The Annals of Statistics
2014, Vol. 42, No. 3, 1145–1165
DOI: 10.1214/14-AOS1216

ASYMPTOTIC EQUIVALENCE OF NONPARAMETRIC DIFFUSION AND EULER SCHEME EXPERIMENTS

BY VALENTINE GENON-CATALOT AND CATHERINE LARÉDO

Université Paris Descartes and INRA

We prove a global asymptotic equivalence of experiments in the sense of Le Cam's theory. The experiments are a continuously observed diffusion with nonparametric drift and its Euler scheme. We focus on diffusions with nonconstant-known diffusion coefficient. The asymptotic equivalence is proved by constructing explicit equivalence mappings based on random time changes. The equivalence of the discretized observation of the diffusion and the corresponding Euler scheme experiment is then derived. The impact of these equivalence results is that it justifies the use of the Euler scheme instead of the discretized diffusion process for inference purposes.

1. Introduction. Proving global asymptotic equivalence of statistical experiments by means of the Le Cam theory of deficiency [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 one. Concretely, in the case of bounded loss functions, a solution to a nonparametric problem in an experiment 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 hold in a globally asymptotically equivalent experiment. The theory also allows to prove asymptotic sufficiency of the restriction of an experiment to a smaller σ-field. When explicit transformations from one experiment to another one are obtained, statistical procedures can be carried over from one experiment to the other one. There is an abundant literature devoted to establishing asymptotic equivalence results. Before considering diffusion experiments, we recall the main contributions in this domain. The first results concern the asymptotic equivalence of density estimation and white noise model [Nussbaum (1996)] and nonparametric regression and white noise [Brown and Low (1996)]. These results were extended to the equivalence of nonparametric regression with random design and white noise [Brown et al. (2002)]. The equivalence between the observation of n independent random variables $X_i, i = 1, \dots, n$ with densities $p(x, \theta_i)$, such that $\theta_i = f(i/n)$ and a nonparametric Gaussian shift experiment with drift linked with f is proved in Grama and Nussbaum (1998, 2002). In Brown et al. (2004), the equivalences concern Poisson

Received June 2013; revised March 2014.

MSC2010 subject classifications. Primary 62B15, 62G20; secondary 62M99, 60J60.

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

processes with nonparametric intensity and white noise. Carter (2006) considers the equivalence of a fixed design regression in two dimensions and a Brownian sheet process with drift. This result is extended to regression experiments with arbitrary dimension in Reiss (2008). The regression model with nonregular errors yields different results, the equivalence being with independent point Poisson processes [Meister and Reiss (2013)]. A step forward in another direction concerns the equivalence of nonparametric autoregression and nonparametric regression [Grama and Neumann (2006)]. Negative results are also important such as the nonequivalence of nonparametric regression and density or white noise when the regression function has smoothness index $1/2$ [Brown and Zhang (1998)]. To our knowledge, the only paper studying the equivalence problem for regression with unknown variances is Carter (2007). More recently, the class of studied models has been enlarged to stationary Gaussian processes with unknown spectral density which are equivalent to white noise [Golubev, Nussbaum and Zhou (2010)]. Another direction concerns inverse problems in regression and white noise [Meister (2011)]. Unusual rates formerly obtained by Gloter and Jacod (2001) find their mathematical understanding with the equivalence result of Reiss (2011) where the discretization of a continuous Gaussian martingale observed with noise on a fixed time interval is equivalent to a Gaussian white noise experiment with the same unusual rate ($n^{-1/4}$ instead of $n^{-1/2}$ in the noise intensity).

Diffusion models defined by stochastic differential equations have also been investigated. References concern nonparametric drift estimation with known constant diffusion coefficient. Genon-Catalot, Larédo and Nussbaum (2002) studied the equivalence of a transient diffusion having positive drift and small constant diffusion coefficient with a white noise model and other related experiments. In the case of recurrent diffusion models, global equivalence with Gaussian white noise no longer holds [Delattre and Hoffmann (2002) for null recurrent diffusions, Dalalyan and Reiss (2006, 2007) for ergodic scalar and multidimensional diffusions]. ARCH-GARCH models exhibit nonstandard equivalence results when compared to their limiting diffusion experiments. In a parametric context, Wang (2002) proves the nonequivalence of the GARCH-experiment with its limiting stochastic volatility model for the natural sampling frequencies. To get the equivalence, suitable frequencies of observations are required [Brown, Wang and Zhao (2003)].

Inference for continuously observed diffusion processes is well developed [e.g., Kutoyants (2004)]. As the diffusion coefficient is identified from a continuous time observation, it is assumed to be known and inference concerns the drift coefficient. On the contrary, inference for discretely observed diffusions is more difficult as the transition densities are generally untractable. Statistical procedures based on the Euler scheme corresponding to the one-step discretization of the diffusion have been successfully carried over to the discretized diffusion observations. In parametric inference, we may quote Genon-Catalot (1990), Larédo (1990) for small diffusion coefficient, Kessler (1997) for positive recurrent diffusions and for nonparametric inference, Hoffmann (1999), Comte, Genon-Catalot and Rozenholc (2007). Therefore, a natural issue for understanding these results is to prove the equivalence of the discretized observation of a diffusion and the corresponding Euler scheme experiment. Such a result has been proved by Milstein and Nussbaum (1998) for diffusions with small-known constant diffusion coefficient and by Dalalyan and Reiss (2006) for positive recurrent diffusions with constant diffusion coefficient. Our aim here is to extend this result to the case of a nonconstant-known diffusion coefficient using random time changes which yield models with diffusion coefficient equal to 1. This provides a canonical way for solving the equivalence problem. The time changed experiment coming from the Euler scheme does not lead to an autonomous diffusion but to an Itô process with predictable drift which induces the main difficulties.

More precisely, we consider the diffusion process (ξ_t) given by

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

where $(W_t)_{t\geq 0}$ is a Brownian motion defined on a filtered probability space $(\Omega, \mathcal{A}, (\mathcal{A}_t)_{t\geq 0}, \mathbb{P})$, η is a real valued random variable, $\mathcal{A}_0$-measurable, $b(\cdot), \sigma(\cdot)$ are real-valued functions defined on $\mathbb{R}$. The diffusion coefficient $\sigma(\cdot)$ is a known nonconstant function. The drift function $b(\cdot)$ is unknown and belongs to a nonparametric class. The sample path of (ξ_t) is continuously observed on a time interval $[0, T]$. We also consider the discrete observation of (ξ_t) at the times $t_i = ih, i \leq n$ with $n = [T/h]$. For simplicity, we assume in what follows that T/h is an integer. The Euler scheme corresponding to (1), with sampling interval h is

$$Z_0 = \eta, \qquad Z_i = Z_{i-1} + hb(Z_{i-1}) + \sqrt{h}\sigma(Z_{i-1})\varepsilon_i, \tag{2}$$

where, for $i \geq 1$, $t_i = ih$ and $\varepsilon_i = (W_{t_i} - W_{t_{i-1}})/\sqrt{h}$. For performing the comparisons, we consider $(Z_0, Z_1, \ldots, Z_n)$ with $n = T/h$. We prove the asymptotic equivalences assuming that n tends to infinity with $h = h_n$ and $nh_n^2 = T^2/n$ tending to 0. This includes both cases $T = nh_n$ bounded and $T \to +\infty$. Note that, for inference in diffusion models from discrete observations, the constraint $nh_n^2 \to 0$ is the standard condition for Lipschitz drift functions [e.g., Kessler (1997), Dalalyan and Reiss (2006), Comte, Genon-Catalot and Rozenholc (2007)]. We can also observe that statistical procedures for estimating the drift generally do not use the knowledge of the diffusion coefficient which appears as a nuisance parameter. Carter (2007) did a noteworthy improvement in this direction: he proves the asymptotic equivalence of the regression experiment with unknown variances with an experiment having two components, the first containing information about the variance, the second containing information on the mean. An important open problem which has never been tackled concerns the similar result for diffusion processes with unknown diffusion coefficient $\sigma(\cdot)$.

The paper is organized as follows. Assumptions and main results are given in Section 2. Theorem 2.1 states the equivalence result of (1) and (2) and Corollary 2.1 states the equivalence of the discrete observation of the diffusion and its Euler scheme. The proof of Theorem 2.1 is developed in Section 3. We consider random time changes on the diffusion and on the Euler scheme leading to processes with diffusion coefficient equal to 1. First, the classical random time change on the diffusion which leads to an autonomous diffusion process with drift $f = b/\sigma^2$ and diffusion coefficient equal to 1 is recalled (Proposition 3.1). We prove the exact equivalence between the diffusion experiment (1) and the random time changed experiment (Proposition 3.2). For the Euler scheme, we build a continuous time accompanying experiment (Proposition 3.3). Then we introduce a random time change leading to a process with unit diffusion coefficient. This process characterized in Proposition 3.4 has a predictable path-dependent drift term. The exact equivalence between the corresponding experiment and the Euler scheme experiment is proved in Theorem 3.1. Finally, for $n \to \infty$, the asymptotic equivalence of the two randomly stopped experiments is proved (Proposition 3.5) under the condition $h = h_n \to 0$, $nh_n^2 \to 0$, thus completing the proof of Theorem 2.1. Concluding remarks and extensions are given in Section 4. Proofs are gathered in Section 5. Appendix contains a short recap on the Le Cam deficiency distance Δ and some useful auxiliary results.

2. Assumptions and main results. We assume that the diffusion coefficient $\sigma(\cdot)$ of (1) is known, belongs to $C^2(\mathbb{R})$ and satisfies:

(C) $\forall x \in \mathbb{R}, 0 < \sigma_0^2 \le \sigma^2(x) \le \sigma_1^2, |\sigma'(x)| + |\sigma''(x)| \le K_\sigma$.

The function $b(\cdot)$ is unknown and such that, for K a positive constant:

(H1) $b(\cdot) \in \mathcal{F}_K = \{b(\cdot) \in C^1(\mathbb{R})$ and for all $x \in \mathbb{R}, |b(x)| + |b'(x)| \le K\}$.

The constant K has to exist but may be unknown.

Condition (C) and assumption (H1) ensure that the stochastic differential equation (1) has a unique strong solution process $(\xi_t)_{t\ge 0}$. The assumptions on b, σ are rather strong but allow to shorten technical proofs. Note that (H1) and (C) include models with or without ergodicity properties. The distribution of the initial variable η of (1) may be known or unknown.

Let $C(\mathbb{R}^+, \mathbb{R})$ be the space of continuous real functions defined on $\mathbb{R}^+$, and denote by $(X_t, t \ge 0)$ the canonical process of $C(\mathbb{R}^+, \mathbb{R})$ given by $(X_t(x) = x(t), t \ge 0)$ for $x \in C(\mathbb{R}^+, \mathbb{R})$, $\mathcal{C}_t^{0,X} = \sigma(X_s, s \le t)$, $\mathcal{C}_t^X = \bigcap_{s>t} \mathcal{C}_s^{0,X}$ and $\mathcal{C}^X = \sigma(\mathcal{C}_t^X, t \ge 0)$. Denote by P_b the distribution of $(\xi_t, t \ge 0)$ defined by (1) on $(C(\mathbb{R}^+, \mathbb{R}), \mathcal{C}^X)$ and consider the experiment associated with the continuous observation of the diffusion

$$\mathcal{E}_0 = \big(C(\mathbb{R}^+, \mathbb{R}), \mathcal{C}^X, (P_b, b \in \mathcal{F}_K)\big).$$

If T is fixed or is a $(\mathcal{C}_t^X)$-stopping time, we define the restriction $P_b/_{\mathcal{C}_T^X}$ of P_b to the σ-field $\mathcal{C}_T^X$. The experiment associated with the continuous observation of (ξ_t) stopped at T is

$$\mathcal{E}_0^T = \big(C(\mathbb{R}^+, \mathbb{R}), \mathcal{C}_T^X, (P_b/_{\mathcal{C}_T^X}, b \in \mathcal{F}_K)\big). \tag{3}$$

Consider now the Euler scheme corresponding to (1), with sampling interval h, defined in (2). This experiment is an autoregression model but we have rather call it Euler scheme as it is associated with the one-step discretization of (1). Let $(\pi_i)_{i\geq 0}$ denote 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 $\mathcal{G}_n = \sigma(\pi_0, \pi_1, \dots, \pi_n)$, $\mathcal{G} = \sigma(\mathcal{G}_n, n \geq 0)$. We denote by Q_b^h the distribution of $(Z_i, i \geq 0)$ defined by (2) 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 stopped at N is

$$\mathcal{G}^{h,N} = \big(\mathbb{R}^{\mathbb{N}}, \mathcal{G}_N, (Q_b^h/_{\mathcal{G}_N}, b \in \mathcal{F}_K)\big). \tag{4}$$

We now state the main result.

THEOREM 2.1. *Assume* (H1)–(C). *For deterministic $N = n$, $h = h_n$, the sequences of experiments $(\mathcal{E}_0^{nh_n})$ and $(\mathcal{G}^{h_n,n})$ are asymptotically equivalent for the Le Cam distance Δ as $n \to \infty$, if $h_n \to 0$ and $nh_n^2 \to 0$: $\Delta(\mathcal{E}_0^{nh_n}, \mathcal{G}^{h_n,n}) \to 0$.*

An important consequence is the comparison of the experiment associated with the discrete observation $(\xi_{ih}, i \leq n)$ of the diffusion with sampling interval h and the experiment $\mathcal{G}^{h,n}$. Let P_b^h denote the distribution of $(\xi_{ih})_{i\geq 0}$ defined by equation (1) on $(\mathbb{R}^{\mathbb{N}}, \mathcal{B}(\mathbb{R}^{\mathbb{N}}))$. For N a $(\mathcal{G}_n)$-stopping time, let $P_b^h/_{\mathcal{G}_N}$ be the restriction of P_b^h to $\mathcal{G}_N$. The experiment associated with the discrete observations (ξ_{ih}) with sampling h stopped at N is

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

COROLLARY 2.1. *Assume* (H1)–(C). *For deterministic $N = n$, $h = h_n$, the sequences of experiments $(\mathcal{E}^{h_n,n})$ and $(\mathcal{G}^{h_n,n})$ are asymptotically equivalent for the Le Cam distance Δ as $n \to \infty$, if $h_n \to 0$ and $nh_n^2 \to 0$: $\Delta(\mathcal{E}^{h_n,n}, \mathcal{G}^{h_n,n}) \to 0$.*

Milstein and Nussbaum (1998), Dalalyan and Reiss (2006) proved that when $\sigma(\cdot)$ is constant and nh_n^2 tends to 0, the discrete observation $(\xi_{ih_n}, i \leq n)$ is an asymptotically sufficient statistic for $(\xi_t, t \leq nh_n)$, that is, $\Delta(\mathcal{E}_0^{nh_n}, \mathcal{E}^{h_n,n}) \to 0$. For nonconstant diffusion coefficient, the latter asymptotic sufficiency result can be deduced using the change of function $F(x) = \int_0^x du/\sigma(u)$. Therefore, applying Theorem 2.1 yields the corollary.

3. Random time changed experiments. To deal with the nonconstant diffusion coefficient $\sigma(\cdot)$, we define experiments obtained by random time changes. For this, set

$$f(x) = \frac{b(x)}{\sigma^2(x)}, \qquad L = \frac{K}{\sigma_0^2}\left(1 + 2\frac{K_\sigma \sigma_1}{\sigma_0^2}\right). \tag{5}$$

Under (H1)–(C), f is bounded and globally Lipschitz with constant L.

3.1. *Time change on the diffusion.* Define for $x \in C(\mathbb{R}^+, \mathbb{R})$, $t, u \geq 0$,

$$\rho_t(x) = \int_0^t \sigma^2(x(s))\, ds, \qquad \tau_u(x) = \inf\{t \geq 0, \rho_t(x) \geq u\}. \tag{6}$$

Since $\sigma(\cdot)$ is known, the functions ρ_t and τ_u are known as well. Therefore, one is allowed to use these functions in the construction of Markov kernels. By (C), $\rho_{+\infty}(x) = +\infty$, $\frac{u}{\sigma_1^2} \leq \tau_u(x) \leq \frac{u}{\sigma_0^2}$, $\rho_{\tau_u(x)}(x) = u$, $\tau_{\rho_t(x)}(x) = t$. Note that $\tau_u(X)$ is a stopping time with respect to the canonical filtration $(\mathcal{C}_s^X, s \geq 0)$. We introduce now a classical time changed process.

PROPOSITION 3.1. *Assume* (H1)–(C). *Let ξ be the solution of* (1) *and set* $(\zeta_u = \xi_{\tau_u(\xi)}, u \geq 0)$ *and* $(\mathcal{G}_u = \mathcal{A}_{\tau_u(\xi)}, u \geq 0)$. *Then*

$$d\zeta_u = f(\zeta_u)\, du + dB_u, \qquad \zeta_0 = \eta, \tag{7}$$

with (B_u) *Brownian motion w.r.t.* $(\mathcal{G}_u)$ *which satisfies the usual conditions.*

The proof relies on classical tools [e.g., Karatzas and Shreve (2000), Chapter 3, Section 4 and Chapter 5, Section 5] and implies $\delta(\mathcal{E}_0^{\tau_a(X)}, \widetilde{\mathcal{E}}_0^a) = 0$ (see Appendix). The main difficulty lies in studying the other deficiency. Denote by $\widetilde{P}_b$ the distribution of $(\zeta_u, u \geq 0)$ on $C(\mathbb{R}^+, \mathbb{R})$. We associate to the time changed process $(\zeta_u, u \geq 0)$ an experiment with sample space $C(\mathbb{R}^+, \mathbb{R})$. For sake of clarity, we use a distinct notation for the canonical process and filtration. Let $(Y_u, u \geq 0)$ be defined by $Y_u(y(\cdot)) = y(u)$ with $y(\cdot) \in C(\mathbb{R}^+, \mathbb{R})$, $(\mathcal{C}_u^Y, u \geq 0)$ be the associated right-continuous canonical filtration and $\mathcal{C}^Y = \sigma(\mathcal{C}_u^Y, u \geq 0)$. Set

$$\widetilde{\mathcal{E}}_0 = \bigl(C(\mathbb{R}^+, \mathbb{R}), \mathcal{C}^Y, (\widetilde{P}_b, b \in \mathcal{F}_K)\bigr).$$

For $A > 0$ a $(\mathcal{C}_u^Y)$-stopping time, define the experiment

$$\widetilde{\mathcal{E}}_0^A = \bigl(C(\mathbb{R}^+, \mathbb{R}), \mathcal{C}_A^Y, (\widetilde{P}_b/_{\mathcal{C}_A^Y}, b \in \mathcal{F}_K)\bigr).$$

Define for $y \in C(\mathbb{R}^+, \mathbb{R})$,

$$T_u(y) = \int_0^u \frac{dv}{\sigma^2(y(v))}, \qquad A_t(y) = \inf\{u \geq 0, T_u(y) \geq t\} = T_{\cdot}(y)^{-1}(t). \tag{8}$$

Thus, for all $t \geq 0$, $A_t(Y)$ is a $(\mathcal{C}_u^Y)$-stopping time.

PROPOSITION 3.2. *Assume* (H1)–(C). *If* $x = (x(t), t \geq 0)$, $(y(u) = x(\tau_u(x)), u \geq 0)$, *then* $A_t(y) = \rho_t(x)$, $T_u(y) = \tau_u(x)$. *For* a, T *deterministic* $\Delta(\mathcal{E}_0^T, \widetilde{\mathcal{E}}_0^{A_T(Y)}) = 0$ *and* $\Delta(\mathcal{E}_0^{\tau_a(X)}, \widetilde{\mathcal{E}}_0^a) = 0$.

The experiments $\mathcal{E}_0$ and $\widetilde{\mathcal{E}}_0$ are linked by the mapping $(x(t), t \geq 0) \to (y(u) = x(\tau_u(x)), u \geq 0)$. For the stopped experiments, noting that $\{u, \tau_u(x) \leq T\} = \{u, u \leq A_T(y)\}$, the previous mapping links $\mathcal{E}_0^T$ and $\widetilde{\mathcal{E}}_0^{A_T(Y)}$. Similarly, the experiments $\widetilde{\mathcal{E}}_0$ and $\mathcal{E}_0$ are linked by the mapping $(y(u), u \geq 0) \to (x(t) = y(A_t(y)), t \geq 0)$ and, for stopped experiments, noting that $\{t, A_t(y) \leq a\} = \{t, t \leq \tau_a(x)\}$, this mapping links $\widetilde{\mathcal{E}}_0^a$ and $\mathcal{E}_0^{\tau_a(X)}$.

3.2. *Time change on the Euler scheme*. As the discrete Euler scheme experiment (4) has not the same sample space as the diffusion experiment (3), an essential tool is to use the accompanying experiment of (4) which is the continuous-time Euler scheme. Given a path $x = x(\cdot) \in \mathcal{C}(\mathbb{R}^+, \mathbb{R})$ and a sampling scheme $t_i = ih, i \geq 1$, we define the diffusion-type process $\bar{\xi}_t$,

$$d\bar{\xi}_t = \bar{b}_h(t, \bar{\xi})\, dt + \bar{\sigma}_h(t, \bar{\xi})\, dW_t, \qquad \bar{\xi}_0 = \eta, \tag{9}$$

with

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

Let Q_b denote the distribution of $(\bar{\xi}_t, t \geq 0)$ on $(C(\mathbb{R}^+, \mathbb{R}), \mathcal{C}^X)$ and, for T a $(\mathcal{C}_t^X)$-stopping time, $Q_b/_{\mathcal{C}_T^X}$ the restriction of Q_b to $\mathcal{C}_T^X$. Set

$$\mathcal{G}_0^T = \big(C\big(\mathbb{R}^+, \mathbb{R}\big), \mathcal{C}_T^X, (Q_b/_{\mathcal{C}_T^X}, b \in \mathcal{F}_K)\big). \tag{10}$$

PROPOSITION 3.3. *For* $h > 0$, N *a* $(\mathcal{G}_n)$*-stopping time, the Le Cam distance between* $\mathcal{G}^{h,N}$ *and* $\mathcal{G}_0^{Nh}$ [(4), (10)] *is equal to* 0: $\Delta(\mathcal{G}^{h,N}, \mathcal{G}_0^{Nh}) = 0$.

Let us define a time changed process associated with the continuous Euler scheme $(\bar{\xi}_t)$. The study of this time changed process is more difficult because the drift term and the diffusion coefficient of the continuous-time Euler scheme are time and path dependent. Let

$$\bar{\rho}_t(x) = \int_0^t \bar{\sigma}_h^2(s, x)\, ds, \qquad \bar{\tau}_u(x) = \inf\{t \geq 0, \bar{\rho}_t(x) \geq u\}.$$

Analogously, $\bar{\tau}_u(X)$ is a stopping time of the canonical filtration $\mathcal{C}^X$. With the convention $\sum_{j=0}^{i-1} = 0$ for $i = 0$, we have, for $i \geq 0$ and $t_i < t \leq t_{i+1}$,

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

Hence, $(\bar{\rho}_t(x), t \geq 0)$ is continuous, increasing on $\mathbb{R}^+$ and maps $(t_i, t_{i+1}]$ on $(\bar{\rho}_{t_i}(x), \bar{\rho}_{t_{i+1}}(x)]$. By (C), $\bar{\rho}_{+\infty}(x) = +\infty$, $u/\sigma_1^2 \leq \bar{\tau}_u(x) \leq (u/\sigma_0^2) + \Delta$, and $\{t \to \bar{\rho}_t(x)\}$, $\{u \to \bar{\tau}_u(x)\}$ are inverse. In particular, for all i, x, $t_i = \bar{\tau}_{\bar{\rho}_{t_i}(x)}(x)$. For $\bar{\xi}$ solution of (9), set $(\overline{\mathcal{G}}_u = \mathcal{A}_{\bar{\tau}_u(\bar{\xi})})$, and define the process

$$(\bar{\zeta}_u = \bar{\xi}_{\bar{\tau}_u(\bar{\xi})}, u \geq 0), \tag{11}$$

which is adapted to the filtration $(\overline{\mathcal{G}}_u)$ which satisfies the usual conditions. Denote by $\widetilde{Q}_b$ the distribution of $(\bar{\zeta}_u)$.

PROPOSITION 3.4. *The process* $(\bar{\zeta}_u)$ *defined in* (11) *has unit diffusion coefficient and drift term given by* [*see* (5)]:

$$\bar{f}(v) = \sum_{i \geq 0} f(\bar{\zeta}_{\bar{\rho}_{t_i}(\bar{\xi})}) 1_{(\bar{\rho}_{t_i}(\bar{\xi}), \bar{\rho}_{t_{i+1}}(\bar{\xi})]}(v), \tag{12}$$

where $(\bar{\rho}_{t_i}(\bar{\xi}))$ *are* $(\overline{\mathcal{G}}_u)$*-stopping times and so,* $\bar{f}(v)$ *is predictable w.r.t.* $(\overline{\mathcal{G}}_u)$.

We associate to the time changed process $(\bar{\zeta}_u, u \geq 0)$ an experiment with sample space $C(\mathbb{R}^+, \mathbb{R})$ and canonical process $(Y_u, u \geq 0)$ with associated canonical filtration $(\mathcal{C}_u^Y, u \geq 0)$. Set

$$\widetilde{\mathcal{G}}_0 = \big(C\big(\mathbb{R}^+, \mathbb{R}\big), \big(\mathcal{C}_u^Y\big), (\widetilde{Q}_b, b \in \mathcal{F}_K)\big).$$

For $A > 0$ a $(\mathcal{C}_u^Y)$-stopping time, define the experiment

$$\widetilde{\mathcal{G}}_0^A = \big(C\big(\mathbb{R}^+, \mathbb{R}\big), \mathcal{C}_A^Y, (\widetilde{Q}_b/_{\mathcal{C}_A^Y}, b \in \mathcal{F}_K)\big).$$

For $y \in C(\mathbb{R}^+, \mathbb{R})$, set $\overline{A}_0(y) = 0$, for $t \in (t_{i-1}, t_i]$,

$$\overline{A}_t(y) = \overline{A}_{t_{i-1}}(y) + \sigma^2\big(y\big(\overline{A}_{t_{i-1}}(y)\big)\big)(t - t_{i-1}). \tag{13}$$

Let $\overline{T}_u(y) = \inf\{t, \overline{A}_t(y) \geq u\}$.

LEMMA 3.1. *Set* $(y(u) = x(\bar{\tau}_u(x)), u \geq 0)$. *Then,* $\overline{A}_t(y) = \bar{\rho}_t(x)$ *and* $\overline{T}_u(y) = \bar{\tau}_u(x)$. *Consequently, for all* $t \geq 0$, $\overline{A}_t(Y)$ *is a* $(\mathcal{C}_u^Y)$*-stopping time.*

Thus, the drift term in Proposition 3.4 is $\bar{f}(v) = \bar{f}(v, \bar{\zeta})$ with

$$\bar{f}(v, y) = \sum_{i \geq 1} f\big(y\big(\overline{A}_{t_{i-1}}(y)\big)\big) 1_{(\overline{A}_{t_{i-1}}(y), \overline{A}_{t_i}(y)]}(v).$$

The following result parallel of Proposition 3.2 contains the main difficulties.

THEOREM 3.1. *Assume* (H1) *and* (C). *For deterministic* $a > 0$ *and* $T = nh$, $\Delta(\mathcal{G}_0^T, \widetilde{\mathcal{G}}_0^{\overline{A}_T(Y)}) = 0$ *and* $\Delta(\mathcal{G}_0^{\bar{\tau}_a(X)}, \widetilde{\mathcal{G}}_0^a) = 0$.

The proof uses the following devices. If x and y are linked by $(x(t), t \geq 0) \to (y(u) = x(\bar{\tau}_u(x)), u \geq 0)$, then $\{u, \bar{\tau}_u(x) \leq T\} = \{u, u \leq \overline{A}_T(y)\}$. Similarly, for $(x(t) = y(\overline{A}_t(y)), t \geq 0)$, then $\{t, \overline{A}_t(y) \leq a\} = \{t, t \leq \bar{\tau}_a(x)\}$.

3.3. *Asymptotic equivalence of randomly stopped experiments*. At this point, the triangle inequality implies that, for fixed T, n, h such that $T = nh$,

$$\Delta(\mathcal{E}_0^T, \mathcal{G}^{h,n}) \le \Delta(\mathcal{E}_0^T, \mathcal{G}_0^T) \le \Delta(\widetilde{\mathcal{E}}_0^{A_T(Y)}, \widetilde{\mathcal{G}}_0^{\overline{A}_T(Y)}).$$

We now introduce the asymptotic framework. Set $T_n = T = nh_n$ and consider the stopping times

$$A_n = A_{nh_n}(Y), \qquad \overline{A}_n = \overline{A}_{nh_n}(Y), \qquad S_n = \overline{A}_n \wedge A_n. \tag{14}$$

It remains to study $\Delta(\widetilde{\mathcal{E}}_0^{A_n}, \widetilde{\mathcal{G}}_0^{\overline{A}_n})$. These two experiments have the same sample space but are observed up to distinct stopping times.

LEMMA 3.2. *Assume* (H1) *and* (C). *There exists a constant D depending only on $K, K_\sigma, \sigma_0, \sigma_1$ such that $E_{\widetilde{P}_b}|A_n - \overline{A}_n| \le Dnh_n^2$.*

Using (14), the triangle inequality yields

$$\Delta(\widetilde{\mathcal{E}}_0^{A_n}, \widetilde{\mathcal{G}}_0^{\overline{A}_n}) \le \Delta(\widetilde{\mathcal{E}}_0^{A_n}, \widetilde{\mathcal{E}}_0^{S_n}) + \Delta(\widetilde{\mathcal{E}}_0^{S_n}, \widetilde{\mathcal{E}}_0^{\overline{A}_n}) + \Delta(\widetilde{\mathcal{E}}_0^{\overline{A}_n}, \widetilde{\mathcal{G}}_0^{\overline{A}_n}). \tag{15}$$

Therefore, we have to study the Le Cam distances, respectively, for the same experiment observed up to two distinct times and for two experiments observed up to the random time $\overline{A}_n$. The following holds.

PROPOSITION 3.5. *Assume* (H1) *and* (C). *There exist constants K_1, K_2 depending only on $K, K_\sigma, \sigma_0, \sigma_1$ such that*

$$\Delta(\widetilde{\mathcal{E}}_0^{A_n}, \widetilde{\mathcal{E}}_0^{S_n}) + \Delta(\widetilde{\mathcal{E}}_0^{\overline{A}_n}, \widetilde{\mathcal{E}}_0^{S_n}) \le K_1(nh_n^2)^{1/2}, \tag{16}$$

$$\Delta(\widetilde{\mathcal{E}}_0^{\overline{A}_n}, \widetilde{\mathcal{G}}_0^{\overline{A}_n}) \le K_2(nh_n^2)^{1/2}. \tag{17}$$

Therefore, if nh_n^2 goes to 0 *as n tends to infinity*, $\Delta(\widetilde{\mathcal{E}}_0^{A_n}, \widetilde{\mathcal{G}}_0^{\overline{A}_n}) \to 0$.

Joining Propositions 3.2, 3.3, Theorem 3.1 and Proposition 3.5 completes the proof of Theorem 2.1.

4. Concluding remarks. In this paper, we have obtained the asymptotic equivalence of the continuous time diffusion (1) observed on the time interval $[0, T]$ and (2) the corresponding Euler scheme with sampling interval h and $T = nh$ in the case of a nonconstant diffusion coefficient. The discrete Euler scheme model is often used in applications instead of the diffusion itself. It is broadly accepted as an appropriate substitute to the diffusion because of its weak convergence to the diffusion. The equivalence result obtained here was known for a constant diffusion coefficient. Our contribution is the extension to the case of a nonconstant diffusion coefficient by means of random time changed experiments.

The constant K in the definition of the class $\mathcal{F}_K$ is not used for building the Markov kernels contrary to the diffusion coefficient $\sigma(\cdot)$. The asymptotic framework is $n \to +\infty$, $h = h_n \to 0$ and $nh_n^2 = T^2/n \to 0$. In our result, $T = nh_n$ may be fixed or tend to infinity. We have no assumption concerning the existence of a stationary regime for the diffusion or for the Euler scheme. This comes from the assumption that b is bounded which allows to substantially shorten proofs. For unbounded drift functions, the two cases "T bounded" and "T tending to infinity" have to be distinguished. In the latter case, the diffusion model must be positive recurrent with moment assumptions on the stationary distribution.

Compared with other equivalence results, the regularity assumption for b might seem too strong. However, a classical assumption for existence and uniqueness of a strong solution to (1) is b locally Lipschitz with linear growth. Generally, authors assume that b is C^1 with linear growth. Dalalyan and Reiss (2006) consider a special class of drift functions: b is locally Lipschitz, *known* outside a compact interval I, and Hölder with exponent $\alpha \in (0, 1)$ *inside* I. They obtain a global asymptotic equivalence of a stationary diffusion and a mixed Gaussian experiment as $T \to +\infty$.

An interesting issue concerns multidimensional diffusions and their associated Euler scheme. If the diffusion matrix is constant, the problem is solved [Dalalyan and Reiss (2007)]. Otherwise, consider a d-dimensional process $d\xi_t = b(\xi_t)\,dt + \Sigma(\xi_t)\,dW_t$, where $b:\mathbb{R}^d \to \mathbb{R}^d$, $\Sigma:\mathbb{R}^d \to \mathbb{R}^d \otimes \mathbb{R}^d$, (W_t) is a d-dimensional Brownian motion. If $\Sigma(x)$ has the special form $\Sigma(x) = \sigma(x)P(x)$ where $\sigma:\mathbb{R}^d \to (0, +\infty)$ and the $d \times d$-matrix $P(x)$ satisfies, for all $x \in \mathbb{R}^d$, $P(x)P(x)^t = I$, the equivalence result is obtained similarly. Indeed, setting $\rho_t(x) = \int_0^t \sigma^2(x(s))\,ds$ with inverse $\tau_u(x)$, the time changed process $\zeta_u = \xi_{\tau_u(\xi)}$ has a diffusion matrix equal to the identity matrix and a drift equal to $b(u)/\sigma^2(u)$. As for the continuous Euler scheme, we can define analogously $\bar{\rho}_t(x)$ and $\bar{\tau}_u(x)$.

Statistical procedures for estimating the drift generally do not use the knowledge of the diffusion coefficient which appears as a nuisance parameter. It is an open question to know whether the equivalence proved here holds when the diffusion coefficient is unknown.

5. Proofs.

PROOF OF PROPOSITION 3.2. By Proposition 3.1, $\widetilde{\mathcal{E}}_0^a$ is the image of $\mathcal{E}_0^{\tau_a(X)}$ by the measurable mapping $(x(t), t \in [0, \tau_a(x)]) \to (y(u) = x(\tau_u(x)), u \in [0, a])$, which implies $\delta(\mathcal{E}_0^{\tau_a(X)}, \widetilde{\mathcal{E}}_0^a) = 0$.

Now, we look at $\mathcal{E}_0^T$. As $T = \tau_{\rho_T(x)}(x)$ [(6)], the image of $(x(t), t \le T)$ is $(y(u) = x(\tau_u(x)), u \le \rho_T(x))$. We must express $\rho_T(x)$ in terms of the path y and prove that $\rho_T(x) = A_T(y)$. Since $(\rho_T(X) \ge u) = (\tau_u(X) \le T)$, $\rho_T(X)$ is a stopping time of $(\mathcal{C}^X_{\tau_u(X)}, u \ge 0)$. The continuity of $u \to \tau_u(X)$ implies

$$\sigma(X_{\tau_v(X)}, v \le u) = \sigma\bigl(X_s, s \le \tau_u(X)\bigr).$$

Thus, $\rho_T(X)$ is a stopping time of $\sigma(Y_v, v \le u)$ with $Y_v = X_{\tau_v(X)}$. Observe that, using the change of variable $\tau_v(X) = s \Leftrightarrow v = \rho_s(X)$, we have

$$T_u(Y) = \int_0^u \bigl(dv/\sigma^2(Y_v)\bigr)\, dv = \int_0^{\tau_u(X)} ds = \tau_u(X).$$

This implies $\rho_T(X) = A_T(Y)$ which yields $\delta(\mathcal{E}_0^T, \widetilde{\mathcal{E}}_0^{A_T(Y)}) = 0$.

Consider now the reverse operation. Let $(B_u, u \ge 0)$ be a standard Brownian motion with respect to a filtration $(\mathcal{G}_u)$ satisfying the usual conditions and ζ_0 be a $\mathcal{G}_0$-measurable random variable. We define, for $u \ge 0$,

$$\zeta_u = \zeta_0 + \int_0^u \frac{b(\zeta_v)}{\sigma^2(\zeta_v)}\, dv + B_u \quad \text{and} \quad T_u = T_u(\zeta) = \int_0^u \frac{dv}{\sigma^2(\zeta_v)}. \tag{18}$$

Clearly, the mapping $u \to T_u$ is a bijection from $[0, a]$ onto $[0, T_a]$ with inverse $t \to T^{-1}(t) := A_t(\zeta)$. Therefore, we can define, for $0 \le t \le T_a$, the process $\xi_t = \zeta_{A_t(\zeta)}$. The change of variable $v = A_s(\zeta) \Leftrightarrow s = T_v$ yields that $ds = dv/\sigma^2(\zeta_v) = dv/\sigma^2(\zeta_{A_s(\zeta)}) = dv/\sigma^2(\xi_s)$ and equation (18) becomes

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

Now, $(M_t = B_{A_t(\zeta)})$ is a martingale w.r.t. $(\mathcal{G}_{A_t(\zeta)})$ satisfying

$$\langle M \rangle_t = A_t(\zeta) = \int_0^{A_t(\zeta)} ds = \int_0^t \sigma^2(\zeta_{A_s(\zeta)})\, ds = \int_0^t \sigma^2(\xi_s)\, ds.$$

Hence, $\tau_u(\xi) = A_.(\zeta)^{-1}(u) = T_u$ and (ξ_t) has distribution P_b. As $(A_t(\zeta))$ is continuous, $(\mathcal{G}_{A_t(\zeta)})$ inherits the usual conditions from $(\mathcal{G}_u)$.

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

$$\Psi : y \in C(\mathbb{R}^+, \mathbb{R}) \to \bigl(x := y(A_t(y)), t \ge 0\bigr) \in C(\mathbb{R}^+, \mathbb{R}).$$

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, $(X_t, t \le \tau_a(X))$ is the image of $(Y(u), u \le a)$ by the measurable mapping Ψ. Hence, $\delta(\widetilde{\mathcal{E}}_0^a, \mathcal{E}_0^{\tau_a(X)}) = 0$. Analogously, $(X_t, t \le T)$ is the image of $(Y(u), u \le A_T(Y))$ which implies $\delta(\widetilde{\mathcal{E}}_0^{A_T(Y)}, \mathcal{E}_0^T)$. $\square$

PROOF OF PROPOSITION 3.3. This proof relies on Lemma 5.1 below. Define the linear interpolation between the points $((t_i, Z_i), i \ge 0)$:

$$y(t) = Z_i + \frac{t - t_i}{t_{i+1} - t_i}(Z_{i+1} - Z_i) \qquad \text{if } t \in [t_i, t_{i+1}] \text{ and } i \ge 0. \tag{19}$$

LEMMA 5.1. *The solution $(\bar{\xi}_t)$ of (9) satisfies $(\bar{\xi}_{t_i}, i \ge 0) = (Z_i, i \ge 0)$ where $(Z_i, i \ge 0)$ is the discrete Euler scheme (2). Moreover,*

$$\bar{\xi}_t = y(t) + \sigma(Z_i) B_i(t) \qquad \textit{if } t \in [t_i, t_{i+1}] \textit{ and } i \ge 0, \tag{20}$$

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* $(\bar{\xi}_t)$ *is adapted to* $(\mathcal{A}_t)$, $((B_i(t), t \in [t_i, t_{i+1}]), i \geq 0)$ *are independent Brownian bridges and the sequence* $((B_i(t), t \in [t_i, t_{i+1}]), i \geq 0)$ *is independent of* $(Z_j, j \geq 0)$.

This is a classical result obtained with standard tools. We may now complete the proof of Proposition 3.3. Since $(Z_i, i \geq 0)$ is the image of $(\bar{\xi}_t, t \geq 0)$ by the mapping $x(\cdot) \to (x(t_i), i \geq 0)$, $\delta(\mathcal{G}_0^{Nh}, \mathcal{G}^{h,N}) = 0$.

Consider, for $\omega \in \Omega$, the application $\Phi_{\sigma,h} = \Phi : \mathbb{R}^{\mathbb{N}} \to \mathcal{C}(\mathbb{R}^+, \mathbb{R})$ defined by $(x_i, i \geq 0) \to x(\cdot)$ with $x(t) = x_{i-1} + \frac{t-t_{i-1}}{t_i - t_{i-1}}(x_i - x_{i-1}) + \sigma(x_{i-1})B_{i-1}(t, \omega)$ for $t \in [t_{i-1}, t_i]$. As σ is known, Φ is a randomization and, by Lemma 5.1, $\mathcal{G}_0^{Nh}$ is the image by Φ of $\mathcal{G}^{h,N}$. Hence, $\delta(\mathcal{G}^{h,N}, \mathcal{G}_0^{Nh}) = 0$. $\square$

PROOF OF PROPOSITION 3.4. By definition of $(\bar{\zeta}_u)$, we have

$$\bar{\zeta}_u = \bar{\xi}_0 + \int_0^{\bar{\tau}_u(\bar{\xi})} \sum_{i \geq 0} b(\bar{\xi}_{t_i}) 1_{t_i < s \leq t_{i+1}} \, ds + \overline{B}_u, \tag{21}$$

where $\overline{B}_u = \int_0^{\bar{\tau}_u(\bar{\xi})} \sum_{i \geq 0} \sigma(\bar{\xi}_{t_i}) 1_{t_i < s \leq t_{i+1}} \, dW_s$ is a martingale w.r.t. $\overline{\mathcal{G}}_u = \mathcal{A}_{\bar{\tau}_u(\bar{\xi})}$ with quadratic variations $\langle \overline{B} \rangle_u = \int_0^{\bar{\tau}_u(\bar{\xi})} \sum_{i \geq 0} \sigma^2(\bar{\xi}_{t_i}) 1_{t_i < s \leq t_{i+1}} \, ds = u$. Therefore, $(\overline{B}_u)$ is a Brownian motion with respect to $(\overline{\mathcal{G}}_u)$.

In the integral of (21), the change of variable $s = \bar{\tau}_v(\bar{\xi}) \Leftrightarrow v = \bar{\rho}_s(\bar{\xi})$ yields, noting that $dv = \sigma^2(\bar{\xi}_{t_i})\, ds$ for $v \in (\bar{\rho}_{t_i}(\bar{\xi}), \bar{\rho}_{t_{i+1}}(\bar{\xi})]$, and that $t_i = \bar{\tau}_{\bar{\rho}_{t_i}(x)}(x)$,

$$\bar{\zeta}_u = \bar{\xi}_0 + \int_0^u \sum_{i \geq 0} \frac{b(\bar{\xi}_{t_i})}{\sigma^2(\bar{\xi}_{t_i})} 1_{\bar{\rho}_{t_i}(\bar{\xi}) < v \leq \bar{\rho}_{t_{i+1}}(\bar{\xi})} \, dv + \overline{B}_u, \tag{22}$$

where $\bar{\xi}_{t_i} = \bar{\zeta}_{\bar{\rho}_{t_i}(\bar{\xi})} = Z_i$ is the discrete Euler scheme (Lemma 5.1).

Thus, $(\overline{Y}_u)$ defined in (11) is a process with diffusion coefficient equal to 1 and drift term $\bar{f}(v)$. We now check that $\bar{f}(v)$ is predictable w.r.t. $(\overline{\mathcal{G}}_u)$, that is, $\forall i$, $\bar{\rho}_{t_i}(\bar{\xi})$ is a $(\overline{\mathcal{G}}_u)$-stopping time and $\bar{\zeta}_{\bar{\rho}_{t_i}(\bar{\xi})}$ is $\overline{\mathcal{G}}_{\bar{\rho}_{t_i}(\bar{\xi})}$-measurable. Noting that $(\bar{\rho}_{t_i}(\bar{\xi}) \leq u) = (\bar{\tau}_u(\bar{\xi}) \geq t_i)$ belongs to $\overline{\mathcal{G}}_u = \mathcal{A}_{\bar{\tau}_u(\bar{\xi})}$ yields that $\bar{\rho}_{t_i}(\bar{\xi})$ is a $(\overline{\mathcal{G}}_u)$-stopping time. We know that $\bar{\zeta}_{\bar{\rho}_{t_i}(\bar{\xi})} = \bar{\xi}_{t_i}$ is $\mathcal{A}_{t_i}$-measurable, which achieves the proof since $\mathcal{A}_{t_i} = \overline{\mathcal{G}}_{\bar{\rho}_{t_i}(\bar{\xi})}$. $\square$

PROOF OF LEMMA 3.1. The relation $(y(u) = x(\bar{\tau}_u(x))$ is equivalent to $(y(\bar{\rho}_t(x)) = x(t))$. First, note that $\overline{A}_{t_1}(y) = \sigma^2(y(0))t_1 = \sigma^2(x(0))t_1 = \bar{\rho}_{t_1}(x)$. By induction, assume that $\overline{A}_{t_j}(y) = \bar{\rho}_{t_j}(x)$ for $j \leq i-1$. Then

$$\begin{aligned}\overline{A}_{t_i}(y) &= \bar{\rho}_{t_{i-1}}(x) + \sigma^2\big(y(\bar{\rho}_{t_{i-1}}(x))\big)(t_i - t_{i-1}) \\ &= \bar{\rho}_{t_{i-1}}(x) + \sigma^2\big(x(t_{i-1})\big)(t_i - t_{i-1}) = \bar{\rho}_{t_i}(x).\end{aligned}$$

Thus, the two inverse functions coincide: $\overline{T}_u(y) = \bar{\tau}_u(x)$. As above, we deduce that $A_t(y)$ is a stopping time w.r.t. $(\mathcal{C}_u^Y)$ with $Y_u = X_{\bar{\tau}_u(X)}$. $\square$

PROOF OF THEOREM 3.1. The proof is divided in several steps.

First, as $\widetilde{\mathcal{G}}_0^a$ is the image of $\mathcal{G}_0^{\bar{\tau}_a(X)}$ by the measurable mapping $(x(t), t \le \bar{\tau}_a(x)) \to (y(u) = x(\bar{\tau}_u(x)), u \in [0, a])$, $\delta(\mathcal{G}_0^{\bar{\tau}_a(X)}, \widetilde{\mathcal{G}}_0^a) = 0$.

Now consider $\mathcal{G}_0^T$. We have $T = \bar{\tau}_{\bar{\rho}_T(x)}(x)$. Hence, the image of $(x(t), t \le T)$ is $(y(u) = x(\bar{\tau}_u(x)), u \le \bar{\rho}_T(x) = \overline{A}_T(y))$ according to Lemma 3.1. This proves that $\delta(\mathcal{G}_0^T, \widetilde{\mathcal{G}}_0^{\overline{A}_T(y)}) = 0$.

Let us study the other deficiencies. We first construct a process $(\bar{\zeta}_u)$ with distribution $\widetilde{Q}_b$ (step 1), then a process $(\bar{\xi}_t)$ with distribution Q_b obtained from $(\bar{\zeta}_u)$ by the mapping $(y(u), u \ge 0) \to (y(\overline{A}_t(y)), t \ge 0)$ (step 2).

Step 1. Let $(\overline{B}_u)$ be a Brownian motion w.r.t. a filtration $(\overline{\mathcal{G}}_u)$ satisfying the usual conditions. Assume that $\bar{\zeta}_0$ is $\overline{\mathcal{G}}_0$-measurable. Then we define recursively a sequence of random times (T_i) and a continuous process $(\bar{\zeta}_u)$. First, set $T_0 = 0$, then

$$T_1 = T_1(\bar{\zeta}) = \sigma^2(\bar{\zeta}_0)t_1, \qquad \bar{\zeta}_u = \bar{\zeta}_0 + f(\bar{\zeta}_0)u + \overline{B}_u \qquad \text{for } 0 < u \le T_1,$$

$$T_i = T_i(\bar{\zeta}) = T_{i-1} + \sigma^2(\bar{\zeta}_{T_{i-1}})(t_i - t_{i-1}), \tag{23}$$

$$\bar{\zeta}_u = \bar{\zeta}_{T_{i-1}} + f(\bar{\zeta}_{T_{i-1}})(u - T_{i-1}) + \overline{B}_u - \overline{B}_{T_{i-1}} \qquad \text{for } T_{i-1} < u \le T_i. \tag{24}$$

Note that $T_i = \overline{A}_{t_i}(\zeta)$ [see (13)].

LEMMA 5.2. *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 $(\bar{\zeta}_u)$ defined in* (23), (24) *is a diffusion-type process adapted to $(\overline{\mathcal{G}}_u)$ with diffusion coefficient equal to* 1 *and drift coefficient*

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

where $(T_i(y) = \overline{A}_{t_i}(y), i \ge 0)$ are recursively defined as in (13) *using $y(\cdot)$ and $f = b/\sigma^2$* [*see* (5)]. *Hence, the process $(\bar{\zeta}_u)$ has distribution $\widetilde{Q}_b$.*

PROOF. First, T_1 is $\overline{\mathcal{G}}_0$-measurable, 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, $\bar{\zeta}_u = \bar{\zeta}_0 + f(\bar{\zeta}_0)u + \overline{B}_u$ is $\overline{\mathcal{G}}_u$-measurable. Thus, T_1 and $\bar{\zeta}_{T_1}$ 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, T_j is a $(\overline{\mathcal{G}}_u)$-stopping time, and $(\bar{\zeta}_u, u \le T_i)$ is $\overline{\mathcal{G}}_u$-measurable. Now, for $u > T_i$, $\bar{\zeta}_u = \bar{\zeta}_{T_i} + f(\bar{\zeta}_{T_i})(u - T_i) + \overline{B}_u - \overline{B}_{T_i}$ defined by (24) is $\overline{\mathcal{G}}_u$-measurable. As $T_{i+1} =$

$T_i + \sigma^2(\bar{\zeta}_{T_i})(t_{i+1} - t_i)$, the induction assumption yields that T_{i+1} is $\overline{\mathcal{G}}_{T_i}$-measurable and, since $T_i < T_{i+1}$ by (C),

$$\forall v \geq u \qquad \{T_{i+1} \leq u\} = \{T_{i+1} \leq u\} \cap \{T_i \leq v\} \in \overline{\mathcal{G}}_v.$$

This implies that $\{T_{i+1} \leq u\} = \{T_{i+1} \leq u\} \cap \bigcap_{v>u}\{T_i \leq v\} \in \bigcap_{v>u} \overline{\mathcal{G}}_v = \overline{\mathcal{G}}_u$ which proves that T_{i+1} is a $(\overline{\mathcal{G}}_u)$-stopping time. Thus, T_{i+1} and $\bar{\zeta}_{T_{i+1}}$ are $\overline{\mathcal{G}}_{T_{i+1}}$-measurable. The proof of Lemma 5.2 is now complete. □

Step 2. Let us study the distribution of $\bar{\xi}_t$ defined as

$$\bar{\xi}_t = \bar{\zeta}_{\overline{A}_t(\bar{\zeta})}. \tag{25}$$

By Lemma 5.2, $\overline{A}_{t_i}(\bar{\zeta}) = T_i$ is a $(\overline{\mathcal{G}}_u)$-stopping time. For $t_i \leq t \leq t_{i+1}$, $\overline{A}_t(\bar{\zeta}) = T_i + (t - t_i)\sigma^2(\bar{\zeta}_{T_i})$ is $\overline{\mathcal{G}}_{T_i}$-measurable, so

$$\forall v > u \qquad \{\overline{A}_t(\bar{\zeta}) \leq u\} = \{\overline{A}_t(\bar{\zeta}) \leq u\} \cap \{T_i \leq v\} \in \overline{\mathcal{G}}_v.$$

Hence, $\{\overline{A}_t(\bar{\zeta}) \leq u\} = \{\overline{A}_t(\bar{\zeta}) \leq u\} \cap \bigcap_{v>u}\{T_i \leq v\} \in \bigcap_{v>u} \overline{\mathcal{G}}_v = \overline{\mathcal{G}}_u$ which proves that $\overline{A}_t(\bar{\zeta})$ is a $(\overline{\mathcal{G}}_u)$-stopping time.

Thus, we can define the filtration $(\overline{\mathcal{A}}_t := \overline{\mathcal{G}}_{\overline{A}_t(\bar{\zeta})})$ to which $(\bar{\xi}_t)$ is adapted.

LEMMA 5.3. *The sequence* $(\bar{\xi}_{t_i} = \bar{\zeta}_{T_i}, i \geq 0)$, *with* $(\bar{\zeta}_u)$ *defined by* (23)–(24), $(\bar{\xi}_t)$ *in* (25), *has the distribution of the discrete Euler scheme* (2).

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

$$\left(\overline{B}^{(i)}_v = \overline{B}_{T_i+v} - \overline{B}_{T_i}, v \geq 0\right) \tag{26}$$

is a Brownian motion independent of $\overline{\mathcal{G}}_{T_i} = \mathcal{A}_{t_i}$, adapted to $(\overline{\mathcal{G}}_{T_i+v})$. As $\bar{\xi}_{t_i} = \bar{\zeta}_{T_i}$ is $\overline{\mathcal{G}}_{T_i}$-measurable, this r.v. is independent of $(\overline{B}^{(i)}_v, v \geq 0)$. Define

$$\varepsilon_{i+1} = \frac{\overline{B}_{T_{i+1}} - \overline{B}_{T_i}}{\sqrt{T_{i+1} - T_i}} = \frac{\overline{B}^{(i)}_{\sigma^2(\overline{Y}_{T_i})(t_{i+1}-t_i)}}{\sigma(\overline{Y}_{T_i})\sqrt{t_{i+1} - t_i}}. \tag{27}$$

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

$$\bar{\zeta}_{T_{i+1}} = \bar{\zeta}_{T_i} + b(\bar{\zeta}_{T_i})(t_{i+1} - t_i) + \sigma(\bar{\zeta}_{T_i})\sqrt{t_{i+1} - t_i}\,\varepsilon_{i+1}, \qquad i \geq 0. \tag{28}$$

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 A.1 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 $\bar{\zeta}_0 \sim \eta$. As $(\bar{\zeta}_0, \varepsilon_k, k \leq i - 1)$ is $\overline{\mathcal{G}}_{T_i}$-measurable, we get that ε_{i+1} is a standard Gaussian variable independent of $(\bar{\zeta}_0, \varepsilon_k, k \le i-1)$. This completes the proof of Lemma 5.3. □

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

LEMMA 5.4. *For $t \in [t_i, t_{i+1}]$, $\bar{\xi}_t = \bar{x}(t) + \sigma(\bar{\xi}_{t_i})\overline{C}_i(t)$, where $((\overline{C}_i(t), t_i \le t \le t_{i+1}), i \ge 0)$ is a sequence of independent Brownian bridges adapted to $(\overline{\mathcal{A}}_t)$, independent of $(\bar{\xi}_{t_j}, j \ge 0)$.*

PROOF. We have $\bar{y}(u) = \bar{\zeta}_{T_i} + \frac{u-T_i}{T_{i+1}-T_i}(\bar{\zeta}_{T_{i+1}} - \bar{\zeta}_{T_i})$. Using (27)–(28), we obtain, for $u \in [T_i, T_{i+1}]$,

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

with

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

For $t_i \le t \le t_{i+1}$, using (13) and (23), we get $\bar{x}(t) = \bar{y}(\overline{A}_t(\bar{\zeta}_.))$. Thus, $\bar{\xi}_t - \bar{x}(t) = D_i(\overline{A}_t(\bar{\zeta}_.))$, and define, using (26), $\overline{C}_i(t)$ by

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

Proving that $(\bar{\xi}_{t_i}, i \ge 0)$ is independent of $((\overline{C}_i(t), t \in [t_i, t_{i+1}]), i \ge 0)$ is equivalent to proving that $(\bar{\xi}_0, \varepsilon_i, i \ge 1)$ is independent of $((\overline{C}_i(t), t \in [t_i, t_{i+1}]), i \ge 0)$. We now show that, $\forall i \ge 1$, $(\bar{\xi}_0, \varepsilon_1, \dots, \varepsilon_i)$ is independent of $(\overline{C}_0, \dots, \overline{C}_{i-1})$ and that the latter processes are independent Brownian bridges. Using Proposition A.1 with $B = \overline{B}^{(i-1)}$, $\mathcal{F}_. = \overline{\mathcal{G}}_{T_{i-1}+.}$, $\tau = \sigma^2(\bar{\xi}_{t_{i-1}})$, $i \ge 1$ yields that $W_i(t - t_{i-1}) = \frac{1}{\sigma(\bar{\xi}_{t_{i-1}})}\overline{B}^{(i-1)}_{\sigma^2(\bar{\xi}_{t_{i-1}})(t-t_{i-1})}$, $t \ge t_{i-1}$, is a Brownian motion independent of $\overline{\mathcal{G}}_{T_{i-1}}$. Thus, $(\overline{C}_{i-1}(t), t \in [t_{i-1}, t_i])$ 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 $(\overline{C}_{i-1}(t), t \in [t_{i-1}, t_i])$ are independent.

For $i = 1$, as $\bar{\xi}_0$ is $\overline{\mathcal{G}}_0$-measurable, we get that $\bar{\xi}_0, \varepsilon_1, \overline{C}_0$ are independent and $\overline{C}_0$ is a Brownian bridge. By induction, let us assume that $\bar{\xi}_0, \varepsilon_1, \dots, \varepsilon_i, \overline{C}_0, \dots, \overline{C}_{i-1}$ are independent and that $\overline{C}_0, \dots, \overline{C}_{i-1}$ are Brownian bridges (on their respective interval of definition). As $Z = (\bar{\xi}_0, \varepsilon_1, \dots, \varepsilon_i, \overline{C}_0, \dots, \overline{C}_{i-1})$ is $\overline{\mathcal{G}}_{T_{i-1}}$-measurable, we get that $Z, \varepsilon_{i+1}, \overline{C}_i$ are independent. The proof of Lemma 5.4 is complete. □

Thus, we have constructed a process $(\bar{\xi}_t)$ with distribution Q_b obtained by the mapping $(y(u), u \geq 0) \to (y(\overline{A}_t(y)), t \geq 0)$. Hence, $(x(t) = y(\overline{A}_t(y)), t \leq \overline{T}_a(y) = \bar{\tau}_a(x))$ is the image of $(y(u), u \leq a)$. This proves $\delta(\widetilde{\mathcal{G}}_0^a, \mathcal{G}_0^{\bar{\tau}_a(X)}) = 0$.

Moreover, $(x(t), t \leq T)$ is the image of $(y(u), u \leq \overline{A}_T(y))$. This yields $\delta(\widetilde{\mathcal{G}}_0^{\overline{A}_T(Y)}, \mathcal{G}_0^T) = 0$. This completes the proof of Theorem 3.1. □

PROOF OF LEMMA 3.2. Using (8), $A_t(y) = u \Leftrightarrow T_u(y) = t$ yields that $A_n = \int_0^{nh_n} \sigma^2(y(A_s(y)))\, ds$. Combining with (13), we get

$$A_n - \overline{A}_n = \sum_{i=1}^n \int_{t_{i-1}}^{t_i} \bigl(\sigma^2\bigl(y\bigl(A_s(y)\bigr)\bigr) - \sigma^2\bigl(y\bigl(\overline{A}_{t_{i-1}}(y)\bigr)\bigr)\bigr)\, ds.$$

Under $\widetilde{P}_b$, $(Y(A_t(Y)) = X_t)$ has distribution P_b (see proof of Proposition 3.2). Hence, $E_{\widetilde{P}_b}|A_n - \overline{A}_n| = E_{P_b}|\sum_{i=1}^n \int_{t_{i-1}}^{t_i} (\sigma^2(X_s) - \sigma^2(X_{t_{i-1}}))\, ds|$. Denoting by $\mathcal{L}$ the generator of the diffusion (X_t) ($\mathcal{L}h = (1/2)\sigma^2 h'' + bh'$), the Itô formula yields $\int_{t_{i-1}}^{t_i} (\sigma^2(X_s) - \sigma^2(X_{t_{i-1}}))\, ds = B_1(i) + B_2(i)$, with

$$B_1(i) = \int_{t_{i-1}}^{t_i} dv \int_{t_{i-1}}^{s} \mathcal{L}\sigma^2(X_u)\, du,$$

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

Condition (C) and (H1) ensure that $\|\mathcal{L}\sigma^2\|_\infty$ is bounded by D_1 depending on K, K_σ, σ_1, so that, $|B_1(i)| \leq D_1 h_n^2/2$. For the second term,

$$\begin{aligned} B_2(i) &= \int_{t_{i-1}}^{t_i} ds \int_{t_{i-1}}^{s} (\sigma^2)'(X_u)\sigma(X_u)\, dB_u \\ &= \int_{t_{i-1}}^{t_i} (t_i - u)\bigl(\sigma^2\bigr)'(X_u)\sigma(X_u)\, dB_u, \end{aligned}$$

$$\sum_{i=1}^n B_2(i) = \int_0^{nh_n} H_u^{(n)}\, dB_u,$$

where

$$H_u^{(n)} = \sum_{i=1}^n 1_{]t_{i-1}, t_i]}(u)(t_i - u)\bigl(\sigma^2\bigr)'(X_u)\sigma(X_u).$$

This yields $E_{P_b}(\sum_{i=1}^n B_2(i))^2 = E_{P_b} \int_0^{nh_n} (H_u^{(n)})^2\, du \leq D_2 n h_n^3$ with D_2 a constant. Therefore, $E_{\widetilde{P}_b}|A_n - \overline{A}_n| \leq D'(nh_n^2 + (nh_n^3)^{1/2}) \leq Dnh_n^2$. □

PROOF OF PROPOSITION 3.5. *Proof of inequality* (16). As $\widetilde{\mathcal{E}}_0^{S_n}$ is a restriction of $\widetilde{\mathcal{E}}_0^{A_n}$ to a smaller σ-algebra, $\delta(\widetilde{\mathcal{E}}_0^{A_n}, \widetilde{\mathcal{E}}_0^{S_n}) = 0$. To evaluate the other deficiency, we introduce a kernel from $\widetilde{\mathcal{E}}_0^{S_n}$ to $\widetilde{\mathcal{E}}_0^{A_n}$. Let $B \in \mathcal{C}_{A_n}^Y$, and set $N(\omega, B) =$

$E_{\widetilde{P}_0}(1_B|\mathcal{C}^Y_{S_n})(\omega)$, where $\widetilde{P}_0$, corresponding to $b=0$, is the distribution of $(\eta + B_u, u \geq 0)$. Now, $N(\widetilde{P}_b|\mathcal{C}^Y_{S_n})$ defines a probability on $(C(\mathbb{R}^+,\mathbb{R}), \mathcal{C}^Y_{A_n})$ with density w.r.t. $\widetilde{P}_0|\mathcal{C}^Y_{A_n}$, $(d\widetilde{P}_b/d\widetilde{P}_0)|\mathcal{C}^Y_{S_n}$. Indeed, for $B \in \mathcal{C}_{A_n}$,

$$N(\widetilde{P}_b|\mathcal{C}_{S_n})(B) = \int_\Omega N(\omega, B)\, d(\widetilde{P}_b|\mathcal{C}^Y_{S_n}) = E_{\widetilde{P}_0}\left(\frac{d\widetilde{P}_b}{d\widetilde{P}_0}\bigg|\mathcal{C}^Y_{S_n} E_{\widetilde{P}_0}(1_B|\mathcal{C}^Y_{S_n})\right)$$

$$= E_{\widetilde{P}_0}\left(\frac{d\widetilde{P}_b}{d\widetilde{P}_0}\bigg|\mathcal{C}^Y_{S_n} 1_B\right).$$

For T a bounded stopping time,

$$\frac{d\widetilde{P}_b}{d\widetilde{P}_0}\bigg|\mathcal{C}^Y_T = \widetilde{L}_T(b) = \exp\left(\int_0^T f(Y_u)\, dY_u - \int_0^T \frac{1}{2} f^2(Y_u)\, du\right).$$

Thus, $(d\widetilde{P}_b/d\widetilde{P}_0)|\mathcal{C}^Y_{A_n} = \widetilde{L}_{A_n}(b) = \widetilde{L}_{S_n}(b)V_n$, with $\log V_n = \int_{S_n}^{A_n} f(Y_u)\, dY_u - \int_{S_n}^{A_n} \frac{1}{2} f^2(Y_u)\, du$. Hence, $d\widetilde{P}_b|\mathcal{C}^Y_{A_n}/dN(\widetilde{P}_b|\mathcal{C}^Y_{S_n}) = V_n$. By the Pinsker inequality (Appendix) and Lemma 3.2, we have

$$\|N(\widetilde{P}_b|\mathcal{C}^Y_{S_n}) - \widetilde{P}_b|\mathcal{C}^Y_{A_n}\|_{\mathrm{TV}} = \frac{1}{2}\int_\Omega d\widetilde{P}_0 |\widetilde{L}_{S_n}(b) - \widetilde{L}_{A_n}(b)|$$

$$\leq \sqrt{K(\widetilde{P}_b|\mathcal{C}^Y_{A_n}, N(\widetilde{P}_b|\mathcal{C}^Y_{S_n}))/2},$$

$$K(\widetilde{P}_b|\mathcal{C}_{\tau_n}, N(\widetilde{P}_b|\mathcal{C}_{S_n})) = E_{\widetilde{P}_b|\mathcal{C}^Y_{A_n}} \int_{S_n}^{A_n} \frac{1}{2} f^2(X_u)\, du$$

$$\leq \frac{K^2}{2\sigma_0^4} E_{\widetilde{P}_b}|A_n - \overline{A}_n| \leq \frac{K^2}{\sigma_0^4} cnh_n^2.$$

Using that $\delta(\widetilde{\mathcal{E}}_0^{S_n}, \widetilde{\mathcal{E}}_0^{A_n}) \leq \sup_{b\in\mathcal{F}_K} \|N(\widetilde{P}_b|\mathcal{C}^Y_{S_n}) - \widetilde{P}_b|\mathcal{C}^Y_{A_n}\|_{\mathrm{TV}}$ yields the first inequality. We proceed analogously for the other one.

Proof of inequality (17). These experiments have the same sample space and are, respectively, associated with the distributions $\widetilde{P}_b$ (resp., $\widetilde{Q}_b$) on $C(\mathbb{R}^+,\mathbb{R})$ of $(\zeta_u, u \geq 0)$ given by (7) [resp., $(\bar{\zeta}_u, u \geq 0)$ given by (11)]. Hence,

$$\Delta(\widetilde{\mathcal{E}}_0^{\overline{A}_n}, \widetilde{\mathcal{G}}_0^{\overline{A}_n}) \leq \sup_{b\in\mathcal{F}} \|\widetilde{P}_b/_{\mathcal{C}^Y_{\overline{A}_n}} - \widetilde{Q}_b/_{\mathcal{C}^Y_{\overline{A}_n}}\|_{\mathrm{TV}} = \Delta_0(\widetilde{\mathcal{E}}_0^{\overline{A}_n}, \widetilde{\mathcal{G}}_0^{\overline{A}_n}).$$

Using the bound of Proposition A.2 yields

$$2\|\widetilde{P}_b/_{\mathcal{C}^Y_{\overline{A}_n}} - \widetilde{Q}_b/_{\mathcal{C}^Y_{\overline{A}_n}}\|^2_{\mathrm{TV}} \leq K(\widetilde{P}_b/_{\mathcal{C}^Y_{\overline{A}_n}}, \widetilde{Q}_b/_{\mathcal{C}^Y_{\overline{A}_n}})$$

$$= E_{\widetilde{P}_b/_{\mathcal{C}^Y_{\overline{A}_n}}}\left(\int_0^{\overline{A}_n} (f(Y_v) - \bar{f}(v, Y))^2\, dv\right).$$

Setting $T_i = T_i(Y)$ and using that, for $i = 1, \dots, n$, $T_i = T_i(Y) = \overline{A}_{t_i}(Y)$ [see (5.2)] and that f is Lipschitz with constant L [see (5)], we get

$$\int_0^{\overline{A}_n} \big(f(Y_v) - \bar{f}(v, Y)\big)^2 \, dv = \sum_{i=1}^n \int_{T_{i-1}}^{T_i} \big(f(Y_v) - f(Y_{T_{i-1}})\big)^2 \, dv$$
$$\leq L^2 \sum_{i=1}^n \int_{T_{i-1}}^{T_i} (Y_v - Y_{T_{i-1}})^2 \, dv.$$

Under $\widetilde{P}_b$, $Y_v - Y_{T_{i-1}} = \int_{T_{i-1}}^v f(Y_u)\, du + B_v - B_{T_{i-1}}$, with (B_v) Brownian motion. So

$$(Y_v - Y_{T_{i-1}})^2 \leq 2\left[\left(\int_{T_{i-1}}^v f(Y_u)\, du\right)^2 + (B_v - B_{T_{i-1}})^2\right].$$

This yields $\int_0^{\overline{A}_n} (f(Y_v) - \bar{f}(v, Y))^2 \, dv \leq 2L^2(R_1 + R_2)$, with

$$R_1 = \sum_{i=1}^n \int_{T_{i-1}}^{T_i} \left(\int_{T_{i-1}}^u f(Y_v)\, dv\right)^2 du, \qquad R_2 = \sum_{i=1}^n \int_{T_{i-1}}^{T_i} (B_u - B_{T_{i-1}})^2 \, du.$$

Using (5) and $T_i - T_{i-1} \leq \sigma_1^2 h_n$ by (23),

$$R_1 \leq \frac{K^2}{\sigma_0^4} \sum_{i=1}^n (T_i - T_{i-1})^3 \leq \frac{K^2}{\sigma_0^4} n\big(\sigma_1^2 h_n\big)^3.$$

For the second term, using definition (26),

$$E_{\widetilde{P}_b}(R_2) = E_{\widetilde{P}_b}\left(\sum_{i=1}^n \int_{T_{i-1}}^{T_i} (B_u - B_{T_{i-1}})^2 \, du\right) \leq \sum_{i=1}^n \int_0^{\sigma_1^2 h_n} E_{\widetilde{P}_b}\big(\overline{B}_v^{(i)}\big)^2 \, dv$$
$$= n \frac{(\sigma_1^2 h_n)^2}{2}.$$

Thus, the result follows from

$$K(\widetilde{P}_b/\mathcal{C}_{\overline{A}_n^Y}, \widetilde{Q}_b/\mathcal{C}_{\overline{A}_n^Y}) \leq 2L^2\left(\frac{K^2}{3\sigma_0^4} n\big(\sigma_1^2 h_n\big)^3 + n\frac{(\sigma_1^2 h_n)^2}{2}\right).$$

%upqed Joining (16), (17) and (15) completes the proof of Proposition 3.5. □

APPENDIX

Let us recall properties of the Le Cam deficiency distance Δ. 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 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 $(\Omega, \mathcal{A})$, and 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 randomization of $\mathcal{E}$ by the kernel M. If the kernel is deterministic, that is, 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 T.

DEFINITION A.1. $\Delta(\mathcal{E}, \mathcal{G}) = \max\{\delta(\mathcal{E}, \mathcal{G}), \delta(\mathcal{G}, \mathcal{E})\}$ where $\delta(\mathcal{E}, \mathcal{G}) = \inf_{M \in \mathcal{M}_{\Omega:\mathcal{X}}} \sup_{f \in \mathcal{F}} \|MP_f - Q_f\|_{\mathrm{TV}}$, $\|\cdot\|_{\mathrm{TV}}$ is the total variation distance and $\mathcal{M}_{\Omega:\mathcal{X}}$ the set of Markov kernels from $(\Omega, \mathcal{A})$ to $(\mathcal{X}, \mathcal{C})$.

When $\Delta(\mathcal{E}, \mathcal{G}) = 0$, the two experiments are said to be equivalent. When the experiments have the same sample space: $(\Omega, \mathcal{A}) = (\mathcal{X}, \mathcal{C})$, it is possible to define $\Delta_0(\mathcal{E}, \mathcal{G}) = \sup_{f \in \mathcal{F}} \|P_f - Q_f\|_{\mathrm{TV}}$, which satisfies $\Delta(\mathcal{E}, \mathcal{G}) \le \Delta_0(\mathcal{E}, \mathcal{G})$. Consider an asymptotic framework $\varepsilon \to 0$ and 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.

DEFINITION A.2. The families $\mathcal{E}^\varepsilon$, $\mathcal{G}^\varepsilon$ are asymptotically equivalent as $\varepsilon \to 0$ if $\Delta(\mathcal{E}^\varepsilon, \mathcal{G}^\varepsilon)$ tends to 0. The σ-algebra $\mathcal{B}^\varepsilon$ is asymptotically sufficient if $\Delta(\mathcal{E}^\varepsilon, \mathcal{E}^\varepsilon/_{\mathcal{B}^\varepsilon})$ tends to 0, where $\mathcal{E}^\varepsilon/_{\mathcal{B}^\varepsilon}$ is the restriction of $\mathcal{E}^\varepsilon$ to $\mathcal{B}^\varepsilon$.

We state now two auxiliary results used in proofs.

PROPOSITION A.1. *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 $(W(t) = \frac{1}{\sqrt{\tau}} B_{\tau t}, t \ge 0)$ is a standard Brownian motion, independent of $\mathcal{F}_0$.*

This result follows from a straightforward application of Paul Lévy's characterisation of the Brownian motion [see, e.g., Karatzas and Shreve (2000)]. Next, we recall the first Pinsker inequality [see, e.g., Tsybakov (2009)] for the total variation distance between probability measures. 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: $\|P - Q\|_{\mathrm{TV}} = \sup_{A \in \mathcal{A}} |P(A) - Q(A)| = \frac{1}{2} \int |p - q|\, d\nu$. The Kullback divergence of P w.r.t. Q is $K(P, Q) = \int \log \frac{dP}{dQ}\, dP$ if $P \ll Q$, $= +\infty$ otherwise.

PROPOSITION A.2. $\|P - Q\|_{\mathrm{TV}} \le \sqrt{K(P, Q)/2}$.

The remarkable feature of this inequality is that the left-hand side is a symmetric 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 $K(Q, P)$. The Pinsker inequality is particularly useful when P, Q are associated with diffusion type processes. Let P (resp., Q) be the distribution $C(\mathbb{R}^+, \mathbb{R})$ of the diffusion type process $d\xi_t = p(t, \xi_.)\,dt + dW_t$ with predictable drift $p(t, X_.)$ [resp., $d\eta_t = q(t, \eta_.)\,dt + dW_t$ with drift $q(t, X_.)$] and constant diffusion coefficient equal to 1, with the same initial condition $\xi_0 = \eta_0$. Let $T = T(X_.)$ be a finite stopping time under P and Q. Then the Girsanov formula stopped at T yields [with (X_v) the canonical process of $C(\mathbb{R}^+, \mathbb{R})$]

$$\frac{dP_T}{dQ_T} = \exp\left(\int_0^T (p(s, X_.) - q(s, X.))\,dX_s - \frac{1}{2}\int_0^T (p^2(s, X_.) - q^2(s, X_.))\,ds\right),$$

where $P_T = P/_{\mathcal{C}_T}, Q_T = Q/_{\mathcal{C}_T}$ are the restriction of P, Q to the σ-field $\mathcal{C}_T$. Hence, using that under $P\,dX_t - p(t, X_.)\,dt = dB_t$, with (B_t) a Brownian motion, yields $K(P_T, Q_T) = (1/2)E_P(\int_0^T (p(s, X_.) - q(s, X_.))^2\,ds)$.

REFERENCES

BROWN, L. D. and LOW, M. G. (1996). Asymptotic equivalence of nonparametric regression and white noise. *Ann. Statist.* **24** 2384–2398. MR1425958

BROWN, L. D., WANG, Y. and ZHAO, L. H. (2003). On the statistical equivalence at suitable frequencies of GARCH and stochastic volatility models with the corresponding diffusion model. *Statist. Sinica* **13** 993–1013. MR2026059

BROWN, L. D. and ZHANG, C.-H. (1998). Asymptotic nonequivalence of nonparametric experiments when the smoothness index is 1/2. *Ann. Statist.* **26** 279–287. MR1611772

BROWN, L. D., CAI, T. T., LOW, M. G. and ZHANG, C.-H. (2002). Asymptotic equivalence theory for nonparametric regression with random design. *Ann. Statist.* **30** 688–707. MR1922538

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. MR2102503

CARTER, A. V. (2006). A continuous Gaussian approximation to a nonparametric regression in two dimensions. *Bernoulli* **12** 143–156. MR2202326

CARTER, A. V. (2007). Asymptotic approximation of nonparametric regression experiments with unknown variances. *Ann. Statist.* **35** 1644–1673. MR2351100

COMTE, F., GENON-CATALOT, V. and ROZENHOLC, Y. (2007). Penalized nonparametric mean square estimation of the coefficients of diffusion processes. *Bernoulli* **13** 514–543. MR2331262

DALALYAN, A. and REISS, M. (2006). Asymptotic statistical equivalence for scalar ergodic diffusions. *Probab. Theory Related Fields* **134** 248–282. MR2222384

DALALYAN, A. and REISS, M. (2007). Asymptotic statistical equivalence for ergodic diffusions: The multidimensional case. *Probab. Theory Related Fields* **137** 25–47. MR2278451

DELATTRE, S. and HOFFMANN, M. (2002). Asymptotic equivalence for a null recurrent diffusion. *Bernoulli* **8** 139–174. MR1895888

GENON-CATALOT, V. (1990). Maximum contrast estimation for diffusion processes from discrete observations. *Statistics* **21** 99–116. MR1056065

GENON-CATALOT, V., LARÉDO, C. and NUSSBAUM, M. (2002). Asymptotic equivalence of estimating a Poisson intensity and a positive diffusion drift. *Ann. Statist.* **30** 731–753. MR1922540

GLOTER, A. and JACOD, J. (2001). Diffusions with measurement errors. I. Local asymptotic normality. *ESAIM Probab. Stat.* **5** 225–242 (electronic). MR1875672

GOLUBEV, G. K., NUSSBAUM, M. and ZHOU, H. H. (2010). Asymptotic equivalence of spectral density estimation and Gaussian white noise. *Ann. Statist.* **38** 181–214. MR2589320

GRAMA, I. G. and NEUMANN, M. H. (2006). Asymptotic equivalence of nonparametric autoregression and nonparametric regression. *Ann. Statist.* **34** 1701–1732. MR2283714

GRAMA, I. and NUSSBAUM, M. (1998). Asymptotic equivalence for nonparametric generalized linear models. *Probab. Theory Related Fields* **111** 167–214. MR1633574

GRAMA, I. and NUSSBAUM, M. (2002). Asymptotic equivalence for nonparametric regression. *Math. Methods Statist.* **11** 1–36. MR1900972

HOFFMANN, M. (1999). Adaptive estimation in diffusion processes. *Stochastic Process. Appl.* **79** 135–163. MR1670522

KARATZAS, I. and SHREVE, S. E. (2000). *Brownian Motion and Stochastic Calculus*, 2nd ed. Springer, New York.

KESSLER, M. (1997). Estimation of an ergodic diffusion from discrete observations. *Scand. J. Stat.* **24** 211–229. MR1455868

KUTOYANTS, Y. A. (2004). *Statistical Inference for Ergodic Diffusion Processes*. Springer, London. MR2144185

LARÉDO, C. F. (1990). A sufficient condition for asymptotic sufficiency of incomplete observations of a diffusion process. *Ann. Statist.* **18** 1158–1171. MR1062703

LE CAM, L. and YANG, G. L. (2000). *Asymptotics in Statistics: Some Basic Concepts*, 2nd ed. Springer, New York. MR1784901

MEISTER, A. (2011). Asymptotic equivalence of functional linear regression and a white noise inverse problem. *Ann. Statist.* **39** 1471–1495. MR2850209

MEISTER, A. and REISS, M. (2013). Asymptotic equivalence for nonparametric regression with non-regular errors. *Probab. Theory Related Fields* **155** 201–229. MR3010397

MILSTEIN, G. and NUSSBAUM, M. (1998). Diffusion approximation for nonparametric autoregression. *Probab. Theory Related Fields* **112** 535–543. MR1664703

NUSSBAUM, M. (1996). Asymptotic equivalence of density estimation and Gaussian white noise. *Ann. Statist.* **24** 2399–2430. MR1425959

REISS, M. (2008). Asymptotic equivalence for nonparametric regression with multivariate and random design. *Ann. Statist.* **36** 1957–1982. MR2435461

REISS, M. (2011). Asymptotic equivalence for inference on the volatility from noisy observations. *Ann. Statist.* **39** 772–802. MR2816338

TSYBAKOV, A. B. (2009). *Introduction to Nonparametric Estimation*. Springer, New York. MR2724359

WANG, Y. (2002). Asymptotic nonequivalence of Garch models and diffusions. *Ann. Statist.* **30** 754–783. MR1922541

MAP5
CNRS UMR 8145
UNIVERSITÉ PARIS DESCARTES
PRES SORBONNE PARIS CITÉ
45, RUE DES SAINTS-PÈRES
75006 PARIS
FRANCE
E-MAIL: valentine.genon-catalot@parisdescartes.fr

MIA AND LPMA
CNRS-UMR 7599
INRA AND UNIVERSITÉ PARIS DIDEROT
INRA, 78350, JOUY-EN-JOSAS
FRANCE
E-MAIL: catherine.laredo@jouy.inra.fr