PROCEEDINGS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 145, Number 5, May 2017, Pages 2131–2138
http://dx.doi.org/10.1090/proc/13354
Article electronically published on October 27, 2016

SEPARABILITY OF REPRODUCING KERNEL SPACES

HOUMAN OWHADI AND CLINT SCOVEL

(Communicated by Mark M. Meerschaert)

Abstract. We demonstrate that a reproducing kernel Hilbert or Banach space of functions on a separable absolute Borel space or an analytic subset of a Polish space is separable if it possesses a Borel measurable feature map.

1. Introduction

Reproducing kernel Hilbert spaces (RKHS) (see e.g. Berlinet and Thomas-Agnan [2] and Steinwart and Christmann [33, Sec. 4]) are important in statistics and learning theory. Moreover, when using these spaces in probability and statistics, separability has powerful effects. For example, for any separable metrizable space X we have: $\mathcal{B}(X \times X) = \mathcal{B}(X) \times \mathcal{B}(X)$ for the Borel σ-algebras [12, Prop. 4.1.7], the Ky-Fan metric can be defined so as to metrize convergence in probability [12, Thm. 9.2.2], convergence in probability implies convergence in law [12, Prop. 9.3.5], convergence in law is metrized by the Prokhorov metric [12, Thm. 11.3.3], the space of probability measures with the weak topology is separable and metrizable [1, Thm. 15.12], and the Kantorovich-Rubinstein and Strassen theorems have sharp forms [12, Thms. 11.8.2 and 11.6.2]. Moreover, separable Hilbert spaces are Polish so that we have all the machinery of descriptive set theory available, regular conditional probabilities exist [12, Thm. 10.2.2], Bochner integration is simple [1, Lems. 11.37 and 11.39], and all probability measures on them are tight [11, Thm. 69, 77-III]. Most importantly, by a classical result (see e.g. Halmos [19, Prob. 17]), all separable Hilbert spaces are isomorphic with $\ell^2(\mathbb{N})$.

According to Montgomery [27], "Separability is a property which greatly facilitates work in metric spaces, but it may be of some interest to point out that this property has been unnecessarily assumed in the proofs of certain theorems concerning such spaces and concerning functions defined on them." Indeed, many works do assume separability of the RKHS. For example, Steinwart and Christmann's [33, Thm. 7.22] oracle inequality for SVMs, Christmann and Steinwart [8, Thms. 7 and 12], [7], [6], Steinwart and Christmann [34], [35], De Vito, Rosasco and Toigo [9], Hable and Christmann [18, Thm. 3.2], Lukić and Beder [25], Steinwart [32] and Vovk [38, Thm. 3]. De Vito, Umanità and Villa [10] assume it in their generalization of Mercer's theorem to matrix valued kernels, and Christmann, Van Messem and Steinwart [8] assert that Support Vector Machines (SVMs) are known to be consistent and robust for classification and regression if they are based on a Lipschitz continuous loss function and on a bounded kernel with a separable reproducing

Received by the editors July 13, 2015 and, in revised form, March 3, 2016, May 31, 2016, and July 5, 2016.

2010 *Mathematics Subject Classification.* Primary 46E22.

kernel Hilbert space which is dense in $L^1(\mu)$, where μ is the marginal distribution of the data-generating distribution. Cambanis [3] proves that a stochastic process with index set a Borel subset of a Polish space has a measurable modification if and only if the reproducing kernel corresponding to the autocorrelation function is measurable and its corresponding RKHS is separable, and that a second order process with index set the real line is oscillatory if and only if its RKHS is separable. Nashed and Walter [28] require a separable RKHS in their development of sampling theorems for functions in reproducing kernel Hilbert spaces, and Zhang and Zhang [41] in reproducing kernel Banach spaces. Hein and Bousquet [22] require it and give some sufficient conditions for it. For an example of a non-separable RKHS, see Canu, Mary, and Rakotomamonjy [4, Ex. 8.1.6].

Let us now briefly discuss the topological spaces under consideration. A Polish space is a separable completely metrizable space and a Suslin space a Hausdorff continuous image of a Polish space. Following Frolik [16], a metrizable space X is said to be absolute Borel if $X \subset Z$ is a Borel subset for all metrizable Z for which it is a subspace. Moreover, Frolik [15] introduces *bianalytic spaces* as Suslin spaces such that their complement in their Čech compactification is also Suslin and, in [15, Thm. 12], shows that a metrizable space is separable absolute Borel if and only if it is bianalytic. On the other hand, a subset of a Polish space is called analytic if it is Suslin. Indeed, the two types of spaces considered here, separable absolute Borel spaces and analytic subsets of Polish spaces, are very general. For example, for a Borel subset of a Polish space, Frolik [16, Thm. 1] asserts that it is separable absolute Borel, and the famous Suslin theorem (see e.g. Kechris [23, Thm. 13.7]) asserts that it is analytic. That is, they both include any Borel subset of a Polish space, in particular any Borel subset of a separable Banach space, so any Borel, open, or closed subset of $\mathbb{R}^n$. Since $\mathbb{R}^n$ is Polish, it follows that this class also includes any analytic subset of $\mathbb{R}^n$. Counterexamples to these spaces include non-separable spaces, non-metrizable spaces and non-Suslin spaces. Moreover, Lusin's theorem (see e.g. Kechris [23, Thm. 21.10]) asserts that all analytic subsets of a Polish space are universally measurable; that is, for every σ-finite measure it is trapped between two Borel subsets of the same measure. Consequently, any subset of a Polish space which is not universally measurable is a counterexample.

Reproducing kernel Hilbert spaces are Hilbert spaces of real-valued functions such that pointwise evaluation is continuous. In their generalization to reproducing kernel Banach spaces (RKBS), Zhang, Xu, and Zhang [40] stipulate that a RKBS on a set X is a reflexive Banach space of real-valued functions on X whose dual space is isometric with a Banach space of functions on X such that pointwise evaluation is continuous for both the Banach space and its dual. They then proceed to develop the theory much along the lines of RKHSs. In particular, in [40, Thm. 2] they show that RKBSs possess reproducing kernels. Moreover, in [40, Thm. 3] they demonstrate that if $\Phi : X \to \mathcal{W}$ is a map to a reflexive Banach space $\mathcal{W}$ and $\Phi^* : X \to \mathcal{W}^*$ is a map to its dual such that the linear span of the image of each map is dense, then a RKBS is determined with reproducing kernel $K(x,x') = [\Phi^*(x), \Phi(x')]$, where $[\,,\,]$ is the dual pairing between $\mathcal{W}^*$ and $\mathcal{W}$. Moreover, in [40, Thm. 4] they assert that all RKBSs possess such maps. Consequently, we refer to such maps $\Phi : X \to \mathcal{W}$ and $\Phi^* : X \to \mathcal{W}^*$ as primary and secondary feature maps for the RKBS.

This generalization to RKBSs has generated much interest lately; for example see recent results of Fukumizu, Lanckriet and Sriperumbudur [17], Zhang and Zhang [42], Fasshauer, Hickernell and Ye [13] in machine learning, in particular of Song, Zhang and Hickernell [30] on sparse learning, and the recent results of Zhang and Zhang [41], Han, Nashed, and Sun [20] and Christensen [5] concerning sampling expansions, frames and Riesz bases in Banach spaces.

It is the purpose of this paper to establish separability for both RKHSs and RKBSs when the domain is a separable absolute Borel space or an analytic subset of a Polish space, in particular when it is a Borel subset of a Polish space, under the simple assumption that the reproducing kernel space possesses a Borel measurable feature map.

2. Main results

Before our main results, we review some existing results regarding the separability of RKHSs. We will consider both when X is not a topological space and when it is. When X is not topological, Berlinet and Thomas-Agnan [2, Thm. 15, pg. 33] shows that a RKHS H is separable if there is a countable subset $X_0 \subset X$ such that $f \in H$ and $f(x) = 0, x \in X_0$ implies that $f = 0$. Moreover, a result of Fortet [14, Thm. 1.2] asserts that a RKHS with kernel k is separable if and only if for all $\varepsilon > 0$ there exists a countable partition $B_j, j \in \mathbb{N}$, of X such that for all $j \in \mathbb{N}$ and all $x_1, x_2 \in B_j$ we have

$$k(x_1, x_1) + k(x_2, x_2) - k(x_1, x_2) - k(x_2, x_1) < \varepsilon\,.$$

Regarding the separability of RKBSs, an "if and only if" characterization is obtained through a generalization of Fortet's theorem from RKHSs to RKBSs. We suspect the proof of our version, Theorem 2.2, is similar to Fortet's [14, Thm. 2.1] for RKHSs, but it is not written down there. Indeed, Fortet's result is a regularity condition on the pullback (pseudo) metric

$$d_\Phi(x_1, x_2) := \|\Phi(x_1) - \Phi(x_2)\|_{H_1} = \sqrt{k(x_1, x_1) + k(x_2, x_2) - k(x_1, x_2) - k(x_2, x_1)}$$

to X determined by a feature map $\Phi : X \to H_1$. In particular, Fortet's condition then becomes: for all $\epsilon > 0$ there exists a countable partition $B_j, j \in \mathbb{N}$, of X such that

$$d_\Phi(x_1, x_2) < \sqrt{\varepsilon}, \quad x_1, x_2 \in B_j, j \in \mathbb{N}\,. \tag{2.1}$$

We begin with a preparatory lemma asserting that the separability of the image of the feature map implies the separability of the corresponding RKHS or RKBS. This lemma is used in both the proof of our generalization of Fortet's result, Theorem 2.2, which is valid when X is not a topological space, and our main result, Theorem 2.4, valid when X is a separable absolute Borel space or an analytic subset of a Polish space.

Lemma 2.1. Consider a (RKBS) RKHS $\mathcal{K}$ of functions on a set X with feature (Banach) Hilbert space $\mathcal{W}$ and (primary) feature map $\Phi : X \to \mathcal{W}$. If $\Phi(X) \subset \mathcal{W}$ is a separable subspace, then $\mathcal{K}$ is separable.

We can now present our generalization of Fortet's result to RKBSs expressed in terms of the pseudometric space (X, d_Φ).[1]

[1] We would like to thank one of the referees for pointing out the possible connection between Fortet's condition (2.1) and a Lindelöf type condition on the pseudometric space (X, d_Φ).

Theorem 2.2. A RKBS $\mathcal{K}$ of functions on a set X is separable if and only if there exists a feature Banach space $\mathcal{W}$ and feature map $\Phi : X \to \mathcal{W}$ such that the topological space (X, d_Φ) determined by the pullback pseudometric

$$d_\Phi(x_1, x_2) := \|\Phi(x_1) - \Phi(x_2)\|_{\mathcal{W}}, \quad x_1, x_2 \in X,$$

is separable.

Now let us consider the case when X is a topological space. Since separability is preserved under continuous maps (see e.g. [39, Thm. 16.4]), Lemma 2.1 implies the RKBS version of Steinwart and Christmann [33, Lem. 4.33] when combined with [33, Lem. 4.29]: A RKBS of functions on a separable space X is separable if it has a continuous feature map. Steinwart and Christmann [33, Lem. 4.33] assert that if X is separable and the kernel k corresponding to the RKHS H is continuous, then H is separable. More generally, Steinwart and Scovel [36, Cor. 3.6] show that if there exists a finite and strictly positive Borel measure on X, then every bounded and separately continuous kernel k has a separable RKHS. However, to obtain our main result, our primary tool to derive separability comes from theorems of Stone [37, Thm. 16, pg. 32] when X is separable absolute Borel, and Srivastava's [31, Thm. 4.3.8] version of Simpson [29] when X is an analytic subset of a Polish space. It is interesting to note that Srivastava's proof is different from Simpson's in that it does not use Stone's theorem [37, Thm. 16, pg. 32].

Lemma 2.3. Let X be separable absolute Borel or an analytic subset of a Polish space and let Y be a metric space, and suppose that $f : X \to Y$ is Borel measurable. Then $f(X) \subset Y$ is separable.

Steinwart and Christmann [33, Lem. 4.25] show that separate measurability of the kernel combined with separability of the corresponding RKHS implies that the canonical feature map is measurable. Our main result is a kind of converse when X is separable absolute Borel or an analytic subset of a Polish space.

Theorem 2.4. Let X be separable absolute Borel or an analytic subset of a Polish space and let $\mathcal{K}$ be a RKHS with measurable feature map or a RKBS with measurable primary feature map of real-valued functions on X. Then $\mathcal{K}$ is separable.

3. Proofs

3.1. **Proof of Lemma 2.1.** For RKHSs this assertion is contained in the proof of Steinwart and Christmann [33, Lem. 4.33]. Roughly, the argument is that rational linear combinations are dense in the linear span of $\Phi(X)$ and the linear span is dense in the closed linear span in the metric defined in the proof of [33, Thm. 4.21]. For the RKBS case, since $\Phi : X \to \mathcal{W}$ is a primary feature map it satisfies $\overline{span(\Phi(X))} = \mathcal{W}$. Moreover, since $\Phi(X) \subset \mathcal{W}$ is separable, the same argument as used in the RKHS case shows that the closed linear span $\overline{span(\Phi(X))} = \mathcal{W}$ is separable, so we conclude that $\mathcal{W}$ is separable. Since $\mathcal{W}$ is reflexive it follows from [26, Cor. 1.12.12] that $\mathcal{W}^*$ is separable. Moreover, Zhang, Xu and Zhang [40, Thm. 3] imply that the dual Banach space is

$$\mathcal{K}^* := \{[\Phi(\cdot), u^*] : u^* \in \mathcal{W}^*\}$$

with norm

$$\|[\Phi(\cdot), u^*]\|_{\mathcal{K}^*} := \|u^*\|_{\mathcal{W}^*},$$

so that the mapping from $\mathcal{W}^*$ to $\mathcal{K}^*$ defined by $u^* \mapsto [\Phi(\cdot), u^*]$ is an isometry. Consequently, the separability of $\mathcal{W}^*$ implies the separability of $\mathcal{K}^*$. Since $\mathcal{K}$ and therefore $\mathcal{K}^*$ are reflexive it follows from [26, Cor. 1.12.12] that $\mathcal{K}$ is separable.

3.2. **Proof of Theorem 2.2.** Let us first demonstrate the equivalence between the separability of the RKBS and Fortet's condition (2.1). Then we will demonstrate the equivalence between Fortet's condition and the separability of the pseudometric space (X, d_Φ). Let us begin with "if". To that end, let us show that condition (2.1) implies that $\Phi(X)$ is separable. Indeed, fix $\varepsilon > 0$ and for each $\frac{\varepsilon}{2^k}, k \in \mathbb{N}$, let $B^k_j, j \in \mathbb{N}$, denote the corresponding partition and let $x^k_j \in B^k_j$ denote a selection. Then the set $\Phi(x^k_j), k \in \mathcal{N}, j \in \mathcal{N}$, is countable, and it is easy to show that it is dense in $\Phi(X)$. That is, $\Phi(X)$ is separable, and the separability of $\mathcal{K}$ follows from Lemma 2.1. Now for the "only if", suppose that $\mathcal{K}$ is separable. Then the canonical feature space $\mathcal{W} := \mathcal{K}$ is separable, and since $\mathcal{K}$ is metric, by e.g. [39, Thm. 16.8], it is second countable. Therefore, since second countability is inherited by subspaces (see e.g. [39, Thm. 16.2]), it follows for the corresponding canonical feature map $\Phi : X \to \mathcal{K}$ that $\Phi(X) \subset \mathcal{K}$ is second countable, and therefore, by e.g. [39, Thm. 16.9], it is separable. Therefore there exists a countable dense set $\Phi(x_j) \in \Phi(X), j \in \mathbb{N}$. Therefore, if for each $\varepsilon > 0$ and for each $j \in \mathbb{N}$ we define $B_j = \{x \in X : \|\Phi(x_j) - \Phi(x)\|_{\mathcal{K}} < \frac{\varepsilon}{2}\}$, it follows that $\bigcup_{j\in\mathbb{N}} B_j = X$ and $\|\Phi(x_1) - \Phi(x_2)\|_{\mathcal{K}} < \varepsilon$ for all $x_1, x_2 \in B_j$. Therefore, we have established the equivalence between the separability of the RKBS and Fortet's condition (2.1).

Now let us demonstrate the equivalence between Fortet's condition and the separability of the pseudometric space (X, d_Φ). To that end, suppose that the pseudometric space (X, d_Φ) is separable. Then Willard [39, Thm. 16.11] asserts that in a pseudometric space, the conditions of being Lindelöf, second countable, and separable are equivalent. Therefore (X, d_Φ) is Lindelöf in that every open cover has a countable subcover. For $x \in X$, let $B_\Phi(x, \epsilon) := \{x' \in X : d_\Phi(x, x') < \epsilon\}$ denote the open ball about x, and for each $\epsilon > 0$ consider the open cover $\{B_\Phi(x, \frac{\epsilon}{2}), x \in X\}$. Then since (X, d_Φ) is Lindelöf it follows that there exists a countable subcover $\{B_\Phi(x, \frac{\epsilon}{2}), x \in X_0\}$ where X_0 is countable. This cover satisfies Fortet's condition (2.1) for the value ϵ, and since ϵ was arbitrary it follows that the map $\Phi : X \to \mathcal{W}$ satisfies Fortet's condition (2.1). In the other direction, suppose that the map $\Phi : X \to \mathcal{W}$ satisfies Fortet's condition (2.1). Fix $\varepsilon > 0$ and for each $\frac{\varepsilon}{2^k}, k \in \mathbb{N}$, let $B^k_j, j \in \mathbb{N}$, denote the corresponding partition and let $x^k_j \in B^k_j$ denote a selection. Then the set $\Phi(x^k_j), k \in \mathcal{N}, j \in \mathcal{N}$, is countable, and it is easy to show that it is dense in $\Phi(X)$. That is, for $x \in X$, the countable set $\{\Phi(x^k_j), k \in \mathcal{N}, j \in \mathcal{N}\}$ comes arbitrarily close to $\Phi(x)$. It follows that the countable set $\{x^k_j, k \in \mathcal{N}, j \in \mathcal{N}\}$ comes arbitrarily close to x in the pseudometric d_Φ. Consequently (X, d_Φ) is separable.

3.3. **Proof of Lemma 2.3.** The case when X is an analytic subset of a Polish space follows directly from Srivastava [31, Thm. 4.3.8]. When X is separable absolute Borel, it follows from Stone's theorem [37, Thm. 16, pg. 32] that when Y is a metric space and $\Phi : X \to Y$ is a measurable bijection, the image Y is separable. However, when Φ is not surjective, since $\Phi(X) \subset Y$ is a metric space, the assertion that the metric subspace $\Phi(X) \subset Y$ is separable follows, assuming that Φ is a measurable injection. Moreover, injectivity is also unnecessary. To see this, extend to the injective map $\hat{\Phi} : X \to X \times Y$ defined by $\hat{\Phi}(x) := \big(x, \Phi(x)\big)$. Then it follows from Hansell's [21, Thm. 1] generalization of Kuratowski [24, Thm. 1, Sec. 31, VI] to the non-separable case that $\hat{\Phi}$ is measurable. To see how this is obtained, since X is assumed to be separable and metrizable, it is second countable (see e.g. [39, Thm. 16.11]), so that it has a countable base $\{G_n, n \in \mathbb{N}\}$ of open sets generating its topology. Let $W \subset X \times Y$ be open and define

$$V_n = \bigcup \{V : V \text{ open}, G_n \times V \subset W\}.$$

Then

$$W = \bigcup_{n \in \mathbb{N}} G_n \times V_n,$$

and therefore

$$\hat{\Phi}^{-1}(W) = \bigcup_{n \in \mathbb{N}} G_n \cap \Phi^{-1}(V_n).$$

Since G_n and V_n are open and therefore measurable and Φ is measurable it follows that $\hat{\Phi}^{-1}(W)$ is measurable. Consequently, since the open sets generate the Borel σ-algebra, it follows that $\hat{\Phi}$ is Borel measurable. Moreover, since $\hat{\Phi}$ is injective the above discussion shows that $\hat{\Phi}(X) \subset X \times Y$ is separable. Since $\Phi(X) = P_Y \hat{\Phi}(X)$ where P_Y is the projection onto the second component and P_Y is continuous, and separability is preserved under continuous maps (see e.g. [39, Thm. 16.4]), it follows that $\Phi(X) \subset Y$ is separable.

3.4. **Proof of Theorem 2.4.** Since the feature space is metric, Lemma 2.3 implies that the image $\Phi(X)$ is separable for any measurable feature map Φ. The assertion then follows from Lemma 2.1.

Acknowledgments

The authors would like to thank the referees for comments and suggestions which substantially improved both the content and the presentation of this work. The authors gratefully acknowledge the support of the Air Force Office of Scientific Research under award number FA9550-12-1-0389 (Scientific Computation of Optimal Statistical Estimators) and AFOSR/DARPA EQUiPS grant number FA9550-16-1-0054 (Computational Information Games).

References

[1] Charalambos D. Aliprantis and Kim C. Border, *Infinite dimensional analysis: A hitchhiker's guide*, 3rd ed., Springer, Berlin, 2006. MR2378491

[2] Alain Berlinet and Christine Thomas-Agnan, *Reproducing kernel Hilbert spaces in probability and statistics*, with a preface by Persi Diaconis, Kluwer Academic Publishers, Boston, MA, 2004. With a preface by Persi Diaconis. MR2239907

[3] Stamatis Cambanis, *The measurability of a stochastic process of second order and its linear space*, Proc. Amer. Math. Soc. **47** (1975), 467–475. MR0356206

[4] S. Canu, X. Mary, and A. Rakotomamonjy, *Functional learning through kernels*, Advances in Learning Theory: Methods, Models and Applications, NATO Science Series, vol. 190, IOS Press, 2003, pp. 89–110.

[5] Jens Gerlach Christensen, *Sampling in reproducing kernel Banach spaces on Lie groups*, J. Approx. Theory **164** (2012), no. 1, 179–203, DOI 10.1016/j.jat.2011.10.002. MR2855776

[6] Andreas Christmann and Ingo Steinwart, *Consistency of kernel-based quantile regression*, Appl. Stoch. Models Bus. Ind. **24** (2008), no. 2, 171–183, DOI 10.1002/asmb.700. MR2406112

[7] Andreas Christmann and Ingo Steinwart, *Universal kernels on non-standard input spaces*, Advances in Neural Information Processing Systems 23, NIPS 2010, pp. 406–414.

[8] Andreas Christmann, Arnout Van Messem, and Ingo Steinwart, *On consistency and robustness properties of support vector machines for heavy-tailed distributions*, Stat. Interface **2** (2009), no. 3, 311–327, DOI 10.4310/SII.2009.v2.n3.a5. MR2540089

[9] Ernesto De Vito, Lorenzo Rosasco, and Alessandro Toigo, *Learning sets with separating kernels*, Appl. Comput. Harmon. Anal. **37** (2014), no. 2, 185–217, DOI 10.1016/j.acha.2013.11.003. MR3223462

[10] Ernesto De Vito, Veronica Umanità, and Silvia Villa, *An extension of Mercer theorem to matrix-valued measurable kernels*, Appl. Comput. Harmon. Anal. **34** (2013), no. 3, 339–351, DOI 10.1016/j.acha.2012.06.001. MR3027907

[11] Claude Dellacherie and Paul-André Meyer, *Probabilités et potentiel* (French). Chapitres I à IV; Édition entièrement refondue; Publications de l'Institut de Mathématique de l'Université de Strasbourg, No. XV; Actualités Scientifiques et Industrielles, No. 1372, Hermann, Paris, 1975. MR0488194

[12] R. M. Dudley, *Real analysis and probability*, Cambridge Studies in Advanced Mathematics, vol. 74, Cambridge University Press, Cambridge, 2002. Revised reprint of the 1989 original. MR1932358

[13] Gregory E. Fasshauer, Fred J. Hickernell, and Qi Ye, *Solving support vector machines in reproducing kernel Banach spaces with positive definite functions*, Appl. Comput. Harmon. Anal. **38** (2015), no. 1, 115–139, DOI 10.1016/j.acha.2014.03.007. MR3273290

[14] R. Fortet, *Espaces à noyau reproduisant et lois de probabilités des fonctions aléatoires* (French), Ann. Inst. H. Poincaré Sect. B (N.S.) **9** (1973), 41–58. MR0345193

[15] Zdeněk Frolík, *On bianalytic spaces* (English, with Russian summary), Czechoslovak Math. J. **13 (88)** (1963), 561–573. MR0162216

[16] Zdeněk Frolík, *Absolute Borel and Souslin sets*, Pacific J. Math. **32** (1970), 663–683. MR0265171

[17] K. Fukumizu, G. R. Lanckriet, and B. K. Sriperumbudur, *Learning in Hilbert vs. Banach spaces: A measure embedding viewpoint*, Advances in Neural Information Processing Systems 24, NIPS 2011, pp. 1773–1781.

[18] R. Hable and A. Christmann, *Estimation of scale functions to model heteroscedasticity by regularised kernel-based quantile methods*, J. Nonparametr. Stat. **26** (2014), no. 2, 219–239, DOI 10.1080/10485252.2013.875547. MR3195981

[19] Paul Richard Halmos, *A Hilbert space problem book*, 2nd ed., Encyclopedia of Mathematics and its Applications, 17, Graduate Texts in Mathematics, vol. 19, Springer-Verlag, New York-Berlin, 1982. MR675952

[20] Deguang Han, M. Zuhair Nashed, and Qiyu Sun, *Sampling expansions in reproducing kernel Hilbert and Banach spaces*, Numer. Funct. Anal. Optim. **30** (2009), no. 9-10, 971–987, DOI 10.1080/01630560903408606. MR2589760

[21] R. W. Hansell, *Sums, products and continuity of Borel maps in nonseparable metric spaces*, Proc. Amer. Math. Soc. **104** (1988), no. 2, 465–471, DOI 10.2307/2046996. MR962814

[22] M. Hein and O. Bousquet, *Kernels, associated structures and generalizations*, Max-Planck-Institut fuer biologische Kybernetik, Technical Report (2004).

[23] Alexander S. Kechris, *Classical descriptive set theory*, Graduate Texts in Mathematics, vol. 156, Springer-Verlag, New York, 1995. MR1321597

[24] K. Kuratowski, *Topology. Vol. I*, New edition, revised and augmented. Translated from the French by J. Jaworowski, Academic Press, New York-London; Państwowe Wydawnictwo Naukowe, Warsaw, 1966. MR0217751

[25] Milan N. Lukić and Jay H. Beder, *Stochastic processes with sample paths in reproducing kernel Hilbert spaces*, Trans. Amer. Math. Soc. **353** (2001), no. 10, 3945–3969, DOI 10.1090/S0002-9947-01-02852-5. MR1837215

[26] Robert E. Megginson, *An introduction to Banach space theory*, Graduate Texts in Mathematics, vol. 183, Springer-Verlag, New York, 1998. MR1650235

[27] D. Montgomery, *Non-separable metric spaces*, Fundamenta Mathematicae **1** (1935), no. 25, 527–533.

[28] M. Zuhair Nashed and Gilbert G. Walter, *General sampling theorems for functions in reproducing kernel Hilbert spaces*, Math. Control Signals Systems **4** (1991), no. 4, 363–390, DOI 10.1007/BF02570568. MR1128261

[29] S. Simpson, *BQO theory and Fraïssé's conjecture*, Recursive Aspects of Descriptive Set Theory, Oxford University Press, Oxford (1985). MR786122

[30] Guohui Song, Haizhang Zhang, and Fred J. Hickernell, *Reproducing kernel Banach spaces with the ℓ^1 norm*, Appl. Comput. Harmon. Anal. **34** (2013), no. 1, 96–116, DOI 10.1016/j.acha.2012.03.009. MR2981335

[31] S. M. Srivastava, *A course on Borel sets*, Graduate Texts in Mathematics, vol. 180, Springer-Verlag, New York, 1998. MR1619545

[32] I. Steinwart, *Convergence types and rates in generic Karhunen-Loève expansions with applications to sample path properties*, arXiv:1403.1040 (2014).

[33] Ingo Steinwart and Andreas Christmann, *Support vector machines*, Information Science and Statistics, Springer, New York, 2008. MR2450103

[34] Ingo Steinwart and Andreas Christmann, *Sparsity of SVMs that use the ϵ-insensitive loss*, Advances in Neural Information Processing Systems 22, NIPS 2009, pp. 1569–1576.

[35] Ingo Steinwart and Andreas Christmann, *Estimating conditional quantiles with the help of the pinball loss*, Bernoulli **17** (2011), no. 1, 211–225, DOI 10.3150/10-BEJ267. MR2797989

[36] Ingo Steinwart and Clint Scovel, *Mercer's theorem on general domains: on the interaction between measures, kernels, and RKHSs*, Constr. Approx. **35** (2012), no. 3, 363–417, DOI 10.1007/s00365-012-9153-3. MR2914365

[37] A. H. Stone, *Non-separable Borel sets*, Rozprawy Mat. **28** (1962), 41 pp. MR0152457

[38] Vladimir Vovk, *On-line regression competitive with reproducing kernel Hilbert spaces (extended abstract)*, Theory and applications of models of computation, Lecture Notes in Comput. Sci., vol. 3959, Springer, Berlin, 2006, pp. 452–463, DOI 10.1007/11750321_43. MR2277266

[39] Stephen Willard, *General topology*, Addison-Wesley Publishing Co., Reading, Mass.-London-Don Mills, Ont., 1970. MR0264581

[40] Haizhang Zhang, Yuesheng Xu, and Jun Zhang, *Reproducing kernel Banach spaces for machine learning*, J. Mach. Learn. Res. **10** (2009), 2741–2775, DOI 10.1109/IJCNN.2009.5179093. MR2579912

[41] Haizhang Zhang and Jun Zhang, *Frames, Riesz bases, and sampling expansions in Banach spaces via semi-inner products*, Appl. Comput. Harmon. Anal. **31** (2011), no. 1, 1–25, DOI 10.1016/j.acha.2010.09.007. MR2795872

[42] Haizhang Zhang and Jun Zhang, *Regularized learning in Banach spaces as an optimization problem: representer theorems*, J. Global Optim. **54** (2012), no. 2, 235–250, DOI 10.1007/s10898-010-9575-z. MR2979626

Department of Computing and Mathematical Sciences, California Institute of Technology, MC 9-94, Pasadena, California 91125

E-mail address: owhadi@caltech.edu

Department of Computing and Mathematical Sciences, California Institute of Technology, MC 9-94, Pasadena, California 91125

E-mail address: clintscovel@gmail.com