

# Asymptotic stability of empirical processes and related functionals<sup>☆</sup>

José L. Fernández<sup>a</sup>, Enrico Ferri<sup>b</sup>, Carlos Vázquez<sup>b</sup>

<sup>a</sup>*Departamento de Matemáticas. Universidad Autónoma de Madrid, Francisco Tomás y Valiente, 7, 28049-Madrid (Spain)*

<sup>b</sup>*Department of Mathematics, Faculty of Informatics, Campus Elviña s/n, 15071-A Coruña (Spain)*

---

## Abstract

Let  $E$  be a space of observables in a sequence of trials  $\xi_n$  and define  $m_n$  to be the empirical distributions of the outcomes. We discuss the almost sure convergence of the sequence  $m_n$  in terms of the  $\psi$ -weak topology of measures, when the sequence  $\xi_n$  is assumed to be stationary. In this respect, the limit variable is naturally described as a certain canonical conditional distribution. Then, given some functional  $\tau$  defined on a space of laws, the consistency of the estimators  $\tau(m_n)$  is investigated. Moreover, a result on the asymptotic stability of the distribution of  $\tau(m_n)$  w.r.t. certain transformations of the data is given.

*Keywords:*

Stationarity, weak convergence of empirical process, random measures, universal Glivenko-Cantelli classes, consistency, asymptotic stability,  $\psi$ -weak topology

---

## 1. Introduction

Let  $E$  be a space in which any element encodes an observable in a sequence of trials  $\xi_1, \xi_2, \dots$ , and let  $E^{\mathbb{N}}$  be the entire space of the sequences of

---

<sup>☆</sup>Partially funded by EU H2020-MSCA-ITN-2014 (WAKEUPCALL Grant Agreement 643045), Spanish MINECO (Grant MTM2016-76497-R) and by Xunta de Galicia (Grant CN2014/044).

*Email addresses:* [joseluis.fernandez@uam.es](mailto:joseluis.fernandez@uam.es) (José L. Fernández), [enrico.ferri@udc.es](mailto:enrico.ferri@udc.es) (Enrico Ferri), [carlosv@udc.es](mailto:carlosv@udc.es) (Carlos Vázquez)

outcomes, endowed with a background probability measure  $\mathbb{P}$ . In particular, throughout this paper, we assume that the observations  $\xi_1, \xi_2, \dots$  form a stationary sequence with respect to the measure  $\mathbb{P}$ .

For any  $n \geq 1$ , let  $m_n \triangleq n^{-1} \sum_{i \leq n} \delta_{\xi_i}$  be the *empirical distribution* generated by the observations. Given the entire class  $\mathfrak{M}_1(E)$  of laws on  $E$ , endowed with some proper measurable structure, each empirical mean  $m_n$  may be always understood as a random element of  $\mathfrak{M}_1(E)$ .

This paper is motivated by the study of the asymptotic stability of the random sequence  $\tau_n \triangleq \tau(m_n)$ ,  $n \geq 1$ , when the variables  $\xi_n$  encode the *historical data* of a certain *financial risk factor* and  $\tau : \mathfrak{M}_1(E) \rightarrow T$  is a certain statistic assessing the downside risk of the related exposure. Indeed, the study of the asymptotic behaviour of the estimators  $\tau_n$  is crucial to gauge the risk properly, see Cont et al. [3], Föllmer and Weber [8] and Krätschmer et al. [18, 19].

In this respect, a key aspect is whether we have *consistency* of the risk estimators  $\tau_n$ , i.e. whether such a sequence admits a proper limit in some stochastic sense.

If the variables  $\xi_n$  are independent and with common distribution  $\mu$ , Varadarajan theorem guarantees the  $\mathbb{P}$ -almost sure convergence of the empirical process  $m_n$  to the law  $\mu$  in the weak topology, and the consistency of the estimators  $\tau_n$  is directly obtained from the continuity property of the statistic  $\tau$ .

Nevertheless, as highlighted by Cont et al. [3] and Kou et al. [17] some commonly used risk functionals, more precisely the functionals associated to the entire class of *law-invariant convex risk measures*, fail to be continuous with respect to the weak topology of measures. Briefly, the reason lies behind the fact that the weak topology is not sensitive enough to the tail behaviour of the distributions, which, by the way, is the main issue in risk analysis.

An approach to overcome this lack of sensitive has been proposed in Krätschmer et al. [18, 19] and Zähle [25, 26]. Its main ingredient is to introduce a proper refinement of the topological structure, fine enough to control the distributions of the tails, via the subspaces  $\mathfrak{M}_1^\psi(E)$  of  $\mathfrak{M}_1(E)$  defined in terms of *gauge* functions  $\psi$  and the associated  $\psi$ -weak topologies.

Theorem 1 of Section 4 describes the  $\mathbb{P}$ -almost sure convergence of the empirical process  $m_n$ , under the  $\psi$ -weak topology. In particular, the limit variable is naturally described as the random measure obtained as the conditional distribution of  $\xi_1$  given the *shift-invariant*  $\sigma$ -algebra associated to the variables  $\xi_n$ . We point out that in the specific case when in addition the

variables  $\xi_n$  form an ergodic sequence, the consistency result as described in our Corollary 1 has been developed by Krättschmer et al. [19].

When assessing the downside risk associated to some financial exposure, besides consistency, *robustness* is a desirable property of asymptotic stability. Following Hampel, an estimation is said to be robust if small changes of the law related to the outcomes  $\xi_n$  only result in small changes of the distribution characterizing the estimators  $\tau_n$ . While the notion of qualitative robustness has been classically developed, cf. [4, 11, 12, 13, 21], by means of the weak topology of measures, Krättschmer et al. [18, 19] and Zähle [25, 26], provide a similar version that applies to the  $\psi$ -weak topology as the basic topological background.

The main goal of Section 5 is to develop a criterion for the asymptotic stability of the estimators  $\tau_n$ , which turns out to be related to the notion of qualitative robustness, by exploiting the consistency result discussed in Corollary 1. With this aim in mind, we formulate there a notion of asymptotic stability in terms of the modulus of continuity of the statistic  $\tau$ . Such a formulation naturally arises when dealing with random measures, and hence in the particular case of the canonical conditional distribution defined by the variables  $\xi_n$ , if stationarity holds.

The paper is organized as follows. In Section 2 we describe the topological structure of the workspace that we consider throughout the paper. In Section 3 we present some useful measure theoretical results. In section 4 we propose a criterion for consistency that is exploited in Section 5 in order to assess the asymptotic stability of the estimators  $\tau_n$ .

## 2. Background

Let  $E$  be a Polish space and let  $\mathcal{E}$  be its Borel  $\sigma$ -algebra.

We denote by  $\mathfrak{M}_1(E)$  the family of Borel probability measures on  $E$  and by  $\mathfrak{C}_b(E)$  the Banach space of bounded continuous functions defined on  $E$ , endowed with the supremum norm.

Here and in the sequel we use the notation

$$\mu f \triangleq \int_E f(x) \mu(dx),$$

wherever the measure  $\mu \in \mathfrak{M}_1(E)$  and the Borel function  $f$  are such that  $\int_E |f(x)| \mu(dx) < +\infty$ .

The *weak topology*  $\sigma(\mathfrak{M}_1(E), \mathfrak{C}_b(E))$  is the coarsest topology on  $\mathfrak{M}_1(E)$  that renders continuous each map  $\mu \in \mathfrak{M}_1(E) \mapsto \mu f$ , when  $f$  runs over  $\mathfrak{C}_b(E)$ .

Since  $E$  is Polish, the space  $\mathfrak{M}_1(E)$  endowed with the weak topology is metrized as a complete and separable metric space by means of the Prohorov distance

$$\pi(\mu, \nu) \triangleq \inf\{\varepsilon > 0 : \mu(B) \leq \nu(B^\varepsilon) + \varepsilon, \text{ for any } B \in \mathcal{E}\}, \quad (1)$$

for any  $\mu, \nu \in \mathfrak{M}_1(E)$ ,

where  $B^\varepsilon \triangleq \{x \in E : \inf_{y \in B} d(y, x) < \varepsilon\}$  stands for the  $\varepsilon$ -hull of  $B \in \mathcal{E}$ , and  $d$  denotes a *consistent distance* i.e. a metric on  $E$  that is consistent with its topological structure.

### 2.1. Bounded Lipschitz functions $\mathfrak{BL}(E)$

Let  $\mathfrak{BL}(E)$  denote the linear space of Lipschitz bounded functions on  $E$ . For a function  $f \in \mathfrak{BL}(E)$  we define

$$\|f\|_{\mathfrak{BL}(E)} \triangleq \|f\|_\infty + \|f\|_{\mathfrak{L}(E)}$$

where  $\|f\|_{\mathfrak{L}(E)}$  is given by  $\|f\|_{\mathfrak{L}(E)} \triangleq \sup_{x \neq y} |f(x) - f(y)|/d(x, y)$ , where  $d$  is a consistent metric on  $E$ . The space  $\mathfrak{BL}(E)$  endowed with the norm  $\|\cdot\|_{\mathfrak{BL}(E)}$  is a Banach space, (cf. [5], Proposition 11.2.1).

The weak topology on  $\mathfrak{M}_1(E)$  may be alternatively generated by means of the space  $\mathfrak{BL}(E)$  of bounded Lipschitz functions on  $E$ , instead of the space  $\mathfrak{C}_b(E)$  of bounded continuous functions: the weak topology is also the coarsest topology which renders continuous each of the mappings  $\mu \in \mathfrak{M}_1(E) \mapsto \mu f$ , when  $f$  runs over  $\mathfrak{BL}(E)$ .

We recall that

$$\beta(\mu, \nu) \triangleq \sup\{|\mu f - \nu f| : \|f\|_{\mathfrak{BL}(E)} \leq 1\}, \quad \text{for any } \mu, \nu \in \mathfrak{M}_1^\psi(E), \quad (2)$$

defines a metric on  $\mathfrak{M}_1(E)$  equivalent to the Prohorov metric (1). Hence, given a sequence  $\mu_0, \mu_1, \dots$  in  $\mathfrak{M}_1(E)$ , one has that  $\mu_n \rightarrow \mu_0$  in the weak topology if and only if  $\beta(\mu_n, \mu_0) \rightarrow 0$ , as  $n \rightarrow +\infty$ , (cf. [5] Theorem 11.3.3).

*Remark 1.* Observe that the distance  $\beta$  as well as the Prohorov distance  $\pi$  depend on the distance  $d$ . Besides, if  $d'$  is a metric on  $E$  equivalent to  $d$ , then, with obvious notation, the corresponding distance  $\pi'$  is equivalent to  $\pi$  and the same for  $\beta'$  and  $\beta$ . This allows us to consider the consistent metric  $d$  that turns out to be more useful for our purposes. In particular, among all the distances consistent with the topology on  $E$ , there is one that is *totally bounded*, which will be convenient to use later on; see, e.g., Theorem 2.8.2 in [5]. On the other hand, note that the specific choice of the consistent metric  $d$  does not affect the separability of  $E$ .

*Separability issues.* We recall that the space  $\mathfrak{C}_b(E)$  endowed with the supremum norm  $\|\cdot\|_\infty$  is, in general, not separable. Likewise, the space  $\mathfrak{BL}(E)$  is, in general, not separable for the topology induced by the norm  $\|\cdot\|_{\mathfrak{BL}(E)}$ .

Let us define  $\mathfrak{BL}_1(E)$  to be the unit ball in  $\mathfrak{BL}(E)$ , i.e. the set of functions  $f \in \mathfrak{BL}(E)$  such that  $\|f\|_{\mathfrak{BL}(E)} \leq 1$ . Recall that the family  $\mathfrak{BL}_1(E)$  of  $\mathfrak{BL}(E)$  depends on the actual distance  $d$  on  $E$  used in the definition of the norm  $\|\cdot\|_{\mathfrak{BL}(E)}$ .

The following proposition will play a relevant role later on in this paper.

**Proposition 1.** *If the consistent distance  $d$  on  $E$  is totally bounded, then the unit ball  $\mathfrak{BL}_1(E)$  is separable for the supremum norm  $\|\cdot\|_\infty$ .*

For the proof of Proposition 1 we shall use the following particular case.

**Lemma 1.** *Assume further that the space  $E$  is compact, then  $\mathfrak{BL}_1(E)$  is separable for the supremum norm  $\|\cdot\|_\infty$ .*

Note that the consistent metric  $d$  in Proposition 1 might fail to be complete in general, since completeness is not a topological invariant. In the special case, when  $d$  is indeed complete, then  $E$  turns out to be compact (cf. [5], Theorem 2.3.1), and Proposition 1 boils down to Lemma 1.

*Proof of Lemma 1.* The result follows directly from *Ascoli-Arzelá theorem*, (cf. [5], Theorem 2.4.7). Indeed, since the family  $\mathfrak{BL}_1(E)$  is uniformly bounded and equicontinuous and  $E$  is compact, then the family  $\mathfrak{BL}_1(E)$  is compact with respect to the  $\|\cdot\|_\infty$ -topology, and, consequently, separable.  $\square$

The following proof of Proposition 1 is modelled upon ideas contained in the proof of Theorem 11.4.1 in [5].

*Proof of Proposition 1.* Let  $\bar{E}$  be the completion of  $E$  with respect to the metric  $d$ , and define  $\mathfrak{BL}_1(\bar{E})$  to be the unit ball in  $\mathfrak{BL}(\bar{E})$  defined by the norm  $\|\cdot\|_{\mathfrak{BL}(\bar{E})}$ .

Fix  $g \in \mathfrak{BL}_1(E)$  and denote by  $\bar{g}$  the unique extension of  $g$  (see, e.g. Proposition 11.2.3 in [5]) defined on the entire  $\bar{E}$  such that  $\|\bar{g}\|_{\mathfrak{BL}(\bar{E})} = \|g\|_{\mathfrak{BL}(E)}$  and hence so that  $\bar{g} \in \mathfrak{BL}_1(\bar{E})$ .

Note that  $\bar{E}$  is compact, since  $E$  is assumed to be totally bounded, (cf. [5], Theorem 2.3.1). Lemma 1 gives us a subset  $\mathfrak{N}$  of  $\mathfrak{BL}_1(\bar{E})$  which is countable and dense in  $\mathfrak{BL}_1(\bar{E})$  with respect to the supremum norm  $\|\cdot\|_\infty$ .

Thus, given  $g \in \mathfrak{BL}_1(E)$ , for any  $\varepsilon > 0$  one can find  $f \in \mathfrak{N}$  so that  $\|\bar{g} - f\|_\infty \leq \varepsilon$ . On the other hand, letting  $f|_E$  be the restriction of  $f$  to the domain  $E$ , one has  $\|g - f|_E\|_\infty \leq \|\bar{g} - f\|_\infty \leq \varepsilon$ . Hence, the family  $\mathfrak{N}|_E$  of the functions in  $\mathfrak{N}$  restricted to  $E$  provides a dense and countable subset of  $\mathfrak{BL}_1(E)$  in the norm  $\|\cdot\|_\infty$ .  $\square$

## 2.2. Gauges $\psi$ .

Let  $\psi$  be a continuous function on  $E$ , satisfying  $\psi \geq 1$  everywhere on  $E$ . Throughout the paper,  $\psi$  will play the role of *gauge function*. In particular, following Follmer and Schied, see [7], and also [18, 19], we associate to such  $\psi$  the space of functions  $\mathfrak{C}_\psi(E)$  given by

$$\mathfrak{C}_\psi(E) \triangleq \{f \in \mathfrak{C}(E) : \|f/\psi\|_\infty < \infty\},$$

and the space of probability measures  $\mathfrak{M}_1^\psi(E)$  defined by

$$\mathfrak{M}_1^\psi(E) \triangleq \{\mu \in \mathfrak{M}_1(E) : \mu\psi < +\infty\}.$$

Observe that  $\mathfrak{C}_b(E) \subseteq \mathfrak{C}_\psi(E)$  and that  $\mathfrak{M}_1^\psi(E) \subseteq \mathfrak{M}_1(E)$ .

**Definition 1** ( $\psi$ -weak topology). The  $\psi$ -weak topology  $\sigma(\mathfrak{M}_1^\psi(E), \mathfrak{C}_\psi(E))$  is the coarsest topology on  $\mathfrak{M}_1^\psi(E)$  that renders continuous the maps  $\mu \in \mathfrak{M}_1^\psi(E) \mapsto \mu f$ , varying  $f \in \mathfrak{C}_\psi(E)$ .

Besides the  $\psi$ -weak topology, in  $\mathfrak{M}_1^\psi(E)$  we need to consider also the *relative weak topology* induced on  $\mathfrak{M}_1^\psi(E)$  as a subspace of  $\mathfrak{M}_1(E)$ , endowed with the weak topology as defined above. This relative weak topology is actually  $\sigma(\mathfrak{M}_1^\psi(E), \mathfrak{C}_b(E))$ , the coarsest topology so that for each  $f \in \mathfrak{C}_b(E)$ , the mapping  $\mu \in \mathfrak{M}_1^\psi(E) \mapsto \mu f$  is continuous, see, e.g., Lemma 2.53 in [1]. The  $\psi$ -weak topology is in general finer than the relative weak topology.

If  $\psi \equiv 1$  or simply if  $\psi$  is bounded above, then  $\mathfrak{C}_b(E) = \mathfrak{C}_\psi(E)$ ,  $\mathfrak{M}_1^\psi(E) = \mathfrak{M}_1(E)$  and the  $\psi$ -weak topology and the relative weak topology coincide.

We introduce now a distance  $d_\psi$  on  $\mathfrak{M}_1^\psi(E)$  by

$$d_\psi(\mu, \nu) \triangleq \pi(\mu, \nu) + |(\mu - \nu)\psi|, \quad \text{for any } \mu, \nu \in \mathfrak{M}_1^\psi(E). \quad (3)$$

The following Proposition combines the results of [7] and [18].

**Proposition 2.**  $\mathfrak{M}_1^\psi(E)$  endowed with the  $\psi$ -weak topology is a Polish space and its topology is generated by the distance  $d_\psi$ .

$\mathfrak{M}_1^\psi(E)$  endowed with the relative weak topology is separable.

*Proof.* Corollary A.45 of [7] gives us that  $\mathfrak{M}_1^\psi(E)$  endowed with the  $\psi$ -weak topology is Polish, in particular, metrizable. Now, Lemma 3.4 in [18], gives us that a sequence converges  $\psi$ -weakly if and only if it converges in the distance  $d_\psi$ .

If  $(e_j)_{j \geq 1}$  is a sequence dense in  $E$ , then the family of convex combinations (with rational weights) of  $\delta_{e_j}$  is contained in  $\mathfrak{M}_1^\psi(E)$  and it is dense in  $\mathfrak{M}_1(E)$ , with respect to Prohorov distance. Hence, the space  $\mathfrak{M}_1^\psi(E)$  is separable when endowed with the relative weak topology.  $\square$

*Remark 2.* Note that, given a sequence  $\mu_0, \mu_1, \dots$  in  $\mathfrak{M}_1^\psi(E)$ , then  $\mu_n \rightarrow \mu_0$  in the  $\psi$ -weak topology, as  $n \rightarrow +\infty$ , if and only if  $\mu_n \rightarrow \mu_0$  in the weak topology and  $\mu_n \psi \rightarrow \mu_0 \psi$ , as  $n \rightarrow +\infty$ .

*The measurable structure of  $\mathfrak{M}_1^\psi(E)$ .* We denote by  $\mathcal{M}$  the Borel  $\sigma$ -algebra on  $\mathfrak{M}_1(E)$  generated by the weak topology and by  $\mathcal{M}^\psi$  the Borel  $\sigma$ -algebra on  $\mathfrak{M}_1^\psi(E)$  generated by the  $\psi$ -weak topology.

Next, we collect some properties of  $\mathcal{M}$  and  $\mathcal{M}^\psi$ . We recall that  $\mathcal{M}$  has the following characterization.

**Lemma 2.** The  $\sigma$ -algebra  $\mathcal{M}$  is generated by the projections  $\pi_B : \mu \mapsto \mu(B)$ , defined for  $\mu \in \mathfrak{M}_1(E)$ , letting  $B$  vary in  $\mathcal{E}$ .

*Proof.* See, e.g., Proposition 2.2.2. in [9].  $\square$

**Lemma 3.** *Let  $\mathfrak{H}$  be a family of functions defined on a set  $H$  and taking values in a measurable space  $(G, \mathcal{G})$ . Let  $\phi$  be a  $H$ -valued map defined on some set  $H_0$ , then  $\phi^{-1}(\sigma(\mathfrak{H})) = \sigma(\mathfrak{H} \circ \phi)$  on  $H_0$ , where  $\mathfrak{H} \circ \phi \triangleq \{h \circ \phi : h \in \mathfrak{H}\}$ .*

*Proof.* First of all, note that  $\phi^{-1}(\sigma(\mathfrak{H}))$  is a  $\sigma$ -algebra on  $H_0$ , since the map  $\phi^{-1}$  preserves all the set operations. Thus, the inclusion  $\sigma(\mathfrak{H} \circ \phi) \subseteq \phi^{-1}(\sigma(\mathfrak{H}))$  is immediate, since  $h \circ \phi$  is  $\phi^{-1}(\sigma(\mathfrak{H}))$ -measurable for any  $h \in \mathfrak{H}$ .

Let now  $\mathcal{H}_0$  be a  $\sigma$ -algebra on  $H_0$  with respect to which  $h \circ \phi$  is  $(\mathcal{H}_0, \mathcal{G})$ -measurable, for any  $h \in \mathfrak{H}$ . Clearly  $\phi^{-1}(\sigma(\mathfrak{H})) \subseteq \mathcal{H}_0$ . Thus, the proof concludes by considering  $\mathcal{H}_0 = \sigma(\mathfrak{H} \circ \phi)$ .  $\square$

The next lemma states that the relative weak topology and the  $\psi$ -weak topology generate the same Borel  $\sigma$ -algebra on  $\mathfrak{M}_1^\psi(E)$ .

**Lemma 4.** *The  $\sigma$ -algebra  $\mathcal{M}^\psi$  is generated by the relative weak topology  $\sigma(\mathfrak{M}_1^\psi(E), \mathfrak{C}_b(E))$ .*

Recall that a  $\sigma$ -algebra is said to be (i) *countably generated* if it is generated by a countable family of sets and (ii) *countably separated* if it admits a countable family of sets separating points. Moreover, a measurable space is said to be *standard* if it is Borel-isomorphic to a Polish space.

*Proof of Lemma 4.* Let us define  $\mathcal{B}^\psi$  to be the Borel  $\sigma$ -algebra associated to the relative weak topology  $\sigma(\mathfrak{M}_1^\psi(E), \mathfrak{C}_b(E))$ . Since the  $\psi$ -weak topology is finer than the weak topology on  $\mathfrak{M}_1^\psi(E)$ , one has  $\mathcal{B}^\psi \subseteq \mathcal{M}^\psi$ .

On the other hand, Proposition 2 gives us that  $\mathfrak{M}_1^\psi(E)$  is separable when endowed with the relative weak topology. As a result, the  $\sigma$ -algebra  $\mathcal{B}^\psi$  is countably generated and countably separated. Indeed, any countable base  $\mathfrak{B}^\psi$  of open sets in  $\sigma(\mathfrak{M}_1^\psi(E), \mathfrak{C}_b(E))$  generates the  $\sigma$ -algebra  $\mathcal{B}^\psi$  and separates points, (cf. [2], §6.5).

Finally, Proposition 2 again implies that the space  $(\mathfrak{M}_1^\psi(E), \mathcal{M}^\psi)$  is standard, and thus, the  $\sigma$ -algebra  $\mathcal{M}^\psi$  coincides with  $\sigma(\mathfrak{B}^\psi)$ , thanks to Theorem 3.3 in [20].  $\square$

The Borel  $\sigma$ -algebra  $\mathcal{M}^\psi$  admits a characterization in terms of the projections  $\pi_B$  analogous to that of Lemma 2 for  $\mathcal{M}$ . This is the content of the next proposition.

**Proposition 3.** *The Borel  $\sigma$ -algebra  $\mathcal{M}^\psi$  is generated by the projections  $\pi_B : \mu \mapsto \mu(B)$ , defined for  $\mu \in \mathfrak{M}_1^\psi(E)$ , letting  $B$  vary in  $\mathcal{E}$ .*

*Proof.* Let  $\phi : \mathfrak{M}_1^\psi(E) \hookrightarrow \mathfrak{M}_1(E)$  be the inclusion of  $\mathfrak{M}_1^\psi(E)$  into  $\mathfrak{M}_1(E)$  and define  $\mathfrak{H}$  to be the family consisting of the projection maps  $\pi_B : \mu \in \mathfrak{M}_1(E) \rightarrow \mu(B)$ , letting  $B$  vary in  $\mathcal{E}$ . The family  $\mathfrak{H} \circ \phi \triangleq \{\pi_B \circ \phi : B \in \mathcal{E}\}$  consists of the projections defined on  $\mathfrak{M}_1^\psi(E)$ .

Let us define  $\mathcal{B}^\psi$  to be the Borel  $\sigma$ -algebra generated by the restriction  $\sigma(\mathfrak{M}_1^\psi(E), \mathfrak{C}_b(E))$  of the weak topology to  $\mathfrak{M}_1^\psi(E)$ . Equality  $\mathcal{B}^\psi = \phi^{-1}(\sigma(\mathfrak{H}))$  holds true, since  $\sigma(\mathfrak{H}) = \mathcal{M}$  due to Lemma 2 and  $\mathcal{B}^\psi = \phi^{-1}(\mathcal{M})$ . Hence, applying Lemma 3, one deduces that  $\mathcal{B}^\psi = \sigma(\mathfrak{H} \circ \phi)$ . The stated result now follows from Lemma 4.  $\square$

### 3. Setup

We now introduce our reference probability space  $(\Omega, \mathcal{F}, \mathbb{P})$  for our asymptotic stability results.

We let  $\Omega$  denote the set  $\Omega = E^{\mathbb{N}}$  of all the sequences  $\omega = (\omega_1, \omega_2, \dots)$  of elements of a Polish space  $E$ . The projections  $\xi_1, \xi_2, \dots$  are the mappings  $\omega \mapsto \xi_n(\omega) \triangleq \omega_n$ , for  $\omega \in \Omega$  and  $n \geq 1$ .

We let  $\mathcal{F}$  denote the  $\sigma$ -algebra in  $\Omega$  generated by the projections  $\xi_1, \xi_2, \dots$ . This family  $\mathcal{F}$  is also the Borel  $\sigma$ -algebra associated to the product topology in  $E^{\mathbb{N}}$ ; it also coincides, since  $E$  is separable, with the product  $\sigma$ -algebra  $\mathcal{E}^{\mathbb{N}}$ , (cf. [23], Theorem 1.10).

We let  $\mathbb{P}$  be a probability measure defined on  $(\Omega, \mathcal{F})$ , so that  $(\Omega, \mathcal{F}, \mathbb{P})$  is a complete probability space.

*Random measures.* By a random measure  $\chi$  on  $(E, \mathcal{E})$  with support in  $\mathfrak{M}_1^\psi(E)$  we understand a probability kernel

$$(\omega, B) \in \Omega \times \mathcal{E} \mapsto \chi(\omega, B) \in [0, 1],$$

such that

- (i) the assignment  $B \in \mathcal{E} \mapsto \chi(\omega, B)$  defines a probability measure in  $\mathfrak{M}_1^\psi(E)$ , for each fixed  $\omega \in \Omega$ ,
- (ii) the mapping  $\omega \in \Omega \mapsto \chi(\omega, B)$  is  $\mathcal{F}$ -measurable, for each fixed  $B \in \mathcal{E}$ .

Besides, Proposition 3 allows to understand  $\chi$  as a random variable on  $(\Omega, \mathcal{F}, \mathbb{P})$  and taking values in  $(\mathfrak{M}_1^\psi(E), \mathcal{M}^\psi)$ .

We shall denote by  $\mathcal{L}(\chi)$  the distribution induced by  $\chi$  on  $(\mathfrak{M}_1^\psi(E), \mathcal{M}^\psi)$  as a pullback in the usual way:

$$\mathcal{L}(\chi)(M) = \mathbb{P} \circ \chi^{-1}(M), \quad \text{for any } M \in \mathcal{M}^\psi.$$

*Empirical process.* The empirical process associated to the projections  $\xi$  is the sequence  $m_1, m_2, \dots$  of random measures defined, for each  $n \geq 1$ , by

$$m_n(\omega, B) \triangleq \frac{1}{n} \sum_{i=1}^n \delta_{\xi_i(\omega)}(B), \quad \text{for any } \omega \in \Omega \text{ and any } B \in \mathcal{E}. \quad (4)$$

Moreover, we say that the empirical process  $m_1, m_2, \dots$  is directed by the variables  $\xi_1, \xi_2, \dots$ . We shall always understand each  $m_n$  as a random variable defined on  $(\Omega, \mathcal{F}, \mathbb{P})$  and with values in  $(\mathfrak{M}_1^\psi(E), \mathcal{M}^\psi)$ .

*Statistics and estimators.* Let  $T$  be a further Polish space endowed with its Borel  $\sigma$ -algebra  $\mathcal{T}$  and with a metric  $d_T$  which induces its topological structure.

Any  $(\mathcal{M}^\psi, \mathcal{T})$ -measurable functional  $\tau : \mathfrak{M}_1^\psi(E) \rightarrow T$  is termed a *statistic* on  $\mathfrak{M}_1^\psi(E)$ .

The sequence of random variables  $\tau_1, \tau_2, \dots$  from  $(\Omega, \mathcal{F}, \mathbb{P})$  into  $(T, \mathcal{T})$  obtained by setting, for each  $n \geq 1$ ,

$$\tau_n \triangleq \tau(m_n),$$

is called the sequence of *estimators* induced by  $\tau$ .

The statistic  $\tau : \mathfrak{M}_1^\psi(E) \rightarrow T$  is said to be  $\psi$ -continuous if it is continuous with respect to the  $\psi$ -weak topology on  $\mathfrak{M}_1^\psi(E)$  and the topology defined on  $T$ . Besides,  $\tau$  is said to be *uniformly*  $\psi$ -continuous if for any  $\varepsilon > 0$  there exists  $\delta(\varepsilon) > 0$  such that  $d_T(\tau(\mu_1), \tau(\mu_2)) < \varepsilon$  if  $d_\psi(\mu_1, \mu_2) < \delta(\varepsilon)$ . Note that, given a random measure  $\chi$  on  $(E, \mathcal{E})$  with support in  $\mathfrak{M}_1^\psi(E)$  and a  $\psi$ -continuous statistic  $\tau$  on  $\mathfrak{M}_1^\psi(E)$ , the composition  $\tau(\chi)$  is  $(\mathcal{F}, \mathcal{T})$ -measurable.

**Definition 2** (Strong Consistency). Given a statistic  $\tau : \mathfrak{M}_1^\psi(E) \rightarrow T$  and a random measure  $\chi$  on  $(E, \mathcal{E})$ , we say that  $\tau$ , or equivalently, the sequence of estimators  $(\tau_n)_n$  induced by  $\tau$ , is *strongly consistent* for  $\tau(\chi)$  if one has  $\mathbb{P}$ -almost surely that  $\tau_n \rightarrow \tau(\chi)$ , as  $n \rightarrow +\infty$ .

#### 4. Consistency

Let  $(\Omega, \mathcal{F}, \mathbb{P})$  be the complete probability space introduced in the previous section.

*Stationarity.* We denote by  $\Sigma$  the shift operator on  $E^\mathbb{N}$ , i.e.

$$\Sigma(x_1, x_2, \dots) = (x_2, x_3, \dots), \quad \text{for any } (x_1, x_2, \dots) \in E^\mathbb{N}.$$

The random sequence  $\xi \triangleq (\xi_1, \xi_2, \dots)$  in  $(E, \mathcal{E})$  given by the canonical projections is said to be *stationary* if one has  $\mathcal{L}(\xi) = \mathcal{L}(\Sigma\xi)$ .

Here and in what follows,  $\mathcal{L}(\xi) \triangleq \mathbb{P} \circ \xi^{-1}$  denotes the distribution on  $(E^{\mathbb{N}}, \mathcal{E}^{\mathbb{N}})$  induced by the random sequence  $\xi$ .

*The shift invariant  $\sigma$ -algebra.* The *shift invariant  $\sigma$ -algebra* is defined to be the collection  $\mathcal{I}$  of the Borel sets  $I \in \mathcal{E}^{\mathbb{N}}$  such that  $\Sigma^{-1}(I) = I$ , and plays a crucial role in what follows.

Observe that if the variables  $\xi_n$  are i.i.d, then the  $\sigma$ -algebra  $\mathcal{I}$  turns out to be  $\mathbb{P}$ -trivial, see, e.g., Corollary 1.6 in [16].

*Canonical random measure.* We now introduce the canonical random measure associated to the sequence  $\xi$ .

**Lemma 5.** *There exists an essentially unique regular version  $\nu$  of the conditional distribution  $\mathbb{P}[\xi_1 \in \cdot | \mathcal{I}]$ .*

Recall that  $\nu$  is by definition a  $\mathcal{I}$ -measurable random probability measure over  $(E, \mathcal{E})$ , i.e. in particular a probability kernel  $(\omega, B) \mapsto \nu(\omega, B)$  from the probability space  $(\Omega, \mathcal{F}, \mathbb{P})$  to  $(E, \mathcal{E})$ . In other terms,  $\nu$  is what we have termed a random measure. We refer to §10.4 of [2] for background and relevance of regular version of conditional distributions.

*Proof of Lemma 5.* See, e.g., Lemma 10.4.3 and Corollary 10.4.6 in [2]. Recall that  $E$  is Polish and  $\mathcal{E}$  is Borel, hence countably generated.  $\square$

In the remainder of this paper, we refer to  $\nu$  as the *canonical random measure* associated to  $\xi$ .

Note that, in the case when the projections  $\xi_n$  are independent with common distribution  $\mu$ , so that  $\mathcal{L}(\xi_n) = \mu$  for each  $n \geq 1$ , then  $\mathbb{P}$ -almost surely  $\nu = \mu$ . This is so because  $\mathcal{I}$  is  $\mathbb{P}$ -trivial and then

$$\mathbb{P}[\xi_1 \in \cdot | \mathcal{I}] = \mathbb{P}[\xi_1 \in \cdot] = \mu, \quad \mathbb{P}\text{-a.s.} \quad (5)$$

Observe that when  $\mathcal{L}(\xi_1) \in \mathfrak{M}_1^\psi(E)$ , i.e. when  $\int_E \psi(x) \mathbb{P} \circ \xi_1^{-1}(dx) < +\infty$ , we have  $\mathbb{P}$ -almost surely that  $\nu \in \mathfrak{M}_1^\psi(E)$ .

In next paragraphs we address the convergence of estimators.

When the random variables  $\xi_1, \xi_2, \dots$  are independent and identically distributed, Varadarajan's theorem (which we record below as Proposition 4) asserts the convergence in the weak topology of the empirical process.

**Proposition 4.** *Assume that the variables  $\xi_1, \xi_2, \dots$  are independent with common law  $\mu \in \mathfrak{M}_1(E)$ , then,  $\mathbb{P}$ -almost surely  $m_n \rightarrow \mu$  in the weak topology, as  $n \rightarrow \infty$ .*

*Proof.* See, e.g., Theorem 11.4.1 in [5]. □

Analogously, for the  $\psi$ -weak topology in  $\mathfrak{M}_1^\psi(E)$  we also have the following Varadarajan type theorem, which will be proved later by using Propositions 5 and 6.

**Theorem 1.** *If  $\xi$  is stationary and such that  $\mathcal{L}(\xi_1) \in \mathfrak{M}_1^\psi(E)$ , then  $\mathbb{P}$ -almost surely  $m_n \rightarrow \nu$  in the  $\psi$ -weak topology, as  $n \rightarrow +\infty$ .*

Recall that a family  $\mathfrak{G}$  of Borel functions on  $E$  is said to be an *universal Glivenko-Cantelli class* if

$$\sup\{|(m_n - \mu)f| : f \in \mathfrak{G}\} \rightarrow 0, \quad \mathbb{P}\text{-a.s. as } n \rightarrow +\infty,$$

whenever the variables  $\xi_1, \xi_2, \dots$  that direct the empirical process  $(m_n)_n$  are independent with common generic distribution  $\mu \in \mathfrak{M}_1(E)$ .

**Proposition 5.** *The unit ball  $\mathfrak{BL}_1(E)$  constitutes an universal Glivenko-Cantelli class.*

*Proof.* Firstly, if  $\xi_1, \xi_2, \dots$  are independent with common generic distribution  $\mu \in \mathfrak{M}_1(E)$ , then, according to Lemma 4,  $\mathbb{P}$ -almost surely  $m_n \rightarrow \mu$  in the weak topology as  $n \rightarrow +\infty$ . Therefore, as discussed in Section 2.1, we obtain that  $\beta(m_n, \mu) \rightarrow 0$ , as  $n \rightarrow +\infty$ . □

**Proposition 6.** *If  $\xi$  is stationary, then  $\mathbb{P}$ -almost surely  $m_n \rightarrow \nu$  in the weak topology, as  $n \rightarrow +\infty$ .*

*Proof.* According to Remark 1, we now use a totally bounded metric on  $E$  to define the norm  $\|\cdot\|_{\mathfrak{BL}_1(E)}$ . The unit ball  $\mathfrak{BL}_1(E)$  is an uniformly bounded family of Borel functions on  $E$ . Moreover, it is separable for the supremum norm, due to Proposition 1.

Thus, since  $\mathfrak{BL}_1(E)$  forms an universal Glivenko-Cantelli class according to Proposition 5, Theorem 1.3 combined to Corollary 1.4 in [24] apply, and in particular we have

$$\sup\{|(m_n - \nu)f| : f \in \mathfrak{BL}_1(E)\} \rightarrow 0, \quad \mathbb{P}\text{-a.s. as } n \rightarrow +\infty \quad (6)$$

The proof now concludes since (6) gives that  $\beta(m_n, v) \rightarrow 0$  almost surely, as  $n \rightarrow +\infty$ , and so that  $\mathbb{P}$ -almost surely  $m_n \rightarrow v$  in the weak topology, as  $n \rightarrow +\infty$ .  $\square$

*Proof of Theorem 1.* According to Proposition 6, we have that  $\mathbb{P}$ -almost surely  $m_n \rightarrow v$  in the relative weak topology, as  $n \rightarrow +\infty$ . Thus, in view of Remark 2, it remains to show that

$$|m_n \psi - v \psi| \rightarrow 0, \quad \mathbb{P}\text{-almost surely as } n \rightarrow +\infty.$$

We apply now *von Neumann's version of Birkhoff's ergodic theorem* as stated in [15], Theorem 9.6. Using the notation therein, consider as space  $S \triangleq \Omega = E^{\mathbb{N}}$ , as transformation  $T \triangleq \Sigma$ , the shift operator defined on it, and as measurable function  $f \triangleq \psi \circ \xi_1$ .

Notice that the shift operator  $T$  preserves the measure  $\mathbb{P}$  since  $\xi$  is assumed to be stationary. If we write  $\mu$  for the common law  $\mu = \mathcal{L}(\xi_1)$ , then a change of variables gives that

$$\int_S f d\mathbb{P} = \int_{\Omega} \psi(\xi_1) d\mathbb{P} = \int_E \psi(x) \mathbb{P} \circ \xi_1^{-1}(dx) = \int_E \psi d\mu.$$

Since by hypothesis  $\mathcal{L}(\xi_1) = \mu \in \mathfrak{M}_1^{\psi}(E)$ , we have  $\int_E \psi d\mu = \mu\psi < +\infty$  and therefore  $f \in L^1(\Omega, \mathcal{F}, \mathbb{P})$ .

Hence, we conclude that

$$m_n \psi \rightarrow \mathbb{E}[\psi(\xi_1) | \mathcal{I}], \quad \mathbb{P}\text{-almost surely as } n \rightarrow +\infty.$$

Finally, according to the *disintegration theorem* (cf. [15], Theorem 5.4), we may recast the limit variable and write

$$\mathbb{E}[\psi(\xi_1) | \mathcal{I}] = \int_E \psi d\mathbb{P}[\xi_1 \in \cdot | \mathcal{I}] = v\psi, \quad \mathbb{P}\text{-almost surely.}$$

$\square$

*Remark 3 (Ergodicity).* Under the same hypotheses of Theorem 1, but assuming additionally that  $\xi$  is ergodic, or equivalently that its distribution  $\mathcal{L}(\xi)$  is ergodic with respect to the shift operator  $\Sigma$ , i.e.  $\mathbb{P}\{\xi \in I\} \in \{0, 1\}$  for any  $I \in \mathcal{I}$ , we easily get that  $\mathbb{P}$ -almost surely the empirical process  $(m_n)_n$  converges to  $\mu = \mathcal{L}(\xi_1)$  in the  $\psi$ -weak topology, as  $n \rightarrow +\infty$ , since the  $\sigma$ -field  $\xi^{-1}\mathcal{I}$  turns out to be  $\mathbb{P}$ -trivial in such a case.

The next result is an immediate consequence of Theorem 1.

**Corollary 1** (*Strong Consistency*). *If  $\xi$  is stationary such that  $\mathcal{L}(\xi_1) \in \mathfrak{M}_1^\psi(E)$  and  $\tau : \mathfrak{M}_1^\psi(E) \rightarrow T$  is  $\psi$ -continuous, then the sequence of estimators  $(\tau_n)_n$  is strongly consistent for  $\tau(v)$ .*

*Remark 4.* The proofs of Theorem 1 and Corollary 1 are in line with the one of Theorem 2.6 in [19], which further assume the sequence  $\xi$  to be ergodic. In particular, within the cited work the role of the map  $\tau$  in Corollary 1 is recovered by a law-invariant convex risk measure and Theorem 6.6 in [23] is considered instead of Corollary 1.4 in [24].

Suppose that  $\xi$  describes the outcome in a sequence of trials. According to the present framework, any element of the sample space  $\Omega \triangleq E^{\mathbb{N}}$  may be understood as a path of  $\xi$ . On the other hand, the random measure  $v$  is a regular version of the distribution induced by the single variable within the process  $\xi$ , conditioned on the information encoded by the shift invariant  $\sigma$ -algebra  $\mathcal{S}$ . In this respect, each  $\omega \in \Omega$  completely describes the limit distribution  $v(\omega, \cdot)$ , which may be understood as the best available description of the outcomes.

## 5. Asymptotic stability

Let  $\theta$  be a  $\mathcal{E}$ -measurable endomorphism over  $E$ , i.e. a  $(\mathcal{E}, \mathcal{E})$ -measurable map from  $E$  to itself, and set

$$\lambda_{\mathbb{P},\theta}(\alpha) \triangleq \mathbb{P}\{d_\psi(v, v \circ \theta^{-1}) > \alpha\}, \quad \text{for any } \alpha > 0, \quad (7)$$

where the random variable  $v \circ \theta^{-1}$  is defined by

$$(\omega, B) \in \Omega \times \mathcal{E} \mapsto (v \circ \theta^{-1})(\omega, B) \triangleq v(\omega, \theta^{-1}(B)). \quad (8)$$

Note that the function (7) is well defined, since  $d_\psi$  is trivially  $(\mathcal{M}^\psi \otimes \mathcal{M}^\psi)$ -measurable. Moreover, observe that  $\lambda_\theta$  is a decreasing function in  $\alpha > 0$  and that  $\lambda_{\mathbb{P},\theta}(\alpha) \rightarrow 1$  as  $\alpha \rightarrow 0$  and  $\lambda_{\mathbb{P},\theta}(\alpha) \rightarrow 0$  as  $\alpha \rightarrow +\infty$ , via monotonicity arguments.

Assume that the statistic  $\tau$  is uniformly  $\psi$ -continuous and that  $\kappa$  is a modulus of continuity of  $\tau$ , i.e. a continuous strictly increasing map  $[0, +\infty] \rightarrow [0, +\infty]$  such that

$$d_\tau(\tau(\mu_1), \tau(\mu_2)) \leq \kappa(d_\psi(\mu_1, \mu_2)), \quad \text{for any } \mu_1, \mu_2 \in \mathfrak{M}_1^\psi(E).$$

Note that, since  $\lambda_{\mathbb{P},\theta}(\alpha) \leq \kappa(\alpha)$  for  $\alpha$  large enough, we are allowed to set

$$\|\theta\|_{\mathbb{P},\kappa} \triangleq \inf\{\alpha > 0 : \lambda_{\mathbb{P},\theta}(\alpha) \leq \kappa(\alpha)\}. \quad (9)$$

*Example 1.* Assume  $\xi$  to be stationary and ergodic, i.e.  $\mathbb{P}\{\xi \in I\} \in \{0, 1\}$  for any  $I \in \mathcal{I}$ . As a consequence, the sequence  $\theta(\xi_1), \theta(\xi_2), \dots$  turns out to be stationary and ergodic (see Lemma 9.1 combined with Lemma 9.5 in [15]) and hence

$$\begin{aligned} v &= \mathbb{P}[\xi_1 \in \cdot | \mathcal{I}] = \mathbb{P} \circ \xi_1^{-1} \quad \text{a.s.} \\ v \circ \theta^{-1} &= \mathbb{P}[\theta(\xi_1) \in \cdot | \mathcal{I}] = \mathbb{P} \circ \theta(\xi_1)^{-1} \quad \text{a.s.} \end{aligned}$$

Next, assume  $\mathbb{P} \circ \xi_1^{-1} = \delta_{x_1}$  for some  $x_1 \in E$ , and suppose that  $\theta(x_1) = x_2$ , which implies  $\mathbb{P} \circ \theta(\xi_1)^{-1} = \delta_{x_2}$ . Then, one has  $d_\psi(v, v \circ \theta^{-1}) = d_\psi(\delta_{x_1}, \delta_{x_2})$  a.s. where

$$d_\psi(\delta_{x_1}, \delta_{x_2}) = \min\{d_E(x_1, x_2), 1\} + |\psi(x_1) - \psi(x_2)| \quad \text{a.s.}$$

and as a direct consequence, the following bound holds

$$\|\theta\|_{\mathbb{P},\kappa} \leq \min\{d_E(x_1, x_2), 1\} + |\psi(x_1) - \psi(x_2)|.$$

**Lemma 6.** *If  $\tau$  is uniformly  $\psi$ -continuous and it admits  $\kappa$  as modulus of continuity, then  $\pi(\mathbb{P} \circ \tau(v)^{-1}, \mathbb{P} \circ \tau(v \circ \theta^{-1})^{-1}) \leq \kappa(\|\theta\|_{\mathbb{P},\kappa})$ .*

*Proof.* Let  $C \in \mathcal{T}$  and fix  $\alpha > 0$  such that  $\lambda_{\mathbb{P},\theta}(\alpha) \leq \kappa(\alpha)$ . Since  $\tau$  is  $\psi$ -continuous, then  $\tau^{-1}(C) \in \mathcal{M}^\psi$ . In particular, for any  $A \in \mathcal{M}^\psi$ , we denote by  $A^\varepsilon \triangleq \{\mu \in \mathfrak{M}_1^\psi(E) : d_\psi(\mu, \nu) \leq \varepsilon, \text{ for some } \nu \in A\}$  the  $\varepsilon$ -hull of  $A$  defined in terms of the metric  $d_\psi$ .

Notice that  $[\tau^{-1}(C)]^\alpha \subseteq \tau^{-1}(C^{\kappa(\alpha)})$  in  $\mathfrak{M}_1^\psi(E)$ , since  $\tau$  is uniformly  $\psi$ -continuous and admits  $\kappa$  as modulus of continuity, where the  $\kappa(\alpha)$ -hull  $C^{\kappa(\alpha)}$  of  $C$  is defined in terms of the metric  $d_T$ . Hence,  $v \circ \theta^{-1} \in [\tau^{-1}(C)]^\alpha$  implies  $v \circ \theta^{-1} \in \tau^{-1}(C^{\kappa(\alpha)})$ , and in particular one has that  $\mathbb{P}\{v \circ \theta^{-1} \in [\tau^{-1}(C)]^\alpha\} \leq \mathbb{P} \circ \tau(v \circ \theta^{-1})^{-1}(C^{\kappa(\alpha)})$ . Thus,

$$\begin{aligned} \mathbb{P} \circ \tau(v)^{-1}(C) &\leq \mathbb{P}\{d_\psi(v, v \circ \theta^{-1}) > \alpha\} + \mathbb{P}\{v \circ \theta^{-1} \in [\tau^{-1}(C)]^\alpha\} \\ &\leq \kappa(\alpha) + \mathbb{P} \circ \tau(v \circ \theta^{-1})^{-1}(C^{\kappa(\alpha)}). \end{aligned}$$

Then, since the choice of  $C \in \mathcal{T}$  is arbitrary, one has that

$$\pi(\mathbb{P} \circ \tau(v)^{-1}, \mathbb{P} \circ \tau(v \circ \theta^{-1})^{-1}) \leq \kappa(\alpha).$$

The proof is concluded by letting  $\alpha$  tend to  $\|\theta\|_{\mathbb{P}, \kappa}$ , while invoking the continuity of  $\kappa$ .  $\square$

*Remark 5.* Note that, if  $\kappa$  is defined to be the identity on  $(0, +\infty)$ , then (9) boils down to the *Ky Fan distance* between  $v$  and  $v \circ \theta$ , which are understood as random variables with values in  $(\mathfrak{M}_1^\psi(E), \mathcal{M}^\psi)$ , (cf. [5], §9.2). In particular, when looking at Lemma 6, this is the case when  $\tau$  is a contraction.

**Theorem 2** (Asymptotic stability). *Let  $\xi$  be stationary such that  $\mathcal{L}(\xi_1) \in \mathfrak{M}_1^\psi(E)$  and assume that  $\mathcal{L}(\theta(\xi_1)) \in \mathfrak{M}_1^\psi(E)$ . If  $\tau$  is uniformly  $\psi$ -continuous and it admits  $\kappa$  as modulus of continuity, then*

$$\limsup_{n \geq 1} \pi(\mathbb{P} \circ \tau(m_n)^{-1}, \mathbb{P} \circ \tau(m_n \circ \theta^{-1})^{-1}) \leq \kappa(\|\theta\|_{\mathbb{P}, \kappa}). \quad (10)$$

*Proof.* By the triangle inequality,

$$\begin{aligned} \pi(\mathbb{P} \circ \tau(m_n)^{-1}, \mathbb{P} \circ \tau(m_n \circ \theta^{-1})^{-1}) &\leq \\ &\pi(\mathbb{P} \circ \tau(m_n)^{-1}, \mathbb{P} \circ \tau(v)^{-1}) \\ &+ \pi(\mathbb{P} \circ \tau(v)^{-1}, \mathbb{P} \circ \tau(v \circ \theta^{-1})^{-1}) \\ &+ \pi(\mathbb{P} \circ \tau(v \circ \theta^{-1})^{-1}, \mathbb{P} \circ \tau(m_n \circ \theta^{-1})^{-1}). \end{aligned} \quad (11)$$

Since  $\tau$  is uniformly  $\psi$ -continuous and it admits  $\kappa$  as modulus of continuity, Lemma 6 applies. Thus, we deduce from inequality (11) that

$$\begin{aligned} \limsup_{n \geq 1} \pi(\mathbb{P} \circ \tau(m_n)^{-1}, \mathbb{P} \circ \tau(m_n \circ \theta^{-1})^{-1}) &\leq \kappa(\|\theta\|_{\mathbb{P}, \kappa}) \\ &+ \limsup_{n \geq 1} \pi(\mathbb{P} \circ \tau(m_n)^{-1}, \mathbb{P} \circ \tau(v)^{-1}) \\ &+ \limsup_{n \geq 1} \pi(\mathbb{P} \circ \tau(m_n \circ \theta^{-1})^{-1}, \mathbb{P} \circ \tau(v \circ \theta^{-1})^{-1}). \end{aligned}$$

Since  $\xi$  is assumed to be stationary such that  $\mathcal{L}(\xi_1) \in \mathfrak{M}_1^\psi(E)$  and  $\tau$  is  $\psi$ -continuous, the result described in Corollary 1 guarantees that  $\mathbb{P}$ -almost surely  $\tau(m_n) \rightarrow \tau(v)$ , as  $n \rightarrow +\infty$ , and hence

$$\limsup_{n \geq 1} \pi(\mathbb{P} \circ \tau(m_n)^{-1}, \mathbb{P} \circ \tau(v)^{-1}) = 0.$$

On the other hand, the sequence  $\theta(\xi_1), \theta(\xi_2), \dots$  is stationary since  $\xi$  is stationary (see Lemma 9.1 in [15]). Moreover, note that

$$m_n \circ \theta^{-1} \triangleq \frac{1}{n} \sum_{i \leq n} \delta_{\xi_i} \circ \theta^{-1} = \frac{1}{n} \sum_{i \leq n} \delta_{\theta(\xi_i)}.$$

Then, since  $\mathcal{L}(\theta(\xi_1)) \in \mathfrak{M}_1^\psi(E)$ , Corollary 1 guarantees that  $\mathbb{P}$ -almost surely  $\tau(m_n \circ \theta^{-1}) \rightarrow \tau(v_\theta)$ , as  $n \rightarrow +\infty$ , where we write  $v_\theta$  to denote a regular version of the conditional distribution  $\mathbb{P}[\theta(\xi_1) \in \cdot | \mathcal{S}]$ . On the other hand, notice that  $\mathbb{P}$ -almost surely  $\mathbb{P}[\theta(\xi_1) \in \cdot | \mathcal{S}] = \mathbb{P}[\xi_1 \in \theta^{-1}(\cdot) | \mathcal{S}]$  and hence  $v_\theta = v \circ \theta^{-1}$ . Thus,

$$\limsup_{n \geq 1} \pi(\mathbb{P} \circ \tau(m_n \circ \theta^{-1})^{-1}, \mathbb{P} \circ \tau(v \circ \theta^{-1})^{-1}) = 0.$$

□

*Remark 6.* In the case when  $\xi$  describes the outcomes in a sequence of trials, we may understand the action of the endomorphism  $\theta$  as a perturbation of the available dataset and the function  $\lambda_\theta$  defined in (7) measures the impact of such a perturbation in terms of the random measure  $v$ .

It is easy to realize that, when the perturbation procedure encoded by the action of the map  $\theta$  does not change appreciably the random measure  $v$  in the stochastic sense provided by (7), then one should expect  $\|\theta\|_{\mathbb{P}, \kappa}$  to be small. In particular, this form of continuity is properly assessed in terms of  $\kappa$ .

In the particular case when  $\tau$  admits  $\kappa$  as modulus of continuity, Theorem 2 guarantees that any changes in the law of  $\tau_n$  due to small perturbations at the level of the dataset encoded by the map  $\theta$  are asymptotically gauged by the relation described in (10) by means of the terms  $\kappa(\|\theta\|_{\mathbb{P}, \kappa})$ .

For instance, **fix a function  $f$  such that  $f - \psi \in \mathfrak{BL}_1(E)$**  and consider  $\tau_f$  to be the functional on  $\mathfrak{M}_1^\psi(E)$  defined by setting  $\mu \in \mathfrak{M}_1^\psi(E) \mapsto \tau_f(\mu) \triangleq \mu f$ . Notice that

$$|\tau_f(\mu) - \tau_f(\nu)| \leq \beta(\mu, \nu) + |(\mu - \nu)\psi| \leq C d_\psi(\mu, \nu), \quad \text{for any } \mu, \nu \in \mathfrak{M}_1^\psi(E),$$

for some positive constant  $C$ , since the distance  $\beta$  as given in (2) is equivalent to Prohorov metric  $\pi$ . **Hence,  $\tau$  is uniformly  $\psi$ -continuous and it admits**

$\kappa(\alpha) \triangleq C\alpha$ , for  $\alpha \geq 0$ , as **modulus of continuity**. According to Theorem 2, the changes in the law of the sample mean of  $f$ ,

$$m_n f = \frac{1}{n} \sum_{i=1}^n f(\xi_i), \quad (12)$$

with respect to the metric  $\pi$ , when considering the sequence  $\theta(\xi_1), \theta(\xi_2), \dots$  are asymptotically controlled by the term

$$\|\theta\|_{\mathbb{P}, \kappa} = \inf\{\alpha > 0 : \lambda_{\mathbb{P}, \theta}(\alpha) \leq C\alpha\}.$$

*Remark 7* (Qualitative Robustness). Theorem 2 and its interpretation in Remark 6 are in line with the notion of qualitative robustness as discussed in Krättschmer et al. [18, 19] and Zähle [25, 26]. Following these cited authors, any statistic  $\tau$  is said to be qualitative robust if small changes of the law related to the outcomes  $\xi_n$  only result in small changes of the distribution characterizing the estimators  $\tau_n$ , for  $n$  large enough.

*Remark 8* (Asymptotic Stability and Elicitability). *Elicitability* provides a widely discussed aspect in evaluating point forecasts; for background see for instance [6, 10, 22, 27]. In this respect, assume that the statistic  $\tau$  is *elicitable*, relative to the class  $\mathfrak{M}_1^\psi(E)$ , by considering some *strictly consistent scoring function*  $S : T \times E \rightarrow [0, +\infty)$ . Moreover assume that  $\tau$  is uniformly continuous with respect to the functional  $(\mu, \nu) \mapsto \tilde{S}(\mu, \nu) \triangleq \int_E S(\tau(\mu), x) \nu(dx)$  in the sense that

$$d_T(\tau(\mu), \tau(\nu)) \leq \kappa(\tilde{S}(\mu, \nu)), \quad \text{for any } \mu, \nu \in \mathfrak{M}_1^\psi(E), \quad (13)$$

for some non-negative continuous and increasing function  $\kappa$  vanishing at zero.

Recall that  $\|\theta\|_{\mathbb{P}, \kappa}$  as defined in (9) implicitly depends on the metric  $d_\psi$ . In a similar way, if  $\tilde{S}$  is  $(\mathcal{M}^\psi \otimes \mathcal{M}^\psi)$ -measurable, we may define

$$\|\theta\|_{\mathbb{P}, \kappa}^{(1)} \triangleq \inf\{\alpha > 0 : \mathbb{P}\{\tilde{S}(v, v \circ \theta^{-1}) > \alpha\} < \kappa(\alpha)\}. \quad (14)$$

Hence, under condition (13) a similar estimate as provided in Lemma 6 may be assessed in terms of (14), and if in addition  $\tau$  is assumed to be  $\psi$ -continuous,  $\xi$  is stationary and  $\mathbb{P}$  is quasi-invariant under  $\theta$ , then, the arguments in the proof of Theorem 2 still remain in force and give

$$\limsup_{n \geq 1} \pi(\mathbb{P} \circ \tau(m_n)^{-1}, \mathbb{P} \circ \tau(m_n \circ \theta^{-1})^{-1}) \leq \kappa(\|\theta\|_{\mathbb{P}, \kappa}^{(1)}).$$

As an example, when considering  $E$  and  $T$  to be the real line endowed with the euclidean metric and  $\psi$  the identity, if  $\tau : \mu \in \mathfrak{M}_1^\psi(E) \mapsto \int_{\mathbb{R}} x \mu(dx)$  defines the mean and  $S(x, y) \triangleq (x - y)^2$ , for any  $(x, y) \in \mathbb{R}^2$ , then condition (13) is guaranteed when for instance  $\kappa(z) \triangleq \sqrt{z}$ , for any  $z \geq 0$ .

Observe also that, in the case when  $\psi$  is strictly increasing and  $\tau(\mu)$  is defined as the  $\alpha$ -quantile of the law  $\mu \in \mathfrak{M}_1^\psi(E)$ , for some fixed  $\alpha \in (0, 1)$ , and the related scoring function is given by  $S(x, y) \triangleq (\mathbb{1}_{\{x \geq y\}} - \alpha)(\psi(x) - \psi(y))$ , (see, e.g., Theorem 3.3 in [10]), condition (13) fails for any  $\kappa$ .

## 6. Concluding Remarks

Theorem 1, as well as Corollary 1 and Theorem 2, still remain in force when the sequence of projections  $\xi_1, \xi_2, \dots$  displays some other forms of probabilistic symmetries.

Recall that the random sequence  $\xi = (\xi_1, \xi_2, \dots)$  is said to be *exchangeable* if and only if  $\mathcal{L}(\xi_i : i \in \mathfrak{I}) = \mathcal{L}(\xi_{\pi_{\mathfrak{J}}(i)} : i \in \mathfrak{I})$ , for any finite family  $\mathfrak{I}$  of indices and any permutation  $\pi_{\mathfrak{J}}$  on it. A numerable sequence of exchangeable random variables is always stationary, (cf. [14], Proposition 2.2). In particular, we get that  $\mathbb{P}$ -almost surely  $\mathcal{S} = \sigma(v)$ , (cf. [16], Corollary 1.6). In addition, each of the previous  $\sigma$ -algebras turns out to be  $\mathbb{P}$ -trivial in the independence setup. In this respect, we are allowed to recast the limit random variable in Theorem 1 by writing  $v = \mathbb{P}[\xi_1 \in \cdot | v]$ , where the equality shall be intended in the  $\mathbb{P}$ -almost surely sense. On the other hand, according to de Finetti's Theorem (cf. [16], Theorem 1.1), when dealing with a numerable random sequence  $\xi = (\xi_1, \xi_2, \dots)$  in  $E$ , the notion of exchangeability equals a conditional form of independence, i.e. one has that  $\mathbb{P}$ -almost surely  $\mathbb{P}[\xi \in \cdot | v] = v^{\mathbb{N}}$ .

Exchangeability provides the main pillar of the Bayesian approach to the inferential analysis. More precisely, when dealing with the non parametric setup, the law induced by the random measure  $v$  over the space  $(\mathfrak{M}_1^\psi(E), \mathcal{M}^\psi)$  may be regarded as the *prior distribution* of the statistical model  $\xi_1, \xi_2, \dots | v \sim_{iid} v$ , where the latter form of independence is to be understood in terms of de Finetti's theorem.

According to such a formulation, Theorem 2 may be regarded as a form of stability obtained when the prior distribution of the model is forced to change, **by considering the random measure defined by identity (8)**.

## Acknowledgements

The authors would like to thank an anonymous reviewer for the suggestions which clearly contributed to improve the article.

## References

- [1] Charalambos D Aliprantis and Kim Border. *Infinite dimensional analysis: a hitchhiker's guide*. Springer Science & Business Media, 2006.
- [2] Vladimir I Bogachev. *Measure theory*, volume 2. Springer Science & Business Media, 2007.
- [3] Rama Cont, Romain Deguest, and Giacomo Scandolo. Robustness and sensitivity analysis of risk measurement procedures. *Quantitative Finance*, 10(6):593–606, 2010.
- [4] Antonio Cuevas and Juan Romo. On robustness properties of bootstrap approximations. *Journal of Statistical Planning and Inference*, 37(2):181–191, 1993.
- [5] Richard M Dudley. *Real analysis and probability*, volume 74. Cambridge University Press, 2002.
- [6] Tobias Fissler and Johanna F Ziegel. Higher order elicibility and osband's principle. *The Annals of Statistics*, 44(4):1680–1707, 2016.
- [7] Hans Föllmer and Alexander Schied. *Stochastic finance: an introduction in discrete time*. Walter de Gruyter, 2011.
- [8] Hans Föllmer and Stefan Weber. The axiomatic approach to risk measures for capital determination. *Annual Review of Financial Economics*, 7:301–337, 2015.
- [9] Jiyanta Kumar Ghosh and R.V. Ramamoorthi. *Bayesian nonparametrics*. Springer, 2003.
- [10] Tilmann Gneiting. Making and evaluating point forecasts. *Journal of the American Statistical Association*, 106(494):746–762, 2011.
- [11] Frank R Hampel. *Contributions to the theory of robust estimation*. PhD thesis, University of California, Berkeley, 1969.

- [12] Frank R Hampel. A general qualitative definition of robustness. *The Annals of Mathematical Statistics*, 42(6):1887–1896, 1971.
- [13] Peter J Huber. *Robust statistics*. Springer, 2011.
- [14] Olav Kallenberg. Spreading and predictable sampling in exchangeable sequences and processes. *The Annals of Probability*, 16(2):508–534, 1988.
- [15] Olav Kallenberg. *Foundations of modern probability*. Springer Science & Business Media, 1997.
- [16] Olav Kallenberg. *Probabilistic symmetries and invariance principles*. Springer Science & Business Media, 2006.
- [17] Steven Kou, Xianhua Peng, and Chris C Heyde. External risk measures and basel accords. *Mathematics of Operations Research*, 38(3):393–417, 2013.
- [18] Volker Krätschmer, Alexander Schied, and Henryk Zähle. Qualitative and infinitesimal robustness of tail-dependent statistical functionals. *Journal of Multivariate Analysis*, 103(1):35–47, 2012.
- [19] Volker Krätschmer, Alexander Schied, and Henryk Zähle. Comparative and qualitative robustness for law-invariant risk measures. *Finance and Stochastics*, 18(2):271–295, 2014.
- [20] George W Mackey. Borel structure in groups and their duals. *Transactions of the American Mathematical Society*, 85(1):134–165, 1957.
- [21] Ivan Mizera. Qualitative robustness and weak continuity: the extreme function. *Nonparametrics and Robustness in Modern Statistical Inference and Time Series Analysis: A Festschrift in honor of Professor Jana Jurecková*, 1:169, 2010.
- [22] Kent Osband. *Providing incentives for better cost forecasting*. PhD thesis, University of California, Berkeley, 1985.
- [23] Kalyanapuram Rangachari Parthasarathy. *Probability measures on metric spaces*, volume 352. American Mathematical Society, 1967.
- [24] Ramon van Handel. The universal Glivenko–Cantelli property. *Probability Theory and Related Fields*, 155(3-4):911–934, 2013.

- [25] Henryk Zähle. Qualitative robustness of statistical functionals under strong mixing. *Bernoulli*, 21(3):1412–1434, 2015.
- [26] Henryk Zähle. A definition of qualitative robustness for general point estimators, and examples. *Journal of Multivariate Analysis*, 143:12–31, 2016.
- [27] Johanna F Ziegel. Coherence and elicibility. *Mathematical Finance*, 26(4):901–918, 2016.