



**HAL**  
open science

# Rates of convergence for the weighted bootstrap of empirical and raking-ratio processes

Mickael Albertus, Philippe Berthet

► **To cite this version:**

Mickael Albertus, Philippe Berthet. Rates of convergence for the weighted bootstrap of empirical and raking-ratio processes. 2019. hal-01979703

**HAL Id: hal-01979703**

**<https://hal.science/hal-01979703>**

Preprint submitted on 13 Jan 2019

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Rates of convergence for the weighted bootstrap of empirical and raking-ratio processes

M. Albertus and P. Berthet

January 13, 2019

## Abstract

We study the weighted bootstrap of the empirical process indexed by a class of functions, when the weights are allowed to be data dependent. In addition to the classical one, we also consider three weighted bootstrap new methods based on the raking-ratio process using an auxiliary information on  $N$  partitions. Assuming entropy conditions like VC dimension, we use nonasymptotic strong approximation arguments to characterize the joint limiting Gaussian processes of  $b_n$  bootstrap experiments and to evaluate the rate of weak uniform convergence as  $b_n \rightarrow +\infty$  with the initial sample size  $n \rightarrow +\infty$ . Berry-Esseen bounds for bootstrapped statistics follows. This justifies the weighted bootstrap methodology to estimate the distribution of raked statistics, in particular their lower variance and smaller confident bands.

*Keywords:* Weighted bootstrap, Uniform central limit theorems, Non-parametric statistics, empirical processes, raking ratio process.

*MSC Classification:* 62G09, 62G20, 60F17, 60F05.

## Contents

<b>1</b>	<b>Introduction</b>	<b>2</b>
1.1	The classical bootstrap . . . . .	2
1.2	The weighted bootstrap . . . . .	4
1.3	Weighted bootstraps . . . . .	6
<b>2</b>	<b>Main results</b>	<b>9</b>
2.1	The class $\mathcal{F}$ . . . . .	9
2.2	Strong approximations . . . . .	10
2.3	Estimation of variance and distribution function . . . . .	12
2.4	Rates of weak convergence . . . . .	15

<b>3</b>	<b>Raking-Ratio results</b>	<b>15</b>
3.1	Strong approximation of $\alpha_n^{*(N)}$	15
3.2	Strong approximation of $\alpha_n^{(N)*}$	17
<b>4</b>	<b>Proofs</b>	<b>18</b>
4.1	Decomposition of $\alpha_n^*$	18
4.2	Proof of Propositions 1, 4 and 5	20
4.3	Construction of limit Gaussian processes	23
4.4	Proof of Theorem 2.1 and 2.2	26
4.5	Proof of Theorem 3.1 and 3.2	30
4.6	Proof of Proposition 6	33
4.7	Proof of Corollaries 1 and 2	36
4.8	Proof of Corollaries (3) and (4)	39

## 1 Introduction

### 1.1 The classical bootstrap

**Presentation.** The bootstrap is a very popular method of statistical inference introduced by Efron [15, 16] that could be viewed as a generalization of the older jackknife method or leave  $k$ -out methods. Given  $n \geq 1$  independent random variables  $X_1, \dots, X_n$  with common law  $P$  on a measurable space  $(\mathcal{X}, \mathcal{A})$  let  $\mathbb{P}_n = \frac{1}{n} \sum_{i=1}^n \delta_{X_i}$  denote the associated empirical measure where  $\delta_{X_i}$  are the Dirac measures. Any statistic of interest  $S_n(X_1, \dots, X_n)$  being symmetric in its arguments can be written  $\varphi(\mathbb{P}_n)$ . Whenever  $S_n$  is not known to satisfy good estimation or test properties, it is unfortunately observed only once. The classical bootstrap aims to learn about unobserved properties of  $S_n$  by re-sampling at will  $X_1^*, \dots, X_n^*$  among  $X_1, \dots, X_n$  uniformly with replacement then estimating by Monte-Carlo methods the bias, variance or distribution of  $S_n^* = S_n(X_1^*, \dots, X_n^*) = \varphi(\mathbb{P}_n^*)$  centered at  $S_n(X_1, \dots, X_n)$ , with  $\mathbb{P}_n^* = \frac{1}{n} \sum_{i=1}^n \delta_{X_i^*}$ . The paradigm of Efron is that without any information on  $P$  the best way to mimic the unknown product measure  $(P)^n = P \times \dots \times P$  of the original sample is to use the product empirical measure  $(\mathbb{P}_n)^n = \mathbb{P}_n \times \dots \times \mathbb{P}_n$  and to center  $\mathbb{P}_n^*$  at  $\mathbb{P}_n$  instead of  $P$ .

**Motivation.** From the mathematical statistics viewpoint a crucial question that has not been investigated in general is to quantify the information one really gets about the distributions of  $S_n$  when bootstrapping  $b_n$  times, with  $b_n \rightarrow +\infty$  as  $n \rightarrow +\infty$ . Towards this aim we address two unusual problems. The first problem is to find a sufficient condition on how to choose  $b_n$  to control the joint weak convergence of bootstrapped experiments. The second problem is how to add auxiliary information on  $P$  while bootstrapping.

**Monte-Carlo bootstrap.** The mathematical justification of the bootstrap methodology is not obvious, even for a single explicit targeted statistic  $S_n$  since it strongly depends on  $\varphi$  itself. Strictly speaking, one should evaluate how close the random experiments  $\varphi(\mathbb{P}_n^*)$  and  $\varphi(\mathbb{P}_n)$  are as  $n \rightarrow +\infty$  and derive

the consequences on the subsequent estimation procedure from  $b_n$  conditionally independent bootstraps. The known answers are mostly asymptotic and don't involve  $n$  and  $b_n$  together. Usually  $b_n = 1$  when  $n \rightarrow +\infty$  and, for fixed  $n$ , it is implicit that  $b_n \rightarrow +\infty$  allows to numerically learn  $\varphi(\mathbb{P}_n^*)$ . However, the learned probability distribution is conditional to the initial sample, and the difference with  $\varphi(\mathbb{P}_n)$  could be misleading. Thus letting  $b_n \rightarrow +\infty$  is not so useful due to the overfitting phenomenon. In other words, a rigorous compromise with the bias involving jointly the initial sample and the bootstrap samples is missing in the theory. We look for  $b_n$  small enough to provide non-asymptotic joint results.

**Asymptotic justification.** The statistics  $S_n$  we consider are sensitive to deviations between empirical and true expectations over a class of functions  $\mathcal{F} \subset L_2(P)$ . They are determined by  $\alpha_n(f) = \sqrt{n}(\mathbb{P}_n(f) - P(f))$  where  $\mathbb{P}_n(f) = n^{-1} \sum_{i=1}^n f(X_i)$  and  $P(f) = \mathbb{E}(f(X))$ ,  $f \in \mathcal{F}$ . The collection  $\alpha_n(\mathcal{F}) = \{\alpha_n(f) : f \in \mathcal{F}\}$  is called the empirical process  $\alpha_n$  indexed by  $\mathcal{F}$ . Its bootstrapped version is  $\alpha_n^*(\mathcal{F}) = \{\alpha_n^*(f) : f \in \mathcal{F}\}$ ,  $\alpha_n^*(f) = \sqrt{n}(\mathbb{P}_n^*(f) - \mathbb{P}_n(f))$  where  $\mathbb{P}_n^*$  is the bootstrapped empirical measure introduced below. What is usually established is that conditionally to  $X_1, \dots, X_n$  the bootstrap process  $\alpha_n^*(\mathcal{F})$  has the same behavior as the empirical process  $\alpha_n(\mathcal{F})$  for  $n$  large. Basically, with probability one  $\alpha_n^*$  weakly converges to the weak limit of  $\alpha_n$  as  $n \rightarrow +\infty$ , most often the  $P$ -Brownian bridge  $\mathbb{G}(\mathcal{F})$  indexed by  $\mathcal{F}$  if  $\mathcal{F}$  is a Donsker class. The bootstrap method is therefore justified at the first order if  $S_n = \varphi(\mathbb{P}_n) \in \mathbb{R}^d$ ,  $S_n^* = \varphi(\mathbb{P}_n^*)$  and  $\varphi$  is Fréchet-differentiable at  $P$  since then the distribution of  $Y_n = \sqrt{n}(S_n - \varphi(P))$  can be estimated by the distribution of  $Y_n^* = \sqrt{n}(S_n^* - S_n)$  which is in smooth cases asymptotically the same random vector  $\varphi'(P) \cdot \mathbb{G}$  as for  $Y_n$  provided that the differential distortion  $\varphi'(\mathbb{P}_n) - \varphi'(P)$  vanishes. Whenever  $Y_n^*$  is simulated  $b_n$  times, the distortion generates a bias and  $b_n$  should be calibrated to avoid overfitting and learning the bias through  $\varphi'(\mathbb{P}_n)$ . In particular cases the weak distance between the distributions of  $Y_n$ ,  $Y_n^*$  and  $\varphi'(P) \cdot \mathbb{G}$  or between the distributions of  $S_n^*$  and  $S_n$  are known to vanish. No general estimate is available at fixed  $n$ , hence the balanced choice of  $b_n$  to guaranty well controlled Monte-Carlo estimates can not be discussed. This is the main motivation of this paper, however we work with the weighted bootstrap version of  $\alpha_n^*$  and  $S_n^*$  together with three new variants exploiting some auxiliary information.

**Weak convergence.** Giné and Zinn [17] proved that for any class of functions  $\mathcal{F}$  with envelope in  $L_2(P)$  the weak convergence of  $\alpha_n(\mathcal{F})$  to a Gaussian or not - process  $\mathbb{G}$  indexed by  $\mathcal{F}$  is necessary and sufficient for the Efron's bootstrap empirical process  $\alpha_n^*(\mathcal{F})$  to almost surely converge weakly to  $\mathbb{G}(\mathcal{F})$  also. This very nice statement is one of the most general results of the huge literature on the bootstrap methodology. For a single real valued and regular statistic  $S_n$  a common approach is through Edgeworth expansions, which exploits the cumulant expansion of the distribution function, see e.g. Hall [19] or Shao and Tu [29]. Other approaches rely on Berry-Esseen bounds, like in Singh [31] or Mallows distances, as in Bickel and Freedman [9]. Under the name of Bayesian bootstrap, Rubin [28] defined an analogue of Efron's bootstrap by resampling according to exchangeable weights that are independent of  $X_1, \dots, X_n$  rather

than uniformly according to  $\mathbb{P}_n$ . In the case  $\mathcal{X} = \mathbb{R}$ , Mason and Newton [26] further generalized the bootstrap by independently assigning self-normalized random weights to the original data. If the weights are drawn independently from a multinomial distribution this reduces to the Efron’s bootstrap whereas if the weights come independently from a Dirichlet distribution this reduces to the Bayesian bootstrap. They established the weak convergence of this weighted real empirical process to a Brownian bridge provided the positive exchangeable weights satisfy a weak convergence condition. Unlike Giné and Zinn [17], their result does not handle the case indexed by  $\mathcal{F}$ . Præstgaard and Wellner [27] fills this gap, by still assuming that the exchangeable weights are independent from the data, with again a common weak limit  $\mathbb{G}(\mathcal{F})$  for  $\alpha_n(\mathcal{F})$  and  $\alpha_n^*(\mathcal{F})$ .

**About rates.** The above weak convergence of Efron’s and weighted bootstrap processes are usually formulated in the sense of Hoffmann-Jørgensen to handle carefully measurability problems – see [4, 20]. The main results are assembled in chapter 3.6 of Van der Vaart and Wellner [35] – see also Kosorok [23]. The weak convergence is usually quantified in terms of the bounded Lipschitz norm between the processes  $\alpha_n^*(\mathcal{F})$  and  $\mathbb{G}(\mathcal{F})$ , with no explicit rate. Obviously the rates could be arbitrarily slow for large  $\mathcal{F}$  or inadequate couples  $(P, \mathcal{F})$ . A nice feature of our approach is that it provides general and explicit rates at the empirical process level for typical  $P$ -Donsker or universal Donsker classes  $\mathcal{F}$ , with quantified statistical consequences. A few authors considered the distance between the probability measures themselves, like  $\mathbb{P}_n^*$ ,  $\mathbb{P}_n$  or  $P$ . For instance Shao [30] proved that the bounded Lipschitz distance between the uniform empirical measure  $\mathbb{P}_n$  on the  $d$ -dimensional unit cube and the Efron’s bootstrap empirical measure  $\mathbb{P}_n^*$  is bounded by  $O(n^{-1/d})$  if  $d > 2$  and  $O(n^{-1/2}(\log n)^{(d-1)/2})$  if  $d = 1, 2$ . This improved Beran, Le Cam and Millar’s result [7] which only implies the convergence to zero, however in the more general indexed by sets setting. Other metrics have been studied in the indexed by  $\mathcal{F}$  setting. For instance, Barbe and Bertail [6] showed that various supremum type distances between the weighted bootstrap measure  $\mathbb{P}_n^*$  and  $P$  on  $\mathcal{F}$  are  $O(n^{-1/2}(\log n)^{1/2})$  in probability where the extra  $\log n$  term can be removed by following [27]. Likewise, we derive almost sure first order rates  $n^{-1/2}$  together with second order rates, so that the weak distance between the distribution of  $S_n^*$  and  $S_n$  may be evaluated.

## 1.2 The weighted bootstrap

Our primary goal is to sharpen the above results for the self-normalized weighted bootstrap empirical measure in terms of distance in distribution. As a consequence, general answers to the two problems of the initial motivation follow. Extensions to Efron’s bootstrap will be studied elsewhere.

**A strong approximation approach.** By strong approximation we mean a coupling of the process  $\alpha_n^*(\mathcal{F})$  and its Gaussian limiting process. This avoids cumbersome weak convergence measurability considerations while providing a sufficient control over the weak convergence metrics and being easy to use. Such a Brownian coupling has been established in the very specific setting where

KMT [22] can be applied, that is when  $P$  is the uniform law on  $(0, 1)$ . Alvarez-Andrade and Bouzebda [3] derived the usual almost sure rate  $O(n^{-1/2} \log n)$  of strong approximation by a sequence of Brownian bridges for the weighted bootstrap process  $\alpha_n^*(u) = \sqrt{n}(\mathbb{F}_n^*(u) - \mathbb{F}_n(u))$ , where

$$\mathbb{F}_n(u) = \frac{1}{n} \sum_{i=1}^n \mathbb{1}_{\{X_i \leq u\}}, \quad \mathbb{F}_n^*(u) = \frac{1}{n} \sum_{i=1}^n W_{i,n} \mathbb{1}_{\{X_i \leq u\}}, \quad u \in (0, 1),$$

and, for independent and identically distributed (*i.i.d.*) weights  $Z_1, \dots, Z_n$  also independent from  $X_1, \dots, X_n$ ,

$$W_{i,n} = \frac{Z_i}{T_n}, \quad T_n = \sum_{i=1}^n Z_i, \quad (1.1)$$

provided  $\mathbb{E}(Z) = \mathbb{V}(Z) = 1$  and  $Z_1$  has a Laplace transform in a neighborhood of 0. Following the forthcoming arguments this induces a distance  $O(n^{-1/2} \log n)$  between  $\mathbb{P}_n$  and  $\mathbb{P}_n^*$  in various weak convergence metrics.

**Main result in the classical setting.** In this paper we revisit the above mentioned results for the self-normalized weighted bootstrap empirical process

$$\alpha_n^*(f) = \sqrt{n}(\mathbb{P}_n^*(f) - \mathbb{P}_n(f)), \quad \mathbb{P}_n^*(f) = \frac{1}{n} \sum_{i=1}^n W_{i,n} f(X_i), \quad f \in \mathcal{F}, \quad (1.2)$$

with the weights  $W_{i,n}$  from (1.1). Clearly for  $\mathcal{F} = \{\mathbb{1}_{\{\cdot \leq t\}} : t \in \mathbb{R}\}$  and the uniform distribution on  $(0, 1)$  we recover the bootstrapped empirical process  $\alpha_n^*(u)$  of [3] defined above, however we allow any distribution  $P$  on  $\mathbb{R}$  for this class  $\mathcal{F}$ , not only absolutely continuous ones. Furthermore we relax the usual assumption that the resampling weights  $Z_i/T_n$  are independent from the original sample by allowing  $(X_i, Z_i)$  to be *i.i.d.* with some distribution  $P^{(X,Z)}$  while still controlling the marginal laws  $P$  of  $X$  and  $P^Z$  of  $Z$ . Our main result is a nonasymptotic joint strong approximation of  $b_n$  bootstrap iterations of  $\alpha_n^*(\mathcal{F})$  by independent versions of the same Gaussian process  $\mathbb{G}(\mathcal{F})$  as the weak limit of  $\alpha_n(\mathcal{F})$ , jointly to the approximation of  $\alpha_n(\mathcal{F})$  itself. This allows to turn these conditionally independent, and orthogonal, empirical processes into independent Gaussian processes. It ensues an uniform central limit theorem for the bootstrap procedure (1.2) with rates in various weak convergence metrics, including uniform Berry-Esseen type results and distances between distributions of  $S_n^*$  and  $S_n$ . The rigorous control of the bootstrap Monte-Carlo procedure itself is our most innovative contribution.

**Bootstrapping under auxiliary information.** Our secondary goal is to extend the bootstrap procedure to a less classical setting where an auxiliary information is known or learned about  $P$ . The motivation comes from the hasty development of distributed data. In this context it is realistic to consider a global statistical model where several sources learn deeply about one partial aspect of  $P$  and only communicate their conclusions rather than their too large

or confidential samples. In the next section we define the raking-ratio empirical process that combines these informations through  $N$  iterations of the procedure in order to improve the inference from  $X_1, \dots, X_n$  as studied in details in [1]. The underlying intuition connecting the bootstrap and the raking-ratio is twofold. On the one hand a better knowledge of  $P$  may help the bootstrap by either improving the initial  $\mathbb{P}_n$  and/or the redrawn  $\mathbb{P}_n^*$ . On the other hand bootstrapping the raked empirical measure denoted below  $\mathbb{P}_n^{(N)}$  may give access to the distribution of a raked statistic observed only once, in particular to evaluate its lower variance, small bias and reduced risk.

**Organization.** The paper is organized as follows. The four bootstrap procedures to be studied are presented in section 1.3 together with our paradigm of auxiliary information from partitions. We present the main results in parts 2 and 3. More precisely, section 2.2 provides the strong approximation of the weighted bootstrap empirical process iterated  $b_n$  times, then statistical consequences are derived in sections 2.3 and 2.4. Results of section 3.1 show that we can apply the Raking-Ratio method after bootstrapping a sample in order to simulate the asymptotic law of the empirical Raking-Ratio process. The results stated in sections 3.2 and ?? show how the performances of the basic bootstrap are improved by using a true information on  $P$ . Finally, the proof of all results are postponed until section 4, focusing mainly on the case without information then avoiding straightforward details.

### 1.3 Weighted bootstraps

**The weighted bootstrap empirical process.** Let  $(X_1, Z_1), \dots, (X_n, Z_n)$  be independent random variables with unknown law  $P^{(X,Z)}$  on  $(\mathcal{X}, \mathcal{A}) \times (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ . The conditional distributions  $P^{(Z|X=x)}$  are assumed to exist and satisfy

$$\mathbb{E}(Z|X = x) = \text{Var}(Z|X = x) = 1, \quad x \in \mathcal{X}. \quad (1.3)$$

For sake of simplicity we shall assume

$$\mathbb{P}(0 \leq Z \leq M_Z) = 1, \quad M_Z < +\infty. \quad (1.4)$$

This is not restrictive when  $n$  is fixed since for  $M_Z = F_Z^{-1}(1 - 1/n^{\theta+1}b_n)$  the random variable  $\mathbb{1}_{\{|Z| \leq M_Z\}}Z + \mathbb{1}_{\{|Z| > M_Z\}}M_Z$  behaves like  $Z$  over all bootstrapped samples with probability of order  $1 - 1/n^\theta$ . However, as  $n \rightarrow +\infty$ , truncation arguments would be cumbersome.

The first weighted bootstrap process  $\alpha_n^*$  to be considered is defined at (1.1) and (1.2). We study the joint convergence of  $(\alpha_n(\mathcal{F}), \alpha_n^*(\mathcal{F}))$  with rates. Evaluating the weak distance between  $\alpha_n^*$  and its limit is crucial since each new bootstrap sample  $X_1^*, \dots, X_n^*$  is affected by it. Likewise, in order to control the global distortion in play by not using  $P^n$  when bootstrapping a collection of estimators  $S_n$  with law  $\varphi(\mathbb{P}_n)$  we shall approximate jointly the  $b_n$  bootstrapped empirical processes. The coupling error being quantified in the very strong sup-norm over  $\mathcal{F}$ ,  $b_n$  has to be sufficiently small to guaranty that the confident bands for infinitely many estimated parameters are uniformly not over-biased.

Beyond the classical one, let us introduce three other weighted bootstraps - only the first two are studied, the third one satisfies similar results however with heavier notation.

**The raking-ratio empirical process.** What we call the raking-ratio algorithm was introduced by Deming and Stephan [14] and rectified by Stephan [34] then justified by Lewis [25], Brown [12], Sinkhorn [32, 33] and finally Ireland and Kullback [21]. This procedure consists in changing iteratively the weights of each  $X_i$  to match known probabilities of discrete marginals. Special cases or closely related methods are stratification, calibration, fitting after sampling, iterative proportions or matrix scaling. A rather general study at the empirical measure level was conducted in [1] from the viewpoint of auxiliary information of partitions. Let us briefly introduce the suitable notation and a few results in our setting where entries of the algorithm are random.

For all  $N \in \mathbb{N}^*$  assume that  $m_N \geq 1$  and  $A_1^{(N)}, \dots, A_{m_N}^{(N)} \subset \mathcal{A}$  is a partition of  $\mathcal{X}$  such that the discrete marginal  $P(\mathcal{A}^{(N)}) = (P(A_1^{(N)}), \dots, P(A_{m_N}^{(N)}))$  is known and  $p_N = \min_{1 \leq j \leq m_N} P(A_j^{(N)}) > 0$ . Let  $\mathbb{P}_n^{(N)}$  and  $\alpha_n^{(N)}$  denote respectively the empirical measure and process associated with the raking-ratio method, defined to be  $\mathbb{P}_n^{(0)} = \mathbb{P}_n$ ,  $\alpha_n^{(0)} = \alpha_n$  and, for  $N \geq 1$  and on the event  $B_n^{1,N} = \{\min_{1 \leq j \leq m_N} \mathbb{P}_n^{(N-1)}(A_j^{(N)}) > 0\}$ ,

$$\mathbb{P}_n^{(N)}(f) = \sum_{j=1}^{m_N} \frac{P(A_j^{(N)})}{\mathbb{P}_n^{(N-1)}(A_j^{(N)})} \mathbb{P}_n^{(N-1)}(f 1_{A_j^{(N)}}), \quad (1.5)$$

$$\alpha_n^{(N)}(f) = \sqrt{n}(\mathbb{P}_n^{(N)}(f) - P(f)), \quad f \in \mathcal{F}. \quad (1.6)$$

In particular,  $\mathbb{P}_n^{(N)}(A_j^{(N)}) = P(A_j^{(N)})$  and  $\alpha_n^{(N)}(A_j^{(N)}) = 0$  for all  $N \in \mathbb{N}$  and  $j = 1, \dots, m_N$ . Notice that  $\mathbb{P}_n^{(N)}$  is the Kullback projection of  $\mathbb{P}_n^{(N-1)}$  satisfying the  $N$ -th step constraint  $P(\mathcal{A}^{(N)})$  according to Proposition 1 of [1]. We proved that under classical entropy conditions on  $\mathcal{F}$ ,  $\alpha_n^{(N)}(\mathcal{F})$  converges weakly as  $n \rightarrow +\infty$  to a centered Gaussian process  $\mathbb{G}^{(N)}$  defined similarly to  $\alpha_n^{(N)}$  - see Proposition 4 of [1]. More precisely, write  $\mathbb{E}[f|A] = P(f 1_A)/P(A)$  and set  $\mathbb{G}^{(0)} = \mathbb{G}$  to be the  $P$ -Brownian bridge indexed by  $\mathcal{F}$  then define the  $P$ -raked Brownian bridge to be, for  $N \in \mathbb{N}^*$ ,

$$\mathbb{G}^{(N)}(f) = \mathbb{G}^{(N-1)}(f) - \sum_{j=1}^{m_N} \mathbb{E}[f|A_j^{(N)}] \mathbb{G}^{(N-1)}(A_j^{(N)}), \quad f \in \mathcal{F}. \quad (1.7)$$

We established that the covariance of the limiting process  $\mathbb{G}^{(N)}$  is

$$\text{Cov}(\mathbb{G}^{(N)}(f), \mathbb{G}^{(N)}(g)) = \text{Cov}(\mathbb{G}^{(0)}(f), \mathbb{G}^{(0)}(g)) - R_N(P, f, g) \quad (1.8)$$

where  $R_N(P, f, g)$  has the following closed form expression given at Proposition



7 in [1]:

$$R_N(P, f, g) = \sum_{k=1}^N \Phi_k^{(N)}(P, f) \cdot \text{Var}(\mathbb{G}[\mathcal{A}^{(k)}]) \cdot \Phi_k^{(N)}(P, g), \quad (1.9)$$

$$\Phi_k^{(N)}(P, f) = \sum_{l=1}^N \mathbf{P}(P, l) \cdot P[f|\mathcal{A}^{(l)}],$$

where  $P[f|\mathcal{A}^{(l)}] = (P(f|A_1^{(l)}), \dots, P(f|A_{m_l}^{(l)}))$ ,  $\mathbf{P}(P, l) \in \mathbb{R}^{m_l}$ . The asymptotic uniform variance reduction is induced by the fact that  $R_N(P, f, f) \geq 0$  for all  $N \geq 1$ ,  $f \in \mathcal{F}$ . Likewise all finite dimensional covariance matrices of  $\mathbb{G}^{(N)}(\mathcal{F})$  are decreasing compared to the initial one. The strong approximation of  $\alpha_n^{(N)}$  by  $\mathbb{G}^{(N)}$  established by Theorem 2.1 of [1] further shows that the bias  $\mathbb{E}(\mathbb{P}_n^{(N)}(f)) - P(f)$  is uniformly small and provides rates of uniform quadratic risk reduction over  $\mathcal{F}$ . If recursive loops are performed among  $p$  partitions with known probabilities, for  $n$  sufficiently large the finite covariance matrices of  $\alpha_n^{(kp)}$  decrease at each loop  $k$ . We also compute – see Theorem 2.2 of [1] – a simple expression for  $\mathbb{G}^{(N)}$  as  $N \rightarrow +\infty$  when raking with two partitions alternatively – the basic case of a two-way contingency table.

**Raking the bootstrapped empirical process.** According to the bootstrap paradigm, in order to estimate the distribution of  $\alpha_n^{(N)}$  one has to re-sample according to the weighted  $\mathbb{P}_n$  then apply the  $N$ -th order raking-ratio procedure to  $\mathbb{P}_n^*$ . This gives access by Monte-Carlo approaches to the unknown distribution of  $\mathbb{G}^{(N)}$  – useful since  $P$  is unknown. Define the raked bootstrapped empirical measure to be  $\mathbb{P}_n^{*(0)} = \mathbb{P}_n^*$  then, recursively and conditionally to  $X_1, \dots, X_n$  on the event  $B_n^{2,N} = \{\min_{1 \leq j \leq m_N} \mathbb{P}_n^{*(N-1)}(A_j^{(N)}) > 0\}$ ,

$$\mathbb{P}_n^{*(N)}(f) = \sum_{j=1}^{m_N} \frac{\mathbb{P}_n(A_j^{(N)})}{\mathbb{P}_n^{*(N-1)}(A_j^{(N)})} \mathbb{P}_n^{*(N-1)}(f \mathbf{1}_{A_j^{(N)}}), \quad (1.10)$$

$$\alpha_n^{*(N)}(f) = \sqrt{n}(\mathbb{P}_n^{*(N)}(f) - \mathbb{P}_n(f)), \quad f \in \mathcal{F}. \quad (1.11)$$

The centering with respect to  $\mathbb{P}_n$  in (1.11) should be discussed. In (1.2) the centering  $\mathbb{P}_n$  stands for the conditional expectation of  $\mathbb{P}_n^*$  and plays the role of the expectation  $P$  of  $\mathbb{P}_n$  in  $\alpha_n$ . On the opposite, there is a bias inherent to the raking-ratio procedure so that  $P$  is no more the expectation of  $\mathbb{P}_n^{(N)}$  in (1.6). This bias was established by Proposition 5 of [1] to be uniformly small, hence  $P$  is confirmed as the targeted probability measure and the limiting process is centered. In order to simulate the influence of this bias we center the bootstrap (1.11) on  $\mathbb{P}_n$ . Therefore  $\mathbb{P}_n^{*(N)}$  and  $\alpha_n^{*(N)}$  use the auxiliary information  $\mathbb{P}_n(\mathcal{A}^{(N)}) = (\mathbb{P}_n(A_1^{(N)}), \dots, \mathbb{P}_n(A_{m_n}^{(N)}))$  instead of the original  $P(\mathcal{A}^{(N)})$  and satisfy

$$\mathbb{P}_n^{*(N)}(A_j^{(N)}) = \mathbb{P}_n(A_j^{(N)}), \quad \alpha_n^{*(N)}(A_j^{(N)}) = 0, \quad j = 1, \dots, m_N.$$

**Bootstrapping the raked empirical process.** A way to exploit directly the information of partitions is to bootstrap by using a probability that is possibly

closer to  $P$  than  $\mathbb{P}_n$  is, namely  $\mathbb{P}_n^{(N)}$ . Let  $T_n^{(N)}/n$  denote the mean of  $Z_1, \dots, Z_n$  under the discrete measure  $\mathbb{P}_n^{(N)}$ , that is  $T_n^{(N)} = \sum_{i=1}^n n\mathbb{P}_n^{(N)}(\{X_i\})Z_i$ . In particular,  $T_n^{(0)} = T_n$ . Given  $N \in \mathbb{N}$ , define the bootstrapped  $N$ -th order raked empirical measure and process to be, on the event  $B_n^{1,N}$ ,

$$\begin{aligned}\mathbb{P}_n^{(N)*}(f) &= \sum_{i=1}^n \frac{n\mathbb{P}_n^{(N)}(\{X_i\})Z_i}{T_n^{(N)}} f(X_i), \\ \alpha_n^{(N)*}(f) &= \sqrt{n}(\mathbb{P}_n^{(N)*}(f) - \mathbb{P}_n^{(N)}(f)), \quad f \in \mathcal{F}.\end{aligned}\tag{1.12}$$

This reproducible imitation of  $\alpha_n$  is a variant of (1.2).

## 2 Main results

### 2.1 The class $\mathcal{F}$

In all this paper, one assume that  $\mathcal{X}$  is measurable with  $\sigma$ -field  $\mathcal{A}$ . For approximations results  $\mathcal{X}$  is not required to be metric separable nor  $\mathcal{A}$  be Borel sets, however this may helps differentiability questions for the statistics to be bootstrapped. Let  $\mathcal{M}$  be the set of measurable real valued functions on  $(\mathcal{X}, \mathcal{A})$  and  $\tilde{\mathcal{F}} \subset \mathcal{F} \subset \mathcal{M}$  be such that  $\sup_{f \in \mathcal{F}} |f| \leq M_{\mathcal{F}} < +\infty$ ,  $\tilde{\mathcal{F}}$  is countable and each  $f \in \mathcal{F}$  is the point-wise limit of a sequence belonging to  $\tilde{\mathcal{F}}$ . This condition ensures that the empirical process  $\alpha_n$  and the variants defined in the previous section are point-wise separable and hence ball measurable, which allows to restrict their weak convergence to ball measurable test maps and avoid outer probabilities – see example 2.3.4 of [35].

For a probability measure  $Q$  on  $(\mathcal{X}, \mathcal{A})$  we endow  $\mathcal{M}$  with the semi-metric  $dQ$  defined by  $dQ^2(f, g) = \int_{\mathcal{X}} (f - h)^2 dQ$ . Let  $N(\mathcal{F}, \varepsilon, d)$  be the minimum number of balls of radius  $\varepsilon$  for the semi-distance  $d$  needed to cover  $\mathcal{F}$ . Let  $N_{[\ ]}(\mathcal{F}, \varepsilon, d)$  be the minimum number of  $\varepsilon$ -brackets for  $d$  necessary to cover  $\mathcal{F}$ . One assume throughout the paper that  $\mathcal{F}$  satisfies either (VC) or (BR) below, which encompass many useful examples. Typically (VC) concerns small classes well behaved with respect to all  $P$ , like VC-classes. On the opposite (BR) concerns classes that are very large or well behaved only with respect to a few  $P$ .

**Hypothesis (VC).** *There exists  $c_0 > 0$ ,  $\nu_0 > 0$  such that  $\sup_Q N(\mathcal{F}, \varepsilon, dQ) \leq c_0 \varepsilon^{-\nu_0}$  where the supremum is taken over all probability measure  $Q$  on  $(\mathcal{X}, \mathcal{A})$ . Define  $\alpha_0 = 1/(2 + 5\nu_0) \in (0, 1/2)$  and  $\beta_0 = (4 + 5\nu_0)/(4 + 10\nu_0)$ .*

**Hypothesis (BR).** *There exists  $b_0 > 0$ ,  $0 < r_0 < 1$  such that  $N_{[\ ]}(\mathcal{F}, \varepsilon, dP) \leq \exp(b_0^2 \varepsilon^{-2r_0})$ . Define  $\gamma_0 = (1 - r_0)/2r_0$ .*

Under (VC) or (BR) the process  $\alpha_n(\mathcal{F})$  converges weakly to the  $P$ -Brownian bridge  $\mathbb{G}(\mathcal{F})$  in  $\ell^\infty(\mathcal{F})$  endowed with the sup-norm  $\|H\|_{\mathcal{F}} = \sup_{f \in \mathcal{F}} |H(f)|$ . According to Propositions 1 and 2 of Berthet and Mason [8] the rate of weak convergence is of order at most  $n^{-\alpha_0}(\log n)^{\beta_0}$  under (VC) and  $(\log n)^{-\gamma_0}$  under

(BR). A slightly improvement and interpolation of these rates and conditions are possible, however we focus on our main topic. Thus we keep the original  $\alpha_0, \beta_0, \gamma_0$  and uniform boundedness – instead of square integrable envelope function – and it would suffice to substitute improved rates of approximation with no change in our statements. Recall that  $\mathbb{G}(\mathcal{F})$  is a centered Gaussian linear process with covariance

$$\text{Cov}(\mathbb{G}(f), \mathbb{G}(g)) = P(fg) - P(f)P(g), \quad f, g \in \mathcal{F}.$$

Write  $\sigma_f^2 = \text{Var}(f(X)) = P(f^2) - P(f)^2$  and  $\sigma_{\mathcal{F}}^2 = \sup_{f \in \mathcal{F}} \sigma_f^2 < +\infty$ .

## 2.2 Strong approximations

For  $n \in \mathbb{N}$ , we write  $LL(n) = L(L(n))$  with  $L(x) = \max(1, \log(x))$ .

**Proposition 1.** *Assume (1.3), (1.4) and either (VC) or (BR). There exists a finite  $K = K(\mathcal{F}, P^{(X,Z)}) > 0$  such that*

$$\limsup_{n \rightarrow +\infty} \sqrt{\frac{n}{LL(n)}} \|\mathbb{P}_n^* - P\|_{\mathcal{F}} \leq K \quad a.s.$$

Next define  $u_n = n, v_n = n^{-\alpha_0}(\log n)^{\beta_0}$  if  $\mathcal{F}$  satisfies (VC) and  $u_n = \log n, v_n = (\log n)^{-\gamma_0}$  if  $(\mathcal{F}, P)$  satisfy (BR). Berthet and Mason (see Propositions 1 and 2 of [8]) proved that we can construct a probability space on which the sequence of empirical process  $\alpha_n(\mathcal{F})$  can be defined together with a coupling sequence of  $P$ -Brownian bridges  $\mathbb{G}_n(\mathcal{F})$  such that, almost surely,  $\|\alpha_n - \mathbb{G}_n\|_{\mathcal{F}} \leq C v_n$  for some  $C > 0$  and all  $n$  large enough. By applying this strong approximation to  $P^{(X,Z)}$  we obtain the following version for the weighted bootstrap empirical process  $\alpha_n^*$ , which we formulate in the same way as in [8].

**Theorem 2.1.** *Assume (1.3), (1.4) and either (VC) or (BR). For all  $\theta > 0$  there exists  $C_\theta > 0, n_\theta > 0$  and a probability space supporting a sequence  $\{(X_n, Z_n)\}$  of i.i.d. random variables with distribution  $P^{(X,Z)}$  and a sequence  $\{(\mathbb{G}_n(\mathcal{F}), \mathbb{G}_n^*(\mathcal{F}))\}$  of pairs of independent  $P$ -Brownian bridges such that for all  $n > n_\theta$ ,*

$$\mathbb{P}(\{\|\alpha_n - \mathbb{G}_n\|_{\mathcal{F}} > C_\theta v_n\} \cup \{\|\alpha_n^* - \mathbb{G}_n^*\|_{\mathcal{F}} > C_\theta v_n\}) \leq \frac{1}{n^\theta}. \quad (2.1)$$

*On the same probability space as above there also exists independent sequences  $\{\mathbb{G}'_n(\mathcal{F})\}$  and  $\{\mathbb{G}^{*\prime}_n(\mathcal{F})\}$  of independent versions of  $\mathbb{G}(\mathcal{F})$  such that*

$$\mathbb{P}\left(\frac{1}{\sqrt{n}} \max_{1 \leq k \leq n} \|\sqrt{k}\alpha_k - \sum_{i=1}^k \mathbb{G}'_i\|_{\mathcal{F}} > C_\theta v_n\right) \leq \frac{1}{u_n^\theta}, \quad (2.2)$$

$$\mathbb{P}\left(\frac{1}{\sqrt{n}} \max_{k_{n,\theta} \leq k \leq n} \|\sqrt{k}\alpha_k^* - \sum_{i=1}^k \mathbb{G}^{*\prime}_i\|_{\mathcal{F}} > C_\theta v_n\right) \leq \frac{1}{u_n^\theta}, \quad (2.3)$$

*for all  $n \geq \max(n_\theta, k_{n,\theta})$ , where  $k_{n,\theta} = \lceil 2M_{\mathbb{Z}}^2(\ln(8) + (1 + \theta) \ln n) \rceil$ .*

The independence of  $\mathbb{G}_n$  and  $\mathbb{G}_n^*$  comes from Lemma 2. The bridges  $\mathbb{G}_n^*$  (resp.  $\mathbb{G}_n$ ) are not pair-wise independent, whereas by construction the  $\mathbb{G}'_n^*$  (resp.  $\mathbb{G}'_n$ ) are mutually independent. Notice that Theorem 2.1 applied with  $\theta > 1$  not only implies that  $\alpha_n^*$  almost surely weakly converge to  $\mathbb{G}$  but also provides rates of weak convergence. In particular this establishes the asymptotic independence of the processes  $\alpha_n$  and  $\alpha_n^*$ . Moreover, by substituting  $\mathbb{G}_n$  and  $\mathbb{G}_n^*$  their residual orthogonality is quantified through

$$\limsup_{n \rightarrow +\infty} \frac{1}{v_n} \sup_{f, g \in \mathcal{F}} |\text{Cov}(\alpha_n(f), \alpha_n^*(g))| < +\infty \quad a.s.$$

Consider now the Monte-Carlo experiment of  $b_n$  iterated bootstraps. Let  $\mathbb{P}_{n,(j)}^*$  denote the  $j^{\text{th}}$  independent bootstrapped empirical measure built from weights  $(Z_{1,(j)}, \dots, Z_{n,(j)})$  drawn from  $P^{(Z|X=X_i)} \times \dots \times P^{(Z|X=X_n)}$  conditionally to  $X_1, \dots, X_n$ . Write  $\alpha_{n,(j)}^*(f) = \sqrt{n}(\mathbb{P}_{n,(j)}^*(f) - \mathbb{P}_n(f))$  the associated empirical process. For  $b_n = 1$  the following result reduces to Theorem 2.1, otherwise the rates of approximation are slowed down. If  $\mathcal{F}$  satisfies (VC) then define

$$w_n = \left( \frac{b_n^5}{n} (\log n)^2 \right)^{\alpha_0} \left( \log \left( \frac{n}{b_n^5} \right) \right)^{5v/(4+10v)} \leq \left( \frac{b_n^5}{n} \right)^{\alpha_0} (\log n)^{\beta_0}.$$

If  $(\mathcal{F}, P)$  satisfy (BR) then define

$$w_n = \left( \frac{1}{\log(n/b_n^5)} \right)^{\gamma_0} \geq \left( \frac{1}{\log n} \right)^{\gamma_0}.$$

**Theorem 2.2.** *Assume (1.3), (1.4) and either (VC) or (BR). Let  $b_n \in \mathbb{N}_*$  be such that  $b_n/n^{1/5} \rightarrow 0$ . For all  $\theta > 0$  there exists  $C_\theta > 0$ ,  $n_\theta > 0$  and a probability space supporting a triangular array  $\{(X_n, Z_{n,(1)}, \dots, Z_{n,(b_n)})\}$  of i.i.d. random vectors distributed as  $P^{(X_1, Z_{1,(1)}, \dots, Z_{1,(b_n)})}$  and a triangular array  $\{(\mathbb{G}_{n,(0)}^*(\mathcal{F}), \mathbb{G}_{n,(1)}^*(\mathcal{F}), \dots, \mathbb{G}_{n,(b_n)}^*(\mathcal{F}))\}$  of  $(b_n + 1)$ -uplets of mutually independent  $P$ -Brownian bridges such that, for  $n \geq n_\theta$ ,*

$$\mathbb{P} \left( \left\{ \|\alpha_n - \mathbb{G}_{n,(0)}^*\|_{\mathcal{F}} > C_\theta w_n \right\} \cup \bigcup_{j=1}^{b_n} \left\{ \|\alpha_{n,(j)}^* - \mathbb{G}_{n,(j)}^*\|_{\mathcal{F}} > C_\theta w_n \right\} \right) \leq \frac{1}{n^\theta}. \quad (2.4)$$

The coupling we perform implies that  $\mathbb{G}_{m,(j)}^*$  and  $\mathbb{G}_{n,(k)}^*$  are dependent if  $m \neq n$  and independent if  $m = n$  and  $j \neq k$ . The following uniform central limit theorem immediately follows, if  $b_n = b$  is fixed. For  $b$  bootstraps define the  $\mathbb{R}^{b+1}$ -valued empirical process indexed by  $\mathcal{F}^{b+1}$  to be

$$\Lambda_{n,b}(\mathcal{F}^{b+1}) = \{(\alpha_n(f_0), \alpha_{n,(1)}^*(f_1), \dots, \alpha_{n,(b)}^*(f_b)) : f = (f_0, f_1, \dots, f_b) \in \mathcal{F}^{b+1}\}.$$

Consider any norm  $\|\cdot\|$  on  $\mathbb{R}^{b+1}$  then endow  $\ell^\infty(\mathcal{F}^{b+1} \rightarrow \mathbb{R}^{b+1})$  with the distance associated to the sup-norm  $\|\Lambda\|_{b+1} = \sup_{f \in \mathcal{F}^{b+1}} \|\Lambda(f)\|$ .

**Proposition 2.** *Under the assumptions of Theorem 2.2, for any fixed  $b \in \mathbb{N}_*$  the sequence  $\Lambda_{n,b}(\mathcal{F}^{b+1})$  converges weakly in  $\ell^\infty(\mathcal{F}^{b+1} \rightarrow \mathbb{R}^{b+1})$  to*

$$\mathbb{G}_b(\mathcal{F}^{b+1}) = \{(\mathbb{G}_{(0)}^*(f_0), \mathbb{G}_{(1)}^*(f_1), \dots, \mathbb{G}_{(b)}^*(f_b)) : f = (f_0, f_1, \dots, f_b) \in \mathcal{F}^{b+1}\}$$

where  $\mathbb{G}_{(0)}^*, \mathbb{G}_{(1)}^*, \dots, \mathbb{G}_{(b)}^*$  are mutually independent  $P$ -Brownian bridges.

Theorem 2.2 also implies that the distance in distribution between  $\Lambda_{n,b_n}(\mathcal{F}^{b_n+1})$  and  $\mathbb{G}_{n,b_n}(\mathcal{F}^{b_n+1}) = \{(\mathbb{G}_{(0)}^*(f_0), \mathbb{G}_{(1)}^*(f_1), \dots, \mathbb{G}_{(b)}^*(f_b)) : f \in \mathcal{F}^{b_n+1}\}$  is at most  $O(w_n)$ , which is severely impacted by  $b_n$  and requires that  $b_n/n^{1/5} \rightarrow 0$ . Let  $d_{PL,n}(\mu, \nu)$  denote the Prokhorov-Lévy distance between two probability measures  $(\mu, \nu)$  on  $\ell^\infty(\mathcal{F}^{b_n+1} \rightarrow \mathbb{R}^{b_n+1})$ .

**Proposition 3.** *Under the assumptions of Theorem 2.2 we have*

$$d_{PL,n}(\Lambda_{n,b_n}(\mathcal{F}^{b_n+1}), \mathbb{G}_{n,b_n}(\mathcal{F}^{b_n+1})) = O(w_n).$$

**Comments.** Proposition 2 shows that for one bootstrap experiment the biased empirical process  $\sqrt{n}(\mathbb{P}_{n,(1)}^*(f) - P(f))$  is asymptotically distributed as  $\mathbb{G}_{(0)}^* + \mathbb{G}_{(1)}^*$  and hence  $\sqrt{2}\mathbb{G}$  so that its asymptotic variance is  $2\text{Var}(f)$ . Moreover, thanks to Theorem 2.2 the study of weak convergence of functionals estimated by  $b_n$  bootstrap experiments which are conditionally independent versions of  $\alpha_n^*$  is made easier by substituting the  $\mathbb{G}_{n,(j)}^*$  to the  $\alpha_{n,(j)}^*$ . Furthermore, when  $n$  is large the distributions of the statistics of interest are typically very concentrated and nearly Gaussian, thus bootstrapping only  $b_n$  times with  $b_n/n^{1/5} \rightarrow 0$  is not so penalizing once the uniform performance is guaranteed by Theorem 2.2. The assumption (1.4) could be relaxed in Theorems 2.1 and 2.2 by taking into account the tail behavior of  $Z$  through the function  $\psi_Z(z) = -\log \mathbb{P}(Z > z)$ ,  $z > 0$ , however at the cost of additional technicalities.

The above coupling results can be applied to control in general estimators bootstrapped  $b_n$  times and expressed as smooth transforms of  $\mathbb{P}_n^*(\mathcal{F})$ . In 2.3 we study in particular the variance and the distribution function uniformly over a class of such estimators. In 2.4 we derive uniform over  $\mathcal{F}$  Berry-Esseen type bounds.

### 2.3 Estimation of variance and distribution function

**Motivation.** In a bootstrap Monte-Carlo procedure the  $\mathbb{P}_n^*(f)$  have to be reevaluated  $b_n$  times by redrawing the  $n$  weights  $Z_i$  according to the conditional distribution  $P^{(Z|X=X_i)}$ . Classically  $b_n = 1$  and the bootstrapped moments of a smooth transform of several mean estimators  $\mathbb{P}_n^*(f)$  is evaluated by Edgeworth expansions. When  $b_n = b > 1$  estimators of the moments of  $\mathbb{P}_n^*(f)$  have also been studied, but typically not jointly. For instance, Booth and Sarkar [11] showed that the distribution of the bootstrap estimator of the variance of a statistic is approximately a chi-squared distribution with  $b - 1$  degrees of freedom. They deduce  $b > 1$  necessary to obtain a relative error less than a fixed bound with some probability, assuming  $n$  large – for an error less than 10% with probability

0.95 about  $b = 800$  are required. Chandra and Ghosh [13] studied the distributions of smooth functions of the empirical mean by using Edgeworth expansions and proved that these statistics converge to a centered Chi-squared distribution. In the same spirit, Babu [5] showed that the bootstrapped version of smooth transforms of a single empirical mean has a similar weak asymptotic behavior by assuming that the functional is three times continuously differentiable.

Let us extend these results to the case of functionals of the empirical measure itself instead of one or several empirical means, under a generic assumption of first order differentiability on the space of measures.

**Hypothesis.** We study the bootstrap estimation of moments and distribution of a statistic  $S_n = \varphi(\mathbb{P}_n)$  with  $\varphi$  differentiable in the following weak sense. Let  $\varphi : \ell^\infty(\mathcal{F}) \rightarrow \mathbb{R}$  be a real-valued function such that, for all  $Q \in \ell^\infty(\mathcal{F})$ ,

$$\varphi(P + Q) = \varphi(P) + \varphi'(P) \cdot Q + R(Q), \quad (2.5)$$

where  $\varphi'(P) \in \mathcal{L}(\ell^\infty(\mathcal{F}), \mathbb{R})$  is a linear application satisfying  $|\varphi'(P) \cdot Q| \leq \|Q\|_{\mathcal{F}}$ ,  $R : \ell^\infty(\mathcal{F}) \rightarrow \mathbb{R}$  is an application such that in a ball  $\mathcal{B}$  centered on the zero function  $|R(Q)| \leq \|Q\|_{\mathcal{F}}^q$  for some  $q > 1$ . The local expansion (2.5) includes Frechet differentiability and examples pages 11-16 in Barbe and Bertail [6]. The key factor showing up in this setting is  $\sigma^2 = \text{Var}(\varphi'(P) \cdot \mathbb{G})$ , with  $\mathbb{G}$  the  $P$ -Brownian bridge. One could refer to  $\sigma^2$  as a conditional coefficient of variation. The following table makes (2.5) explicit for some classical estimators, in particular the associated finite classes and the differential  $\varphi'(P) \cdot \mathbb{G}$ .

$S_n = \varphi(\mathbb{P}_n)$	$\mathcal{F}$	$\varphi'(P) \cdot \mathbb{G}$	$\sigma^2$
Mean $\mathbb{P}_n(f_0)$	$\{f_0\}$	$\mathbb{G}(f_0)$	$\text{Var}(f_0(X))$
Variance $\mathbb{P}_n(f_0^2) - \mathbb{P}_n(f_0)^2$	$\{f_0^2, f_0\}$	$\mathbb{G}(f_0^2)$	$\text{Var}(f_0^2(X))$
Inverse mean $\frac{1}{\mathbb{P}_n(f_0)}$	$\{f_0\}$	$\frac{-1}{P^2(f_0)} \mathbb{G}(f_0)$	$\frac{\text{Var}(f_0(X))}{P^4(f_0)}$
Conditional mean $\frac{\mathbb{P}_n(f_0 \mathbb{1}_A)}{\mathbb{P}_n(A)}$	$\{f_0 \mathbb{1}_A, \mathbb{1}_A\}$	$\frac{\mathbb{G}(f_0 \mathbb{1}_A) - P(f_0 A)\mathbb{G}(A)}{P(A)}$	$\frac{\text{Var}((f_0(X) - \mathbb{E}[f_0(X) A]) \mathbb{1}_{X \in A})}{P(A)^2}$

**Variance and distribution estimations by bootstrap.** Write  $S_{n,(j)}^* = \varphi(\mathbb{P}_{n,(j)}^*)$ , for  $1 \leq j \leq b_n$ . Whenever  $S_n$  is asymptotically normal, or regular in terms of Edgeworth expansion,  $\text{Var}(S_n)$  is crucial to provide confident bands. The variance and distribution function of  $S_n$  could be estimated by the following bootstrap estimators

$$\widehat{\text{Var}}(S_n^*) = \frac{1}{b_n} \sum_{j=1}^{b_n} (S_{n,(j)}^* - S_n)^2 = \frac{1}{b_n} \sum_{j=1}^{b_n} \left( \varphi(\mathbb{P}_{n,(j)}^*) - \varphi(\mathbb{P}_n) \right)^2,$$

$$\widehat{F}_{S_n^*}(x) = \frac{1}{b_n} \sum_{j=1}^{b_n} \mathbb{1}_{\{S_{n,(j)}^* \leq x\}}.$$

**Results.** The next statement provides a confidence interval of  $\text{Var}(S_n)$  from the value  $\widehat{\text{Var}}(S_n^*)$ . In particular the statistic  $(nb_n/\sigma^2)\widehat{\text{Var}}(S_n^*)$  is asymptotically close to a  $\chi^2(b_n)$ .

**Corollary 1.** Assume (1.3), (1.4) and (VC). Let  $b_n \in \mathbb{N}_*$  be such that  $b_n/n^{1/5} \rightarrow 0$ . There exists  $C_0, n_0 > 0$  such that for all  $n > n_0$ ,

$$\mathbb{P}\left(\frac{n}{\sigma^2}|\widehat{\text{Var}}(S_n^*) - \text{Var}(S_n)| \geq \delta + C_0 w_n n^{\alpha_0/2}\right) \leq \alpha + \frac{1}{n^2}, \quad (2.6)$$

with  $\alpha, \delta$  such that  $\mathbb{P}\left(\left|\frac{\chi^2(b_n)}{b_n} - 1\right| \geq \delta\right) \leq \alpha$ . In particular, almost surely there exists  $C'_0 > 0, n'_0 = n'_0(\omega) > 0$  such that for all  $n > n'_0$ ,

$$|\widehat{\text{Var}}(S_n^*) - \text{Var}(S_n)| \leq \frac{C'_0}{\sqrt{b_n}}. \quad (2.7)$$

**Remark.** If  $b_n = b$  is fixed and the class  $\mathcal{F}$  is finite we get the same approximation for  $b$  as Booth and Sarkar [11] to have a relative error for  $\widehat{\text{Var}}(S_n^*)$  less than  $\delta$  with a probability greater than  $1 - \alpha$ , namely  $b \simeq 2|\Phi^{-1}(\alpha/2)|^2/\delta^2$ . Moreover Corollary 1 provides a second order control of both the width and probability of the confident interval for the variance, uniformly in  $n$  and for infinite classes.

The following result is a DKW-type inequality for bootstrap statistics. It evaluates the uniform deviation between the unknown distribution function of  $S_n$  and the estimated empirical distribution function  $\widehat{F}_{S_n^*}$  by also taking into account the shift due to the unavoidable bias  $B_n = S_n - \varphi(P)$ .

**Corollary 2.** Assume (1.3), (1.4) and either (BR) or (VC). Let  $b_n \in \mathbb{N}_*$  be such that  $b_n/n^{1/5} \rightarrow 0$ . Almost surely there exists  $C_0 > 0, n_0 = n_0(\omega) > 0$  such that for all  $n > n_0$ ,

$$\sup_{x \in A_n^C} |\widehat{F}_{S_n^*}(x + B_n) - F_{S_n}(x)| \leq C_0 \sqrt{\frac{\log n}{b_n}}, \quad (2.8)$$

with  $A_n = [\varphi(P) - a_n; \varphi(P) + a_n]$ ,  $a_n = \sigma \sqrt{\log(\log n / b_n w_n^2)}$ .

**Remark.** Let  $\Lambda$  denote the class of strictly increasing continuous mappings of  $\mathbb{R}$  onto itself. The  $J_1$ -topology is defined by the Skorokhod metric (see Billingsley [10])

$$d(F, G) = \inf_{\lambda \in \Lambda} \max\left(\sup_{x \in \mathbb{R}} |\lambda(x) - x|, \sup_{x \in \mathbb{R}} |G \circ \lambda(x) - F(x)|\right).$$

Let  $d_A(F, G)$  denote the above distance when the second supremum is only evaluated on  $x \in A$ . Since  $\sqrt{n}B_n \rightarrow_{\text{weak}} \phi'(P) \cdot \mathbb{G}$  then  $B_n = O_{p.s.}(\sqrt{\log(n)/n})$  and it follows from (2.8) that

$$\mathbb{P}\left(d_{A_n^C}(F_{S_n}, \widehat{F}_{S_n^*}) > C \sqrt{\frac{\log n}{b_n}}\right) \leq \frac{1}{n^2}.$$

## 2.4 Rates of weak convergence

**Local Berry-Esseen bounds.** Let  $\Phi$  be the distribution function of the standard normal law. Singh[31] – see also section 3.1.3. of [29] - established several Berry-Esseen type inequalities for the distribution of the sum of bootstrapped variables of a given sample with respect to the distribution function of the sum of the variables of this sample and then with respect to  $\Phi$ . In particular, under certain conditions he established that the uniform deviation between these distributions is almost-surely at most  $O(n^{-1/2})$ . With Edgeworth expansion techniques, Hall [18, 19] studied the leading term of the expansion of the uniform deviation conditionally to the sample and found the same rate as the statistic  $n^{-3/2}|\sum_{j=1}^n X_j^3| + n^{-2}\sum_{j=1}^n X_j^4$  converges to 0 – that is also  $O(n^{-1/2})$  if the fourth moment exists. In this section, we derive a Berry-Esseen inequality for the bootstrapped empirical process indexed by functions, i.e. uniform results among large classes of statistics, however with slower rates than the  $O(n^{-1/2})$  for a single smooth statistic.

**Uniform Berry-Esseen bounds.** Let  $\mathcal{L}$  be a set of Lipschitz functions defined on  $\ell^\infty(\mathcal{F})$  such that all  $\phi \in \mathcal{L}$  has a Lipschitz constant bounded by  $C_0 < +\infty$  and the density  $\phi(\mathbb{G}(f))$  is bounded by  $C_1 < +\infty$  for all  $f \in \mathcal{F}$ .

**Corollary 3.** *Assume (1.3), (1.4) and either (VC) or (BR). There exists  $C > 0, n_0 > 0$  such that for all  $n > n_0$ ,*

$$\sup_{\phi \in \mathcal{L}} \sup_{f \in \mathcal{F}} \sup_{x \in \mathbb{R}} |\mathbb{P}(\phi(\alpha_n^*(f)) \leq x) - \mathbb{P}(\phi(\mathbb{G}(f)) \leq x)| \leq CC_0 C_1 v_n, \quad (2.9)$$

$$\sup_{\phi \in \mathcal{L}} \sup_{f \in \mathcal{F}} \sup_{x \in \mathbb{R}} |\mathbb{P}(\phi(\alpha_n^*(f)) \leq x) - \mathbb{P}(\phi(\alpha_n(f)) \leq x)| \leq CC_0 C_1 v_n. \quad (2.10)$$

In particular, if  $\tilde{\sigma}_{\mathcal{F}}^2 = \inf_{f \in \mathcal{F}} \text{Var}(f(X)) > 0$  then

$$\begin{aligned} \sup_{x \in \mathbb{R}} \sup_{f \in \mathcal{F}} \left| \mathbb{P} \left( \frac{\alpha_n^*(f)}{\sqrt{\text{Var}(f(X))}} \leq x \right) - \Phi(x) \right| &\leq \frac{C}{\sqrt{2\pi\tilde{\sigma}_{\mathcal{F}}}} v_n, \\ \sup_{x \in \mathbb{R}} \sup_{f \in \mathcal{F}} \left| \mathbb{P} \left( \frac{\alpha_n^*(f)}{\sqrt{\text{Var}(f(X))}} \leq x \right) - \mathbb{P} \left( \frac{\alpha_n(f)}{\sqrt{\text{Var}(f(X))}} \leq x \right) \right| &\leq \frac{C}{\sqrt{2\pi\tilde{\sigma}_{\mathcal{F}}}} v_n. \end{aligned}$$

## 3 Raking-Ratio results

### 3.1 Strong approximation of $\alpha_n^{*(N)}$

Fix  $N_0 \in \mathbb{N}^*$  and denote  $P_{N_0} = \prod_{N=1}^{N_0} p_N, M_{N_0} = \prod_{N=1}^{N_0} m_N$ . The bootstrapped empirical measure associated with the Raking-Ratio method assumes the following law of iterated logarithm:

**Proposition 4.** *If  $\mathcal{F}$  satisfies (VC) or (BR) then there exists  $K = K(\mathcal{F}, Z) > 0$  such that,*

$$\limsup_{n \rightarrow +\infty} \max_{1 \leq N \leq N_0} \sqrt{\frac{n}{LL(n)}} \|\mathbb{P}_n^{*(N)} - P\|_{\mathcal{F}} \leq Kb_{N_0} \quad a.s.,$$



where  $b_0 = 1$  and  $b_{N_0} = \prod_{N=1}^{N_0} (1 + M_{\mathcal{F}}/p_N)$ .

The bootstrapped empirical process  $\alpha_n^{*(N)}$  also satisfies the following Talagrand type concentration inequalities:

**Proposition 5.** *Let  $K_{\mathcal{F}} = \max(1, M_{\mathcal{F}})$  and assume (1.3), (1.4) and that  $\alpha_n^{(N)}$  and  $\alpha_n^{*(N)}$  are both defined for every  $0 \leq N \leq N_0$ . For all  $n > 0, \lambda > 0$  we have*

$$\begin{aligned} \mathbb{P} \left( \max_{0 \leq N \leq N_0} \|\alpha_n^{*(N)}\|_{\mathcal{F}} > \lambda \right) &\leq 4^{N_0+1} N_0^2 M_{N_0} \mathbb{P} \left( \|\alpha_n\|_{\mathcal{F}} > \frac{\lambda P_{N_0}^2}{3M_{N_0}(4K_{\mathcal{F}})^{N_0}(1 + \lambda/\sqrt{n})^{2N_0}} \right) \\ &\quad + 2^{N_0+2} N_0 \exp \left( -\frac{\lambda^2 P_{N_0}^4}{18M_{N_0}^2 M_{\mathcal{F}}^2 M_{\mathcal{F}}^2 (4K_{\mathcal{F}})^{2N_0} (1 + \lambda/\sqrt{n})^{4N_0}} \right). \end{aligned} \quad (3.1)$$

Under (BR) it holds, for  $n > n_0$  and  $\lambda_0 < \lambda < D_0\sqrt{n}$ ,

$$\mathbb{P} \left( \max_{0 \leq N \leq N_0} \|\alpha_n^{*(N)}\|_{\mathcal{F}} > \lambda \right) \leq D_1 \exp(-D_2\lambda^2), \quad (3.2)$$

for some positive values  $\lambda_0, D_0, D_1, D_2$ . Under (VC) it holds, for  $n > n_0$  and  $\lambda_0 < \lambda < 2M_{\mathcal{F}}\sqrt{n}$ ,

$$\mathbb{P} \left( \max_{0 \leq N \leq N_0} \|\alpha_n^{*(N)}\|_{\mathcal{F}} > \lambda \right) \leq D_3 \lambda_0^{\nu_0} \exp(-D_2\lambda^2), \quad (3.3)$$

for some positive values  $\lambda_0, D_3, D_4$ .

The following theorem establishes the strong approximation of  $\alpha_n^{*(N)}$  to  $\mathbb{G}^{(N)}$  defined by (1.7).

**Theorem 3.1.** *Assume (1.3), (1.4) and either (VC) or (BR). Let  $\theta > 0$ . There exists  $C_{\theta}, n_{\theta} > 0$  and a probability space supporting a sequence  $\{(X_n, Z_n)\}$  of i.i.d. random variables with distribution  $P^{(X,Z)}$  and sequences  $\{(\mathbb{G}_n^{(N)}(\mathcal{F}), \mathbb{G}_n^{*(N)}(\mathcal{F}))\}$  of pairs of independent versions of  $\mathbb{G}^{(N)}(\mathcal{F})$  for all  $0 \leq N \leq N_0$  such that for all  $n > n_{\theta}$ ,*

$$\mathbb{P} \left( \max_{1 \leq N \leq N_0} \|\alpha_n^{*(N)} - \mathbb{G}_n^{*(N)}\|_{\mathcal{F}} > C_{\theta} v_n \right) < \frac{1}{n^{\theta}}, \quad (3.4)$$

$$\mathbb{P} \left( \max_{1 \leq N \leq N_0} \|\alpha_n^{(N)} - \mathbb{G}_n^{(N)}\|_{\mathcal{F}} > C_{\theta} v_n \right) < \frac{1}{n^{\theta}}. \quad (3.5)$$

A stronger result of this result keeping the same spirit as Theorem ?? is the following:

**Theorem 3.2.** *Assume (1.3), (1.4) and either (VC) or (BR). Let  $\theta > 0$  and  $b_n \in \mathbb{N}_*$  be such that  $b_n/n^{1/5} \rightarrow 0$ . There exists  $C_{\theta} > 0, n_{\theta} > 0$  and a probability space supporting a triangular array  $\{(X_n, Z_{n,(1)}, \dots, Z_{n,(b_n)})\}$  of*

*i.i.d.* random vectors distributed as  $P^{(X_1, Z_{1,(1)}, \dots, Z_{1,(b_n)})}$  and a triangular array  $\{(\mathbb{G}_{n,(0)}^*(\mathcal{F}), \mathbb{G}_{n,(1)}^*(\mathcal{F}), \dots, \mathbb{G}_{n,(b_n)}^*(\mathcal{F}))\}$  of  $(b_n + 1)$ -uplets of mutually independent  $P$ -Brownian bridges such that, for  $n \geq n_\theta$ ,

$$\begin{aligned} \mathbb{P} \left( \max_{j=1, \dots, b_n} \max_{1 \leq N \leq N_0} \|\alpha_{n,(j)}^{*(N)} - \mathbb{G}_{n,(j)}^{*(N)}\|_{\mathcal{F}} > C_\theta w_n \right) &< \frac{1}{n^\theta}, \\ \mathbb{P} \left( \max_{j=1, \dots, b_n} \max_{1 \leq N \leq N_0} \|\alpha_{n,(j)}^{(N)} - \mathbb{G}_{n,(j)}^{(N)}\|_{\mathcal{F}} > C_\theta w_n \right) &< \frac{1}{n^\theta}. \end{aligned}$$

In particular, Theorem 3.1 implies that the bootstrapped empirical process associated with this method  $\alpha_n^{*(N)}$  converges weakly in  $\ell^\infty(\mathcal{F})$  to the Gaussian process  $\mathbb{G}^{(N)}$  as the same way as  $\alpha_n^{(N)}$ . A simple way to simulate the law of the Raking-Ratio process is to bootstrap the initial available sample and apply the Raking-Ratio method. We can therefore estimate the covariance or variance of the raking-ratio empirical process by Monte-Carlo method without applying (1.8). The next result is a corollary of Theorem 3.2 and gives details about the precision of these estimators:

**Corollary 4.** *Assume (1.3), (1.4) and (VC). Let  $b_n \in \mathbb{N}_*$  be such that  $b_n/n^{1/5} \rightarrow 0$ . Almost surely there exists  $C_0 > 0, n_0 = n_0(\omega) > 0$  such that for all  $n > n_0$  it holds*

$$\begin{aligned} \max_{0 \leq N \leq N_0} \sup_{f, g \in \mathcal{F}} \left| \frac{1}{b_n} \sum_{j=1}^{b_n} (\mathbb{P}_{n,(j)}^{*(N)}(f) - \mathbb{P}_n(f)) (\mathbb{P}_{n,(j)}^{*(N)}(g) - \mathbb{P}_n(g)) - \frac{1}{n} \text{Cov}(\mathbb{G}^{(N)}(f), \mathbb{G}^{(N)}(g)) \right| \\ \leq \frac{C_0 \log n}{n}, \quad (3.6) \end{aligned}$$

$$\max_{0 \leq N \leq N_0} \sup_{f \in \mathcal{F}} \left| \frac{1}{b_n} \sum_{j=1}^{b_n} (\mathbb{P}_{n,(j)}^{*(N)}(f) - \mathbb{P}_n(f))^2 - \frac{1}{n} \text{Var}(\mathbb{G}^{(N)}(f)) \right| \leq \frac{C_0 \log n}{n}. \quad (3.7)$$

**Remark.** *Proposition 4 is an interesting result since we could estimate the variance and covariance of the raking-ratio process, and so the efficiency of an auxiliary information, without knowing probabilities  $P(\mathcal{A}^{(N)})$  for  $N \in \mathbb{N}_*$ . If the auxiliary information has a cost – for example if this information is provided by experts, statistical learning or by data purchasing – the statistician could test the efficiency of the information before paying it.*

### 3.2 Strong approximation of $\alpha_n^{(N)*}$

We defined our raked bootstrapped empirical process  $\alpha_n^{*(N)}$  by bootstrapping  $\alpha_n$  in a first time and raking among known partitions after and we have shown that this process has the same asymptotic behavior as the Raking-Ratio Gaussian process  $\mathbb{G}^{(N)}$ . So a natural question is whether the behavior is the same for  $\alpha_n^{(N)*}$ , if we apply the Raking-Ratio at first and bootstrapped in a second time by adding random weights. The answer to this question is negative, main reason

being that our process  $\alpha_n^{*(N)}$  no longer satisfies the constraints given by the auxiliary information  $\mathbb{P}_n(\mathcal{A}^{(N)})$  and has no reason to have the same asymptotic limit as that the non-bootstrapped process  $\alpha_n^{(N)}$ . However, we can prove that  $\alpha_n^{(N)*}$  has the same asymptotic behavior as  $\alpha_n^*$ .

**Proposition 6.** *Assume (1.3), (1.4) and either (VC) or (BR). For all  $\theta > 0$  there exists positive constants  $C_\theta, n_\theta$  and a probability space supporting a sequence  $\{(X_n, Z_n)\}$  of i.i.d. random variables with distribution  $P^{(X,Z)}$  and a sequence  $\{(\mathbb{G}_n(\mathcal{F}), \mathbb{G}_n^{(0)*}(\mathcal{F}), \dots, \mathbb{G}_n^{(N_0)*}(\mathcal{F}))\}$  of  $P$ -Brownian bridges such that for all  $n \geq n_\theta$ ,*

$$\mathbb{P} \left( \{ \|\alpha_n - \mathbb{G}_n\|_{\mathcal{F}} > C_\theta v_n \} \cup \bigcup_{N=0}^{N_0} \{ \|\alpha_n^{(N)*} - \mathbb{G}_n^{(N)*}\|_{\mathcal{F}} > C_\theta v_n \} \right) \leq \frac{1}{n^\theta}. \quad (3.8)$$

## 4 Proofs

By definition of  $v_n$  there exists  $\varepsilon \in (0, 1/2)$  such that  $v_n n^{1/2-\varepsilon}$  is of the form  $n^\alpha$  or  $n^\alpha (\log n)^\beta$  with some  $\alpha > 0, \beta \in \mathbb{R}$ . We fix this value of  $\varepsilon$  for this section.

### 4.1 Decomposition of $\alpha_n^*$

We introduce at Step 1 a simpler definition of the bootstrapped empirical process. This new process can be expressed through the classical empirical process indexed by a special function class. In fact, this class is an extension of  $\mathcal{F}$  and is defined at Step 2. At Step 3 we give the decomposition of this new process via the empirical process indexed by this class of function.

**Step 1.** In order to study  $\alpha_n^*(\mathcal{F})$  from (1.2) it is more convenient to first work with the conditionally centered version  $\tilde{\alpha}_n^*(\mathcal{F})$  that we define to be

$$\tilde{\alpha}_n^*(f) = \sqrt{n}(\tilde{\mathbb{P}}_n^*(f) - \bar{Z}_n \mathbb{P}_n(f)) = \frac{1}{\sqrt{n}} \sum_{i=1}^n (Z_i - \bar{Z}_n) f(X_i), \quad f \in \mathcal{F}, \quad (4.1)$$

where, by recalling also (1.1) and (1.2),

$$\tilde{\mathbb{P}}_n^*(f) = \frac{1}{n} \sum_{i=1}^n Z_i f(X_i), \quad \bar{Z}_n = \frac{T_n}{n} = \frac{1}{n} \sum_{i=1}^n Z_i, \quad \alpha_n^*(f) = \frac{n}{T_n} \tilde{\alpha}_n^*(f). \quad (4.2)$$

**Step 2.** Define  $\phi_M(z) = \max(-M, \min(M, z))$  for  $z \in \mathbb{R}$  and  $M > 0$  and let  $\mathbf{1}$  denote the application equal to 1 on  $\mathbb{R}$  or  $\mathcal{X}$ . Given  $f \in \mathcal{F}$  write  $h_f = h_f^{(1)} - P(f)h^{(2)} - h_f^{(3)}$  where

$$h_f^{(1)}(x, z) = \phi_{M_z}(z) f(x), \quad h^{(2)}(x, z) = \phi_{M_z}(z), \quad h_f^{(3)}(x, z) = f(x), \quad (4.3)$$

for  $(x, z) \in \mathcal{X} \times \mathbb{R}$ . The functions of (4.3) belong to the following extension of  $\mathcal{F}$

$$\mathcal{F}_1 = \{h(x, z) = f(x)g(z) : f \in \mathcal{F} \cup \{\mathbf{1}\}, g \in \{\phi_{M_Z}, \mathbf{1}\}\}. \quad (4.4)$$

The following lemma proves that the enlarged class  $\mathcal{F}_1$  has the same entropy as  $\mathcal{F}$ .

**Lemma 1.** *The class  $\mathcal{F}_1$  defined by (4.4) is uniformly bounded, point-wise measurable and satisfies (VC) or (BR) with respect to  $P^{(X,Z)}$  with the same powers as  $\mathcal{F}$  with respect to  $P$ , that is  $c_0, \nu_0, b_0, r_0$ .*

*Proof.* Let  $h = fg \in \mathcal{F}_1$  with  $f \in \mathcal{F} \cup \{\mathbf{1}\}$  and  $g \in \{\mathbf{1}, \phi_M\}$ . Clearly  $h$  is bounded by  $M_1 = \max(1, M_Z) \cdot \max(1, M_{\mathcal{F}}) < +\infty$ . For  $(x, z) \in \mathcal{X} \times \mathbb{R}$ ,  $f(x)$  is the limit of a sequence  $\{f_n(x)\}$  with  $f_n \in \tilde{\mathcal{F}}$ , so  $h(x, z)$  is the limit of  $f_n(x)g(z)$  where  $gf_n \in \tilde{\mathcal{F}} \times \{\mathbf{1}, \phi_M\}$  which is countable and independent of  $h$ . If  $\mathcal{F}$  is covered by  $N(\mathcal{F}, \varepsilon, d_P)$  balls then  $\mathcal{F}_1$  can be covered by  $2N(\mathcal{F}, \varepsilon/M_Z, d_{P(x,z)})$  because  $d_Q^2(h, f_0g) \leq M_Z^2 d_Q^2(f, f_0)$ . If  $\mathcal{F}$  is covered by  $N_{[\cdot]}(\mathcal{F}, \varepsilon, d_P)$   $\varepsilon$ -brackets then  $\mathcal{F}_1$  can be covered by  $2N_{[\cdot]}(\mathcal{F}, \varepsilon, d_{P(x,z)})$   $\varepsilon$ -brackets since  $f_-(x) < f(x) < f_+(x)$  implies  $f_-(x)g(z) < h(x, z) < f_+(x)g(z)$  for  $g$  positive. Hence if  $\mathcal{F}$  satisfies (VC) or (BR) it is the same for  $\mathcal{F}_1$  with larger constants  $c_0$  or  $b_0$ .  $\square$

**Step 3.** By (1.3) we have  $\mathbb{E}(Zf(X)) = P(f) = \mathbb{E}(ZP(f))$ . By (4.1),  $\tilde{\alpha}_n^*$  can be split into

$$\begin{aligned} \sqrt{n}\tilde{\alpha}_n^*(f) &= \sum_{i=1}^n f(X_i)Z_i - \frac{T_n}{n} \left( \sum_{i=1}^n f(X_i) \right) \\ &= \sum_{i=1}^n Z_i(f(X_i) - P(f)) - \frac{T_n}{\sqrt{n}}\alpha_n(f) \\ &= \sum_{i=1}^n (Z_i f(X_i) - \mathbb{E}(Zf(X))) - \sum_{i=1}^n (Z_i P(f) - \mathbb{E}(ZP(f))) \\ &\quad - \sqrt{n}\alpha_n(f) \left( 1 + \left( \frac{T_n}{n} - 1 \right) \right). \end{aligned}$$

The last expression is a linear functional of the empirical process  $\alpha_n^{(X,Z)}$  built from the sample  $(X_1, Z_1), \dots, (X_n, Z_n)$  and indexed by  $\mathcal{F}_1$  from (4.4) since

$$\begin{aligned} \tilde{\alpha}_n^*(f) &= \alpha_n^{(X,Z)}(h_f^{(1)}) - P(f) \cdot \alpha_n^{(X,Z)}(h_f^{(2)}) - \alpha_n^{(X,Z)}(h_f^{(3)}) \left( 1 + \left( \frac{T_n}{n} - 1 \right) \right) \\ &= \alpha_n^{(X,Z)}(h_f) - \alpha_n^{(X,Z)}(f) \left( \frac{T_n}{n} - 1 \right), \end{aligned} \quad (4.5)$$

where  $h_f^{(1)}, h_f^{(2)}, h_f^{(3)} \in \mathcal{F}_1$  are defined by (4.3).

## 4.2 Proof of Propositions 1, 4 and 5

Let  $t_n = \sqrt{n/LL(n)}$  and fix  $\varepsilon > 0$ . At Step 1 we prove a law of iterated logarithm for the simpler empirical bootstrapped measure  $\mathbb{P}_n^*$ . Proposition 1 is proved at Step 2 as a direct consequence of Step 1. At Step 3 we use Proposition 1 to prove Proposition 4. Finally, we prove Proposition 5 at Step 4.

**Step 1.** By Lemma 1,  $\mathcal{F}$  and  $\mathcal{F}_1$  satisfy both (VC) or (BR). The law of the iterated logarithm holds – see Alexander [2] – then with probability one there exists  $n_0 = n_0(\omega) > 0$  such that for  $n > n_0$ ,

$$t_n \|\mathbb{P}_n - P\|_{\mathcal{F}} \leq \sqrt{2}\sigma_{\mathcal{F}}(1 + \varepsilon), \quad t_n \|\mathbb{P}_n - P\|_{\mathcal{F}_1} \leq \sqrt{2}\sigma_{\mathcal{F}_1}(1 + \varepsilon). \quad (4.6)$$

According to the law of large numbers with probability one there exists  $n_1 = n_1(\omega) > 0$  such that for  $n > n_1$ ,

$$\left| \frac{T_n}{n} - 1 \right| = |(\mathbb{P}_n - P)(Z)| \leq \varepsilon. \quad (4.7)$$

Notice that

$$t_n \|\tilde{\mathbb{P}}_n^* - P\|_{\mathcal{F}} \leq t_n \|\tilde{\mathbb{P}}_n^* - \bar{Z}_n \mathbb{P}_n\|_{\mathcal{F}} + t_n \|\bar{Z}_n \mathbb{P}_n - P\|_{\mathcal{F}}. \quad (4.8)$$

Decomposition of  $\tilde{\alpha}_n^*$  given by (4.5) and (4.6) imply that for all  $n \geq n_0$

$$\begin{aligned} t_n \|\tilde{\mathbb{P}}_n^* - \bar{Z}_n \mathbb{P}_n\|_{\mathcal{F}} &\leq t_n \|\mathbb{P}_n - P\|_{\mathcal{F}_1} (2 + M_{\mathcal{F}} + \varepsilon) \\ &\leq \sqrt{2}\sigma_{\mathcal{F}_1}(1 + \varepsilon)(2 + M_{\mathcal{F}} + \varepsilon). \end{aligned} \quad (4.9)$$

Using that  $\mathcal{F} \subset \mathcal{F}_1$ , (4.6) and (4.7) we can write for all  $n \geq \max(n_0, n_1)$  by:

$$\begin{aligned} t_n \|\bar{Z}_n \mathbb{P}_n - P\|_{\mathcal{F}} &= t_n \left\| \frac{\sum_{i=1}^n Z_i}{n} \left( \frac{1}{n} \sum_{i=1}^n f(X_i) \right) - \mathbb{E}[f(X)] \right\|_{\mathcal{F}} \\ &\leq t_n \|\mathbb{P}_n(Z)(\mathbb{P}_n(f) - P(f))\|_{\mathcal{F}_1} + M_{\mathcal{F}} t_n \|\mathbb{P}_n(Z) - P(Z)\| \\ &\leq (M_Z + M_{\mathcal{F}})(1 + \varepsilon). \end{aligned} \quad (4.10)$$

We have shown by (4.8), (4.9) and (4.10) that for all  $n > \max(n_0, n_1)$ ,  $t_n \|\tilde{\mathbb{P}}_n^* - P\|_{\mathcal{F}} < K_0(1 + \varepsilon)$  with  $K_0 = K_0(\mathcal{F}, P^{(X,Z)})\sqrt{2}(\sigma_{\mathcal{F}_1}(2 + M_{\mathcal{F}}) + M_Z + M_{\mathcal{F}}) > 0$ .

**Step 2.** By (4.2) we have for all  $n > 0$ ,

$$t_n \|\mathbb{P}_n^* - P\|_{\mathcal{F}} = \frac{1}{1 + (T_n/n - 1)} t_n \|\tilde{\mathbb{P}}_n^* - P\|_{\mathcal{F}}.$$

According to the conclusion of Step 1 and (4.7) for all  $n > \max(n_0, n_1)$  and  $\varepsilon \in (0, 1/2)$  it holds,

$$t_n \|\mathbb{P}_n^* - P\|_{\mathcal{F}} \leq \frac{K_0}{1 - \varepsilon}(1 + \varepsilon) \leq 2K_0(1 + \varepsilon).$$

Proposition 1 is proved with  $K = K(\mathcal{F}, P^{(X,Z)}) = 2K_0(\mathcal{F}, P^{(X,Z)}) > 0$ .

**Step 3.** Let  $b = \sqrt{2}\sigma_{\mathcal{F}}$ . A sufficient condition to prove Proposition 4 is to show that with probability one there exists  $K = K(\mathcal{F}, P^{(X,Z)}) > 0, n_2 = n_2(\omega) > 0$  such that for all  $n > n_2$ ,

$$t_n \|\mathbb{P}_n^{*(N)} - P\|_{\mathcal{F}} \leq (Kb_N - b)(1 + \varepsilon). \quad (4.11)$$

The case  $N = 0$  is proved by Proposition 1. Now let assume that this condition is satisfied for some fixed  $0 \leq N < N_0$ . Then by (4.11) for all  $n \geq n_2$ ,

$$\begin{aligned} t_n \|\mathbb{P}_n^{*(N+1)} - P\|_{\mathcal{F}} &\leq t_n \|\mathbb{P}_n^{*(N+1)} - \mathbb{P}_n^{*(N)}\|_{\mathcal{F}} + t_n \|\mathbb{P}_n^{*(N)} - P\|_{\mathcal{F}} \\ &\leq t_n \|\mathbb{P}_n^{*(N+1)} - \mathbb{P}_n^{*(N)}\|_{\mathcal{F}} + (Kb_N - b)(1 + \varepsilon) \end{aligned} \quad (4.12)$$

Notice that  $m_{N+1} \leq p_{(N_0)}$  and  $\|\mathbb{P}_n^{*(N)}(f1_A)\|_{\mathcal{F}} \leq M_{\mathcal{F}}\mathbb{P}_n^{*(N)}(A)$ . By (4.6) and (4.11) we have for all  $n > \max(n_0, n_2)$ ,

$$\begin{aligned} t_n \|\mathbb{P}_n^{*(N+1)} - \mathbb{P}_n^{*(N)}\|_{\mathcal{F}} &= t_n \left\| \sum_{j \leq m_{N+1}} \frac{\mathbb{P}_n^{*(N)}(f1_{A_j^{(N+1)}})}{\mathbb{P}_n^{*(N)}(A_j^{(N+1)})} (\mathbb{P}_n^{*(N)}(A_j^{(N+1)}) - \mathbb{P}_n(A_j^{(N+1)})) \right\|_{\mathcal{F}} \\ &\leq \frac{M_{\mathcal{F}}}{p_{N+1}} (t_n \|\mathbb{P}_n^{*(N)} - P\|_{\mathcal{F}} + t_n \|\mathbb{P}_n - P\|_{\mathcal{F}}) \\ &\leq \frac{M_{\mathcal{F}}}{p_{N+1}} Kb_N(1 + \varepsilon). \end{aligned} \quad (4.13)$$

Remind that  $b_{N+1} = b_N(1 + M_{\mathcal{F}}/p_{N+1})$ . By (4.12) and (4.13) we have for all  $n > \max(n_0, n_2)$ ,  $t_n \|\mathbb{P}_n^{*(N+1)} - P\|_{\mathcal{F}} \leq (Kb_{N+1} - b)(1 + \varepsilon)$  which proves Proposition 4 by induction.

**Step 4.** Denote  $\beta_n^{(N)}(\mathcal{F})$  the process defined by  $\beta_n^{(N)}(f) = \sqrt{n}(\mathbb{P}_n^{*(N)}(f) - \mathbb{P}_n^{(N)}(f))$ . Notice that  $\alpha_n^{*(N)} = \beta_n^{(N)} + \alpha_n^{(N)} - \alpha_n$ . For all  $\lambda > 0$ ,

$$\mathbb{P} \left( \|\alpha_n^{*(N)}\|_{\mathcal{F}} > 3\lambda \right) \leq \mathbb{P}(\|\alpha_n\|_{\mathcal{F}} > \lambda) + \mathbb{P} \left( \|\alpha_n^{(N)}\|_{\mathcal{F}} > \lambda \right) + \mathbb{P} \left( \|\beta_n^{(N)}\|_{\mathcal{F}} > \lambda \right). \quad (4.14)$$

Remind that  $K_{\mathcal{F}} = \max(1, M_{\mathcal{F}})$ . For  $K \leq p_N$  and  $K' \leq p_N - K$  we have

$$\begin{aligned}
& \mathbb{P} \left( \|\beta_n^{(N)}\|_{\mathcal{F}} > \lambda \right) \\
& \leq \mathbb{P} \left( \sqrt{n} \sum_{j=1}^{m_N} \left| \frac{\mathbb{P}_n(A_j^{(N)})}{\mathbb{P}_n^{*(N-1)}(A_j^{(N)})} \mathbb{P}_n^{*(N-1)}(f \mathbf{1}_{A_j^{(N)}}) - \frac{P(A_j^{(N)})}{\mathbb{P}_n^{(N-1)}(A_j^{(N)})} \mathbb{P}_n^{(N-1)}(f \mathbf{1}_{A_j^{(N)}}) \right| > \lambda \right) \\
& \leq \mathbb{P} \left( \sum_{j=1}^{m_N} \frac{\|\alpha_n\|_{\mathcal{F}} + \|\beta_n^{(N-1)}\|_{\mathcal{F}}}{\mathbb{P}_n^{*(N-1)}(A_j^{(N)}) \mathbb{P}_n^{(N-1)}(A_j^{(N)})} > \frac{\lambda}{2K_{\mathcal{F}}} \right) \\
& \leq \mathbb{P} \left( \|\alpha_n\|_{\mathcal{F}} > \frac{\lambda K^2}{4m_N K_{\mathcal{F}}} \right) + \mathbb{P} \left( \|\beta_n^{(N-1)}\|_{\mathcal{F}} > \frac{\lambda K^2}{4m_N K_{\mathcal{F}}} \right) \\
& \quad + \mathbb{P} \left( \min_{j=1, \dots, m_N} \mathbb{P}_n^{*(N-1)}(A_j^{(N)}) < K \right) + \mathbb{P} \left( \min_{j=1, \dots, m_N} \mathbb{P}_n^{(N-1)}(A_j^{(N)}) < K \right) \\
& \leq \mathbb{P} \left( \|\alpha_n\|_{\mathcal{F}} > \frac{\lambda K^2}{4m_N K_{\mathcal{F}}} \right) + \mathbb{P} \left( \|\beta_n^{(N-1)}\|_{\mathcal{F}} > \frac{\lambda K^2}{4m_N K_{\mathcal{F}}} \right) \\
& \quad + \mathbb{P} \left( \|\alpha_n^{(N-1)}\|_{\mathcal{F}} + \|\beta_n^{(N-1)}\|_{\mathcal{F}} > K' \sqrt{n} \right) + \mathbb{P} \left( \|\alpha_n^{(N-1)}\|_{\mathcal{F}} > K' \sqrt{n} \right) \\
& \leq \mathbb{P} \left( \|\alpha_n\|_{\mathcal{F}} > \frac{\lambda K^2}{4m_N K_{\mathcal{F}}} \right) + 2\mathbb{P} \left( \|\beta_n^{(N-1)}\|_{\mathcal{F}} > \frac{\lambda K^2}{4m_N K_{\mathcal{F}}} \right) + 2\mathbb{P} \left( \|\alpha_n^{(N-1)}\|_{\mathcal{F}} > \frac{\lambda K^2}{4m_N K_{\mathcal{F}}} \right),
\end{aligned}$$

where the last inequality holds provided that  $K' \sqrt{n} \geq \lambda K^2 / 2m_N K_{\mathcal{F}}$ . Define

$$\beta = \frac{1}{1 + \lambda / \sqrt{n}} \in (0, 1), \quad K = \beta p_N, \quad K' = p_N(1 - \beta).$$

Since  $p_N \leq 1/2$  for any  $N \geq 1$  it holds  $K' > 0$  and  $K' \sqrt{n} \geq \lambda K^2 / m_N(1 + K_{\mathcal{F}})$ . By iteration we obtain that

$$\begin{aligned}
\mathbb{P} \left( \|\beta_n^{(N)}\|_{\mathcal{F}} > \lambda \right) & \leq 2^{N-1} \mathbb{P} \left( \|\alpha_n\|_{\mathcal{F}} > \frac{\lambda P_N^2}{M_N (4K_{\mathcal{F}})^N (1 + \lambda / \sqrt{n})^{2N}} \right) \\
& \quad + 2^N \mathbb{P} \left( \|\beta_n^{(0)}\|_{\mathcal{F}} > \frac{\lambda P_N^2}{M_N (4K_{\mathcal{F}})^N (1 + \lambda / \sqrt{n})^{2N}} \right) \\
& \quad + 2^N \mathbb{P} \left( \max_{K=1, \dots, N} \|\alpha_n^{(K)}\|_{\mathcal{F}} > \frac{\lambda P_N^2}{M_N (4K_{\mathcal{F}})^N (1 + \lambda / \sqrt{n})^{2N}} \right).
\end{aligned} \tag{4.15}$$

Notice that  $P_N^2 / M_N (4K_{\mathcal{F}})^N (1 + \lambda / \sqrt{n})^{2N} \leq 1$ . By using (4.14), (4.15) and Proposition 3 of [1] we have, under the event that  $\alpha_n^{(N)}$  are defined for every  $0 \leq N \leq N_0$ , that

$$\begin{aligned}
\mathbb{P} \left( \|\alpha_n^{*(N)}\|_{\mathcal{F}} > \lambda \right) & \leq 4^{N+1} N M_N \mathbb{P} \left( \|\alpha_n\|_{\mathcal{F}} > \frac{\lambda P_N^2}{3M_N (4K_{\mathcal{F}})^N (1 + \lambda / \sqrt{n})^{2N}} \right) \\
& \quad + 2^{N+1} \mathbb{P} \left( \|\beta_n^{(0)}\|_{\mathcal{F}} > \frac{\lambda P_N^2}{3M_N (4K_{\mathcal{F}})^N (1 + \lambda / \sqrt{n})^{2N}} \right).
\end{aligned} \tag{4.16}$$

Since  $\beta_n^{(0)} = \alpha_n^*$  and according to Hoeffding inequality we have

$$\begin{aligned} \mathbb{P}(\|\beta_n^{(0)}\|_{\mathcal{F}} > \lambda) &\leq \mathbb{P}\left(\left|\frac{n}{T_n} - 1\right| > \frac{\lambda}{M_Z M_{\mathcal{F}} \sqrt{n}}\right) = \mathbb{P}\left(|T_n - n| > \frac{\lambda \sqrt{n}}{M_Z M_{\mathcal{F}}}\right) \\ &\leq 2 \exp\left(-\frac{\lambda^2}{2M_Z^4 M_{\mathcal{F}}^2}\right). \end{aligned} \quad (4.17)$$

With (4.16) and (4.17) we have

$$\begin{aligned} \mathbb{P}\left(\|\alpha_n^{*(N)}\|_{\mathcal{F}} > \lambda\right) &\leq 4^{N+1} N M_N \mathbb{P}\left(\|\alpha_n\|_{\mathcal{F}} > \frac{\lambda P_N^2}{3M_N (4K_{\mathcal{F}})^N (1 + \lambda/\sqrt{n})^{2N}}\right) \\ &\quad + 2^{N+2} \exp\left(-\frac{\lambda^2 P_N^4}{18M_N^2 M_Z^4 M_{\mathcal{F}}^2 (4K_{\mathcal{F}})^{2N} (1 + \lambda/\sqrt{n})^{4N}}\right). \end{aligned}$$

The latter bound is increasing with  $N$  and  $\mathbb{P}(\max_{0 \leq N \leq N_0} \|\alpha_n^{*(N)}\|_{\mathcal{F}} > \lambda) \leq \sum_{N=0}^{N_0} \mathbb{P}(\|\alpha_n^{*(N)}\|_{\mathcal{F}} > \lambda)$  which leads to (3.1). Inequality (3.2) is a consequence of Theorem 2.14.2 and 2.14.25 of van der Vaart and Wellner [35] whereas (3.3) is a consequence of Theorem 2.14.9 of [35].

### 4.3 Construction of limit Gaussian processes

**Bootstrapped Gaussian processes.** The Gaussian processes leading the asymptotic behavior of the weighted bootstrap empirical processes are as follows. Let  $\mathbb{G}^0(\mathcal{F}_1)$  be the  $P^{(X,Z)}$ -Brownian bridge indexed by  $\mathcal{F}_1$ . The bootstrapped Gaussian process  $\mathbb{G}^*(\mathcal{F})$  and the standard  $P$ -Brownian bridge  $\mathbb{G}(\mathcal{F})$  – which is actually in this section the margin process of  $\mathbb{G}^0(\mathcal{F}_1)$  – are defined to be

$$\mathbb{G}(f) = \mathbb{G}^0(h_f^{(3)}) = \mathbb{G}^0(f \times \mathbf{1}), \quad (4.18)$$

$$\mathbb{G}^*(f) = \mathbb{G}^0(h_f^{(1)}) - P(f)\mathbb{G}^0(h^{(2)}) - \mathbb{G}^0(h_f^{(3)}), \quad f \in \mathcal{F}. \quad (4.19)$$

Note that by linearity one can define  $\mathbb{G}^0(h_f) = \mathbb{G}^*(f)$ . The following lemma establishes the distribution of  $\mathbb{G}^*(\mathcal{F})$  and  $\mathbb{G}(\mathcal{F})$  and the independence between them.

**Lemma 2.** *If (1.3) and (1.4) hold then  $\mathbb{G}^*(\mathcal{F})$  and  $\mathbb{G}(\mathcal{F})$  are independent  $P$ -Brownian bridges.*

*Proof.* By definition of  $\mathbb{G}^0(\mathcal{F}_1)$ , the two processes  $\mathbb{G}^*(\mathcal{F})$  and  $\mathbb{G}(\mathcal{F})$  are centered linear Gaussian, indexed by  $\mathcal{F}$  such that, for all  $f, g \in \mathcal{F}$ ,

$$\begin{aligned} \text{Cov}(\mathbb{G}(f), \mathbb{G}(g)) &= P^{(X,Z)}(h_f^{(3)} h_g^{(3)}) - P^{(X,Z)}(h_f^{(3)}) P^{(X,Z)}(h_g^{(3)}) \\ &= P(fg) - P(f)P(g). \end{aligned}$$

By (1.4) we have  $h_f(X, Z) = (Z-1)(f(X) - P(f)) - P(f)$  a.s. thus (1.3) implies

$$\begin{aligned} P^{(X,Z)}(h_f) &= \mathbb{E}(Z(f(X) - P(f)) - f(X)) \\ &= \mathbb{E}(\mathbb{E}(Z|X)(f(X) - P(f))) - P(f) = -P(f). \end{aligned}$$



Since  $\mathbb{G}^0$  is also linear and (1.3) implies  $\mathbb{E}((Z-1)^2|X) = 1$  we get

$$\begin{aligned}\text{Cov}(\mathbb{G}^*(f), \mathbb{G}^*(g)) &= P^{(X,Z)}((h_f + P(f))(h_g + P(g))) \\ &= \mathbb{E}((Z-1)^2(f(X) - P(f))(g(X) - P(g))) \\ &= P(fg) - P(f)P(g).\end{aligned}$$

Moreover, for all  $f, g \in \mathcal{F}$  we get, since (1.3) implies  $\mathbb{E}(Z-1|X) = 0$ ,

$$\begin{aligned}\text{Cov}(\mathbb{G}^*(f), \mathbb{G}(g)) &= \text{Cov}(\mathbb{G}^0(h_f), \mathbb{G}^0(h_g^{(3)})) \\ &= P^{(X,Z)}((h_f + P(f))(h_g^{(3)} - P(g))) \\ &= \mathbb{E}((Z-1)(f(X) - P(f))(g(X) - P(g))) = 0,\end{aligned}$$

which proves the independence.  $\square$

**Raked bootstrapped Gaussian process.** Let  $\mathbb{G}^{(N)*}(\mathcal{F})$  be the raked bootstrapped  $P$ -Brownian bridge defined recursively by:

$$\mathbb{G}^{*(0)}(f) = \mathbb{G}^*(f), \quad \mathbb{G}^{*(N+1)}(f) = \mathbb{G}^{*(N)}(f) - \sum_{j=1}^{m_{N+1}} \mathbb{E}[f|A_j^{(N+1)}] \mathbb{G}^{*(N)}(A_j^{(N+1)}), \quad (4.20)$$

where  $\mathbb{G}^*(\mathcal{F})$  is defined by (4.19) through  $\mathbb{G}^0(\mathcal{F}_1)$ . The following lemma establishes the distribution of  $\mathbb{G}^{*(N)}$  and the independence between this process and  $\mathbb{G}^{(N)}(\mathcal{F})$ .

**Lemma 3.** *If (1.3) and (1.4) hold then  $\mathbb{G}^{*(N)}(\mathcal{F})$  and  $\mathbb{G}^{(N)}(\mathcal{F})$  are independent Gaussian processes with same distribution.*

*Proof.* Let  $f, g \in \mathcal{F}$ . By Lemma 2 we have

$$\begin{aligned}\text{Cov}(\mathbb{G}^{*(0)}(f), \mathbb{G}^{(0)}(g)) &= 0, \\ \text{Cov}(\mathbb{G}^{*(0)}(f), \mathbb{G}^{*(0)}(g)) &= \text{Cov}(\mathbb{G}^{(0)}(f), \mathbb{G}^{(0)}(g)).\end{aligned}$$

If we assume that

$$\begin{aligned}\text{Cov}(\mathbb{G}^{*(N)}(f), \mathbb{G}^{(N)}(g)) &= 0, \\ \text{Cov}(\mathbb{G}^{*(N)}(f), \mathbb{G}^{*(N)}(g)) &= \text{Cov}(\mathbb{G}^{(N)}(f), \mathbb{G}^{(N)}(g)),\end{aligned}$$

for some  $N > 0$  then by construction of  $\mathbb{G}^{(N)}(\mathcal{F})$ ,  $\mathbb{G}^{*(N)}(\mathcal{F})$  respectively defined by (1.7) and (4.20) we have necessary

$$\begin{aligned}\text{Cov}(\mathbb{G}^{*(N+1)}(f), \mathbb{G}^{(N+1)}(g)) &= 0, \\ \text{Cov}(\mathbb{G}^{*(N+1)}(f), \mathbb{G}^{*(N+1)}(g)) &= \text{Cov}(\mathbb{G}^{(N+1)}(f), \mathbb{G}^{(N+1)}(g)),\end{aligned}$$

which prove the lemma by induction the same distribution.  $\square$

**Bootstrapped raked Gaussian process.** Let  $\mathbb{G}^{0,(N)}(\mathcal{F}_1)$  be the  $P^{(X,Z)}$ -raked Brownian bridge defined recursively as in (1.7) from  $\mathbb{G}^{0,(0)} = \mathbb{G}^0$  and the following auxiliary information, for all  $N > 0$  and  $j = 1, \dots, m_N$ ,

$$\mathcal{A}^{0,(N)} = \{A_1^{0,(N)}, \dots, A_{m_N}^{0,(N)}\}, \quad A_j^{0,(N)} = A_j^{(N)} \times [0, M_Z]. \quad (4.21)$$

Notice that  $P^{(X,Z)}(\mathcal{A}^{0,(N)}) = P(\mathcal{A}^{(N)})$ . The bootstrapped raking-ratio Gaussian process  $\mathbb{G}^{(N)*}(\mathcal{F})$  is defined through

$$\mathbb{G}^{(N)*}(f) = \mathbb{G}^{0,(N)}(h_f^{(1)}) - P(f)\mathbb{G}^{0,(N)}(h^{(2)}) - \mathbb{G}^{0,(N)}(h_f^{(3)}), \quad f \in \mathcal{F}. \quad (4.22)$$

By linearity, one can define  $\mathbb{G}^{0,(N)}(h_f) = \mathbb{G}^{(N)*}(f)$ . The following lemma gives the distribution of  $\mathbb{G}^{(N)*}(\mathcal{F})$ .

**Lemma 4.** *If (1.3) and (1.4) hold then for all  $\mathbb{G}^{(N)*}(\mathcal{F})$  and  $\mathbb{G}(\mathcal{F})$  are independent  $P$ -Brownian bridges.*

*Proof.* By (1.8), Lemma 2 and definition of  $\mathbb{G}^{(N)*}(\mathcal{F})$  given by (4.22) and we have for all  $f, g \in \mathcal{F}$ ,

$$\begin{aligned} \text{Cov}(\mathbb{G}^{(N)*}(f), \mathbb{G}^{(N)*}(g)) &= \text{Cov}(\mathbb{G}^{0,(N)}(h_f), \mathbb{G}^{0,(N)}(h_g)) \\ &= \text{Cov}(\mathbb{G}^0(h_f), \mathbb{G}^0(h_g)) - R_N(P^{(X,Z)}, h_f, h_g) \\ &= \text{Cov}(\mathbb{G}(f), \mathbb{G}(g)) - R_N(P^{(X,Z)}, h_f, h_g). \end{aligned}$$

Let  $h_1, h_2 \in \mathcal{F}_1$  defined by  $h_i(x, z) = f_i(x)g_i(z)$  for  $i = 1, 2$ . By the definition of  $R_N$  given by (1.9) we have  $\Phi_k^{(N)}(P^{(X,Z)}, h) = \Phi_k^{(N)}(P, f_1)$  and  $\text{Var}(\mathbb{G}[\mathcal{A}^{(N)}]) = \text{Var}(\mathbb{G}^0[\mathcal{A}^{0,(N)}])$  since for every  $1 \leq j_1, j_2 \leq m_N$ ,

$$\begin{aligned} \text{Cov}(\mathbb{G}^0(A_{j_1}^{0,(N)}), \mathbb{G}^0(A_{j_2}^{0,(N)})) &= P^{(X,Z)}(A_{j_1}^{0,(N)} \cap A_{j_2}^{0,(N)}) - P^{(X,Z)}(A_{j_1}^{0,(N)})P^{(X,Z)}(A_{j_2}^{0,(N)}) \\ &= P(A_{j_1}^{(N)} \cap A_{j_2}^{(N)}) - P(A_{j_1}^{(N)})P(A_{j_2}^{(N)}) \\ &= \text{Cov}(\mathbb{G}(A_{j_1}^{(N)}), \mathbb{G}(A_{j_2}^{(N)})). \end{aligned}$$

These remarks lead to  $R_N(P^{(X,Z)}, h_1, h_2) = R_N(P, f_1, f_2)$ . Then

$$\begin{aligned} R_N(P^{(X,Z)}, h_f, h_g) &= R_N(P^{(X,Z)}, h_f^{(1)} - P(f)h^{(2)} - h_f^{(3)}, h_g^{(1)} - P(g)h^{(2)} - h_g^{(3)}) \\ &= -P(f)P(g)R_N(P, \mathbf{1}, \mathbf{1}). \end{aligned}$$

By using (1.8) and the fact that  $\mathbb{G}(\mathbf{1}) = 0, \mathbb{G}^{(N)}(\mathbf{1}) = 0$  we have

$$R_N(\mathbf{1}, \mathbf{1}) = \text{Cov}(\mathbb{G}(\mathbf{1}), \mathbb{G}(\mathbf{1})) - \text{Cov}(\mathbb{G}^{(N)}(\mathbf{1}), \mathbb{G}^{(N)}(\mathbf{1})) = 0.$$

We have shown that  $\text{Cov}(\mathbb{G}^{(N)*}(f), \mathbb{G}^{(N)*}(g)) = \text{Cov}(\mathbb{G}(f), \mathbb{G}(g))$ . By Lemma 2  $\mathbb{G}^{0,(0)} = \mathbb{G}^*$  is independent of  $\mathbb{G}$  and if for all  $f, g \in \mathcal{F}$ ,

$$\text{Cov}(\mathbb{G}^{(N)*}(f), \mathbb{G}(g)) = \text{Cov}(\mathbb{G}^{0,(N)}(h_f), \mathbb{G}(g)) = 0,$$

then

$$\text{Cov}(\mathbb{G}^{(N+1)*}(f), \mathbb{G}(g)) = \text{Cov}(\mathbb{G}^{0,(N+1)}(h_f), \mathbb{G}(g)) = 0,$$

since  $\mathbb{G}^{0,(N+1)}(h_f)$  is a linear combination of elements of  $\mathbb{G}^{0,(N)}(\mathcal{F}_1)$ . By induction we have proved the independence between  $\mathbb{G}^{(N)*}(\mathcal{F})$  and  $\mathbb{G}(\mathcal{F})$ .  $\square$

#### 4.4 Proof of Theorem 2.1 and 2.2

At Step 1 we prove Proposition 7 below. At Step 2 we deduce Theorem 2.1 from Proposition 7 and derive at Step 3 Theorem 2.2 from Theorem 2.1. Fix  $\theta > 0$ .

**Proposition 7.** *Assume (1.3), (1.4) and either (VC) or (BR). There exists constants  $C_\theta, n_\theta > 0$  and sequences  $\{(X_n, Z_n)\}$  of i.i.d. random variables with distribution  $P^{(X,Z)}$  and a sequence  $\{(\mathbb{G}_n(\mathcal{F}), \mathbb{G}_n^*(\mathcal{F}))\}$  of pairs of independent  $P$ -Brownian bridges, all of these sequences being defined on the same probability space, such that for all  $n > n_\theta$ ,*

$$\mathbb{P}(\{\|\alpha_n - \mathbb{G}_n\|_{\mathcal{F}} > C_\theta v_n\} \cup \{\|\tilde{\alpha}_n^* - \mathbb{G}_n^*\|_{\mathcal{F}} > C_\theta v_n\}) \leq \frac{1}{n^\theta}. \quad (4.23)$$

On the above probability space there also exists independent sequences  $\{\mathbb{G}'_n(\mathcal{F})\}$  and  $\{\mathbb{G}_n^{*\prime}(\mathcal{F})\}$  of independent versions of  $\mathbb{G}(\mathcal{F})$  such that, for all  $n > n_\theta$ ,

$$\mathbb{P}\left(\frac{1}{\sqrt{n}} \max_{1 \leq k \leq n} \|\sqrt{k}\alpha_k - \sum_{i=1}^k \mathbb{G}'_i\|_{\mathcal{F}} > C_\theta v_n\right) \leq \frac{1}{u_n^\theta}, \quad (4.24)$$

$$\mathbb{P}\left(\frac{1}{\sqrt{n}} \max_{1 \leq k \leq n} \|\sqrt{k}\tilde{\alpha}_k^* - \sum_{i=1}^k \mathbb{G}_i^{*\prime}\|_{\mathcal{F}} > C_\theta v_n\right) \leq \frac{1}{u_n^\theta}. \quad (4.25)$$

**Step 1.** By Lemma 1 one can apply Propositions 1,2 and Theorems 1,2 of Berthet and Mason [8] to  $\alpha_n^{(X,Z)}(\mathcal{F}_1)$ . There exists  $C_0 = C_0(\theta) > 0, n_0 = n_0(\theta) > 0$  and a probability space where one can construct a sequences of independent random variables  $\{(X_n, Z_n)\}$  distributed as  $P^{(X,Z)}$  and two sequences  $\{\mathbb{G}_n^0(\mathcal{F}_1)\}, \{\mathbb{G}_n^{0'}(\mathcal{F})\}$  of independent  $P^{(X,Z)}$ -Brownian bridge satisfying for all  $n > n_0$ ,

$$\mathbb{P}\left(\|\alpha_n^{(X,Z)} - \mathbb{G}_n^0\|_{\mathcal{F}_1} > C_0 v_n\right) \leq \frac{1}{n^\theta} \quad (4.26)$$

$$\mathbb{P}\left(\frac{1}{\sqrt{n}} \max_{1 \leq k \leq n} \|\sqrt{k}\alpha_k^{(X,Z)} - \sum_{i=1}^k \mathbb{G}_i^{0'}\|_{\mathcal{F}_1} > C_0 v_n\right) \leq \frac{1}{6u_n^\theta}. \quad (4.27)$$

We only prove (4.24) and (4.25) by using (4.27). We can adapt the following proof to show (4.23) by using (4.26). Let denote  $\{\mathbb{G}_n(\mathcal{F})\}$  and  $\{\mathbb{G}_n^*(\mathcal{F})\}$  the processes defined respectively by (4.18) and (4.19) both built through  $\{\mathbb{G}_n^0(\mathcal{F}_1)\}$ . According to Lemma 2,  $\{\mathbb{G}_n(\mathcal{F})\}, \{\mathbb{G}_n^*(\mathcal{F})\}$  are independent  $P$ -Brownian bridges. Since  $\alpha_k^{(X,Z)}(h_f^{(3)}) = \alpha_k(f)$ , inequality (4.24) is satisfied in particular. Write

$C'_0 = C_0 \max(2, M_{\mathcal{F}})$ . By (4.27) one can write that for all  $n > n_0$ ,

$$\begin{aligned}
& \mathbb{P} \left( \frac{1}{\sqrt{n}} \max_{1 \leq k \leq n} \|\sqrt{k} \tilde{\alpha}_k^* - \sum_{i=1}^k \mathbb{G}_i^*\|_{\mathcal{F}} > C'_0 v_n \right) \\
& \leq \mathbb{P} \left( \frac{1}{\sqrt{n}} \sup_{f \in \mathcal{F}} \max_{1 \leq k \leq n} \left| \sqrt{k} \alpha_k^{(X,Z)}(h_f^{(1)}) - \sum_{i=1}^k \mathbb{G}_i^{0'}(h_f^{(1)}) \right| > C'_0 v_n \right) \\
& \quad + \mathbb{P} \left( \frac{1}{\sqrt{n}} \sup_{f \in \mathcal{F}} \max_{1 \leq k \leq n} \left| \mathbb{E}[f] \left( \sqrt{k} \alpha_k^{(X,Z)}(h_f^{(2)}) - \sum_{i=1}^k \mathbb{G}_i^{0'}(h_f^{(2)}) \right) \right| > C'_0 v_n \right) \\
& \quad + \mathbb{P} \left( \frac{1}{\sqrt{n}} \sup_{f \in \mathcal{F}} \max_{1 \leq k \leq n} \left| \sqrt{k} \alpha_k^{(X,Z)}(h_f^{(3)}) - \sum_{i=1}^k \mathbb{G}_i^{0'}(h_f^{(3)}) \right| > C'_0 v_n \right) \\
& \quad + \mathbb{P} \left( \frac{1}{\sqrt{n}} \sup_{f \in \mathcal{F}} \max_{1 \leq k \leq n} \left| \sqrt{k} \alpha_k^{(X,Z)}(h_f^{(3)}) \cdot \frac{1}{k} \sum_{i=1}^k (Z_i - 1) \right| > C'_0 v_n \right) \\
& \leq \frac{1}{2u_n^\theta} + \mathbb{P} \left( \frac{1}{\sqrt{n}} \max_{1 \leq k \leq n} \left[ \|\sqrt{k} \alpha_k^{(X,Z)}\|_{\mathcal{F}_1} \cdot \left| \frac{1}{k} \sum_{i=1}^k (Z_i - 1) \right| \right] > C'_0 v_n \right). \quad (4.28)
\end{aligned}$$

The last right-hand side member of (4.28) can be bounded by

$$\begin{aligned}
& \mathbb{P} \left( \frac{1}{\sqrt{n}} \max_{1 \leq k \leq n} \left[ \|\sqrt{k} \alpha_k^{(X,Z)}\|_{\mathcal{F}_1} \cdot \left| \frac{1}{k} \sum_{i=1}^k (Z_i - 1) \right| \right] > C'_0 v_n \right) \quad (4.29) \\
& \leq \mathbb{P} \left( \max_{1 \leq k \leq n} \|\alpha_k^{(X,Z)}\|_{\mathcal{F}_1} > 2C_0 v_n n^{1/2-\varepsilon} \right) + \mathbb{P} \left( \max_{1 \leq k \leq n} \left| \frac{1}{\sqrt{k}} \sum_{i=1}^k (Z_i - 1) \right| > n^\varepsilon \right).
\end{aligned}$$

By Hoeffding inequality and definition of  $\varepsilon$  there exists  $n_1 = n_1(\theta) > 0$  such that for all  $n > n_1$ ,

$$\mathbb{P} \left( \max_{1 \leq k \leq n} \left| \frac{1}{k} \sum_{i=1}^k (Z_i - 1) \right| > n^\varepsilon \right) \leq 2n \exp \left( -\frac{2n^{2\varepsilon}}{M_Z^2} \right) \leq \frac{1}{6u_n^\theta}. \quad (4.30)$$

Moreover by (4.27) we for all  $n > n_0$ ,

$$\begin{aligned}
& \mathbb{P} \left( \max_{1 \leq k \leq n} \|\alpha_k^{(X,Z)}\|_{\mathcal{F}_1} > 2C_0 v_n n^{1/2-\varepsilon} \right) \\
& \leq \mathbb{P} \left( \max_{1 \leq k \leq n} \left\| \alpha_k^{(X,Z)} - \frac{1}{\sqrt{k}} \sum_{i=1}^k \mathbb{G}_i^{0'} \right\|_{\mathcal{F}_1} > C_0 v_n n^{1/2-\varepsilon} \right) \\
& \quad + \mathbb{P} \left( \max_{1 \leq k \leq n} \left\| \frac{1}{\sqrt{k}} \sum_{i=1}^k \mathbb{G}_i^{0'} \right\|_{\mathcal{F}_1} > C_0 v_n n^{1/2-\varepsilon} \right) \\
& \leq \frac{1}{6u_n^\theta} + \mathbb{P} \left( \max_{1 \leq k \leq n} \left\| \frac{1}{\sqrt{k}} \sum_{i=1}^k \mathbb{G}_i^{0'} \right\|_{\mathcal{F}_1} > C_0 v_n n^{1/2-\varepsilon} \right), \quad (4.31)
\end{aligned}$$

By Borell-Sudakov inequality – see annex A.2.1 of [35] – there exists  $n_2 = n_2(\theta) > 0$  such that for all  $n > n_2$ ,

$$\begin{aligned} \mathbb{P} \left( \max_{1 \leq k \leq n} \left\| \frac{1}{\sqrt{k}} \sum_{i=1}^k \mathbb{G}_i^{0'} \right\|_{\mathcal{F}_1} > C_0 v_n n^{1/2-\varepsilon} \right) &\leq \sum_{k=1}^n \mathbb{P} \left( \left\| \frac{1}{\sqrt{k}} \sum_{i=1}^k \mathbb{G}_i^{0'} \right\|_{\mathcal{F}_1} > C_0 v_n n^{1/2-\varepsilon} \right) \\ &\leq 2n \exp \left( -\frac{(C_0 v_n n^{1/2-\varepsilon})^2}{8\mathbb{E}[\|\mathbb{G}_1^{0'}\|_{\mathcal{F}_1}]^2} \right) \leq \frac{1}{6u_n^\theta}. \end{aligned} \quad (4.32)$$

So (4.29), (4.30), (4.31) and (4.32) imply that for all  $n > \max(n_0, n_1)$ ,

$$\mathbb{P} \left( \frac{1}{\sqrt{n}} \max_{1 \leq k \leq n} \left[ \|\sqrt{k} \alpha_k^{(X,Z)}\|_{\mathcal{F}_1} \cdot \left| \frac{1}{k} \sum_{i=1}^k (Z_i - 1) \right| \right] > C'_\theta v_n \right) \leq \frac{1}{2u_n^\theta}. \quad (4.33)$$

Inequality (4.25) is proved by (4.28) and (4.33) with  $n_\theta = \max(n_0, n_1, n_2)$  and  $C_\theta = C'_0$ .

**Step 2.** Now we prove (2.2) and (2.3) of Theorem 2.1. By adapting the following proof we can also show (2.1). According to Proposition 7, there exists a constant  $C_0 = C_0(\theta)$ ,  $n_0 = n_0(\theta) > 0$ , a sequence  $\{(X_n, Z_n)\}$  of i.i.d. random variables with distribution  $P^{(X,Z)}$ , a sequence  $\{\mathbb{G}'_n(\mathcal{F})\}, \{\mathbb{G}^{*'}_n(\mathcal{F})\}$  of independent  $P$ -brownian bridge satisfying, for all  $n > n_0$ , (2.2) and

$$\mathbb{P} \left( \frac{1}{\sqrt{n}} \max_{1 \leq k \leq n} \|\sqrt{k} \tilde{\alpha}_k^* - \sum_{i=1}^k \mathbb{G}'_i\|_{\mathcal{F}} > C_0 v_n \right) \leq \frac{1}{4n^\theta}. \quad (4.34)$$

By definition of  $\tilde{\alpha}_n^*$  given by (4.2),

$$\begin{aligned} &\mathbb{P} \left( \frac{1}{\sqrt{n}} \max_{k_{n,\theta} \leq k \leq n} \|\sqrt{k} \alpha_k^* - \sum_{i=1}^k \mathbb{G}'_i\|_{\mathcal{F}} > 3C_0 v_n \right) \\ &= \mathbb{P} \left( \frac{1}{\sqrt{n}} \max_{k_{n,\theta} \leq k \leq n} \left\| \frac{k}{T_k} \sqrt{k} \tilde{\alpha}_k^* - \sum_{i=1}^k \mathbb{G}'_i\|_{\mathcal{F}} > 3C_0 v_n \right\| \right) \\ &\leq \mathbb{P} \left( \frac{1}{\sqrt{n}} \max_{k_{n,\theta} \leq k \leq n} \left\| \frac{k}{T_k} (\sqrt{k} \tilde{\alpha}_k^* - \sum_{i=1}^k \mathbb{G}'_i)\right\|_{\mathcal{F}} > 2C_0 v_n \right) \\ &\quad + \mathbb{P} \left( \frac{1}{\sqrt{n}} \max_{k_{n,\theta} \leq k \leq n} \left\| \left( \frac{k}{T_k} - 1 \right) \sum_{i=1}^k \mathbb{G}'_i\right\|_{\mathcal{F}} > C_0 v_n \right). \end{aligned} \quad (4.35)$$

According to Hoeffding inequality there exists  $n_1 = n_1(\theta) > 0$  such that for all  $n > n_1$ ,

$$\begin{aligned} \mathbb{P} \left( \max_{k_{n,\theta} \leq k \leq n} \frac{k}{T_k} \geq 2 \right) &\leq \sum_{k=k_{n,\theta}}^n \mathbb{P} \left( \frac{T_k}{k} \leq \frac{1}{2} \right) \leq \sum_{k=k_{n,\theta}}^n \mathbb{P} (|T_k - k| \geq k/2) \\ &\leq 2 \sum_{k=k_{n,\theta}}^n \exp \left( -\frac{2(k^2/4)}{kM_Z^2} \right) \leq \frac{1}{4n^\theta}. \end{aligned} \quad (4.36)$$

Using (4.34) and (4.36), we found that for all  $n > \max(n_0, n_1)$ ,

$$\mathbb{P} \left( \frac{1}{\sqrt{n}} \max_{k_{n,\theta} \leq k \leq n} \left\| \frac{k}{T_k} (\sqrt{k} \tilde{\alpha}_k^* - \sum_{i=1}^k \mathbb{G}_i^{*'}) \right\|_{\mathcal{F}} > 2C_0 v_n \right) \leq \frac{1}{2n^\theta}. \quad (4.37)$$

By definition of  $\varepsilon$  one can write that

$$\begin{aligned} & \mathbb{P} \left( \frac{1}{\sqrt{n}} \max_{k_{n,\theta} \leq k \leq n} \left\| \left( \frac{k}{T_k} - 1 \right) \sum_{i=1}^k \mathbb{G}_i^{*'} \right\|_{\mathcal{F}} > C_0 v_n \right) \\ & \leq \mathbb{P} \left( n^{-1/2+\varepsilon} \max_{k_{n,\theta} \leq k \leq n} \left\| \frac{1}{\sqrt{k}} \sum_{i=1}^k \mathbb{G}_i^{*'} \right\|_{\mathcal{F}} > C_0 v_n \right) \\ & \quad + \mathbb{P} \left( \max_{k_{n,\theta} \leq k \leq n} \left| \sqrt{k} \left( \frac{k}{T_k} - 1 \right) \right| > n^\varepsilon \right). \end{aligned} \quad (4.38)$$

With the same calculations as (4.32) – but on the class  $\mathcal{F}$  – we show that there exists  $n_2 = n_2(\theta) > 0$  such that for all  $n > n_2$  :

$$\mathbb{P} \left( n^{-1/2+\varepsilon} \max_{k_{n,\theta} \leq k \leq n} \left\| \frac{1}{\sqrt{k}} \sum_{i=1}^k \mathbb{G}_i^{*'} \right\|_{\mathcal{F}} > C_0 v_n \right) \leq \frac{1}{4n^\theta}, \quad (4.39)$$

By (4.36) and Hoeffding inequality, there exists  $n_3 = n_3(\theta) > 0$  such that for all  $n > n_3$ ,

$$\begin{aligned} & \mathbb{P} \left( \max_{k_{n,\theta} \leq k \leq n} \left| \sqrt{k} \left( \frac{k}{T_k} - 1 \right) \right| > n^\varepsilon \right) = \mathbb{P} \left( \max_{k_{n,\theta} \leq k \leq n} \left| \frac{k}{T_k} \cdot \frac{T_k - k}{\sqrt{k}} \right| > n^\varepsilon \right) \\ & \leq \mathbb{P} \left( \max_{1 \leq k \leq n} \left| \frac{T_k - k}{\sqrt{k}} \right| > \frac{n^\varepsilon}{2} \right) + \mathbb{P} \left( \max_{k_{n,\theta} \leq k \leq n} \frac{k}{T_k} > 2 \right) \leq \frac{1}{4n^\theta}. \end{aligned} \quad (4.40)$$

By (4.35), (4.37), (4.38), (4.39) and (4.40), we have proved (2.3) with  $n_\theta = \max(n_0, n_1, n_2, n_3)$  and  $C_\theta = 3C_0$ .

**Step 3.** Let  $\mathcal{F}_{b_n}$  be an extension of  $\mathcal{F}_1$  defined by

$$\mathcal{F}_{b_n} = \{h(x, z_1, \dots, z_{b_n}) = f(x)g(z_i) : 1 \leq i \leq b_n, f \in \mathcal{F} \cup \{\mathbf{1}\}, g \in \{\phi_{M_Z}, \mathbf{1}\}\}$$

The fact that  $\mathcal{F}_{b_n}$  satisfies the same metric entropy conditions than  $\mathcal{F}$  is easy to check. Let  $\mathbb{G}^{0,b_n}(\mathcal{F}_{b_n})$  be the  $P(X, Z^{(1)}, \dots, Z^{(b_n)})$ -Brownian bridge. For  $i \leq b_n$ , we define the bootstrapped  $i$ -marginal process  $\mathbb{G}^{*,i}(\mathcal{F})$  by

$$\mathbb{G}^{*(i)}(f) = \mathbb{G}^{0,b_n}(h_{f,b_n}^{(1,i)}) - P(f)\mathbb{G}^{0,b_n}(h_{b_n}^{(2,i)}) - \mathbb{G}^{0,b_n}(h_{f,b_n}^{(3)}), \quad f \in \mathcal{F}_{b_n},$$

where

$$\begin{aligned} h_{f,b_n}^{(1,i)}(x, z_1, \dots, z_{b_n}) &= f(x)\phi_{M_Z}(z_i), \quad h_{b_n}^{(2,i)}(x, z_1, \dots, z_{b_n}) = \phi_{M_Z}(z_i), \\ h_{f,b_n}^{(3)}(x, z_1, \dots, z_{b_n}) &= f(x). \end{aligned}$$

In fact  $\mathbb{G}^{*,1}, \dots, \mathbb{G}^{*,b_n}$  are independent  $P$ -Brownian bridges. Theorem 2.1 implies that there exists positive constants  $C = C(b_n), n_0$ , a sequence  $\{(X_i, Z_i^{(1)}, \dots, Z_i^{(b_n)})\}$  of i.i.d. random variables with law  $P^{(X, Z^{(1)}, \dots, Z^{(b_n)})}$  and a sequence  $\{(\mathbb{G}_n(\mathcal{F}), \mathbb{G}_n^{*,1}(\mathcal{F}), \dots, \mathbb{G}_n^{*,b_n}(\mathcal{F}))\}$  of independent  $P$ -Brownian bridge such that for  $n \geq n_0$ ,

$$\mathbb{P}(B_n) = \mathbb{P} \left( \{ \|\alpha_n - \mathbb{G}_n\|_{\mathcal{F}} > Cv_n \} \cup \bigcup_{j=1}^{b_n} \{ \|\alpha_n^{*(j)} - \mathbb{G}_n^{*,j}\|_{\mathcal{F}} > Cv_n \} \right) \leq \frac{1}{n^2}, \quad (4.41)$$

where  $\alpha_n^{*(i)}(\mathcal{F})$  is defined by  $\alpha_n^{*(i)}(f) = \alpha_n^*(h_{f, b_n}^{(1,i)})$  is the bootstrapped empirical process with weights constructed from  $(Z_1^{(i)}, \dots, Z_n^{(i)})$ .

#### 4.5 Proof of Theorem 3.1 and 3.2

Process  $\alpha_n^{*(N)}(\mathcal{F})$  could be seen as the process  $\alpha_n^{*(N-1)}(\mathcal{F})$  corrected with partition  $\mathcal{A}^{(N)}$  since

$$\begin{aligned} \alpha_n^{*(N)}(f) &= \sqrt{n}(\mathbb{P}_n^{*(N)}(f) - \mathbb{P}_n(f)) \\ &= \sqrt{n} \left( \sum_{j=1}^{m_N} \frac{\mathbb{P}_n(A_j^{(N)})}{\mathbb{P}_n^{*(N-1)}(A_j^{(N)})} \mathbb{P}_n^{*(N-1)}(f \mathbf{1}_{A_j^{(N)}}) - \mathbb{P}_n(f \mathbf{1}_{A_j^{(N)}}) \right) \\ &= \sum_{j=1}^{m_N} \frac{\mathbb{P}_n(A_j^{(N)})}{\mathbb{P}_n^{*(N-1)}(A_j^{(N)})} (\alpha_n^{*(N-1)}(f \mathbf{1}_{A_j^{(N)}}) - \mathbb{E}_n[f | A_j^{(N)}] \alpha_n^{*(N-1)}(A_j^{(N)})), \end{aligned}$$

where  $\mathbb{E}_n[f | A] = \frac{\mathbb{P}(f \mathbf{1}_A)}{\mathbb{P}_n(A)}$ . We denote the following quantities:

$$p_{n,j}^{(N)} = \frac{\mathbb{P}_n(A_j^{(N)})}{\mathbb{P}_n^{*(N-1)}(A_j^{(N)})} - 1, \quad q_{n,j}^{(N)}(f) = \mathbb{E}_n[f | A_j^{(N)}] - \mathbb{E}[f | A_j^{(N)}].$$

Then,

$$\begin{aligned} \alpha_n^{*(N)}(f) &= \alpha_n^{*(N-1)} \circ \phi_{(N)}(f) + \Gamma_n^{(N)}(f), \\ \Gamma_n^{(N)}(f) &= \sum_{j=1}^{m_N} p_{n,j}^{(N)} \alpha_n^{*(N-1)} \circ \phi_{(N),j}(f) - (1 + p_{n,j}^{(N)}) q_{n,j}^{(N)}(f) \alpha_n^{*(N-1)}(A_j^{(N)}). \end{aligned}$$

By successive iterations we get

$$\begin{aligned} \alpha_n^{*(N)}(f) &= \alpha_n^{*(0)} \circ \phi_{(1)} \circ \dots \circ \phi_{(N)}(f) + F_n^{(N)}(f), \quad (4.42) \\ F_n^{(N)}(f) &= \sum_{k=1}^N \Gamma_n^{(k)}(\phi_{(k+1)} \circ \dots \circ \phi_{(N)}(f)). \end{aligned}$$

Let introduce the same notations than [1]. For all  $f \in \mathcal{F}$ ,  $N \leq N_0$  for some fixed  $N_0$  and  $j = 1, \dots, m_N$  we denote

$$\phi_{(N,j)}(f) = (f - \mathbb{E}[f|A_j^{(N)}])1_{A_j^{(N)}}, \quad \phi_{(N)}(f) = \sum_{j=1}^{m_N} \phi_{(N,j)}(f),$$

and function classes

$$\begin{aligned} \mathcal{F}_{(N)} &= \phi_{(1)} \circ \dots \circ \phi_{(N)}(\mathcal{F}), \\ \mathcal{H}_{(N)} &= \bigcup_{1 \leq k \leq N} \bigcup_{1 \leq j \leq m_k} \phi_{(j,k)} \circ \phi_{(k+1)} \circ \dots \circ \phi_{(N)}(\mathcal{F}), \\ F_0 &= \bigcup_{0 \leq N \leq N_0} F_{(N)}, \quad \mathcal{H}_0 = \bigcup_{0 \leq N \leq N_0} \mathcal{H}_{(N)}. \end{aligned}$$

The following lemma establishes that the bootstrapped process  $\alpha_n^*(\mathcal{F}_{(N)})$  is the main process contributing for  $\alpha_n^{*(N)}(\mathcal{F})$  and  $F_n^{(N)}$  is an error process. It is the equivalent of Lemma 2 of [1].

**Lemma 5.** *Assume (1.3), (1.4) and either (VC) or (BR). For any  $\xi, \theta > 0$  there exists  $n_{\xi, \theta} > 0$  such that, for all  $n > n_{\xi, \theta}$ ,*

$$\mathbb{P} \left( \max_{0 \leq N \leq N_0} \|F_n^{(N)}\|_{\mathcal{F}} > \xi v_n \right) \leq \frac{1}{n^\theta}.$$

Fix  $\theta, \xi > 0$ . Now we prove Lemma 5 at Step 1 and apply it at Step 2 to prove Theorem 3.1. In Step 3, we describe how to prove Theorem 3.2.

**Step 1.** Let bound the terms of error:

$$\begin{aligned} |p_{n,j}^{(N)}| &= \left| \frac{\mathbb{P}_n(A_j^{(N)}) - \mathbb{P}_n^{*(N-1)}(A_j^{(N)})}{\mathbb{P}_n^{*(N-1)}(A_j^{(N)})} \right| \leq \frac{\|\alpha_n\|_{\mathcal{F}} + \|\alpha_n^{*(N-1)}\|_{\mathcal{F}}}{\sqrt{n}p_N - \|\alpha_n^{*(N-1)}\|_{\mathcal{F}}}, \\ \|q_{n,j}^{(N)}\|_{\mathcal{F}} &= \left\| \frac{\mathbb{P}_n(f1_{A_j^{(N)}}) - \mathbb{E}[f|A_j^{(N)}]\mathbb{P}_n(A_j^{(N)})}{\mathbb{P}_n(A_j^{(N)})} \right\|_{\mathcal{F}} \leq \frac{2M_{\mathcal{F}}\|\alpha_n\|_{\mathcal{F}}}{\sqrt{n}p_N - \|\alpha_n\|_{\mathcal{F}}}, \\ \|F_n^{(N)}\|_{\mathcal{F}} &\leq \sum_{k=1}^N \max_{1 \leq j \leq m_k} |p_{n,j}^{(k)}| \sum_{j=1}^{m_k} |\alpha_n^{*(k-1)} \circ \phi_{k,j} \circ \phi_{(k+1)} \circ \dots \circ \phi_{(N)} f| \\ &\quad + \max_{1 \leq j \leq m_k} \left( |1 + p_{n,j}^{(k)}| |q_{n,j}^{(k)}| |\alpha_n^{*(k-1)}(A_j)| \right) \\ &\leq S_N \max_{0 \leq k \leq N-1} \|\alpha_n^{*(k)}\|_{\mathcal{H}_0} \left( P_n^{(N)} + (1 + P_n^{(N)})Q_n^{(N)} \right), \\ \max_{1 \leq N \leq N_0} \|F_n^{(N)}\|_{\mathcal{F}} &\leq S_{N_0} \max_{0 \leq N \leq N_0-1} \|\alpha_n^{*(N)}\|_{\mathcal{H}_0} \left( P_n^{(N_0)} + Q_n^{(N_0)} + P_n^{(N_0)}Q_n^{(N_0)} \right), \end{aligned}$$

where  $P_n^{(N)} = \max_{0 \leq k \leq N} \max_{1 \leq j \leq m_k} |p_{n,j}^{(k)}|$ ,  $Q_n^{(N)} = \max_{0 \leq k \leq N} \max_{1 \leq j \leq m_k} |q_{n,j}^{(k)}|$ . For  $\varepsilon \in (0, 1/2)$  there exists  $n_\varepsilon > 0$  such that for all  $n > n_\varepsilon$ ,  $\xi v_n > 4S_{N_0}(1 +$



$2M_{\mathcal{F}}n^\varepsilon/(\sqrt{np_{N_0}})$ . Then, for every  $n > n_\xi$  we have

$$\begin{aligned} \mathbb{P}\left(\max_{0 \leq N \leq N_0} \|F_n^{(N)}\|_{\mathcal{F}} > \xi v_n\right) &\leq \mathbb{P}\left(\max_{0 \leq N \leq N_0} \|\alpha_n^{*(N)}\|_{\mathcal{H}_0} > n^{\varepsilon/2}\right) \\ &+ 2\mathbb{P}\left(P_n^{(N_0)} > \frac{4n^{\varepsilon/4}}{\sqrt{np_{N_0}}}\right) + 2\mathbb{P}\left(Q_n^{(N_0)} > \frac{4M_{\mathcal{F}}n^{\varepsilon/4}}{\sqrt{np_{N_0}}}\right). \end{aligned} \quad (4.43)$$

According to Lemma 1 of [1] and Proposition 5 there exists  $n_0 = n_0(\theta) > 0$  such that for all  $n > n_0$  we have

$$\mathbb{P}\left(\max_{0 \leq N \leq N_0} \|\alpha_n^{*(N)}\|_{\mathcal{H}_0} > n^{\varepsilon/2}\right) \leq \frac{1}{3n^\theta}. \quad (4.44)$$

Moreover there exists  $n_1 = n_1(\theta) > 0$  such that for all  $n > n_1$ ,

$$\begin{aligned} \mathbb{P}\left(P_n^{(N_0)} > \frac{4n^{\varepsilon/4}}{\sqrt{np_{N_0}}}\right) &\leq \mathbb{P}\left(\|\alpha_n\|_{\mathcal{F}} > n^{\varepsilon/4}\right) + \mathbb{P}\left(\max_{0 \leq N \leq N_0} \|\alpha_n^{*(N)}\|_{\mathcal{F}} > n^{\varepsilon/4}\right) \\ &+ \mathbb{P}\left(\max_{0 \leq N \leq N_0} \|\alpha_n^{*(N)}\|_{\mathcal{F}} > \frac{\sqrt{np_{N_0}}}{2}\right) \leq \frac{1}{6n^\theta}, \end{aligned} \quad (4.45)$$

$$\begin{aligned} \mathbb{P}\left(Q_n^{(N_0)} > \frac{4M_{\mathcal{F}}n^{\varepsilon/4}}{\sqrt{np_{N_0}}}\right) &\leq \mathbb{P}\left(\|\alpha_n\|_{\mathcal{F}} > n^{\varepsilon/4}\right) + \mathbb{P}\left(\|\alpha_n\|_{\mathcal{F}} > \frac{\sqrt{np_{N_0}}}{2}\right) \\ &\leq \frac{1}{6n^\theta} \end{aligned} \quad (4.46)$$

By (4.43), (4.44), (4.45) and (4.46), Lemma 5 is proved with  $n_{\xi, \theta} = \max(n_\xi, n_0, n_1)$ .

**Step 2.** By Lemma 1 of [1] we can apply Lemma 5, Theorem 2.1 and Theorem 2.1 of [1] to  $\mathcal{F}_0$ . Since these two last theorems are based on the Berthet and Mason strong approximation of  $\alpha_n(\mathcal{F})$  to  $\mathbb{G}(\mathcal{F})$ , the following Gaussian approximation is satisfied. For some constant  $C_0 = C_0(\theta) > 0$  and  $n_\theta > 0$  one can build on a probability space  $(\Omega, \mathcal{T}, \mathbb{P})$  a version of the sequence  $\{(X_n, Z_n)\}$  of i.i.d. random variables with distribution  $P^{(X, Z)}$  and a sequence  $\{\mathbb{G}_n^*(\mathcal{F})\}$  of independent  $P$ -Brownian bridges in such a way that, that for all  $n > n_\theta$ , (3.5) holds and

$$\mathbb{P}\left(\left\{\|\alpha_n^* - \mathbb{G}_n^*\|_{\mathcal{F}_0} > C_0 v_n\right\} \cup \left\{\max_{0 \leq N \leq N_0} \|F_n^{(N)}\|_{\mathcal{F}} > C_0 v_n\right\}\right) \leq \frac{1}{2n^\theta}. \quad (4.47)$$

Let denote  $\{\mathbb{G}_n^{*(N)}(\mathcal{F})\}$  the  $N$ -th raked  $P$ -Brownian bridge defined by (4.20) built through  $\{\mathbb{G}_n^*\}$ . This construction implies  $\mathbb{G}_n^{*(0)} = \mathbb{G}_n^*$ . Successive iterations given by (4.42) give  $\max_{1 \leq N \leq N_0} \|\alpha_n^{*(N)} - \mathbb{G}_n^{*(N)}\| \leq \|\alpha_n^* - \mathbb{G}_n^*\|_{\mathcal{F}_0} + \max_{0 \leq N \leq N_0} \|F_n^{(N)}\|_{\mathcal{F}}$ . By (4.47) we have for all  $n > n_\theta$ ,

$$\begin{aligned} &\mathbb{P}\left(\max_{1 \leq N \leq N_0} \|\alpha_n^{*(N)} - \mathbb{G}_n^{*(N)}\|_{\mathcal{F}} > 2C_0 v_n\right) \\ &\leq \mathbb{P}\left(\|\alpha_n^* - \mathbb{G}_n^*\|_{\mathcal{F}_0} > C_0 v_n\right) + \mathbb{P}\left(\max_{0 \leq N \leq N_0} \|F_n^{(N)}\|_{\mathcal{F}} > C_0 v_n\right) \leq \frac{1}{n^\theta}. \end{aligned}$$

We have proved (3.4) of Theorem 3.1 with  $C_\theta = 2C_0 > 0$ .

**Step 3.** To prove Theorem 3.2 we only need to extend the notations of previous steps and use Theorem 2.2 instead of Theorem 2.1 at Step 2.

#### 4.6 Proof of Proposition 6

Let  $\theta > 0$ . Proposition 8 below is proved at Step 1 then Proposition 6 is proved at Step 2. Like  $\tilde{\alpha}_n^*(\mathcal{F})$  given by (4.2) let us define the process  $\tilde{\alpha}_n^{(N)*}(\mathcal{F})$  by  $\tilde{\alpha}_n^{(N)*}(f) = (T_n^{(N)}/n)\alpha_n^{(N)*}(f)$ . The following proposition establishes the strong approximation of  $\tilde{\alpha}_n^{(N)*}(\mathcal{F})$  to  $\mathbb{G}^{(N)*}(\mathcal{F})$ .

**Proposition 8.** *Assume (1.3), (1.4) and either (VC) or (BR). There exists  $C_\theta, n_\theta > 0$ , a sequence  $\{(X_n, Z_n)\}$  of i.i.d. random variables with distribution  $P^{(X,Z)}$  and a sequence  $\{(\mathbb{G}_n(\mathcal{F}), \mathbb{G}_n^{(0)*}(\mathcal{F})), \dots, \mathbb{G}_n^{(N_0)*}(\mathcal{F})\}$  of  $P$ -Brownian bridges, all of these sequences being defined on the same probability space, such that for all  $n > n_\theta$ ,*

$$\mathbb{P}\left(\{\|\alpha_n - \mathbb{G}_n\|_{\mathcal{F}} > C_\theta v_n\} \cup \bigcup_{N=0}^{N_0} \{\|\tilde{\alpha}_n^{(N)*} - \mathbb{G}_n^{(N)*}\|_{\mathcal{F}} > C_\theta v_n\}\right) \leq \frac{1}{n^\theta}. \quad (4.48)$$

**Step 1.** By Lemma 1 one can apply Theorem 2.1 of [1] to  $(\alpha_n^{(X,Z)})^{(N)}(\mathcal{F}_1)$  the  $N$ -th raked empirical process of  $\alpha_n^{(X,Z)}$  defined by (1.6), the law  $P^{(X,Z)}$  and the auxiliary information  $\mathcal{A}^{0,(N)}$  defined by (4.21). There exists  $C_0 = C_0(\theta), n_0 = n_0(\theta) > 0$  and a probability space which supports a sequence of independent random variables  $\{(X_n, Z_n)\}$  distributed as  $P^{(X,Z)}$  and a sequence  $\{\mathbb{G}_n^{0,(0)}(\mathcal{F}_1), \dots, \mathbb{G}_n^{0,(N_0)}(\mathcal{F}_1)\}$  of process of independent  $P^{(X,Z)}$ -raked Brownian bridge satisfying for all  $n > n_0$ ,

$$\mathbb{P}\left(\max_{0 \leq N \leq N_0} \|(\alpha_n^{(X,Z)})^{(N)} - \mathbb{G}_n^{0,(N)}\|_{\mathcal{F}_1} > C_0 v_n\right) \leq \frac{1}{7n^\theta}. \quad (4.49)$$

Since  $\alpha_n(f) = (\alpha_n^{(X,Z)})^{(0)}(h_f^{(3)})$  then (4.49) implies immediately

$$\mathbb{P}(\|\alpha_n - \mathbb{G}_n\|_{\mathcal{F}} > C_\theta v_n) \leq \frac{1}{7n^\theta}. \quad (4.50)$$

Let decompose  $\tilde{\alpha}_n^{(N)*}(f)$  as the same way as (4.5):

$$\begin{aligned} \tilde{\alpha}_n^{(N)*}(f) &= (\alpha_n^{(X,Z)})^{(N)}(h_f^{(1)}) - \mathbb{E}[f] (\alpha_n^{(X,Z)})^{(N)}(h_f^{(2)}) \\ &\quad - (\alpha_n^{(X,Z)})^{(N)}(h_f^{(3)}) \left(1 + \left(T_n^{(N)}/n - 1\right)\right), \end{aligned} \quad (4.51)$$

where  $h_f^{(1)}, h_f^{(2)}, h_f^{(3)}$  are defined by (4.3). Let  $C'_0 = C'_0(\theta) = C_0 \max(1, M_{\mathcal{F}})$ . By (4.49), (4.51) and definition of  $\mathbb{G}^{(N)*}(\mathcal{F})$  given by (4.22) it holds for all

$n > n_0$ ,

$$\begin{aligned} & \mathbb{P} \left( \max_{0 \leq N \leq N_0} \|\tilde{\alpha}_n^{(N)*} - \mathbb{G}_n^{(N)*}\|_{\mathcal{F}} > C'_0 v_n \right) \\ & \leq \frac{3}{7n^\theta} + \mathbb{P} \left( \max_{0 \leq N \leq N_0} \left| \frac{T_n^{(N)}}{n} - 1 \right| \|(\alpha_n^{(X,Z)})^{(N)}\|_{\mathcal{F}_1} > C_0 v_n \right). \end{aligned} \quad (4.52)$$

Last term of (4.52) can be bounded as the same way as (4.29). By (4.49) and definition of  $\varepsilon$  we have for all  $n > n_0$ ,

$$\begin{aligned} & \mathbb{P} \left( \max_{0 \leq N \leq N_0} \left| \frac{T_n^{(N)}}{n} - 1 \right| \|(\alpha_n^{(X,Z)})^{(N)}\|_{\mathcal{F}_1} > C_0 v_n \right) \\ & \leq \mathbb{P} \left( \|(\alpha_n^{(X,Z)})^{(N)}\|_{\mathcal{F}_1} > C_0 v_n n^{1/2-\varepsilon} \right) + \mathbb{P} \left( \max_{0 \leq N \leq N_0} \sqrt{n} \left| \frac{T_n^{(N)}}{n} - 1 \right| > n^\varepsilon \right) \\ & \leq \frac{1}{7n^\theta} + \mathbb{P} \left( \|\mathbb{G}_n^{0,(N)}\|_{\mathcal{F}_1} > C'_0 v_n n^{1/2-\varepsilon} \right) + \mathbb{P} \left( \max_{0 \leq N \leq N_0} \sqrt{n} \left| \frac{T_n^{(N)}}{n} - 1 \right| > n^\varepsilon \right) .. \end{aligned} \quad (4.53)$$

We do the same calculations as (4.32). By Borell-Sudakov there exists  $n_1 = n_1(\theta) > 0$  such that for all  $n > n_1$ ,

$$\mathbb{P}(\|\mathbb{G}_n^{0,(N)}\|_{\mathcal{F}_1} > C_\theta v_n n^{1/2-\varepsilon}) \leq \frac{1}{7n^\theta}. \quad (4.54)$$

Since  $\sqrt{n}(T_n^{(N)}/n - 1) = (\alpha_n^{(X,Z)})^{(N)}(\mathbf{1} \times \phi_{M_Z})$ , Proposition 3 of [1] implies that there exists  $n_2 = n_2(\theta) > 0$  such that for all  $n > n_2$ ,

$$\mathbb{P} \left( \max_{0 \leq N \leq N_0} \sqrt{n} \left| \frac{T_n^{(N)}}{n} - 1 \right| > n^\varepsilon \right) \leq \mathbb{P} \left( \max_{0 \leq N \leq N_0} \|(\alpha_n^{(X,Z)})^{(N)}\|_{\mathcal{F}_1} > n^\varepsilon \right) \leq \frac{1}{7n^\theta}. \quad (4.55)$$

By (4.50), (4.52), (4.53), (4.54) and (4.55) we have shown that (4.48) is satisfied with  $n_\theta = \max(n_0, n_1, n_2)$  and  $C_\theta = C'_0$ .

**Step 2.** According to Proposition 8 there exists  $C_0 = C_0(\theta) > 0, n_0 = n_0(\theta) > 0$ , a sequence  $\{(X_n, Z_n)\}$  of i.i.d. random variables with distribution  $P^{(X,Z)}$  and a sequence  $\{(\mathbb{G}_n(\mathcal{F}), \mathbb{G}_n^{(0)*}(\mathcal{F}), \dots, \mathbb{G}_n^{(N_0)*}(\mathcal{F}))\}$  of  $P$ -Brownian bridges, all of these sequences being defined on the same probability space, such that for all  $n > n_0$ ,

$$\mathbb{P} \left( \{ \|\alpha_n - \mathbb{G}_n\|_{\mathcal{F}} > C_0 v_n \} \cup \bigcup_{N=0}^{N_0} \{ \|\tilde{\alpha}_n^{(N)*} - \mathbb{G}_n^{(N)*}\|_{\mathcal{F}} > C_0 v_n \} \right) \leq \frac{1}{5n^\theta}. \quad (4.56)$$

In particular for all  $n > n_0$ ,

$$\mathbb{P}(\|\alpha_n - \mathbb{G}_n\|_{\mathcal{F}} > C_0 v_n) \leq \frac{1}{5n^\theta}. \quad (4.57)$$

To prove (3.8) we use the same method as the proof of Theorem 2.1. We write

$$\begin{aligned} & \mathbb{P}\left(\max_{0 \leq N \leq N_0} \|\alpha_n^{(N)*} - \mathbb{G}_n^{(N)*}\|_{\mathcal{F}} > 3C_0 v_n\right) \\ & \leq \mathbb{P}\left(\max_{0 \leq N \leq N_0} \frac{n}{T_n^{(N)}} \|\tilde{\alpha}_n^{(N)*} - \mathbb{G}_n^{(N)*}\|_{\mathcal{F}} > 2C_0 v_n\right) \\ & \quad + \mathbb{P}\left(\max_{0 \leq N \leq N_0} \left|\frac{n}{T_n^{(N)}} - 1\right| \|\mathbb{G}_n^{(N)*}\|_{\mathcal{F}} > C_0 v_n\right). \end{aligned} \quad (4.58)$$

From one hand, using (4.56) one can say that for all  $n > n_0$ ,

$$\begin{aligned} & \mathbb{P}\left(\max_{0 \leq N \leq N_0} \frac{n}{T_n^{(N)}} \|\tilde{\alpha}_n^{(N)*} - \mathbb{G}_n^{(N)*}\|_{\mathcal{F}} > 2C_0 v_n\right) \\ & \leq \mathbb{P}\left(\max_{0 \leq N \leq N_0} \|\tilde{\alpha}_n^{(N)*} - \mathbb{G}_n^{(N)*}\|_{\mathcal{F}} > C_0 v_n\right) + \mathbb{P}\left(\max_{0 \leq N \leq N_0} \left|\frac{T_n^{(N)}}{n} - 1\right| > 1/2\right) \\ & \leq \frac{1}{5n^\theta} + \mathbb{P}\left(\max_{0 \leq N \leq N_0} \|(\alpha_n^{(X,Z)})^{(N)}\|_{\mathcal{F}_1} > \sqrt{n}/2\right). \end{aligned} \quad (4.59)$$

By Proposition 3 of [1] there exists  $n_1 = n_1(\theta) > 0$  such that for all  $n > n_1$ ,

$$\mathbb{P}\left(\max_{0 \leq N \leq N_0} \|(\alpha_n^{(X,Z)})^{(N)}\|_{\mathcal{F}_1} > \sqrt{n}/2\right) \leq \frac{1}{5n^\theta}. \quad (4.60)$$

By (4.59) and (4.60) we have shown that for any  $n > \max(n_0, n_1)$ ,

$$\mathbb{P}\left(\max_{0 \leq N \leq N_0} \frac{n}{T_n^{(N)}} \|\tilde{\alpha}_n^{(N)*} - \mathbb{G}_n^{(N)*}\|_{\mathcal{F}} > 2C_0 v_n\right) \leq \frac{2}{5n^\theta}. \quad (4.61)$$

From the other hand, we write

$$\begin{aligned} & \mathbb{P}\left(\max_{0 \leq N \leq N_0} \left|\frac{n}{T_n^{(N)}} - 1\right| \|\mathbb{G}_n^{(N)*}\|_{\mathcal{F}} > C_0 v_n\right) \\ & \leq \mathbb{P}\left(\max_{0 \leq N \leq N_0} \|\mathbb{G}_n^{(N)*}\|_{\mathcal{F}} > C_0 v_n n^{1/2-\varepsilon}\right) + \mathbb{P}\left(\max_{0 \leq N \leq N_0} \sqrt{n} \left|\frac{n}{T_n^{(N)}} - 1\right| > n^\varepsilon\right). \end{aligned} \quad (4.62)$$

By definition of  $\varepsilon$  and Borell-Sudakov inequality, there  $n_2 = n_2(\theta) > 0$  such that for all  $n > n_2$ ,

$$\mathbb{P}\left(\|\mathbb{G}_n^*\|_{\mathcal{F}} > C v_n n^{1/2-\varepsilon}\right) \leq \frac{1}{5n^\theta}. \quad (4.63)$$

There also exists  $n_3 = n_3(\theta) > 0$  such that for all  $n > n_3$ ,

$$\mathbb{P} \left( \max_{0 \leq N \leq N_0} \sqrt{n} \left| \frac{n}{T_n^{(N)}} - 1 \right| > n^\varepsilon \right) \leq \mathbb{P} \left( \max_{0 \leq N \leq N_0} \|(\alpha_n^{(X,Z)})^{(N)}\|_{\mathcal{F}_1} > 2n^\varepsilon \right) \leq \frac{1}{5n^\theta}. \quad (4.64)$$

By (4.62), (4.63) and (4.64) it holds that for any  $n > \max(n_2, n_3)$ ,

$$\mathbb{P} \left( \max_{0 \leq N \leq N_0} \left| \frac{n}{T_n^{(N)}} - 1 \right| \|\mathbb{G}_n^{(N)*}\|_{\mathcal{F}} > C_0 v_n \right) \leq \frac{2}{5n^\theta}. \quad (4.65)$$

By (4.57) (4.58), (4.61) and (4.65) we have shown (3.8) with  $n_\theta = \max(n_0, n_1, n_2, n_3)$  and  $C_\theta = 3C_0 > 0$ .

#### 4.7 Proof of Corollaries 1 and 2

For the proofs of these corollaries we need to introduce the following quantities

$$R_{n,0} = \sqrt{n}R \left( \frac{\alpha_n}{\sqrt{n}} \right), \quad R_{n,j} = \sqrt{n}R \left( \frac{\alpha_{n,(j)}^* + \alpha_n}{\sqrt{n}} \right), \quad j = 1, \dots, b_n. \quad (4.66)$$

By definition of  $R$  these quantities are bounded by

$$\|R_{n,0}\|_{\mathcal{F}} \leq \frac{\|\alpha_n\|_{\mathcal{F}}^q}{n^{(q-1)/2}}, \quad \|R_{n,j}\|_{\mathcal{F}} \leq \frac{(2^{q-1}\|\alpha_n\|_{\mathcal{F}}^q + 2^{q-1}\|\alpha_{n,(j)}^*\|_{\mathcal{F}}^q)}{n^{(q-1)/2}}.$$

Talagrand inequalities given by Proposition 5 and Proposition 3 of [1] applied to  $\alpha_n$  and  $\alpha_{n,(j)}^*$  prove that for any  $\xi, \theta > 0$  there exists  $n_{\xi,\theta} > 0$  such that for all  $n > n_{\xi,\theta}$ ,

$$\mathbb{P}(\|R_{n,j}\|_{\mathcal{F}} > \xi) \leq \frac{1}{n^\theta}, \quad j = 0, 1, \dots, b_n. \quad (4.67)$$

According to (4.67), Borell-Sudakov inequality and Theorem 2.2, there exists  $C_1, n_1 > 0$ , an array  $\{(X_n, Z_{n,(1)}, \dots, Z_{n,(b_n)})\}$  of i.i.d. random vectors distributed as  $P^{(X_1, Z_{1,(1)}, \dots, Z_{1,(b_n)})}$  and an array  $\{(\mathbb{G}_{n,(0)}^*(\mathcal{F}), \mathbb{G}_{n,(1)}^*(\mathcal{F}), \dots, \mathbb{G}_{n,(b_n)}^*(\mathcal{F}))\}$  of mutually independent  $P$ -Brownian bridges such that for  $n > n_1$ ,

$$\mathbb{P}(e_n > C_1 w_n) \leq \frac{1}{3n^2}, \quad (4.68)$$

with  $e_n = \max_{j=1, \dots, b_n} \left( \|\alpha_n - \mathbb{G}_{n,(0)}^*\|_{\mathcal{F}}, \|\alpha_{n,(j)}^* - \mathbb{G}_{n,(j)}^*\|_{\mathcal{F}}, \|R_{n,0}\|_{\mathcal{F}}, \|R_{n,j}\|_{\mathcal{F}} \right)$ . Now we prove Corollary 1 at Step 1 and Corollary 2 at Step 2.

**Step 1.** Since  $\mathcal{F}$  satisfies (VC), by Borell-Sudakov inequality there exists  $n_2 > 0$  such that for all  $n > n_2$ ,

$$\mathbb{P}(e'_n > n^{\alpha_0/2}) \leq \frac{1}{3n^2}, \quad (4.69)$$

with  $e'_n = \max_{j=0,1,\dots,b_n} \|\mathbb{G}_{n,(j)}^*\|_{\mathcal{F}}$ . By (4.66) one can write

$$\begin{aligned}
& n(\widehat{\text{Var}}(S_n^*) - \text{Var}(S_n)) \\
&= \frac{n}{b_n} \sum_{j=1}^{b_n} (\varphi(\mathbb{P}_{n,(j)}^*) - \varphi(\mathbb{P}_n))^2 - \text{Var}(\varphi(\mathbb{P}_n)) \\
&= \frac{n}{b_n} \sum_{j=1}^{b_n} \left[ \varphi\left(P + \frac{\alpha_{n,(j)}^* + \alpha_n}{\sqrt{n}}\right) - \varphi\left(P + \frac{\alpha_n}{\sqrt{n}}\right) \right]^2 - \text{Var}\left(\varphi\left(P + \frac{\alpha_n}{\sqrt{n}}\right)\right) \\
&= \frac{1}{b_n} \sum_{j=1}^{b_n} \left( \varphi'(P) \cdot \alpha_{n,(j)}^* + R_{n,j} + R_{n,0} \right)^2 - \text{Var}\left(\varphi'(P) \cdot \alpha_n + R_{n,0}\right),
\end{aligned}$$

which implies

$$n|\widehat{\text{Var}}(S_n^*) - \text{Var}(S_n)| \leq \left| \frac{1}{b_n} \sum_{j=1}^{b_n} \varphi'^2(P) \cdot \mathbb{G}_{n,(j)}^* - \sigma^2 \right| + 13e_n^2 + 10e_n e'_n. \quad (4.70)$$

There exists  $n_3 > 0$  such that for all  $n > n_3$ ,  $13(C_1 w_n)^2 + 10C_1 w_n n^{\alpha_0/2} \leq 11C_1 w_n n^{\alpha_0/2}$ . Notice that all  $\mathbb{G}_{n,(j)}^*$  are independent so  $\sum_{j=1}^{b_n} \varphi'^2(P) \cdot \mathbb{G}_{n,(j)}^*$  has the same distribution as  $\sigma^2 \chi^2(b_n)$ . By (4.68), (4.69) and (4.70) for all  $n > \max(n_1, n_2, n_3)$  it holds

$$\begin{aligned}
& \mathbb{P}\left(\frac{n}{\sigma^2} \left| \widehat{\text{Var}}(S_n^*) - \text{Var}(S_n) \right| \geq \delta + 11C_1 w_n n^{\alpha_0/2}\right) \\
& \leq \mathbb{P}\left(\frac{1}{\sigma^2} \left| \frac{1}{b_n} \sum_{j=1}^{b_n} \varphi'^2(P) \cdot \mathbb{G}_{n,(j)}^* - \sigma^2 \right| \geq \delta\right) + \mathbb{P}(e_n^2 > (C_1 w_n)^2) + \mathbb{P}(e_n e'_n > C_1 w_n n^{\alpha_0/2}) \\
& \leq \mathbb{P}\left(\left| \frac{\chi^2(b_n)}{b_n} - 1 \right| \geq \delta\right) + \frac{1}{n^2} \leq \alpha + \frac{1}{n^2}.
\end{aligned}$$

We have shown that (2.6) is satisfied for  $C_0 = 11C_1$  and  $n_0 = \max(n_1, n_2, n_3)$ . According to Hoeffding inequality,  $\mathbb{P}\left(\left| \frac{\chi^2(b_n)}{b_n} - 1 \right| > \delta\right) \leq 2/b_n \delta^2$ . There exists  $n_4 > 0$  such that for  $n > n_4$ ,  $\sigma^2 \left(\frac{1}{\sqrt{b_n}} + C_0 w_n n^{\alpha_0/2-1}\right) \leq C'_0/\sqrt{b_n}$  for some  $C'_0 > \sigma^2$ . For  $\delta = n/\sqrt{b_n}$  and  $n > n_4$ , (2.6) implies

$$\begin{aligned}
& \mathbb{P}\left(\left| \widehat{\text{Var}}(S_n^*) - \text{Var}(S_n) \right| > C'_0/\sqrt{b_n}\right) \\
& \leq \mathbb{P}\left(\left| \widehat{\text{Var}}(S_n^*) - \text{Var}(S_n) \right| > \sigma^2 \left(\frac{1}{\sqrt{b_n}} + C_0 w_n n^{\alpha_0/2-1}\right)\right) \leq \frac{3}{n^2}.
\end{aligned}$$

The last inequality and Borel-Cantelli implies (2.7).

**Step 2.** We denote  $G_n(x) = \widehat{F}_{S_n^*}(x + B_n)$ . Since all  $\varphi'(P) \cdot \mathbb{G}_{n,(j)}^*$  have the

same law we have  $|G_n - F_{S_n}| \leq |F_1| + |F_2| + F_3$  with

$$\begin{aligned}
F_1(x) &= \frac{1}{b_n} \sum_{j=1}^{b_n} \mathbb{1}_{\varphi(P) + \frac{1}{\sqrt{n}} \varphi'(P) \cdot \mathbb{G}_{n,(j)}^* \in I[x, x - E_{n,j}]}, \\
F_2(x) &= \frac{1}{b_n} \sum_{j=1}^{b_n} \mathbb{1}_{\varphi(P) + \frac{1}{\sqrt{n}} \varphi'(P) \cdot \mathbb{G}_{n,(j)}^* \leq x} - \mathbb{P} \left( \varphi(P) + \frac{1}{\sqrt{n}} \varphi'(P) \cdot \mathbb{G}_{n,(0)}^* \leq x \right), \\
F_3(x) &= \mathbb{P} \left( \varphi(P) + \frac{1}{\sqrt{n}} \varphi'(P) \cdot \mathbb{G}_{n,(0)}^* \in I[x, x - E_{n,0}] \right), \\
E_{n,j} &= \frac{1}{\sqrt{n}} \varphi'(P) \cdot (\alpha_{n,(j)}^* - \mathbb{G}_{n,(j)}^*) + R_{n,j} - R_{n,0}, \\
E_{n,0} &= \frac{1}{\sqrt{n}} \varphi'(P) \cdot (\alpha_n - \mathbb{G}_{n,(0)}^*) + R_{n,0}.
\end{aligned}$$

where  $R_{n,0}, R_{n,j}$  are defined by (4.66) and  $I[a, b]$  means the interval of the real contained between  $a$  and  $b$ . Notice that  $\|E_{n,j}\|_{\mathcal{F}} \leq 3e_n$  for  $j = 0, 1, \dots, b_n$ . We denote  $C'_1 = C_1/3$  and  $I_n(x) = [x - C'_1 w_n, x + C'_1 w_n]$ . Remind that  $A_n = [\varphi(P) - a_n; \varphi(P) + a_n]$  with  $a_n = \sigma \sqrt{\log(\log n / b_n w_n^2)}$ . Then for  $C = \sqrt{2} C'_1 / \pi \sigma > 0$  we have for any  $n > 1$ ,

$$\begin{aligned}
\sup_{x \in A_n^C} \mathbb{P} \left( \varphi(P) + \frac{1}{\sqrt{n}} \varphi'(P) \cdot \mathbb{G}_{n,(0)}^* \in I_n(x) \right) &\leq C w_n \exp \left( -\frac{a_n^2}{2\sigma^2} \right) \\
&\leq C \sqrt{\frac{\log n}{b_n}}. \tag{4.71}
\end{aligned}$$

Under the event  $B_n = \{e_n \leq C_1 w_n\}$  we have

$$\begin{aligned}
|F_1(x)| &\leq \frac{1}{b_n} \sum_{i=1}^{b_n} \mathbb{1}_{\varphi(P) + \frac{1}{\sqrt{n}} \varphi'(P) \cdot \mathbb{G}_{n,(j)}^* \in I_n(x)} \\
&\leq \left| \frac{1}{b_n} \sum_{i=1}^{b_n} \mathbb{1}_{\varphi(P) + \frac{1}{\sqrt{n}} \varphi'(P) \cdot \mathbb{G}_{n,(j)}^* \in I_n(x)} - \mathbb{P} \left( \varphi(P) + \frac{1}{\sqrt{n}} \varphi'(P) \cdot \mathbb{G}_{n,(0)}^* \in I_n(x) \right) \right| \\
&\quad + \mathbb{P} \left( \varphi(P) + \frac{1}{\sqrt{n}} \varphi'(P) \cdot \mathbb{G}_{n,(0)}^* \in I_n(x) \right). \tag{4.72}
\end{aligned}$$

Since all  $G_{n,(j)}^*$  are i.i.d. we have according to (4.71), (4.72) and DKW inequality that for all  $\varepsilon > 0$  and  $n > 1$ ,

$$\mathbb{P} \left( \sup_{x \in A_n^C} |F_1(x)| > 2\varepsilon + C \sqrt{\frac{\log n}{b_n}} \right) \leq 4e^{-2b_n \varepsilon^2}. \tag{4.73}$$

By DWK inequality we have also

$$\mathbb{P} \left( \sup_{x \in A_n^C} |F_2(x)| > \varepsilon \right) \leq \mathbb{P} \left( \sup_{x \in \mathbb{R}} |F_2(x)| > \varepsilon \right) \leq 2e^{-2b_n \varepsilon^2}. \tag{4.74}$$

Under the event  $B_n$ ,  $|F_3(x)| \leq \mathbb{P}(\varphi(P) + \frac{1}{\sqrt{n}}\varphi'(P) \cdot \mathbb{G}_{n,(0)}^* \in I_n(x))$  then (4.71) implies that for  $n > 1$ ,

$$\sup_{x \in A_n^c} |F_3(x)| \leq C \sqrt{\frac{\log n}{b_n}}. \quad (4.75)$$

We have shown by (4.68), (4.73), (4.74) and (4.75) that for all  $\varepsilon > 0$  and  $n > n_1$ ,

$$\mathbb{P} \left( \sup_{x \in A_n^c} |G_n(x) - F_{S_n}(x)| > 3\varepsilon + 2C \sqrt{\frac{\log n}{b_n}} \right) \leq \frac{1}{3n^2} + 6e^{-2b_n\varepsilon^2},$$

which implies in particular for  $\varepsilon = \sqrt{\log(3n)/b_n} \leq 2\sqrt{\log n/b_n}$  and  $n > n_1$ ,

$$\mathbb{P} \left( \sup_{x \in A_n^c} |G_n(x) - F_{S_n}(x)| > (6 + 2C) \sqrt{\frac{\log n}{b_n}} \right) \leq \frac{1}{n^2}.$$

We have proved by Borel-Cantelli (2.8) with  $C_0 = 6 + 2C > 0$ .

#### 4.8 Proof of Corollaries (3) and (4)

At Step 1 we prove Corollary (3) and at Step 2 we prove Corollary (4).

**Step 1.** According to Theorem 2.1, there exists  $C > 0, n_1 > 0$ , a sequence  $\{(X_n, Z_n)\}$  of i.i.d. random variables with distribution  $P^{(X,Z)}$  and a sequence of  $\{\mathbb{G}_n^*(\mathcal{F})\}$  of  $P$ -Brownian bridge such that for all  $n > n_1$ ,  $\mathbb{P}(A_n) \leq 1/n^2$  with

$$A_n = \{|\alpha_n - \mathbb{G}_n|_{\mathcal{F}} \geq Cv_n/4\} \cup \{|\alpha_n^* - \mathbb{G}_n^*|_{\mathcal{F}} \geq Cv_n/4\}.$$

Then we have for all  $n > n_0$

$$\begin{aligned} \mathbb{P}(\phi(\alpha_n^*(f)) \leq x) &\leq \mathbb{P} \left( \phi(\mathbb{G}_n^*(f)) \leq x + \frac{1}{4}Cv_n \right) + P(A_n) \\ &\leq \mathbb{P}(\phi(\mathbb{G}_n^*(f)) \leq x) + \frac{1}{4}CC_0C_1v_n + \frac{1}{n^2}. \end{aligned}$$

Similarly,

$$\begin{aligned} \mathbb{P}(\phi(\mathbb{G}_n^*(f)) \leq x) - CC_0C_1v_n &\leq \mathbb{P} \left( \phi(\mathbb{G}_n^*(f)) \leq x - \frac{1}{4}Cv_n \right) \\ &\leq \mathbb{P}(\phi(\alpha_n^*(f)) \leq x) + P(A_n) \\ &\leq \mathbb{P}(\phi(\alpha_n^*(f)) \leq x) + \frac{1}{n^2}. \end{aligned}$$

By definition of  $v_n$  there exists  $n_2 > 0$  such that for all  $n > n_2$ ,  $1/n^2 \leq CC_0C_1v_n/4$ . So for  $n > \max(n_1, n_2)$ ,

$$\sup_{\phi \in \mathcal{L}} \sup_{f \in \mathcal{F}} \sup_{x \in \mathbb{R}} |\mathbb{P}(\phi(\alpha_n^*(f)) \leq x) - \mathbb{P}(\phi(\mathbb{G}(f)) \leq x)| \leq \frac{1}{2}CC_0C_1v_n. \quad (4.76)$$



We have shown that (2.9) holds for  $n_0 = \max(n_1, n_2)$ . As the same way, we can show that for  $n > n_0$ ,

$$\sup_{\phi \in \mathcal{L}} \sup_{f \in \mathcal{F}} \sup_{x \in \mathbb{R}} |\mathbb{P}(\phi(\alpha_n(f)) \leq x) - \mathbb{P}(\phi(\mathbb{G}(f)) \leq x)| \leq \frac{1}{2} C C_0 C_1 v_n. \quad (4.77)$$

Inequality (2.10) is a consequence of (4.76) and (4.77).

**Step 2.** We prove (3.6) and (3.7) is a direct consequence of (3.6). We decompose the sum of the left-hand side of (3.6) as

$$\begin{aligned} & \frac{1}{b_n} \sum_{j=1}^{b_n} \alpha_{n,(j)}^{*(N)}(f) \alpha_{n,(j)}^{*(N)}(g) - \text{Cov}(\mathbb{G}_{n,(j)}^{*(N)}(f), \mathbb{G}_{n,(j)}^{*(N)}(g)) \\ & \leq \frac{1}{b_n} \sum_{j=1}^{b_n} \sup_{f, g \in \mathcal{F}} \left| \mathbb{G}_{n,(j)}^{*(N)}(f) \mathbb{G}_{n,(j)}^{*(N)}(g) - \text{Cov}(\mathbb{G}_{n,(j)}^{*(N)}(f), \mathbb{G}_{n,(j)}^{*(N)}(g)) \right| \\ & \quad + \max_{j=1, \dots, b_n} \|\alpha_{n,(j)}^{*(N)} - \mathbb{G}_{n,(j)}^{*(N)}\|_{\mathcal{F}} \left( \|\alpha_{n,(j)}^{*(N)}\|_{\mathcal{F}} + \|\mathbb{G}_{n,(j)}^{*(N)}\|_{\mathcal{F}} \right) \end{aligned} \quad (4.78)$$

According to Theorem 3.2 applied with  $\theta = 2$  there exists  $C_1 > 0$ ,  $n_1 > 0$  and a probability space supporting a triangular array  $\{(X_n, Z_{n,(1)}, \dots, Z_{n,(b_n)})\}$  of i.i.d. random vectors distributed as  $P^{(X_1, Z_{1,(1)}, \dots, Z_{1,(b_1)})}$  and a triangular array  $\{(\mathbb{G}_{n,(0)}^*(\mathcal{F}), \mathbb{G}_{n,(1)}^*(\mathcal{F}), \dots, \mathbb{G}_{n,(b_n)}^*(\mathcal{F}))\}$  of  $(b_n + 1)$ -uplets of mutually independent  $P$ -Brownian bridges such that, for  $n \geq n_1$ ,

$$\mathbb{P} \left( \max_{j=1, \dots, b_n} \max_{1 \leq N \leq N_0} \|\alpha_{n,(j)}^{*(N)} - \mathbb{G}_{n,(j)}^{*(N)}\|_{\mathcal{F}} > C_1 w_n \right) < \frac{1}{4n^2}. \quad (4.79)$$

Borell-Sudakov inequality implies that there exists  $C_2 > 0$ ,  $n_2 > 0$  such that for all  $n > n_2$ ,

$$\begin{aligned} & \mathbb{P} \left( \max_{j=1, \dots, b_n} \max_{0 \leq N \leq N_0} \|\mathbb{G}_{n,(j)}^{*(N)}\|_{\mathcal{F}} > C_1 \log n \right) \\ & \leq b_n \max_{j=1, \dots, b_n} \max_{0 \leq N \leq N_0} \mathbb{P} \left( \|\mathbb{G}_{n,(j)}^{*(N)}\|_{\mathcal{F}} > C_2 \log n \right) \leq \frac{1}{4n^2}. \end{aligned} \quad (4.80)$$

According to Proposition 5 there exists also  $C_3 > 0$ ,  $n_3 > 0$  such that for all  $n > n_3$  it holds,

$$\mathbb{P} \left( \max_{j=1, \dots, b_n} \max_{0 \leq N \leq N_0} \|\alpha_{n,(j)}^{*(N)}\|_{\mathcal{F}} > C_3 \log n \right) \leq \frac{1}{4n^2}. \quad (4.81)$$

For all  $f, g \in \mathcal{F}$  and  $j = 1, \dots, b_n$ ,  $\mathbb{G}_{n,(j)}^{*(N)}(f) \mathbb{G}_{n,(j)}^{*(N)}(g) - \text{Cov}(\mathbb{G}_{n,(j)}^{*(N)}(f), \mathbb{G}_{n,(j)}^{*(N)}(g))$  is a square-Gaussian random variable – see Definition 3.1 of Kozachenko and Moklyachuk [24]. Since  $\mathcal{F}$  satisfies (VC), let apply Theorem 4.1 of [24] with the following parameters:  $A^+ = A^- = 1$ ,  $r(x) = \log x$ ,  $\delta_0 = \sigma_{\mathcal{F}} \leq t_0 = M_{\mathcal{F}}$ ,  $M = 1$ ,  $R(s) = (1 - |s|)^{-1/2} \exp(-|s|/2)$  – see Lemma 3.2. For  $x > 2\nu_0 \sigma_{\mathcal{F}}$ , we find

$$D_1^+(x) = D_1^-(x) \leq \frac{22e^{\nu_0}}{(2\nu_0 \sigma_{\mathcal{F}})^{\nu_0}} x^{\nu_0} \exp\left(-\frac{x}{2\sigma_{\mathcal{F}}}\right).$$

This Theorem implies that there exists  $C_4 > 0, n_4 > 0$  such that for all  $n > n_4$ ,  $2\nu_0\sigma_{\mathcal{F}} < C_4 \log n$  and

$$\begin{aligned} & \mathbb{P} \left( \frac{1}{b_n} \sum_{j=1}^{b_n} \max_{0 \leq N \leq N_0} \sup_{f, g \in \mathcal{F}} \left| \mathbb{G}_{n,(j)}^{*(N)}(f) \mathbb{G}_{n,(j)}^{*(N)}(g) - \text{Cov}(\mathbb{G}_{n,(j)}^{*(N)}(f), \mathbb{G}_{n,(j)}^{*(N)}(g)) \right| > C_4 \log n \right) \\ & \leq b_n \max_{j=1, \dots, b_n} \max_{0 \leq N \leq N_0} \mathbb{P} \left( \sup_{f, g \in \mathcal{F}} \left| \mathbb{G}_{n,(j)}^{*(N)}(f) \mathbb{G}_{n,(j)}^{*(N)}(g) - \text{Cov}(\mathbb{G}_{n,(j)}^{*(N)}(f), \mathbb{G}_{n,(j)}^{*(N)}(g)) \right| > C_4 \log n \right) \\ & \leq \frac{1}{4n^2}. \end{aligned} \tag{4.82}$$

Clearly, there exists  $C_5 > C_4, n_5 > 0$  such that  $C_4 \log n + C_0(C_1 + C_2)w_n \log n < C_5 \log n$  for all  $n > n_5$ . Finally, according to (4.78), (4.79), (4.80) and (4.82), we have for all  $n > \max(n_1, n_2, n_3, n_4, n_5)$ ,

$$\mathbb{P} \left( \max_{0 \leq N \leq N_0} \sup_{f, g \in \mathcal{F}} \left| \frac{1}{b_n} \sum_{j=1}^{b_n} \alpha_{n,(j)}^{*(N)}(f) \alpha_{n,(j)}^{*(N)}(g) - \text{Cov}(\mathbb{G}_{n,(j)}^{*(N)}(f), \mathbb{G}_{n,(j)}^{*(N)}(g)) \right| > C_5 \log n \right) \leq \frac{1}{n^2},$$

which prove (3.6) with  $n_0 = \max(n_1, n_2, n_3, n_4, n_5)$  and  $C_0 = C_5$  by Borel-Cantelli.

## References

- [1] Mickael Albertus and Philippe Berthet. “Auxiliary information : the raking-ratio empirical process”. In: (). URL: <https://arxiv.org/abs/1803.06907>.
- [2] Kenneth S. Alexander. “Probability inequalities for empirical processes and a law of the iterated logarithm”. In: *Ann. Probab.* 12.4 (1984), pp. 1041–1067. ISSN: 0091-1798. URL: [http://links.jstor.org/sici?sici=0091-1798\(198411\)12:4%3C1041:PIFEP%3E2.0.CO;2-7&origin=MSN](http://links.jstor.org/sici?sici=0091-1798(198411)12:4%3C1041:PIFEP%3E2.0.CO;2-7&origin=MSN).
- [3] Sergio Alvarez-Andrade and Salim Bouzebda. “Strong approximations for weighted bootstrap of empirical and quantile processes with applications”. In: *Stat. Methodol.* 11 (2013), pp. 36–52. ISSN: 1572-3127. URL: <https://doi.org/10.1016/j.stamet.2012.09.001>.
- [4] Niels Trolle Andersen. “The central limit theorem for nonseparable valued functions”. In: *Z. Wahrsch. Verw. Gebiete* 70.3 (1985), pp. 445–455. ISSN: 0044-3719. URL: <https://doi.org/10.1007/BF00534875>.
- [5] Gutti Jogesh Babu. “Bootstrapping statistics with linear combinations of chi-squares as weak limit”. In: *Sankhyā Ser. A* 46.1 (1984), pp. 85–93. ISSN: 0581-572X.
- [6] Philippe Barbe and Patrice Bertail. *The weighted bootstrap*. Vol. 98. Lecture Notes in Statistics. Springer-Verlag, New York, 1995, pp. viii+230. ISBN: 0-387-94478-8. URL: <https://doi.org/10.1007/978-1-4612-2532-4>.

- [7] RJ Beran, L Le Cam, and PW Millar. “Convergence of stochastic empirical measures”. In: Journal of multivariate analysis 23.1 (1987), pp. 159–168.
- [8] Philippe Berthet and David M. Mason. “Revisiting two strong approximation results of Dudley and Philipp”. In: IMS Lecture Notes Monogr. Ser. 51 (2006), pp. 155–172. URL: <https://doi.org/10.1214/074921706000000824>.
- [9] Peter J. Bickel and David A. Freedman. “Some asymptotic theory for the bootstrap”. In: Ann. Statist. 9.6 (1981), pp. 1196–1217. ISSN: 0090-5364. URL: [http://links.jstor.org/sici?sici=0090-5364\(198111\)9:6%3C1196:SATFTB%3E2.0.CO;2-R&origin=MSN](http://links.jstor.org/sici?sici=0090-5364(198111)9:6%3C1196:SATFTB%3E2.0.CO;2-R&origin=MSN).
- [10] Patrick Billingsley. Convergence of probability measures. Second. Wiley Series in Probability and Statistics: Probability and Statistics. A Wiley-Interscience Publication. John Wiley & Sons, Inc., New York, 1999, pp. x+277. ISBN: 0-471-19745-9. DOI: 10.1002/9780470316962. URL: <https://doi.org/10.1002/9780470316962>.
- [11] James G Booth and Somnath Sarkar. “Monte Carlo approximation of bootstrap variances”. In: The American Statistician 52.4 (1998), pp. 354–357.
- [12] David T. Brown. “A note on approximations to discrete probability distributions”. In: Information and Control 2 (1959), pp. 386–392. ISSN: 0890-5401.
- [13] Tapas K. Chandra and J. K. Ghosh. “Valid asymptotic expansions for the likelihood ratio statistic and other perturbed chi-square variables”. In: Sankhyā Ser. A 41.1-2 (1979), pp. 22–47. ISSN: 0581-572X.
- [14] W. Edwards Deming and Frederick F. Stephan. “On a least squares adjustment of a sampled frequency table when the expected marginal totals are known”. In: Ann. Math. Statistics 11 (1940), pp. 427–444. ISSN: 0003-4851.
- [15] B. Efron. “Bootstrap methods: another look at the jackknife”. In: Ann. Statist. 7.1 (1979), pp. 1–26. ISSN: 0090-5364. URL: [http://links.jstor.org/sici?sici=0090-5364\(197901\)7:1%3C1:BMALAT%3E2.0.CO;2-6&origin=MSN](http://links.jstor.org/sici?sici=0090-5364(197901)7:1%3C1:BMALAT%3E2.0.CO;2-6&origin=MSN).
- [16] Bradley Efron. The jackknife, the bootstrap and other resampling plans. Vol. 38. CBMS-NSF Regional Conference Series in Applied Mathematics. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, Pa., 1982, pp. vi+92. ISBN: 0-89871-179-7.
- [17] Evarist Giné and Joel Zinn. “Bootstrapping general empirical measures”. In: Ann. Probab. 18.2 (1990), pp. 851–869. ISSN: 0091-1798. URL: [http://links.jstor.org/sici?sici=0091-1798\(199004\)18:2%3C851:BGEM%3E2.0.CO;2-6&origin=MSN](http://links.jstor.org/sici?sici=0091-1798(199004)18:2%3C851:BGEM%3E2.0.CO;2-6&origin=MSN).
- [18] Peter Hall. “Rate of convergence in bootstrap approximations”. In: Ann. Probab. 16.4 (1988), pp. 1665–1684. ISSN: 0091-1798. URL: [http://links.jstor.org/sici?sici=0091-1798\(198810\)16:4%3C1665:ROCIBA%3E2.0.CO;2-Z&origin=MSN](http://links.jstor.org/sici?sici=0091-1798(198810)16:4%3C1665:ROCIBA%3E2.0.CO;2-Z&origin=MSN).

- [19] Peter Hall. The bootstrap and Edgeworth expansion. Springer Series in Statistics. Springer-Verlag, New York, 1992, pp. xiv+352. ISBN: 0-387-97720-1. URL: <https://doi.org/10.1007/978-1-4612-4384-7>.
- [20] Jørgen Hoffmann-Jørgensen. Stochastic processes on Polish spaces. 39. Aarhus Universitet. Matematisk Institut, 1991.
- [21] C. T. Ireland and S. Kullback. “Contingency tables with given marginals”. In: Biometrika 55 (1968), pp. 179–188. ISSN: 0006-3444. URL: <https://doi.org/10.1093/biomet/55.1.179>.
- [22] J. Komlós, P. Major, and G. Tusnády. “An approximation of partial sums of independent RV’s and the sample DF. I”. In: Z. Wahrscheinlichkeitstheorie und Verw. Gebiete 32 (1975), pp. 111–131. URL: <https://doi.org/10.1007/BF00533093>.
- [23] Michael R. Kosorok. Introduction to empirical processes and semiparametric inference. Springer Series in Statistics. Springer, New York, 2008, pp. xiv+483. ISBN: 978-0-387-74977-8. URL: <https://doi.org/10.1007/978-0-387-74978-5>.
- [24] Yuri Kozachenko and Oksana Moklyachuk. “Large deviation probabilities for square-Gaussian stochastic processes”. In: Extremes 2.3 (1999), 269–293 (2000). ISSN: 1386-1999. DOI: 10.1023/A:1009907019950. URL: <https://doi.org/10.1023/A:1009907019950>.
- [25] P. M. Lewis II. “Approximating probability distributions to reduce storage requirements”. In: Information and Control 2 (1959), pp. 214–225. ISSN: 0890-5401.
- [26] David M. Mason and Michael A. Newton. “A rank statistics approach to the consistency of a general bootstrap”. In: Ann. Statist. 20.3 (1992), pp. 1611–1624. ISSN: 0090-5364. URL: <https://doi.org/10.1214/aos/1176348787>.
- [27] Jens Præstgaard and Jon A. Wellner. “Exchangeably weighted bootstraps of the general empirical process”. In: Ann. Probab. 21.4 (1993), pp. 2053–2086. ISSN: 0091-1798. URL: [http://links.jstor.org/sici?sici=0091-1798\(199310\)21:4%3C2053:EWBTG%3E2.0.CO;2-W&origin=MSN](http://links.jstor.org/sici?sici=0091-1798(199310)21:4%3C2053:EWBTG%3E2.0.CO;2-W&origin=MSN).
- [28] Donald B. Rubin. “The Bayesian bootstrap”. In: Ann. Statist. 9.1 (1981), pp. 130–134. ISSN: 0090-5364. URL: [http://links.jstor.org/sici?sici=0090-5364\(198101\)9:1%3C130:TBB%3E2.0.CO;2-0&origin=MSN](http://links.jstor.org/sici?sici=0090-5364(198101)9:1%3C130:TBB%3E2.0.CO;2-0&origin=MSN).
- [29] Jun Shao and Dong Sheng Tu. The jackknife and bootstrap. Springer Series in Statistics. Springer-Verlag, New York, 1995, pp. xviii+516. ISBN: 0-387-94515-6. URL: <https://doi.org/10.1007/978-1-4612-0795-5>.
- [30] Yongzhao Shao. “Rate of convergence of bootstrapped empirical measures”. In: Statist. Probab. Lett. 53.3 (2001), pp. 293–298. ISSN: 0167-7152. URL: [https://doi.org/10.1016/S0167-7152\(01\)00073-6](https://doi.org/10.1016/S0167-7152(01)00073-6).
- [31] Kesar Singh. “On the asymptotic accuracy of Efron’s bootstrap”. In: Ann. Statist. 9.6 (1981), pp. 1187–1195. ISSN: 0090-5364. URL: [http://links.jstor.org/sici?sici=0090-5364\(198111\)9:6%3C1187:OTAAOE%3E2.0.CO;2-6&origin=MSN](http://links.jstor.org/sici?sici=0090-5364(198111)9:6%3C1187:OTAAOE%3E2.0.CO;2-6&origin=MSN).

- [32] Richard Sinkhorn. “A relationship between arbitrary positive matrices and doubly stochastic matrices”. In: Ann. Math. Statist. 35 (1964), pp. 876–879. ISSN: 0003-4851. URL: <https://doi.org/10.1214/aoms/1177703591>.
- [33] Richard Sinkhorn. “Diagonal equivalence to matrices with prescribed row and column sums”. In: Amer. Math. Monthly 74 (1967), pp. 402–405. ISSN: 0002-9890. URL: <https://doi.org/10.2307/2314570>.
- [34] Frederick F Stephan. “An iterative method of adjusting sample frequency tables when expected marginal totals are known”. In: The Annals of Mathematical Statistics 13.2 (1942), pp. 166–178.
- [35] Aad W. van der Vaart and Jon A. Wellner. Weak convergence and empirical processes. Springer Series in Statistics. With applications to statistics. Springer-Verlag, New York, 1996, pp. xvi+508. ISBN: 0-387-94640-3. URL: <https://doi.org/10.1007/978-1-4757-2545-2>.