ESAIM: Probability and Statistics December 2001, Vol. 5, 225–242
URL: http://www.emath.fr/ps/

DIFFUSIONS WITH MEASUREMENT ERRORS. I. LOCAL ASYMPTOTIC NORMALITY

Arnaud Gloter[1] and Jean Jacod[2]

Abstract. We consider a diffusion process X which is observed at times i/n for $i = 0, 1, \ldots, n$, each observation being subject to a measurement error. All errors are independent and centered Gaussian with known variance ρ_n. There is an unknown parameter within the diffusion coefficient, to be estimated. In this first paper the case when X is indeed a Gaussian martingale is examined: we can prove that the LAN property holds under quite weak smoothness assumptions, with an explicit limiting Fisher information. What is perhaps the most interesting is the rate at which this convergence takes place: it is $1/\sqrt{n}$ (as when there is no measurement error) when ρ_n goes fast enough to 0, namely $n\rho_n$ is bounded. Otherwise, and provided the sequence ρ_n itself is bounded, the rate is $(\rho_n/n)^{1/4}$. In particular if $\rho_n = \rho$ does not depend on n, we get a rate $n^{-1/4}$.

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

Received February 16, 2001. Revised October 24, 2001.

1. Introduction

1) In this paper as well as in the companion paper [6], we are interested in the following problem: let X be a diffusion process on the time interval $[0, 1]$, whose law depends on a parameter θ. We observe this process at times i/n for $i = 0, 1, \ldots, n$, and each observation is blurred by an error which is centered normal with variance ρ_n; in other words we observe $X_{\frac{i}{n}} + \sqrt{\rho_n}\, U_i$ where the U_i's are i.i.d. $\mathcal{N}(0, 1)$, independent of the process. Our aim is to estimate the parameter θ, knowing the noise level ρ_n.

This problem is clearly of pratical relevance but does not seem to have been studied so far, with the exception of a recent paper by Malyutov and Bayborodin [12] where no attempt towards optimality is made. The model is a hidden Markov model for which a lot is known (for "optimal" inference for such models, see *e.g.* Bickel and Ritov [1], [2], Jensen and Pedersen [9], Leroux [10] or Ryden [13]). However the situation at hand differs from ordinary Markov hidden models in which the hidden Markov chain is typically homogeneous ergodic, and time goes to infinity. Here the hidden Markov chain, *i.e.* the sequence $X_{i/n}, i = 0, 1, \ldots$, has a transition kernel depending on n which degenerates as $n \to \infty$, while no ergodic property is relevant here.

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

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

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

2) When there is no observation noise (*i.e.* $\rho_n = 0$) the question is completely solved for diffusions of the form

$$\mathrm{d}X_t \;=\; b_t \mathrm{d}t + a(\theta, t, X_t)\mathrm{d}W_t, \qquad \mathcal{L}(X_0) = \eta. \tag{1.1}$$

Here W is a standard Brownian motion, and η is an arbitrary initial law on $\mathbb{R}$, and b is (non-anticipative) drift term which may depend on the path of X or W. In fact, and of course assuming some smoothness of a, Donhal [3] proved the LAMN (= Local Asymptotic Mixed Normal) property in the 1-dimensional case for a bounded away from 0 and $b_t = b(t, X_t)$, and in [4,5] we have proved the LAMN property in the d-dimensional case when a is a gradient and exhibited optimal estimators in all cases, and Gobet [7] has shown the LAMN property in general, when a is non-degenerate. It turns out that the convergence rate in the LAMN property and for all optimal sequences of estimators is $1/\sqrt{n}$.

Observe that in (1.1) the parameter θ does not appear in the drift b_t; it could appear but this would not improve the estimation since when θ appears in the drift only, it is not identifiable on the basis of even the full observation of X over any bounded time interval.

3) When noise is present, several questions are natural. First, for a fixed noise level $\rho_n = \rho$ for all n, is it possible to have consistent estimators for θ and, if yes, at which rate? Second, if ρ_n is small enough the noise is not going to affect the estimation procedures evoked above; what does exactly "small enough" mean? Third, what happens in between, when ρ_n is small, but not "small enough"? Further, it is one thing to exhibit reasonable estimators, it is better to be able to study their asymptotic properties (such as their rates of convergence), it is even better to have "optimal" estimators and, for this, we need to describe the asymptotic behaviour of the relative likelihoods.

So an "optimal" program looks as follow: 1) study likelihoods (hopefully getting the LAMN property, as in the non-noisy case); 2) exhibit optimal sequences of estimators. We are unable so far to carry out this program completely. What we can do is as follows (under suitable assumptions, and with the notation $c(\theta, t, x) = a(\theta, t, x)^2$):

- Prove the LAMN (and in fact the LAN) property in the Gaussian martingale case, that is when $b_t(\omega) = 0$ and $c(\theta, t, x) = c(\theta, t)$ does not depend on x in (1.1). The corresponding Fisher information is $I(\theta) = \int_0^1 \iota(c(\theta, s), \dot{c}(\theta, s))\mathrm{d}s$ for a suitable function ι on $\mathbb{R}^2$ ($\dot{c}$ is the derivative of $\theta \rightsquigarrow c(\theta, s)$), and the rate u_n depends only on the sequence ρ_n.
- In the case above we exhibit estimators $\widehat{\theta}_n$ which are asymptotically optimal, *i.e.* $\frac{1}{u_n}(\widehat{\theta}_n - \theta)$ converges to an $\mathcal{N}(0, 1/I(\theta))$ variable if the true value of the parameter is θ.
- In the general case of (1.1), we can exhibit estimators $\widehat{\theta}_n$ such that $\frac{1}{u_n}(\widehat{\theta}_n - \theta)$ converges (with the same u_n as above) to a mixed normal variable which, conditionally on the path of X, is $\mathcal{N}(0, 1/I(\theta))$, where $I(\theta) = \int_0^1 \iota(c(\theta, s, X_s), \dot{c}(\theta, s, X_s))\mathrm{d}s$ for the same function ι as above.

This first paper is devoted to the first problem above, while the second and third problems are studied in [6]. For simplicity we restrict ourselves to the 1-dimensional case for the parameter θ (not a real restriction), and also for the process X: this could be relaxed at the cost of many more calculations.

The paper is organized as follows: in Section 2 we state the result and give a short description of the proof. In Section 3 we prove the result in the very simple case where $c(\theta, t) = c(\theta)$ depends on θ only. Section 4 is devoted to an auxiliary result which might be of independent interest and connects the LAN property for the initial experiments with the LAN property for "superexperiments" and "subexperiments". In Section 5 we give some technical results useful for this paper and/or [6]. Sections 6 and 7 are devoted to constructing the subexperiments and superexperiments respectively, and to proving the LAN property for both of them, while some results on matrices are gathered in an Appendix.

2. The result

1) As said before, this paper is concerned with the following special case of (1.1); recall that $c = a^2$:

$$X_t = \int_0^t \sqrt{c(\theta,s)}\,\mathrm{d}W_s. \tag{2.1}$$

Here, Θ is any interval (bounded or not) of $\mathbb{R}$ and c is a function: $\Theta\times[0,1]\mapsto\mathbb{R}_+$. The assumptions will be at least (H1$_\theta$) and sometimes (H2$_\theta$) or (H3$_\theta$) below, for a given θ:

Hypothesis (H1$_\theta$): The function $\zeta\mapsto c(\zeta,x)$ is twice differentiable with partial first and second derivatives denoted by $\dot c$ and $\ddot c$, and c, $\dot c$ and $\ddot c$ are continuous on $\Theta\times[0,1]$. Further the function $c(\theta,.)$ does not vanish, and the function $\dot c(\theta,.)$ is not identically 0.

Hypothesis (H2$_\theta$): The function $\dot c(\theta,.)$ does not vanish.

Hypothesis (H3$_\theta$): The set $F=\{s\in[0,1]:\dot c(s,\theta)=0\}$ is the union of its connected components with positive length, plus a Borel set with Lebesgue measure equal to 0. Moreover the function $\dot c(\theta,.)$ is Hölder-continuous with some index $\alpha\in(0,1]$.

Hypothesis (H1$_\theta$) is a standard smoothness assumption, plus some non-degeneracy and identifiability at point θ. (H2$_\theta$) and (H3$_\theta$) are additional identifiability assumptions at point θ. Condition (H3$_\theta$) rules out the case of a pathological set F, and the Hölder condition enables us to control $\dot c$ near F. (H2$_\theta$) is indeed a strong assumption, while (H3$_\theta$) is perhaps ugly-looking, but rather mild.

In this first paper we are interested in the LAN property but not in estimators, so we do not need a "global" identifiability assumption here.

Next we are given an i.i.d. sequence of $\mathcal{N}(0,1)$ variables (U_i), independent of W. Our observations at stage n consist in the finite sequence

$$X_{i/n}+\sqrt{\rho_n}U_i,\qquad i=0,\dots,n, \tag{2.2}$$

where ρ_n is a *known* positive number for each n (the observation for $i=0$ above gives no information about θ, so one could as well take $i=1,\dots,n$ in (2.2)).

At stage n the simplest statistical experiment describing the previous scheme of observations consists in taking the state space to be $\mathbb{R}^{n+1}$ with the Borel σ-field, and for each θ the probability measure P^n_θ which is the law of the sequence in (2.2) when X is given by (2.1). The measures P^n_θ are all equivalent, and we set $Z^n_{\zeta/\theta}=\mathrm{d}P^n_\zeta/\mathrm{d}P^n_\theta$ for the Radon–Nikodym derivative.

According to LeCam and Yang [11] for example, let us recall what the LAN property at point θ, with rate u_n, is: this means that for any sequence h_n of numbers going to a limit $h\in\mathbb{R}$ the sequence $Z^n_{\theta+u_nh_n/\theta}$ converges in law under P^n_θ to a limit of the form

$$\exp\left(hU\sqrt{I(\theta)}-\frac{h^2I(\theta)}{2}\right),\qquad \text{where } U\in\mathcal{N}(0,1)\quad\text{and } I(\theta)>0. \tag{2.3}$$

Then $I(\theta)$ is the (asymptotic) *Fisher information.*

The rate of convergence u_n actually depends on the behaviour of the sequence $n\rho_n$. In fact, up to taking subsequences it is no real restriction to assume that this sequence converges in $[0,\infty]$ and, ruling out the totally uninteresting case where the sequence ρ_n itself is unbounded, we see that three cases can occur:

$$\left.\begin{array}{llll} \textbf{Case 1} & n\rho_n \to u = 0: & & \text{take} \quad u_n = 1/\sqrt{n} \\ \textbf{Case 2} & n\rho_n \to u \in (0,\infty): & & \text{take} \quad u_n = 1/\sqrt{n} \\ \textbf{Case 3} & n\rho_n \to u = \infty, & \sup_n \rho_n < \infty: & \text{take} \quad u_n = (\rho_n/n)^{1/4} \end{array}\right\}. \tag{2.4}$$

Set also

$$\iota_u(x,y) = \begin{cases} \dfrac{y^2}{2x^2} & \text{if } u = 0 \\ \dfrac{y^2(2+x/u)}{2\sqrt{u}x^{3/2}(4+x/u)^{3/2}} & \text{if } 0 < u < \infty \\ \dfrac{y^2}{8x^{3/2}} & \text{if } u = \infty. \end{cases} \tag{2.5}$$

The main result of this paper is then:

Theorem 2.1. *Assume (H1$_\theta$). We have the LAN property at point θ and with the rates u_n given above and the Fisher information*

$$I(\theta) = \int_0^1 \iota_u(c(\theta,s), \dot{c}(\theta,s))\mathrm{d}s, \tag{2.6}$$

in Cases 1 and 2 and also in Case 3 if further we have either (H2$_\theta$), or (H3$_\theta$) and the sequence $n^{1-4\alpha}\rho_n$ is bounded (where α appears in (H3$_\theta$)).

The additional condition in Case 3 under (H3$_\theta$) is of course automatically satisfied if $\alpha > 1/4$. This theorem is also valid with slightly different observation schemes: namely if we observe the variables in (2.2) for $i = 1, \ldots, n$ only (perhaps a more natural setting), or if we observe the increments $X_{i/n} + \sqrt{\rho_n}\, U_i - (X_{(i-1)/n} + \sqrt{\rho_n}\, U_{i-1})$ for $i = 1, \ldots, n$.

Recall also the well known fact that, if there is no measurement error, the LAN property with rate $1/\sqrt{n}$ and asymptotic Fisher information $\int_0^1 \frac{\dot{c}(\rho,s)^2}{2c(\theta,s)^2}\mathrm{d}s$ holds; this is in accordance with the above: take $\rho_n = 0$, so we are in Case 1.

2) Main steps of the proof. Although all involved variables are Gaussian, the problem is not simple, because of the complicated dependence structure of the observations.

Of course the case where $c(\theta,s) = c(\theta)$ does not depend on time is significantly easier, and we treat this case (referred to as the "homogeneous case") first, in Section 3 below: the method consists in making a linear transformation on the observed variables (2.2) to obtain independent variables; the key observation is that, due to the homogeneous structure, the orthogonal matrices which diagonalize the covariances are in fact independent of θ: then so is the linear transformation mentionned above, and the problem reduces to the classical situation of independent (non-identically distributed) centered Gaussian variables with unknown variances.

When c depends on time, things are more complicated, and hinge upon two different ideas: the first idea is to split the sequence $1, 2, \ldots, n$ into l_n subsequences of length k_n, so that the variances of the increments $X_{i/n} - X_{(i-1)/n}$ are "almost" constant in i within any given block of length k_n when θ and ζ are close (an approximation is made here, but since c is continuous in time this approximation is good if k_n/n is small enough). Then each block can be treated as in the simple case above.

However, these blocks are *not* independent, so we need a second idea: we make the blocks independent by deleting the last observation in each of them: doing so we loose information, and get what one can call a "subexperiment". We can also make them independent by adding some observations, namely the values of $X_{i/n}$ and U_i (instead of just $X_{i/n} + \sqrt{\rho_n}\, U_i$) at the end of each block: doing so increases information, and get what one can call a "superexperiment".

The difference in information between these experiments is so small that both the sub- and superexperiments have the LAN property with the same rate and the same asymptotic Fisher information: this yields (as shown in Sect. 4) that our original experiments, which are "in between" the sub- and superexperiments, share the same property.

It should be emphasized that even under (H2$_\theta$), and even when (H2$_\theta$) is strenghtened so that the function $\dot{c}$ is uniformly away from 0, we still need to consider sub- and superexperiments.

3) Some general notation. Observe that there is no need to take $\mathbb{R}^{n+1}$ as our basic space. We stay closer to the structure of our processes by taking the state space $\Omega_n = \mathbb{R}^{2(n+1)}$, with the canonical variables $V_0, \dots, V_n, U_0, \dots, V_n$ and the Borel σ-field $\mathcal{H}_n$. Then P_θ^n is the unique probability measure under which the canonical variables are all independent, and the U_i's is $\mathcal{N}(0,1)$, and $V_0 = 0$ a.s., and the $V_i's$ for $i \geq 1$ are $\mathcal{N}(0, c_i^n(\theta))$, with

$$c_i^n(\theta) \;=\; \int_{(i-1)/n}^{i/n} c(\theta, s)\mathrm{d}s. \tag{2.7}$$

That is, the variables in (2.2) (for $i = 0, 1, \dots, n$) have the same joint law than the variables $V_0 + \ldots + V_i + \sqrt{\rho_n}\, U_i$ under P_θ^n. The σ-field corresponding to the observations (2.2) is

$$\mathcal{F}_n \;=\; \sigma(V_0 + \ldots + V_i + \sqrt{\rho_n}\, U_i : i = 0, \dots, n). \tag{2.8}$$

We also have $\mathcal{F}_n = \sigma(U_0, R_1, \dots, R_n)$, where

$$R_i = V_i + \sqrt{\rho_n}\, (U_i - U_{i-1}), \qquad i = 1, \dots, n. \tag{2.9}$$

3. The homogeneous case

In this section we wish to prove our result in the homogeneous case, because it minimizes the technicalities and clearly show why the rates are given by (2.4). In fact, since this is a particular example of the general case proved below, we feel free to slightly modify the observation scheme, in order to have even more simplicity: instead of observing the σ-field $\mathcal{F}_n$ of (2.8), we observe $\mathcal{F}'_n = \mathcal{F}_n \bigvee \sigma(U_n)$ (or equivalently, the variables in (2.2) and also the variable X_1, or the variable U_n: this will thus be a particular use of the "superexperiments" studied later).

So, in this section we suppose that $c(\zeta, s) = c(\zeta)$ for all ζ, s. We fix θ and assume (H1$_\theta$), which here implies (H2$_\theta$), and in (2.8) we have $c_i^n(\zeta) = c(\zeta)/n$. Our observation consists in the pair (U_0, U_n) whose law in the same under all P_ζ^n, and on the vector $S_n = (V_1 + \sqrt{\rho_n} U_1, R_2, \dots, R_{n-1}, V_n - \sqrt{\rho_n} U_{n-1})$ which under P_ζ^n is independent of (U_0, U_n) and is Gaussian centered with covariance matrix $C^n(\zeta)$ given by

$$C^n(\zeta)_{i,j} = \begin{cases} \dfrac{c(\zeta)}{n} + \rho_n & \text{if } i = j = 1 \text{ or } i = j = n \\ \dfrac{c(\zeta)}{n} + 2\rho_n & \text{if } 2 \leq i = j \leq n-1 \\ -\rho_n & \text{if } |i - j| = 1 \\ 0 & \text{otherwise.} \end{cases}$$

This matrix can be diagonalized by an $n \times n$ orthogonal matrix P^n which does not depend on ζ, and its eigenvalues $\lambda_u^n(\zeta)$, increasingly ordered, can be explicitely computed: all this is proved in the Appendix, once observed that with the notation (8.4) we have $C^n(\zeta) = C(a_n(\zeta), 1, \rho_n)$ with $k = n$ and $b_i = 1$ and $c_i = \frac{c(\theta)}{n}$ and $a_n(\zeta) = \frac{c(\zeta)-c(\theta)}{n}$. In particular, combining (8.2) and Lemma 8.1, we get

$$\lambda_i^n(\zeta) = \frac{c(\zeta)}{n} + 2\rho_n \left(1 - \cos\left(\frac{(i-1)\pi}{n}\right)\right).$$

Consider now the random vector $T^n = (T_i^n)_{1\le i\le n} = P^{n\star} S^n$, and set $T_i'^n = T_i^n/\sqrt{\lambda_i^n(\theta)}$: under P_ζ^n, the components $T_i'^n$ are independent with laws $\mathcal{N}(0, \lambda_i^n(\zeta)/\lambda_i^n(\theta))$ and independent of (U_0, U_n), and in particular they are $\mathcal{N}(0,1)$ under P_θ^n. Also, $\mathcal{F}'_n$ is the σ-field generated by U_0, U_n, and $T_i'^n$ for $i = 1, \ldots, n$. Then if (h_n) is a sequence of reals going to a limit h, and if we set

$$\delta_i^n = \frac{\lambda_i^n(\theta + u_n h_n)}{\lambda_i^n(\theta)} - 1 = \frac{c(\theta + u_n h_n) - c(\theta)}{c(\theta) + 2n\rho_n \left(1 - \cos\left(\frac{(i-1)\pi}{n}\right)\right)}, \tag{3.1}$$

the likelihood $Z^n_{\theta+u_nh_n/\theta} = (\mathrm{d}P^n_{\theta+u_nh_n}/\mathrm{d}P^n_\theta)_{|\mathcal{F}'_n}$ is given by:

$$\log Z^n_{\theta+u_nh_n/\theta} = -\frac{1}{2}\sum_{i=1}^n \left(\log(1+\delta_i^n) - (T_i'^n)^2 \frac{\delta_i^n}{1+\delta_i^n}\right).$$

Since the $T_i'^n$ are i.i.d. $\mathcal{N}(0,1)$ under P_θ^n, so in particular $(T_i'^n)^2$ has expectation 1 and variance 2, a well known result (see *e.g.* Th. VIII-3.32 of [8]) shows that $Z^n_{\theta+u_nh_n/\theta}$ converges in law under P_θ^n to the variable defined in (2.3) (*i.e.*, we have the desired LAN property) as soon as

$$\sup_{1\le i\le n} |\delta_i^n| \to 0, \qquad \sum_{i=1}^n (\delta_i^n)^2 \to 2h^2 I(\theta) \tag{3.2}$$

where $I(\theta)$ is given by (2.6), *i.e.* $I(\theta) = \iota_u(c(\theta), \dot c(\theta))$ here.

First, hypothesis (H1$_\theta$) and (3.1) gives $|\delta_i^n| \le C u_n$, hence giving the first part of (3.2). Second, $c(\theta + u_n h_n) - c(\theta) = u_n h_n(\dot c(\theta) + O(u_n))$, so with the notation (8.7) and using (3.1) again, we have

$$\sum_{i=1}^n (\delta_i^n)^2 = (\dot c(\theta) + O(u_n))^2 \ \frac{u_n^2 h_n^2}{\pi n \rho_n^2}\ J_2'\left(\frac{c(\theta)}{n\rho_n}, n\right). \tag{3.3}$$

Then (8.8) yields

$$\frac{u_n^2 h_n^2}{\pi n \rho_n^2}\ I_2\left(\frac{c(\theta)}{n\rho_n}\right) \le \frac{u_n^2 h_n^2}{\pi n \rho_n^2}\ J_2'\left(\frac{c(\theta)}{n\rho_n}, n\right) \le \frac{u_n^2 h_n^2}{\pi n \rho_n^2}\ I_2\left(\frac{c(\theta)}{n\rho_n}\right) + \frac{u_n^2 h_n^2}{c(\theta)^2}. \tag{3.4}$$

Since $u_n \to 0$, using (8.6) and studying separately the three cases in (2.4), we immediately deduce from $h_n \to h$ and (3.3) and (3.4) that the second part of (3.2) holds with $I(\theta) = \iota_u(c(\theta), \dot c(\theta))$.

4. Subexperiments and superexperiments

The result of this section might be of independent interest. We have a sequence of statistical experiments $(\Omega_n, \mathcal{F}_0^n, \mathcal{F}_1^n, \mathcal{F}_2^n, (P_\theta^n))$, with increasing σ-fields $\mathcal{F}_0^n \subset \mathcal{F}_1^n \subset \mathcal{F}_2^n$. Let also $Z^{n,i}_{\zeta/\theta}$ be the $\mathcal{F}_n^i$-relative likelihood, *i.e.* $Z^{n,i}_{\zeta/\theta} = (\mathrm{d}P_\zeta^n/\mathrm{d}P_\theta^n)_{|\mathcal{F}_i^n}$.

Theorem 4.1. *Suppose that $Z^{n,i}_{\zeta_n/\theta}$ converges in law under P^n_θ to a limit Y with $0 < Y < \infty$ and $E(Y) = 1$, for both $i = 0$ and $i = 2$ (the sequence of parameters ζ_n is completely arbitrary). Then the same convergence holds for $i = 1$.*

Proof. Set $Z^n_i = Z^{n,i}_{\zeta_n/\theta}$. Since $E^n_\theta(Z^n_i) = 1$, the sequence (Z^n_0, Z^n_1, Z^n_2) is uniformly tight. Up to taking subsequences, we may assume that (Z^n_0, Z^n_1, Z^n_2) converges in law under P^n_θ to a limit denoted by (Z_0, Z_1, Z_2). Our assumption yields $\mathcal{L}(Z_0) = \mathcal{L}(Z_2) = \mathcal{L}(Y)$.

Set $\mathcal{F}_0 = \sigma(Z_0)$ and $\mathcal{F}_1 = \sigma(Z_0, Z_1)$. Let ψ_p be continuous functions on $\mathbb{R}$ with $0 \le \psi_p \le 1_{[-p,p]}$ and $\psi_p(x) \uparrow 1$. Finally let ϕ be a bounded continuous function on $\mathbb{R}^2$. For $i = 0, 1, 2$ we have

$$E^n_\theta(\phi(Z^n_0, Z^n_1)\psi_p(Z^n_i)Z^n_i) \ \to_{(n)} \ E(\phi(Z_0, Z_1)\psi_p(Z_i)Z_i) \ \to_{(p)} \ E(\phi(Z_0, Z_1)Z_i),$$

$$\begin{aligned} E^n_\theta(\phi(Z^n_0, Z^n_1)(1 - \psi_p(Z^n_i))Z^n_i) \ &\le \ CE^n_\theta(Z^n_i - \psi_p(Z^n_i)Z^n_i) \ = \ C(1 - E^n_\theta(\psi_p(Z^n_i)Z^n_i)) \\ &\to_{(n)} \ C(1 - E(\psi_p(Z_i)Z_i)) \ \to_{(p)} \ 0, \end{aligned}$$

hence $E^n_\theta(\phi(Z^n_0, Z^n_1)Z^n_i) \to E(\phi(Z_0, Z_1)Z_i)$. Now $E^n_\theta(\phi(Z^n_0, Z^n_1)Z^n_1) = E^n_\theta(\phi(Z^n_0, Z^n_1)Z^n_2)$ for each n, so $E(\phi(Z_0, Z_1)Z_1) = E(\phi(Z_0, Z_1)Z_2)$, which in turn yields $Z_1 = E(Z_2|\mathcal{F}_1)$. One shows in a similar way that $Z_0 = E(Z_1|\mathcal{F}_0)$.

Therefore $\sqrt{Z_0} \le E(\sqrt{Z_1}|\mathcal{F}_0) \le E(\sqrt{Z_2}|\mathcal{F}_0)$. Since Z_0 and Z_2 have the same law, $E(\sqrt{Z_0}) = E(\sqrt{Z_2})$ and thus $Z_0 = Z_2$ a.s. But $Z_1 = E(Z_2|\mathcal{F}_1)$ a.s., hence $\mathcal{L}(Z_1) = \mathcal{L}(Y)$. □

5. Change of probability

Our aim is to prove Theorem 2.1, so we always assume (H1$_\theta$) at least. That is, we fix θ and a sequence h_n going to a limit $h \in \mathbb{R}$; if Z_n is the relative likelihood $\mathrm{d}P^n_{\theta+u_nh_n}/\mathrm{d}P^n_\theta\big|_{\mathcal{F}_n}$ (in restriction to $\mathcal{F}_n$), we need to prove that Z_n converges in law under P^n_θ to (2.3). By virtue of (H1$_\theta$) there exists an $\varepsilon > 0$ such that

$$\frac{\varepsilon}{n} \le c^n_i(\theta) \le \frac{1}{n\varepsilon}, \qquad \frac{\varepsilon}{n} \le c^n_i(\theta + u_nh_n) \le \frac{1}{n\varepsilon}, \qquad n \ge 1,\ 1 \le i \le n. \tag{5.1}$$

As said before we wish to construct blocks such that within each of them one can consider $c^n_i(\theta)$ (resp. $c^n_i(\theta + u_nh_n)$) in (2.7) as "almost" independent of i. For this, we first change $c^n_i(\theta + u_nh_n)$ in such a way that the asymptotic behaviour of the likelihood is not modified, and we will use the following well known result (see *e.g.* [4]): let $P'n$ be an arbitrary probability measure on $(\Omega_n, \mathcal{H}_n)$ which is absolutely continuous w.r.t. P^n_θ; let also $Z'_n = \mathrm{d}P'^n/\mathrm{d}P^n_\theta|_{\mathcal{F}_n}$ in restriction to $\mathcal{F}_n$.

Lemma 5.1. *Assume that the variation distance (on $\mathcal{H}_n$) $||P'^n - P^n_{\theta+h_nu_n}||_v$ goes to 0. If the sequence Z'_n converges in law under P^n_θ to a limit Y satisfying $0 < Y < \infty$ and $E(Y) = 1$, then the sequence Z_n converges in law under P^n_θ to the same limit Y.*

We will apply this lemma to the measure P'^n under which the variables U_i and V_i are all independent, with U_i being $\mathcal{N}(0, 1)$ and $V_0 = 0$ a.s. and V_i for $i = 1, \ldots, n$ being $\mathcal{N}(0, c'^n_i)$, for numbers $c'^n_i > 0$ to be constructed later.

Lemma 5.2. *If*

$$n^2 \sum_{i=1}^{n} |c'^n_i - c^n_i(\theta + u_nh_n)|^2 \to 0, \tag{5.2}$$

then $||P'^n - P^n_{\theta+u_nh_n}||_v \to 0$.

Proof. Let $\bar{Z}_n = \mathrm{d}P^n_{\theta+u_nh_n}/\mathrm{d}P'^n$ on the σ-field $\mathcal{H}_n$. Setting $c''^n_i = c^n_i(\theta + u_nh_n)$, we have

$$\bar{Z}_n \;=\; \prod_{i=1}^n \left(\sqrt{\frac{c'^n_i}{c''^n_i}}\,\exp -\frac{V_i^2}{2}\left(\frac{1}{c''^n_i} - \frac{1}{c'^n_i}\right)\right).$$

Hence if U is an $\mathcal{N}(0,1)$ variable, and since $E(\mathrm{e}^{aU^2}) = 1/\sqrt{1-2a}$ for $a < 1/2$, we get

$$\begin{aligned} E'^n\left(\sqrt{\bar{Z}_n}\right) \;&=\; \textstyle\prod_{i=1}^n \left(\left(\frac{c'^n_i}{c''^n_i}\right)^{1/4}\; E\left(\exp -\frac{U^2}{4}\left(\frac{c'^n_i}{c''^n_i} - 1\right)\right)\right)\\ &=\; \textstyle\prod_{i=1}^n \left(\left(\frac{c'^n_i}{c''^n_i}\right)^{1/4}\left(1 + \frac{1}{2}\left(\frac{c'^n_i}{c''^n_i} - 1\right)\right)^{-1/2}\right)\\ &=\; \textstyle\prod_{i=1}^n \left((1+\delta^n_i)^{1/4}\left(1 + \frac{1}{2}\delta^n_i\right)^{-1/2}\right), \end{aligned} \tag{5.3}$$

where $\delta^n_i = \frac{c'^n_i - c''^n_i}{c''^n_i}$.

The variation distance and the Hellinger distance induce the same topology (see *e.g.* [8], Chap. V), hence the result will be proved if we show that $E'^n((1-\sqrt{\bar{Z}_n})^2) \to 0$, which amounts to $E'^n(\sqrt{\bar{Z}_n}) \to 1$ (because $E'^n(\bar{Z}_n) = 1$). This convergence holds as soon as $\sum_{i=1}^n (\delta^n_i)^2 \to 0$, and this is implied by (5.2) because of (5.1). □

Now we choose c'^n_i. Recall once more that θ and the sequence h_n are fixed, while u_n is given by (2.4). We set also

$$c^n_i = c^n_i(\theta), \qquad \dot{c}^n_i = \dot{c}^n_i(\theta).$$

Then we choose a subdivision $0 = s^n_1 < s^n_2 < \ldots < s^n_{l_n+1} = 1$ in such a way that

$$\Delta_n := \sup_{1\le m\le l_n} (s^n_{m+1} - s^n_m) \;\to\; 0, \qquad \inf_{1\le m\le l_n} (s^n_{m+1} - s^n_m) \;\ge\; \frac{1}{n}\cdot \tag{5.4}$$

The choice of l_n and of the s^n_m's will be made later. Then we set

$$I_{n,m} = (s^n_m, s^n_{m+1}], \quad J_{n,m} = \left\{i : \frac{i}{n} \in I_{n,m}\right\}, \quad k_{n,m} = \#J_{n,m}, \quad i_{n,m} = \inf J_{n,m}, \tag{5.5}$$

$$\left.\begin{array}{ll} c^{n,m}_{\min} = \inf_{s\in I_{n,m}} c(\theta,s), & c^{n,m}_{\max} = \sup_{s\in I_{n,m}} c(\theta,s),\\ |\dot{c}|^{n,m}_{\min} = \inf_{s\in I_{n,m}} |\dot{c}(\theta,s)|, & |\dot{c}|^{n,m}_{\max} = \sup_{s\in I_{n,m}} |\dot{c}(\theta,s)| \end{array}\right\}.$$

Observe that $i \in J_{n,m} \;\Rightarrow\; |\dot{c}^n_i| \le |\dot{c}|^{n,m}_{\max}/n$. We denote by $s_{n,m}$ any point achieving the maximum of $s \rightsquigarrow |\dot{c}(s,\theta)|$ on the closure of $I_{n,m}$, and we set $\dot{c}_{n,m} = \dot{c}(\theta, s_{n,m})/n$, and $\ddot{c}_{n,m} = \ddot{c}^n_{i_{n,m}}(\theta)$, and

$$i \in J_{n,m} \quad\Rightarrow\quad c'^n_i = \begin{cases} c^n_i + u_nh_n\dot{c}^n_i + \dfrac{u_n^2h_n^2}{2}\,\ddot{c}_{n,m}\,\dfrac{\dot{c}^n_i}{\dot{c}_{n,m}} & \text{if } \dot{c}_{n,m} \ne 0,\\ c^n_i + \dfrac{u_n^2h_n^2}{2}\,\ddot{c}_{n,m} & \text{if } \dot{c}_{n,m} = 0. \end{cases} \tag{5.6}$$

Recall (5.1) and also that $|\dot{c}^n_i| \le C/n$ and $|\ddot{c}_{n,m}| \le C/n$ for some constant C (below, C denotes a constant which may change from line to line and depend on the function c, but not on n and i). Hence, since $u_nh_n \to 0$, for all n large enough $c'^n_i > 0$ for all i.

Observe that (5.6) is more complicated than just considering the function c as constant on each interval $I(n,m)$. If we had chosen $c_i'^n$ independent of i within each interval $I(n,m)$, then getting (5.2) would necessitate restrictive conditions on the regularity of $\dot c$ and on the rate $\Delta_n \to 0$, implying in turn that the LAN property cannot be proved for all bounded sequence ρ_n with this "simple" choice of $c_i'^n$. Our choice (5.6) is a modification of the Taylor expansion of $c_i^n(\theta+u_nh_n)$ chosen such that results of the Appendix apply to the covariance matrix of the observation under P'^n.

Lemma 5.3. *The above-defined numbers $c_i'^n$ satisfy (5.2).*

Proof. Taylor's formula yields for $i \in J_{n,m}$:

$$c_i^n(\theta+u_nh_n) - c_i'^n = \frac{u_n^2h_n^2}{2}\left(\ddot c_i^n(\theta_i^n) - \ddot c_{n,m} + \ddot c_{n,m}\left(1 - \frac{\dot c_i^n}{\dot c_{n,m}}\right)\right),$$

where θ_i^n is in between θ and $\theta+u_nh_n$, and with the convention $0/0=1$ (recall that if $\dot c_{n,m}=0$ we also have $\dot c_i^n(\theta)=0$). In Cases 1 or 2, we have $|c_i^n(\theta+u_nh_n) - c_i'^n| \le Cn^{-2}$ and (5.2) follows.

Case 3 needs more attention. By (H1$_\theta$) and (5.4) there is a sequence $\varepsilon_n \to 0$ having

$$s_m^n \le s \le t \le s_{m+1}^n,\ |\zeta-\theta| \le u_n|h_n| \quad\Rightarrow\quad |\ddot c(\zeta,t) - \ddot c(\theta,s)| + |\dot c(\theta,t) - \dot c(\theta,s)| \le \varepsilon_n.$$

Then

$$i \in J_{n,m} \qquad\Rightarrow\qquad |\ddot c_i^n(\theta_i^n) - \ddot c_{n,m}| + |\dot c_i^n - \dot c_{n,m}| \le \frac{\varepsilon_n}{n}\cdot$$

Since $|\ddot c_{n,m}| \le C/n$ and $|\dot c_i^n| \le |\dot c_{n,m}|$ and $u_n^4 \le C/n$, we deduce that

$$n^2|c_i^n(\theta+u_nh_n) - c_i'^n|^2 \le \begin{cases} \frac{C}{n}\varepsilon_n^2 & \text{if } \dot c_{n,m}=0 \\ \frac{C}{n} & \text{if } 0<|\dot c_{n,m}|<\sqrt{\varepsilon_n}/n \\ \frac{C}{n}\varepsilon_n & \text{if } |\dot c_{n,m}| \ge \sqrt{\varepsilon_n}/n. \end{cases}$$

Hence, in order to obtain (5.2) it suffices to prove that $\frac{1}{n}\sum_{m\in K_n} k_{n,m} \to 0$, where $K_n = \{m : 0 < |\dot c_{n,m}| < \sqrt{\varepsilon_n}/n\}$. This is immediate under (H2$_\theta$).

Now we complement the notation of (H3$_\theta$). The open set F^c is the finite or countable union of its connected components (a_j,b_j), the number of which being denoted by $M \le \infty$. The closed set F is the finite or countable union of its connected components $[a_j',b_j']$ (with $a_j' < b_j'$, the number of which is denoted by $M' \le \infty$), plus possibly a set F' having $\lambda(F')=0$, where λ is the Lebesgue measure.

Observe that by (5.4) we have $k_{n,m} \le 2n\lambda(I_{n,m})$. For all i and $\eta>0$, for all n large enough we have $|\dot c(\theta,s)| \ge \sqrt{\varepsilon_n}$ when $s \in [a_i+\eta, b_i-\eta]$ (recall that $\varepsilon_n \to 0$); hence for all n large enough $|\dot c_{n,m}| \ge \sqrt{\varepsilon_n}/n$ for any interval $I_{n,m} \subset]a_i+\eta, b_i-\eta[$. We also have $\dot c(s,\theta)=0$ if $s \in [a_i',b_i']$, hence $\dot c_{n,m}=0$ as soon as $I_{n,m} \subset [a_i',b_i']$. Therefore, since each interval $I_{n,m}$ has length smaller than Δ_n, we get for any integer N, any $\eta>0$ and any n bigger than some $n_{N,\eta}$:

$$\frac{1}{n}\sum_{m\in K_n} k_{n,m} \le 2\left(1 - \sum_{i=1}^{N\wedge M}(b_i - a_i - 2\Delta_n - 2\eta) - \sum_{i=1}^{N\wedge M'}(b_i' - a_i' - 2\Delta_n)\right)\cdot$$

Since $\Delta_n \to 0$ we obtain

$$\limsup_n \frac{1}{n}\sum_{m\in K_n} k_{n,m} \le 2\left(1 - \sum_{i=1}^{N\wedge M}(b_i - a_i - 2\eta) - \sum_{i=1}^{N\wedge M'}(b_i' - a_i')\right).$$

The union of all (a_j, b_j) and $[a'_i, b'_i]$ being F'^c, and $\lambda(F'^c) = 1$, and $\eta > 0$ is arbitrary: henceforth $\frac{1}{n} \sum_{m \in K_n} k_{n,m} \to 0$. □

Let us end this section by describing how to gather all previous results for proving Theorem 2.1. We first fix θ. Then for each n we construct two σ-fields $\mathcal{F}_n^-$ and $\mathcal{F}_n^+$ with $\mathcal{F}_n^- \subset \mathcal{F}_n^+ \subset \mathcal{H}_n$ (corresponding to the sub- and superexperiment) and subdivisions $(s_m^{n,-})$ and $(s_m^{n,+})$ satisfying (5.4); then with each sequence h_n converging to a limit h we associate $c_i'^{n-}$ (resp. $c_i'^{n+}$) by (5.6) and the corresponding measures P'^{n-} (resp. P'^{n+}) described above, and finally the relative likelihoods $Z_n^- = \mathrm{d}P'^{n-}/\mathrm{d}P_\theta^n|_{\mathcal{F}_n^-}$ (resp. $Z_n^+ = \mathrm{d}P'^{n+}/\mathrm{d}P_\theta^n|_{\mathcal{F}_n^+}$).

Putting together Theorem 4.1 and Lemmas 5.1, 5.2 and 5.3, we readily obtain the following corollary:

Corollary 5.4. *Suppose that both Z_n^- and Z_n^+ converge in law under P_θ^n to the same limit given by (2.3) with $I(\theta)$ given by (2.6), for any choice of sequences h_n converging to h. Then the experiments $(\Omega_n, \mathcal{F}_n, (P_\zeta^n))$ have the LAN property at point θ with Fisher information $I(\theta)$, for any choice of σ-fields $\mathcal{F}_n$ such that $\mathcal{F}_n^- \subset \mathcal{F}_n \subset \mathcal{F}_n^+$.*

Therefore for proving Theorem 2.1 it is enough to exhibit the σ-fields $\mathcal{F}_n^-$ and $\mathcal{F}_n^+$ with $\mathcal{F}_n^- \subset \mathcal{F}_n \subset \mathcal{F}_n^+$ for every observed σ-field $\mathcal{F}_n$ of interest (and in particular the one given by (2.8)), and such that the convergence assumptions of the previous corollary hold true.

6. LAN property for subexperiments

First we need to construct the subexperiments and the subdivisions satisfying (5.4). We choose $\gamma \in (\frac{1}{2}, 1)$ and set

$$k_n = [n^\gamma], \qquad l_n = \left[\frac{n}{k_n}\right].$$

The following is obvious:

$$k_n \to \infty, \qquad l_n \to \infty, \qquad u_n^2 l_n \to 0. \tag{6.1}$$

Then we set $s_m^n = mk_n/n$ if $0 \le m \le l_n$ and $s_{l_n+1}^n = 1$. Clearly (5.4) holds, and we use all notation of Section 5, in particular $c_i'^n$ given by (5.6) and P'^n as defined before Lemma 5.2. Observe that $k_{n,m} = k_n$ if $1 \le m \le l_n - 1$.

Next, denote by M_n the set of all indices m in $\{1, \dots, l_n - 1\}$ such that $s \rightsquigarrow \dot{c}(\theta, s)$ does not vanish on the interval $I_{n,m}$. Then set (recall (2.9) for R_i):

$$\mathcal{F}_n^- = \sigma(R_i : k_n(m-1) + 1 \le i \le k_n m - 1, m \in M_n). \tag{6.2}$$

Our aim is to prove the following, where $Z_n^- = \mathrm{d}P'^n/\mathrm{d}P_\theta^n|_{\mathcal{F}_n^-}$:

Proposition 6.1. *Under (H1$_\theta$), and with the previous notation, the sequence Z_n^- converges in law under P_θ^n to the limit described in (2.3), with u_n given by (2.4) and $I(\theta)$ given by (2.6).*

Proof. 1) The observations corresponding to the σ-field $\mathcal{F}_n^-$ are naturally divided into $\#M_n$ blocks: if $m \in M_n$, let $R^{n,m}$ denote the column vector whose components are $R_i^{n,m} = R_{(m-1)k_n+i}$ for $i = 1, \dots, k_n - 1$. By construction, the vectors $R^{n,m}$ for $m \in M_n$ are centered Gaussian *and independent* under both P_θ^n and P'^n.

Let us fix $m \in M_n$. The covariance matrix of $R^{n,m}$ under P_θ^n (resp. P'^n) is denoted by $C^{n,m}$ (resp. $C'^{n,m}$): observe that all $\dot{c}_i^n$ for $i \in J_{n,m}$ have the same sign, say α_n ($= +1$ or $= -1$); so with the notation (8.4) we have $C^{n,m} = C(0, 2, \rho_n)$ and $C'^{n,m} = C(v_{n,m}, 2, \rho_n)$, provided $k = k_{n,m} - 1$ and $c_i = c_{(m-1)k_n+i}^n$ and $b_i = b_i^{n,m} = |\dot{c}_{(m-1)k_n+i}^n|$ and $v_{n,m} = \alpha_n u_n h_n \left(1 + \frac{u_n h_n \ddot{c}_{n,m}}{2\dot{c}_{n,m}}\right)$ (recall (5.6)).

Next, consider the vectors $S^{n,m}$ with components $S_i^{n,m} = R_i^{n,m}/\sqrt{b_i^{n,m}}$: under P_θ^n (resp. P'^n) it is centered Gaussian with covariance $\widehat{C}^{n,m}$ (resp. $\widehat{C}'^{n,m}$) associated with $C^{n,m}$ and $C'^{n,m}$ as in Lemma 8.1. Write $\widehat{\lambda}_i^{n,m}$ and $\widehat{\lambda}_i'^{n,m}$, $i = 1, \dots, k_n - 1$, for the increasingly ordered eigenvalues of these matrices. From Lemma 8.1 we know that the orthogonal matrix which diagonalizes $\widehat{C}^{n,m}$ and $\widehat{C}'^{n,m}$ is the same, say $P^{n,m}$, and also that (use notation (8.2)):

$$\left.\begin{array}{c}\widehat{\lambda}_i'^{n,m} = \widehat{\lambda}_i^{n,m} + v_{n,m}, \\ \frac{c_{\min}^{n,m}}{|\dot{c}|_{\max}^{n,m}}\left(1+\frac{n\lambda_i(2,\rho_n)}{c_{\max}^{n,m}}\right) \le \widehat{\lambda}_i^{n,m} \le \frac{c_{\max}^{n,m}}{|\dot{c}|_{\min}^{n,m}}\left(1+\frac{n\lambda_i(2,\rho_n)}{c_{\min}^{n,m}}\right)\end{array}\right\}. \tag{6.3}$$

Hence the vector $Y^{n,m} = (P^{n,m})^\star S^{n,m}$ has components $Y_i^{n,m}$ which are independent normal centered with variances $\widehat{\lambda}_i^{n,m}$ and $\widehat{\lambda}_i'^{n,m}$ under P_θ^n and P'^n respectively. Then the variables $Y_i'^{n,m} = Y_i^{n,m}/\sqrt{\widehat{\lambda}_i^{n,m}}$ are i.i.d. $\mathcal{N}(0,1)$ under P_θ^n, and independent $\mathcal{N}(0,\widehat{\lambda}_i'^{n,m}/\widehat{\lambda}_i^{n,m})$ under P'^n (when m and i vary); furthermore these variables generate the σ-field $\mathcal{F}_n^-$: so exactly as in Section 3, and if we set

$$\delta_i^{n,m} \;=\; \frac{\widehat{\lambda}_i'^{n,m}-\widehat{\lambda}_i^{n,m}}{\widehat{\lambda}_i^{n,m}} = \frac{v_{n,m}}{\widehat{\lambda}_i^{n,m}}, \tag{6.4}$$

we have

$$\log Z_n^- \;=\; -\frac{1}{2}\sum_{m\in M_n}\sum_{i=1}^{k_{n,m}-1}\left(\log(1+\delta_i^{n,m}) - (Y_i'^{n,m})^2\frac{\delta_i^{n,m}}{1+\delta_i^{n,m}}\right)\cdot \tag{6.5}$$

Therefore, as in Section 3 again, for obtaining the desired claim it is enough to prove the following two conditions:

$$\sup_{1\le i\le k_{n,m}-1,\ m\in M_n} |\delta_i^{n,m}| \;\to\; 0, \tag{6.6}$$

$$F_n \;:=\; \sum_{m\in M_n}\widehat{F}_{n,m} \;\to\; 2I(\theta)h^2, \qquad \text{where } \widehat{F}_{n,m} = \sum_{i=1}^{k_{n,m}-1}|\delta_i^{n,m}|^2. \tag{6.7}$$

2) First, $|v_{n,m}| \le Cu_n(1+u_n/|\dot{c}|_{\min}^{n,m})$, and (6.3) and (H1$_\theta$) yield $0 < 1/\widehat{\lambda}_i^{n,m} \le C|\dot{c}|_{\max}^{n,m} \le C$. Hence (6.4) yields $|\delta_i^{n,m}| \le Cu_n$, and we have (6.6).

Next, with the notation $W_n = \sup_{1\le m\le l_n}(c_{\max}^{n,m}/c_{\min}^{n,m})^2$ and $\phi(x,a) = a + 2(1-\cos x)$ we have by (8.2):

$$\frac{v_{n,m}^2\ (|\dot{c}|_{\min}^{n,m})^2}{n^2\rho_n^2 W_n\phi\left(\frac{i\pi}{k_{n,m}},\frac{c_{\min}^{n,m}}{n\rho_n}\right)^2} \le |\delta_i^{n,m}|^2 \le \frac{v_{n,m}^2\ (|\dot{c}|_{\max}^{n,m})^2 W_n}{n^2\rho_n^2\phi\left(\frac{i\pi}{k_{n,m}},\frac{c_{\max}^{n,m}}{n\rho_n}\right)^2}.$$

Taking into account the value of $v_{n,m}$, hypothesis (H1$_\theta$) and the fact that $|\dot{c}|_{\min}^{n,m} \le |\dot{c}|_{\max}^{n,m} = n|\dot{c}_{n,m}|$, we deduce that $|v_{n,m}^2 - u_n^2h_n^2| \le Cu_n^2h_n^2\left(\frac{u_n}{|\dot{c}|_{\max}^{n,m}} + \frac{u_n^2}{(|\dot{c}|_{\max}^{n,m})^2}\right)$, hence

$$\begin{aligned}&\frac{u_n^2h_n^2}{n^2\rho_n^2W_n\phi\left(\frac{i\pi}{k_{n,m}},\frac{c_{\min}^{n,m}}{n\rho_n}\right)^2}\left((|\dot{c}|_{\min}^{n,m})^2 - Cu_n|\dot{c}|_{\min}^{n,m} - Cu_n^2\right)\\ &\qquad\le\; |\delta_i^{n,m}|^2 \;\le\; \frac{u_n^2h_n^2W_n}{n^2\rho_n^2\phi\left(\frac{i\pi}{k_{n,m}},\frac{c_{\max}^{n,m}}{n\rho_n}\right)^2}\left((|\dot{c}|_{\max}^{n,m})^2 + Cu_n|\dot{c}|_{\max}^{n,m} + Cu_n^2\right)\cdot\end{aligned} \tag{6.8}$$

3) Now we use the notation I_2 and J_2 of (8.6) and (8.7). From (6.7) and (6.8) we deduce

$$\begin{aligned}&\frac{u_n^2h_n^2k_{n,m}}{n^2\rho_n^2W_n\pi}J_2\left(\frac{c_{\min}^{n,m}}{n\rho_n},k_{n,m}\right)\left((|\dot c|_{\min}^{n,m})^2-Cu_n|\dot c|_{\min}^{n,m}-Cu_n^2\right)^+\\&\qquad\le\ \widehat F_{n,m}\ \le\ \frac{u_n^2h_n^2k_{n,m}W_n}{n^2\rho_n^2\pi}J_2\left(\frac{c_{\max}^{n,m}}{n\rho_n},k_{n,m}\right)\left((|\dot c|_{\max}^{n,m})^2+Cu_n|\dot c|_{\max}^{n,m}+Cu_n^2\right).\end{aligned}\tag{6.9}$$

Thus if we set $w_n(y)=\frac{u_n^2}{n\rho_n^2\pi}I_2(y/n\rho_n)$, we deduce from (8.8) and (H1$_\theta$) that

$$\begin{aligned}&\frac{h_n^2}{W_n}\left(\frac{k_{n,m}}{n}w_n(c_{\min}^{n,m})-Cu_n^2\right)^+\left((|\dot c|_{\min}^{n,m})^2-Cu_n|\dot c|_{\min}^{n,m}-Cu_n^2\right)^+\\&\qquad\le\ \widehat F_{n,m}\ \le\ h_n^2W_n\frac{k_{n,m}}{n}w_n(c_{\max}^{n,m})\left((|\dot c|_{\max}^{n,m})^2+Cu_n|\dot c|_{\max}^{n,m}+Cu_n^2\right).\end{aligned}\tag{6.10}$$

Notice that if we set $\widehat F_{n,m}=0$ for $m\in\{1,\dots,l_n-1\}\backslash M_n$, and since then we also have $|\dot c|_{\min}^{n,m}=0$, the estimates (6.10) are also valid in this case. Furthermore, with this convention, we have $F_n=\sum_{m=1}^{l_n-1}\widehat F_{n,m}$.

4) Using (8.6), one checks that the sequence w_n converges to the following function w, uniformly on the interval $[\varepsilon,1/\varepsilon]$:

$$w(y)\ =\ \begin{cases}\dfrac{1}{y^2} & \text{in Case 1}\\ \dfrac{2+y/u}{\sqrt{u}y^{3/2}(4+y/u)^{3/2}} & \text{in Case 2}\\ \dfrac{1}{4y^{3/2}} & \text{in Case 3.}\end{cases}\tag{6.11}$$

Let r_m^n and $r_m'^n$ be points in the closed interval $\bar I_{n,m}$ such that $c(\theta,r_m^n)=c_{\max}^{n,m}$ and $c(\theta,r_m'^n)=c_{\min}^{n,m}$. By (H1$_\theta$) there exists $\zeta_n\to0$ such that $u_n\le\zeta_n$ and

$$\left.\begin{array}{ll}1-\zeta_n\le\dfrac{1}{W_n}\le W_n\le1+\zeta_n, & y\in\left[\varepsilon,\dfrac{1}{\varepsilon}\right]\ \Rightarrow\ |w_n(y)-w(y)|\le\zeta_n,\\ ||\dot c(\theta,r_m^n)|-|\dot c|_{\max}^{m,n}|\le\zeta_n, & ||\dot c(\theta,r_m'^n)|-|\dot c|_{\min}^{m,n}|\le\zeta_n\end{array}\right\}.\tag{6.12}$$

Hence by (6.10) and $\varepsilon\le c(\theta,s)\le\frac{1}{\varepsilon}$ and the fact that $w_n(y)$ for $y\in[\varepsilon,1/\varepsilon]$ and $|\dot c|_{\max}^{n,m}$ and u_n are uniformly bounded, and also because $k_{n,m}=k_n$ for $m\le l_n-1$:

$$\begin{aligned}&\frac{(1-\zeta_n)h_n^2}{l_n-1}\sum_{m=1}^{l_n-1}\frac{(l_n-1)k_n}{n}\ w(c(\theta,r_m'^n))\dot c(\theta,r_m'^n)^2-C\zeta_n-Cu_n^2l_n\\&\qquad\le\ F_n\ \le\ \frac{(1+\zeta_n)h_n^2}{l_n-1}\sum_{m=1}^{l_n-1}\frac{(l_n-1)k_n}{n}\ w(c(\theta,r_m^n))\dot c(\theta,r_m^n)^2+C\zeta_n.\end{aligned}\tag{6.13}$$

Observe that $\frac{(l_n-1)k_n}{n}\to1$, while $u_n^2l_n\to0$ by (6.1). Since the sums in (6.13) are Riemann sums, using the uniform convergence of w_n to w on $[\varepsilon,1/\varepsilon]$ and recalling (5.1), we deduce (6.7) and our proof is finished. □

7. LAN property for superexperiments

First we need to construct the superexperiments and the subdivisions satisfying (5.4). Here again we choose $\gamma \in (\frac{1}{2}, 1)$ and set

$$k_n = [n^\gamma], \qquad l'_n = \left[\frac{n}{k_n}\right], \qquad k'_n = n - (l'_n - 1)k_n. \tag{7.1}$$

Set $t^n_m = k_n m/n$ if $0 \le m \le l'_n - 1$, and $t^n_{l'_n} = 1$, and $I'_{n,m} = (t^n_{m-1}, t^n_m]$ for $1 \le m \le l'_n$. Set also $J'_{n,m} = \{i : i/n \in I'_{n,m}\}$.

Denote by K'_n the set of all indices m such that either $s \rightsquigarrow \dot{c}(\theta, s)$ does not vanish on $I'_{n,m}$, or $\dot{c}(\theta, s) = 0$ for all $s \in I'_{n,m}$, and $K''_n = \{1, \ldots, l'_n\} \backslash K'_n$. If $m \in K''_n$, we divide $J'_{n,m}$ into sub-intervals of integers which are of maximal length and such that either $\dot{c}^n_i > 0$ or $\dot{c}^n_i < 0$ for all i in any of these sub-intervals, plus all the "intervals" $\{i\}$ such that $\dot{c}^n_i = 0$. We have thus divided $\{1, \ldots, n\}$ into l_n intervals (the $J'_{n,m}$'s for $m \in K'_n$, and the sub-intervals of the $J'_{n,m}$'s for $m \in K''_n$): these intervals, ordered according to the natural order, are denoted $J_{n,m}$, for $m = 1, \ldots, l_n$. Let $k_{n,m} = \# J_{n,m}$, and call $i_{n,m}$ and $j_{n,m}$ the smallest and the biggest points in $J_{n,m}$, and by convention $i_{n,0} = j_{n,0} = 0$: we have $i_{n,m} \le j_{n,m}$ and $i_{n,1} = 1$ and $i_{n,m+1} = j_{n,m} + 1$ and $j_{n,l_n} = n$. It remains to set $s^n_m = \frac{j_{n,m-1}}{n}$ for $m = 1, \ldots, l_n$, and $I_{n,m} = (s^n_m, s^n_{m+1}]$: we have (5.4) and (5.5). Observe also that under (H2$_\theta$) we have $K'_n = \{1, \ldots, l'_n\}$ and $l_n = l'_n$ and $I_{n,m} = I'_{n,m}$.

Finally, we set

$$\mathcal{F}^+_n = \mathcal{F}_n \bigvee \sigma(U_{j_{n,m}} : m = 0, \ldots, l_n). \tag{7.2}$$

Our aim is to prove the following, where $Z^+_n = \mathrm{d}P'^n / \mathrm{d}P^n_\theta |_{\mathcal{F}^+_n}$:

Proposition 7.1. *Assume (H1$_\theta$) and, in Case 3, either (H2$_\theta$) or (H3$_\theta$) plus the boundedness of the sequence $\rho_n n^{1-4\alpha}$. Then with the previous notation, the sequence Z^+_n converges in law under P^n_θ to the limit described in (2.3), with u_n given by (2.4) and $I(\theta)$ given by (2.6).*

Then Propositions 6.1 and 7.1, together with Corollary 5.4, will end the proof of Theorem 2.1, once noticed that $\mathcal{F}^-_n \subset \mathcal{F}_n \subset \mathcal{F}^+_n$.

Proof. 1) Set

$$R'_i = \begin{cases} V_i + \sqrt{\rho_n}\, U_i & \text{if } i = i_{n,m} < j_{n,m} \text{ for some } m \\ V_i - \sqrt{\rho_n}\, U_{i-1} & \text{if } i = j_{n,m} > i_{n,m} \text{ for some } m \\ V_i & \text{if } i = i_{n,m} = j_{n,m} \text{ for some } m \\ R_i & \text{otherwise.} \end{cases}$$

Then, comparing with (2.8) and (7.2), we obtain

$$\mathcal{F}^+_n = \sigma(U_{j_{n,m}} : m = 0, \ldots, l_n; R'_i, i = 1, \ldots, n)\cdot$$

Denote by K_n the set of indices m such that $J_{n,m}$ is one of the initial sets $J'_{n,m'}$ on which $\dot{c}(\theta, .)$ does not vanish, and by H_n the set of indices m such that $J_{n,m}$ is one of the initial sets $J'_{n,m'}$ on which $\dot{c}(\theta, .)$ is identically 0, and L_n the set of indices m such that $i_{n,m} = j_{n,m}$, and finally $M_n = \{1, \ldots, l_n\} \backslash (L_n \cup K_n \cup H_n)$. Under both

P^n_θ and P'^n, the observations are divided into independent blocks, namely

$$\left.\begin{array}{l} \text{1. the } U_{j_{n,m}} \text{ for } m=0,\dots,l_n \text{ (they are } \mathcal{N}(0,1) \text{ under both probabilities),} \\ \text{2. the } R'^{n,m} = V_{i_{n,m}} \text{ for } m\in L_n, \\ \text{3. the column vectors } R'^{n,m} \text{ whose components are } R'^{n,m}_i = R'_{i_{n,m}+i-1} \text{ for} \\ i=1,\dots,k_{n,m} \text{ for } m\in K_n\cup H_n\cup M_n \end{array}\right\}\cdot \tag{7.3}$$

Note that under (H2$_\theta$) we have $H_n = L_n = M_n = \emptyset$.

2) In order to study $R'_{n,m}$, we essentielly repeat what was done for Proposition 6.1, with some slight modifications.

First suppose that $m \in K_n \cup H_n \cup M_n$. Under P^n_θ (resp. P'^n) the vector $R'^{n,m}$ is centered Gaussian with covariance matrix $C^{n,m}$ (resp. $C'^{n,m}$). If $m \in K_n \cup M_n$ (resp. $m \in H_n$) the $\dot c^n_i$ for $i \in J_{n,m}$ all have the same sign, say α_n ($= +1$ or $= -1$) (resp. $\dot c^n_i = 0$ for $i \in J_{n,m}$), hence $C^{n,m} = C(0,1,\rho_n)$ (*cf.* (8.4)) and $C'^{n,m} = C(v_{n,m},1,\rho_n)$, provided $k = k_{n,m}$ and $c_i = c^n_{i_{n,m}+i-1}$ and $b_i = b^{n,m}_i = |\dot c^n_{i_{n,m}+i-1}|$ and $v_{n,m} = \alpha_n u_n h_n \left(1+\frac{u_n h_n \ddot c_{n,m}}{2\dot c_{n,m}}\right)$ (resp. $v_{n,m} = u_n^2 h_n^2 \ddot c_{n,m}/2$).

Now, for $m \in K_n \cup M_n$, we introduce the vector $S^{n,m}$ with components $S^{n,m}_i = R'^{n,m}_i/\sqrt{b^{n,m}_i}$: on the one side the σ-fields generated by $R'^{n,m}$ and by $S^{n,m}$ coincide. On the other side these vectors are centered Gaussian under both P^n_θ and P'^n, and their covariance matrices are the matrices $\widehat C^{n,m}$ and $\widehat C'^{n,m}$ associated with $C^{n,m}$ and $C'^{n,m}$ as in Lemma 8.1. Write $\widehat\lambda^{n,m}_i$ and $\widehat\lambda'^{n,m}_i$, $i = 1,\dots,k_{n,m}$ for the increasingly ordered eigenvalues of these matrices. From Lemma 8.1 the orthogonal matrix which diagonalizes $\widehat C^{n,m}$ and $\widehat C'^{n,m}$ is the same, say $P^{n,m}$. Hence the vector $Y^{n,m} = (P^{n,m})^\star S^{n,m}$ has components $Y^{n,m}_i$ which are independent normal centered with variances $\widehat\lambda^{n,m}_i$ and $\widehat\lambda'^{n,m}_i$ under P^n_θ and P'^n respectively. Hence the variables $Y'^{n,m}_i = Y^{n,m}_i/\sqrt{\widehat\lambda^{n,m}_i}$ are independent, $\mathcal{N}(0,1)$ underP^n_θ and $\mathcal{N}(0,\widehat\lambda'^{n,m}_i/\widehat\lambda^{n,m}_i)$ under P'^n, and they generate the same σ-field than $R'^{n,m}$. Further, we have (6.3) with $\lambda_i(1,\rho_n)$ instead of $\lambda_i(2,\rho_n)$, and we define $\delta^{n,m}_i$ by (6.4).

On the other hand, if $m \in L_n$ the 1-dimensional variable $Y'^{n,m}_{i_{n,m}} = R'^{n,m}/\sqrt{c^n_{i_{n,m}}}$ is $\mathcal{N}(0,1)$ (resp. $\mathcal{N}(0, c'^n_{i_{n,m}}/c^n_{i_{n,m}})$) under P^n_θ (resp. P'^n); in this case we set $\delta^{n,m}_{i_{n,m}} = (c'^n_{i_{n,m}} - c^n_{i_{n,m}})/c^n_{i_{n,m}}$.

Therefore, using the independence of all variables occuring in (7.3), we obtain as in Section 3 that

$$\log Z^+_n = -\frac12 \sum_{m=1}^{l_n}\sum_{i=1}^{k_{n,m}} \left(\log(1+\delta^{n,m}_i) - (Y'^{n,m}_i)^2 \frac{\delta^{n,m}_i}{1+\delta^{n,m}_i}\right),$$

and here again it remains to prove the analogues of (6.6) and (6.7), except that M_n is replaced by $\{1,2,\dots,l_n\}$ and that we must take $\widehat F_{n,m} = \sum_{i=1}^{k_{n,m}} |\delta^{n,m}_i|^2$. Again as in Proposition 6.1 we have $|\delta^{n,m}_i| \le C u_n$, hence (6.6) holds.

If $m \in K_n \cup H_n \cup M_n$, and in view of (8.2), we get the estimate (6.8) except that we should replace $i\pi/k_{n,m}$ by $(i-1)\pi/k_{n,m}$. Therefore (6.9) holds with J'_2 instead of J_2, and by (8.8) we have the following estimate, analogous to (6.10) (with the same notation W_n, w_n as in Proposition 6.1; of course if $m \in H_n$ we have $\dot c^{n,m}_{\min} = \dot c^{n,m}_{\max} = 0$)

$$\begin{aligned} \frac{h_n^2 k_{n,m}}{W_n n} w_n(c^{n,m}_{\min}) \left((|\dot c|^{n,m}_{\min})^2 - C u_n |\dot c|^{n,m}_{\min} - C u_n^2\right)^+ &\le \widehat F_{n,m} \\ \le h_n^2 W_n \left(\frac{k_{n,m}}{n} w_n(c^{n,m}_{\max}) + C u_n^2\right) &\left((|\dot c|^{n,m}_{\max})^2 + C u_n |\dot c|^{n,m}_{\max} + C u_n^2\right)\cdot \end{aligned} \tag{7.4}$$

In particular since W_n and $|\dot c|^{n,m}_{\max}$ and $w_n(y)$ for $y\in[\varepsilon,1/\varepsilon]$ are uniformly bounded, we get

$$m\in K_n\cup H_n\cup M_n \quad\Rightarrow\quad \widehat F_{n,m}\le C\frac{k_{n,m}}{n}\,. \tag{7.5}$$

4) Suppose that (H2$_\theta$) holds. Then $l_n=l'_n$ and $H_n=L_n=M_n=\emptyset$, and $k_{n,m}=k_n$ for all $m\le l_n-1$. Then using (6.10) and with ζ_n as in Proposition 6.1, we obtain (compare to (6.13); we use the same notation for r^n_m and r'^n_m and w; the last extra term on the right below comes from (7.1, 7.5) and the fact that $k_{n,l_n}=k'_n\le 2k_n$):

$$\frac{(1-\zeta_n)h_n^2}{l'_n-1}\sum_{m=1}^{l'_n-1}\frac{(l'_n-1)k_n}{n}\,w(c(\theta,r'^n_m))\dot c(\theta,r'^n_m)^2-C\zeta_n\le\ F_n$$

$$\le\ \frac{(1+\zeta_n)h_n^2}{l'_n-1}\sum_{m=1}^{l'_n-1}\frac{(l'_n-1)k_n}{n}\,w(c(\theta,r^n_m))\dot c(\theta,r^n_m)^2+C\zeta_n+Cu_n^2l'_n+\frac{C}{n^{1-\gamma}}\cdot \tag{7.6}$$

Then we conclude exactly as in Proposition 6.1 that (6.7) holds and our result is proved.

5) It remains to examine the situation when (H2$_\theta$) fails. Let $\eta(r)=\sup(|\dot c(\theta,t)-\dot c(\theta,s)|:s,t\in[0,1],|s-t|\le r)$. Then $\eta(r)$ decreases to 0 when $r\downarrow 0$, and under (H3$_\theta$) we even have $\eta(r)\le Cr^\alpha$.

When $m\in L_n$, we have $|\dot c|^{n,m}_{\max}\le C(1/n)^\alpha$ because $\dot c(\theta,.)$ vanishes at a distance less than $1/n$ from the point $i_{n,m}/n$. Hence, in view of (5.6), $\widehat F_{n,m}\le C\left(u_n^2(1/n)^{2\alpha}+u_n^4\right)$. This and (7.5) yield (since $k_{n,m}=1$ when $m\in L_n$):

$$\sum_{m\in L_n\cup M_n}\widehat F_{n,m}\le C\beta_n\left(1+u_n^2n\left(\frac1n\right)^{2\alpha}+nu_n^4\right),\quad\text{where}\quad \beta_n=\frac1n\sum_{m\in L_n\cup M_n}k_{n,m}.$$

Note that β_n is the Lebesgue measure of the union of all $I'_{n,m}$ (for $m=1,\dots,l'_n$) on which $\dot c(\theta,.)$ vanishes but is not identically 0, and for such an m we have $0<\sup_{s\in I'_{n,m}}|\dot c(\theta,s)|\le\sqrt{\varepsilon_n}$, where $\varepsilon_n=C(1/n^{1-\gamma})^{2\alpha}$. Then we prove exactly as in Lemma 5.3 that $\beta_n\to0$. The sequence nu_n^4 is always bounded; the sequence $u_n^2n^{1-2\alpha}$ tends to 0 in Cases 1 and 2 and is bounded in Case 3 as soon as $\rho_nn^{1-4\alpha}$ stays bounded, an assumption which is made in Theorem 2.1 under (H3$_\theta$). Therefore we deduce that

$$\gamma_n:=\sum_{m\in L_n\cup M_n}\widehat F_{n,m}\to 0. \tag{7.7}$$

Now we combine (6.10) for $m\in K_n\cup H_n$ (then m corresponds to one of the l_n original sets $I'_{n,m'}$) with (7.7), to obtain estimates similar to (7.6): an upper bound for F_n is clearly

$$\frac{(1+\zeta_n)h_n^2}{l'_n-1}\sum_{m=1}^{l'_n-1}\frac{(l'_n-1)k_n}{n}\,w(c(\theta,r^n_m))\dot c(\theta,r^n_m)^2+C\zeta_n+Cu_n^2l'_n+\frac{C}{n^{1-\gamma}}+\gamma_n. \tag{7.8}$$

A lower bound for F_n is like in (7.6) except that the sum is taken over all $m\le l'_n-1$ which belong to K'_n (that is, such that the original interval $I'_{n,m}$ equals one of the $I_{n,j}$ for $j\in K_n\cup H_n$). But since when $m\notin K'_n$ we have $|\dot c(\theta,r^n_m)|\le Cn^{(1-\gamma)\alpha}$, a lower bound for F_n is also given by

$$\frac{(1-\zeta_n)h_n^2}{l'_n-1}\sum_{m=1}^{l'_n-1}\frac{(l_n-1)k_n}{n}\,w(c(\theta,r'^n_m))\dot c(\theta,r'^n_m)^2-C\zeta_n-C\beta_nn^{(1-\gamma)\alpha}. \tag{7.9}$$

Then, once more like in Proposition 6.1, we conclude from (7.7) and (7.8) and (7.9) that (6.7) holds, and we are finished.

8. Appendix: Some results on matrices

Here we give some elementary results on tridiagonal matrices which come naturally as covariance matrices of our observations.

Let ρ be a positive number and β be either 1 or 2. We introduce the following $k \times k$ nonnegative symmetric and tridiagonal matrix $D(\beta,\rho)$:

$$D(\beta,\rho)_{i,j} = \begin{cases} \beta\rho & \text{if } i=j=1 \text{ or } i=j=k \\ 2\rho & \text{if } 2 \le i=j \le k-1 \\ -\rho & \text{if } |i-j|=1 \\ 0 & \text{otherwise.} \end{cases} \tag{8.1}$$

The eigenvalues $(\lambda_i(\beta,\rho): i=1,\dots,k)$ of $D(\beta,\rho)$, increasingly ordered, can be explicitely computed

$$\lambda_i(2,\rho) = 2\rho\left(1-\cos\left(\frac{i\pi}{k+1}\right)\right), \qquad \lambda_i(1,\rho) = 2\rho\left(1-\cos\left(\frac{(i-1)\pi}{k}\right)\right)\cdot \tag{8.2}$$

Next, let $b_1,\dots,b_k$ and $c_1,\dots,c_k$ be positive numbers, and set

$$\left.\begin{array}{ll} b_{\min} = \inf_i b_i, & b_{\max} = \sup_i b_i, \\ c_{\min} = \inf_i c_i, & c_{\max} = \sup_i c_i \end{array}\right\}. \tag{8.3}$$

Let also $a \in \mathbb{R}$ be such that $c_i + ab_i > 0$ for all i. With all these we associate the following $k \times k$ nonnegative symmetric and tridiagonal matrix $C(a,\beta,\rho) = C((b_i),(c_i),a,\beta,\rho)$:

$$C(a,\beta,\rho)_{i,j} = \begin{cases} c_i + ab_i + \beta\rho & \text{if } i=j=1 \text{ or } i=j=k \\ c_i + ab_i + 2\rho & \text{if } 2 \le i=j \le k-1 \\ -\rho & \text{if } |i-j|=1 \\ 0 & \text{otherwise.} \end{cases} \tag{8.4}$$

Finally let Δ be the matrice diagonal matrix with entries $\Delta_{i,i} = b_i^{-1/2}$, and the two (symmetric nonnegative) matrices $\widehat{C} = \Delta C(0,\beta,\rho)\Delta$ and $\widehat{C}' = \Delta C(a,\beta,\rho)\Delta$. We denote by $\widehat{\lambda}_i$ and $\widehat{\lambda}'_i$ the eigenvalues of $\widehat{C}$ and $\widehat{C}'$, increasingly ordered, and by Λ and Λ' the associated diagonal matrices, and by P and P' orthogonal matrices such that $\widehat{C} = P\Lambda P^*$ and $\widehat{C}' = P'\Lambda' P'^*$.

Lemma 8.1. *With the above notation, we have*

(i) $\widehat{\lambda}'_i = \widehat{\lambda}_i + a$;
(ii) $\frac{c_{\min}}{b_{\max}}\left(1+\frac{\lambda_i(\beta,\rho)}{c_{\max}}\right) \le \widehat{\lambda}_i \le \frac{c_{\max}}{b_{\min}}\left(1+\frac{\lambda_i(\beta,\rho)}{c_{\min}}\right)$;
(iii) *we can choose P' as $P' = P$.*

Proof. a) We first prove an auxiliary result: let D be a nonnegative symmetric $k \times k$ matrix, and Γ be a diagonal matrix with $\Gamma_{i,i} > 0$, and $D' = \Gamma D \Gamma$. Denote by λ_i and λ'_i the eigenvalues of D and D', increasingly ordered. Letting $\Gamma_{\max} = \sup_i \Gamma_{i,i}$ and $\Gamma_{\min} = \inf_i \Gamma_{i,i}$, we have

$$\Gamma^2_{\min}\lambda_i \le \lambda'_i \le \Gamma^2_{\max}\lambda_i. \tag{8.5}$$

To see this, let $\mathcal{E}_i$ be the family of all linear subspaces of $\mathbb{R}^k$ with dimension i. We have $\lambda_i = \inf_{E\in\mathcal{E}_i} \sup(x^* Dx : x \in E, ||x|| = 1)$, and similarly for λ'_i. The map $E \mapsto \Gamma E$ (image of E by Γ) is one-to-one and onto from $\mathcal{E}_i$ into itself, hence

$$\begin{aligned}\lambda'_i &= \inf_{E\in\mathcal{E}_i} \sup((\Gamma x)^* D\Gamma x : x \in E, ||x|| = 1)\\ &= \inf_{E\in\mathcal{E}_i} \sup(y^* Dy : y \in E, ||\Gamma^{-1}y|| = 1).\end{aligned}$$

Now if $||\Gamma^{-1}y|| = 1$ we have $\Gamma_{\min} \le ||y|| \le \Gamma_{\max}$, and (8.5) follows.

b) Set $D = D(\beta,\rho)$. Denote by M the diagonal matrix with entries $M_{i,i} = \sqrt{c_i/b_i}$, and $L = M^{-1}\Delta D\Delta M^{-1}$ and $N = I + L$, so that

$$\widehat{C} = M^2 + \Delta D\Delta = MNM.$$

Next, denote by $\lambda_i(D) = \lambda_i(\beta,\rho)$, $\lambda_i(L)$ and $\lambda_i(N) = \lambda_i(L) + 1$ the eigenvalues of D, L and N respectively, increasingly ordered. Since $(\Delta M^{-1})_{i,i} = \sqrt{1/c_i}$, applying twice (8.5) yields

$$\frac{1}{c_{\max}}\lambda_i(D) \le \lambda_i(L) \le \frac{1}{c_{\min}}\lambda_i(D),$$

$$\frac{c_{\min}}{b_{\max}}\lambda_i(N) \le \widehat{\lambda}_i \le \frac{c_{\max}}{b_{\min}}\lambda_i(N),$$

hence (ii). Next, (i) readily follows from $\widehat{C}' = \widehat{C} + aI$. Finally, $\Lambda' = \Lambda + aI$ yields for any possible choice of P and P':

$$P'\Lambda P'^* + aI = \widehat{C}' = \widehat{C} + aI = P\Lambda P^* + aI,$$

and (iii) follows. □

Finally, in connection with the eigenvalues given in (8.2), we introduce the function $\phi(x,a) = 2(1-\cos(x))+a$ on $(0,\infty)\times\mathbb{R}$. We need the following simple properties of integrals and Riemann sums:

First, for any $a > 0$ the following integral may be explicitely computed:

$$I_2(a) := \int_0^\pi \frac{1}{\phi(x,a)^2}\mathrm{d}x = \frac{\pi(2+a)}{a^{3/2}(4+a)^{3/2}}\cdot \tag{8.6}$$

Next, define the Riemann sums, for $k = 2,3,\ldots$:

$$J_2(a,k) = \frac{\pi}{k}\sum_{i=1}^{k-1}\frac{1}{\phi\left(\frac{i\pi}{k},a\right)^2}, \qquad J'_2(a,k) = \frac{\pi}{k}\sum_{i=0}^{k-1}\frac{1}{\phi\left(\frac{i\pi}{k},a\right)^2}\cdot \tag{8.7}$$

We then have the following inequalities:

$$I_2(a) - \frac{\pi}{ka^2} \le J_2(a,k) \le I_2(a) \le J'_2(a,k) \le I_2(a) + \frac{\pi}{ka^2}\cdot \tag{8.8}$$

References

[1] P. Bickel and Y. Ritov, Inference in hidden Markov models. I. Local asymptotic normality in the stationary case. *Bernoulli* **2** (1996) 199-228.

[2] P. Bickel and Y. Ritov, Asymptotic normality for the maximum likelihood estimator for general hidden Markov models. *Ann. Statist.* **26** (1998) 1614-1635.

[3] G. Dohnal, On estimating the diffusion coefficient. *J. Appl. Probab.* **24** (1987) 105-114.

[4] V. Genon–Catalot and J. Jacod, On the estimation of the diffusion coefficient for multidimensional diffusion processes. *Ann. Inst. H. Poincaré Probab. Statist.* **29** (1993) 119-153.

[5] V. Genon–Catalot and J. Jacod, Estimation of the diffusion coefficient for diffusion processes: random sampling. *Scand. J. Statist.* **21** (1994) 193-221.

[6] A. Gloter and J. Jacod, *Diffusion with measurement error. II. Optimal estimators* (2000).

[7] E. Gobet, *LAMN property for elliptic diffusions* (2000).
[8] J. Jacod and A. Shiryaev, *Limit Theorems for Stochastic Processes*. Springer-Verlag, Berlin (1987).
[9] J.L. Jensen and N. Petersen, Asymptotic normality of the Maximum likelihood estimator in state space models. *Ann. Statist.* **27** (1999) 514-535.
[10] B. Leroux, Maximum likelihood estimation for hidden Markov models. *Stochastic Process. Appl.* **40** (1992) 127-143.
[11] L. LeCam and G.L. Yang, *Asymptotics in Statistics*. Springer-Verlag, Berlin (1990).
[12] M.B. Malyutov and O. Bayborodin, Fitting diffusion and trend in noise via Mercer expansion, in *Proc. 7th Int. Conf. on Analytical and Stochastic Modeling Techniques*. Hamburg (2000).
[13] T. Ryden, Consistent and asymptotically normal estimators for hidden Markov models. *Ann. Statist.* **22** (1994) 1884-1895.