

Local asymptotic normality for shape and periodicity of a signal in the drift of a degenerate diffusion with internal variables

Simon Holbach*

Fakultät für Mathematik, Universität Bielefeld, Postfach 10 01 31, 33501 Bielefeld, Germany
e-mail: sholbach@math.uni-bielefeld.de

Abstract: Taking a multidimensional time-homogeneous dynamical system and adding a randomly perturbed time-dependent deterministic signal to some of its components gives rise to a high-dimensional system of stochastic differential equations which is driven by possibly very low-dimensional noise. Equations of this type are commonly used in biology for modeling neurons or in statistical mechanics for certain Hamiltonian systems. Assuming that the signal depends on an unknown shape parameter ϑ and also has an unknown periodicity T , we prove Local Asymptotic Normality (LAN) jointly in ϑ and T for the statistical experiment arising from (partial) observation of this diffusion in continuous time. The local scale turns out to be $n^{-1/2}$ for ϑ and $n^{-3/2}$ for T . Our approach unifies and extends existing results on LAN in variants of the signal in noise model where the parameters ϑ and T are treated separately. Consequently, we can establish the same efficiency bounds in our more complex model and make use of efficient estimators known from these submodels.

MSC 2010 subject classifications: 62F12, 60J60.

Keywords and phrases: Local asymptotic normality, parametric signal estimation, degenerate diffusion, periodic drift, neuron models.

Received February 2019.

1. Introduction of the model and the problem

Let $U \subset \mathbb{R}^{N+L}$ be a σ -compact set and let $f: U \rightarrow \mathbb{R}^N$ and $g: U \rightarrow \mathbb{R}^L$ be locally Lipschitz continuous functions. Finally, let $S: [0, \infty) \rightarrow \mathbb{R}^N$ be a continuous periodic signal and consider the deterministic dynamical system

$$\begin{aligned}dX_t &= f(X_t, Y_t)dt + S(t)dt, \\dY_t &= g(X_t, Y_t)dt.\end{aligned}\tag{1}$$

This system is divided into two groups of variables: The N components of X whose dynamics depend directly on the signal and the L components of Y which

*This article features results from the author's PhD thesis [14] at Johannes Gutenberg-Universität Mainz. The author would like to thank Reinhard Höpfner for fruitful discussions and helpful remarks and suggestions.

are affected by the signal only indirectly through the influence of X . Intuitively speaking, we can think of (1) as a dynamical system with no intrinsic time-inhomogeneity which then receives an additional time-dependent *external input* S in some of its variables, while the remaining variables merely describe an interior mechanism. This is why we sometimes refer to X as the *adjustable variable(s)* and Y as the *internal variable(s)*. Note that the only source of time-inhomogeneity is indeed the signal – if the system receives constant external input $S \equiv c \in \mathbb{R}^N$ (or none at all, i.e. $c = 0$), it is homogeneous in time. Systems of this kind frequently arise in the context of neuroscience and statistical mechanics (see Examples 1.1 and 1.2 below).

We construct a stochastic model by following the idea that the signal is not actually received in its original shape, but is subject to random perturbations by external noise (i.e. noise that is independent of the rest of the system). To take account of this notion, it seems natural to substitute the signal term $S(t)dt$ in (1) with the increment dZ_t of a process taking values in a closed set $U' \subset \mathbb{R}^N$ and satisfying an SDE of the type

$$dZ_t = [S(t) + b(Z_t)]dt + \sigma(Z_t)dW_t,$$

where W is an M -dimensional standard Brownian Motion, while $b: U' \rightarrow \mathbb{R}^N$ and $\sigma: U' \rightarrow \mathbb{R}^{N \times M}$ are locally Lipschitz continuous drift and volatility functions. Note that this SDE can be viewed as a generalized Ornstein-Uhlenbeck type process with time-dependent mean-reversion level (think of $b(Z_t) = -\beta Z_t$ with $\beta \in (0, \infty)$). A particularly prominent special case is the classical signal in noise model (take $M = N = 1$, $b \equiv 0$, and $\sigma \equiv 1$, see for example [24, Example I.7.3, Chapter III.5]), which arises in a wide variety of fields including communication, radiolocation, seismic signal processing, or computer-aided diagnosis and has been the subject of extensive study.

Perturbing $S(t)$ randomly in this way leads to the stochastic dynamical system

$$\begin{aligned} dX_t &= f(X_t, Y_t)dt + dZ_t, \\ dY_t &= g(X_t, Y_t)dt, \\ dZ_t &= [S(t) + b(Z_t)]dt + \sigma(Z_t)dW_t, \end{aligned} \tag{2}$$

with state space

$$\mathbf{E} := U \times U' \subset \mathbb{R}^{N+L+N}.$$

This system can be thought of as degenerate in the following sense: Firstly, the equation for Y does not incorporate the driving Brownian Motion W explicitly, making it rather unclear which effect noise has on these components. Secondly, the dimension M of the driving Brownian Motion can (and will usually) be much lower than the dimension $N + L + N$ of the system. This is why we call a stochastic process satisfying a system of stochastic differential equations of the type (2) a *degenerate diffusion with internal variables and randomly perturbed time-inhomogeneous deterministic input*.

We now have three groups of variables: The entirely autonomous external input governed by dZ_t (the “noisy signal”), the components of X that are directly adjusted by the noisy signal, and the components of the internal variable Y whose dynamics are only indirectly affected by noise, since the respective differential equations incorporate neither Z nor the driving Brownian Motion W explicitly. Note that, for this reason, Y is conditionally deterministic given X and has continuously differentiable trajectories.

The system (2) is a generalization of the one introduced in equation (18) of Section 4.1 of [22], which is a probabilistic version of a class of dynamical systems that are well-known in the mathematical modeling of neurons (see Example 1.1 below). In [13] (which can be viewed as a companion article to the present one), we study the model (2) from a purely probabilistic standpoint and use methods from [23] to discuss sufficient conditions for the process (X, Y, Z) to be positive Harris recurrent. Before we explain the focus of the current article, let us introduce two major examples.

Example 1.1. Let $N = 1$, $L = 3$, $U = \mathbb{R} \times [0, 1]^3$ and consider the coefficient functions

$$f(x, y) = -36y_1^4(x + 12) - 120y_2^3y_3(x - 120) - 0.3(x - 10.6)$$

and

$$g(x, y) = \begin{pmatrix} \alpha_1(x)(1 - y_1) - \beta_1(x)y_1 \\ \alpha_2(x)(1 - y_2) - \beta_2(x)y_2 \\ \alpha_3(x)(1 - y_3) - \beta_3(x)y_3 \end{pmatrix}$$

with

$$\begin{aligned} \alpha_1(x) &= \begin{cases} \frac{0.1-0.01x}{\exp(1-0.1x)-1}, & x \neq 10, \\ 0.1, & \text{else,} \end{cases} & \beta_1(x) &= 0.125 \exp(-x/80), \\ \alpha_2(x) &= \begin{cases} \frac{2.5-0.1x}{\exp(2.5-0.1x)-1}, & x \neq 25, \\ 1, & \text{else,} \end{cases} & \beta_2(x) &= 4 \exp(-x/18), \\ \alpha_3(x) &= 0.07 \exp(-x/20), & \beta_3(x) &= \frac{1}{\exp(3-0.1x)+1}. \end{aligned}$$

for all $(x, y) = (x, y_1, y_2, y_3)^\top \in U$. The corresponding dynamical system (1) is known as the *Hodgkin-Huxley system* and it was first introduced by Hodgkin and Huxley in 1952 (see [11], note however that we use the slightly different model constants from [26]) with the aim of describing the initiation and propagation of action potentials in the cell membrane of a neuron in response to an external stimulus. While X is the membrane potential itself (usually labeled V in the literature), the internal variables Y_1 , Y_2 , and Y_3 (commonly denoted by n , m , and h) correspond to the ionic mechanism underlying its evolution. The two predominant ion currents in the cell membrane are import of sodium Na^+ and export of potassium K^+ through the membrane. Each of the internal variables signifies the probability that a specific type of gate in the respective ion channel is open at a given time. It is for this reason that n , m , and h are often called

gating variables. In the context of this model, the signal S represents the dendritic input which the neuron receives from a large number of other neurons, transported by an even larger number of synapses located on the respective dendritic tree. The resulting “total dendritic input” can then be thought of as an average of interdependent and repeating similar currents, which is why S is usually assumed to be periodic (or even constant). When modeling neurons, particular interest lies in the typical spiking behaviour of the membrane potential, a feature that is commonly agreed upon to be adequately described by the Hodgkin-Huxley model. For a more detailed modern introduction, interpretation, and an in-depth comparison with other neuron models, see for example [26] and [6].

Adding noise in the sense of (2) by choosing $\sigma \in C^\infty(U')$ and $b(Z_t) = -\beta Z_t$ with $\beta \in (0, \infty)$, we acquire the so-called *stochastic Hodgkin-Huxley model (with mean reverting Ornstein-Uhlenbeck type input)*. It was first introduced and studied by Höpfner, Löcherbach, and Thieullen in the series of the three papers [21], [22], and [23]. The constant β is determined by the so-called time constant of the membrane which represents spontaneous voltage decay not related to the input. For many types of neurons, the time constant is known from experiments (see [7]). A degree of freedom lies in the choice of the volatility σ which reflects the nature of the influence of noise. In the past, mean reverting Ornstein-Uhlenbeck type equations with various volatilities have been used to model the membrane potential itself (see for example [30] or [16]), and in a sense our stochastic Hodgkin-Huxley model can be viewed as a refinement of this kind of model. If σ is Lipschitz continuous, existence of a unique non-exploding strong solution taking values in $\mathbf{E} = \mathbb{R} \times [0, 1]^3 \times U'$ follows from the same arguments as in [21, Proposition 1] and [22, Proposition 2].

Analogously, one can introduce stochastic versions of simpler neuron models such as the FitzHugh-Nagumo model (see [26, equations (4.11) and (4.12)]) or the Morris-Lecar model (see [33] or, for a modern version, [35]).

Example 1.2. Systems of coupled oscillators are particularly intuitive Hamiltonian systems and several different stochastic models have been subject to research in the past (see e.g [10], [2], [34], [3]). The following example is inspired by the model from [4] to which we add a time-inhomogeneity and the corresponding external variables.

Let us think of three rotors, each given by their angle $q_i(t) \in \mathbb{R}$ and momentum $p_i(t) \in \mathbb{R}$ at the time $t \in [0, \infty)$ for each $i \in \{1, 2, 3\}$. Assuming their respective masses to be all equal to 1 and not taking into account units, the laws of classical mechanics imply

$$\dot{q}_i = p_i \quad \text{for all } i \in \{1, 2, 3\}. \quad (3)$$

We suppose that these rotors are coupled in row, i.e.

$$\begin{aligned} \dot{p}_1 &= w_1(q_2 - q_1) - u_1(q_1), \\ \dot{p}_2 &= -[w_1(q_2 - q_1) + w_3(q_2 - q_3)] - u_2(q_2), \\ \dot{p}_3 &= w_3(q_2 - q_3) - u_3(q_3), \end{aligned} \quad (4)$$

where $w_1, w_2, w_3: \mathbb{R} \rightarrow \mathbb{R}$ and $u_1, u_2, u_3: \mathbb{R} \rightarrow \mathbb{R}$ are related to interaction potentials and pinning potentials, respectively. A classical model is the one that arises if we let one or both of the outer rotors receive external torques and interact with Langevin type heat baths. In order to give a mathematical description of this, we fix $i \in \{1, 3\}$ for the remainder of this paragraph. Applying an external time-dependent torque $S_i: [0, \infty) \rightarrow \mathbb{R}$ to the i -th rotor means expanding the equation for p_i to

$$dp_i = [w_i(q_2 - q_i) - u_i(q_i)] dt + S_i dt,$$

which turns (3) and (4) into a system like (1). On top of that, we want to add interaction with a heat bath, i.e. for a temperature $\tau_i \in (0, \infty)$ and a dissipation constant $\delta_i \in (0, \infty)$, the equation for p_i is further expanded to

$$\begin{aligned} dp_i &= [w_i(q_2 - q_i) - u_i(q_i)] dt + S_i dt - \delta_i p_i dt + \sqrt{2\delta_i \tau_i} dW_t^{(i)} \\ &= [w_i(q_2 - q_i) - u_i(q_i) - \delta_i p_i] dt + \left[S_i dt + \sqrt{2\delta_i \tau_i} dW_t^{(i)} \right], \end{aligned}$$

where the last term in parentheses is the total sum of external influences. Following the spirit of (2), we may replace this term with the increments of a more general random perturbation of the torque: We take

$$dp_i = [w_i(q_2 - q_i) - u_i(q_i) - \delta_i p_i] dt + dZ_t^{(i)}$$

with

$$dZ_t^{(i)} = \left[S_i(t) + b_i(Z_t^{(i)}) \right] dt + \sigma_i(Z_t^{(i)}) dW_t^{(i)}$$

for some volatility $\sigma_i: \mathbb{R} \rightarrow \mathbb{R}$ and a drift $b_i: \mathbb{R} \rightarrow \mathbb{R}$. What we end up with is indeed a degenerate diffusion with internal variables and randomly perturbed time-inhomogeneous deterministic input as in (2). If only the first rotor in the chain receives an external input, the dimensions are $M = N = 1$ and $L = 5$, $U = \mathbb{R}^6$, $U' = \mathbb{R}$. If both of the outer rotors receive an external input, the dimensions are $M = N = 2$ and $L = 4$, $U = \mathbb{R}^6$, $U' = \mathbb{R}^2$.

In this article, we want to study a statistical model in which the deterministic signal S depends on a set of parameters. More precisely, we assume that there is an open set $\Theta \subset \mathbb{R}^D$ such that

$$S = S_{(\vartheta, T)} \quad \text{with } (\vartheta, T) \in \Theta \times (0, \infty),$$

where T is the signal's periodicity and ϑ is a D -dimensional shape parameter. A natural goal is to estimate ϑ and T simultaneously from continuous observation of the process. However, observing the process (X, Y, Z) entirely may not make sense in many models: The external variable Z can be of a rather abstract nature and, for example, in the Hodgkin-Huxley model from Example 1.1 the only variable that is arguably observable is the membrane potential X . In spite of that, Section 3.1 shows:

Result 1. As long as the initial configuration (X_0, Y_0, Z_0) is deterministic and known, it does not matter whether we can observe the entire process (X, Y, Z) , only the adjustable variable X , or only the external variable Z .

This is the content of Remark 3.1 and Proposition 3.2. Since Z is the most convenient process to handle statistically among all of these, our considerations in the sequel are confined to this external variable. Being able to relate statistical problems entirely to Z means that as long as this variable fits our setting, we can treat any example of (2) (including in particular those that were introduced in Examples 1.1 and 1.2). In Section 3.2, we prove an LAN result for the external variable (Theorem 3.7), generalizing [12, Theorem 2.3] in which we only treated the case $M = N = 1$. This can then be combined with the previous results in order to obtain:

Result 2. Under reasonable regularity conditions on the parametrization and under some non-degeneracy and ergodicity of the external variable Z , the sequence of statistical experiments corresponding to continuous observation of (X, Y, Z) over growing time intervals $[0, n]$ for $n \rightarrow \infty$ has the LAN property. The local scales are identified as $n^{-1/2}$ for the shape and $n^{-3/2}$ for the periodicity.

The rigorous and precise corresponding statement is Theorem 2.3. It allows for application to simultaneous estimation of shape and periodicity, as under LAN we can use Hájek's Convolution Theorem and the Local Asymptotic Minimax Theorem in order to establish optimality for estimators when the rescaled estimation errors are stochastically asymptotically equivalent to the central statistic of the experiment (see [31], [5], [29] or [15] for a detailed presentation of the relevant theory). While this is not the focus of the present article, we briefly comment on specific estimators in relevant submodels in Remark 3.10 and 3.11.

2. Main results and applications

First, let us recall and collect the basic assumptions that were mentioned in the introduction.

(A0) Basic setting: The state space is $\mathbf{E} = U \times U'$ where $U \subset \mathbb{R}^{N+L}$ is σ -compact and $U' \subset \mathbb{R}^N$ is closed. All of the coefficient functions f, g, b, σ are locally Lipschitz continuous and the signal $S_{(\vartheta, T)}$ is continuous, T -periodic with $T \in (0, \infty)$ and depends on some parameter ϑ taken from an open set $\Theta \subset \mathbb{R}^D$.

Throughout this article, (A0) will be a tacit standing assumption.

Using the notation $\Phi_t = (X_t, Y_t, Z_t)$ for all $t \in [0, \infty)$ and incorporating the

parameters, we rewrite the equation (2) as

$$d\Phi_t = B_{(\vartheta, T)}(t, \Phi_t)dt + \Sigma(\Phi_t)dW_t, \quad (5)$$

where

$$B_{(\vartheta, T)}: [0, \infty) \times \mathbf{E} \rightarrow \mathbb{R}^{N+L+N}, \quad (t, x, y, z) \mapsto \begin{pmatrix} f(x, y) + S_{(\vartheta, T)}(t) + b(z) \\ g(x, y) \\ S_{(\vartheta, T)}(t) + b(z) \end{pmatrix},$$

for each $(\vartheta, T) \in \Theta \times (0, \infty)$, while

$$\Sigma: \mathbf{E} \rightarrow \mathbb{R}^{(N+L+N) \times M}, \quad (x, y, z) \mapsto \begin{pmatrix} \sigma(z) \\ 0_{L \times M} \\ \sigma(z) \end{pmatrix}.$$

We fix some probability space $(\Omega, \mathcal{F}, \mathbb{P})$, and we consider the following assumptions about the SDE (5):

- (A1) Unique solvability:** For all $(\vartheta, T) \in \Theta \times (0, \infty)$ and all deterministic starting points $\Phi_0 \in \mathbf{E}$, the SDE (5) has a unique strong solution $\Phi^{(\vartheta, T)} = (X^{(\vartheta, T)}, Y^{(\vartheta, T)}, Z^{(\vartheta, T)}): [0, \infty) \rightarrow \mathbf{E}$ under \mathbb{P} .
- (A2) Bounded diffusion matrix:** The mapping $\sigma\sigma^\top: U' \rightarrow \mathbb{R}^{N \times N}$ is uniformly bounded away from 0 and from ∞ in the sense that there are $\sigma_0, \sigma_\infty \in (0, \infty)$ such that

$$\sigma_0 |x|^2 \leq x^\top (\sigma\sigma^\top(z)) x \leq \sigma_\infty |x|^2 \quad \text{for all } x \in \mathbb{R}^N \text{ and } z \in U'.$$

- (A3) Transition densities for the external variable:** For all $(\vartheta, T) \in \Theta \times (0, \infty)$ and $t > s \geq 0$, there is a measurable function $p_{s,t}^{(\vartheta, T)}: U' \times U' \rightarrow [0, \infty)$ such that

$$\mathbb{P} \left(Z_t^{(\vartheta, T)} \in B \mid Z_s^{(\vartheta, T)} = z \right) = \int_B p_{s,t}^{(\vartheta, T)}(z, w) dw$$

for all $z \in U'$ and measurable sets $B \subset U'$.

- (A4) Periodic recurrence of the external variable:** For all $(\vartheta, T) \in \Theta \times (0, \infty)$, the grid chain $(Z_{nT}^{(\vartheta, T)})_{n \in \mathbb{N}_0}$ is positive Harris recurrent.

Remark 2.1. 1.) As we know from linear algebra, (A2) also yields that the inverse $(\sigma\sigma^\top(z))^{-1}$ exists for all $z \in U'$, is symmetric and positive definite (and hence possesses a square root $(\sigma\sigma^\top(z))^{-1/2} \in \mathbb{R}^{N \times N}$), and we have

$$\sigma_\infty^{-1} |x|^2 \leq x^\top (\sigma\sigma^\top(z))^{-1} x \leq \sigma_0^{-1} |x|^2 \quad \text{for all } x \in \mathbb{R}^N. \quad (6)$$

2.) Note that $\sigma^\top (\sigma\sigma^\top)^{-1}(z) \in \mathbb{R}^{M \times N}$ is a right inverse of $\sigma(z)$. Thus, the linear mapping $\sigma(z): \mathbb{R}^M \rightarrow \mathbb{R}^N$ is surjective and hence $M \geq N$. In this sense, (A2) is a non-degeneracy condition on the external equation for Z . It is also

“almost sufficient” for (A3) (it is sufficient e.g. in the case that b and σ are smooth with bounded derivatives of any order, compare [9]).

3.) Together with (A3), the recurrence assumption (A4) allows us to make use of certain variants of classical limit theorems (see [18], [19]) which we will need for Lemma 3.4 below. Note that (A4) is weaker than the assertion that the entire process $\Phi^{(\vartheta, T)}$ is positive Harris-recurrent (compare [13]).

Let $(\vartheta, T) \in \Theta \times (0, \infty)$. We define the probability measure

$$\mathbb{P}^{(\vartheta, T)} := \mathcal{L} \left([0, \infty) \ni t \mapsto \Phi_t^{(\vartheta, T)} \mid \mathbb{P} \right)$$

on $\mathcal{B}(C([0, \infty); \mathbf{E}))$ such that for the canonical process $\pi = (\pi_t)_{t \in [0, \infty)}$ on $C([0, \infty); \mathbf{E})$ we have

$$\mathcal{L} \left(\pi \mid \mathbb{P}^{(\vartheta, T)} \right) = \mathcal{L} \left(\Phi^{(\vartheta, T)} \mid \mathbb{P} \right).$$

Observing the process continuously then means working with the filtration given by

$$\mathcal{F}_t := \bigcap_{r \in (t, \infty)} \sigma(\pi_s \mid s \in [0, r]) \subset \mathcal{B}(C([0, \infty); \mathbf{E})) \quad \text{for all } t \in [0, \infty),$$

and gives rise to the sequence of statistical experiments defined by

$$\mathcal{E}_{(X, Y, Z)} := \left(C([0, \infty); \mathbf{E}), \mathcal{F}_n, \left\{ \mathbb{P}^{(\vartheta, T)} \mid_{\mathcal{F}_n} \mid (\vartheta, T) \in \Theta \times (0, \infty) \right\}_{n \in \mathbb{N}} \right).$$

As is proved in Section 3, for all $(\tilde{\vartheta}, \tilde{T}) \in \Theta \times (0, \infty)$ the corresponding log-likelihood ratios are given by

$$\begin{aligned} \log \frac{d\mathbb{P}^{(\tilde{\vartheta}, \tilde{T})} \mid_{\mathcal{F}_t}}{d\mathbb{P}^{(\vartheta, T)} \mid_{\mathcal{F}_t}} &= \int_0^t \left((\sigma \sigma^\top (\pi_s^Z))^{-1/2} (S_{(\tilde{\vartheta}, \tilde{T})} - S_{(\vartheta, T)})(s) \right)^\top dB_s^{(\vartheta, T)} \\ &\quad - \frac{1}{2} \int_0^t (S_{(\tilde{\vartheta}, \tilde{T})} - S_{(\vartheta, T)})^\top(s) (\sigma \sigma^\top (\pi_s^Z))^{-1} \\ &\quad \quad \quad (S_{(\tilde{\vartheta}, \tilde{T})} - S_{(\vartheta, T)})(s) ds, \end{aligned}$$

where $B^{(\vartheta, T)}$ is a Brownian Motion and $\pi^Z = (\pi^{(N+L+1)}, \dots, \pi^{(N+L+N)})$. Examining its structure suggests that in order to find a suitable quadratic expansion for LAN we have to impose appropriate smoothness conditions on the signal with respect to the parameters. The following set of conditions (S1)–(S5) turns out to be sufficient:

(S1) Basic regularity: For each $\vartheta \in \Theta$, we have a 1-periodic function

$$S_\vartheta = \begin{pmatrix} S_\vartheta^{(1)} \\ \vdots \\ S_\vartheta^{(N)} \end{pmatrix} \in C^2([0, \infty); \mathbb{R}^N)$$

such that

$$S(s) \in C^1(\Theta; \mathbb{R}^N) \quad \text{for every } s \in [0, \infty)$$

and

$$\partial_{\vartheta_i} S_{\vartheta}(\cdot) \in \mathbb{L}_{\text{loc}}^2([0, \infty); \mathbb{R}^N) \quad \text{for every } \vartheta \in \Theta \text{ and } i \in \{1, \dots, D\}.$$

(S2) $\mathbb{L}_{\text{loc}}^2$ -differentiability with respect to (ϑ, T) : The mapping

$$\begin{aligned} S: \Theta \times (0, \infty) &\rightarrow \mathbb{L}_{\text{loc}}^2([0, \infty); \mathbb{R}^N), \\ (\vartheta, T) &\mapsto S_{(\vartheta, T)} := S_{\vartheta} \left(\frac{\cdot}{T} \right), \end{aligned}$$

is $\mathbb{L}_{\text{loc}}^2$ -differentiable with the derivative

$$\begin{aligned} \dot{S}: \Theta \times (0, \infty) &\rightarrow \mathbb{L}_{\text{loc}}^2([0, \infty); \mathbb{R}^{N \times (D+1)}), \\ (\vartheta, T) &\mapsto \dot{S}_{(\vartheta, T)} := \begin{pmatrix} \partial_{\vartheta_1} S_{(\vartheta, T)}^{(1)} & \cdots & \partial_{\vartheta_D} S_{(\vartheta, T)}^{(1)} & \partial_T S_{(\vartheta, T)}^{(1)} \\ \vdots & \ddots & \vdots & \vdots \\ \partial_{\vartheta_1} S_{(\vartheta, T)}^{(N)} & \cdots & \partial_{\vartheta_D} S_{(\vartheta, T)}^{(N)} & \partial_T S_{(\vartheta, T)}^{(N)} \end{pmatrix}, \end{aligned}$$

in the sense that for every $t \in (0, \infty)$ and $(\vartheta, T) \in \Theta \times (0, \infty)$ we have¹

$$\lim_{(\tilde{\vartheta}, \tilde{T}) \rightarrow (\vartheta, T)} \int_0^t \left| \frac{S_{(\tilde{\vartheta}, \tilde{T})}(s) - S_{(\vartheta, T)}(s) - \dot{S}_{(\vartheta, T)}(s)((\tilde{\vartheta}, \tilde{T}) - (\vartheta, T))}{|(\tilde{\vartheta}, \tilde{T}) - (\vartheta, T)|} \right|^2 ds = 0.$$

(S3) $\mathbb{L}_{\text{loc}}^2$ -continuity of the (ϑ, T) -derivative: The mapping \dot{S} is continuous in the sense that for all $t \in (0, \infty)$ and $(\vartheta, T) \in \Theta \times (0, \infty)$ we have

$$\lim_{(\tilde{\vartheta}, \tilde{T}) \rightarrow (\vartheta, T)} \int_0^t \left| \dot{S}_{(\tilde{\vartheta}, \tilde{T})}(s) - \dot{S}_{(\vartheta, T)}(s) \right|^2 ds = 0,$$

where the notation $|\cdot|$ is used for the Frobenius norm of a matrix.

(S4) $\mathbb{L}_{\text{loc}}^2$ -Hölder condition with respect to T for the ϑ -derivative: For any fixed $\vartheta \in \Theta$, the mapping

$$(0, \infty) \ni T \mapsto D_{\vartheta} S_{(\vartheta, T)} \in \mathbb{L}_{\text{loc}}^2([0, \infty); \mathbb{R}^{N \times D})$$

with

$$D_{\vartheta} S_{(\vartheta, T)} := \begin{pmatrix} \partial_{\vartheta_1} S_{(\vartheta, T)}^{(1)} & \cdots & \partial_{\vartheta_D} S_{(\vartheta, T)}^{(1)} \\ \vdots & \ddots & \vdots \\ \partial_{\vartheta_1} S_{(\vartheta, T)}^{(N)} & \cdots & \partial_{\vartheta_D} S_{(\vartheta, T)}^{(N)} \end{pmatrix}$$

¹In the context of vector operations, we often write (ϑ, T) instead of the formally correct but awkward $(\vartheta^\top, T)^\top$.

satisfies the following local Hölder condition: For each $T \in (0, \infty)$, there are

$$\alpha \in (0, 2] \quad \text{and} \quad \beta \in [0, 1 + 3\alpha/2)$$

such that for suitable $\varepsilon > 0$ and $t_0 \in [0, \infty)$ we have

$$\int_{t_0}^t \left| D_\vartheta S_{(\vartheta, \tilde{T})}(s) - D_\vartheta S_{(\vartheta, T)}(s) \right|^2 ds \leq Ct^\beta \left| \tilde{T} - T \right|^\alpha$$

for all $t > t_0$, $\tilde{T} \in (T - \varepsilon, T + \varepsilon)$, and for some constant $C \in (0, \infty)$ that does not depend on \tilde{T} or t .

(S5) Linearly independent derivatives: For all $\vartheta \in \Theta$, the functions

$$\partial_{\vartheta_1} S_\vartheta, \dots, \partial_{\vartheta_D} S_\vartheta, S'_\vartheta$$

are linearly independent.

Remark 2.2. 1.) If (S1) holds and $\dot{S}_{(\vartheta, T)}(s)$ is continuous (and thus also locally bounded) with respect to ϑ , T , and s , (S2) and (S3) follow by dominated convergence. Note that in general, (S1) does not require that for example $\partial_{\vartheta_1} S_{(\vartheta, T)}(s)$ is continuous (or even locally bounded) in T or s .

2.) Suppose that (S1) holds and that for every $\vartheta \in \Theta$ and $t \in (0, \infty)$ there are $\delta = \delta(\vartheta) \in (0, 1]$ and $C(\vartheta, t) \leq \text{cst } t^\zeta$ with $\zeta \in [0, \delta/2)$ such that the mapping

$$[0, \infty) \ni s \mapsto D_\vartheta S_\vartheta(s) := \begin{pmatrix} \partial_{\vartheta_1} S_\vartheta^{(1)}(s) & \cdots & \partial_{\vartheta_D} S_\vartheta^{(1)}(s) \\ \vdots & \ddots & \vdots \\ \partial_{\vartheta_1} S_\vartheta^{(N)}(s) & \cdots & \partial_{\vartheta_D} S_\vartheta^{(N)}(s) \end{pmatrix} \in \mathbb{R}^{N \times d}$$

is Hölder- δ -continuous on $[0, t]$ with Hölder-constant $C(\vartheta, t)$. If $T \in (0, \infty)$, we get that for sufficiently small $\varepsilon > 0$ and for all $\tilde{T} \in (T - \varepsilon, T + \varepsilon)$

$$\begin{aligned} \int_0^t \left| D_\vartheta S_{(\vartheta, \tilde{T})}(s) - D_\vartheta S_{(\vartheta, T)}(s) \right|^2 ds &= \int_0^t \left| D_\vartheta S_\vartheta \left(\frac{s}{\tilde{T}} \right) - D_\vartheta S_\vartheta \left(\frac{s}{T} \right) \right|^2 ds \\ &\leq \sup_{T' \in (T - \varepsilon, T + \varepsilon)} C \left(\vartheta, \frac{t}{T'} \right)^2 \int_0^t \left| \frac{s}{\tilde{T}} - \frac{s}{T} \right|^{2\delta} ds \\ &\leq \text{cst} \left(\frac{t}{T - \varepsilon} \right)^{2\zeta} \left(\frac{|\tilde{T} - T|}{(T - \varepsilon)^2} \right)^{2\delta} \int_0^t s^{2\delta} ds \\ &\leq \text{cst } t^{2\zeta + 2\delta + 1} \left| \tilde{T} - T \right|^{2\delta}. \end{aligned}$$

Setting $\alpha := 2\delta$, we can choose

$$\beta := 2(\delta + \zeta) + 1 < 2 \left(\delta + \frac{\delta}{2} \right) + 1 = 1 + 3\alpha/2,$$

and hence the Hölder condition (S4) is fulfilled.

3.) As a consequence of the two preceding observations, all of the hypotheses (S1)–(S4) are fulfilled if the mapping $\Theta \times [0, \infty) \ni (\vartheta, s) \mapsto S_\vartheta(s)$ is in $C_b^2(\Theta \times [0, \infty); \mathbb{R}^N)$ and 1-periodic with respect to s . Existence and boundedness of $\partial_s D_\vartheta S_\vartheta(s)$ ensure that we can choose $\delta = 1$ and $\zeta = 0$ above.

4.) Note that the choice of the matrix norm in (S3) and (S4) is of course arbitrary. We decided to go with the Frobenius norm, because it is commonly used and it is convenient to handle in our calculations.

The main result is the following one. For a detailed explanation and proof, as well as an explicit introduction of the Fisher Information, we refer to Section 3.

Theorem 2.3 (Local Asymptotic Normality for $\mathcal{E}_{(X,Y,Z)}$). *Grant all of the hypotheses (A1)–(A4) and (S1)–(S5) and fix $(\vartheta, T) \in \Theta \times (0, \infty)$. Set*

$$\delta_n := \begin{pmatrix} n^{-1/2} & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & n^{-1/2} & 0 \\ 0 & \dots & 0 & n^{-3/2} \end{pmatrix} \in \mathbb{R}^{(D+1) \times (D+1)} \quad \text{for all } n \in \mathbb{N},$$

and fix any bounded sequence $(h_n)_{n \in \mathbb{N}} \subset \mathbb{R}^{D+1}$. Then $\mathbb{P}^{(\vartheta, T)}$ -almost surely we have

$$\log \frac{d\mathbb{P}^{(\vartheta, T) + \delta_n h_n} |_{\mathcal{F}_n}}{d\mathbb{P}^{(\vartheta, T)} |_{\mathcal{F}_n}} = h_n^\top \mathcal{S}_n^{(\vartheta, T)} - \frac{1}{2} h_n^\top \mathcal{I}_{(\vartheta, T)} h_n + o_{\mathbb{P}^{(\vartheta, T)}}(1), \quad \text{as } n \rightarrow \infty, \quad (7)$$

with Fisher Information $\mathcal{I}_{(\vartheta, T)} = \mathcal{I}_{(\vartheta, T)}(1)$ as introduced in (21) and score

$$\mathcal{S}_n^{(\vartheta, T)} = \delta_n \int_0^n \left((\sigma \sigma^\top)^{-1/2} (\pi_s^Z) \dot{S}_{(\vartheta, T)}(s) \right)^\top dB_s^{(\vartheta, T)} \quad \text{for all } n \in \mathbb{N}$$

such that weak convergence

$$\mathcal{L} \left(\mathcal{S}_n^{(\vartheta, T)} \Big| \mathbb{P}^{(\vartheta, T)} \right) \xrightarrow{n \rightarrow \infty} \mathcal{N} \left(0, \mathcal{I}_{(\vartheta, T)} \right)$$

holds.

Proof of Theorem 2.3. This is an immediate consequence of Theorem 3.7, Proposition 3.2 and (17). In particular, the assumptions (A2) and (S5) can in fact be replaced by the slightly weaker but more technical conditions (A2') and (S5') which are introduced in Section 3 below and are discussed in Remark 3.5. \square

Note that other than the basic existence and uniqueness assumption (A1), the conditions for Theorem 2.3 incorporate only the external variable and the deterministic signal. Before we proceed to the proof section, we would like to collect some comments on relevant examples in which these conditions are fulfilled.

Example 2.4. A simple yet important example for the external variable is the multidimensional Ornstein-Uhlenbeck process with time-dependent mean

reversion level $S_{(\vartheta, T)}$. This process corresponds to (12) with $b(z) = -\beta z$ for all $z \in U' = \mathbb{R}^N$ with some positive definite $\beta \in \mathbb{R}^{N \times N}$ and a constant volatility $\sigma \in \mathbb{R}^{N \times M}$ such that $\sigma\sigma^\top \in \mathbb{R}^{N \times N}$ is positive definite. Assumption (A2) is then trivially fulfilled, and in complete analogy to the case $M = N = 1$ (see [18, Example 2.3]), one can calculate explicitly its transition densities, yielding (A3). These can then be used to apply Theorem 3.2 and Theorem 4.6 (with $f \equiv 1$ and $V(z) = |z|^2$) from [32] in order to check (A4).

Example 2.5. 1.) Let $S_\vartheta(s) = F(\vartheta, \varphi(s))$, where $\varphi \in C^2([0, \infty); \mathbb{R}^K)$ is 1-periodic and

$$F: \Theta \times \mathbb{R}^K \ni (\vartheta, \xi) = (\vartheta_1, \dots, \vartheta_D, \xi_1, \dots, \xi_K) \mapsto F(\vartheta, \xi) = \begin{pmatrix} F_1(\vartheta, \xi) \\ \vdots \\ F_N(\vartheta, \xi) \end{pmatrix} \in \mathbb{R}^N$$

is continuously differentiable with respect to $\vartheta \in \Theta$ and twice continuously differentiable with respect to $\xi \in \mathbb{R}^K$. Clearly, the property (S1) holds, and since $\dot{S}_{(\vartheta, T)}(s)$ is given by

$$\begin{pmatrix} (\partial_{\vartheta_1} F_1)(\vartheta, \varphi(\frac{s}{T})) & \cdots & (\partial_{\vartheta_D} F_1)(\vartheta, \varphi(\frac{s}{T})) & -sT^{-2}(\nabla_\xi F_1)(\vartheta, \varphi(\frac{s}{T}))^\top \varphi'(\frac{s}{T}) \\ \vdots & \ddots & \vdots & \vdots \\ (\partial_{\vartheta_1} F_N)(\vartheta, \varphi(\frac{s}{T})) & \cdots & (\partial_{\vartheta_D} F_N)(\vartheta, \varphi(\frac{s}{T})) & -sT^{-2}(\nabla_\xi F_N)(\vartheta, \varphi(\frac{s}{T}))^\top \varphi'(\frac{s}{T}) \end{pmatrix}$$

which is continuous with respect to ϑ, T , and s , we also have (S2) and (S3). Moreover, we see that the Hölder property from part 2.) of Remark 2.2 is fulfilled if it is fulfilled by the mapping

$$\mathbb{R}^K \ni \xi \mapsto \begin{pmatrix} (\partial_{\vartheta_1} F_1)(\vartheta, \xi) & \cdots & (\partial_{\vartheta_D} F_1)(\vartheta, \xi) \\ \vdots & \ddots & \vdots \\ (\partial_{\vartheta_1} F_N)(\vartheta, \xi) & \cdots & (\partial_{\vartheta_D} F_N)(\vartheta, \xi) \end{pmatrix}.$$

In that case, all of the hypotheses (S1)–(S4) hold.

2.) If the signal has a product structure $S_\vartheta(s) = G(\vartheta)\varphi(s)$ with some $\varphi \in C^2([0, \infty); \mathbb{R}^K)$ 1-periodic and $G \in C^1(\Theta; \mathbb{R}^{N \times K})$, we can treat it as a special case of the preceding example. As for all $s, \tilde{s} \in [0, \infty)$ we have

$$\begin{aligned} |D_\vartheta S_\vartheta(s) - D_\vartheta S_\vartheta(\tilde{s})|^2 &= \sum_{n=1}^N \sum_{d=1}^D \left(\sum_{k=1}^K (\partial_{\vartheta_d} G_{n,k})(\vartheta) (\varphi_k(s) - \varphi_k(\tilde{s})) \right)^2 \\ &\leq \left(\sum_{n=1}^N \sum_{d=1}^D \sum_{k=1}^K (\partial_{\vartheta_d} G_{n,k})^2(\vartheta) \right) |\varphi(s) - \varphi(\tilde{s})|^2 \\ &\leq \left(\sum_{n=1}^N \sum_{d=1}^D \sum_{k=1}^K (\partial_{\vartheta_d} G_{n,k})^2(\vartheta) \right) \|\varphi'\|_\infty^2 |s - \tilde{s}|^2, \end{aligned}$$

no further conditions are needed to ensure the Hölder property from part 2.) of Remark 2.2 to hold with $\delta = 1$ and $\zeta = 0$.

3.) In particular, the example above secures that (S1)–(S4) are fulfilled for signals of the form

$$S_{\vartheta}(s) = \sum_{k=1}^K (\sin(2k\pi s)G_k(\vartheta) + \cos(2k\pi s)H_k(\vartheta)) \quad \text{for all } s \in [0, \infty) \quad (8)$$

with $K \in \mathbb{N}$ and $G_k, H_k \in C^1(\Theta; \mathbb{R}^N)$ for all $k \in \{1, \dots, K\}$.

4.) Taking $K = D$, $N = 1$ and $G_k(\vartheta) = \vartheta_k$, $H_k(\vartheta) = 0$ for all $\vartheta \in \Theta$ and $k \in \{1, \dots, K\}$, the signal from (8) clearly also satisfies (S5), as long as $0 \notin \Theta$.

3. Proofs and supplementary results

3.1. Observing (X, Y, Z) , X , or Z

We start this section with a fundamental observation: If the starting point is known, observing only the adjustable variable X is actually no restriction, since we can successively reconstruct the remaining variables Y and Z . Let us explain this step for step in the following remark.

Remark 3.1. Assume that the starting point $(X_0, Y_0, Z_0) \in \mathbf{E}$ is known. Fix a finite time horizon $t_0 \in (0, \infty)$ and assume that the trajectory $(X_t)_{t \in [0, t_0]}$ has been observed and is thus also known. Then the function $(t, y) \mapsto g(X_t, y)$ is completely known, and given the structure of the internal equation in (2), the trajectory $(Y_t)_{t \in [0, t_0]}$ is now given as the solution to the ordinary differential equation

$$dY_t = g(X_t, Y_t)dt \quad \text{for all } t \in [0, t_0].$$

Now we know both $(X_t)_{t \in [0, t_0]}$ and $(Y_t)_{t \in [0, t_0]}$, and by rearranging the first line of (2), this information allows us to calculate

$$Z_t = Z_0 + X_t - X_0 - \int_0^t f(X_s, Y_s)ds \quad \text{for all } t \in [0, t_0].$$

All in all, we have reconstructed every component of $(X_t, Y_t, Z_t)_{t \in [0, t_0]}$ just from $(X_t)_{t \in [0, t_0]}$ and the starting point (X_0, Y_0, Z_0) .

Remark 3.1 is the legitimation for us to work with the idealized assumption that we can in fact observe the entire process (X, Y, Z) even in situations where realistically one could only observe the adjustable variable X . This enables us to lift known results about estimators (compare Remarks 3.10 and 3.11). Next, we will formally describe the corresponding statistical experiment.

In order to make Proposition 3.2 more apprehensible, we will do this very carefully and with much attention to measure-theoretic subtleties. A look at (5) reveals that the drift coefficient depends on the parameter $(\vartheta, T) \in \Theta \times (0, \infty)$,

while the volatility does not. Hence, we can use [15, Theorem 6.10]² in order to determine the log-likelihood ratios. Let $(t, x, y, z) \in [0, \infty) \times \mathbf{E}$. Comparing the drift coefficients of (5) with different parameters $(\tilde{\vartheta}, \tilde{T}), (\vartheta, T) \in \Theta \times (0, \infty)$, we see that

$$(B_{(\tilde{\vartheta}, \tilde{T})} - B_{(\vartheta, T)})(t, x, y, z) = \Sigma \Sigma^\top(x, y, z) \Gamma(t, x, y, z),$$

where

$$\Gamma(t, x, y, z) := \begin{pmatrix} 0 \\ (\sigma \sigma^\top)^{-1}(z) (S_{(\tilde{\vartheta}, \tilde{T})} - S_{(\vartheta, T)})(t) \end{pmatrix} \in \mathbb{R}^{N+L+N}.$$

Thanks to (A2) and (6),

$$\int_0^t (\Gamma^\top \Sigma \Sigma^\top \Gamma)(s, \pi_s) ds \leq \sigma_0^{-1} \int_0^t |S_{(\tilde{\vartheta}, \tilde{T})}(s) - S_{(\vartheta, T)}(s)|^2 ds < \infty, \quad (9)$$

because the signals are continuous. Thence, both conditions (+) and (++) of [15, Theorem 6.10] are fulfilled. Writing $m^{\Phi, (\vartheta, T)}$ for the local martingale part of π under $\mathbb{P}^{(\vartheta, T)}$, we can conclude that

$$\log \frac{d\mathbb{P}^{(\tilde{\vartheta}, \tilde{T})}|_{\mathcal{F}_t}}{d\mathbb{P}^{(\vartheta, T)}|_{\mathcal{F}_t}} = \int_0^t \Gamma(s, \pi_s)^\top dm_s^{\Phi, (\vartheta, T)} - \frac{1}{2} \int_0^t (\Gamma^\top \Sigma \Sigma^\top \Gamma)(s, \pi_s) ds.$$

Setting $\pi^Z := (\pi^{(N+L+1)}, \dots, \pi^{(N+L+N)})^\top$ and writing $m^{Z, (\vartheta, T)}$ for its local martingale part under $\mathbb{P}^{(\vartheta, T)}$, the expression for the log-likelihood ratio can be rewritten as

$$\begin{aligned} & \int_0^t \left((\sigma \sigma^\top(\pi_s^Z))^{-1} (S_{(\tilde{\vartheta}, \tilde{T})} - S_{(\vartheta, T)})(s) \right)^\top dm_s^{Z, (\vartheta, T)} \\ & - \frac{1}{2} \int_0^t (S_{(\tilde{\vartheta}, \tilde{T})} - S_{(\vartheta, T)})^\top(s) (\sigma \sigma^\top(\pi_s^Z))^{-1} (S_{(\tilde{\vartheta}, \tilde{T})} - S_{(\vartheta, T)})(s) ds. \end{aligned}$$

In order to eliminate the rather unintuitive integral with respect to $m^{Z, (\vartheta, T)}$, we introduce the local $(\mathbb{P}^{(\vartheta, T)}, (\mathcal{F}_t)_{t \in [0, \infty)})$ -martingale $B^{(\vartheta, T)} := (B_t^{(\vartheta, T)})_{t \in [0, \infty)}$ given by

$$B_t^{(\vartheta, T)} = \int_0^t (\sigma \sigma^\top)^{-1/2}(\pi_s^Z) dm_s^{Z, (\vartheta, T)} \quad \text{for all } t \in [0, \infty). \quad (10)$$

²Note that we do not assume – as in this theorem – that B and Σ are defined on the entire euclidean space and are globally Lipschitz continuous. By our assumptions, $\mathbf{E} = U \times U'$ is σ -compact and hence we can find a sequence $(K_n)_{n \in \mathbb{N}}$ of compact sets increasing to \mathbf{E} . Using Kirszbraun's Theorem ([28, Hauptsatz I]), the restriction to each K_n of $B(t, \cdot)$ and Σ can be extended to globally Lipschitz continuous functions on \mathbb{R}^{N+L+N} (which also satisfy a linear growth condition). Hence, the proof of [15, Theorem 6.10] needs only a slight adjustment to work in our case: Using the notation from there, the stopping time ϱ_n has to be replaced by $\varrho_n \wedge \inf\{t > 0 \mid \eta_t \notin K_n\}$, and in equation (II⁽ⁿ⁾) and thereafter the coefficients b , σ and c have to be altered in analogy to γ . The rest of the proof then needs no further changes.

Its quadratic variation process is

$$\begin{aligned} \left\langle \int_0^\cdot (\sigma\sigma^\top)^{-1/2}(\pi_s^Z) dm_s^{Z,(\vartheta,T)} \right\rangle_t &= \int_0^t (\sigma\sigma^\top)^{-1}(\pi_s^Z) d \left(\int_0^s \sigma\sigma^\top(\pi_r^Z) dr \right) \\ &= t \cdot \mathbf{1}_{N \times N} \end{aligned}$$

for all $t \in [0, \infty)$, so Lévy’s Characterization Theorem [25, Theorem II.6.1] yields that $B^{(\vartheta,T)}$ is an N -dimensional $(\mathbb{P}^{(\vartheta,T)}, (\mathcal{F}_t)_{t \in [0, \infty)})$ -Brownian Motion. Incorporating this process, we can write

$$\begin{aligned} \log \frac{d\mathbb{P}^{(\tilde{\vartheta}, \tilde{T})}|_{\mathcal{F}_t}}{d\mathbb{P}^{(\vartheta,T)}|_{\mathcal{F}_t}} &= \int_0^t \left((\sigma\sigma^\top(\pi_s^Z))^{-1/2} (S_{(\tilde{\vartheta}, \tilde{T})} - S_{(\vartheta,T)})(s) \right)^\top dB_s^{(\vartheta,T)} \\ &\quad - \frac{1}{2} \int_0^t (S_{(\tilde{\vartheta}, \tilde{T})} - S_{(\vartheta,T)})^\top(s) (\sigma\sigma^\top(\pi_s^Z))^{-1} \\ &\quad \quad \quad (S_{(\tilde{\vartheta}, \tilde{T})} - S_{(\vartheta,T)})(s) ds. \end{aligned} \tag{11}$$

We note immediately that the only component of π that is featured explicitly in this expression is the π^Z -component. It seems plausible that we should get the same expression for the log-likelihood ratio in an experiment that does not even know that any variables other than Z exist. Let us make this formally rigorous.

Let $\eta = (\eta_t)_{t \in [0, \infty)}$ be the canonical process on $C([0, \infty); U')$, and write

$$\mathbb{Q}^{(\vartheta,T)} := \mathcal{L} \left([0, \infty) \ni t \mapsto Z_t^{(\vartheta,T)} \mid \mathbb{P} \right)$$

for the law on $\mathcal{B}(C([0, \infty); U'))$ of the unique strong solution $Z^{(\vartheta,T)}$ on (Ω, \mathcal{F}) under \mathbb{P} of

$$dZ_t = [S_{(\vartheta,T)}(t) + b(Z_t)]dt + \sigma(Z_t)dW_t, \tag{12}$$

when issued from $Z_0 \in \mathbb{R}^N$ with the parameter $(\vartheta, T) \in \Theta \times (0, \infty)$. For any $t \in [0, \infty)$, let

$$\mathcal{G}_t := \bigcap_{r \in (t, \infty)} \sigma(\eta_s \mid s \in [0, r]) \subset \mathcal{B}(C([0, \infty); U'))$$

and consider the sequence of experiments given by

$$\mathcal{E}_Z := \left(C([0, \infty); U'), \mathcal{G}_n, \left\{ \mathbb{Q}^{(\vartheta,T)}|_{\mathcal{G}_n} \mid (\vartheta, T) \in \Theta \times (0, \infty) \right\} \right)_{n \in \mathbb{N}}. \tag{13}$$

Using the same arguments as above and writing $\tilde{m}^{Z,(\vartheta,T)}$ for the local martingale part of η under $\mathbb{Q}^{(\vartheta,T)}$, we can again use [15, Theorem 6.10] and conclude

$$\begin{aligned} \log \frac{d\mathbb{Q}^{(\tilde{\vartheta}, \tilde{T})}|_{\mathcal{G}_t}}{d\mathbb{Q}^{(\vartheta,T)}|_{\mathcal{G}_t}} &= \int_0^t \left((\sigma\sigma^\top)^{-1/2}(\eta_s) (S_{(\tilde{\vartheta}, \tilde{T})} - S_{(\vartheta,T)})(s) \right)^\top d\tilde{B}_s^{(\vartheta,T)} \\ &\quad - \frac{1}{2} \int_0^t (S_{(\tilde{\vartheta}, \tilde{T})} - S_{(\vartheta,T)})^\top(s) (\sigma\sigma^\top(\eta_s))^{-1} \\ &\quad \quad \quad (S_{(\tilde{\vartheta}, \tilde{T})} - S_{(\vartheta,T)})(s) ds, \end{aligned} \tag{14}$$

where the process $\tilde{B}^{(\vartheta, T)} := (\tilde{B}_t^{(\vartheta, T)})_{t \in [0, \infty)}$ given by

$$\tilde{B}_t^{(\vartheta, T)} = \int_0^t (\sigma \sigma^\top)^{-1/2} (\eta_s) d\tilde{m}_s^{Z, (\vartheta, T)} \quad \text{for all } t \in [0, \infty) \quad (15)$$

is again an N -dimensional $(\mathbb{Q}^{(\vartheta, T)}, (\mathcal{G}_t)_{t \in [0, \infty)})$ -Brownian Motion.

We now have calculated the log-likelihood ratios for both $\mathcal{E}_{(X, Y, Z)}$ and \mathcal{E}_Z . Comparing them leads to the following result.

Proposition 3.2. *Grant assumptions (A1) and (A2). The sequences $\mathcal{E}_{(X, Y, Z)}$ and \mathcal{E}_Z corresponding to continuous observation of (X, Y, Z) or Z respectively, with the same deterministic starting point $(X_0, Y_0, Z_0) \in \mathbf{E}$, are statistically equivalent in the sense that*

$$\mathcal{L} \left(\left(\log \frac{d\mathbb{Q}^{(\tilde{\vartheta}, \tilde{T})} |_{\mathcal{G}_t}}{d\mathbb{Q}^{(\vartheta, T)} |_{\mathcal{G}_t}} \right)_{t \in [0, \infty)} \middle| \mathbb{Q}^{(\vartheta, T)} \right) = \mathcal{L} \left(\left(\log \frac{d\mathbb{P}^{(\tilde{\vartheta}, \tilde{T})} |_{\mathcal{F}_t}}{d\mathbb{P}^{(\vartheta, T)} |_{\mathcal{F}_t}} \right)_{t \in [0, \infty)} \middle| \mathbb{P}^{(\vartheta, T)} \right) \quad (16)$$

for all $(\vartheta, T), (\tilde{\vartheta}, \tilde{T}) \in \Theta \times (0, \infty)$. In particular, we have LAN for $\mathcal{E}_{(X, Y, Z)}$ if and only if we have it for \mathcal{E}_Z with the same local scale, the same Fisher Information and an identically distributed Score.

Proof. Due to the definition of $\mathbb{P}^{(\vartheta, T)}$ and $\mathbb{Q}^{(\vartheta, T)}$, we have

$$\mathcal{L}(\eta | \mathbb{Q}^{(\vartheta, T)}) = \mathcal{L}(\pi^Z | \mathbb{P}^{(\vartheta, T)}), \quad (17)$$

and in view of (10), (11), (14), and (15), this implies (16) from which the second statement of this proposition follows immediately. \square

In view of Theorem 2.3, Proposition 3.2 is the justification for us to restrict ourselves to studying the simpler process Z instead of the more complex (X, Y, Z) in the following section.

3.2. Local Asymptotic Normality for Z

This section centres around the sequence of statistical experiments defined by \mathcal{E}_Z in (13) which corresponds to continuous observation over growing time intervals of the N -dimensional diffusion Z following the parameter-dependent SDE (12). As mentioned in Section 1, taking $M = N = 1$, $b \equiv 0$, and $\sigma \equiv 1$ leads to the classical “signal in white noise” model. For this special case, Ibragimov and Khasminskii proved LAN with rate $n^{-3/2}$ for a smooth signal with known ϑ and unknown T , and discussed asymptotic efficiency for certain estimators (see [24, Sections II.7 and III.5]). In [8], Golubev extended their approach with \mathbb{L}^2 -methods in order to estimate T at the same rate for unknown shape (see Remark 3.11 below) which in turn was the basis for Castillo, Lévy-Leduc and Matias for non-parametric estimation of the shape under unknown T (see [1]). For our more general diffusion (12), we will stay within the confines of parametric estimation.

The main result of this section is LAN for the sequence of experiments \mathcal{E}_Z with unknown ϑ and unknown T (Theorem 3.7). For $M = N = 1$, Höpfner and Kutoyants had already solved this problem both for known T with unknown ϑ (see [17]) and for known ϑ with unknown T (see [19]). A result on LAN jointly in ϑ and T was presented more recently in [12], but still only in dimension one. Theorem 3.7 extends all of these results (compare Remark 3.9) and allows for application to simultaneous estimation of the shape and the periodicity in any dimension.

In the context of this subsection, we replace the assumption (A1) with the following weaker analogue.

(A1') Unique solvability: For all $(\vartheta, T) \in \Theta \times (0, \infty)$ and all deterministic starting points $Z_0 \in U'$, the SDE (12) has a unique strong solution $Z^{(\vartheta, T)}: [0, \infty) \rightarrow U'$ under \mathbb{P} .

We also work with the following slight relaxation of (A2).

(A2') Uniform ellipticity: The mapping $\sigma\sigma^\top: U' \rightarrow \mathbb{R}^{N \times N}$ is uniformly elliptic, i.e. there is some $\sigma_0 \in (0, \infty)$ such that

$$x^\top (\sigma\sigma^\top(z)) x \geq \sigma_0 |x|^2 \quad \text{for all } x \in \mathbb{R}^N \text{ and } z \in U'.$$

Note that so far, the only use of (A2) occurred in (9), and there (A2') would also suffice. Let us also give an equivalent reformulation of (A4) which incorporates the notation we introduced in the previous section.

(A4) Periodic recurrence of (12): For all $(\vartheta, T) \in \Theta \times (0, \infty)$, the grid chain $(\eta_{kT})_{k \in \mathbb{N}_0}$ under $\mathbb{Q}^{(\vartheta, T)}$ is positive Harris recurrent with invariant probability measure $\mu^{(\vartheta, T)}$.

Periodicity of the signal is the reason why (A4) even makes sense at all: Since $S_{(\vartheta, T)}$ and therefore the entire drift term of (12) is T -periodic, the grid chain is a U' -valued time-homogeneous discrete-time Markov process. Another important process that is embedded in η in a similar way is the $C([0, T]; U')$ -valued time-homogeneous *path segment chain* $\eta^{\text{ps}} := (\eta_k^{\text{ps}})_{k \in \mathbb{N}_0}$ defined by taking an arbitrary $\eta_0^{\text{ps}} \in C([0, T]; U')$ with $\eta_0^{\text{ps}}(T) = Z_0$ and then setting

$$\eta_k^{\text{ps}} := ([0, T] \ni t \mapsto \eta_{(k-1)T+t}) \quad \text{for all } k \in \mathbb{N}.$$

As we know from [18, Theorem 2.1 (a)]³, the path segment chain η^{ps} inherits positive Harris recurrence under $\mathbb{Q}^{(\vartheta, T)}$ from the grid chain, and its invariant distribution $m^{(\vartheta, T)}$ is the unique measure on $\mathcal{B}(C([0, T]; U'))$ such that for all $l \in \mathbb{N}$, $0 = t_0 < t_1 < \dots < t_l = T$, and $B_0, \dots, B_l \in \mathcal{B}(U')$ we have

$$\begin{aligned} & m^{(\vartheta, T)}(\eta_{t_i} \in B_i \text{ for all } i \in \{0, \dots, l\}) \\ &= \int_{B_0} \mu^{(\vartheta, T)}(dx_0) \int_{B_1} Q_{t_0, t_1}^{(\vartheta, T)}(x_0, dx_1) \dots \int_{B_l} Q_{t_{l-1}, t_l}^{(\vartheta, T)}(x_{l-1}, dx_l), \end{aligned} \tag{18}$$

³Note that even though this theorem is only explicitly stated for \mathbb{R} -valued processes, the authors remark at the beginning of the section that it remains valid for any polish state space, in particular for the cloed set $U' \subset \mathbb{R}^N$.

where $(Q_{s,t}^{(\vartheta,T)})_{t>s\geq 0}$ is the transition semi-group of η under $\mathbb{Q}^{(\vartheta,T)}$.

We will make use of the following strong law of large numbers for the path segment chain which we cite from [18, Theorem 2.1 (b)].

Proposition 3.3. *Let (A1'), (A3) and (A4) hold and fix some $(\vartheta, T) \in \Theta \times (0, \infty)$. Assume that $(A_t)_{t \in [0, \infty)}$ is a $(\mathbb{Q}^{(\vartheta,T)}, (\mathcal{G}_t)_{t \in [0, \infty)})$ -increasing process. If there is a non-negative function $F \in \mathbb{L}^1(m^{(\vartheta,T)})$ such that*

$$A_{kT} = \sum_{j=1}^k F(\eta_j^{\text{ps}}) \quad \mathbb{Q}^{(\vartheta,T)}\text{-almost surely for all } k \in \mathbb{N},$$

then

$$\frac{1}{t} A_t \xrightarrow{t \rightarrow \infty} \frac{1}{T} \int_{C([0,T]; U')} F(\varphi) m^{(\vartheta,T)}(d\varphi) \quad \mathbb{Q}^{(\vartheta,T)}\text{-almost surely.}$$

Proof. See Section 2 of [18]. □

Proposition 3.3 is the key to the following Lemma 3.4 which is a slightly modified multi-dimensional version of Lemmas 2.1 and 2.2 from [19].

Lemma 3.4. *Grant assumptions (A1'), (A3) and (A4). Further assume that the measurable mapping $G: U' \rightarrow \mathbb{R}^{N \times N}$ has values only in the set of symmetric matrices and is uniformly elliptic. We define the mapping*

$$\begin{aligned} \mathbb{B}_G^{(\vartheta,T)}: (\mathbb{L}^2([0, 1]; \mathbb{R}^N))^2 &\rightarrow \mathbb{R}, \\ (u, v) &\mapsto \int_0^1 u(s)^\top \left(\mu^{(\vartheta,T)} Q_{0,sT}^{(\vartheta,T)}(G^{-1}) \right) v(s) ds, \end{aligned} \tag{19}$$

where

$$\mu^{(\vartheta,T)} Q_{0,sT}^{(\vartheta,T)}(G^{-1}) = \int_{U'} \mu^{(\vartheta,T)}(dz) \int_{U'} Q_{0,sT}^{(\vartheta,T)}(z, d\tilde{z}) G^{-1}(\tilde{z}) \in \mathbb{R}^{N \times N}$$

is understood as a matrix-valued integral. Then the following statements are true.

- (i) $\mathbb{B}_G^{(\vartheta,T)}$ is a non-negative definite and symmetric bilinear form.
- (ii) If we consider $u, v \in \mathbb{L}^2([0, 1]; \mathbb{R}^N)$ as 1-periodic functions on $[0, \infty)$, then for any $k \in \mathbb{N}_0$ we have

$$\frac{k+1}{t^{k+1}} \int_0^t s^k u(s/T)^\top G^{-1}(\eta_s) v(s/T) ds \xrightarrow{t \rightarrow \infty} \mathbb{B}_G^{(\vartheta,T)}[u, v] \tag{20}$$

$\mathbb{Q}^{(\vartheta,T)}$ -almost surely.

Proof. For the sake of simplicity and as (ϑ, T) is fixed anyway, we drop all corresponding superscripts. First, we check that \mathbb{B}_G is indeed a well-defined mapping with values in \mathbb{R} . Let the lower bound for the eigenvalues of $G(\cdot)$ be denoted by

$G_0 \in (0, \infty)$. Recall that $G^{-1}(\cdot)$ always exists, is positive definite, and G_0^{-1} is an upper bound for its eigenvalues. Then by linearity and contractivity of the operator $\mu Q_{0,sT}$, we can estimate

$$0 \leq \mathbb{B}_G[u, u] = \int_0^1 \mu Q_{0,sT} (u(s)^\top G^{-1}(\cdot) u(s)) ds \leq G_0^{-1} \int_0^1 |u(s)|^2 ds < \infty.$$

Thanks to the symmetry of G^{-1} , we can polarize the integrand and thus the whole expression, which allows us to use the above in order to conclude that

$$|\mathbb{B}_G[u, v]| = \frac{1}{2} |\mathbb{B}_G[u, u] + \mathbb{B}_G[v, v] - \mathbb{B}_G[u + v, u + v]| < \infty,$$

and hence \mathbb{B}_G is well-defined. It is then trivial to see that it is a non-negative definite and symmetric bilinear form, and the proof for (i) is complete.

We note that the left hand side of (20) is bilinear in u and v as well. Thanks to this and (i), the proof of the second statement of the lemma can be reduced to the case $u = v$, since the general case then follows by polarization.

Let us fix $u \in \mathbb{L}^2([0, 1]; \mathbb{R}^N)$ and define the process $A := (A_t)_{t \in [0, \infty)}$ with

$$A_t := \int_0^t u(s/T)^\top G^{-1}(\eta_s) u(s/T) ds \quad \text{for all } t \in [0, \infty).$$

Since $G^{-1}(\cdot)$ is positive definite, the integrand is non-negative, and therefore A is an increasing process whose trajectories are obviously continuous. Note that the expression on the left hand side of (20) can be rewritten as

$$\frac{k+1}{t^{k+1}} \int_0^t s^k dA_s.$$

For $k = 0$, this is simply $\frac{1}{t} A_t$, which we will handle with the help of Proposition 3.3. The general statement then follows from this special case by elementary calculus (compare Lemma 3.17 of [14]).

In order to establish the functional relation between A and η that is needed in Proposition 3.3, we define the function

$$F: C([0, T]; U') \rightarrow [0, \infty), \quad \varphi \mapsto \int_0^T u(s/T)^\top G^{-1}(\varphi(s)) u(s/T) ds,$$

which is bounded by $TG_0^{-1} \|u\|_{\mathbb{L}^2([0, 1])}$, and thus it is integrable with respect to the probability measure m . Since u is 1-periodic, we see that

$$\begin{aligned} \sum_{j=1}^k F(\eta_j^{\text{ps}}) &= \sum_{j=1}^k \int_0^T u(s/T)^\top G^{-1}(\eta_{(j-1)T+s}) u(s/T) ds \\ &= \int_0^{kT} u(s/T)^\top G^{-1}(\eta_s) u(s/T) ds \\ &= A_{kT} \end{aligned}$$

for all $k \in \mathbb{N}$, and consequently Proposition 3.3 allows to deduce \mathbb{Q} -almost sure convergence

$$\begin{aligned} \lim_{t \rightarrow \infty} \frac{1}{t} A_t &= \frac{1}{T} \int_{C([0,T];U')} \int_0^T u(s/T)^\top G^{-1}(\varphi(s)) u(s/T) ds m(d\varphi) \\ &= \frac{1}{T} \int_0^T u(s/T)^\top \left(\int_{C([0,T];U')} G^{-1}(\varphi(s)) m(d\varphi) \right) u(s/T) ds \\ &= \frac{1}{T} \int_0^T u(s/T)^\top \left(\int_{U'} G^{-1}(x) \mu_{Q_{0,s}}(dx) \right) u(s/T) ds \\ &= \mathbb{B}_G[u, u], \end{aligned}$$

where the use of Fubini's Theorem in the second step is justified by the non-negativity of the integrand, and the third step makes use of (18). This completes the proof. \square

Using the notation from (19), for each $(\vartheta, T) \in \Theta \times (0, \infty)$ and $t \in [0, \infty)$ we define the symmetric $(D+1) \times (D+1)$ -dimensional block matrix

$$\mathcal{I}_{(\vartheta, T)}(t) := \begin{pmatrix} t \left(\mathbb{B}_{\sigma\sigma^\top}^{(\vartheta, T)} [\partial_{\vartheta_i} S_\vartheta, \partial_{\vartheta_j} S_\vartheta] \right)_{i,j=1}^D & -\frac{t^2}{2T^2} \left(\mathbb{B}_{\sigma\sigma^\top}^{(\vartheta, T)} [\partial_{\vartheta_i} S_\vartheta, S'_\vartheta] \right)_{i=1}^D \\ \dots & \frac{t^3}{3T^4} \mathbb{B}_{\sigma\sigma^\top}^{(\vartheta, T)} [S'_\vartheta, S'_\vartheta] \end{pmatrix}. \quad (21)$$

Its derivative with respect to t is given by

$$\mathcal{I}'_{(\vartheta, T)}(t) = \begin{pmatrix} \left(\mathbb{B}_{\sigma\sigma^\top}^{(\vartheta, T)} [\partial_{\vartheta_i} S_\vartheta, \partial_{\vartheta_j} S_\vartheta] \right)_{i,j=1}^D & -tT^{-2} \left(\mathbb{B}_{\sigma\sigma^\top}^{(\vartheta, T)} [\partial_{\vartheta_i} S_\vartheta, S'_\vartheta] \right)_{i=1}^D \\ \dots & t^2 T^{-4} \mathbb{B}_{\sigma\sigma^\top}^{(\vartheta, T)} [S'_\vartheta, S'_\vartheta] \end{pmatrix}.$$

We make the following assumption.

(S5') **Regularity of the signal with respect to $\mathbb{B}_{\sigma\sigma^\top}^{(\vartheta, T)}$:** For all $(\vartheta, T) \in \Theta \times (0, \infty)$ and $t \in (0, \infty)$, we have

- (i) $\mathcal{I}_{(\vartheta, T)}(t)$ is invertible, (ii) $\mathcal{I}'_{(\vartheta, T)}(t)$ is invertible.

While part (ii) of (S5') is merely needed for technical reasons (as will become clear in the proof of Theorem 3.7 below), part (i) is of more general importance, since $\mathcal{I}_{(\vartheta, T)}(1)$ will turn out to be the Fisher Information. We will discuss these conditions in detail in the following remark.

Remark 3.5. 1.) Note that $\mathcal{I}'_{(\vartheta, T)}(t)$ is the Gramian matrix of the vectors $\partial_{\vartheta_1} S_\vartheta, \dots, \partial_{\vartheta_D} S_\vartheta, -tT^{-2} S'_\vartheta$ with respect to the non-negative definite symmetric bilinear form $\mathbb{B}_{\sigma\sigma^\top}^{(\vartheta, T)}$. Hence, it is non-negative definite. The same is true for $\mathcal{I}_{(\vartheta, T)}(t)$, since it is ‘‘almost a Gramian matrix’’. Indeed, setting

$$u_1 := t^{1/2} \partial_{\vartheta_1} S_\vartheta, \dots, u_D := t^{1/2} \partial_{\vartheta_D} S_\vartheta, u_{D+1} := -\frac{t^{3/2}}{2T^2} S'_\vartheta,$$

we can write

$$\mathcal{I}_{(\vartheta,T)}(t) = \begin{pmatrix} \mathbb{B}_{\sigma\sigma^\top}^{(\vartheta,T)}[u_1, u_1] & \cdots & \cdots & \mathbb{B}_{\sigma\sigma^\top}^{(\vartheta,T)}[u_1, u_{D+1}] \\ \vdots & \ddots & & \vdots \\ \vdots & & \mathbb{B}_{\sigma\sigma^\top}^{(\vartheta,T)}[u_D, u_D] & \mathbb{B}_{\sigma\sigma^\top}^{(\vartheta,T)}[u_D, u_{D+1}] \\ \mathbb{B}_{\sigma\sigma^\top}^{(\vartheta,T)}[u_{D+1}, u_1] & \cdots & \mathbb{B}_{\sigma\sigma^\top}^{(\vartheta,T)}[u_{D+1}, u_D] & \frac{4}{3}\mathbb{B}_{\sigma\sigma^\top}^{(\vartheta,T)}[u_{D+1}, u_{D+1}] \end{pmatrix},$$

and we see that for all $x \in \mathbb{R}^{D+1}$

$$\begin{aligned} x^\top \mathcal{I}_{(\vartheta,T)}(t)x &= \sum_{i,j=1}^{D+1} x_i \mathbb{B}_{\sigma\sigma^\top}^{(\vartheta,T)}[u_i, u_j] x_j + \frac{1}{3} x_{D+1}^2 \mathbb{B}_{\sigma\sigma^\top}^{(\vartheta,T)}[u_{D+1}, u_{D+1}] \\ &\geq \mathbb{B}_{\sigma\sigma^\top}^{(\vartheta,T)} \left[\sum_{i=1}^{D+1} x_i u_i, \sum_{j=1}^{D+1} x_j u_j \right] \end{aligned}$$

which is non-negative.

2.) In particular, 1.) implies that $\mathcal{I}_{(\vartheta,T)}(t)$ and $\mathcal{I}'_{(\vartheta,T)}(t)$ are invertible if and only if they are positive definite.

3.) If $\mathbb{B}_{\sigma\sigma^\top}^{(\vartheta,T)}$ is positive definite (and hence an inner product), the same reasoning as in 1.) yields that linear independence of $\partial_{\vartheta_1} S_\vartheta, \dots, \partial_{\vartheta_D} S_\vartheta, S'_\vartheta$ is equivalent to invertibility of $\mathcal{I}'_{(\vartheta,T)}(t)$, and sufficient for invertibility of $\mathcal{I}_{(\vartheta,T)}(t)$.

4.) If (A2) holds, for all $u \in \mathbb{L}^2([0, 1]; \mathbb{R}^N)$ we can use (6) and estimate

$$\mathbb{B}_{\sigma\sigma^\top}^{(\vartheta,T)}[u, u] = \int_0^1 \mu^{(\vartheta,T)} Q_{0,sT}^{(\vartheta,T)} \left(u(s)^\top (\sigma\sigma^\top)^{-1} (\cdot) u(s) \right) ds \geq \sigma_\infty^{-1} \int_0^1 |u(s)|^2 ds,$$

i.e. $\mathbb{B}_{\sigma\sigma^\top}^{(\vartheta,T)}$ is positive definite (in fact even coercive). Thus, (A2) and (S5) together imply (S5').

5.) A very simple and seemingly natural sufficient condition for (S5') is orthogonality of the functions $\partial_{\vartheta_1} S_\vartheta, \dots, \partial_{\vartheta_D} S_\vartheta, S'_\vartheta$ with respect to $\mathbb{B}_{\sigma\sigma^\top}^{(\vartheta,T)}$ (without assuming this bilinear form to be positive definite). This is equivalent to both $\mathcal{I}_{(\vartheta,T)}(t)$ and $\mathcal{I}'_{(\vartheta,T)}(t)$ being diagonal matrices with non-vanishing diagonal entries and as such they are invertible. However, this is not a very likely scenario, since S_ϑ has D degrees of freedom, determines the D functions $\partial_{\vartheta_1} S_\vartheta, \dots, \partial_{\vartheta_D} S_\vartheta$, and then S'_ϑ – while adding no further degree of freedom – would have to be orthogonal to these as well.

Example 3.6. 1.) If the signal is of the form

$$S_\vartheta = \sum_{i=1}^D \vartheta_i \varphi_i,$$

where $\varphi_1, \dots, \varphi_D \in \mathbb{L}^2([0, \infty); \mathbb{R}^N)$ are 1-periodic and orthonormal with respect

to $\mathbb{B}_{\sigma\sigma^\top}^{(\vartheta,T)}$, we have

$$\mathcal{I}_{(\vartheta,T)}(t) = \begin{pmatrix} t \cdot 1_{D \times D} & -\frac{t^2}{2T^2} \left(\sum_{j=1}^D \vartheta_j \mathbb{B}_{\sigma\sigma^\top}^{(\vartheta,T)}[\varphi_i, \varphi'_j] \right)_{i=1, \dots, D} \\ \dots & \frac{t^3}{3T^4} \sum_{i,j=1}^D \vartheta_i \vartheta_j \mathbb{B}_{\sigma\sigma^\top}^{(\vartheta,T)}[\varphi'_i, \varphi'_j] \end{pmatrix}$$

which is invertible for all $t \in (0, \infty)$ whenever

$$\frac{4}{3} \sum_{i,j=1}^D \vartheta_i \vartheta_j \mathbb{B}_{\sigma\sigma^\top}^{(\vartheta,T)}[\varphi'_i, \varphi'_j] \neq \sum_{i=1}^D \left(\sum_{j=1}^D \vartheta_j \mathbb{B}_{\sigma\sigma^\top}^{(\vartheta,T)}[\varphi_i, \varphi'_j] \right)^2. \quad (22)$$

Similarly,

$$\mathcal{I}'_{(\vartheta,T)}(t) = \begin{pmatrix} 1_{D \times D} & -tT^{-2} \left(\sum_{j=1}^D \vartheta_j \mathbb{B}_{\sigma\sigma^\top}^{(\vartheta,T)}[\varphi_i, \varphi'_j] \right)_{i=1, \dots, D} \\ \dots & t^2 T^{-4} \sum_{i,j=1}^D \vartheta_i \vartheta_j \mathbb{B}_{\sigma\sigma^\top}^{(\vartheta,T)}[\varphi'_i, \varphi'_j] \end{pmatrix}$$

is invertible for all $t \in (0, \infty)$ whenever

$$\sum_{i,j=1}^D \vartheta_i \vartheta_j \mathbb{B}_{\sigma\sigma^\top}^{(\vartheta,T)}[\varphi'_i, \varphi'_j] \neq \sum_{i=1}^D \left(\sum_{j=1}^D \vartheta_j \mathbb{B}_{\sigma\sigma^\top}^{(\vartheta,T)}[\varphi_i, \varphi'_j] \right)^2. \quad (23)$$

If $\mathbb{B}_{\sigma\sigma^\top}^{(\vartheta,T)}$ is positive definite, part 3.) of Remark 3.5 gives the condition

$$S'_\vartheta = \sum_{i=1}^D \vartheta_i \varphi'_i \neq \sum_{i,j=1}^D \vartheta_j \mathbb{B}_{\sigma\sigma^\top}^{(\vartheta,T)}[\varphi_i, \varphi'_j] \varphi_i$$

for invertibility of both $\mathcal{I}_{(\vartheta,T)}(t)$ and $\mathcal{I}'_{(\vartheta,T)}(t)$.

2.) For $M = N$, let $\sigma \equiv 1_{N \times N}$, then $\mathbb{B}_{\sigma\sigma^\top}^{(\vartheta,T)}$ is just the standard \mathbb{L}^2 -inner product with respect to Lebesgue's measure. If $N = 1$, $D = 2d$ with $d \in \mathbb{N}$, and the signal has a finite Fourier expansion

$$S_\vartheta(s) = \sum_{k=1}^d \sqrt{2} (\vartheta_k \sin(2k\pi s) + \vartheta_{d+k} \cos(2k\pi s)) \quad \text{for all } s \in [0, \infty),$$

it is both of the type from the first part of this example and of the type introduced in part 3.) of Example 2.5 (so in particular it satisfies (S1)–(S4)). Elementary calculations show that the conditions (22) and (23) then become

$$\sum_{k=1}^d k(\vartheta_k^2 + \vartheta_{k+d}^2) \neq \alpha \sum_{k=1}^d k^2 \vartheta_{k+d}^2 \quad \text{for all } \alpha \in \{3, 4\}.$$

If for example there are no cos-terms involved, i.e. $\vartheta_{d+1} = \dots = \vartheta_D = 0$, these inequalities are valid for all $(\vartheta_1, \dots, \vartheta_d) \neq 0$.

Having introduced all relevant objects and assumptions, and having illustrated them by examples, we can now give the main result of this section.

Theorem 3.7 (Local Asymptotic Normality for \mathcal{E}_Z). *Grant all of the hypotheses (A1'), (A2'), (A3), (A4), (S1)–(S4) and (S5') and fix $(\vartheta, T) \in \Theta \times (0, \infty)$. Set*

$$\delta_n := \begin{pmatrix} n^{-1/2} & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & n^{-1/2} & 0 \\ 0 & \cdots & 0 & n^{-3/2} \end{pmatrix} \in \mathbb{R}^{(D+1) \times (D+1)} \quad \text{for all } n \in \mathbb{N},$$

and fix any bounded sequence $(h_n)_{n \in \mathbb{N}} \subset \mathbb{R}^{D+1}$. Then $\mathbb{Q}^{(\vartheta, T)}$ -almost surely we have

$$\log \frac{d\mathbb{Q}^{(\vartheta, T) + \delta_n h_n} | \mathcal{G}_n}{d\mathbb{Q}^{(\vartheta, T)} | \mathcal{G}_n} = h_n^\top \mathcal{S}_n^{(\vartheta, T)} - \frac{1}{2} h_n^\top \mathcal{I}_{(\vartheta, T)} h_n + o_{\mathbb{Q}^{(\vartheta, T)}}(1), \quad \text{as } n \rightarrow \infty, \quad (24)$$

with Fisher Information

$$\mathcal{I}_{(\vartheta, T)} = \begin{pmatrix} \left(\mathbb{B}_{\sigma\sigma^\top}^{(\vartheta, T)} [\partial_{\vartheta_i} S_\vartheta, \partial_{\vartheta_j} S_\vartheta] \right)_{i,j=1}^D & -\frac{1}{2} T^{-2} \left(\mathbb{B}_{\sigma\sigma^\top}^{(\vartheta, T)} [\partial_{\vartheta_i} S_\vartheta, S'_\vartheta] \right)_{i=1}^D \\ \cdots & \frac{1}{3} T^{-4} \mathbb{B}_{\sigma\sigma^\top}^{(\vartheta, T)} [S'_\vartheta, S'_\vartheta] \end{pmatrix}.$$

and score

$$\mathcal{S}_n^{(\vartheta, T)} = \delta_n \int_0^n \left((\sigma\sigma^\top)^{-1/2} (\eta_s) \dot{S}_{(\vartheta, T)}(s) \right)^\top d\tilde{B}_s^{(\vartheta, T)} \quad \text{for all } n \in \mathbb{N}$$

such that weak convergence

$$\mathcal{L} \left(\mathcal{S}_n^{(\vartheta, T)} \middle| \mathbb{Q}^{(\vartheta, T)} \right) \xrightarrow{n \rightarrow \infty} \mathcal{N} \left(0, \mathcal{I}_{(\vartheta, T)} \right)$$

holds.

Proof of Theorem 3.7. We fix $(\vartheta, T) \in \Theta \times (0, \infty)$, and in order to reduce notational complexity we drop corresponding indices whenever there is no risk of ambiguity: We write $\mathbb{Q} := \mathbb{Q}^{(\vartheta, T)}$, $\tilde{B} := \tilde{B}^{(\vartheta, T)}$ (see (15)), $\mathcal{S}_n := \mathcal{S}_n^{(\vartheta, T)}$, $\mathcal{I} := \mathcal{I}_{(\vartheta, T)}$, $\mathcal{I}(t) := \mathcal{I}_{(\vartheta, T)}(t)$ for all $t \in [0, \infty)$ (see (21)), and $\mathbb{B} := \mathbb{B}_{\sigma\sigma^\top}^{(\vartheta, T)}$ (see (19)). Moreover, we set

$$(\vartheta_n, T_n) := (\vartheta, T) + \delta_n h_n \quad \text{for all } n \in \mathbb{N}. \quad (25)$$

We now proceed to give the proof, divided into several steps.

1.) The main idea is to introduce a time step size $t \in (0, \infty)$ into the log-likelihood ratio and then interpret

$$\left(\log \frac{d\mathbb{Q}^{(\vartheta, T) + \delta_n h_n} | \mathcal{G}_{tn}}{d\mathbb{Q}^{(\vartheta, T)} | \mathcal{G}_{tn}} \right)_{t \in [0, \infty)}, \quad n \in \mathbb{N},$$

as a sequence of continuous-time stochastic processes. Splitting them into several parts and applying Lemma 3.4 together with tools from continuous-time martingale theory will eventually lead to the desired quadratic expansion. Indeed, adding and subtracting the term $\dot{S}_{(\vartheta,T)}(s)\delta_n h_n$ to the difference of the signals yields

$$\begin{aligned} & \log \frac{d\mathbb{Q}^{(\vartheta,T)+\delta_n h_n} |_{\mathcal{G}_{tn}}}{d\mathbb{Q}^{(\vartheta,T)} |_{\mathcal{G}_{tn}}} \\ &= \int_0^{tn} ((\sigma\sigma^\top)^{-1/2}(\eta_s)(S_{(\vartheta_n,T_n)} - S_{(\vartheta,T)})(s))^\top d\tilde{B}_s \\ & \quad - \frac{1}{2} \int_0^{tn} (S_{(\vartheta_n,T_n)} - S_{(\vartheta,T)})^\top(s) (\sigma\sigma^\top(z))^{-1} (S_{(\vartheta_n,T_n)} - S_{(\vartheta,T)})(s) ds \\ &= h_n^\top \left(\delta_n \int_0^{tn} ((\sigma\sigma^\top)^{-1/2}(\eta_s)\dot{S}_{(\vartheta,T)}(s))^\top d\tilde{B}_s \right) \\ & \quad - \frac{1}{2} h_n^\top \left(\delta_n \int_0^{tn} \dot{S}_{(\vartheta,T)}(s)^\top (\sigma\sigma^\top(\eta_s))^{-1} \dot{S}_{(\vartheta,T)}(s) ds \delta_n \right) h_n \\ & \quad + \int_0^{tn} \left((\sigma\sigma^\top)^{-1/2}(\eta_s)(S_{(\vartheta_n,T_n)} - S_{(\vartheta,T)} - \dot{S}_{(\vartheta,T)}\delta_n h_n)(s) \right)^\top d\tilde{B}_s \\ & \quad - \frac{1}{2} \int_0^{tn} (S_{(\vartheta_n,T_n)} - S_{(\vartheta,T)} - \dot{S}_{(\vartheta,T)}\delta_n h_n)^\top(s) (\sigma\sigma^\top(\eta_s))^{-1} \\ & \quad \quad \quad (S_{(\vartheta_n,T_n)} - S_{(\vartheta,T)} - \dot{S}_{(\vartheta,T)}\delta_n h_n)(s) ds \\ & \quad - \int_0^{tn} (S_{(\vartheta_n,T_n)} - S_{(\vartheta,T)} - \dot{S}_{(\vartheta,T)}\delta_n h_n)^\top(s) (\sigma\sigma^\top(\eta_s))^{-1} (\dot{S}_{(\vartheta,T)}\delta_n h_n) ds \\ &=: h_n^\top \mathcal{S}_n(t) - \frac{1}{2} h_n^\top \mathcal{I}_n(t) h_n + R_n(t) - \frac{1}{2} U_n(t) - V_n(t), \end{aligned}$$

and in order to prove the theorem, we will study convergence in distribution of $\mathcal{S}_n(t)$ for $n \rightarrow \infty$ and show almost sure convergence of $\mathcal{I}_n(1)$ to $\mathcal{I} = \mathcal{I}(1)$. Finally, we show that $R_n(t)$, $U_n(t)$, and $V_n(t)$ converge to zero in probability.

2.) For any fixed $n \in \mathbb{N}$, the process

$$M_n := (\mathcal{S}_n(t))_{t \in [0, \infty)} = \left(\delta_n \int_0^{tn} ((\sigma\sigma^\top)^{-1/2}(\eta_s)\dot{S}_{(\vartheta,T)}(s))^\top d\tilde{B}_s \right)_{t \in [0, \infty)}$$

is obviously an \mathbb{R}^{D+1} -valued local martingale with respect to \mathbb{Q} . In order to determine its weak limit for $n \rightarrow \infty$ in the Skorohod space $\mathcal{D}([0, \infty); \mathbb{R}^{D+1})$, we study its quadratic variation process $\langle M_n \rangle := (\langle M_n \rangle_t)_{t \in [0, \infty)}$ with

$$\langle M_n \rangle_t := \begin{pmatrix} \langle M_n^{(1)}, M_n^{(1)} \rangle_t & \cdots & \langle M_n^{(1)}, M_n^{(D+1)} \rangle_t \\ \vdots & \ddots & \vdots \\ \langle M_n^{(D+1)}, M_n^{(1)} \rangle_t & \cdots & \langle M_n^{(D+1)}, M_n^{(D+1)} \rangle_t \end{pmatrix} \in \mathbb{R}^{(D+1) \times (D+1)}.$$

As follows from basic stochastic calculus, $\langle M_n \rangle$ is equal to $(\mathcal{I}_n(t))_{t \in [0, \infty)}$. Consequently, for $i, j \in \{1, \dots, D\}$ we have

$$\begin{aligned} \langle M_n^{(i)}, M_n^{(j)} \rangle_t &= \frac{1}{n} \int_0^{tn} (\partial_{\vartheta_i} S_{(\vartheta, T)}(s))^\top (\sigma \sigma^\top(\eta_s))^{-1} \partial_{\vartheta_j} S_{(\vartheta, T)}(s) ds \\ &= t \cdot \frac{1}{tn} \int_0^{tn} (\partial_{\vartheta_i} S_\vartheta(s/T))^\top (\sigma \sigma^\top(\eta_s))^{-1} \partial_{\vartheta_j} S_\vartheta(s/T) ds, \end{aligned}$$

and due to the periodicity of S_ϑ and by part (ii) of Lemma 3.4 with $g = \sigma \sigma^\top$ and $k = 0$, this expression converges to

$$t \cdot \mathbb{B}[\partial_{\vartheta_i} S_\vartheta, \partial_{\vartheta_j} S_\vartheta] = \mathcal{I}_{i,j}(t)$$

\mathbb{Q} -almost surely for $n \rightarrow \infty$. Since

$$\partial_T S_{(\vartheta, T)}(s) = \partial_T S_\vartheta(s/T) = -sT^{-2} S'_\vartheta(s/T) \quad \text{for all } s \in (0, \infty),$$

the same argument with $k = 1$ yields

$$\begin{aligned} \langle M_n^{(i)}, M_n^{(D+1)} \rangle_t &= \langle M_n^{(D+1)}, M_n^{(i)} \rangle_t \\ &= \frac{1}{n^2} \int_0^{tn} (\partial_{\vartheta_i} S_{(\vartheta, T)}(s))^\top (\sigma \sigma^\top(\eta_s))^{-1} \partial_T S_{(\vartheta, T)}(s) ds \\ &= \frac{-t^2}{2T^2} \cdot \frac{1}{\frac{1}{2}(tn)^2} \int_0^{tn} s \cdot (\partial_{\vartheta_i} S_\vartheta(s/T))^\top (\sigma \sigma^\top(\eta_s))^{-1} S'_\vartheta(s/T) ds \\ &\xrightarrow{n \rightarrow \infty} \frac{-t^2}{2T^2} \cdot \mathbb{B}[\partial_{\vartheta_i} S_\vartheta, S'_\vartheta] = \mathcal{I}_{i, D+1}(t) = \mathcal{I}_{D+1, i}(t) \end{aligned}$$

\mathbb{Q} -almost surely, and analogously (with $k = 2$)

$$\begin{aligned} \langle M_n^{(D+1)}, M_n^{(D+1)} \rangle_t &= \frac{1}{n^3} \int_0^{tn} (\partial_T S_{(\vartheta, T)}(s))^\top (\sigma \sigma^\top(\eta_s))^{-1} \partial_T S_{(\vartheta, T)}(s) ds \\ &= \frac{t^3}{3T^4} \cdot \frac{1}{\frac{1}{3}(tn)^3} \int_0^{tn} s^2 \cdot S'_\vartheta(s/T)^\top (\sigma \sigma^\top(\eta_s))^{-1} S'_\vartheta(s/T) ds \\ &\xrightarrow{n \rightarrow \infty} \frac{t^3}{3T^4} \cdot \mathbb{B}[S'_\vartheta, S'_\vartheta] = \mathcal{I}_{D+1, D+1}(t) \end{aligned}$$

\mathbb{Q} -almost surely. In other words,

$$\langle M_n \rangle_t \xrightarrow{n \rightarrow \infty} \mathcal{I}(t) \quad \mathbb{Q}\text{-almost surely for all } t \in [0, \infty),$$

and hence the Martingale Convergence Theorem [27, Corollary VIII.3.24] implies weak convergence

$$\mathcal{L}(M_n | \mathbb{Q}) \xrightarrow{n \rightarrow \infty} \mathcal{L}(M | \mathbb{Q}) \quad \text{in } \mathcal{D}([0, \infty); \mathbb{R}^{D+1}) \quad (26)$$

to some limit martingale $M = (M(t))_{t \in [0, \infty)}$ with quadratic variation process $\langle M \rangle = (\mathcal{I}(t))_{t \in [0, \infty)}$.⁴ As noted in Remark 3.5, $\mathcal{I}'(t)$ is symmetric and non-negative definite, so it possesses a square root $\sqrt{\mathcal{I}'(t)} \in \mathbb{R}^{(D+1) \times (D+1)}$. By (S5'), $\mathcal{I}'(t)$ is invertible and hence $\sqrt{\mathcal{I}'(t)}$ is invertible as well. Thus, the Representation Theorem [25, Theorem II.7.1] yields that M can be expressed as

$$M(t) = \int_0^t \sqrt{\mathcal{I}'(s)} dB'_s \quad \text{for all } t \in [0, \infty)$$

with some $(D + 1)$ -dimensional Brownian Motion B' . Together with (26), this also implies weak convergence

$$\mathcal{L}(M_n(t)|\mathbb{Q}) \xrightarrow{n \rightarrow \infty} \mathcal{L}(M(t)|\mathbb{Q}) = \mathcal{N}\left(0, \int_0^t \mathcal{I}'(s) ds\right) = \mathcal{N}(0, \mathcal{I}(t))$$

for all $t \in [0, \infty)$. In particular, choosing $t = 1$ yields weak convergence of the score

$$\mathcal{L}(\mathcal{S}_n|\mathbb{Q}) = \mathcal{L}(M_n(1)|\mathbb{Q}) \xrightarrow{n \rightarrow \infty} \mathcal{N}(0, \mathcal{I}(1)) = \mathcal{N}(0, \mathcal{I}),$$

which completes this step of the proof.

3.) In the second step, we have shown on the fly that

$$\mathcal{I}_n(1) = \langle M_n \rangle_1 \xrightarrow{n \rightarrow \infty} \langle M \rangle_1 = \mathcal{I}(1)$$

\mathbb{Q} -almost surely.

4.) It remains to show convergence to zero in \mathbb{Q} -probability of the remainder terms $R_n(t)$, $U_n(t)$, and $V_n(t)$ introduced at the very beginning of this proof. Therefore, we consider the sequence $(R_n)_{n \in \mathbb{N}}$ of the local \mathbb{Q} -martingales defined by

$$R_n(t) = \int_0^{tn} \left((\sigma\sigma^\top)^{-1/2}(\eta_s) (S_{(\vartheta_n, T_n)} - S_{(\vartheta, T)} - \dot{S}_{(\vartheta, T)} \delta_n h_n)(s) \right)^\top d\tilde{B}_s$$

for all $t \in [0, \infty)$. Their quadratic variation processes are obviously given by $(U_n(t))_{t \in [0, \infty)}$. Exploiting the uniform ellipticity assumption (A2'), we can estimate

$$\begin{aligned} \langle R_n \rangle_t &= \int_0^{tn} (S_{(\vartheta_n, T_n)} - S_{(\vartheta, T)} - \dot{S}_{(\vartheta, T)} \delta_n h_n)^\top(s) (\sigma\sigma^\top(\eta_s))^{-1} \\ &\quad (S_{(\vartheta_n, T_n)} - S_{(\vartheta, T)} - \dot{S}_{(\vartheta, T)} \delta_n h_n)(s) ds \\ &\leq \sigma_0^{-1} \int_0^{tn} \left| S_{(\vartheta_n, T_n)} - S_{(\vartheta, T)} - \dot{S}_{(\vartheta, T)} \delta_n h_n \right|^2 ds \\ &= \sigma_0^{-1} \int_0^{tn} \left| S_{(\vartheta_n, T_n)} - S_{(\vartheta, T)} - D_\vartheta S_{(\vartheta, T)}(\vartheta_n - \vartheta) \right. \\ &\quad \left. - \partial_T S_{(\vartheta, T)}(s)(T_n - T) \right|^2 ds. \end{aligned} \tag{27}$$

⁴To be exact, M is actually defined on some arbitrary probability space, but in order to avoid making things more complicated than necessary, we assume without loss of generality that M is in fact defined on (a standard extension of) the same probability space as the sequence $(M_n)_{n \in \mathbb{N}}$.

Note that this upper bound is entirely deterministic. In order to prove that it in fact converges to zero, we will separate the dependence on the parameters ϑ and T in such a way that we can use the periodicity and (S1)–(S4) efficiently. This can be achieved by continuing the inequality (27) with

$$\begin{aligned} \langle R_n \rangle_t &\leq 3\sigma_0^{-1} \left(\int_0^{tn} |S_{(\vartheta_n, T_n)}(s) - S_{(\vartheta, T_n)}(s) - D_{\vartheta}S_{(\vartheta, T_n)}(s)(\vartheta_n - \vartheta)|^2 ds \right. \\ &\quad + \int_0^{tn} |(D_{\vartheta}S_{(\vartheta, T_n)} - D_{\vartheta}S_{(\vartheta, T)}(s))(\vartheta_n - \vartheta)|^2 ds \\ &\quad \left. + \int_0^{tn} |S_{(\vartheta, T_n)}(s) - S_{(\vartheta, T)}(s) - \partial_T S_{(\vartheta, T)}(s)(T_n - T)|^2 ds \right) \\ &=: 3\sigma_0^{-1}(A_n + B_n + C_n). \end{aligned}$$

We will treat convergence of A_n , B_n , and C_n step for step. For this purpose, set $H := \sup_{n \in \mathbb{N}} |h_n|$ and note that due to (25) we have

$$|\vartheta_n - \vartheta| \leq Hn^{-1/2} \quad \text{and} \quad |T_n - T| \leq Hn^{-3/2}$$

for all $n \in \mathbb{N}$.

Starting with A_n , we observe that for sufficiently large $n \in \mathbb{N}$ we have $T_n \in [T/2, 2T]$ and thus A_n can be estimated by

$$\begin{aligned} &\left(\frac{tn}{T_n} + 1 \right) \int_0^{T_n} |S_{(\vartheta_n, T_n)}(s) - S_{(\vartheta, T_n)}(s) - D_{\vartheta}S_{(\vartheta, T_n)}(s)(\vartheta_n - \vartheta)|^2 ds \\ &= \left(\frac{tn}{T_n} + 1 \right) |\vartheta_n - \vartheta|^2 \int_0^{T_n} \left| \frac{S_{(\vartheta_n, T_n)}(s) - S_{(\vartheta, T_n)}(s) - D_{\vartheta}S_{(\vartheta, T_n)}(s)(\vartheta_n - \vartheta)}{|\vartheta_n - \vartheta|} \right|^2 ds \\ &\leq \left(\frac{tn}{T/2} + 1 \right) H^2 n^{-1} \int_0^{2T} \left| \frac{S_{(\vartheta_n, T_n)}(s) - S_{(\vartheta, T_n)}(s) - D_{\vartheta}S_{(\vartheta, T_n)}(s)(\vartheta_n - \vartheta)}{|\vartheta_n - \vartheta|} \right|^2 ds, \end{aligned}$$

where the factor in front of the integral is obviously convergent. Using the \mathbb{L}^2 -continuity condition (S3) and a simple application of the Mean Value Theorem (compare Lemma 3.18 of [14]), one sees that the integral itself tends to zero.

Next, using the Hölder condition (S4), we obtain for sufficiently large $n \in \mathbb{N}$ that

$$\begin{aligned} B_n &\leq |\vartheta_n - \vartheta|^2 \int_0^{tn} |D_{\vartheta}S_{(\vartheta, T_n)}(s) - D_{\vartheta}S_{(\vartheta, T)}(s)|^2 ds \\ &\leq H^2 n^{-1} \left(\int_0^{t_0} |D_{\vartheta}S_{(\vartheta, T_n)}(s) - D_{\vartheta}S_{(\vartheta, T)}(s)|^2 ds + C(tn)^{\beta} |T_n - T|^{\alpha} \right) \\ &\leq H^2 n^{-1} \int_0^{t_0} \left| \dot{S}_{(\vartheta, T_n)}(s) - \dot{S}_{(\vartheta, T)}(s) \right|^2 ds + CH^{2+\alpha} t^{\beta} n^{\beta - (1+3\alpha/2)}. \end{aligned}$$

The particular conditions on α and β from (S4) make the second summand vanish for $n \rightarrow \infty$, while the first summand converges to zero because of (S3).

In order to estimate C_n , we make explicit use of the C^2 -property (S1) which is readily translated into the condition that the mapping

$$(0, \infty) \ni T \mapsto S_{(\vartheta, T)}(s)$$

is twice continuously differentiable for any fixed $s \in (0, \infty)$. Consequently, for every $s \in (0, \infty)$ and any $i \in \{1, \dots, N\}$ Taylor expansion with the Lagrange form of the remainder provides a $\varrho_i = \varrho_i(s, \vartheta, T, T_n, h_n)$ between T and T_n such that for sufficiently large $n \in \mathbb{N}$ we can infer that

$$\begin{aligned} & |S_{(\vartheta, T_n)}(s) - S_{(\vartheta, T)}(s) - (T_n - T)\partial_T S_{(\vartheta, T)}(s)|^2 \\ &= \sum_{i=1}^N \left(\frac{1}{2} (T_n - T)^2 \partial_T^2 S_{(\vartheta, T)}^{(i)}(s)|_{T=\varrho_i} \right)^2 \\ &= \frac{1}{4} (T_n - T)^4 \sum_{i=1}^N \left(\frac{s^2}{\varrho_i^4} \left(S_{\vartheta}^{(i)} \right)''(s/\varrho_i) + \frac{2s}{\varrho_i^3} \left(S_{\vartheta}^{(i)} \right)'(s/\varrho_i) \right)^2 \\ &\leq \frac{1}{4} H^4 n^{-6} 2N \left[\left(s^2 \frac{\|S_{\vartheta}''\|_{\infty}}{(T - n^{-3/2}H)^4} \right)^2 + \left(s \frac{2\|S_{\vartheta}'\|_{\infty}}{(T - n^{-3/2}H)^3} \right)^2 \right] \\ &\leq \text{cst } n^{-6} (s^4 + s^2) \end{aligned}$$

for some positive constant not depending on s or n . Integrating yields

$$C_n \leq \text{cst } n^{-6} \int_0^{tn} (s^4 + s^2) ds$$

and hence C_n vanishes for $n \rightarrow \infty$.

So far, we have shown that the sequence of random variables $(U_n(t))_{n \in \mathbb{N}}$ not only vanishes in probability under \mathbb{Q} for $n \rightarrow \infty$, but is even bounded by a deterministic sequence which goes to zero. Therefore,

$$\mathbb{E}_{\mathbb{Q}}[R_n(t)^2] = \mathbb{E}_{\mathbb{Q}}[\langle R_n \rangle_t] = \mathbb{E}_{\mathbb{Q}}[U_n(t)] \xrightarrow{n \rightarrow \infty} 0, \tag{28}$$

and in particular, $R_n(t)$ also vanishes in probability under \mathbb{Q} for $n \rightarrow \infty$. Finally, the same is true for the last remainder variable $V_n(t)$, as by the Cauchy-Schwarz inequality we get that

$$|V_n(t)|^2 \leq U_n(t) h_n^{\top} \mathcal{I}_n(t) h_n \leq U_n(t) H^2 |\mathcal{I}_n(t)| \xrightarrow{n \rightarrow \infty} 0, \tag{29}$$

since $\mathcal{I}_n(t)$ converges and $U_n(t)$ goes to zero. Taking $t = 1$ completes the proof. \square

Remark 3.8. The convergence in probability for $n \rightarrow \infty$ of the remainder terms $R_n(t)$, $U_n(t)$, and $V_n(t)$ (which determine the term $o_{\mathbb{Q}(\vartheta, T)}(1)$ in (24)) is in fact even uniform with respect to $t \in [0, t_0]$ for every $t_0 \in (0, \infty)$. This is clear for $U_n(t)$, since it only increases with t . Using the Burkholder-Davis-Gundy inequality, the estimation (28) can be improved to

$$\mathbb{E}_{\mathbb{Q}} \left[\sup_{t \in [0, t_0]} |R_n(t)|^2 \right] \leq 4\mathbb{E}_{\mathbb{Q}}[\langle R_n \rangle_{t_0}] = 4\mathbb{E}_{\mathbb{Q}}[U_n(t_0)] \xrightarrow{n \rightarrow \infty} 0,$$

which also takes care of $R_n(t)$. For $V_n(t)$, we notice that the bound given in (29) only depends on t via $\mathcal{I}_n(t)$ and $U_n(t)$ which are both non-decreasing with respect to t .

Remark 3.9. In the one-dimensional case $M = N = 1$, there are already known variants of Theorem 3.7 that treat shape and periodicity separately.

For the submodel in which the shape parameter ϑ is known and the only parameter of interest is the periodicity T , Theorem 3.7 includes [19, Theorem 1.1] as a special case (for this purpose, our assumptions can be reduced to (A1'), (A2'), (A3), (A4) and the C^2 -property from (S1)).

For the submodel for known periodicity and unknown shape on the other hand, Theorem 3.7 leads to the same conclusion as [17, Theorem 2.1] (note that Score and Fisher Information are given on the time scale of the known periodicity T there). In [17], the \mathbb{L}^2 -smoothness conditions on the signal are not assumed to hold with respect to Lebesgue's measure ds as in (S2) and (S3) but with respect to the measure

$$\nu^\vartheta(ds) := \mu^{(\vartheta, T)} P_{0, sT}^{(\vartheta, T)}(\sigma^{-2}) ds$$

which is assumed to be locally finite. In practice, the latter property will essentially always be checked via uniform ellipticity (A2'). Assuming (A2'), the density of ν^ϑ is bounded, and hence the corresponding \mathbb{L}^2 -smoothness assumptions can most naturally be checked via our conditions (S2) and (S3). In this sense, the difference between our assumptions and those in [17] is negligible.

The key to bringing these results together in Theorem 3.7 is the Hölder condition (S4) which is crucial for dealing with the term B_n in step 3.) of its proof. This is the only instant where (in contrast to the terms A_n and C_n) we have to impose more than just "joint smoothness", but a more specific relation of the interplay between T and ϑ . It should also be noted that (A2') is essential for this step, as it yields a purely deterministic bound for the remainder terms. Otherwise, even if we restated (S2)–(S4) with \mathbb{L}^2 -convergence replaced by convergence with respect to the semi-norm induced by $\mathbb{B}_{\sigma\sigma^\top}^{(\vartheta, T)}$, we could not treat this term with the methods used in Lemma 3.4 due to the occurrence of different periodicities in the integrand.

Remark 3.10. In the two submodels corresponding to a separate treatment of shape and periodicity, the following is known about estimators achieving the efficiency bounds following from Theorem 3.7. Keep in mind that by Remark 3.1 the same estimators can be used for the richer process (X, Y, Z) , even if only X is actually observable, and by Theorem 2.3 the same efficiency bounds hold in this case.

First assume that T is known, and that we would like to estimate the unknown shape parameter ϑ . In this case, Section 4 of [17] (the general assumptions of which are covered by ours, as already noted in Remark 3.9) provides a minimum distance estimator that can be made efficient via one-step-modification (see [17, Theorem 4.1], and discusses efficiency of the maximum likelihood and Bayesian estimator (see [17, Remark 4.2]).

Next assume that ϑ is known, and that we would like to estimate the unknown periodicity T . In this situation, [20, Theorem 1] states that the corresponding maximum likelihood and Bayesian estimator are efficient in the sense of [20, Equation (21)]. Note that this property is slightly different from efficiency in the sense of LeCam (see [15, p. 194]) and that it is proved under the assumption that the convergence of certain functionals (similar to (20) in our setting) is locally uniform with respect to T .

Remark 3.11. For $M = N$, $\sigma \equiv 1_{N \times N}$, and any drift $b: \mathbb{R}^N \rightarrow \mathbb{R}^N$, the process $\tilde{Z} = (\tilde{Z}_t)_{t \in [0, \infty)}$ defined by

$$\tilde{Z}_t := Z_t - \int_0^t b(Z_s) ds \quad \text{for all } t \in [0, \infty)$$

fulfills the “signal in white noise” equation

$$d\tilde{Z}_t = S_{(\vartheta, T)}(t)dt + dW_t.$$

This transformation removes some complexity from the model and relates it to a well-studied equation. Note however that even if Z satisfies the recurrence assumption (A4), \tilde{Z} does not.

For \tilde{Z} with $M = N = 1$, Golubev (see [8], or compare [1] for a more detailed probabilistic explanation) proposes an estimator for T under unknown infinite-dimensional ϑ (the vector of the Fourier-coefficients of the signal) which he proves to be asymptotically normal and efficient, where the normalisation factor (translated into our notation) is given by

$$n^{3/2} \left(\frac{1}{12T^4} \int_0^1 (S'_\vartheta(s))^2 ds \right)^{1/2} = (\delta_n)_{d+1, d+1}^{-1} \left(\frac{1}{4} \mathcal{I}_{d+1, d+1} \right)^{-1/2}.$$

This corresponds to what is indeed the optimal limit variance in Theorem 3.7: The factor $\frac{1}{4}$ stems from the fact that [8] deals with a two-sided model which is statistically equivalent to two independent copies of our one-sided model (compare [12, Example 2.7]).

References

- [1] I. CASTILLO, C. LÉVY-LEDUC, C. MATIAS: Exact adaptive estimation of the shape of a periodic function with unknown period corrupted by white noise. In: *Mathematical Methods of Statistics* Vol. 15 (2006), pp. 1–30. [MR2256473](#)
- [2] N. CUNEO, J. P. ECKMANN: Non-equilibrium steady states for chains of four rotors. In: *Communications in Mathematical Physics* Issue 1 (2016), pp. 185–221. [MR3509013](#)
- [3] N. CUNEO, J. P. ECKMANN, M. HAIRER, L. REY-BELLET: Non-equilibrium steady states for networks of oscillators. In: *Electronic Journal of Probability* Vol. 23 (2018), no. 55, pp. 1–28. [MR3814249](#)

- [4] N. CUNEO, J. P. ECKMANN, C. POQUET: Non-equilibrium steady state and subgeometric ergodicity for a chain of three coupled rotors. In: *Non-linearity* Vol. 28 (2015), pp. 2397–2421. [MR3366649](#)
- [5] R. DAVIES: Asymptotic inference when the amount of information is random. In: *Proceedings of the Berkeley Symposium in Honour of J. Neyman and J. Kiefer* Vol. II, Wadsworth, 1985. [MR0822069](#)
- [6] A. DESTEXHE: Conductance-based integrate and fire models. In: *Neural Computation* Vol. 9 (1997), pp. 503–514.
- [7] S. DITLEVSEN, P. LÁNSKÝ: Estimation of the input parameters in the Feller neuronal model. In: *Physical Review E* Vol. 73 (2006), 061910. [MR2276280](#)
- [8] G. GOLUBEV: Estimating the period of a signal of unknown shape corrupted by white noise. In: *Problems in Information Transmission* Vol. 24 (1988), pp. 38–52. [MR0979694](#)
- [9] M. HAIRER: On Malliavin’s proof of Hörmander’s Theorem. In: *Bulletin des Sciences Mathématiques* Vol. 135 (2011), pp. 650–666. [MR2838095](#)
- [10] M. HAIRER, J. C. MATTINGLY: Slow energy dissipation in anharmonic oscillator chains. In: *Communications on Pure and Applied Mathematics* Vol. 62 (2009), pp. 999–1032. [MR2531551](#)
- [11] A. L. HODGKIN, A. F. HUXLEY: A quantitative description of membrane current and its application to conduction and excitation in nerve. In: *Journal of Physiology* Vol. 117 (1952), pp. 500–544.
- [12] S. HOLBACH: Local asymptotic normality for shape and periodicity in the drift of a time inhomogeneous diffusion. In: *Statistical Inference for Stochastic Processes* Vol. 21 (2018), pp. 527–538. [MR3846990](#)
- [13] S. HOLBACH: Positive Harris recurrence for degenerate diffusions with internal variables and randomly perturbed time-periodic input. [arXiv:1907.13585](#) [math.PR], 2019.
- [14] S. HOLBACH: Recurrence and parameter estimation for degenerate diffusions with internal variables and randomly perturbed time-inhomogeneous deterministic input. Dissertation, Johannes Gutenberg-Universität Mainz, 2018, available online at <http://publications.ub.uni-mainz.de/theses/volltexte/2018/100002300/pdf/100002300.pdf>
- [15] R. HÖPFNER: Asymptotic Statistics with a View to Stochastic Processes. de Gruyter, 2014. [MR3185373](#)
- [16] R. HÖPFNER: On a set of data for the membrane potential in a neuron. In: *Mathematical Biosciences* Vol. 207 (2007), pp. 275–301. [MR2331416](#)
- [17] R. HÖPFNER, Y. A. KUTOYANTS: On LAN for parametrized continuous periodic signals in a time inhomogeneous diffusion. In: *Statistics & Decisions* Vol. 27 (2009), pp. 309–326. [MR2791380](#)
- [18] R. HÖPFNER, Y. A. KUTOYANTS: Estimating discontinuous periodic signals in a time inhomogeneous diffusion. In: *Statistical Inference for Stochastic Processes* Vol. 13 (2010), pp. 193–230. [MR2729649](#)
- [19] R. HÖPFNER, Y. A. KUTOYANTS: Estimating a periodicity parameter in the drift of a time inhomogeneous diffusion. In: *Mathematical Methods of Statistics* Vol. 20 (2011), pp. 58–74. [MR2811031](#)

- [20] R. HÖPFNER, Y. A. KUTOYANTS: On frequency estimation for a periodic ergodic diffusion process. In: *Problems of Information Transmission* Vol. 48(2) (2012), pp. 127–141. [MR3226954](#)
- [21] R. HÖPFNER, E. LÖCHERBACH, M. THIEULLEN: Ergodicity for a stochastic Hodgkin-Huxley model driven by Ornstein-Uhlenbeck type input. In: *Annales de l'Institut Henri Poincaré* Vol. 1 (2016), pp. 483–501. [MR3449311](#)
- [22] R. HÖPFNER, E. LÖCHERBACH, M. THIEULLEN: Strongly degenerate time inhomogeneous SDEs: densities and support properties. Application to a Hodgkin-Huxley system with periodic input. In: *Bernoulli* Vol. 23(4A) (2017), 2587–2616. [MR3648039](#)
- [23] R. HÖPFNER, E. LÖCHERBACH, M. THIEULLEN: Ergodicity and limit theorems for degenerate diffusions with time periodic drift. Application to a stochastic Hodgkin-Huxley model. In: *ESAIM P&S* Vol. 20 (2016), pp. 527–554. [MR3581833](#)
- [24] I. A. IBRAGIMOV, R. Z. KHASHINSKII: Statistical Estimation. Springer, 1981. [MR0620321](#)
- [25] N. IKEDA, S. WATANABE: Stochastic Differential Equations and Diffusion Processes. North-Holland Library, 2nd edition, 1989. [MR1011252](#)
- [26] E. M. IZHIKEVICH: Dynamical Systems in Neuroscience. The MIT Press, 2007. [MR2263523](#)
- [27] J. JACOD, A. SHIRYAEV: Limit Theorems for Stochastic Processes. Springer, 2nd edition, 2002. [MR1943877](#)
- [28] M. KIRSZBRAUN: Über die zusammenziehende und Lipschitzsche Transformationen. In: *Fundamenta Mathematicae* Vol. 22 (1935), pp. 77–108.
- [29] Y. A. KUTOYANTS: Statistical Inference for Ergodic Diffusion Processes. Springer, 2004. [MR2144185](#)
- [30] P. LÁNSKÝ, L. SACERDOTE, F. TOMASSETTI: On the comparison of Feller and Ornstein-Uhlenbeck models for neural activity. In: *Biological Cybernetics* Vol. 73 (1995), pp. 457–465.
- [31] L. LECAM, G. YANG: Asymptotics in Statistics. Some Basic Concepts. Springer, 1990. [MR1066869](#)
- [32] S. MEYN, R. TWEEDIE: Stability of Markovian processes I: criteria for discrete-time chains. In: *Advances in Applied Probability* Vol. 24 (1992), pp. 542–574. [MR1174380](#)
- [33] C. MORRIS, H. LECAR: Voltage oscillations in the barnacle giant muscle fiber. In: *Biophysical Journal* Vol. 35 (1981), pp. 193–213.
- [34] L. REY-BELLET, L. E. THOMAS: Exponential convergence to non-equilibrium stationary states in classical statistical mechanics. In: *Communications in Mathematical Physics* Vol. 225 (2002), pp. 309–329. [MR1889227](#)
- [35] J. RINZEL, B. ERMENTROUT: Analysis of neural excitability and oscillations. In: *Methods in Neuronal Modeling: From Ions to Networks, 2nd edition* (1998), pp. 251–291.