



HAL
open science

An extended continuous mapping theorem for outer almost sure weak convergence

Davit Varron

► **To cite this version:**

Davit Varron. An extended continuous mapping theorem for outer almost sure weak convergence. Theory of Probability and Its Applications c/c of Teoriiia Veroiatnostei i Ee Primenenie, 2019, 64 (2), pp.304-323. 10.1137/S0040585X97T989507. hal-03207677

HAL Id: hal-03207677

<https://hal.science/hal-03207677>

Submitted on 26 Apr 2021

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.

An extended continuous mapping theorem for outer almost sure weak convergence

Davit Varron, Université de Bourgogne-Franche-Comté, UMR CNRS 6623

October 20, 2018

Abstract:

We prove an extended continuous mapping theorem for outer almost sure weak convergence in a metric space, a notion that is used in bootstrap empirical processes theory. Then we make use of those results to establish the consistency of several bootstrap procedures in empirical likelihood theory for functional parameters.

1 Introduction and main results

1.1 Introduction

The asymptotic theory of bootstrap empirical processes has been the subject of important investigations. The archetype of result in this theory is a Donsker type theorem for bootstrap empirical measures, which holds *conditionally to the original sampling sequence*, a theory started by Giné and Zinn [9] - see also [8]. Such results are in the vein of the conditional multiplier Donsker theorem - see [11] - and describe an "outer almost sure weak convergence of the bootstrap empirical process" - a mathematical notion that has to be handled with particular care, as pointed out by Giné [8]. The classical notion of weak convergence in a metric space (see, e.g., [4, 13, 14]) is a central probabilistic tool for statistical purposes. Such a type of convergence benefits from crucial stability properties, the most general and useful one being the continuous mapping theorem: if a sequence of random elements (\mathcal{Z}_n) converges weakly (in a metric space (D, d_D)) to a Borel separable probability measure - represented by a random variable \mathcal{Z} - and if g is continuous from on (D, d_D) to a metric space (E, d_E) , then $g(\mathcal{Z}_n)$ converges weakly to $g(\mathcal{Z})$ in (E, d_E) . This theorem was then extended by Prokhorov to a more general case, establishing the weak convergence of $g_n(\mathcal{Z}_n)$ with the sequence of maps (g_n) converging to g in a somewhat minimal sense - see [13]. Gill *et al.* [7, p. 125] finally extended this result to the framework of *weak convergence of random elements in a metric space*, in the sense of Hoffman-Jørgensen (see [10]). That notion of weak convergence of random elements is more general than that of weak convergence of probabilities (when (D, d_D) is not separable). Indeed it is possible to consider sequences \mathcal{Z}_n that

fail to be Borel measurable for each $n \geq 1$. Instead, sequences are only required to be asymptotically Borel measurable. This subtlety turns out to be crucial for general empirical processes, as this type of lack of measurability often occurs - as first pointed out by Chibisov [5].

In this article, we prove a similar refined continuous mapping theorem for the notion of "outer almost sure weak convergence" which occurs in bootstrap Donsker limit theorems. To the best of our knowledge, this problem was never treated in the literature. Our main motivation for proving such a result is to apply it to a general framework of statistical applications: bootstrapping the limit law of some self normalised processes that arise in empirical likelihood theory for the estimation of trajectories- see §2. Such a result is of course far from being surprising and completely complies to the intuition of any statistician. Its proof is however more involved than first expected, due to some important nonmeasurability issues. The best way to explain this is to recall the rigorous definition of "outer almost sure weak convergence" that is used in bootstrap empirical processes theory. This is the aim of the forthcoming subsection.

1.2 The probabilistic framework

Weak convergence of random elements in a metric space can itself be metrized by the bounded Lipschitz distance. However it is not guaranteed that this distance between the bootstrap empirical process and its Gaussian counterpart is measurable with respect to the observed (i.e. non bootstrap) sample. For this reason, the notion of outer probability/ outer expectation will be used. Here is the most concise reminder of the notion of inner/outer expectation that will be adapted to our purposes - for further details see, e.g. [15, Chapter 1.2]. Given a probability space $(\Omega, \mathcal{A}, \mathbb{P})$ and given a bounded map \mathcal{H} from Ω to \mathbb{R} we shall define

$$\mathbb{E}^*(\mathcal{H}) := \mathbb{E}(\mathcal{H}^*),$$

where \mathcal{H}^* denotes (any version of) the measurable cover function of \mathcal{H} , namely any measurable function \mathcal{H}^* which satisfies $\mathcal{H}^* \geq \mathcal{H}$ everywhere on Ω , and which satisfies the following property: for any map \mathcal{H}' from Ω to \mathbb{R} that is Borel measurable with respect to \mathcal{A} , if $\mathcal{H} \leq \mathcal{H}'$ everywhere, then one has $\mathcal{H}' \geq \mathcal{H}^*$ \mathbb{P} -almost everywhere. In order to maintain some meaningful precision in our notations, we shall frequently write $\mathbb{E}_\Omega^*(\mathcal{H}) = \mathbb{E}(\mathcal{H}_\Omega^*)$ instead of $\mathbb{E}^*/\mathcal{H}^*$ to keep track of the dependency upon $(\Omega, \mathcal{A}, \mathbb{P})$. Similarly, one can define $\mathbb{E}_*(\mathcal{H}) := -\mathbb{E}^*(-\mathcal{H})$. We shall also write $\mathbb{P}^*(A)$ for $\mathbb{E}^*(\mathbb{1}_A)$. In the particular case of a product probability space

$$(\Omega, \mathcal{A}, \mathbb{P}) = \left(\prod_{\ell=1}^L \Omega_\ell, \bigotimes_{\ell=1}^L \mathcal{A}_\ell, \bigotimes_{\ell=1}^L \mathbb{P}_\ell \right), \quad (1)$$

we shall also use the notion of partial outer expectation: for fixed $\ell \in \llbracket 1, L \rrbracket$ and for any $\omega_{-\ell} = (\omega_1, \dots, \omega_{\ell-1}, \omega_{\ell+1}, \dots, \omega_L)$ one can define the partial map

$$\mathcal{H}_{\omega_{-\ell}}^{(\ell)} : \omega_\ell \rightarrow \mathcal{H}(\omega_1, \dots, \omega_{\ell-1}, \omega_\ell, \omega_{\ell+1}, \dots, \omega_L),$$

on Ω_ℓ . We shall then denote by $(\mathcal{H}_{\omega_{-\ell}})_{\Omega_\ell}^*$ the measurable cover function of $\mathcal{H}_{\omega_{-\ell}}^{(\ell)}(\cdot)$ (implicitly on the probability space $(\Omega_\ell, \mathcal{A}_\ell, \mathbb{P}_\ell)$), and we will write

$$\mathbb{E}_{\Omega_\ell}^*(\mathcal{H}) : \omega_{-\ell} \rightarrow \mathbb{E}\left((\mathcal{H}_{\omega_{-\ell}})_{\Omega_\ell}^*\right),$$

where the last expectation is implicitly done on the probability space $(\Omega_\ell, \mathcal{A}_\ell, \mathbb{P}_\ell)$. Note that, when $L = 1$, our first notation $\mathbb{E}_{\Omega_1}^*$ does not clash with our definition of partial outer expectations: simply consider both $\mathcal{H}_{\Omega_1}^*$ and $\mathbb{E}_{\Omega_1}^*(\mathcal{H})$ as functions with no argument - which can be seen as constants.

From now on we shall consider $(\Omega, \mathcal{A}, \mathbb{P})$ as in (1), with $L = 2$. In order to stay focused on the possible applications in bootstrap theory, it may be convenient to the reader to keep in mind the following: the probability space $(\Omega_1, \mathcal{A}_1, \mathbb{P}_1)$ shall be implicitly understood as that of the random sample, while $(\Omega_2, \mathcal{A}_2, \mathbb{P}_2)$ must be considered as that of the "resampling" mechanisms generating the bootstrap empirical measure. Denote by $BL_1(D)$ the set of all reals valued functions L on D such that

$$\sup_{(z, z') \in D^2, d_D(z, z') > 0} \frac{|L(z) - L(z')|}{d_D(z, z')} + \sup_{z \in D} |L(z)| \leq 1.$$

We shall also freely use the notation BL_1 for any metric space. Our baseline assumption is as follows: we consider a sequence of maps \mathcal{Z}_n from $\Omega_1 \times \Omega_2$ to D satisfying

$$\left(\sup_{L \in BL_1(D)} \left| \mathbb{E}_{\Omega_2}^*(L(\mathcal{Z}_n)) - \mathbb{E}(L(\mathcal{Z})) \right| \right)_{\Omega_1}^* \rightarrow 0, \mathbb{P}_1 - a.s., \quad (2)$$

for a generic d_D -Borel measurable random variable \mathcal{Z} admitting a separable support $D_0 \subset D$. The use of a minimal measurable cover in (2) can be interpreted in other words: the maps of bounded Lipschitz distances of interest are converging to zero $\mathbb{P}_1 - a.s.*$ (see [15, p. 52] for more details). This notion of *a.s.** (or outer almost sure) convergence is of crucial importance: it is well known that, for sequences of nonmeasurable maps, the notion of almost everywhere (or even everywhere) convergence to zero is of no use - see [15, p. 52-53]. The so far most parcimonious set of assumptions upon (g_n) for a continuous mapping theorem to hold can be found in [4, p. 34] or [15, p. 67]. We shall here adapt this set of assumptions to naturally fit into (2). Let $D_0 \subset D$ be a Borel and separable set such that $\mu(D_0) = 1$. Let D_n be a sequence of Borel subsets of D , let (g_n) be a sequence of maps respectively from $\Omega_1 \times D_n$ to E and let g be Borel from D to E . Our assumption upon (g_n) is as follows.

(Hg) For \mathbb{P}_1 - almost all $\omega_1 \in \Omega_1$, the following property holds: for any $z \in D_0$ and for any sequence $z_n \in D_n$ converging to z , we have $g_n(\omega_1, z_n) \rightarrow g(z)$.

Here " \mathbb{P}_1 - almost all" has to be understood as: for all $\omega_1 \in A$, where $\mathbb{P}_1^*(A^C) = 0$.

1.3 A general continuous mapping theorem

Assuming that

$$\mathcal{Z}_n(\Omega) \subset D_n \text{ for each } n \in \mathbb{N}^*, \quad (3)$$

a satisfying formulation of a continuous mapping theorem must be

$$\left(\sup_{L \in BL_1(E)} \left| \mathbb{E}_{\Omega_2}^* \left(L(g_n(\cdot_1, \mathcal{Z}_n(\cdot_1, \cdot_2))) \right) - \mathbb{E} \left(L(g(\mathcal{Z})) \right) \right| \right)_{\Omega_1}^* \rightarrow 0, \mathbb{P}_1 - a.s. \quad (4)$$

If the involved maps on Ω_1 were measurable - hence \mathbb{P}_1 - a.s. equal to their measurable covers - then (4) would be equivalent to a pointwise convergence to zero of those maps, except on the complement of a set $\tilde{\Omega}_1$ fulfilling $\mathbb{P}_1(\tilde{\Omega}_1) = 1$. It would then be an immediate consequence of (2) by fixing $\omega_1 \in \tilde{\Omega}_1 \subset \mathcal{A}_1$ and invoking the usual continuous mapping theorem for the sequence $\mathcal{Z}_n(\omega_1, \cdot_2)$. But the measurability in (4) typically fails to hold, especially in empirical processes theory, and hence those simple arguments cannot be invoked. The approach that we take in this paper is to use arguments that are similar to that of Egorov's theorem : *a.s.** convergence is equivalent to almost uniform convergence - see [15, p. 53]. This approach succeeds, but at the price the following measurability assumption: we shall suppose that all the maps

$$\omega_1 \rightarrow \sup_{y \in D_n, d_D(y, z) < \delta} d_E(g_n(\omega_1, y), g(z)), \quad z \in D_0, \quad n \in \mathbb{N}, \quad \delta > 0, \quad (5)$$

are Borel measurable from $(\Omega_1, \mathcal{A}_1)$ to \mathbb{R} . This condition is far from being automatically satisfied. We will however show that it is the case in several statistical applications - see §2.

Theorem 1 *Assume that (2), (3), (Hg) and (5) hold. Then (4) holds.*

A straightforward consequence of Theorem 1 is the following continuous mapping theorem for "weak convergence in probability" which - despite of being weaker - is expected to have a wider range of statistical applications, yet still benefiting from a clear practical interpretation. It can be proved by a subsequence argument and holds under a less stringent version of (Hg), namely:

(Hg)' For fixed $\delta > 0$ we have, as $n \rightarrow \infty$:

$$\mathbb{P}_1^* \left(\left\{ \omega_1 \in \Omega_1, \exists z \in D_0, \forall \tau > 0, \sup_{y \in D_n, d_D(z', z) < \tau} d_E(g_n(\omega_1, y), g(z)) > \delta \right\} \right) \rightarrow 0.$$

Corollary 1 *Assume that*

$$\left(\sup_{L \in BL_1(D)} \left| \mathbb{E}_{\Omega_2}^* \left(L(\mathcal{Z}_n(\cdot_1, \cdot_2)) \right) - \mathbb{E} \left(L(\mathcal{Z}) \right) \right| \right)_{\Omega_1}^* \rightarrow 0, \text{ in } \mathbb{P}_1\text{-probability.}$$

Then, under the assumptions (Hg)' and (5) we have

$$\left(\sup_{L \in BL_1(E)} \left| \mathbb{E}_{\Omega_2}^* \left(L(g_n(\cdot_1, \mathcal{Z}_n(\cdot_1, \cdot_2))) \right) - \mathbb{E} \left(L(g(\mathcal{Z})) \right) \right| \right)_{\Omega_1}^* \rightarrow 0, \text{ in } \mathbb{P}_1\text{-probability.}$$

1.4 Application to the bootstrap multisample empirical process

We will now show the implications of Theorem 1 - and more specifically Corollary 1 - to the theory of bootstrap empirical processes. We also seize the opportunity to formulate some unsurprising extensions of the classical theory to the multisample empirical process. From now on we shall be focusing on the general problem of inference through the use of one or more ($k \geq 1$) i.i.d. samples drawn from k respective populations. The k samples may have different sizes n_1, \dots, n_k and are assumed to be mutually independent. To make the link with the probabilistic framework of §1.2, we shall consider a finite collection of probability spaces $(\mathfrak{X}_j, \mathcal{X}_j, P_{0,j})$, $j \in \llbracket 1, k \rrbracket$, and we shall denote by $\mathbf{P}_0 = (P_{0,1}, \dots, P_{0,k})$ the corresponding k -tuple of probability laws. We then take $(\Omega_1, \mathcal{A}_1, \mathbb{P}_1)$ as the canonical probability space formalizing our model, namely:

$$\Omega_1 := \prod_{j=1}^k \Omega_{1,j}, \mathcal{A}_1 := \bigotimes_{j=1}^k \mathcal{A}_{1,j}, \mathbb{P}_1 := \bigotimes_{j=1}^k \mathbb{P}_{1,j}, \text{ where}$$

$$\Omega_{1,j} := \mathfrak{X}_j^{\mathbb{N}^*}, \mathcal{A}_{1,j} := \mathcal{X}_j^{\otimes \mathbb{N}^*} \text{ and } \mathbb{P}_{1,j} := P_{0,j}^{\otimes \mathbb{N}^*}.$$

Now the probability space $(\Omega_1, \mathcal{A}_2, \mathbb{P}_2)$ generating the bootstrap mechanism is canonically defined as:

$$\Omega_2 := \prod_{j=1}^k \Omega_{2,j}, \mathcal{A}_2 := \bigotimes_{j=1}^k \mathcal{A}_{2,j}, \mathbb{P}_2 := \bigotimes_{j=1}^k Q_j, \text{ where}$$

$$\Omega_{2,j} := \prod_{n_j=1}^{\infty} \llbracket 1, n_j \rrbracket^{n_j}, \mathcal{A}_{2,j} = \mathcal{P}(\Omega_{2,j}), Q_j := \bigotimes_{n_j=1}^{\infty} Mult_{n_j},$$

and where $Mult_{n_j}$ stands for the multinomial distribution with n_j experiments, each of one having a vector probabilities $(n_j^{-1}, \dots, n_j^{-1})$. Now let us write any element of Ω_1 as

$$\omega_1 = (\omega_1^{(1)}, \dots, \omega_1^{(k)}), \text{ with each } \omega_1^{(j)} = (\omega_{i,j})_{i \geq 1} \in \Omega_{1,j},$$

and let us write any element of Ω_2 as

$$\omega_2 = (\omega_2^{(1)}, \dots, \omega_2^{(k)}) \text{ with each } \omega_2^{(j)} = (r_{n_j}^{(j)})_{n_j \geq 1}, \text{ and } r_{n_j}^{(j)} = (r_{1,n_j}^{(j)}, \dots, r_{n_j,n_j}^{(j)}) \in \llbracket 1, n_j \rrbracket^{n_j}.$$

With these notations we then define $X_{i,j}(\omega_1, \omega_2) := \omega_{i,j}$ - for each $j \in \llbracket 1, k \rrbracket$ and $i \in \mathbb{N}^*$ - so that each sequence $(X_{i,j})_{i \geq 1}$ is $P_{0,j}$ i.i.d. We then define the random weights of Efron's bootstrap procedure - see [6] - as follows:

$$W_{i,n_j}^{(j)}(\omega_1, \omega_2) := r_{i,n_j}^{(j)}.$$

For any k -tuple of non null integers $\mathbf{n} := (n_1, \dots, n_k)$ define the multivariate empirical measure and its bootstrap counterpart as the following maps on

$\Omega_1 \times \Omega_2$:

$$\mathbf{P}_{\mathbf{n}}(\omega_1, \omega_2) := \left(\frac{1}{n_j} \sum_{i=1}^{n_j} \delta_{X_{i,j}}(\omega_1, \omega_2) \right)_{j \leq k} =: \left(P_{n_1,1}(\omega_1^{(1)}, \omega_2^{(1)}), \dots, P_{n_k,k}(\omega_1^{(k)}, \omega_2^{(k)}) \right),$$

$$\hat{\mathbf{P}}_{\mathbf{n}}(\omega_1, \omega_2) := \left(\frac{1}{n_j} \sum_{i=1}^{n_j} W_{i,n_j}^{(j)}(\omega_1, \omega_2) \delta_{X_{i,j}(\omega_1, \omega_2)} \right)_{j \leq k} =: \left(\hat{P}_{n_1,1}(\omega_1^{(1)}, \omega_2^{(1)}), \dots, \hat{P}_{n_k,k}(\omega_1^{(k)}, \omega_2^{(k)}) \right),$$

where " δ_x " stands for the Dirac point mass at x . Note that each $P_{n_j,j}$ (resp. $\hat{P}_{n_j,j}$) depends on (ω_1, ω_2) through $\omega_1^{(j)}$ (resp. $(\omega_1^{(j)}, \omega_2^{(j)})$) only. We shall now consider, for each $j \in \llbracket 1, k \rrbracket$ a class of real valued functions \mathcal{F}_j that is $P_{0,j}$ Donsker, namely: there exists a tight, centered, \mathcal{F}_j -indexed Gaussian process \mathbb{G}_j such that

$$\mathbb{G}_{n_j,j} := \sqrt{n_j} \left(P_{n_j,j} - P_{0,j} \right) \rightarrow_{\mathcal{L}} \mathbb{G}_j, \text{ as } n_j \rightarrow \infty,$$

in the Banach space $(\ell^\infty(\mathcal{F}_j), \|\cdot\|_{\mathcal{F}_j})$ of all real bounded functions on \mathcal{F}_j endowed with the sup norm. We shall also generically use the symbol $\|\phi\|_A$ for the sup norm of a real valued map ϕ over an index set A . The limit process \mathbb{G}_j induces a Borel probability law on the subspace $\mathcal{C}(\mathcal{F}_j, \|\cdot\|_{P_{0,j,2}})$ of functions on \mathcal{F}_j that are continuous with respect to $\|\cdot\|_{P_{0,j,2}}$ - here we use the general notation $\|\cdot\|_{Q,r}$ for the $L^r(Q)$ norm on a probability space. In order to unburden the notations, we shall simply write

$$\begin{aligned} \mathbb{D}_j &:= (\ell^\infty(\mathcal{F}_j), \|\cdot\|_{\mathcal{F}_j}), \mathbb{D}_{0,j} := (\mathcal{C}(\mathcal{F}_j, \|\cdot\|_{P_{0,j}}), \|\cdot\|_{\mathcal{F}_j}), \text{ and} \\ \mathbb{D} &:= \mathbb{D}_1 \times \dots \times \mathbb{D}_k, \mathbb{D}_0 := \mathbb{D}_{0,1} \times \dots \times \mathbb{D}_{0,k}, \end{aligned}$$

for their natural cartesian product as Banach spaces. By independence we have

$$(\mathbb{G}_{n_1,1}, \dots, \mathbb{G}_{n_k,k}) \rightarrow_{\mathcal{L}} (\mathbb{G}_1, \dots, \mathbb{G}_k), \text{ as } \mathbf{n} \rightarrow \infty, \quad (6)$$

where the family $(\mathbb{G}_1, \dots, \mathbb{G}_k)$ is mutually independent, and where " $\mathbf{n} \rightarrow \infty$ " stands for " $\min\{n_1, \dots, n_k\} \rightarrow \infty$ ". At the price of a small technicality, the following result holds as a consequence of the bootstrap central limit theorem in probability, writing

$$\hat{\mathbb{G}}_{n_j,j}(\omega) := \sqrt{n_j} \left(\hat{P}_{n_j,j}(\omega_1^{(j)}, \omega_2^{(j)}) - P_{n_j,j}(\omega_1^{(j)}, \omega_2^{(j)}) \right), \omega \in \Omega.$$

Theorem 2 *We have, as $\mathbf{n} \rightarrow \infty$*

$$\left(\sup_{L \in BL_1(\mathbb{D})} \left| \mathbb{E}_{\Omega_2} \left(L \left(\hat{\mathbb{G}}_{n_1,1}, \dots, \hat{\mathbb{G}}_{n_k,k} \right) \right) - \mathbb{E} \left(L \left(\mathbb{G}_1, \dots, \mathbb{G}_k \right) \right) \right| \right)_{\Omega_1}^* \rightarrow 0,$$

in \mathbb{P}_1 -probability, where \mathbb{E} stands for the expectation on any probability space on which $(\mathbb{G}_1, \dots, \mathbb{G}_k)$ is mutually independent.

Note that \mathbb{E}_{Ω_2} is not an outer partial expectation but rather a true partial expectation.

Remark: The same arguments hold for the general case of the *exchangeable bootstrap*, when the laws $Mult_{n_j}$ are replaced by exchangeable laws on $[0, +\infty[^{n_j}$ - see [12]. A version of Theorem 2 holds under the assumption that each \mathcal{F}_j is $\mathbf{P}_{0,j}$ -Donsker and that the triangular arrays $(W_{n_j,i}^{(j)})_{n_j \geq 1, i \in [1, n_j]}$ satisfy the conditions of Theorem 2.2 in [12]. We omit to explicitly state it for concision only: otherwise we would have to recall in details the mathematical framework of the exchangeable bootstrap Donsker theorem.

It is known that Donsker theorems have applications in statistics through the functional delta-method, when the statistic is a function of the empirical measure that is Hadamard differentiable tangentially to a suitable subspace. For the multisample version (6), the theoretical framework of the functional delta method can be formulated as follows: we assume that the multisample statistic has expression $\Phi(\mathbf{P}_n)$, where Φ takes values in a Banach space $(D, \|\cdot\|)$ and is defined on a subset $\tilde{\mathbb{D}} \subset \mathbb{D}$ containing \mathbf{P}_0 and all the possible values of \mathbf{P}_n , $\mathbf{n} \in \mathbb{N}^k$. That map Φ is assumed to be Hadamard differentiable at $\mathbf{P}_0 := (P_{0,1}, \dots, P_{0,k})$ tangentially to \mathbb{D}_0 , namely:

($H\Phi$) There exist continuous linear maps $d\Phi_{\mathbf{P}_0}^{(j)}$ from \mathbb{D}_j to D such that,

for each $\boldsymbol{\psi} = (\psi_1, \dots, \psi_k) \in \prod_{j=1}^k \mathbb{D}_{0,j}$ and for each sequence

$\mathbf{P}^{(n)} = \mathbf{P}_0 + \epsilon_n \mathbf{Q}_n \in \tilde{\mathbb{D}}$ fulfilling $\epsilon_n \rightarrow 0$ and $\mathbf{Q}_n \rightarrow \boldsymbol{\psi}$ in \mathbb{D} we have

$$\lim_{n \rightarrow \infty} \epsilon_n^{-1} \{ \Phi(\mathbf{P}^{(n)}) - \Phi(\mathbf{P}_0) \} = \sum_{j=1}^k d\Phi_{\mathbf{P}_0}^{(j)}(\psi_j).$$

The notion of Hadamard differentiability tangentially to a suitable (more regular) subspace is often more useful - since less stringent - than the notion of Hadamard differentiability alone (i.e. when $\mathbb{D}_0 = \mathbb{D}$) and is known to open a wider spectrum of applications. In a series of recent works, Beutner and Zähle [2, 3, 1] showed that a more refined notion of *quasi Hadamard differentiability* also widens the scope of statistical applications. In our next result, we shall use the notation $\boldsymbol{\alpha}_n = (\alpha_{1,n}, \dots, \alpha_{k,n}) := N^{-1} \mathbf{n}$ where $N := n_1 + \dots + n_k$.

Theorem 3 *Assume that each \mathcal{F}_j is $P_{0,j}$ -Donsker and that ($H\Phi$) holds. Then, when $\mathbf{n} \rightarrow \infty$ together with*

$$\boldsymbol{\alpha}_n \rightarrow \boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_k) \in (0, 1]^k, \quad (7)$$

we have

$$\sqrt{N} \left(\Phi(\mathbf{P}_n) - \Phi(\mathbf{P}_0) \right) \rightarrow_{\mathcal{L}} \sum_{j=1}^k \alpha_j^{-1/2} d\Phi_{\mathbf{P}_0}^{(j)}(\mathbb{G}_j). \quad (8)$$

With small efforts, it is possible to prove the following bootstrap version of Theorem 3.

Theorem 4 *Under the assumptions of Theorem 3 we have*

$$\sup_{L \in BL_1(D)} \left| \mathbb{E}_{\Omega_2} \left(L \left(\sqrt{N} \left(\Phi(\hat{\mathbf{P}}_n) - \Phi(\mathbf{P}_n) \right) \right) \right) - \mathbb{E} \left(L \left(\sum_{j=1}^k \alpha_j^{-1/2} d\Phi_{\mathbf{P}_0}(\mathbb{G}_j) \right) \right) \right|_{\Omega_1}^* \rightarrow 0,$$

in \mathbb{P}_1 -probability.

Remark: From now on we shall only focus on bootstrap results that hold in (outer) probability, and not (outer) almost surely. Our choice is motivated as follows: it is well known that the delta method for the bootstrap almost surely requires that Φ satisfies a stronger form of Hadamard differentiability - see e.g. [15, Assertion (3.9.12), p. 379]. This stronger form is less frequently met in practice - for example this is not the case for the inverse map on distribution functions. On the other hand, bootstrap results in probability are more than sufficient for statistical purposes.

An interesting addition that Theorem 1 brings to Theorem 3 is the possibility to prove bootstrap convergences in probability when $\sqrt{N}(\Phi(\hat{\mathbf{P}}_n) - \Phi(\mathbf{P}_n))$ is replaced by $\hat{g}(\sqrt{N}(\Phi(\hat{\mathbf{P}}_n) - \Phi(\mathbf{P}_n)))$, where \hat{g} is an "estimated map" - typically a renormalization by an estimation of the variance function of the limit process - that is defined through the non bootstrap observations only. Such a setup naturally arises in empirical likelihood methods for the estimation of a function/trajectory - see below. We can formalise this framework by a net $(g_n)_{n \in \mathbb{N}^k}$ of maps from $\Omega_1 \times D$ to a Banach space E . A direct application of Corollary 1 in conjunction with Theorem 4 is the following.

Corollary 2 *Let $(\mathbf{n}(n))_{n \in \mathbb{N}^*}$ be a \mathbb{N}^k valued sequence. Write $N(n) := \sum_{j=1}^k n_j(n)$, and assume that*

$$\min\{n_1(n), \dots, n_k(n)\} \rightarrow \infty, \text{ and } N(n)^{-1}(n_1(n), \dots, n_k(n)) \rightarrow (\alpha_1, \dots, \alpha_k) \in (0, 1]^k. \quad (9)$$

Now assume that $(g_n) = (g_{\mathbf{n}(n)})_{n \in \mathbb{N}^*}$ satisfies $(Hg)'$ and (5). Then we have, as $n \rightarrow \infty$

$$\begin{aligned} \sup_{L \in BL_1(E)} \left| \mathbb{E}_{\Omega_2} \left(L \left(g_{\mathbf{n}(n)} \left(\sqrt{N(n)} \left(\phi(\hat{\mathbf{P}}_{\mathbf{n}(n)}) - \phi(\mathbf{P}_{\mathbf{n}(n)}) \right) \right) \right) \right) \right. \\ \left. - \mathbb{E} \left(L \left(g \left(\sum_{j=1}^k \alpha_j^{-1/2} d\Phi_{\mathbf{P}_0}^{(j)}(\mathbb{G}_j) \right) \right) \right) \right|_{\Omega_1}^* \rightarrow 0, \text{ in } \mathbb{P}_1 \text{ - probability.} \end{aligned}$$

In the following section, we exhibit statistical applications where the net (g_n) meets the requirements $(Hg)'$ and (5).

Note : in Corollary 2 we drop the formalism " $\mathbf{n} \rightarrow \infty$ and (7)" to an indexation by n coupled with (9). This change is purely formal and is intended only towards a coherence with the notations of Theorem 1.

2 Statistical applications: empirical likelihood for a functional parameter in explicit plug-in estimation

In [16], Varron proved the consistency of a general empirical likelihood method for a functional parameter $(\boldsymbol{\theta}_0(t))_{t \in \mathfrak{T}}$ which can be expressed as $\Phi(\mathbf{P}_0)$, where the explicitly known map Φ takes its values in the Banach space $D := \ell^\infty(\mathfrak{T})^p$ of all \mathbb{R}^p valued and bounded functions on \mathfrak{T} , endowed with the multivariate sup norm

$$\|\boldsymbol{\phi}\|_{\mathfrak{T},p} := \sup_{t \in \mathfrak{T}} \|\boldsymbol{\phi}(t)\|_p, \quad \boldsymbol{\phi} \in \ell^\infty(\mathfrak{T})^p, \quad \text{with } \|\mathbf{u}\|_p^2 := \mathbf{u}^\top \mathbf{u}, \quad \mathbf{u} \in \mathbb{R}^p,$$

where " \mathbf{u}^\top " stands for the transposition of \mathbf{u} (unless otherwise specified, each element $\mathbf{u} \in \mathbb{R}^p$ will be represented by a column vector). This framework encompasses several applications, in particular for survival analysis - see [16, Section 2]. Let us write \mathbb{I}_p for the $p - p$ identity matrix and GL_p for the group of all invertible real $p - p$ matrices. Varron proved (see [16, Theorem 1]) that one can build asymptotic confidence tubes, for which the calibration is determined by the law of a \mathfrak{T} -indexed, \mathbb{R}^+ -valued stochastic process \mathcal{W} that takes its values in $[0, +\infty)$. Unfortunately whereas the formal expression of \mathcal{W} is known, its dependency upon \mathbf{P}_0 makes it practically non implementable. Hence, a critical question to make those confidence tubes fully implementable in practice is to estimate this limit law by a bootstrap procedure.

In this section we will place ourselves in the set of assumptions of [16, Theorem 1]. We will then propose a bootstrap procedure and use Corollary 1 to establish its consistency.

A careful look at the arguments of the proof of [16, Theorem 1] shows that \mathcal{W} is the weak limit, as $\mathbf{n} \rightarrow \infty$ together with (7), of the \mathfrak{T} -indexed processes

$$\mathcal{W}_{\mathbf{n}} : t \rightarrow \sqrt{N}(\Phi(\mathbf{P}_{\mathbf{n}}) - \Phi(\mathbf{P}_0))(t)^\top M_{\mathbf{n}}(t) \sqrt{N}(\Phi(\mathbf{P}_{\mathbf{n}}) - \Phi(\mathbf{P}_0))(t), \quad \text{where}$$

$$M_{\mathbf{n}}(t) := S_{\mathbf{n}}(t)^{-1} \mathbb{1}_{\text{GL}_p}(S_{\mathbf{n}}(t)) + \mathbb{I}_p \mathbb{1}_{\text{GL}_p^c}(S_{\mathbf{n}}(t)), \quad \text{and}$$

$$S_{\mathbf{n}}(t) := \frac{1}{N} \sum_{j=1}^k \alpha_{j,\mathbf{n}}^{-1} \sum_{i=1}^{n_j} \hat{m}_{i,j}(t) \hat{m}_{i,j}(t)^\top,$$

$$\hat{m}_{i,j}(t) := d\Phi_{\mathbf{P}_{\mathbf{n}}}^{(j)}(\delta_{X_{i,j}} - P_{n_{j,j}})(t), \quad t \in \mathfrak{T},$$

under the assumption that the differentiability of Φ can be extended to the values of $\mathbf{P}_{\mathbf{n}}$ in a weaker (Gâteaux) sense, and with an additional continuity property at \mathbf{P}_0 - see assumptions (HT2) and (HT3) in [16]. The weak convergence of $\mathcal{W}_{\mathbf{n}}$ to \mathcal{W} can be seen as a combination of Theorem 3 with the consistency of $S_{\mathbf{n}}(\cdot)$ in the sense that

$$\|S_{\mathbf{n}} - V\|_{\mathfrak{T},p^2} \rightarrow 0, \quad \text{in outer probability, where } V(t) := \sum_{j=1}^k \alpha_j^{-2} V_j(t),$$

$$V_j(t) := \text{Var}(m_{1,j}(t)), \quad m_{i,j}(t) := [d\Phi_{\mathbf{P}_0}^{(j)}(\delta_{X_{i,j}} - P_{0,j})], \quad t \in \mathfrak{T},$$

and under the assumption that $\inf\{\det(V(t)), t \in \mathfrak{T}\} > 0$. As one can then see, \mathcal{W} can be expressed as

$$\mathcal{W} = \left\{ t \rightarrow \tilde{\mathcal{W}}(t)^\top V^{-1}(t) \tilde{\mathcal{W}}(t) \right\}, \quad \text{where}$$

$$\tilde{\mathcal{W}} := \sum_{j=1}^k \alpha_j^{-1/2} d\Phi_{\mathbf{P}_0}(\mathbb{G}_j).$$

Hence a possible bootstrap method to estimate the law of \mathcal{W} could be to use $M_{\mathbf{n}}(\cdot)$ as an estimator of $V(\cdot)^{-1}$, then to bootstrap $\sqrt{N}(\Phi(\mathbf{P}_{\mathbf{n}}) - \Phi(\mathbf{P}_0))$, and finally combine them to use the bootstrap law (conditionnally to the original data) of

$$\mathcal{W}_{\mathbf{n}}^{\text{Boot}} := \left\{ t \rightarrow \sqrt{N}(\Phi(\hat{\mathbf{P}}_{\mathbf{n}}) - \Phi(\mathbf{P}_{\mathbf{n}}))(t)^\top M_{\mathbf{n}}(t) \sqrt{N}(\Phi(\hat{\mathbf{P}}_{\mathbf{n}}) - \Phi(\mathbf{P}_{\mathbf{n}}))(t) \right\},$$

as an estimation of the law of \mathcal{W} . We prove the consistency of this bootstrap procedure through Corollary 1 by proceeding as follows: we first fix $\boldsymbol{\alpha} \in (0, 1]^k$ and we take an arbitrary sequence $(\mathbf{n}(n))_{n \geq 1}$ such that $N(n)^{-1} \mathbf{n}(n) \rightarrow \boldsymbol{\alpha}$ as $n \rightarrow \infty$. We then investigate the validity of $(Hg)'$ and (5) for the net $g_{\mathbf{n}} := g_{\mathbf{n}(n)}$, where

$$g_{\mathbf{n}} : \Omega_1 \times D \rightarrow E$$

$$(\omega_1, \phi) \rightarrow \left\{ t \rightarrow \phi(t)^\top M_{\mathbf{n}}(t) (\omega_1) \phi(t) \right\}.$$

Note that, for fixed t and \mathbf{n} , the map $M_{\mathbf{n}}(t)$ is in fact defined on $\Omega_1 \times \Omega_2$. We allow ourselves to use the notation $M_{\mathbf{n}}(t)(\omega_1)$ since $M_{\mathbf{n}}(t)$ is constant with respect to ω_2 .

It is clear that $(Hg)'$ is guaranteed for the choice of limit function

$$g : \phi \rightarrow \left\{ t \rightarrow \phi(t)^\top V^{-1}(t) \phi(t) \right\},$$

whenever the following consistency holds

$$\left(\|M_{\mathbf{n}(n)} - V^{-1}\|_{\mathfrak{T}, p^2} \right)_{\Omega_1}^* \rightarrow 0, \quad \text{in } \mathbb{P}_1\text{-probability,}$$

which is true under the assumptions of [16, Theorem 1] - see p. 114 -155 of the just cited article. Let us finally focus on the measurability assumption (5) and discuss on its possible validity. As we will show - see Proposition 1 below - the minimal measurability requirement that

$$d\Phi_{\mathbf{P}_{\mathbf{n}}}^{(j)}(\delta_{X_{i,j}} - P_{n_j,j})(t) \text{ is Borel measurable for each } t \in \mathfrak{T}, \quad \mathbf{n} \in \mathbb{N}^{*k}$$

is already useful when the possible values of $V_{\mathbf{n}}$ and $\Phi(\mathbf{P}_{\mathbf{n}}) - \Phi(\mathbf{P}_0)$ are trajectories that all benefit from a *common separability property*. Let $\mathfrak{T}_0 \subset \mathfrak{T}$ be

countable. A bounded map ϕ from \mathfrak{T} to \mathbb{R}^p is said to be \mathfrak{T}_0 separable if, for each $t \in \mathfrak{T}$ there exists a sequence (t_m) of elements of \mathfrak{T}_0 such that $\phi(t_m) \rightarrow \phi(t)$ as $m \rightarrow \infty$. We shall denote by $\ell_{\mathfrak{T}_0, p}$ the (closed) subspace of $\ell^\infty(\mathfrak{T})^p$ consisting of all such maps.

Proposition 1 *Assume that $V \in \ell_{\mathfrak{T}_0, p^2}$ and that each $d\Phi_{\mathbf{P}_n}^{(j)}(\delta_{X_{i,j}} - P_{n_j, j})$ takes its values in $\ell_{\mathfrak{T}_0, p}$. Assume that Φ takes its values in $\ell_{\mathfrak{T}_0, p}$. Then the maps $g_{\mathbf{n}(n)}$ satisfy (5) for the choices of $D = D_n := \ell_{\mathfrak{T}_0, p}$ and $E := \ell_{\mathfrak{T}_0, 1}$. As a consequence if, in addition, all the assumptions of [16, Theorem 1] are fulfilled, then we have as $n \rightarrow \infty$:*

$$\left(\sup_{L \in BL_1(\ell_{\mathfrak{T}_0, 1})} \left| \mathbb{E}_{\Omega_2} \left(L(\mathcal{W}_{\mathbf{n}(n)}^{Boot}) \right) - \mathbb{E} \left(L(\mathcal{W}) \right) \right| \right)_{\Omega_1}^* \rightarrow 0, \text{ in } \mathbb{P}_1\text{-probability.}$$

A typical situation is when $\mathfrak{T} \subset \mathbb{R}$ is an interval, $\mathfrak{T}_0 = \mathfrak{T} \cap \mathbb{Q}$ and the observed trajectories are almost surely right-continuous. This is for example the case of all the examples treated in [16, Section 2]. Hence, a consequence of Proposition 1 is that each of those examples of statistical applications can benefit from a consistent bootstrap approximation of \mathcal{W} . Consequently, simultaneous confidence regions can be fully implemented in practice.

3 Proofs

3.1 Proof of theorem 1

3.1.1 Some preliminary properties

Let us first deduce some regularity properties of \mathcal{Z} from assumption (2). Because $\mathbb{P}_1(\emptyset) = 0$, there must exist at least one $\omega_1 \in \Omega_1$ such that

$$\sup_{L \in BL_1(D)} \left| \mathbb{E}_{\Omega_2}^* \left(L(\mathcal{Z}_n(\omega_1, \cdot)) \right) - \mathbb{E} \left(L(\mathcal{Z}) \right) \right| \rightarrow 0, \text{ as } n \rightarrow \infty. \quad (10)$$

As a consequence, \mathcal{Z} is the limit (in the sense of weak convergence) of at least one sequence of random elements on $(\Omega_2, \mathcal{A}_2)$ that respectively take their values in D_n . We can then use the same arguments as those of the beginning of the proof of [15, Theorem 1.11.1, p. 67] to state that, without loss of generality, one can assume that D_0 has the following properties:

$$\text{For any } z \in D_0, \text{ there exists a sequence } z_n \in D_n \text{ such that } z_n \rightarrow z. \quad (11)$$

$$\text{The restriction } g|_{D_0} \text{ is } d_D\text{-continuous.} \quad (12)$$

Let us first point out that (2) can be formulated as

$$(\rho(X_n, X))_{\Omega_1}^* \rightarrow 0, \mathbb{P}_1\text{-a.s., where} \quad (13)$$

$$X_n : \Omega_1 \mapsto \ell^\infty(BL_1(D))$$

$$\begin{aligned}
\omega_1 &\rightarrow \left\{ L \rightarrow \mathbb{E}_{\Omega_2}^* \left(L(\mathcal{Z}_n(\omega_1, \cdot_2)) \right) \right\}, \\
X : \Omega_1 &\mapsto \ell^\infty(BL_1(D)) \\
\omega_1 &\rightarrow \left\{ L \rightarrow \mathbb{E} \left(L(\mathcal{Z}) \right) \right\},
\end{aligned}$$

and where ρ is the distance induced by the sup norm $\|\cdot\|_{BL_1(D)}$. As a consequence, Egorov's theorem applies (see, e.g. [15, Lemma 1.9.2, p. 53]): for each $\epsilon > 0$ there exists a set $\Omega_{1,\epsilon} \in \mathcal{A}_1$ such that

$$\begin{aligned}
\mathbb{P}_1(\Omega_{1,\epsilon}) &\geq 1 - \epsilon, \text{ and} \\
\sup_{\omega_1 \in \Omega_{1,\epsilon}} \sup_{L \in BL_1(D)} \left| \mathbb{E}_{\Omega_2}^* \left(L(\mathcal{Z}_n(\omega_1, \cdot_2)) \right) - \mathbb{E} \left(L(\mathcal{Z}) \right) \right| &\rightarrow 0. \quad (14)
\end{aligned}$$

From now on we shall fix $\epsilon > 0$ and we shall show the existence of $\Omega'_{1,\epsilon} \in \mathcal{A}_1$ such that $\mathbb{P}(\Omega'_{1,\epsilon}) \geq 1 - 2\epsilon$ and

$$\sup_{\omega_1 \in \Omega'_{1,\epsilon}} \sup_{L \in BL_1(E)} \left| \mathbb{E}_{\Omega_2}^* \left(L(g_n(\omega_1, \mathcal{Z}_n(\omega_1, \cdot_2))) \right) - \mathbb{E} \left(L(g(\mathcal{Z})) \right) \right| \rightarrow 0, \quad (15)$$

which would prove (4) since $\epsilon > 0$ is arbitrary (and by a use of the converse part of Egorov's theorem). Let us now choose $\Omega_{1,\epsilon} \in \mathcal{A}_1$ such that (14) holds. The latter assertion can be seen as a uniform weak convergence of a $\Omega_{1,\epsilon}$ -indexed family of random elements on $(\Omega_2, \mathcal{A}_2, \mathbb{P}_2)$. That uniformity is straightforwardly carried over some usual basic properties of weak convergence, as we will show in our two first lemmas. From now on ∂B will denote the topological frontier of a set $B \subset D$.

Lemma 1 *For any closed set F of (D, d_D) we have:*

$$\overline{\lim}_{n \rightarrow \infty} \sup_{\omega_1 \in \Omega_{1,\epsilon}} \mathbb{P}_{\Omega_2}^* (\mathcal{Z}_n(\omega_1, \cdot_2) \in F) \leq \mathbb{P}(\mathcal{Z} \in F). \quad (16)$$

As a consequence, for any $B \subset D$ fulfilling $\mathbb{P}(\mathcal{Z} \in \partial B) = 0$ we have

$$\lim_{n \rightarrow \infty} \sup_{\omega_1 \in \Omega_{1,\epsilon}} \left| \mathbb{P}_{\Omega_2}^* (\mathcal{Z}_n(\omega_1, \cdot_2) \in B) - \mathbb{P}(\mathcal{Z} \in B) \right| = 0. \quad (17)$$

Proof: Fix $\delta > 0$ and approximate $\mathbb{1}_F$ from above by a bounded Lipschitz function L_δ fulfilling $\mathbb{E}(L_\delta(\mathcal{Z})) \leq \mathbb{E}(\mathbb{1}_F(\mathcal{Z})) + \delta$, which is possible since \mathcal{Z} induces a Borel measure on D . Then use (14) to obtain

$$\begin{aligned}
&\overline{\lim}_{n \rightarrow \infty} \sup_{\omega_1 \in \Omega_{1,\epsilon}} \mathbb{E}_{\Omega_2}^* \left(\mathbb{1}_F(\mathcal{Z}_n(\omega_1, \cdot_2)) \right) \\
&\leq \overline{\lim}_{n \rightarrow \infty} \sup_{\omega_1 \in \Omega_{1,\epsilon}} \mathbb{E}_{\Omega_2}^* \left(L_\delta(\mathcal{Z}_n(\omega_1, \cdot_2)) \right) \\
&\leq \mathbb{E} \left(L_\delta(\mathcal{Z}) \right)
\end{aligned}$$

$$\leq \mathbb{E}(\mathbf{1}_F(\mathcal{Z})) + \delta.$$

This concludes the proof, since δ is arbitrary. \square

We shall generically denote by $B(z, \delta)$ the open ball with center z and radius δ in the metric space (D, d_D) or (E, d_E) - depending on the context. We shall also use the generic notation

$$K^\delta := \bigcup_{z \in K} B(z, \delta).$$

Lemma 2 *For any $\epsilon_1 > 0$ there exists a compact set $K_{\epsilon_1} \subset D_0$ such that*

$$\forall \delta_1 > 0, \overline{\lim}_{n \rightarrow \infty} \sup_{\omega_1 \in \Omega_{1,\epsilon}} \mathbb{P}_{\Omega_2}^* \left(\mathcal{Z}_n(\omega_1, \cdot_2) \notin K_{\epsilon_1}^{\delta_1} \right) < \epsilon_1, \quad (18)$$

$$\mathbb{P}(\mathcal{Z} \notin K_{\epsilon_1}) < \epsilon_1. \quad (19)$$

Proof: Fix $\epsilon_1 > 0$. Since \mathcal{Z} induces a tight probability measure on D_0 , one can find K_{ϵ_1} such that (19) holds. Now, for fixed δ_1 , apply Lemma 1 to the closed set $F = D \setminus K_{\epsilon_1}^{\delta_1}$. \square

3.1.2 Construction of $\Omega'_{1,\epsilon}$

From now on, we shall consider a family $(K_{\epsilon_1})_{\epsilon_1 > 0}$ of compacts as exhibited in Lemma 2. Our next lemma is the crucial step of the proof of Theorem 1: by using arguments inspired from Egorov's theorem, we construct the set $\Omega'_{1,\epsilon}$ on which the sequence (g_n) satisfies a crucial "uniform local convergence" property. That property will be our main tool to establish (15). Note that we cannot directly use Egorov's theorem since it seems impossible to express (Hg) by a convergence in a semimetric space.

Lemma 3 *There exists $\Omega'_{1,\epsilon} \in \mathcal{A}_1$ such that $\Omega'_{1,\epsilon} \subset \Omega_{1,\epsilon}$, $\mathbb{P}_1(\Omega'_{1,\epsilon}) \geq 1 - 2\epsilon$ and such that*

$$\begin{aligned} & \forall \delta > 0, \forall \epsilon_1 > 0, \exists \tau > 0, \exists n_0, \forall n \geq n_0, \\ & \forall z \in K_{\epsilon_1}, \forall y \in B(z, \tau) \cap D_n, \sup_{\omega_1 \in \Omega'_{1,\epsilon}} d_E(g_n(y), g(z)) < \delta. \end{aligned} \quad (20)$$

Proof: Let \tilde{D}_0 be a dense and countable subset of D_0 . We first notice that (Hg) has the following consequence :

$$\begin{aligned} & \mathbb{P} \left(\bigcap_{k_1 \in \mathbb{N}^*} \bigcap_{k_2 \in \mathbb{N}^*} \bigcup_{k \in \mathbb{N}^*} \bigcup_{n_0 \in \mathbb{N}^*} A_{k_1, k_2, k, n_0} \right) = 1, \text{ where} \\ & A_{k_1, k_2, k, n_0} := \left\{ \omega_1 \in \Omega_1, \forall n \geq n_0, \forall z \in K_{k_2}^{k-1} \cap \tilde{D}_0, \right. \end{aligned} \quad (21)$$

$$\forall y \in B(z, k^{-1}) \cap D_n, d_E(g_n(\omega_1, y), g(z)) < k_1^{-1}\}.$$

Indeed, if (21) did not hold, then there would exist $\omega_1 \in \Omega_1$, $k_1 \in \mathbb{N}^*$, $k_2 \in \mathbb{N}^*$ as well as two sequences $z_{n_k} \in K_{k_2}^{k_1^{-1}} \cap \tilde{D}_0$ and $y_{n_k} \in D_{n_k}$ such that

$$d_D(y_{n_k}, z_{n_k}) \leq k^{-1}, \text{ but } d_E(g_{n_k}(\omega_1, y_{n_k}), g(z_{n_k})) \geq k_1^{-1}, \text{ for all } k \geq 1. \quad (22)$$

Because $K_{k_2}^{k_1^{-1}}$ is compact, one would be able to extract a subsequence $(z_{n'_k})$ of (z_{n_k}) converging to some $z \in K_{k_2}^{k_1^{-1}} \subset D_0$. In addition, since the corresponding subsequence $(y_{n'_k})$ would in turn converge to z , assumption (Hg) would imply that $g_{n'_k}(\omega_1, y_{n'_k}) \rightarrow g(z)$. But the continuity of $g|_{D_0}$ also entails that $g(z_{n'_k}) \rightarrow g(z)$, which would contradict (22). Hence (21) holds. Also notice that (5) ensures that each A_{k_1, k_2, k, n_0} belongs to \mathcal{A}_1 . Now fix $(k_1, k_2) \in \mathbb{N}^{*2}$. By (21) one has

$$0 = \mathbb{P}_1 \left(\bigcap_{k \in \mathbb{N}^*} \bigcap_{n_0 \in \mathbb{N}^*} A_{k_1, k_2, k, n_0}^C \right).$$

Moreover we have $A_{k_1, k_2, k, n_0} \supset A_{k_1, k_2, k', n_0}$ for $k \leq k'_3$ and for each $n_0 \geq 1$. As a consequence the sequence of elements of \mathcal{A}_1

$$\left(\bigcap_{n_0 \in \mathbb{N}^*} A_{k_1, k_2, k, n_0}^C \right)_{k \in \mathbb{N}^*}$$

is decreasing (in the sense of inclusion). Hence one can choose $\mathbf{k}(k_1, k_2)$ such that

$$\mathbb{P}_1 \left(\bigcap_{n_0 \in \mathbb{N}^*} A_{k_1, k_2, \mathbf{k}(k_1, k_2), n_0}^C \right) \leq \epsilon 2^{-k_1 - k_2 - 1}. \quad (23)$$

Now $A_{k_1, k_2, \mathbf{k}(k_1, k_2), n_0}^C$ is also decreasing in n_0 . Hence (23) implies the existence of $\mathbf{n}_0(k_1, k_2)$ satisfying.

$$\mathbb{P}_1 \left(A_{k_1, k_2, \mathbf{k}(k_1, k_2), \mathbf{n}_0(k_1, k_2)}^C \right) \leq \epsilon 2^{-k_1 - k_2}.$$

Hence, by the union bound one has

$$\mathbb{P}_1 \left(\bigcup_{k_1, k_2 \in \mathbb{N}^*} A_{k_1, k_2, \mathbf{k}(k_1, k_2), \mathbf{n}_0(k_1, k_2)}^C \right) \leq \sum_{k_1, k_2 \in \mathbb{N}^*} \epsilon 2^{-k_1 - k_2} = \epsilon.$$

Recalling that $\mathbb{P}_1(\Omega_{1, \epsilon}) \geq 1 - \epsilon$ we readily conclude that

$$\mathbb{P}_1 \left(\Omega_{1, \epsilon}^C \cup \bigcup_{k_1, k_2 \in \mathbb{N}^*} A_{k_1, k_2, \mathbf{k}(k_1, k_2), \mathbf{n}_0(k_1, k_2)}^C \right) \leq 2\epsilon.$$

We then choose $\Omega'_{1, \epsilon}$ as the complement of the latter, namely

$$\Omega'_{1, \epsilon} := \Omega_{1, \epsilon} \cap \bigcap_{k_1, k_2 \in \mathbb{N}^*} A_{k_1, k_2, \mathbf{k}(k_1, k_2), \mathbf{n}_0(k_1, k_2)} \subset \bigcap_{k_1, k_2 \in \mathbb{N}^*} A_{k_1, k_2, \mathbf{k}(k_1, k_2), \mathbf{n}_0(k_1, k_2)}.$$

This inclusion is then read as

$$\begin{aligned}
& \forall \omega_1 \in \Omega'_{1,\epsilon}, \forall k_1 \geq 1, \forall k_2 \geq 1, \forall n \geq \mathbf{n}_0(k_1, k_2), \forall z \in K_{k_2}^{\mathbf{k}(k_1, k_2)^{-1}} \cap \tilde{D}_0, \\
& \forall y \in B(z, \mathbf{k}(k_1, k_2)^{-1}), d_E(g_n(y), g(z)) < k_1^{-1}, \text{ or equivalently} \quad (24) \\
& \forall k_1 \geq 1, \forall k_2 \geq 1, \forall n \geq \mathbf{n}_0(k_1, k_2), \forall z \in K_{k_2}^{\mathbf{k}(k_1, k_2)^{-1}} \cap \tilde{D}_0, \forall y \in B(z, \mathbf{k}(k_1, k_2)^{-1}) \\
& \forall \omega_1 \in \Omega'_{1,\epsilon}, d_E(g_n(y), g(z)) < k_1^{-1},
\end{aligned}$$

by moving the first term in the chain of "∀" to the last position in that chain. As a consequence we have

$$\begin{aligned}
& \forall \delta > 0, \forall \epsilon_1 > 0, \exists \tau > 0, \exists n_0, \forall n \geq n_0, \\
& \forall z \in K_{\epsilon_1}^\tau \cap \tilde{D}_0, \forall y \in B(z, \tau) \cap D_n, \sup_{\omega_1 \in \Omega'_{1,\epsilon}} d_E(g_n(y), g(z)) < \delta. \quad (25)
\end{aligned}$$

The proof is concluded as follows: take $\delta > 0$ and $\epsilon_1 > 0$ and choose $\tau > 0$ and n_0 as in (25). Now fix $n \geq n_0$, $z \in K_{\epsilon_1}$ and $y \in B(z, \tau/2) \cap D_n$. Next, choose (by (11)) a sequence (z_m) in D_0 converging to z . Then, for all m large enough one has $d_D(y, z_m) < \tau$ together with $z_m \in K_{\epsilon_1}^\tau \cap \tilde{D}_0$, which are combined to obtain, by (25) :

$$\sup_{\omega_1 \in \Omega'_{1,\epsilon}} d_E(g_n(\omega_1, y), g(z_m)) < \delta.$$

But, since $(z_m) \rightarrow z$ and $g|_{D_0}$ is continuous, one has $d_E(g(z_m), g(z)) \rightarrow 0$. By the triangle inequality together with $m \rightarrow \infty$ we conclude that

$$\sup_{\omega_1 \in \Omega'_{1,\epsilon}} d_E(g_n(\omega_1, y), g(z)) < \delta.$$

This proves that (20) holds. \square

3.1.3 Proof of (15)

We shall now prove that $\Omega'_{1,\epsilon}$ fulfills the requirements of (15), which would conclude the proof of Theorem 1.

Lemma 4 *For any $\epsilon_2 > 0$, the d_E -compact set $\tilde{K}_{\epsilon_2} := g(K_{\epsilon_2})$ satisfies*

$$\forall \delta_2 > 0, \overline{\lim}_{n \rightarrow \infty} \sup_{\omega_1 \in \Omega'_{1,\epsilon}} \mathbb{P}_{\Omega_2}^* \left(g_n(\omega_1, \mathcal{Z}_n(\omega_1, \cdot_2)) \notin \tilde{K}_{\epsilon_2}^{\delta_2} \right) < \epsilon_2. \quad (26)$$

Proof: Fix $\epsilon_2 > 0$ and take K_{ϵ_2} as the corresponding member of the family $(K_{\epsilon_1})_{\epsilon_1 > 0}$. Now fix $\delta_2 > 0$. Using Lemma 3 (with the formal replacement of δ by δ_2) we can choose $\tau > 0$ and $n_0 \geq 1$ such that, for all $n \geq n_0$ we have

$$\forall z \in K_{\epsilon_2}, \forall y \in B(z, \tau) \cap D_n, \sup_{\omega_1 \in \Omega'_{1,\epsilon}} d_E(g_n(\omega_1, y), g(z)) < \delta_2.$$

This entails the implication " $y \in K_{\epsilon_2}^\tau \cap D_n \Rightarrow g_n(\omega_1, y) \in g(K_{\epsilon_2})^{\delta_2}$ " for each $n \geq n_0$ and $\omega_1 \in \Omega'_{1,\epsilon}$. As a consequence, recalling (3) we have

$$\sup_{\omega_1 \in \Omega'_{1,\epsilon}} \mathbb{P}_{\Omega_2}^* \left(g_n(\omega_1, \mathcal{Z}_n(\omega_1, \cdot)) \notin g(K_{\epsilon_2})^{\delta_2} \right) \leq \sup_{\omega_1 \in \Omega'_{1,\epsilon}} \mathbb{P}_{\Omega_2}^* \left(\mathcal{Z}_n(\omega_1, \cdot) \notin K_{\epsilon_2}^\tau \right).$$

By Lemma 2, the right hand side of the preceding inequality is strictly less than ϵ_2 for all large enough n . Moreover, since $K_{\epsilon_2} \subset D_0$ and $g|_{D_0}$ is continuous, the set $\tilde{K}_{\epsilon_2} := g(K_{\epsilon_2})$ is compact for the metric d_E , which concludes the proof. \square

Lemma 5 *For each $\epsilon_3 > 0$ and for each compact $\tilde{K} \subset E$, there exists $\delta_3 > 0$ and $\{L_1, \dots, L_M\} \subset BL_1(E)$ such that*

$$\sup_{L \in BL_1(E)} \min_{m=1, \dots, M} \|L - L_m\|_{\tilde{K}^{\delta_3}} < \epsilon_3. \quad (27)$$

Proof: Fix $\epsilon_3 > 0$ and a compact $\tilde{K} \subset E$. Since the family of functions $BL_1(E)$ is uniformly bounded and uniformly d_E -equicontinuous on the compact set \tilde{K} , it is therefore totally bounded for the seminorm $\|\cdot\|_{\tilde{K}}$. Let L_1, \dots, L_M be the centers of $\|\cdot\|_{\tilde{K}}$ balls with radius $\epsilon_3/2$ for which the union covers $BL_1(E)$. Now choose $\delta_3 := \epsilon_3/3$. For any $L \in BL_1(E)$ there exists $L_m, m \in \llbracket 1, M \rrbracket$ such that

$$\sup_{e \in \tilde{K}} |L(e) - L_m(e)| < \epsilon_3/3.$$

Now since both L and L_m are 1-Lipschitz, we then have, by the triangle inequality

$$\sup_{e \in \tilde{K}^{\delta_3}} |L(e) - L_m(e)| < \epsilon_3/3 + \epsilon_3/3 + \epsilon_3/3,$$

which concludes the proof. \square .

Lemma 6 *Let $L \in BL_1(E)$ and let $K \subset D_0$ be a member of the family $(K_{\epsilon_1})_{\epsilon_1 > 0}$. For each $\epsilon_4 > 0$ there exists $\delta_4 > 0$ and a real valued bounded Lipschitz map L' on D for which we have:*

$$\begin{aligned} \overline{\lim}_{n \rightarrow \infty} \sup_{\omega_1 \in \Omega'_{1,\epsilon}} \left| \mathbb{E}_{\Omega_2}^* \left(L \left(g_n(\omega_1, \mathcal{Z}_n(\omega_1, \cdot)) \mathbf{1}_{K^{\delta_4}}(\mathcal{Z}_n(\omega_1, \cdot)) \right) \right) \right. \\ \left. - \mathbb{E}_{\Omega_2}^* \left(L'(\mathcal{Z}_n(\omega_1, \cdot)) \mathbf{1}_{K^{\delta_4}}(\mathcal{Z}_n(\omega_1, \cdot)) \right) \right| < 2\epsilon_4, \end{aligned} \quad (28)$$

$$\left| \mathbb{E} \left(L(g(\mathcal{Z})) \mathbf{1}_{K^{\delta_4}}(\mathcal{Z}) \right) - \mathbb{E} \left(L'(\mathcal{Z}) \mathbf{1}_{K^{\delta_4}}(\mathcal{Z}) \right) \right| < 2\epsilon_4. \quad (29)$$

Proof: Fix $L \in BL_1(E)$ and let K be a member of $(K_{\epsilon_1})_{\epsilon_1 > 0}$. By (20) we know that there exists $\delta_4 > 0$ and $n_0 \geq 1$ such that, for each $z \in K$:

$$n \geq n_0, y \in B(z, 3\delta_4) \cap D_n \Rightarrow \sup_{\omega_1 \in \Omega'_{1,\epsilon}} d_E(g_n(\omega_1, y), g(z)) \leq \epsilon_4. \quad (30)$$

Since K is totally bounded there exists $\{z_1, \dots, z_k\} \subset K$ such that

$$K^{\delta_4} \subset \bigcup_{i=1}^k B(z_i, 2\delta_4), \text{ and}$$

$$\forall n \geq n_0, \forall i \in \llbracket 1, k \rrbracket, \forall y \in B(z_i, 3\delta_4) \cap D_n, \sup_{\omega_1 \in \Omega'_{1,\epsilon}} d_E(g_n(\omega_1, y), g(z_i)) \leq \epsilon_4. \quad (31)$$

Note that (31) implies:

$$\forall n \geq n_0, \forall i \in \llbracket 1, k \rrbracket, \forall y \in B(z_i, 3\delta_4) \cap D_n, \sup_{\omega_1 \in \Omega'_{1,\epsilon}} |L(g_n(\omega_1, y)) - L(g(z_i))| \leq \epsilon_4, \quad (32)$$

since $L \in BL_1(E)$. Using (11) and taking limits in n we also obtain

$$\forall i \in \llbracket 1, k \rrbracket, \forall y \in B(z_i, 3\delta_4) \cap D_0, \sup_{\omega_1 \in \Omega'_{1,\epsilon}} |L(g(y)) - L(g(z_i))| \leq \epsilon_4. \quad (33)$$

We will now show that there exists $\delta \in [2\delta_4, 3\delta_4[$ such that

$$\forall i \in \llbracket 1, k \rrbracket, \mathbb{P}(\mathcal{Z} \in \partial B(z_i, \delta) \cap K^{\delta_4}) = 0. \quad (34)$$

To prove this, let us first notice that, for fixed $i \in \llbracket 1, k \rrbracket$, the family

$$\left(\mathbb{P}(\mathcal{Z} \in \partial B(z_i, \delta) \cap K^{\delta_4}) \right)_{\delta \in [2\delta_4, 3\delta_4[}$$

is summable (since \mathbb{P} is a finite measure). As a consequence, for fixed i , the set

$$\left\{ \delta \in [2\delta_4, 3\delta_4[, \mathbb{P}(\mathcal{Z} \in \partial B(z_i, \delta) \cap K^{\delta_4}) > 0 \right\}$$

is countable. By union, this is in turn the case for the set

$$\left\{ \delta \in [2\delta_4, 3\delta_4[, \exists i \in \llbracket 1, k \rrbracket, \mathbb{P}(\mathcal{Z} \in \partial B(z_i, \delta) \cap K^{\delta_4}) > 0 \right\}.$$

Hence the complement of this set in $[2\delta_4, 3\delta_4[$ is nonempty. This proves the existence of $\delta \in [2\delta_4, 3\delta_4[$ fulfilling (34).

Let us now define the following partition of K^{δ_4} :

$$B_i := B(z_i, \delta) \cap K^{\delta_4} \setminus \left(\bigcup_{j=1}^{i-1} B(z_j, \delta) \right), \quad i \in \llbracket 1, k \rrbracket.$$

Since each B_i is the difference between two open sets, and since \mathcal{Z} induces a regular probability measure on (D, d_D) , there exist Lipschitz functions L'_1, \dots, L'_k such that

$$\max_{i \in \llbracket 1, k \rrbracket} \mathbb{E} \left(|L'_i(\mathcal{Z}) - \mathbf{1}_{B_i}(\mathcal{Z})| \right) < \frac{\epsilon_4}{k}. \quad (35)$$

By (32) and since $B_i \subset B(z_i, \delta) \subset B(z_i, 3\delta_4)$ for each $i \in \llbracket 1, k \rrbracket$, we have, for all $n \geq n_0$:

$$\sup_{\omega_1 \in \Omega'_{1,\epsilon}} \left\| \sum_{i=1}^k L(g(z_i)) \mathbf{1}_{B_i} - L(g_n(\omega_1, \cdot)) \mathbf{1}_{B_i} \right\|_D < \epsilon_4.$$

Likewise, using (33) we have

$$\left\| \sum_{i=1}^k L(g(z_i)) \mathbf{1}_{B_i} - L(g(\cdot)) \mathbf{1}_{B_i} \right\|_{D_0} < \epsilon_4. \quad (36)$$

But by (35) we also have

$$\begin{aligned} \mathbb{E} \left(\left| \sum_{i=1}^k L(g(z_i)) \mathbf{1}_{B_i}(\mathcal{Z}) - L(g(z_i)) L'_i(\mathcal{Z}) \right| \right) &\leq \max_{i \in \llbracket 1, k \rrbracket} |L(g(z_i))| \sum_{i=1}^k \frac{\epsilon_4}{k} \\ &\leq \epsilon_4. \end{aligned} \quad (37)$$

Now since $\mathbb{P}(\mathcal{Z} \in \partial B_i) = 0$ (by (34)) and by Lemma 1 we have

$$\lim_{n \rightarrow \infty} \sup_{\omega_1 \in \Omega'_{1,\epsilon}} \left| \mathbb{E}_{\Omega_2}^* \left(\sum_{i=1}^k L(g(z_i)) \mathbf{1}_{B_i}(\mathcal{Z}_n(\omega_1, \cdot)) \right) - \mathbb{E} \left(\sum_{i=1}^k L(g(z_i)) \mathbf{1}_{B_i}(\mathcal{Z}) \right) \right| = 0. \quad (38)$$

Moreover by (14) we have

$$\lim_{n \rightarrow \infty} \sup_{\omega_1 \in \Omega'_{1,\epsilon}} \left| \mathbb{E}_{\Omega_2}^* \left(\sum_{i=1}^k L(g(z_i)) L'_i(\mathcal{Z}_n(\omega_1, \cdot)) \right) - \mathbb{E} \left(\sum_{i=1}^k L(g(z_i)) L'_i(\mathcal{Z}) \right) \right| = 0. \quad (39)$$

The proof is concluded as follows: set $L' := \sum_{i=1}^k L(g(z_i)) L'_i$. Combining (36) and (37) we obtain (29) recalling that $\{B_i, i \in \llbracket 1, k \rrbracket\}$ is a partition of K^{δ_4} . Finally, a combination of (35), (37), (38) and (39) proves (28). \square

We now have all the tools to establish (15). First fix $\epsilon > 0$, and choose $\tilde{K}_\epsilon = g(K_\epsilon)$ as in Lemma 4. Next, apply Lemma 5 with the choice of $\epsilon_3 := \epsilon$. This exhibits $\delta_3 > 0$ and a finite sequence L_1, \dots, L_M fulfilling (27). Then, for each fixed $m \in \llbracket 1, M \rrbracket$, apply Lemma 6 to $L := L_m$, K_ϵ and $\epsilon_4 := \epsilon$ in order to exhibit $\delta_{4,m} > 0$ fulfilling (28) and (29). Combine those assertions with (14) to obtain, for the choice of $\delta_4 := \min\{\delta_3, \delta_{4,1}, \dots, \delta_{4,M}\}$.

$$\begin{aligned} \overline{\lim}_{n \rightarrow \infty} \sup_{\omega_1 \in \Omega'_{1,\epsilon}} \max_{L \in \{L_1, \dots, L_M\}} \left| \mathbb{E}_{\Omega_2}^* \left(L(g_n(\omega_1, \mathcal{Z}_n(\omega_1, \cdot))) \mathbf{1}_{K_\epsilon^{\delta_4}}(\mathcal{Z}_n(\omega_1, \cdot)) \right) \right. \\ \left. - \mathbb{E} \left(L(g(\mathcal{Z})) \mathbf{1}_{K_\epsilon^{\delta_4}}(\mathcal{Z}) \right) \right| < 4\epsilon. \end{aligned}$$

This, combined with (27) entails

$$\overline{\lim}_{n \rightarrow \infty} \sup_{\omega_1 \in \Omega'_{1,\epsilon}} \sup_{L \in BL_1(E)} \left| \mathbb{E}_{\Omega_2}^* \left(L(g_n(\omega_1, \mathcal{Z}_n(\omega_1, \cdot))) \mathbf{1}_{K_\epsilon^{\delta_4}}(\mathcal{Z}_n(\omega_1, \cdot)) \right) \right|$$

$$- \mathbb{E} \left(L(g(\mathcal{Z})) \mathbf{1}_{K_\varepsilon^{\delta_4}}(\mathcal{Z}) \right) \Big| < 5\varepsilon. \quad (40)$$

Finally, recall that K_ε fulfills the conditions of Lemma 2 for $\varepsilon_1 := \varepsilon$. We hence make use of this lemma with $\delta_1 := \delta_4$ to obtain

$$\begin{aligned} \overline{\lim}_{n \rightarrow \infty} \sup_{\omega_1 \in \Omega'_{1,\varepsilon}} \sup_{L \in BL_1(E)} \left| \mathbb{E}_{\Omega_2}^* \left(L \left(g_n(\omega_1, \mathcal{Z}_n(\omega_1, \cdot, 2)) \right) \mathbf{1}_{D \setminus K_\varepsilon^{\delta_4}}(\mathcal{Z}_n(\omega_1, \cdot, 2)) \right) \right| &\leq \varepsilon, \\ \sup_{L \in BL_1(E)} \left| \mathbb{E} \left(L(g(\mathcal{Z})) \mathbf{1}_{D \setminus K_\varepsilon^{\delta_4}}(\mathcal{Z}) \right) \right| &\leq \varepsilon. \end{aligned} \quad (41)$$

Combine those two assertions with (40) to obtain

$$\overline{\lim}_{n \rightarrow \infty} \sup_{\omega_1 \in \Omega'_{1,\varepsilon}} \sup_{L \in BL_1(E)} \left| \mathbb{E}_{\Omega_2}^* \left(L \left(g_n(\omega_1, \mathcal{Z}_n(\omega_1, \cdot, 2)) \right) \right) - \mathbb{E} \left(L(g(\mathcal{Z})) \right) \right| < 7\varepsilon.$$

This proves (15) and hence completes the proof of Theorem 1. \square

3.2 Proof of Corollary 1

Fix an arbitrary subsequence (n_k) . It is sufficient to show that there exists a further subsequence (n'_k) of (n_k) for which

$$\left(\sup_{L \in BL_1(E)} \left| \mathbb{E}_{\Omega_2}^* \left(L \left(g_{n'_k}(\cdot, \mathcal{Z}_{n'_k}(\cdot, \cdot, 2)) \right) \right) - \mathbb{E} \left(L(g(\mathcal{Z})) \right) \right| \right)_{\Omega_1}^* \rightarrow 0, \quad (42)$$

in \mathbb{P}_1 -probability as $k \rightarrow \infty$. By assumption there exists a subsequence (n'_k) of (n_k) satisfying, for all $k \geq 1$:

$$\mathbb{P}_1 \left(\left(\sup_{L \in BL_1(D)} \left| \mathbb{E}_{\Omega_2}^* \left(L \left(\mathcal{Z}_{n'_k}(\cdot, \cdot, 2) \right) \right) - \mathbb{E} \left(L(\mathcal{Z}) \right) \right| \right)_{\Omega_1}^* \geq 2^{-k} \right) \leq 2^{-k}.$$

Now, by $(Hg)'$ we can suppose without loss of generality that (n'_k) also fulfills

$$\mathbb{P}_1^* \left(\left\{ \omega_1 \in \Omega_1, \exists z \in D_0, \forall \tau > 0, \sup_{y \in D_{n'_k}, d_D(z', z) < \tau} d_E(g_{n'_k}(\omega_1, y), g(z)) > 2^{-k} \right\} \right) \leq 2^{-k},$$

for all $k \geq 1$. Apply the Borel Cantelli Lemma for outer probabilities to conclude that $(\mathcal{Z}_{n'_k})$ and $(g_{n'_k})$ satisfy all the assumptions of Theorem 1, which proves the stronger almost sure version of (42) along (n'_k) and hence concludes the proof. \square

3.3 Proof of Theorem 2

We will prove Theorem 2 for the case $k = 2$. The general case is then obtained by an elementary recursion. Fix $\omega_1 = (\omega_1^{(1)}, \omega_1^{(2)}) \in \Omega_1$. We have

$$\sup_{L \in BL_1(\mathbb{D})} \left| \mathbb{E}_{\Omega_2} \left(L \left(\hat{\mathbb{G}}_{n_1,1}(\omega_1, \cdot, 2), \hat{\mathbb{G}}_{n_2,2}(\omega_1, \cdot, 2) \right) \right) - \mathbb{E} \left(L(\mathbb{G}_1, \mathbb{G}_2) \right) \right|$$

$$\begin{aligned}
&= \sup_{L \in BL_1(\mathbb{D})} \left| \int L\left(\hat{\mathbb{G}}_{n_1,1}(\omega_1, \omega_2), \hat{\mathbb{G}}_{n_2,2}(\omega_1, \omega_2)\right) d\mathbb{P}_2(\omega_2) - \mathbb{E}\left(L(\mathbb{G}_1, \mathbb{G}_2)\right) \right| \\
&\hspace{25em} (43) \\
&= \sup_{L \in BL_1(\mathbb{D})} \left| \int L\left(\hat{\mathbb{G}}_{n_1,1}(\omega_1^{(1)}, \omega_2^{(1)}), \hat{\mathbb{G}}_{n_2,2}(\omega_1^{(2)}, \omega_2^{(2)})\right) dQ_1(\omega_2^{(1)}) dQ_2(\omega_2^{(2)}) - \mathbb{E}\left(L(\mathbb{G}_1, \mathbb{G}_2)\right) \right| \\
&= \sup_{L \in BL_1(\mathbb{D})} \left| \int \left(\int L\left(\hat{\mathbb{G}}_{n_1,1}(\omega_1^{(1)}, \omega_2^{(1)}), \hat{\mathbb{G}}_{n_2,2}(\omega_1^{(2)}, \omega_2^{(2)})\right) dQ_1(\omega_2^{(1)}) \right) dQ_2(\omega_2^{(2)}) - \mathbb{E}\left(L(\mathbb{G}_1, \mathbb{G}_2)\right) \right| \\
&\leq \sup_{L \in BL_1(\mathbb{D})} \left| \int \left(\int L\left(\hat{\mathbb{G}}_{n_1,1}(\omega_1^{(1)}, \omega_2^{(1)}), \hat{\mathbb{G}}_{n_2,2}(\omega_1^{(2)}, \omega_2^{(2)})\right) dQ_1(\omega_2^{(1)}) \right) dQ_2(\omega_2^{(2)}) \right. \\
&\quad \left. - \int \mathbb{E}\left(L(\mathbb{G}_1, \hat{\mathbb{G}}_{n_2,2}(\omega_1^{(2)}, \omega_2^{(2)}))\right) dQ_2(\omega_2^{(2)}) \right| \\
&\quad + \sup_{L \in BL_1(\mathbb{D})} \left| \int \mathbb{E}\left(L(\mathbb{G}_1, \hat{\mathbb{G}}_{n_2,2}(\omega_1^{(2)}, \omega_2^{(2)}))\right) dQ_2(\omega_2^{(2)}) - \mathbb{E}\left(L(\mathbb{G}_1, \mathbb{G}_2)\right) \right| \\
&=: \Delta_{n,1}(\omega_1) + \Delta_{n,2}(\omega_1),
\end{aligned}$$

where in (43) we employ the abuse of notations $\hat{\mathbb{G}}_{n_1,1}(\omega_1^{(1)}, \omega_2^{(1)})$ instead of $\hat{\mathbb{G}}_{n_1,1}(\omega_1, \omega_2)$ by recalling that the involved map is constant in $\omega_1^{(2)}, \omega_2^{(2)}$ (the same argument holds for $\hat{\mathbb{G}}_{n_2,2}(\omega_1^{(2)}, \omega_2^{(2)})$). Recalling that, if $L \in BL_1(\mathbb{D})$, then the partial map $x \rightarrow L(x, y)$ belongs to $BL_1(\mathbb{D}_1)$ for each $y \in \mathbb{D}_2$, the triangle inequality yields

$$\begin{aligned}
&\Delta_{n,1}(\omega_1) \\
&\leq \sup_{\substack{L \in BL_1(\mathbb{D}), \\ \omega_2^{(2)} \in \Omega_{2,2}}} \left| \int L\left(\hat{\mathbb{G}}_{n_1,1}(\omega_1^{(1)}, \omega_2^{(1)}), \hat{\mathbb{G}}_{n_2,2}(\omega_1^{(2)}, \omega_2^{(2)})\right) dQ_1(\omega_2^{(1)}) - \mathbb{E}\left(L(\mathbb{G}_1, \hat{\mathbb{G}}_{n_2,2}(\omega_1^{(2)}, \omega_2^{(2)}))\right) \right| \\
&\leq \sup_{\tilde{L} \in BL_1(\mathbb{D}_1)} \left| \int \tilde{L}\left(\hat{\mathbb{G}}_{n_1,1}(\omega_1^{(1)}, \omega_2^{(1)})\right) dQ_1(\omega_2^{(1)}) - \mathbb{E}\left(\tilde{L}(\mathbb{G}_1)\right) \right| \\
&=: \tilde{\Delta}_{n,1}(\omega_1^{(1)}).
\end{aligned}$$

This upper bound defines a function on $\Omega_{1,1}$ only (since its expression does not depend upon $\omega_1^{(2)}$). As a consequence, since the coordinate map $\pi_1 : (\omega_1^{(1)}, \omega_1^{(2)}) \rightarrow \omega_1^{(1)}$ on Ω_1 is perfect - see e.g. [15, p. 10] - we have, for all $(\omega_1^{(1)}, \omega_1^{(2)}) \in \Omega_{1,1} \times \Omega_{1,2}$

$$(\Delta_{n,1})_{\Omega_1}^*(\omega_1^{(2)}, \omega_1^{(1)}) \leq (\tilde{\Delta}_{n,1} \circ \pi_1)_{\Omega_1}^*(\omega_1^{(2)}, \omega_1^{(1)}) = (\tilde{\Delta}_{n,1})_{\Omega_{1,1}}^*(\omega_1^{(1)}). \quad (44)$$

Now since \mathcal{F}_1 is $P_{0,1}$ -Donsker, the bootstrap Donsker theorem applies - see e.g. [15, p. 347, Theorem 3.6.1] - namely: $(\tilde{\Delta}_{n,1})_{\Omega_{1,1}}^* \rightarrow 0$ in $\mathbb{P}_{1,1}$ -probability, which in turn implies that $(\Delta_{n,1})_{\Omega_1}^* \rightarrow 0$ in \mathbb{P}_1 -probability by (44). The proof will then be concluded if we prove that the same convergence holds for $(\Delta_{n,2})_{\Omega_1}^*$. To prove this, notice that $L \in BL_1(\mathbb{D})$ then the partial integration map $y \rightarrow$

$\int_x L(x, y) d\nu(x)$ belongs to $BL_1(\mathbb{D}_2)$, for any Borel probability measure ν . This entails

$$\begin{aligned} \Delta_{n,2}(\omega_1) &\leq \sup_{\tilde{L} \in BL_1(\mathbb{D}_2)} \left| \int \tilde{L}(\hat{\mathbb{G}}_{n_2,2}(\omega_1^{(2)}, \omega_2^{(2)})) dQ_2(\omega_2^{(2)}) - \mathbb{E}(\tilde{L}(\mathbb{G}_2)) \right| \\ &=: \tilde{\Delta}_{n,2}(\omega_1^{(1)}). \end{aligned}$$

Now apply the bootstrap Donsker Theorem to the sequence $\tilde{\Delta}_{n,2}$ of maps on $\Omega_{1,2}$: since \mathcal{F}_2 is $P_{0,2}$ Donsker, we have $(\tilde{\Delta}_{n,2})_{\Omega_{1,2}}^* \rightarrow 0$ in $\mathbb{P}_{1,2}$ probability. Using a similar comparison argument as (44) we obtain that $(\Delta_{n,2})_{\Omega_1}^* \rightarrow 0$ in \mathbb{P}_1 -probability. This concludes the proof. \square

3.4 Proof of Theorem 3

Fix $\alpha = (\alpha_1, \dots, \alpha_k) \in (0, 1]^k$, write \mathbf{G}_n for $(\mathbb{G}_{n_1,1}, \dots, \mathbb{G}_{n_k,k})$, write $\hat{\mathbf{G}}_n$ for $(\hat{\mathbb{G}}_{n_1,1}, \dots, \hat{\mathbb{G}}_{n_k,k})$ and write \mathbf{G} for $(\mathbb{G}_1, \dots, \mathbb{G}_k)$. The proof follows the same line as in [15, p. 380]. We can first assume without loss of generality that the map $d\Phi_{\mathbf{P}_0}$ is defined and continuous on the whole space \mathbb{D} , and hence so are the maps $d\Phi_{\mathbf{P}_0}^\alpha$ and $d\Phi_{\mathbf{P}_0}^{\alpha_n}$, where, for any $\beta = (\beta_1, \dots, \beta_k) \in (0, 1]^k$ we write

$$d\Phi_{\mathbf{P}_0}^\beta := (\psi_1, \dots, \psi_k) \rightarrow \sum_{j=1}^k \beta_j^{-1/2} d\Phi_{\mathbf{P}_0}^{(j)}(\psi_j).$$

For any L in $BL_1(D)$, the map $L \circ d\Phi_{\mathbf{P}_0}^\alpha$ belongs to $BL_{\|\cdot\|_{d\Phi_{\mathbf{P}_0}^\alpha}}(\mathbb{D})$, where $\|\cdot\|$ stands for the operator norm. The same statement holds for α_n instead of α . Hence by Theorem 2 and by (7) we have

$$\left(\sup_{L \in BL_1(D)} \left| \mathbb{E}_{\Omega_2} \left(L \left(d\Phi_{\mathbf{P}_0}^{\alpha_n}(\hat{\mathbf{G}}_n) \right) \right) - \mathbb{E} \left(L \left(d\Phi_{\mathbf{P}_0}^\alpha(\mathbf{G}) \right) \right) \right| \right)_{\Omega_1}^* \rightarrow 0, \quad (45)$$

in \mathbb{P}_1 -probability. On the other hand one has, for any $\epsilon > 0$ and n , everywhere on Ω_1 :

$$\begin{aligned} &\sup_{L \in BL_1(D)} \left| \mathbb{E}_{\Omega_2} \left(L \left(\sqrt{N}(\Phi(\hat{\mathbf{P}}_n) - \Phi(\mathbf{P}_n)) \right) \right) - \mathbb{E}_{\Omega_2} \left(L \left(d\Phi_{\mathbf{P}_0}^{\alpha_n}(\hat{\mathbf{G}}_n) \right) \right) \right| \\ &\leq \epsilon + 2\mathbb{P}_{\Omega_2} \left(\left\| \sqrt{N}(\Phi(\hat{\mathbf{P}}_n) - \Phi(\mathbf{P}_n)) - d\Phi_{\mathbf{P}_0}^{\alpha_n}(\hat{\mathbf{G}}_n) \right\| > \epsilon \right). \end{aligned} \quad (46)$$

We already know that

$$\mathbf{G}_n \rightarrow_{\mathcal{L}} \mathbf{G}.$$

In addition, by mutual independence in $j \in \llbracket 1, k \rrbracket$ and by using [15, Corollary 2.9.4, p. 180] we also have

$$\hat{\mathbf{G}}_n \rightarrow_{\mathcal{L}} \mathbf{G} + \mathbf{G}',$$

where \mathbf{G}' is an independent copy of \mathbf{G} . As a consequence, both sequences $\sqrt{N}(\mathbf{P}_n - \mathbf{P}_0)$ and $\sqrt{N}(\hat{\mathbf{P}}_n - \mathbf{P}_0)$ are weakly convergent as $n \rightarrow \infty$ together with (7). One can hence apply [15, Theorem 3.9.4, p. 374] - using e.g. an arbitrary sequence $(n(n))$ as we did in §2 - to obtain

$$\begin{aligned} & \left(\left\| \sqrt{N} \left(\Phi(\mathbf{P}_n) - \Phi(\mathbf{P}_0) \right) - d\Phi_{\mathbf{P}_0}(\sqrt{N}(\mathbf{P}_n - \mathbf{P}_0)) \right\| \right)^* \rightarrow 0, \text{ in } \mathbb{P}\text{-probability,} \\ & \left(\left\| \sqrt{N} \left(\Phi(\hat{\mathbf{P}}_n) - \Phi(\mathbf{P}_0) \right) - d\Phi_{\mathbf{P}_0}(\sqrt{N}(\hat{\mathbf{P}}_n - \mathbf{P}_0)) \right\| \right)^* \rightarrow 0, \text{ in } \mathbb{P}\text{-probability.} \end{aligned}$$

Substracting these convergences entails

$$\left(\left\| \sqrt{N} \left(\Phi(\hat{\mathbf{P}}_n) - \Phi(\mathbf{P}_n) \right) - d\Phi_{\mathbf{P}_0}(\sqrt{N}(\hat{\mathbf{P}}_n - \mathbf{P}_n)) \right\| \right)^* \rightarrow 0, \text{ in } \mathbb{P}\text{-probability,}$$

which implies, by Fubini's Theorem for outer expectations (see [15, Lemma 1.2.6, p. 11])

$$\mathbb{E}_{\Omega_1}^* \left(\mathbb{P}_{\Omega_2} \left(\left\| \sqrt{N}(\Phi(\hat{\mathbf{P}}_n) - \Phi(\mathbf{P}_n)) - d\Phi_{\mathbf{P}_0}^{\alpha_n}(\hat{\mathbf{G}}_n) \right\| > \epsilon \right) \right) \rightarrow 0,$$

for any given $\epsilon > 0$. This combined with (46) proves that

$$\left(\sup_{L \in BL_1(D)} \left| \mathbb{E}_{\Omega_2} \left(L \left(\sqrt{N}(\Phi(\hat{\mathbf{P}}_n) - \Phi(\mathbf{P}_n)) \right) \right) - \mathbb{E}_{\Omega_2} \left(L \left(d\Phi_{\mathbf{P}_0}^{\alpha_n}(\hat{\mathbf{G}}_n) \right) \right) \right| \right)_{\Omega_1}^* \rightarrow 0,$$

in \mathbb{P}_1 -probability. This concludes the proof by using (45) together with the triangle inequality for outer expectations. \square

3.5 Proof of Proposition 1

Take $z = \phi \in \ell_{\mathfrak{X}_0,1}$, $n \in \mathbb{N}^*$, $\delta > 0$. Then we have the following equalities as maps on Ω_1 , (47) being a consequence of the fact that the all the involved \mathfrak{X} -indexed trajectories are elements of $\ell_{\mathfrak{X}_0,1}$

$$\begin{aligned} & \sup_{y \in D_n, d_D(y,z) < \delta} d_E(g_{n(n)}(\cdot, y), g(z)) \\ & \equiv \sup_{\tilde{\phi}, \|\tilde{\phi} - \phi\|_{\mathfrak{X}} < \delta} \sup_{t \in \mathfrak{X}} \left| \tilde{\phi}(t)^\top M_{n(n)}(t) \tilde{\phi}(t) - \phi(t)^\top V(t)^{-1} \phi(t) \right|, \\ & \equiv \sup_{\tilde{\phi}, \|\tilde{\phi} - \phi\|_{\mathfrak{X}} < \delta} \sup_{t \in \mathfrak{X}_0} \left| \tilde{\phi}(t)^\top M_{n(n)}(t) \tilde{\phi}(t) - \phi(t)^\top V(t)^{-1} \phi(t) \right| \quad (47) \\ & \equiv \sup_{t \in \mathfrak{X}_0} \sup_{\tilde{\phi}, \|\tilde{\phi} - \phi\|_{\mathfrak{X}} < \delta} \left| \tilde{\phi}(t)^\top M_{n(n)}(t) \tilde{\phi}(t) - \phi(t)^\top V(t)^{-1} \phi(t) \right| \\ & \equiv \sup_{t \in \mathfrak{X}_0} \sup_{\mathbf{u} \in A_{t,\delta}} \left| \mathbf{u}^\top M_{n(n)}(t) \mathbf{u} - \phi(t)^\top V(t)^{-1} \phi(t) \right|, \end{aligned}$$

where, for each $t \in \mathfrak{T}_0$ we wrote $A_{t,\delta} := \{\mathbf{u} \in \mathbb{R}^p, \|\mathbf{u} - \phi(t)\|_p < \delta\}$. Since \mathfrak{T}_0 is countable it will be sufficient to prove - for fixed $t \in \mathfrak{T}_0$ - the measurability of the second supremum in the preceding display. To see this, take $t \in \mathfrak{T}_0$ and first note that $\inf\{\det(V_{\mathbf{n}}(t)), t \in \mathfrak{T}\} = \inf\{\det(V_{\mathbf{n}}(t)), t \in \mathfrak{T}_0\}$ - and is hence Borel measurable - because $V_{\mathbf{n}}$ takes its values in $\ell_{\mathfrak{T}_0, p^2}$ and since the determinant is continuous function. Since $S_{\mathbf{n}(n)}^{-1}(t)$ is itself Borel measurable, we first conclude that $M_{\mathbf{n}(n)}(t)$ is measurable. Finally the pointwise supremum of all the continuous functions $M \rightarrow u^\top M u$ for $u \in A_{t,\delta}$ is Borel, which concludes the proof. \square

References

- [1] E. Beutner and H. Zähle. A modified functional delta method and its application to the estimation of risk functionals. *Journ. Multiv. Anal.*, 101(10):2452–2463, 2010.
- [2] E. Beutner and H. Zähle. Deriving the asymptotic distribution of U-and V-statistics of dependent data using weighted empirical processes. *Bernoulli*, 18(3):803–822, 2012.
- [3] E. Beutner and H. Zähle. Functional delta-method for the bootstrap of quasi-Hadamard differentiable functionals. *Electron. Journ. Statist.*, 10(1):1181–1222, 2016.
- [4] P. Billingsley. *Convergence of probability measures*. Wiley, New York, 1968.
- [5] D. Chibisov. An investigation of the asymptotic power of the tests of fit. *Theory Probab. Appl.*, 10(3):421–437, 1965.
- [6] B. Efron. Bootstrap methods: Another look at the jackknife. *Ann. Statist.*, 7(1):1–26, 1979.
- [7] R. D. Gill, J. Præstgaard, and J. A. Wellner. Non and semi-parametric maximum likelihood estimators and the von Mises method (part 1)[with discussion and reply]. *Scand. J. of Statist.*, 14(2):97–128, 1989.
- [8] E. Giné. Lectures on some aspects of the bootstrap. In *Lectures on Probability Theory and Statistics*, pages 37–151. Springer, 1997.
- [9] E. Giné and J. Zinn. Bootstrapping general empirical measures. *Ann. Probab.*, pages 851–869, 1990.
- [10] J. Hoffmann-Jorgensen. *Stochastic processes on Polish spaces*. Number 39 in Various Publication Series. Aarhus Universitet, Aarhus, Denmark, 1991.
- [11] M. Ledoux and M. Talagrand. Un critère sur les petites boules dans le théorème limite central. *Prob. Theo. Related Fields*, 77(1):29–47, 1988.

- [12] J. Praestgaard and J. A Wellner. Exchangeably weighted bootstraps of the general empirical process. *Ann. Probab.*, 21(4):2053–2086, 1993.
- [13] Y. V. Prokhorov. Convergence of random processes and limit theorems in probability theory. *Theory Probab. Appl.*, 1(2):157–214, 1956.
- [14] A. V. Skorokhod. Limit theorems for stochastic processes. *Theory Probab. Appl.*, 1(3):261–290, 1956.
- [15] A.W. Van der Vaart and J.A. Wellner. *Weak Convergence and Empirical Processes*. Springer, New-York, 1996.
- [16] D. Varron. Empirical likelihood confidence tubes for functional parameters in plug-in estimation. *Journ. Multiv. Anal.*, 152:100–118, 2016.