Biased Brownian Coupling of the Empirical Process of Stationary Weakly Dependent Data

Philippe BERTHET (Toulouse, France)

Peligrad conf., June 23, 2010

- Sequence X_n of r.v.

Introduction

General setting

- Sequence X_n of r.v.
  - $(\Omega, \mathcal{T}, \mathbb{P}) \to (\mathcal{X}, \mathcal{A})$, same law $P = \mathbb{P}^{X_1}$

Introduction

General setting

- Sequence X_n of r.v.
  - $(\Omega, \mathcal{T}, \mathbb{P}) \to (\mathcal{X}, \mathcal{A})$, same law $P = \mathbb{P}^{X_1}$
  - independent or stationary mixing (β, ρ or ϕ)

- Sequence X_n of r.v.
  - $(\Omega, \mathcal{T}, \mathbb{P}) \to (\mathcal{X}, \mathcal{A})$, same law $P = \mathbb{P}^{X_1}$
  - independent or stationary mixing (β, ρ or ϕ)
  - Ω rich enough to allow our Gaussian coupling

- Sequence X_n of r.v.
  - $(\Omega, \mathcal{T}, \mathbb{P}) \to (\mathcal{X}, \mathcal{A})$, same law $P = \mathbb{P}^{X_1}$
  - independent or stationary mixing (β, ρ or ϕ)
  - Ω rich enough to allow our Gaussian coupling
- Class $\mathcal{F}$ of bounded functions $f : \mathcal{X} \to \mathbb{R}$

- Sequence X_n of r.v.
  - $(\Omega, \mathcal{T}, \mathbb{P}) \to (\mathcal{X}, \mathcal{A})$, same law $P = \mathbb{P}^{X_1}$
  - independent or stationary mixing (β, ρ or ϕ)
  - Ω rich enough to allow our Gaussian coupling
- Class $\mathcal{F}$ of bounded functions $f : \mathcal{X} \to \mathbb{R}$
  - mean and variance

$$Pf = \int_{\mathcal{X}} f dP, \quad \int_{\mathcal{X}} (f - Pf)^2 dP < \infty$$

- Sequence X_n of r.v.
  - $(\Omega, \mathcal{T}, \mathbb{P}) \to (\mathcal{X}, \mathcal{A})$, same law $P = \mathbb{P}^{X_1}$
  - independent or stationary mixing (β, ρ or ϕ)
  - Ω rich enough to allow our Gaussian coupling
- Class $\mathcal{F}$ of bounded functions $f : \mathcal{X} \to \mathbb{R}$
  - mean and variance

$$Pf = \int_{\mathcal{X}} f dP, \quad \int_{\mathcal{X}} (f - Pf)^2 dP < \infty$$

  - sup-norm

$$\|\mathbb{G}\|_{\mathcal{F}} = \sup_{f \in \mathcal{F}} |\mathbb{G}(f)|$$

Introduction

Emprical process indexed by functions

- Based on $X_1, ..., X_n$

Introduction

Emprical process indexed by functions

- Based on $X_1, ..., X_n$
  - empirical measure

$$P_n = \frac{1}{n} \sum_{i=1}^{n} \delta_{X_i}$$

Introduction

Emprical process indexed by functions

- Based on $X_1, ..., X_n$
  - empirical measure

$$P_n = \frac{1}{n} \sum_{i=1}^{n} \delta_{X_i}$$

  - empirical means

$$P_n f = \int_{\mathcal{X}} f dP_n = \frac{1}{n} \sum_{i=1}^{n} f(X_i)$$

Introduction

Emprical process indexed by functions

- Based on $X_1, ..., X_n$
  - empirical measure

$$P_n = \frac{1}{n}\sum_{i=1}^{n}\delta_{X_i}$$

  - empirical means

$$P_n f = \int_{\mathcal{X}} f dP_n = \frac{1}{n}\sum_{i=1}^{n} f(X_i)$$

  - the $(P, \mathcal{F})$-empirical process

$$\alpha_n(f) = \sqrt{n}(P_n f - Pf) = \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\{f(X_i) - Pf\}$$

Introduction

Emprical process indexed by functions

- Based on $X_1, ..., X_n$
  - empirical measure

$$P_n = \frac{1}{n} \sum_{i=1}^{n} \delta_{X_i}$$

  - empirical means

$$P_n f = \int_{\mathcal{X}} f dP_n = \frac{1}{n} \sum_{i=1}^{n} f(X_i)$$

  - the $(P, \mathcal{F})$-empirical process

$$\alpha_n(f) = \sqrt{n}(P_n f - Pf) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \{f(X_i) - Pf\}$$

- Note that α_n is a sum of $\mathbb{L}_\infty(\mathcal{F})$-valued r.v.'s

- **Space** but no time involved in

$$f \in \mathcal{F} \rightarrow \; \alpha_n(f) = \sqrt{n}(P_n f - Pf)$$

Introduction

Time and space

- **Space** but no time involved in

$$f \in \mathcal{F} \rightarrow \ \alpha_n(f) = \sqrt{n}(P_n f - Pf)$$

- **Time** but no space involved in (only one f),

$$t \in [0,1] \rightarrow \ S_n(t) = \frac{1}{\sqrt{n}} \sum_{i=1}^{[nt]} (f(X_i) - Pf)$$

Introduction

Time and space

- **Space** but no time involved in

$$f \in \mathcal{F} \rightarrow \alpha_n(f) = \sqrt{n}(P_n f - Pf)$$

- **Time** but no space involved in (only one f),

$$t \in [0,1] \rightarrow S_n(t) = \frac{1}{\sqrt{n}} \sum_{i=1}^{[nt]} (f(X_i) - Pf)$$

- Possible to mix methods to study the sequential

$$\alpha_n(t,f) = \sqrt{n}(P_{[nt]} f - Pf) = \frac{1}{\sqrt{n}} \sum_{i=1}^{[nt]} (f(X_i) - \mathbb{E}(f(X)))$$

(nothing to do with $\mathcal{F} = \left\{ f_t = \mathbb{I}_{(-\infty,t]} \right\}$ in α_n on $\mathbb{R}$).

- More generally,

$$\alpha_n(t,f,\varphi) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} c_{i,n}(t)\varphi(X_i,Y_i)f(X_i).$$

for deterministic weights $c_{i,n}(t)$ and functions φ_j acting as random weights (or contrasts) based on a single sample (X_i, Y_i). If countable φ this is to study the joint behavior.

- More generally,

$$\alpha_n(t,f,\varphi)=\frac{1}{\sqrt{n}}\sum_{i=1}^{n}c_{i,n}(t)\varphi(X_i,Y_i)f(X_i).$$

  for deterministic weights $c_{i,n}(t)$ and functions φ_j acting as random weights (or contrasts) based on a single sample (X_i,Y_i). If countable φ this is to study the joint behavior.
- Likewise U-processes versions, by crossing indexes of two samples (X_i) and (Y_i),

$$\alpha_n(s,t,f)=\frac{1}{\sqrt{n}}\sum_{i=1}^{[sn]}\sum_{j=1}^{[tn]}c_{i,j,n}f(X_i,Y_j).$$

- More generally,

$$\alpha_n(t, f, \varphi) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} c_{i,n}(t)\varphi(X_i, Y_i) f(X_i).$$

for deterministic weights $c_{i,n}(t)$ and functions φ_j acting as random weights (or contrasts) based on a single sample (X_i, Y_i). If countable φ this is to study the joint behavior.

- Likewise U-processes versions, by crossing indexes of two samples (X_i) and (Y_i),

$$\alpha_n(s, t, f) = \frac{1}{\sqrt{n}} \sum_{i=1}^{[sn]} \sum_{j=1}^{[tn]} c_{i,j,n} f(X_i, Y_j).$$

- Typical limiting processes $\mathbb{G}$ are combinations of **Brownian Motions** and **Brownian Bridges** indexed by functions $\neq \mathcal{F}$.

- Provide tractable **Brownian coupling**

Introduction

Goal : Brownian paradigm

- Provide tractable **Brownian coupling**
  - $(n \to \infty)$ to derive **functional CLT** for α_n and **CLT** for $\Psi_n(\alpha_n)$

- Provide tractable **Brownian coupling**
  - ($n \to \infty$) to derive **functional CLT** for α_n and **CLT** for $\Psi_n(\alpha_n)$
  - (fixed n) to plugin a **limiting** or **intermediate** Gaussian process

- Provide tractable **Brownian coupling**
  - $(n \to \infty)$ to derive **functional CLT** for α_n and **CLT** for $\Psi_n(\alpha_n)$
  - (fixed n) to plugin a **limiting** or **intermediate** Gaussian process
  - fixed n is sometimes a shorter way to handle $n \to \infty$

- Provide tractable **Brownian coupling**
  - $(n \to \infty)$ to derive **functional CLT** for α_n and **CLT** for $\Psi_n(\alpha_n)$
  - (fixed n) to plugin a **limiting** or **intermediate** Gaussian process
  - fixed n is sometimes a shorter way to handle $n \to \infty$
- To study a statistic $T_n(X_1, ..., X_n) = \Psi_n(P_n)$

- Provide tractable **Brownian coupling**
  - $(n \to \infty)$ to derive **functional CLT** for α_n and **CLT** for $\Psi_n(\alpha_n)$
  - (fixed n) to plugin a **limiting** or **intermediate** Gaussian process
  - fixed n is sometimes a shorter way to handle $n \to \infty$
- To study a statistic $T_n(X_1, ..., X_n) = \Psi_n(P_n)$
  - approximate α_n with a version $\mathbb{G}_n$ of its limit $\mathbb{G}$ **OR** close to $\mathbb{G}$

$$\Psi_n \left(P + \frac{\mathbb{G}_n}{\sqrt{n}} + \frac{\alpha_n - \mathbb{G}_n}{\sqrt{n}} \right) \sim \Psi_\infty(P) + \phi_n \left(\frac{\mathbb{G}_n}{\sqrt{n}} \right) + o_{a.s.} \left(\frac{v_n}{\sqrt{n}} \right)$$

- Provide tractable **Brownian coupling**
  - $(n \to \infty)$ to derive **functional CLT** for α_n and **CLT** for $\Psi_n(\alpha_n)$
  - (fixed n) to plugin a **limiting** or **intermediate** Gaussian process
  - fixed n is sometimes a shorter way to handle $n \to \infty$
- To study a statistic $T_n(X_1, ..., X_n) = \Psi_n(P_n)$
  - approximate α_n with a version $\mathbb{G}_n$ of its limit $\mathbb{G}$ **OR** close to $\mathbb{G}$

$$\Psi_n\left(P + \frac{\mathbb{G}_n}{\sqrt{n}} + \frac{\alpha_n - \mathbb{G}_n}{\sqrt{n}}\right) \sim \Psi_\infty(P) + \phi_n\left(\frac{\mathbb{G}_n}{\sqrt{n}}\right) + o_{a.s.}\left(\frac{v_n}{\sqrt{n}}\right)$$

  - easy proof of weak convergence, exploit Gaussianity at finite n

- Let $\mathbb{G}$ be a P-**Brownian Bridge** indexed by $\mathcal{F}$

- Let $\mathbb{G}$ be a P-**Brownian Bridge** indexed by $\mathcal{F}$
  - mean zero Gaussian process, covariance

$$cov(\mathbb{G}(f),\mathbb{G}(g)) = cov(\alpha_n(f),\alpha_n(g)) = Pfg - Pf\,Pg$$

- Let $\mathbb{G}$ be a P-**Brownian Bridge** indexed by $\mathcal{F}$
  - mean zero Gaussian process, covariance

$$cov(\mathbb{G}(f), \mathbb{G}(g)) = cov(\alpha_n(f), \alpha_n(g)) = Pfg - Pf\, Pg$$

  - $\mathbb{G}$ is uniformly continuous on $\mathcal{F}$ with respect to

$$d_P(f,g) = \|f-g\|_{L_2(P)}$$

- Let $\mathbb{G}$ be a P-**Brownian Bridge** indexed by $\mathcal{F}$
  - mean zero Gaussian process, covariance

$$cov(\mathbb{G}(f), \mathbb{G}(g)) = cov(\alpha_n(f), \alpha_n(g)) = Pfg - Pf\,Pg$$

  - $\mathbb{G}$ is uniformly continuous on $\mathcal{F}$ with respect to

$$d_P(f, g) = \|f - g\|_{L_2(P)}$$

- **Weak invariance principle**

- Let $\mathbb{G}$ be a P-**Brownian Bridge** indexed by $\mathcal{F}$
  - mean zero Gaussian process, covariance

$$cov(\mathbb{G}(f), \mathbb{G}(g)) = cov(\alpha_n(f), \alpha_n(g)) = Pfg - Pf\,Pg$$

  - $\mathbb{G}$ is uniformly continuous on $\mathcal{F}$ with respect to

$$d_P(f, g) = \|f - g\|_{L_2(P)}$$

- **Weak invariance principle**
  - $\mathcal{F}$ is P-Donsker if

$$\alpha_n \to \mathbb{G} \quad \text{weakly in } \mathbb{L}_\infty(\mathcal{F})$$

Introduction

Weak invariance, independent case

- Let $\mathbb{G}$ be a P-**Brownian Bridge** indexed by $\mathcal{F}$
  - mean zero Gaussian process, covariance

$$cov(\mathbb{G}(f), \mathbb{G}(g)) = cov(\alpha_n(f), \alpha_n(g)) = Pfg - Pf\,Pg$$

  - $\mathbb{G}$ is uniformly continuous on $\mathcal{F}$ with respect to

$$d_P(f,g) = \|f-g\|_{L_2(P)}$$

- **Weak invariance principle**
  - $\mathcal{F}$ is P-Donsker if

$$\alpha_n \to \mathbb{G} \quad \text{weakly in } \mathbb{L}_\infty(\mathcal{F})$$

  - well known sufficient conditions on $(P, \mathcal{F})$

Introduction

Weak invariance, independent case

- Let $\mathbb{G}$ be a P-**Brownian Bridge** indexed by $\mathcal{F}$
  - mean zero Gaussian process, covariance

$$cov(\mathbb{G}(f), \mathbb{G}(g)) = cov(\alpha_n(f), \alpha_n(g)) = Pfg - Pf\,Pg$$

  - $\mathbb{G}$ is uniformly continuous on $\mathcal{F}$ with respect to

$$d_P(f,g) = \|f-g\|_{L_2(P)}$$

- **Weak invariance principle**
  - $\mathcal{F}$ is P-Donsker if

$$\alpha_n \to \mathbb{G} \quad \text{weakly in } \mathbb{L}_\infty(\mathcal{F})$$

  - well known sufficient conditions on $(P, \mathcal{F})$
  - uniform CLT, a few rates (only one is optimal)

- Dependent stationnary $\mathbb{X} = \{X_n\}$ with small mixing coefficients (β, ρ or ϕ) has a weak limit, the $(P, \mathbb{X})$-Brownian Bridge $\mathbb{G}_\infty$ with covariance

$$\begin{aligned}\Gamma_\infty(f,g) = & \mathbb{E}(f(X)g(X)) - \mathbb{E}(f(X))\mathbb{E}(g(X)) \\ & + \sum_{i=2}^{\infty} \left(\mathbb{E}\left(f(X_i)g(X_1)\right) - \mathbb{E}(f(X))\mathbb{E}(g(X))\right) \\ & + \sum_{i=2}^{\infty} \left(\mathbb{E}\left(f(X_1)g(X_i)\right) - \mathbb{E}(f(X))\mathbb{E}(g(X))\right).\end{aligned}$$

- Dependent stationnary $\mathbb{X} = \{X_n\}$ with small mixing coefficients (β, ρ or ϕ) has a weak limit, the $(P, \mathbb{X})$-Brownian Bridge $\mathbb{G}_\infty$ with covariance

$$\begin{aligned}\Gamma_\infty(f,g) =& \mathbb{E}(f(X)g(X)) - \mathbb{E}(f(X))\mathbb{E}(g(X)) \\ &+ \sum_{i=2}^{\infty} \left(\mathbb{E}\left(f(X_i)g(X_1)\right) - \mathbb{E}(f(X))\mathbb{E}(g(X))\right) \\ &+ \sum_{i=2}^{\infty} \left(\mathbb{E}\left(f(X_1)g(X_i)\right) - \mathbb{E}(f(X))\mathbb{E}(g(X))\right).\end{aligned}$$

- It is not easy to learn accurately the dependence structure of $\mathbb{X}$ from $X_1, ..., X_n$ so to learn $\mathbb{G}_\infty$.

Introduction

Strong invariance principle, independent case

- Last 50 years, on $\mathcal{X} = \mathbb{R}$ a lot of work to strengthen Donsker

Introduction

Strong invariance principle, independent case

- Last 50 years, on $\mathcal{X} = \mathbb{R}$ a lot of work to strengthen Donsker
  - Skorokhod embedding (allows up to $n^{-1/4+\varepsilon}$ rates for sums)

Introduction

Strong invariance principle, independent case

- Last 50 years, on $\mathcal{X} = \mathbb{R}$ a lot of work to strengthen Donsker
  - Skorokhod embedding (allows up to $n^{-1/4+\varepsilon}$ rates for sums)
  - Csörgő-Révész construction

- Last 50 years, on $\mathcal{X} = \mathbb{R}$ a lot of work to strengthen Donsker
  - Skorokhod embedding (allows up to $n^{-1/4+\varepsilon}$ rates for sums)
  - Csörgő-Révész construction
  - Komlós-Major-Tusnády "Hungarian construction"

Introduction

Strong invariance principle, independent case

- Last 50 years, on $\mathcal{X} = \mathbb{R}$ a lot of work to strengthen Donsker
  - Skorokhod embedding (allows up to $n^{-1/4+\varepsilon}$ rates for sums)
  - Csörgő-Révész construction
  - Komlós-Major-Tusnády "Hungarian construction"
  - refinements

Introduction

Strong invariance principle, independent case

- Last 50 years, on $\mathcal{X} = \mathbb{R}$ a lot of work to strengthen Donsker
  - Skorokhod embedding (allows up to $n^{-1/4+\varepsilon}$ rates for sums)
  - Csörgő-Révész construction
  - Komlós-Major-Tusnády "Hungarian construction"
  - refinements
- The **strong invariance principle**

Introduction

Strong invariance principle, independent case

- Last 50 years, on $\mathcal{X} = \mathbb{R}$ a lot of work to strengthen Donsker
  - Skorokhod embedding (allows up to $n^{-1/4+\varepsilon}$ rates for sums)
  - Csörgő-Révész construction
  - Komlós-Major-Tusnády "Hungarian construction"
  - refinements
- The **strong invariance principle**
  - on $(\Omega, \mathcal{T}, \mathbb{P})$ construct $\{X_n\}$ and **dependent** copies $\mathbb{G}_n$ of $\mathbb{G}$

$$\|\alpha_n - \mathbb{G}_n\|_{\mathcal{F}} = O(v_n) \quad a.s.$$

Introduction

Strong invariance principle, independent case

- Last 50 years, on $\mathcal{X} = \mathbb{R}$ a lot of work to strengthen Donsker
  - Skorokhod embedding (allows up to $n^{-1/4+\varepsilon}$ rates for sums)
  - Csörgő-Révész construction
  - Komlós-Major-Tusnády "Hungarian construction"
  - refinements
- The **strong invariance principle**
  - on $(\Omega, \mathcal{T}, \mathbb{P})$ construct $\{X_n\}$ and **dependent** copies $\mathbb{G}_n$ of $\mathbb{G}$

$$\|\alpha_n - \mathbb{G}_n\|_{\mathcal{F}} = O(v_n) \quad a.s.$$

  - implies the weak invariance principle

Introduction

Strong invariance principle, independent case

- Last 50 years, on $\mathcal{X} = \mathbb{R}$ a lot of work to strengthen Donsker
  - Skorokhod embedding (allows up to $n^{-1/4+\varepsilon}$ rates for sums)
  - Csörgő-Révész construction
  - Komlós-Major-Tusnády "Hungarian construction"
  - refinements
- The **strong invariance principle**
  - on $(\Omega, \mathcal{T}, \mathbb{P})$ construct $\{X_n\}$ and **dependent** copies $\mathbb{G}_n$ of $\mathbb{G}$

$$\|\alpha_n - \mathbb{G}_n\|_{\mathcal{F}} = O(v_n) \quad a.s.$$

  - implies the weak invariance principle
  - if sufficient probabilities, Lévy-Prohorov distance $d_L(\alpha_n, \mathbb{G}_n) = O(v_n)$

- The **strong approximation**

- The **strong approximation**
  - on $(\Omega, \mathcal{T}, \mathbb{P})$ construct $\{X_n\}$ and **independent** copies $\mathbb{G}_i^*$ of $\mathbb{G}$

$$\max_{1 \leq m \leq n} \left\| \sqrt{m}\, \alpha_m - \sum_{i=1}^{m} \mathbb{G}_i^* \right\|_{\mathcal{F}} = O(V_n) \quad a.s.$$

- The **strong approximation**
  - on $(\Omega, \mathcal{T}, \mathbb{P})$ construct $\{X_n\}$ and **independent** copies $\mathbb{G}_i^*$ of $\mathbb{G}$

$$\max_{1\leq m\leq n} \left\| \sqrt{m}\,\alpha_m - \sum_{i=1}^{m} \mathbb{G}_i^* \right\|_{\mathcal{F}} = O(V_n) \quad a.s.$$

  - implies

$$\|\alpha_n - \mathbb{G}_n\|_{\mathcal{F}} = O\left(\frac{V_n}{\sqrt{n}}\right) \quad a.s.$$

with a version $\mathbb{G}_n$ of $\mathbb{G}$ having a Kiefer-type structure

$$\mathbb{G}_n = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \mathbb{G}_i^*$$

- The **strong approximation**
  - on $(\Omega, \mathcal{T}, \mathbb{P})$ construct $\{X_n\}$ and **independent** copies $\mathbb{G}_i^*$ of $\mathbb{G}$

$$\max_{1 \leq m \leq n} \left\| \sqrt{m}\, \alpha_m - \sum_{i=1}^{m} \mathbb{G}_i^* \right\|_{\mathcal{F}} = O(V_n) \quad a.s.$$

  - implies

$$\|\alpha_n - \mathbb{G}_n\|_{\mathcal{F}} = O\left(\frac{V_n}{\sqrt{n}}\right) \quad a.s.$$

with a version $\mathbb{G}_n$ of $\mathbb{G}$ having a Kiefer-type structure

$$\mathbb{G}_n = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \mathbb{G}_i^*$$

  - slower $V_n / \sqrt{n} \gg v_n \to 0$ versus useful independence of $\mathbb{G}_i^*$

Introduction

Gaussian coupling, independent case

- Invariance/strong approximation are implied by a **Brownian coupling**

Introduction

Gaussian coupling, independent case

- Invariance/strong approximation are implied by a **Brownian coupling**
  - for deterministic $c_\theta = c_\theta(\mathcal{F}, P) > 0$

- Invariance/strong approximation are implied by a **Brownian coupling**
  - for deterministic $c_\theta = c_\theta(\mathcal{F}, P) > 0$
  - for deterministic $v_n = v_n(\mathcal{F}, P) \to 0$

- Invariance/strong approximation are implied by a **Brownian coupling**
    - for deterministic $c_\theta = c_\theta(\mathcal{F}, P) > 0$
    - for deterministic $v_n = v_n(\mathcal{F}, P) \to 0$
    - for any $\theta > 0$, $n \geq 1$ on Ω we can construct versions of $X_1, ..., X_n$ and $\mathbb{G}$ such that

$$\mathbb{P}\left(\|\alpha_n - \mathbb{G}\|_{\mathcal{F}} > c_\theta v_n\right) \leq \frac{1}{n^\theta}$$

- Limiting process $\mathbb{G}_\infty$ hard to couple with $X_1, ..., X_n$ and makes few sense at finite n

Introduction

Problem in the dependent case

- Limiting process $\mathbb{G}_\infty$ hard to couple with $X_1, \dots, X_n$ and makes few sense at finite n
- Covariances Γ_n (of α_n) $\to \Gamma_\infty$ (of $\mathbb{G}_\infty$) where

$$\begin{aligned}\Gamma_n(f,g) &= \mathbb{E}(f(X)g(X)) - \mathbb{E}(f(X))\mathbb{E}(g(X)) \\ &+ \sum_{i=2}^{n} \frac{n-i+1}{n} \left(\mathbb{E}\left(f(X_i)g(X_1)\right) - \mathbb{E}(f(X))\mathbb{E}(g(X \right. \\ &+ \sum_{i=2}^{n} \frac{n-i+1}{n} \left(\mathbb{E}\left(f(X_1)g(X_i)\right) - \mathbb{E}(f(X))\mathbb{E}(g(X \right.\end{aligned}$$

Introduction

Problem in the dependent case

- Limiting process $\mathbb{G}_\infty$ hard to couple with $X_1, \ldots, X_n$ and makes few sense at finite n
- Covariances Γ_n (of α_n) $\rightarrow \Gamma_\infty$ (of $\mathbb{G}_\infty$) where

$$\begin{aligned}\Gamma_n(f,g) &= \mathbb{E}(f(X)g(X)) - \mathbb{E}(f(X))\mathbb{E}(g(X)) \\ &+ \sum_{i=2}^{n} \frac{n-i+1}{n} \left(\mathbb{E}\left(f(X_i)g(X_1)\right) - \mathbb{E}(f(X))\mathbb{E}(g(X\right. \\ &+ \sum_{i=2}^{n} \frac{n-i+1}{n} \left(\mathbb{E}\left(f(X_1)g(X_i)\right) - \mathbb{E}(f(X))\mathbb{E}(g(X\right.\end{aligned}$$

- Find an intermediate Gaussian process $\mathbb{G}_n^*$ with covariance close to Γ_n and Γ_∞ such that

Introduction

Problem in the dependent case

- Limiting process $\mathbb{G}_\infty$ hard to couple with $X_1, ..., X_n$ and makes few sense at finite n
- Covariances Γ_n (of α_n) $\to \Gamma_\infty$ (of $\mathbb{G}_\infty$) where

$$\begin{aligned}\Gamma_n(f,g) &= \mathbb{E}(f(X)g(X)) - \mathbb{E}(f(X))\mathbb{E}(g(X)) \\ &+ \sum_{i=2}^{n} \frac{n-i+1}{n} \left(\mathbb{E}\left(f(X_i)g(X_1)\right) - \mathbb{E}(f(X))\mathbb{E}(g(X \\ &+ \sum_{i=2}^{n} \frac{n-i+1}{n} \left(\mathbb{E}\left(f(X_1)g(X_i)\right) - \mathbb{E}(f(X))\mathbb{E}(g(X\end{aligned}$$

- Find an intermediate Gaussian process $\mathbb{G}_n^*$ with covariance close to Γ_n and Γ_∞ such that
  - for c_θ, $v_n \to 0$, any $\theta > 0$, all $n \geq 1$ on Ω construct α_n and $\mathbb{G}_n^*$

$$\mathbb{P}\left(\|\alpha_n - \mathbb{G}_n^*\|_{\mathcal{F}} > c_\theta v_n\right) \leq \frac{1}{n^\theta}$$

Introduction

Gaussian coupling, dependent case

- Find an intermediate Gaussian process $\mathbb{G}_n^*$ with covariance Γ_n^* close to Γ_n and Γ_∞ such that

- Find an intermediate Gaussian process $\mathbb{G}_n^*$ with covariance Γ_n^* close to Γ_n and Γ_∞ such that
    - for deterministic c_θ and $v_n \to 0$

- Find an intermediate Gaussian process $\mathbb{G}_n^*$ with covariance Γ_n^* close to Γ_n and Γ_∞ such that
  - for deterministic c_θ and $v_n \to 0$
  - for any $\theta > 0$, $n \geq 1$ on Ω construct $X_1, \dots, X_n$ and version $\mathbb{G}_n^*$

$$\mathbb{P}\left(\|\alpha_n - \mathbb{G}_n^*\|_{\mathcal{F}} > c_\theta v_n\right) \leq \frac{1}{n^\theta}$$

- Find an intermediate Gaussian process $\mathbb{G}_n^*$ with covariance Γ_n^* close to Γ_n and Γ_∞ such that
  - for deterministic c_θ and $v_n \longrightarrow 0$
  - for any $\theta > 0$, $n \geq 1$ on Ω construct $X_1, ..., X_n$ and version $\mathbb{G}_n^*$

$$\mathbb{P}\left(\|\alpha_n - \mathbb{G}_n^*\|_{\mathcal{F}} > c_\theta v_n\right) \leq \frac{1}{n^\theta}$$

- Try to keep the **covariance bias** small,

$$\max\left(\sup_{f,g\in\mathcal{F}} |\Gamma_n^* - \Gamma_n|(f,g), \sup_{f,g\in\mathcal{F}} |\Gamma_n^* - \Gamma_\infty|(f,g)\right) < b_n$$

Introduction

Gaussian coupling, dependent case

- Find an intermediate Gaussian process $\mathbb{G}_n^*$ with covariance Γ_n^* close to Γ_n and Γ_∞ such that
  - for deterministic c_θ and $v_n \to 0$
  - for any $\theta > 0$, $n \geq 1$ on Ω construct $X_1, ..., X_n$ and version $\mathbb{G}_n^*$

$$\mathbb{P}\left(\|\alpha_n - \mathbb{G}_n^*\|_{\mathcal{F}} > c_\theta v_n\right) \leq \frac{1}{n^\theta}$$

- Try to keep the **covariance bias** small,

$$\max\left(\sup_{f,g\in\mathcal{F}} |\Gamma_n^* - \Gamma_n|(f,g), \sup_{f,g\in\mathcal{F}} |\Gamma_n^* - \Gamma_\infty|(f,g)\right) < b_n$$

- Independent case only : $\Gamma_n^* = \Gamma_n = \Gamma_\infty$.

Introduction

Gaussian coupling, dependent case

- Find an intermediate Gaussian process $\mathbb{G}_n^*$ with covariance Γ_n^* close to Γ_n and Γ_∞ such that
  - for deterministic c_θ and $v_n \to 0$
  - for any $\theta > 0$, $n \geq 1$ on Ω construct $X_1, \ldots, X_n$ and version $\mathbb{G}_n^*$

$$\mathbb{P}\left(\|\alpha_n - \mathbb{G}_n^*\|_{\mathcal{F}} > c_\theta v_n\right) \leq \frac{1}{n^\theta}$$

- Try to keep the **covariance bias** small,

$$\max\left(\sup_{f,g\in\mathcal{F}} |\Gamma_n^* - \Gamma_n|(f,g), \sup_{f,g\in\mathcal{F}} |\Gamma_n^* - \Gamma_\infty|(f,g)\right) < b_n$$

- Independent case only : $\Gamma_n^* = \Gamma_n = \Gamma_\infty$.
- Otherwise, **trade off** between v_n, $n^{-\theta}$, b_n.

- **Application 1.** Helps in studying the estimation of

$$f_* = \arg\min_{\mathcal{F}} Pf$$

by means of

$$f_n = \arg\min_{\mathcal{F}} P_n f = \arg\min_{\mathcal{F}} \left\{ Pf + \frac{\mathbb{G}}{\sqrt{n}} f + \frac{\alpha_n - \mathbb{G}}{\sqrt{n}} f \right\}$$

- **Application 1.** Helps in studying the estimation of

$$f_* = \arg\min_{\mathcal{F}} Pf$$

by means of

$$f_n = \arg\min_{\mathcal{F}} P_n f = \arg\min_{\mathcal{F}} \left\{ Pf + \frac{\mathbb{G}}{\sqrt{n}} f + \frac{\alpha_n - \mathbb{G}}{\sqrt{n}} f \right\}$$

- **Contrasts** $\mathcal{F} = \{f(\cdot, g) : g \in \mathcal{G}\}$, **risk** Pf

- **Application 1.** Helps in studying the estimation of

$$f_* = \arg\min_{\mathcal{F}} Pf$$

by means of

$$f_n = \arg\min_{\mathcal{F}} P_n f = \arg\min_{\mathcal{F}} \left\{ Pf + \frac{\mathbb{G}}{\sqrt{n}} f + \frac{\alpha_n - \mathbb{G}}{\sqrt{n}} f \right\}$$

- **Contrasts** $\mathcal{F} = \{f(\cdot, g) : g \in \mathcal{G}\}$, **risk** Pf
  - classification : $X = (Y, Z)$, risk $Pf = \mathbb{P}(g(Y) \neq Z)$

- **Application 1.** Helps in studying the estimation of

$$f_* = \arg\min_{\mathcal{F}} Pf$$

by means of

$$f_n = \arg\min_{\mathcal{F}} P_n f = \arg\min_{\mathcal{F}} \left\{ Pf + \frac{\mathbb{G}}{\sqrt{n}} f + \frac{\alpha_n - \mathbb{G}}{\sqrt{n}} f \right\}$$

- **Contrasts** $\mathcal{F} = \{f(\cdot, g) : g \in \mathcal{G}\}$, **risk** Pf
  - classification : $X = (Y, Z)$, risk $Pf = \mathbb{P}(g(Y) \neq Z)$
  - regression : $X = (Y, Z)$, risk $Pf = \mathbb{E}(g(Y) - Z)^2$

Applications

Empirical Risk Minimization

- **Application 1.** Helps in studying the estimation of

$$f_* = \arg\min_{\mathcal{F}} Pf$$

by means of

$$f_n = \arg\min_{\mathcal{F}} P_n f = \arg\min_{\mathcal{F}} \left\{ Pf + \frac{\mathbb{G}}{\sqrt{n}} f + \frac{\alpha_n - \mathbb{G}}{\sqrt{n}} f \right\}$$

- **Contrasts** $\mathcal{F} = \{f(\cdot, g) : g \in \mathcal{G}\}$, **risk** Pf
  - classification : $X = (Y, Z)$, risk $Pf = \mathbb{P}(g(Y) \neq Z)$
  - regression : $X = (Y, Z)$, risk $Pf = \mathbb{E}(g(Y) - Z)^2$
  - likelihood : X has density g_θ, risk $Pf = -\mathbb{E}(\log g_\theta(X))$

- **Application 1 (continued).** The excess risk is

$$P(f_n - f_*) = P_n(f_n - f_*) - \frac{\mathbb{G}}{\sqrt{n}}(f_n - f_*) + \mathit{errors}$$

- **Application 1 (continued).** The excess risk is

$$P(f_n - f_*) = P_n(f_n - f_*) - \frac{\mathbb{G}}{\sqrt{n}}(f_n - f_*) + errors$$

  - allows to find **excess risk limiting distribution**

- **Application 1 (continued).** The excess risk is

$$P(f_n - f_*) = P_n(f_n - f_*) - \frac{\mathbb{G}}{\sqrt{n}}(f_n - f_*) + errors$$

  - allows to find **excess risk limiting distribution**
  - alternative approach of **model selection**

- **Application 1 (continued).** The excess risk is

$$P(f_n - f_*) = P_n(f_n - f_*) - \frac{\mathbb{G}}{\sqrt{n}}(f_n - f_*) + \mathit{errors}$$

  - allows to find **excess risk limiting distribution**
  - alternative approach of **model selection**
  - helps understanding **optimal penalties** (to compete increments of $\mathbb{G}/\sqrt{n}$)

- **Application 2.** Given $\xi_n < 1/\sqrt{n}$ and consider the almost minimizers

$$\mathcal{F}_n = \{f \in \mathcal{F} : P_n f < P_n f_n + \xi_n\}$$

Under entropy and/or margin conditions Berthet-Saumard studied

$$\begin{aligned} \mathbb{P}\left(diam_{\mathbb{L}_2}\mathcal{F}_n > d_n^+\right) &< n^{-\theta} \\ \mathbb{P}\left(diam_{\mathbb{L}_2}\mathcal{F}_n < d_n^-\right) &< n^{-\theta} \\ \mathbb{P}\left(d_n^- < \sup_{f,g\in\mathcal{F}_n} |Pf - Pg| > d_n^+\right) &< n^{-\theta} \end{aligned}$$

- **Application 2.** Given $\xi_n < 1/\sqrt{n}$ and consider the almost minimizers

$$\mathcal{F}_n = \{f \in \mathcal{F} : P_n f < P_n f_n + \xi_n\}$$

Under entropy and/or margin conditions Berthet-Saumard studied

$$\begin{aligned} \mathbb{P}\left(diam_{\mathbb{L}_2}\mathcal{F}_n > d_n^+\right) &< n^{-\theta} \\ \mathbb{P}\left(diam_{\mathbb{L}_2}\mathcal{F}_n < d_n^-\right) &< n^{-\theta} \\ \mathbb{P}\left(d_n^- < \sup_{f,g\in\mathcal{F}_n} |Pf - Pg| > d_n^+\right) &< n^{-\theta} \end{aligned}$$

- The Gaussian coupling starts the proof but covariance is needed later.

- **Application 3a.** For $f \in \mathcal{F}$ find the limiting process of

$$\frac{1}{n_1 \dots n_k} \sum_{1 \leqslant i_1 \leqslant n_1} \dots \sum_{1 \leqslant i_k \leqslant n_k} f(X_{i_1}^{(1)}, \dots, X_{i_k}^{(k)}) - \mathbb{E}(f(X_1^{(1)}, \dots, X_1^{(k)}))$$

at rate $\sqrt{\min_{j \leq k} n_j}$ in Berthet-Paroux current work

- **Application 3a.** For $f \in \mathcal{F}$ find the limiting process of

$$\frac{1}{n_1 \dots n_k} \sum_{1 \leqslant i_1 \leqslant n_1} \dots \sum_{1 \leqslant i_k \leqslant n_k} f(X_{i_1}^{(1)}, \dots, X_{i_k}^{(k)}) - \mathbb{E}(f(X_1^{(1)}, \dots, X_1^{(k)}))$$

at rate $\sqrt{\min_{j \leq k} n_j}$ in Berthet-Paroux current work

- gives finite n_j approximations by the limiting or intermediate Gaussian process

- **Application 3a.** For $f \in \mathcal{F}$ find the limiting process of

$$\frac{1}{n_1 \dots n_k} \sum_{1 \leqslant i_1 \leqslant n_1} \dots \sum_{1 \leqslant i_k \leqslant n_k} f(X_{i_1}^{(1)}, \dots, X_{i_k}^{(k)}) - \mathbb{E}(f(X_1^{(1)}, \dots, X_1^{(k)}))$$

at rate $\sqrt{\min_{j \leq k} n_j}$ in Berthet-Paroux current work

  - gives finite n_j approximations by the limiting or intermediate Gaussian process
  - limiting processes indexed by **classes of conditional expectations** of $f \in \mathcal{F}$

Applications

k-sample U-processes

- **Application 3a.** For $f \in \mathcal{F}$ find the limiting process of

$$\frac{1}{n_1 \dots n_k} \sum_{1 \leqslant i_1 \leqslant n_1} \cdots \sum_{1 \leqslant i_k \leqslant n_k} f(X_{i_1}^{(1)}, \dots, X_{i_k}^{(k)}) - \mathbb{E}(f(X_1^{(1)}, \dots, X_1^{(k)}))$$

  at rate $\sqrt{\min_{j \leq k} n_j}$ in Berthet-Paroux current work
  - gives finite n_j approximations by the limiting or intermediate Gaussian process
  - limiting processes indexed by **classes of conditional expectations** of $f \in \mathcal{F}$
- **Application 3b.** Likewise for hybrid or randomly weighted empirical processes, convergence to a sum of Brownian motions and Bridges indexed by conditional expectations.

- **Application 4.** Sets $\mathcal{C}$ in $\mathbb{R}^d$, sample $X_1, ..., X_n$ density f

- **Application 4.** Sets $\mathcal{C}$ in $\mathbb{R}^d$, sample $X_1, ..., X_n$ density f
  - target a λ-level set $C_* = \{f > \lambda\}$

$$
\begin{aligned}
C_* &= \arg\max_{C \in \mathcal{C}} \{P(C) - \lambda\mu(C)\} \\
&= \arg\max_{C \in \mathcal{C}} \{P(C) : \mu(C) \leq v_\lambda\} \\
&= \arg\min_{C \in \mathcal{C}} \{\mu(C) : P(C) \geq p_\lambda\}
\end{aligned}
$$

- **Application 4.** Sets $\mathcal{C}$ in $\mathbb{R}^d$, sample $X_1, ..., X_n$ density f
  - target a λ-level set $C_* = \{f > \lambda\}$

$$
\begin{aligned}
C_* &= \arg\max_{C \in \mathcal{C}} \{P(C) - \lambda\mu(C)\} \\
&= \arg\max_{C \in \mathcal{C}} \{P(C) : \mu(C) \leq v_\lambda\} \\
&= \arg\min_{C \in \mathcal{C}} \{\mu(C) : P(C) \geq p_\lambda\}
\end{aligned}
$$

  - Berthet-Einmahl are studying the **joint limiting shapes** of $C_{k,n} \Delta C_*$

$$
\begin{aligned}
C_{1,n} &= \arg\max_{C \in \mathcal{C}} \{P_n(C) - \lambda\mu(C)\} \\
C_{2,n} &= \arg\max_{C \in \mathcal{C}} \{P_n(C) : \mu(C) \leq v_\lambda\} \\
C_{3,n} &= \arg\min_{C \in \mathcal{C}} \{\mu(C) : P_n(C) \geq p_\lambda\}
\end{aligned}
$$

- **Application 5a.** Robust algorithms for learning theory in adversarial environment (bandits, exploit or explore compromise). Decisions are randomized but depend from the past whereas regret involves also the close unknown futur.

- **Application 5a.** Robust algorithms for learning theory in adversarial environment (bandits, exploit or explore compromise). Decisions are randomized but depend from the past whereas regret involves also the close unknown futur.
- **Application 5b.** Control of bias, risk and regret (when a cost is involved) when using model selection among random small dimension subfields to estimate regression.

Couplings

KMT, optimal rate

- Let $P = U(0,1)$, $\mathcal{F} = \left\{ \mathbb{I}_{[0,t]} : 0 < t < 1 \right\}$, $v_n = \dfrac{\log n}{\sqrt{n}}$

Couplings

KMT, optimal rate

- Let $P = U(0,1)$, $\mathcal{F} = \left\{ \mathbb{I}_{[0,t]} : 0 < t < 1 \right\}$, $v_n = \dfrac{\log n}{\sqrt{n}}$
- By KMT on Ω there exists $\{X_n\}$ and Brownian Bridges $\{\mathbb{G}_n\}$

$$\begin{aligned} \mathbb{P}\left(\sqrt{n}\,\|\alpha_n - \mathbb{G}_n\|_{\mathcal{F}} \geq c_1\lambda + \log n\right) &\leq c_2 \exp\left(-c_3\lambda\right) \\ \mathbb{P}\left(\|\alpha_n - \mathbb{G}_n\|_{\mathcal{F}} \geq c_\theta v_n\right) &\leq \frac{1}{n^\theta} \end{aligned}$$

Couplings

KMT, optimal rate

- Let $P = U(0,1)$, $\mathcal{F} = \left\{ \mathbb{I}_{[0,t]} : 0 < t < 1 \right\}$, $v_n = \dfrac{\log n}{\sqrt{n}}$
- By KMT on Ω there exists $\{X_n\}$ and Brownian Bridges $\{\mathbb{G}_n\}$

$$\begin{aligned} \mathbb{P}\left(\sqrt{n}\left\|\alpha_n - \mathbb{G}_n\right\|_{\mathcal{F}} \geq c_1\lambda + \log n\right) &\leq c_2 \exp\left(-c_3\lambda\right) \\ \mathbb{P}\left(\left\|\alpha_n - \mathbb{G}_n\right\|_{\mathcal{F}} \geq c_\theta v_n\right) &\leq \frac{1}{n^\theta} \end{aligned}$$

- Indeed

$$\text{for } \textbf{some } \{X_n, \mathbb{G}_n\} \quad \limsup_{n\to\infty} \frac{\sqrt{n}}{\log n}\left\|\alpha_n - \mathbb{G}_n\right\|_{\mathcal{F}} \leq 12 \quad a.s.$$

$$\text{for } \textbf{any } \{X_n, \mathbb{G}_n\} \quad \liminf_{n\to\infty} \frac{\sqrt{n}}{\log n}\left\|\alpha_n - \mathbb{G}_n\right\|_{\mathcal{F}} \geq \frac{1}{6} \quad a.s.$$

Couplings

Tools

- Two approaches

Couplings

Tools

- Two approaches
  - i) $\mathcal{F}$ only controled by **entropy** and $P = \mathbb{P}^X$ **free**

Couplings

Tools

- Two approaches
  - i) $\mathcal{F}$ only controled by **entropy** and $P = \mathbb{P}^X$ **free**
  - entropy (uniform, bracketing, random) : if **large** then **more mixing** asked

Couplings

Tools

- Two approaches
  - i) $\mathcal{F}$ only controled by **entropy** and $P = \mathbb{P}^X$ **free**
  - entropy (uniform, bracketing, random) : if **large** then **more mixing** asked
  - ii) better specified **geometry** of $\mathcal{F}$ with respect to P

Couplings

Tools

- Two approaches
  - i) $\mathcal{F}$ only controled by **entropy** and $P = \mathbb{P}^X$ **free**
  - entropy (uniform, bracketing, random) : if **large** then **more mixing** asked
  - ii) better specified **geometry** of $\mathcal{F}$ with respect to P
  - new Donsker classes ($\mathcal{F}$ built to control the rate by P) : more tricky under mixing

Couplings

Tools

- Two approaches
  - i) $\mathcal{F}$ only controled by **entropy** and $P = \mathbb{P}^X$ **free**
  - entropy (uniform, bracketing, random) : if **large** then **more mixing** asked
  - ii) better specified **geometry** of $\mathcal{F}$ with respect to P
  - new Donsker classes ($\mathcal{F}$ built to control the rate by P) : more tricky under mixing
- **Tools** : Gaussian coupling in $\mathbb{R}^d$, moment inequalities, concentration inequalities, symetrization, Dudley-Philipp's strong embedding, blocking to apply Berbee lemma.

Couplings
Tools

- Two approaches
  - i) $\mathcal{F}$ only controled by **entropy** and $P = \mathbb{P}^X$ **free**
  - entropy (uniform, bracketing, random) : if **large** then **more mixing** asked
  - ii) better specified **geometry** of $\mathcal{F}$ with respect to P
  - new Donsker classes ($\mathcal{F}$ built to control the rate by P) : more tricky under mixing
- **Tools** : Gaussian coupling in $\mathbb{R}^d$, moment inequalities, concentration inequalities, symetrization, Dudley-Philipp's strong embedding, blocking to apply Berbee lemma.
- **Simple technique** : relies on previous tools from Dudley, Philipp, Berkes, Kuelbs, Dehling, Pollard, Giné, Talagrand, Zaitsev, Massart, Koltchinskii, Rio, Einmahl, Mason, among many others (independent case) and the authors in dependent data...

Uniform entropy

Independent case

- By Dudley-Philipp if $\mathcal{F}$ is P-Donsker then

$$v_n = o(\sqrt{\log \log n})$$

Uniform entropy

Independent case

- By Dudley-Philipp if $\mathcal{F}$ is P-Donsker then

$$v_n = o(\sqrt{\log\log n})$$

- This is the case if

Uniform entropy

Independent case

- By Dudley-Philipp if $\mathcal{F}$ is P-Donsker then

$$v_n = o(\sqrt{\log \log n})$$

- This is the case if
  - $\mathcal{F}$ is pointwise measurable, with envelope $F \in \mathbb{L}_2(P)$

Uniform entropy

Independent case

- By Dudley-Philipp if $\mathcal{F}$ is P-Donsker then

$$v_n = o(\sqrt{\log \log n})$$

- This is the case if
  - $\mathcal{F}$ is pointwise measurable, with envelope $F \in \mathbb{L}_2(P)$
  - for any probability measure P

$$N\left(\varepsilon\sqrt{PF^2}, \mathcal{F}, d_P\right) \leq \phi(\varepsilon), \quad 0 < \varepsilon < 1$$

Uniform entropy

Independent case

- By Dudley-Philipp if $\mathcal{F}$ is P-Donsker then

$$v_n = o(\sqrt{\log \log n})$$

- This is the case if
  - $\mathcal{F}$ is pointwise measurable, with envelope $F \in \mathbb{L}_2(P)$
  - for any probability measure P

$$N\left(\varepsilon\sqrt{PF^2}, \mathcal{F}, d_P\right) \leq \phi(\varepsilon), \quad 0 < \varepsilon < 1$$

  - ϕ is increasing, $\phi(\varepsilon) \to \infty$ as $\varepsilon \to 0$ and

$$\int_{[0,1]} \sqrt{\log \phi} < \infty$$

- Introduce

$$J_\phi(\varepsilon) = \int_0^\varepsilon \sqrt{\log \phi}, \quad \Psi_\phi(\varepsilon) = \frac{\varepsilon}{\phi^{5/2}(\varepsilon)}$$

so that $J_\phi \circ \Psi_\phi^{-1}(\varepsilon) \gg \varepsilon \quad$ as $\varepsilon \to 0$

Uniform entropy

Independence case

- Introduce

$$J_\phi(\varepsilon) = \int_0^\varepsilon \sqrt{\log \phi}, \quad \Psi_\phi(\varepsilon) = \frac{\varepsilon}{\phi^{5/2}(\varepsilon)}$$

so that $J_\phi \circ \Psi_\phi^{-1}(\varepsilon) \gg \varepsilon \quad$ as $\varepsilon \to 0$

- **Theorem U1.** We can construct $\{X_n\}$ and versions $\{\mathbb{G}_n\}$ of $\mathbb{G}$ on Ω such that

$$\|\alpha_n - \mathbb{G}_n\|_{\mathcal{F}} = O_{a.s.}\left(J_\phi \circ \Psi_\phi^{-1}\left(\sqrt{\frac{\log n}{n}}\right)\right)$$

Uniform entropy

Independence case

- Introduce

$$J_\phi(\varepsilon) = \int_0^\varepsilon \sqrt{\log \phi}, \quad \Psi_\phi(\varepsilon) = \frac{\varepsilon}{\phi^{5/2}(\varepsilon)}$$

so that $J_\phi \circ \Psi_\phi^{-1}(\varepsilon) \gg \varepsilon \quad \text{as } \varepsilon \to 0$

- **Theorem U1.** We can construct $\{X_n\}$ and versions $\{\mathbb{G}_n\}$ of $\mathbb{G}$ on Ω such that

$$\|\alpha_n - \mathbb{G}_n\|_{\mathcal{F}} = O_{a.s.}\left(J_\phi \circ \Psi_\phi^{-1}\left(\sqrt{\frac{\log n}{n}}\right)\right)$$

- This is uniform in P for a **fixed** $\mathcal{F}$

Uniform entropy

Polynomialy decreasing, independent case

- Assume

Uniform entropy

Polynomialy decreasing, independent case

- Assume
  - $\mathcal{F}$ is pointwise measurable, uniformly bounded

Uniform entropy

Polynomialy decreasing, independent case

- Assume
  - $\mathcal{F}$ is pointwise measurable, uniformly bounded
  - for $\nu > 0$ and any probability measure P

$$N\left(\varepsilon\sqrt{PF^2}, \mathcal{F}, d_P\right) \leq \frac{1}{\varepsilon^\nu}, \quad 0 < \varepsilon < 1$$

Uniform entropy

Polynomialy decreasing, independent case

- Assume
  - $\mathcal{F}$ is pointwise measurable, uniformly bounded
  - for $\nu > 0$ and any probability measure P

$$N\left(\varepsilon\sqrt{PF^2}, \mathcal{F}, d_P\right) \leq \frac{1}{\varepsilon^{\nu}}, \quad 0 < \varepsilon < 1$$

  - if $\mathcal{F} = \{\mathbb{I}_C : C \in \mathcal{C}\}$, $\mathcal{C}$ *VC-class* then $\nu = VC - 1$

Uniform entropy

Polynomialy decreasing, independent case

- Assume
  - $\mathcal{F}$ is pointwise measurable, uniformly bounded
  - for $\nu > 0$ and any probability measure P

$$N\left(\varepsilon\sqrt{PF^2}, \mathcal{F}, d_P\right) \leq \frac{1}{\varepsilon^\nu}, \quad 0 < \varepsilon < 1$$

  - if $\mathcal{F} = \{\mathbb{I}_C : C \in \mathcal{C}\}$, $\mathcal{C}$ *VC-class* then $\nu = VC - 1$
- **Corollary.** We can construct $\{X_n\}$ and $\{\mathbb{G}_n\}$ on Ω such that

$$\mathbb{P}\left(\|\alpha_n - \mathbb{G}\|_{\mathcal{F}} > c_\theta \frac{(\log n)^{\tau_0}}{n^\tau}\right) \leq \frac{1}{n^\theta}$$

where

$$\tau = \frac{1}{2+5\nu}, \quad \tau_0 = \frac{4+5\nu}{4+10\nu}$$

Uniform entropy

Polynomialy decreasing, independent case

- Assume

Uniform entropy

Polynomialy decreasing, independent case

- Assume
  - $\mathcal{F}$ is pointwise measurable, uniformly bounded

Uniform entropy

Polynomialy decreasing, independent case

- Assume
  - $\mathcal{F}$ is pointwise measurable, uniformly bounded
  - for $\nu_0 > 0$ and any probability measure P

$$N\left(\varepsilon\sqrt{PF^2}, \mathcal{F}, d_P\right) \leq \frac{1}{\varepsilon^{\nu}}, \quad 0 < \varepsilon < 1$$

- Assume
  - $\mathcal{F}$ is pointwise measurable, uniformly bounded
  - for $\nu_0 > 0$ and any probability measure P

$$N\left(\varepsilon\sqrt{PF^2}, \mathcal{F}, d_P\right) \leq \frac{1}{\varepsilon^{\nu}}, \quad 0 < \varepsilon < 1$$

- **Corollary.** We can construct $\{X_n\}$ and **i.i.d.** $\{\mathbb{G}_n^*\}$ on Ω

$$\frac{1}{\sqrt{n}} \max_{1 \leq m \leq n} \left\| \sqrt{m}\,\alpha_m - \sum_{i=1}^{m} \mathbb{G}_i^* \right\|_{\mathcal{F}} = O_{a.s.}\left(\frac{(\log n)^{\tau_0}}{n^{\tau(\alpha)}}\right)$$

$$\tau(\alpha) = \frac{\alpha\tau - 1/2}{1+\alpha} < \tau = \frac{1}{2+5\nu_0}, \quad \alpha \in \left(\frac{1}{2\tau}, \frac{1}{\tau}\right)$$

Uniform entropy, dependent case

Polynomialy decreasing, exponentialy mixing

- Assume

$$\phi_k < \exp(-\theta k), \quad k \in \mathbb{N}$$

Uniform entropy, dependent case

Polynomialy decreasing, exponentialy mixing

- Assume

$$\phi_k < \exp(-\theta k), \quad k \in \mathbb{N}$$

- **Theorem U2.** We can construct $\{X_n\}$ and $\{\mathbb{G}_n^*\}$ on Ω,

$$\mathbb{P}\left(\|\alpha_n - \mathbb{G}_n^*\|_{\mathcal{F}} > c_\theta v_n\right) \leq \frac{1}{n^\theta}$$

with rate

$$v_n = C_\lambda n^{-\frac{1}{6+5v}} (\log n)^{\frac{12+5v}{12+10v}}$$

and covariance bias

$$b_n = C_0 n^{-\frac{2}{6+5v}} (\log n)^{\frac{6}{6+5v}}, \quad C_0 > 0.$$

Uniform entropy, dependent case

Polynomialy decreasing, exponentialy mixing

- Assume
$$\phi_k < \exp(-\theta k), \quad k \in \mathbb{N}$$
- **Theorem U2.** We can construct $\{X_n\}$ and $\{\mathbb{G}_n^*\}$ on Ω,
$$\mathbb{P}\left(\|\alpha_n - \mathbb{G}_n^*\|_{\mathcal{F}} > c_\theta v_n\right) \leq \frac{1}{n^\theta}$$
with rate
$$v_n = C_\lambda n^{-\frac{1}{6+5v}} (\log n)^{\frac{12+5v}{12+10v}}$$
and covariance bias
$$b_n = C_0 n^{-\frac{2}{6+5v}} (\log n)^{\frac{6}{6+5v}}, \quad C_0 > 0.$$
- Even very strong mixing do not interpolate with independence,
$$n^{-\frac{1}{6+5v}} \gg n^{-\frac{1}{2+5v}}$$

Uniform entropy, dependent case

Polynomialy decreasing, polynomialy mixing

- Assume $(\gamma > 1)$

$$\phi_k < k^{-\gamma}, \quad k \in \mathbb{N}$$

Uniform entropy, dependent case

Polynomialy decreasing, polynomialy mixing

- Assume ($\gamma > 1$)

$$\phi_k < k^{-\gamma}, \quad k \in \mathbb{N}$$

- **Theorem U3.** We can construct $\{X_n\}$ and $\{\mathbb{G}_n^*\}$ on Ω,

$$\mathbb{P}\left(\|\alpha_n - \mathbb{G}_n^*\|_{\mathcal{F}} > c_\theta v_n\right) \leq n^{1-(1+\gamma)w}$$

with rate

$$v_n = C_\lambda n^{-\frac{1-2w}{6+5v}} (\log n)^{\frac{12+5v}{12+10v}}$$

and covariance bias

$$b_n = n^{-\frac{2-4w}{6+5v}} (\log n)^{\frac{6}{6+5v}}$$

where

$$w = \frac{2}{17v^2 + 17v + 10}$$

whence a strong invariance principle if $\gamma > 17v^2 + 17v + 9$.

Uniform entropy

Class of cubes, any P

- **Exemple.** Cubes $\mathcal{F} = \left\{ \mathbb{I}_{[s,t]} : s, t \in \mathcal{X} = [0,1]^d \right\}$

Uniform entropy

Class of cubes, any P

- **Exemple.** Cubes $\mathcal{F} = \left\{ \mathbb{I}_{[s,t]} : s, t \in \mathcal{X} = [0,1]^d \right\}$
  - approximated by quadrants, VC-index $2d + 1$

Uniform entropy

Class of cubes, any P

- **Exemple.** Cubes $\mathcal{F} = \left\{ \mathbb{I}_{[s,t]} : s, t \in \mathcal{X} = [0,1]^d \right\}$
  - approximated by quadrants, VC-index $2d + 1$
  - by Theorem U1

$$\|\alpha_n - \mathbb{G}_n\|_{\mathcal{F}} = O_{a.s.}\left(\frac{(\log n)^{\cdots}}{n^{1/(7+10d)}} \right)$$

Uniform entropy

Class of cubes, any P

- **Exemple.** Cubes $\mathcal{F} = \left\{ \mathbb{I}_{[s,t]} : s,t \in \mathcal{X} = [0,1]^d \right\}$
  - approximated by quadrants, VC-index $2d+1$
  - by Theorem U1

$$\|\alpha_n - \mathbb{G}_n\|_{\mathcal{F}} = O_{a.s.}\left(\frac{(\log n)^{\cdots}}{n^{1/(7+10d)}}\right)$$

  - by Theorem U2

$$\|\alpha_n - \mathbb{G}_n^*\|_{\mathcal{F}} = O_{a.s.}\left(\frac{(\log n)^{\cdots}}{n^{1/(11+10d)}}\right)$$

Uniform entropy

Class of cubes, any P

- **Exemple.** Cubes $\mathcal{F} = \left\{ \mathbb{I}_{[s,t]} : s,t \in \mathcal{X} = [0,1]^d \right\}$
  - approximated by quadrants, VC-index $2d+1$
  - by Theorem U1

$$\|\alpha_n - \mathbb{G}_n\|_{\mathcal{F}} = O_{a.s.}\left(\frac{(\log n)^{\cdots}}{n^{1/(7+10d)}}\right)$$

  - by Theorem U2

$$\|\alpha_n - \mathbb{G}_n^*\|_{\mathcal{F}} = O_{a.s.}\left(\frac{(\log n)^{\cdots}}{n^{1/(11+10d)}}\right)$$

  - **distribution free**, but not dimension free, first of the kind

Uniform entropy

Class of spheres, any P

- **Exemple.** Spheres $\mathcal{F} = \{\mathbb{I}_C : C \in \mathcal{S}_d\}$ in $\mathcal{X} = [0,1]^d$

Uniform entropy

Class of spheres, any P

- **Exemple.** Spheres $\mathcal{F} = \{\mathbb{I}_C : C \in \mathcal{S}_d\}$ in $\mathcal{X} = [0,1]^d$
  - VC-index $d + 2$

Uniform entropy

Class of spheres, any P

- **Exemple.** Spheres $\mathcal{F} = \{\mathbb{I}_C : C \in \mathcal{S}_d\}$ in $\mathcal{X} = [0,1]^d$
  - VC-index $d + 2$
  - by Theorem U1, for **any** P

$$\|\alpha_n - \mathbb{G}_n\|_{\mathcal{F}} = O_{a.s.}\left(\frac{(\log n)^{\cdots}}{n^{1/(12+10d)}}\right)$$

Uniform entropy

Class of spheres, any P

- **Exemple.** Spheres $\mathcal{F} = \{\mathbb{I}_C : C \in \mathcal{S}_d\}$ in $\mathcal{X} = [0,1]^d$
  - VC-index $d+2$
  - by Theorem U1, for **any** P

$$\|\alpha_n - \mathbb{G}_n\|_{\mathcal{F}} = O_{a.s.}\left(\frac{(\log n)^{\cdots}}{n^{1/(12+10d)}}\right)$$

  - by Theorem U2, for **any** P

$$\|\alpha_n - \mathbb{G}_n^*\|_{\mathcal{F}} = O_{a.s.}\left(\frac{(\log n)^{\cdots}}{n^{1/(16+10d)}}\right)$$

Uniform entropy

Class of spheres, any P

- **Exemple.** Spheres $\mathcal{F} = \{\mathbb{I}_C : C \in \mathcal{S}_d\}$ in $\mathcal{X} = [0,1]^d$
  - VC-index $d+2$
  - by Theorem U1, for **any** P

$$\|\alpha_n - \mathbb{G}_n\|_{\mathcal{F}} = O_{a.s.}\left(\frac{(\log n)^{\cdots}}{n^{1/(12+10d)}}\right)$$

  - by Theorem U2, for **any** P

$$\|\alpha_n - \mathbb{G}_n^*\|_{\mathcal{F}} = O_{a.s.}\left(\frac{(\log n)^{\cdots}}{n^{1/(16+10d)}}\right)$$

  - **distribution free**, far from best possible rate $n^{-1/2d}$ (Beck 85)

Uniform entropy

Class of spheres, any P

- **Exemple.** Spheres $\mathcal{F} = \{\mathbb{I}_C : C \in \mathcal{S}_d\}$ in $\mathcal{X} = [0,1]^d$
  - VC-index $d+2$
  - by Theorem U1, for **any** P

$$\|\alpha_n - \mathbb{G}_n\|_{\mathcal{F}} = O_{a.s.}\left(\frac{(\log n)^{\cdots}}{n^{1/(12+10d)}}\right)$$

  - by Theorem U2, for **any** P

$$\|\alpha_n - \mathbb{G}_n^*\|_{\mathcal{F}} = O_{a.s.}\left(\frac{(\log n)^{\cdots}}{n^{1/(16+10d)}}\right)$$

  - **distribution free**, far from best possible rate $n^{-1/2d}$ (Beck 85)
  - far from $(\log n)^{3/2} n^{-1/2d}$ for **uniform** P (Massart 89)

Uniform Entropy

Weighted sup-norm

- **Exemple.** Consider a weight ω on $\mathcal{X} = [0,1]^d$

- **Exemple.** Consider a weight ω on $\mathcal{X} = [0,1]^d$
  - $\inf_{\mathcal{X}} \omega = 0, \quad \int_{\mathcal{X}} \omega^{-2} \, dP < \infty$

- **Exemple.** Consider a weight ω on $\mathcal{X} = [0,1]^d$
  - $\inf_{\mathcal{X}} \omega = 0, \quad \int_{\mathcal{X}} \omega^{-2}\, dP < \infty$
  - let $\mathcal{X}_n \uparrow \mathcal{X} : 1 \gg \inf_{\mathcal{X}_n} \omega \gg \sqrt{(\log n)/n}$

- **Exemple.** Consider a weight ω on $\mathcal{X} = [0,1]^d$
  - $\inf_{\mathcal{X}} \omega = 0, \quad \int_{\mathcal{X}} \omega^{-2}\, dP < \infty$
  - let $\mathcal{X}_n \uparrow \mathcal{X}$: $1 \gg \inf_{\mathcal{X}_n} \omega \gg \sqrt{(\log n)/n}$
  - $\mathcal{F}_{\omega,n} = \left\{ \mathbb{I}_{[0,x]}/\omega(x) : x \in \mathcal{X}_n \right\}$ (VC index $d+2$)

Uniform Entropy

Weighted sup-norm

- **Exemple.** Consider a weight ω on $\mathcal{X} = [0,1]^d$
  - $\inf_{\mathcal{X}} \omega = 0, \quad \int_{\mathcal{X}} \omega^{-2}\, dP < \infty$
  - let $\mathcal{X}_n \uparrow \mathcal{X}$: $1 \gg \inf_{\mathcal{X}_n} \omega \gg \sqrt{(\log n)/n}$
  - $\mathcal{F}_{\omega,n} = \left\{ \mathbb{I}_{[0,x]}/\omega(x) : x \in \mathcal{X}_n \right\}$ (VC index $d+2$)
  - write $\alpha_n'(t) = \alpha_n(\mathbb{I}_{[0,t]})$, by Theorem U1, for **any** P

$$\left\| \frac{\alpha_n' - \mathbb{G}_n'}{\omega} \right\|_{\mathcal{X}_n} = \|\alpha_n - \mathbb{G}_n\|_{\mathcal{F}_{\omega,n}} = O_{a.s.}\left(\left(\frac{(\log n)^{7+5d}}{n \inf_{\mathcal{X}_n} \omega^2} \right)^{1/(12+10d)} \right)$$

Uniform Entropy

Weighted sup-norm

- **Exemple.** Consider a weight ω on $\mathcal{X} = [0,1]^d$
  - $\inf_{\mathcal{X}} \omega = 0, \quad \int_{\mathcal{X}} \omega^{-2}\, dP < \infty$
  - let $\mathcal{X}_n \uparrow \mathcal{X} : 1 \gg \inf_{\mathcal{X}_n} \omega \gg \sqrt{(\log n)/n}$
  - $\mathcal{F}_{\omega,n} = \left\{ \mathbb{I}_{[0,x]} / \omega(x) : x \in \mathcal{X}_n \right\}$ (VC index $d+2$)
  - write $\alpha_n'(t) = \alpha_n(\mathbb{I}_{[0,t]})$, by Theorem U1, for **any** P

$$\left\| \frac{\alpha_n' - \mathbb{G}_n'}{\omega} \right\|_{\mathcal{X}_n} = \|\alpha_n - \mathbb{G}_n\|_{\mathcal{F}_{\omega,n}} = O_{a.s.}\left(\left(\frac{(\log n)^{7+5d}}{n \inf_{\mathcal{X}_n} \omega^2} \right)^{1/(12+10d)} \right)$$

  - **general** and **tractable** statement relating P to ω

Uniform Entropy

Weighted sup-norm

- **Exemple.** Consider a weight ω on $\mathcal{X} = [0,1]^d$
  - $\inf_{\mathcal{X}} \omega = 0, \quad \int_{\mathcal{X}} \omega^{-2}\, dP < \infty$
  - let $\mathcal{X}_n \uparrow \mathcal{X} : 1 \gg \inf_{\mathcal{X}_n} \omega \gg \sqrt{(\log n)/n}$
  - $\mathcal{F}_{\omega,n} = \left\{ \mathbb{I}_{[0,x]}/\omega(x) : x \in \mathcal{X}_n \right\}$ (VC index $d+2$)
  - write $\alpha_n'(t) = \alpha_n(\mathbb{I}_{[0,t]})$, by Theorem U1, for **any** P

$$\left\| \frac{\alpha_n' - \mathbb{G}_n'}{\omega} \right\|_{\mathcal{X}_n} = \|\alpha_n - \mathbb{G}_n\|_{\mathcal{F}_{\omega,n}} = O_{a.s.}\left(\left(\frac{(\log n)^{7+5d}}{n \inf_{\mathcal{X}_n} \omega^2} \right)^{1/(12+10d)} \right)$$

  - **general** and **tractable** statement relating P to ω
  - By Theorem U2, power becomes $1/(16+10d)$

- **Exemple (continued).** In contrast, if $d = 1$ and P has continuous d.f. F

- **Exemple (continued).** In contrast, if $d = 1$ and P has continuous d.f. F
  - best known **weighted approximation** (Csörgő-Horváth 93)

$$\left\| \frac{\alpha_n' - \mathbb{G}_n'}{\omega_\alpha} \right\|_{\mathcal{X}_n} = O_{\mathbb{P}}\left(\frac{1}{n^{1/2-\alpha}} \right)$$

- **Exemple (continued).** In contrast, if $d = 1$ and P has continuous d.f. F
  - best known **weighted approximation** (Csörgő-Horváth 93)

$$\left\| \frac{\alpha_n' - \mathbb{G}_n'}{\omega_\alpha} \right\|_{\mathcal{X}_n} = O_{\mathbb{P}}\left(\frac{1}{n^{1/2-\alpha}} \right)$$

  - restricted to iid case and **intrinsic** "weight and set"

$$\omega_\alpha = F^\alpha (1-F)^\alpha, \quad \mathcal{X}_n = \left[F^{-1}\left(\frac{c}{n}\right), F^{-1}\left(1 - \frac{c}{n}\right) \right], \quad c > 0$$

- **Exemple (continued).** In contrast, if $d = 1$ and P has continuous d.f. F
  - best known **weighted approximation** (Csörgő-Horváth 93)

$$\left\| \frac{\alpha_n' - \mathbb{G}_n'}{\omega_\alpha} \right\|_{\mathcal{X}_n} = O_{\mathbb{P}}\left(\frac{1}{n^{1/2-\alpha}} \right)$$

  - restricted to iid case and **intrinsic** "weight and set"

$$\omega_\alpha = F^\alpha (1-F)^\alpha, \quad \mathcal{X}_n = \left[F^{-1}(\frac{c}{n}), F^{-1}(1-\frac{c}{n}) \right], \quad c > 0$$

  - for this weight, but also for **more** P, Theorem U1 yields

$$\left\| \frac{\alpha_n' - \mathbb{G}_n'}{\omega_\alpha} \right\|_{\mathcal{X}_n} = O_{\mathbb{P}}\left(\frac{(\log n)^{6/11}}{n^{(1/2-\alpha)/11}} \right)$$

Uniform Entropy

Squetch of proof

- Algebra of proof for uniform entropy

Uniform Entropy

Squetch of proof

- Algebra of proof for uniform entropy
  - large blocks, size p, separated by small blocks, size q

- Algebra of proof for uniform entropy
  - large blocks, size p, separated by small blocks, size q
  - apply Berbee lemma, making blocs independent

Uniform Entropy

Squetch of proof

- Algebra of proof for uniform entropy
  - large blocks, size p, separated by small blocks, size q
  - apply Berbee lemma, making blocs independent
  - build concentration inequality also using Berbee (Talagrand, moment inequalities)

Uniform Entropy

Squetch of proof

- Algebra of proof for uniform entropy
  - large blocks, size p, separated by small blocks, size q
  - apply Berbee lemma, making blocs independent
  - build concentration inequality also using Berbee (Talagrand, moment inequalities)
  - neglect small bloks, covariance inequalities

Uniform Entropy

Squetch of proof

- Algebra of proof for uniform entropy
  - large blocks, size p, separated by small blocks, size q
  - apply Berbee lemma, making blocs independent
  - build concentration inequality also using Berbee (Talagrand, moment inequalities)
  - neglect small bloks, covariance inequalities
  - sum the dependent r.v. inside large blocks (p-dim law P')

- Algebra of proof for uniform entropy
  - large blocks, size p, separated by small blocks, size q
  - apply Berbee lemma, making blocs independent
  - build concentration inequality also using Berbee (Talagrand, moment inequalities)
  - neglect small bloks, covariance inequalities
  - sum the dependent r.v. inside large blocks (p-dim law P')
  - $\mathbb{L}_2(P')$ entropy of $\mathcal{F}$ controled by initial entropy

Uniform Entropy

Squetch of proof

- Algebra of proof for uniform entropy
  - large blocks, size p, separated by small blocks, size q
  - apply Berbee lemma, making blocs independent
  - build concentration inequality also using Berbee (Talagrand, moment inequalities)
  - neglect small bloks, covariance inequalities
  - sum the dependent r.v. inside large blocks (p-dim law P')
  - $\mathbb{L}_2(P')$ entropy of $\mathcal{F}$ controled by initial entropy
  - take a covering of $\mathcal{F}$, form indep. vectors (1 per block)

Uniform Entropy

Squetch of proof

- Algebra of proof for uniform entropy
  - large blocks, size p, separated by small blocks, size q
  - apply Berbee lemma, making blocs independent
  - build concentration inequality also using Berbee (Talagrand, moment inequalities)
  - neglect small bloks, covariance inequalities
  - sum the dependent r.v. inside large blocks (p-dim law P')
  - $\mathbb{L}_2(P')$ entropy of $\mathcal{F}$ controled by initial entropy
  - take a covering of $\mathcal{F}$, form indep. vectors (1 per block)
  - apply Zaitzev coupling in finite dimension

- Algebra of proof for uniform entropy
  - large blocks, size p, separated by small blocks, size q
  - apply Berbee lemma, making blocs independent
  - build concentration inequality also using Berbee (Talagrand, moment inequalities)
  - neglect small bloks, covariance inequalities
  - sum the dependent r.v. inside large blocks (p-dim law P')
  - $\mathbb{L}_2(P')$ entropy of $\mathcal{F}$ controled by initial entropy
  - take a covering of $\mathcal{F}$, form indep. vectors (1 per block)
  - apply Zaitzev coupling in finite dimension
  - rebuild processes from coupled pointwise values

Bracketing Entropy

General rate, independent case

- Previous results are uniform in P : what about P **fixed** ?

Bracketing Entropy

General rate, independent case

- Previous results are uniform in P : what about P **fixed** ?
  - $\mathcal{F}$ is pointwise measurable, with envelope $F \in \mathbb{L}_2(P)$

Bracketing Entropy

General rate, independent case

- Previous results are uniform in P : what about P **fixed** ?
  - $\mathcal{F}$ is pointwise measurable, with envelope $F \in \mathbb{L}_2(P)$
  - increasing $\varphi(\varepsilon) \to \infty$, $\sqrt{\varphi}$ integrable, $\mathcal{F}$ is P-Donsker if

$$\log N_{[\,]}(\varepsilon, \mathcal{F}, d_P) \leq \varphi(\varepsilon), \quad 0 < \varepsilon < 1$$

$$J_{[\varphi]}(\varepsilon) = \int_0^{\varepsilon} \sqrt{\varphi}, \quad \Psi_{[\varphi]}(\varepsilon) = \varepsilon \exp\left(-\frac{5}{2}\varphi(\varepsilon)\right)$$

Bracketing Entropy

General rate, independent case

- Previous results are uniform in P : what about P **fixed** ?
  - $\mathcal{F}$ is pointwise measurable, with envelope $F \in \mathbb{L}_2(P)$
  - increasing $\varphi(\varepsilon) \to \infty$, $\sqrt{\varphi}$ integrable, $\mathcal{F}$ is P-Donsker if

$$\log N_{[\,]}(\varepsilon, \mathcal{F}, d_P) \leq \varphi(\varepsilon), \quad 0 < \varepsilon < 1$$

$$J_{[\varphi]}(\varepsilon) = \int_0^{\varepsilon} \sqrt{\varphi}, \quad \Psi_{[\varphi]}(\varepsilon) = \varepsilon \exp\left(-\frac{5}{2}\varphi(\varepsilon)\right)$$

  - again $\mathcal{F}_n = \{f \in \mathcal{F} : \|f\|_{\mathcal{X}} < M_n\}$

Bracketing Entropy

General rate, independent case

- Previous results are uniform in P : what about P **fixed** ?
  - $\mathcal{F}$ is pointwise measurable, with envelope $F \in \mathbb{L}_2(P)$
  - increasing $\varphi(\varepsilon) \to \infty$, $\sqrt{\varphi}$ integrable, $\mathcal{F}$ is P-Donsker if

$$\log N_{[\,]}(\varepsilon, \mathcal{F}, d_P) \leq \varphi(\varepsilon), \quad 0 < \varepsilon < 1$$

$$J_{[\varphi]}(\varepsilon) = \int_0^\varepsilon \sqrt{\varphi}, \quad \Psi_{[\varphi]}(\varepsilon) = \varepsilon \exp\left(-\frac{5}{2}\varphi(\varepsilon)\right)$$

  - again $\mathcal{F}_n = \{f \in \mathcal{F} : \|f\|_{\mathcal{X}} < M_n\}$
- **Theorem B1.** If $M_n\sqrt{(\log n)/n} \to 0$ we can construct $\{X_n\}$ and $\{\mathbb{G}_n\}$ on Ω such that

$$\|\alpha_n - \mathbb{G}_n\|_{\mathcal{F}_n} = O_{a.s.}\left(J_{[\varphi]} \circ \Psi_{[\varphi]}^{-1}\left(M_n\sqrt{\frac{\log n}{n}}\right)\right)$$

Bracketing entropy

Polynomialy log-entropy

- Assume

Bracketing entropy

Polynomialy log-entropy

- Assume
  - $\mathcal{F}$ is pointwise measurable, uniformly bounded

Bracketing entropy

Polynomialy log-entropy

- Assume
  - $\mathcal{F}$ is pointwise measurable, uniformly bounded
  - $N_{[\,]}$ brackets $\{f \in \mathcal{F} : l \leq f \leq u\}$, $d_P\left(l, u\right) < \varepsilon$,

$$\log N_{[\,]}\left(\varepsilon, \mathcal{F}, d_P\right) \leq \frac{1}{\varepsilon^{2r_0}}, \quad r_0 < 1.$$

Bracketing entropy

Polynomialy log-entropy

- Assume
  - $\mathcal{F}$ is pointwise measurable, uniformly bounded
  - $N_{[\,]}$ brackets $\{f \in \mathcal{F} : l \leq f \leq u\}$, $d_P(l, u) < \varepsilon$,

$$\log N_{[\,]}(\varepsilon, \mathcal{F}, d_P) \leq \frac{1}{\varepsilon^{2r_0}}, \quad r_0 < 1.$$

- **Corollary.** For any $\theta > 0$, all n we can construct $(\alpha_n, \mathbb{G})$

$$\mathbb{P}\left(\|\alpha_n - \mathbb{G}\|_{\mathcal{F}} > \frac{c_\theta}{(\log n)^r} \right) \leq \frac{1}{n^\theta}$$

where

$$r = \frac{1 - r_0}{2r_0}$$

Bracketing entropy, dependent case

Polynomialy log-entropy, exponentialy mixing

- Assume $(r_0 < 1)$

$$\log N_{[\,]}(\varepsilon, \mathcal{F}, d_P) \leq \frac{1}{\varepsilon^{2r_0}} \quad \text{and} \quad \phi_k < \exp(-\gamma k)$$

Bracketing entropy, dependent case

Polynomialy log-entropy, exponentialy mixing

- Assume $(r_0 < 1)$

$$\log N_{[\,]}(\varepsilon, \mathcal{F}, d_P) \leq \frac{1}{\varepsilon^{2r_0}} \quad \text{and} \quad \phi_k < \exp(-\gamma k)$$

- **Theorem B2.** We can construct $\{X_n\}$ and $\{\mathbb{G}_n^*\}$ on Ω,

$$\mathbb{P}\left(\|\alpha_n - \mathbb{G}_n^*\|_{\mathcal{F}} > c_\theta v_n\right) \leq \frac{1}{n^\theta}$$

with rate

$$v_n = \frac{1}{(\log n)^r}, \quad r = \frac{1 - r_0}{2r_0}.$$

and covariance bias

$$b_n = \frac{1}{(\log n)^{1/r_0}}.$$

Bracketing entropy, dependent case

Polynomialy log-entropy, exponentialy mixing

- Assume $(r_0 < 1)$

$$\log N_{[\,]}(\varepsilon, \mathcal{F}, d_P) \leq \frac{1}{\varepsilon^{2r_0}} \quad \text{and} \quad \phi_k < \exp(-\gamma k)$$

- **Theorem B2.** We can construct $\{X_n\}$ and $\{\mathbb{G}_n^*\}$ on Ω,

$$\mathbb{P}\left(\|\alpha_n - \mathbb{G}_n^*\|_{\mathcal{F}} > c_\theta v_n\right) \leq \frac{1}{n^\theta}$$

with rate

$$v_n = \frac{1}{(\log n)^r}, \quad r = \frac{1 - r_0}{2r_0}.$$

and covariance bias

$$b_n = \frac{1}{(\log n)^{1/r_0}}.$$

- To interpolate with independence, heavy **covariance cost**.

Bracketing entropy, dependent case

Polynomialy log-entropy, polymialy mixing

- Assume $(r_0 < 1 < \gamma)$

$$\log N_{[\,]}\left(\varepsilon, \mathcal{F}, d_P\right) \leq \frac{1}{\varepsilon^{2r_0}} \quad \text{and} \quad \phi_k < k^{-\gamma}$$

Bracketing entropy, dependent case

Polynomialy log-entropy, polymialy mixing

- Assume $(r_0 < 1 < \gamma)$

$$\log N_{[\,]}\left(\varepsilon, \mathcal{F}, d_P\right) \leq \frac{1}{\varepsilon^{2r_0}} \quad \text{and} \quad \phi_k < k^{-\gamma}$$

- **Theorem B3.** We can construct $\{X_n\}$ and $\{\mathbb{G}_n^*\}$ on Ω,

$$\mathbb{P}\left(\|\alpha_n - \mathbb{G}_n^*\|_{\mathcal{F}} > c v_n\right) \leq n^{1-\frac{1+\gamma}{2+r_0}}$$

  with rate

$$v_n = \frac{1}{(\log n)^r}, \quad r = \frac{1-r_0}{2r_0}$$

  and covariance bias

$$b_n = \frac{1}{(\log n)^{1/r_0}}.$$

Bracketing entropy, dependent case

Polynomialy log-entropy, polymialy mixing

- Assume $(r_0 < 1 < \gamma)$

$$\log N_{[\,]}\left(\varepsilon, \mathcal{F}, d_P\right) \leq \frac{1}{\varepsilon^{2r_0}} \quad \text{and} \quad \phi_k < k^{-\gamma}$$

- **Theorem B3.** We can construct $\{X_n\}$ and $\{\mathbb{G}_n^*\}$ on Ω,

$$\mathbb{P}\left(\|\alpha_n - \mathbb{G}_n^*\|_{\mathcal{F}} > cv_n\right) \leq n^{1-\frac{1+\gamma}{2+r_0}}$$

with rate

$$v_n = \frac{1}{(\log n)^r}, \quad r = \frac{1-r_0}{2r_0}$$

and covariance bias

$$b_n = \frac{1}{(\log n)^{1/r_0}}.$$

- To interpolate with independence, heavy **probability** and **covariance cost**. If $\gamma > 2r_0 + 3$ then a.s. invariance principle

Bracketing entropy

Smooth sets

- **Exemple.** Class $\mathcal{C}_d$ of sets in $\mathcal{X} = [0,1]^d$, P uniform

Bracketing entropy

Smooth sets

- **Exemple.** Class $\mathcal{C}_d$ of sets in $\mathcal{X} = [0,1]^d$, P uniform
  - bracket entropy b_0/ε^{r_0}, $r_0 < 1/2$, Minkowski-regular boundary

Bracketing entropy

Smooth sets

- **Exemple.** Class $\mathcal{C}_d$ of sets in $\mathcal{X} = [0,1]^d$, P uniform
  - bracket entropy b_0/ε^{r_0}, $r_0 < 1/2$, Minkowski-regular boundary
  - Massart (89) improved Révesz (76) into

$$\|\alpha_n - \mathbb{G}_n\|_{\mathcal{C}_d} = O_{a.s.}\left(\frac{1}{(\log n)^{(1-2r_0)/2d}}\right)$$

Bracketing entropy

Smooth sets

- **Exemple.** Class $\mathcal{C}_d$ of sets in $\mathcal{X} = [0,1]^d$, P uniform
  - bracket entropy b_0/ε^{r_0}, $r_0 < 1/2$, Minkowski-regular boundary
  - Massart (89) improved Révesz (76) into

$$\|\alpha_n - \mathbb{G}_n\|_{\mathcal{C}_d} = O_{a.s.}\left(\frac{1}{(\log n)^{(1-2r_0)/2d}}\right)$$

  - Theorem B1 improves, even if $d = 1$

$$\|\alpha_n - \mathbb{G}_n\|_{\mathcal{C}_d} = O_{a.s.}\left(\frac{1}{(\log n)^{(1-r_0)/2r_0}}\right)$$

Bracketing entropy

Smooth sets

- **Exemple.** Class $\mathcal{C}_d$ of sets in $\mathcal{X} = [0,1]^d$, P uniform
  - bracket entropy b_0/ε^{r_0}, $r_0 < 1/2$, Minkowski-regular boundary
  - Massart (89) improved Révesz (76) into

$$\|\alpha_n - \mathbb{G}_n\|_{\mathcal{C}_d} = O_{a.s.}\left(\frac{1}{(\log n)^{(1-2r_0)/2d}}\right)$$

  - Theorem B1 improves, even if $d = 1$

$$\|\alpha_n - \mathbb{G}_n\|_{\mathcal{C}_d} = O_{a.s.}\left(\frac{1}{(\log n)^{(1-r_0)/2r_0}}\right)$$

  - if $C \in \mathcal{C}_d$ are boundary differentiable through order $\alpha > d-1$ (then $r_0 = (d-1)/2\alpha$) **far better** as regularity $\alpha \to \infty$,

$$\|\alpha_n - \mathbb{G}_n\|_{\mathcal{C}_d} = O_{a.s.}\left(\frac{1}{(\log n)^{\alpha/(d-1)-1/2}}\right)$$

Bracketing entropy

Smooth sets

- **Exemple.** Class $\mathcal{C}_d$ of sets in $\mathcal{X} = [0,1]^d$, P uniform
  - bracket entropy b_0/ε^{r_0}, $r_0 < 1/2$, Minkowski-regular boundary
  - Massart (89) improved Révesz (76) into

$$\|\alpha_n - \mathbb{G}_n\|_{\mathcal{C}_d} = O_{a.s.}\left(\frac{1}{(\log n)^{(1-2r_0)/2d}}\right)$$

  - Theorem B1 improves, even if $d = 1$

$$\|\alpha_n - \mathbb{G}_n\|_{\mathcal{C}_d} = O_{a.s.}\left(\frac{1}{(\log n)^{(1-r_0)/2r_0}}\right)$$

  - if $C \in \mathcal{C}_d$ are boundary differentiable through order $\alpha > d-1$ (then $r_0 = (d-1)/2\alpha$) **far better** as regularity $\alpha \to \infty$,

$$\|\alpha_n - \mathbb{G}_n\|_{\mathcal{C}_d} = O_{a.s.}\left(\frac{1}{(\log n)^{\alpha/(d-1)-1/2}}\right)$$

  - in case of polynomial mixing, requires $\gamma > (d-1)/\alpha + 3$

- Versions under random entropy conditions (Settati 09).

Other conditions

Random entropy or geometric construction

- Versions under random entropy conditions (Settati 09).
- Versions for classes of functions with controled coordinates on a basis (Haar, wavelets, features) well suited for P or a Q close to P.

Other conditions

Random entropy or geometric construction

- Versions under random entropy conditions (Settati 09).
- Versions for classes of functions with controled coordinates on a basis (Haar, wavelets, features) well suited for P or a Q close to P.
- Versions for classes of functions well constructed from an alphabet of functions. Implies a d-dimensional **dimension free** rate $n^{-1/10}$ for KMT (cadrants, uniform law). Becomes a function of the mixing rate in dependent case.

Other conditions

Random entropy or geometric construction

- Versions under random entropy conditions (Settati 09).
- Versions for classes of functions with controled coordinates on a basis (Haar, wavelets, features) well suited for P or a Q close to P.
- Versions for classes of functions well constructed from an alphabet of functions. Implies a d-dimensional **dimension free** rate $n^{-1/10}$ for KMT (cadrants, uniform law). Becomes a function of the mixing rate in dependent case.
- Each time : the smaller is $\mathcal{F}$, the stronger the mixing should be to minimize the loss of rate.

Basic tool

Finite dimensional coupling

- We use the finite dimensional coupling of Zaitzev

- We use the finite dimensional coupling of Zaitzev
  - $Y_1, \ldots, Y_n$ independent, centered, in $\mathbb{R}^d$ and

$$|Y_i|_d < M$$

- We use the finite dimensional coupling of Zaitzev
  - $Y_1, \ldots, Y_n$ independent, centered, in $\mathbb{R}^d$ and

$$\left|Y_i\right|_d < M$$

  - for all λ we can define $Z_1, \ldots, Z_n$ independent centered Gaussian, each Z_i having same covariance matrix as Y_i, such that

$$\mathbb{P}\left(\left|\sum_{i=1}^{n}(Y_i - Z_i)\right|_d > \lambda\right) \leq c_0 d^2 \exp\left(-\frac{c_2 \lambda}{d^2 M}\right)$$

Random entropy

Typical rates, independent case

- Random entropy condition is close of necessity to Donsker property *(it is for sets)*

Random entropy

Typical rates, independent case

- Random entropy condition is close of necessity to Donsker property *(it is for sets)*
- Settati 09 weakened $\mathbb{L}_2$ bracketing entropy condition by using

$$\mathbb{P}\left(\log N\left(\frac{\sigma\varphi_n}{\sqrt{n}}, \mathcal{F}, \mathbb{L}_1(P_n)\right) > \varphi_n\right) \leq \frac{1}{n^2}$$

then the rate v_n of approximation is

$$\begin{array}{ll} (\log n)^{1-1/2\delta} & \varphi_n = n^\delta,\ \delta < 1/2 \\ \exp(-(c\log n)^{1/\delta}) & \varphi_n = (\log n)^\delta,\ \delta > 1 \\ \text{polynomial } n^{-1/(5v+2)} & \varphi_n = v\log n \\ \text{close to } n^{-1/2} & \varphi_n = (\log n)^\delta,\ \delta < 1 \end{array}$$

Geometrical conditions

Classes of sets

- Let $\mathcal{C} \subset \mathcal{A}$

Geometrical conditions

Classes of sets

- Let $\mathcal{C} \subset \mathcal{A}$
  - empirical process

$$\alpha_n(C) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \{\mathbb{I}_C(X_i) - P(C)\}, \quad C \in \mathcal{C}$$

Geometrical conditions

Classes of sets

- Let $\mathcal{C} \subset \mathcal{A}$
  - empirical process

$$\alpha_n(C) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \{\mathbb{I}_C(X_i) - P(C)\}, \quad C \in \mathcal{C}$$

  - d_P-continuous P-Brownian bridge $\mathbb{G}$ indexed by $\mathcal{C}$

$$cov(\mathbb{G}(C), \mathbb{G}(D)) = P(C \cap D) - P(C)P(D)$$

Geometrical conditions

Classes of sets

- Let $\mathcal{C} \subset \mathcal{A}$
  - empirical process

$$\alpha_n(C) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \{\mathbb{I}_C(X_i) - P(C)\}, \quad C \in \mathcal{C}$$

  - d_P-continuous P-Brownian bridge $\mathbb{G}$ indexed by $\mathcal{C}$

$$cov(\mathbb{G}(C), \mathbb{G}(D)) = P(C \cap D) - P(C)P(D)$$

- Given P we construct a P-Donsker $\mathcal{C}$ while controling the rate through a notion of **complexity** of order m

Geometrical conditions

Classes of sets

- Let $\mathcal{C} \subset \mathcal{A}$
  - empirical process

$$\alpha_n(C) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \{\mathbb{I}_C(X_i) - P(C)\}, \quad C \in \mathcal{C}$$

  - d_P-continuous P-Brownian bridge $\mathbb{G}$ indexed by $\mathcal{C}$

$$cov(\mathbb{G}(C), \mathbb{G}(D)) = P(C \cap D) - P(C)P(D)$$

- Given P we construct a P-Donsker $\mathcal{C}$ while controling the rate through a notion of **complexity** of order m
- Open : work out the **appropriate mixing coefficients** (on $\mathcal{C}$ only?).

Geometrical conditions

Strong chaining for sets

- Take a countable $\mathcal{C}_\infty \subset \mathcal{C}$, choose $\mathcal{C}_n^*$ **finite** in $\mathcal{C}_\infty$

Geometrical conditions

Strong chaining for sets

- Take a countable $\mathcal{C}_\infty \subset \mathcal{C}$, choose $\mathcal{C}_n^*$ **finite** in $\mathcal{C}_\infty$
  - **continuity moduli operator** $q_n^* : \mathcal{C} \to \mathcal{C}_n^*$, $q_*(C) \subset C$

$$\mathbb{P}\left(\sqrt{n}Y \geq \varrho(n)\right) \leq \frac{1}{n^2}$$

for both $Y = \|\alpha_n - \alpha_n \circ q_n^*\|_{\mathcal{C}}$ and $Y = \|\mathbb{G} - \mathbb{G} \circ q_n^*\|_{\mathcal{C}}$

Geometrical conditions

Strong chaining for sets

- Take a countable $\mathcal{C}_\infty \subset \mathcal{C}$, choose $\mathcal{C}_n^*$ **finite** in $\mathcal{C}_\infty$
  - **continuity moduli operator** $q_n^* : \mathcal{C} \to \mathcal{C}_n^*$, $q_*(C) \subset C$

$$\mathbb{P}\left(\sqrt{n}Y \geq \varrho(n)\right) \leq \frac{1}{n^2}$$

for both $Y = \|\alpha_n - \alpha_n \circ q_n^*\|_\mathcal{C}$ and $Y = \|\mathbb{G} - \mathbb{G} \circ q_n^*\|_\mathcal{C}$

  - **chaining operator** $q_m^{**} : \mathcal{C}_n^* \to \mathcal{C}_m^{**}$ where

$$\mathcal{C}_m^{**} = \{C_1 \cup \ldots \cup \ C_m : C_j \in \mathcal{P}_j, C_j \cap C_l = \emptyset \text{ if } j \neq l\}$$

and $\mathcal{P}_j = \{C_{j,k} : k \leq k(j)\}$ are m **finite partitions**

$$\sup_{C \in \mathcal{C}_n^*} P(C \backslash q_m^{**}(C)) \leq h_n(m) \downarrow 0 \text{ as } m \to \infty$$

Geometrical conditions

Strong chaining for sets

- **Geometrical intrinsic dimension** *m* hidden in Ψ

$$c_n = card(\mathcal{C}_n^*) \qquad d_m = \sum_{j \leq m} k(j)$$

$$S_m = \sum_{j \leq m} k(j) \left\| P^2 \right\|_{\mathcal{P}_j} \qquad \Psi(m) = \frac{h_n(m)}{m^2 d_m^4 (\log d_m)^2}$$

Geometrical conditions

Strong chaining for sets

- **Geometrical intrinsic dimension** m hidden in Ψ

$$c_n = card(\mathcal{C}_n^*) \qquad d_m = \sum_{j \leq m} k(j)$$

$$S_m = \sum_{j \leq m} k(j) \left\| P^2 \right\|_{\mathcal{P}_j} \qquad \Psi(m) = \frac{h_n(m)}{m^2 d_m^4 (\log d_m)^2}$$

- **Theorem G1.** If $S_m \leq a_\infty m$ we can construct $(\alpha_n, \mathbb{G}_n)$

$$\|\alpha_n - \mathbb{G}_n\|_{\mathcal{C}} =$$

$$O_{a.s.}\left(\frac{\varrho(n)}{\sqrt{n}} \vee \frac{\log c_n}{n} \vee \sqrt{h_n \circ \Psi^{-1}\left(\frac{(\log n)^2}{n \log c_n} \right) \log c_n} \right)$$

Geometrical Conditions

KMT in dimension d>2

- **Example.** Cubes $\mathcal{F} = \left\{ \mathbb{I}_{[s,t]} : s, t \in \mathcal{X} = [0,1]^d \right\}$

Geometrical Conditions

KMT in dimension d>2

- **Example.** Cubes $\mathcal{F} = \left\{ \mathbb{I}_{[s,t]} : s, t \in \mathcal{X} = [0,1]^d \right\}$
  - if P has a bounded density

$$\|\alpha_n - \mathbb{G}_n\|_{\mathcal{F}} = O_{a.s.}\left(\frac{(\log n)^{d-1/2}}{n^{1/10}} \right)$$

- **Example.** Cubes $\mathcal{F} = \left\{ \mathbb{I}_{[s,t]} : s, t \in \mathcal{X} = [0,1]^d \right\}$
  - if P has a bounded density

$$\|\alpha_n - \mathbb{G}_n\|_{\mathcal{F}} = O_{a.s.}\left(\frac{(\log n)^{d-1/2}}{n^{1/10}} \right)$$

  - almost **dimension free**, improves $n^{-1/12}$ (Rio 96)

Geometrical Conditions

KMT in dimension d>2

- **Example.** Cubes $\mathcal{F} = \left\{ \mathbb{I}_{[s,t]} : s,t \in \mathcal{X} = [0,1]^d \right\}$
  - if P has a bounded density

$$\|\alpha_n - \mathbb{G}_n\|_{\mathcal{F}} = O_{a.s.}\left(\frac{(\log n)^{d-1/2}}{n^{1/10}} \right)$$

  - almost **dimension free**, improves $n^{-1/12}$ (Rio 96)
  - partitions are dyadic, $m \approx \log n$, $d_m \approx 2^m$, S_m bounded, $h_n(m) = h(m) \approx 2^{-m}$, $\Psi(m) \approx 2^{-5m}$, $h \circ \Psi^{-1}(x) \approx x^{1/5}$, $\varrho(n) \approx 1/\sqrt{n}$, $c_n \approx n^d$

Uniform Entropy

Weighted sup-norm

- Conclusion

Uniform Entropy

Weighted sup-norm

- Conclusion
  - among the first strong invariance principles in the mixing setting

- Conclusion
  - among the first strong invariance principles in the mixing setting
  - **trade off** between **covariance bias** and **rate** of approximation

Uniform Entropy

Weighted sup-norm

- Conclusion
  - among the first strong invariance principles in the mixing setting
  - **trade off** between **covariance bias** and **rate** of approximation
  - the approximating Gaussian process becomes the limiting one after normalization but the price is the covariance bias...