Risk Estimation under Probabilistic Symmetries

Enrico Ferri * Carlos Vázquez * José Luis Fernández †

March 13, 2017

Abstract

Consistency and qualitative robustness, two of the main forms of stability usually required when dealing with risk estimators, are presented in an overall perspective by considering different notions of probabilistic symmetries at the level of the stochastic drivers. For such a purpose, a general topological framework is properly introduced by extending the common notion of the weak topology of measures. In particular, the main objective of the work is to state all the results in terms of exchangeability, the celebrated notion of distributional invariance arising when dealing with arbitrary permutations of finite order. In this respect, a new concept of qualitative robustness is defined in terms of the perturbations that affect the system when data are slightly perturbed, and a refined version of the celebrated Hampel's theorem is provided. Such an overture turns out to be strongly appealing from the conceptual point of view, since data are the main drivers of the empirical analysis.

1. INTRODUCTION

According to the standard practice, the assessment of the risk associated to some market exposure is generally fulfilled throughout two different steps. Firstly, a predictive law describing the expected behaviour of the exposure into the future is inferred by calibrating an *a priori* defined statistical model, on the basis of the available historical information. Thereafter, such a law is used in order to compute the downside risk, by performing a suitable risk functional. As a result, the estimation obtained in such a way turns out to be strongly sensitive to the amount of data processed at the stage of the statistical inference, i.e. different amount of data typically generate different estimates. That is the reason why we cannot shy away from the question regarding the behaviour displayed by the estimates when enlarging the informative dataset, by asking if some form of asymptotic stability may be expected when imposing certain regularity conditions on the stochastic drivers. Convergence results of this type have been classically worked out under the notion of *consistency*. On the other hand, Cont et. al [5] pointed out that risk can be properly assessed only when dealing with robust estimators, taking into account the classical notion of *qualitative robustness* introduced by Hampel [20]. More precisely, robustness might be understood as a form of stability displayed by the estimators when the predictive distribution is forced to be slightly perturbed. On the mathematical level, qualitative robustness can be properly formalized as the equicontinuity of the family of laws induced by the risk estimators, which are to be understood as functions of the predicting distribution, when the amount of the processed data is larger than a fixed threshold depending on the impact of the perturbation itself.

Although the estimation process previously described is generally granted to be effective, it ostensibly makes sense only when dealing with law-invariant risk measures, saying informally that the information encoded by the stochastic factors may be modelled just by considering the related distributions. In this respect, the natural workspace is provided by a suitable family of probability measures defined on the space

*Department of Mathematics, Faculty of Informatics, University of Coruña, enrico.ferri@ucd.es/carlosv@udc.es

†Department of Mathematics, Faculty of Sciences, Autonomous University of Madrid, joseluis.fernandez@uam.es

of the empirical outcomes. It is thus nothing short of clear that the topological structure of such a domain plays a crucial role when investigating the forms of stability previously discussed. Notwithstanding the notion of qualitative robustness is classically investigated in terms of the topology of the weak convergence of measures, cf. [5, 7, 19, 20, 22, 36], it has been recently pointed out how such a formulation might generate a number of debated complications, cf. [30, 31]. First of all, since the weak topology looks at the body of the distributions, two laws could be rather close and, on the contrary, still display completely different behaviours on tails. Worse still, all the risk estimators that are, by their nature, insensitive to the catastrophic events unavoidably appear robust according to this setup.
In this respect, we prefer to initially maintain the formulation as general as possible, since the problems previously discussed can be easily overcome by introducing some topological refinement, which is able to capture the deviation of distributions on their tails, cf. [30, 31].

One of our intents is to investigate the performance of the risk estimates when the available dataset is enlarged, by analysing under which stochastic hypotheses they display the asymptotic forms of stability previously discussed. Indeed, while consistency is generally assessed according to classical ergodic arguments [31], the more restrictive *i.i.d.* setting is usually recovered in order to investigate the notion of qualitative robustness, cf. [19, 20, 30, 31]. In particular Krätschmer et. al [30, 31] restate the celebrated Hampel's theorem in a more general framework, while Zähle [47, 46] achieved a similar result by allowing the stochastic process modelling the empirical outcomes to display an internal dependence structure.
In the present work, our goal is to provide a common framework in order to properly get an overall perspective and thus develop the theory concerning the consistency of the risk estimators together with the qualitative robustness in an elegant and self-consistent setup. For these purposes, we start by analysing the notion of consistency by taking into account different forms of distributional dependence. Indeed, although the stationary structure provides the main pillar of the ergodic analysis, there are plenty of different probabilistic symmetries that may be properly exploited for such a purpose. The symmetry structures which act in the drama of probability are generally introduced in terms of families of random elements, whose joint distribution does not change when considering specific classes of transformations. In this respect, stationarity is defined as the form of distributional invariance arising when dealing with shift operators. Our objective is to develop the main results in terms of exchangeability, the celebrated notion of probabilistic symmetry corresponding to the distributional invariance exhibited under finite permutations, and that can be properly defined in terms of optional shifts. In most of the cases, exchangeability ostensibly provides a more convenient assumption due to many aspects, first and foremost the possibility to exploit the notion of conditional independence, as classically assessed by de Finetti's theorem (cf. [25], Theorem 9.16). Moreover, it ostensibly appears a suitable assumption also in view of many applications, by encoding the feeling that the empirical analysis is not sensitive to the order by which data are processed. Furthermore, our refined overture allows to assess the analysis in terms of random probability measures. This alternative turns out to be strongly appealing from the conceptual point of view, as the law of the generic random variable modelling the empirical observations may be properly described as a function depending on numerable sequences of outcomes. Stated differently, any infinitely large dataset determines the law that may be understood as the best available statistical description of the data. Nevertheless, the main strength of the our formulation lies in the possibility to restate the notion of qualitative robustness in terms of the perturbations that affect the system when data are slightly stressed. Assessing this notion of stability in terms of corrupted dataset outwardly provides a better interpretation also at the practical level, since according to the standard practice, by their nature data are the main drivers of the empirical analysis.

Our article is organized as follows. In Section 2 we define the setup according to which we develop our approach. Section 3 is entirely dedicated to the study of consistent risk estimates, by exploiting the notions of probabilistic symmetries of major concern, and thus assessing the main results in terms of random probability measures. In Section 4 we introduce our refined notion of qualitative robustness in terms of measurable action of group, providing an alternative version of the celebrated Hampel's theorem. In Section 5 we discuss how our results easily reduce when dealing the same topological setting considered by Krätschmer et. al [30, 31].

2. BASIC NOTIONS AND SETUP

Let $(E, \mathfrak{E})$ be a Polish space endowed with the Borel σ-algebra $\mathscr{E}$ generated by the topology $\mathfrak{E}$ as usual, and denote by $\mathfrak{M}_1(E)$ the entire family of laws defined on $(E, \mathscr{E})$. Through the paper we largely focus on the analytical properties of the generic functional defined on $\mathfrak{M}_1(E)$. For this purpose, characterizing the topological structure of such a space turns out to be crucial in order to properly get an overall perspective. It's worth noting that there exist plenty of topologies that can be classically defined on the space $\mathfrak{M}_1(E)$. In order to maintain the setup as general as possible, let $\mathfrak{F}(E)$ be a family of measurable functions on $(E, \mathscr{E})$. We shall implicitly assume that every function of this type takes real values, and thus the measurablility is to be understood in terms of the common Borel structure on the real line. Moreover, we denote by $\mathfrak{M}_1^{\mathfrak{F}}(E)$ the family of measures $\mu \in \mathfrak{M}_1(E)$ such that $\int_E |f| d\mu < +\infty$, for any $f \in \mathfrak{F}(E)$.
The main idea is to properly define a topological structure on $\mathfrak{M}_1^{\mathfrak{F}}(E)$, by associating it with the family $\mathfrak{F}(E)$ via the canonical bilinear form

$$(\mu, f) \mapsto \mu f \triangleq \int_E f d\mu, \qquad \text{for any } \mu \in \mathfrak{M}_1^{\mathfrak{F}}(E) \text{ and any } f \in \mathfrak{F}(E). \tag{1}$$

At this aim, assume that $\mathfrak{F}(E)$ separates the points of $\mathfrak{M}_1^{\mathfrak{F}}(E)$, i.e. given any couple $\mu, \nu \in \mathfrak{M}_1^{\mathfrak{F}}(E)$ one has $\mu = \nu$ if and only if $\mu f = \nu f$ for any $f \in \mathfrak{F}(E)$. Such an assumption is not too restrictive, since the following result generally holds true.

Lemma 1. *Let $\mathfrak{A}(E)$ denote the family of all bounded and uniformly continuous real functions on E, then $\mathfrak{F}(E)$ separates the points of $\mathfrak{M}_1^{\mathfrak{F}}(E)$ whenever $\mathfrak{F}(E) \supseteq \mathfrak{D}(E)$ for some uniformly dense subset $\mathfrak{D}(E)$ of $\mathfrak{A}(E)$.*

Proof. Given $\mu, \nu \in \mathfrak{M}_1^{\mathfrak{F}}(E)$ such that $\mu = \nu$, one trivially obtains that $\mu f = \nu f$, for any $f \in \mathfrak{F}(E)$. Conversely, assume that $\mu f = \nu f$, for any $f \in \mathfrak{F}(E)$, and thus for any $f \in \mathfrak{D}(E)$. Since $\mathfrak{D}(E)$ is uniformly dense in $\mathfrak{A}(E)$, the equality $\mu = \nu$ easily follows from Theorem 15.1 in [2]. □

In this respect, the space $\mathfrak{M}_1^{\mathfrak{F}}(E)$ may be considered as a subspace of the product space $\mathbb{R}^{\mathfrak{F}(E)}$, since according to the previous assumption we can associate any $\mu \in \mathfrak{M}_1^{\mathfrak{F}}(E)$ to the linear map $f \mapsto \mu f$, varying $f \in \mathfrak{F}(E)$. As a result, the space $\mathbb{R}^{\mathfrak{F}(E)}$ induces a topological structure on $\mathfrak{M}_1^{\mathfrak{F}}(E)$ in a natural way.

Definition 1 ($\mathfrak{F}(E)$-weak topology). We call $\mathfrak{F}(E)$-*weak topology* the topology on $\mathfrak{M}_1^{\mathfrak{F}}(E)$ inherited from the product topology defined on $\mathbb{R}^{\mathfrak{F}(E)}$, and generally denoted by $\sigma\big(\mathfrak{M}_1^{\mathfrak{F}}(E), \mathfrak{F}(E)\big)$.

According to Definition 1, the $\mathfrak{F}(E)$-weak topology is thus the projective topology on $\mathfrak{M}_1^{\mathfrak{F}}(E)$ defined by the linear forms on $\mathfrak{M}_1^{\mathfrak{F}}(E)$ belonging to $\mathfrak{F}(E)$, i.e. the coarsest topology on $\mathfrak{M}_1^{\mathfrak{F}}(E)$ that renders continuous the maps $\mu \mapsto \mu f$, varying $f \in \mathfrak{F}(E)$. It is worth noting that, if $\langle \mathfrak{ca}(E), \mathfrak{F}(E) \rangle$ provides a dual pair in the duality (1), where $\mathfrak{ca}(E)$ denotes the family of all signed measures of bounded variation on $(E, \mathscr{E})$, and $\mathfrak{M}_1^{\mathfrak{F}}(E)$ is a $\sigma\big(\mathfrak{ca}(E), \mathfrak{F}(E)\big)$-closed and convex subspace of $\mathfrak{ca}(E)$, the $\mathfrak{F}(E)$-weak topology on $\mathfrak{M}_1^{\mathfrak{F}}(E)$ is nothing but the relativization of $\sigma\big(\mathfrak{ca}(E), \mathfrak{F}(E)\big)$ to $\mathfrak{M}_1^{\mathfrak{F}}(E)$. Hence, according to the previous arguments, we also say that $\sigma\big(\mathfrak{M}_1^{\mathfrak{F}}(E), \mathfrak{F}(E)\big)$ is *weakly generated* by $\mathfrak{F}(E)$, or *generated by the duality* with $\mathfrak{F}(E)$. Moreover, a sequence in $\mathfrak{M}_1^{\mathfrak{F}}(E)$ is $\mathfrak{F}(E)$-*weak convergent*, if (and only if) it converges with respect to the $\mathfrak{F}(E)$-weak topology on $\mathfrak{M}_1^{\mathfrak{F}}(E)$.
It's nothing short of clear that the regularitity properties displayed by the elements of $\mathfrak{F}(E)$ strongly impact on the nature of the resulting topology on $\mathfrak{M}_1^{\mathfrak{F}}(E)$. Besides, not every choice for the family $\mathfrak{F}(E)$

is suitable in our context. Indeed, we would like to obtain a topological space that may be metrized by some complete and separable distance, and such a result is not guaranteed in general.

Definition 2 (Dual Normal Consistence). We shall say that the family $\mathfrak{F}(E)$ is *dual normal consistent* if (and only if) the space $\mathfrak{M}_1^{\mathfrak{F}}(E)$ jointly with the $\mathfrak{F}(E)$-weak topology turns out to be metrized by some complete and separable distance $d_{\mathfrak{F}(E)}$ on it.

By means of Definition 2, note that when letting $\mathfrak{F}(E)$ be the family $\mathfrak{C}_b(E)$ of continuous and bounded functions on $(E, \mathscr{E})$, we easily obtain that $\mathfrak{M}_1^{\mathfrak{F}}(E)$ equals the entire space $\mathfrak{M}_1(E)$. In particular, the topology on it generated by the duality with $\mathfrak{C}_b(E)$ is nothing but the standard weak topology of measures. Thus, the family $\mathfrak{C}_b(E)$ clearly turns out to be dual normal consistent. Indeed, since $(E, \mathfrak{E})$ is assumed to be Polish, $\mathfrak{M}_1(E)$ endowed with weak topology of measures is classically metrized as a complete and separable metric space, by considering for instance the Prohorov distance

$$\pi(\mu, \nu) \triangleq \inf\{\varepsilon > 0 : \mu(B) \leq \nu(B^\varepsilon) + \varepsilon, \text{ for any } B \in \mathscr{E}\}, \qquad \text{for any } \mu, \nu \in \mathfrak{M}_1(E). \tag{2}$$

where $B^\varepsilon \triangleq \{x \in E : \inf_{y \in B} d_E(y, x) < \varepsilon\}$ denotes the ε-hull of $B \in \mathscr{E}$. Moreover, the same topological structure may be equally achieved when enforcing the analytical regularity of the functions in $\mathfrak{C}_b(E)$, by considering the family $\mathfrak{BL}(E)$ of bounded and Lipschitz continuous functions on $(E, \mathscr{E})$, (cf. [12], Theorem 11.3.3). More generally, letting $\mathfrak{A}(E)$ denote the family of all bounded and uniformly continuous real function on E, the standard weak topology of measures on $\mathfrak{M}_1(E)$ is generated by the duality with any uniformly dense subset of $\mathfrak{A}(E)$, (cf. [2], Theorem 15.2). Nevertheless, continuity can not be ostensibly omitted, since the topology generated on $\mathfrak{M}_1(E)$ by the duality with the entire family of measurable and bounded functions on $(E, \mathscr{E})$ equals the topology associated to the *setwise convergence*, that is not metrizable, (cf. [18], Proposition 2.2.1). Thus, the natural question is whether the boundedness of the functions in $\mathfrak{F}(E)$ is necessary in order to get a dual normal consistent family. As illustrated below, the answer is no.

Definition 3 (ψ-weak topology). Let ψ be a real valued and non negative $\mathscr{E}$-measurable function on E, satisfying $\psi \geq 1$ outside some compact set. We call *ψ-weak topology* the topology on $\mathfrak{M}_1^\psi(E) \triangleq \{\mu \in \mathfrak{M}_1(E) : \mu\psi < \infty\}$ generated by the duality with $\mathfrak{C}_\psi(E) \triangleq \{f \in \mathfrak{C}(E) : \|f/(1+\psi)\|_\infty < \infty\}$.

Note that the map $\Psi : \mathfrak{M}_1(E) \to \mathfrak{M}_1^\psi(E)$ defined by means of the equality $d\Psi(\mu) = (1+\psi)^{-1} d\mu$ defines an homeomorphism between the two spaces. Thus, due to Theorem 6.2 in [37] combined with Corollary 11.5.5. in [12], since $(E, \mathscr{E})$ is assumed to be Polish, the space $\mathfrak{M}_1^\psi(E)$ jointly with the ψ-weak topology can be metrized as a separable and complete space. Furthermore, the previous arguments can be naturally generalized in the following lemma.

Lemma 2. *Endowed with the ψ-weak topology, the space $\mathfrak{M}_1^\psi(E)$ is metrized as a complete and separable space, when considering the metric $d_\psi(\mu, \nu) \triangleq \pi(\mu, \nu) + |(\mu - \nu)\psi|$, for any $\mu, \nu \in \mathfrak{M}_1^\psi(E)$, where $(\mu - \nu)\psi \triangleq \mu\psi - \nu\psi$.*

It's worth noting that the Prohorov measure π in the statement of Lemma 2 can be properly replaced by any other distance metrizing the weak topology of measures, like for instance the *Fortet-Mourier* distance induced by $\|\mu\|^\star_{\mathfrak{BL}} \triangleq \sup\{|\mu f| : \|f\|_{\mathfrak{BL}} \leq 1\}$, for $\mu \in \mathfrak{M}_1^\psi(E)$, the norm classically associated to the topological dual of the space $\mathfrak{BL}(E)$.

Proof of Lemma 2. The proof trivially follows from Definition 3, noting that a sequence $(\mu_n)_n$ in $\mathfrak{M}_1^\psi(E)$ converges to some law $\mu \in \mathfrak{M}_1^\psi(E)$ in the ψ-weak topology if and only if it weakly converges to μ and $\mu_n\psi \to \mu\psi$, as $n \to +\infty$. In other worlds, the ψ-weak topology on $\mathfrak{M}_1^\psi(E)$ is generated by the duality with $\mathfrak{C}_b(E)$, or equally with $\mathfrak{BL}(E)$, jointly with the function ψ. □

Given some dual normal consistent family of functions $\mathfrak{F}(E)$, the space $\mathfrak{M}_1^{\mathfrak{F}}(E)$ endowed with the $\mathfrak{F}(E)$-weak topology provides the natural domain according to which we state all the results, since in the law invariance framework the downside risk may be properly assessed by considering a suitable functional defined on it. More precisely, any application of this type naturally defines a sequence of risk estimators, by considering a certain family of random measures on $(E,\mathscr{E})$, i.e. a family of probability kernels from the basic underlying probability space $(\Omega,\mathscr{F},\mathbb{P})$ to $(E,\mathscr{E})$. In this respect, it is worth noting that any random probability measure on $(E,\mathscr{E})$ can be also properly described as a random element of the space $\mathfrak{M}_1(E)$, endowed with the σ-algebra generated as usual by the projection maps $\pi_B:\mu\mapsto\pi_B(\mu)\triangleq\mu(B)$, letting B vary in $\mathscr{E}$. As a result, since any sequence of risk estimators turns out to be a family of random variables taking values in a suitable space, any form of regularity is to be assessed in probabilistic terms.

3. STRONG CONSISTENCY OF RISK ESTIMATORS

It is a well-known fact that the downside risk of an exposure is assessed on the basis of the available informative set of historical values. Thus, we implicitly assume that the collected observations are encoded by the elements of the Borel space $(E,\mathscr{E})$. Therefore, given some underlying non atomic probability space $(\Omega,\mathscr{F},\mathbb{P})$, any empirical outcome may be modeled by considering some random variable defined on it and taking values in $(E,\mathscr{E})$. For sake of simplicity, we sometimes refer to any variable of this type as a random element in $(E,\mathscr{E})$. As usual, we shall denote by $\mathscr{L}(\xi)$ the law induced by the generic random element ξ in $(E,\mathscr{E})$.
Without loss of generality, we're allowed to set $\Omega\triangleq E^{\mathbb{N}}$, the family of numerable sequences in E, and thus to let $\mathscr{F}\triangleq\mathscr{E}^{\mathbb{N}}$ be the product σ-algebra on it, classically generated by the canonical projections $\xi_n:\omega\mapsto\xi_n(\omega)\triangleq\omega_n$, for any $\omega\in\Omega$, varying $n\geq 1$. Since both at a conceptual and technical level, the standard procedure is to look at data as the realizations of the random sequence $\xi\triangleq(\xi_n)_n$, we shall refer to ξ as the process governing the data or, more simply, the *data process*. Besides, such a sequence naturally induces the family of random measures $(m_n)_n$ on $(E,\mathscr{E})$ defined by setting

$$m_n(\omega,B)\triangleq\frac{1}{n}\sum_{i=1}^{n}\delta_{\xi_i(\omega)}(B),\qquad\text{for any }\omega\in\Omega\text{ and }B\in\mathscr{E},\tag{3}$$

varying $n\geq 1$. In particular, let $\mathfrak{M}_{1,\mathrm{emp}}(E)\subset\mathfrak{M}_1(E)$ be the space of the *empirical measures* on $(E,\mathscr{E})$ of the form $n^{-1}\sum_{i\leq n}\delta_{x_i}$ for any $(x_1,...,x_n)\in E^n$, varying $n\geq 1$. Since for any $n\geq 1$ the random measure m_n is nothing but a random element in the space $\mathfrak{M}_{1,\mathrm{emp}}(E)$, endowed with the σ-algebra generated by the canonical projections $\pi_B:\mu\mapsto\mu(B)$, for $\mu\in\mathfrak{M}_{1,\mathrm{emp}}(E)$, varying $B\in\mathscr{E}$, we shall refer to (3) as the family of *empirical kernels* directed by ξ.
Let M be some subspace of $\mathfrak{M}_1(E)$ such that $\mathfrak{M}_{1,\mathrm{emp}}(E)\subset M$. Given an *a priori* defined Borel space $(T,\mathscr{T})$, the risk associated to an exposure is usually assessed by performing a certain measurable functional $\tau:M\to T$, usually called *statistic*. In particular, we shall suppose that T is endowed with some metric d_T consistent with its topological structure, i.e. the topology $\mathfrak{T}$ on T generating the Borel σ-algebra $\mathscr{T}$ is metrized by d_T. We also assume that the pair $(T,\mathfrak{T})$ provides a Polish space and we refer to such a space as the *action domain* of the statistic τ. In general, letting $(\mathscr{F}_n)_n$ be the filtration in $\mathscr{F}$ generated by the data process ξ, any sequence $(\mu_n)_n$ of random measures on $(E,\mathscr{E})$ adapted to $(\mathscr{F}_n)_n$, i.e. such that μ_n is $\mathscr{F}_n$-measurable for any $n\geq 1$, induces the family of estimators $(\tau(\mu_n))_n$. As we mainly focus on the family of empirical kernels (3) in order to perform the statistic τ, the latter arguments easily lead to the following definition.

Definition 4 (Risk Estimators). Given the family $(m_n)_n$ of empirical kernels (3) and a statistic $\tau:M\to T$, we shall refer to the random sequence $(\tau_n)_n$ in $(T,\mathscr{T})$ obtained by setting $\tau_n\triangleq\tau(m_n)$, for any $n\geq 1$, as the family of *historical risk estimators*, or simply the *estimators*, induced by τ.

The obvious problem that arises at this step is to characterize the asymptotic behavior of the family of estimators $(\tau_n)_n$ when n is large. In particular, it appears crucial to understand under which stochastic

assumptions the sequence $(\tau_n)_n$ admits a consistent limit with respect to a certain notion of convergence. For this purpose, the main idea is to deal with a data process ξ whose internal structure display a certain probabilistic symmetry, saying that the random elements $\xi_1, \xi_2, ...$ turn out to be jointly distributionally invariant under the action of some measurable endomorphism on $(E^N, \mathscr{E}^{\mathbb{N}})$.
In particular, we mainly focus on stationarity and exchangeablility, two of the stochastic symmetries of major concern, that respectively arise when dealing with operations of shift and arbitrary permutations of finite order. More precisely, we recall that a random sequence $\xi \triangleq (\xi_1, \xi_2, ...)$ in $(E, \mathscr{E})$ is said stationary if (and only if) $\mathscr{L}(\xi) = \mathscr{L}(\Sigma(\xi))$, where $\Sigma(\omega_1, \omega_2, ...) \triangleq (\omega_2, \omega_3, ...)$ denotes the *shift* operator on $E^{\mathbb{N}}$, while it's called exchangeable if (and only if) $\mathscr{L}(\xi_n : n \in \mathfrak{I}) = \mathscr{L}(\xi_{\pi_{\mathfrak{I}}(n)} : n \in \mathfrak{I})$ for any finite family of indices $\mathfrak{I}$ and any permutation $\pi_{\mathfrak{I}}$ on it. In this respect, we also denote by $\mathscr{E}^{\mathbb{N}}_{\text{inv}(\Sigma)}$ the shift invariant σ-algebra in $(E^{\mathbb{N}}, \mathscr{E}^{\mathbb{N}})$, i.e. the collection of all the Borel sets $I \in \mathscr{E}^{\mathbb{N}}$ such that $\Sigma^{-1}(I) = I$. It's nothing short of clear that the notion of exchanegability provides a stronger form of dependence than stationarity. On the other hand, as stated by Proposition 9.18 in [25], it can be properly described as a form of stronger stationarity involving optional shift. Nevertheless, the two forms of probabilistic symmetry can be both suitably exploited in order to investigate the asymptotic nature of the estimators $(\tau_n)_n$. Besides, the analytical properties of the statistic $\tau : M \to T$ also play a crucial role, in this setting. In particular, letting $\mathfrak{F}(E)$ be some *a priori* defined dual normal consistent family on measurable functions on $(E, \mathscr{E})$, we shall consider the related space $\mathfrak{M}^{\mathfrak{F}}_1(E)$ as the domain of the statistic τ. Thus, we shall say that the statistic τ is $(d_{\mathfrak{F}(E)}, d_T)$-continuous if it is continuous with respect to the topology generated on $\mathfrak{M}^{\mathfrak{F}}_1(E)$ by the duality with $\mathfrak{F}(E)$ and the topology $\mathfrak{T}$ on T generating the Borel σ-algebra $\mathscr{T}$ respectively. Moreover, we denote by $\xi^{-1}\mathscr{E}^{\mathbb{N}}_{\text{inv}(\Sigma)}$ the family of sets $\xi^{-1}(B) \in \mathscr{F}$ obtained by varying $B \in \mathscr{E}^{\mathbb{N}}_{\text{inv}(\Sigma)}$, which trivially admits the structure of σ-algebra since ξ^{-1} preserves all the set operations. The following result explains how the asymptotic behaviour of the estimators $(\tau_n)_n$ can be assessed in terms of stationarity.

Theorem 1 (*Consistency under Stationarity*)**.** *Suppose that the statistic τ is $(d_{\mathfrak{F}(E)}, d_T)$-continuous and that the sequence ξ is stationary and such that $\mathscr{L}(\xi_1) \in \mathfrak{M}^{\mathfrak{F}}_1(E)$. Then, considering the random probability measure $\upsilon \triangleq \mathbb{P}[\xi_1 \in \cdot\ |\xi^{-1}\mathscr{E}^{\mathbb{N}}_{\text{inv}(\Sigma)}]$, we get that $\tau_n \to \tau(\upsilon)$ almost certainty with respect to $\mathbb{P}$, as $n \to +\infty$.*

Proof. First of all, note that the stationary structure of the random sequence ξ implies that $\mu \triangleq \mathscr{L}(\xi_1) = \mathscr{L}(\xi_n)$ for any $n \geq 1$. Next, let us fix $f \in \mathfrak{F}(E)$. Since $\mu \in \mathfrak{M}^{\mathfrak{F}}_1(E)$, the von Neumann's version of Birkhoff's theorem (cf. [25], Theorem 9.6) combined with the celebrated disintegration theorem (cf. [25], Theorem 5.4), assures that the asymptotic behavior $m_n(\omega, \cdot\,)f \to \upsilon(\omega, \cdot\,)f$ holds true for $\mathbb{P}$-almost any $\omega \in \Omega$ and in L^1 norm as $n \to +\infty$, where we set $\upsilon \triangleq \mathbb{P}[\xi_1 \in \cdot\ |\xi^{-1}\mathscr{E}^{\mathbb{N}}_{\text{inv}(\Sigma)}]$. We also highlight that the conditional distribution υ always exists, since E is Polish and endowed with the Borel σ-algebra $\mathscr{E}$. Given a dual normal consistent family $\mathfrak{F}(E)$ of measurable functions on $(E, \mathscr{E})$, letting f vary in $\mathfrak{F}(E)$, the previous arguments assure that $m_n(\omega, \cdot\,) \to \upsilon(\omega, \cdot\,)$ in the $\mathfrak{F}(E)$-weak topology on $\mathfrak{M}^{\mathfrak{F}}_1(E)$, for $\mathbb{P}$-almost any $\omega \in \Omega$, as $n \to +\infty$. Letting now $d_{\mathfrak{F}(E)}$ be some metric consistent with the topology weakly generated on $\mathfrak{M}^{\mathfrak{F}}_1(E)$ by the duality with $\mathfrak{F}(E)$, since τ is assumed to be $(d_{\mathfrak{F}(E)}, d_T)$-continuous on $\mathfrak{M}^{\mathfrak{F}}_1(E)$, we easily conclude the proof. □

It is worth noting that a regular version of υ always exists since E is Polish and endowed with the Borel σ-algebra $\mathscr{E}$. Moreover, since the stationary structure of the sequence ξ implies that $\mathscr{L}(\xi_n) = \mathscr{L}(\xi_1)$ for any $n \geq 2$, if in addition ξ is ergodic, i.e. its distribution on $(E^{\mathbb{N}}, \mathscr{E}^{\mathbb{N}})$ is ergodic with respect to the shift operator Σ on $E^{\mathbb{N}}$, the asymptotic result stated by Theorem 1 boils down to $\tau_n \to \tau(\mu)$, $\mathbb{P}$-almost surely as $n \to +\infty$, where we set $\mu \triangleq \mathscr{L}(\xi_1)$. Indeed the probability measure $\mathbb{P}$ turns out to be trivial on $\mathscr{E}^{\mathbb{N}}_{\text{inv}(\Sigma)}$ in such a case.
Theorem 1 still remains valid when enforcing the stochastic structure of the data process, by considering a stronger form of internal dependence. Indeed, it can be naturally reformulated in terms of exchangeability, as stated below.

Theorem 2 (*Consistency under Exchangeability*)**.** *Suppose that τ is $(d_{\mathfrak{F}(E)}, d_T)$-continuous and that the sequence ξ is exchangeable and such that $\mathscr{L}(\xi_1) \in \mathfrak{M}_1^{\mathfrak{F}}(E)$. Then, letting υ be a regular version of the random measure $\mathbb{P}[\xi_1 \in \cdot \, | \xi^{-1}\mathscr{E}^{\mathbb{N}}_{\mathrm{inv}(\Sigma)}]$, we get that $\tau_n \to \tau(\upsilon)$ almost certainty with respect to $\mathbb{P}$, as $n \to +\infty$.*

The following proof is based on the circle of ideas lying behind de Finetti's theorem (cf. [25], Theorem 9.16). Such a result may be formulated in terms of the ergodic theory, and thus exploiting the discussion present in the proof of Theorem 1, cf. [24, 26]. Nevertheless, we prefer to provide an alternative proof based on simple martingale arguments, by generalizing the discussion in [27], since we found it more illuminating from the probabilistic point of view.

Proof of Theorem 2. First of all, note that the internal exchangeable structure of the sequence ξ implies that $\mu \triangleq \mathscr{L}(\xi_1) = \mathscr{L}(\xi_n)$ for any $n \geq 2$. Thus, set $f \in L^1(E, \mathscr{E}, \mu)$ and for any $n \geq 1$ let $\mathscr{S}_n$ be the σ-algebra generated by the random elements ξ_i when $i \geq n+1$ together with the variables of the form $\phi(\xi_1, ..., \xi_n)$, for some symmetric Borel function $\phi : E^n \to E$. Whenever we consider a $\mathscr{S}_n$-measurable random element η in $(E, \mathscr{E})$, we easily get that $\mathbb{E}[f(\xi_i)\eta] = \mathbb{E}[f(\xi_1)\eta]$ for $1 \leq i \leq n$, and in particular by linearity we have

$$\mathbb{E}\left[\frac{1}{n}\sum_{i=1}^{n} f(\xi_i)\eta\right] = \mathbb{E}[f(\xi_1)\eta].$$

As a result, since $n^{-1}\sum_{i=1}^{n} f(\xi_i)$ is clearly $\mathscr{S}_n$-measurable, it turns out to be a version of $\mathbb{E}[f(\xi_1)|\mathscr{S}_n]$. Noting that $\mathscr{S}_n \supseteq \mathscr{S}_{n+1}$, the celebrated Lévy's downward theorem (cf. [44], Theorem 14.4) combined with the disintegration theorem (cf. [25], Theorem 5.4) implies that $m_n(\omega, \cdot\,)f \to \upsilon^1(\omega, \cdot\,)f$, as $n \to +\infty$, where we denoted by υ^1 the conditional distribution $\mathbb{P}[\xi_1 \in \cdot \, |\mathscr{S}_\infty]$ obtained by setting $\mathscr{S}_\infty \triangleq \bigcap_n \mathscr{S}_n$ as usual. Note also that the conditional distribution υ^1 always exists, as E is Polish and endowed with the Borel σ-algebra $\mathscr{E}$. In particular, since the numerable sequence ξ admits an exchangeable structure, Corollary 1.6 in [26] assures that $\mathscr{S}_\infty = \xi^{-1}\mathscr{E}^{\mathbb{N}}_{\mathrm{inv}(\Sigma)} = \sigma(\upsilon)$, and thus $\upsilon = \upsilon^1$, almost surely with respect to $\mathbb{P}$, where we denoted by υ the random measure obtained in the proof of Theorem 1. Given now a dual normal consistent family $\mathfrak{F}(E)$ of measurable functions on $(E, \mathscr{E})$, since $\mu \in \mathfrak{M}_1^{\mathfrak{F}}(E)$, letting f vary in $\mathfrak{F}(E)$, the previous arguments imply that $m_n(\omega, \cdot\,) \to \upsilon(\omega, \cdot\,)$ $\mathfrak{F}(E)$-weakly, for $\mathbb{P}$-almost any $\omega \in \Omega$, as $n \to +\infty$. The proof is concluded by considering the same arguments as in the proof of Theorem 1. □

It is nothing short of clear from the proof of Theorem 2, that when looking at the random measure υ, the representation $\upsilon = \mathbb{P}[\xi \in \cdot \, |\upsilon]$ holds true, by means of the equality $\upsilon = \mathbb{P}[\xi \in \cdot \, |\sigma(\upsilon)]$, where $\sigma(\upsilon) \subset \mathscr{F}$ denotes the σ-algebra generated by the random measure υ. Moreover, due to Proposition 1.4 in [26], it can be $\mathbb{P}$-almost everywhere uniquely described in terms of the limits of the empirical kernels (3), by means of the $\mathbb{P}$-almost sure setwise convergence $m_n(\,\cdot\,, B) \to \upsilon(\,\cdot\,, B)$, for any $B \in \mathscr{E}$, as $n \to +\infty$. In this respect, we shall refer to υ as the random measure directing the sequence ξ.
It is worth recalling that the concept of exchangeability turns out to be very close to the notion of *spreadablility*, the distributional form of invariance that arises when looking at any sub-sequence of ξ. More precisely, the random sequence ξ is said *spreadable* (or *contractable*) if (and only if) $\mathscr{L}(\xi_1, \xi_2, ...) = \mathscr{L}(\xi_{k_1}, \xi_{k_2}, ...)$, for any strictly increasing sequence $(k_n)_n$ of positive integers. Hence, since due to Theorem 1.1 in [26] the notion of exchangeability classically equals the notion of spreadability, when dealing with numerable random sequences, Theorem 2 may be naturally restated as follows.

Corollary 1 (*Consistency under Spreadability*)**.** *Suppose that τ is $(d_{\mathfrak{F}(E)}, d_T)$-continuous and that the sequence ξ directed by the random measure υ is spreadable and such that $\mathscr{L}(\xi_1) \in \mathfrak{M}_1^{\mathfrak{F}}(E)$. Then, $\tau_n \to \tau(\upsilon)$ almost certainty with respect to $\mathbb{P}$, as $n \to +\infty$.*

The latter results outwardly provide a characterization of the random sequence of estimators $(\tau_n)_n$, assessing its asymptotic behaviour in terms of ones of the most celebrated notions of probabilistic symmetries. In particular, we shall refer to Theorem 1, 2 and thus to Corollary 1, by means of the following classical terminology.

Definition 5 (Strong Consistency). If $\tau_n \to \tau(\upsilon)$ almost surely with respect to $\mathbb{P}$ as $n \to +\infty$, we shall say that the random sequence $(\tau_n)_n$ provides a family of *strongly consistent* estimators for $\tau(\upsilon)$, or simply that the statistic τ is *strongly consistent* with respect to the data process ξ or equivalently to the directing random measure υ.

Note that in our setup, whenever we fix an element ω in Ω, we obtain a law $\upsilon(\omega, \cdot\,)$ on the space of the data $(E, \mathscr{E})$. Recalling that the space Ω has been defined as the family of numerable sequences in E, we get that any (infinite) dataset determines a probability measure on the Borel space $(E, \mathscr{E})$ that may be understood as the best probabilistic description of the data. In other words, when enlarging the dataset, the model implicitly elects the limit distribution on the basis of the experience.

4. QUALITATIVE ROBUSTNESS UNDER GROUP ACTION

Roughly speaking, the celebrated Hampel's notion of robustness refers to the property of a family of estimators to be stable with respect to little changes affecting the predicting distribution, cf. [19, 20]. More precisely, some authors refer to such a form of stability saying that the laws induced by the family of the risk estimators do not notably change when the distribution from which data are independently sampled is forced to be slightly modified, [30, 31]. Zähle provides a refined version of qualitative robustness, by allowing the data process to admit an internal dependence structure, [47, 45, 46]. Main objective of the current section is to discuss a refined notion of qualitative robustness which naturally arises when dealing with the setup previously discussed. Indeed, we witnessed how exchangeability provide a suitable form of probabilistic structure when investigating the strong consistency of the generic family of risk estimators. On the other hand, we also highlighted that such an approach turns out to be strongly appealing also due to the possibility to assess the analysis in terms of random probability measures. In this respect, the predicting distribution can be seen as a function of the available dataset. In particular, any numerable sequence of outcomes determines a probability measure on the Borel space $(E, \mathscr{E})$. In a setting where empirical observations are the real drivers of the study, it appears natural to assess the notion of robustness as a form of stability that risk estimators display when data are corrupted. In this respect, we shall provide a revised notion of qualitative robustness by exploiting the celebrated de Finetti's theorem, which classically assess the notion of exchangeability in terms of conditional independence.

Let us maintain the same notation from the previous section. In particular, the underlying triple $(\Omega, \mathscr{F}, \mathbb{P})$ shall be defined by setting $\Omega \triangleq E^{\mathbb{N}}$, jointly with the product σ-algebra $\mathscr{F} \triangleq \mathscr{E}^{\mathbb{N}}$ defined in the common way. Moreover, given some measurable group $(G, \mathscr{G})$, suppose that the probability measure $\mathbb{P}$ is invariant under its measurable action $\varphi : (g, \omega) \mapsto \varphi(g, \omega)$ on $(\Omega, \mathscr{F})$, i.e. $\mathbb{P} \circ \varphi(g^{-1}, \cdot\,) = \mathbb{P}$, for any $g \in G$. In this respect, we get that the action φ does not corrupt the probabilistic structure of the space $(\Omega, \mathscr{F}, \mathbb{P})$, saying that if an event in $\mathscr{F}$ occurs almost surely under $\mathbb{P}$, then it maintains the same occurrence also with respect to all the measures $\mathbb{P} \circ \varphi(g^{-1}, \cdot)$, varying $g \in G$. According to the natural interpretation, since Ω is the family of numerable sequence in E, we shall refer to G as the *perturbation group* and thus we shall call φ the *measure preserving perturbation action* of G on the space $(\Omega, \mathscr{F}, \mathbb{P})$. In addition, for notation simplicity we shall sometimes write $g\omega$ instead of $\varphi(g, \omega)$, for any $g \in G$ and $\omega \in \Omega$, as usual.

Recall that the equality $\mathbb{P}[\xi \in \cdot\,|\upsilon](\omega) = \upsilon^{\mathbb{N}}(\omega, \cdot\,)$, that is guaranteed by de Finetti's theorem when dealing with numerable exchangeable random sequences ξ in $(E, \mathscr{E})$, holds true for $\mathbb{P}$-almost any $\omega \in \Omega$. As a result, the zero probability set of $\omega \in \Omega$ such that the equality $\upsilon_g^{\mathbb{N}}(\omega, \cdot\,) = \mathbb{P}[\xi \in \cdot\,|\upsilon](g\omega)$ fails may depend on the single $g \in G$, and thus, when varying $g \in G$, the union of such sets might cover the entire sample space Ω. In order to overcome such a problem, we assume the existence of some subspace $\Omega_G \subset \Omega$ with $\mathbb{P}(\Omega_G) = 1$, that turns out to be invariant under the action of G, i.e. $\varphi(g, \Omega_G) = \Omega_G$, for any $g \in G$.

Given the random measures υ and $\upsilon^{\mathbb{N}}$ on $(E, \mathscr{E})$ and $(\Omega, \mathscr{F})$ respectively, we set $\upsilon_g(\omega, \cdot\,) \triangleq \upsilon(g\omega, \cdot\,)$ and

$\upsilon_g^{\mathbb{N}}(\omega,\cdot\,) \triangleq \upsilon^{\mathbb{N}}(g\omega,\cdot\,)$, for any $\omega \in \Omega$ and $g \in G$. In particular, since $\varphi(e,\cdot\,)$ is the identity operator on Ω, where e denotes the unitary element of G, we enforce the notation by writing υ_e and $\upsilon_e^{\mathbb{N}}$ instead of υ and $\upsilon^{\mathbb{N}}$ respectively.
Let M be some subset of the space $\mathfrak{M}_1(E)$ such that $\mathfrak{M}_{1,\mathrm{emp}}(E) \subset M$ and that for any $g \in G$ the random measure υ_g takes values in it, i.e. $\upsilon_g(\omega,\cdot\,) \in M$ for any $\omega \in \Omega$. Let also suppose that M and $\mathfrak{M}_1(T)$ are endowed with two respective metrics denoted by $d_{E,1}$ and $d_{T,1}$, consistent with their topological structures.

Definition 6 (G-Qualitative Robustness). A family of estimators $(\tau_n)_n$ defined by some statistic $\tau : M \to T$ is said *G-marginal robust* for υ with respect to $(d_{E,1}, d_{T,1})$ if (and only if) for any $\varepsilon > 0$ there exists $\delta(\varepsilon) > 0$ and $n(\varepsilon) \in \mathbb{N}$ such that the following statement

$$\text{if } g \in G \text{ with } d_{E,1}\big(\upsilon_g(\omega,\cdot\,), \upsilon_e(\omega,\cdot\,)\big) < \delta(\varepsilon), \text{ then } d_{T,1}\big(\upsilon_g^{\mathbb{N}}(\omega,\cdot\,) \circ \tau_n^{-1}, \upsilon_e^{\mathbb{N}}(\omega,\cdot\,) \circ \tau_n^{-1}\big) < \varepsilon, \text{ for any } n \geq n(\varepsilon),$$

holds true for $\mathbb{P}$-almost any $\omega \in \Omega$. In this respect, we also say that the statistic τ is G-marginal robust for υ with respect to $(d_{E,1}, d_{T,1})$.

The celebrated Hampel's theorem provides sufficient conditions for the qualitative robustness of the estimators $(\tau_n)_n$ in terms of the weak topology of measures, and in particular by letting $d_{E,1}$ and $d_{T,1}$ be the Prohorov metric (2). In such a setup, since almost surely pointwise convergence classically implies convergence in probability, Varadarajan's theorem (cf. [12], Theorem 11.4.1) guarantees that the following statement

$$\kappa_\mu(m_n, \mu) \triangleq \inf\big\{\varepsilon > 0 : \mu^{\mathbb{N}}\big\{\omega \in \Omega : \pi\big(m_n(\omega,\cdot\,), \mu\big) > \varepsilon\big\} \leq \varepsilon\big\} \to 0, \qquad \text{as } n \to +\infty, \tag{4}$$

holds true for any $\mu \in \mathfrak{M}_1(E)$, where κ_μ denotes the Ky Fan metric associated to $\mu^{\mathbb{N}}$ and defined on the random elements in the space $\mathfrak{M}_1(E)$, endowed with the measurable structure generated by the projection maps $\pi_B : \mu \mapsto \mu(B)$, for $\mu \in \mathfrak{M}_1(E)$, varying $B \in \mathscr{E}$.
However, such a result generally fails when replacing the Prohorov metric π with distances on $\mathfrak{M}_1(E)$ inducing a topological structure that is finer that the weak topology of measures. In this respect, the previous arguments naturally lead to the following notion.

Definition 7 (UGC Property). We shall say that a measurable group $(G, \mathscr{G})$ measurably acting on $(\Omega, \mathscr{F})$ admits the *UGC property* with respect to some distance $d_{E,1}$ on M jointly with the random measure υ if (and only if) the following statement

$$\sup_{g \in G} \inf\big\{\varepsilon > 0 : \upsilon_g^{\mathbb{N}}\big\{\omega \in \Omega : d_{E,1}\big(m_n(\omega,\cdot\,), \upsilon_g(\bar{\omega},\cdot\,)\big) > \varepsilon\big\}(\bar{\omega}) \leq \varepsilon\big\} \to 0, \qquad \text{as } n \to +\infty \tag{5}$$

holds true for $\mathbb{P}$-almost any $\bar{\omega} \in \Omega$.

In the latter definition, we maintained the notation classically used in literature, cf. [31, 47, 46], where the acronym *UGC* stands for *"Uniformly Glivenko Cantelli"*, even though it's sufficient to take a quick look to realize that Glivenko-Cantelli theorem (cf. [12], Theorem 11.4.2.) has almost nothing to share with property in (5). Indeed, while Glivenko-Cantelli theorem refers to the uniformly convergence of the cumulative distributions associated to the family of empirical kernels (3) defined on the real line in the i.i.d. setup, the UGC property in Definition 7 encodes the convergence in measure of such a family, which is to be understood as a random sequence in the space $\mathfrak{M}_1(E)$, where the stochastic process ξ directing the empirical kernels $(m_n)_n$ is allowed to display an internal dependence structure. Since the notation we used may generate confusing misunderstandings, we found important to highlight this issue.

Theorem 3 (*Hampel-type Theorem*)**.** *Let $\mathfrak{F}(E)$ be a dual normal consistent family of measurable functions and denote by $d_{\mathfrak{F}(E)}$ some metric consistent with the $\mathfrak{F}(E)$-weak topology on $\mathfrak{M}_1^{\mathfrak{F}}(E)$. Suppose that the data process ξ displays an internal exchangeable structure and that $\mathscr{L}(\xi_1) \in \mathfrak{M}_1^{\mathfrak{F}}(E)$. Given a measurable group*

$(G,\mathscr{G})$ satisfying the UGC property (5) with respect to $d_{\mathfrak{F}(E)}$ jointly with the random measure υ directing ξ, if the statistic $\tau:\mathfrak{M}_1^{\mathfrak{F}}(E)\to T$ is $(d_{\mathfrak{F}(E)},d_T)$-continuous and the $\mathfrak{F}(E)$-weak topology is as fine as the weak topology of measures, then the sequence $(\tau_n)_n$ turns out to be G-marginal robust for υ with respect to $(d_{\mathfrak{F}(E)},\pi)$.

We restate the proof of Theorem 2.4 in [47], which is adapted to the discussion of Hampel's theorem as stated in [22, 23].

Proof of Theorem 3. Let $\bar\Omega$ be a measurable subset of Ω such that $\mathbb{P}(\bar\Omega)=1$, fix $\bar\omega\in\bar\Omega$ and $\varepsilon>0$. We have to prove that there exists $\delta(\varepsilon)>0$ such that for any fixed $g\in G$ for which $d_{\mathfrak{F}(E)}(\upsilon_e(\bar\omega,\cdot\,),\upsilon_g(\bar\omega,\cdot\,))\le\delta(\varepsilon)$, we get $\pi(\upsilon_e^{\mathbb{N}}(\bar\omega,\cdot\,)\circ\tau_n^{-1},\upsilon_g^{\mathbb{N}}(\bar\omega,\cdot\,)\circ\tau_n^{-1})\le\varepsilon$ when n is large enough. The main idea is to exploit the following inequality,

$$\pi\big(\upsilon_e^{\mathbb{N}}(\bar\omega,\cdot\,)\circ\tau_n^{-1},\upsilon_g^{\mathbb{N}}(\bar\omega,\cdot\,)\circ\tau_n^{-1}\big)\le\pi\big(\upsilon_e^{\mathbb{N}}(\bar\omega,\cdot\,)\circ\tau_n^{-1},\delta_{\tau_\upsilon(\bar\omega)}\big)+\pi(\delta_{\tau_\upsilon(\bar\omega)},\upsilon_g^{\mathbb{N}}(\bar\omega,\cdot\,)\circ\tau_n^{-1})$$

where $\delta_{\tau_\upsilon(\bar\omega)}$ denotes the Dirac delta distribution on $(T,\mathscr{T})$ with $\tau_\upsilon(\bar\omega)$ as the point mass, and where we set $\tau_\upsilon(\bar\omega)\triangleq\tau(\upsilon_e(\bar\omega,\cdot\,))$, which turns out to be well defined since $\upsilon(\bar\omega,\cdot\,)\in\mathfrak{M}_1^{\mathfrak{F}}(E)$ due to Theorem 2. Let Ω_1 be a measurable subset of Ω such that $\mathbb{P}(\Omega_1)=1$ and fix $\bar\omega_1\in\Omega_1$. Note that, for large n we get

$$\begin{aligned}1-\varepsilon/2 &\le \upsilon_g^{\mathbb{N}}\big\{\omega\in\Omega:d_{\mathfrak{F}(E)}\big(\upsilon_g(\bar\omega_1,\cdot\,),m_n(\omega,\cdot\,)\big)\le\delta\big\}(\bar\omega_1)\\ &\le \upsilon_g^{\mathbb{N}}\big\{\omega\in\Omega:d_{\mathfrak{F}(E)}\big(\upsilon_e(\bar\omega_1,\cdot\,),m_n(\omega,\cdot\,)\big)\le d_{\mathfrak{F}(E)}\big(\upsilon_g(\bar\omega_1,\cdot\,),\upsilon_e(\bar\omega_1,\cdot\,)\big)+\delta\big\}(\bar\omega_1)\\ &\le \upsilon_g^{\mathbb{N}}\big\{\omega\in\Omega:d_{\mathfrak{F}(E)}\big(\upsilon_e(\bar\omega_1,\cdot\,),m_n(\omega,\cdot\,)\big)\le 2\delta\big\}(\bar\omega_1)\\ &\le \upsilon_g^{\mathbb{N}}\big\{\omega\in\Omega:d_T\big(\tau_\upsilon(\bar\omega_1),\tau_n(\omega)\big)\le\varepsilon/2\big\}(\bar\omega_1)\end{aligned}\tag{6}$$

due to the UGC Property (5), since it implies that $\upsilon_g^{\mathbb{N}}\{\omega\in\Omega:d_{\mathfrak{F}(E)}(\upsilon_g(\bar\omega,\cdot\,),m_n(\omega,\cdot\,))\ge\delta\}(\bar\omega)\to 0$ for $\mathbb{P}$-almost any $\bar\omega\in\Omega$ and any $\delta>0$, as $n\to+\infty$, the implication below

$$\text{if }\ d_{\mathfrak{F}(E)}\big(\upsilon_g(\bar\omega_1,\cdot\,),m_n(\omega,\cdot\,)\big)\le\delta,\quad\text{then }\ d_{\mathfrak{F}(E)}\big(\upsilon_e(\bar\omega_1,\cdot\,),m_n(\omega,\cdot\,)\big)\le d_{\mathfrak{F}(E)}\big(\upsilon_g(\bar\omega_1,\cdot\,),\upsilon_e(\bar\omega_1,\cdot\,)\big)+\delta,$$

and the $(d_{\mathfrak{F}(E)},d_T)$-continuity of τ at $\upsilon_e(\bar\omega_1,\cdot\,)$. Restating the resulting inequality (6) as follows,

$$(\delta_{\tau_\upsilon(\bar\omega_1)}\otimes\upsilon_g^{\mathbb{N}}(\bar\omega_1,\cdot)\circ\tau_n^{-1})\{(t_1,t_2)\in T\times T:d_T(t_1,t_2)\le\varepsilon/2\}\ge 1-\varepsilon/2,$$

since the space T is Polish and endowed with the related Borel σ-algebra $\mathscr{T}$ by assumption, every law in $\mathfrak{M}_1(T)$ is tight due to Ulam's theorem (cf. [12], Theorem 7.1.4). Hence, it is possible to exploit the celebrated Strassen's theorem (cf. [12], Theorem 11.6.2.) in order to get

$$\delta_{\tau_\upsilon(\bar\omega_1)}(C)\le\upsilon_g^{\mathbb{N}}(\bar\omega_1,\cdot)\circ\tau_n^{-1}(C^{\varepsilon/2})+\varepsilon/2,\qquad\text{for any }C\in\mathscr{T},$$

and thus

$$\pi\big(\delta_{\tau_\upsilon(\bar\omega_1)},\upsilon_g^{\mathbb{N}}(\bar\omega_1,\cdot\,)\circ\tau_n^{-1}\big)\le\varepsilon/2.$$

In order to conclude the proof, consider a second measurable subset Ω_2 of the sample space such that $\mathbb{P}(\Omega_2)=1$. Now, since the weak convergence of $(m_n)_n$ to $\upsilon(\bar\omega_2,\cdot)$ holds true $\upsilon^{\mathbb{N}}(\bar\omega_2,\cdot)$-a.s. for any $\bar\omega_2\in\Omega_2$ due to the exchangeability of the sequence ξ jointly with de Finetti's and Varadarajan's theorems (cf. Theorem 9.16. in [25] and Theorem 11.4.1. in [12]), then for any $\varepsilon>0$ we get $\pi(\upsilon^{\mathbb{N}}(\bar\omega_2,\cdot\,)\circ\tau_n^{-1},\delta_{\tau_\upsilon(\bar\omega_2)})\le\varepsilon/2$, when n is large. Letting $\bar\Omega\triangleq\Omega_1\cap\Omega_2$ we conclude the proof. □

Theorem 3 may be naturally generalized by providing a sort of converse. For this purpose, we shall say that the sequence of estimators $(\tau_n)_n$ is *weakly consistent* for the random measure υ with respect to the metric d_T if (and only if) it converges to $\tau(\upsilon(\bar\omega,\cdot\,))$ in $\upsilon^{\mathbb{N}}(\bar\omega,\cdot\,)$-probability, i.e. $\inf\{\varepsilon>0:\upsilon^{\mathbb{N}}\big\{\omega\in\Omega:d_T(\tau(\upsilon(\bar\omega,\cdot\,)),\tau_n(\omega))>\varepsilon\}(\bar\omega)\le\varepsilon\big\}\to 0$, as $n\to 0$, for $\mathbb{P}$-almost any $\bar\omega\in\Omega$.

Theorem 4 (*Converse of Hampel-type Theorem*)**.** *Let us maintain the same notation from Theorem 3. Suppose that $(\tau_n)_n$ is G-marginal robust for υ with respect to $(d_{\mathfrak{F}(E)}, \pi)$ and that $(\tau_n)_n$ is weakly consistent for υ with respect to d_T, then $\tau : \mathfrak{M}_1^{\mathfrak{F}}(E) \to T$ is $(d_{\mathfrak{F}(E)}, d_T)$-continuous at $\upsilon_e(\omega, \cdot\,)$ for $\mathbb{P}$-almost any $\omega \in \Omega$.*

Proof. Let $\bar{\Omega}$ be some subset of Ω such that $\mathbb{P}(\bar{\Omega}) = 1$ and fix $\bar{\omega} \in \bar{\Omega}$. We have to prove that for any $\varepsilon > 0$ there exists $\delta(\varepsilon) > 0$ such that if $d_{\mathfrak{F}(E)}(\upsilon_g(\bar{\omega}, \cdot\,), \upsilon_e(\bar{\omega}, \cdot\,)) < \delta(\epsilon)$ for some $g \in G$, then $d_T(\tau_{\upsilon_g}(\bar{\omega}), \tau_{\upsilon_e}(\bar{\omega})) < \varepsilon$. For this purpose, the proof of Theorem 2.6 in [30] can be naturally adapted in the present setup by considering the following triangular inequality

$$\begin{aligned} d_T(\tau_{\upsilon_g}(\bar{\omega}), \tau_{\upsilon_e}(\bar{\omega})) &= \pi\big(\delta_{\tau_{\upsilon_g}(\bar{\omega})}, \delta_{\tau_{\upsilon_e}(\bar{\omega})}\big) \\ &\leq \pi\big(\delta_{\tau_{\upsilon_g}(\bar{\omega})}, \upsilon_g^{\mathbb{N}}(\bar{\omega}, \cdot\,) \circ \tau_n^{-1}\big) + \pi\big(\upsilon_g^{\mathbb{N}}(\bar{\omega}, \cdot\,) \circ \tau_n^{-1}, \upsilon_e^{\mathbb{N}}(\bar{\omega}, \cdot\,) \circ \tau_n^{-1}\big) + \pi\big(\upsilon_e^{\mathbb{N}}(\bar{\omega}, \cdot\,) \circ \tau_n^{-1}, \delta_{\tau_{\upsilon_e}(\bar{\omega})}\big) \end{aligned}$$

The proof concludes by noting that, while the first and the third summand tend to zero as $n \to +\infty$, since $(\tau_n)_n$ is assumed to be weakly consistent for υ with respect to d_T, the second summand can be bounded from above by any $\varepsilon > 0$ for n large enough, whenever we fix $\delta(\varepsilon) > 0$ such that $d_{\mathfrak{F}(E)}(\upsilon_g(\bar{\omega}, \cdot\,), \upsilon_e(\bar{\omega}, \cdot\,)) < \delta(\varepsilon)$, due to the G-qualitative robustness of $(\tau_n)_n$ for υ with respect to $(d_{\mathfrak{F}(E)}, d_T)$. □

Note that, setting $\mathfrak{F}(E) = \mathfrak{C}_b(E)$, and thus letting $d_{\mathfrak{F}(E)}$ be some metric for the weak convergence of measures, the $(d_{\mathfrak{F}(E)}, d_T)$-continuity of τ turns out to be equivalent to the G-qualitative robustness of $(\tau_n)_n$ for the random measure υ with respect to $(d_{\mathfrak{F}(E)}, \pi)$, since both the UGC property (5) displayed by the group G and the weak consistency of the sequence $(\tau_n)_n$ are always guaranteed in such a case. The reason of these additional assumptions lies behind our desire to deal with some topological structure that turns out to be finer than the topology for the weak convergence.
Moreover, developing the present theory in a overall setup, by considering random measures on $(E, \mathscr{E})$ instead of the generic couple of law in $\mathfrak{M}_1(E)$, allows us to assess the notion of robustness in terms of corrupted dataset. Indeed, since every sample ω encodes a certain sequence of observations, assessing the impact of the perturbation in terms of the distance between the laws $\upsilon_g(\omega, \cdot\,)$ and $\upsilon_e(\omega, \cdot\,)$, by varying $g \in G$, allowed us to restate the notion of robustness when altering the historical observations. In this respect, robustness turns out to be a form of stability that the laws induced on $(T, \mathscr{T})$ by the family $(\tau_n)_n$ asymptotically display when data are perturbed, and it strongly depends on the continuity of the statistic τ, due to Theorem 3. On the other hand, continuity represents a suitable form of analytical regularity also on a practical level. Indeed, since from the regulatory point of view every monetary risk measure can be seen as the minimal amount of money to put aside in order to hedge the exposure, the continuity of the predictor statistic implies that little forecasting mistakes can be controlled by injecting just limited amounts of capital as safety fund. As a result, Theorem 3 claims essentially that the performance of the risk measure from the regulatory point of view strongly impacts on the stability of the associated family of estimators in terms of robustness. More than that, when looking at it jointly with Theorem 4, one can interpret these two forms of regularity as the two side of the same coin.

5. CONVEX LAW INVARIANT RISK MEASURES

Everyone owns an innate feeling of how the risk associated to some market exposure is to be understood. Nevertheless, it is not ostensibly possible to properly assess it in quantitative terms, unless a suitable risk evaluation procedure has been *a priori* defined. Within the drama of the modern risk management, such a role is recovered by statistics that are defined by considering certain risk measures. In particular, given a law invariant monetary risk measure $\rho : \mathfrak{X} \to \mathbb{R}$, where $\mathfrak{X}$ denotes some family of random elements in $\mathbb{R}$ endowed with the Borel σ-algebra $\mathscr{B}(\mathbb{R})$, a possible statistic is obtained by considering the associated distribution-based risk functional

$$\mathfrak{R}_\rho : \mu \mapsto \mathfrak{R}_\rho(\mu) \triangleq \rho(\xi), \qquad \text{for any } \xi \in \mathfrak{X} \text{ such that } \mathscr{L}(\xi) = \mu, \tag{7}$$

defined for any μ lying in $\mathfrak{M}(\mathfrak{X}) \triangleq \{\mathscr{L}(\xi) : \xi \in \mathfrak{X}\}$. Hence, assuming that $\mathfrak{M}_{1,\mathrm{emp}}(\mathbb{R}) \subset \mathfrak{M}(\mathfrak{X})$, the sequence of real valued random variables $\rho_n \triangleq \mathfrak{R}_\rho(m_n)$, obtained by varying $n \geq 1$, classically defines a family of estimators for the risk measure ρ, as stated in Definition 4.
Notwithstanding the analytical properties of the functional (7) have been classically assessed in terms of the standard weak topology of measures, we already claimed that such a formulation may generate a number of debated complications. First of all, since the tail behaviours of the distributions is completely neglected in such a setup, it does not appear ostensibly possible to formulate a notion of qualitative robustness in a meaningful way. Worst still, the distribution-based risk functionals associated to most of the common law invariant risk measures lack to be continuous when endowing its domain with such a topological structure. Nevertheless, according to Theorem 1, 2 and 3, the continuity of the functional (7) turns out to be the main pillar when assessing the asymptotic stability of the risk estimators $(\rho_n)_n$ in terms of consistency and robustness.
As observed in [30, 31, 47], it is possible to easily overcome such problems by taking into consideration some suitable refinement of the standard weak topology, which is able to capture the tail behaviour of the distributions. In this respect, the performing idea is to look at the topology generated by the duality with some family $\mathfrak{F}(\mathbb{R})$ including functions that may be unbounded outside a certain compact interval of the real line.
According to Definition 3, the most appealing setting of this type is probably achieved in [31] by considering the ψ-weak topology on the space $\mathfrak{M}_1^\psi(\mathbb{R})$ obtained by setting $\psi \triangleq \phi(|\cdot|)$, for some left-continuous increasing and convex function $\phi : [0, +\infty) \to [0, +\infty)$ such that $\phi(0) = 0$ and $\phi(x) \to +\infty$, as $x \to +\infty$. In this respect, the role of the function ψ is to be understood as a form of penalization of the distributions on their tails, with the main attempt to recover the information regarding the extreme events that are completely neglected in the standard setup. Besides, the asymptotic nature of such a function clearly impacts on the analytical properties displayed by the generic functional defined on the space $\mathfrak{M}_1^\psi(\mathbb{R})$. Indeed, letting ρ be some convex and law-invariant risk measure such that the related distribution-based risk functional $\mathfrak{R}_\rho$ can be properly defined on the space $\mathfrak{M}_1^\psi(\mathbb{R})$, according to Theorem 2.8 in [31] the continuity of $\mathfrak{R}_\rho$ with respect to the ψ-weak topology is guaranteed only when the function ϕ satisfies the so-called Δ_2-condition,

$$\text{there exists a constant } C > 0 \text{ and } x_0 \in \mathbb{R} \text{ such that } \phi(2x) \leq C\phi(x), \text{ for any } x \geq x_0. \tag{8}$$

More precisely, since we implicitly assume to deal with the representation $\psi \triangleq \phi(|\cdot|)$, with notational abuse we say that the function ψ satisfies the Δ_2-condition when the statement (8) holds true for the function ϕ. In the present framework, Theorem 1, 2 and thus Corollary 1, may be properly restated as follows.

Theorem 5 (Consistency in the ψ-weak topology)**.** *Let ρ be a convex law invariant risk measure such that the related distribution-based functional (7) can be properly defined on the space $\mathfrak{M}_1^\psi(\mathbb{R})$. Then, in each of the following cases*

i. ξ is a stationary sequence of real valued random variables,

ii. ξ is an exchangeable sequence of real valued random variables,

iii. ξ is a spreadable sequence of real valued random variables,

if $\mathscr{L}(\xi_1) \in \mathfrak{M}_1^\psi(\mathbb{R})$ and ψ satisfies the Δ_2-condition (8), we get that the sequence $(\rho_n)_n$ provides a family of strong consistent estimators for $\mathfrak{R}_\rho(\upsilon)$, where υ is a regular version of the conditional distribution $\mathbb{P}[\xi_1 \in \cdot \,|\xi^{-1}\mathscr{B}(\mathbb{R})^{\mathbb{N}}_{\mathrm{inv}(\Sigma)}]$.

Since the continuity of the statistic defining the risk estimators $(\rho_n)_n$ constitutes the main pillar also when investigating the notion of qualitative robustness, condition (8) can be properly exploited in order to restate Theorem 3 in the current setup. Besides, since according to the circle of ideas lying behind Lemma 2, the ψ-weak topology on $\mathfrak{M}_1^\psi(\mathbb{R})$ can be properly defined as the topology generated by the duality with the family $\mathfrak{C}_b(\mathbb{R})\cup\{\psi\}$, the UGC property (5) strongly simplifies in the current setup. Indeed, it is nothing short of clear that when looking at the metric $d_\psi(\mu,\nu) \triangleq \pi(\mu,\nu) + |(\mu-\nu)\psi|$, for $\mu,\nu \in \mathfrak{M}_1^\psi(\mathbb{R})$, jointly with the random measure υ, it easily boils down to the following condition,

$$\sup_{g\in G} \inf\big\{\varepsilon > 0 : \upsilon_g^{\mathbb{N}}\big\{\omega\in\Omega : \big|\big(m_n(\omega,\cdot\,) - \upsilon_g(\bar\omega,\cdot\,)\big)\psi\big| > \varepsilon\big\}(\bar\omega) \le \varepsilon\big\} \to 0, \qquad \text{as } n\to+\infty, \tag{9}$$

for $\mathbb{P}$-almost any $\bar\omega\in\Omega$. Hence, in the present setting, Theorem 3 is restated as follows, by providing the natural generalization of Theorem 2.15 in [31] when dealing with our refined framework, since the notion of independence has been assessed in conditional terms.

Theorem 6 (Hampel-type Theorem for the ψ-weak topology)**.** *Let ρ be a convex law invariant risk measure such that the related distribution-based functional (7) can be properly defined on the space $\mathfrak{M}_1^\psi(\mathbb{R})$. Suppose that the data process ξ displays an internal exchangeable structure and that $\mathscr{L}(\xi_1)\in\mathfrak{M}_1^\psi(E)$. Given a measurable group $(G,\mathscr{G})$ satisfying the UGC property (9) with respect to d_ψ jointly with the random measure υ directing ξ, if ψ satisfies the Δ_2-condition (8), then the sequence $(\rho_n)_n$ turns out to be G-marginal robust for υ with respect to (d_ψ,π).*

REFERENCES

[1] Aldous, D. J. *Exchangeability and related topics.* Springer, 1985.

[2] Aliprantis, C. D., and Border, K. *Infinite dimensional analysis: a hitchhiker's guide.* Springer Science & Business Media, 2006.

[3] Belomestny, D., Krätschmer, V., et al. Central limit theorems for law-invariant coherent risk measures. *Journal of Applied Probability 49*, 1 (2012), 1–21.

[4] Bogachev, V. I. *Measure theory*, vol. 2. Springer Science & Business Media, 2007.

[5] Cont, R., Deguest, R., and Scandolo, G. Robustness and sensitivity analysis of risk measurement procedures. *Quantitative Finance 10*, 6 (2010), 593–606.

[6] Cuevas, A. Qualitative robustness in abstract inference. *Journal of statistical planning and inference 18*, 3 (1988), 277–289.

[7] Cuevas, A., and Romo, J. On robustness properties of bootstrap approximations. *Journal of statistical planning and inference 37*, 2 (1993), 181–191.

[8] de Finetti, B. Sull'approssimazione empirica di una legge di probabilità. *Giornale dell'Istituto Italiano degli Attuati 4*, 3 (1933), 415–420.

[9] de Finetti, B. La prévision: ses lois logiques, ses sources subjectives. In *Annales de l'institut Henri Poincaré* (1937), vol. 7, pp. 1–68.

[10] de Finetti, B. La probabilitá e la statistica nei rapporti con l'induzione, secondo i diversi punti di vista. In *Induzione e statistica.* Springer, 2011, pp. 1–122.

[11] Dudley, R. M. *Uniform central limit theorems*, vol. 23. Cambridge Univ Press, 1999.

[12] Dudley, R. M. *Real analysis and probability*, vol. 74. Cambridge University Press, 2002.

[13] Engelking, R. *General Topology*, vol. 47. Polish Sci. Publ., 1992.

[14] Föllmer, H., and Schied, A. *Stochastic finance: an introduction in discrete time.* Walter de Gruyter, 2011.

[15] Föllmer, H., and Weber, S. The axiomatic approach to risk measures for capital determination. *Annual Review of Financial Economics 7*, 1 (2015).

[16] Freedman, D. A. Invariants under mixing which generalize de finetti's theorem. *The Annals of Mathematical Statistics 33*, 3 (1962), 916–923.

[17] Gänssler, P., and Stute, W. Empirical processes: a survey of results for independent and identically distributed random variables. *The Annals of Probability* (1979), 193–243.

[18] Ghosh J, R. R. *Bayesian Nonparametrics.* Springer, 2003.

[19] Hampel, F. R. Contributions to the theory of robust estimation. *Ph.D. thesis, Univ. California, Berkeley* (1969), 1887–1896.

[20] Hampel, F. R. A general qualitative definition of robustness. *The Annals of Mathematical Statistics* (1971), 1887–1896.

[21] Hewitt, E., and Savage, L. J. Symmetric measures on cartesian products. *Transactions of the American Mathematical Society 80*, 2 (1955), 470–501.

[22] Huber, P. J. *Robust statistics.* Springer, 2011.

[23] Huber, P. J., and Ronchetti, E. M. Robustness of design. *Robust Statistics, Second Edition* (1975), 239–248.

[24] Kallenberg, O. Spreading and predictable sampling in exchangeable sequences and processes. *The Annals of Probability* (1988), 508–534.

[25] Kallenberg, O. *Foundations of modern probability.* Springer Science & Business Media, 2006.

[26] Kallenberg, O. *Probabilistic symmetries and invariance principles.* Springer Science & Business Media, 2006.

[27] Kingman, J. F. Uses of exchangeability. *The Annals of Probability* (1978), 183–197.

[28] Koch-Medina, P., and Munari, C. Law-invariant risk measures: Extension properties and qualitative robustness. *Statistics & Risk Modeling 31*, 3-4 (2014), 215–236.

[29] Kou, S., Peng, X., and Heyde, C. C. External risk measures and basel accords. *Mathematics of Operations Research 38*, 3 (2013), 393–417.

[30] Krätschmer, V., Schied, A., and Zähle, H. Qualitative and infinitesimal robustness of tail-dependent statistical functionals. *Journal of Multivariate Analysis 103*, 1 (2012), 35–47.

[31] Krätschmer, V., Schied, A., and Zähle, H. Comparative and qualitative robustness for law-invariant risk measures. *Finance and Stochastics 18*, 2 (2014), 271–295.

[32] Krätschmer, V., Schied, A., and Zähle, H. Quasi-hadamard differentiability of general risk functionals and its application. *Statistics & Risk Modeling 32*, 1 (2015), 25–47.

[33] Lacker, D. Law invariant risk measures and information divergences. *arXiv preprint arXiv:1510.07030* (2015).

[34] Lacker, D. Liquidity, risk measures, and concentration of measure. *arXiv preprint arXiv:1510.07033* (2015).

[35] Lauer, A., and Zähle, H. Nonparametric estimation of risk measures of collective risks. *Statistics & Risk Modeling* (2015).

[36] Mizera, I., et al. Qualitative robustness and weak continuity: the extreme unction. *Nonparametrics and Robustness in Modern Statistical Inference and Time Series Analysis: A Festschrift in honor of Professor Jana Jurecková 1* (2010), 169.

[37] Parthasarathy, K. R. *Probability measures on metric spaces*, vol. 352. American Mathematical Soc., 1967.

[38] Prokhorov, Y. V. Convergence of random processes and limit theorems in probability theory. *Theory of Probability & Its Applications 1*, 2 (1956), 157–214.

[39] Rachev, S. T., and Rüschendorf, L. *Mass Transportation Problems: Volume I: Theory*, vol. 1. Springer Science & Business Media, 1998.

[40] Rao, R. R. Relations between weak and uniform convergence of measures with applications. *The Annals of Mathematical Statistics* (1962), 659–680.

[41] Ryll-Nardzewski, C. On stationary sequences of random variables and the de finetti's equivalence. In *Colloquium Mathematicae* (1957), vol. 4, Institute of Mathematics Polish Academy of Sciences, pp. 149–156.

[42] Sinai, Ya, S. *Dynamic System II: Ergodic Theory with Applications to Dynamical Systems and Statistical Mechanics*. Springer-Verlag Berlin Heidelberg, 1989.

[43] Strassen, V. The existence of probability measures with given marginals. *The Annals of Mathematical Statistics* (1965), 423–439.

[44] Williams, D. *Probability with Martingales*. Cambridge University Pres, 1991.

[45] Zähle, H. Qualitative robustness of von mises statistics based on strongly mixing data. *Statistical Papers 55*, 1 (2014), 157–167.

[46] Zähle, H. Qualitative robustness of statistical functionals under strong mixing. *Bernoulli 21*, 3 (2015), 1412–1434.

[47] Zähle, H. A definition of qualitative robustness for general point estimators, and examples. *Journal of Multivariate Analysis 143* (2016), 12–31.