Bernoulli **8**(2), 2002, 139–174

Asymptotic equivalence for a null recurrent diffusion

SYLVAIN DELATTRE[1] and MARC HOFFMANN[2]

[1]*Université Paris XII-Val de Marne, Département de Mathématiques, 61 avenue du Général de Gaulle, 94010 Créteil Cedex, France. E-mail: delat@ccr.jussieu.fr*
[2]*Laboratoire de Probabilités et Modèles Aléatoires, CNRS-UMR 7599, Université Paris VII, UFR de Mathématiques, case 7012, 2 place jussieu, 75251 Paris Cedex 05, France. E-mail: hoffmann@math.jussieu.fr*

We establish that the model generated by the observation of the path of a one-dimensional null recurrent diffusion, when the parameter is the compactly supported drift, is asymptotically equivalent to a mixed Gaussian white noise experiment as the observation time $T \to \infty$. The approximation is given in the sense of Le Cam's deficiency Δ-distance over Sobolev balls of smoothness order $\beta > \frac{1}{2}$.

Keywords: deficiency distance; diffusion processes; mixed Gaussian white noise; mixed normality; nonparametric experiments

1. Introduction

1.1. Motivation

In this paper we study the *asymptotic structure* of a null recurrent Markov model. By 'asymptotic structure' we mean the approximation of a given family of measures by other families which are better known or more tractable.

It is certainly the basic principle of Le Cam's (1986) asymptotic decision theory to approximate general statistical experiments by simpler ones. Although the term *simple experiment* is not defined precisely, one should be able in a simple experiment to find optimal estimators, or at least asymptotically optimal ones, in the minimax sense, say. Mathematically, Le Cam's Δ-distance provides us with a strong enough notion to achieve such an aim. For a review of the concepts of Le Cam's theory, see Nussbaum (1996), Le Cam (1986), Strasser (1985) and Le Cam and Yang (2000).

Gaussian experiments are a primary type of such simple experiments. The known results about asymptotic equivalence include the asymptotic equivalence of Gaussian regression and Gaussian white noise (Brown and Low 1996), density estimation from independent and identically distributed observations and Gaussian white noise (Nussbaum 1996; Klemela and Nussbaum 1998), asymptotic equivalence for nonparametric generalized linear models (Grama and Nussbaum 1998).

A broad guiding principle is that experiments that are parametrically locally asymptotically normal (see, for example, Ibragimov and Has'minskii 1981) can be approximated

1350–7265

in the nonparametric case by Gaussian white noise models. The next logical step, intended here, is to investigate a class of nonparametric experiments that can be approximated by mixed white noise experiments. In the parametric case, these should correspond to locally asymptotically mixed normal (LAMN) models (see, for example, Le Cam and Yang 2000). Emerging results of this kind have been proposed in Delattre and Hoffmann (2001).

We consider in this paper the following pilot model: for $T > 0$, we observe $(X_t)_{t\in[0,T]}$, where

$$X_t = x_0 + \int_0^t f(X_s)\mathrm{d}s + W_t, \qquad t \in \mathbb{R}_+, \tag{1}$$

with $x_0 \in \mathbb{R}$ and $(W_t)_{t\in\mathbb{R}_+}$ a standard Brownian motion. The unknown parameter f is a bounded function with support in $[-r, r]$ and belongs to a Besov ball $\Sigma = \Sigma(r, \beta, \rho)$ of smoothness β and radius ρ. Asymptotics are taken as $T \to \infty$.

Under the compact support assumption on f, the process X is Markov, null recurrent in the sense of Harris (see, for example, Revuz and Yor 1991) with invariant measure $m_f(x)\mathrm{d}x$ given by

$$m_f(x) = \frac{2}{\exp(-2\int_{-r}^x f(y)\mathrm{d}y) + \exp(2\int_x^r f(y)\mathrm{d}y)}. \tag{2}$$

1.2. Basic facts in the parametric case

In order to better understand the statistical model to hand, let us first consider the simplest parametric model, defined by $f(x) = \theta g(x)$, where g is a given known function. The log-likelihood is then

$$\theta \int_0^T g(X_t)\mathrm{d}X_t - \frac{\theta^2}{2}\int_0^T g^2(X_t)\mathrm{d}t.$$

We further restrict the model to the ergodic case, i.e. when the integral $\int_{\mathbb{R}} \exp(2\theta \int_0^x g(y)\mathrm{d}y)\mathrm{d}x < +\infty$. If θ_0 denotes the true value of the parameter, the ergodic theorem yields

$$\frac{1}{T}\int_0^T g^2(X_t)\mathrm{d}t \to \int g^2(x)\mu_{\theta_0}(x)\mathrm{d}x \text{ almost surely,}$$

if the integral $\int g^2(x)\mu_{\theta_0}(x)\mathrm{d}x$ is finite, where μ_{θ_0} is the probability measure proportional to $m_{\theta_0 g}$. This allows the local asymptotic normality property to be derived at any $\theta_0 > 0$ with the rate $\sqrt{T}$, thanks to classical limit theorem for continuous martingales (see Liptser and Shiryaev 1977; or Kutoyants 1984).

Returning to our original problem, when g has compact support, we have $\int_{\mathbb{R}} \exp(2\theta \int_0^x g(y)\mathrm{d}y)\mathrm{d}x = \infty$ and the previous results are no longer true. Instead, we have the local asymptotic mixed normality property with the rate $T^{1/4}$. This can be easily seen when θ is close to 0. Indeed, if $\theta_0 = 0$, then $X = W$ and we have

$$\int_0^T g^2(X_s)\mathrm{d}s = T\int_0^1 g^2(\sqrt{T}\beta_u^T)\mathrm{d}u,$$

where $\beta_u^T = T^{-1/2}W_{Tu}$ is a Brownian motion,

$$= T\int_{\mathbb{R}} g^2(\sqrt{T}x)L(\beta^T)_1^x\,\mathrm{d}x,$$

where $L(\beta^T)_x^u$ denotes the local time of β^T at time u and level x,

$$= \sqrt{T}\int_{\mathbb{R}} g^2(y)L(\beta^T)_1^{y/\sqrt{T}}\,\mathrm{d}y,$$

and hence

$$\frac{1}{\sqrt{T}}\int_0^T g^2(X_s)\mathrm{d}s \to L(\beta)_1^0\int g^2(x)\mathrm{d}x \text{ in law,}$$

where β is a standard Brownian motion. More generally, the Papanicolaou–Stroock–Varadhan theorem (Revuz and Yor 1991, p. 482) yields

$$\left(T^{-1/4}\int_0^T g(W_s)\mathrm{d}W_s,\, T^{-1/2}\int_0^T g^2(W_s)\mathrm{d}s\right) \to \left(B_{\left(\int g^2(x)\mathrm{d}x\right)L(\beta)_0^1},\, \int g^2(x)\mathrm{d}xL(\beta)_0^1\right),$$

where (B, β) are two independent Brownian motions. As we see, the local asymptotic mixed normality property holds at $\theta = 0$ with rate $T^{1/4}$ and Fisher information $|\mathcal{N}(0, \int g^2(x)\mathrm{d}x)|$.

1.3. Results

Our first result is the local asymptotic equivalence of the diffusion model (1) with parameter space $\Sigma(r, \beta, \rho)$ for sufficiently large β to the mixed Gaussian white noise experiment given by the observation of the random measure $Y_{f_0,T}^f(\mathrm{d}x)$ on $[-r, r]$, defined by

$$Y_{f_0,T}^f(\mathrm{d}x) = f(x)\mathrm{d}x + T^{-1/4}\Lambda^{-1/2}m_{f_0}(x)^{-1/2}B(\mathrm{d}x), \tag{3}$$

where $B(\mathrm{d}x)$ is centred and Gaussian, with intensity $\mathrm{d}x$, and independent of the random variable Λ, equal to the absolute value of a standard Gaussian random variable in law. Here, local equivalence means equivalence for f in a neighbourhood of f_0 in a proper topology, shrinking to 0 as $T \to \infty$ (Theorem 1).

This result is already 'nonparametric' in spirit since the rate of shrinking is slower than the 'parametric rate' $T^{-1/4}$, but is unsatisfactory for general purposes, since the result only holds for vanishing sets. It remains to piece together the parameter-local approximation by means of a preliminary estimator of f_0 to globalize the local version given by (3).

The form of globalization we propose is essentially different from the ones that are usually proposed in the literature (Brown and Low 1996; Nussbaum 1996; Grama and Nussbaum 1998), due to the absence of a simple variance-stabilizing transform in the sense of (6) – see Section 2.2. It can be described as follows. For $T > 0$, let $u_T > 0$ be such that $u_T \to 0$ and $u_T T^{1/4} \to \infty$. The diffusion model (1) with parameter f in the whole space Σ

is then globally equivalent to the following experiment: we observe the random measures Y_T^f and Z_T^f defined on $[-r, r]$ by

$$Y_T^f(\mathrm{d}x) = f(x)\mathrm{d}x + T^{-1/4}\Lambda^{-1/2}\hat{m}_{f,T}(x)^{-1/2}B(\mathrm{d}x),$$

$$Z_T^f(\mathrm{d}x) = f(x)\mathrm{d}x + u_T\tilde{B}(\mathrm{d}x),$$

where

$$\hat{m}_{f,T}(x) = 2\left\{\exp\left(-2\int_{-r}^{x} Z_T^f(\mathrm{d}y)\right) + \exp\left(2\int_{x}^{r} Z_T^f(\mathrm{d}y)\right)\right\}^{-1},$$

Λ, $B(\mathrm{d}x)$ are defined as in (3), and $\tilde{B}(\mathrm{d}x)$ is a centred Gaussian measure with intensity $\mathrm{d}x$, independent of (Λ, B).

Thus, we see that we have a form close to the classical 'signal plus noise' model, with noise intensity $T^{-1/4}$, with a random component in the variance. The variance process itself depends on f through the function $\hat{m}_{f,T}$, which converges to m_f as $T \to \infty$, at a rate u_T strictly slower than $T^{-1/4}$.

1.4. Organization of the paper

In Section 2, we give a precise definition of the statistical model. We state the local equivalence (Theorem 1) and its global analogue (Theorem 2). As in Nussbaum (1996) and Grama and Nussbaum (1998), the key idea for the local equivalence is a coupling technique for the likelihood processes of the two experiments: this is the scope of Section 3 (Proposition 1). We derive a first (weaker) version of Theorem 1, to be considered as an intermediate step. We develop the globalization in Section 4, by means of preliminary estimators based on Nadaraya–Watson techniques together with a key lemma (Lemma 1) which extends a former result of Nussbaum (1996). Thorough proofs of Theorems 1 and 2 are given in Section 5; the proof of Proposition 1 is given in Section 6. Some of the technical results that can be omitted in an initial reading are postponed until the Appendix.

2. Main results

Our approximation results will be stated in terms of deficiency between experiments (the Δ-distance) – see Le Cam (1986) and Strasser (1985) for a definition; see also Nussbaum (1996).

We will often define statistical experiments in term of observed random variables: if, on some probability space $(\Omega, \mathcal{A}, P^\theta)$, $\xi_1^\theta, \ldots, \xi_n^\theta$ are random variables with values in some measurable spaces

$$(A_1, \mathcal{A}_1), \ldots, (A_n, \mathcal{A}_n)$$

the experiment generated by the observation of $(\xi_1^\theta, \ldots, \xi_n^\theta)$ is

$$(A_1 \times \cdots \times A_n, \mathcal{A}_1 \otimes \cdots \otimes \mathcal{A}_n, (\Pi^\theta)_{\theta\in\Theta}),$$

where Π^θ is the law of $(\xi_1^0, \ldots, \xi_n^\theta)$ under P^θ, and Θ will be a parameter space to be specified.

2.1. Statistical model

The nonparametric class $\Sigma = \Sigma(r, \beta, \rho)$ is defined by

$$\Sigma = \{f : \mathbb{R} \to \mathbb{R} | f \text{ bounded}, \|f\|_{\beta,2,2} \leqslant \rho \text{ and } \operatorname{supp}(f) \subset [-r, r]\},$$

where supp(f) denotes the support of f, and the constants $r, \rho > 0$ and $\beta \in (0, 1)$ are given; we denote by $\|f\|_{\beta,2,2}$ the Besov norm

$$\|f\|_{\beta,2,2} = \left(\int_{\mathbb{R}} f^2(x)\mathrm{d}x\right)^{1/2} + \left(\int_{\mathbb{R}} \mathrm{d}h \frac{1}{|h|^{1+2\beta}} \int_{\mathbb{R}} (f(x+h) - f(x))^2 \, \mathrm{d}x\right)^{1/2},$$

which coincides with the classical Sobolev norm. Let $P_{x_0}^f$ be the unique law of the solution of (1) on the canonical space $C(\mathbb{R}_+, \mathbb{R})$ endowed with the canonical filtration $(\mathcal{F}_t^0)_{t\geqslant 0}$. We consider the model given by the observation, until time T, of the diffusion process X^{f,x_0} defined by (1). Formally, for $x_0 \in \mathbb{R}$, we study the family of experiments

$$\mathbb{E}^{x_0,T} = (C(\mathbb{R}_+, \mathbb{R}), \mathcal{F}_T^0, (P_{x_0,T}^f, f \in \Sigma)), \qquad T > 0, \tag{4}$$

where $P_{x_0,T}^f$ denotes the restriction of $P_{x_0}^f$ to the σ-field $\mathcal{F}_T^0$.

2.2. Local equivalence

The first approximation will be given when the invariant measure m_f is close to some known function uniformly on $[-r, r]$: let $\eta_T > 0$ be a family of real numbers, and for $f_0 \in \Sigma$ define

$$\Sigma_T(f_0) = \left\{ f \in \Sigma | \sup_{x\in[-r,r]} \left| \int_0^x (f(y) - f_0(y))\mathrm{d}y \right| \leqslant \eta_T \right\}.$$

Let us define precisely the accompanying family of experiments: let $B(\mathrm{d}x)$ be a centred Gaussian random measure on $[-r, r]$ with intensity $\mathrm{d}x$, i.e. a centred Gaussian process $(B(\varphi), \varphi \in L^2([-r, r], \mathrm{d}x))$ with covariance $(\varphi, \psi) \mapsto \int_{[-r,r]} \varphi(x)\psi(x)\mathrm{d}x$, and let Λ be a positive random variable whose law is the absolute value of an $\mathcal{N}(0, 1)$ random variable, independent of $B(\mathrm{d}x)$. For two bounded Borel functions, $f : [-r, r] \to \mathbb{R}$ and $v : [-r, r] \to \mathbb{R}_+$, denote by $Q(f, v)$ the law on the canonical space

$$(G, \mathcal{G}) = \left(\mathbb{R}^{L^2([-r,r])}, \mathcal{B}(\mathbb{R})^{\otimes L^2([-r,r])}\right)$$

of the process

$$f(x)\mathrm{d}x + \Lambda^{-1/2} v(x)^{1/2} B(\mathrm{d}x) \left(= \int_{[-r,r]} f(x)\varphi(x)\mathrm{d}x + \Lambda^{-1/2} B(v^{1/2}\varphi), \, \varphi \in L^2([-r, r]) \right). \tag{5}$$

Then the statistical model associated with the observation of the random measure (3) is

$$\mathbb{F}^{T,f_0} = \left(G, \mathcal{G}, \left(Q\left(f, T^{-1/2} m_{f_0}^{-1/2}\right)\right)_{f \in \Sigma}\right).$$

If $\mathbb{E}$ is an experiment with parameter space Θ and Θ' is a subset of Θ, denote by $\mathbb{E}_{\Theta'}$ the experiment with parameter space Θ' which is the restriction of $\mathbb{E}$ to Θ'. The local equivalence can then be stated as follows.

Theorem 1. *If $\beta > 1/2$ and $\eta_T \to 0$ then for all $x_0 \in \mathbb{R}$ the experiments $\mathbb{E}^{x_0,T}_{\Sigma_T(f_0)}$ and $\mathbb{F}^{T,f_0}_{\Sigma_T(f_0)}$ are asymptotically equivalent.*

Remark 1. The restriction $\beta > 1/2$ is made for technical reasons. We do not know whether the equivalence holds for lower orders of smoothness. However, the threshold $1/2$ is optimal for the asymptotic equivalence between white noise and nonparametric regression: this follows from a simple extension of the classical result of Brown and Low (1996), as mentioned to us by a referee; we therefore conjecture that our result is optimal.

Remark 2. Note also that the restriction $\beta < 1$ in the definition of Σ is not essential, since the equivalence also holds for subsets of Σ, therefore for all Besov classes with smoothness order greater than or equal to 1.

As a corollary, we obtain that any regular parametric submodel is locally asymptotically mixed normal:

Corollary 1. *Assume that $\beta > 1/2$. If $(f_\theta)_{\theta \in \mathbb{R}}$ is a family of functions such that*

- *$f_\theta \in \Sigma(r, \beta, \rho)$ for all $\theta \in \mathbb{R}$,*
- *the map $\theta \mapsto f_\theta$ from $\mathbb{R}$ into $L^2(\mathbb{R}, \mathrm{d}x)$ is differentiable at $\theta = 0$, with derivative $\dot{f}_0$,*

then the parametric models $(P^{f_\theta}_{x_0,T})_{\theta \in \mathbb{R}}$, $T > 0$, have the local asymptotic mixed normality property at $\theta = 0$ at the rate $T^{1/4}$ and with random Fisher information

$$I = \left| \mathcal{N}\left(0, \int_{-r}^{r} \dot{f}_0(x)^2 m_{f_0}(x) \mathrm{d}x\right)\right| \text{ in law.}$$

Proof of Corollary 1. Since

$$\sup_{|\theta| \leqslant CT^{-1/4}} \sup_{x \in [-r,r]} \left| \int_0^x (f_\theta(y) - f_0(y)) \mathrm{d}y \right| \to 0,$$

by Theorem 1, the model $(P^{f_\theta}_{x_0,T})_{|\theta| \leqslant CT^{-1/4}}$ is asymptotically equivalent to the observation of

$$(f_\theta - f_0)(x)\mathrm{d}x + T^{-1/4} \Lambda^{-1/2} m_{f_0}(x)^{-1/2} B(\mathrm{d}x)$$

with the same notation as in (3). The latter model is also asymptotically equivalent to

$$Y(\mathrm{d}x) = \theta \dot{f}_0(x)\mathrm{d}x + T^{-1/4} \Lambda^{-1/2} B(\mathrm{d}x)$$

because $\int_{\mathbb{R}}(f_\theta - f_0 - \theta\dot{f}_0)^2(x)m_{f_0}(x)\mathrm{d}x \to 0$. Finally,

$$\frac{\int \dot{f}_0(x)m_{f_0}(x)Y(\mathrm{d}x)}{\int \dot{f}_0(x)^2 m_{f_0}(x)\mathrm{d}x} = \theta + \frac{T^{-1/4}\xi}{\sqrt{I}}$$

is a sufficient statistic, with $\xi = \mathcal{N}(0, 1)$ in law and independent of I. □

Remark 3. We were not able to obtain a homoscedastic form of the local approximation; the local equivalence to (3) is not changed if we subtract $f_0(x)\mathrm{d}x$ from (3) and multiply by $\sqrt{m_{f_0}}$, since those terms do not depend on the unknown parameter f. Thus we also have local equivalence to

$$m_{f_0}(x)^{1/2}(f - f_0)(x)\mathrm{d}x + T^{-1/4}\Lambda^{-1/2}B(\mathrm{d}x),$$

with the same properties as in (3). Usually (Nussbaum 1996; Grama and Nussbaum 1998), a first approach would consist in looking for a variance-stabilizing transform in the following sense: find a functional $f \mapsto \mathcal{T}(f)$ solution to

$$D\mathcal{T}(f)[h] = (m_f)^{1/2}h \qquad \text{for all } h \in L^2([-r, r]), \tag{6}$$

where $\mathrm{D}\mathcal{T}(f)$ is the differential operator of $\mathcal{T}$ at point f. Heuristically, using

$$m_{f_0}^{1/2}(f - f_0) = \mathrm{D}\mathcal{T}(f_0)[f - f_0] \simeq \mathcal{T}(f) - \mathcal{T}(f_0) \qquad \text{for } f \text{ close to } f_0,$$

and the fact that we may ignore the term $\mathcal{T}(f_0)$ (again, this simply amounts to a translation of the observed process by a known quantity), we would eventually obtain a global equivalence of the form

$$\mathcal{T}(f)(x)\mathrm{d}x + T^{-1/4}\Lambda^{-1/2}B(\mathrm{d}x).$$

But this simple approach fails, since there is no solution to (6). Assume, on the contrary, that such a functional $\mathcal{T}$ exists. From (6) and the definition of $\sqrt{m_f}$, this implies that the mapping $f \mapsto D\mathcal{T}(f)[h]$ is also differentiable, for $h \in L^2([-r, r])$; but from the explicit form of $\sqrt{m_f}$ one can readily ascertain that, for two functions h and k in $L^2([-r, r])$ with $h \neq k$,

$$\frac{\partial^2}{\partial a \partial b}\mathcal{T}(ah(\cdot) + bk(\cdot)) \neq \frac{\partial^2}{\partial b \partial a}\mathcal{T}(ah(\cdot) + bk(\cdot))$$

since $m_{f_0}(x)$ is not local in f_0, i.e. involves all the values of f_0. By Schwarz's lemma for mixed derivatives, this is in contradiction to the function $(a, b) \mapsto \mathcal{T}(ah(\cdot) + bk(\cdot))$ being twice differentiable.

2.3. Global equivalence

We now state an equivalence result for f lying in the whole parameter space $\Sigma(r, \beta, \rho)$.

Consider two centred Gaussian random measures $B(\mathrm{d}x)$ and $\tilde{B}(\mathrm{d}x)$ with intensity $\mathrm{d}x$, and a positive random variable $\Lambda = |\mathcal{N}(0, 1)|$ in law such that B, $\tilde{B}$ and Λ are independent. Let $u_T > 0$, $T > 0$, and let $\mathbb{G}^T$ be the experiment, with parameter $f \in \Sigma$, generated by the observation of the following two random measures on $[-r, r]$:

$$Y_T^f(\mathrm{d}x) = f(x)\mathrm{d}x + T^{-1/4}\Lambda^{-1/2}\hat{m}_{f,T}(x)^{-1/2}B(\mathrm{d}x), \tag{7}$$

$$Z_T^f(\mathrm{d}x) = f(x)\mathrm{d}x + u_T\tilde{B}(\mathrm{d}x),$$

where

$$\hat{m}_{f,T}(x) = 2\left\{\exp\left(-2\int_{-r}^{x} Z_T^f(\mathrm{d}y)\right) + \exp\left(2\int_{x}^{r} Z_T^f(\mathrm{d}y)\right)\right\}^{-1}.$$

Here $(\hat{m}_{f,T}(x))_{x\in[-r,r]}$ denotes a continuous modification and, for $\varphi \in L^2([-r, r])$, $Y_T^f(\varphi)$ is a stochastic integral. (We observe that equation (7) defines $(Y_T^f(\varphi), \varphi \in L^2([-r, r])$ up to a modification, but this suffices to characterize the law of (Y_T^f, Z_T^f).)

Theorem 2. *If $\beta > 1/2$, $u_T \to 0$ and $u_T T^{1/4} \to +\infty$, then for all $x_0 \in \mathbb{R}$ the experiments $\mathbb{E}^{x_0,T}$ and $\mathbb{G}^T$ are asymptotically equivalent.*

Remark 4. A by-product of the proof is that we can replace $T^{-1/4}$ by $(T')^{-1/4}$ in the definition of $\mathbb{G}^T$, where $T' \sim T$. This effect is due to the mixed Gaussian character of the model.

3. Local approximation of the likelihood ratio

In this section we give a result which claims that the likelihood processes of $\mathbb{E}^{x_0,T}$ and $\mathbb{F}^{f_0,T}$ can be coupled (Proposition 1 below). Then we deduce a version of Theorem 1 (Corollary 2 below) which is a local asymptotic equivalence result weaker than Theorem 1.

Let X be the canonical process on $C(\mathbb{R}_+, \mathbb{R})$. Recall from Girsanov's formula that

$$\frac{\mathrm{d}P_{x_0,T}^f}{\mathrm{d}P_{x_0,T}^{f_0}} = Z_T^{f_0,f}\, P_{x_0}^{f_0}\text{-a.s.},$$

where

$$\log Z_T^{f_0,f} = \int_0^T (f - f_0)(X_s)\mathrm{d}W_s^{f_0} - \frac{1}{2}\int_0^T (f - f_0)^2(X_s)\mathrm{d}s, \tag{8}$$

$$W_t^{f_0} = X_t - X_0 - \int_0^t f_0(X_s)\mathrm{d}s.$$

The local approximation will be given on the following shrinking balls: for $f_0 \in \Sigma(r, \beta, \rho)$ and $\varepsilon_T, \eta_T > 0$, let $\mathcal{V}_T(f_0) = \mathcal{V}(f_0, \varepsilon_T, \eta_T)$ be defined by

$$\mathcal{V}_T(f_0) = \left\{ f \in \Sigma : \int_{-r}^{r} |f - f_0|^2 \leqslant \varepsilon_T^2,\ \sup_{x\in[-r,r]}\left|\int_0^x (f - f_0)\right| \leqslant \eta_T \right\}$$

where ε_T and η_T satisfy the following assumption:

Assumption A. $\varepsilon_T = O(T^{-p_1})$ *for some* $p_1 > 1/8$, *and* $\eta_T = O(T^{-p_2})$ *for some* $p_2 > 1/2 - 2p_1$. □

Proposition 1. *For all* $x_0 \in \mathbb{R}$, $T > 0$, $f_0 \in \Sigma$, *there exists a* $P_{x_0}^{f_0}$*-Gaussian centred random measure* $Y_T^{f_0}$ *on* $[-r, r]$ *with intensity* m_{f_0} *and a non-negative random variable* $\Lambda(f_0, T)$ *such that*:

(i) $Y_T^{f_0}$ *and* $\Lambda(f_0, T)$ *are independent under* $P_{x_0}^{f_0}$;
(ii) $\mathcal{L}(\Lambda(f_0, T)|P_{x_0}^{f_0}) \stackrel{\text{TV}}{\rightarrow} |\mathcal{N}(0, 1)|$ *as* $T^{-1/2}x_0 \rightarrow 0$ *uniformly in* $f_0 \in \Sigma$, *and where* $\stackrel{\text{TV}}{\rightarrow}$ *denotes convergence in the total variational norm.*

Moreover, under Assumption A *we have*

$$\mathrm{E}_{x_0}^{f_0}(|\tilde{Z}_T^{f_0,f} - Z_T^{f_0,f}|) \rightarrow 0 \qquad \text{as } T \rightarrow \infty, \tag{9}$$

uniformly in $x_0 \in \mathbb{R}$, $f_0 \in \Sigma$ *and* $f \in \mathcal{V}_T(f_0)$, *where*

$$\log \tilde{Z}_T^{f_0,f} = T^{1/4}\Lambda(f_0, T)^{1/2} Y_T^{f_0}(f - f_0) - \frac{T^{1/2}}{2}\Lambda(f_0, T)\int_{-r}^{r} (f - f_0)^2(x) m_{f_0}(x)\mathrm{d}x. \tag{10}$$

The proof is postponed until Section 6.

Corollary 2. *Suppose that Assumption* A *holds. Let* $R_T > 0$ *be such that* $T^{-1/2}R_T \rightarrow 0$. *Then we have*

$$\Delta\left(\mathbb{E}_{\mathcal{V}_T(f_0)}^{x_0,T}, \mathbb{F}_{\mathcal{V}_T(f_0)}^{T,f_0}\right) \rightarrow 0 \qquad \text{as } T \rightarrow \infty,$$

uniformly in $|x_0| \leqslant R_T$ *and uniformly in* $f_0 \in \Sigma$.

Proof. Let $Y_T^{f_0}$ and $\Lambda(f_0, T)$ be as in Proposition 1. Consider a random variable $\tilde{\Lambda}(f_0, T)$ defined on some probability space such that $\mathcal{L}(\tilde{\Lambda}(f_0, T)) = \mathcal{L}(\Lambda(f_0, T)|P_{x_0}^{f_0})$. Let $\tilde{\mathbb{F}}^{T,f_0}$ be the experiment with parameter $f \in \mathcal{V}_T(f_0)$ generated by the observation of the pair $(\tilde{\Lambda}(f_0, T), \tilde{Y}_{f_0,T}^f)$, where $\tilde{Y}_{f_0,T}^f$ denotes the random measure on $[-r, r]$ which equals

$$f(x)\mathrm{d}x + T^{-1/4}\tilde{\Lambda}(f_0, T)^{-1/2} m_{f_0}(x)^{-1/2} B(\mathrm{d}x)$$

on the set $\{\tilde{\Lambda}(f_0, T) > 0\}$ and which vanishes on $\{\tilde{\Lambda}(f_0, T) = 0\}$. Here $B(\mathrm{d}x)$ is Gaussian with intensity $\mathrm{d}x$ and is independent of $\tilde{\Lambda}(f_0, T)$. If $\Pi_{f_0,T}^f$ denotes the law of $(\tilde{\Lambda}(f_0, T), \tilde{Y}_{f_0,T}^f)$, then we clearly have

$$\mathcal{L}\left(\left(\frac{\mathrm{d}\Pi_{f_0,T}^f}{\mathrm{d}\Pi_{f_0,T}^{f_0}}, f \in \mathcal{V}_T(f_0)\right)\middle|\Pi_{f_0,T}^{f_0}\right) = \mathcal{L}\Big(\Big(\tilde{Z}_T^{f_0,f}, f \in \mathcal{V}_T(f_0)\Big)|P_{x_0}^{f_0}\Big).$$

From (9), we deduce the asymptotic equivalence (see Le Cam and Lo Yang 2000) of $\mathbb{E}_{\mathcal{V}_T(f_0)}^{x_0,T}$ and $\tilde{\mathbb{F}}_{\mathcal{V}_T(f_0)}^{f_0,T}$.

Next, Proposition 1(ii) and the fact that $T^{-1/2}R_T \rightarrow 0$ yield

$$\Big\|\mathcal{L}\Big(\Big(\tilde{\Lambda}(f_0, T), \tilde{Y}_{f_0,T}^f\Big)\Big) - \mathcal{L}\Big(\Big(\Lambda, Y_{f_0,T}^f\Big)\Big)\Big\|_{\mathrm{TV}} \rightarrow 0$$

uniformly in $x_0 \in [-R_{\mathrm{T}}, R_T]$, $f_0 \in \Sigma$ and $f \in \mathcal{V}_T(f_0)$, where $(\Lambda, Y^f_{f_0,T})$ generates $\mathbb{F}^{T,f_0}$ and is defined by (3). To conclude the proof, it suffices to remark that $\mathbb{F}^{f_0,T}$ is the experiment given by $(\Lambda, Y^f_{f_0,T})$ since Λ is a function of $Y^f_{f_0,T}$, namely $m_{f_0}(r)$ times the quadratic variation on $[-r, r]$ of $\int_{-r}^{\cdot} Y^f_{f_0,T}(\mathrm{d}x)$. □

4. Globalization

Following Le Cam's principle, in order to obtain a global equivalence result, it remains to piece together the parameter-local approximation of Corollary 2 by means of a preliminary estimator of f_0.

4.1. Preliminary estimator

We first investigate the properties of a preliminary estimator $\hat{f}_T$ (being $\mathcal{F}^0_T$-measurable) constructed using Nadaraya–Watson type techniques. Consider a non-negative compactly supported kernel K of class C^1 such that $\int_{\mathbb{R}} K(x)\mathrm{d}x = 1$ and $\int_{\mathbb{R}} xK(x)\mathrm{d}x = 0$. For $h > 0$, set $K_h(x) = h^{-1}K(h^{-1}x)$. Let $h(T) = (\sqrt{T})^{-1/(2\beta+1)}$. Recall that X denotes the canonical process on $C(\mathbb{R}_+, \mathbb{R})$. For each $T > 0$, thanks to Kolmogorov's criterion, the process $(\int_0^T K_{h(T)}(X_t - x)\mathrm{d}X_t,\ x \in \mathbb{R})$ has a $P^0_{x_0,T}$-modification which is continuous and with which we will work. For $x \in [-r, r]$ set

$$\hat{f}_T(x) = \frac{\int_0^T K_{h(T)}(X_t - x)\mathrm{d}X_t}{\int_0^T K_{h(T)}(X_t - x)\mathrm{d}t}, \tag{11}$$

where, by convention, $0/0 = 0$. Moreover, we set $\hat{f}_T(x) = 0$ for $x \notin [-r, r]$. The performance of $\hat{f}_T$ is summarized in the following result, the proof of which is postponed to the Appendix:

Proposition 2. *The random variables*

$$(\sqrt{T})^{2\beta/(2\beta+1)} \int_{-r}^{r} (\hat{f}_T(x) - f(x))^2\,\mathrm{d}x, \tag{12}$$

$$T^{1/4} \sup_{x\in[-r,r]} \left| \int_{-r}^{x} (\hat{f}_T(y) - f(y))\mathrm{d}y \right|, \tag{13}$$

$$\|\hat{f}_T\|_{\beta,2,2}, \tag{14}$$

$T > 0$, *are* $P^f_{x_0}$*-tight, uniformly with respect to* $f \in \Sigma$.

4.2. Globalization

This enables us to state our first global result. Let $T' = T'(T) > 0$ such that $T' \to \infty$ and $T'/T \to 0$ as $T \to \infty$. Consider (B, Λ) as in Section 2.3 and independent of the diffusion X^{f,x_0}. Let $\tilde{\mathbb{E}}^{x_0,T}$ be the experiment generated by the observation of $((X_t^{f,x_0})_{0\leq t\leq T'}, \tilde{Y}_T^f) \in C([0, T'], \mathbb{R}) \times G$, where $\tilde{Y}_T^f$ is the random measure

$$\tilde{Y}_T^f(\mathrm{d}x) = f(x)\mathrm{d}x + [T - T']^{-1/4}\Lambda^{-1/2} m_{\hat{f}_{T'}}(x)^{-1/2} B(\mathrm{d}x), \tag{15}$$

now with $f \in \Sigma$. Here $\hat{f}_{T'}$ is shorthand for $\hat{f}_{T'} \circ X^{f,x_0}$.

Proposition 3. *If $\beta > 1/2$ and $T' = T^q$ with $1 - 2\beta/(1 + 6\beta) < q < 1$, then, for all $x_0 \in \mathbb{R}$, the experiments $\mathbb{E}^{x_0,T}$ and $\tilde{\mathbb{E}}^{x_0,T}$ are asymptotically equivalent.*

To prepare for the proof of Proposition 3, we first give an extension to a lemma due to Nussbaum (1996). Let $(A, \mathcal{A}, (P^\theta)_{\theta\in\Theta})$ be a statistical experiment. For $i = 1, 2$, consider a measurable space $(A_i, \mathcal{A}_i)$, and for each $\theta \in \Theta$ a Markov kernel $N_i^\theta(w, \mathrm{d}w_i)$ from $(A, \mathcal{A})$ into $(A_i, \mathcal{A}_i)$. We can define the two statistical experiments

$$\mathbb{E}_i = (A \times A_i, \mathcal{A} \otimes \mathcal{A}_i, (P_i^\theta)_{\theta\in\Theta}),$$

where $P_i^\theta(\mathrm{d}w, \mathrm{d}w_i) = P^\theta(\mathrm{d}w)N_i^\theta(w, \mathrm{d}w_i)$. Moreover, for each $\omega \in A$, let $\mathcal{V}(\omega)$ be a subset of Θ and let $\mathbb{F}_i(\omega)$ be the experiment

$$\mathbb{F}_i(\omega) = (A_i, \mathcal{A}_i, (N_i^\theta(w, \cdot))_{\theta\in\mathcal{V}(\omega)}).$$

Lemma 1. *Suppose that the following assumptions hold:*

(i) *For each $\theta \in \Theta$, the set $\{\omega \in A | \theta \in \mathcal{V}(\omega)\}$ belongs to $\mathcal{A}$.*
(ii) *For $i = 1, 2$, for all finite subsets $\Theta' \subset \Theta$, there exists a countably generated σ-field $\mathcal{A}_i' \subset \mathcal{A}_i$ such that, for all $\theta_0 \in \Theta'$, for P^{θ_0}-almost all ω, $\mathcal{A}_i'$ is $(N_i^\theta(w, \cdot), \theta \in \Theta')$-sufficient.*

Then, for all $B \in \mathcal{A}$, we have

$$\Delta(\mathbb{E}_1, \mathbb{E}_2) \leq \sup_{\omega\in B} \Delta(\mathbb{F}_1(\omega), \mathbb{F}_2(\omega)) + \sup_{\theta\in\Theta} P^\theta(\{\theta \notin \mathcal{V}\} \cup (A\backslash B)). \tag{16}$$

The proof is postponed to the Appendix.

Proof of Proposition 3. Recall that X denotes the canonical process on $C(\mathbb{R}, \mathbb{R}_+)$, $\mathcal{F}_t^0 = \sigma(X_s, s \leq t)$ and $Q(f, v)$ is the law of (5). It is easily seen that $Q(f, v)(\mathrm{d}y)$ is a Markov kernel from $\mathcal{L}_{\mathbb{R}}^\infty([-r, r]) \times \mathcal{L}_{\mathbb{R}+}^\infty([-r, r])$ (endowed with the Borel σ-field associated with the uniform norm) into $(G, \mathcal{G})$.

We will apply Lemma 1 with

$$\begin{aligned}
&\Theta = \Sigma(\beta, r, \rho), && (A, \mathcal{A}, (P^\theta)_{\theta\in\Theta}) = (C(\mathbb{R}_+, \mathbb{R}), \mathcal{F}_{T'}^0, (P_{x_0,T'}^f)_{f\in\Sigma}),\\
&(A_1, \mathcal{A}_1) = (C(\mathbb{R}_+, \mathbb{R}), \mathcal{F}_{T-T'}^0), && N_1^f(\omega, \mathrm{d}\omega_1) = P_{X_{T'}(\omega),T-T'}^f(\mathrm{d}\omega_1),\\
&(A_2, \mathcal{A}_2) = (G, \mathcal{G}), && N_2^f(\omega, \mathrm{d}\omega_2) = Q(f, (T - T')^{-1/2} m_{\hat{f}_{T'}(\omega)}^{-1})(\mathrm{d}\omega_2).
\end{aligned}$$

Observe that, since $\hat{f}_{T'} \in C([-r, r], \mathbb{R})$ and is $\mathcal{F}_{T'}^0$-measurable, $N_2^f(\omega, \mathrm{d}\omega_2)$ is still a Markov kernel from $(A, \mathcal{A})$ into $(A_2, \mathcal{A}_2)$. Then $\mathbb{E}_2 = \tilde{\mathbb{E}}^{x_0,T}$. Moreover, thanks to the Markov property, $\mathbb{E}_1$ is the image of $\mathbb{E}^{x_0,T}$ under the map $(\Omega, \mathcal{F}_T^0) \to (\Omega \times \Omega, \mathcal{F}_{T'}^0 \otimes \mathcal{F}_{T-T'}^0)$, $\omega(\cdot) \mapsto (\omega(\cdot), \omega(T' + \cdot))$ and $\mathbb{E}^{x_0,T}$ is the image of $\mathbb{E}_1$ under the map $(\Omega \times \Omega, \mathcal{F}_{T'}^0 \otimes \mathcal{F}_{T-T'}^0) \to (\Omega, \mathcal{F}_T^0)$, $(\omega, \omega_1) \mapsto \omega 1_{[0,T']} + \omega_1(\cdot - T')1_{[T',T]}$. Hence $\Delta(\mathbb{E}_1, \mathbb{E}^{x_0,T}) = 0$.

Assumption (ii) is satisfied with $\mathcal{A}_1' = \mathcal{A}_1$ and $\mathcal{A}_2' = \sigma(Y([-r, x]),\ x \in [-r, r] \cap \mathbb{Q})$, where Y denotes the canonical process on A_2. Set

$$\mathcal{V}(\omega) = \left\{ f \in \Sigma : \int_{-r}^{r} |f - \hat{f}_{T'}|^2 \leqslant T^{-2p_1},\ \sup_{x \in [-r,r]} \left| \int_0^x (f - \hat{f}_{T'}) \right| \leqslant T^{-p_2} \right\},$$

where p_1, p_2 are such that

$$\frac{1}{8} < p_1 < \frac{q\beta}{2(2\beta+1)}, \qquad \frac{1}{2} - 2p_1 < p_2 < \frac{q}{4},$$

which is possible in view of the condition on q and β. Assumption (i) is obviously satisfied. Moreover, thanks to Proposition 2, since $p_1 < q\beta/2(2\beta+1)$ and $p_2 < q/4$, we have $P^f_{x_0,T'}(f \notin \mathcal{V}(\omega)) \to 0$ as $T \to \infty$, uniformly with respect to $f \in \Sigma$.

Let $K > 0$ (arbitrarily large) and set

$$B = \{|X_{T'}| \leqslant K(T')^{1/2}\} \cap \{\|\{\hat{f}_{T'}\|_{\beta,2,2} \leqslant K\}.$$

Since $T' = o(T - T')$, we can apply Corollary 2 (with $\rho + K$ in place of ρ) and we deduce that $\sup_{\omega \in B} \Delta(\mathbb{F}_1(\omega), \mathbb{F}_2(\omega)) \to 0$ as $T \to \infty$. Therefore, upper bound (16) yields $\limsup_T \Delta(\mathbb{E}_1, \mathbb{E}_2) \leqslant \limsup_T \sup_{f \in \Sigma} P^f_{x_0}(B)$. Because $\inf_{T>1, f \in \Sigma} P^f_{x_0}(\|\hat{f}_{T'}\|_{\beta,2,2} \leqslant K) \to 1$ as $K \to \infty$ by Proposition 2 and $\inf_{T>0, f \in \Sigma} P^f_{x_0}(|X_{T'}| \leqslant K(T')^{1/2}) \to 1$ as $K \to \infty$ (see Section 6), the proof is complete. □

5. Proof of Theorems 1 and 2

In this section we assume that $\beta > 1/2$.

5.1. A key lemma

Let $(A, \mathcal{A})$ be a measurable space. For $f \in \Sigma$ and $T > 0$, consider an $(A, \mathcal{A})$-valued random variable ξ_T^f, continuous positive random processes $(V_T^f(x), x \in [-r, r])$, $(\tilde{V}_T^f(x), x \in [-r, r])$, $(U_T^f(x), x \in [-r, r])$, and random measures $\mathcal{Y}_T^f$, $\tilde{\mathcal{Y}}_T^f$ and $\mathcal{Z}_T^f$ on $[-r, r]$, defined on the same probability space. Assume that

$$\mathcal{Y}_T^f(\mathrm{d}x) = f(x)\mathrm{d}x + \Lambda^{-1/2} V_T^f(x)^{1/2} B(\mathrm{d}x), \tag{17}$$

$$\tilde{\mathcal{Y}}_T^f(\mathrm{d}x) = f(x)\mathrm{d}x + \Lambda^{-1/2} \tilde{V}_T^f(x)^{1/2} B(\mathrm{d}x), \tag{18}$$

$$\mathcal{Z}_T^f(\mathrm{d}x) = f(x)\mathrm{d}x + U_T^f(x)^{1/2} B'(\mathrm{d}x), \tag{19}$$

where $\Lambda \overset{\mathcal{L}}{=} |\mathcal{N}(0, 1)|$, B and B' are Gaussian centred with intensity dx, and Λ, B, B', $(\xi_T^f, V_T^f, \tilde{V}_T^f, U_T^f)$ are independent.

Lemma 2. *For each $T > 0$, consider a non-empty subset Σ_T of Σ and assume that the following conditions are fullfilled:*

(i) $\sup_{x\in[-r,r]} |V_T^f(x)/\tilde{V}_T^f(x) - 1| \to 0$ *in probability, uniformly in $f \in \Sigma_T$.*
(ii) $\inf_{x\in[-r,r]} U_T^f(x)\}/\sup_{x\in[-r,r]} V_T^f(x) \to \infty$ *in probability, uniformly in $f \in \Sigma_T$.*
(iii) *The probability measures on $C([-r, r], \mathbb{R})$,*

$$\mathcal{L}\left(\frac{V_T^f}{\inf_{x\in[-r,r]}\tilde{V}_T^f(x)}\right), \qquad T > 0,\ f \in \Sigma_T,$$

are tight.
(iv) $\liminf_T \inf_{\varepsilon>0, f\in\Sigma_T} P(\inf_{x\in[-r,r]} V_T^f(x) > \varepsilon) = 1.$

Then the experiment generated by the observation of $(\xi_T^f, V_T^f, \tilde{V}_T^f, U_T^f, \mathcal{Y}_T^f)$, and the one generated by $(\xi_T^f, V_T^f, \tilde{V}_T^f, U_T^f, \tilde{\mathcal{Y}}_T^f, \mathcal{Z}_T^f)$, both with parameter space Σ_T, are asymptotically equivalent.

Proof. It is clear that the experiment generated by the observation of $(\xi_T^f, V_T^f, \tilde{V}_T^f, U_T^f, \tilde{\mathcal{Y}}_T^f, \mathcal{Z}_T^f)$ is more informative than the one generated by $(\xi_T^f, V_T^f, \tilde{V}_T^f, U_T^f, \tilde{\mathcal{Y}}_T^f)$. Moreover, if $(V_T^f, \tilde{V}_T^f, U_T^f)$ satisfy conditions (i)–(iv) then so does $(\tilde{V}_T^f, V_T^f, U_T^f)$. Therefore it suffices to show that the experiment generated by the observation of $(\xi_T^f, V_T^f, \tilde{V}_T^f, U_T^f, \mathcal{Y}_T^f)$ is asymptotically more informative than the one generated by the observation of $(\xi_T^f, V_T^f, \tilde{V}_T^f, U_T^f, \tilde{\mathcal{Y}}_T^f, \mathcal{Z}_T^f)$.

We will split $\mathcal{Y}_T^f$ into two parts. Introduce the following subsets of $[-r, r]$:

$$I_T = \bigcup_{i=1}^{n_T}\left[(i-1)\frac{2r}{n_T}, (i-1+a_T)\frac{2r}{n_T}\right], \qquad J_T = \bigcup_{i=1}^{n_T}\left[(i-1+a_T)\frac{2r}{n_T}, i\frac{2r}{n_T}\right],$$

where $n(T)$ is a positive integer and a_T a real number in (0, 1) which will be specified later. Define the functions Φ_T and Ψ_T from $[-r, r]$ into $[-r, r]$ by

$$\Phi_T(x) = \frac{1}{a_T}\int_0^x 1_{I_T}(y)\mathrm{d}y, \qquad \Psi_T(x) = \frac{1}{1-a_T}\int_0^x 1_{J_T}(y)\mathrm{d}y,$$

and the random measures $\hat{\mathcal{Y}}_T^f$ and $\hat{\mathcal{Z}}_T^f$ on $[-r, r]$ by

$$\hat{\mathcal{Z}}_T^f(\varphi) = \frac{1}{a_T}\int \varphi\circ\Phi_T(x)1_{I_T}(x)\mathcal{Y}_T^f(\mathrm{d}x),$$

$$\hat{\mathcal{Y}}_T^f(\varphi) = \frac{1}{1-a_T}\int \varphi\circ\Psi_T(x)1_{J_T}(x)\mathcal{Y}_T^f(\mathrm{d}x).$$

From representation (17) it follows that

$$\hat{\mathcal{Y}}_T^f(\mathrm{d}x) = f \circ \Psi_T^{-1}(x)\mathrm{d}x + \frac{1}{\sqrt{1-a_T}}\Lambda^{-1/2}(V_T^f \circ \Psi_T^{-1}(x))^{1/2} N_T(\mathrm{d}x), \tag{20}$$

$$\hat{\mathcal{Z}}_T^f(\mathrm{d}x) = f \circ \Phi_T^{-1}(x)\mathrm{d}x + \frac{1}{\sqrt{a_T}}\Lambda^{-1/2}(V_T^f \circ \Phi_T^{-1}(x))^{1/2} M_T(\mathrm{d}x) \tag{21}$$

where $N_T(\varphi) := \sqrt{1-a_T}B[1_{J_T}(\varphi \circ \Psi_T)]$ and $M_T(\varphi) := \sqrt{a_T}B[1_{I_T}(\varphi \circ \Phi_T)]$ are two independent Gaussian random measures with intensity dx, independent of $(\Lambda, \xi_T^f, V_T^f, \tilde{V}_T^f, U_T^f)$.

Observe that Λ is a function of $(V_T^f, \mathcal{Y}_T^f)$ since $\Lambda^{-1/2}\int_{-r}^r V_T^f(x)\mathrm{d}x$ is the quadractic variation of $\mathcal{Y}_T^f$ on $[-r, r]$. Consider two independent centred Gaussian random measures W_1 and W_2 on $[-r, r]$ with intensity dx, independent of $(\xi_T^f, V_T^f, \tilde{V}_T^f, U_T^f)$. We define

$$\overline{\mathcal{Z}}_T^f(\mathrm{d}x) = \hat{\mathcal{Z}}_T^f(\mathrm{d}x) + \left(U_T^f(x)^{1/2} - \frac{1}{\sqrt{a_T}}\Lambda^{-1/2}(V_T^f \circ \Phi_T^{-1}(x))^{1/2}\right)^+ W_1(\mathrm{d}x),$$

$$\overline{\mathcal{Y}}_T^f(\mathrm{d}x) = \hat{\mathcal{Y}}_T^f(\mathrm{d}x) + \frac{\Lambda^{-1/2}}{\sqrt{1-a_T}}\left(\rho(f, T)^{1/2}\tilde{V}_T^f(x)^{1/2} - (V_T^f \circ \Psi_T^{-1}(x))^{1/2}\right) W_2(\mathrm{d}x),$$

where

$$\rho(f, T) = 1 \wedge \sup_{x\in[-r,r]} \frac{V_T^f \circ \Phi_T^{-1}(x)}{\tilde{V}_T^f(x)}.$$

A regular version of the conditional law of $(\overline{\mathcal{Y}}_T^f - \hat{\mathcal{Y}}_T^f, \overline{\mathcal{Z}}_T^f - \hat{\mathcal{Z}}_T^f)$ with respect to the σ-field generated by $(\xi_T^f, V_T^f, \tilde{V}_T^f, U_T^f, \mathcal{Y}_T^f)$ is

$$R\left(0, \frac{\Lambda^{-1}}{1-a_T}((\xi_T^{1/2}\tilde{V}^{1/2} - (V \circ \Psi_T^{-1})^{1/2}))\otimes R\left(0, \left|\left(U^{1/2} - \frac{\Lambda^{-1/2}}{\sqrt{a_T}}(V \circ \Phi_T^{-1})^{1/2}\right)^+\right|^2\right)\right),$$

which does not depend on the parameter f. Here, we denote by $R(f, v)$ the law on $(G, \mathcal{G})$ of the Gaussian measure with drift f and intensity v, for two continuous functions f and $v \geqslant 0$ on $[-r, r]$. Consequently, the experiment generated by the observation of $(\xi_T^f, V_T^f, \tilde{V}_T^f, U_T^f, \mathcal{Y}_T^f)$ is more informative than the one generated by $(\xi_T^f, V_T^f, \tilde{V}_T^f, U_T^f, \overline{\mathcal{Y}}_T^f, \overline{\mathcal{Z}}_T^f)$.

We choose a_T, $T > 0$, such that $a_T \to 0$ and

$$\inf_{f\in\Sigma_T} \overline{P}\left(\inf U_T^f \geqslant \frac{1}{a_T}\Lambda^{-1}\sup V_T^f\right) \to 1, \tag{22}$$

which is possible thanks to condition (ii). Due to condition (iv), there exists $\varepsilon_T > 0$ such that $\inf_{f\in\Sigma_T} P(\inf V_T^f \geqslant \varepsilon_T) \to 1$ as $T \to \infty$. Moreover, choose n_T such that

$$\varepsilon_T^{-2}\int_{-r}^r (f - f\circ\Psi_T^{-1})^2 \to 0, \qquad \varepsilon_T^{-2}\int_{-r}^r (f - f\circ\Phi_T^{-1})^2 \to 0 \tag{23}$$

uniformly in $f \in \Sigma$, a choice which is possible since

$$|\Phi_T^{-1}(x) - x| \leqslant \frac{2r}{n_T}, \qquad |\Psi_T^{-1}(x) - x| \leqslant \frac{2r}{n_T} \tag{24}$$

and Σ_T is relatively compact for the topology of uniform convergence ($\beta > 1/2$). Let us show that the experiment generated by the observation of $(\xi_T^f, V_T^f, \tilde{V}_T^f, U_T^f, \bar{\mathcal{Y}}_T^f, \bar{\mathcal{Z}}_T^f)$ and the one generated by the observation of $(\xi_T^f, V_T^f, \tilde{V}_T^f, U_T^f, \tilde{\mathcal{Y}}_T^f, \mathcal{Z}_T^f)$ are asymptotically total variation equivalent, which will complete the proof.

From (20), it follows that a regular version of the conditional law of $(\bar{\mathcal{Y}}_T^f, \bar{\mathcal{Z}}_T^f)$ with respect to the σ-field generated by $(\xi_T^f, V_T^f, \tilde{V}_T^f, U_T^f, \Lambda)$ is

$$R\left(f \circ \Psi_T^{-1}, \frac{\rho(f, T)}{1 - a_T}\Lambda^{-1}\tilde{V}_T^f\right) \otimes R\left(f \circ \Phi_T^{-1}, U_T^f \vee \left(\frac{1}{a_T}\Lambda^{-1}V_T^f \circ \Phi_T^{-1}\right)\right). \qquad (25)$$

Using (22) and the well-known upper bound

$$\|R(f_1, v) - R(f_2, v)\|_{\mathrm{TV}} \leqslant 2\sqrt{2}\left(1 - \exp\left(-\frac{1}{8}\int_{-r}^{r}\frac{(f_1 - f_2)^2}{v}\right)\right)^{1/2},$$

together with (23), it is easily seen that the total variation distance between (25) and

$$R\left(f, \frac{\rho(f, T)}{1 - a_T}\Lambda^{-1}\tilde{V}_T^f\right) \otimes R\left(f, U_T^f\right)$$

converges to 0 in probability, uniformly in $f \in \Sigma_T$. Integrating with respect to Λ, we deduce that the conditional law of $(\bar{\mathcal{Y}}_T^f, \bar{\mathcal{Z}}_T^f)$ with respect to the σ-field generated by $(\xi_T^f, V_T^f, \tilde{V}_T^f, U_T^f)$ is close to

$$Q\left(f, \frac{\rho(f, T)}{1 - a_T}\tilde{V}_T^f\right) \otimes R\left(f, U_T^f\right)$$

for the total variational norm, in probability. In view of the definition of $Q(f, v)$, we clearly have

$$\|Q(f, \rho\tilde{v}) - Q(f, \tilde{v})\|_{\mathrm{TV}} \leqslant \|\mathcal{L}(\rho\Lambda^{-1}) - \mathcal{L}(\Lambda^{-1})\|_{\mathrm{TV}}$$
$$= \|\mathcal{N}(0, 1/\rho^2) - \mathcal{N}(0, 1)\|_{\mathrm{TV}} \to 0 \qquad \text{as } \rho \to 1.$$

Since, furthermore, $\rho(f, T)/(1 - a_T) \to 1$ in probability uniformly in $f \in \Sigma_T$ by (i), (iii) and (24), we deduce that

$$\|\mathcal{L}(\xi_T^f, V_T^f, \tilde{V}_T^f, U_T^f, \bar{\mathcal{Y}}_T^f, \bar{\mathcal{Z}}_T^f) - \mathcal{L}(\xi_T^f, V_T^f, \tilde{V}_T^f, U_T^f, \tilde{\mathcal{Y}}_T^f, \mathcal{Z}_T^f)\|_{\mathrm{TV}} \to 0$$

uniformly in $f \in \Sigma_T$. □

5.2. Proof of Theorem 2

Recall that experiment $\tilde{\mathbb{E}}^{x_0,T}$ is defined by (15). We choose $T' = T^q$ with $q = 3/4$ satisfying the assumption of Proposition 3, in virtue of which it suffices to show that $\tilde{\mathbb{E}}^{x_0,T}$ and $\mathbb{G}^T$ are asymptotically equivalent.

Consider B, $\tilde{B}$, Λ, Y_T^f, Z_t^f and $\hat{m}_{f,T}$ as in Section 2.3. Assume, moreover, that the diffusion X^{f,x_0} given by (1) is defined on the same probability space and that B, $\tilde{B}$, Λ and

X^{f,x_0} are independent. Let $\tilde{Y}_T^f$ be defined by (15). Then $\mathbb{G}^T$ corresponds with the observation of (Y_T^f, Z_T^f) and $\tilde{\mathbb{E}}^{x_0,T}$ corresponds with the observation of $((X_t^{f,x_0})_{t\leq T'}, \tilde{Y}_T^f)$.

First we prove that $\tilde{\mathbb{E}}^{x_0,T}$ is asymptotically more informative than $\mathbb{G}^T$. Applying Lemma 2 with $\Sigma_T = \Sigma$, $\xi_T^f = (X_t^{f,x_0})_{t\leq T'}$, $V_T^f = \tilde{V}_T^f = (T-T')^{-1/2} m^{-1}_{\hat{f}_{T'}\circ X^{f,x_0}}$ and $U_T^f = u_T^2$ (assumption (ii) holds since $T^{1/4}u_T \to \infty$), yields that $\tilde{\mathbb{E}}^{x_0,T}$ is asymptotically equivalent to the experiment generated by $((X_t^{f,x_0})_{t\leq T'}, \tilde{Y}_T^f, Z_T^f)$.

Next, we apply Lemma 2 with $\Sigma_T = \Sigma$, $\xi_T^f = ((X_t^{f,x_0})_{t\leq T'}, Z_T^f)$, $V_T^f = (T - T')^{-1/2} m^{-1}_{\hat{f}_{T'}\circ X^{f,x_0}}$ and $\tilde{V}_T^f(x) = \hat{m}_{f,T}(x)^{-1}$. Conditions (i)–(iv) are easily verified thanks to Proposition 2 and the fact that $u_T \to 0$. It follows that the experiment generated by $((X_t^{f,x_0})_{t\leq T'}, \tilde{Y}_T^f, Z_T^f)$ is asymptotically equivalent to the one generated by $((X_t^{f,x_0})_{t\leq T'}, \tilde{Y}_T^f, Z_T^f)$ and consequently asymptotically more informative than $\mathbb{G}^T$.

It remains to show that $\mathbb{G}^T$ is asymptotically more informative than $\tilde{\mathbb{E}}^{x_0,T}$. Let $\tilde{\Lambda} \stackrel{\mathcal{L}}{=} |\mathcal{N}(0,1)|$ independent of $(\Lambda, B, \tilde{B})$ (and consequently of (Y_T^f, Z_T^f) as well). Clearly, $\mathbb{G}^T$ is of the same type as the experiment generated by $(\tilde{\Lambda}, Y_T^f, Z_T^f)$. Lemma 2 applied to $\xi_T^f = (\tilde{\Lambda}, Z_T^f)$, $V_T = \tilde{V}_T = T^{-1/2}(\hat{m}_{f,T})^{-1}$ and $U_T = \tilde{\Lambda}^{-1}(T')^{-1/2}(\tilde{m}_{f,T})^{-1}$ yields that $\mathbb{G}^T$ is asymptotically equivalent to the experiment given by the observation of $(\tilde{\Lambda}, Z_T^f, Y_T^f, \overline{Z}_T^f)$, where

$$\overline{Z}_t^f(\mathrm{d}x) = f(x)\mathrm{d}x + (T')^{-1/4}\tilde{\Lambda}^{-1/2}[\hat{m}_{f,T}]^{-1/2}\overline{B}(\mathrm{d}x)$$

and $\overline{B}$ is Gaussian independent of $(B, \tilde{B}, \Lambda, \tilde{\Lambda})$.

At this stage, we need a preliminary estimator of f constructed from Y_T^f. Using the notation of Section 4.1, for all $T > 0$ there exists a continuous process $(\hat{g}_T(x), x \in [-r, r])$ defined on $(G, \mathcal{G})$ such that, for all continuous functions f and $v \geq 0$ on $[-r, r]$, $(\hat{g}_T)$ is a $Q(f, v)$-modification of

$$\left(\int_{-r}^{r} K_{h(T)}(y-x)Y(\mathrm{d}y),\ x \in [-r, r]\right)$$

where Y denotes the canonical process on G. Moreover, along the same lines as the proof of Proposition 2, $\hat{g}_T \circ Y_T^f$ satisfies properties (12) and (14) of Proposition 2. We can apply Lemma 2 to $\xi_T^f = (\tilde{\Lambda}, Z_T^f, Y_T^f)$, $V_T^f(x) = (T')^{-1/2}[\hat{m}_{f,T}(x)]^{-1}$, $\tilde{V}_T^f(x) = (T')^{-1/2} m_{\hat{g}_T\circ Y_T^f}(x)^{-1}$ and $\mathcal{Y}_T^f = \overline{Z}_T^f$. It follows that $\mathbb{G}^T$ is asymptotically equivalent to the observation of $(\tilde{\Lambda}, Z_T^f, Y_T^f, \tilde{Z}_T^f)$, where

$$\tilde{Z}_T^f(\mathrm{d}x) = f(x)\mathrm{d}x + (T')^{-1/4}\tilde{\Lambda}^{-1} m_{\hat{g}_T\circ Y_T^f}(x)^{-1/2}\overline{B}(\mathrm{d}x).$$

To finish the proof, it suffices to show that the experiment generated by $(Z_T^f, Y_T^f, \overline{Z}_T^f)$ is asymptotically equivalent to the one generated by the observation of $(Z_T^f, Y_T^f, (X_t^{f,x_0})_{t\leq T'})$ since we have already seen that this last experiment is asymptotically equivalent to $\tilde{\mathbb{E}}^{x_0,T}$. We will apply Lemma 1 with $\Theta = \Sigma$,

$$\begin{aligned}
(A, \mathcal{A}) &= (G\times G, \mathcal{G}\otimes\mathcal{G}), & P^f &= P_T^f = \mathcal{L}((Z_T^f, Y_T^f))\\
(A_1, \mathcal{A}_1) &= (C(\mathbb{R}_+, \mathbb{R}), \mathcal{F}^0_{T'}), & N_1^f(z, y, \mathrm{d}\omega_1) &= P^f_{x_0,T'}(\mathrm{d}\omega_1),\\
(A_2, \mathcal{A}_2) &= (G, \mathcal{G}), & N_2^f(z, y, \mathrm{d}\omega_2) &= Q(f, (T')^{-1/2} m^{-1}_{\hat{g}_T(y)})(\mathrm{d}\omega_2).
\end{aligned}$$

Let $1/8 < p < \beta/2(2\beta+1)$ and set

$$\mathcal{V}(z, y) = \left\{ f \in \Sigma : \int_{-r}^{r} |f - \hat{g}_T(y)|^2 \leqslant T^{-2p}, \sup_{x \in [-r, r]} \left| \int_0^x (f - \hat{g}_T(y)) \right| \leqslant T^{-p} \right\}.$$

Due to the result for (12) in Proposition 2 applied to $\hat{g}_T(y)$ under P_T^f, we have $P_T^f(f \notin \mathcal{V}) \to 0$ as $T \to \infty$, uniformly with respect to $f \in \Sigma$. Moreover, since $p/q > 1/2 - 2p/q$ $(q = 3/4)$, we can apply Corollary 2 (with $\rho + K$ in place of ρ and T' in place of T) and deduce that $\sup_{(z,y) \in B_K} \Delta(\mathbb{F}_1(z, y), \mathbb{F}_2(z, y)) \to 0$ as $T \to \infty$, where $B_K = \{(z, y) \in G \times G \| \|\{\hat{g}_T(y)\|_{\beta,2,2} \leqslant K\}$. Since K is arbitrarily large, upper bound (16) yields the result.

5.3. Proof of Theorem 1

By Lemma 2 applied to $V_T^f = \tilde{V}_T^f = m_{f_0}^{-1}$, $\mathbb{F}_{\Sigma_T(f_0)}^{T, f_0}$ is asymptotically equivalent to the experiment with parameter space $\Sigma_T(f_0)$ generated by the observation of $(Y_{f_0,T}^f, Z_T^f)$. Lemma 2, again with $\xi_T^f = Z_T^f$, $V_T^f = m_{f_0}^{-1}$, $\tilde{V}_T^f = \hat{m}_{f,T}$, gives the result.

6. Proof of Proposition 1

6.1. Notation and structure of the proof

Recall that the likelihood ratio $Z_T^{f_0, f}$ is given by (8), where X denotes the canonical process. First, we introduce additional notation in order to write the likelihood ratio in a proper way. We define the right-continuous filtration $(\mathcal{F}_t)_{t \geqslant 0}$ by $\mathcal{F}_t = \bigcap_{s>t} \mathcal{F}_s^0$. For $f \in \Sigma$, we have

$$m_f(x) = \frac{2 \exp[2H_f(x)]}{\exp[2H_f(-r)] + \exp[2H_f(r)]}, \tag{26}$$

where $H_f(x) = \int_0^x f(y) \mathrm{d}y$. Define

$$\Phi_f(x) = \int_0^x \exp(-2H_f(y)) \mathrm{d}y, \qquad \Phi_{f,x_0}(x) = \Phi_f(x) - \Phi_f(x_0).$$

One can easily verify that $\Phi_f(X_t) = \Phi_f(X_0) + \int_0^t \Phi'_f(X_s) \mathrm{d}W_s^f$; thus if

$$A_t^f = \int_0^t \Phi'_f(X_s)^2 \, \mathrm{d}s, \qquad B_t^f = \Phi_{f,X_0}(X_{\tau_t^f}),$$

where $(\tau_t^f)_{t \geqslant 0}$ is the inverse of $(A_t^f)_{t \geqslant 0}$, then B^f is a $(P_{x_0}^f, (\mathcal{F}_{\tau_t^f})_t)$-standard Brownian motion according to the Dambis–Dubins–Schwarz theorem (see, for example, Revuz and Yor 1991). Hence, for each $T > 0$, the process $(\beta_u^{f,T})_{u \geqslant 0}$ defined by

$$\beta_u^{f,T} = T^{-1/2} B_{Tu}^f$$

is a $(P_{x_0}^f, (\mathcal{F}_{\tau_{Tu}^f})_u)$-standard Brownian motion as well.

If ψ is a Borel function on $\mathbb{R}$, set

$$V_T^f[\psi](u) := \frac{1}{T^{1/2}} \int_0^{Tu} \psi(B_s^f)\mathrm{d}s = T^{1/2} \int_0^u \psi(T^{1/2}\beta_v^{f,T})\mathrm{d}v, \tag{27}$$

$$M_T^f[\psi](u) := \frac{1}{T^{1/4}} \int_0^{Tu} \psi(B_s^f)\mathrm{d}B_s^f = T^{1/4} \int_0^u \psi(T^{1/2}\beta_v^{f,T})\mathrm{d}\beta_v^{f,T}. \tag{28}$$

Remark that the stochastic integral (28) is also meaningful if ψ is a random function which is $\mathcal{F}_0 \otimes \mathcal{B}(\mathbb{R})$-measurable. Using this notation, we can write

$$\log Z_T^{f_0,f} = T^{1/4} M_T^{f_0}[\psi_{f,f_0}]\Big(\alpha_1^{f_0,T}\Big) - \frac{T^{1/2}}{2} V_T^{f_0}[\psi_{f,f_0}^2]\Big(\alpha_1^{f_0,T}\Big), \tag{29}$$

where

$$\psi_{f,f_0}(y) = \frac{f - f_0}{\Phi'_{f_0}} \circ \Phi_{f_0,X_0}^{-1}(y) \quad \text{and} \quad \alpha_u^{f,T} = \frac{A_{Tu}^f}{T}.$$

Finally, we introduce $\gamma_T^f(\psi)$, the Dambis–Dubins–Schwarz Brownian motion of the $P_{x_0}^f$-martingale $M_T^f[\psi]$. In particular, we have

$$M_T^f[\psi](u) = \gamma_T^f(\psi)(\langle M_T^f[\psi]\rangle_u).$$

To conclude this section, let us briefly sketch our proof strategy. In (29), we have

$$M_T^{f_0}[\psi_{f,f_0}](\alpha_1^{f_0,T}) = \gamma_T^{f_0}(\psi_{f,f_0}\Big((V_T^{f_0}[\psi_{f,f_0}^2](\alpha_1^{f_0,T})\Big).$$

We 'wish' to construct an approximation of $V_T^{f_0}[\psi_{f,f_0}^2](\alpha_1^{f_0,T})$ that is independent of $\gamma_T^{f_0}(\psi_{f,f_0})$. We proceed as follows. First, we have

$$\begin{aligned} V_T^{f_0}[\psi_{f,f_0}^2]_u &= \int_{\mathbb{R}} L(\beta^{f_0,T})_u^{T^{-1/2}y} \psi_{f,f_0}(y)^2 \,\mathrm{d}y, \\ &\simeq L(\beta^{f_0,T})_u^0 \int_{\mathbb{R}} \psi_{f,f_0}(y)^2 \,\mathrm{d}y. \end{aligned}$$

Next, introducing the $P_{x_0}^{f_0}$-martingale

$$N_u^{f_0,T} = \int_0^u \Big(\exp(2H_{f_0}(r))1_{\{T^{1/2}\beta_v^{f_0,T} > \Phi_{f_0,X_0}(r)\}} - \exp(2H_{f_0}(-r))1_{\{T^{1/2}\beta_v^{f_0,T} \leqslant \Phi_{f_0,X_0}(-r)\}}\Big)\mathrm{d}\beta_v^{f_0,T} \tag{30}$$

and

$$\tilde{\alpha}_u^{f_0,T} = \inf\{v \geqslant 0 | \langle N^{f_0,T}\rangle_v > u\}, \tag{31}$$

with the usual convention $\inf \varnothing = +\infty$, we obtain the following:

(i) $\tilde{\alpha}_u^{f_0,T}$ is 'close' to $\alpha_u^{f_0,T}$.

(ii) $N^{f_0,T}_{\tilde{\alpha}^{f_0},T}$ is independent of $\gamma_T^{f_0}(\psi_{f,f_0})$ by Knight's theorem.

(iii) We can approximate $L(\beta^{f_0,T})^0_{\tilde{\alpha}_u^{f_0,T}}$ by a random variable measurable with respect to the σ-filed generated by $(N^{f_0,T}_{\tilde{\alpha}_u^{f_0,T}})_{u\geqslant 0}$ (by an application of Skorokhod's lemma) and thus independent of $\gamma_T^{f_0}(\psi_{f,f_0})$.

6.2. Some approximations

This subsection is devoted to a key lemma (Lemma 6) which is the essential ingredient in obtaining (9). Define the probability measure $P_{x_0}^{f,f_0,T} = Z_T^{f_0,f}.P_{x_0}^{f_0}$, where $Z_T^{f_0,f}$ is given by (8).

First we study the modulus of continuity of the local time of $\beta_u^{f_0,T}$ under $P_{x_0}^{f,f_0,T}$.

Lemma 3. *There exist constants C_p, $p \geqslant 0$, such that, for all $T > 0$, $x_0 \in \mathbb{R}$, and $(f_0, f) \in \Sigma \times \Sigma$:*

(i) *if U_1 and U_2 are two $P_{x_0}^{f_0}$-a.s. finite stopping times with respect to the filtration $(\mathcal{F}_{\tau^{f_0}_{Tu}})_{u\geqslant 0}$ then, for all $y \in \mathbb{R}$,*

$$\mathrm{E}_{x_0}^{f,f_0,T}[|L(\beta^{f_0,T})^y_{U_1} - L(\beta^{f_0,T})^y_{U_2}|^p] \leqslant C_p \mathrm{E}_{x_0}^{f,f_0,T}[|U_1 - U_2|^{p/2}];$$

(ii) *for all $(x, y) \in [\Phi_{f_0,x_0}(-r), \Phi_{f_0,x_0}(r)]$, for all $u \geqslant 0$,*

$$\mathrm{E}_{x_0}^{f,f_0,T}[\sup_{v\leqslant u}|L(\beta^{f_0,T})^x_v - L(\beta^{f_0,T})^y_v|^p] \leqslant C_p\left(u^{p/4}|x-y|^{p/2} + u^{p/2}\sup_{z\in[-r,r]}\left|\int_0^z (f-f_0)(z')\mathrm{d}z'\right|^p\right).$$

Proof. Elementary computations yield that

$$L(\beta^{f_0,T})^y_u = L(\beta^{f,T})^{T^{-1/2}\Phi_{f,X_0}\circ\Phi^{-1}_{f_0,X_0}(T^{1/2}y)}_{\alpha^{f,T}\circ\theta_u^{f_0,T}}\left(\frac{\Phi'_{f_0}}{\Phi'_f}\right)\circ\Phi^{-1}_{f_0,X_0}(T^{1/2}y),$$

where $\theta_u^{f_0,T} := T^{-1}\tau^{f_0}_{Tu}$ is the inverse of $\alpha^{f_0,T}$.

Since $\alpha_1^{f_0,T}$ is an $(\mathcal{F}_{\tau^{f_0}_{Tu}})_{u\geqslant 0}$-stopping time, we can (and will) assume that $U_i \leqslant \alpha_1^{f_0,T}$, $i = 1, 2$, or $U_i \geqslant \alpha_1^{f_0,T}$, $i = 1, 2$. First,

$$|\alpha^{f,T}\circ\theta^{f_0,T}_{U_1} - \alpha^{f,T}\circ\theta^{f_0,T}_{U_2}| \leqslant C|U_1 - U_2|.$$

Second, for f_0, $f \in \Sigma$, we have

$$\frac{x}{C} \leqslant \Phi_f(x) \leqslant Cx \quad \text{and} \quad \left|\frac{\Phi'_{f_0}}{\Phi'_f} - 1\right| \leqslant C \sup_{x\in[-r,r]}\left|\int_0^x (f-f_0)(y)\mathrm{d}y\right|;$$

therefore, it suffices to show (i) and (ii) with $(L(\beta^{f,T})^y_{\alpha^{f,T}\circ\theta_u^{f_0,T}})_{y,u}$ in place of $(L(\beta^{f_0,T})^y_u)_{y,u}$. The proof is easily completed using well-known properties of the local time of a Brownian motion, and the following facts:

- $\alpha^{f,T} \circ \theta^{f_0,T}_{U_i}$ is a stopping time with respect to $(\mathcal{F}_{\tau^f_{Tu}})_{u\geqslant 0}$, since $\tau^f_t \uparrow\uparrow +\infty$ as $t \uparrow\uparrow \infty$;
- under $P^{f,f_0,T}_{x_0}$ the process $(\beta^{f,T}_{u\wedge\alpha_1^{f,T}})_{u\geqslant 0}$ is a standard Brownian motion with respect to $(\mathcal{F}_{\tau^f_{Tu}})_{u\geqslant 0}$, stopped at time $\alpha_1^{f,T}$;
- for all $(\mathcal{F}_{\tau^{f_0}_{Tu}})_{u\geqslant 0}$-stopping times U such that $U \geqslant \alpha_1^{f_0,T}$,

$$L(\beta^{f_0,T})^y_U = L(\beta^{f_0,T})^y_{\alpha_1^{f_0,T}} + L(\hat{\beta}^{f_0,T})^{y+\beta^{f_0,T}_{\alpha_1^{f_0,T}}}_{U-\alpha_1^{f_0,T}}$$

where

$$\hat{\beta}^{f_0,T}_u := \beta^{f_0,T}_{\alpha_1^{f_0,T}+u} - \beta^{f_0,T}_{\alpha_1^{f_0,T}}$$

is a $(P^{f,f_0,T}_{x_0}, (\mathcal{F}_{T\alpha_1^{f_0,T}+Tu})_u))$-standard Brownian motion. □

Our second result states that $\alpha_1^{f_0,T}$ is close to $\tilde{\alpha}_1^{f_0,T}$, which is technically more convenient to work with.

Lemma 4. *We have*

$$\mathrm{E}^{f,f_0,T}_{x_0}(|\tilde{\alpha}_1^{f_0,T} - \alpha_1^{f_0,T}|^p) \leqslant C_p T^{-p/2}$$

for all $0 \leqslant p < \infty$, $T > 0$, $x_0 \in \mathbb{R}$, $(f_0, f) \in \Sigma \times \Sigma$.

Proof. We first need an auxiliary result (see the appendix for a proof).

Lemma 5. *Let a and b be two real-valued functions on $\mathbb{R}_+$ satisfying the following conditions:*

(*i*) *a and b are non-decreasing, continuous and vanish at time 0, and $a_\infty = b_\infty = +\infty$.*
(*ii*) *The function a is of class C^1, and there exists $\varepsilon > 0$ such that $a'_u \geqslant \varepsilon$ for all $u \geqslant 0$.*

Then

$$|a_v^{-1} - b_v^{-1}| \leqslant \varepsilon^{-1}\varphi(\rho_v^\varepsilon) \leqslant \rho_v^\varepsilon - v/\varepsilon,$$

where

$$\rho_v^\varepsilon = \inf\{u \geqslant 0 : v + \varphi(u) < \varepsilon u\} \quad \textit{and} \quad \varphi(u) = \sup_{v\leqslant u}|a_v - b_v|.$$

On the one hand, the occupation time formula yields

$$\begin{aligned} S_u^{f_0,T} &:= \theta_u^{f_0,T} - \langle N^{f_0,T}\rangle_u \\ &= T^{-1/2}\int_{\Phi_{f_0,X_0(-r)}}^{\Phi_{f_0,X_0(r)}} L(\beta^{f_0,T})_u^{T^{-1/2}y} \frac{1}{\Phi'_{f_0}\circ\Phi^{-1}_{f_0,X_0}(y)^2}\,\mathrm{d}y; \end{aligned}$$

thus, in view of Lemma 3(i), $\mathrm{E}^{f,f_0,T}_{x_0}(|S_u^{f_0,T}|^q) \leqslant C_q u^{q/2} T^{-q/2}$ for all $0 \leqslant q < \infty$. On the other hand, since $(\partial/\partial u)\theta_u^{f_0,T} \geqslant \varepsilon > 0$, we may apply Lemma 5 to the functions $a_u = \theta_u^{f_0,T}$ and $b_u = \langle N^{f_0,T}\rangle_u$ to obtain

$$|\tilde{\alpha}_1^{f_0,T} - \alpha_1^{f_0,T}| \leqslant \rho(f_0, T) - 1/\varepsilon$$

with $\rho(f_0, T) = \inf\{u : 1 + S_U^{f_0,T} < \varepsilon u\}$. Since $\{\rho(f_0, T) > 1/\varepsilon + u\} \subset \{S_{1/\varepsilon+u}^{f_0,T} \geqslant \varepsilon u\}$, Chebyshev's inequality yields

$$\begin{aligned}
\mathrm{E}_{x_0}^{f,f_0,T}(|\tilde{\alpha}_1^{f_0,T} - \alpha_1^{f_0,T}|^p) &\leqslant \mathrm{E}_{x_0}^{f,f_0,T}((\rho(f_0, T) - 1/\varepsilon)^p),\\
&\leqslant T^{-p/2} + \int_{T^{-1/2}}^{\infty} pu^{p-1}(\varepsilon u)^{-3p}\mathrm{E}_{x_0}^{f,f_0,T}(|S_{1/\varepsilon+u}^{f_0,T}|^{3p})\mathrm{d}u,\\
&\leqslant C_p T^{-p/2}.
\end{aligned}$$

□

Next, we have an approximation result for $V_T^{f_0}$ at time $\alpha_1^{f_0,T}$. Define the random variables

$$\lambda(f, T) = \frac{\exp(2H_f(r))L(\beta^{f,T})_{\tilde{\alpha}_1^{f,T}}^{T^{-1/2}\Phi_{f,X_0}(r)} + \exp(2H_f(-r))L(\beta^{f,T})_{\tilde{\alpha}_1^{f,T}}^{T^{-1/2}\Phi_{f,X_0}(-r)}}{\exp(2H_f(r)) + \exp(2H_f(-r))} \tag{32}$$

(see (26) for the definition of H_f).

Lemma 6. *There exist constants C_p, $p \geqslant 0$, such that, for all $x_0 \in \mathbb{R}$, $T > 0$, f_0, $f \in \Sigma$, for all random function ψ $\mathcal{F}_0 \otimes \mathcal{B}(\mathbb{R})$-measurable and supported by $[\Phi_{f,X_0}(-r), \Phi_{f,X_0}(r)]$, we have $P_{x_0}^{f,f_0,T}$-a.s.:*

$$\mathrm{E}_{x_0}^{f,f_0,T}\left(\left|V_T^{f_0}[\psi](\alpha_1^{f_0,T}) - \lambda(f_0, T)\int_{\mathbb{R}}\psi(x)\mathrm{d}x\right|^p\right)$$
$$\leqslant C_p\left(\int_{\mathbb{R}}|\psi(x)|\mathrm{d}x\right)^p\left(T^{-p/4} + \sup_{x\in[-r,r]}\left|\int_0^x (f - f_0)(y)\mathrm{d}y\right|^p\right).$$

Proof. Let $\mu(f, T)$ be the random variable defined by (32) with $\alpha_1^{f,T}$ in place of $\tilde{\alpha}_1^{f,T}$. Lemma 3(i) applied to $U_1 = \alpha_1^{f_0,T}$ and $U_2 = \tilde{\alpha}_1^{f_0,T}$, together with Lemma 4, yields $\mathrm{E}_{x_0}^{f,f_0,T}(|\lambda(f_0, T) - \mu(f_0, T)|^p) \leqslant C_p T^{-p/4}$ for all $p \geqslant 0$. Therefore it suffices to prove Lemma 6 with $\mu(f_0, T)$ in place of $\lambda(f_0, T)$

From the occupation time formula it follows that

$$V_T^{f_0}[\psi](\alpha_1^{f_0,T}) = \int_{\Phi_{f_0,X_0}(-r)}^{\Phi_{f_0,X_0}(r)} L(\beta^{f_0,T})_{\alpha_1^{f_0,T}}^{T^{-1/2}}\psi(y)\mathrm{d}y.$$

Furthermore, since $\alpha_1^{f_0,T} \leqslant C$, the conclusion easily follows using Lemma 3(ii) and the definition of $\mu(f_0, T)$. □

6.3. Construction of the Gaussian measure

We define here the random variable $\Lambda(f_0, T)$ and the Gaussian random measure $Y_T^{f_0}$, and we prove properties (i) and (ii) of Proposition 1. Set

$$\Lambda(f, T) = \Big(-\inf_{u\leqslant 1} N^{f,T}_{\tilde{\alpha}^{f,T}_u} - T^{-1/2}\exp(2H_f(r))\Phi^-_{f,X_0}(r) - T^{-1/2}\exp(2H_f(-r))\Phi^+_{f,X_0}(-r)\Big)^+. \tag{33}$$

Lemma 7. *We have*

$$\Lambda(f, T) = \frac{\exp(2H_f(r)) + \exp(2H_f(-r))}{2}\lambda(f, T).$$

Proof. From Tanaka's formula it follows that

$$\begin{aligned}
&\exp(2H_f(r))(\beta^{f,T}_u - T^{-1/2}\Phi_{f,X_0}(r))^+ + \exp(2H_f(-r))(\beta^{f,T}_u - T^{-1/2}\Phi_{f,X_0}(-r))^- \\
&\quad = N^{f,T}_u + T^{-1/2}\exp(2H_f(r))\Phi^-_{f,X_0}(r) + T^{-1/2}\exp(2H_f(-r))\Phi^+_{f,X_0}(-r) \\
&\qquad + \tfrac{1}{2}\exp(2H_f(r))L(\beta^{f,T})^{T^{-1/}\Phi_{f,X_0}(r)}_u + \tfrac{1}{2}\exp(2H_f(-r))L(\beta^{f,T})^{T^{-1/2}\Phi_{f,X_0}(-r)}_u.
\end{aligned}$$

Skorokhod's lemma (Revuz and Yor, 1991, p. 222) implies that

$$\tfrac{1}{2}\exp(2H_f(r))L(\beta^{f,T})^{T^{-1/}\Phi_{f,X_0}(r)}_{\tilde{\alpha}^{f_0,T}_1} + \tfrac{1}{2}\exp(2H_f(-r))L(\beta^{f,T})^{T^{-1/2}\Phi_{f,X_0}(-r)}_{\tilde{\alpha}^{f_0,T}_1}$$

is equal to

$$\sup_{u\leqslant\tilde{\alpha}^{f,T}_1}\{0\vee[-N^{f,T}_u - T^{-1/2}\exp(2H_f(r))\Phi^-_{f,X_0}(r) - T^{-1/2}\exp(2H_f(-r))\Phi^+_{f,X_0}(-r)]\},$$

which is the desired result. □

At this stage, the preceding results give the construction of an approximation of $\log Z_T^{f_0,f}$ sufficiently accurate for our purpose, and which has the law of $\log \tilde{Z}_T^{f_0,f}$ – recall (10). However, this equality in law is valid for fixed f only, and we need it as a random process indexed by f. Therefore, we introduce the following discretization technique: let $n \geqslant 1$ be a positive integer. For $k = 1, \ldots, n$, define

$$\begin{aligned}
&a(f_0, T)_i = \Phi_{f_0,X_0}(-r + i2r/n),\\
&\varphi_i^{f_0,T}(y) = (a(f_0, T)_i - a(f_0, T)_{i-1})^{-1/2}1_{(a(f_0,T)_{i-1},a(f_0,T)_i)}(y).
\end{aligned}$$

The functions $\varphi_i^{f_0,T}$ enjoy the following properties:

$$\int_{\mathbb{R}} \varphi_i^{f_0,T}(y)^2 \mathrm{d}y = 1;$$

$$\mathrm{supp}(\varphi_i^{f_0,T}) \subset [\Phi_{f_0,X_0}(-r), \Phi_{f_0,X_0}(r)];$$

$$\mathrm{supp}(\varphi_i^{f_0,T}) \cap \mathrm{supp}(\varphi_j^{f_0,T}) = \varnothing, \qquad \text{if } i \neq j.$$

Let $\tilde{B}(\mathrm{d}x)$ be a centred Gaussian random measure on $\mathbb{R}$ with intensity $\mathrm{d}x$ and independent of X. For a Borel function ψ on $\mathbb{R}$, we set

$$\langle \psi, \varphi_i^{f_0,T} \rangle = \int_{\mathbb{R}} \varphi_i^{f_0,T}(y)\psi(y)\mathrm{d}y, \qquad \mathrm{proj}_T^{f_0} \psi = \sum_{i=1}^{n} \left\langle \psi, \varphi_i^{f_0,T} \right\rangle \varphi_i^{f_0,T},$$

$$\chi_T^{f_0}(\psi) = \lambda(f_0, T)^{-1/2} \sum_i \langle \psi, \varphi_i^{f_0,T} \rangle \gamma_T^{f_0}(\varphi_i^{f_0,T})_{\lambda(f_0,T)} + \tilde{B}(\psi - P_T^{f_0}\psi),$$

and

$$Y_T^{f_0}(h) = \left(\frac{2}{\exp(2H_{f_0}(r)) + \exp(2H_{f_0}(-r))} \right)^{1/2} \chi_T^{f_0}\left(\frac{h}{\Phi'_{f_0}} \circ \Phi_{f_0,X_0}^{-1} \right), \qquad h \in L^2([-r, r]).$$

Lemma 8. *Under $P_{x_0}^{f_0}$, $Y_T^{f_0}$ is a centred Gaussian random measure on $[-r, r]$ with intensity m_{f_0}, independent of $\Lambda(f_0, T)$. Moreover, $\mathcal{L}(\Lambda(f_0, T)|P_{x_0}^{f_0}) \xrightarrow{\mathrm{TV}} |\mathcal{N}(0, 1)|$ as $T^{-1/2}x_0 \to 0$ uniformly in $f_0 \in \Sigma$.*

Proof. Since $\langle N^{f_0,T}, M_T^{f_0}[\varphi_i^{f_0,T}] \rangle = 0$ and $\langle M_T^{f_0}[\varphi_i^{f_0,T}], M^{f_0,T}[\varphi_j^{f_0,T}] \rangle = 0$ for $i \neq j$, Knight's theorem (Revuz and Yor 1991, p. 172) implies that the $(n+1)$-dimensional process

$$\left(N_{\tilde{\alpha}_u^{f_0,T}}^{f_0,T}, \left(\gamma_T^{f_0}(\varphi_i^{f_0,T})_u \right)_{1 \leqslant i \leqslant n} \right)_{u \geqslant 0}$$

is a Brownian motion independent of $\mathcal{F}_0$, under $P_{x_0}^{f_0}$. Moreover, Lemma 7 shows that $(\Lambda(f_0, T), \lambda(f_0, T))$ is a function of $((N_{\tilde{\alpha}_u^{f_0,T}}^{f_0,T})_u, X_0)$. Thus, under $P_{x_0}^{f_0}$, $\chi_T^{f_0}(\mathrm{d}y)$ is a Gaussian centred random measure on $\mathbb{R}$ with intensity $\mathrm{d}y$, independent of $\lambda(f_0, T)$. The first claim follows because of $\Phi'_{f_0}(x) = m_{f_0}(0)/m_{f_0}(x)$.

The second assertion follows from the fact that $\mathcal{L}(-\inf_{u \leqslant 1} N_{\tilde{\alpha}_u^{f_0,T}}^{f_0,T} | P_{x_0}^{f_0}) = |\mathcal{N}(0, 1)|$ and $\exp(2H_{f_0}(r))\Phi_{f_0,x_0}^{-}(r) + \exp(2H_{f_0}(-r))\Phi_{f_0,x_0}^{+}(-r) \leqslant C(1 + |x_0|)$. □

6.4. Conclusion of the proof

We are now ready to prove (9), from which Proposition 1 will follow. For simplicity, we will assume without loss of generality that Assumption A holds with $p_1 < 1/4$ and $p_2 < 1/4$.

Lemma 9 (Scheffé, uniform version). *For $n \geqslant 1$, let X_n be a non-negative random variable defined on a measurable space equipped with a probability measure P_n. If $X_n \to 1$ in P_n-probability and $\int X_n \mathrm{d}P_n \to 1$, then $\int |X_n - 1| \mathrm{d}P_n \to 0$.*

Proof. We write $|X_n - 1| = (1 - X_n)1_{X_n \leqslant 1} + (X_n - 1)1_{X_n > 1}$. The sequence $(1 - X_n)1_{X_n \leqslant 1}$ converges to 0 in P_n-probability and is dominated by 1, so $\int (1 - X_n)1_{X_n \leqslant 1} \mathrm{d}P_n \to 0$. Then

$$\int (X_n - 1)1_{X_n > 1} \mathrm{d}P_n = \int (X_n - 1)\mathrm{d}P_n + \int (1 - X_n)1_{X_n \leqslant 1} \mathrm{d}P_n$$

and the first term tends to 0 by assumption. The conclusion follows. □

Recall that $P_{x_0}^{f, f_0, T} = Z_T^{f_0, f}.P_{x_0}^{f_0}$. Applying Scheffé's lemma to the family of random variables $(\tilde{Z}_T^{f_0, f}/Z_T^{f_0, f})_{T \geqslant 0}$ and the family of probability measures $(P_{x_0, T}^{f_0, f})_{T \geqslant 0}$, (9) will follow if

$$\frac{\tilde{Z}_T^{f_0, f}}{Z_T^{f_0, f}} \to 1 \qquad \text{in } P_{x_0, T}^{f, f_0}\text{-probability as } T \to \infty, \tag{34}$$

uniformly in $x_0 \in \mathbb{R}$, $f_0 \in \Sigma$ and $f \in \mathcal{V}_T(f_0)$. In view of decomposition (29), Lemma 7 and the definition of m_{f_0} (see (26)), expression (34) can be deduced from

$$T^{1/2}\left\{ V_T^{f_0}\left[\psi_{f, f_0}^2\right](\alpha_1^{f_0, T}) - \lambda(f_0, T)\int_{\mathbb{R}} \psi_{f, f_0}^2(y)\mathrm{d}y \right\} \to 0, \tag{35}$$

$$T^{1/4}\left\{ M_T^{f_0}\left[\psi_{f, f_0}\right](\alpha_1^{f_0, T}) - \Lambda(f_0, T)^{1/2} Y_T^{f_0}(f - f_0) \right\} \to 0 \tag{36}$$

in $P_{x_0}^{f, f_0, T}$-probability uniformly in $x_0 \in \mathbb{R}$, $f_0 \in \Sigma$ and $f \in \mathcal{V}_T(f_0)$.

Further, according to Lemma 6, the random variables

$$\eta_T^{-1}\varepsilon_T^{-2}\left\{ V_T^{f_0}\left[\psi_{f, f_0}^2\right](\alpha_1^{f_0, T}) - \lambda(f_0, T)\int_{\mathbb{R}} \psi_{f, f_0}^2(y)\mathrm{d}y \right\},$$

are tight under $P_{x_0}^{f, f_0, T}$ for $T \geqslant 1$ and $f_0 \in \Sigma$, uniformly in (x_0, T, f_0, f) with $f \in \mathcal{V}_T(f_0)$. Since $\eta_T \varepsilon_T^2 T^{1/2} \to 0$ by Assumption A, we have (35). It remains to prove (36).

By Girsanov's theorem,

$$M_T^{f_0}[\psi](u) = T^{1/4} V_T^{f_0}\left[\psi\psi_{f, f_0}\right]\left(\alpha_1^{f_0} \wedge u\right) + \tilde{M}_T^{f, f_0}[\psi](u), \tag{37}$$

with $\tilde{M}_T^{f, f_0}[\psi]$ being a continuous $P_{x_0}^{f, f_0, T}$-martingale. Thus we have

$$\gamma_T^{f_0}(\psi)_u = T^{1/4} V_T^{f_0}\left[\psi\psi_{f, f_0}\right]\left(\alpha_1^{f_0} \wedge \langle M_T^{f_0}[\psi]\rangle_u^{-1}\right) + \tilde{\gamma}_T^{f, f_0}(\psi)_u, \tag{38}$$

where $\langle M_T^{f_0}[\psi]\rangle^{-1}$ is the right-continuous generalized inverse of $\langle M_T^{f_0}[\psi]\rangle$ and $\tilde{\gamma}_T^{f, f_0}(\psi)(u) = M_T^{f_0}[\psi](\langle M_T^{f_0}[\psi]\rangle_u^{-1})$. In this subsection we write φ_i in place of $\varphi_i^{f_0, T}$. On the one hand, from the definition of $Y_T^{f_0}$, we have

$$\Lambda(f_0, T)^{1/2} Y_T^{f_0}(f - f_0) = \sum_{i=1}^{n} \langle \psi_{f, f_0}, \varphi_i \rangle \gamma_T^{f_0}(\varphi_i)_{\lambda(f_0, T)} + \lambda(f_0, T)^{1/2} \tilde{B}(\psi_{f, f_0} - \mathrm{proj}_T^{f_0} \psi_{f, f_0}).$$

From (38), we readily derive

$$\Lambda(f_0, T)^{1/2} Y_T^{f_0}(f - f_0) = T^{1/4} \sum_{i=1}^n \langle \psi_{f,f_0} \varphi_i \rangle V_T^{f_0}[\varphi_i \psi_{f,f_0}](\alpha_1^{f_0,T} \wedge \langle M_T^{f_0}(\varphi_i)\rangle^{-1}_{\lambda(f_0,T)})$$
$$+ \sum_{i=1}^n \langle \psi_{f,f_0}, \varphi_i \rangle \tilde{\gamma}_T^{f,f_0}(\varphi_i)_{\lambda(f_0,T)} + \lambda(f_0, T)^{1/2} \tilde{B}(\psi_{f,f_0} - \mathrm{proj}_T^{f_0} \psi_{f,f_0}. \tag{39}$$

On the other hand, we have

$$M_T^{f_0}[\psi_{f,f_0}](\alpha_1^{f_0,T}) = T^{1/4} V_T^{f_0}[\psi^2_{f,f_0}](\alpha_1^{f_0,T}) + \sum_{i=1}^n \langle \psi_{f,f_0}, \varphi_i \rangle \tilde{\gamma}_T^{f,f_0}(\varphi_i)_{\langle M_T^{f_0}(\varphi_i)\rangle(\alpha_1^{f_0,T})}$$
$$+ \tilde{M}_T^{f,f_0}[\psi_{f_0,f} - \mathrm{proj}_T^{f_0} \psi_{f_0,f}](\alpha_1^{f_0,T}).$$

In view of these two decompositions and (35), in order to prove (36) it suffices to show that the following convergences hold in $P_{x_0}^{f,f_0,T}$-probability, uniformly:

$$T^{1/2}\left(\sum_{i=1}^n \langle \psi_{f,f_0}, \varphi_i \rangle V_T^{f_0}[\varphi_i \psi_{f,f_0}]\Big(\alpha_1^{f_0,T} \wedge \langle M_T^{f_0}(\varphi_i)\rangle^{-1}_{\lambda(f_0,T)}\Big) - \lambda(f_0, T) \int \psi^2_{f,f_0}(x)\mathrm{d}x \right) \to 0, \tag{40}$$

$$T^{1/4} \sum_{i=1}^n \langle \psi_{f,f_0}, \varphi_i \rangle \Big(\tilde{\gamma}_T^{f,f_0}(\varphi_i)(\lambda(f_0, T)) - \tilde{\gamma}_T^{f,f_0}(\varphi_i)\Big(\langle M_T^{f_0}[\varphi_i]\rangle_{\alpha_1^{f_0,T}}\Big)\Big) \to 0, \tag{41}$$

$$T^{1/4} \tilde{B}\Big(\psi_{f,f_0} - \mathrm{proj}_T^{f_0} \psi_{f,f_0}\Big) \to 0, \tag{42}$$

$$T^{1/4} \tilde{M}_T^{f,f_0}\Big[\psi_{f,f_0} - \mathrm{proj}_T^{f_0} \psi_{f,f_0}\Big]\Big(\alpha_1^{f_0,T}\Big) \to 0. \tag{43}$$

Since the choice of n is free, we now let $n = n_T$. Using the fact that $\|\psi_{f,f_0}\|_{\beta,2,2} \leqslant C$, we readily derive the bounds

$$\int_{\mathbb{R}} \Big| \psi_{f,f_0}(x) - \mathrm{proj}_T^{f_0} \psi_{f,f_0}(x) \Big|^2 \mathrm{d}x \leqslant C\, n_T^{-2\beta}.$$

Choose n_T of polynomial growth such that

$$T^{1/2} \varepsilon_T n_T^{-\beta} \to 0 \quad \text{and} \quad T^{1/2} n_T^{-2\beta} \to 0.$$

We readily derive (42). Applying Lemma 6 to $\psi = \psi_{f,f_0} - \mathrm{proj}_T^{f_0} \psi_{f,f_0}$, we deduce that $T^{1/2} \langle M_T^{f,f_0} - \mathrm{proj}_T^{f_0} \psi_{f,f_0}] \rangle(\alpha_1^{f_0,T}) \to 0$ which implies (43). It remains to prove (40) and (41).

Let us first focus on (40). Since $V_T^{f_0}[\varphi_i \varphi_j] = 0$ if $i \neq j$, we have

$$V_T^{f_0}[\varphi_i \psi_{f,f_0}] = \langle \varphi_i, \psi_{f_0,f} \rangle V_T^{f_0}[\varphi_i^2] + V_T^{f_0}[\varphi_i(\psi_{f,f_0} - \mathrm{proj}_T^{f_0} \psi_{f,f_0})].$$

Moreover, $V_T^{f_0}[\varphi_i^2] = \langle M_T^{f_0}(\varphi_i) \rangle$ is increasing; therefore

$$V_T^{f_0}\left[\varphi_i^2\right]\left(\alpha_1^{f_0,T}\right) \wedge \lambda(f_0, T) \leqslant V_T^{f_0}\left[\varphi_i^2\right]\left(\alpha_1^{f_0,T} \wedge \langle M_T^{f_0}(\varphi_i)\rangle_{\lambda(f_0,T)}^{-1}\right)$$
$$\leqslant \lambda(f_0, T).$$

It follows that

$$V_T^{f_0}[\varphi_i\psi_{f,f_0}]\left(\alpha_1^{f_0,T} \wedge \langle M_T^{f_0}[\varphi_i]\rangle_{(f_0,T)}^{-1}\right) - \langle\varphi_i, \psi_{f,f_0}\rangle\lambda(f_0, T)$$
$$= O\left\{ V_T^{f_0}\left[\varphi_i\left|\psi_{f,f_0} - \mathrm{proj}_T^{f_0}\psi_{f,f_0}\right|\right]\left(\alpha_1^{f_0,T}\right) + |\langle\varphi_i, \psi_{f,f_0}\rangle| \left| V_T^{f_0}\left[\varphi_i^2\right]\left(\alpha_1^{f_0,T}\right) - \lambda(f_0, T)\right| \right\}.$$

Thus

$$\sum_{i=1}^{n}\langle\psi_{f,f_0}, \varphi_i\rangle V_T^{f_0}[\varphi_i\psi_{f,f_0}](\langle M_T^{f_0}(\varphi_i)\rangle_{\lambda(f_0,T)}^{-1})$$
$$= \lambda(f_0, T)\int_{\mathbb{R}} |\mathrm{proj}_T^{f_0}\psi_{f,f_0}(x)|^2 \mathrm{d}x + O\left\{ V_T^{f_0}\left[\mathrm{proj}_T^{f_0}|\psi_{f,f_0}|\,|\psi_{f,f_0} - \mathrm{proj}_T^{f_0}\psi_{f,f_0}|\right](\alpha_1^{f_0,T}).\right.$$
$$\left. + \sum_i \langle\varphi_i, \psi_{f,f_0}\rangle^2 \left| V_T^{f_0}\left[\varphi_i^2\right]\left(\alpha_1^{f_0,T}\right) - \lambda(f_0, T)\right| \right\}.$$

By the Cauchy–Schwarz inequality, we have

$$\left| \int_{\mathbb{R}} \mathrm{proj}_T^{f_0}\psi_{f,f_0}(x)|^2\,\mathrm{d}x - \int_{\mathbb{R}} \psi_{f,f_0}(x)^2\,\mathrm{d}x \right| \leqslant C\,n_T^{-\beta}\varepsilon_T = o(T^{-1/2}).$$

Using the occupation time formula and Lemma 3(ii) (with $U_1 = \alpha_1^{f_0,T}$ and $U_2 = 0$), together with the fact that $\alpha_1^{f_0,T} \leqslant C$, we derive

$$\mathrm{E}_{x_0}^{f,f_0,T}\left(V_T^{f_0}\left[\mathrm{proj}_T^{f_0}|\psi_{f,f_0}|\left|\psi_{f,f_0} - \mathrm{proj}_T^{f_0}\psi_{f,f_0}\right|\right](\alpha_1^{f_0,T}) \right)$$
$$\leqslant C\int_{\mathbb{R}} \mathrm{proj}_T^{f_0}|\psi_{f,f_0}|\left|\psi_{f,f_0} - \mathrm{proj}_T^{f_0}\psi_{f,f_0}\right| \leqslant C\varepsilon_T n_T^{-\beta}.$$

Hence (40) will follow from

$$T^{1/2}\sum_i \langle\varphi_i, \psi_{f,f_0}\rangle^2 \left| V_T^{f_0}\left[\varphi_i^2\right]\left(\alpha_1^{f_0,T}\right) - \lambda(f_0, T)\right| \to 0.$$

Lemma 6 applied to $\psi = \varphi_i^2$ entails that the expectation of the above sum is less than $CT^{1/2}\varepsilon_T^2(T^{-1/4} + \eta_T) \to 0$ since $2p_1 + p_2 > 1/2$.

We now prove (41). Let $q \in (0, p_2)$ such that $2p_1 + q > 1/2$. Lemma 6 applied to $\psi = \varphi_i^2$ yields that, for all $p \geqslant 0$,

$$P_{x_0}^{f,f_0,T}\left(|\lambda(f_0, T) - \langle M_T^{f_0}(\varphi_i)\rangle_{\alpha_1^{f_0,T}}| > T^{-q}\right) \leqslant C_p T^{p(q-p_2)}.$$

Thus

$$P_{x_0}^{f,f_0,T}\left(\sup_i |\lambda(f_0, T) - \langle M_T^{f_0}(\varphi_i)\rangle_{\alpha_1^{f_0,T}}| > T^{-q}\right) \leqslant C_p n_T T^{-p(p_2-q)}.$$

Define the random variables

$$\lambda(q, f_0, T) = T^{-q}\lfloor T^q \lambda(f_0, T)\rfloor,$$

where $\lfloor\cdot\rfloor$ denotes the integer part,

$$S(i, f_0, T) = (\lambda(q, f_0, T) - 2T^{-q} \vee \langle M_T^{f_0}(\varphi_i)\rangle_{\tilde{\alpha}_1^{f_0,T}} \wedge (\lambda(q, f_0, T) + 2T^{-q}).$$

For p large enough we have $n_T T^{-p(p_2-q)} \to 0$, hence

$$\sum_{i=1}^{n(T)} \rangle \psi_{f,f_0}, \varphi_i\rangle \left(\tilde{\gamma}^{f,f_0}(\varphi_i)_{S(i,f_0,T)} - \tilde{\gamma}^{f,f_0}(\varphi_i)_{\langle M_T^{f_0}(\varphi_i)\rangle \circ \alpha_1^{f_0,T}}\right) \to 0.$$

Consequently, to prove (41), it is enough to show that $T^{1/4}D(f_0, f)_T \to 0$ and $T^{1/4}D'(f_0, f)_T \to 0$, where

$$D(f_0, f)_T = \sum_{i=1}^{n(T)} \langle \psi_{f,f_0}, \varphi_i\rangle \big(\tilde{\gamma}^{f,f_0}(\varphi_i)_{S(i,f_0,T)} - \tilde{\gamma}^{f,f_0}(\varphi_i)_{\lambda(q,f_0,T)}\big),$$

$$D'(f_0, f)_T = \sum_{i=1}^{n(T)} \langle \psi_{f,f_0}, \varphi_i\rangle \big(\tilde{\gamma}^{f,f_0}(\varphi_i)_{\lambda(f_0,T)} - \tilde{\gamma}^{f,f_0}(\varphi_i)_{\lambda(q,f_0,T)}\big).$$

On the one hand, since $N^{f_0,T}$ is also a $P_{x_0}^{f,f_0,T}$-martingale, in the same way as in Lemma 8, one shows that $(\tilde{\gamma}^{f,f_0}(\varphi_i))_i$ is an n_T-dimensional Brownian motion independent of $\lambda(f_0, T)$ under $P_{x_0}^{f,f_0,T}$. Therefore, we have

$$\begin{aligned} \mathrm{E}_{x_0}^{f,f_0,T}\big(D'(f_0, f)_T^2 | \lambda(f_0, T)\big) &= \sum_i \langle \psi_{f,f_0}, \varphi_i\rangle^2 |\lambda(f_0, T) - \lambda(q, f_0, T)|, \\ &\leqslant C\varepsilon_T^2 T^{-q}. \end{aligned}$$

Since $2p_1 + q > 1/2$, we deduce that $T^{1/4}D'(f_0, f)_T \to 0$.

On the other hand, for $l \in \mathbb{N}$, on the set $\{\lambda(q, f_0, T) = lT^{-q}\}$ we have $D(f_0, f)_T = D_l(f_0, f)_T$ with

$$D_l(f_0, f)_T := \sum_{i=1}^{n(T)} \langle \psi_{f,f_0}, \varphi_i\rangle \big(\tilde{\gamma}^{f,f_0}(\varphi_i)_{S(l,i,f_0,T)} - \tilde{\gamma}^{f,f_0}(\varphi_i)_{lT^{-q}}\big),$$

$$S(l, i, f_0, T) = (lT^{-q} - 2T^{-q}) \vee \langle M_T^{f_0}[\varphi_i]\rangle_{\tilde{\alpha}_1^{f_0,T}} \wedge (lT^{-q} + 2T^{-q}).$$

Hölder's inequality yields that, for all $p \geqslant 1$,

$$\mathrm{E}_{x_0}^{f,f_0,T}(|D(f_0, f)_T|^{2/p}) \leqslant \sum_{l=0}^{\infty} P_{x_0}^{f,f_0,T}[\lambda(q, f_0, T) = lT^{-q}]^{1-1/p}\, \mathrm{E}_{x_0}^{f,f_0,T}\big(|D_l(f_0, f)_T|^2\big)^{1/p}.$$

Furthermore,

$$\mathrm{E}_{x_0}^{f,f_0,T}\left(|D_l(f_0, f)_T|^2\right) \leqslant C\varepsilon_T^2 T^{-q}. \tag{44}$$

Indeed, the random variables

$$\tilde{\gamma}^{f,f_0}(\varphi_i)_{S(l,i,f_0,t)} - \tilde{\gamma}^{f,f_0}(\varphi_i)_{lT^{-q}}, \qquad 1 \leqslant i \leqslant n(T),$$

are uncorrelated since: first,

$$\tilde{\gamma}^{f,f_0}(\varphi_i)_{S(l,i,f_0,T)} - \tilde{\gamma}^{f,f_0}(\varphi_i)_{lT^{-q}} = \tilde{M}_T^{f,f_0}[\varphi_i](\sigma(l, i, f_0, T)) - \tilde{M}_T^{f,f_0}[\varphi_i]\Big(\langle M_T^{f_0}[\varphi_i]\rangle_{lT^{-q}}^{-1}\Big),$$

where

$$\sigma(l, i, f_0, T) = \langle M_T^{f_0}[\varphi_i]\rangle_{(l-2)^+T^{-q}}^{-1} \vee \alpha_1^{f_0,T} \wedge \langle M_T^{f_0}[\varphi_i]\rangle_{(l-2)T^{-q}}^{-1};$$

second, $\langle M_T^{f_0}[\varphi_i]\rangle_{lT^{-q}}^{-1}$ and $\sigma(l, i, f_0, T)$ are both stopping times with respect to the filtration $(\mathcal{F}_{\tau_{Tu}^{f_0}})_{u\geqslant 0}$; and finally,

$$\langle M_T^f[\varphi_i], M_T^{f_0}[\varphi_j]\rangle = 0 \qquad \text{if } i \neq j.$$

Recall that $\lambda(f_0, T) = m_{f_0}(0)\Lambda(f_0, T)$ by Lemma 7. In view of (33), on the set $\{\Lambda(f_0, T) > 0\}$ we have $\Lambda(f_0, T) = \tilde{\Lambda}(f_0, T) - a(T, f_0, X_0)$, where

$$\tilde{\Lambda}(f_0, T) = -\inf_{u\leqslant 1} N_{\tilde{\alpha}_u^{f_0,T}}^{f_0,T},$$

$$a(T, f_0, X_0) := T^{-1/2}\exp(2H_{f_0}(r))\Phi_{f_0,X_0}^-(r) + T^{-1/2}\exp(2H_{f_0}(-r))\Phi_{f_0,X_0}^+(-r) \geqslant 0.$$

For $l \geqslant 1$, we have

$$P_{x_0}^{f,f_0,T}[\lambda(q, f_0, T) = lT^{-q}] \leqslant P_{x_0}^{f,f_0,T}\left[\frac{lT^{-q}}{m_{f_0}(0)} \leqslant \tilde{\Lambda}(f_0, T) - a(T, f_0, x_0) \leqslant \frac{(l+1)T^{-q}}{m_{f_0}(0)}\right]$$
$$\leqslant CT^{-q}\exp\left(-\frac{l^2T^{-2q}}{2m_{f_0}(0)^2}\right)$$

because $\tilde{\Lambda}(f_0, T) = |\mathcal{N}(0, 1)|$ in law under $P_{x_0}^{f,f_0,T}$. Thus

$$\sum_{l=0}^{\infty} P_{x_0}^{f,f_0,T}[\lambda(q, f_0, T) = lT^{-q}]^{1-1/p} \leqslant C_p T^{q/p}.$$

In view of (44), to deduce (43), it suffices to pick p large enough so that $p_1 + (1/2 - 1/p)q > 1/4$. The proof of Proposition 1 is complete.

Appendix

A.1. Proof of Lemma 5.

First, assume that $a_u = u$ for all $u \geqslant 0$. For $v \geqslant 0$, we have $b \circ b_v^{-1} = v$, hence $|b_v^{-1} - v| = |b_v^{-1} - b \circ b_v^{-1}| \leqslant \varphi(b_v^{-1})$. Moreover, if $u \geqslant 0$ is such that $v + \varphi(u) < u$, then, for all $t \in (v + \varphi(u), u)$, we have $b_t \geqslant t - \varphi(t) \geqslant t - \varphi(u) > v$, thus $b_v^{-1} \leqslant t$; consequently, $b_v^{-1} \leqslant v + \varphi(u)$. We deduce that, for real such u, we have $|b_v^{-1} - v| \leqslant \varphi(u)$. Therefore, if $\{u : v + \varphi(u) \leqslant u\} \neq \varnothing$, by continuity of φ, we have $|b_v^{-1} - v| \leqslant \varphi(\rho_v^1)$; if $\{u : v + \varphi(u) < u\} = \varnothing$, then $\varphi(\infty) = \infty$ and $\rho_v^1 = \infty$. The result follows.

General case: Let $v \geqslant 0$. Since a is continuous and increasing, we can write

$$|b_\beta^{-1} - a_v^{-1}| = |(a^{-1} \circ b)_{a_v^{-1}}^{-1} - a_v^{-1}|.$$

But $|a^{-1} \circ b_u - u| \leqslant \varepsilon^{-1}\varphi(u)$ for all u, because

$$\varphi(u) \geqslant |b_u - a_u| = |a(a^{-1} \circ b_u) - a_u| \geqslant \varepsilon |a^1 \circ b_u - u|.$$

We derive $|b_v^{-1} - a_v^{-1}| \leqslant \varepsilon^{-1}\varphi(\tilde{\rho}_{a_v^{-1}})$ with

$$\tilde{\rho}_s = \inf\{u : s + \varepsilon^{-1}\varphi(u) < u\}.$$

It remains to note that $a_v^{-1} \leqslant \varepsilon^{-1} v$ and that $\tilde{\rho}_{\varepsilon^{-1}v} = \rho_v^\varepsilon$.

A.2. Proof of Proposition 2.

We will use the notation and results of Section 6. In particular, recall that the random variables $\Lambda(f, T)$ are given by (33) and satisfy $\Lambda(f, T) \to |\mathcal{N}(0, 1)|$ in law under $P_{x_0}^f$ uniformly in $f \in \Sigma$, by Lemma 8. Moreover, we have that $\Lambda(f, T) \leqslant -\inf_{u \leqslant 1} N_{\tilde{\alpha}_u^f}^{f,T} = |\mathcal{N}(0, 1)|$ in law under $P_{x_0}^f$. For a test function φ, we write $U_T[\varphi] = T^{-1/2} \int_0^T \varphi(X_s) \mathrm{d}s$ and $V_T[\varphi] = T^{-1/4} \int_0^T \varphi(X_s) \mathrm{d}W_s^f$. In what follows, the constants depend only on β, ρ, r and the function K.

Lemma 10. (i) *For all $p \geqslant 0$, $x_0 \in \mathbb{R}$, $f \in \Sigma$ and for all functions φ with support in $[-2r, 2r]$, we have*

$$\mathrm{E}_{x_0}^f \left(\left| U_T[\varphi] - \Lambda(f, T) \int_{-2r}^{2r} \varphi(y) m_f(y) \mathrm{d}y \right|^{2p} \right) \leqslant C_p T^{-p/2} \left(\int_{-2r}^{2r} |\varphi(y)| \mathrm{d}y \right)^{2p}.$$

(ii) *We have*

$$\limsup_T \sup_{(f, x_0) \in \Sigma \times [-R_T, R_T]} P_{x_0}^f \left(\inf_{x \in [-r, r]} U_T(K_{h(T)}(\cdot - x)) \leqslant \varepsilon \right) \to 0 \qquad \textit{as } \varepsilon \to 0.$$

Proof. One can easily verify that

$$U_T[\varphi] = V_T^f[\psi_f](\alpha_1^{f,T}) \quad \text{with} \quad \psi_f = \left(\frac{\varphi}{(\Phi_f')^2}\right) \circ \Phi_{f,X_0}^{-1}.$$

Since ψ_f is supported by $[\Phi_{f,X_0}(-2r), \Phi_{f,X_0}(2r)]$, (i) follows from Lemmas 6 (with $f = f_0$ and $2r$ in place of r) and 7.

Next, the occupation time formula yields

$$V_T^f[\psi_f](\alpha_1^{f,T}) = \int_{\mathbb{R}} L(\beta^{f,T})_{\alpha_1^{f,T}}^{y/\sqrt{T}} \psi_f(y)\mathrm{d}y,$$

$$\geqslant \int_{\mathbb{R}} \psi_f(y)\mathrm{d}y \inf_{y \in [\Phi_{f,X_0}(-2r), \Phi_{f,X_0}(2r)]} L(\beta^{f,T})_\varepsilon^{y/\sqrt{T}},$$

because $\alpha_1^{f,T} \geqslant \varepsilon$ for some constant $\varepsilon > 0$. Taking $\varphi = K_{h(T)}(\cdot - x)$, we deduce that there exists a constant $C > 0$ such that, for T large enough,

$$\inf_{x \in [-r,r]} U_T(K_{h(T)}(\cdot - x)) \geqslant C \inf_{y \in [\Phi_{f,X_0}(-2r), \Phi_{f,X_0}(2r)]} L(\beta^{f,T})_\varepsilon^{y/\sqrt{T}}.$$

Moreover, since $\beta^{f,T}$ is a standard Brownian motion under $P_{x_0}^f$, we have

$$\inf_{y \in [\Phi_{f,X_0}(-2r), \Phi_{f,X_0}(2r)]} L(\beta^{f,T})_\varepsilon^{y/\sqrt{T}} \to |\mathcal{N}(0, \varepsilon)|$$

in law under $P_{x_0}^f$, uniformly in $f \in \Sigma$. The proof of (ii) is easily completed. □

A.2.1. Proof of (12)

First, note that by Lemma 10(ii), it suffices to show the tightness of the family

$$(\sqrt{T})^{2\beta/(2\beta+1)} \int_{-r}^{r} U_T[K_{h(T)}(\cdot - x)]^2 (\hat{f}_T(x) - f(x))^2 \,\mathrm{d}x, \qquad T > 0.$$

Moreover,

$$U_T[K_{h(T)}(\cdot - x)](\hat{f}_T(x) - f(x))$$
$$= U_T[K_{h(T)}(\cdot - x) f] - U_T[K_{h(T)}(\cdot - x)] f(x) + T^{-1/4} V_T[K_{h(T)}(\cdot - x)].$$

We further decompose the bias term $U_T[K_{h(T)}(\cdot - x) f] - U_T[K_{h(T)}(\cdot - x)] f(x)$ as

$$\Lambda(f, T) \int_{\mathbb{R}} K_{h(T)}(y - x)(f(y) - f(x)) m_f(y)\mathrm{d}y + R_T(x),$$

where

$$R_T(x) = U_T[K_{h(T)}(\cdot - x) f] - \Lambda(f, T) \int_{\mathbb{R}} K_{h(T)}(y - x) f(y) m_f(y)\mathrm{d}y$$
$$+ U_T[K_{h(T)}(\cdot - x)] - \Lambda(f, T) \int_{\mathbb{R}} K_{h(T)}(y - x) m_f(y)\mathrm{d}y.$$

We first bound the remainder term R_T. Applying Lemma 10(i), we obtain that

$$\mathrm{E}_{x_0}^f(R_T^2(x)) < CT^{-1/2}\left\{\left(\int K_{h(T)}(y-x)|f(y)|\mathrm{d}y\right)^2 + \left(\int K_{h(T)}(y-x)\mathrm{d}y\right)^2\right\}$$

$$\leqslant CT^{-1/2}\left\{\int K_{h(T)}(y-x)f(y)^2\mathrm{d}y + 1\right\}$$

by Jensen's inequality. Integrating with respect to x over $[-r, r]$ yields

$$\mathrm{E}_{x_0}^f\left(\int_{-r}^{r} R_T^2(x)\mathrm{d}x\right) \leqslant CT^{-1/2}\left(\int f^2 + 1\right).$$

The above term is thus asymptotically negligible. Using the bound

$$\int_{\mathbb{R}}\left(\int_{\mathbb{R}} K_{h(T)}(y-x)(f(y)-f(x))m_f(y)\mathrm{d}y\right)^2 \mathrm{d}x \leqslant Ch(T)^{2\beta},$$

together with the fact that $\Lambda(f, T)$ has a second-order moment uniformly bounded, we derive from the specification of $h(T)$ that the bias term has the right order.

Let us turn to the variance term. By definition of $h(T)$, it suffices to show that

$$h(T)\int_{-r}^{r} V_T^2[K_{h(T)}(\cdot - x)]\mathrm{d}x \tag{45}$$

is $P_{x_0}^f$-tight, uniformly in f. Since

$$\mathrm{E}_{x_0}^f(V_T^2(K_{h(T)}(\cdot - x))) = \mathrm{E}_{x_0}^f(U_T[K_{h(T)}(\cdot - x)^2]),$$

Lemma 10(i) applied to $\varphi = K_{h(T)}(\cdot - x)^2$ for $x \in [-r, r]$, together with the fact that

$$\int_{\mathbb{R}} K_{h(T)}(y-x)^2\,\mathrm{d}y \leqslant C/h(T),$$

implies that the expectation of (45) is bounded, which completes the proof of (12).

A.2.2. Proof of (13)

Using Lemma 10, it is easily seen that the random variables

$$T^{1/4}\int_{-r}^{r}\left|\frac{U_T[K_{h(T)}(x-\cdot)f]}{U_T[K_{h(T)}(x-\cdot)]} - \frac{\int K_{h(T)}(y-x)f(y)m_f(y)\mathrm{d}y}{\int K_{h(T)}(y-x)m_f(y)}\right|\mathrm{d}x$$

are tight under $P_{x_0}^f$, uniformly in $f \in \Sigma$. Furthermore, since

$$\int_{-r}^{r}\mathrm{d}x\left(\int \mathrm{d}y K_{h(T)}(y-x)(m_f(y)-m_f(x))\right)^2 \leqslant Ch(T)^{2(\beta+1)} = o(T^{-1/2}),$$

it suffices to show the tightness of

$$T^{1/4} \sup_{x\in[-r,r]} \left| \int_{-r}^{x} \left(\frac{\int K_{h(T)}(z-y)f(z)m_f(z)\mathrm{d}z}{m_f(y)} - f(y) \right) \mathrm{d}y \right|$$

and

$$\sup_{x\in[-r,r]} \left| \int_{-r}^{x} V_T[K_{h(T)}[K_{h(T)}(\cdot - y)]\mathrm{d}y \right|.$$

Set $G_f(x) = \int_{-r}^{x} f(y)m_f(y)\mathrm{d}y$. We have

$$\int_{-r}^{x} \frac{\int K_{h(T)}(z-y)f(z)m_f(z)\mathrm{d}z}{m_f(y)} \mathrm{d}y$$
$$= \int \mathrm{d}z\, G_f(z) \left(\frac{K_{h(T)}(z-x)}{m_f(x)} - \frac{K_{h(T)}(z+r)}{m_f(-r)} + \int_{-r}^{x} \mathrm{d}y\, K_{h(T)}(z-y) \frac{m'_f(y)}{m_f(y)^2} \right)$$

and

$$\sup_{x\in[-r,r]} \left| \int K_{h(T)}(z-x)(G_f(z) - G_f(x))\mathrm{d}z \right| \leqslant Ch(T)^{\beta+1};$$

hence

$$\int_{-r}^{x} \frac{\int K_{h(T)}(z-y)f(z)m_f(z)\mathrm{d}z}{m_f(y)} \mathrm{d}y = \frac{G_f(x)}{m_f(x)} - \frac{G_f(-r)}{m_f(-r)} + \int_{-r}^{x} \mathrm{d}y\, G_f(y) \frac{m'_f(y)}{m_f(y)^2} + O(h(T)^{\beta+1}),$$
$$= \int_{-r}^{x} f(y)\mathrm{d}y + o(T^{-1/4})$$

from the choice of $h(T)$.

Let us now treat the variance term. We have, by Fubini's theorem for stochastic integrals,

$$\int_{-r}^{x_1} V_T[K_{h(T)}(\cdot - y)]\mathrm{d}y - \int_{-r}^{x_2} V_T[K_{h(T)}(\cdot - y)\mathrm{d}y = T^{-1/4} \int_0^T \left(\int_{x_2}^{x_1} K_{h(T)}(X_s - y)\mathrm{d}y \right) \mathrm{d}W_s^f.$$

Therefore, for all $p \geqslant 0$, the Burkholder–Davis–Gundy inequality yields

$$\begin{aligned}
&\mathrm{E}^f_{x_0}\left\{\left|\int_{-r}^{x_1} V_T[K_{h(T)}(\cdot - y)]\mathrm{d}y - \int_{-r}^{x_2} V_T[K_{h(T)}(\cdot - y)]\mathrm{d}y\right|^p\right\} \\
&\leqslant C_p T^{-p/4}\mathrm{E}^f_{x_0}\left\{\left(\int_0^T\left(\int_{x_2}^{x_1} K_{h(T)}(X_s - y)\mathrm{d}y\right)^2\mathrm{d}s\right)^{p/2}\right\} \\
&\leqslant C_p\left(\int_{\mathbb{R}}\left(\int_{x_2}^{x_1} K_{h(T)}(z - y)\mathrm{d}y\right)^2\mathrm{d}z\right)^{p/2} \qquad \text{by Lemma 10(i),} \\
&\leqslant C_p|x_1 - x_2|^{p/2}.
\end{aligned}$$

The Kolmogorov criterion gives the result.

A.2.3. Proof of (14)

We will use the definition of $\|\cdot\|_{\beta,2,2}$ in terms of the modulus of continuity given in Section 2.1. Let us first show the tightness of the family

$$\|x \mapsto 1_{[-r,r]}(x)U_T[K_{h(T)}(\cdot - x)]\|_{\beta,2,2}, \qquad T > 0. \tag{46}$$

Clearly, a family of random functions $x \mapsto g_T(x)$ for $T > 0$, such that $(\int_{-r}^r |g'_T(x)|^2\mathrm{d}x,\ T > 0)$ is tight, is such that $(\|1_{[-r,r]}g_T\|_{\beta,2,2},\ T > 0)$ is also tight. By Lebesgue's theorem, the functions $x \mapsto U_T[K_{h(T)}(\cdot - x)]$ and $x \mapsto \int_{\mathbb{R}} K_{h(T)}(y - x)m_f(y)\mathrm{d}y$ are differentiable, with derivatives

$$h(T)^{-1}T^{-1/2}\int_0^T (K')_{h(T)}(X_s - x)f(X_s)\mathrm{d}s,$$

$$h(T)^{-1}\int_{\mathbb{R}} (K')_{h(T)}(y - x)f(y)m_f(y)\mathrm{d}y.$$

Applying Lemma 10(i) and the specification of $h(T)$, we see that the family

$$\int_{-r}^r \mathrm{d}x\left|\partial_x\left(U_T[K_{h(T)}(\cdot - x)f] - \Lambda(f,\, T)\int_{\mathbb{R}} K_{h(T)}(y - x)f(y)m_f(y)\mathrm{d}y\right)\right|^2,$$

$T > 0$, is tight uniformly in $f \in \Sigma$. Moreover, we clearly have that the Besov norm of $x \mapsto 1_{[-r,r]}(x)\int_{\mathbb{R}} K_{h(T)}(y - x)m_f(y)\mathrm{d}y$ is uniformly bounded, so (46) is proved.

Therefore, by Lemma 10(ii), it suffices to show the tightness of

$$\|x \mapsto 1_{[-r,r]}(x)U_T[K_{h(T)}(\cdot - x)f]\|_{\beta,2,2} \tag{47}$$

and

$$T^{-1/4}\|x \mapsto 1_{[-r,r]}(x)V_T[K_{h(T)}(\cdot - x)]\|_{\beta,2,2}. \tag{48}$$

Repeating the argument developed for (46), with $U_T[K_{h(T)}]$ replaced by $U_T[K_{h(T)}f]$, we readily obtain the tightness of (47).

Let us turn to the variance term. For this, it is enough to show the boundedness of

$$T^{-1}\mathrm{E}^f_{x_0}\left\{\int_{-2r}^{2r}\frac{\mathrm{d}h}{h^{1+2\beta}}\int_{-2r}^{2r}\left(\int_0^T[K_{h(T)}(X_s-x-h)-K_{h(T)}(X_s-x)]\mathrm{d}W^f_s\right)^2\mathrm{d}x\right\}.$$

Applying Lemma 10(i), the latter quantity is less than

$$CT^{-1/2}\int_{-2r}^{2r}\frac{\mathrm{d}h}{h^{1+2\beta}}\int_{-2r}^{2r}\mathrm{d}x\int_{-2r}^{2r}\mathrm{d}y\, m_f(y)[K_{h(T)}(y-x-h)-K_{h(T)}(y-x)]^2.$$

Pick $\alpha\in[1,2]$ such that $2\beta<\alpha<2\beta+1$. Using the fact that K is of class C^1, we see that the above term is of order $T^{-1/2}h(T)^{-\alpha}$, which goes to 0 as $T\to\infty$.

A.3. Proof of Lemma 1

By symmetry it suffices to show that

$$\delta(\mathbb{E}_1,\mathbb{E}_2)\leqslant\sup_{\omega\in B}\delta(\mathbb{F}_1(\omega),\mathbb{F}_2(\omega))+\sup_{\theta\in\Theta}P^\theta(\{\theta\notin\mathcal{V}\}\cup(A\backslash B)) \tag{49}$$

for all $B\in\mathcal{A}$. Since

$$\delta(\mathbb{E}_1,\mathbb{E}_2)=\sup\{\delta(\mathbb{E}_{1,\Theta'},\mathbb{E}_{2,\Theta'});\ \Theta'\text{finite subset of }\Theta\}$$

(see Le Cam 1986, Theorem 2.3.2; or Strasser 1985, Corollary 59.4) we can proceed as if Θ is finite and shall do so below.

Define the probability measure $\nu(\mathrm{d}\omega)$ on $(A,\mathcal{A})$ and the Markov kernel $\nu_i(\omega,\mathrm{d}\omega_i)$ from $(A,\mathcal{A})$ into $(A_i,\mathcal{A}_i)$ by

$$\nu(\mathrm{d}\omega)=\frac{1}{|\Theta|}\sum_{\theta\in\Theta}P^\theta(\mathrm{d}\omega),\qquad \nu_i(\omega,\mathrm{d}\omega_i)=\frac{1}{|\Theta|}\sum_{\theta\in\Theta}N_i^\theta(\omega,\mathrm{d}\omega_i).$$

By assumption (ii), there exists a countably generated sub-σ-field $\mathcal{A}'_i\subset\mathcal{A}_i$ which is $(N_i^\theta(\omega,\cdot),\theta\in\Theta)$-sufficient $\nu(\mathrm{d}\omega)$-a.s. ($i=1,2$). Since $\mathcal{A}'_i$ is countably generated, there exists a sequence of finite $\mathcal{A}_i$-measurable partitions of A_i, say $(B^n_{i,N})_{1\leqslant n\leqslant m(N,i)}$, $N\geqslant 1$, such that

$$\mathcal{A}'_i=\bigvee_{N\geqslant 1}\mathcal{A}'_{i,N}\quad\text{and}\quad\mathcal{A}'_{i,N}\subset\mathcal{A}'_{i,N+1},$$

where $\mathcal{A}'_{i,N}$ is the σ-field generated by $(B^n_{i,N})_{1\leqslant n\leqslant m(N,i)}$. Set

$$Z^\theta_{i,N}(\omega,\omega_i)=\sum_{n=1}^{m(N,i)}1_{B^n_{i,N}}(\omega_i)\frac{N_i^\theta(\omega,B^n_{i,n})}{\nu_i(\omega,B^n_{i,N})},$$

where, by convention, $0/0=0$. It is well known that for all $\theta\in\Theta$, $\omega\in A$, the sequence $Z^\theta_{i,N}(\omega,\cdot)$ converges $\nu_i(\omega,\cdot)$-a.s. and in $L^1(\nu_i(\omega,\cdot))$ to $Z^\theta_i(\omega,\cdot):=\limsup_N Z^\theta_{i,N}(\omega,\cdot)$, which is a version of the Radon–Nikodym derivative of $N_i^\theta(\omega, \cdot)$ with respect to $\nu_i(\omega, \cdot)$ restricted to $\mathcal{A}_i'$ and actually on the whole σ-field $\mathcal{A}_i$ by sufficiency.

Define the Markov kernel $N_{i,N}^\theta(\omega, \mathrm{d}\omega_i) = Z_{i,N}^\theta(\omega, \omega_i)\nu_i(\omega, \mathrm{d}\omega_i)$. Let us verify that it suffices to show, for all N, Lemma 1 with $(N_{i,N}^\theta, \mathcal{A}_{i,N}')$ in place of $(N_i^\theta, \mathcal{A}_i)$. Define the statistical experiments

$$\mathbb{F}_{i,N}(\omega) = (\mathcal{A}_i, \mathcal{A}_i, (N_{i,N}^\theta(\omega, \cdot))_{\theta\in\mathcal{V}(\omega)}),$$

$$\mathbb{E}_{i,N} = (A \times A_i, \mathcal{A} \otimes \mathcal{A}_i, (P^\theta(\mathrm{d}\omega)N_{i,N}^\theta(\omega, \mathrm{d}\omega_i))_{\theta\in\Theta})$$

and the random variables

$$R_{i,N}^\theta(\omega) = \int_{A_i} |Z_i^\theta(\omega, \omega_i) - Z_{i,N}^\theta(\omega, \omega_i)| \upsilon_i(\omega, \mathrm{d}\omega_i).$$

Then we have $\Delta(\mathbb{F}_i(\omega), \mathbb{F}_{i,N}(\omega)) \leqslant \frac{1}{2}\max_{\theta\in\Theta} R_{i,N}^\theta(\omega)$, which yields

$$\delta(\mathbb{F}_{1,N}(\omega), \mathbb{F}_{2,N}(\omega)) \leqslant \delta(\mathbb{F}_1(\omega), \mathbb{F}_2(\omega) + \tfrac{1}{2}\max_{\theta\in\Theta} R_{1,N}^\theta(\omega) + \tfrac{1}{2}\max_{\theta\in\Theta} R_{2,N}^\theta(\omega).$$

Likewise, we have

$$\delta(\mathbb{E}_1, \mathbb{E}_2) \leqslant \delta(\mathbb{E}_{1,N}, \mathbb{E}_{2,N}) + \tfrac{1}{2}\max_{\theta\in\Theta} \int R_{1,N}^\theta(\omega)P_\theta(\mathrm{d}\omega) + \tfrac{1}{2}\max_{\theta\in\Theta} \int R_{2,N}^\theta(\omega)P_\theta(\mathrm{d}\omega).$$

Using the fact that $R_{i,N}^\theta \to 0$ pointwise and in $L^1(P^\theta)$ for all θ, together with the fact that Θ is finite, we see that (49) will be deduced from the upper bounds

$$\delta(\mathbb{E}_{1,N}, \mathbb{E}_{2,N}) \leqslant \sup_{\omega\in B_N^\varepsilon} \delta(\mathbb{F}_{1,N}(\omega), \mathbb{F}_{2,N}(\omega)) + \sup_{\theta\in\Theta} P_\theta(\{\theta \notin \mathcal{V}\} \cup (A \backslash B_N^\varepsilon)),$$

where $B_N^\varepsilon = B \cap \{\max_{i=1,2}\max_{\theta\in\Theta} R_{i,N}^\theta \leqslant \varepsilon\}$ and $\varepsilon > 0$ is arbitrarily small. Furthermore, since $\mathcal{A} \otimes \mathcal{A}_{i,N}'$ is $\mathbb{E}_{i,N}$-sufficient and $\mathcal{A}_{i,N}'$ is $\mathbb{F}_{i,N}(\omega)$-sufficient for all $\omega \in A$, it is enough to prove (49) when $\mathcal{A}_i$ is generated by a finite partition.

In that case it is easily seen that for each $\omega \in A$ there exists a Markov kernel K_ω from $(A_1, \mathcal{A}_1)$ into $(A_2, \mathcal{A}_2)$ such that

$$\delta(\mathbb{F}_1(\omega), \mathbb{F}_2(\omega) = \tfrac{1}{2}\max_{\Theta} \| N_1^\theta(\omega, \cdot)K_\omega - N_2^\theta(\omega, \cdot)\|_{\mathrm{TV}}.$$

Moreover, the family $(K_\omega)_{\omega\in A}$ can be chosen so that

$$(\omega, \omega_1) \mapsto K_\omega(\omega_1, A_2)$$

is $\mathcal{A} \otimes \mathcal{A}_1$-measurable for all $A_2 \in \mathcal{A}_2$ (Strasser 1985, Theorem 6.10). Set $\tilde{B} = \{\theta \in \mathcal{V}\} \cap B$ and define the Markov kernel $\tilde{K}$ from $(A \times A_1, \mathcal{A} \otimes \mathcal{A}_1)$ into $(A \times A_2, \mathcal{A} \otimes \mathcal{A}_2)$ by

$$\tilde{K}(\omega, \omega_1, \mathrm{d}\omega', d\omega_2) = 1_{\tilde{B}}(\omega)\varepsilon_\omega(\mathrm{d}\omega') \otimes K_\omega(\omega_1, \mathrm{d}\omega_2) + 1_{A\backslash\tilde{B}}(\omega)Q_2(\mathrm{d}\omega', \mathrm{d}\omega_2),$$

where Q_2 is any probability measure on $A \times A_2$. Then we have

$$\|P_1^\theta\tilde{K} - P_2^\theta\|_{TV} \leqslant 2\int_{\tilde{B}} \delta(\mathbb{F}_1(\omega), \mathbb{F}_2(\omega))P^\theta(\mathrm{d}\omega) + P^\theta(A\backslash\tilde{B}),$$

which readily yields (49).

Remark 5. The result of Lemma 1 does not hold in general without condition (i): take $\Theta = \mathbb{R}$, $(\Omega, \mathcal{A}) = ([0, 1], \mathcal{B}([0, 1]0))$, $(\Omega_1, \mathcal{A}_1) = (\mathbb{R}^{[0,1]}, \mathcal{B}(\mathbb{R})^{\otimes[0,1]})$, $(\Omega_2, \mathcal{A}_2) = (\mathbb{R}, \mathcal{B}(\mathbb{R}))$, P^θ the Lebesgue measure, $N_1^\theta(\omega, \mathrm{d}\omega_1)$ the Dirac mass at point $t \mapsto \theta 1_{\{t=\omega\}}$, $N_2^\theta(\omega, \mathrm{d}\omega_2)$ the Dirac mass at θ, and $\mathcal{V}(\omega) = \Theta$. Then $\mathbb{F}_1(\omega)$, $\mathbb{F}_2(\omega)$ and $\mathbb{E}_2$ are equivalent to the perfect experiment but $\mathbb{E}_1$ is equivalent to the trivial experiment since $\omega_1(t) = 0$ P_1^θ-a.s., for all $t \in [0, 1]$.

References

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

Delattre, S. and Hoffmann, M. (2001) Mixed Gaussian white noise. To appear in *Math. Methods Statist.*

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

Ibragimov, I.A. and Has'minskii, R.Z. (1981) *Statistical Estimation: Asymptotic Theory.* Berlin: Springer-Verlag.

Klemela, J. and Nussbaum, M. (1998) Constructive asymptotic equivalence of density estimation and Gaussian white noise. Preprint, Weierstrass Institute.

Kutoyants, A.Y. (1984) *Parameter Estimation for Stochastic Processes*. Berlin: Heldermann Verlag.

Le Cam, L. (1986) *Asymptotic Methods in Statistical Decision Theory.* Berlin: Springer-Verlag.

Le Cam, L. and Yang, L.G. (2000) *Asymptotics in Statistics: Some Basic Concepts,* 2nd edition. New York: Springer-Verlag.

Liptser, R.S. and Shiryaev, A.N. (1977) *Statistics of Random Processes Vol. I.* Berlin: Springer-Verlag.

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

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

Revuz, D. and Yor, M. (1991) *Continuous Martingales and Brownian Motion*. Berlin: Springer-Verlag.

Strasser, H. (1985) *Mathematical Theory of Statistics*. Berlin: Walter de Gruyter.

Received March 2000 and revised August 2001