ATHENS UNIVERSITY OF ECONOMICS AND BUSINESS

DEPARTMENT OF ECONOMICS

WORKING PAPER SERIES 9-2015

Saddle-Type Functionals for Continuous Processes with Applications to Tests for Stochastic Spanning

Stelios Arvanitis

76 Patission Str., Athens 104 34, Greece
Tel. (++30) 210-8203911 - Fax: (++30) 210-8203301
www.econ.aueb.gr

Saddle-Type Functionals for Continuous Processes with Applications to Tests for Stochastic Spanning

Stelios Arvanitis

Department of Economics, Athens University of Economics and Business.

Patision 76, P.O. Box 10434, Athens, Greece.

Abstract

We derive the continuity properties of the cdf of a random variable defined as a saddle-type point of a real valued continuous stochastic process on a compact metric space. This result facilitates the derivation of first order asymptotic properties of tests for stochastic spanning w.r.t. some stochastic dominance relation based on subsampling. As an illustration we define the concept of Markowitz stochastic spanning, derive an analytic representation upon the empirical analog of which we construct a relevant statistical test. The aforementioned result enables derivation of asymptotic exactness for the relevant procedure based on subsampling, when the metric space has the form of a simplicial complex, the spanning set is a compact subset and the significance level is chosen according to the number of extreme points of the complex inside the spanning set. Consistency is also derived. Such tests are of interest in financial economics since they can provide reductions of portfolio sets.

Keywords: Continuous Process, Malliavin Derivative, Nested Optimizations, Saddle-Type Point, Connected Support, Atom, Absolute Continuity, Markowitz Stochastic Dominance, Stochastic Spanning, Spanning Test, Subsampling, Gaussian Process, Brownian Bridge, Asymptotic Exactness, Consistency.

1 Introduction

We derive the continuity properties of the cdf of a random variable defined by a saddle-type functional of a real valued continuous stochastic process defined on a compact metric space. Our motivation stems from the fact that the limit theory of tests for *stochastic spanning* usually involves weak limits represented as a finite recursion of optimization functionals applied on some relevant Gaussian process. The possibility of the existence of atoms in this distribution directly affects the issue of asymptotic exactness of the aforementioned tests when those are based on resampling procedures such as bootstrap and subsampling.

The notion of stochastic spanning is a brilliant idea of Thierry Post, influenced by the notion of M-V spanning in Huberman and Kandell [6], that was formulated in the context of second order stochastic dominance in Arvanitis et al. [1]. It can be easily generalized on the framework of an arbitrary preorder defined on some set of probability distributions. Given such a preorder, and if the efficient set of the preorder, the set that maximal elements is, non-empty, a spanning subset of the preorder is essentially any superset of the efficient set.[1] As such a spanning set can either be used to provide an "outer approximation" of the underlying efficient set, and/or, when small enough, to provide with a desirable reduction of the initial set of distributions which could be very large. In such a case the issue of optimal choice could be reduced to a potentially computationally easier problem. Both issues could be of great interest to financial economics since in the usual applications the underlying distributions represent returns of financial assets and the preorders are stochastic dominance rules that reflect classes of utility functions (e.g. for the first and second order, as well as the Prospect and Markowitz stochastic dominance rules and their relations to classes of utilities see Levy and Levy [7]). Obviously those notions could also be of potential interest in any field of economic theory or decision science that examines optimal choice under uncertainty.

Given the previous a natural question arises. Assume that we are given some subset of the underlying set of distributions and we want to ascertain whether the former spans the latter w.r.t. the preorder. When the two sets are not equal, it is in some cases the fact that spanning occurs if and only if a functional defined by a complex recursion of optimizations w.r.t. the given sets is zero (see for example the discussion in page 6 of Arvanitis et al. [1] for the case of second order stochastic dominance, or Proposition 1 below for the case of Markowitz stochastic dominance). The verification of the above is usually analytically intractable due to the dependence of the functional on the generally unknown underlying distributions and/or due to the complexity of the optimizations involved. Hence this cannot be directly used. However using the principle of analogy statistical tests can be designed for testing the null hypothesis of spanning. Due to analogous difficulties the asymptotic critical values are not analytically tractable, but can be approximated by resampling procedures, whence the usefullness of the core result of the paper arises, since it becomes relevant for the establishment of asymptotic exactness.[2]

In this respect the second section of the paper sets up the probabilistic framework and derives the aforementioned result about properties of the law of a random variable defined by a finite number of nested optimizations on a continuous process w.r.t. possibly interdependent parameter spaces. More specifically, under weak conditions involving Malliavin differentiability, existence of moments for suprema as well as a countability property for the singular points of the derivative, we derive connectedness for the support of the law, a countable number of atoms, and absolute continuity when restricted between successive atoms. The present result is a non-trivial

[1] The exact definition of spanning, enables the existence of spanning sets even if the preorder has no maximal elements.

[2] Notice that spanning tests subsume as special cases the relevant test of efficiency w.r.t. the underlying preorder. Hence procedures developed in papers such as Post and Versijp [12], Scaillet and Topaloglou [14], Arvanitis and Topaloglou [2] etc, can be considered as spanning tests for singleton spanning sets.

extension that can be used for the derivation of analogous results w.r.t. more complex preorders and it simultaneously extends relevant results concerning suprema of analogous stochastic processes (see section 2 for references).

As an illustration of the previous remark, in the third section of the paper we derive a testing procedure along with its first order limit theory for the concept of spanning w.r.t. the Markowitz stochastic dominance preorder. We are doing so, by defining the notion and providing with an original characterization of spanning by the zero of an analogous to the aforementioned functionals. Using this, we define the test statistic, derive its limit distribution under the null, define a subsampling algorithm for the approximation of the asymptotic critical values and among others *use the main result* for the derivation of asymptotic exactness. We also derive consistency. Notice that the arguments for the derivation of the limit theory involve among others the determination of the asymptotic behavior of several random elements when restricted to complex partitions of the parameter spaces involved. In this respect the arguments are similar, though not the same, to the ones used in Arvanitis, Hallam and Post [1]. One major difference is the fact that in the present case we do not need the assumption of the boundendess of the support of the involved distributions. Hence in this sense the present results could also be considered as generalizations to the ones derived in the aforementioned paper. In the final section we conclude.

2 Assumption Framework and Main Result

Suppose that $\Lambda_1, \Lambda_2, \ldots, \Lambda_s$ are compact metric spaces, and consider $\Lambda = \prod_{i=1}^{s} \Lambda_i$ equipped with the product topology. Furthermore, consider the functional oper $\doteqdot$ $\text{opt}_1 \circ \text{opt}_2 \circ \cdots \circ \text{opt}_s$ where $\text{opt}_i = \sup$ or $\inf$ w.r.t. to some non-empty compact $\Lambda_i^{\star} \subseteq \Lambda_i$ for $i = 1, \ldots, s$. When $i > 1$ then $\Lambda_i^{\star}$ is allowed to depend on the elements of $\prod_{j=1}^{i-1} \Lambda_{i-j}^{\star}$, $i > 1$.

The probabilistic framework follows closely Chapter 2 of Nualart [10]. In this respect it consists of $(\Omega, \mathcal{F}, \mathbb{P})$, a complete probability space, where $\mathcal{F}$ is generated by some isonormal Gaussian process $W = \{W(h), h \in H\}$ where H is an appropriate Hilbert space. X is some real valued stochastic process on Λ with continuous sample paths (i.e. its paths are $\mathbb{P}$ a.s. elements of $C(\Lambda, \mathbb{R})$). In many applications X is a Gaussian weak limit for some net of appropriate processes. D denotes the Malliavin derivative operator and $\mathbb{D}^{1,2}$ the completion of the family of Malliavin differentiable random variables w.r.t. the norm $\sqrt{\mathbb{E}\left(z^2 + (Dz)^2\right)}$.

We are interested in the form of the support and the continuity properties of the cdf of the law of the random variable $\xi := \text{oper} X_{\lambda}$. The following assumption describes sufficient conditions for the aforementioned law to have a countable number of atoms while being absolutely continuous when restricted between their successive pairs. Given this, the result to be established below, allows, first for the random variable at hand to be defined by complex saddle type functionals,[3] and second for discontinuities. Hence it generalizes established results concerning the absolute con-

[3] The term "saddle-type" is obviously used in an abusive manner, since in general commutativity between the successive optimization functionals does not hold.

tinuity of the distribution of suprema of stochastic processes. For an excellent treatment of those results see, inter alia, Propositions 2.1.7 and 2.1.10 of Nualart[10] as well as the literature on the fibering method and its probabilistic applications, e.g. Lifshits [8].

Assumption 1. *For the stochastic process $X : \Omega \to C(\Lambda, \mathbb{R})$ suppose that:*

1. $\mathbb{E}\left[\sup_\Lambda X_\lambda^2\right] < +\infty$.
2. *For all $\lambda \in \Lambda$, $X(\lambda) \in \mathbb{D}^{1,2}$, and the H -valued process DX has a continuous version and*
$$\mathbb{E}\left[\sup_\Lambda \|DX_\lambda\|^2\right] < +\infty.$$
3. *There exists a countable subset of $\mathbb{R}$, say T, such that the relation*
$$\mathbb{P}\left(\{\xi = t\} \cap \Omega_t\right) \geq 0$$
holds if and only if $t \in T$, where
$$\Omega_t = \{\omega \in \Omega : DX_\lambda(\omega) = 0 \, \textit{for some } \lambda \textit{ such that } t = X_\lambda(\omega)\}.$$

Remark 1. *In the usual case, i.e. where X is zero-mean Gaussian, the first condition can be established by strong results that imply the subexponentiality of the distribution of $\sup_\Lambda X_\lambda$, such as Lemma A.2.7 of Van Der Vaart and Wellner [16]. This would follow from conditions that restrict the packing numbers of $\Lambda \times \mathbb{R}$ metrized as a totally bounded metric space by the use of the covariance function of X, to be polynomially bounded, something that is easily established if the Λ_i are subsets of Euclidean spaces for all i. In the same respect, the second condition is easily established as in Nualart [10] (see page 110). More specifically, if $K(\lambda_1, \lambda_2)$ is the aforementioned covariance function, then H is the closed span of $\{h_\lambda(\cdot) = K(\lambda, \cdot), \lambda \in \Lambda\}$, with inner product $\left\langle h_{\lambda_1}, h_{\lambda_2}\right\rangle_H = K(\lambda_1, \lambda_2)$, whence $DX_\lambda = K(\lambda, \lambda)$. In this case the previous along with dominated convergence would imply the existence of $\mathbb{E}\left[\sup_\Lambda \|DX_\lambda\|^2\right]$. The third condition is the most difficult to establish. In the cases that we have in mind, "outer approximations" of T can be derived by analogous, as well as easier to establish, properties of random variables that are stochastically dominated by ξ, see for example the corollary below.*

We are now able to state and prove the main result.

Theorem 2. *Under Assumption 1 the law of ξ has connected support, say $\mathrm{supp}(\xi)$, that contains T. If $t \in T$ then the cdf of the law evaluated at t has a jump discontinuity of size at most $\mathbb{P}(\Omega_t)$. If t_1, t_2 are successive elements of T then the law restricted to (t_1, t_2) is absolutely continuous w.r.t. the Lebesgue measure. If T has infimum then the law restricted to $(-\infty, \inf T)$ is absolutely continuous w.r.t. the Lebesgue measure. Dually if T has supremum then the law restricted to $(\sup T, +\infty)$ is absolutely continuous w.r.t. the Lebesgue measure.*

Proof. First notice that that $\xi \in \mathbb{D}^{1,2}$. This follows by the use of similar arguments to the ones in the proof of Proposition 2.1.10 of Nualart [10]. Precisely, consider a countable dense subset of Λ , say Λ_∞ as well as $\xi_n := \text{oper}X_\lambda$ where opt_i is considered w.r.t. $\Lambda^\star_{i,n}(\lambda_{i-1}) = \left\{\text{the first } n \text{elements of } \Lambda^\star_i(\lambda_{i-1}) \cap \text{pr}_i\Lambda_\infty\right\}$ and $\lambda_{i-1} \in \Lambda^\star_{i-1,n}$ when $i > 1$. The function $\text{oper} : C(\Lambda, \mathbb{R}) \to \mathbb{R}$ is Lipschitz, hence due to Proposition 1.2.4 of Nualart [10] $\eta_n \in \mathbb{D}^{1,2}$. Furthermore, due to Assumption 1.1 $\xi_n \to \xi$ in $L^2(\Omega)$ and therefore the preliminary result will follow if $(D\xi_n)_{n\in\mathbb{N}}$ is $L^2(\Omega)$ bounded. Define

$$A_n = \left\{\omega \in \Omega : \xi_n = X_{\lambda_n}, \xi_n \neq X_{\lambda_k}, \forall k < n\right\}.$$

Using the local property of D we have that

$$D\xi_n = \sum_{n\in\mathbb{N}} 1_{A_n} DX_{\lambda_n}$$

and thereby

$$\mathbb{E}\|D\xi_n\|^2_H < +\infty$$

due to Assumption 1.2. Then Assumption 1.3 as well as Proposition 2.1.7 of Nualart [10] imply the first part of the Theorem. For the rest assume first that T is empty. Then the result will follow from a series of arguments almost identical to the ones in the proof of Proposition 2.1.11 of Nualart[10]. Specifically, consider the set

$$G = \{\omega \in \Omega : \text{ there exists } \lambda \in \Lambda \text{ such that } DX_\lambda \geq D\xi \text{ and } X_\lambda = \xi\}$$

and using Λ_∞ above H_∞ a countable dense subset of the unit ball of H, and $B_r(\lambda)$ the ball in Λ with center λ and radius $r > 0$ we have that

$$G \subseteq \cup_{\lambda\in\Lambda_\infty, r\in\mathbb{Q}_{++}, k\in\mathbb{N}_0, h\in H_\infty} G_{\lambda,r,k,h}$$

i.e. a countable union, where

$$G_{\lambda,r,k,h} := \left\{\omega \in \Omega : \langle DX_{\lambda'} - D\eta, h\rangle > \frac{1}{k} \text{ for all } \lambda' \in B_r(\lambda)\right\}$$

$$\cap \{\text{oper}X_{\lambda'} = \xi\}$$

For some λ, r, k, h as above, define $\xi' = \text{oper}X_{\lambda'}$, where now opt_i is considered w.r.t. $\Lambda^\star_i(\lambda_{i-1}) \cap \text{pr}_1 B_r(\lambda)$ choose a countable dense subset of $B_r(\lambda)$, say $B^\infty_r(\lambda)$ and using

$$\Lambda^\infty_{i,n}(\lambda_{i-1}) = \left\{\text{the first } n \text{elements of } \Lambda^\star_i(\lambda_{i-1}) \cap \text{pr}_i B^\infty_r(\lambda)\right\}$$

define $\xi^{'}_n = \text{oper}X_\lambda$ analogously. We have that as $n \to \infty$ $\xi^{'}_n \to \xi'$ in $L^2(\Omega)$ norm due to 1.1. Due to Lemma 1.2.3 of Nualart [10] and 1.2 we also have that $D\xi^{'}_n \to D\xi'$ in the weak topology of $L^2(\Omega, H)$. Using again the local property argument as above we have that for any $\omega \in G_{\lambda,r,k,h}$ $D\xi^{'}_n = DX_{\lambda'}$ for some $\lambda' \in B^\infty_r(\lambda)$. But for such ω we have that $\left\langle D\xi^{'}_n - D\xi^{'}, h\right\rangle > \frac{1}{k}$ for all n. This directly implies that $\mathbb{P}\left(G_{\lambda,r,k,h}\right) = 0$

which due to the countability implies that $\mathbb{P}(G) = 0$. Then the result follows from Theorem 2.1.3 of Nualart [10]. Now suppose that $t \in T$ and consider

$$\mathbb{P}(\xi = t) = \mathbb{P}(\{\xi = t\} \cap \Omega_t) + \mathbb{P}(\{\xi = t\} \cap \Omega_t^c)$$

If for some $t \in T$, $\mathbb{P}(\Omega_t^c) > 0$ notice that

$$\mathbb{P}(\{\xi = t\} \cap \Omega_t^c) = \mathbb{P}(\xi\text{=t}/\Omega_t^c)\,\mathbb{P}(\Omega_t^c),$$

and consider the process $X^\star := X\big|_{\Omega - \cup_{t\in T}\Omega_t^c}$ that obviously satisfies Assumption 1 with $T^\star = \emptyset$ along with the obvious change of notation. Hence $\xi^\star$ has an absolutely continuous law something that implies that $\mathbb{P}(\xi\text{=t}/\Omega_t^c) = \mathbb{P}(\xi^\star\text{=t}) = 0$. If $\mathbb{P}(\Omega_t^c) = 0$ then trivially $\mathbb{P}(\{\xi = t\} \cap \Omega_t^c) = 0$ establishing that $\mathbb{P}(\xi = t) = \mathbb{P}(\{\xi = t\} \cap \Omega_t^c)$ in any case. Now suppose that t_1, t_2 are successive elements of T and consider $\Omega_{t_1,t_2} = \{\omega \in \Omega : \xi \in (t_1, t_2)\}$. The previous imply that $\mathbb{P}\left(\Omega_{t_1,t_2}\right) > 0$, hence the process $X_\star := X\big|_{\Omega_{t_1,t_2}}$ satisfies Assumption 1 with $T_\star = \emptyset$ and thereby $\xi_\star$ has an absolutely continuous law. The other cases follow analogously when the intersections apperaring in the Theorem are non empty. When empty the results are trivial. □

Notice that the previous result encompasses the standard absolute continuity results in the aforementioned literature that hold when oper is a composition of suprema, the parameter spaces are independent, and $\mathbb{P}(\Omega_t) = 0$ for all $t \in T$. Even in the special case where T is a singleton, the result is a generalization of Theorem 2 of Lifshits [8] since it allows for non-Gaussianity, dependence between the factors of the parameter space, as well as saddle-type functionals. The following corollary focuses on this particular case and estimates the size of the potential jump discontinuity by assuming the existence of an auxiliary first order stochastically dominated random variable.

Corollary 1. *Suppose that Assumption 1 is satisfied. Furthermore suppose that* $T = \{c\}$, $\xi \geq \eta$ $\mathbb{P}$ *a.s. and that supp* $(\eta) = [c, +\infty)$. *Then, supp* $(\xi) = [c, +\infty)$, *its cdf is absolutely continuous on* $(c, +\infty)$ *and it may have a jump discontinuity of size at most* $\mathbb{P}(\eta = c)$ *at* c.

Proof. It follows simply by Theorem 2 by noticing that the relation between ξ, η implies that supp (ξ) is the closure of $(c, +\infty)$ and also that $\mathbb{P}(\xi = c) \leq \mathbb{P}(\eta = c)$. □

The latter corollary is to our view the most useful result for the establishment of the limit theory for tests of stochastic spanning. In such frameworks, it is usually the case that X is Gaussian, that it is derived as a weak limit of processes used in the defiition of the test statistics while ξ can be conveniently defined as a difference between infima of X defined on different regions of Λ with easily derivable properties.

3 Application: A Test for Stochastic Spanning of the Markowitz Type

In this section we introduce the concept of stochastic spanning for the Markowitz dominance rule. We first provide some order theoretic characterization of the concept, and derive an analytical representation using a functional defined by recursive optimizations. We then define a statistical testing procedure using the principle of conditioning based on subsampling, and derive its first order limit theory, among others via the use of the corollary 1.

3.1 Markowitz Stochastic Dominance and Stochastic Spanning

Given $(\Omega, \mathcal{F}, \mathbb{P})$ suppose that F denotes the cdf of some probability measure on $\mathbb{R}^n$ with finite first moment. Let $G(z, \boldsymbol{\lambda}, F)$ be $\int_{\mathbb{R}^n} \mathbb{I}\{\lambda^{Tr}u \leq z\} dF(u)$, i.e. the cdf of the linear transformation $\mathbb{R}^n \ni x \to \lambda^{Tr}x$ where λ assumes its values in $\mathbb{L}$ which is a closed non-empty subset of $\mathbb{S} = \{\boldsymbol{\lambda} \in \mathbb{R}^n_+ : 1^{Tr}\lambda = 1, \}$. Analogously let $\mathbb{K}$ denote some *distinguished* subcollection of $\mathbb{L}$. In the context of financial econometrics, F usually represents the joint distribution of n asset returns, and $\mathbb{S}$ the space of linear portfolios that can be constructed upon the previous, if shortselling is prohibited. The parameter set $\mathbb{L}$ represents the portfolio collection at hand, consisted for example by, in some particular sense, economically feasible portfolios. We will denote generic elements of $\mathbb{L}$ by λ, κ etc. In order to define the concepts of Markowitz stochastic dominance and subsequently of spanning consider

$$\mathcal{J}(z_1, z_2, \lambda; F) := \int_{z_1}^{z_2} G(u, \lambda, F)\, du. \tag{1}$$

Notice that the existence of the mean of the underlying distribution implies that we can allow the limits of integration above to assume extended values, hence we can obtain the following definition.

Definition 1. κ *weakly Markowitz-dominates* λ, denoted by $\kappa \succcurlyeq_M \lambda$, iff

$$\Delta_1(z, \lambda, \kappa, F) \doteqdot \mathcal{J}(-\infty, z, \kappa, F) - \mathcal{J}(-\infty, z, \lambda, F) \leq 0 \; \forall z \in \mathbb{R}_-$$

and

$$\Delta_2(z, \lambda, \kappa, F) \doteqdot \mathcal{J}(z, +\infty, \kappa, F) - \mathcal{J}(z, +\infty, \lambda, F) \leq 0 \; \forall z \in \mathbb{R}_{++}.$$

Levy and Levy [7] show that $\kappa \succcurlyeq_M \lambda$ iff the expected utility of κ is greater than or equal to the expected utility of λ *for any* utility function in the set of increasing and, concave on the negative part and convex on the positive part real functions (termed as reverse S-shaped (at zero) utility functions).

It is easy to see that $\succcurlyeq_M$ is a preorder on $\mathbb{L}$ since it does not generally satisfy antisymmetry due to the fact that first $G(u, \kappa, F) = G(u, \lambda, F)$ does not imply that $\kappa = \lambda$ and second, even if the inequalities appearing in the previous definition are satisfied as equalities the relation $G(\cdot, \kappa, F) = G(\cdot, \lambda, F)$ is not guaranteed. If the inequalities above are satisfied as equalities then the pair (κ, λ) belongs to the generally non-trivial equivalence part of the preorder. Strict dominance $\kappa \succ_M \lambda$ is the irreflexive part of the preorder and it holds iff at least one of the previous inequalities holds strictly for some $z \in \mathbb{R}$. Finally notice that since it is possible for some $z \in \mathbb{R}$ at least one of the inequalities defining the preorder to change orientation, the relation is not generally total. When this is the case κ and λ are incomparable w.r.t. $\succcurlyeq_M$. The following definition and the subsequent lemma clarify the concept of spanning and part of its structure.

Definition 2. $\mathbb{K}$ Markowitz-spans $\mathbb{L}$ (say $\mathbb{K} \succcurlyeq_M \mathbb{L}$with abuse of notation) iff for any $\lambda \in \mathbb{L}$, $\exists \kappa \in \mathbb{K} : \kappa \succcurlyeq_M \lambda$. If $\mathbb{K} = \{\kappa\}$ then κ is termed as Markowitz super-efficient.

It is easy to see that if the set of maximal elements of the preorder is non-empty, i.e. the efficient set $\mathcal{E}_M$ of the preorder, then $\mathbb{K} \supseteq \mathcal{E}_M$ implies that $\mathbb{K} \succcurlyeq_M \mathbb{L}$. Since $\mathbb{L} \succcurlyeq_M \mathbb{L}$ the existence of a spanning set needs not the non-emptyness of the efficient set. If $\mathbb{K} \succcurlyeq_M \mathbb{L}$ then the optimal choice of *every* agent with preferences represented by a reverse S-shaped utility function lies nessecarily inside $\mathbb{K}$. Hence if $\mathbb{K} \subset \mathbb{L}$ and spanning occurs, then the problem of optimal choice within $\mathbb{L}$ can be reduced to the analogous problem within $\mathbb{K}$, and the latter could be less complex than the former. Therefore the interest in the verification of spanning can be motivated by reasons of tractability to the problem of optimal choice in such frameworks. Furthermore, in cases where $\mathcal{E}_M$ is non-empty, any spanning set can be perceived as an outer approximation of the efficient set. Hence the notion becomes relevant to the problem of the examination of the properties of the efficient set, which in most cases is also complex. The above, naturally raise the following question. Given $\mathbb{K}$ a non empty subset of $\mathbb{L}$, is $\mathbb{K} \succcurlyeq_M \mathbb{L}$? The following lemma provides with an analytical characterization by means of nested optimizations.

Lemma 1. *Suppose that $\mathbb{K}$ is closed and there exists a $z^\star \in \mathbb{R}$ which is a continuity point for $G(z, \kappa, F)$ for any $\kappa \in \mathbb{K}$ and that $\sup_{\kappa\in\mathbb{K}} \int_{A(z^\star)} G(u, \kappa, F)\, du < +\infty$, where $A(z^\star) = \begin{cases} (-\infty, z^\star], & z \in \mathbb{R}_- \\ [z^\star, +\infty), & z \in \mathbb{R}_{++} \end{cases}$. Then $\mathbb{K} \succcurlyeq_M \mathbb{L}$ iff*

$$\xi(F) \doteqdot \sup_{\lambda\in\mathbb{L}} \inf_{\kappa\in\mathbb{K}} \max_{i=1,2} \sup_{z\in A_i} \Delta_i(z, \lambda, \kappa, F) = 0, \tag{2}$$

where $A_1 = \mathbb{R}_-$, $A_2 = \mathbb{R}_{++}$.

Proof. ($\Leftarrow$) If $\mathbb{K} \succcurlyeq_M \mathbb{L}$ then for any λ there exists some κ such that $\sup_{z\leq 0} \Delta_1(z, \lambda, \kappa, F) \leq 0$ and $\sup_{z>0} \Delta_2(z, \lambda, \kappa, F) \leq 0$ which implies that

$$\inf_{\kappa\in\mathbb{K}} \max_{i=1,2} \sup_{z\in A_i} \Delta_i(z, \lambda, \kappa, F) \leq 0. \tag{3}$$

Suppose without loss of generality that $z^\star \in \mathbb{R}_-$. Its existence implies that $G(z^\star, \kappa, F)$ is continuous w.r.t. $\kappa \in \mathbb{K}$. The condition $\sup_{\kappa\in\mathbb{K}} \int_{-\infty}^{z^\star} G(u, \kappa, F)\, du < +\infty$ along with the Dominated Convergence Theorem imply that $\mathcal{J}(-\infty, z^\star, \kappa, F)$ is continuous w.r.t. κ. This along with the compactness of $\mathbb{K}$ imply that $\arg\min_{\kappa\in\mathbb{K}} \mathcal{J}(-\infty, z^\star, \kappa, F)$ is non empty. Let $\kappa^\star$ be an element of the latter. Then, the first equality follows from that

$$\xi(F) \geq \inf_{\kappa\in\mathbb{K}} \mathcal{J}(-\infty, z, \kappa, F) - \mathcal{J}(-\infty, z, \kappa^\star, F) = 0.$$

($\Rightarrow$) Suppose now that $\xi(F) = 0$ and consider an arbitrary λ. This implies that 3 holds and thereby there exists some element of $\mathbb{K}$ for which $\Delta_i(z, \lambda, \kappa, F) \leq 0$ for every $z \in A_i$, $i = 1, 2$. □

Notice that the existence of a common point of continuity for the underlying set of cdf's will be implied if the support of F has no atomes. Furthermore, the case of super-efficiency is trivially implied by the previous result via the commutativity of $\max - \sup$.

Corollary 2. *Under the scope of the previous lemma, κ is Markowitz super-efficient iff*

$$\max_{i=1,2} \sup_{\lambda \in \mathbb{L}} \sup_{z \in A_i} \Delta_i \left(z, \boldsymbol{\lambda}, \kappa, F\right) = 0.$$

The previous lemma cannot be directly used if F is unknown and/or the optimizations involved are infeasible as is usually the case. However in conjunction with the principle of analogy it provides the backbone for the construction of statistical inferential procedures for the question above.

3.2 An Asymptotically Exact and Consistent Statistical Test

We employ Lemma 1 in order to construct a statistical test for the question above. If $\mathbb{K} \succcurlyeq_M \mathbb{L}$ is chosen as the null hypothesis, then in the framework of the aforementioned result, the hypothesis structure takes the following form:[4]

$$\mathbf{H_0} : \xi\left(F\right) = 0,$$
$$\mathbf{H_a} : \xi\left(F\right) > 0.$$

In order to proceed with the development of the decision process we extend our framework as follows. Consider a process $\left(Y_t\right)_{t \in \mathbb{Z}}$ taking values in $\mathbb{R}^n$. Y_{t_i} denotes the i^{th} element of Y_t. The sample is the random element $\left(Y_t\right)_{t=1,\ldots,T}$. In a financial framework it usually represents returns of n financial basis assets upon which portfolios can be constructed via convex combinations. F is the cdf of Y_0 and $\widehat{F}_T$ the *empirical* cdf associated with the random element $\left(Y_t\right)_{t=1,\ldots,T}$. Given the previous and using the principle of analogy we consider the following random variable that will assume the role of the test statistic

$$\xi_T \doteqdot \xi\left(\sqrt{T}F_T\right) = \sup_{\lambda \in \mathbb{L}} \inf_{\kappa \in \mathbb{K}} \max_{i=1,2} \sup_{z \in A_i} \Delta_i \left(z, \boldsymbol{\lambda}, \kappa, \sqrt{T}F_T\right),$$

which is obviously the empirical analog of $\xi\left(F\right)$. Notice again that when $\mathbb{K}$ is a singleton then the test statistic coincides with the one used in Arvanitis and Topaloglou [2]. Now, the following assumption enables the derivation of the limit distribution of ξ_T under $\mathbf{H_0}$.[5]

Assumption 2. $\left(Y_t\right)_{t \in \mathbb{Z}}$ *is strictly stationary and a-mixing with mixing coefficients a_t such that $a_T = O(T^{-a})$ for some $a > 1 + \frac{2}{\delta}$. F is absolutely continuous with convex support and such that for some $\delta > 0$,*

$$\mathbb{E}\left\|Y_0\right\|^{2+\delta} < +\infty.$$

Furthermore,

$$\mathbb{V} = \mathbb{E}\left[\left(Y_0 - \mathbb{E}Y_0\right)\left(Y_0 - \mathbb{E}Y_0\right)^T\right] + 2\sum_{t=1}^{\infty} \mathbb{E}\left[\left(Y_0 - \mathbb{E}Y_0\right)\left(Y_t - \mathbb{E}Y_t\right)^T\right]$$

is positive definite.

[4] Notice that corollary 2 implies that the hypotheses are, in the special case of superefficiency, as in Arvanitis and Topaloglou [2].

[5] $\left(x\right)_+ = \max\left\{x, 0\right\}$ and $\left(x\right)_- = \min\left\{x, 0\right\}$ and when x is a vector they are to be interpreted in the coordinatewise sense.

Remark 3. *The mixing part of the previous assumption is readily implied by concepts such as geometric ergodicity which holds for many stationary models used in the context of financial econometrics under parameter restrictions and restrictions on the properties of the innovation processes involved. Prominent examples are the strictly stationary versions of (possibly multivariate) ARMA or several GARCH and stochastic volatility type of models (see for example Francq and Zacoian [4]). Counter-examples are stationary models that exhibit long memory, etc. The moment condition enables the validity of a mixing CLT. It is readily established in models such as the ones mentioned above usually in the form of stricter restrictions on the properties of building blocks and the parameters of the processes involved. The condition on the definition of $\mathbb{V}$ can be easily established via parameter restrictions on models such as the aforementioned. Notice that due to the compactness of Λ, the previous imply that*

$$\sup_{\lambda\in\Lambda}\int_{-\infty}^{+\infty}\sqrt{G\left(u,\lambda,F\right)\left(1-G\left(u,\lambda,F\right)\right)}du<+\infty,$$

which is a uniform version of the analogous condition used in Horvath, Kokoszka, and Zitikis [5]. It also implies that $\sup_{\kappa\in\mathbb{K}}\int_{-\infty}^{z^{\star}}G\left(u,\kappa,F\right)du<+\infty$, hence it conforms with the relevant restriction appearing in Lemma 1.

The derivation of the limit theory of ξ_T under the null, is crucially based on partitions of the "parameter sets" $\mathbb{K}$ and $\mathbb{L}$ and the approximation of the test statistic by auxiliary random elements in the context of this hypothesis. Given those, the limit theory can be readily based on standard results and the application of the Continuous Mapping Theorem. We describe the partitioning and then present the result. Consider

$$\mathbb{L}^{=}=\left\{\lambda\in\mathbb{L}:\inf_{\kappa\in\mathbb{K}}\max_{i=1,2}\sup_{z\in A_i}\Delta_i\left(z,\boldsymbol{\lambda},\kappa,F\right)=0\right\},$$

and notice that under $\mathbf{H_0}$ it is non-empty. Furthermore for an arbitrary λ consider

$$\mathbb{K}^{\leqslant}\left(\lambda\right)=\left\{\kappa\in K:\max_{i=1,2}\sup_{z\in A_i}\Delta_i\left(z,\boldsymbol{\lambda},\kappa,F\right)\leq 0\right\}.$$

Again for any λ and under $\mathbf{H_0}$, $\mathbb{K}^{\leqslant}\left(\lambda\right)$ is also non empty. In what follows $\rightsquigarrow$ denotes convergence in distribution.

Proposition 1. *Suppose that $\mathbb{K}$ is closed, Assumption 2 holds and that $\mathbf{H_0}$ is true. Then as $T\to\infty$*

$$\xi_T\rightsquigarrow\xi_\infty,$$

where

$$\xi_\infty\doteqdot\sup_{\lambda\in\mathbb{L}^{=}}\inf_{\kappa\in\mathbb{K}^{\leq}(\lambda)}\max_{i=1,2}\sup_{z\in A_i}\Delta_i\left(z,\boldsymbol{\lambda},\kappa,\mathcal{G}_F\right),$$

and $\mathcal{G}_F$ is a centered Gaussian process with covariance kernel given by $\mathrm{Cov}(\mathcal{G}_F(x),\mathcal{G}_F(y))=\sum_{t\in\mathbb{Z}}\mathrm{Cov}\left(\mathbb{I}_{X_0\leq x},\mathbb{I}_{X_t\leq y}\right)$ and $\mathbb{P}$ almost surely uniformly continuous sample paths defined on $\mathbb{R}^n$.[6]

[6] see Theorem 7.3 of Rio [13].

Proof. The result follows directly from Lemmata 3, 4 and 5 in the Appendix. □

In the case of superefficiency we obtain the limit distribution of ξ_T directly from above in the following corollary. This is an improvement of the relevant results in Arvanitis and Topaloglou [2] who only derive an upper bound that leads to an asymptotically conservative test based on block-bootstrap. Hence our present results can also be used to provide with an asymptotic improvement of the Arvanitis and Topaloglou[2] testing procedure.

Corollary 3. *In the case of super-efficiency*

$$\xi_\infty \doteqdot \max_{i=1,2} \sup_{\lambda\in\mathbb{L}^=} \sup_{z\in A_i} \Delta_i\left(z, \lambda, \kappa, \mathcal{G}_F\right).$$

Notice that one cannot directly use the results of the previous lemma, in order to construct an asymptotic decision procedure since ξ_∞ depends on the generally unknown F. However, a feasible decision rule can be established by the use of some resampling procedure. We consider resampling using the method of subsampling, exactly as in Linton, Post and Wang [9].

Algorithm. *The testing procedure consists of the following steps:*

1. *Evaluate* ξ_T at the original sample value.
2. *For* $0 < b_T \leq T$ *generate subsamples from the original observations* $(Y_i)_{i=t,\ldots t+b_T-1}$ *for all* $t = 1, 2, \ldots, T - b_T + 1$.
3. *Evaluate the test statistic on each subsample thereby obtaining* $\xi_{T,b_T,t}$ *for all* $t = 1, 2, \ldots, T - b_T + 1$.
4. *Approximate the cdf of the asymptotic distribution under the null of* ξ_T by $s_{T,b}(y) = \frac{1}{T-b_T+1}\sum_{t=1}^{T-b_T+1} 1\left(\xi_{T,b_T,t} \leq y\right)$ and calculate
$$q_{T,b_T}\left(1-\alpha\right) = \inf_y \left\{s_{T,b}(y) \geq 1-\alpha\right\}.$$
5. *Reject* $\mathbf{H_0}$ *iff* $\xi_T > q_{T,b_T}(1-\alpha)$.

We derive asymptotic exactness and consistency for this testing procedure by utilizing Theorem 3.5.1.i of Politis et al. [11]. In order to do so we first use the following standard assumption that restricts the asymptotic behaviour of b_T.

Assumption 3. *Suppose that* (b_T)*, possibly depending on* $(Y_t)_{t=1,\ldots,T}$, satisfies

$$\mathbb{P}\left(l_T \leq b_T \leq u_T\right) \to 1$$

where (l_T) and (u_T) are real sequences such that $1 \leq l_T \leq u_T$ for all T, $l_T \to \infty$ and $\frac{u_T}{T} \to 0$ as $T \to \infty$.

The Politis et al. [11] 3.5.1.i Theorem also requires continuity of the limit cdf at the relevant quantile. In order to achieve this we use the following assumption that restricts the form of $\mathbb{L}$ as a simplicial complex.

Assumption 4. *$\mathbb{L}$ is a simplicial complex comprised of a finite number of subsimplices of $\mathbb{S} = \{\lambda \in \mathbb{R}^n_+ : \mathbf{e}'\lambda = 1,\}$. It contains all the extreme points of $\mathbb{S}$. $\mathbb{K}$ contains $0 \leq m < n$ extreme points of $\mathbb{S}$.*

The assumption generalizes the "parameter space" structure compared to Arvanitis et al. [1], where $\mathbb{K}$ is simply a strict subcomplex of $\mathbb{L}$. The first part of the assumption is trivially satisfied when $\mathbb{L} = \mathbb{S}$ which is usually the case in most applications. Furthermore, mainly due to reasons of computational facilitation, $\mathbb{K}$ is usually also a subcomplex of $\mathbb{L}$ something that implies that $m < n$ for in the opposite case we would have that $\mathbb{K} = \mathbb{L}$, whence the null hypothesis would be trivially satisfied. As we show in the auxiliary results in the appendix, proposition 1 implies parts 1 and 2 of assumption 1. The previous assumption along with proposition 1 imply that the third part of the particular assumption is satisfied with $T = \{0\}$, and $\mathbb{P}(\Omega_0) \leq \frac{m}{n}$. Theorem 2 then implies that the cdf of the null limit distribution is (absolutely) continuous at the quantile evaluated on $1-\alpha$ when $1-\alpha > \frac{m}{n}$. *This is essentially the part of the asymptotic exactness derivation for which the main theorem 2 (actually Corollary 1) becomes useful.* Given this we obtain the folowing result that establishes the required limit theory.

Theorem 4. *Suppose that $\mathbb{K}$ is closed, Assumptions 2, 3 and 4 hold and that $\alpha < \frac{n-m}{n}$. For the testing procedure described in 3.2 we have that*

1. *If $\mathbf{H_0}$ is true then*
$$\lim_{T\to\infty} \mathbb{P}\left(\xi_T > q_{T,b_T}(1-\alpha)\right) = \alpha.$$
2. *If $\mathbf{H_a}$ is true then*
$$\lim_{T\to\infty} \mathbb{P}\left(\xi_T > q_{T,b_T}(1-\alpha)\right) = 1.$$

Proof. The first result follows by a direct application of Theorem 3.5.1.i of Politis et al. [11] enabled by the results of Lemma 6. For the second result notice that if $\mathbf{H_a}$ is true then the set

$$\mathbb{L}^{>} = \left\{\lambda \in \mathbb{L} : \inf_{\kappa\in\mathbb{K}} \max_{i=1,2} \sup_{z\in A_i} \Delta_i(z, \lambda, \kappa, F) > 0\right\}$$

is non empty. Then we have that

$$\xi_T \geq \xi_T^\star \doteqdot \sqrt{T} \sup_{\lambda\in\mathbb{L}^{>}} \inf_{\kappa\in\mathbb{K}} \max_{i=1,2} \sup_{z\in A_i} \Delta_i(z, \lambda, \kappa, F_T),$$

and due to Lemma 2 we have that

$$\sup_{\lambda\in\mathbb{L}^{>}} \inf_{\kappa\in\mathbb{K}} \max_{i=1,2} \sup_{z\in A_i} \Delta_i(z, \lambda, \kappa, F_T) \rightsquigarrow \sup_{\lambda\in\mathbb{L}^{>}} \inf_{\kappa\in\mathbb{K}} \max_{i=1,2} \sup_{z\in A_i} \Delta_i(z, \lambda, \kappa, F)$$

which implies that under $\mathbf{H_a}$, $\xi_T^\star$ diverges to $+\infty$ in probability. The result follows. □

The restriction on the significance level is in usual applications non-binding. For example when $\mathbb{K}$ is a singleton, i.e. when the test is applied for super-efficiency, then it implies that $\alpha < 1/2$ something that is most usually satisfied. Notice that the restriction becomes closer to binding the more extreme points exist inside $\mathbb{K}$. An example of an extreme case, is when n is large, $\mathbb{K}$ is finite and contains $n-1$ extreme points. In such cases the result leads to subsampling tests that tend to asymptotically favor the spanning null. Such cases could be *partially* handled presently, by breaking up $\mathbb{K}$ is ”smaller pieces” and iterating the testing procedure w.r.t. them. For example the procedure could be applied for any subset of $\mathbb{K}$ that contains m points, for m sufficiently small in order to obtain a meaningful significance level. If for some subset, spanning cannot be rejected, then it can be infered that spanning cannot be rejected for the initial $\mathbb{K}$, since supersets of spanning sets are due to Definition 2 spanning sets.

4 Conclusions

We have derived properties of the cdf of a random variable defined by recursive optimizations applied on a continuous stochastic process w.r.t. possibly dependent parameter spaces. Those properties extend previous results and can be useful for the derivation of the limit theory of tests for stochastic spanning w.r.t. preorders defined by stochastic dominance rules. As an illustration we have defined the concept of spanning, constructed an analogous test based on subsampling, and derived the first order limit theory for the case of the Markowitz stochastic dominance.

The scope of the present paper does not contain the issue of the numerical implementation of the test. The optimizations involved on the computation of the statistic as well as the critical values are not trivial. Given the representation of the Markowitz dominance by the set of the reverse S-shaped utilities, we conjecture that there exists a representation of the test statistic involving this class of utilities that is similar to the representation by utilities of the spanning test for the second order stochastic dominance appearing in Proposition 3.2 of Arvanitis, Hallam and Post [1]. If this is true it is possible that a feasible numerical algorithm can be designed via the use of empirical supports and piecewise linear approximations of the aforementioned utilities defined by a finite number of parameters in the spirit of Section 7 of the aforementioned paper.

The preorder used is simply illustrative. Analogous results can be derived for other forms of stochastic dominance rules, such as first or third order, or Prospect stochastic dominance. We leave issues such as the derivation of such results, and/or the numerical implementation of testing procedures such as the above for future research.

References

[1] Arvanitis, Stelios and Hallam, Mark S and Post, Thierry and Topaloglou, Nikolas (October 4, 2015). Stochastic Spanning. Available at SSRN: http://ssrn.com/abstract=2575165 or http://dx.doi.org/10.2139/ssrn.2575165 2, 3, 12, 13

[2] Arvanitis, S., & Topaloglou, N. (2013). Testing For Prospect And Markowitz Stochastic Dominance Efficiency. Dept. of Economics Working Paper Series, http://www.econ.aueb.gr/en/uploadfiles/Alldp182013.pdf. 2, 9, 11, 15, 18

[3] Davidson, R. and Duclos, J.-Y. (2000). Statistical Inference for Stochastic Dominance and for the Measurement of Poverty and Inequality. *Econometrica* 68, 1435-1464. 18

[4] Francq, C., & Zakoian, J. M. (2011). *GARCH models: structure, statistical inference and financial applications*. John Wiley & Sons. 10

[5] Horvath, L. and Kokoszka, P. and R. Zitikis (2006). Testing for Stochastic Dominance Using the Weighted McFadden-type Statistic. *Journal of Econometrics* 133, 191-205. 10

[6] Huberman, G. and S. Kandel, (1987). Mean-Variance Spanning. *Journal of Finance* 42, 873-888. 2

[7] M. Levy and Levy, H. (2002). Prospect Theory: Much Ado about Nothing?, *Management Science* 48, 1334-1349. 2, 7

[8] Lifshits, M. A. (1983). On the absolute continuity of distributions of functionals of random processes. *Theory of Probability & Its Applications*, 27(3), 600-607. 4, 6

[9] Linton, O., Post, T., & Whang, Y. J. (2014). Testing for the stochastic dominance efficiency of a given portfolio. *The Econometrics Journal*, 17(2), S59-S74. 11

[10] Nualart, D. (2006). *The Malliavin calculus and related topics*. Berlin: Springer. 3, 4, 5, 6, 17

[11] Politis, D. N., J. P. Romano and M. Wolf (1999). *Subsampling*. Springer New York. 11, 12

[12] Post, Thierry, and P. Versijp, 2007, Multivariate tests for stochastic dominance efficiency of a given portfolio, *Journal of Financial and Quantitative Analysis*, 42, 489-516. 2

[13] Rio, E. (2000), Théorie Asymptotique des Processus Aléatoires Faiblement Dépendants. *Mathématiques et Applications*, Vol. 31, Berlin: Springer- Verlag. 10

[14] Scaillet, O., & Topaloglou, N. (2010). Testing for stochastic dominance efficiency. *Journal of Business & Economic Statistics*, 28(1), 169-180. ⍰ 2

[15] Van Der Vaart, Aad W (1997). *Asymptotic Statistics*. CUP.

[16] Van Der Vaart, A. W., & Wellner, J. A. (1996). *Weak Convergence*. Springer New York. 4, 15, 17

[17] Sidak, Zbynek, Pranab K. Sen, and Jaroslav Hajek (1999). *Theory of rank tests*. Academic Press. 18

Appendix-Auxiliary Lemmata

The following are auxiliary lemmata used for the derivation of the proofs of Proposition 1 and Theorem 4.

Lemma 2. *Under Assumption 2*

$$\begin{pmatrix} \Delta_1\left(z,\lambda,\kappa,\sqrt{T}\left(F_T-F\right)\right) \\ \Delta_2\left(z,\lambda,\kappa,\sqrt{T}\left(F_T-F\right)\right) \end{pmatrix} \rightsquigarrow \begin{pmatrix} \Delta_1\left(z,\lambda,\kappa,\mathcal{G}_F\right) \\ \Delta_2\left(z,\lambda,\kappa,\mathcal{G}_F\right) \end{pmatrix}$$

as random elements defined on the space of bounded functions on $\mathbb{L}\times\mathbb{K}\times\mathbb{R}$ *equiped with the* $\sup$*-norm.*

Proof. Given the compactness of $\mathbb{K}\times\mathbb{L}$, the proof follows by a trivial extension of the proof of Lemma AL.2 of Arvanitis and Topaloglou [2]. □

Lemma 3. *Under Assumption 2*

$$\sup_{\mathbb{L}^=}\inf_{\mathbb{K}}\max_{i=1,2}\sup_{z\in A_i}\Delta_i\left(z,\lambda,\kappa,\sqrt{T}F_T\right)\rightsquigarrow\sup_{\mathbb{L}^=}\inf_{\mathbb{K}^{\leq}(\lambda)}\max_{i=1,2}\sup_{z\in A_i}\Delta_i\left(z,\lambda,\kappa,\mathcal{G}_F\right).$$

Proof. Consider $\mathbb{K}^{>}(\lambda)=\left\{\kappa\in\mathbb{K}:\max_{i=1,2}\sup_{z\in A_i}\Delta_i\left(z,\lambda,\kappa,F\right)>0\right\}$. For any $\lambda\in\mathbb{L}^=$ for which $\mathbb{K}^{>}(\lambda)\neq\emptyset$

$$\inf_{\mathbb{K}}\max_{i=1,2}\sup_{z\in A_i}\Delta_i\left(z,\lambda,\kappa,\sqrt{T}F_T\right)$$

$$=\min\left\{\inf_{\mathbb{K}^{\leq}(\lambda)}\max_{i=1,2}\sup_{z\in A_i}\Delta_i\left(z,\lambda,\kappa,\sqrt{T}\left(F_T-F\right)\right),R_T(\lambda)\right\}$$

where $R_T(\lambda)=\max_{i=1,2}\sup_{z\in A_i}\Delta_i\left(z,\lambda,\kappa,\sqrt{T}\left(F_T-F\right)\right)$ for any element of the underlying probability space for which $\inf_{\mathbb{K}^{>}(\lambda)}\max_{i=1,2}\sup_{z\in A_i}\Delta_i\left(z,\lambda,\kappa,\sqrt{T}F_T\right)$ is achieved on the boundary of $\mathbb{K}^{\leq}(\lambda)$ or $R_T(\lambda)=\max_{i=1,2}\sup_{z\in A_i}\Delta_i\left(z,\lambda,\kappa_T,\sqrt{T}F_T\right)$ for κ_T measurable and inside $\mathbb{K}^{>}(\lambda)$. Using a Skorokhod representation argument (see inter alia Theorem 1.10.4 of van der Vaart and Wellner [16]), for any such element of the underlying (or potentially enlarged) probability space, the sequence $(R_T(\lambda))$ can be partitioned to subsequences which (if any) nessecarily diverge to $+\infty$, and to subsequences which (if any) converge to the limit of

$$\max_{i=1,2}\sup_{z\in A_i}\Delta_i\left(z,\lambda,\kappa,\sqrt{T}\left(F_T-F\right)\right)$$

evaluated on the boundary of $\mathbb{K}^{\leq}(\lambda)$. In any case the minimum above weakly converges to

$$\inf_{\mathbb{K}^{\leq}(\lambda)}\max_{i=1,2}\sup_{z\in A_i}\Delta_i\left(z,\lambda,\kappa,\mathcal{G}_F\right)$$

and then the CMT and Lemma 2 establish the required result. □

Lemma 4. *For $\epsilon_T \to 0$, $\sqrt{T}\epsilon_T \to +\infty$, as $T \to \infty$, consider the set*

$$\mathbb{L}^{<}_{\epsilon_T} = \left\{ \lambda : \inf_{\mathbb{K}} \max_{i=1,2} \sup_{z \in A_i} \Delta_i \left(z, \lambda, \kappa, F \right) \leq -\epsilon_T \right\}$$

and define the random variable

$$\xi^{\epsilon}_T = \sup_{\mathbb{L}^{=} \cup \mathbb{L}^{<}_{\epsilon_T}} \inf_{\mathbb{K}} \max_{i=1,2} \sup_{z \in A_i} \Delta_i \left(z, \lambda, \kappa, \sqrt{T} F_T \right).$$

Then under Assumption 1 and if the null hypothesis is true, for any $\varepsilon > 0$

$$\lim_{T \to \infty} \mathbb{P} \left(\xi_T - \xi^{\epsilon}_T > \varepsilon \right) = 0.$$

Proof. For an arbitrary λ and $0 < \delta < 1$ consider the subset of $\mathbb{K}^{\leq} (\lambda)$

$$\mathbb{K}^{\leq} (\lambda)^{\delta \epsilon_T} = \left\{ \kappa \in K : \max_{i=1,2} \sup_{z \in A_i} \Delta_i \left(z, \lambda, \kappa, F \right) \leq -\delta \epsilon_T \right\}.$$

Furthermore define

$$\mathbb{L}^{<} = \left\{ \lambda \in \mathbb{L} : \inf_{\mathbb{K}} \max_{i=1,2} \sup_{z \in A_i} \Delta_i \left(z, \lambda, \kappa, F \right) < 0 \right\}$$

and notice that under the null, $\mathbb{L} = \mathbb{L}^{=} \cup \mathbb{L}^{<}$. Pick $\varepsilon > 0$ as before and under the null hypothesis consider

$$\begin{aligned} \mathbb{P} \left(\xi_T - \xi^{\epsilon}_T > \varepsilon \right) &= \mathbb{P} \left(\max \left\{ \xi^{\epsilon}_T, \sup_{\mathbb{L}^{<} - \mathbb{L}^{<}_{\epsilon_T}} \inf_{\mathbb{K}} \max_{i=1,2} \sup_{z \in A_i} \Delta_i \left(z, \lambda, \kappa, \sqrt{T} F_T \right) \right\} - \xi^{\epsilon}_T > \varepsilon \right) \\ &\leq \mathbb{P} \left(\xi^{\epsilon}_T < \sup_{\mathbb{L}^{<} - \mathbb{L}^{<}_{\epsilon_T}} \inf_{\mathbb{K}} \max_{i=1,2} \sup_{z \in A_i} \Delta_i \left(z, \lambda, \kappa, \sqrt{T} F_T \right) \right) \\ &\leq \mathbb{P} \left(\sup_{\mathbb{L}^{=}} \inf_{\mathbb{K}} \max_{i=1,2} \sup_{z \in A_i} \Delta_i \left(z, \lambda, \kappa, \sqrt{T} F_T \right) < \sup_{\mathbb{L}^{<}} \sup_{\mathbb{K}^{\leq}(\lambda)^{\delta \epsilon_{\mathrm{T}}}} \max_{i=1,2} \sup_{z \in A_i} \Delta_i \left(z, \lambda, \kappa, \sqrt{T} F_T \right) \right). \end{aligned} \tag{4}$$

From Lemmas 2, 3 the lhs of the inequality inside the previous probability weakly converges to $\sup_{\mathbb{L}^{=}} \inf_{\mathbb{K}^{\leq}(\lambda)} \max_{i=1,2} \sup_{z \in A_i} \Delta_i \left(z, \lambda, \kappa, \mathcal{B} \circ F \right)$. For the rhs we obtain

$$\begin{aligned} \sup_{\mathbb{L}^{<}} \sup_{\mathbb{K}^{\leq}(\lambda)^{\delta \epsilon_{\mathrm{T}}}} \max_{i=1,2} \sup_{z \in A_i} \Delta_i \left(z, \lambda, \kappa, \sqrt{T} F_T \right) &\leq \sup_{\mathbb{L}^{<}} \sup_{\mathbb{K}^{\leq}(\lambda)^{\delta \epsilon_{\mathrm{T}}}} \max_{i=1,2} \sup_{z \in A_i} \Delta_i \left(z, \lambda, \kappa, \sqrt{T} \left(F_T - F \right) \right) \\ &+ \sup_{\mathbb{L}^{<}} \sup_{\mathbb{K}^{\leq}(\lambda)^{\delta \epsilon_{\mathrm{T}}}} \max_{i=1,2} \sup_{z \in A_i} \Delta_i \left(z, \lambda, \kappa, \sqrt{T} F \right). \end{aligned}$$

Due to Lemma 2 and the CMT the first term on the rhs of the last display weakly converges to

$$\sup_{\mathbb{L}^{<}} \sup_{\mathbb{K}^{\leq}(\lambda)} \max_{i=1,2} \sup_{z \in A_i} \Delta_i \left(z, \lambda, \kappa, \mathcal{G}_F \right)$$

and the second diverges to $-\infty$ due to the construction of $\mathbb{K}^{\leq} (\lambda)^{\delta \epsilon_T}$ and the properties of ϵ_T. Hence the probability in 4 converges to zero. □

Lemma 5. *Under Assumption 2 and for ϵ_T, $\mathbb{L}^{<}_{\epsilon_T}$, ξ^{ϵ}_T as in Lemma 4*

$$\left|\mathbb{P}\left(\xi^{\epsilon}_T \leq y\right) - \mathbb{P}\left(\sup_{\mathbb{L}^{=}} \inf_{\kappa\in\mathbb{K}} \max_{i=1,2} \sup_{z\in A_i} \Delta_i\left(z,\lambda,\kappa,\sqrt{T}F_T\right) \leq y\right)\right| \to 0.$$

Proof. For any $y \in \mathbb{R}$ we have that

$$\limsup_{T\to\infty}\left|\mathbb{P}\left(\xi^{\epsilon}_T \leq y\right) - \mathbb{P}\left(\sup_{\mathbb{L}^{=}} \inf_{\kappa\in\mathbb{K}} \max_{i=1,2} \sup_{z\in A_i} \Delta_i\left(z,\lambda,\kappa,\sqrt{T}F_T\right) \leq y\right)\right|$$

$$\leq \limsup_{T\to\infty} \mathbb{P}\left(\sup_{\mathbb{L}^{<}_{\epsilon_T}} \inf_{\kappa\in\mathbb{K}} \max_{i=1,2} \sup_{z\in A_i} \sup_{z\in A_i} \Delta_i\left(z,\lambda,\kappa,\sqrt{T}F_T\right) > y\right),$$

due to the elementary inequality $|\mathbb{P}\left(\max(Y,Z) \leq y\right) - \mathbb{P}\left(Y \leq y\right)| \leq \mathbb{P}\left(Z > y\right)$. Now, the previous limsup is less than or equal

$$\limsup_{T\to\infty} \mathbb{P}\left(\sup_{\mathbb{L}^{<}_{\epsilon_T}} \max_{i=1,2} \sup_{z\in A_i} \Delta_i\left(z,\lambda,\kappa,\sqrt{T}\left(F_T - F\right)\right) > y + \sqrt{T}\epsilon_T\right)$$

and the latter is zero due to Lemma 2 and the assumed properties of ϵ_T as defined in the statement of Lemma 4. □

Lemma 6. *Under Assumptions 2 and 4 the distribution of ξ_∞ has support $[0,+\infty)$, its cdf is absolutely continuous on $(0,+\infty)$ and it may have a jump discontinuity of size at most $\frac{m}{n}$ at zero.*

Proof. The result stems from Corollary 1 as long as the requirements of Assumption 1 are satisfied and an appropriately bounding η is found. First notice that

$$\mathbb{E}\sup_{\Lambda}\left\|X_\lambda\right\|^2 = \sup_{\lambda\in\mathbb{L}^{=}} \sup_{\kappa\in\mathbb{K}^{\leq}(\lambda)} \sup_{z\in A_i} \max_{i=1,2} \Delta^2_i\left(z,\lambda,\kappa,\mathcal{G}_F\right)$$

$$\leq \sum_{i=1,2} \mathbb{E}\sup_{\lambda\in\mathbb{L}} \sup_{\kappa\in\mathbb{K}} \sup_{z\in A_i} \Delta^2_i\left(z,\lambda,\kappa,\mathcal{G}_F\right)$$

which is less than or equal to a positive constant times

$$\mathbb{E}\left(\sup_{\lambda\in\mathbb{L},z<0}\int_{-\infty}^{z}\left(z-\lambda^{Tr}u\right)^{+} d\mathcal{G}_F\left(u\right)\right)^2 + \mathbb{E}\left(\sup_{\lambda\in\mathbb{L},z<0}\int_{-\infty}^{z} -\left(z-\lambda^{Tr}u\right)^{+} d\mathcal{G}_F\left(u\right)\right)^2$$

$$+\mathbb{E}\left(\sup_{\lambda\in\mathbb{L},z\geq 0}\int_{z}^{+\infty}\left(z-\lambda^{Tr}u\right)^{+} d\mathcal{G}_F\left(u\right)\right)^2 + \mathbb{E}\left(\sup_{\lambda\in\mathbb{L},z\geq 0}\int_{z}^{+\infty} -\left(z-\lambda^{Tr}u\right)^{+} d\mathcal{G}_F\left(u\right)\right)^2.$$

Due to zero mean Gaussianity of the processes involved, the fact that the packing numbers of $\Lambda \times \mathbb{R}$ are bounded by a polynomial w.r.t. the inverted radii, Lemma A.2.7 of Van Der Vaart and Wellner [16] implies the subexponentiality of the distributions of the suprema above, and thereby the existence of their second moments. Hence hypothesis 1 of Assumption 1 holds. Using the discussion in Nualart [10], immediately after the proof of Proposition 2.1.11 (p. 109) we have that hypothesis 2 of Assumption 1 also holds due to Assumption 2. Notice now that due to the convexity of the sets

$$\{y \in \text{supp}F : \lambda^T y \geq x\}, \ \{y \in \text{supp}F : \lambda^T y < x\}$$

for all $\lambda \in \Lambda$ and x we have that excluding $\mathbb{P}$-negligible events $\Delta_i(z, \lambda, \kappa, \mathcal{G}_F)$ is zero only when $\kappa = \lambda$ and it is at most only then that ξ_∞ has degenerate variance. Thereby $T = \{0\}$ and we will try to obtain a lower bound for ξ_∞. We have that due to Davidson and Duclos [3] Equation (2), Arvanitis and Topaloglou [2] Lemma AL.1 and due to Assumption 2

$$\xi_T \geq \sup_{\lambda \in \mathbb{L}} \inf_{\kappa \in \mathbb{K}} \max_i \Delta_i \left(0, \lambda, \kappa, \sqrt{T} F_T\right)$$

$$\geq \frac{1}{2} \frac{1}{\sqrt{T}} \left(\sup_{\lambda \in \mathbb{L}} \lambda^{Tr} - \sup_{\kappa \in \mathbb{K}} \kappa^{Tr}\right) \sum_{i=1}^{T} (Y_i - \mathbb{E}(Y_0))$$

$$\rightsquigarrow \frac{1}{2} \sup_{\lambda \in \mathbb{L}} \lambda^{Tr} Z - \frac{1}{2} \sup_{\kappa \in \mathbb{K}} \kappa^{Tr} Z$$

where $Z \sim N(0_{n \times 1}, \mathbb{V})$. Hence

$$\xi_\infty \geq \eta \doteqdot \sup_{\lambda \in \mathbb{L}} \lambda^{Tr} Z - \sup_{\kappa \in \mathbb{K}} \kappa^{Tr} Z \geq 0.$$

The previous inequality implies the applicability of Corollary 1 for $c = 0$. We obtain the result by estimating an upper bound for $\mathbb{P}(\eta = 0)$. Due to Assumption 2 and the non-degeneracy of $\mathbb{V}$ the latter probability equals exactly the probability that the minimum of the random vector Z occurs at a coordinate that corresponds to a common extreme point for $\mathbb{L}$ and $\mathbb{K}$. Using Theorem 2 in chapter 3 (p. 37) of Sidak et al. [17] by (in their notation) letting p be the density of the n-variate standard normal distribution it is easy to see that

$$\mathbb{P}(\eta = 0) \leq \frac{m}{n}.$$

$\square$