

DISCRETE APPROXIMATION OF INTEGRAL OPERATORS

CARSTEN FRANZ

(Communicated by Edward C. Waymire)

ABSTRACT. A method to approximate the eigenvalues of linear operators depending on an unknown distribution is introduced and applied to weighted sums of squared normally distributed random variables. This area of application includes the approximation of the asymptotic null distribution of certain degenerated U- and V-statistics.

1. INTRODUCTION

Degenerated U- and V-statistics commonly possess asymptotic null distributions which are a weighted sum of squared normally distributed random variables. The weights are the eigenvalues of a certain integral operator which might (see Baringshaus and the author [1, 2]) but need not (e.g. the Cramér-von Mises-statistic) depend on an underlying, unknown distribution. This article deals with a method to approximate the eigenvalues of operators in the case of dependence on an unknown distribution by substituting it with a sequence of distributions which are known to converge weakly to the unknown distribution, e.g. the empirical distribution of observations made. This problem differs from the usual approximation of linear operators due to the different spaces on which the approximating operators are defined. Koltchinskii and Giné [5] used random matrices to lay a solid foundation of the treatment of this problem.

This article presents another way which can be used to cover some aspects of this problem. By the price of stronger assumptions on the kernel functions the way offers two extensions to the existing solution by Koltchinskii and Giné: first, the probability distributions converging weakly are not required to be empirical measures of iid random variables, and second, the kernel functions of the approximating operators are allowed to vary, even to be dependent on the sequence of converging probability measures, e.g. empirical distributions of observations made.

The approach presented is built on work of Stummel [7, 8, 9], a summary of which is given within the next subsections. The second section deals with the application to statistics; an example which particularly highlights the second extension to existing methods can be found in section 3.

Received by the editors March 9, 2005.

2000 *Mathematics Subject Classification*. Primary 62G20, 45C05.

Key words and phrases. Linear operators, approximation of eigenvalues, asymptotic null distribution of degenerated U-statistics.

©2006 American Mathematical Society
Reverts to public domain 28 years from publication

1.1. Approximation by discrete spaces. The theory of discrete approximation as introduced by Stummel [7, 8, 9] allows the approximating operators to be defined on different normed spaces E_ι with $\iota \in I \subset \mathbb{N}$ being an infinite countable index set and thus delivers the framework for the treatment of the problem of interest.

In order to make statements on the convergence, so-called discrete convergence, of functions $u_\iota \rightarrow u$, where $u_\iota \in E_\iota$, $\iota \in I$, and $u \in E$ is an element of the limiting normed space E , the notion of equivalent sequences is needed:

Definition 1.1. Two sequences $(u_\iota)_{\iota \in I}$, $(v_\iota)_{\iota \in I}$ with $u_\iota, v_\iota \in E_\iota$ are called *equivalent* iff

$$\|u_\iota - v_\iota\|_{E_\iota} \xrightarrow{\iota \rightarrow \infty} 0.$$

The equivalence classes of the equivalence relation on the space of sequences $(u_\iota)_\iota \in \prod_\iota E_\iota$ defined by this notion are connected to the limiting space E with a mapping Q which in turn is needed to possess continuity and linearity properties.

Definition 1.2. A sequence of normed spaces $(E_\iota)_\iota$, $\iota \in I$, is called *discrete approximation* of the space E (with mapping Q) iff

- (i) E, E_ι , $\iota \in I$, are defined above the same field \mathbb{C} , and there exists a mapping Q which associates each element $u \in E$ with an equivalence class $Q(u) \subset \prod_\iota E_\iota$.
- (ii) [**Continuity of the norm**] For all $u \in E$ and every sequence $(u_\iota)_\iota \in Q(u)$ we have

$$\|u_\iota\|_{E_\iota} \xrightarrow{\iota \rightarrow \infty} \|u\|_E.$$

- (iii) [**Linearity**] For all $\alpha, \beta \in \mathbb{C}$, all $u, v \in E$ and all $(u_\iota)_\iota \in Q(u)$, $(v_\iota)_\iota \in Q(v)$, $(w_\iota)_\iota \in Q(\alpha u + \beta v)$,

$$\|\alpha u_\iota + \beta v_\iota - w_\iota\|_{E_\iota} \xrightarrow{\iota \rightarrow \infty} 0.$$

Having defined such a mapping Q , a sequence $(u_\iota)_\iota$ is said to *converge discretely* to $u \in E$ iff $(u_\iota)_\iota \in Q(u)$.

1.2. Discrete compactness and consistency. Let the spaces $(E_\iota)_\iota$ be a discrete approximation of the space E (with corresponding mapping Q). We are now turning our attention to operators defined on these spaces. Approximating operators defined on E_ι are connected with the operator E , the eigenvalues of which we are interested in, with a “consistency condition”.

Definition 1.3. A bounded operator $B : E \rightarrow E$ and a sequence of bounded operators $(B_\iota)_{\iota \in I}$, $B_\iota : E_\iota \rightarrow E_\iota$ are called *consistent* iff there exists a dense subset $F \subset E$, such that for every $f \in F$ there exists a sequence $(f_\iota)_{\iota \in I}$, $f_\iota \in E_\iota$, with

$$f_\iota \rightarrow f \quad \text{and} \quad B_\iota f_\iota \rightarrow Bf.$$

Stummel approximated eigenvalue problems of the kind

$$Au = \lambda Bu$$

by “discrete” eigenvalue problems

$$A_\iota u_\iota = \lambda^\iota B_\iota u_\iota.$$

For the verification of the convergence of the eigenvalues, projection operators $P : E \rightarrow E$, $P_\iota : E_\iota \rightarrow E_\iota$, which are defined by means of integrals of resolvents $A(z)^{-1} := (A - zB)^{-1}$, are examined. For the discrete convergence of the integrals of resolvents, the discrete compactness of $(B_\iota)_\iota$ is crucial.

Definition 1.4. Let $(B_\iota)_{\iota \in I}$ be a sequence of bounded operators $B_\iota : E_\iota \rightarrow E_\iota$. The sequence is called *discretely compact* iff for each $\|\cdot\|_{E_\iota}$ -bounded subsequence $(u_\iota)_{\iota \in I'}$ of $(u_\iota)_{\iota \in I}$, i.e.

$$\|u_\iota\|_{E_\iota} \leq \gamma \in (0, \infty) \quad \text{for all } \iota \in I',$$

a sub-subsequence $(u_\iota)_{\iota \in I''}$ and a $w \in E$ exists, such that

$$B_\iota u_\iota \xrightarrow[\iota \in I'', \iota \rightarrow \infty]{} w.$$

1.3. Convergence of eigenvalues. Given the eigenvalue problem $Au = \lambda Bu$, $\lambda \in \mathbb{C}$, $u \in E$, Stummel's main theorem on the convergence of eigenvalues of discrete operators [8, II.3.2.(8), p. 257] mainly requires consistency and discrete compactness.

Theorem 1.5. *Let $A, (A_\iota)_{\iota \in I}$ and $B, (B_\iota)_{\iota \in I}$ be consistent linear operators. Let $B, B_\iota, \iota \in I$, be compact and let $(B_\iota)_{\iota \in I}$ be discretely compact. Let the resolvent set and the bounding set of A, B be nonempty. Let λ be an eigenvalue of A, B with algebraic multiplicity m and let U be an arbitrary closed neighborhood of λ , which is bounded, and the only element it has in common with the spectrum $\Sigma(A, B)$ of A, B is the eigenvalue λ . Then, there exists $\iota_0 \in \mathbb{N}$, such that for $\iota \geq \iota_0, \iota \in I$, there are exactly m eigenvalues $\lambda_1^\iota, \dots, \lambda_m^\iota$ of A_ι, B_ι (counting algebraic multiplicity) with*

$$\Sigma(A_\iota, B_\iota) \cap U = \{\lambda_1^\iota, \dots, \lambda_m^\iota\}, \quad \iota \in I, \quad \iota \geq \iota_0.$$

The *resolvent set* is the set of all $z \in \mathbb{C}$ for which the resolvent $A(z)^{-1} = (A - zB)^{-1}$ is a bounded linear mapping. The resolvent set is the complement of the spectrum of A, B . The *bounding set* is the set of $z \in \mathbb{C}$, for which there exists a constant $\gamma > 0$ and an $\iota_0 \in I$, such that for every $\iota \geq \iota_0$ z is an element of the resolvent set of A_ι, B_ι and the operator norm of $A_\iota(z)^{-1}$ is bounded uniformly in $\iota \geq \iota_0$ by γ .

The proof of the theorem Stummel shows that the resolvent set and the bounding set coincide, as well as the convergence of the resolvents [8, II.2.2.(1), p. 245]. It is shown that the boundary of an arbitrary compact neighborhood of an eigenvalue, which contains no further eigenvalues, is completely contained in the resolvent set of A_ι, B_ι for ι greater than some ι_0 [8, II.2.1.(3), p. 243]. The positively oriented integrals of the resolvents on this boundary are hence well defined and converge discretely [8, II.2.3.(4), p. 249]. They can be used to discretely define converging projection operators P, P_ι which project the corresponding domain on the algebraic eigenspace of the eigenvalue surrounded. As a consequence $P(E)$ is discretely approximated by $P_\iota(E_\iota)$, which eventually has the same dimension as $P(E)$. Thus, the initially chosen neighborhood eventually contains m eigenvalues of A_ι, B_ι . By choosing smaller and smaller neighborhoods of λ , convergence can be deduced. Even the discrete convergence of the corresponding eigenfunctions can be inferred. For a more detailed treatment the reader is referred to the work of Stummel.

2. APPLICATION IN ASYMPTOTIC STATISTICS

Stummel's theory is to be applied to the approximation of the eigenvalues of an operator

$$(2.1) \quad \begin{aligned} E &\rightarrow E, \\ H : f &\mapsto \int h(\cdot, y) f(y) d\mu(y), \end{aligned}$$

by operators

$$(2.2) \quad \begin{aligned} L_2(\mathbb{R}^d, \mathcal{B}^d, \mu_\iota) &\rightarrow L_2(\mathbb{R}^d, \mathcal{B}^d, \mu_\iota), \\ H_\iota : f &\mapsto \int h_\iota(\cdot, y) f(y) d\mu_\iota(y), \end{aligned}$$

where μ, μ_ι are probability measures on $(\mathbb{R}^d, \mathcal{B}^d)$ with μ_ι converging weakly to μ , $\mu_\iota \Rightarrow \mu$. $E := L_2(\mathbb{R}^d, \mathcal{B}^d, \mu) \cap C_b(R)$ denotes the normed space of square μ -integrable functions which are continuous and bounded with respect to the uniform norm on the support R of μ . E is furnished with the $L_2(\mathbb{R}^d, \mathcal{B}^d, \mu)$ -norm. The conditions we will impose on the kernel h of H will guarantee that H indeed maps E into itself. The restriction on continuous functions will be no great limitation of the result, unlike the restriction on bounded functions.

Theorem 2.1. *Let $E := L_2(\mathbb{R}^d, \mathcal{B}^d, \mu) \cap C_b(R)$ be furnished with the $L_2(\mathbb{R}^d, \mathcal{B}^d, \mu)$ -norm, and let R be the support of μ and $\mu_\iota \Rightarrow \mu$ with*

$$(2.3) \quad \text{supp}(\mu_\iota) \subset R \quad \text{for all } \iota \in I.$$

Then, the spaces $E_\iota := L_2(\mathbb{R}^d, \mathcal{B}^d, \mu_\iota)$ form a discrete approximation of E with the mapping Q being defined by $(u_\iota)_\iota \in Q(u)$ iff

$$\int |u_\iota - u|^2 d\mu_\iota \xrightarrow{\iota \rightarrow \infty} 0$$

for $u \in E$. For the integration of u with respect to μ_ι choose the restriction of u on the support $\text{supp}(\mu_\iota)$ of μ_ι vanishing on $\mathbb{R}^d \setminus \text{supp}(\mu_\iota)$.

Proof. For simplicity, we denote the norm on E with $\|\cdot\|$ and the norm on E_ι with $\|\cdot\|_\iota$. For any $u \in E$ the equivalence class is nonempty. This can be seen by letting u_ι be the restriction of u to the support of μ_ι . Condition (ii) is fulfilled, due to

$$\| \|u_\iota\|_\iota - \|u\| \| \leq \|u_\iota - u\|_\iota + \| \|u\|_\iota - \|u\| \|$$

for $(u_\iota)_\iota \in Q(u)$, $u \in E$, and the weak convergence of μ_ι to μ . The verification of (iii) is straightforward. \square

Given the situation of Theorem 2.1 under certain conditions, the operators H , $(H_\iota)_\iota$ are consistent and $(H_\iota)_\iota$ discretely compact.

Lemma 2.2. *Let the conditions of Theorem 2.1 be fulfilled. Let $(h_\iota)_\iota$ be a uniformly $\|\cdot\|_\infty$ -bounded sequence of functions*

$$h_\iota : (R \times R, \mathcal{R} \otimes \mathcal{R}) \rightarrow (\mathbb{R}, \mathcal{B})$$

symmetric in its arguments with \mathcal{R} the Borel- σ -algebra restricted to R . For every $\varepsilon > 0$ let there be a $\delta > 0$ such that for any $\iota \in I$

$$(2.4) \quad \sup\{|h_\iota(x, y) - h_\iota(x', y)| : |x - x'| < \delta, x, x' \in R, y \in R\} < \varepsilon.$$

Then, the operators $(H_\iota)_\iota$ defined above are discretely compact.

Since h_ι is defined on $R \times R$, we can evaluate $(H_\iota u_\iota)(x)$ with $u_\iota \in L_2(\mathbb{R}^d, \mathcal{B}^d, \mu_\iota)$ not only for $x \in \text{supp}(\mu_\iota)$ but also for $x \in R$ by

$$(H_\iota u_\iota)(x) := \int h_\iota(x, y) u_\iota(y) \, d\mu_\iota(y).$$

Proof. Since the kernels are uniformly $\|\cdot\|_\infty$ -bounded the operators $(H_\iota)_\iota$ are bounded operators. Let $(R_k)_k, k \in \mathbb{N}$, be an increasing sequence of compact sets with $R_k := R \cap [-\kappa, +\kappa]^d$ for some $\kappa > k$ with $\mu(R \setminus R_k) \leq \frac{1}{2^k}$ and $\mu_\iota(R \setminus R_k) \leq \frac{1}{2^k}$ for every $\iota \in I$. Now, let $(u_\iota), \iota \in I'$, be a subsequence of $(u_\iota), \iota \in I$, with $\|u_\iota\|_\iota \leq \gamma$ for $\iota \in I'$.

The functions $(g_\iota), \iota \in I', g_\iota : R \rightarrow \mathbb{R}$ with $g_\iota := H_\iota u_\iota$ are uniformly $\|\cdot\|_\infty$ -bounded and by the continuity condition (2.4) are equicontinuous on R . By the Arzela-Ascoli theorem there exists a subsequence $(g_\iota)_{\iota \in I''_k}$ of $(g_\iota)_{\iota \in I'}$ and a continuous, complex-valued function w_k being $\|\cdot\|_\infty$ -bounded on R_k such that

$$\max_{x \in R_k} |g_\iota(x) - w_k(x)| \xrightarrow{\iota \in I''_k, \iota \rightarrow \infty} 0.$$

By successive selection of the $(I''_k)_{k \in \mathbb{N}}$ the sequence can be created such that $I''_{k+1} \subset I''_k$ for every $k \in \mathbb{N}$. Let I'' be the diagonal sequence of the sequences $(I''_k)_{k \in \mathbb{N}}$. Then,

$$w(x) := \lim_{\iota \in I'', \iota \rightarrow \infty} g_\iota(x) = \lim_{\iota \in I'', \iota \rightarrow \infty} H_\iota u_\iota(x), \quad x \in R,$$

is $\|\cdot\|_\infty$ -bounded, and due to the selection of the $(I''_k)_{k \in \mathbb{N}}$ it is a continuous function on R . Hence, we have found $w \in E$ with

$$\begin{aligned} & \|H_\iota u_\iota - w\|_\iota^2 \\ &= \underbrace{\int_{R \setminus R_k} |H_\iota u_\iota(x) - w(x)|^2 d\mu_\iota(x)}_{\leq \text{const.} \cdot \mu_\iota(R \setminus R_k) \leq \text{const.} \cdot 2^{-k}} + \underbrace{\int_{R_k} |H_\iota u_\iota(x) - w_k(x)|^2 d\mu_\iota(x)}_{\xrightarrow{\iota \in I'', \iota \rightarrow \infty} 0} \end{aligned}$$

i.e. for $k \rightarrow \infty$ discrete compactness follows. □

The consistency proof mainly uses the same procedures as the proof of discrete-compactness.

Lemma 2.3. *Let the conditions of the preceding lemma be fulfilled. For every $x \in R$ let $(h_\iota(x, \cdot))_\iota$ converge uniformly on every compact set to the symmetric function $h(x, \cdot)$ where*

$$h : (R \times R, \mathcal{R} \otimes \mathcal{R}) \rightarrow (\mathbb{R}, \mathcal{B})$$

denotes a bounded function such that for every $\varepsilon > 0$ there exists a $\delta > 0$ with

$$(2.5) \quad \sup\{|h(x, y) - h(x', y)| : |x - x'| < \delta, x, x' \in R, y \in R\} < \varepsilon.$$

Then, the operator H and the sequence of operators $(H_\iota)_\iota$ are consistent, and H is a compact operator in E .

Proof. We make use of the notations introduced in the proof of the preceding lemma. Due to the bounded kernel the operator H is bounded. To prove consistency, let $u \in E$ and let the functions u_ι of the sequence $(u_\iota)_\iota$ be defined as the restriction of u on $\text{supp}(\mu_\iota)$ (which is vanishing on $\mathbb{R}^d \setminus \text{supp}(\mu_\iota)$). $(u_\iota)_\iota$ is uniformly $\|\cdot\|_\iota$ -bounded, and by the proof of the last theorem $(g_\iota)_\iota$ with $g_\iota := H_\iota u_\iota$ is

a uniformly $\|\cdot\|_\infty$ -bounded sequence being equicontinuous on R_k for every $k \in \mathbb{N}$. For $x \in R$

$$g_\iota(x) := \int h_\iota(x, y)u(y) d\mu_\iota(y) \xrightarrow{\iota \rightarrow \infty} \int h(x, y)u(y) d\mu(y) = Hu(x)$$

by the uniform convergence and the $\|\cdot\|_\infty$ -boundedness of h, u . This pointwise convergence together with the equicontinuity yields the uniform convergence and hence discrete convergence of g_ι to Hu on R .

We will now show that the bounded and linear operator H is compact in E . We need to show (see e.g. Dunford and Schwartz [3, Theorem 15, p. 22]) that for a sequence $(f_n)_n$ of functions $f_n \in E$ which are uniformly $\|\cdot\|$ -bounded by 1, the sequence $(Hf_n)_n$ has a point of accumulation $f \in E$: Following the proof of Lemma 2.2, the sequence $(Hf_n)_n$ is $\|\cdot\|_\infty$ -bounded on R and by using (2.5) is equicontinuous on R . The Arzela-Ascoli theorem with diagonal sequence procedure gives a subsequence of $(Hf_n)_n$ which converges uniformly on every compact set R_k to an $\|\cdot\|_\infty$ -bounded function f which is continuous on R . Due to the $\|\cdot\|_\infty$ -boundedness of every subsequence, L_2 -convergence of a certain sub-subsequence and thus compactness of H follows. \square

In order to apply Stummel's main theorem, Theorem 1.5, to the integral operators H, H_ι , we define the operators A, A_ι , $\iota \in I$, as identity on E, E_ι , $\iota \in I$. Thus, the resolvent set and the bounding set contain 0 and are therefore nonempty. The operators $A, (A_\iota)_\iota$ are consistent by Definition 1.2. Let $B := H$ and $B_\iota := H_\iota$, $\iota \in I$. Under the conditions of Lemma 2.3 the operators $(B_\iota)_\iota$ are discretely compact and $B, (B_\iota)_\iota$ are consistent. The operators B_ι , $\iota \in I$, are Hilbert-Schmidt-operators and therefore compact, as is B by Lemma 2.3. Due to the continuity conditions on h the eigenfunctions to nonvanishing eigenvalues of B taken as an operator on $L_2(R, \mathcal{R}, \mu)$ are continuous. Hence, the restriction of B from $L_2(R, \mathcal{R}, \mu)$ to E does not change the nonvanishing eigenvalues or their multiplicity. The operators B, B_ι are self-adjoint and thus can be understood as operators over the field \mathbb{R} without changing the nonvanishing eigenvalues. Hence, Theorem 1.5 can be applied.

Stummel's eigenvalue problem $Au = \lambda Bu$, here $u = \lambda Hu$, leads to the inverse eigenvalues of the usual eigenvalue problem $\lambda u = Hu$ that we are interested in. The following result, Stummel's main theorem applied to integral operators, is valid for either definition.

Corollary 2.4. *Let the conditions of Lemma 2.2 and Lemma 2.3 be fulfilled. Then, for every nonvanishing eigenvalue λ of H with multiplicity $m \in \mathbb{N}$ and every sufficiently small neighborhood U of λ there exists an ι_0 such that the operator H_ι has exactly m eigenvalues (counting multiplicity) in U for every $\iota \geq \iota_0$.*

An important application of this theory in the field of asymptotic statistics is the possibility to approximate the distribution of weighted sums of squares of independent standard normals with the weights being the eigenvalues of a positive trace-class operator H . The approximation is a weighted sum of squares of independent standard normals, where the weights are the eigenvalues of the positive trace-class operator H_ι .

To obtain this result we demand, in addition to the requirements of the last corollary that the operators $(H_\iota)_\iota, H$ are positive and of trace-class.

Theorem 2.5. *Let the requirements of Lemma 2.2 and Lemma 2.3 be fulfilled and let (H_ι) , H be positive and of trace class. Then,*

$$S^\iota := \sum_{\sigma \in \Sigma_\iota} \lambda_\sigma^\iota Z_\sigma^2 \xrightarrow[\iota \rightarrow \infty]{\mathcal{D}} \sum_{\sigma \in \Sigma} \lambda_\sigma Z_\sigma^2 =: S,$$

where $(Z_\sigma)_\sigma$ is an iid sequence of standard normally distributed random variables. The sequences $(\lambda_\sigma^\iota)_\sigma$, $\sigma \in \Sigma_\iota$, denote the positive eigenvalues of H_ι and $(\lambda_\sigma)_\sigma$, $\sigma \in \Sigma \subset \mathbb{N}$, denote the positive eigenvalues of H . Each eigenvalue appears according to its multiplicity.

Proof. By (2.4) and (2.5) h_ι converges uniformly to h on every compact set. Due to the convergence in distribution $\mu_\iota \implies \mu$ the trace H_ι converges to the trace of H .

In the following let $\lambda_\sigma := 0$ for $\sigma \in \mathbb{N}$ if $\sigma \notin \Sigma$ and let $\lambda_\sigma^\iota := 0$ for $\sigma \in \mathbb{N}$ if $\sigma \notin \Sigma_\iota$. To show convergence in distribution, we show the stronger L_1 -convergence: Let $\varepsilon > 0$ be arbitrary. Let $\sigma_0 \in \mathbb{N}$ be such that

$$\sum_{\substack{\sigma > \sigma_0 \\ \sigma \in \Sigma}} \lambda_\sigma \leq \varepsilon.$$

Let the positive eigenvalues (λ_σ^ι) of H_ι be indexed by $\Sigma_\iota \subset \mathbb{N}$ such that for every $\sigma \leq \sigma_0$, $\sigma \in \mathbb{N}$, we have $\lambda_\sigma^\iota \rightarrow \lambda_\sigma$ for $\iota \rightarrow \infty$. With an appropriate change of indexing this follows from Corollary 2.4. According to this indexing let $(Z_\sigma)_\sigma$ be assigned to $(\lambda_\sigma^\iota)_\sigma$, $(\lambda_\sigma)_\sigma$, respectively. By the pointwise convergence of the eigenvalues there exists an $\iota_0 \in \mathbb{N}$ such that for every $\iota > \iota_0$

$$|\lambda_\sigma^\iota - \lambda_\sigma| \leq \frac{\varepsilon}{2^\sigma} \quad \text{for } 1 \leq \sigma \leq \sigma_0,$$

and by the convergence of the trace we can increase ι_0 until

$$\left| \sum_{\sigma \in \Sigma_\iota} \lambda_\sigma^\iota - \sum_{\sigma \in \Sigma} \lambda_\sigma \right| \leq \varepsilon \quad \text{for every } \iota > \iota_0.$$

Then,

$$\begin{aligned} E|S^\iota - S| &\leq \sum_{\sigma \leq \sigma_0} |\lambda_\sigma^\iota - \lambda_\sigma| + \sum_{\sigma > \sigma_0} |\lambda_\sigma^\iota - \lambda_\sigma| \\ &\leq 2\varepsilon + \sum_{\sigma > \sigma_0} \lambda_\sigma^\iota + \sum_{\sigma \leq \sigma_0} \lambda_\sigma^\iota - \sum_{\sigma \leq \sigma_0} \lambda_\sigma + \sum_{\sigma \leq \sigma_0} \lambda_\sigma + \sum_{\sigma > \sigma_0} \lambda_\sigma - \sum_{\sigma} \lambda_\sigma \\ &\leq 5\varepsilon, \end{aligned}$$

i.e. we have L_1 -convergence. □

3. EXAMPLE

As an example, the preceding theorem will be applied to approximate the limiting null distribution of the two-sample V-statistic

$$\begin{aligned} T_{m,n} = \frac{mn}{m+n} &\left[\frac{2}{mn} \sum_{j=1}^m \sum_{k=1}^n \frac{|X_j - Y_k|^2}{1 + |X_j - Y_k|^2} - \frac{1}{m^2} \sum_{j=1}^m \sum_{k=1}^m \frac{|X_j - X_k|^2}{1 + |X_j - X_k|^2} \right. \\ &\left. - \frac{1}{n^2} \sum_{j=1}^n \sum_{k=1}^n \frac{|Y_j - Y_k|^2}{1 + |Y_j - Y_k|^2} \right] \end{aligned}$$

used to measure the distance between the d -variate distributions P^X and P^Y on the basis of realizations of independent random variables $X_1, X_2, \dots \sim P^X$ and $Y_1, Y_2, \dots \sim P^Y$. The statistic can be written as

$$T_{m,n} = \frac{mn}{m+n} \int \cdots \int \tilde{h}(x_1, y_1; x_2, y_2) dP_n^Y(y_1) dP_n^Y(y_2) dP_m^X(x_1) dP_m^X(x_2)$$

with

$$\tilde{h}(x_1, y_1; x_2, y_2) = \frac{|x_1 - y_2|^2}{1 + |x_1 - y_2|^2} + \frac{|x_2 - y_1|^2}{1 + |x_2 - y_1|^2} - \frac{|x_1 - x_2|^2}{1 + |x_1 - x_2|^2} - \frac{|y_1 - y_2|^2}{1 + |y_1 - y_2|^2},$$

where $|\cdot|$ denotes the euclidian norm and P_m^X, P_n^Y are the random empirical measures of X_1, \dots, X_m and Y_1, \dots, Y_n , respectively.

In order to carry out a two sample tests of $H_0 : P^X = P^Y$ against $H_1 : P^X \neq P^Y$, the null distribution of $T_{m,n}$ is needed. Let $F_0 = P^X = P^Y$ be the nondegenerate distribution under the null hypothesis. Applying results of Neuhaus [6] it can be shown that the limiting null distribution for $m, n \rightarrow \infty$ is that of a weighted sum of squares of independent standard normals. The weights are the eigenvalues divided by two of the positive trace class Hilbert-Schmidt operator

$$\begin{aligned} L_2(\mathbb{R}^d \times \mathbb{R}^d, \mathcal{B}^d \otimes \mathcal{B}^d, F_0 \otimes F_0) &\rightarrow L_2(\mathbb{R}^d \times \mathbb{R}^d, \mathcal{B}^d \otimes \mathcal{B}^d, F_0 \otimes F_0), \\ \tilde{H} : f &\mapsto \int \tilde{h}(\cdot, \cdot; x, y) f(x, y) d(F_0 \otimes F_0)(x, y). \end{aligned}$$

It can be further shown (see [4]) that these weights equal the eigenvalues of the positive trace class Hilbert-Schmidt operator

$$\begin{aligned} E &\rightarrow E, \\ H : f &\mapsto \int h(\cdot, y) f(y) dF_0(y), \end{aligned}$$

with $h(x_1, x_2) := \int \tilde{h}(x_1, y_1; x_2, y_2) d(F_0 \otimes F_0)(y_1, y_2)$ and E being the space of continuous bounded functions on $R = \text{supp}(F_0)$ equipped with the L_2 -norm with respect to F_0 . This correspondence leads our interest to the approximation of the eigenvalues of H which can be achieved with the methods introduced in the last section. For that matter let

$$\begin{aligned} E_N &\rightarrow E_N, \\ H_N : f &\mapsto \int h_N(\cdot, y) f(y) dF_N(y) \end{aligned}$$

with kernel $h_N(x_1, x_2) := \int \tilde{h}(x_1, y_1; x_2, y_2) d(F_N \otimes F_N)(y_1, y_2)$, $N := m + n$, $F_N = \frac{m}{m+n} P_m^X + \frac{n}{m+n} P_n^Y$ being the empirical distribution of the pooled sample and $E_N := L_2(\mathbb{R}^d, \mathcal{B}^d, F_N)$. Omitting tedious notation concerning the sequence of sample sizes we proceed as follows: First, a sequence of sample sizes is fixed. Then, the spaces E_N are constructed. Almost surely F_N converges weakly to F_0 and thus, by Theorem 2.1 the spaces E_N almost surely form a discrete approximation of E . Finally, given a sequence of observations such that $(F_N)_N$ converges weakly to F_0 , the sequence of operators $(H_N)_N$ is derived. This sequence will now be shown to be consistent with H and discretely compact.

Corollary 3.1. *For iid random variables $X_1, X_2, \dots, Y_1, Y_2, \dots \sim F_0$ the operators $H, (H_N)_N$ almost surely are consistent, $(H_N)_N$ almost surely is discretely compact and H is compact in E .*

Proof. As noted, the empirical distributions $(F_N)_N$ converge almost surely weakly to F_0 (and condition (2.3) is obviously satisfied). Let $(F_N)_N$ be such a sequence. Using

$$\left| \frac{|x-y|^2}{1+|x-y|^2} - \frac{|x'-y|^2}{1+|x'-y|^2} \right| \leq 2|x-x'|$$

gives $|h_N(x, y) - h_N(x', y)| \leq 4|x-x'|$, and thus conditions (2.4) and analogously (2.5) are fulfilled. The uniform convergence of $(h_N(x, \cdot))_N$ to $h(x, \cdot)$ on every compact set follows from the pointwise convergence due to the weak convergence of laws and the just established equicontinuity together with the symmetry of the kernel functions. Hence, the corollary follows from Lemma 2.2 and Lemma 2.3. \square

As mentioned above, the operators H and as a consequence the operators $(H_N)_N$ can be shown to be positive. H is of trace class and so are $(H_N)_N$, since they only have a finite number of eigenvalues. The requirements of Theorem 2.5 being fulfilled, we almost surely have

$$S^N := \sum_{\sigma \in \Sigma_N} \lambda_\sigma^N Z_\sigma^2 \xrightarrow[N \rightarrow \infty]{\mathcal{D}} \sum_{\sigma \in \Sigma} \lambda_\sigma Z_\sigma^2 =: S,$$

where $(\lambda_\sigma^N)_\sigma, (\lambda_\sigma)_\sigma$ are the eigenvalues of H_N, H , respectively.

The eigenvalues of H_N can easily be obtained by a linear equation system: Because F_N is an empirical measure it is sufficient to evaluate the operator H_N at Z_1, \dots, Z_N with $Z_1 := X_1, \dots, Z_m := X_m, Z_{m+1} := Y_1, \dots, Z_N := Y_n$. For $u \in L_2(\mathbb{R}^d, \mathcal{B}^d, F_N)$ this gives

$$(H_N u)(Z_i) = \frac{1}{N} \sum_{j=1}^N h_N(Z_i, Z_j) u(Z_j) = \frac{1}{N^3} \sum_{j,k,\ell=1}^N \tilde{h}(Z_i, Z_k; Z_j, Z_\ell) u(Z_j),$$

$i = 1, \dots, N$, and leads to the eigenvalue problem

$$\lambda \vec{u} = A \vec{u}$$

with $\vec{u} = (u(Z_1), \dots, u(Z_N))'$ and $A = (a_{i,j})_{i,j}$,

$$a_{i,j} := \frac{1}{N^3} \sum_{k,\ell=1}^N \tilde{h}(Z_i, Z_k; Z_j, Z_\ell), \quad i, j \in \{1, \dots, N\}.$$

Without going into detail we note that the test statistic works well for scale alternatives (and in the univariate case for normally distributed random variables is far better than Cramér-von Mises and Kolmogorov-Smirnov).

REMARKS AND CONCLUSION

The mentioned example is one element of a class of two-sample tests introduced by Baringhaus and the author [1, 2] and was chosen to briefly outline a possible application of this approximation method. For results on the performance of the tests see [2]. For additional information on the operators H, \hat{H} and their positivity, trace class property as well as correspondence of eigenvalues, the reader is referred to [4].

Although using this way of approximation the finite sample distribution of the V-statistic is not taken into account, the procedure was shown in [2] to have the same properties as Monte-Carlo bootstrap methods. An advantage of the procedure which might be transferred to other tests is that no Monte-Carlo bootstrap

methods are needed to obtain an approximation of the critical value even though the mentioned test statistic is not asymptotically distribution free.

ACKNOWLEDGMENTS

The author is grateful to the referee for pointing out the work of V. Koltchinskii and E. Giné.

REFERENCES

- [1] L. Baringhaus and C. Franz, On a new multivariate two-sample test. *J. Multivariate Anal.* **88** (2004), 190-206. MR2021870 (2004i:62092)
- [2] ———, A family of multivariate two-sample tests. *submitted*.
- [3] N. Dunford and J. T. Schwartz, *Linear operators. Part I*. John Wiley and Sons Inc., New York, 1988. MR1009162 (90g:47001a)
- [4] C. Franz, "Asymptotische Eigenschaften einer neuen Klasse von nichtparametrischen Zwei-Stichproben-Tests." Doctoral thesis, University of Hannover, 2004.
- [5] V. Koltchinskii and E. Giné, Random matrix approximation of spectra of integral operators. *Bernoulli* **6** (2000), 113-167. MR1781185 (2001e:47080)
- [6] G. Neuhaus, Functional limit theorems for U-statistics in the degenerate case. *J. Multivariate Anal.* **7** (1977), 424-439. MR0455084 (56:13324)
- [7] F. Stummel, Diskrete Konvergenz linearer Operatoren. I. *Math. Ann.* **190** (1970), 45-92. MR0291870 (45:959)
- [8] ———, Diskrete Konvergenz linearer Operatoren. II. *Math. Z.* **120** (1971), 231-264. MR0291871 (45:960)
- [9] ———, "Diskrete Konvergenz linearer Operatoren. III.", Linear operators and approximation (Proc. Conf., Oberwolfach, 1971), Internat. Ser. Numer. Math. **20**, 196-216, Birkhäuser, Basel, 1972. MR0410431 (53:14180)

INSTITUT FÜR MATHEMATISCHE STOCHASTIK, UNIVERSITÄT HANNOVER, WELFENGARTEN 1,
30167 HANNOVER, GERMANY
E-mail address: carsten.franz@gmail.com