Ann. Inst. Statist. Math.
Vol. 50, No. 3, 493–502 (1998)

# ON MINIMUM DISTANCE ESTIMATION IN RECURRENT MARKOV STEP PROCESSES II

R. HÖPFNER[1]* AND YU. A. KUTOYANTS[2]

[1] *Institut für Mathematische Stochastik, Universität Freiburg, Hebelstrasse 27, D-79104 Freiburg im Breisgau, Germany*
[2] *Laboratoire de Statistique et Processus, Universite du Maine, F-72085 Le Mans, Cédex 9, France*

(Received May 8, 1996; revised June 10, 1997)

**Abstract.** Consider a Markov step process $X = (X_t)_{t\geq 0}$ whose generator depends on an unknown parameter $\vartheta$. We are interested in estimation of $\vartheta$ by a class of minimum distance estimators (MDE) based on observation of $X$ up to time $S_n$, with $(S_n)_n$ a sequence of stopping times increasing to $\infty$. We give a precise description of the MDE error at stage $n$, for $n$ fixed, i.e. a stochastic expansion in terms of powers of a norming constant and suitable coefficients (which can be calculated explicitly from the observed path of $X$ up to time $S_n$).

*Key words and phrases.* Markov step processes, minimum distance estimators, stochastic expansions.

## 1. Introduction

Minimum distance estimators (MDE) have been studied for statistical models in various settings. General features of asymptotics of minimum distance estimators have been formulated by Millar (1984) who applies these results to the case where i.i.d random variables or certain discrete-time processes are observed. Models for diffusion-type processes are treated in Kutoyants (1984, 1994), and we refer to the references therein. In the present note we continue our study (see Höpfner and Kutoyants (1997)) of minimum distance estimation of an unknown (here: one-dimensional) parameter $\vartheta$ in a recurrent Markov step process $X = (X_t)_{t\geq 0}$, under more restrictive assumptions on the parameterization, and with a different aim. In Höpfner and Kutoyants (1997), we were interested in asymptotic properties of MDE's if at stage $n$, the process is observed continuously up to time $S_n$, $(S_n)_n$ a sequence of stopping times increasing to $\infty$ with $n$. Here, as a complement to

* Now at FB 17 Mathematik-Informatik, Universität-GH Paderborn, D-33095 Paderborn, Germany.

these asymptotic results, we give an accurate description of the MDE $\vartheta^*_{S_n}$ at fixed level $n$, in terms of an expansion

$$P_\vartheta\left(\left|\vartheta^*_{S_n} - \left\{\vartheta + \sum_{j=1}^{r}\frac{1}{j!}Y_{n,j}(\vartheta)\delta_n(\vartheta)^j\right\}\right| \geq \frac{C_{r+1}(\vartheta)}{(r+1)!}\delta_n(\vartheta)^{r+1/2}\right) \leq P_\vartheta(B^c_{n,\vartheta})$$

where $(\delta_n(\vartheta))_{n\geq 1}$ is a sequence of norming constants for the MDE error at $\vartheta$, where the coefficients $Y_{n,j}(\vartheta)$ can be calculated from the trajectory of the process $X$ up to time $S_n$, and where $B^c_{n,\vartheta}$ is a small exceptional set such that bounds on $P_\vartheta(B^c_{n,\vartheta})$ are available. Expansions of this type have been considered for maximum likelihood estimators of parameters in diffusion processes by Kutoyants (1994) and in Poisson process intensities by Kutoyants (1983). In our Markov step process setting, a particular choice of an observation scheme $(S_n)_n$ allows to take $\delta_n(\vartheta) = n^{-1/2}$ for all $\vartheta \in \Theta$, and leads to an expansion of type

$$P_\vartheta\left(\left|\vartheta^*_{S_n} - \left\{\vartheta + \sum_{j=1}^{r}\frac{1}{j!}Y_{n,j}(\vartheta)n^{-j/2}\right\}\right| \geq \frac{Cn^{-r/2-1/4}}{(r+1)!}\right) \leq C'_p n^{-p/4(2r+1)}$$

valid uniformly in $\vartheta \in \Theta$, for all $n \geq 1$ and arbitrary $p > 2$, where $C$, $C'_p$ are suitable constants. Almost no conditions on $\Theta$ (except open and one-dimensional) are needed; so we have very accurate information on the MDE error at every fixed level $n$.

Consider a Markov step process $X = (X_t)_{t\geq 0}$ with state space $(E, \mathcal{E})$, fixed initial point $X_0 \equiv x_0$, with jump times $T_j$, $j \geq 1$, $T_0 := 0$, whose generator $\Pi_\vartheta(\cdot,\cdot)$ depends on an unknown parameter $\vartheta \in \Theta$ ranging over an open subset $\Theta$ of $\mathbb{R}$. For $x, y \in E$, $\Pi_\vartheta(x, dy) = \mu_\vartheta(x)\pi_\vartheta(x, dy)$ where the function $\mu_\vartheta : E \to (0,\infty)$ is measurable, and where $\pi_\vartheta(\cdot,\cdot)$ is a transition probability on $(E,\mathcal{E})$: so under $\vartheta$, $\mathcal{L}(X_{T_{j+1}} \mid X_{T_j} = x) = \pi_\vartheta(x,\cdot)$ and $\mathcal{L}(T_{j+1} - T_j \mid X_{T_j} = x)$ is the exponential law with parameter $\mu_\vartheta(x)$. We assume that

> for all $\vartheta \in \Theta$, the process $X$ is recurrent in the sense of Harris under $\vartheta$, positive or null, with invariant measure $m_\vartheta$.

By Harris recurrence (see Azéma *et al.* (1969)), subsets $U \in \mathcal{E}$ of positive invariant measure are visited a.s. infinitely often by the process $X$, for arbitrary choice of the starting point $x_0$. The state space $E$ is a Borel subset of $\bar{E} = \mathbb{R}^l$, with $\mathcal{E}$ the trace of $\bar{\mathcal{E}} = \mathcal{B}(\mathbb{R}^l)$ on $E$. $(\Omega, \mathcal{F}_\infty)$ denotes the canonical path space of $X$ (the space of all piecewise constant right-continuous functions $\omega : [0,\infty) \to E$ without accumulation of jumps in finite time intervals, starting at $\omega(0) = x_0$) endowed with the filtration $\mathbb{F} = (\mathcal{F}_t)_{t\geq 0}$ generated by the canonical process; for all $\vartheta \in \Theta$, there is a unique probability $P_\vartheta$ on $(\Omega, \mathcal{F}_\infty)$ such that the canonical process on $(\Omega, \mathbb{F}, \mathcal{F}_\infty)$ is Markov with generator $\Pi_\vartheta(\cdot,\cdot)$, and we identify $X$ with this canonical process.

Let $(S_n)_{n\geq 1}$ denote a sequence of $\mathbb{F}$-stopping times such that

$$P_\vartheta(S_n < \infty) = 1, \quad n \geq 1,\ \vartheta \in \Theta, \quad S_n(\omega) \uparrow \infty \quad \text{as} \quad n \to \infty, \quad \forall \omega \in \Omega.$$

At each stage $n$, the trajectory of $X$ being observed up to time $S_n$, we construct a minimum distance estimator (MDE) $\vartheta^*_{S_n}$ for the unknown parameter $\vartheta \in \Theta$. Select a subset $U \in \mathcal{E}$ (we assume that such subsets exist) such that for all $\vartheta \in \Theta$

$$0 < m_\vartheta(U) < \infty, \qquad \sup_{x \in U} \mu_\vartheta(x) < \infty \tag{1.1}$$

and a finite measure $F$ on $(\bar{E}, \bar{\mathcal{E}}) = (\mathbb{R}^l, \mathcal{B}(\mathbb{R}^l))$; w.l.o.g., we take $F(\bar{E}) = 1$. Write $\|\cdot\|$, $\langle\cdot,\cdot\rangle$ for the norm and the scalar product in the space $L^2(F)$ of measurable functions $f : \bar{E} \to \mathbb{R}$ with $\int f^2(v)F(dv) < \infty$. Introduce

$$\tau^* = \inf\{t > 0 : X_t \in U\}, \qquad A^*_t = \int_0^t 1_U(X_s)ds, \qquad t \geq 0$$

and define for $t \geq 0$, $v \in \bar{E}$, $\xi \in \Theta$, $\omega \in \Omega$

$$L_t(\xi, v) = \left( \int_0^t 1_U(X_s)\Pi_\xi(X_s, v)ds / A^*_t \right) 1_{\{\tau^* < t\}}$$

$$\hat{L}_t(v) = \left( \sum_{j \geq 1} 1_{\{T_j \leq t\}} 1_{U \times (-\infty, v]}(X_{T_{j-1}}, X_{T_j}) / A^*_t \right) 1_{\{\tau^* < t\}}$$

(as in classical statistics of i.i.d. variables, we write indistinctly $H$ for a finite measure on $(E, \mathcal{E})$ and for its distribution function $v \to H(v) := H((-\infty, v] \cap E)$ on $\bar{E}$). By choice of $\Omega$ and by (1.1), $v \to L_t(\xi, v)(\omega)$ and $v \to \hat{L}_t(v)(\omega)$ are distribution functions of finite measures on $(E, \mathcal{E})$, for all $\omega$, $t$, $\xi$. Given $U$, $F$, $(S_n)_n$, we estimate the unknown parameter governing the generator of $X$ at stage $n$ by

$$\vartheta^*_{S_n} = \arg \inf_{\xi \in \Theta} \|\hat{L}_{S_n}(\cdot) - L_{S_n}(\xi, \cdot)\|_{L^2(F)} \tag{1.2}$$

with arbitrary $\mathcal{F}_{S_n}$-measurable choice (such choices do exist) of an argmin on the set $A_n$ where $S_n$ is finite and where $\xi \to \|\hat{L}_{S_n}(\cdot) - L_{S_n}(\xi, \cdot)\|_{L^2(F)}$ has a minimum inside $\Theta$, and with $\vartheta^*_{S_n} = \vartheta_0$ on $A^c_n$ for an arbitrary fixed value $\vartheta_0$ in $\Theta$; note that the set $A_n$ belongs to $\mathcal{F}_{S_n}$. For details, see Höpfner and Kutoyants ((1997), Lemma 1).

## 2. An expansion

We list the conditions to be used in the sequel.

Regularity Condition (R). For some $r \geq 1$ and all $\vartheta \in \Theta$, we have signed finite kernels $\Pi^{(m)}_\vartheta(\cdot,\cdot)$ on $(E, \mathcal{E})$, $1 \leq m \leq r+2$, which are $m$-th derivatives of $\Pi^{(0)}_\vartheta(\cdot,\cdot) := \Pi_\vartheta(\cdot,\cdot)$ in the sense of the following properties i)–iii):

i) for $x \in E$ and $0 \leq m \leq r+2$, both positive and negative parts of the finite measures $\Pi^{(m)}_\vartheta(x,\cdot)$ have continuous distribution function on $\bar{E} = \mathbb{R}^l$;

ii) with $|H|$ denoting total variation of a finite measure $H$, there are $M_\vartheta < \infty$ and $\varrho_\vartheta > 0$ such that

$$\sup_{\xi:|\xi-\vartheta|\le\varrho_\vartheta} \sup_{x\in U} |\Pi_\xi^{(m)}|(x,E) \le M_\vartheta, \qquad 0\le m\le r+2;$$

iii) for $0 \le m \le r+1$ and $\xi \to \vartheta$,

$$\sup_{x\in U, v\in\bar{E}} |\Pi_\xi^{(m)}(x,v) - \Pi_\vartheta^{(m)}(x,v) - (\xi-\vartheta)\Pi_\vartheta^{(m+1)}(x,v)| = o(|\xi-\vartheta|)$$

and $\sup_{x\in U, v\in\bar{E}} |\Pi_\xi^{(r+2)}(x,v) - \Pi_\vartheta^{(r+2)}(x,v)| \to 0$.

TIGHTNESS CONDITION (T). For all $\vartheta\in\Theta$, there is a sequence $\delta_n = \delta_n(\vartheta) \downarrow 0$ such that $\mathcal{L}(\delta_n^{-1}\|\hat{L}_{S_n}(\cdot) - L_{S_n}(\vartheta,\cdot)\| \mid P_\vartheta)$, $n\ge 1$, is tight in $\mathbb{R}$.

For a large class of sequences $(S_n)_n$, this condition is verified by exploiting the martingale structure of $(\hat{L}_t(v) - L_t(\vartheta)(v))_{t\ge0}$ under $P_\vartheta$ for $v\in\bar{E}$ fixed, using arguments similar to Höpfner and Kutoyants ((1997), proof of Theorem 3).

MONOTONICITY CONDITION (M). Let $V$ denote the support of the measure $F$ in $\bar{E}$. For all $\vartheta\in\Theta$, there is some $\kappa_\vartheta > 0$ such that

$$\sup_{\xi:|\xi-\vartheta|\le\varrho_\vartheta} \sup_{x\in U, v\in V} \Pi_\xi^{(1)}(x,v) \le -\kappa_\vartheta < 0.$$

Condition (R) ensures that for all $n\ge1$, $\omega\in\Omega$ fixed, the map $\vartheta\to L_{S_n}(\vartheta,\cdot)(\omega)$ taking values in the Banach space of continuous bounded functions on $\mathbb{R}$ has $r+2$ continuous derivatives on $\Theta$:

LEMMA 2.1. *Under Condition* (R), *consider processes*

$$L_t^{(m)}(\xi,\cdot) := \left(\int_0^t 1_U(X_s)\Pi_\xi^{(m)}(X_s,\cdot)ds/A_t^*\right) 1_{\{\tau^*<t\}}, \qquad t\ge0$$

*for* $1\le m\le r+2$, *write* $L_t^{(0)}(\vartheta,\cdot) := L_t(\vartheta,\cdot)$. *Then* $L_t^{(m)}(\xi,\cdot)(\omega)$ *are signed finite measures on* $(E,\mathcal{E})$, *with total variation bounded by* $M_\vartheta$ *if* $|\xi-\vartheta|\le\varrho_\vartheta$, *and with continuous distribution function. Uniformly in* $t\ge0$, $\omega\in\Omega$, *as* $\xi\to\vartheta$

$$\sup_{v\in\bar{E}} |L_t^{(m)}(\xi,v)J - L_t^{(m)}(\vartheta,v) - (\xi-\vartheta)L_t^{(m+1)}(\vartheta,v)| = o(|\xi-\vartheta|)$$

*if* $0\le m\le r+1$, *and* $\sup_{v\in\bar{E}} |L_t^{(r+2)}(\xi,v)J - L_t^{(r+2)}(\vartheta,v)| \to 0$.

LEMMA 2.2. *Assume Conditions* (R), (T), (M) *for* $U$, $F$, $(S_n)_n$. *Fix a point* $\vartheta\in\Theta$, *choose sequences* $\varepsilon_n\downarrow0$, $u_n\downarrow0$ (*depending possibly on* $\vartheta$) *such that*

$\delta_n(\vartheta) < \varepsilon_n < u_n \leq 1$ *and* $u_n/\varepsilon_n \to \infty$, *write* $Z_n(\vartheta, \cdot) = \delta_n^{-1}(\hat{L}_{S_n}(\cdot) - L_{S_n}(\vartheta, \cdot))$. *Define*

$$B_{n,\vartheta} := \{\tau^* < S_n < \infty\} \cap \left\{ \|Z_n(\vartheta, \cdot)\| \leq \frac{u_n}{\varepsilon_n} \frac{\kappa_\vartheta^2}{4M_\vartheta} \right\}$$

*and take* $n_0 = n_0(\vartheta)$ *large enough to have*

$$u_n M_\vartheta^2 \leq \frac{\kappa_\vartheta^2}{4}, \qquad u_n \leq \varrho_\vartheta, \quad \forall n \geq n_0.$$

*For* $\omega \in B_{n,\vartheta}$ *fixed and suppressed from the notation, consider*

$$(-u_n, u_n) \times (-\varepsilon_n, \varepsilon_n) \supset (u, \varepsilon) \mapsto \Phi_n(u, \varepsilon)$$

*defined by*

$$\Phi_n(u, \varepsilon) := \varepsilon \langle Z_n(\vartheta), L_{S_n}^{(1)}(\vartheta + u) \rangle - \langle L_{S_n}(\vartheta + u) - L_{S_n}(\vartheta), L_{S_n}^{(1)}(\vartheta + u) \rangle.$$

*Then the following* i)–ii) *hold for all* $n \geq n_0$, $\omega \in B_{n,\vartheta}$.

i) *For* $\varepsilon \in (-\varepsilon_n, \varepsilon_n)$ *fixed*, $(-u_n, u_n) \ni u \to \Phi_n(u, \varepsilon)$ *has derivative bounded away from* 0 *on* $(-u_n, u_n)$, *and has a zero at some* $u_n^*(\varepsilon) \in (-u_n, u_n)$.

ii) *The function* $(-\varepsilon_n, \varepsilon_n) \ni \varepsilon \to u_n^*(\varepsilon)$ *has continuous derivatives* $\chi_{n,j} = \frac{d^j u_n^*}{d\varepsilon^j}$ *on* $(-\varepsilon_n, \varepsilon_n)$ *up to order* $j = r + 1$, *and*

$$\sup_{|\eta| < \varepsilon_n} |\chi_{n,j}(\eta)| \leq \frac{u_n^{2j-1}}{\varepsilon_n^{2j-1}} C_j(\vartheta) \tag{2.1}$$

*for functions* $C_j : \Theta \to (0, \infty)$ *which are deterministic and independent of* $n$, $u_n$, $\varepsilon_n$, ..., *for* $j \leq r + 1$. *If*

$$\sup_{\vartheta \in \Theta} M_\vartheta < \infty, \qquad \inf_{\vartheta \in \Theta} \kappa_\vartheta > 0, \tag{2.2}$$

*the functions* $C_j(\cdot)$ *can be replaced by constants not depending on* $\vartheta$.

Note that by Condition (R), the MDE error $\vartheta_{S_n}^* - \vartheta$ provides one root of

$$\Theta - \vartheta \ni u \to \Phi_n(u, \delta_n) = \langle \hat{L}_{S_n} - L_{S_n}(\vartheta + u), L_{S_n}^{(1)}(\vartheta + u) \rangle$$

if $\omega \in A_n$, the set occurring in the definition (1.2) of $\vartheta_{S_n}^*$.

THEOREM 2.1. *Assume Conditions* (R), (T), (M) *for* $U$, $F$, $(S_n)_n$. *Fix* $\vartheta \in \Theta$. *With notations of Lemma* 2.2, *we choose* $u_n$, $\varepsilon_n$ *such that* $\frac{u_n^{2r+1}}{\varepsilon_n^{2r+1}} = \delta_n^{-1/2}$, *with* $\delta_n = \delta_n(\vartheta)$. *Then the following holds for all* $n \geq n_0(\vartheta)$.

$$P_\vartheta \left( \left| \vartheta_{S_n}^* - \left\{ \vartheta + \sum_{j=1}^{r} \frac{1}{j!} Y_{n,j}(\vartheta) \delta_n^j \right\} \right| \geq \frac{C_{r+1}(\vartheta)}{(r+1)!} \delta_n^{r+1/2} \right) \leq P_\vartheta(B_{n,\vartheta}^c) \tag{2.3}$$

*where $Y_{n,j}(\vartheta)$ are $\mathcal{F}_{S_n}$-measurable random variables such that $Y_{n,j}(\vartheta) = \chi_{n,j}(0)$ on $B_{n,\vartheta}$, for $1 \le j \le r$: for $j = 1, 2$,*

$$Y_{n,1}(\vartheta) = \frac{\langle Z_n(\vartheta), L^{(1)}_{S_n}(\vartheta)\rangle}{\|L^{(1)}_{S_n}(\vartheta)\|^2} 1_{\{\tau^* < S_n < \infty\}},$$

$$Y_{n,2}(\vartheta) = 2\frac{\langle Z_n(\vartheta), L^{(2)}_{S_n}(\vartheta)\rangle}{\|L^{(1)}_{S_n}(\vartheta)\|^2} Y_{n,1}(\vartheta) - 3\frac{\langle L^{(1)}_{S_n}(\vartheta), L^{(2)}_{S_n}(\vartheta)\rangle}{\|L^{(1)}_{S_n}(\vartheta)\|^2} Y^2_{n,1}(\vartheta)$$

*and expressions for $3 \le j \le r$ can be obtained from* (3.2)–(3.6) *below.*

In the next result we specialize to a particular observation scheme which satisfies Condition (T), already used in Höpfner and Kutoyants ((1997), Section 4), and put

$$S_n = U_n := \inf\{t > 0 : A^*_t > n\}, \qquad \delta_n := n^{-1/2}, \ \ n \ge 1. \tag{2.4}$$

THEOREM 2.2. *With notations and assumptions of Lemma 2.2 and Theorem 2.1, consider $(S_n)_n$ and $(\delta_n)_n$ given by (2.4). Then*

i) *for every $p > 2$, there are (deterministic) functions $C'_p : \Theta \to (0, \infty)$ such that*

$$P_\vartheta(B^c_{n,\vartheta}) \le n^{-p/(4(2r+1))} C'_p(\vartheta), \qquad n \ge 1, \ \ \vartheta \in \Theta;$$

*under (2.2), these $C'_p(\cdot)$ can be replaced by constants depending only on $p$;*

ii) *under (2.2) and if $\inf_{\vartheta\in\Theta} \varrho_\vartheta > 0$, $n_0(\cdot)$ in Lemma 2.2 can be choosen independent of $\vartheta$, yielding*

$$P_\vartheta\left( \left| \vartheta^*_{S_n} - \left\{ \vartheta + \sum_{j=1}^{r} \frac{1}{j!} Y_{n,j}(\vartheta) n^{-j/2} \right\} \right| \ge \frac{C n^{-r/2-1/4}}{(r+1)!} \right) \le C'_p n^{-p/(4(2r+1))} \tag{2.5}$$

*for certain constants $C$, $C'_p$, uniformly in $\vartheta \in \Theta$ and $n \ge 1$.*

## 3. Proofs

PROOF OF LEMMA 2.1. Immediate from Condition (R) and definition of $L^{(m)}_t(\vartheta, \cdot)$.

PROOF OF LEMMA 2.2. First, we prove i). For $\omega \in \{\tau^* < S_n < \infty\}$ fixed, for $|\varepsilon| < \varepsilon_n$, $\Phi_n(\cdot, \varepsilon)$ has derivative

$$\frac{\partial \Phi_n(u, \varepsilon)}{\partial u} = \varepsilon \langle Z_n(\vartheta), L^{(2)}_{S_n}(\vartheta + u)\rangle - \|L^{(1)}_{S_n}(\vartheta + u)\|^2$$
$$\langle L_{S_n}(\vartheta + u) - L_{S_n}(\vartheta), L^{(2)}_{S_n}(\vartheta + u)\rangle.$$

By Conditions (R), (M) and the mean value theorem, uniformly in $|\varepsilon| < \varepsilon_n$, $|u| < u_n$,

$$\frac{\partial \Phi_n(u,\varepsilon)}{\partial u} \le \varepsilon_n M_\vartheta \|Z_n(\vartheta)\| - \kappa_\vartheta^2 + u_n M_\vartheta^2$$

and thus, by choice of $n_0$ and $B_{n,\vartheta}$,

$$\sup_{|\varepsilon|<\varepsilon_n} \sup_{|u|<u_n} \frac{\partial \Phi_n(u,\varepsilon)}{\partial u} \le -\frac{\kappa_\vartheta^2}{2}.$$

The same argument shows that

$$\sup_{|\varepsilon|<\varepsilon_n} |\Phi_n(0,\varepsilon)| \le \varepsilon_n M_\vartheta \|Z_n(\vartheta)\| \le u_n \frac{\kappa_\vartheta^2}{4}$$

on $B_n$, for $n \ge n_0$. Both last assertions yield i). We turn to ii). For $\omega \in \{S_n < \infty\}$ fixed, $(u,\varepsilon) \to \Phi_n(u,\varepsilon)$ has continuous derivatives up to order $r+1$, by Condition (R). The implicit function theorem shows that for $\omega \in B_n$, the function $(-\varepsilon_n, \varepsilon_n) \ni \varepsilon \to u_n^*(\varepsilon)$ has $r+1$ continuous derivatives on $(-\varepsilon_n, \varepsilon_n)$. Since $u_n^*(0) = 0$, we have an expansion

$$u_n^*(\varepsilon) = \sum_{j=1}^{r} \frac{1}{j!} \frac{d^j u_n^*}{d\varepsilon^j}(0)\varepsilon^j + \frac{1}{(r+1)!} \frac{d^{r+1} u_n^*}{d\varepsilon^{r+1}}(\eta_{n,r+1,\varepsilon})\varepsilon^{r+1} \tag{3.1}$$

on $(-\varepsilon_n, \varepsilon_n)$, valid for $\omega \in B_n$ and $n \ge n_0$, where $\eta_{n,r+1,\varepsilon}$ is between 0 and $\varepsilon$. Write $\chi_{n,j} := \frac{\partial^j u_n^*}{\partial \varepsilon^j}$, $j \le r+1$. For $\omega \in B_n$ fixed and $|\varepsilon| < \varepsilon_n$,

$$\frac{du_n^*}{d\varepsilon}(\varepsilon) = -\frac{\dfrac{\partial \Phi_n}{\partial \varepsilon}}{\dfrac{\partial \Phi_n}{\partial u}}(u_n^*(\varepsilon),\varepsilon). \tag{3.2}$$

Write shortly $\varphi_{n,i,j} = \frac{\partial^{i+j}\Phi_n}{\partial u^i \partial \varepsilon^j}$, then

$$\chi_{n,2}(\varepsilon) = \frac{2\varphi_{n,1,1}\varphi_{n,1,0}\varphi_{n,0,1} - \varphi_{n,0,2}\varphi_{n,1,0}^2 - \varphi_{n,2,0}\varphi_{n,0,1}^2}{(\varphi_{n,1,0})^3}(u_n^*(\varepsilon),\varepsilon) \tag{3.3}$$

and higher derivatives have the structure

$$\chi_{n,k}(\varepsilon) = \sum_{I_k} c_{i_1,\dots,i_{2k-1},j_1,\dots,j_{2k-1}} \left( \prod_{l=1}^{2k-1} \frac{\varphi_{n,i_l,j_l-i_l}}{\varphi_{n,1,0}} \right) (u_n^*(\varepsilon),\varepsilon) \tag{3.4}$$

where summation is over the set $I_k$ of nonnegative $i_1, \dots, i_{2k-1}, j_1, \dots, j_{2k-1}$ such that $i_1 + \cdots + i_{2k-1} = 2k-2$, $j_1 + \cdots + j_{2k-1} = 3k-2$, and where $c_{i_1,\dots,i_{2k-1},j_1,\dots,j_{2k-1}}$

are suitable constants. For $k = 1, 2$, this is immediate from (3.2) and (3.3), and follows in general from

$$\frac{d}{d\varepsilon}\frac{\varphi_{n,i_l,j_l-i_l}}{\varphi_{n,1,0}}(u_n^*(\cdot),\cdot) = \varphi_{n,1,0}^{-2}\left\{\varphi_{n,1,0}\left[\varphi_{n,i_l+1,j_l-i_l}\left(-\frac{\varphi_{n,0,1}}{\varphi_{n,1,0}}\right)+\varphi_{n,i_l,j_l-i_l+1}\right]\right.$$
$$\left.-\varphi_{n,i_l,j_l-i_l}\left[\varphi_{n,2,0}\left(-\frac{\varphi_{n,0,1}}{\varphi_{n,1,0}}\right)+\varphi_{n,1,1}\right]\right\}$$
$$= \varphi_{n,1,0}^{-3}\{\varphi_{n,i_l,j_l-i_l}(\varphi_{n,2,0}\varphi_{n,0,1}-\varphi_{n,1,0}\varphi_{n,1,1})$$
$$+\varphi_{n,i_l+1,j_l-i_l}(-\varphi_{n,1,0}\varphi_{n,0,1})$$
$$+\varphi_{n,i_l,j_l-i_l+1}(\varphi_{n,1,0})^2\}.$$

Here, by definition of $\Phi_n(u,\varepsilon)$, one has $\varphi_{n,i,j}\equiv 0$ whenever $j \geq 2$,

$$\varphi_{n,l,1}(u,\varepsilon) = \langle Z_n(\vartheta), L_{S_n}^{(l+1)}(\vartheta+u)\rangle \tag{3.5}$$

for all $l$, and with $C_{l,j} = \frac{l!}{j!(l-j)!}$,

$$\begin{aligned}\varphi_{n,l,0}(u,\varepsilon) &= \varepsilon\langle Z_n(\vartheta), L_{S_n}^{(l+1)}(\vartheta+u)\rangle \\ &\quad - \langle L_{S_n}(\vartheta+u) - L_{S_n}(\vartheta), L_{S_n}^{(l+1)}(\vartheta+u)\rangle \\ &\quad - \sum_{j=1}^{l} C_{l,j}\langle L_{S_n}^{(j)}(\vartheta+u), L_{S_n}^{(l+1-j)}(\vartheta+u)\rangle.\end{aligned} \tag{3.6}$$

It remains to show (2.1). Using the above representation of $\chi_{n,k}$ and $\varphi_{n,l,j}$, where every $\varphi_{n,l,0}$, $\varphi_{n,l,1}$ contains exactly one factor $\langle Z_n(\vartheta), L_{S_n}^{l+1}(\vartheta)\rangle$, it is clear from Condition (R) that for every $k \leq r+1$ there is some polynomial $(x,y)\to p_k(x,y)$ with nonnegative coefficients, with $x \to p_k(x,y)$ having degree at most $2k-1$, such that

$$\sup_{|\eta|<\varepsilon_n}|\chi_{n,k}(\eta)| \leq \frac{p_k(\|Z_n(\vartheta)\|M_\vartheta, M_\vartheta)}{\inf_{|\varepsilon|<\varepsilon_n,|v|<u_n}|\varphi_{n,1,0}(v,\varepsilon)|^{2k-1}} \tag{3.7}$$

where $p_k(\cdot,\cdot)$ depends only on the structure of the derivatives of the implicit function, and not on $\omega\in B_n$, $n\geq n_0$, or on choice of $u_n\leq 1$, $\varepsilon_n\leq 1$. From step 1 above, $\inf_{|\varepsilon|<\varepsilon_n,|v|<u_n}|\varphi_{n,1,0}(v,\varepsilon)| \geq \frac{\kappa_\vartheta^2}{2}$ on $B_n$, for $n\geq n_0$, and also $\|Z_n(\vartheta)\| \leq \frac{u_n}{\varepsilon_n}\frac{\kappa_\vartheta^2}{4M_\vartheta}$. Since $p_k(\cdot,y)$ has degree at most $2k-1$, we obtain

$$\sup_{|\eta|<\varepsilon_n}|\chi_{n,k}(\eta)| \leq \frac{u_n^{2k-1}}{\varepsilon_n^{2k-1}}\frac{p_k\left(\frac{\kappa_\vartheta^2}{4}, M_\vartheta\right)}{\left(\frac{\kappa_\vartheta^2}{2}\right)^{2k-1}} =: \frac{u_n^{2k-1}}{\varepsilon_n^{2k-1}}C_k(\vartheta); \tag{3.8}$$

clearly $C_k(\cdot)$ can be replaced by a constant if (2.2) holds. This is (2.1), and Lemma 2.2 is proved.

PROOF OF THEOREM 2.1. Put $C_{n,\vartheta} := A_n \cap \{|\vartheta^*_{S_n} - \vartheta| < u_n\}$, where $A_n$ is as in (1.2). On the event $B_{n,\vartheta} \cap C_{n,\vartheta}$, since $\delta_n < \varepsilon_n$, the MDE estimation error $\vartheta^*_{S_n} - \vartheta$ is the unique zero $u^*_n(\delta_n)$ of the function $u \to \Phi_n(u, \delta_n)$ in $(-u_n, u_n)$. By assumption, $\frac{u_n^{2r+1}}{\varepsilon_n^{2r+1}} = \delta_n^{-1/2}$. So (3.1) + (3.8) yield (2.3), but with $P_\vartheta(B^c_{n,\vartheta} \cup C^c_{n,\vartheta})$ replacing $P_\vartheta(B^c_{n,\vartheta})$ on the r.h.s of (2.3). We show that under our conditions, $C^c_{n,\vartheta}$ is contained in $B^c_{n,\vartheta}$ for $n \geq n_0$. By definition of $C_{n,\vartheta}$, a standard argument (cf. Höpfner and Kutoyants (1997), Proof of Lemma 2) shows that $C^c_{n,\vartheta}$ is contained in the event

$$\{2\|Z_n(\vartheta)\| \geq \inf_{\xi:|\xi-\vartheta|>u_n} \delta_n^{-1}\|L_{S_n}(\xi) - L_{S_n}(\vartheta)\|\} \cup \{S_n = \infty\}$$

which by Conditions (M) and (R) is a subset of

$$\{2\|Z_n(\vartheta)\| \geq \delta_n{}^{1} u_n \kappa_\vartheta\} \cup \{S_n = \infty\};$$

since $\varepsilon_n > \delta_n$, this shows $C^c_{n,\vartheta} \subset B^c_{n,\vartheta}$.

PROOF OF THEOREM 2.2. We prove i); ii) is then immediate from Lemma 2.2 and (2.3). Choose $S_n = U_n$ and $\delta_n = n^{-1/2}$ as defined in (2.4). Below, we shall show that for every $p > 2$, there is some finite constant $C''_p$ depending only on $p$ such that

$$E_\vartheta(|n^{1/2}(\hat{L}_{U_n}(v) - L_{U_n}(\vartheta, v))|^p) \leq C''_p(1 + M_\vartheta)^{p/2} \quad \forall v \in \mathbb{R}^l J, \qquad n \geq 1. \tag{3.9}$$

Since $F$ is a probability measure,

$$\begin{aligned} E_\vartheta(\|Z_n(\vartheta)\|^p) &= E_\vartheta(\|n^{1/2}(\hat{L}_{U_n} - L_{U_n}(\vartheta))\|^p_{L^2(F)}) \\ &\leq E_\vartheta\left(\int |n^{1/2}(\hat{L}_{U_n}(v) - L_{U_n}(\vartheta, v))|^p F(dv)\right) \leq C''_p(1 + M_\vartheta)^{p/2} \end{aligned}$$

and thus, for every $p > 2$ (clearly $\tau^* < U_n < \infty$),

$$P_\vartheta(B^c_{n,\vartheta}) \leq \frac{E_\vartheta(\|Z_n(\vartheta)\|^p)}{\left(\dfrac{u_n}{\varepsilon_n}\dfrac{\kappa^2_\vartheta}{4M_\vartheta}\right)^p} \leq \frac{\varepsilon^p_n}{u^p_n} \frac{C''_p(1 + M_\vartheta)^{p/2}}{\left(\dfrac{\kappa^2_\vartheta}{4M_\vartheta}\right)^p} =: \frac{\varepsilon^p_n}{u^p_n} C'_p(\vartheta)$$

which (by choice of $u_n$, $\varepsilon_n$ in Theorem 2.1) is the assertion; under (2.2), $C'_p(\cdot)$ can be replaced by a constant not depending on $\vartheta$. It remains to prove (3.9). Write $\mu(ds, d\tilde{x})$ for the random measure $\sum_{j\geq 1, T_j<\infty} \varepsilon_{(T_j,(X_{T_{j-1}}, X_{T_j}))}(ds, d\tilde{x})$ on $(0, \infty) \times (E \times E)$, and $\nu_\vartheta(ds, d\tilde{x})$ for its compensator relative to $\mathbb{F}$ under $P_\vartheta$; let $M^v$ denote the martingale

$$M^v_t = \int_0^t \int_E \int_E 1_{U\times(-\infty,v]}(s, \tilde{x})(\mu - \nu_\vartheta)(ds, d\tilde{x}) = (1_{U\times(-\infty,v]} * (\mu - \nu_\vartheta))_t, \qquad t \geq 0.$$

As in Höpfner and Kutoyants ((1997), Section 4), write $U_{tn} := \inf\{s > 0 : A^*_s > tn\}$, $\tilde{M}^{(n)}_t := n^{-1/2} M^v_{U_{tn}}$; then $n^{1/2}(\hat{L}_{U_n}(v) - L_{U_n}(\vartheta, v)) = \tilde{M}^{(n)}_1$. The Burkholder-Davies-Gundy inequality (see Dellacherie and Meyer (1980), VII.92) gives some constant $C'''_p$ depending only on $p$ such that

$$\begin{aligned} E_\vartheta\left(\left|\sup_{t\le 1} M^{(n)}_t\right|^p\right) &\le C'''_p E_\vartheta([M^{(n)}]_1^{p/2}) \\ &= C'''_p E_\vartheta\left(\left(\frac{(1_{U\times(-\infty,v]} * \mu)_{U_n}}{n}\right)^{p/2}\right) \\ &\le C'''_p E_\vartheta\left(\left(\frac{(1_{U\times E} * \mu)_{U_n}}{n}\right)^{p/2}\right) \\ &\le C'''_p (1+M_\vartheta)^{p/2} \sup_{t\ge 1} E\left(\left(\frac{N_t}{t}\right)^{p/2}\right) =: C''_p (1+M_\vartheta)^{p/2} \end{aligned}$$

where $[\tilde{M}^{(n)}]$ denotes the quadratic variation process of $\tilde{M}^{(n)}$, and $N$ is a standard Poisson process. The last chain of inequalities—establishing (3.9)—uses that the counting process $1_{U\times E} * \mu$ is compensated by $1_{U\times E} * \nu_\vartheta = \int 1_U(X_s)\mu_\vartheta(X_s)ds \le M_\vartheta A^*$ under $P_\vartheta$. Defining $\tau_t = \inf\{s > 0 : (1_{U\times E} * \nu_\vartheta)_s > t\}$ such that $N = ((1_{U\times E} * \mu)_{\tau_t})_{t\ge 0}$ is standard Poisson, one has $U_n \le \tau_{n(1+M_\vartheta)}$. This gives (3.9), and Theorem 2.2 is proved.

## References

Azéma, J., Duflo, M. and Revuz, D. (1969). Measures invariantes des processus de Markov récurrents, Sém. Proba. III, *Lecture Notes in Math.*, **88**, 24–33, Springer, Berlin.

Dellacherie, C. and Meyer, P. A. (1980). *Probabilités et Potentiel*, Vol. 2, Hermann, Paris.

Höpfner, R. and Kutoyants, Yu. A. (1997). On minimum distance estimation in recurrent Markov step processes I, *Scand. J. Statist.*, **24**, 61–79.

Kutoyants, Yu. A. (1983). Asymptotic expansion of the maximum likelihood estimate of the intensity parameter for inhomogeneous Poisson observations, *Transactions of 9th Prague Conference 1982*, 35–41, Academia, Prague.

Kutoyants, Yu. A. (1984). Expansion of a maximum likelihood estimate by diffusion powers, *Theory of Probability and Its Applications*, **29**, 465–477.

Kutoyants, Yu. A. (1994). *Identification of Dynamical Systems with Small Noise*, Kluwer, Dordrecht.

Millar, P. W. (1984). A general approach to the optimality of minimum distance estimators, *Trans. Amer. Math. Soc.*, **286**, 377–418.