

# *Polynomials under Ornstein–Uhlenbeck noise and an application to inference in stochastic Hodgkin–Huxley systems*

**Reinhard Höpfner**

## **Statistical Inference for Stochastic Processes**

An International Journal devoted to  
Time Series Analysis and the Statistics  
of Continuous Time Processes and  
Dynamical Systems

ISSN 1387-0874

Volume 24

Number 1

Stat Inference Stoch Process (2021)

24:35-59

DOI 10.1007/s11203-020-09226-0

**Your article is published under the Creative Commons Attribution license which allows users to read, copy, distribute and make derivative works, as long as the author of the original work is cited. You may self-archive this article on your own website, an institutional repository or funder's repository and make it publicly available immediately.**



# Polynomials under Ornstein–Uhlenbeck noise and an application to inference in stochastic Hodgkin–Huxley systems

Reinhard Höpfner<sup>1</sup> 

Received: 31 March 2020 / Accepted: 3 August 2020 / Published online: 13 September 2020  
© The Author(s) 2020

## Abstract

We discuss estimation problems where a polynomial  $s \rightarrow \sum_{i=0}^{\ell} \vartheta_i s^i$  with strictly positive leading coefficient is observed under Ornstein–Uhlenbeck noise over a long time interval. We prove local asymptotic normality (LAN) and specify asymptotically efficient estimators. We apply this to the following problem: feeding noise  $dY_t$  into the classical (deterministic) Hodgkin–Huxley model in neuroscience, with  $Y_t = \vartheta t + X_t$  and  $X$  some Ornstein–Uhlenbeck process with backdriving force  $\tau$ , we have asymptotically efficient estimators for the pair  $(\vartheta, \tau)$ ; based on observation of the membrane potential up to time  $n$ , the estimate for  $\vartheta$  converges at rate  $\sqrt{n^3}$ .

**Keywords** Diffusion models · Local asymptotic normality · Asymptotically efficient estimators · Degenerate diffusions · Stochastic Hodgkin–Huxley model

**Mathematics Subject Classification** 62F12 · 60J60

## 1 Introduction

Problems of parametric inference when we observe over a long time interval a process  $Y$  of type

$$dY_t = \left( \sum_{j=1}^m \vartheta_j f_j(t) - \tau Y_t \right) dt + \sqrt{c} dW_t, \quad \tau > 0$$

with unknown parameters  $(\vartheta_1, \dots, \vartheta_m)$  or  $(\vartheta_1, \dots, \vartheta_m, \tau)$  and with  $(f_1, \dots, f_m)$  a given set of functions have been considered in a number of papers; alternatively, such models can be written as

---

✉ Reinhard Höpfner  
hoepfner@mathematik.uni-mainz.de

<sup>1</sup> Institut für Mathematik, Universität Mainz, Staudingerweg 9, 55099 Mainz, Germany

$$Y_t = \sum_{j=1}^m \vartheta_j g_j(t) + X_t, \quad dX_t = -\tau X_t dt + \sqrt{c} dW_t$$

with related functions  $(g_1, \dots, g_m)$ . Also, driving Brownian motion  $W$  in the Ornstein–Uhlenbeck type equations above has been replaced by certain Lévy processes or by fractional Brownian motion. Many papers focus on orthonormal sets of periodic functions  $[0, \infty) \rightarrow \mathbb{R}$  with known periodicity. To determine estimators and limit laws for rescaled estimation errors in this case, periodicity allows to exploit ergodicity or stationarity with respect to the time grid of multiples of the periodicity. We mention Dehling et al. (2010), Franke and Kott (2013) and Dehling et al. (2017) where limit distributions for least squares estimators and maximum likelihood estimators are obtained. Rather than in asymptotic properties, Pchelintsev (2013) is interested in methods which allow to reduce squared risk—i.e. risk defined with respect to one particular loss function—uniformly over determined subsets of the parameter space, at fixed and finite sample size. Asymptotic efficiency of estimators is the topic of Höpfner and Kutoyants (2009), where sums  $\sum \vartheta_j f_j$  as above are replaced by periodic functions  $S$  of known periodicity whose shape depends on parameters  $(\vartheta_1, \dots, \vartheta_m)$ . When the parametrization is smooth enough, local asymptotic normality in the sense of LeCam (see Le Cam 1969; Hajek 1970; Davies 1985; Pfanzagl 1994; Le Cam and Yang 2002; with a different notion of local neighbourhood see Ibragimov and Khasminskii (1981) and Kutoyants (2004)) allows to identify a limit experiment with the following property: risk—asymptotically as the time of observation tends to  $\infty$ , and with some uniformity over small neighbourhoods of the true parameter—is bounded below by a corresponding minimax risk in a limit experiment. This assertion holds with respect to a broad class of loss functions.

With a view to an estimation problem which arises in stochastic Hodgkin–Huxley models and which we explain below, the present paper deals with parameter estimation when one observes a process  $Y$

$$Y_t = \sum_{j=0}^p \vartheta_j t^j + X_t, \quad dX_t = -\tau X_t dt + \sqrt{c} dW_t, \quad \tau > 0 \quad (1)$$

with leading coefficient  $\vartheta_p > 0$  so that paths of  $Y$  almost surely tend to  $\infty$ . Then good estimators for the parameters based on observation of  $Y$  up to time  $n$  show the following behaviour: whereas estimation of parameters  $\tau$  and  $\vartheta_0$  works at the ‘usual’ rate  $\sqrt{n}$ , parameters  $\vartheta_j$  with  $1 \leq j \leq p$  can be estimated at rate  $\sqrt{n^{2j+1}}$  as  $n \rightarrow \infty$ . With rescaled time  $(tn)_{t \geq 0}$ , we prove local asymptotic normality as  $n \rightarrow \infty$  in the sense of LeCam with local scale

$$\psi_n := \begin{pmatrix} \frac{1}{\sqrt{n}} & 0 & \dots & 0 \\ 0 & \frac{1}{\sqrt{n^3}} & \dots & 0 \\ 0 & \ddots & \ddots & 0 \\ 0 & \dots & \dots & \frac{1}{\sqrt{n^{2p+1}}} \\ 0 & \dots & 0 & \frac{1}{\sqrt{n}} \end{pmatrix}$$

and with limit information process  $J = (J_t)_{t \geq 0}$

$$J(t) = \frac{1}{c} \begin{pmatrix} \frac{\tau^2}{1} t & \frac{\tau^2}{2} t^2 & \cdots & \frac{\tau^2}{p+1} t^{p+1} & 0 \\ \frac{\tau^2}{2} t^2 & \frac{\tau^2}{3} t^3 & \cdots & \frac{\tau^2}{p+2} t^{p+2} & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \frac{\tau^2}{p+1} t^{p+1} & \frac{\tau^2}{p+2} t^{p+2} & \cdots & \frac{\tau^2}{2p+1} t^{2p+1} & 0 \\ 0 & \cdots & \cdots & 0 & \frac{c}{2\tau} t \end{pmatrix}, \quad t \geq 0$$

at every  $\theta := (\vartheta_0, \dots, \vartheta_p, \tau)$ . As a consequence of local asymptotic normality, there is a local asymptotic minimax theorem (Ibragimov and Khasminskii 1981; Davies 1985; Le Cam and Yang 2002; Kutoyants 2004; Höpfner 2014) which allows to identify optimal limit distributions for rescaled estimation errors in the statistical model (1); the theorem also specifies a particular expansion of rescaled estimation errors (in terms of the central sequence in local experiments at  $\theta$ ) which characterizes asymptotic efficiency. We can construct asymptotically efficient estimators for the model (1), and these estimators have a simple and explicit form.

We turn to an application of the results obtained for model (1). Consider the problem of parameter estimation in a stochastic Hodgkin–Huxley model for the spiking behaviour of a single neuron belonging to an active network

$$\begin{cases} dV_t = dY_t - F(V_t, n_t, m_t, h_t) dt \\ dn_t = \{\alpha_n(V_t)(1 - n_t) - \beta_n(V_t)(1 - n_t)\} dt \\ dm_t = \{\alpha_m(V_t)(1 - m_t) - \beta_m(V_t)(1 - m_t)\} dt \\ dh_t = \{\alpha_h(V_t)(1 - h_t) - \beta_h(V_t)(1 - h_t)\} dt \end{cases} \quad (2)$$

where input  $dY_t$  received by the neuron is modelled by the increments of the stochastic process

$$Y_t = \vartheta t + X_t, \quad dX_t = -\tau X_t dt + \sqrt{c} dW_t, \quad (\vartheta, \tau) \in (0, \infty)^2. \quad (3)$$

The functions  $F(\cdot, \cdot, \cdot, \cdot)$  and  $\alpha_j(\cdot), \beta_j(\cdot), j \in \{n, m, h\}$  are those of Izhikevich (2007) pp. 37–39. The stochastic model (2) extends the classical deterministic model with constant rate of input  $a > 0$

$$\begin{cases} dV_t = a dt - F(V_t, n_t, m_t, h_t) dt \\ dn_t = \{\alpha_n(V_t)(1 - n_t) - \beta_n(V_t)(1 - n_t)\} dt \\ dm_t = \{\alpha_m(V_t)(1 - m_t) - \beta_m(V_t)(1 - m_t)\} dt \\ dh_t = \{\alpha_h(V_t)(1 - h_t) - \beta_h(V_t)(1 - h_t)\} dt \end{cases} \quad (4)$$

by taking into account ‘noise’ in the dendritic tree where incoming excitatory or inhibitory spike trains emitted by a large number of other neurons in the network add up and decay. See Hodgkin and Huxley (1952), Izhikevich (2007), Ermentrout and Terman (2010) and the literature quoted there for the role of this model in neuroscience. Stochastic Hodgkin–Huxley models have been considered in Höpfner et al. (2016a), (2016b), (2017) and Holbach (2020). For suitable data sets, membrane potential data hint to the existence of a quadratic variation which indicates the need for a stochastic modelization.

In systems (2) or (4), the variable  $V = (V_t)_{t \geq 0}$  represents the membrane potential in the neuron; the variables  $j = (j_t)_{t \geq 0}, j \in \{n, m, h\}$ , are termed gating variables and represent—in the sense of averages over a large number of channels—opening and closing of ion channels of certain types. The membrane potential can be measured intracellularly in good time resolution whereas the gating variables in the Hodgkin–Huxley model are not accessible to direct measurement.

In a sense of equivalence of experiments as in Holbach (2019), the stochastic Hodgkin Huxley model (2)+(3) corresponds to a submodel of (1). This is of biological importance. Under the assumption that the stochastic model admits a fixed starting point which does not depend on  $\theta := (\vartheta, \tau)$ , we can estimate the components  $\vartheta > 0$  and  $\tau > 0$  of the unknown parameter  $\theta = (\vartheta, \tau)$  in equations (2)+(3) from the evolution of the membrane potential alone, and have at our disposal simple and explicit estimators  $\check{\theta}(n) = (\check{\vartheta}(n), \check{\tau}(n))$  with the following two properties (i) and (ii).

(i) With local parameter  $h = (h_1, h_2)$  parametrizing shrinking neighbourhoods of  $\theta = (\vartheta, \tau)$ , risks

$$\sup_{|h| \leq C} E_{(\vartheta+h_1/\sqrt{n^3}, \tau+h_2/\sqrt{n})} \left( L \left( \frac{\sqrt{n^3}(\check{\vartheta}(n) - (\vartheta + h_1/\sqrt{n^3}))}{\sqrt{n}(\check{\tau}(n) - (\tau + h_2/\sqrt{n}))} \right) \right) \quad (5)$$

converge as  $n \rightarrow \infty$  to

$$E \left( L \left( \frac{\frac{3\sqrt{c}}{\tau} \int_0^1 s d\tilde{W}_s^{(1)}}{\sqrt{2\tau} \tilde{W}_1^{(2)}} \right) \right) \quad (6)$$

where  $\tilde{W} = (\tilde{W}^{(1)}, \tilde{W}^{(2)})$  is two-dimensional standard Brownian motion. Here  $C$  is an arbitrary constant, and  $L: \mathbb{R}^2 \rightarrow [0, \infty)$  any loss function which is continuous, subconvex and bounded.

(ii) We can compare the sequence of estimators  $\check{\theta}(n) = (\check{\vartheta}(n), \check{\tau}(n))$  for  $\theta = (\vartheta, \tau)$  in (5) to arbitrary estimator sequences  $T(n) = (T^{(1)}(n), T^{(2)}(n))$  which can be defined from observation of the membrane potential up to time  $n$ , provided their rescaled estimation errors—using the same norming as in (5)—are tight. For all such estimator sequences,

$$\sup_{C \uparrow \infty} \liminf_{n \rightarrow \infty} \sup_{|h| \leq C} E_{(\vartheta+h_1/\sqrt{n^3}, \tau+h_2/\sqrt{n})} \left( L \left( \frac{\sqrt{n^3}(T^{(1)}(n) - (\vartheta + h_1/\sqrt{n^3}))}{\sqrt{n}(T^{(2)}(n) - (\tau + h_2/\sqrt{n}))} \right) \right)$$

is always greater or equal than the limit in (6). This is the assertion of the local asymptotic minimax theorem. It makes sure that asymptotically as  $n \rightarrow \infty$ , it is impossible to outperform the simple and explicit estimator sequence  $\check{\theta}(n) = (\check{\vartheta}(n), \check{\tau}(n))$  which we have at hand.

The paper is organized as follows. Section 2 collects for later use convergence results for certain functionals of the Ornstein–Uhlenbeck process. Section 3 deals with local asymptotic normality (LAN) for the model (1): proposition 1 and theorem 1 in Sect. 3.1 prove LAN, the local asymptotic minimax theorem is corollary 1 in Sect. 3.1; we introduce and investigate estimators for  $\theta = (\vartheta, \tau)$  in Sects. 3.2 and 3.3; theorem 2 in Sect. 3.4 states their asymptotic efficiency. The application to parameter estimation in the stochastic Hodgkin–Huxley model (2)+(3) based on observation of the membrane potential is the topic of the final Sect. 4: see theorem 3 and corollary 2 there.

## 2 Functionals of the Ornstein Uhlenbeck process

We state for later use properties of some functionals of the Ornstein Uhlenbeck process

$$dX_t = -\tau X_t dt + \sigma dW_t, \quad t \geq 0 \quad (7)$$

with fixed starting point  $x_0 \in \mathbb{R}$ .  $\tau > 0$  and  $\sigma > 0$  are fixed, and  $\nu := \mathcal{N}(0, \frac{\sigma^2}{2\tau})$  is the invariant measure of the process in (7);  $X$  is defined on some  $(\Omega, \mathcal{A}, P)$ .

**Lemma 1** For  $X$  defined by (7), for every  $f \in L^1(\nu)$  and  $\ell \in \mathbb{N}$ , we have almost sure convergence as  $r \rightarrow \infty$

$$\frac{\ell}{r^\ell} \int_0^r s^{\ell-1} f(X_s) ds \longrightarrow \nu(f).$$

**Proof** (Höpfner and Kutoyants 2011, lemma 2.2; Holbach 2019, lemma 2.5, compare to Bingham et al. 1987 theorem 1.6.4 p. 33)

(1) We consider functions  $f \in L^1(\nu)$  which satisfy  $\nu(f) \neq 0$ . The case  $\ell = 1$  is the well known ratio limit theorem for additive functionals of the ergodic diffusion  $X$  (Kutoyants 2004; Höpfner 2014 p. 214). Assuming that the assertion holds for  $\ell = \ell_0 \in \mathbb{N}$ , define  $A_r = \int_0^r s^{\ell_0-1} f(X_s) ds$ . Stieltjes product formula for semimartingales with paths of locally bounded variation yields

$$\int_0^r s dA_s = r A_r - \int_0^r A_s ds, \quad 0 < r < \infty.$$

Under our assumption, both terms on the right hand side are of stochastic order  $O(r^{\ell_0+1})$ : since  $\frac{\ell_0}{s^{\ell_0}} A_s$  converges to  $\nu(f) \neq 0$  almost surely as  $s \rightarrow \infty$ , the second term on the right hand side behaves as  $\nu(f) \int_0^r \frac{s^{\ell_0}}{\ell_0} ds = \nu(f) \frac{r^{\ell_0+1}}{\ell_0(\ell_0+1)}$  as  $r \rightarrow \infty$ ; the first term on the right hand side behaves as  $\nu(f) \frac{r^{\ell_0+1}}{\ell_0}$ . This proves the assertion for  $\ell_0 + 1$ .

(2) We consider functions  $f \in L^1(\nu)$  such that  $\nu(f) = 0$ ,  $f \neq 0$ . Writing  $f^+ \geq 0$  for the positive part of  $f$  and  $f^- \geq 0$  for the negative part, we have  $f = f^+ - f^-$  and  $\alpha := \nu(f^+) = \nu(f^-) > 0$ . For  $N$  arbitrarily large but fixed, step (1) applied to functions  $h_N := [f^+ - (1 - \frac{1}{N})f^-]$  and  $g_N := [(1 - \frac{1}{N})f^+ - f^-]$  yields almost sure convergence

$$\begin{aligned} \lim_{r \rightarrow \infty} \frac{\ell}{r^\ell} \int_0^r s^{\ell-1} h_N(X_s) ds &= \int h_N d\nu = \frac{\alpha}{N} \\ \lim_{r \rightarrow \infty} \frac{\ell}{r^\ell} \int_0^r s^{\ell-1} g_N(X_s) ds &= \int g_N d\nu = -\frac{\alpha}{N} \end{aligned}$$

as  $r \rightarrow \infty$ . Thus, letting  $N$  tend to  $\infty$ , comparison of trajectories

$$\int_0^r s^{\ell-1} g_N(X_s) ds \leq \int_0^r s^{\ell-1} f(X_s) ds \leq \int_0^r s^{\ell-1} h_N(X_s) ds$$

establishes the desired result for functions  $f \in L^1(\nu)$  with  $\nu(f) = 0$ .  $\square$

**Lemma 2** For  $X$  as above we have for every  $\ell \in \mathbb{N}$

$$\frac{1}{r^\ell} \int_0^r s^\ell dX_s = X_r + \rho_\ell(r), \quad \lim_{r \rightarrow \infty} \rho_\ell(r) = 0 \quad \text{almost surely}.$$

**Proof** This is integration by parts

$$\int_0^r s^\ell dX_s = r^\ell X_r - \int_0^r X_s \ell s^{\ell-1} ds, \quad r > 0 \quad (8)$$

and lemma 1 (with  $f(x) = x$  and  $\nu = \mathcal{N}(0, \frac{\sigma^2}{2\tau})$ ) applied to the right hand side.  $\square$

**Lemma 3** For  $X$  defined by (7), for every  $\ell \in \mathbb{N}_0$ , we have convergence in law

$$\frac{1}{\sqrt{n^{2\ell+1}}} \int_0^n s^\ell X_s ds$$

as  $n \rightarrow \infty$  to the limit

$$\frac{\sigma}{\tau} \int_0^1 s^\ell dB_s$$

where  $B$  is standard Brownian motion.

**Proof** Rearranging SDE (7) we write

$$\tau X_s ds = -dX_s + \sigma dW_s$$

and have for every  $\ell \in \mathbb{N}_0$

$$\tau \int_0^r s^\ell X_s ds = - \int_0^r s^\ell dX_s + \sigma \int_0^r s^\ell dW_s, \quad r \geq 0. \quad (9)$$

In case  $\ell = 0$ , the right hand side is  $-(X_r - X_0) + \sigma W_r$ , and the scaling property of Brownian motion combined with ergodicity of  $X$  yields weak convergence as asserted. In case  $\ell \geq 1$ , lemma 2 transforms the first term on the right hand side of (9), and we have

$$\tau \int_0^r s^\ell X_s ds = -r^\ell (X_r + \rho_\ell(r)) + \sigma \int_0^r s^\ell dW_s. \quad (10)$$

The martingale convergence theorem (Jacod and Shiryaev 1987, VIII.3.24) shows that

$$\left( \frac{1}{\sqrt{n^{2\ell+1}}} \int_0^{tn} s^\ell dW_s \right)_{t \geq 0} \quad (11)$$

converges weakly in the Skorohod path space  $D([0, \infty)\mathbb{R})$  to a continuous limit martingale with angle bracket  $t \rightarrow \frac{1}{2\ell+1} t^{2\ell+1}$ , i.e. to

$$\left( \int_0^t s^\ell dB_s \right)_{t \geq 0}.$$

Scaled in the same way, the first term on the right hand side of (10)

$$-\frac{1}{\sqrt{n}} (X_{tn} + \rho_\ell(tn)) t^\ell$$

is negligible in comparison to (11), uniformly on compact  $t$ -intervals, by ergodicity of  $X$ .  $\square$

**Lemma 4** (a) For every  $\ell \in \mathbb{N}$  we have an expansion

$$\frac{1}{\sqrt{n^{2\ell+1}}} \int_0^n s^\ell X_s ds = \frac{\sigma}{\tau} \frac{1}{\sqrt{n^{2\ell+1}}} \int_0^n s^\ell dW_s - \frac{1}{\tau \sqrt{n}} (X_n + \rho_\ell(n)) \quad (12)$$

where  $\lim_{n \rightarrow \infty} \rho_\ell(n) = 0$  almost surely. In case  $\ell = 0$  we have

$$\frac{1}{\sqrt{n}} \int_0^n X_s ds = \frac{\sigma}{\tau} \frac{1}{\sqrt{n}} W_n - \frac{1}{\tau \sqrt{n}} (X_n - X_0).$$

(b) For every  $\ell \in \mathbb{N}$ , we have joint weak convergence as  $n \rightarrow \infty$

$$\left( \frac{1}{\sqrt{n}} \int_0^n s^0 X_s ds, \dots, \frac{1}{\sqrt{n^{2\ell+1}}} \int_0^n s^\ell X_s ds \right) \quad (13)$$

with limit law

$$\frac{\sigma}{\tau} \left( \int_0^1 s^0 dB_s, \dots, \int_0^1 s^\ell dB_s \right).$$



**Proof** Part (a) is (10) plus scaling as in the proof of lemma 3. For different  $\ell \in \mathbb{N}_0$ , the expansions (12) hold with respect to the same driving Brownian motion  $W$  from SDE (7): this gives (b).  $\square$

### 3 The statistical model of interest

Consider now a more general problem of parameter estimation from continuous-time observation of

$$Y_t = R(t) + X_t, \quad t \geq 0 \quad (14)$$

where  $R(\cdot)$  is a sufficiently smooth deterministic function which depends on some finite-dimensional parameter  $\vartheta$ , and where the Ornstein Uhlenbeck process  $X = (X_t)_{t \geq 0}$ , unique strong solution to

$$dX_t = -\tau X_t dt + \sqrt{c} dW_t, \quad (15)$$

depends on a parameter  $\tau > 0$ . The starting point  $X_0 \equiv x_0$  is deterministic. Then  $Y$  solves the SDE

$$dY_t = (S(t) - \tau Y_t) dt + \sqrt{c} dW_t, \quad t \geq 0 \quad (16)$$

where  $S$  depending on  $\vartheta$  and  $\tau$  is given by

$$S(t) = [R' + \tau R](t) \quad \text{where} \quad R' := \frac{d}{dt} R. \quad (17)$$

Conversely, if a process  $Y$  is solution to an SDE of type (16), then solving  $R' = -\tau R + S$  with suitable initial point  $r_0$  we get a representation (14) for  $Y$  where

$$R(t) = r_0 e^{-\tau t} + \int_0^t S(s) e^{-\tau(t-s)} ds. \quad (18)$$

For examples of parametric models of this type, see e.g. Dehling et al. (2010), Franke and Kott (2013), Höpfner and Kutoyants (2009), Pchelintsev (2013), and example 2.3 in Höpfner and Kutoyants (2011). The constant  $c > 0$  in (15) is fixed and known: observing the trajectory of  $Y$  continuously in time,  $\langle Y \rangle_t = c t$  equals the limit in probability  $\lim_{n \rightarrow \infty} \sum_{j=1}^{2^n} (Y_{tj/2^n} - Y_{t(j-1)/2^n})^2$  of squared increments, thus  $c$  is known almost surely and cannot be considered as a parameter.

We wish to estimate the unknown parameter  $\theta := (\vartheta, \tau)$  based on time-continuous observation of  $Y$  in (14) over a long time interval, in the model

$$R_\vartheta(t) = \sum_{j=0}^p \vartheta_j t^j \quad (19)$$

with real coefficients where we assume a leading coefficient  $\vartheta_p > 0$  such that trajectories of  $Y$  tend to  $\infty$  almost surely as  $t \rightarrow \infty$ . Thus our parametrization is

$$\theta := (\vartheta_0, \vartheta_1, \dots, \vartheta_{p-1}, \vartheta_p, \tau)^\top \in \Theta := \mathbb{R}^p \times (0, \infty) \times (0, \infty) \quad (20)$$

and in SDE (16) which governs the observation  $Y$ ,  $S$  depending on  $\theta = (\vartheta, \tau)$  has the form

$$S_\theta(t) = [R'_\vartheta + \tau R_\vartheta](t) = \tau \vartheta_0 + \sum_{j=1}^p \vartheta_j t^{j-1} (j + \tau t). \quad (21)$$

### 3.1 Local asymptotic normality for the model (14)+(19)

Let  $C := C([0, \infty), \mathbb{R})$  denote the canonical path space for continuous processes; with  $\pi = (\pi_t)_{t \geq 0}$  the canonical process (i.e.  $\pi_t(f) = f(t)$  for  $f \in C$ ,  $t \geq 0$ ) and  $\mathcal{C} = \sigma(\pi_t : t \geq 0)$ ,

$$\mathbb{G} = (\mathcal{G}_t)_{t \geq 0}, \quad \mathcal{G}_t := \bigcap_{r > t} \sigma(\pi_s : s \leq r), \quad t \geq 0$$

is the canonical filtration. Let  $Q_\theta$  denote the law on  $(C, \mathcal{C}, \mathbb{G})$  of the process  $Y$  in (14) under  $\theta \in \Theta$ , cf. (20). By (14), (15), (16), (19) and (21), the canonical process  $\pi = (\pi_t)_{t \geq 0}$  on  $(C, \mathcal{C})$  under  $Q_\theta$  solves

$$d\pi_t = \left( \left[ \tau \vartheta_0 + \sum_{j=1}^p \vartheta_j t^{j-1} (j + \tau t) \right] - \tau \pi_t \right) dt + \sqrt{c} dW_t. \quad (22)$$

For pairs  $\theta' \neq \theta$  in  $\Theta$ , probability measures  $Q_{\theta'}$ ,  $Q_\theta$  are locally equivalent relative to  $\mathbb{G}$ , and we write

$$\begin{aligned} \gamma'(s, y) &= \frac{S_{\theta'}(s) - \tau'y}{c} = \frac{[R'_{\theta'} + \tau'R_{\theta'}](s) - \tau'y}{c} \\ \gamma(s, y) &= \frac{S_\theta(s) - \tau y}{c} = \frac{[R'_\theta + \tau R_\theta](s) - \tau y}{c}. \end{aligned}$$

With  $m^{\pi, \theta} = \sqrt{c} dW_s$  the martingale part of  $\pi$  under  $\theta$ , the likelihood ratio process of  $Q_{\theta'}$  w.r.t.  $Q_\theta$  relative to  $\mathbb{G}$  (Lipster and Shiryaev 2001; Ibragimov and Khasminskii 1981; Jacod and Shiryaev 1987; Kutoyants 2004; Höpfner 2014 p. 162) is

$$L_t^{\theta'/\theta} = \exp \left( \int_0^t (\gamma' - \gamma)(s, \pi_s) \sqrt{c} dW_s - \frac{1}{2} \int_0^t (\gamma' - \gamma)^2(s, \pi_s) c ds \right). \quad (23)$$

In the integrand,

$$\begin{aligned} c(\gamma' - \gamma)(s, \pi_s) &= (R'_{\theta'} - R'_\theta)(s) - (\tau' - \tau)\pi_s + \tau'R_{\theta'}(s) - \tau R_\theta(s) \\ &= (R'_{\theta'} - R'_\theta)(s) - (\tau' - \tau)(\pi_s - R_\theta(s)) + \tau(R_{\theta'} - R_\theta)(s) + (\tau' - \tau)(R_{\theta'} - R_\theta)(s), \end{aligned}$$

so we exploit (14) to write for short

$$c(\gamma' - \gamma)(s, \pi_s) = [(R'_{\theta'} - R'_\theta) + \tau(R_{\theta'} - R_\theta)](s) - (\tau' - \tau)X_s + (\tau' - \tau)(R_{\theta'} - R_\theta)(s) \quad (24)$$

where  $X$  under  $\theta = (\vartheta, \tau)$  is the Ornstein Uhlenbeck process (15), and where

$$[(R'_{\theta'} - R'_\theta) + \tau(R_{\theta'} - R_\theta)](s) = \tau(\vartheta'_0 - \vartheta_0) + \sum_{j=1}^p (\vartheta'_j - \vartheta_j) s^{j-1} [j + \tau s].$$

Localization at  $\theta \in \Theta$  will be as follows: with notation

$$\begin{aligned} \vartheta'_0(n, h) &:= \vartheta_0 + \frac{1}{\sqrt{n}} h_0, \quad \tau'(n, h) := \tau + \frac{1}{\sqrt{n}} h_{p+1}, \\ \vartheta'_j(n, h) &:= \vartheta_j + \frac{1}{\sqrt{n^{2j+1}}} h_j, \quad 1 \leq j \leq p \end{aligned}$$

we insert

$$\theta'(n, h) = \begin{pmatrix} \vartheta'_0(n, h) \\ \vartheta'_1(n, h) \\ \dots \\ \vartheta'_p(n, h) \\ \tau'(n, h) \end{pmatrix} \quad \text{where} \quad h = \begin{pmatrix} h_0 \\ h_1 \\ \dots \\ h_p \\ h_{p+1} \end{pmatrix} \quad \text{is such that} \quad \theta'(n, h) \in \Theta$$

in place of  $\theta'$  into (23); finally we rescale time. Define

$$\psi_n := \begin{pmatrix} \frac{1}{\sqrt{n}} & 0 & \dots & 0 \\ 0 & \frac{1}{\sqrt{n^3}} & \dots & 0 \\ 0 & \ddots & \ddots & 0 \\ 0 & \dots & \dots & \frac{1}{\sqrt{n^{2p+1}}} \\ 0 & \dots & 0 & \frac{1}{\sqrt{n}} \end{pmatrix}. \quad (25)$$

With local parameter  $h$  and local scale (25) at  $\theta$ , we obtain from (23)+(24)

$$L_{tn}^{(\theta + \psi_n h) / \theta} = \exp \left( h^\top S_{n, \theta}(t) - \frac{1}{2} h^\top J_{n, \theta}(t) h + \rho_{n, \theta, h}(t) \right) \quad (26)$$

where  $\rho_{n, \theta, h}$  is some process of remainder terms,  $S_{n, \theta}$  a martingale with respect to  $\mathcal{Q}_\theta$  and  $(\mathcal{G}_{tn})_{t \geq 0}$

$$S_{n, \theta}(t) := \frac{1}{\sqrt{c}} \begin{pmatrix} \frac{1}{\sqrt{n}} \int_0^{tn} \tau \, dW_s \\ \frac{1}{\sqrt{n^3}} \int_0^{tn} (1 + \tau s) \, dW_s \\ \vdots \\ \frac{1}{\sqrt{n^{2j+1}}} \int_0^{tn} s^{j-1} (j + \tau s) \, dW_s \\ \vdots \\ \frac{1}{\sqrt{n^{2p+1}}} \int_0^{tn} s^{p-1} (p + \tau s) \, dW_s \\ -\frac{1}{\sqrt{n}} \int_0^{tn} X_s \, dW_s \end{pmatrix} \quad (27)$$

(again by (14),  $X_s$  stands for  $\pi_s - R_\theta(s)$  under  $\theta$ ), and  $J_{n, \vartheta}$  the angle bracket of  $S_{n, \theta}$  under  $\theta$ .

**Proposition 1** (a) For fixed  $0 < t < \infty$ , components of  $J_{n, \theta}(t)$  converge  $\mathcal{Q}_\theta$ -almost surely as  $n \rightarrow \infty$  to those of the deterministic process

$$J(t) = \frac{1}{c} \begin{pmatrix} \frac{\tau^2}{1} t & \frac{\tau^2}{2} t^2 & \dots & \frac{\tau^2}{p+1} t^{p+1} & 0 \\ \frac{\tau^2}{2} t^2 & \frac{\tau^2}{3} t^3 & \dots & \frac{\tau^2}{p+2} t^{p+2} & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \frac{\tau^2}{p+1} t^{p+1} & \frac{\tau^2}{p+2} t^{p+2} & \dots & \frac{\tau^2}{2p+1} t^{2p+1} & 0 \\ 0 & \dots & \dots & 0 & \frac{c}{2\tau} t \end{pmatrix}, \quad t \geq 0. \quad (28)$$

For every  $0 < t < \infty$ , the matrix  $J(t)$  is invertible.

(b) Let  $\tilde{W}$  denote a two-dimensional standard Brownian motion with components  $\tilde{W}^{(1)}$  and  $\tilde{W}^{(2)}$ . In the cadlag path space  $D = D([0, \infty), \mathbb{R}^{p+2})$  (Jacod and Shiryaev 1987, chapters VI and VIII), martingales  $S_{n,\theta}$  under  $\mathcal{Q}_\theta$  converge weakly as  $n \rightarrow \infty$  to the limit martingale

$$S(t) = \begin{pmatrix} \frac{\tau}{\sqrt{c}} \int_0^t s^0 d\tilde{W}_s^{(1)} \\ \frac{\tau}{\sqrt{c}} \int_0^t s^1 d\tilde{W}_s^{(1)} \\ \vdots \\ \frac{\tau}{\sqrt{c}} \int_0^t s^p d\tilde{W}_s^{(1)} \\ \frac{1}{\sqrt{2\tau}} \tilde{W}_t^{(2)} \end{pmatrix}, \quad t \geq 0. \quad (29)$$

**Proof** The proof is in several steps. (1) We specify the angle bracket (or predictable quadratic covariation) process  $J_{n,\theta}$  of  $S_{n,\theta}$  under  $\mathcal{Q}_\theta$ . Its state at time  $t$

$$J_{n,\theta}(t) := \left( J_{n,\theta}^{(i,j)}(t) \right)_{i,j=0,1,\dots,p+1}$$

is a symmetric matrix of size  $(p+2) \times (p+2)$ . Taking into account the norming factor in front of  $S_{n,\theta}$  in (27) we consider throughout  $cJ_{n,\theta}$ . The entries are given as follows. We have

$$cJ_{n,\theta}^{(i,j)}(t) = \frac{1}{n^{i+j+1}} \int_0^{tn} s^{i+j-2} (i+\tau s)(j+\tau s) ds \xrightarrow{n \rightarrow \infty} \tau^2 \frac{1}{i+j+1} t^{i+j+1}$$

for all  $1 \leq i, j \leq p$ . In the first line of  $cJ_{n,\theta}(t)$  we have

$$cJ_{n,\theta}^{(0,0)}(t) = \tau^2 t, \quad cJ_{n,\theta}^{(0,p+1)}(t) = -\frac{1}{n} \int_0^{tn} \tau X_s ds$$

in first and last position, and in-between for  $1 \leq j \leq p$

$$cJ_{n,\theta}^{(0,j)}(t) = \frac{1}{n^{j+1}} \int_0^{tn} \tau s^{j-1} (j+\tau s) ds \xrightarrow{n \rightarrow \infty} \tau^2 \frac{1}{j+1} t^{j+1}.$$

For the last column of  $cJ_{n,\theta}(t)$ , the first entry  $cJ_{n,\theta}^{(0,p+1)}(t)$  has been given above, the last entry is

$$cJ_{n,\theta}^{(p+1,p+1)}(t) = \frac{1}{n} \int_0^{tn} X_s^2 ds,$$

in-between we have for  $1 \leq j \leq p$

$$cJ_{n,\theta}^{(j,p+1)}(t) = -\frac{1}{n^{j+1}} \int_0^{tn} s^{j-1} (j+\tau s) X_s ds.$$

It remains to consider the three integrals which are not deterministic: here lemma 1 establishes almost sure convergence

$$\frac{1}{n} \int_0^{tn} X_s^2 ds \rightarrow \frac{c}{2\tau} t, \quad \frac{1}{n} \int_0^{tn} \tau X_s ds \rightarrow 0, \quad \frac{1}{n^{j+1}} \int_0^{tn} s^{j-1} (j+\tau s) X_s ds \rightarrow 0$$

as  $n \rightarrow \infty$  under  $\mathcal{Q}_\theta$ . This proves almost sure convergence of the components of  $J_{n,\theta}(t)$  to the corresponding components of  $J(t)$  defined in (28).

(2) We prove that for every  $0 < t < \infty$ , the matrix  $J(t)$  defined in (28) is invertible. For this it is sufficient to check invertibility of  $(p+1) \times (p+1)$  matrices

$$\tilde{J}(t) := \left( \frac{1}{i+j+1} t^{i+j+1} \right)_{i,j=0,\dots,p}. \quad (30)$$

which up to the factor  $\frac{\tau}{c}$  represent the upper left block in  $J(t)$ . We have to show that  $\min_{|u|=1} u^\top \tilde{J}(t) u$  is strictly positive. In case  $t = 1$ , the rows of  $\tilde{J}(1)$  are linearly independent vectors in  $\mathbb{R}^{p+1}$ , thus the assertion holds. For  $t \neq 1$ , associate  $v(u) = \sqrt{t} (u_i t^i)_{0 \leq i \leq p}$  to  $u \in \mathbb{R}^{p+1}$ . Then we have  $u^\top \tilde{J}(t) u = v(u)^\top \tilde{J}(1) v(u)$ , from which we deduce the assertion. Part a) of the proposition is proved.

(3) As an auxiliary step, we determine a martingale  $\tilde{S}$  which admits  $\tilde{J}$  defined in (30)

$$\tilde{J} = (\tilde{J}(t))_{t \geq 0}, \quad \tilde{J}(t) = \left( \tilde{J}^{(i,j)}(t) \right)_{i,j=0,1,\dots,p}$$

as its angle bracket. In integral representation we have

$$\tilde{J}^{(i,j)}(t) = \int_0^t s^{i+j} ds = \int_0^t (\Psi_s \Psi_s)^{(i,j)} ds$$

where  $\Psi_s$  is a square root

$$\Psi_s := \frac{1}{\sqrt{q(s)}} (s^{i+\ell})_{i,\ell=0,\dots,p} \quad \text{with} \quad q(s) := \sum_{\ell=0}^p s^{2\ell}$$

for the matrix  $(s^{i+j})_{i,j=0,\dots,p}$  (note that for fixed  $s$ , this matrix is not invertible). From this representation for the angle brackets  $\tilde{J}$  we obtain a representation of the martingale  $\tilde{S}$  (Ikeda and Watanabe 1989, theorem 7.1' on p. 90):

$$\tilde{S} = (\tilde{S}^0, \dots, \tilde{S}^p), \quad \tilde{S}^i(t) = \sum_{j=0}^p \int_0^t \Psi_s^{(i,j)} d\tilde{B}_s^j, \quad 0 \leq i \leq p$$

with some  $(p+1)$ -dimensional standard Brownian motion

$$\tilde{B} = (\tilde{B}^0, \dots, \tilde{B}^p).$$

Given the simple structure of the  $\Psi_s$ , we can define a new one-dimensional Brownian motion  $\tilde{W}^{(1)}$  by

$$\tilde{W}_s^{(1)} := \sum_{\ell=0}^p \int_0^t \frac{1}{\sqrt{q(s)}} s^\ell d\tilde{B}_s^\ell$$

and end up with

$$\tilde{S} = (\tilde{S}^0, \dots, \tilde{S}^p), \quad \tilde{S}^i(t) = \int_0^t s^i d\tilde{W}_s^{(1)}, \quad 0 \leq i \leq p. \quad (31)$$

(4) Now we can determine the martingale  $S$  which admits  $J$  defined in (28) as angle bracket. Since  $J(t)$  has a diagonal block structure where (up to multiplication with a constant in every block) the upper left block has been considered in step 3 whereas we have for the lower right block

$$\lim_{n \rightarrow \infty} \frac{1}{n} \int_0^{tn} X_s^2 ds = \frac{c}{2\tau} t$$

$Q_\theta$ -almost surely by lemma 1, the desired representation is

$$S(t) = \begin{pmatrix} \frac{\tau}{\sqrt{c}} \int_0^t s^0 d\tilde{W}_s^{(1)} \\ \frac{\tau}{\sqrt{c}} \int_0^t s^1 d\tilde{W}_s^{(1)} \\ \vdots \\ \frac{\tau}{\sqrt{c}} \int_0^t s^p d\tilde{W}_s^{(1)} \\ \frac{1}{\sqrt{2\tau}} \tilde{W}_t^{(2)} \end{pmatrix}, \quad t \geq 0$$

with  $\tilde{W}^{(1)}$  from (31), and with another one-dimensional Brownian motion  $\tilde{W}^{(2)}$  which is independent from  $\tilde{W}^{(1)}$ . This is the form appearing in (29) of the proposition.

(5) On the basis of part (a) of the proposition, the martingale convergence theorem (Jacod and Shiryaev 1987, VIII.3.24) establishes weak convergence (in the path space  $D([0, \infty), \mathbb{R}^{p+2})$ , under  $Q_\theta$ , as  $n \rightarrow \infty$ ) of the martingales  $S_{n,\theta}$  under  $Q_\theta$  to the limit martingale  $S$  which has been determined in step (4). This finishes the proof of proposition 1.  $\square$

As a consequence of proposition 1, we obtain local asymptotic normality (Le Cam 1969; Hajek 1970; Ibragimov and Khasminskii 1981; Davies 1985; Le Cam and Yang 2002; Pfanzagl 1994; Kutoyants 2004; Höpfner 2014, section 7.1).

**Theorem 1** (a) At  $\theta \in \Theta$ , with local scale at  $\theta$  given by  $(\psi_n)_n$  from (25), quadratic expansions

$$\log L_n^{(\theta + \psi_n h_n) / \theta} = h_n^\top S_{n,\theta}(1) - \frac{1}{2} h_n^\top J_{n,\theta}(1) h_n + o_{(Q_\theta)}(1), \quad n \rightarrow \infty$$

hold for arbitrary bounded sequences  $(h_n)_n$  in  $\mathbb{R}^{p+2}$ ; since  $\Theta$  is open,  $\theta + \psi_n h_n$  belongs to  $\Theta$  for  $n$  large enough. Eventually as  $n \rightarrow \infty$ ,  $J_{n,\theta}(1)$  takes its values in the set of invertible  $(p+2) \times (p+2)$ -matrices,  $Q_\theta$ -almost surely.

(b) For every  $\theta \in \Theta$ , we have weak convergence in  $D([0, \infty), \mathbb{R}^{(p+2)} \times \mathbb{R}^{(p+2) \times (p+2)})$  as  $n \rightarrow \infty$

$$\mathcal{L}((S_{n,\theta}, J_{n,\theta}) | Q_\theta) \longrightarrow \mathcal{L}(S, J)$$

with  $S$  the martingale in (29) and  $J$  its angle bracket in (28).

(c) There is a Gaussian shift limit experiment  $\mathcal{E}(S, J)$  with likelihood ratios

$$\exp\left(h^\top S(1) - \frac{1}{2} h^\top J(1) h\right), \quad h \in \mathbb{R}^{p+2}.$$

**Proof** (1) As a first step, weak convergence of  $S_{n,\theta}$  to  $S$  under  $Q_\theta$  in proposition 1 implies (Jacod and Shiryaev 1987, theorem VI.6.1) joint weak convergence of the martingale together with its angle bracket. This is part (b) of the theorem. For  $0 < t < \infty$  fixed, invertibility of  $J_{n,\theta}(t)$ ,  $Q_\theta$ -almost surely for sufficiently large  $n$ , follows from invertibility of  $J(t)$  and componentwise almost surely convergence  $J_{n,\theta}(t) \rightarrow J(t)$  by proposition 1.

(2) We can represent the limit experiment  $\mathcal{E}(S, J)$  in (c) as  $\{\mathcal{N}(J(1)h, J(1)) : h \in \mathbb{R}^{p+2}\}$ .

(3) Fix a bounded sequence  $(h_n)_n$  in  $\mathbb{R}^{p+2}$ , take  $n$  large enough so that  $\theta + \psi_n h_n$  is in  $\Theta$ , and define

$$\rho_{n,\theta,h_n}(t) := \log L_{tn}^{(\theta + \psi_n h_n) / \theta} - \left\{ h_n^\top S_{n,\theta}(t) - \frac{1}{2} h_n^\top J_{n,\theta}(t) h_n \right\}, \quad t \geq 0. \quad (32)$$

Using notation  $\theta'(n, h) = \theta + \psi_n h$  as in (23)–(26), we split  $\theta'(n, h_n) = (\vartheta'(n, h_n), \tau'(n, h_n))$  into a bloc  $\vartheta'(n, h_n) = (\vartheta'_0(n, h_n), \dots, \vartheta'_p(n, h_n))$  and the last component  $\tau'(n, h_n)$ . We

write  $h_{n,0}, h_{n,1}, \dots, h_{n,p+1}$  for the components of the local parameter  $h_n$ . Comparing (26) to (23), we see that out of

$$c(\gamma' - \gamma)(s, \pi_s) = G(n, h_n)(s) + H(n, h_n)(s)$$

to be considered in (23)+(24) we did consider

$$\begin{aligned} H(n, h_n)(s) &:= \left[ (R'_{\vartheta'(n, h_n)} - R'_{\vartheta}) + \tau(R_{\vartheta'(n, h_n)} - R_{\vartheta}) \right](s) - (\tau'(n, h_n) - \tau)X_s \\ &= h_{n,0} \frac{1}{\sqrt{n}} \tau + \sum_{i=1}^p h_{n,i} \frac{1}{\sqrt{n^{2i+1}}} s^{i-1}(i + \tau s) - h_{n,p+1} \frac{1}{\sqrt{n}} X_s \end{aligned}$$

under the integral signs, whereas we did neglect contributions

$$G(n, h_n)(s) := (\tau'(n, h_n) - \tau)(R_{\vartheta'(n, h_n)} - R_{\vartheta})(s)$$

under the integral signs, both in the martingales and in the quadratic variations. With these notations, the remainder terms (32) have the form

$$\rho_{n,\theta,h_n}(t) = \frac{1}{\sqrt{c}} \int_0^{tn} G(n, h_n)(s) dW_s - \frac{1}{2c} \int_0^{tn} [2G(n, h_n)H(n, h_n) + G^2(n, h_n)](s) ds.$$

Recall that  $(h_n)_n$  is a bounded sequence. By choice of the localization and by (19) we have

$$G(n, h_n)(s) = O\left(\frac{1}{\sqrt{n}}\right) \cdot \sum_{i=0}^p h_{n,i} \frac{1}{\sqrt{n^{2i+1}}} s^i.$$

Transforming the convergence arguments in the proof of proposition 1 into tightness arguments, the random objects

$$\int_0^{tn} H^2(n, h_n)(s) ds = O_{Q_\theta}(1)$$

remain tight under  $Q_\theta$  as  $n \rightarrow \infty$ , for every  $t$  fixed. The deterministic sequence

$$\int_0^{tn} G^2(n, h_n)(s) ds = O\left(\frac{1}{n}\right)$$

vanishing as  $n \rightarrow \infty$ ,

$$\int_0^{tn} [G(n, h_n)H(n, h_n)](s) ds = O_{Q_\theta}\left(\frac{1}{\sqrt{n}}\right)$$

vanishes under  $Q_\theta$  as  $n \rightarrow \infty$  by Cauchy-Schwarz. The sequence of martingales

$$\left( \int_0^{tn} G(n, h_n)(s) dW_s \right)_{t \geq 0}$$

has angle brackets which vanish as  $n \rightarrow \infty$  for every  $t$  fixed, so the martingales itself vanish in  $Q_\theta$ -probability, uniformly over compact  $t$ -intervals as  $n \rightarrow \infty$ . With  $0 < t_0 < \infty$  arbitrary, this proves

$$\sup_{0 \leq t \leq t_0} |\rho_{n,\theta,h_n}(t)| \quad \text{vanishes in } Q_\theta\text{-probability as } n \rightarrow \infty$$

for the remainder terms (32). Since we did consider arbitrary bounded sequences  $(h_n)_n$ , we can reformulate the last assertion in the form

$$\sup_{|h| \leq C} \sup_{0 \leq t \leq t_0} |\rho_{n,\theta,h}(t)| \quad \text{vanishes in } Q_\theta\text{-probability as } n \rightarrow \infty$$

for arbitrary  $0 < C < \infty$ . We thus have proved part (a) of the theorem. The proof is finished.  $\square$

The local asymptotic minimax theorem arises as a consequence of theorem 1, see Ibragimov and Khasminskii (1981), Davies (1985), Le Cam and Yang (2002), Kutoyants (2004), or Höpfner (2014) theorem 7.12. Note that it is interesting to consider quite arbitrary  $\mathcal{G}_n$ -measurable random variables  $T_n$  taking values in  $\mathbb{R}^{(p+2)}$  as possibly useful estimators for the unknown parameter  $\theta \in \Theta$ .

**Corollary 1** For  $\theta \in \Theta$ , for arbitrary estimator sequences  $(T_n)_n$  whose rescaled estimation errors

$$\mathcal{L}(\psi_n^{-1}(T_n - \theta) \mid Q_\theta)$$

at  $\theta$  are tight as  $n \rightarrow \infty$ , for arbitrary loss functions  $L : \mathbb{R}^{(p+2)} \rightarrow [0, \infty)$  which are continuous, bounded and subconvex, the following local asymptotic minimax bound holds:

$$\sup_{C \uparrow \infty} \liminf_{n \rightarrow \infty} \sup_{|h| \leq C} E_{\theta + \psi_n h} (L(\psi_n^{-1}(T_n - (\theta + \psi_n h)))) \geq E(L(J^{-1}(1)S(1))) .$$

Estimator sequences whose rescaled estimation errors at  $\theta$  admit as  $n \rightarrow \infty$  a representation

$$\psi_n^{-1}(T_n - \theta) = [J_{n,\theta}(1)]^{-1} S_{n,\theta}(1) + o_{Q_\theta}(1) = J^{-1}(1) S_{n,\theta}(1) + o_{Q_\theta}(1) \quad (33)$$

have the property

$$\lim_{n \rightarrow \infty} \sup_{|h| \leq C} E_{\theta + \psi_n h} (L(\psi_n^{-1}(T_n - (\theta + \psi_n h)))) = E(L(J^{-1}(1)S(1)))$$

for every  $0 < C < \infty$  fixed, and thus attain the local asymptotic minimax bound at  $\theta$ .

**Remark 1** In theorem 1, the limit experiment  $\mathcal{E}(S, J)$  at  $\theta = (\vartheta, \tau) \in \Theta$  depends on the component  $\tau$  (the constant  $c$  is not a parameter), by (28), but not on  $\vartheta = (\vartheta_0, \vartheta_1, \dots, \vartheta_p)$ . The  $\tau$ -component of  $\mathcal{E}(S, J)$  is the well-known limit experiment when an ergodic Ornstein Uhlenbeck process (15) with backdriving force  $\tau$  is observed over a long time interval (Kutoyants 2004; Höpfner 2014 section 8.1).

### 3.2 Estimating $(\vartheta_0, \dots, \vartheta_p)$ in the model (14)+(19)

By abuse of language, we write in this subsection  $Y$  for  $\pi$  on  $(C, \mathcal{C})$  under  $Q_\theta$ ,  $X$  for  $\pi - R_\theta$  under  $Q_\theta$ ; as before  $\sqrt{c}W$  denotes the martingale part of  $Y$  or  $X$  under  $Q_\theta$  relative to  $\mathbb{G}$ . To estimate  $\vartheta = (\vartheta_0, \dots, \vartheta_p) \in \mathbb{R}^p \times (0, \infty)$  in the model (14)+(19), consider

$$\tilde{\vartheta}(t) := \operatorname{arginf}_{\vartheta' = (\vartheta'_0, \dots, \vartheta'_p)} \int_0^t (Y(s) - R_{\vartheta'}(s))^2 ds . \quad (34)$$



Least squares estimators (34) are uniquely determined—see (37) below—and have an explicit and easy-to-calculate form; we discuss their asymptotics under  $\theta = (\vartheta, \tau)$ . Define martingales  $\tilde{S}_{n,\theta}$  with respect to  $\mathcal{Q}_\theta$  and  $(\mathcal{G}_{tn})_{t \geq 0}$

$$\tilde{S}_{n,\theta}(t) := \frac{1}{\sqrt{c}} \begin{pmatrix} \frac{1}{\sqrt{n}} \int_0^{tn} \tau dW_s \\ \frac{1}{\sqrt{n^3}} \int_0^{tn} (1 + \tau s) dW_s \\ \vdots \\ \frac{1}{\sqrt{n^{2p+1}}} \int_0^{tn} s^{p-1} (p + \tau s) dW_s \end{pmatrix}, \quad t \geq 0 \quad (35)$$

which coincide with  $S_{n,\theta}$  of (27) whose last component has been suppressed. Let  $\tilde{J}_{n,\theta}$  denote the angle bracket of  $\tilde{S}_{n,\theta}$  under  $\mathcal{Q}_\theta$ . We consider also

$$\tilde{\psi}_n := \begin{pmatrix} \frac{1}{\sqrt{n}} & 0 & \dots & 0 \\ 0 & \frac{1}{\sqrt{n^3}} & \dots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \dots & \frac{1}{\sqrt{n^{2p+1}}} \end{pmatrix} \quad (36)$$

which coincides with local scale  $\psi_n$  of (25) whose last row and last column have been suppressed, and invertible deterministic  $(p+1) \times (p+1)$  matrices as defined in (30) in the proof of proposition 1:

$$\tilde{J}(t) := \left( \frac{1}{i+j+1} t^{i+j+1} \right)_{i,j=0,\dots,p}.$$

**Proposition 2** For every  $\theta = (\vartheta, \tau) \in \Theta$ , rescaled estimation errors of the least squares estimator (34) admit a representation

$$\tilde{\psi}_n^{-1} (\tilde{\vartheta}(n) - \vartheta) = [\tilde{J}_{n,\theta}(1)]^{-1} \tilde{S}_{n,\theta}(1) + o_{\mathcal{Q}_\theta}(1) = \left[ \frac{\tau^2}{c} \tilde{J}(1) \right]^{-1} \tilde{S}_{n,\theta}(1) + o_{\mathcal{Q}_\theta}(1)$$

as  $n \rightarrow \infty$ .  $\square$

**Proof** (1) Almost surely as  $n \rightarrow \infty$ , angle brackets  $\tilde{J}_{n,\theta}$  of  $\tilde{S}_{n,\theta}$  under  $\mathcal{Q}_\theta$  converge to

$$\frac{\tau^2}{c} \left( \int_0^t s^{i+j} ds \right)_{i,j=0,\dots,p} = \frac{\tau^2}{c} \tilde{J}(t)$$

for fixed  $0 < t < \infty$ . This has been proved in proposition 1.

(2) Least squares estimators  $\tilde{\vartheta}(t)$  in (34) are uniquely defined and have the explicit form

$$\begin{pmatrix} \int_0^t Y_s ds \\ \int_0^t s Y_s ds \\ \vdots \\ \int_0^t s^p Y_s ds \end{pmatrix} = \tilde{J}(t) \begin{pmatrix} \tilde{\vartheta}_0(t) \\ \tilde{\vartheta}_1(t) \\ \vdots \\ \tilde{\vartheta}_p(t) \end{pmatrix} = \tilde{J}(t) \tilde{\vartheta}(t) : \quad (37)$$

to check this, take derivatives under the integral sign in (34), use (19) for  $i = 0, 1, \dots, p$

$$\frac{d}{d\vartheta'_i} (Y(s) - R_{\vartheta'}(s))^2 = -2 (Y(s) - R_{\vartheta'}(s)) \frac{d}{d\vartheta'_i} R_{\vartheta'}(s) = -2 (Y(s) - R_{\vartheta'}(s)) s^i,$$

put integrals equal to zero and use the definition (30) of  $\tilde{J}(t)$ . On the other hand, (19) shows

$$\begin{pmatrix} \int_0^t R_{\vartheta}(s) ds \\ \int_0^t s R_{\vartheta}(s) ds \\ \dots \\ \int_0^t s^p R_{\vartheta}(s) ds \end{pmatrix} = \tilde{J}(t) \begin{pmatrix} \vartheta_0 \\ \vartheta_1 \\ \dots \\ \vartheta_p \end{pmatrix}. \quad (38)$$

Thus (14) allows to write

$$\tilde{J}(t) (\tilde{\vartheta}(t) - \vartheta) = \begin{pmatrix} \int_0^t X_s ds \\ \int_0^t s X_s ds \\ \dots \\ \int_0^t s^p X_s ds \end{pmatrix}. \quad (39)$$

The scaling property

$$\tilde{\psi}_n \tilde{J}(n) \tilde{\psi}_n = \tilde{J}(1) \quad (40)$$

applied to (39) then yields the representation

$$\tilde{\psi}_n^{-1} (\tilde{\vartheta}(n) - \vartheta) = [\tilde{J}(1)]^{-1} \begin{pmatrix} \frac{1}{\sqrt{n}} \int_0^n X_s ds \\ \frac{1}{\sqrt{n^3}} \int_0^n s X_s ds \\ \dots \\ \frac{1}{\sqrt{n^{2p+1}}} \int_0^n s^p X_s ds \end{pmatrix}. \quad (41)$$

(3) Representations (12) in lemma 4 combined with the definition of  $\tilde{S}_{n,\theta}$  in (35) show that under  $Q_\theta$  as  $n \rightarrow \infty$ , the vector on the right hand side of (41) can be written as

$$\frac{\sqrt{c}}{\tau} \begin{pmatrix} \frac{1}{\sqrt{n}} \int_0^n dW_s \\ \frac{1}{\sqrt{n^3}} \int_0^n s dW_s \\ \dots \\ \frac{1}{\sqrt{n^{2p+1}}} \int_0^n s^p dW_s \end{pmatrix} + o_{Q_\theta}(1) = \frac{\sqrt{c}}{\tau} \left[ \frac{\sqrt{c}}{\tau} \tilde{S}_{n,\theta}(1) + o_{Q_\theta}(1) \right] + o_{Q_\theta}(1).$$

Taking into account step (1) this allows to write representation (41) of rescaled estimation errors as

$$\tilde{\psi}_n^{-1} (\tilde{\vartheta}(n) - \vartheta) = \left[ \frac{\tau^2}{c} \tilde{J}(1) \right]^{-1} \tilde{S}_{n,\theta}(1) + o_{Q_\theta}(1) = [\tilde{J}_{n,\theta}(1)]^{-1} \tilde{S}_{n,\theta}(1) + o_{Q_\theta}(1)$$

which concludes the proof.  $\square$

### 3.3 Estimating $\tau$ in the model (14)+(19)

Also in this subsection,  $Y$  stands for  $\pi$  on  $(C, \mathcal{C})$  under  $Q_\theta$ ,  $X$  for  $\pi - R_\theta$  under  $Q_\theta$ , and  $\sqrt{c}W$  for the martingale part of  $Y$  or  $X$  under  $Q_\theta$  relative to  $\mathbb{G}$ . To estimate  $\tau > 0$  in the model (14)+(19) based on observation of  $Y$  up to time  $n$ , define

$$\tilde{\tau}(n) = \frac{\sum_{i=0}^p \tilde{\vartheta}_i(n) \int_0^n s^i dY_s - \int_0^n Y_s dY_s}{\int_0^n Y_s^2 ds - \tilde{\vartheta}(n)^\top \tilde{J}(n) \tilde{\vartheta}(n)} \quad (42)$$

where  $\tilde{\vartheta}(n)$  is the least squares estimator (34), and  $\tilde{J}(n)$  is given by (30).

A motivation is as follows. With notations of Sect. 3.1 write the log-likelihood surface  $\theta' \rightarrow \log L_n^{\theta'/\theta}$  under  $Q_\theta$  in the form

$$\log L_t^{(\vartheta', \tau')/(\vartheta, \tau)} = \int_0^t (\gamma' - \gamma)(s, Y_s) dY_s - \frac{1}{2} \int_0^t ([\gamma']^2 - \gamma^2)(s, Y_s) c ds ;$$

neglect contributions which do not depend on  $\theta' = (\vartheta', \tau')$ ; maximize in  $\tau' > 0$  on  $\vartheta'$ -sections  $\Theta_{\vartheta'} := \{(\vartheta', \tau') : \tau' > 0\} \subset \Theta$  on which  $\vartheta'$  remains fixed; finally, insert the estimate  $\tilde{\vartheta}(n)$  in place of  $\vartheta'$ . Making use of (37) the resulting estimator for the parameter  $\tau > 0$  is  $\tilde{\tau}(n)$  as specified in (42).

**Proposition 3** As  $n \rightarrow \infty$ , rescaled estimation errors of the estimator (42) admit an expansion

$$\sqrt{n} (\tilde{\tau}(n) - \tau) = -2\tau \frac{1}{\sqrt{cn}} \int_0^n X_s dW_s + o_{Q_\theta}(1) = \frac{-\frac{1}{\sqrt{cn}} \int_0^n X_s dW_s}{\frac{1}{cn} \int_0^n X_s^2 ds} + o_{Q_\theta}(1)$$

under  $\theta = (\vartheta, \tau) \in \Theta$ , with  $X$  solution to the Ornstein Uhlenbeck SDE (15) driven by  $W$ .

**Proof** Combining (42) with (37) we have

$$\tilde{\tau}(n) - \tau = \frac{\sum_{i=0}^p \tilde{\vartheta}_i(n) \int_0^n s^i (dY_s + \tau Y_s ds) - \int_0^n Y_s (dY_s + \tau Y_s ds)}{\int_0^n Y_s^2 ds - \tilde{\vartheta}(n)^\top \tilde{J}(n) \tilde{\vartheta}(n)}. \quad (43)$$

(1) Consider the numerator

$$\sum_{i=0}^p \tilde{\vartheta}_i(n) \int_0^n s^i (dY_s + \tau Y_s ds) - \int_0^n Y_s (dY_s + \tau Y_s ds) \quad (44)$$

on the right hand side of (43). (37) and (38) allow to write

$$\begin{aligned} \int_0^n Y_s \tau R_\vartheta(s) ds &= \sum_{i=0}^p \vartheta_i \int_0^n s^i \tau Y_s ds = \tau \vartheta^\top \tilde{J}(n) \tilde{\vartheta}(n) = \tau \tilde{\vartheta}(n)^\top \tilde{J}(n) \vartheta \\ &= \sum_{i=0}^p \tilde{\vartheta}_i(n) \int_0^n s^i \tau R_\vartheta(s) ds. \end{aligned}$$

Adding and subtracting this expression, (44) takes the form

$$\sum_{i=0}^p \tilde{\vartheta}_i(n) \int_0^n s^i (dY_s + \tau Y_s ds - \tau R_\vartheta(s) ds) - \int_0^n Y_s (dY_s + \tau Y_s ds - \tau R_\vartheta(s) ds). \quad (45)$$

Exploiting first (19) and then (37) and (30) we can write

$$\begin{aligned} \int_0^n Y_s R'_\vartheta(s) ds &= \sum_{j=1}^p \vartheta_j \int_0^n j s^{j-1} Y_s ds = \sum_{j=1}^p \vartheta_j j \sum_{k=0}^p \frac{1}{(j-1) + k + 1} n^{(j-1)+k+1} \tilde{\vartheta}_k(n) \\ &= \sum_{k=0}^p \tilde{\vartheta}_k(n) \sum_{j=0}^{p-1} \frac{1}{k+j+1} n^{k+j+1} (j+1) \vartheta_{j+1} \\ &= \sum_{k=0}^p \tilde{\vartheta}_k(n) \int_0^n s^k \sum_{j=1}^p j \vartheta_j s^{j-1} ds = \tilde{\vartheta}(n)^\top \begin{pmatrix} \int_0^n s^0 R'_\vartheta(s) ds \\ \int_0^n s^1 R'_\vartheta(s) ds \\ \vdots \\ \int_0^n s^p R'_\vartheta(s) ds \end{pmatrix}. \end{aligned}$$

Adding and subtracting this expression to (45) we thus can write (44) as

$$\sum_{i=0}^p \tilde{\vartheta}_i(n) \int_0^n s^i (dY_s + \tau Y_s ds - [R'_\vartheta(s) + \tau R_\vartheta(s)] ds) - \int_0^n Y_s (dY_s + \tau Y_s ds - [R'_\vartheta(s) + \tau R_\vartheta(s)] ds)$$

which in virtue of (21) and (16) equals

$$\sum_{i=0}^p \tilde{\vartheta}_i(n) \int_0^n s^i \sqrt{c} dW_s - \int_0^n Y_s \sqrt{c} dW_s.$$

Using again (14), we have reduced the numerator (44) on the right hand side of (43) to

$$\sum_{i=0}^p (\tilde{\vartheta}_i(n) - \vartheta_i) \int_0^n s^i \sqrt{c} dW_s - \int_0^n X_s \sqrt{c} dW_s. \quad (46)$$

As in lemma 4, joint laws

$$\mathcal{L} \left( \left( \frac{1}{\sqrt{n}} \int_0^{tn} s^0 dW_s, \dots, \frac{1}{\sqrt{n^{2p+1}}} \int_0^{tn} s^p dW_s \right)_{t \geq 0} \right)$$

do not depend on  $n$ , whereas by proposition 2 rescaled estimation errors

$$\left( \sqrt{n^{2i+1}} (\tilde{\vartheta}_i(n) - \vartheta_i) \right)_{i=0,1,\dots,p} \quad \text{under } Q_\vartheta$$

converge in law as  $n \rightarrow \infty$ , and thus are tight as  $n \rightarrow \infty$ . Terms  $\int_0^n X_s dW_s$  in (46) are of stochastic order  $O_{Q_\vartheta}(\sqrt{n})$  as  $n \rightarrow \infty$ , by proposition 1. As a consequence, our final representation (46) of the numerator (44) on the right hand side of (43) allows to write the rescaled estimation error as

$$\tilde{\tau}(n) - \tau = \frac{-\int_0^n X_s \sqrt{c} dW_s + O_{Q_\vartheta}(1)}{\int_0^n Y_s^2 ds - \tilde{\vartheta}(n)^\top \tilde{J}(n) \tilde{\vartheta}(n)}. \quad (47)$$

(2) We consider the denominator

$$\int_0^n Y_s^2 ds - \tilde{\vartheta}(n)^\top \tilde{J}(n) \tilde{\vartheta}(n) \quad (48)$$

on the right hand side of (47)—i.e. on the right hand side of (43)—which we write as

$$\int_0^n Y_s^2 ds - (\tilde{\vartheta}(n) - \vartheta)^\top \tilde{J}(n) (\tilde{\vartheta}(n) - \vartheta) - 2\vartheta^\top \tilde{J}(n) (\tilde{\vartheta}(n) - \vartheta) - \vartheta^\top \tilde{J}(n) \vartheta.$$

From (14)+(19) we have

$$\int_0^n Y_s^2 ds = \int_0^n X_s^2 ds + 2 \sum_{i=0}^p \vartheta_i \int_0^n s^i X_s ds + \vartheta^\top \tilde{J}(n) \vartheta$$

whereas (39) shows

$$\vartheta^\top \tilde{J}(n) (\tilde{\vartheta}(n) - \vartheta) = \sum_{i=0}^p \vartheta_i \int_0^n s^i X_s ds.$$

Thus we have reduced the denominator (48) to

$$\int_0^n X_s^2 ds - (\tilde{\vartheta}(n) - \vartheta)^\top \tilde{J}(n) (\tilde{\vartheta}(n) - \vartheta) .$$

The first summand in this expression is  $O_{Q_\theta}(n)$ , by lemma 1, whereas the second summand

$$(\tilde{\vartheta}(n) - \vartheta)^\top \tilde{J}(n) (\tilde{\vartheta}(n) - \vartheta) = (\tilde{\psi}_n^{-1}(\tilde{\vartheta}(n) - \vartheta))^\top \tilde{J}(1) (\tilde{\psi}_n^{-1}(\tilde{\vartheta}(n) - \vartheta))$$

converges in law as  $n \rightarrow \infty$  under  $Q_\theta$ , by (40) and proposition 2, and thus is tight as  $n \rightarrow \infty$ . Taking all this together, the denominator (48) on the right hand side of (43) under  $Q_\theta$  satisfies

$$\int_0^n Y_s^2 ds - \tilde{\vartheta}(n)^\top \tilde{J}(n) \tilde{\vartheta}(n) = \int_0^n X_s^2 ds + O_{Q_\theta}(1), \quad n \rightarrow \infty . \quad (49)$$

(3) The proof is finished: taking together (43), (47) and (49), we have

$$(\tilde{\tau}(n) - \tau) = \frac{-\int_0^n X_s \sqrt{c} dW_s + O_{Q_\theta}(1)}{\int_0^n X_s^2 ds + O_{Q_\theta}(1)}$$

and thus

$$\sqrt{n} (\tilde{\tau}(n) - \tau) = \frac{-\frac{1}{\sqrt{cn}} \int_0^n X_s dW_s + O_{Q_\theta}(\frac{1}{\sqrt{n}})}{\frac{1}{cn} \int_0^n X_s^2 ds + O_{Q_\theta}(\frac{1}{n})} .$$

By lemma 1,  $\frac{1}{n} \int_0^n X_s^2 ds$  converges  $Q_\theta$ -almost surely to  $\frac{c}{2\tau}$ : so proposition 3 is proved.  $\square$

### 3.4 Efficiency in the model (14)+(19)

We can put together the results of subsections 3.2 and 3.3 to prove that for every  $\theta \in \Theta$  as  $n \rightarrow \infty$ ,

$$\tilde{\theta}(n) := (\tilde{\vartheta}(n), \tilde{\tau}(n))$$

is an asymptotically efficient estimator sequence in the sense of the local asymptotic minimax theorem.

**Theorem 2** *Observing  $Y$  in (14)+(19) over the time interval  $[0, n]$  as  $n \rightarrow \infty$ , the sequence*

$$\tilde{\theta}(n) := (\tilde{\vartheta}(n), \tilde{\tau}(n))$$

*defined by (34) and (42) is such that representation (33) of corollary 1 in Sect. 3.1 holds as  $n \rightarrow \infty$ :*

$$\psi_n^{-1}(\tilde{\theta}(n) - \theta) = [J_{n,\theta}(1)]^{-1} S_{n,\theta}(1) + o_{Q_\theta}(1) = J^{-1}(1) S_{n,\theta}(1) + o_{Q_\theta}(1) .$$

*The estimator sequence  $(\tilde{\theta}(n))_n$  is thus efficient at  $\theta$  in the sense of the local asymptotic minimax theorem. This holds for all  $\theta = (\vartheta, \tau) \in \Theta$ .*

**Proof** If we compare the set of definitions for  $S_{n,\theta}$  in (27),  $J$  in (28),  $\psi_n$  in (25) to the set of definitions for  $\tilde{S}_{n,\theta}$  in (35),  $\tilde{J}$  in (30),  $\tilde{\psi}_n$  in (36), we can merge the assertions of propositions 2 and 3

$$\tilde{\psi}_n^{-1}(\tilde{\vartheta}(n) - \vartheta) = \left[ \frac{\tau^2}{c} \tilde{J}(1) \right]^{-1} \tilde{S}_{n,\theta}(1) + o_{Q_\theta}(1)$$

$$\sqrt{n} (\tilde{\tau}(n) - \tau) = -2\tau \frac{1}{\sqrt{cn}} \int_0^n X_s dW_s + o_{Q_\theta}(1)$$

under  $Q_\theta$  as  $n \rightarrow \infty$  into one assertion

$$\psi_n^{-1}(\tilde{\theta}(n) - \theta) = J^{-1}(1) S_{n,\theta}(1) + o_{Q_\theta}(1).$$

Together with proposition 1 (a) in Sect. 3.1, this shows that condition (33) of corollary 1 in Sect. 3.1 is satisfied. But the last condition implies asymptotic efficiency of an estimator sequence for the unknown parameter in the model (14)+(19) at  $\theta = (\vartheta, \tau) \in \Theta$ .  $\square$

## 4 Application: inference in stochastic Hodgkin–Huxley models

Hodgkin–Huxley models play an important role in neuroscience and are considered as realistic models for the spiking behaviour of neurons (see Hodgkin and Huxley 1952; Izhikevich 2007; Ermentrout and Terman 2010). The classical deterministic model with constant rate of input is a 4-dimensional dynamical system with variables  $(V, n, m, h)$

$$\begin{cases} dV_t = a dt - F(V_t, n_t, m_t, h_t) dt \\ dn_t = [\alpha_n(V_t)(1 - n_t) - \beta_n(V_t)n_t] dt \\ dm_t = [\alpha_m(V_t)(1 - m_t) - \beta_m(V_t)m_t] dt \\ dh_t = [\alpha_h(V_t)(1 - h_t) - \beta_h(V_t)h_t] dt \end{cases} \quad (50)$$

where  $a > 0$  is a constant. The functions  $(V, n, m, h) \rightarrow F(V, n, m, h)$  and  $V \rightarrow \alpha_j(V)$ ,  $V \rightarrow \beta_j(V)$ ,  $j \in \{n, m, h\}$ , are those of Izhikevich (2007) pp. 37–38 (i.e. the same as in Höpfner et al. (2016a), section 2.1).  $V$  takes values in  $\mathbb{R}$  and models the membrane potential in the single neuron. The variables  $n, m, h$  are termed gating variables and take values in  $[0, 1]$ . Write  $E_4 := \mathbb{R} \times [0, 1]^3$  for the state space.

Depending on the value of the constant  $a > 0$ , the following behaviour of the deterministic dynamical system is known, see Ermentrout and Terman (2010) pp. 63–66. On some interval  $(0, a_1)$  there is a stable equilibrium point for the system. There is a bistability interval  $\mathbb{I}_{bs} = (a_1, a_2)$  on which a stable orbit coexists with a stable equilibrium point. There is an interval  $(a_2, a_3)$  on which a stable orbit exists together with an unstable equilibrium point. At  $a = a_3$  orbits collapse into equilibrium; for  $a > a_3$  the equilibrium point is again stable. Here  $0 < a_1 < a_2 < a_3 < \infty$  are suitably determined<sup>1</sup> endpoints for intervals. Equilibrium points and orbits depend on the value of  $a$ . Evolution of the system along an orbit yields a remarkable excursion of the membrane potential  $V$  which we interpret as a spike.

In simulations, the equilibrium point appears to be globally attractive on  $(0, a_1)$ , the orbit appears to be globally attractive on  $(a_2, a_3)$ ; on the bistability interval  $\mathbb{I}_{bs} = (a_1, a_2)$ , the behaviour of the system depends on the choice of the starting value: simulated trajectories with randomly chosen starting point either spiral into the stable equilibrium, or are attracted by the stable orbit.

We feed noise into the system. Prepare an Ornstein–Uhlenbeck process (15) with parameter  $\tau > 0$

$$dX_t = -\tau X_t dt + \sqrt{c} dW_t \quad (51)$$

<sup>1</sup> Note that the constants of Ermentrout and Terman (2010) are different from the constants of Izhikevich (2007) which we use for the Hodgkin–Huxley model (50). With constants from Izhikevich (2007), simulations localize  $a_1 = \inf \mathbb{I}_{bs}$  between 5.24 and 5.25, and  $a_2 = \sup \mathbb{I}_{bs}$  close to 8.4; the value of  $a_3$  is  $\approx 163.5$  and thus far beyond any ‘biologically relevant’ value for the parameter  $a$ . Numerical calculations and simulations related to the bistability interval have been done in Hummel (2019).

and replace input  $a dt$  in the deterministic system (50) above by increments (16)

$$dY_t = \vartheta(1 + \tau t) dt - \tau dY_t + \sqrt{c} dW_t \quad (52)$$

of the stochastic process  $Y$  in (14) which depends on the parameter  $\vartheta > 0$ :

$$Y_t = \vartheta t + X_t, \quad t \geq 0. \quad (53)$$

This yields a stochastic Hodgkin–Huxley model

$$\begin{cases} dV_t = dY_t - F(V_t, n_t, m_t, h_t) dt \\ dn_t = [\alpha_n(V_t)(1 - n_t) - \beta_n(V_t)n_t] dt \\ dm_t = [\alpha_m(V_t)(1 - m_t) - \beta_m(V_t)m_t] dt \\ dh_t = [\alpha_h(V_t)(1 - h_t) - \beta_h(V_t)h_t] dt \end{cases} \quad (54)$$

with parameters  $\vartheta > 0$  and  $\tau > 0$ . By (54), the 5-dimensional stochastic system

$$\mathbb{X} = (\mathbb{X}_t)_{t \geq 0}, \quad \mathbb{X}_t := (V_t, n_t, m_t, h_t, Y_t)$$

is strongly Markov with state space  $E_5 := \mathbb{R} \times [0, 1]^3 \times \mathbb{R}$ . Stochastic Hodgkin–Huxley models where stochastic input encodes a periodic signal have been considered in Höpfner et al. (2016a), (2016b), (2017) and in Holbach (2019). A biological interpretation of the model (54) is as follows. The structure  $dY_t = \vartheta dt + dX_t$  of input reflects superposition of some global level  $\vartheta > 0$  of excitation through the network with ‘noise’ in the single neuron. Noise arises out of accumulation and decay of a large number of small postsynaptic charges, due to the spikes—excitatory or inhibitory, and registered at synapses along the dendritic tree—in spike trains which the neuron receives from a large number of other neurons in the network.

In simulations, the stochastic Hodgkin–Huxley model (54) which we consider in this section exhibits the following behaviour. For values of  $\vartheta$  in neighbourhoods of the bistability interval  $\mathbb{I}_{bs}$  of (50), the system (54) alternates (possibly long) time periods of seemingly regular spiking with (possibly long) time periods of quiet behaviour. ‘Quiet’ means that the system performs small random oscillations in neighbourhoods of some typical point. For smaller values of  $\vartheta$ , quiet behaviour prevails, for larger values of  $\vartheta$  we see an almost regular spiking.

The aim of the present section is estimation of an unknown parameter

$$\theta = (\vartheta, \tau) \in \Theta, \quad \Theta := (0, \infty)^2$$

in the system (54) with stochastic input (52), based on observation of the membrane potential  $V$  over a long time interval. For this, our standing assumption will be:

a starting value  $\mathbb{X}_0 \equiv (V_0, n_0, m_0, h_0, Y_0) \in \text{int}(E_5)$  is deterministic, fixed and known. (55)

Assuming (55) we recover first, for the internal variables  $j \in \{n, m, h\}$ , the state  $j_t$  at time  $t$  from the trajectory of  $V$  up to time  $t$

$$\check{j}_t := j_0 e^{-\int_0^t (\alpha_j + \beta_j)(V_r) dr} + \int_0^t \alpha_j(V_s) e^{-\int_s^t (\alpha_j + \beta_j)(V_r) dr} ds, \quad t \geq 0, \quad (56)$$

and then, in virtue of the first equation in (54), the state  $Y_t$  at time  $t$  of the process (53) of accumulated dendritic input from the trajectory of  $V$  up to time  $t$ :

$$\check{Y}_t = Y_0 + (V_t - V_0) + \int_0^t F(V_t, \check{n}_t, \check{m}_t, \check{h}_t) dt, \quad t \geq 0; \quad (57)$$

here and below we write  $\check{\cdot}$  to distinguish reconstructed variables. Thus  $(V, \check{n}, \check{m}, \check{h}, \check{Y})$  reconstructs the trajectory of the 5-dimensional stochastic system  $\mathbb{X}$  from observation of the membrane potential  $V$  and given starting point satisfying assumption (55). The motivation is from biology. For single neurons in active networks, the membrane potential can be measured intracellularly with very good time resolution, whereas the gating variables  $j \in \{n, m, h\}$  in the stochastic system (54) represent averages over large numbers of certain ion channels and are not accessible to measurement.

Under assumption (55), the problem of estimating the unknown parameter  $\theta = (\vartheta, \tau) \in \Theta$  in the stochastic Hodgkin–Huxley system (54) based on observation of the membrane potential can be formulated as follows. Consider the canonical path space  $(C, \mathcal{C})$ ,  $C := C([0, \infty), \mathbb{R}^5)$ , equipped with the canonical process  $\pi = (\pi_t)_{t \geq 0}$  and the canonical filtration  $\mathbb{G} = (\mathcal{G}_t)_{t \geq 0}$ , and also the smaller filtration

$$\mathbb{G}^{(1)} = (\mathcal{G}_t^{(1)})_{t \geq 0}, \quad \mathcal{G}_t^{(1)} := \bigcap_{r > t} \sigma \left( \pi_0, \pi_s^{(1)} : 0 \leq s \leq r \right)$$

generated by observation of the first component  $\pi^{(1)}$  of the canonical process  $\pi$  knowing the starting point  $\pi_0$  of  $\pi$ . For  $\theta \in \Theta$ , let  $Q_\theta$  denote the law of the process  $\mathbb{X}$  under  $\theta = (\vartheta, \tau)$  on  $(C, \mathcal{C})$ , with starting point (55) not depending on  $\theta$ . On  $(C, \mathcal{C})$  we write for short

$$\zeta = (\zeta_t)_{t \geq 0}, \quad \zeta := \check{\pi}^{(5)} \quad (58)$$

for the reconstruction  $\check{\pi}^{(5)}$  of the fifth component  $\pi^{(5)}$  of  $\pi$  (which under  $Q_\theta$  represents accumulated dendritic input  $Y_t = \vartheta t + X_t$ ,  $t \geq 0$ ) from the first component  $\pi^{(1)}$  (which under  $Q_\theta$  represents the membrane potential  $V$ ) and the starting point  $\pi_0$ ; on the lines of (57) we have

$$\zeta_t = \pi_0^{(5)} + (\pi_t^{(1)} - \pi_0^{(1)}) + \int_0^t F(\pi_s^{(1)}, \check{\pi}_s^{(2)}, \check{\pi}_s^{(3)}, \check{\pi}_s^{(4)}) ds, \quad t \geq 0. \quad (59)$$

By definition of  $\mathbb{G}^{(1)}$ , the observed process  $\pi^{(1)}$  and the reconstructed processes  $\zeta$ ,  $\check{\pi}^{(j)}$ ,  $j \in \{2, 3, 4\}$ , are  $\mathbb{G}^{(1)}$ -semimartingales. Write  $\sqrt{c} W$  for the  $\mathbb{G}^{(1)}$ -martingale part of  $\zeta$  or of  $\pi^{(1)}$  under  $Q_\theta$ . The likelihood ratio process of  $Q_{\theta'}$  with respect to  $Q_\theta$  relative to  $\mathbb{G}^{(1)}$  is obtained in analogy to (23)+(24), special case  $R_\vartheta(s) = \vartheta s$ . Then the following is proposition 3.2 in Holbach (2020):

**Proposition 4** (Holbach 2020) *For pairs  $\theta' = (\vartheta', \tau')$ ,  $\theta = (\vartheta, \tau)$  in  $\Theta = (0, \infty)^2$ , writing*

$$M_t^{\theta'/\theta} := \frac{1}{\sqrt{c}} \int_0^t \{(\vartheta' - \vartheta)(1 + \tau s) - (\tau' - \tau)(\zeta_s - \vartheta s) + (\tau' - \tau)(\vartheta' - \vartheta)s\} dW_s, \quad t \geq 0,$$

*likelihood ratios in the statistical model*

$$\left( C, \mathcal{C}, \mathbb{G}^{(1)}, \{Q_\theta : \theta \in \Theta\} \right)$$

*are given by*

$$L_t^{\theta'/\theta} = \exp \left( M_t^{\theta'/\theta} - \frac{1}{2} \langle M^{\theta'/\theta} \rangle_t \right), \quad t \geq 0 \quad (60)$$

*where  $\langle M^{\theta'/\theta} \rangle$  denotes the angle bracket of the martingale  $M^{\theta'/\theta}$  under  $Q_\theta$  relative to  $\mathbb{G}^{(1)}$ .*

Note that under  $Q_\theta$ , the  $\mathbb{G}^{(1)}$ -adapted process  $(\zeta_t - \vartheta t)_{t \geq 0}$  in the integrand of  $M^{\theta'/\theta}$  represents the Ornstein Uhlenbeck process  $X$  of equation (51); the constant  $c$  is known from quadratic variation of  $\zeta$ .



We know everything about the likelihoods (60): they are the likelihoods in the submodel where  $\vartheta_0 \equiv 0$  is fixed of the model considered in Sect. 3, case  $p := 1$ . As a consequence, in the statistical model associated to the stochastic Hodgkin Huxley model, we have LAN at  $\theta$ , with local scale  $\frac{1}{\sqrt{n^3}}$  for the  $\vartheta$ -component and  $\frac{1}{\sqrt{n}}$  for the  $\tau$ -component,  $\theta = (\vartheta, \tau) \in \Theta$ . We have a characterization of efficient estimators by the local asymptotic minimax theorem, and we did construct asymptotically efficient estimators. Since  $\zeta$  is a  $\mathbb{G}^{(1)}$ -semimartingale, common  $\mathbb{G}^{(1)}$ -adapted determinations for  $\theta \in \Theta$  of the stochastic integrals  $\int s d\zeta_s$  and  $\int \zeta_s d\zeta_s$  exist. According to (34), (37) and (42), we estimate the first component of the unknown parameter  $\theta = (\vartheta, \tau)$  in  $\Theta = (0, \infty)^2$  in the system (54) by

$$\check{\vartheta}(t) := \operatorname{arginf}_{\vartheta'} \int_0^t (\zeta_s - \vartheta' s)^2 ds = \frac{3}{t^3} \int_0^t s \zeta_s ds \quad (61)$$

and then the second component by

$$\check{\tau}(t) := \frac{\check{\vartheta}(t) \int_0^t s d\zeta_s - \int_0^t \zeta_s d\check{\zeta}_s}{\int_0^t \check{\zeta}_s^2 ds - [\check{\vartheta}(t)]^2 t^3/3}. \quad (62)$$

The estimators  $\check{\vartheta}(t)$ ,  $\check{\tau}(t)$  are  $\mathbb{G}_t^{(1)}$ -measurable,  $t \geq 0$ . The structure of the likelihoods (60) is the structure of the likelihoods in Sect. 3 with  $p := 1$ , submodel  $\vartheta_0 \equiv 0$ . The structure of the pair  $(\check{\vartheta}(t), \check{\tau}(t))$  in (61)+(62) is the structure of the estimators  $(\tilde{\vartheta}(t), \tilde{\tau}(t))$  in Sect. 3 with  $p := 1$ , submodel  $\vartheta_0 \equiv 0$ . Under  $Q_\theta$ , we have from (27) and (28) and theorem 2 in Sect. 3.4 a representation

$$\begin{pmatrix} \sqrt{n^3} (\check{\vartheta}(n) - \vartheta) \\ \sqrt{n} (\check{\tau}(n) - \tau) \end{pmatrix} = \begin{pmatrix} \frac{\tau^2}{3c} & 0 \\ 0 & \frac{1}{2\tau} \end{pmatrix}^{-1} \begin{pmatrix} \frac{1}{\sqrt{cn^3}} \int_0^n (1 + \tau s) dW_s \\ \frac{-1}{\sqrt{cn}} \int_0^n (\zeta_s - \vartheta s) dW_s \end{pmatrix} + o_{Q_\theta}(1) \quad (63)$$

of rescaled estimation errors as  $n \rightarrow \infty$ , and proposition 1 in Sect. 3.1 shows convergence in law

$$\begin{pmatrix} \sqrt{n^3} (\check{\vartheta}(n) - \vartheta) \\ \sqrt{n} (\check{\tau}(n) - \tau) \end{pmatrix} \longrightarrow \begin{pmatrix} \frac{3\sqrt{c}}{\tau} \int_0^1 s d\tilde{W}_s^{(1)} \\ \sqrt{2\tau} \tilde{W}_1^{(2)} \end{pmatrix}$$

under  $Q_\theta$  as  $n \rightarrow \infty$ , with some two-dimensional standard Brownian motion  $\tilde{W}$ . Consider on  $(C, \mathcal{C}, (\mathcal{G}_{tn}^{(1)})_{t \geq 0}, \{Q_\theta : \theta \in \Theta\})$  martingales

$$\check{S}_{n,\theta}(t) := \begin{pmatrix} \frac{1}{\sqrt{cn^3}} \int_0^{tn} (1 + \tau s) dW_s \\ \frac{-1}{\sqrt{cn}} \int_0^{tn} (\zeta_s - \vartheta s) dW_s \end{pmatrix}$$

under  $Q_\theta$ , and let  $\check{J}_{n,\theta}$  denote their angle brackets under  $Q_\theta$ . Define local scale

$$\check{\psi}_n := \begin{pmatrix} \frac{1}{\sqrt{n^3}} & 0 \\ 0 & \frac{1}{\sqrt{n}} \end{pmatrix}$$

and limit information

$$\check{J}(t) := \begin{pmatrix} \frac{\tau^2}{3c} t^3 & 0 \\ 0 & \frac{1}{2\tau} t \end{pmatrix}.$$

With these notations, theorem 1 in Sect. 3.1 and theorem 2 in Sect. 3.4 yield:

**Theorem 3** *In the sequence of statistical models*

$$(C, \mathcal{C}, (\mathcal{G}_{tn}^{(1)})_{t \geq 0}, \{Q_\theta : \theta \in \Theta\})$$

the following holds at every point  $\theta = (\vartheta, \tau)$  in  $\Theta = (0, \infty)^2$ :

(a) we have LAN at  $\theta$  with local scale  $(\check{\psi}_n)_n$  and local parameter  $h \in \mathbb{R}^2$ :

$$\log L_{tn}^{\theta + \check{\psi}_n h / \theta} = h^\top \check{S}_{n,\theta}(t) - \frac{1}{2} h^\top \check{J}_{n,\theta}(t) h + o_{Q_\theta}(1), \quad n \rightarrow \infty;$$

(b) by (63), rescaled estimation errors of  $\check{\theta}(n) := (\check{\vartheta}(n), \check{\tau}(n))$  admit the expansion

$$\check{\psi}_n^{-1}(\check{\theta}(n) - \theta) = [\check{J}(1)]^{-1} \check{S}_{n,\theta}(1) + o_{Q_\theta}(1) = [\check{J}_{n,\theta}(1)]^{-1} \check{S}_{n,\theta}(1) + o_{Q_\theta}(1), \quad n \rightarrow \infty.$$

When we observe—for some given starting point of the system—the membrane potential in a stochastic Hodgkin–Huxley model (54) up to time  $n$ , with accumulated stochastic input defined by (53) together with (51) which depends on an unknown parameter  $\theta = (\vartheta, \tau)$  in  $\Theta = (0, \infty)^2$ , the following resumes in somewhat loose language the assertion of the local asymptotic minimax theorem, corollary 1 in Sect. 3.1:

**Corollary 2** *For loss functions  $L : \mathbb{R}^2 \rightarrow [0, \infty)$  which are continuous, subconvex and bounded, for  $0 < C < \infty$  arbitrary, maximal risk over shrinking neighbourhoods of  $\theta$*

$$\lim_{n \rightarrow \infty} \sup_{|h| \leq C} E_{\left(\vartheta + h_1/\sqrt{n^3}, \tau + h_2/\sqrt{n}\right)} \left( L \left( \frac{\sqrt{n^3}(\check{\vartheta}(n) - (\vartheta + h_1/\sqrt{n^3}))}{\sqrt{n}(\check{\tau}(n) - (\tau + h_2/\sqrt{n}))} \right) \right)$$

converges as  $n \rightarrow \infty$  to

$$E \left( L \left( \frac{\frac{3\sqrt{C}}{\tau} \int_0^1 s d\tilde{W}_s^{(1)}}{\sqrt{2}\tau \tilde{W}_1^{(2)}} \right) \right).$$

Within the class of  $(\mathcal{G}_n^{(1)})_n$ -adapted estimator sequences  $(T_n)_n$  whose rescaled estimation errors at  $\theta = (\vartheta, \tau)$  are tight—at rate  $\sqrt{n^3}$  for the  $\vartheta$ -component, and at rate  $\sqrt{n}$  for the  $\tau$ -component—it is impossible to outperform the sequence  $(\check{\vartheta}(n), \check{\tau}(n))$  defined by (61)+(62), asymptotically as  $n \rightarrow \infty$ .

Note that we are free to measure risk through any loss function which is continuous, subconvex and bounded.

**Funding** Open Access funding provided by Projekt DEAL.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Bingham N, Goldie C, Teugels J (1987) Regular variation. Cambridge University Press, Cambridge
- Davies R (1985) Asymptotic inference when the amount of information is random. In LeCam L, Olshen R (eds), Proceedings of the Berkeley symposium in honor of J. Neyman and J. Kiefer. Vol. 2, Wadsworth Belmont
- Dehling H, Franke B, Kott T (2010) Drift estimation for a periodic mean reversion process. *Stat Inference Stoch Process* 13:175–192
- Dehling H, Franke B, Woerner J (2017) Estimating drift parameters in a fractional Ornstein Uhlenbeck process with periodic mean. *Stat Inference Stoch Process* 20:1–14
- Ermentrout G, Terman D (2010) Mathematical foundations of neuroscience. Springer, Berlin
- Franke B, Kott T (2013) Parameter estimation for the drift of a time-inhomogeneous jump diffusion process. *Stat Neerl* 76:145–168
- Hajek J (1970) A characterization of limiting distributions for regular estimators. *Z Wahrscheinlichkeitstheorie und Verw Geb* 14:323–330
- Hodgkin A, Huxley A (1952) A quantitative description of ion currents and its application to conduction and excitation in nerve membranes. *J Physiol* 117:500–544
- Holbach S (2020) Positive Harris recurrence for degenerate diffusions with internal variables and randomly perturbed time-periodic input. *Stoch. Processes Applications*, in press, <https://doi.org/10.1016/j.spa.2020.07.005>. [arXiv:1907.13585](https://arxiv.org/abs/1907.13585)
- Holbach S (2019) Local asymptotic normality for shape and periodicity of a signal in the drift of a degenerate diffusion with internal variables. *Electron J Stat* 13:4884–4915
- Höpfner R (2014) Asymptotic statistics. DeGruyter, Berlin
- Höpfner R, Kutoyants Yu (2009) On LAN for parametrized continuous periodic signals in a time inhomogeneous diffusion. *Stat Decis* 27:309–326
- Höpfner R, Kutoyants Yu (2011) Estimating a periodicity parameter in the drift of a time inhomogeneous diffusion. *Math Methods Stat* 20:58–74
- Höpfner R, Löcherbach E, Thieullen M (2016) Ergodicity for a stochastic Hodgkin-Huxley model driven by Ornstein-Uhlenbeck type input. *Ann Instit Henri Poincaré* 52(1):483–501
- Höpfner R, Löcherbach E, Thieullen M (2016) Ergodicity and limit theorems for degenerate diffusions with time periodic drift. Application to a stochastic Hodgkin-Huxley model. *ESAIM PS* 20:527–554
- Höpfner R, Löcherbach E, Thieullen M (2017) Strongly degenerate time inhomogeneous SDEs: Densities and support properties. Application to Hodgkin-Huxley type systems. *Bernoulli* 23(4A):2587–2616
- Hummel C (2019) Netzwerke von Hodgkin-Huxley Neuronen. Institut für Mathematik, Universität Mainz, Masterarbeit
- Ibragimov A, Khasminskii R (1981) Statistical estimation. Springer, New York
- Ikeda N, Watanabe S (1989) Stochastic differential equations and diffusion processes, 2nd edn. North-Holland / Kodansha, Tokyo
- Izhikevich E (2007) Dynamical systems in neuroscience: the geometry of excitation and bursting. MIT press, Cambridge
- Jacod J, Shiryaev A (1987) Limit theorems for stochastic processes. Springer, Berlin
- Kutoyants Yu (2004) Statistical inference for ergodic diffusion processes. Springer, London
- Le Cam L (1969) Théorie asymptotique de la décision statistique. Les Presses de l'Université de Montréal, Montréal
- Le Cam L, Yang G (2002) Asymptotics in statistics, 2nd edn. Springer, Berlin
- Lipster R, Shiryaev A (2001) Statistics of random processes, vol I+II, 2nd edn. Springer, Berlin
- Pchelintsev E (2013) Improved estimation in a non-Gaussian parametric regression. *Stat Inference Stoch Process* 16:15–28
- Pfanzagl J (1994) Parametric statistical theory. deGruyter, Berlin

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