GLASNIK MATEMATIČKI
Vol. 36(56)(2001), 139 – 153

A GENERAL THEOREM ON APPROXIMATE MAXIMUM LIKELIHOOD ESTIMATION

Miljenko Huzak
University of Zagreb,Croatia

Abstract. In this paper a version of the general theorem on approximate maximum likelihood estimation is proved. We assume that there exists a log-likelihood function $L(\vartheta)$ and a sequence $(L_n(\vartheta))$ of its estimates defined on some statistical structure parameterized by ϑ from an open set $\Theta \subseteq \mathbb{R}^d$, and dominated by a probability $\mathbb{P}$. It is proved that if $L(\vartheta)$ and $L_n(\vartheta)$ are random functions of class $C^2(\Theta)$ such that there exists a unique point $\hat{\vartheta} \in \Theta$ of the global maximum of $L(\vartheta)$ and the first and second derivatives of $L_n(\vartheta)$ with respect to ϑ converge to the corresponding derivatives of $L(\vartheta)$ uniformly on compacts in Θ with the order $O_{\mathbb{P}}(\gamma_n)$, $\lim_n \gamma_n = 0$, then there exists a sequence of Θ-valued random variables $\hat{\vartheta}_n$ which converges to $\hat{\vartheta}$ with the order $O_{\mathbb{P}}(\gamma_n)$ and such that $\hat{\vartheta}_n$ is a stationary point of $L_n(\vartheta)$ in asymptotic sense. Moreover, we prove that under two more assumption on L and L_n, such estimators could be chosen to be measurable with respect to the σ-algebra generated by $L_n(\vartheta)$.

1. Introduction

This paper is concerned with a generalization of the so called general theorem on approximate maximum likelihood estimation given in [9] as Theorem 3. A general problem is as follows. Let $L(\vartheta)$ be a log-likelihood function (or any contrast function) defined on some statistical structure parameterized with ϑ and dominated by a probability measure $\mathbb{P}$, and let $(L_n(\vartheta))$ be a sequence of its approximations. If we know that the maximum likelihood estimator (MLE for short) $\hat{\vartheta}$ exists and that $L_n(\vartheta)$ converge to $L(\vartheta)$ uniformly on compacts with known order, the problem is under what conditions there

2000 *Mathematics Subject Classification.* 62F10, 62F12.

Key words and phrases. Parameter estimation, consistent estimators, approximate likelihood function.

Supported in part by the MZT Grant 037058 of the Republic of Croatia.

exists a sequence of estimators $(\hat{\vartheta}_n)$ which converges to $\hat{\vartheta}$ and such that $\hat{\vartheta}_n$ is a stationary point of $L_n(\vartheta)$ in asymptotic sense. In addition, we are also concerned with the problem when $\hat{\vartheta}_n$ could be chosen to be measurable with respect to the σ-algebra generated by $L_n(\vartheta)$. For example, these problems arise when only discrete time observations of a continuous time process are available over bounded time interval (see for instance [6] or [7], and [9]).

Le Breton in [9] proved the theorem (Theorem 3 in [9]) of this kind and applied it to the problem of estimation of the drift parameter in a linear stochastic differential equation with constant coefficients. The main assumption was that the unknown parameter could be any value from $\mathbb{R}^d$. Moreover, some of the conditions on $L(\vartheta)$ were that the second derivative of L with respect to ϑ did not depend on ϑ (condition (A2) in [9] Theorem 3) and that L has a unique stationary point $\hat{\vartheta}$ (condition (A3) in the same theorem). These conditions together with the assumption about parameter space are somewhat restrictive. For example, if the drift of a diffusion is nonlinear in its parameters, then the second derivative of the log-likelihood function generally depends on the parameters. In this paper we prove a version of Le Breton's theorem assuming that the parameter space is an open set Θ in the Euclidean space $\mathbb{R}^d$ and omitting the restrictive condition (A2) from [9]. Moreover, we weaken (A3) from [9] by assuming that $\hat{\vartheta}$ is a unique point of maximum of L on Θ. To obtain appropriate measurability of $\hat{\vartheta}_n$ we have to impose some additional conditions on L and L_n. Note that these considerations were not present in Le Breton's approach. In this sense generalized version of Le Breton's theorem can be applied to more general diffusion models. For example (see [6, 7]), let us consider the diffusion growth model defined by the Itô stochastic differential equation

$$dX_t = (\alpha - \beta h(\gamma, X_t))X_t\, dt + \sigma X_t\, dW_t, \quad X_0 = x_0 > 0,$$

where $h(\gamma, x) = (x^\gamma - 1)/\gamma$ if $\gamma \neq 0$ and $h(\gamma, x) = \log x$ if $\gamma = 0$ ($\gamma \in \mathbb{R}$, $x > 0$), $\sigma > 0$ being fixed and where $\vartheta = (\alpha, \beta, \gamma)$ belongs to a relatively compact parameter space Θ, a neighborhood of the true value of ϑ with the specific property (see [7]) and such that

$$\Theta \subset \{(\alpha, \beta, \gamma) \in \mathbb{R}^3 \ : \ \beta > 0,\ \gamma(\alpha - \frac{\sigma^2}{2}) + \beta > 0\}.$$

The log-likelihood function based on continuous observation $X = (X_t, 0 \leq t \leq T)$ for ϑ is

$$L(\vartheta) = \frac{1}{\sigma^2}\int_0^T (\alpha - \beta h(\gamma, X_t))\frac{1}{X_t}\, dX_t - \frac{1}{2\sigma^2}\int_0^T (\alpha - \beta h(\gamma, X_t))^2\, dt.$$

Let us suppose that we observe X at time moments $0 = t_0 < t_1 < \ldots < t_n = T$, T being fixed, $n \in \mathbb{N}$. For any such choice of time moments we approximate

$L(\vartheta)$ with

$$L_n(\vartheta) = \frac{1}{\sigma^2}\sum_{i=1}^{n}\Big((\alpha-\beta h(\gamma, X_{t_{i-1}}))\frac{1}{X_{t_{i-1}}}(X_{t_i}-X_{t_{i-1}}) - \frac{1}{2}(\alpha-\beta h(\gamma, X_{t_{i-1}}))^2(t_i-t_{i-1})\Big).$$

$L(\vartheta)$ and $L_n(\vartheta)$, $n \in \mathbb{N}$, with $\gamma_n = \max_{1\le i\le n}(t_i - t_{i-1}) \to 0$, satisfy the conditions of the theorem (see Section 3) in the sense described in [7]. This model can be used in e.g. modeling growth of tumor spheroids where it is not possible for an individual spheroid to be observed arbitrarily long, but, theoretically, it is possible to observe it at every time moments in some bounded time interval (see [6]). It should be stressed that the asymptotic concept considered here and in [9] (i.e. where we discretely observe trajectories at n distinct time moments over the same bounded time interval in a way that the maximum of the times between successive observations Δ_n tends to zero), is different of the concept used by many other authors since in many other applications it is not possible to take observations at any time moment, but it is possible to take it arbitrarily long. For example, Yosida in [10] and Florens-Zmirnou in [5] proved that the estimators based on an approximation to the continuous-time likelihood function of an ergodic diffusion, obtained by replacing Lebesgue integrals and Itô integrals by Riemann-Itô sums as in the above example, are consistent and asymptotically efficient when $n\Delta_n \to +\infty$ for equidistant time points such that $\Delta_n \to 0$. If $\Delta_n = \Delta > 0$ is a constant, then the estimators are not consistent (see [2]). On the other hand, Dacunha-Castelle and Florens-Zmirou in [3] proved that for equidistant time points, the condition $n\Delta_n \to +\infty$ is sufficient (and necessary) for the discrete-time MLE of the drift parameters being consistent and asymptotically efficient. Since the transition densities cannot be obtained explicitly in many examples, some authors proposed other methods of estimations and proved the same good asymptotic properties of the obtained estimators under assumption that $\Delta_n = \Delta > 0$ is constant (for example see [1],[8] and [2]).

The paper is organized in the following way. Next section contains the definitions and notation used throughout in the paper. The generalized version of Le Breton's theorem is stated in Section 3 and the proofs are in Section 4.

2. Notation

We denote by $\langle\cdot|\cdot\rangle$ the scalar product in d-dimensional Euclidean space $\mathbb{R}^d$ and by $|\cdot|$ the induced norm.

If $x \mapsto f(x)$ is a real-valued function defined on an open subset of $\mathbb{R}^d$, then we will denote by $Df(x)$, $D^2f(x)$ its first and second derivatives with respect to x respectively. The notation $D^2 f(x) < 0$ means that the Hessian $D^2 f(x)$ is a negatively definite matrix.

Let U and Θ be open sets in $\mathbb{R}^d$ and $U \subset \Theta$. The closure of U will be denoted by $\mathcal{C}\ell(U)$ and the σ-algebra of Borel subsets of Θ by $\mathcal{B}(\Theta)$. We will say that U is a relatively compact set in Θ if U is an open set such that $\mathcal{C}\ell(U)$ is compact in Θ. If $\varepsilon > 0$ is a real number then $U + \varepsilon$ stands for the set $\{x \in \mathbb{R}^d : (\exists y \in U)\, |x - y| < \varepsilon\}$. Moreover, $K(x_0, \varepsilon)$ denotes the open ball in $\mathbb{R}^d$ with the center x_0 and radius ε.

Let $(\gamma_n, n \in \mathbb{N})$ be a sequence of positive numbers and let $X = (X_n, n \in \mathbb{N})$ be a sequence of random variables on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. We will say that X is $O_{\mathbb{P}}(\gamma_n)$, $n \in \mathbb{N}$, and write $X_n = O_{\mathbb{P}}(\gamma_n)$, $n \in \mathbb{N}$, if the sequence $(X_n/\gamma_n, n \in \mathbb{N})$ is bounded in probability, i.e. if

$$\lim_{A \to +\infty} \overline{\lim_{n}}\, \mathbb{P}\{\gamma_n^{-1} |X_n| > A\} = 0.$$

Generally, if $A \subset \Omega$ and $B \subset \mathbb{R}^d$ then A^c and B^c stand for the complement sets $\Omega \setminus A$ and $\mathbb{R}^d \setminus B$ respectively.

3. The result

Let Θ be an open subset of the Euclidean space $\mathbb{R}^d$, let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, and let $(\mathcal{F}_n, n \in \mathbb{N})$ be a family of sub-σ-algebras of $\mathcal{F}$. Moreover, let $(\gamma_n; n \in \mathbb{N})$ be a sequence of positive numbers such that $\lim_n \gamma_n = 0$, and let $L, L_n : \Omega \times \Theta \to \mathbb{R}$, $n \in \mathbb{N}$, be functions satisfying the following assumptions.

(A1): For all $\vartheta \in \Theta$, $\omega \mapsto L(\omega, \vartheta)$ is $\mathcal{F}$-measurable and $\omega \mapsto L_n(\omega, \vartheta)$ is $\mathcal{F}_n$-measurable, $n \in \mathbb{N}$. For all $\omega \in \Omega$, $\vartheta \mapsto L(\vartheta) \equiv L(\omega, \vartheta)$ and $\vartheta \mapsto L_n(\vartheta) \equiv L_n(\omega, \vartheta)$, $n \in \mathbb{N}$, are of class $C^2(\Theta)$.

(A2): For all $\omega \in \Omega$, the function $\vartheta \mapsto L(\vartheta) \equiv L(\omega, \vartheta)$ has a unique point of global maximum $\hat{\vartheta} \equiv \hat{\vartheta}(\omega)$ in Θ, and $D^2 L(\hat{\vartheta}) < 0$.

(A3): For any relatively compact set $\mathcal{K} \subset \Theta$,

$$\sup_{\vartheta \in \mathcal{K}} |D^l L_n(\vartheta) - D^l L(\vartheta)| = O_{\mathbb{P}}(\gamma_n),\ n \in \mathbb{N},\ \ l = 1, 2.$$

To obtain $\mathcal{F}_n$-measurability of the estimators, we need a few more assumptions.

(A4): For all $\omega \in \Omega$ and $r > 0$, $L(\omega, \hat{\vartheta}(\omega)) > \sup_{|x| \geq r} L(\omega, \hat{\vartheta}(\omega) + x)$;

(A5): $\sup_{\vartheta \in \Theta} |L_n(\vartheta) - L(\vartheta)| \xrightarrow{\mathbb{P}} 0$, $n \to \infty$.

Note that the measurability of the suprema in (A3-5) follows from (A1).

Theorem 3.1. *Let* (A1-3) *hold. Then there exists a sequence* $(\hat{\vartheta}_n, n \in \mathbb{N})$ *of* Θ*-valued random variables such that*

(i) $\lim_n \mathbb{P}\{DL_n(\hat{\vartheta}_n) = 0\} = 1$;

(ii) $\hat{\vartheta}_n \xrightarrow{\mathbb{P}} \hat{\vartheta}$, $n \to +\infty$;

(iii) *If* $(\tilde{\vartheta}_n; n \in \mathbb{N})$ *is any other sequence of random variables which satisfies* (i) *and* (ii), *then* $\lim_n \mathbb{P}\{\tilde{\vartheta}_n = \hat{\vartheta}_n\} = 1$;

(iv) *The sequence* $\gamma_n^{-1}(\hat{\vartheta}_n - \hat{\vartheta})$, $n \in \mathbb{N}$, *is bounded in probability.*

If in addition (A4-5) *hold, then* $\hat{\vartheta}_n$ *could be chosen to be* $\mathcal{F}_n$*-measurable,* $n \in \mathbb{N}$.

We conclude this section with another result about the existence of the $\mathcal{F}_n$-measurable estimators having the properties (i-iv) from the theorem. This result refers to the functions L_n, $n \in \mathbb{N}$, with very specific property and is based only on the assumptions (A1-3).

Corollary 3.2. *Let* (A1-3) *hold. If for all* $n \in \mathbb{N}$ *and* $\omega \in \Omega$, *the function* $\vartheta \mapsto L_n(\omega, \vartheta)$ *has a unique point* $\tilde{\vartheta}_n(\omega)$ *of local maximum in* Θ *which is a point of global maximum as well, then random variable* $(\tilde{\vartheta}_n, n \in \mathbb{N})$ *is a sequence of* Θ*-valued random variables such that for all* $n \in \mathbb{N}$, $\tilde{\vartheta}_n$ *is a* $\mathcal{F}_n$*-measurable and* (i-iv) *from the theorem hold.*

4. Proofs

We need the following lemmas. Some of them can be considered as standard but since we have not found them in literature as stated here, we provide proofs for readers' convenience.

Lemma 4.1. *Let* $\mathcal{G}$ *be a* σ*-algebra on* Ω *and let* $G : \Omega \times \Theta \to \mathbb{R}$ *be a function such that for all* $\omega \in \Omega$, $\vartheta \mapsto G(\omega, \vartheta)$ *is continuous in* Θ, *and for all* $\vartheta \in \Theta$, $\omega \mapsto G(\omega, \vartheta)$ *is* $\mathcal{G}$*-measurable. Then*

(i) G *is a* $\mathcal{G} \otimes \mathcal{B}(\Theta)$*-measurable function;*

(ii) *If for all* $\omega \in \Omega$, $\vartheta \mapsto G(\omega, \vartheta)$ *has a unique point* $\hat{\vartheta}(\omega)$ *of global maximum in* Θ, *then* $\omega \mapsto \hat{\vartheta}(\omega)$ *is a* $\mathcal{G}$*-measurable mapping.*

Proof. Let $(\Theta_n, n \in \mathbb{N})$ be a sequence of relatively compact sets in Θ such that $\Theta = \cup_{n \in \mathbb{N}} \Theta_n$ and for all $n \in \mathbb{N}$, $\mathcal{C}\ell(\Theta_n) \subset \Theta_{n+1}$. Let $n \in \mathbb{N}$ be fixed. Since $\mathcal{C}\ell(\Theta_n)$ is a compact set, there exist a finite number of points $(\vartheta_i^{(n)}, 1 \le i \le k_n)$ such that

$$\Theta_n \subset \mathcal{C}\ell(\Theta_n) \subseteq \bigcup_{i=1}^{k_n} K(\vartheta_i^{(n)}, \frac{1}{n}).$$

Let us define a sequence $(G^{(n)}, n \in \mathbb{N})$ of simple functions in $\Omega \times \Theta$ by

$$G^{(n)}(\omega,\vartheta) := \sum_{i=1}^{k_n} G(\omega,\vartheta_i^{(n)}) \cdot 1_{K(\vartheta_i^{(n)},\frac{1}{n})}(\vartheta) - \frac{1}{2} \sum_{1\le i<j\le k_n} \Big(G(\omega,\vartheta_i^{(n)}) + G(\omega,\vartheta_j^{(n)})\Big) \cdot 1_{K(\vartheta_i^{(n)},\frac{1}{n})\cap K(\vartheta_j^{(n)},\frac{1}{n})}(\vartheta) + \cdots + (-1)^{k_n}\frac{1}{k_n}(G(\omega,\vartheta_{i_1}^{(n)}) + \cdots + G(\omega,\vartheta_{i_{k_n}}^{(n)})) \cdot 1_{K(\vartheta_{i_1}^{(n)},\frac{1}{n})\cap\cdots\cap K(\vartheta_{i_{k_n}}^{(n)},\frac{1}{n})}(\vartheta).$$

Trivially, the functions $G^{(n)}$ $(n \in \mathbb{N})$ are $\mathcal{G} \otimes \mathcal{B}(\Theta)$-measurable and for all $(\omega,\vartheta) \in \Omega \times \Theta$, $G(\omega,\vartheta) = \lim_n G^{(n)}(\omega,\vartheta)$ by continuity of $\vartheta \mapsto G(\omega,\vartheta)$. Hence (i) follows.

Let $K \subset \Theta$ be an open ball. Since Θ is an open set in $\mathbb{R}^d$ and $\vartheta \mapsto G(\omega,\vartheta)$ is a continuous function for any $\omega \in \Omega$,

$$\begin{aligned}\{\omega \in \Omega : \hat{\vartheta}(\omega) \in K^c\} &= \{\omega \in \Omega : \forall\vartheta \in K, \exists\vartheta' \in K^c,\ G(\omega,\vartheta) < G(\omega,\vartheta')\} \\ &= \bigcap_{\vartheta\in K\cap\mathbb{Q}^d} \bigcup_{\vartheta'\in K^c\cap\mathbb{Q}^d} \{\omega \in \Omega : G(\omega,\vartheta) < G(\omega,\vartheta')\} \in \mathcal{G},\end{aligned}$$

which proves (ii). ∎

Lemma 4.2. *Let* (A1) *hold. Then for a positive random variable* ϵ *and a* Θ*-valued random variable* $\tilde{\vartheta}$*, the functions* $\delta^{(j)}(\tilde{\vartheta},\epsilon)$*,* $j = 1,2$*, defined by*

$$\delta^{(1)}(\tilde{\vartheta},\epsilon) := \sup\{\delta > 0 : \forall x \in \mathbb{R}^d,\ |x| \le \delta \Rightarrow |D^2L(\tilde{\vartheta}+x) - D^2L(\tilde{\vartheta})| \le \epsilon\},$$

$$\delta^{(2)}(\tilde{\vartheta},\epsilon) := \sup\{\delta > 0 : \forall x \in \mathbb{R}^d,\ |x| \le \delta \Rightarrow |DL(\tilde{\vartheta}+x) - DL(\tilde{\vartheta}) - D^2L(\tilde{\vartheta})x| \le \epsilon|x|\},$$

are positive not necessarily finite random variables.

Proof. (A1) implies $\delta^{(j)} > 0$, $j = 1,2$. Moreover, (A1) and Lemma 4.1 imply that the functions $(\omega,\vartheta) \mapsto D^lL(\omega,\vartheta)$, $l = 1,2$, are $\mathcal{F} \otimes \mathcal{B}(\Theta)$-measurable, and $\omega \mapsto \tilde{\vartheta}(\omega)$ is a $\mathcal{F}$-measurable mapping, by the assumption. Hence, $\omega \mapsto D^lL(\tilde{\vartheta}) \equiv D^lL(\omega,\tilde{\vartheta}(\omega))$, $l = 1,2$, are $\mathcal{F}$-measurable mappings. Let $h > 0$ be an arbitrary number. Then

$$\{\omega \in \Omega : \delta^{(1)}(\omega) \ge h\} = \{\omega \in \Omega : \forall x \in \mathbb{R}^d,\ |x| \le h \Rightarrow |D^2L(\omega,\tilde{\vartheta}(\omega)+x) - D^2L(\omega,\tilde{\vartheta}(\omega))| \le \epsilon(\omega)\}.$$

Because $\mathbb{Q}^d$ is a dense set in $\mathbb{R}^d$, and because for all $\omega \in \Omega$, the function $\vartheta \mapsto D^2L(\omega,\vartheta)$ is continuous, the set on the right side of the above equation is equal to the set

$$\bigcap_{x\in\mathbb{Q}^d,\ |x|\le h} \{\omega \in \Omega : |D^2L(\omega,\tilde{\vartheta}(\omega)+x) - D^2L(\omega,\tilde{\vartheta}(\omega))| \le \epsilon(\omega)\}$$

which is in $\mathcal{F}$. Hence, $\delta^{(1)}$ is a random variable. The proof of the $\mathcal{F}$-measurability of $\delta^{(2)}$ goes in a similar way. □

The following topological result is needed for proving the existence of the estimators.

LEMMA 4.3. *Let $K(0,r)$ $(r > 0)$ be an open ball in $\mathbb{R}^d$, and let $F : \mathcal{C}\ell(K(0,r)) \to \mathbb{R}^d$ be a continuous function. If for all $x \in \mathbb{R}^d$ such that $|x| = r$, $\langle x|F(x)\rangle < 0$, then there exists a point $x_* \in K(0,r)$ such that $F(x_*) = 0$.*

PROOF. Let us assume that the assertion of the lemma is false, i.e. for all $x \in K(0,r)$, $F(x) \neq 0$. Then the function $x \mapsto \frac{1}{|F(x)|}\langle x|F(x)\rangle$ is a well-defined continuous function. Hence, there exists a number $\epsilon > 0$ such that

$$r - \epsilon \leq |x| \leq r \quad \Rightarrow \quad \frac{1}{|F(x)|}\langle x|F(x)\rangle \leq -\epsilon. \tag{4.1}$$

Namely, if this is not true then there exists a sequence $(x_n; n \in \mathbb{N})$ in $\mathcal{C}\ell(K(0,r))$ such that

$$r - \frac{1}{n} \leq |x_n| \leq r \ \& \ \frac{1}{|F(x_n)|}\langle x_n|F(x_n)\rangle > -\frac{1}{n}, \quad n \in \mathbb{N}.$$

Since $(x_n; n \in \mathbb{N})$ is bounded, there exists a convergent subsequence $(x_{n_k}; k \in \mathbb{N})$, $\lim_k x_{n_k} = x'$. From the above conditions and the continuity, it follows that

$$|x'| = r \ \& \ \frac{1}{|F(x')|}\langle x'|F(x')\rangle \geq 0 \quad \Rightarrow \quad |x'| = r \ \& \ \langle x'|F(x')\rangle \geq 0,$$

which contradicts the assumption of the lemma. Hence, there exists $\epsilon > 0$ such that (4.1) holds. For such ϵ, $x \mapsto G(x) := x + \epsilon F(x)/|F(x)|$ is a $\mathbb{R}^d$-valued continuous function on $\mathcal{C}\ell(K(0,r))$ too. If $|x| < r - \epsilon$ then $|G(x)| \leq \epsilon + |x| \leq r$, and if $r - \epsilon \leq |x| \leq r$ then

$$|G(x)|^2 = \epsilon^2 + 2\frac{\epsilon}{|F(x)|}\langle x|F(x)\rangle + |x|^2 \leq r^2 - \epsilon^2 \leq r^2,$$

i.e. $|G(x)| \leq r$. Hence, G maps $\mathcal{C}\ell(K(0,r))$ into $\mathcal{C}\ell(K(0,r))$. By the Brower fixed point theorem (see e.g. [4], p. 144), there exists $x_0 \in K(0,r)$ such that $G(x_0) = x_0$. This implies $F(x_0) = 0$ which is in contradiction with the assumption. □

The last lemma is used in the proof of the $\mathcal{F}_n$-measurability of the estimators $\hat{\vartheta}_n$, $n \in \mathbb{N}$.

LEMMA 4.4. *Let* (A1-2) *and* (A4) *hold.*

(i) *There exists a positive and finite random variable Q such that*

$$(\forall x \in \mathbb{R}^d)\ |x| < Q \ \Rightarrow\ \hat{\vartheta} + x \in \Theta \text{ and } D^2 L(\hat{\vartheta} + x) < 0.$$

(ii) *For any random variable ε such that $0 < \varepsilon \leq Q$, there exists a random variable $s(\varepsilon)$ such that $0 < s(\varepsilon) < \varepsilon$ and*

$$\inf_{|x| \leq s(\varepsilon)} L(\hat{\vartheta} + x) > \sup_{|x| \geq \varepsilon} L(\hat{\vartheta} + x).$$

Moreover, $\Delta(\varepsilon) := \inf_{|x| \leq s(\varepsilon)} L(\hat{\vartheta} + x) - \sup_{|x| \geq \varepsilon} L(\hat{\vartheta} + x)$ is a positive random variable.

Proof. For fixed $n \in \mathbb{N}$, let

$$Q_n := \sup\{r > 0 : K(\hat{\vartheta}, r) \subseteq \Theta \ \ \& \sup_{\vartheta \in K(\hat{\vartheta}, r)} \sup_{|\eta|=1} \langle D^2 L(\vartheta)\eta|\eta\rangle \leq -\frac{1}{n}\}.$$

Since $\vartheta \mapsto D^2 L(\vartheta)$ is continuous and $\mathbb{Q}^d$ is dense in $\mathbb{R}^d$, for all $r > 0$,

$$\begin{aligned} \{Q_n < r\} \quad &= \quad \textstyle\bigcup_{\vartheta \in \mathbb{Q}^d \cap \Theta} \{|\vartheta - \hat{\vartheta}| < r,\ \sup_{|\eta|=1} \langle D^2 L(\vartheta)\eta|\eta\rangle > -\frac{1}{n}\} \cup \\ &\quad\ \cup \textstyle\bigcup_{y \in \mathbb{Q}^d \setminus \Theta} \{|y - \hat{\vartheta}| < r\} \in \mathcal{F}. \end{aligned}$$

Hence, for all $n \in \mathbb{N}$, Q_n is a random variable. Let $M > 0$ be an arbitrary real number and $Q = \sup_n Q_n \wedge M$. Then Q is a finite random variable, and $Q > 0$ by (A2). Moreover, let $x \in \mathbb{R}^d$ be such that $|x| < Q$. For $\epsilon = Q - |x| > 0$ there exists $n \in \mathbb{N}$ such that $|x| = Q - \epsilon < Q_n$, which implies that $\hat{\vartheta} + x \in \Theta$ and $\sup_{|\eta|=1} \langle D^2 L(\hat{\vartheta} + x)\eta|\eta\rangle \leq -\frac{1}{n} < 0$ by the definition of Q_n. This proves (i).

Let ε be a random variable such that $0 < \varepsilon \leq Q$. Note that $\sup_{|x| \geq \varepsilon} L(\hat{\vartheta} + x)$ is a finite random variable since for all $r \in \mathbb{R}$,

$$\{\sup_{|x| \geq \varepsilon} L(\hat{\vartheta} + x) > r\} = \bigcup_{\vartheta \in \mathbb{Q}^d \cap \Theta} \{|\vartheta - \hat{\vartheta}| \geq \varepsilon, L(\vartheta) > r\} \in \mathcal{F}$$

and (A2) holds. Let

$$S_\varepsilon := \sup\{0 < r \leq \varepsilon/2 : \inf_{|x| \leq r} L(\hat{\vartheta} + x) \geq \sup_{|x| \geq \varepsilon} L(\hat{\vartheta} + x)\}.$$

(A4) implies $S_\varepsilon > 0$. Moreover, for all $r > 0$,

$$\{S_\varepsilon \geq r\} = \{\inf_{|x| \leq r} L(\hat{\vartheta} + x) \geq \sup_{|x| \geq \varepsilon} L(\hat{\vartheta} + x)\} \in \mathcal{F}$$

by (A1). Hence, S_ε is a random variable. Let $s(\varepsilon) := S_\varepsilon/2$. Then $s(\varepsilon)$ is a random variable such that $0 < s(\varepsilon) \leq \varepsilon/4 < \varepsilon$ and

$$\inf_{|x| \leq s(\varepsilon)} L(\hat{\vartheta} + x) \geq \inf_{|x| \leq S_\varepsilon} L(\hat{\vartheta} + x) \geq \sup_{|x| \geq \varepsilon} L(\hat{\vartheta} + x)$$

by the definition. To prove that $\inf_{|x|\le s(\varepsilon)} L(\hat\vartheta+x) > \sup_{|x|\ge\varepsilon} L(\hat\vartheta+x)$ it is sufficient to show that $\inf_{|x|\le s(\varepsilon)} L(\hat\vartheta+x) > \inf_{|x|\le S_\varepsilon} L(\hat\vartheta+x)$. Let us suppose that this is not true, i.e. $\inf_{|x|\le S_\varepsilon/2} L(\hat\vartheta+x) = \inf_{|x|\le S_\varepsilon} L(\hat\vartheta+x)$. Since

$$\inf_{|x|\le S_\varepsilon/2} L(\hat\vartheta+x) \ge \inf_{|x|\le 3S_\varepsilon/4} L(\hat\vartheta+x) \ge \inf_{|x|\le S_\varepsilon} L(\hat\vartheta+x),$$

$\inf_{|x|\le S_\varepsilon/2} L(\hat\vartheta+x) = \inf_{|x|\le 3S_\varepsilon/4} L(\hat\vartheta+x)$. These minima are obtained at points in $K(0,\varepsilon)$. Since $K(0,\varepsilon)\subseteq K(0,Q)$ and $x=0$ is a unique stationary point and a point of maximum of $x\mapsto L(\hat\vartheta+x)$ in $K(0,\varepsilon)$, the points of minima have to be in spheres $\{x : |x| = S_\varepsilon/2\}$ and $\{x : |x| = 3S_\varepsilon/4\}$ respectively. Let x_* be a point of the minimum of the function $x\mapsto L(\hat\vartheta+x)$ defined in the set $\{x : |x| \le S_\varepsilon/2\}$. Then $|x_*| = S_\varepsilon/2$ and x_* is a point of the minimum of the same function but defined in $\{x : |x|\le 3S_\varepsilon/4\}$. Hence, x_* is another stationary point of the function $x\mapsto L(\hat\vartheta+x)$ in $K(0,\varepsilon)$, which is a contradiction.

Finally, it follows easily that $\Delta(\varepsilon)$ is a positive random variable. □

Proof of the theorem. Let $\{\mathcal{K}_m; m\in\mathbb{N}\}$ be a family of relatively compact sets and $(\varepsilon_m; m\in\mathbb{N})$ be a sequence of positive numbers such that

$$(\forall m\in\mathbb{N})\ \mathcal{C}\ell(\mathcal{K}_m)\subset\mathcal{C}\ell(\mathcal{K}_m+\varepsilon_m)\subset\mathcal{K}_{m+1} \text{ and } \bigcup_{m=1}^{\infty}\mathcal{K}_m=\Theta,$$
$$(\forall m\in\mathbb{N})\ \varepsilon_m>\varepsilon_{m+1} \text{ and } \lim_m \varepsilon_m = 0,$$

and let

$$\Omega_{nm} := \{\hat\vartheta\in\mathcal{K}_m,\ \sup_{|x|=\varepsilon_m}\langle DL_n(\hat\vartheta+x)|x\rangle<0,\ \sup_{|x|\le\varepsilon_m}\sup_{|\xi|=1}\langle D^2L_n(\hat\vartheta+x)\xi|\xi\rangle<0\}.$$

By Lemma 4.3 there exists a point $\hat\vartheta_{nm}\in K(\hat\vartheta,\varepsilon_m)$ such that $DL_n(\hat\vartheta_{nm})=0$ on the event Ω_{nm}. On the same event, $\hat\vartheta_{nm}$ is a unique stationary point and so a unique point of the maximum of the function $\vartheta\mapsto L_n(\vartheta)$ resticted to $K(\hat\vartheta,\varepsilon_m)$ since this function is strictly concave (see [9]). By Lemma 4.1 (ii) $\omega\mapsto\hat\vartheta_{nm}(\omega)$ is a $\mathcal{F}\cap\Omega_{nm}$-measurable random variable. By the same lemma, part (i), $\Omega_{nm}\in\mathcal{F}$. Moreover, if $\omega\in\Omega_{nm}\cap\Omega_{nm'}$ for some $m,m'\in\mathbb{N}$ such that, say, $m<m'$, then $\hat\vartheta_{nm},\hat\vartheta_{nm'}\in K(\hat\vartheta,\varepsilon_m)$ and $\hat\vartheta_{nm}=\hat\vartheta_{nm'}$ by the uniqueness. Hence, the function $\omega\mapsto\hat\vartheta_n(\omega)$ defined by

$$\hat\vartheta_n(\omega) := \begin{cases} \hat\vartheta_{nm}(\omega) & \text{if } (\exists m\in\mathbb{N})\ \omega\in\Omega_{nm} \\ \hat\vartheta(\omega) & \text{if } (\forall m\in\mathbb{N})\ \omega\notin\Omega_{nm}, \end{cases}$$

is a Θ-valued random variable. From the discussion given above it follows that for any $n\in\mathbb{N}$ and $\eta>0$,

$$\Omega_{nm}\subseteq\{DL_n(\hat\vartheta_n)=0\},\quad m\in\mathbb{N} \text{ and}$$
$$\Omega_{nm}\subseteq\{|\hat\vartheta_n-\hat\vartheta|<\varepsilon_m\}\subseteq\{|\hat\vartheta_n-\hat\vartheta|<\eta\},\quad m\ge m_0,$$

where $m_0 \in \mathbb{N}$ is such that $\eta \geq \varepsilon_{m_0}$. If

$$\lim_m \overline{\lim_n} \mathbb{P}(\Omega_{nm}^c) = 0 \tag{4.2}$$

holds, then (i) and (ii) from the theorem follow. Moreover, let $(\tilde{\vartheta}_n; n \in \mathbb{N})$ be any other sequence of Θ-valued random variables. By the same uniqueness argument it follows that

$$\begin{aligned} & \Omega_{nm} \cap \{|\tilde{\vartheta}_n - \hat{\vartheta}| < \varepsilon_m\} \cap \{DL_n(\tilde{\vartheta}_n) = 0\} \subseteq \{\tilde{\vartheta}_n = \hat{\vartheta}_n\} \\ \Rightarrow \quad & \mathbb{P}\{\tilde{\vartheta}_n \neq \hat{\vartheta}_n\} \leq \mathbb{P}(\Omega_{nm}^c) + \mathbb{P}\{|\tilde{\vartheta}_n - \hat{\vartheta}| \geq \varepsilon_m\} + \mathbb{P}\{DL_n(\tilde{\vartheta}_n) \neq 0\}. \end{aligned}$$

If $\tilde{\vartheta}_n$, $n \in \mathbb{N}$, satisfy (i) and (ii) from the theorem, and if (4.2) holds, then $\lim_n \mathrm{P}\{\tilde{\vartheta}_n = \hat{\vartheta}_n\} = 1$. Hence to prove the statements (i-iii) from the theorem it is sufficient to prove (4.2).

(A1-2) and the continuity of the quadratic form $\xi \mapsto \langle D^2L(\hat{\vartheta})\xi|\xi\rangle$ imply that $\omega \mapsto q(\omega) := -\max_{|\xi|=1}\langle D^2L(\omega, \hat{\vartheta}(\omega))\xi|\xi\rangle$ is a positive random variable. By definition, for all $x \in \mathbb{R}^d$,

$$\langle D^2L(\hat{\vartheta})x|x\rangle \leq -q|x|^2. \tag{4.3}$$

For $\omega \in \Omega$ and $x \in \mathbb{R}^d$, let $\phi \equiv \phi_{\omega,x} : [0,1] \to \mathbb{R}$ be the function defined by

$$\phi(t) := \langle DL_n(\hat{\vartheta} + tx)|x\rangle.$$

(A1) implies that the mean value theorem could be applied on ϕ. Hence, there exists a number $\lambda \equiv \lambda(\omega, x) \in \langle 0, 1\rangle$ such that

$$\begin{aligned} & \phi(1) - \phi(0) = \phi'(\lambda) \\ \Rightarrow \quad & \langle DL_n(\hat{\vartheta} + x)|x\rangle - \langle DL_n(\hat{\vartheta})|x\rangle = \langle D^2L_n(\bar{\vartheta}(x))x|x\rangle, \end{aligned}$$

$$\begin{aligned} \Rightarrow \quad \langle DL_n(\hat{\vartheta} + x)|x\rangle \quad = \quad & \langle D^2L(\hat{\vartheta})x|x\rangle + \langle DL_n(\hat{\vartheta})|x\rangle + \\ & + \langle (D^2L(\bar{\vartheta}(x)) - D^2L(\hat{\vartheta}))x|x\rangle + \\ & + \langle (D^2L_n(\bar{\vartheta}(x)) - D^2L(\bar{\vartheta}(x)))x|x\rangle. \end{aligned} \tag{4.4}$$

where $\bar{\vartheta}(x) = \hat{\vartheta} + \lambda x$.

Let C_{nm} be the event defined by the following inequalities:

$$\sup_{\vartheta \in \mathcal{K}_m} |DL_n(\vartheta) - DL(\vartheta)| \quad < \quad \frac{\varepsilon_m}{2} q \tag{4.5}$$

$$\sup_{\vartheta \in \mathcal{K}_{m+1}} |D^2L_n(\vartheta) - D^2L(\vartheta)| \quad < \quad \frac{1}{6} q \tag{4.6}$$

$$\hat{\vartheta} \in \mathcal{K}_m \tag{4.7}$$

$$\delta^{(1)}(\hat{\vartheta}, \frac{1}{6} q) \quad \geq \quad \varepsilon_m \tag{4.8}$$

where $\delta^{(1)}$ is the random function from Lemma 4.2. Let $x \in \mathbb{R}^d$ be such that $|x| \leq \varepsilon_m$ and let $\bar{\vartheta}(x)$ be from (4.4).

On the event C_{nm}, $\bar{\vartheta}(x) \in \mathcal{K}_{m+1}$. Moreover, (4.5) and (4.7) imply

$$|DL_n(\hat{\vartheta})| < \frac{\varepsilon_m}{2} q. \tag{4.9}$$

Since $|\bar{\vartheta}(x) - \hat{\vartheta}| = |\lambda x| \leq |x| \leq \varepsilon_m$, (4.8) implies

$$\sup_{|x| \leq \varepsilon_m} |D^2 L(\bar{\vartheta}(x)) - D^2 L(\hat{\vartheta})| \leq \frac{1}{6} q \tag{4.10}$$

by the definition of the random function $\delta^{(1)}$, and (4.6) implies

$$\sup_{|x| \leq \varepsilon_m} |D^2 L_n(\bar{\vartheta}(x)) - D^2 L(\bar{\vartheta}(x))| \leq \frac{1}{6} q. \tag{4.11}$$

For all $x \in \mathbb{R}^d$ such that $|x| = \varepsilon_m$, (4.9-4.11), (4.4) and (4.3) imply

$$\langle x | DL_n(\hat{\vartheta} + x) \rangle < -\frac{\varepsilon_m^2}{6} q < 0 \tag{4.12}$$

on C_{nm}. Moreover, (4.6) and (4.8) imply

$$|D^2 L_n(\hat{\vartheta} + x) - D^2 L(\hat{\vartheta})| < \frac{1}{3} q,$$

and this and (4.3) imply that $D^2 L_n(\hat{\vartheta} + x) < 0$ for all $x \in \mathbb{R}^d$ such that $|x| \leq \varepsilon_m$. Hence $C_{nm} \subseteq \Omega_{nm}$ and

$$\begin{aligned} \mathbb{P}(\Omega_{nm}^c) \leq\ & \mathbb{P}\{\sup_{\vartheta \in \mathcal{K}_m} |DL_n(\vartheta) - DL(\vartheta)| \geq \tfrac{\varepsilon_m}{4} q\} + \\ & + \mathbb{P}\{\sup_{\vartheta \in \mathcal{K}_{m+1}} |D^2 L_n(\vartheta) - D^2 L(\vartheta)| \geq \tfrac{1}{6} q\} + \\ & + \mathbb{P}\{\hat{\vartheta} \in \mathcal{K}_m^c\} + \mathbb{P}\{\delta^{(1)}(\hat{\vartheta}, \tfrac{1}{6} q) < \varepsilon_m\}. \end{aligned}$$

which implies (4.2), since $\hat{\vartheta} \in \Theta$, $\delta^{(1)}(\hat{\vartheta}, \frac{1}{6} q) > 0$ and (A3) holds.

Let us prove the statement (iv). Let $n, m \in \mathbb{N}$, $A > 0$ and $K > 0$ be arbitrary numbers, and let $\delta^{(2)}$ be the random function from Lemma 4.2. On the event

$$\begin{aligned} B = \{ & |(D^2 L(\hat{\vartheta}))^{-1}| \leq K, \ \ \delta^{(2)}(\hat{\vartheta}, \tfrac{1}{2K}) \geq \varepsilon_m, \ \ DL_n(\hat{\vartheta}_n) = 0, \\ & \hat{\vartheta} \in \mathcal{K}_m, \ \ |\hat{\vartheta}_n - \hat{\vartheta}| \leq \varepsilon_m, \ \ \sup_{\vartheta \in \mathcal{K}_{m+1}} |DL_n(\vartheta) - DL(\vartheta)| \leq \tfrac{1}{2K} A \gamma_n \}, \end{aligned}$$

$\hat{\vartheta}_n \in \mathcal{K}_{m+1}$ and

$$\begin{aligned} |\hat{\vartheta}_n - \hat{\vartheta}| & \leq |(D^2 L(\hat{\vartheta}))^{-1}| \cdot |D^2 L(\hat{\vartheta})(\hat{\vartheta}_n - \hat{\vartheta})| \leq \\ & \leq |(D^2 L(\hat{\vartheta}))^{-1}| \cdot |DL(\hat{\vartheta}_n) - DL(\hat{\vartheta}) - D^2 L(\hat{\vartheta})(\hat{\vartheta}_n - \hat{\vartheta})| + \\ & \quad + |(D^2 L(\hat{\vartheta}))^{-1}| \cdot |DL_n(\hat{\vartheta}_n) - DL(\hat{\vartheta}_n)| \leq \\ & \leq K \tfrac{1}{2K} |\hat{\vartheta}_n - \hat{\vartheta}| + K \tfrac{1}{2K} \gamma_n A \leq \\ & \leq \tfrac{1}{2} |\hat{\vartheta}_n - \hat{\vartheta}| + \tfrac{1}{2} \gamma_n A \\ \Rightarrow \ |\hat{\vartheta}_n - \hat{\vartheta}| & \leq \gamma_n A. \end{aligned}$$

Hence, $B \subseteq \{\gamma_n^{-1}|\hat{\vartheta}_n - \hat{\vartheta}| \leq A\}$. This implies

$$\begin{aligned}\mathbb{P}\{\gamma_n^{-1}|\hat{\vartheta}_n - \hat{\vartheta}| > A\} \quad \leq \quad & \mathbb{P}\{|(D^2L(\hat{\vartheta}))^{-1}| > K\} + \\ & + \mathbb{P}\{\delta^{(2)}(\hat{\vartheta}, \tfrac{1}{2K}) < \varepsilon_m\} + \mathbb{P}\{\hat{\vartheta} \in \mathcal{K}_m^c\} + \\ & + \mathbb{P}\{|\hat{\vartheta}_n - \hat{\vartheta}| > \varepsilon_m\} + \mathbb{P}\{DL_n(\hat{\vartheta}_n) \neq 0\} + \\ & + \mathbb{P}\{\sup_{\vartheta \in \mathcal{K}_{m+1}} |DL_n(\vartheta) - DL(\vartheta)| > \tfrac{1}{2K}\gamma_n A\}.\end{aligned}$$

Since (A3), (i) and (ii) hold,

$$\begin{aligned}& \lim_{A \to +\infty} \overline{\lim}_n \mathbb{P}\{\gamma_n^{-1}|\hat{\vartheta}_n - \hat{\vartheta}| > A\} \leq \\ \leq \quad & \mathbb{P}\{|(D^2L(\hat{\vartheta}))^{-1}| > K\} + \mathbb{P}\{\delta^{(2)}(\hat{\vartheta}, \tfrac{1}{2K}) < \varepsilon_m\} + \mathbb{P}\{\hat{\vartheta} \in \mathcal{K}_m^c\}.\end{aligned}$$

Letting first $m \to +\infty$ and then $K \to +\infty$,

$$\lim_{A \to +\infty} \overline{\lim_n} \, \mathbb{P}_{\theta_0}\{\gamma_n^{-1}|\hat{\vartheta}_n - \hat{\vartheta}| > A\} = 0$$

which proves (iv).

Finally, let us prove the last statement of the theorem about the existence of a $\mathcal{F}_n$-measurable version of the estimators. Assume that (A4) and (A5) hold. Let $\mathcal{K}_m$, ε_m, $m \in \mathbb{N}$, be as above. For $n, m \in \mathbb{N}$, the event

$$\begin{aligned}\tilde{\Omega}_{nm} \quad := \quad & \bigcup_{\vartheta \in \mathcal{K}_m \cap \mathbb{Q}^d} \bigcup_{\eta \in \langle 0, \varepsilon_m \rangle \cap \mathbb{Q}} (\{\sup_{|x| = \eta} \langle DL_n(\vartheta + x) | x \rangle < 0\} \cup \\ & \cup \{\sup_{|x| \leq \varepsilon_m} \sup_{|\xi| = 1} \langle D^2L_n(\vartheta + x)\xi | \xi \rangle < 0\} \cup \\ & \cup \{\inf_{|x| \leq \eta} L_n(\vartheta + x) > \sup_{|x| \geq \varepsilon_m} L_n(\vartheta + x)\})\end{aligned}$$

is in $\mathcal{F}_n$. By the similar arguments to those used for $\hat{\vartheta}_n$, there exists a unique point $\hat{\vartheta}'_{nm} \in \Theta$ of global maximum of the function $\vartheta \mapsto L_n(\vartheta)$, $L_n : \Theta \to \mathbb{R}$, on the event $\tilde{\Omega}_{nm}$. Hence $\omega \mapsto \hat{\vartheta}'_{nm}(\omega)$ is a $\mathcal{F}_n$-measurable random variable on $\tilde{\Omega}_{nm}$ by Lemma 4.1 (ii), and $\hat{\vartheta}'_n : \Omega \to \Theta$ defined by

$$\hat{\vartheta}'_n(\omega) := \begin{cases} \hat{\vartheta}'_{nm}(\omega) & \text{if } (\exists m \in \mathbb{N})\, \omega \in \tilde{\Omega}_{nm} \\ \vartheta_* & \text{if } (\forall m \in \mathbb{N})\, \omega \notin \tilde{\Omega}_{nm} \end{cases}$$

is $\mathcal{F}_n$-measurable random variable. Here, ϑ_* is an arbitrary fixed point in Θ.

Since (A4) holds, Lemma 4.4 could be applied. Let Q, $s(\varepsilon_m \wedge Q)$ and $\Delta(\varepsilon_m \wedge Q)$ be random variables from that lemma. Obviously, $\hat{\vartheta}'_n = \hat{\vartheta}_n$ on $\Omega_{nm} \cap \tilde{\Omega}_{nm} \cap \{\inf_{|x| \leq s(\varepsilon_m \wedge Q)} L_n(\hat{\vartheta} + x) > \sup_{|x| \geq \varepsilon_m \wedge Q} L_n(\hat{\vartheta} + x)\}$ for any $n, m \in \mathbb{N}$. Therefore, to prove that $\hat{\vartheta}'_n$, $n \in \mathbb{N}$, satisfies (i-iv) it is sufficient to prove that

(4.13)

$$\lim_m \overline{\lim_n} \mathbb{P}(\tilde{\Omega}_{nm}^c) = \lim_m \overline{\lim_n} \mathbb{P}\{\inf_{|x| \leq s(\varepsilon_m \wedge Q)} L_n(\hat{\vartheta} + x) \leq \sup_{|x| \geq \varepsilon_m \wedge Q} L_n(\hat{\vartheta} + x)\} = 0.$$

Let $\eta > 0$ be a rational number, and let $C'_{nm\eta}$ be the event defined by (4.5) with ε_m being replaced with η, (4.6-4.8) and

$$s(\varepsilon_m \wedge Q) \geq \eta \tag{4.14}$$
$$Q \geq \varepsilon_m \tag{4.15}$$
$$\sup_{\vartheta\in\Theta} |L_n(\vartheta) - L(\vartheta)| \leq \frac{\Delta(\varepsilon_m \wedge Q)}{4}. \tag{4.16}$$

By using the same arguments as in the proof of (4.12) we conclude that

$$\sup_{|x|=\eta} \langle x|DL_n(\hat\vartheta + x)\rangle < -\frac{\eta^2}{6}q < 0$$

on the event $C'_{nm\eta}$. Moreover, since (4.3), (4.6) and (4.8) hold,

$$\sup_{|x|\leq\varepsilon_m} \sup_{|\xi|=1} \langle D^2L_n(\hat\vartheta + x)\xi|\xi\rangle < 0$$

on $C'_{nm\eta}$ too. Finally, let $x, y \in \mathbb{R}^d$ be such that $|x| \leq \eta$, $|y| \geq \varepsilon_m$ and $\hat\vartheta + x$, $\hat\vartheta + y \in \Theta$. Since $0 < \eta \leq s(\varepsilon_m \wedge Q) = s(\varepsilon_m) < \varepsilon_m \leq Q$ by (4.14-4.15) and Lemma 4.4 (ii),

$$\begin{aligned}
L_n(\hat\vartheta + x) &= L_n(\hat\vartheta + x) - L(\hat\vartheta + x) + +L(\hat\vartheta + x) - L(\hat\vartheta + y) + \\
&\quad + L(\hat\vartheta + y) - L_n(\hat\vartheta + y) + L_n(\hat\vartheta + y) \\
&\overset{(4.16)}{\geq} -\frac{\Delta(\varepsilon_n)}{4} + \inf_{|x|\leq\eta} L(\hat\vartheta + x) - \sup_{|y|\geq\varepsilon_m} L(\hat\vartheta + y) \\
&\quad - \frac{\Delta(\varepsilon_m)}{4} + L_n(\hat\vartheta + y) \\
&\geq \frac{\Delta(\varepsilon_m)}{2} + L_n(\hat\vartheta + y)
\end{aligned}$$

which implies

$$\inf_{|x|\leq\eta} L_n(\hat\vartheta + x) \geq \frac{\Delta(\varepsilon_m)}{2} + \sup_{|x|\geq\varepsilon_m} L_n(\hat\vartheta + x)$$

on $C'_{nm\eta}$. In particular,

$$\inf_{|x|\leq s(\varepsilon_m\wedge Q)} L_n(\hat\vartheta + x) \geq \frac{\Delta(\varepsilon_m \wedge Q)}{2} + \sup_{|x|\geq\varepsilon_m\wedge Q} L_n(\hat\vartheta + x)$$

holds on $C'_{nm\eta}$ for $\eta = s(\varepsilon_m \wedge Q)$. Since the functions $\vartheta \mapsto \sup_{|x|=\eta}\langle DL_n(\vartheta + x)|x\rangle$, $\sup_{|x|\leq\varepsilon_m}\sup_{|\xi|=1}\langle D^2L_n(\vartheta+x)\xi|\xi\rangle$, $\inf_{|x|\leq\eta} L_n(\vartheta+x) - \sup_{|x|\geq\varepsilon_m} L_n(\vartheta+x)$ are continuous at $\vartheta = \hat\vartheta$, $C'_{nm\eta} \subseteq \tilde\Omega_{nm}$. Because $Q > 0$, $s(\varepsilon_m \wedge Q) > 0$ and $\Delta(\varepsilon_m \wedge Q) > 0$ by Lemma 4.4, and $\delta^{(1)}(\hat\vartheta, \frac{1}{6}q) > 0$ by Lemma 4.2, and because (A3) and (A5) hold,

$$\lim_m \lim_{\eta\to 0} \overline{\lim_n} \mathbb{P}(C'^{c}_{nm\eta}) = 0 \text{ and}$$
$$\lim_m \overline{\lim_n} \mathbb{P}\{ \inf_{|x|\le s(\varepsilon_m\wedge Q)} L_n(\hat\vartheta + x) \le \sup_{|x|\ge \varepsilon_m\wedge Q} L_n = (\hat\vartheta + x)\} = 0$$

which implies (4.13). □

Proof of the corollary. Let $n \in \mathbb{N}$. Since $\tilde\vartheta_n$ is a unique point of global maximum of $\mathcal{F}_n$-measurable function L_n, $\tilde\vartheta_n$ is an $\mathcal{F}_n$-measurable random variable by Lemma 4.1 (ii). Moreover, on the events Ω_{nm}, $m \in \mathbb{N}$, $\tilde\vartheta_n$ coincides with the estimator $\hat\vartheta_n$ defined in the first part of the proof of Theorem 3.1 since $\tilde\vartheta_n$ is a unique point of local maximum too. From this, (i-iv) follow just as in the proof of the theorem. □

Remark 4.5. Note that the same assertion as in the corollary holds under the following weaker condition: (A1-3) hold and for any $n \in \mathbb{N}$, there exists a sequence $(A_{nm}; m \in \mathbb{N})$ of events in $\mathcal{F}_n$ such that (i) $\lim_m \overline{\lim}_n \mathbb{P}(A^c_{nm}) = 0$, (ii) on A_{nm}, there exists a unique point of global maximum $\tilde\vartheta_n$ of $L_n(\vartheta)$, $\vartheta \in \Theta$, and (iii) on A_{nm}, there exists an open ball $U_m \subseteq \Theta$ with center in $\hat\vartheta$, such that $\tilde\vartheta_n \in U_m$ and $\tilde\vartheta_n$ is a unique stationary point of L_n on U_m.

Acknowledgements.
The author would like to thank Professors Hrvoje Šikić and Zoran Vondraček for their support and many helpful comments and suggestions.

References

[1] B. M. Bibby and M. Sørensen, Martingale estimation functions for discretely observed diffusion processes, *Bernoulli*, **1** (1/2) (1995) 17-39
[2] B. M. Bibby and M. Sørensen, On estimation of discretely sampled diffusions: a review, *Theory of Stochastic Processes*, **2** (1996) 49-56
[3] D. Dacunha-Castelle and D. Florens-Zmirou, Estimation of the coefficients of a diffusion from discrete observations, *Stochastics*, **19** (1986) 263-284
[4] Yu. Borisovich and N. Bliznyakov, Ya. Izrailevich, T. Fomenko, *Introduction to Topology*, Mir Publishers, Moscow, 1985
[5] D. Florens-Zmirou, Approximate discrete time shemes for statistics of diffusion processes, *Statistics*, **20** (1989) 547-557
[6] M. Huzak, *Selection of diffusion growth process and parameter estimation from discrete observation*, Ph.D. thesis, University of Zagreb, 1997 (in Croatian)
[7] M. Huzak, Parameter estimation of diffusion models from discrete observations, *Mathematical Communications* **3**(1998) 221-225
[8] M. Kessler and M. Sørensen, Estimating equations based on eigenfunctions for a discretely observed diffusion processes, *Bernoulli*, **5** (2)(1999) 299-314
[9] A. Le Breton, On continuous and discrete sampling for parameter estimation in diffusion type processes, *Math. Prog. Study* **5**(1976) 124-144

[10] N. Yoshida, Estimation for diffusion processes from discrete observations, *J. Multivar. Anal.*, **41** (1992) 220-242

Department of Mathematics,
University of Zagreb,
Bijenička 30,
10000 Zagreb, Croatia
E-mail: huzak@math.hr

Received: 12.06.2000.
Revised: 05.07.2000.