

DEEP COMPUTATIONS AND NIP

EDUARDO DUEÑEZ¹ JOSÉ IOVINO¹ TONATIUH MATOS-WIEDERHOLD²
 LUCIANO SALVETTI² FRANKLIN D. TALL²

¹Department of Mathematics, University of Texas at San Antonio

²Department of Mathematics, University of Toronto

ABSTRACT. This paper revisits and extends a bridge between functional analysis and model theory, emphasizing its relevance to the theoretical foundations of machine learning. We show that the compactness behavior of families of Baire class 1 functions mirrors the learnability conditions in the sense of *Probably Approximately Correct* (PAC) learning, and that the failure of compactness corresponds to the presence of infinite *Vapnik-Chervonenkis* (VC) dimension. From this perspective, Rosenthal compacta emerge as the natural topological counterpart of PAC-learnable concept classes, while NIP vs. IP structures capture the precise boundary between analytical regularity and combinatorial intractability. These parallels suggest a unified framework linking compactness, definability, and learnability, exemplifying how the topology of function spaces encodes the algorithmic and epistemic limits of prediction.

1. INTRODUCTION

Suppose that A is a subset of the real line \mathbb{R} and that \bar{A} is its *closure*. It is a well-known fact that any point of closure of A , say $x \in \bar{A}$, can be *approximated* by points inside of A , in the sense that a sequence $\{x_n\}_{n \in \mathbb{N}} \subseteq A$ must exist with the property that $\lim_{n \rightarrow \infty} x_n = x$. For most applications we wish to approximate objects more complicated than points, such as functions.

Suppose we wish to build a neural network that decides, given an 8 by 8 black-and-white image of a hand-written scribble, what single decimal digit the scribble represents. Maybe there exists f , a function representing an optimal solution to this classifier. Thus if X is the set of all (possible) images, then for $I \in X$, $f(I) \in \{0, 1, 2, \dots, 9\}$ is the “best” (or “good enough” for whatever deployment is needed) possible guess. Training the neural network involves approximating f until its guesses are within an acceptable error range. In general, f might be a function defined on a more complicated topological space X .

Often computers’ viable operations are restricted (addition, subtraction, multiplication, division, etc.) and so we want to approximate a complicated function using simple functions (like polynomials). The problem is that, in contrast with mere points, functions in the closure of a set of functions need not be approximable (meaning the pointwise limit of a sequence of functions) by functions in the set.

Functions that are the pointwise limit of continuous functions are *Baire class 1 functions*, and the set of all of these is denoted by $B_1(X)$. Notice that these are not necessarily continuous themselves! A set of Baire class 1 functions, A , will be relatively compact if its closure consists of just Baire class 1 functions (we delay the formal definition of *relatively compact* until Section 2, but the fact mentioned here is sufficient). The Bourgain-Fremlin-Talagrand (BFT) theorem reveals a precise correspondence between relative compactness in $B_1(X)$ and the model-theoretic

32 notion of *Non-Independence Property* (NIP). This was realized by Pierre Simon in
 33 [Sim15b].

34 Simon’s insight was to view definable families of functions as sets of real-valued
 35 functions on type spaces and to interpret relative compactness in $B_1(X)$ as a form
 36 of “tame behavior” under ultrafilter limits. From this perspective, NIP theories are
 37 those whose definable families behave like relatively compact sets of Baire class 1
 38 functions, avoiding the wild, $\beta\mathbb{N}$ -like configurations that witness instability. This
 39 observation opened a new bridge between analysis and logic: topological compact-
 40 ness corresponds to the absence of combinatorial independence. Simon’s later de-
 41 velopments connected these ideas to *Keisler measures* and *empirical averages*, al-
 42 lowing tools from functional analysis to be used to study learnability and definable
 43 types. This reinterpretation of model-theoretic tameness through the lens of the
 44 BFT theorem has made NIP a central notion not only in stability theory but also
 45 in contemporary connections with learning theory and ergodic analysis.

Historically, the notion of NIP arises from Shelah’s foundational work on the
 classification theory of models. In his seminal book *Unstable Theories* [She78],
 Shelah introduced the independence property as a key dividing line within unstable
 structures, identifying the class of *stable* theories inside those in which this property
 fails. Fix a first-order formula $\varphi(x, y)$ in a language L and a model M of an L -
 theory T . We say that $\varphi(x, y)$ has the *independence property* (IP) in M if there is
 a sequence $(c_i)_{i \in \mathbb{N}} \subseteq M^{|x|}$ such that for every $S \subseteq \mathbb{N}$ there is $a_S \in M^{|y|}$ with

$$\forall i \in \mathbb{N}, \quad M \models \varphi(c_i, a_S) \iff i \in S.$$

46 The formula $\varphi(x, y)$ has the IP if it does so in some model M , and the formula
 47 has the *non-independence property* (NIP) if it does not have the IP. The latter
 48 notion of NIP generalizes stability by forbidding the full combinatorial indepen-
 49 dence pattern while allowing certain controlled forms of instability. Thus, Simon’s
 50 interpretation of the BFT theorem can be viewed as placing Shelah’s dividing line
 51 into a topological-analytic framework, connecting the earliest notions of stability
 52 to compactness phenomena in spaces of Baire class 1 functions.

53 One of the most important innovations in Machine Learning is the mathemati-
 54 cal notion, introduced by Turing Awardee Leslie Valiant in the 1980s, of ‘probably
 55 approximately correct learning’, or PAC-learning for short [BD19]. We give a stan-
 56 dard but short overview of these concepts in the context that is relevant to this
 57 work.

58 Consider the following important idea in data classification. Suppose that A is
 59 a set and that \mathcal{C} is a collection of sets. We say that \mathcal{C} *shatters* A if every subset
 60 of A is of the form $C \cap A$ for some $C \in \mathcal{C}$. For a classical geometric example, if
 61 A is the set of four points on the Euclidean plane of the form $(\pm 1, \pm 1)$, then the
 62 collection of all half-planes does not shatter A , the collection of all open balls does
 63 not shatter A , but the collection of all convex sets shatters A . While A need not be
 64 finite, it will usually be assumed to be so in Machine Learning applications. A finer
 65 way to distinguish collections of sets that shatter a given set from those that do
 66 not is by the *Vapnik-Chervonenkis dimension* (*VC-dimension*), which is equal to
 67 the cardinality of the largest finite set shattered by the collection, in case it exists,
 68 or to infinity otherwise.

69 A concrete illustration of these ideas appears when considering threshold clas-
 70 sifiers on the real line. Let \mathcal{H} be the collection of all indicator functions h_t given

by $h_t(x) = 1$ if $x \leq t$ and $h_t(x) = 0$ otherwise. Each h_t is a Baire class 1 function, and the family \mathcal{H} is relatively compact in $B_1(\mathbb{R})$. In model-theoretic terms, \mathcal{H} is NIP, since no configuration of points and thresholds can realize the full independence pattern of a binary matrix. By contrast, the family of parity functions $\{x \mapsto (-1)^{\langle w, x \rangle} : w \in \{0, 1\}^n\}$ on $\{0, 1\}^n$ (here $\langle w, x \rangle$ is the usual vector dot product) has the independence property and fails relative compactness in $B_1(X)$, capturing the analytical meaning of instability. This dichotomy mirrors the behavior of concept classes with finite versus infinite VC dimension in statistical learning theory.

Going back to the model theoretic framework, let

$$\mathcal{F}_\varphi(M) := \{\varphi(M, a) : a \in M^{|y|}\}$$

be the family of subsets of $M^{|x|}$ defined by instances of the formula φ , where $\varphi(M, a)$ is the set of $|x|$ -tuples c in M for which $M \models \varphi(c, a)$. The fundamental theorem of statistical learning states that a binary hypothesis class is PAC-learnable if and only if it has finite VC-dimension, and the subsequent theorem connects the rest of the concepts presented in this section.

Theorem 1.1 (Laskowski). *The formula $\varphi(x, y)$ has the NIP if and only if $\mathcal{F}_\varphi(M)$ has finite VC-dimension.*

For two simple examples of formulas satisfying the NIP, consider first the language $L = \{<\}$ and the model $M = (\mathbb{R}, <)$ of the reals with their usual linear order. Take the formula $\varphi(x, y)$ to mean $x < y$, then $\varphi(M, a) = (-\infty, a)$, and so $\mathcal{F}_\varphi(M)$ is just the set of left open rays. The VC-dimension of this collection is 1, since it can shatter a single point, but no two point set can be shattered since the rays are downwards closed. Now in contrast, the collection of open intervals, given by the formula $\varphi(x; y_1, y_2) := (y_1 < x) \wedge (x < y_2)$, has VC-dimension 2.

In this work, we study the corresponding notions of NIP (and hence PAC-learnability) in the context of Compositional Computation Structures (CCS) introduced in [ADIW24].

2. GENERAL TOPOLOGICAL PRELIMINARIES

In this section we give preliminaries from general topology and function space theory. We include some of the proofs for completeness but a reader familiar with these topics may skip them.

A *Polish space* is a separable and completely metrizable topological space. The most important examples are the reals \mathbb{R} , the Cantor space $2^{\mathbb{N}}$ (the set of all infinite binary sequences, endowed with the product topology), and the Baire space $\mathbb{N}^{\mathbb{N}}$ (the set of all infinite sequences of naturals, also with the product topology). Countable products of Polish spaces are Polish; this includes spaces like $\mathbb{R}^{\mathbb{N}}$, the space of sequences of real numbers. A subspace of a Polish space is itself Polish if and only if it is a G_δ -set, that is, it can be written as the intersection of a countable family of open subsets; in particular, closed subsets and open subsets of Polish spaces are also Polish spaces.

In this work we talk a lot about subspaces, and so there is a pertinent subtlety of the definitions worth mentioning: *completely metrizable space* is not the same as *complete metric space*; for an illustrative example, notice that $(0, 1)$ is homeomorphic to the real line, and thus a Polish space (being Polish is a topological property), but with the metric inherited from the reals, as a subspace, $(0, 1)$ is **not** a complete metric space. In summary, a Polish space has its topology generated by

115 *some* complete metric, but other metrics generating the same topology might not
 116 be. In practice, such as when studying descriptive set theory, one finds that we can
 117 often keep the metric implicit.

118 Given two topological spaces X and Y we denote by $B_1(X, Y)$ the set of all func-
 119 tions $f : X \rightarrow Y$ such that for all open $U \subseteq Y$, $f^{-1}[U]$ is an F_σ subset of X (that
 120 is, a countable union of closed sets); we call these types of functions *Baire class*
 121 *1 functions*. When $Y = \mathbb{R}$ we simply denote this collection by $B_1(X)$. We en-
 122 dow $B_1(X, Y)$ with the topology of pointwise convergence (the topology inherited
 123 from the product topology of Y^X). By $C_p(X, Y)$ we denote the set of all contin-
 124 uous functions $f : X \rightarrow Y$ with the topology of pointwise convergence. Similarly,
 125 $C_p(X) := C_p(X, \mathbb{R})$. A natural question is, how do topological properties of X
 126 translate to $C_p(X)$ and vice versa? These questions, and in general the study of
 127 these spaces, are the concern of C_p -theory, an active field of research in general
 128 topology which was pioneered by A. V. Arhangel'skii and his students in the 1970's
 129 and 1980's. This field has found many exciting applications in model theory and
 130 functional analysis. Good recent surveys on the topics include [HT23] and [Tka11].
 131 We begin with the following:

132 **Fact 2.1.** *If all open subsets of X are F_σ (in particular if X is metrizable), then*
 133 $C_p(X, Y) \subseteq B_1(X, Y)$.

134 The proof of the following fact (due to Baire) can be found in Section 10 of
 135 [Tod97].

136 **Fact 2.2** (Baire). *If X is a complete metric space, then the following are equivalent:*

- 137 (i) f is a Baire class 1 function, that is, $f \in B_1(X)$.
- 138 (ii) f is a pointwise limit of continuous functions.
- 139 (iii) For every closed $F \subseteq X$, the restriction $f|_F$ has a point of continuity.

140 Moreover, if X is Polish and $f \notin B_1(X)$, then there exists countable $D_0, D_1 \subseteq X$ and
 141 reals $a < b$ such that $\overline{D_0} = \overline{D_1}$, $D_0 \subseteq f^{-1}(-\infty, a]$ and $D_1 \subseteq f^{-1}[b, \infty)$.

142 A subset $L \subseteq X$ is *relatively compact* in X if the closure of L in X is compact.
 143 Relatively compact subsets of $B_1(X)$ (for X Polish space) have been objects of
 144 interest to many people working in Analysis and Topological Dynamics. We begin
 145 with the following well-known result. Recall that a set $A \subseteq \mathbb{R}^X$ of real-valued
 146 functions is *pointwise bounded* if for every $x \in X$ there is $M_x > 0$ such that $|f(x)| <$
 147 M_x for all $f \in A$. We include the proof for the reader's convenience:

148 **Lemma 2.3.** *Let X be a Polish space and $A \subseteq B_1(X)$ be pointwise bounded. The*
 149 *following are equivalent:*

- 150 (i) A is relatively compact in $B_1(X)$.
- 151 (ii) A is relatively countably compact in $B_1(X)$, i.e., every countable subset of
 152 A has an accumulation point in $B_1(X)$.
- 153 (iii) $\overline{A} \subseteq B_1(X)$, where \overline{A} denotes the closure in \mathbb{R}^X .

154 *Proof.* By definition, being pointwise bounded means that there is, for each $x \in X$,
 155 $M_x > 0$ such that, for every $f \in A$, $|f(x)| \leq M_x$.

156 (i) \Rightarrow (ii) holds in general.

157 (ii) \Rightarrow (iii) Assume that A is relatively countably compact in $B_1(X)$ and that
 158 $f \in \overline{A} \setminus B_1(X)$. By Fact 2.2, there are countable $D_0, D_1 \subseteq X$ with $\overline{D_0} = \overline{D_1}$, and
 159 $a < b$ such that $D_0 \subseteq f^{-1}(-\infty, a]$ and $D_1 \subseteq f^{-1}[b, \infty)$. We claim that there is a

sequence $\{f_n\}_{n \in \mathbb{N}} \subseteq A$ such that for all $x \in D_0 \cup D_1$, $\lim_{n \rightarrow \infty} f_n(x) = f(x)$. Indeed, use the countability to enumerate $D_0 \cup D_1$ as $\{x_n\}_{n \in \mathbb{N}}$. Then find, for each positive n , $f_n \in A$ with $|f_n(x_i) - f(x_i)| < \frac{1}{n}$ for all $i \leq n$. The claim follows.

By relative countable compactness of A , there is an accumulation point $g \in B_1(X)$ of $\{f_n\}_{n \in \mathbb{N}}$. It is straightforward to show that since f and g agree on $D_0 \cup D_1$, g does not have a point of continuity on the closed set $\overline{D_0} = \overline{D_1}$, which contradicts Fact 2.2.

(iii) \Rightarrow (i) Suppose that $\overline{A} \subseteq B_1(X)$. Then $\overline{A} \cap B_1(X) = \overline{A}$ is a closed subset of $\prod_{x \in X} [-M_x, M_x]$; Tychonoff's theorem states that the product of compact spaces is always compact, and since closed subsets of compact spaces are compact, \overline{A} must be compact, as desired. \square

2.1. From Rosenthal's dichotomy to NIP. The fundamental idea that connects the rich theory here presented to real-valued computations is the concept of an *approximation*. In the reals, points of closure from some subset can always be approximated by points inside the set, via a convergent sequence. For more complicated spaces, such as $C_p(X)$, this fails in a remarkably intriguing way. Let us show an example that is actually the protagonist of a celebrated result. Consider the Cantor space $X = 2^{\mathbb{N}}$ and let $p_n(x) = x(n)$ define a continuous mapping $X \rightarrow \{0, 1\}$. Then one can show (see Chapter 1.1 of [Tod97] for details) that, perhaps surprisingly, the only continuous functions in the closure of $\{p_n\}_{n \in \mathbb{N}}$ are the functions p_n themselves; moreover, none of the subsequences of $\{p_n\}_{n \in \mathbb{N}}$ converge. In some sense, this example is the worst possible scenario for convergence. The topological space obtained from this closure is well-known. Topologists refer to it as the Stone-Ćech compactification of the discrete space of natural numbers, or $\beta\mathbb{N}$ for short, and it is an important object of study in general topology.

Theorem 2.4 (Rosenthal's Dichotomy). *If X is Polish and $\{f_n\} \subseteq C_p(X)$ is pointwise bounded, then either $\{f_n\}_{n \in \mathbb{N}}$ contains a convergent subsequence or a subsequence whose closure (in \mathbb{R}^X) is homeomorphic to $\beta\mathbb{N}$.*

In other words, a pointwise bounded set of continuous functions will either contain a subsequence that converges or a subsequence whose closure is essentially the same as the example mentioned in the previous paragraphs (the worst possible scenario). Note that in the preceding example, the functions are trivially pointwise bounded in \mathbb{R}^X as the functions can only take values 0 and 1.

If we intend to generalize our results from $C_p(X)$ to the bigger space $B_1(X)$, we find a similar dichotomy. Either every point of closure of the set of functions will be a Baire class 1 function, or there is a sequence inside the set that behaves in the worst possible way (which in this context, is the IP!). The theorem is usually not phrased as a dichotomy but rather as an equivalence (with the NIP instead):

Theorem 2.5 (Bourgain-Fremlin-Talagrand, Theorem 4G in [BFT78]). *Let X be a Polish space and $A \subseteq C_p(X)$ be pointwise bounded. The following are equivalent:*

- (i) A is relatively compact in $B_1(X)$, i.e., $\overline{A} \subseteq B_1(X)$.
- (ii) For every $\{f_n\}_{n \in \mathbb{N}} \subseteq A$ and for every $a < b$ there is $I \subseteq \mathbb{N}$ such that

$$\bigcap_{n \in I} f_n^{-1}(-\infty, a] \cap \bigcap_{n \notin I} f_n^{-1}[b, \infty) = \emptyset.$$

Our goal now is to characterize relatively compact subsets of $B_1(X, Y)$ when $Y = \mathbb{R}^{\mathcal{P}}$ with \mathcal{P} countable. Given $P \in \mathcal{P}$ we denote the *projection map* onto the

203 \mathcal{P} -coordinate by $\pi_{\mathcal{P}} : \mathbb{R}^{\mathcal{P}} \rightarrow \mathbb{R}$. From a high-level topological interpretation, the
 204 subsequent lemma states that, in this context, the spaces \mathbb{R} and $\mathbb{R}^{\mathcal{P}}$ are really not
 205 that different, and that if we understand the Baire class 1 functions of one space,
 206 then we also understand the functions of both. In fact, \mathbb{R} and any other Polish
 207 space is embeddable as a closed subspace of $\mathbb{R}^{\mathcal{P}}$.

208 **Lemma 2.6.** *Let X be a Polish space and \mathcal{P} be a countable set. Then, $f \in B_1(X, \mathbb{R}^{\mathcal{P}})$
 209 if and only if $\pi_{\mathcal{P}} \circ f \in B_1(X)$ for all $\mathcal{P} \in \mathcal{P}$.*

Proof. Only one implication needs a proof. Suppose that $\pi_{\mathcal{P}} \circ f \in B_1(X)$ for all $\mathcal{P} \in \mathcal{P}$. Let V be a basic open subset of $\mathbb{R}^{\mathcal{P}}$. That is, there exists a finite $\mathcal{P}' \subseteq \mathcal{P}$ such that $V = \bigcap_{\mathcal{P} \in \mathcal{P}'} \pi_{\mathcal{P}}^{-1}[\mathcal{U}_{\mathcal{P}}]$ where $\mathcal{U}_{\mathcal{P}}$ is open in \mathbb{R} . Finally,

$$f^{-1}[V] = \bigcap_{\mathcal{P} \in \mathcal{P}'} (\pi_{\mathcal{P}} \circ f)^{-1}[\mathcal{U}_{\mathcal{P}}]$$

210 is an F_{σ} set. Since \mathcal{P} is countable, $\mathbb{R}^{\mathcal{P}}$ is second countable so every open set \mathcal{U} in
 211 $\mathbb{R}^{\mathcal{P}}$ is a countable union of basic open sets. Hence, $f^{-1}[\mathcal{U}]$ is F_{σ} . \square

212 Below we consider \mathcal{P} with the discrete topology. For each $f : X \rightarrow \mathbb{R}^{\mathcal{P}}$ denote
 213 $\hat{f}(\mathcal{P}, x) := \pi_{\mathcal{P}} \circ f(x)$ for all $(\mathcal{P}, x) \in \mathcal{P} \times X$. Similarly, for each $g : \mathcal{P} \times X \rightarrow \mathbb{R}$ denote
 214 $\check{g}(x)(\mathcal{P}) := g(\mathcal{P}, x)$. Given $A \subseteq (\mathbb{R}^{\mathcal{P}})^X$, we denote \hat{A} as the set of all \hat{f} such that
 215 $f \in A$.

216 The map $(\mathbb{R}^{\mathcal{P}})^X \rightarrow \mathbb{R}^{\mathcal{P} \times X}$ given by $f \mapsto \hat{f}$ is a homeomorphism and its inverse is
 217 given by $g \mapsto \check{g}$.

218 **Lemma 2.7.** *Let X be a Polish space and \mathcal{P} be countable. Then, $f \in B_1(X, \mathbb{R}^{\mathcal{P}})$ if
 219 and only if $\hat{f} \in B_1(\mathcal{P} \times X)$.*

Proof. (\Rightarrow) Given an open set of reals \mathcal{U} , we have that for every $\mathcal{P} \in \mathcal{P}$, $f^{-1}[\pi_{\mathcal{P}}^{-1}[\mathcal{U}]]$ is F_{σ} by Lemma 2.6. Given that \mathcal{P} is a discrete countable space, we observe that

$$\hat{f}^{-1}[\mathcal{U}] = \bigcup_{\mathcal{P} \in \mathcal{P}} (\{\mathcal{P}\} \times f^{-1}[\pi_{\mathcal{P}}^{-1}[\mathcal{U}]])$$

is also an F_{σ} set. (\Leftarrow) By lemma 2.6 it suffices to show that $\pi_{\mathcal{P}} \circ f \in B_1(X)$ for all $\mathcal{P} \in \mathcal{P}$. Fix an open $\mathcal{U} \subseteq \mathbb{R}$. Write $\hat{f}^{-1}[\mathcal{U}] = \bigcup_{n \in \mathbb{N}} F_n$ where F_n is closed in $\mathcal{P} \times X$. Then,

$$(\pi_{\mathcal{P}} \circ f)^{-1}[\mathcal{U}] = \bigcup_{n \in \mathbb{N}} \{x \in X : (\mathcal{P}, x) \in F_n\}$$

220 which is F_{σ} . \square

221 We now direct our attention to a notion of the NIP that is more general than
 222 the one from the introduction. It can be interpreted as a sort of continuous version
 223 of the one presented in the preceding section.

Definition 2.8. We say that $A \subseteq \mathbb{R}^X$ has the *Non-Independence Property* (NIP) if and only if for every $\{f_n\}_{n \in \mathbb{N}} \subseteq A$ and for every $a < b$ there are finite disjoint sets $E, F \subseteq \mathbb{N}$ such that

$$\bigcap_{n \in E} f_n^{-1}(-\infty, a] \cap \bigcap_{n \in F} f_n^{-1}[b, \infty) = \emptyset.$$

Note that if X is compact and $A \subseteq C_p(X)$, then A has the NIP if and only if for every $\{f_n\}_{n \in \mathbb{N}} \subseteq A$ and for every $a < b$ there is $I \subseteq \mathbb{N}$ such that

$$\bigcap_{n \in I} f_n^{-1}(-\infty, a] \cap \bigcap_{n \notin I} f_n^{-1}[b, \infty) = \emptyset.$$

224 Given $A \subseteq Y^X$ and $K \subseteq X$ we write $A|_K := \{f|_K : f \in A\}$, i.e., the set of
 225 all restrictions of functions in A to K . The following Theorem is a slightly more
 226 general version of Theorem 2.5.

227 **Theorem 2.9.** *Assume that \mathcal{P} is countable, X is a Polish space, and $A \subseteq C_p(X, \mathbb{R}^{\mathcal{P}})$
 228 is such that $\pi_P \circ A$ is pointwise bounded for all $P \in \mathcal{P}$. The following are equivalent
 229 for every compact $K \subseteq X$:*

- 230 (1) $\overline{A|_K} \subseteq B_1(K, \mathbb{R}^{\mathcal{P}})$.
 231 (2) $\pi_P \circ A|_K$ has the NIP for every $P \in \mathcal{P}$.

Proof. (1) \Rightarrow (2). Let $P \in \mathcal{P}$. Fix $\{f_n\}_{n \in \mathbb{N}} \subseteq A$ and $a < b$. By (1) we have that $\overline{A|_K} \subseteq B_1(K, \mathbb{R}^{\mathcal{P}})$. Applying the homeomorphism $f \mapsto \hat{f}$ and using lemma 2.7 we get $\widehat{\overline{A|_K}} \subseteq B_1(\mathcal{P} \times K)$. By Theorem 2.5, there is $I \subseteq \mathbb{N}$ such that

$$(\mathcal{P} \times K) \cap \bigcap_{n \in I} \hat{f}_n^{-1}(-\infty, a] \cap \bigcap_{n \notin I} \hat{f}_n^{-1}[b, \infty) = \emptyset$$

Hence,

$$K \cap \bigcap_{n \in I} (\pi_P \circ f_n)^{-1}(-\infty, a] \cap \bigcap_{n \notin I} (\pi_P \circ f_n)^{-1}[b, \infty) = \emptyset$$

By compactness, there are finite $E \subseteq I$ and $F \subseteq \mathbb{N} \setminus I$ such that

$$K \cap \bigcap_{n \in E} (\pi_P \circ f_n)^{-1}(-\infty, a] \cap \bigcap_{n \in F} (\pi_P \circ f_n)^{-1}[b, \infty) = \emptyset$$

232 Thus, $\pi_P \circ A|_L$ has the NIP.

233 (2) \Rightarrow (1) Fix $f \in \overline{A|_K}$. By lemma 2.6 it suffices to show that $\pi_P \circ f \in B_1(K)$
 234 for all $P \in \mathcal{P}$. By (2), $\pi_P \circ A|_K$ has the NIP. Hence, by Theorem 2.5 we have
 235 $\overline{\pi_P \circ A|_K} \subseteq B_1(K)$. But then $\pi_P \circ f \in \overline{\pi_P \circ A|_K} \subseteq B_1(K)$. \square

236 Lastly, a simple but significant result that helps understand the operation of
 237 restricting a set of functions to a specific subspace of the domain space X , of course
 238 in the context of the NIP, is that we may always assume that said subspace is
 239 closed. Concretely, whether we take its closure or not has no effect on the NIP:

240 **Lemma 2.10.** *Assume that X is Hausdorff and that $A \subseteq C_p(X)$. The following
 241 are equivalent for every $L \subseteq X$:*

- 242 (i) $A|_L$ has the NIP.
 243 (ii) $A|_{\overline{L}}$ has the NIP.

Proof. It suffices to show that (i) \Rightarrow (ii). Suppose that (ii) does not hold, i.e., that there are $\{f_n\}_{n \in \mathbb{N}} \subseteq A$ and $a < b$ such that for all finite disjoint $E, F \subseteq \mathbb{N}$:

$$\overline{L} \cap \bigcap_{n \in E} f_n^{-1}(-\infty, a] \cap \bigcap_{n \in F} f_n^{-1}[b, \infty) \neq \emptyset.$$

Pick $a' < b'$ such that $a < a' < b' < b$. Then, for any finite disjoint $E, F \subseteq \mathbb{N}$ we can choose

$$x \in \overline{L} \cap \bigcap_{n \in E} f_n^{-1}(-\infty, a') \cap \bigcap_{n \in F} f_n^{-1}(b', \infty)$$

By definition of closure:

$$L \cap \bigcap_{n \in E} f_n^{-1}(-\infty, a'] \cap \bigcap_{n \in F} f_n^{-1}[b', \infty) \neq \emptyset.$$

244 This contradicts (i). □

245 3. NIP IN THE CONTEXT OF COMPOSITIONAL COMPUTATION STRUCTURES

246 In this section, we study what the NIP tell us in the context of deep compu-
 247 tations as defined in [ADIW24]. We say a structure (L, \mathcal{P}, Γ) is a *Compositional*
 248 *Computation Structure* (CCS) if $L \subseteq \mathbb{R}^{\mathcal{P}}$ is a subspace of $\mathbb{R}^{\mathcal{P}}$, with the pointwise
 249 convergence topology, and $\Gamma \subseteq L^L$ is a semigroup under composition. The motiva-
 250 tion for CCS comes from (continuous) model theory, where \mathcal{P} is a fixed collection
 251 of predicates and L is a (real-valued) structure. Every point in L is identified with
 252 its “type”, which is the tuple of all values the point takes on the predicates from
 253 \mathcal{P} , i.e., an element of $\mathbb{R}^{\mathcal{P}}$. In this context, elements of \mathcal{P} are called *features*. In the
 254 discrete model theory framework, one views the space of complete-types as a sort of
 255 compactification of the structure L . In this context, we don’t want to consider only
 256 points in L (realized types) but in its closure \bar{L} (possibly unrealized types). The
 257 problem is that the closure \bar{L} is not necessarily compact, an assumption that turns
 258 out to be very useful in the context of continuous model theory. To bypass this
 259 problem in a framework for deep computations, Alva, Dueñez, Iovino and Walton
 260 introduced in [ADIW24] the concept of *shards*, which essentially consists in cover-
 261 ing (a large fragment) of the space \bar{L} by compact, and hence pointwise-bounded,
 262 subspaces (shards). We shall give the formal definition next.

263 A *sizer* is a tuple $\mathbf{r}_\bullet = (r_p)_{p \in \mathcal{P}}$ of positive real numbers indexed by \mathcal{P} . Given a
 264 sizer \mathbf{r}_\bullet , we define the \mathbf{r}_\bullet -*shard* as:

$$L[\mathbf{r}_\bullet] = L \cap \prod_{p \in \mathcal{P}} [-r_p, r_p]$$

265 For an illustrative example, we can frame Newton’s polynomial root approxima-
 266 tion method in the context of a CCS (see Example 5.6 of [ADIW24] for details) as
 267 follows. Begin by considering the extended complex numbers $\hat{\mathbb{C}} := \mathbb{C} \cup \{\infty\}$ with
 268 the usual Riemann sphere topology that makes it into a compact space (where
 269 unbounded sequences converge to ∞). In fact, not only is this space compact
 270 but it is covered by the shard given by the sizer $(1, 1, 1)$ (the unit sphere is con-
 271 tained in the cube $[-1, 1]^3$). The space $\hat{\mathbb{C}}$ is homeomorphic to the usual unit
 272 sphere $S^2 := \{(x, y, z) : x^2 + y^2 + z^2 = 1\}$ of \mathbb{R}^3 , by means of the stereographic
 273 projection and its inverse $\hat{\mathbb{C}} \rightarrow S^2$. This function is regarded as a triple of pred-
 274 icates $x, y, z : \hat{\mathbb{C}} \rightarrow [-1, 1]$ where each will map an extended complex number to
 275 its corresponding real coordinate on the cube $[-1, 1]^3$. Now fix the cubic com-
 276 plex polynomial $p(s) := s^3 - 1$, and consider the map which performs one step
 277 in Newton’s method at a particular (extended) complex number s , for finding
 278 a root of p , $\gamma_p : \hat{\mathbb{C}} \rightarrow \hat{\mathbb{C}}$. The explicit inner workings of γ_p are irrelevant for
 279 this example, except for the fact that it is a continuous mapping. It follows that
 280 $(S^2, \{x, y, z\}, \{\gamma_p^k : k \in \mathbb{N}\})$ is a CCS. The idea is that repeated applications of
 281 $\gamma_p(s), \gamma_p \circ \gamma_p(s), \gamma_p \circ \gamma_p \circ \gamma_p(s), \dots$ would approximate a root of p provided s was
 282 a good enough initial guess.

283 The \mathbf{r}_\bullet -type-shard is defined as $\mathcal{L}[\mathbf{r}_\bullet] = \overline{\mathcal{L}[\mathbf{r}_\bullet]}$ and \mathcal{L}_{sh} is the union of all type-
 284 shards. Notice that \mathcal{L}_{sh} is not necessarily equal to $\mathcal{L} = \overline{\mathcal{L}}$, unless \mathcal{P} is countable
 285 (see [ADIW24]). A *transition* is a map $f : \mathcal{L} \rightarrow \mathcal{L}$, in particular, every element
 286 in the semigroup Γ is a transition (these are called *realized computations*). In
 287 practice, one would like to work with “definable” computations, i.e., ones that can
 288 be described by a computer. In this topological framework, being continuous is an
 289 expected requirement. However, as in the case of complete-types in model theory,
 290 we will work with “unrealized computations”, i.e., maps $f : \mathcal{L}_{\text{sh}} \rightarrow \mathcal{L}_{\text{sh}}$. Note that
 291 continuity of a computation does not imply that it can be continuously extended
 292 to \mathcal{L}_{sh} .

293 Suppose that a transition map $f : \mathcal{L} \rightarrow \mathcal{L}$ can be extended continuously to
 294 a map $\mathcal{L} \rightarrow \mathcal{L}$. Then, the Stone-Weierstrass theorem implies that the feature
 295 $\pi_{\mathbf{P}} \circ f$ (here \mathbf{P} is a fixed predicate, and the feature is hence continuous) can be
 296 uniformly approximated by polynomials on the compact set $\mathcal{L}[\mathbf{r}_\bullet]$. Theorem 2.2 in
 297 [ADIW24] formalizes the converse of this fact, in the sense that transitions maps
 298 that are not continuously extendable in this fashion cannot be obtained from simple
 299 constructions involving predicates. Under this framework, the features $\pi_{\mathbf{P}} \circ f$ of such
 300 transitions f are not approximable by polynomials, and so they are understood as
 301 “non-computable” since, again, we expect the operations computers carry out to be
 302 determined by elementary algebra corresponding to polynomials (namely addition
 303 and multiplication). Therefore it is crucial we assume some extendibility conditions.

304 We say that the CCS $(\mathcal{L}, \mathcal{P}, \Gamma)$ satisfies the *Extendibility Axiom* if for all $\gamma \in \Gamma$,
 305 there is $\tilde{\gamma} : \mathcal{L}_{\text{sh}} \rightarrow \mathcal{L}_{\text{sh}}$ such that for every sizer \mathbf{r}_\bullet , there is an \mathbf{s}_\bullet such that
 306 $\tilde{\gamma}|_{\mathcal{L}[\mathbf{r}_\bullet]} : \mathcal{L}[\mathbf{r}_\bullet] \rightarrow \mathcal{L}[\mathbf{s}_\bullet]$ is continuous. For a deeper discussion about this axiom, we
 307 refer the reader to [ADIW24].

308 A collection \mathbf{R} of sizers is called *exhaustive* if $\mathcal{L}_{\text{sh}} = \bigcup_{\mathbf{r}_\bullet \in \mathbf{R}} \mathcal{L}[\mathbf{r}_\bullet]$. We say that
 309 $\Delta \subseteq \Gamma$ is *\mathbf{R} -confined* if $\gamma|_{\mathcal{L}[\mathbf{r}_\bullet]} : \mathcal{L}[\mathbf{r}_\bullet] \rightarrow \mathcal{L}[\mathbf{r}_\bullet]$ for every $\mathbf{r}_\bullet \in \mathbf{R}$ and $\gamma \in \Delta$. Elements in
 310 Δ are called *real-valued computations* (in this article we will refer to them simply as
 311 *computations*) and elements in $\overline{\Delta} \subseteq \mathcal{L}_{\text{sh}}^{\mathcal{L}}$ are called (real-valued) *deep computations*
 312 or *ultracomputations*. By $\tilde{\Delta}$ we denote the set of all extensions $\tilde{\gamma}$ for $\gamma \in \Delta$. For a
 313 more complete description of this framework, we refer the reader to [ADIW24].

314 **3.1. NIP and Baire-1 definability of deep computations.** Under what con-
 315 ditions are deep computations Baire class 1, and thus well-behaved according to
 316 our framework, on type-shards? The next Theorem says that, again under the
 317 assumption that \mathcal{P} is countable, the space of deep computations is a Rosenthal
 318 compactum (when restricted to shards) if and only if the set of computations has
 319 the NIP on features. Hence, we can import the theory of Rosenthal compacta into
 320 this framework of deep computations.

321 **Theorem 3.1.** *Let $(\mathcal{L}, \mathcal{P}, \Gamma)$ be a CCS satisfying the Extendibility Axiom with \mathcal{P}*
 322 *countable. Let \mathbf{R} be an exhaustive collection of sizers. Let $\Delta \subseteq \Gamma$ be \mathbf{R} -confined. The*
 323 *following are equivalent.*

- 324 (1) $\tilde{\Delta}|_{\mathcal{L}[\mathbf{r}_\bullet]} \subseteq \mathbf{B}_1(\mathcal{L}[\mathbf{r}_\bullet], \mathcal{L}[\mathbf{r}_\bullet])$ for all $\mathbf{r}_\bullet \in \mathbf{R}$.
 (2) $\pi_{\mathbf{P}} \circ \Delta|_{\mathcal{L}[\mathbf{r}_\bullet]}$ has the NIP for all $\mathbf{P} \in \mathcal{P}$ and $\mathbf{r}_\bullet \in \mathbf{R}$, that is, for all $\mathbf{P} \in \mathcal{P}$,
 $\mathbf{r}_\bullet \in \mathbf{R}$, $\mathbf{a} < \mathbf{b}$, $\{\gamma_n\}_{n \in \mathbb{N}} \subseteq \Delta$ there are finite disjoint $\mathbf{E}, \mathbf{F} \subseteq \mathbb{N}$ such that

$$\mathcal{L}[\mathbf{r}_\bullet] \cap \bigcap_{n \in \mathbf{E}} (\pi_{\mathbf{P}} \circ \gamma_n)^{-1}(-\infty, \mathbf{a}] \cap \bigcap_{n \in \mathbf{F}} (\pi_{\mathbf{P}} \circ \gamma_n)^{-1}[\mathbf{b}, \infty) = \emptyset.$$

Moreover, if any (hence all) of the preceding conditions hold, then every deep computation $f \in \bar{\Delta}$ can be extended to a Baire-1 function on shards, i.e., there is $\tilde{f} : \mathcal{L}_{\text{sh}} \rightarrow \mathcal{L}_{\text{sh}}$ such that $\tilde{f}|_{\mathcal{L}[\mathbf{r}_\bullet]} \in B_1(\mathcal{L}[\mathbf{r}_\bullet], \mathcal{L}[\mathbf{r}_\bullet])$ for all $\mathbf{r}_\bullet \in \mathbf{R}$. In particular, on each shard every deep computation is the pointwise limit of a countable sequence of computations.

Proof. Since \mathcal{P} is countable, then $\mathcal{L}[\mathbf{r}_\bullet] \subseteq \mathbb{R}^{\mathcal{P}}$ is Polish. Also, the Extendibility Axiom implies that $\pi_{\mathcal{P}} \circ \tilde{\Delta}|_{\mathcal{L}[\mathbf{r}_\bullet]}$ is a pointwise bounded set of continuous functions for all $\mathbf{P} \in \mathcal{P}$. Hence, Theorem 2.9 and Lemma 2.10 prove the equivalence of (1) and (2). If (1) holds and $f \in \bar{\Delta}$, then write $f = \mathcal{U}\lim_i \gamma_i$ as an ultralimit. Define $\tilde{f} := \mathcal{U}\lim_i \tilde{\gamma}_i$. Hence, for all $\mathbf{r}_\bullet \in \mathbf{R}$ we have $\tilde{f}|_{\mathcal{L}[\mathbf{r}_\bullet]} \in \tilde{\Delta}|_{\mathcal{L}[\mathbf{r}_\bullet]} \subseteq B_1(\mathcal{L}[\mathbf{r}_\bullet], \mathcal{L}[\mathbf{r}_\bullet])$. That every deep computation is a pointwise limit of a countable sequence of computations follows from the fact that for a Polish space X every compact subset of $B_1(X)$ is Fréchet-Urysohn (that is, a space where topological closures coincide with sequential closures, see Theorem 3F in [BFT78] or Theorem 4.1 in [Deb13]). \square

Given a countable set Δ of computations satisfying the NIP on features and shards (condition (2) of Theorem 3.1) we have that $\tilde{\Delta}|_{\mathcal{L}[\mathbf{r}_\bullet]}$ (for a fixed sizer \mathbf{r}_\bullet) is a separable *Rosenthal compactum* (compact subset of $B_1(\mathcal{P} \times \mathcal{L}[\mathbf{r}_\bullet])$). The work of Todorčević ([Tod99]) and Argyros, Dodos, Kanellopoulos ([ADK08]) culminates in a trichotomy theorem for separable Rosenthal Compacta. Inspired by the work of Glasner and Megrelishvili ([GM22]), we are interested to see how this allows us to classify and obtain different levels of PAC-learnability (NIP).

Recall that a topological space X is *hereditarily separable* (HS) if every subspace is separable and that X is *first countable* if every point in X has a countable local basis. Every separable metrizable space is hereditarily separable and it is a result of R. Pol that every hereditarily separable Rosenthal compactum is first countable (see section 10 in [Deb13]). This suggests the following definition:

Definition 3.2. Let (L, \mathcal{P}, Γ) be a CCS satisfying the Extendibility Axiom and \mathbf{R} be an exhaustive collection of sizers. Let $\Delta \subseteq \Gamma$ be an \mathbf{R} -confined countable set of computations satisfying the NIP on shards and features (condition (2) in Theorem 3.1). We say that Δ is:

- (i) NIP₁ if $\tilde{\Delta}|_{\mathcal{L}[\mathbf{r}_\bullet]}$ is first countable for every $\mathbf{r}_\bullet \in \mathbf{R}$.
- (ii) NIP₂ if $\tilde{\Delta}|_{\mathcal{L}[\mathbf{r}_\bullet]}$ is hereditarily separable for every $\mathbf{r}_\bullet \in \mathbf{R}$.
- (iii) NIP₃ if $\tilde{\Delta}|_{\mathcal{L}[\mathbf{r}_\bullet]}$ is metrizable for every $\mathbf{r}_\bullet \in \mathbf{R}$.

Observe that $\text{NIP}_3 \Rightarrow \text{NIP}_2 \Rightarrow \text{NIP}_1 \Rightarrow \text{NIP}$. A natural question that would continue this work is to find examples of CCS that separate these levels of NIP. In [Tod99], Todorčević isolates 3 canonical examples of Rosenthal compacta that witness the failure of the converse implications above. We now present some separable and non-separable examples of Rosenthal compacta:

- (1) *Alexandroff compactification of a discrete space of size continuum.* For each $\mathbf{a} \in 2^{\mathbb{N}}$ consider the map $\delta_{\mathbf{a}} : 2^{\mathbb{N}} \rightarrow \mathbb{R}$ given by $\delta_{\mathbf{a}}(\mathbf{x}) = 1$ if $\mathbf{x} = \mathbf{a}$ and $\delta_{\mathbf{a}}(\mathbf{x}) = 0$ otherwise. Let $A(2^{\mathbb{N}}) = \{\delta_{\mathbf{a}} : \mathbf{a} \in 2^{\mathbb{N}}\} \cup \{0\}$, where 0 is the zero map. Notice that $A(2^{\mathbb{N}})$ is a compact subset of $B_1(2^{\mathbb{N}})$, in fact $\{\delta_{\mathbf{a}} : \mathbf{a} \in 2^{\mathbb{N}}\}$ is a discrete subspace of $B_1(2^{\mathbb{N}})$ and its pointwise closure is precisely $A(2^{\mathbb{N}})$. Hence, this is a Rosenthal compactum which is not first countable. Notice that this space is also not separable.

- 370 (2) *Extended Alexandroff compactification.* For each finite binary sequence $s \in$
 371 $2^{<\mathbb{N}}$, let $v_s : 2^{\mathbb{N}} \rightarrow \mathbb{R}$ be given by $v_s(x) = 1$ if x extends s and $v_s(x) =$
 372 0 otherwise. Let $\hat{A}(2^{\mathbb{N}})$ be the pointwise closure of $\{v_s : s \in 2^{<\mathbb{N}}\}$, i.e.,
 373 $\hat{A}(2^{\mathbb{N}}) = A(2^{\mathbb{N}}) \cup \{v_s : s \in 2^{<\mathbb{N}}\}$. Note that this space is a separable
 374 Rosenthal compactum which is not first countable.
- 375 (3) *Split Cantor.* Let $<$ be the lexicographic order in the space of infinite
 376 binary sequences, i.e., $2^{\mathbb{N}}$. For each $a \in 2^{\mathbb{N}}$ let $f_a^- : 2^{\mathbb{N}} \rightarrow \mathbb{R}$ be given by
 377 $f_a^-(x) = 1$ if $x < a$ and $f_a^-(x) = 0$ otherwise. Let $f_a^+ : 2^{\mathbb{N}} \rightarrow \mathbb{R}$ be given
 378 by $f_a^+(x) = 1$ if $x \leq a$ and $f_a^+(x) = 0$ otherwise. The split Cantor is the
 379 space $S(2^{\mathbb{N}}) = \{f_a^- : a \in 2^{\mathbb{N}}\} \cup \{f_a^+ : a \in 2^{\mathbb{N}}\}$. This is a separable Rosenthal
 380 compactum. One example of a countable dense subset is the set of all f_a^+
 381 and f_a^- where a is an infinite binary sequence that is eventually constant.
 382 Moreover, it is hereditarily separable but it is not metrizable.
- (4) *Alexandroff Duplicate.* Let K be any compact metric space and consider
 the Polish space $X = C(K) \sqcup K$, i.e., the disjoint union of $C(K)$ (with its
 supremum norm topology) and K . For each $a \in K$ define $g_a^0, g_a^1 : X \rightarrow \mathbb{R}$ as
 follows:

$$g_a^0(x) = \begin{cases} x(a), & x \in C(K) \\ 0, & x \in K \end{cases}$$

$$g_a^1(x) = \begin{cases} x(a), & x \in C(K) \\ \delta_a(x), & x \in K \end{cases}$$

383 Let $D(K) = \{g_a^0 : a \in K\} \cup \{g_a^1 : a \in K\}$. Notice that $D(K)$ is a first
 384 countable Rosenthal compactum. It is not separable if K is uncountable.
 385 The interesting case will be when $K = 2^{\mathbb{N}}$.

- (5) *Extended Alexandroff Duplicate of the split Cantor.* For each finite binary
 sequence $t \in 2^{<\mathbb{N}}$ let $a_t \in 2^{\mathbb{N}}$ be the sequence starting with t and ending
 with 0's and let $b_t \in 2^{\mathbb{N}}$ be the sequence starting with t and ending with
 1's. Define $h_t : 2^{\mathbb{N}} \rightarrow \mathbb{R}$ by

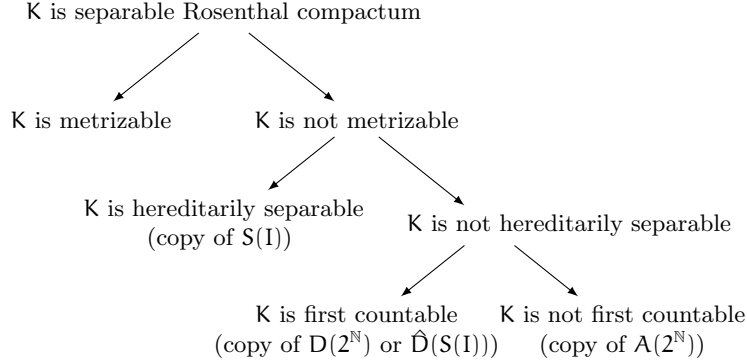
$$h_t(x) = \begin{cases} 0, & x < a_t \\ 1/2, & a_t \leq x \leq b_t \\ 1, & b_t < x \end{cases}$$

386 Let $\hat{D}(S(2^{\mathbb{N}}))$ be the pointwise closure of the set $\{h_t : t \in 2^{<\mathbb{N}}\}$. Hence,
 387 $\hat{D}(S(2^{\mathbb{N}}))$ is a separable first countable Rosenthal compactum which is not
 388 hereditarily separable. In fact, it contains an uncountable discrete subspace
 389 (see Theorem 5 in [Tod99]).

390 **Theorem 3.3** (Todorćević's Trichotomy, [Tod99], Theorem 3 in [ADK08]). *Let K*
 391 *be a separable Rosenthal Compactum.*

- 392 (i) *If K is hereditarily separable but non-metrizable, then $S(2^{\mathbb{N}})$ embeds into K .*
 393 (ii) *If K is first countable but not hereditarily separable, then either $D(2^{\mathbb{N}})$ or*
 394 *$\hat{D}(S(2^{\mathbb{N}}))$ embeds into K .*
 395 (iii) *If K is not first countable, then $A(2^{\mathbb{N}})$ embeds into K .*

396 In other words, we have the following classification:



397

398 Lastly, the definitions provided here for NIP_i ($i = 1, 2, 3$) are topological.

399 **Question 3.4.** Is there a non-topological characterization for NIP_i , $i = 1, 2, 3$?

400 More can be said about the nature of the embeddings in Todorćević's Trichotomy.
 401 Given a separable Rosenthal compactum K , there is typically more than one countable
 402 dense subset of K . We can view a separable Rosenthal compactum as the accumulation
 403 points of a countable family of pointwise bounded real-valued functions.
 404 The choice of the countable families is not important when a bijection between
 405 them can be lifted to a homeomorphism of their closures. To be more precise:

406 **Definition 3.5.** Given a Polish space X , a countable set I and two pointwise
 407 bounded families $\{f_i : i \in I\} \subseteq \mathbb{R}^X$, $\{g_i : i \in I\} \subseteq \mathbb{R}^X$ indexed by I . We say that
 408 $\{f_i : i \in I\}$ and $\{g_i : i \in I\}$ are *equivalent* if and only if the map $f_i \mapsto g_i$ is extended
 409 to a homeomorphism from $\overline{\{f_i : i \in I\}}$ to $\overline{\{g_i : i \in I\}}$.

410 Notice that in the separable examples discussed before ($\hat{A}(2^{\mathbb{N}})$, $S(2^{\mathbb{N}})$ and $\hat{D}(S(2^{\mathbb{N}}))$)
 411 the countable dense subsets are indexed by the binary tree $2^{<\mathbb{N}}$. This choice of index
 412 is useful because the Ramsey theory of perfect subsets of the Cantor space $2^{\mathbb{N}}$
 413 can be imported to analyze the behavior of the accumulation points. Since $2^{<\mathbb{N}}$ is
 414 countable, we can always choose this index for the countable dense subsets. This
 415 is done in [ADK08].

416 **Definition 3.6.** Given a Polish space X and a pointwise bounded family $\{f_t : t \in$
 417 $2^{<\mathbb{N}}\}$. We say that $\{f_t : t \in 2^{<\mathbb{N}}\}$ is *minimal* if and only if for every dyadic subtree
 418 $\{s_t : t \in 2^{<\mathbb{N}}\}$ of $2^{<\mathbb{N}}$, $\{f_{s_t} : t \in 2^{<\mathbb{N}}\}$ is equivalent to $\{f_t : t \in 2^{<\mathbb{N}}\}$.

419 One of the main results in [ADK08] is that there are (up to equivalence) seven
 420 minimal families of Rosenthal compacta and that for every relatively compact $\{f_t :$
 421 $t \in 2^{<\mathbb{N}}\} \subseteq B_1(X)$ there is a dyadic subtree $\{s_t : t \in 2^{<\mathbb{N}}\}$ such that $\{f_{s_t} : t \in 2^{<\mathbb{N}}\}$
 422 is equivalent to one of the minimal families. We shall describe the minimal families
 423 next. We will follow the same notation as in [ADK08]. For any node $t \in 2^{<\mathbb{N}}$, we
 424 denote by $t \smallfrown 0^\infty$ ($t \smallfrown 1^\infty$) the infinite binary sequence starting with t and ending
 425 in 0's (1's). Fix a regular dyadic subtree $R = \{s_t : t \in 2^{<\mathbb{N}}\}$ of $2^{<\mathbb{N}}$ (i.e., a dyadic
 426 subtree such that every level of R is contained in a level of $2^{<\mathbb{N}}$) with the property
 427 that for all $s, s' \in R$, $s \smallfrown 0^\infty \neq s' \smallfrown 0^\infty$ and $s \smallfrown 1^\infty \neq s' \smallfrown 1^\infty$. Given $t \in 2^{<\mathbb{N}}$, let
 428 v_t be the characteristic function of the set $\{x \in 2^{\mathbb{N}} : x \text{ extends } t\}$. Let $<$ be the
 429 lexicographic order in $2^{\mathbb{N}}$. Given $\alpha \in 2^{\mathbb{N}}$, let $f_\alpha^+ : 2^{\mathbb{N}} \rightarrow \{0, 1\}$ be the characteristic
 430 function of $\{x \in 2^{\mathbb{N}} : \alpha \leq x\}$ and let $f_\alpha^- : 2^{\mathbb{N}} \rightarrow \{0, 1\}$ be the characteristic function of

431 $\{x \in 2^{\mathbb{N}} : a < x\}$. Given two maps $f, g : 2^{\mathbb{N}} \rightarrow \mathbb{R}$ we denote by $(f, g) : 2^{\mathbb{N}} \sqcup 2^{\mathbb{N}} \rightarrow \mathbb{R}$
 432 the function which is f on the first copy of $2^{\mathbb{N}}$ and g on the second copy of $2^{\mathbb{N}}$.

- 433 (1) $D_1 = \{\frac{1}{|t|+1}v_t : t \in 2^{<\mathbb{N}}\}$. This is discrete in $\overline{D_1} = A(2^{\mathbb{N}})$.
 434 (2) $D_2 = \{s_t \widehat{0}^\infty : t \in 2^{<\mathbb{N}}\}$. This is discrete in $\overline{D_2} = 2^{\leq \mathbb{N}}$.
 435 (3) $D_3 = \{f_{s_t \widehat{0}^\infty}^+ : t \in 2^{<\mathbb{N}}\}$. This is a discrete in $\overline{D_3} = S(2^{\mathbb{N}})$.
 436 (4) $D_4 = \{f_{s_t \widehat{1}^\infty}^- : t \in 2^{<\mathbb{N}}\}$. This is discrete in $\overline{D_4} = S(2^{\mathbb{N}})$.
 437 (5) $D_5 = \{v_t : t \in 2^{<\mathbb{N}}\}$. This is discrete in $\overline{D_5} = \widehat{A}(2^{\mathbb{N}})$.
 438 (6) $D_6 = \{(v_{s_t}, s_t \widehat{0}^\infty) : t \in 2^{<\mathbb{N}}\}$. This is discrete in $\overline{D_6} = \widehat{D}(2^{\mathbb{N}})$.
 439 (7) $D_7 = \{(v_{s_t}, x_{s_t \widehat{0}^\infty}^+) : t \in 2^{<\mathbb{N}}\}$. This is discrete in $\overline{D_7} = \widehat{D}(S(2^{\mathbb{N}}))$

440 **Theorem 3.7** (Heptacotomy of minimal families, Theorem 2 in [ADK08]). *Let*
 441 *X be Polish. For every relatively compact $\{f_t : t \in 2^{<\mathbb{N}}\} \subseteq B_1(X)$, there exists $i =$*
 442 *$1, 2, \dots, 7$ and a regular dyadic subtree $\{s_t : t \in 2^{<\mathbb{N}}\}$ of $2^{<\mathbb{N}}$ such that $\{f_{s_t} : t \in 2^{<\mathbb{N}}\}$*
 443 *is equivalent to D_i . Moreover, all D_i are minimal and mutually non-equivalent.*

444 **3.2. NIP and definability by universally measurable functions.** We now
 445 turn to the question: what happens when \mathcal{P} is uncountable? Notice that the
 446 countability assumption is crucial in the proof of Theorem 2.9 essentially because it
 447 makes $\mathbb{R}^{\mathcal{P}}$ a Polish space. For the uncountable case, we may lose Baire-1 definability
 448 so we shall replace $B_1(X)$ by a bigger class. Recall that the purpose of studying the
 449 class of Baire-1 functions is that a pointwise limit of continuous functions is not
 450 necessarily continuous. In [BFT78], J. Bourgain, D.H. Fremlin and M. Talagrand
 451 characterized the Non-Independence Property of a set of continuous functions with
 452 various notions of compactness in function spaces containing $C(X)$, such as $B_1(X)$.
 453 In this section we will replace $B_1(X)$ with the larger space $M_r(X)$ of universally
 454 measurable functions. The development of this section is based on Theorem 2F in
 455 [BFT78]. We now give the relevant definitions. Readers with little familiarity with
 456 measure theory can review the appendix for standard definitions appearing in this
 457 subsection.

458 Given a Hausdorff space X and a measurable space (Y, Σ) , we say that $f : X \rightarrow Y$
 459 is *universally measurable* (with respect to Σ) if $f^{-1}(E)$ is universally measurable
 460 for every $E \in \Sigma$, i.e., $f^{-1}(E)$ is μ -measurable for every Radon probability measure
 461 μ on X . When $Y = \mathbb{R}$ we will always take $\Sigma = \mathcal{B}(\mathbb{R})$, the Borel σ -algebra of \mathbb{R} .
 462 In that case, a function $f : X \rightarrow \mathbb{R}$ is universally measurable if and only if $f^{-1}(U)$
 463 is μ -measurable for every Radon probability measure μ on X and every open set
 464 $U \subseteq \mathbb{R}$. Following [BFT78], the collection of all universally measurable real-valued
 465 functions will be denoted by $M_r(X)$. In the context of deep computations, we will
 466 be interested in transition maps from a state space $L \subseteq \mathbb{R}^{\mathcal{P}}$ to itself. There are two
 467 natural σ -algebras one can consider in the product space $\mathbb{R}^{\mathcal{P}}$: the Borel σ -algebra,
 468 i.e., the σ -algebra generated by open sets in $\mathbb{R}^{\mathcal{P}}$; and the cylinder σ -algebra, i.e.,
 469 the σ -algebra generated by Borel cylinder sets or equivalently basic open sets in
 470 $\mathbb{R}^{\mathcal{P}}$. Note that when \mathcal{P} is countable, both σ -algebras coincide but in general the
 471 cylinder σ -algebra is strictly smaller. We will use the cylinder σ -algebra to define
 472 universally measurable maps $f : \mathbb{R}^{\mathcal{P}} \rightarrow \mathbb{R}^{\mathcal{P}}$. The reason for this choice is because of
 473 the following characterization:

474 **Lemma 3.8.** *Let X be a Hausdorff space and $Y = \prod_{i \in I} Y_i$ be any product of*
 475 *measurable spaces (Y_i, Σ_i) for $i \in I$. Let Σ_Y be the cylinder σ -algebra generated by*
 476 *the measurable spaces (Y_i, Σ_i) . Let $f : X \rightarrow Y$. The following are equivalent:*

- 477 (i) $f : X \rightarrow Y$ is universally measurable (with respect to Σ_Y).
 478 (ii) $\pi_i \circ f : X \rightarrow Y_i$ is universally measurable (with respect to Σ_i) for all $i \in I$.

479 *Proof.* (i) \Rightarrow (ii) is clear since the projection maps π_i are measurable and the com-
 480 position of measurable functions is measurable. To prove (ii) \Rightarrow (i), suppose that
 481 $C = \prod_{i \in I} C_i$ is a measurable cylinder and let J be the finite set of $i \in I$ such that
 482 $C_i \neq Y_i$. Then, $C = \bigcap_{i \in J} \pi_i^{-1}(C_i)$ so $f^{-1}(C) = \bigcap_{i \in J} (\pi_i \circ f)^{-1}(C_i)$ is a universally
 483 measurable set by assumption. \square

484 The previous lemma says that a transition map is universally measurable if and
 485 only if it is universally measurable on all its features. In other words, we can check
 486 measurability of a transition just by checking measurability in all its features. We
 487 will denote by $M_r(X, \mathbb{R}^P)$ the collection of all universally measurable functions
 488 $f : X \rightarrow \mathbb{R}^P$ (with respect to the cylinder σ -algebra), endowed with the topology of
 489 pointwise convergence.

490 **Definition 3.9.** Let (L, \mathcal{P}, Γ) be a CCS. We say that a transition $f : L \rightarrow L$ is
 491 *universally measurable shard-definable* if and only if there exists $\tilde{f} : \mathcal{L}_{sh} \rightarrow \mathcal{L}_{sh}$
 492 extending f such that for every sizer r_\bullet there is a sizer s_\bullet such that the restriction
 493 $\tilde{f}|_{\mathcal{L}[r_\bullet]} : \mathcal{L}[r_\bullet] \rightarrow \mathcal{L}[s_\bullet]$ is universally measurable, i.e. $\pi_P \circ \tilde{f}|_{\mathcal{L}[r_\bullet]} : \mathcal{L}[r_\bullet] \rightarrow [-s_P, s_P]$
 494 is μ -measurable for every Radon probability measure μ on $\mathcal{L}[r_\bullet]$.

495 We will need the following result about NIP and universally measurable func-
 496 tions:

497 **Theorem 3.10** (Bourgain-Fremlin-Talagrand, Theorem 2F in [BFT78]). *Let X be a*
 498 *Hausdorff space and $A \subseteq C(X)$ be pointwise bounded. The following are equivalent:*

- 499 (i) $\overline{A} \subseteq M_r(X)$.
 500 (ii) For every compact $K \subseteq X$, $A|_K$ has the NIP.
 501 (iii) For every Radon measure μ on X , A is relatively countably compact in
 502 $\mathcal{L}^0(X, \mu)$, i.e., every countable subset of A has an accumulation point in
 503 $\mathcal{L}^0(X, \mu)$.

504 Theorem 2.5 immediately yields the following.

505 **Theorem 3.11.** *Let (L, \mathcal{P}, Γ) be a CCS satisfying the Extendibility Axiom. Let R*
 506 *be an exhaustive collection of sizers. Let $\Delta \subseteq \Gamma$ be R -confined. If $\pi_P \circ \Delta|_{\mathcal{L}[r_\bullet]}$ has*
 507 *the NIP for all $P \in \mathcal{P}$ and all $r_\bullet \in R$, then every deep computation is universally*
 508 *measurable shard-definable.*

509 *Proof.* By the Extendibility Axiom, Theorem 2.5 and lemma 2.10 we have that
 510 $\pi_P \circ \tilde{\Delta}|_{\mathcal{L}[r_\bullet]} \subseteq M_r(\mathcal{L}[r_\bullet])$ for all $r_\bullet \in R$ and $P \in \mathcal{P}$. Let $f \in \overline{\Delta}$ be a deep computation.
 511 Write $f = \mathcal{U} \lim_i \gamma_i$ as an ultralimit of computations in Δ . Define $\tilde{f} := \mathcal{U} \lim_i \tilde{\gamma}_i$.
 512 Then, for all $r_\bullet \in R$ and $P \in \mathcal{P}$ $\pi_P \circ \tilde{\gamma}_i|_{\mathcal{L}[r_\bullet]} \in M_r(\mathcal{L}[r_\bullet])$ for all i so $\pi_P \circ \tilde{f}|_{\mathcal{L}[r_\bullet]} \in$
 513 $\overline{M_r(\mathcal{L}[r_\bullet])} \subseteq M_r(\mathcal{L}[r_\bullet])$. \square

514 **Question 3.12.** Under the same assumptions of the previous Theorem, suppose
 515 that every deep computation of Δ is universally measurable shard-definable. Must
 516 $\pi_P \circ \Delta|_{\mathcal{L}[r_\bullet]}$ have the NIP for all $P \in \mathcal{P}$ and all $r_\bullet \in R$?

517 **3.3. Talagrand stability and definability by universally measurable func-**
 518 **tions.** There is another notion closely related to NIP, introduced by Talagrand
 519 in [Tal84] while studying Pettis integration. Suppose that X is a compact Haus-
 520 dorff space and $A \subseteq \mathbb{R}^X$. Let μ be a Radon probability measure on X . Given a
 521 μ -measurable set $E \subseteq X$, a positive integer k and real numbers $a < b$, we write:

$$D_k(A, E, a, b) = \bigcup_{f \in A} \{x \in E^{2k} : f(x_{2i}) \leq a, f(x_{2i+1}) \geq b \text{ for all } i < k\}$$

522 We say that A is *Talagrand μ -stable* if and only if for every μ -measurable
 523 set $E \subseteq X$ of positive measure and for every $a < b$ there is $k \geq 1$ such that
 524 $(\mu^{2k})^*(D_k(A, E, a, b)) < (\mu(E))^{2k}$. Notice that we work with the outer measure
 525 because it is not necessarily true that the sets $D_k(A, E, a, b)$ are μ -measurable.
 526 This is certainly the case when A is a countable set of continuous (or μ -measurable)
 527 functions.

528 The following lemma establishes that Talagrand stability is a way to ensure that
 529 deep computations are definable by measurable functions. We include the proof for
 530 the reader's convenience.

531 **Lemma 3.13.** *If A is Talagrand μ -stable, then \overline{A} is also Talagrand μ -stable and*
 532 *$\overline{A} \subseteq \mathcal{L}^0(X, \mu)$.*

533 *Proof.* First, observe that a subset of a μ -stable set is μ -stable. To show that \overline{A}
 534 is μ -stable, observe that $D_k(\overline{A}, E, a, b) \subseteq D_k(A, E, a', b')$ where $a < a' < b' <$
 535 b and E is a μ -measurable set with positive measure. It suffices to show that
 536 $\overline{A} \subseteq \mathcal{L}^0(X, \mu)$. Suppose that there exists $f \in \overline{A}$ such that $f \notin \mathcal{L}^0(X, \mu)$. By a
 537 characterization of measurable functions (see 413G in [Fre03]), there exists a μ -
 538 measurable set E of positive measure and $a < b$ such that $\mu^*(P) = \mu^*(Q) = \mu(E)$
 539 where $P = \{x \in E : f(x) \leq a\}$ and $Q = \{x \in E : f(x) \geq b\}$. Then, for any $k \geq 1$:
 540 $(P \times Q)^k \subseteq D_k(\{f\}, E, a, b)$ so $(\mu^{2k})^*(D_k(\{f\}, E, a, b)) = (\mu^*(P)\mu^*(Q))^k = (\mu(E))^{2k}$.
 541 Thus, $\{f\}$ is not μ -stable, but we argued before that a subset of a μ -stable set must
 542 be μ -stable. \square

543 We say that A is *universally Talagrand stable* if A is Talagrand μ -stable for
 544 every Radon probability measure μ on X . A similar argument as before, yields the
 545 following:

546 **Theorem 3.14.** *Let (L, \mathcal{P}, Γ) be a CCS satisfying the Extendibility Axiom. If*
 547 *$\pi_P \circ \Delta|_{L[r_\bullet]}$ is universally Talagrand stable for all $P \in \mathcal{P}$ and all sizers r_\bullet , then*
 548 *every deep computation is universally measurable sh-definable.*

549 It is then natural to ask: what is the relationship between Talagrand stability
 550 and the NIP? The following dichotomy will be useful.

551 **Lemma 3.15** (Fremlin's Dichotomy, 463K in [Fre03]). *If (X, Σ, μ) is a perfect σ -*
 552 *finite measure space (in particular, for X compact and μ a Radon probability measure*
 553 *on X) and $\{f_n : n \in \mathbb{N}\}$ be a sequence of real-valued measurable functions on X , then*
 554 *either:*

- 555 (i) $\{f_n : n \in \mathbb{N}\}$ has a subsequence that converges μ -almost everywhere, or
- 556 (ii) $\{f_n : n \in \mathbb{N}\}$ has a subsequence with no μ -measurable accumulation point in
 557 \mathbb{R}^X .

558 The preceding lemma can be considered as the measure theoretic version of
 559 Rosenthal's Dichotomy. Combining this dichotomy with the Theorem 3.10 we get
 560 the following result:

561 **Theorem 3.16.** *Let X be a Hausdorff space and $A \subseteq C(X)$ be pointwise bounded.*
 562 *The following are equivalent:*

- 563 (i) $\overline{A} \subseteq M_r(X)$.
- 564 (ii) For every compact $K \subseteq X$, $A|_K$ has the NIP.
- 565 (iii) For every Radon measure μ on X , A is relatively countably compact in
 566 $\mathcal{L}^0(X, \mu)$, i.e., every countable subset of A has an accumulation point in
 567 $\mathcal{L}^0(X, \mu)$.
- 568 (iv) For every Radon measure μ on X and every sequence $\{f_n : n \in \mathbb{N}\} \subseteq A$,
 569 there is a subsequence that converges μ -almost everywhere.

570 *Proof.* Notice that the equivalence (i)-(iii) is Theorem 3.10. Notice that the equiv-
 571 alence of (iii) and (iv) is Fremlin's Dichotomy Theorem. \square

572 **Lemma 3.17.** *Let X be a compact Hausdorff space and $A \subseteq C(X)$ be pointwise*
 573 *bounded. If A is universally Talagrand stable, then A has the NIP.*

574 *Proof.* By Theorem 3.10, it suffices to show that A is relatively countably compact
 575 in $\mathcal{L}^0(X, \mu)$ for all Radon probability measure μ on X . Since A is Talagrand μ -stable
 576 for any such μ , then $\overline{A} \subseteq \mathcal{L}^0(X, \mu)$. In particular, A is relatively countably compact
 577 in $\mathcal{L}^0(X, \mu)$. \square

578 **Question 3.18.** Is the converse true?

579 There is a delicate point in this question, as it may be sensitive to set-theoretic
 580 axioms (even assuming countability of A).

581 **Theorem 3.19** (Talagrand, Theorem 9-3-1(a) in [Tal84]). *Let X be a compact*
 582 *Hausdorff space and $A \subseteq M_r(X)$ be countable and pointwise bounded. Assume that*
 583 *$[0, 1]$ is not the union of $< \mathfrak{c}$ closed measure zero sets. If A has the NIP, then A is*
 584 *universally Talagrand stable.*

585 **Theorem 3.20** (Fremlin, Shelah, [FS93]). *It is consistent that there exists a count-*
 586 *able pointwise bounded set of Lebesgue measurable functions with the NIP which is*
 587 *not Talagrand stable with respect to Lebesgue measure.*

588 APPENDIX: MEASURE THEORY

589 Given a set X , a collection Σ of subsets of X is called a σ -algebra if Σ contains
 590 X and is closed under complements and countable unions. Hence, for example, a
 591 σ -algebra is also closed under countable intersections. Intuitively, a σ -algebra is
 592 a collection of sets in which we can define a σ -additive measure. We call sets in
 593 a σ -algebra Σ *measurable sets* and the pair (X, Σ) a measurable space. If X is a
 594 topological space, there is a natural σ -algebra of subsets of X , namely the *Borel*
 595 σ -algebra $\mathcal{B}(X)$, i.e., the smallest σ -algebra containing all open subsets of X . Given
 596 two measurable spaces (X, Σ_X) and (Y, Σ_Y) , we say that a function $f : X \rightarrow Y$ is
 597 *measurable* if and only if $f^{-1}(E) \in \Sigma_X$ for every $E \in \Sigma_Y$. In particular, we say that
 598 $f : X \rightarrow \mathbb{R}$ is measurable if $f^{-1}(E) \in \Sigma_X$ for all $E \in \mathcal{B}(\mathbb{R})$ (equivalently, E open in
 599 \mathbb{R}).

Given a measurable space (X, Σ) , a σ -additive measure is a non-negative function $\mu : \Sigma \rightarrow \mathbb{R}$ with the property that $\mu(\emptyset) = 0$ and $\mu(\bigcup_{n=0}^{\infty} A_n) = \sum_{n=0}^{\infty} \mu(A_n)$ whenever $\{A_n : n \in \mathbb{N}\} \subseteq \Sigma$ is pairwise disjoint. We call (X, Σ, μ) a *measure space*. A σ -additive measure is called a *probability measure* if $\mu(X) = 1$. A measure μ is *complete* if for every $A \subseteq B \in \Sigma$, $\mu(B) = 0$ implies $A \in \Sigma$. In words, subsets of measure-zero sets are always measurable (and hence, by the monotonicity of μ , have measure zero as well). A measure μ is σ -finite if $X = \bigcup_{n=1}^{\infty} X_n$ where $\mu(X_n) < \infty$ for all $n \in \mathbb{N}$ (i.e., X can be decomposed into countably many finite measure sets). A measure μ is *perfect* if for every measurable $f : X \rightarrow \mathbb{R}$ and every measurable set E with $\mu(E) > 0$, there exists a compact $K \subseteq f(E)$ such that $\mu(f^{-1}(K)) > 0$. We say that a property $\phi(x)$ about $x \in X$ holds μ -almost everywhere if $\mu(\{x \in X : \phi(x) \text{ does not hold}\}) = 0$.

A special example of the preceding concepts is that of a *Radon measure*. If X is a Hausdorff topological space, then a measure μ on the Borel sets of X is called a *Radon measure* if

- for every open set U , $\mu(U)$ is the supremum of $\mu(K)$ over all compact $K \subseteq U$, that is, the measure of open sets may be approximated via compact sets; and
- every point of X has a neighborhood $U \ni x$ for which $\mu(U)$ is finite.

Perhaps the most famous example of a Radon measure on \mathbb{R} is the Lebesgue measure of Borel sets. If X is finite, $\mu(A) := |A|$ (the cardinality of A) defines a Radon measure on X . Every Radon measure is perfect (see 451A, 451B and 451C in [Fre03]).

While not immediately obvious, sets can be measurable according to one measure, but non-measurable according to another. Given a measure space (X, Σ, μ) we say that a set $E \subseteq X$ is μ -measurable if there are $A, B \in \Sigma$ such that $A \subseteq E \subseteq B$ and $\mu(B \setminus A) = 0$. The set of all μ -measurable sets is a σ -algebra containing Σ and it is denoted by Σ_μ . A set $E \subseteq X$ is *universally measurable* if it is μ -measurable for every Radon probability measure on X . It follows that Borel sets are universally measurable. We say that $f : X \rightarrow \mathbb{R}$ is μ -measurable if $f^{-1}(E) \in \Sigma_\mu$ for all $E \in \mathcal{B}(\mathbb{R})$ (equivalently, E open in \mathbb{R}). The set of all μ -measurable functions is denoted by $\mathcal{L}^0(X, \mu)$.

Recall that if $\{X_i : i \in I\}$ is a collection of topological spaces indexed by some set I , then the product space $X := \prod_{i \in I} X_i$ is endowed with the topology generated by *cylinders*, that is, sets of the form $\prod_{i \in I} U_i$ where each U_i is open in X_i , and $U_i = X_i$ except for finitely many indices $i \in I$. If each space is measurable, say we pair X_i with a σ -algebra Σ_i , then there are multiple ways to interpret the product space X as a measurable space, but the interpretation we care about in this paper is the so called *cylinder σ -algebra*, as used in Lemma 3.8. Namely, let Σ be the σ -algebra generated by sets of the form

$$\prod_{i \in I} C_i, \quad C_i \in \Sigma_i, \quad C_i = X_i \text{ for all but finitely many } i \in I.$$

We remark that when I is uncountable and $\Sigma_i = \mathcal{B}(X_i)$ for all $i \in I$, then Σ is, in general, strictly **smaller** than $\mathcal{B}(X)$.

REFERENCES

- [ADIW24] Samson Alva, Eduardo Dueñez, Jose Iovino, and Claire Walton. Approximability of deep equilibria. *arXiv preprint arXiv:2409.06064*, 2024. Last revised 20 May 2025, version 3.
- [ADK08] Spiros A. Argyros, Pandelis Dodos, and Vassilis Kanellopoulos. A classification of separable Rosenthal compacta and its applications. *Dissertationes Mathematicae*, 449:1–55, 2008.
- [Ark91] A. V. Arkhangel’skii. *Topological Function Spaces*. Springer, New York, 1st edition, 1991.
- [BD19] Shai Ben-David. Understanding machine learning through the lens of model theory. *The Bulletin of Symbolic Logic*, 25(3):319–340, 2019.
- [BFT78] J. Bourgain, D. H. Fremlin, and M. Talagrand. Pointwise compact sets of Baire-measurable functions. *American Journal of Mathematics*, 100(4):845–886, 1978.
- [Deb13] Gabriel Debs. Descriptive aspects of Rosenthal compacta. In *Recent Progress in General Topology III*, pages 205–227. Springer, 2013.
- [Fre00] David H. Fremlin. *Measure Theory, Volume 1: The Irreducible Minimum*. Torres Fremlin, Colchester, UK, 2000. Second edition 2011.
- [Fre01] David H. Fremlin. *Measure Theory, Volume 2: Broad Foundations*. Torres Fremlin, Colchester, UK, 2001.
- [Fre03] David H. Fremlin. *Measure Theory, Volume 4: Topological Measure Spaces*. Torres Fremlin, Colchester, UK, 2003.
- [FS93] David H. Fremlin and Saharon Shelah. Pointwise compact and stable sets of measurable functions. *Journal of Symbolic Logic*, 58(2):435–455, 1993.
- [GM22] Eli Glasner and Michael Megrelishvili. Todorčević’ trichotomy and a hierarchy in the class of tame dynamical systems. *Transactions of the American Mathematical Society*, 375(7):4513–4548, 2022.
- [HT23] Clovis Hamel and Franklin D. Tall. C_p -theory for model theorists. In Jose Iovino, editor, *Beyond First Order Model Theory, Volume II*, chapter 5, pages 176–213. Chapman and Hall/CRC, 2023.
- [Kec95] Alexander S. Kechris. *Classical Descriptive Set Theory*, volume 156 of *Graduate Texts in Mathematics*. Springer-Verlag, 1995.
- [Kha20] Karim Khanaki. Stability, nip, and nsop; model theoretic properties of formulas via topological properties of function spaces. *Mathematical Logic Quarterly*, 66(2):136–149, 2020.
- [She78] Saharon Shelah. *Unstable Theories*, volume 187 of *Studies in Logic and the Foundations of Mathematics*. North-Holland, 1978.
- [Sim15a] Pierre Simon. *A Guide to NIP Theories*, volume 44 of *Lecture Notes in Logic*. Cambridge University Press, 2015.
- [Sim15b] Pierre Simon. Rosenthal compacta and NIP formulas. *Fundamenta Mathematicae*, 231(1):81–92, 2015.
- [Tal84] Michel Talagrand. *Pettis Integral and Measure Theory*, volume 51 of *Memoirs of the American Mathematical Society*. American Mathematical Society, Providence, RI, USA, 1984. Includes bibliography (pp. 220–224) and index.
- [Tal87] Michel Talagrand. The Glivenko-Cantelli problem. *The Annals of Probability*, 15(3):837–870, 1987.
- [Tka11] Vladimir V. Tkachuk. *A C_p -Theory Problem Book: Topological and Function Spaces*. Problem Books in Mathematics. Springer, 2011.
- [Tod97] Stevo Todorčević. *Topics in Topology*, volume 1652 of *Lecture Notes in Mathematics*. Springer Berlin, Heidelberg, 1997.
- [Tod99] Stevo Todorčević. Compact subsets of the first Baire class. *Journal of the American Mathematical Society*, 12(4):1179–1212, 1999.