ASYMPTOTIC INference FOR SPATIAL AUTOREGRESSION AND ORTHOGONALITY OF ORNSTEIN-UHLENBECK SHEETS

Mátyás Arató, Gyula Pap, Martien M.C. van Zuijlen

Report No. 9927 (July 1999)
Asymptotic inference for spatial autoregression and orthogonality of Ornstein–Uhlenbeck sheets

MÁTYÁS ARATÓ and GYULA PAP
Institute of Mathematics and Informatics, Lajos Kossuth University
Pf. 12, H-4010 Debrecen, Hungary

MARTIEN C. A. VAN ZUIJLEN
Department of Mathematics, University of Nijmegen
Toernooiveld 1, 6525 ED Nijmegen, The Netherlands

Abstract

The relationship between the discrete time nearly unit root doubly geometric spatial model and the continuous time Ornstein–Uhlenbeck random field is investigated. It turns out that the Ornstein–Uhlenbeck sheet can be considered as the continuous time counterpart of the doubly geometric model. However, it is shown that in contrast to the one-dimensional case and due to the orthogonality of the measures generated by the Ornstein–Uhlenbeck sheets with different parameters, this connection does not hold on the level of estimators.

Key words: Nearly unit root spatial autoregression, Wiener sheet, Ornstein–Uhlenbeck sheet, equivalence and orthogonality of measures, Radon-Nikodym derivative.

1 Introduction

Consider the AR(1) time series model

\[
\begin{align*}
X_k &= \alpha X_{k-1} + \varepsilon_k, & k = 1, 2, \ldots, n \\
X_0 &= 0.
\end{align*}
\]

(1)

It is well known that in the asymptotically stationary case when $|\alpha| < 1$, the sequence $(\hat{\alpha}_n)_{n \geq 1}$ of least squares estimators (LSE) of $\alpha$ is asymptotically normal (Mann and Wald [27], Anderson [1]).

---

This research has been supported by NWO, The Netherlands (Nederlandse Organisatie voor Wetenschappelijk Onderzoek).
In the unit root case when \( \alpha = 1 \), the sequence \((\hat{\alpha}_n)_{n \geq 1}\) is not asymptotically normal but

\[
n(\hat{\alpha}_n - 1) \xrightarrow{D} \frac{\int_0^1 W(t) dW(t)}{\int_0^1 W^2(t) dt},
\]

where \( \{W(t) : t \in [0, 1]\} \) denotes a standard Wiener process (White [36]).

In the nearly unit root model

\[
\begin{align*}
X_k^{(n)} &= \alpha^{(n)} X_{k-1}^{(n)} + \varepsilon_k^{(n)}, & k &= 1, 2, \ldots, n \\
X_0^{(n)} &= 0,
\end{align*}
\]

(2)

where

\[
\alpha^{(n)} = 1 - \frac{\gamma^{(n)}}{n}, \quad \gamma^{(n)} \to \gamma,
\]

the sequence \((\hat{\alpha}_n^{(n)})_{n \geq 1}\) of LSE of \(\alpha^{(n)}\) satisfies the following asymptotic result:

\[
n(\hat{\alpha}_n^{(n)} - \alpha^{(n)}) \xrightarrow{D} \frac{\int_0^1 Y(t) dW(t)}{\int_0^1 Y^2(t) dt},
\]

(3)

where \( \{Y(t) : t \in [0, 1]\} \) is a continuous time AR(1) process, i.e., an Ornstein-Uhlenbeck process, defined as the solution of the stochastic differential equation

\[
\begin{align*}
dY(t) &= -\gamma Y(t) dt + dW(t), & t &\in [0, 1] \\
Y(0) &= 0
\end{align*}
\]

(4)

(Bobkowki [14], Phillips [29], Chan and Wei [16]). The process \( \{Y(t) : t \in [0, 1]\} \) can be written as

\[
Y(t) = \int_0^t e^{\gamma(v-t)} dW(v).
\]

Arató, Kolmogorov and Siny [5] and Arató [2], [4] has drawn the attention to the connection between discrete and continuous time models. The result (3) can also be formulated as

\[
\hat{\gamma}_n^{(n)} \xrightarrow{D} \hat{\gamma},
\]

where \( \hat{\gamma}_n^{(n)} \) is the LSE of \(\gamma^{(n)}\) in the discrete time model (2) and \( \hat{\gamma} \) is the maximum likelihood estimator of \(\gamma\) in the continuous time model (4) (e.g., Arató [3]).

Similar results hold for the AR(p) model

\[
\begin{align*}
X_k &= \alpha_1 X_{k-1} + \cdots + \alpha_p X_{k-p} + \varepsilon_k, & k &= 1, 2, \ldots, n \\
X_0 &= X_{-1} = \ldots = X_{1-p} = 0.
\end{align*}
\]

(5)

In the asymptotically stationary case when all roots of the characteristic polynomial

\[
\phi(z) = 1 - \alpha_1 z - \ldots - \alpha_p z^p
\]
are outside the unit circle, it is well known that the sequence of the LSE of the coefficients is asymptotically normal (Mann and Wald [27], Anderson [1]). In the unit root case when all roots of the characteristic polynomial \( \varphi \) are on the unit circle, Chan and Wei [17] proved that the LSE \( \hat{\alpha}_n = (\hat{\alpha}_{1,n}, \ldots, \hat{\alpha}_{p,n})^\top \) of the coefficients \( \alpha = (\alpha_1, \ldots, \alpha_p)^\top \) is not asymptotically normal but with suitable normalizing matrices \( \{\delta_n\} \), the sequence \( \delta_n^{-1}(\hat{\alpha}_n - \alpha) \) converges in distribution. Moreover, they gave a representation of the limit distribution involving multiple stochastic integrals with respect to Wiener processes.

Jeganathan [21] considered the nearly unit root AR(\( p \)) model

\[
\begin{align*}
X_k^{(n)} &= \alpha_1^{(n)} X_{k-1}^{(n)} + \cdots + \alpha_p^{(n)} X_{k-p}^{(n)} + \varepsilon_k^{(n)}, \quad k = 1, 2, \ldots, n, \\
X_0^{(n)} &= X_{-1}^{(n)} = \ldots = X_{-p}^{(n)} = 0, \tag{6}
\end{align*}
\]

where the coefficients \( \alpha^{(n)} = (\alpha_1^{(n)}, \ldots, \alpha_p^{(n)})^\top \) are given by \( \alpha^{(n)} = \alpha + \delta_n \gamma_n \), where \( \gamma_n \rightarrow \gamma \), and \( \{\delta_n\} \) are the same normalizing matrices used by Chan and Wei [17]. Jeganathan [21] proved that the sequence \( \delta_n^{-1}(\hat{\alpha}_n^{(n)} - \alpha^{(n)}) \) converges in distribution and gave a complicated representation of the limit distribution. Van der Meer, Pap and Van Zuijlen [28] have given a simpler representation showing that there is an appropriate continuous time AR(\( p \)) model related to the discrete time model in the same way as in the case of the AR(1) model. In fact, it is easier to handle the LSE of the roots of the characteristic polynomials rather than the LSE of the coefficients.

The roots of the characteristic polynomials converge to unit roots in the following way:

\[
\varphi_n(z) = 1 - \alpha_1^{(n)} z - \cdots - \alpha_p^{(n)} z^p = \prod_{j=1}^q \prod_{k=1}^{r_j} (1 - e^{-\gamma_j^{(n)}/n + i\theta_j z}) \rightarrow \prod_{j=1}^q (1 - e^{i\theta_j z})^{r_j},
\]

where \( \theta_1, \ldots, \theta_q \in (-\pi, \pi] \) are all different and \( \gamma_j^{(n)} \rightarrow \gamma_j \). Then the sequence \( (\hat{\gamma}_{j,k}^{(n)})_{n \geq 1} \) of LSE of \( \gamma_j^{(n)} \) satisfies the following asymptotic result:

\[
\hat{\gamma}_{j,k}^{(n)} \xrightarrow{D} \gamma_{j,k}
\]

where \( \{\hat{\gamma}_{j,k}\} \) are the maximum likelihood estimators in the following continuous time AR(\( p \)) model:

\[
\begin{align*}
\prod_{j=1}^{r_j} (d + \gamma_{j,k}) Y(t) &= dW_j(t), \quad t \in [0,1], j = 1, \ldots, q, \\
Y_j(0) &= \ldots = Y_{j,(r_j-1)}(0) = 0, \quad j = 1, \ldots, q,
\end{align*}
\]

where \( \{W_j(t) : t \in [0,1]\}, \ j = 1, \ldots, q \) are independent standard Wiener processes, real-valued for \( \vartheta_j = 0 \) or \( \vartheta_j = \pi \), otherwise complex-valued.

Similar connections between certain discrete and continuous time multivariate autoregressive schemes has been shown in Kormos and Pap [23], Pap and Zuijlen [30], [31], Varga [35].
Now, consider the so-called doubly geometric spatial autoregressive model

\[
\begin{align*}
X_{k, \ell} &= \alpha_1 X_{k-1, \ell} + \alpha_2 X_{k-1, \ell-1} - \alpha_1 \alpha_2 X_{k-1, \ell-1} + \varepsilon_{k, \ell}, \quad k, \ell = 1, 2, \ldots, n \\
X_{0, \ell} &= X_{1, 0} = 0
\end{align*}
\]  

(7)

introduced by Martin [25]. The model has been used by Jain [20] in the study of image processing, by Martin [26], Cullis and Gleeson [18], Basu and Reinsel [10] in agricultural trials and by Tjostheim [33] in digital filtering.

In the asymptotically stationary case when \(|\alpha_1| < 1\) and \(|\alpha_2| < 1\), asymptotic normality of several estimators of \((\alpha_1, \alpha_2)\) has been shown (e.g., Tjostheim [32], Basu [7], Khalil [22], Basu and Reinsel [8], [9]).

In the unit root case when \(\alpha_1 = \alpha_2 = 1\), contrary to the AR(1) model, the sequence of Gauss–Newton estimators of \((\alpha_1, \alpha_2)\) has been shown to be again asymptotically normal (Bhattacharyya, Khalil and Richardson [11]). Bhattacharyya, Richardson and Franklin [12] have recently investigated the nearly unit root case

\[
\begin{align*}
X_{k, \ell}^{(n)} &= \alpha_1^{(n)} X_{k-1, \ell}^{(n)} + \alpha_2^{(n)} X_{k-1, \ell-1}^{(n)} - \alpha_1^{(n)} \alpha_2^{(n)} X_{k-1, \ell-1}^{(n)} + \varepsilon_{k, \ell}^{(n)}, \quad k, \ell = 1, 2, \ldots, n \\
X_{0, \ell}^{(n)} &= X_{1, 0}^{(n)} = 0,
\end{align*}
\]

(8)

where

\[
\alpha_j^{(n)} = 1 - \frac{\gamma_j^{(n)}}{n}, \quad \gamma_j^{(n)} \to \gamma_j, \quad j = 1, 2,
\]

and proved asymptotic normality of Gauss–Newton estimators of \((\alpha_1^{(n)}, \alpha_2^{(n)})\).

The purpose of this paper is to clarify the relation between the discrete time nearly unit root doubly geometric spatial model (8) and the continuous time Ornstein–Uhlenbeck random field

\[
Y(s, t) = \int_0^s \int_0^t e^{\gamma_1(u-s) + \gamma_2(v-t)} dW(u, v), \quad s, t \in [0, 1],
\]

(9)

where \(\{W(s, t) : s, t \in [0, 1]\}\) is a standard Wiener sheet. It turns out that although the Ornstein–Uhlenbeck sheet (9) can be considered again as the continuous time counterpart of the doubly geometric model (8), this connection does not hold on the level of estimators. In fact, there is no MLE of \((\gamma_1, \gamma_2)\) in (9), since the measures generated by Ornstein–Uhlenbeck sheets with different parameters are orthogonal with respect to each other (in other words, there is no Radon–Nikodym derivative). This phenomena is related to the fact that the parameters \((\gamma_1, \gamma_2)\) can be estimated in a strongly consistent way (compare with Ying [38], who gave strongly consistent estimators in the stationary case).

2 Convergence of doubly geometric spatial models to Ornstein–Uhlenbeck sheet

Define \(D([0, 1]^2 \to \mathbb{R}^k)\) to be the set of all functions \(f : [0, 1]^2 \to \mathbb{R}^k\) for which \(\lim_{u \uparrow s, v \uparrow t} f(u, v), \lim_{u \downarrow s, v \downarrow t} f(u, v), \lim_{u \uparrow s, v \downarrow t} f(u, v), \lim_{u \downarrow s, v \uparrow t} f(u, v)\) exist and \(\lim_{u \uparrow s, v \downarrow t} f(u, v) = \lim_{u \downarrow s, v \uparrow t} f(u, v) = \lim_{u \downarrow s, v \downarrow t} f(u, v) = \lim_{u \uparrow s, v \uparrow t} f(u, v) = \cdots\)
The set $D([0,1]^2 \to \mathbb{R}^k)$ can be endowed with a metric making it a complete and separable space (Bickel and Wichura [15]). Convergence of probability measures on $D([0,1]^2 \to \mathbb{R}^k)$ will be understood relative to this metric. We denote by $C([0,1]^2 \to \mathbb{R}^k)$ the space of continuous functions endowed with the supremum norm $\| \cdot \|_\infty$. For measurable mappings $\Phi, \Phi_n : D([0,1]^2 \to \mathbb{R}^k) \to D([0,1]^2 \to \mathbb{R}^\ell)$, $n = 1, 2, \ldots$ we shall write $\Phi_n \rightharpoonup \Phi$ if $\| \Phi_n(x) - \Phi(x) \|_\infty \to 0$ for all $x, x_n \in D([0,1]^2 \to \mathbb{R}^k)$ with $\|x_n - x\|_\infty \to 0$. We shall need the following simple lemma, which gives a sufficient condition for convergence to a functional of a continuous process.

**Lemma 1** Let $\Phi, \Phi_n : D([0,1]^2 \to \mathbb{R}^k) \to D([0,1]^2 \to \mathbb{R}^\ell)$, $n = 1, 2, \ldots$ be measurable mappings such that $\Phi_n \rightharpoonup \Phi$. Let $Z, Z_n$, $n = 1, 2, \ldots$ be stochastic processes with values in $D([0,1]^2 \to \mathbb{R}^k)$ such that $Z_n \rightharpoonup Z$ in $D([0,1]^2 \to \mathbb{R}^k)$ and almost all trajectories of $Z$ are continuous. Then $\Phi_n(Z_n) \rightharpoonup \Phi(Z)$ in $D([0,1]^2 \to \mathbb{R}^\ell)$.

**Proof.** According to Billingsley [13, Theorem 5.5], it suffices to prove that $\mathbb{P}(Z \in H) = 1$, where

$$H := \{ x \in D([0,1]^2 \to \mathbb{R}^k) : \Phi_n(x_n) \to \Phi(x) \text{ for } \forall x_n \in D([0,1]^2 \to \mathbb{R}^k) \text{ with } x_n \to x \},$$

where $x_n \to x$ denotes convergence in the space $D([0,1]^2 \to \mathbb{R}^k)$. Let $x \in C([0,1]^2 \to \mathbb{R}^k)$ and $x_n \in D([0,1]^2 \to \mathbb{R}^k)$, $n = 1, 2, \ldots$ with $x_n \to x$. Then we have $\|x_n - x\|_\infty \to 0$ (see Billingsley [13, (14.11)], Bickel and Wichura [15]), hence by the assumptions, $\|\Phi_n(x_n) - \Phi(x)\|_\infty \to 0$, which implies $\Phi_n(x_n) \to \Phi(x)$ in the space $D([0,1]^2 \to \mathbb{R}^k)$. Hence $H \subset C([0,1]^2 \to \mathbb{R}^k)$ and from the continuity of $Z$ we obtain $\mathbb{P}(Z \in H) = 1$.

Consider the discrete time doubly geometric nearly unit root spatial autoregressive model (8). The random step functions

$$Y^{(n)}(s,t) := \frac{1}{n} X^{(n)}_{\lfloor nt \rfloor}, \quad s, t \in [0,1],$$

$$M^{(n)}(s,t) := \frac{1}{n} \sum_{i=1}^{\lfloor nt \rfloor} \sum_{j=1}^{\lfloor i/n \rfloor} \xi_{i,j}^{(n)} \quad s, t \in [0,1]$$

can be considered as random elements in the space $D([0,1]^2 \to \mathbb{R})$. In the following proposition the relationship between asymptotic behavior of the sequence $(Y^{(n)})_{n \geq 1}$ (derived from the spatial model) and that of the “noise process” $(M^{(n)})_{n \geq 1}$ is clarified.

**Proposition 1** The following statements are equivalent:

(i) $M^{(n)} \rightharpoonup W$ in $D([0,1]^2 \to \mathbb{R})$,

(ii) $Y^{(n)} \rightharpoonup Y$ in $D([0,1]^2 \to \mathbb{R})$,
(iii) \((M^{(n)}, Y^{(n)}) \xrightarrow{P} (W, Y)\) in \(D([0, 1]^2 \rightarrow \mathbb{R}^2)\).

**Proof.** The purpose of the following discussion is to show that there exists measurable mappings \(\Phi_n : D([0, 1]^2 \rightarrow \mathbb{R}) \rightarrow D([0, 1]^2 \rightarrow \mathbb{R}), \quad n = 1, 2, \ldots\) such that \(Y^{(n)} = \Phi_n(M^{(n)})\). The doubly geometric spatial process can be expressed as a moving average:

\[
X_{k,\ell}^{(n)} = \sum_{i=1}^{k} \sum_{j=1}^{\ell} \left( \alpha_1^{(n)} \right)^{k-i} \left( \alpha_2^{(n)} \right)^{\ell-j} \varepsilon_{i,j}^{(n)}.
\]

Moreover, the coefficients in (8) can be written as follows:

\[
\alpha_j^{(n)} = e^{-\gamma_j^{(n)}/n}, \quad j = 1, 2,
\]

where \(\gamma_j^{(n)} \rightarrow \gamma_j, \quad j = 1, 2\). Hence

\[
Y^{(n)}(s, t) = \frac{1}{n} \sum_{i=1}^{[ns]} \sum_{j=1}^{[nt]} e^{-\gamma_1^{(n)} i/n - \gamma_2^{(n)} j/n} \sum_{i=1}^{[ns]} \sum_{j=1}^{[nt]} e^{\gamma_1^{(n)} i/n + \gamma_2^{(n)} j/n} \varepsilon_{i,j}^{(n)}.
\]

Using the expression

\[
e^{\gamma_1^{(n)} i/n} = e^{\gamma_1^{(n)} [ns]/n} + \sum_{k=i+1}^{[ns]} \left( e^{\gamma_1^{(n)} (k-1)/n} - e^{\gamma_1^{(n)} k/n} \right)
\]

and a similar expression for \(e^{\gamma_2^{(n)} j/n}\) with \(t\) instead of \(s\), we obtain

\[
Y^{(n)}(s, t) = \frac{1}{n} \sum_{i=1}^{[ns]} \sum_{j=1}^{[nt]} \varepsilon_{i,j}^{(n)} + \frac{1}{n} \sum_{i=1}^{[ns]} \sum_{j=1}^{[nt]} e^{\gamma_1^{(n)} i/n} \sum_{k=2}^{[ns]} \sum_{l=1}^{[nt]} \left( e^{\gamma_1^{(n)} (k-1)/n} - e^{\gamma_1^{(n)} k/n} \right)\varepsilon_{i,j}^{(n)}
\]

\[
+ \frac{1}{n} \sum_{i=1}^{[ns]} \sum_{j=1}^{[nt]} \sum_{\ell=1}^{[nt]} \left( e^{\gamma_2^{(n)} (\ell-1)/n} - e^{\gamma_2^{(n)} \ell/n} \right)\varepsilon_{i,j}^{(n)}
\]

\[
+ \frac{1}{n} \sum_{i=1}^{[ns]} \sum_{j=1}^{[nt]} \sum_{\ell=1}^{[nt]} \left( e^{\gamma_2^{(n)} (k-1)/n} - e^{\gamma_2^{(n)} k/n} \right)\varepsilon_{i,j}^{(n)}
\]

\[
+ \left( e^{\gamma_2^{(n)} (\ell-1)/n} - e^{\gamma_2^{(n)} \ell/n} \right)\varepsilon_{i,j}^{(n)}.
\]
Thus we can derive

\[
Y^{(n)}(s,t) = M^{(n)}(s,t) - \gamma_1^{(n)} e^{-\gamma_1^{(n)}|u|/n} \sum_{k=2}^{[ns]} M^{(n)} \left( \frac{s-\frac{k-1}{n}}{n} \right) \int_{(k-1)/n}^{k/n} e^{\gamma_1^{(n)}u} du \\
- \gamma_2^{(n)} e^{-\gamma_2^{(n)}|v|/n} \sum_{\ell=2}^{[nt]} M^{(n)} \left( \frac{t-\frac{\ell-1}{n}}{n} \right) \int_{(\ell-1)/n}^{\ell/n} e^{\gamma_2^{(n)}v} dv \\
+ \gamma_1^{(n)} \gamma_2^{(n)} e^{-\gamma_1^{(n)}|u|/n - \gamma_2^{(n)}|v|/n} \sum_{k=2}^{[ns]} \sum_{\ell=2}^{[nt]} \int_{(k-1)/n}^{k/n} \int_{(\ell-1)/n}^{\ell/n} e^{\gamma_1^{(n)}u+\gamma_2^{(n)}v} dv du,
\]

which implies \( Y^{(n)} = \Phi_n(M^{(n)}) \) with

\[
\Phi_n(x)(s,t) = x(s,t) - \gamma_1^{(n)} \int_0^{[ns]/n} e^{\gamma_1^{(n)}(u-\frac{u-1}{n})} x(u, \frac{u-1}{n}) du \\
- \gamma_2^{(n)} \int_0^{[nt]/n} e^{\gamma_2^{(n)}(v-\frac{v-1}{n})} x(\frac{v-1}{n}, v) dv \\
+ \gamma_1^{(n)} \gamma_2^{(n)} \int_0^{[ns]/n} \int_0^{[nt]/n} e^{\gamma_1^{(n)}(u-\frac{u-1}{n}) + \gamma_2^{(n)}(v-\frac{v-1}{n})} x(u,v) dv du.
\]

Clearly \( \Phi_n \to \Phi \) where

\[
\Phi(x)(s,t) = x(s,t) - \gamma_1 \int_0^s e^{\gamma_1(u-s)} x(u,t) du - \gamma_2 \int_0^t e^{\gamma_2(v-t)} x(s,v) dv \\
+ \gamma_1 \gamma_2 \int_0^s \int_0^t e^{\gamma_1(u-s) + \gamma_2(v-t)} x(u,v) dv du.
\]

Suppose \( M^{(n)} \overset{p}{\to} W \). Applying Lemma 1 we obtain that \( Y^{(n)} \overset{p}{\to} \bar{Y} \) in \( D([0,1]^2 \to \mathbb{R}) \), where \( \bar{Y} = \Phi(W) \). Now we show that \( \bar{Y} = Y \) \( \mathbb{P} \)-a.s. First we mention that \( Y(s,t) \) is the \( L^2 \)-limit of the sequence \( (Z_n(s,t))_{n \geq 1} \), where

\[
Z_n(s,t) := \sum_{i,j=1}^n e^{\gamma_1(i-1)/n - \gamma_2(j-1)/n} \Delta_1 \Delta_2 W \left( \frac{i-1}{n} s, \frac{j-1}{n} t \right) ,
\]

where \( \Delta_1 \) and \( \Delta_2 \) denotes the usual differences:

\[
\Delta_1 W \left( \frac{i}{n} s, \frac{j}{n} t \right) := W \left( \frac{i}{n} s, \frac{j}{n} t \right) - W \left( \frac{i-1}{n} s, \frac{j}{n} t \right) ,
\]

\[
\Delta_2 W \left( \frac{i}{n} s, \frac{j}{n} t \right) := W \left( \frac{i}{n} s, \frac{j+1}{n} t \right) - W \left( \frac{i}{n} s, \frac{j}{n} t \right) .
\]
Clearly
\[ Z_n(s, t) = e^{-\gamma_t s/n - \gamma_n t/n} W(s, t) + \left( e^{-\gamma_t s/n} - 1 \right) \sum_{i=1}^{n-1} e^{\gamma_i (s/n - i)} W \left( \frac{s}{n}, t \right) \\
+ \left( e^{-\gamma_n t/n} - 1 \right) \sum_{j=1}^{n-1} e^{\gamma_j (t/n + j)} W \left( s, \frac{t}{n} \right) \\
+ \left( e^{-\gamma_t s/n} - 1 \right) \left( e^{-\gamma_n t/n} - 1 \right) \sum_{i,j=1}^{n-1} e^{\gamma_i (s/n - i) + \gamma_j (t/n - j)} W \left( \frac{s}{n}, \frac{t}{n} \right). \]

Hence \( \bar{Y} = \Phi(W) \) is also an \( L^2 \) limit of the sequence \( (Z_n)_{n \geq 1} \), which implies \( \bar{Y} = Y \) \( \mathbb{P} \)-a.s., thus (i) \( \implies \) (ii) is proved. In a similar way we obtain (i) \( \implies \) (iii).

Next we start with
\[ M^{(n)}(s, t) = M^{(n)} \left( \frac{[ns]}{n}, \frac{[nt]}{n} \right) = \frac{1}{n} \sum_{k=1}^{[ns]} \sum_{t=1}^{[nt]} e_{k,t}^{(n)}. \]

Using the expressions
\[ 1 = e^{\tilde{\gamma}_1^{[n]}(k-[ns])/(s-[nt])} + \sum_{i=1}^{[ns]} e^{\tilde{\gamma}_i^{[n]}(k-[ns])/(s-[nt])}, \]
and
\[ 1 - e^{\tilde{\gamma}_1^{[n]}(k-[ns])/(s-[nt])}, \]
and a similar expression for \( 1 - e^{\tilde{\gamma}_2^{[n]}(k-[ns])/(s-[nt])}, \) we obtain
\[ M^{(n)}(s, t) = M^{(n)} \left( \frac{[ns]}{n}, \frac{[nt]}{n} \right) + \left( 1 - e^{\tilde{\gamma}_1^{[n]}/(s-[nt])} \right) \sum_{i=2}^{[ns]} M^{(n)} \left( \frac{[ns]}{n}, \frac{[nt]}{n} \right) \]
\[ + \left( 1 - e^{\tilde{\gamma}_2^{[n]}/(s-[nt])} \right) \sum_{j=2}^{[nt]} M^{(n)} \left( \frac{[ns]}{n}, \frac{[nt]}{n} \right) \]
\[ + \left( 1 - e^{\tilde{\gamma}_1^{[n]}/(s-[nt])} \right) \left( 1 - e^{\tilde{\gamma}_2^{[n]}/(s-[nt])} \right) \sum_{k=2}^{[ns]} \sum_{j=2}^{[nt]} M^{(n)} \left( \frac{[ns]}{n}, \frac{[nt]}{n} \right). \]
Hence,
\[
M^{(n)}(s,t) = Y^{(n)}(s,t) + n \left(1 - e^{-\gamma_i s/n}\right) \int_0^{[n]t/n} Y^{(n)}(u,t) \, du
\]
\[
+ n \left(1 - e^{-\gamma_2 s/n}\right) \int_0^{[n]t/n} Y^{(n)}(s,v) \, dv
\]
\[
+ n^2 \left(1 - e^{-\gamma_i s/n}\right) \left(1 - e^{-\gamma_2 s/n}\right) \int_0^{[n]t/n} \int_0^{[n]t/n} Y^{(n)}(u,v) \, du \, dv,
\]
which implies \( M^{(n)} = \Psi_n(Y^{(n)}) \), where \( \Psi_n \rightarrow \Psi \),

\[
\Psi(x)(s,t) = x(s,t) + \gamma_1 \int_0^s x(u,t) \, du + \gamma_2 \int_0^t x(s,v) \, dv + \gamma_1 \gamma_2 \int_0^s \int_0^t x(u,v) \, du \, dv.
\]
Suppose \( Y^{(n)} \xrightarrow{D} \bar{Y} \). Applying Lemma 1 we obtain that \( M^{(n)} \xrightarrow{D} \bar{W} \) in \( D([0,1]^2 \to \mathbb{R}) \), where \( \bar{W} = \Psi(Y) \). Now, we show that \( \bar{W} = W \) \( \mathbb{P} \)-a.s. Let us put
\[
Z_n(s,t) := \sum_{i,j=1}^n \bar{\Delta}_1 \bar{\Delta}_2 Y^i \left( \frac{s}{n}, \frac{t}{n} \right),
\]
where \( \bar{\Delta}_1 \) and \( \bar{\Delta}_2 \) denote the modified differences:
\[
\bar{\Delta}_1 W^i \left( \frac{s}{n}, \frac{t}{n} \right) := W^i \left( \frac{s}{n}, \frac{t}{n} \right) - e^{-\gamma_i s/n} W^i \left( \frac{s}{n}, \frac{t}{n} \right),
\]
\[
\bar{\Delta}_2 W^i \left( \frac{s}{n}, \frac{t}{n} \right) := W^i \left( \frac{s}{n}, \frac{t}{n} \right) - e^{-\gamma_2 t/n} W^i \left( \frac{s}{n}, \frac{t}{n} \right).
\]
Then, on the one hand
\[
Z_n(s,t) = \sum_{i,j=1}^n \left( \Delta_1 \Delta_2 Y^i \left( \frac{s}{n}, \frac{t}{n} \right) + \left(1 - e^{-\gamma_i s/n}\right) \Delta_2 Y^i \left( \frac{s-i}{n}, \frac{t}{n} \right)
\]
\[
+ \left(1 - e^{-\gamma_2 t/n}\right) \Delta_1 Y^i \left( \frac{s}{n}, \frac{t-j}{n} \right) + \left(1 - e^{-\gamma_2 t/n}\right) \left(1 - e^{-\gamma_2 t/n}\right) Y^i \left( \frac{s}{n}, \frac{t-j}{n} \right)
\]
\[
= Y(s,t) + \left(1 - e^{-\gamma_i s/n}\right) \sum_{i=1}^n Y^i \left( \frac{s-i}{n}, \frac{t}{n} \right) + \left(1 - e^{-\gamma_2 t/n}\right) \sum_{j=1}^n Y^i \left( \frac{s}{n}, \frac{t-j}{n} \right)
\]
\[
+ \sum_{i,j=1}^n \left(1 - e^{-\gamma_i s/n}\right) \left(1 - e^{-\gamma_2 t/n}\right) Y^i \left( \frac{s-i}{n}, \frac{t-j}{n} \right),
\]
hence, \( \bar{W} = \Psi(Y) \) is the \( L^2 \)-limit of the sequence \( (Z_n)_{n \geq 1} \). On the other hand,
\[
Z_n(s,t) = \sum_{i,j=1}^n \int_{s(i-1)/n}^{s(i)/n} \int_{t(j-1)/n}^{t(j)/n} e^{\gamma_1 (u-s)/n + \gamma_2 (v-t)/n} dW(u,v)
\]
9
and
\[
W(s, t) = \sum_{i,j=1}^{n} \int_{s(i-1)/n}^{s(i)/n} \int_{t(j-1)/n}^{t(j)/n} dW(u, v)
\]

imply that \( W \) is also an \( L^2 \)-limit of the sequence \( (Z_n)_{n \geq 1} \), since
\[
||Z_n(s, t) - W(s, t)||_{L^2}^2 = \sum_{i,j=1}^{n} \int_{s(i-1)/n}^{s(i)/n} \int_{t(j-1)/n}^{t(j)/n} \left(e^{\gamma_1(u-si/n) + \gamma_2(v-tj/n) - 1}\right)^2 dudv \to 0
\]
as \( n \to \infty \). Consequently, \( \hat{W} = W \) \( \mathbb{P} \)-a.s., hence (ii) \( \implies \) (i) is proved. In a similar way we obtain (iii) \( \implies \) (iii). \( \square \)

Applying the functional central limit theorem on the space \( D([0, 1]^2 \to \mathbb{R}) \) (see Bickel and Wichura [15, Theorem 5]) we obtain the following corollary.

**Corollary 1** Suppose that \( \{\xi_{k, \ell}^{(n)}\} \) are i.i.d., mean zero and variance 1. Then,
\[
(M^{(n)}, Y^{(n)}) \overset{D}{\to} (W, Y) \quad \text{in} \quad D([0, 1]^2 \to \mathbb{R}^2).
\]

We remark that Corollary 1 can be also proved directly by the help of the expression
\[
Y^{(n)}(s, t) = \frac{1}{n} \sum_{i=1}^{[ns]} \sum_{j=1}^{[nt]} e^{\gamma_1(i-\lfloor ns \rfloor)/n + \gamma_2(j-\lfloor nt \rfloor)/n} \xi_{i,j}^{(n)}
\]
and by proving finite dimensional convergence (applying the Central Limit Theorem) and showing tightness (applying the technique used in Bharatcharyya, Richardson and Franklin [12]).

## 3 Orthogonality of Ornstein–Uhlenbeck sheets

For \( \gamma \in \mathbb{R} \) and \( \sigma > 0 \), let
\[
Y_{\gamma, \sigma^2}(t) := \sigma \int_{0}^{t} e^{\gamma(v-t)} dW(v), \quad t \in [0, 1].
\]
The process \( \{Y_{\gamma, \sigma^2}(t) : t \in [0, 1]\} \) is an Ornstein–Uhlenbeck process with parameters \((\gamma, \sigma^2)\), and it is the solution of the following stochastic differential equation:
\[
\begin{aligned}
\left\{ 
\begin{array}{l}
    dY(t) = -\gamma Y(t) dt + \sigma dW(t), \quad t \in [0, 1], \\
    Y(0) = 0.
\end{array}
\right.
\end{aligned}
\tag{10}
\]

Let \( \mathbb{P}_{Y_{\gamma, \sigma^2}} \) be the measure generated by the process \( Y_{\gamma, \sigma^2} \) on \( C([0, 1] \to \mathbb{R}) \). For probability measures \( \mathbb{P}_1 \) and \( \mathbb{P}_2 \), equivalence or orthogonality will be denoted by \( \sim \) and by \( \perp \), respectively. The following dichotomy is well known (see Arató [3]):
\[
\begin{aligned}
\mathbb{P}_{Y_{\gamma, \sigma^2}} \sim \mathbb{P}_{Y_{\gamma, \sigma^2}} & \quad \text{if} \quad \sigma^2 = \sigma^2, \\
\mathbb{P}_{Y_{\gamma, \sigma^2}} \perp \mathbb{P}_{Y_{\gamma, \sigma^2}} & \quad \text{if} \quad \sigma^2 \neq \sigma^2.
\end{aligned}
\]
The orthogonality in case $\sigma^2 \neq \bar{\sigma}^2$ is based on the fact that
\[
\sum_{j=1}^{n} \left( Y_{\gamma, \sigma^2} \left( \frac{j}{n} \right) - Y_{\gamma, \bar{\sigma}^2} \left( \frac{j+1}{n} \right) \right)^2 \rightarrow \sigma^2 \quad \mathbb{P}_{Y_{\gamma, \sigma^2}} \text{-a.s.} \tag{11}
\]
which can be proved by the help of the following representation (derived from 10):
\[
Y_{\gamma, \sigma^2}(t) = -\gamma \int_{0}^{t} Y_{\gamma, \sigma^2}(v) \, dv + \sigma W(t)
\]
using the following almost sure limit results:
\[
\sum_{j=1}^{n} \left( W \left( \frac{j}{n} \right) - W \left( \frac{j+1}{n} \right) \right)^2 \rightarrow 1, \quad \sum_{j=1}^{n} \left( \int_{(j-1)/n}^{j/n} Y_{\gamma, \sigma^2}(v) \, dv \right)^2 \rightarrow 0,
\]
where the second convergence is a consequence of the almost sure continuity of the process $Y_{\gamma, \sigma^2}$. In other words, the parameter $\sigma^2$ can be estimated in a strongly consistent way (see also Ying [37] in the stationary case).

Now, we turn to the Ornstein–Uhlenbeck sheet with parameters $(\gamma_1, \gamma_2, \sigma^2)$ which is defined as follows:
\[
Y_{\gamma_1, \gamma_2, \sigma^2}(s, t) = \sigma \int_{0}^{s} \int_{0}^{t} e^{\gamma_1(u-s)+\gamma_2(v-t)} \, dW(u, v), \quad s, t \in [0, 1]^2,
\]
where $\gamma_1, \gamma_2 \in \mathbb{R}$, $\sigma > 0$. Let $\mathbb{P}_{Y_{\gamma_1, \gamma_2, \sigma^2}}$ be the measure generated by the process $Y_{\gamma_1, \gamma_2, \sigma^2}$ on $C([0, 1]^2 \to \mathbb{R})$. The following dichotomy can be proved (compare with Kurchenko [24]):

**Proposition 2**

\[
\begin{cases}
\mathbb{P}_{Y_{\gamma_1, \gamma_2, \sigma^2}} \sim \mathbb{P}_{Y_{\gamma_1, \gamma_2, \sigma^2}} & \text{if } (\gamma_1, \gamma_2, \sigma^2) = (\gamma_1^*, \gamma_2^*, \bar{\sigma}^2), \\
\mathbb{P}_{Y_{\gamma_1, \gamma_2, \sigma^2}} \not\sim \mathbb{P}_{Y_{\gamma_1, \gamma_2, \sigma^2}} & \text{if } (\gamma_1, \gamma_2, \sigma^2) \neq (\gamma_1^*, \gamma_2^*, \bar{\sigma}^2).
\end{cases}
\]

**Proof.** It is known that the convergence
\[
\sum_{j,k=1}^{n} \left( Y_{\gamma_1, \gamma_2, \sigma^2} \left( \frac{j}{n}, \frac{k}{n} \right) - Y_{\gamma_1, \gamma_2, \sigma^2} \left( \frac{j+1}{n}, \frac{k+1}{n} \right) \right)^2 \rightarrow \sigma^2
\]
holds $\mathbb{P}_{Y_{\gamma_1, \gamma_2, \sigma^2}}$-a.s. (see e.g., Deo and Wong [19]). It is easy to show that the process $\{Y_{\gamma_1, \gamma_2, \sigma^2}(s, 1) : s \in [0, 1]\}$ is a one-dimensional Ornstein–Uhlenbeck process with parameters
\[
\left( \gamma_1, \frac{1-e^{-2\gamma_2}}{2\gamma_2} \sigma^2 \right),
\]
so that we obtain from the one-dimensional consideration in (11) that
\[
\sum_{j=1}^{n} \left( Y_{\gamma_1, \gamma_2, \sigma^2} \left( \frac{j}{n}, 1 \right) - Y_{\gamma_1, \gamma_2, \sigma^2} \left( \frac{j+1}{n}, 1 \right) \right)^2 \rightarrow \frac{1-e^{-2\gamma_2}}{2\gamma_2} \sigma^2 \quad \mathbb{P}_{Y_{\gamma_1, \gamma_2, \sigma^2}} \text{-a.s.}
\]

11
Similarly,

$$\sum_{k=1}^{n} \left( Y_{\gamma_1, \gamma_2, \sigma^2} \left( 1, \frac{k}{n} \right) - Y_{\gamma_1, \gamma_2, \sigma^2} \left( 1, \frac{k-1}{n} \right) \right)^2 \xrightarrow{\text{a.s.}} \frac{1 - e^{-2\gamma_1}}{2\gamma_1} \sigma^2$$

The function

$$\gamma \mapsto \frac{1 - e^{-2\gamma}}{2\gamma}$$

is strictly monotone, hence at least one of the limits

$$\sigma^2, \quad \frac{1 - e^{-2\gamma_2}}{2\gamma_2} \sigma^2, \quad \frac{1 - e^{-2\gamma_1}}{2\gamma_1} \sigma^2$$

is different in case $$(\gamma_1, \gamma_2, \sigma^2) \neq (\gamma_1, \gamma_2, \sigma^2)$$). Since $${\mathbb{P}}_{Y_{\gamma_1, \gamma_2, \sigma^2}}$$ and $${\mathbb{P}}_{Y_{\gamma_1, \gamma_2, \sigma^2}}$$ are Gaussian measures, the Feldman–Hajek dichotomy theorem yields that they are either orthogonal or equivalent. In case $$(\gamma_1, \gamma_2, \sigma^2) \neq (\gamma_1, \gamma_2, \sigma^2)$$ they can not be equivalent, since at least one of the above sequences of random variables has different limit with respect to these measures.

In fact, it turned out that all the parameters $\gamma_1, \gamma_2, \sigma^2$ can be estimated in a strongly consistent way (see also Ying [38] in the stationary case).

We remark that if we consider the shifted processes

$$Z_{m, \gamma_1, \gamma_2, \sigma^2}(s, t) := Y_{\gamma_1, \gamma_2, \sigma^2}(s, t) + m$$

where $m \in \mathbb{R}$, then the measures generated by the processes $Z_{m, \gamma_1, \gamma_2, \sigma^2}$, $m \in \mathbb{R}$ (with fixed $\gamma_1, \gamma_2, \sigma^2$) are equivalent (see Baran, Pap, Zuijen [6]).

References


