Stochastic Processes and their Applications 97 (2002) 147–170
stochastic processes and their applications
www.elsevier.com/locate/spa

Rate of convergence for parametric estimation in a stochastic volatility model

Marc Hoffmann
Laboratoire de Probabilités et Modèles Aléatoires, CNRS-UMR 7599, et Université Paris VII, UFR de Mathématiques, case 7012, 2 place Jussieu, 75251 Paris, France

Received 19 January 2000; received in revised form 4 April 2001; accepted 12 July 2001

Abstract

We consider the following hidden Markov chain problem: estimate the finite-dimensional parameter θ in the equation $v_t = v_0 + \int_0^t \sigma(\theta, v_s)\,\mathrm{d}W_s + \text{drift}$, when we observe discrete data $X_{i/n}$ at times $i = 0, \ldots, n$ from the diffusion $X_t = x_0 + \int_0^t v_s\,\mathrm{d}B_s + \text{drift}$. The processes $(W_t)_{t\in[0,1]}$ and $(B_t)_{t\in[0,1]}$ are two independent Brownian motions; asymptotics are taken as $n\to\infty$. This stochastic volatility model has been paid some attention lately, especially in financial mathematics.

We prove in this note that the unusual rate $n^{-1/4}$ is a lower bound for estimating θ. This rate is indeed optimal, since Gloter (CR Acad. Sci. Paris, t330, Série I, pp. 243–248), exhibited $n^{-1/4}$ consistent estimators. This result shows in particular the significant difference between "high frequency data" and the ergodic framework in stochastic volatility models (compare Genon-Catalot, Jeantheau and Laredo (Bernoulli 4 (1998) 283; Bernoulli 5 (2000) 855; Bernoulli 6 (2000) 1051) and also Sørensen (Prediction-based estimating functions. Technical report, Department of Theoretical Statistics, University of Copenhagen, 1998)).

MSC: 62G99; 62F99; 62M99

Keywords: Stochastic volatility models; Discrete sampling; High frequency data; Non-parametric Bayesian estimation

1. Introduction and main result

Stochastic volatility lies in the scope of hidden Markov chains. In the context of diffusion models, one basically attempts to recover information from n data extracted from a diffusion process, with diffusion coefficient driven itself by a stochastic differential equation which contains the unknown parameter of interest. The observation is no more a Markov chain and classical tools do not simply carry over to a non-Markovian set-up. The seminal work of Genon-Catalot et al. (1998, 2000a,b) led the path: their study started with the ergodic case, when the time discretization step Δ_n between two

E-mail address: hoffmann@math.jussieu.fr (M. Hoffmann).

0304-4149/02/$ - see front matter
PII: S0304-4149(01)00130-2

data tends to 0, but $n\Delta_n \to \infty$. They proceeded further to the case of a constant discretization step Δ_n where they showed in particular that the classical rate $n^{-1/2}$ holds. See also Sørensen (1998).

In this paper, we study an alternative asymptotic framework: the case of "small" Δ_n, i.e. when Δ_n is of order $1/n$. In this context, we can relax the ergodicity assumption which plays no role: the sample size increases not because of a longer observation period but, rather, because of more frequent observations. We prove that if $\Delta_n = 1/n$, the rate $n^{-1/4}$ is a lower bound for estimating a parameter $\theta \in \Theta \subset \mathbb{R}^k$ in the diffusion coefficient. Recently, Gloter (2000) exhibited $n^{-1/4}$ consistent estimator and this rate is therefore optimal.

Our approach is based on a Bayesian angle: we treat the hidden diffusion process as an unknown parameter living on the Wiener space $\mathscr{C}_0$ of continuous functions vanishing at the origin, equipped with the Wiener measure as prior probability. This enables us to substantially simplify the stochastic structure of the model, at the cost of enlarging the parameter space: the finite dimensional parameter space Θ is embedded in $\Theta \times \mathscr{C}_0$. We are thus led to a non-parametric problem. In the resulting simpler stochastic model, we can develop classical non-parametric tools and obtain the desired lower bound for the minimax risk. The essential difficulty consists in proving that the Wiener measure acts like a least favourable prior on an appropriate subset of $\mathscr{C}_0$.

1.1. Statistical setting

We observe $X^n = (X_{i/n},\ i = 0, \dots, n)$, where $(X_t)_{t\in[0,1]}$ is a 1-dimensional diffusion process of the form

$$X_t = x_0 + \int_0^t v_s \,\mathrm{d}B_s + \int_0^t b(s,\, v_s, X_s)\,\mathrm{d}s, \quad t \in [0,1], \tag{1}$$

where $x_0 \in \mathbb{R}$, $(B_t)_{t\in[0,1]}$ is a standard Brownian motion, the function b is smooth and the diffusion coefficient $(v_t)_{t\in[0,1]}$—the so-called stochastic volatility—solves the 1-dimensional equation

$$v_t = v_0 + \int_0^t \sigma(\theta,\, v_s)\,\mathrm{d}W_s + \int_0^t \rho(s,\, v_s)\,\mathrm{d}s, \quad t \in [0,1], \tag{2}$$

with $v_0 \in \mathbb{R}$, $(W_t)_{t\in[0,1]}$ a standard Brownian motion independent of $(B_t)_{t\in[0,1]}$ and ρ smooth. The function $\sigma(\theta, x)$ is known upto the parameter $\theta \in \Theta$, where $\Theta \subset \mathbb{R}^k$, $k \geqslant 1$ is given. Our aim is to estimate θ from the data X^n. Asymptotics are taken as $n \to \infty$. In this set-up, the drifts b and ρ cannot be identified from the data and are considered as nuisance parameters. Likewise, the initial condition v_0 is unknown. The assumptions on Eqs. (1) and (2) are the following.

Assumption A. A1. For all $\theta \in \Theta$, the map $x \to \sigma(\theta,x)$ is either, Lipschitz continuous and satisfies $\inf_x \sigma(\theta,x) > 0$, or has the multiplicative form $\sigma(\theta,x) = \sigma(\theta)x$.

A2. We have $\sup_{t,x} |b(t,x,y)| \leqslant C_1(1+|y|)$ and $\sup_t |\rho(t,x)| \leqslant C_2(1+|x|)$ for two constants $C_1, C_2 > 0$.

By Assumption A1 and A2, Eq. (2) admits a strong solution, with (almost surely on an appropriate probability space) locally bounded paths. This with Assumption A2 ensures the existence of a weak solution to (1) (on an appropriate probability space). Clearly, the assumptions on b and ρ are not minimal, but will be sufficient for the level of generalization intended here.

1.1.1. Main result

We assess the quality of an estimation procedure in squared-error loss, uniformly over the parameter set Θ. If $\hat{\theta}_n = \hat{\theta}_n(X^n)$ is an estimator of θ, we introduce its maximal quadratic risk

$$\mathbf{R}_n(\hat{\theta}_n, \varphi_n) = \sup_{b,\rho,v_0 \in V} \sup_{\theta \in \Theta} E_\theta^n\{\varphi_n^{-2}|\hat{\theta}_n - \theta|^2\},$$

where φ_n is a normalizing factor and $|\cdot|$ is the Euclidean norm on $\mathbb{R}^k$. The notation $E_\theta^n (= E^n_{\theta,b,\rho,v_0})$ denotes integration w.r.t. the law of the n-dimensional random vector $(X_{i/n},\ i = 1,\dots,n)$ on $\mathbb{R}^n$. The supremum in b and ρ is taken over all admissible drifts which satisfy Assumption A2. The supremum in v_0 is taken over a given subset $V \subseteq \mathbb{R}$ of possible initial conditions for the process v. Of course, the finiteness of $\mathbf{R}_n$ will be meaningful only if $\varphi_n \to 0$ as $n \to \infty$.

We further impose the following restrictions on V, Θ and σ, which assess in some sense the non-degeneracy of the model:

Assumption B. B1. $V \neq \{0\}$.

B2. The set Θ is bounded and contains an open set Θ_0.

B3. There exists $v_0 \in V$ such that $\{\sigma(\theta, v_0),\ \theta \in \Theta\} \subset \mathbb{R}$ contains an open set, and such that the map $\theta \to \sigma(\theta, v_0)$ is Lipschitz continuous on Θ_0.

The aim of this note is to prove the following:

Theorem 1. *Grant Assumptions A and B. The rate $\varphi_n = n^{-1/4}$ is a lower bound for the estimation of θ*:

$$\liminf_{n\to\infty} \inf_F \mathbf{R}_n(F, n^{-1/4}) > 0,$$

where the infimum is taken over all estimators.

Several remarks are in order: Theorem 1 is still valid if we take $V = \mathbb{R}$ and if we relax assumption A to arbitrary ρ and b—provided that Eqs. (1) and (2) are still meaningful—but it is dubious that one can find an estimator providing the finiteness of $\mathbf{R}_n$ for any $\varphi_n \to 0$ in this case. Second, the restriction to quadratic loss is inessential: Theorem 1 can be extended to any convex and bowl-shaped loss functions vanishing at the origin. Finally, this rate can be attained and is therefore optimal. Indeed, Gloter (2000), constructed convergent estimators for the normalization $n^{-1/4}$.

1.1.2. Translation into mathematical finance

The following type of models have been proposed, with $\Theta \subset \mathbb{R}$: one has data from the path of the diffusion

$$dX_t = \sqrt{V_t}\, dB_t + \text{drift}, \quad dV_t = c(\theta, V_t)\, dW_t + \text{drift}.$$

Hull and White (1988), and later Heston (1993), considered the function $c(\theta, V_t) = \theta\sqrt{V_t}$. Chesney and Scott (1989), Wiggins (1987), Melino and Turnbull (1990), proposed the model $c(\theta, V_t) = \sqrt{\theta} V_t$. In our setting, this corresponds to the change of variable $v_t = \sqrt{V_t}$, and under regularity assumptions, by Itô's formula, we have the following correspondence:

$$\sigma(\theta, x) = \frac{1}{2x} c(\theta, x^2).$$

In particular, the model of Hull, White and Heston turns to $\sigma(\theta, x) = \theta/2$ in our case and the model of Chesney, Scott, Wiggins, Melino and Turnbull corresponds to $\sigma(\theta, x) = 1/2\sqrt{\theta} x$. Note that these models are fully compatible with assumptions A and B.

1.2. Sketch of proof of Theorem 1

We first describe our method of proof in an informal way, emphasizing the analogy to a non-parametric Bayesian problem. The thorough proof of Theorem 1 is given in Sections 2 and 3.

Step 1: *reduction to the simplest case.* We first restrict our attention to the special case $\Theta \subset (0, +\infty)$, $\sigma(\theta, x) = \sqrt{\theta}$. Assume further that $\rho = 0$, $b = 0$, and set $\tilde{W}_t = W_t + v_0/\sqrt{\theta}$. Thus

$$v_t = \sqrt{\theta} \tilde{W}_t \quad \text{and} \quad X_t = x_0 + \sqrt{\theta} \int_0^t \tilde{W}_s\, dB_s.$$

By Itô's formula, $d\langle X \rangle_t = \theta \tilde{W}_t^2\, dt$, $dv_t^2 = 2v_t\, dv_t$+a bounded variation part, and $d\langle v^2 \rangle_t = 4\theta^2 \tilde{W}_t^2\, dt$. We thus obtain the following representation:

$$\theta = \frac{1}{4} \frac{\langle v^2 \rangle_1}{\langle X \rangle_1}. \tag{3}$$

It is well known that the quadratic variation $\langle X \rangle_1$ can be approximated from X^n with the rate $n^{-1/2}$ (e.g. Revuz and Yor, 1990), which is superoptimal for our purpose, i.e. if we *believe* that the correct rate is $n^{-1/4}$. The complexity of our problem must then be linked to the estimation of $\langle v^2 \rangle_1$. This shall explain heuristically the inflation of the risk bound. More precisely, we have the representation

$$\langle v^2 \rangle_1 =: \phi(\theta, W) = \langle (v_0 + \sqrt{\theta} W)^2 \rangle_1,$$

with W being an auxiliary "unknown" infinite dimensional parameter. This suggests to enlarge the parameter space by considering simultaneously (θ, W) instead of θ solely. To exploit this point of view, we must use the prior information we have on W, namely that W is a typical Brownian path.

Step 2: *equivalence to a non-parametric Bayesian estimation problem.* We endow Θ with its Borel σ-field $\mathscr{B}(\Theta)$. Let us recall a well-known fact from the minimax theory:

$$\inf_F \sup_{\theta\in\Theta} E_\theta^n\{|F-\theta|^2\} = \inf_F \sup_\mu \int_\Theta \mu(\mathrm{d}\theta) E_\theta^n\{|F-\theta|^2\},$$

where the supremum is taken over all probability measures μ on Θ. By the minimax theorem (possibly under restrictions on Θ and the set of estimators considered), we can interchange the inf and sup and the last quantity is equal to

$$\sup_\mu \left(\inf_F \int_\Theta \mu(\mathrm{d}\theta) E_\theta^n\{|F-\theta|^2\} \right). \tag{4}$$

Thus, evaluating the minimax risk is equivalent to finding the least favourable prior μ on Θ. Now, let $\mathscr{C}_0$ denote the Wiener space of continuous functions on $[0,1]$ vanishing at the origin, equipped with the norm of uniform convergence and its Borel σ-field $\mathscr{F}$. Let $P_{\theta,\omega}^n$ be the law of $(X_{i/n},\ i=1,\dots,n)$, conditional on $W=\omega$:

$$P_\theta^n(\mathrm{d}x_1 \dots \mathrm{d}x_n) = \int_{\mathscr{C}_0} \mathbf{W}(\mathrm{d}\omega) P_{\theta,\omega}(\mathrm{d}x_1 \dots \mathrm{d}x_n),$$

where $\mathbf{W}$ denotes the Wiener measure on $(\mathscr{C}_0, \mathscr{F})$. The minimax risk (4) is then greater than

$$\inf_F \int_{\Theta\times\mathscr{C}_0} \mu(\mathrm{d}\theta)\mathbf{W}(\mathrm{d}\omega) E_{\theta,\omega}^n\{|F-\theta|^2\} \tag{5}$$

for any prior μ on Θ. The worse μ, the closer we get to the optimal minimax risk. So in our new enlarged parameter space, we must understand how a clever choice of μ on Θ makes $\mu\otimes\mathbf{W}$ a least favourable prior on $\Theta\times\mathscr{C}_0$.

Though it certainly does not look so, the second problem is actually easier than the original one: computations are straightforward with $P_{\theta,\omega}^n$ which is the law of a Gaussian process with independent increments. It is, however, not clear how to work explicitly with P_θ^n.

Step 3: *the general case by renormalization of experiments*. We use a renormalization argument, based on the scaling properties of Brownian motion. For $0<\delta<1$, denote by $\mathscr{G}_\delta^n(\sigma)$ the experiment generated by the data $(X_{i/n},\ i=0,\dots,\lfloor n\delta\rfloor)$, where X solves (1) with diffusion coefficient σ for the component v; $\lfloor x\rfloor$ denotes the integer part of x. The proof then goes as follows:

1. For ε_n converging to 0 at a certain rate, the experiments $\mathscr{G}_1^{\lfloor n\varepsilon_n\rfloor}(\sigma)$ and $\mathscr{G}_{\varepsilon_n}^n(\varepsilon_n^{-1/2}\sigma)$ are equivalent.
2. The rate $(n\varepsilon_n)^{-1/4}$ is a lower bound for $\mathscr{G}_{\varepsilon_n}^n(\varepsilon_n^{-1/2}\sigma(\theta,\, v_0))$ by a reparametrization argument and Steps 1 and 2.
3. $\mathscr{G}_{\varepsilon_n}^n(\varepsilon_n^{-1/2}\sigma)$ and $\mathscr{G}_{\varepsilon_n}^n(\varepsilon_n^{-1/2}\sigma(\theta, v_0))$ are close as $n\to\infty$.
4. If the rate $n^{-1/4}$ could be strictly improved in $\mathscr{G}_1^n(\sigma)$, the rate $(n\varepsilon_n)^{-1/4}$ could also be improved in $\mathscr{G}_{\varepsilon_n}^n(\varepsilon_n^{-1/2}\sigma)$ which contradicts 2 by 3.

2. The fundamental case $\sigma(\theta,x)=\theta^{1/2}$

In this section, we prove the following:

Proposition 1. *Assume* $\Theta=(\theta_-,\theta_+)\subset(0,+\infty)$, $\sigma(x,\theta)=\sqrt{\theta}$, $b=0$, $\rho=0$ *and* $V=\{v_0\}$, *with* $v_0>0$. *In this setting, we have Theorem* 1.

We evaluate a lower bound for the Bayes risk

$$\mathbf{B}_n(v_n\otimes\mathbf{W},n^{-1/4})=\inf_F\int_\Theta v_n(\mathrm{d}\theta)E_\theta^n\{n^{1/2}(F-\theta)^2\}$$

$$=\inf_F\int_{\Theta\times\mathscr{C}_0}v_n(\mathrm{d}\theta)\mathbf{W}(\mathrm{d}\omega)E_{\theta,\omega}^n\{n^{1/2}(F-\theta)^2\},$$

where the infimum is taken over all estimators for an appropriate sequence of probability measures v_n on Θ. The conclusion follows from

$$\inf_F\mathbf{R}_n(F,n^{-1/4})\geqslant\mathbf{B}_n(v_n\otimes\mathbf{W},n^{-1/4}).$$

In the following, the expression "for large enough n" is implicitly understood as "for n bigger than a constant possibly depending on all the characteristics of the problem", i.e. Θ, V, and all other intermediate ancillary quantities. As soon as a quantity is stated as "being fixed", further dependence on that quantity may be omitted in the notation. Abusing notation slightly, we will write $a_n\leqslant b_n(1+\mathrm{o}(1))$ if a_n and b_n are two sequences (possibly depending on the characteristics of the problem) that satisfy $\limsup_{n\to\infty}a_n/b_n\leqslant 1$.

2.1. Preliminaries: a two-points inequality

Let us consider an apparently different problem: we first look for a bound in the ideal Bayes problem

$$\mathbf{B}_n(v_n\otimes\Lambda_n,n^{-1/4})=\inf_F\int_{\Theta\times\mathscr{C}_0}v_n(\mathrm{d}\theta)\Lambda_n(\mathrm{d}\omega)E_{\theta,\omega}^n\{n^{1/2}(F-\theta)^2\},$$

where now Λ_n is an *arbitrary* sequence of probability measures on $(\mathscr{C}_0,\mathscr{F})$. We give the following "mild" construction of $v_n\otimes\Lambda_n$ which shows that the rate $n^{-1/4}$ cannot be improved in the ideal Bayes problem. Of course, the construction is artificial, in the sense that we could find a "hard" pairing $v_n\otimes\Lambda_n$ so that there exists *no convergent estimator* in this ideal context. However, proving that the rate $n^{-1/4}$ is a lower bound in the ideal Bayes problem is instructive to understand for which choice of "mild" $v_n(\mathrm{d}\theta)$, the prior $\mathbf{W}(\mathrm{d}\omega)$ in the true problem is not essentially more favourable than the "mild" $\Lambda_n(\mathrm{d}\omega)$ in the ideal problem.

Remember that v_0 is fixed throughout this section. For $\xi>0$, define

$$t_n(\theta)=(1-\xi n^{-1/4})\theta,$$

$$T_n(\omega)=\sqrt{1+\xi n^{-1/4}}\,\omega.$$

Let (θ_0, ω_0) be an element of $\Theta \times \mathscr{C}_0$ such that for large enough n, $t_n(\theta_0) \in \Theta$ and:

$$\eta_- \leqslant v_0/\sqrt{\theta_0} + \omega_0(t) \leqslant \eta+, \tag{6}$$

for some constants $0 < \eta_- < \eta_+$, a choice which is obviously possible. The numbers η_- and η_+ are fixed from now on. Pick the following priors:

$$v_n(\mathrm{d}\theta) = \tfrac{1}{2}(\delta_{\theta_0}(\mathrm{d}\theta) + \delta_{t_n(\theta_0)}(\mathrm{d}\theta))$$

and

$$\Lambda_n(\mathrm{d}\omega) = \tfrac{1}{2}(\delta_{\omega_0}(\mathrm{d}\omega) + \delta_{T_n(\omega_0)}(\mathrm{d}\omega)).$$

Proposition 2. *There exists an explicit choice* [1] *of* ξ *such that*

$$\liminf_{n \to \infty} \mathbf{B}_n(v_n \otimes \Lambda_n, n^{-1/4}) > 0.$$

Proof. Let F be an arbitrary estimator. The Bayes risk of F for the prior $v_n \otimes \mathbf{W}$ and with normalization $n^{-1/4}$ is greater than

$$\tfrac{1}{4}(E^n_{\theta_0,\omega_0}\{n^{1/2}(\theta_0 - F)^2\} + E^n_{t_n(\theta_0),T_n(\omega_0)}\{n^{1/2}(t_n(\theta_0) - F)^2\}).$$

For $\lambda > 0$, by the triangle inequality, this last quantity is greater than

$$\frac{\mathrm{e}^{-\lambda}}{8} n^{1/2}[\theta_0 - t_n(\theta_0)]^2 P^n_{\theta_0,\omega_0}\left\{\frac{\mathrm{d}P^n_{t_n(\theta_0),T_n(\omega_0)}}{\mathrm{d}P^n_{\theta_0,\omega_0}} \geqslant \mathrm{e}^{-\lambda}\right\}. \tag{7}$$

Next, we need a control of the above likelihood ratio. We first introduce some notation: for a continuous function φ on $[0,1]$, let Q^n_φ denote the law of $Y^n = (Y_{i/n},\ i = 1,\dots,n)$, where

$$Y_t = x_0 + \int_0^t \varphi(s)\,\mathrm{d}B_s. \tag{8}$$

(The initial condition x_0 is fixed througout this section.) We set as usual $\|f\|_\infty = \sup_{t\in[0,1]} |f(t)|$ and $\|f\|_p = (\int_0^1 |f(s)|^p\,\mathrm{d}s)^{1/p}$ for $1 \leqslant p < \infty$. We plan to apply the following result, which controls the separation rate between $P^n_{\theta_0,\omega_0}$ and $P^n_{t_n(\theta_0),T_n(\omega_0)}$ as $n \to \infty$. The proof is delayed until appendix.

Lemma 1. *Let* φ_0, φ_1 *be two continuous functions on* $[0,1]$ *such that*:
(i) $\inf_t \varphi_0(t) \geqslant \mu > 0$,
(ii) $\|\varphi_1^2 - \varphi_0^2\|_\infty \leqslant \frac{1}{5}\mu^2$,
(iii) *For some* $L > 0$:

$$\|\varphi_1^2 - \varphi_0^2\|_2 \leqslant \frac{\mu^2 L}{\sqrt{n}}.$$

[1] Namely $\xi = \frac{1}{2}\eta_-/\eta_+$ but other choices are obviously possible, see the proof below.

Let $\lambda > 0$. *For* $n \geqslant 1$, *we have*

$$Q^n_{\varphi_1}\left\{\frac{\mathrm{d}Q^n_{\varphi_0}}{\mathrm{d}Q^n_{\varphi_1}} \geqslant \mathrm{e}^{-\lambda}\right\} \geqslant 1 - \frac{L}{\lambda}(L/3+\sqrt{3}/2).$$

Note that $P^n_{\theta,\omega} = Q^n_{v_0+\sqrt{\theta}\omega}$. By definition of t_n and T_n, we have:

$$\begin{aligned}
&|[v_0+\sqrt{\theta_0}\omega_0(t)]^2 - [v_0+\sqrt{t_n(\theta_0)}T_n(\omega_0)(t)]^2| \\
&\quad\leqslant [\sqrt{\theta_0}\omega_0(t) - \sqrt{t_n(\theta_0)}T_n(\omega_0)(t)]2[v_0+\sqrt{\theta_0}\omega(t)](1+\mathrm{o}(1)) \\
&\quad\leqslant [1-(1-\xi n^{-1/4})^{1/2}(1+\xi n^{-1/4})^{1/2}]2\sqrt{\theta_0}|\omega_0(t)|\,|\sqrt{\theta_0}\omega_0(t)+v_0|(1+\mathrm{o}(1)) \\
&\quad\leqslant \xi^2\theta_0\eta_+^2 n^{-1/2}(1+\mathrm{o}(1)).
\end{aligned}$$

We also have that $v_0 + t_n(\theta_0)T_n(\omega_0)(t)$ is greater than

$$v_0 + (1-\xi n^{-1/4})^{1/2}(1+\xi n^{-1/4})\sqrt{\theta_0}\omega_0(t) \geqslant \sqrt{\theta_0}\eta_-(1+\mathrm{o}(1)). \tag{9}$$

We then apply Lemma 1 for large enough n, with, for instance $\mu = \sqrt{\theta_0}\eta_-/2$ and $L = 2\xi^2\theta_0\eta_+^2/(\theta_0\eta_-^2/2) = 4\xi^2\eta_+^2/\eta_-^2 = 1$ for the choice $\xi = (1/2)\eta_-/\eta_+$. We derive

$$P^n_{\theta_0,\omega_0}\left\{\frac{\mathrm{d}P^n_{t_n(\theta_0),T_n(\omega_0)}}{\mathrm{d}P^n_{\theta_0,\omega_0}} \geqslant \mathrm{e}^{-\lambda}\right\} \geqslant 1 - \frac{2+3\sqrt{3}}{6\lambda}. \tag{10}$$

Back to (7), we have

$$n^{1/2}[\theta_0 - t_n(\theta_0)]^2 \geqslant \xi^2\theta_0$$

which proves the proposition in view of (10) and for large enough λ. □

2.2. *From* Λ_n *to* **W**: *a key inequality*

We start with an abstract result which is the key ingredient to understand how the Wiener measure **W** acts like a least-favourable prior on $\mathscr{C}_0$, i.e. like Λ_n in the ideal context of Section 2.1.

Let $(E,\mathscr{E},\mu)$ be a probability space, let G and H be two positive measurable mappings on E, let T be a measurable map on E. We denote by $T\mu$ the image measure of μ by T on $(E,\mathscr{E})$, defined by $T\mu(\mathscr{A}) = \mu[T^{-1}(\mathscr{A})]$ for $\mathscr{A}\in\mathscr{E}$. We also denote by $D(P,Q) = \int |\log \mathrm{d}P/\mathrm{d}Q|\,\mathrm{d}P$ ($\leqslant +\infty$) the following Kullback–Leiber type pseudo-distance between the probability measures P and Q.

Lemma 2. *Let* $\mathscr{A} = T^{-1}(\mathscr{B})$ *for some measurable* $\mathscr{B}\in\mathscr{E}$. *Assume that*

(i) $G(x) + H(T(x)) \geqslant \alpha$ *for* $x\in\mathscr{A}$ *and for some* $\alpha > 0$,

(ii) $T\mu\sim\mu$ *and* $D(T\mu,\mu)\leqslant\beta$, *for some* $\beta < +\infty$. ($\sim$ *denotes equivalence between measures.*) *Then*

$$\int_{\mathscr{A}} G(x)\mu(\mathrm{d}x) + \int_{\mathscr{B}} H(x)\mu(\mathrm{d}x) \geqslant \alpha \sup_{\lambda>0} \mathrm{e}^{-\lambda}\left(\mu(\mathscr{A}) - \frac{1}{\lambda}\beta\right).$$

Proof. Let $\lambda > 0$. We have

$$\int_{\mathscr{A}} G(x)\mu(\mathrm{d}x) + \int_{\mathscr{B}} H(x)\mu(\mathrm{d}x)$$

$$= \int_{T^{-1}(\mathscr{B})} \left\{ G(x) + H(T(x)) \frac{\mathrm{d}\mu}{\mathrm{d}T\mu}(T(x)) \right\} \mu(\mathrm{d}x)$$

$$\geqslant \mathrm{e}^{-\lambda} \int_{\mathscr{A}} \mu(\mathrm{d}x)[G(x) + H(T(x))] 1_{\frac{\mathrm{d}\mu}{\mathrm{d}T\mu}(T(x)) \geqslant \mathrm{e}^{-\lambda}}$$

$$\geqslant \alpha \mathrm{e}^{-\lambda} \int_{\mathscr{A}} \mu(\mathrm{d}x) 1_{\frac{\mathrm{d}\mu}{\mathrm{d}T\mu}(T(\cdot)) \geqslant \mathrm{e}^{-\lambda}} \quad \text{by (i)}$$

$$= \alpha \mathrm{e}^{-\lambda} T\mu(\mathscr{B}) \int_{\mathscr{B}} 1_{\frac{\mathrm{d}\mu}{\mathrm{d}T\mu}(x) \geqslant \mathrm{e}^{-\lambda}} \frac{T\mu(\mathrm{d}x)}{T\mu(\mathscr{B})}$$

$$\geqslant \alpha \mathrm{e}^{-\lambda} T\mu(\mathscr{B}) \left\{ 1 - \frac{1}{\lambda} \int_{\mathscr{B}} \left| \log \frac{\mathrm{d}\mu}{\mathrm{d}T\mu}(x) \right| \frac{T\mu(\mathrm{d}x)}{T\mu(\mathscr{B})} \right\} \quad \text{by Chebyshev's inequality}$$

$$\geqslant \alpha \mathrm{e}^{-\lambda} \left\{ T\mu(\mathscr{B}) - \frac{1}{\lambda} \int_{E} \left| \log \frac{\mathrm{d}T\mu}{\mathrm{d}\mu}(x) \right| T\mu(\mathrm{d}x) \right\}$$

$$\geqslant \alpha \mathrm{e}^{-\lambda} \left\{ \mu(\mathscr{A}) - \frac{1}{\lambda} \beta \right\} \quad \text{by (ii) and since } T\mu(\mathscr{B}) = \mu(\mathscr{A}). \qquad \square$$

Let us interpret Lemma 2 in our context: we know from Proposition 2 that if we pick a pair $(\theta_0, \omega_0) \in \Theta \times \mathscr{C}_0$ such that $(t_n[\theta_0], T_n[\omega_0])$ is well defined and satisfies the assumptions of Section 2.1, then we have for large enough n

$$E_{\theta_0,\omega_0}\{(F - \theta_0)^2\} + E_{t_n(\theta_0),T_n(\omega_0)}\{[F - t_n(\theta_0)]^2\} \geqslant \text{constant}(\theta_0) n^{-1/2}, \tag{11}$$

or, in a statistical-like language, by testing

$$P^n_{\theta_0,\omega_0} \text{ versus } P^n_{t_n(\theta_0),T_n(\omega_0)},$$

we achieve the desired bound. In a Bayesian framework, such a test is obtained by equipping $\Theta \times \mathscr{C}_0$ with the sequence of priors $v_n \otimes \Lambda_n$ of Section 2.1. Although we can put an arbitrary prior on Θ, we do not have this freedom on $\mathscr{C}_0$: the structure of our model imposes the Wiener measure $\mathbf{W}$ as prior probability. Thus, by picking v_n as before and setting

$$G(\omega) = \tfrac{1}{2} E^n_{\theta_0,\omega}\{(F - \theta_0)^2\},$$

$$H(\omega) = \tfrac{1}{2} E^n_{t_n(\theta_0),\omega}\{(F - \theta_0)^2\},$$

for an arbitrary estimator F, we wish to bound from below the quantity

$$\int_{\mathscr{C}_0} [G(\omega) + H(\omega)] \mathbf{W}(\mathrm{d}\omega),$$

knowing that thanks to the specific choice of v_n, we have by (11)

$$G(\omega) + H(T_n[\omega]) \geqslant \text{constant}(\theta_0) n^{-1/2}$$

for ω in an appropriate set. A lower bound of this kind is precisely the target of Lemma 2. However, Lemma 2 cannot be applied directly:

The map $\omega \mapsto T_n(\omega) = constant_n\ \omega$ is such that $\mathbf{W}$ *and* $T_n\mathbf{W}$ *are mutually orthogonal, and therefore the crucial assumption* (ii) *of Lemma* 2 *is not satisfied.*

To circumvent this objection, we need to modify T_n. We approximate

$$T_n(\omega) = \sqrt{1 + \xi n^{-1/4}}\, \omega$$

by a "smooth" transformation $\tilde{T}_n(\omega)$ such that $\tilde{T}_n\mathbf{W} \sim \mathbf{W}$ while preserving the properties of T_n that were used in Proposition 2.

More precisely, let ϕ be an infinitely many times differentiable function, with support in $[0, \delta]$ for some $\delta > 0$ and such that $\int \phi(x)\,\mathrm{d}x = 1$. The function ϕ is fixed throughout the paper, and dependence on ϕ in the constants may be omitted from now on. Define

$$\tilde{T}_n(\omega) = \omega + \phi_n(\omega),$$

where

$$\phi_n[\omega](t) := \frac{\xi}{2} \int_0^1 n^{1/4} \phi(n^{1/2}(t - s))\omega_s\,\mathrm{d}s.$$

Also

$$\frac{\mathrm{d}}{\mathrm{d}t}\phi_n[\omega](t) = \frac{\xi}{2} \int_0^1 n^{3/4} \phi'(n^{1/2}(t - s))\omega_s\,\mathrm{d}s.$$

Under $\mathbf{W}$, the process $t \to \mathrm{d}/\mathrm{d}t\phi_n[\omega](t)$ is adapted w.r.t. the canonical filtration, Gaussian and continuous (in the mean square). By Girsanov's theorem, $\tilde{T}_n\mathbf{W} \sim \mathbf{W}$. The properties of the transformation $\tilde{T}_n$ are summarized in the next lemma.

We first introduce some notation. For a function $f \in \mathscr{C}_0$ and $s > 0$, let

$$|f|_{s,2,\infty} = \sup_{t \in [0,1]} t^{-s} \sup_{h \leqslant t} \left(\int_{I_h} (f(u + h) - f(u))^2\,\mathrm{d}u \right)^{1/2} \leqslant +\infty,$$

with $I_h = \{u \in [0, 1]\colon u - h \in [0, 1]\}$, denote the Besov semi-norm of f on $[0, 1]$. The Besov space $B^s_{2,\infty}$ on the interval $[0, 1]$ is defined as

$$B^s_{2,\infty} = \{f : [0, 1] \to \mathbb{R}\colon \|f\|_{s,2,\infty} := \|f\|_2 + |f|_{s,2,\infty} < +\infty\}.$$

It is noteworthy (see Roynette, 1993) and will be exploited later that $\omega \in B^{1/2}_{2,\infty}$, $\mathbf{W}(\mathrm{d}\omega)$ almost surely.

Lemma 3. (i) *For $\omega_0 \in \mathscr{C}_0$, let*

$$\alpha(\omega_0, \xi, \theta_0) = \sqrt{\theta_0}\xi \left(\frac{5\xi}{8} \|\omega_0\|_\infty + c(\phi)|\omega_0|_{1/2,2,\infty} \right) 2(\sqrt{\theta_0}\|\omega_0\|_\infty + v_0),$$

where $c(\phi)=\int_{\mathbb{R}}(1+u)|\phi(u)|\,\mathrm{d}u$. *We have*

$$\|[\sqrt{\theta_0}\omega_0+v_0]^2-[\sqrt{t_n(\theta_0)}\tilde{T}_n(\omega_0)+v_0]^2\|_2\leqslant \alpha(\omega_0,\xi,\theta_0)n^{-1/2}(1+\mathrm{o}(1)).$$

(ii) *There exists an explicit*[2] $\beta(\xi)>0$ *such that for large enough* n

$$D(\tilde{T}_n\mathbf{W},\mathbf{W})\leqslant\beta(\xi).$$

Proof. Note first that

$$\|\tilde{T}_n(\omega_0)\|_\infty\leqslant\|\omega_0\|_\infty+\frac{\xi}{2}n^{-1/4}\|\omega_0\|_\infty\|\phi\|_1\leqslant\|\omega_0\|_\infty(1+\mathrm{o}(1)).$$

Thus

$$\begin{aligned}&\|(\sqrt{\theta_0}\omega_0+v_0)^2-(\sqrt{t_n(\theta_0)}\tilde{T}_n(\omega_0)+v_0)^2\|_2\\&\quad\leqslant\|\sqrt{\theta_0}\omega_0-\sqrt{t_n(\theta_0)}\tilde{T}_n(\omega_0)\|_2\,2(\sqrt{\theta_0}\|\omega_0\|_\infty+v_0)(1+\mathrm{o}(1)).\end{aligned}$$

We plan to use the following decomposition:

$$\sqrt{\theta_0}\omega_0-\sqrt{t_n(\theta_0)}\tilde{T}_n(\omega_0)=A_1^{(n)}+A_2^{(n)}+A_3^{(n)},\tag{12}$$

with

$$\begin{aligned}A_1^{(n)}&=\sqrt{\theta_0}\omega_0-\sqrt{t_n(\theta_0)}T_n(\omega_0),\\A_2^{(n)}&=\sqrt{t_n(\theta_0)}T_n(\omega_0)-\sqrt{t_n(\theta_0)}\left(1+\frac{\xi}{2}n^{-1/4}\right)\omega_0,\\A_3^{(n)}&=\sqrt{t_n(\theta_0)}\left(1+\frac{\xi}{2}n^{-1/4}\right)\omega_0-\sqrt{t_n(\theta_0)}\tilde{T}_n(\omega_0),\end{aligned}$$

where T_n is defined in Section 2.1. We already know from Section 2.1 that

$$\|A_1^{(n)}\|_2\leqslant\sqrt{\theta_0}\|\omega_0\|_\infty\frac{\xi^2}{2}n^{-1/2}(1+\mathrm{o}(1)).$$

Next

$$\begin{aligned}\|A_2^{(n)}\|_2&=\sqrt{t_n(\theta_0)}\|\omega_0\|_2\left|\sqrt{1+\xi n^{-1/4}}-\left(1+\frac{\xi}{2}n^{-1/4}\right)\right|\\&\leqslant\sqrt{\theta_0}\|\omega_0\|_\infty\frac{\xi^2}{8}n^{-1/2}(1+\mathrm{o}(1)).\end{aligned}$$

In order to bound $A_3^{(n)}$, we will use the following classical inverse estimate (see e.g. Kerkyacharian and Picard, 1997):

$$\left\|\omega_0(\cdot)-\int_0^1 n^{1/2}\phi(n^{1/2}(\cdot-s))\omega_0(s)\,\mathrm{d}s\right\|_2\leqslant c(\phi)|\omega_0|_{1/2,2,\infty}n^{-1/4}.\tag{13}$$

[2] Specified in (14) below.

Thus

$$\|A_3^{(n)}\|_2 = \sqrt{t_n(\theta_0)}\frac{\xi}{2}n^{-1/4}\left\|\omega_0(\cdot) - \int_0^1 n^{1/2}\phi(n^{1/2}(\cdot - s))\omega_0(s)\,\mathrm{d}s\right\|_2$$

$$\leqslant \sqrt{\theta_0}\frac{\xi}{2}c(\phi)n^{-1/2}(1+\mathrm{o}(1)).$$

Summing the error estimates of $A_i^{(n)}$, $i=1,2,3$, we obtain (i).

We next turn to (ii). We denote by $y_t(\omega)=\omega_t$ the canonical process on $(\mathscr{C}_0,\mathscr{F})$. By Girsanov theorem

$$D(\tilde{T}_n\mathbf{W},\mathbf{W}) = E_{\tilde{T}_n\mathbf{W}}\left\{\left|\int_0^1 \frac{\mathrm{d}}{\mathrm{d}t}\phi_n[y](t)\,\mathrm{d}y_t - \frac{1}{2}\int_0^1 \frac{\mathrm{d}}{\mathrm{d}t}\phi_n[y](t)^2\,\mathrm{d}t\right|\right\}$$

$$\leqslant E_{\tilde{T}_n\mathbf{W}}\left\{\left|\int_0^1 \frac{\mathrm{d}}{\mathrm{d}t}\phi_n[y](t)\,\mathrm{d}\tilde{W}_t\right|\right\} + \frac{1}{2}\int_0^1 E_{\tilde{T}_n\mathbf{W}}\left\{\frac{\mathrm{d}}{\mathrm{d}t}\phi_n[y](t)^2\right\}\mathrm{d}t,$$

where $\tilde{W}_t = y_t - \phi_n[y](t)$ is a $(\tilde{T}_n\mathbf{W},\mathscr{F})$ Brownian motion. By the Burckholder–Davis–Gundy inequality, the last quantity is less than

$$h\left(\int_0^1 E_{\tilde{T}_n\mathbf{W}}\left\{\frac{\mathrm{d}}{\mathrm{d}t}\phi_n[y](t)^2\right\}\mathrm{d}t\right)$$

with $h(x)=c_\star\sqrt{x}+\frac{1}{2}x$. ($c_\star$ is the constant in the BDG inequality.) We will use the following decomposition: $(\mathrm{d}/\mathrm{d}t)\phi_n[y](t)=B_1^{(n)}+B_2^{(n)}$, with

$$B_1^{(n)} = \frac{\xi}{2}n^{3/4}\int_0^1 \phi'(n^{1/2}(t-s))\tilde{W}_s\,\mathrm{d}s,$$

$$B_2^{(n)} = \frac{\xi}{2}n^{3/4}\int_0^1 \phi'(n^{1/2}(t-s))\phi_n[y](s)\,\mathrm{d}s.$$

Integrating by part and using that $\phi(n^{1/2}(t-1))=0$ for all $t\in[0,1]$ since ϕ is compactly supported in $[0,\delta]$, we have

$$B_1^{(n)} = \frac{\xi}{2}n^{1/4}\int_0^1 \phi(n^{1/2}(t-s))\,\mathrm{d}\tilde{W}_s.$$

It follows that

$$E_{\tilde{T}_n\mathbf{W}}\{(B_1^{(n)})^2\} \leqslant \frac{\xi^2}{4}E_{\tilde{T}_n\mathbf{W}}\left\{\left(n^{1/4}\int_0^1 \phi(n^{1/2}(t-s))\,\mathrm{d}\tilde{W}_s\right)^2\right\}$$

$$= \frac{\xi^2}{4}n^{1/2}\int_0^1 \phi(n^{1/2}(t-s))^2\,\mathrm{d}s = \frac{\xi^2}{4}\|\phi\|_2^2$$

by Itô's isometry. Next, by definition of $\phi_n[y]$,

$$|B_2^{(n)}| \leqslant \sup_{t\in[0,1]}|y_t|\frac{\xi}{2}n\int_0^1\int_0^1 |\phi'(n^{1/2}(t-s))\phi(n^{1/2}(s-u))|\,\mathrm{d}u\,\mathrm{d}s$$

$$= \sup_{t\in[0,1]}|y_t|\frac{\xi}{2}\|\phi\|_1\|\phi'\|_1.$$

Since $|y_t| \leqslant \sup_{t\in[0,1]}|\tilde{W}_t| + \xi/2\int_0^1 n^{1/4}|\phi(n^{1/2}(t-s))|\,|y_s|\,\mathrm{d}s$, Gronwall's lemma implies

$$|y_t| \leqslant \sup_{t\in[0,1]}|\tilde{W}_t|\exp\left\{\frac{\xi}{2}\int_0^1 n^{1/4}|\phi(n^{1/2}(t-s))|\,\mathrm{d}s\right\}$$

$$= \sup_{t\in[0,1]}|\tilde{W}_t|\exp\left(\frac{\xi}{2}n^{-1/4}\|\phi\|_1\right)$$

which is less than $\sup_{t\in[0,1]}|\tilde{W}_t|(1+\mathrm{o}(1))$. Thus

$$E_{\tilde{T}_n\mathbf{W}}\{(B_2^{(n)})^2\} \leqslant \frac{\xi^2}{2}E_{\tilde{T}_n\mathbf{W}}\left\{\sup_{t\in[0,1]}\tilde{W}_t^2\right\}\|\phi\|_1^2\|\phi'\|_1^2(1+\mathrm{o}(1))$$

$$\leqslant 2\xi^2\|\phi\|_1^2\|\phi'\|_1^2(1+\mathrm{o}(1))$$

by Doob's inequality. Putting together the two estimates for $E_{\tilde{T}_n\mathbf{W}}\{(B_i^{(n)})^2\}$, $i=1,2$, we obtain

$$E_{\tilde{T}_n\mathbf{W}}\left\{\left[\frac{\mathrm{d}}{\mathrm{d}t}\phi_n[y](t)\right]^2\right\} \leqslant \xi^2\left(\frac{\|\phi\|_2^2}{2}+4\|\phi\|_1^2\|\phi'\|_1^2\right)(1+\mathrm{o}(1)).$$

Finally, for large enough n

$$D(\tilde{T}_n\mathbf{W},\mathbf{W}) \leqslant h(\xi^2(\|\phi\|_2^2+8\|\phi\|_1^2\|\phi'\|_1^2)) =: \beta(\xi) \tag{14}$$

say, and the conclusion follows. □

2.3. Proof of Proposition 1

In keeping with the notation of Sections 2.1 and 2.2. The initial condition $v_0>0$ is fixed throughout this section. Pick $\theta_0\in\Theta$ so that for n large enough, we have $t_n(\theta_0)\in\Theta$. The number θ_0 is fixed from now on. Pick $0<\eta_-<\eta_+$ with η_- small enough so that

$$\eta_- - \frac{v_0}{\sqrt{\theta_0}} < 0. \tag{15}$$

The numbers η_- and η_+ are fixed from now on. For $M>0$, define

$$\mathscr{A}(M) = \{\omega\in\mathscr{C}_0 : \eta_- \leqslant v_0/\sqrt{\theta_0}+\omega(t) \leqslant \eta_+\} \cap \{|\omega|_{1/2,2,\infty} \leqslant M\}.$$

For a function $f\in\mathscr{C}_0$, let $\mathscr{F}(f) = (\int_0^1 f(x)\mathrm{e}^{\mathrm{i}2\pi xk}\,\mathrm{d}x)_{k\in\mathbf{Z}}$ denote its Fourier transform. We have $\mathscr{F}(\tilde{T}_n[\omega]) = (1+(\xi n^{-1/4}/2)\mathscr{F}[n^{1/2}\phi(n^{1/2}\cdot)])\mathscr{F}(\omega)$, and since

$$\left|\frac{\xi n^{-1/4}}{2}\mathscr{F}[n^{1/2}\phi(n^{1/2}\cdot)](k)\right| \leqslant \frac{\xi\|\phi\|_1 n^{-1/4}}{2} < 1$$

for large enough n uniformly in k, the mapping $\omega\mapsto\tilde{T}_n(\omega)$ is continuous bijective, with continuous inverse. It follows that $\mathscr{A}(M)$ can be written as $\tilde{T}_n^{-1}(\mathscr{B})$, with a measurable $\mathscr{B}$, namely $\mathscr{B} = \tilde{T}_n(\mathscr{A}(M)) = (\tilde{T}_n^{-1})^{-1}(\mathscr{A}(M))$.

(a) Let

$$\mu_n(\mathrm{d}\theta) = \tfrac{1}{2}[\delta_{\theta_0}(\mathrm{d}\theta) + \delta_{t_n(\theta_0)}(\mathrm{d}\theta)]$$

be defined as in Section 2.1 and let F be an arbitrary estimator. We have

$$\begin{aligned}
\mathbf{R}_n(F, n^{-1/4}) &= \sup_{\theta\in\Theta} E_\theta^n\{n^{1/2}(F-\theta)^2\} \\
&\geqslant \int_{\Theta\times\mathscr{C}_0} v_n(\mathrm{d}\theta)\mathbf{W}(\mathrm{d}\omega)E_{\theta,\omega}^n\{n^{1/2}(F-\theta)^2\} \\
&\geqslant \frac{1}{2}\int_{\mathscr{A}(M)} E_{\theta_0,\omega}^n\{n^{1/2}(F-\theta_0)^2\}\mathbf{W}(\mathrm{d}\omega) \\
&\quad + \frac{1}{2}\int_{\tilde{T}_n(\mathscr{A}(M))} E_{t_n(\theta_0),\omega}^n\{n^{1/2}[F-t_n(\theta_0)]^2\}\mathbf{W}(\mathrm{d}\omega).
\end{aligned}$$

Define

$$\begin{aligned}
G(\omega) &= \tfrac{1}{2}E_{\theta_0,\omega}^n\{n^{1/2}(F-\theta_0)^2\}, \\
H(\omega) &= \tfrac{1}{2}E_{t_n(\theta_0),\omega}^n\{n^{1/2}[F-t_n(\theta_0)]^2\}.
\end{aligned}$$

Repeating the proof of Proposition 2, we have, for $\lambda > 0$

$$G(\omega) + H(\tilde{T}_n[\omega]) \geqslant \frac{\mathrm{e}^{-\lambda}}{4} n^{1/2}[\theta_0 - t_n(\theta_0)]^2 P_{\theta_0,\omega}^n\left(\frac{\mathrm{d}P^n_{t_n(\theta_0),\tilde{T}_n(\omega)}}{\mathrm{d}P^n_{\theta_0,\omega}} \geqslant \mathrm{e}^{-\lambda}\right). \tag{16}$$

(b) For $\omega \in \mathscr{A}(M)$, we have

$$v_0 + \sqrt{\theta_0}\omega(t) \geqslant \sqrt{\theta_0}\eta_-$$

and by (9):

$$v_0 + t_n(\theta_0)T_n[\omega](t) \geqslant \sqrt{\theta_0}\eta_-(1+\mathrm{o}(1)).$$

Since

$$\|\tilde{T}_n(\omega) - T_n(\omega)\|_\infty \leqslant \frac{\xi}{2}n^{-1/4}\|\omega\|_\infty(1+\|\phi\|_1) \leqslant \frac{\xi}{2}n^{-1/4}\eta_+(1+\|\phi\|_1) = \mathrm{o}(1)$$

on $\mathscr{A}(M)$, we derive

$$v_0 + t_n(\theta_0)\tilde{T}_n[\omega](t) \geqslant \sqrt{\theta_0}\eta_-(1+\mathrm{o}(1)).$$

(c) We plan to apply (i) of Lemma 3. For $\omega \in \mathscr{A}(M)$, we have

$$\alpha(\omega, \xi, \theta_0) \leqslant \sqrt{\theta_0}\xi\left(\frac{5\xi}{8}\eta_+ + c(\phi)M\right)2\eta_+ =: \alpha_1(\xi, M),$$

say, therefore:

$$\|[\sqrt{\theta_0}\omega + v_0]^2 - [\sqrt{t_n(\theta_0)}\tilde{T}_n(\omega) + v_0]^2\|_2 \leqslant \alpha_1(\xi, M)n^{-1/2}(1 + \mathrm{o}(1)).$$

(d) Let us pick [3] ξ so that

$$\frac{2\alpha_1(\xi, M)}{(1/4)\theta_0\eta_-^2} \leqslant 1. \tag{17}$$

The number ξ is fixed from now on. Using (b) and (c), we can now apply Lemma 1 for large enough n, with, for instance $\mu = \frac{1}{2}\sqrt{\theta_0}\eta_-$, $L = 2\alpha_1(\xi, M)/(1/4)\theta_0\eta_-^2 \leqslant 1$ and using that $n^{1/2}[\theta_0 - t_n(\theta_0)]^2 \geqslant \xi^2\theta_0^2$, we can refine (16):

$$G(\omega) + H(\tilde{T}_n[\omega]) \geqslant \frac{\xi^2\theta_0^2}{4}\sup_{\lambda>0} \mathrm{e}^{-\lambda}[1 - (2 + 3\sqrt{3})/6\lambda] =: \alpha_2(M) > 0. \tag{18}$$

(e) We plan to apply Lemma 2 with $\mu = \mathbf{W}$. We first need a lower bound for $\mathbf{W}\{\mathscr{A}(M)\}$. Let us first note that thanks to (15):

$$\mathbf{W}\left\{\omega \in \mathscr{C}_0\colon \eta_- - \frac{v_0}{\sqrt{\theta_0}} \leqslant \omega \leqslant \eta_+\right\} \geqslant c_* > 0. \tag{19}$$

(See Revuz and Yor, 1990, p. 71) for instance; we simply express that the Brownian motion has a positive probability to stay within a given strip around the origin between time 0 and 1.) Besides, for all $z > 0$, there exists $M = M(z)$ such that

$$\mathbf{W}\{|\omega|_{1/2,2,\infty} \leqslant M(z)\} \geqslant 1 - z. \tag{20}$$

Therefore, taking $z = c_*/2$, using (19), we derive

$$\mathbf{W}\{\mathscr{A}(M(z))\} \geqslant c_*/2. \tag{21}$$

The constant z is fixed from now on and so is $M = M(z)$ satisfying (20). By (ii) of Lemma 3, we have (ii) of Lemma 2 with $\beta = \beta(\xi) =: \beta_1$. By (18), we have (i) of Lemma 2 with $\alpha = \alpha_2(M) =: \alpha_3$ on $\mathscr{A}(M(z)) =: \mathscr{A}_1$. Applying Lemma 2, we obtain, for large enough n:

$$\mathbf{R}_n(F, n^{-1/4}) \geqslant \alpha_3 \sup_{\lambda>0} \mathrm{e}^{-\lambda}\left(c_*/2 - \frac{1}{\lambda}\beta_1\right). \qquad \square$$

Corollary 1. *Grant Assumptions A and B and assume that* $\Theta \subset \mathbb{R}^1$. *In the case* $\sigma(x, \theta) = \sqrt{|\theta|}$, *we have Theorem* 1.

Proof. By assumption B1, there exists $v_1 \neq 0$ in V. By assumption B2, Θ contains an open interval, say (θ_-, θ_+). With no loss of generality, we may assume that $v_1 > 0$ and $(\theta_-, \theta_+) \subset (0, +\infty)$. (Otherwise, we can take $v_1 < 0$ and $(\theta_-, \theta_+) \subset (-\infty, 0)$ and apply the same subsequent argument with obvious changes.) By assumption A2, the

[3] The reader can check that $\xi := \min\{1, \sqrt{\theta_0}\eta_-^2/[\eta_+(10\eta_+ + 16c(\phi)M)]\}$ solves (17).

class of admissible drifts contains $b=0$ and $\rho=0$. It follows that

$$\mathbf{R}_n(F,n^{-1/4}) \geqslant \sup_{b=0,\rho=0,\,v_0=v_1} \sup_{\theta\in(\theta_-,\theta_+)} E_\theta^n\{n^{1/2}(F-\theta)^2\}$$

and we conclude by applying Proposition 1. □

We actually proved a stronger result, which will be useful for the proof of the general case: define, for any positive σ-finite measure $\mathbf{M}$ on $\Theta\times\mathscr{C}_0$, the following "Bayes"-type risk:

$$\tilde{\mathbf{B}}_n(\mathbf{M},\varphi_n) = \inf_F \sup_{b,\rho,\,v_0} \int_{\Theta\times\mathscr{C}_0} \mathbf{M}(\mathrm{d}\theta,\mathrm{d}\omega) E_{\theta,\omega}^n\{\varphi_n^{-1}(F-\theta)^2\},$$

where $\varphi_n>0$ is a normalizing factor and the infimum is taken over all estimators.

Corollary 2. *Grant Assumptions A and B, assume that $\Theta\subset\mathbb{R}^1$, and $\sigma(x,\theta)=\sqrt{|\theta|}$. For any θ_0 in the interior of Θ and $\mathscr{A}=\mathscr{A}_1$ (defined in* (e) *in the proof of Proposition* 1 *above), define*

$$\mathbf{M}^{(n)}(\mathrm{d}\theta,\mathrm{d}\omega) = \delta_{\theta_0}(\mathrm{d}\theta)\otimes 1_{\mathscr{A}}(\omega)\mathbf{W}(\mathrm{d}\omega) + \delta_{t_n(\theta_0)}(\mathrm{d}\theta)\otimes 1_{\tilde{T}_n(\mathscr{A})}(\omega)\mathbf{W}(\mathrm{d}\omega).$$

We have

$$\liminf_{n\to\infty} \tilde{\mathbf{B}}_n(\mathbf{M}^{(n)},n^{-1/4})>0.$$

3. The general case

3.1. Preliminaries

We say that two sequences of experiments $\mathscr{E}_i^n=(\Omega_n,\mathscr{A}_n,(\mathbf{P}_{\theta,n}^i,\ \theta\in\Theta))$, for $i=1,2$ defined on common probability spaces and with same parameter space Θ are *strongly equivalent* if there exists a measurable and invertible sequence of mappings $\phi_n:\Omega_n\to\Omega_n$ such that for all $\theta\in\Theta: \phi_n\mathbf{P}_{\theta,n}^1=\mathbf{P}_{\theta,n}^2$ and $\phi_n^{-1}\mathbf{P}_{\theta,n}^2=\mathbf{P}_{\theta,n}^1$. Clearly, minimax and Bayes risk bounds coincide in two strongly equivalent sequences of experiments.

Let $\delta\in(0,1]$ and denote by $\mathscr{G}_\delta^n(\sigma)$ the experiment generated by the observation $(X_{i/n},\ i=0,\dots,\lfloor n\delta\rfloor)$, where X solves (1) with diffusion coefficient σ for the hidden component v_t, where $\lfloor x\rfloor$ denotes the integer part of x. Note that $\mathscr{G}_\delta^n(\cdot)$ and $\mathscr{G}_1^{\lfloor n\delta\rfloor}(\cdot)$ involve the same number of data, namely $\lfloor n\delta\rfloor$.

Lemma 4. *Let $0<\varepsilon_n\leqslant 1$ be such that $n\varepsilon_n$ is a sequence of integers such that $n\varepsilon_n\to\infty$ as $n\to\infty$. Assume that $b=0$ and $\rho=0$. The experiments $\mathscr{G}_1^{n\varepsilon_n}(\sigma)$ and $\mathscr{G}_{\varepsilon_n}^n(\varepsilon_n^{-1/2}\sigma)$ are strongly equivalent.*

Proof. For simplicity, we omit any reference to the initial conditions x_0 and v_0 which are not affected by the transformations under consideration. Let $\tilde{B}_t=(1/\sqrt{\varepsilon_n})B_{\varepsilon_n t}$. By scaling, $\tilde{B}$ is a standard Brownian motion. Thus $\mathscr{G}_{\varepsilon_n}^n(\varepsilon_n^{-1/2}\sigma)$ is given by the observation of $(X_{i/n},\ i=0,\dots,n\varepsilon_n)$, where

$$\mathrm{d}X_t = v_t \mathrm{d}B_t = \sqrt{\varepsilon_n} v_t\, \mathrm{d}\tilde{B}_{t/\varepsilon_n}, \quad v_t = v_0 + \varepsilon_n^{-1/2} \int_0^t \sigma(\theta,\, v_s)\, \mathrm{d}W_s, \qquad t \in [0, \varepsilon_n].$$

Putting $s = t/\varepsilon_n$ and $Y_s = (1/\sqrt{\varepsilon_n})X_{\varepsilon_n s}$, we equivalently observe $(Y_{i/n\varepsilon_n},\ i=0,\dots,n\varepsilon_n)$, with Y satisfying

$$\mathrm{d}Y_s = v_{\varepsilon_n s}\, \mathrm{d}\tilde{B}_s, \quad s \in [0,1]$$

and

$$\mathrm{d}v_{\varepsilon_n s} = \frac{1}{\sqrt{\varepsilon_n}} \sigma(\theta,\, v_{\varepsilon_n s})\, \mathrm{d}W_{\varepsilon_n s} = \sigma(\theta,\, v_{\varepsilon_n s})\, \mathrm{d}\tilde{W}_s,$$

where $\tilde{W}_s = (1/\sqrt{\varepsilon_n})W_{\varepsilon_n s}$ is a standard Brownian motion by scaling. Finally, setting $\tilde{v}_s = v_{s\varepsilon_n}$, the experiment $\mathscr{G}^n_{\varepsilon_n}(\varepsilon_n^{-1/2}\sigma)$ is equivalent to observing $(Y_{i/n\varepsilon_n},\ i=0,\dots,n\varepsilon_n)$, where

$$\mathrm{d}Y_s = \tilde{v}_s\, \mathrm{d}\tilde{B}_s, \quad \mathrm{d}\tilde{v}_s = \sigma(\theta, \tilde{v}_s)\, \mathrm{d}\tilde{W}_s, \qquad s \in [0,1]. \qquad \square$$

3.2. Proof of Theorem 1

With no loss of generality, we assume that $b=0$ and $\rho=0$ and that $V=\{v_0\}$, with $v_0 \neq 0$ satisfying Assumption B3 . The number v_0 is fixed throughout this section.

(a) Let $0<\varepsilon_n<1$. We restrict the experiment to the time interval $[0,\varepsilon_n]$ and dilate the diffusion coefficient σ by the amount $\varepsilon_n^{-1/2}$, i.e. we consider the experiment $\mathscr{G}^n_{\varepsilon_n}(\varepsilon_n^{-1/2}\sigma)$. We denote by $(\mathscr{C}_0^{\varepsilon_n}, \mathscr{F}^{\varepsilon_n}, \mathbf{W}^{\varepsilon_n})$ the Wiener space and measure restricted to the time interval $[0,\varepsilon_n]$. We use the natural correspondence between $\mathscr{C}_0^{\varepsilon_n}$ and $\mathscr{C}_0$:

$$(\omega_t)_{t\in[0,\varepsilon_n]} \in \mathscr{C}_0^{\varepsilon_n} \leftrightarrow (\omega_{\varepsilon_n t})_{t\in[0,1]} \in \mathscr{C}_0. \tag{22}$$

Define the scaling operator

$$\tau^{\varepsilon_n}[(\omega_t)_{t\in[0,1]}] := (\varepsilon_n^{-1/2}\omega_{\varepsilon_n t})_{t\in[0,1]}.$$

If $\mathscr{B} \subset \mathscr{C}_0$, set $\mathscr{B}^{\varepsilon_n} := (\tau^{\varepsilon_n})^{-1}(\mathscr{B})$ and by the correspondence (22), we can inject $\mathscr{B}^{\varepsilon_n}$ in $\mathscr{C}_0^{\varepsilon_n}$ in a canonical way. Since $\mathbf{W}$ is τ^{ε_n}-invariant, we thus have, for any positive functional F on $\mathscr{C}_0$:

$$\int_{\mathscr{B}} F(\omega)\mathbf{W}(\mathrm{d}\omega) = \int_{\mathscr{B}^{\varepsilon_n}} F(\tau^{\varepsilon_n}\omega)\mathbf{W}(\mathrm{d}\omega) = \int_{\mathscr{B}^{\varepsilon_n}} F(\varepsilon_n^{-1/2}\omega)\mathbf{W}^{\varepsilon_n}(\mathrm{d}\omega),$$

the last integral being understood as an integral on $\mathscr{C}_0^{\varepsilon_n}$.

(b) For clarity, we further denote by $P^n_{\sigma(\theta,\cdot)}$ and $P^{n,\varepsilon_n}_{\varepsilon_n^{-1/2}\sigma(\theta,\cdot)}$ the law of the observation in the experiments $\mathscr{G}^n_1(\sigma)$ and $\mathscr{G}^n_{\varepsilon_n}(\varepsilon_n^{-1/2}\sigma)$, respectively. We will eventually need the disintegration of $P^{n,\varepsilon_n}_{\varepsilon_n^{-1/2}\sigma(\theta,\cdot)}$ conditional on $\mathbf{W}^{\varepsilon_n}$, which we write as

$$P^{n,\varepsilon_n}_{\varepsilon_n^{-1/2}\sigma(\theta,\cdot)}(\mathrm{d}x_1\dots\mathrm{d}x_{n\varepsilon_n}) = \int_{\mathscr{C}_0^{\varepsilon_n}} \mathscr{P}^{n,\varepsilon_n}_{\varepsilon_n^{-1/2}\sigma(\theta,\cdot)}(\omega, \mathrm{d}x_1\dots\mathrm{d}x_{n\varepsilon_n})\mathbf{W}^{\varepsilon_n}(\mathrm{d}\omega).$$

Let F_{ε_n} be an estimator in $\mathscr{G}^n_{\varepsilon_n}(\varepsilon_n^{-1/2}\sigma)$ for which there exists $K>0$ independent of θ such that

$$|\theta - F_{\varepsilon_n}| \leqslant K, \tag{23}$$

a choice which is obviously possible since Θ is bounded by Assumption B2. For two arbitrary $\theta_1,\theta_2 \in \Theta$ and $\mathscr{B}_1,\mathscr{B}_2 \subset \mathscr{C}_0^{\varepsilon_n}$, we successively have

$$\begin{aligned}
&\sup_{\theta\in\Theta} E^{n,\varepsilon_n}_{\varepsilon_n^{-1/2}\sigma(\theta,\cdot)}\{|F_{\varepsilon_n}-\theta|^2\}\\
&\quad\geqslant \frac{1}{2}\sum_{i=1}^{2}\int_{\mathscr{B}_i}\mathbf{W}^{\varepsilon_n}(\mathrm{d}\omega)\int_{\mathbb{R}^{\lfloor n\varepsilon_n\rfloor}}|F_{\varepsilon_n}(x_1,\ldots,x_{n\varepsilon_n})-\theta_i|^2\mathscr{P}^{n,\varepsilon_n}_{\varepsilon_n^{-1/2}\sigma(\theta_i,\cdot)}(\omega,\mathrm{d}x_1\ldots\mathrm{d}x_{n\varepsilon_n})\\
&\quad\geqslant \frac{1-\delta}{2}\sum_{i=1}^{2}\int_{\mathscr{B}_i}\mathbf{W}^{\varepsilon_n}(\mathrm{d}\omega)\int_{\mathbb{R}^{n\varepsilon_n}}|F_{\varepsilon_n}(x_1,\ldots,x_{n\varepsilon_n})-\theta_i|^2\\
&\qquad\times 1_{\zeta_{n,i}(\omega;\,x_1,\ldots,x_{n\varepsilon_n})\geqslant 1-\delta}\}\mathscr{P}^{n,\varepsilon_n}_{\varepsilon_n^{-1/2}\sigma(\theta_i,\,v_0)}(\omega,\mathrm{d}x_1\ldots\mathrm{d}x_{n\varepsilon_n}),
\end{aligned}$$

for all $0<\delta<1$ and where

$$\zeta_{n,i}(\omega;x_1,\ldots,x_{n\varepsilon_n}) = \frac{\mathrm{d}\mathscr{P}^{n,\varepsilon_n}_{\varepsilon_n^{-1/2}\sigma(\theta_i,\cdot)}(\omega,\cdot)}{\mathrm{d}\mathscr{P}^{n,\varepsilon_n}_{\varepsilon_n^{-1/2}\sigma(\theta_i,\,v_0)}(\omega,\cdot)}(x_1,\ldots,x_{n\varepsilon_n})$$

denotes the conditional likelihood process between the scaled experiments with diffusion coefficients $\varepsilon_n^{-1/2}\sigma(\theta,\cdot)$ and $\varepsilon_n^{-1/2}\sigma(\theta,v_0)$, respectively. We introduce the following parametrization:

$$\tilde{\theta} := \sigma(\theta,v_0), \quad \tilde{\Theta} := \sigma(\Theta,v_0) \subset \mathbb{R}^1$$

and consider the subexperiment with parameter $\tilde{\theta}$ instead of θ from now on. By Assumption B3, there exists $c_2 = c_2(\Theta,V,C_1,C_2)$ such that

$$(F_{\varepsilon_n}-\theta)^2 \geqslant c_2^2(\sigma(\theta,v_0)-\sigma(F_{\varepsilon_n},v_0))^2 =: c_2^2(\tilde{\theta}-G_{\varepsilon_n})^2, \text{ say.}$$

We further omit the dependence in $(x_1,\ldots,x_{n\varepsilon_n})$ in the notation. It follows that

$$\begin{aligned}
&\frac{1-\delta}{2}\sum_{i=1}^{2}\int_{\mathscr{B}_i}\mathbf{W}^{\varepsilon_n}(\mathrm{d}\omega)\int_{\mathbb{R}^{n\varepsilon_n}}|F_{\varepsilon_n}-\theta_i|^2 1_{\zeta_{n,i}(\omega,\cdot)\geqslant 1-\delta}\,\mathrm{d}\mathscr{P}^{n,\varepsilon_n}_{\varepsilon_n^{-1/2}\sigma(\theta_i,\,v_0)}(\omega,\cdot)\\
&\quad\geqslant c_2^2\frac{1-\delta}{2}\sum_{i=1}^{2}\int_{\mathscr{B}_i}\mathbf{W}^{\varepsilon_n}(\mathrm{d}\omega)\int_{\mathbb{R}^{n\varepsilon_n}}(\tilde{\theta}_i-G_{\varepsilon_n})^2\,\mathrm{d}\mathscr{P}^{n,\varepsilon_n}_{\varepsilon_n^{-1/2}\tilde{\theta}_i}(\omega,\cdot)\\
&\qquad - K^2c_2^2\frac{(1-\delta)}{2}\int_{\mathscr{B}_i}\mathbf{W}^{\varepsilon_n}(\mathrm{d}\omega)\int_{\mathbb{R}^{n\varepsilon_n}}1_{\zeta_{n,i}(\omega,\cdot)<1-\delta}\,\mathrm{d}\mathscr{P}^{n,\varepsilon_n}_{\varepsilon_n^{-1/2}\tilde{\theta}_i}(\omega,\cdot) \quad \text{by (23).}
\end{aligned}$$

(c) For the choice $\mathscr{B}_1 = \mathscr{A}^{\varepsilon_n}$—recall the definition of $\mathscr{A}$ in Section 2.3, taking $\mathscr{B}_2 = \tilde{T}_{n\varepsilon_n}(\mathscr{A}^{\varepsilon_n})$, $\tilde{\theta}_1$ in the interior of $\tilde{\Theta}$ and $\tilde{\theta}_2 = t_{n\varepsilon_n}(\tilde{\theta}_1)$, the last quantity is equal to

$$c_2^2\frac{(1-\delta)}{2}\int_{\tilde{\Theta}\times\mathscr{C}_0^{\varepsilon_n}} \mathbf{M}^{(n\varepsilon_n)}(\mathrm{d}\tilde{\theta},\mathrm{d}\omega)\int_{\mathbb{R}^{n\varepsilon_n}} (\tilde{\theta}-G_{\varepsilon_n})^2\,\mathrm{d}\mathscr{P}^{n,\varepsilon_n}_{\varepsilon_n^{-1/2}\tilde{\theta}}(\omega,\cdot)$$

$$-K^2c_2^2\frac{(1-\delta)}{2}\sum_{i=1}^{2}\int_{\mathscr{B}_i}\mathbf{W}^{\varepsilon_n}(\mathrm{d}\omega)\int_{\mathbb{R}^{n\varepsilon_n}} 1_{\zeta_{n,i}(\omega,\cdot)<1-\delta}\,\mathrm{d}\mathscr{P}^{n,\varepsilon_n}_{\varepsilon_n^{-1/2}\sigma(\theta_i,v_0)}(\omega,\cdot), \quad (24)$$

where the measure $\mathbf{M}^{(n\varepsilon_n)}$ is defined in Corollary 2 above and is specified by $\tilde{\theta}_1$ and $\mathscr{A}^{\varepsilon_n}$. Moreover, there exists a constant $c_3 = c_3(\theta_1, V, C_1, C_2)$ such that for large enough n:

$$\int_{\tilde{\Theta}\times\mathscr{C}_0^{\varepsilon_n}} \mathbf{M}^{(n\varepsilon_n)}(\mathrm{d}\tilde{\theta},\mathrm{d}\omega)\int_{\mathbb{R}^{n\varepsilon_n}} (\tilde{\theta}-G_{\varepsilon_n})^2\,\mathrm{d}\mathscr{P}^{n,\varepsilon_n}_{\varepsilon_n^{-1/2}\tilde{\theta}}(\omega,\cdot) \geqslant c_3(n\varepsilon_n)^{-1/2},$$

otherwise, the rate $(n\varepsilon_n)^{-1/4}$ could be improved in $\mathscr{G}^n_{\varepsilon_n}(\varepsilon_n^{-1/2}\tilde{\theta})$, and by the equivalence Lemma 4, the rate $(n\varepsilon_n)^{-1/4}$ could be improved in $\mathscr{G}_1^{n\varepsilon_n}(\tilde{\theta})$, but, since $n\varepsilon_n \to \infty$, that would contradict Corollary 2.

(d) We now turn to the study of the remainder term. Recall that $\mathscr{B}_1 = \mathscr{A}^{\varepsilon_n}$. We have

$$\begin{aligned}\mathscr{A}^{\varepsilon_n} &= \{\omega\in\mathscr{C}_0 : \eta_- \leqslant v_0/\sqrt{\tilde{\theta}_1} + \tau^{\varepsilon_n}\omega \leqslant \eta_+\} \cap \{|\tau^{\varepsilon_n}\omega|_{1/2,2,\infty} \leqslant M\}\\ &= \{\omega\in\mathscr{C}_0^{\varepsilon_n} : \eta_- \leqslant v_0/\sqrt{\tilde{\theta}_1} + \varepsilon_n^{-1/2}\omega \leqslant \eta_+\} \cap \{|\varepsilon_n^{-1/2}\omega|_{1/2,2,\infty} \leqslant M\}.\end{aligned}$$

Thus, on $\mathscr{A}^{\varepsilon_n}$, we have

$$\inf_{t\in[0,\varepsilon_n]} (v_0 + \varepsilon_n^{-1/2}\sqrt{\tilde{\theta}_1}\omega_t) \geqslant \sqrt{\tilde{\theta}_1}\eta_-.$$

We need the following modification of Lemma 1—recall Section 2—, which proof[4] is delayed until appendix:

Lemma 5. *Let φ_0, φ_1 be two continuous functions on $[0,1]$ such that*:
(i) $\inf_t \varphi_1(t) \geqslant \mu > 0$, *for some* $\mu > 0$,
(ii) $\|\varphi_1^2 - \varphi_0^2\|_\infty \leqslant \frac{1}{5}\mu^2$,
(iii) *For some* $L > 0$:

$$\|\varphi_1^2 - \varphi_0^2\|_2 \leqslant \frac{\mu^2 L}{\sqrt{n}}.$$

Let $0 < \delta < 1$. *For* $n \geqslant 1$, *we have*

$$Q^n_{\varphi_1}\left\{\frac{\mathrm{d}Q^n_{\varphi_0}}{\mathrm{d}Q^n_{\varphi_1}} \leqslant 1-\delta\right\} \leqslant \left(\log\frac{1}{1-\delta}\right)^{-1} L(3L/2 + \sqrt{3}).$$

[4] Lemma 5 is slightly different from Lemma 1: the lower estimate assumption is now on φ_1 and not φ_0 while the upper bound is still given under $Q^n_{\varphi_1}$.

Now, letting $L = L_n$ depend on n, replacing $Q^n_{\varphi_i}$ by $Q^{n\varepsilon_n}_{\varphi_i}$, with parameters $\varphi_0 = v_0 + \varepsilon_n^{-1/2}\int_0^t \sigma(\theta_1, v_s)\,\mathrm{d}y_s(\omega)$ and $\varphi_1 = v_0 + \varepsilon_n^{-1/2}\sigma(\theta_1, v_0)\omega_t$, we readily obtain, for $\omega \in \mathscr{A}^{\varepsilon_n}$ and large enough n:

$$\int_{\mathbb{R}^{n\varepsilon_n}} 1_{\zeta_{n,1}(\omega,\cdot)<1-\delta}\,\mathrm{d}\mathscr{P}^{n,\varepsilon_n}_{\varepsilon_n^{-1/2}\sigma(\theta_1,v_0)}(\omega,\cdot) \leqslant \left(\log\frac{1}{1-\delta}\right)^{-1} L_n(3L_n/2 + \sqrt{3}),$$

with

$$L_n^2 = \frac{n}{\tilde{\theta}_1^2\eta_-^4}\int_0^{\varepsilon_n}\left\{[v_0 + \varepsilon_n^{-1/2}\sigma(\theta_1, v_0)\omega_t]^2 - \left[v_0 + \varepsilon_n^{-1/2}\int_0^t \sigma(\theta_1, v_s)\,\mathrm{d}y_s(\omega)\right]^2\right\}^2 \mathrm{d}t.$$

Integrating w.r.t. $\mathbf{W}^{\varepsilon_n}$, using $1_{\mathscr{A}^{\varepsilon_n}} \leqslant 1$ and Cauchy–Schwarz, we obtain

$$\int_{\mathscr{A}^{\varepsilon_n}} \mathbf{W}^{\varepsilon_n}(\mathrm{d}\omega)\int_{\mathbb{R}^{n\varepsilon_n}} 1_{\zeta_{n,1}(\omega,\cdot)<1-\delta}\,\mathrm{d}\mathscr{P}^{n,\varepsilon_n}_{\varepsilon_n^{-1/2}\sigma(\theta_1,v_0)}(\omega,\cdot)$$
$$\leqslant \frac{3}{2}\left(\log\frac{1}{1-\delta}\right)^{-1}\{[\mathbf{E}^{\varepsilon_n}(L_n^2)]^{1/2} + \mathbf{E}^{\varepsilon_n}(L_n^2)\}. \tag{25}$$

We see that the order of (25) is given by the term $[\mathbf{E}^{\varepsilon_n}(L_n^2)]^{1/2}$. We need the following preliminary decomposition:

$$[v_0 + \varepsilon_n^{-1/2}\sigma(\theta_1, v_0)\omega_t]^2 - \left[v_0 + \varepsilon_n^{-1/2}\int_0^t \sigma(\theta_1, v_s)\,\mathrm{d}y_s(\omega)\right]^2$$
$$= \varepsilon_n^{-1/2}\int_0^t [\sigma(\theta_1, v_0) - \sigma(\theta_1, v_s)]\,\mathrm{d}y_s(\omega) \times \left\{\varepsilon_n^{-1/2}\int_0^t [\sigma(\theta_1, v_0) + \sigma(\theta_1, v_s)]\,\mathrm{d}y_s(\omega) + 2v_0\right\}$$
$$=: \varepsilon_n^{-1/2}\int_0^t [\sigma(\theta_1, v_0) - \sigma(\theta_1, v_s)]\,\mathrm{d}y_s(\omega) Z_t^{(n)}(\omega) \quad \text{say}.$$

First, the Burckholder–Davis–Gundy inequality yields the following estimate:

$$\mathbf{E}^{\varepsilon_n}\left\{\sup_{t\in[0,\varepsilon_n]} (Z_t^{(n)})^4\right\} \leqslant c_4 = c_4(\theta_1, C_1, C_2).$$

Second, applying repeatedly Cauchy–Schwarz, BDG and Jensen, the term $\mathbf{E}^{\varepsilon_n}\{L_n^2\}$ is less than

$$\frac{n}{\tilde{\theta}_1^2\eta_-^4}\int_0^{\varepsilon_n}\left(\mathbf{E}^{\varepsilon_n}\left\{\varepsilon_n^{-2}\left(\int_0^t [\sigma(\theta_1, v_0) - \sigma(\theta_1, v_s)]\,\mathrm{d}y_s(\omega)\right)^4\right\}\right)^{1/2}\mathrm{d}t$$
$$\times \left(\mathbf{E}^{\varepsilon_n}\left\{\sup_{t\in[0,\varepsilon_n]} (Z_t^{(n)})^4\right\}\right)^{1/2}$$

$$\leqslant \frac{n}{\tilde{\theta}_1^2 \eta_-^4} \sqrt{c_4} \varepsilon_n^{-1} \int_0^{\varepsilon_n} \left(\mathbf{E}^{\varepsilon_n} \left\{ \left(\int_0^t [\sigma(\theta_1, v_0) - \sigma(\theta_1, v_s)]^2 \, \mathrm{d}s \right)^2 \right\} \right)^{1/2} \mathrm{d}t$$

$$\leqslant \frac{n}{\tilde{\theta}_1^2 \eta_-^4} \sqrt{c_4} \varepsilon_n^{-1} \int_0^{\varepsilon_n} \left(t \int_0^t \mathbf{E}^{\varepsilon_n} \left\{ [\sigma(\theta_1, v_0) - \sigma(\theta_1, v_s)]^4 \right\} \mathrm{d}s \right)^{1/2} \mathrm{d}t$$

$$\leqslant \frac{n}{\tilde{\theta}_1^2 \eta_-^4} \sqrt{c_4} c_5(\theta_1, C_1, C_2) \varepsilon_n^{-1} \int_0^{\varepsilon_n} \left(t \int_0^t s^2 \, \mathrm{d}s \right)^{1/2} \mathrm{d}t$$

$$\leqslant c_6 n \varepsilon_n^2$$

for some $c_6 = c_6(\theta_1, C_1, C_2) > 0$. In view of (25),

$$\int_{\mathscr{A}^{\varepsilon_n}} \mathbf{W}^{\varepsilon_n}(\mathrm{d}\omega) \int_{\mathbb{R}^{n\varepsilon_n}} 1_{\zeta_{n,1}(\omega,\,\cdot) < 1-\delta} \, \mathrm{d}\mathscr{P}^{n,\varepsilon_n}_{\varepsilon_n^{-1/2}\sigma(\theta_1,\, v_0)}(\omega, \cdot) \quad \text{is of order } \sqrt{n}\varepsilon_n.$$

Picking ε_n of the order n^{-a}, with $\frac{2}{3} < a < 1$, we see that $\sqrt{n}\varepsilon_n = \mathrm{o}((n\varepsilon_n)^{-1/2})$, so the second term in (24) is asymptotically negligible. The same technique applies for $\mathscr{B}_2$. We omit the details. In conclusion, for ε_n of the order n^{-a} with $\frac{2}{3} < a < 1$, we have:

$$\liminf_{n \to \infty} \inf_{F_{\varepsilon_n}} \sup_{\theta \in \Theta} E^{n,\varepsilon_n}_{\varepsilon_n^{-1/2}\sigma(\theta,\,\cdot)} \{ (n\varepsilon_n)^{1/2} |F_{\varepsilon_n} - \theta|^2 \} > 0.$$

Since Θ is bounded, the infimum is clearly attained among the estimators satisfying (23) so the above infimum is taken over all estimators. By the equivalence Lemma 4, the rate $(n\varepsilon_n)^{-1/4}$ is also a lower bound in $\mathscr{G}_1^{n\varepsilon_n}(\sigma)$. Since $n\varepsilon_n \to \infty$, the conclusion follows.

Acknowledgements

Thanks to Sylvain Delattre, Emmanuel Gobet and Dominique Picard for fruitful discussions. The careful reading and comments of Arnaud Gloter helped to improve considerably a former version of the paper. This work was partially supported by the Institute of Statistical Mathematics, Tokyo.

Appendix A.

A.1. Proof of Lemma 1

Proof. By Chebyshev's inequality

$$Q^n_{\varphi_1} \left\{ \frac{\mathrm{d}Q^n_{\varphi_0}}{\mathrm{d}Q^n_{\varphi_1}} \geqslant \mathrm{e}^{-\lambda} \right\} \geqslant 1 - \frac{1}{\lambda} E_{Q^n_{\varphi_1}} \left\{ \left| \log \frac{\mathrm{d}Q^n_{\varphi_0}}{\mathrm{d}Q^n_{\varphi_1}} \right| \right\}. \tag{A.1}$$

Under $Q^n_{\varphi_1}$, the process $Y_{i/n},\ i=0,\dots,n$ is an inhomogeneous Markov chain with initial value $Y_0=x_0$, and transition probability

$$Q^n_{\varphi_1}\{Y_{i/n}\in \mathrm{d}y \mid Y_{(i-1)/n}=x\}=\frac{\sqrt{n}}{\sqrt{2\pi\delta_i^n(\varphi_1)}}\exp\left(-\frac{n(y-x)^2}{2\delta_i^n(\varphi_1)}\right)\mathrm{d}y,$$

where we denote $\delta_i^n(f)=n\int_{(i-1)/n}^{i/n}f^2(s)\,\mathrm{d}s$. We thus have, under $Q^n_{\varphi_1}$:

$$\log\frac{\mathrm{d}Q^n_{\varphi_0}}{\mathrm{d}Q^n_{\varphi_1}}=\frac{1}{2}\sum_{i=1}^n\left[\log\frac{\delta_i^n(\varphi_1)}{\delta_i^n(\varphi_0)}-\left(\frac{1}{\delta_i^n(\varphi_0)}-\frac{1}{\delta_i^n(\varphi_1)}\right)(\Delta_i^nY)^2\right],$$

where $\Delta_i^nY=\sqrt{n}(Y_{i/n}-Y_{(i-1)/n})$. Writing $(\Delta_i^nY)^2=\delta_i^n(\varphi_1)+\varepsilon_i^n(\varphi_1)$, we thus obtain, under $Q^n_{\varphi_1}$:

$$\log\frac{\mathrm{d}Q^n_{\varphi_0}}{\mathrm{d}Q^n_{\varphi_1}}=\frac{1}{2}\sum_{i=1}^n\left[\log\frac{\delta_i^n(\varphi_1)}{\delta_i^n(\varphi_0)}+1-\frac{\delta_i^n(\varphi_1)}{\delta_i^n(\varphi_0)}+\left(\frac{1}{\delta_i^n(\varphi_1)}-\frac{1}{\delta_i^n(\varphi_0)}\right)\varepsilon_i^n(\varphi_1)\right].$$

By (A.1), it suffices to bound the expectation of the above sum. Define $\xi_i^n(\varphi_1,\varphi_0)$ by $\delta_i^n(\varphi_1)/\delta_i^n(\varphi_0)=1+\xi_i^n(\varphi_1,\varphi_0)$. By Taylor's formula

$$\left|\log\frac{\delta_i^n(\varphi_1)}{\delta_i^n(\varphi_0)}+1-\frac{\delta_i^n(\varphi_1)}{\delta_i^n(\varphi_0)}\right|=|\log(1+\xi_i^n)-\xi_i^n|\leqslant\frac{2}{3}(\xi_i^n)^2$$

since $|\log(1+x)-x|\leqslant\frac{2}{3}x^2$ for $|x|\leqslant\frac{1}{5}$ and

$$|\xi_i^n|=\left|\frac{\delta_i^n(\varphi_1)}{\delta_i^n(\varphi_0)}-1\right|\leqslant\mu^{-2}n\int_{(i-1)/n}^{i/n}|\varphi_1^2(s)-\varphi_0^2(s)|\,\mathrm{d}s\quad\text{by (i)}$$

$$\leqslant\mu^{-2}\|\varphi_1^2-\varphi_0^2\|_\infty\leqslant\frac{1}{5}\quad\text{by (ii)}.$$

It follows that

$$\frac{1}{3}\sum_{i=1}^n(\xi_i^n)^2\leqslant\frac{1}{3\mu^4}\sum_{i=1}^n[\delta_i^n(\varphi_1)-\delta_i^n(\varphi_0)]^2.$$

By Jensen's inequality, this last quantity is less than

$$\frac{n}{3\mu^4}\|\varphi_0^2-\varphi_1^2\|_2^2\leqslant L^2/3. \tag{A.2}$$

By Cauchy–Schwarz, the second term is controlled by its variance, namely, since the $\varepsilon_i^n(\varphi_1)$ are independent, the following quantity:

$$\frac{1}{4}\sum_{i=1}^n\left(\frac{1}{\delta_i^n(\varphi_1)}-\frac{1}{\delta_i^n(\varphi_0)}\right)^2E_{Q^n_{\varphi_1}}\{[\varepsilon_i^n(\varphi_1)]^2\}=\frac{1}{4}\sum_{i=1}^n(\xi_i^n)^2E_{Q^n_{\varphi_1}}\left\{\left(\frac{\varepsilon_i^n(\varphi_1)}{\delta_i^n(\varphi_1)}\right)^2\right\}.$$

Since $[\varepsilon_i^n(\varphi_1)]^2=[\delta_i^n(\varphi_1)]^2[\mathbf{N}(0,1)]^4$ in law, the expectation is equal to 3. Therefore, by the same argument used to get (A.2), the above variance is less than $\frac{3}{4}L^2$. After the normalization due to Cauchy–Schwarz, we eventually derive

$$E_{Q^n_{\varphi_1}}\left\{\left|\log\frac{\mathrm{d}Q^n_{\varphi_0}}{\mathrm{d}Q^n_{\varphi_1}}\right|\right\}\leqslant L(L/3+\sqrt{3}/2).\qquad\square$$

A.2. Proof of Lemma 5

Proof. By Chebyshev's inequality

$$Q^n_{\varphi_1}\left\{\frac{\mathrm{d}Q^n_{\varphi_0}}{\mathrm{d}Q^n_{\varphi_1}} \leqslant 1-\delta\right\} \leqslant \left(\log\frac{1}{1-\delta}\right)^{-1} E_{Q^n_{\varphi_1}}\left\{\left|\log\frac{\mathrm{d}Q^n_{\varphi_1}}{\mathrm{d}Q^n_{\varphi_0}}\right|\right\}.$$

Under $Q^n_{\varphi_1}$:

$$\log\frac{\mathrm{d}Q^n_{\varphi_1}}{\mathrm{d}Q^n_{\varphi_0}} = \frac{1}{2}\sum_{i=1}^{n}\left[\log\frac{\delta^n_i(\varphi_0)}{\delta^n_i(\varphi_1)} - \left(\frac{1}{\delta^n_i(\varphi_1)} - \frac{1}{\delta^n_i(\varphi_0)}\right)(\Delta^n_i Y)^2\right],$$

where $\Delta^n_i Y = \sqrt{n}(Y_{i/n} - Y_{(i-1)/n})$. Writing again $(\Delta^n_i Y)^2 = \delta^n_i(\varphi_1) + \varepsilon^n_i(\varphi_1)$, we thus obtain, under $Q^n_{\varphi_1}$:

$$\log\frac{\mathrm{d}Q^n_{\varphi_1}}{\mathrm{d}Q^n_{\varphi_0}} = \frac{1}{2}\sum_{i=1}^{n}\left\{\log\frac{\delta^n_i(\varphi_0)}{\delta^n_i(\varphi_1)} - 1 + \frac{\delta^n_i(\varphi_1)}{\delta^n_i(\varphi_0)} + \left(\frac{1}{\delta^n_i(\varphi_1)} - \frac{1}{\delta^n_i(\varphi_0)}\right)\varepsilon^n_i(\varphi_1)\right\}.$$

Define $\xi^n_i(\varphi_0,\varphi_1)$ by $\delta^n_i(\varphi_0)/\delta^n_i(\varphi_1) = 1 + \xi^n_i(\varphi_0,\varphi_1)$. Since

$$|\xi^n_i| = \left|\frac{\delta^n_i(\varphi_0)}{\delta^n_i(\varphi_1)} - 1\right| \leqslant \mu^{-2} n\int_{(i-1)/n}^{i/n} |\varphi_0^2(s) - \varphi_1^2(s)|\,\mathrm{d}s \quad \text{by (i)}$$

$$\leqslant \mu^{-2}\|\varphi_0^2 - \varphi_1^2\|_\infty \leqslant \frac{1}{5} \quad \text{by (ii)}$$

and using that $|\log(1+x) - 1 + (1+x)^{-1}| \leqslant 3x^2$ for $|x| \leqslant \frac{1}{5}$, it follows that the determinist term is bounded by

$$\frac{3}{2}\sum_{i=1}^{n}(\xi^n_i)^2 \leqslant \frac{1}{3\mu^4}\sum_{i=1}^{n}[\delta^n_i(\varphi_1) - \delta^n_i(\varphi_0)]^2.$$

By Jensen's inequality, this last quantity is less than

$$\frac{3n}{2\mu^4}\|\varphi_0^2 - \varphi_1^2\|_2^2 \leqslant 3L^2/2. \tag{A.3}$$

By Cauchy–Schwarz, the stochastic term is controlled by its variance, namely, since the $\varepsilon^n_i(\varphi_1)$ are independent:

$$\frac{1}{4}\sum_{i=1}^{n}\left(\frac{1}{\delta^n_i(\varphi_1)} - \frac{1}{\delta^n_i(\varphi_0)}\right)^2 E_{Q^n_{\varphi_1}}\{[\varepsilon^n_i(\varphi_1)]^2\}$$

$$= \frac{1}{4}\sum_{i=1}^{n}[1 - 1/(1+\xi^n_i)]^2 E_{Q^n_{\varphi_1}}\left\{\left(\frac{\varepsilon^n_i(\varphi_1)}{\delta^n_i(\varphi_1)}\right)^2\right\}.$$

Since $[\varepsilon^n_i(\varphi_1)]^2 = [\delta^n_i(\varphi_1)]^2[\mathbf{N}(0,1)]^4$ in law, the expectation is equal to 3. Therefore, using $|1-(1+x)^{-1}| \leqslant 2|x|$ for $|x| \leqslant \frac{1}{5}$ and the same arguments as for (A.3), the above variance is less than $3L^2$. After the normalization due to Cauchy–Schwarz, we eventually derive

$$E_{Q^n_{\varphi_1}}\left\{\left|\log\frac{\mathrm{d}Q^n_{\varphi_0}}{\mathrm{d}Q^n_{\varphi_1}}\right|\right\} \leqslant L(3L/2 + \sqrt{3}). \qquad \square$$

References

Chesney, M., Scott, L., 1989. Pricing European currency options: A comparison of the modified Black-Scholes model and a random variance model. J. Financial Quantitive Anal. 24, 267–289.

Genon-Catalot, V., Jeantheau, T., Laredo, C., 1998. Limit theorems for discretely observed stochastic volatility models. Bernoulli 4, 283–303.

Genon-Catalot, V., Jeantheau, T., Laredo, C., 2000a. Parameter estimation for discretely observed stochastic volatility models. Bernoulli 5, 855–872.

Genon-Catalot, V., Jeantheau, T., Laredo, C., 2000b. Stochastic volatility models as hidden Markov models and statistical applications. Bernoulli 6, 1051–1079.

Gloter, A., 2000. Estimation du coefficient de diffusion de la volatilité d'un modèle à volatilité stochastique. CR Acad. Sci. Paris, t330, Série I, pp. 243–248.

Heston, S.L., 1993. A closed-form solution for options with stochastic volatility with applications to bond and currency options. Rev. Financial Studies 6, 327–343.

Hull, J., White, A., 1988. An analysis of the bias in option pricing caused by a stochastic volatility. Adv. Futures Options Res. 3, 29–61.

Kerkyacharian, G., Picard, D., 1997. Limit of the quadratic risk in density estimation using linear methods. Statist. Probab. Lett. 31, 299–312.

Melino, A., Turnbull, S.M., 1990. Pricing foreign currency options with stochastic volatility. J. Econometrics 45, 239–265.

Revuz, D., Yor, M., 1990. Continuous Martingales and Brownian Motion. Springer, Berlin.

Roynette, B., 1993. Mouvement brownien et espaces de Besov. Stochastics Stochastic Reports 43, 221–260.

Sørensen, M., 1998. Prediction-based estimating functions. Technical report, Department of Theoretical Statistics, University of Copenhagen.

Wiggins, J.B., 1987. Option values under stochastic volatility: theory and empirical estimates. J. Financial Econom. 19, 351–372.