# A SIMPLE CONTINUOUS THEORY

#### JAMES HANSON

ABSTRACT. In the context of continuous first-order logic, special attention is often given to theories that are somehow continuous in an 'essential' way. A common feature of such theories is that they do not interpret any infinite discrete structures. We investigate a stronger condition that is easier to establish and use it to give an example of a strictly simple continuous theory that does not interpret any infinite discrete structures: the theory of richly branching  $\mathbb{R}$ -forests with generic binary predicates. We also give an example of a superstable theory that fails to satisfy this stronger condition but nevertheless does not interpret any infinite discrete structures.

#### Introduction

In continuous first-order model theory, there is a common desire to find examples of various phenomena occurring in 'essentially continuous' theories, where 'essential continuity' is the informally defined property of resembling a discrete theory as little as possible. Most of the standard motivating examples of metric structures—infinite-dimensional Hilbert spaces, atomless probability algebras, operator algebras, Banach lattices, the Urysohn sphere, and  $\mathbb{R}$ -trees—are regarded as essentially continuous [9]. Despite not settling on a precise definition of essential continuity, continuous logicians have been asking the following question for a long time.

# Question 0.1. Is there an essentially continuous strictly simple theory?

As discussed in the introduction of [9], natural attempts to produce an example fail: The randomization of an IP theory is TP<sub>2</sub> [4, Cor. 4.13], implying that no randomization is strictly simple, and the theories produced by adding generic automorphisms to the standard examples of essentially continuous stable theories are themselves stable (whereas in discrete logic this often produces a strictly simple theory).

Obviously we would like to address Question 0.1 in this paper. The issue of course is that 0.1 is not fully specified. In all honesty, it is unclear that there will ever be a definitively compelling notion of essential continuity, and as such we will not attempt to provide one here.

In earlier work of the author studying uncountable (or inseparable) categoricity in continuous logic, the condition of not interpreting the theory of any infinite discrete structure (which by an abuse of terminology we will refer to as not 'interpreting an infinite discrete structure') was used as a precisely defined proxy for essential continuity [12]. In that context, this was a rather natural choice, as any

Date: June 26, 2023.

<sup>2020</sup> Mathematics Subject Classification. 03C66, 03C45.

Key words and phrases. continuous logic, simplicity, non-discreteness notions.

 $\omega$ -stable theory that interprets an infinite discrete structure has a strongly minimal imaginary sort, allowing for a Baldwin-Lachlan characterization of uncountable categoricity for those theories in particular. Outside of the context of  $\omega$ -stable theories (or, more generally, outside of the context of dictionaric theories), this is perhaps too strong. The difficulty is that interpretation of an infinite discrete structure relies on the ability to isolate the relevant discrete behavior in a definable set and, as is well known at this point, definable sets are generally quite poorly behaved in continuous logic. As we will show in Section 1.3, this can in fact be a fatal obstruction to interpreting an infinite discrete structure in a theory that is seemingly rather discrete in nature. Given this issue, in this paper we study a stronger precisely defined proxy for essential continuity, which is roughly speaking equivalent to having no uniformly discrete 'hyperimaginary sorts.' This notion seems to be more robust; in particular, it is always witnessed by 1-types (Theorem 1.5). Furthermore, as evidenced by the relative complexity of the proofs of Proposition 1.17 and Theorem 2.38, it seems to be generally easier to work with. Finally, we show that, under the assumption of dictionaricity (and so a fortiori  $\omega$ -stability), it is equivalent to the condition of interpreting an infinite discrete structure (Proposition 1.9).

In Section 2, we present an example of a strictly simple theory that exhibits this stronger non-discreteness property (and therefore also does not interpret any infinite discrete structure). The idea is itself fairly simple—add a generically chosen binary predicate to a generic ' $\mathbb{R}$ -forest.' Aside from verifying that the example works, the only real issue is choosing the modulus of uniform continuity of the predicate correctly so that the resulting class of structures has the amalgamation property. It turns out that 1-Lipschitzness with regards to the  $\ell^1$ -metric on pairs (i.e., d(xy, zw) = d(x, z) + d(y, w)) is precisely what is necessary to make this work.

Finally, in Section 3, we describe a stronger non-discreteness property that separates Hilbert spaces, atomless probability algebras, and the Urysohn sphere from  $\mathbb{R}$ -trees and ask whether it is incompatible with strict simplicity.

**Notation.** In continuous logic, often the minimum and maximum function are used to clamp a given value to a certain interval, such as in the expression  $\max(0,\min(1,x))$ . In the author's opinion representing such a frequently used operation with 12 characters is a bit ridiculous, so we will employ the following notation.

**Notation 0.2.** For any  $r \leq s$ , we write  $[x]_r^s$  for the function  $\max(r, \min(s, x))$ .

# 1. Crisp imaginary types

**Definition 1.1.** In a theory T, a type p(x) is *crisp* if the metric is  $\{0,1\}$ -valued on the set of realizations of p(x) in the monster model.

While it might be natural to use the word 'discrete' for the previous notion, there is an issue with this, which is that discreteness in metric spaces is a topological condition and is strictly weaker than uniform discreteness (even in arbitrarily saturated metric structures). 'Uniform discreteness,' while accurate, is also long and still fails to specify the scale of uniform discreteness. (In particular, the class of infinite structures with uniformly discrete metrics is not elementary.) As such, we have opted to use the adjective 'crisp,' which is commonly used in the real-valued logic literature to mean  $\{0,1\}$ -valuedness.

<sup>&</sup>lt;sup>1</sup>Although note that the converse does not hold.

**Definition 1.2.** An *imaginary type* is a type in an imaginary sort.

#### 1.1. Characterization.

**Proposition 1.3.** A theory T has a non-algebraic crisp imaginary type if and only if there is a type p(x) in some real (possibly infinitary) product sort and a formula E(x,y) such that on the set of realizations of p(x), E(x,y) is  $\{0,1\}$ -valued and defines an equivalence relation with infinitely many classes.

Proof. Let q(z) be a non-algebraic crisp type in some imaginary sort I. By the infinitary pigeonhole principle, we may assume that q(z) is a complete type over some set of parameters A. By [14, Prop. 3.2.13], there is a definable pseudo-metric  $\rho(x,y)$  on  $H^n$  (for some  $n \leq \omega$  where H is the home sort) and a definable subset  $D \subseteq H^n/\rho$  such that I and D are isomorphic as sorts (with the metric on I corresponding to the metric  $\rho$  on D). Therefore we may assume that q(z) is a type in D. Let p(x) be the type of elements of  $H^n$  that project to realizations of q(z). We clearly now have that  $\rho(x,y)$  is  $\{0,1\}$ -valued on the realizations of p(x) and defines an equivalence relation with infinitely many classes.

For the other direction, we have that  $d(x,y)=0 \lor (p(x) \land p(y) \land E(x,y)=0)$  is a type-definable equivalence relation. By [3, Thm. 2.20], there is a definable pseudo-metric  $\rho(x,y)$  with the property that  $\rho(x,y)=0$  if and only if  $d(x,y)=0 \lor (p(x) \land p(y) \land E(x,y)=0)$ . Note that for realizations of p(x), we have by compactness that there is some  $\varepsilon>0$  such that if E(x,y)=1, then  $\rho(x,y)\geq\varepsilon$ . By replacing  $\rho(x,y)$  with  $\min(\frac{1}{\varepsilon}\rho(x,y),1)$ , we may assume that  $\rho(x,y)$  is  $\{0,1\}$ -valued on realizations of p(x). In particular, we have that  $\rho(x,y)=E(x,y)$  for realizations of p(x). Let  $f:H^\omega\to H^\omega/\rho$  be the definable quotient map. Let q(y) be the type of f(x) for any x realizing p(x). By construction we have that q(y) is a non-algebraic crisp imaginary type.

**Lemma 1.4.** If T has a non-algebraic crisp imaginary type, then there is a type p(x) in a finitary product sort and a formula E(x,y) satisfying the conclusion of Proposition 1.3.

Proof. Assume that we have p(x) and E(x,y) satisfying the conclusion of Proposition 1.3 with  $x = x_0x_1x_2...$  an  $\omega$ -tuple of variables. Since E(x,y) is  $\{0,1\}$ -valued on p(x), there is a restricted formula  $E_0(x,y)$  that is equal to E(x,y) for realizations of p(x). This restricted formula can only involve finitely many variables in the tuple x, so we may write the formula  $E_0$  as  $E_0(z,w)$ , where z and w are finite initial segments of x and y, respectively, with |z| = |w|. Finally, if we take q(z) to be the restriction of p(x) to the tuple z, then we have that for realizations of q(x),  $E_0(z,w)$  is a  $\{0,1\}$ -valued equivalence relation with infinitely many classes.

**Theorem 1.5.** Fix a theory T. The following are equivalent.

- (1) T has a non-algebraic crisp imaginary type.
- (2) There is a countable set of parameters A, a 1-type  $p(x) \in S_x(A)$ , and a formula E(x,y) over A such that E defines a  $\{0,1\}$ -valued equivalence relation on the realizations of p(x) with infinitely many equivalence classes.
- (3) There is a model  $M \models T$  with  $|M| \leq |T|$ , a 1-type  $p(x) \in S_x(M)$ , and a formula E(x,y) over M such that E defines a  $\{0,1\}$ -valued equivalence relation on the realizations of p(x) with more than one equivalence class.

<sup>&</sup>lt;sup>2</sup>See also [7, Prop. 4.8] for an easier proof.

*Proof.* 2 obviously implies 1. To see that 1 implies 2, assume that we have a complete non-algebraic crisp imaginary type  $p_0(\bar{x})$  witnessed by the formula  $E_0(\bar{x}, \bar{y})$  over some countable set of parameters  $A_0$  where  $\bar{x}$  is an  $n_0$ -tuple of variables, with  $n_0$  finite. We will inductively reduce the number of variables until we only have

At stage i, given  $p_i$ ,  $E_i$ ,  $A_i$ , and  $n_i$ , if  $n_i = 1$ , then we are done. Otherwise, write  $p_i$  as  $p_i(\bar{x}, z)$  and  $E_i$  as  $E_i(\bar{x}, z, \bar{y}, w)$ . Let q(z) be the restriction of  $p_i$  to the variable y. Let c be some realization of q. At this point there are two cases.

If the number of  $E_i$ -classes of elements of the form bc is infinite, then let  $p_{i+1}(\bar{x}) = p_i(\bar{x},c) \in S_{\bar{x}}(A_ic)$ , let  $E_{i+1}(\bar{x},\bar{y}) = E_i(\bar{x},c,\bar{y},c)$ , let  $A_{i+1} = A_ic$ , and let  $n_{i+1} = n_i - 1$ .

If the number of  $E_i$ -classes of elements of the form  $\bar{b}c$  is k for some finite k, then the same will be true of any  $c' \models q$  (since q is a complete type). Let F(z, w) be the relation on realizations of q(z) defined by the following: F(c, c') holds if and only if

- for every tuple  $\bar{b}$  such that  $\bar{b}c \models p_i$ , there is a tuple  $\bar{b}'$  such that  $\bar{b}'c' \models p_i$  and  $E_i(\bar{b}c, \bar{b}'c')$  and
- for every tuple  $\bar{b}'$  such that  $\bar{b}'c' \models p_i$ , there is a tuple  $\bar{b}$  such that  $\bar{b}c \models p_i$  and  $E_i(\bar{b}c,\bar{b}'c')$ ,

where quantification is taken over the monster model. This is clearly an  $A_i$ -invariant equivalence relation on realizations of q(z). We need to show that it is actually type-definable and co-type-definable on realizations of q(z). To see that it is type-definable, it is sufficient to show that if  $(r_j(z,w))_{j\in J}$  is a net of types in  $S_{zw}(A_i)$  converging to r(z,w) such that  $r_j(z,w)\vdash q(z)\land q(w)\land F(z,w)$  for each  $j\in J$ , then  $r(z,w)\vdash F(z,w)$  as well. For each  $j\in J$ , let  $M_j$  be an  $|\mathcal{L}|^+$ -saturated model containing  $c_j$  and  $c_j'$  such that  $c_jc_j'\models r_j$ . For each  $j\in J$ , since  $M_j$  is  $|\mathcal{L}|^+$ -saturated, we know that there are  $\bar{b}_{j,0},\ldots,\bar{b}_{j,k-1}$  such that  $\bar{b}_{j,\ell}c_j\models p_i$  for each  $\ell< k$  and such that the  $E_i$ -classes of  $\bar{b}_{j,0}c_j,\ldots,\bar{b}_{j,k-1}c_j$  are distinct and represent all  $E_i$ -classes of tuples containing  $c_j$ . We also know that there are similar tuples  $\bar{b}'_{j,0},\ldots,\bar{b}'_{j,k-1}$  for  $c'_j$ . Since  $F(c_j,c'_j)$  holds, we may assume that  $E_i(\bar{b}_{j,\ell}c_j,\bar{b}'_{j,\ell}c'_j)$  holds for each  $\ell< k$ .

Let  $\mathcal{U}$  be an ultrafilter on J extending the filter generated by  $\{\{\underline{j}' \geq j : k \in J\}: j \in J\}$ . (This is a filter because J is a directed set.) Let  $(M_{\mathcal{U}}, c_{\mathcal{U}}, \bar{b}_{\mathcal{U},0}, \dots, \bar{b}_{\mathcal{U},k-1}, c'_{\mathcal{U}}, \bar{b}'_{\mathcal{U},0}, \dots, \bar{b}'_{\mathcal{U},k-1})$  be the ultraproduct

$$\prod_{j \in J} (M_j, c_j, \bar{b}_{j,0}, \dots, \bar{b}_{j,k-1}, c'_j, \bar{b}'_{j,0}, \dots, \bar{b}'_{j,k-1}) / \mathcal{U}.$$

Since  $E_i$  is type-definable, we have that  $E(\bar{b}_{\mathcal{U},\ell}c_{\mathcal{U}},\bar{b}'_{\mathcal{U},\ell}c'_{\mathcal{U}})$  holds for each  $\ell < k$ . Furthermore, by the choice of k, we know that there are no more  $E_i$ -equivalence classes of tuples containing  $c_{\mathcal{U}}$  and of tuples containing  $c'_{\mathcal{U}}$ , so  $F(c_{\mathcal{U}},c'_{\mathcal{U}})$  holds. Therefore we have that F(z,w) is type-definable.

Now to see that it is co-type-definable, we need to show that if  $(r_j(z,w))_{j\in J}$  is a net of types in  $S_{zw}(A_i)$  converging to r(z,w) such that  $r_j(z,w)\vdash q(z)\land q(w)\land \neg F(z,w)$  for each  $j\in J$ , then  $r(z,w)\vdash \neg F(z,w)$ . To this end, find a family of models and tuples  $(M_j,c_j,\bar{b}_{j,0},\ldots,\bar{b}_{j,k-1},c'_j,\bar{b}'_{j,0},\ldots,\bar{b}'_{j,k-1})_{j\in J}$  as before, with  $\bar{b}_{j,0}c_j,\ldots,\bar{b}_{j,k-1}c_j$  representing all  $E_i$ -classes of tuples containing  $c_j$  and  $\bar{b}'_{j,0}c'_j,\ldots,\bar{b}'_{j,k-1}c'_j$  representing all  $E_i$ -classes of tuples containing  $c_j$ . Let  $\mathcal{U}$  be an ultrafilter that is compatible with the directed set J as before. Since  $\neg F(c_j,c'_j)$  for each  $j\in J$ , we have that for each  $j\in J$ , either there is an  $\ell< k$  such that

 $\neg E_i(\bar{b}_{j,\ell}c_j,\bar{b}'_{j,m}c'_j)$  for all m < k, or there is an m < k such that  $\neg E_i(\bar{b}_{j,\ell}c_j,\bar{b}'_{j,m}c'_j)$  for all  $\ell < k$ . Since there are finitely many possibilities, there is an  $X \in \mathcal{U}$  such that this behaves uniformly on X. Without loss of generality, we may assume that for all  $j \in X$ ,  $\neg E_i(\bar{b}_{j,0}c_j,\bar{b}'_{j,m}c'_j)$  for all m < k. Since this is a type-definable fact, we have that in the ultraproduct  $(M_{\mathcal{U}},c_{\mathcal{U}},\bar{b}_{\mathcal{U},0},\ldots,\bar{b}_{\mathcal{U},k-1},c'_{\mathcal{U}},\bar{b}'_{\mathcal{U},0},\ldots,\bar{b}'_{\mathcal{U},k-1})$ ,  $\neg E_i(\bar{b}_{\mathcal{U},0}c_{\mathcal{U}},\bar{b}'_{\mathcal{U},m}c'_{\mathcal{U}})$  for all m < k. Once again, since these are the only  $E_i$ -classes containing tuples extending  $c_j$  and  $c'_j$ , we have that  $\neg F(c_{\mathcal{U}},c'_{\mathcal{U}})$ . Therefore F(z,w) is co-type-definable as well.

Since F(z, w) is type-definable and co-type-definable, we can find an formula  $E_{i+1}(y, z)$  that is  $\{0, 1\}$ -valued and agrees with F(z, w) on realizations of q(y). Let  $p_{i+1}(z) = q(z)$ ,  $A_{i+1} = A_i$ , and  $n_{i+1} = 1$ .

By induction, since  $n_{i+1} < n_i$  at each stage, we eventually arrive at the required p(x), E(x, y), and A.

To see that 2 implies 3, assume 2 and pass to a model M containing A with  $|M| \leq |T|$ . By the pigeonhole principle, there is some completion of p(x) over M with infinitely many E-classes.

Finally, assume that 3 holds. Let a and b be two realizations of p(x) such that E(a,b)=0. Fix an M-coheir  $q(x)\supset p(x)$  and let  $c\models q\upharpoonright Mab$ . Since E is an equivalence relation, it must be the case that E(c,a)=E(c,b)=0. Therefore any Morley sequence generated by q is pairwise E-inequivalent and so E has infinitely many equivalence classes among the realizations of p.

The following corollary gives a nice test for establishing that a theory has no non-algebraic crisp imaginary types. Recall that  $\#^{dc}M$  is the metric density character of M.

**Corollary 1.6.** Suppose that for any  $M \models T$  with  $\#^{dc}M \leq |T|$  and any 1-type  $p(x) \in S_1(M)$ , the set  $F_p := \{q(x,y) \in S_2(M) : q(x,y) \supseteq p(x) \cup p(y)\}$  is connected. Then T has no non-algebraic crisp imaginary types.

Proof. By Theorem 1.5, if T has a non-algebraic crisp imaginary type, then there is a countable set of parameters A, a 1-type  $r(x) \in S_x(A)$ , and an A-formula E(x,y) such that E defines a  $\{0,1\}$ -valued equivalence relation on the realizations of r(x) with infinitely many classes. If we find a model  $M \supseteq A$  with  $\#^{dc}M \le |T|$ , then there is a completion  $p(x) \supseteq r(x)$  such that among the realizations of p(x) there are infinitely many E-classes (by the pigeonhole principle). Therefore the set  $F_p$  has a non-trivial clopen subset (i.e., the set of types  $q(x,y) \supseteq p(x) \cup p(y)$  such that  $q(x,y) \vdash E(x,y) = 1$ ) and is not connected.

In particular, Corollary 1.6 can be used to quickly verify that the theories of infinite-dimensional Hilbert spaces, atomless probability algebras, and the Urysohn sphere do not have any non-algebraic crisp imaginary types (although for Hilbert spaces at least this was established in [5]). Perhaps more surprisingly, this also implies the same for the randomization of any discrete or continuous theory.

It is unclear at the moment whether the condition in Corollary 1.6 characterizes the non-existence of non-algebraic crisp imaginary types, although it seems unlikely.

**Question 1.7.** If T has no non-algebraic crisp imaginary types, does it follow that for every model  $M \models T$  and every 1-type  $p(x) \in S_x(M)$ , the set  $\{q(x,y) \in S_2(M) : q(x,y) \supseteq p(x) \cup p(y)\}$  is connected?

1.2. Relationship with earlier notions. Here we will show that under a non-trivial, but relatively common, niceness condition, a theory with a non-algebraic crisp imaginary type interprets the theory of an infinite discrete structure. A theory T is dictionaric if for any set of parameters A and any finite tuple of variables x, the type space  $S_{\bar{x}}(A)$  has a basis of definable neighborhoods. As established in [12], this condition allows one to generalize many constructions from discrete logic to the context of continuous logic.

**Lemma 1.8.** If  $p(x) \in S_x(A)$  is a crisp type in some sort, then there is an open neighborhood  $U \ni p(x)$  such that for any a and b with  $\operatorname{tp}(a/A), \operatorname{tp}(b/A) \in U$ , either  $d(a,b) \leq \frac{1}{4}$  or  $d(a,b) \geq \frac{3}{4}$ .

*Proof.* Assume that for every open neighborhood  $U\ni p(x)$ , there are a and b realizing types in U such that  $\frac{1}{4}< d(a,b)<\frac{3}{4}$ . Then, by compactness, there are a and b realizing p(x) such that  $\frac{1}{4}\le d(a,b)\le\frac{3}{4}$ , contradicting crispness of p(x).  $\square$ 

**Proposition 1.9.** If T is dictionaric and has a non-algebraic crisp type (in a real sort), then it interprets the theory of an infinite discrete structure.

Proof. Let  $p(x) \in S_x(A)$  be a non-algebraic crisp type. By Lemma 1.8, there is an open set  $U \ni p(x)$  such that for any a and b realizing types in U, either  $d(a,b) \le \frac{1}{4}$  or  $d(a,b) \ge \frac{3}{4}$ . Since T is dictionaric, we can find an A-definable set D such that  $p \in D \subseteq U$ . Therefore the same fact holds for elements of D. The function  $E(x,y) = [4d(x,y)-1]_0^1$  now defines a  $\{0,1\}$ -valued equivalence relation on the set of realizations of D. Since D contains p(x), there are infinitely many E-classes.  $\square$ 

Note however that it is not generally true that  $T^{\text{eq}}$  is dictionaric for dictionaric T.<sup>3</sup> We will show in Proposition 1.17 that without the assumption of dictionaricity, these conditions are not equivalent.

## **Proposition 1.10.** Fix a theory T.

- (1) If T has formulas  $\varphi(\bar{x})$  and  $E(\bar{x}, \bar{y})$  such that
  - E is a  $\{0,1\}$ -valued equivalence relation on  $\{\bar{a}: \varphi(\bar{a})<1\}$ ,
  - E has infinitely many equivalence classes, and
  - every E-class of some  $\bar{a}$  with  $\varphi(\bar{a}) < 1$  contains a  $\bar{b}$  such that  $\varphi(\bar{b}) = 0$ , then T has an infinite crisp imaginary sort I and a bijection f between I and E-classes of elements with  $\varphi < 1$  such that any automorphism of the monster fixes  $a \in I$  if and only if it fixes f(a).
- (2) Such a pair of formulas exists for any infinite crisp imaginary sort.
- (3) The above formulas may be taken to be restricted formulas.
- (4) If the above formulas are quantifier-free, they may be taken to be quantifier-free restricted formulas.

*Proof.* For 1, assume that we are given  $\varphi(\bar{x})$  and  $E(\bar{x}, \bar{y})$  such that the bullet points hold. We may assume that  $\bar{x}$  is an  $\omega$ -tuple. By replacing  $\varphi(\bar{x})$  with  $[\varphi(\bar{x})]_0^1$  and  $E(\bar{x}, \bar{y})$  with  $[E(\bar{x}, \bar{y})]_0^1$  if necessary, we may assume that  $\varphi(\bar{x})$  and  $E(\bar{x}, \bar{y})$  are [0, 1]-valued everywhere. Consider the formulas

$$\begin{split} &\psi(\bar{x},\bar{z}) = [\min(2-3\varphi(\bar{x}),E(\bar{x},\bar{z}))]_0^1,\\ &\rho(\bar{x},\bar{y}) = \max\left(|\varphi(\bar{x})-\varphi(\bar{y})|,\sup_{\bar{z}}|\psi(\bar{x},\bar{z})-\psi(\bar{y},\bar{z})|\right). \end{split}$$

<sup>&</sup>lt;sup>3</sup>See the discussion after Corollary 4.12 in [12].

 $\rho(\bar{x}, \bar{y})$  is the supremum of a family of pseudo-metrics and so is a pseudo-metric. Furthermore, since  $\rho(\bar{x}, \bar{y}) \geq |\varphi(\bar{x}) - \varphi(\bar{y})|$ , we may regard  $\varphi(\bar{x})$  as a formula on the imaginary sort  $H^{\omega}/\rho$ , which we will write as  $\varphi(x)$ . Let  $\pi: H^{\omega} \to H^{\omega}/\rho$  be the quotient map.

We need to argue the following: The set  $D = \{x \in H^{\omega}/\rho : \varphi(x) = 0\}$  is infinite and definable and  $\rho$  is  $\{0,1\}$ -valued on D. First to see that there are infinitely many elements in D, note that if  $\bar{b}$  and  $\bar{b}'$  are two elements of  $H^{\omega}$  in different E-classes with  $\varphi(\bar{b}) = \varphi(\bar{b}') = 0$ , then  $\psi(\bar{b}, \bar{b}) = 1$  and  $\psi(\bar{b}', \bar{b}) = 0$ , so  $\rho(\bar{b}, \bar{b}') = 1$ . Therefore  $\pi(b)$  and  $\pi(b')$  are distinct elements of D. Since there are infinitely many E-classes, this establishes that D is infinite. To show that that D is definable, fix  $\varepsilon > 0$  with  $\varepsilon < 1$  and note that if  $\varphi(\pi(\bar{a})) < \varepsilon$ , then  $\varphi(\bar{a}) = \varphi(\pi(\bar{a})) < \varepsilon$ , so we can find b in the same E class as  $\bar{a}$  such that  $\varphi(\bar{b}) = 0$ . We then have that  $\varphi(\pi(\bar{b})) = 0$  so  $\pi(b) \in D$ . Now consider  $\rho(b,\bar{a})$ . We have that  $|\varphi(b) - \varphi(\bar{a})| < \varepsilon$ . Furthermore, for any  $\bar{c} \in H^{\omega}$ , we either have that  $\varphi(\bar{c}) = 1$  or  $\varphi(\bar{c}) < 1$ . If  $\varphi(\bar{c}) = 1$ , then  $\psi(\bar{b},\bar{c})=\psi(\bar{a},\bar{c})=1$ . On the other hand if  $\varphi(\bar{c})<1$ , then  $E(\bar{b},\bar{c})$  and  $E(\bar{a},\bar{c})$  are equal, since E is a  $\{0,1\}$ -valued equivalence relation on  $\{\bar{x}:\varphi(\bar{x})<1\}$  and  $\bar{b}$  and  $\bar{a}$  are E-equivalent. Since we can do this for any sufficiently small  $\varepsilon > 0$ , we have that D is definable. We can now treat D as the required imaginary sort. It is clear that for any  $\bar{a}$  and  $\bar{b}$  with  $\varphi(\bar{a}) = \varphi(\bar{b}) = 0$ ,  $\pi(\bar{a}) = \pi(\bar{b})$  if and only if  $E(\bar{a}, \bar{b})$ , so we have that the required function f exists as well.

For 2, by [12, Lem. 2.17], we know that if T has an infinite crisp imaginary sort, then there is a definable pseudo-metric  $\rho$  on  $H^\omega$  (where H is the home sort) and an infinite definable set  $D\subseteq H^\omega/\rho$  on which  $\rho$  is  $\{0,1\}$ -valued. Let  $\delta(x)$  be the distance predicate of  $\delta$  in the imaginary sort  $H^\omega/\rho$ . Let  $q:H^\omega\to H^\omega/\rho$  be the (definable) quotient map. Let  $\varphi(\bar x)=\min(6\delta(q(\bar x)),1)$  and let  $E(\bar x,\bar y)=[1-3\rho(\bar x,y)]_0^1$ . We need to show that  $\varphi(x)$  and E(x,y) satisfy the bullet points in the statement of the proposition. For any  $\bar a$ , we have that  $\varphi(\bar a)<1$  if and only if  $\delta(q(\bar a))<\frac{1}{6}$  (i.e.,  $d(q(\bar a),D)<\frac{1}{6}$ ). For any  $\bar a$  and  $\bar b$  satisfying  $\varphi(\bar x)<1$ , we either have that  $q(\bar a)$  and  $q(\bar b)$  have  $\rho$ -distance less than  $\frac{1}{6}$  to the same point in D, in which case  $\rho(\bar a,\bar b)<1$  or they have  $\rho$ -distance less than  $\frac{1}{6}$  to different points in D, in which case  $\rho(\bar a,\bar b)>1-\frac{1}{6}-\frac{1}{6}=\frac{2}{3}$  and so  $E(\bar a,\bar b)=0$ . Evidently, this means that  $E(\bar x,\bar y)$  is a  $\{0,1\}$ -valued equivalence relation on  $\{\bar a:\varphi(\bar a)<1\}$ .

On the other hand, if  $\varphi(\bar{a}) < 1$ , then there is a  $\bar{b}$  such that  $q(\bar{b}) \in D$  (and therefore  $\varphi(\bar{b}) = 0$ ) and  $\rho(\bar{a}, \bar{b}) < \frac{1}{6}$  (and therefore  $E(\bar{a}, \bar{b}) = 1$ ). Therefore the last bullet point holds as well.

For 3 and 4, to see that we may assume that  $\varphi(\bar{x})$  and  $E(\bar{x},\bar{y})$  are restricted formulas, find restricted formulas  $\psi(\bar{x})$  and  $F(\bar{x},\bar{y})$  such that  $|\varphi(\bar{x})-\psi(\bar{x})|<\frac{1}{3}$  and  $|E(\bar{x},\bar{y})-F(\bar{x},\bar{y})|<\frac{1}{3}$  for all  $\bar{x}$  and  $\bar{y}$ . (Note that we can take  $\psi(\bar{x})$  and  $F(\bar{x},\bar{y})$  to be quantifier-free if  $\varphi(\bar{x})$  and  $E(\bar{x},\bar{y})$  are quantifier-free.) Now let  $\varphi'(\bar{x})=[3\psi(\bar{x})-1]_0^1$  and  $E'(\bar{x},\bar{y})=[3E(\bar{x},\bar{y})-1]_0^1$ . For any  $\bar{a}$ , we have that if  $\varphi(\bar{a})=0$ , then  $\varphi'(\bar{a})=0$  and if  $\varphi'(\bar{a})<1$ , then  $\varphi(\bar{x})<\frac{2}{3}$ . Furthermore, we have that for any  $\bar{a}$  and  $\bar{b}$ , if  $E(\bar{a},\bar{b})\in\{0,1\}$ , then  $E(\bar{a},\bar{b})=E'(\bar{a},\bar{b})$  as well. Therefore  $\varphi'(\bar{x})$  and  $E'(\bar{x},\bar{y})$  satisfy the bullet point as required.

The following corollary is more or less a rephrasing of part of Proposition 1.10, but it is worth emphasizing.

Corollary 1.11. A continuous theory T interprets a discrete theory with infinite models if and only if it has a pair of restricted formulas  $\varphi(\bar{x})$  and  $E(\bar{x}, \bar{y})$  satisfying

that E is a  $\{0,1\}$ -valued equivalence relation on  $\{\bar{a}: \varphi(\bar{a}) < 1\}$ , E has infinitely many equivalence classes, and every E-class of some  $\bar{a}$  with  $\varphi(\bar{a}) < 1$  contains a  $\bar{b}$  such that  $\varphi(\bar{b}) = 0$ .

One thing to note about Proposition 1.9 is that, together with Theorem 1.5, it implies that if a dictionaric theory interprets an infinite discrete structure, then it has a one-dimensional interpretation of an infinite discrete structure (i.e., an infinite discrete definable subset of a quotient of the home sort). It seems unlikely that this is always true, but we do not at present know an example.

**Question 1.12.** Does there exist a theory that interprets an infinite discrete structure but does not have a one-dimensional interpretation of an infinite discrete structure?

### 1.3. An example separating the two notions.

**Definition 1.13.** Let  $\mathcal{L}_{\heartsuit}$  be the language with a [0,1]-valued metric and a constant symbol \*. Let  $M_{\heartsuit}$  be the  $\mathcal{L}_{\heartsuit}$ -structure whose universe is [0,1] with  $d(x,y) = \max(x,y)$  for  $x \neq y$  and  $*^M = 0$ . Let  $T_{\heartsuit} = \operatorname{Th}(M_{\heartsuit})$ .

It is immediate that  $T_{\heartsuit}$  has a non-algebraic crisp type, namely the type of  $1 \in M_{\heartsuit}$  over the empty set.

### Lemma 1.14.

- (1) An  $\mathcal{L}_{\heartsuit}$ -structure M is a model of  $T_{\heartsuit}$  if and only if  $\{d(x,*^M): x \in M\}$  is dense in [0,1] and for any distinct a and b in M,  $d(a,b) = \max(d(a,*^M),d(b,*^M))$ .
- (2) Every non-algebraic type over any set of parameters is minimal. In particular,  $T_{\heartsuit}$  is weakly minimal and therefore superstable.
- (3)  $T_{\heartsuit}$  has quantifier elimination.
- (4) Any n-type  $p(\bar{x}) \in S(A)$  is uniquely determined by the values  $d(x_i, *)$  for i < n, the values of  $d(x_i, x_j)$  for i < j < n, and the values of  $d(x_i, a)$  for i < n and  $a \in A$ .
- (5) For each  $n < \omega$ , the type space  $S_n(T_{\heartsuit})$  is locally path-connected.

Proof. For 1, first note that for any formula  $\varphi(x)$  and any closed  $F \subseteq \mathbb{R}$ , there is a continuous first-order theory that is satisfied by a structure M if and only if  $\{\varphi^M(x):x\in M\}$  is dense in F. Secondly, note that the second condition is clearly satisfied by any ultrapower of  $M_{\heartsuit}$ . Therefore if  $M\models T_{\heartsuit}$ , then there is an ultrapower of  $M_{\heartsuit}$  into which M elementarily embeds, implying that for any distinct a and b in M,  $d(a,b)=\max(d(a,*^M),d(b,*^M))$ . Conversely, assume that M is an  $\mathcal{L}_{\heartsuit}$ -structure satisfying these conditions. It's easy to see that there are winning strategies for the 'approximate Ehrenfeucht-Fraïssé games' between  $M_{\heartsuit}$  and M (i.e., for a given  $\varepsilon>0$  fixed in advanced, duplicator wins if the predicates agree to within  $\varepsilon$  on the finite partial map from M to  $M_{\heartsuit}$  built in the game). This is enough to establish that  $M\equiv M_{\heartsuit}$ .

For 2, given  $A \subseteq N$ , it is clear that for any  $r \in (0,1]$ , any permutation of  $\{b \in N \setminus A : d(b,*) = r\}$  extends to an automorphism of N fixing A pointwise (namely, the automorphism that fixes the rest of N). This implies that  $T_{\heartsuit}$  is weakly minimal and therefore superstable.

For 3, note that the argument for 2 implies that for any n-tuple  $\bar{a} \in N \succeq M$ ,  $\operatorname{tp}(\bar{a}/M)$  is uniquely determined by the indices i < n for which  $a_i \in M$ , the indices

i < j < n for which  $a_i = a_j$ , and the values of  $d(a_i, *)$  for each i < n. This implies that quantifier-free types over M determine full types over M. Since we can do this for any  $M \models T_{\heartsuit}$ , we have that  $T_{\heartsuit}$  has quantifier elimination.

Finally, 4 is clear from quantifier elimination and 5 follows from 4.  $\Box$ 

**Lemma 1.15.** If  $T_{\heartsuit}$  interprets a discrete theory with infinite models, then it does so without parameters.

Proof. By Proposition 1.10, we know that if  $T_{\heartsuit}$  interprets a discrete theory with infinite models, then it has an infinite crisp imaginary sort I definable over some finite tuple  $\bar{c}$  of parameters. Let  $M \sqcup \bar{e}$  be a structure consisting of a model M of  $T_{\heartsuit}$  together with a tuple of new elements  $\bar{e}$  of the same size as  $\bar{c}$  with pairwise distance 1. Suppose furthermore  $M \sqcup \bar{e}$  has constants for the elements of  $\bar{e}$ .  $M \sqcup \bar{e}$  interprets  $(T_{\heartsuit})_{\bar{c}}$  by defining a new metric d' on its home sort given by the ordinary metric on M,  $d'(x,e_i)=|d(x,*)-d(c_i,*)|$  for each  $x\in M$ , and  $d'(e_i,e_j)=d(c_i,c_j)$ . By [14, Prop. 3.4.1, 3.4.3],  $^4T_{\heartsuit}$  interprets  $\mathrm{Th}(M\sqcup \bar{e})$  without parameters, so we have that  $T_{\heartsuit}$  interprets  $(T_{\heartsuit})_{\bar{c}}$  for any finite tuple  $\bar{c}$  of parameters.

**Lemma 1.16.** If I is a crisp imaginary sort of  $T_{\heartsuit}$  definable over  $\varnothing$ , then for every  $a \in I$ , there is a unique smallest finite set  $\bar{b}$  in the home sort such that  $a \in \operatorname{dcl}(\bar{b})$  and  $\bar{b} \in \operatorname{acl}(a)$ .

*Proof.* (This is almost identical to the proof of the analogous fact for the theory of the infinite set.)

First note that by Proposition 1.10, we may assume that I is a definable subset of a quotient of a finite power of the home sort. For any  $a \in I$ , there is some finite  $\bar{b}$  such that  $a \in \operatorname{dcl}(\bar{b})$ . We may clearly assume that  $*\notin \bar{b}$ .

Now assume that  $\bar{b}\bar{e}$  and  $\bar{b}\bar{f}$  are two different finite tuples of parameters such  $\bar{e} \cap \bar{f} = \varnothing$  and  $a \in \operatorname{dcl}(\bar{b}\bar{e}) \cap \operatorname{dcl}(\bar{b}\bar{f})$ . We need to show that  $a \in \operatorname{dcl}(\bar{b})$ . To do this it will be sufficient to show that if  $\bar{g}$  and  $\bar{h}$  are two tuples of elements realizing the same type over  $\bar{b}$ , then there is an automorphism  $\sigma \in \operatorname{Aut}(\mathbb{M}/\bar{b})$  taking  $\bar{g}$  to  $\bar{h}$  and fixing a. Clearly we may assume that  $\bar{g} \cap \bar{b} = \bar{h} \cap \bar{b} = \varnothing$ .

Let  $\bar{g}_0 = \bar{g} \cap \bar{e}$ ,  $\bar{g}_1 = \bar{g} \cap \bar{f}$ , and  $\bar{g}_2 = \bar{g} \setminus (\bar{e}\bar{f})$ . Let  $\bar{h}_0\bar{h}_1\bar{h}_2$  be a partitioning of  $\bar{f}$  satisfying that  $\bar{g}_0\bar{g}_1\bar{g}_2 \equiv_{\bar{b}} \bar{h}_0\bar{h}_1\bar{h}_2$ . Let  $\bar{g}' = \bar{g}'_0\bar{g}'_1\bar{g}'_2$  be a realization of  $\operatorname{tp}(\bar{g}_0\bar{g}_1\bar{g}_2/\bar{b})$  satisfying that  $\bar{g}'_0\bar{g}'_1\bar{g}'_2 \cup_{\bar{b}} \bar{e}\bar{f}\bar{g}\bar{h}$  (i.e., that any elements in common between  $\bar{g}'$  and  $\bar{e}\bar{f}\bar{g}$  are contained in  $\bar{b}$ ).

By Lemma 1.14, we have that  $\{\bar{g}'_0, \bar{g}'_1, \bar{g}'_2, \bar{e}\bar{g}_0, f\bar{g}_1, \bar{g}_2\}$  is pairwise independent over  $\bar{b}$ , implying that we can find an automorphism  $\sigma_0 \in \operatorname{Aut}(\mathbb{M}/\bar{b}\bar{e})$  such that  $\sigma_0(\bar{g}_1) = \bar{g}'_1$ ,  $\sigma_0$  fixes everything in  $\{\bar{g}'_0, \bar{g}'_2, \bar{e}\bar{g}_0, \bar{g}_2\}$ , and  $\{\bar{g}'_0, \bar{g}'_1\sigma_0(\bar{f}), \bar{g}'_2, \bar{e}\bar{g}_0, \bar{g}_2\}$  is still independent over  $\bar{b}$ . We can then find  $\sigma_1 \in \operatorname{Aut}(\mathbb{M}/\bar{b} \cup \sigma_0(\bar{f}))$  such that  $\sigma_1(\bar{g}_0) = \bar{g}'_0$ ,  $\sigma_1(\bar{g}_2) = \bar{g}'_2$ , and  $\sigma_1$  fixes  $\bar{g}'_1$ .

We then have that  $\sigma_1(\sigma_0(\bar{g})) = \bar{g}'$  and  $\sigma_1 \circ \sigma_0$  fixes a. By the symmetric argument, we can find an automorphism  $\tau \in \operatorname{Aut}(\mathbb{M}/\bar{b})$  taking  $\bar{h}$  to  $\bar{g}'$ , so we get that  $\tau^{-1} \circ \sigma_1 \circ \sigma_0$  is the required automorphism.

Since we can do this for any  $b\bar{e}$  and  $b\bar{f}$ , we get that a is definable over the intersection of all such finite tuples. Let  $\bar{b}$  be this minimal tuple. Since the set

<sup>&</sup>lt;sup>4</sup>The definition of imaginary sort used in [14] is slightly more permissive than the standard definition, but it is equivalent for non-trivial theories, specifically theories whose models have more than one element in some sort [14, Rem. 3.2.9].

 $\{b_0,\ldots,b_{n-1}\}$  is fixed by any automorphism fixing a and since  $\bar{b}$  is a finite tuple, we have that  $\bar{b} \in \operatorname{acl}(a)$ .

**Proposition 1.17.**  $T_{\heartsuit}$  does not interpret a discrete theory with infinite models. In particular, there is a superstable theory with a non-algebraic crisp type that does not interpret a discrete theory with infinite models.

*Proof.* By Proposition 1.10 and Lemma 1.15, if  $T_{\heartsuit}$  interprets a discrete theory with infinite models, then there is a pair of quantifier-free restricted formulas  $\varphi(\bar{x})$  and  $E(\bar{x}, \bar{y})$  such that  $E(\bar{x}, \bar{y})$  is a  $\{0, 1\}$ -valued equivalence relation on the set of tuples  $\bar{a}$  satisfying  $\varphi(\bar{a}) < 1$  with infinitely many classes such that each E-class contains an  $\bar{a}$  with  $\varphi(\bar{a}) = 0$ . Let  $n = |\bar{x}|$ .

Let I be the imaginary sort corresponding to  $\varphi(\bar{x})$  and  $E(\bar{x},\bar{y})$ . Let  $\pi:H^n\to I$  be the partial function taking each tuple  $\bar{a}$  satisfying  $\varphi(\bar{a})<1$  to the corresponding element of I. Note that  $\pi(\bar{a})\in\operatorname{dcl}(\bar{a})$ . By Lemma 1.16, we have that for any  $\bar{a}$  satisfying  $\varphi(\bar{a})<1$ , there is a unique minimal set  $B_{\bar{a}}\subseteq\bar{a}$  such that  $\pi(\bar{a})\in\operatorname{dcl}(B_{\bar{a}})$  and  $B_{\bar{a}}\subseteq\operatorname{acl}(\pi(\bar{a}))$ . For any  $\bar{a}$  satisfying  $\varphi(\bar{a})<1$ , let  $I_{\bar{a}}\subseteq n=\{0,1,\ldots,n-1\}$  be the set of indices such that  $a_i\in B_{\bar{a}}$ .

Call an ordered pair  $\langle \bar{a}, \bar{b} \rangle$  of elements critical if  $\varphi(\bar{a}) < 1$ ,  $\varphi(\bar{b}) < 1$ ,  $\bar{a} \equiv \bar{b}$ ,  $|B_{\bar{a}} \setminus B_{\bar{b}}| = 1$ , and  $a_i = b_i$  for all  $i \in n \setminus I_{\bar{a}}$ . (Note that by Lemma 1.16, critical pairs exist.) Given a critical pair  $\langle \bar{a}, \bar{b} \rangle$ , let  $c_{\bar{a},\bar{b}}$  be the unique element of  $B_{\bar{a}} \setminus B_{\bar{b}}$ . Note that for any critical pair  $\langle \bar{a}, \bar{b} \rangle$ ,  $d(\bar{a}, \bar{b}) = d(c_{\bar{a},\bar{b}}, *)$ . Also note that for any critical pair  $\langle \bar{a}, \bar{b} \rangle$ ,  $E(\bar{a}, \bar{b}) = 0$  (otherwise  $B_{\bar{a}}$  and  $B_{\bar{b}}$  would not be minimal). Finally, note that if  $\langle \bar{a}, \bar{b} \rangle$  is a critical pair, then  $\langle \bar{b}, \bar{a} \rangle$  is a critical pair as well.

Let  $r = \inf\{d(\bar{a}, \bar{b}) : \langle \bar{a}, \bar{b} \rangle$  a critical pair}. Let  $(p_k(\bar{x}, \bar{y}))_{k < \omega}$  be a sequence of 2n-types of critical pairs satisfying  $p_k(\bar{x}, \bar{y}) \vdash \varphi(\bar{x}) = 0 \land \varphi(\bar{y}) = 0 \land d(\bar{x}, \bar{y}) < r + 2^{-k}$  for each  $k < \omega$ . By compactness, we may assume that the sequences is converging to some 2n-type  $q(\bar{x}, \bar{y})$ . Clearly we have that  $q(\bar{x}, \bar{y}) \vdash \varphi(\bar{x}) = 0 \land \varphi(\bar{y}) = 0$ .

r cannot be 0. To see this, assume that r is 0. This implies that  $q(\bar{x}, \bar{y}) \vdash d(\bar{x}, \bar{y}) = 0$  and so  $q(\bar{x}, \bar{y}) \vdash E(\bar{x}, \bar{y}) = 1$  (since  $q(\bar{x}, \bar{y}) \vdash \varphi(\bar{x}) = 0 \land \varphi(\bar{y}) = 0$ ), which contradicts the fact that  $E(\bar{x}, \bar{y})$  is continuous on type space.

For any  $s \in (0,1)$ , let  $q_s(\bar{x},\bar{y})$  be the unique type satisfying  $q_s(\bar{x},\bar{y}) \vdash d(t_0,t_1) = 0$  if and only if  $p(\bar{x},\bar{y}) \vdash d(t_0,t_1) = 0$  for each pair of terms  $t_0,t_1 \in \{x_0,\ldots,x_{n-1},y_0,\ldots,y_{n-1}\}$  and satisfying that  $d(x_i,*)^{q_s} = s \cdot d(x_i,*)^p$  and  $d(y_i,*)^{q_s} = s \cdot d(y_i,*)^p$  for each i < n (where  $\psi^q$  is the value of  $\psi$  assigned by the type q).

Note that as s approaches  $1, q_s \to q$  topologically. Therefore, for sufficiently large s < 1, we have by continuity that  $q_s(\bar{x}, \bar{y}) \vdash \varphi(\bar{x}) < 1 \land \varphi(\bar{y}) < 1$ . Furthermore, for sufficiently large s < 1, we have that  $q_s(\bar{x}, \bar{y}) \vdash E(\bar{x}, \bar{y}) = 0$ .

For each  $p_k(\bar{x}, \bar{y})$ , we have for each i < j < n that

$$p_k(\bar{x}, \bar{y}) \vdash d(x_i, y_i) = 0 \lor d(x_i, y_i) = 0 \lor (d(x_i, x_i)) = 0 \land d(y_i, y_i) = 0$$

which is a closed condition and therefore preserved under taking a limit. In particular,  $q(\bar{x}, \bar{y})$  satisfies the same thing. Furthermore, each  $q_s(\bar{x}, \bar{y})$  satisfies the same thing by construction. In particular, this implies that  $q_s(\bar{x}, \bar{y})$  is the type of a critical pair. Choose some s < 1 that is sufficiently large in the preceding senses. Let  $\bar{a}\bar{b}$  be a realization of  $q_s(\bar{x}, \bar{y})$ .

By assumption, there must be some  $\bar{a}'$  satisfying  $\varphi(\bar{a}') = 0$  such that  $E(\bar{a}, \bar{a}') = 1$ . Furthermore, we can choose this  $\bar{a}'$  such that  $\bar{a}' \downarrow_{\bar{a}} \bar{b}$  (i.e., such that  $\bar{a}' \cap \bar{b} \subseteq \bar{a}$ ), implying that  $b_{\bar{b},\bar{a}} \notin \bar{a}'$ . We need to find  $\bar{b}'$  such that  $\bar{a}\bar{a}' \equiv \bar{b}\bar{b}'$  and  $\langle \bar{a}',\bar{b}' \rangle$  is a critical pair. Define  $\bar{b}'$  as follows: For each i < n, if  $a'_i = c_{\bar{a},\bar{b}}$ , let  $b'_i = c_{\bar{b},\bar{a}}$ . Otherwise, let  $b'_i = a'_i$ .

By Lemma 1.14, we have that  $\bar{a}\bar{a}' \equiv \bar{b}\bar{b}'$  and so  $\bar{a}' \equiv \bar{b}'$  and therefore  $\varphi(\bar{b}') = 0$ . Furthermore, for each i < n, if  $i \in n \setminus I_{\bar{a}'}$ , then  $a'_i \neq c_{\bar{a},\bar{b}} \in C_{\bar{a}'}$  and so  $b'_i = a'_i$ . Therefore we have that  $\langle \bar{a}', \bar{b}' \rangle$  is a critical pair.

Since  $\langle \bar{a}', \bar{b}' \rangle$  is a critical pair satisfying  $\varphi(\bar{a}') = \varphi(\bar{b}') = 0$ , we must have that  $d(\bar{a}', \bar{b}') = d(c_{\bar{a}', \bar{b}'}, *) \geq r$ , but  $d(c_{\bar{a}', \bar{b}'}, *) = sr < r$  by construction and since  $C_{\bar{a}} = C_{\bar{a}'}$  and  $C_{\bar{b}} = C_{\bar{b}'}$ , we must have that  $c_{\bar{a}', \bar{b}'} = c_{\bar{a}, \bar{b}}$ , which is a contradiction.

Therefore no such crisp imaginary sort can exist and  $T_{\heartsuit}$  does not interpret a discrete theory with infinite models.

#### 2. Generic binary predicates on $\mathbb{R}$ -forests

Here we will give our example of a strictly simple theory that has no non-algebraic crisp imaginary types.

2.1. The  $\mathbb{R}$ -forest for the  $\mathbb{R}$ -trees. The techniques used here are similar to those in [6] and [15], but neither of them are precisely applicable. Morally speaking, all of the continuous model theory results of this section are due to Carlisle and Henson in [6], which studied the model companion of the theory of rooted  $\mathbb{R}$ -trees of a given height. There the model companion was characterized as being the theory of richly branching  $\mathbb{R}$ -trees. While a similar characterization of  $\mathbb{R}$ F\* is possible, we won't need it for our purposes.

The results regarding  $\mathbb{R}$ -forests in this section are also direct adaptations of standard results in the  $\mathbb{R}$ -tree literature (such as those in [10]), but these are often stated in terms of finite sets of points and for  $\mathbb{R}$ -trees in particular. Given these issues, we will give a partially self-contained presentation.

Recall than an  $\mathbb{R}$ -tree is a uniquely arc-connected complete metric space (X, d) such that for any  $x, y \in X$ , the unique arc between x and y, written [x, y], is isometric to  $[0, d(x, y)] \subseteq \mathbb{R}$ .

**Definition 2.1.** An  $\mathbb{R}$ -forest is a complete extended metric space (X, d) with the property that for any  $x \in X$ ,  $\{y \in X : d(x, y) < \infty\}$  is an  $\mathbb{R}$ -tree.

For the sake of this section we will allow metrics to be extended (i.e.,  $[0,\infty]$ -valued). Officially we could say that our metrics are [0,1]-valued by taking  $\frac{d(x,y)}{1+d(x,y)}$  as our metric, but this is inconvenient and unnecessary to do in practice. As long as we are careful that the connectives we use with d(x,y) are continuous on  $[0,\infty]$ , there is no issue. In particular, note that addition on  $[0,\infty]$  defined in the standard way is continuous.

In any extended metric space X, given  $x \in X$ , we will write  $[x]_{<\infty}$  for the set  $\{y \in X : d(x,y) < \infty\}$ .

**Fact 2.2** ([10, Thm. 3.38]). A metric space (X, d) can be isometrically embedded into an  $\mathbb{R}$ -tree if and only if it satisfies the 4-point condition: For all  $x, y, z, w \in X$ ,  $d(x, y) + d(z, w) \leq \max(d(x, z) + d(y, w), d(y, z) + d(x, w))$ .

Let 4PC(x, y, z, w) denote the closed condition  $d(x, y) + d(z, w) \le \max(d(x, z) + d(y, w), d(y, z) + d(x, w))$ .

**Proposition 2.3.** The class of  $\mathbb{R}$ -forests is axiomatizable.

*Proof.* By the continuous version of the Keisler-Shelah theorem [11, Cor. C.5], we know that in order to establish that the class of  $\mathbb{R}$ -forests is axiomatizable, it is sufficient to show that any ultraproduct of  $\mathbb{R}$ -forests is an  $\mathbb{R}$ -forest and that for any complete extended metric space X and any ultrafilter  $\mathcal{U}$ , if the ultrapower  $X^{\mathcal{U}}$  is an  $\mathbb{R}$ -forest, then X is an  $\mathbb{R}$ -forest as well.

First let  $(X_i)_{i\in I}$  be a family of  $\mathbb{R}$ -forests and let  $\mathcal{U}$  be an ultrafilter on I. Let  $X_{\mathcal{U}}$  be the ultraproduct of the  $X_i$ 's with regards to  $\mathcal{U}$ . We need to show that for each  $x\in X_{\mathcal{U}}$ ,  $[x]_{<\infty}$  is an  $\mathbb{R}$ -tree. Suppose that for some distinct  $x,y\in X_{\mathcal{U}}$ ,  $d(x,y)<\infty$ . Let  $x=(x_i)_{i\in I}/\mathcal{U}$  and  $y=(y_i)_{i\in I}/\mathcal{U}$ . We have that for a large set of indices,  $d(x_i,y_i)<\infty$  in  $X_i$ . For any such index and any  $r\in [0,d(x,y)]$ , let  $z_{i,r}$  be the unique element of  $[x_i,y_i]$  satisfying  $d(x_i,z_{i,r})=\frac{r}{d(x,y)}\cdot d(x_i,y_i)$ . For each r, let  $z_r=(z_i,r)_{i\in I}/\mathcal{U}$ . It is straightforward to show that the function  $r\mapsto z_r$  is an isometric embedding of [0,d(x,y)] into X satisfying  $z_0=x$  and  $z_{d(x,y)}=y$ . Since we can do this for any  $x,y\in X$  with  $d(x,y)<\infty$ , we have that each  $[x]_{<\infty}$  is arc-connected.

To finish showing that X is an  $\mathbb{R}$ -forest, we need to establish that each  $[x]_{<\infty} \subseteq X$  is uniquely arc-connected. By Fact 2.2, for each  $r \in (0, \infty)$ , every  $\mathbb{R}$ -forest satisfies the following truncated form of the 4-point condition:

$$\forall x_0 x_1 x_2 x_3 \left( \bigwedge_{i < j < 4} d(x_i, x_j) < r \to 4 \text{PC}(x_0, x_1, x_2, x_3) \right).$$

These conditions are expressible as closed conditions in continuous logic, so  $X_{\mathcal{U}}$  satisfies these as well. Since this holds for each  $r \in (0, \infty)$ , we have that for any  $x \in X_{\mathcal{U}}$ ,  $[x]_{<\infty}$  can be embedded in an  $\mathbb{R}$ -tree. This implies that for any  $x, y \in X$  with  $d(x, y) < \infty$ , there is at most one arc between x and y. Since we have already established that there is an arc between x and y, we have that each  $[x]_{<\infty} \subseteq X_{\mathcal{U}}$  is uniquely arc-connected. Therefore  $X_{\mathcal{U}}$  is an  $\mathbb{R}$ -forest.

Now fix a complete extended metric space X and an ultrafilter  $\mathcal{U}$  and suppose that  $X^{\mathcal{U}}$  is an  $\mathbb{R}$ -forest. Identify X with its image in  $X^{\mathcal{U}}$ . Clearly we have that for each  $x \in X$ ,  $[x]_{<\infty}$  embeds into an  $\mathbb{R}$ -tree, so we just need to show that X is arc connected. Fix  $x, y \in X$  with d(x, y) and consider  $[x, y] \subseteq X^{\mathcal{U}}$ . For all ultrapowers Y of  $X^{\mathcal{U}}$ , Y is an  $\mathbb{R}$ -forest by the above. Therefore for any automorphism of any such Y fixing x and y, the set  $[x, y] \subseteq X^{\mathcal{U}}$  is fixed pointwise. Therefore  $[x, y] \subseteq \operatorname{dcl}\{x, y\}$  in  $X^{\mathcal{U}}$ . Since Y is an elementary submodel containing x and y, it therefore must be the case that  $[x, y] \subseteq Y$  as well. Therefore X is arc-connected. Since we can do this for any  $x, y \in X$  with  $d(x, y) < \infty$ , we have that X is an  $\mathbb{R}$ -forest.  $\square$ 

**Definition 2.4.** Let  $\mathbb{R}\mathsf{F}$  denote the common theory of all  $\mathbb{R}$ -forests.

**Proposition 2.5.** For any  $\mathbb{R}$ -forests  $X \subseteq Y$  and any  $y \in Y$ , if  $d(y, X) < \infty$ , then there is a unique  $x \in X$  such that d(y, x) = d(y, X).

Proof. Fix  $z \in X$  with finite distance to y. By compactness of [z,y], there is a unique  $x_z \in [z,y] \cap X$  of minimal distance to y. To finish the proposition, all we need to do is show that  $x_{z_0} = x_{z_1}$  for any  $z_0, z_1 \in X$  with finite distance to Y. By [10, Lem. 3.20, 3.22], for any such  $z_0$  and  $z_1$ , there is a unique  $w \in Y$  such that  $[x_{z_0},y] \cap [x_{z_1},y] = [w,y]$  which furthermore satisfies that  $w \in [x_{z_0},x_{z_1}]$ . Therefore  $w \in X$ , which implies that  $x_{z_0} \in [w,y]$  and  $x_{z_1} \in [w,y]$ , but this is only possible if  $x_{z_0} = x_{z_1} = w$ .

**Definition 2.6.** Given  $\mathbb{R}$ -forests  $X \subseteq Y$ , we write  $\pi_X$  for the partially defined function that takes an element of Y to the unique closest point in X if it exists.

**Proposition 2.7.**  $Mod(\mathbb{RF})$  has the joint embedding property and the amalgamation property.

*Proof.* Given two  $\mathbb{R}$ -forests X and Y, the space  $X \sqcup Y$  (where all new distances are set to  $\infty$ ) is clearly an  $\mathbb{R}$ -forest. Therefore  $\operatorname{Mod}(\mathbb{R}\mathsf{F})$  has the joint embedding property.

To see that  $\operatorname{Mod}(\mathbb{RF})$  has the amalgamation property, fix an  $\mathbb{R}$ -forest X and two  $\mathbb{R}$ -forest extensions  $Y_0 \supseteq X$  and  $Y_1 \supseteq X$ . We may assume that  $Y_0 \setminus X$  and  $Y_1 \setminus X$  are disjoint. For any  $y_0 \in Y_0$  and  $y_1 \in Y_1$ , let  $\rho(y_0, y_1) = \inf_{x_0, x_1 \in X} d(y_0, x_0) + d(x_0, x_1) + d(x_1, y_1)$ . Clearly if either  $d(y, X_0) = \infty$  or  $d(y, X_1) = \infty$ , then  $\rho(y_0, y_1) = \infty$ . By Proposition 2.5, we have that if  $d(y_0, X) < \infty$  and  $d(y_1, X) < \infty$ , then

$$\rho(y_0, y_1) \le d(y_0, \pi_X(y_0)) + d(\pi_X(y_0), \pi_X(y_1)) + d(\pi_X(y_1), y_1),$$

but by [10, Lem. 3.20, 3.22]  $\rho(y_0, y_1)$  is actually equal to the second quantity. This implies that the unique metric on  $Y_0 \cup Y_1$  extending the existing metrics and  $\rho$  is a metric and that the resulting space is an  $\mathbb{R}$ -forest.

The following Fact 2.8 is a direct generalization of some classic preservation results, and the proofs are essentially the same (see for example Theorem 6.5.9 and Exercise 9 in Section 6.5 in [16]). A continuous theory T has an  $\forall \exists$ -axiomatization if it can be axiomatized by conditions of the form  $\sup_x \inf_y \varphi \leq 0$  with  $\varphi$  quantifier-free.

## Fact 2.8. Fix a continuous theory T.

- (1) The following are equivalent:
  - (a) T has an  $\forall \exists$ -axiomatization.
  - (b) For any chain  $(M_i)_{i<\alpha}$  of models of T,  $\overline{\bigcup_{i<\alpha} M_i} \models T$ .
- (2) Suppose that for any models  $A, B, C \models T$  with  $B, C \subseteq A$  and  $B \cap C \neq \emptyset$ ,  $B \cap C \models T$ . Then the following hold:
  - (a) If  $A \models T$  and  $(B_i)_{i \in I}$  is a family of substructures of A such  $B_i \models T$  for each  $i \in I$  and  $\bigcap_{i \in I} B_i$  is non-empty, then  $\bigcap_{i \in I} B_i \models T$ .
  - (b) T has an  $\forall \exists$ -axiomatization.

### Proposition 2.9.

- (1) Given an  $\mathbb{R}$ -forest Y and a family  $(X_i)_{i\in I}$  of substructures of Y that are themselves  $\mathbb{R}$ -forests,  $\bigcap_{i\in I} X_i$  is an  $\mathbb{R}$ -forest.
- (2) For any  $\mathbb{R}$ -forest Y and any  $X_0 \subseteq Y$ , there is a unique smallest  $\mathbb{R}$ -forest  $X \subseteq Y$  such that  $X_0 \subseteq X$ .
- (3)  $\mathbb{R}\mathsf{F}$  has an  $\forall \exists$ -axiomatization.

*Proof.* 1 is immediate from the definition of  $\mathbb{R}$ -forest. 2 follows by considering the intersection of all  $\mathbb{R}$ -forests  $Z \subseteq Y$  such that  $X_0 \subseteq Z$ . 3 follows from 1 and Fact 2.8.

**Definition 2.10.** Given an  $\mathbb{R}$ -forest Y and a set  $X_0 \subseteq Y$ , we will write  $\langle X_0 \rangle$  for the unique smallest  $\mathbb{R}$ -forest  $X \subseteq Y$  such that  $X_0 \subseteq X$ .

Call an  $\mathbb{R}$ -forest X finitely generated if there is a finite set  $X_0 \subseteq X$  such that  $X = \langle X_0 \rangle$ .

Technically in the above definition we ought to specify the ambient space in our notation, but in Proposition 2.12 we will see that this is unnecessary.

**Lemma 2.11.** For any  $\mathbb{R}$ -forest A and set  $B \subseteq A$ ,

$$\langle B \rangle = \overline{ \left[ \ \, \int \{ [b_0,b_1] : b_0,b_1 \in B, \ d(b_0,b_1) < \infty \} }.$$

*Proof.* First we will establish the statement of the lemma in the special case that B is finite. We will show this by induction. If  $|B| \leq 2$ , then it is clearly true, so assume that for some  $n \geq 2$ , the lemma holds for all B with  $|B| \leq n$ . Let B be a subset of A with |B| = n and fix some  $b \in B$ . If  $d(b, B) = \infty$ , then  $\langle Bb \rangle = \langle B \rangle \cup \{b\}$ , so we are done. Otherwise, if  $d(b, B) < \infty$ , let  $c \in B$  be an element with  $d(b, c) < \infty$ . We now have that  $\langle B \rangle \cup [c, b]$  is the unique minimal  $\mathbb{R}$ -forest containing Bb.

Now to finish the argument we will show that for any  $B \subseteq A$ ,  $\langle B \rangle = \overline{\bigcup \{\langle B_0 \rangle : B_0 \subseteq B, B_0 \text{ finite}\}}$ . Let  $C = \bigcup \{\langle B_0 \rangle : B_0 \subseteq B, B_0 \text{ finite}\}$ . Clearly  $B \subseteq C \subseteq \langle B \rangle$ . By Fact 2.8 and Proposition 2.9,  $\overline{C}$  is an  $\mathbb{R}$ -forest. Therefore  $\overline{C} = \langle B \rangle$ .

**Proposition 2.12.** For any  $\mathbb{R}$ -forests A and B and any tuples  $\bar{a} \in A$  and  $\bar{b} \in B$  of the same (possibly infinite) length  $\alpha$ , if  $d(a_i, a_j) = d(b_i, b_j)$  for all  $i < j < \alpha$ , then there is a unique isometry  $f: \langle \bar{a} \rangle \to \langle \bar{b} \rangle$  satisfying  $f(a_i) = b_i$  for each  $i < \alpha$ .

*Proof.* It is a standard fact and not hard to show that for any subset A of an  $\mathbb{R}$ -tree,  $\langle A \rangle$  is the injective envelope<sup>5</sup> of A, written e(A). By [17, Thm. 2.1], for any metric spaces X and Y with an isometry  $f: X \to Y$ , there is a unique isometry from e(X) to e(Y) extending f. Since  $\langle \bar{a} \rangle = \bigcup_{i < |\bar{a}|} \langle \bar{a} \cap [a_i]_{<\infty} \rangle$ , the full result follows immediately.

**Definition 2.13.** Call an  $\mathbb{R}$ -forest X type-existentially closed or t.e.c. if for any finitely generated  $\mathbb{R}$ -forest  $Y \subseteq X$  and any finitely generated  $\mathbb{R}$ -forest  $Z \supseteq Y$ , there is an isometric embedding  $f: Z \to X$  fixing Y pointwise.

#### Proposition 2.14.

- (1) Every  $\mathbb{R}$ -forest A has an  $\mathbb{R}$ -forest extension  $B \supseteq A$  that is t.e.c.
- (2) If A and B are t.e.c.  $\mathbb{R}$ -forests and  $\bar{a} \in A$  and  $\bar{b} \in B$  are finite n-tuples such that  $d(a_i, a_j) = d(b_i, b_j)$  for each i < j < n, then  $(A, \bar{a})$  and  $(B, \bar{b})$  are back-and-forth equivalent and therefore  $(A, \bar{a}) \equiv (B, \bar{b})$ . In particular, there is a common theory of all t.e.c.  $\mathbb{R}$ -forests.
- (3) Any ultrapower of a t.e.c.  $\mathbb{R}$ -forest is t.e.c.
- (4) The common theory of t.e.c.  $\mathbb{R}$ -forests has quantifier-elimination and is the model companion of  $\mathbb{R}$ F.

*Proof.* For 1, first note that by Fact 2.8 and Proposition 2.9, closures of unions of chains of  $\mathbb{R}$ -forests are  $\mathbb{R}$ -forests. By Proposition 2.7, this means that given an  $\mathbb{R}$ -forest A, we can build a chain  $(A_i)_{i<|A|}$  of  $\mathbb{R}$ -forests with the property that any finitely generated extension C of any finitely generated substructure B of A has an embedding into  $\overline{\bigcup_{i<|A|}A_i}$  that fixes B. Repeating this, we can build a chain  $(A'_j)_{j<\omega_1}$  of  $\mathbb{R}$ -forests extending A with the property that  $A'_{j+1}$  embeds any finitely generated substructure of a finitely generated substructure of  $A'_i$ . The union is then

<sup>&</sup>lt;sup>5</sup>Also known as the metric envelope, tight span, or hyperconvex hull.

metrically closed (since  $\omega_1$  has uncountable cofinality) and so is a t.e.c. extension of A.

For 2, fix  $(A, \bar{a})$  and  $(B, \bar{b})$  as in the statement of the proposition. By Proposition 2.12,  $\langle \bar{a} \rangle$  and  $\langle \bar{b} \rangle$  are isometric by an isometry f satisfying  $f(a_i) = b_i$  for each  $i < |\bar{a}|$ . Back-and-forth equivalence is immediate from the definition of type-existential closure and Proposition 2.7. Finally the fact that back-and-forth implies elementary equivalence is proven in the same manner as it is in discrete logic.

For 3, fix an  $\mathbb{R}$ -forest A and an ultrafilter  $\mathcal{U}$  on an index set I. Fix a finite n-tuple  $\bar{a} \in A^{\mathcal{U}}$  and assume that  $d(a_j, a_k) < \infty$  for each j < k < n. Let  $(\bar{a}^i)_{i \in I}$  be a family of tuples in A corresponding to  $\bar{a}$ . We may assume without loss of generality that for each  $i \in I$ ,  $d(a_j^i, a_k^i) < \infty$  for each  $j < k < \omega$ . Let B be a finitely generated  $\mathbb{R}$ -tree extending  $\langle \bar{a} \rangle$ . By Proposition 2.5, we know that  $B \setminus \langle \bar{a} \rangle$  has finitely many connected components  $B_0, \ldots, B_{m-1}$  with the property that for each j < m,  $\overline{B_j}$  is  $B_j$  together with a single point  $c_j \in \langle \bar{a} \rangle$ . For each  $c_j$ , we can find k(j,0) and k(j,1) less than n such that  $c_j \in [a_{k(j,0)}, a_{j(j,1)}]$ . Let  $r_j \in [0,1]$  be the unique value such that  $d(a_{k(j,0)}, c_j) = r_j \cdot d(a_{k(j,1)}, c_j)$ .

For each  $i \in I$ , we can find points  $c^i_j$  such that  $d(a^i_{k(j,0)}, c^i_j) = r_j \cdot d(a^i_{k(j,1)}, c^i_j)$ . By the amalgamation property and the fact that A is t.e.c., we can find  $B^i_0, \ldots, B^i_{m-1}$  disjoint from  $\langle \bar{a}^i \rangle$  such that for each  $j < m, B^i_j \cup \{c^i_j\}$  is an  $\mathbb{R}$ -tree. Let  $\bar{b}$  be a finite tuple in B such that  $\langle \bar{a}\bar{b} \rangle = B$  and  $\bar{b} \cap \langle \bar{a} \rangle = \varnothing$ . For  $i \in I$ , let  $\bar{b}^i$  be the images of  $\bar{b}$  under the isometries between  $B_j$  and  $B^i_j$ . Finally let  $\bar{b}^{\mathcal{U}}$  be the element of  $A^{\mathcal{U}}$  corresponding to the family  $(\bar{b}^i)_{i \in I}$ . By construction we have that  $\bar{a}\bar{b}$  and  $\bar{a}\bar{b}^{\mathcal{U}}$  have the same pairwise distances, so by Proposition 2.12, there is an isometry taking  $B = \langle \bar{a}\bar{b} \rangle$  to  $\bar{a}\bar{b}^{\mathcal{U}}$  that fixes every element of  $\langle \bar{a} \rangle$ .

Now finally for an arbitrary tuple  $\bar{a}$  perform the above construction for each finite distance class of  $\bar{a}$  and for any connected component C of an extension  $B \supseteq \langle \bar{a} \rangle$  not connected to  $\langle \bar{a} \rangle$ , embed C into some finite distance class of A not used elsewhere in the embedding. Since we can do this for any finite  $\bar{a}$  and any finitely generated  $B \supseteq \langle \bar{b} \rangle$ , we have that  $A^{\mathcal{U}}$  is t.e.c.

For 4, note that by 2 and Proposition 2.12, we have that the type of any tuple  $\bar{a}$  in a t.e.c.  $\mathbb{R}$ -forest is determined by its quantifier-free type. By 3 this happens in some  $\aleph_0$ -saturated model of the theory. Therefore the theory has quantifier elimination. Since this theory has quantifier elimination, it is the model companion of  $\mathbb{R}\mathsf{F}$  by 1.

# **Definition 2.15.** Let $\mathbb{R}F^*$ denote the model companion of $\mathbb{R}F$ .

Recall that in a type space  $S_1(A)$ , the induced metric on  $S_1(A)$ , written  $\partial$ , is defined by  $\partial(p,q) = \inf\{d(a,b) : a \models p, b \models q\}$ . This concept is still well defined when we allow metric structures with extended metrics.

# **Proposition 2.16.** $\mathbb{R}F^*$ is strictly stable.

*Proof.* By Propositions 2.5 and 2.14, we know that given any model A of  $\mathbb{R}F^*$  we know that there is a unique type in  $S_1(A)$  of an element with infinite distance to any element of A and for any b in the monster with finite distance to A,  $\operatorname{tp}(b/A)$  is uniquely determined by  $\pi_A(b)$  and d(b,A). This immediately gives that  $\mathbb{R}F^*$  is stable. To see that it is strictly stable, note that for any  $b_0, b_1$  both with finite distance to A, we have that  $\partial(\operatorname{tp}(b_0/A), \operatorname{tp}(b_1/A)) = |d(b_0, A) - d(b_1, A)|$  if  $\pi_A(b_0) = \pi_A(b_1)$  and  $\partial(\operatorname{tp}(b_0/A), \operatorname{tp}(b_1/A)) = d(b_0, A) + d(\pi_A(b_0), \pi_A(b_1)) + d(b_1, A) = d(b_0, b_1)$  if

 $\pi_A(b_0) \neq \pi_A(b_1)$ . This immediately allows us to find for any infinite  $\kappa$  a model A of  $\mathbb{R}\mathsf{F}^*$  and a family  $(b_i)_{i<\kappa^{\aleph_0}}$  of elements such that  $\partial(\operatorname{tp}(b_i/A),\operatorname{tp}(b_j/A)) \geq 1$  for each  $i < j < \kappa^{\aleph_0}$ , implying that  $\mathbb{R}\mathsf{F}^*$  is not superstable.

**Proposition 2.17.** For any sets A, B, and C of parameters,  $B \bigcup_A C$  if and only if  $\langle AB \rangle \cap \langle AC \rangle = \langle A \rangle$ .

*Proof.* Given Proposition 2.7, we clearly have that if  $B \downarrow_A C$ , then  $\langle ABC \rangle$  needs to be the free amalgamation of  $\langle AB \rangle$  and  $\langle AC \rangle$  over  $\langle A \rangle$ , implying that  $\langle AB \rangle \cap \langle AC \rangle = \langle A \rangle$ . By quantifier elimination and stationarity of independence, this implies the converse as well.

### 2.2. Generic 1-1-Lipschitz binary predicates.

**Definition 2.18.** A function  $f: X^2 \to X$  is 1-1-Lipschitz if for any  $a \in X$ ,  $x \mapsto f(a,x)$  and  $x \mapsto f(x,a)$  are 1-Lipschitz functions.

Note that this condition is equivalent to being 1-Lipschitz in the  $\ell^1$ -metric (i.e.,  $d(x_0x_1,y_0y_1)=d(x_0,y_0)+d(x_1,y_1)$ ), which is strictly weaker than being 1-Lipschitz in the max metric. Clearly any 1-1-Lipschitz function is 2-Lipschitz with regards to the max metric. A fact which we will use several times is that if  $(f_i)_{i\in I}$  is any family of 1-1-Lipschitz [0,1]-valued functions on a metric space, then  $\sup_{i\in I} f_i(x,y)$  and  $\inf_{i\in I} f_i(x,y)$  are also 1-1-Lipschitz and [0,1]-valued.

**Definition 2.19.** Let  $\mathcal{L}_{\mathbb{RFR}}$  be the language containing an extended metric and a single 2-Lipschitz [0,1]-valued binary predicate R.

Let  $\mathbb{R}FR$  be the theory that says that  $(A, d) \models \mathbb{R}F$  and R is 1-1-Lipschitz.

Note that by Proposition 2.9,  $\mathbb{R}FR$  has an  $\forall \exists$ -axiomatization.

The proof of the following proposition makes it clear that we need the 1-1-Lipschitz condition to ensure that the class of models of  $\mathbb{R}\mathsf{FR}$  has the amalgamation property.

**Proposition 2.20.**  $Mod(\mathbb{R}FR)$  has the joint embedding property and the amalgamation property.

*Proof.* Now suppose that A,  $B_0$ , and  $B_1$  are models of  $\mathbb{R}\mathsf{FR}$  with  $A \subseteq B_0$  and  $A \subseteq B_1$ . Let C be the free amalgamation of the underlying  $\mathbb{R}$ -forests of  $B_0$  and  $B_1$  over A. We just need to show that we can extend R to a 1-1-Lipschitz function on C. Let

$$Q(x,y) = [\inf\{R(z,w) + d(x,z) + d(y,w) : R(z,w) \text{ defined}\}]_0^1$$

Since the function  $(x,y) \mapsto R(z,w) + d(x,z) + d(y,w)$  is 1-1-Lipschitz for each fixed z and w, we immediately have that Q(x,y) is 1-1-Lipschitz on C. It is also [0,1]-valued by construction. We just need to argue that for any pair (x,y), if R(x,y) is defined, then Q(x,y) = R(x,y). To do this we just need to show that  $R(x,y) \leq R(z,w) + d(y,w) + d(z,w)$  for each (z,w) at which R is defined. If  $\{x,y,z,w\} \subseteq B_0$  or  $\{x,y,z,w\} \subseteq B_1$ , then this is immediate from the fact that R is 1-1-Lipschitz, so assume that  $\{x,y\} \subseteq B_0$  and  $\{z,w\} \subseteq B_1$ . If any of the elements of  $\{x,y,z,w\}$  have infinite distance to A, then the inequality is trivial, so assume that all four of them have finite distances. We now have that  $d(x,z) = \frac{1}{2} \sum_{i=1}^{n} \frac{1}{2} \sum_{i=1$ 

 $d(x, \pi_A(x)) + d(\pi_A(x), \pi_A(z)) + d(\pi_A(z), z)$  and likewise for y and w. Finally, we have that

$$|R(x,y) - R(\pi_A(x), \pi_A(y))| \le d(x, \pi_A(x)) + d(y, \pi_A(y)),$$

$$|R(\pi_A(x), \pi_A(y)) - R(\pi_A(z), \pi_A(w))| \le d(\pi_A(x), \pi_A(z)) + d(\pi_A(y), \pi_A(w)),$$

$$|R(\pi_A(z), \pi_A(w)) - R(z, w)| \le d(\pi_A(z), z) + d(\pi_A(w), w),$$

whereby |R(x,y) - R(z,w)| is less than or equal to the sum of the three quantities on the right, which is precisely d(x,z) + d(y,w). Therefore  $R(x,y) \leq R(z,w) + d(x,z) + d(y,w)$ , as required.

The joint embedding property follows by amalgamating over the empty model of  $\mathbb{R}\mathsf{FR}$ .

**Definition 2.21.** Call a model A of  $\mathbb{R}\mathsf{FR}$  type-existentially closed or t.e.c. if for any finitely generated  $B \subseteq A$  with  $B \models \mathbb{R}\mathsf{FR}$  and any finitely generated  $C \supseteq B$  with  $C \models \mathbb{R}\mathsf{FR}$ , there is an  $\mathcal{L}_{\mathbb{R}\mathsf{FR}}$ -isomorphism  $f: C \to A$  fixing B pointwise.

#### Proposition 2.22.

- (1) Any model of  $\mathbb{R}FR$  has an extension that is a t.e.c. model of  $\mathbb{R}FR$ .
- (2) If A and B are t.e.c. models of  $\mathbb{R}\mathsf{FR}$  and  $\bar{a} \in A$  and  $\bar{b} \in B$  are finite n-tuples such that there is an  $\mathcal{L}_{\mathbb{R}\mathsf{FR}}$ -isomorphism  $f: \langle \bar{a} \rangle \to \langle \bar{b} \rangle$  satisfying  $f(a_i) = b_i$  for each i < n, then  $(A, \bar{a})$  and  $(B, \bar{b})$  are back-and-forth equivalent and therefore  $(A, \bar{a}) \equiv (B, \bar{b})$ . In particular, there is a common theory of all t.e.c. models of  $\mathbb{R}\mathsf{FR}$ .

*Proof.* 1 and 2 follow by the same argument as parts 1 and 2 of Proposition 2.14, using Proposition 2.20.  $\Box$ 

**Definition 2.23.** Given two sets A and B, a correlation between A and B is a total, surjective relation  $O \subseteq A \times B$  (i.e., a relation such that  $(\forall a \in A)(\exists b \in B)O(a,b)$  and  $(\forall b \in B)(\exists a \in A)O(a,b)$ ). Given sets A and B and n-tuples  $\bar{a} \in A$  and  $\bar{b} \in B$ , we'll write  $\operatorname{cor}(A, \bar{a}; B, \bar{b})$  for the set of correlations O between A and B satisfying that for each i < n,  $O(a_i, b_i)$ . We will write  $\operatorname{cor}(A, B)$  for  $\operatorname{cor}(A, \varnothing; B, \varnothing)$ .

Given two extended metric structures A and B and a correlation  $O \in cor(A, B)$ , the K-truncated d-distortion<sup>6</sup> of O is the quantity

$$\operatorname{dis}_{d}^{K}(O) = \sup_{O(a,b), O(a',b')} \left| [d(a,a')]_{0}^{K} - [d(b,b')]_{0}^{K} \right|.$$

Given two  $\mathcal{L}_{\mathbb{RFR}}$ -structures A and B and a correlation  $O \in \text{cor}(A, B)$ , the Ktruncated distortion of O is the quantity

$$\operatorname{dis}^K(O) = \max\left(\operatorname{dis}_d^K(O), \sup_{O(a,b),O(a',b')} |R(a,a') - R(b,b')|\right).$$

Given tuples  $\bar{a} \in A$  and  $\bar{b} \in B$  of the same length, the *K-truncated*  $\mathcal{L}_{\mathbb{RFR}}$ -distance between  $(A, \bar{a})$  and  $(B, \bar{b})$  is the quantity

$$\rho_K(A, B) = \inf_{O \in \operatorname{cor}(A, \bar{a}; B, \bar{b})} \operatorname{dis}^K(A, B).$$

Given an n-tuple  $\bar{a}$  in an  $\mathbb{R}$ -forest, we write  $\langle \bar{a} \rangle_K$  for the set  $\bigcup \{[a_j, a_k] : j, k < n, \ d(a_j, a_k) < K\}$ . We will say that K is large for  $\bar{a}$  if  $K \ge 1$  and for any  $j, k < n, \ d(a_j, a_k) < \infty$  implies that  $d(a_j, a_k) < K$ .

<sup>&</sup>lt;sup>6</sup>Distortion for metrics is normally defined with a factor of  $\frac{1}{2}$ , but we will not need this here.

Note that for any finite tuple  $\bar{a}$  in an  $\mathbb{R}$ -forest, there is a K such that  $\langle \bar{a} \rangle_K = \langle \bar{a} \rangle$ .

**Definition 2.24.** Given  $a, b \in A \models \mathbb{R}FR$  with  $d(a, b) < \infty$  and  $r \in [0, 1]$ , we will write a:r:b for the unique element of [a, b] satisfying  $d(a, a:r:b) = r \cdot d(a, b)$ .

Recall that by Lemma 2.11, for any tuple  $\bar{a}$ , elements of the form  $a_j:r:a_k$  cover  $\langle \bar{a} \rangle$ . Also, note that for any K > 0 and  $r \in [0,1]$ ,  $(a,b) \mapsto a:r:b$  is a definable function on the definable set<sup>7</sup> of pairs (a,b) with  $d(a,b) \leq K$ .

**Lemma 2.25.** Fix extended metric spaces A and B.

- (1) For any subset  $A_0 \subseteq A$ , correlation  $O \in \text{cor}(A_0, B)$ , and 1-Lipschitz function  $f: B \to [0, 1]$ , there is a 1-Lipschitz function  $g: A \to [0, 1]$  such that for any  $(a, b) \in O$  and any  $K \ge 1$ ,  $|g(a) f(b)| \le \text{dis}_d^K(O)$ .
- (2) For any subset  $A_0 \subseteq A$  and 1-Lipschitz functions  $f: A_0 \to [0,1]$  and  $g: A \to [0,1]$ , there is a 1-Lipschitz function  $h: A \to [0,1]$  extending f such that  $\sup_{a \in A} |h(a) g(a)| = \sup_{a \in A_0} |f(a) g(a)|$ .
- (3) For any  $O \in \text{cor}(A, B)$ , let  $O^{(2)} = \{((a, a'), (b, b')) : (a, b), (a', b') \in O\}$  and let  $d^{+2}$  be the metric  $d^{+2}(xy, zw) = d(x, z) + d(y, w)$ . For any K,  $\operatorname{dis}_{d+2}^{K}(O^{(2)}) \leq 2 \operatorname{dis}_{d}^{K}(O)$  (where  $\operatorname{dis}_{d+2}^{K}$  is computed with regards to the metric  $d^{+2}$ ).

Proof. For 1, let  $g(x) = \inf_{(a,b) \in O} [d(x,a) + f(b)]_0^1$ . g(x) is 1-Lipschitz and [0,1]-valued by construction. We also clearly have that for any  $(a,b) \in O$ ,  $g(a) \leq f(b)$ . Fix  $(a',b') \in O$ . If  $d(a,a') \geq 1$ , then  $d(a,a') + f(b') \geq 1$ . Otherwise, if d(a,a') < 1, we have that  $[d(b,b')]_0^K \leq d(a,a') + \operatorname{dis}_d^K(O)$ . Since f is 1-Lipschitz, it is 1-Lipschitz with regards to the metric  $[d]_0^K$ , so we have that  $f(b') \geq f(b) - [d(b,b')]_0^K$ . Therefore  $f(b') \geq f(b) - d(a,a') - \operatorname{dis}_d^K(O)$  and so  $d(a,a') + f(b') \geq f(b) - \operatorname{dis}_d^K(O)$ . Since we can do this for any  $(a',b') \in O$ , we have that  $g(a) \geq f(b) - \operatorname{dis}_d^K(O)$ , whence  $|g(a) - f(b)| \leq \operatorname{dis}_d^K(O)$ .

For 2, let  $r = \sup_{a \in A_0} |f(a) - g(a)|$ . Let

$$h(x) = \min\left(g(x) + r, \inf_{a \in A_0} [d(x, a) + f(a)]_0^1\right).$$

Clearly we have that  $h(a) \leq g(a) + r$  for any  $a \in A$ . Also note that h(x) is 1-Lipschitz and [0,1]-valued by construction. By the same argument as in 1, we have that for any  $a \in A_0$ ,  $\inf_{a' \in A_0} [d(a,a') + f(a')]_0^1 = f(a)$ . By the choice of r, this implies that h(x) extends f(x). We just need to show that  $h(a) \geq g(a) - r$  for each  $a \in A$ . For any  $a' \in A_0$ , we have that  $f(a') \geq g(a') - r$ , whence

$$\inf_{a' \in A_0} [d(a, a') + f(a')]_0^1 \ge \inf_{a' \in A_0} [d(a, a') + g(a') - r]_0^1$$

$$\ge \inf_{a' \in A_0} [d(a, a') + g(a')]_0^1 - r$$

$$\ge [g(a)]_0^1 - r = g(a) - r.$$

Therefore, since  $g(a) + r \ge g(a) - r$ , we have that  $h(a) \ge g(a) - r$ , as required. 3 is immediate from the relevant definitions.

<sup>&</sup>lt;sup>7</sup>This is not in general a definable set in an arbitrary metric structure but it is always definable in  $\mathbb{R}$ -forests with a distance predicate of  $\max(d(x,y)-K,0)$ .

**Lemma 2.26.** Fix an n-tuple  $\bar{a}$  and an (n+1)-tuple  $\bar{b}c$  in models of  $\mathbb{R}\mathsf{FR}$ . If K is large for  $\bar{b}c$ , then for any  $\varepsilon > 0$ , there is a model  $\langle \bar{a}e \rangle \models \mathbb{R}\mathsf{FR}$  extending  $\langle \bar{a} \rangle$  such that  $d(e, \langle \bar{a} \rangle) = \infty$  if and only if  $d(c, \langle \bar{b} \rangle) = \infty$  and

$$\rho_K(\langle \bar{a}e \rangle_K, \bar{a}e; \langle \bar{b}c \rangle, \bar{b}c) \le 4\rho_K(\langle \bar{a} \rangle_K, \bar{a}; \langle \bar{b} \rangle, \bar{b}) + \varepsilon.$$

*Proof.* Let  $\rho = \rho_K(\langle \bar{a} \rangle_K, \bar{a}; \langle \bar{b} \rangle, \bar{b})$ . Fix  $\varepsilon > 0$  and fix a correlation  $O \in \text{cor}(\langle \bar{a} \rangle_K, \bar{a}; \langle \bar{b} \rangle, \bar{b})$  with  $\text{dis}^K(O) \leq \rho + \frac{1}{4}\varepsilon$ . There are two cases to consider; either  $d(c, \langle \bar{b} \rangle) = \infty$  or  $d(c, \langle \bar{b} \rangle) < \infty$ .

If  $d(c, \langle \bar{b} \rangle) = \infty$ , then we will make  $\langle \bar{a} \rangle \cup \{e\}$  into a model of  $\mathbb{R}\mathsf{FR}$  by first setting  $d(e, x) = \infty$  for any  $x \in \langle \bar{a} \rangle$ . We then immediately have that  $\mathrm{dis}_d^K(O \cup \{(e, c)\}) = \mathrm{dis}_d^K(O)$ . The functions  $x \mapsto R(x, c)$  and  $x \mapsto R(c, x)$  are 1-Lipschitz and [0, 1]-valued on  $\langle \bar{b} \rangle$ . By Lemma 2.25 part 1, we can find 1-Lipschitz function  $f : \langle \bar{a} \rangle \to [0, 1]$  and  $g : \langle \bar{a} \rangle \to [0, 1]$  satisfying that for any  $(a, b) \in O$ ,  $|f(a) - R(b, c)| \leq \rho$  and  $|g(a) - R(c, b)| \leq \rho$ . (In particular, we take  $A = \langle \bar{a} \rangle$ ,  $A_0 = \langle \bar{a} \rangle_K$ , and  $B = \langle \bar{b} \rangle$ .) We then have that defining R(a, e) = f(a), R(e, a) = g(a), and R(e, e) = R(c, c) gives a model of  $\mathbb{R}\mathsf{FR}$  satisfying  $\rho_K(\langle \bar{a}e \rangle_K, \bar{a}e; \langle \bar{b}c \rangle, \bar{b}c) < \rho + \varepsilon$ .

If  $d(c, \langle \bar{b} \rangle) < \infty$ , first note that  $d(c, \langle \bar{b} \rangle) < K$ . Fix some  $a^{\dagger} \in \langle \bar{a} \rangle_K$  such that  $(a^{\dagger}, \pi_{\langle \bar{b} \rangle}(c)) \in O$ . Let  $b^{\dagger} = \pi_{\langle \bar{b} \rangle}(c)$ . Build an  $\mathbb{R}$ -forest  $\langle \bar{a}e \rangle$  extending  $\langle \bar{a} \rangle$  by taking the free amalgamation that is a copy of  $[0, d(b^{\dagger}, c)]$  glued to  $\langle \bar{a} \rangle$  by identifying 0 with  $a^{\dagger}$  (and letting e be the point  $d(b^{\dagger}, c)$  in  $[0, d(b^{\dagger}, c)]$ ). Let O' be  $O \cup \{(a^{\dagger}: r: e, b^{\dagger}: r: c) : r \in [0, 1]\}$ . O' is a correlation between  $\langle \bar{a}e \rangle_K$  and  $\langle \bar{b}c \rangle$ . For any  $a \in \langle \bar{a} \rangle_K$ ,  $a' \in \langle \bar{a}e \rangle_K \setminus \langle \bar{a} \rangle_K$ ,  $b \in \langle \bar{b} \rangle$ , and  $b' \in \langle \bar{b}c \rangle \setminus \langle \bar{b} \rangle$  with  $(a, b), (a', b') \in O'$ , we have that

$$\begin{split} |[d(a,a')]_0^1 - [d(b,b')]_0^1| &\leq |[d(a,a^\dagger)]_0^1 - [d(b,b^\dagger)]_0^1| + |[d(a^\dagger,a')]_0^1 - [d(b^\dagger,b')]_0^1| \\ &\leq |[d(a,a^\dagger)]_0^1 - [d(b,b^\dagger)]_0^1| + 0 \leq \operatorname{dis}_d^K(O). \end{split}$$

Therefore  $\operatorname{dis}_d^K(O') \leq \operatorname{dis}_d^K(O)$ . Since  $\operatorname{dis}_d^K(O) \leq \operatorname{dis}_d^K(O')$ , we have that  $\operatorname{dis}_d^K(O') = \operatorname{dis}_d^K(O)$ .

Let  $d^{+2}$  be the metrics on  $A=\langle \bar{a}e\rangle^2$  and  $B=\langle \bar{b}c\rangle^2$  defined by  $d^{+2}(xy,zw)=d(x,z)+d(y,w)$ . Let  $O'^{(2)}=\{((a,a'),(b,b')):(a,b),(a',b')\in O'\}$ . By Lemma 2.25 part 3, we have that  $\operatorname{dis}_{d^{+2}}^K(O'^{(2)})\leq 2\operatorname{dis}_d^K(O')\leq 2\rho+\varepsilon$ . Recall that 1-1-Lipschitz functions on  $\langle \bar{a}e\rangle$  are the same thing as 1-Lipschitz functions on  $(A,d^{+2})$ . By Lemma 2.25 part 1, we can find a 1-1-Lipschitz function  $f:\langle \bar{a}e\rangle^2\to [0,1]$  satisfying that for any  $(a,b),(a',b')\in O',|f(a,a')-R(b,b')|\leq 2\rho+\frac{1}{2}\varepsilon$ . By Lemma 2.25 part 2, we can find a 1-1-Lipschitz function  $Q:\langle \bar{a}e\rangle^2$  extending  $R:\langle \bar{a}\rangle^2\to [0,1]$  such that  $\sup_{a,a'\in\langle \bar{a}e\rangle}|Q(a,a')-f(a,a')|=\sup_{a,a'\in\langle \bar{a}e\rangle_K}|R(a,a')-f(a,a')|\leq 2\rho+\frac{1}{2}\varepsilon$ . Therefore, we now have that for any  $(a,b),(a',b')\in O',$ 

$$\begin{aligned} |Q(a,a') - R(b,b')| &\leq |Q(a,a') - f(a,a')| + |f(a,a') - R(b,b')| \\ &\leq 2\rho + \frac{1}{2}\varepsilon + 2\rho + \frac{1}{2}\varepsilon = 4\rho + \varepsilon. \end{aligned}$$

Since we can do this for any such (a,b) and (a',b'), we have that  $\operatorname{dis}^K(O') \leq \max(\operatorname{dis}_d^K(O'), 4\rho + \varepsilon) = 4\rho + \varepsilon$ . Therefore  $\rho_K(\langle \bar{a}e \rangle_K, \bar{a}e; \langle \bar{b}c \rangle, \bar{b}c) \leq 4\rho + \varepsilon$ , as required.

**Lemma 2.27.** For any  $A \models \mathbb{R}\mathsf{FR}$ , index set I, ultrafilter  $\mathcal{U}$  on I, and family  $(\bar{a}^i)_{i \in I}$  of n-tuples with limit  $\bar{a}$  in the ultrapower  $A^{\mathcal{U}}$ , if K > 0 is large for  $\bar{a}$ , then

$$\lim_{i \to \mathcal{U}} \rho_K(\langle \bar{a}^i \rangle_K, \bar{a}^i; \langle \bar{a} \rangle; \bar{a}) = 0.$$

*Proof.* We may assume without loss of generality that for any  $i \in I$  and j, k < n, if  $d(a_j, a_k) < \infty$ , then  $d(a_j^i, a_k^i) < K$  and if  $d(a_j, a_k) = \infty$ , then  $d(a_j^i, a_k^i) > 2K$ .

Given  $i \in I$ , let  $O^i \subseteq \langle \bar{a}^i \rangle_K \times \langle \bar{a} \rangle$  be given by

$$O^i = \{(a_j^i : r : a_k^i, a_j : r : a_k) : j, k < n, \ d(a_j, a_k) < \infty, \ r \in [0, 1]\}.$$

By Lemma 2.11,  $O^i$  is a correlation between  $\langle \bar{a}^i \rangle_K$  and  $\langle \bar{a} \rangle$ . The result now follows from the fact that  $(x,y,r) \mapsto x:r:y$  is uniformly definable and Lipschitz on  $\{(x,y):d(x,y)\leq K\}\times [0,1]$  (with the max metric and a Lipschitz constant that depends on K) as well as the fact that  $\bar{a}$  is a finite tuple. In particular, for each quadruple  $j,k,\ell,m < n$  with  $d(a_j,a_k) < \infty$  and  $d(a_\ell,a_m) < \infty$ ,

$$\begin{cases} n \text{ with } d(a_{j}, a_{k}) < \infty \text{ and } d(a_{\ell}, a_{m}) < \infty, \\ \lim_{i \to \mathcal{U}} \sup_{r, s \in [0, 1]} |[d(a_{j}^{i} : r : a_{k}^{i}, a_{\ell}^{i} : s : a_{k}^{i})]_{0}^{K} - [d(a_{j} : r : a_{k}, a_{\ell} : s : a_{k})]_{0}^{K}| = 0, \\ \lim_{i \to \mathcal{U}} \sup_{r, s, \in [0, 1]} |R(a_{j}^{i} : r : a_{k}^{i}, a_{\ell}^{i} : s : a_{k}^{i}) - R(a_{j} : r : a_{k}, a_{\ell} : s : a_{k})| = 0.$$

For part of the following proof note that if  $\rho_K(\langle \bar{a} \rangle_K, \bar{a}; \langle \bar{b} \rangle, \bar{b}) = 0$  for some finite tuples  $\bar{a}$  and  $\bar{b}$ , then by compactness it is actually the case that  $(\langle \bar{a} \rangle_K, [d]_0^K, R, \bar{a})$  and  $(\langle \bar{b} \rangle, [d]_0^K, R, \bar{b})$  (where  $[d]_0^K$  is the metric  $[d(x, y)]_0^K$ ) are isomorphic as pointed  $\mathcal{L}_{\mathbb{RFR}}$ -structures [13, Cor. 1.14 (iii)].

**Proposition 2.28.** Any ultrapower of a t.e.c. model of  $\mathbb{R}FR$  is t.e.c.

*Proof.* Fix a t.e.c. model A, an index set I, and an ultrafilter  $\mathcal{U}$  on I. Fix an n-tuple  $\bar{a} \in A^{\mathcal{U}}$  corresponding to a family  $(\bar{a}^i)_{i \in I}$  and fix some finitely generated extension  $\langle \bar{a}\bar{b} \rangle \supseteq \langle \bar{a} \rangle$ . Let  $\bar{b}$  be an m-tuple. Fix a K that is large for  $\bar{a}\bar{b}$ . By passing to a smaller set of indices if necessary, we may assume that for any j,k < n, if  $d(a_j,a_k) < \infty$ , then  $d(a_j^i,a_k^i) < K$  and if  $d(a_j,a_k) = \infty$ , then  $d(a_j^i,a_k^i) > 2K$  for every  $i \in I$ .

Fix  $i \in I$ . There are three cases.

- Case 1. If  $\langle \bar{a}^i \rangle$  is isomorphic to  $\langle \bar{a} \rangle$ , then since A is t.e.c., we can find a  $\bar{b}^i \in A$  such that  $(\langle \bar{a}^i \bar{b}^i \rangle, \bar{a}^i \bar{b}^i)$  is isomorphic to  $(\langle \bar{a} \bar{b} \rangle, \bar{a} \bar{b})$ . In this case let  $L^i = 0$ .
- Case 2. If  $\langle \bar{a}^i \rangle$  is not isomorphic to  $\langle \bar{a} \rangle$  but  $\rho_K(\langle \bar{a}^i \rangle_K, \bar{a}^i; \langle \bar{a} \rangle, \bar{a}) = 0$ , then it must be the case that there are j,k < n such that  $d(a_j,a_k) = \infty$  but  $d(a_j^i,a_k^i) < \infty$  (by the discussion before the statement of the proposition). Let  $L^i$  be  $\min\{d(a_j^i,a_k^i):d(a_j,a_k)=\infty\}$ . (Note that  $L^i>2K$ .) By applying Lemma 2.26 inductively m times, we can find a model  $\langle \bar{a}^i \bar{b}^i \rangle \models \mathbb{R}$ FR extending  $\langle \bar{a}^i \rangle$  such that  $\rho_K(\langle \bar{a}^i \bar{b}^i \rangle_K, \bar{a}^i \bar{b}^i; \langle \bar{a}\bar{b} \rangle, \bar{a}\bar{b}) \leq \frac{1}{L^i}$  and for any j < m, if  $d(b_j, \langle \bar{a} \rangle) = \infty$ , then  $d(b_j^i, \langle \bar{a}^i \rangle) = \infty$ .
- Case 3. If  $\rho_K(\langle \bar{a}^i \rangle_K, \bar{a}^i; \langle \bar{a} \rangle, \bar{a}) > 0$ , then by applying Lemma 2.26 inductively m times, we have that for each  $i \in I$ , there is a model  $\langle \bar{a}^i \bar{b}^i \rangle \models \mathbb{R} \mathsf{FR}$  extending  $\langle \bar{a}^i \rangle$  such that  $\rho_K(\langle \bar{a}^i \bar{b}^i \rangle_K, \bar{a}^i \bar{b}^i; \langle \bar{a} \bar{b} \rangle, \bar{a} \bar{b}) \leq 5^m \rho_K(\langle \bar{a}^i \rangle_K, \bar{a}^i; \langle \bar{a} \rangle, \bar{a})$  and for any j < m, if  $d(b_j, \langle \bar{a} \rangle) = \infty$ , then  $d(b_j^i, \langle \bar{a}^i \rangle) = \infty$ .

Since A is t.e.c., we may furthermore assume that each  $\langle \bar{a}^i \bar{b}^i \rangle$  is actually a substructure of A extending  $\langle \bar{a}^i \rangle$ . Let  $\bar{b}^{\mathcal{U}}$  be the elements of  $A^{\mathcal{U}}$  corresponding to the family  $(\bar{b}^i)_{i \in I}$ . One of the 3 cases must happen on a  $\mathcal{U}$ -large set of indices. If case 1 happens on a  $\mathcal{U}$ -large set, then we clearly have that  $(\langle \bar{a}\bar{b}^{\mathcal{U}}\rangle, \bar{a}\bar{b}^{\mathcal{U}})$  is isomorphic to  $(\langle \bar{a}\bar{b}\rangle, \bar{a}\bar{b})$ . If case 2 happens on a  $\mathcal{U}$ -large set, then we must have that  $\lim_{i \to \mathcal{U}} L^i = \infty$  and therefore  $\lim_{i \to \mathcal{U}} \rho_K(\langle \bar{a}^i\bar{b}^i\rangle_K, \bar{a}^i\bar{b}^i; \langle \bar{a}\bar{b}\rangle, \bar{a}\bar{b}) = 0$ . Likewise, if case 3 happens on

a  $\mathcal{U}$ -large set, then  $\lim_{i\to\mathcal{U}} \rho_K(\langle \bar{a}^i \rangle_K, \bar{a}^i; \langle \bar{a} \rangle, \bar{a}) = 0$  by Lemma 2.27 and therefore  $\lim_{i\to\mathcal{U}} \rho_K(\langle \bar{a}^i \bar{b}^i \rangle_K, \bar{a}^i \bar{b}^i; \langle \bar{a} \bar{b} \rangle, \bar{a} \bar{b}) = 0$  as well.

By the discussion just before the statement of the proposition, we have that in these second two cases, the pointed  $\mathcal{L}_{\mathbb{RFR}}$ -structures  $(\langle \bar{a}\bar{b}^{\mathcal{U}}\rangle, [d]_0^K, R, \bar{a}\bar{b}^{\mathcal{U}})$  and  $(\langle \bar{a}\bar{b}\rangle, [d]_0^K, R, \bar{a}\bar{b})$  are isomorphic. Furthermore, we know that for any  $b_j \in \bar{b}$ , if  $d(b_j, \langle \bar{a}b_{< j}\rangle) = \infty$ , then we ensured that  $d(b_j^i, \langle \bar{a}b_{< j}^i\rangle) = \infty$ . This implies that  $(\langle \bar{a}\bar{b}^{\mathcal{U}}\rangle, \bar{a}\bar{b}^{\mathcal{U}})$  and  $(\langle \bar{a}\bar{b}\rangle, \bar{a}\bar{b})$  are actually isomorphic as pointed  $\mathcal{L}_{\mathbb{RFR}}$ -structures. Since  $\langle \bar{a}\rangle \subseteq \operatorname{dcl}(\bar{a})$ , we have that the isomorphism witnessing this fixes  $\langle \bar{a}\rangle$  pointwise.

So, since we can do this for any  $\langle \bar{a} \rangle \subseteq A^{\mathcal{U}}$  and any finite extension  $\langle \bar{a}\bar{b} \rangle \supseteq \langle \bar{a} \rangle$ , we have that  $A^{\mathcal{U}}$  is t.e.c.

**Definition 2.29.** Let  $\mathbb{R}FR^*$  be the common theory of t.e.c. models of  $\mathbb{R}FR$ .

**Corollary 2.30.**  $\mathbb{R}\mathsf{FR}^*$  is the model companion of  $\mathbb{R}\mathsf{FR}$ . In any model of this theory, the type of a tuple  $\bar{a}$  is uniquely determined by the pointed  $\mathcal{L}_{\mathbb{R}\mathsf{FR}}$ -isomorphism type of  $(\langle \bar{a} \rangle, \bar{a})$ .

*Proof.* By Proposition 2.28, there is an  $\aleph_0$ -saturated t.e.c. model A of  $\mathbb{RFR}^*$ . By Proposition 2.22, for any finite  $\bar{a} \in A$ , we have that  $\operatorname{tp}(\bar{a})$  is uniquely determined by the pointed isomorphism type of  $(\langle \bar{a} \rangle, \bar{a})$ .

Fix an arbitrary model B of  $\mathbb{R}\mathsf{FR}^*$ . By the continuous Keisler-Shelah theorem [11, Cor. C.5], there is an ultrapower  $B^{\mathcal{U}}$  of B that is isomorphic to an ultrapower of A. Therefore  $B^{\mathcal{U}}$  is t.e.c., by Proposition 2.28. For any finite tuple  $\bar{b} \in B$ , we have that  $\langle \bar{b} \rangle$  computed in B is the same as  $\langle \bar{b} \rangle$  computed in  $B^{\mathcal{U}}$  (since  $\langle \bar{b} \rangle \subseteq \operatorname{dcl}(\bar{b})$ ). Therefore  $\operatorname{tp}(\bar{b})$  is again uniquely determined by the pointed isomorphism type of  $(\langle \bar{b} \rangle, \bar{b})$ . This implies that if C and D are models of  $\mathbb{R}\mathsf{FR}^*$  such that  $C \subseteq D$ , then  $C \subseteq D$ . Therefore, by Proposition 2.22 part 1,  $\mathbb{R}\mathsf{FR}^*$  is the model companion of  $\mathbb{R}\mathsf{FR}$ .

**Proposition 2.31.** The reduct of  $\mathbb{R}FR^*$  to the pure metric language is  $\mathbb{R}F^*$ .

*Proof.* If  $A \models \mathbb{R}\mathsf{FR}^*$  and  $B \supseteq A$  is an  $\mathbb{R}$ -forest, then by Lemma 2.26, for any finite  $\bar{a} \in A$  and  $\bar{b} \in B$ ,  $\langle \bar{a}\bar{b} \rangle$  can be expanded to a model of  $\mathbb{R}\mathsf{F}$  extending  $\langle \bar{a} \rangle$ . (Specifically, we can consider the model of  $\mathbb{R}\mathsf{F}$  whose underlying  $\mathbb{R}$ -forest is  $\langle \bar{a}\bar{b} \rangle$  which has R(x,y)=0 for all x and  $y.^8$ ) By compactness and Proposition 2.9, this implies that B can be expanded to a model of  $\mathbb{R}\mathsf{F}\mathsf{R}$ .

This implies that for any model  $A_0 \models \mathbb{RFR}^*$  we can build two sequence  $(A_i)_{i < \omega}$  and  $(B_i)_{i < \omega}$  such that  $A_0 \subseteq B_0 \subseteq A_1 \subseteq B_1 \subseteq \ldots$  and for each  $i < \omega$ ,  $A_i \models \mathbb{RFR}^*$  and  $B_i \models \mathbb{RF}^*$ . We then have that the completion of the union is a model of  $\mathbb{RFR}^*$  whose pure metric reduct is a model of  $\mathbb{RFR}^*$ , whence the theory  $\mathbb{RF}^*$  is the pure metric reduct of the theory  $\mathbb{RFR}^*$ .

# 2.3. Simplicity itself.

**Proposition 2.32.**  $\mathbb{R}FR^*$  is unstable.

*Proof.* The  $\mathcal{L}_{\mathbb{RFR}}$ -structure whose universe is  $\{a_i, b_i : i < \omega\}$  with all non-zero distances  $\infty$ ,  $R(a_i, a_j) = R(b_i, b_j) = 0$  for all  $i, j < \omega$ ,  $R(a_i, b_j) = 0$  if i < j, and  $R(a_i, b_j) = 1$  if  $i \geq j$  is a model of  $\mathbb{RFR}$ . It can be embedded in a model of  $\mathbb{RFR}^*$ , which is therefore unstable.

<sup>&</sup>lt;sup>8</sup>This can also be proven directly without using Lemma 2.26 relatively easily, but we have already proven a strictly harder statement, so there is no point is bothering to write this out.

Simplicity in the context of compact abstract theories was defined in [2]. This of course applies to the special case of continuous logic. In particular, we have the following.

Fact 2.33 ([2, Thm. 1.51]). A theory T is simple if and only if there is a ternary relation  $\downarrow^*$  of small subsets of the monster  $\mathbb O$  satisfying the following:

- (Invariance) For any  $\sigma \in Aut(\mathbb{O})$  and any A, B, and C,  $B \downarrow^* {}_{\Delta}C$  if and only if  $\sigma B \bigcup_{\sigma A}^* \sigma C$ .
- (Existence)  $B \bigcup_{M}^{*} M$  for all models M.
- (Finite character)  $B \bigcup_{A}^{*} C$  if and only if  $B_0 \bigcup_{A}^{*} C_0$  for all finite  $B_0 \subseteq B$
- (Symmetry)  $B \bigcup_A^* C$  if and only if  $C \bigcup_A^* B$ . (Restricted transitivity) For models  $M \preceq N$ ,  $B \bigcup_M NC$  if and only if  $B \downarrow_M N$  and  $B \downarrow_N C$ .
- (Local character) There exists a  $\lambda$  such that whenever  $cf(\mu) > \lambda + |B|$  and  $(M_i)_{i<\mu}$  is an increasing sequence of models, there is some  $j<\mu$  such that  $B \bigcup_{M_i}^* \bigcup_{i < \mu} M_i$ .
- (Extension) If  $B \bigcup_{A}^{*} C$  and  $C' \supseteq C$ , then there is  $B' \equiv_{A} B$  such that
- (The independence theorem over models) For any model M, if  $B_0 \bigcup_{M}^* B_1$ ,  $C_0 \downarrow_M^* B_0$ ,  $C_1 \downarrow_M^* B_1$ , and  $C_0 \equiv_M C_1$ , then there is a  $C \downarrow_M^* B_0 B_1$  with  $B \equiv_{MA_i} B_i$  for both i < 2.

Furthermore, in this case \(\)\* coincides with non-dividing.

**Lemma 2.34.** If A is an  $\mathbb{R}$ -forest,  $B, C \subseteq A$  are  $\mathbb{R}$ -forests such that  $A = B \cup C$ , and  $f: A \to [0,1]$  is a function that is 1-Lipschitz on B and 1-Lipschitz in C, then f is 1-Lipschitz on all of A.

*Proof.* Fix  $b \in B$  and  $c \in C$ . If  $d(b, B \cap C) = \infty$  or  $d(c, B \cap C) = \infty$ , then trivially  $|f(b)-f(c)| \leq d(b,c) = \infty$ , so assume that  $d(b,B\cap C) < \infty$  and  $d(c,B\cap C) < \infty$ . Let  $D = B \cap C$ . We now have that

$$|f(b) - f(c)| \le |f(b) - f(\pi_D(b))| + |f(\pi_D(b)) - f(\pi_D(c))| + |f(\pi_D(c)) - f(c)|$$
  
 
$$\le d(b, \pi_D(b)) + d(\pi_D(b), \pi_D(c)) + d(\pi_D(c), c) = d(b, c).$$

Therefore f is 1-Lipschitz on all of A.

**Proposition 2.35.** Let  $B \downarrow_A^* C$  be the ternary relation  $\langle AB \rangle \cap \langle AC \rangle = \langle A \rangle$ .  $\downarrow^*$ is a simple independence relation in  $\mathbb{R}FR^*$  and so  $\mathbb{R}FR^*$  is simple and  $B \bigcup_A C$  if and only if  $\langle AB \rangle \cap \langle AC \rangle = \langle A \rangle$ .

*Proof.* By Proposition 2.31, the pure metric reduct of  $\mathbb{R}FR^*$  is  $\mathbb{R}F^*$ . By Proposition 2.31, the pure metric reduct of  $\mathbb{R}FR^*$  is  $\mathbb{R}F^*$ . tion 2.16,  $\mathbb{R}\mathsf{F}^*$  is stable, and by Proposition 2.17,  $\downarrow^*$  is the relation of non-dividing in  $\mathbb{R}F^*$ . This immediately implies that  $\int_0^*$  satisfies all of the properties in Fact 2.33 except for the independence theorem over models.

So fix a model M and sets  $B_0$ ,  $B_1$ ,  $C_0$ , and  $C_1$  such that  $B_0 \downarrow_M^* B_1$ ,  $C_i \downarrow_M^* B_i$  for both i < 2, and  $C_0 \equiv_M C_1$ . Let C be an  $\mathbb{R}$ -forest extending  $\langle MB_0B_1 \rangle$  in the same way that  $C_0$  and  $C_1$  do. (This exists by stationarity of non-dividing in the stable reduct.) There is clearly a well-defined function  $Q: \langle MB_0B_1C\rangle^2 \to [0,1]$  defined in such a way that  $(\langle MB_iC \rangle, Q, MB_iC)$  is isomorphic as a pointed  $\mathcal{L}_{\mathbb{RFR}}$ -structure to  $(\langle MB_iC_i\rangle,R)$  for both i<2. Therefore the only thing we need to check is that Q is 1-1-Lipschitz. By Lemma 2.34, the functions  $x\mapsto Q(x,b)$  and  $x\mapsto Q(b,x)$  are 1-Lipschitz for any  $b\in\langle MB_0B_1C\rangle$  (since they are 1-Lipschitz on  $\langle MB_0C\rangle$  and  $\langle MB_1C\rangle$ ). Therefore Q is 1-1-Lipschitz and we can assume that Q is R for some model of  $\mathbb{R}FR$  extending  $\langle MB_0B_1\rangle$ .

Therefore  $\downarrow^*$  is a simple independence relation and  $\mathbb{R}FR^*$  is simple.

2.4.  $\mathbb{R}\mathsf{FR}^*$  has no non-algebraic crisp imaginary types. We will now use Corollary 1.6 to show that  $\mathbb{R}\mathsf{FR}^*$  has no non-algebraic crisp imaginary types. For any type  $p(x) \in S_1(A)$ , let  $F_p$  be  $\{q(x,y) \in S_2(A) : q(x,y) \supseteq p(x) \cup p(y)\}$ .

**Lemma 2.36.** For any A, any non-algebraic 1-type  $p(x) \in S_1(A)$ , and any  $q(x,y) \in F_p$ , there is a type  $r(x,y) \in F_p$  and a continuous path  $f: [0,1] \to F_p$  such that f(0) = q and f(1) = r and for any  $bc \models r$ ,  $b \downarrow_A c$ .

Proof. Since  $\langle A \rangle \subseteq \operatorname{dcl}(A)$ , we may assume that  $A = \langle A \rangle$ . By Proposition 2.7 and Corollary 2.30, we know that given b,  $\operatorname{tp}(b/A)$  is uniquely determined by the pointed isomorphism type of  $(\langle Ab \rangle, Ab)$ . Fix b and c not in A such that  $b \equiv_A c$ . Since  $b \equiv_A c$ , we either have that  $d(A,b) = d(A,c) = \infty$  or  $d(A,b) = d(A,c) < \infty$  and  $\pi_A(b) = \pi_A(c)$ . In the first case, however, we have that  $\langle Ab \rangle \cap \langle Ac \rangle = \langle A \rangle$ , so  $b \downarrow_A c$  and we can take f to be a constant path. In the second case, if  $\langle Ab \rangle \cap \langle Ac \rangle = \langle A \rangle$ , then again  $b \downarrow_A c$  and so we can again take f to be a constant path. Otherwise, there is a unique  $e \in \langle Ab \rangle \cap \langle Ac \rangle$  such that  $\langle Ab \rangle \cap \langle Ac \rangle = \langle Ae \rangle$ . Let  $a = \pi_A(b) = \pi_A(c)$ .

For each  $r \in [0, d(a, e))$ , build  $b_r$  and  $c_r$  realizing  $\operatorname{tp}(b/A) = \operatorname{tp}(c/A)$  and  $e_r$  satisfying that  $d(a, e_r) = r$ ,  $\langle Ab_r \rangle \cap \langle Ac_r \rangle = \langle Ae_r \rangle$ , and for any  $s \in [0, 1]$  and  $t \in [0, 1]$ ,  $R(a:s:b_r, a:t:c_r) = R(a:s:b, a:t:c)$ . Note that the above completely specifies a completion of  $\operatorname{tp}(b/A) \cup \operatorname{tp}(c/A)$ . It is also straightforward to check that the described behavior of R is 1-1-Lipschitz. Furthermore, we have that  $b_0 \downarrow_A c_0$ , since  $\langle Ab_0 \rangle \cap \langle Ac_0 \rangle = \langle Ae_0 \rangle = \langle A \rangle$ .

So now we just need to verify that the function  $f(r) = \operatorname{tp}(b_r c_r/A)$  is continuous (where  $b_{d(a,e)} c_{d(a,e)} = bc$ ). If  $(r_i)_{i<\omega}$  is a sequence of elements of [0,1] and  $\mathcal{U}$  is an ultrafilter on  $\omega$ , then it is fairly immediate that in an ultrapower of any model containing A and  $\{b_r, c_r : r \in [0,1]\}$ , a 2-tuple b'c' corresponding to the family  $(b_{r_i}c_{r_i})_{i<\omega}$  will have the property that  $(\langle Ab'c'\rangle, b'c')$  is isomorphic to  $(\langle Ab_sc_s\rangle, Ab_sc_s)$  as a pointed  $\mathcal{L}_{\mathbb{RFR}}$ -structure, where  $s = \lim_{i \to \mathcal{U}} r_i$ . By Corollary 2.30, this implies that  $b'c' \equiv_A b_sc_s$ . Since we can do this for any  $s \in [0,1]$ , we have that f is continuous.

For the following lemma, recall that any convex combination of 1-Lipschitz functions is 1-Lipschitz. This implies that any convex combination of 1-1-Lipschitz functions is 1-1-Lipschitz as well.

**Lemma 2.37.** Fix  $p(x) \in S_x(A)$ . Fix  $q_0(x,y), q_1(x,y) \in F_p$  such that for both i < 2, if  $bc \models q_i$ , then  $b \downarrow_A c$ . There is a continuous function  $g : [0,1] \to F_p$  such that  $g(0) = q_0$  and  $g(1) = q_1$ .

Proof. Find  $b_0c_0 \models q_0$  and  $b_1c_1 \models q_1$ . If  $d(A,b_0) = \infty$ , then for each  $r \in (0,1)$ , find  $b_r$  and  $c_r$  both realizing p such that  $R(b_r,c_r) = (1-r)R(b_0,c_0) + rR(b_1,c_1)$  and  $R(c_r,b_r) = (1-r)R(c_0,b_0) + rR(c_1,b_1)$ . If  $d(A,b_0) < \infty$ , then let  $a = \pi_{\langle A \rangle}(b_0) = \pi_{\langle A \rangle}(c_0) = \pi_{\langle A \rangle}(b_1) = \pi_{\langle A \rangle}(c_1)$ . For each  $r \in (0,1)$ , find  $b_r$  and  $c_r$  both realizing p

such that for any  $s, t \in [0, 1]$ ,

$$R(a:s:b_r, a:t:c_r) = (1-r)R(a:s:b_0, a:t:c_0) + rR(a:s:b_1, a:t:c_0).$$

Since convex combinations of 1-1-Lipschitz functions are 1-1-Lipschitz, this specifies a complete type in  $F_p$  for each  $r \in (0,1)$ .

In both cases we can use the same argument as in Lemma 2.36 to show that the function  $g(r) = \operatorname{tp}(b_r c_r/A)$  is continuous.

We don't need this but one thing to note about Lemmas 2.36 and 2.37 is that they do not actually require that b and c realize the same type over A.

**Theorem 2.38.**  $\mathbb{R}FR^*$  is a simple, unstable continuous theory with no non-algebraic crisp imaginary types.

*Proof.* RFR\* is unstable by Proposition 2.32 and simple by Proposition 2.35. By Lemmas 2.36 and 2.38,  $F_p$  is path connected for any 1-type  $p(x) \in S_1(A)$  for any set of parameters A. Therefore by Corollary 1.6, RFR\* has no non-algebraic crisp imaginary types.

Given the discussion in the introduction of [9], it is natural to wonder whether adding a generic automorphism to  $\mathbb{R}F^*$  results in a strictly simple theory.

**Question 2.39.** Does the theory of models of  $\mathbb{R}F^*$  with an automorphism have a model companion? If it does, is this theory strictly simple?

#### 3. A STRONGER NOTION

There is a sense in which models of  $\mathbb{R}\mathsf{F}^*$  are 'less essentially continuous' than, say, Hilbert spaces or the Urysohn sphere. In particular, given any point a in a model of  $\mathbb{R}\mathsf{F}^*$  and  $r \in (0, \infty)$ , the set  $\{x: d(x, a) = r\}$  is definable and is an ultrametric space with a dense distance set. While as we have shown, the induced structure on this ultrametric space does not interpret an infinite discrete structure, there is still an abundance of type-definable and co-type-definable equivalence relations whose equivalence classes are uniformly metrically separated (specifically,  $d(x,y) \leq r$  and d(x,y) < r for any r). Similarly, there is also the discrete equivalence relation given by  $d(x,y) < \infty$  in any  $\mathbb{R}$ -forest. In some sense, rather than having too few uniformly metrically discrete equivalence relations to interpret an infinite discrete structure,  $\mathbb{R}\mathsf{F}\mathsf{R}^*$  has too many. By contrast, non-compact ultrametric structures with nowhere dense distance sets<sup>9</sup> interpret infinite discrete structures (see [12, Sec. 5.2.3]).

A salient feature differentiating the  $\mathbb{R}$ -arboreal theories (i.e.,  $\mathbb{R}$ -trees and  $\mathbb{R}$ -forests) from Hilbert spaces, atomless probability algebras, and the Urysohn sphere is connectivity. In particular, the theories of infinite dimensional Hilbert spaces and atomless probability algebras enjoy the following property:

(\*) For any small partial type  $\Sigma(\bar{x})$  (in finitely many variables), the metric space of realizations of  $\Sigma(\bar{x})$  in the monster has a bounded number of connected components.

One thing to note is that  $(\star)$  rules out any theory (with non-compact models) whose models are all locally compact, including the 'inherently non-discrete' strongly minimal theories of [12, Sec. 5.1].

<sup>&</sup>lt;sup>9</sup>The distance set of T is the set of distances d(a,b) of pairs of elements of models of T.

Regardless,  $(\star)$  seems reasonably natural but also harder to characterize than the property of having no non-algebraic crisp imaginary types. We can say a little about it, though.

**Proposition 3.1.** Let T be a theory with monster model  $\mathbb{O}$ . The following are equivalent.

- (1) T has  $(\star)$ .
- (2) For every model M and every complete type  $p(\bar{x}) \in S(M)$ ,  $p(\mathbb{O})$  has a bounded number of connected components.
- (3) For every model M and every complete type  $p(\bar{x}) \in S(M)$ ,  $p(\mathbb{O})$  has a single connected component.

Furthermore, if T has  $(\star)$ , then it has no non-algebraic crisp imaginary types.

*Proof.*  $1 \Rightarrow 2$  is obvious.

 $abla 3 \Rightarrow \neg 2$ . Suppose that for some model M and some complete type  $p(\bar{x}) \in S(M)$ , there are realizations a and b of p such that a and b are in different connected components of  $p(\mathbb{O})$ . Let c realize some M-coheir  $q(\bar{x})$  of  $p(\bar{x})$  over Mab. Since being in the same connected component of  $p(\mathbb{O})$  is an invariant equivalence relation, we must have that either c is not in the same connected component as a or that c is not in the connected component as b. By invariance, this implies that c is not in the connected components of either a or b. Now any Morley sequence  $(c_i)_{i < \kappa}$  in  $q(\bar{x})$  will have the property that for any  $i < j < \kappa$ ,  $c_i$  and  $c_j$  are not in the same connected component of  $p(\mathbb{O})$ , implying that  $p(\mathbb{O})$  has an unbounded number of connected components.

 $egline{1}{1}{1}{2}{3}{2}{3}$ . Suppose that there is a small partial type  $\Sigma(\bar{x})$  such that  $\Sigma(\mathbb{O})$  has a large number of connected components. Fix a model M containing all of the parameters of  $\Sigma(\bar{x})$ . Let  $(a_i)_{i<(2^{|M|+|T|})^+}$  be a collection of elements of  $\mathbb{O}$  in pairwise distinct connected components of  $\Sigma(\mathbb{O})$ . By the pigeonhole principle, there is a completion  $p(\bar{x}) \supseteq \Sigma(\bar{x})$  realized by  $a_i$  and  $a_j$  for  $i < j < (2^{|M|+|T|})^+$ .

For the final statement, note that if T has a non-algebraic crisp imaginary type, then by Theorem 1.5, there is a complete type p(x) and a  $\{0,1\}$ -valued definable equivalence relation E(x,y) on  $p(\mathbb{O})$  with a large number of E-classes. Any  $a,b \models p$  with  $\neg E(a,b)$  must be in distinct connected components of  $p(\mathbb{O})$ , so T has  $(\star)$ .  $\square$ 

One thing to note is that the first part of Proposition 3.1 doesn't use anything special about the relation 'a and b are in the same connected component of  $p(\mathbb{O})$ ' other than the fact that it is an invariant equivalence relation on  $p(\mathbb{O})$ . By comparison to Theorem 1.5, a reasonable test question for the robustness of  $(\star)$  would be this:

#### **Question 3.2.** Is the failure of $(\star)$ always witnessed by 1-types?

And given the results of Section 2 and the fact that  $(\star)$  separates Hilbert spaces, atomless probability algebras, and the Urysohn sphere<sup>10</sup> from the  $\mathbb{R}$ -arboreal theories, the next question is natural.

**Question 3.3.** Is there a strictly simple theory with  $(\star)$ ?

Characterizing  $(\star)$  more precisely might require an analysis of the topological properties of sets of types of the form  $\{q(\bar{x},\bar{y})\in S_{\bar{x}\bar{y}}(M):q(\bar{x},\bar{y})\supseteq p(\bar{x})\cup S_{\bar{x}\bar{y}}(M)\}$ 

 $<sup>^{10}</sup>L^p$ -Banach lattices probably have  $(\star)$  too, but we have not checked this.

 $p(\bar{y})$ ,  $(\forall ab \models q)a$  and b are in the same connected component of  $p(\mathbb{O})$ , which seems difficult. By [8], the complexity of the class of connected closed subset of a Polish space can be as high as  $\Pi_2^1$ , but even worse it was established in [1] that connectivity between points in a complete metric space is not always witnessed by separable subspaces. It's not even clear that  $(\star)$  is set-theoretically absolute. Perhaps this will be the kind of ugliness that is ironed out by considering the class of all models of a theory, but perhaps not.

**Question 3.4.** Is  $(\star)$  set-theoretically absolute? What is the topological complexity of the set of complete theories satisfying  $(\star)$ ?

One can write down a number of reasonable conditions intermediate between  $(\star)$  and having no non-algebraic crisp imaginary types—for instance, we could consider path-connected components or quasi-components instead of connected components or we could think about  $\varepsilon$ -chainability as a uniform proxy for connectivity—but given that there are no strongly motivating examples separating these intermediate notions, it seems premature to commit them to print.

#### References

- [1] T. Banakh, M. Vovk, and M. R. Wójcik. Connected, not separably connected complete metric spaces, 2009.
- [2] Itaï Ben Yaacov. Simplicity in compact abstract theories. Journal of Mathematical Logic, 03(02):163-191, November 2003.
- [3] Itaï Ben Yaacov. Uncountable dense categoricity in cats. Journal of Symbolic Logic, 70(3):829–860, September 2005.
- [4] Itaï Ben Yaacov. On theories of random variables. Israel Journal of Mathematics, 194(2):957– 1012, March 2013.
- [5] Itaï Ben Yaacov and Alexander Berenstein. Imaginaries in Hilbert spaces. Archive for Mathematical Logic, 43(4):459–466, May 2004.
- [6] Sylvia Carlisle and C. Ward Henson. Model theory of R-trees. Journal of Logic and Analysis, 12, December 2020.
- [7] Gabriel Conant and James Hanson. Separation for isometric group actions and hyperimaginary independence. Fundamenta Mathematicae, 259(1):97-109, 2022.
- [8] Gabriel Debs and Jean Saint Raymond. The descriptive complexity of connectedness in Polish spaces. Fundamenta Mathematicae, 249(3):261–286, 2020.
- [9] Clifton Ealy and Isaac Goldbring. Thorn-forking in continuous logic. The Journal of Symbolic Logic, 77(1):63-93, March 2012.
- [10] Steven Neil Evans. Probability and Real Trees. Springer Berlin Heidelberg, 2008.
- [11] Isaac Goldbring and H. Jerome Keisler. Continuous sentences preserved under reduced products. *Journal of Symbolic Logic*, 87(2):649–681, September 2020.
- [12] James Hanson. Strongly Minimal Sets and Categoricity in Continuous Logic.  $Memoirs\ of\ the\ AMS.$  Accepted.
- [13] James Hanson. Approximate Isomorphism of Metric Structures. arXiv e-prints, page arXiv:2011.00588, November 2020. Submitted.
- [14] James Hanson. Definability and categoricity in continuous logic. PhD thesis, University of Wisconsin–Madison, 2020.
- [15] James Hanson. Topometric characterization of type spaces in continuous logic. arXiv e-prints, page arXiv:2106.13261. June 2021. Submitted.
- [16] Wilfrid Hodges. Model Theory. Cambridge University Press, March 1993.
- [17] J. R. Isbell. Six theorems about injective metric spaces. Commentarii Mathematici Helvetici, 39(1):65–76, December 1964.

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF MARYLAND, COLLEGE PARK, MD 20742, USA *Email address*: jhanson9@umd.edu