$$

Using the property $\sum_{i=1}^{k_n-1} c_i^2 = 1$, we deduce from (4.17) again that

$$E \left(\left(\sum_{i=1}^{k_n-1} E(c_i^2 V_i^2 | \mathcal{G}_{i-1}^{n,m}) \right)^{p/2} \middle| \mathcal{G}_m^n \right) \leq C \frac{k_n^{p/2}}{n^p}.$$

On the other hand, Bürkhölder–Davis–Gundy inequality and (HS') and (4.17) yield by the same type of argument:

$$E(|c_i V_i|^p | \mathcal{G}_m^n) \leq C_p c_i^p \frac{k_n^{p/2}}{n^p}.$$

Since $c_i^p \leq c_i^2$ when $p \geq 2$, putting all these together yields (4.23).

3) We have $S_i^{n,m} = V_i' + \sqrt{\rho_n}(U_j^{n,m} - U_{j-1}^{n,m})$, where $V_i' = \int_{u_{i-1}}^{u_i} a(\theta, s_m^n, \widehat{X}_m^n) dW_s$. Then $E(V_i S_i^{n,m} | \mathcal{G}_m^n) = \delta_{i,l} a(\theta, s_m^n, \widehat{X}_m^n) E(\int_{u_{i-1}}^{u_i} (a(\theta, s, X_s) - a(\theta, s_m^n, \widehat{X}_m^n)) ds | \mathcal{G}_m^n)$, and Ito's formula and (1.1) with $b_t = 0$ yield

$$E(a(\theta, s, X_s) - a(\theta, s_m^n, \widehat{X}_m^n) | \mathcal{G}_m^n) = E \left(\int_{u_0}^s \left(a'_t(\theta, w, X_w) + \frac{1}{2} a''_{xx}(\theta, w, X_w) a^2(\theta, w, X_w) \right) dw \middle| \mathcal{G}_m^n \right),$$

whose absolute value is smaller than $C k_n/n$. Hence $|E(V_i V_i' | \mathcal{G}_m^n)| \leq C k_n/n^2$, and

$$|E(G_j^{n,m}(G_j^{n,m} - F_j^{n,m}) | \mathcal{G}_m^n)| = \left| \sum_{i,l=1}^{k_n-1} c_i c_l E(V_i V_l' | \mathcal{G}_m^n) \right| \leq C \frac{k_n}{n^2} \sum_{i=1}^{k_n-1} c_i^2 = C \frac{k_n}{n^2}.$$

In order to obtain (4.24) it remains to observe that $x^2 - y^2 = 2x(x - y) - (x - y)^2$ and to use (4.23).

4) Since $(x^2 - y^2)^2 \leq 8x^2(x - y)^2 + 2(x - y)^4$, we readily deduce (4.25) from (4.22) and (4.23) and Cauchy–Schwarz inequality. \square

Combining (4.22) and (4.23) also yields:

$$E(|F_j^{n,m}|^p | \mathcal{G}_m^n) \leq C_p \left(\frac{k_n^{p/2}}{n^p} + \rho_n^{p/2} \Psi_j^{n,m}(\theta)^{p/2} \right). \tag{4.26}$$

Finally, let us give the following easy consequences of (4.2) and of the definitions of $\eta_{n,j}$ and $\Phi_j^{n,m}$ and $\Psi_j^{n,m}$:

$$|n\rho_n \Phi_j^{n,m}(\zeta)| \leq \frac{C}{\eta_{n,j}}, \quad |n\rho_n \Psi_j^{n,m}(\zeta)| \leq \frac{C}{\eta_{n,j}}. \tag{4.27}$$

5. ESTIMATES FOR $\Delta_n - \Gamma_n$

We wish here to give estimates for the difference $\Delta_n(\zeta) - \Gamma_n(\zeta)$ and its first two derivatives. With ζ being fixed, let us set

$$\psi_{n,m,j}(x) = n\rho_n \phi \left(\frac{j\pi}{k_n}, \frac{c(\zeta, s_m^n, x)}{n\rho_n} \right).$$

In view of (2.10, 4.12–4.15) and (4.16), we see that the differences above are of the form $A_n + B_n + D_n$, where

$$\left. \begin{aligned} A_n &= \sum_{m=2}^{l_n} \alpha_m^n, & \alpha_m^n &= \sum_{j=1}^{k_n-1} \left(g_{n,m,j}(\widehat{Y}_m^n) - g_{n,m,j}(\widehat{X}_m^n) \right) \\ B_n &= \sum_{m=2}^{l_n} \beta_m^n, & \beta_m^n &= \sum_{j=1}^{k_n-1} n(G_j^{n,m})^2 \left(h_{n,m,j}(\widehat{Y}_m^n) - h_{n,m,j}(\widehat{X}_m^n) \right) \\ D_n &= \sum_{m=2}^{l_n} \delta_m^n, & \delta_m^n &= \sum_{j=1}^{k_n-1} n \left((F_j^{n,m})^2 - (G_j^{n,m})^2 \right) h_{n,m,j}(\widehat{Y}_m^n) \end{aligned} \right\} \tag{5.1}$$

provided we take the following functions:

$$\left. \begin{aligned} \text{for } \Delta_n - \Gamma_n : & \quad g_{n,m,j} = \log \psi_{n,m,j}, & h_{n,m,j} &= \frac{1}{\psi_{n,m,j}}, \\ \text{for } \dot{\Delta}_n - \dot{\Gamma}_n : & \quad g_{n,m,j} = \frac{\dot{c}(\zeta, s_m^n, \cdot)}{\psi_{n,m,j}}, & h_{n,m,j} &= -\frac{\dot{c}(\zeta, s_m^n, \cdot)}{\psi_{n,m,j}^2}, \\ \text{for } \ddot{\Delta}_n - \ddot{\Gamma}_n : & \quad g_{n,m,j} = -\frac{\dot{c}(\zeta, s_m^n, \cdot)^2}{\psi_{n,m,j}^2} + \frac{\ddot{c}(\zeta, s_m^n, \cdot)}{\psi_{n,m,j}}, & h_{n,m,j} &= \frac{2\dot{c}(\zeta, s_m^n, \cdot)^2}{\psi_{n,m,j}^3} - \frac{\ddot{c}(\zeta, s_m^n, \cdot)}{\psi_{n,m,j}^2} \end{aligned} \right\}. \tag{5.2}$$

In view of (HS') and (4.2) and of the fact that the derivative of $\psi_{n,m,j}$ is $c'_x(\zeta, s_m^n, \cdot)$, a common feature of all these functions is that they are twice continuously differentiable, with the following estimates, uniform in n , in m, m' between 1 and l_n , in j between 1 and $k_n - 1$, and in $\zeta \in \Theta$ (recall (4.7) and (4.8)):

$$\left. \begin{aligned} |g_{n,m,j}^{(k)}(x)| &\leq C\eta_{n,j} & \text{for } k = 1, 2 \\ |h_{n,m,j}^{(k)}(x)| &\leq C(\eta_{n,j})^{k+1}, & \text{for } k = 0, 1, 2 \end{aligned} \right\}. \tag{5.3}$$

In fact, for further reference we will slightly extends our definition of A_n by letting it be of the form (5.1) with

$$\alpha_m^n = \sum_{j=1}^{k_n-1} Z_{n,m,j} \left(g_{n,m,j}(\widehat{Y}_m^n) - g_{n,m,j}(\widehat{X}_m^n) \right) \tag{5.4}$$

where $Z_{n,m,j}$ is \mathcal{G}_{m-1}^n -measurable, and instead of the first line in (5.3) we have

$$|Z_{n,m,j} g_{n,m,j}^{(k)}(x)| \leq C\eta_{n,j} \quad \text{for } k = 1, 2. \tag{5.5}$$

Lemma 5.1. *Under (5.4) and (5.5) we have $E(A_n^2) \leq C\sigma_n^2$.*

Proof. By Taylor's formula $\alpha_m^n = \alpha_m'^n + \alpha_m''^n$, where

$$\alpha_m'^n = \sum_{j=1}^{k_n-1} Z_{n,m,j} (\hat{Y}_m^n - \hat{X}_m^n) g'_{n,m,j}(\hat{X}_{m-1}^n),$$

$$|\alpha_m''^n| \leq \sum_{j=1}^{k_n-1} |Z_{n,m,j}| (|\hat{Y}_m^n - \hat{X}_{m-1}^n|^2 + |\hat{X}_m^n - \hat{X}_{m-1}^n|^2) \sup_x |g''_{n,m,j}(x)|.$$

If $A_n' = \sum_{m=2}^{l_n} \alpha_m'^n$ and $A_n'' = \sum_{m=2}^{l_n} \alpha_m''^n$ it is clearly enough to prove our estimate separately for A_n' and A_n'' .

In view of (4.18) we have $E(\alpha_m'^n | \mathcal{G}_{m-1}^n) = 0$, hence $E(A_n'^2) = \sum_{m=2}^{l_n} E((\alpha_m'^n)^2)$. But (5.5) and (4.9) yield

$$|\alpha_m'^n| \leq C |\hat{Y}_m^n - \hat{X}_m^n| \sum_{j=1}^{k_n-1} \eta_{n,j} \leq C |\hat{Y}_m^n - \hat{X}_m^n| \sigma_n.$$

Then (4.20) and (4.1) yield $E((A_n')^2) \leq C\sigma_n^2$.

By (5.5) and (4.9) again we have

$$|\alpha_m''^n| \leq C (|\hat{Y}_m^n - \hat{X}_m^n|^2 + |\hat{X}_m^n - \hat{X}_{m-1}^n|^2) \sigma_n.$$

Then (4.17) and (4.20) give

$$E(A_n''^2) \leq l_n \sum_{m=2}^{l_n} E((\alpha_m''^n)^2) \leq C \frac{l_n^2 k_n^2}{n^2} \sigma_n^2 \leq C\sigma_n^2. \quad \square$$

Lemma 5.2. *We have $E(B_n^2) \leq C\sigma_n^2$.*

Proof. If $\beta_m^n = E(\beta_m^n | \mathcal{G}_m^n)$ and since β_m^n is \mathcal{G}_{m+1}^n -measurable, we have

$$E(B_n^2) \leq 2 \sum_{m=2}^{l_n} E((\beta_m^n - \beta_m'^n)^2) + 2E \left(\left(\sum_{m=2}^{l_n} \beta_m'^n \right)^2 \right). \tag{5.6}$$

By (4.21) we have

$$\begin{aligned} E((\beta_m^n - \beta_m'^n)^2 | \mathcal{G}_m^n) &= E \left(\left(\sum_{j=1}^{k_n-1} (n(G_j^{n,m})^2 - n\rho_n \Psi_j^{n,m}(\theta)) (h_{n,m,j}(\hat{Y}_m^n) - h_{n,m,j}(\hat{X}_m^n)) \right)^2 | \mathcal{G}_m^n \right) \\ &= 2 \sum_{j=1}^{k_n-1} (n\rho_n \Psi_j^{n,m}(\theta))^2 (h_{n,m,j}(\hat{Y}_m^n) - h_{n,m,j}(\hat{X}_m^n))^2. \end{aligned}$$

Then (5.3) and (4.9) and (4.27) give $E((\beta_m^n - \beta_m'^n)^2 | \mathcal{G}_m^n) \leq C |\widehat{Y}_m^n - \widehat{X}_m^n|^2 \sigma_n$, and we readily deduce from (4.20) and (4.1) that

$$\sum_{m=2}^{l_n} E((\beta_m^n - \beta_m'^n)^2) \leq Cl_n \left(\frac{k_n}{n} \right) \sigma_n \leq C \sigma_n^2. \quad (5.7)$$

Next we have $\beta_m'^n = u_m^n + v_m^n$, where

$$u_m^n = \sum_{j=1}^{k_n-1} n \rho_n \Psi_j^{n,m-1}(\theta) (h_{n,m,j}(\widehat{Y}_m^n) - h_{n,m,j}(\widehat{X}_m^n)),$$

$$v_m^n = \sum_{j=1}^{k_n-1} n \rho_n (\Psi_j^{n,m}(\theta) - \Psi_j^{n,m-1}(\theta)) (h_{n,m,j}(\widehat{Y}_m^n) - h_{n,m,j}(\widehat{X}_m^n)).$$

We can apply Lemma 5.1 with $Z_{n,m,j} = n \rho_n \Psi_j^{n,m-1}(\theta)$ and $g_{n,m,j} = h_{n,m,j}$, hence (5.5) holds by (5.3) and we get

$$E \left(\left(\sum_{m=2}^{l_n} u_m^n \right)^2 \right) \leq C \sigma_n^2. \quad (5.8)$$

On the other hand, $n \rho_n (\Psi_j^{n,m}(\theta) - \Psi_j^{n,m-1}(\theta)) = c(\theta, s_m^n, \widehat{X}_m^n) - c(\theta, s_{m-1}^n, \widehat{X}_{m-1}^n)$. Thus (4.9) and (HS') give

$$|v_m^n| \leq C \left(\frac{k_n}{n} + |\widehat{X}_m^n - \widehat{X}_{m-1}^n| \right) |\widehat{Y}_m^n - \widehat{X}_m^n| \sigma_n,$$

and by (4.17) and (4.20) we get

$$E \left(\left(\sum_{m=2}^{l_n} v_m^n \right)^2 \right) \leq l_n E \left(\sum_{m=2}^{l_n} (v_m^n)^2 \right) \leq C \frac{l_n^2 k_n^2}{n^2} \sigma_n^2 \leq C \sigma_n^2.$$

This and (5.6, 5.7) and (5.8), give the result. \square

Lemma 5.3. *We have $E(D_n^2) \leq C \sigma_n'^2$.*

Proof. If $\delta_m'^n = E(\delta_m^n | \mathcal{G}_m^n)$ and since δ_m^n is \mathcal{G}_{m+1}^n -measurable, we have

$$E(D_n^2) \leq 2 \sum_{m=2}^{l_n} E((\delta_m^n)^2) + 2E \left(\left(\sum_{m=2}^{l_n} \delta_m'^n \right)^2 \right). \quad (5.9)$$

First, equations (4.24) and (5.3) yield $|\delta_m'^n| \leq C \frac{k_n}{n} \sum_{j=1}^{k_n-1} \eta_{n,j} \leq C \frac{k_n}{n} \sigma_n$, hence

$$E \left(\left(\sum_{m=2}^{l_n} \delta_m'^n \right)^2 \right) \leq C \sigma_n^2. \quad (5.10)$$

Secondly,

$$(\delta_m^n)^2 = n^2 \sum_{i,j=1}^{k_n-1} ((F_i^{n,m})^2 - (G_i^{n,m})^2) ((F_j^{n,m})^2 - (G_j^{n,m})^2) h_{n,m,i}(\widehat{Y}_m^n) h_{n,m,j}(\widehat{Y}_m^n),$$

hence if $w_j^{n,m} = \frac{k_n}{n^2} \left(\frac{k_n}{n^2} + \frac{n\rho_n \Psi_j^{n,m}(\theta)}{n} \right)$, we get by (4.25):

$$\begin{aligned} E((\delta_m^n)^2 | \mathcal{G}_m^n) &\leq Cn^2 \sum_{i,j=1}^{k_n-1} |h_{n,m,i}(\widehat{Y}_m^n) h_{n,m,j}(\widehat{Y}_m^n)| \sqrt{w_i^{n,m} w_j^{n,m}} \\ &\leq Cn^2 \left(\sum_{j=1}^{k_n-1} |h_{n,m,j}(\widehat{Y}_m^n)| \sqrt{w_j^{n,m}} \right)^2. \end{aligned}$$

But (4.27) and (5.3) yields $|h_{n,m,j}(\widehat{Y}_m^n)| \sqrt{w_j^{n,m}} \leq C \left(\frac{k_n}{n^2} \eta_{n,j} + \frac{\sqrt{k_n}}{n^{3/2}} \sqrt{\eta_{n,j}} \right)$. Since $k_n \leq n$, we deduce from (4.9) that $E((\delta_m^n)^2 | \mathcal{G}_m^n) \leq C \frac{k_n}{n} \sigma_n'^2$, and

$$\sum_{m=2}^{l_n} E((\delta_m^n)^2 | \mathcal{G}_m^n) \leq C \frac{k_n l_n}{n} \sigma_n'^2 \leq C \sigma_n'^2.$$

This, together with (5.9) and (5.10), gives the result. □

If we gather these three lemmas, we obtain the

Proposition 5.4. *There is a constant C such that for all $\theta \in \Theta$:*

$$E((\Delta_n(\zeta) - \Gamma_n(\zeta))^2) + E((\dot{\Delta}_n(\zeta) - \dot{\Gamma}_n(\zeta))^2) + E((\ddot{\Delta}_n(\zeta) - \ddot{\Gamma}_n(\zeta))^2) \leq C \sigma_n'^2. \tag{5.11}$$

6. ASYMPTOTIC BEHAVIOUR OF Δ_n

Here is a list of lemmas which will provide our final result.

Lemma 6.1. *We have*

$$\sup_n E \left(\sup_{\zeta} u_n^2 |\dot{\Delta}_n(\zeta)| \right) < \infty. \tag{6.1}$$

$$\sup_n E \left(\sup_{\zeta, \zeta': |\zeta - \zeta'| \leq \eta} u_n^2 |\ddot{\Delta}_n(\zeta) - \ddot{\Delta}_n(\zeta')| \right) \xrightarrow{\eta \rightarrow 0} 0. \tag{6.2}$$

Proof. Set

$$V_n = \sum_{m=2}^{l_n} \sum_{j=1}^{k_n-1} (n(F_j^{n,m})^2 \eta_{n,j}^2 + \eta_{m,j}).$$

Due to (HS') and (4.13) and (4.7), and also due to (4.15) and to the uniform continuity of c , \dot{c} and \ddot{c} , we have

$$\sup_{\zeta} |\dot{\Delta}_n(\zeta)| \leq CV_n, \quad \sup_{\zeta, \zeta': |\zeta - \zeta'| \leq \eta} |\ddot{\Delta}_n(\zeta) - \ddot{\Delta}_n(\zeta')| \leq CV_n \nu(\eta),$$

for a suitable continuous increasing function ν having $\nu(0) = 0$. Then it is enough to prove that $E(V_n) \leq C/u_n^2$. But (4.26) and (4.27) and (4.9) yield $E(V_n) \leq Cl_n\sigma_n = Cl_nk_n/(nu_n^2) \leq C/u_n^2$. \square

Lemma 6.2. *For each $\zeta \in \Theta$ we have the following convergence:*

$$u_n^2(\Delta_n(\zeta) - \Delta_n(\theta)) \xrightarrow{P} \Gamma(\theta, \zeta), \quad (6.3)$$

where $\Gamma(\theta, \theta) = 0$ and, for every $\zeta \neq \theta$, $\Gamma(\theta, \zeta) > 0$ a.s.

Proof. 1) First, by virtue of (5.11), the absolute value of the difference between the left side of (6.3) and the variable $V_n = u_n^2(\Gamma_n(\zeta) - \Gamma_n(\theta))$ has an expectation smaller than $C\sigma'_n u_n^2$, and this goes to 0 by (4.1). Hence it suffices to prove that V_n converges in probability to a $\Gamma(\theta, \zeta)$ having the required properties.

2) Next, set $\delta_j^{n,m} = \Psi_j^{n,m}(\theta)/\Psi_j^{n,m}(\zeta)$. We have $V_n = \sum_{m=2}^{l_n} \alpha_m^n$, with

$$\alpha_m^n = u_n^2 \sum_{j=1}^{k_n-1} \left(\frac{(G_j^{n,m})^2}{\rho_n \Psi_j^{n,m}(\theta)} (\delta_j^{n,m} - 1) - \log \delta_j^{n,m} \right).$$

Then if $\alpha_m'^n = E(\alpha_m^n | \mathcal{G}_m^n)$, we have by (4.21):

$$\alpha_m'^n = u_n^2 \sum_{j=1}^{k_n-1} (\delta_j^{n,m} - 1 - \log \delta_j^{n,m}), \quad (6.4)$$

$$E((\alpha_m^n - \alpha_m'^n)^2 | \mathcal{G}_m^n) = 2u_n^4 \sum_{j=1}^{k_n-1} (\delta_j^{n,m} - 1)^2. \quad (6.5)$$

Hypothesis (HS') gives $|\delta_j^{n,m} - 1| \leq C\eta_{n,j}$, hence (6.5) is smaller than $C\sigma_n u_n^4 = Ck_n u_n^2/n$ by (4.9). Thus

$$E \left(\left(\sum_{m=2}^{l_n} (\alpha_m^n - \alpha_m'^n) \right)^2 \right) = \sum_{m=2}^{l_n} \mathbb{E}((\alpha_m^n - \alpha_m'^n)^2) \leq C u_n^2 \rightarrow 0.$$

So it suffices to prove the convergence in probability of $K'_n = \sum_{m=2}^{l_n} \alpha_m'^n$ to $\Gamma(\theta, \zeta)$.

3) Set

$$\begin{aligned} I'(a, b) &= \frac{1}{\pi} ((a-b)I_1(b) - I(a, b)) \\ &= \frac{1}{\pi} \int_0^\pi \left(\frac{\phi(x, a)}{\phi(x, b)} - 1 - \log \frac{\phi(x, a)}{\phi(x, b)} \right) dx = \frac{a-b}{\sqrt{b(4+b)}} - 2 \log \frac{\sqrt{a} + \sqrt{4+a}}{\sqrt{b} + \sqrt{4+b}} \end{aligned} \quad (6.6)$$

(recall (4.3)) and $J'(a, b, k) = \frac{1}{\pi} ((a-b)J_1(b, k) - J(a, b, k))$. Equation (6.4) yields that $\alpha_m'^n = k_n u_n^2 J'(a, b, k_n)$ if $a = c(\theta, s_m^n, \widehat{X}_m^n)/n\rho_n$ and $b = c(\zeta, s_m^n, \widehat{X}_m^n)/n\rho_n$. Thus, setting

$$\alpha_m'^n = k_n u_n^2 I' \left(\frac{c(\theta, s_m^n, \widehat{X}_m^n)}{n\rho_n}, \frac{c(\zeta, s_m^n, \widehat{X}_m^n)}{n\rho_n} \right)$$

and $K''^n = \sum_{m=2}^{l_n} \alpha_m''^n$, we deduce from (4.5), (HS') and (4.1) that

$$|K_n' - K_n''| \leq \sum_{m=2}^{l_n} |\alpha_m'^n - \alpha_m''^n| \leq Cl_n u_n^2 \rightarrow 0. \tag{6.7}$$

Thus it is enough to prove that $K_n'' \xrightarrow{P} \Gamma(\theta, \zeta)$.

4) Suppose first that we are in Case 2: $n\rho_n \rightarrow u \in (0, \infty)$. The function I' being continuous, Riemann approximation of integrals and the fact that $k_n l_n u_n^2 \rightarrow 1$ immediately yield pathwise convergence of K_n'' to

$$\Gamma(\theta, \zeta) = \int_0^1 I' \left(\frac{c(\theta, t, X_t)}{u}, \frac{c(\zeta, t, X_t)}{u} \right) dt.$$

The integrand in the first term of (6.6) being nonnegative, and being equal to 0 iff $a = b$, we obtain $\Gamma(\theta, \theta) = 0$ and $\Gamma(\theta, \zeta) > 0$ if $\zeta \neq \theta$ by (HI): hence the result.

5) Suppose next that we are in Case 1: $n\rho_n \rightarrow 0$. From the last expression in (6.6) we have that $I'(a/x, b/x) \rightarrow I'_0(a, b) = \frac{a-b}{b} - \log \frac{a}{b}$ as $x \rightarrow 0$, uniformly in a, b when a and b range through $[1/C, C]$ for some C . We then deduce from Riemann approximation of integrals and the fact that $k_n l_n u_n^2 \rightarrow 1$ the pathwise convergence of K_n'' to

$$\Gamma(\theta, \zeta) = \int_0^1 I'_0(c(\theta, t, X_t), c(\zeta, t, X_t)) dt.$$

Since $I'_0(a, b)$ equals 0 if $a = b$ and is positive otherwise, we obtain $\Gamma(\theta, \theta) = 0$ and $\Gamma(\theta, \zeta) > 0$ if $\zeta \neq \theta$ by (HI): hence the result.

6) Finally suppose that we are in Case 3: $n\rho_n \rightarrow \infty$. From the last expression in (6.6) we have that $\sqrt{x} I'(a/x, b/x) \rightarrow I'_1(a, b) = \frac{(\sqrt{a}-\sqrt{b})^2}{2\sqrt{b}}$ as $x \rightarrow \infty$, uniformly in a, b when a and b range through $[1/C, C]$ for some C . We then deduce from Riemann approximation of integrals and the fact that $k_n l_n \sim n = u_n^{-2} \sqrt{n\rho_n}$ the pathwise convergence of K_n'' to

$$\Gamma(\theta, \zeta) = \int_0^1 I'_1(c(\theta, t, X_t), c(\zeta, t, X_t)) dt.$$

Since $I'_1(a, b)$ equals 0 if $a = b$ and is positive otherwise, we obtain $\Gamma(\theta, \theta) = 0$ and $\Gamma(\theta, \zeta) > 0$ if $\zeta \neq \theta$ by (HI): hence the result. □

Lemma 6.3. *The sequence $u_n^2 \ddot{\Delta}_n(\theta)$ converges in probability to $2I(\theta)$, as given by (1.5).*

Proof. Using (5.11) we see, exactly as in the previous lemma, that it suffices to prove the convergence of $u_n^2 \ddot{\Gamma}_n(\theta)$. We have $u_n^2 \ddot{\Gamma}_n(\theta) = \sum_{m=2}^{l_n} \alpha_m^n$, where

$$\begin{aligned} \alpha_m^n &= u_n^2 \sum_{j=1}^{k_n-1} \left(\frac{(G_j^{n,m})^2}{n\rho_n^2 \Psi_j^{n,m}(\theta)^2} \left(\frac{2\dot{c}(\theta, s_m^n, \widehat{X}_m^n)^2}{n\rho_n \Psi_j^{n,m}(\theta)} - \ddot{c}(\theta, s_m^n, \widehat{X}_m^n) \right) \right. \\ &\quad \left. - \frac{1}{n\rho_n \Psi_j^{n,m}(\theta)} \left(\frac{\dot{c}(\theta, s_m^n, \widehat{X}_m^n)^2}{n\rho_n \Psi_j^{n,m}(\theta)} - \ddot{c}(\theta, s_m^n, \widehat{X}_m^n) \right) \right). \end{aligned}$$

Then if $\alpha_m'^n = E(\alpha_m^n | \mathcal{G}_m^n)$, we have by (4.21):

$$\alpha_m'^n = u_n^2 \sum_{j=1}^{k_n-1} \frac{\dot{c}(\theta, s_m^n, \widehat{X}_m^n)^2}{(n\rho_n \Psi_j^{n,m}(\theta))^2}, \tag{6.8}$$

$$E((\alpha_m^n - \alpha_m'^n)^2 | \mathcal{G}_m^n) = 2u_n^4 \sum_{j=1}^{k_n-1} \left(\frac{2\dot{c}(\theta, s_m^n, \widehat{X}_m^n)^2}{(n\rho_n \Psi_j^{n,m}(\theta))^2} - \frac{\ddot{c}(\theta, s_m^n, \widehat{X}_m^n)}{n\rho_n \Psi_j^{n,m}(\theta)} \right)^2. \tag{6.9}$$

Exactly as in the previous lemma, equation (6.9) is smaller than $C\sigma_n u_n^4$, and we are left to prove the convergence of $K_n' = \sum_{m=2}^{l_n} \alpha_m'^n$ to $2I(\theta)$.

But (6.8) yields $\alpha_m'^n = \frac{k_n u_n^2 \dot{c}(\theta, s_m^n, \widehat{X}_m^n)^2}{\pi n^2 \rho_n^2} J_2 \left(\frac{c(\theta, s_m^n, \widehat{X}_m^n)}{n\rho_n}, k_n \right)$. Then if

$$\alpha_m''^n = \frac{k_n u_n^2 \dot{c}(\theta, s_m^n, \widehat{X}_m^n)^2}{\pi n^2 \rho_n^2} I_2 \left(\frac{c(\theta, s_m^n, \widehat{X}_m^n)}{n\rho_n} \right)$$

and $K_n'' = \sum_{m=2}^{l_n} \alpha_m''^n$, we deduce (6.7) from (4.5) and (2.1). Thus it is enough to prove that $K_n'' \xrightarrow{P} 2I(\theta)$. But for this we use (4.3) and Riemann approximation for integrals. \square

Lemma 6.4. *The sequence $(u_n \dot{\Delta}_n(\theta))$ converges stably in law to a variable which, conditionally on (X, W, W') , is centered Gaussian with variance $4I(\theta)$.*

Proof. 1) As a consequence of (4.1) and (5.11), we have $E(u_n |\dot{\Delta}(\theta) - \dot{\Gamma}(\theta)|) \leq C u_n \sigma_n' \rightarrow 0$, so it suffices to prove the stable convergence of $u_n \dot{\Gamma}_n(\theta)$.

We have $u_n \dot{\Gamma}_n(\theta) = \sum_{m=2}^{l_n} z_m^n$, with

$$z_m^n = u_n \dot{c}(\theta, s_m^n, \widehat{X}_m^n) \sum_{j=1}^{k_n-1} \left(\frac{1}{n\rho_n \Psi_j^{n,m}(\theta)} - \frac{(G_j^{n,m})^2}{n\rho_n^2 \Psi_j^{n,m}(\theta)^2} \right). \tag{6.10}$$

Clearly $E(z_m^n | \mathcal{G}_m^n) = 0$, so by applying Theorem 3-2 of [5] our result will follow from the following four convergences:

$$\sum_{m=2}^{[l_n t]} E((z_m^n)^2 | \mathcal{G}_m^n) \xrightarrow{P} 4 \int_0^t \iota_u(c(\theta, s, X_s), \dot{c}(\theta, s, X_s)) ds \quad \forall t \in [0, 1], \tag{6.11}$$

$$\sum_{m=2}^{l_n} E((z_m^n)^4 | \mathcal{G}_m^n) \xrightarrow{P} 0, \tag{6.12}$$

$$\sum_{m=2}^{[l_n t]} E(z_m^n (W_{k_n(m-1)/n} - W_{k_n(m-2)/n}) | \mathcal{G}_m^n) \xrightarrow{P} 0 \quad \forall t \in [0, 1], \tag{6.13}$$

$$\sum_{m=2}^{[l_n t]} E(z_m^n (W'_{k_n(m-1)/n} - W'_{k_n(m-2)/n}) | \mathcal{G}_m^n) \xrightarrow{P} 0 \quad \forall t \in [0, 1]. \tag{6.14}$$

Observe that by (4.11) the variables in (6.13) and (6.14) whose conditional expectation are taken are sums of terms, all of them containing an odd power of increments of either W or W' , so the left sides in these are identically 0. Next by (4.21) yields $E((z_m^n)^2 | \mathcal{G}_m^n) = 2\alpha_m^n$, with the notation (6.8): therefore (6.11) is proved exactly as in the previous lemma.

Finally, the j^{th} summand in (6.10) is $V_j = \gamma_j V'_j$, where $\gamma_j = 1/n\rho_n\Psi_j^{n,m}(\theta)$ and $V'_j = 1 - n\gamma_j(G_j^{n,m})^2$. Conditionally on \mathcal{G}_m^n the γ_j are deterministic, and the V'_j are centered and i.i.d.; hence taking the conditional expectation of the fourth power leaves us with only the expectations of the following, where we put $\delta = \sum_{j=1}^{k_n-1} \gamma_j^2$:

$$\sum_{j,j'=1}^{k_n-1} V_j^2 V_{j'}^2 = \left(\sum_{j=1}^{k_n-1} V_j^2 \right)^2 = \delta^2 \left(\frac{1}{\delta} \sum_{j=1}^{k_n-1} \gamma_j^2 V_j'^2 \right)^2 \leq \delta \sum_{j=1}^{k_n-1} \gamma_j^2 V_j'^4,$$

whose conditional expectation is $C\delta^2$. Now by (4.7, 4.8) and (4.9) we have $\delta \leq C\sigma_n$. Hence $E((z_m^n)^4 | \mathcal{G}_m^n) \leq C u_n^4 \sigma_n^2 = C k_n^2/n^2$, so the left side of (6.12) is smaller than $C l_n k_n^2/n^2 \leq C k_n/n \rightarrow 0$. \square

7. PROOF OF THEOREM 1.1 AND REMARK 4

1) By Proposition 3.1 it is enough to prove the result under (HS') and (HI), so that all lemmas of the previous section hold true.

1) Recalling that $\hat{\theta}_n$ is the (or one of the) absolute minimum (absolute minima) of $\Delta_n(\cdot)$, or equivalently of $u_n^2(\Delta_n(\cdot) - \Delta_n(\theta))$, combining Lemma 6.2 with the first statement of Lemma 6.1 classically yields that $\hat{\theta}_n$ converges in probability to the unique minimum of the function $\Gamma(\theta, \cdot)$, which is θ .

2) Assuming further that θ is in the interior of Θ , and in view of what precedes, we get that

$$P(\dot{\Delta}_n(\hat{\theta}_n) = 0) \rightarrow 1. \tag{7.1}$$

Further, on the set $\Omega_n = \{\dot{\Delta}_n(\hat{\theta}_n) = 0\}$ we have

$$\frac{1}{u_n}(\hat{\theta}_n - \theta) = - \frac{u_n \dot{\Delta}_n(\theta)}{u_n^2 \ddot{\Delta}_n(\theta) + \int_0^1 u_n^2 \left(\ddot{\Delta}_n(\theta + w(\hat{\theta}_n - \theta)) - \ddot{\Delta}_n(\theta) \right) dw}. \tag{7.2}$$

By Lemmas 6.1 and 6.3 and Step 1 above, the denominator in (7.2) converges in probability to $2I(\theta)$. By Lemma 6.4 the numerator converges stably in law to a variable which conditionally on (X, W, W') is centered Gaussian with variance $4I(\theta)$. Combining this with (7.1) and with well known properties of stable convergence, the desired result is straightforward.

2) Now let us examine more closely Remark 4. We have $X = \sqrt{\theta}W$, and the estimators $\hat{\theta}_n$ above are asymptotically optimal (we are in the Gaussian case, so the LAN property holds).

Now, when there is no noise, the optimal estimators (in all possible sense) for θ , at stage n , are

$$\tilde{\theta}_n = \sum_{i=1}^n (X_{i/n} - X_{(i-1)/n})^2 \tag{7.3}$$

and we know that $\sqrt{n}(\tilde{\theta}_n - \theta)$ converges to a variable $\mathcal{N}(0, 2\theta^2)$ when the true value of the parameter is θ .

If we are in the noisy case, we can try the same estimator, but with the noisy observations. It takes the form

$$\begin{aligned}\tilde{\theta}_n &= \sum_{i=1}^n (X_{i/n} + \sqrt{\rho_n} U_i - X_{(i-1)/n} - \sqrt{\rho_n} U_{i-1})^2 \\ &= \theta \sum_{i=1}^n (\Delta_i^n W)^2 + n\rho_n \sum_{i=1}^n (\Delta_i^n W')^2 + 2\sqrt{n\rho_n\theta} \sum_{i=1}^n \Delta_i^n W \Delta_i^n W',\end{aligned}\tag{7.4}$$

where $\Delta_i^n W$ and $\Delta_i^n W'$ are the increments of W and W' between $(i-1)/n$ and i/n , and under P_θ . Then the following facts are almost obvious:

1. $\tilde{\theta}_n$ converges in probability to θ iff we are in Case 1;
2. $\sqrt{n}(\tilde{\theta}_n - \theta)$ converges in law to an $\mathcal{N}(0, 2\theta^2)$ variable iff $n^{3/2}\rho_n \rightarrow 0$ (hence $\tilde{\theta}_n$ is asymptotically optimal);
3. $\sqrt{n}(\tilde{\theta}_n - \theta)$ converges in law to an $\mathcal{N}(v, 2\theta^2)$ variable iff $n^{3/2}\rho_n \rightarrow v \in (0, \infty)$;
4. in all other cases, the sequence $\sqrt{n}(\tilde{\theta}_n - \theta)$ is not even tight.

REFERENCES

- [1] G. Dohnal, On estimating the diffusion coefficient. *J. Appl. Probab.* **24** (1987) 105-114.
- [2] V. Genon-Catalot and J. Jacod, On the estimation of the diffusion coefficient for multidimensional diffusion processes. *Ann. Inst. H. Poincaré Probab. Statist.* **29** (1993) 119-153.
- [3] A. Gloter and J. Jacod, *Diffusion with measurement error. I. Local Asymptotic Normality* (2000).
- [4] J. Jacod and A. Shiryaev, *Limit Theorems for Stochastic Processes*. Springer-Verlag, Berlin (1987).
- [5] J. Jacod, On continuous conditional Gaussian martingales and stable convergence in law, Séminaire Proba. XXXI. Springer-Verlag, Berlin, *Lecture Notes in Math.* **1655** (1997) 232-246.
- [6] M.B. Malyutov and O. Bayborodin, Fitting diffusion and trend in noise via Mercer expansion, in *Proc. 7th Int. Conf. on Analytical and Stochastic Modeling Techniques*. Hamburg (2000).
- [7] A. Renyi, On stable sequences of events. *Sankhyā Ser. A* **25** (1963) 293-302.