# Invariance Principle for weakly dependent random fields

Jana Klicnarová

University of South Bohemia
České Budějovice
Czech Republic

CINCINNATI SYMPOSIUM 2014

Joint work with Dalibor Volný and Yizao Wang

Supported by Czech Science Foundation (project n. P201/11/P164).

# Problem

$(X_{\mathbf{i}})_{\{\mathbf{i}\in\mathbb{Z}^d\}}$ – a stationary random field with a zero mean and a finite second moment.

## CLT problem

Put

$$S_{\Gamma_n} = \sum_{\mathbf{i}\in\Gamma_n} X_{\mathbf{i}},$$

where $\Gamma_n \subset \mathbb{Z}^d$, $|\Gamma_n| \to \infty$ as $n \to \infty$.

**When does**

$$S_{\Gamma_n}/\sqrt{|\Gamma_n|} \to \ N(0, \sigma^2)?$$

## IP PROBLEM

Let

$$S_n(A) = \sum_{\mathbf{i}\in[\mathbf{0},\mathbf{n}]} \lambda(nA \cap R_{\mathbf{i}})X_{\mathbf{i}},$$

where $R_{\mathbf{i}} = (\mathbf{i} - \mathbf{1}, \mathbf{i}]$ and $nA = \{n\mathbf{x} : \mathbf{x} \in A\}$.

**When does**

$$\{n^{-d/2}S_n(A); A \in \mathcal{A}\} \to \sigma W \quad \text{in} \quad C(\mathcal{A}),$$

**where $W$ is a standard Brownian motion indexed by $\mathcal{A}$?**

# Basic notation

$$(\Omega, \mathcal{A}, P) \quad = \quad (\mathbb{R}^{\mathbb{Z}^d}, \mathcal{B}^{\mathbb{Z}^d}, P^{\mathbb{Z}^d})$$

$(\epsilon_{\mathbf{k}})_{\mathbf{k}\in\mathbb{Z}^d}$ – iid random variables, $\epsilon_{\mathbf{k}}(\omega) = \omega_{\mathbf{k}}$

$f$ – a measurable function on $\Omega$ such that $f \in L_2(\mu)$, regular and with a zero mean.

$\{T^{\mathbf{k}}\}_{\mathbf{k}\in\mathbb{Z}^d}$ – shift operators on $\mathbb{R}^d$: $(T^{\mathbf{k}}\omega)_{\mathbf{l}} = \omega_{\mathbf{k}+\mathbf{l}}$,

$(X_{\mathbf{i}})_{\mathbf{i}}$ – it is a (strictly) stationary process ($X_{\mathbf{i}} = f \circ T^{\mathbf{i}}$).

$U$ – a unitary operator on $L^2$, such that $U^{\mathbf{i}} f = f \circ T^{\mathbf{i}}$,

$(\mathcal{F}_{\mathbf{k}})_{\mathbf{k}}$ – a filtration, $\mathcal{F}_{\mathbf{k}} = \sigma(\epsilon_{\mathbf{l}} : \mathbf{l} \le \mathbf{k})$

$$\sigma^2_{\Gamma_n} \quad = \quad \mathrm{E}(S_{\Gamma_n}(f))^2.$$

# Projection and Hannan's condition

In a 1-dimensional case:

$$P_i(X) = \mathsf{E}(X|\mathcal{F}_i) - \mathsf{E}(X|\mathcal{F}_{i-1}).$$

## Hannan's condition (1973)

$$\sum_{i=1}^{\infty} ||P_0(X_i)||_2 < \infty.$$

Hannan (1973) proved a CLT and an IP for stationary processes under this condition and with some more conditions. Later, Dedecker, Merlevede and Volný (2007) proved an IP for stationary processes under only this condition.

# Hannan's condition in a high dimension

The definition of $P_{\mathbf{i}}$ in a d-dimensional case is more complicated, for more details see Volný and Wang (2014). The idea of the definition:

- we suppose a **commuting filtration** (see Khoshnevisan (2002)): for every bounded $\mathcal{F}_l$-measurable r.v. Y:

$$\mathsf{E}(Y|\mathcal{F}_k) = \mathsf{E}(Y|\mathcal{F}_{k \wedge l}) \ a.s.$$

- the projection operator $P_{\mathbf{i}}$ is defined as:

$$P_{\mathbf{i}} = \Pi_{q=1}^{d} P_{j_q}^{q},$$

  where $P_{j_q}^{q}$ are "marginal" projections.

# Hannan's condition in a multidimensional case and limit theorems

Volný and Wang (2014) established an Invariance Principle under Hannan's condition (they suppose a finite second moment):

$$\sum_{\mathbf{i}\in\mathbb{Z}^d} ||P_{\mathbf{0}}X_{\mathbf{i}}||_2 < \infty.$$

Their result is for summation over rectangles. We are interested in limit theorems where a summation is over more general sets.

# El Machkouri, Volný, Wu (2013)

## p-stability – notation

Let us have

$$X_{\mathbf{i}} = g(\varepsilon_{\mathbf{i}-\mathbf{j}}; \mathbf{j} \in \mathbb{Z}^d),\ \mathbf{i} \in \mathbb{Z}^d, \tag{1}$$

where $(\varepsilon_{\mathbf{i}})_{\mathbf{i}}$ are i.i.d., and $(\varepsilon_{\mathbf{i}}^{'})_{\mathbf{i}}$ are i.i.d. copies of $(\varepsilon_{\mathbf{i}})_{\mathbf{i}}$.
Then $X_{\mathbf{i}}^{*}$ is a version of $X_{\mathbf{i}}$ such that

$$X_i^{*} = g(\varepsilon_{\mathbf{i}-\mathbf{j}}^{*}; \mathbf{j} \in \mathbb{Z}^d),\ \mathbf{i} \in \mathbb{Z}^d,$$

where

$$\begin{aligned} \varepsilon_{\mathbf{i}}^{*} &= \varepsilon_{\mathbf{i}} \text{ for all } \mathbf{i} \neq \mathbf{0}, \\ &= \varepsilon_{\mathbf{0}}^{'} \text{ for } \mathbf{i} = \mathbf{0}. \end{aligned}$$

Then we can put $\delta_{i,p} = ||X_{\mathbf{i}} - X_{\mathbf{i}}^{*}||_p$ and $\Delta_p = \sum_{\mathbf{i} \in \mathbb{Z}^d} \delta_{\mathbf{i},p}$.

## Definition

We say, that the process is p-stable if $\Delta_p < \infty$.

# El Machkouri, Volný, Wu (2013) – notation

## Ψ-stability – Young function

A function $\Psi$ is a **Young function** if it is a real convex nondecreasing function defined on $\mathbb{R}^+$ which satisfies

$$\begin{aligned} \lim_{t\to\infty} \Psi(t) &= \infty \\ \Psi(0) &= 0. \end{aligned}$$

The Orlicz space $L_\Psi$ is defined as a space of real random variables $Z$ defined on a probability space $(\Omega, \mathcal{A}, P)$ such that

$$\mathrm{E}[\Psi(|Z|/c)] < \infty$$

for some $c > 0$. For more details see for example Ledoux and Talagrand (1991).

# El Machkouri, Volný, Wu (2013) – notation

## Luxembourg norm

The Orlicz space $L_\Psi$ is equipped with Luxemburg norm $||\cdot||_\Psi$ defined for real random variable by

$$||Z||_\Psi = \inf\{c > 0; \mathrm{E}[\Psi(|Z|/c)] \leq 1\}.$$

So, it is possible to generalize the definition of $p$-stable processes to $\Psi$-stable processes. Then we can put $\delta_{i,\Psi} = ||X_{\mathbf{i}} - X_{\mathbf{i}}^*||_\Psi$ and $\Delta_\Psi = \sum_{\mathbf{i}\in\mathbb{Z}^d} \delta_{\mathbf{i},\Psi}$.

## Definition

We say, that the process is $\Psi$-stable if $\Delta_\Psi < \infty$.

# CLT – El Machkouri, Volný, Wu (2013)

**Theorem**

Let $(X_{\mathbf{i}})_{\mathbf{i}\in\mathbb{Z}^d}$ be a stationary centred random field defined by (1) satisfying $\Delta_2 < \infty$. Assume that $(\Gamma_n)_n$ is a sequence of finite subsets of $\mathbb{Z}^d$ such that $|\Gamma_n| \to \infty$ and $\sigma_{\Gamma_n} = \mathrm{E}(S^2_{\Gamma_n}) \to \infty$, then Levy distance

$$L\left(S_{\Gamma_n}/\sqrt{|\Gamma_n|}, \mathcal{N}(0, \sigma_{\Gamma_n^2}/|\Gamma_n|)\right) \to 0 \text{ as } n \to \infty.$$

**Corollary**

If $|\partial\Gamma_n|/|\Gamma_n| \to 0$ and $\sigma^2 = \sum_{\mathbf{k}\in\mathbb{Z}^d} \mathrm{E}(X_{\mathbf{0}}X_{\mathbf{k}}) > 0$ then

$$\frac{S_{\Gamma_n}}{\sqrt{|\Gamma_n|}} \to \mathcal{N}(0, \sigma^2).$$

# El Machkouri, Volný, Wu (2013) – notation

To introduce an Invariance Principle given by El Machkouri, Volný and Wu we need also to recall some definitions about entropy and VC-classes. For more details see for example van der Vaart and Wellner (1996).

## Covering number and entropy

Let us have a collection $\mathcal{A}$ of Borel subsets on $[0,1]^d$. We can equip a collection with a pseudometric $\rho$:

$$\rho(A,B) = \sqrt{\lambda(A \Delta B)}.$$

To measure a size of $\mathcal{A}$ it is possible to use a metric entropy. Let us recall, that the entropy $H(\mathcal{A},\rho,\varepsilon)$ is the logarithm of $N(\mathcal{A},\rho,\varepsilon)$, where $N(\mathcal{A},\rho,\varepsilon)$ is so called a covering number – it is the smallest number of open balls of radius $\varepsilon$ with respect to $\rho$ which cover $\mathcal{A}$.

# El Machkouri, Volný, Wu (2013) – notation

## Vapnik-Chervonenkis classes

Let $\mathcal{C}$ be a collection of subsets of a set $\mathcal{X}$. And let $F \subset \mathcal{X}$. We say that $\mathcal{C}$ picks out a certain subset of $F$ if this can be formed as $F \cap C$ for some $C \in \mathcal{C}$. The collection $\mathcal{C}$ is said to shatter $F$ if it picks out each of its $2^{|F|}$ subsets. The VC-index $V(\mathcal{C})$ of the class $\mathcal{C}$ is the smallest $n$ for which no set of size $n$ is shattered by $\mathcal{C}$. Formally,

$$V(\mathcal{C}) = \inf \left\{ n; \max_{x_1,\dots,x_n} \Delta_n(\mathcal{C}, x_1, \dots, x_n) < 2^n \right\},$$

where $\Delta_n(\mathcal{C}, x_1, \dots, x_n) = |\{C \cap \{x_1, \dots, x_n\}; C \in \mathcal{C}\}|$.

# El Machkouri, Volný, Wu (2013) – notation

### Young function

Let $\beta > 0$ and

$$h_\beta = ((1-\beta)\beta)^{\frac{1}{\beta}} \mathrm{I}_{\{0<\beta<1\}}.$$

Then we denote by $\psi_\beta$ the Young function:

$$\psi_\beta(x) = \exp\{(x + h_\beta)^\beta\} - \exp\{h_\beta^\beta\}, \quad x \in \mathbb{R}^+.$$

# IP – El Machkouri, Volný, Wu (2013)

## IP

Let $(U_{\mathbf{i}} f)_{\mathbf{i}\in\mathbb{Z}^d}$ be a stationary centered random field and let $\mathcal{A}$ be a collection of regular Borel subsets of $[0,1]^d$. Assume that one of the following conditions holds:

(I) The collection $\mathcal{A}$ is a Vapnik-Chervonenkis class with an index V and there exists $p > 2(V-1)$ such that $f \in L_p$ and $\Delta_p < \infty$.

(II) There exists a positive $\theta$ and $0 < q < 2$: $E[\exp(\theta|f|^{\beta(q)})] < \infty$, where $\beta(q) = 2q/(2-q)$ and $\Delta_{\Psi(\beta(q))} < \infty$ and such that the class $\mathcal{A}$ satisfies condition

$$\int_0^1 (H(\mathcal{A},\rho,\varepsilon))^{1/q}\, d\varepsilon < \infty.$$

# IP – El Machkouri, Volný, Wu (2013)

(iii) $f \in L^\infty$, $\int_0^1 (H(\mathcal{A}, \rho, \varepsilon))^{1/2}\, d\varepsilon < \infty$ and

$$\Delta_\infty < \infty.$$

Then the sequence of processes $\{n^{-d/2} S_n(A); A \in \mathcal{A}\}$, where

$$S_n(A) = \sum_{\mathbf{i} \in [\mathbf{0}, \mathbf{n}]} \lambda(nA \cap R_{\mathbf{i}}) U_{\mathbf{i}} f$$

with $R_{\mathbf{i}} = (\mathbf{i} - \mathbf{1}, \mathbf{i}]$, converges in distribution in $C(\mathcal{A})$ to $\sigma W$, where $W$ is a standard Brownian motion indexed by $\mathcal{A}$ and $\sigma^2 = \sum_{\mathbf{i} \in \mathbb{Z}^d} \mathrm{E}(f U_{\mathbf{i}} f)$.

# CLT with general summation under Hannan's condition

## Problem

Can we formulate a limit theorem for general $(\Gamma_n)$ also under Hannan's condition?

## CLT for martingale differences

If $(f \circ T^{\mathbf{i}})_{\mathbf{i} \in \mathbb{Z}^d}$ is a martingale difference field and $f \in L_2$ then the Central Limit Theorem holds: for $\Gamma_n \subset Z^d$: $|\Gamma_n| \to \infty$ we have

$$\frac{S_{\Gamma_n}(f)}{\sqrt{|\Gamma_n|}} \to \mathcal{N}(0, ||f||_2^2).$$

# Important condition on $(\Gamma_n)$

To obtain a Central Limit Theorem for general sets $(\Gamma_n)$, we need a collection of $(\Gamma_n)$ to satisfy some condition.

- If we suppose $(\Gamma_n)$ such that $\lim_{n\to\infty} \frac{|\partial\Gamma_n|}{|\Gamma_n|} = 0$ then under $(L_2)$ Hannnan's condition:

$$\lim_{n\to\infty} \frac{\mathrm{E}(S^2_{\Gamma_n}(f))}{|\Gamma_n|} = \sum_{\mathbf{k}\in\mathbb{Z}^d} \mathrm{E}(fU_{\mathbf{k}}f).$$

- At least we need

$$\liminf_{n\to\infty} \frac{\mathrm{E}(S^2_{\Gamma_n}(f))}{|\Gamma_n|} > 0.$$

# CLT under Hannan's condition with $|\partial\Gamma_n|/|\Gamma_n| \to 0$

**Theorem**

Let a zero-mean $f \in L_2$ be regular and satisfy Hannnan's condition. Let $(\Gamma_n)_{n\in\mathbb{N}}$ be such that

$$\lim_{n\to\infty} \frac{|\partial\Gamma_n|}{|\Gamma_n|} = 0 \ \text{ and } \ |\Gamma_n| \to \infty.$$

then $S_{\Gamma_n}(f)/\sqrt{|\Gamma_n|} \to N(0,\sigma^2)$, where $\sigma^2 = \sum_{\mathbf{k}\mathbb{Z}^d} \mathsf{E}(fU_{\mathbf{k}}f)$.

# CLT under Hannan's condition

**Theorem**

Let a zero-mean $f \in L_2$ be regular and satisfy Hannan's condition. Let $(\Gamma_n)_n$ be a sequence of finite subsets of $\mathbb{Z}^d$ such that $|\Gamma_n| \to \infty$, $\sigma_n = \mathrm{E}(S_{\Gamma_n}^2) \to \infty$ and

$$\liminf_{n\to\infty} \sigma_n / |\Gamma_n|^{1/2} > 0,$$

then $S_{\Gamma_n}(f)/\sigma_n$ converge in distribution to $\mathcal{N}(0,1)$.

**Corollary**

For a regular $f \in L^2$ satisfying the Hannan's condition we have

$$L(S_{\Gamma_n}(f)/|\Gamma_n|^{1/2}, \mathcal{N}(0, \sigma_n^2/|\Gamma_n|)) \to 0.$$

# Hannan's condition

A random field $(f \circ T^{\mathbf{i}})_{\mathbf{i}\in\mathbb{Z}^d}$ satisfies $L_p$-Hannan's condition if

$$\sum_{\mathbf{i}\in\mathbb{Z}^d} ||P_0 X_{\mathbf{i}}||_p < \infty.$$

And, more generally, we say that a random field satisfies $L_\psi$-Hannan's condition if

$$\sum_{\mathbf{i}\in\mathbb{Z}^d} ||P_0 X_{\mathbf{i}}||_\psi < \infty,$$

where $\psi$ is any Young function and $||\cdot||_\psi$ is a related Luxemburg norm, see Ledoux and Talagrand (1991).

For short of notation we will use

$$\Theta_p = \sum_{\mathbf{i}\in\mathbb{Z}^d} ||P_0 X_{\mathbf{i}}||_p$$

and

$$\Theta_\psi = \sum ||P_0 X_{\mathbf{i}}||_\psi.$$

# p-stability and Hannan's condition

## Hannan's condition is weaker then p-stability condition in $L_p$

Wu (2005) proved, in a 1-dimensional case, that $\Delta_p \geq \Theta_p$. In other words, that p-stability of a process implies $L_p$-Hannan's condition for this process. This result can be extend into a high dimensional case, too. Volný and Wang (2014) showed an example of a process such that it satisfies Hannan's condition but p-stability does not take a place.

# Invariance Principle

**Theorem**

Let $(U_{\mathbf{i}} f)_{\mathbf{i}\in\mathbb{Z}^d}$ be a stationary centered random field and let $\mathcal{A}$ be a collection of regular Borel subsets of $[0,1]^d$. Assume that one of the following conditions holds:

(I) The collection $\mathcal{A}$ is a Vapnik-Chervonenkis class with an index V and there exists $p > 2(V-1)$ such that $f \in L_p$ and $L_p$-Hannan's condition is satisfied.

(II) There exists a positive $\beta$ such that $0 < \beta < 1$ and

$$\int_0^1 (H(\mathcal{A}, \rho, \varepsilon))^{1/\beta} \, d\varepsilon < \infty. \tag{2}$$

For $\gamma = \beta/(1-\beta)$: $||f||_{\psi_\gamma} < \infty$ and $\Theta_{\psi_\gamma} < \infty$.

# Invariance Principle

(III) $f \in L^{\infty}$, $\int_0^1 (H(\mathcal{A}, \rho, \varepsilon))^{1/2}\, d\varepsilon < \infty$ and
$\Theta_{\infty} := \sum_{\mathbf{i} \in \mathbb{Z}^d} \|P_0 U_{\mathbf{i}} f\|_{\infty} < \infty$

Then the sequence of processes $\{n^{-d/2} S_n(A); A \in \mathcal{A}\}$, where

$$S_n(A) = \sum_{\mathbf{i} \in [\mathbf{0},\mathbf{n}]} \lambda(nA \cap R_{\mathbf{i}}) U_{\mathbf{i}} f$$

with $R_{\mathbf{i}} = (\mathbf{i} - \mathbf{1}, \mathbf{i}]$, converges in distribution in $C(\mathcal{A})$ to $\sigma W$, where $W$ is a standard Brownian motion indexed by $\mathcal{A}$ and $\sigma^2 = \sum_{\mathbf{i} \in \mathbb{Z}^d} \mathrm{E}(f U_{\mathbf{i}} f)$.

# Key step on a way to IP

## Lemma

For $p \geq 2$ and a regular zero-mean $f \in L_p$:

$$\left\|\sum_{\mathbf{i}\in\mathbb{Z}^d} c_{\mathbf{i}} X_{\mathbf{i}}\right\|_p \leq B_p \sqrt{\sum_{\mathbf{i}\in\mathbb{Z}^d} c_{\mathbf{i}}^2} \sum_{\mathbf{i}\in\mathbb{Z}^d} ||P_{\mathbf{0}} X_{\mathbf{i}}||_p = B_p \Theta_p \sqrt{\sum_{\mathbf{i}\in\mathbb{Z}^d} c_{\mathbf{i}}^2},$$

where $P_{\mathbf{0}} f = P_{\{0,0,\dots,0\}} f$, $\Theta_p = \sum_{\mathbf{i}\in\mathbb{Z}^d} ||P_{\mathbf{0}} U_{\mathbf{i}} f||_p$ and $B_p$ is a constant depending on $p$, more precisely $B_p = 18p\sqrt{q}$, where $q$ is such that $1/p + 1/q = 1$.

## Remark

The constant $B_p$ can be found in Hall and Heyde (1980) and for $p = 2$ we are able to prove this inequality with $B_2 = 1$.

Thank you for your attention!