© 2011 by Shichang Song. All rights reserved.

# MODEL THEORY AND PROBABILITY

BY

SHICHANG SONG

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Mathematics
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2011

Urbana, Illinois

Doctoral Committee:

Professor Sławomir Solecki, Chair
Professor Emeritus C. Ward Henson, Director of Research
Professor Lou van den Dries
Professor Richard Sowers

# Abstract

This thesis presents a systematic study of the model theory of probability algebras, random variable structures, and adapted structures, with an emphasis on their atomless counterparts.

In this thesis, the author uses a continuous version of first order logic that has been developed recently and that is better suited for applications to metric structures than classical first order logic. The set of truth values in continuous logic is the interval $[0, 1]$ instead of the truth values {True, False} in classical logic.

The author studies axioms, type spaces, quantifier elimination, separable categoricity, saturated models, stability, and $d$-finiteness for the theories of atomless probability algebras and atomless random variable structures. Explicit formulas for the $d^*$-metric between types in the theory of atomless random variable structures are given. For the theory of atomless adapted structures, the author studies axioms, type spaces, quantifier elimination, separably categoricity, and $d$-finiteness.

*To my family.*

# Acknowledgements

This thesis would not have been possible without the support of many people. Many thanks to my thesis adviser, Professor Henson, who introduced this area to me and helped me to make many corrections not only in mathematics but also in English. The author is also grateful to professors and students in the department of mathematics, especially in the logic group for their help.

# Preface

This thesis is a systematic presentation of the model theory of probability algebras, random variable structures, and adapted structures, with an emphasis on their atomless counterparts.

The treatment of probability spaces and stochastic processes using ideas from model theory was started by Keisler and his students. They developed an alternative framework to treat these structures, because classical first order model theory does not apply adequately to these structures. They defined notions of universality, homogeneity, and saturation for probability spaces, and proved many results about these concepts. They also defined adapted distributions for stochastic processes. The book [13] is a survey of this topic.

In this thesis, the author uses a continuous version of first order logic that has been developed recently (see [4] and [6]) and that is better suited for applications to metric structures than classical first order logic. The set of truth values in continuous logic is the interval $[0, 1]$ instead of the truth values {True, False} in classical logic. Continuous logic has improved the effectiveness of first order model theory for structures from analysis; for instance, see [1], [2], [3] and [7].

Ben Yaacov [1] studied probability algebras in the setting of compact abstract theories. His work on probability algebras was carried into continuous logic by Berenstein and Henson as part of the background of their paper [7]. In [2], Ben Yaacov studied the theories of ($[0, 1]$-valued) random variable structures and their atomless counterparts. He axiomatized the theory of random variable structures by $\mathsf{RV}(\mathsf{BY})$ and showed that the theory of probability algebras and the theory $\mathsf{RV}(\mathsf{BY})$ are biinterpretable. For the theory of atomless random variable structures, he studied separable categoricity, type spaces, quantifier elimination, stability, *etc* in [2].

In an early draft of [2] in 2006 with the title *On the theory of random variables and adapted spaces*, Ben Yaacov initiated a study of adapted structures, which are random variable structures equipped with a $T$-indexed family of conditional expectations, where $T$ is a linear ordering. Ben Yaacov axiomatized the theory of adapted structures with one conditional expectation. He also presented axioms for the theories of (totally atomless) adapted structures with a $T$-indexed family of conditional expectations. He announced basic

model theoretic properties of the theory of totally atomless adapted structures with a $T$-indexed family of conditional expectations; *e.g.*, it admits quantifier elimination and it is the model-completion of the theory of adapted structures with a $T$-indexed family of conditional expectations. In the case when $T$ is finite, he stated that the theory of totally atomless adapted structures with a $T$-indexed family of conditional expectations is separably categorical; hence this theory is complete for arbitrary $T$. Moreover, he pointed out connections with work by Fajardo and Keisler [13]. He gave a detailed discussion of presentations of canonical bases in atomless random variable structures. Further, he stated that the theory of totally atomless adapted structures is the theory of beautiful tuples of random variable structures, and that model theoretic independence for that theory can be recovered from the general theory of beautiful tuples. The content of Ben Yaacov's paper evolved rapidly after 2006 and later versions of [2] do not contain any discussion of adapted structures.

In this thesis, the author studies the theory of adapted structures mainly using measure theoretic tools, rather than the advanced model theoretic tools sketched by Ben Yaacov in his early versions of [2].

In the theory of atomless probability algebras, axiomatized by $\mathsf{APr}$, we study notions of universality, homogeneity and saturation in the setting of continuous logic, and connect them with analogous notions studied by Keisler, Hoover, and Fajardo. Our main new result here is Theorem 3.4.2, which characterizes the $\kappa$-saturated models of $\mathsf{APr}$ for all infinite cardinals $\kappa$. Our main tool in proving this result is Maharam's structure theory for measure algebras.

We axiomatize the theory of random variable structures by $\mathsf{RV}$, which is different from but equivalent to Ben Yaacov's theory $\mathsf{RV(BY)}$. Our approach to axiomatizing the theory of random variable structures only needs elementary measure theoretic probability theory, while Ben Yaacov's axiomatization is based on the completeness theorem for Łukasiewicz's $[0,1]$-valued propositional logic, which is somewhat difficult and not widely known. We study the metric topology on types of the theory of atomless random variable structures (axiomatized by $\mathsf{ARV}$). An explicit formula for the $d^*$-metric between types is given; see Theorem 5.3.9. We also make some connections to Wasserstein distances in optimal transport. Theorem 5.5.3 describes the $\kappa$-saturated models of $\mathsf{ARV}$ for all infinite cardinals $\kappa$ and Theorem 5.7.3 characterizes the $d$-finite tuples in models of $\mathsf{ARV}$. We connect our result on $d$-finiteness in $\mathsf{ARV}$ with work due to Keisler, Hoover, Fajardo, and Sun. Motivated by the correspondence between type spaces over $\emptyset$ and Wasserstein spaces (see Theorem 5.3.6), we define conditional Wasserstein spaces (see Definition 5.8.2). As an application of the Ryll-Nardzewski theorem in continuous logic, we determine when two natural topologies on these spaces are the same (see Theorem 5.8.7).

The theory of adapted structures provides the framework for a future study of stochastic processes. We

discuss it in two steps. The first step is to consider the theory of adapted structures with one conditional expectation, axiomatized by $\mathsf{RVE}$, and the theory of *totally atomless* (for definition, see page 138) adapted structures with one conditional expectation, axiomatized by $\mathsf{ARVE}$. The second step is to consider the theory of adapted structures with a $T$-indexed family of conditional expectations, axiomatized by $\mathsf{RVE_T}$, and the theory of totally atomless adapted structures with a $T$-indexed family of conditional expectations, axiomatized by $\mathsf{ARVE_T}$, where $T$ is a linear ordering. The first step builds basic tools in a simpler setting, which are applied in the second step for general $T$. In general, the theories $\mathsf{ARVE}$ and $\mathsf{ARVE_T}$ behave similarly when $T$ is finite, but are quite different when $T$ is infinite; *e.g.*, $\mathsf{ARVE_T}$ is separably categorical iff $T$ is finite. For the theories $\mathsf{ARVE}$ and $\mathsf{ARVE_T}$, we give proofs of basic model theoretic properties. Our new results are as follows. We investigate the type spaces of $\mathsf{ARVE_T}$ and make the connection between adapted distributions of stochastic processes and types in $\mathsf{ARVE_T}$, and we characterize the $d$-finite tuples in models of $\mathsf{ARVE_T}$ for finite $T$.

The thesis is organized as follows:

In Chapter 1, we outline the continuous logic background that is needed in this thesis. Most of this material is adapted from [4]. In Section 1.7, we present inductive theories, adapted from [29]. In Section 1.8, we prove Lindström's test for model completeness in continuous logic, which provides a sufficient condition for model completeness. In Chapters 6 and 7, we use Lindström's test to prove the theories $\mathsf{ARVE}$ and $\mathsf{ARVE_T}$ are model complete.

In Chapter 2, we present the background from analysis that is necessary for this thesis. There are four different topics. Section 2.1 provides the basics of measure theoretic probability theory. Section 2.2 presents results about measure algebras from [14]. Section 2.3 discusses the symmetric rearrangement of a Borel set and the symmetric-decreasing rearrangement of a Borel measurable function. Section 2.4 gives background on Wasserstein distances in optimal transport. These are related to the type spaces of $\mathsf{ARV}$ and are used in the proofs of Theorems 5.3.9 and 5.3.14, which give explicit formulas for the $d$-metric and the $d^*$-metric between types in $\mathsf{ARV}$.

In Chapter 3, we investigate the theory of probability algebras, axiomatized by $\mathsf{Pr}$, and the theory of atomless probability algebras, axiomatized by $\mathsf{APr}$. Section 3.1 presents material from [1] and [7] about the axioms for these theories and their basic model theoretic properties. Section 3.2 presents the relationships among several notions of universality. Section 3.3 discusses the relationships among several notions of homogeneity. Section 3.4 characterizes the $\kappa$-saturated models of $\mathsf{APr}$ for all infinite cardinals $\kappa$ and connects continuous model theoretic saturation with Hoover-Keisler saturation. Section 3.5 gives the definition of $d$-finiteness and presents a result from [5] that all finite tuples in $\mathsf{APr}$ are $d$-finite.

In Chapter 4, we provide axioms for the theory of random variable structures. In Section 4.1, we axiomatize the theory of random variable structures by $\mathsf{RV}$ and by $\mathsf{ARV}$ for its atomless counterpart. In Section 4.2, we present Ben Yaacov's axiomatizations given in [2], which are equivalent to ours. In Section 4.3, we consider the class of structures obtained by adding the multiplication function to random variable structures. We axiomatize this class by $\mathsf{RV}(\times)$ and it follows that $\mathsf{RV}(\times)$ is an extension by definitions of $\mathsf{RV}$.

In Chapter 5, we investigate the theory of random variable structures and the theory of atomless random variable structures. This chapter is the most important one of this thesis. In Section 5.1, we show how to generate a natural random variable structure from a probability algebra. In Section 5.2, basic model theoretic properties of $\mathsf{ARV}$ are presented with detailed proofs. In Section 5.3, we study the type spaces of $\mathsf{ARV}$. The main results are Theorem 5.3.9 and Theorem 5.3.14, which give explicit formulas for the $d^*$-metric between types in $\mathsf{ARV}$. In Section 5.4, using some of the model theoretic results in the preceding section, we give a new proof of Theorem 2.4.4, which gives a formula for the Wasserstein distance between probability measures on the real line. In Section 5.5, we describe the $\kappa$-saturated models of $\mathsf{ARV}$ for all infinite cardinals $\kappa$. In Section 5.6, we give a proof of a fact from Ben Yaacov [2] that $\mathsf{ARV}$ is $\omega$-stable and that its model theoretic independence relation coincides with probabilistic independence. In Section 5.7, we characterize the $d$-finite tuples in models of $\mathsf{ARV}$. In Section 5.8, we define conditional Wasserstein spaces and prove some results about these spaces.

In Chapters 6 and 7, we investigate the model theory of adapted structures, which are random variable structures equipped with a $T$-indexed family of conditional expectations. In Chapter 6, we study the theories of (totally atomless) adapted structures with only one conditional expectation. In Section 6.1, we axiomatize the theory of adapted structures with one conditional expectation by $\mathsf{RVE}$ and the theory of *totally atomless* adapted structures with one conditional expectation by $\mathsf{ARVE}$. In Section 6.2, we prove several model theoretic properties of $\mathsf{ARVE}$. In Section 6.3, we characterize the $d$-finite tuples in models of $\mathsf{ARVE}$. In Chapter 7, we study the theories of (totally atomless) adapted structures with a $T$-indexed family of conditional expectations, where $T$ is a linear ordering, parallel to Chapter 6. In Section 7.1, we axiomatize the theory of adapted structures with a $T$-indexed family of conditional expectations by $\mathsf{RVE}_\mathsf{T}$, and the theory of totally atomless adapted structures with a $T$-indexed family of conditional expectations by $\mathsf{ARVE}_\mathsf{T}$. In Section 7.2, we prove several model theoretic properties of $\mathsf{ARVE}_\mathsf{T}$. In Section 7.3, we characterize the $d$-finite tuples in models of $\mathsf{ARVE}_\mathsf{T}$ when $T$ is finite.

# Table of Contents

**List of Abbreviations** . . . . . **xi**

**Chapter 1 Basics of continuous logic** . . . . . **1**
1.1 Metric structures and signatures . . . . . 1
1.2 Continuous logic: syntax and semantics . . . . . 3
1.3 Basic model theoretic concepts . . . . . 8
1.4 Constructions of models . . . . . 9
1.5 Spaces of types . . . . . 11
1.6 Definability in metric structures . . . . . 12
1.7 Inductive theories . . . . . 16
1.8 Model completeness . . . . . 17
1.9 Separable categoricity . . . . . 19
1.10 Quantifier elimination . . . . . 21
1.11 Stability and independence . . . . . 22

**Chapter 2 Background from analysis** . . . . . **25**
2.1 Basics of probability theory . . . . . 25
2.2 Background for measure algebras . . . . . 27
2.3 Rearrangements . . . . . 41
2.4 Wasserstein distances . . . . . 43

**Chapter 3 Model theory of probability algebras** . . . . . **45**
3.1 Basics of Pr and APr . . . . . 45
3.2 Universality of probability spaces . . . . . 48
3.3 Homogeneous models of APr . . . . . 49
3.4 Saturated models of APr . . . . . 51
3.5 On $d$-finite tuples in APr . . . . . 54

**Chapter 4 Axioms for the theory of random variable structures** . . . . . **55**
4.1 Axioms for RV and ARV . . . . . 55
4.2 Ben Yaacov's axiomatizations . . . . . 79
4.3 Extension of the theory by multiplication . . . . . 81

**Chapter 5 Model theory of random variable structures** . . . . . **90**
5.1 Probability algebras and random variable structures . . . . . 90
5.2 Basic model theoretic properties of RV and ARV . . . . . 92
5.3 On type spaces of ARV . . . . . 96
5.4 A new proof of an integral formula for the Wasserstein distance . . . . . 107
5.5 Saturated models of ARV . . . . . 110
5.6 Stability and independence . . . . . 113
5.7 On $d$-finite tuples in ARV . . . . . 115
5.8 Conditional Wasserstein distances . . . . . 120

**Chapter 6 Model theory of adapted structures with one conditional expectation . . . . 124**
6.1 Axioms for RVE and ARVE . . . . . 125
6.2 Basics of ARVE . . . . . 127
6.3 On $d$-finite tuples in ARVE . . . . . 134

**Chapter 7 Model theory of adapted structures with $T$-indexed conditional expectations 138**
7.1 Axioms for $\mathsf{RVE}_\mathsf{T}$ and $\mathsf{ARVE}_\mathsf{T}$ . . . . . 138
7.2 Basics of $\mathsf{ARVE}_\mathsf{T}$ . . . . . 140
7.3 On $d$-finite tuples in $\mathsf{ARVE}_\mathsf{T}$ . . . . . 151

**References . . . . . 155**

**Vita . . . . . 157**

# List of Abbreviations

| | |
|---|---|
| $\mathbb{N}$ | The set of natural numbers $\{1, 2, \cdots\}$. |
| $\mathbb{D}$ | The set of dyadic numbers in $[0, 1]$. |
| $d(a, b)$ | $\max_{1 \leq i \leq n} d(a_i, b_i)$, where $a = (a_1, \cdots, a_n)$ and $b = (b_1, \cdots, b_n)$ |
| $d^*(a, b)$ | $\sum_{i=1}^{n} d(a_i, b_i)$, where $a = (a_1, \cdots, a_n)$ and $b = (b_1, \cdots, b_n)$ |
| $d(p, q)$ | The $d$-metric on $n$-types, induced by $d$ on $n$-tuples. |
| $d^*(p, q)$ | The $d^*$-metric on $n$-types, induced by $d^*$ on $n$-tuples. |
| e.c. | Existentially closed. |
| $(A_\kappa, \mu_\kappa)$ | The measured algebra associated to the product measure space $([0, 1]^\kappa, \mu_\kappa)$. |
| a.s. | Almost surely. |
| $\mathcal{L}^1\big((\Omega, \mathcal{F}, \mu), [0, 1]\big)$ | The space of measurable functions from the probability space $(\Omega, \mathcal{F}, \mu)$ to $[0, 1]$ with the $L^1$-pseudometric. |
| $L^1\big((\Omega, \mathcal{F}, \mu), [0, 1]\big)$ | The $L^1$-space of equivalence classes of the random variables in $\mathcal{L}^1\big((\Omega, \mathcal{F}, \mu), [0, 1]\big)$ with the $L^1$-metric. |
| $L^1\big((\Omega, \mathcal{F}, \mu), \{0, 1\}\big)$ | The set of equivalence classes of characteristic functions in $\mathcal{L}^1\big((\Omega, \mathcal{F}, \mu), [0, 1]\big)$. |
| $L^1\big((\Omega, \mathcal{F}, \mu), \mathbb{D}\big)$ | The set of equivalence classes of $\mathbb{D}$-valued simple functions in $\mathcal{L}^1\big((\Omega, \mathcal{F}, \mu), [0, 1]\big)$. |
| $L^1(\mathcal{F}, [0, 1])$ | $L^1\big((\Omega, \mathcal{F}, \mu), [0, 1]\big)$, where $\mathcal{F}$ is the $\sigma$-algebra of measurable sets of the probability space $(\Omega, \mathcal{F}, \mu)$. |
| $L^1(\mu, [0, 1])$ | $L^1\big((\Omega, \mathcal{F}, \mu), [0, 1]\big)$, where $\mu$ is the probability measure of the probability space $(\Omega, \mathcal{F}, \mu)$. |
| $L^1(\mathscr{A}, [0, 1])$ | $L^1\big((\Omega, \mathcal{F}, \mu), [0, 1]\big)$, where $\mathscr{A}$ is the probability algebra associated to the probability space $(\Omega, \mathcal{F}, \mu)$. |
| $\chi_A$ | The characteristic function of the set $A$. |
| $\mathbf{1}_A$ | $[\chi_A]_{a.s.}$, where $A$ is a measurable set of a probability space. |
| $\mathrm{dist}(f)$ | The distribution induced by the random variable $f$. |
| $\mathrm{dist}(f \mid \mathscr{A})$ | The conditional distribution over $\mathscr{A}$ induced by the random variable $f$. |
| $\mathrm{supp}(f)$ | The support of a function $f$, which is the set of points where the function $f$ is not zero. |

| | |
|---|---|
| $f^*$ | The decreasing rearrangement of a Borel function $f\colon [0,1] \to [0,1]$. |
| $\sigma(A)$ | The $\sigma$-algebra (resp. $\sigma$-complete boolean algebra) $\sigma$-generated by $A$, where $A$ is a subset of a $\sigma$-algebra (resp. $\sigma$-complete boolean algebra). |
| $\mathscr{A} \vee \mathscr{B}$ | The $\sigma$-algebra (resp. $\sigma$-complete boolean algebra) $\sigma$-generated by the $\sigma$-algebras (resp. $\sigma$-complete boolean algebras) $\mathscr{A}$ and $\mathscr{B}$. |
| $\sigma(X)$ | The $\sigma$-algebra of measurable sets generated by the random variables in $X$, where $X$ is a set of random variables. |
| $W_p$ | The Wasserstein distance of order $p$ between probability measures. |
| $P_p(\mathcal{X})$ | The Wasserstein space of order $p$, which is a subspace of the space of probability measures on a Polish metric space $\mathcal{X}$; it is equipped with the Wasserstein distance $W_p$. |
| $L_{\mathsf{Pr}}$ | The signature for probability algebras. |
| $\mathsf{Pr}$ | Axioms for the $L_{\mathsf{Pr}}$-theory of probability algebras. (p. 45) |
| $\mathsf{APr}$ | Axioms for the $L_{\mathsf{Pr}}$-theory of *atomless* probability algebras. (p. 46) |
| $L_{\mathsf{RV}}$ | The signature for random variable structures. |
| $\mathsf{RV}$ | Axioms for the $L_{\mathsf{RV}}$-theory of random variable structures. (p. 58) |
| $\mathsf{ARV}$ | Axioms for the $L_{\mathsf{RV}}$-theory of *atomless* random variable structures. (p. 59) |
| $L_{\mathsf{RV}(\times)}$ | The signature for random variable structures with multiplication in the signature. |
| $\mathsf{RV}(\times)$ | Axioms for the $L_{\mathsf{RV}(\times)}$-theory of random variable structures with multiplication in the signature. (p. 81) |
| $\mathsf{ARV}(\times)$ | Axioms for the $L_{\mathsf{RV}(\times)}$-theory of *atomless* random variable structures with multiplication in the signature. (p. 89) |
| $L_{\mathsf{RVE}}$ | The signature for adapted structures with one conditional expectation. |
| $\mathsf{RVE}$ | Axioms for the $L_{\mathsf{RVE}}$-theory of adapted structures with one conditional expectation. (p. 125) |
| $\mathsf{ARVE}$ | Axioms for the $L_{\mathsf{RVE}}$-theory of *totally atomless* adapted structures with one conditional expectation. (p. 126) |
| $L_{\mathsf{RVE}_\mathsf{T}}$ | The signature for adapted structures with a $T$-indexed family of conditional expectations. |
| $\mathsf{RVE}_\mathsf{T}$ | Axioms for the $L_{\mathsf{RVE}_\mathsf{T}}$-theory of adapted structures with a $T$-indexed family of conditional expectations. (p. 139) |
| $\mathsf{ARVE}_\mathsf{T}$ | Axioms for the $L_{\mathsf{RVE}_\mathsf{T}}$-theory of *totally atomless* adapted structures with a $T$-indexed family of conditional expectations. (p. 139) |

# Chapter 1

# Basics of continuous logic

In this chapter, we follow the articles [4] and [6] to give the background for basics of continuous logic, which provides the framework in Chapters 3, 4, 5, 6, and 7. Most of the notations, definitions and results are from [4].

## 1.1 Metric structures and signatures

In this section, we present the definitions of metric structures and signatures. This material is taken from [4].

### Metric structures

Let $(M, d)$ be a complete metric space. We say it is *bounded* if there is a $B > 0$ such that for all $x, y \in M$, we have $d(x, y) \leq B$. The smallest such number $B$ is called the *diameter* of $(M, d)$. Suppose $(M_i, d_i)$ are metric spaces for every $i = 1, \cdots, n$. Unless otherwise specified, the metric on the product space $M_1 \times \cdots \times M_n$ is taken to be the maximum metric

$$d\big((x_1, \cdots, x_n), (y_1, \cdots, y_n)\big) := \max_{1 \leq i \leq n} d_i(x_i, y_i).$$

Another natural metric is defined as

$$d^*\big((x_1, \cdots, x_n), (y_1, \cdots, y_n)\big) := \sum_{i=1}^{n} d_i(x_i, y_i).$$

This $d^*$ will be used mainly in Section 5.3. Note that these two metrics are equivalent to each other.

Let $(M, d_M)$ and $(N, d_N)$ be metric spaces, and suppose $f\colon M \to N$ is a function. We say $\Delta\colon (0, 1] \to (0, 1]$ is a *modulus of uniform continuity for* $f$ if for each $\epsilon \in (0, 1]$ and for all $x, y \in M$ we have

$$d_M(x, y) < \Delta(\epsilon) \Longrightarrow d_N\big(f(x), f(y)\big) \leq \epsilon.$$

We say $f$ is *uniformly continuous* if it has a modulus of uniform continuity.

Let $(M, d)$ be a bounded complete metric space. A *predicate* on $M$ is a uniformly continuous function from $M^n$ to a closed bounded interval in $\mathbb{R}$ (for some $n \in \mathbb{N}$) and a *function* on $M$ is a uniformly continuous function from $M^n$ into $M$ (for some $n \in \mathbb{N}$). When a predicate or function is defined on $M^n$, we call it *$n$-ary* and we say $n$ is its *arity*.

A *metric structure* $\mathcal{M}$ based on a bounded metric space $(M, d)$ consists of:

- a family $(R_i \mid i \in I)$ of predicates on $M$;
- a family $(F_j \mid j \in J)$ of functions on $M$;
- a family $(a_k \mid k \in K)$ of distinguished elements of $M$.

where each of the sets $I, J, K$ is possibly empty.

We often denote the metric structure $\mathcal{M}$ as

$$\mathcal{M} = \big(M, R_i, F_j, a_k \mid i \in I, j \in J, k \in K\big).$$

## Signatures

For a metric structure $\mathcal{M} = \big(M, R_i, F_j, a_k \mid i \in I, j \in J, k \in K\big)$, the *signature* $L$ associated to $\mathcal{M}$ consists of:

- a family of predicate symbols $(P_i \mid i \in I)$ such that to each $P_i$ we associate an integer $a(P_i)$ which is the arity of $R_i$, a modulus of uniform continuity $\Delta_{P_i}$ which is the same for $R_i$, and a closed bounded interval $I_{P_i}$ which is the same for $R_i$. We denote $R_i$ by $P_i^{\mathcal{M}}$ for each $i \in I$.
- a family of function symbols $(f_j \mid j \in J)$ such that to each $f_j$ we associate an integer $a(f_j)$ which is the arity of $F_j$ and a modulus of uniform continuity $\Delta_{f_j}$ which is the same for $F_j$. We denote $F_j$ by $f_j^{\mathcal{M}}$ for each $j \in J$.
- a family of constant symbols $(c_k \mid k \in K)$. We denote $a_k$ by $c_k^{\mathcal{M}}$ for each $k \in K$.
- the metric symbol $d$ which is given by the metric on $M$. We use the same notation here for simplicity.
- a non-negative real number $D_L$ which is a bound on the diameter of $(M, d)$.

Sometimes we let $d^{\mathcal{M}}$ denote the metric $d$ given by $\mathcal{M}$. Unless otherwise specified, $D_L = 1$ and $I_P = [0, 1]$ for every predicate symbol $P$ in the signature.

## 1.2 Continuous logic: syntax and semantics

In this section, we introduce the syntax and semantics of continuous logic. Most of the presentation is taken from [4]. We fix a signature $L$ for metric structures as described in last section.

### Syntax

The *symbols* of $L$ consist of the *nonlogical* symbols of $L$, which are the predicate, function, and constant symbols, and also the so-called *logical* symbols of $L$. The logical symbols include the symbol $d$ for the metric on the underlying metric space, an infinite set of variables which usually is countable, the connective symbols, which are continuous functions from $[0,1]^n$ to $[0,1]$ for all $n \in \mathbb{N}$, and the quantifier symbols which are sup and inf.

The *cardinality* of $L$, denoted by $\mathrm{Card}(L)$, is the least infinite cardinal $\geq$ the cardinality of the set of nonlogical symbols of $L$.

The *$L$-terms* are formed inductively as in first order logic. That is, every variable and every constant symbol is an $L$-term; if $f$ is an $n$-ary function symbol and $t_1, \cdots, t_n$ are $L$-terms, then $f(t_1, \cdots, t_n)$ is also an $L$-term.

For an $n$-ary predicate $P$ (including $d$ when $n = 2$) and $L$-terms $t_1, \cdots, t_n$, the expression of the form $P(t_1, \cdots, t_n)$ is an *atomic formula of $L$*. (The logical symbol $d$ for the metric is treated as a binary predicate symbol just as the equality symbol $=$ is treated in first order logic.)

Then using connective symbols and quantifier symbols, we construct *$L$-formulas* (or *formulas of $L$*) inductively as in first order logic. More precisely:

**1.2.1 Definition.** *The class of $L$-formulas is the smallest class satisfying the following:*

*(i) Atomic formulas of $L$ are $L$-formulas.*

*(ii) If $u\colon [0,1]^n \to [0,1]$ is continuous and $\varphi_1, \cdots, \varphi_n$ are $L$-formulas, then $u(\varphi_1, \cdots, \varphi_n)$ is also an $L$-formula.*

*(iii) If $\varphi$ is an $L$-formula and $x$ is a variable, then $\sup_x \varphi$ and $\inf_x \varphi$ are also $L$-formulas.*

A *quantifier free* $L$-formula is one constructed without using quantifiers sup or inf. An $L$-formula is called an *$L$-sentence* if it has no free variables. For a term $t$ or a formula $\varphi$, we write $t(x_1, \cdots, x_n)$ or $\varphi(x_1, \cdots, x_n)$ to indicate that its free variables are among $x_1, \cdots, x_n$.

## Prestructures and structures

Let $(M_0, d_0)$ be a pseudometric space whose diameter is $\leq D_L$. An *$L$-prestructure* $\mathcal{M}_0$ based on $(M_0, d_0)$ is a structure consisting of the following:

(i) for every predicate symbol $P$ of $L$, a function $P^{\mathcal{M}_0}$ from $M_0^n$ to $I_P$, which has $\Delta_P$ as a modulus of uniform continuity, where $n$ is the arity of $P$.

(ii) for every function symbol $f$ of $L$, a function $f^{\mathcal{M}_0}$ from $M_0^n$ to $M_0$, which has $\Delta_f$ as a modulus of uniform continuity, where $n$ is the arity of $f$.

(iii) for every constant symbol $c$ of $L$, an element $c^{\mathcal{M}_0}$ of $M_0$.

An $L$-prestructure $\mathcal{M}$ is an *$L$-structure*, if its underlying space is a complete metric space.

Starting from an $L$-prestructure $\mathcal{M}_0$, there is a procedure to construct an $L$-structure associated to it. First, we define its quotient prestructure $\mathcal{M}$. Let $(M, d)$ be the quotient metric space of $(M_0, d_0)$ and let $\pi\colon M_0 \to M$ be the quotient map. Then we define the $L$-prestructure $\mathcal{M}$ as follows:

(i) for every predicate symbol $P$ of $L$, define $P^{\mathcal{M}}\colon M^n \to I_P$ by $P^{\mathcal{M}}(\pi(x_1), \cdots, \pi(x_n)) = P^{\mathcal{M}_0}(x_1, \cdots, x_n)$ for all $x_1, \cdots, x_n \in M_0$, where $n$ is the arity of $P$.

(ii) for every function symbol $f$ of $L$, define $f^{\mathcal{M}}\colon M^n \to M$ by $f^{\mathcal{M}}(\pi(x_1), \cdots, \pi(x_n)) = \pi(f^{\mathcal{M}_0}(x_1, \cdots, x_n))$ for all $x_1, \cdots, x_n \in M_0$, where $n$ is the arity of $f$.

(iii) for every constant symbol $c$ of $L$, define $c^{\mathcal{M}} := \pi(c^{\mathcal{M}_0})$.

This makes $\mathcal{M}$ an $L$-prestructure.

Then we define an $L$-structure $\overline{\mathcal{M}}$ by taking a completion of $\mathcal{M}$. Let $(\overline{M}, d)$ be the completion of the metric space $(M, d)$. We define $\overline{\mathcal{M}}$ as follow:

(i) for every predicate symbol $P$ of $L$, define $P^{\overline{\mathcal{M}}}\colon \overline{M}^n \to I_P$ to be the unique continuous extension of $P^{\mathcal{M}}$, where $n$ is the arity of $P$.

(ii) for every function symbol $f$ of $L$, define $f^{\overline{\mathcal{M}}}\colon \overline{M}^n \to \overline{M}$ to be the unique continuous extension of $f^{\mathcal{M}}$, where $n$ is the arity of $f$.

(iii) for every constant symbol $c$ of $L$, define $c^{\overline{\mathcal{M}}} := c^{\mathcal{M}}$.

This makes $\overline{\mathcal{M}}$ an $L$-structure.

## Semantics

Let $\mathcal{M}$ be an $L$-prestructure and let $A \subseteq M$. By adding a new constant symbol $c(a)$ to $L$ for every $a \in A$, we extend $L$ to $L(A)$. Then by defining $c(a)^{\mathcal{M}} := a$ for every $a \in A$, we extend $\mathcal{M}$ to an $L(A)$-prestructure canonically, denoted by $(\mathcal{M}, a)_{a\in A}$.

Let $t(x_1, \cdots, x_n)$ be an $L(M)$-term. Then as in first order logic, the interpretation of $t$ in $\mathcal{M}$ is a function $t^{\mathcal{M}} \colon M^n \to M$. Next, for an $L(M)$-sentence $\sigma$ we define the *value of* $\sigma$ *in* $\mathcal{M}$ by induction. The value of $\sigma$ is a real number in $[0, 1]$ and is denoted by $\sigma^{\mathcal{M}}$.

**1.2.2 Definition.** *(i)* $\big(d(t_1, t_2)\big)^{\mathcal{M}} = d^{\mathcal{M}}(t_1^{\mathcal{M}}, t_2^{\mathcal{M}})$ *for all terms* $t_1$ *and* $t_2$*;*

*(ii)* $\big(P(t_1, \cdots, t_n)\big)^{\mathcal{M}} = P^{\mathcal{M}}(t_1^{\mathcal{M}}, \cdots, t_n^{\mathcal{M}})$ *for every* $n$*-ary predicate symbol* $P$ *and terms* $t_1, \cdots, t_n$*;*

*(iii) for* $L(M)$*-sentences* $\sigma_1, \cdots, \sigma_n$ *and every continuous function* $u \colon [0, 1]^n \to [0, 1]$*,*

$$\big(u(\sigma_1, \cdots, \sigma_n)\big)^{\mathcal{M}} = u(\sigma_1^{\mathcal{M}}, \cdots, \sigma_n^{\mathcal{M}});$$

*(iv) for every* $L(M)$*-formula* $\varphi(x)$*,*

$$\big(\sup_x \varphi(x)\big)^{\mathcal{M}}$$

*is the supremum in* $[0, 1]$ *of the set* $\{\varphi(a)^{\mathcal{M}} \mid a \in M\}$*;*

*(v) for every* $L(M)$*-formula* $\varphi(x)$*,*

$$\big(\inf_x \varphi(x)\big)^{\mathcal{M}}$$

*is the infimum in* $[0, 1]$ *of the set* $\{\varphi(a)^{\mathcal{M}} \mid a \in M\}$*.*

Given an $L(M)$-formula $\varphi(x_1, \cdots, x_m)$ we let $\varphi^{\mathcal{M}}$ denote the function from $M^n$ to $[0, 1]$ defined by

$$\varphi^{\mathcal{M}}(a_1, \cdots, a_n) = \big(\varphi(a_1, \cdots, a_n)\big)^{\mathcal{M}},$$

for every $(a_1, \cdots, a_n) \in M^n$.

For every $L$-term $t(x)$ and every $L$-formula $\varphi(x)$, the functions $t^{\mathcal{M}}$ and $\varphi^{\mathcal{M}}$ are uniformly continuous functions from $M^n$ to $[0, 1]$ and the their moduli of continuity do not depend on the structure $\mathcal{M}$ but only on the data given by the signature $L$; see [4, Theorem 3.5].

**1.2.3 Theorem** ([4, Theorem 3.7])**.** *Let* $\mathcal{M}_0$ *be an* $L$*-prestructure with underlying pseudometric space* $(M_0, d_0)$*; let* $\mathcal{M}$ *be its quotient* $L$*-prestructure with quotient map* $\pi \colon M_0 \to M$ *and let* $\mathcal{N}$ *be the* $L$*-structure*

*coming from completing $\mathcal{M}$ (as explained on page 4). Let $t(x)$ be an $L$-term and $\varphi(x)$ be an $L$-formula. Then:*

*(i)* $t^{\mathcal{M}}(\pi(a_1), \cdots, \pi(a_n)) = t^{\mathcal{M}_0}(a_1, \cdots, a_n)$ *for all* $a_1, \cdots, a_n \in M_0$*;*

*(ii)* $t^{\mathcal{N}}(b) = t^{\mathcal{M}}(b)$ *for all* $b \in M^n$*;*

*(iii)* $\varphi^{\mathcal{M}}(\pi(a_1), \cdots, \pi(a_n)) = \varphi^{\mathcal{M}_0}(a_1, \cdots, a_n)$ *for all* $a_1, \cdots, a_n \in M_0$*;*

*(iv)* $\varphi^{\mathcal{N}}(b) = \varphi^{\mathcal{M}}(b)$ *for all* $b \in M^n$*.* □

Two $L$-formulas $\varphi(x_1, \cdots, x_n)$ and $\psi(x_1, \cdots, x_n)$ are *logically equivalent* if

$$\varphi^{\mathcal{M}}(a_1, \cdots, a_n) = \psi^{\mathcal{M}}(a_1, \cdots, a_n)$$

for every $L$-structure $\mathcal{M}$ and all $a_1, \cdots, a_n \in M$. The *logical distance* between $\varphi$ and $\psi$ is

$$\sup\{|\varphi^{\mathcal{M}}(a) - \psi^{\mathcal{M}}(a)| \mid \mathcal{M} \text{ is an } L\text{-structure and } a \in M^n\}.$$

Since we consider all continuous functions from $[0,1]^n$ to $[0,1]$ as connectives, the set of all $L$-formulas is very large. Indeed, only the size of a set of $L$-formulas that is dense in the set of all $L$-formulas with respect to the logical distance really matters in continuous logic. The *density character* of a topological space is the smallest cardinality of a dense subset of the space. In [4], it was shown that the density character of the set of all $L$-formulas with respect to the logical distance is always $\leq \text{Card}(L)$.

An *$L$-condition* is a formal expression of the form $\varphi = 0$, where $\varphi$ is an $L$-formula. We call $\varphi = 0$ a *closed $L$-condition* if $\varphi$ is a sentence. If $x_1, \cdots, x_n$ are distinct variables, we write $E(x_1, \cdots, x_n)$ to indicate an $L$-condition $\varphi(x_1, \cdots, x_n) = 0$. Given an $L(M)$-condition $E(x_1, \cdots, x_n)$ of the form $\varphi(x_1, \cdots, x_n) = 0$ and $a_1, \cdots, a_n \in M$, we say $E$ is *true of* $a_1, \cdots, a_n$ *in* $\mathcal{M}$ if $\varphi^{\mathcal{M}}(a_1, \cdots, a_n) = 0$. We write $\mathcal{M} \models E[a_1, \cdots, a_n]$.

**1.2.4 Remark.** *Here we give some useful abbreviations for $L$-conditions. Let $\varphi$ and $\psi$ be $L$-formulas.*

*(i)* $\varphi = \psi$ *is an abbreviation for the $L$-condition $|\varphi - \psi| = 0$; here we are using the connective $u\colon [0,1]^2 \to [0,1]$ defined by $u(s,t) = |s - t|$.*

*(ii)* $\varphi = r$ *is an abbreviation for the $L$-condition $|\varphi - r| = 0$, where $r \in [0,1]$, and we are using the connective $u\colon [0,1] \to [0,1]$ defined by $u(t) = |t - r|$.*

*(iii)* $\varphi \leq \psi$ *and* $\psi \geq \varphi$ *are abbreviations for the $L$-condition $\varphi \dot{-} \psi = 0$, where $\dot{-}\colon [0,1]^2 \to [0,1]$ is the connective defined by $\dot{-}(s,t) = \max(s - t, 0)$.*

*(iv)* $\varphi \leq r$ *is an abbreviation for the* $L$*-condition* $\varphi \dot{-} r = 0$ *where* $r \in [0,1]$.

**1.2.5 Definition.** *Let* $\varphi$ *be an* $L$*-formula.*

1. *We say that* $\varphi$ *is an* $\forall$ $L$-formula *if for every* $\epsilon > 0$ *there is a formula* $\sigma$ *of the form* $\sup_x \psi$ *where* $\psi$ *is a quantifier free formula, such that for every* $L$*-structure* $\mathcal{M}$ *and* $a \in M^n$*, one has* $|\varphi^{\mathcal{M}}(a) - \sigma^{\mathcal{M}}(a)| \leq \epsilon$.

2. *We say that* $\varphi$ *is an* $\forall\exists$ $L$-formula *if for every* $\epsilon > 0$ *there is a formula* $\sigma$ *of the form* $\sup_x \inf_y \psi$ *where* $\psi$ *is a quantifier free formula, such that for every* $L$*-structure* $\mathcal{M}$ *and* $a \in M^n$*, one has* $|\varphi^{\mathcal{M}}(a) - \sigma^{\mathcal{M}}(a)| \leq \epsilon$.

**1.2.6 Definition.** *Let* $E$ *be an* $L(M)$*-condition.*

1. *We say that* $E$ *is an* $\forall$ $L$-condition *if* $E$ *is of the form* $\varphi(x) = 0$*, where* $\varphi(x)$ *is an* $\forall$ $L$*-formula.*

2. *We say that* $E$ *is an* $\forall\exists$ $L$-condition *if* $E$ *is of the form* $\varphi(x) = 0$*, where* $\varphi(x)$ *is an* $\forall\exists$ $L$*-formula.*

## Connectives

A *system of connectives* is a family $\mathcal{F} = (F_n \mid n \geq 1)$ where each $F_n$ is a set of connectives $u\colon [0,1]^n \to [0,1]$. We say that $\mathcal{F}$ is *closed* if it satisfies the following properties:

- For each $n$ and each $j = 1, \cdots, n$, the canonical projection $\pi_j^n\colon [0,1]^n \to [0,1]$ onto the $j^{th}$ coordinate is in $\mathcal{F}_n$.

- For each $m$ and $n$, if $u \in F_n$ and $v_1, \cdots, v_n \in F_m$, then $u(v_1, \cdots, v_n)$ is in $F_m$.

Let $\overline{\mathcal{F}}$ denote the smallest closed system of connectives generated by the system $\mathcal{F}$ of connectives. We say that $\mathcal{F}$ is *full* if for every $\epsilon > 0$ and every connective $u\colon [0,1]^n \to [0,1]$, there is a connective $v \in \overline{F}_n$ such that $|u(t) - v(t)| \leq \epsilon$ for every $t \in [0,1]^n$.

For a system $\mathcal{F}$ of connectives, the following collection of formulas, called $\mathcal{F}$*-restricted* formulas, is defined by induction:

- Atomic formulas are $\mathcal{F}$-restricted formulas.

- If $u \in F_n$ and $\varphi_1, \cdots, \varphi_n$ are $\mathcal{F}$-restricted formulas, then $u(\varphi_1, \cdots, \varphi_n)$ is also an $\mathcal{F}$-restricted formula.

- If $\varphi(x)$ is an $\mathcal{F}$-restricted formula, then $\sup_x \varphi(x)$ and $\inf_x \varphi(x)$ are also $\mathcal{F}$-restricted formulas.

If $\mathcal{F}$ is a full system of connectives, then the set of $\mathcal{F}$-restricted formulas is dense in the set of all formulas with respect to the logical distance between formulas; see [4, Theorem 6.3].

There is indeed a very simple full system of connectives. Let $\mathcal{F}_0 = (F_n \mid n \geq 1)$ where $F_1 = \{0, 1, \frac{t}{2}\}$ (we treat $0, 1$ as constant functions of one variable here), $F_2 = \{\dot{-}\}$, and all the other $F_n$'s are empty. Then [4, Proposition 6.3] shows that the system of connectives $\mathcal{F}_0$ is full.

## 1.3 Basic model theoretic concepts

In this section, we introduce several of the most basic model theoretic concepts and results including the compactness theorem, taken from [4]. We fix a signature $L$ for metric structures.

Let $\mathcal{M}$ and $\mathcal{N}$ be $L$-structures. An *embedding* $T$ from $\mathcal{M}$ into $\mathcal{N}$ is an isometric mapping of $M$ into $N$ such that $T$ commutes with the interpretations of the functions and predicates of $L$ in the following way:

- $f^{\mathcal{N}}\big(T(a_1), \cdots, T(a_n)\big) = T\big(f^{\mathcal{M}}(a_1, \cdots, a_n)\big)$ for every $n$-ary function symbol $f$ and all $a_1, \cdots, a_n \in M$;
- $P^{\mathcal{N}}\big(T(a_1), \cdots, T(a_n)\big) = P^{\mathcal{M}}(a_1, \cdots, a_n)$ for every $n$-ary predicate symbol $P$ and all $a_1, \cdots, a_n \in M$;
- $c^{\mathcal{N}} = T(c^{\mathcal{M}})$ for every constant symbol $c$.

A surjective embedding is called an *isomorphism*. If there is an isomorphism between $\mathcal{M}$ and $\mathcal{N}$, we say that $\mathcal{M}$ and $\mathcal{N}$ are *isomorphic*, and write $\mathcal{M} \cong \mathcal{N}$. An isomorphism between $\mathcal{M}$ and itself is called an *automorphism*.

If $M \subseteq N$ and the inclusion map from $M$ into $N$ is an embedding from $\mathcal{M}$ into $\mathcal{N}$, we say $\mathcal{M}$ is a *substructure* of $\mathcal{N}$, and write $\mathcal{M} \subseteq \mathcal{N}$.

An *$L$-theory* $T$ is a set of closed $L$-conditions. We call an $L$-structure $\mathcal{M}$ a *model* of $T$ and write $\mathcal{M} \models T$ if $\mathcal{M} \models E$ for every condition $E$ in $T$. For an $L$-structure $\mathcal{M}$, the *theory* of $\mathcal{M}$ is the set of closed $L$-conditions that are true in $\mathcal{M}$; it is denoted by $Th(\mathcal{M})$. A theory is called *complete* if it is of the form $Th(\mathcal{M})$ for some structure $\mathcal{M}$.

Let $T$ be an $L$-theory. A closed $L$-condition $E$ is called a *logical consequence of* $T$ and is written as $T \models E$ if $\mathcal{M} \models E$ for every model $\mathcal{M}$ of $T$. We say that a set $\Sigma$ of closed $L$-conditions *axiomatizes* $T$ if every closed $L$-condition in $\Sigma$ is a logical consequence of $T$ and vice versa. We say that $T$ is an *$\forall$ $L$-theory*, also called a *universal $L$-theory*, if there is a set $\Sigma$ of $\forall$ $L$-conditions such that $\Sigma$ axiomatizes $T$. Similarly, we say that $T$ is an *$\forall\exists$ $L$-theory* if there is a set $\Sigma$ of $\forall\exists$ $L$-conditions such that $\Sigma$ axiomatizes $T$.

Given $L$-structures $\mathcal{M}$ and $\mathcal{N}$, we say that $\mathcal{M}$ and $\mathcal{N}$ are *elementarily equivalent*, denoted by $\mathcal{M} \equiv \mathcal{N}$, if $Th(\mathcal{M}) = Th(\mathcal{N})$. Equivalently, $\sigma^{\mathcal{M}} = \sigma^{\mathcal{N}}$ for every $L$-sentence $\sigma$. A map $F$ from a subset of $M$ into $N$ is called an *elementary map from $\mathcal{M}$ into $\mathcal{N}$* if for all $L$-formulas $\varphi(x_1, \cdots, x_n)$ and $a_1, \cdots, a_n$ in the domain

of $F$, we have

$$\varphi^{\mathcal{M}}(a_1, \cdots, a_n) = \varphi^{\mathcal{N}}(F(a_1), \cdots, F(a_n)).$$

An elementary map from $\mathcal{M}$ into $\mathcal{N}$ whose domain is $M$ is said to be an *elementary embedding from $\mathcal{M}$ into $\mathcal{N}$*. Suppose $\mathcal{M} \subseteq \mathcal{N}$. We say that $\mathcal{M}$ is an *elementary substructure of $\mathcal{N}$*, denoted by $\mathcal{M} \preceq \mathcal{N}$, if the inclusion map from $M$ into $N$ is an elementary embedding from $M$ into $N$. We also call $\mathcal{N}$ an *elementary extension of $\mathcal{M}$* in this case.

Let $L' \supseteq L$ be another signature. If $\mathcal{M}'$ is an $L'$-structure, then by ignoring the interpretations of the symbols in $L' \backslash L$, we get an $L$-structure $\mathcal{M}$. We call $\mathcal{M}$ a *reduct* of $\mathcal{M}'$, denoted by $\mathcal{M}'|_L$, and call $\mathcal{M}'$ an *expansion* of $\mathcal{M}$.

A set $\mathcal{S}$ of $L$-formulas is *dense with respect to logical distance* if for every $L$-formula $\varphi(x_1, \cdots, x_n)$ and every $\epsilon > 0$ there is $\psi(x_1, \cdots, x_n) \in \mathcal{S}$ such that the logical distance between $\varphi$ and $\psi$ is less than $\epsilon$.

**1.3.1 Proposition** (Tarski-Vaught Test for $\preceq$, [4, Proposition 4.5])**.** *Let $\mathcal{S}$ be a set of $L$-formulas which is dense with respect to logical distance. For $L$-structures $\mathcal{M}$ and $\mathcal{N}$ with $\mathcal{M} \subseteq \mathcal{N}$, the following are equivalent:*

*(i)* $\mathcal{M} \preceq \mathcal{N}$.

*(ii) For every $L$-formula $\varphi(x_1, \cdots, x_n, y) \in \mathcal{S}$ and $a_1, \cdots, a_n \in M$, we have*

$$\inf\{\varphi^{\mathcal{N}}(a_1, \cdots, a_n, b) \mid b \in N\} = \inf\{\varphi^{\mathcal{N}}(a_1, \cdots, a_n, c) \mid c \in M\}.$$

□

**1.3.2 Theorem** (Compactness Theorem, [4, Theorem 5.8])**.** *Let $T$ be an $L$-theory. If every finite subset of $T$ has a model, then $T$ has a model.* □

For a class $\mathcal{C}$ of $L$-structures, we say that $\mathcal{C}$ is *axiomatizable* in $L$ if there is an $L$-theory $T$ such that $\mathcal{C}$ is the class of all models of $T$. In this situation, we call $T$ a set of *axioms for $\mathcal{C}$* in $L$.

## 1.4 Constructions of models

In this section, we list some results around constructions of models. Some of this material is taken from [4, Chapter 7]. We fix a signature $L$ for metric structures.

Let $\Lambda$ be a linear ordering. A $\Lambda$-*chain* of $L$-structures is a family of $L$-structures $(\mathcal{M}_\lambda \mid \lambda \in \Lambda)$ such that $\mathcal{M}_\lambda \subseteq \mathcal{M}_\mu$ whenever $\lambda < \mu$. The *union* of the $\Lambda$-chain, denoted by $\bigcup_{\lambda \in \Lambda} \mathcal{M}_\lambda$, is the $L$-structure whose

underlying set is the completion of the metric space $\bigcup_{\lambda\in\Lambda} M_\lambda$. A $\Lambda$-chain is called an *elementary chain* if $\mathcal{M}_\lambda \preceq \mathcal{M}_\mu$ whenever $\lambda < \mu$.

**1.4.1 Proposition** ([4, Proposition 7.2])**.** *If $(\mathcal{M}_\lambda \mid \lambda \in \Lambda)$ is an elementary chain, then $\mathcal{M}_\lambda \preceq \bigcup_{\lambda\in\Lambda} \mathcal{M}_\lambda$ for each $\lambda \in \Lambda$.* □

**1.4.2 Definition.** *Suppose that $\mathcal{M}$ is an $L$-structure. The **atomic diagram** of $\mathcal{M}$ is $\{\varphi(a) = r \mid \varphi$ is an atomic $L$-formula, $r \in [0,1]$, $a \in M^n$, and $\mathcal{M} \models \varphi(a) = r\}$. The **elementary diagram** of $\mathcal{M}$ is $\{\varphi(a) = r \mid \varphi$ is an $L$-formula, $r \in [0,1]$, $a \in M^n$, and $\mathcal{M} \models \varphi(a) = r\}$. We write $\mathrm{Diag}(\mathcal{M})$ and $\mathrm{Diag}_{\mathrm{el}}(\mathcal{M})$ for the atomic and elementary diagram of $\mathcal{M}$, respectively.*

For an atomic diagram and an elementary diagram, we have the following lemma:

**1.4.3 Lemma.** *Let $\mathcal{M}$ be an $L$-structure. Suppose that $\mathcal{N}$ is an $L(M)$-structure.*

*(i) If $\mathcal{N} \models \mathrm{Diag}(\mathcal{M})$, then there is an $L$-embedding of $\mathcal{M}$ into $\mathcal{N}|_L$.*

*(ii) If $\mathcal{N} \models \mathrm{Diag}_{\mathrm{el}}(\mathcal{M})$, then there is an elementary $L$-embedding of $\mathcal{M}$ into $\mathcal{N}|_L$.*

*Proof.* The proof is similar to the first order case. See [25, Lemma 2.3.3]. □

**1.4.4 Proposition** (Downward Löwenheim-Skolem Theorem, [4, Proposition 7.3])**.** *Let $L$ be a signature for metric structures and let $\kappa$ be an infinite cardinal such that $\mathrm{Card}(L) \leq \kappa$. Let $\mathcal{M}$ be an $L$-structure and let $A$ be a subset of $M$ whose density character is $\leq \kappa$. Then there is a substructure $\mathcal{N}$ of $\mathcal{M}$ satisfying:*

*(i) $A \subseteq N \subseteq M$;*

*(ii) $\mathcal{N} \preceq \mathcal{M}$;*

*(iii) the density character of $N$ is $\leq \kappa$.* □

Let $\kappa$ be an infinite cardinal and let $\mathcal{M}$ be an $L$-structure. For a set $\Gamma(x)$ of $L$-conditions, we say that $\Gamma(x)$ is *satisfiable in* $\mathcal{M}$ if there is $a \in M^n$ such that $\mathcal{M} \models \Gamma[a]$. We say that $\mathcal{M}$ is *$\kappa$-saturated* if the following holds: for every subset $A$ of $M$ with cardinality $< \kappa$ and every set $\Gamma(x)$ of $L(A)$-conditions, if every finite subset of $\Gamma(x)$ is satisfiable in $(\mathcal{M}, a)_{a\in A}$, then $\Gamma(x)$ is satisfiable in $(\mathcal{M}, a)_{a\in A}$.

**1.4.5 Proposition** ([4, Proposition 7.10])**.** *Let $\mathcal{M}$ be an $L$-structure. Then for every infinite cardinal $\kappa$, the structure $\mathcal{M}$ has a $\kappa$-saturated elementary extension.* □

We say that $\mathcal{M}$ is *strongly $\kappa$-homogeneous* if the following holds: for every subset $A$ of $M$ with $\mathrm{Card}(A) < \kappa$ and all $f$ and $g\colon A \to M$ such that

$$(\mathcal{M}, f(a))_{a\in A} \equiv (\mathcal{M}, g(a))_{a\in A},$$

we have

$$(\mathcal{M}, f(a))_{a\in A} \cong (\mathcal{M}, g(a))_{a\in A}.$$

**1.4.6 Proposition.** *Let $\mathcal{M}$ be an $L$-structure. Then for every infinite cardinal $\kappa$, the structure $\mathcal{M}$ has a $\kappa$-saturated elementary extension $\mathcal{N}$ that is strongly $\kappa$-homogeneous.*

*Proof.* This follows from [4, Proposition 7.12]. □

For a complete theory $T$ in $L$, a *$\kappa$-universal domain for $T$* (also called a monster model) is a $\kappa$-saturated and strongly $\kappa$-homogeneous model of $T$. Suppose that $\mathbb{M}$ is a $\kappa$-universal domain for $T$. For a subset $A$ of $\mathbb{M}$, we say that $A$ is *small* if $\mathrm{Card}(A) < \kappa$.

**1.4.7 Proposition.** *For every complete theory $T$ in $L$, there is a $\kappa$-universal domain for $T$ for every infinite cardinal $\kappa$.*

*Proof.* This follows from Proposition 1.4.6. □

## 1.5 Spaces of types

We fix a signature $L$ for metric structures and an $L$-theory $T$. In this section, we introduce the spaces of types. In particular, we define the logic topology on types and the metric topology on types, taken from [4].

Let $\mathcal{M}$ be a model of $T$ and $A$ a subset of $M$. Let $\mathcal{M}_A$ denote the $L(A)$-structure $(\mathcal{M}, a)_{a\in A}$ and let $T_A$ denote the $L(A)$-theory of $\mathcal{M}_A$. For $b = (b_1, \cdots, b_n) \in M^n$, the *$n$-type of $b$ over $A$*, denoted by $p = \mathrm{tp}_{\mathcal{M}}(b/A)$, is the set of all $L(A)$-conditions $E(x_1, \cdots, x_n)$ such that $\mathcal{M}_A \models E[b_1, \cdots, b_n]$. In this situation, we also say that *$b$ realizes $p$ in $\mathcal{M}$*. The set of all these kinds of $n$-types over $A$ is denoted by $S_n(T_A)$ or $S_n^T(A)$ (or simply just by $S_n(A)$ if the context makes $T$ clear).

### The logic topology on types

Let $T_A$ be as above. For an $L(A)$-formula $\varphi(x)$ and $\epsilon > 0$, we define $[\varphi < \epsilon]$ to be the set:

$$\{p \in S_n(T_A) \mid \text{ the condition } (\varphi \leq \delta) \text{ is in } p \text{ for some } 0 \leq \delta < \epsilon\}.$$

The *logic topology* on $S_n(T_A)$ is the topology generated by the following basic open sets. For every $p \in S_n(T_A)$, the basic open neighborhoods of $p$ are the sets of the form $[\varphi < \epsilon]$ where $(\varphi = 0)$ is in $p$ and $\epsilon > 0$.

**1.5.1 Proposition** ([4, Proposition 8.6]). *The type space $S_n(T_A)$ equipped with the logic topology is hausdorff and compact for each $n \geq 1$.* □

### The metric topology on types

Let $T_A$ be as above, and let $\mathcal{M}_A = (\mathcal{M}, a)_{a \in A}$ be a model of $T_A$ in which every type in $S_n(T_A)$ is realized. Such a model exists by Proposition 1.4.7. For all $p, q \in S_n(T_A)$, we define $d(p, q)$ to be

$$\inf\big\{ \max_{1 \leq i \leq n} d(b_i, c_i) \mid \mathcal{M}_A \models p[b_1, \cdots, b_n] \text{ and } \mathcal{M}_A \models q[c_1, \cdots, c_n] \big\}.$$

Note that $d(p, q)$ does not depend on $\mathcal{M}_A$. The *metric topology* on $S_n(T_A)$ is the topology induced by the metric $d$ between types. This metric $d$ is also called the *d-metric* between types.

**1.5.2 Proposition** ([4, Proposition 8.7]). *The metric topology is finer than the logic topology on $S_n(T_A)$ for each $n \geq 1$.* □

**1.5.3 Proposition** ([4, Proposition 8.8]). *The metric space $\big(S_n(T_A), d\big)$ is complete for each $n \geq 1$.* □

## 1.6 Definability in metric structures

This section presents the notion of definability in continuous first order model theory, taken from [4]. We fix a signature $L$ for metric structures.

### Definable predicates, sets and functions

Let $\mathcal{M}$ be an $L$-structure and let $A$ be a subset of $M$. For a predicate $P\colon M^n \to [0, 1]$, we say that $P$ *is definable in $\mathcal{M}$ over $A$* if there is a sequence $\{\varphi_k(x)\}_{k \in \mathbb{N}}$ of $L(A)$-formulas such that the predicates $\varphi_k^{\mathcal{M}}(x)$ converge to $P(x)$ uniformly on $M^n$; *i.e.*,

$$\forall \epsilon > 0 \ \exists N > 0 \ \forall k > N \ \forall x \in M^n \ (|\varphi_k^{\mathcal{M}}(x) - P(x)| \leq \epsilon).$$

Let $\mathcal{M}$ be an $L$-structure and $D \subseteq M^n$. The predicate giving the distance to $D$ in $M^n$ is given by

$$\operatorname{dist}(x, D) = \inf_{y \in D} d(x, y).$$

Recall that $d(x, y) := \max_{1 \leq i \leq n} d(x_i, y_i)$ for tuples $x, y \in M^n$. The *distance predicate* $\operatorname{dist}(x, D)$ in model

theory for metric structures is important and is characterized by the following two axioms in continuous logic.

Let $L(P)$ be a signature obtained from $L$ by adding a new predicate symbol $P$. Consider the following two $L(P)$-conditions:

$(E_1)$ $\sup_x \inf_y \max\big(P(y), |P(x) - d(x,y)|\big) = 0$

$(E_2)$ $\sup_x |P(x) - \inf_y \min\big(P(y) + d(x,y), 1\big)| = 0$

Note that for every nonempty $D \subseteq M^n$, if $F(x)$ is defined to be $\mathrm{dist}(x, D)$ in the $L$-structure $\mathcal{M}$, then the $L(P)$-structure $(\mathcal{M}, F)$ satisfies $(E_1)$ and $(E_2)$.

**1.6.1 Theorem** ([4, Theorem 9.12])**.** *Let $(\mathcal{M}, F)$ be an $L(P)$-structure satisfying $(E_1)$ and $(E_2)$. Let $D = \{x \in M^n \mid F(x) = 0\}$ be the zeroset of $F$. Then $D$ is nonempty and $F(x) = \mathrm{dist}(x, D)$ for every $x \in M^n$.* □

A subset $D \subseteq M^n$ is called *type-definable in $\mathcal{M}$ over $A$* if there is a set $\Gamma(x)$ of $L(A)$-formulas such that for every $a \in M^n$ we have $a \in D$ if and only if $\varphi^{\mathcal{M}}(a) = 0$ for every $\varphi(x) \in \Gamma(x)$. In this case, we also say *$D$ is type-defined by $\Gamma(x)$*. We say that $D$ is a *zeroset in $\mathcal{M}$ over $A$* if there is a definable predicate $P\colon M^n \to [0,1]$ over $A$ such that $D = \{x \in M^n \mid P(x) = 0\}$.

**1.6.2 Proposition** ([4, Proposition 9.14])**.** *Let $D$ be a subset of $M^n$ and $A \subseteq M$. Then $D$ is a zeroset in $\mathcal{M}$ over $A$ if and only if there is a sequence $\{\varphi_k(x)\}_{k\in\mathbb{N}}$ of $L(A)$-formulas such that $D$ is type-defined by $\{\varphi_k(x)\}_{k\in\mathbb{N}}$.* □

A closed subset $D$ of $M^n$ is called *definable in $\mathcal{M}$ over $A$* (or $D$ is an *$A$-definable set*) if the distance predicate $\mathrm{dist}(x, D)$ is definable in $\mathcal{M}$ over $A$. The following result shows an important property of a definable set.

**1.6.3 Theorem** ([4, Theorem 9.17])**.** *Let $\mathcal{M}$ be an $L$-structure and let $A$ be a subset of $M$. Let $D$ be a closed subset of $M^n$. Then $D$ is definable in $\mathcal{M}$ over $A$ if and only if the following holds: for every predicate $P\colon M^m \times M^n \to [0,1]$ which is definable in $\mathcal{M}$ over $A$, the predicate $Q\colon M^m \to [0,1]$ defined by*

$$Q(x) := \inf\{P(x,y) \mid y \in D\}$$

*is definable in $\mathcal{M}$ over $A$.* □

**1.6.4 Proposition** ([4, Proposition 9.19]). *Suppose $P$ is a definable predicate in $\mathcal{M}$ over $A$. The zeroset $D$ of $P$ is definable in $\mathcal{M}$ over $A$ if and only if $P$ satisfies the following statement:*

$$\forall \epsilon > 0 \ \exists \delta > 0 \ \forall x \in M^n \ \big(P(x) \leq \delta \Rightarrow \operatorname{dist}(x, D) \leq \epsilon\big).$$

□

A function $f\colon M^n \to M$ is called *definable in $\mathcal{M}$ over $A$* if the function $d(f(x), y)$ is a definable predicate in $\mathcal{M}$ over $A$.

**1.6.5 Proposition** ([4, Proposition 9.25]). *Let $\mathcal{M}$ be an $L$-structure and let $A$ be a subset of $M$. If $f\colon M^n \to M$ is definable in $\mathcal{M}$ over $A$, then*

*(i) If $\mathcal{N} \preceq \mathcal{M}$ and $A \subseteq N$, then the restriction of $f$ to $N^n$ is into $N$ and is definable in $\mathcal{N}$ over $A$.*

*(ii) If $\mathcal{N} \succeq \mathcal{M}$, then $f$ can be extended to $g\colon N^n \to N$ which is definable in $\mathcal{N}$ over $A$.* □

## Extension by definitions

Let $L$ be a signature for metric structures and let $T$ be a consistent $L$-theory. Let $L'$ be a signature such that $L' \supseteq L$ and let $T'$ be an $L'$-theory. We say that $T$ is the *restriction of $T'$ to $L$* (or $T'$ is a *conservative extension of $T$*) if for every closed $L$-condition $E$ we have $T' \models E$ iff $T \models E$.

For an $L'$-formula $\varphi(x)$, we say that $\varphi(x)$ is *defined in $T'$ over $L$* if for every $\epsilon > 0$ there is an $L$-formula $\psi(x)$ such that

$$T' \models (\sup_x |\varphi(x) - \psi(x)|) \leq \epsilon.$$

For an $n$-ary predicate symbol $P \in L'$, we say that $P$ is *defined in $T'$ over $L$* if the formula $P(x)$ is defined in $T'$ over $L$. For an $n$-ary function symbol $f \in L'$, we say that $f$ is *defined in $T'$ over $L$* if the formula $d(f(x), y)$ is defined in $T'$ over $L$. For a constant symbol $c$ in $L'$, we say that $c$ is *defined in $T'$ over $L$* if the formula $d(c, x)$ is defined in $T'$ over $L$.

For a conservative extension $T'$ of $T$, we say that $T'$ is *an extension by definitions of $T$* if every nonlogical symbol in $L'$ is defined in $T'$ over $L$.

**1.6.6 Proposition** ([4, Proposition 9.28]). *Let $L' \supseteq L$. Let $T$ be an $L$-theory and let $T'$ be an $L'$-theory. Suppose that $T'$ is an extension by definitions of $T$. Then every $L'$-formula is defined in $T'$ over $L$.* □

**1.6.7 Corollary** ([4, Corollary 9.31]). *Let $L'$, $L$, $T$, $T'$ be as in the above proposition. Then every model of $T$ has a (unique) expansion to a model of $T'$.* □

The following theorem generalizes Beth's Definability Theorem to continuous logic.

**1.6.8 Theorem** ([4, Theorem 9.32]). *Let $T'$ be an $L'$-theory and $L \subseteq L'$. Let $S$ be a nonlogical symbol in $L'$. Suppose that $S$ is implicitly defined in $T'$ over $L$. That is, assume that if $\mathcal{M}$ and $\mathcal{N}$ are models of $T'$ for which $\mathcal{M}|_L = \mathcal{N}|_L$, then one always has $S^{\mathcal{M}} = S^{\mathcal{N}}$. Then $S$ is defined in $T'$ over $L$.* □

**1.6.9 Corollary** ([4, Corollary 9.33]). *Let $L \subseteq L'$, let $T$ be a consistent $L$-theory and let $T'$ be a consistent $L'$-theory. Then $T'$ is an extension by definitions of $T$ if and only if every model of $T$ has a unique expansion which is a model of $T'$.* □

## Definable and algebraic closures

Let $\mathcal{M}$ be an $L$-structure and $A \subseteq M$. For $a \in M^n$, we say that $a$ is *definable in $\mathcal{M}$ over $A$* (or *$A$-definable in $\mathcal{M}$*) if the singleton set $\{a\}$ is definable in $\mathcal{M}$ over $A$. We say that $a$ is *algebraic in $\mathcal{M}$ over $A$* (or *$A$-algebraic in $\mathcal{M}$*) if there is a compact set $C \subseteq M^n$ such that $a \in C$ and $C$ is definable in $\mathcal{M}$ over $A$.

Let $\mathrm{dcl}_{\mathcal{M}}(A)$ denote the set of all $A$-definable elements of $M$; we call it the *definable closure of $A$ in $\mathcal{M}$*. Similarly, the *algebraic closure of $A$ in $\mathcal{M}$*, denoted by $\mathrm{acl}_{\mathcal{M}}(A)$, is the set of all $A$-algebraic elements of $M$.

**1.6.10 Proposition** ([4]). *Let $\mathcal{M}$ be an $L$-structure, $a \in M^n$, and $A$ a subset of $M$. Then*

- *for every $\mathcal{N} \succeq \mathcal{M}$*
$$\mathrm{dcl}_{\mathcal{M}}(A) = \mathrm{dcl}_{\mathcal{N}}(A) \subseteq \mathrm{acl}_{\mathcal{N}}(A) = \mathrm{acl}_{\mathcal{M}}(A).$$
- *$a$ is $A$-definable if and only if for every $\mathcal{N} \succeq \mathcal{M}$ the only realization of $\mathrm{tp}_{\mathcal{M}}(a/A)$ in $\mathcal{N}$ is $a$.*
- *$a$ is $A$-algebraic if and only if for every $\mathcal{N} \succeq \mathcal{M}$ the set of realizations of $\mathrm{tp}_{\mathcal{M}}(a/A)$ in $\mathcal{N}$ is compact. Moreover, if $\mathcal{N}$ is $\kappa$-saturated with $\kappa$ uncountable, then $a$ is $A$-algebraic if and only if the set of realizations of $\mathrm{tp}_{\mathcal{M}}(a/A)$ in $\mathcal{N}$ has density character $< \kappa$.*

*Proof.* See [4, Corollary 10.5, Exercise 10.7, and Exercise 10.8]. □

**1.6.11 Proposition** (Properties of *dcl* and *acl*, [4]). *Let $\mathcal{M}$ be an $L$-structure and let $A$, $B$ be subsets of $M$. Then*

*(i)* $A \subseteq \mathrm{dcl}(A)$ *and* $A \subseteq \mathrm{acl}(A)$.

*(ii) If* $A \subseteq \mathrm{dcl}(B)$, *then* $\mathrm{dcl}(A) \subseteq \mathrm{dcl}(B)$; *if* $A \subseteq \mathrm{acl}(B)$, *then* $\mathrm{acl}(A) \subseteq \mathrm{acl}(B)$.

*(iii)* $\mathrm{dcl}(A) = \mathrm{dcl}(\mathrm{dcl}(A))$ *and* $\mathrm{acl}(\mathrm{acl}(A)) = \mathrm{acl}(A)$.

*(Here we simply write* dcl *and* acl *for* $\mathrm{dcl}_{\mathcal{M}}$ *and* $\mathrm{acl}_{\mathcal{M}}$ *respectively.)*

*Proof.* $(i)$ and $(ii)$ are [4, Exercise 10.10 and Proposition 10.11]. $(iii)$ follows from $(i)$ and $(ii)$. □

## 1.7 Inductive theories

In this section, we give definitions and basic results about inductive theories, adapted from the thesis of K. Schoretsanitis [29]. Let $L$ be a signature for metric structures.

**1.7.1 Definition.** *(i) Let $\mathcal{A}$ and $\mathcal{B}$ be $L$-structures and $\mathcal{A} \subseteq \mathcal{B}$. We say that $\mathcal{A}$ is* ***existentially closed (e.c.)*** *in $\mathcal{B}$, write $\mathcal{A} \subseteq_{e.c.} \mathcal{B}$, if for every quantifier free $L$-formula $\varphi(x, y)$ and $a \in A^n$ we have*

$$(\inf_x \varphi(x, a))^{\mathcal{A}} = (\inf_x \varphi(x, a))^{\mathcal{B}}.$$

*(ii) Let $K$ be a class of $L$-structures. For $\mathcal{A} \in K$, we say that $\mathcal{A}$ is* ***e.c.*** *in $K$ if for every $\mathcal{B} \in K$ such that $\mathcal{A} \subseteq \mathcal{B}$, we have that $\mathcal{A} \subseteq_{e.c.} \mathcal{B}$.*

*(iii) If $K$ is the class of all models of an $L$-theory $T$, then we say that a model $\mathcal{M}$ of $T$ is an* ***e.c.*** *model if $\mathcal{M}$ is e.c. in $K$.*

**1.7.2 Definition.** *A class of $L$-structures $K$ is called an* ***inductive class*** *if it satisfies the following:*

*(i) $K$ is closed under isomorphism.*

*(ii) $K$ is closed under unions of chains.*

An $L$-theory $T$ is *inductive* if the class $K$ of all models of $T$ is inductive.

**1.7.3 Theorem** ([29, Theorem 4.9])**.** *Let $K$ be an inductive class of $L$-structures and let $\mathcal{A} \in K$. Then there is an e.c. structure $\mathcal{B} \in K$ such that $\mathcal{A} \subseteq \mathcal{B}$.* □

**1.7.4 Theorem** ([29, Theorem 4.10])**.** *Let $T$ be an $L$-theory. Then $T$ is inductive if and only if $T$ is an $\forall\exists$ $L$-theory.* □

For an $L$-theory, let $T_\forall$ denote the set of all logical consequences $E$ of $T$ such that $E$ is a closed $\forall$ $L$-condition.

**1.7.5 Lemma** ([29, Lemma 4.12])**.** *Let $T$ be an $L$-theory. If $\mathcal{M} \models T_\forall$, then there is $\mathcal{N} \models T$ such that $\mathcal{M} \subseteq \mathcal{N}$.* □

**1.7.6 Theorem.** *Let $T$ be an $L$-theory. Then $T$ is a universal theory if and only if for every model $\mathcal{M}$ of $T$ one has that every $L$-substructure $\mathcal{N}$ of $\mathcal{M}$ is a model of $T$.*

*Proof.* $\Longrightarrow$: This is trivial.

$\Longleftarrow$: Use Lemma 1.7.5. □

**1.7.7 Proposition** ([29, Proposition 4.13])**.** *Let $T$ be an $\forall\exists$ theory and let $\mathcal{M} \models T$. Then $\mathcal{M}$ is an e.c. model of $T$ if and only if $\mathcal{M}$ is an e.c. model of $T_\forall$.* □

## 1.8 Model completeness

This section presents *model completeness* and Lindström's test for model completeness in continuous logic.

### Model completeness

Let $L$ be a signature for metric structures. A theory is *model complete* if every embedding between its models is elementary.

**1.8.1 Proposition** ([29, Proposition 4.16])**.** *Let $T$ be an $L$-theory. If $T$ is model complete, then $T$ is an $\forall\exists$ theory.* □

Let $\kappa$ be an infinite cardinal. We say that an $L$-theory $T$ is *$\kappa$-categorical* if $T$ has exactly one model with density character $\kappa$ up to isomorphism.

**1.8.2 Theorem** (Lindström's test for model completeness)**.** *Let $T$ be an $\forall\exists$ $L$-theory which has no compact models. If $T$ is $\kappa$-categorical for some cardinal $\kappa \geq \mathrm{Card}(L)$, then $T$ is model complete.*

Toward the proof of Theorem 1.8.2, we need the following lemmas.

**1.8.3 Lemma.** *Given $L$-structures $\mathcal{A}, \mathcal{B}, \mathcal{C}$, suppose that $\mathcal{A} \subseteq \mathcal{C}$ and $\mathcal{A} \subseteq_{e.c.} \mathcal{B}$. Then there is an $L$-structure $\mathcal{D}$ with an embedding $f\colon \mathcal{B} \to \mathcal{D}$ and an elementary embedding $g\colon \mathcal{C} \to \mathcal{D}$ such that $f(a) = g(a)$ for every $a \in \mathcal{A}$.*

*Proof.* We may assume that $B \cap C = A$. Let $L(B \cup C)$ be the extension of $L$ by adding to $L$ constant symbols $d$ for every $d \in B \cup C$. We will use Theorem 1.3.2 (Compactness Theorem) to show that $\mathrm{Diag}(\mathcal{B}) \cup \mathrm{Diag}_{\mathrm{el}}(\mathcal{C})$ is consistent. Take any finite subset $\Gamma$ of $\mathrm{Diag}(\mathcal{B}) \cup \mathrm{Diag}_{\mathrm{el}}(\mathcal{C})$. We may assume that $\Gamma$ is $\{\varphi_1(a_1, b_1) = 0, \cdots, \varphi_k(a_k, b_k) = 0, \psi_1(c_1) = 0, \cdots, \psi_l(c_l) = 0\}$, where $\varphi_1, \cdots, \varphi_k$ are atomic $L$-formulas, $a_1, \cdots, a_k$ are tuples in $A$, $b_1, \cdots, b_k$ are tuples in $B \setminus A$, $\psi_1, \cdots, \psi_l$ are $L$-formulas, and $c_1, \cdots, c_l$ are tuples in $C$. Let

$\varphi$ denote $\max_{1\leq i\leq k}\varphi_i$ and let $\psi$ denote $\max_{1\leq j\leq l}\{\psi_j\}$. Then $\Gamma$ is equivalent to $\max(\varphi(b,a),\psi(c))=0$ for some tuples $a$ in $A$, $b$ in $B\setminus A$ and $c$ in $C$. Since $\varphi$ is quantifier free and $\mathcal{A}\subseteq_{e.c.}\mathcal{B}$, for each $\epsilon>0$ there is a tuple $a_\epsilon$ in $A$ such that $\mathcal{A}\models\varphi(a_\epsilon,a)\leq\epsilon$. Then because $\mathcal{A}\subseteq\mathcal{C}$ and $\varphi$ is quantifier free, we have $\mathcal{C}\models\varphi(a_\epsilon,a)\leq\epsilon$ for each $\epsilon>0$. Hence there is an elementary extension $\mathcal{C}'$ of $\mathcal{C}$ and a tuple $b'$ in $C'\setminus C$ such that $\mathcal{C}'\models\varphi(b',a)=0$. In the $L$-structure $\mathcal{C}'$, we interpret $b$ as $b'$, interpret every $f\in B$ which is not in $A\cup b$ as arbitrary element in $C'$, and interpret every $g\in C$ as itself in $C'$. Then $\mathcal{C}'$ is expanded to be an $L(B\cup C)$-structure and $\mathcal{C}'\models\max\big(\varphi(b^{\mathcal{C}'},a^{\mathcal{C}'}),\psi(c^{\mathcal{C}'})\big)$. Hence $\Gamma$ is satisfiable, whence $\mathrm{Diag}(\mathcal{B})\cup\mathrm{Diag}_{\mathrm{el}}(\mathcal{C})$ is consistent. Let $\mathcal{D}$ be a model of $\mathrm{Diag}(\mathcal{B})\cup\mathrm{Diag}_{\mathrm{el}}(\mathcal{C})$. Then $\mathcal{D}$ is an $L_{B\cup C}$-structure. Since $A=B\cap C$, there is an embedding $f\colon\mathcal{B}\to\mathcal{D}$ and an elementary embedding $g\colon\mathcal{C}\to\mathcal{D}$ such that $f(a)=g(a)=a^{\mathcal{D}}$ for every $a\in\mathcal{A}$. □

**1.8.4 Lemma.** *Given $L$-structures $\mathcal{A}$ and $\mathcal{B}$ with $\mathcal{A}\subseteq_{e.c.}\mathcal{B}$, there is an $L$-structure $\mathcal{C}$ such that $\mathcal{B}\subseteq\mathcal{C}$ and $\mathcal{A}\preceq\mathcal{C}$.*

*Proof.* By Lemma 1.8.3, there is an $L$-structure $\mathcal{C}$ with an embedding $f\colon\mathcal{B}\to\mathcal{C}$ and an elementary embedding $g\colon\mathcal{A}\to\mathcal{C}$ such that $f(a)=g(a)$ for every $a\in\mathcal{A}$. Therefore, $f[\mathcal{B}]\subseteq\mathcal{C}$ and $f[\mathcal{A}]=g[\mathcal{A}]\preceq\mathcal{C}$. After identifying $\mathcal{B}$ with its isomorphic copy $f[\mathcal{B}]$ in $\mathcal{C}$, we have $\mathcal{B}\subseteq\mathcal{C}$ and $\mathcal{A}\preceq\mathcal{C}$. □

**1.8.5 Lemma.** *Let $T$ be an $L$-theory. If $T$ is $\lambda$-categorical for some infinite cardinal $\lambda$ and has an e.c. model $\mathcal{M}$ with density character $\geq\lambda$, then every model of $T$ with density character $\geq\lambda$ is e.c.*

*Proof.* First, we show that the unique model $\mathcal{A}$ of $T$ of density character $\lambda$ is e.c. By Theorem 1.4.4 (Downward Löwenheim-Skolem Theorem), we may assume that $\mathcal{A}\preceq\mathcal{M}$. Then, $\mathcal{A}\subseteq_{e.c.}\mathcal{M}$. Let $\mathcal{B}\supseteq\mathcal{A}$ be a model of $T$. By Lemma 1.8.3, there is a $\mathcal{C}\models T$ with an embedding $f\colon\mathcal{M}\to\mathcal{C}$ and an elementary embedding $g\colon\mathcal{B}\to\mathcal{C}$ such that $f(a)=g(a)$ for every $a\in A$. Since $\mathcal{M}$ is an e.c. model and $\mathcal{C}\models T$, we have $f[\mathcal{M}]\subseteq_{e.c.}\mathcal{C}$. Then for every quantifier free formula $\varphi$ and every $a\in A$, we have

$$(\inf_x\varphi(x,a))^{\mathcal{A}}=(\inf_x\varphi(x,a))^{\mathcal{M}}=(\inf_x\varphi(x,f(a)))^{\mathcal{C}}=(\inf_x\varphi(x,g(a)))^{\mathcal{C}}=(\inf_x\varphi(x,a))^{\mathcal{B}},$$

which implies $\mathcal{A}\subseteq_{e.c.}\mathcal{B}$.

Second, let $\mathcal{N}$ be a model of $T$ with density character $\geq\lambda$. Let $\mathcal{B}\supseteq\mathcal{N}$ be a model of $T$. Consider any quantifier free formula $\varphi(x,y)$ where $y$ is $n$-ary, and any $a\in N^n$. By $\lambda$-categoricity, we may assume that $a\in\mathcal{A}\preceq\mathcal{N}$ where $\mathcal{A}$ is the unique model of density character $\lambda$. Then as shown above, $\mathcal{A}\subseteq_{e.c.}\mathcal{B}$, whence $(\inf_x\varphi(x,a))^{\mathcal{B}}=(\inf_x\varphi(x,a))^{\mathcal{A}}$. Since $(\inf_x\varphi(x,a))^{\mathcal{A}}=(\inf_x\varphi(x,a))^{\mathcal{N}}$, we have $(\inf_x\varphi(x,a))^{\mathcal{B}}=(\inf_x\varphi(x,a))^{\mathcal{N}}$. □

*Proof of Theorem 1.8.2.* By Theorem 1.7.4 and Theorem 1.7.3, the theory $T$ has an e.c. model of density character $\geq \kappa$. Let $\mathcal{M}_1$ and $\mathcal{M}_2$ be models of $T$ with density character $\geq \kappa$ such that $\mathcal{M}_1 \subseteq \mathcal{M}_2$. By Lemma 1.8.5, we have $\mathcal{M}_1, \mathcal{M}_2$ are e.c. models, whence $\mathcal{M}_1 \subseteq_{e.c.} \mathcal{M}_2$. By Lemma 1.8.4, there is $\mathcal{M}_3 \models T$ such that $\mathcal{M}_2 \subseteq \mathcal{M}_3$ and $\mathcal{M}_1 \preceq \mathcal{M}_3$. Then by Lemma 1.8.5, $\mathcal{M}_2 \subseteq_{e.c.} \mathcal{M}_3$. Inductively, we get a chain $(\mathcal{M}_n \mid n \in \mathbb{N})$ such that $\mathcal{M}_n \preceq \mathcal{M}_{n+2}$ for every $n \in \mathbb{N}$. Therefore, $\mathcal{M}_1 \subseteq \mathcal{M}_2 \preceq \overline{\bigcup \mathcal{M}_n}$ and $\mathcal{M}_1 \preceq \overline{\bigcup \mathcal{M}_n}$, whence $\mathcal{M}_1 \preceq \mathcal{M}_2$.

Now let $\mathcal{M}_1$ and $\mathcal{M}_2$ be arbitrary models of $T$ with $\mathcal{M}_1 \subseteq \mathcal{M}_2$. Because $T$ has no compact models, there is a model $\mathcal{M}_3$ of $T$ with density character $\geq \kappa$ such that $\mathcal{M}_1 \preceq \mathcal{M}_3$, therefore $\mathcal{M}_1 \subseteq_{e.c.} \mathcal{M}_3$. By Lemma 1.8.3, there is $\mathcal{M}_4 \models T$ with an embedding $f\colon \mathcal{M}_3 \to \mathcal{M}_4$ and an elementary embedding $g\colon \mathcal{M}_2 \to \mathcal{M}_4$ such that $f(a) = g(a)$ for every $a \in M_1$. Note that $\mathcal{M}_3$ and $\mathcal{M}_4$ are models of $T$ with density character $\geq \kappa$. Then as proved above, we have $f[\mathcal{M}_3] \preceq \mathcal{M}_4$, whence $g[\mathcal{M}_1] = f[\mathcal{M}_1] \preceq \mathcal{M}_4$. Since $g[\mathcal{M}_2] \preceq \mathcal{M}_4$, we have $g[\mathcal{M}_1] \preceq g[\mathcal{M}_2]$, whereby $\mathcal{M}_1 \preceq \mathcal{M}_2$. □

**1.8.6 Remark.** *This proof of Theorem 1.8.2 follows C. Ward Henson's unpublished manuscript, which is similar to the first order case; e.g., see [17, Theorem 7.3.4].*

## 1.9 Separable categoricity

Let $L$ be a signature for metric structures which has only a countable number of nonlogical symbols. Let $T$ be a complete $L$-theory.

An $L$-theory is called *separably categorical*, if it has exactly one separable model up to isomorphism. Since a theory might only have finite models, a separably categorical theory is not necessarily $\omega$-categorical. A separably categorical theory is always complete. If it has a compact model, then by the downward Löwenheim-Skolem theorem, it has exactly one model up to isomorphism, and hence it is complete. If it has no compact models, then it is complete by Vaught's test. An $\omega$-categorical theory is complete if and only if it has no finite models.

Let $p \in S_n(T)$. We say $p$ is *principal* if for every model $\mathcal{M}$ of $T$, the set of realizations of $p$ in $\mathcal{M}$, denoted by $p(\mathcal{M})$, is definable in $\mathcal{M}$ over $\emptyset$.

**1.9.1 Proposition** ([4, Proposition 12.4])**.** *Let $p \in S_n(T)$. Then $p$ is principal if and only if $p$ has the same neighborhoods in the logic topology as in the metric topology.* □

**1.9.2 Theorem** (Omitting Types Theorem, [4, Theorem 12.6])**.** *Let $T$ be a complete countable theory and let $p \in S_n(T)$. Then $p$ is principal if and only if $p$ is realized in every model of $T$.* □

A model $\mathcal{M}$ of $T$ is *atomic* if every type over the empty set realized in $\mathcal{M}$ is principal. A structure $\mathcal{M}$ is *approximately $\aleph_0$-saturated* if for every $n$-tuple $a$ in $M$, every type $p(x,a) \in S_1(a)$, and every $\epsilon > 0$, there is an $n$-tuple $a'$ in $M$ such that $d(a,a') \leq \epsilon$ and the type $p(x,a')$ is realized in $\mathcal{M}$.

**1.9.3 Theorem** (Ryll-Nardzewski Theorem)**.** *Let $T$ be a complete countable theory. Then the following are equivalent:*

*(i) $T$ is separably categorical.*

*(ii) Every $n$-type over the empty set is principal for all $n \in \mathbb{N}$.*

*(iii) Every model of $T$ is approximately $\aleph_0$-saturated.*

*(iv) The d-metric topology on $S_n(T)$ coincides with the logic topology for all $n \in \mathbb{N}$.*

*(v) The d-metric topology on $S_n(T)$ is compact for all $n \in \mathbb{N}$.*

*Proof.* This is [5, Fact 1.14]. □

**1.9.4 Corollary** ([4, Corollary 12.11])**.** *Suppose that $T$ is a separably categorical theory and $\mathcal{M}$ is the separable model of $T$. Then for all $a, b \in M^n$ and for every $\epsilon > 0$ there is an automorphism $F$ of $\mathcal{M}$ such that*

$$d\big(F(a), b\big) \leq d\big(\mathrm{tp}(a), \mathrm{tp}(b)\big) + \epsilon.$$

□

**1.9.5 Corollary** ([4, Corollary 12.13])**.** *Suppose $L \subseteq L'$ are countable signatures, $T'$ is complete in $L'$ and $T$ is its restriction to $L$. If $T'$ is $\omega$-categorical, then so is $T$.* □

**1.9.6 Proposition.** *Let $T$ be a separably categorical theory and let $\mathcal{M}$ be a separable model of $T$. For every closed subset $X \subseteq M^n$, if $F(X) = X$ for all $F \in \mathrm{Aut}(\mathcal{M})$, then $X$ is $\emptyset$-definable. In particular, a $\emptyset$-type-definable set in $\mathcal{M}$ is $\emptyset$-definable.*

*Proof.* Fix a closed set $X \subseteq M^n$ such that $F(X) = X$ for all $F \in \mathrm{Aut}(\mathcal{M})$. Define $D(a) := \mathrm{dist}(a, X)$ for all $a \in M^n$. We will show that $D$ is a $\emptyset$-definable predicate. For all $a, b \in M^n$, if $\mathrm{tp}(a) = \mathrm{tp}(b)$ then by Corollary 1.9.4, for all $\epsilon > 0$ there is an automorphism $F$ of $M$ such that $d\big(F(a), b\big) \leq \epsilon$. Since $F$ fixes $X$ setwise, we have $D(a) = D\big(F(a)\big)$. Consequently, we have $\big|D(a) - D(b)\big| \leq \epsilon$ for all $\epsilon$. Therefore, $D(a) = D(b)$. Now, we define $\tilde{D}\colon S_n(T) \to [0,1]$ by $\tilde{D}(p) = D(a)$, where $a \in M^n$ is any realization of the type $p \in S_n(T)$. As shown above, $\tilde{D}$ is well-defined. Let $C$ be the set $\big\{\mathrm{tp}(x) \mid x \in X\big\}$. Define $\Phi(p) := \mathrm{dist}(p, C)$ for all $p \in S_n(T)$. It

follows from the definition that $\Phi(p)$ is $d$-continuous. Since $T$ is separably categorical, the $d$-metric topology coincides with the logic topology. So $\Phi(p)$ is also continuous under the logic topology.

For every $r \in [0,1]$ and every $\mathrm{tp}(a)$ in $\tilde{D}^{-1}([0,r])$, we have $D(a) \leq r$ which means $\mathrm{dist}(a,X) \leq r$. Hence, for all $\epsilon > 0$, there exists $b \in X$ such that $d(a,b) < r+\epsilon$. Then $\Phi\big(\mathrm{tp}(a)\big) = \mathrm{dist}(\mathrm{tp}(a), C) \leq d\big(\mathrm{tp}(a), \mathrm{tp}(b)\big) < r+\epsilon$ for all $\epsilon > 0$. So $\Phi\big(\mathrm{tp}(a)\big) \leq r$. Therefore, $\tilde{D}^{-1}\big([0,r]\big) \subseteq \Phi^{-1}\big([0,r]\big)$.

Now for all $\mathrm{tp}(a) \in \Phi^{-1}\big([0,r]\big)$, we have $\Phi\big(\mathrm{tp}(a)\big) \leq r$, which means $\mathrm{dist}\big(\mathrm{tp}(a), C\big) \leq r$. For all $\epsilon > 0$, by Corollary 1.9.4 there exists $c \in X$ such that $d(a,c) < r+\epsilon$. Therefore, we get $\tilde{D}\big(\mathrm{tp}(a)\big) = D(a) = \mathrm{dist}(a,X) \leq r$. Then we have $\Phi^{-1}\big([0,r]\big) \subseteq \tilde{D}^{-1}\big([0,r]\big)$.

Overall, we get $\tilde{D}^{-1}\big([0,r]\big) = \Phi^{-1}\big([0,r]\big)$ for all $r \in [0,1]$. Since $\Phi$ is continuous with respect to the logic topology, $\tilde{D}$ is also continuous with respect to the logic topology. Hence $D$ is a $\emptyset$-definable predicate. □

## 1.10 Quantifier elimination

Let $L$ be a signature for metric structures and let $T$ be an $L$-theory. An $L$-formula $\varphi(x)$ is *approximable in $T$ by quantifier free formulas* if for every $\epsilon > 0$ there is a quantifier free formula $\psi(x)$ such that for every model $\mathcal{M}$ of $T$ and every $a \in M^n$,

$$|\varphi^{\mathcal{M}}(a) - \psi^{\mathcal{M}}(a)| \leq \epsilon.$$

The theory $T$ *admits quantifier elimination* if every $L$-formula is approximable in $T$ by quantifier free formulas.

**1.10.1 Remark.** *(i) Let $T$ be an $L$-theory and let $L(C)$ be an extension of $L$ by constants. If $T$ admits quantifier elimination in the signature $L$, then $T$ also admits quantifier elimination in the signature $L(C)$.*

*(ii) Let $T$ be an $L$-theory and let the $L$-theory $T' \supseteq T$. If $T$ admits quantifier elimination, then $T'$ also admits quantifier elimination.*

**1.10.2 Proposition** ([4, Proposition 13.6])**.** *Let $T$ be an $L$-theory. Then $T$ admits quantifier elimination if and only if the following statement holds:*

*for all models $\mathcal{M}$ and $\mathcal{N}$ of $T$, every embedding of a substructure of $\mathcal{M}$ into $\mathcal{N}$ can be extended to an embedding of $\mathcal{M}$ into an elementary extension of $\mathcal{N}$.*

*Moreover, if* $\mathrm{Card}(L) \leq \kappa$ *where $\kappa$ is an infinite cardinal, then it suffices to consider models of $T$ with density character $\leq \kappa$ in the above statement.* □

**1.10.3 Definition** ([6, Definition 4.15])**.** *We say that an $L$-theory $T$ has the* ***back-and-forth property*** *if for all $\aleph_0$-saturated models $\mathcal{M}$ and $\mathcal{N}$ of $T$, all $n$-tuples $a \in M^n$ and $b \in N^n$, and all $c \in M$, if $a$ and $b$ have the same quantifier free type (i.e., $\varphi^{\mathcal{M}}(a) = \varphi^{\mathcal{N}}(b)$ for all quantifier free formulas $\varphi$), then there is $d \in N$ such that $(a, c)$ and $(b, d)$ have the same quantifier free type.*

**1.10.4 Theorem.** *For an $L$-theory $T$, the following are equivalent:*

*(i) The theory $T$ admits quantifier elimination.*

*(ii) The theory $T$ has the back-and-forth property.*

*(iii) Every type in $S_n(T)$ is determined by quantifier free formulas.*

*Proof.* $(i) \Leftrightarrow (ii)$: This is [6, Theorem 4.16].

$(i) \Rightarrow (iii)$: This is trivial.

$(iii) \Rightarrow (ii)$: For all $\aleph_0$-saturated models $\mathcal{M}$ and $\mathcal{N}$ of $T$, all $n$-tuples $a \in \mathcal{M}^n$ and $b \in \mathcal{N}^n$, and all $c \in \mathcal{M}$, if $a$ and $b$ have the same quantifier free type, then by assumption we have $\mathrm{tp}(a) = \mathrm{tp}(b)$. Let $p(x, y) = \mathrm{tp}(c, a)$. Then $p(x, b)$ is consistent and is realized by some $d \in \mathcal{N}$ by $\aleph_0$-saturation. Hence, $(a, c)$ and $(b, d)$ have the same quantifier free type. □

### Model companion and model completion

Fix a signature $L$ for metric structures. An $L$-theory $T^*$ is called a *companion* of an $L$-theory $T$ if every model of $T$ can be embedded into a model of $T^*$ and vice versa. A *model companion* of a theory $T$ is a companion of $T$ that is model complete. A *model completion* for a theory $T$ is a model companion $T^*$ such that for every model $\mathcal{M}$ of $T$, the theory $T^* \cup \mathrm{Diag}(\mathcal{M})$ is complete.

**1.10.5 Proposition.** *Suppose $T^*$ is a model companion of $T$. Then $T^*$ is a model completion of $T$ if and only if $T^*$ admits quantifier elimination.*

*Proof.* Similar to the first order case; *e.g.*, see [10, Proposition 3.5.19]. □

## 1.11 Stability and independence

We fix a signature $L$ for metric structures and a complete $L$-theory $T$. Throughout this section, $\kappa$ is an infinite cardinal $> \mathrm{Card}(L)$ and $\lambda$ is an infinite cardinal.

**1.11.1 Definition.** • *A theory $T$ is $\lambda$-**stable with respect to the discrete metric** if for every $\mathcal{M} \models T$ and every $A \subseteq M$ of cardinality $\leq \lambda$, the type space $S_1(T_A)$ has cardinality $\leq \lambda$. We say that $T$ is*

***stable with respect to the discrete metric*** *if $T$ is $\lambda$-stable with respect to the discrete metric for some $\lambda$.*

- *A theory $T$ is $\lambda$-**stable** if for every $\mathcal{M} \models T$ and every $A \subseteq M$ of cardinality $\leq \lambda$, the type space $S_1(T_A)$ under metric topology has density character $\leq \lambda$. We say that $T$ is **stable** if $T$ is $\lambda$-stable for some $\lambda$.*

**1.11.2 Theorem** ([4, Theorem 14.6])**.** *A theory is stable if and only if it is stable with respect to the discrete metric.* □

**1.11.3 Definition.** *Let $\mathbb{M}$ be a $\kappa$-universal domain for $T$ and let $B, C \subseteq \mathbb{M}$ be small subsets. Let $p(X, B)$ be a partial type over $B$ where $X$ could be infinite. We say that $p(X, B)$ **divides over** $C$ if there is a $C$-indiscernible sequence $\{B_i\}_{i\in\mathbb{N}}$ in $\mathrm{tp}(B/C)$ such that $\bigcup_{i\in\mathbb{N}} p(X, B_i)$ is inconsistent with $T$.*

*Moreover, if $A, B, C \subseteq \mathbb{M}$ are small such that $\mathrm{tp}(A/B \cup C)$ does not divide over $C$, then we say that $A$ **is independent from** $B$ **over** $C$, denoted by $A \downarrow_C B$.*

*We say a type $\mathrm{tp}(A/C)$ **stationary** if for every $A'$ such that $\mathrm{tp}(A/C) = \mathrm{tp}(A'/C)$, $A \downarrow_C B$, and $A' \downarrow_C B$, we have $\mathrm{tp}(A/B \cup C) = \mathrm{tp}(A'/B \cup C)$.*

**1.11.4 Theorem** ([4, Theorem 14.12])**.** *Let $\mathbb{M}$ be a $\kappa$-universal domain for $T$. Let $A, B, \cdots \subseteq \mathbb{M}$ be small and let $M \preceq \mathbb{M}$ be small. If $T$ is stable, then the independence relation $\downarrow$ satisfies the following properties:*

*(i) Invariance under automorphisms of $\mathbb{M}$.*

*(ii) Symmetry: $A \downarrow_C B \iff B \downarrow_C A$.*

*(iii) Transitivity: $A \downarrow_C (B \cup D)$ if and only if $A \downarrow_C B$ and $A \downarrow_{B\cup C} D$.*

*(iv) Finite character: $A \downarrow_C B$ if and only if $a \downarrow_C B$ for all finite tuples $a$ from $A$.*

*(v) Extension: for all $A, B, C$, there is $A'$ such that $A' \downarrow_C B$ and $\mathrm{tp}(A/C) = \mathrm{tp}(A'/C)$.*

*(vi) Local character: for every finite tuple $a$, there is $B' \subseteq B$ of cardinality $\leq |T|$ such that $a \downarrow_{B'} B$.*

*(vii) Stationarity of types: if $\mathrm{tp}(A/M) = \mathrm{tp}(A'/M)$, $A \downarrow_M B$, and $A' \downarrow_M B$, then we have $\mathrm{tp}(A/B \cup M) = \mathrm{tp}(A'/B \cup M)$.* □

**1.11.5 Definition.** *A relation satisfying properties (i)–(vii) in the above theorem is called a **stable independence relation**.*

**1.11.6 Theorem** ([4, Theorem 14.14])**.** *Let $\mathbb{M}$ be a $\kappa$-universal domain for $T$. If $T$ is stable, then there is precisely one stable independence relation on $\mathbb{M}$. Moreover, if there is a stable independence relation on $\mathbb{M}$ on $\mathbb{M}$, then $T$ is stable.* □

**1.11.7 Theorem** (Stability Spectrum Theorem, [4, Theorem 14.7]). *Let $T$ be a stable theory and let $\mu(T)$ be the first cardinal such that $T$ is stable. Then there is a cardinal $\kappa = \kappa(T)$ such that $T$ is $\lambda$-stable if and only if $\lambda = \mu(T) + \lambda^{<\kappa}$.* □

**1.11.8 Definition** ([4, Definition 14.9]). *A theory is **superstable** if $\kappa$ (defined in the above theorem) is $\aleph_0$.*

**1.11.9 Proposition** ([4, Proposition 14.10]). *A theory $T$ is superstable if and only if it is $\lambda$-stable for all $\lambda \geq 2^{|T|}$.* □

**1.11.10 Proposition** ([4, Corollary 14.17]). *Suppose a theory $T$ is stable. Then $T$ is $\lambda$-stable with respect to the discrete metric for all $\lambda$ such that $\lambda = \lambda^{|T|}$.* □

# Chapter 2

# Background from analysis

In this chapter, we present the background from analysis necessary for later chapters. In Section 2.1, we discuss basic probability theory. Section 2.2 plays an essential role in the following chapters. In this section, we introduce measure algebras and the property of one measure algebra being atomless over another one. We also present Maharam's Theorem about the decomposability of measure spaces. In Section 2.3, we discuss the symmetric rearrangement of a Borel set and the symmetric-decreasing rearrangement of a Borel measurable function. In Section 2.4, we give background on Wasserstein distances in optimal transport.

## 2.1 Basics of probability theory

In this section, we present some basic probability theory results. The book [12] is a good source for most of them.

A *probability space* is a triple $(\Omega, \mathcal{F}, P)$, where $\Omega$ is a set of "outcomes", $\mathcal{F}$ is a set of "events", and $P\colon \mathcal{F} \to [0,1]$ is a function that assigns probabilities to events. Moreover, $\mathcal{F}$ is a $\sigma$-algebra of subsets of $\Omega$ and $P$ is a *probability measure* on $\mathcal{F}$ which means that $P$ is $\sigma$-additive and $P(\Omega) = 1$. If $\Omega$ is a countable set, then the probability space $(\Omega, \mathcal{F}, P)$ is called a *discrete probability space*. For any $F \in \mathcal{F}$ with positive measure, if for every $G \in \mathcal{F}$, either $P(F \cap G) = 0$ or $P(F \cap G) = P(F)$, then $F$ is called an *atom* in $\mathcal{F}$. If for every $F \in \mathcal{F}$ with positive measure there is an atom $F_0 \in \mathcal{F}$ such that $F_0 \subseteq F$, then such a $\sigma$-algebra $\mathcal{F}$ is called *atomic*. Note that if $(\Omega, \mathcal{F}, P)$ is a discrete probability space, then $\mathcal{F}$ is atomic.

A real valued function $X$ on $\Omega$ is called a *random variable* if for every Borel $B \subseteq \mathbb{R}$, we have $X^{-1}(B) \in \mathcal{F}$. A random variable $X$ induces a natural probability measure on $\mathbb{R}$, called its *distribution* and denoted by $\mathrm{dist}(X)$; it is defined by setting $\mathrm{dist}(X)(B) = P(X^{-1}(B)) = P(X \in B)$ for all Borel $B \subseteq \mathbb{R}$. Let $A$ be a set of random variables : $\Omega \to \mathbb{R}$. Let $\sigma(A)$ denote the smallest $\sigma$-algebra $\mathcal{A} \subseteq \mathcal{F}$ such that all random variables in $A$ are $\mathcal{A}$-measurable. When $A$ only has one element $X$, we also call $\sigma(X)$ the *$\sigma$-algebra (of measurable sets) generated by the random variable $X$*. Let $\{X_n\}_{n\in\mathbb{N}}$ be a sequence of random variables on $\Omega$. We say

that $X_n$ *converges almost surely* if

$$P\big(\{\omega \mid \lim_{n\to\infty} X_n(\omega) \text{ exists } \}\big) = 1.$$

Two events $A, B \in \mathcal{F}$ are *independent* if $P(A \cap B) = P(A)P(B)$. Two random variables $X, Y \colon \Omega \to \mathbb{R}$ are *independent* if for all Borel $C, D \subseteq \mathbb{R}$, the events $X^{-1}(C)$ and $Y^{-1}(D)$ are independent. Two $\sigma$-algebras $\mathcal{A} \subseteq \mathcal{F}$ and $\mathcal{B} \subseteq \mathcal{F}$ are *independent* if for all $A \in \mathcal{A}$ and all $B \in \mathcal{B}$, the events $A$ and $B$ are independent. Note that random variables $X$ and $Y$ are independent if and only if the $\sigma$-algebras $\sigma(X)$ and $\sigma(Y)$ are.

**2.1.1 Definition.** *Consider a probability space $(\Omega, \mathcal{F}_0, P)$, a $\sigma$-subalgebra $\mathcal{F}$ of $\mathcal{F}_0$, and a random variable $X$ with $\int_\Omega |X| dP < \infty$. The* ***conditional expectation of $X$ given $\mathcal{F}$****, denoted by $\mathbb{E}(X|\mathcal{F})$, is any random variable $Y$ satisfying the following two conditions:*

- *The random variable $Y$ is $\mathcal{F}$-measurable.*
- *For every $A \in \mathcal{F}$, we have $\int_A X dP = \int_A Y dP$.*

*If $X$ is a characteristic function, say $X = \mathbf{1}_A$ where $A \in \mathcal{F}_0$, then $\mathbb{E}(\mathbf{1}_A|\mathcal{F})$ is called the* ***conditional probability of $A$ given $\mathcal{F}$****, denoted also by $\mathbb{P}(A|\mathcal{F})$.*

Any $Y$ satisfies the above two conditions is said to be a *version* of $\mathbb{E}(X|\mathcal{F})$.

Given probability measures $\mu$ and $\nu$ on $(\Omega, \mathcal{F})$, we say that $\nu$ *is absolutely continuous with respect to* $\mu$, denoted by $\nu \ll \mu$, if for all $A \in \mathcal{F}$, we have $\nu(A) = 0$ whenever $\mu(A) = 0$.

The following theorem is well-known in probability theory.

**2.1.2 Theorem** (Radon-Nikodym Theorem)**.** *Let $\mu$ and $\nu$ be probability measures on $(\Omega, \mathcal{F})$. If $\nu \ll \mu$, then there is a $\mathcal{F}$-measurable function $f$ such that for every $A \in \mathcal{F}$, $\int_A f d\mu = \nu(A)$. Such $f$ is called a* Radon-Nikodym derivative of $\nu$ with respect to $\mu$*, and is denoted by $\frac{d\nu}{d\mu}$.* □

**2.1.3 Remark.** *Given that $\int_\Omega |X| dP < \infty$, the conditional expectation of $X$ given $\mathcal{F}$ exists and is unique up to $=_{a.s.}$. The existence follows from the Radon-Nikodym Theorem.*

Conditional expectations have the following properties:

**2.1.4 Proposition.** *Consider a probability space $(\Omega, \mathcal{F}_0, P)$ and a $\sigma$-subalgebra $\mathcal{F}$ of $\mathcal{F}_0$. Let $X, Y, \cdots$ be random variables $: \Omega \to \mathbb{R}$ with $\int_\Omega |X| dP < \infty, \int_\Omega |Y| dP < \infty, \cdots$.*

1. *(Linearity) $\mathbb{E}(aX + bY|\mathcal{F}) = a\mathbb{E}(X|\mathcal{F}) + b\mathbb{E}(Y|\mathcal{F})$ for all $a, b \in \mathbb{R}$.*
2. *(Monotonicity) If $X \leq Y$, then $\mathbb{E}(X|\mathcal{F}) \leq \mathbb{E}(Y|\mathcal{F})$.*

3. *(Monotone convergence theorem) If $X_n \geq 0$ for every $n \in \mathbb{N}$, the sequence $\{X_n\}_{n\in\mathbb{N}}$ is increasing, and $X_n \to X$ a.s. with $\int_\Omega X dP < \infty$, then $\{\mathbb{E}(X_n|\mathcal{F})\}_{n\in\mathbb{N}}$ is increasing and $\mathbb{E}(X_n|\mathcal{F}) \to \mathbb{E}(X|\mathcal{F})$ a.s. as $n \to \infty$.*

4. *If $\varphi\colon \mathbb{R} \to \mathbb{R}$ is a convex function and $\int_\Omega |X|dP, \int_\Omega |\varphi(X)|dP < \infty$, then $\varphi\big(\mathbb{E}(X|\mathcal{F})\big) \leq \mathbb{E}\big(\varphi(X)|\mathcal{F}\big)$.*

5. *$\int_\Omega \mathbb{E}(X|\mathcal{F})dP = \int_\Omega X dP$.*

6. *(Transitivity) If $\mathcal{F}_2 \subseteq \mathcal{F}_1 \subseteq \mathcal{F}_0$, then*

$$\mathbb{E}\big(\mathbb{E}(X|\mathcal{F}_2)|\mathcal{F}_1\big) = \mathbb{E}\big(\mathbb{E}(X|\mathcal{F}_1)|\mathcal{F}_2\big) = \mathbb{E}(X|\mathcal{F}_2).$$

*Proof.* See [12, pages 224–227]. □

Next, we give the definition for transformations between two probability spaces.

**2.1.5 Definition** ([33, Definition 1.1])**.** *Let $(\Omega_1, \mathcal{F}_1, P_1)$ and $(\Omega_2, \mathcal{F}_2, P_2)$ be probability spaces.*

1. *A transformation $T\colon \Omega_1 \to \Omega_2$ is **measurable** if $T^{-1}(\mathcal{F}_2) \subseteq \mathcal{F}_1$; i.e., if $F \in \mathcal{F}_2$ then $T^{-1}(F) \in \mathcal{F}_1$.*

2. *A transformation $T\colon \Omega_1 \to \Omega_2$ is **measure-preserving** if $T$ is measurable and $P_1\big(T^{-1}(F)\big) = P_2(F)$ for all $F \in \mathcal{F}_2$.*

3. *A transformation $T\colon \Omega_1 \to \Omega_2$ is **invertible measure-preserving** if $T$ is measure-preserving, bijective, and its inverse $T^{-1}$ is also measure-preserving.*

## 2.2 Background for measure algebras

In this section, we give the background for measure algebras, adapted from the books [12] and [14].

### Dynkin's $\pi$-$\lambda$ theorem

In this subsection, we introduce Dynkin's well-known $\pi$-$\lambda$ Theorem and then use it to get some results on $\sigma$-algebras.

Let $(\Omega, \mathcal{F}, \mu)$ be a probability space. Let $\mathcal{P}$ and $\mathcal{L}$ be subsets of $\mathcal{F}$. We say that $\mathcal{P}$ is a *$\pi$-system* if it is closed under intersection; *i.e.*, if $A, B \in \mathcal{P}$ then $A \cap B \in \mathcal{P}$. We say that $\mathcal{L}$ is a *$\lambda$-system* if: $(i)$ $\Omega \in \mathcal{L}$; $(ii)$ If $A, B \in \mathcal{L}$ and $A \subseteq B$ then $B \setminus A \in \mathcal{L}$; $(iii)$ If for all $n \in \mathbb{N}$ one has $A_n \in \mathcal{L}$, and $\{A_n\}_{n\in\mathbb{N}}$ is an increasing sequence so that $\bigcup_{n\in\mathbb{N}} A_n = A$, then $A \in \mathcal{L}$.

**2.2.1 Theorem** ($\pi$-$\lambda$ Theorem, [12, page 25]). *If $\mathcal{P}$ is a $\pi$-system and $\mathcal{L}$ is a $\lambda$-system that contains $\mathcal{P}$, then $\sigma(\mathcal{P}) \subseteq \mathcal{L}$.* □

Let $\mathscr{B}$ be a $\sigma$-algebra. For every $A \subseteq \mathscr{B}$, *the $\sigma$-subalgebra $\sigma$-generated by $A$* in $\mathscr{B}$, denoted by $\sigma(A)$, is the smallest $\sigma$-subalgebra of $\mathscr{B}$ including $A$. For all $\sigma$-subalgebras $\mathscr{A}_1$ and $\mathscr{A}_2$ of $\mathscr{B}$, the $\sigma$-subalgebra $\sigma$-generated by $\mathscr{A}_1$ and $\mathscr{A}_2$ is denoted by $\mathscr{A}_1 \vee \mathscr{A}_2$.

Let $(\Omega, \mathcal{F}, \mu)$ be a probability space and let $\mathscr{A} \subseteq \mathcal{F}$ be an atomic $\sigma$-subalgebra. A set $\{a_i \mid i \in I\}$ is called a *maximal family of pairwise inequivalent atoms of $\mathscr{A}$* if $(i)$ $a_i$ is an atom of $\mathscr{A}$ for each $i \in I$; $(ii)$ $\mu(a_i \cap a_j) = 0$ whenever $i \neq j \in I$; $(iii)$ $\bigcup_{i\in I} a_i = \Omega$.

**2.2.2 Fact.** *Let $(\Omega, \mathcal{F}, \mu)$ be a probability space.*

*(i) Every maximal family of pairwise inequivalent atoms is countable.*

*(ii) Every atomic $\sigma$-subalgebra of $\mathcal{F}$ has a maximal family of pairwise inequivalent atoms.*

*(iii) Let $\mathscr{A} \subseteq \mathcal{F}$ be an atomic $\sigma$-subalgebra. If $\{a_i \mid i \in I\}$ is a maximal family of pairwise inequivalent atoms of $\mathscr{A}$, then for every $a \in \mathscr{A}$ there is $b \in \sigma(\{a_i \mid i \in I\})$ such that $\mu(a \triangle b) = 0$.*

*Proof.* $(i)$: This follows from the defining properties of maximal families of pairwise inequivalent atoms because $\mu$ is a probability measure.

$(ii)$: Fix an atomic $\sigma$-subalgebra $\mathscr{A}$ of $\mathcal{F}$. For all atoms $a, b$ of $\mathscr{A}$, we say $a \sim b$ if $\mu(a \cap b) > 0$. Since $a, b$ are atoms, it follows that $\mu(a \triangle b) = 0$. Note that $\sim$ defines an equivalence relation on the set of atoms of $\mathscr{A}$. Let $\{\Gamma_i \mid i \in I\}$ denote the set of all $\sim$ equivalence classes. Since $\mu$ is a probability measure, we know that $I$ is a countable set. For each $i \in I$, take $s_i \in \Gamma_i$. Then $s_i$ is an atom for each $i \in I$. Whenever $i \neq j \in I$, since $s_i$ and $s_j$ belong to different equivalence classes, we have $\mu(s_i \cap s_j) = 0$. If $\mu(\bigcup_{i\in I} s_i) \neq 1$, then since $\mathscr{A}$ is atomic, there is an atom $a \subseteq \Omega \setminus (\bigcup_{i\in I} s_i)$ in $\mathscr{A}$. Then $a \in \Gamma_j$ for some $j \in I$, and thus $\mu(a \cap s_j) > 0$, which contradicts the fact that $a \subseteq \Omega \setminus (\bigcup_{i\in I} s_i)$. Hence, $\mu(\bigcup_{i\in I} s_i) = 1$. Take $i_0 \in I$. Let $a_{i_0} = s_{i_0} \cup \big(\Omega \setminus (\bigcup_{i\in I} s_i)\big)$ and $a_i = s_i$ for all $i \neq i_0$. Then $\{a_i \mid i \in I\}$ is a maximum family of pairwise inequivalent atoms of $\mathscr{A}$.

$(iii)$: Suppose $\{a_i \mid i \in I\}$ is a maximal family of pairwise inequivalent atoms of $\mathscr{A}$. Then for every $a \in \mathscr{A}$, let $J$ denote the set $\{j \mid j \in I, \mu(a_j \cap a) > 0\}$. Since $I$ is a countable set, $J$ is also countable. Let $b$ be $\bigcup_{j\in J} a_j$. Then $b \in \sigma(\{a_i \mid i \in I\})$. Since $a_i$'s are atoms, for every $j \in J$, we have that $\mu(a_j \triangle (a_j \cap a)) = 0$. Then it follows that $\mu(a \triangle b) = 0$. □

**2.2.3 Proposition.** *Let $(\Omega, \mathscr{C}, \mu)$ be a probability space and let $\mathscr{A}, \mathscr{B} \subseteq \mathscr{C}$ be $\sigma$-algebras. Assume that $\mathscr{A}$ contains the set of all sets of measure zero and $\mathscr{B}$ is atomic. Let $\{b_i \mid i \in I\}$ be a maximal family of pairwise*

*inequivalent atoms of $\mathscr{B}$, where $I$ is a countable set. Then, we have*

$$\mathscr{A} \vee \mathscr{B} = \{\bigcup_{i\in I} a_i \cap b_i \mid a_i \in \mathscr{A} \text{ for every } i \in I\}.$$

*Proof.* Let $\mathcal{L}$ denote the set $\{\bigcup_{i\in I} a_i \cap b_i \mid a_i \in \mathscr{A}$ for every $i \in I\}$. For the set $\mathcal{L}$, we have

(i) $\Omega \in \mathcal{L}$.

(ii) Let $A_1 = \bigcup_{i\in I} a_i \cap b_i$ and $A_2 = \bigcup_{i\in I} a_i' \cap b_i$ be in $\mathcal{L}$. Then

$$A_1 \setminus A_2 = \bigcup_{i\in I} (a_i \setminus a_i') \cap b_i \in \mathcal{L}.$$

(iii) Let $\{A_j = \bigcup_{i\in I} a_{ij} \cap b_i\}_{j\in\mathbb{N}}$ be a sequence in $\mathcal{L}$. Then

$$\bigcup_{j=1}^{\infty} A_j = \bigcup_{j=1}^{\infty} \bigcup_{i\in I} a_{ij} \cap b_i = \bigcup_{i\in I} (\bigcup_{j=1}^{\infty} a_{ij}) \cap b_i \in \mathcal{L}.$$

Therefore $\mathcal{L}$ is a $\lambda$-system. Let $\mathcal{P}$ be $\{c_1 \cap c_2 \mid c_1, c_2 \in \{b_i \mid i \in I\} \cup \mathscr{A}\}$. Since whenever $i \neq i' \in I$ we have $\mu(b_i \cap b_{i'}) = 0$, the set of all sets of measure zero is contained in $\mathscr{A}$, and $\mathscr{A}$ is closed under intersections, we have that $\mathcal{P}$ is a $\pi$-system. By Theorem 2.2.1 ($\pi$-$\lambda$ Theorem), we have $\sigma(\mathcal{P}) \subseteq \mathcal{L}$. Because $\mathscr{A}$ contains the set of all sets of measure zero and Fact 2.2.2(*iii*), we have that $\mathscr{A} \vee \mathscr{B} = \sigma(\{b_i \mid i \in I\} \cup \mathscr{A})$. Hence,

$$\mathscr{A} \vee \mathscr{B} = \sigma(\{b_i \mid i \in I\} \cup \mathscr{A}) \subseteq \sigma(\mathcal{P}) \subseteq \mathcal{L} \subseteq \sigma(\{b_i \mid i \in I\} \cup \mathscr{A}).$$

Therefore $\mathscr{A} \vee \mathscr{B} = \mathcal{L}$. □

## Measure algebras

In this subsection, we present the necessary background for measure algebras. Some properties of the relation "atomless over" are given, which play an important role in the next chapters. Finally, we prove Theorem 2.2.21, which tells us how to extend a partial measure-preserving isomorphism between probability algebras to a global one.

### Basics of measure algebras

A boolean algebra $B$ is *$\sigma$-complete* if the supremum of every countable subset of $B$ exists in $B$. A boolean algebra $B$ is *complete* if the supremum of every subset of $B$ exists in $B$. Let $A$ be a subset of a boolean

algebra (resp. $\sigma$-complete boolean algebra) $B$. Then the *boolean subalgebra generated by $A$ (resp. $\sigma$-complete boolean subalgebra $\sigma$-generated by $A$)* in $B$, denoted by $\langle A\rangle$ (resp. $\sigma(A)$), is the smallest boolean subalgebra (resp. $\sigma$-complete boolean subalgebra) of $B$ including $A$. For $\sigma$-complete boolean subalgebras $A_1$ and $A_2$ of a $\sigma$-complete algebra $B$, the $\sigma$-complete boolean subalgebra $\sigma$-generated by $A_1$ and $A_2$ is denoted by $A_1 \vee A_2$. Note that a $\sigma$-algebra is a $\sigma$-complete boolean algebra.

A *measure space* is a triple $(X, A, \mu)$ where $X$ is a set, $A$ is a $\sigma$-algebra of subsets of $X$, and $\mu\colon A \to [0,\infty)$ is a countably additive finite-valued measure.

A *measured algebra* is a pair $(A, \mu)$ where $A$ is a $\sigma$-complete boolean algebra and $\mu\colon A \to [0,\infty)$ is a finite-valued function such that $\mu(a) = 0$ iff $a = \mathbf{0}$, and $\mu$ is countably additive, which means that if $\{a_n\}_{n\in\mathbb{N}}$ is a countable subset of $A$ and $a_m \cap a_n = \mathbf{0}$ whenever $m \neq n$, then

$$\mu(\sum_{n\in\mathbb{N}} a_n) = \sum_{n\in\mathbb{N}} \mu(a_n).$$

A boolean algebra $A$ is said to be a *measure algebra* if there is a finite-valued $\mu$ for which $(A, \mu)$ is a measured algebra. A measured algebra $(A, \mu)$ is called a *probability algebra* if $\mu(\mathbf{1}) = 1$. Let $(A, \mu)$ be a measured algebra. For all $a, b \in A$, define

$$d(a, b) = \mu(a \triangle b),$$

where $\triangle$ is the symmetric difference of those two sets. In [15, Lemma 323F], it is shown that $A$ is complete under this metric.

Let $A$ be a measure algebra. Let $A^+$ denote the set of nonzero elements in $A$. For every $a \in A^+$, if for every $b \in A$, either $a \cap b = \mathbf{0}$ or $a \cap b = a$, then $a$ is called an *atom* in $A$. If for every $a \in A^+$ there is an atom $a_0 \in A^+$ such that $a_0 \subseteq a$, then such a measure algebra $A$ is called *atomic*.

Let $(X, A, \mu)$ be a measure space. For all $a, b \in A$, we write $a \equiv_\mu b$ if $\mu(a \triangle b) = 0$. Note that $\equiv_\mu$ defines an equivalence relation. We denote the equivalence class of $a$ under $\equiv_\mu$ by $[a]_\mu$. Let $\widehat{A}$ denote the set $\{[a]_\mu \mid a \in A\}$. Then it is routine to verify that the operations of complement, union and intersection on $A$ induce operations on $\widehat{A}$, which make $\widehat{A}$ a $\sigma$-complete boolean algebra. Moreover, $\mu$ induces a countably additive, strictly positive measure on $\widehat{A}$. We call $(\widehat{A}, \mu)$ the *measured algebra associated to* $(X, A, \mu)$ and we call $\widehat{A}$ the *measure algebra associated to* $(X, A, \mu)$. If the measure $\mu$ on $A$ is a probability measure, then we call $(\widehat{A}, \mu)$ the *probability algebra associated to* $(X, A, \mu)$.

**2.2.4 Definition** ([33, Definition 2.1 and 2.2])**.** *Let $(X_1, A_1, \mu_1)$ and $(X_2, A_2, \mu_2)$ be probability spaces.*

(i) *Probability spaces* $(X_1, A_1, \mu_1)$ *and* $(X_2, A_2, \mu_2)$ *are said to be* ***isomorphic*** *if there are* $M_1 \in A_1, M_2 \in A_2$ *with* $\mu_1(M_1) = \mu_2(M_2) = 1$*, and an invertible measure-preserving transformation* $\varphi\colon M_1 \to M_2$*. If the probability space* $(X_1, A_1, \mu_1)$ *is the same as* $(X_2, A_2, \mu_2)$*, then such an isomorphism is called an* ***automorphism****.*

(ii) *Let* $(\widehat{A}_1, \mu_1)$ *and* $(\widehat{A}_2, \mu_2)$ *be the measured algebras associated to* $(X_1, A_1, \mu_1)$ *and* $(X_2, A_2, \mu_2)$ *respectively. The measured algebras are* ***isomorphic*** *if there is a bijection* $\Phi\colon \widehat{A}_1 \to \widehat{A}_2$ *which preserves complements, countable unions and intersections and satisfies* $\mu_1(\Phi^{-1}(b)) = \mu_2(b)$ *for all* $b \in \widehat{A}_2$*. If the probability space* $(X_1, A_1, \mu_1)$ *is the same as* $(X_2, A_2, \mu_2)$*, then such an isomorphism is called an* ***automorphism****. We say that the probability spaces are* ***conjugate*** *if their associated probability algebras are isomorphic.*

Since point maps induce set maps, we have that if two probability spaces are isomorphic, then they are conjugate. The other direction is not true in general. For the case when set maps are always induced by point maps for probability spaces, there is the following result.

**2.2.5 Theorem** ([33, Theorem 2.2])**.** *Let* $X_1$ *and* $X_2$ *be complete separable metric spaces, let* $B(X_1)$ *and* $B(X_2)$ *be their* $\sigma$*-algebras of Borel subsets, and let* $\mu_1$ *and* $\mu_2$ *be probability measures on* $B(X_1)$ *and* $B(X_2)$ *respectively. Let* $\Phi\colon \widehat{B}(X_2) \to \widehat{B}(X_1)$ *be an isomorphism of probability algebras. Then there are* $M_1 \in B(X_1), M_2 \in B(X_2)$ *with* $\mu_1(M_1) = \mu_2(M_2) = 1$*, and an invertible measure-preserving transformation* $\varphi\colon M_1 \to M_2$ *such that* $\Phi([b]_{\mu_2}) = [\varphi^{-1}(b \cap M_2)]_{\mu_1}$ *for every* $b \in B(X_2)$*. If* $\phi$ *is any other isomorphism from* $(X_1, B(X_1), \mu_1)$ *to* $(X_2, B(X_2), \mu_2)$ *that induces* $\Phi$*, then*

$$\mu_1(\{x \in X_1 \mid \varphi(x) \neq \phi(x)\}) = 0.$$

□

**2.2.6 Theorem** ([14, Theorem 2.6])**.** *Let* $(A, \mu)$ *be a measured algebra. Then there is a measure space* $(X, B, \nu)$ *such that* $(A, \mu) \cong (\widehat{B}, \nu)$ *as measured algebras.* □

**2.2.7 Proposition.** *Let* $A, B$ *be* $\sigma$*-complete subalgebras in a probability algebra* $(C, \mu)$*. Suppose* $B$ *is atomic. Let* $\{b_i \mid i \in I\}$ *be an enumeration of the atoms in* $B$*. Then*

$$A \vee B = \{\bigcup_{i \in I} a_i \cap b_i \mid a_i \in A \text{ for every } i \in I\}.$$

*Proof.* By Theorem 2.2.6, there is a probability space $(X, \mathscr{C}, \nu)$ such that $(C, \mu) \cong (\widehat{\mathscr{C}}, \nu)$ as probability

algebras. Without loss of generality, we identify $(C,\mu)$ with $(\widehat{\mathscr{C}},\nu)$. Thus, $(C,\mu)$ is the probability algebra associated to $(X,\mathscr{C},\nu)$. As explained on page 30, there is a mapping $\pi$ from the $\sigma$-algebra $\mathscr{C}$ to the $\sigma$-complete boolean algebra $\pi(\mathscr{C}) := \widehat{\mathscr{C}} = C$. Let $\mathscr{A}$ denote $\pi^{-1}(A)$. Then because $A$ is a $\sigma$-complete boolean algebra and $\mathscr{C}$ is a $\sigma$-algebra, it follows that $\mathscr{A}$ is a $\sigma$-subalgebra of $\mathscr{C}$ and $\mathscr{A}$ contains the set of all sets of measure zero. Further, we know that $(A,\mu)$ is the probability algebra associated to $(\Omega,\mathscr{A},\nu)$. Let $\mathscr{B}$ denote $\pi^{-1}(B)$. Then $\mathscr{B}$ is a $\sigma$-subalgebra of $\mathscr{C}$. Since $B$ is atomic, we have that $\mathscr{B}$ is also atomic. Since $\mu$ is finite-valued, $B$ has at most countably many atoms, and thus $I$ is a countable set. For each $i \in I$, we have that $\pi^{-1}(b_i)$ is an equivalence class of measurable sets in $\mathscr{B}$. Take $C_i$ in the class $\pi^{-1}(b_i)$ for each $i \in I$. Then for each $i \in I$ we have $C_i \subseteq \mathscr{B}$ is an atom, $\mu(C_i \triangle C_j) = 0$ whenever $i \neq j \in I$, and $\mu(\bigcup_{i\in I} C_i) = 1$. Take $i_0 \in I$. Let $B_{i_0} = C_{i_0} \cup (\Omega \setminus (\bigcup_{i\in I} C_i))$ and $B_j = C_j$ for all $j \neq i_0 \in I$. Then $\{B_i \mid i \in I\}$ is a maximal family of pairwise inequivalent atoms of $\mathscr{B}$. The rest follows from Proposition 2.2.3. □

Let $A$ be a boolean algebra and let $A^+$ denote the set of nonzero elements in $A$. For every $a \in A^+$, *the relative algebra of $A$ on $a$*, denoted by $A \restriction a$, is the set $\{a \cap b \mid b \in A\}$. Note that $A \restriction a$ is also a boolean algebra, where $a$ becomes the $\mathbf{1}$. For a subset $X$ of a complete boolean algebra $A$, we say $A$ is *completely generated by $X$* if the smallest complete subalgebra of $A$ containing $X$ is $A$.

**2.2.8 Definition.** *Let $A$ be a complete boolean algebra. Then*

*(i)* $\tau(A) := \inf\{|X| : X \subseteq A$ *and $A$ is completely generated by* $X\}$.

*(ii) ($\tau$-homogeneous) We say that $A$ is* $\tau$***-homogeneous*** *if* $\tau(A \restriction a) = \tau(A)$ *for all* $a \in A^+$.

*(iii) (homogeneous) We say that $A$ is* ***homogeneous*** *if* $A \cong A \restriction a$ *for all* $a \in A^+$.

*(iv) Suppose $B$ is a complete subalgebra of $A$. For every $a \in A$, we define*

$$\tau_B(a) := \inf\{|X| : X \subseteq A \text{ and } A \restriction a \text{ is completely generated by } \{a \cap b \mid b \in X \cup B\}\}.$$

**2.2.9 Remark.** *By [14, page 889, 2.2(h)], any $\sigma$-complete subalgebra of a measure algebra is a complete subalgebra. Therefore, when $A$ is a measure algebra, we have*

$$\tau(A) = \inf\{|X| : X \subseteq A \text{ and } A \text{ is } \sigma\text{-generated by } X\},$$

*and*

$$\tau_B(a) = \inf\{|X| : X \subseteq A \text{ and } A \restriction a \text{ is } \sigma\text{-generated by } \{a \cap b \mid b \in X \cup B\}\}.$$

**2.2.10 Theorem** ([14, Theorem 3.5])**.** *Let $(A, \mu)$ and $(B, \nu)$ be $\tau$-homogeneous measured algebras such that $\tau(A) = \tau(B)$ and $\mu(\mathbf{1}) = \nu(\mathbf{1})$. Then there is a measure-preserving isomorphism between $A$ and $B$.* □

**2.2.11 Proposition.** *Let $A$ be a measure algebra. Then $A$ is $\tau$-homogeneous if and only if $A$ is homogeneous as a boolean algebra.*

*Proof.* [14, Corollary 3.6] shows that if $A$ is $\tau$-homogeneous then it is homogeneous. The other direction is trivial. □

**2.2.12 Theorem** ([14, Theorem 3.13])**.** *Let $(A, \mu)$ and $(B, \nu)$ be measured algebras and $C$ a $\sigma$-complete subalgebra of $A$. Let $g\colon C \to B$ be a measure-preserving homomorphism. Set $\kappa = \min\{\tau(B \restriction b) \mid b \in B^+\}$. Suppose $\tau(C) < \kappa$. Then we have:*

*(i) If $\tau(A) < \kappa$, then there is a measure-preserving homomorphism $f\colon A \to B$ extending $g$.*

*(ii) If $A$ and $B$ are both homogeneous and $\tau(A) = \tau(B) = \kappa$, then $f$ can be taken to be an isomorphism.*

□

**2.2.13 Corollary** ([14, Corollary 3.19])**.** *Let $A$ be a homogeneous measured algebra, $C \subseteq A$ be a $\sigma$-complete subalgebra with $\tau(C) < \tau(A)$, and $g\colon C \to A$ be a $\sigma$-complete monomorphism (i.e., it preserves countable unions of elements in $C$). Then $g$ extends to an automorphism of $A$.* □

**Atomlessness**

**2.2.14 Definition.** *(i) A measure space $(X, A, \mu)$ is **atomless** if for every $a \in A$ of positive measure, there exists $b \in A$ such that $b \subseteq a$ and $0 < \mu(b) < \mu(a)$.*

*(ii) Let $(X, A, \mu)$ be a measure space and let $B$ be a $\sigma$-subalgebra of $A$. We say that $A$ **is atomless over** $B$, if for every $a \in A$ of positive measure, there exists $b \in A$ such that for all $c \in B$, we have $a \cap b \neq a \cap c$.*

*(iii) Let $(X, A, \mu)$ be a measure space and let $B$ be a $\sigma$-subalgebra of $A$. Given an infinite cardinal $\kappa$, we say that $A$ **is $\kappa$-atomless over** $B$, if for every $\sigma$-subalgebra $B'$, which is $\sigma$-generated by $B \cup S$, where $S$ is a set of cardinality $< \kappa$ in $A$, we have that $A$ is atomless over $B'$. When $B$ is trivial and $A$ is $\kappa$-atomless over $B$, we say simply that $(X, A, \mu)$ **is $\kappa$-atomless**.*

Analogously, we define these notions of atomlessness for measure algebras.

**2.2.15 Definition.** *(i) A measure algebra $A$ is **atomless** if for every $a \in A^+$, there exists $b \in A^+$ such that $b \cap a = b$ and $b \neq a$.*

*(ii)* *Let $A$ be a measure algebra and let $B$ be a $\sigma$-complete subalgebra of $A$. We say that $A$* ***is atomless over*** *$B$, if for every $a \in A^+$, there exists $a' \in A$ such that for all $b \in B$, we have $a \cap a' \neq a \cap b$.*

*(iii)* *Let $A$ be a measure algebra and let $B$ be a $\sigma$-complete subalgebra of $A$. Given an infinite cardinal $\kappa$, we say that $A$* ***is $\kappa$-atomless over*** *$B$, if for every $\sigma$-complete subalgebra $B'$, which is $\sigma$-generated by $B \cup S$, where $S$ is a set of cardinality $< \kappa$ in $A$, we have that $A$ is atomless over $B'$. When $B$ is trivial and $A$ is $\kappa$-atomless over $B$, we say simply that $A$* ***is $\kappa$-atomless****.*

**2.2.16 Lemma** (Maharam's Lemma)**.** *Let $(X, A, \mu) \supseteq (X, B, \mu)$ be measure spaces and let $\widehat{A}$ and $\widehat{B}$ be the measure algebras associated to $(X, A, \mu)$ and $(X, B, \mu)$ respectively. Then the following are equivalent:*

*(i)* *The measure space $(X, A, \mu)$ is atomless over $(X, B, \mu)$.*

*(ii)* *For every $a \in A$ of positive measure and for every $B$-measurable function $f \colon X \to \mathbb{R}$ such that $0 \leq f \leq \mathbb{E}(a \mid B)$, there is a set $b \in A$ such that $b \subseteq a$ and $\mathbb{E}(b \mid B) = f$.*

*(iii)* *For every $a \in A$ of positive measure, there is a set $b \in A$ such that $b \subseteq a$ and*

$$\mathbb{E}(b \mid B) = \frac{\mathbb{E}(a \mid B)}{2}.$$

*(iv)* *For every $a \in A$ of positive measure and for all $\epsilon > 0$, there is a set $b \in A$ such that $b \subseteq a$ and*

$$\left\|\mathbb{E}(b \mid B) - \frac{\mathbb{E}(a \mid B)}{2}\right\|_1 \leq \epsilon.$$

*(v)* *The measure algebra $\widehat{A}$ is atomless over $\widehat{B}$.*

*Proof.* $(i) \Rightarrow (ii)$: This is [8, Theorem 1.3].

$(ii) \Rightarrow (iii)$: This is trivial.

$(iii) \Rightarrow (iv)$: This is trivial.

$(iv) \Rightarrow (i)$: We need only prove that for every $a \in A$ with $\mu(a) \neq 0$, there is a set $b \in A$ such that $a \cap b \neq a \cap c$ for all $c \in B$. Fix $a \in A$ with $\mu(a) > 0$. Choose $\epsilon$ so that $0 < \epsilon < \frac{\mu(a)}{12}$. By $(iv)$, there is a set $b \in A$ such that $b \subseteq a$ and

$$\left\|\mathbb{E}(b \mid B) - \frac{\mathbb{E}(a \mid B)}{2}\right\|_1 \leq \epsilon.$$

If for some $c \in B$ we have $a \cap b = a \cap c$, then $b = a \cap b = a \cap c$, whereby $b \cap c = (a \cap c) \cap c = a \cap c = b$. Hence

$$\int_c \mathbb{E}(a \mid B) d\mu = \int_c \chi_a d\mu = \mu(a \cap c) = \mu(b).$$

Also

$$\int_c \mathbb{E}(b \mid B)d\mu = \mu(b \cap c) = \mu(b).$$

Then we have

$$\int_c (\mathbb{E}(b \mid B) - \frac{\mathbb{E}(a \mid B)}{2})d\mu = \frac{\mu(b)}{2}.$$

Thus

$$\frac{\mu(b)}{2} \leq \big\|\mathbb{E}(b \mid B) - \frac{\mathbb{E}(a \mid B)}{2}\big\|_1 \leq \epsilon.$$

Then

$$\frac{\mu(a)}{2} = \big\|\frac{\mathbb{E}(a \mid B)}{2}\big\|_1 \leq \big\|\mathbb{E}(b \mid B) - \frac{\mathbb{E}(a \mid B)}{2}\big\|_1 + \|\mathbb{E}(b \mid B)\|_1 \leq \epsilon + \mu(b).$$

Hence $\frac{\mu(a)}{2} \leq 3\epsilon$, contradicting the fact that $\epsilon < \frac{\mu(a)}{12}$. Therefore $a \cap b = b \neq a \cap c$ for all $c \in B$.

$(v) \Rightarrow (i)$: Since $\widehat{A}$ is atomless over $\widehat{B}$, for every $a \in A$ with $\mu(a) > 0$, there exists $b \in A$ such that $[a]_\mu \cap [b]_\mu \neq [a]_\mu \cap [c]_\mu$ for all $c \in B$, whereby $\mu\big((a \cap b) \triangle (a \cap c)\big) \neq 0$ for all $c \in B$. Hence $a \cap b \neq a \cap c$ for all $c \in B$, whereby $(X, A, \mu)$ is atomless over $(X, B, \mu)$.

$(iii) \Rightarrow (v)$: By $(iii)$, for every $a \in A$ with $\mu(a) > 0$, there exists $b \in A$ such that $b \subseteq a$ and

$$\mathbb{E}(b \mid B) = \frac{\mathbb{E}(a \mid B)}{2}.$$

Hence, $\mu(b) = \frac{\mu(a)}{2} > 0$ and for all $c \in B$,

$$\mu(b \cap c) = \frac{\mu(a \cap c)}{2}.$$

If $[a]_\mu \cap [b]_\mu = [a]_\mu \cap [c]_\mu$ for some $c \in B$, then $([a]_\mu \cap [b]_\mu) \cap [c]_\mu = ([a]_\mu \cap [c]_\mu) \cap [c]_\mu$, and thus $[(a \cap b) \cap c]_\mu = [(a \cap c) \cap c]_\mu$. Since $b \subseteq a$, we have $[b \cap c]_\mu = [a \cap c]_\mu = [a \cap b]_\mu = [b]_\mu$, whereby $\mu(b \cap c) = \mu(a \cap c) = \mu(b) > 0$. This contradicts the fact that

$$\mu(b \cap c) = \frac{\mu(a \cap c)}{2}.$$

Therefore, $\widehat{A}$ is atomless over $\widehat{B}$. □

For the notion of "atomless over", we have the following technical lemma:

**2.2.17 Lemma.** *Let $A \supseteq B$ be two probability algebras. If $A$ is atomless over $B$, then $\tau_B(a) \geq \omega$ for all $a \in A^+$.*

*Proof.* Suppose not. Then there exists $a \in A^+$ such that $A \restriction a = \{a \cap b \mid b \in \sigma(\{b_1, \cdots, b_n\} \cup B)\}$ for some

$b_1, \cdots, b_n \in A$. Then by Proposition 2.2.7, one has

$$\sigma(\{b_1, \cdots, b_n\} \cup B) = \{ \bigcup_{\epsilon_1, \cdots, \epsilon_n} (b_1^{\epsilon_1} \cap \cdots \cap b_n^{\epsilon_n} \cap c_\epsilon) \mid \epsilon_1, \cdots, \epsilon_n \in \{+1, -1\}, c_\epsilon \in B\}$$

where $b_i^{+1} = b_i$ and $b_i^{-1} = b_i^{\complement}$ for each $i = 1, \cdots, n$. Since $A$ is atomless over $B$ and $a \in A^+$, there exists $b_0 \in \sigma(\{b_1, \cdots, b_n\} \cup B)$ such that $a \cap b_0 \neq a \cap c$ for all $c \in B$. Then for some $\epsilon_1, \cdots, \epsilon_n \in \{+1, -1\}$ (without loss of generality, assume $\epsilon_1 = \cdots = \epsilon_n = +1$), one has $\mu(a \cap b_1 \cap \cdots \cap b_n) \neq 0$ and for all $c \in B$

$$a \cap b_1 \cap \cdots \cap b_n \cap b_0 \neq a \cap b_1 \cap \cdots \cap b_n \cap c.$$

Since $b_0 \in \sigma(\{b_1, \cdots, b_n\} \cup B)$, one has that $a \cap b_1 \cap \cdots \cap b_n \cap b_0 = a \cap b_1 \cap \cdots \cap b_n \cap c_0$ for some $c_0 \in B$, which is a contradiction. □

**2.2.18 Corollary.** *Let $A \supseteq B$ be two probability algebras. Suppose that $A$ is atomless over $B$. For all $n \in \mathbb{N}$, let $a_1, \cdots, a_n$ be in $A$. Then $A$ is atomless over $\sigma(B \cup \{a_1, \cdots, a_n\})$.* □

The following lemma gives more techniques for dealing with atomlessness.

**2.2.19 Lemma** ([19, Lemma 4.4])**.** *Let $M$ be a Polish space. Let $(\Omega, \mathcal{F}, P)$ be a probability space and let $\mathscr{A}, \mathscr{B}, \mathscr{C} \subseteq \mathcal{F}$ be $\sigma$-subalgebras.*

*(i) Suppose $\mathscr{A} \subseteq \mathscr{B}$ and $\mathscr{B}$ is atomless over $\mathscr{A}$. Let $G$ be a conditional distribution over $\mathscr{A}$ on $M$; i.e., for every Borel set $U \subseteq M$ the function $G(U)\colon \Omega \to [0,1]$ is $\mathscr{A}$-measurable, $G(\emptyset) = 0$ almost surely, $G(M) = 1$ almost surely, and if $U = \bigcup_{i \in \mathbb{N}} U_i$, where $U_i \cap U_j = \emptyset$ whenever $i \neq j \in \mathbb{N}$, then $G(U) = \sum_{i \in \mathbb{N}} G(U_i)$ almost surely. Then there is a $\mathscr{B}$-measurable function $f\colon \Omega \to M$ such that for every Borel $U \subseteq M$,*

$$\mathbb{E}(f \in U \mid \mathscr{A}) = G(U) \text{ a.s.}$$

*(ii) Suppose $\mathscr{A}$ and $\mathscr{B}$ are $\sigma$-subalgebras of $\mathscr{C}$. If $\mathscr{B}$ is atomless and $\mathscr{A}, \mathscr{B}$ are independent, then $\mathscr{C}$ is atomless over $\mathscr{A}$.* □

**Extending isomorphisms**

We first give a lemma from Fremlin [14].

**2.2.20 Lemma** ([14])**.** *Let $(A, \mu), (B, \nu)$ be measured algebras and $C \subseteq A$ a $\sigma$-complete subalgebra. Suppose that $f\colon C \to B$ is a measure-preserving homomorphism such that $B$ is atomless over $f[C]$, the image of $f$.*

*Then for every $a \in A$ there is a measure-preserving homomorphism $g$ from $\sigma(C \cup \{a\})$ to $B$ which extends $f$. Moreover, if $b \in B$ satisfies $\mu\big(a \cap f^{-1}(d)\big) = \nu(b \cap d)$ for every $d \in f[C]$, then $g$ can be taken to satisfy $g(a) = b$.*

*Proof.* The first part is [14, Lemma 3.4] and the second part is in the proof of it. □

The following theorem is an extension of Theorem 2.2.12.

**2.2.21 Theorem.** *Let $A$ and $B$ be probability algebras and let $C \subseteq A$ be a $\sigma$-complete subalgebra. Let $f\colon C \to B$ be a measure-preserving homomorphism. Suppose $\tau_C(a) = \tau_{f[C]}(b) = \kappa \geq \omega$ for all $a \in A^+$ and $b \in B^+$. Given $\eta < \kappa$, let $(a_\zeta)_{\zeta<\eta} \subseteq A$ and $(b_\zeta)_{\zeta<\eta} \subseteq B$ be two families. Let $C_0$ denote the $\sigma$-complete subalgebra $\sigma$-generated by $C \cup \{a_\zeta\}_{\zeta<\eta}$. If there is a measure-preserving homomorphism $f_0\colon C_0 \to B$ such that $f_0(a_\zeta) = b_\zeta$ for all $\zeta < \eta$ and $f_0$ extends $f$, then there is a measure-preserving isomorphism $g\colon A \to B$ which extends $f_0$.*

*Proof.* Let $(c_\xi)_{\xi<\kappa} \subseteq A$ and $(d_\xi)_{\xi<\kappa} \subseteq B$ be two families such that $A$ is $\sigma$-generated by $C \cup \{c_\xi\}_{\xi<\kappa}$ and $B$ is $\sigma$-generated by $f[C] \cup \{d_\xi\}_{\xi<\kappa}$. We will use transfinite induction to define $\sigma$-complete subalgebras $C_\xi, D_\xi$, and measure-preserving isomorphisms $g_\xi$ for all $\xi \leq \kappa$, where for some families $\{c'_\xi\}_{\xi<\kappa} \subseteq A, \{d'_\xi\}_{\xi<\kappa} \subseteq B$ to be determined later,

- the $\sigma$-complete subalgebra $C_\xi$ of $A$ is $\sigma$-generated by $\{c_\rho\}_{\rho<\xi} \cup \{c'_\rho\}_{\rho<\xi} \cup C \cup \{a_\zeta\}_{\zeta<\eta}$.
- the $\sigma$-complete subalgebra $D_\xi$ of $B$ is $\sigma$-generated by $\{d_\rho\}_{\rho<\xi} \cup \{d'_\rho\}_{\rho<\xi} \cup f[C] \cup \{b_\zeta\}_{\zeta<\eta}$.
- the measure-preserving isomorphism $g_\xi$ is from $C_\xi$ to $D_\xi$.

The family $(g_\xi)_{\xi\leq\kappa}$ will also satisfy: $g_0 = f_0$; $g_{\rho_2}$ extends $g_{\rho_1}$ whenever $\rho_1 < \rho_2 \leq \kappa$.

*Zero case:* Let $C_0 = \sigma(C \cup \{a_\zeta\}_{\zeta<\eta})$, $D_0 = f_0[C_0]$, and $g_0 = f_0$.

*Successor case:* Suppose that $C_\rho$, $D_\rho$, and $g_\rho$ have been defined for all ordinals $\rho \leq \xi$. Also $c'_\rho$ and $d'_\rho$ have been defined for all ordinals $\rho < \xi$.

(i) Let $C'_\xi$ be $\sigma(C_\xi \cup \{c_\xi\})$, a $\sigma$-complete subalgebra of $A$. For every $b \in B^+$, if $B \restriction b = D_\xi \restriction b$, then

$$\tau_{f[C]}(b) = \inf\big\{|X| : X \subseteq B \text{ and } D_\xi \restriction b \text{ is } \sigma\text{-generated by } \{b \cap x \mid x \in X \cup f[C]\}\big\} \leq |\xi| + |\xi| + |\eta| < \kappa.$$

This contradicts the fact that $\tau_{f[C]}(b) = \kappa$. Therefore $B \restriction b \neq \{b \cdot d \mid d \in D_\xi\}$, and thus $B$ is atomless over $D_\xi = g_\xi[C_\xi]$. By Lemma 2.2.20, there is a measure-preserving homomorphism $g'_\xi\colon C'_\xi \to B$ extending $g_\xi$. Let $d'_\xi = g'_\xi(c_\xi)$, so that $g'_\xi[C'_\xi] = D'_\xi = \sigma(D_\xi \cup \{d'_\xi\})$.

(ii) Now $g_\xi'^{-1}\colon D'_\xi \to C'_\xi$ is a measure-preserving isomorphism. For every $a \in A^+$, if $A \restriction a = C'_\xi \restriction a$, then

$$\tau_C(a) = \inf\big\{|X| : X \subseteq A \text{ and } C'_\xi \restriction a \text{ is } \sigma\text{-generated by } \{a \cap x \mid x \in X \cup C\}\big\} \leq |\xi| + |\xi| + |\eta| < \kappa.$$

This contradicts the fact that $\tau_C(a) = \kappa$. Then by the same argument, there is a measure-preserving homomorphism $f_\xi\colon D_{\xi+1} \to A$ extending $g_\xi'^{-1}$, where $D_{\xi+1} := \sigma(D'_\xi \cup \{d_\xi\})$. Let $c'_\xi = f_\xi(d_\xi)$ and $C_{\xi+1} = \sigma(C'_\xi \cup \{c'_\xi\})$. Then $g_{\xi+1} = f_\xi^{-1}\colon C_{\xi+1} \to D_{\xi+1}$ is a measure-preserving isomorphism extending $g_\xi$. This completes the inductive step to $\xi + 1$.

*Limit case:* For a limit ordinal $\xi > 0$, we consider

$$C^*_\xi = \bigcup_{\zeta<\xi} C_\zeta, D^*_\xi = \bigcup_{\zeta<\xi} D_\zeta, g^*_\xi = \bigcup_{\zeta<\xi} g_\zeta.$$

Then $C^*_\xi$ and $D^*_\xi$ are subalgebras of $A$ and $B$ respectively, and $g^*_\xi$ is a measure-preserving isomorphism. Moreover, $g^*_\xi$ preserves the metrics on $C^*_\xi$ and $D^*_\xi$. Therefore, $g^*_\xi$ has a unique extension to a measure-preserving isomorphism $g_\xi\colon C_\xi = \overline{C^*_\xi} \to D_\xi = \overline{D^*_\xi}$. This completes the inductive step to $\xi$.

This transfinite induction finishes at $\kappa$, when $C_\kappa = A$, $D_\kappa = B$, and $g = g_\kappa\colon A \to B$ is the desired measure-preserving isomorphism. □

**2.2.22 Remark.** *This proof is a revision of the proof of [14, Theorem 3.5].*

**2.2.23 Corollary.** *Let $A, B, C$ be countably $\sigma$-generated probability algebras with $C \subseteq A$. Let $f\colon C \to B$ be a measure-preserving homomorphism. Suppose $A$ is atomless over $C$ and $B$ is atomless over $f[C]$. Then there is a measure-preserving isomorphism $g\colon A \to B$ extending $f$.*

*Proof.* Use Lemma 2.2.17 and Theorem 2.2.21. □

The following result shows how to extend measure-preserving isomorphisms with a different assumption:

**2.2.24 Proposition.** *Let $A \subseteq B$ be countably $\sigma$-generated probability algebras and let the $\sigma$-complete subalgebra $C \subseteq B$ be atomic. Suppose $B$ is atomless over $A$. Let $f\colon A \to A$ be a measure-preserving automorphism. If there is a measure-preserving homomorphism $f_0\colon \sigma(A \cup C) \to B$ extending $f$, then there is a measure-preserving automorphism $g\colon B \to B$ extending $f_0$.*

*Proof.* Note that the $\sigma$-complete subalgebra $\sigma(A \cup C)$ is also countably $\sigma$-generated. By Corollary 2.2.23, it suffices to show that $B$ is atomless over $f_0[\sigma(A \cup C)]$. Let the set $\{c_i \mid i \in I\}$ be an enumeration of the

atoms in $C$. Since $C \subseteq B$ is a probability algebra, $I$ is a countable set. By Proposition 2.2.7,

$$\sigma(A \cup C) = \{\bigcup_{i \in I} a_i \cap c_i \mid a_i \in A \text{ for every } i \in I\}.$$

By Corollary 2.2.18, for each $i \in I$, we know that $B$ is atomless over $\sigma(A \cup c_i)$. Since $C$ is atomic, for every $b \in B^+$, we have that $b = \bigcup_{i \in I} b \cap c_i$, whereby for some $j \in I$ we have $b \cap c_j \in B^+$. Since $B$ is atomless over $\sigma(A \cup c_j)$, there is $d \in B$ such that $d \cap (b \cap c_j) \neq d' \cap (b \cap c_j)$ for all $d' \in \sigma(A \cup c_j)$. By Proposition 2.2.7, we know that $\sigma(A \cup c_j) = \{(a_1 \cap c_j) \cup (a_2 \cap c_j^{\complement}) \mid a_1, a_2 \in A\}$. Hence, $d \cap (b \cap c_j) \neq d' \cap (b \cap c_j)$ for all $d' \in \sigma(A \cup c_j)$ is equivalent to $d \cap (b \cap c_j) \neq a \cap (b \cap c_j)$ for all $a \in A$. Suppose there is $d'' \in \sigma(A \cup C)$ so that $d \cap b = d'' \cap b$. By Proposition 2.2.7, we may assume $d'' = \bigcup_{i \in I} a_i \cap c_i$, where $a_i \in A$ for each $i \in I$. Then since $\{c_i \mid i \in I\}$ is an enumeration of the atoms in $C$, we have $(d \cap b) \cap c_j = (d'' \cap b) \cap c_j = a_j \cap b \cap c_j$, which contradicts the fact that $d \cap (b \cap c_j) \neq a \cap (b \cap c_j)$ for all $a \in A$. Therefore $B$ is atomless over $\sigma(A \cup C)$. $\square$

## Maharam's Theorem

In this subsection, we present Maharam's celebrated theorem (Theorem 2.2.28), which characterizes the structure of measure algebras.

If $(\Omega_n, \mathcal{F}_n, \mu_n)$ is a sequence of probability spaces and $\alpha_n > 0$ for all $n \in \mathbb{N}$ such that $\sum_{n \in \mathbb{N}} \alpha_n = 1$, then the *(countable) convex combination* $(\Omega, \mathcal{F}, \mu) = \bigoplus_{n \in \mathbb{N}} (\Omega_n, \mathcal{F}_n, \alpha_n \mu_n)$ is defined as follows: replacing $\Omega_n$ by an isomorphic copy, we may assume that the sets $\Omega_n$ are pairwise disjoint. Let $\Omega = \bigoplus_{n \in \mathbb{N}} \Omega_n = \bigcup_{n \in \mathbb{N}} \Omega_n$. A set $B \subseteq \Omega$ is measurable (*i.e.*, in $\mathcal{F} := \bigoplus_{n \in \mathbb{N}} \mathcal{F}_n$) if $B \cap \Omega_n \in \mathcal{F}_n$ for all $n \in \mathbb{N}$. For every $B \in \mathcal{F}$, we define $\mu(B) := \bigoplus_{n \in \mathbb{N}} \alpha_n \mu_n(B) = \sum_{n \in \mathbb{N}} \alpha_n \mu_n(B \cap \Omega_n)$. Clearly, $\mu(\Omega) = 1$.

Analogously, a (countable) convex combination of probability algebras is defined.

Let $(\widehat{A}, \mu_1)$ and $(\widehat{B}, \mu_2)$ be measured algebras associated to measure spaces $(X_1, A, \mu_1)$ and $(X_2, B, \mu_2)$ respectively. The *free product* of measured algebras $(\widehat{A}, \mu_1)$ and $(\widehat{B}, \mu_2)$ is the measured algebra associated to the product measure space $(X_1 \times X_2, A \times B, \mu_1 \times \mu_2)$. We write $(\widehat{A}, \mu_1) \widehat{\times} (\widehat{B}, \mu_2)$, or $(\widehat{A} \widehat{\times} \widehat{B}, \mu_1 \widehat{\times} \mu_2)$, for this free product.

Let $[0,1]^\kappa$ denote the product of $\kappa$ copies of $[0,1]$ with Lebesgue measure and let $\mu_\kappa$ be the product measure on it. Let $(A_\kappa, \mu_\kappa)$ denote the measured algebra associated to the product measure space $([0,1]^\kappa, \mu_\kappa)$. For all subsets $\Delta \supseteq \Gamma$ of $\kappa$, let $\mathrm{pr}_\Gamma^\Delta \colon [0,1]^\Delta \to [0,1]^\Gamma$ be the canonical projection mapping. We write simply $\mathrm{pr}_\Gamma$ if $\Delta = \kappa$. The following proposition describes the Borel measurable subsets of $[0,1]^\kappa$.

**2.2.25 Proposition.** *Let $\kappa$ be an infinite ordinal and let $(A_\kappa, \mu_\kappa)$ denote the measured algebra associated to the product measure space $([0,1]^\kappa, \mu_\kappa)$. Then*

*(i) The set of Borel measurable subsets of* $[0,1]^\kappa$ *is*

$$B_\kappa = \{\mathrm{pr}_\Gamma^{-1}(G) \mid \Gamma \text{ is a countable subset of } \kappa \text{ and } G \text{ is a Borel subset of } [0,1]^\Gamma\}.$$

*(ii) For all countable subsets* $\Gamma$ *of* $\kappa$ *and all Borel subsets* $G$ *of the product measure space* $([0,1]^\Gamma, \mu_\Gamma)$*, we have*

$$\mu_\kappa(\mathrm{pr}_\Gamma^{-1}(G)) = \mu_\Gamma(G).$$

*(iii) For every* $A \in A_\kappa$*, there is* $B \in B_\kappa$ *such that* $[B]_{\mu_\kappa} = A$*, where* $[B]_{\mu_\kappa}$ *is the equivalence class of* $B$ *under* $\equiv_{\mu_\kappa}$*.*

*Proof.* $(i)$ Note that $B_\kappa$ is a $\sigma$-algebra and it contains the basic open sets in the product topological space $[0,1]^\kappa$. Hence $B_\kappa$ is the set of Borel measurable subsets of the product measure space $[0,1]^\kappa$.

$(ii)$ and $(iii)$ follow from [14, Theorem 1.11]. □

**2.2.26 Proposition** ([14])**.** *Let* $\kappa$ *be an infinite cardinal. Then*

*(i)* $(A_\kappa, \mu_\kappa)$ *is homogeneous and* $\tau(A_\kappa) = \kappa$*.*

*(ii) If* $(A, \mu)$ *is a homogeneous probability algebra with* $\tau(A) = \kappa$*, then it is isomorphic to the probability algebra* $(A_\kappa, \mu_\kappa)$*.*

*Proof.* $(i)$ is [14, Proposition 3.7(b)].

$(ii)$ is [14, Corollary 3.8]. □

**2.2.27 Theorem** ([14, Proposition 3.11])**.** *Let* $(A, \mu)$ *be a probability algebra and let* $C$ *be a* $\sigma$*-complete subalgebra of* $A$*. Then*

*(i) Suppose* $\kappa \geq \max\big(\omega, \tau_C(\mathbf{1})\big)$*. Let* $(C \widehat{\times} A_\kappa, \nu)$ *be the free product of probability algebras* $C$ *and* $A_\kappa$*, and let* $e\colon C \to C \widehat{\times} A_\kappa$ *be the canonical measure-preserving homomorphism. Then there exists an extension of* $e$ *to a measure-preserving homomorphism* $f\colon A \to C \widehat{\times} A_\kappa$*.*

*(ii) If, in (i),* $\tau_C(a) = \kappa$ *for all* $a \in A^+$*, then* $f$ *can be taken to be an isomorphism.* □

Maharam's Theorem [24] gives us the characterization of measure algebras.

**2.2.28 Theorem** (Maharam's Theorem, [24, Theorem 2]; see also [14, Theorem 3.9])**.** *For all atomless probability spaces* $\Omega$*, there is a countable set of distinct infinite cardinals* $S = \{\kappa_i \mid i \in I\}$ *such that the measure algebra of* $\Omega$ *is isomorphic to a convex combination of the homogeneous probability algebras* $[0,1]^{\kappa_i}$*. The set* $S$ *is uniquely determined by* $\Omega$ *and is called the Maharam spectrum of* $\Omega$*.* □

**2.2.29 Remark.** *The Maharam spectrum of* $[0,1]^\kappa$ *is* $\{\max(\aleph_0, \kappa)\}$.

## 2.3 Rearrangements

In this section, we discuss the notion of symmetric rearrangement of a Borel set and symmetric-decreasing rearrangement of a Borel measurable function. Then Lemma 2.3.3 is introduced, which is used in proving Theorem 5.3.9.

**2.3.1 Definition** ([23, Chapter 3])**.** *(i) For a Borel subset* $A \subseteq \mathbb{R}$ *of finite Lebesgue measure,* $A^*$ *is called the* ***symmetric rearrangement of the set*** $A$ *if it is the open interval centered at the origin whose length is the measure of* $A$*. Hence,*

$$A^* = \big\{x \mid |x| < \frac{\lambda(A)}{2}\big\},$$

*where* $\lambda$ *is the standard Lebesgue measure on* $\mathbb{R}$*.*

*(ii) For a Borel measurable function* $f\colon \mathbb{R} \to \mathbb{R}$*, vanishing at infinity,* $f^*$ *is the called the* ***symmetric-decreasing rearrangement of*** $f$ *if*

$$f^*(x) = \int_0^\infty \chi_{\{|f|>t\}^*}(x)dt.$$

The rearrangement $f^*$ has a number of easy properties. See more on pages 80–81 in [23].

**2.3.2 Fact.** *(i)* $f^*(x)$ *is nonnegative.*

*(ii)* $f^*(x)$ *is radially symmetric and nonincreasing; i.e.,* $f^*(x) = f^*(y)$ *if* $|x| = |y|$ *and* $f^*(x) \geq f^*(y)$ *if* $|x| \leq |y|$*.*

*(iii)* $f^*(x)$ *is lower semi-continuous.*

*(iv) For every* $t > 0$*,*

$$\big\{x \mid f^*(x) > t\big\} = \big\{x \mid |f(x)| > t\big\}^*.$$

*An easy, but important, consequence of this fact is that* $\operatorname{dist}\big(f^*(x)\big) = \operatorname{dist}\big(|f(x)|\big)$.

**2.3.3 Lemma** (Nonexpansivity of rearrangement, [23, Theorem 3.5])**.** *Let* $J\colon \mathbb{R} \to \mathbb{R}$ *be a nonnegative*

*convex function such that $J(0) = 0$. Let $f$ and $g$ be nonnegative functions on $\mathbb{R}$, vanishing at infinity. Then*

$$\int_{\mathbb{R}} J\big(f^*(x) - g^*(x)\big)dx \le \int_{\mathbb{R}} J\big(f(x) - g(x)\big)dx. \tag{2.1}$$

*If we also assume that $J$ is strictly convex, $f = f^*$, and $f$ is strictly decreasing, then equality in (2.1) implies that $g = g^*$.* □

Here we also define the rearrangements of Borel functions from $[0,1]$ to $[0,1]$, which are used to give a new proof of Theorem 2.4.4 in Section 5.4, which gives a formula for the Wasserstein distance between probability measures on the real line.

For a Borel measurable function $f\colon [0,1] \to [0,1]$, we define $\tilde{f}\colon \mathbb{R} \to \mathbb{R}$ as follows:

$$\tilde{f}(x) = \begin{cases} f(x) & x \in [0,1] \\ f(-x) & x \in [-1,0] \\ 0 & \text{otherwise} \end{cases}$$

Then the symmetric-decreasing rearrangement $\tilde{f}^*$ of $\tilde{f}$ is as follows:

$$\tilde{f}^*(x) = \int_0^\infty \chi_{\{|\tilde{f}|>t\}^*}(x)dt = \int_0^1 \chi_{\{\tilde{f}>t\}^*}(x)dt$$

Let $f^*(x)$ be $\tilde{f}^* \restriction_{[0,1]}$. Then $f^*(x)$ is from $[0,1]$ to $[0,1]$. We call $f^*$ the *decreasing rearrangement* of $f$. By Fact 2.3.2, we get

$$\tilde{f}^*(x) = \begin{cases} f^*(x) & x \in [0,1] \\ f^*(-x) & x \in [-1,0] \\ 0 & \text{otherwise} \end{cases}$$

By Fact 2.3.2, we have the following.

**2.3.4 Fact.** *Let $f$ be a Borel measurable function from $[0,1]$ to $[0,1]$. Let $f^*$ be the decreasing rearrangement of $f$. Then*

*(i) $f^*(x)$ is nonincreasing and lower semi-continuous.*

*(ii)* $\operatorname{dist}\big(f^*(x)\big) = \operatorname{dist}\big(f(x)\big)$.

**2.3.5 Corollary** (Nonexpansivity of rearrangement)**.** *Let $J\colon [0,1] \to [0,1]$ be a nonnegative convex function*

*such that* $J(0) = 0$. *Let* $f, g\colon [0,1] \to [0,1]$ *be Borel. Then*

$$\int_0^1 J\big(f^*(x) - g^*(x)\big)dx \le \int_0^1 J\big(f(x) - g(x)\big)dx.$$

*Proof.* This follows directly from Lemma 2.3.3. □

## 2.4 Wasserstein distances

In this section, we present the necessary background on Wasserstein distances, which is used in Section 5.3. Historically, Wasserstein distance was introduced by various people and thus had several names. It was suggested in Rüschendorf [27] that it should be called the Gini-Dall'Aglio-Kantorovich-Vasershtein-Mallows distance and the name "minimal $L_p$-metric" was preferred in Rüschendorf's survey [28]. In the rest of this section, we will mostly follow the notions and results in Villani's book [32, Chapter 6].

**2.4.1 Definition** (Wasserstein distance, [32, Definition 6.1])**.** *Let* $(\mathcal{X}, d)$ *be a Polish metric space, and let* $p \in [1, \infty)$. *For probability measures* $\mu, \nu$ *on* $\mathcal{X}$, *the* ***Wasserstein distance of order p*** *between* $\mu$ *and* $\nu$ *is defined as follows:*

$$W_p(\mu,\nu) = \Big(\inf_{\pi\in\Pi(\mu,\nu)} \int_{\mathcal{X}} d(x,y)^p d\pi(x,y)\Big)^{1/p} = \inf\Big\{\big[\mathbb{E}d(X,Y)^p\big]^{\frac{1}{p}} \mid \mathrm{dist}(X) = \mu, \mathrm{dist}(Y) = \nu\Big\},$$

*where* $\Pi(\mu,\nu)$ *denote the collection of all probability measures on* $\mathcal{X}\times\mathcal{X}$ *with marginals* $\mu$ *and* $\nu$ *on the first and second factors respectively.*

**2.4.2 Definition** (Wasserstein space, [32, Definition 6.4])**.** *Let* $(\mathcal{X}, d)$ *be a Polish metric space, and let* $p \in [1, \infty)$. *The* ***Wasserstein space of order*** $p$ *is defined as*

$$P_p(\mathcal{X}) := \big\{\mu \mid \mu \text{ is a probability measure on } \mathcal{X} \text{ and } \int_{\mathcal{X}} d(x_0,x)^p \mu(dx) < +\infty \text{ for some } x_0 \in \mathcal{X}\big\}.$$

This space does not depend on the choice of the point $x_0$ and $W_p$ defines a (finite) complete metric on $P_p(\mathcal{X})$. See [16, Propositions 2 and 6] for a proof.

In this thesis, a Wasserstein space is always equipped with the distance $W_p$ for corresponding $p$.

For $p = 1$, we have the following formula:

**2.4.3 Theorem** (Kantorovich-Rubinstein duality formula)**.** *Let* $(\mathcal{X}, d)$ *be a Polish metric space. Then for*

*all* $\mu, \nu \in P_1(\mathcal{X})$,

$$W_1(\mu,\nu) = \sup_{\|\psi\|_{Lip}\leq 1} \{\int_{\mathcal{X}} \psi d\mu - \int_{\mathcal{X}} \psi d\nu\}, \tag{2.2}$$

*where the Lipschitz norm of a real valued function* $\psi$ *on* $\mathcal{X}$ *is defined as*

$$\|\psi\|_{Lip} := \sup_{x\neq y} \frac{|\psi(x)-\psi(y)|}{d(x,y)}.$$

*Proof.* See [32, (6.3)]. □

The following theorem gives an integral formula for the Wasserstein distance between probability distributions on the line, which is used in the proof of Theorem 5.3.9.

**2.4.4 Theorem.** *For all* $\mu, \nu \in P_1(\mathbb{R})$, *one has*

$$W_1(\mu,\nu) = \int_{-\infty}^{+\infty} |F(x) - G(x)| dx,$$

*where* $F$ *and* $G$ *are distribution functions for* $\mu$ *and* $\nu$ *respectively.*

*Proof.* For a proof using the Kantorovich-Rubinstein duality formula, see Proposition 20.11 in Dudley's book [11]. □

**2.4.5 Remark.** *Theorem 2.4.4 was discovered independently by the following people at least: Dall'Aglio (1956) [9], Vallander (1973) [31], Szulga (1978) [30], and the author. Later, the author will revisit this theorem and give his proof using rearrangements and model theoretic results; see Theorem 5.4.2.*

Now we discuss convergence in Wasserstein spaces. The notation $\mu_i \to \mu$ means that $\mu_i$ *converges weakly* to $\mu$; *i.e.*, $\int_{\mathcal{X}} \varphi(x)\mu_i(dx) \to \int_{\mathcal{X}} \varphi(x)\mu(dx)$ for all bounded continuous functions $\varphi\colon \mathcal{X} \to \mathbb{R}$.

**2.4.6 Definition** (Weak convergence in $P_p$, [32, Definition 6.8])**.** *Let* $(\mathcal{X}, d)$ *be a Polish metric space and let* $p \in [1,\infty)$. *Let* $(\mu_i)_{i\in\mathbb{N}}$ *be a sequence in* $P_p(\mathcal{X})$ *and let* $\mu$ *be in* $P_p(\mathcal{X})$. *Then we say* $(\mu_i)_{i\in\mathbb{N}}$ ***converges weakly in*** $P_p(\mathcal{X})$ *if for some (and thus every)* $x_0 \in \mathcal{X}$, *one has* $\mu_i \to \mu$ *and*

$$\int_{\mathcal{X}} d(x_0,x)^p \mu_i(dx) \to \int_{\mathcal{X}} d(x_0,x)^p \mu(dx).$$

**2.4.7 Theorem** ([32, Theorem 6.9])**.** *Let* $(\mathcal{X}, d)$ *be a Polish space and let* $p \in [1,\infty)$. *Then the topology defined by the Wasserstein distance* $W_p$ *is the same as the topology of weak convergence in* $P_p(\mathcal{X})$. *In other words, if* $(\mu_i)_{i\in\mathbb{N}}$ *is a sequence in* $P_p(\mathcal{X})$ *and* $\mu$ *is in* $P_p(\mathcal{X})$, *then* $\mu_i$ *converges weakly in* $P_p(\mathcal{X})$ *to* $\mu$ *if and only if* $W_p(\mu_i,\mu) \to 0$. □

# Chapter 3

# Model theory of probability algebras

In this chapter, we discuss the model theory of probability algebras with an emphasis on atomless probability algebras. The pioneering work in this general direction was done by Keisler, Hoover, and Fajardo in the 80s. They did not work in an explicit model theoretic framework, but the results they proved were essentially model theoretic; see [19], [18], and [13]. Recently, Ben Yaacov [1] studied (atomless) probability algebras using his framework of compact abstract theories. Later his ideas were carried into continuous logic in the background sections of Berenstein and Henson [7], where probability algebras were treated as metric structures. Here we follow the work in [7].

In Section 3.1, we present background material for the theories $\mathsf{Pr}$ and $\mathsf{APr}$. Our new results are given in Sections 3.2–3.4. In Section 3.2, we discuss different notions of universality. In Section 3.3, we discuss several notions of homogeneity. In Section 3.4, we describe the saturated models of $\mathsf{APr}$; see Theorem 3.4.2, one of our main results in this chapter. We also compare continuous model theoretic saturation with the saturation property introduced by Hoover-Keisler [19]. In Section 3.5, we present material from Ben Yaacov and Usvyatsov [5] about $d$-finite tuples in models of $\mathsf{APr}$.

## 3.1 Basics of Pr and APr

In this section, we give axioms for the classes of (atomless) probability algebras and present basic model theoretic properties of these theories. These results are adapted from [7], and some of them rely on [1].

Let the signature $L_{\mathsf{Pr}}$ denote the set $\{\mathbf{0}, \mathbf{1}, {\cdot}^{\complement}, \cap, \cup, \mu\}$, where $\mathbf{0}$ and $\mathbf{1}$ are constant symbols, ${\cdot}^{\complement}$ is a unary function symbol, $\cap$ and $\cup$ are binary function symbols, and $\mu$ is a unary predicate symbol. Among those symbols, ${\cdot}^{\complement}$ and $\mu$ are 1-Lipschitz, and $\cap$ and $\cup$ are 2-Lipschitz. The theory $\mathsf{Pr}$ consists of the following axioms:

(i) Boolean algebra axioms:

$$\sup_x \sup_y \big(d(x \cup y, y \cup x)\big) = 0; \quad \sup_x \sup_y \big(d(x \cap y, y \cap x)\big) = 0$$
$$\sup_x \sup_y \sup_z d\big(x \cup (y \cup z), (x \cup y) \cup z\big) = 0$$

$\sup_x \sup_y \sup_z d\big(x \cap (y \cap z), (x \cap y) \cap z\big) = 0$

$\sup_x \sup_y d\big(x \cup (x \cap y), x\big) = 0$; $\quad \sup_x \sup_y d\big(x \cap (x \cup y), x\big) = 0$

$\sup_x \sup_y \sup_z d\big(x \cup (y \cap z), (x \cup y) \cap (x \cup z)\big) = 0$

$\sup_x \sup_y \sup_z d\big(x \cap (y \cup z), (x \cap y) \cup (x \cap z)\big) = 0$

$\sup_x d(x \cup x^{\complement}, \mathbf{1}) = 0$; $\quad \sup_x d(x \cap x^{\complement}, \mathbf{0}) = 0$

(ii) Measure axioms:

$\mu(\mathbf{0}) = 0$ and $\mu(\mathbf{1}) = 1$

$\sup_x \sup_y \big| \frac{\mu(x \cup y) + \mu(x \cap y)}{2} - \frac{\mu(x) + \mu(y)}{2} \big| = 0$

(iii) Connections between $\mu$ and $d$:

$\sup_x \sup_y \big| d(x, y) - \mu(x \triangle y) \big| = 0$, where $x \triangle y$ denotes the symmetric difference: $x \triangle y = (x \cap y^{\complement}) \cup (x^{\complement} \cap y)$.

Recall the definitions on page 30. Let $(\Omega, \mathcal{F}, \mu)$ be a probability space and let $(\widehat{\mathcal{F}}, \mu)$ be the probability algebra associated to $(\Omega, \mathcal{F}, \mu)$. Now we interpret $\widehat{\mathcal{F}}$ as an $L_{\mathsf{Pr}}$-structure $\mathcal{M} = (\widehat{\mathcal{F}}, \mathbf{0}, \mathbf{1}, \cdot^{\complement}, \cap, \cup, \mu, d)$. We interpret $\mathbf{0}$ as $[\emptyset]_\mu$, $\mathbf{1}$ as $[\Omega]_\mu$, the unary function $\cdot^{\complement}$ as the complement of equivalence classes of events, and $\cap, \cup$ as the intersection and union of equivalence classes of events. The predicate $\mu$ is interpreted as the measure of equivalence classes of events, and $d$, the metric on $\widehat{\mathcal{F}}$, is defined as $d(A, B) := \mu(A \triangle B)$ for all $A, B \in \widehat{\mathcal{F}}$. Note that the interpretations of all function and predicate symbols satisfy the moduli of uniform continuity of those symbols. Also $(\widehat{\mathcal{F}}, d)$ is a complete metric space; see [15, Lemma 323F]. Hence $\mathcal{M}$ is an $L_{\mathsf{Pr}}$-structure. A structure $\mathcal{M}$ as above is called a *probability structure*. More specifically, it is called the *probability structure associated to* $(\Omega, \mathcal{F}, \mu)$.

**3.1.1 Theorem** ([7])**.** *Let $\mathcal{M}$ be an $L_{\mathsf{Pr}}$-structure. Then $\mathcal{M}$ is a model of* $\mathsf{Pr}$ *if and only if $\mathcal{M}$ is isomorphic to the probability structure associated to $(\Omega, \mathcal{F}, \mu)$ for a probability space $(\Omega, \mathcal{F}, \mu)$.*

*Proof.* This is [7, Theorem 5.2]. □

From the form of its axioms, the theory $\mathsf{Pr}$ is universal. It is not complete. For example, the closed statement $\inf_x |\mu(x) - \frac{1}{2}| = 0$ is valid in some probability algebras but not all of them. To axiomatize the theory of atomless probability algebras, we add the following axiom, which states that every event splits into two disjoint events with approximately the same measure.

$(iv)$ $\sup_x \inf_y \big| \mu(x \cap y) - \mu(x \cap y^{\complement}) \big| = 0$

Let $\mathsf{APr}$ denote the set of axioms in $\mathsf{Pr}$ together with Axiom $(iv)$ above.

**3.1.2 Corollary** ([7])**.** *The models of* APr *are isomorphic to the probability structures associated to atomless probability spaces.*

*Proof.* This is [7, Corollary 6.1]. □

**3.1.3 Theorem** ([2] and [7])**.** *Let $\mathcal{M}$ be a model of* APr*. Then:*

*(i) The theory* APr *is separably categorical and complete.*

*(ii) The theory* APr *admits quantifier elimination.*

*(iii) The universal part of* APr *is* Pr*, and* APr *is the model completion of* Pr*.*

*(iv) Let $C \subseteq M$. Then $\mathrm{dcl}(C) = \mathrm{acl}(C) =$ the $\sigma$-subalgebra of $M$ $\sigma$-generated by $C$.*

*(v) Assume that $\mathcal{M}$ is the probability structure associated to a probability space $(\Omega, \mathcal{F}, \mu)$. Let $C \subseteq M$ and let $\mathcal{C}$ be the $\sigma$-algebra $\sigma$-generated by the events in $C$. For any two $n$-tuples $a, b \in M^n$, we have $\mathrm{tp}(a/C) = \mathrm{tp}(b/C)$ if and only if $\mathbb{P}(a_1^{\epsilon_1} \cap \cdots \cap a_n^{\epsilon_n} | \mathcal{C}) = \mathbb{P}(b_1^{\epsilon_1} \cap \cdots \cap b_n^{\epsilon_n} | \mathcal{C})$ a.s. for all $\epsilon_1, \cdots, \epsilon_n \in \{-1, 1\}$.*

*(vi) The theory* APr *is $\omega$-stable and its independence relation coincides with the probabilistic independence relation. That is, if $A, B, C$ are subsets of a model of* APr*, then*

$$A \mathop{\downarrow}_C B \quad \Longleftrightarrow \quad \mathbb{P}\big(a \cap b \mid \sigma(C)\big) = \mathbb{P}\big(a \mid \sigma(C)\big)\mathbb{P}\big(b \mid \sigma(C)\big) \quad \text{a.s., for all } a \in \sigma(A) \text{ and } b \in \sigma(B).$$

*Moreover, types over arbitrary sets of parameters are stationary.*

*Proof.* $(i)$ and $(ii)$ are [7, Corollary 6.2].

$(iii)$ As noted above, Pr is a universal theory; APr is the model completion of Pr by $(i)$, $(ii)$ and Proposition 1.10.5.

$(iv)$ This is [7, Lemma 6.7].

$(v)$ This is [7, Lemma 6.5].

$(vi)$ This is [2, Fact 2.10]. □

**3.1.4 Remark.** *By [20, Proposition 6.6], one also has*

$$A \mathop{\downarrow}_C B \quad \Longleftrightarrow \quad \mathbb{P}\big(a \mid \sigma(B) \vee \sigma(C)\big) = \mathbb{P}\big(a \mid \sigma(C)\big) \quad \text{a.s., for all } a \in \sigma(A).$$

The following is a technical lemma that will be used in the following chapters.

**3.1.5 Lemma.** *Suppose $(X, \mathscr{B}, \mu)$ is an atomless probability space where $X$ is a complete separable metric space and $\mathscr{B}$ is the $\sigma$-algebra of Borel subsets of $X$. For any two measurable partitions $P_1$ and $P_2$ of $X$ where $P_1 = (A_n)_{n\in\mathbb{N}}$ and $P_2 = (B_n)_{n\in\mathbb{N}}$, if for every $i \in \mathbb{N}$ we have $\mu(A_i) = \mu(B_i)$, then there is an automorphism $\phi$ of $(X, \mathscr{B}, \mu)$ satisfying $\phi(A_i) = B_i$ up to a null set for every $i \in \mathbb{N}$.*

*Proof.* Since $\mathscr{B}$ is atomless, $\mathscr{B} \restriction A_i$ and $\mathscr{B} \restriction B_i$ are atomless for every $i \in \mathbb{N}$. By the fact that $\mu(A_i) = \mu(B_i)$ and using Theorem 3.1.3(i), there is an $L_{\mathsf{Pr}}$-isomorphism $\Phi_i \colon \widehat{\mathscr{B} \restriction B_i} \to \widehat{\mathscr{B} \restriction A_i}$ for every $i \in \mathbb{N}$.

We define $\Phi \colon \widehat{\mathscr{B}} \to \widehat{\mathscr{B}}$ as follows:

$$\begin{aligned} &\Phi \colon \widehat{\mathscr{B}} \to \widehat{\mathscr{B}} \\ &[C]_\mu \mapsto \bigcup_{i\in\mathbb{N}} \Phi_i([C \cap B_i]_\mu) \text{ for every } C \in \mathscr{B} \end{aligned}$$

Then $\Phi([B_i]_\mu) = [A_i]_\mu$ for every $i \in \mathbb{N}$ and $\Phi$ is an $L_{\mathsf{Pr}}$-automorphism of $\widehat{\mathscr{B}}$. Then by Theorem 2.2.5, up to a null set, there is an automorphism $\phi$ of $(X, \mathscr{B}, \mu)$ satisfying $\phi(A_i) = B_i$ for every $i \in \mathbb{N}$. □

## 3.2 Universality of probability spaces

In this section, we present the equivalence of several notions of universality.

**3.2.1 Definition** (Hoover-Keisler, [19])**.** *Let $\mathcal{X}$ be a Polish metric space. A probability space $\Omega$ is called $\mathcal{X}$-**universal** if for all $\mathcal{X}$-valued random variables $y$ on an arbitrary probability space $\Gamma$, there is an $\mathcal{X}$-valued random variable $x$ on $\Omega$ such that $x$ and $y$ have the same distribution. We call $\Omega$ **universal** if it is $\mathcal{X}$-universal for every Polish metric space $\mathcal{X}$.*

**3.2.2 Theorem.** *For a probability space $\Omega$, the following are equivalent:*

*(i)* $\Omega$ *is universal.*

*(ii)* $\Omega$ *is* $[0,1]$*-universal.*

*(iii)* $\Omega$ *is* $[0,1]^{\mathbb{N}}$*-universal.*

*(iv)* $\Omega$ *is atomless.*

*Proof.* The equivalence $(i) \iff (iv)$ is [13, Theorem 1E.2].

The implications $(i) \implies (ii)$ and $(i) \implies (iii)$ are trivial.

The implication $(ii) \Longrightarrow (iv)$ is easy. If $\Omega$ is not atomless, then it has at least one atom. Consequently, the identity map from $[0,1]$ to $[0,1]$, which gives the uniform distribution, will not be realized by any $[0,1]$-valued random variable on $\Omega$.

The implication $(iii) \Longrightarrow (iv)$ is similar to the the implication $(ii) \Longrightarrow (iv)$. □

## 3.3 Homogeneous models of APr

In this section, we discuss several notions of homogeneity coming from measure theory and model theory. Then we will show connections between those notions.

First, let us recall some notions from Fremlin's book [14]; also see Section 2.2 for more details.

**3.3.1 Definition.** *Let $A$ be a complete boolean algebra and $\tau(A) := \min\{|X| : A \text{ completely generated by } X\}$.*

- *($\tau$-homogeneous) We say that $A$ is $\tau$-homogeneous if $\tau(A \restriction a) = \tau(A)$ for all $a \in A^+$.*
- *(homogeneous) We say that $A$ is homogeneous if $A \cong A \restriction a$ for all $a \in A^+$.*

**3.3.2 Remark.** *When $A$ is a measure algebra, $\tau(A) = \min\{|X| : A \text{ is } \sigma\text{-generated by } X\}$.*

**3.3.3 Proposition.** *Let $(M, \mu, d)$ be a probability structure. Then $\tau(M) + \aleph_0 = \|M\| + \aleph_0$ where $\|M\|$ is the density character of $M$.*

*Proof.* This is trivial when $M$ is finite. Assume $M$ is infinite, so $\tau(M)$ and $\|M\|$ are both infinite.

Let $X$ be a subset of $M$ such that $M$ is $\sigma$-generated by $X$. Then consider the smallest boolean algebra $A$ containing $X$. Since $X$ is infinite, $|X| = |A|$. Note that $A$ is a dense subset of $M$. Thus $\|M\| \leq |A| = |X|$, whereby $\|M\| \leq \tau(M)$. Let $Y$ be a dense subset of $M$. Since the $\sigma$-complete subalgebra $\sigma(Y)$ is a measure algebra, it is complete under the metric $d(a,b) = \mu(a \triangle b)$. Hence, we have $\sigma(Y) \supseteq \overline{Y} = M$. Thus $\tau(M) \leq |Y|$, whereby $\tau(M) \leq \|M\|$. Therefore, $\tau(M) = \|M\|$. □

Recall that model theoretic homogeneity is defined on page 10. The following result shows the connection among several notions of homogeneity.

**3.3.4 Proposition.** *Let $\mathcal{A} = (A, \mu, d)$ be an atomless probability structure. Then the following are equivalent:*

*(i) $A$ is $\tau$-homogeneous as a complete boolean algebra.*

*(ii) $A$ is homogeneous as a complete boolean algebra.*

*(iii) $\mathcal{A}$ is strongly $\|A\|$-homogeneous as an $L_{\mathsf{Pr}}$-structure.*

*(iv) $\mathcal{A}$ is strongly $\omega$-homogeneous as an $L_{\mathsf{Pr}}$-structure.*

*Proof.* The implication $(i) \Rightarrow (ii)$ is Proposition 2.2.11.

For $(ii) \Rightarrow (iii)$, there are two cases.

(a) When $\|A\| = \tau(A) = \aleph_0$, the proof is easy, since the $\sigma$-subalgebra $\sigma$-generated by a finite set is still finite.

(b) Suppose $\|A\| = \tau(A) > \aleph_0$. For a subset $C$ of $A$ with $\aleph_0 \leq \mathrm{Card}(C) < \|A\|$, let $\sigma(C)$ be the $\sigma$-subalgebra $\sigma$-generated by elements in $C$. Then we know $\tau(\sigma(C)) = \|\sigma(C)\| < \|A\|$. By Corollary 2.2.13, we get an automorphism of $A$ fixing $\sigma(C)$ pointwise, which shows that $A$ is strongly $\|A\|$-homogeneous.

The implication $(iii) \Rightarrow (iv)$ is trivial.

We complete the proof by showing $(iv) \Rightarrow (i)$. We argue as follows:

(1) For all $a, b \in A^+$, if $\mu(a) = \mu(b)$, then there exists an automorphism $f$ sending $a$ to $b$. Hence $\tau(A \restriction a) = \tau(A \restriction b)$.

(2) The function $\tau$ is an increasing function with respect to the measures, that is if $\mu(a) \leq \mu(b)$, then $\tau(A \restriction a) \leq \tau(A \restriction b)$. Because there always exists $b' \in A$ such that $a \subseteq b'$ and $\mu(b') = \mu(b)$, then by the above Fact (1), we have $\tau(A \restriction b') = \tau(A \restriction b)$. Since $a \subseteq b'$, we have $\tau(A \restriction a) \leq \tau(A \restriction b')$. Then we get $\tau(A \restriction a) \leq \tau(A \restriction b)$.

(3) If $a \cap b = 0$, then $\tau(A \restriction (a \cup b)) \leq \tau(A \restriction a) + \tau(A \restriction b) = \max\big(\tau(A \restriction a), \tau(A \restriction b)\big)$, when $\tau(A \restriction a)$ and $\tau(A \restriction b)$ are both infinite. Since $A$ is atomless, for all $a \in A^+$ we have that $\tau(A \restriction a)$ is infinite, which is actually a special case of Lemma 2.2.17.

From Facts (1), (2), and (3), we get that $\tau(A \restriction a)$ has a constant value for all $a \in A^+$. Hence we have proved that $A$ is $\tau$-homogeneous. □

By Proposition 2.2.26, for a homogeneous probability algebra $(A, \mu)$, one has $(A, \mu) \cong (A_\kappa, \mu_\kappa)$ where $\kappa = \tau(A)$.

**3.3.5 Remark.** *Let $(A, \mu, d)$ be an infinite probability structure. Then*

*(i) If $A$ is $\tau$-homogeneous as a complete boolean algebra, then $A$ is atomless.*

*(ii) The fact that $A$ is strongly $\omega$-homogeneous does not necessarily imply that $A$ is atomless.*

## 3.4 Saturated models of APr

In this section, we describe the $\kappa$-saturated models of $\mathsf{APr}$ for each infinite cardinal $\kappa$ and then we show that for atomless probability spaces, $\aleph_1$-saturation is equivalent to Hoover-Keisler saturation.

**3.4.1 Definition.** *Let $\kappa$ be an infinite cardinal. A probability space $\Omega$ is $\kappa$-saturated if its probability algebra is $\kappa$-saturated as an $L_{\mathsf{Pr}}$-structure.*

The following theorem is the main result of this chapter.

**3.4.2 Theorem.** *If $\Omega$ is an atomless probability space and $\kappa$ is an infinite cardinal, then the following are equivalent:*

*(i) $\Omega$ is $\kappa$-saturated.*

*(ii) Every cardinal in the Maharam spectrum of $\Omega$ is $\geq \kappa$.*

*(iii) $\Omega$ is $\kappa$-atomless.*

*Proof.* $(i) \implies (ii)$: Let $(\mathscr{B}, \mu)$ be the probability algebra associated to $\Omega$; hence $\mathscr{B} \models \mathsf{APr}$. Suppose the least cardinal $\lambda$ in the Maharam spectrum of $\mathscr{B}$ is less than $\kappa$. By Theorem 2.2.28 (Maharam's Theorem), $\Omega$ can be decomposed into two parts: $(\mathscr{B}, \mu) \cong (A_\lambda, \alpha\mu_\lambda) \oplus (\mathscr{B}', \mu')$ where $(A_\lambda, \mu_\lambda)$ is the probability algebra of $([0,1]^\lambda, \mu_\lambda)$, every cardinal in the Maharam spectrum of $\mathscr{B}'$ is $> \lambda$, the real number $\alpha \in (0,1]$, and $\mu'(1_{\mathscr{B}'}) = 1 - \alpha$. Note that $A_\lambda$ as a metric space with the metric $d_\lambda(a,b) := \mu_\lambda(a \triangle b)$ has a dense subset of cardinality $\lambda$. Let $\{a_\gamma \mid \gamma \in \lambda\}$ be such a dense subset. Without loss of generality, let $a_0$ be $[0,1]^\lambda$. Let $p(x/\{a_\gamma \oplus \mathbf{0}_{\mathscr{B}'} \mid \gamma \in \lambda\})$ consist of the following conditions:

- $\mu(x) = \frac{1}{2}$;
- $\mu(x \cap (a_\gamma \oplus \mathbf{0}_{\mathscr{B}'})) = \mu(x)\mu(a_\gamma \oplus \mathbf{0}_{\mathscr{B}'})$ for all $\gamma \in \lambda$.

We consider the probability algebra of $([0,1]^\lambda \times [0,1], \alpha\mu_\lambda \times \mu_1) \oplus ([0,1], (1-\alpha)\mu_1)$. It is a model of $\mathsf{APr}$. Let $a'_\gamma := (a_\gamma \times [0,1]) \oplus \mathbf{0}_{[0,1]}$ and $a := ([0,1]^\lambda \times [0, \frac{1}{2}]) \oplus \mathbf{0}_{[0,1]}$. Then we have that $\mathrm{tp}(a'_\gamma : \gamma \in \lambda) = \mathrm{tp}(a_\gamma \oplus \mathbf{0}_{[0,1]} : \gamma \in \lambda)$ and $a$ realizes $p(x/\{a'_\gamma : \gamma \in \lambda\})$. Hence, $p(x/\{a_\gamma \oplus \mathbf{0}_{\mathscr{B}'} \mid \gamma \in \lambda\})$ is consistent. Since $\mathscr{B}$ is $\kappa$-saturated, $p(x/\{a_\gamma \oplus \mathbf{0}_{\mathscr{B}'} \mid \gamma \in \lambda\})$ is realized by an element $a$ in $\mathscr{B}$. Say $a = A \oplus B$, with $A \in A_\lambda$ and $B \in \mathscr{B}'$. Then

$$\mu((A \oplus B) \cap (a_\gamma \oplus \mathbf{0}_{\mathscr{B}'})) = \mu(A \oplus B)\mu(a_\gamma \oplus \mathbf{0}_{\mathscr{B}'}), \forall \gamma \in \lambda.$$

Thus $\mu((A \cap a_\gamma) \oplus (\mathbf{0}_{\mathscr{B}'})) = (\alpha\mu_\lambda(A) + \mu'(B))\alpha\mu_\lambda(a_\gamma)$. Then, we get

$$\alpha\mu_\lambda(A \cap a_\gamma) = (\alpha\mu_\lambda(A) + \mu'(B))\alpha\mu_\lambda(a_\gamma), \forall \gamma \in \lambda.$$

Letting $\gamma = 0$ and recalling $a_0 = [0,1]^\lambda$, we have $\alpha\mu_\lambda(A) = (\alpha\mu_\lambda(A) + \mu'(B))\alpha$. Therefore, $\mu(A \oplus B) = \alpha\mu_\lambda(A) + \mu'(B) = \mu_\lambda(A)$. Also $\mu(A \oplus B) = \mu(a) = \frac{1}{2}$, whence $\mu_\lambda(A) = \frac{1}{2}$. Then

$$\alpha\mu_\lambda(A \cap a_\gamma) = \mu_\lambda(A)\alpha\mu_\lambda(a_\gamma), \forall\gamma \in \lambda,$$

whereby

$$\mu_\lambda(A \cap a_\gamma) = \mu_\lambda(A)\mu_\lambda(a_\gamma), \forall\gamma \in \lambda.$$

Note that the set $\{C \in A_\lambda \mid \mu_\lambda(A \cap C) = \mu_\lambda(A)\mu_\lambda(C)\}$ is a closed subset in $(A_\lambda, d_\lambda)$. Since $\{a_\gamma \mid \gamma \in \lambda\}$ is dense in $(A_\lambda, d_\lambda)$, we have $\mu_\lambda(A \cap C) = \mu_\lambda(A)\mu_\lambda(C)$ for all $C \in A_\lambda$. Then since $A \in A_\lambda$, we have $\mu_\lambda(A) = \mu_\lambda(A)^2$. Hence $\mu_\lambda(A) = 0$ or 1, which contradicts the fact that $\mu_\lambda(A) = \frac{1}{2}$.

$(ii) \Longrightarrow (iii)$: First, consider the case where $(\mathscr{B}, \mu) \cong (A_\lambda, \mu_\lambda)$ for some $\lambda \geq \kappa$, where $(A_\lambda, \mu_\lambda)$ is the measured algebra of the product measure space $([0,1]^\lambda, \mu_\lambda)$. Since $\Omega$ is atomless, by Lemma 2.2.17 we know that $\Omega$ is $\aleph_0$-atomless. Suppose $\kappa$ is uncountable. Take $A \subseteq \mathscr{B}$ with $|A| < \kappa$, and let $\mathscr{A} = \sigma(A)$. We will show that $\mathscr{B}$ is atomless over $\mathscr{A}$. Let $B_\lambda$ denote the set of Borel subsets of the measure space $([0,1]^\lambda, \mu_\lambda)$. By Proposition 2.2.25, for each $a \in A$, there is $c(a) \in B_\lambda$ of the form $\mathrm{pr}_\Gamma^{-1}(c_\Gamma(a))$ such that $[c(a)]_{\mu_\lambda} = a$, where $\Gamma \subseteq \lambda$ is countable, $c_\Gamma(a)$ is a Borel subset of $[0,1]^\Gamma$, and $\mathrm{pr}_\Gamma \colon [0,1]^\lambda \to [0,1]^\Gamma$ is the canonical projection. We define $\mathrm{Supp}(c(a)) := \Gamma$. Let $\mathcal{C}$ denote the $\sigma$-algebra $\sigma$-generated by $\{c(a) \mid a \in A\}$ in $B_\lambda$. Note that $\mathscr{B}$ and $\mathscr{A}$ are probability algebras associated to $([0,1]^\lambda, B_\lambda, \mu_\lambda)$ and $([0,1]^\lambda, \mathcal{C}, \mu_\lambda)$ respectively. By Lemma 2.2.16$(v)$, to show $\mathscr{B}$ is atomless over $\mathscr{A}$, it suffices to show $B_\lambda$ is atomless over $\mathcal{C}$. Define $\mathrm{Supp}(\mathcal{C}) := \bigcup_{a \in A} \mathrm{Supp}(c(a))$. Since $|\mathrm{Supp}(c(a))| \leq \aleph_0$ for each $a \in A$, we know that $\mathrm{Supp}(\mathcal{C})$ is of cardinality $|A| \times \aleph_0 < \kappa \leq \lambda$. Take any $\gamma \in \lambda \backslash \mathrm{Supp}(\mathcal{C})$. Let $\mathcal{B}$ denote the set of Borel subsets of $[0,1]^{\{\gamma\}} = [0,1]$. Then $\mathrm{pr}_{\{\gamma\}}^{-1}(\mathcal{B})$ is an atomless $\sigma$-subalgebra of $B_\lambda$. Since elements in $\mathcal{C}$ only depend on the coordinates in $\mathrm{Supp}(\mathcal{C})$ and elements in $\mathrm{pr}_{\{\gamma\}}^{-1}(\mathcal{B})$ only depend on the coordinate $\gamma$ which is not in $\mathrm{Supp}(\mathcal{C})$, we know that $\mathrm{pr}_{\{\gamma\}}^{-1}(\mathcal{B})$ and $\mathcal{C}$ are independent. By Lemma 2.2.19, $B_\lambda$ is atomless over $\mathcal{C}$.

For the general case, by Theorem 2.2.28 (Maharam's Theorem) $\mathscr{B}$ is isomorphic to a convex combination of probability algebras $(A_{\kappa_n})_{n \in \mathbb{N}}$, where $\{\kappa_n \mid n \in \mathbb{N}\}$ is the Maharam spectrum of $\Omega$. Let $p_n \colon \mathscr{B} \to A_{\kappa_n}$ be the canonical projection map for each $n \in \mathbb{N}$. Take $A \subseteq \mathscr{B}$ with $|A| < \kappa$, and let $\mathscr{A} = \sigma(A)$. As shown above, for each $\kappa_n$ in the Maharam spectrum of $\Omega$, we have that $A_{\kappa_n}$ is atomless over $p_n(\mathscr{A})$. Let $B \in \mathscr{B}^+$. Then for some $n \in \mathbb{N}$, we have $p_n(B) \in A_{\kappa_n}^+$. Since $A_{\kappa_n}$ is atomless over $p_n(\mathscr{A})$, there is $C_n \in A_{\kappa_n}$ such that $p_n(B) \cap C_n \neq p_n(B) \cap D_n$ for all $D_n \in p_n(\mathscr{A})$. Hence there is $C \in \mathscr{B}$ such that $B \cap C \neq B \cap D$ for all $D \in \mathscr{A}$, whereby $\mathscr{B}$ is atomless over $\mathscr{A}$.

$(iii) \Longrightarrow (i)$: Fix $A \subseteq \mathscr{B}$ with $|A| < \kappa$; let $\mathscr{A} := \sigma(A)$. By assumption, $\mathscr{B}$ is atomless over $\mathscr{A}$. Let $p(x)$

be any type in $S_1(A)$. By Theorem 3.1.3(v), for all $a, b$ in any elementary extension of $\mathscr{B}$, we have that

$$\mathrm{tp}(a/A) = \mathrm{tp}(b/A) \iff \mathbb{P}(a \mid \mathscr{A}) = \mathbb{P}(b \mid \mathscr{A}) \text{ and } \mathbb{P}(a\complement \mid \mathscr{A}) = \mathbb{P}(b\complement \mid \mathscr{A}).$$

If $p(x)$ is realized by $a$ in some elementary extension of $\mathscr{B}$, then $\mathbb{P}(a \mid \mathscr{A})$ and $\mathbb{P}(a\complement \mid \mathscr{A})$ are $\mathscr{A}$-measurable functions. By Lemma 2.2.16 (Maharam's Lemma), there exists $b \in \mathscr{B}$ such that

$$\mathbb{P}(a \mid \mathscr{A}) = \mathbb{P}(b \mid \mathscr{A}) \text{ and } \mathbb{P}(a\complement \mid \mathscr{A}) = \mathbb{P}(b\complement \mid \mathscr{A}).$$

Therefore $\mathrm{tp}(b/A) = \mathrm{tp}(a/A) = p(x)$, which means that $p(x)$ is realized in $\mathscr{B}$. □

The following definition is due to Hoover and Keisler [19].

**3.4.3 Definition** (Hoover-Keisler, [19])**.**

- *Let $\Omega$ and $\Gamma$ be probability spaces and let $\mathcal{X}$ be a Polish metric space. Let $x, y\colon \Gamma \to \mathcal{X}$ be random variables. The probability space $\Omega$ is said to have the* ***saturation property for*** $\mathrm{dist}(x, y)$ *if for every random variable $x'\colon \Omega \to \mathcal{X}$ with $\mathrm{dist}(x) = \mathrm{dist}(x')$, there is a random variable $y'\colon \Omega \to \mathcal{X}$ such that $\mathrm{dist}(x, y) = \mathrm{dist}(x', y')$.*
- *A probability space $\Omega$ is said to be* ***Hoover-Keisler saturated*** *if for any two random variables $x, y\colon \Gamma \to \mathcal{X}$, where $\Gamma$ is an arbitrary probability space and $\mathcal{X}$ is an arbitrary Polish metric space, $\Omega$ has the saturation property of $\mathrm{dist}(x, y)$.*

Here we show the relation between their definitions and ours.

**3.4.4 Corollary.** *An atomless probability space is Hoover-Keisler saturated if and only if it is $\aleph_1$-saturated.*

*Proof.* By [13, Theorem 3B.7], an atomless probability space $\Omega$ is Hoover-Keisler saturated if and only if every cardinal in its Maharam spectrum is uncountable. By Theorem 3.4, $\Omega$ is $\aleph_1$ saturated if and only if every cardinal in its Maharam spectrum is uncountable. □

**3.4.5 Theorem.** *For all infinite cardinals $\kappa$, the probability algebra of $[0,1]^\kappa$ is strongly $\kappa$-homogeneous and $\kappa$-saturated. Moreover, it is the unique model of $\mathsf{APr}$ of density character $\kappa$ with these properties.*

*Proof.* The first part follows directly from Proposition 3.3.4, Theorem 2.2.28 (Maharam's Theorem) and Theorem 3.4.2. The uniqueness follows from the following argument: for a complete theory $T$ and $\kappa \geq \mathrm{Card}(T)$, by a standard back-and-forth argument, any two $\kappa$-saturated models of $T$ of density character $\kappa$ are isomorphic to each other. □

## 3.5 On $d$-finite tuples in APr

Many first order model theory results fail in the context of continuous logic. This is partly because finite tuples in continuous logic may behave like infinite tuples in first order model theory. The notion of $d$-finite tuple was defined by Ben Yaacov and Usvyatsov in [5] to capture some aspects of finite tuples in first order model theory. Many results in first order model theory mentioning finite tuples will still hold in continuous logic after replacing finite tuples with $d$-finite tuples. Roughly speaking, $d$-finite tuples in continuous logic behave like finite tuples in first order model theory. See more details in [5].

**3.5.1 Definition** ([5]). *Let $a$ and $c$ be tuples and let $p = \mathrm{tp}(a/c)$. Then $a$ is called $d$-**finite over** $c$, equivalently, the type $p$ is called $d$-**finite**, if for every tuple $b$ and every $\epsilon > 0$, there is $\delta = \delta^{a,c}_{b,\epsilon} > 0$ such that whenever $a' \equiv_c a$ and $d(a, a') \leq \delta$, there is $b'$ such that $d(b, b') \leq \epsilon$ and $a'b' \equiv_c ab$. If $c = \emptyset$, we omit $c$ from the notation.*

For the definition when $a$ or $c$ is infinite, see [5].

Let $L$ be a countable signature. For a separably categorical $L$-theory $T$, the following theorem, which will be used in later chapters, characterizes the $d$-finite tuples in models of $T$.

**3.5.2 Theorem** ([5]). *Let $T$ be a separably categorical $L$-theory for a countable signature $L$. Let $\mathcal{M} \models T$ and let $f$ be a finite tuple in $M$. We let $T(f)$ denote the theory $Th(\mathcal{M}, f)$. Then $f$ is $d$-finite if and only if the theory $T(f)$ is separably categorical.*

*Proof.* This is [5, Proposition 2.9]. □

The following theorem identifies the $d$-finite tuples in models of APr.

**3.5.3 Theorem** ([5]). *Every finite tuple in a model of* APr *is $d$-finite.*

*Proof.* This is [5, Example 2.13]. □

# Chapter 4

# Axioms for the theory of random variable structures

The study of $[0,1]$-valued (atomless) random variable structures was initiated by Ben Yaacov [2], where he axiomatized the theory of random variable structures by $\mathsf{RV}(\mathsf{BY})$ via propositional continuous logic, and proved some model theoretic properties of the theory (see Theorem 5.2.3 in the next chapter). In this chapter, we axiomatize the theory of random variable structures with a different approach. In Section 4.1, using only basic measure theoretic probability theory, we provide an elementary approach to axiomatizing the theory of random variable structures by $\mathsf{RV}$ and the theory of atomless random variable structures by $\mathsf{ARV}$. In the next chapter, we make a further study of $\mathsf{ARV}$. In Section 4.2, we present Ben Yaacov's axiomatizations. In Section 4.3, we consider the class of structures obtained by adding the multiplication function to all models of $\mathsf{RV}$. Ben Yaacov [2] showed that the multiplication function is uniformly definable in models of $\mathsf{RV}$ and it follows from the general theory of extensions by definitions that this class of expansions of models of $\mathsf{RV}$ is axiomatizable. We axiomatize this class explicitly by $\mathsf{RV}(\times)$.

## 4.1 Axioms for RV and ARV

In this section, we define the (atomless) random variable structures. Then we present an elementary approach to axiomatizing these structures. Only basic measure theoretic probability theory is required. The main result is Theorem 4.1.10.

Let $(\Omega, \mathcal{F}, \mu)$ be a probability space. Consider the set of all $\mathcal{F}$-measurable functions $f\colon \Omega \to [0,1]$. Define the $L^1$-*metric* $d_1(f,g) := \int_\Omega |f-g| d\mu$ for all $\mathcal{F}$-measurable $f, g\colon \Omega \to [0,1]$. The set of such functions together with $d_1$ forms a pseudometric space, which is denoted by $\mathcal{L}^1\big((\Omega, \mathcal{F}, \mu), [0,1]\big)$, or simply by $\mathcal{L}^1(\mu, [0,1])$. For all $f, g \in \mathcal{L}^1(\mu, [0,1])$, we say that $f$ is equal to $g$ *almost surely*, and write $f =_{a.s.} g$ (or $f = g \quad a.s.$), if $f$ is equal to $g$ up to a null set. We denote the equivalence class of $f$ under $=_{a.s.}$ by $[f]_{a.s.}$. A random variable in $[f]_{a.s.}$ is called a *version* of the class. For each $F \in \mathcal{F}$, let $\chi_F$ denote the characteristic function of $F$, and

let $\mathbf{1}_F$ denote $[\chi_F]_{a.s.}$. Note that $N := \{f \mid \int_\Omega |f| d\mu = 0\} = \{f \mid f = 0 \quad a.s.\}$. Then the quotient space

$$L^1\big((\Omega, \mathcal{F}, \mu), [0,1]\big) = \mathcal{L}^1\big((\Omega, \mathcal{F}, \mu), [0,1]\big)/N$$

is a metric space with the $L^1$-metric $d_1$, called an $L^1$-*space*. It is well known that $L^1\big((\Omega, \mathcal{F}, \mu), [0,1]\big)$ with the $L^1$-metric $d_1$ is a complete metric space. When the underlying probability space is clear, $L^1\big((\Omega, \mathcal{F}, \mu), [0,1]\big)$ is often abbreviated as $L^1(\mu, [0,1])$, or just $L^1(\mathcal{F}, [0,1])$ when the underlying set $\Omega$ and the probability measure $\mu$ are clear. We write $L^1\big((\Omega, \mathcal{F}, \mu), \{0,1\}\big)$ for the set of equivalence classes of characteristic functions in $\mathcal{L}^1\big((\Omega, \mathcal{F}, \mu), [0,1]\big)$. Let $\mathbb{D}$ denote the dyadic numbers in $[0,1]$. We write $L^1\big((\Omega, \mathcal{F}, \mu), \mathbb{D}\big)$ for the set of equivalence classes of $\mathbb{D}$-valued simple functions in $\mathcal{L}^1\big((\Omega, \mathcal{F}, \mu), [0,1]\big)$. Clearly,

$$L^1\big((\Omega, \mathcal{F}, \mu), \{0,1\}\big) \subseteq L^1\big((\Omega, \mathcal{F}, \mu), \mathbb{D}\big) \subseteq L^1\big((\Omega, \mathcal{F}, \mu), [0,1]\big).$$

Moreover, we have that $L^1\big((\Omega, \mathcal{F}, \mu), \{0,1\}\big)$ is closed in $L^1\big((\Omega, \mathcal{F}, \mu), [0,1]\big)$, and $L^1\big((\Omega, \mathcal{F}, \mu), \mathbb{D}\big)$ is dense in $L^1\big((\Omega, \mathcal{F}, \mu), [0,1]\big)$. Let $A$ be a subset of $L^1\big((\Omega, \mathcal{F}, \mu), [0,1]\big)$. Let $\sigma(A) \subseteq \mathcal{F}$ denote the $\sigma$-subalgebra of $\mathcal{F}$-measurable sets generated by the random variables in the equivalence classes in $A$. We call $\sigma(A)$ *the $\sigma$-algebra generated by $A$.*

The elements in $L^1\big((\Omega, \mathcal{F}, \mu), [0,1]\big)$ are not $\mathcal{F}$-measurable functions, but equivalence classes of them. In probability theory, most useful functions, relations, and maps (such as continuous functions, integrals, inequality relations, conditional expectations) on measurable functions are well-defined on the equivalence classes of those functions. For example, for all $f, g \in L^1\big((\Omega, \mathcal{F}, \mu), [0,1]\big)$, we define the following:

$$\varphi\big([f]_{a.s.}\big) := [\varphi(f)]_{a.s.} \text{ for all continuous functions } \varphi\colon [0,1] \to [0,1]$$

$$\varphi\big([f]_{a.s.}, [g]_{a.s.}\big) := [\varphi(f,g)]_{a.s.} \text{ for all continuous functions } \varphi\colon [0,1]^2 \to [0,1]$$

$$\int_\Omega [f]_{a.s.} d\mu := \int_\Omega f d\mu$$

$$[f]_{a.s.} \leq [g]_{a.s.} \text{ iff } f \leq g \quad a.s.$$

$$\mathbb{E}([f]_{a.s.} \mid \mathscr{A}) := [\mathbb{E}(f \mid \mathscr{A})]_{a.s.} \text{ for all } \sigma\text{-subalgebras } \mathscr{A} \subseteq \mathcal{F}.$$

The readers can check that these definitions are well-defined. Therefore, it causes no harm (and is more readable) to denote an equivalence class in $L^1\big((\Omega, \mathcal{F}, \mu), [0,1]\big)$ by a member of the class. For example, when

we say

"Take $f, g \in L^1(\mu, [0,1])$ with the same distribution",

what we really mean is

"Take $[f]_{a.s.}, [g]_{a.s.} \in L^1(\mu, [0,1])$, where $f, g \in \mathcal{L}^1(\mu, [0,1])$ have the same distribution".

A ($[0,1]$-valued) *random variable structure* is based on a set of the form $M = L^1\big((\Omega, \mathcal{F}, \mu), [0,1]\big)$, where $(\Omega, \mathcal{F}, \mu)$ is a probability space. It is called an *atomless random variable structure*, if its underlying probability space is atomless. We use the setting of continuous logic [4] to discuss the model theory of random variable structures. Here we consider the signature $L_{\mathsf{RV}} = \{\mathbf{0}, \neg, \dot{-}, \frac{1}{2}, I\}$, where $\mathbf{0}$ is a constant symbol, $\dot{-}$ is a binary function symbol, $\neg$ and $\frac{1}{2}$ are unary function symbols, and $I$ is a unary predicate symbol. Recall that on $M^n$, we take the maximum metric. Among those symbols, $\neg$ is 1-Lipschitz, $\frac{1}{2}$ is $\frac{1}{2}$-Lipschitz, $\dot{-}$ is 2-Lipschitz and $I$ is 1-Lipschitz.

We interpret the symbols of $L_{\mathsf{RV}}$ in $M$ as follows:

- $\mathbf{0}^{\mathcal{M}}(\omega) = 0$ for all $\omega \in \Omega$
- $\neg^{\mathcal{M}}(f) = 1 - f$ for all $f \in M$
- $(\dot{-})^{\mathcal{M}}(f, g) = f \dot{-} g = \max(f - g, 0)$ for all $f, g \in M$
- $(\frac{1}{2})^{\mathcal{M}} f = f/2$ for all $f \in M$
- $I^{\mathcal{M}}(f) = \int_\Omega f d\mu$ for all $f \in M$
- $d^{\mathcal{M}}(f, g) = \int_\Omega |f - g| d\mu$ for all $f, g \in M$

Then $\mathcal{M} = \Big(L^1\big((\Omega, \mathcal{F}, \mu), [0,1]\big), \mathbf{0}, \neg, \dot{-}, \frac{1}{2}, I, d\Big)$ is an $L_{\mathsf{RV}}$-structure and is denoted by $L^1\big((\Omega, \mathcal{F}, \mu), [0,1]\big)$. Note that $\Big(\mathcal{L}^1\big((\Omega, \mathcal{F}, \mu), [0,1]\big), \mathbf{0}, \neg, \dot{-}, \frac{1}{2}, I, d\Big)$ is an $L_{\mathsf{RV}}$-prestructure whose associated (quotient) $L_{\mathsf{RV}}$-structure is $\Big(L^1\big((\Omega, \mathcal{F}, \mu), [0,1]\big), \mathbf{0}, \neg, \dot{-}, \frac{1}{2}, I, d\Big)$. Let $\mathcal{RV}$ denote the class of all random variable structures as $L_{\mathsf{RV}}$-structures and let $\mathcal{ARV}$ denote the class of all atomless random variable structures as $L_{\mathsf{RV}}$-structures. We will show that the classes $\mathcal{RV}$ and $\mathcal{ARV}$ are elementary.

In the signature $L_{\mathsf{RV}}$, we will also use the following symbols as shorthand for expressions built from symbols in $L_{\mathsf{RV}}$:

$$\mathbf{1} = \neg\mathbf{0}$$

$$x \dotplus y = \neg(\neg x \dot{-} y)$$

$$x \wedge y = x \dot{-} (x \dot{-} y)$$

$$x \vee y = \neg(\neg x \wedge \neg y)$$

$$\frac{x}{2} = \frac{1}{2}x$$

By induction on $n$, we define

$$\frac{1}{2^n}x = \frac{x}{2^n} = \frac{1}{2}\frac{x}{2^{n-1}} \text{ for all } n \in \mathbb{N},$$

$$x_1 \dot{+} x_2 \dot{+} \cdots \dot{+} x_n = (x_1 \dot{+} x_2 \dot{+} \cdots \dot{+} x_{n-1}) \dot{+} x_n,$$

$$x_1 \wedge x_2 \wedge \cdots \wedge x_n = (x_1 \wedge x_2 \wedge \cdots \wedge x_{n-1}) \wedge x_n,$$

and

$$x_1 \vee x_2 \vee \cdots \vee x_n = (x_1 \vee x_2 \vee \cdots \vee x_{n-1}) \vee x_n.$$

Let $\mathbb{D}$ denote the set of dyadic numbers in [0,1]. Consider $r \in \mathbb{D}$. Suppose $r = \frac{m}{2^n}$, where $m, n \in \mathbb{N}$, $0 < m < 2^n$, and $2 \nmid m$. We define

$$rx = \underbrace{\frac{x}{2^n} \dot{+} \frac{x}{2^n} \dot{+} \cdots \dot{+} \frac{x}{2^n}}_{m \text{ times}}.$$

When $r = 0$ or 1, we write $0x$ for $\mathbf{0}$ and $1x$ for $x$.

The theory $\mathsf{RV}$ consists of the following axioms:

(E1) $\sup_x \inf_y \max\Big(I(y \wedge \neg y), \big|I(x \wedge \neg x) - d(x,y)\big|\Big) = 0$

(E2) $\sup_x \Big|I(x \wedge \neg x) - \inf_y \big(I(y \wedge \neg y) \dot{+} d(x,y)\big)\Big| = 0$

(APPR) $\sup_x \inf_{y_1,\cdots,y_{2^n}} \big(d(x, \frac{1}{2^n}y_1 \dot{+} \frac{1}{2^n}y_2 \dot{+} \cdots \dot{+} \frac{1}{2^n}y_{2^n}) \dot{+} \max_{1 \le i \le 2^n} I(y_i \wedge \neg y_i)\big) \dot{-} \frac{1}{2^n} = 0 \quad \forall n \in \mathbb{N}$

(ADD) $\sup_x \sup_y \frac{1}{2}\big|I(x) - \big(I(x \dot{-} y) + I(y \wedge x)\big)\big| = 0$

(C) $\sup_x I(\mathbf{0} \dot{-} x) = 0; \quad \sup_x d(x \dot{-} \mathbf{0}, x) = 0; \quad |I(\mathbf{1}) - 1| = 0$

(H1) $\sup_x \sup_y d\big(\frac{x \dot{-} y}{2}, \frac{x}{2} \dot{-} \frac{y}{2}\big) = 0$

(H2) $\sup_x \sup_y d((\frac{x}{2} \dot{+} \frac{y}{2}) \dot{-} \frac{x}{2}, \frac{y}{2}) = 0; \quad \sup_x d(\frac{x}{2} \dot{+} \frac{x}{2}, x) = 0$

(H3) $\sup_x \sup_y d\big(\frac{x}{2} \vee \frac{y}{2}, \frac{1}{2}(x \vee y)\big) = 0$

(H4) $\sup_x \sup_y I\big((\frac{1}{2}x \wedge y) \dot{-} (x \wedge y)\big) = 0$

(H5) $\sup_x \sup_y \frac{1}{2}(\frac{x}{2} \dot{+} \frac{y}{2}) = \frac{x}{4} \dot{+} \frac{y}{4}$

(MET) $\sup_x \sup_y \frac{1}{2}\big|d(x,y) - \big(I(x \dot{-} y) + I(y \dot{-} x)\big)\big| = 0$

(N) $d(\neg \mathbf{1}, \mathbf{0}) = 0; \quad \sup_x \sup_y d(x \dot{-} y, \neg y \dot{-} \neg x) = 0$

(P1) $\sup_{x_1} \sup_{x_2} \sup_{y_1} \sup_{y_2} d(x_1 \dot{+} y_1, x_2 \dot{+} y_2) \dot{-} \big(d(x_1, x_2) \dot{+} d(y_1, y_2)\big) = 0$

(P2) $\sup_x \sup_y \Big(d(x \dot{+} y, x \vee y) \dot{-} \max\big(I(x \wedge \neg x), I(y \wedge \neg y)\big)\Big) = 0$

(P3) $\sup_x \sup_y \sup_z d\big((x \dot{+} y) \dot{+} z, x \dot{+} (y \dot{+} z)\big) = 0$

(S1) $\sup_x \sup_y \sup_z d\big((z \vee y) \dot{-} x, (z \dot{-} x) \vee y\big) \dot{-} I(x \wedge y) = 0$

(S2) $\sup_x \sup_y \sup_z I\Big(\big((x \dot{+} y) \wedge z\big) \dot{-} \big((x \wedge z) \dot{+} (y \wedge z)\big)\Big) = 0$

(L1) $\sup_x \sup_y d(x \vee y, y \vee x) = 0; \quad \sup_x \sup_y d(x \wedge y, y \wedge x) = 0$

(L2) $\sup_x \sup_y \sup_z d\big(x \vee (y \vee z), (x \vee y) \vee z\big) = 0$

(L3) $\sup_x \sup_y \sup_z d\big(x \wedge (y \wedge z), (x \wedge y) \wedge z\big) = 0$

(L4) $\sup_x \sup_y d\big(x \vee (x \wedge y), x\big) = 0; \quad \sup_x \sup_y d\big(x \wedge (x \vee y), x\big) = 0$

(L5) $\sup_x \sup_y \sup_z d\big(x \vee (y \wedge z), (x \vee y) \wedge (x \vee z)\big) = 0$

(L6) $\sup_x \sup_y \sup_z d\big(x \wedge (y \vee z), (x \wedge y) \vee (x \wedge z)\big) = 0$

Axioms (L1) to (L6) are the axioms for distributive lattices.

Let $\mathsf{ARV}$ be $\mathsf{RV}$ together with the following axiom:

(NA) $\sup_x \inf_y \big(\max(I(y \wedge \neg y), \big|I(y \wedge x) - \frac{I(x)}{2}\big|)\big) = 0$

**4.1.1 Proposition.** *Every random variable structure $\mathcal{M} = (M, \mathbf{0}, \neg, \dot{-}, \frac{1}{2}, I, d)$ is a model of $\mathsf{RV}$. Further, if $\mathcal{M}$ is an atomless random variable structure, then it is a model of $\mathsf{ARV}$.*

*Proof.* Assume $M = L^1\big((\Omega, \mathcal{F}, \mu), [0,1]\big)$ for some probability space $(\Omega, \mathcal{F}, \mu)$. By Theorem 1.2.3, it suffices to consider axioms in the $L_{\mathsf{RV}}$-prestructure $M_0 = \mathcal{L}^1\big((\Omega, \mathcal{F}, \mu), [0,1]\big)$. Note that $f \wedge g = \min(f, g)$ and $f \vee g = \max(f, g)$ for all $f, g \in M_0$. We will check the axioms one by one.

(E1) and (E2): We consider

$$\mathcal{X} = \{f \in M_0 \mid f \text{ is a characteristic function}\}.$$

For all $f \in M_0$, we have

$$|f(\omega) - \chi_{\{f \geq \frac{1}{2}\}}(\omega)| \leq |f(\omega) - \chi_A(\omega)| \text{ for all } A \in \mathcal{F} \text{ and all } \omega \in \Omega,$$

whereby $\mathrm{dist}(f,\mathcal{X}) = d(f,\chi_{\{f\geq\frac{1}{2}\}})$. Also we note that

$$d(f,\chi_{\{f\geq\frac{1}{2}\}}) = \int_\Omega |f-\chi_{\{f\geq\frac{1}{2}\}}|d\mu = \int_\Omega |f\wedge(1-f)|d\mu = I^{\mathcal{M}_0}(f\wedge\neg f),$$

whereby $\mathrm{dist}(f,\mathcal{X}) = I^{\mathcal{M}}(f\wedge\neg f)$. Then to verify Axioms (E1) and (E2), we need only check that

$$\sup_x \inf_y \max\big(\mathrm{dist}(y,\mathcal{X}), \big|\mathrm{dist}(x,\mathcal{X}) - d(x,y)\big|\big) = 0$$

and

$$\sup_x \big|\mathrm{dist}(x,\mathcal{X}) - \inf_y \big(\mathrm{dist}(y,\mathcal{X}) \dot{+} d(x,y)\big)\big| = 0.$$

Both are clear here.

(APPR): This axiom is an approximation result from real analysis. For all $n\in\mathbb{N}$ and $f\in M_0$, let $g_i = \chi_{\{f\geq\frac{i-1}{2^n}\}}$, for every $1\leq i\leq 2^n$. Then

$$\frac{1}{2^n}g_1 \dot{+}\cdots\dot{+}\frac{1}{2^n}g_{2^n} = \sum_{i=1}^{2^n}\frac{1}{2^n}g_i = \frac{1}{2^n}\chi_{\{0\leq f<\frac{1}{2^n}\}} + \frac{2}{2^n}\chi_{\{\frac{1}{2^n}\leq f<\frac{2}{2^n}\}} + \cdots + \frac{2^n}{2^n}\chi_{\{\frac{2^n-1}{2^n}\leq f\}}.$$

Then

$$\begin{aligned} d\big(f,\frac{1}{2^n}g_1\dot{+}\cdots\dot{+}\frac{1}{2^n}g_{2^n}\big) = d\big(f,\sum_{i=1}^{2^n}\frac{i}{2^n}\chi_{\{\frac{i-1}{2^n}\leq f<\frac{i}{2^n}\}}\big) &= \int_\Omega \big|f-\sum_{i=1}^{2^n}\frac{i}{2^n}\chi_{\{\frac{i-1}{2^n}\leq f<\frac{i}{2^n}\}}\big|d\mu \\ &\leq \int_\Omega \sum_{i=1}^{2^n}\frac{1}{2^n}\chi_{\{\frac{i-1}{2^n}<f\leq\frac{i}{2^n}\}}d\mu = \frac{1}{2^n}\mu(\Omega) = \frac{1}{2^n}\end{aligned}$$

Also note that $I(g_i\wedge\neg g_i) = 0$ for every $1\leq i\leq 2^n$. Consequently, (APPR) is true in $\mathcal{M}$.

(ADD): For all $a,b\in[0,1]$, we have $a = \max(a-b,0)+\min(a,b)$, *i.e.*, $a = a\dot{-}b + a\wedge b$.

(C): For all $a\in[0,1]$, we have $0\dot{-}a = 0$ and $a\dot{-}0 = 0$.

(H1): For all $a,b\in[0,1]$, we have $\frac{\max(a-b,0)}{2} = \max(\frac{a}{2}-\frac{b}{2},0)$

(H2): For all $a,b\in[0,1]$, we have $(\frac{a}{2}\dot{+}\frac{b}{2})\dot{-}\frac{a}{2} = (\frac{a}{2}+\frac{b}{2})-\frac{a}{2} = \frac{b}{2}$ and $\frac{a}{2}\dot{+}\frac{a}{2} = \frac{a}{2}+\frac{a}{2} = a$.

(H3): For all $a,b\in[0,1]$, we have $\max(\frac{a}{2},\frac{b}{2}) = \frac{\max(a,b)}{2}$.

(H4): For all $a,b\in[0,1]$, we have $\frac{a}{2}\wedge b\leq a\wedge b$.

(H5): For all $a,b\in[0,1]$, we have $\frac{a}{2}\dot{+}\frac{b}{2} = \frac{a}{2}+\frac{b}{2}$ and $\frac{a}{4}\dot{+}\frac{b}{4} = \frac{a}{4}+\frac{b}{4}$.

(MET): For all $f,g\in M_0$, we have $f\dot{-}g+g\dot{-}f = |f-g|$. Hence $I^{\mathcal{M}_0}(f\dot{-}g)+I^{\mathcal{M}_0}(g\dot{-}f) = I^{\mathcal{M}_0}(|f-g|) = \int_\Omega |f-g|d\mu$.

(N): For all $a,b\in[0,1]$, we have $a-b = (1-b)-(1-a) = \neg b-\neg a$, and thus $\max(a-b,0) = \max(\neg b-\neg a,0)$.

(P1): Let $f_1, f_2, g_1, g_2 \in M_0$. Since $d(f_1, f_2) \dot{+} d(g_1, g_2) = \min(1, d(f_1, f_2) + d(g_1, g_2))$ and $d(f_1 \dot{+} g_1, f_2 \dot{+} g_2) \leq 1$, to show $d(f_1 \dot{+} g_1, f_2 \dot{+} g_2) \leq d(f_1, f_2) \dot{+} d(g_1, g_2)$, it suffices to show $d(f_1 \dot{+} g_1, f_2 \dot{+} g_2) \leq d(f_1, f_2) + d(g_1, g_2)$. Then it suffices to show for all $a, b, c, d \in [0, 1]$, we have $|(a \dot{+} b) - (c \dot{+} d)| \leq |a - c| + |b - d|$. Without loss of generality, we may assume $a \dot{+} b \geq c \dot{+} d$. If $c \dot{+} d = 1$ then $a \dot{+} b = 1$, and thus $|(a \dot{+} b) - (c \dot{+} d)| = 0 \leq |a - c| + |b - d|$. Otherwise, $c \dot{+} d = c + d$. Then $|(a \dot{+} b) - (c \dot{+} d)| = |(a \dot{+} b) - (c + d)| \leq |(a + b) - (c + d)| \leq |a - c| + |b - d|$.

(P2): We will show $|(a \dot{+} b) - (a \vee b)| \leq \max(a \wedge \neg a, b \wedge \neg b)$, for all $a, b \in [0, 1]$. Then it follows that $\int_\Omega |(f \dot{+} g) - (f \vee g)| d\mu \leq \max(\int_\Omega f \wedge \neg f d\mu, \int_\Omega g \wedge \neg g d\mu)$, for all $f, g \in M_0$. Without loss of generality, we may assume $a \wedge \neg a \leq b \wedge \neg b$.

Suppose $b \wedge \neg b = b$, and thus $b \leq \frac{1}{2}$. Since $a \wedge \neg a \leq b \wedge \neg b = b$, either $a \leq b$ or $\neg a \leq b$. If $a \leq b \leq \frac{1}{2}$, then $a + b \leq 1$ and $a \vee b = b$. Hence, $|(a \dot{+} b) - (a \vee b)| = |a + b - b| = a \leq b$. If $\neg a \leq b$, then $a + b \geq 1$ and thus $a \geq b$, since $b \leq \frac{1}{2}$. Hence, $|(a \dot{+} b) - (a \vee b)| = |1 - a| = \neg a \leq b$.

Suppose $b \wedge \neg b = \neg b$, and thus $\neg b \leq \frac{1}{2}$. Since $a \wedge \neg a \leq b \wedge \neg b = \neg b$, either $a \leq \neg b$ or $\neg a \leq \neg b$. If $a \leq \neg b$, then $a + b \leq 1$, and thus $a \leq b$, since $\neg b \leq \frac{1}{2}$. Hence, $|(a \dot{+} b) - (a \vee b)| = |a + b - b| = a \leq \neg b$. If $\neg a \leq \neg b$, then $a \geq b$. Since $\neg b \leq \frac{1}{2}$, we have $a + b \geq 1$. Hence, $|(a \dot{+} b) - (a \vee b)| = |1 - a| = \neg a \leq \neg b$.

(P3): For all $a, b, c \in [0, 1]$, we have $(a \dot{+} b) \dot{+} c = \min(a + b, 1) \dot{+} c = \min\big(\min(a + b, 1) + c, 1\big)$. Similarly, $a \dot{+} (b \dot{+} c) = \min\big(a + \min(b + c, 1), 1\big)$.

If $a \dot{+} b = 1$, then $(a \dot{+} b) \dot{+} c = 1$ and $a + \min(b + c, 1) \geq a + b \geq 1$, and thus $a \dot{+} (b \dot{+} c) = 1$.

If $b \dot{+} c = 1$, then $a \dot{+} (b \dot{+} c) = 1$ and $\min(a + b, 1) + c \geq b + c \geq 1$, and thus $(a \dot{+} b) \dot{+} c = 1$.

Otherwise (*i.e.*, $a + b < 1$ and $b + c < 1$), we have $(a \dot{+} b) \dot{+} c = \min(a + b + c, 1)$ and $a \dot{+} (b \dot{+} c) = \min(a + b + c, 1)$.

Overall, we have $(a \dot{+} b) \dot{+} c = a \dot{+} (b \dot{+} c)$.

(S1): It suffices to show $|[(c \vee b) \dot{-} a] - [(c \dot{-} a) \vee b]| \leq a \wedge b$ for all $a, b, c \in [0, 1]$.

Suppose $a \leq b$. Then $a \wedge b = a$ and $c \vee b \geq a$. Thus $(c \vee b) \dot{-} a = (c \vee b) - a$. Suppose $c < a \leq b$. Then $(c \dot{-} a) \vee b = b$, and thus

$$|[(c \vee b) \dot{-} a] - [(c \dot{-} a) \vee b]| = |(b - a) - b| = a = a \wedge b.$$

Suppose $c \geq a$. Then $(c \dot{-} a) \vee b = (c - a) \vee b$. If $c - a \geq b$, then $c \geq b$, and thus

$$|[(c \vee b) \dot{-} a] - [(c \dot{-} a) \vee b]| = |(c - a) - (c - a)| = 0 \leq a \wedge b.$$

If $c - a < b$, then $|[(c \vee b) \dot{-} a] - [(c \dot{-} a) \vee b]| = |[(c \vee b) - a] - b| = |a + b - c \vee b|$. Since $c - a < b$, we have

$a+b>c$, and thus $a+b\geq c\vee b$. Hence $|a+b-c\vee b|=a+b-c\vee b\leq a=a\wedge b$, since $b\leq c\vee b$.

Suppose $a>b$. Then $a\wedge b=b$. If $c<a$, then $(c\dot{-}a)\vee b=b$, and $(c\vee b)\dot{-}a=0$, since $a>b$. Hence, $|[(c\vee b)\dot{-}a]-[(c\dot{-}a)\vee b]|=|0-b|=b=a\wedge b$. If $c\geq a>b$, then $(c\dot{-}a)\vee b=(c-a)\vee b$ and $(c\vee b)\dot{-}a=c\dot{-}a=c-a$. Hence, $|[(c\vee b)\dot{-}a]-[(c\dot{-}a)\vee b]|=|(c-a)-(c-a)\vee b|\leq b=a\wedge b$.

(S2): It suffices to show $(a\dotplus b)\wedge c\leq(a\wedge c)\dotplus(b\dotplus c)$ for all $a,b,c\in[0,1]$. Suppose $(a\dotplus b)\wedge c=a\dotplus b$. Then $c\geq a$ and $c\geq b$. Hence $(a\wedge c)\dotplus(b\wedge c)=a\dotplus b$, and thus $(a\dotplus b)\wedge c\leq(a\wedge c)\dotplus(b\dotplus c)$. Suppose $(a\dotplus b)\wedge c=c$. Note that $(a\wedge c)\dotplus(b\dotplus c)$ has four possibilities $a\dotplus b$, $a\dotplus c$, $c\dotplus b$, and $c\dotplus c$. Since $(a\dotplus b)\wedge c=c$, we see that $c$ is smaller than each of them. Hence, $c\leq(a\wedge c)\dotplus(b\dotplus c)$, and thus $(a\dotplus b)\wedge c\leq(a\wedge c)\dotplus(b\dotplus c)$.

(L1)–(L6): It is clear that all numbers in $[0,1]$ satisfy the axioms for distributive lattices. Then it follows that all elements in $M_0$ will satisfy these axioms.

When $(\Omega,\mathcal{F},\mu)$ is atomless, it is clear that $\Big(L^1\big((\Omega,\mathcal{F},\mu),[0,1]\big),\mathbf{0},\neg,\dot{-},\frac{1}{2},I,d\Big)$ satisfies (NA). Hence it is a model of $\mathsf{ARV}$. □

Indeed, $\mathsf{RV}$ also axiomatizes the class $\mathcal{RV}$ (see Theorem 4.1.10) and then $\mathsf{ARV}$ axiomatizes the class $\mathcal{ARV}$ (see Corollary 4.1.13), which are the main results from this section. Toward the proof of Theorem 4.1.10, we will prove the following results about the models of $\mathsf{RV}$. In the following arguments, we will interpret symbols of $L_{\mathsf{RV}}$ in a given model $\mathcal{M}$ of $\mathsf{RV}$ without putting $\mathcal{M}$ explicitly into the notations, for easier readability.

**4.1.2 Fact.** *Let $\mathcal{M}$ be a model of $\mathsf{RV}$. For all $x,y\in M$, we have the following properties:*

*(i)* $I(x)=0$ *if and only if* $x=\mathbf{0}$.

*(ii)* $\mathbf{0}\dot{-}x=\mathbf{0}$ *and* $x\dot{-}\mathbf{1}=\mathbf{0}$.

*(iii)* $\neg x=\mathbf{1}\dot{-}x$ *and* $\neg\neg x=x$.

*(iv)* $x\wedge\mathbf{0}=\mathbf{0}$ *and* $x\wedge\mathbf{1}=x$.

*(v)* $x\vee\mathbf{0}=x$ *and* $x\vee\mathbf{1}=\mathbf{1}$.

*(vi)* $x\dot{-}x=\mathbf{0}$.

*(vii)* $I(\neg x)=1-I(x)$.

*(viii)* $\frac{I(x)}{2}=I(\frac{x}{2})$.

*(ix)* $x\dotplus y=y\dotplus x$.

*(x)* $d(\frac{x}{2},\frac{y}{2})=\frac{1}{2}d(x,y)$.

*(xi)* $\mathbf{1} \dot{+} x = x \dot{+} \mathbf{1} = \mathbf{1}$ *and* $\mathbf{0} \dot{+} x = x \dot{+} \mathbf{0} = x$.

*(xii) If* $\frac{x}{2} \dot{+} \frac{y}{2} = \frac{\mathbf{1}}{2}$, *then* $x = \neg y$.

*Proof.* $(i)$ By (C), we have $\mathbf{0} \dot{-} \mathbf{0} = \mathbf{0}$ and $0 = I(\mathbf{0} \dot{-} \mathbf{0}) = I(\mathbf{0})$. For the converse, suppose $I(x) = 0$. Using (MET) and (C), we have $d(x, \mathbf{0}) = I(x \dot{-} \mathbf{0}) + I(\mathbf{0} \dot{-} x) = I(x) + I(\mathbf{0}) = I(x) = 0$, and thus $x = \mathbf{0}$.

$(ii)$ By $(i)$ and (C), we have $\mathbf{0} \dot{-} x = \mathbf{0}$ for all $x$. In particular, $\mathbf{0} \dot{-} \neg x = \mathbf{0}$, so using (N), we have $x \dot{-} \mathbf{1} = \neg\mathbf{1} \dot{-} \neg x = \mathbf{0} \dot{-} \neg x = \mathbf{0}$.

$(iii)$ By (N) and (C), we have $\mathbf{1} \dot{-} x = \neg x \dot{-} \neg\mathbf{1} = \neg x \dot{-} \mathbf{0} = \neg x$. Using (L1), $(ii)$ and (C), we get $\neg\neg x = \mathbf{1} \dot{-} (\mathbf{1} \dot{-} x) = \mathbf{1} \wedge x = x \wedge \mathbf{1} = x \dot{-} (x \dot{-} \mathbf{1}) = x$.

$(iv)$ By (L1) and $(ii)$, we have $x \wedge \mathbf{0} = \mathbf{0} \wedge x = \mathbf{0} \dot{-} (\mathbf{0} \dot{-} x) = \mathbf{0}$. By $(ii)$ and (C), we get $x \wedge \mathbf{1} = x \dot{-} (x \dot{-} \mathbf{1}) = x \dot{-} \mathbf{0} = x$.

$(v)$ Using $(iv)$ and $(iii)$, we have

$$x \vee \mathbf{0} = \neg(\neg x \wedge \neg\mathbf{0}) = \neg(\neg x \wedge \mathbf{1}) = \neg(\neg x) = x.$$

Using (N) and $(iv)$, we have

$$x \vee \mathbf{1} = \neg(\neg x \wedge \neg\mathbf{1}) = \neg(\neg x \wedge \mathbf{0}) = \neg\mathbf{0} = \mathbf{1}.$$

$(vi)$ Setting $x = y$ in (MET), we have $I(x \dot{-} x) + I(x \dot{-} x) = 0$; this yields $x \dot{-} x = \mathbf{0}$ using $(i)$.

$(vii)$ By (C), (ADD), $(iii)$, and $(iv)$, we have $1 = I(\mathbf{1}) = I(\mathbf{1} \dot{-} x) + I(x \wedge \mathbf{1}) = I(\neg x) + I(x)$, whereby $I(\neg x) = 1 - I(x)$.

$(viii)$ By (H2), we have $x \dot{-} \frac{x}{2} = (\frac{x}{2} \dot{+} \frac{x}{2}) \dot{-} \frac{x}{2} = \frac{x}{2}$. Then by (ADD) and (L1), we have

$$I(x) = I(x \dot{-} \frac{x}{2}) + I(\frac{x}{2} \wedge x) = I(\frac{x}{2}) + I(x \wedge \frac{x}{2}) = I(\frac{x}{2}) + I\big(x \dot{-} (x \dot{-} \frac{x}{2})\big) = I(\frac{x}{2}) + I(x \dot{-} \frac{x}{2}) = I(\frac{x}{2}) + I(\frac{x}{2}).$$

$(ix)$ To show $x \dot{+} y = y \dot{+} x$, it suffices to show $\neg(\neg x \dot{-} y) = \neg(\neg y \dot{-} x)$, which follows from (N) and $(iii)$.

$(x)$ By (MET), (H1), and $(viii)$, we have

$$d(\frac{x}{2}, \frac{y}{2}) = I(\frac{x}{2} \dot{-} \frac{y}{2}) + I(\frac{y}{2} \dot{-} \frac{x}{2}) = I(\frac{x \dot{-} y}{2}) + I(\frac{y \dot{-} x}{2}) = \frac{I(x \dot{-} y)}{2} + \frac{I(y \dot{-} x)}{2} = \frac{d(x, y)}{2}.$$

$(xi)$ By (N) and $(ii)$, we have $\mathbf{1} \dot{+} x = \neg(\neg\mathbf{1} \dot{-} x) = \neg(\mathbf{0} \dot{-} x) = \neg\mathbf{0} = \mathbf{1}$. Then by $(ix)$, we have $x \dot{+} \mathbf{1} = \mathbf{1} \dot{+} x = \mathbf{1}$.

By $(iii)$, $\mathbf{0} \dot{+} x = \neg(\neg\mathbf{0} \dot{-} x) = \neg(\mathbf{1} \dot{-} x) = \neg\neg x = x$. By $(ix)$, we have $x \dot{+} \mathbf{0} = \mathbf{0} \dot{+} x = x$.

$(xii)$ By (H2), we have $\frac{\mathbf{1}}{2} \dot{-} \frac{x}{2} = (\frac{x}{2} \dot{+} \frac{y}{2}) \dot{-} \frac{x}{2} = \frac{y}{2}$. Then by (H1), we have $\mathbf{1} \dot{-} x = y$, whereby $x = \neg y$ by $(iii)$. □

**4.1.3 Proposition.** *Let $\mathcal{M}$ be a model of* RV. *Let $D = \{x \in M \mid I(x \wedge \neg x)\} = 0$. For all $x, y \in D$, define $x^{\complement} := \neg x$, $x \cap y := x \wedge y$, $x \cup y := x \vee y$, and $\mu(x) = I(x)$. Then $D$ is a uniformly definable set and $(D, \mathbf{0}, \mathbf{1}, {}^{\complement}, \cap, \cup, \mu)$ is a model of* Pr. *Moreover, if $\mathcal{M}$ is of the form $L^1\big((\Omega, \mathcal{F}, \mu), [0,1]\big)$ for the probability space $(\Omega, \mathcal{F}, \mu)$, then $D$ is $L^1\big((\Omega, \mathcal{F}, \mu), \{0,1\}\big)$.*

*Proof.* Let $\mathcal{M}$ be a model of RV and let $D = \{y \in M \mid I(y \wedge \neg y) = 0\}$. By Fact 4.1.2$(i)$, we know $D = \{y \in M \mid y \wedge \neg y = \mathbf{0}\}$. By Theorem 1.6.1, (E1), and (E2), we know that $I(x \wedge \neg x) = \operatorname{dist}(x, D)$. Hence, $D$ is a uniformly definable set in $M$.

First, we want to show that $D$ is closed under $\neg, \wedge, \vee$ and also $\mathbf{0}, \mathbf{1} \in D$. For all $x, y \in D$, we have $x \wedge \neg x = y \wedge \neg y = \mathbf{0}$. Since $\neg\neg x = x$ and $x \wedge \neg x = \neg x \wedge x$, we get $\neg x \wedge \neg(\neg x) = \mathbf{0}$, whence $\neg x \in D$. To show $x \wedge y \in D$, it suffices to show $(x \wedge y) \wedge \neg(x \wedge y) = (x \wedge y) \wedge (\neg x \vee \neg y) = \mathbf{0}$. By the fact that $\vee$ and $\wedge$ satisfy the distributive lattice axioms, we need only show that $\big((x \wedge y) \wedge \neg x\big) \vee \big((x \wedge y) \wedge \neg y\big) = \mathbf{0}$, which is true since $x \wedge \neg x = y \wedge \neg y = \mathbf{0}$. Then since $x \vee y = \neg(\neg x \wedge \neg y)$, we know $x \vee y \in D$ as well. By Fact 4.1.2$(iv)$, we know $\mathbf{0} \wedge \neg\mathbf{0} = \mathbf{0}$, and thus $\mathbf{0} \in D$. Hence, $\mathbf{1} = \neg\mathbf{0} \in D$.

Second, for all $x, y \in D$, define $x \cap y := x \wedge y$, $x \cup y := x \vee y$, and $x^{\complement} = \neg x$. We show that $(D, \mathbf{0}, \mathbf{1}, {}^{\complement}, \cap, \cup, \mu)$ is a model of Pr. For all $x \in D$, we have $x \wedge \neg x = \mathbf{0}$, and then $\neg(x \wedge \neg x) = \neg\mathbf{0} = \mathbf{1}$. Then by Fact 4.1.2$(iii)$ and (L1), we have $\mathbf{1} = \neg(x \wedge \neg x) = \neg x \vee \neg\neg x = \neg x \vee x = x \vee \neg x$. Because $\wedge, \vee$ also satisfy the axioms for distributive lattices, we see that $(D, \mathbf{0}, \mathbf{1}, {}^{\complement}, \cup, \cap)$ satisfies all boolean algebra axioms in Pr.

For all $x \in D$, define $\mu(x) := I(x)$. By Fact 4.1.2$(i)$ and $(vii)$, we have $\mu(\mathbf{0}) = 0$ and $\mu(\mathbf{1}) = 1$. For all $x, y \in D$, we have $I(x \vee y) = I\big(\neg(\neg x \wedge \neg y)\big) = 1 - I(\neg x \wedge \neg y)$, by Fact 4.1.2$(vii)$. By (ADD) and (N), we have $I(\neg y) = I(\neg y \dot{-} \neg x) + I(\neg x \wedge \neg y) = I(x \dot{-} y) + I(\neg x \wedge \neg y)$, and thus $I(\neg x \wedge \neg y) = I(\neg y) - I(x \dot{-} y)$. Hence, $I(x \vee y) = 1 - I(\neg x \wedge \neg y) = 1 - (I(\neg y) - I(x \dot{-} y))$, whence $I(x \vee y) = I(y) + I(x \dot{-} y)$ by Fact 4.1.2 $(vii)$. By (ADD), we have $I(x) = I(x \dot{-} y) + I(y \wedge x)$. Then by eliminating the term $I(x \dot{-} y)$, we get $I(x \vee y) + I(y \wedge x) = I(x) + I(y)$, whence $I(x \cup y) + I(x \cap y) = I(x) + I(y)$. Therefore $\mu(x \cup y) + \mu(x \cap y) = \mu(x) + \mu(y)$. Consequently, $(D, \mathbf{0}, \mathbf{1}, {}^{\complement}, \cap, \cup, \mu)$ satisfies the measure axioms in Pr.

Next, for all $x, y \in D$, by (P2) we know $d(x \dot{+} y, x \vee y) = 0$, and thus $x \dot{+} y = x \vee y$. Since $(x \dot{+} y) = \neg(\neg x \dot{-} y)$, by Fact 4.1.2$(iii)$ we have $x \dot{-} y = \neg(\neg x \dot{+} y)$. Then by Fact 4.1.2$(iii)$, we have $x \dot{-} y = \neg(\neg x \dot{+} y) = \neg(\neg x \vee y) = \neg\neg x \wedge \neg y = x \wedge \neg y$. By (MET), we have

$$d(x, y) = I(x \dot{-} y) + I(y \dot{-} x) = I(x \wedge \neg y) + I(y \wedge \neg x) = \mu(x \cap y^{\complement}) + \mu(y \cup x^{\complement}) = \mu(x \triangle y).$$

Hence, $(D, \mathbf{0}, \mathbf{1}, \cdot^{\complement}, \cap, \cup, \mu)$ satisfies Axiom $(iii)$ in $\mathsf{Pr}$. Since $d$ is a complete metric on $M$ and $D$ is a zeroset (thus it is closed), the metric $d$ is complete on $D$.

Since $x^{\complement} = \neg x$ for all $x \in D$ and $\neg$ is 1-Lipschitz, we get $\cdot^{\complement}$ is 1-Lipschitz. By (P2) and Fact 4.1.2$(iii)$, for all $x, y \in D$, we have

$$x \cap y = x \wedge y = \neg\neg(\neg\neg x \wedge \neg\neg y) = \neg(\neg x \vee \neg y) = \neg(\neg x \dot{+} \neg y) = \neg\neg(\neg\neg x \dot{-} \neg y) = x \dot{-} \neg y.$$

Since $\neg$ is 1-Lipschitz and $\dot{-}$ is 2-Lipschitz, we have that $\cap$ is 2-Lipschitz. Since $x \cup y = (x \cap y)^{\complement}$ for all $x, y \in D$, we know that $\cup$ is 2-Lipschitz. Since $\mu(x) = I(x)$ for all $x \in D$ and $I$ is 1-Lipschitz, we know that $\mu$ is 1-Lipschitz. Hence, $(D, \mathbf{0}, \neg, \cdot^{\complement}, \cap, \cup, \mu)$ is an $L_{\mathsf{Pr}}$-structure. Therefore, $(D, \mathbf{0}, \neg, \cdot^{\complement}, \cap, \cup, \mu)$ is a model of $\mathsf{Pr}$.

Suppose $\mathcal{M}$ is of the form $L^1\big((\Omega, \mathcal{F}, \mu), [0,1]\big)$, where $(\Omega, \mathcal{F}, \mu)$ is a probability space. Then for every $f \in L^1\big((\Omega, \mathcal{F}, \mu), \{0,1\}\big)$, there is $A \in \mathcal{F}$, such that $f = [\chi_A]_{a.s.}$. Thus $f \wedge \neg f = \mathbf{0}$, whereby $f \in D$. For the converse, take $x \in D$ with $I(x \wedge \neg x) = 0$. Suppose $x = [f]_{a.s.}$ for an $\mathcal{F}$-measurable $f \colon \Omega \to [0,1]$. Then $\int_\Omega \min(f, 1-f) d\mu = 0$, whereby $f$ is *a.s.* a characteristic function. Hence $x \in L^1\big((\Omega, \mathcal{F}, \mu), \{0,1\}\big)$, and thus $D = L^1\big((\Omega, \mathcal{F}, \mu), \{0,1\}\big)$. □

The following lemmas will be used in the proofs of Proposition 4.1.7 and Theorem 4.1.10.

**4.1.4 Lemma.** *Let $\mathcal{M} \models \mathsf{RV}$. Then:*

*(i) For all $m, n \in \mathbb{N}$ and all $x_1, \cdots, x_n \in M$, we have*

$$\frac{1}{2}\Big(\frac{x_1}{2^m} \dot{+} \cdots \dot{+} \frac{x_n}{2^m}\Big) = \frac{x_1}{2^{m+1}} \dot{+} \cdots \dot{+} \frac{x_n}{2^{m+1}}.$$

*(ii) For all $m, n \in \mathbb{N}$ and all $x_1, \cdots, x_m, y_1, \cdots, y_n \in M$, we have*

$$(x_1 \dot{+} \cdots \dot{+} x_m) \dot{+} (y_1 \dot{+} \cdots \dot{+} y_n) = x_1 \dot{+} \cdots \dot{+} x_m \dot{+} y_1 \dot{+} \cdots \dot{+} y_n.$$

*(iii) For all $n \in \mathbb{N}$ and all $x \in M$, we have*

$$\underbrace{\frac{x}{2^n} \dot{+} \cdots \dot{+} \frac{x}{2^n}}_{2^n \text{ times}} = x.$$

*(iv) For all $m, n \in \mathbb{N}$ and all $x_1, \cdots, x_n \in M$, we have*

$$\frac{1}{2^m}(x_1 \vee \cdots \vee x_n) = \frac{x_1}{2^m} \vee \cdots \vee \frac{x_n}{2^m}.$$

*Proof.* $(i)$: Use induction on $n$ and (H5).

$(ii)$: Use induction on $n$ and (P3).

$(iii)$: Use induction on $n$, $(ii)$, and (H2).

$(iv)$: Use induction on $n$ and (H3). □

**4.1.5 Lemma.** *Let $\mathcal{M} \models \mathsf{RV}$ and let $x, y, z \in M$ be such that $x \wedge y = y \wedge z = z \wedge x = 0$. Then for all $n \in \mathbb{N}$, all $x_0, x_1, \cdots, x_n \in M$ such that $x_i \wedge x_j = \mathbf{0}$ when $0 \leq i \neq j \leq n$, and all $r, s, t, r_0, r_1, \cdots, r_n \in \mathbb{D}$, we have:*

*(i)* $rx \wedge sy = \mathbf{0}$.

*(ii)* $rx \dot{-} sy = rx$.

*(iii)* $rx \dot{+} sy = rx \vee sy$.

*(iv)* $(r_0x_0 \dot{+} \cdots \dot{+} r_{n-1}x_{n-1}) \wedge r_nx_n = \mathbf{0}$.

*(v)* $r_1x_1 \dot{+} \cdots \dot{+} r_nx_n = r_1x_1 \vee \cdots \vee r_nx_n$.

*(vi)* $I(r_1x_1 \dot{+} \cdots \dot{+} r_nx_n) = I(r_1x_1) + \cdots + I(r_nx_n)$.

*Proof.* $(i)$: By (H4), Fact 4.1.2$(i)$, and induction, we have $\frac{x}{2^n} \wedge y = \mathbf{0}$ for all $n \in \mathbb{N}$. Then for all $n_1, n_2 \in \mathbb{N}$, by (S2), Fact 4.1.2$(i, xi)$, and (C), we have

$$\mathbf{0} = ((\frac{x}{2^{n_1}} \dot{+} \frac{x}{2^{n_2}}) \wedge y) \dot{-} ((\frac{x}{2^{n_1}} \wedge y) \dot{+} (\frac{x}{2^{n_2}} \wedge y)) = ((\frac{x}{2^{n_1}} \dot{+} \frac{x}{2^{n_2}}) \wedge y) \dot{-} \mathbf{0} = (\frac{x}{2^{n_1}} \dot{+} \frac{x}{2^{n_2}}) \wedge y.$$

By induction, we have for all $k \in \mathbb{N}$ and all $n_1, \cdots, n_k \in \mathbb{N}$,

$$(\frac{x}{2^{n_1}} \dot{+} \cdots \dot{+} \frac{x}{2^{n_k}}) \wedge y = \mathbf{0}.$$

Hence for all $r \in \mathbb{D}$, we have $rx \wedge y = \mathbf{0}$. By (L1), we have $y \wedge rx = \mathbf{0}$. Then by what we just proved, we have $sy \wedge rx = \mathbf{0}$, and thus by (L1), $rx \wedge sy = \mathbf{0}$, for all $s \in \mathbb{D}$.

$(ii)$: By $(i)$, for all $r, s \in \mathbb{D}$, we have $rx \wedge sy = \mathbf{0}$; *i.e.*, $rx \dot{-} (rx \dot{-} sy) = \mathbf{0}$. Then by (ADD) and Fact 4.1.2$(i)$, we have $I(rx) = I(rx \dot{-} sy) + I(rx \wedge sy) = I(rx \dot{-} sy)$. Then by (MET), the fact that $rx \dot{-} (rx \dot{-} sy) = \mathbf{0}$,

(ADD), (L1), (C), and the fact that $I(rx) = I(rx \dot{-} sy)$, we have

$$\begin{aligned} d(rx \dot{-} sy, rx) &= I\big((rx \dot{-} sy) \dot{-} rx\big) + I\big(rx \dot{-} (rx \dot{-} sy)\big) = I\big((rx \dot{-} sy) \dot{-} rx\big) \\ &= I(rx \dot{-} sy) - I\big((rx \dot{-} sy) \wedge rx\big) = I(rx \dot{-} sy) - I\big(rx \wedge (rx \dot{-} sy)\big) \\ &= I(rx \dot{-} sy) - I\big(rx \dot{-} (rx \dot{-} (rx \dot{-} sy))\big) = I(rx \dot{-} sy) - I\big(rx \dot{-} \mathbf{0}\big) \\ &= I(rx) - I(rx) = 0. \end{aligned}$$

Hence $rx \dot{-} sy = rx$.

$(iii)$: Note that $rx \dot{+} sy = \neg(\neg(rx) \dot{-} sy)$ and $rx \vee sy = \neg(\neg(rx) \wedge \neg(sy))$. To show $rx \dot{+} sy = rx \vee sy$, by Fact 4.1.2$(iii)$ it suffices to show $\neg(rx) \dot{-} sy = \neg(rx) \wedge \neg(sy)$. By (N) and $(ii)$, we have $\neg(rx) \wedge \neg(sy) = \neg(rx) \dot{-} (\neg(rx) \dot{-} \neg(sy)) = \neg(rx) \dot{-} (sy \dot{-} rx) = \neg(rx) \dot{-} sy$.

$(iv)$: We use induction on $n$. When $n = 1$, this follows from $(i)$. When $n = 2$, by $(iii)$, (L1), (L6), $(i)$, and Fact 4.1.2$(v)$ we have

$$(r_0x_0 \dot{+} r_1x_1) \wedge r_2x_2 = (r_0x_0 \vee r_1x_1) \wedge r_2x_2 = (r_0x_0 \wedge r_2x_2) \vee (r_1x_1 \wedge r_2x_2) = \mathbf{0} \vee \mathbf{0} = \mathbf{0}.$$

Suppose the conclusion holds for all $2 \leq m < n$. By the induction assumption for $m = n - 1$, we have $(r_0x_0 \dot{+} \cdots \dot{+} r_{n-2}x_{n-2}) \wedge x_{n-1} = (r_0x_0 \dot{+} \cdots \dot{+} r_{n-2}x_{n-2}) \wedge x_n = \mathbf{0}$. By assumption, we also have $x_{n-1} \wedge x_n = \mathbf{0}$. Then by the induction assumption for $m = 2$, we have

$$(r_0x_0 \dot{+} \cdots \dot{+} r_{n-1}x_{n-1}) \wedge r_nx_n = \big((r_0x_0 \dot{+} \cdots \dot{+} r_{n-2}x_{n-2}) \dot{+} r_{n-1}x_{n-1}\big) \wedge r_nx_n = \mathbf{0}.$$

$(v)$: We use induction on $n$. When $n = 1$, this is trivial. Suppose the conclusion holds for all $m < n$. By $(iv)$, we have $(r_1x_1 \dot{+} \cdots \dot{+} r_{n-1}x_{n-1}) \wedge r_nx_n = \mathbf{0}$. Then by $(iii)$ and the induction assumption for $m = n-1$, we have

$$(r_1x_1 \dot{+} \cdots \dot{+} r_{n-1}x_{n-1}) \dot{+} r_nx_n = (r_1x_1 \dot{+} \cdots \dot{+} r_{n-1}x_{n-1}) \vee r_nx_n = (r_1x_1 \vee \cdots \vee r_{n-1}x_{n-1}) \vee r_nx_n,$$

whereby $r_1x_1 \dot{+} \cdots \dot{+} r_nx_n = r_1x_1 \vee \cdots \vee r_nx_n$.

$(vi)$: We use induction on $n$. When $n = 1$, this is trivial. When $n = 2$, by $(iii)$, Fact 4.1.2$(vii)$, (L1),

(ADD), (N), and $(ii)$ we have

$$\begin{aligned} I(r_1x_1 \dot{+} r_2x_2) &= I(r_1x_1 \vee r_2x_2) = I\Big(\neg\big(\neg(r_1x_1) \wedge \neg(r_2x_2)\big)\Big) = 1 - I\big(\neg(r_1x_1) \wedge \neg(r_2x_2)\big) \\ &= 1 - I\big(\neg(r_2x_2) \wedge \neg(r_1x_1)\big) = 1 - \big(I(\neg(r_1x_1)) - I(\neg(r_1x_1) \dot{-} \neg(r_2x_2))\big) \\ &= 1 - \big(1 - I(r_1x_1) - I(r_2x_2 \dot{-} r_1x_1)\big) = I(r_1x_1) + I(r_2x_2 \dot{-} r_1x_1) = I(r_1x_1) + I(r_2x_2). \end{aligned}$$

Suppose the conclusion holds for all $2 \leq m < n$. By $(iv)$, we have $(r_1x_1 \dot{+} \cdots \dot{+} r_{n-1}x_{n-1}) \wedge r_nx_n = \mathbf{0}$. By the induction assumption for $m = 2$ and $m = n - 1$, we have

$$\begin{aligned} I(r_1x_1 \dot{+} \cdots \dot{+} r_nx_n) &= I((r_1x_1 \dot{+} \cdots \dot{+} r_{n-1}x_{n-1}) \dot{+} r_nx_n) = I(r_1x_1 \dot{+} \cdots \dot{+} r_{n-1}x_{n-1}) + I(r_nx_n) \\ &= I(r_1x_1) + \cdots + I(r_{n-1}x_{n-1}) + I(r_nx_n). \end{aligned}$$

□

**4.1.6 Lemma.** *Let $\mathcal{M} \models \mathsf{RV}$ and let $D = \{x \in M \mid I(x \wedge \neg x) = 0\}$. Then for all $r, s \in \mathbb{D}$, all $x \in M$, and all $a, a_1, \cdots, a_k \in D$, where $k \in \mathbb{N}$ and $a_i \wedge a_j = \mathbf{0}$ if $i \neq j$, we have*

*(i)* $\frac{rx}{2} = \frac{r}{2}x$.

*(ii)* $ra \dot{+} sa = (r \dot{+} s)a$.

*(iii)* $\neg(ra) = (\neg r)a \dot{+} \neg a$.

*(iv)* $ra \dot{-} sa = (r \dot{-} s)a$.

*(v)* $r(a_1 \dot{+} \cdots \dot{+} a_k) = ra_1 \dot{+} \cdots \dot{+} ra_k$.

*(vi)* $ra \wedge sa = (r \wedge s)a$.

*(vii)* $I(ra) = rI(a)$, *and thus* $ra = \mathbf{0}$ *if and only if* $r = 0$ *or* $a = \mathbf{0}$.

*Proof.* We assume familiarity with Fact 4.1.2.

$(i)$: If $r = 0$ or 1, this is trivial. Suppose $r = \frac{m}{2^n}$, where $m, n \in \mathbb{N}$, $0 < m < 2^n$ and $2 \nmid m$. Then $rx = \underbrace{\frac{x}{2^n} \dot{+} \cdots \dot{+} \frac{x}{2^n}}_{m \text{ times}}$. By (H5), Lemma 4.1.4$(i)$, we have

$$\frac{rx}{2} = \frac{1}{2}(\underbrace{\frac{x}{2^n} \dot{+} \cdots \dot{+} \frac{x}{2^n}}_{m \text{ times}}) = \underbrace{\frac{x}{2^{n+1}} \dot{+} \cdots \dot{+} \frac{x}{2^{n+1}}}_{m \text{ times}} = \frac{m}{2^{n+1}}x = \frac{r}{2}x.$$

($ii$): If $r$ or $s$ is zero, then this follows from Fact 4.1.2($xi$). If $r$ or $s$ is 1, then by Fact 4.1.2($ix$) it suffices to show $a \dot{+} ta = a$ for all $t \in \mathbb{D}$. Since $a \in D$, by Fact 4.1.2($i$) we have $a \wedge \neg a = \mathbf{0}$. By Lemma 4.1.5(i), we have $ta \wedge \neg a = \mathbf{0}$. By Fact 4.1.2($v$), (L1), (S1), Fact 4.1.2($ii$), and Fact 4.1.2($iii$), we have

$$a \dot{+} ta = \neg(\neg a \dot{-} ta) = \neg((\mathbf{0} \vee \neg a) \dot{-} ta) = \neg((\mathbf{0} \dot{-} ta) \vee \neg a) = \neg(\neg a) = a.$$

Next, suppose $r = \frac{m_1}{2^{n_1}}$, $s = \frac{m_2}{2^{n_2}}$, where $n_1, n_2 \in \mathbb{N}$, $0 < m_1 < 2^{n_1}$, $0 < m_2 < 2^{n_2}$, and $n_1 \leq n_2$. By Lemma 4.1.4($iii$), we have

$$\frac{a}{2^{n_1}} = \underbrace{\frac{a}{2^{n_2}} \dot{+} \cdots \dot{+} \frac{a}{2^{n_2}}}_{2^{n_2-n_1} \text{ times}}.$$

Then by Lemma 4.1.4($ii$) and induction, we have

$$ra = \underbrace{\frac{a}{2^{n_1}} \dot{+} \cdots \dot{+} \frac{a}{2^{n_1}}}_{m_1 \text{ times}} = \underbrace{(\underbrace{\frac{a}{2^{n_2}} \dot{+} \cdots \dot{+} \frac{a}{2^{n_2}}}_{2^{n_2-n_1} \text{ times}}) \dot{+} \cdots \dot{+} (\underbrace{\frac{a}{2^{n_2}} \dot{+} \cdots \dot{+} \frac{a}{2^{n_2}}}_{2^{n_2-n_1} \text{ times}})}_{m_1 \text{ times}} = \underbrace{\frac{a}{2^{n_2}} \dot{+} \cdots \dot{+} \frac{a}{2^{n_2}}}_{m_1 2^{n_2-n_1}},$$

and thus

$$ra \dot{+} sa = (\underbrace{\frac{a}{2^{n_2}} \dot{+} \cdots \dot{+} \frac{a}{2^{n_2}}}_{m_1 2^{n_2-n_1}}) \dot{+} (\underbrace{\frac{a}{2^{n_2}} \dot{+} \cdots \dot{+} \frac{a}{2^{n_2}}}_{m_2}) = \underbrace{\frac{a}{2^{n_2}} \dot{+} \cdots \dot{+} \frac{a}{2^{n_2}}}_{m_1 2^{n_2-n_1}+m_2}.$$

Suppose $r \dot{+} s < 1$. If $2 \nmid m_1 2^{n_2-n_1} + m_2$, then

$$\underbrace{\frac{a}{2^{n_2}} \dot{+} \cdots \dot{+} \frac{a}{2^{n_2}}}_{m_1 2^{n_2-n_1}+m_2} = \frac{m_1 2^{n_2-n_1} + m_2}{2^{n_2}} a = (r \dot{+} s)a.$$

Otherwise, say $\frac{m_1 2^{n_2-n_1}+m_2}{2^{n_2}} = \frac{m_3}{2^{n_3}}$, where $n_3 \in \mathbb{N}$ and $0 < m_3 < 2^{n_3}$. Then by Lemma 4.1.4($ii, iii$) and induction, we have

$$\underbrace{\frac{a}{2^{n_2}} \dot{+} \cdots \dot{+} \frac{a}{2^{n_2}}}_{m_1 2^{n_2-n_1}+m_2} = \underbrace{\frac{a}{2^{n_3}} \dot{+} \cdots \dot{+} \frac{a}{2^{n_3}}}_{m_3 \text{ times}} = (r \dot{+} s)a.$$

Hence $ra \dot{+} sa = (r \dot{+} s)a$ if $r \dot{+} s < 1$.

Suppose $r \dot{+} s \geq 1$. By Lemma 4.1.4($ii, iii$) and induction, it suffices to prove $a \dot{+} ta = a$ for all $t \in \mathbb{D}$, which was already showed.

($iii$): Since $a \in D$, by Proposition 4.1.3 we have $a \wedge \neg a = \mathbf{0}$ and $a \vee \neg a = \mathbf{1}$. By Lemma 4.1.5($iii$), (H3)

and $(i)$, we have

$$\frac{(\neg r)a \dot{+} \neg a}{2} = \frac{(\neg r)a \vee \neg a}{2} = \frac{(\neg r)a}{2} \vee \frac{\neg a}{2} = \frac{\neg r}{2}a \vee \frac{\neg a}{2} = \frac{\neg r}{2}a \dot{+} \frac{\neg a}{2}. \tag{4.1}$$

Then by $(i)$, Equation (4.1), (P3), $(ii)$, Lemma 4.1.5$(iii)$, (H3), and the fact that $a \vee \neg a = \mathbf{1}$, we have

$$\begin{aligned}\frac{ra}{2} \dot{+} \frac{(\neg r)a \dot{+} \neg a}{2} &= \frac{r}{2}a \dot{+} (\frac{\neg r}{2}a \dot{+} \frac{\neg a}{2}) = (\frac{r}{2}a \dot{+} \frac{\neg r}{2}a) \dot{+} \frac{\neg a}{2} \\ &= (\frac{r}{2} \dot{+} \frac{\neg r}{2})a \dot{+} \frac{\neg a}{2} = \frac{1}{2}a \dot{+} \frac{\neg a}{2} = \frac{a}{2} \vee \frac{\neg a}{2} \\ &= \frac{1}{2}(a \vee \neg a) = \frac{\mathbf{1}}{2}.\end{aligned}$$

By Fact 4.1.2$(xii)$, we have $\neg(ra) = (\neg r)a \dot{+} \neg a$.

$(iv)$: Note that $\neg(t_1 \dot{-} t_2) = \neg t_1 \dot{+} t_2$ for all $t_1, t_2 \in [0,1]$. Then by $(iii)$ and $(ii)$, we have

$$\neg\big((r \dot{-} s)a\big) = (\neg r \dot{+} s)a \dot{+} \neg a = ((\neg r)a \dot{+} sa) \dot{+} \neg a.$$

By Fact 4.1.2$(ix)$, (P3), $(iii)$, and Fact 4.1.2$(iii)$, we have

$$((\neg r)a \dot{+} sa) \dot{+} \neg a = ((\neg r)a \dot{+} \neg a) \dot{+} sa = \neg(ra) \dot{+} sa = \neg(\neg\neg(ra) \dot{-} sa) = \neg(ra \dot{-} sa),$$

and thus $\neg\big((r \dot{-} s)a\big) = \neg(ra \dot{-} sa)$. Hence, by Fact 4.1.2$(iii)$, $ra \dot{-} sa = (r \dot{-} s)a$.

$(v)$: If $r = 0$ or 1, this is trivial. Suppose $r = \frac{m}{2^n}$, where $n \in \mathbb{N}$, $0 < m < 2^n$, and $2 \nmid m$. Then by Lemma 4.1.5$(v)$, Lemma 4.1.4$(iv)$, (P3), Fact 4.1.2$(ix)$, and induction, we have

$$\begin{aligned}&r(a_1 \dot{+} \cdots \dot{+} a_k) = \underbrace{\frac{1}{2^n}(a_1 \dot{+} \cdots \dot{+} a_k) \dot{+} \cdots \dot{+} \frac{1}{2^n}(a_1 \dot{+} \cdots \dot{+} a_k)}_{m \text{ times}} \\ &= \underbrace{\frac{1}{2^n}(a_1 \vee \cdots \vee a_k) \dot{+} \cdots \dot{+} \frac{1}{2^n}(a_1 \vee \cdots \vee a_k)}_{m \text{ times}} = \underbrace{(\frac{1}{2^n}a_1 \vee \cdots \vee \frac{1}{2^n}a_k) \dot{+} \cdots \dot{+} (\frac{1}{2^n}a_1 \vee \cdots \vee \frac{1}{2^n}a_k)}_{m \text{ times}} \\ &= \underbrace{(\frac{1}{2^n}a_1 \dot{+} \cdots \dot{+} \frac{1}{2^n}a_k) \dot{+} \cdots \dot{+} (\frac{1}{2^n}a_1 \dot{+} \cdots \dot{+} \frac{1}{2^n}a_k)}_{m \text{ times}} = \underbrace{(\frac{1}{2^n}a_1 \dot{+} \cdots \dot{+} \frac{1}{2^n}a_1)}_{m \text{ times}} \dot{+} \cdots \dot{+} \underbrace{(\frac{1}{2^n}a_k \dot{+} \cdots \dot{+} \frac{1}{2^n}a_k)}_{m \text{ times}} \\ &= ra_1 \dot{+} \cdots \dot{+} ra_k.\end{aligned}$$

$(vi)$: Since $ra \wedge sa = ra \dot{-} (ra \dot{-} sa)$, this follows from $(iv)$.

$(vii)$: If $r = 0$ or $r = 1$, this follows directly. Suppose $r = \frac{m}{2^n}$, where $m, n \in \mathbb{N}$, $0 < m < 2^n$, and $2 \nmid m$.

We claim that $I(\frac{k}{2^n}a) = \frac{k}{2^n}I(a)$ for all $1 \leq k \leq 2^n$. We use induction on $k$. When $k = 1$, by Fact 4.1.2$(viii)$ and induction we have $I(\frac{1}{2^n}a) = \frac{I(a)}{2^n}$. Suppose our claim holds for all $k < l \leq 2^n$. By (ADD), $(iv)$, $(vi)$ and the induction assumption for $k = l - 1$, we have

$$\begin{aligned} I(\frac{l}{2^n}a) &= I\big((\frac{l-1}{2^n} \dot{+} \frac{1}{2^n})a\big) = I\big((\frac{l-1}{2^n} \dot{+} \frac{1}{2^n})a \dot{-} \frac{1}{2^n}a\big) + I\big((\frac{l-1}{2^n} \dot{+} \frac{1}{2^n})a \wedge \frac{1}{2^n}a\big) \\ &= I(\frac{l-1}{2^n}a) + I(\frac{1}{2^n}a) = \frac{l-1}{2^n}I(a) + \frac{1}{2^n}I(a) = \frac{l}{2^n}I(a). \end{aligned}$$

Hence, $I(\frac{k}{2^n}a) = \frac{k}{2^n}I(a)$ for all $1 \leq k \leq 2^n$, and thus $I(ra) = rI(a)$. □

**4.1.7 Proposition.** *Let $\mathcal{M}$ be a model of* RV *and let $D = \{x \in M \mid x \wedge \neg x = \mathbf{0}\}$. Let $\mathcal{M}_0 \subseteq \mathcal{M}$ be the smallest $L_{\mathsf{RV}}$-prestructure containing $D$. Then $M_0$ is the set $\{r_1a_1 \dot{+} \cdots \dot{+} r_ka_k \mid k \in \mathbb{N}, r_1, \cdots, r_k \in \mathbb{D}, a_1, \cdots, a_k \in D,$ and $a_i \wedge a_j = \mathbf{0}$ if $i \neq j\}$. Moreover, every nonzero element in $M_0$ has a unique decomposition $r_1a_1 \dot{+} \cdots \dot{+} r_ka_k$, where*

- *$k \in \mathbb{N}$;*
- *$r_1, \cdots, r_k \in \mathbb{D}$ with $0 < r_1 < \cdots < r_k$;*
- *$a_1, \cdots, a_k \in D$ such that $a_i \neq \mathbf{0}$ for each $i$, and $a_i \wedge a_j = \mathbf{0}$ whenever $i \neq j$.*

*Proof.* By Proposition 4.1.3, $(D, \mathbf{0}, \mathbf{1}, \neg, \wedge, \vee)$ is a boolean algebra. Let $S$ denote the set $\{r_1a_1 \dot{+} \cdots \dot{+} r_ka_k \mid k \in \mathbb{N}, r_1, \cdots, r_k \in \mathbb{D}, a_1, \cdots, a_k \in D,$ and $a_i \wedge a_j = \mathbf{0}$ if $i \neq j\}$. Clearly, $S \subseteq M_0$. We will show that $(S, \mathbf{0}, \neg, \dot{-}, \frac{1}{2}, I, d)$ is an $L_{\mathsf{RV}}$-prestructure. Taking $k = 1$, $r_1 = 0$, and $a_1 = \mathbf{0}$ in the definition of membership shows that $\mathbf{0} \in S$. By Fact 4.1.2$(iii)$, for all $x, y \in M$ we have $x \dot{-} y = \neg(\neg x \dot{+} y)$. Hence, we need only show that $S$ is closed under $\neg, \dot{+}$, and $\frac{1}{2}$.

Take $x = r_1a_1 \dot{+} \cdots \dot{+} r_ka_k \in S$, where $r_1, \cdots, r_k \in \mathbb{D}$, $a_1, \cdots, a_k \in D$, and $a_i \wedge a_j = \mathbf{0}$ if $i \neq j$. By Lemma 4.1.5$(v)$, we have $x = r_1a_1 \vee \cdots \vee r_ka_k$. Then by Lemma 4.1.4$(iv)$, Lemma 4.1.6$(i)$, and Lemma 4.1.5$(v)$, we have

$$\frac{x}{2} = \frac{r_1a_1}{2} \vee \cdots \vee \frac{r_ka_k}{2} = \frac{r_1}{2}a_1 \vee \cdots \vee \frac{r_k}{2}a_k = \frac{r_1}{2}a_1 \dot{+} \cdots \dot{+} \frac{r_k}{2}a_k \in S.$$

Hence, $S$ is closed under $\frac{1}{2}$. Let $y = (\neg r_1)a_1 \dot{+} \cdots \dot{+} (\neg r_k)a_k \dot{+} \neg(a_1 \vee \cdots \vee a_k)$. Because $D$ is a boolean algebra and $a_i \wedge a_j = \mathbf{0}$ if $i \neq j$, we know $y \in S$. We will show $y = \neg x$. Similar to the calculation of $\frac{x}{2}$, we have $\frac{y}{2} = \frac{\neg r_1}{2}a_1 \dot{+} \cdots \dot{+} \frac{\neg r_k}{2}a_k \dot{+} \frac{\neg(a_1 \vee \cdots \vee a_k)}{2}$. Then by (P3), Fact 4.1.2$(ix)$, induction, Lemma 4.1.6$(ii)$,

Lemma 4.1.5($v$), Lemma 4.1.4($iv$), and the fact that $(D, \mathbf{0}, \mathbf{1}, \neg, \wedge, \vee)$ is a boolean algebra, we have

$$\begin{aligned}
\frac{x}{2} \dot{+} \frac{y}{2} &= (\frac{r_1}{2}a_1 \dot{+} \cdots \dot{+} \frac{r_k}{2}a_k) \dot{+} (\frac{\neg r_1}{2}a_1 \dot{+} \cdots \dot{+} \frac{\neg r_k}{2}a_k \dot{+} \frac{\neg(a_1 \vee \cdots \vee a_k)}{2}) \\
&= (\frac{r_1}{2} \dot{+} \frac{\neg r_1}{2})a_1 \dot{+} \cdots \dot{+} (\frac{r_k}{2} \dot{+} \frac{\neg r_k}{2})a_k \dot{+} \frac{\neg(a_1 \vee \cdots \vee a_k)}{2} \\
&= \frac{1}{2}a_1 \dot{+} \cdots \dot{+} \frac{1}{2}a_k \dot{+} \frac{\neg(a_1 \vee \cdots \vee a_k)}{2} \\
&= \frac{1}{2}a_1 \vee \cdots \vee \frac{1}{2}a_k \vee \frac{\neg(a_1 \vee \cdots \vee a_k)}{2} \\
&= \frac{1}{2}(a_1 \vee \cdots \vee a_k \vee \neg(a_1 \vee \cdots \vee a_k)) \\
&= \frac{1}{2}\mathbf{1} = \frac{\mathbf{1}}{2}.
\end{aligned}$$

Hence by Fact 4.1.2($xii$), we have $\neg x = y \in S$. That is, $S$ is closed under $\neg$.

Take $x = r_1a_1 \dot{+} \cdots \dot{+} r_ka_k, y = s_1b_1 \dot{+} \cdots \dot{+} s_lb_l \in S$, where $r_1, \cdots, r_k, s_1, \cdots, s_l \in \mathbb{D}$, $a_1, \cdots, a_k, b_1, \cdots, b_l \in D$, $a_i \wedge a_{i'} = \mathbf{0}$ if $1 \leq i \neq i' \leq k$, and $b_j \wedge b_{j'} = \mathbf{0}$ if $1 \leq j \neq j' \leq l$. Let $a_0$ be $\neg(a_1 \vee \cdots \vee a_k)$ and let $b_0$ be $\neg(b_1 \vee \cdots \vee b_l)$. Since $(D, \mathbf{0}, \mathbf{1}, \neg, \wedge, \vee)$ is a boolean algebra, we have that $\{a_0, \cdots, a_k\}$ and $\{b_0, \cdots, b_l\}$ are two partitions of $\mathbf{1}$. Then there is a finer partition $\{c_1, \cdots, c_m\}$ of $\mathbf{1}$ satisfying (a) for all $0 \leq i \leq k$, there is a subset $P(i) \subseteq \{1, \cdots, m\}$ such that $a_i = \bigvee_{p \in P(i)} c_p$; (b) for all $0 \leq j \leq l$, there is a subset $Q(j) \subseteq \{1, \cdots, m\}$ such that $b_j = \bigvee_{q \in Q(j)} c_q$. Then by Lemma 4.1.6($v$), (P3), Fact 4.1.2($ix$), and induction, we may assume that $x = r'_1c_1 \dot{+} \cdots \dot{+} r'_mc_m$ and $y = s'_1c_1 \dot{+} \cdots \dot{+} s'_mc_m$, where $r'_1, \cdots, r'_m, s'_1, \cdots, s'_m \in \mathbb{D}$ (could be 0), $c_1, \cdots, c_m \in D$, and $c_i \wedge c_j = \mathbf{0}$ if $i \neq j$. Then by Lemma 4.1.6, we have $x \dot{+} y = (r'_1 \dot{+} s'_1)c_1 \dot{+} \cdots \dot{+} (r'_m \dot{+} s'_m)c_m \in S$. Thus $S$ is closed under $\dot{+}$. Therefore, $S$ is an $L_{\mathsf{RV}}$-prestructure, and thus $S \supseteq M_0$. Hence, $S = M_0$, the smallest $L_{\mathsf{RV}}$-prestructure containing $D$ in $M$.

Consider a nonzero element $x$ in $M_0$. Then $x$ has a decomposition $x = t_1c_1 \dot{+} \cdots \dot{+} t_kc_k$, where $t_1, \cdots, t_k \in \mathbb{D}$, $c_1, \ldots, c_k \in D$, and $c_i \wedge c_j = \mathbf{0}$ if $i \neq j$. Suppose $k$ is the smallest integer for such decomposition. By (P3), Fact 4.1.2($ix$), and induction, we may reorder those terms such that $x = r_1a_1 \dot{+} \cdots \dot{+} r_ka_k$, where $r_1 \leq r_2 \leq \cdots \leq r_k \in \mathbb{D}$ and $a_1, \cdots, a_k \in D$, and $a_i \wedge a_j = \mathbf{0}$ if $i \neq j$. If $r_i = r_{i+1}$ for some $i$ and $j$, by Lemma 4.1.6($v$) we have $r_ia_i \dot{+} r_{i+1}a_{i+1} = r_i(a_i \dot{+} a_{i+1})$. Then it contradicts the fact that $k$ is the smallest integer for such decomposition. So we have $t_1 < t_2 < \cdots < t_k$. By Fact 4.1.2($xi$), if $r_i = 0$ or $a_i = \mathbf{0}$ for some $i$, we may delete the term $r_ia_i$. Then it also contradicts the fact that $k$ is the smallest integer for such decomposition. So we have $0 < r_1 < \cdots < r_k$ and $a_i \neq \mathbf{0}$ for each $1 \leq i \leq k$. We will show that this decomposition is unique. Suppose $x$ has another decomposition $s_1b_1 \dot{+} \cdots \dot{+} s_lb_l$, where $l \in \mathbb{N}$, $s_1, \cdots, s_l \in \mathbb{D}$ with $0 < s_1 < \cdots < s_l$, $b_1, \cdots, b_l \in D$, $b_j \neq \mathbf{0}$ for each $1 \leq j \leq l$, and $b_j \wedge b_{j'} = \mathbf{0}$ whenever $1 \leq j \neq j' \leq l$. For every $1 \leq i \leq k$, if $a_i \wedge b_j = \mathbf{0}$ for all $1 \leq j \leq l$, then by Lemma 4.1.5, we have

$(r_1a_1 \dot{+} \cdots \dot{+} r_ka_k) \wedge b_j = (r_1a_1 \vee \cdots \vee r_ka_k) \wedge b_j = (r_1a_1 \wedge b_j) \vee \cdots \vee (r_ka_k \wedge b_j) = \mathbf{0}$. But by Lemmas 4.1.5 and 4.1.6, we also have $(s_1b_1 \dot{+} \cdots \dot{+} s_lb_l) \wedge b_j = (s_1b_1 \vee \cdots \vee s_lb_l) \wedge b_j = s_jb_j \wedge b_j = s_jb_j \neq \mathbf{0}$, whereby there is $j$ such that $a_i \wedge b_j \neq \mathbf{0}$. Since $a_{i'} \wedge (a_i \wedge b_j) = \mathbf{0}$ for all $i' \neq i$, by Lemmas 4.1.5 and 4.1.6 we have

$$\begin{aligned}
&(r_1a_1 \dot{+} \cdots \dot{+} r_ka_k) \wedge (a_i \wedge b_j) = (r_1a_1 \vee \cdots \vee r_ka_k) \wedge (a_i \wedge b_j) = r_ia_i \wedge (a_i \wedge b_j) \\
&= r_i((a_i \wedge b_j) \vee (a_i \wedge \neg b_j)) \wedge (a_i \wedge b_j) = r_i((a_i \wedge b_j) \dot{+} (a_i \wedge \neg b_j)) \wedge (a_i \wedge b_j) \\
&= (r_i(a_i \wedge b_j) \dot{+} r_i(a_i \wedge \neg b_j)) \wedge (a_i \wedge b_j) = (r_i(a_i \wedge b_j) \vee r_i(a_i \wedge \neg b_j)) \wedge (a_i \wedge b_j) \\
&= r_i(a_i \wedge b_j) \wedge (a_i \wedge b_j) = r_i(a_i \wedge b_j).
\end{aligned}$$

Similarly, $(s_1b_1 \dot{+} \cdots \dot{+} s_lb_l) \wedge (a_i \wedge b_j) = s_j(a_i \wedge b_j)$. Thus, $r_i(a_i \wedge b_j) = s_j(a_i \wedge b_j)$. By Lemma 4.1.6(*vii*) and $a_i \wedge b_j \neq \mathbf{0}$, we have $r_i = s_j$. Since $s_j \neq s_{j'}$ if $j \neq j'$, we get that there is exactly one $1 \leq j \leq l$ such that $r_i = s_j$. Also, we have $a_i \wedge b_{j'} = \mathbf{0}$ if $j \neq j'$. Therefore, $\{r_1, \cdots, r_k\} \subseteq \{s_1, \cdots, s_l\}$, and thus $k \leq l$. By symmetry, $l \leq k$, whereby $k = l$. Since for every $i$ there is $j$ such that $r_i = s_j$, and $0 < r_1 < \cdots < r_k, 0 < s_1 < \cdots < s_l = s_k$, we also get that $r_1 = s_1, \cdots, r_k = s_k$. Now, we need only show $a_i = b_i$ for every $1 \leq i \leq k$. We already know that $a_i \wedge b_i \neq \mathbf{0}$ but $a_i \wedge b_j = \mathbf{0}$ if $i \neq j$. Then for every $1 \leq i \leq k$, by Lemmas 4.1.5 and 4.1.6 we have $(r_1a_1 \dot{+} \cdots \dot{+} r_ka_k) \wedge b_i = r_ia_i \wedge b_i = r_i((a_i \wedge b_i) \wedge (a_i \wedge \neg b_i)) \wedge b_i = r_i(a_i \wedge b_i)$, also $(r_1b_1 \dot{+} \cdots \dot{+} r_kb_k) \wedge b_i = r_ib_i$, and thus $r_i(a_i \wedge b_i) = r_ib_i$. Similarly, we have $r_i(b_i \wedge a_i) = r_ia_i$. Therefore for every $m \in \mathbb{N}$, we have

$$\underbrace{r_i(a_i \wedge b_i) \dot{+} \cdots \dot{+} r_i(a_i \wedge b_i)}_{m \text{ times}} = \underbrace{r_ib_i \dot{+} \cdots \dot{+} r_ib_i}_{m \text{ times}}.$$

By Lemma 4.1.6, we have $(\underbrace{r_i \dot{+} \cdots \dot{+} r_i}_{m \text{ times}})(a_i \wedge b_i) = (\underbrace{r_i \dot{+} \cdots \dot{+} r_i}_{m \text{ times}})b_i$. Take a large $m$, we get $\underbrace{r_i \dot{+} \cdots \dot{+} r_i}_{m \text{ times}} = 1$, and thus $a_i \wedge b_i = b_i$. Similarly, we get $b_i \wedge a_i = a_i$, whereby $a_i = b_i$. □

**4.1.8 Proposition.** *Let $\mathcal{M}$ be a model of* RV *and let $D = \{x \in M \mid x \wedge \neg x = \mathbf{0}\}$. Let $\mathcal{M}_0 \subseteq \mathcal{M}$ be the smallest $L_{\mathsf{RV}}$-prestructure containing $D$. Then $M_0$ is the set $\{r_1a_1 \dot{+} \cdots \dot{+} r_na_n \mid n \in \mathbb{N}, r_1, \cdots, r_n \in \mathbb{D}, a_1, \cdots, a_n \in D$, and $a_i \wedge a_j = \mathbf{0}$ if $i \neq j\}$. Moreover, every $L_{\mathsf{Pr}}$-isomorphism $\phi\colon D \to L^1\big((\Omega, \mathcal{F}, \mu), \{0,1\}\big)$, where $(\Omega, \mathcal{F}, \mu)$ is a probability space, can be uniquely extended to an $L_{\mathsf{RV}}$-embedding $\Phi\colon M_0 \to L^1\big((\Omega, \mathcal{F}, \mu), \mathbb{D}\big)$ which is defined by*

$$\Phi(r_1a_1 \dot{+} \cdots \dot{+} r_na_n) = r_1\phi(a_1) \dot{+} \cdots \dot{+} r_n\phi(a_n),$$

*where $n \in \mathbb{N}$, $r_i \in \mathbb{D}$ and $a_i \in D$ for all $1 \leq i \leq n$, and $a_i \wedge a_j = \mathbf{0}$ for all $1 \leq i \neq j \leq n$.*

*Proof.* Suppose $\phi$ can be extended to an $L_{\mathsf{RV}}$-embedding $\Phi\colon M_0 \to L^1\big((\Omega,\mathcal{F},\mu),\mathbb{D}\big)$. Then

$$\Phi(r_1a_1 \dotplus \cdots \dotplus r_na_n) = r_1\Phi(a_1) \dotplus \cdots \dotplus r_n\Phi(a_n),$$

where $n \in \mathbb{N}$, $r_i \in \mathbb{D}$ and $a_i \in D$ for all $1 \leq i \leq n$, and $a_i \wedge a_j = \mathbf{0}$ for all $1 \leq i \neq j \leq n$. Since $\Phi$ is an extension of $\phi$, we have

$$\Phi(r_1a_1 \dotplus \cdots \dotplus r_na_n) = r_1\phi(a_1) \dotplus \cdots \dotplus r_n\phi(a_n).$$

Hence such an extension $\Phi$ is uniquely determined by $\phi$.

Let $D = \{x \in M \mid I(x \wedge \neg x)\} = 0$. By Proposition 4.1.3, we know that $(D, \mathbf{0}, \neg, \cdot^{\complement}, \cap, \cup, \mu)$ is a model of $\mathsf{Pr}$. By Theorem 3.1.1, we know that there is a probability space $(\Omega,\mathcal{F},\mu)$ such that $D$ as an $L_{\mathsf{Pr}}$-structure is isomorphic to $\widehat{\mathcal{F}}$. Let $\mathcal{N} = \big(L^1((\Omega,\mathcal{F},\mu),[0,1]), \mathbf{0}, \neg, \dot{-}, \frac{1}{2}, I, d\big)$. By Proposition 4.1.1, we have $\mathcal{N} \models \mathsf{RV}$. Let $\mathcal{X}$ denote $L^1((\Omega,\mathcal{F},\mu),\{0,1\})$. By Proposition 4.1.3, we have $(\mathcal{X}, 0, 1, \neg, \wedge, \vee, \bar{\mu}) \models \mathsf{Pr}$ and it is isomorphic to $\widehat{\mathcal{F}}$. Hence, $D$ is $L_{\mathsf{Pr}}$-isomorphic to $\mathcal{X}$. We call this isomorphism $\phi\colon D \to \mathcal{X}$. Then for all $x, y \in D$, we have that $\phi(x \dot{-} y) = \phi(x) \dot{-} \phi(y)$, $\phi(\mathbf{0}) = \mathbf{0}$, $\phi(\mathbf{1}) = \mathbf{1}$, $\int_\Omega \phi(x)d\mu = I^{\mathcal{N}}\big(\phi(x)\big) = I(x)$, and $d(x,y) = d^{\mathcal{N}}\big(\phi(x),\phi(y)\big)$. Let $\mathcal{M}_0$ be the smallest $L_{\mathsf{RV}}$-prestructure containing $D$. By Proposition 4.1.7, we know that every nonzero element $x \in M_0$ has a unique decomposition of the form $x = r_1a_1 \dotplus \cdots \dotplus r_na_n$, where $n \in \mathbb{N}$, $0 < r_1 < \cdots < r_n \in \mathbb{D}$, $a_1, \cdots, a_n \in D$, $a_i \neq \mathbf{0}$ for each $i$, and $a_i \wedge a_j = \mathbf{0}$ whenever $i \neq j$. We extend $\phi\colon D \to \mathcal{X}$ to a mapping $\Phi\colon M_0 \to N$, by defining

$$\Phi(r_1a_1 \dotplus \cdots \dotplus r_na_n) := r_1\phi(a_1) \dotplus \cdots \dotplus r_n\phi(a_n),$$

where $n \in \mathbb{N}$, $0 < r_1 < \cdots < r_n \in \mathbb{D}$, $a_1, \cdots, a_n \in D$, $a_i \neq \mathbf{0}$ for each $i$, and $a_i \wedge a_j = \mathbf{0}$ whenever $i \neq j$. Clearly, $\Phi$ is uniquely determined by $\phi$.

Next, we will check that $\Phi$ preserves $\mathbf{0}, \mathbf{1}, \neg, \frac{1}{2}, \dotplus, I$ and $d$. We already know that $\Phi(\mathbf{0}) = \mathbf{0}$ and $\Phi(\mathbf{1}) = \mathbf{1}$. To show that $\Phi$ preserves $\neg, \frac{1}{2}, \dotplus$, we need the following claim:

**4.1.9 Claim.** *Take any nonzero $x \in M_0$. Suppose $x$ has the form $r_1a_1 \dotplus \cdots \dotplus r_na_n$, where $n \in \mathbb{N}$, $0 < r_1 < \cdots < r_n \in \mathbb{D}$, $a_1, \cdots, a_n \in D$, $a_i \neq \mathbf{0}$ for each $i$, and $a_i \wedge a_j = \mathbf{0}$ whenever $i \neq j$. Suppose $x$ has another form $s_1b_1 \dotplus \cdots \dotplus s_mb_m$, where $m \in \mathbb{N}$, $s_1, \cdots, s_m \in \mathbb{D}$, $b_1, \cdots, b_m \in D$, and $b_k \wedge b_l = \mathbf{0}$ whenever $k \neq l$. Then*

$$\Phi(x) = r_1\Phi(a_1) + \cdots + r_n\Phi(a_n) = s_1\Phi(b_1) + \cdots + s_m\Phi(b_m).$$

*Proof of Claim 4.1.9*: By definition, we know that $\Phi(x) = r_1\Phi(a_1) \dotplus \cdots \dotplus r_n\Phi(a_n)$. Then because $\Phi\colon D \to \mathcal{X} \subseteq N = L^1(\mu,[0,1])$ is an $L_{\mathsf{Pr}}$-isomorphism, we know that $\Phi(a_i) \wedge \Phi(a_j) = \mathbf{0}$ whenever $i \neq j$.

Hence, $\Phi(x) = r_1\Phi(a_1) \dot{+} \cdots \dot{+} r_n\Phi(a_n) = r_1\Phi(a_1) + \cdots + r_n\Phi(a_n)$.

By (P3), Fact 4.1.2($ix$) and induction, after reordering, we may assume that $s_1 \leq \cdots \leq s_m$. Note that during the reordering process, $s_1\Phi(b_1) + \cdots + s_m\Phi(b_m)$ is not changed; let $C$ denote this sum. Since $x \neq \mathbf{0}$, by Fact 4.1.2($xi$) we know that these $s_k$'s can not all be 0. Then there is a $k \leq m$ such that $s_l = 0$ iff $l < k$. Then by Fact 4.1.2($xi$), we have $x = s_1b_1 \dot{+} \cdots \dot{+} s_mb_m = s_kb_k \dot{+} \cdots \dot{+} s_mb_m$. Also $C = s_1\Phi(b_1) + \cdots + s_m\Phi(b_m) = s_k\Phi(b_k) + \cdots + s_m\Phi(b_m)$. Let $Z \subseteq \{k, \cdots, m\}$ be such that $b_l = \mathbf{0}$ iff $l \in Z$. Let $p = |\{k, \cdots, m\} \setminus Z|$ and let $\{m(1), \cdots, m(p)\} = \{k, \cdots, m\} \setminus Z$, where $0 < m(1) \leq \cdots \leq m(p)$. Then by Lemma 4.1.6($vii$) and Fact 4.1.2($xi$), we have $x = s_kb_k \dot{+} \cdots \dot{+} s_mb_m = s_{m(1)}b_{m(1)} \dot{+} \cdots \dot{+} s_{m(p)}b_{m(p)}$. Also $C = s_k\Phi(b_k) + \cdots + s_m\Phi(b_m) = s_{m(1)}\Phi(b_{m(1)}) + \cdots + s_{m(p)}\Phi(b_{m(p)})$, since $\Phi(\mathbf{0}) = \mathbf{0}$. Let $q = |\{s_{m(1)}, \cdots, s_{m(p)}\}|$ and let $\{t_1, \cdots, t_q\}$ denote the set $\{s_{m(1)}, \cdots, s_{m(p)}\}$, where $0 < t_1 < \cdots < t_q$. For all $1 \leq l \leq q$, let $f(l)$ be the first number $h$ such that $t_l = s_{m(h)}$ and let $g(l)$ be the last number $h$ such that $t_l = s_{m(h)}$. Clearly, $g(l) + 1 = f(l + 1)$ for all $1 \leq l < q$, $f(1) = 1$, and $g(q) = p$. Let $c_l$ denote $b_{m(f(l))} \vee \cdots \vee b_{m(g(l))}$ for all $1 \leq l \leq q$. Clearly, $c_l \neq \mathbf{0}$ for all $1 \leq l \leq q$. By Lemma 4.1.5($v$), we know that $c_l = b_{m(f(l))} \vee \cdots \vee b_{m(g(l))} = b_{m(f(l))} \dot{+} \cdots \dot{+} b_{m(g(l))}$, for all $1 \leq l \leq q$. Also, since $\Phi\colon D \to \mathcal{X} \subseteq N$ is an $L_{Pr}$-isomorphism, we have

$$\Phi(c_l) = \Phi(b_{m(f(l))}) \dot{+} \cdots \dot{+} \Phi(b_{m(g(l))}) = \Phi(b_{m(f(l))}) + \cdots + \Phi(b_{m(g(l))}).$$

Then by Lemma 4.1.6($v$), for all $1 \leq l \leq q$, we have

$$s_{m(f(l))}b_{m(f(l))} \dot{+} \cdots \dot{+} s_{m(g(l))}b_{m(g(l))} = t_lb_{m(f(l))} \dot{+} \cdots \dot{+} t_lb_{m(g(l))} = t_l(b_{m(f(l))} \dot{+} \cdots \dot{+} b_{m(g(l))}) = t_lc_l.$$

Also for all $1 \leq l \leq q$, we have

$$s_{m(f(l))}\Phi(b_{m(f(l))}) + \cdots + s_{m(g(l))}\Phi(b_{m(g(l))}) = t_l(\Phi(b_{m(f(l))}) + \cdots + \Phi(b_{m(g(l))})) = t_l\Phi(c_l).$$

Therefore by (P3) and induction we have

$$\begin{aligned} x &= s_{m(1)}b_{m(1)} \dot{+} \cdots \dot{+} s_{m(p)}b_{m(p)} \\ &= (s_{m(f(1))}b_{m(f(1))} \dot{+} \cdots \dot{+} s_{m(g(1))}b_{m(g(1))}) \dot{+} \cdots \dot{+} (s_{m(f(q))}b_{m(f(q))} \dot{+} \cdots \dot{+} s_{m(g(q))}b_{m(g(q))}) \\ &= t_1c_1 \dot{+} \cdots \dot{+} t_qc_q. \end{aligned}$$

Also

$$\begin{aligned}C &= s_{m(1)}\Phi(b(m_1)) + \cdots + s_{m(p)}\Phi(b(m_p)) \\ &= (s_{m(f(1))}\Phi(b_{m(f(1))}) + \cdots + s_{m(g(1))}\Phi(b_{m(g(1))})) + \cdots + (s_{m(f(q))}\Phi(b_{m(f(q))}) + \cdots + s_{m(g(q))}\Phi(b_{m(g(q))})) \\ &= t_1\Phi(c_1) + \cdots + t_q\Phi(c_q).\end{aligned}$$

Since $D$ is a boolean algebra and $b_k \wedge b_l = \mathbf{0}$ whenever $k \neq l$, we have $c_i \wedge c_j = \mathbf{0}$ whenever $1 \leq i \neq j \leq q$. Hence, we have $x = t_1c_1 \dot{+} \cdots \dot{+} t_qc_q$ such that $0 < t_1 < \cdots < t_l$, $c_i \neq \mathbf{0}$ for all $1 \leq i \leq q$, and $c_i \wedge c_j = \mathbf{0}$ whenever $1 \leq i \neq j \leq q$. By Proposition 4.1.7, every nonzero $x$ has a unique such decomposition. Thus $q = n$, $t_i = r_i$ for all $1 \leq i \leq n$, and $c_i = a_i$ for all $1 \leq i \leq n$. Hence, $C = t_1\Phi(c_1) + \cdots + t_q\Phi(c_q) = r_1\Phi(a_1) + \cdots + r_n\Phi(a_n)$. This completes the proof of Claim 4.1.9.

Next, we will show that $\Phi$ preserves $\frac{1}{2}$, $\neg$, and $\dot{+}$. Take any nonzero $x \in M_0$. Suppose $x$ has the form $r_1a_1 \dot{+} \cdots \dot{+} r_na_n$, where $n \in \mathbb{N}$, $0 < r_1 < \cdots < r_n \in \mathbb{D}$, $a_1, \cdots, a_n \in D$, $a_i \neq \mathbf{0}$ for each $i$, and $a_i \wedge a_j = \mathbf{0}$ whenever $i \neq j$. As shown in the proof of Proposition 4.1.7, we know that $\frac{x}{2} = \frac{r_1}{2}a_1 \dot{+} \cdots \dot{+} \frac{r_n}{2}a_n$, and $\neg x = \neg r_1a_1 \dot{+} \cdots \dot{+} \neg r_na_n \dot{+} \neg(a_1 \vee \cdots \vee a_n)$. By Claim 4.1.9, we have $\Phi(\frac{x}{2}) = \frac{r_1}{2}\Phi(a_1) \dot{+} \cdots \dot{+} \frac{r_n}{2}\Phi(a_n) = \frac{1}{2}(r_1\Phi(a_1) \dot{+} \cdots \dot{+} r_n\Phi(a_n)) = \frac{1}{2}\Phi(x)$; *i.e.*, $\Phi$ preserves $\frac{1}{2}$. Also

$$\Phi(\neg x) = \Phi(\neg r_1a_1 \dot{+} \cdots \dot{+} \neg r_na_n \dot{+} \neg(a_1 \vee \cdots \vee a_n)) = \neg r_1\Phi(a_1) \dot{+} \cdots \dot{+} \neg r_n\Phi(a_n) \dot{+} \Phi(\neg(a_1 \vee \cdots \vee a_n)).$$

Then because $\Phi(a_n), \cdots, \Phi(a_1), \Phi(\neg(a_1 \vee \cdots \vee a_n))$ are in $L^1(\mu, [0,1])$, $a_i \wedge a_j = \mathbf{0}$ if $i \neq j$, and $\Phi\colon D \to \mathcal{X}$ is an $L_{\mathsf{P_r}}$-isomorphism, we have that

$$\begin{aligned}&\neg r_1\Phi(a_1) \dot{+} \cdots \dot{+} \neg r_n\Phi(a_n) \dot{+} \Phi(\neg(a_1 \vee \cdots \vee a_n)) \\ &= (1 - r_1)\Phi(a_1) + \cdots + (1 - r_n)\Phi(a_n) + (1 - \Phi(a_1) - \cdots - \Phi(a_n)) \\ &= 1 - r_1\Phi(a_1) - \cdots - r_n\Phi(a_n) = \neg(r_1\Phi(a_1) + \cdots + r_n\Phi(a_n)) \\ &= \neg(r_1\Phi(a_1) \dot{+} \cdots \dot{+} r_n\Phi(a_n)).\end{aligned}$$

Hence, $\Phi(\neg x) = \neg\Phi(x)$; *i.e.*, $\Phi$ preserves $\neg$.

Take any nonzero $x, y \in M_0$. Suppose they have the form $x = r_1a_1 \dot{+} \cdots \dot{+} r_ka_k$ and $y = s_1b_1 \dot{+} \cdots \dot{+} s_lb_l$, where $k, l \in \mathbb{N}$, $0 < r_1 < \cdots < r_k \in \mathbb{D}$, $0 < s_1 < \cdots < s_l \in \mathbb{D}$, $a_1, \cdots, a_k, b_1, \cdots, b_l \in D$, $a_i \neq \mathbf{0}$ for all $1 \leq i \leq k$, $b_j \neq \mathbf{0}$ for all $1 \leq j \leq l$, $a_i \wedge a_{i'} = \mathbf{0}$ whenever $1 \leq i \neq i' \leq k$, and $b_j \wedge b_{j'} = \mathbf{0}$ whenever $1 \leq j \neq j' \leq l$. Let $a_0$ be $\neg(a_1 \vee \cdots \vee a_k)$ and let $b_0$ be $\neg(b_1 \vee \cdots \vee b_l)$. Since $(D, \mathbf{0}, \mathbf{1}, \neg, \wedge, \vee)$ is a boolean

algebra, we have $a_0 \vee a_1 \vee \cdots \vee a_k = b_0 \vee b_1 \vee \cdots \vee b_l = \mathbf{1}$, $a_i \wedge a_{i'} = \mathbf{0}$ for all $0 \leq i \neq i' \leq k$, and $b_j \wedge b_{j'} = \mathbf{0}$ for all $0 \leq j \neq j' \leq l$. That is, $\{a_0, \cdots, a_k\}$ and $\{b_0, \cdots, b_l\}$ are two partitions of $\mathbf{1}$. Then there is a finer partition $\{c_1, \cdots, c_m\}$ of $\mathbf{1}$ satisfying (a) for all $0 \leq i \leq k$, there is a subset $P(i) \subseteq \{1, \cdots, m\}$ such that $a_i = \bigvee_{p \in P(i)} c_p$; (b) for all $0 \leq j \leq l$, there is a subset $Q(j) \subseteq \{1, \cdots, m\}$ such that $b_j = \bigvee_{q \in Q(j)} c_q$. Then by Lemma 4.1.6($v$), (P3), Fact 4.1.2($ix$), and induction, we may assume that $x = r'_1 c_1 \dot{+} \cdots \dot{+} r'_m c_m$ and $y = s'_1 c_1 \dot{+} \cdots \dot{+} s'_m c_m$, where $r'_1, \cdots, r'_m, s'_1, \cdots, s'_m \in \mathbb{D}$ (could be 0), $c_1, \cdots, c_m \in D$, and $c_i \wedge c_j = \mathbf{0}$ if $i \neq j$. Then by Lemma 4.1.6, we have $x \dot{+} y = (r'_1 \dot{+} s'_1)c_1 \dot{+} \cdots \dot{+} (r'_m \dot{+} s'_m)c_m$. Hence

$$\begin{aligned}
\Phi(x \dot{+} y) &= (r'_1 \dot{+} s'_1)\Phi(c_1) + \cdots + (r'_m \dot{+} s'_m)\Phi(c_m) \\
&= (r'_1\Phi(c_1) + \cdots + r'_m) \dot{+} (s'_1\Phi(c_1) + \cdots + s'_m\Phi(c_m)) \\
&= \Phi(x) \dot{+} \Phi(y).
\end{aligned}$$

Thus, $\Phi$ preserves $\dot{+}$.

Now, we will prove $\Phi$ preserves $I$ and $d$. Take a nonzero $x \in M_0$. Suppose $x$ has the form $r_1 a_1 \dot{+} \cdots \dot{+} r_n a_n$, where $n \in \mathbb{N}$, $0 < r_1 < \cdots < r_n \in \mathbb{D}$, $a_1, \cdots, a_n \in D$, $a_i \neq \mathbf{0}$ for each $i$, and $a_i \wedge a_j = \mathbf{0}$ whenever $i \neq j$. Since $\Phi \colon D \to \mathcal{X}$ is an isomorphism, we have $\Phi(a_i) \wedge \Phi(a_j) = \mathbf{0}$ whenever $i \neq j$. For all $f \in N$, we have $I^{\mathcal{N}}(f) = \int_\Omega f d\mu$. Therefore,

$$\begin{aligned}
I^{\mathcal{N}}(\Phi(x)) = \int_\Omega \Phi(x) d\mu &= \int_\Omega \big(r_1\Phi(a_1) \dot{+} \cdots \dot{+} r_n\Phi(a_n)\big) d\mu \\
&= r_1 \int_\Omega \Phi(a_1) d\mu + \cdots + r_n \int_\Omega \Phi(a_n) d\mu \\
&= r_1 I^{\mathcal{N}}(\Phi(a_1)) + \cdots + r_n I^{\mathcal{N}}(\Phi(a_n)) \\
&= r_1 I(a_1) + \cdots + r_n I(a_n).
\end{aligned}$$

By Lemma 4.1.6($vii$), we have $I^{\mathcal{N}}(\Phi(x)) = I(r_1 a_1) + \cdots + I(r_n a_n)$. Then by Lemma 4.1.5, we have $I(r_1 a_1) + \cdots + I(r_n a_n) = I(r_1 a_1 \dot{+} \cdots \dot{+} r_n a_n) = I(x)$, and thus $I(x) = I^{\mathcal{N}}(\Phi(x))$. That is, $\Phi$ preserves $I$. Since $d(x, y) = I(x \dot{-} y) + I(y \dot{-} x)$ for all $x, y \in M$ and $\Phi$ preserves $I$ and $\dot{-}$, it follows that $\Phi$ preserves $d$.

Therefore, $\Phi$ is an $L_{\mathsf{RV}}$-embedding from $\mathcal{M}_0$ to $\mathcal{N}$. □

**4.1.10 Theorem.** *Let $\mathcal{M}$ be a model of* RV. *Then $\mathcal{M}$ is isomorphic to* $\Big(L^1\big((\Omega, \mathcal{F}, \mu), [0,1]\big), \mathbf{0}, \neg, \dot{-}, \frac{1}{2}, I, d\Big)$ *for some probability space $(\Omega, \mathcal{F}, \mu)$.*

*Proof.* Let $D = \{x \in M \mid I(x \wedge \neg x)\} = 0$. By Proposition 4.1.3, we know that $(D, \mathbf{0}, \neg, \cdot^{\complement}, \cap, \cup, \mu)$ is a model of Pr. By Theorem 3.1.1, we know that there is a probability space $(\Omega, \mathcal{F}, \mu)$ such that $D$ as an

$L_{\mathsf{Pr}}$-structure is isomorphic to $\widehat{\mathcal{F}}$. Let $\mathcal{N} = \big(L^1((\Omega, \mathcal{F}, \mu), [0,1]), \mathbf{0}, \neg, \dot{-}, \frac{1}{2}, I, d\big)$. By Proposition 4.1.1, we have $\mathcal{N} \models \mathsf{RV}$. Let $\mathcal{X}$ denote $L^1((\Omega, \mathcal{F}, \mu), \{0,1\})$. By Proposition 4.1.3, we have $(\mathcal{X}, 0, 1, \neg, \wedge, \vee, \bar{\mu}) \models \mathsf{Pr}$ and it is isomorphic to $\widehat{\mathcal{F}}$. Hence, $D$ is $L_{\mathsf{Pr}}$-isomorphic to $\mathcal{X}$. We call this isomorphism $\phi\colon D \to \mathcal{X}$. Then by Proposition 4.1.8, we extend $\phi\colon D \to \mathcal{X}$ to an $L_{\mathsf{RV}}$-embedding $\Phi\colon M_0 \to N$.

Let $(M', d)$ be the completion of $(M_0, d)$ in $M$. Because $\Phi$ is isometric, we know that $\Phi$ is extended uniquely to an embedding $\overline{\Phi}$ from $\mathcal{M}'$ to $\mathcal{N}$. Note that dyadic number valued simple functions are dense in $N$. Hence $\overline{\Phi}$ is a surjective embedding; that is, $\overline{\Phi}$ is an isomorphism between $L_{\mathsf{RV}}$-structures $\mathcal{M}'$ and $\mathcal{N}$. Then we will show $\mathcal{M}'$ is $\mathcal{M}$.

For every $x \in M$ and $n \in \mathbb{N}$, using (APPR), there are elements $y_1, \cdots, y_{2^n} \in M$ such that $\max_i I(y_i \wedge \neg y_i) \leq \frac{1}{2^n}$ and $d(x, \frac{1}{2^n} y_1 \dot{+} \cdots \dot{+} \frac{1}{2^n} y_{2^n}) \leq \frac{1}{2^n}$. For every $1 \leq i \leq 2^n$, since $I(y_i \wedge \neg y_i) = \mathrm{dist}(y_i, D)$, there is $z_i \in D$ such that $d(y_i, z_i) \leq \frac{1}{2^n}$. Then by (P1),

$$\begin{aligned} d(\frac{1}{2^n} y_1 \dot{+} \cdots \dot{+} \frac{1}{2^n} y_{2^n}, \frac{1}{2^n} z_1 \dot{+} \cdots \dot{+} \frac{1}{2^n} z_{2^n}) &\leq d(\frac{1}{2^n} y_1, \frac{1}{2^n} z_1) + \cdots + d(\frac{1}{2^n} y_{2^n}, \frac{1}{2^n} z_{2^n}) \\ &= \frac{1}{2^n}\big(d(y_1, z_1) + \cdots + d(y_{2^n}, z_{2^n})\big) \\ &\leq \max_{1 \leq i \leq 2^n} d(y_i, z_i) \leq \frac{1}{2^n} \end{aligned}$$

where the equality is by Fact 4.1.2 $(ix)$. Then

$$\begin{aligned} d(x, \frac{1}{2^n} z_1 \dot{+} \cdots \dot{+} \frac{1}{2^n} z_{2^n}) &\leq d(x, \frac{1}{2^n} y_1 \dot{+} \cdots \dot{+} \frac{1}{2^n} y_{2^n}) + d(\frac{1}{2^n} y_1 \dot{+} \cdots \dot{+} \frac{1}{2^n} y_{2^n}, \frac{1}{2^n} z_1 \dot{+} \cdots \dot{+} \frac{1}{2^n} z_{2^n}) \\ &\leq \frac{1}{2^n} + \frac{1}{2^n} = \frac{1}{2^{n-1}}. \end{aligned}$$

Since $\frac{1}{2^n} z_1 \dot{+} \cdots \dot{+} \frac{1}{2^n} z_{2^n} \in M_0$, it follows that $M_0$ is dense in $M$, whereby $M' = M$. Therefore $\overline{\Phi}$ is an isomorphism from $\mathcal{M}$ to $\mathcal{N}$. □

Although from the form of its axioms, the theory $\mathsf{RV}$ is not directly a universal theory, it is in fact universal by the following lemma:

**4.1.11 Lemma** ([2, Lemma 2.14])**.** *Let $\mathcal{M} \models \mathsf{RV}$ be of the form $L^1\big((\Omega, \mathcal{F}, \mu), [0,1]\big)$ where $(\Omega, \mathcal{F}, \mu)$ is a probability space. Suppose $\mathcal{N}$ is an $L_{\mathsf{RV}}$-substructure in $\mathcal{M}$. Then $N = L^1\big((\Omega, \sigma(N), \mu), [0,1]\big)$, where $\sigma(N)$ is the $\sigma$-algebra of measurable sets generated by the elements in $N$.*

*Proof.* Obviously, $N \subseteq L^1\big((\Omega, \sigma(N), \mu), [0,1]\big)$. We show that $N$ contains a dense subset of the space $L^1\big((\Omega, \sigma(N), \mu), [0,1]\big)$. Then by the fact that $\mathcal{N}$ is a $L_{\mathsf{RV}}$-structure, we know that it is a complete metric space. Therefore, $N = L^1\big((\Omega, \sigma(N), \mu), [0,1]\big)$.

Let $\mathbb{D}$ denote the dyadic numbers in $[0,1]$. For every $t \in \mathbb{D}$, we have that $t\mathbf{1} \in N$ and note that it is a constant function with value $t$. For all $[f]_{a.s.} \in N$ and all $t \in \mathbb{D}$, we have $g_t = [f]_{a.s.} \dot{-} t\mathbf{1} \in N$. Let $g_t(m)$ denote $\underbrace{g_t \dot{+} \cdots \dot{+} g_t}_{m}$. Note that

$$\lim_{m\to\infty} g_t(m) = [\chi_{\{f>t\}}]_{a.s.}$$

where the limit is the almost sure pointwise limit of random variables. By the dominated convergence theorem, the limit of $(g_t(m))_{m\in\mathbb{N}}$ in the sense of the metric of $M$ (*i.e.*, $L^1$-metric) is $[\chi_{\{f>t\}}]_{a.s.}$. Since $\mathcal{N}$ is an $L_{\mathsf{RV}}$-structure (and thus its underlying space is a complete metric space), we have that $g_t(m) \in N$ and $\lim_{m\to\infty} g_t(m) = [\chi_{\{f>t\}}]_{a.s.} \in N$. Let $A = \big\{\{f > t\} \mid [f]_{a.s.} \in N, t \in \mathbb{D}\big\}$. Note that $\sigma(N) = \sigma(A)$ and for every $B \in \sigma(A)$, we have $[\chi_B]_{a.s.} \in N$. Hence, every function in $L^1\big((\Omega, \sigma(A), \mu), \mathbb{D}\big)$ is in $N$. Since the set $L^1\big((\Omega, \sigma(A), \mu), \mathbb{D}\big)$ is a dense subset in $L^1\big((\Omega, \sigma(N), \mu), [0,1]\big)$, we have $N = L^1\big((\Omega, \sigma(N), \mu), [0,1]\big)$. □

**4.1.12 Corollary.** *The theory* RV *is a universal theory.*

*Proof.* Use Lemma 4.1.11 and Theorem 1.7.6. □

**4.1.13 Corollary.** *Let $\mathcal{M}$ be an $L_{\mathsf{RV}}$-structure. Then $\mathcal{M}$ is a model of* ARV *if and only if $\mathcal{M}$ is isomorphic to $\Big(L^1\big((\Omega, \mathcal{F}, \mu), [0,1]\big), \mathbf{0}, \neg, \dot{-}, \frac{1}{2}, I, d\Big)$ for some atomless probability space $(\Omega, \mathcal{F}, \mu)$.*

*Proof.* By Proposition 4.1.1, we know that $\Big(L^1\big((\Omega, \mathcal{F}, \mu), [0,1]\big), \mathbf{0}, \neg, \dot{-}, \frac{1}{2}, I, d\Big)$, where $(\Omega, \mathcal{F}, \mu)$ is an atomless probability space, is a model of ARV.

For the other direction, by Theorem 4.1.10, $\mathcal{M}$ as a model of RV is isomorphic to $\Big(L^1\big((\Omega, \mathcal{F}, \mu), [0,1]\big), \mathbf{0}, \neg, \dot{-}, \frac{1}{2}, I, d\Big)$ for a probability space $(\Omega, \mathcal{F}, \mu)$. Let $D$ denote $L^1\big((\Omega, \mathcal{F}, \mu), \{0,1\}\big)$. By Proposition 4.1.3, $D$ is a model of $Pr$ and $D$ is $L_{\mathsf{Pr}}$-isomorphic to $\widehat{\mathcal{F}}$. For all $x \in D$, using (NA) we know that for every $\epsilon > 0$, there is $y_\epsilon$ such that $\mathrm{dist}(y_\epsilon, D) \leq \epsilon$ and $|I(y_\epsilon \wedge x) - \frac{I(x)}{2}| \leq \epsilon$. Then there is $y \in D$ such that $|I(y \wedge x) - \frac{I(x)}{2}| \leq 2\epsilon$. Thus $D$ as a $L_{\mathsf{Pr}}$-structure satisfies Axiom $(iv)$ in APr, whereby $D$ is a model of APr. By Theorem 3.1.2, $(\Omega, \mathcal{F}, \mu)$ is atomless. □

## 4.2 Ben Yaacov's axiomatizations

The theory of the class $\mathcal{RV}$ was first studied by Ben Yaacov in [2]. He axiomatized the class via propositional continuous logic, after Rose, Rosser and Church. However, the axiomatizations given in the preceding section only use basic measure theoretic probability theory. We follow the notation from the last section. Ben Yaacov's axioms [2] for the theory of the class $\mathcal{RV}$ consist of the following conditions; we let RV(BY) denote

the set of these axioms to distinguish them from our axioms given in the preceding section. (Ben Yaacov denoted his axioms in [2] by RV.)

(RV1) $\sup_x \sup_y \frac{1}{2}\big|I(x) - \big(I(x \dot{-} y) + I(y \wedge x)\big)\big| = 0$

(RV2) $|I(\mathbf{1}) - 1| = 0$

(RV3) $\sup_x \sup_y \frac{1}{2}\big|d(x,y) - \big(I(x \dot{-} y) + I(y \dot{-} x)\big)\big| = 0$

(RV4.1) $\sup_x \sup_y d\big((x \dot{-} y) \dot{-} x, \mathbf{0}\big) = 0$

(RV4.2) $\sup_x \sup_y \sup_z d\Big(\big((x \dot{-} z) \dot{-} (x \dot{-} y)\big) \dot{-} (y \dot{-} z), \mathbf{0}\Big) = 0$

(RV4.3) $\sup_x \sup_y d\big((x \wedge y) \dot{-} (y \wedge x), \mathbf{0}\big) = 0$

(RV4.4) $\sup_x \sup_y d\big((x \dot{-} y) \dot{-} (\neg y \dot{-} \neg x), \mathbf{0}\big) = 0$

(RV4.5) $\sup_x d\big(\frac{x}{2} \dot{-} (x \dot{-} \frac{x}{2}), \mathbf{0}\big) = 0$

(RV4.6) $\sup_x d\big((x \dot{-} \frac{x}{2}) \dot{-} \frac{x}{2}, \mathbf{0}\big) = 0$

From the form of its axioms, the theory RV(BY) is universal. The theory ARV(BY) is obtained by adding the following axiom to RV(BY):

(NA) $\sup_x \inf_y \max(I(y \wedge \neg y), \big|I(y \wedge x) - \frac{I(x)}{2}\big|) = 0$

Ben Yaacov proved the following result:

**4.2.1 Proposition** ([2])**.** *Let $\mathcal{M}$ be an $L_{\mathsf{RV}}$-structure. Then $\mathcal{M}$ is a model of* RV(BY) *if and only if $\mathcal{M}$ is isomorphic to $\Big(L^1\big((\Omega, \mathcal{F}, \mu), [0,1]\big), \mathbf{0}, \neg, \dot{-}, \frac{1}{2}, I, d\Big)$ for some probability space $(\Omega, \mathcal{F}, \mu)$. Moreover, $\mathcal{M}$ is a model of* ARV(BY) *if and only if $(\Omega, \mathcal{F}, \mu)$ is an atomless probability space.*

*Proof.* This is [2, Corollary 2.8]. □

**4.2.2 Corollary.** *The theory* RV *(resp.* ARV*) is equivalent to the theory* RV(BY) *(resp.* ARV(BY)*).*

*Proof.* Use Theorem 4.1.10, Corollary 4.1.13 and Proposition 4.2.1. □

## 4.3 Extension of the theory by multiplication

Multiplication between random variables is a basic and useful operation. Thus it is convenient if we have the multiplication function in the signature. Let $L_{\mathsf{RV}(\times)}$ denote the signature consisting of $L_{\mathsf{RV}}$ and a new binary function symbol "$\times$", whose modulus of uniform continuity is $\Delta_\times(\epsilon) := \frac{\epsilon}{2}$ for every $\epsilon \in (0, 1]$. We often omit the symbol "$\times$" when no confusion will occur; *e.g.*, we often write $xy$ instead of $x \times y$. Every model of $\mathsf{RV}$ has a natural expansion to an $L_{\mathsf{RV}(\times)}$-structure, obtained by interpreting "$\times$" as multiplication between random variables. Ben Yaacov showed [2, Lemma 2.13] that the multiplication function is uniformly definable in every model of his theory $\mathsf{RV}(\mathsf{BY})$, and hence the same is true in every model of our theory $\mathsf{RV}$. By Corollary 1.6.9, the class of expansions discussed above is axiomatizable, and its theory is an extension by definitions of the $L_{\mathsf{RV}}$-theory $\mathsf{RV}$.

In this section, we give an explicit set of axioms, denoted by $\mathsf{RV}(\times)$, for the theory of this class of expansions. The $L_{\mathsf{RV}(\times)}$-theory $\mathsf{RV}(\times)$ consist of the axioms in $\mathsf{RV}$ together with the following axioms:

(M1) $\sup_x \sup_y \sup_z d\big((x \dot{-} y)z, xz \dot{-} yz\big) = 0$

(M2) $\sup_x d(\frac{x}{2}, \frac{\mathbf{1}}{2} \times x) = 0$; $\quad \sup_x d(\mathbf{1} \times x, x) = 0$

(M3) $\sup_x \sup_y d\big(xy \dot{-} (x \wedge y), \mathbf{0}\big) = 0$

(M4) $\sup_x \sup_y d(xy, yx) = 0$

(M5) $\sup_x \sup_y \sup_z d\big((x \vee y)z, xz \vee yz\big) = 0$

(M6) $\sup_x \sup_y \sup_z d\big((x \times y) \times z, x \times (y \times z)\big) = 0$

Throughout this section, we will interpret symbols of $L_{\mathsf{RV}(\times)}$ in a given model $\mathcal{M}$ of $\mathsf{RV}(\times)$ without putting $\mathcal{M}$ explicitly into the notations, for easier readability. Also note that $\mathsf{RV}(\times) \supseteq \mathsf{RV}$, so for a statement without "$\times$", Theorem 4.1.10 provides a powerful tool to justify whether the statement holds or not.

**4.3.1 Fact.** *Let $\mathcal{M}$ be a model of $\mathsf{RV}(\times)$. For all $x, y, z \in M$, we have the following properties:*

*(i)* $\mathbf{0} \times x = x \times \mathbf{0} = \mathbf{0}$.

*(ii)* $xy \dot{-} x = xy \dot{-} y = \mathbf{0}$.

*(iii)* $(x \wedge y)z = xz \wedge yz$.

*(iv)* $x \vee xy = x$ *and* $x \wedge xy = xy$.

*Proof.* $(i)$ By Fact 4.1.2$(vi)$, we have $x \dot{-} x = \mathbf{0}$. By (M1), we have $\mathbf{0} \times x = (x \dot{-} x) \times x = xx \dot{-} xx$, which is $\mathbf{0}$ by Fact 4.1.2$(vi)$. Then by (M4), we have $x \times \mathbf{0} = \mathbf{0}$.

$(ii)$ By Fact 4.1.2$(ii)$, we have $x \dot{-} \mathbf{1} = y \dot{-} \mathbf{1} = \mathbf{0}$. Then by (M1) and $(i)$, we have $xy \dot{-} \mathbf{1}y = xy \dot{-} x\mathbf{1} = \mathbf{0}$. By (M2) and (M4), we have $\mathbf{1}y = y$ and $x\mathbf{1} = \mathbf{1}x = x$, and thus $xy \dot{-} y = xy \dot{-} x = \mathbf{0}$.

$(iii)$ By (M1), we have $(x \wedge y)z = \big(x \dot{-} (x \dot{-} y)\big)z = xz \dot{-} (xz \dot{-} yz) = xz \wedge yz$.

$(iv)$ By (M2), (M4), Fact 4.1.2$(v)$ and (M5), we have $x = \mathbf{1}x = x\mathbf{1} = x(\mathbf{1} \vee y) = x\mathbf{1} \vee xy = x \vee xy$.

By Fact 4.1.2$(iv)$, (L1), $(iii)$, (M4), and (M2), we have

$$xy = x(y \wedge \mathbf{1}) = x(\mathbf{1} \wedge y) = x\mathbf{1} \wedge xy = \mathbf{1}x \wedge xy = x \wedge xy.$$

□

**4.3.2 Lemma.** *Let $\mathcal{M}$ be a model of $\mathsf{RV}(\times)$. For every $x \in M$, the following are equivalent:*

*(i)* $xx = x$.

*(ii)* $x \times \neg x = \mathbf{0}$.

*(iii)* $x \wedge \neg x = \mathbf{0}$.

*(iv)* $x \vee \neg x = \mathbf{1}$.

*Proof.* $(iii) \Leftrightarrow (iv)$: Since "$\times$" is not involved in the statements and $(iii) \Leftrightarrow (iv)$ holds in models of $\mathsf{RV}$, we know that it also holds in all models of $\mathsf{RV}(\times)$.

$(i) \Rightarrow (ii)$: By (MET), $xx = x$ implies $x \dot{-} xx = \mathbf{0}$. Hence by Fact 4.1.2$(iii)$ and (M1), we have $x \times \neg x = x(\mathbf{1} \dot{-} x) = x \dot{-} xx = \mathbf{0}$.

$(ii) \Rightarrow (iii)$: By Fact 4.3.1$(ii)$ and Fact 4.1.2$(iv)$, we have $x(x \wedge \neg x) = xx \wedge (x \times \neg x) = xx \wedge \mathbf{0} = \mathbf{0}$. By (M4), $\neg x \times x = x \times \neg x = \mathbf{0}$. Then by Fact 4.3.1$(ii)$ and Fact 4.1.2$(iv)$, we have $\neg x(x \wedge \neg x) = (\neg x \times x) \wedge (\neg x)(\neg x) = \mathbf{0} \wedge (\neg x)(\neg x) = \mathbf{0}$. Hence, $x(x \wedge \neg x) = \neg x(x \wedge \neg x) = \mathbf{0}$. Then by (M5) and Fact 4.1.2$(v)$, we have $(x \vee \neg x) \times (x \wedge \neg x) = x(x \wedge \neg x) \vee \neg x(x \wedge \neg x) = \mathbf{0} \vee \mathbf{0} = \mathbf{0}$. Note that $\mathsf{RV} \models d(\frac{\mathbf{1}}{2} \dot{-} (x \vee \neg x), \mathbf{0}) = 0$. Since $\mathsf{RV}(\times) \models \mathsf{RV}$, we have $\frac{\mathbf{1}}{2} \dot{-} (x \vee \neg x) = \mathbf{0}$. Multiplying $\frac{\mathbf{1}}{2} \dot{-} (x \vee \neg x) = \mathbf{0}$ by $(x \wedge \neg x)$ from the right hand side, by (M1) we have $\frac{\mathbf{1}}{2} \times (x \wedge \neg x) \dot{-} (x \vee \neg x) \times (x \wedge \neg x) = \mathbf{0}$, whence $\frac{\mathbf{1}}{2} \times (x \wedge \neg x) = \mathbf{0}$ by Fact 4.3.1$(i)$ and (C). Then by (H2), (M2) and Fact 4.1.2$(xi)$, we have that

$$x \wedge \neg x = \frac{1}{2}(x \wedge \neg x) \dotplus \frac{1}{2}(x \wedge \neg x) = \frac{\mathbf{1}}{2} \times (x \wedge \neg x) \dotplus \frac{\mathbf{1}}{2} \times (x \wedge \neg x) = \mathbf{0} \dotplus \mathbf{0} = \mathbf{0}.$$

$(iii) \Rightarrow (i)$: By (M3), we have $(x \times \neg x) \dot{-} (x \wedge \neg x) = \mathbf{0}$, whence $x \times \neg x = \mathbf{0}$ by (C). Then by (M2), (M1), Fact 4.1.2$(iii)$, and (M4), we have $x \dot{-} xx = \mathbf{1}x \dot{-} xx = (\mathbf{1} \dot{-} x)x = \neg x \times x = x \times \neg x = \mathbf{0}$. By Fact 4.3.1$(ii)$, $xx \dot{-} x = \mathbf{0}$. Then by (MET), we have $xx = x$. □

**4.3.3 Remark.** *Let $\mathcal{M}$ be a model of $\mathsf{RV}(\times)$. By Theorem 4.1.10, we may assume that $M$ is of the form $L^1(\mu, [0,1])$ for the probability space $(\Omega, \mathcal{F}, \mu)$. Then the set $\mathcal{X}_M = \{x \in M \mid x \wedge \neg x = \mathbf{0}\}$ is the set of characteristic functions in $M$. Lemma 4.3.2 implies that $\mathcal{X}_M = \{x \in M \mid xx = x\}$.*

**4.3.4 Lemma.** *Let $\mathcal{M}$ be a model of $\mathsf{RV}(\times)$ and let $x, y$ be elements in $M$. If $xx = x$ and $yy = y$, then $x \times y = x \wedge y$.*

*Proof.* By (M4), Fact 4.3.1$(iii)$ and $(iv)$, we have that $x(x \wedge y) = xx \wedge xy = x \wedge xy = xy$. By Lemma 4.3.2, $xx = x$ implies $x \times \neg x = \mathbf{0}$, whence by (M4), $\neg x \times x = \mathbf{0}$. Then by (M4), Fact 4.3.1$(ii)$ and Fact 4.1.2$(iv)$, we have $\neg x(x \wedge y) = (\neg x \times x) \wedge (\neg x \times y) = \mathbf{0} \wedge (\neg x \times y) = \mathbf{0}$. Then by (M5) and Fact 4.1.2$(v)$, we have $(x \vee \neg x) \times (x \wedge y) = x(x \wedge y) \vee \neg x(x \wedge y) = xy \vee \mathbf{0} = xy$. By Lemma 4.3.2, $xx = x$ implies $x \vee \neg x = \mathbf{1}$. Then by (M2), $(x \vee \neg x) \times (x \wedge y) = \mathbf{1} \times (x \wedge y) = x \wedge y$. Hence $xy = x \wedge y$. □

Let $\mathcal{M}$ be a model of $\mathsf{RV}(\times)$. By Theorem 4.1.10, we may assume that $M$ is of the form $L^1(\mu, [0,1])$ for the probability space $(\Omega, \mathcal{F}, \mu)$. Throughout this section, let $\mathcal{X}_M$ denote the set $\{x \in M \mid xx = x\}$. By Lemma 4.3.2, we know that $\mathcal{X}_M$ is the set of characteristic functions in $M$. Lemma 4.3.4 tells us that "$\times$" defines multiplication between characteristic functions in $M$. Next, we will extend this fact to all random variables. For all $x, y \in \mathcal{X}_M$ with $x \times y = \mathbf{0}$, by Lemma 4.3.4 we have $x \times y = x \wedge y = \mathbf{0}$. Hence, $x$ and $y$ correspond two disjoint measurable sets.

**4.3.5 Lemma.** *Let $\mathcal{M}$ be a model of $\mathsf{RV}(\times)$.*

*(i) For all $n \in \mathbb{N}$ and all $x \in M$, we have $\frac{x}{2^n} = \frac{\mathbf{1}}{2^n} \times x$.*

*(ii) If $zz = z$, then for all $x, y \in M$, we have $(x \dot{+} y)z = xz \dot{+} yz$.*

*(iii) If $xx = x$, then for all $r \in \mathbb{D}$, we have $rx = (r\mathbf{1}) \times x$.*

*(iv) For all $r, s \in \mathbb{D}$, we have $r\mathbf{1} \times s\mathbf{1} = (rs)\mathbf{1}$.*

*(v) If $xx = x$ and $yy = y$, then for all $r, s \in \mathbb{D}$, we have $(rx) \times (sy) = (rs)(x \times y)$.*

*Proof.* $(i)$: We use induction on $n$. When $n = 1$, this is (M2). Suppose the conclusion holds for all $m \leq n$. Then by (M2), (M6) and the induction assumption, we have

$$\frac{\mathbf{1}}{2^{n+1}} \times x = (\frac{1}{2}\frac{\mathbf{1}}{2^n}) \times x = (\frac{\mathbf{1}}{2} \times \frac{\mathbf{1}}{2^n}) \times x = \frac{\mathbf{1}}{2} \times (\frac{\mathbf{1}}{2^n} \times x) = \frac{\mathbf{1}}{2} \times \frac{x}{2^n} = \frac{1}{2}\frac{x}{2^n} = \frac{x}{2^{n+1}}.$$

$(ii)$: By Lemma 4.3.2 and (M2), it suffices to show $(z \vee \neg z) \times \big((x \dot{+} y)z\big) = (z \vee \neg z) \times (xz \dot{+} yz)$. Then by (M5), it suffices to show $z \times \big((x \dot{+} y)z\big) = z \times (xz \dot{+} yz)$ and $\neg z \times \big((x \dot{+} y)z\big) = \neg z \times (xz \dot{+} yz)$.

By (M4), (M6), Fact 4.1.2$(iii)$, (M1), and (M2), we have

$$\begin{aligned} z \times \big((x \dot{+} y)z\big) &= \big((x \dot{+} y)z\big) \times z = (x \dot{+} y) \times (zz) = (x \dot{+} y)z = \big(\neg(\neg x \dot{-} y)\big)z \\ &= \big(\mathbf{1} \dot{-} ((\mathbf{1} \dot{-} x) \dot{-} y)\big)z = \mathbf{1}z \dot{-} ((\mathbf{1}z \dot{-} xz) \dot{-} yz) \\ &= z \dot{-} ((z \dot{-} xz) \dot{-} yz). \end{aligned}$$

By Fact 4.1.2$(iii)$, (M4), (M1), (M2), and (M6), we have

$$\begin{aligned} z \times (xz \dot{+} yz) &= z \times \big(\neg(\neg(xz) \dot{-} yz)\big) = \big(\mathbf{1} \dot{-} ((\mathbf{1} \dot{-} xz) \dot{-} yz)\big) \times z \\ &= \mathbf{1}z \dot{-} ((\mathbf{1}z \dot{-} xzz) \dot{-} yzz) \\ &= z \dot{-} ((z \dot{-} xz) \dot{-} yz). \end{aligned}$$

Hence, $z \times \big((x \dot{+} y)z\big) = z \times (xz \dot{+} yz)$.

By Lemma 4.3.2, $zz = z$ implies $z \times \neg z = \mathbf{0}$. Then by (M4), (M6), and Fact 4.3.1$(i)$, we have

$$\neg z \times \big((x \dot{+} y)z\big) = \big((x \dot{+} y)z\big) \times \neg z = (x \dot{+} y) \times (z \times \neg z) = (x \dot{+} y) \times \mathbf{0} = \mathbf{0}.$$

By Fact 4.1.2$(iii)$, (M4), (M1), (M2), (M6), Fact 4.3.1$(i)$, (C), and Fact 4.1.2$(vi)$, we have

$$\begin{aligned} \neg z \times (xz \dot{+} yz) &= \neg z \times \big(\neg(\neg(xz) \dot{-} yz)\big) = \big(\mathbf{1} \dot{-} ((\mathbf{1} \dot{-} xz) \dot{-} yz)\big) \times \neg z \\ &= \mathbf{1}\neg z \dot{-} \big((\mathbf{1}\neg z \dot{-} (xz)\neg z) \dot{-} (yz)\neg z\big) = \neg z \dot{-} \big((\neg z \dot{-} x(z \times \neg z)) \dot{-} y(z \times \neg z)\big) \\ &= \neg z \dot{-} \big((\neg z \dot{-} x \times \mathbf{0}) \dot{-} y \times \mathbf{0}\big) = \neg z \dot{-} ((\neg z \dot{-} \mathbf{0}) \dot{-} \mathbf{0}) = \neg z \dot{-} \neg z = \mathbf{0}. \end{aligned}$$

Hence, $\neg z \times \big((x \dot{+} y)z\big) = \neg z \times (xz \dot{+} yz)$.

$(iii)$: If $r = 0$ or 1, by Fact 4.3.1$(i)$ and (M2) this is trivial. Suppose $r = \frac{m}{2^n}$, where $n \in \mathbb{N}$, $0 < m < 2^n$, and $2 \nmid m$. Then by $(i)$ and $(ii)$, we have

$$rx = \underbrace{\frac{x}{2^n} \dot{+} \cdots \dot{+} \frac{x}{2^n}}_{m \text{ times}} = \underbrace{\frac{\mathbf{1}}{2^n} \times x \dot{+} \cdots \dot{+} \frac{\mathbf{1}}{2^n} \times x}_{m \text{ times}} = (\underbrace{\frac{\mathbf{1}}{2^n} \dot{+} \cdots \dot{+} \frac{\mathbf{1}}{2^n}}_{m \text{ times}})x = (r\mathbf{1}) \times x.$$

$(iv)$: If either $r$ or $s$ is 0 or 1, by Fact 4.3.1$(i)$ and (M2) this is trivial. Suppose $r = \frac{m_1}{2^{n_1}}, s = \frac{m_2}{2^{n_2}}$, where

$n_1, n_2 \in \mathbb{N}$, $0 < m_1 < 2^{n_1}$, $0 < m_2 < 2^{n_2}$, $2 \nmid m_1$, and $2 \nmid m_2$. By $(iii)$ and (H5), we have

$$\begin{aligned}(r\mathbf{1}) \times (s\mathbf{1}) = r(s\mathbf{1}) &= r(\underbrace{\frac{\mathbf{1}}{2^{n_2}} \dot{+} \cdots \dot{+} \frac{\mathbf{1}}{2^{n_2}}}_{m_2 \text{ times}}) \\ &= \underbrace{\frac{1}{2^{n_1}}(\underbrace{\frac{\mathbf{1}}{2^{n_2}} \dot{+} \cdots \dot{+} \frac{\mathbf{1}}{2^{n_2}}}_{m_2 \text{ times}}) \dot{+} \cdots \dot{+} \frac{1}{2^{n_1}}(\underbrace{\frac{\mathbf{1}}{2^{n_2}} \dot{+} \cdots \dot{+} \frac{\mathbf{1}}{2^{n_2}}}_{m_2 \text{ times}})}_{m_1 \text{ times}} \\ &= \underbrace{\frac{\mathbf{1}}{2^{n_1+n_2}} \dot{+} \cdots \dot{+} \frac{\mathbf{1}}{2^{n_1+n_2}}}_{m_1 m_2 \text{ times}} = \frac{m_1 m_2}{2^{n_1+n_2}}\mathbf{1} = (rs)\mathbf{1}.\end{aligned}$$

$(v)$: By (M6) and (M4), we have $(x \times y) \times (x \times y) = (xx) \times (yy) = x \times y$. Then by $(iii)$, (M6), (M4), and $(iv)$, we have

$$(rx) \times (sy) = \big((r\mathbf{1}) \times x\big) \times \big((s\mathbf{1}) \times y\big) = \big((r\mathbf{1}) \times (s\mathbf{1})\big) \times (x \times y) = \big((rs)\mathbf{1}\big) \times (xy) = (rs)(x \times y).$$

□

**4.3.6 Lemma.** *Let $\mathcal{M}$ be a model of $\mathsf{RV}(\times)$ and let $x, y \in M$. If $xx = x$ and $xy = \mathbf{0}$, then $(rx \dot{+} y)z = (rx)z \dot{+} yz$, for all $z \in M$ and all $r \in \mathbb{D}$.*

*Proof.* By Lemma 4.3.2 and (M2), it suffices to show $(x \vee \neg x) \times \big((rx \dot{+} y)z\big) = (x \vee \neg x) \times ((rx)z \dot{+} yz)$. Then by (M5), it suffices to show $x \times \big((rx \dot{+} y)z\big) = x \times ((rx)z \dot{+} yz)$ and $\neg x \times \big((rx \dot{+} yz\big) = \neg x \times ((rx)z \dot{+} yz)$.

First, we will show $x \times \big((rx \dot{+} y)z\big) = x \times ((rx)z \dot{+} yz) = (rx)z$. By Fact 4.1.2$(iii)$, (M1), (M2), and (M6), we have

$$\begin{aligned}x \times \big((rx \dot{+} y)z\big) &= x \times \Big(\big(\neg(\neg(rx) \dot{-} y)\big)z\Big) = x \times \Big(\big(\mathbf{1} \dot{-} ((\mathbf{1} \dot{-} rx) \dot{-} y)\big)z\Big) \\ &= x \times (\mathbf{1}z \dot{-} ((\mathbf{1}z \dot{-} (rx)z) \dot{-} yz)) = x \times (z \dot{-} ((z \dot{-} (rx)z) \dot{-} yz)) \\ &= xz \dot{-} \big((xz \dot{-} x((rx)z)) \dot{-} x(yz)\big) = xz \dot{-} \big((xz \dot{-} (x \times (rx))z) \dot{-} (xy)z\big).\end{aligned}$$

Then by $xx = x$, Lemma 4.3.5$(v)$, $xy = \mathbf{0}$, Fact 4.3.1$(i)$, (C), Lemma 4.3.5$(iii)$, (M6), and Fact 4.3.1$(iv)$,

we have

$$\begin{aligned} xz \dot{-} \big((xz \dot{-} (x \times (rx))z) \dot{-} (xy)z\big) &= xz \dot{-} \big((xz \dot{-} (rx)z) \dot{-} \mathbf{0}z\big) = xz \dot{-} \big((xz \dot{-} (rx)z) \dot{-} \mathbf{0}\big) \\ &= xz \dot{-} \big((xz \dot{-} (rx)z)\big) = (xz) \wedge (rx)z = (xz) \wedge ((r\mathbf{1}) \times x)z \\ &= (xz) \wedge ((r\mathbf{1})(xz)) = (r\mathbf{1})(xz) = (r\mathbf{1} \times x)z = (rx)z. \end{aligned}$$

Thus, $x \times \big((rx \dot{+} y)z\big) = (rx)z$.

Next, let us show $x \times ((rx)z \dot{+} yz) = (rx)z$. By Fact 4.1.2($iii$), (M1), (M4), (M2), and (M6), we have

$$\begin{aligned} x \times ((rx)z \dot{+} yz) &= x \times \Big(\neg\big(\neg((rx)z) \dot{-} yz\big)\Big) \\ &= x \times \Big(\mathbf{1} \dot{-} \big((\mathbf{1} \dot{-} (rx)z) \dot{-} yz\big)\Big) \\ &= x\mathbf{1} \dot{-} \big((x\mathbf{1} \dot{-} x((rx)z)) \dot{-} x(yz)\big) \\ &= x \dot{-} \big((x \dot{-} (x \times (rx))z) \dot{-} (xy)z\big) \end{aligned}$$

Then by $xx = x$, Lemma 4.3.5($v$), $xy = \mathbf{0}$, Fact 4.3.1($i$), (C), Lemma 4.3.5($iii$), (M6), (M4), and Fact 4.3.1($iv$), we have

$$\begin{aligned} x \dot{-} \big((x \dot{-} (x \times (rx))z) \dot{-} (xy)z\big) &= x \dot{-} \big((x \dot{-} (rx)z) \dot{-} \mathbf{0}z\big) = x \dot{-} \big((x \dot{-} (rx)z) \dot{-} \mathbf{0}\big) \\ &= x \dot{-} \big(x \dot{-} (rx)z\big) = x \wedge (r\mathbf{1} \times x)z \\ &= x \wedge \big(x \times (r\mathbf{1}) \times z\big) = x \times (r\mathbf{1}) \times z = (r\mathbf{1} \times x)z = (rx)z. \end{aligned}$$

Therefore, we have $x \times \big((rx \dot{+} y)z\big) = x \times ((rx)z \dot{+} yz) = (rx)z$.

Second, we will show $\neg x \times \big((rx \dot{+} yz\big) = \neg x \times ((rx)z \dot{+} yz) = (\neg x)(yz)$. By Fact 4.1.2($iii$), (M1), (M2), and (M6), we have

$$\begin{aligned} (\neg x) \times \big((rx \dot{+} y)z\big) &= (\neg x) \times \Big(\big(\neg(\neg(rx) \dot{-} y)\big)z\Big) = (\neg x) \times \Big(\big(\mathbf{1} \dot{-} ((\mathbf{1} \dot{-} rx) \dot{-} y)\big)z\Big) \\ &= (\neg x) \times \big(\mathbf{1}z \dot{-} ((\mathbf{1}z \dot{-} (rx)z) \dot{-} yz)\big) = (\neg x) \times \big(z \dot{-} ((z \dot{-} (rx)z) \dot{-} yz)\big) \\ &= (\neg x)z \dot{-} \big(((\neg x)z \dot{-} (\neg x)((rx)z)) \dot{-} (\neg x)(yz)\big) \\ &= (\neg x)z \dot{-} \big(((\neg x)z \dot{-} ((\neg x) \times (rx))z) \dot{-} (\neg x)(yz)\big). \end{aligned}$$

By Lemma 4.3.2, $xx = x$ implies $x \times \neg x = \mathbf{0}$. Then by $xx = x$, Lemma 4.3.5($iii$), (M6), (M4), and Fact

4.3.1($i$), we have

$$(\neg x) \times (rx) = (\neg x) \times (r\mathbf{1} \times x) = (\neg x \times r\mathbf{1}) \times x = (r\mathbf{1}) \times \big((\neg x) \times x\big) = (r\mathbf{1}) \times \mathbf{0} = \mathbf{0}. \tag{4.2}$$

Then by Equation (4.2), Fact 4.3.1($i$), (C), (M6), (M4), and Fact 4.3.1($iv$), we have

$$\begin{aligned}
&(\neg x)z \dot{-} \big(((\neg x)z \dot{-} ((\neg x) \times (rx))z) \dot{-} (\neg x)(yz)\big) = (\neg x)z \dot{-} \big(((\neg x)z \dot{-} \mathbf{0}z) \dot{-} (\neg x)(yz)\big) \\
&= (\neg x)z \dot{-} \big(((\neg x)z \dot{-} \mathbf{0}) \dot{-} (\neg x)(yz)\big) = (\neg x)z \dot{-} \big((\neg x)z \dot{-} (\neg x)(yz)\big) \\
&= (\neg x)z \wedge (\neg x)(yz) = (\neg x)z \wedge \big((\neg x)z\big)y = \big((\neg x)z\big)y = (\neg x)(yz).
\end{aligned}$$

Thus, $(\neg x) \times \big((rx \dot{+} y)z\big) = (\neg x)(yz)$.

Next, let us show $(\neg x) \times ((rx)z \dot{+} yz) = (\neg x)(yz)$. By Fact 4.1.2($iii$), (M1), (M4), (M2), (M6), we have

$$\begin{aligned}
(\neg x) \times ((rx)z \dot{+} yz) &= (\neg x) \times \Big(\neg\big(\neg((rx)z) \dot{-} yz\big)\Big) \\
&= (\neg x) \times \Big(\mathbf{1} \dot{-} \big((\mathbf{1} \dot{-} (rx)z) \dot{-} yz\big)\Big) \\
&= (\neg x)\mathbf{1} \dot{-} \big(((\neg x)\mathbf{1} \dot{-} (\neg x)((rx)z)) \dot{-} (\neg x)(yz)\big) \\
&= (\neg x) \dot{-} \big(((\neg x) \dot{-} ((\neg x) \times (rx))z) \dot{-} (\neg x)(yz)\big)
\end{aligned}$$

Then by Equation (4.2), Fact 4.3.1($i$), (C), Lemma 4.3.5($iii$), (M6), (M4), and Fact 4.3.1($iv$), we have

$$\begin{aligned}
&(\neg x) \dot{-} \big(((\neg x) \dot{-} ((\neg x) \times (rx))z) \dot{-} (\neg x)(yz)\big) = (\neg x) \dot{-} \big(((\neg x) \dot{-} \mathbf{0}z) \dot{-} (\neg x)(yz)\big) \\
&= (\neg x) \dot{-} \big(((\neg x) \dot{-} \mathbf{0}) \dot{-} (\neg x)(yz)\big) = (\neg x) \dot{-} \big((\neg x) \dot{-} (\neg x)(yz)\big) \\
&= (\neg x) \wedge (\neg x)(yz) = (\neg x)(yz).
\end{aligned}$$

Therefore, we have $(\neg x) \times \big((rx \dot{+} y)z\big) = (\neg x) \times ((rx)z \dot{+} yz) = (\neg x)(yz)$. □

**4.3.7 Lemma.** *Let $\mathcal{M}$ be a model of $\mathsf{RV}(\times)$. For all $n \in \mathbb{N}$, all $r_1, \cdots, r_n \in \mathbb{D}$, and all $x_1, \cdots, x_n, y \in M$ satisfying $x_i x_i = x_i$ for all $1 \leq i \leq n$ and $x_i x_j = \mathbf{0}$ for all $1 \leq i \neq j \leq n$, we have*

$$(r_1 x_1 \dot{+} \cdots \dot{+} r_n x_n)y = r_1(x_1 y) \dot{+} \cdots \dot{+} r_n(x_n y).$$

*Proof.* We use induction on $n$. When $n = 1$, by Lemma 4.3.5($iii$) and (M6) we have $(r_1 x_1)y = (r_1\mathbf{1} \times x_1)y = (r_1\mathbf{1}) \times (x_1 y) = r(x_1 y)$. Suppose the conclusion holds for all $m \leq n-1$. By (M4), the induction assumption,

Lemma 4.1.6($vii$), and Fact 4.1.2($xi$), we have

$$x_n(r_1x_1 \dot{+} \cdots \dot{+} r_{n-1}x_{n-1}) = (r_1x_1 \dot{+} \cdots \dot{+} r_{n-1}x_{n-1})x_n = r_1(x_1x_n) \dot{+} \cdots \dot{+} r_{n-1}(x_{n-1}x_n)$$
$$= r_1\mathbf{0} \dot{+} \cdots \dot{+} r_{n-1}\mathbf{0} = \mathbf{0} \dot{+} \cdots \dot{+} \mathbf{0} = \mathbf{0}.$$

Then by (P3), Fact 4.1.2($ix$), Lemma 4.3.6 and the induction assumption, we have

$$\begin{aligned}(r_1x_1 \dot{+} \cdots \dot{+} r_nx_n)y &= \big(r_nx_n \dot{+} (r_1x_1 \dot{+} \cdots \dot{+} r_{n-1}x_{n-1})\big)y = (r_nx_n)y \dot{+} (r_1x_1 \dot{+} \cdots \dot{+} r_{n-1}x_{n-1})y \\ &= r_n(x_ny) \dot{+} (r_1(x_1y) \dot{+} \cdots \dot{+} r_{n-1}(x_{n-1}y)) \\ &= (r_1(x_1y) \dot{+} \cdots \dot{+} r_{n-1}(x_{n-1}y)) \dot{+} r_n(x_ny) \\ &= r_1(x_1y) \dot{+} \cdots \dot{+} r_n(x_ny).\end{aligned}$$

□

**4.3.8 Lemma.** *Let $\mathcal{M}$ be a model of $\mathsf{RV}(\times)$. Let $x = (x_1, \cdots, x_m)$ and $y = (y_1, \cdots, y_n)$ be tuples in $M$. Suppose $x$ and $y$ are partitions of $\mathbf{1}$. That is, $x_ix_i = x_i$ for all $1 \leq i \leq m$, $y_jy_j = y_j$ for all $1 \leq j \leq n$, $x_i \times x_{i'} = y_j \times y_{j'} = \mathbf{0}$ for all $1 \leq i \neq i' \leq m$, $1 \leq j \neq j' \leq n$, and $x_1 \dot{+} \cdots \dot{+} x_m = y_1 \dot{+} \cdots \dot{+} y_n = \mathbf{1}$. Then for all $r_1, \cdots, r_m, s_1, \cdots, s_n \in \mathbb{D}$, we have*

$$\begin{aligned}&(r_1x_1 \dot{+} \cdots \dot{+} r_mx_m) \times (s_1y_1 \dot{+} \cdots \dot{+} s_ny_n) \\ &= (r_1s_1)(x_1y_1) \dot{+} \cdots \dot{+} (r_1s_n)(x_1y_n) \dot{+} \cdots \dot{+} (r_ms_1)(x_my_1) \dot{+} \cdots \dot{+} (r_ms_n)(x_my_n).\end{aligned}$$

*On the right hand side of the equation, there are $mn$ many terms, being of the forms $(r_is_j)(x_iy_j)$ for all $1 \leq i \leq m$ and $1 \leq j \leq n$.*

*Proof.* For all $1 \leq i \leq m$, by (M4) and Lemma 4.3.7 we have

$$(r_ix_i) \times (s_1y_1 \dot{+} \cdots \dot{+} s_ny_n) = (s_1y_1 \dot{+} \cdots \dot{+} s_ny_n) \times (r_ix_i) = s_1(y_1(r_ix_i)) \dot{+} \cdots \dot{+} s_n(y_n(r_ix_i)).$$

For all $1 \leq j \leq n$, by Lemma 4.3.5($iii$), ($v$), (M2), and (M4) we have

$$s_j(y_j(r_ix_i)) = (s_j\mathbf{1}) \times (y_j \times (r_ix_i)) = (s_j\mathbf{1}) \times (r_i(y_j \times x_i)) = (s_jr_i)(\mathbf{1} \times (y_j \times x_i)) = (r_is_j)(x_iy_j).$$

Hence,

$$(r_i x_i) \times (s_1 y_1 \dot{+} \cdots \dot{+} s_n y_n) = (r_i s_1)(x_i y_1) \dot{+} \cdots \dot{+} (r_i s_n)(x_i y_n).$$

The rest follows from Lemma 4.3.7, Lemma 4.1.4($ii$) and induction. □

**4.3.9 Theorem.** *Let $\mathcal{M}$ be a model of $\mathsf{RV}(\times)$. Then $\mathcal{M}$ is $L_{\mathsf{RV}(\times)}$-isomorphic to*

$$\Big(L^1\big(\mu, [0,1]\big), \mathbf{0}, \neg, \dot{-}, \frac{1}{2}, \times, I, d\Big),$$

*where $(\Omega, \mathcal{F}, \mu)$ is a probability space.*

*Proof.* By Theorem 4.1.10, we may assume $M = L^1(\mu, [0,1])$, where $(\Omega, \mathcal{F}, \mu)$ is a probability space. By Lemmas 4.3.4 and 4.3.8, we know that "$\times$" on $M$ is multiplication for all $\mathbb{D}$-valued simple functions. From real analysis, we know that the set of all $\mathbb{D}$-valued simple functions is dense in $M = L^1(\mu, [0,1])$. Because "$\times$" and the usual multiplication function on $M$ are continuous, we know that "$\times$" on $M$ is exactly the multiplication function between random variables. □

Let $\mathsf{ARV}(\times)$ denote the theory consisting of axioms in $\mathsf{RV}(\times)$ together with Axiom (NA) in $\mathsf{ARV}$.

**4.3.10 Corollary.** *Let $\mathcal{M}$ be a model of $\mathsf{ARV}(\times)$. Then $\mathcal{M}$ is $L_{\mathsf{RV}(\times)}$-isomorphic to*

$$\Big(L^1\big(\mu, [0,1]\big), \mathbf{0}, \neg, \dot{-}, \frac{1}{2}, \times, I, d\Big),$$

*where $(\Omega, \mathcal{F}, \mu)$ is an atomless probability space.*

*Proof.* This follows from Theorem 4.3.9 and Corollary 4.1.13. □

**4.3.11 Remark.** *By Beth's Definability Theorem (Theorem 1.6.8), Theorem 4.3.9 and Corollary 4.3.10, "$\times$" is defined over $L_{\mathsf{RV}}$ in both $\mathsf{RV}(\times)$ and $\mathsf{ARV}(\times)$. Thus, $\mathsf{RV}(\times)$ (resp. $\mathsf{ARV}(\times)$) is an extension by definitions of $\mathsf{RV}$ (resp. $\mathsf{ARV}$). The fact that "$\times''$ is uniformly definable was showed in Ben Yaacov [2, Lemma 2.13] in every model of his theory $\mathsf{RV}(\mathsf{BY})$, and hence the same is true in every model of our theory $\mathsf{RV}$. By Corollary 1.6.9, the class of $L_{\mathsf{RV}(\times)}$-expansions of models of $\mathsf{RV}$ is axiomatizable, and its theory is an extension by definitions of the $L_{\mathsf{RV}}$-theory $\mathsf{RV}$. But here, we give an explicit set of axioms for the theory of this class of expansions.*

# Chapter 5

# Model theory of random variable structures

This chapter is a continuation of Chapter 4. In this chapter, we make a comprehensive study of the model theory of random variable structures, which was initiated by Ben Yaacov in [2]. It is the most important chapter of this thesis. In Section 5.1, Theorem 5.1.1 shows how to generate a natural random variable structure from a probability algebra. In Section 5.2, we give a detailed proof of Theorem 5.2.3 from [2], which presents basic properties of $\mathsf{ARV}$. In Section 5.3, we study the type spaces of $\mathsf{ARV}$. Our main results are two explicit formulas for the $d^*$-metric between types in $\mathsf{ARV}$; see Theorems 5.3.9 and 5.3.14. A new proof of an integral formula for the Wasserstein distance between probability measures on the real line is given in Section 5.4. In Section 5.5, we characterize the saturated models of $\mathsf{ARV}$. In Section 5.6, we give a proof of a result from Ben Yaacov [2] that $\mathsf{ARV}$ is $\omega$-stable and that describes its model theoretic independence notion in probabilistic terms. In Section 5.7, we characterize the $d$-finite tuples in $\mathsf{ARV}$ and give some applications. In Section 5.8, we define conditional Wasserstein distances. Our definition of conditional Wasserstein distances generalizes the definition of Wasserstein distances.

## 5.1 Probability algebras and random variable structures

In this section, we will prove that every probability algebra determines a unique random variable structure up to isomorphism.

Let $(\mathscr{A}, \mu)$ be the probability algebra associated to the probability space $(\Omega, \mathcal{F}, \mu)$. Then for all $a \in \mathscr{A}$, there exists $F \in \mathcal{F}$ such that $a = [F]_\mu$. Let $\mathbb{D}$ denote the set of the dyadic numbers in $[0,1]$ and let $L^1\big((\Omega, \mathcal{F}, \mu), \mathbb{D}\big)$ denote the set of equivalence classes of $\mathbb{D}$-valued simple functions in $\mathcal{L}^1\big((\Omega, \mathcal{F}, \mu), \mathbb{D}\big)$. Note that $L^1\big((\Omega, \mathcal{F}, \mu), \mathbb{D}\big)$ is an $L_{\mathsf{RV}}$-prestructure which is dense in the $L_{\mathsf{RV}}$-structure $L^1\big((\Omega, \mathcal{F}, \mu), [0,1]\big)$.

**5.1.1 Theorem.** *Let $(\mathscr{A}, \mu)$ and $(\mathscr{A}', \mu')$ be probability algebras associated to probability spaces $(\Omega, \mathcal{F}, \mu)$ and $(\Omega', \mathcal{F}', \mu')$ respectively. Every $L_{\mathsf{Pr}}$-isomorphism $\psi\colon \mathscr{A} \to \mathscr{A}'$ induces an $L_{\mathsf{RV}}$-isomorphism*

$$R(\psi)\colon L^1\big((\Omega, \mathcal{F}, \mu), [0,1]\big) \to L^1\big((\Omega', \mathcal{F}', \mu'), [0,1]\big).$$

*On the other hand, every $L_{\mathsf{RV}}$-isomorphism*

$$\Psi\colon L^1\big((\Omega,\mathcal{F},\mu),[0,1]\big)\to L^1\big((\Omega',\mathcal{F}',\mu'),[0,1]\big)$$

*induces an $L_{\mathsf{Pr}}$-isomorphism $P(\Psi)\colon \mathscr{A}\to\mathscr{A}'$. Moreover, $R(P(\Psi))=\Psi$ and $P(R(\psi))=\psi$.*

*Proof.* We define $\pi\colon \mathscr{A}\to L^1\big((\Omega,\mathcal{F},\mu),\{0,1\}\big)$ by $\pi\big([a]_\mu\big)=[\chi_a]_{a.s.}$ for all $[a]_\mu\in\mathscr{A}$, where $a\in\mathcal{F}$. It is clear that $\pi$ is well-defined and injective. For each equivalence class $[f]_{a.s.}$ of a characteristic function $f\colon\Omega\to[0,1]$, there exists $F\in\mathcal{F}$ such that $f=\chi_F$, whereby $\pi([F]_\mu)=[\chi_F]_{a.s.}=[f]_{a.s.}$. Therefore $\pi$ defines a bijective mapping from $\mathscr{A}$ to $L^1\big((\Omega,\mathcal{F},\mu),\{0,1\}\big)$. By Proposition 4.1.3, we know that $\pi$ is an $L_{\mathsf{Pr}}$-isomorphism. Similarly, the mapping $\pi'\colon\mathscr{A}'\to L^1\big((\Omega',\mathcal{F}',\mu'),\{0,1\}\big)$ defined by $\pi'\big([a']_\mu\big)=[\chi_{a'}]_{a.s.}$ for all $[a']_\mu\in\mathscr{A}'$, where $a'\in\mathcal{F}'$, is also an $L_{\mathsf{Pr}}$-isomorphism.

Then for every $L_{\mathsf{Pr}}$-isomorphism $\psi\colon\mathscr{A}\to\mathscr{A}'$, there is the following $L_{\mathsf{Pr}}$-isomorphism

$$\varphi:=\pi'\circ\psi\circ\pi^{-1}\colon L^1\big((\Omega,\mathcal{F},\mu),\{0,1\}\big)\to L^1\big((\Omega',\mathcal{F}',\mu'),\{0,1\}\big).$$

Note that

$$L^1\big((\Omega,\mathcal{F},\mu),\mathbb{D}\big)=\{\sum_{i=1}^n d_i[\chi_{F_i}]_{a.s.}\mid n\in\mathbb{N},d_i\in\mathbb{D}\text{ and }F_i\in\mathcal{F}\text{ for each }i,\mu(F_i\cap F_j)=0\text{ whenever }i\neq j\}.$$

We define

$$\Psi_0\colon L^1\big((\Omega,\mathcal{F},\mu),\mathbb{D}\big)\to L^1\big((\Omega',\mathcal{F}',\mu'),\mathbb{D}\big)$$

by $\Psi_0(f):=\sum_{i=1}^n d_i\varphi([\chi_{F_i}]_{a.s.})$, where $f=\sum_{i=1}^n d_i[\chi_{F_i}]_{a.s.}$, $n\in\mathbb{N}$, $d_i\in\mathbb{D}$ and $F_i\in\mathcal{F}$ for all $1\leq i\leq n$, and $\mu(F_i\cap F_j)=0$ for all $1\leq i\neq j\leq n$. By Proposition 4.1.8, $\Psi_0$ is an $L_{\mathsf{RV}}$-embedding. Since $\varphi$ is an $L_{\mathsf{Pr}}$-isomorphism, we have that $\Psi_0$ is surjective, whereby it is an $L_{\mathsf{RV}}$-isomorphism between $L_{\mathsf{RV}}$-prestructures. Since $L^1\big((\Omega,\mathcal{F},\mu),\mathbb{D}\big)$ and $L^1\big((\Omega',\mathcal{F}',\mu'),\mathbb{D}\big)$ are dense in $L^1\big((\Omega,\mathcal{F},\mu),[0,1]\big)$ and $L^1\big((\Omega',\mathcal{F}',\mu'),[0,1]\big)$ respectively, $\Psi_0$ induces an $L_{\mathsf{RV}}$-isomorphism $\Psi\colon L^1\big((\Omega,\mathcal{F},\mu),[0,1]\big)\to L^1\big((\Omega',\mathcal{F}',\mu'),[0,1]\big)$. Then $\Psi$ is our desired $R(\psi)$. Note that $\Psi$ is uniquely determined by $\varphi$.

On the other hand, for all $L_{\mathsf{RV}}$-isomorphisms

$$\Psi\colon L^1\big((\Omega,\mathcal{F},\mu),[0,1]\big)\to L^1\big((\Omega',\mathcal{F}',\mu'),[0,1]\big),$$

we write $\varphi$ for the restriction of $\Psi$ to $L^1\big((\Omega,\mathcal{F},\mu),\{0,1\}\big)$. Then note that $L^1\big((\Omega,\mathcal{F},\mu),\{0,1\}\big)$ and $L^1\big((\Omega',\mathcal{F}',\mu'),\{0,1\}\big)$ are zeroset of $I(x\wedge\neg x)=0$ in $L^1\big((\Omega,\mathcal{F},\mu),[0,1]\big)$ and $L^1\big((\Omega',\mathcal{F}',\mu'),[0,1]\big)$ re-

spectively, so we have that

$$\varphi\colon L^1\big((\Omega,\mathcal{F},\mu),\{0,1\}\big)\to L^1\big((\Omega',\mathcal{F}',\mu'),\{0,1\}\big)$$

is a bijection. By Proposition 4.1.3, we know that

$$(L^1\big((\Omega,\mathcal{F},\mu),\{0,1\}\big),\mathbf{0},\mathbf{1},\neg,\wedge,\vee,\mu)$$

and

$$(L^1\big((\Omega',\mathcal{F}',\mu'),\{0,1\}\big),\mathbf{0},\mathbf{1},\neg,\wedge,\vee,\mu')$$

are models of $\mathsf{Pr}$. Since $\Psi$ is an $L_{\mathsf{RV}}$-isomorphism, we have that $\varphi$ is an $L_{\mathsf{Pr}}$-isomorphism and it is uniquely determined by $\Psi$. We define $\psi\colon \mathscr{A}\to\mathscr{A}'$ by $\psi := \pi'^{-1}\circ\varphi\circ\pi$. Because $\Psi$ is an $L_{\mathsf{RV}}$-isomorphism, $\varphi$ is an $L_{\mathsf{RV}}$-isomorphism, and $\pi,\pi'$ are $L_{\mathsf{Pr}}$-isomorphism, we have that $\psi$ is an $L_{\mathsf{Pr}}$-isomorphism, which is our desired $P(\Psi)$.

From the construction of $R(\psi)$ and $P(\Psi)$, we see that

$$\varphi=\pi'\circ\psi\circ\pi^{-1}\colon L^1\big((\Omega,\mathcal{F},\mu),\{0,1\}\big)\to L^1\big((\Omega',\mathcal{F}',\mu'),\{0,1\}\big)$$

uniquely determines $\Psi\colon L^1\big((\Omega,\mathcal{F},\mu),[0,1]\big)\to L^1\big((\Omega',\mathcal{F}',\mu'),[0,1]\big)$ and vice versa, and thus $P(R(\psi))=\psi$ and $R(P(\Psi))=\Psi$. □

**5.1.2 Remark.** *Let $(\mathscr{A},\mu)$ be the probability algebra associated to the probability space $(\Omega,\mathcal{F},\mu)$. By Theorem 5.1.1, the $L_{\mathsf{RV}}$-structure $L^1\big((\Omega,\mathcal{F},\mu),[0,1]\big)$ is determined by the probability algebra $(\mathscr{A},\mu)$ up to $L_{\mathsf{RV}}$-isomorphism. We define $L^1(\mathscr{A},[0,1])$ to be the $L_{\mathsf{RV}}$-structure $L^1\big((\Omega,\mathcal{F},\mu),[0,1]\big)$.*

## 5.2 Basic model theoretic properties of RV and ARV

In this section, we prove basic properties of the theory of atomless random variable structures that are taken from Ben Yaacov [2].

**5.2.1 Lemma** ([2])**.** *For all continuous functions $\varphi\colon[0,1]^n\to[0,1]$, there is a sequence of $L_{\mathsf{RV}}$-terms $\big(t_k(x)\big)_{k\in\mathbb{N}}$ such that for all models $\mathcal{M}$ of $\mathsf{RV}$, we have that $t_k^{\mathcal{M}}(x)$ uniformly converges to $\varphi^{\mathcal{M}}(x)$, where $\varphi^{\mathcal{M}}\colon M^n\to[0,1]$ is defined by $\varphi^{\mathcal{M}}(f)=\varphi(f)$ for all $f\in M^n$.*

*Proof.* This follows from [2, Lemma 2.13]. □

**5.2.2 Lemma** ([2]). *Let $\mathcal{M}$ be a model of* RV *and $f, g \in M^n$. Then $\mathrm{tp}^{\mathrm{qf}}_{\mathcal{M}}(f) = \mathrm{tp}^{\mathrm{qf}}_{\mathcal{M}}(g)$ if and only if $f$ and $g$ have the same joint distribution as random variables.*

*Proof.* This is [2, Lemma 2.15]. □

The following theorem includes basic properties of the theory ARV.

**5.2.3 Theorem** ([2, Theorem 2.17]). *Let $\mathcal{M} = L^1\big((\Omega, \mathcal{F}, \mu), [0,1]\big)$ be a model of* ARV. *Then*

*(i) The theory* ARV *is complete and separably categorical.*

*(ii) The theory* ARV *admits quantifier elimination.*

*(iii) The universal part of* ARV *is* RV, *and* ARV *is the model completion of* RV.

*(iv) If $A \subseteq M$, then $\mathrm{dcl}(A) = \mathrm{acl}(A) = L^1\big((\Omega, \sigma(A), \mu), [0,1]\big) \subseteq M$, where $\sigma(A)$ is the $\sigma$-algebra of measurable sets generated by the random variables in $A$.*

*(v) Two tuples $f$ and $g$ in $M^n$ have the same type over a set $A \subseteq M$ if and only if they have the same joint conditional distribution over $\sigma(A)$.*

*Proof.* $(i)$ Let $\mathcal{M}' = L^1\big((\Omega', \mathcal{F}', \mu'), [0,1]\big)$ and $\mathcal{M} = L^1\big((\Omega, \mathcal{F}, \mu), [0,1]\big)$ be separable models of ARV. Let $\widehat{\mathcal{F}'}$ and $\widehat{\mathcal{F}}$ be the probability algebras associated to the probability spaces $(\Omega', \mathcal{F}', \mu')$ and $(\Omega, \mathcal{F}, \mu)$ respectively. Then $\widehat{\mathcal{F}'}$ and $\widehat{\mathcal{F}}$ are separable models of APr. By Theorem 3.1.3 $(i)$, $\widehat{\mathcal{F}}$ is $L_{\mathsf{Pr}}$-isomorphic to $\widehat{\mathcal{F}'}$. By Theorem 5.1.1, we have that $\mathcal{M}$ is $L_{\mathsf{RV}}$-isomorphic to $\mathcal{M}'$. Hence ARV is separably categorical, whereby complete.

$(ii)$ Let $f$ and $g$ be $n$-tuples in a $\kappa$-universal domain $\mathcal{M}$ for ARV for an uncountable cardinal $\kappa$, and assume $\mathrm{tp}^{\mathrm{qf}}(f) = \mathrm{tp}^{\mathrm{qf}}(g)$. Suppose that $\mathcal{M} = L^1\big((\Omega, \mathcal{F}, \mu), [0,1]\big)$. By Lemma 5.2.1, for all continuous $\psi\colon [0,1]^n \to [0,1]$, we have $I(\psi(f)) = I(\psi(g))$, whereby $f$ and $g$ have the same joint distribution. Let $\mathbb{D}$ denote the set of the dyadic numbers in $[0,1]$. Let $\mathbb{D} \times n$ denote the set $\{(r,i) \mid r \in \mathbb{D}, 1 \leq i \leq n\}$. We write $a_{r,i}$ for the event $\{f_i \leq r\}$ and $b_{r,i}$ for the event $\{g_i \leq r\}$, where $(r,i) \in \mathbb{D} \times n$. Since $f$ and $g$ have the same joint distribution, $\mathrm{tp}^{\mathsf{APr},\mathrm{qf}}\big(\{a_{r,i}\}_{(r,i)\in\mathbb{D}\times n}\big) = \mathrm{tp}^{\mathsf{APr},\mathrm{qf}}\big(\{b_{r,i}\}_{(r,i)\in\mathbb{D}\times n}\big)$. By Theorem 3.1.3, the theory APr eliminates quantifiers. Thus, $\mathrm{tp}^{\mathsf{APr}}\big(\{a_{r,i}\}_{(r,i)\in\mathbb{D}\times n}\big) = \mathrm{tp}^{\mathsf{APr}}\big(\{b_{r,i}\}_{(r,i)\in\mathbb{D}\times n}\big)$. Since $\mathcal{M}$ is a $\kappa$-universal domain and $\widehat{\mathcal{F}}$ is a uniformly definable subset by Proposition 4.1.3, we know that $\widehat{\mathcal{F}}$ is a $\kappa$-universal domain for APr. Because $\kappa$ is uncountable, there is $\varphi \in \mathrm{Aut}(\widehat{\mathcal{F}})$ such that $\varphi(a_{r,i}) = b_{r,i}$ for every $(r,i) \in \mathbb{D} \times n$. By Theorem 5.1.1, $\varphi$ induces an $L_{\mathsf{RV}}$-isomorphism $\Phi \in \mathrm{Aut}(\mathcal{M})$ such that $\mathbf{1}_{\varphi(A)} = \Phi(\mathbf{1}_A)$ for every $A \in \widehat{\mathcal{F}}$. Then for every $(r,i) \in \mathbb{D} \times n$, we know that $\Phi(\mathbf{1}_{a_{r,i}}) = \mathbf{1}_{b_{r,i}}$; *i.e.*, $\Phi(\mathbf{1}_{\{f_i \leq r\}}) = \mathbf{1}_{\{g_i \leq r\}}$. By the definition of the Lebesgue integral, we have that $\int_\Omega |\Phi(f_i) - g_i| d\mu = 0$ for each $i \leq n$. Thus

$d\big(\Phi(f), g\big) = \max_{1\leq i\leq n} d\big(\Phi(f_i), g_i\big) = 0$, whereby $\Phi(f) = g$. Hence $\mathrm{tp}(f) = \mathrm{tp}(g)$. Then by Theorem 1.10.4, $\mathsf{ARV}$ admits quantifier elimination.

$(iii)$ By Corollary 4.1.12, we know that $\mathsf{RV}$ is a universal theory. By Theorem 4.1.10 and Corollary 4.1.13, we have that the $\mathsf{RV}$-substructures of models of $\mathsf{ARV}$ are precisely the models of $\mathsf{RV}$. Therefore the theory $\mathsf{RV}$ is the universal part of the theory $\mathsf{ARV}$. Then by $(ii)$ and Proposition 1.10.5, we know that $\mathsf{ARV}$ is the model completion of $\mathsf{RV}$.

$(iv)$ By Lemma 5.2.1, one has that for every $n \in \mathbb{N}$ and every continuous $\varphi\colon [0,1]^n \to [0,1]$, there is a sequence of $L_{\mathsf{RV}}$-terms $(t_{k,\varphi})_{k\in\mathbb{N}}$ such that $t_{k,\varphi}^{\mathcal{M}}$ uniformly converges to $\varphi^{\mathcal{M}}$. For all $a_1, \cdots, a_n \in A$ and all $k \in \mathbb{N}$, note that $d(x, t_{k,\varphi}(a_1, \cdots, a_n))$ is an $L_{\mathsf{RV}}(A)$-formula. Since $t_{k,\varphi}^{\mathcal{M}}(a_1, \cdots, a_n)$ converges to $\varphi^{\mathcal{M}}(a_1, \cdots, a_n)$, we have that $d^{\mathcal{M}}(x, t_{k,\varphi}^{\mathcal{M}}(a_1, \cdots, a_n))$ uniformly converges to $d^{\mathcal{M}}(x, \varphi^{\mathcal{M}}(a_1, \cdots, a_n))$. That is $d(x, \varphi(a_1, \cdots, a_n))$ is definable over $A$. Hence, $\varphi(a_1, \cdots, a_n)$ is in $\mathrm{dcl}(A)$. Therefore

$$\{\varphi(a_1, \cdots, a_n) \mid n \in \mathbb{N}, \varphi \in C([0,1]^n, [0,1]), a_1, \cdots, a_n \in A\} \subseteq \mathrm{dcl}(A).$$

Then by Proposition 1.6.11, one has $\mathrm{dcl}\big(\{\varphi(a_1, \cdots, a_n) \mid n \in \mathbb{N}, \varphi \in C([0,1]^n, [0,1]), a_1, \cdots, a_n \in A\}\big) \subseteq \mathrm{dcl}(A)$. Since the set $\{\varphi(a_1, \cdots, a_n) \mid n \in \mathbb{N}, \varphi \in C([0,1]^n, [0,1]), a_1, \cdots, a_n \in A\}$ is dense in the space $L^1\big((\Omega, \sigma(A), \mu), [0,1]\big)$, one has

$$\mathrm{dcl}\big(\{\varphi(a_1, \cdots, a_n) \mid n \in \mathbb{N}, \varphi \in C([0,1]^n, [0,1]), a_1, \cdots, a_n \in A\}\big) \supseteq L^1\big((\Omega, \sigma(A), \mu), [0,1]\big),$$

whereby $L^1\big((\Omega, \sigma(A), \mu), [0,1]\big) \subseteq \mathrm{dcl}(A) \subseteq \mathrm{acl}(A)$.

Let $f$ be in $\mathrm{acl}(A)$. As shown above, $L^1\big((\Omega, \sigma(f), \mu), [0,1]\big) \subseteq \mathrm{dcl}(f) \subseteq \mathrm{dcl}(\mathrm{acl}(A))$. For every $B \in \sigma(f)$, one has $\mathbf{1}_B \in \mathrm{dcl}(\mathrm{acl}(A))$. By Proposition 1.6.11, one has $\mathrm{acl}(A) \subseteq \mathrm{dcl}(\mathrm{acl}(A)) \subseteq \mathrm{acl}(\mathrm{acl}(A)) = \mathrm{acl}(A)$, whereby $\mathbf{1}_B \in \mathrm{acl}(A)$. Suppose that $B$ is not in $\sigma(A)$. Consider a $\kappa$-universal domain $\mathscr{B} \models \mathsf{APr}$ where $\kappa$ is uncountable, $\sigma(A)$ is small, and $\sigma(A) \preceq \mathscr{B}$. Let $\mathrm{Aut}(\mathscr{B} \mid \sigma(A))$ denote the set of all $L_{\mathsf{Pr}}$-automorphisms of $\mathscr{B}$ that fix $\sigma(A)$ pointwise. By Theorem 3.1.3 and Proposition 1.6.10, one has that the orbit set $\{\Phi(B) \mid \Phi \in \mathrm{Aut}(\mathscr{B} \mid \sigma(A))\}$ has density character $\geq \kappa$. Every $\Phi \in \mathrm{Aut}(\mathscr{B} \mid \sigma(A))$ induces an $L_{\mathsf{RV}}$-automorphism of $L^1(\mathscr{B}, [0,1])$ that fixes $A$ pointwise. Let $Aut\big(L^1(\mathscr{B}, [0,1]) \mid A\big)$ denote those induced $L_{\mathsf{RV}}$-automorphism. Then the orbit set $\{\psi(\mathbf{1}_B) \mid \psi \in Aut\big(L^1(\mathscr{B}, [0,1]) \mid A\big)\}$ also has density character $\geq \kappa$. Consequently, this orbit set is not compact in $L^1(\mathscr{B}, [0,1])$, whereby the realization of $\mathrm{tp}(\mathbf{1}_B/A)$ is not compact in $L^1(\mathscr{B}, [0,1])$. Hence $\mathbf{1}_B$ is not $A$-algebraic; *i.e.*, $\mathbf{1}_B \notin \mathrm{acl}(A)$, which is a contradiction. Therefore, $B \in \sigma(A)$ for every

$B \in \sigma(f)$, whereby $f$ is $\sigma(A)$-measurable. Hence

$$\mathrm{acl}(A) \subseteq L^1\big((\Omega, \sigma(A), \mu), [0,1]\big) \subseteq \mathrm{dcl}(A) \subseteq \mathrm{acl}(A).$$

$(v)$ By $(iv)$, one has $\mathrm{tp}(h/A) = \mathrm{tp}(h/\sigma(A))$ for every $h \in M^n$. When $\sigma(A)$ is finite, say $\sigma(A) = \{B_1, \cdots, B_m\}$. Let $\mathrm{dist}(h \mid \sigma(A))$ denote the joint conditional distribution of $h$ over $\sigma(A)$ for $h \in M^n$. Then $\mathrm{dist}(f \mid \sigma(A)) = \mathrm{dist}(g \mid \sigma(A))$ if and only if $\mathrm{dist}(f, \mathbf{1}_{B_1}, \cdots, \mathbf{1}_{B_m}) = \mathrm{dist}(g, \mathbf{1}_{B_1}, \cdots, \mathbf{1}_{B_m})$. By Lemma 5.2.2 and $(ii)$, one has $\mathrm{dist}(f, \mathbf{1}_{B_1}, \cdots, \mathbf{1}_{B_m}) = \mathrm{dist}(g, \mathbf{1}_{B_1}, \cdots, \mathbf{1}_{B_m})$ if and only if $\mathrm{tp}(f, \mathbf{1}_{B_1}, \cdots, \mathbf{1}_{B_m}) = \mathrm{tp}(g, \mathbf{1}_{B_1}, \cdots, \mathbf{1}_{B_m})$, which is equivalent to $\mathrm{tp}(f/A) = \mathrm{tp}(g/A)$. In the case when $\sigma(A)$ is infinite, we consider the set of all finite subalgebras of $\sigma(A)$, denoted by $\{\mathscr{A}_t \mid t \in T\}$ where $T$ is an index set. One has that $\mathrm{tp}(f/\sigma(A)) = \mathrm{tp}(g/\sigma(A))$ if and only if $\mathrm{tp}(f/\mathscr{A}_t) = \mathrm{tp}(g/\mathscr{A}_t)$ for all $t \in T$. We just proved that $\mathrm{tp}(f/\mathscr{A}_t) = \mathrm{tp}(g/\mathscr{A}_t)$ if and only if $\mathrm{dist}(f \mid \mathscr{A}_t) = \mathrm{dist}(g \mid \mathscr{A}_t)$. Also one has $\mathrm{dist}(f \mid \sigma(A)) = \mathrm{dist}(g \mid \sigma(A))$ if and only if $\mathrm{dist}(f \mid \mathscr{A}_t) = \mathrm{dist}(g \mid \mathscr{A}_t)$ for all $t \in T$. Hence $\mathrm{tp}(f/\sigma(A)) = \mathrm{tp}(g/\sigma(A))$ if and only if $\mathrm{dist}(f \mid \sigma(A)) = \mathrm{dist}(g \mid \sigma(A))$. □

**5.2.4 Remark.** *We follow [2] for the detailed proofs of $(i)$, $(ii)$, $(iii)$, and $(v)$, while our proof for $(iv)$ is different.*

For all $t = (t_1, \cdots, t_n) \in [0,1]^n$, define $B_t := [0, t_1) \times \cdots \times [0, t_n)$.

**5.2.5 Lemma.** *Let $\mathcal{M}_1$ and $\mathcal{M}_2$ be models of $\mathsf{RV}$ and let $\Psi\colon M_1 \to M_2$ be an $L_{\mathsf{RV}}$-embedding. Then for all $n \in \mathbb{N}$ and all $a \in M_1^n$, we have*

*(i)* $\Psi(\mathbf{1}_{a \in B_t}) = \mathbf{1}_{\Psi(a) \in B_t}$ *for all* $t \in [0,1]^n$.

*(ii)* $\Psi(\varphi(a)) = \varphi(\Psi(a))$ *for all continuous* $\varphi\colon [0,1]^n \to [0,1]$.

*Proof.* $(i)$: For all $t = (t_1, \cdots, t_n) \in [0,1]^n$, and all variables $x = (x_1, \cdots, x_n)$, define $g^t(x)$ to be $\min\{t_1 \mathbf{1} \dot{-} x_1, \cdots, t_n \mathbf{1} \dot{-} x_n\}$. Let $g_m^t(x)$ denote $\underbrace{g^t(x) \dot{+} \cdots \dot{+} g^t(x)}_{m \text{ times}}$ for all $m \in \mathbb{N}$. Then $g_m^t(x)$ is an $L_{\mathsf{RV}}$-term. Note that for every model $\mathcal{M}$ of $\mathsf{RV}$ and $a = (a_1, \cdots, a_n) \in M^n$, we have

$$\lim_{m \to \infty} g_m^t(a) = \mathbf{1}_{\{a_1 < t_1, \cdots, a_n < t_n\}} = \mathbf{1}_{a \in B_t}$$

where the limit is the almost sure pointwise limit of random variables. Then by the dominated convergence theorem, the limit of $(g_m^t(a))_{m \in \mathbb{N}}$ in the sense of the metric of $M$ (*i.e.*, $L^1$-metric) is $\mathbf{1}_{a \in B_t}$.

Since $\Psi$ is an $L_{\mathsf{RV}}$-embedding, we have $\Psi(g_m^t(a)) = g_m^t(\Psi(a))$ for all $t \in [0,1]^n$, all $m \in \mathbb{N}$ and all $a \in M_1^n$. Since $\Psi$ is continuous, we have $\Psi(\lim_{m\to\infty} g_m^t(a)) = \lim_{m\to\infty} g_m^t(\Psi(a))$, whereby $\Psi(\mathbf{1}_{a\in B_t}) = \mathbf{1}_{\Psi(a)\in B_t}$ for all $t \in [0,1]^n$ and all $a \in M_1^n$.

$(ii)$: By Lemma 5.2.1, we know that for every continuous $\varphi\colon [0,1]^n \to [0,1]$, there is a sequence of $L_{\mathsf{RV}}$-terms $(t_{k,\varphi})_{k\in\mathbb{N}}$ such that $t_{k,\varphi}^{\mathcal{M}}$ uniformly converges to $\varphi^{\mathcal{M}}$ for all models $\mathcal{M}$ of $\mathsf{RV}$. Since $\Psi$ is an $L_{\mathsf{RV}}$-embedding, for every continuous $\varphi\colon [0,1]^n \to [0,1]$ and every $k \in \mathbb{N}$, we have $d(\Psi(b), t_{k,\varphi}(\Psi(a))) = d(b, t_{k,\varphi}(a))$ for all $a \in M_1^n$ and all $b \in M_1$. Letting $k$ go to $\infty$, we have $d(\Psi(b), \varphi(\Psi(a))) = d(b, \varphi(a))$, and thus $\Psi(\varphi(a)) = \varphi(\Psi(a))$ for all $a \in M_1^n$. □

## 5.3 On type spaces of ARV

In this section we investigate the type spaces of ARV. First, we present the results from [3] about the logic topology on the type spaces of ARV. Then, we study the metric topology on types. Our main results are Theorems 5.3.9 and 5.3.14.

### The logic topology on types

Here, we introduce the results from [3] to connect the logic topology with the topology of weak convergence. We begin by defining some notions of convergence of distributions from probability theory, *e.g.*, weak convergence and strong convergence. We follow the definitions in [8].

**5.3.1 Definition.** *Let $\mathscr{A}$ be a probability algebra. An $n$-**dimensional conditional distribution over** $\mathscr{A}$, $\mu$, is an $L^1\big(\mathscr{A}, [0,1]\big)$-valued Borel probability measure on $\mathbb{R}^n$. More precisely, it satisfies the following:*

- $\mu(B) \geq 0$ *a.s. for all Borel sets* $B \subseteq \mathbb{R}^n$.
- $\mu(\mathbb{R}^n) = 1$ *a.s.*
- $\mu(\bigcup_{i=1}^{\infty} B_i) = \sum_{i=1}^{\infty} \mu(B_i)$ *a.s. for all disjoint Borel sets* $B_1, B_2, \cdots \subseteq \mathbb{R}^n$.

*Let $\mathfrak{D}_{\mathbb{R}^n}(\mathscr{A})$ denote the space of all $n$-dimensional conditional distributions over $\mathscr{A}$. For every Borel set $B \subseteq \mathbb{R}^n$, we denote by $\mathfrak{D}_B(\mathscr{A})$ the space of all $n$-dimensional conditional distributions over $\mathscr{A}$ which, as measures, are supported by $B$.*

Note that the $n$-dimensional conditional distributions over $\emptyset$ are exactly the $n$-dimensional distributions.

Let $f$ be an $n$-tuple of real-valued random variables. The *(joint) conditional distribution of $f$ over $\mathscr{A}$*, denoted here by $\mu = \operatorname{dist}(f \mid \mathscr{A})$, is the $n$-dimensional conditional distribution over $\mathscr{A}$ given by

$$\mu(B) = \mathbb{P}(f \in B \mid \mathscr{A}) \text{ for all Borel sets } B \subseteq \mathbb{R}^n.$$

A net $(f_i)_{i\in I} \subseteq L^1\big(\mathscr{A}, [0,1]\big)$ *converges weakly to $f$* if for every $g \in L^\infty\big(\mathscr{A}, [0,1]\big)$, we have $E(f_i g) \to E(fg)$. We say the net $(f_i)_{i\in I}$ *converges strongly to $f$* if it converges in $L^1$.

Now we introduce the definitions of weak and strong convergence of conditional distributions.

**5.3.2 Definition.** *Let $\mu, (\mu_i)_{i\in\mathbb{N}}$ be $n$-dimensional conditional distributions over $\mathscr{A}$. We say the sequence $(\mu_i)_{i\in\mathbb{N}}$* ***converges weakly (resp. strongly)*** *to $\mu$ if for every $x$ with $\mu(\{x\}) = 0$, we have $\mu_i\big((-\infty, x]\big)$ converges weakly (resp. strongly) to $\mu\big((-\infty, x]\big)$.*

**5.3.3 Theorem** ([3])**.** *Let $f$ be an $n$-tuple in a model $\mathcal{M}$ of* ARV *and let $A$ be a subset of $M$. Let $\mathscr{A}$ denote $\sigma(A)$, the $\sigma$-algebra of measurable sets generated by the random variables in $A$. Then the joint conditional distribution* $\operatorname{dist}(f \mid \mathscr{A})$ *only depends on* $\operatorname{tp}(f/A)$*. Moreover, the mapping*

$$\begin{aligned} \zeta\colon S_n(A) &\to \mathfrak{D}_{[0,1]^n}(\mathscr{A}) \\ \operatorname{tp}(f/A) &\mapsto \operatorname{dist}(f \mid \mathscr{A}) \end{aligned}$$

*is a homeomorphism between $S_n(A)$ equipped with the logic topology and $\mathfrak{D}_{[0,1]^n}(\mathscr{A})$ equipped with the topology of weak convergence.*

*Proof.* This is [3, Theorem 4.3]. □

From now on, because of the above theorem, we may use $\zeta$ to identify $S_n(A)$, the set of $n$-types over $A$ with $\mathfrak{D}_{[0,1]^n}\big(\sigma(A)\big)$, the set of $n$-dimensional conditional distributions over $\sigma(A)$.

**5.3.4 Theorem** ([3])**.** *Let $A$ be a set of parameters, and identify $S_n(A)$ with $\mathfrak{D}_{[0,1]^n}\big(\sigma(A)\big)$. Then the metric topology on types and the topology of strong convergence of conditional distributions agree.*

*Proof.* This is [3, Theorem 4.9]. □

## The metric topology on types

Here, we study the $d$-metric and the $d^*$-metric between types with connections to Wasserstein distances. The main results are two explicit formulas for the $d$-metric and the $d^*$-metric between types. Theorem 5.3.9 is for a special case and Theorem 5.3.14 is for the most general case.

We first show a theorem concerning the relation between Wasserstein spaces and type spaces of ARV. Fix a signature $L$ for metric structures and an consistent $L$-theory $T$. Let $\mathcal{M}$ be a model of $T$ and $A \subseteq M$. Denote the $L(A)$-structure $(\mathcal{M}, a)_{a\in A}$ by $\mathcal{M}_A$, and set $T_A$ to be the $L(A)$-theory of $\mathcal{M}_A$.

**5.3.5 Definition.** *For each $n \geq 1$, let $a = (a_1, \cdots, a_n)$ and $b = (b_1, \cdots, b_n)$ be in a metric structure $M^n$, we define $d^*(a, b)$ to be $\sum_{i=1}^n d(a_i, b_i)$. Let $T_A$ be as above and let $\mathcal{M}_A = (\mathcal{M}, a)_{a\in A}$ be any model of $T_A$ in which every type in $S_n(T_A)$ is realized. For all $p, q \in S_n(T_A)$, we define $d^*(p, q)$ to be*

$$\inf\Big\{\sum_{i=1}^n d(b_i, c_i) \mid \mathcal{M}_A \models p[b_1, \cdots, b_n] \text{ and } \mathcal{M}_A \models q[c_1, \cdots, c_n]\Big\}.$$

Note that $d(a, b) \leq d^*(a, b) \leq nd(a, b)$ and $d(p, q) \leq d^*(p, q) \leq nd(p, q)$. This new metric therefore defines the same topology on type spaces as the usual metric. Model theoretic results such as the Ryll-Nardzewski Theorem for metric structures only depend on the metric topology on type spaces, so this new metric is adequate for model theoretic purposes. The reason for us to prefer this metric is the following:

**5.3.6 Theorem.** *Let $f$ be an $n$-tuple in a model $\mathcal{M}$ of ARV. Suppose $M = L^1\big(\mu, [0, 1]\big)$ for the atomless probability space $(\Omega, \mathcal{F}, \mu)$. Then $\mathrm{dist}(f)$ is a probability measure on $[0, 1]^n$. Moreover, the mapping*

$$\begin{aligned} \eta_n \colon S_n(ARV) &\to P_1\big([0, 1]^n\big) \\ \mathrm{tp}(f) &\mapsto \mathrm{dist}(f) \end{aligned}$$

*is an isometric isomorphism between $\big(S_n(ARV), d^*\big)$ and $\Big(P_1\big([0, 1]^n\big), W_1\Big)$, where the metric on which the definition of $W_1$ depends is defined by*

$$d^*_{[0,1]^n}(a, b) = \sum_{i=1}^n |a_i - b_i|, \text{ for all } a, b \in [0, 1]^n.$$

*Proof.* By Theorem 5.3.3, we know that $\eta_n$ is well-defined and bijective. Let $f$ and $g$ be $n$-tuples in $M$.

$$\begin{aligned} W_1\big(\mathrm{dist}(f), \mathrm{dist}(g)\big) &= \inf\Big\{\mathbb{E}\big(\sum_{i=1}^n |X_i - Y_i|\big) : \mathrm{dist}(X) = \mathrm{dist}(f), \mathrm{dist}(Y) = \mathrm{dist}(g)\Big\} \\ &= \inf\Big\{\sum_{i=1}^n \mathbb{E}\big(|X_i - Y_i|\big) : \mathrm{dist}(X) = \mathrm{dist}(f), \mathrm{dist}(Y) = \mathrm{dist}(g)\Big\} \\ &= d^*\big(\mathrm{tp}(f), \mathrm{tp}(g)\big). \end{aligned}$$

□

**5.3.7 Corollary.** *Let $f$ be an $n$-tuple in a model $\mathcal{M}$ of* ARV*. Suppose $M = L^1\big(\mu, [0,1]\big)$ for the atomless probability space $(\Omega, \mathcal{F}, \mu)$. Then* $\operatorname{dist}(f)$ *is a probability measure on $[0,1]^n$. Moreover, the mapping*

$$\eta_n \colon S_n(ARV) \to P_1\big([0,1]^n\big)$$
$$\operatorname{tp}(f) \mapsto \operatorname{dist}(f)$$

*is bijective and bi-Lipschitz between $\big(S_n(ARV), d\big)$ and $\Big(P_1\big([0,1]^n\big), W_1\Big)$, where the metric on which the definition of $W_1$ depends is defined by*

$$d_{[0,1]^n}(a,b) = \max_{1 \le i \le n} |a_i - b_i|, \text{ for all } a, b \in [0,1]^n.$$

*In particular, when $n = 1$, $\eta_1$ is an isometry.*

*Proof.* This follows from Theorem 5.3.6 and the fact that $d(p,q) \le d^*(p,q) \le nd(p,q)$ for all $p, q \in S_n(ARV)$. □

Now we apply model theory to prove the following result, which will be used in the proof of Theorem 5.3.9.

**5.3.8 Proposition.** *Let $(\Omega, \mathcal{F}, \mu)$ be an atomless probability space and let $f, g \colon \Omega \to [0,1]$ be random variables. Suppose $\mu\big(\operatorname{supp}(f)\big) \le \mu\big(\operatorname{supp}(g)\big)$.*

*(i) There are random variables $\tilde{f}, \tilde{g} \colon \Omega \to [0,1]$ such that $\operatorname{dist}(\tilde{f}) = \operatorname{dist}(f)$, $\operatorname{dist}(\tilde{g}) = \operatorname{dist}(g)$, and $d(\tilde{f}, \tilde{g}) = W_1\big(\operatorname{dist}(f), \operatorname{dist}(g)\big)$. Moreover, for any such random variables $\tilde{f}$ and $\tilde{g}$, one has $\operatorname{supp}(\tilde{f}) \subseteq \operatorname{supp}(\tilde{g})$ up to a null set.*

*(ii) Let $A \subseteq B$ be in $\mathcal{F}$ with $\mu(A) = \mu\big(\operatorname{supp}(f)\big)$ and $\mu(B) = \mu\big(\operatorname{supp}(g)\big)$. Then there are $f', g' \colon \Omega \to [0,1]$ satisfying:*

- $\operatorname{supp}(f') = A$ *and* $\operatorname{supp}(g') = B$*;*
- $\operatorname{dist}(f) = \operatorname{dist}(f')$ *and* $\operatorname{dist}(g) = \operatorname{dist}(g')$*;*
- $d(f', g') = W_1\big(\operatorname{dist}(f'), \operatorname{dist}(g')\big)$.

*Proof.* $(i)$: By Theorem 5.2.3, the theory ARV is separably categorical. Hence there are $\tilde{f}, \tilde{g} \colon \Omega \to [0,1]$ such that $\operatorname{tp}(f) = \operatorname{tp}(\tilde{f})$, $\operatorname{tp}(g) = \operatorname{tp}(\tilde{g})$, and $d(\tilde{f}, \tilde{g}) = d(\operatorname{tp}(f), \operatorname{tp}(g))$. Then by Theorem 5.2.3 and Theorem 5.3.6, one has that $\operatorname{dist}(\tilde{f}) = \operatorname{dist}(f)$, $\operatorname{dist}(\tilde{g}) = \operatorname{dist}(g)$, and $d(\tilde{f}, \tilde{g}) = d^*(\tilde{f}, \tilde{g}) = W_1\big(\operatorname{dist}(f), \operatorname{dist}(g)\big)$.

Suppose $\mathrm{supp}(\tilde{f})$ is not contained in $\mathrm{supp}(\tilde{g})$ up to a null set. Let $D = \mathrm{supp}(\tilde{f})\backslash\mathrm{supp}(\tilde{g})$, so $\mu(D) > 0$. Since $\mathrm{dist}(f) = \mathrm{dist}(\tilde{f})$ and $\mathrm{dist}(g) = \mathrm{dist}(\tilde{g})$, one has that $\mu(\mathrm{supp}(\tilde{f})) = \mu(\mathrm{supp}(f)) \leq \mu(\mathrm{supp}(\tilde{g})) = \mu(\mathrm{supp}(g))$. Because $\Omega$ is atomless, there is a measurable set $C \subseteq \mathrm{supp}(\tilde{g})\backslash\mathrm{supp}(\tilde{f})$ such that $\mu(C) = \mu(D)$. Let $\Omega_C$ and $\Omega_D$ denote the probability spaces $(C, \mathcal{F} \restriction C, \tilde{\mu})$ and $(D, \mathcal{F} \restriction D, \tilde{\mu})$ respectively, where $\tilde{\mu} = \frac{\mu}{\mu(C)} = \frac{\mu}{\mu(D)}$. Then $\tilde{f} \restriction D \in L^1(\Omega_D, [0,1])$. By Theorem 5.2.3, the theory $\mathsf{ARV}$ is separably categorical, so all types in $S_n(ARV)$ are realized in any model of $\mathsf{ARV}$. Thus $\mathrm{tp}(\tilde{f} \restriction D)$ is realized in $L^1(\Omega_C, [0,1])$, say by $h\colon C \to [0,1]$. Then by Theorem 5.2.3, $\mathrm{dist}(h) = \mathrm{dist}(\tilde{f} \restriction D)$. We define $\tilde{f}'\colon \Omega \to [0,1]$ as follows:

$$\tilde{f}'(\omega) = \begin{cases} h(\omega) & \omega \in C \\ 0 & \omega \in D \\ \tilde{f}(\omega) & \text{elsewhere} \end{cases}$$

Then $\mathrm{dist}(\tilde{f}) = \mathrm{dist}(\tilde{f}')$, since $\mu(C) = \mu(D)$ and $\mathrm{dist}(h) = \mathrm{dist}(\tilde{f} \restriction D)$. Furthermore,

$$\begin{aligned} d(\tilde{f}', \tilde{g}) = \int_\Omega |\tilde{f}' - \tilde{g}| d\mu &= \int_{\Omega\backslash(C\cup D)} |\tilde{f}' - \tilde{g}| d\mu + \int_C |\tilde{f}' - \tilde{g}| d\mu + \int_D |\tilde{f}' - \tilde{g}| d\mu \\ &= \int_{\Omega\backslash(C\cup D)} |\tilde{f} - \tilde{g}| d\mu + \int_C |h - \tilde{g}| d\mu + \int_D |0 - \tilde{g}| d\mu \\ &< \int_{\Omega\backslash(C\cup D)} |\tilde{f} - \tilde{g}| d\mu + \int_C (h + \tilde{g}) d\mu + 0 \\ &= \int_{\Omega\backslash(C\cup D)} |\tilde{f} - \tilde{g}| d\mu + \int_C \tilde{g} d\mu + \int_D \tilde{f} d\mu \\ &= \int_{\Omega\backslash(C\cup D)} |\tilde{f} - \tilde{g}| d\mu + \int_C |\tilde{f} - \tilde{g}| d\mu + \int_D |\tilde{f} - \tilde{g}| d\mu \\ &= \int_\Omega |\tilde{f} - \tilde{g}| d\mu = d(\tilde{f}, \tilde{g}) = W_1(\mathrm{dist}(f), \mathrm{dist}(g)). \end{aligned}$$

By Corollary 5.3.7, this contradicts the fact that $\mathrm{dist}(\tilde{f}') = \mathrm{dist}(f)$ and $\mathrm{dist}(\tilde{g}) = \mathrm{dist}(g)$.

$(ii)$: By $(i)$, there are random variables $\tilde{f}, \tilde{g}\colon \Omega \to [0,1]$ such that $\mathrm{dist}(\tilde{f}) = \mathrm{dist}(f)$, $\mathrm{dist}(\tilde{g}) = \mathrm{dist}(g)$, and $d(\tilde{f}, \tilde{g}) = W_1\big(\mathrm{dist}(f), \mathrm{dist}(g)\big)$. Moreover, $\mathrm{supp}(\tilde{f}) \subseteq \mathrm{supp}(\tilde{g})$ up to a null set. Let $C = \mathrm{supp}(\tilde{f})$, $D = \mathrm{supp}(\tilde{g})$, and assume that $C \subseteq D$. Then, one has that $\mu(C) = \mu(A)$ and $\mu(D) = \mu(B)$. Since $\mathcal{F}$ is atomless, there is an atomless and countably $\sigma$-generated $\sigma$-subalgebra $\mathcal{F}_0$ of $\mathcal{F}$ which contains $A, B, C, D$. Let $\mathcal{A} = \sigma(C \cup D)$ be the $\sigma$-subalgebra $\sigma$-generated by $C$ and $D$ in $\mathcal{F}_0$. Then we define $\varphi\colon \mathcal{A} \to \mathcal{F}_0$ by $\varphi(C) = A$ and $\varphi(D) = B$. Since $\mathcal{F}_0$ is atomless and countably $\sigma$-generated, $\mathcal{F}_0$ is a homogeneous probability algebra. By Corollary 2.2.13, $\varphi$ extends to an $L_{\mathsf{Pr}}$-automorphism $\Phi$ of $\mathcal{F}_0$. Naturally, $\Phi$ induces an $L_{\mathsf{RV}}$-automorphism of $L^1\big((\Omega, \mathcal{F}_0, \mu), [0,1]\big)$; we still call it $\Phi$. Let $f' = \Phi(\tilde{f})$ and $g' = \Phi(\tilde{g})$. Then $\mathrm{dist}(f') = \mathrm{dist}(\tilde{f})$, $\mathrm{dist}(g') = \mathrm{dist}(\tilde{g})$, and $d(f', g') = d(\tilde{f}, \tilde{g})$. Moreover, $\mathrm{supp}(f') = A$ and $\mathrm{supp}(g') = B$. $\square$

Next, for a special case, we prove an explicit formula for the $d$-metric between types.

**5.3.9 Theorem.** *Let $\kappa$ be an uncountable cardinal. Let $\mathcal{M} \models \mathsf{ARV}$ be a $\kappa$-universal domain of the form $\big(L^1(\mu,[0,1]),\neg,\frac{1}{2},\cdot,\dot{-},I\big)$, where $(\Omega,\mathcal{F},\mu)$ is an atomless probability space. Suppose $C \subseteq M = L^1(\mu,[0,1])$ is a small subset; let $\mathcal{C}$ be the $\sigma$-algebra of measurable sets generated by the random variables in $C$. Let $a=(a_1,\cdots,a_n)\in M^n$ and $b=(b_1,\cdots,b_n)\in M^n$ be disjointly supported; i.e., they satisfy $a_i\cdot a_j = b_i\cdot b_j = 0$ whenever $i\neq j$ and $1\leq i,j\leq n$. Then*

$$d\big(\mathrm{tp}(a/C),\mathrm{tp}(b/C)\big) = \max_{1\leq i\leq n}\int_0^1 \big\|\mathbb{P}(a_i > t \mid \mathcal{C}) - \mathbb{P}(b_i > t \mid \mathcal{C})\big\|_1 dt, \tag{5.1}$$

*where $\|\cdot\|_1$ is the $L^1$-norm.*

*Proof.* **Case 1**: We first consider the case in which $\mathcal{C}=\emptyset$. By definition,

$$d\big(\mathrm{tp}(a),\mathrm{tp}(b)\big) = \inf\big\{\max_{1\leq i\leq n} I\big(|x_i-y_i|\big) : x,y\in M, x\models \mathrm{tp}(a), \text{ and } y\models \mathrm{tp}(b)\big\}.$$

Then

$$\begin{aligned} d\big(\mathrm{tp}(a),\mathrm{tp}(b)\big) &= \inf\big\{\max_{1\leq i\leq n} I\big(|x_i-y_i|\big) : x,y\in M, x\models \mathrm{tp}(a), \text{ and } y\models \mathrm{tp}(b)\big\} \\ &\geq \max_{1\leq i\leq n} \inf\big\{I\big(|x_i-y_i|\big) : x,y\in M, x\models \mathrm{tp}(a), \text{ and } y\models \mathrm{tp}(b)\big\} \\ &= \max_{1\leq i\leq n}\int_0^1 \big|\mu(a_i>t)-\mu(b_i>t)\big|dt. \end{aligned}$$

So we need only show that

$$d\big(\mathrm{tp}(a),\mathrm{tp}(b)\big) \leq \max_{1\leq i\leq n}\int_0^1 \big|\mu(a_i>t)-\mu(b_i>t)\big|dt.$$

Let $\big([0,1],\mathscr{B},\lambda\big)$ be the standard Lebesgue space. Let $f:=(f_1,\cdots,f_n)$ and $g:=(g_1,\cdots,g_n)$ whose coordinates are elements of $L^1\big(\lambda,[0,1]\big)$, such that $\mathrm{dist}(f)=\mathrm{dist}(a)$ and $\mathrm{dist}(g)=\mathrm{dist}(b)$. Let $A_i$ be the support set for $f_i$ and $B_i$ for $g_i$, for every $i\leq n$. Reordering the indices if necessary, we may assume that there is $l\leq n$ such that $\mu(A_i)\geq\mu(B_i)$ iff $i\leq l$. Let $A_i'$ be the closed interval $\big[\sum_{j<i}\mu(A_j),\sum_{j\leq i}\mu(A_j)\big]$ for $i\leq n$. For $i\leq l$, let $B_i' := \big[\sum_{j<i}\mu(A_j),\sum_{j<i}\mu(A_j)+\mu(B_i)\big]$. Thus, $B_i'\subseteq A_i'$, for $i\leq l$. Since $a_1,\cdots,a_n$ are disjointly supported and the same for $b_1,\cdots,b_n$, we know that $f_1,\cdots,f_n$ are disjointly supported and the same for $g_1,\cdots,g_n$. For each $l<i\leq n$, let $B_i'$ be a union of closed intervals satisfying, $B_i'\supseteq A_i'$, $\lambda(B_i')=\mu(B_i)$, and $B_i'$ is disjoint from $B_j'$ for all $j\neq i$. For every $i\leq n$, by Proposition 5.3.8 there are

$f_i', g_i'\colon [0,1] \to [0,1]$ satisfying the following:

- $\mathrm{dist}(f_i') = \mathrm{dist}(f_i) = \mathrm{dist}(a_i)$, $\mathrm{dist}(g_i') = \mathrm{dist}(g_i) = \mathrm{dist}(b_i)$;
- The support set of $f_i'$ and $g_i'$ are $A_i'$ and $B_i'$, respectively;
- $d(f_i', g_i') = d\big(\mathrm{tp}(a_i), \mathrm{tp}(b_i)\big)$.

By Corollary 5.3.7 and Theorem 2.4.4, we have:

$$d(f_i', g_i') = d\big(\mathrm{tp}(a_i), \mathrm{tp}(b_i)\big) = \int_0^1 \big|\mu(a_i > t) - \mu(b_i > t)\big| dt.$$

Let $f' := (f_1', \cdots, f_n')$ and $g' := (g_1', \cdots, g_n')$. Note that $\mathrm{dist}(f') = \mathrm{dist}(f) = \mathrm{dist}(a)$ and $\mathrm{dist}(g') = \mathrm{dist}(g) = \mathrm{dist}(b)$ by the fact that they are disjointly supported. By Theorem 5.2.3, we have $\mathrm{tp}(f') = \mathrm{tp}(f) = \mathrm{tp}(a)$ and $\mathrm{tp}(g') = \mathrm{tp}(g) = \mathrm{tp}(b)$. Then,

$$d\big(\mathrm{tp}(a), \mathrm{tp}(b)\big) \leq \max_{1\leq i\leq n} d(f_i', g_i') = \max_{1\leq i\leq n} \int_0^1 \big|\mu(a_i > t) - \mu(b_i > t)\big| dt.$$

**Case 2**: We next consider the case in which $\mathcal{C}$ is finite with atoms $C_1, \cdots, C_m$. Let $a_{ij} = a_i \cdot \chi_{C_j}$ and $b_{ij} = b_i \cdot \chi_{C_j}$, for every $1 \leq j \leq m$ and every $1 \leq i \leq n$. Also for every $1 \leq j \leq m$, let $\bar{a}_j = (a_{1j}, \cdots, a_{nj})$ and $\bar{b}_j = (b_{1j}, \cdots, b_{nj})$. For each $1 \leq j \leq m$, by Proposition 5.3.8 there are $\bar{a}_j' = (a_{1j}', \cdots, a_{nj}')$ and $\bar{b}_j' = (b_{1j}', \cdots, b_{nj}')$, whose coordinates are supported in $C_j$ such that $\mathrm{tp}(\bar{a}_j') = \mathrm{tp}(\bar{a}_j)$, $\mathrm{tp}(\bar{b}_j') = \mathrm{tp}(\bar{b}_j)$, and $d(a_{ij}', b_{ij}') = W_1\big(\mathrm{dist}(a_{ij}'), \mathrm{dist}(b_{ij}')\big)$ for all $1 \leq i \leq n$. Then by Corollary 5.3.7,

$$d(a_{ij}', b_{ij}') = d\big(\mathrm{tp}(a_{ij}'), \mathrm{tp}(b_{ij}')\big) = \int_0^1 |\mu(a_{ij}' > t) - \mu(b_{ij}' > t)| dt \text{ for all } i \leq n.$$

Let $a_i' = \sum_{j=1}^m a_{ij}'$ and $b_i' = \sum_{j=1}^m b_{ij}'$. Let $a' = (a_1', \cdots, a_n')$ and $b' = (b_1', \cdots, b_n')$. For each $1 \leq j \leq m$,

$$\mathbb{P}(a_i > t \mid \mathcal{C}) = \sum_{j=1}^m \frac{\mu\big(\{a_i > t\} \cap C_j\big)\chi_{C_j}}{\mu(C_j)} = \sum_{j=1}^m \frac{\mu(a_{ij} > t)\chi_{C_j}}{\mu(C_j)}.$$

Thus we have $\mathrm{dist}\big(a_i \mid \sigma(C)\big) = \mathrm{dist}\big(a_i' \mid \sigma(C)\big)$ and $\mathrm{dist}\big(b_i \mid \sigma(C)\big) = \mathrm{dist}\big(b_i' \mid \sigma(C)\big)$. Then by Theorem 5.2.3($v$), we have $\mathrm{tp}(a_i/C) = \mathrm{tp}(a_i'/C)$ and $\mathrm{tp}(b_i/C) = \mathrm{tp}(b_i'/C)$. By the fact that the coordinates of $a$, $a'$, $b'$, and $b$ are disjointly supported, we have $\mathrm{tp}(a/C) = \mathrm{tp}(a'/C)$ and $\mathrm{tp}(b/C) = \mathrm{tp}(b'/C)$, since $\mathrm{dist}\big(a \mid \sigma(C)\big) = \mathrm{dist}\big(a' \mid \sigma(C)\big)$ and $\mathrm{dist}\big(b \mid \sigma(C)\big) = \mathrm{dist}\big(b' \mid \sigma(C)\big)$. For all $1 \leq i \leq n$, we have

$$d(a_i', b_i') = \int_0^1 |a_i' - b_i'| dt = \sum_{j=1}^m \int_{C_j} |a_i' - b_i'| dt = \sum_{j=1}^m \int_{C_j} |a_{ij}' - b_{ij}'| dt = \sum_{j=1}^m \int_0^1 |a_{ij}' - b_{ij}'| dt = \sum_{j=1}^m d(a_{ij}', b_{ij}').$$

By the fact that $d(a'_{ij}, b'_{ij}) = \int_0^1 |\mu(a'_{ij} > t) - \mu(b'_{ij} > t)|dt$, and $\mathrm{dist}\big(a \mid \sigma(C)\big) = \mathrm{dist}\big(a' \mid \sigma(C)\big)$ and $\mathrm{dist}\big(b \mid \sigma(C)\big) = \mathrm{dist}\big(b' \mid \sigma(C)\big)$, we have

$$d(a'_i, b'_i) = \sum_{j=1}^{m} d(a'_{ij}, b'_{ij}) = \sum_{j=1}^{m} \int_0^1 \big|\mu(a'_{ij} > t) - \mu(b'_{ij} > t)\big|dt = \sum_{j=1}^{m} \int_0^1 \big|\mu(a_{ij} > t) - \mu(b_{ij} > t)\big|dt. \quad (5.2)$$

For all $f, g \in M$ with $\mathrm{tp}(f/C) = \mathrm{tp}(a/C)$ and $\mathrm{tp}(g/C) = \mathrm{tp}(b/C)$, we have that for each $1 \leq i \leq n$,

$$\begin{aligned} d(f_i, g_i) = \sum_{j=1}^{m} d(f_{ij}, g_{ij}) \geq \sum_{j=1}^{m} d\big(\mathrm{tp}(f_{ij}), \mathrm{tp}(g_{ij})\big) &= \sum_{j=1}^{m} \int_0^1 \big|\mu(f_{ij} > t) - \mu(g_{ij} > t)\big|dt \\ &= \sum_{j=1}^{m} \int_0^1 \big|\mu(a_{ij} > t) - \mu(b_{ij} > t)\big|dt, \end{aligned}$$

where $f_{ij} = f_i \cdot \chi_{C_j}$ and $g_{ij} = g_i \cdot \chi_{C_j}$. By Equation (5.2), we have $d(f, g) \geq d(a', b')$, whereby $d(a', b') = d\big(\mathrm{tp}(a/C), \mathrm{tp}(b/C)\big)$. Then by

$$\mathbb{P}(a_i > t \mid \mathcal{C}) = \sum_{j=1}^{m} \frac{\mu(a_{ij} > t)\chi_{C_j}}{\mu(C_j)},$$

we have

$$\begin{aligned} \big\|\mathbb{P}(a_i > t \mid \mathcal{C}) - \mathbb{P}(b_i > t \mid \mathcal{C})\big\|_1 &= \int_\Omega \Big|\sum_{j=1}^{m} \big(\mu(a_{ij} > t) - \mu(b_{ij} > t)\big) \frac{\chi_{C_j}}{\mu(C_j)}\Big| d\mu \\ &= \sum_{j=1}^{m} \int_{C_j} \frac{\big|\mu(a_{ij} > t) - \mu(b_{ij} > t)\big|}{\mu(C_j)} d\mu \\ &= \sum_{j=1}^{m} \big|\mu(a_{ij} > t) - \mu(b_{ij} > t)\big|. \end{aligned} \quad (5.3)$$

By Equations (5.2) and (5.3), we have

$$d(a'_i, b'_i) = \int_0^1 \big\|\mathbb{P}(a_i > t \mid \mathcal{C}) - \mathbb{P}(b_i > t \mid \mathcal{C})\big\|_1 dt,$$

whereby

$$d\big(\mathrm{tp}(a/C), \mathrm{tp}(b/C)\big) = d(a', b') = \max_{1 \leq i \leq n} \int_0^1 \big\|\mathbb{P}(a_i > t \mid \mathcal{C}) - \mathbb{P}(b_i > t \mid \mathcal{C})\big\|_1 dt.$$

**Case 3**: Finally, suppose $\mathcal{C}$ is an infinite $\sigma$-algebra. Let $\{\mathcal{C}_j \mid j \in J\}$ be the family of all finite subalgebras of $\mathcal{C}$. Then $\mathcal{C} = \cup_{j \in J} \mathcal{C}_j$. Since all measurable functions are approximated by simple functions, $\mathbb{P}(a_i > t \mid \mathcal{C}) = \sup_{j \in J} \mathbb{P}(a_i > t \mid \mathcal{C}_j)$ *a.s.* and $\mathbb{P}(b_i > t \mid \mathcal{C}) = \sup_{j \in J} \mathbb{P}(b_i > t \mid \mathcal{C}_j)$ *a.s.* for each $1 \leq i \leq n$. Further,

$$\big\|\mathbb{P}(a_i > t \mid \mathcal{C}) - \mathbb{P}(b_i > t \mid \mathcal{C})\big\|_1 = \sup_{j \in J} \big\|\mathbb{P}(a_i > t \mid \mathcal{C}_j) - \mathbb{P}(b_i > t \mid \mathcal{C}_j)\big\|_1.$$

By the compactness theorem 1.3.2, we know that

$$d\big(\mathrm{tp}(a/C), \mathrm{tp}(b/C)\big) = \sup_{j\in J} d\big(\mathrm{tp}(a/\mathcal{C}_j), \mathrm{tp}(b/\mathcal{C}_j)\big).$$

Thus, by **Case 2** we have

$$d\big(\mathrm{tp}(a/C), \mathrm{tp}(b/C)\big) = \max_{1\le i\le n} \int_0^1 \big\|\mathbb{P}(a_i > t \mid \mathcal{C}) - \mathbb{P}(b_i > t \mid \mathcal{C})\big\|_1 dt.$$

□

**5.3.10 Remark.** *(i) In Case 3, one has* $d\big(\mathrm{tp}(a/C), \mathrm{tp}(b/C)\big) = \sup_{j\in J} d\big(\mathrm{tp}(a/\mathcal{C}_j), \mathrm{tp}(b/\mathcal{C}_j)\big)$ *holds in general and does not need assumptions on* $a$ *and* $b$ *in the lemma.*

*(ii) If* $a$ *and* $b$ *are just elements of* $M$*, then Theorem 5.3.9 yields the following formula:*

$$d\big(\mathrm{tp}(a/C), \mathrm{tp}(b/C)\big) = \int_0^1 \big\|\mathbb{P}(a > t|\mathcal{C}) - \mathbb{P}(b > t|\mathcal{C})\big\|_1 dt.$$

*(iii) It follows from the proof of Theorem 5.3.9 that*

$$d^*\big(\mathrm{tp}(a/C), \mathrm{tp}(b/C)\big) = \sum_{i=1}^{n} \int_0^1 \big\|\mathbb{P}(a_i > t|\mathcal{C}) - \mathbb{P}(b_i > t|\mathcal{C})\big\|_1 dt,$$

*for all disjointly supported* $a = (a_1, \cdots, a_n), b = (b_1, \cdots, b_n) \in M^n$*. Indeed, every occurrence of* $\max_{1\le i\le n}$ *in the proof can be replaced by* $\sum_{i=1}^n$*.*

**5.3.11 Question.** *We see that by Theorem 5.3.9, for* $a, b \in M^n$*, being disjointly supported is a sufficient condition for Equation (5.1). Does Equation (5.1) hold for a more general class of* $a, b \in M^n$*? We know that Equation (5.1) does not hold for all* $a, b \in M^n$*, since in general the joint distributions of* $a$ *and* $b$ *are not determined by the individual distributions of their coordinates, while the right hand side of Equation (5.1) only depends on the distributions of* $a$ *and* $b$*'s coordinates.*

Now we give some results that will be used to prove Theorem 5.3.14, which considers arbitrary $n$-types over parameters in ARV. The following lemma is needed for Proposition 5.3.13.

**5.3.12 Lemma.** *Let* $\mathcal{M} \models \mathsf{ARV}$ *be of the form* $M = L^1\big(\mu, [0,1]\big)$ *where* $(\Omega, \mathcal{F}, \mu)$ *is an atomless probability space, and let* $C$ *be a subset of* $M$*. Let* $\mathcal{C}$ *denote the* $\sigma$*-algebra of measurable sets generated by the random variables in* $C$*. Suppose* $\mathcal{C}$ *is finite with atoms* $C_1, \cdots, C_m$*. Let* $a = (a_1, \cdots, a_n)$ *and* $b = (b_1, \cdots, b_n)$ *be*

*tuples in $M$. Then*

$$\mathrm{tp}(a/C) = \mathrm{tp}(b/C) \iff \mathrm{tp}\Big(\big(\{a_{ij}\}_{1\le i\le n, 1\le j\le m}\big)\Big) = \mathrm{tp}\Big(\big(\{b_{ij}\}_{1\le i\le n, 1\le j\le m}\big)\Big),$$

*where $a_{ij} = a_i \cdot \chi_{C_j}$ and $b_{ij} = b_i \cdot \chi_{C_j}$ for all $1 \le i \le n$ and $1 \le j \le m$.*

*Proof.* We need only show from right to left. For all Borel $B \subseteq [0,1]^n$, we have

$$\begin{aligned}
\mathbb{P}\big((a_1, \cdots, a_n) \in B \mid \mathcal{C}\big) &= \frac{\mu\big(C_1 \cap \{(a_1, \cdots, a_n) \in B\}\big)}{\mu(C_1)}\chi_{C_1} + \cdots + \frac{\mu\big(C_n \cap \{(a_1, \cdots, a_n) \in B\}\big)}{\mu(C_n)}\chi_{C_n} \\
&= \frac{\mu((a_{11}, \cdots, a_{n1}) \in B)}{\mu(C_1)}\chi_{C_1} + \cdots + \frac{\mu((a_{1m}, \cdots, a_{nm}) \in B)}{\mu(C_m)}\chi_{C_m} \\
&= \frac{\mu((b_{11}, \cdots, b_{n1}) \in B)}{\mu(C_1)}\chi_{C_1} + \cdots + \frac{\mu((b_{1m}, \cdots, b_{nm}) \in B)}{\mu(C_m)}\chi_{C_m} \\
&= \frac{\mu\big(C_1 \cap \{(b_1, \cdots, b_n) \in B\}\big)}{\mu(C_1)}\chi_{C_1} + \cdots + \frac{\mu\big(C_n \cap \{(b_1, \cdots, b_n) \in B\}\big)}{\mu(C_n)}\chi_{C_n} \\
&= \mathbb{P}\big((b_1, \cdots, b_n) \in B \mid \mathcal{C}\big).
\end{aligned}$$

Then by Theorem 5.3.3, we know $\mathrm{tp}\big((a_1, \cdots, a_n)/C\big) = \mathrm{tp}\big((b_1, \cdots, b_n)/C\big)$. □

The following proposition will be used in the proof of Theorem 5.3.14.

**5.3.13 Proposition.** *Let $\mathcal{M} \models \mathsf{ARV}$ be an $\aleph_0$-universal domain and let $C$ be a subset of $M$. Suppose that $M = L^1\big(\mu, [0,1]\big)$ for the atomless probability space $(\Omega, \mathcal{F}, \mu)$. Let $\mathcal{C}$ denote the $\sigma$-algebra of measurable sets generated by all the random variables in $C$. Suppose $\mathcal{C}$ is finite with atoms $C_1, \cdots, C_m$. For all $n$-tuples $a = (a_1, \cdots, a_n)$ in $M$, define $a_{ij} := a_i \cdot \chi_{C_j}$ for all $1 \le i \le n$ and $1 \le j \le m$. The $mn$-tuple $(a_{11}, \cdots, a_{1m}, \cdots, a_{n1}, \cdots, a_{nm})$ is denoted by $\big(\{a_{ij}\}_{1\le i\le n, 1\le j\le m}\big)$. Then the mapping*

$$\begin{aligned}
\theta_{\mathcal{C}} \colon S_n(ARV_C) &\to S_{mn}(ARV) \\
\mathrm{tp}(a/C) &\mapsto \mathrm{tp}\Big(\big(\{a_{ij}\}_{1\le i\le n, 1\le j\le m}\big)\Big)
\end{aligned}$$

*is isometric from $(S_n(ARV_C), d^*)$ into $(S_{mn}(ARV), d^*)$.* □

*Proof.* By Lemma 5.3.12, $\theta_{\mathcal{C}}$ is well-defined. We will show that $\theta_{\mathcal{C}}$ is isometric. Let $a = (a_1, \cdots, a_n)$ and $b = (b_1, \cdots, b_n)$ be $n$-tuples in $M$. We define $a_{ij} = a_i \cdot \chi_{C_j}$ and $b_{ij} = b_i \cdot \chi_{C_j}$ for all $1 \le i \le n$ and $1 \le j \le m$. Since $C_1, \cdots, C_m$ are the atoms of $\mathcal{C}$, for all $1 \le i \le n$ we have $a_i = a_{i1} + \cdots + a_{im}$, $b_i = b_{i1} + \cdots + b_{im}$, and

$d^*(a_i, b_i) = \sum_{j=1}^m d^*(a_{ij}, b_{ij})$. Then

$$d^*(a,b) = \sum_{i=1}^n \sum_{j=1}^m d^*(a_{ij}, b_{ij}) = d^*\Big(\big(\{a_{ij}\}_{1\le i\le n, 1\le j\le m}\big), \big(\{b_{ij}\}_{1\le i\le n, 1\le j\le m}\big)\Big).$$

By Lemma 5.3.12, we get

$$d^*\big(\mathrm{tp}(a/C), \mathrm{tp}(b/C)\big) = d^*\Big(\mathrm{tp}\big(\{a_{ij}\}_{1\le i\le n, 1\le j\le m}\big), \mathrm{tp}\big(\{b_{ij}\}_{1\le i\le n, 1\le j\le m}\big)\Big),$$

whereby

$$d^*(p,q) = d^*\big(\theta_{\mathcal{C}}(p), \theta_{\mathcal{C}}(q)\big) \text{ for all } p, q \in S_n(ARV_C).$$

□

Now, we use Theorem 2.4.3 (the Kantorovich-Rubinstein duality formula) to give a formula for the $d^*$-metric between $n$-types over parameters in the most general situation,

**5.3.14 Theorem.** *Let $\kappa$ be an uncountable cardinal. Let $\mathcal{M} \models \mathsf{ARV}$ be a $\kappa$-universal domain of the form $\big(L^1(\mu, [0,1]), \neg, \frac{1}{2}, \cdot, \dot{-}, I\big)$, where $(\Omega, \mathcal{F}, \mu)$ is an atomless probability space. Suppose $C \subseteq M = L^1(\mu, [0,1])$ is a small subset; let $\mathcal{C}$ be the $\sigma$-algebra of measurable sets generated by the random variables in $C$. Let $\{\mathcal{C}_j \mid j \in J\}$ be the family of all finite subalgebras of $\mathcal{C}$. Then $\mathcal{C} = \cup_{j\in J}\mathcal{C}_j$. For all $p, q \in S_n(ARV_C)$ and for all $j \in J$, let $p_j$ and $q_j$ be the restriction of $p$ and $q$ in $S_n(ARV_{\mathcal{C}_j})$ respectively. Then let $\nu_j = \eta_{mn} \circ \theta_{\mathcal{C}_j}(p_j)$ and $\xi_j = \eta_{mn} \circ \theta_{\mathcal{C}_j}(q_j)$, where $\eta_{mn}$ and $\theta_{\mathcal{C}}$ are from Theorem 5.3.6 and Proposition 5.3.13. Then*

$$d^*(p,q) = \sup_{\|\psi\|_{Lip}\le 1} \sup_{j\in J} \Big\{ \int_{[0,1]^{mn}} \psi d\nu_j - \int_{[0,1]^{mn}} \psi d\xi_j \Big\},$$

*where $\|\psi\|_{Lip} \le 1$ means*

$$|\psi(x) - \psi(y)| \le \sum_{i=1}^{mn} |x_i - y_i|, \text{ for all } x = (x_1, \cdots, x_{mn}), y = (y_1, \cdots, y_{mn}) \in [0,1]^{mn}.$$

*Proof.* Let $j \in J$. By Proposition 5.3.13, we have $d^*(p_j, q_j) = d^*\big(\theta_{\mathcal{C}_j}(p_j), \theta_{\mathcal{C}_j}(q_j)\big)$. By Theorem 5.3.6, we have $d^*\big(\theta_{\mathcal{C}_j}(p_j), \theta_{\mathcal{C}_j}(q_j)\big) = W_1\big(\eta_{mn} \circ \theta_{\mathcal{C}_j}(p_j), \eta_{mn} \circ \theta_{\mathcal{C}_j}(q_j)\big) = W_1(\nu_j, \xi_j)$. By Theorem 2.4.3, we have

$$W_1(\nu_j, \xi_j) = \sup_{\|\psi\|_{Lip}\le 1} \Big\{ \int_{[0,1]^{mn}} \psi d\nu_j - \int_{[0,1]^{mn}} \psi d\xi_j \Big\},$$

where $\|\psi\|_{Lip} \le 1$ means $|\psi(x) - \psi(y)| \le \sum_{i=1}^{mn} |x_i - y_i|$, for all $x = (x_1, \cdots, x_{mn})$ and $y = (y_1, \cdots, y_{mn}) \in$

$[0,1]^{mn}$.

Then, by Remark 5.3.10$(i)$ we have

$$d^*(p,q) = \sup_{j\in J} d^*(p_j,q_j) = \sup_{\|\psi\|_{Lip}\le 1} \sup_{j\in J}\Big\{\int_{[0,1]^{mn}} \psi d\nu_j - \int_{[0,1]^{mn}} \psi d\xi_j\Big\}.$$

□

## 5.4 A new proof of an integral formula for the Wasserstein distance

In this section, using rearrangements and model theoretic results from Section 5.3, we give a new proof of Theorem 2.4.4 (see Theorem 5.4.2 below).

First, we need the following lemma, which is interesting in itself.

**5.4.1 Lemma.** *Let $f, g$ be simple, nonincreasing, right-continuous functions from the standard Lebesgue space $\big([0,1], \mathscr{B}, \lambda\big)$ to $[0,1]$. Then*

$$\int_0^1 \big|f(x) - g(x)\big|dx = \int_0^1 \big|\lambda(f > t) - \lambda(g > t)\big|dt.$$

*Proof.* Let $D$ be the set

$$\big\{(x,y) \mid x,y \in [0,1], f(x) \le y \le g(x) \text{ or } g(x) \le y \le f(x)\big\},$$

and let $I(D)$ be the set

$$I(D) = \big\{(x,y) \mid x,y \in [0,1], f(x) < y < g(x) \text{ or } g(x) < y < f(x)\big\}.$$

Since $f, g$ are simple, nonincreasing, right-continuous functions, we know that $I(D) \subseteq D$, and there is a set $S$, which is a finite union of rectangles, such that $I(D)$ and $D$ are equal to $S$ up to null sets. Therefore, $(\lambda\times\lambda)(D) = (\lambda\times\lambda)\big(I(D)\big)$. Hence,

$$\int_0^1 \big|f(x) - g(x)\big|dx = (\lambda\times\lambda)(D) = (\lambda\times\lambda)\big(I(D)\big).$$

Note that $\lambda(f > s)$ and $\lambda(g > s)$ are also simple, nonincreasing, right-continuous functions. Let $D'$ be the

set

$$\big\{(s,t) \mid s,t \in [0,1], \lambda(f > s) \le t \le \lambda(g > s) \text{ or } \lambda(g > s) \le t \le \lambda(f > s)\big\},$$

and let $I(D')$ be the set

$$I(D') = \big\{(s,t) \mid s,t \in [0,1], \lambda(f > s) < t < \lambda(g > s) \text{ or } \lambda(g > s) < t < \lambda(f > s)\big\}.$$

Since $\lambda(f > s)$ and $\lambda(g > s)$ are simple, nonincreasing, right-continuous functions, we know that $I(D') \subseteq D'$, and there is a set $S'$, which is a finite union of rectangles, such that $I(D')$ and $D'$ are equal to $S'$ up to null sets. Therefore, $(\lambda \times \lambda)(D') = (\lambda \times \lambda)\big(I(D')\big)$. Hence

$$\int_0^1 \big|\lambda(f > t) - \lambda(g > t)\big| dt = (\lambda \times \lambda)(D') = (\lambda \times \lambda)\big(I(D')\big).$$

We will show that if $(x,y) \in I(D)$ then $(y,x) \in D'$, and if $(s,t) \in I(D')$ then $(t,s) \in D$.

If $(x,y) \in I(D)$, then $f(x) < y < g(x)$ or $g(x) < y < f(x)$. For all $u, v \in [0,1]$ and all simple, nonincreasing, right-continuous functions $h\colon [0,1] \to [0,1]$, elementary arguments yield the following facts:

$$\lambda(h > v) \le u \text{ if and only if } h(u) \le v; \tag{5.4}$$

$$\lambda(h > v) > u \text{ if and only if } h(u) > v. \tag{5.5}$$

Suppose $f(x) < y < g(x)$. Then, by (5.4), $\lambda(f > y) \le x$; by (5.5), $\lambda(g > y) > x$. Hence $\lambda(f > y) \le x < \lambda(g > y)$, whereby $(y,x) \in D'$. If $g(x) < y < f(x)$, then by (5.4), $\lambda(g > y) \le x$; by (5.5), $\lambda(f > y) > x$. Hence, $\lambda(g > y) \le x < \lambda(f > y)$, whereby $(y,x) \in D'$. Therefore, $I(D) \subseteq D'$.

If $(s,t) \in I(D')$, then $\lambda(f > s) < t < \lambda(g > s)$ or $\lambda(g > s) < t < \lambda(f > s)$. Suppose $\lambda(f > s) < t < \lambda(g > s)$. By (5.4), $f(t) \le s$ and by (5.5), $g(t) > s$. Hence, $f(t) \le s < g(t)$, whereby $(t,s) \in D$. If $\lambda(g > s) < t < \lambda(f > s)$, then by (5.4), $g(t) \le s$ and by (5.5), $f(t) > s$. Hence, $g(t) \le s < f(t)$, whereby $(t,s) \in D$. Therefore, $I(D') \subseteq D$.

Hence $(\lambda \times \lambda)\big(I(D)\big) \le (\lambda \times \lambda)(D')$ and $(\lambda \times \lambda)\big(I(D')\big) \le (\lambda \times \lambda)(D)$. Since $(\lambda \times \lambda)(D) = (\lambda \times \lambda)\big(I(D)\big)$ and $(\lambda \times \lambda)(D') = (\lambda \times \lambda)\big(I(D')\big)$, we have $(\lambda \times \lambda)(D) = (\lambda \times \lambda)(D')$, whereby

$$\int_0^1 \big|f(x) - g(x)\big| dx = \int_0^1 \big|\lambda(f > t) - \lambda(g > t)\big| dt.$$

□

Now we are ready to prove:

**5.4.2 Theorem** (Theorem 2.4.4 revisited)**.** *For all $\mu, \nu \in P_1(\mathbb{R})$, one has*

$$W_1(\mu, \nu) = \int_{-\infty}^{+\infty} |F(x) - G(x)| dx, \tag{5.6}$$

*where $F$ and $G$ are distribution functions for $\mu$ and $\nu$ respectively.*

*Proof.* **Case 1**: We consider the case in which $\mu, \nu \in P_1([0,1]) \subseteq P_1(\mathbb{R})$. Let $([0,1], \mathscr{B}, \lambda)$ be the standard Lebesgue space. Let $f, g\colon [0,1] \to [0,1]$ be measurable functions such that $\mathrm{dist}(f) = \mu$ and $\mathrm{dist}(g) = \nu$. Then by Theorem 5.3.6, $W_1(\mu, \nu) = d(\mathrm{tp}(f), \mathrm{tp}(g))$. By Proposition 5.3.8, we may further assume that $d(f,g) = d(\mathrm{tp}(f), \mathrm{tp}(g)) = I\big(|f-g|\big) = \int_0^1 |f(x) - g(x)| dx$. Let $F(t) = \lambda(f \leq t)$ and $G(t) = \lambda(g \leq t)$.

First, suppose $f$ and $g$ are simple functions. We consider the decreasing rearrangement $f^*$ of $f$ and the decreasing rearrangement $g^*$ of $g$ from $[0,1]$ to $[0,1]$; for the definition, see Section 2.3. By Fact 2.3.4, $\mathrm{dist}(f^*) = \mathrm{dist}(f)$, $\mathrm{dist}(g^*) = \mathrm{dist}(g)$, and $f^*$ and $g^*$ are nonincreasing and lower semi-continuous. Because $f$ and $g$ are simple and $\mathrm{dist}(f) = \mathrm{dist}(f^*), \mathrm{dist}(g) = \mathrm{dist}(g^*)$, we have $f^*$ and $g^*$ are also simple. Thus, by Fact 2.3.4 we have that $f^*$ and $g^*$ are simple, nonincreasing and lower semi-continuous, whereby they are right-continuous. Since $J(x) = |x|$ is a nonnegative convex function with $J(0) = 0$, by Corollary 2.3.5 we have

$$\int_0^1 \big|f^*(x) - g^*(x)\big| dx \leq \int_0^1 \big|f(x) - g(x)\big| dx = d\big(\mathrm{tp}(f), \mathrm{tp}(g)\big).$$

Since $\mathrm{dist}(f^*) = \mathrm{dist}(f) = \mu$ and $\mathrm{dist}(g^*) = \mathrm{dist}(g) = \nu$, we have $W_1(\mu, \nu) \leq I\big(|f^* - g^*|\big)$, whereby $I\big(|f^* - g^*|\big) = d\big(\mathrm{tp}(f), \mathrm{tp}(g)\big) = W_1(\mu, \nu)$. Therefore, Equation (5.6) holds by Lemma 5.4.1.

Second, we consider general $[0,1]$-valued random variables $f$ and $g$. For all $h_1, h_2 \in L^1\big([0,1], [0,1]\big)$,

$$W_1(\mathrm{dist}(h_1), \mathrm{dist}(h_2)) \leq d(h_1, h_2) = \int_0^1 |h_1(x) - h_2(x)| dx.$$

Since the set of simple functions is dense in $L^1\big(\lambda, [0,1]\big)$, one has that the set $S = \{\mathrm{dist}(h) \mid h\colon [0,1] \to [0,1]$ is a simple function$\}$ is dense in $(P_1([0,1]), W_1)$. In the product space $P_1([0,1]) \times P_1([0,1])$, we consider the maximum metric induced by $W_1$. Then the set of pairs $(\mu, \nu)$ satisfying Equation (5.6) is a closed subset of $P_1([0,1]) \times P_1([0,1])$. We already know that all pairs in $S \times S$ satisfy Equation (5.6). Then by the fact that $S$ is dense in $(P_1([0,1]), W_1)$, we know that for all $\mu, \nu$ in $P_1([0,1])$, Equation (5.6) holds.

**Case 2**: We consider the case in which $\mu, \nu \in P_1([-n,n]) \subseteq P_1(\mathbb{R})$ for some $n \in \mathbb{N}$. We define a function $\varphi\colon \mathbb{R} \to \mathbb{R}$ by $\varphi(x) = \frac{x+n}{2n}$. This function $\varphi$ induces $\Phi\colon P_1([-n,n]) \to P_1([0,1])$ by defining $\Phi(\mu)(A) = \mu(\varphi^{-1}(A))$, where $\mu$ is in $P_1([-n,n])$ and $A$ is a Lebesgue measurable subset of $[0,1]$. Let $f$

and $g$ be any two random variables with $\mathrm{dist}(f) = \mu$ and $\mathrm{dist}(g) = \nu$. Then $\mathrm{dist}(\varphi \circ f) = \mathrm{dist}(\Phi(\mu))$ and $\mathrm{dist}(\varphi \circ g) = \mathrm{dist}(\Phi(\nu))$. Also $\|f - g\|_1 = 2n\|\varphi \circ f - \varphi \circ g\|_1$. Hence $2nW_1(\Phi(\mu), \Phi(\nu)) \leq W_1(\mu, \nu)$. Similarly, one has $2nW_1(\Phi(\mu), \Phi(\nu)) \geq W_1(\mu, \nu)$, whereby $2nW_1(\Phi(\mu), \Phi(\nu)) = W_1(\mu, \nu)$. Let $\Phi(F)$ and $\Phi(G)$ denote the distribution functions for $\Phi(\mu)$ and $\Phi(\nu)$ respectively. Then $\Phi(F)(x) = \Phi(\mu)((-\infty, x]) = \mu(\varphi^{-1}(-\infty, x]) = \mu((-\infty, 2nx - n]) = F(2nx - n)$ for $x \in \mathbb{R}$. Similarly, $\Phi(G)(x) = G(2nx - n)$. Then by substitution, one has

$$\int_{\mathbb{R}} |F - G| dx = 2n \int_{\mathbb{R}} |\Phi(F) - \Phi(G)| dx.$$

By Case 1, one has

$$W_1(\Phi(\mu), \Phi(\nu)) = \int_{\mathbb{R}} |\Phi(F) - \Phi(G)| dx.$$

Therefore, Equation (5.6) holds for $\mu$ and $\nu$.

**Case 3**: The most general case. Let $T$ denote $\bigcup_{n\in\mathbb{N}} P_1([-n, n])$. Then $T$ is dense in $(P_1(\mathbb{R}), W_1)$. In the product space $P_1(\mathbb{R}) \times P_1(\mathbb{R})$, we consider the maximum metric induced by $W_1$. Then the set of pairs $(\mu, \nu)$ satisfying Equation (5.6) is a closed subset of $P_1(\mathbb{R}) \times P_1(\mathbb{R})$. By Case 2, we know that all pairs in $T \times T$ satisfy Equation (5.6). Then by the fact that $T$ is dense in $(P_1(\mathbb{R}), W_1)$, we know that for all $\mu, \nu$ in $P_1(\mathbb{R})$, Equation (5.6) holds. □

## 5.5 Saturated models of ARV

In this section, we characterize the saturated models of ARV.

The theory *ARV* is a separably categorical theory, then by Theorem 1.9.3 (Ryll-Nardzewski Theorem), we know that the unique separable model of *ARV* is approximately $\aleph_0$-saturated. However, the following theorem shows that it is not $\aleph_0$-saturated.

**5.5.1 Theorem.** *No separable model of* ARV *is* $\aleph_0$*-saturated.*

*Proof.* By Theorem 5.2.3, the theory ARV is $\aleph_0$-categorical. We may assume the separable model $\mathcal{M}$ is of the form $L^1(\lambda, [0, 1])$, where $\big([0, 1], \mathscr{B}, \lambda\big)$ is the standard Lebesgue space. Then $(\Omega, \mathcal{F}, \mu) = \big([0, 1] \times \{0, 1\}, \mathscr{B} \otimes P(\{0, 1\}), \lambda \otimes \mu_0\big)$ is also a separable probability space, where $\mu_0\big(\{0\}\big) = \mu_0\big(\{1\}\big) = \frac{1}{2}$ and $P(\{0, 1\})$ is the power set of $\{0, 1\}$. Let $\{C_n \mid n \in \mathbb{N}\}$ be a countable open basis for $\mathscr{B}$ and let $\alpha$ be a measure-preserving isomorphism between probability spaces $\prod_{n\in\mathbb{N}}(\{0, 1\}, P(\{0, 1\}), \mu_0)$ and $([0, 1], \mathscr{B}, \lambda)$. Define $f\colon [0, 1] \to [0, 1]$ as follows:

$$f(t) = \alpha\Big(\big(\chi_{C_1}(t), \chi_{C_2}(t), \cdots, \chi_{C_n}(t), \cdots\big)\Big) \text{ for all } t \in [0, 1].$$

Then we know $f \in L^1(\lambda, [0,1])$. Now we define functions $f', g'$ from $\Omega$ to $[0,1]$ by $f'(t,i) = f(t)$ and $g'(t,i) = i$ for all $(t,i) \in \Omega = [0,1] \times \{0,1\}$. Clearly, they are both in $L^1\big(\lambda, [0,1]\big)$. Note that $\mathrm{dist}(f') = \mathrm{dist}(f)$, which implies $\mathrm{tp}(f') = \mathrm{tp}(f)$ by Theorem 5.2.3. Let $p_0(x,y)$ stand for $\mathrm{tp}(f', g')$. Since $\mathrm{tp}(f') = \mathrm{tp}(f)$, we get that $p_0(f,y)$ is consistent. If $p_0(f,y)$ is realized in $M$ by $g$, then because $g'^2 = g'$, $I(g') = \frac{1}{2}$ and $\mathrm{tp}(g) = \mathrm{tp}(g')$, we have $g = \chi_D$ for some $D \in \mathscr{B}$ and $\lambda(D) = \frac{1}{2}$. For all $n \in \mathbb{N}$, let $D_n$ be the set $\big\{x \in \{0,1\}^{\mathbb{N}} \mid x(n) = 1\big\}$. Then $\alpha(D_n) \in \mathscr{B}$. Now we have

$$\begin{aligned}\mu\big(f' \in \alpha(D_n); g' = 1\big) &= \mu\big((t,i) \in \Omega \mid t \in C_n, i = 1\big) = \frac{1}{2}\lambda(C_n) = \lambda(D)\lambda(C_n) \\ &= \lambda\big(f \in \alpha(D_n); g = 1\big) = \lambda(D \cap C_n).\end{aligned}$$

Therefore $D$ is independent from $C_n$ for all $n \in \mathbb{N}$. Similarly, we get that for all finite $J \subset \mathbb{N}$, the set $D$ is independent from $\bigcap_{j\in J} C_j$. Then by [12, page 26], $D$ is independent from $\sigma\big(\{C_n\}_{n\in\mathbb{N}}\big) = \mathscr{B}$. Hence, $\lambda(D \cap D) = \lambda(D)^2$. Consequently, $\lambda(D) = 0$ or 1, which contradicts $\lambda(D) = \frac{1}{2}$. □

**5.5.2 Remark.** *This proof borrows ideas from [13, Theorem 3B.1].*

**5.5.3 Theorem.** *Let $\mathcal{M}$ be a model of* ARV. *Suppose $M = L^1\big(\mu, [0,1]\big)$ for the atomless probability space $(\Omega, \mathscr{B}, \mu)$. For every uncountable cardinal $\kappa$, the structure $\mathcal{M} \models$ ARV is $\kappa$-saturated if and only if $\mathscr{B}$ is a $\kappa$-saturated probability algebra.*

*Proof.* $\Longrightarrow$: Suppose $\mathcal{M}$ is $\kappa$-saturated. We want to show that $\mathscr{B}$ is also $\kappa$-saturated. For all $A \subseteq \mathscr{B}$ with $|A| < \kappa$, and all $p(x) = \mathrm{tp}(x/A) \in S_1^{\mathsf{APr}}(A)$, let $\widehat{A} := \{\chi_a \mid a \in A\} \subseteq M$. Then $|\widehat{A}| < \kappa$. Now, let $\widehat{p}(\widehat{x}) = \mathrm{tp}(\widehat{x}/\widehat{A})$ such that $\widehat{x}^2 = \widehat{x}$. If $\phi(x, A) \in p(x)$, then $\phi(\widehat{x}, \widehat{A}) \in \widehat{p}(\widehat{x})$. By the assumption, we know that $\widehat{p}(\widehat{x}) \in S_1^{\mathsf{ARV}}(\widehat{A})$ is realized in $M$ by $a$. As $a^2 = a$, we know that $a$ is an characteristic function for a measurable set $C \in \mathscr{B}$, that is $\chi_C = a$. Then it is clear that $C \models \mathrm{tp}(x/A)$. Hence, $\mathscr{B}$ is $\kappa$-saturated.

$\Longleftarrow$: Let $A \subseteq M$ with $|A| < \kappa$ and $\mathrm{tp}(x/A) \in S_1^{\mathsf{ARV}}(A)$. Suppose $\mathrm{tp}(x/A)$ is realized by $f$ in $\mathcal{N}$, which is an elementary extension of $\mathcal{M}$. For every $t \in [0,1]$, the conditional probability $\mathbb{P}\big(f > t \mid \sigma(A)\big)$ is a $\sigma(A)$-measurable function. Since $\kappa$ is uncountable, $\sigma(A)$ is $\sigma$-generated by $\aleph_0 + |A| < \kappa$ many elements. By Theorem 3.4.2, $\mathscr{B}$ is $\kappa$-atomless, therefore it is atomless over $\sigma(A)$. Thus by Lemma 2.2.19, there exists a $\mathscr{B}$-measurable random variable $g\colon \Omega \to [0,1]$ such that for every $t \in [0,1]$,

$$\mathbb{P}\big(g > t \mid \sigma(A)\big) = \mathbb{P}\big(f > t \mid \sigma(A)\big) \text{ almost surely.}$$

By Theorem 5.2.3, $\mathrm{tp}(f/A) = \mathrm{tp}(g/A)$. Therefore, $\mathrm{tp}(x/A) = \mathrm{tp}(f/A)$ is realized by $g$ in $M$. □

**5.5.4 Theorem.** *Let $\mathcal{M} = L^1\big(\mu, [0,1]\big) \models \mathsf{ARV}$, where $(\Omega, \mathscr{B}, \mu)$ is an atomless probability space. Then the following are equivalent:*

*(i)* $\Omega$ *is Hoover-Keisler saturated.*

*(ii)* $\Omega$ *is* $\aleph_1$*-saturated.*

*(iii) Every cardinal in the Maharam spectrum of* $\Omega$ *is uncountable.*

*(iv)* $\mathcal{M}$ *is* $\aleph_1$*-saturated.*

*(v)* $\mathcal{M}$ *is* $\aleph_0$*-saturated.*

*(vi) For all elements* $a, b, c \in M$ *with* $\mathrm{tp}(a) = \mathrm{tp}(b)$*, there exists* $d \in M$ *such that* $\mathrm{tp}(a,c) = \mathrm{tp}(b,d)$*.*

*Proof.* The equivalence $(i) \Longleftrightarrow (ii)$ is Corollary 3.4.4.

The equivalence $(ii) \Longleftrightarrow (iii)$ is Theorem 3.4.2.

The equivalence $(ii) \Longleftrightarrow (iv)$ is Theorem 5.5.3.

The implication $(iv) \Rightarrow (v)$ is trivial.

$(v) \Rightarrow (vi)$: Let $p(x,y)$ be $\mathrm{tp}(a,c)$; it is in $S_2(ARV)$. Then $\mathrm{tp}(a) = \mathrm{tp}(b)$ implies that $p(b,y)$ is consistent. Since $\mathcal{M}$ is $\aleph_0$-saturated, $p(b,y)$ is realized in $M$ by some element, say $d$. Then $\mathrm{tp}(b,d) = \mathrm{tp}(a,c) = p(x,y)$.

$(vi) \Rightarrow (iii)$: Suppose not, then by Theorem 2.2.28 (Maharam's Theorem), $\Omega$ has a decomposition

$$(\Omega, \mathscr{B}, \mu) \cong r \cdot ([0,1], \mathcal{L}, \lambda) \oplus (1-r) \cdot (\Omega', \mathscr{B}', \mu') \text{ for some } r \in (0,1],$$

where $([0,1], \mathcal{L}, \lambda)$ is the standard Lebesgue space, $(\Omega', \mathscr{B}', \mu')$ is a probability space, and every cardinal in the Maharam spectrum of the probability space $\Omega'$ is uncountable. Take a countable open basis for $[0,1]$, denoted by $\{C_n \mid n \in \mathbb{N}\}$, and let $\alpha$ be a measure-preserving isomorphism between probability spaces $\prod_{n\in\mathbb{N}}(\{0,1\}, P(\{0,1\}), \mu_0)$ and $([0,1], \mathscr{B}, \lambda)$. Define $f\colon \Omega = [0,1] \sqcup \Omega' \to [0,1]$ by $f(\omega') = 0$ for all $\omega' \in \Omega'$ and $f(t) = \alpha((\chi_{C_1}(t), \chi_{C_2}(t), \cdots, \chi_{C_n}(t), \cdots))$ for all $t \in [0,1]$. Then we know that $f \in L^1(\mu, [0,1])$. Now define $(\overline{\Omega}, \bar{\mu})$ as $(\Omega, \mu) \times (\{0,1\}, \mu_0)$ where $\mu_0(\{0\}) = \mu_0(\{1\}) = \frac{1}{2}$. Then define functions $f', g'$ from $\overline{\Omega} = \Omega \times \{0,1\}$ to $[0,1]$ by $f'(\omega, i) = f(\omega)$ and $g'(\omega, i) = i$ for all $(\omega, i) \in \Omega \times \{0,1\}$. They are in $L^1(\bar{\mu}, [0,1])$. Note that $\mathrm{dist}(f') = \mathrm{dist}(f)$, which implies $\mathrm{tp}(f') = \mathrm{tp}(f)$ by Theorem 5.2.3. Let $p(x,y)$ stand for $\mathrm{tp}(f', g')$. By Theorem 5.2.3, the theory $\mathsf{ARV}$ is separably categorical. Therefore $p(x,y)$ is realized in $M$, say by $(f'', g'')$. Then $\mathrm{tp}(f'') = \mathrm{tp}(f)$. By the assumption of $(vi)$, we know that $p(f,y)$ is realized in $M$ by an element $g$, whence $\mathrm{tp}(f,g) = \mathrm{tp}(f', g')$. Because $g'^2 = g'$, $I(g') = \frac{1}{2}$, and $\mathrm{tp}(g) = \mathrm{tp}(g')$, we have that $g = \chi_D$ for some

$D \in \mathscr{B}$ and $\mu(D) = \frac{1}{2}$. Say $D = D_0 \sqcup D'$ with $D_0 \subseteq [0,1]$ and $D' \subseteq \Omega'$. For all $n \in \mathbb{N}$, let $D_n$ be the set $\{x \in \{0,1\}^{\mathbb{N}} \mid x(n) = 1\}$. Then $\alpha(D_n)$ is a measurable subset of $[0,1]$. Thus we have

$$\bar{\mu}\big(f' \in \alpha(D_n); g' = 1\big) = \bar{\mu}\big((t,i) \in \bar{\Omega} \mid t \in C_n, i = 1\big) = \frac{1}{2}\mu(C_n) = \mu(D)\mu(C_n)$$
$$= \mu\big(f \in \alpha(D_n); g = 1\big) = \mu(D \cap C_n).$$

Therefore $D$ is independent from $C_n$ for all $n \in \mathbb{N}$. Similarly, we get that for all finite $J \subset \mathbb{N}$, the set $D$ is independent from $\bigcap_{j \in J} C_j$. Then by [12, page 26], the set $D$ is independent from $\sigma(\{C_n\}_{n \in \mathbb{N}})$. Therefore $D$ is independent from $D_0$. Hence, $\mu(D \cap D_0) = \mu(D_0) = \mu(D)\mu(D_0)$. Since $\mu(D) \neq 0$, we have $\mu(D_0) = 0$. Thus we have $f \cdot g = 0$, but $f' \cdot g' \neq 0$, which is a contradiction to $\mathrm{tp}(f,g) = \mathrm{tp}(f',g')$. □

**5.5.5 Theorem.** *For every uncountable cardinal $\kappa$, we have that $L^1\big(([0,1]^\kappa, A_\kappa, \mu_\kappa), [0,1]\big) \models \mathsf{ARV}$ is $\kappa$-saturated and strongly $\kappa$-homogeneous. Moreover, it is the unique model of $\mathsf{ARV}$ of density character $\kappa$ with these properties.*

*Proof.* The first part follows from Theorem 3.4.5 and Theorem 5.5.3. The uniqueness follows from the following argument: for a complete theory $T$ and $\kappa \geq \mathrm{Card}(T)$, by the standard back-and-forth argument any two $\kappa$-saturated models of $T$ of density character $\kappa$ are isomorphic to each other. □

## 5.6 Stability and independence

In this section, we prove the theory $\mathsf{ARV}$ is $\omega$-stable and then characterize its model theoretic independence relation in familiar probabilistic terms. These results were proved by Ben Yaacov [2], using the analogous properties of $\mathsf{APr}$ via biinterpretability.

**5.6.1 Theorem** ([2, Theorem 2.17 (vi)])**.** *The theory $\mathsf{ARV}$ is $\omega$-stable, independence coinciding with probabilistic independence. More specifically, $A \mathop{\smile\!\!\!\!\!\!|}^{\mathsf{ARV}}_{B} C$ if and only if $\sigma(A) \mathop{\smile\!\!\!\!\!\!|}^{\mathsf{APr}}_{\sigma(B)} \sigma(C)$. Moreover, types over arbitrary sets are stationary.*

*Proof.* First, we prove that $\mathsf{ARV}$ is $\omega$-stable. Let $C$ be a countable subset of a model of $\mathsf{ARV}$. By the Downward Löwenheim-Skolem Theorem (Proposition 1.4.4), there is a separable submodel containing $C$. By Theorem 5.2.3, $\mathsf{ARV}$ is separably categorical. So we may assume that $C$ is contained in $\mathcal{N} \models \mathsf{ARV}$ of the form $N = L^1\big(([0,1]^2, p_1^{-1}(\mathscr{B}), \lambda \times \lambda), [0,1]\big)$, where $p_1 \colon [0,1]^2 \to [0,1]$ is defined by $p_1(x,y) = x$ for all $(x,y) \in [0,1]^2$, and $([0,1], \mathscr{B}, \lambda)$ is the standard Lebesgue space. Let $\mathcal{M} \models \mathsf{ARV}$ be of the form $M = L^1\big(([0,1]^2, \mathscr{B} \times \mathscr{B}, \lambda \times \lambda), [0,1]\big)$. Then $\mathcal{N}$ is a submodel of $\mathcal{M}$. Note that $\mathscr{B} \times \mathscr{B}$ is atomless over

$p_1^{-1}(\mathscr{B})$. Consider a $p(x/C) \in S_1(C)$. Then $p(x/C)$ is realized by $f$ in some elementary extension of $\mathcal{N}$. Then $\text{dist}(f \mid \sigma(C))$ is a conditional distribution over $\sigma(C)$. By Lemma 2.2.19, there exists $g \in M$ such that $\text{dist}(g \mid \sigma(C)) = \text{dist}(f \mid \sigma(C))$. Then by Theorem 5.2.3, we have $\text{tp}(g/C) = \text{tp}(f/C) = p(x/C)$. Since $p(x/C)$ is chosen arbitrarily, we know that all types in $S_1(C)$ are realized in $\mathcal{M}$. Because $(M, d)$ is separable, $S_1(C)$ is separable with respect to the $d$-metric. Hence $\mathsf{ARV}$ is $\omega$-stable.

Second, we show the characterization of the stable independence relation for $\mathsf{ARV}$. The stable independence relation $\mathop{\smile\!\!\!\!\!\!|}^{\mathsf{APr}}$ for $\mathsf{APr}$ is defined in Theorem 3.1.3. For simplicity, we will use $\mathop{\smile\!\!\!\!\!\!|}^{*}$ to denote $\mathop{\smile\!\!\!\!\!\!|}^{\mathsf{APr}}$. Let $\mathbb{M}$ be a $\kappa$-universal domain for $\mathsf{ARV}$ with $\kappa$ uncountable. Assume $\mathbb{M}$ is of the form $L^1\big(\mu, [0,1]\big)$ for the atomless probability space $(\Omega, \mathbb{F}, \mu)$. Define a relation $\mathop{\smile\!\!\!\!\!\!|}$ by $A \mathop{\smile\!\!\!\!\!\!|}_B C$ if $\sigma(A) \mathop{\smile\!\!\!\!\!\!|}^{\mathsf{APr}}_{\sigma(B)} \sigma(C)$, for all small $A, B, C \subseteq \mathbb{M}$. We will show that $\mathop{\smile\!\!\!\!\!\!|}$ is a stable independence relation for $\mathsf{ARV}$. Then by Theorem 1.11.6, $\mathop{\smile\!\!\!\!\!\!|} = \mathop{\smile\!\!\!\!\!\!|}^{\mathsf{ARV}}$.

In the following, the subsets $A, B, C, \cdots \subseteq \mathbb{M}$ are all small.

(i) Invariance under automorphisms of $\mathbb{M}$.

By Theorem 5.1.1, every $L_{\mathsf{RV}}$-automorphism of $\mathbb{M}$ induces an $L_{\mathsf{Pr}}$-automorphism of the probability algebra $\widehat{\mathbb{F}}$ and $\mathop{\smile\!\!\!\!\!\!|}^{*}$ is invariant under automorphisms of $\widehat{\mathbb{F}}$.

(ii) Symmetry: $A \mathop{\smile\!\!\!\!\!\!|}_C B \iff B \mathop{\smile\!\!\!\!\!\!|}_C A$.

This follows directly from the symmetry of $\mathop{\smile\!\!\!\!\!\!|}^{*}$.

(iii) Transitivity: $A \mathop{\smile\!\!\!\!\!\!|}_C (B \cup D) \iff A \mathop{\smile\!\!\!\!\!\!|}_C B$ and $A \mathop{\smile\!\!\!\!\!\!|}_{B \cup C} D$.

Note that $\sigma(A_1 \cup A_2) = \sigma(\sigma(A_1) \cup \sigma(A_2))$ for all subsets $A_1, A_2 \subseteq \mathbb{M}$ and

$$B_1 \mathop{\smile\!\!\!\!\!\!|}^{*}_{B_2} B_3 \iff \sigma(B_1) \mathop{\smile\!\!\!\!\!\!|}^{*}_{\sigma(B_2)} \sigma(B_3),$$

for all $\sigma$-subalgebras $B_1, B_2, B_3$. Then

$$\begin{aligned}
A \mathop{\smile\!\!\!\!\!\!|}_C (B \cup D) &\iff \sigma(A) \mathop{\smile\!\!\!\!\!\!|}^{*}_{\sigma(C)} \sigma(B \cup D) \iff \sigma(A) \mathop{\smile\!\!\!\!\!\!|}^{*}_{\sigma(C)} \sigma(B) \cup \sigma(D) \\
&\iff \sigma(A) \mathop{\smile\!\!\!\!\!\!|}^{*}_{\sigma(C)} \sigma(B) \text{ and } \sigma(A) \mathop{\smile\!\!\!\!\!\!|}^{*}_{\sigma(B) \cup \sigma(C)} \sigma(D) \\
&\iff \sigma(A) \mathop{\smile\!\!\!\!\!\!|}^{*}_{\sigma(C)} \sigma(B) \text{ and } \sigma(A) \mathop{\smile\!\!\!\!\!\!|}^{*}_{\sigma(B \cup C)} \sigma(D) \\
&\iff A \mathop{\smile\!\!\!\!\!\!|}_C B \text{ and } A \mathop{\smile\!\!\!\!\!\!|}_{B \cup C} D.
\end{aligned}$$

(iv) Finite character: $A \mathop{\smile\!\!\!\!\!\!|}_C B$ iff $a \mathop{\smile\!\!\!\!\!\!|}_C B$ for all finite tuples $a$ from $A$.

This follows directly from the finite character property of $\mathop{\smile\!\!\!\!\!\!|}^{*}$.

(v) Extension: For all $A, B, C$, there is $A'$ such that $A' \mathop{\downarrow}_C B$ and $\mathrm{tp}(A/C) = \mathrm{tp}(A'/C)$.

Assume $A \supseteq C$. There is a $\sigma$-subalgebra $D \subseteq \mathbb{F}$ such that $D \mathop{\downarrow}^*_{\sigma(C)} \sigma(B)$ and $\mathrm{tp}^{\mathsf{APr}}(D/\sigma(C)) = \mathrm{tp}^{\mathsf{APr}}(\sigma(A)/\sigma(C))$. Then there is an $L_{\mathsf{Pr}}$-elementary mapping $i\colon \sigma(A) \to D$ over $\sigma(C)$. Further, this mapping $i$ induces a subset $A' \subseteq \mathbb{M}$ such that $\sigma(A') = D$ and $\mathrm{tp}^{\mathsf{ARV}}(A'/C) = \mathrm{tp}^{\mathsf{ARV}}(A/C)$. By $\sigma(A') \mathop{\downarrow}^*_{\sigma(C)} \sigma(B)$, we have $A' \mathop{\downarrow}_C B$.

(vi) Local character: For finite tuple $a$, there is $B' \subseteq B$ of cardinality $\leq |T|$ such that $a \mathop{\downarrow}_{B'} B$.

Note that the space $\mathbb{E}\big(\mathrm{dcl}(a) \mid \sigma(B)\big)$ is separable. Let $B'$ be a countable dense subset of the space $\mathbb{E}\big(\mathrm{dcl}(a) \mid \sigma(B)\big)$. Then $|B'| = \aleph_0 = |\mathsf{ARV}|$. For all $C \in \sigma(a)$, one has

$$\mathbb{P}(C \mid \sigma(B) \vee \sigma(B')) = \mathbb{P}(C \mid \sigma(B')),$$

whereby $C \mathop{\downarrow}^*_{\sigma(B')} \sigma(B)$ by Theorem 3.1.3. Hence $\sigma(a) \mathop{\downarrow}^*_{\sigma(B')} \sigma(B)$, whereby $a \mathop{\downarrow}_{B'} B$.

(vii) Stationarity of types over arbitrary sets: If $\mathrm{tp}(A/C) = \mathrm{tp}(A'/C)$, $A \mathop{\downarrow}_C B$, and $A' \mathop{\downarrow}_C B$, then we have $\mathrm{tp}(A/B \cup C) = \mathrm{tp}(A'/B \cup C)$.

By the finite character property, it suffices to show the stationarity of types for finite $A$ and $A'$. Assume that $A = a$ and $A' = a' \in M^n$. Let $D$ be a Borel subset of $[0,1]^n$. Then $\mathrm{tp}(a/C) = \mathrm{tp}(a'/C)$ implies $\mathbb{P}\big(a \in D \mid \sigma(C)\big) = \mathbb{P}\big(a' \in D \mid \sigma(C)\big)$ by Theorem 5.2.3. Since $a \mathop{\downarrow}_C B$ iff $\sigma(a) \mathop{\downarrow}^*_{\sigma(C)} \sigma(B)$, one has $\mathbb{P}\big(a \in D \mid \sigma(C) \vee \sigma(B)\big) = \mathbb{P}\big(a \in D \mid \sigma(C)\big)$, by Theorem 3.1.3. Similarly,

$$\mathbb{P}\big(a' \in D \mid \sigma(C) \vee \sigma(B)\big) = \mathbb{P}(a' \in D \mid \sigma(C)\big).$$

Hence $\mathbb{P}\big(a \in D \mid \sigma(C) \vee \sigma(B)\big) = \mathbb{P}\big(a' \in D \mid \sigma(C) \vee \sigma(B)\big)$, whereby we have $\mathbb{P}\big(a \in D \mid \sigma(B \cup C)\big) = \mathbb{P}\big(a' \in D \mid \sigma(B \cup C)\big)$; *i.e.*, $a$ and $a'$ have the same joint conditional distribution over $\sigma(B \cup C)$. Hence $\mathrm{tp}(a/B \cup C) = \mathrm{tp}(a'/B \cup C)$, by Theorem 5.2.3.

□

**5.6.2 Remark.** *By Remark 3.1.4, one has*

$$A \mathop{\downarrow}^{\mathsf{ARV}}_C B \quad \Longleftrightarrow \quad \mathbb{E}\big(f \mid \sigma(B \cup C)\big) = \mathbb{E}\big(f \mid \sigma(C)\big) \text{ a.s. for all } f \in L^1(\sigma(A), [0,1]).$$

## 5.7 On $d$-finite tuples in ARV

In this section, we study the $d$-finite tuples in models of $\mathsf{ARV}$ and then show some applications.

At the beginning, we recall the definition of $d$-finite tuples.

**5.7.1 Definition** ([5]). *Let $a$ and $c$ be tuples and let $p = \mathrm{tp}(a/c)$. Then $a$ is called $d$-**finite over** $c$, equivalently, the type $p$ is called $d$-**finite**, if for every tuple $b$ and every $\epsilon > 0$, there is $\delta = \delta^{a,c}_{b,\epsilon} > 0$ such that whenever $a' \equiv_c a$ and $d(a, a') \leq \delta$, there is $b'$ such that $d(b, b') \leq \epsilon$ and $a'b' \equiv_c ab$. If $c = \emptyset$, we omit $c$ from the notation.*

## Characterization of the $d$-finite tuples in ARV

Now, we characterize the $d$-finite tuples in models of ARV. We know that the theory ARV is an $\aleph_0$-categorical theory, by Theorem 5.2.3. By Theorem 1.9.3 (Ryll-Nardzewski Theorem), its unique separable model is approximately $\aleph_0$-saturated, but it is not $\aleph_0$-saturated by Theorem 5.5.1. For any approximately $\aleph_0$-saturated model, Ben Yaacov and Usvyatsov [5] proved the following:

**5.7.2 Proposition** ([5]). *If $\mathcal{M}$ is an approximately $\aleph_0$-saturated model of $T$ and $a \in M$ is $d$-finite, then every type over $a$ in at most countably many variables is realized in $M$.*

*Proof.* This is [5, Corollary 2.5]. □

Let $\mathcal{M} \models \mathsf{ARV}$. For tuple $f = (f_1, \cdots, f_n) \in M^n$, let $\mathsf{ARV(f)}$ denote $Th(\mathcal{M}, f)$. By Theorem 3.5.2, one has that $f$ is $d$-finite (over $\emptyset$) if and only if $\mathsf{ARV(f)}$ is $\aleph_0$-categorical. We use this result to characterize the $d$-finite tuples in $\mathcal{M}$.

**5.7.3 Theorem.** *Let $\mathcal{M} \models \mathsf{ARV}$ and $f = (f_1, \cdots, f_n) \in M^n$. Then $f$ is $d$-finite if and only if $f_i$ is a discrete random variable for every $1 \leq i \leq n$.*

*Proof.* $\Longleftarrow$: Let $\mathcal{N} = L^1(\lambda, [0,1]) \models \mathsf{ARV}$, where $([0,1], \mathcal{L}, \lambda)$ is the standard Lebesgue space. To show $\mathsf{ARV(f)}$ is $\aleph_0$-categorical, we need only show that for all $g, h \in N^n$ such that $\mathrm{tp}(f) = \mathrm{tp}(g) = \mathrm{tp}(h)$, there exists $\varphi \in \mathrm{Aut}(\mathcal{N})$ such that $\varphi(g) = h$.

Let $g = (g_1, \cdots, g_n) \in N^n$ such that $\mathrm{tp}(f) = \mathrm{tp}(g)$. Since all $f_i$'s are discrete random variables, all $g_i$'s are also discrete. Therefore, we can write $g_i$'s as $g_i = \sum_{j=1}^{\infty} r_{ij}\chi_{A_{ij}}$ for all $i = 1, \cdots, n$, where all $r_{ij} \in [0,1]$ and $(A_{i1}, \cdots, A_{in}, \cdots)$ is a measurable partition of $[0,1]$. Then $\{A_{1i_1} \cap A_{2i_2} \cap \cdots \cap A_{ni_n} \mid \forall i_1, i_2, \cdots, i_n \in \mathbb{N}\}$ also forms a partition of $[0,1]$, although some might be null sets. Take any other $h = (h_1, \cdots, h_n) \in N^n$ such that $\mathrm{tp}(f) = \mathrm{tp}(g) = \mathrm{tp}(h)$. Then for every $i = 1, \cdots, n$, every $h_i$ can be written as $\sum_{j=1}^{\infty} r_{ij}\chi_{B_{ij}}$ such that $(B_{i1}, \cdots, B_{in}, \cdots)$ is a measurable partition of $[0,1]$ with $\mu(A_{ij}) = \mu(B_{ij})$ for all $j \in \mathbb{N}$ and $\mu(A_{1i_1} \cap A_{2i_2} \cap \cdots \cap A_{ni_n}) = \mu(B_{1i_1} \cap B_{2i_2} \cap \cdots \cap B_{ni_n})$ for all $i_1, i_2, \cdots, i_n \in \mathbb{N}$. Then by Lemma 3.1.5, there is an automorphism $\varphi$ of $([0,1], \mathcal{L}, \lambda)$ such that $\varphi(A_{1i_1} \cap A_{2i_2} \cap \cdots \cap A_{ni_n}) = B_{1i_1} \cap B_{2i_2} \cap \cdots \cap B_{ni_n}$

(up to null sets) for all $i_1, i_2, \cdots, i_n \in \mathbb{N}$. We then extend $\varphi$ to an automorphism $\widehat{\varphi}$ of $\mathcal{N}$ such that $\widehat{\varphi}(g) = h$. Hence $f$ is $d$-finite.

$\Longrightarrow$: Assume that $f$ is a $d$-finite element in $\mathcal{N} = L^1(\lambda, [0,1]) \models \mathsf{ARV}$, where $([0,1], \mathcal{L}, \lambda)$ is the standard Lebesgue space. Then for all $g \in N$ such that $\mathrm{tp}(f) = \mathrm{tp}(g)$, there is a $\varphi \in \mathrm{Aut}(\mathcal{N})$ satisfying $\varphi(f) = g$. Suppose $f$ is not a discrete random variable, then $f$ has a decomposition into two parts, one is a discrete random variable $f_d$, and the other is a nonzero continuous random variable $f_c$. We write $f = f_d + f_c$; note that $f_d \cdot f_c = 0$. By Theorem 5.2.3, $f_c \in \mathrm{dcl}(f)$ and $\mathsf{ARV}(\mathsf{f})$ is separably categorical. Therefore, $\mathsf{ARV}(\mathsf{f}, \mathsf{f_c})$ is also separably categorical, whereby $\mathsf{ARV}(\mathsf{f_c})$ is separably categorical by Proposition 1.9.5. Hence $f_c$ is $d$-finite, by Theorem 3.5.2.

Define $F(t) = \lambda(f_c \leq t)$ for all $t \in [0,1]$. Because $f_c$ is nonzero, $F(t)$ is continuous and strictly increasing. Thus, the inverse function $F^{-1}(t)$ exists; we write $g(t) = F^{-1}(t)$ for all $t \in [0,1]$. Because $F(t)\colon [0,1] \to [0,1]$ is continuous and strictly increasing, we have that $\sigma(g)$ is the set $\mathscr{B}$ of all Borel sets in $[0,1]$. Also, note that $\mathrm{dist}(g) = \mathrm{dist}(f_c)$, whereby $\mathrm{tp}(g) = \mathrm{tp}(f_c)$ by Theorem 5.2.3. We know that $L^1(([0,1], \mathcal{L}, \lambda), [0,1])$ is isomorphic to $L^1\big(([0,1]^2, \mathcal{L} \times \mathcal{L}, \lambda \times \lambda), [0,1]\big)$, since $\mathsf{ARV}$ is separably categorical by Theorem 5.2.3. Let $\alpha$ be such an isomorphism. Let $g_1 = \alpha(g)$. Because $\sigma(g)$ is the set of all Borel sets in $[0,1]$, we know that $L^1(([0,1], \sigma(g), \lambda), [0,1]) = L^1(([0,1], \mathcal{L}, \lambda), [0,1])$. By Theorem 5.2.3, we have that $\mathrm{dcl}(g) = L^1(([0,1], \mathcal{L}, \lambda), [0,1])$. Since $\alpha$ is an isomorphism and $\alpha(g) = g_1$, we know that

$$\mathrm{dcl}(g_1) = L^1\big(([0,1]^2, \sigma(g_1), \lambda \times \lambda), [0,1]\big) = L^1\big(([0,1]^2, \mathcal{L} \times \mathcal{L}, \lambda \times \lambda), [0,1]\big).$$

We define $g_2(t,s) := g(t)$ for all $t, s \in [0,1]$. Then we know that $\mathrm{dist}(g_2) = \mathrm{dist}(g)$, which implies $\mathrm{tp}(g_2) = \mathrm{tp}(g)$ by Theorem 5.2.3. But clearly, the completion of $\sigma(g_2)$ is not $\mathcal{L} \times \mathcal{L}$. Thus, $\mathrm{dcl}(g_2) \subsetneq \mathrm{dcl}(g_1)$. If there is an automorphism $\Phi$ of $L^1\big(([0,1]^2, \mathcal{L} \times \mathcal{L}, \lambda \times \lambda), [0,1]\big)$ sending $g_1$ to $g_2$. Then $\Phi(\mathrm{dcl}(g_1)) = \mathrm{dcl}(g_2)$. Since $\mathrm{dcl}(g_1) = L^1\big(([0,1]^2, \mathcal{L} \times \mathcal{L}, \lambda \times \lambda), [0,1]\big)$, we have that $\Phi(\mathrm{dcl}(g_1)) = \mathrm{dcl}(g_1)$, and thus $\mathrm{dcl}(g_2) = \mathrm{dcl}(g_1)$, which contradicts the fact that $\mathrm{dcl}(g_2) \subsetneq \mathrm{dcl}(g_1)$.

Therefore $\mathsf{ARV}(\mathsf{f})$ is not $\aleph_0$-categorical, which contradicts the fact that $f_c$ is $d$-finite. Hence, $f$ is a discrete random variable.

If $f = (f_1, \cdots, f_n)$ is a $d$-finite tuple, then by definition of $d$-finiteness, $f_i$ is $d$-finite for every $i = 1, \cdots, n$. Thus we know that $f_i$ is a discrete random variable for every $i = 1, \cdots, n$. □

Let $\mathcal{M} \models \mathsf{ARV}$. If $f$ is not a $d$-finite tuple in $M$, then $\mathsf{ARV}(\mathsf{f})$ is not $\aleph_0$-categorical. But how many separable models would it have?

**5.7.4 Proposition.** *Let $\mathcal{M} \models \mathsf{ARV}$ and let $f$ be a tuple in $M$. Then*

*(i) If $f$ is d-finite, then* $\mathsf{ARV}(\mathsf{f})$ *has a unique separable model up to isomorphism.*

*(ii) If $f$ is not d-finite, then* $\mathsf{ARV}(\mathsf{f})$ *has continuum many non-isomorphic separable models.*

*Proof.* $(i)$: This is Theorem 3.5.2.

$(ii)$: Without loss of generality, we may assume that $\mathcal{M} = L^1\big(\lambda, [0,1]\big)$, where $([0,1], \mathcal{L}, \lambda))$ is the standard Lebesgue space, and $f$ is a finite tuple in $M$, which is not $d$-finite. By Theorem 5.7.3, $f = (f_1, \cdots, f_n)$ is not a discrete random variable on $[0,1]^n$. Since $\sigma(f)$ is countably $\sigma$-generated, there exists $g \in M$ such that $\sigma(g) = \sigma(f)$. Then $g$ is not discrete and hence $g$ has a decomposition into two parts: $g = g_c + g_d$, where $g_c$ is a nonzero continuous random variable, $g_d$ is a discrete random variable, and $g_c \cdot g_d = 0$. For all $r, s \in [0,1]$, let $\mathscr{B}([r,s])$ denote the set of all the Borel sets in $[r,s]$. Then $\mathscr{B} = \mathscr{B}([0,1])$. In $[0,1] \times [0,1]$, define $\mathscr{B}_r$ as $\sigma(\{A_1 \times [0,1], A_2 \times A_3 : A_1 \in \mathscr{B}([0,r]), A_2 \in \mathscr{B}([r,1]), A_3 \in \mathscr{B}\})$. Then $\mathscr{B}_0$ is just the set of all the Borel sets in $[0,1] \times [0,1]$. Without loss of generality, we may assume that $\sigma(g_d) = \mathscr{B}([0,t])$ for some $t \in (0,1)$ and let $\mathscr{B}_g$ denote $\sigma(g) = \sigma(g_c) \vee \sigma(g_d)$.

For $([0,1] \times [0,1], \mathscr{B}_r)$, there is a natural probability measure $\lambda_r$ satisfying:

- $\lambda_r(A_1 \times [0,1]) = \lambda(A_1)$ for all $A_1 \in \mathscr{B}([0,r])$.
- $\lambda_r(A_2 \times A_3) = \lambda(A_2)\lambda(A_3)$ for all $A_2 \in \mathscr{B}([r,1])$ and all $A_3 \in \mathscr{B}$.

Let $\mathcal{L}_r$ denote the completion of $\mathscr{B}_r$ under $\lambda_r$. Define $M_r := L^1\big(([0,1]\times[0,1], \mathcal{L}_r, \lambda_r), [0,1]\big)$. Then $M_r \models \mathsf{ARV}$ and it is separable. Define $g_r \colon [0,1] \times [0,1] \to [0,1]$ by $g_r(s,t) = g(s)$ for $s, t \in [0,1]$. Note that $\mathrm{dist}(g_r) = \mathrm{dist}(g) = \mathrm{dist}(f)$, which implies $\mathrm{tp}(g_r) = \mathrm{tp}(g) = \mathrm{tp}(f)$ by Theorem 5.2.3. Therefore, $(M_r, g_r) \models \mathsf{ARV}(\mathsf{f})$ for all $r \in [0,1]$. Since $\sigma(g) = \mathscr{B}_g$, we have $\sigma(g_r) = \mathscr{B}_g \times [0,1] := \{B \times [0,1] : B \in \mathscr{B}_g\}$.

Now we show for distinct $r \leq t$, those models are not isomorphic to each other. For all $r_1, r_2 \in [0,t]$ with $r_1 > r_2$, suppose $(M_{r_1}, g_{r_1}) \cong (M_{r_2}, g_{r_2})$; let $\alpha$ be such isomorphism, then $\alpha(g_{r_1}) = g_{r_2}$. Note that $\alpha$ induces an isomorphism $\beta$ between $(\mathscr{B}_{r_1}, \sigma(g_{r_1}), \lambda_{r_1})$ and $(\mathscr{B}_{r_2}, \sigma(g_{r_2}), \lambda_{r_2})$ such that $\beta(\mathscr{B}_{r_1}) = \mathscr{B}_{r_2}$, $\beta(\sigma(g_{r_1})) = \sigma(g_{r_2})$, and $\beta$ preserves the measure. Since $\sigma(g_{r_1}) = \sigma(g_{r_2}) = \mathscr{B}_g \times [0,1]$, we have $\beta(\mathscr{B}_g \times [0,1]) = \mathscr{B}_g \times [0,1]$. For any two measure algebras $\mathscr{A}_1 \subseteq \mathscr{A}_2$, an element $A \in \mathscr{A}_2$ is called an atom over $\mathscr{A}_1$ if for every $B \in \mathscr{A}_2$ there exists $C \in \mathscr{A}_1$ such that $A \cap B = A \cap C$. Note that $[0,r] \times [0,1] \in \mathscr{B}_r$ is an atom over $\mathscr{B}_g \times [0,1]$ for all $r \in [0,t]$. Also $\lambda_r([0,r] \times [0,1]) = r$ and it has the largest measure among atoms over $\mathscr{B}_g \times [0,1]$ in $\mathscr{B}_r$. Because $\beta$ is an isomorphism, $\beta([0,r_1] \times [0,1])$ is also an atom over $\beta(\mathscr{B}_g \times [0,1]) = \mathscr{B}_g \times [0,1]$. But $\lambda_{r_2}(\beta([0,r_1] \times [0,1])) = \lambda_{r_1}([0,r_1] \times [0,1]) = r_1 > r_2$. This contradicts the fact that each atom in $\mathscr{B}_{r_2}$ over $\mathscr{B}_g \times [0,1]$ has measure at most $r_2$.

Therefore, $\mathsf{ARV}(\mathsf{f})$ has continuum many separable models. □

## Applications

Now, we connect $d$-finite tuples with work by Keisler, Hoover, Fajardo, and Sun. Starting in the 80s, Keisler, Hoover, Fajardo, and Sun published several papers concerning model theoretic results about probability spaces. The $d$-finiteness in ARV would give some of their results a more precise view.

**5.7.5 Theorem** ([13])**.** *No ordinary probability space is saturated.*

*Proof.* This is [13, Theorem 3B.1]. □

Since no ordinary probability space is saturated, they proved the following weak saturation property for atomless probability spaces.

**5.7.6 Proposition** ([13])**.** *Let $\Omega$ be an atomless probability space and let $\Gamma$ be another probability space. Let $M$ and $N$ be Polish spaces. Then for every simple $M$-valued random variable $x$ on $\Omega$ and every pair of random variables $\bar{x}, \bar{y}$ on $\Gamma$ with values in $M$ and $N$ such that $\mathrm{dist}(\bar{x}) = \mathrm{dist}(x)$, there exists a random variable $y$ on $\Omega$ such that $\mathrm{dist}(x, y) = \mathrm{dist}(\bar{x}, \bar{y})$.*

*Proof.* This is [13, Proposition 3B.3]. □

Now, in the continuous model theory setting, those two results correspond to the following:

**5.7.7 Theorem.** *No separable model of* ARV *is $\aleph_0$-saturated.*

*Proof.* This is Theorem 5.5.1. □

**5.7.8 Proposition.** *Let $\mathcal{M} \models$ ARV. Then for any $d$-finite tuple $f$ in $M$ and any type $p(f, y)$ in $S_1(f)$, there is $g \in M$ realizing $p(f, y)$.*

*Proof.* This is Proposition 5.7.2. □

**5.7.9 Remark.** *By Proposition 5.7.8 and Theorem 5.7.3, Proposition 5.7.6 is strengthened so the conclusion is true for discrete random variables (in place of simple ones).*

In [22], Keisler and Sun proved the following theorem, which gives a local sufficient condition for the saturation property to hold.

**5.7.10 Theorem** ([22])**.** *Let $f, g$ be random variables with values in Polish spaces $X, Y$ respectively and assume that $\mathrm{dist}(f)$ is atomless. Suppose $(\Omega, \mathcal{F}, \mu)$ has the saturation property for $\mathrm{dist}(f, g)$ but the standard Lebesgue unit interval $([0, 1], \mathcal{L}, \lambda)$ does not. Then $(\Omega, \mathcal{F}, \mu)$ is Hoover-Keisler saturated.*

*Proof.* This is [22, Theorem 2.7]. □

In the continuous model theory setting, the preceding theorem corresponds the following:

**5.7.11 Theorem.** *Let $M$ be $L^1(\lambda, [0,1])$, where $([0,1], \mathcal{L}, \lambda)$ is the standard Lebesgue space. For every non-d-finite tuple $f$ in $M$, there is a type $p(x,y)$ in $\mathsf{ARV}$ and an $f'$ in $M$ with $\mathrm{tp}(f') = \mathrm{tp}(f)$, such that $p(f',y)$ is not realized in $M$. For every model $\mathcal{N}$ of $\mathsf{ARV}$, if for every $f'$ in $N$ with $\mathrm{tp}(f') = \mathrm{tp}(f)$, the type $p(f',y)$ is realized in $N$, then $\mathcal{N}$ is $\aleph_0$-saturated.*

**5.7.12 Remark.** *By Theorem 5.7.3, Theorem 5.7.10 is strengthened so the conclusion is true for non-atomic distributions* $\mathrm{dist}(f)$ *(in place of atomless ones).*

## 5.8 Conditional Wasserstein distances

By Theorem 5.3.3, the mapping

$$\begin{aligned} \zeta \colon S_n^{\mathsf{ARV}}(A) &\to \mathfrak{D}_{[0,1]^n}(\mathscr{A}) \\ \mathrm{tp}(f/A) &\mapsto \mathrm{dist}(f \mid \mathscr{A}) \end{aligned}$$

is a homeomorphism between the type space $S_n^{\mathsf{ARV}}(A)$ equipped with the logic topology and the space $\mathfrak{D}_{[0,1]^n}(\mathscr{A})$ of conditional distributions over $\mathscr{A}$ equipped with the topology of weak convergence. If we consider the $d^*$-metric on the type spaces over $\emptyset$ in $\mathsf{ARV}$, then by Theorem 5.3.6, the mapping

$$\begin{aligned} \eta_n \colon S_n(\mathsf{ARV}) &\to \mathfrak{D}_{[0,1]^n} \\ \mathrm{tp}(f) &\mapsto \mathrm{dist}(f) \end{aligned}$$

is an isometric isomorphism between the type space $\big(S_n(\mathsf{ARV}), d^*\big)$ and the Wasserstein space $(\mathfrak{D}_{[0,1]^n}, W_1)$. In model theory, types over parameters are as important as types over $\emptyset$, so it is natural and potentially valuable to define a new distance on $\mathfrak{D}_{[0,1]^n}(\mathscr{A})$ which generalizes the Wasserstein distance between distributions, so that $\zeta$ becomes an isometric isomorphism. In this section, we will discuss this new metric, called conditional Wasserstein distance, and define conditional Wasserstein spaces.

Let $(\Omega, \mathcal{F}, m)$ be a $\kappa$-saturated probability space, where $\kappa$ is an uncountable cardinal. Let $\mathscr{A}$ be a $\sigma$-subalgebra of $\mathcal{F}$ that is small with respect to $\kappa$. For a Polish metric space $\mathcal{X}$, an *$n$-dimensional conditional distribution over $\mathscr{A}$ on $\mathcal{X}$*, say $\mu$, is an $L^1\big((\Omega, \mathscr{A}, m), [0,1]\big)$-valued Borel probability measure on $\mathcal{X}$. More precisely, it satisfies the following:

- $\mu(B) \geq 0$ a.s. for all Borel sets $B \subseteq \mathcal{X}$.

- $\mu(\mathcal{X}) = 1$ a.s.
- $\mu(\bigcup_{i=1}^{\infty} B_i) = \sum_{i=1}^{\infty} \mu(B_i)$ a.s. for all disjoint Borel sets $B_1, B_2, \cdots \subseteq \mathcal{X}$.

Every random variable $f$ from $\Omega$ to $\mathcal{X}$ determines a conditional distribution over $\mathscr{A}$, denoted by $\mathrm{dist}(f \mid \mathscr{A})$. It is defined by $\mathrm{dist}(f \mid \mathscr{A})(B) := \mathbb{P}(f \in B \mid \mathscr{A})$ for every Borel set $B \subseteq \mathcal{X}$, where $\mathbb{P}(\cdot \mid \mathscr{A})$ is the conditional probability. Let $\mathfrak{D}_{\mathcal{X}}(\mathscr{A})$ denote the set of all conditional distributions over $\mathscr{A}$ on $\mathcal{X}$. For every Borel set $B \subseteq \mathcal{X}$, we denote by $\mathfrak{D}_B(\mathscr{A})$ the set of all conditional distributions over $\mathscr{A}$ which, as measures, are supported by $B$.

**5.8.1 Definition** (Conditional Wasserstein distance)**.** *Let $(\Omega, \mathcal{F}, m)$ be a $\kappa$-saturated probability space for an uncountable cardinal $\kappa$ and let $\mathscr{A}$ be a small $\sigma$-subalgebra of $\mathcal{F}$. Let $(\mathcal{X}, d)$ be a Polish metric space and let $p \in [1, \infty)$. For conditional distributions $\mu$ and $\nu$ over $\mathscr{A}$ on $\mathcal{X}$, we define the* ***conditional Wasserstein distance of order*** *$p$* ***over*** *$\mathscr{A}$ as follows:*

$$W_p(\mu, \nu) := \inf\Big\{ \big[\mathbb{E}d(f,g)^p\big]^{\frac{1}{p}} \mid f, g \colon \Omega \to \mathcal{X}, \mathrm{dist}(f \mid \mathscr{A}) = \mu, \mathrm{dist}(g \mid \mathscr{A}) = \nu \Big\}.$$

In the definition, we need $\Omega$ to be $\kappa$-saturated with $\kappa > |\mathscr{A}|$. Otherwise, for some $\mu \in \mathfrak{D}_{\mathcal{X}}(\mathscr{A})$, it is possible that no $f \colon \Omega \to \mathcal{X}$ satisfies $\mathrm{dist}(f \mid \mathscr{A}) = \mu$. For example, $(\Omega, \mathcal{F}, m)$ is the standard Lebesgue space $([0,1], \mathcal{L}, \lambda)$. In the proof of Theorem 5.5.1, $f \in M = L^1(\lambda, [0,1])$ and $p_0(x,y) \in S_2(\mathsf{ARV})$ are defined such that the type $p_0(f, y)$ is not realized in $M$. Therefore there exists a type $p(y)$ in $S_1(f)$ such that it is not realized in $M$. By Theorem 5.2.3, the type $p(y)$ corresponds $\mu \in \mathfrak{D}_{[0,1]^n}\big(\sigma(f)\big)$ such that no $g \in M$ satisfies $\mathrm{dist}(g \mid \sigma(f)) = \mu$.

Nonetheless, there is an alternative approach to the definition as follows:

**5.8.2 Definition.** *Let $(\mathcal{X}, d)$ be a Polish metric space and let $\mathscr{A}$ be a probability algebra. Let $p \in [1, \infty)$. For $\mu, \nu \in \mathfrak{D}_{\mathcal{X}}(\mathscr{A})$, the* ***conditional Wasserstein distance of order*** *$p$* ***over*** *$\mathscr{A}$ is defined as*

$$W_p(\mu, \nu) := \inf_{\Omega, f, g} \big\{ \big[\mathbb{E}d(f,g)^p\big]^{\frac{1}{p}} \big\}$$

*where the infimum is over all probability spaces $(\Omega, \mathcal{F}, \mu)$ with $\mathscr{A} \subseteq \mathcal{F}$ and all random variables $f, g \colon \Omega \to \mathcal{X}$ such that $\mathrm{dist}(f \mid \mathscr{A}) = \mu$ and $\mathrm{dist}(g \mid \mathscr{A}) = \nu$.*

**5.8.3 Remark.** *These two definitions are equivalent to each other. Let $W_p^{\Omega}$ denote the distance defined in Definition 5.8.1 and let $W_p$ denote the one defined in Definition 5.8.2. Clearly, for all $\mu, \nu \in \mathfrak{D}_{\mathcal{X}}(\mathscr{A})$, one has $W_p(\mu, \nu) \leq W_p^{\Omega}(\mu, \nu)$. For all $\epsilon > 0$, by definition there are $f, g$ such that $W_p(\mu, \nu) + \epsilon \geq \big[\mathbb{E}d(f,g)^p\big]^{\frac{1}{p}}$.*

*Since $\Omega$ in Definition 5.8.1 is $\kappa$-saturated for some large cardinal $\kappa$ and $\mathscr{A}$ is small, there are $f', g' \colon \Omega \to \mathcal{X}$ such that* $\operatorname{dist}\big((f', g') \mid \mathscr{A}\big) = \operatorname{dist}\big((f, g) \mid \mathscr{A}\big)$. *Therefore,* $\operatorname{dist}(f' \mid \mathscr{A}) = \operatorname{dist}(f \mid \mathscr{A})$, $\operatorname{dist}(g' \mid \mathscr{A}) = \operatorname{dist}(g \mid \mathscr{A})$, *and* $[\mathbb{E}d(f, g)^p]^{\frac{1}{p}} = [\mathbb{E}d(f', g')^p]^{\frac{1}{p}}$. *Consequently,* $W_p(\mu, \nu) + \epsilon \geq [\mathbb{E}d(f', g')^p]^{\frac{1}{p}} \geq W_p^{\Omega}(\mu, \nu)$. *Hence* $W_p^{\Omega}(\mu, \nu) = W_p(\mu, \nu)$.

**5.8.4 Definition** (Conditional Wasserstein space)**.** *Let $(\Omega, \mathcal{F}, m)$, $\mathscr{A}$, $(\mathcal{X}, d)$, $p$ be as above. For every $x \in \mathcal{X}$, let $c_x$ denote the constant function from $\Omega$ to $\mathcal{X}$ whose value is $x$. The* ***conditional Wasserstein space of order*** *$p$* ***over*** *$\mathscr{A}$ is defined as*

$$P_p(\mathcal{X}, \mathscr{A}) := \big\{\mu \in \mathfrak{D}_{\mathcal{X}}(\mathscr{A}) \mid W_p\big(\mu, \operatorname{dist}(c_{x_0} \mid \mathscr{A})\big) < \infty \text{ for some } x_0 \in \mathcal{X}\big\}.$$

Suppose $\mathcal{X} = [0,1]^n$ with the distance defined by $d((a_1, \cdots, a_n), (b_1, \cdots, b_n)) = \sum_{i=1}^n d(a_i, b_i)$ for all $(a_1, \cdots, a_n), (b_1, \cdots, b_n) \in [0,1]^n$. Then $P_p([0,1]^n, \mathscr{A}) = \mathfrak{D}_{[0,1]^n}(\mathscr{A})$. Now we connect conditional Wasserstein spaces to types spaces over set of parameters.

**5.8.5 Theorem.** *Let $\mathcal{M}$ be a model of* ARV *of the form $L^1(\mu, [0,1])$ for the atomless probability space $(\Omega, \mathcal{F}, \mu)$. Suppose $\mathcal{M}$ is a $\kappa$-universal domain for an uncountable cardinal $\kappa$. Let $A$ be a small subset of $M$ and let $\mathscr{A} = \sigma(A)$ be the $\sigma$-algebra of measurable sets generated by the random variables in $A$. Let $f$ be an $n$-tuple in $M$. Then* $\operatorname{dist}(f \mid \mathscr{A})$ *is a conditional distribution over $\mathscr{A}$ on $[0,1]^n$. Moreover, the mapping*

$$\eta_n^A \colon S_n(A) \to P_1\big([0,1]^n, \mathscr{A}\big)$$
$$\operatorname{tp}(f/A) \mapsto \operatorname{dist}(f \mid \mathscr{A})$$

*is an isometric isomorphism between $\big(S_n(A), d^*\big)$ and $\big(P_1([0,1]^n, \mathscr{A}), W_1\big)$, where the metric on which the definition of $W_1$ depends is defined by*

$$d^*_{[0,1]^n}(a, b) = \sum_{i=1}^n |a_i - b_i|, \text{ for all } a, b \in [0,1]^n.$$

*Proof.* This follows from Theorem 5.3.3 and the definition of conditional Wasserstein distances. □

If $\mathscr{B} \subseteq \mathscr{A}$, then we define $\pi_{\mathscr{B}} \colon \mathfrak{D}_{\mathcal{X}}(\mathscr{A}) \to \mathfrak{D}_{\mathcal{X}}(\mathscr{B})$ as follows: for all Borel $B \subseteq \mathcal{X}$, define $\pi_{\mathscr{B}}(\mu)(B) = \mathbb{E}\big(\mu(B) \mid \mathscr{B}\big)$ for all $\mu \in \mathfrak{D}_{\mathcal{X}}(\mathscr{A})$.

**5.8.6 Proposition.** *Let $\mathscr{A}$ be a probability algebra and let $\mu, \nu \in \mathfrak{D}_{[0,1]^n}(\mathscr{A})$. Then*

$$W_1(\mu, \nu) = \sup_{\mathscr{B} \subseteq \mathscr{A}} \big\{W_1\big(\pi_{\mathscr{B}}(\mu), \pi_{\mathscr{B}}(\nu)\big) \mid \mathscr{B} \text{ is a finite subalgebra of } \mathscr{A}\big\}.$$

*Proof.* Let $\mathcal{M}$ be a $\kappa$-universal domain for $\mathsf{ARV}$ with $\kappa > |\mathscr{A}|$. Assume $M$ is of the form $L^1\big((\Omega, \mathcal{F}, m), [0,1]\big)$ such that $\mathscr{A} \subseteq \mathcal{F}$. By Remark 5.3.10$(i)$, for all $f, g \in M^n$, we have

$$d\big(\mathrm{tp}(f/A), \mathrm{tp}(g/A)\big) = \sup_{\mathscr{B} \subseteq \mathscr{A}} \big\{ d\big(\mathrm{tp}(f/\mathscr{B}), \mathrm{tp}(g/\mathscr{B})\big) \mid \mathscr{B} \text{ is a finite subalgebra of } \mathscr{A} \big\}.$$

By Theorem 5.8.5, we have

$$W_1\big(\mathrm{dist}(f \mid \mathscr{A}), \mathrm{dist}(g \mid \mathscr{A})\big) = \sup_{\mathscr{B} \subseteq \mathscr{A}} \big\{ W_1\big(\mathrm{dist}(f \mid \mathscr{B}), \mathrm{dist}(g \mid \mathscr{B}) \mid \mathscr{B} \text{ is a finite subalgebra of } \mathscr{A} \big\}.$$

Hence,

$$W_1\big(\mathrm{dist}(f \mid \mathscr{A}), \mathrm{dist}(g \mid \mathscr{A})\big) = \sup_{\mathscr{B} \subseteq \mathscr{A}} \big\{ W_1\big(\pi_{\mathscr{B}}(\mathrm{dist}(f \mid \mathscr{A})), \pi_{\mathscr{B}}(\mathrm{dist}(g \mid \mathscr{A}))\big) \mid \mathscr{B} \text{ is finite subalgebra} \big\}.$$

Since $(\Omega, \mathcal{F}, \mu)$ is $\kappa$-saturated and $\mathscr{A}$ is small, for all $\mu, \nu \in \mathfrak{D}_{[0,1]^n}(\mathscr{A})$, there are $f, g \in M$ such that $\mathrm{dist}(f \mid \mathscr{A}) = \mu$ and $\mathrm{dist}(g \mid \mathscr{A}) = \nu$. The rest follows. □

**5.8.7 Theorem.** *Let $(\Omega, \mathscr{A}, m)$ be a probability space, where $\mathscr{A}$ is countably $\sigma$-generated. Then the topology defined using the conditional Wasserstein distance $W_1$ on $\mathfrak{D}_{[0,1]^n}(\mathscr{A})$ is the same as the topology of weak convergence if and only if $(\Omega, \mathscr{A}, m)$ is a discrete probability space.*

*Proof.* Suppose $\mathscr{A} = \sigma(f)$ for some random variable $f \colon \Omega \to [0,1]$. Then by Theorem 5.7.3 and Proposition 5.7.4, $\mathsf{ARV}(\mathsf{f})$ is separably categorical if and only if $f$ is discrete. Then by Theorem 1.9.3 (Ryll-Nardzewski Theorem), we get that in $S_n\big(\mathsf{ARV}(\mathsf{f})\big)$ the metric topology and the logic topology coincide if and only if $f$ is discrete. The rest follows from Theorem 5.3.3 and Theorem 5.8.5. □

**5.8.8 Remark.** *Comparing the preceding theorem with Theorem 2.4.7, we see the difference between conditional Wasserstein spaces over non-atomic $\mathscr{A}$ and Wasserstein spaces.*

**5.8.9 Theorem.** *Let $(\Omega, \mathscr{A}, m)$ be a probability space. Suppose that $\mathscr{A}$ is $\sigma$-generated by a subset $A$ of $L^1\big((\Omega, \mathscr{A}, m), [0,1]\big)$. Then the conditional Wasserstein space $\big(\mathfrak{D}_{[0,1]^n}(\mathscr{A}), W_1\big)$ has density character $\leq |A| + \aleph_0$.*

*Proof.* By Theorem 5.8.5, we know that there is an isometric isomorphism between $\big(S_n(A), d^*\big)$ and $\big(\mathfrak{D}_{[0,1]^n}(\mathscr{A}), W_1\big)$. Since $\mathsf{ARV}$ is $\omega$-stable, we have that $(S_n(A), d^*)$ has density character $\leq |A| + \aleph_0$. Hence, $\big(\mathfrak{D}_{[0,1]^n}(\mathscr{A}), W_1\big)$ has density character $\leq |A| + \aleph_0$. □

# Chapter 6

# Model theory of adapted structures with one conditional expectation

In this chapter and next chapter, we study the model theory of adapted structures. The theory of adapted structures provides the framework for a future study of stochastic processes.

Let $T$ be a linear ordering throughout this chapter. An *adapted structure with a $T$-indexed family of conditional expectations* is of the form $(L^1((\Omega, \mathscr{B}, \mu), [0,1]), \big(\mathbb{E}(\cdot \mid \mathscr{A}_t)\big)_{t\in T})$, where $(\Omega, \mathscr{B}, \mu)$ is a probability space, $(\mathscr{A}_t)_{t\in T}$ is an increasing family of $\sigma$-subalgebras of $\mathscr{B}$, and $\mathbb{E}(\cdot \mid \mathscr{A}_t)$ is the conditional expectation with respect to $\mathscr{A}_t$ for each $t \in T$. It is called a *totally atomless adapted structure with a $T$-indexed family of conditional expectations* if for each $t \in T$, the $\sigma$-algebra $\mathscr{A}_t$ is atomless over previous ones (if there is a minimal element 0 in $T$, then $\mathscr{A}_0$ is atomless), and $\mathscr{B}$ is atomless over $\mathscr{A}_t$. It follows from the definition of atomlessness that $\mathscr{B}$ and each $\mathscr{A}_t$ are atomless. When $T$ only contains one element, a (totally atomless) adapted structure with a $T$-indexed family of conditional expectations is called *a (totally atomless) adapted structure with one conditional expectation.*

In an early draft of [2] in 2006, Ben Yaacov augmented $L_{\mathsf{RV}}$, the signature of $[0,1]$-valued random variable structures, with a $T$-indexed family of new 1-Lipschitz unary function symbols to the signature $L_{\mathsf{RVE},0} = L_{\mathsf{RV}} \cup \{E_t\}_{t\in T}$. For each $t \in T$, the function symbol $E_t$ is interpreted as a conditional expectation of random variables. Then he also axiomatized the theories of (totally atomless) adapted structures with a $T$-indexed family of conditional expectations. Ben Yaacov announced that the theory of totally atomless adapted structures with a $T$-indexed family of conditional expectations admits quantifier elimination, is separably categorical when $T$ is finite, and is the model completion of the theory of adapted structures. Moreover, he pointed out connections with work by Fajardo and Keisler [13]. He gave a detailed discussion of presentations of canonical bases in atomless random variable structures. Further, he stated that the theory of totally atomless adapted structures is the theory of beautiful tuples of random variable structures, and that the model theoretic independence for that theory can be recovered from the general theory of beautiful tuples. The content of Ben Yaacov's paper evolved rapidly after 2006 and later versions of [2] did not contain any discussion of adapted structures.

In Chapter 6, we study the theories of (totally atomless) adapted structures with one conditional expec-

tation. In Chapter 7, we study the theories of (totally atomless) adapted structures with a $T$-indexed family of conditional expectations. Chapter 7 is parallel to Chapter 6.

In Section 6.1, we axiomatize the theory of adapted structures with one conditional expectation by RVE, and axiomatize the theory of totally atomless adapted structures with one conditional expectation by ARVE. In Section 6.2, we prove basic model theoretic properties of ARVE such as separable categoricity, quantifier elimination *etc.* In Section 6.3, we characterize the $d$-finite tuples in models of ARVE.

## 6.1 Axioms for RVE and ARVE

In this section, we axiomatize the theory of adapted structures with one conditional expectation by RVE, and axiomatize the theory of totally atomless adapted structures with one conditional expectation by ARVE.

We consider the signature $L_{\mathsf{RVE}}$ as $L_{\mathsf{RV}(\times)} \cup \{E\}$, where $E$ is a new 1-Lipschitz unary function symbol. The theory RVE consists of $\mathsf{RV}(\times)$ together with the following axioms:

(E1) $\sup_x \sup_y d\Big(E(x) \dot{-} E(y), E\big(E(x) \dot{-} E(y)\big)\Big) = 0$

(E2) $\sup_x d\big(E(\neg x), \neg E(x)\big) = 0$

(E3) $\sup_x d\big(E(\frac{x}{2}), \frac{E(x)}{2}\big) = 0$

(E4) $\sup_x \sup_y |I[xE(y)] - I[E(x)E(y)]| = 0$

In the following arguments, we will interpret symbols of $L_{\mathsf{RVE}}$ in a given model $\mathcal{M}$ of RVE without putting $\mathcal{M}$ explicitly into the notations, for easier readability. The following fact is some observations of those axioms.

**6.1.1 Fact.** *Let $\mathcal{M}$ be a model of* RVE*.*

*(i)* $E(\mathbf{0}) = \mathbf{0}$.

*(ii)* $E(E(x)) = E(x)$ *for every* $x \in M$.

*(iii)* $E(M) = \{x \in M \mid E(x) = x\}$. *Hence* $E(M)$ *is closed.*

*Proof.* $(i)$ Using (E3).

$(ii)$ In (E1), let $y = \mathbf{0}$ and then use $(i)$.

$(iii)$ For all $x \in E(M)$, there is $y \in M$ such that $x = E(y)$. By $(i)$, $E(x) = E(E(y)) = E(y) = x$. □

**6.1.2 Proposition.** *Let $\mathcal{M}$ be an $L_{\mathsf{RVE}}$-structure. Then $\mathcal{M}$ is a model of* RVE *if and only if $\mathcal{M}$ is isomorphic to an adapted structure with one conditional expectation.*

*Proof.* $\Longleftarrow$: If $M = L^1\big((\Omega, \mathscr{B}, \mu), [0,1]\big)$, where $(\Omega, \mathscr{B}, \mu)$ is a probability space and $\mathscr{A} \subseteq \mathscr{B}$ is a $\sigma$-subalgebra, then it is easy to see that $\big(M, \mathbb{E}(\cdot \mid \mathscr{A})\big)$ satisfies (E1)–(E4).

$\Longrightarrow$: If $(M, E) \models \mathsf{RVE}$, then we have $\mathcal{M} \models \mathsf{RV}(\times)$. We may assume $M$ is of the form $L^1\big((\Omega, \mathscr{B}, \mu), [0,1]\big)$ for some probability space $(\Omega, \mathscr{B}, \mu)$. Let $N$ be $E(M)$. Then by Fact 6.1.1($iii$), we know that $N$ is a closed subset. Then by the first 3 axioms of $\mathsf{RVE}$, we know that $N$ is an $L_{\mathsf{RV}(\times)}$-substructure of $M$. By Lemma 4.1.11, we know $N = L^1\big((\Omega, \mathscr{A}, \mu), [0,1]\big)$, where $\mathscr{A} = \sigma(N) \subseteq \mathscr{B}$ is a $\sigma$-subalgebra. For all $a \in \mathscr{A}$, we have $\mathbf{1}_a = [\chi_a]_{a.s.} \in N = E(M)$. By Fact 6.1.1($iii$), $E(\mathbf{1}_a) = \mathbf{1}_a$. For all $f \in M$, by (E4) we have $I[fE(\mathbf{1}_a)] = I[E(f)E(\mathbf{1}_a)]$. As $E(\mathbf{1}_a) = \mathbf{1}_a$, we have $I(f\mathbf{1}_a) = I(E(f)\mathbf{1}_a)$, that is $\int_a f d\mu = \int_a E(f) d\mu$. Since $E(f) \in N = L^1\big((\Omega, \mathscr{A}, \mu), [0,1]\big)$, we have $E(f) = \mathbb{E}(f \mid \mathscr{A})$, as required. □

The theory $\mathsf{ARVE}$ consists of the theory $\mathsf{RVE}$ together with the following two axioms:

(A1) $\sup_x \inf_y \max\Big(I(y \wedge (\neg y)), d\big(E(y \wedge x), \frac{E(x)}{2}\big)\Big) = 0$

(A2) $\sup_x \inf_y \max\Big(I(E(y) \wedge (\neg E(y))), \big|I(E(y) \wedge E(x)) - \frac{I(x)}{2}\big|\Big) = 0$

**6.1.3 Theorem.** *Let $\mathcal{M}$ be an $L_{\mathsf{RVE}}$-structure. Then $\mathcal{M}$ is a model of $\mathsf{ARVE}$ if and only if $\mathcal{M}$ is isomorphic to a totally atomless adapted structure with one conditional expectation.*

*Proof.* $\Longleftarrow$: Let $\mathcal{M}$ be the totally atomless adapted structure of the form $\Big(L^1\big((\Omega, \mathscr{B}, \mu), [0,1]\big), \mathbb{E}(\cdot \mid \mathscr{A})\Big)$, where $(\Omega, \mathscr{B}, \mu)$ is a probability space, the $\sigma$-subalgebra $\mathscr{A} \subseteq \mathscr{B}$ is atomless, and $\mathscr{B}$ is atomless over $\mathscr{A}$. By Proposition 6.1.2, we need only show that (A1) and (A2) hold in $\mathcal{M}$. By Lemma 2.2.16($ii$) and ($vi$), (A1) holds. Since $\mathscr{A}$ is atomless, (A2) holds. Hence $\mathcal{M}$ is a model of $\mathsf{ARVE}$.

$\Longrightarrow$: By Proposition 6.1.2, we know $\mathcal{M}$ is isomorphic to an adapted structure. Suppose it is of the form $\Big(L^1\big((\Omega, \mathscr{B}, \mu), [0,1]\big), \mathbb{E}(\cdot \mid \mathscr{A})\Big)$, where $(\Omega, \mathscr{B}, \mu)$ is a probability space and $\mathscr{A}$ is a $\sigma$-subalgebra of $\mathscr{B}$. By (A2), we know that $\mathscr{A}$ is atomless. So we only need show that $\mathscr{B}$ is atomless over $\mathscr{A}$. For every $A \in \mathscr{B}$ and for every $\epsilon > 0$, by Lemma 2.2.16($iv$) (Maharam's Lemma) we want to find $B \in \mathscr{A}$ such that $B \subseteq A$ and

$$\big\|E(\mathbf{1}_B) - \frac{E(\mathbf{1}_A)}{2}\big\|_1 \leq \epsilon.$$

Let $x$ denote $\mathbf{1}_A$. By (A1), there exists $y_\epsilon \in M$ such that $I\big(y_\epsilon \wedge (\neg y_\epsilon)\big) \leq \epsilon$, and

$$d(E(y_\epsilon \cdot x), \frac{E(x)}{2}) \leq \frac{\epsilon}{2}.$$

Let $D$ be the set $L^1\big((\Omega, \mathscr{B}, \mu), \{0,1\}\big)$ in $M$. By Proposition 4.1.3, we get that $\mathrm{dist}(f, D) = I(f \wedge (\neg f))$, for

all $f \in M$. By this fact, there exists $B_\epsilon \in \mathscr{B}$ such that $d(y_\epsilon, \chi_{B_\epsilon}) \leq \frac{\epsilon}{2}$. Let $z_\epsilon$ denote $\mathbf{1}_{B_\epsilon}$. Then

$$\begin{aligned} d(E(y_\epsilon x), E(z_\epsilon x)) &= \int_\Omega |E(y_\epsilon x - z_\epsilon x)| d\mu \leq \int_\Omega E(|y_\epsilon - z_\epsilon||x|) d\mu \\ &\leq \int_\Omega E(|y_\epsilon - z_\epsilon|) d\mu = \int_\Omega |y_\epsilon - z_\epsilon| d\mu = d(y_\epsilon, z_\epsilon) \leq \frac{\epsilon}{2}. \end{aligned}$$

Thus

$$d(E(z_\epsilon x), \frac{E(x)}{2}) \leq d(E(y_\epsilon x), \frac{E(x)}{2}) + d(E(y_\epsilon x), E(z_\epsilon x)) \leq \epsilon.$$

Let $B = B_\epsilon \cap A$. Then

$$\big\| E(\mathbf{1}_B) - \frac{E(\mathbf{1}_A)}{2} \big\|_1 \leq \epsilon.$$

Hence, $\mathscr{B}$ is atomless over $\mathscr{A}$. □

## 6.2 Basics of ARVE

In this section, we prove some basic model theoretic properties of ARVE. We prove that ARVE is separably categorical and admits quantifier elimination. Also we investigate the type spaces of ARVE and definable closure in models of ARVE.

We first give the following lemma.

**6.2.1 Lemma.** *Let $\mathcal{M}$ be a model of* RVE *of the form $\big(L^1((\Omega, \mathscr{B}, \mu), [0,1]), \mathbb{E}(\cdot \mid \mathscr{A})\big)$, where $(\Omega, \mathscr{B}, \mu)$ is a probability space and $\mathscr{A}$ is a $\sigma$-subalgebra of $\mathscr{B}$. Suppose $\Phi$ is an $L_{\mathsf{RV}}$-automorphism of $L^1\big((\Omega, \mathscr{B}, \mu), [0,1]\big)$ such that the restriction of $\Phi$ to $L^1\big((\Omega, \mathscr{A}, \mu), [0,1]\big)$ is an $L_{\mathsf{RV}}$-automorphism of $L^1\big((\Omega, \mathscr{A}, \mu), [0,1]\big)$. Then $\Phi$ is an $L_{\mathsf{RVE}}$-automorphism of $\mathcal{M}$.*

*Proof.* Since $\Psi$ is an $L_{\mathsf{RV}}$-automorphism of $L^1\big((\Omega, \mathscr{B}, \mu), [0,1]\big)$ and $\mathsf{RV}(\times)$ is an extension by definitions of RV, we know that $\Psi$ is also an $L_{\mathsf{RV}(\times)}$-automorphism. We need only show that $E(\Psi(f)) = \Psi(E(f))$, where $E(\cdot)$ is the conditional expectation $\mathbb{E}(\cdot \mid \mathscr{A})$, for all $f \in M$. For all $f, g \in M$, by (E4) and the facts that $\Psi^{-1}(E(g)) \in L^1\big((\Omega, \mathscr{A}, \mu), [0,1]\big)$ and $\Psi$ is an $L_{\mathsf{RV}}$-automorphism, we have

$$I\big(E(\Psi(f))E(g)\big) = I\big(\Psi(f)E(g)\big) = I\big(f \cdot \Psi^{-1}(E(g))\big) = I\big(E(f)\Psi^{-1}(E(g))\big) = I\big(\Psi(E(f))E(g)\big).$$

Hence for all $A \in \mathscr{A}$ and all $f \in M$, we have $\int_\Omega \mathbb{E}(\Psi(f) \mid \mathscr{A}) \mathbf{1}_A d\mu = \int_\Omega \Psi(\mathbb{E}(f \mid \mathscr{A})) \mathbf{1}_A d\mu$. Since the restriction of $\Psi$ to $L^1\big((\Omega, \mathscr{A}, \mu), [0,1]\big)$ is an automorphism, $\Psi(\mathbb{E}(f \mid \mathscr{A}))$ is $\mathscr{A}$-measurable. Therefore, $\Psi(\mathbb{E}(f \mid \mathscr{A})) = \mathbb{E}(\Psi(f) \mid \mathscr{A})$ for all $f \in M$, whereby $E(\Psi(f)) = \Psi(E(f))$. □

Consider a probability algebra $(\mathscr{B}, \mu)$ and a $\sigma$-complete subalgebra $\mathscr{A}$ of $\mathscr{B}$. Suppose $(\mathscr{B}, \mu)$ is the probability algebra associated to the probability space $(\Omega, \mathcal{G}, \mu)$. As explained on page 30, there is a mapping $\pi$ from the $\sigma$-algebra $\mathcal{G}$ to the $\sigma$-complete boolean algebra $\pi(\mathcal{G}) := \widehat{\mathcal{G}} = \mathscr{B}$. Let $\mathcal{F}$ denote $\pi^{-1}(\mathscr{A})$. Then because $\mathscr{A}$ is a $\sigma$-complete boolean algebra and $\mathcal{G}$ is a $\sigma$-algebra, it is routine to check that $\mathcal{F}$ is a $\sigma$-subalgebra of $\mathcal{G}$. Further, we know that $(\mathscr{A}, \mu)$ is the probability algebra associated to $(\Omega, \mathcal{F}, \mu)$. The conditional expectation $\mathbb{E}(\cdot \mid \mathcal{F})$ with respect to $\mathcal{F}$ is defined from $L^1\big((\Omega, \mathcal{G}, \mu), [0,1]\big)$ to $L^1\big((\Omega, \mathcal{F}, \mu), [0,1]\big)$. Therefore, $\mathbb{E}(\cdot \mid \mathcal{F})$ is defined from $L^1(\mathscr{B}, \mu) := L^1\big((\Omega, \mathcal{G}, \mu), [0,1]\big)$ to $L^1(\mathscr{A}, \mu) := L^1\big((\Omega, \mathcal{F}, \mu), [0,1]\big)$, denoted by $\mathbb{E}(\cdot \mid \mathscr{A})$.

By Theorem 5.1.1, we see that a probability algebra $(\mathscr{B}, \mu)$ and a $\sigma$-complete subalgebra $\mathscr{A}$ of $\mathscr{B}$ determine an adapted structure with one conditional expectation, denoted by $\big(L^1(\mathscr{B}, [0,1]), \mathbb{E}(\cdot \mid \mathscr{A})\big)$, up to $L_{\mathsf{RVE}}$-isomorphism.

## Separable categoricity

In his early draft of [2], Ben Yaacov announced that the theory $\mathsf{ARVE}$ is separably categorical.

**6.2.2 Proposition.** *Let $(\mathscr{B}_i, \mu_i)$ be a probability algebra and let $\mathscr{A}_i \subseteq \mathscr{B}_i$ be a $\sigma$-complete subalgebra for each $i = 1, 2$. Suppose $(\mathscr{A}_i, \mu_i)$ is atomless, $(\mathscr{B}_i, \mu_i)$ is atomless over $(\mathscr{A}_i, \mu_i)$, and $L^1(\mathscr{B}_i, [0,1])$ is separable, for each $i = 1, 2$. For every $L_{\mathsf{RV}}$-isomorphism $\varphi\colon L^1(\mathscr{A}_1, [0,1]) \to L^1(\mathscr{A}_2, [0,1])$, there is an $L_{\mathsf{RVE}}$-isomorphism*

$$\Phi\colon \big(L^1(\mathscr{B}_1, [0,1]), \mathbb{E}(\cdot \mid \mathscr{A}_1)\big) \to \big(L^1(\mathscr{B}_2, [0,1]), \mathbb{E}(\cdot \mid \mathscr{A}_2)\big)$$

*extending $\varphi$.*

*Proof.* Let $(A_\omega, \mu_\omega)$ be the probability algebra associated to $([0,1]^\omega, \mu_\omega)$. Note that $(A_\omega, \mu_\omega)$ is separable and atomless. The free product of $(A_\omega, \mu_\omega)$ and $(A_\omega, \mu_\omega)$ is denoted by $(A_\omega \widehat{\times} A_\omega, \nu)$. For each $i = 1, 2$, since $L^1(\mathscr{B}_i, [0,1])$ is separable, we have that $\mathscr{B}_i$ and $\mathscr{A}_i$ are both separable atomless probability algebras. By Theorem 3.1.3, $\mathsf{APr}$ is separably categorical. Hence there is a measure-preserving isomorphism $\psi_2\colon (\mathscr{A}_2, \mu_2) \to (A_\omega, \mu_\omega)$. By Theorem 5.1.1, $\varphi$ induces a natural measure-preserving isomorphism $\phi$ from $(\mathscr{A}_1, \mu_1)$ to $(\mathscr{A}_2, \mu_2)$. Let $\psi_1$ be $\psi_2 \circ \phi$. Then $\psi_1$ is a measure-preserving isomorphism from $(\mathscr{A}_1, \mu_1)$ to $(A_\omega, \mu_\omega)$. Define $\alpha_1\colon [0,1]^\omega \to [0,1]^\omega \times [0,1]^\omega$ by $\alpha(A) := A \times [0,1]^\omega$ for all measurable $A \subseteq [0,1]^\omega$, and define $\alpha_2\colon [0,1]^\omega \to [0,1]^\omega \times [0,1]^\omega$ by $\alpha(A) := [0,1]^\omega \times A$ for all measurable $A \subseteq [0,1]^\omega$. Then $\alpha_i$ induces a measure-preserving homomorphism $\iota_i\colon (A_\omega, \mu_\omega) \to (A_\omega \widehat{\times} A_\omega, \nu)$ for each $i = 1, 2$. Since the image of $\alpha_1$ only depends on the first coordinate and the image of $\alpha_2$ only depends on the second coordinate, we know that the image of $\alpha_1$ and the image of $\alpha_2$ are independent. Then by Lemma 2.2.19($ii$), we know that $[0,1]^\omega \times [0,1]^\omega$

is atomless over the image of $\alpha_1$. Then by Lemma 2.2.16, we have $A_\omega \widehat{\times} A_\omega$ is atomless over $\iota_1[A_\omega]$. Then $e_i := \iota_1\psi_i \colon (\mathscr{A}_i, \mu_i) \to (A_\omega \widehat{\times} A_\omega, \nu)$ is a measure-preserving homomorphism for each $i = 1, 2$. Note that $e_i[\mathscr{A}_i] = \iota_1[A_\omega]$ for each $i = 1, 2$. We have that $\mathscr{B}_i$ is atomless over $\mathscr{A}_i$, and $L^1(\mathscr{B}_i, [0,1])$ is separable for each $i = 1, 2$. By Lemma 2.2.17, $\tau_{\mathscr{A}_1}(b_1) = \tau_{\mathscr{A}_2}(b_2) = \omega$ for all $b_1 \in \mathscr{B}_1^+$ and all $b_2 \in \mathscr{B}_2^+$. By Theorem 2.2.27, $e_i$ can be extended to an isomorphism $f_i \colon (\mathscr{B}_i, \mu_i) \to (A_\omega \widehat{\times} A_\omega, \nu)$ for each $i = 1, 2$. By Theorem 5.1.1, $f_i$ induces a natural $L_{\mathsf{RV}}$-isomorphism $\widehat{f_i}$ from $L^1(\mathscr{B}_i, [0,1])$ to $L^1(A_\omega \widehat{\times} A_\omega, [0,1])$, and $e_i$ induces a natural $L_{\mathsf{RV}}$-isomorphism $\widehat{e_i}$ from $L^1(\mathscr{A}_i, [0,1])$ to $L^1(\iota_1[A_\omega], [0,1])$, for each $i = 1, 2$. Moreover, we have $\widehat{e_2} \circ \varphi = \widehat{e_1}$. Since $f_i$ extends $e_i$ for each $i = 1, 2$, we have that $\widehat{f_i}$ extends $\widehat{e_i}$. Then by Lemma 6.2.1, we know that $\widehat{f_i}$ is an $L_{\mathsf{RVE}}$-isomorphism from $(L^1(\mathscr{B}_i, [0,1]), \mathbb{E}(\cdot \mid \mathscr{A}_i))$ to $(L^1(A_\omega \widehat{\times} A_\omega, [0,1]), \mathbb{E}(\cdot \mid \iota_1[A_\omega]))$. Then $\widehat{f_2}^{-1} \circ \widehat{f_1}$ is our desired $\Phi$. □

**6.2.3 Corollary.** *The theory* ARVE *is separably categorical, and hence it is complete.* □

**6.2.4 Proposition.** *The theory* ARVE *is model complete.*

*Proof.* From the form of its axioms, ARVE is an $\forall\exists$ theory. Then by Theorem 1.8.2 (Lindström's test for model completeness) and Corollary 6.2.3 (ARVE is separably categorical), ARVE is model complete. □

## Definable closures

Let $\mathcal{M}$ be a model of ARVE and let $A$ be a subset. Here, we use the notation $\mathrm{dcl}^E(A)$ for $\mathrm{dcl}^{\mathsf{ARVE}}_{\mathcal{M}}(A)$ and the notation $\mathrm{dcl}(A)$ for $\mathrm{dcl}^{\mathsf{ARV}}_{\mathcal{M}}(A)$.

**6.2.5 Proposition.** *Let* $\mathcal{M} \models$ ARVE *and let* $A$ *be a subset of* $M$. *Then* $\mathrm{dcl}^E(A) \supseteq \mathrm{dcl}\big(E(\mathrm{dcl}(A)) \cup A\big)$.

*Proof.* For any subset $B$ of $M$, if $a \in M$ is definable over $B$ in the reduct of $\mathcal{M}$ to $L_{\mathsf{RV}}$, then $a$ is also definable over $B$ in the $L_{\mathsf{RVE}}$-structure $\mathcal{M}$. Hence $\mathrm{dcl}(B) \subseteq \mathrm{dcl}^E(B)$ for any subset $B \subseteq M$. Then by Proposition 1.6.11,

$$\mathrm{dcl}^E(A) \subseteq \mathrm{dcl}\big(\mathrm{dcl}^E(A)\big) \subseteq \mathrm{dcl}^E\big(\mathrm{dcl}^E(A)\big) = \mathrm{dcl}^E(A),$$

whereby $\mathrm{dcl}\big(\mathrm{dcl}^E(A)\big) = \mathrm{dcl}^E(A)$. Note that $\mathrm{dcl}^E(A) \supseteq A$ and $E\big(\mathrm{dcl}^E(A)\big) \subseteq \mathrm{dcl}^E\big(\mathrm{dcl}^E(A)\big) = \mathrm{dcl}^E(A)$. Then $E\big(\mathrm{dcl}(A)\big) \cup A \subseteq \mathrm{dcl}^E(A) = \mathrm{dcl}\big(\mathrm{dcl}^E(A)\big)$, whence by $\mathrm{dcl}\big(\mathrm{dcl}^E(A)\big) = \mathrm{dcl}^E(A)$ and Proposition 1.6.11,

$$\mathrm{dcl}\big(E(\mathrm{dcl}(A) \cup A)\big) \subseteq \mathrm{dcl}\big(\mathrm{dcl}^E(A)\big) = \mathrm{dcl}^E(A).$$

□

**6.2.6 Question.** *Is it true that* $\mathrm{dcl}^E(A) = \mathrm{dcl}\big(E(\mathrm{dcl}(A)) \cup A\big)$*?*

## Type spaces

Our main results are Theorems 6.2.8 and 6.2.9.

**6.2.7 Lemma.** *Let $\mathcal{M} \models \mathsf{ARVE}$ be of the form $\Big(L^1\big((\Omega, \mathscr{B}, \mu), [0,1]\big), \mathbb{E}(\cdot \mid \mathscr{A})\Big)$, where $(\Omega, \mathscr{B}, \mu)$ is a probability space, the $\sigma$-subalgebra $\mathscr{A} \subseteq \mathscr{B}$ is atomless, and $\mathscr{B}$ is atomless over $\mathscr{A}$. Suppose $\mathscr{A}$ is the $\sigma$-algebra generated by the random variables in $A \subseteq M$. For all $a, b \in M^n$ and $C \subseteq M$, we have:*

$$\mathrm{tp}^{\mathsf{ARV}}(a/C \cup A) = \mathrm{tp}^{\mathsf{ARV}}(b/C \cup A) \Longrightarrow \mathrm{tp}^{\mathsf{ARVE}}(a/C) = \mathrm{tp}^{\mathsf{ARVE}}(b/C).$$

*Proof.* Let $\mathcal{N} \models \mathsf{ARV}(\times)$ be a $\kappa$-universal domain for an uncountable cardinal $\kappa$. Without loss of generality, we may assume that $\mathcal{N}$ is of the form $L^1\big((\Omega, \mathscr{C}, \mu), [0,1]\big)$ where $\mathscr{A} \subseteq \mathscr{B} \subseteq \mathscr{C}$ and $\mathscr{A}, \mathscr{B}$ are small. Then by interpreting $E$ as $\mathbb{E}(\cdot \mid \mathscr{A})$ in $\mathcal{N}$, the $L_{\mathsf{RV}(\times)}$-structure $\mathcal{N}$ is expanded to become an $L_{\mathsf{RVE}}$-structure. Since $\mathscr{A}$ is small, by Theorems 3.4.2 and 5.5.3 we know that $\mathscr{C}$ is also atomless over $\mathscr{A}$, whence $\mathcal{N}$ is a model of $\mathsf{ARVE}$. Note that $\mathcal{M} \subseteq \mathcal{N}$. Then by Proposition 6.2.4 ($\mathsf{ARVE}$ is model complete), we have that $\mathcal{M} \preceq \mathcal{N}$. Since $C$ and $A$ are small and $\mathrm{tp}^{\mathsf{ARV}}(a/C \cup A) = \mathrm{tp}^{\mathsf{ARV}}(b/C \cup A)$, there is $\Phi \in \mathrm{Aut}_{\mathsf{RV}}\big(L^1((\Omega, \mathscr{C}, \mu), [0,1])\big)$ such that $\Phi(a) = b$ and $\Phi$ fixes $C \cup A$ pointwise. Since $\Phi$ fixes $A$ and $\sigma(A) = \mathscr{A}$, we have $\Phi \restriction_{L^1((\Omega, \mathscr{A}, \mu), [0,1])} \in \mathrm{Aut}_{\mathsf{RV}}\big(L^1((\Omega, \mathscr{A}, \mu), [0,1])\big)$. Then by Lemma 6.2.1, $\Phi \in \mathrm{Aut}_{\mathsf{RVE}}\Big(L^1\big((\Omega, \mathscr{C}, \mu), [0,1]\big), \mathbb{E}(\cdot \mid \mathscr{A})\Big)$ such that $\Phi(a) = b$ and $\Phi$ fixes $C$. Therefore $\mathrm{tp}^{\mathsf{ARVE}}(a/C) = \mathrm{tp}^{\mathsf{ARVE}}(b/C)$. □

For all $t = (t_1, \cdots, t_n) \in [0,1]^n$, define $B_t := [0, t_1) \times \cdots \times [0, t_n)$.

**6.2.8 Theorem.** *Let $\mathcal{M} \models \mathsf{ARVE}$ be a $\kappa$-universal domain where $\kappa > 2^\omega$. Assume that $\mathcal{M}$ is of the form $\Big(L^1\big((\Omega, \mathscr{B}, \mu), [0,1]\big), \mathbb{E}(\cdot \mid \mathscr{A})\Big)$, where $(\Omega, \mathscr{B}, \mu)$ is a probability space, the $\sigma$-subalgebra $\mathscr{A} \subseteq \mathscr{B}$ is atomless, and $\mathscr{B}$ is atomless over $\mathscr{A}$. Then for all $a, b \in M^n$, we have $\mathrm{tp}(a) = \mathrm{tp}(b)$ if and only if*

$$\mathrm{dist}\big(\mathbb{E}(a \in B_{t^1} \mid \mathscr{A}), \cdots, \mathbb{E}(a \in B_{t^m} \mid \mathscr{A})\big) = \mathrm{dist}\big(\mathbb{E}(b \in B_{t^1} \mid \mathscr{A}), \cdots, \mathbb{E}(b \in B_{t^m} \mid \mathscr{A})\big)$$

*for all $m \in \mathbb{N}$ and for all $t^1, \cdots, t^m \in [0,1]^n$.*

*Proof.* $\Longrightarrow$: Since $\mathcal{M}$ is a $\kappa$-universal domain and $\mathrm{tp}(a) = \mathrm{tp}(b)$, there is an automorphism $\Phi$ of $\mathcal{M}$ such that $\Phi(a) = b$. By Lemma 5.2.5, we have that

$$\mathrm{tp}(\mathbf{1}_{a \in B_{t^1}}, \cdots, \mathbf{1}_{a \in B_{t^m}}) = \mathrm{tp}(\mathbf{1}_{b \in B_{t^1}}, \cdots, \mathbf{1}_{b \in B_{t^m}})$$

for all $m \in \mathbb{N}$ and all $t^1, \cdots, t^m \in [0,1]^n$, and thus

$$\mathrm{tp}\big(E(a \in B_{t^1}), \cdots, E(a \in B_{t^m})\big) = \mathrm{tp}\big(E(b \in B_{t^1}), \cdots, E(b \in B_{t^m})\big).$$

Then use Theorem 5.2.3($v$).

$\Longleftarrow$: Note that $E(f) = \mathbb{E}(f \mid \mathscr{A})$ for all $f \in M$. By assumption, we know that

$$\mathrm{dist}\big(E(a \in B_{t^1}), \cdots, E(a \in B_{t^m})\big) = \mathrm{dist}\big(E(b \in B_{t^1}), \cdots, E(b \in B_{t^m})\big).$$

Then by Theorem 5.2.3($v$), for all $t^1, \cdots, t^m \in [0,1]^n$ we have that

$$\mathrm{tp}^{\mathsf{ARV}}\big(E(a \in B_{t^1}), \cdots, E(a \in B_{t^m})\big) = \mathrm{tp}^{\mathsf{ARV}}\big(E(b \in B_{t^1}), \cdots, E(b \in B_{t^m})\big).$$

Thus, we have

$$\mathrm{tp}^{\mathsf{ARV}}\big(\{E(a \in B_t)\}_{t \in [0,1]^n}\big) = \mathrm{tp}^{\mathsf{ARV}}\big(\{E(b \in B_t)\}_{t \in [0,1]^n}\big).$$

Let $\lambda$ be a cardinal $> 2^{|\mathscr{B}|}$. Let $\mathcal{N} \models \mathsf{ARV}(\times)$ be a $\lambda$-universal domain and we may assume that $\mathcal{N}$ is of the form $L^1\big((\Omega, \mathscr{C}, \mu), [0,1]\big)$ where $\mathscr{A} \subseteq \mathscr{B} \subseteq \mathscr{C}$ and $\mathscr{A}, \mathscr{B}$ are small. Since $\mathscr{A}$ is small, by Theorems 3.4.2 and 5.5.3 we have that $\mathscr{C}$ is also atomless over $\mathscr{A}$. By interpreting $E$ as $\mathbb{E}(\cdot \mid \mathscr{A})$ in $\mathcal{N}$, we expand $\mathcal{N}$ to a model of $\mathsf{ARVE}$. Note that $\mathcal{M} \subseteq \mathcal{N}$. By Proposition 6.2.4, we have $\mathcal{M} \preceq \mathcal{N}$. Since $\mathcal{M}$ is a $\kappa$-universal domain, $L^1((\Omega, \mathscr{B}, \mu), [0,1])$ and $L^1((\Omega, \mathscr{A}, \mu), [0,1])$ are $\kappa$-universal domains for $\mathsf{ARV}(\times)$. Hence there is an automorphism $\Phi$ of the $L_{\mathsf{RV}(\times)}$-structure $L^1((\Omega, \mathscr{A}, \mu), [0,1])$ such that $\Phi(E(a \in B_t)) = E(b \in B_t)$ for every $t \in [0,1]^n$. Then by the fact that $\mathcal{N}$ as a model of $\mathsf{ARV}(\times)$ is a $\lambda$-universal domain with $\lambda > 2^{|\mathscr{B}|} \geq 2^{|\mathscr{A}|}$, there is an automorphism $\Psi$ of the $L_{\mathsf{RV}(\times)}$-structure $L^1((\Omega, \mathscr{C}, \mu), [0,1])$ extending $\Phi$. Since $\Psi$ extends $\Phi$, by Lemma 6.2.1 we know that $\Psi$ is an automorphism of the $L_{\mathsf{RVE}}$-structure $\mathcal{N}$.

Let $a' = \Psi(a)$. By Lemma 5.2.5, we have $\Psi(\mathbf{1}_{f \in B_t}) = \mathbf{1}_{\Psi(f) \in B_t}$ for all $t \in [0,1]^n$ and all $f \in N^n$. Then because $\Psi$ is an automorphism of the $L_{\mathsf{RVE}}$-structure $\mathcal{N}$, we have

$$E(a' \in B_t) = E(\mathbf{1}_{\Psi(a) \in B_t}) = E(\Psi(\mathbf{1}_{a \in B_t})) = \Psi(E(\mathbf{1}_{a \in B_t})) = \Psi(E(a \in B_t)),$$

for every $t \in [0,1]^n$. Therefore, $E(a' \in B_t) = E(b \in B_t)$ for every $t \in [0,1]^n$. That is, $a'$ and $b$ have the same conditional distribution over $\mathscr{A}$. By Theorem 5.2.3($v$) and Lemma 6.2.7, we have $\mathrm{tp}^{\mathsf{ARVE}}(a') = \mathrm{tp}^{\mathsf{ARVE}}(b)$, whence $\mathrm{tp}^{\mathsf{ARVE}}(a) = \mathrm{tp}^{\mathsf{ARVE}}(b)$. □

**6.2.9 Theorem.** *Let $\mathcal{M} \models \mathsf{ARVE}$ be a $\kappa$-universal domain where $\kappa > 2^\omega$. Assume that $\mathcal{M}$ is of the form*

$\Big(L^1\big((\Omega,\mathscr{B},\mu),[0,1]\big),\mathbb{E}(\cdot\mid\mathscr{A})\Big)$, *where* $(\Omega,\mathscr{B},\mu)$ *is a probability space, the* $\sigma$*-subalgebra* $\mathscr{A}\subseteq\mathscr{B}$ *is atomless, and* $\mathscr{B}$ *is atomless over* $\mathscr{A}$*. For all* $a,b\in M^n$*, we have* $\mathrm{tp}(a)=\mathrm{tp}(b)$ *if and only if*

$$\mathrm{dist}\big(\mathbb{E}(\varphi_1(a)\mid\mathscr{A}),\cdots,\mathbb{E}(\varphi_m(a)\mid\mathscr{A})\big)=\mathrm{dist}\big(\mathbb{E}(\varphi_1(b)\mid\mathscr{A}),\cdots,\mathbb{E}(\varphi_m(b)\mid\mathscr{A})\big),$$

*for all* $m\in\mathbb{N}$ *and all continuous* $\varphi_1,\cdots,\varphi_m\colon[0,1]^n\to[0,1]$.

*Proof.* The proof is a revision of the proof of Theorem 6.2.8.

$\Longrightarrow$: Since $\mathcal{M}$ is a $\kappa$-universal domain and $\mathrm{tp}(a)=\mathrm{tp}(b)$, there is an automorphism $\Phi$ of $\mathcal{M}$ such that $\Phi(a)=b$. By Lemma 5.2.5, we have that $\mathrm{tp}(\varphi_1(a),\cdots,\varphi_m(a))=\mathrm{tp}(\varphi_1(b),\cdots,\varphi_m(b))$ for all $m\in\mathbb{N}$ and for all continuous $\varphi_1,\cdots,\varphi_m\colon[0,1]^n\to[0,1]$, and thus we have $\mathrm{tp}(E(\varphi_1(a)),\cdots,E(\varphi_m(a)))=\mathrm{tp}(E(\varphi_1(b)),\cdots,E(\varphi_m(b)))$. Then use Theorem 5.2.3($v$).

$\Longleftarrow$: Note that $E(f)=\mathbb{E}(f\mid\mathscr{A})$ for all $f\in M$. By assumption, for all $m\in\mathbb{N}$ and all continuous $\varphi_1,\cdots,\varphi_m\colon[0,1]^n\to[0,1]$, we have that

$$\mathrm{dist}\big(E(\varphi_1(a)),\cdots,E(\varphi_m(a))\big)=\mathrm{dist}\big(E(\varphi_1(b)),\cdots,E(\varphi_m(b))\big).$$

Then by Theorem 5.2.3($v$), for all $m\in\mathbb{N}$ and all continuous $\varphi_1,\cdots,\varphi_m\colon[0,1]^n\to[0,1]$, we have that

$$\mathrm{tp}^{\mathsf{ARV}}\big(E(\varphi_1(a)),\cdots,E(\varphi_m(a))\big)=\mathrm{tp}^{\mathsf{ARV}}\big(E(\varphi_1(b)),\cdots,E(\varphi_m(b))\big).$$

Thus, we have

$$\mathrm{tp}^{\mathsf{ARV}}\big(\{E(\varphi(a))\mid\varphi\colon[0,1]\to[0,1]\text{ is continuous}\}\big)=\mathrm{tp}^{\mathsf{ARV}}\big(\{E(\varphi(b))\mid\varphi\colon[0,1]\to[0,1]\text{ is continuous}\}\big).$$

Let $\lambda$ be a cardinal $>2^{|\mathscr{B}|}$. Let $\mathcal{N}\models\mathsf{ARV}(\times)$ be a $\lambda$-universal domain and we may assume that $\mathcal{N}$ is of the form $L^1\big((\Omega,\mathscr{C},\mu),[0,1]\big)$ where $\mathscr{A}\subseteq\mathscr{B}\subseteq\mathscr{C}$ and $\mathscr{A},\mathscr{B}$ are small. Since $\mathscr{A}$ is small, by Theorems 3.4.2 and 5.5.3 we have that $\mathscr{C}$ is also atomless over $\mathscr{A}$. By interpreting $E$ as $\mathbb{E}(\cdot\mid\mathscr{A})$ in $\mathcal{N}$, we expand $\mathcal{N}$ to a model of $\mathsf{ARVE}$. Note that $\mathcal{M}\subseteq\mathcal{N}$. By Proposition 6.2.4, we have $\mathcal{M}\preceq\mathcal{N}$. Since $\mathcal{M}$ is a $\kappa$-universal domain, $L^1((\Omega,\mathscr{B},\mu),[0,1])$ and $L^1((\Omega,\mathscr{A},\mu),[0,1])$ are $\kappa$-universal domains for $\mathsf{ARV}(\times)$. Hence there is an automorphism $\Phi$ of the $L_{\mathsf{RV}(\times)}$-structure $L^1((\Omega,\mathscr{A},\mu),[0,1])$ such that $\Phi(E(\varphi(a)))=E(\varphi(b))$ for all continuous $\varphi\colon[0,1]^n\to[0,1]$. Then by the fact that $\mathcal{N}$ as a model of $\mathsf{ARV}(\times)$ is a $\lambda$-universal domain with $\lambda>2^{|\mathscr{B}|}\geq2^{|\mathscr{A}|}$, there is an automorphism $\Psi$ of the $L_{\mathsf{RV}(\times)}$-structure $L^1((\Omega,\mathscr{C},\mu),[0,1])$ extending $\Phi$. Since $\Psi$ extends $\Phi$, by Lemma 6.2.1 we know that $\Psi$ is an automorphism of the $L_{\mathsf{RVE}}$-structure $\mathcal{N}$.

Let $a' = \Psi(a)$. By Lemma 5.2.5, we have that $\Psi(\varphi(a)) = \varphi(\Psi(a)) = \varphi(a')$ for all continuous $\varphi\colon [0,1]^n \to [0,1]$. Because $\Psi$ extends $\Phi$, for all continuous $\varphi\colon [0,1]^n \to [0,1]$ we have $\Psi(E(\varphi(a))) = E(\varphi(b))$. Thus $E(\Psi(\varphi(a))) = \Psi(E(\varphi(a))) = E(\varphi(b))$, since $\Psi$ is an automorphism of the $L_{\mathsf{RVE}}$-structure $\mathcal{N}$. Then by the fact that $\Psi(\varphi(a)) = \varphi(\Psi(a)) = \varphi(a')$, we have $E(\varphi(a')) = E(\varphi(b))$ for all continuous $\varphi\colon [0,1]^n \to [0,1]$. Using Lusin's theorem and the dominated convergence theorem, we have that $E(a' \in B) = E(b \in B)$ for all Borel $B \subseteq [0,1]^n$. That is, $a'$ and $b$ determine the same conditional distribution over $\mathscr{A}$. By Theorem 5.2.3($v$) and Lemma 6.2.7, we have $\mathrm{tp}^{\mathsf{ARVE}}(a') = \mathrm{tp}^{\mathsf{ARVE}}(b)$, whence $\mathrm{tp}^{\mathsf{ARVE}}(a) = \mathrm{tp}^{\mathsf{ARVE}}(b)$. □

**6.2.10 Corollary.** *Let $\mathcal{M} \models \mathsf{ARVE}$ be a $\kappa$-universal domain where $\kappa > 2^\omega$. Assume that $\mathcal{M}$ is of the form $\Big(L^1\big((\Omega, \mathscr{B}, \mu), [0,1]\big), \mathbb{E}(\cdot \mid \mathscr{A})\Big)$, where $(\Omega, \mathscr{B}, \mu)$ is a probability space, the $\sigma$-subalgebra $\mathscr{A} \subseteq \mathscr{B}$ is atomless, and $\mathscr{B}$ is atomless over $\mathscr{A}$. For all $a, b \in M^n$ and small $C \subseteq M$, the following are equivalent:*

*(i)* $\mathrm{tp}(a/C) = \mathrm{tp}(b/C)$.

*(ii) For all $k, l \in \mathbb{N}$, all $t^1, \cdots, t^k \in [0,1]^{n+l}$, and all $c_1, \cdots, c_l \in C$, one has*

$$\begin{aligned}&\mathrm{dist}\Big(\mathbb{E}\big((a, c_1, \cdots, c_l) \in B_{t^1} \mid \mathscr{A}\big), \cdots, \mathbb{E}\big((a, c_1, \cdots, c_l) \in B_{t^k} \mid \mathscr{A}\big)\Big)\\ =&\mathrm{dist}\Big(\mathbb{E}\big((b, c_1, \cdots, c_l) \in B_{t^1} \mid \mathscr{A}\big), \cdots, \mathbb{E}\big((b, c_1, \cdots, c_l) \in B_{t^k} \mid \mathscr{A}\big)\Big).\end{aligned}$$

*(iii) For all $k, l \in \mathbb{N}$, all continuous $\varphi_1, \cdots, \varphi_k\colon [0,1]^{n+l} \to [0,1]$, and all $c_1, \cdots, c_l \in C$, one has*

$$\begin{aligned}&\mathrm{dist}\Big(\mathbb{E}\big(\varphi_1(a, c_1, \cdots, c_l) \mid \mathscr{A}\big), \cdots, \mathbb{E}\big(\varphi_k(a, c_1, \cdots, c_l) \mid \mathscr{A}\big)\Big)\\ =&\mathrm{dist}\Big(\mathbb{E}\big(\varphi_1(b, c_1, \cdots, c_l) \mid \mathscr{A}\big), \cdots, \mathbb{E}\big(\varphi_k(b, c_1, \cdots, c_l) \mid \mathscr{A}\big)\Big).\end{aligned}$$

*Proof.* $(i) \Rightarrow (ii)$ and $(i) \Rightarrow (iii)$: Since $\mathrm{tp}(a/C) = \mathrm{tp}(b/C)$, we have $\mathrm{tp}(a, c_1, \cdots, c_l) = \mathrm{tp}(b, c_1, \cdots, c_l)$ for all $l \in \mathbb{N}$ and all $c_1, \cdots, c_l \in C$. Then use Theorem 6.2.8 and Theorem 6.2.9 respectively.

$(ii) \Rightarrow (i)$: By Theorem 6.2.8, $\mathrm{tp}(a, c_1, \cdots, c_l) = \mathrm{tp}(b, c_1, \cdots, c_l)$ for all $l \in \mathbb{N}$ and all $c_1, \cdots, c_l \in C$. Therefore $\mathrm{tp}(a/C) = \mathrm{tp}(b/C)$.

$(iii) \Rightarrow (i)$: By Theorem 6.2.9, $\mathrm{tp}(a, c_1, \cdots, c_l) = \mathrm{tp}(b, c_1, \cdots, c_l)$ for all $l \in \mathbb{N}$ and all $c_1, \cdots, c_l \in C$. Therefore $\mathrm{tp}(a/C) = \mathrm{tp}(b/C)$. □

## Quantifier Elimination

**6.2.11 Theorem.** *The theory $\mathsf{ARVE}$ admits quantifier elimination.*

*Proof.* Let $\mathcal{M} \models \mathsf{ARVE}$. Let $a, b \in M^n$ with $\mathrm{tp}^{\mathsf{ARVE},\mathrm{qf}}(a) = \mathrm{tp}^{\mathsf{ARVE},\mathrm{qf}}(b)$. By Lemma 5.2.1, for all $m \in \mathbb{N}$ and all continuous $\varphi_1, \cdots, \varphi_m \colon [0,1]^n \to [0,1]$, there are sequences of $n$-ary $L_{\mathsf{RV}}$-terms $(t_{k,\varphi_1})_{k\in\mathbb{N}}, \cdots, (t_{k,\varphi_m})_{k\in\mathbb{N}}$ such that for each $j = 1, \cdots, m$, we have that $t_{k,\varphi_j}^{\mathcal{M}}$ uniformly converges to $\varphi_j^{\mathcal{M}}$, where $\varphi_j^{\mathcal{M}} \colon M^n \to [0,1]$ is defined by $\varphi_j^{\mathcal{M}}(f) = \varphi_j(f)$ for all $f \in M^n$. Then for every quantifier free $L_{\mathsf{RVE}}$-formula $F(z)$ where $z$ is a tuple of $m$ variables, by the fact that $F^{\mathcal{M}}$ and $E^{\mathcal{M}}$ are continuous, we have that $F^{\mathcal{M}}\big(E^{\mathcal{M}}(t_{k,\varphi_1}^{\mathcal{M}}(f)), \cdots, E^{\mathcal{M}}(t_{k,\varphi_m}^{\mathcal{M}}(f))\big)$ converges to $F^{\mathcal{M}}\big(E^{\mathcal{M}}(\varphi_1(f)), \cdots, E^{\mathcal{M}}(\varphi_m(f))\big)$ for all $f \in M^n$. Then because $F\big(E(t_{k,\varphi_1}), \cdots, E(t_{k,\varphi_m})\big)$ is a quantifier free $L_{\mathsf{RVE}}$-formula for each $k \in \mathbb{N}$ and $\mathrm{tp}^{\mathsf{ARVE},\mathrm{qf}}(a) = \mathrm{tp}^{\mathsf{ARVE},\mathrm{qf}}(b)$, we have that

$$F^{\mathcal{M}}\big(E^{\mathcal{M}}(\varphi_1(a)), \cdots, E^{\mathcal{M}}(\varphi_m(a))\big) = F^{\mathcal{M}}\big(E^{\mathcal{M}}(\varphi_1(b)), \cdots, E^{\mathcal{M}}(\varphi_m(b))\big).$$

Hence,

$$\mathrm{tp}^{\mathsf{ARVE},\mathrm{qf}}\big(E(\varphi_1(a)), \cdots, E(\varphi_m(a))\big) = \mathrm{tp}^{\mathsf{ARVE},\mathrm{qf}}\big(E(\varphi_1(b)), \cdots, E(\varphi_m(b))\big),$$

and thus

$$\mathrm{tp}^{\mathsf{ARV},\mathrm{qf}}\big(E(\varphi_1(a)), \cdots, E(\varphi_m(a))\big) = \mathrm{tp}^{\mathsf{ARV},\mathrm{qf}}\big(E(\varphi_1(b)), \cdots, E(\varphi_m(b))\big),$$

for all $m \in \mathbb{N}$ and all continuous $\varphi_1, \cdots, \varphi_m \colon [0,1]^n \to [0,1]$. By Theorem 5.2.3, we have that

$$\mathrm{dist}\big(E(\varphi_1(a)), \cdots, E(\varphi_m(a))\big) = \mathrm{dist}\big(E(\varphi_1(b)), \cdots, E(\varphi_m(b))\big),$$

for all $m \in \mathbb{N}$ and all continuous $\varphi_1, \cdots, \varphi_m \colon [0,1]^n \to [0,1]$. By Theorem 6.2.9, we know that $\mathrm{tp}^{\mathsf{ARVE}}(a) = \mathrm{tp}^{\mathsf{ARVE}}(b)$. Then by Theorem 1.10.4, we have that $\mathsf{ARVE}$ admits quantifier elimination. □

**6.2.12 Proposition.** *The universal part of* $\mathsf{ARVE}$ *is* $\mathsf{RVE}$*, and* $\mathsf{ARVE}$ *is the model completion of* $\mathsf{RVE}$.

*Proof.* By Corollary 4.1.12, we know that $\mathsf{RV}$ is a universal theory. Then from the form of the axioms in $\mathsf{RVE} \setminus \mathsf{RV}$ (see page 81 for $\mathsf{RV}(\times) \setminus \mathsf{RV}$ and page 125 for $\mathsf{RVE} \setminus \mathsf{RV}(\times)$), we know that $\mathsf{RVE}$ is a universal theory. By Proposition 6.1.2 and Theorem 6.1.3, we have that the $L_{\mathsf{RVE}}$-substructures of models of $\mathsf{ARVE}$ are precisely the models of $\mathsf{RVE}$. Therefore the theory $\mathsf{RVE}$ is the universal part of the theory $\mathsf{ARVE}$. Then by Theorem 6.2.11 and Proposition 1.10.5, we know that $\mathsf{ARVE}$ is the model completion of $\mathsf{RVE}$. □

## 6.3 On $d$-finite tuples in ARVE

In this section, we characterize the $d$-finite tuples in models of $\mathsf{ARVE}$.

Let $\mathcal{M} \models \mathsf{ARVE}$. For a tuple $f = (f_1, \cdots, f_n) \in M^n$, let $\mathsf{ARVE}(\mathsf{f})$ denote the theory $Th(\mathcal{M}, f)$. By

Corollary 6.2.3 ($\mathsf{ARVE}$ is separably categorical) and Theorem 3.5.2, $f$ is $d$-finite (over $\emptyset$) if and only if $\mathsf{ARVE}(\mathsf{f})$ is $\aleph_0$-categorical. We will use this result to characterize the $d$-finite tuples in $M$.

**6.3.1 Theorem.** *Let $\mathcal{M} \models \mathsf{ARVE}$ and $f = (f_1, \cdots, f_n) \in M^n$. Let $\mathcal{F}$ denote $L^1\big(\sigma(f), [0,1]\big)$, where $\sigma(f)$ is the $\sigma$-algebra generated by $f$. Then $f$ is a $d$-finite tuple in $M$ if and only if the $\sigma$-algebras $\sigma(f)$ and $\sigma\big(E(\mathcal{F})\big)$ are atomic.*

*Proof.* $\Longrightarrow$: If $f$ is $d$-finite, then $\mathsf{ARVE}(\mathsf{f})$ is $\omega$-categorical. By Corollary 1.9.5, $\mathsf{ARV}(\mathsf{f})$ is $\omega$-categorical. Then by Theorem 5.7.3, we have that $f_i$ is a discrete random variable for every $1 \leq i \leq n$. Therefore $\sigma(f)$ is atomic. By the fact that $\mathsf{ARVE}(\mathsf{f})$ is $\omega$-categorical, we know that for every $g \in \mathrm{dcl}^{\mathsf{ARVE}}(f)$, the theory $\mathsf{ARVE}(\mathsf{f},\mathsf{g})$ is also $\omega$-categorical. By Corollary 1.9.5 and Theorem 5.7.3, we have that $g$ is discrete. By Theorem 5.2.3($iv$), $\mathrm{dcl}^{\mathsf{ARV}}(E(\mathcal{F})) = L^1\big(\sigma(E(\mathcal{F})), [0,1]\big)$. Since $\sigma(f)$ is atomic, there is $h \in \mathrm{dcl}^{\mathsf{ARV}}(E(\mathcal{F})) = L^1\big(\sigma(E(\mathcal{F})), [0,1]\big)$ such that $\sigma(h) = \sigma\big(E(\mathcal{F})\big)$ up to a set of sets of measure zero. By Proposition 6.2.5, $\mathrm{dcl}^{\mathsf{ARV}}(E(\mathcal{F})) \subseteq \mathrm{dcl}^{\mathsf{ARVE}}(f)$. Hence $h$ is discrete, whereby $\sigma\big(E(\mathcal{F})\big) = \sigma(h)$ is atomic.

$\Longleftarrow$: Let $([0,1], \mathcal{L}, \lambda)$ be the standard Lebesgue space and let $([0,1]\times[0,1], \mathcal{L}\times\mathcal{L}, \lambda\times\lambda)$ be the product space. Define $\pi\colon [0,1]\times[0,1] \to [0,1]$ by $\pi(x,y) := x$ for all $(x,y) \in [0,1]\times[0,1]$. Let $\mathscr{B}$ denote $\mathcal{L}\times\mathcal{L}$. Then $\pi^{-1}(\mathcal{L})$ is atomless and $\mathscr{B}$ is atomless over $\pi^{-1}(\mathcal{L})$. Let $\mathscr{A} \subseteq \mathscr{B}$ be the smallest $\sigma$-algebra containing $\pi^{-1}(\mathcal{L})$ and the set of all sets of measure zero in $\mathscr{B}$. Then we also have $\mathscr{A}$ is atomless and $\mathscr{B}$ is atomless over $\mathscr{A}$. Hence, $\Big(L^1\big(([0,1]\times[0,1], \mathscr{B}, \lambda\times\lambda), [0,1]\big), \mathbb{E}(\cdot \mid \mathscr{A})\Big)$ is a model of $\mathsf{ARVE}$. Suppose $\mathcal{M}$ is a separable model. By Corollary 6.2.3, $\mathcal{M}$ is isomorphic to $\Big(L^1\big(([0,1]\times[0,1], \mathscr{B}, \lambda\times\lambda), [0,1]\big), \mathbb{E}(\cdot \mid \mathscr{A})\Big)$. Without loss of generality, we may identify $\mathcal{M}$ with $\Big(L^1\big(([0,1]\times[0,1], \mathscr{B}, \lambda\times\lambda), [0,1]\big), \mathbb{E}(\cdot \mid \mathscr{A})\Big)$.

For $g \in M^n$ with $(\mathcal{M}, f) \equiv (\mathcal{M}, g)$, we will show there exists $\psi \in \mathrm{Aut}_{\mathsf{ARVE}}(\mathcal{M})$ such that $\psi(f) = g$. Then $\mathsf{ARVE}(\mathsf{f})$ is separably categorical. By Theorem 3.5.2, $f$ is $d$-finite. Let $\mathcal{G}$ denote $L^1\big(\sigma(g), [0,1]\big)$. Note that $\mathcal{F}, \mathcal{G}$ are $L_{\mathsf{RV}(\times)}$-structures. For every $L_{\mathsf{RV}(\times)}$-term $t$, define $\phi_0(E(t(f))) = E(t(g))$. Since $\mathrm{tp}^{\mathsf{ARVE}}(f) = \mathrm{tp}^{\mathsf{ARVE}}(g)$, we have that $\phi_0$ is well-defined and is an $L_{\mathsf{RV}(\times)}$-elementary map in $\mathcal{M}$ from a subset of $E(\mathcal{F})$ into $E(\mathcal{G})$; *i.e.*, for all $L_{\mathsf{RV}(\times)}$-formulas $\tau$ and all $L_{\mathsf{RV}(\times)}$-terms $t_1, \cdots, t_k$, we have

$$\tau^{\mathcal{M}}\big(E(t_1(f)), \cdots, E(t_k(f))\big) = \tau^{\mathcal{M}}\big(E(t_1(g)), \cdots, E(t_k(g))\big) = \tau^{\mathcal{M}}\Big(\phi_0\big(E(t_1(f))\big), \cdots, \phi_0\big(E(t_k(f))\big)\Big).$$

Then by the fact that $\{E(t(f)) \mid t \text{ is an } L_{\mathsf{RV}(\times)}\text{-term}\}$ and $\{E(t(g)) \mid t \text{ is an } L_{\mathsf{RV}(\times)}\text{-term}\}$ are dense in $E(\mathcal{F})$ and $E(\mathcal{G})$ respectively, and $\mathrm{tp}^{\mathsf{ARVE}}(f) = \mathrm{tp}^{\mathsf{ARVE}}(g)$, we have that $\phi_0$ is extended to an $L_{\mathsf{RV}(\times)}$-elementary map $\phi\colon E(\mathcal{F}) \to E(\mathcal{G})$ in $\mathcal{M}$, which is distance preserving and bijective. Then because

$$L^1\big(\sigma(E(\mathcal{F})), [0,1]\big) = \overline{\{t(a_1, \cdots, a_k) \mid k \in \mathbb{N}, a_1, \cdots, a_k \in E(\mathcal{F}), t \text{ is an } L_{\mathsf{RV}(\times)}\text{-term}\}}$$

and

$$L^1\big(\sigma(E(\mathcal{G})),[0,1]\big)=\overline{\{t(b_1,\cdots,b_k)\mid k\in\mathbb{N},b_1,\cdots,b_k\in E(\mathcal{G}),t\text{ is an }L_{\mathsf{RV}(\times)}\text{-term}\}},$$

we have that $\phi$ is extended to an $L_{\mathsf{RV}(\times)}$-isomorphism $\tilde{\phi}\colon L^1\big(\sigma(E(\mathcal{F})),[0,1]\big)\to L^1\big(\sigma(E(\mathcal{G})),[0,1]\big)$. Then by the fact that $\sigma\big(E(\mathcal{F})\big)$ is atomic, we know that $\sigma\big(E(\mathcal{G})\big)$ is also atomic. By Fact 2.2.2, there is a maximal family $\{P_j\mid j\in J\}$ of pairwise inequivalent atoms of $\sigma\big(E(\mathcal{F})\big)$, where $J$ is a countable set. Since $\tilde{\phi}$ is an $L_{\mathsf{RV}(\times)}$-isomorphism, for each $j\in J$ we have that $\tilde{\phi}(\mathbf{1}_{P_j})=\mathbf{1}_{Q_j}$ for some $Q_j\in\sigma\big(E(\mathcal{G})\big)$, for each $j\in J$ we have $Q_j$ is an atom in $\sigma\big(E(\mathcal{G})\big)$, whenever $j\neq j'\in J$ we have $\mu(Q_j\cap Q_{j'})=0$, and $\mu(P_j)=\mu(Q_j)$ for each $j\in J$. Therefore $\sum_{j\in J}\mu(Q_j)=1$. After adding the complement of $\bigcup_{j\in J}Q_j$ to one of the atoms, we may assume that $\{Q_j\mid j\in J\}$ is a maximal family of pairwise inequivalent atoms of $\sigma\big(E(\mathcal{G})\big)$. By Lemma 3.1.5, there is a measure-preserving automorphism $\Phi\colon\mathscr{A}\to\mathscr{A}$ such that $\Phi(P_j)=Q_j$ up to a set of measure zero for each $j\in J$. By Theorem 5.1.1, $\Phi$ determines an $L_{\mathsf{RV}(\times)}$-automorphism $\varphi\colon L^1(\mathscr{A},[0,1])\to L^1(\mathscr{A},[0,1])$, which extends $\tilde{\phi}$.

Since $\mathrm{tp}^{\mathsf{ARVE}}(f)=\mathrm{tp}^{\mathsf{ARVE}}(g)$, we have $\mathrm{tp}^{\mathsf{ARV}}(f)=\mathrm{tp}^{\mathsf{ARV}}(g)$. By Theorem 5.2.3($v$), we have $\mathrm{dist}(f)=\mathrm{dist}(g)$. Therefore, the fact that $\sigma(f)$ is atomic implies that $\sigma(g)$ is atomic, and thus $f$ and $g$ are both discrete random variables. Moreover, by Fact 2.2.2, there is a maximal family $\{F_i\mid i\in I\}$ of pairwise inequivalent atoms of $\sigma(f)$, where $I$ is a countable set. Then for each $i\in I$, we have a sequence of $L_{\mathsf{RV}(\times)}$-terms $t_i^k$ such that $\mathbf{1}_{F_i}=\lim_{k\to\infty}t_i^k(f)$. Since $\mathrm{tp}^{\mathsf{ARV}(\times)}(f)=\mathrm{tp}^{\mathsf{ARV}(\times)}(g)$ and $\lim_{k\to\infty}t_i^k(f)^2=\lim_{k\to\infty}t_i^k(f)$ for each $i\in I$, we have that $\lim_{k\to\infty}t_i^k(g)^2=\lim_{k\to\infty}t_i^k(g)$, and thus $\lim_{k\to\infty}t_i^k(g)=\mathbf{1}_{G_i}$ for some $G_i\in\mathscr{B}$. Moreover, whenever $i\neq i'\in I$ we have $\mu(G_i\cap G_{i'})=0$, and for all $i\in I$ we have $\mu(F_i)=\mu(G_i)$ and $f(x)=g(y)$ for almost all $x\in F_i,y\in G_i$. Thus, $\sum_{i\in I}\mu(G_i)=\sum_{i\in I}\mu(F_i)=1$. Then by the fact that $\mathrm{dist}(f)=\mathrm{dist}(g)$, we have that $G_i$ is an atom in $\sigma(g)$ for each $i\in I$. After adding the complement of $\bigcup_{i\in I}G_i$ to one of the atoms, we may assume that $\{G_i\mid i\in I\}$ is a maximal family of pairwise inequivalent atoms of $\sigma(g)$. Further, for all $i\in I$, by the definition of $\phi$ we have

$$\phi(E(\mathbf{1}_{F_i}))=\phi\big(E(\lim_{k\to\infty}t_i^k(f))\big)=\lim_{k\to\infty}\phi(E(t_i^k(f)))=\lim_{k\to\infty}E(t_i^k(g))=E(\lim_{k\to\infty}t_i^k(g))=E(\mathbf{1}_{G_i}).$$

Hence, $\varphi(E(\mathbf{1}_{F_i}))=\phi(E(\mathbf{1}_{F_i}))=E(\mathbf{1}_{G_i})$ for all $i\in I$. For all $i\in I$ and all $A\in\mathscr{A}$, we have

$$\mu(A\cap F_i)=\int_A\mathbf{1}_{F_i}d\mu=\int_A E(\mathbf{1}_{F_i})d\mu=\int_{\Phi(A)}\varphi\big(E(\mathbf{1}_{F_i})\big)d\mu=\int_{\Phi(A)}E(\mathbf{1}_{G_i})d\mu=\mu(\Phi(A)\cap G_i).$$

Since $\mathscr{B}$ is atomless over $\mathscr{A}$, by Lemma 2.2.20 we may assume there is a measure-preserving homomorphism

$\Phi_i \colon \sigma(\mathscr{A} \cup \{F_i\}) \to \mathscr{B}$ which extends $\Phi$ such that $\Phi_i(F_i) = G_i$. By Proposition 2.2.3,

$$\mathscr{A} \vee \sigma(f) = \{\bigcup_{i \in I} A_i \cap F_i \mid A_i \in \mathscr{A} \text{ for each } i \in I\},$$

and

$$\mathscr{A} \vee \sigma(g) = \{\bigcup_{i \in I} A_i \cap G_i \mid A_i \in \mathscr{A} \text{ for each } i \in I\}.$$

We define $\Phi' \colon \mathscr{A} \vee \sigma(f) \to \mathscr{A} \vee \sigma(g)$ by $\Phi'(\bigcup_{i \in I} A_i \cap F_i) := \bigcup_{i \in I} \Phi_i(A_i \cap F_i) = \bigcup_{i \in I} \Phi(A_i) \cap G_i$, where $A_i \in \mathscr{A}$ for every $i \in I$. Because $\Phi_i$ is a measure-preserving homomorphism for each $i \in I$, we have that $\Phi'$ is a measure-preserving homomorphism from $\mathscr{A} \vee \sigma(f)$ to $\mathscr{A} \vee \sigma(g)$. Since $\Phi$ is surjective, we know that $\Phi'$ is also surjective. Then $\Phi'$ is a measure-preserving isomorphism that extends $\Phi$. Since $\sigma(f)$ is atomic, by Proposition 2.2.24 there is a measure-preserving automorphism $\tilde{\Phi} \colon \mathscr{B} \to \mathscr{B}$ extending $\Phi'$. Therefore, by Lemma 6.2.1 we know that $\tilde{\Phi}$ induces an $L_{\mathsf{RVE}}$-automorphism $\tilde{\varphi} \colon M \to M$ extending $\varphi$. Then because for all $i \in I$ we have $\tilde{\Phi}(F_i) = G_i$ and $f(x) = g(y)$ for almost all $x \in F_i, y \in G_i$, we have that $\tilde{\varphi}(f) = g$. $\square$

# Chapter 7

# Model theory of adapted structures with $T$-indexed conditional expectations

Let $T$ be a linear ordering throughout this chapter. An *adapted structure with a $T$-indexed family of conditional expectations* is of the form $(L^1((\Omega, \mathscr{B}, \mu), [0,1]), \big(\mathbb{E}(\cdot \mid \mathscr{A}_t)\big)_{t\in T})$, where $(\Omega, \mathscr{B}, \mu)$ is a probability space, $(\mathscr{A}_t)_{t\in T}$ is an increasing family of $\sigma$-subalgebras of $\mathscr{B}$, and $\mathbb{E}(\cdot \mid \mathscr{A}_t)$ is the conditional expectation with respect to $\mathscr{A}_t$ for each $t \in T$. It is called a *totally atomless adapted structure with a $T$-indexed family of conditional expectations* if for each $t \in T$, the $\sigma$-algebra $\mathscr{A}_t$ is atomless over previous ones (if there is a minimal element 0 in $T$, then $\mathscr{A}_0$ is atomless), and $\mathscr{B}$ is atomless over $\mathscr{A}_t$. It follows from the definition of atomlessness that $\mathscr{B}$ and each $\mathscr{A}_t$ are atomless. When $T$ only contains one element, a (totally atomless) adapted structure with a $T$-indexed family of conditional expectations is called *a (totally atomless) adapted structure with one conditional expectation*.

In this chapter we study the theories of (totally atomless) adapted structures with a $T$-indexed family of conditional expectations, whereas Chapter 6 studies the theories of (totally atomless) adapted structures with one conditional expectation. Chapter 7 is parallel to Chapter 6. In Section 7.1, we axiomatize the theory of adapted structures with a $T$-indexed family of conditional expectations by $\mathsf{RVE_T}$, and the theory of totally atomless adapted structures with a $T$-indexed family of conditional expectations by $\mathsf{ARVE_T}$. In Section 7.2, we consider basic model theoretic properties of $\mathsf{ARVE_T}$. For instance, we prove that $\mathsf{ARVE_T}$ is separably categorical if and only if $T$ is a finite linear ordering and show that $\mathsf{ARVE_T}$ has continuum many non-isomorphic separable models when $T$ is infinite. In Section 7.3, we characterize the $d$-finite tuples in models of $\mathsf{ARVE_T}$ when $T$ is finite.

## 7.1 Axioms for $\mathsf{RVE_T}$ and $\mathsf{ARVE_T}$

Fix a linear ordering $(T, <)$. Let $L_{\mathsf{RVE_T}}$ denote $L_{\mathsf{RV}(\times)} \cup \{E_t : t \in T\}$, where every $E_t$ is a 1-Lipschitz unary function symbol. Let $\mathsf{RVE_t}$ denote the theory of adapted structure with one conditional expectation where the conditional expectation symbol is $E_t$. The theory $\mathsf{RVE_T}$ is the theory consisting of $\bigcup_{t\in T} \mathsf{RVE_t}$ and the following closed conditions:

$(\mathrm{RVE}_{\mathsf{s<t}})$ $\sup_x \max\big(d(E_s[E_t(x)], E_t[E_s(x)]), d(E_t[E_s(x)], E_s(x))\big) = 0$ for all $s <\in T$.

**7.1.1 Proposition.** *Let $\mathcal{M}$ be an $L_{\mathsf{RVE_T}}$-structure. Then $\mathcal{M}$ is a model of $\mathsf{RVE_T}$ if and only if $\mathcal{M}$ is isomorphic to an adapted structure with a $T$-indexed family of conditional expectations.*

*Proof.* $\Longleftarrow$: By Proposition 6.1.2 and the transitivity property of conditional expectations, this is clear.

$\Longrightarrow$: Since $\mathcal{M} \models \mathsf{RV}(\times)$, we may assume that $M$ is $L^1((\Omega, \mathscr{B}, \mu), [0,1])$ for the probability space $(\Omega, \mathscr{B}, \mu)$. Let $N_t$ denote $E_t(M)$ for every $t \in T$. By $(\mathsf{RVE_{s<t}})$, we have $N_s \subseteq N_t$ for all $s < t \in T$. Since $\mathcal{M} \models \mathsf{RVE_t}$ for every $t \in T$, as shown in the proof of Proposition 6.1.2, there is a $\sigma$-subalgebra $\mathscr{A}_t \subseteq \mathscr{B}$ such that $E_t$ is $\mathbb{E}_t(\cdot \mid \mathscr{A}_t)$ for every $t \in T$. Then because $N_s \subseteq N_t$ for all $s < t \in T$, we know that every $\mathscr{A}_s$-measurable function is also $\mathscr{A}_t$-measurable. Hence we have $\mathscr{A}_s \subseteq \mathscr{A}_t$ whenever $s < t$. Therefore, $\mathcal{M}$ is isomorphic to the adapted structure $(L^1((\Omega, \mathscr{B}, \mu), [0,1]), \big(\mathbb{E}(\cdot \mid \mathscr{A}_t)\big)_{t\in T})$. □

The theory $\mathsf{ARVE_T}$ consists of the theory $\mathsf{RVE_T}$ together with the following conditions:

$(\mathrm{ARVE}_{\mathsf{s<t}})$ $\sup_x \inf_y \max\Big(I\big(E_t(y) \wedge (\neg E_t(y))\big), d(E_s(E_t(y) \wedge E_t(x)), \frac{E_s(x)}{2})\Big) = 0$ for all $s < t \in T$;

$(\mathrm{ARVE}_{\mathsf{t}})$ $\sup_x \inf_y \max\Big(I(y \wedge (\neg y)), d\big(E_t(y \wedge x), \frac{E_t(x)}{2}\big)\Big) = 0$ for all $t \in T$;

if $(T, <)$ has a minimal element 0, then we add:

$(\mathrm{ARVE}_0)$ $\sup_x \inf_y \max\Big(I\big(E_0(y) \wedge (\neg E_0(y))\big), d\big(I(E_0(y) \wedge E_0(x)), \frac{I(x)}{2}\big)\Big) = 0.$

Let $\mathcal{M}$ be an adapted structure with a $T$-indexed family of conditional expectations of the form

$$(L^1((\Omega, \mathscr{B}, \mu), [0,1]), \big(\mathbb{E}(\cdot \mid \mathscr{A}_t)\big)_{t\in T}),$$

where $(\Omega, \mathscr{B}, \mu)$ is a probability space and $(\mathscr{A}_t)_{t\in T}$ is an increasing family of $\sigma$-subalgebras of $\mathscr{B}$. By Proposition 7.1.1, we have that $\mathcal{M}$ is a model of $\mathsf{RVE_T}$. For all $s, t \in T$ with $s < t$, Axiom $(\mathrm{ARVE}_{\mathsf{s<t}})$ holds in $\mathcal{M}$ if and only if $\mathscr{A}_t$ is atomless over $\mathscr{A}_s$. For all $t \in T$, Axiom $(\mathrm{ARVE}_{\mathsf{t}})$ holds in $\mathcal{M}$ if and only if $\mathscr{B}$ is atomless over $\mathscr{A}_t$. When $T$ has a minimal element 0, Axiom $(\mathrm{ARVE}_0)$ holds in $\mathcal{M}$ if and only if $\mathscr{A}_0$ is atomless.

**7.1.2 Theorem.** *Let $\mathcal{M}$ be an $L_{\mathsf{RVE_T}}$-structure. Then $\mathcal{M}$ is a model of $\mathsf{ARVE_T}$ if and only if $\mathcal{M}$ is isomorphic to a totally atomless adapted structure with a $T$-indexed family of conditional expectations.*

*Proof.* $\Longleftarrow$: Suppose that $\mathcal{M}$ is of the form $\big(L^1((\Omega, \mathscr{B}, \mu), [0,1]), \big(\mathbb{E}(\cdot \mid \mathscr{A}_t)\big)_{t\in T}\big)$, where $(\Omega, \mathscr{B}, \mu)$ is a probability space, $(\mathscr{A}_t)_{t\in T}$ is an increasing family of $\sigma$-subalgebras of $\mathscr{B}$, for each $t \in T$ the $\sigma$-algebra $\mathscr{A}_t$ is atomless over all previous ones (if $T$ has a minimal element 0 then $\mathscr{A}_0$ is atomless), and $\mathscr{B}$ is atomless over

$\mathscr{A}_t$ for each $t$. By Proposition 7.1.1, we have that $\mathcal{M}$ is a model of $\mathsf{RVE_T}$. By Lemma 2.2.16 (Maharam's Lemma), the axioms $(\mathrm{ARVE_{s<t}})$, $(\mathrm{ARVE_t})$, and $(\mathrm{ARVE_0})$ follow. Hence $\mathcal{M}$ is a model of $\mathsf{ARVE_T}$.

$\Longrightarrow$: Let $\mathcal{M} \models \mathsf{ARVE_T}$. Then $\mathcal{M}$ is also a model of $\mathsf{RVE_T}$. By Proposition 7.1.1, $\mathcal{M}$ is $L_{\mathsf{RVE_T}}$-isomorphic to $\big(L^1((\Omega,\mathscr{B},\mu),[0,1]),\big(\mathbb{E}(\cdot \mid \mathscr{A}_t)\big)_{t\in T}\big)$, where $(\Omega,\mathscr{B},\mu)$ is a probability space and $(\mathscr{A}_t)_{t\in T}$ is an increasing family of $\sigma$-subalgebras of $\mathscr{B}$. For all $s < t \in T$, by $(\mathrm{ARVE_{s<t}})$ we see that $\big(L^1(\mathscr{A}_t,[0,1]),\mathbb{E}(\cdot \mid \mathscr{A}_s)\big) \models \mathsf{ARVE}$. Hence $\mathscr{A}_t$ is atomless over $\mathscr{A}_s$. For all $t \in T$, by $(\mathrm{ARVE_t})$ we see that $\big(L^1(\mathscr{B},[0,1]),\mathbb{E}(\cdot \mid \mathscr{A}_t)\big) \models \mathsf{ARVE}$, whereby $\mathscr{B}$ is atomless over $\mathscr{A}_t$. Also when $T$ has a minimal element 0, by $(\mathrm{ARVE_0})$ we see that $L^1(\mathscr{A}_0,[0,1]) \models \mathsf{ARV}$, whereby $\mathscr{A}_0$ is atomless. □

**7.1.3 Remark.** *If $T'$ is a subset of $T$, then $\mathsf{ARVE_T} \models \mathsf{ARVE_{T'}}$. (If $T'$ has a minimum element 0 that is* not *the minimum element of $T$, then $\mathsf{ARVE_{T'}}$ contains the axiom* $(\mathrm{ARVE_0})$ *which is* not *explicitly an axiom of $\mathsf{ARVE_T}$. However, it follows from Theorem 7.1.2 that $\mathsf{ARVE_T} \models (\mathrm{ARVE_0})$, and hence $\mathsf{ARVE_T} \models \mathsf{ARVE_{T'}}$. Otherwise, $\mathsf{ARVE_{T'}}$ is a subset of $\mathsf{ARVE_T}$.)*

## 7.2 Basics of $\mathsf{ARVE_T}$

In this section, we prove basic model theoretic properties of $\mathsf{ARVE_T}$. We prove that $\mathsf{ARVE_T}$ is separably categorical if and only if $T$ is finite and $\mathsf{ARVE_T}$ admits quantifier elimination for all $T$. Also we investigate the type spaces of $\mathsf{ARVE_T}$.

First, we prove the following lemma.

**7.2.1 Lemma.** *Let $(\Omega,\mathscr{B},\mu)$ be a probability space and let $(\mathscr{A}_t)_{t\in T}$ be an increasing family of $\sigma$-subalgebras of $\mathscr{B}$. Let $\mathcal{M}$ be a model of $\mathsf{RVE_T}$ of the form $\big(L^1((\Omega,\mathscr{B},\mu),[0,1]),(\mathbb{E}(\cdot \mid \mathscr{A}_t))_{t\in T}\big)$. Suppose that $\Phi$ is an $L_{\mathsf{RV}}$-automorphism of $L^1\big((\Omega,\mathscr{B},\mu),[0,1]\big)$ such that the restriction of $\Phi$ to $L^1\big((\Omega,\mathscr{A}_t,\mu),[0,1]\big)$ is an $L_{\mathsf{RV}}$-automorphism of $L^1\big((\Omega,\mathscr{A}_t,\mu),[0,1]\big)$ for each $t \in T$. Then $\Phi$ is an $L_{\mathsf{RVE_T}}$-automorphism of $\mathcal{M}$.*

*Proof.* By Lemma 6.2.1, we have that $\Phi$ is an $L_{\mathsf{RVE_t}}$-automorphism of $\mathcal{M}$ for each $t \in T$. Then it follows that $\Phi$ is an $L_{\mathsf{RVE_T}}$-automorphism of $\mathcal{M}$. □

Consider a probability algebra $(\mathscr{B},\mu)$ and an increasing $T$-indexed family of $\sigma$-complete subalgebras $(\mathscr{A}_t)_{t\in T}$ of $\mathscr{B}$. Suppose $(\mathscr{B},\mu)$ is the probability algebra associated to the probability space $(\Omega,\mathcal{G},\mu)$. As explained on page 30, there is a mapping $\pi$ from the $\sigma$-algebra $\mathcal{G}$ to the $\sigma$-complete boolean algebra $\pi(\mathcal{G}) := \widehat{\mathcal{G}} = \mathscr{B}$. For all $t \in T$, let $\mathcal{F}_t$ denote $\pi^{-1}(\mathscr{A}_t)$. Then because $\mathscr{A}_t$ is a $\sigma$-complete boolean algebra and $\mathcal{G}$ is a $\sigma$-algebra, it is routine to check that $\mathcal{F}_t$ is a $\sigma$-subalgebra of $\mathcal{G}$. Further, we know that $(\mathscr{A}_t,\mu)$ is the probability algebra associated to $(\Omega,\mathcal{F}_t,\mu)$ for all $t \in T$. The conditional expectation $\mathbb{E}(\cdot \mid \mathcal{F}_t)$ with

respect to $\mathcal{F}_t$ is defined from $L^1\big((\Omega,\mathcal{G},\mu),[0,1]\big)$ to $L^1\big((\Omega,\mathcal{F}_t,\mu),[0,1]\big)$ for all $t\in T$. Therefore, $\mathbb{E}(\cdot\mid\mathcal{F}_t)$ is defined from $L^1(\mathscr{B},\mu):=L^1\big((\Omega,\mathcal{G},\mu),[0,1]\big)$ to $L^1(\mathscr{A}_t,\mu):=L^1\big((\Omega,\mathcal{F}_t,\mu),[0,1]\big)$, denoted by $\mathbb{E}(\cdot\mid\mathscr{A}_t)$, for all $t\in T$.

By Theorem 5.1.1, we see that a probability algebra $(\mathscr{B},\mu)$ and an increasing $T$-indexed family of $\sigma$-complete subalgebras $(\mathscr{A}_t)_{t\in T}$ of $\mathscr{B}$ determine an adapted structure with a $T$-indexed family of conditional expectations, denoted by $\big(L^1(\mathscr{B},[0,1]),(\mathbb{E}(\cdot\mid\mathscr{A}_t))_{t\in T}\big)$, up to $L_{\mathsf{RVE_T}}$-isomorphism.

## Separable categoricity

**7.2.2 Theorem.** *If $T$ is a finite linear ordering, then $\mathsf{ARVE_T}$ has a unique separable model. Hence $\mathsf{ARVE_T}$ is complete for all $T$.*

*Proof.* By induction on $|T|$, we show that $\mathsf{ARVE_T}$ is separably categorical. When $|T|=1$, this is Corollary 6.2.3. Suppose the conclusion holds for $|T|=k-1$. Consider $|T|=k$. Let 0 be the minimal element in $T$. For each $i=1,2$, let $\mathcal{M}_i$ be a separable model of $\mathsf{ARVE_T}$ of the form $\big(L^1((\Omega^i,\mathscr{B}^i,\mu^i),[0,1]),\big(\mathbb{E}(\cdot\mid\mathscr{A}^i_t)\big)_{t\in T}\big)$, where $(\Omega^i,\mathscr{B}^i,\mu^i)$ is a probability space, $(\mathscr{A}^i_t)_{t\in T}$ is an increasing family of $\sigma$-subalgebras of $\mathscr{B}^i$, for each $t\in T$ the $\sigma$-algebra $\mathscr{A}^i_t$ is atomless over all previous ones, $\mathscr{A}^i_0$ is atomless, and $\mathscr{B}^i$ is atomless over $\mathscr{A}^i_t$ for each $t$. Suppose $S$ is the maximal element in $T$. Let $T'$ denote the set $\{t\in T\mid t<S\}$. Then $|T'|=k-1$. By induction, there is an $L_{\mathsf{RVE_{T'}}}$-isomorphism $\varphi$ between $\big(L^1((\Omega^1,\mathscr{A}^1_S,\mu^1),[0,1]),\big(\mathbb{E}(\cdot\mid\mathscr{A}^1_t)\big)_{t\in T'}\big)$ and $\big(L^1((\Omega^2,\mathscr{A}^2_S,\mu^2),[0,1]),\big(\mathbb{E}(\cdot\mid\mathscr{A}^2_t)\big)_{t\in T'}\big)$. By Proposition 6.2.2, there is an $L_{\mathsf{RVE}}$-isomorphism $\Phi$ between $\big(L^1((\Omega^1,\mathscr{B}^1,\mu^1),[0,1]),\big(\mathbb{E}(\cdot\mid\mathscr{A}^1_S)\big)$ and $\big(L^1((\Omega^2,\mathscr{B}^2,\mu^2),[0,1]),\big(\mathbb{E}(\cdot\mid\mathscr{A}^2_S)\big)$ extending $\varphi$. Then because $\Phi$ extends $\varphi$, we know that for each $t\in T$ the restriction of $\Phi$ to $L^1((\Omega^1,\mathscr{A}^1_t,\mu^1),[0,1])$ is an $L_{\mathsf{RV}}$-isomorphism between $L^1((\Omega^1,\mathscr{A}^1_t,\mu^1),[0,1])$ and $L^1((\Omega^2,\mathscr{A}^2_t,\mu^2),[0,1])$. Then by Lemma 7.2.1, there is an $L_{\mathsf{RVE_T}}$-isomorphism $\Phi$ between $\mathcal{M}_1$ and $\mathcal{M}_2$. Hence $\mathsf{ARVE_T}$ is separably categorical, whereby it is a complete theory.

Consider any (possibly infinite) linear ordering $T$. Let $\mathcal{M}\models\mathsf{ARVE_T}$. For every closed $L_{\mathsf{ARVE_T}}$-condition $P$ in $Th(\mathcal{M})$, there is a finite subset $T'$ of $T$ such that $P$ is an $L_{\mathsf{ARVE_{T'}}}$-condition. As shown above, $\mathsf{ARVE_{T'}}$ is complete and $\mathcal{M}\models\mathsf{ARVE_{T'}}$, so $\mathsf{ARVE_{T'}}\models P$. Since $\mathsf{ARVE_T}\models\mathsf{ARVE_{T'}}$, we have $\mathsf{ARVE_T}\models P$, whereby $\mathsf{ARVE_T}\models Th(\mathcal{M})$. Hence, $\mathsf{ARVE_T}$ is complete. □

When $T$ is infinite, $\mathsf{ARVE_T}$ is no longer separably categorical.

**7.2.3 Theorem.** *Let $T$ be a countable infinite linear ordering. Then $\mathsf{ARVE_T}$ has continuum many non-isomorphic separable models.*

*Proof.* An infinite linear ordering $T$ contains either an infinite strictly increasing sequence or an infinite strictly decreasing sequence. In the following two cases, we define $\overline{T} = T \cup \{T_0\}$ for some $T_0$ such that $\overline{T} = T_- \cup \{T_0\} \cup T_+$, where for all $s \in T_-$ and all $t \in T_+$, we have $s < T_0 < t$. In Case 1a, $T_0 \in T$, but in Case 1b and Case 2, $T_0 \notin T$ is a new element.

**Case 1**: $T$ contains an infinite strictly increasing sequence $(t_i)_{i\in\mathbb{N}}$. Let $T_-$ be the subset of $T$ containing all $t \in T$ such that $t \leq t_i$ for some $i \in \mathbb{N}$. **Case 1a**: if $T\backslash T_-$ has a minimum element, then we call it $T_0$. Let $T_+$ be $T\backslash(\{T_0\} \cup T_-)$ and let $\overline{T} = T$. **Case 1b**: if $T\backslash T_-$ has no minimum element, then let $T_+ = T\backslash T_-$. We insert a new element $T_0$ into $T$ such that for all $t \in T_-$ and all $s \in T_+$, one has $t < T_0 < s$. Let $\overline{T} = T \cup \{T_0\}$. Then $\overline{T}$ is also a linear ordering. Note that $T_+$ may be the empty set.

**Case 2**: $T$ contains an infinite strictly decreasing sequence $(t_i)_{i\in\mathbb{N}}$ so that $t_i < t_j$ whenever $i > j$. Let $T_+$ be the subset of $T$ that contains all $t \in T$ such that $t \geq t_i$ for all $i \in \mathbb{N}$. Let $T_- = T\backslash T_+$. We insert a new element $T_0$ into $T$ such that for all $t \in T_-$ and all $s \in T_+$, one has $t < T_0 < s$. Let $\overline{T} = T \cup \{T_0\}$. Then $\overline{T}$ is also a linear ordering. Note that $T_-$ may be the empty set.

Now, we will construct adapted structures with a $\overline{T}$-indexed family of conditional expectations and then consider their reducts to $L_{\mathsf{RVE_T}}$. For all $r \in (0, \frac{1}{2})$, let $(\Gamma_r, G_r, \mu_r)$ denote the probability space $(\{0,1\}, \mathcal{P}(\{0,1\}), \mu_r)$ where $\mu_r(\{0\}) = r$ and $\mu_r(\{1\}) = 1 - r$. For any linear ordering $S$, let $([0,1]^S, \mathcal{L}_S, \lambda_S)$ denote the product of $S$ copies of the standard Lebesgue space $([0,1], \mathcal{L}, \lambda)$. Let $(\Omega_r, \mathcal{F}^r, \nu_r)$ denote the product of $([0,1]^{T_-}, \mathcal{L}_{T_-}, \lambda_{T_-})$, $(\Gamma_r, G_r, \mu_r)$ and $([0,1]^{T_+}, \mathcal{L}_{T_+}, \lambda_{T_+})$, where $\Omega_r = [0,1]^{T_-} \times \Gamma_r \times [0,1]^{T_+}$, $\mathcal{F}^r = \mathcal{L}_{T_-} \times G_r \times \mathcal{L}_{T_+}$, and $\nu_r = \lambda_{T_-} \times \mu_r \times \lambda_{T_+}$. For an element in $\Omega_r$, we write $(\omega(t))_{t\in\overline{T}}$, where $\omega(t) \in [0,1]$ when $t \neq T_0$, and $\omega(T_0) \in G_r$. For every $t \in \overline{T}$, let $\mathcal{F}^r_t$ be the filtration such that $(\omega(t))_{t\in\overline{T}}, (\omega'(t))_{t\in\overline{T}} \in \Omega_r$ belong to the same sets in $\mathcal{F}^r_t$ if and only if $\omega(s) = \omega'(s)$ whenever $s \leq t$. Then we have that $\mathcal{F}^r_s \subseteq \mathcal{F}^r_t$ whenever $s \leq t$. Let $\mathcal{M}_r = \big(L^1(\nu_r, [0,1]), (\mathbb{E}(\cdot \mid \mathcal{F}^r_t))_{t\in T}\big)$. Then $\mathcal{M}_r \models \mathsf{ARVE_T}$. Since $T$ is countable, we know that $\mathcal{M}_r$ is separable.

Let $\mathcal{F}^r_-$ denote the smallest $\sigma$-subalgebra containing $\mathcal{F}^r_t$ for all $t \in T_-$ and let $\mathcal{F}^r_+$ denote $\bigcap_{t\in T_+} \mathcal{F}^r_t$. For $r \neq r' \in (0, \frac{1}{2})$, if $\mathcal{M}_r$ is $L_{\mathsf{RVE_T}}$-isomorphic to $\mathcal{M}_{r'}$, then $\mathcal{F}^r_t \cong \mathcal{F}^{r'}_t$ for all $t \in T$, whereby $\mathcal{F}^r_- \cong \mathcal{F}^{r'}_-$ and $\mathcal{F}^r_+ \cong \mathcal{F}^{r'}_+$. In Case 1, $\mathcal{F}^r_{T_0} = \bigcap_{t\in T\backslash T_-} \mathcal{F}^r_t$, and in Case 2, $\mathcal{F}^r_{T_0} = \mathcal{F}^r_+$ for all $r \in (0, \frac{1}{2})$. Hence, in both cases, we have $\mathcal{F}^r_{T_0} \cong \mathcal{F}^{r'}_{T_0}$. Note that $\mathcal{F}^r_{T_0}$ is not atomless over $\mathcal{F}^r_-$ and the largest atom in $\mathcal{F}^r_{T_0}$ over $\mathcal{F}^r_-$ is of the measure $1 - r$ for every $r \in (0, \frac{1}{2})$. Because of this fact, $\mathcal{M}_r \ncong \mathcal{M}_{r'}$ whenever $r \neq r'$. Therefore $\mathsf{ARVE_T}$ has continuum many non-isomorphic separable models. □

**7.2.4 Proposition.** *The theory* $\mathsf{ARVE_T}$ *is model complete.*

*Proof.* From the form of its axioms, $\mathsf{ARVE_T}$ is an $\forall\exists$-theory. When $T$ is finite, by Theorem 7.2.2 ($\mathsf{ARVE_T}$

is separably categorical, when $T$ is finite) and Theorem 1.8.2 (Lindström's test for model completeness) we have that $\mathsf{ARVE}_\mathsf{T}$ is model complete.

Now we consider the case in which $T$ is infinite. For any finite $T' \subseteq T$, we just proved that $\mathsf{ARVE}_{\mathsf{T}'}$ is model complete. Let $\mathcal{M} \subseteq \mathcal{N}$ be models of $\mathsf{ARVE}_\mathsf{T}$. For any finite $T' \subseteq T$, since $\mathsf{ARVE}_\mathsf{T} \models \mathsf{ARVE}_{\mathsf{T}'}$ we have that $\mathcal{M}$ and $\mathcal{N}$ are also models of $\mathsf{ARVE}_{\mathsf{T}'}$. For any $L_{\mathsf{RVE}_\mathsf{T}}$-formula $\varphi(x_1, \cdots, x_n)$, only finitely many $t \in T$ are involved in $\varphi$. Let $T(\varphi)$ denote the set of those $t$'s. Then $\varphi$ is an $L_{\mathsf{RVE}_{\mathsf{T}(\varphi)}}$-formula. Since $\mathsf{ARVE}_{\mathsf{T}(\varphi)}$ is model complete, for all $a_1, \cdots, a_n \in M$, we have $\varphi^{\mathcal{M}}(a_1, \cdots, a_n) = \varphi^{\mathcal{N}}(a_1, \cdots, a_n)$. Hence, $\mathcal{M} \preceq \mathcal{N}$, whereby $\mathsf{ARVE}_\mathsf{T}$ is model complete. □

## Type spaces

In [18], Hoover introduced the notion of adapted function (AF) and based on it, he defined adapted distributions.

**7.2.5 Definition** ([13]). • *Let $x = (x_t)_{t\in T}$ be a stochastic process defined on an adapted space $\Omega = (\Omega, \{\mathcal{F}_t\}_{t\in T}, \mu)$ (i.e., $x_t$ is $\mathcal{F}_t$-measurable for each $t \in T$). The set **AF of adapted functions** (resp. AF([0,1]) of [0,1]-adapted functions) is the family of expressions obtained by applying the following rules (i)-(iii) finitely many times. The value of an adapted function (resp. $[0,1]$-adapted function) at $(x, \Omega)$ is an $\mathbb{R}$-valued (resp. $[0,1]$-valued) random variable on $\Omega$, which is also given by those rules.*

*(i) (Basis Step) If $\phi \in C_b(\mathbb{R}^n, \mathbb{R})$ (resp. $C(\mathbb{R}^n, [0,1])$) and $\vec{t} = (t_1, \cdots, t_n) \in T^n$, then the expression $\phi_{\vec{t}}$ is in AF (resp. AF([0,1])). Its value at $x$ is the random variable*

$$\phi_{\vec{t}}(x) : \Omega \to \mathbb{R},$$

*defined by $\phi_{\vec{t}}(x)(\omega) = \phi(x_{t_1}(\omega), \cdots, x_{t_n}(\omega))$.*

*(ii) (Composition Step) If $\psi \in C_b(\mathbb{R}^n, \mathbb{R})$ (resp. $C(\mathbb{R}^n, [0,1])$) and $f_1, \cdots, f_n \in AF$ (resp. AF([0,1])), then the expression $\psi(f_1, \cdots, f_n)$ is in AF (resp. AF([0,1])). If the value of $f_i$ at $x$ is $f_i(x)$, then the value of $\psi(f_1, \cdots, f_n)$ at $x$ is the random variable*

$$\psi(f_1, \cdots, f_n)(x) : \Omega \to \mathbb{R},$$

*defined by $\psi(f_1, \cdots, f_n)(x)(\omega) = \psi(f_1(x)(\omega), \cdots, f_n(x)(\omega))$.*

*(iii) (Conditional Expectation Step) If $f$ is in AF (resp. AF([0,1])) and $t \in T$, then the expression*

$\mathbb{E}(f \mid t)$ *is in AF (resp. AF([0,1])). Its value at* $x$ *is the random variable*

$$\mathbb{E}(f \mid t)(x) : \Omega \to \mathbb{R},$$

*defined by* $\mathbb{E}(f \mid t)(x)(\omega) = \mathbb{E}(f(x) \mid \mathcal{F}_t)(\omega)$.

- *Let* $x, y$ *be stochastic processes on adapted spaces* $(\Omega, \mu), (\Gamma, \nu)$ *respectively, then we say* $x$ *and* $y$ *have the* ***same adapted distribution*** *if for all* $f$ *in AF we have* $\int_\Omega f(x)d\mu = \int_\Gamma f(y)d\nu$.

**7.2.6 Remark.** *By scaling all the bounded functions appearing in the Basic Step and Composition Step of* $f$ *in AF, we see that to say* $x$ *and* $y$ *have the same adapted distribution, it suffices to consider all* $f$ *in AF([0,1]).*

**7.2.7 Proposition.** *Let* $\mathcal{M} \models \mathsf{RVE_T}$ *be of the form* $(L^1((\Omega, \mathscr{B}, \mu), [0,1]), \big(\mathbb{E}(\cdot \mid \mathscr{A}_t)\big)_{t\in T})$, *where* $(\Omega, \mathscr{B}, \mu)$ *is a probability space and* $(\mathscr{A}_t)_{t\in T}$ *is an increasing family of* $\sigma$*-subalgebras of* $\mathscr{B}$. *Let* $x = (x_t)_{t\in T}$ *and* $y = (y_t)_{t\in T}$ *be stochastic processes, where* $x_t$ *and* $y_t$ *are in* $L^1(\mathscr{A}_t, [0,1])$ *for each* $t \in T$. *Then:*

$$\mathrm{tp}^{\mathrm{qf}}(x) = \mathrm{tp}^{\mathrm{qf}}(y) \quad \Longleftrightarrow \quad x \text{ and } y \text{ have the same adapted distribution.}$$

*Proof.* For all $n \in \mathbb{N}$, all $\vec{t} = (t_1, \cdots, t_n) \in T^n$ and all $L_{\mathsf{RVE_T}}$-terms $\varphi$ with $n$ variables, define $\varphi_{\vec{t}}((a_t)_{t\in T}) := \varphi(a_{t_1}, \cdots, a_{t_n})$. We call $\varphi_{\vec{t}}$ a $T$-$L_{\mathsf{RVE_T}}$-*term.*

In $L_{\mathsf{RVE_T}}$, there is only one predicate symbol $I$. Thus, $\mathrm{tp}^{\mathrm{qf}}(x) = \mathrm{tp}^{\mathrm{qf}}(y)$ if and only if for all $T$-$L_{\mathsf{RVE_T}}$-terms $\varphi_{\vec{t}}$ where $\vec{t} \in T^n$, one has $I\big(\varphi_{\vec{t}}(x)\big) = I\big(\varphi_{\vec{t}}(y)\big)$.

$\Longleftarrow$: For every $(t_1, \cdots, t_n) \in T^n$ and every $L_{\mathsf{RVE_T}}$-term $\varphi$, there is $f$ in AF such that $\varphi_{\vec{t}}(x) = f(x)$. That is to say, every $T$-$L_{\mathsf{RVE}}$-term is in AF.

$\Longrightarrow$: From the definition of AF([0,1]), we see that in each step, every expression in AF([0,1]) is in the uniform limit of the set of all $T$-$L_{\mathsf{RVE_T}}$-terms. Hence, if $I\big(\varphi_{\vec{t}}(x)\big) = I\big(\varphi_{\vec{t}}(y)\big)$ for all $T$-$L_{\mathsf{RVE_T}}$-terms $\varphi_{\vec{t}}$, we have $I(f(x)) = I(f(y))$ for all $f$ in AF([0,1]). □

Next, we show how types in $\mathsf{ARVE_T}$ are determined. Our treatment is analogous to what we did in Section 6.2 for types in $\mathsf{ARVE}$. Our main results are Theorems 7.2.9 and 7.2.11.

**7.2.8 Lemma.** *Let* $\mathcal{M} \models \mathsf{ARVE_T}$ *be of the form* $(L^1((\Omega, \mathscr{B}, \mu), [0,1]), \big(\mathbb{E}(\cdot \mid \mathscr{A}_t)\big)_{t\in T})$, *where* $(\Omega, \mathscr{B}, \mu)$ *is a probability space,* $(\mathscr{A}_t)_{t\in T}$ *is an increasing family of* $\sigma$*-subalgebras of* $\mathscr{B}$, *each algebra* $\mathscr{A}_t$ *is atomless over all previous ones (if* $T$ *has a minimal element 0, then* $\mathscr{A}_0$ *is atomless), and* $\mathscr{B}$ *is atomless over* $\mathscr{A}_t$ *for each* $t$. *Let* $\mathscr{A}_\infty$ *denote the smallest* $\sigma$*-subalgebra of* $\mathscr{B}$ *that contains* $\mathscr{A}_t$ *for all* $t \in T$. *Suppose that* $\mathscr{A}_\infty$ *is the*

*$\sigma$-algebra generated by the random variables in $A_\infty \subseteq M$. For all $a, b \in M^n$ and $C \subseteq M$, we have:*

$$\mathrm{tp}^{\mathsf{ARV}}(a/C \cup A_\infty) = \mathrm{tp}^{\mathsf{ARV}}(b/C \cup A_\infty) \Longrightarrow \mathrm{tp}^{\mathsf{ARVE_T}}(a/C) = \mathrm{tp}^{\mathsf{ARVE_T}}(b/C).$$

*Proof.* Let $\mathcal{N} \models \mathsf{ARV}(\times)$ be a $\kappa$-universal domain for an uncountable cardinal $\kappa$. Without loss of generality, we may assume that $\mathcal{N}$ is of the form $L^1\big((\Omega, \mathscr{C}, \mu), [0,1]\big)$ where $\mathscr{A}_t \subseteq \mathscr{B} \subseteq \mathscr{C}$ and $\mathscr{A}_t, \mathscr{B}$ are small for each $t \in T$. Hence, $A_\infty$ is also small. Then by interpreting $E_t$ as $\mathbb{E}(\cdot \mid \mathscr{A}_t)$ in $\mathcal{N}$ for each $t \in T$, the $L_{\mathsf{RV}(\times)}$-structure $\mathcal{N}$ is expanded to become an $L_{\mathsf{RVE_T}}$-structure. Since $\mathscr{A}_t$ is small for each $t \in T$, by Theorems 3.4.2 and 5.5.3 we know that $\mathscr{C}$ is also atomless over $\mathscr{A}_t$, whence $\mathcal{N}$ is a model of $\mathsf{ARVE_T}$. Note that $\mathcal{M} \subseteq \mathcal{N}$. Then by Proposition 7.2.4 ($\mathsf{ARVE_T}$ is model complete), we have that $\mathcal{M} \preceq \mathcal{N}$. Since $C$ and $A_\infty$ are small and $\mathrm{tp}^{\mathsf{ARV}}(a/C \cup A_\infty) = \mathrm{tp}^{\mathsf{ARV}}(b/C \cup A_\infty)$, there is $\Phi \in \mathrm{Aut}_{\mathsf{RV}}\big(L^1((\Omega, \mathscr{C}, \mu), [0,1])\big)$ such that $\Phi(a) = b$ and $\Phi$ fixes $C \cup A_\infty$ pointwise. Since $\Phi$ fixes $A_\infty$ and $\sigma(A_\infty) = \mathscr{A}_\infty$, we have $\Phi \restriction_{L^1((\Omega, \mathscr{A}_t, \mu), [0,1])}$ is in $\mathrm{Aut}_{\mathsf{RV}}\big(L^1((\Omega, \mathscr{A}_t, \mu), [0,1])\big)$ for each $t \in T$. By Lemma 7.2.1, we have that $\Phi$ is an $L_{\mathsf{RVE_T}}$-automorphism of $\Big(L^1\big((\Omega, \mathscr{C}, \mu), [0,1]\big), (\mathbb{E}(\cdot \mid \mathscr{A}_t))_{t \in T}\Big)$. Then since $\Phi(a) = b$ and $\Phi$ fixes $C$, we have $\mathrm{tp}^{\mathsf{ARVE_T}}(a/C) = \mathrm{tp}^{\mathsf{ARVE_T}}(b/C)$. □

For all $r = (r_1, \cdots, r_n) \in [0,1]^n$, define $B_r := [0, r_1) \times \cdots \times [0, r_n)$.

**7.2.9 Theorem.** *Let $\mathcal{M} \models \mathsf{ARVE_T}$ be a $\kappa$-universal domain where $\kappa > 2^{|T|}$. Assume $\mathcal{M}$ is of the form $(L^1((\Omega, \mathscr{B}, \mu), [0,1]), \big(\mathbb{E}(\cdot \mid \mathscr{A}_t)\big)_{t \in T})$, where $(\Omega, \mathscr{B}, \mu)$ is a probability space, $(\mathscr{A}_t)_{t \in T}$ is an increasing family of $\sigma$-subalgebras of $\mathscr{B}$, each algebra $\mathscr{A}_t$ is atomless over all previous ones (if $T$ has a minimal element 0, then $\mathscr{A}_0$ is atomless), and $\mathscr{B}$ is atomless over $\mathscr{A}_t$ for each $t$. For each $t \in T$, let $T(t)$ denote the set $\{s \in T \mid s < t\}$. For all $a, b \in M^n$, we have $\mathrm{tp}^{\mathsf{ARVE_T}}(a) = \mathrm{tp}^{\mathsf{ARVE_T}}(b)$ if and only if*

$$\mathrm{tp}^{\mathsf{ARVE_{T(t)}}}\big(\mathbb{E}(a \in B_{r^1} \mid \mathscr{A}_t), \cdots, \mathbb{E}(a \in B_{r^m} \mid \mathscr{A}_t)\big) = \mathrm{tp}^{\mathsf{ARVE_{T(t)}}}\big(\mathbb{E}(b \in B_{r^1} \mid \mathscr{A}_t), \cdots, \mathbb{E}(b \in B_{r^m} \mid \mathscr{A}_t)\big),$$

*for all $t \in T$, all $m \in \mathbb{N}$ and all $r^1, \cdots, r^m \in [0,1]^n$. (When $T(t) = \emptyset$, the theory $\mathsf{ARVE_{T(t)}}$ is defined to be $\mathsf{ARV}(\times)$.)*

*Proof.* The proof is a revision of the proof of Theorem 6.2.8.

$\Longrightarrow$: Since $\mathcal{M}$ is a $\kappa$-universal domain and $\mathrm{tp}^{\mathsf{ARVE_T}}(a) = \mathrm{tp}^{\mathsf{ARVE_T}}(b)$, there is an automorphism $\Phi$ of $\mathcal{M}$ such that $\Phi(a) = b$. By Lemma 5.2.5, we have that

$$\mathrm{tp}^{\mathsf{ARVE_T}}(\mathbf{1}_{a \in B_{r^1}}, \cdots, \mathbf{1}_{a \in B_{r^m}}) = \mathrm{tp}^{\mathsf{ARVE_T}}(\mathbf{1}_{b \in B_{r^1}}, \cdots, \mathbf{1}_{b \in B_{r^m}})$$

for all $m \in \mathbb{N}$ and all $r^1, \cdots, r^m \in [0,1]^n$, and thus for all $t \in T$, we have

$$\mathrm{tp}^{\mathsf{ARVE}_{\mathsf{T}(\mathrm{t})}}(E_t(\mathbf{1}_{a \in B_{r^1}}), \cdots, E_t(\mathbf{1}_{a \in B_{r^m}})) = \mathrm{tp}^{\mathsf{ARVE}_{\mathsf{T}(\mathrm{t})}}(E_t(\mathbf{1}_{b \in B_{r^1}}), \cdots, E_t(\mathbf{1}_{b \in B_{r^m}})).$$

$\Longleftarrow$: For each $t \in T$, let $E_t(f)$ denote $\mathbb{E}(f \mid \mathscr{A}_t)$ for every $f \in M$. First, we show that $\mathrm{tp}^{\mathsf{ARVE}_{\mathsf{T}'}}(a) = \mathrm{tp}^{\mathsf{ARVE}_{\mathsf{T}'}}(b)$ for any finite $T' \subseteq T$. Without loss of generality, we may assume that $T$ is finite. We use induction on $|T|$. When $|T| = 1$, this is Theorem 6.2.8. Suppose it holds when $|T| = k-1$. We will prove it holds when $|T| = k$. Suppose $T = \{1, 2, \cdots, k\}$ and let $T'$ be $\{1, 2, \cdots, k-1\}$. By assumption, for all $r^1, \cdots, r^m \in [0,1]^n$,

$$\mathrm{tp}^{\mathsf{ARVE}_{\mathsf{T}'}}\big(E_k(a \in B_{r^1}), \cdots, E_k(a \in B_{r^m})\big) = \mathrm{tp}^{\mathsf{ARVE}_{\mathsf{T}'}}\big(E_k(b \in B_{r^1}), \cdots, E_k(b \in B_{r^m})\big).$$

Therefore

$$\mathrm{tp}^{\mathsf{ARVE}_{\mathsf{T}'}}\big(\{E_k(a \in B_r)\}_{r \in [0,1]^n}\big) = \mathrm{tp}^{\mathsf{ARVE}_{\mathsf{T}'}}\big(\{E_k(b \in B_r)\}_{r \in [0,1]^n}\big).$$

Let $\lambda$ be a cardinal $> 2^{|\mathscr{B}|}$. Let $\mathcal{N} \models \mathsf{ARV}(\times)$ be a $\lambda$-universal domain and we may assume that $\mathcal{N}$ is of the form $L^1\big((\Omega, \mathscr{C}, \mu), [0,1]\big)$ where $\mathscr{B} \subseteq \mathscr{C}$ and $\mathscr{B}$ is small. Since $\mathscr{A}_k$ is small, by Theorems 3.4.2 and 5.5.3 we have that $\mathscr{C}$ is also atomless over $\mathscr{A}_k$. By interpreting $E_l$ as $\mathbb{E}(\cdot \mid \mathscr{A}_l)$ for each $l \leq k$ in $\mathcal{N}$, we expand $\mathcal{N}$ to a model of $\mathsf{ARVE}_{\mathsf{T}}$. Note that $\mathcal{M} \subseteq \mathcal{N}$. By Proposition 7.2.4, we have $\mathcal{M} \preceq \mathcal{N}$. Since $\mathcal{M}$ is a $\kappa$-universal domain, we get that $L^1\big((\Omega, \mathscr{A}_k, \mu), [0,1], \{\mathbb{E}(\cdot \mid \mathscr{A}_l)\}_{l \in T'}\big)$ is a $\kappa$-universal domain for $\mathsf{ARVE}_{\mathsf{T}'}$. Then there is an $L_{\mathsf{RVE}_{\mathsf{T}'}}$-automorphism $\Phi$ of the $L_{\mathsf{RVE}_{\mathsf{T}'}}$-structure $L^1\big((\Omega, \mathscr{A}_k, \mu), [0,1], \{\mathbb{E}(\cdot \mid \mathscr{A}_l)\}_{l \in T'}\big)$ such that $\Phi(E_k(a \in B_r)) = E_k(b \in B_r)$ for every $r \in [0,1]^n$. Then by the fact that $\mathcal{N}$ as a model of $\mathsf{ARV}(\times)$ is a $\lambda$-universal domain with $\lambda > 2^{|\mathscr{B}|} \geq 2^{|\mathscr{A}_k|}$, we know that there is an $L_{\mathsf{RV}(\times)}$-automorphism $\Psi$ of the $L_{\mathsf{RV}(\times)}$-structure $L^1((\Omega, \mathscr{C}, \mu), [0,1])$ extending $\Phi$. Let $\psi$ be the induced automorphism of $\mathscr{C}$. Then $\psi(\mathscr{A}_l) = \mathscr{A}_l$ and $\psi^{-1}(\mathscr{A}_l) = \mathscr{A}_l$ for each $l \in T$, since $\Psi$ extends $\Phi$. By Lemma 7.2.1, $\Psi$ is an $L_{\mathsf{RVE}_{\mathsf{T}}}$-automorphism of the $L_{\mathsf{RVE}_{\mathsf{T}}}$-structure $\mathcal{N}$.

Let $a' = \Psi(a)$. By Lemma 5.2.5, we have that $\Psi(\mathbf{1}_{f \in B_r}) = \mathbf{1}_{\Psi(f) \in B_r}$ for all $r \in [0,1]^n$ and all $f \in N^n$. Then because $\Psi$ is an automorphism of the $L_{\mathsf{RVE}_{\mathsf{T}}}$-structure $\mathcal{N}$, we have

$$E_k(a' \in B_r) = E_k(\mathbf{1}_{\Psi(a) \in B_r}) = E_k(\Psi(\mathbf{1}_{a \in B_r})) = \Psi(E_k(\mathbf{1}_{a \in B_r})) = \Psi(E_k(a \in B_r)),$$

for every $r \in [0,1]^n$. Therefore $E_k(a' \in B_r) = E_k(b \in B_r)$ for every $r \in [0,1]^n$, whereby $a'$ and $b$ determine the same conditional distribution over $\mathscr{A}_k$. By Theorem 5.2.3($v$) and Lemma 7.2.8, we have $\mathrm{tp}^{\mathsf{ARVE}_{\mathsf{T}}}(a') = \mathrm{tp}^{\mathsf{ARVE}_{\mathsf{T}}}(b)$, whereby $\mathrm{tp}^{\mathsf{ARVE}_{\mathsf{T}}}(a) = \mathrm{tp}^{\mathsf{ARVE}_{\mathsf{T}}}(b)$.

Second, for any $L_{\mathsf{RVE_T}}$-formula $\varphi$, only finitely many $t \in T$ are involved in the formula. Let $T(\varphi)$ denote the set of those $t$'s. Then we just proved $\mathrm{tp}^{\mathsf{ARVE}_{\mathsf{T}(\varphi)}}(a) = \mathrm{tp}^{\mathsf{ARVE}_{\mathsf{T}(\varphi)}}(b)$. Since $\varphi$ is an $L_{\mathsf{RVE}_{\mathsf{T}(\varphi)}}$-formula, we have $\varphi^{\mathcal{M}}(a) = \varphi^{\mathcal{M}}(b)$. Then because $\varphi$ is arbitrarily chosen, we have $\mathrm{tp}^{\mathsf{ARVE_T}}(a) = \mathrm{tp}^{\mathsf{ARVE_T}}(b)$. □

**7.2.10 Remark.** *In the proof, we see that if $T$ has a maximum element $s$, then for all $a, b \in M^n$, we have that $\mathrm{tp}^{\mathsf{ARVE_T}}(a) = \mathrm{tp}^{\mathsf{ARVE_T}}(b)$ if and only if*

$$\mathrm{tp}^{\mathsf{ARVE}_{\mathsf{T}(\mathsf{s})}}\big(\mathbb{E}(a \in B_{r^1} \mid \mathscr{A}_s), \cdots, \mathbb{E}(a \in B_{r^m} \mid \mathscr{A}_s)\big) = \mathrm{tp}^{\mathsf{ARVE}_{\mathsf{T}(\mathsf{s})}}\big(\mathbb{E}(b \in B_{r^1} \mid \mathscr{A}_s), \cdots, \mathbb{E}(b \in B_{r^m} \mid \mathscr{A}_s)\big),$$

*for all $m \in \mathbb{N}$ and for all $r^1, \cdots, r^m \in [0,1]^n$.*

**7.2.11 Theorem.** *Let $\mathcal{M} \models \mathsf{ARVE_T}$ be a $\kappa$-universal domain where $\kappa > 2^{|T|}$. Assume $\mathcal{M}$ is of the form $(L^1((\Omega, \mathscr{B}, \mu), [0,1]), \big(\mathbb{E}(\cdot \mid \mathscr{A}_t)\big)_{t \in T})$, where $(\mathscr{A}_t)_{t \in T}$ is an increasing family of $\sigma$-subalgebras of $\mathscr{B}$, each algebra $\mathscr{A}_t$ is atomless over all previous ones (if $T$ has a minimal element 0, then $\mathscr{A}_0$ is atomless), and $\mathscr{B}$ is atomless over $\mathscr{A}_t$ for each $t$. For each $t \in T$, let $T(t)$ denote the set $\{s \in T \mid s < t\}$. For all $a, b \in M^n$, we have $\mathrm{tp}^{\mathsf{ARVE_T}}(a) = \mathrm{tp}^{\mathsf{ARVE_T}}(b)$ if and only if for all $t \in T$, all $m \in \mathbb{N}$, and all continuous $\varphi_1, \cdots, \varphi_m \colon [0,1]^n \to [0,1]$, one has*

$$\mathrm{tp}^{\mathsf{ARVE}_{\mathsf{T}(\mathsf{t})}}\big(\mathbb{E}(\varphi_1(a) \mid \mathscr{A}_t), \cdots, \mathbb{E}(\varphi_m(a) \mid \mathscr{A}_t)\big) = \mathrm{tp}^{\mathsf{ARVE}_{\mathsf{T}(\mathsf{t})}}\big(\mathbb{E}(\varphi_1(b) \mid \mathscr{A}_t), \cdots, \mathbb{E}(\varphi_m(b) \mid \mathscr{A}_t)\big).$$

*(When $T(t) = \emptyset$, the theory $\mathsf{ARVE}_{\mathsf{T}(\mathsf{t})}$ is defined to be $\mathsf{ARV}(\times)$.)*

*Proof.* The proof is a revision of the proofs of Theorems 7.2.9 and 6.2.9.

$\Longrightarrow$: Since $\mathcal{M}$ is a $\kappa$-universal domain and $\mathrm{tp}^{\mathsf{ARVE_T}}(a) = \mathrm{tp}^{\mathsf{ARVE_T}}(b)$, there is an automorphism $\Phi$ of $\mathcal{M}$ such that $\Phi(a) = b$. By Lemma 5.2.5, we have that

$$\mathrm{tp}^{\mathsf{ARVE_T}}(\varphi_1(a), \cdots, \varphi_m(a)) = \mathrm{tp}^{\mathsf{ARVE_T}}(\varphi_1(b), \cdots, \varphi_m(b))$$

for all $m \in \mathbb{N}$ and all continuous $\varphi_1, \cdots, \varphi_m \colon [0,1]^n \to [0,1]$, and thus for all $t \in T$, we have

$$\mathrm{tp}^{\mathsf{ARVE}_{\mathsf{T}(\mathsf{t})}}(E_t(\varphi_1(a)), \cdots, E_t(\varphi_m(a))) = \mathrm{tp}^{\mathsf{ARVE}_{\mathsf{T}(\mathsf{t})}}(E_t(\varphi_1(b)), \cdots, E_t(\varphi_m(b)))$$

$\Longleftarrow$: For each $t \in T$, let $E_t(f)$ denote $\mathbb{E}(f \mid \mathscr{A}_t)$ for every $f \in M$. First, we show that $\mathrm{tp}^{\mathsf{ARVE}_{\mathsf{T}'}}(a) = \mathrm{tp}^{\mathsf{ARVE}_{\mathsf{T}'}}(b)$ for any finite $T' \subseteq T$. Without loss of generality, we may assume that $T$ is finite. We use induction on $|T|$. When $|T| = 1$, this is Theorem 6.2.9. Suppose it holds when $|T| = k - 1$. We will prove

it holds when $|T| = k$. Suppose $T = \{1, 2, \cdots, k\}$ and let $T'$ be $\{1, 2, \cdots, k-1\}$. By assumption, for all $r^1, \cdots, r^m \in [0,1]^n$,

$$\mathrm{tp}^{\mathsf{ARVE}_{\mathsf{T}'}}\big(E_k(a \in B_{r^1}), \cdots, E_k(a \in B_{r^m})\big) = \mathrm{tp}^{\mathsf{ARVE}_{\mathsf{T}'}}\big(E_k(b \in B_{r^1}), \cdots, E_k(b \in B_{r^m})\big).$$

Therefore

$$\mathrm{tp}^{\mathsf{ARVE}_{\mathsf{T}'}}\big(\{E_k(a \in B_r)\}_{r\in[0,1]^n}\big) = \mathrm{tp}^{\mathsf{ARVE}_{\mathsf{T}'}}\big(\{E_k(b \in B_r)\}_{r\in[0,1]^n}\big).$$

Let $\lambda$ be a cardinal $> 2^{|\mathscr{B}|}$. Let $\mathcal{N} \models \mathsf{ARV}(\times)$ be a $\lambda$-universal domain and we may assume that $\mathcal{N}$ is of the form $L^1\big((\Omega, \mathscr{C}, \mu), [0,1]\big)$ where $\mathscr{B} \subseteq \mathscr{C}$ and $\mathscr{B}$ is small. Since $\mathscr{A}_k$ is small, by Theorems 3.4.2 and 5.5.3 we have that $\mathscr{C}$ is also atomless over $\mathscr{A}_k$. By interpreting $E_l$ as $\mathbb{E}(\cdot \mid \mathscr{A}_l)$ for each $l \leq k$ in $\mathcal{N}$, we expand $\mathcal{N}$ to a model of $\mathsf{ARVE}_{\mathsf{T}}$. Note that $\mathcal{M} \subseteq \mathcal{N}$. By Proposition 7.2.4, we have $\mathcal{M} \preceq \mathcal{N}$. Since $\mathcal{M}$ is a $\kappa$-universal domain, we get that $L^1\big((\Omega, \mathscr{A}_k, \mu), [0,1], \{\mathbb{E}(\cdot \mid \mathscr{A}_l)\}_{l\in T'}\big)$ is a $\kappa$-universal domain for $\mathsf{ARVE}_{\mathsf{T}'}$. Then there is an $L_{\mathsf{RVE}_{\mathsf{T}'}}$-automorphism $\Phi$ of the $L_{\mathsf{RVE}_{\mathsf{T}'}}$-structure $L^1\big((\Omega, \mathscr{A}_k, \mu), [0,1], \{\mathbb{E}(\cdot \mid \mathscr{A}_l)\}_{l\in T'}\big)$ such that $\Phi(E_k(a \in B_r)) = E_k(b \in B_r)$ for every $r \in [0,1]^n$. Then by the fact that $\mathcal{N}$ as a model of $\mathsf{ARV}(\times)$ is a $\lambda$-universal domain with $\lambda > 2^{|\mathscr{B}|} \geq 2^{|\mathscr{A}_k|}$, we know that there is an $L_{\mathsf{RV}(\times)}$-automorphism $\Psi$ of the $L_{\mathsf{RV}(\times)}$-structure $L^1((\Omega, \mathscr{C}, \mu), [0,1])$ extending $\Phi$. Let $\psi$ be the induced automorphism of $\mathscr{C}$. Then $\psi(\mathscr{A}_l) = \mathscr{A}_l$ and $\psi^{-1}(\mathscr{A}_l) = \mathscr{A}_l$ for each $l \in T$, since $\Psi$ extends $\Phi$. By Lemma 7.2.1, $\Psi$ is an $L_{\mathsf{RVE}_{\mathsf{T}}}$-automorphism of the $L_{\mathsf{RVE}_{\mathsf{T}}}$-structure $\mathcal{N}$.

Let $a' = \Psi(a)$. By Lemma 5.2.5, we have that $\Psi(\varphi(a)) = \varphi(\Psi(a)) = \varphi(a')$ for all continuous $\varphi\colon [0,1]^n \to [0,1]$. Because $\Psi$ extends $\Phi$, for all continuous $\varphi\colon [0,1]^n \to [0,1]$ we have $\Psi(E_k(\varphi(a))) = E_k(\varphi(b))$. Thus $E_k(\Psi(\varphi(a))) = \Psi(E_k(\varphi(a))) = E_k(\varphi(b))$, since $\Psi$ is an automorphism of the $L_{\mathsf{RVE}_{\mathsf{T}}}$-structure $\mathcal{N}$. Then by the fact that $\Psi(\varphi(a)) = \varphi(\Psi(a)) = \varphi(a')$, we have $E_k(\varphi(a')) = E_k(\varphi(b))$ for all continuous $\varphi\colon [0,1]^n \to [0,1]$. Using Lusin's theorem and the dominated convergence theorem, we have that $E_k(a' \in B) = E_k(b \in B)$ for all Borel $B \subseteq [0,1]^n$. That is, $a'$ and $b$ determine the same conditional distribution over $\mathscr{A}_k$. By Theorem 5.2.3($v$) and Lemma 7.2.8, we have $\mathrm{tp}^{\mathsf{ARVE}_{\mathsf{T}}}(a') = \mathrm{tp}^{\mathsf{ARVE}_{\mathsf{T}}}(b)$, whereby $\mathrm{tp}^{\mathsf{ARVE}_{\mathsf{T}}}(a) = \mathrm{tp}^{\mathsf{ARVE}_{\mathsf{T}}}(b)$.

Second, for any $L_{\mathsf{RVE}_{\mathsf{T}}}$-formula $\varphi$, only finitely many $t \in T$ are involved in the formula. Let $T(\varphi)$ denote the set of those $t$'s. Then we just proved $\mathrm{tp}^{\mathsf{ARVE}_{\mathsf{T}(\varphi)}}(a) = \mathrm{tp}^{\mathsf{ARVE}_{\mathsf{T}(\varphi)}}(b)$. Since $\varphi$ is an $L_{\mathsf{RVE}_{\mathsf{T}(\varphi)}}$-formula, we have $\varphi^{\mathcal{M}}(a) = \varphi^{\mathcal{M}}(b)$. Then because $\varphi$ is arbitrarily chosen, we have $\mathrm{tp}^{\mathsf{ARVE}_{\mathsf{T}}}(a) = \mathrm{tp}^{\mathsf{ARVE}_{\mathsf{T}}}(b)$. □

**7.2.12 Remark.** *In the proof, we see that if $T$ has a maximum element $s$, then for all $a, b \in M^n$, we have that $\mathrm{tp}^{\mathsf{ARVE}_{\mathsf{T}}}(a) = \mathrm{tp}^{\mathsf{ARVE}_{\mathsf{T}}}(b)$ if and only if*

$$\mathrm{tp}^{\mathsf{ARVE}_{\mathsf{T}(s)}}\big(\mathbb{E}(\varphi_1(a) \mid \mathscr{A}_s), \cdots, \mathbb{E}(\varphi_m(a) \mid \mathscr{A}_s)\big) = \mathrm{tp}^{\mathsf{ARVE}_{\mathsf{T}(s)}}\big(\mathbb{E}(\varphi_1(b) \mid \mathscr{A}_s), \cdots, \mathbb{E}(\varphi_m(b) \mid \mathscr{A}_s)\big),$$

*for all $m \in \mathbb{N}$ and all continuous $\varphi_1, \cdots, \varphi_m \colon [0,1]^n \to [0,1]$.*

**7.2.13 Corollary.** *Let $\mathcal{M} \models \mathsf{ARVE}_\mathsf{T}$ be a $\kappa$-universal domain where $\kappa > 2^{|T|}$. Assume $\mathcal{M}$ is of the form $(L^1((\Omega, \mathscr{B}, \mu), [0,1]), \big(\mathbb{E}(\cdot \mid \mathscr{A}_t)\big)_{t \in T})$, where $(\Omega, \mathscr{B}, \mu)$ is a probability space, $(\mathscr{A}_t)_{t \in T}$ is an increasing family of $\sigma$-subalgebras of $\mathscr{B}$, each algebra $\mathscr{A}_t$ is atomless over all previous ones (if $T$ has a minimal element 0, then $\mathscr{A}_0$ is atomless), and $\mathscr{B}$ is atomless over $\mathscr{A}_t$ for each $t$. For each $t \in T$, let $T(t)$ denote the set $\{s \in T \mid s < t\}$. For all $a, b \in M^n$ and small $C \subseteq M$, the following are equivalent:*

*(i)* $\mathrm{tp}^{\mathsf{ARVE}_\mathsf{T}}(a/C) = \mathrm{tp}^{\mathsf{ARVE}_\mathsf{T}}(b/C)$.

*(ii) For all $t \in T$, all $k, l \in \mathbb{N}$, all $r^1, \cdots, r^k \in [0,1]^{n+l}$, and all $c_1, \cdots, c_l \in C$, one has*

$$\begin{aligned}&\mathrm{tp}^{\mathsf{ARVE}_{\mathsf{T}(\mathsf{t})}}\Big(\mathbb{E}\big((a, c_1, \cdots, c_l) \in B_{r^1} \mid \mathscr{A}_t\big), \cdots, \mathbb{E}\big((a, c_1, \cdots, c_l) \in B_{r^k} \mid \mathscr{A}_t\big)\Big)\\ =&\mathrm{tp}^{\mathsf{ARVE}_{\mathsf{T}(\mathsf{t})}}\Big(\mathbb{E}\big((b, c_1, \cdots, c_l) \in B_{r^1} \mid \mathscr{A}_t\big), \cdots, \mathbb{E}\big((b, c_1, \cdots, c_l) \in B_{r^k} \mid \mathscr{A}_t\big)\Big).\end{aligned}$$

*(iii) For all $t \in T$, all $k, l \in \mathbb{N}$, all continuous $\varphi_1, \cdots, \varphi_k \colon [0,1]^{n+l} \to [0,1]$, and all $c_1, \cdots, c_l \in C$, one has*

$$\begin{aligned}&\mathrm{tp}^{\mathsf{ARVE}_{\mathsf{T}(\mathsf{t})}}\Big(\mathbb{E}\big(\varphi_1(a, c_1, \cdots, c_l) \mid \mathscr{A}_t\big), \cdots, \mathbb{E}\big(\varphi_k(a, c_1, \cdots, c_l) \mid \mathscr{A}_t\big)\Big)\\ =&\mathrm{tp}^{\mathsf{ARVE}_{\mathsf{T}(\mathsf{t})}}\Big(\mathbb{E}\big(\varphi_1(b, c_1, \cdots, c_l) \mid \mathscr{A}_t\big), \cdots, \mathbb{E}\big(\varphi_k(b, c_1, \cdots, c_l) \mid \mathscr{A}_t\big)\Big).\end{aligned}$$

*(When $T(t) = \emptyset$, the theory $\mathsf{ARVE}_{\mathsf{T}(\mathsf{t})}$ is defined to be $\mathsf{ARV}(\times)$.)*

*Proof.* $(i) \Rightarrow (ii)$ and $(i) \Rightarrow (iii)$: Since $\mathrm{tp}(a/C) = \mathrm{tp}(b/C)$, we have $\mathrm{tp}(a, c_1, \cdots, c_l) = \mathrm{tp}(b, c_1, \cdots, c_l)$ for all $l \in \mathbb{N}$ and all $c_1, \cdots, c_l \in C$. Then use Theorem 7.2.9 and Theorem 7.2.11 respectively.

$(ii) \Rightarrow (i)$: By Theorem 7.2.9, $\mathrm{tp}(a, c_1, \cdots, c_l) = \mathrm{tp}(b, c_1, \cdots, c_l)$ for all $l \in \mathbb{N}$ and all $c_1, \cdots, c_l \in C$. Therefore $\mathrm{tp}(a/C) = \mathrm{tp}(b/C)$.

$(iii) \Rightarrow (i)$: By Theorem 7.2.11, $\mathrm{tp}(a, c_1, \cdots, c_l) = \mathrm{tp}(b, c_1, \cdots, c_l)$ for all $l \in \mathbb{N}$ and all $c_1, \cdots, c_l \in C$. Therefore $\mathrm{tp}(a/C) = \mathrm{tp}(b/C)$. □

**7.2.14 Remark.** *Using Remarks 7.2.10 and 7.2.12, if $T$ has a maximum element $s$, then for all $a, b \in M^n$ and small $C \subseteq M$, the following are equivalent:*

*(i)* $\mathrm{tp}^{\mathsf{ARVE}_\mathsf{T}}(a/C) = \mathrm{tp}^{\mathsf{ARVE}_\mathsf{T}}(b/C)$.

*(ii) For all $k, l \in \mathbb{N}$, all $r^1, \cdots, r^k \in [0,1]^{n+l}$, and all $c_1, \cdots, c_l \in C$, one has*

$$\begin{aligned}&\mathrm{tp}^{\mathsf{ARVE}_{\mathsf{T}(\mathsf{s})}}\Big(\mathbb{E}\big((a, c_1, \cdots, c_l) \in B_{r^1} \mid \mathscr{A}_t\big), \cdots, \mathbb{E}\big((a, c_1, \cdots, c_l) \in B_{r^k} \mid \mathscr{A}_t\big)\Big)\\ =&\mathrm{tp}^{\mathsf{ARVE}_{\mathsf{T}(\mathsf{s})}}\Big(\mathbb{E}\big((b, c_1, \cdots, c_l) \in B_{r^1} \mid \mathscr{A}_t\big), \cdots, \mathbb{E}\big((b, c_1, \cdots, c_l) \in B_{r^k} \mid \mathscr{A}_t\big)\Big).\end{aligned}$$

*(iii) For all $k, l \in \mathbb{N}$, all continuous $\varphi_1, \cdots, \varphi_k \colon [0,1]^{n+l} \to [0,1]$, and all $c_1, \cdots, c_l \in C$, one has*

$$\begin{aligned}&\mathrm{tp}^{\mathsf{ARVE}_{\mathsf{T}(\mathsf{s})}}\Big(\mathbb{E}\big(\varphi_1(a, c_1, \cdots, c_l) \mid \mathscr{A}_t\big), \cdots, \mathbb{E}\big(\varphi_k(a, c_1, \cdots, c_l) \mid \mathscr{A}_t\big)\Big)\\ =&\mathrm{tp}^{\mathsf{ARVE}_{\mathsf{T}(\mathsf{s})}}\Big(\mathbb{E}\big(\varphi_1(b, c_1, \cdots, c_l) \mid \mathscr{A}_t\big), \cdots, \mathbb{E}\big(\varphi_k(b, c_1, \cdots, c_l) \mid \mathscr{A}_t\big)\Big).\end{aligned}$$

## Quantifier Elimination

**7.2.15 Theorem.** *The theory $\mathsf{ARVE}_{\mathsf{T}}$ admits quantifier elimination.*

*Proof.* First, suppose that $T$ is finite. We use induction on $|T|$. When $|T| = 0$, we know that $\mathsf{ARVE}_{\mathsf{T}}$ is $\mathsf{ARV}$. Hence, this is Theorem 5.2.3$(ii)$. When $|T| = 1$, this is Theorem 6.2.11. Suppose $\mathsf{ARVE}_{\mathsf{T}}$ admits quantifier elimination when $|T| \leq k-1$, where $k \geq 2$. We prove that $\mathsf{ARVE}_{\mathsf{T}}$ admits quantifier elimination when $|T| = k$. For every $t \in T$, let $T(t)$ denote the set $\{s \in T \mid s < t\}$. Then $|T(t)| \leq k-1$ for all $t \in T$. Let $\mathcal{M} = (L^1((\Omega, \mathscr{B}, \mu), [0,1]), \big(\mathbb{E}(\cdot \mid \mathscr{A}_t)\big)_{t\in T})$ be a $\kappa$-universal domain for $\mathsf{ARVE}_{\mathsf{T}}$, where $(\mathscr{A}_t)_{t\in T}$ is an increasing family of $\sigma$-subalgebras of $\mathscr{B}$. Let $a, b \in M^n$ with $\mathrm{tp}^{\mathsf{ARVE}_{\mathsf{T}},\mathrm{qf}}(a) = \mathrm{tp}^{\mathsf{ARVE}_{\mathsf{T}},\mathrm{qf}}(b)$. By Lemma 5.2.1, for all $m \in \mathbb{N}$ and all continuous $\varphi_1, \cdots, \varphi_m \colon [0,1]^n \to [0,1]$, there are sequences of $n$-ary $L_{\mathsf{RV}}$-terms $(t_{k,\varphi_1})_{k\in\mathbb{N}}, \cdots, (t_{k,\varphi_m})_{k\in\mathbb{N}}$ such that for each $j = 1, \cdots, m$, we have that $t^{\mathcal{M}}_{k,\varphi_j}$ uniformly converges to $\varphi_j^{\mathcal{M}}$, where $\varphi_j^{\mathcal{M}} \colon M^n \to [0,1]$ is defined by $\varphi_j^{\mathcal{M}}(f) = \varphi_j(f)$ for all $f \in M^n$. Then for every $t \in T$ and every quantifier free $L_{\mathsf{RVE}_{\mathsf{T}(\mathsf{t})}}$-formula $F(z)$ where $z$ is a tuple of $m$ variables, by the fact that $F^{\mathcal{M}}$ and $E_t^{\mathcal{M}}$ are continuous, we have that $F^{\mathcal{M}}\big(E_t^{\mathcal{M}}(t^{\mathcal{M}}_{k,\varphi_1}(f)), \cdots, E_t^{\mathcal{M}}(t^{\mathcal{M}}_{k,\varphi_m}(f))\big)$ converges to $F^{\mathcal{M}}\big(E_t^{\mathcal{M}}(\varphi_1(f)), \cdots, E_t^{\mathcal{M}}(\varphi_m(f))\big)$ for all $f \in M^n$. Then because $F\big(E_t(t_{k,\varphi_1}), \cdots, E_t(t_{k,\varphi_m})\big)$ is a quantifier free $L_{\mathsf{RVE}_{\mathsf{T}}}$-formula for all $k \in \mathbb{N}$ and $\mathrm{tp}^{\mathsf{ARVE}_{\mathsf{T}},\mathrm{qf}}(a) = \mathrm{tp}^{\mathsf{ARVE}_{\mathsf{T}},\mathrm{qf}}(b)$, we have that

$$F^{\mathcal{M}}\big(E_t^{\mathcal{M}}(\varphi_1(a)), \cdots, E_t^{\mathcal{M}}(\varphi_m(a))\big) = F^{\mathcal{M}}\big(E_t^{\mathcal{M}}(\varphi_1(b)), \cdots, E_t^{\mathcal{M}}(\varphi_m(b))\big).$$

Hence,

$$\mathrm{tp}^{\mathsf{ARVE}_{\mathsf{T}(\mathsf{t})},\mathrm{qf}}\big(E_t(\varphi_1(a)), \cdots, E_t(\varphi_m(a))\big) = \mathrm{tp}^{\mathsf{ARVE}_{\mathsf{T}(\mathsf{t})},\mathrm{qf}}\big(E_t(\varphi_1(b)), \cdots, E_t(\varphi_m(b))\big),$$

for all $t \in T$, all $m \in \mathbb{N}$, and all continuous functions $\varphi_1, \cdots, \varphi_m \colon [0,1]^n \to [0,1]$. By the induction

hypothesis, we know that $\mathsf{ARVE}_{\mathsf{T}(\mathsf{t})}$ admits quantifier elimination. Hence

$$\mathrm{tp}^{\mathsf{ARVE}_{\mathsf{T}(\mathsf{t})}}\big(E_t(\varphi_1(a)),\cdots,E_t(\varphi_m(a))\big)=\mathrm{tp}^{\mathsf{ARVE}_{\mathsf{T}(\mathsf{t})}}\big(E_t(\varphi_1(b)),\cdots,E_t(\varphi_m(b))\big),$$

for all $t\in T$, for all $m\in\mathbb{N}$, and for all continuous $\varphi_1,\cdots,\varphi_m\colon[0,1]^n\to[0,1]$. By Theorem 7.2.11, $\mathrm{tp}^{\mathsf{ARVE}_{\mathsf{T}}}(a)=\mathrm{tp}^{\mathsf{ARVE}_{\mathsf{T}}}(b)$. By Theorem 1.10.4, $\mathsf{ARVE}_{\mathsf{T}}$ admits quantifier elimination.

Second, we give the proof for arbitrary $T$. We just proved that for any finite $T'\subseteq T$, the theory $\mathsf{ARVE}_{\mathsf{T}'}$ admits quantifier elimination. Clearly, $\mathsf{ARVE}_{\mathsf{T}}\models\mathsf{ARVE}_{\mathsf{T}'}$. For every $L_{\mathsf{RVE}_{\mathsf{T}}}$-formula $\varphi(x)$, there is a finite subset $T(\varphi)\subseteq T$ such that $\varphi$ is an $L_{\mathsf{RVE}_{\mathsf{T}(\varphi)}}$-formula. Since $\mathsf{ARVE}_{\mathsf{T}(\varphi)}$ admits quantifier elimination, for every $\epsilon>0$ there is a quantifier free $L_{\mathsf{RVE}_{\mathsf{T}(\varphi)}}$-formula $\psi(x)$ such that $\mathsf{ARVE}_{\mathsf{T}(\varphi)}\models\sup_x|\varphi(x)-\psi(x)|\leq\epsilon$. Therefore, $\mathsf{ARVE}_{\mathsf{T}}\models\sup_x|\varphi(x)-\psi(x)|\leq\epsilon$. Hence, $\mathsf{ARVE}_{\mathsf{T}}$ admits quantifier elimination. □

**7.2.16 Proposition.** *The universal part of* $\mathsf{ARVE}_{\mathsf{T}}$ *is* $\mathsf{RVE}_{\mathsf{T}}$*, and* $\mathsf{ARVE}_{\mathsf{T}}$ *is the model completion of* $\mathsf{RVE}_{\mathsf{T}}$*.*

*Proof.* By Proposition 6.2.12, we know that $\mathsf{RVE}$ is a universal theory. Therefore, $\bigcup_{t\in T}\mathsf{RVE}_{\mathsf{t}}$ is a universal theory. Then from the form of the axioms in $\mathsf{RVE}_{\mathsf{T}}\setminus\bigcup_{t\in T}\mathsf{RVE}_{\mathsf{t}}$ (see page 139), $\mathsf{RVE}_{\mathsf{T}}$ is a universal theory. By Proposition 7.1.1 and Theorem 7.1.2, we have that the $\mathsf{RVE}_{\mathsf{T}}$-substructures of models of $\mathsf{ARVE}_{\mathsf{T}}$ are precisely the models of $\mathsf{RVE}_{\mathsf{T}}$. Therefore the theory $\mathsf{RVE}_{\mathsf{T}}$ is the universal part of the theory $\mathsf{ARVE}_{\mathsf{T}}$. Then by Theorem 7.2.15 and Proposition 1.10.5, we know that $\mathsf{ARVE}_{\mathsf{T}}$ is the model completion of $\mathsf{RVE}_{\mathsf{T}}$. □

## 7.3 On $d$-finite tuples in $\mathsf{ARVE}_{\mathsf{T}}$

In this section, we give the characterization of $d$-finite tuples in models of $\mathsf{ARVE}_{\mathsf{T}}$ when $T$ is finite.

We may assume that $T$ is finite. Then the theory $\mathsf{ARVE}_{\mathsf{T}}$ is separably categorical, by Theorem 7.2.2. Let $\mathcal{M}\models\mathsf{ARVE}_{\mathsf{T}}$. For an $n$-tuple $f=(f_1,\cdots,f_n)\in M^n$. Let $\mathsf{ARVE}_{\mathsf{T}}(\mathsf{f})$ denote the theory $Th(\mathcal{M},f)$. By Theorem 3.5.2, $f$ is $d$-finite (over $\emptyset$) if and only if $\mathsf{ARVE}_{\mathsf{T}}(\mathsf{f})$ is separably categorical. We use this result to characterize the $d$-finite tuples in $M$.

**7.3.1 Theorem.** *Assume* $T=\{1,\cdots,m\}$ *for some* $m\in\mathbb{N}$*. Let* $\mathcal{M}\models\mathsf{ARVE}_{\mathsf{T}}$ *and* $f=(f_1,\cdots,f_n)\in M^n$*. Let* $\mathcal{F}$ *denote* $L^1\big(\sigma(f),[0,1]\big)$*, where* $\sigma(f)$ *is the* $\sigma$*-algebra generated by* $f$*. Then* $f$ *is a* $d$*-finite tuple in* $M$ *if and only if the* $\sigma$*-algebras* $\sigma(f),\sigma\big(E_1(\mathcal{F})\big),\cdots,\sigma\big(E_m(\mathcal{F})\big)$ *are all atomic.*

*Proof.* $\Longrightarrow$: If $f$ is $d$-finite, then the theory $\mathsf{ARVE}_{\mathsf{T}}(\mathsf{f})$ is separably categorical. For each $1\leq j\leq m$, consider the signature $L_{\mathsf{RVE}_{\mathsf{j}}}\cup\{f\}$ where the only expectation symbol is $E_j$. Let $\mathsf{ARVE}_{\mathsf{j}}(\mathsf{f})$ denote the restriction of

$\mathsf{ARVE_T(f)}$ to $L_{\mathsf{RVE_j}} \cup \{f\}$. By Corollary 1.9.5, $\mathsf{ARVE_j(f)}$ is separably categorical. Then by Theorem 6.3.1, $\sigma(f)$ and $\sigma\big(E_j(\mathcal{F})\big)$ are both atomic.

$\Longleftarrow$: We use the induction on $|T| = m$. When $m = 1$, this is Theorem 6.3.1. Suppose it holds for $|T| = m-1$. We will show it holds for $|T| = m$. Let $T'$ denote $\{1, \cdots, m-1\}$.

Let $([0,1], \mathcal{L}, \lambda)$ denote the standard Lebesgue space. For all $n \in \mathbb{N}$, let $([0,1]^n, \mathcal{L}^n, \lambda^n)$ denote the product space of $n$ copies of $([0,1], \mathcal{L}, \lambda)$. For all $n > l \in \mathbb{N}$, define $\pi_{n,l}\colon [0,1]^n \to [0,1]^l$ by $\pi_{n,l}(x_1, \cdots, x_n) := (x_1, \cdots, x_l)$ for all $(x_1, \cdots, x_n) \in [0,1]^n$. Then for all $1 \leq n < m$, we have that $\pi_{m,n}^{-1}(\mathcal{L}^n)$ is atomless over all previous ones ($\pi_{m,1}^{-1}(\mathcal{L})$ is atomless), and $\mathcal{L}^m$ is atomless over $\pi_{m,n}^{-1}(\mathcal{L}^n)$. For all $n < m$, let $\mathscr{A}_n \subseteq \mathcal{L}^m$ be the smallest $\sigma$-algebra containing $\pi_{m,n}^{-1}(\mathcal{L}^n)$ and the set of all sets of measure zero in $\mathcal{L}^m$. Then we also have $\mathscr{A}_n$ is atomless over all previous ones ($\mathscr{A}_1$ is atomless), and $\mathcal{L}^m$ is atomless over $\mathscr{A}_n$ for all $n < m$. Hence, $\Big(L^1\big(([0,1]^m, \mathcal{L}^m, \lambda^m), [0,1]\big), \mathbb{E}(\cdot \mid \mathscr{A}_1), \cdots, \mathbb{E}(\cdot \mid \mathscr{A}_m)\Big)$ is a model of $\mathsf{ARVE_T}$. Suppose $\mathcal{M}$ is a separable model. By Corollary 7.2.2, $\mathcal{M}$ is isomorphic to $\Big(L^1\big(([0,1]^m, \mathcal{L}^m, \lambda^m), [0,1]\big), \mathbb{E}(\cdot \mid \mathscr{A}_1), \cdots, \mathbb{E}(\cdot \mid \mathscr{A}_m)\Big)$. Without loss of generality, we may identify $\mathcal{M}$ with $\Big(L^1\big(([0,1]^m, \mathcal{L}^m, \lambda^m), [0,1]\big), \mathbb{E}(\cdot \mid \mathscr{A}_1), \cdots, \mathbb{E}(\cdot \mid \mathscr{A}_m)\Big)$.

For $g \in M^n$ with $(\mathcal{M}, f) \equiv (\mathcal{M}, g)$, we will show there exists $\psi \in \mathrm{Aut}_{\mathsf{ARVE_T}}(\mathcal{M})$ such that $\psi(f) = g$. Then $\mathsf{ARVE_T(f)}$ is separably categorical. Let $\mathcal{G}$ denote $L^1\big(\sigma(g), [0,1]\big)$, where $\sigma(g)$ is the $\sigma$-algebra generated by $g$. For every $L_{\mathsf{RV}(\times)}$-term $t$, define $\phi_0(E_m(t(f))) = E_m(t(g))$. Since $\mathrm{tp}^{\mathsf{ARVE_T}}(f) = \mathrm{tp}^{\mathsf{ARVE_T}}(g)$, we have that $\phi_0$ is well-defined and is an $L_{\mathsf{RV}(\times)}$-elementary map in $\mathcal{M}$ from a subset of $E_m(\mathcal{F})$ into $E_m(\mathcal{G})$; *i.e.*, for all $L_{\mathsf{RV}(\times)}$-formulas $\tau$ and all $L_{\mathsf{RV}(\times)}$-terms $t_1, \cdots, t_k$, we have

$$\begin{aligned}\tau^{\mathcal{M}}\big(E_m(t_1(f)), \cdots, E_m(t_k(f))\big) &= \tau^{\mathcal{M}}\big(E_m(t_1(g)), \cdots, E_m(t_k(g))\big)\\ &= \tau^{\mathcal{M}}\Big(\phi_0\big(E_m(t_1(f))\big), \cdots, \phi_0\big(E_m(t_k(f))\big)\Big).\end{aligned}$$

Then by the fact that $\{E_m(t(f)) \mid t \text{ is an } L_{\mathsf{RV}(\times)}\text{-term}\}$ and $\{E_m(t(g)) \mid t \text{ is an } L_{\mathsf{RV}(\times)}\text{-term}\}$ are dense in $E_m(\mathcal{F})$ and $E_m(\mathcal{G})$ respectively, and $\mathrm{tp}^{\mathsf{ARVE_T}}(f) = \mathrm{tp}^{\mathsf{ARVE_T}}(g)$, we have that $\phi_0$ is extended to an $L_{\mathsf{RV}(\times)}$-elementary map $\phi\colon E_m(\mathcal{F}) \to E_m(\mathcal{G})$ in $\mathcal{M}$, which is distance preserving and bijective. Then because

$$L^1\big(\sigma(E_m(\mathcal{F})), [0,1]\big) = \overline{\{t(a_1, \cdots, a_k) \mid k \in \mathbb{N}, a_1, \cdots, a_k \in E_m(\mathcal{F}), t \text{ is an } L_{\mathsf{RV}(\times)}\text{-term}\}}$$

and

$$L^1\big(\sigma(E_m(\mathcal{G})), [0,1]\big) = \overline{\{t(b_1, \cdots, b_k) \mid k \in \mathbb{N}, b_1, \cdots, b_k \in E_m(\mathcal{G}), t \text{ is an } L_{\mathsf{RV}(\times)}\text{-term}\}},$$

we have that $\phi$ is extended to an $L_{\mathsf{RV}(\times)}$-isomorphism $\tilde{\phi}\colon L^1\big(\sigma(E_m(\mathcal{F})), [0,1]\big) \to L^1\big(\sigma(E_m(\mathcal{G})), [0,1]\big)$. Then by the fact that $\sigma\big(E_m(\mathcal{F})\big)$ is atomic, we know that $\sigma\big(E_m(\mathcal{G})\big)$ is also atomic. Since $\sigma(E_m(\mathcal{F}))$ is

atomic, there is a discrete $u \in L^1\big(\sigma(E_m(\mathcal{F})), [0,1]\big)$ such that $\sigma(u) = \sigma(E_m(\mathcal{F}))$ up to a set of sets of measure zero. Let $v$ denote $\tilde{\phi}(u)$. Then $\sigma(v) = \sigma(E_m(\mathcal{G}))$ up to a set of sets of measure zero. Consider $\mathcal{M}' = \Big(L^1\big(([0,1]^m, \mathscr{A}_m, \lambda^m), [0,1]\big), \mathbb{E}(\cdot \mid \mathscr{A}_1), \cdots, \mathbb{E}(\cdot \mid \mathscr{A}_{m-1})\Big)$. Then $\mathcal{M}' \models \mathsf{ARVE}_{\mathsf{T}'}$. Also $u, v \in \mathcal{M}'$ and $(\mathcal{M}', u) \equiv (\mathcal{M}', v)$. For each $1 \leq j \leq m-1$, one has $E_j(L^1(\sigma(u), [0,1])) \subseteq E_j(\mathcal{F})$ and $E_j(L^1(\sigma(v), [0,1])) \subseteq E_j(\mathcal{G})$. Hence, $\sigma(u)$ and $\sigma\big(E_j(L^1(\sigma(u), [0,1]))\big)$ for each $j = 1, \cdots, m-1$ are all atomic. Therefore by the induction step we see that $u$ is $d$-finite in $\mathcal{M}'$, so is $v$. Thus there is $\varphi \in \mathrm{Aut}_{\mathsf{ARVE}_{\mathsf{T}'}}(\mathcal{M}')$ extending $\phi$ such that $\varphi(u) = v$. Define the measure-preserving automorphism $\Phi\colon \mathscr{A}_m \to \mathscr{A}_m$ by $\Phi(A) = \varphi(\mathbf{1}_A)$ for all $A \in \mathscr{A}_m$. Then $\Phi^{-1}(\mathscr{A}_j) = \mathscr{A}_j$ for each $1 \leq j \leq m$.

Since $\mathrm{tp}^{\mathsf{ARVE}_{\mathsf{T}}}(f) = \mathrm{tp}^{\mathsf{ARVE}_{\mathsf{T}}}(g)$, we have $\mathrm{tp}^{\mathsf{ARV}}(f) = \mathrm{tp}^{\mathsf{ARV}}(g)$. By Theorem 5.2.3($v$), we have $\mathrm{dist}(f) = \mathrm{dist}(g)$. Therefore, the fact that $\sigma(f)$ is atomic implies that $\sigma(g)$ is atomic, and thus $f$ and $g$ are both discrete random variables. Moreover, by Fact 2.2.2, there is a maximal family $\{F_i \mid i \in I\}$ of pairwise inequivalent atoms of $\sigma(f)$, where $I$ is a countable set. Then for each $i \in I$, we have a sequence of $L_{\mathsf{RV}(\times)}$-terms $t_i^k$ such that $\mathbf{1}_{F_i} = \lim_{k\to\infty} t_i^k(f)$. Since $\mathrm{tp}^{\mathsf{ARV}(\times)}(f) = \mathrm{tp}^{\mathsf{ARV}(\times)}(g)$ and $\lim_{k\to\infty} t_i^k(f)^2 = \lim_{k\to\infty} t_i^k(f)$ for each $i \in I$, we have that $\lim_{k\to\infty} t_i^k(g)^2 = \lim_{k\to\infty} t_i^k(g)$, and thus $\lim_{k\to\infty} t_i^k(g) = \mathbf{1}_{G_i}$ for some $G_i \in \mathcal{L}^m$. Moreover, whenever $i \neq i' \in I$ we have $\mu(G_i \cap G_{i'}) = 0$, and for all $i \in I$ we have $\mu(F_i) = \mu(G_i)$ and $f(x) = g(y)$ for almost all $x \in F_i, y \in G_i$. Thus, $\sum_{i\in I} \mu(G_i) = \sum_{i\in I} \mu(F_i) = 1$. Then by the fact that $\mathrm{dist}(f) = \mathrm{dist}(g)$, we have that $G_i$ is an atom in $\sigma(g)$ for each $i \in I$. After adding the complement of $\bigcup_{i\in I} G_i$ to one of the atoms, we may assume that $\{G_i \mid i \in I\}$ is a maximal family of pairwise inequivalent atoms of $\sigma(g)$. Further, for all $i \in I$, by the definition of $\phi$ we have

$$\phi(E_m(\mathbf{1}_{F_i})) = \phi\big(E_m(\lim_{k\to\infty} t_i^k(f))\big) = \lim_{k\to\infty} \phi(E_m(t_i^k(f))) = \lim_{k\to\infty} E_m(t_i^k(g)) = E_m(\lim_{k\to\infty} t_i^k(g)) = E_m(\mathbf{1}_{G_i}).$$

Hence, $\varphi(E_m(\mathbf{1}_{F_i})) = \phi(E_m(\mathbf{1}_{F_i})) = E_m(\mathbf{1}_{G_i})$ for all $i \in I$. For all $i \in I$ and all $A \in \mathscr{A}_m$, we have

$$\mu(A \cap F_i) = \int_A \mathbf{1}_{F_i} d\mu = \int_A E_m(\mathbf{1}_{F_i}) d\mu = \int_{\Phi(A)} \varphi\big(E_m(\mathbf{1}_{F_i})\big) d\mu = \int_{\Phi(A)} E_m(\mathbf{1}_{G_i}) d\mu = \mu(\Phi(A) \cap G_i).$$

Since $\mathcal{L}^m$ is atomless over $\mathscr{A}_m$, by Lemma 2.2.20 we may assume there is a measure-preserving homomorphism $\Phi_i\colon \sigma(\mathscr{A}_m \cup \{F_i\}) \to \mathcal{L}^m$ which extends $\Phi$ such that $\Phi_i(F_i) = G_i$. By Proposition 2.2.3,

$$\mathscr{A}_m \vee \sigma(f) = \{\bigcup_{i\in I} A_i \cap F_i \mid A_i \in \mathscr{A}_m \text{ for each } i \in I\},$$

and

$$\mathscr{A}_m \vee \sigma(g) = \{\bigcup_{i\in I} A_i \cap G_i \mid A_i \in \mathscr{A}_m \text{ for each } i \in I\}.$$

We define $\Phi'\colon \mathscr{A}_m \vee \sigma(f) \to \mathscr{A}_m \vee \sigma(g)$ by $\Phi'(\bigcup_{i\in I} A_i \cap F_i) := \bigcup_{i\in I} \Phi_i(A_i \cap F_i) = \bigcup_{i\in I} \Phi(A_i) \cap G_i$, where $A_i \in \mathscr{A}_m$ for every $i \in I$. Because $\Phi_i$ is a measure-preserving homomorphism for each $i \in I$, we have that $\Phi'$ is a measure-preserving homomorphism from $\mathscr{A}_m \vee \sigma(f)$ to $\mathscr{A}_m \vee \sigma(g)$. Since $\Phi$ is surjective, we know that $\Phi'$ is also surjective. Then $\Phi'$ is a measure-preserving isomorphism that extends $\Phi$. Since $\sigma(f)$ is atomic, by Proposition 2.2.24 there is a measure-preserving automorphism $\tilde{\Phi}\colon \mathcal{L}^m \to \mathcal{L}^m$ extending $\Phi'$. Therefore, by Lemma 7.2.1 we know that $\tilde{\Phi}$ induces an $L_{\mathsf{RVE_T}}$-automorphism $\tilde{\varphi}\colon M \to M$ extending $\varphi$. Because for all $i \in I$ we have $\tilde{\Phi}(F_i) = G_i$ and $f(x) = g(y)$ for almost all $x \in F_i, y \in G_i$, we have $\tilde{\varphi}(f) = g$. □

When $T$ is infinite, $\mathsf{ARVE_T}$ is no longer separably categorical, by Theorem 7.2.3. Therefore, Theorem 3.5.2 can no longer be used to characterize the $d$-finite tuples in models of $\mathsf{ARVE_T}$.

**7.3.2 Question.** *Let $\mathcal{M} \models \mathsf{ARVE_T}$ and let $f = (f_1, \cdots, f_n) \in M^n$. Let $\mathcal{F}$ denote $L^1\big(\sigma(f), [0,1]\big)$, where $\sigma(f)$ is the $\sigma$-algebra generated by $f$. Is it true that $f$ is a d-finite tuple in $M$ if and only if the $\sigma$-algebras $\sigma(f), \{\sigma\big(E_t(\mathcal{F})\big)\}_{t\in T}$ are all atomic?*

# References


[1] Ben Yaacov, I., Schrödinger's cat, Israel J. Math. 153 (2006), 157–191.

[2] Ben Yaacov, I., On theories of random variables, submitted, arXiv:0901.1584.

[3] Ben Yaacov, I., Berenstein, A., and Henson, C. W., Almost indiscernible sequences and convergence of canonical bases, submitted, arXiv:0907.4508.

[4] Ben Yaacov, I., Berenstein, A., Henson, C. W., and Usvyatsov, A., Model theory for metric structures, in *Model Theory with Applications to Algebra and Analysis*, Volume 2, London Math. Society Lecture Note Series, 350, Cambridge University Press, 2008, 315–427.

[5] Ben Yaacov, I. and Usvyatsov, A., On $d$-finiteness in continuous structures, Fund. Math. 194 (2007), 67–88.

[6] Ben Yaacov, I. and Usvyatsov, A., Continuous first order logic and local stability, Trans. Amer. Math. Soc. 362 (2010), 5213–5259.

[7] Berenstein, A. and Henson, C. W., Model theory of probability spaces with an automorphism, submitted.

[8] Berkes, I. and Rosenthal, H. P., Almost exchangeable sequences of random variables, Z. Wahrsch. Verw. Gebiete 70 (1985), 473–507.

[9] Dall'Aglio, G., Sugli estremi dei momenti delle funzioni di ripartizione doppia (Italian), Ann. Scuoloa Norm. Sup. Pisa (3) 10 (1956), 35–74.

[10] Chang, C. C. and Keisler, H. J., *Model Theory*, Studies in Logic and Foundations of Mathematics, Volume 73, Third Edition, North-Holland, 1990.

[11] Dudley, R., *Probabilities and Metrics. Convergence of Laws on Metric Spaces, with a View to Statistical Testing*, Lecture Notes in Statistics, Volume 45, Matematisk Institut, Aarhus Universitet, 1976.

[12] Durrett, R., *Probability: Theory and Examples*, Second Edition, Duxbury Press, 1995.

[13] Fajardo, S. and Keisler, H. J., *Model Theory of Stochastic Processes*, Lecture Notes in Logic, 14, Association for Symbolic Logic, 2002.

[14] Fremlin, D. H., Measure algebras, in *Handbook of Boolean Algebras*, Volume 3, North-Holland, 1989, 877–980.

[15] Fremlin, D. H., *Measure algebras*, Volume 3 of *Measure Theory*, Torres Fremlin, 2003–04; for information see http://www.essex.ac.uk/maths/staff/fremlin/mt.htm.

[16] Givens, C. R. and Shortt, R. M., A class of Wasserstein metrics for probability distributions, Michigan Math. J. 31 (1984), 231–240.

[17] Hodges, W., *A Shorter Model Theory*, Cambridge University Press, 1997.

[18] Hoover, D., A characterization of adapted distribution, Ann. Probab. 15 (1987), 1600–1611.

[19] Hoover, D. and Keisler, H. J., Adapted probability distributions, Trans. Amer. Math. Soc. 286 (1984), 159–201.

[20] Kallenberg, O., *Foundations of Modern Probability*, Second Edition, Probability and its Applications, Springer-Verlag, 2002.

[21] Kantorovich, L. V., On the translocation of masses, Dokl. Akad. Nauk. USSR 37 (1942), 199–201. English translation in J. Math. Sci.( N.Y.) 133 (2006), 1381–1382.

[22] Keisler, H. J. and Sun, Y., Why saturated probability spaces are necessary, Adv. Math. 221 (2009), 1584–1607.

[23] Lieb, E. H. and Loss, M., *Analysis*, Second Edition, Graduate Studies in Mathematics, Volume 14, American Mathematical Society, 2001.

[24] Maharam, D., On homogeneous measure algebras, Proc. Natl. Acad. Sci. USA 28 (1942), 108–111.

[25] Marker, D., *Model Theory: an Introduction*, Graduate Texts in Mathematics, 217, Springer-Verlag, 2002.

[26] Monge, G., Mémoire sur la théorie des déblais et des remblais, in Histoire de l'Académie Royale des Sciences de Paris (1781), 666–704.

[27] Rüschendorf, L., Wasserstein-metric, in *Encyclopaedia of Mathematics*, Springer-Verlag, 2001. Available at http://eom.springer.de/W/w120020.htm.

[28] Rüschendorf, L., Monge-Kantorovich transportation problem and optimal couplings, Jahresber. Deutsch. Math.-Verein. 109 (2007), no.3, 113–137.

[29] Schoretsanitis, K., Fraïssé theory for metric structures, PhD Thesis, University of Illinois at Urbana-Champaign, 2007.

[30] Szulga, A., On the Wasserstein metric, Transactions of the Eighth Prague Conference on Information Theory, Statistical Decision Functions, Random Processes (Prague, 1978), Vol. B, 267–273, Reidel, 1978.

[31] Vallander, S., Calculations of the Vasserštein distance between probability distributions on the line (Russian), Teor. Verojatnost. i Primenen. 18 (1973), 824–827.

[32] Villani, C., *Optimal Transport. Old and New*, Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], 338. Springer-Verlag, 2009.

[33] Walters, P., *An Introduction to Ergodic Theory*, Graduate Texts in Mathematics, 79, Springer-Verlag, 1982.

# Vita

Song Shichang was born and grew up in Dalian, Liaoning Province, People's Republic of China. In Chinese culture, a Chinese name is written with the family name first and the given name next. Song prefers to be called by his family name Song if not being called by his full name. Song received a Bachelor of Science degree in Mathematics with a minor in Economics from Peking University in 2002, and a Master of Science degree in Mathematics from Peking University in 2004.