Asymptotic equivalence for nonparametric diffusion and Euler scheme experiments

Valentine Genon-Catalot, Catherine Larédo

Equivalence for nonparametric drift estimation of a diffusion process and its Euler scheme.

V. Genon-Catalot[1], C. Larédo[2]

[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

The main goal of the asymptotic equivalence theory of Le Cam (1986) is to approximate general statistical models by simple ones. We develop here a global asymptotic equivalence result for nonparametric drift estimation of a discretely observed diffusion process and its Euler scheme. The asymptotic equivalences are established by constructing explicit equivalence mappings. The impact of such asymptotic equivalence results is that it justifies the use in many applications of the Euler scheme instead of the diffusion process. We especially investigate the case of diffusions with non constant diffusion coefficient. To obtain asymptotic equivalence, experiments obtained by random change of times are introduced.

October 3, 2012

AMS 2000 subject classification: Primary 62B15, 62G20; secondary 62M99, 60J60.

Keywords and phrases: Diffusion process, discrete observations, Euler scheme, nonparametric experiments, deficiency distance, Le Cam equivalence.

Running title: Equivalence for a diffusion and its Euler scheme

1 Introduction

Global asymptotic equivalence of statistical experiments by means of the Le Cam theory of deficiency (Le Cam 1986, Le Cam and Yang 2000) is an important issue for nonparametric estimation problems. The interest is to obtain asymptotic results for some experiment by means of an equivalent simpler one. Concretely, a solution to a nonparametric problem in a simple experiment automatically yields a corresponding solution in an asymptotically equivalent experiment. For instance, when minimax rates of convergence in a nonparametric estimation problem are obtained in one experiment, the same rates automatically hold in a globally asymptotically equivalent experiment (see *e.g.* Nussbaum 1996, Brown and Low 1996). The theory also allows to prove asymptotic sufficiency of the restriction of an experiment to a smaller σ-field.

In most cases, authors are interested in the asymptotic equivalence of density estimation or nonparametric regression and Gaussian white noise (see *e.g.* Nussbaum 1996, Brown and Low 1996, Grama and Nussbaum 1998, 2002, Brown *et al.* 2004, Reiss 2008). Our concern is here the case of experiments associated with diffusion processes. Diffusion processes defined by stochastic differential equations are widely used for modeling purposes in many fields of applications (stochastic models in finance, pharmacokinetic/pharmacodynamic models in biological sciences, ...). As density estimation and nonparametric regression, diffusion models are complex and looking for simpler equivalent experiments is worthwhile. In Larédo (1990), an asymptotic sufficiency property of some incomplete observation is proved; Genon-Catalot *et al.* (2002) studied the equivalence of a diffusion having positive drift and small constant diffusion coefficient with a white noise model and other related experiments; Delattre and Hoffmann (2002) studied the equivalence of diffusions with compactly supported drift and constant diffusion coefficient (null recurrent model) with a mixed Gaussian white noise. Dalalyan and Reiss (2006, 2007) studied the equivalence of one-dimensional and multidimensional diffusions with ergodic properties and constant diffusion coefficient with Gaussian white noise.

Discrete time approximations to a continuous time stochastic process are ubiquitous in the applied mathematical sciences and engineering . The Euler scheme is a classical discrete time approximation to diffusion processes and possesses the advantage of being an autoregressive Markov model with Gaussian transitions. In the parametric framework, the Euler scheme likelihood is classically used as a contrast process and yields optimal estimators (see *e.g.* Genon-Catalot 1990 for the small variance asymptotics, Kessler 1997 for ergodic diffusions, Gobet 2002 for the LAN property implying the optimality of parametric estimators based on the Euler scheme approximation). In the nonparametric framework, Milstein and Nussbaum (1998) proved the asymptotic equivalence of a diffusion process continuously observed on a fixed time interval $[0, T]$ having unknown drift function and constant small known diffusion coefficient with the corresponding Euler scheme. They also obtained the asymptotic sufficiency of the discretized observation of the diffusion with small sampling interval.

In this paper, we consider the experiments associated with a scalar diffusion (ξ_t) with unknown drift function $b(.)$ and known non constant diffusion coefficient $\sigma(.)$ continuously or discretely observed on a time interval $[0, T]$ whose length T tends to infinity. Our aim is to prove that these nonparametric experiments are equivalent to the corresponding Euler scheme experiment, which is an autoregression model thus simpler. A noteworthy consequence of this equivalence result is that it justifies the use of the Euler scheme statistical model in nonparametric problems too. In Dalalyan and Reiss (2006), this property is proved for a diffusion coefficient equal to 1. The extension to the case of a non constant diffusion coefficient is not straightforward and actually surprisingly difficult. The asymptotic equivalences are established by constructing explicit equivalence mappings through the introduction of experiments obtained by random time changes. In addition to the equivalence results, the equivalence mappings provide recipes to find the correspondence betweeen optimal procedures.

Let us now precise the framework of the paper. We consider the one dimensional diffusion process (ξ_t) given by

$$d\xi_t = b(\xi_t)dt + \sigma(\xi_t)dW_t, \quad \xi_0 = \eta, \tag{1}$$

where $(W_t)_{t\geq 0}$ is a Brownian motion defined on a probability space $(\Omega, \mathcal{A}, (\mathcal{A}_t)_{t\geq 0}, \mathbb{P})$, η is a real valued random variable, $\mathcal{A}_0$-measurable, $b(.), \sigma(.)$ are real-valued functions defined on $\mathbb{R}$. The diffusion coefficient $\sigma(.)$ is known and satisfies the following condition:

(C) $\sigma(.) \in C^2(\mathbb{R})$ and there exist positive constants $\sigma_0,\ \sigma_1,\ K_\sigma$ such that

$$\forall x \in \mathbb{R} \quad \sigma_0^2 \leq \sigma^2(x) \leq \sigma_1^2, \ \ |\sigma'(x)| + |\sigma''(x)| \leq K_\sigma.$$

We present three nonparametric experiments for estimating the drift function $b(.)$ in equation (1). The first experiment $\mathcal{E}_0^T$ is associated with the continuous observation of (ξ_t) up to time T. The second experiment $\mathcal{E}^{h,n}$ is associated with the discrete observations of (ξ_t) with sampling interval h up to time $T = nh$, *i.e.* observation of $(\xi_{ih}), 0 \leq i \leq n$. The third experiment $\mathcal{G}^{h,n}$ is associated with the Euler scheme of (1) with sampling interval h up to time $T = nh$, *i.e.* observation of $(Z_i), 0 \leq i \leq n$ (see (9)) for the precise definition of (Z_i)). Our main result (Theorem 6.1) states that, under Condition (C), for estimating $b(.)$ in a class $\mathcal{F}_K$ detailed later on (see (H1)), the three experiments $\mathcal{E}_0^T$, $\mathcal{E}^{h,n}$ and $\mathcal{G}^{h,n}$ are asymptotically equivalent as $n \to \infty$ for the Le Cam deficiency distance Δ if, simultaneously, $h = h_n \to 0$, $T = nh_n \to \infty$ or is bounded, and $nh_n^2 \to 0$. The equivalence of $\mathcal{E}_0^T$ and $\mathcal{E}^{h,n}$ shows that the discretization is an asymptotically sufficient statistic for $\mathcal{E}_0^T$, a result which can also be deduced from Dalalyan and Reiss (2006). The main difficulties occur for getting the equivalence of $\mathcal{E}_0^T$ and the Euler scheme $\mathcal{G}^{h,n}$.

The paper is organized as follows. Assumptions and notations are given in Section 2. We build, in Section 3, a continuous experiment, the continuous Euler scheme, which is equivalent to the experiment associated with the Euler scheme $\mathcal{G}^{h,n}$ in the sense of the Le Cam deficiency distance Δ (Lemma 3.3). Then, we prove in Section 4 that, as $n \to \infty$, the asymptotic equivalence of the discretized diffusion experiment $\mathcal{E}^{h,n}$ and the continuous diffusion experiment $\mathcal{E}_0^T$, if $h = h_n \to 0$ and $nh_n^2 \to 0$. We consider, in Section 5, the two random time changes on the diffusion and on the continuous Euler scheme leading to processes with constant diffusion coefficient. For non constant diffusion coefficient, these two random times are distinct. We prove the exact Δ-equivalence of $\mathcal{E}_0^T$ (resp. the continuous Euler scheme) with the random time changed experiment. Finally, Section 6 contains the proof of Theorem 6.1. In Section 8, the definition and some properties of the Le Cam deficiency distance Δ between statistical experiments are recalled (8.1), and finally some useful auxiliary results are gathered in 8.2.

2 Assumptions and notations

The function $b(.)$ is unknown and varies in the class $\mathcal{F}_K$ of functions satisfying, for K a given positive constant:

(H1) $b(.) \in C^1(\mathbb{R})$ and for all $x \in \mathbb{R}$, $|b(x)| + |b'(x)| \leq K$.

Condition (C) and Assumption (H1) ensure that the stochastic differential equation (1) has a unique strong solution process $(\xi_t)_{t \geq 0}$. Moreover, the function

$$F(x) = \int_0^x \frac{1}{\sigma(u)}\, du \tag{2}$$

is well defined and one-to-one and the two functions

$$f(x) = \frac{b(x)}{\sigma^2(x)} \qquad \text{and} \qquad \mu(x) = \frac{b(F^{-1}(x))}{\sigma(F^{-1}(x))} - \frac{1}{2}\sigma'(F^{-1}(x)) \tag{3}$$

are Lipschitz and bounded:

$$\forall x, y \in \mathbb{R}, \quad |f(x) - f(y)| \leq L|x - y|, \quad |f(x)| \leq \frac{K}{\sigma_0^2}, \tag{4}$$

$$\forall x, y \in \mathbb{R}, \ |\mu(x) - \mu(y)| \leq M|x - y|, \quad |\mu(x)| \leq C, \tag{5}$$

with $L = \frac{K}{\sigma_0^2}(1 + 2\frac{K_\sigma \sigma_1}{\sigma_0^2})$, $M = \frac{K\sigma_1}{\sigma_0}(1 + \frac{K_\sigma}{\sigma_0} + +\frac{1}{2}\sigma_0)$, $C = \frac{K}{\sigma_0} + \frac{1}{2}K_\sigma$.

We introduce below the transformed process $(\eta_t = F(\xi_t))$ which satisfies:

$$d\eta_t = \mu(\eta_t)dt + dW_t, \quad \eta_0 = F(\eta). \tag{6}$$

Let us construct our experiments. Let $C(\mathbb{R}^+, \mathbb{R})$ be the space of continuous real functions defined on $\mathbb{R}^+$, and denote by $(X_t \ , t \geq 0)$ the canonical process of $C(\mathbb{R}^+, \mathbb{R})$ given by $(X_t(x) = x(t), t \geq 0)$ for $x \in C(\mathbb{R}^+, \mathbb{R})$, $\mathcal{C}_t^0 = \sigma(X_s, s \leq t)$, $\mathcal{C}_t = \cap_{s>t}\mathcal{C}_s^0$ and $\mathcal{C} = \sigma(\mathcal{C}_t, t \geq 0)$. Let us denote by P_b the distribution of $(\xi_t, t \geq 0)$ defined by (1) on $(C(\mathbb{R}^+, \mathbb{R}), \mathcal{C})$.

Now, if T is a $(\mathcal{C}_t)$-stopping time, we can define the restriction $P_b/_{\mathcal{C}_T}$ of P_b to the σ-field $\mathcal{C}_T$. The experiment associated with the continuous observation of (ξ_t) stopped at T is given by:

$$\mathcal{E}_0^T = \left(\mathcal{C}(\mathbb{R}^+, \mathbb{R}), \mathcal{C}_T, (P_b/_{\mathcal{C}_T}, b \in \mathcal{F}_K)\right). \tag{7}$$

We are interested in discrete observations of the diffusion with sampling interval $h > 0$. Let us denote by $(\pi_i)_{i\geq 0}$ the canonical projections of $\mathbb{R}^{\mathbb{N}} \to \mathbb{R}$ given by $(\pi_i(x) = x_i, i \geq 0)$ for $x \in \mathbb{R}^{\mathbb{N}}$ and set $\mathcal{G}_n = \sigma(\pi_0, \pi_1, \ldots, \pi_n)$ and $\mathcal{G} = \sigma(\mathcal{G}_n, n \geq 0)$. Let P_b^h denote the distribution of $(\xi_{ih})_{i\geq 0}$ defined by equation (1) on $(\mathbb{R}^{\mathbb{N}}, \mathcal{B}(\mathbb{R}^{\mathbb{N}}))$. If N is a $(\mathcal{G}_n)$- stopping time, we consider the restriction $P_b^h/_{\mathcal{G}_N}$ of P_b^h to $\mathcal{G}_N$, The experiment associated with the discrete observations (ξ_{ih}), with sampling interval h and stopping at N, is given by:

$$\mathcal{E}^{h,N} = (\mathbb{R}^{\mathbb{N}}, \mathcal{G}_N, (P_b^h/_{\mathcal{G}_N}, b \in \mathcal{F}_K). \tag{8}$$

Then, let us consider the Euler scheme corresponding to (1), with sampling interval h. Let us set $t_i = ih$ and define for $i \geq 1$,

$$Z_0 = \eta, \quad Z_i = Z_{i-1} + hb(Z_{i-1}) + \sigma(Z_{i-1})(W_{t_i} - W_{t_{i-1}}) \tag{9}$$

We denote by Q_b^h the distribution of $(Z_i, i \geq 0)$ defined by (9) on $(\mathbb{R}^{\mathbb{N}}, \mathcal{B}(\mathbb{R}^{\mathbb{N}}))$. For N a $(\mathcal{G}_n)$-stopping time, we consider the restriction $Q_b^h/_{\mathcal{G}_N}$ of Q_b^h to $\mathcal{G}_N$. The experiment associated with the discrete Euler scheme (Z_i) with sampling interval h and stopping at N is:

$$\mathcal{G}^{h,N} = (\mathbb{R}^{\mathbb{N}}, \mathcal{G}_N, (Q_b^h/_{\mathcal{G}_N}, b \in \mathcal{F}_K)). \tag{10}$$

We use in the sequel the Le Cam theory (Le Cam, 1986) for comparing statistical experiments. A short recap of this theory is given in Section 8.1.

3 Equivalence of the discrete and the continuous Euler scheme experiments

Given a path $x(.) \in \mathcal{C}(\mathbb{R}^+, \mathbb{R})$ and a sampling scheme $t_i = ih, i \geq 1$, we can define the diffusion-type process $\overline{\xi}_t$,

$$d\overline{\xi}_t = \overline{b}_h(t, \overline{\xi}_.)dt + \overline{\sigma}_h(t, \overline{\xi}_.)dW_t, \quad \overline{\xi}_0 = \eta. \tag{11}$$

where

$$\overline{b}_h(t, x_.) = \sum_{i\geq 1} b(x(t_{i-1})1_{(t_{i-1},t_i]}(t), \qquad \overline{\sigma}_h(t, x_.) = \sum_{i\geq 1} \sigma(x(t_{i-1})1_{(t_{i-1},t_i]}(t). \tag{12}$$

Let us denote by Q_b the distribution of $(\overline{\xi}_t, t \geq 0)$ defined by (11) on $(C(\mathbb{R}^+, \mathbb{R}), \mathcal{C})$.
Now, if T is a $(\mathcal{C}_t)$-stopping time, we can define the restriction $Q_b/_{\mathcal{C}_T}$ of Q_b to the σ-field $\mathcal{C}_T$. We define:

$$\mathcal{G}_0^T = \left(\mathcal{C}(\mathbb{R}^+, \mathbb{R}), \mathcal{C}_T, (Q_b/_{\mathcal{C}_T}, b \in \mathcal{F}_K)\right). \tag{13}$$

Lemma 3.1. *If $\overline{\xi}(.)$ is solution of (11) on Ω, then, using notation (9), $(\overline{\xi}_{t_i}, i \geq 0) = (Z_i, i \geq 0)$.*

Proof. We have $\overline{\xi}_0 = Z_0 = \eta$. Then

$$\begin{aligned} \overline{\xi}_{t_1} &= (t_1 - 0)b(\overline{\xi}_0) + \sigma(\overline{\xi}_0)(W_{t_1} - W_0) \\ &= b(\eta)(t_1 - 0) + \sigma(\eta)(W_{t_1} - W_0) = Z_1 \end{aligned} \tag{14}$$

By induction, assume that $\overline{\xi}_{t_j} = Z_j$ for $j = 0, 1, \ldots, i$. Then

$$\overline{\xi}_{t_{i+1}} = \overline{\xi}_{t_i} + b(\overline{\xi}_{t_i})(t_{i+1} - t_i) + \sigma(\overline{\xi}_{t_i})(W_{t_{i+1}} - W_{t_i}) = Z_{i+1}$$

Thus, $(\overline{\xi}_{t_i}, i \geq 0) = (Z_i, i \geq 0)$. □

Let us define the linear interpolation of $(Z_i, i \geq 0)$:

$$y(t) = Z_i + \frac{t - t_i}{t_{i+1} - t_i}(Z_{i+1} - Z_i) \quad \text{if } t \in [t_i, t_{i+1}] \quad \text{and } i \geq 0. \tag{15}$$

Lemma 3.2. *The solution $(\overline{\xi}_t, t \geq 0)$ of (11) is equal to:*

$$\overline{\xi}_t = y(t) + \sigma(Z_i)B_i(t), \qquad \text{if } t \in [t_i, t_{i+1}] \quad \text{and } i \geq 0, \tag{16}$$

where $B_i(t) = W_t - W_{t_i} - \frac{t-t_i}{t_{i+1}-t_i}(W_{t_{i+1}} - W_{t_i})$. The process $(\overline{\xi}_t)$ is adapted to $(\mathcal{A}_t)$. The processes $((B_i(t), t \in [t_i, t_{i+1}]), i \geq 0)$ are independent Brownian bridges and the sequence $((B_i(t), t \in [t_i, t_{i+1}]), i \geq 0)$ is independent of the sequence $(Z_j, j \geq 0)$.

Proof. Let $t \in [t_i, t_{i+1}]$,

$$\overline{\xi}_t = Z_i + b(Z_i)(t - t_i) + \sigma(Z_i)(W_t - W_{t_i}).$$

Thus, using (15),

$$\overline{\xi}_t = y(t) + \sigma(Z_i)B_i(t),$$

where $(B_i(t))$ is the Brownian bridge defined in Lemma 3.2 for $t_i \leq t \leq t_{i+1}$. It is such that $(B_i(t_i+u), u \in [0, h])$ has the distribution of $(W_u - \frac{u}{h}W_h, 0 \leq u \leq h)$. Using that, for all $i \geq 0$, $B_i(t)$ is $\mathcal{A}_{t_{i+1}}$- measurable and independent of $\mathcal{A}_{t_i}$ yields that $(B_i, i \geq 0)$ are independent processes. Now, by elementary computations, we get that $(B_i(.), i \geq 0)$ is independent of the vector $(W_{t_{i+1}} - W_{t_i}, i \geq 0)$. Since Z_0 is independent of $(W_t, t \geq 0)$, we first get the independence of Z_0 and $(B_i(.), W_{t_{i+1}} - W_{t_i}, i \geq 0)$. This implies that Z_0, $(B_i(.), i \geq 0)$, $(W_{t_{i+1}} - W_{t_i}, i \geq 0)$ are independent. As $\sigma(Z_i, i \geq 0) \subset \sigma(Z_0, W_{t_{i+1}} - W_{t_i}, i \geq 0)$, we obtain the result. □

Lemma 3.3. *For all $h > 0$ and all $(\mathcal{G}_n)$-stopping time N, the Le Cam deficiency distance Δ between $\mathcal{G}^{h,N}$ and $\mathcal{G}_0^{Nh}$ (see (10) and (13)) is equal to 0, i.e. $\Delta(\mathcal{G}^{h,N}, \mathcal{G}_0^{Nh}) = 0$.*

Proof. Recall that for two experiments $\mathcal{E}, \mathcal{E}'$, $\Delta(\mathcal{E}, \mathcal{E}') = \max(\delta(\mathcal{E}, \mathcal{E}'), \delta(\mathcal{E}', \mathcal{E}))$ where $\delta(\mathcal{E}, \mathcal{E}')$ is the deficiency of $\mathcal{E}$ with respect to $\mathcal{E}'$.
The Euler scheme $(Z_i, i \geq 0)$ is the image of the sample path $(\bar{\xi}_t, t \geq 0)$ by the mapping $x(.) \to (x(t_i), i \geq 0)$. Therefore $\delta(\mathcal{G}_0^{Nh}, \mathcal{G}^{h,N}) = 0$.
Consider now, for each $\omega \in \Omega$, the application

$$\Phi : \quad \mathbb{R}^{\mathbb{N}} \to \mathcal{C}(\mathbb{R}^+, \mathbb{R})$$

$$(x_i, i \geq 0) \to x(.) \quad \text{with}$$

$$\forall i \geq 0,\ t \in [t_{i-1}, t_i], \quad x(t) = x_{i-1} + \frac{t - t_{i-1}}{t_i - t_{i-1}}(x_i - x_{i-1}) + \sigma(x_{i-1}) B_{i-1}(t, \omega).$$

Then, by Lemma 3.2, the experiment $\mathcal{G}_0^{Nh}$ is the image, by the randomization Φ of $\mathcal{G}^{h,N}$. Therefore $\delta(\mathcal{G}^{h,N}, \mathcal{G}_0^{Nh}) = 0$. Hence the result. □

4 Asymptotic sufficiency of the discretized diffusion

In this section, we study the experiments $\mathcal{E}_0^{nh_n}$ (continuous observation $(\xi_t, t \leq nh_n)$) and $\mathcal{E}^{h_n,n}$ (discrete observation $(\xi_{t_i}, i \leq n)$ with $t_i = ih_n$) and prove that their are asymptotically equivalent, hence that the discretization $(\xi_{t_i}, i \leq n)$ is an asymptotically sufficient statistic for $\mathcal{E}_0^{nh_n}$. We use the functions F, μ and the process (η_t) defined in (2), (3) and (6). Let

$$\mathcal{H}_0^T = \left(\mathcal{C}(\mathbb{R}^+, \mathbb{R}), \mathcal{C}_T, (P_b^F/_{\mathcal{C}_T}, b \in \mathcal{F}_K)\right), \tag{17}$$

where P_b^F is the distribution of $(\eta_t, t \geq 0)$. And

$$\mathcal{H}^{h,N} = (\mathbb{R}^{\mathbb{N}}, \mathcal{G}_N, (P_b^{h,F}/_{\mathcal{G}_N}, b \in \mathcal{F}_K)), \tag{18}$$

where $P_b^{h,F}$ is the distribution of the discretization $(\eta_{t_i}, i \geq 0)$. Consider the continuous mapping, that we again denote by F,

$$x = (x(t), t \geq 0) \in C(\mathbb{R}^+, \mathbb{R}) \to F(x) = (F(x(t)), t \geq 0) \in C(\mathbb{R}^+, \mathbb{R}).$$

It is invertible with inverse

$$x = (x(t), t \geq 0) \in C(\mathbb{R}^+, \mathbb{R}) \to F^{-1}(x) = (F^{-1}(x(t)), t \geq 0) \in C(\mathbb{R}^+, \mathbb{R}).$$

Then, $\mathcal{H}_0^T = F\mathcal{E}_0^T$ (resp. $\mathcal{H}^{h,N} = F\mathcal{E}^{h,N}$) is the image of $\mathcal{E}_0^T$ (resp. $\mathcal{E}^{h,N}$) by the invertible mapping F, and $F^{-1}\mathcal{H}_0^T = \mathcal{E}_0^T$ (resp. $F^{-1}\mathcal{H}^{h,N} = \mathcal{E}^{h,N}$) is the image of $\mathcal{E}_0^T$ (resp. $\mathcal{E}^{h,N}$) by the mapping F^{-1}. Thus

$$\Delta(\mathcal{E}_0^T, \mathcal{H}_0^T) = 0, \quad \Delta(\mathcal{E}^{h,N}, \mathcal{H}^{h,N}) = 0. \tag{19}$$

Let us now compare the experiments $\mathcal{H}_0^T$ and $\mathcal{H}^{h,N}$ for $N = n$ and $T = nh$. We need to introduce the continuous Euler scheme associated with the diffusion (6):

$$d\bar{\eta}_t = \mu_h(t, \bar{\eta}_.)dt + dW_t, \quad \bar{\eta}_0 = F(\eta), \tag{20}$$

with (see (3))

$$\mu_h(t, x_.) = \sum_{i\geq 1} \mu(x(t_{i-1}))\ 1_{(t_{i-1},t_i]}(t). \tag{21}$$

We introduce the corresponding experiment stopped at T:

$$\bar{\mathcal{H}}_0^T = \left(\mathcal{C}(\mathbb{R}^+, \mathbb{R}), \mathcal{C}_T, (Q_b^F/_{\mathcal{C}_T}, b \in \mathcal{F}_K)\right),$$

where Q_b^F is the distribution of (20). We also introduce the discrete Euler scheme experiment corresponding to the observation $(\bar{\eta}_{t_i})$ stopped at N:

$$\bar{\mathcal{H}}^{h,N} = (\mathbb{R}^{\mathbb{N}}, \mathcal{G}_N, (Q_b^{h,F}/_{\mathcal{G}_N}, b \in \mathcal{F}_K).$$

The experiments $\mathcal{H}_0^T$ and $\bar{\mathcal{H}}_0^T$ have the same sample space and can be compared using their Δ_0-distance:

$$\Delta(\mathcal{H}_0^T, \bar{\mathcal{H}}_0^T) \leq \Delta_0(\mathcal{H}_0^T, \bar{\mathcal{H}}_0^T) = \sup_{b\in\mathcal{F}_K} ||P_b^F/_{\mathcal{C}_T} - Q_b^F/_{\mathcal{C}_T}||_{TV}$$

Lemma 4.1. *Assume (H1) and condition (C). Using definitions (5) for M and C, we have*

$$||P_b^F/_{\mathcal{C}_T} - Q_b^F/_{\mathcal{C}_T}||_{TV} \leq M h_n \sqrt{n} \left(\frac{2}{3}C^2 h_n + 1\right)^{1/2}.$$

Proof. We first use the Pinsker inequality (see Section 8.2, Proposition 8.4) to get an upper bound for the total variation norm by the Kullback-Leibler divergence for two probability distributions P and Q,

$$||P - Q||_{TV} \leq \sqrt{K(P, Q)/2}. \tag{22}$$

As both processes are diffusion-type processes with diffusion coefficient equal to 1, the two distributions P_b^F and Q_b^F are equivalent on $\mathcal{C}_T$ for T deterministic or T a bounded stopping time w.r.t. the canonical filtration. Applying the Girsanov formula yields, with (X_t) the canonical process of $C(\mathbb{R}^+, \mathbb{R})$),

$$\frac{dP_b^F}{dQ_b^F}/_{\mathcal{C}_T}(X) = \exp\left(\int_0^T (\mu(X_t) - \mu_h(t, X_.))dX_t - \frac{1}{2}\int_0^T (\mu(X_t) - \mu_h(t, X_.))^2 dt\right). \tag{23}$$

Now, (X_t) has drift term μ under P_b^F and,

$$\begin{aligned} K(P_b^F/_{\mathcal{C}_T}, Q_b^F/_{\mathcal{C}_T}) &= E_{P_b^F}\left(\frac{1}{2}\int_0^T (\mu(X_t) - \mu_h(t, X_.))^2 dt + \int_0^T (\mu(X_t) - \mu(t, X_.))dW_t\right) \\ &= \frac{1}{2}E_{P_b^F}\left(\int_0^T (\mu(X_t) - \mu_h(t, X_.))^2 dt\right). \end{aligned}$$

Using (3), (5), (21), this term satisfies,

$$\begin{aligned} K(P_b^F/_{\mathcal{C}_T}, Q_b^F/_{\mathcal{C}_T}) &= \frac{1}{2}\sum_{i=0}^{n-1} E_{P_b^F}\int_{t_i}^{t_{i+1}} (\mu(X_t) - \mu(X_{t_i}))^2 dt \\ &\leq \frac{M^2}{2}\sum_{i=0}^{n-1} E_{P_b^F}\int_{t_i}^{t_{i+1}} (X_t - X_{t_i})^2 dt. \end{aligned}$$

Now, by (6) and (5), $\mu^2(x) \leq C^2$ and, under P_b^F, X_t satisfies that, for $t_i \leq t \leq t_{i+1}$,

$$\begin{aligned}(X_t - X_{t_i})^2 &\leq 2(\int_{t_i}^{t} \mu(X_s)ds)^2 + 2(W_t - W_{t_i})^2 \\ &\leq 2C^2(t-t_i)^2 + 2(W_t - W_{t_i})^2.\end{aligned}$$

Hence,

$$E_{P_b^F} \int_{t_i}^{t_{i+1}} (X_t - X_{t_i})^2 dt \leq K^2 h_n^2 \int_{t_i}^{t_{i+1}} (1 + E_{P_b^F} X_s^2) ds + h_n^2.$$

$\square$

Proposition 4.1. *Assume (H1) and condition (C). If $h = h_n$ and nh_n^2 tend to 0, then*

$$\Delta(\mathcal{E}_0^{nh_n}, \mathcal{E}^{h_n,n}) \to 0.$$

Remark : Recall that $nh_n^2 = Th_n$. Hence the result includes both cases: fixed observation time T or $T = nh_n \to \infty$. $\Diamond$

Proof. The triangle inequality and the fact that the Δ_0-distance between experiments having the same sample space is larger than their Δ-distance yield:

$$\Delta(\mathcal{E}_0^{nh_n}, \mathcal{E}^{h_n,n}) \leq \Delta(\mathcal{E}_0^{nh_n}, \mathcal{H}_0^{nh_n}) + \Delta_0(\mathcal{H}_0^{nh_n}, \bar{\mathcal{H}}_0^{nh_n})$$

$$+\Delta(\bar{\mathcal{H}}_0^{nh_n}, \bar{\mathcal{H}}^{n,h_n}) + \Delta_0(\bar{\mathcal{H}}^{n,h_n}, \mathcal{H}^{h_n,n}) + \Delta(\mathcal{H}^{h_n,n}, \mathcal{E}^{h_n,n}).$$

On the right-hand side, there are nul terms. Indeed, by (19), we have

$$\Delta(\mathcal{E}_0^{nh_n}, \mathcal{H}_0^{nh_n}) = \Delta(\mathcal{E}^{h_n,n}, \mathcal{H}^{h_n,n}) = 0.$$

The continuous and the discrete Euler scheme being equivalent experiments yields:

$$\Delta(\bar{\mathcal{H}}_0^{nh_n}, \bar{\mathcal{H}}^{h_n,n}) = 0 \quad .$$

There remains two terms. As the experiment $\mathcal{H}^{h_n,n}$ (resp. $\bar{\mathcal{H}}^{n,h_n}$)) is a restriction of $\mathcal{H}_0^{nh_n}$ (resp. $\bar{\mathcal{H}}_0^{nh_n}$) to a smaller σ-algebra,

$$\Delta_0(\mathcal{H}^{h_n,n}, \bar{\mathcal{H}}^{nh_n}) \leq \Delta_0(\mathcal{H}_0^{nh_n}, \bar{\mathcal{H}}_0^{nh_n})$$

Using the previous lemma, we have , under the condition $nh_n^2 \to 0$,

$$\Delta_0(\mathcal{H}_0^{nh_n}, \bar{\mathcal{H}}_0^{nh_n}) \leq 2Mh_n\sqrt{n}\left(\frac{2}{3}C^2 h_n + 1\right)^{1/2} \to 0.$$

$\square$

We must now draw some conclusions of these results. First, $\mathcal{E}_0^{nh_n}$ and $\mathcal{E}^{h_n,n}$ are asymptotically equivalent, *i.e.* the discrete observation $(\xi_{t_i}, i \leq n)$ is an asymptotically sufficient statistic for $(\xi_t, t \leq nh_n)$. Second, for a diffusion (η_t) with diffusion coefficient equal to 1 (or constant), the experiment associated with the discrete observation $(\eta_{t_i}, i \leq n)$ is asymptotically equivalent to the experiment corresponding to its continuous Euler scheme $(\bar{\eta}_t, t \leq n\Delta_n)$ or to its discrete Euler scheme $(\bar{\eta}_{t_i}, i \leq n)$. So the result of the paper is achieved for diffusions with constant diffusion coefficients.

However, we have not proved that the experiment $\mathcal{E}_0^{nh_n}$ is asymptotically equivalent to $\mathcal{G}_0^{nh_n}$ as these experiments have non constant and distinct diffusion coefficients, respectively the function $\sigma(.)$ and $\bar{\sigma}_h(t, x_.)$ defined in (12).

5 Random time changed experiments

5.1 Time change on the diffusion

As the function σ is known, we introduce the following functionals defined on $C(\mathbb{R}^+, \mathbb{R})$, for all $t, u \geq 0$:

$$x \in C(\mathbb{R}^+, \mathbb{R}), \;\; \rho_t(x) = \int_0^t \sigma^2(x(s))ds; \quad \tau_u(x) = \inf\{t \geq 0, \rho_t(x) \geq u\}. \tag{24}$$

Note that, by condition (C),

$$\rho_{+\infty}(x) = +\infty, \quad \frac{u}{\sigma_1^2} \leq \tau_u(x) \leq \frac{u}{\sigma_0^2}, \quad \rho_{\tau_u(x)}(x) = u, \quad \tau_{\rho_t(x)}(x) = t. \tag{25}$$

Consider the experiment $\mathcal{E}_0^{\tau_a(X)}$ stopped at time $\tau_a(X)$. In order to obtain a constant diffusion coefficient, we introduce the time changed process associated with ξ solution of (1):

$$\forall u \geq 0, \quad Y_u = \xi_{\tau_u(\xi)}. \tag{26}$$

Let $\tilde{P}_b$ denote the distribution of $(Y_u, u \geq 0)$ on $(C(\mathbb{R}^+, \mathbb{R}), \mathcal{C})$. Then, for $A > 0$ a $(\mathcal{C}_t)$-stopping time, define the experiment

$$\tilde{\mathcal{E}}_0^A = (C(\mathbb{R}^+, \mathbb{R}), \mathcal{C}_A, (\tilde{P}_b/_{\mathcal{C}_A}, b \in \mathcal{F}_K)). \tag{27}$$

Proposition 5.1. *Assume (H1) and (C). Let ξ be the solution of (1) and let Y be defined by (26) using (24). Then,*

$$dY_u = f(Y_u)\, du + dB_u, \;\; Y_0 = \eta,$$

where $f(y) = \frac{b(y)}{\sigma^2(y}$ (see (3)) and (B_u) is a Brownian motion with respect to the filtration $(\mathcal{G}_u = \mathcal{A}_{\tau_u(\xi)}), u \geq 0)$.

Proof. The following proof relies on classical results (see *e.g.* Karatzas and Shreve (2000), 5.5). We have $Y_0 = \xi_0 = \eta$ and

$$Y_u = \xi_{\tau_u(\xi)} = \xi_0 + \int_0^{\tau_u(\xi)} b(\xi_s)ds + \int_0^{\tau_u(\xi)} \sigma(\xi_s)dW_s. \tag{28}$$

The change of variable $s = \tau_v(\xi) \Leftrightarrow v = \rho_s(\xi)$ yields that $dv = \sigma^2(\xi_s)ds = \sigma^2(Y_v)d\tau_v(\xi)$ so that, for $0 \leq v \leq a$,

$$\frac{d\tau_v(\xi)}{dv} = \frac{1}{\sigma^2(Y_v)} \quad \text{and} \quad Y_u = Y_0 + \int_0^u \frac{b(Y_v)}{\sigma^2(Y_v)}dv + B_u,$$

where the process (B_u) is defined by

$$B_u = \int_0^{\tau_u(\xi)} \sigma(\xi_s)dW_s.$$

We have $< B >_u = \int_0^{\tau_u(\xi)} \sigma^2(\xi_s)ds = u$. Hence, (B_u) is a standard Brownian motion with respect to the filtration $(\mathcal{G}_u = \mathcal{A}_{\tau_u(\xi)})$ and (Y_u) defined in (26) is a diffusion process with drift coefficient $\frac{b}{\sigma^2}(.)$, constant $(= 1)$ diffusion coefficient and initial condition $Y_0 = \eta$, which is $\mathcal{G}_0$ measurable. Note that $(\mathcal{G}_u)$ satisfies the usual conditions by the continuity of $\tau_.(\xi)$. □

Proposition 5.2. *Under (H1) and (C), for $a > 0$ deterministic,*

$$\Delta(\mathcal{E}_0^{\tau_a(X)}, \tilde{\mathcal{E}}_0^a) = 0.$$

Analogously, for $T > 0$ deterministic, we have (see (24):

$$\Delta(\mathcal{E}_0^T, \tilde{\mathcal{E}}_0^{\rho_T(X)}) = 0.$$

Proof. We only proof the first point as the second is analogous. By the previous proposition, $\tilde{\mathcal{E}}_0^a$ is the image of $\mathcal{E}_0^{\tau_a(X)}$ by the measurable mapping $(x(t), t \in [0, \tau_a(x)]) \rightarrow (y(u) = x(\tau_u(x)), u \in [0, a])$, which implies $\delta(\mathcal{E}_0^{\tau_a(\xi)}, \tilde{\mathcal{E}}_0^a) = 0$.
Let us consider now the reverse operation. Let $(B_u, u \geq 0)$ be a standard Brownian motion with respect to a filtration $(\mathcal{G}_u)$ satisfying the usual conditions and Y_0 be a $\mathcal{G}_0$-measurable random variable. We define, for $u \geq 0$,

$$Y_u = Y_0 + \int_0^u \frac{b(Y_v)}{\sigma^2(Y_v)}dv + B_u, \tag{29}$$

$$T_u = T_u(Y) = \int_0^u \frac{dv}{\sigma^2(Y_v)}.$$

Clearly, the mapping $u \rightarrow T_u$ is a bijection from $[0, a]$ onto $[0, T_a]$ with inverse $t \rightarrow T^{-1}(t) := A_t$. Therefore, we can define, for $0 \leq t \leq T_a$, the process

$$\xi_t = Y_{A_t}.$$

The change of variable $v = A_s \Leftrightarrow s = T_v$ yields that $ds = dv/\sigma^2(Y_v) = dv/\sigma^2(Y_{A_s}) = dv/\sigma^2(\xi_s)$ and equation (29) becomes

$$\xi_t = \xi_0 + \int_0^{A_t} \frac{b(Y_v)}{\sigma^2(Y_v)} dv + B_{A_t} = \xi_0 + \int_0^t b(\xi_s)ds + B_{A_t}.$$

Now, $(M_t = B_{A_t})$ is a martingale with respect to the filtration $(\mathcal{G}_{A_t})$ satisfying

$$< M >_t = A_t = \int_0^{A_t} ds = \int_0^t \sigma^2(Y_{A_s})ds = \int_0^t \sigma^2(\xi_s)ds.$$

This shows that $\tau_u(\xi) = A^{-1}(u) = T_u$ and that (ξ_t) has distribution P_b. Note that, as (A_t) is continuous, the filtration $(\mathcal{G}_{A_t})$ inherits the usual conditions from $(\mathcal{G}_t)$.

Finally, we can express the above properties on the canonical space. Let $y = (y(v), v \geq 0)$ and set $T_u(y) = \int_0^u dv/\sigma^2(y(v))$ with inverse $A_.(y)$. Consider the measurable mapping

$$\Psi : y \in C(\mathbb{R}^+, \mathbb{R}) \rightarrow (x := y(A_t(y)), t \geq 0) \in C(\mathbb{R}^+, \mathbb{R}).$$

As

$$A_t(y) = \int_0^t \sigma^2(x(s))ds = \rho_t(x),$$

we see that $A_.(y)^{-1}(u) = \tau_u(x)$. Thus, $(\xi(t), t \leq \tau_a(\xi))$ is the image of $(Y(u), u \leq a)$ by the mapping Ψ. Hence, $\delta(\tilde{\mathcal{E}}_0^a, \mathcal{E}_0^{\tau_a(X)}) = 0$. $\square$

5.2 Time change on the continuous Euler scheme

Let us now consider the continuous Euler scheme $(\overline{\xi}_u)$ defined in (11), (12) and define analogously:

$$\overline{\rho}_t(x) = \int_0^t \overline{\sigma}_h^2(s, x_.)ds, \quad \overline{\tau}_u(x) = \inf\{t \geq 0, \overline{\rho}_t(x) \geq u\}. \tag{30}$$

Using (12), $(\overline{\rho}_t(x))$ satisfies, for $i \geq 0$ and $t_i < t \leq t_{i+1}$,

$$\overline{\rho}_t(x) = \overline{\rho}_{t_i}(x) + (t - t_i)\overline{\sigma}_h^2(t_i, x_.) = h \sum_{j=0}^{i-1} \sigma^2(x(t_j)) + (t - t_i)\sigma^2(x(t_i)), \tag{31}$$

(with $\sum_{j=0}^{i-1} = 0$ for $i = 0$). Note that

$$\overline{\rho}_{t_i}(x) = h \sum_{j=0}^{i-1} \sigma^2(x(t_j)). \tag{32}$$

Hence, $(\overline{\rho}_t(x), t \geq 0)$ is continuous, strictly increasing on $\mathbb{R}^+$, and maps each interval $(t_i, t_{i+1}]$ on the interval $(\overline{\rho}_{t_i}(x), \overline{\rho}_{t_{i+1}}(x)]$.
Condition (C) ensures that

$$\overline{\rho}_{+\infty}(x) = +\infty, \quad \frac{u}{\sigma_1^2} \leq \overline{\tau}_u(x) \leq \frac{u}{\sigma_0^2} + \Delta, \tag{33}$$

and that $t \to \overline{\rho}_t(x)$ and $u \to \overline{\tau}_u(x)$ are inverse. In particular, for all i and all trajectory x

$$t_i = \overline{\tau}_{\overline{\rho}_{t_i}(x)}(x). \tag{34}$$

For $\overline{\xi}$ the solution of (11), we define, for $u \geq 0$,

$$\overline{Y}_u = \overline{\xi}_{\overline{\tau}_u(\overline{\xi})}, \tag{35}$$

which is adapted to the filtration

$$(\overline{\mathcal{G}}_u = \mathcal{A}_{\overline{\tau}_u(\overline{\xi})}). \tag{36}$$

Denote by $\widetilde{Q}_b$ the distribution of $(\overline{Y}_u, u \geq 0)$. We now describe the distribution $\widetilde{Q}_b$ with more precision.

Proposition 5.3. *The process $(\overline{Y}_u)$ defined in (35) is a process with constant diffusion coefficient equal to* 1 *and drift term given by:*

$$\overline{f}(v, .) = \sum_{i \geq 0} f(\overline{Y}_{\overline{\rho}_{t_i}(\overline{\xi})}) \, 1_{(\overline{\rho}_{t_i}(\overline{\xi}), \overline{\rho}_{t_{i+1}}(\overline{\xi})]}(v), \tag{37}$$

where $\overline{f}(v, .)$ is predictable w.r.t. the filtration $(\overline{\mathcal{G}}_u)$ and $f = b/\sigma^2$ (see (3)).

Proof. By definition of $(\overline{Y}_u)$, we have

$$\overline{Y}_u = \overline{\xi}_0 + \int_0^{\overline{\tau}_u(\overline{\xi})} \sum_{i \geq 0} b(\overline{\xi}_{t_i}) 1_{t_i < s \leq t_{i+1}} ds + \overline{B}_u \quad \text{where} \tag{38}$$

$$\overline{B}_u = \int_0^{\overline{\tau}_u(\overline{\xi})} \sum_{i \geq 0} \sigma(\overline{\xi}_{t_i}) 1_{t_i < s \leq t_{i+1}} dW_s.$$

The process $(\overline{B}_u)$ is a martingale with respect to $\overline{\mathcal{G}}_u = \mathcal{A}_{\overline{\tau}_u(\overline{\xi})}$ with quadratic variations:

$$< \overline{B} >_u = \int_0^{\overline{\tau}_u(\overline{\xi})} \sum_{i \geq 0} \sigma^2(\overline{\xi}_{t_i}) 1_{t_i < s \leq t_{i+1}} ds = u.$$

Therefore $(\overline{B}_u)$ is a Brownian motion with respect to $(\overline{\mathcal{G}}_u)$.
In the ordinary integral of (38), the change of variable $s = \overline{\tau}_v(\overline{\xi}) \Leftrightarrow v = \overline{\rho}_s(\overline{\xi})$ yields, noting that for $v \in (\overline{\rho}_{t_i}(\overline{\xi}), \overline{\rho}_{t_{i+1}}(\overline{\xi})]$, $dv = \sigma^2(\overline{\xi}_{t_i})ds$ and using (34),

$$\overline{Y}_u = \overline{\xi}_0 + \int_0^u \sum_{i \geq 0} \frac{b(\overline{\xi}_{t_i})}{\sigma^2(\overline{\xi}_{t_i})} 1_{\overline{\rho}_{t_i}(\overline{\xi}) < v \leq \overline{\rho}_{t_{i+1}}(\overline{\xi})} \, dv + \overline{B}_u, \tag{39}$$

where $\overline{\xi}_{t_i} = \overline{Y}_{\overline{\rho}_{t_i}(\overline{\xi})} = Z_i$ is the discrete Euler scheme (see Lemma 3.1).
Thus, $(\overline{Y}_u)$ defined in (35) is a process with constant diffusion coefficient equal to 1 and drift term given by $\overline{f}(v, .)$. We now check that $\overline{f}(v, .)$ is predictable w.r.t. the filtration $(\overline{\mathcal{G}}_u)$ *i.e.* that, for all i, $\overline{\rho}_{t_i}(\overline{\xi})$ is a $(\overline{\mathcal{G}}_u)$-stopping time and that $\overline{Y}_{\overline{\rho}_{t_i}(\overline{\xi})}$ is $\overline{\mathcal{G}}_{\overline{\rho}_{t_i}(\overline{\xi})}$-measurable.
Noting that, for all u,

$$(\overline{\rho}_{t_i}(\overline{\xi}) \leq u) = (\overline{\tau}_u(\overline{\xi}) \geq t_i)$$

belongs to $\overline{\mathcal{G}}_u = \mathcal{A}_{\overline{\tau}_u(\overline{\xi})}$ yields that $\overline{\rho}_{t_i}(\overline{\xi})$ is a $(\overline{\mathcal{G}}_u)$-stopping time. We know that

$$\overline{Y}_{\overline{\rho}_{t_i}(\overline{\xi})} = \overline{\xi}_{t_i}$$

is $\mathcal{A}_{t_i}$-measurable, which achieves the proof since $\mathcal{A}_{t_i} = \overline{\mathcal{G}}_{\overline{\rho}_{t_i}(\overline{\xi})}$.

□

We consider, for $a > 0$, the experiments $\mathcal{G}_0^{\overline{\tau}_a(X)}$ (see (13)) and

$$\widetilde{\mathcal{G}}_0^a = (C(\mathbb{R}^+, \mathbb{R}), \mathcal{C}_a, (\widetilde{Q}_b/_{\mathcal{C}_a}, b \in \mathcal{F})). \tag{40}$$

Proposition 5.4. *For all $a > 0$ deterministic,*

$$\Delta(\mathcal{G}_0^{\overline{\tau}_a(X)}, \widetilde{\mathcal{G}}_0^a) = 0.$$

Analogously, for all $T > 0$ deterministic,

$$\Delta(\mathcal{G}_0^T, \widetilde{\mathcal{G}}_0^{\overline{\rho}_T(X)}) = 0.$$

Proof. We only prove the first point. The proof is divided in several steps.
First, as the experiment $\widetilde{\mathcal{G}}_0^a$ is the image of the experiment $\mathcal{G}_0^{\overline{\tau}_a(X)}$ by the measurable mapping $(x(t), t \leq \overline{\tau}_a(x)) \to (y(u) = x(\overline{\tau}_u(x)), u \in [0, a])$, $\delta(\mathcal{G}_0^{\overline{\tau}_a(X)}, \widetilde{\mathcal{G}}_0^a) = 0$.
We now prove that $\delta(\widetilde{\mathcal{G}}_0^a, \mathcal{G}_0^{\overline{\tau}_a(X)}) = 0$. For this, we first construct a process $(\overline{Y}_u)$ with distribution $\widetilde{Q}_b$ (step 1). Then, we construct a process $(\overline{X}_t)$ with distribution Q_b obtained from $(\overline{Y}_u)$ by the measurable mapping $(y(u), u \in [0, a] \to (y(\overline{\rho}_t(y.)),\ t \leq \overline{\tau}_a(y.))$ (step 2).

Step 1.
Let $(\overline{B}_u)$ be a standard Brownian motion with respect to a filtration $(\overline{\mathcal{G}}_u)$ satisfying the usual conditions and assume that X_0 is $\overline{\mathcal{G}}_0$ - measurable. Then, we define recursively a

sequence of random times (T_i) and a continuous process $(\overline{Y}_u)$ as follows. We first set $T_0 = 0$ and $\overline{Y}_0 = X_0$. Then

$$T_1 = T_1(\overline{Y}) = \sigma^2(\overline{Y}_0)t_1, \qquad \overline{Y}_u = \overline{Y}_0 + \frac{b(\overline{Y}_0)}{\sigma^2(\overline{Y}_0)}\, u + \overline{B}_u, \text{ for } 0 < u \le T_1.$$

$$T_i = T_i(\overline{Y}) = T_{i-1} + \sigma^2(\overline{Y}_{T_{i-1}})(t_i - t_{i-1}), \tag{41}$$

$$\overline{Y}_u = \overline{Y}_{T_{i-1}} + \frac{b(\overline{Y}_{T_{i-1}})}{\sigma^2(\overline{Y}_{T_{i-1}})}(u - T_{i-1}) + \overline{B}_u - \overline{B}_{T_{i-1}}, \text{ for } T_{i-1} < u \le T_i. \tag{42}$$

□

Lemma 5.1. *The sequence (T_i) is an increasing sequence of $(\overline{\mathcal{G}}_u)$-stopping times such that, for all $i \ge 1$, T_i is $\overline{\mathcal{G}}_{T_{i-1}}$ measurable. Moreover, the process $(\overline{Y}_u)$ defined in (41),(42) is adapted to $(\overline{\mathcal{G}}_u)$ and is a diffusion-type process with diffusion coefficient equal to* 1 *and drift coefficient equal to*

$$\overline{f}(u, y_.) = \sum_{i\ge 1} f(y(T_{i-1}(y_.)))1_{T_{i-1}(y_.)<u\le T_i(y_.)},$$

where $(T_i(y_.), i \ge 0)$ are recursively defined as in (41)-(42) using the trajectory $y_.$ and $f = b/\sigma^2$ (see (3)). Hence, the process $(\overline{Y}_u)$ has distribution $\widetilde{Q}_b$.

Proof. First note that T_1 is $\overline{\mathcal{G}}_0$-measurable, and thus $\{T_1 \le u\} \in \overline{\mathcal{G}}_0 \subset \overline{\mathcal{G}}_u$. Hence T_1 is a $(\overline{\mathcal{G}}_u)$ -stopping time. Now, $\overline{Y}_u = \overline{Y}_0 + \frac{b(\overline{Y}_0)}{\sigma^2(\overline{Y}_0)}u + \overline{B}_u$ is, for all u, $\overline{\mathcal{G}}_u$-measurable. Thus, T_1 and $\overline{Y}_{T_1}$ are $\overline{\mathcal{G}}_{T_1}$ measurable.
By induction, assume that, for $1 \le j \le i$, T_j is $\overline{\mathcal{G}}_{T_{j-1}}$-measurable, that T_j is a $(\overline{\mathcal{G}}_u)$-stoppping time, and that $\overline{Y}_u$ given by (42) for $u \le T_i$ is $\overline{\mathcal{G}}_u$-measurable.
Recall that, for $u > T_i$, $\overline{Y}_u = \overline{Y}_{T_i} + \frac{b(\overline{Y}_{T_i})}{\sigma^2(\overline{Y}_{T_i})}(u - T_i) + \overline{B}_u - \overline{B}_{T_i}$ defined by (42) is $\overline{\mathcal{G}}_u$-measurable. As $T_{i+1} = T_i + \sigma^2(\overline{Y}_{T_i})(t_{i+1} - t_i)$, the induction assumption yields that T_{i+1} is $\overline{\mathcal{G}}_{T_i}$-measurable. By Condition (C), $T_i < T_{i+1}$. Hence, using the induction assumption,

$$\forall v \ge u, \quad \{T_{i+1} \le u\} = \{T_{i+1} \le u\} \cap \{T_i \le v\} \in \overline{\mathcal{G}}_v.$$

This implies that $\{T_{i+1} \le u\} = \{T_{i+1} \le u\} \bigcap \cap_{v>u}\{T_i \le v\} \in \cap_{v>u}\overline{\mathcal{G}}_v = \overline{\mathcal{G}}_u$ which proves that T_{i+1} is a $(\overline{\mathcal{G}}_u)$-stopping time. Thus, T_{i+1} and $\overline{Y}_{T_{i+1}}$ are $\overline{\mathcal{G}}_{T_{i+1}}$-measurable.
The proof is now complete. □

Step 2.
Let us define

$$\overline{\sigma}(t, \overline{Y}_.) = \sum_{i\ge 1} \sigma(\overline{Y}_{T_{i-1}})1_{]t_{i-1},t_i]}(t), \tag{43}$$

$$\overline{\rho}_t(\overline{Y}_.) = \int_0^t \overline{\sigma}^2(s, \overline{Y}_.)ds \quad \text{ and } \quad \overline{\tau}_a(\overline{Y}_.) = \inf\{t; \overline{\rho}_t(\overline{Y}_.) \ge a\}. \tag{44}$$

We set

$$\overline{X}_t = \overline{Y}_{\overline{\rho}_t(\overline{Y}_.)} \tag{45}$$

and look at the distribution of $(\overline{X}_t, t \leq \overline{\tau}_a(\overline{Y}_.))$. Note that, for all $i \geq 0$,

$$\overline{\rho}_{t_i}(\overline{Y}_.) = T_i. \tag{46}$$

is a $(\overline{\mathcal{G}}_u)$-stopping time (see Lemma 5.1). We need to check that, for all t, $\overline{\rho}_t(\overline{Y}_.)$ is a $(\overline{\mathcal{G}}_u)$ -stopping time. Noting that, for $t_i \leq t \leq t_{i+1}$, $\overline{\rho}_t(\overline{Y}_.) = T_i + (t - t_i)\sigma^2(\overline{Y}_{T_i})$ is $\overline{\mathcal{G}}_{T_i}$-measurable, we can write

$$\forall v > u, \quad \{\overline{\rho}_t(\overline{Y}_.) \leq u\} = \{\overline{\rho}_t(\overline{Y}_.) \leq u\} \cap \{T_i \leq v\} \in \overline{\mathcal{G}}_v.$$

This implies that $\{\overline{\rho}_t(\overline{Y}_.) \leq u\} = \{\overline{\rho}_t(\overline{Y}_.) \leq u\} \bigcap \cap_{v>u}\{T_i \leq v\} \in \cap_{v>u}\overline{\mathcal{G}}_v = \overline{\mathcal{G}}_u$ which proves that $\overline{\rho}_t(\overline{Y}_.)$ is a $(\overline{\mathcal{G}}_u)$-stopping time.

Thus we can define the filtration $(\overline{\mathcal{A}}_t := \overline{\mathcal{G}}_{\overline{\rho}_t(\overline{Y}_.)})$ to which $(\overline{X}_t)$ is adapted. By construction, $\overline{\tau}_a(\overline{Y}_.)$ is a $(\overline{\mathcal{A}}_t)$-stopping time.

Lemma 5.2. *The sequence $(\overline{X}_{t_i} = \overline{Y}_{T_i}, i \geq 0)$, where $(\overline{Y}_u)$ is defined by (41)-(42) and $(\overline{X}_t)$ is defined in (45), has the distribution of the discrete Euler scheme (9).*

Proof. For all $i \geq 0$, the process

$$(\bar{B}^{(i)}_v = \overline{B}_{T_i+v} - \overline{B}_{T_i}, v \geq 0) \tag{47}$$

is a Brownian motion independent of $\overline{\mathcal{G}}_{T_i} = \mathcal{A}_{t_i}$, adapted to the filtration $(\overline{\mathcal{G}}_{T_i+v})$. As $\overline{X}_{t_i} = \overline{Y}_{T_i}$ is $\overline{\mathcal{G}}_{T_i}$-measurable, this r.v. is independent of $(\bar{B}^{(i)}_v, v \geq 0)$. Let us define

$$\varepsilon_{i+1} = \frac{\overline{B}_{T_{i+1}} - \overline{B}_{T_i}}{\sqrt{T_{i+1} - T_i}} = \frac{\bar{B}^{(i)}_{\sigma^2(\overline{Y}_{T_i})(t_{i+1}-t_i)}}{\sigma(\overline{Y}_{T_i})\sqrt{t_{i+1} - t_i}}. \tag{48}$$

The random variable ε_{i+1} is $\overline{\mathcal{G}}_{T_{i+1}}$-measurable. We can write,

$$\overline{Y}_{T_{i+1}} = \overline{Y}_{T_i} + b(\overline{Y}_{T_i})(t_{i+1} - t_i) + \sigma(\overline{Y}_{T_i})\sqrt{t_{i+1} - t_i}\, \varepsilon_{i+1}, \quad i \geq 0. \tag{49}$$

To conclude, it is enough to prove that $(\varepsilon_i, i \geq 1)$ is a sequence of *i.i.d.* standard Gaussian random variables, independent of $\overline{\mathcal{G}}_0$.
Applying Proposition 8.3 of the Appendix yields that, for all $i \geq 0$, ε_{i+1} is a standard Gaussian variable independent of $\overline{\mathcal{G}}_{T_i}$. This holds for $i = 0$ and proves that ε_1 is independent of $\overline{\mathcal{G}}_0$ and has distribution $\mathcal{N}(0, 1)$. By induction, assume that $(\varepsilon_k, k \leq i - 1)$ are *i.i.d.* standard Gaussian random variables, independent of $\overline{\mathcal{G}}_0$. Consider Z a $\overline{\mathcal{G}}_0$-measurable random variable. As $(Z, \varepsilon_k, k \leq i - 1))$ is $\overline{\mathcal{G}}_{T_i}$-measurable, we get that ε_{i+1} is a standard Gaussian variable independent of $(Z, \varepsilon_k, k \leq i - 1))$. This completes the proof of Lemma 5.2. □

Define now $(\overline{x}(t))$ as the linear interpolation between the points $(t_i, \overline{X}_{t_i})$. We now describe the processes $(\overline{X}_t - \overline{x}(t))$ for $t_i \leq t \leq t_{i+1}$.

Lemma 5.3. *For $t \in [t_i, t_{i+1}]$,*

$$\overline{X}_t = \overline{x}(t) + \sigma(\overline{X}_{t_i})\bar{C}_i(t),$$

where $(\bar{C}_i(t), t_i \leq t \leq t_{i+1})$ is a sequence of independent Brownian bridges adapted to the filtration $(\overline{\mathcal{A}}_t)$. The sequence of processes $((\bar{C}_i(t), t_i \leq t \leq t_{i+1}), i \geq 0)$ is independent of the sequence $(\overline{X}_{t_j}, j \geq 0)$.

Proof.

$$\overline{y}(u) = \overline{Y}_{T_i} + + \frac{u - T_i}{T_{i+1} - T_i}(\overline{Y}_{T_{i+1}} - \overline{Y}_{T_i}).$$

Using (48)-(49), we obtain, for $u \in [T_i, T_{i+1}]$, :

$$\begin{aligned}\overline{Y}_u &= \overline{y}(u) + \overline{B}_u - \overline{B}_{T_i} - \frac{u - T_i}{T_{i+1} - T_i}\sigma(\overline{Y}_{T_i})\sqrt{t_{i+1} - t_i}\,\frac{\overline{B}_{T_{i+1}} - \overline{B}_{T_i}}{\sqrt{T_{i+1} - T_i}} \\ &= \overline{y}(u) + D_i(u) \text{ with}\end{aligned}$$

$$D_i(u) = \overline{B}_u - \overline{B}_{T_i} - \frac{u - T_i}{T_{i+1} - T_i}(\overline{B}_{T_{i+1}} - \overline{B}_{T_i}). \tag{50}$$

Now, for $t_i \leq t \leq t_{i+1}$, using (41),(43),(44),(46), we get $\overline{x}(t) = \overline{y}(\overline{\rho}_t(\overline{Y}_.))$, Thus,

$$\overline{X}_t - \overline{x}(t) = D_i(\overline{\rho}_t(\overline{Y}_.)).$$

Using the Brownian motion defined in (47), we get

$$\overline{X}_t - \overline{x}(t) = \overline{B}^{(i)}_{\sigma^2(\overline{X}_{t_i})(t-t_i)} - \frac{t - t_i}{t_{i+1} - t_i}\overline{B}^{(i)}_{\sigma^2(\overline{X}_{t_i})(t_{i+1}-t_i)} = \sigma(\overline{X}_{t_i})\bar{C}_i(t), \tag{51}$$

which defines $\bar{C}_i(t)$.
Proving that the sequence $(\overline{X}_{t_i}, i \geq 0)$ is independent of the sequence of processes $((\bar{C}_i(t), t \in [t_i, t_{i+1}]), i \geq 0)$ is equivalent to proving that $(\overline{X}_0, \varepsilon_i, i \geq 1)$ is independent of $((\bar{C}_i(t), t \in [t_i, t_{i+1}]), i \geq 0)$. We now prove that, for all $i \geq 1$, $(\overline{X}_0, \varepsilon_1, \ldots, \varepsilon_i)$ is independent of $(\bar{C}_0, \ldots, \bar{C}_{i-1})$ and that the latter processes are independent Brownian bridges. For this, we use Proposition 8.3 of the Appendix. With $B = \overline{B}^{(i-1)}$, $\mathcal{F}_. = \overline{\mathcal{G}}_{T_{i-1}+.}$, $\tau = \sigma^2(\overline{X}_{t_{i-1}})$, $i \geq 1$, we deduce that

$$W_i(t - t_{i-1}) = \frac{\overline{B}^{(i-1)}_{\sigma^2(\overline{X}_{t_{i-1}})(t-t_{i-1})}}{\sigma(\overline{X}_{t_{i-1}})},\ t \geq t_{i-1},$$

is a Brownian motion, independent of $\overline{\mathcal{G}}_{T_{i-1}}$. Thus, $(\bar{C}_{i-1}(t), t \in [t_{i-1}, t_i])$ is a Brownian bridge independent of $W_i(t_i - t_{i-1}) = \varepsilon_i\sqrt{t_i - t_{i-1}}$. Moreover, $\overline{\mathcal{G}}_{T_{i-1}}$, $W_i(t_i - t_{i-1})$, and $(\bar{C}_{i-1}(t), t \in [t_{i-1}, t_i])$ are independent.

For $i = 1$, as $\overline{X}_0$ is $\overline{\mathcal{G}}_0$-measurable, we get that $\overline{X}_0$, ε_1, $\bar{C}_0$ are independent and $\bar{C}_0$ is a Brownian bridge. By induction, let us assume that $\overline{X}_0, \varepsilon_1, \ldots, \varepsilon_i, \bar{C}_0, \ldots, \bar{C}_{i-1}$ are independent and that $\bar{C}_0, \ldots, \bar{C}_{i-1}$ are Brownian bridges (on their respective interval of definition). As $Z = (\overline{X}_0, \varepsilon_1, \ldots, \varepsilon_i, \bar{C}_0, \ldots, \bar{C}_{i-1})$ is $\overline{\mathcal{G}}_{T_{i-1}}$-measurable, we get that $Z, \varepsilon_{i+1}, \bar{C}_i$ are independent. This achieves the proof. □

6 Main result

We can now state the main result of this paper.

Theorem 6.1. *Assume (H1) and (C). For deterministic $N = n$, for $h = h_n$, the two sequences of experiments $(\mathcal{E}^{h_n,n})$ and $(\mathcal{G}^{h_n,n})$ are asymptotically equivalent, in the sense of the Le Cam deficiency distance Δ as $n \to \infty$ if $h_n \to 0$, $nh_n \to \infty$ and $nh_n^2 \to 0$,*

$$\Delta(\mathcal{E}^{h_n,n}, \mathcal{G}^{h_n,n}) \to 0$$

Set $T = nh_n$ and consider the stopping times

$$\tau_n = \rho_{nh_n}(X), \quad \bar{\tau}_n = \overline{\rho}_{nh_n}(X), \quad S_n = \bar{\tau}_n \wedge \tau_n. \tag{52}$$

Before going into the last steps, let us make a brief summary using the time changed experiments. We have the following inequality:

$$\begin{aligned}&\Delta(\mathcal{E}^{h_n,n}, \mathcal{G}^{h_n,n}) \leq \\ &\Delta(\mathcal{E}^{h_n,n}, \mathcal{E}_0^{nh_n}) + \Delta(\mathcal{E}_0^{nh_n}, \tilde{\mathcal{E}}_0^{\tau_n}) + \Delta(\tilde{\mathcal{E}}_0^{\tau_n}, \tilde{\mathcal{G}}_0^{\bar{\tau}_n}) + \Delta(\tilde{\mathcal{G}}_0^{\bar{\tau}_n}, \mathcal{G}_0^{nh_n}) + \Delta(\mathcal{G}_0^{nh_n}, \mathcal{G}^{h_n,n}).\end{aligned}$$

By Lemma 3.3 and 4.1, we have, as $n \to +\infty$, $h_n \to 0$, $nh_n \to +\infty$, $nh_n^2 \to 0$,

$$\Delta(\mathcal{G}_0^{nh_n}, \mathcal{G}^{h_n,n}) = 0, \quad \Delta(\mathcal{E}_0^{nh_n}, \mathcal{E}^{h_n,n}) \to 0.$$

By Propositions 5.2 and 5.4,

$$\Delta(\mathcal{E}_0^{nh_n}, \tilde{\mathcal{E}}_0^{\tau_n}) \to 0, \quad \Delta(\mathcal{G}_0^{nh_n}, \tilde{\mathcal{G}}_0^{\bar{\tau}_n}) \to 0.$$

Hence, to achieve the proof of Theorem 6.1, it remains to study $\Delta(\tilde{\mathcal{E}}_0^{\tau_n}, \tilde{\mathcal{G}}_0^{\bar{\tau}_n})$. Here , these two experiments are observed up to distinct stopping times, which leads to an additional difficulty. Using (52), the triangle inequality yields:

$$\Delta(\tilde{\mathcal{E}}_0^{\tau_n}, \tilde{\mathcal{G}}_0^{\bar{\tau}_n}) \leq \Delta(\tilde{\mathcal{E}}_0^{\tau_n}, \tilde{\mathcal{E}}_0^{S_n}) + \Delta(\tilde{\mathcal{E}}_0^{S_n}, \tilde{\mathcal{E}}_0^{\bar{\tau}_n}) + \Delta(\tilde{\mathcal{E}}_0^{\bar{\tau}_n}, \tilde{\mathcal{G}}_0^{\bar{\tau}_n}). \tag{53}$$

Therefore, we have to study the Δ- deficiency distance respectively for the same experiment observed up to two distinct times and for two experiments observed up to the random time $\bar{\tau}_n$.

The result will follow from Propositions 6.1 and 6.2 which are proved afterwards.

6.1 Experiments stopped at $\bar{\tau}_n$

Let us first study the last term of (53). These two experiments of have the same sample space and are respectively associated with the family of distributions $\tilde{P}_b$ (resp $\tilde{Q}_b$) on $C(\mathbb{R}^+, \mathbb{R})$ of $(Y_u, u \geq 0)$ given by (26) (resp. $(\overline{Y}_u, u \geq 0)$ given by (35)). Hence,

$$\Delta(\tilde{\mathcal{E}}_0^{\bar{\tau}_n}, \tilde{\mathcal{G}}_0^{\bar{\tau}_n}) \leq \sup_{b \in \mathcal{F}} ||\tilde{P}_b/_{\mathcal{C}_{\bar{\tau}_n}} - \tilde{Q}_b/_{\mathcal{C}_{\bar{\tau}_n}}||_{TV} = \Delta_0(\tilde{\mathcal{E}}_0^{\bar{\tau}_n}, \tilde{\mathcal{G}}_0^{\bar{\tau}_n}).$$

Proposition 6.1. *Under (H1) and (C), we have,*

$$\Delta_0(\tilde{\mathcal{E}}_0^{\bar{\tau}_n}, \tilde{\mathcal{G}}_0^{\bar{\tau}_n}) \leq L\sigma_1^2 (nh_n^2)^{1/2} \left(\frac{K^2 \sigma_1^2 h_n}{3\sigma_0^4} + \frac{1}{2} \right)^{1/2}.$$

Proof. Using the bound of Proposition (8.4) yields

$$||\tilde{P}_b/_{\mathcal{C}_A} - \tilde{Q}_b/_{\mathcal{C}_A}||_{TV} \leq \sqrt{\frac{1}{2} K(\tilde{P}_b/_{\mathcal{C}_A},\ \tilde{P}_b/_{\mathcal{C}_A})}.$$

These two distributions are associated with diffusion type processes so , noting that $Y_0 = \overline{Y}_0$, an application of the Girsanov formula up to the a.s. finite random time $A = A(X_.) = \bar{\tau}_n(X_.)$ (with (X_v) the canonical process of $C(\mathbb{R}^+, \mathbb{R})$) yields,

$$\frac{d\tilde{P}_b/_{\mathcal{C}_A}}{d\tilde{Q}_b/_{\mathcal{C}_A}} = \exp\left(\int_0^A (f(X_v - \bar{f}(v, X_.))dX_v - \frac{1}{2} \int_0^A (f(X_v) - \bar{f}(v, X_.))^2 dv \right). \tag{54}$$

Hence, using that (X_t) has drift term f under $\tilde{P}_b$ yields,

$$\begin{aligned} K(\tilde{P}_b/_{\mathcal{C}_A}, \tilde{Q}_b/_{\mathcal{C}_A}) &= E_{\tilde{P}_b} \left(\frac{1}{2} \int_0^A (f(X_v) - \bar{f}(v, X_.))^2 dX_v + \int_0^A (f(X_v) - \bar{f}(v, X_.))dB_v \right) \\ &= \frac{1}{2} E_{\tilde{P}_b/_{\mathcal{C}_A}} \left(\int_0^A (f(X_v) - \bar{f}(v, X_.))^2 dv \right). \end{aligned}$$

Setting $T_i = T_i(X)$ and using (5.1), we get,

$$\int_0^A (f(X_v) - \bar{f}(v, X_.))^2 dv = \sum_{i \geq 1} \int_{A \wedge T_{i-1}}^{A \wedge T_i} \left(f(X_v) - f(X_{T_{i-1}} \right)^2 dv.$$

Using now that $f = b/\sigma^2$ is Lipschitz with constant L (see (4)) and that , for $i = 1, \ldots, n$, $T_i = T_i(X_.) = \overline{\rho}_{t_i}(X_.)$ yields

$$\int_0^{\bar{\tau}_n} (f(X_v) - \bar{f}(v, X_.))^2 dv \leq L^2 \sum_{i=1}^n \int_{T_{i-1}}^{T_i} \left(X_v - X_{T_{i-1}} \right)^2 dv. \tag{55}$$

Under $\tilde{P}_b$,

$$X_v - X_{T_{i-1}} = \int_{T_{i-1}}^v f(X_u)du + B_v - B_{T_{i-1}},$$

where (B_v) is a Brownian motion. Therefore,

$$\left(X_v - X_{T_{i-1}})\right)^2 \leq 2\left[\left(\int_{T_{i-1}}^{v} f(X_u)du\right)^2 + (B_v - B_{T_{i-1}})^2\right].$$

This yields

$$\int_0^{\bar{\tau}_n} (f(X_v) - \bar{f}(v, X_.))^2 dv \leq 2L^2(A_1(\bar{\tau}_n) + A_2(\bar{\tau}_n)),$$

with

$$A_1(\bar{\tau}_n) = \sum_{i=1}^{n} \int_{T_{i-1}}^{T_i} \left(\int_{T_{i-1}}^{u} f(X_v)dv\right)^2 du;\ A_2(\bar{\tau}_n) = \sum_{i=1}^{n} \int_{T_{i-1}}^{T_i} \left(B_u - B_{T_{i-1}}\right)^2 du \tag{56}$$

Using (4) and that $T_i - T_{i-1} \leq \sigma_1^2 h_n$ by (41),

$$A_1(\bar{\tau}_n) \leq \frac{K^2}{\sigma_0^4} \sum_{i=1}^{n} (T_i - T_{i-1})^3 \leq \frac{K^2}{\sigma_0^4} n(\sigma_1^2 h_n)^3,$$

For the second term, using definition (47),

$$A_2(\bar{\tau}_n) = \sum_{i=1}^{n} \int_{T_{i-1}}^{T_i} \left(B_u - B_{T_{i-1}}\right)^2 du \leq \sum_{i=1}^{n} \int_0^{\sigma_1^2 h_n} (\bar{B}_v^{(i)})^2 dv. \tag{57}$$

Thus,

$$K(\tilde{P}_b/_{\mathcal{C}_{\bar{\tau}_n}}, \tilde{Q}_b/_{\mathcal{C}_{\bar{\tau}_n}}) \leq 2L^2\left(\frac{K^2}{3\sigma_0^4} n(\sigma_1^2 h_n)^3 + n\frac{(\sigma_1^2 h_n)^2}{2}\right)$$

Hence, the result. □

6.2 An asymptotic sufficiency property

Let us recall that $\tau_n = \rho_{nh_n}(X), \bar{\tau_n} = \overline{\rho}_{nh_n}(X)$ and $S_n = inf(\tau_n \bar{\tau_n})$ (see (52)) . It remains to prove (see (53)) that

$$\Delta(\tilde{\mathcal{E}}_0^{S_n}, \tilde{\mathcal{E}}_0^{\tau_n}) \to 0, \quad \Delta(\tilde{\mathcal{E}}_0^{S_n}, \tilde{\mathcal{E}}_0^{\bar{\tau_n}}) \to 0$$

i.e. to prove the asymptotic sufficiency of $\mathcal{C}_{S_n}$ for the experiments $\tilde{\mathcal{E}}_0^{\tau_n}$ and $\tilde{\mathcal{E}}_0^{\bar{\tau_n}}$.

Lemma 6.1. *Under (H1), there exists a constant D depending on $K, K_\sigma, \sigma_1, \sigma_0$ such that,*

$$E_{\tilde{P}_b}|\tau_n - \bar{\tau_n}| \leq Dnh_n^2.$$

($D = 2K_\sigma^2 + \frac{2KK_\sigma}{\sigma_0} + \left(\frac{2}{3}(K_\sigma^2 + \sigma_1 K_\sigma)\right)^{1/2}$).

Proof. By (24) and (32),

$$\tau_n - \bar{\tau}_n = \sum_{i=1}^{n} \int_{t_{i-1}}^{t_i} \left(\sigma^2(X_v) - \sigma^2(X_{t_{i-1}})\right) dv.$$

Under $\tilde{P}_b$, we get, denoting by $\mathcal{L}$ the generator of the diffusion (Y_u) ($\mathcal{L}f = \frac{1}{2}f'' + \frac{b}{\sigma^2}f'$),

$$\sigma^2(X_v) = \sigma^2(X_{t_{i-1}}) + \int_{t_{i-1}}^{v} \mathcal{L}\sigma^2(X_u)du + \int_{t_{i-1}}^{v} (\sigma^2)'(X_u)dB_u,$$

Thus, $\int_{t_{i-1}}^{t_i} \left(\sigma^2(X_v) - \sigma^2(X_{t_{i-1}})\right) dv = B_1(i) + B_2(i)$, with

$$B_1(i) = \int_{t_{i-1}}^{t_i} dv \int_{t_{i-1}}^{v} \mathcal{L}\sigma^2(X_u)du \ ; \ B_2(i) = \int_{t_{i-1}}^{t_i} dv \int_{t_{i-1}}^{v} (\sigma^2)'(X_u)dB_u.$$

Now Condition (C) and (H1) ensure that $\|\mathcal{L}\sigma^2\|_\infty$ is bounded by $D_1 = K_\sigma(\sigma_1 + K_\sigma + \frac{2K}{\sigma_0})$ so that,

$$|B_1(i)| \le D_1 \frac{h_n^2}{2}.$$

For the second term changing integrations yields,

$$B_2(i) = \int_{t_{i-1}}^{t_i} dv \int_{t_{i-1}}^{v} (\sigma^2)'(X_u)dB_u = \int_{t_{i-1}}^{t_i} (t_i - u)(\sigma^2)'(X_u)dB_u.$$

Thus,

$$\sum_{i=1}^{n} B_2(i) = \int_0^{nh_n} H_u^{(n)} dB_u, \text{ where } H_u^{(n)} = \sum_{i=1}^{n} 1_{]t_{i-1},t_i]}(u)(t_i - u)(\sigma^2)'(X_u).$$

This stochastic integral is well defined with finite quadratic variation and satisfies under Condition (C),

$$E_{\tilde{P}_b} \left(\sum_{i=1}^{n} B_2(i) \right)^2 = E_{\tilde{P}_b} \int_0^{nh_n} (H_u^{(n)})^2 du \le \|(\sigma^2)'\|_\infty^2 \, n \frac{h_n^3}{3}.$$

Finally, joining these two results and applying the Burkholder-Davies-Gundy inequality with $D' = \frac{D_1}{2} \vee (\frac{2\sigma_1 K_\sigma}{3})^{1/2}$

$$E_{\tilde{P}_b}|\tau_n - \bar{\tau}_n| \le D'(nh_n^2 + (nh_n^3)^{1/2}) = D' \, nh_n^2(1 + \frac{1}{\sqrt{nh_n}}) \le D' \, nh_n^2.$$

This achieves the proof if nh_n is bounded or goes to infinity with n. $\square$

Proposition 6.2. *Under (H1) and (C), we have,*

$$\Delta(\mathcal{E}_0^{\tau_n}, \mathcal{E}_0^{S_n}) \le \frac{K}{2^{1/2}\sigma_0^2}(cnh_n^2)^{1/2}, \qquad \Delta(\mathcal{E}_0^{\bar{\tau}_n}, \mathcal{E}_0^{S_n}) \le \frac{K}{2^{1/2}\sigma_0^2}(cnh_n^2)^{1/2}.$$

Proof. As $\mathcal{E}_0^{S_n}$ is a restriction of $\mathcal{E}_0^{\tau_n}$ to a smaller σ-algebra, $\delta(\mathcal{E}_0^{\tau_n}, \mathcal{E}_0^{S_n}) = 0$.

To evaluate the other deficiency, we introduce a kernel from $\mathcal{E}_0^{S_n}$ to $\mathcal{E}_0^{\tau_n}$. Let $A \in \mathcal{C}_{\tau_n}$, and set

$$N(\omega, A) = E_{\tilde{P}_0}(1_A|\mathcal{C}_{S_n})(\omega).$$

Using the kernel N, $N(\tilde{P}_b|\mathcal{C}_{S_n})$ defines a probability on $(\Omega, \mathcal{C}_{\tau_n})$. Let us compute the density of the probability $N(\tilde{P}_b|\mathcal{C}_{S_n})$ w.r.t. $\tilde{P}_0|\mathcal{C}_{\tau_n}$: for $A \in \mathcal{C}_{\tau_n}$,

$$N(\tilde{P}_b|\mathcal{C}_{S_n})(A) = \int_\Omega N(\omega, A)d(\tilde{P}_b|\mathcal{C}_{S_n}) = E_{\tilde{P}_0}\left(\frac{d\tilde{P}_b}{d\tilde{P}_0}|\mathcal{C}_{S_n} E_{\tilde{P}_0}(1_A|\mathcal{C}_{S_n})\right) = E_{\tilde{P}_0}\left(\frac{d\tilde{P}_b}{d\tilde{P}_0}|\mathcal{C}_{S_n} 1_A\right).$$

Thus, the probability $N(\tilde{P}_b|\mathcal{C}_{S_n})$ on $(\Omega, \mathcal{C}_{\tau_n})$ admits a density w.r.t. $\tilde{P}_0|\mathcal{C}_{\tau_n}$ equal to

$$\frac{d\tilde{P}_b}{d\tilde{P}_0}|\mathcal{C}_{S_n}.$$

(Here, $\tilde{P}_0$ corresponding to $b = 0$, is the distribution of $(\eta + B_u, u \geq 0)$). Denote by $\tilde{L}_T(b)$ the density $\frac{d\tilde{P}_b}{d\tilde{P}_0}|\mathcal{C}_T$ for any bounded stopping time T. We have

$$\tilde{L}_T(b) = \exp\left(\int_0^T f(X_u)dX_u - \int_0^T \frac{1}{2}f^2(X_u)du\right)$$

with $f = b/\sigma^2$. Thus,

$$\frac{d\tilde{P}_b}{d\tilde{P}_0}|\mathcal{C}_{\tau_n} = \tilde{L}_{\tau_n}(b) = \tilde{L}_{S_n}(b)V_n = \frac{d\tilde{P}_b}{d\tilde{P}_0}|\mathcal{C}_{S_n}V_n,$$

with

$$U_n = \log V_n = \int_{S_n}^{\tau_n} f(X_u)dX_u - \int_{S_n}^{\tau_n} \frac{1}{2}f^2(X_u)du.$$

We deduce:

$$\frac{d\tilde{P}_b|\mathcal{C}_{\tau_n}}{dN(\tilde{P}_b|\mathcal{C}_{S_n})} = V_n.$$

By the Pinsker inequality (see Section 8.2 and Tsybakov (2009)), we have:

$$\|N(\tilde{P}_b|\mathcal{C}_{S_n}) - \tilde{P}_b|\mathcal{C}_{\tau_n}\|_{TV} = \frac{1}{2}\int_\Omega d\tilde{P}_0|\tilde{L}_{S_n}(b) - \tilde{L}_{\tau_n}(b)|$$

$$\leq \sqrt{K(\tilde{P}_b|\mathcal{C}_{\tau_n}, N(\tilde{P}_b|\mathcal{C}_{S_n}))/2}.$$

Now, by Lemma 6.1

$$K(\tilde{P}_b|\mathcal{C}_{\tau_n}, N(\tilde{P}_b|\mathcal{C}_{S_n})) = E_{\tilde{P}_b|\mathcal{C}_{\tau_n}} U_n = E_{\tilde{P}_b|\mathcal{C}_{\tau_n}} \int_{S_n}^{\tau_n} \frac{1}{2}f^2(X_u)du \leq \frac{K^2}{2\sigma_0^4}E_{\tilde{P}_b}|\tau_n - \bar{\tau}_n|$$

$$\leq \frac{K^2}{\sigma_0^4} cnh_n^2.$$

As

$$\delta(\mathcal{E}_0^{S_n}, \mathcal{E}_0^{\tau_n}) \leq \sup_{b \in \mathcal{F}_K} \|N(\tilde{P}_b|\mathcal{C}_{S_n}) - \tilde{P}_b|\mathcal{C}_{\tau_n}\|_{TV},$$

we obtain the first inequality. We proceed analogously for the other inequality.

□

7 Extensions

If we do not assume that b is bounded, then, we have to make additional assumptions on f, μ, (Y_u) and (η_t) defined in (3),(28) and (6):

(H2) f and μ are Lipschitz.

(H3) $\sup_{u \geq 0} \mathbb{E} Y_u^2 + \sup_{t \geq 0} \mathbb{E} \eta_t^2 \leq K$.

Indeed, instead of using the bounds for f, μ, we have to use their linear growth ($f^2(x) + \mu^2(x) \leq C(1+x^2)$ for a constant C). Then, when taking expectations in the proofs of Lemma 4.1, Proposition 6.1, Lemma 6.1 and Proposition 6.2, uniform upper bounds of second-order moments are required. Assumption (H3) is fulfilled for instance when the diffusion model for ξ is ergodic with $\mathbb{E}\eta^2 < +\infty$ ($\eta = \xi_0$).

8 Apppendix

8.1 The Le Cam deficiency distance

In this section, we recall the main definitions and properties used in relation with the Le Cam deficiency distance. The Le Cam distance between statistical experiments is generally denoted by Δ. In what follows, all measurable spaces are supposed to be Polish metric spaces equipped with their Borel σ-algebras. The Le Cam distance was introduced to compare experiments having the same parameter space $\mathcal{F}$, but possibly different sample spaces. Consider two statistical experiments $\mathcal{E} = (\Omega, \mathcal{A}, (P_f)_{f \in \mathcal{F}})$ and $\mathcal{G} = (\mathcal{X}, \mathcal{C}, (Q_f)_{f \in \mathcal{F}})$ and assume that the two families $(P_f)_{f \in \mathcal{F}}$, $(Q_f)_{f \in \mathcal{F}}$ are dominated. A Markov kernel $M(\omega, dx)$ from $(\Omega, \mathcal{A})$ to $(\mathcal{X}, \mathcal{C})$ is a mapping from Ω into the set of probability measures on $(\mathcal{X}, \mathcal{C})$ such that:

- For all $C \in \mathcal{C}$, $\omega \to M(\omega, C)$ is measurable on $(\Omega, \mathcal{A})$,

- For all $\omega \in \Omega$, $M(\omega, dx)$ is probability measure on $(\mathcal{X}, \mathcal{C})$.

The image MP_f of P_f under M is defined by:

$$MP_f(C) = \int_\Omega M(\omega, C) dP_f(\omega).$$

The experiment $M\mathcal{E} = (\mathcal{X}, \mathcal{C}, (MP_f)_{f \in \mathcal{F}})$ is called a randomisation of $\mathcal{E}$ by the kernel M. If the kernel is deterministic, *i.e.* for $T : (\Omega, \mathcal{A}) \to (\mathcal{X}, \mathcal{C})$ a random variable, $T(\omega, C) = 1_C(T(\omega))$, the experiment $T\mathcal{E}$ is called the image experiment by the random variable T.

Let $\mathcal{M}_{\Omega:\mathcal{X}}$ denote the set of Markov kernels from $(\Omega, \mathcal{A})$ to $(\mathcal{X}, \mathcal{C})$.

Definition 8.1. *(1) The* deficiency *of $\mathcal{E}$ with respect to $\mathcal{G}$ is given by*

$$\delta(\mathcal{E}, \mathcal{G}) = \inf_{M \in \mathcal{M}_{\Omega:\mathcal{X}}} \sup_{f \in \mathcal{F}} ||MP_f - Q_f||_{TV},$$

where $||.||_{TV}$ denotes the total variation norm for measures.

(2) $\Delta(\mathcal{E}, \mathcal{G}) = max\{\delta(\mathcal{E}, \mathcal{G}), \delta(\mathcal{G}, \mathcal{E})\}$.

Actually, Δ is a pseudo-distance. When $\Delta(\mathcal{E}, \mathcal{G}) = 0$, the two experiments are said to be equivalent.
When the two experiments have the same sample space : $(\Omega, \mathcal{A}) = (\mathcal{X}, \mathcal{C})$, it is possible to define

$$\Delta_0((\mathcal{E}, \mathcal{G}) = \sup_{f \in \mathcal{F}} ||P_f - Q_f||_{TV}.$$

The following inequality is useful:

$$\Delta(\mathcal{E}, \mathcal{G}) \leq \Delta_0(\mathcal{E}, \mathcal{G}).$$

The sufficiency of a σ-algebra or a statistic can be expressed in terms of the Δ-distance.

Proposition 8.1. *Consider the experiment $\mathcal{E}$ and let $\mathcal{B} \subset \mathcal{A}$ be a σ-algebra on Ω. Consider the experiment experiment $\mathcal{E}/_\mathcal{B} = (\Omega, \mathcal{B}, P_f/_\mathcal{B})$ which is the restriction of $\mathcal{E}$ to $\mathcal{B}$. Then, $\mathcal{B}$ is sufficient for $\mathcal{E}$ if and only if $\Delta(\mathcal{E}, \mathcal{E}/_\mathcal{B}) = 0$.*
Let $T : (\Omega, \mathcal{A}) \to (\mathcal{X}, \mathcal{C})$ be a random variable. The statistic T is sufficient for $\mathcal{E}$ if and only if $\Delta(\mathcal{E}, T\mathcal{E}) = 0$.

Proof. Consider first the deterministic kernel from $(\Omega, \mathcal{A})$ to $(\Omega, \mathcal{B})$ given by $M(\omega, B) = 1_B(\omega)$. As $MP_f = P_f/_\mathcal{B}$, $\delta((\mathcal{E}, \mathcal{E}/_\mathcal{B}) = 0$. Then, consider the kernel N from $(\Omega, \mathcal{B})$ to $(\Omega, \mathcal{A})$ given by:

$$N(\omega, A) = E_{P_f}(1_A|\mathcal{B}).$$

By the sufficiency of $\mathcal{B}$, $N(\omega, A) = E_{P_*}(1_A|\mathcal{B})$ admits a version independent of the parameter f. By the assumption on the sample spaces, there is a regular version of the conditional

probability given $\mathcal{B}$. Hence, $\delta((\mathcal{E}/_{\mathcal{B}}, \mathcal{E}) = 0$.
By definition, the statistic T is sufficient is the σ-algebra $\sigma(T)$ is sufficient. As $\Delta(T\mathcal{E}, \mathcal{E}/_{\sigma(T)}) = 0$, it follows immediately that T is sufficient if and only if $\Delta(\mathcal{E}, T\mathcal{E}) = 0$.

Note that, when the family (P_f) is dominated by a probability of the model P_{f_0} for some $f_0 \in \mathcal{F}$, then, we can take $N(\omega, A) = E_{P_{f_0}}(1_A|\mathcal{B})$. □

Now, we introduce an asymptotic framework $\varepsilon \to 0$ and consider the families of experiments $\mathcal{E}^\varepsilon = (\Omega^\varepsilon, \mathcal{A}^\varepsilon, (P_f^\varepsilon)_{f\in\mathcal{F}})$, $\mathcal{G}^\varepsilon = (\mathcal{X}^\varepsilon, \mathcal{C}^\varepsilon, (Q_f^\varepsilon)_{f\in\mathcal{F}})$, $\mathcal{B}^\varepsilon \subset \mathcal{A}^\varepsilon$ a σ-algebra, $T^\varepsilon : (\Omega^\varepsilon, \mathcal{A}^\varepsilon) \to (\mathcal{X}^\varepsilon, \mathcal{C}^\varepsilon)$ a statistic.

Definition 8.2.

- *The families $\mathcal{E}^\varepsilon$, $\mathcal{G}^\varepsilon$ are said to be asymptotically equivalent as ε tends to* 0 *if $\Delta(\mathcal{E}^\varepsilon, \mathcal{G}^\varepsilon)$ tends to* 0.
- *The σ-algebra $\mathcal{B}^\varepsilon$ is said to be asymptotically sufficient if $\Delta(\mathcal{E}^\varepsilon, \mathcal{E}^\varepsilon/_{\mathcal{B}^\varepsilon})$ tends to* 0.
- *The statistic T^ε is asymptotically sufficient if $\Delta(\mathcal{E}^\varepsilon, T^\varepsilon\mathcal{E}^\varepsilon)$ tends to* 0.

The comparison of statistical experiments with different sample spaces is difficult. So, it is often the case that an accompanying experiment $\mathcal{H}^\varepsilon = ((\Omega^\varepsilon, \mathcal{A}^\varepsilon, (R_f^\varepsilon)_{f\in\mathcal{F}})$ having the same sample space as $\mathcal{E}^\varepsilon$ is introduced. Then, we have the following result.

Proposition 8.2.

- *If, as ε tends to* 0*, $\Delta_0(\mathcal{E}^\varepsilon, \mathcal{H}^\varepsilon)$ tends to* 0 *and $\Delta(\mathcal{H}^\varepsilon, \mathcal{G}^\varepsilon)$ tends to* 0*, then, the families $\mathcal{E}^\varepsilon$ and $\mathcal{G}^\varepsilon$ are asymptotically equivalent.*
- *If, as ε tends to* 0*, $\Delta_0(\mathcal{E}^\varepsilon, \mathcal{H}^\varepsilon)$ tends to* 0 *and $\Delta(\mathcal{H}^\varepsilon, T^\varepsilon\mathcal{H}^\varepsilon)$ tends to* 0*, then T^ε is asymptotically sufficient for $\mathcal{E}^\varepsilon$.*

Proof. For the first point, we use the inequality:

$$\Delta(\mathcal{E}^\varepsilon, \mathcal{G}^\varepsilon) \le \Delta(\mathcal{E}^\varepsilon, \mathcal{H}^\varepsilon) + \Delta(\mathcal{H}^\varepsilon, \mathcal{G}^\varepsilon) \le \Delta_0(\mathcal{E}^\varepsilon, \mathcal{H}^\varepsilon) + \Delta(\mathcal{H}^\varepsilon, \mathcal{G}^\varepsilon).$$

For the second point,

$$\Delta(\mathcal{E}^\varepsilon, T^\varepsilon\mathcal{E}^\varepsilon) \le \Delta(\mathcal{E}^\varepsilon, \mathcal{H}^\varepsilon) + \Delta(\mathcal{H}^\varepsilon, T^\varepsilon\mathcal{H}^\varepsilon) + \Delta(T^\varepsilon\mathcal{H}^\varepsilon, T^\varepsilon\mathcal{E}^\varepsilon) \le 2\Delta_0(\mathcal{E}^\varepsilon, \mathcal{H}^\varepsilon) + \Delta(\mathcal{H}^\varepsilon, T^\varepsilon\mathcal{H}^\varepsilon).$$

□

8.2 Auxiliary results

The following results are used above.

Proposition 8.3. *Let $(B_t, t \ge 0)$ be a Brownian motion with respect to a filtration $(\mathcal{F}_t, t \ge 0)$ (satisfying the usual conditions) and let τ be a positive $\mathcal{F}_0$-measurable random variable. Then, the process $(W(t) = \frac{1}{\sqrt{\tau}} B_{\tau t}, t \ge 0)$ is a standard Brownian motion, independent of $\mathcal{F}_0$.*

Proof. As τ is $\mathcal{F}_0$-measurable, for all $t \geq 0$, τt is a $(\mathcal{F}_t)$- stopping time. By the optional sampling theorem, we deduce that the processes $(B_{\tau t}, t \geq 0)$ and $(B^2_{\tau t} - \tau t, t \geq 0)$ are local martingales with respect to the filtration $(\mathcal{F}_{\tau t})$, with continuous sample paths, null at 0. As τ is $\mathcal{F}_0$-measurable, the same holds for the two processes $(W(t))$ and $(W^2(t) - t)$. Thus, by Paul Lévy's characterization, we deduce that $(W(t))$ is a Brownian motion, with respect to the same filtration. Thus, $(W(t))$ is independent of $\mathcal{F}_0$. □

We recall useful inequalities for the total variation distance between probability measures. The following one is called the first Pinsker inequality (see *e.g* Tsybakov, 2009).

Let $(\mathcal{X}, \mathcal{A})$ be a measurable space, P, Q two probability measures on $(\mathcal{X}, \mathcal{A})$, ν a σ-finite measure on $(\mathcal{X}, \mathcal{A})$ such that $P \ll \nu$, $Q \ll \nu$ and set $p = dP/d\nu, q = dQ/d\nu$. The total variation distance between P and q is defined by:

$$||P - Q||_{TV} = \sup_{A \in \mathcal{A}} |P(A) - Q(A)| = \frac{1}{2} \int |p - q| d\nu.$$

The Kullback divergence of P w.r.t. Q is given by:

$$K(P, Q) = \int \log \frac{dP}{dQ} dP \quad \text{if } P \ll Q, \quad = +\infty \quad \text{otherwise.}$$

Proposition 8.4.

$$||P - Q||_{TV} \leq \sqrt{K(P, Q)/2}.$$

The remarkable feature of this inequality is that the left-hand side is a symetric quantity whereas the right-hand-side is not. The noteworthy consequence is that it is possible to choose, for the right-hand-side, $K(P, Q)$ or $K(Q, P)$.

References

[1] Brown, L.D., Carter, A.V., Low, M.G. and Zhang, C.H. (2004). Equivalence theory for density estimation, Poisson processes and Gaussian white noise with drift. *Ann. Statist.* **32** 2074-2097.

[2] Brown, L.D. and Low, M.G. (1996). Asymptotic equivalence of nonparametric regression and white noise. *Ann. Statist.* **24** 2384-2398.

[3] Dalalyan, A. and Reiss, M. (2006). Asymptotic statistical equivalence for scalar ergodic diffusions. *Probab. Theory Related Fields* **134** 248-282.

[4] Dalalyan, A. and Reiss, M. (2007). Asymptotic statistical equivalence for ergodic diffusions: the multidimensional case. *Probab. Theory Related Fields* **137** 25-47.

[5] Delattre, S. and Hoffmann, M. (2002). Asymptotic equivalence for nul recurrent diffusion. *Bernoulli* **8** 139-174.

[6] Genon-Catalot, V. (1990). Maximum contrast estimation for diffusion processes from discrete observations. *Statistics* **21** 99-116.

[7] Genon-Catalot, V., Larédo, C. and Nussbaum, M. (2002). Asymptotic Equivalence for Estimating a Poisson Intensity and a Positive diffusion drift. *Ann. Statist.* **30** 731-753.

[8] Gobet, E. (2002). LAN property for ergodic diffusions with discrete observations. *Ann. lnstit. Henri Poincaré, Probab. Stat.* **38** 711-737.

[9] Grama, I. and Nussbaum, M. (1998). Asymptotic equivalence for nonparametric generalized linear models. *Probab. Theory Related Fields* **111** 167-214.

[10] Grama, I. and Nussbaum, M. (2002). Asymptotic equivalence of nonparametric regression. *Math. Methods Statist.* **11** 1-36.

[11] Kessler, M. (1997). Estimation of an ergodic diffusion from discrete observations. *Scand. J. Statist.* **24** 211-229.

[12] Larédo, C. (1990). A sufficient condition for asymptotic sufficiency of incomplete observations of a diffusion process. *Ann. Statist.* **18** 1158-1171.

[13] Le Cam, L.(1986) *Asymptotic Methods in Statistical Decision Theory.* Springer, New York.

[14] Le Cam, L. and Yang, L.G. (2000). *Asymptotics in Statistics. Some Basic Concepts* . 2nd Edition. Springer, New York.

[15] Milstein, G. and Nussbaum, M. (1998). Diffusion approximation for nonparametric autoregression. *Probab. Theory Related Fields* **112** 535-543.

[16] Nussbaum, M. (1996). Asymptotic equivalence of density estimation and Gaussian white noise. *Ann. Statist.* **24** 2399-2430.

[17] Reiss, M. (2008). Asymptotic equivalence for nonparametric regression with multivariate and random design. *Ann. Statist.* **36** 1957-1982.

[18] Tsybakov, A. B. (2009). *Introduction to nonparametric estimation.* Revised and extended from the 2004 French original. Translated by Vladimir Zaiats. Springer Series in Statistics. Springer, New York.