A 2-Dimensional Functional Central Limit Theorem for Non-stationary Dependent Random Fields

Michael C. Tseng
Department of Economics
University of Central Florida
October 2019

Abstract

We obtain an elementary invariance principle for multi-dimensional Brownian sheet where the underlying random fields are not necessarily independent or stationary. Possible applications include unit-root tests for spatial as well as panel data models.

MSC classification: 60F17.
JEL classification: C10.
Keywords: Brownian sheet; m-dependent fields.

†Department of Economics, University of Central Florida. Email: michael.tseng@ucf.edu
1 Introduction

We obtain an elementary invariance principle for dependent random fields that does not require stationarity. Invariance principles under stationarity and mixing-type of conditions have also been considered by Berkes and Morrow (1981), Wang and Woodroofe (2013), and Volný and Wang (2014). Our current setting requires elements of the underlying field \( \{x_{i,j}\}_{i \geq 1, j \geq 1} \) to have the property that \( x_{i,j} \) and \( x_{i',j'} \) are uncorrelated whenever \( \max\{|i - i'|, |j - j'|\} \) is greater than some finite integer, say \( m \). Such random fields will be referred to as \( m \)-dependent. In dimension one, \( m \)-dependence generalizes finite-order moving-average time series whose the underlying innovation is a martingale difference sequence. For ease of exposition, we present our result for dimension two. Extending to higher dimensions is straightforward.

The rest of this paper is organized as follows. Section 2 states preliminary facts regarding tightness and asymptotic Gaussianity for certain random elements on the Skorohod space \( D([0,1]^2) \). Section 3 defines \( m \)-dependent random fields and establishes a maximal inequality. Section 4 proves the main result, an invariance principle for \( m \)-dependent random fields.

2 Preliminary Facts

2.1 \( D([0,1]^2) \)

We recall relevant properties of the Skorohod metric space \( D = D([0,1]^2) \) (see Bickel and Wichura (1971)). A step function on \([0,1]^2\) is an indicator function of the form \( 1_{E_1 \times E_2} \) where \( E_i \) is either a left-closed, right-open subset of \([0,1]\) or \( \{1\} \). As a set, \( D \) is the uniform closure of the vector space generated by step functions. Let \( \pi \) denote a rectangular partition of the form \( \{\pi_{ij} = [\pi_{1i}, \pi_{1i+1}] \times [\pi_{2j}, \pi_{2j+1}], 1 \leq i \leq n_1, 1 \leq j \leq n_2\} \) of \([0,1]^2\) where the south and west (resp. north and east) edges of each rectangle are closed (resp. open). Let \( G_\delta \) denote the family of all rectangular partition \( \pi \) such that where \( \max\{|\pi_{1i} - \pi_{1i+1}|, |\pi_{2j} - \pi_{2j+1}|\} > \delta \) for all \( i, j \). For \( x \in D \) and \( 0 < \delta < 1 \), define

\[
w'(x, \delta) = \inf_{\pi \in G_\delta} \max_{s,t \in \pi_{ij}} |x(s) - x(t)|.
\]

As in the case of \( D[0,1] \), \( x \) lies in \( D \) if and only if \( \lim_{\delta \to 0} w'(x, \delta) \to 0 \). A related quantity, the modulus of continuity, is defined by

\[
w(x, \delta) = \sup_{\|s-t\|_\infty < \delta} |x(s) - x(t)|.
\]
It is clear that \( x \) lies in \( C = C([0,1]^2) \) if and only if \( \lim_{\delta \to 0} w(x, \delta) \to 0 \). In general, \( w'(x, \delta) \leq w(x, 2\delta) \). If \( x \in C \), then \( w(x, \delta) \leq 2w'(x, \delta) \).

The Skorohod topology on \( D \) is defined as follows. Let \( \Lambda \) denote the class of maps \( \lambda : [0, 1]^2 \to [0,1]^2 \) such that \( \lambda(t_1, t_2) = (\lambda_1(t_1), \lambda_2(t_2)) \) where \( \lambda_1, \lambda_2 : [0,1] \to [0,1] \) are strictly increasing, continuous, \( \lambda_1(0) = \lambda_2(0) = 0 \), \( \lambda_1(1) = \lambda_2(1) = 1 \). For \( x, y \in D \), the Skorohod metric \( d(x, y) \) is defined by

\[
d(x, y) = \inf\{\epsilon : \sup_t \|\lambda(t) - t\|_\infty \leq \epsilon, \sup_t |x(t) - y(\lambda(t))| \leq \epsilon\}.
\]

The Skorohod topology coincides with the uniform topology on \( C \subset D \).

\( D \) is not complete under the Skorohod metric \( d \). The Skorohod topology is also induced by another metric, under which \( D \) is complete. If one restricts \( \lambda \) to maps satisfying

\[
\max_{i=1,2} \|\lambda_i\| < \infty, \|\lambda_i\| \equiv \sup_{t_i > s_i} \log \frac{\lambda_i t_i}{t_i - s_i},
\]

the Billingsley’s metric \( d_0 \) is defined by

\[
d_0(x, y) = \inf\{\epsilon : \|\lambda\| \leq \epsilon, \sup_t |x(t) - y(\lambda(t))| \leq \epsilon\}.
\]

From the Taylor expansion estimate

\[
\log 1 - 2\epsilon \leq -\epsilon \leq \log \frac{\lambda_i t_i}{t_i} \leq \epsilon \leq \log 1 + 2\epsilon,
\]

it follows immediately the definition that \( d(x, y) \leq 2d_0(x, y) \). On the other hand, if \( d(x, y) < \delta^2 \) for \( 0 < \delta < \frac{1}{4} \), then

\[
d_0(x, y) \leq 4\delta + w'(x, \delta).
\]

Therefore the metric \( d_0 \) is equivalent to \( d \).

The argument for completeness of \( D \) under \( d_0 \) is the same as that in Theorem 14.2 on p115 of [Billingsley (1968)](#), applied to each coordinate. The difference between \( d_0 \) and \( d \) is the extra rigidity requirement on \( \lambda \), which implies that certain sequences which are Cauchy under \( d \) are not Cauchy under \( d_0 \). Next we have a characterization of compactness in \( D \) of Arzelà-Ascoli type.

**Proposition 1.** \( A \subset D \) is precompact if and only if the following conditions hold:

(i) \( \sup_{x \in A} \sup_t |x(t)| < \infty \).

\[\text{This can be proved by applying the same argument as in Lemma 2 on p113 of [Billingsley (1968)](#) to each coordinate.}\]
\[(ii) \lim_{\delta \to 0} \sup_{x \in A} w'(x, \delta) = 0.\]

This can be proved by applying the same argument as Theorem 14.3 of Billingsley (1968). The Borel \(\sigma\)-algebra \(\mathcal{D}\) on \(D\) is generated by the coordinate maps \(\pi_{t_1, t_2, \ldots, t_k}(x) = (x(t_1), \ldots, x(t_k)), t_1, \ldots, t_k \in [0, 1]^2\). Proposition \(\ref{prop2}\) immediately leads to the following characterization of tightness on \(D\).

**Proposition 2.** A sequence of probability measures \(\{P_n\}\) on \((D, \mathcal{D})\) is tight if and only if the following conditions hold:

(i) The family \(\{P_n\}\) pushed forward to the real line by \(\|\cdot\|_{\infty}\) is tight, i.e. for all \(\eta > 0\), there exists \(a > 0\) such that

\[P_n(x : \|x\|_{\infty} > a) < \eta, \forall n \geq 1.\]

(ii) For all \(\epsilon > 0\) and \(\eta > 0\), there exists a \(\delta \in (0, 1)\) and \(n_0\) such that

\[P_n(x : w'(x, \delta) \geq \epsilon) \leq \eta, \forall n \geq n_0.\]

See Theorems 8.2 and 15.2 of Billingsley (1968); the same argument as Theorem 8.2 goes through. Necessity is immediate consequence of Prop. \(\ref{prop2}\). Necessity implies that, in condition (ii), \(n_0\) can be taken to be equal 1 without loss of generality, since any finite set of probability measures is tight. From this strengthened condition (ii), sufficiency follows. We are interested in (limit) probability measures whose support lie in \(C = C([0, 1]^2)\). The following is the two dimensional analogue of Theorem 15.5 of Billingsley (1968). It provides sufficient conditions that guarantee tightness as well as any limit measure having support in \(C\).

**Proposition 3.** A sequence of probability measures \(\{P_n\}\) on \((D, \mathcal{D})\) is tight if the following conditions hold:

(i) For all \(\eta > 0\), there exists \(a > 0\) such that

\[P_n(x : x(0, 0) > a) < \eta, \forall n \geq 1.\]

(ii) For all \(\epsilon > 0\) and \(\eta > 0\), there exists a \(\delta \in (0, 1)\) and \(n_0\) such that

\[P_n(x : w(x, \delta) \geq \epsilon) \leq \eta, \forall n \geq n_0.\]

Moreover, for any \(P\) is a weak limit point of \(\{P_n\}\), \(P(C) = 1\).
Proof. Since \( w'(x, \delta) \leq w(x, 2\delta) \), condition (ii) of Proposition 2 follows from condition (ii).

Let \( \epsilon > 0 \) and \( \eta > 0 \), choose a \( \delta \in (0, 1) \) such that

\[
P_n(x : w(x, \delta) \leq \epsilon) \geq 1 - \frac{\eta}{2}, \ \forall n \geq 1,
\]

and \( a > 0 \) such that

\[
P_n(x : x(0, 0) \leq a) \geq 1 - \frac{\eta}{2}, \ \forall n \geq 1.
\]

We have

\[
P_n((x : w(x, \delta) \leq \epsilon) \cap (x : x(0, 0) \leq a)) \geq 1 - \eta,
\]

and (partitioning \([0, 1]^2 \) into \( \delta \times \delta \) regular grids),

\[
(x : w(x, \delta) \leq \epsilon) \cap (x : x(0, 0) \leq a) \subset (x : \|x\|_\infty \leq a + \epsilon \cdot \frac{1}{\delta^2}).
\]

So condition (i) of Proposition 2 holds. This proves tightness.

If \( w(y, \delta) \geq 2\epsilon \), then \( y \) is interior to \( w(x, \delta) \geq \epsilon \). By characterization of weak convergence, a subsequence \( P_{n'} \Rightarrow P \) therefore implies

\[
P(y : w(y, \frac{\delta}{2}) \geq 2\epsilon) \leq \lim \inf P_{n'}(x : w(x, \delta) \geq \epsilon).
\]

Let \( \epsilon_k \to 0 \), condition (ii) implies that there exists a sequence \( \delta_k \to 0 \) such that

\[
P(y : w(y, \delta_k) \leq \epsilon_k) \to 1.
\]

Let \( A \) be the set defined in the proposition.

Following Theorem 8.3 of Billingsley (1968), we obtain a sufficient condition for condition (ii) of Proposition 3 that can be applied to the random elements we will consider.

Proposition 4. If for all \( \epsilon > 0 \) and \( \eta > 0 \), there exists a \( \delta \in (0, 1) \) and \( n_0 \) such that for all \((t_1, t_2) \in [0, 1]^2\),

\[
\frac{1}{\delta^2} P_n(x : \sup_{t_1 \leq s_1 \leq t_1 + \delta, t_2 \leq s_2 \leq t_2 + \delta} |x(s_1, s_2) - x(t_1, t_2)| \geq \epsilon) \leq \eta, \ \forall n \geq n_0,
\]

then Condition (ii) of Proposition 3 holds.

A proof can be obtained extending that for Theorem 8.3 of Billingsley (1968) to the two dimensional setting. We now apply the above results to random elements. Let \( \{\xi_{i,j}\} \) be
a field of random variables defined on a common probability space \((\Omega, \mathcal{F}, P)\). Define the double partial sum by

\[
S_{n_1, n_2} = \sum_{i,j=1}^{n_1, n_2} \xi_{i,j}.
\]

Consider the random element in \(D\) defined by

\[
X_n(t_1, t_2) = \frac{1}{n} S_{[nt_1], [nt_2]}.
\]

**Proposition 5.** The sequence \(\{X_n\}\) is tight on \((D, \mathcal{D})\) if for all \(\epsilon > 0\), there exists \(\lambda > 0\) and \(n_0\) such that for all \(n \geq n_0\) and all \(k_1, k_2 \geq 1\),

\[
P(\max_{i \leq n, j \leq n} |S_{k_1+i, k_2+j} - S_{k_1, k_2}| \geq \lambda n) \leq \frac{\epsilon}{\lambda^2}.
\]

Moreover, if \(P\) is the weak limit of a subsequence of \(\{X_n\}\), \(P(C) = 1\).

Condition (i) of \(\ref{condition}\) holds by definition of \(X_n\). Condition (ii) of \(\ref{condition}\) applied to \(X_n\) says that

\[
\frac{1}{\delta^2} P(\max_{i \leq n\delta, j \leq n\delta} \frac{|S_{k_1+i, k_2+j} - S_{k_1, k_2}|}{n} \geq \epsilon) \leq \eta
\]

for all \((k_1, k_2)\) and \(n\) uniformly large. Put \(m = n\delta\), the expression becomes

\[
P(\max_{i \leq m, j \leq m} |S_{k_1+i, k_2+j} - S_{k_1, k_2}| \geq \frac{m}{\delta} \epsilon) \leq \delta^2 \eta.
\]

Let \(\lambda = \frac{\epsilon}{\delta}\), then the expression becomes

\[
P(\max_{i \leq m, j \leq m} |S_{k_1+i, k_2+j} - S_{k_1, k_2}| \geq m \lambda) \leq \frac{\epsilon^2}{\lambda^2} \eta.
\]

\(\epsilon^2 \eta\) can be collapsed into \(\epsilon\) and we arrive at

\[
P(\max_{i \leq m, j \leq m} \frac{|S_{k_1+i, k_2+j} - S_{k_1, k_2}|}{m} \geq \lambda) \leq \frac{\epsilon}{\lambda^2}.
\]

which is what appears in Proposition 5.

**Corollary 1.** The sequence \(\{X_n\}\) is tight and any limiting measure is supported on \(C\) if the family

\[
\{ \max_{i \leq n, j \leq n} \frac{|S_{k_1+i, k_2+j} - S_{k_1, k_2}|^2}{n^2}, n \geq 1, k_1, k_2 \geq 1 \}
\]

is uniformly integrable.
2.2 Asymptotic Gaussianity

We now give conditions under which a sequence of random elements in $D$ converges to the Brownian sheet in finite dimensional distributions. The condition below extends those in Section 19 of Billingsley [1968]. For $(s_1, t_1] \times (s_2, t_2] \subset [0, 1]^2$, define

$$X(\Delta(s_1, t_1] \times (s_2, t_2]) = X(t_1, t_2) - X(s_1, t_2) - X(t_1, s_2) + X(s_1, s_2).$$

Also, for $(s, t) \in (0, 1)^2$, define

$$I(s, t) = \{(s', t') \in (0, 1)^2 : s' < s \text{ or } t' < t\}.$$

The following conditions on infinitesimal moments are sufficient for Gaussianity in the limit.

Condition 1\(^a\).

Let $t \in (0, 1)$, $(\hat{s}, \hat{t}] \subset [0, 1]$, $t_1, \ldots, t_k \in I(t, \hat{s})$. For all $u_1, \ldots, u_k \in \mathbb{R}$,

$$\lim_{n \to 0} \frac{1}{h} |E[e^{i \sum_{j=1}^{k} u_j X_n(t_j)} X_n(\Delta(\hat{s}, \hat{t}] \times (t, t+h))]| = 0,$$

$$\lim_{n \to 0} \frac{1}{h} |E[e^{i \sum_{j=1}^{k} u_j X_n(t_j)} (X_n^2(\Delta(\hat{s}, \hat{t}] \times (t, t+h))] - h \cdot (\hat{t} - \hat{s}))]| = 0,$$

Condition 2\(^a\).

$$\lim_{\alpha \to \infty} \sup_{t \in [0, 1]^2} \lim_{n \to \infty} E[X_n(t) \mathbb{1}_{X_n(t) \geq \alpha}] \to 0.$$ 

Condition 3\(^a\).

Let $t \in (0, 1)$, $(\hat{s}, \hat{t}] \subset [0, 1]$, 

$$\lim_{\alpha \to \infty} \limsup_{h \to 0} \limsup_{n \to 0} \frac{1}{h} E[X_n^2(\Delta(\hat{s}, \hat{t}] \times (t, t+h))] \mathbb{1}_{X_n^2(\Delta(\hat{s}, \hat{t}] \times (t, t+h)) \geq \alpha h}] \to 0.$$

The following proposition can be shown by induction, on $k$, and the Cramer-Wold device.

**Proposition 6.** Let $W$ denote the Brownian sheet on $D$. Suppose a sequence of random elements $\{X_n\}$ satisfies Conditions 1\(^a\), 2\(^a\), and 3\(^a\), then

$$P_n \circ \pi_{-1}^{t_1, t_2, \ldots, t_k} \Rightarrow W \circ \pi_{-1}^{t_1, t_2, \ldots, t_k}$$

for all $t_1, \ldots, t_k \in [0, 1]^2$.

The lemma below summarizes the fact that, under weak convergence of finite dimensional law and tightness, Brownian sheet is the unique weak limit.
Lemma 1. Suppose a sequence of random elements $X_n, n = 1, 2, \cdots \subset D$ is tight and $P(X \in C) = 1$ for any weak limit $X$. If $\{X_n\}$ satisfies Conditions 1°a, 2°a, and 3°a, then $\{X_n\}$ converges weakly to the Brownian sheet.

3 $m$-dependent Fields

On $\mathbb{N}^2$, consider the component-wise order defined by $(i, j) \leq (i', j')$ if $i \leq i'$ and $j \leq j'$.

Definition 1. (Walsh (1986), p336)

Let $(\Omega, \mathcal{F}, P)$ be a probability space, a family $\sigma$-subalgebras $\{\mathcal{F}_{i,j}\}_{i \geq 1, j \geq 1}$ is a (2-dimensional) filtration if $\mathcal{F}_{i,j} \subset \mathcal{F}_{i',j'}$ whenever $(i, j) \leq (i', j')$.

A family of random variables $\{x_{i,j}\}_{i \geq 1, j \geq 1} \subset L^1(\Omega, \mathcal{F}, P)$ is a martingale with respect to filtration $\{\mathcal{F}_{i,j}\}_{i \geq 1, j \geq 1}$ if the following holds:

(i) $x_{i,j}$ is $\mathcal{F}_{i,j}$-measurable for all $(i, j)$,

(ii) for all $(i, j) \geq (i', j')$, $E[x_{i,j}|\mathcal{F}_{i',j'}] = x_{i',j'}$.

We will restrict to $L^2$-martingales. The conditional expectation operator $E[\cdot | \mathcal{F}_{i,j}]$ will be denoted by $E_{i,j}[\cdot]$. For a random variable $y \in L^2$, define the $L^2$-increment, with respect to a given filtration $\{\mathcal{F}_{i,j}\}$,

$$
\Delta y(i, j) = E_{i,j}[y] - E_{i-1,j}[y] - E_{i,j-1}[y] + E_{i-1,j-1}[y].
$$

Given an $L^2$-martingale $\{x_{i,j}\}$, for each element $x_{i,j}$ define

$$
\hat{x}_{i,j}^{i',j'} = \Delta x_{i,j}(i', j').
$$

It follows from the definition of a martingale that

$$
x_{i,j} = \sum_{(i',j') \leq (i,j)} \hat{x}_{i,j}^{i',j'}, \text{ a.s.} \tag{1}
$$

Next we define $m$-dependent random fields.

Definition 2. A family of random variables $\{x_{i,j}\}_{i \geq 1, j \geq 1} \subset L^2(\Omega, \mathcal{F}, P)$ is $m$-dependent with respect to filtration $\{\mathcal{F}_{i,j}\}_{i \geq 1, j \geq 1}$ if the following holds:

(i) $x_{i,j}$ is $\mathcal{F}_{i+m,j+m}$-measurable for all $(i, j)$,

(ii) For any $k > m$, $E_{i-k,j}[x_{i,j}] = E_{i,j=k}[x_{i,j}] = E_{i-k,j-k}[x_{i,j}] = 0$. 

7
It follows from the definition that, if \( \{x_{i,j}\} \) is a \( m \)-dependent field,

\[
x_{i,j} = \sum_{-m \leq k_1 \leq m, -m \leq k_2 \leq m} \hat{x}_{i,j}^{k_1, k_2}, \text{ a.s.}
\]  \( (2) \)

**Lemma 2.** Let \( \{x_{i,j}\}_{i \geq 1, j \geq 1} \) be \( m \)-dependent with respect to filtration \( \mathcal{F}_{i,j} \). For any \(-m \leq k_1, k_2 \leq m\), define \( \sigma \)-subalgebras \( \mathcal{G}_{i,j} = \mathcal{F}_{i-k_1, j-k_2} \). Then

\[
Y_{(i,j)}^{(k_1, k_2)} = \sum_{1 \leq i' \leq i, 1 \leq j' \leq j} \hat{x}_{i', j'}^{k_1, k_2}
\]

is a martingale with respect to \( \{\mathcal{G}_{i,j}\} \).

**Proof.** For any \((i', j')\), \( \hat{x}_{i', j'}^{k_1, k_2} \) is \( \mathcal{G}_{i', j'} \)-measurable by definition. Since \( \{\mathcal{G}_{i,j}\} \) is a filtration, \( Y_{(i,j)}^{(k_1, k_2)} = \mathcal{G}_{i,j} \)-measurable. Let \((i', j') \geq (i, j)\),

\[
E[\hat{x}_{i', j'}^{k_1, k_2} | \mathcal{G}_{i,j}] = E[\hat{x}_{i', j'}^{k_1, k_2} | \mathcal{F}_{i-k_1, j-k_2}] = 0.
\]

This proves the martingale property. \( \square \)

The following theorem due to Walsh extends the \( L^p \)-inequality to 2-dimensional martingales.

**Theorem 3.1.** (Walsh (1986), p351) Let \( \{x_{i,j}\}_{i \geq 1, j \geq 1} \) be a martingale, \( p > 1 \), and \( \frac{1}{p} + \frac{1}{q} = 1 \). Define

\[
S_{n_1, n_2} = \sum_{i,j=1}^{n_1, n_2} x_{i,j}, \quad S^*_{n_1, n_2} = \sup_{(i,j) \leq (n_1, n_2)} |S_{i,j}|
\]

Then

\[
E[(S^*_{n_1, n_2})^p] \leq \left( \frac{p}{p-1} \right)^{2p} \sup_{(i,j) \leq (n_1, n_2)} E[|S_{i,j}|^p].
\]

For our purposes, we need to extend Theorem 3.1 to \( m \)-dependent fields.

**Lemma 3.** Let \( \{z_{i,j}\}_{1 \leq i \leq m, 1 \leq j \leq m} \) be complex numbers and \( p > 1 \), then

\[
|\sum_{-m \leq i,j \leq m} z_{i,j}|^p \leq (2m+1)^{2(p-1)} \sum_{-m \leq i,j \leq m} |z_{i,j}|^p.
\]

**Proposition 7.** Let \( \{x_{i,j}\}_{i \geq 1, j \geq 1} \) be an \( m \)-dependent field, \( p > 1 \), \( \frac{1}{p} + \frac{1}{q} = 1 \), and

\[
Y_{(i,j)}^{(k_1, k_2)} = \sum_{1 \leq i' \leq i, 1 \leq j' \leq j} \hat{x}_{i', j'}^{k_1, k_2}.
\]
Then
\[ E[(S_{n_1,n_2}^*)^p] \leq q^{2p} \cdot (2m + 1)^{2(p-1)} \sum_{-m \leq k_1, k_2 \leq m} \max_{(i,j) \leq (n_1,n_2)} E[|Y^{(k_1,k_2)}_{i,j}|^p]. \]

**Proof.** By Equation 2

\[ x_{i,j} = \sum_{-m \leq k_1 \leq m, -m \leq k_2 \leq m} \hat{x}_{i,j}^{i-k_1,j-k_2}, \text{ a.s.} \]

So, re-arranging the finite sum gives

\[
S_{i,j} = \sum_{1 \leq i' \leq i, 1 \leq j' \leq j} x_{i',j'}
\]

\[
= \sum_{1 \leq i' \leq i, 1 \leq j' \leq j} \left( \sum_{-m \leq k_1 \leq m, -m \leq k_2 \leq m} \hat{x}_{i',j'}^{i'\!-\!k_1,j'\!-\!k_2} \right)
\]

\[
= \sum_{-m \leq k_1 \leq m, -m \leq k_2 \leq m} \left( \sum_{1 \leq i' \leq i, 1 \leq j' \leq j} \hat{x}_{i',j'}^{i'\!-\!k_1,j'\!-\!k_2} \right)
\]

\[
= \sum_{-m \leq k_1 \leq m, -m \leq k_2 \leq m} Y^{(k_1,k_2)}_{(i,j)}.
\]

By Lemma 3

\[
|S_{i,j}|^p = \left| \sum_{-m \leq k_1, k_2 \leq m} Y^{(k_1,k_2)}_{(i,j)} \right|^p
\]

\[
\leq (2m + 1)^{2(p-1)} \sum_{-m \leq i,j \leq m} |Y^{(k_1,k_2)}_{(i,j)}|^p.
\]

This in turn implies

\[
E\left[ \max_{(i,j) \leq (n_1,n_2)} |S_{i,j}|^p \right] \leq E\left[ \sum_{-m \leq k_1, k_2 \leq m} \max_{(i,j) \leq (n_1,n_2)} |Y^{(k_1,k_2)}_{(i,j)}|^p \right]
\]

\[
= \sum_{-m \leq k_1, k_2 \leq m} E\left[ \max_{(i,j) \leq (n_1,n_2)} |Y^{(k_1,k_2)}_{(i,j)}|^p \right]
\]

\[
\leq q^{2p} (2m + 1)^{2(p-1)} \sum_{-m \leq i,j \leq m} \max_{(i,j) \leq (n_1,n_2)} E[|Y^{(k_1,k_2)}_{(i,j)}|^p],
\]

where the second inequality follows from applying Theorem 3.1 to the martingale

\[
\{Y^{(k_1,k_2)}_{(i,j)}, i \leq n_1, j \leq n_2\}.
\]
4 Invariance Principle

Lemma 4. Let \( \{\nu_{i,j}\}_{i \geq 1, j \geq 1} \) be an \( m \)-dependent field such that \( \{\nu_{i,j}^2\}_{i \geq 1, j \geq 1} \) is uniformly integrable and \( E[\frac{S_{n,n}^2}{n}] \to \sigma^2 \). Then the sequence \( \{X_n\} \) is tight and any limiting measure is supported on \( C \).

Proof. By Corollary 1, it suffices to show that the family

\[
\left\{ \max_{i \leq n, j \leq n} \frac{|S_{k_1+i,k_2+j} - S_{k_1,k_2}|^2}{n^2}, n \geq 1, k_1, k_2 \geq 1 \right\}
\]

is uniformly integrable. The argument below goes through for any \((k_1, k_2)\) and for notational simplicity, we will omit the \((k_1, k_2)\) subscript (or set them equal to zero).

For \( c > 0 \), consider the \( m \)-dependent fields \( \{\nu_{i,j}^c\} \) and \( \{Z_{i,j}\} \) defined by

\[
\nu_{i,j}^c = E_{i-m,j+m}[\nu_{i,j}1_{|\nu_{i,j}| \leq c}] - E_{i-m,j-m}[\nu_{i,j}1_{|\nu_{i,j}| \leq c}],
\]

\[
Z_{i,j} = \nu_{i,j} - \nu_{i,j}^c.
\]

We define the notation

\[
E^y[\xi] \equiv E[\xi 1_{|\xi| \geq y}], \quad \nu_{i,j}' = \sum_{i' \leq i, j' \leq j} \nu_{i',j}', \quad Z_{i,j}' = \sum_{i' \leq i, j' \leq j} Z_{i',j}'.
\]

Since \( S_{i,j}^2 \leq 2\nu_{i,j}'^2 + 2Z_{i,j}'^2 \),

\[
E^y[\max_{i \leq n, j \leq n} \frac{|S_{i,j}|^2}{n^2}] \leq 4E^y[\max_{i \leq n, j \leq n} \frac{\nu_{i,j}'^2}{n^2}] + 4E[\max_{i \leq n, j \leq n} \frac{|Z_{i,j}'|^2}{n^2}].
\]

By uniform square integrability of \( \{\nu_{i,j}\} \), \( E[|Z_{i,j}|^2] \leq g(c) \to 0 \) as \( c \to \infty \), for all \( i, j \).

Therefore, in the notation of Proposition 7

\[
E[|Y_{(i,j)}^{(0,0)}|^2] \leq n^2g(c), \quad \forall (i, j) \leq (n, n),
\]

which implies

\[
E[\max_{(i,j) \leq (n,n)} \frac{|Z_{i,j}|^2}{n^2}] \leq 2^4 \cdot (2m + 1)^2 \sum_{-m \leq k_1, k_2 \leq m} g(c).
\]

The right-hand side can be made arbitrarily small by choosing \( c \) sufficiently large.

We now consider the term \( \max_{(i,j) \leq (n,n)} \frac{|\nu_{i,j}|^2}{n^2} \). Applying Proposition 7 to the case \( p = 4 \)
gives
\[
E\left[ \max_{(i,j) \leq (n,n)} |\nu_{ij}|^4 \right] \leq \frac{4^8}{3} \cdot (2m + 1)^6 \sum_{-m \leq k_1, k_2 \leq m} \max_{(i,j) \leq (n,n)} E[|Y_{(i,j)},(k_1,k_2)|^4],
\]
where
\[
Y_{(i,j)}^{(k_1,k_2)} = \sum_{(i',j') \leq (i,j)} E[\nu_{i',j'} | F_{k_1+k_1+k_2}].
\]
Each summand in \(Y_{(i,j)}^{(k_1,k_2)}\) is bounded in absolute value by \(c\) a.s. By a similar counting argument as in inequality 23.7 of Billingsley (1968), one can show that
\[
E[|Y_{(i,j)},(k_1,k_2)|^4] \leq n^4 K_c
\]
where \(K_c\) is a constant that only depends on \(c\). This shows \(\{\max_{(i,j) \leq (n,n)} |\nu_{ij}|^2, n \geq 1\}\) is bounded in \(L^2\), therefore uniformly integrable. One can then choose \(y\) sufficiently large so that
\[
E[y^2 \max_{i \leq n, j \leq n} |\nu_{ij}|^2] < \frac{n}{2}
\]
is sufficiently small, uniform in \(n\).

In other words, for all \(\eta > 0\), there exists \(c\) such that \(4E[\max_{i \leq n, j \leq n} |\nu_{ij}|^2] < \frac{n}{2}\). With this given choice of \(c\), there exist \(y > 0\) such that \(E[y^2 \max_{i \leq n, j \leq n} |\nu_{ij}|^2] < \frac{n}{2}\). This proves the proposition.

**Theorem 4.1.** Suppose, in addition to assumptions of Lemma 4,
\[
E\left[ \frac{(S_{k_1+n,k_2+n} - S_{k_1,k_2})^2}{n^2} | \mathcal{F}_{k_1-m_1,k_2-m_2} \right] \to \sigma^2,
\]
as \(n \to \infty\) for all \((k_1, k_2)\), then \(\{X_n\}\) converges weakly to the Brownian sheet.

**Proof.** According to Lemma 4, it suffices to verify Conditions 1\(^\circ\)a, 2\(^\circ\)a, and 3\(^\circ\)a.

Uniform integrability of the family,
\[
\{ \max_{i \leq n, j \leq n} \frac{|S_{k_1+i,k_2+j} - S_{k_1,k_2}|^2}{n^2}, n \geq 1, k_1, k_2 \geq 1 \}
\]
follows from Lemma 4. This in turn implies that
\[
\{ X_n^2(\Delta((\hat{s},\hat{t}) \times (t, t + h))) | \frac{h(t - \hat{s})}{h(t - \hat{s})}, h \geq \frac{1}{n}, t \in (0, 1 - h), (\hat{s},\hat{t}) \subset [0, 1], n = 1, 2, \ldots \}
\]
is uniformly integrable. Condition 2\(^\circ\)a now follows by taking \(t = \hat{s} = 0\). Condition 3\(^\circ\)a is also immediate (since \(0 < \hat{t} - \hat{s} < 1\).
To verify Condition 1°a, let \( t \in (0, 1), (\hat{s}, \hat{t}) \subset [0, 1], t_1, \ldots, t_k \subset I(t, \hat{s}), \) and \( u_1, \ldots, u_k \in \mathbb{R}. \) Define
\[
U_n = E\left[\sum_{l=1}^{k} iu_l X_n(t_l) \right| \mathcal{F}_{[n(t, \hat{s})]}^w]
\]
where the "weak past" \( \sigma \)-algebra \( \mathcal{F}_{k_1, k_2}^w \) is defined to be
\[
\mathcal{F}_{k_1, k_2}^w = \sigma(\mathcal{F}_{i,j}, i \leq k_1, \text{ or } j \leq k_2).
\]
Then
\[
\|U_n - \sum_{l=1}^{k} iu_l X_n(t_l)\|_2 \leq \max\{|u_1|, \ldots, |u_k|\} \frac{1}{n} \sigma \sum_{l=1}^{k} \sum_{(1,1) \leq (i,j) \leq [nt_l]} \|E[\nu_{i,j}|\mathcal{F}_{[n(t, \hat{s})]}^w] - \nu_{i,j}\|_2
\]
\[
\leq K \cdot \frac{1}{n},
\]
for some constant \( K \) that depends on \( m \) and \( \sup \|\nu_{i,j}\|_2 \), since the sum
\[
\sum_{l=1}^{k} \sum_{(1,1) \leq (i,j) \leq [nt_l]} \|E[\nu_{i,j}|\mathcal{F}_{[n(t, \hat{s})]}^w] - \nu_{i,j}\|_2
\]
is finite by \( m \)-dependence and \( L^2 \)-boundedness.

Together with uniform integrability of \( X_n(\Delta((\hat{s}, \hat{t}) \times (t, t+h))) \), this in turn implies
\[
E\left[\sum_{j=1}^{k} iu_j X_n(t_j) - e^{iU_n} X(\Delta_{t,t+h})\right] \to 0.
\]
But
\[
|E[e^{iU_n} X_n(\Delta_{t,t+h})]| \leq \|E[X(\Delta_{t,t+h})|\mathcal{F}_{[n(t, \hat{s})]}^w]\|_2 = O(\frac{1}{n}) \to 0
\]
as \( n \to \infty. \) Similarly, uniform integrability of \( X_n^2(\Delta_{t,t+h}) \) implies
\[
E\left[\sum_{j=1}^{k} iu_j X_n(t_j) - e^{iU_n}(X_n^2(\Delta_{t,t+h}) - h(\hat{t} - \hat{s}))\right] \to 0.
\]
Since
\[
|E[e^{iU_n}(X^2(\Delta_{t,t+h}) - h(\hat{t} - \hat{s}))]| \leq |(X_n^2(\Delta_{t,t+h}) - h(\hat{t} - \hat{s}))| \to 0,
\]
by (2), both conditions in 1°a are verified. This proves the theorem.
5 Conclusion

We proved an elementary invariance principle for the Brownian sheet where strong or wide-sense stationarity is not required. It is of interest in applications where stationarity may be too strong an assumption. An immediately application is unit root testing for spatial models, the detailed discussion of which will be given in a separate paper.

References

Berkes, I. and G. J. Morrow (1981). Strong invariance principles for mixing random fields. *Probability Theory and Related Fields* 57(1), 15–37.

Bickel, P. J. and M. J. Wichura (1971). Convergence criteria for multiparameter stochastic processes and some applications. *Annals of Mathematical Statistics* 42(5), 1656–1670.

Billingsley, P. (1968). *Convergence of Probability Measures*. Wiley.

Volný, D. and Y. Wang (2014). An invariance principle for stationary random fields under Hannans condition. *Stochastic Processes and Their Applications* 124(12), 4012–4029.

Walsh, J. B. (1986). Martingales with a multidimensional parameter and stochastic integrals in the plane. In *Lectures in Probability and Statistics*, pp. 329–491. Springer.

Wang, Y. and M. Woodroofe (2013). A new condition for the invariance principle for stationary random fields. *Statistica Sinica*, 1673–1696.