



# Efficient parametric estimation for a signal-plus-noise Gaussian model from discrete time observations

Dominique Dehay<sup>1</sup> · Khalil El Waled<sup>2,3</sup> · Vincent Monsan<sup>4</sup>

Received: 24 April 2019 / Accepted: 28 July 2020 / Published online: 5 August 2020  
© Springer Nature B.V. 2020

## Abstract

This paper deals with the parametric inference for integrated continuous time signals embedded in an additive Gaussian noise and observed at deterministic discrete instants which are not necessarily equidistant. The unknown parameter is multidimensional and compounded of a signal-of-interest parameter and a variance parameter of the noise. We state the consistency and the minimax efficiency of the maximum likelihood estimator and of the Bayesian estimator when the time of observation tends to infinity and the delays between two consecutive observations tend to 0 or are only bounded. The class of signals in consideration contains among others, almost periodic signals and also non-continuous periodic signals. However the problem of frequency estimation is not considered here. Furthermore, in this paper the signal-plus-noise discretely observed in time model is considered as a particular case of a more general model of independent Gaussian observations forming a triangular array.

**Keywords** Maximum likelihood estimation · Bayesian estimation · High frequency sampling · Low frequency sampling · Minimax efficiency · Asymptotic properties of estimators · Triangular Gaussian array · Hellinger distance

## 1 Introduction

Consider the following integrated signal-plus-noise model

$$dX_t = f(\boldsymbol{\alpha}, t) dt + \sigma(\boldsymbol{\beta}, t) dW_t, \quad t \geq 0 \quad (1)$$

---

✉ Dominique Dehay  
dominique.dehay@univ-rennes2.fr

Khalil El Waled  
khalil.elwaled@gmail.com

Vincent Monsan  
vmonsan@yahoo.fr

<sup>1</sup> CNRS, IRMAR – UMR 6625, Univ Rennes, 35000 Rennes, France

<sup>2</sup> University of Nouakchott Al Aasriya, Nouakchott, Mauritania

<sup>3</sup> Qassim University, Buraydah, Saudi Arabia

<sup>4</sup> Université Félix Houphouët-Boigny, Abidjan, Côte d'Ivoire

where the functions  $f : \mathbf{A} \times \mathbb{R}^+ \rightarrow \mathbb{R}$  and  $\sigma : \mathbf{B} \times \mathbb{R}^+ \rightarrow \mathbb{R}^+$  are measurable,  $f(\boldsymbol{\alpha}, t)$ , respectively  $\sigma(\boldsymbol{\beta}, t)$ , is continuous in the first component  $\boldsymbol{\alpha} \in \mathbf{A}$ , respectively in  $\boldsymbol{\beta} \in \mathbf{B}$ ;  $\mathbf{A}$  is a bounded open convex subset of  $\mathbb{R}^a$ ,  $\mathbf{B}$  is a bounded open convex subset of  $\mathbb{R}^b$ ,  $a, b \geq 0$ ,  $a + b > 0$ , and  $\{W_t\}$  is a Wiener process defined over a probability space  $(\Omega, \mathcal{F}, \mathbb{P})$ . We assume that the initial random variable  $X_0$  is independent on the Wiener process  $\{W_t\}$  and does not depend on the unknown parameter  $\boldsymbol{\theta} := (\boldsymbol{\alpha}, \boldsymbol{\beta})$ .

Since very long time, this model has received a considerable amount of investigation. The statistical analysis of such signals has attracted much interest, its applications ranging from telecommunications, mechanics, to econometrics and financial studies. For the continuous time observation framework, we cite the well-known work by Ibragimov and Has'minskii (1981) as well as the contributions by Kutoyants (1984) who studied the consistency and the minimax efficiency of the maximum likelihood estimator and the Bayesian estimator.

However, in practice it is difficult to record numerically a continuous time process and generally the observations take place at discrete moments (Mishra and Prakasa Rao 2001). Most of the publications on discrete time observation concern regular sampling, that is the instants of observation are usually equally spaced. Nevertheless, many applications make use of non equidistant sampling. The sampled points can be associated with quantiles of some distribution (see e.g. in another context Blanke and Vial 2014; see also Sacks and Ylvisaker 1968) or can be perturbed by some jitter effect (see e.g. Dehay et al. 2017).

In the present paper we study the asymptotic optimality of the maximum likelihood estimator and the Bayesian estimator of the unknown parameter  $\boldsymbol{\theta} = (\boldsymbol{\alpha}, \boldsymbol{\beta})$  built from a discrete time observation  $\{X_{t_0}, \dots, X_{t_n}\}$ ,  $0 = t_0 < t_1 < \dots < t_n = T_n$  of the process  $\{X_t\}$  as  $n$  and  $T_n \rightarrow \infty$ , when the delays between two consecutive observations tend to 0 or are only bounded. The non uniform sampling scheme is scarcely taken in consideration in the usual literature on the inference of such a model (1) of integrated signal-plus-noise. For this scheme of observation we state that the rate of convergence of the maximum likelihood estimator and the Bayesian estimator for the parameter  $\boldsymbol{\alpha}$  of the signal-of-interest is  $\sqrt{T_n}$  while the rate of convergence for the parameter  $\boldsymbol{\beta}$  of the noise variance is  $\sqrt{n}$ , without any condition on the convergence to 0 of the delay between two observations as  $n \rightarrow \infty$ , in contrary to the model of an ergodic diffusion (Florens-Zmirou 1989; Genon-Catalot 1990; Mishra and Prakasa Rao 2001; Uchida and Yoshida 2012). This fact is due to the non-randomness of the signal-of-interest  $f(\boldsymbol{\alpha}, t)$  and of the variance  $\sigma^2(\boldsymbol{\beta}, t)$ . Notice that model (1) is not ergodic, and the signal-of-interest is not necessarily continuous or periodic in time. The problem of frequency estimation is not tackled in this work.

The paper is organized as follows. The framework and the assumptions on the model are introduced in Sect. 2. In the following we are going to consider a general parametric estimation problem for a triangular Gaussian array with independent components. The signal-plus-noise discretely observed in time model (1) is a particular case of the general setting. Next specific assumptions for the general model that fit this special case are presented. These conditions are fulfilled for almost periodic signal-plus-noise Gaussian models.

The uniform local asymptotic normality of the log-likelihood in the setting of independent random variables has been stated in Theorem II. 6.2 of Ibragimov and Has'minskii (1981) under a Lindeberg condition. However, in our framework this condition can be simplified using Hellinger-type distances (see Theorem 4.1 and relation (5.6) in Dzhaparidze and Valkeila 1990). Moreover the Hellinger-type distances between two distributions of the observation according two different values of the parameter can be easily estimated ("Appendix A", see also Gushchin and K uchler 2003). Then the local asymptotic normality of the log-likelihood for our parametric Gaussian model is established in Sect. 3 following a methodology presented in Gushchin and Valkeila (2017).

Sections 4.1 and 4.2 are devoted to prove that the maximum likelihood estimator and the Bayesian estimator are consistent, asymptotically normal and asymptotically optimal, following the method of minimax efficiency from Chapter III in Ibragimov and Has’minskii (1981). Again, using Hellinger-type distances allow us to simplify the computations.

To illustrate the previous results, the trivial example of the linear model is given in Sect. 5.

In “Appendix A”, some elementary results for Hellinger-type distances between products of normal distributions and Hellinger integral of them are provided. These results generalize results stated in Gushchin and Küchler (2003) for zero-mean normal distributions.

We complete this work in “Appendix B” by stating some expressions of the Fisher information matrices and of the identifiability functions in the cases of almost periodic and of periodic functions.

## 2 Framework

### 2.1 Triangular Gaussian array model

First consider that the observations come from a triangular random array formed for each  $n > 0$  by independent Gaussian variables  $Y_{n,i}$ ,  $i = 1, \dots, n$  with mean  $F_{n,i}(\alpha)$  and variance  $G_{n,i}^2(\beta)$  which depend on the unknown parameter  $\theta = (\alpha, \beta) \in \Theta = A \times B \subset \mathbb{R}^{a+b}$ . Hence the likelihood of the observation  $\{Y_{n,i} : i = 1 \dots, n\}$  is equal to

$$Z_n(\theta) = \exp \left( \frac{-n \ln(2\pi)}{2} - \sum_{i=1}^n \ln G_{n,i}(\beta) - \sum_{i=1}^n \frac{(Y_{n,i} - F_{n,i}(\alpha))^2}{2G_{n,i}^2(\beta)} \right).$$

Hereafter the properties of the likelihood  $Z_n(\theta)$  will be obtained under the following assumptions.

**A1** For every  $n \geq 1$  and every  $i \in \{1, \dots, n\}$ , the functions  $\alpha \mapsto F_{n,i}(\alpha)$  and  $\beta \mapsto G_{n,i}^2(\beta)$  are differentiable. Moreover there exist a bounded family of positive numbers  $\{h_{n,i} : n \geq 1, i = 1, \dots, n\}$ , and for every  $\varepsilon > 0$  there is  $\eta > 0$  such that for  $|\alpha - \alpha'| \leq \eta$  and  $|\beta - \beta'| \leq \eta$  we have that

$$|\nabla_{\alpha} F_{n,i}(\alpha) - \nabla_{\alpha} F_{n,i}(\alpha')| \leq \varepsilon h_{n,i} \quad \text{and} \quad |\nabla_{\beta} G_{n,i}^2(\beta) - \nabla_{\beta} G_{n,i}^2(\beta')| \leq \varepsilon h_{n,i}.$$

Here the  $a$ -dimensional vector  $\nabla_{\alpha} F_{n,i}(\alpha)$  is the gradient function of  $F_{n,i}(\alpha)$  with respect to  $\alpha$  and the  $b$ -dimensional vector  $\nabla_{\beta} G_{n,i}^2(\beta)$  is the gradient function of  $G_{n,i}^2(\beta)$  with respect to  $\beta$ .

**A2** There exist constant values  $0 < \gamma_1 < \gamma_2 < \infty$  such that

$$0 < \gamma_1 h_{n,i} \leq \inf_{\beta} G_{n,i}^2(\beta) \leq \sup_{\beta} G_{n,i}^2(\beta) < \gamma_2 h_{n,i},$$

$$\sup_{\alpha} |\nabla_{\alpha} F_{n,i}(\alpha)| < \gamma_2 h_{n,i} \quad \text{and} \quad \sup_{\beta} |\nabla_{\beta} G_{n,i}^2(\beta)| < \gamma_2 h_{n,i}.$$

In the following we also assume that  $T_n := \sum_{i=1}^n h_{n,i} \rightarrow \infty$ , as  $n \rightarrow \infty$ .

**A3** There exist two positive definite matrices  $J_a^{(\theta)}$  and  $J_b^{(\beta)}$  such that

$$J_a^{(\theta)} = \lim_{n \rightarrow \infty} \frac{1}{T_n} \sum_{i=1}^n \frac{\nabla_{\alpha}^* F_{n,i}(\alpha) \nabla_{\alpha} F_{n,i}(\alpha)}{G_{n,i}^2(\beta)}$$

and

$$J_b^{(\beta)} = \lim_{n \rightarrow \infty} \frac{1}{2n} \sum_{i=1}^n \nabla_{\beta}^* \ln G_{n,i}^2(\beta) \nabla_{\beta} \ln G_{n,i}^2(\beta),$$

the convergences being uniform with respect to  $\theta$  varying in  $\Theta = A \times B$ . Here and henceforth the superscript\* designates the transpose operator for vectors and matrices.

**A4** For every  $\nu > 0$  there exists  $\mu_{\nu} > 0$  and  $n_{\nu} > 0$  such that

$$\frac{1}{T_n} \sum_{i=1}^n \frac{(F_{n,i}(\alpha) - F_{n,i}(\alpha'))^2}{h_{n,i}} \geq \mu_{\nu} \quad \text{and} \quad \frac{1}{n} \sum_{i=1}^n \frac{(G_{n,i}^2(\beta) - G_{n,i}^2(\beta'))^2}{(h_{n,i})^2} \geq \mu_{\nu}$$

for any  $n \geq n_{\nu}$  and all  $\theta = (\alpha, \beta), \theta' = (\alpha', \beta')$  in  $\Theta$  with  $|\alpha - \alpha'| \geq \nu$  and  $|\beta - \beta'| \geq \nu$ .

**Remarks** 1. Assumptions A1 and A2 are technical conditions on the smoothness of the functions  $F_{n,i}(\alpha)$  and  $G_{n,i}(\beta)$ .

2. With assumption A3 the asymptotic Fisher information  $d \times d$ -matrix  $J^{(\theta)}$  of the model,  $d := a + b$ , is equal to

$$J^{(\theta)} := \text{diag} \left[ J_a^{(\theta)}, J_b^{(\beta)} \right] = \begin{bmatrix} J_a^{(\theta)} & \mathbf{0}_{a \times b} \\ \mathbf{0}_{b \times a} & J_b^{(\beta)} \end{bmatrix}.$$

Under assumptions A1, A2 and A3, the function  $\theta \mapsto J^{(\theta)}$  is continuous on  $\Theta = A \times B$ . Furthermore as  $J^{(\theta)}$  is a positive definite matrix, the square root  $(J^{(\theta)})^{\frac{-1}{2}} = \text{diag} \left[ (J_a^{(\alpha, \beta)})^{\frac{-1}{2}}, (J_b^{(\beta)})^{\frac{-1}{2}} \right]$  of its inverse  $(J^{(\theta)})^{-1}$  is well defined and continuous in  $\Theta$ .

3. Assumption A4 is an identifiability condition. Assume that the following limits are finite:

$$\mu_a(\alpha, \alpha') := \liminf_{n \rightarrow \infty} \frac{1}{T_n} \sum_{i=1}^n \frac{(F_{n,i}(\alpha) - F_{n,i}(\alpha'))^2}{h_{n,i}}, \tag{2}$$

$$\mu_b(\beta, \beta') := \liminf_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n \frac{(G_{n,i}^2(\beta) - G_{n,i}^2(\beta'))^2}{h_{n,i}^2}, \tag{3}$$

the convergences being uniform with respect to  $\alpha, \alpha' \in A, \beta, \beta' \in B$  with  $|\alpha - \alpha'| \geq \nu$  and  $|\beta - \beta'| \geq \nu$ , and assume that

$$\mu_{\nu} := \frac{1}{2} \min \left\{ \inf_{|\alpha - \alpha'| \geq \nu} \mu_a(\alpha, \alpha'), \inf_{|\beta - \beta'| \geq \nu} \mu_b(\beta, \beta') \right\} > 0$$

for any  $\nu > 0$ , then assumption A4 is fulfilled.

## 2.2 Discrete time observation of the signal-plus-noise model

As an application of the previous model, come back to the continuous time integrated signal-plus-noise model (1) for which we assume that the observations during a time interval  $[0, T_n]$  occur at instants  $0 = t_{n,0} < t_{n,1} < \dots < t_{n,n} = T_n$  of the interval  $[0, T_n]$ , where  $0 < h_{n,i} := t_{n,i} - t_{n,i-1} \leq h_n := \max_i h_{n,i}$ . We also assume that  $T_n \rightarrow \infty$  as  $n \rightarrow \infty$  and the set  $\{h_n : n \geq 1\}$  is bounded.

Then the observation of the sequence  $X_{t_{n,i}}, i \in \{0, \dots, n\}$  corresponds to the observation of  $Y_{n,0} := X_0$  and of the increments defined by  $Y_{n,i} := X_{t_{n,i}} - X_{t_{n,i-1}}, i \in \{1, \dots, n\}$ . Since

the initial random variable  $X_0$  is independent of the Wiener process  $\{W_t\}$  and its distribution does not depend on the parameter  $\theta = (\alpha, \beta)$ , we do not take care of  $Y_{n,0}$  in the following.

For  $i \in 1, \dots, n$ , denote

$$F_{n,i}(\alpha) := \int_{t_{n,i-1}}^{t_{n,i}} f(\alpha, t) dt \quad \text{and} \quad G_{n,i}(\beta) := \left( \int_{t_{n,i-1}}^{t_{n,i}} \sigma^2(\beta, t) dt \right)^{1/2}.$$

Thus, when the true value of the parameter is  $\theta = (\alpha, \beta)$ , the increment  $Y_{n,i}, i \geq 1$ , is equal to

$$Y_{n,i}^{(\theta)} := F_{n,i}(\alpha) + G_{n,i}(\beta)W_{n,i}^{(\beta)}$$

where

$$W_{n,i}^{(\beta)} := \frac{1}{G_{n,i}(\beta)} \int_{t_{n,i-1}}^{t_{n,i}} \sigma(\beta, t) dW_t.$$

Notice that the random variables  $W_{n,i}^{(\beta)}, i = 1, \dots, n$ , are independent with the same standard normal distribution  $\mathcal{N}(0, 1)$ . Hence the triangular Gaussian array model introduced in Sect. 2.1 applies in the present situation.

Moreover, assumptions A1 and A2 are fulfilled when the following specific to model (1) assumptions A1' and A2' are satisfied.

**A1'** The functions  $f : A \times \mathbb{R}^+ \rightarrow \mathbb{R}$ , and  $\sigma : B \times \mathbb{R}^+ \rightarrow \mathbb{R}^+$ , are measurable. The functions  $\alpha \mapsto f(\alpha, t)$  and  $\beta \mapsto \sigma^2(\beta, t)$  are differentiable. Furthermore, for every  $\varepsilon > 0$  there exists  $\eta > 0$  such that for  $|\alpha - \alpha'| \leq \eta$  and  $|\beta - \beta'| \leq \eta$  we have:

$$\sup_t |\nabla_\alpha f(\alpha, t) - \nabla_\alpha f(\alpha', t)| \leq \varepsilon \quad \text{and} \quad \sup_t |\nabla_\beta \sigma^2(\beta, t) - \nabla_\beta \sigma^2(\beta', t)| \leq \varepsilon.$$

**A2'** The function  $t \mapsto f(\alpha, t)$  is locally integrable in  $\mathbb{R}$  for any  $\alpha \in A$ ; moreover

$$0 < \inf_{\beta,t} \sigma^2(\beta, t) \leq \sup_{\beta,t} \sigma^2(\beta, t) < \infty,$$

$$\sup_{\alpha,t} |\nabla_\alpha f(\alpha, t)| < \infty \quad \text{and} \quad \sup_{\beta,t} |\nabla_\beta \sigma^2(\beta, t)| < \infty.$$

**Remarks** 1. We readily see that assumptions A1' and A2' (as well as A1 and A2) are fulfilled when the parameter set  $\Theta = A \times B$  is compact and the functions  $(\alpha, t) \mapsto (f(\alpha, t), \nabla_\alpha f(\alpha, t))$  and  $(\beta, t) \mapsto (\sigma^2(\beta, t), \nabla_\beta \sigma^2(\beta, t))$  are almost periodic in  $t$  uniformly with respect to  $\theta = (\alpha, \beta) \in \Theta$  (see ‘‘Appendix B’’).

2. Neither assumption A1' nor assumption A1 are satisfied when we consider the problem of frequency estimation for model (1). For example, the function  $f(\alpha, t) = \sin(\alpha t)$  with  $\alpha \in A, A \subset \mathbb{R}$  does not satisfied assumption A1 since  $\sup_t |\cos(\alpha t) - \cos(\alpha' t)| = 2$  when  $\alpha \neq \alpha'$ .
3. For model (1), the limits  $J_a^{(\theta)}$  and  $J_b^{(\beta)}$  in assumption A3 exist when the functions  $\nabla_\alpha f(\alpha, t), \sigma^2(\beta, t)$  and  $\nabla_\beta \sigma^2(\beta, t)$  are almost periodic in time  $t$  and  $h_n \rightarrow 0$ . This is also true when these functions are periodic with the same period  $P > 0$  and the delay between two consecutive observations is constant and equal to  $h = P/\nu$  for some fixed positive integer  $\nu$  (see ‘‘Appendix B’’).
4. For model (1), when the functions  $f(\alpha, t), f(\alpha', t), \sigma^2(\beta, t)$  and  $\sigma^2(\beta', t)$  are almost periodic in time  $t$ , then the limits  $\mu_a(\alpha, \alpha')$  and  $\mu_b(\beta, \beta')$  in relations (2) and (3) exist. If in addition for  $\alpha \neq \alpha'$  we have  $\sup_t |f(\alpha, t) - f(\alpha', t)| > 0$ , and for  $\beta \neq \beta'$ ,  $\sup_t |\sigma^2(\beta, t) - \sigma^2(\beta', t)| > 0$ , then  $\mu_a(\alpha, \alpha')$  and  $\mu_b(\beta, \beta')$  are positive. See ‘‘Appendix B’’.

- 5. Some explicit expressions of  $J_a^{(\theta)}$ ,  $J_b^{(\beta)}$ ,  $\mu_a(\alpha, \alpha')$  and  $\mu_b(\beta, \beta')$  are given in “Appendix B” for functions  $f(\theta, t)$  and  $\sigma^2(\beta, t)$  periodic in time  $t$  when the delays between two consecutive observations tend to 0 as well as when the delays are constant.

### 3 LAN property of the triangular Gaussian array model

From now on, we consider the general model of triangular Gaussian array presented in Sect. 2.1. As already mentioned, the results stated in the following apply directly to the integrated signal-plus-noise model described in Sect. 2.2. Under assumptions A1, A2 and A3, the model of triangular Gaussian array is regular in the sense given in Ibragimov IA, Has’minskii (1981, pp 65). To establish the asymptotic normality and the asymptotic efficiency of the maximum likelihood estimator and of the Bayesian estimator, we will apply the minimax efficiency theory from Ibragimov and Has’minskii (1981). Thus, we study the asymptotic behaviour of the likelihood of the observation  $\{Y_{n,1}, \dots, Y_{n,n}\}$  in the neighbourhood of the true value of the parameter. For this purpose let  $Z_n^{(\theta, \mathbf{w})}$  be the likelihood ratio defined by

$$Z_n^{(\theta, \mathbf{w})} := \frac{dP_n^{\theta + \mathbf{w}\Phi_n^{(\theta)}}}{dP_n^\theta}(Y_{n,1}, \dots, Y_{n,n}) = \prod_{i=1}^n \frac{dP_{n,i}^{\theta + \mathbf{w}\Phi_n^{(\theta)}}}{dP_{n,i}^\theta}(Y_{n,i}).$$

for  $\theta \in \Theta$  and  $\mathbf{w} \in \mathcal{W}_{\theta,n} := \{\mathbf{w} \in \mathbb{R}^d : \theta + \mathbf{w}\Phi_n^{(\theta)} \in \Theta\}$ . Recall that  $d := a + b$ . Here the invertible  $d \times d$ -matrix (local normalizing matrix)  $\Phi_n^{(\theta)}$  is equal to  $\Phi_n^{(\theta)} := \text{diag}[\varphi_n^{(\theta)}, \psi_n^{(\beta)}]$ , setting  $\varphi_n^{(\alpha,\beta)} := (T_n J_a^{(\theta)})^{-\frac{1}{2}}$  and  $\psi_n^{(\beta)} := (n J_b^{(\beta)})^{-\frac{1}{2}}$ . Furthermore  $P_n^{(\theta')}$ , respectively  $P_{n,i}^{(\theta')}$ , is the distribution of  $(Y_{n,1}, \dots, Y_{n,n})$ , respectively of  $Y_{n,i}$ , when the value of the parameter is  $\theta' \in \Theta$ . Next we state that the log-likelihood ratio  $\ln Z_n^{(\theta, \mathbf{w})}$  is asymptotically normal as  $n \rightarrow \infty$ .

**Proposition 1** *Assume that  $\Theta = A \times B$  is open and convex, and conditions A1, A2 and A3 are fulfilled. Then the family  $\{P_n^\theta : \theta \in \Theta\}$  is uniformly locally asymptotically normal (uniformly LAN) in any compact subset  $\mathcal{K}$  of  $\Theta$ . That is for any compact subset  $\mathcal{K}$  of  $\Theta$ , for arbitrary sequences  $\{\theta_n\} \subset \mathcal{K}$  and  $\{\mathbf{w}_n\} \subset \mathbb{R}^d$  such that  $\theta_n + \mathbf{w}_n \Phi_n^{(\theta_n)} \in \mathcal{K}$  and  $\mathbf{w}_n \rightarrow \mathbf{w} \in \mathbb{R}^d$  as  $n \rightarrow \infty$ , the log-likelihood ratio  $\ln Z_n^{(\theta_n, \mathbf{w}_n)}$  can be decomposed as*

$$\ln Z_n^{(\theta_n, \mathbf{w}_n)} = \Delta_n^{(\theta_n)} \mathbf{w}^* - \frac{1}{2} |\mathbf{w}|^2 + r_n(\theta_n, \mathbf{w}_n) \tag{4}$$

where the random vector  $\Delta_n^{(\theta_n)}$  converges in law to the  $d$ -dimensional standard normal distribution and the random variable  $r_n(\theta_n, \mathbf{w}_n)$  converges in  $P_{\theta_n}$ -probability to 0.

**Proof** The weak convergence of log-likelihood to a Gaussian limit holds if the Hellinger process (of order 1/2) converges to a corresponding constant and the Lindeberg condition holds, see Chapter 10 in Jacod and Shiryaev (1987). Dzhaparidze and Valkeila (1990, Theorem 4.1) have shown that the Lindeberg condition can be replaced by convergence of the so-called  $p$ -divergency process to 0. In the case of independent observations the Hellinger process and the  $p$ -divergency process are just the sum of squared Hellinger distances and the corresponding Hellinger  $p$ -distances,  $p > 2$  integer (Jacod and Shiryaev 1987, Proposition IV.1.73). Moreover, in the Gaussian case, these distances can be expressed via the means and variances, and the Hellinger  $p$ -distances can be estimated via the Hellinger distance

(Hellinger 2-distance  $\rho_2$ ). See inequality (18) in “Appendix A” below (see also, Gushchin and Küchler 2003). In the following we do not apply directly the notion of Hellinger process and  $p$ -divergency process. Thus we do not define them in the paper. For more information on these notions the reader is referred to Jacod and Shiryaev (1987) and Dzhariparidze and Valkeila (1990).

Hence, following Gushchin and Valkeila (2017), to prove that the family  $\{P_n^{(\theta)} : \theta \in \Theta\}$  is LAN uniformly in any compact subset  $\mathcal{K}$  of  $\Theta$ , it suffices to establish that

$$\limsup_n \sup_i \rho_2^2(P_{n,i}^{(\theta_n)}, P_{n,i}^{(\theta_n + \mathbf{w}_n \Phi_n^{(\theta_n)})}) = 0 \tag{5}$$

and

$$\lim_{n \rightarrow \infty} \sum_{i=1}^n \rho_2^2(P_{n,i}^{(\theta_n + \mathbf{w}_{n,1} \Phi_n^{(\theta_n)})}, P_{n,i}^{(\theta_n + \mathbf{w}_{n,2} \Phi_n^{(\theta_n)})}) - \frac{1}{4} |\mathbf{w}_{n,1} - \mathbf{w}_{n,2}|^2 = 0 \tag{6}$$

for any  $\{\theta_n\} \subset \mathcal{K}$  and any bounded  $\{\mathbf{w}_{n,j}\} \subset \mathbb{R}^d$  such that  $\theta_n + \mathbf{w}_{n,i} \Phi_n^{(\theta_n)} \in \Theta, j = 1, 2$ . The definition of the Hellinger distance  $\rho_2(\cdot, \cdot)$  is recalled in “Appendix A”.

Let  $\theta_n = (\alpha_n, \beta_n) \in \mathcal{K}$ , and let  $\{\mathbf{w}_{n,j}\} \subset \mathbb{R}^d$  be bounded for  $j = 1, 2$ . Denote  $\theta_n^{(j)} = (\alpha_n^{(j)}, \beta_n^{(j)}) := \theta_n + \mathbf{w}_{n,j} \Phi_n^{(\theta_n)}$  for  $j = 1, 2$ , and  $\mathbf{w}_n = (\mathbf{u}_n, \mathbf{v}_n) := \mathbf{w}_{n,2} - \mathbf{w}_{n,1}$ . Thanks to equality (12) in “Appendix A”, the Hellinger distance between the two normal distributions  $P_{n,i}^{(\theta_n^{(j)})} = \mathcal{N}(F_{n,i}(\alpha_n^{(j)}), G_{n,i}^2(\beta_n^{(j)})), j = 1, 2$ , is equal to

$$\rho_2^2(P_{n,i}^{(\theta_n^{(1)})}, P_{n,i}^{(\theta_n^{(2)})}) = 2 - 2 \left( 1 + \frac{y_{n,i}^2}{4(1 + y_{n,i})} \right)^{-1/4} \exp \left( \frac{-x_{n,i}^2}{4(2 + y_{n,i})} \right)$$

where

$$\begin{aligned} x_{n,i} &:= \frac{F_{n,i}(\alpha_n^{(2)}) - F_{n,i}(\alpha_n^{(1)})}{G_{n,i}(\beta_n^{(1)})} \\ &= \frac{\mathbf{u}_n \phi_n^{(\alpha_n, \beta_n)}}{G_{n,i}(\beta_n^{(1)})} \int_0^1 \nabla_{\alpha}^* F_{n,i}(\alpha_n^{(1)} + t \mathbf{u}_n \phi_n^{(\alpha_n, \beta_n)}) dt \end{aligned} \tag{7}$$

and

$$\begin{aligned} y_{n,i} &:= \frac{G_{n,i}^2(\beta_n^{(1)} + \mathbf{v}_n \phi_n^{(\beta_n)}) - G_{n,i}^2(\beta_n^{(1)})}{G_{n,i}^2(\beta_n^{(1)})} \\ &= \frac{\mathbf{v}_n \phi_n^{(\beta_n)}}{G_{n,i}^2(\beta_n^{(1)})} \int_0^1 \nabla_{\beta}^* G_{n,i}^2(\beta_n^{(1)} + t \mathbf{v}_n \phi_n^{(\beta_n)}) dt. \end{aligned} \tag{8}$$

From assumption A2, we readily obtain that

$$x_{n,i}^2 \leq T_n^{-1} h_{n,i} |\mathbf{u}_n|^2 |(\mathbf{J}_a^{(\alpha_n, \beta_n)})^{-1/2}|^2 \gamma_2^2 \gamma_1^{-1} \tag{9}$$

and

$$y_{n,i}^2 \leq n^{-1} |\mathbf{v}_n|^2 |(\mathbf{J}_b^{(\beta_n)})^{-1/2}|^2 \gamma_2^2 \gamma_1^{-2}. \tag{10}$$

By remark 2 in Sect. 2.1, we know that the functions  $\theta \mapsto |(\mathbf{J}_a^{(\theta)})^{-1/2}|$  and  $\theta \mapsto |(\mathbf{J}_b^{(\theta)})^{-1/2}|$  are well defined and continuous in  $\Theta$ , so they are bounded in the compact subset  $\mathcal{K}$ . Hence when  $\{\theta_n\} \subset \mathcal{K}$  and the sequence  $\{\mathbf{w}_n = (\mathbf{u}_n, \mathbf{v}_n)\}$  is bounded, we deduce that condition (5) is fulfilled.

Next, thanks to the Taylor expansion formula with integral reminder in expressions (7) and (8), and to the asymptotic equality (15) in “Appendix A” applied to  $\rho_2^2(\mathbf{P}_{n,i}^{(\theta_n^{(1)})}, \mathbf{P}_{n,i}^{(\theta_n^{(2)})})$ , assumptions (A1), (A2) and (A3) imply that condition (6) is satisfied.

Hence the proposition is proved. □

**Remarks** By a direct but boring computation, we can prove the LAN decomposition (4) with the random vector  $\Delta_n^{(\theta)} = \sum_{i=1}^n \Delta_{n,i}^{(\theta)}$  where

$$\Delta_{n,i}^{(\theta)} := \left( \frac{\nabla_{\alpha} F_{n,i}(\alpha) \varphi_n^{(\alpha, \beta)}}{G_{n,i}(\beta)} W_{n,i}^{(\theta)}, \frac{(\nabla_{\beta} \ln G_{n,i}^2(\beta)) \psi_n^{(\beta)}}{2} ((W_{n,i}^{(\theta)})^2 - 1) \right)$$

and  $W_{n,i}^{(\theta)} := (Y_{n,i} - F_{n,i}(\alpha)) / G_{n,i}(\beta)$ .

### 4 Efficient estimation

Cramér-Rao lower bound of the mean square risk is not entirely satisfactory to define the asymptotic efficiency of a sequence of estimators. See e.g. Section I.9 in Ibragimov and Has’minskii (1981), see also Section 1.2 in Höpfner (2014). Then we consider here the asymptotic optimality in the sense of local asymptotic minimax lower bound of the risk for the sequence  $\{\bar{\theta}_n\} := \{\bar{\theta}_n, n > 0\}$  for the estimation of  $\theta$ , that is, of the risk

$$\mathcal{R}_{\theta}(\{\bar{\theta}_n\}) := \lim_{\epsilon \rightarrow 0} \liminf_{n \rightarrow \infty} \sup_{|\theta' - \theta| \leq \epsilon} E_{\theta'} \left[ L \left( \sqrt{T_n}(\bar{\alpha}_n - \alpha'), \sqrt{n}(\bar{\beta}_n - \beta') \right) \right]$$

where  $\bar{\theta}_n$  is any statistic function of the observation  $\{Y_{n,i}, i = 1, \dots, n\}$ . The loss function  $L(\cdot)$  belongs to the set  $\mathcal{L}$  of non-negative Borel functions on  $\mathbb{R}^d$  which are continuous at 0 with  $L(\mathbf{0}_d) = 0$ ,  $L(-x) = L(x)$ , the set  $\{x : L(x) < c\}$  is a convex set for any  $c > 0$ , and we also assume that the function  $L(\cdot) \in \mathcal{L}$  admits a polynomial majorant.

Clearly all functions  $L(\theta) = |\theta|^r$ ,  $r > 0$ , as well as  $L(\theta) = 1_{\{|\theta| > r\}}$ ,  $r > 0$ , belong to  $\mathcal{L}$ . (Here  $1_{\{x > r\}}$  denotes the indicator function of  $(r, \infty)$ .)

Since the model of observation is locally asymptotically normal then the local asymptotic minimax risk  $\mathcal{R}_{\theta}(\{\bar{\theta}_n\})$  for any sequence  $\{\theta_n = (\bar{\alpha}_n, \bar{\beta}_n)\}$  of estimators of  $\theta = (\alpha, \beta)$  admits a lower bound for any loss function  $L \in \mathcal{L}$ . More precisely

$$\mathcal{R}_{\theta}(\{\bar{\theta}_n\}) \geq E[L(\xi^{(\theta)})] \tag{11}$$

where  $\xi^{(\theta)}$  is a random  $d$ -dimensional vector whose distribution is zero-mean Gaussian with  $d \times d$ -matrix variance equal to  $(\mathbf{J}^{(\theta)})^{-1}$  (see Le Cam 1969 and Hájek 1972; see also Ibragimov and Has’minskii 1981).

#### 4.1 Maximum likelihood estimator

Recall that the maximum likelihood estimator  $\widehat{\theta}_n$  is any statistics defined from the observation such that

$$\widehat{\theta}_n \in \arg \sup_{\theta \in \Theta} Z_n(\theta).$$

In the next theorem we establish that  $\widehat{\theta}_n$  is an efficient estimator of  $\theta$  in the sense that its asymptotic minimax risk  $\mathcal{R}_{\theta}(\{\widehat{\theta}_n\})$  is equal to the lower bound  $E[L(\xi^{(\theta)})]$ .



**Theorem 1** *Let  $\Theta = \mathbf{A} \times \mathbf{B}$  be open, convex and bounded. Assume that conditions A1–A4 are fulfilled. Then the maximum likelihood estimator  $\widehat{\boldsymbol{\theta}}_n = (\widehat{\boldsymbol{\alpha}}_n, \widehat{\boldsymbol{\beta}}_n)$  of  $\boldsymbol{\theta} = (\boldsymbol{\alpha}, \boldsymbol{\beta})$  is consistent and asymptotically normal uniformly with respect to  $\boldsymbol{\theta}$  varying in any compact subset  $\mathcal{K}$  of  $\Theta = \mathbf{A} \times \mathbf{B}$ ,*

$$\lim_{n \rightarrow \infty} \mathcal{L}_{\boldsymbol{\theta}} \left[ \left( \sqrt{T_n}(\widehat{\boldsymbol{\alpha}}_n - \boldsymbol{\alpha}), \sqrt{n}(\widehat{\boldsymbol{\beta}}_n - \boldsymbol{\beta}) \right) \right] = \mathcal{N}_d(\mathbf{0}_d, (\mathbf{J}^{(\boldsymbol{\theta})})^{-1})$$

setting  $\mathbf{J}^{(\boldsymbol{\theta})} = \text{diag}[\mathbf{J}_a^{(\boldsymbol{\theta})}, \mathbf{J}_b^{(\boldsymbol{\beta})}]$ . Moreover it is locally asymptotically minimax at any  $\boldsymbol{\theta} \in \Theta$  for any loss function  $L(\cdot) \in \mathcal{L}$ , in the sense that inequality (11) becomes an equality for  $\widehat{\boldsymbol{\theta}}_n = \widehat{\boldsymbol{\theta}}_n$ .

Accordingly, under assumptions A1', A2', A3 and A4, the conclusions of the theorem are also valid for the integrated signal-plus-noise model (1) observed at discrete instants as presented in Sect. 2.2.

**Proof** To prove this theorem, we state that in our framework the conditions in Theorem 1.1 and Corollary 1.1 from Chapter III in Ibragimov and Has'minskii (1981) are fulfilled. Then we are going to establish the following properties.

- N1 The family  $\{\mathbf{P}_{(n)}^{\boldsymbol{\theta}}, \boldsymbol{\theta} \in \Theta\}$  is uniformly LAN in any compact subset of  $\Theta$ .
- N2 For every  $\boldsymbol{\theta} \in \Theta$ , the  $d \times d$ -matrix  $\boldsymbol{\Phi}_n^{(\boldsymbol{\theta})}$  is positive definite,  $d = a + b$ , and there exists a continuous  $d \times d$ -matrix valued function  $(\boldsymbol{\theta}, \boldsymbol{\theta}') \mapsto B(\boldsymbol{\theta}, \boldsymbol{\theta}')$  such that for every compact subset  $\mathcal{K}$  of  $\Theta$

$$\lim_{n \rightarrow \infty} |\boldsymbol{\Phi}_n^{(\boldsymbol{\theta})}| = 0 \quad \text{and} \quad \lim_{n \rightarrow \infty} (\boldsymbol{\Phi}_n^{(\boldsymbol{\theta})})^{-1} \boldsymbol{\Phi}_n^{(\boldsymbol{\theta}')} = B(\boldsymbol{\theta}, \boldsymbol{\theta}')$$

where the convergences are uniform with respect to  $\boldsymbol{\theta}$  and  $\boldsymbol{\theta}'$  varying in  $\mathcal{K}$ .

- N3 For every compact subset  $\mathcal{K}$  of  $\Theta$ , there exist  $s > a + b$ ,  $p \geq 1$ ,  $B = B(\mathcal{K}) > 0$  and  $q = q(\mathcal{K}) \in \mathbb{R}$ , such that

$$\sup_{\boldsymbol{\theta} \in \mathcal{K}} \sup_{\mathbf{w}_1, \mathbf{w}_2 \in \mathcal{W}_{\boldsymbol{\theta}, r, n}} |\mathbf{w}_1 - \mathbf{w}_2|^{-s} \rho_p^p \left( \mathbf{P}_n^{(\boldsymbol{\theta} + \mathbf{w}_1 \boldsymbol{\Phi}_n^{(\boldsymbol{\theta})})}, \mathbf{P}_n^{(\boldsymbol{\theta} + \mathbf{w}_2 \boldsymbol{\Phi}_n^{(\boldsymbol{\theta})})} \right) < B(1 + r^q)$$

for any  $r > 0$ . Here  $\mathcal{W}_{\boldsymbol{\theta}, r, n} := \{\mathbf{w} \in \mathbb{R}^d : |\mathbf{w}| < r \text{ and } \boldsymbol{\theta} + \mathbf{w} \boldsymbol{\Phi}_n^{(\boldsymbol{\theta})} \in \Theta\}$ .

- N4 For any compact subset  $\mathcal{K}$  of  $\Theta$ , and for every  $N > 0$ , there exists  $n_1 = n_1(N, \mathcal{K}) > 0$  such that

$$\sup_{\boldsymbol{\theta} \in \mathcal{K}} \sup_{n > n_1} \sup_{\mathbf{w} \in \mathcal{W}_{\boldsymbol{\theta}, n}} |\mathbf{w}|^N H \left( \mathbf{P}_n^{(\boldsymbol{\theta})}, \mathbf{P}_n^{(\boldsymbol{\theta} + \mathbf{w} \boldsymbol{\Phi}_n^{(\boldsymbol{\theta})})} \right) < \infty.$$

Here  $\rho_p(\cdot, \cdot)$  is the Hellinger  $p$ -distance, and  $H(\cdot, \cdot)$  is the Hellinger integrale. Their definitions are recalled in ‘‘Appendix A’’. Recall that  $\mathcal{W}_{\boldsymbol{\theta}, n} := \{\mathbf{w} \in \mathbb{R}^d : \boldsymbol{\theta} + \mathbf{w} \boldsymbol{\Phi}_n^{(\boldsymbol{\theta})} \in \Theta\}$ .

- (1) In Proposition 1 we have stated that the family  $\{\mathbf{P}_{\boldsymbol{\theta}}^{(n)}, \boldsymbol{\theta} \in \Theta\}$  is uniformly LAN in any compact subset of  $\Theta$  (condition N1). In addition, as  $\boldsymbol{\Phi}_n^{(\boldsymbol{\theta})} = \text{diag}[\boldsymbol{\varphi}_n^{(\boldsymbol{\theta})}, \boldsymbol{\psi}_n^{(\boldsymbol{\beta})}]$ ,  $\boldsymbol{\varphi}_n^{(\boldsymbol{\theta})} := (T_n \mathbf{J}_a^{(\boldsymbol{\theta})})^{\frac{-1}{2}}$  and  $\boldsymbol{\psi}_n^{(\boldsymbol{\beta})} := (n \mathbf{J}_b^{(\boldsymbol{\beta})})^{\frac{-1}{2}}$ , from the continuity of  $\boldsymbol{\theta} \mapsto \mathbf{J}_a^{(\boldsymbol{\theta})}$  and  $\boldsymbol{\beta} \mapsto \mathbf{J}_b^{(\boldsymbol{\beta})}$  we deduce that condition N2 is fulfilled with

$$B(\boldsymbol{\theta}, \boldsymbol{\theta}') = \text{diag} \left[ \left( \mathbf{J}_a^{(\boldsymbol{\theta})} \right)^{\frac{1}{2}} \left( \mathbf{J}_a^{(\boldsymbol{\alpha}', \boldsymbol{\beta}')} \right)^{\frac{-1}{2}}, \left( \mathbf{J}_b^{(\boldsymbol{\beta})} \right)^{\frac{1}{2}} \left( \mathbf{J}_b^{(\boldsymbol{\beta}')} \right)^{\frac{-1}{2}} \right].$$

- (2) Now we check condition N3. Let be fixed the compact subset  $\mathcal{K} \subset \Theta$ , the integer  $p \geq 2$  and  $r > 0$ . Denote  $\theta_n^{(i)} = (\alpha_n^{(i)}, \beta_n^{(i)}) := \theta + \mathbf{w}_i \Phi_n^{(\theta)}$ , and  $\mathbf{w} = (\mathbf{u}, \mathbf{v}) := \mathbf{w}_2 - \mathbf{w}_1$ . Thanks to the independence of the Gaussian random variables  $Y_{n,i}, i = 1, \dots, n$ , we have that (see inequality (18))

$$\rho_p^p(\mathbb{P}_n^{(\theta_n^{(1)})}, \mathbb{P}_n^{(\theta_n^{(2)})}) \leq C_p \rho_2^p(\mathbb{P}_n^{(\theta_n^{(1)})}, \mathbb{P}_n^{(\theta_n^{(2)})}).$$

From inequalities (19) as well as (9) and (10) we obtain that

$$\rho_2^2(\mathbb{P}_n^{(\theta_n^{(1)})}, \mathbb{P}_n^{(\theta_n^{(2)})}) \leq C \left( |\mathbf{u}|^2 |(J_a^{(\theta)})^{-1/2}|^2 + |\mathbf{v}|^2 |(J_b^{(\beta)})^{-1/2}|^2 \right)$$

for some  $C > 0$ . We know that the functions  $\theta \mapsto |(J_a^{(\theta)})^{-1/2}|$  and  $\theta \mapsto |(J_b^{(\beta)})^{-1/2}|$  are continuous (remark 2 Sect. 2.1), they are bounded on the compact subset  $\mathcal{K} \subset \Theta$ . Hence there exists a constant  $b(\mathcal{K}) > 0$  such that

$$\rho_2^2(\mathbb{P}_n^{(\theta_n^{(1)})}, \mathbb{P}_n^{(\theta_n^{(2)})}) \leq b(\mathcal{K}) |\mathbf{w}|^2$$

and condition N3 is fulfilled with  $p$  even integer,  $s = p > \min\{3, a + b\}$ ,  $q = 0$ , and  $B = C_p b(\mathcal{K})^{p/2}$ .

- (3) It remains to prove that condition N4 is fulfilled. For that purpose we study the Hellinger integral  $H(\mathbb{P}_n^{(\theta)}, \mathbb{P}_n^{(\theta_n)})$  where  $\theta = (\alpha, \beta)$  and  $\theta_n = (\alpha_n, \beta_n) := \theta + \mathbf{w} \Phi_n^{(\theta)}$ . The independence of the Gaussian random variables  $Y_{n,i}, i = 1, \dots, n$ , implies the equality:

$$H(\mathbb{P}_n^{(\theta)}, \mathbb{P}_n^{(\theta_n)}) = \prod_{i=1}^n \left( 1 + \frac{y_{n,i}^2}{2(1 + y_{n,i})} \right)^{-1/4} \exp \left( \frac{-x_{n,i}^2}{4(2 + y_{n,i})} \right)$$

where  $x_{n,i}$  and  $y_{n,i}$  are defined according to relations (7) and (8) with evident modifications.

- (i) Assume that  $\theta \in \mathcal{K}$  and  $|\mathbf{w} \Phi_n^{(\theta)}| \leq \gamma_1 \gamma_2^{-1} / 2$ . Then equality (8) and assumption (A2) imply that  $|y_{n,i}| \leq 1/2$  for every  $i$ . Thanks to inequality  $(1 + a)^{-1/4} \leq e^{-a/8}$  for  $0 \leq a \leq 1/2$ , we deduce that

$$\begin{aligned} H(\mathbb{P}_n^{(\theta)}, \mathbb{P}_n^{(\theta_n)}) &\leq \prod_{i=1}^n \exp \left( \frac{-y_{n,i}^2}{16(1 + y_{n,i})} \right) \exp \left( \frac{-x_{n,i}^2}{4(2 + y_{n,i})} \right) \\ &\leq \exp \left( \frac{-1}{24} \sum_{i=1}^n (x_{n,i}^2 + y_{n,i}^2) \right). \end{aligned}$$

First, Taylor expansion formula applied to equalities (7) and (8) with assumptions (A1) and (A2), entail that for any  $\epsilon > 0$ , there exists  $\nu_\epsilon > 0$  such that for  $|\mathbf{w} \Phi_n^{(\theta)}| \leq \nu_\epsilon$  we have the inequality:

$$\left| x_{n,i}^2 - \left( \frac{\mathbf{u} \varphi_n^{(\theta)} \nabla_{\alpha}^* F_{n,i}(\alpha)}{G_{n,i}(\beta)} \right)^2 \right| \leq T_n^{-1} h_{n,i} |\mathbf{u}|^2 |(J_a^{(\theta)})^{-1/2}|^2 \epsilon.$$

Second, by assumption (A3), for every  $\epsilon > 0$ , there exists  $n_\epsilon > 0$  such that for  $n \geq n_\epsilon$ ,

$$\left| \sum_{i=1}^n \left( \frac{\mathbf{u} \varphi_n^{(\theta)} \nabla_{\alpha}^* F_{n,i}(\alpha)}{G_{n,i}(\beta)} \right)^2 - |\mathbf{u}|^2 \right| \leq |\mathbf{u}|^2 |(J_a^{(\theta)})^{-1/2}|^2 \epsilon.$$

In the same way we can study  $y_{n,i}^2$ . Then, by the continuity of the positive functions  $\theta \mapsto |(\mathbf{J}_a^{(\theta)})^{-1/2}|$  and  $\theta \mapsto |(\mathbf{J}_b^{(\beta)})^{-1/2}|$ , we deduce the existence of  $n_1 = n_1(\mathcal{K}) > 0$  and  $\nu > 0$  such that for every  $\theta \in \mathcal{K}, n \geq n_1$  and  $\mathbf{w} \in \mathcal{W}_{\theta,n}$  with  $|\mathbf{w}\Phi_n^{(\theta)}| \leq \nu$  we have:

$$H(\mathbf{P}_n^{(\theta)}, \mathbf{P}_n^{(\theta_n)}) \leq \exp\left(\frac{-1}{24} \sum_{i=1}^n (x_{n,i}^2 + y_{n,i}^2)\right) \leq e^{-|\mathbf{w}|^2/48}.$$

(ii) Assumption (A2) implies that

$$|x_{n,i}|^2 \geq \gamma_1^{-1} h_{n,i}^{-1} (F_{n,i}(\boldsymbol{\alpha}) - F_{n,i}(\boldsymbol{\alpha}_n))^2.$$

Then, from the identifiability assumption A4, for  $\nu > 0$  there exist  $\mu_{v,1} > 0$  and  $n_{v,1} > 0$  such that for  $n > n_{v,1}$ , with  $|\mathbf{u}\varphi_n^{(\theta)}| = |\boldsymbol{\alpha}_n - \boldsymbol{\alpha}| \geq \nu$  we have:

$$\frac{1}{T_n} \sum_{i=1}^n |x_{n,i}|^2 \geq \mu_{v,1}.$$

Since  $|\mathbf{u}\phi_n^{(\theta)}| = |\boldsymbol{\alpha}_n - \boldsymbol{\alpha}| \leq \text{diam}(\mathbf{A})$ , we deduce that

$$\sum_{i=1}^n |x_{n,i}|^2 \geq \frac{T_n \mu_{v,1} |\mathbf{u}\phi_n^{(\theta)}|^2}{\text{diam}(\mathbf{A})^2} \geq \frac{|\mathbf{u}|^2 \mu_{v,1} |(\mathbf{J}_a^{(\theta)})^{\frac{1}{2}}|^2}{\text{diam}(\mathbf{A})^2}.$$

Notice that for the last inequality we have used the relation:  $|\mathbf{u}\phi_n^{(\theta)}|^2 = T_n^{-1} |\mathbf{u}(\mathbf{J}_a^{(\theta)})^{\frac{1}{2}}|^2 \geq T_n^{-1} |\mathbf{u}|^2 |(\mathbf{J}_a^{(\theta)})^{\frac{1}{2}}|^2$ . Next, from the equality  $\ln(1+a) = a \int_0^1 (1+ta)^{-1} dt$  for  $a > -1$ , and thanks to assumption (A2) we easily establish that

$$\begin{aligned} \ln\left(1 + \frac{y_{n,i}^2}{1 + y_{n,i}}\right) &= \int_0^1 \frac{(G_{n,i}^2(\boldsymbol{\beta}_n) - G_{n,i}^2(\boldsymbol{\beta}))^2 dt}{t G_{n,i}^4(\boldsymbol{\beta}_n) + t G_{n,i}^4(\boldsymbol{\beta}) + (1-t) G_{n,i}^2(\boldsymbol{\beta}_n) G_{n,i}^2(\boldsymbol{\beta})} \\ &\geq \frac{(G_{n,i}^2(\boldsymbol{\beta}_n) - G_{n,i}^2(\boldsymbol{\beta}_n^{(1)}))^2}{2 G_{n,i}^2(\boldsymbol{\beta}_n^{(2)}) G_{n,i}^2(\boldsymbol{\beta}_n^{(1)})} \geq \frac{(G_{n,i}^2(\boldsymbol{\beta}_n) - G_{n,i}^2(\boldsymbol{\beta}))^2}{2 \gamma_1^2 h_{n,i}^2}. \end{aligned}$$

The identifiability assumption (A4) entails that for  $\nu > 0$  there exist  $\mu_{v,2} > 0$  and  $n_{v,2} > 0$  such that for  $n > n_{v,2}$  and  $|\mathbf{v}\psi_n^{(\beta)}| = |\boldsymbol{\beta}_n - \boldsymbol{\beta}| \geq \nu$  we have:

$$\frac{1}{n} \sum_{i=1}^n \ln\left(1 + \frac{y_{n,i}^2}{1 + y_{n,i}}\right) \geq \mu_{v,2}.$$

As previously we deduce that

$$\sum_{i=1}^n \ln\left(1 + \frac{y_{n,i}^2}{1 + y_{n,i}}\right) \geq \frac{|\mathbf{v}|^2 \mu_{v,2} |(\mathbf{J}_b^{(\beta)})^{\frac{1}{2}}|^2}{\text{diam}(\mathbf{B})^2}.$$

Then, by the continuity of the positive functions  $\theta \mapsto |(\mathbf{J}_a^{(\theta)})^{-1/2}|$  and  $\theta \mapsto |(\mathbf{J}_b^{(\beta)})^{-1/2}|$ , we deduce the existence of  $\mu_\nu = \mu(\nu, \mathcal{K}) > 0$  and  $n_2 = n_2(\nu, \mathcal{K}) > 0$  such that for every  $\theta \in \mathcal{K}, n \geq n_2$  and  $\mathbf{w} \in \mathcal{W}_{\theta,n}$  and  $|\mathbf{w}\Phi_n^{(\theta)}| \geq \nu$  we have

$$H(\mathbf{P}_n^{(\theta)}, \mathbf{P}_n^{(\theta_n)}) \leq e^{-|\mathbf{w}|^2 \mu_\nu}.$$

(iii) From previous parts (i) and (ii), there exists  $n_o > 0$  and  $\mu > 0$  such that for every  $\theta \in \mathcal{K}$ ,  $n \geq n_o$  and  $w \in \mathcal{W}_{\theta,n}$ , we have:

$$H(\mathbf{P}_n^{(\theta)}, \mathbf{P}_n^{(\theta_n)}) \leq e^{-|w|^2\mu}.$$

Hence condition N4 is satisfied. This achieves the proof of the theorem. □

### 4.2 Bayesian estimator

In this section, the unknown parameter  $\theta = (\alpha, \beta)$  is supposed to be a random vector with known prior density distribution  $\pi(\cdot)$  on the parameter set  $\Theta = \mathbf{A} \times \mathbf{B}$ . We are going to study the property of the Bayesian estimator  $\tilde{\theta}_n$  that minimizes the mean Bayesian risk defined as

$$R_n(\tilde{\theta}_n) := \int_{\Theta} E_{\theta} [l((\tilde{\theta}_n - \theta)\delta_n)] \pi(\theta) d\theta,$$

where for simplicity of presentation the loss function  $l(\cdot)$  is equal to  $l(\theta) = |\theta|^r$  for some  $r > 0$  (see e.g. Ibragimov and Has'minskii 1981). Here  $\delta_n = \text{diag}[\sqrt{T_n}\mathbb{I}_{a \times a}, \sqrt{n}\mathbb{I}_{b \times b}]$ . Then, we state that

**Theorem 2** *Let  $\Theta = \mathbf{A} \times \mathbf{B}$  be open convex and bounded. Assume that the conditions of Theorem 1 are fulfilled. Assume that the prior density  $\pi(\theta)$  is continuous and positive on  $\Theta$  and that the loss function  $l(\theta) = |\theta|^a$  for some  $a > 0$ . Then, uniformly with respect to  $\tilde{\theta} = (\tilde{\alpha}, \tilde{\beta})$  varying in any compact subset  $\mathcal{K}$  of  $\Theta$ , the corresponding Bayesian estimator  $\tilde{\theta}_n = (\tilde{\alpha}_n, \tilde{\beta}_n)$  converges in probability and is asymptotically normal:*

$$\lim_{n \rightarrow \infty} \mathcal{L}_{\theta} \left[ \left( \sqrt{T_n}(\tilde{\alpha}_n - \alpha), \sqrt{n}(\tilde{\beta}_n - \beta) \right) \right] = \mathcal{N}_d(\mathbf{0}_d, (\mathbf{J}^{(\theta)})^{-1}).$$

Moreover, the Bayesian estimator  $\tilde{\theta}_n$  is locally asymptotically minimax at any  $\theta \in \Theta$  for any loss function  $L(\cdot) \in \mathcal{L}$ , in the sense that inequality (11) becomes an equality for  $\tilde{\theta}_n = \tilde{\theta}_n$ .

**Proof** This is a direct consequence of Theorem 2.1 in Chapter III of Ibragimov and Has'minskii (1981) and the proof of Theorem 1. □

### 5 Linear parameter models

Consider the specific case where  $f(\alpha, t) = \alpha f(t)^* = \alpha_1 f_1(t) + \dots + \alpha_a f_a(t)$ , and  $\sigma^2(\beta, t) = \beta \sigma^2(t)$ .  $\theta = (\alpha, \beta)$  and  $\Theta = \mathbf{A} \times \mathbf{B} \subset \mathbb{R}^a \times \mathbb{R}^+$ :

$$dX_t = \alpha f(t)^* dt + \sqrt{\beta} \sigma(t) dW_t.$$

In this case  $a \geq 1$ ,  $b = 1$ . We assume that

$$\sup_t |f_j(t)| < \infty, \quad 0 < \inf_t \sigma^2(t) \leq \sup_t \sigma^2(t) < \infty$$

and there exists a positive definite  $a \times a$ -matrix  $\mathbf{J}_a$  which fulfils:

$$\mathbf{J}_a = \lim_{n \rightarrow \infty} \frac{1}{T_n} \sum_{i=1}^n \frac{\mathbf{F}_{n,i}^* \mathbf{F}_{n,i}}{G_{n,i}^2}.$$

Here  $f(\cdot) := (f_1(\cdot), \dots, f_a(\cdot))$ ,  $\mathbf{F}_{n,i} := \int_{t_{n,i-1}}^{t_{n,i}} f(t) dt$  and  $G_{n,i}^2 := \int_{t_{n,i-1}}^{t_{n,i}} \sigma^2(t) dt$ .

Using the notation of Sect. 2 we have that  $F_{n,i}(\alpha) = \alpha F_{n,i}$  and  $G_{n,i}^2(\beta) = \beta G_{n,i}^2$ ,  $J_a^{(\theta)} = \beta^{-2} J_a$ ,  $J_b^{(\beta)} = \beta^{-2}/2$ ,

$$\mu_a(\alpha, \alpha') = \liminf_{n \rightarrow \infty} (\alpha - \alpha') \left( \frac{1}{T_n} \sum_{i=1}^n \frac{F_{n,i}^* F_{n,i}}{t_{n,i} - t_{n,i-1}} \right) (\alpha - \alpha')^*,$$

and

$$\mu_b(\beta, \beta') = \left( \liminf_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n \frac{G_{n,i}^4}{(t_{n,i} - t_{n,i-1})^2} \right) (\beta - \beta')^2.$$

Then Theorems 1 and 2 can be applied in this case. Furthermore the maximum likelihood estimator  $\hat{\theta}_n = (\hat{\alpha}_n, \hat{\beta}_n)$  has an explicit expression

$$\hat{\alpha}_n = \left( \sum_{i=1}^n \frac{F_{n,i}^* F_{n,i}}{G_{n,i}^2} \right)^{-1} \left( \sum_{i=1}^n \frac{Y_{n,i} F_{n,i}}{G_{n,i}^2} \right), \quad \hat{\beta}_n = \frac{1}{n} \sum_{i=1}^n \frac{(Y_{n,i} - \hat{\alpha}_n F_{n,i}^*)^2}{G_{n,i}^2},$$

and we readily obtain that  $\hat{\theta}_n$  converges in norm  $L^r$  to  $\theta$  as  $n \rightarrow \infty$  for any  $r \geq 1$ .

**Acknowledgements** The authors express their gratitude to an anonymous referee for a thorough reading of the manuscript and for the valuable suggestion to deduce the main results using Hellinger-type distances. The authors would like also to thank Pr. Yury Kutoyants for helpful discussions.

## 6 Appendix A

First we recall the definitions of some notions that are needed in the paper. For more information the readers are referred to Jacod and Shiryaev (1987) and Dzhaparidze and Valkeila (1990).

Let  $P$  and  $\tilde{P}$  be two probability measures on a measurable space  $(\Omega, \mathcal{F})$ , and let  $Q$  be any measure dominating  $P$  and  $\tilde{P}$ . Put  $Z = dP/dQ$  and  $\tilde{Z} = d\tilde{P}/dQ$ . The Hellinger  $p$ -distance  $\rho_p(P, \tilde{P})$  of order  $p \geq 1$  between  $P$  and  $\tilde{P}$ , is defined by  $\rho_p^p(P, \tilde{P}) := E_Q[|Z|^{1/p} - \tilde{Z}|^{1/p}|^p]$ , and  $\rho_2(P, \tilde{P})$  is called simply Hellinger distance. The Hellinger integral of order  $\alpha \in (0, 1)$  of  $P$  and  $\tilde{P}$  is equal to  $H(\alpha; P, \tilde{P}) := E_Q[Z^\alpha \tilde{Z}^{1-\alpha}]$ . When  $\alpha = 1/2$ ,  $H(P, \tilde{P}) := H(1/2; P, \tilde{P})$  is the Hellinger integral of  $P$  and  $\tilde{P}$ , and  $H(P, \tilde{P}) = 1 - \frac{1}{2} \rho_2^2(P, \tilde{P})$ . Finally introduce the symmetrized Kullback-Leibler divergence by  $J(P, \tilde{P}) := E_Q[Z \ln(Z/\tilde{Z}) + \tilde{Z} \ln(\tilde{Z}/Z)]$  if  $P \sim \tilde{P}$ , and  $J(P, \tilde{P}) = \infty$  otherwise. It is easy to verify that all the previous definitions do not depend on the choice of  $Q$ .

Next, we establish some technical results for non-centered normal distributions. These results have been stated in Lemma 4.1 and 4.2 of Gushchin and K uchler (2003) in the case of zero-mean normal distributions.

First we consider two normal distributions on  $\mathbb{R}$ .

**Lemma 1** Consider two normal distributions  $q = \mathcal{N}(\mu, \sigma^2)$  and  $\tilde{q} = \mathcal{N}(\tilde{\mu}, \tilde{\sigma}^2)$  on  $\mathbb{R}$ . Let  $q_o = \mathcal{N}(0, 1)$  and  $\tilde{q}' = \mathcal{N}(x, 1 + y)$  where  $x = (\tilde{\mu} - \mu)/\sigma$  and  $y = (\tilde{\sigma}^2 - \sigma^2)/\sigma^2$ . Then,

$$H(q, \tilde{q}) = H(q_o, \tilde{q}') = \left( 1 + \frac{y^2}{4(1+y)} \right)^{-1/4} \exp\left( \frac{-x^2}{4(2+y)} \right), \tag{12}$$

$$J(q, \tilde{q}) = J(q_o, \tilde{q}') = \frac{y^2}{2(1+y)} + \frac{x^2(2+y)}{2(1+y)} \tag{13}$$

and for every  $p \geq 1$  there exists  $C_p > 0$ , such that

$$\rho_p^p(\varrho, \tilde{\varrho}) \leq C_p(|x|^p + |y|^p). \tag{14}$$

Moreover,

$$\rho_2^2(\varrho, \tilde{\varrho}) = \frac{x^2}{4} + \frac{y^2}{8} + o(x^2 + y^2) \tag{15}$$

as  $|x| + |y| \rightarrow 0$ .

**Proof** The lemma can be established with elementary computations. The proof is left to the reader (see Gushchin and K uchler 2003). Notice that  $1 + y = \tilde{\sigma}_2/\sigma_2 > 0$ .  $\square$

When  $\Omega = \mathbb{R}^n$ ,  $P_n = \varrho_1 \times \dots \times \varrho_n$  and  $\tilde{P}_n = \tilde{\varrho}_1 \times \dots \times \tilde{\varrho}_n$  where  $\varrho_i$  and  $\tilde{\varrho}_i, i = 1, \dots, n$  are probability measures on  $\mathbb{R}$ , it is easy to verify

$$H(P_n, \tilde{P}_n) = \prod_{i=1}^n H(\varrho_i, \tilde{\varrho}_i) \quad \text{and} \quad J(P_n, \tilde{P}_n) = \sum_{i=1}^n J(\varrho_i, \tilde{\varrho}_i). \tag{16}$$

**Lemma 2** If  $\varrho_i$  and  $\tilde{\varrho}_i$  are any normal distributions, then

$$\exp\left(\frac{-1}{8}J(P_n, \tilde{P}_n)\right) \leq H(P_n, \tilde{P}_n) \leq \left(1 + \frac{1}{2}J(P_n, \tilde{P}_n)\right)^{-1/4}. \tag{17}$$

For every even integer  $p \geq 2$  there exists  $C_p > 0$  depending only of  $p$  such that

$$\rho_p^p(P_n, \tilde{P}_n) \leq C_p \rho_2^p(P_n, \tilde{P}_n). \tag{18}$$

Moreover, there exists  $C > 0$  such that if  $\varrho_i = \mathcal{N}(\mu_i, \sigma_i^2)$ ,  $\tilde{\varrho}_i = \mathcal{N}(\tilde{\mu}_i, \tilde{\sigma}_i^2)$  and where  $x_i = (\tilde{\mu}_i - \mu_i)/\sigma_i$  and  $y_i = (\tilde{\sigma}_i^2 - \sigma_i^2)/\sigma_i^2$ , we have:

$$\rho_2^2(P_n, \tilde{P}_n) \leq C \sum_{i=1}^n (x_i^2 + y_i^2). \tag{19}$$

**Proof** (i) Inequalities (17) are direct consequence of Lemma 1 as well as the inequalities,  $1 + \sum_i a_i \leq \prod_i (1 + a_i) \leq \exp(\sum_i a_i)$  for  $a_i \geq 0, i = 1, \dots$

(ii) To prove inequality (18) we adapt the proof given by Gushchin and K uchler (2003) when the normal distributions have zero means.

First, let  $p > 2$ , integer and even. From relation (4.1) in Gushchin and K uchler (2003) (see also Corollary 3.1 in Dzhaparidze and Valkeila 1990), we know that

$$\rho_p^p(P_n, \tilde{P}_n) \leq C_p \left( \left( \frac{1}{2} \sum_{i=1}^n \rho_2^2(\varrho_i, \tilde{\varrho}_i) \right)^{p/2} + \sum_{i=1}^n \rho_p^p(\varrho_i, \tilde{\varrho}_i) \right)$$

for some  $C_p > 0$ . Then thanks to equalities (12) and (13), we can consider that  $\varrho_i = \mathcal{N}(0, 1)$  and  $\tilde{\varrho}_i = \mathcal{N}(x_i, y_i)$ ,  $y_i > -1$ , without loss of generality. Then inequality (14) and the fact that  $p \geq 2$  imply that

$$\rho_p^p(P_n, \tilde{P}_n) \leq C'_p \left( \sum_{i=1}^n (x_i^2 + y_i^2) \right)^{p/2}. \tag{20}$$

for some  $C'_p > 0$ . Hence inequality (19) is proved. From relation (13) applied to  $J(\varrho_i, \tilde{\varrho}_i)$ , we also deduce that

$$\rho_p^p(P_n, \tilde{P}_n) \leq C'_p (1 + \sup_i y_i)^{p/2} (J(P_n, \tilde{P}_n))^{p/2}.$$

If  $\rho_2^2(P_n, \tilde{P}_n) \geq 1$  we trivially have that  $\rho_p^p(P_n, \tilde{P}_n) \leq 2 \leq 2\rho_2^p(P_n, \tilde{P}_n)$ . Assume now that  $\rho_2^2(P_n, \tilde{P}_n) \leq 1$ , that is,  $H(P, \tilde{P}) \geq \frac{1}{2}$ . In this case, relation (17) entails that  $J(P_n, \tilde{P}_n) \leq 30\rho_2^2(P_n, \tilde{P}_n)$ . Moreover, the Hellinger integrals being bounded by 1, from equalities (16) and (12) we obtain that

$$2^4 \geq H^{-4}(P_n, \tilde{P}_n) \geq H^{-4}(\varrho_i, \tilde{\varrho}_i) \geq 1 + \frac{y_i^2}{4(1 + y_i)}$$

and  $y_i \leq y_o := 30 + 8\sqrt{15}$ ,  $i = 1, \dots, n$ . Then inequality (20) entails that

$$\rho_p^p(P, \tilde{P}) \leq 60C'_p(1 + y_o)^{p/2}\rho_2^p(P, \tilde{P}).$$

Hence inequality (18) is proved. □

### 7 Appendix B

To illustrate assumptions A3 and A4 we state the following results the proofs of which are left to the reader. (See also the particular case of the linear parameter model in Sect. 5).

*Almost periodic functions* First recall that a function  $t \mapsto \phi(\chi, t)$  is almost periodic in  $\mathbb{R}$  uniformly with respect to  $\chi$  varying in a set  $X$  when for every  $\epsilon > 0$ , there exists  $l_\epsilon > 0$  such that for any  $a \in \mathbb{R}$  there is  $\rho \in [a, a + l_\epsilon]$  for which

$$\sup_{\chi, t} |\phi(\chi, t + \rho) - \phi(\chi, t)| \leq \epsilon.$$

See e.g chapters II and IV in Corduneanu (1968). As an example, let  $k$  be a positive number and let  $\lambda_1, \dots, \lambda_k$  be  $k$  distinct real numbers. Then the function  $\phi(\chi, t) = \chi_1 \cos(\lambda_1 t) + \dots + \chi_k \cos(\lambda_k t)$  is almost periodic in  $t$  uniformly with respect to  $\chi = (\chi_1, \dots, \chi_k)$  varying in any compact subset  $X$  of  $\mathbb{R}^k$ .

Now assume that the function  $t \mapsto (f(\alpha, t), \nabla_\alpha f(\alpha, t))$  is almost periodic in  $\mathbb{R}$  uniformly with respect to  $\alpha$  varying in  $A$ , the function  $t \mapsto (\sigma^2(\beta, t), \nabla_\beta \sigma^2(\beta, t))$  is almost periodic in  $\mathbb{R}$  uniformly with respect to  $\beta$  varying in  $B$ ,  $\inf_{\beta, t} \sigma^2(\beta, t) > 0$  and  $h_n \rightarrow 0$  as  $n \rightarrow \infty$ . Then  $J_a^{(\theta)}$  and  $\mu_a(\alpha, \alpha')$  exists and

$$J_a^{(\theta)} = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T \frac{\nabla_\alpha^* f(\alpha, t) \nabla_\alpha f(\alpha, t)}{\sigma^2(\beta, t)} dt,$$

$$\mu_a(\alpha, \alpha') = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T (f(\alpha, t) - f(\alpha', t))^2 dt.$$

the convergences being uniform with respect to  $\alpha$  and  $\alpha'$  varying in  $A$  and with respect to  $\beta$  varying in  $B$ . Notice that  $\mu_a(\alpha, \alpha') > 0$  as soon as there exists  $t$  such that  $f(\alpha, t) \neq f(\alpha', t)$ . If in addition,  $t_{n,i} - t_{n,i-1} = h_n$  then  $J_b^{(\beta)}$  and  $\mu_b(\beta, \beta')$  exist and

$$J_b^{(\beta)} = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_0^T \nabla_\beta^* \ln \sigma^2(\beta, t) \nabla_\beta \ln \sigma^2(\beta, t) dt,$$

$$\mu_b(\beta, \beta') = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T (\sigma^2(\beta, t) - \sigma^2(\beta', t))^2 dt.$$

the convergences being uniform with respect to  $\beta$  and  $\beta'$  varying in  $B$ .

*Periodic functions* When the functions  $f(\alpha, t)$  and  $\sigma^2(\beta, t)$  are periodic in  $t$  with the same period  $P > 0$ , we obtain expressions for  $J_a^{(\theta)}$ ,  $\mu_a(\alpha, \alpha')$ ,  $J_b^{(\beta)}$  and  $\mu_b(\beta, \beta')$ . For continuous functions when  $h_n \rightarrow 0$  we have the relations:

$$J_a^{(\theta)} = \frac{1}{P} \int_0^P \frac{\nabla_\alpha^* f(\alpha, t) \nabla_\alpha f(\alpha, t)}{\sigma^2(\beta, t)} dt,$$

and

$$\mu_a(\alpha, \alpha') = \frac{1}{P} \int_0^P (f(\alpha, t) - f(\alpha', t))^2 dt.$$

If in addition,  $t_{n,i} - t_{n,i-1} = h_n$  then

$$J_b^{(\beta)} = \frac{1}{2P} \int_0^P \nabla_\beta^* \ln \sigma^2(\beta, t) \nabla_\beta \ln \sigma^2(\beta, t) dt,$$

and

$$\mu_b(\beta, \beta') = \frac{1}{P} \int_0^P (\sigma^2(\beta, t) - \sigma^2(\beta', t))^2 dt.$$

Now we no longer assume that the delays between two observations tend to 0, but we assume that the sampling scheme has some periodic feature, that is the instants of observation are defined as  $0 \leq t_0 < \dots < t_\nu = P$ , and  $t_{i+j\nu} = t_i + jP$ , for some  $\nu \in \mathbb{N}$ , and for any  $i = 0, \dots, \nu$  and any  $j \in \mathbb{N}$ . Then we obtain that

$$J_a^{(\theta)} = \frac{1}{P} \sum_{i=1}^\nu \frac{\nabla_\alpha^* F_i(\alpha) \nabla_\alpha F_i(\alpha)}{G_i^2(\beta)}$$

and

$$\mu_a(\alpha, \alpha') = \frac{1}{P} \sum_{i=1}^\nu \frac{(F_i(\alpha) - F_i(\alpha'))^2}{t_i - t_{i-1}}$$

where we have omitted the unnecessary index  $n$ . Furthermore

$$J_b^{(\beta)} = \frac{1}{2\nu} \sum_{i=1}^\nu \nabla_\beta^* \ln G_i^2(\alpha) \nabla_\beta \ln G_i^2(\beta)$$

and

$$\mu_b(\beta, \beta') = \frac{1}{\nu} \sum_{i=1}^\nu \frac{(G_i^2(\beta) - G_i^2(\beta'))^2}{(t_i - t_{i-1})^2}.$$

Furthermore when we assume that  $t_i - t_{i-1} = h > 0$  fixed and  $P = \nu h$ ,  $\nu \in \mathbb{N}$ , we obtain that

$$\mu_a(\alpha, \alpha') = \frac{1}{Ph} \sum_{i=1}^\nu (F_i(\alpha) - F_i(\alpha'))^2$$

and

$$\mu_b(\beta, \beta') = \frac{1}{Ph} \sum_{i=1}^\nu (G_i^2(\beta) - G_i^2(\beta'))^2.$$



## References

- Blanke D, Vial C (2014) Global smoothness estimation of a Gaussian process from general sequence designs. *Electron J Stat* 8:1152–1187
- Corduneanu C (1968) Almost periodic functions. Interscience Publishers Wiley, Hoboken
- Dehay D, Dudek AE, El Badaoui M (2017) Bootstrap for almost cyclostationary processes with jitter effect. *Digital Signal Proc* 73:93–105
- Dzhaparidze K, Valkeila E (1990) On the Hellinger type distances for filtered experiments. *Probab Theory Relat Fields* 85:105–117
- Florens-Zmirou D (1989) Approximate discrete-time schemes for statistics of diffusion processes. *Statistics* 20(4):547–557
- Genon-Catalot V (1990) Maximum contrast estimation for diffusion process from discrete observations. *Statistics* 21:99–116
- Gushchin AA, Küchler U (2003) On parametric statistical models for stationary solutions of affine stochastic delay differential equations. *Math Methods Stat* 12(1):31–61
- Gushchin A, Valkeila E (2017) Quadratic approximation for log-likelihood ratio processes. In: Panov Vladimir (ed) *Modern problems of stochastic analysis and statistics, selected contributions in Honor of Valentin Konakov*, Springer Proceedings in Mathematics & Statistics 208, pp 179–215. Springer International Publishing AG, Cham
- Hájek J (1972) Local asymptotic minimax and admissibility in estimation. In: *Proceedings of 6-th Berkeley symposium mathematics statistical probabilities*
- Höpfner R (2014) *Asymptotic statistics with a view to stochastic processes*. De Gruyter, Berlin
- Ibragimov IA, Has'minskii RZ (1981) *Statistical estimation, asymptotic theory*. Springer, New York
- Jacod J, Shiryaev AN (1987) *Limit theorems for stochastic processes*. Springer, Berlin
- Kutoyants Y (1984) Parameter estimation for stochastic processes. In: *Research and exposition in mathematics*, vol 6. Heldermann-Verlag, Berlin
- Le Cam L (1969) *Théorie Asymptotique de la Décision Statistique*. Presses de l'Université de Montréal, Montréal
- Mishra MN, Prakasa Rao BLS (2001) Asymptotic minimax estimation in nonlinear stochastic differential equations from discrete observations. *Stat Decis* 19:121–136
- Sacks J, Ylvisaker D (1968) Designs for regression problems with correlated errors: many parameters. *Ann Math Stat* 39:49–69
- Uchida M, Yoshida N (2012) Adaptive estimation of an ergodic diffusion process based on sampled data. *Stoch Process Appl* 122:2885–2924

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.