March 26, 2010

DRIFT ESTIMATION FOR A PERIODIC MEAN REVERSION PROCESS

HEROLD DEHLING, BRICE FRANKE, AND THOMAS KOTT

ABSTRACT. In this paper we propose a periodic, mean-reverting Ornstein-Uhlenbeck process of the form
$$dX_t = (L(t) - \alpha\, X_t)\, dt + \sigma\, dB_t, \quad t \geq 0,$$
where $L(t)$ is a periodic, parametric function. We apply maximum likelihood estimation for the drift parameters based on time-continuous observations. The estimator is given explicitly and we prove strong consistency and asymptotic normality as the observed number of periods tends to infinity. The essential idea of the asymptotic study is the interpretation of the stochastic process as a sequence of random variables that take values in some function space.

1. INTRODUCTION

The ordinary Ornstein-Uhlenbeck process is defined as solution to the stochastic differential equation
$$dX_t = \alpha(\mu - X_t)dt + \sigma dB_t, \quad t \geq 0,$$
where α and σ are positive constants, $\mu \in \mathbb{R}$ and where X_0 with $\mathrm{E}(X_0^2) < \infty$ is a real-valued random variable which is independent of the standard Brownian motion $(B_t)_{t \geq 0}$. Originally introduced by Ornstein and Uhlenbeck (1932) as a model for particle motion in a fluid, this process is now widely used in many areas of application. The main characteristic of the Ornstein-Uhlenbeck process is the tendency to return towards the long-term equilibrium μ. This property, known as *mean-reversion*, is found in many real life processes, e.g. in commodity and energy price processes, see e.g. Geman (2005).

In many real-life applications, however, the assumption of a constant mean level is not adequate due to seasonality patterns or a long-term trend of the process. Thus we want to consider the more general process satisfying the stochastic differential equation

$$(1) \qquad dX_t = (L(t) - \alpha X_t)dt + \sigma dB_t, \quad t \geq 0,$$

where $L(t)$ is a time-dependent mean reversion level and where α, σ are positive constants. Note that model (1) differs from the original Ornstein-Uhlenbeck process in the position of α within the drift term. However, model (1) can easily be transformed to a process with drift term $\alpha(\tilde{L}(t) - X_t)dt$ where $\tilde{L}(t) = L(t)/\alpha$. The advantage of (1) compared with the process provided with the drift $\alpha(L(t) - X_t)dt$ is the simplification of the study of the estimators.

In this paper we make a parametric model for the mean reversion function $L(t)$. We assume that

$$(2) \qquad L(t) = \sum_{i=1}^{p} \mu_i \varphi_i(t),$$

where the basis functions $\varphi_1(t), \ldots, \varphi_p(t)$ are known and $\mu_1, \ldots, \mu_p$ and α are unknown parameters. In contrast, the diffusion parameter σ is assumed to be known which is a

common assumption in the field of drift parameter estimation for a time-continuous diffusion. This is due to the fact that the measures corresponding to different diffusion parameters are singular so that σ can be computed, rather than estimated, from a single continuous-time observation path.

The conditions on the drift coefficient, here

$$(3) \qquad S(\theta, t, X_t) = \sum_{i=1}^{p} \mu_i \varphi_i(t) - \alpha X_t, \quad \theta = (\mu_1, \ldots \mu_p, \alpha)^t,$$

that ensure existence and uniqueness of a solution of equation (1) are well known, see Kuo [4] (Theorem 10.3.5, p. 192), for example. Due to the linear form of $S(\theta, t, \cdot)$ the global Lipschitz condition is satisfied in our setting such that there exists at most one solution of (1). If the basis functions $\varphi_1(t), \ldots, \varphi_p(t)$ are bounded on compact sets in $\mathbb{R}$, for instance, then the linear growth condition which implies the uniqueness of an existing solution is fulfilled.

Drift parameter estimation for time-continuously observed diffusion processes is a well-established area of research, for which a variety of techniques has been proposed. For example, Kutoyants [5] investigates several estimation techniques for the drift term of ergodic, time-homogenous diffusion processes. The analysis of asymptotic properties of drift estimates for time-inhomogeneous diffusion models has been paid much less attention. Among some other authors, Bishwal [1] studies the maximum likelihood estimator for a time-inhomogeneous diffusion provided with drift function $f(\theta, t, X)$ defined on $\Theta \times [0, T] \times C[0, T]$, where Θ is the parameter space and $C[0, T]$ is the space of all continuous, real-valued functions on $[0, T]$. Note that our model presented above belongs to this class of diffusion processes and that our results comply with Bishwal's findings. However, in order to apply Bishwal's results directly, one has to verify the required conditions, among others the convergence of

$$(4) \qquad \frac{1}{m_T} \int_0^T \left(\frac{\partial}{\partial \theta} f(\theta, t, X_t) \right)^2 dt$$

where m_T is an increasing, non-random sequence, see Bishwal [1] condition (A6) on p. 65. The verification of this condition does not go without saying in our model and requires some auxiliary ideas and results. In contrast to Bishwal's general setting, we obtain an explicit representation of the estimator which is used for the study of its asymptotic behavior. By investigating this representation, the convergence of the multi-dimensional version of (4) is shown, see Proposition 4.5 where $\frac{1}{T}Q_T$ is the analog to the term in (4).

2. MAXIMUM LIKELIHOOD ESTIMATOR

Let us denote by P_X the measure induced by the observable realizations $X^T = \{X_t, 0 \leq t \leq T\}$ on the measurable space $(C[0, T], \mathcal{B}[0, T])$, $C[0, T]$ being the space of continuous, real-valued functions on $[0, T]$ and $\mathcal{B}[0, T]$ the associated Borel σ-field. Moreover, let P_B be the measure generated by the Brownian motion on $(C[0, T], \mathcal{B}[0, T])$. Then the likelihood function of observations X^T of the process with stochastic differential (1) is defined as

$$\mathcal{L}(\theta, X^T) := \frac{dP_X}{dP_B}$$

where dP_X/dP_B is the Radon-Nikodym derivative. The maximum likelihood estimator is defined as the maximum of the functional $\theta \mapsto \mathcal{L}(\theta, X^T)$, i.e.

$$\hat{\theta}_{ML} := \arg\max_{\theta} \mathcal{L}(\theta, X^T).$$

A corollary to Girsanov's theorem, see Theorem 7.6 on p. 246 in [7] by Lipster and Shiryayev, gives an explicit expression of the likelihood function of a diffusion process provided that

(5)
$$P\left(\int_0^T S(\theta, t, X_t)^2 dt < \infty\right) = 1$$

for all $0 \leq T < \infty$ and all θ.

Lemma 2.1. *Let $\mathcal{L}(\theta, X^T)$ denote the likelihood function of observations $X^T = \{X_t, 0 \leq t \leq T\}$ of the process introduced in (1) provided with the mean reversion function (2). If the drift term given in (3) satisfies condition (5) then*

$$\hat{\theta}_{ML} = Q_T^{-1} P_T.$$

The objects $Q_T \in \mathbb{R}^{(p+1)\times(p+1)}$ and $P_T \in \mathbb{R}^{p+1}$ are defined as

$$Q_T = \begin{pmatrix} G_T & -a_T \\ -a_T^t & b_T \end{pmatrix},$$

$$P_T = \left(\int_0^T \varphi_1(t)dX_t, \ldots, \int_0^T \varphi_p(t)dX_t - \int_0^T X_t\, dX_t\right)^t$$

where $G_T = (\int_0^T \varphi_j(t)\varphi_k(t)dt)_{1\leq j,k\leq p} \in \mathbb{R}^{p\times p}$, $a_T = (\int_0^T \varphi_1(t)X_t\, dt, \ldots, \int_0^T \varphi_p(t)X_t\, dt)^t$ and $b_T = \int_0^T X_t^2 dt$.

Proof. The likelihood function of a diffusion process of the form

$$dX_t = S(\theta, t, X_t)dt + \sigma dB_t, \quad 0 \leq t \leq T,$$

is given by

(6)
$$\mathcal{L}(\theta, X^T) = \frac{dP_X}{dP_B} = \exp\left(\frac{1}{\sigma^2}\int_0^T S(\theta, t, X_t)dX_t - \frac{1}{2\sigma^2}\int_0^T S(\theta, t, X_t)^2 dt\right)$$

if condition (5) is fulfilled, see Lipster and Shiryayev [7] (Theorem 7.6, p. 246). The partial derivatives of the logarithm of this functional are

(7)
$$\frac{\partial}{\partial\theta_i}\ln(\mathcal{L}(\theta, X^T)) = \frac{1}{\sigma^2}\int_0^T \frac{\partial}{\partial\theta_i}S(\theta, t, X^T)dX_t - \frac{1}{\sigma^2}\int_0^T S(\theta, t, X_t)\frac{\partial}{\partial\theta_i}S(\theta, t, X_t)dt.$$

The drift function of our mean reversion model is given in (3) and the derivatives are

$$\frac{\partial}{\partial\theta_i}S(\theta, t, X_t) = \begin{cases} \varphi_i(t), & i = 1\ldots, p; \\ -X_t, & i = p+1. \end{cases}$$

Setting the partial derivatives of the log-likelihood function in (7) equal zero gives a system of linear equations which yields the assertion.

$\square$

Remark 1. *Note that the matrix Q_T introduced in the previous lemma is not a priori invertible. However, we will see later that it is invertible for T large enough in the periodic mean reversion model, see Remark 3 for more details on that.*

Remark 2. *The maximum likelihood estimator introduced above can be motivated by an alternative derivation: Interpreting Euler's discretization of the stochastic differential equation (1) as a linear model and applying the ordinary least squares estimation method to this discrete version provides an estimator $\hat{\theta}_T^{\Delta t}$ containing Riemann and Itô sums. It can be then seen that $\hat{\theta}_T^{\Delta t} \to \hat{\theta}_{ML}$ as $\Delta t \to 0$.*

3. MAXIMUM LIKELIHOOD ESTIMATION FOR A PERIODIC MEAN REVERSION FUNCTION

In many applications, the data display regular seasonal effects. These can be modeled by assuming that the mean-reversion function $L(t)$ is periodic, i.e. that

$$(8) \qquad L(t + \nu) = L(t)$$

where ν is the period observed in the data. The resulting stochastic process exhibits a cyclical evolution due to the periodicity of this mean reversion mechanism. Combining the assumption of periodicity with the parametric model (2) leads to the requirement

$$(9) \qquad \varphi_j(t + \nu) = \varphi_j(t).$$

By applying Gram-Schmidt orthogonalization, we may assume without loss of generality that $\varphi_1(t), \ldots, \varphi_p(t)$ form an orthonormal system in $L_2([0, \nu], \frac{1}{\nu}d\lambda)$, i.e. that

$$(10) \qquad \int_0^\nu \varphi_j(t)\varphi_k(t)dt = \begin{cases} \nu, & j = k \\ 0, & j \neq k. \end{cases}$$

In the rest of this paper we will assume that we observe an integral multiple of the period length, i.e. that

$$T = N\nu,$$

for some integer N. Moreover, we will assume without loss of generality that $\nu = 1$.

Under the above assumptions, the matrix Q_T, defined in (6), simplifies to

$$(11) \qquad Q_T = \begin{pmatrix} T\,I_p & -a_T \\ -a_T^t & b_T \end{pmatrix}$$

where I_p denotes the $(p \times p)$-identity matrix. The inverse of a matrix of this special form can be explicitly computed by the following lemma.

Lemma 3.1. *The inverse of the matrix Q_T, given in (11), is given by*

$$(12) \qquad Q_T^{-1} = \frac{1}{T}\begin{pmatrix} I_p + \gamma_T \Lambda_T \Lambda_T^t & -\gamma_T \Lambda_T \\ -\gamma_T \Lambda_T^t & \gamma_T \end{pmatrix}$$

where

$$(13) \qquad \Lambda_{T,i} = \frac{1}{T}\int_0^T \varphi_i(t)X_t dt, \quad i = 1, \ldots, p$$

$$(14) \qquad \gamma_T = \left(\frac{1}{T}\int_0^T X_t^2 dt - \sum_{i=1}^p \Lambda_{T,i}^2\right)^{-1}$$

and $\Lambda_T = (\Lambda_{T,1}, \ldots, \Lambda_{T,p})^t$.

Proof. We make use of the following formula for the inverse of a partitioned matrix which can be deduced from the Frobenius matrix inversion formula, cf. Gantmacher [2], p. 73. Alternatively the formula can also be verified directly. We have for $a \in \mathbb{R}^p, b \in \mathbb{R}$

$$(15) \qquad \begin{pmatrix} I_p & a \\ a^t & b \end{pmatrix}^{-1} = \begin{pmatrix} I_p + \frac{1}{b-\|a\|^2} aa^t & -\frac{1}{b-\|a\|^2} a \\ -\frac{1}{b-\|a\|^2} a^t & \frac{1}{b-\|a\|^2} \end{pmatrix},$$

where $\|\cdot\|$ denotes the usual Euclidean norm on $\mathbb{R}^p$. With the notation introduced above, we can write Q_T as follows,

$$Q_T = T \begin{pmatrix} I_p & -\Lambda_T \\ -\Lambda_T^t & \frac{1}{T} \int_0^T X_t^2 dt \end{pmatrix}$$

and thus apply the above formula for the calculation of Q_T^{-1}. □

Remark 3. *Note that the Frobenius matrix inversion formula holds if and only if the entries of the matrix on the right hand side of (15) are well-defined. We will see in the proof of Proposition 4.5 that the limit of $\frac{1}{T} Q_T^{-1}$ is well defined since we show that the limit of γ_T denoted by γ is greater than zero. Consequently, $\frac{1}{T} Q_T^{-1}$ exists almost surely if T is large enough.*

We can now formulate our main results about the asymptotic behavior of the maximum likelihood estimator in the periodic Ornstein-Uhlenbeck model.

Theorem 1. *Let $\{X_t, 0 \le t \le T\}$ be observations of the periodic mean reversion process as introduced in (2), satisfying (9) and (10). Then the maximum likelihood estimator given in Lemma 2.1 is consistent, i.e.*

$$\hat{\theta}_{ML} \to \theta, \ almost \ surely,$$

as $T \to \infty$.

For the description of the asymptotic distribution of $\hat{\theta}_{ML}$, we have to introduce the $(p+1) \times (p+1)$ matrix

$$(16) \qquad C = \begin{pmatrix} I_p + \gamma \Lambda \Lambda^t & -\gamma \Lambda \\ -\gamma \Lambda^t & \gamma \end{pmatrix}$$

where the entries are defined by

$$(17) \qquad \Lambda_i = \int_0^1 \varphi_i(t) \tilde{h}(t) dt, \quad i = 1, \ldots, p$$

$$(18) \qquad \gamma = \left(\int_0^1 (\tilde{h}(t))^2 dt + \frac{\sigma^2}{2\alpha} - \sum_{i=1}^p \Lambda_i^2 \right)^{-1}$$

and $\Lambda = (\Lambda_1, \ldots, \Lambda_p)^t$. Here, the function $\tilde{h} : [0, \infty) \to \mathbb{R}$ is defined by

$$(19) \qquad \tilde{h}(t) = e^{-\alpha t} \sum_{j=1}^p \mu_j \int_{-\infty}^t e^{\alpha s} \varphi_j(s) ds.$$

Theorem 2. *Let $\{X_t, 0 \leq t \leq T\}$ be observations of the periodic mean reversion process as introduced in (2), satisfying (9) and (10). Then the maximum likelihood estimator given in Lemma 2.1 is asymptotically normal. More precisely,*

$$\sqrt{T}\sigma^{-1}(\hat{\theta}_{ML} - \theta) \to N(0, C),$$

where C is defined as in (16).

The proofs of these theorems require a number of auxiliary results, which will be given in the next two sections.

4. PROOF OF THEOREM 1

The proofs of Theorem 1 and Theorem 2 make use of a representation of the maximum likelihood estimator that will be established in the following proposition.

Proposition 4.1. *The maximum likelihood estimator $\hat{\theta}_{ML}$, defined in Lemma 2.1, can be written as*

$$(20) \qquad \hat{\theta}_{ML} = \theta + \sigma Q_T^{-1} R_T,$$

where

$$(21) \qquad R_T := \begin{pmatrix} \int_0^T \varphi_1(t) dB_t \\ \vdots \\ \int_0^T \varphi_p(t) dB_t \\ -\int_0^T X_t dB_t \end{pmatrix},$$

and where Q_T is defined in Lemma 2.1.

Proof. By definition, we have

$$\hat{\theta}_{ML} = Q_T^{-1} P_T,$$

where Q_T and P_T are defined as in Lemma 2.1. We rewrite this by making use of (1). In fact, the stochastic integrals in P_T are understood in accordance to

$$dX_t = \Big(\sum_{j=1}^p \mu_j \varphi_j(t) - \alpha X_t\Big) dt + \sigma dB_t$$

as

$$\int_0^T \varphi_i(t) dX_t = \sum_{j=1}^p \mu_j \int_0^T \varphi_i(t)\varphi_j(t) dt - \alpha \int_0^T \varphi_i(t) X_t dt + \sigma \int_0^T \varphi_i(t) dB_t, \quad i = 1, \dots, p,$$

$$\int_0^T X_t dX_t = \sum_{j=1}^p \mu_j \int_0^T X_t \varphi_j(t) dt - \alpha \int_0^T X_t^2 dt + \sigma \int_0^T X_t dB_t.$$

Hence, it follows that

$$P_T = \begin{pmatrix} \int_0^T \varphi_1(t) dX_t \\ \vdots \\ \int_0^T \varphi_p(t) dX_t \\ -\int_0^T X_t dX_t \end{pmatrix} = Q_T \theta + \sigma R_T$$

so that $\hat{\theta}_{ML} = \theta + \sigma Q_T^{-1} R_T$. $\qquad\qquad\square$

In what follows, we will show that $Q_T^{-1} R_T$ converges to zero almost surely, as $T \to \infty$. In order to do so, we write

$$Q_T^{-1} R_T = \left(T Q_T^{-1}\right) \left(\frac{1}{T} R_T\right).$$

We will show that $T Q_T^{-1}$ converges almost surely to a finite limit and that $\frac{1}{T} R_T$ converges almost surely to zero. Both of these results require some auxiliary results which will be proved first.

Lemma 4.2. *The solution of the stochastic differential equation (1) has the explicit representation*

$$(22) \qquad\qquad X_t = e^{-\alpha t} X_0 + h(t) + Z_t,$$

where

$$h(t) = e^{-\alpha t} \int_0^t e^{\alpha s} L(s) ds = e^{-\alpha t} \sum_{i=1}^p \mu_i \int_0^t e^{\alpha s} \varphi_i(s) ds$$

and

$$Z_t = \sigma e^{-\alpha t} \int_0^t e^{\alpha s} dB_s.$$

Proof. The Itô lemma states for $Y_t = g(t, X_t)$ that

$$dY_t = \frac{\partial g}{\partial t}(t, X_t) dt + \frac{\partial g}{\partial x}(t, X_t) dX_t + \frac{1}{2} \frac{\partial^2 g}{\partial x^2}(t, X_t)(dX_t)^2$$

which reduces for $g(t, x) = e^{\alpha t} x$ to

$$dY_t = \alpha e^{\alpha t} X_t dt + e^{\alpha t} dX_t.$$

Plugging (1) in this equation gives

$$dY_t = e^{\alpha t}(L(t) dt + \sigma dB_t).$$

Integrating and multiplying by $e^{-\alpha t}$ finishes the proof of the lemma. $\qquad\square$

The process $(X_t)_{t \geq 0}$ is not stationary, since we have chosen an arbitrary initial random variable. Thus we are unable to apply the ergodic theorem. In order to solve this problem, we will next introduce a stationary solution to the stochastic differential equation (1) for $t \in \mathbb{R}$ instead of $t \geq 0$. We define the process

$$(23) \qquad\qquad \tilde{X}_t = \tilde{h}(t) + \tilde{Z}_t$$

where $\tilde{h}(t)$ is defined in (19) and $\tilde{Z}_t$ is defined as

$$(24) \qquad\qquad \tilde{Z}_t = \sigma e^{-\alpha t} \int_{-\infty}^t e^{\alpha s} d\tilde{B}_s$$

where $(\tilde{B}_s)_{s \in \mathbb{R}}$ denotes a bilateral Brownian motion, i.e.

$$\tilde{B}_s := B_s \mathbf{1}_{\mathbb{R}_+}(s) + \bar{B}_{-s} \mathbf{1}_{\mathbb{R}_-}(s)$$

with $(B_s)_{s \geq 0}$ and $(\bar{B}_s)_{s \geq 0}$ two independent standard Brownian motions. Thereby, $\mathbf{1}_A$ denotes the indicator function of the set A.

In the following lemma we take the process introduced above as a sequence of function-valued random variables. This method originates from probability theory on Banach spaces.

Lemma 4.3. *The sequence of $C[0,1]$-valued random variables*

(25) $$W_k(s) := \tilde{X}_{k-1+s}, \ 0 \le s \le 1, k \in \mathbb{N}$$

is stationary and ergodic.

Proof. We denote by $\tilde{h}_0$ the restriction of the function $\tilde{h}$ to $[0,1]$. Since the function $\tilde{h}$ is periodic, we have the decomposition

$$
\begin{aligned}
W_k(t) &= \tilde{h}(k-1+t) + \sigma e^{-\alpha(k-1+t)} \int_{-\infty}^{k-1+t} e^{\alpha s} d\tilde{B}_s \\
&= \tilde{h}_0(t) + \sigma e^{-\alpha(k-1+t)} \int_{k-1}^{k-1+t} e^{\alpha s} d\tilde{B}_s + \sigma \sum_{l=-\infty}^{k-1} e^{-\alpha(k-1+t)} \int_{l-1}^{l} e^{\alpha s} d\tilde{B}_s.
\end{aligned}
$$

Making use of the time shifted Brownian motion

$$\tilde{B}_s^{(l)} := \tilde{B}_{s+l}$$

yields

$$
\begin{aligned}
W_k(t) &= \tilde{h}_0(t) + \sigma e^{-\alpha t} \int_0^t e^{\alpha s} d\tilde{B}_s^{(k)} + \sigma \sum_{l=-\infty}^{k-1} e^{-\alpha(k-l+t)} \int_0^1 e^{\alpha s} d\tilde{B}_s^{(l)} \\
&= \tilde{h}_0(t) + \sigma e^{-\alpha t} \int_0^t e^{\alpha s} d\tilde{B}_s^{(k)} + \sigma \sum_{j=-\infty}^{0} e^{-\alpha(1+t-j)} \int_0^1 e^{\alpha s} d\tilde{B}_s^{(j+k-1)}.
\end{aligned}
$$

Consequently, this can be written as

$$W_k(\cdot) = \tilde{h}_0(\cdot) + F_0(Y_k) + \sum_{l=-\infty}^{0} e^{\alpha(j-1)} F(Y_{j+k-1})$$

where we used the a.s. defined functionals

$$F_0 : C[0,1] \to C[0,1]; \ \omega \mapsto \left(t \mapsto \sigma e^{-\alpha t} \int_0^t e^{\alpha s} d\omega(s) \right),$$

$$F : C[0,1] \to C[0,1]; \ \omega \mapsto \left(t \mapsto \sigma e^{-\alpha t} \int_0^1 e^{\alpha s} d\omega(s) \right)$$

and the $C[0,1]$-valued random variables

$$Y_l = (s \mapsto \tilde{B}_s^{(l)} - \tilde{B}_0^{(l)}, \ 0 \le s < 1).$$

The series $(Y_l)_{l \in \mathbb{Z}}$ consists of independent and identically distributed random variables. This implies that $(W_k)_{k \in \mathbb{N}}$ is stationary and ergodic since each element of this sequence can be represented as a measurable function $G : (C[0,1])^{\mathbb{N}} \to C[0,1]$ of elements of the iid sequence $(Y_l)_{l \in \mathbb{Z}}$, i.e.

$$W_k = G(Y_k, Y_{k-1}, \dots).$$

$\square$

Lemma 4.4. *As $t \to \infty$ one has*

(26) $$\left| \tilde{X}_t - X_t \right| \to 0, \quad a.s.$$

Proof. We have

$$\left|\tilde{X}_t - X_t\right| \leq e^{-\alpha t}|X_0| + \left|\tilde{h}(t) - h(t)\right| + \left|\tilde{Z}_t - Z_t\right|$$

$$\leq e^{-\alpha t}|X_0| + e^{-\alpha t}\sum_{i=1}^p \mu_i \int_{-\infty}^0 e^{\alpha s}|\varphi_i(s)|ds + e^{-\alpha t}\int_{-\infty}^0 e^{\alpha s}d\tilde{B}_s.$$

Obviously, the three terms on the right side converge toward zero as $t \to \infty$. □

Proposition 4.5. *As $T \to \infty$, we have*

$$T Q_T^{-1} \to C, \quad \text{almost surely,}$$

where C is the matrix defined in (16).

Proof. We first consider the entries of the vector Λ_T, i.e. $\frac{1}{T}\int_0^T X_t\varphi_j(t)dt$. From Lemma 4.4 we may conclude that

$$\frac{1}{T}\int_0^T \tilde{X}_t\varphi_j(t)dt - \frac{1}{T}\int_0^T X_t\varphi_j(t)dt \to 0,$$

almost surely. Moreover, we get by the ergodic theorem

$$\frac{1}{T}\int_0^T \tilde{X}_t\varphi_j(t)dt = \frac{1}{T}\sum_{k=1}^T \int_{k-1}^k \tilde{X}_t\varphi_j(t)dt \longrightarrow E\left(\int_0^1 \tilde{X}_t\varphi_j(t)\,dt\right) = \int_0^1 \tilde{h}(t)\varphi_j(t)dt.$$

Thus we have established convergence of $\Lambda_{T,j}$, $1 \leq j \leq p$. In order to determine the limit of γ_T, it suffices to consider the term $\frac{1}{T}\int_0^T X_t^2 dt$. Due to representation (22) it holds that

$$(27) \qquad \limsup_{T\to\infty} \frac{1}{T}\int_0^T X_t dt = \limsup_{T\to\infty} \frac{1}{T}\int_0^T e^{-\alpha t}X_0 + h(t) + Z_t\,dt < \infty,$$

almost surely, since $h(t)$ is bounded and X_0 and Z_t are bounded almost surely, compare the proof of Lemma 4.6. It follows from (27) and Lemma 4.4 that

$$\frac{1}{T}\int_0^T \tilde{X}_t^2 dt - \frac{1}{T}\int_0^T X_t^2 dt = \frac{1}{T}\int_0^T (\tilde{X}_t + X_t)(\tilde{X}_t - X_t)dt \to 0.$$

Consequently, again by the ergodic theorem, we get

$$\begin{aligned}
\frac{1}{T}\int_0^T \tilde{X}_t^2 dt &= \frac{1}{T}\sum_{k=1}^T \int_{k-1}^k \tilde{X}_t^2 dt \\
&\to E\left(\int_0^1 \tilde{X}_t^2 dt\right) \\
&= E\left(\int_0^1 (\tilde{h}(t) + \tilde{Z}_t)^2 dt\right) \\
&= E\left(\int_0^1 (\tilde{h}(t))^2 dt + 2\int_0^1 \tilde{h}(t)\tilde{Z}_t dt + \int_0^1 \tilde{Z}_t^2 dt\right) \\
&= \int_0^1 (\tilde{h}(t))^2 dt + E(\tilde{Z}_1)^2 = \int_0^1 (\tilde{h}(t))^2 dt + \frac{\sigma^2}{2\alpha}.
\end{aligned}$$

By Bessel's inequality, we have

$$\sum_{i=1}^{p} \Lambda_i^2 \leq \int_0^1 (\tilde{h}(t))^2 dt$$

and thus $(\int_0^1 (\tilde{h}(t))^2 dt + E(\tilde{Z}_1)^2 - \sum_{i=1}^{p} \Lambda_i^2) \geq E(\tilde{Z}_1)^2 = \frac{\sigma^2}{2\alpha} > 0$. This proves the assertion of the proposition. $\qquad \square$

Lemma 4.6. *The sequence $\frac{1}{\sqrt{T}} R_T$ is bounded in L^2.*

Proof. Note that

$$\frac{1}{\sqrt{T}} \int_0^T \varphi_i(t) dB_t$$

is L_2-bounded because

$$(28) \qquad \mathrm{Var}\left(\frac{1}{\sqrt{T}} \int_0^T \varphi_i(t) dB_t \right) = \mathrm{Var}\left(\frac{1}{\sqrt{T}} \int_0^T \varphi_i(t) dB_t \right) = \frac{1}{T} \int_0^T \varphi_i^2(t) dt = 1.$$

For the last entry of $\frac{1}{\sqrt{T}} R_T$ we have to prove the boundedness of

$$\mathrm{Var}\left(\frac{1}{\sqrt{T}} \int_0^T X_t dB_t \right)$$
$$= \frac{1}{T} E\left(\int_0^T X_t^2 dt \right)$$
$$= \frac{1}{T} E\left(\int_0^T e^{-\alpha t} X_0 h(t) + e^{-\alpha t} X_0 Z_t + e^{-2\alpha t} X_0^2 + h(t) Z_t + h(t)^2 + Z_t^2 \, dt \right).$$

Since Z_t is a zero-mean random variable the expectation of the second and fourth term is zero. Moreover, the variance

$$E(Z_t^2) = \frac{\sigma^2}{2\alpha}(1 - e^{-2\alpha t})$$

is bounded and justifies

$$\sup_{T \geq 0} \frac{1}{T} E\left(\int_0^T Z_t^2 dt \right) < \infty.$$

Moreover, the function

$$h(t) = e^{-\alpha t} \sum_{i=1}^{p} \mu_i \int_0^t e^{\alpha s} \varphi_i(s) ds$$

is bounded due to the periodicity of $\varphi_i(t)$, $i = 1, \ldots, p$. The boundedness of $h(t)$ gives

$$\sup_{T \geq 0} \frac{1}{T} E\left(\int_0^T e^{-\alpha t} X_0 h(t) dt \right) < \infty$$

and

$$\sup_{T \geq 0} \frac{1}{T} E\left(\int_0^T h(t)^2 dt \right) < \infty.$$

This finishes the proof of the L^2-boundedness of $\frac{1}{\sqrt{T}} R_T$. $\qquad \square$

Proposition 4.7. *As $T \to \infty$, we have*

(29) $$\lim_{T \to \infty} \frac{1}{T} R_T = 0, \quad \text{almost surely.}$$

Proof. Observe that R_T is a martingale; thus we get by using Doob's maximal inequality for submartingales that for any $\epsilon > 0$

$$P(\sup_{2^k \leq T \leq 2^{k+1}} \frac{1}{T} |R_T| \geq \epsilon) \leq P(\sup_{2^k \leq T \leq 2^{k+1}} |R_T| \geq \epsilon 2^k)$$

$$\leq \frac{4}{\epsilon^2 2^{2k}} E |R_{2^{k+1}}|^2 = O(2^{-k}).$$

Applying the Borel-Cantelli theorem, we obtain $\limsup_{T \to \infty} \frac{1}{T} |R_T| \leq \epsilon$, almost surely, and thus we have shown that $R_T/T \to 0$. $\square$

Proof of Theorem 1. This follows directly from Proposition 4.5 and Proposition 4.7. $\square$

5. PROOF OF THEOREM 2

In the proof of Theorem 2 we use again the representation (20), i.e. $\hat{\theta}_{ML} - \theta = \sigma Q_T^{-1} R_T$, which we rewrite as

$$\sqrt{T} \frac{\hat{\theta}_{ML} - \theta}{\sigma} = \sqrt{T} Q_T^{-1} R_T = (T Q_T^{-1}) \frac{1}{\sqrt{T}} R_T.$$

By Proposition 4.5, $T Q_T^{-1}$ converges almost surely to the matrix C. Then, by Slutsky's theorem, Theorem 2 is an immediate corollary of the following proposition. Note that $\Sigma_0^{-1} = C$ by applying the same formula as in the proof of Lemma 3.1 to Σ_0. $\square$

Proposition 5.1. *Under the assumptions of Theorem 2, we have, as $T \to \infty$,*

$$\frac{1}{\sqrt{T}} R_T \xrightarrow{\mathcal{D}} N(0, \Sigma_0),$$

where the $(p+1) \times (p+1)$ matrix Σ_0 is defined as

$$\Sigma_0 = \begin{pmatrix} I_p & \Lambda \\ \Lambda^t & \omega \end{pmatrix}$$

where $\omega = \int_0^1 (\tilde{h}(t))^2 dt + \frac{\sigma^2}{2\alpha}$. The entries of the vector Λ are specified in (18).

The remaining part of this section is devoted to the proof of this proposition. Recall that

$$\frac{1}{\sqrt{T}} R_T = \begin{pmatrix} \frac{1}{\sqrt{T}} \int_0^T \varphi_1(t) dB_t \\ \vdots \\ \frac{1}{\sqrt{T}} \int_0^T \varphi_p(t) dB_t \\ -\frac{1}{\sqrt{T}} \int_0^T X_t dB_t \end{pmatrix}.$$

Since the basis functions $\varphi_1, \ldots, \varphi_p$ are orthonormal, the first p entries of the vector $\frac{1}{\sqrt{T}} R_T$ are independent, normally distributed random variables with mean zero and variance 1. Thus it remains to investigate the asymptotic distribution of the last entry

$$\frac{1}{\sqrt{T}} \int_0^T X_t dB_t,$$

and its joint distribution with the first p components.

By Lemma 4.2, the process $(X_t)_{t \geq 0}$ can be expressed as

$$X_t = e^{-\alpha t} X_0 + h(t) + \sigma e^{-\alpha t} \int_0^t e^{\alpha s} dB_s,$$

and thus we have

$$(30) \quad \frac{1}{\sqrt{T}} \int_0^T X_t dB_t = \frac{X_0}{\sqrt{T}} \int_0^T e^{-\alpha t} dB_t + \frac{1}{\sqrt{T}} \int_0^T h(t) dB_t + \sigma \frac{1}{\sqrt{T}} \int_0^T \int_0^t e^{\alpha(s-t)} dB_s dB_t.$$

The first term on the right hand side converges to 0 in probability, as

$$\text{Var}\left(\frac{1}{\sqrt{T}} \int_0^T e^{-\alpha t} dB_t\right) = \frac{1}{T} \int_0^T e^{-2\alpha t} dt \longrightarrow 0.$$

The second term is normally distributed with mean zero and variance

$$\frac{1}{T} \int_0^T (h(t))^2 dt \longrightarrow \int_0^1 (\tilde{h}(t))^2 dt.$$

The asymptotic distribution of the third term, as well as its joint distribution with any stochastic integral $\int_0^T \varphi(t)\, dB_t$, will be evaluated next.

Proposition 5.2. *Let $\varphi : [0, \infty) \to \mathbb{R}$ be an L_2-function, for which*

$$\sigma_\varphi^2 := \lim_{T \to \infty} \frac{1}{T} \int_0^T (\varphi(t))^2 dt$$

exists. Then, as $T \to \infty$,

$$\frac{1}{\sqrt{T}} \left(\int_0^T \int_0^t e^{\alpha(s-t)} dB_s dB_t, \int_0^T \varphi(t) dB_t \right) \xrightarrow{\mathcal{D}} N\left(0, \begin{pmatrix} \frac{1}{2\alpha} & 0 \\ 0 & \sigma_\varphi^2 \end{pmatrix}\right),$$

where $N(0, A)$ denotes a bivariate normal distribution with mean vector 0 and covariance matrix A.

Proof. Application of the time change formula for stochastic integrals twice, cf. Øksendal [8] (Theorem 8.5.7, p. 148), for $g(\tau) := T\tau$, $g'(\tau) = T$, results in

$$\frac{1}{\sqrt{T}} \int_0^T \int_0^t e^{\alpha(s-t)} dB_s dB_t = \int_0^1 \int_0^t e^{\alpha(s-Tt)} dB_s dB_t^{(T)}$$

$$= \sqrt{T} \int_0^1 \int_0^t e^{\alpha T(s-t)} dB_s^{(T)} dB_t^{(T)}$$

where $B_t^{(T)} = \frac{1}{\sqrt{T}} B_{Tt}$. Therefore, it is sufficient to study the asymptotic distribution of

$$\sqrt{T} \int_0^1 \int_0^t e^{\alpha T(s-t)} dW_s dW_t$$

where $(W_t)_{t \geq 0}$ denotes a Brownian motion with the same distribution as $(B_t^{(T)})_{t \geq 0}$. The symmetrization theorem for double Wiener integrals, cf. Kuo [4] (Theorem 9.2.8, p. 154), provides the identity

$$(31) \qquad \sqrt{T} \int_0^1 \int_0^t e^{\alpha T(s-t)} dW_s dW_t = \frac{\sqrt{T}}{2} \int_0^1 \int_0^1 e^{-\alpha T|s-t|} dW_s dW_t.$$

By Lemma 5.3 we obtain

$$(32) \qquad \sqrt{T} \int_0^1 \int_0^1 e^{-\alpha T |s-t|} dW_s dW_t \overset{\mathcal{D}}{=} \sum_{j=1}^\infty \lambda_{T,j}(\xi_{T,j}^2 - 1)$$

where $(\lambda_{T,j})_{j \in \mathbb{N}}$ is the set of eigenvalues of the integral operator with kernel $f_T(s,t) = \sqrt{T} e^{-\alpha T |s-t|}$ and where $\xi_{T,j} = \int_0^1 e_{T,j}(t) dW_t$. Here we denote by $e_{T,j}(t)$ the eigenfunction associated to the eigenvalue λ_i. By Lemma 5.4 the eigenvalues have the properties

$$\lim_{T \to \infty} \sum_{j=1}^\infty \lambda_{T,j}^2 = \frac{1}{\alpha}$$

$$\lim_{T \to \infty} \max_{j \geq 1} |\lambda_{T,j}| = 0.$$

Define $\xi_T := \frac{1}{\sqrt{T}} \int_0^T \varphi(t) dB_t$. Note that $\xi_T, \xi_{T,j}, j \geq 1$ are jointly normally distributed and that $(\xi_{T,j})_{j \geq 1}$ are iid standard normally distributed random variables. Projecting ξ_T onto the space spanned by the random variables $(\xi_{T,j})_{j \geq 1}$, we can write

$$\xi_T = \xi_{T,0} + \sum_{j=1}^\infty \alpha_{T,j} \xi_{T,j},$$

where $\xi_{T,0}$ is independent of $(\xi_{T,j})_{j \geq 1}$. Define $\sigma_T^2 := \frac{1}{T} \int_0^T \varphi^2(t) dt$ and $\sigma_{T,0}^2 := \mathrm{Var}(\xi_{T,0})$ and note that

$$\sigma_T^2 = \sigma_{T,0}^2 + \sum_{j=1}^\infty \alpha_{T,j}^2 \longrightarrow \sigma_\varphi^2.$$

We will now apply the Cramér-Wold device to prove convergence of the joint distribution of ξ_T and $\sum_{j=1}^\infty \lambda_{T,j}(\xi_{T,j}^2 - 1)$. Let $\mu_1, \mu_2 \in \mathbb{R}$; we will show that

$$\mu_1 \xi_T + \mu_2 \sum_{j=1}^\infty \lambda_{T,j}(\xi_{T,j}^2 - 1) \overset{\mathcal{D}}{\longrightarrow} N(0, \mu_1^2 \sigma_\varphi^2 + 2\mu_2^2 \frac{1}{\alpha}).$$

In order to do so, we compute the characteristic function of the left hand side and note that

$$\mu_1 \xi_T + \mu_2 \sum_{j=1}^\infty \lambda_{T,j}(\xi_{T,j}^2 - 1) = \mu_1 \xi_{T,0} + \sum_{j=1}^\infty (\mu_1 \alpha_{T,j} \xi_{T,j} + \mu_2 \lambda_{T,j}(\xi_{T,j}^2 - 1)).$$

If Z is standard normally distributed, the characteristic function of $a Z + b (Z^2 - 1)$ is given by

$$\psi(t) = (1 - 2itb)^{-1/2} \exp\left(-ibt - \frac{a^2 t^2}{2(1 - 2ibt)} \right).$$

Thus the characteristic function of $\mu_1 \xi_{T,0} + \sum_{j=1}^\infty (\mu_1 \alpha_{T,j} \xi_{T,j} + \mu_2 \lambda_{T,j}(\xi_{T,j}^2 - 1))$ equals

$$\psi_T(t) = e^{-\frac{1}{2} \mu_1^2 \sigma_{T,0}^2 t^2} \prod_{j=1}^\infty \left\{ (1 - 2i\mu_2 \lambda_{T,j} t)^{-1/2} \exp\left(-i\mu_2 \lambda_{T,j} t - \frac{(\mu_1 \alpha_{T,j})^2 t^2}{2(1 - 2i\mu_2 \lambda_{T,j} t)} \right) \right\}.$$

Taking logarithms and using Taylor expansion, we obtain

$$
\begin{aligned}
\log \psi_T(t) &= -\frac{1}{2}\mu_1^2\sigma_{T,0}^2 t^2 - \sum_{j=1}^{\infty}\left(\frac{1}{2}\log(1 - 2i\mu_2\lambda_{T,j}t) + i\mu_2\lambda_{T,j}t + \frac{\mu_1^2\alpha_{T,j}^2 t^2}{2(1 - 2i\mu_2\lambda_{T,j}t)}\right) \\
&= -\frac{1}{2}\mu_1^2\sigma_{T,0}^2 t^2 - \sum_{j=1}^{\infty}(\mu_2^2\lambda_{T,j}^2 + \frac{1}{2}\mu_1^2\alpha_{T,j}^2)t^2 + o(1) \\
&= -\frac{1}{2}\left(\mu_1^2\sigma_{T,0}^2 + \sum_{j=1}^{\infty}\mu_1^2\alpha_{T,j}^2 + 2\sum_{j=1}^{\infty}\mu_2^2\lambda_{T,j}^2\right)t^2 + o(1) \\
&\to -\frac{1}{2}\left(\mu_1^2\sigma_\varphi^2 + \mu_2^2\frac{2}{\alpha}\right)t^2.
\end{aligned}
$$

Note that the right hand side is the logarithm of the characteristic function of a normal distribution with mean 0 and variance $\mu_1^2\sigma_\varphi^2 + \mu_2^2\frac{2}{\alpha}$. $\qquad\square$

Lemma 5.3. *Let $f : [0,1]^2 \to \mathbb{R}$ be a symmetric continuous kernel and let $(\lambda_i)_{i\geq 1}$ and $(e_i(t))_{i\geq 1}$ denote the set of eigenvalues and corresponding eigenfunctions of the integral operator $G_f : L^2[0,1] \to L^2[0,1]$ with kernel f, i.e. $G_f g(x) = \int_0^1 g(y)f(x,y)dy$. Then*

$$
\int_0^1\int_0^1 f(s,t)dW_s dW_t = \sum_{i=1}^{\infty}\lambda_i(\xi_i^2 - 1),
$$

where

$$
\xi_i = \int_0^1 e_i(t)dW_t.
$$

The random variables $(\xi_i)_{i\in\mathbb{N}}$ are independent and standard normally distributed random variables.

Proof. Since the kernel f is continuous and symmetric the operator G_f is self-adjoint and compact. By Mercer's Theorem it holds that the kernel can be represented as

$$
(33) \qquad\qquad f(s,t) = \sum_{i=1}^{\infty}\lambda_i e_i(s)e_i(t)
$$

where λ_i and e_i, $i \in \mathbb{N}$, are the eigenvalues and eigenfunctions of the integral operator G_T, i.e.

$$
\int_0^1 f(s,t)e_i(s)ds = \lambda_i e_i(t), \qquad i \in \mathbb{N}.
$$

Moreover, it holds that the functions e_i, $i \in \mathbb{N}$, form an orthonormal basis of $L^2[0,1]$. Define the random variables

$$
\xi_i := \int_0^1 e_i(t)\,dW_t, \qquad i \in \mathbb{N},
$$

and note that $(\xi_i)_{i\geq 1}$ is an iid sequence of standard normally distributed random variables. It follows by (33) that

$$
\int_0^1\int_0^1 f(s,t)dW_s dW_t = \sum_{i=1}^{\infty}\lambda_i\int_0^1\int_0^1 e_i(s)e_i(t)dW_s dW_t = \sum_{i=1}^{\infty}\lambda_i(\xi_i^2 - 1).
$$

The last equality follows by Itô's Theorem which states that

$$\int_0^1 \int_0^1 e_i(s)e_i(t)dW_s dW_t = H_2\left(\int_0^1 e_i(t)\, dW_t\right),$$

where H_2 is the second Hermite polynomial, i.e. $H_2(x) = x^2 - 1$. □

We now consider the kernel $f_T : [0,1] \to \mathbb{R}$, defined by

$$(34) \qquad f_T(s,t) = \sqrt{T}e^{-\alpha T|s-t|}, \quad s,t \in [0,1].$$

Lemma 5.4. *Let* $(\lambda_{T,i})_{i\geq 1}$, *denote the set of eigenvalues of the integral operator with kernel (34). Then we have*

$$(35) \qquad \lim_{T\to\infty} \sum_{i=1}^{\infty} \lambda_{T,i}^2 = \frac{1}{\alpha}$$

$$(36) \qquad \lim_{T\to\infty} \max_{i\geq 1} |\lambda_{T,i}| = 0.$$

Proof. Note that the operator G_{f_T} is self-adjoint and bounded so that its eigenvalues are real-valued, and

$$\max_{i\geq 1} \lambda_{T,i}^2 = \sup_{g\in L^2[0,1]:\|g\|=1} \|G_{f_T}g\|^2$$

where $\|\cdot\|$ denotes the standard L_2-norm on $L^2[0,1]$. By an equality in Lax [6] (Theorem 2, p. 176) we get

$$\sup_{g\in L^2[0,1]:\|g\|=1} \|G_{f_T}g\| \leq \sup_{t\in[0,1]} \int_0^1 |f_T(s,t)|ds.$$

Simple integration yields

$$\int_0^1 |f_T(s,t)|ds = \int_0^1 \sqrt{T}e^{-\alpha T|s-t|}ds = \frac{1}{\alpha T}(2 - e^{-\alpha Tt} - e^{-\alpha T(1-t)}),$$

and thus it follows that

$$(37) \qquad \max_{i\in\mathbb{N}} \lambda_{T,i} \leq \frac{2}{\alpha T} \to 0 \quad \text{as } T \to \infty.$$

The assertion of Mercer's Theorem given in (33) and the orthonormality of the eigenvalues provide the identity

$$\sum_{i=1}^{\infty} \lambda_i^2 = \int_0^1 \int_0^1 f_T(s,t)^2 ds dt = \frac{1}{2\alpha}\left(2 + \frac{1}{T\alpha}(e^{-2\alpha T} - 1)\right)$$

where the last equality is obtained by simple integration. □

References

[1] J. P. N. Bishwal (2008): *Parameter Estimation in Stochastic Differential Equations.* Springer-Verlag, Berlin.

[2] F. R. Gantmacher (1986): *Matrizentheorie.* VEB Deutscher Verlag der Wissenschaften, Berlin.

[3] H. Geman (2005): *Commodities and Commodity Derivatives.* J. Wiley, Chichester.

[4] H. H. Kuo (2006): *Introduction to Stochastic Integrals.* Springer-Verlag.

[5] Y. A. Kutoyants (2004): *Statistical Inference for Ergodic Diffusion Processes.* Springer-Verlag, London.

[6] P. D. Lax (2002): *Functional Analysis.* John Wiley, New York.

[7] R. S. Lipster and A. N. Shiryayev (1977): *Statistics of Random Processes I.* Springer-Verlag, Berlin.

[8] B. Øksendal (2003): *Stochastic Differential Equations.* Springer-Verlag, Berlin.

[9] L. S. Ornstein and G. E. Uhlenbeck (1930): On the theory of Brownian Motion. *Physical Review* **36**, 823–841.

Fakultät für Mathematik, Ruhr-Universität Bochum, 44780 Bochum, Germany

E-mail address: herold.dehling@ruhr-uni-bochum.de

E-mail address: brice.franke@ruhr-uni-bochum.de

E-mail address: thomas.kott@ruhr-uni-bochum.de