Inf. Comput. 310 (2026) 105441

Contents lists available at ScienceDirect

Information and Computation

journal homepage: www.elsevier.com/locate/yinco

A convergence law for continuous logic and continuous structures with finite domains

Vera Koponen

Department of Mathematics, Uppsala University, Sweden

ARTICLE INFO

Keywords:
Finite model theory
Continuous logic
Continuous model theory
Convergence law
Statistical relational artificial intelligence

ABSTRACT

We consider continuous relational structures with finite domain $[n] := \{1, \dots, n\}$ and a many valued logic, CLA, with values in the unit interval and which uses continuous connectives and continuous aggregation functions. CLA subsumes first-order logic on "conventional" finite structures. To each relation symbol R and identity constraint ic on a tuple the length of which matches the arity of R we associate a continuous probability density function $\mu_R^{ic} : [0,1] \to [0,\infty)$.

We also consider a probability distribution on the set $\mathbf{W}_n$ of continuous structures with domain $[n]$ which is such that for every relation symbol R, identity constraint ic, and tuple $\bar{a}$ satisfying ic, the distribution of the value of $R(\bar{a})$ is given by μ_R^{ic}, independently of the values for other relation symbols or other tuples.

In this setting we prove that every formula in CLA is asymptotically equivalent to a formula without any aggregation function. This is used to prove a convergence law for CLA which reads as follows for formulas without free variables: If $\varphi \in CLA$ has no free variable and $I \subseteq [0,1]$ is an interval, then there is $\alpha \in [0,1]$ such that, as n tends to infinity, the probability that the value of φ is in I tends to α.

1. Introduction

Logical convergence laws have been proved, or disproved, in many different contexts since the seminal work of Glebskii et al. [1] in the late1960′ies and, independently, Fagin [2]. The results have considered various kinds of structures (e.g. graphs, trees, relational structures, permutations), various logics (e.g. first-order logic and extensions of it, many valued logics) and various probability distributions (typically on a set of structures with a common finite domain). A far from complete list, only meant to illustrate some of the diversity in the results, includes the following studies: [3–15]. The notion of "almost sure (or asymptotic)" elimination of quantifiers or of aggregation functions is closely related to the notion of logical convergence law and indeed in several cases a logical convergence law is obtained as a consequence of almost sure (or asymptotic) elimination of quantifiers or of aggregation functions.

With few exceptions (mentioned below), the studies of logical convergence laws have considered "ordinary" structures in the sense that for every relation symbol R, say of arity ν, and any choice of elements $a_1, \dots, a_\nu$ from the domain of a structure, either $R(a_1, \dots, a_\nu)$ is true (has value 1) or false (has value 0). However, many phenomena are naturally described by continuously varying magnitudes. Therefore much of the methodology in data science, machine learning and artificial intelligence considers continuously varying data and continuous random variables. For example, neural networks typically have continuous real valued inputs and outputs, and the output often comes together with a probability distribution.

E-mail address: vera.koponen@math.uu.se

https://doi.org/10.1016/j.ic.2026.105441
Received 14 April 2025; Received in revised form 26 February 2026; Accepted 13 March 2026

This motivates making investigations about convergence laws for formal logics that “talk about” real valued relations (from the logician’s point of view) or random variables (from the probability theorist’s point of view) where each relation (random variable) comes together with a probability distribution. To deal with real valued relations we will use the concept of continuous structure that has been studied in continuous model theory for quite some time; see for example [16] or [17, Sections 2–4]. However, continuous model theory focuses on continuous structures with infinite domains because the main applications are in analysis and algebra, and here we consider finite domains.

Given a finite signature of relation symbols σ and a finite domain, which can as well be $[n] := \{1, \dots, n\}$ for some positive integer n, we can now let $\mathbf{W}_n$ be the (uncountably infinite) set of all continuous σ-structures (or “possible worlds”) with domain $[n]$ such that all relation symbols take values in the unit interval $[0, 1]$. To each relation symbol $R \in \sigma$, of arity ν_R say, and tuple $\bar{a} \in [n]^{\nu_R}$ we can associate a random variable $X_n^{R,\bar{a}} : \mathbf{W}_n \to [0, 1]$ such that for $\mathcal{A} \in \mathbf{W}_n$, $X_n^{R,\bar{a}}(\mathcal{A}) = R^{\mathcal{A}}(\bar{a})$, where $R^{\mathcal{A}}(\bar{a})$ is the value of $R(\bar{a})$ in $\mathcal{A}$. To each $X_n^{R,\bar{a}}$ we will associate a continuous probability density function $\mu_R^{ic} : [0, 1] \to [0, \infty)$ that depends only on R and the identity constraints that are satisfied by $\bar{a}$, represented by the notation ic, and such that μ_R^{ic} describes the distribution of the values of $X_n^{R,\bar{a}}$. We will then consider the product space of all $X_n^{R,\bar{a}}$ as R ranges over σ and $\bar{a}$ ranges over $[n]^{\nu_R}$. We let $\mathbb{P}_n$ denote the associated probability distribution on $\mathbf{W}_n$, which will have the property that for all $R \in \sigma$ and $\bar{a} \in [n]^{\nu_R}$, $X_n^{R,\bar{a}}$ is independent of all $X_n^{Q,\bar{a}}$ where $Q \neq R$ or $\bar{a} \neq \bar{b}$.

The logic that will be used, which will be called *Continuous Logic with Aggregation functions*, or CLA for short, is a many-valued logic with (truth) values in $[0, 1]$. CLA allows any continuous function $f : [0, 1]^k \to [0, 1]$ as a k-ary connective and it allows the use of any aggregation function $F : \bigcup_{k=1}^{\infty} [0, 1]^k \to [0, 1]$ that is continuous according to Definition 3. In this sense the aggregation functions maximum, minimum, and average are continuous, so CLA subsumes first-order logic on continuous structures with finite domain.

Challenges and results, informally. Let $\varphi(\bar{x})$ be a formula of CLA where $\bar{x} = (x_1, \dots, x_k)$ is a sequence of distinct free (object) variables. Also let $n \in \mathbb{N}^+$ (the set of positive integers), let $\bar{a} \in [n]^k$ and let $I \subseteq [0, 1]$. The question now is how to compute, or estimate, preferably in an efficient way, the probability that the value of $\varphi(\bar{a})$ in a random $\mathcal{A} \in \mathbf{W}_n$ is in I, and if this probability converges as n tends to infinity. Recall that the space $\mathbf{W}_n$ of continuous σ-structures with domain $[n]$ is (uncountably) infinite and each particular member of $\mathbf{W}_n$ has probability 0.

We will see that if a formula $\psi(\bar{x})$ (from CLA) does not use any aggregation function, then, for every $\bar{a} \in [n]^{|\bar{x}|}$, the probability that $\psi(\bar{a})$ belongs to I is equal to an integral that depends only on $\psi(\bar{x})$, I and the distributions μ_R^{ic} as R ranges over σ and ic ranges over the possible identity constraints that can be formed with ν_R variables. In particular, the probability does not depend on the domain size.

Our main result (Theorem 1) is that if $\varphi(\bar{x})$ is a formula in CLA and $ic(\bar{x})$ is an identity constraint, then there is a formula $\psi(\bar{x})$ without any aggregation function, such that as the domain size n tends to infinity, the probability that $\varphi(\bar{a})$ and $\psi(\bar{a})$ have approximately the same values for all $\bar{a} \in [n]^k$ that satisfy $ic(\bar{x})$ tends to 1. The proof shows how to “eliminate” aggregation functions from $\varphi(\bar{x})$, one by one, until one gets a formula $\psi(\bar{x})$ without any aggregation function which “almost surely” gives “almost the same” value. And as said above, the probability that the value of $\psi(\bar{a})$ belongs to an interval I can be computed without reference to the domain size. Thus we have a procedure for estimating the probability that the value of $\varphi(\bar{a})$ belongs to an interval I which depends only on $\varphi(\bar{x})$, I and the distributions of the form μ_R^{ic}; in particular, we have a convergence law for CLA.

A comparison with probabilistic logic programming. To see similarities and differences between the set-up of this article and a well-studied concept we make an informal comparison to *Probabilistic Logic Programming* (PLP) [18], a formalism used in for example *Statistical Relational Artificial Intelligence*, a subfield of Artificial Intelligence (AI) and Machine Learning which combines the logical and probabilistic approaches to AI; see e.g. [19–21].

Suppose that P, Q and R are relation symbols of arity 1. A probabilistic logic program (PL-program) consists partly of a set of “probabilistic facts” which assign probabilities to some relations (atomic formulas). The probabilistic facts could, for example, say that $P(x)$ and $Q(x)$ have probabilities α and β, respectively. This is interpreted as meaning that for every grounding of $P(x)$, say $P(a)$ where a is a member of some domain, the probability that $P(a)$ holds is α, independently of whether or not other ground atoms involving P or Q are true. The other part of a PL-program is a set of rules, for example $(P(x) \wedge Q(x)) \to R(x)$, which is implicitly understood to be universally quantified. So if the meaning of $P(x)$, $Q(x)$, and $R(x)$ is “x is wealthy”, “x is beautiful”, and “x is influencial”, respectively, then the program expresses that the probability that (a random) x is influencial is (at least) $\alpha \cdot \beta$.

Rather than to consider wealth, beauty, and influentiality as binary properties we can consider the degree, on a continuous scale, of wealth, beauty, and influence. We normalize the magnitudes so that for example the degree of wealth is always a number in the unit interval. Suppose that continuous probability density functions $\mu_P, \mu_Q : [0, 1] \to [0, \infty)$ are associated to P and Q, respectively. We interpret this as meaning that for any grounding $P(a)$ of $P(x)$ the value of $P(a)$ is distributed according to μ_P, independently of the values of other ground atoms involving P or Q. Suppose that we have found a continuous function $\mathsf{C} : [0, 1]^2 \to [0, 1]$ such that the composition $g(x) := \mathsf{C}(P(x), Q(x))$ gives a (more or less) reliable measure of the degree of influence of a random x. This means that C outputs the influence of an individual x if it gets the degrees of wealth and beauty of x as input. The expression $\mathsf{C}(P(x), Q(x))$ is a formula of CLA which can be seen as a “continuous rule” which outputs the degree of influence of x. In this context we do not need an explicit symbol R for “influence”.

To complicate things, whether x is influencial may not only depend on properties of x alone (wealth, beauty) but also on degrees of relationships that x has with others. Suppose that $E(x, y)$ expresses the degree to which y is exposed to x (e.g. in media) and that $f_E : [0, 1] \to [0, \infty)$ is a continuous probability density function, interpreted as saying that for every grounding $E(a, b)$, where a and b are different, the degree to which b is exposed to a is distributed according to f_E. Suppose that the “degree of influence of a” can be estimated (in a “continuous way”) by knowing only the degrees of wealth and beauty of a, and the average of the sequence of values of $E(a, b)$ as b ranges over other individuals of the domain. With CLA we can then express “the degree of influence of x” by a formula $\varphi(x)$ which uses average to aggregate over the domain.

Related work. It was mentioned above that almost all work on logical convergence laws have considered “conventional” structures with true/false (or 1/0) valued relations. One exception is [22] by Grädel et. al. where the authors prove logical convergence laws for first-order logic with semiring semantics, which partly means that grounded atomic formulas have values in a semiring and for every grounded atomic formula either itself or its negation must have the value 0. Another exception is [23] by Goldbring et. al. who prove an “approximate 0–1 law” for finite metric spaces in which nontrivial distances are in the interval $[\frac{1}{2}, 1]$. These metric spaces can be viewed as relational structures with a relation symbol of arity 2 that takes (truth) values in $[\frac{1}{2}, 1]$ for unordered pairs of distinct elements. Finally, the featured graph neural networks considered by Adam-Day et. al. in [24] and [25] can be viewed as graphs expanded by some unary relation symbols that can take any nonnegative real values, because the so-called node features can also be modeled by unary relation symbols that take nonnegative real values.

The last two references have a clear connection to machine learning and AI. Other studies of logical convergence phenomena in contexts where the formal logic or the probability distribution is inspired by concepts or methods in machine learning and AI include [26–35], in order of publication year.

Notation and terminology. We let $\mathbb{N}$ denote the set of all nonnegative integers and $\mathbb{N}^+$ the set of all positive integers. For all $n \in \mathbb{N}^+$, $[n] := \{1, \dots, n\}$. Logical (object) variables are denoted by x, y or z, sometimes with subscripts, and finite sequences of logical variables are denoted by $\bar{x}$. The length of $\bar{x}$ is denoted by $|\bar{x}|$. Real numbers will be denoted by p, q or r, sometimes with subscripts, and finite sequences of reals by $\bar{p}$, $\bar{q}$ or $\bar{r}$. For a set S, $|S|$ denotes its cardinality. Elements from some domain of a (continuous) structure are denoted by a, b, c, possibly by subscripts, and $\bar{a}$, $\bar{b}$ etc denote finite sequences of such elements. By $\mathrm{rng}(\bar{a})$ we denote the set of all elements in $\bar{a}$. If $\bar{a} = (a_1, \dots, a_k)$ then $\bar{a}b$ denotes the sequence $(a_1, \dots, a_k, b)$. Although this study is concerned with continuous structures with finite domain it is expected that the reader has some familiarity with “ordinary” (discrete) finite model theory as found in for example [40].

*For the rest of the article σ denotes a nonempty finite set of relation symbols (in the sense of first-order logic) which we call a **signature**. We assume that each $R \in \sigma$ has **arity** ν_R.*

2. Continuous structures and continuous logic with aggregation functions

In this section we define the logical concepts that we will work with: continuous σ-structures, continuous aggregation functions, and continuous logic with aggregation functions. We also prove some basic results about the concepts.

Definition 1. A ***continuous σ-structure*** $\mathcal{A}$ is determined by the following information:

1. A nonempty set A, which we call the ***domain*** of $\mathcal{A}$.
2. For every $R \in \sigma$ (with arity ν_R), a function $R^{\mathcal{A}} : A^{\nu_R} \to [0, 1]$ (so $R^{\mathcal{A}}(\bar{a}) \in [0, 1]$ for all $\bar{a} \in A^{\nu_R}$).

A continuous σ-structure is called ***finite*** it its domain is finite.

Definition 2. (a) Let $[0, 1]^{<\omega}$ denote the set of all finite nonempty sequences of reals from the unit interval $[0, 1]$, or equivalently, $[0, 1]^{<\omega} = \bigcup_{k=1}^{\infty} [0, 1]^k$.
(b) A function $F : [0, 1]^{<\omega} \to [0, 1]$ will be called an ***aggregation function*** if it is symmetric in the sense that if $(r_1, \dots, r_k) \in [0, 1]^k$ and $(q_1, \dots, q_k)$ is a permutation (reordering) of $(r_1, \dots, r_k)$ then $F(r_1, \dots, r_k) = F(q_1, \dots, q_k)$.

Example 1. For $\bar{p} = (p_1, \dots, p_n) \in [0, 1]^{<\omega}$, define

1. $\max(\bar{p})$ ($\min(\bar{p})$) to be the *maximum* (*minimum*) of $p_1, \dots, p_n$, and
2. $\mathrm{am}(\bar{p}) = (p_1 + \dots + p_n)/n$.

Thus ‘am’ is the *arithmetic mean*, or *average*. Definition 3 below defines the notion of “continuity” for aggregation functions. For an aggregation function $F : [0, 1]^{<\omega} \to [0, 1]$, the intuition behind the technical part (2) of Definition 3 is that, for every $\varepsilon > 0$, if $(r_1, \dots, r_n), (q_1, \dots, q_m) \in [0, 1]^{<\omega}$ are sufficiently long sequences that are sufficiently similar regarding the proportion of entries in sufficently small subintervals of $[0, 1]$ and if none of the sequences have too large gaps (if the entries are ordered from smaller to larger), then $|F(r_1, \dots, r_n) - F(q_1, \dots, q_m)| \le \varepsilon$.

Definition 3. Let $F : [0, 1]^{<\omega} \to [0, 1]$ be an aggregation function. We call F ***continuous*** if the following two conditions hold:

1. For every $\varepsilon > 0$ there is $\delta > 0$ such that for all n, if $(q_1, \dots, q_n)$, $(q'_1, \dots, q'_n) \in [0, 1]^n$ and $|q_i - q'_i| \le \delta$ for all $i = 1, \dots, n$, then $|F(q_1, \dots, q_n) - F(q'_1, \dots, q'_n)| \le \varepsilon$.
2. For every $\varepsilon > 0$ there are $\delta > 0$ and $M, N \in \mathbb{N}^+$ such that if $\alpha_0, \dots, \alpha_{M-1} \in [0, 1]$ and $(q_1, \dots, q_n), (q'_1, \dots, q'_m) \in [0, 1]^{<\omega}$ (where we may have $n \neq m$) are such that conditions (a)–(d) below hold, then $|F(q_1, \dots, q_n) - F(q'_1, \dots, q'_m)| \le \varepsilon$:
   (a) $n, m \ge N$,

(b) for all $i = 0, \dots, M-1$, if $\alpha_i > 0$ then $\alpha_i > \delta$, and (regardless of whether $\alpha_i > 0$ or not)

$$\frac{\left|\left\{l : q_l \in \left[\frac{i}{M}, \frac{i+1}{M}\right]\right\}\right|}{n} \in (\alpha_i - \delta, \alpha_i + \delta),$$

$$\frac{\left|\left\{l : q'_l \in \left[\frac{i}{M}, \frac{i+1}{M}\right]\right\}\right|}{m} \in (\alpha_i - \delta, \alpha_i + \delta),$$

(c) if $i < j < k$ or $i > j > k$, and in addition $\alpha_i > 0$ and $\alpha_j = 0$, then $\alpha_k = 0$, and

(d) if $i < j < k$ or $i > j > k$, and in addition $\alpha_i = \alpha_j = 0$ and $\alpha_k > 0$, then

$$\{l : q_l \in \left[\tfrac{i}{M}, \tfrac{i+1}{M}\right]\} = \{l : q'_l \in \left[\tfrac{i}{M}, \tfrac{i+1}{M}\right]\} = \emptyset,$$

Lemma 1. *The aggregation functions maximum, minimum and arithmetic mean are continuous.*

Proof. We start with maximum. It is immediately clear that max satisfies condition (1) of Definition 3 so we only verify that (2) of the same definition holds. Let $\varepsilon > 0$. Suppose that conditions (a)–(d) of part (2) of Definition 3 are satisfied by $\delta > 0$, $M, N \in \mathbb{N}^+$, $\alpha_0, \dots, \alpha_{M-1} \in [0, 1]$, and $(q_1, \dots, q_n), (q'_1, \dots, q'_m)$. For $i = 0, \dots, M-1$, let $P_i = \left\{l : q_l \in \left[\frac{i}{M}, \frac{i+1}{M}\right]\right\}$ and $P'_i = \left\{l : q'_l \in \left[\frac{i}{M}, \frac{i+1}{M}\right]\right\}$. Let s be maximal such that $P_s \neq \emptyset$. Then $s/M \leq \max(\bar{q}) \leq (s+1)/M$.

Suppose for a contradiction that $i > s$ and $\alpha_i > 0$. By assumption (2)(a) (of Definition 3) we get $\alpha_i > \delta$, so $P_i \neq \emptyset$ by assumption (2)(b) and this contradicts the choice of s. We conclude that $\alpha_i = 0$ for all $i = s+1, \dots, M-1$.

Assuming that $\delta > 0$ is small enough we have $\alpha_0 + \dots + \alpha_{M-1} \approx 1$ so there is $i \leq s$ such that $\alpha_i > 0$. Let t be maximal such that $\alpha_t > 0$; hence $t \leq s$. By (2)(d) we have $P'_i = \emptyset$ for all $i > s+1$. Hence $\max(\bar{q}') \leq (s+2)/M$.

Suppose, for a contradiction, that $P'_{s-1} = \emptyset$. Then $\alpha_{s-1} = 0$, because $\alpha_{s-1} > 0$ would imply $\alpha_{s-1} > \delta$ and hence $P'_{s-1} \neq \emptyset$. If $\alpha_s > 0$ then (2)(c) implies that $\alpha_i = 0$ for all $i < s$ which contradicts that $t \leq s$ and $\alpha_t > 0$. Thus $\alpha_{s-1} = \alpha_s = 0$, $\alpha_t > 0$, and $t < s-1 < s$, so (2)(d) implies that $P_s = \emptyset$ which contradicts the choice of s.

Hence we conclude that $P'_{s-1} \neq \emptyset$ and we get $(s-1)/M \leq \max(\bar{q}') \leq (s+2)/M$. It follows that $|\max(\bar{q}) - \max(\bar{q}')| \leq 2/M$. So if $\delta > 0$ is chosen small enough, M and N large enough, and conditions (2)(a)–(2)(d) hold, then $|\max(\bar{q}) - \max(\bar{q}')| \leq \varepsilon$.

The proof that minimum is continuous is similar, so we next consider the arithmetic mean (average), abbreviated 'am'. It is again clear that am has property (1) of Definition 3, so we only verify property (2). Suppose that $\bar{q} = (q_1, \dots, q_n)$ and $\bar{q}' = (q'_1, \dots, q'_m)$ satisfy conditions (a)–(d) of part (2) of Definition 3. Then

$$\sum_{i=0}^{M-1} \frac{i}{M}(\alpha_i - \delta) \leq \mathrm{am}(\bar{q}) \leq \sum_{i=0}^{M-1} \frac{i+1}{M}(\alpha_i + \delta)$$

and similarly if $\bar{q}$ is replaced by $\bar{q}'$. Hence $|\mathrm{am}(\bar{q}) - \mathrm{am}(\bar{q}')| \leq \sum_{i=0}^{M-1} \frac{\alpha_i + 2i\delta + \delta}{M}$, and if δ is small enough (relative to M) then this is approximately $\sum_{i=0}^{M-1} \frac{\alpha_i}{M}$. For sufficiently small δ, the sum $\alpha_0 + \dots + \alpha_{M-1}$ is approximately 1, so $\sum_{i=0}^{M-1} \frac{\alpha_i}{M} \approx \frac{1}{M}$ which can be made as small as we like if M is chosen large enough. ☐

One may note that in the context of true/false-valued relations studied in [33] the aggregation functions maximum and minimum are not continuous (also called "strongly admissible" in [33]) according to the definitions used there (but only "semi-continuous", or "admissible").

Definition 4. **(Syntax of CLA)** We define CLA, the set of formulas of ***continuous logic with aggregation functions***, as follows, where for every $\varphi \in CLA$ we simultaneously define the set of its free variables, denoted $Fv(\varphi)$:

1. For each $c \in [0, 1]$, $c \in CLA$ and $Fv(c) = \emptyset$.
2. For all logical variables x and y, '$x = y$' belongs to CLA and $Fv(x = y) = \{x, y\}$.
3. For every $R \in \sigma$ (of arity ν_R) and any choice of logical variables $x_1, \dots, x_{\nu_R}$, $R(x_1, \dots, x_{\nu_R})$ belongs to CLA and $Fv(R(x_1, \dots, x_{\nu_R})) = \{x_1, \dots, x_{\nu_R}\}$.
4. If $k \in \mathbb{N}^+$, $\varphi_1, \dots, \varphi_k \in CLA$ and $\mathsf{C} : [0, 1]^k \to [0, 1]$ is a continuous function, then the expression $\mathsf{C}(\varphi_1, \dots, \varphi_k)$ is a formula of CLA and its set of free variables is $Fv(\varphi_1) \cup \dots \cup Fv(\varphi_k)$.
5. Suppose that $\varphi \in CLA$, y is a logical variable, and that $F : [0, 1]^{<\omega} \to [0, 1]$ is a continuous aggregation function. Then the expression $F(\varphi : y)$ is a formula of CLA and its set of free variables is $Fv(\varphi) \setminus \{y\}$. Hence this construction binds the variable y.

When denoting a formula (of CLA) by $\varphi(\bar{x})$ (or $\psi(\bar{x})$ etc), *it will be assumed that all free variables of $\varphi(\bar{x})$ appear in the sequence $\bar{x}$* (but we do not insist that every variable in $\bar{x}$ appears in $\varphi(\bar{x})$). Formulas of the forms 1, 2 and 3 in Definition 4 will be called ***atomic***.

Definition 5. **(Semantics of CLA)** For every $\varphi \in CLA$ and every sequence of distinct logical variables $\bar{x} = (x_1, \dots, x_k)$ such that $Fv(\varphi) \subseteq \{x_1, \dots, x_k\}$ we associate a mapping from pairs $(\mathcal{A}, \bar{a})$, where $\mathcal{A}$ is a finite continuous σ-structure and $\bar{a} = (a_1, \dots, a_k) \in A^k$, to $[0, 1]$. The number in $[0, 1]$ to which $(\mathcal{A}, \bar{a})$ is mapped is denoted $\mathcal{A}(\varphi(\bar{a}))$ and is defined by induction on the complexity of formulas, as follows:

1. If $\varphi(\bar{x})$ is a constant c from $[0, 1]$, then $\mathcal{A}(\varphi(\bar{a})) = c$.
2. If $\varphi(\bar{x})$ has the form $x_i = x_j$, then $\mathcal{A}(\varphi(\bar{a})) = 1$ if $a_i = a_j$, and otherwise $\mathcal{A}(\varphi(\bar{a})) = 0$.

3. For every $R \in \sigma$, if $\varphi(\bar{x})$ has the form $R(x_{i_1}, \dots, x_{i_{\nu_R}})$, then $\mathcal{A}(\varphi(\bar{a})) = R^{\mathcal{A}}(a_{i_1}, \dots, a_{i_{\nu_R}})$.
4. If $\varphi(\bar{x})$ has the form $\mathsf{C}(\varphi_1(\bar{x}), \dots, \varphi_k(\bar{x}))$, where $\mathsf{C} : [0,1]^k \to [0,1]$ is continuous, then
$$\mathcal{A}\big(\varphi(\bar{a})\big) \ = \ \mathsf{C}\big(\mathcal{A}(\varphi_1(\bar{a})), \dots, \mathcal{A}(\varphi_k(\bar{a}))\big).$$
5. Suppose that $\varphi(\bar{x})$ has the form $F(\psi(\bar{x}, y) : y)$ where y is a logical variable that does not occur in $\bar{x}$ and $F : [0,1]^{<\omega} \to [0,1]$ is a continuous aggregation function. Then
$$\mathcal{A}\big(\varphi(\bar{a})\big) = F(r_1, \dots, r_n)$$
where, for $i = 1, \dots, n$, $r_i = \mathcal{A}\big(\psi(\bar{a}, b_i)\big)$ and $b_1, \dots, b_n$ is an enumeration (without repetition) of all elements in $A \setminus \mathrm{rng}(\bar{a})$ where A is the domain of $\mathcal{A}$.

Remark 1. (i) Note that we have $\mathcal{A}(R(\bar{a})) = R^{\mathcal{A}}(\bar{a})$ for every $R \in \sigma$, every continuous σ-structure $\mathcal{A}$, and every $\bar{a} \in A^{\nu_R}$.

(ii) The classical connectives $\neg$, $\wedge$, $\vee$ and $\to$, which can be seen as functions from $\{0,1\}$ or $\{0,1\}^2$ to $\{0,1\}$, can be extended, for example by Lukasiewicz semantics [36], to continuous functions from $[0,1]$ or $[0,1]^2$ to $[0,1]$, so these connectives can be expressed with CLA. Likewise, the aggregation functions max and min can be used to express existential and universal quantifications, so CLA subsumes first-order logic on "ordinary" σ-structures $\mathcal{A}$ in which $R^{\mathcal{A}}(\bar{a})$ is either 0 or 1 for all $R \in \sigma$ and $\bar{a} \in A^{\nu_R}$. This may not be immediately clear because according to the semantics of CLA, the construction $F(\varphi(x_1, \dots, x_k, y) : y)$ only aggregates only over y that are different from all $x_1, \dots, x_k$. But the first-order construction $\exists y \varphi(x_1, \dots, x_k, y)$ can be expressed in CLA by
$$\max\nolimits_{k+1}\big(\varphi(x_1, \dots, x_k, x_1), \dots, \varphi(x_1, \dots, x_k, x_k), \max(\varphi(x_1, \dots, x_k, y) : y)\big)$$
where $\max_{k+1} : [0,1]^{k+1} \to [0,1]$ is the continuous function which outputs the largest entry in a $(k+1)$-tuple.

Definition 6. (a) A formula of CLA is called ***aggregation-free*** if it contains no aggregation function, that is, if it was constructed by only using parts 1 – 4 of Definition 4.
(b) The formulas $\varphi(\bar{x}), \psi(\bar{x}) \in CLA$ are ***equivalent*** if for every finite continuous σ-structure $\mathcal{A}$ and every $\bar{a} \in A^{|\bar{x}|}$, $\mathcal{A}(\varphi(\bar{a})) = \mathcal{A}(\psi(\bar{a}))$.

The next results allows us to simplify aggregation-free formulas to such which use exactly one connective (i.e. continuous function $\mathsf{C} : [0,1]^k \to [0,1]$ for some k), and exactly once.

Lemma 2. *If $\psi(\bar{x}) \in CLA$ is aggregation-free then, for some m, it is equivalent to a formula $\mathsf{C}(\psi_1(\bar{x}), \dots, \psi_m(\bar{x}))$ where $\mathsf{C} : [0,1]^m \to [0,1]$ is continuous, for all i, $\psi_i(\bar{x})$ is atomic, and if $i \neq j$ then $\psi_i(\bar{x})$ and $\psi_j(\bar{x})$ are different formulas.*

Proof. We use induction on the complexity of aggregation-free formulas. Suppose that $\psi(\bar{x})$ is aggregation-free. If $\psi(\bar{x})$ is atomic then it is equivalent to $\mathsf{C}(\psi(\bar{x}))$ where $\mathsf{C} : [0,1] \to [0,1]$ is the identity function.

Now suppose that $\psi(\bar{x})$ has the form $\mathsf{D}(\psi_1(\bar{x}), \dots, \psi_t(\bar{x}))$ where $\mathsf{D} : [0,1]^t \to [0,1]$ is continuous. By the induction hypothesis, for every $i = 1, \dots, t$, $\psi_i(\bar{x})$ is equivalent to a formula of the form $\mathsf{D}_i(\psi_{i,1}(\bar{x}), \dots, \psi_{i,s_i}(\bar{x}))$ where each $\psi_{i,j}(\bar{x})$ is atomic and $\mathsf{D}_i : [0,1]^{s_i} \to [0,1]$ is continuous. Now let $m = s_1 + \dots + s_t$ and let $\mathsf{C} : [0,1]^m \to [0,1]$ be the composition defined by
$$\mathsf{C}(r_1, \dots, r_m) = \mathsf{D}(\mathsf{D}_1(r_1, \dots, r_{s_1}), \dots, \mathsf{D}_t(r_{m-s_t}, \dots, r_m)),$$
so C is continuous. Then $\psi(\bar{x})$ is equivalent to
$$\mathsf{C}(\psi_{1,1}(\bar{x}), \dots, \psi_{1,s_1}(\bar{x}), \dots, \psi_{t,1}(\bar{x}), \dots, \psi_{t,s_t}(\bar{x})).$$
To simplify notation let us rename the formulas $\psi_{i,j}$ so that $\psi(\bar{x})$ is equivalent to $\mathsf{C}(\psi_1(\bar{x}), \dots, \psi_m(\bar{x}))$. Suppose that for $1 \le i < j \le m$, $\psi_i(\bar{x})$ is the same formula as $\psi_j(\bar{x})$. For notational simplicity and without loss of generality, suppose that $j = m$. Then define $\mathsf{C}' : [0,1]^{m-1} \to [0,1]$ by $\mathsf{C}'(r_1, \dots, r_{m-1}) = \mathsf{C}(r_1, \dots, r_{m-1}, r_{m-1})$. Then C' is continuous and $\mathsf{C}'(\psi_1(\bar{x}), \dots, \psi_{m-1}(\bar{x}))$ is equivalent to $\mathsf{C}(\psi_1(\bar{x}), \dots, \psi_m(\bar{x}))$. By continuing in this way as long as necessary we can make sure to get a formula like $\mathsf{C}(\psi_1(\bar{x}), \dots, \psi_m(\bar{x}))$ that satisfies the additional condition that $\psi_i(\bar{x})$ is different from $\psi_j(\bar{x})$ if $i \neq j$. ☐

The idea with an identity constraint, that we define next, is that it specifies which pairs of elements in a tuple/sequence are equal and which are not.

Definition 7. Let $\bar{x} = (x_1, \dots, x_k)$ be a sequence of (not necessarily distinct) variables.
(i) Let $K = \{(i,j) : 1 \le i < j \le k\}$. An ***identity constraint for*** $\bar{x}$ is a first-order formula of the following form for some $I \subseteq K$: $\bigwedge_{(i,j) \in I} x_i = x_j \ \wedge \ \bigwedge_{(i,j) \in K \setminus I} x_i \neq x_j$.
(ii) Let $ic(\bar{x})$ be an identity constraint for $\bar{x}$. Suppose that $\bar{a} = (a_1, \dots, a_k)$ where $a_1, \dots, a_k$ may be any objects. We say that $\bar{a}$ ***satisfies*** $ic(\bar{x})$ if for all $1 \le i < j \le k$, $a_i = a_j$ if and only if $ic(\bar{x}) \models x_i = x_j$, where '$\models$' has the usual meaning for first-order logic.
(iii) Suppose that $ic(\bar{x})$ is an identity constraint for $\bar{x}$ and let $\bar{x}'$ be a sequence of variables such that $\mathrm{rng}(\bar{x}') \subseteq \mathrm{rng}(\bar{x})$. The ***restriction of*** $ic(\bar{x})$ ***to*** $\bar{x}'$ is an identity constraint $ic'(\bar{x}')$ for $\bar{x}'$ such that $ic(\bar{x}) \wedge ic'(\bar{x}')$ is consistent (or equivalently, such that $ic'(\bar{x}')$ is a logical consequence of $ic(\bar{x})$). Note that, syntactically, the restriction of $ic(\bar{x})$ to $\bar{x}'$ is not unique, but it is unique up to logical equivalence, so as *the* restriction we can just choose one of the logically equivalent candidates.

Observe that an identity constraint $ic(\bar{x})$ may be inconsistent (as it may violate the transitivity of '$=$', or the reflexivity of '$=$' if x_i and x_j are the same variable for some $i \neq j$) but every result that will follow and mentions an identity constraint is vacuously true if the identity constraint is inconsistent. To allow repetitions in the sequence $\bar{x}$ may seem odd at this point but it will be technically convenient later. Informally, the next lemma says that for every aggregation-free formula $\varphi(\bar{x})$ and identity constraint $ic(\bar{x})$ there is an aggregation-free formula $\psi(\bar{x})$ which uses only atomic formulas of the form (3) in Definition 4 and such that, conditioned on $\bar{a}$ satisfying $ic(\bar{x})$, $\varphi(\bar{a})$ and $\psi(\bar{a})$ have he same value.

Lemma 3. *Let $\bar{x} = (x_1, \dots, x_k)$ be a sequence of distinct variables, let $ic(\bar{x})$ be an identity constraint for $\bar{x}$, and let $\varphi(\bar{x})$ be an aggregation-free formula. Then there is an aggregation-free formula $\psi(\bar{x})$ such that*

1. *if $\mathcal{A}$ is a finite continuous σ-structure and $\bar{a} \in A^k$ satisfies $ic(\bar{x})$, then $\mathcal{A}(\varphi(\bar{a})) = \mathcal{A}(\psi(\bar{a}))$, and*
2. *$\psi(\bar{x})$ is either a constant (i.e. of the form (1) in Definition 4) or $\psi(\bar{x})$ has the form $\mathsf{C}(R_1(\bar{x}_1), \dots, R_m(\bar{x}_m))$ where $\mathsf{C} : [0,1]^m \to [0,1]$ is continuous, for each $i = 1, \dots, m$, $R_i \in \sigma$, $\mathrm{rng}(\bar{x}_i) \subseteq \mathrm{rng}(\bar{x})$, $|\bar{x}_i| = \nu_{R_i}$, and if $i \neq j$ then $R_i \neq R_j$ or $\bar{x}_i \neq \bar{x}_j$.*

Proof. Let $\bar{x} = (x_1, \dots, x_k)$ be a sequence of distinct variables, let $ic(\bar{x})$ be an identity constraint for $\bar{x}$, and let $\varphi(\bar{x})$ be an aggregation-free formula. By Lemma 2 we may assume that $\varphi(\bar{x})$ has the form

$$\mathsf{D}(c_1, \dots, c_t, \theta_{t+1}(\bar{x}_{t+1}), \dots, \theta_{t+s}(\bar{x}_{t+s}), R_{t+s+1}(\bar{x}_{t+s+1}), \dots, R_{t+s+u}(\bar{x}_{t+s+u}))$$

where $\mathsf{D} : [0,1]^{t+s+u} \to [0,1]$ is continuous, each c_i is a constant, each $\theta_i(\bar{x}_i)$ has the form $x_j = x_l$ for some distinct $j, l \in \{1, \dots, k\}$ (if $j = l$ we can replace $x_j = x_l$ by 1), $R_i \in \sigma$, $\mathrm{rng}(\bar{x}_i) \subseteq \mathrm{rng}(\bar{x})$, and $|\bar{x}_i| = \nu_{R_i}$. Moreover, by the same lemma, we may assume that if $t + s < i < j \leq t + s + u$ then $R_i \neq R_j$ or $\bar{x}_i \neq \bar{x}_j$.

Let $t + 1 \leq i \leq t + s$ and suppose that θ_i is the formula $x_j = x_l$ where $j \neq l$. Then let $c_i = 1$ if $ic(\bar{x}) \models x_j = x_l$ and otherwise let $c_i = 0$. It follows that if $u = 0$ then $\varphi(\bar{x})$ is $\mathsf{D}(c_1, \dots, c_t, \theta_{t+1}(\bar{x}_{t+1}), \dots, \theta_{t+s}(\bar{x}_{t+s}))$ and if $c := \mathsf{D}(c_1, \dots, c_{t+s})$, then for every finite continuous σ-structure $\mathcal{A}$ and every $\bar{a} \in A^k$ that satisfies $ic(\bar{x})$ we have $\mathcal{A}(\varphi(\bar{a})) = c = \mathcal{A}(c)$.

Now suppose that $u > 0$ and define $\mathsf{C} : [0,1]^u \to [0,1]$ by

$$\mathsf{C}(r_1, \dots, r_u) = \mathsf{D}(c_1, \dots, c_{t+s}, r_1, \dots, r_u).$$

Then C is continuous and for all finite continuous σ-structures $\mathcal{A}$ and all $\bar{a} \in [n]^k$ that satisfy $ic(\bar{x})$, $\mathcal{A}(\varphi(\bar{a})) = \mathcal{A}(\mathsf{C}(R_{t+s+1}(\bar{a}_{t+s+1}), \dots, R_{t+s+u}(\bar{a}_{t+s+u})))$ where $\bar{a}_i$ is the subsequence of $\bar{a}$ that corresponds to $\bar{x}_i$. □

3. Probability distributions

In this section we define the probability theoretic concepts of this study.

Definition 8. Let $n \in \mathbb{N}^+$.
(a) $\mathbf{W}_n$ denotes the set of all continuous σ-structures with domain $[n] = \{1, \dots, n\}$.
(b) $\xi_n := \prod_{R \in \sigma} n^{\nu_R}$.
(c) $\Xi_n := \{(R, \bar{a}) : R \in \sigma \text{ and } \bar{a} \in [n]^{\nu_R}\}$ (so $\xi_n = |\Xi_n|$).
(d) Let $\prec_n$ be a linear order on Ξ_n such that if $n < m$, then $\prec_n$ is the restriction of $\prec_m$ to Ξ_n.
(e) Define $\mathfrak{s}_n : \mathbf{W}_n \to [0,1]^{\xi_n}$ by, for all $\mathcal{A} \in \mathbf{W}_n$, letting $\mathfrak{s}_n(\mathcal{A}) = (r_1, \dots, r_{\xi_n})$ where, for all $i = 1, \dots, \xi_n$, $r_i = R^{\mathcal{A}}(\bar{a})$ where $(R, \bar{a})$ is the i^{th} element of Ξ_n with respect to the order $\prec_n$.

From the above definition and the definition of continuous σ-structure we immediately get:

Lemma 4. *For all $n \in \mathbb{N}^+$, $\mathfrak{s}_n : \mathbf{W}_n \to [0,1]^{\xi_n}$ is bijective.*

Thus we can identify $\mathbf{W}_n$ with $[0,1]^{\xi_n}$ where the latter is a compact metric space. The next couple of definitions show how we can view $\mathbf{W}_n$ as a probability space. The intuition is that for each $R \in \sigma$ and each identity constraint $ic(x_1, \dots, x_{\nu_R})$, if $(a_1, \dots, a_{\nu_R}) \in [n]^{\nu_R}$ satisfies $ic(x_1, \dots, x_{\nu_R})$ then the distribution of the random variable $X_n^{R,\bar{a}} : \mathbf{W}_n \to [0,1]$ defined by $X_n^{R,\bar{a}}(\mathcal{A}) = R^{\mathcal{A}}(\bar{a})$ for all $\mathcal{A} \in \mathbf{W}_n$ is given by a continuous probability density function $\mu_R^{ic} : [0,1] \to [0,\infty)$ which depends only on R and ic, where we can freely choose μ_R^{ic} as long as it is continuous and $\int_{[0,1]} \mu_R^{ic}(x)dx = 1$. So if $I \subseteq [0,1]$ is (Lebesgue) measurable then the probability that $X_n^{R,\bar{a}}$ belongs to I equals $\int_I \mu_R^{ic}(x)dx$. Moreover, we want this probability to be independent of the value of $X_n^{Q,\bar{b}}$ if $Q \neq R$ or if $\bar{b} \neq \bar{a}$.

Definition 9. Let $n \in \mathbb{N}^+$.
(a) To every pair (R, ic), where $R \in \sigma$ and $ic(x_1, \dots, x_{\nu_R})$ is an identity constraint for $(x_1, \dots, x_{\nu_R})$, we associate a continuous probability density function (also called probability distribution) $\mu_R^{ic} : [0,1] \to [0,\infty)$.
(b) For all $\mathcal{A} \in \mathbf{W}_n$ define

$$\pi_n(\mathcal{A}) = \prod_{R \in \sigma} \prod_{\substack{ic \text{ is an} \\ \text{identity constraint} \\ \text{for } (x_1,\dots,x_{\nu_R})}} \prod_{\substack{\bar{a} \in [n]^{\nu_R} \text{ and} \\ \bar{a} \text{ satisfies } ic}} \mu_R^{ic}(R^{\mathcal{A}}(\bar{a})).$$

(c) For every $i = 1, \dots, \xi_n$ we define $\lambda_i : [0,1] \to [0,\infty)$ by choosing the i^{th} pair $(R, \bar{a})$ in Ξ_n according to the order $\prec_n$ and, for every $r \in [0,1]$, letting $\lambda_i(r) = \mu_R^{ic}(r)$ where ic is the identity constraint for sequences of length ν_R that $\bar{a}$ satisfies.
(d) For every $(r_1, \dots, r_{\xi_n}) \in [0,1]^{\xi_n}$ define $\pi_n^*(r_1, \dots, r_{\xi_n}) = \lambda_1(r_1) \cdot \ldots \cdot \lambda_{\xi_n}(r_{\xi_n})$.

Lemma 5. *Let $n \in \mathbb{N}^+$.*
(a) For each $i \in [\xi_n]$, λ_i is continuous and $\int_{[0,1]} \lambda_i(r)dr = 1$.
(b) π_n^* is a continuous function from $[0,1]^{\xi_n}$ to $[0,1]$ and

$$\int_{[0,1]^{\xi_n}} \pi_n^*(r_1, \dots, r_{\xi_n})dr_1 \dots dr_{\xi_n} = 1,$$

so π_n^* is a probability density function on $[0,1]^{\xi_n}$.
(c) For all $\mathcal{A} \in \mathbf{W}_n$, $\pi_n(\mathcal{A}) = \pi_n^*(\mathfrak{s}_n(\mathcal{A}))$.

Proof. (a) For each $i \in [\xi_n]$ there are $R \in \sigma$ and an identity constraint ic such that $\lambda_i = \mu_R^{ic}$ where $\mu_R^{ic} : [0,1] \to [0,\infty)$ is a probability density function and hence $\int_{[0,1]} \mu_R^{ic}(r)dr = 1$.

(b) We have

$$\int_{[0,1]^{\xi_n}} \pi_n^*(r_1,\dots,r_{\xi_n})dr_1 \dots dr_{\xi_n} = \int_{[0,1]^{\xi_n}} \lambda_1(r_1)\cdot \ldots \cdot \lambda_{\xi_n}(r_{\xi_n})dr_1 \dots dr_{\xi_n} =$$
$$\left(\int_{[0,1]} \lambda_1(r_1)dr_1\right)\cdot \ldots \cdot \left(\int_{[0,1]} \lambda_{\xi_n}(r_{\xi_n})dr_{\xi_n}\right) = 1\cdot \ldots \cdot 1 = 1.$$

Part (c) follows directly from the definitions. □

Definition 10. Let $n \in \mathbb{N}^+$.
(a) For every (Lebesgue) measurable $X \subseteq [0,1]^{\xi_n}$, let

$$\mathbb{P}_n^*(X) = \int_X \pi_n^*(x_1,\dots,x_{\xi_n})dx_1 \dots dx_{\xi_n}.$$

(b) We call $\mathbf{X} \subseteq \mathbf{W}_n$ ***measurable*** if $\mathfrak{s}_n(\mathbf{X}) := \{\mathfrak{s}_n(\mathcal{A}) : \mathcal{A} \in \mathbf{X}\}$ is a measurable subset of $[0,1]^{\xi_n}$.
(c) For every measurable $\mathbf{X} \subseteq \mathbf{W}_n$ define $\mathbb{P}_n(\mathbf{X}) = \mathbb{P}_n^*(\mathfrak{s}_n(\mathbf{X}))$.

We now make sure that certain natural subsets of $\mathbf{W}_n$ are measurable.

Lemma 6. *Let $n \in \mathbb{N}^+$, $\varphi(\bar{x}) \in CLA$, and $\bar{a} \in [n]^{|\bar{x}|}$.*
(a) Define $f : [0,1]^{\xi_n} \to [0,1]$ by $f(\bar{r}) = \mathfrak{s}_n^{-1}(\bar{r})(\varphi(\bar{a}))$, so $f(\bar{r}) = \mathcal{A}(\varphi(\bar{a}))$ where $\mathcal{A} \in \mathbf{W}_n$ is such that $\mathfrak{s}_n(\mathcal{A}) = \bar{r}$. Then f is continuous.
(b) If $I \subseteq [0,1]$ is an interval then $\{\mathcal{A} \in \mathbf{W}_n : \mathcal{A}(\varphi(\bar{a})) \in I\}$ is a measurable subset of $\mathbf{W}_n$.

Proof. We begin by showing that (b) follows from (a). By Definition 10, to prove (b) it suffices to prove that $X \subseteq [0,1]^{\xi_n}$, where $X := \{\mathfrak{s}_n(\mathcal{A}) : \mathcal{A} \in \mathbf{W}_n$ and $\mathcal{A}(\varphi(\bar{a})) \in I\}$, is measurable. By part (a), f as defined in that part is continuous, hence a measurable function. Since every interval is measurable and we have $X = f^{-1}(I)$ it follows that X is measurable. Hence it remains to prove (a).

Let f be defined as in part (a). We use induction on the complexity of $\varphi(\bar{x})$. The base case concerns formulas according to cases (1)–(3) of Definition 4.

Suppose that $\varphi(\bar{x})$ is a constant $c \in [0,1]$. Then $f(\bar{r}) = c$ for all $\bar{r} \in [0,1]^{\xi_n}$ so f is continuous.

Suppose that $\varphi(\bar{x})$ has the form $x_i = x_j$ for some x_i and x_j in $\bar{x}$. If x_i and x_j are the same variable then $f(\bar{r}) = 1$ for all $\bar{r} \in [0,1]^{\xi_n}$ so f is continuous. Now suppose that x_i and x_j are different variables. If $a_i = a_j$ (where a_i and a_j are the i^{th} respectively j^{th} entries in $\bar{a}$) then $f(\bar{r}) = 1$ for all $\bar{r} \in [0,1]^{\xi_n}$; otherwise $f(\bar{r}) = 0$ for all $\bar{r} \in [0,1]^{\xi_n}$. In either case, f is continuous.

Suppose that $\varphi(\bar{x})$ has the form $R(x_{i_1},\dots,x_{i_{\nu_R}})$ for some $R \in \sigma$ and $x_{i_1},\dots,x_{i_{\nu_R}}$ from $\bar{x}$. For some $j \in [\xi_n]$, $(R,(a_{i_1},\dots,a_{i_{\nu_R}}))$ is the j^{th} element of Ξ_n according to the order $\prec_n$. Then $f(r_1,\dots,r_{\xi_n}) = r_j$ for all $r_1,\dots,r_{\xi_n} \in [0,1]$, so f is continuous.

Now we turn to the inductive case which concerns parts (4) and (5) of Definition 4. Suppose that $\varphi(\bar{x})$ has the form $\mathsf{C}(\varphi_1(\bar{x}),\dots,\varphi_k(\bar{x}))$ where $\mathsf{C} : [0,1]^k \to [0,1]$ is continuous. By the induction hypothesis there are, for all $i = 1,\dots,k$, continuous $f_i : [0,1]^{\xi_n} \to [0,1]$ such that for all $\mathcal{A} \in \mathbf{W}_n$, $f_i(\mathfrak{s}_n(\mathcal{A})) = \mathcal{A}(\varphi_i(\bar{a}))$. Then, for all $\mathcal{A} \in \mathbf{W}_n$, $\mathcal{A}(\varphi(\bar{a})) = \mathsf{C}(f_1(\mathfrak{s}_n(\mathcal{A})),\dots,f_k(\mathfrak{s}_n(\mathcal{A})))$. It follows that for all $\bar{r} \in [0,1]^{\xi_n}$, $f(\bar{r}) = \mathsf{C}(f_1(\bar{r}),\dots,f_k(\bar{r}))$, so f is continuous.

Finally, suppose that $\varphi(\bar{x})$ has the form $F(\psi(\bar{x},y), : y)$ where $F : [0,1]^{<\omega} \to [0,1]$ is a continuous aggregation function. Let $m := n - |\mathrm{rng}(\bar{a})|$ and let $b_1,\dots,b_m$ enumerate $[n] \setminus \mathrm{rng}(\bar{a})$. By the induction hypothesis, for each $i = 1,\dots,m$, there is continuous $f_i : [0,1]^{\xi_n} \to [0,1]$ such that for all $\mathcal{A} \in \mathbf{W}_n$, $\mathcal{A}(\psi(\bar{a},b_i)) = f_i(\mathfrak{s}_n(\mathcal{A}))$. It follows that for all $\mathcal{A} \in \mathbf{W}_n$, $\mathcal{A}(\varphi(\bar{a})) = F(\mathcal{A}(\psi(\bar{a},b_1)),\dots,\mathcal{A}(\psi(\bar{a},b_m))) = F(f_1(\mathfrak{s}_n(\mathcal{A})),\dots,f_m(\mathfrak{s}_n(\mathcal{A})))$. Since F is continuous as an aggregation function according to Definition 3 it follows from part (1) of that definition that F is continuous (in the usual sense of analysis in several variables) when restricted to sequences from $[0,1]^m$. Hence the function $f : [0,1]^{\xi_n} \to [0,1]$ defined by $f(\bar{r}) = F(f_1(\bar{r}),\dots,f_m(\bar{r}))$ for all $\bar{r} \in [0,1]^{\xi_n}$ is continuous and for all $\mathcal{A} \in \mathbf{W}_n$, $f(\mathfrak{s}_n(\mathcal{A})) = \mathcal{A}(\varphi(\bar{a}))$. This concludes the proof. □

Remark 2. Let $\mathsf{C} : [0,1]^2 \to [0,1]$ be defined by $\mathsf{C}(r_1,r_2) = |r_1 - r_2|$, so C is continuous. Let $\bar{x}$ be a sequence of distinct variables and let $\varphi_1(\bar{x}),\varphi_2(\bar{x}) \in CLA$. Furthermore, let $\varphi(\bar{x})$ denote the formula $\mathsf{C}(\varphi_1(\bar{x}),\varphi_2(\bar{x}))$ of CLA. By Lemma 6, it follows that, for all $n \in \mathbb{N}^+$, all $\bar{a} \in [n]^{|\bar{x}|}$, and all $\varepsilon > 0$, the set

$$\left\{\mathcal{A} \in \mathbf{W}_n : \Big|\mathcal{A}(\varphi_1(\bar{a})) - \mathcal{A}(\varphi_2(\bar{a}))\Big| \le \varepsilon\right\}$$

is a measurable subset of $\mathbf{W}_n$. As the set $[n]^{|\bar{x}|}$ is finite it follows that for every identity constraint $ic(\bar{x})$ the set

$$\bigcap_{\substack{\bar{a}\in[n]^{|\bar{x}|} \\ ic(\bar{a})\ holds}} \left\{\mathcal{A} \in \mathbf{W}_n : \Big|\mathcal{A}(\varphi_1(\bar{a})) - \mathcal{A}(\varphi_2(\bar{a}))\Big| \le \varepsilon\right\}$$

is a measurable subset of $\mathbf{W}_n$. In particular it makes sense to talk about the probability of such a set, using the probability distribution $\mathbb{P}_n$ on $\mathbf{W}_n$.

Definition 11. Let $\bar{x}$ be a sequence of length k of distinct variables and let $ic(\bar{x})$ be an identity constraint for $\bar{x}$. The formulas $\varphi(\bar{x}),\psi(\bar{x}) \in CLA$ are ***asymptotically equivalent with respect to*** ic if for all $\varepsilon > 0$,

$$\lim_{n\to\infty} \mathbb{P}_n\Big(\Big\{\mathcal{A} \in \mathbf{W}_n : \text{ for all } \bar{a} \in [n]^k \text{ satisfying } ic(\bar{x}),$$
$$\Big|\mathcal{A}(\varphi(\bar{a})) - \mathcal{A}(\psi(\bar{a}))\Big| \le \varepsilon\Big\}\Big) = 1.$$

If both φ and ψ have no free variables, that is, if $\bar{x}$ is empty, then we omit the references to an identity constraint.

4. Main results

The main results are given by the following theorem, which is subsequently proved in the rest of the section:

Theorem 1. *Let $\bar{x} = (x_1, \dots, x_k)$ be a sequence of distinct variables, let $ic(\bar{x})$ be an identity constraint for $\bar{x}$, and let $\varphi(\bar{x}) \in CLA$. (We allow the possiblility that $\bar{x}$ is empty in which case we can omit references to the identity constraint.)*
(a) Then there is an aggregation-free formula $\psi(\bar{x}) \in CLA$ such that $\varphi(\bar{x})$ and $\psi(\bar{x})$ are asymptotically equivalent with respect to $ic(\bar{x})$.
(b) For every interval $I \subseteq [0,1]$ there is $\alpha \in [0,1]$ such that, for all $\varepsilon > 0$, if n is large enough and $\bar{a} \in [n]^{|\bar{x}|}$ satisfies $ic(\bar{x})$, then

$$\Big|\mathbb{P}_n\big(\big\{\mathcal{A} \in \mathbf{W}_n : \mathcal{A}(\varphi(\bar{a})) \in I\big\}\big) - \alpha\Big| \leq \varepsilon.$$

Remark 3. It has been assumed that formulas in CLA take values in $[0,1]$ which is natural if we think of the values as probabilities or degrees of truth or belief. But for any constant $C > 1$ we can as well replace $[0,1]$ by $[0,C]$ in all definitions, lemmas and proofs, and the proofs will still work out.

The rest of this section is devoted to proving Theorem 1. We prove it by first proving a sequence of lemmas, then proving part (a) of Theorem 1, and then using part (a) to prove part (b) of the same theorem. There are two key lemmas: Lemma 8 shows that the probability that an aggregation-free formula takes a value in an interval J can be computed exactly by computing an integral that depends only on the formula, an identity constraint and J. Lemma 12 shows how to asymptotically eliminate one aggregation function, which is later used in a proof that uses induction on complexity of formulas.

Lemma 7. *Let $n, s \in \mathbb{N}^+$, let $l_1, \dots, l_s \in [\xi_n]$ be distinct, let $\mathsf{C} : [0,1]^s \to [0,1]$ be continuous, and let $J \subseteq [0,1]$ be an interval. Then*

$$\mathbb{P}^*_n\Big(\big\{\bar{r} \in [0,1]^{\xi_n} : \mathsf{C}(r_{l_1}, \dots, r_{l_s}) \in J\big\}\Big) =$$
$$\int_Y \lambda_{l_1}(r_1) \cdot \ldots \cdot \lambda_{l_s}(r_s) dr_1 \dots dr_s \quad \text{where } Y := \mathsf{C}^{-1}(J).$$

Proof. Fix $n, s \in \mathbb{N}^+$ and let $l_1, \dots, l_s \in [\xi_n]$, C, and J be as assumed. For $i \in [\xi_n]$ let $I_i = [0,1]$. Let $Y = \mathsf{C}^{-1}(J)$. As $\mathsf{C} : [0,1]^s \to [0,1]$ is continuous it follows that Y is a (Lebesgue) measurable subset of $[0,1]^s$. As $I_i = [0,1]$ for all $i \in [\xi_n]$ we may also view C as a function from $I_{l_1} \times \ldots \times I_{l_s}$ to $[0,1]$ and Y may be identified with a measurable subset of $I_{l_1} \times \ldots \times I_{l_s}$. Now we get (where $\bar{r} = (r_1, \dots, r_{\xi_n})$)

$$\mathbb{P}^*_n\Big(\big\{\bar{r} \in [0,1]^{\xi_n} : \mathsf{C}(r_{l_1}, \dots, r_{l_s}) \in J\big\}\Big) =$$
$$\mathbb{P}^*_n\Big(\big\{\bar{r} \in [0,1]^{\xi_n} : (r_{l_1}, \dots, r_{l_s}) \in Y\big\}\Big) =$$
$$\int_{Y \times \prod_{i \in [\xi_n] \setminus \{l_1, \dots, l_s\}} I_i} \pi^*_n(r_1, \dots, r_{\xi_n}) dr_1 \dots dr_{\xi_n} =$$
$$\int_{Y \times \prod_{i \in [\xi_n] \setminus \{l_1, \dots, l_s\}} I_i} \lambda_1(r_1) \cdot \ldots \cdot \lambda_{\xi_n}(r_{\xi_n}) dr_1 \dots dr_{\xi_n} =$$
$$\int_Y \lambda_{l_1}(r_{l_1}) \cdot \ldots \cdot \lambda_{l_s}(r_{l_s}) dr_{l_1} \dots dr_{l_s} \cdot \prod_{i \in [\xi_n] \setminus \{l_1, \dots, l_s\}} \int_{I_i} \lambda_i(r_i) dr_i =$$
$$\int_Y \lambda_{l_1}(r_{l_1}) \cdot \ldots \cdot \lambda_{l_s}(r_{l_s}) dr_{l_1} \dots dr_{l_s}$$

because $I_i = [0,1]$ and $\int_{[0,1]} \lambda_i(r_i) dr_i = 1$ for all i (by Lemma 5). □

Lemma 8. *Let $\bar{x} = (x_1, \dots, x_k)$ be a sequence of distinct variables, let $ic(\bar{x})$ be an identity constraint for $\bar{x}$, let $\mathsf{C} : [0,1]^s \to [0,1]$ be continuous, and let $\varphi(\bar{x})$ denote the formula $\mathsf{C}\big(R_1(\bar{x}_1), \dots, R_s(\bar{x}_s)\big)$, where, for all $i = 1, \dots, s$, $R_i \in \sigma$, $\mathrm{rng}(\bar{x}_i) \subseteq \mathrm{rng}(\bar{x})$, $|\bar{x}_i| = \nu_{R_i}$, and if $1 \leq i < j \leq s$ then $R_i \neq R_j$ or $\bar{x}_i \neq \bar{x}_j$. (We allow that $\bar{x}_i$ contains repetitions of variables.) Let $J \subseteq [0,1]$ be an interval, let $n \in \mathbb{N}^+$, let $\bar{a} = (a_1, \dots, a_k) \in [n]^k$ satisfy $ic(\bar{x})$, and let, for all $i = 1, \dots, s$, $\bar{a}_i$ be the subsequence of $\bar{a}$ that corresponds to $\bar{x}_i$ (so if $\bar{x}_i = (x_{l_1}, \dots, x_{l_{\nu_{R_i}}})$ then $\bar{a}_i = (a_{l_1}, \dots, a_{l_{\nu_{R_i}}})$), and let $ic_i(\bar{x}_i)$ be the restriction of $ic(\bar{x})$ to $\bar{x}_i$. Then*

$$\mathbb{P}_n\Big(\big\{\mathcal{A} \in \mathbf{W}_n : \ \mathcal{A}(\varphi(\bar{a})) \in J\big\}\Big) = \int_Y \mu^{ic_1}_{R_1}(r_1) \cdot \ldots \cdot \mu^{ic_s}_{R_s}(r_s) dr_1 \dots dr_s$$

where $Y := \mathsf{C}^{-1}(J)$. Hence the integral depends only on the formula $\varphi(\bar{x})$, the indentity constraint $ic(\bar{x})$ and J; in particular there is no dependence on n.

Proof. We adopt the assumptions of the lemma, so in particular some n is fixed and $\bar{a} \in [n]^k$ is a sequence that satisfies $ic(\bar{x})$. By assumption, if $1 \leq i < j \leq s$ then $R_i \neq R_j$ or $\bar{x}_i \neq \bar{x}_j$. Let $l_1, \dots, l_s$ be the places of the distinct pairs $(R_1, \bar{a}_1), \dots, (R_s, \bar{a}_s) \in \Xi_n$ according to the order $\prec_n$, so $l_1, \dots, l_s$ are distinct numbers in $[\xi_n]$.

We now get (where $\bar{r} = (r_1, \dots, r_{\xi_n})$)

$$\mathbb{P}_n\Big(\big\{\mathcal{A} \in \mathbf{W}_n : \ \mathcal{A}(\varphi(\bar{a})) \in J\big\}\Big) = \mathbb{P}^*_n\Big(\big\{\bar{r} \in [0,1]^{\xi_n} : \mathsf{C}(r_{l_1}, \dots, r_{l_s}) \in J\big\}\Big)$$

$$= \int_Y \lambda_{l_1}(r_1) \cdot \ldots \cdot \lambda_{l_s}(r_s) dr_1 \ldots dr_s \quad \text{(by Lemma 7 where } Y := \mathsf{C}^{-1}(J))$$

$$= \int_Y \mu_{R_1}^{ic_1}(r_1) \cdot \ldots \cdot \mu_{R_s}^{ic_s}(r_s) dr_1 \ldots dr_s$$

(by Definition 9 of λ_i and the choice of $l_1, \ldots, l_s$).

□

Lemma 9. *Let $\bar{x} = (x_1, \ldots, x_k)$ be a sequence of distinct variables, let y be a variable that does not occur in $\bar{x}$ and let $ic(\bar{x})$ be an identity constraint for $\bar{x}$. Let $\mathsf{C} : [0,1]^s \to [0,1]$ be continuous and let $\varphi(\bar{x}, y)$ denote the formula $\mathsf{C}\big(R_1(\bar{x}_1, y), \ldots, R_s(\bar{x}_s, y)\big)$, where, for all $i = 1, \ldots, s$, $R_i \in \sigma$, $\mathrm{rng}(\bar{x}_i) \subseteq \mathrm{rng}(\bar{x})$, $|\bar{x}_i| = \nu_{R_i} - 1$, and if $1 \le i < j \le s$ then $R_i \neq R_j$ or $\bar{x}_i \neq \bar{x}_j$.*

Let $J \subseteq [0,1]$ be an interval, let $n \in \mathbb{N}^+$, let $\bar{a} \in [n]^{|\bar{x}|}$ satisfy $ic(\bar{x})$, and suppose that $b_1, \ldots, b_m \in [n]$ are distinct and do not occur in $\bar{a}$. For all $i = 1, \ldots, m$, the event $\{\mathcal{A} \in \mathbf{W}_n : \mathcal{A}(\varphi(\bar{a}, b_i)) \in J\}$ is independent from all events $\{\mathcal{A} \in \mathbf{W}_n : \mathcal{A}(\varphi(\bar{a}, b_j)) \in J\}$ where $j \neq i$.

Proof. Fix $n \in \mathbb{N}^+$ and a sequence $\bar{a} = (a_1, \ldots, a_k) \in [n]^k$ that satisfies $ic(\bar{x})$. Let $\bar{a}_i$ be the subsequence of $\bar{a}$ that corresponds to the subsequence $\bar{x}_i$ of $\bar{x}$. For each $i = 1, \ldots, m$, let $l_{i,1}, \ldots, l_{i,s}$ be the places of the (by assumption) distinct pairs $(R_1, \bar{a}_1 b_i), \ldots, (R_s, \bar{a}_s b_i) \in \Xi_n$ according to the order $\prec_n$, so $l_{i,1}, \ldots, l_{i,s}$ are distinct numbers in $[\xi_n]$. Since $b_i \neq b_j$ if $i \neq j$ it follows that $l_{i,l} \neq l_{j,l'}$ if $i \neq j$, $1 \le l \le s$ and $1 \le l' \le s$.

Note that for all $l = 1, \ldots, s$ and $i, j = 1, \ldots, m$, $\bar{a}_l b_i$ and $\bar{a}_l b_j$ satisfy the same identity constraint. For each $l = 1, \ldots, s$, let $id_l(\bar{x}_l, y)$ be the (up to equivalence unique) identity constraint that is satified by $\bar{a}_l b_1$.

For each $i \in [\xi_n]$ let $I_i = [0,1]$. Let $Y = \mathsf{C}^{-1}(J)$, so $Y \subseteq [0,1]^s$. For all $i = 1, \ldots, m$ let $Y_i \subseteq I_{l_{i,1}} \times \ldots \times I_{l_{i,s}}$ be a copy of Y. (That is $Y_i = \mathsf{C}^{-1}(J)$ if C is viewed as a function from $I_{l_{i,1}} \times \ldots \times I_{l_{i,s}}$ to $[0,1]$.) Let $L = \bigcup_{i=1}^m \{l_{i,1}, \ldots, l_{i,s}\}$. For $i = 1, \ldots, m$, let $\mathbf{E}_n^i = \{\mathcal{A} \in \mathbf{W}_n : \mathcal{A}(\varphi(\bar{a}, b_i)) \in J\}$. Now we have (where $\bar{r} = (r_1, \ldots, r_{\xi_n})$)

$$\mathbb{P}_n\Big(\bigcap_{i=1}^m \mathbf{E}_n^i\Big) = \mathbb{P}_n^*\Big(\big\{\bar{r} \in [0,1]^{\xi_n} : \text{ for all } i = 1, \ldots, m, \ (r_{l_{i,1}}, \ldots, r_{l_{i,s}}) \in J\big\}\Big) =$$

$$\int_{Y_1 \times \ldots \times Y_m \times \prod_{i \in [\xi_n] \setminus L} I_i} \pi_n^*(r_1, \ldots, r_{\xi_n}) dr_1 \ldots dr_{\xi_n} =$$

$$\int_{Y_1 \times \ldots \times Y_m \times \prod_{i \in [\xi_n] \setminus L} I_i} \lambda_1(r_1) \cdot \ldots \cdot \lambda_{\xi_n}(r_{\xi_n}) dr_1 \ldots dr_{\xi_n} =$$

$$\left(\prod_{i=1}^m \int_{Y_i} \lambda_{l_{i,1}}(r_{l_{i,1}}) \cdot \ldots \cdot \lambda_{l_{i,s}}(r_{l_{i,s}}) dr_{l_{i,1}} \ldots dr_{l_{i,s}}\right) \cdot \left(\prod_{i \in [\xi_n] \setminus L} \int_{I_i} \lambda_i(r_i) dr_i\right) =$$

$$\prod_{i=1}^m \int_{Y_i} \lambda_{l_{i,1}}(r_{l_{i,1}}) \cdot \ldots \cdot \lambda_{l_{i,s}}(r_{l_{i,s}}) dr_{l_{i,1}} \ldots dr_{l_{i,s}} =$$

$$\prod_{i=1}^m \int_{Y_i} \mu_{R_1}^{id_1}(r_{l_{i,1}}) \cdot \ldots \cdot \mu_{R_s}^{id_s}(r_{l_{i,s}}) dr_{l_{i,1}} \ldots dr_{l_{i,s}} = \prod_{i=1}^m \mathbb{P}_n\big(\mathbf{E}_n^i\big) \quad \text{(by Lemma 8).}$$

Hence the events $\mathbf{E}_n^1, \ldots, \mathbf{E}_n^m$ are independent. □

The following is a direct consequence of [37, Corollary A.1.14] which in turn follows from a bound given by Chernoff [38]:

Lemma 10. *Let Z be the sum of n independent 0/1-valued random variables, each one with probability p of having the value 1, where $p > 0$. For every $\varepsilon > 0$ there is $c_\varepsilon > 0$, depending only on ε, such that the probability that $|Z - pn| > \varepsilon pn$ is less than $2e^{-c_\varepsilon pn}$. (If $p = 0$ then the same statement holds if '$2e^{-c_\varepsilon pn}$' is replaced by 'e^{-n}'.)*

Lemma 11. *Let $\bar{x} = (x_1, \ldots, x_k)$ be a sequence of distinct variables, let y be a variable that does not occur in $\bar{x}$, let $ic(\bar{x})$ be an identity constraint for $\bar{x}$, and let $ic'(\bar{x}, y)$ be the (up to equivalence unique) identity constraint for $(x_1, \ldots, x_k, y)$ which implies $ic(\bar{x})$ and $y \neq x_i$ for all $i = 1, \ldots, k$. Let $\mathsf{C} : [0,1]^s \to [0,1]$ be continuous and let $\varphi(\bar{x}, y)$ denote the formula $\mathsf{C}\big(R_1(\bar{x}_1, y), \ldots, R_s(\bar{x}_s, y)\big)$, where, for all $i = 1, \ldots, s$, $R_i \in \sigma$, $\mathrm{rng}(\bar{x}_i) \subseteq \mathrm{rng}(\bar{x})$, $|\bar{x}_i| = \nu_{R_i} - 1$, and if $1 \le i < j \le s$ then $R_i \neq R_j$ or $\bar{x}_i \neq \bar{x}_j$.*

Let $J_1, \ldots, J_t \in [0,1]$ be intervals. For $i = 1, \ldots, t$, let $Y_i = \mathsf{C}^{-1}(J_i)$ and let

$$\alpha_i = \int_{Y_i} \mu_{R_1}^{ic_1}(r_1) \cdot \ldots \cdot \mu_{R_s}^{ic_s}(r_s) dr_1 \ldots dr_s$$

where, for $j = 1, \ldots, s$, ic_j is the restriction of $ic'(\bar{x}, y)$ to $\bar{x}_j y$. Let $\varepsilon > 0$. Then there is $\beta > 0$ such that for all sufficiently large n the probability that the following holds for a random $\mathcal{A} \in \mathbf{W}_n$ is at least $1 - e^{-\beta n}$: For all $\bar{a} \in [n]^k$ that satisfy $ic(\bar{x})$ and all $i = 1, \ldots, t$,

$$\alpha_i - \varepsilon \le \frac{\Big|\big\{b \in [n] \setminus \mathrm{rng}(\bar{a}) : \mathcal{A}(\varphi(\bar{a}, b)) \in J_i\big\}\Big|}{n - |\mathrm{rng}(\bar{a})|} \le \alpha_i + \varepsilon. \tag{4.1}$$

Proof. Fix some $n \in \mathbb{N}^+$ and a sequence $\bar{a} \in [n]^k$ satisfying $ic(\bar{x})$. Also fix some $i \in \{1, \ldots, t\}$. Let $b_1, \ldots, b_m$ be a list, without repetition, of all members of $[n] \setminus \mathrm{rng}(\bar{a})$, so $m = n - |\mathrm{rng}(\bar{a})|$. For all $j = 1, \ldots, m$, let $\mathbf{E}_n^j = \{\mathcal{A} \in \mathbf{W}_n : \mathcal{A}(\varphi(\bar{a}, b_j)) \in J_i\}$ and let $Z_j : \mathbf{W}_n \to \{0,1\}$ be the random variable defined by $Z_j(\mathcal{A}) = 1$ if $\mathcal{A} \in \mathbf{E}_n^j$ and $Z_i(\mathcal{A}) = 0$ otherwise. Let $Z = Z_1 + \ldots + Z_m$. By Lemma 8, the probability that $Z_j = 1$ is α_i. By Lemma 9, $Z_1, \ldots, Z_m$ are independent. As $m = n - |\mathrm{rng}(\bar{a})|$ it follows from Lemma 10 that for every $\varepsilon > 0$ there is $c > 0$ (depending only on ε) such that the probability that $|Z - \alpha_i(n - |\mathrm{rng}(\bar{a})|)| > \varepsilon\alpha_i(n - |\mathrm{rng}(\bar{a})|)$ is less than $2e^{-c\alpha_i(n-|\mathrm{rng}(\bar{a})|)}$. By assuming that n is large enough and adjusting c to some (slightly smaller) $c' > 0$ it follows that the probability that $|Z - \alpha_i(n - |\mathrm{rng}(\bar{a})|)| > \varepsilon\alpha_i(n - |\mathrm{rng}(\bar{a})|)$ is less than $2e^{-c'\alpha_i n}$. This is equivalent to saying that the probability that a random $\mathcal{A} \in \mathbf{W}_n$ does *not* satisfy (4.1) is less than $2e^{-c'\alpha_i n}$. Since there are n^k sequences $\bar{a} \in [n]^k$ it follows that the probability that there is some $\bar{a} \in [n]^k$ satisfying $ic(\bar{x})$ such that (4.1) is not satisfied is less than $n^k \cdot 2e^{-c'\alpha_i n}$. Hence the probability that there is some $i \in \{1, \ldots, t\}$ and $\bar{a}$ such that (4.1) is not satisfied is less than $n^k \cdot 2te^{-c'\alpha_i n}$. Then there is $\beta > 0$ such that $n^k \cdot 2e^{-c'\alpha_i n} \leq e^{-\beta n}$ for all sufficiently large n. Thus, the probability that, for sufficiently large n, a random $\mathcal{A} \in \mathbf{W}_n$ satisfies (4.1) for all $i = 1, \ldots, t$ and all sequences $\bar{a} \in [n]^k$ that satisfy $ic(\bar{x})$ is at least $1 - e^{-\beta n}$. □

Lemma 12. *Let $\bar{x} = (x_1, \ldots, x_k)$ be a sequence of distinct variables, let $ic(\bar{x})$ be an identity constraint for $\bar{x}$, suppose that the variable y does not occur in $\bar{x}$. Suppose that $\varphi(\bar{x}, y) \in CLA$ is aggregation-free and that $F : [0,1]^{<\omega} \to [0,1]$ is a continuous aggregation function. Then the formula $F(\varphi(\bar{x}, y) : y)$ is asymptotically equivalent to an aggregation-free formula with respect to $ic(\bar{x})$.*

Proof. By Lemma 3, we may assume that $\varphi(\bar{x}, y)$ is a constant or has the form

$$\mathsf{C}(R_1(\bar{x}_1), \ldots, R_t(\bar{x}_t), R_{t+1}(\bar{x}_{t+1}, y), \ldots, R_{t+s}(\bar{x}_{t+s}, y))$$

where $\mathsf{C} : [0,1]^{t+s} \to [0,1]$ is continuous, for all i, $R_i \in \sigma$, $\mathrm{rng}(\bar{x}_i) \subseteq \mathrm{rng}(\bar{x})$, $|\bar{x}_i| = \nu_{R_i}$ if $i \leq t$ and $|\bar{x}_i| = \nu_{R_i} - 1$ otherwise, and if $1 \leq i < j \leq t + s$ then $R_i \neq R_j$ or $\bar{x}_i \neq \bar{x}_j$. The first case, when $\varphi(\bar{x}, y)$ is a constant, is covered by the second case since C may be a constant function. Therefore we only consider the second case.

Let $\psi(\bar{x}) := F(\varphi(\bar{x}, y) : y)$. It suffices to find continuous $\mathsf{D} : [0,1]^t \to [0,1]$ such that if $\theta(\bar{x}) := \mathsf{D}(R_1(\bar{x}_1), \ldots, R_t(\bar{x}_t))$ then, for all $\varepsilon > 0$,

$$\lim_{n\to\infty} \mathbb{P}_n\Big(\big\{\mathcal{A} \in \mathbf{W}_n : \text{ for all } \bar{a} \in [n]^{|\bar{x}|} \text{ satisfying } ic(\bar{x}) \\ |\mathcal{A}(\psi(\bar{a})) - \mathcal{A}(\theta(\bar{a}))| \leq \varepsilon\big\}\Big) = 1. \tag{4.2}$$

Fix $\bar{r} = (r_1, \ldots, r_t) \in [0,1]^t$. We now find out what $\mathsf{D}(\bar{r})$ should be for (4.2) to hold. First we define $\mathsf{C}_{\bar{r}} : [0,1]^s \to [0,1]$ by

$$\mathsf{C}_{\bar{r}}(p_1, \ldots, p_s) = \mathsf{C}(r_1, \ldots, r_t, p_1, \ldots, p_s)$$

so $\mathsf{C}_{\bar{r}}$ is continuous. Let $\varphi_{\bar{r}}(\bar{x}, y) := \mathsf{C}_{\bar{r}}(R_{t+1}(\bar{x}_{t+1}, y), \ldots, R_{t+s}(\bar{x}_{t+s}, y))$. Let $ic'(\bar{x}, y)$ be the (unique up to equivalence) identity constraint that implies $ic(\bar{x})$ and $y \neq x_i$ for all $i = 1, \ldots, k$.

Let $M \in \mathbb{N}^+$ and $J_i := [\frac{i}{M}, \frac{i+1}{M}]$ for $i = 0, \ldots, M-1$. By Lemma 8, for all $i = 0, \ldots, M-1$, there is α_i depending only on $ic'(\bar{x}, y)$ (which depends only on $ic(\bar{x})$), J_i and $\varphi_{\bar{r}}$ (which depends only on φ and $\bar{r}$) such that for all n, all tuples $\bar{a} \in [n]^k$ satisfying $ic(\bar{x})$ and all $b \in [n] \setminus \mathrm{rng}(\bar{a})$,

$$\mathbb{P}_n\big(\{\mathcal{A} \in \mathbf{W}_n : \mathcal{A}(\varphi_{\bar{r}}(\bar{a}, b)) \in J_i\}\big) = \alpha_i.$$

For all n and $\delta > 0$ let $\mathbf{X}_n^\delta$ be the set of all $\mathcal{A} \in \mathbf{W}_n$ such that for all tuples $\bar{a} \in [n]^k$ that satisfy $ic(\bar{x})$ and all $i = 0, \ldots, M-1$,

$$\alpha_i - \delta \leq \frac{\big|\{b \in [n] \setminus \mathrm{rng}(\bar{a}) : \mathcal{A}(\varphi_{\bar{r}}(\bar{a}, b)) \in J_i\}\big|}{n - |\mathrm{rng}(\bar{a})|} \leq \alpha_i + \delta. \tag{4.3}$$

According to Lemma 11, $\lim_{n\to\infty} \mathbb{P}_n\big(\mathbf{X}_n^\delta\big) = 1$.

Let $\varepsilon > 0$. By assumption, F is a continuous aggregation function. This means that $\delta > 0$ and $M, N \in \mathbb{N}$ can be chosen so that if $\bar{q} = (q_1, \ldots, q_n), \bar{q}' = (q'_1, \ldots, q'_m) \in [0,1]^{<\omega}$ satisfy conditions (2)(a)–(d) of Definition 3 of continuity of aggregation functions, then $|F(\bar{q}) - F(\bar{q}')| \leq \varepsilon$. Suppose that δ, M and N are such numbers. Without loss of generality we may assume that, for all $i = 1, \ldots, M-1$, if $\alpha_i > 0$ then $\alpha_i > \delta$. Without loss of generality we may assume that $N \in \mathbb{N}$ is large enough so that if $n \geq N + |\mathrm{rng}(\bar{a})|$ then $\mathbf{X}_n^\delta \neq \emptyset$. Let $n \geq N + |\mathrm{rng}(\bar{a})|$, $\mathcal{A} \in \mathbf{X}_n^\delta$, and let $\bar{a} \in [n]^k$ satisfy $ic(\bar{x})$. Enumerate $[n] \setminus \mathrm{rng}(\bar{a})$ as $b_1, \ldots, b_m$ where $m := n - |\mathrm{rng}(\bar{a})|$ so $m \geq N$. Define $q_i := \mathcal{A}(\varphi_{\bar{r}}(\bar{a}, b_i))$ and $\bar{q} := (q_1, \ldots, q_m)$. Then conditions (2)(a) and (2)(b) of Definition 3 hold for $\bar{q}$ (because $m \geq N$, $\alpha_i > \delta$ if $\alpha > 0$, and $\mathcal{A} \in \mathbf{X}_n^\delta$).

We now verify that conditions (2)(c) and (2)(d) of Definition 3 hold for $\bar{q}$. For all $j = 1, \ldots, s$ let $ic'_j(\bar{x}_j, y)$ be the restriction of $ic'(\bar{x}, y)$ to $\bar{x}_j y$. By Lemma 8 we have that, for all $i = 0, \ldots, M-1$,

$$\alpha_i = \int_{Y_i} \mu_{R_{t+1}}^{ic'_1}(r_1) \cdot \ldots \cdot \mu_{R_{t+s}}^{ic'_s}(r_s) dr_1 \ldots dr_s$$

where $Y_i = \mathsf{C}_{\bar{r}}^{-1}(J_i) \subseteq [0,1]^s$ and $g(r_1, \ldots, r_s) := \mu_{R_{t+1}}^{ic'_1}(r_1) \cdot \ldots \cdot \mu_{R_{t+s}}^{ic'_s}(r_s)$ is a continuous function, so Y_i is a measurable subset of $[0,1]^s$. Note that by the continuity of g, $g([0,1]^s)$ is an interval. Hence $g([0,1]^s) \cap J_i$ is an interval for all i. Also, if $\alpha_i = 0$ then $g([0,1]^s) \cap J_i$ contains at most one point; hence, if $g([0,1]^s) \cap J_i$ contains more than one point then $\alpha_i > 0$. It follows that if $i < j < k$ (or $i > j > k$), $\alpha_i > 0$, $\alpha_j = 0$ and $\alpha_k > 0$ then $g([0,1]^s)$ is not an interval, so g is *not* continuous. Thus we conclude that condition (2)(c) of Definition 3 holds for $\bar{q}$. Now suppose that $i < j < k$ (or $i > j > k$) $\alpha_i = \alpha_j = 0$ and $\alpha_k > 0$. For a contradiction, suppose that $\{l : q_l \in J_i\} \neq \emptyset$. From $\alpha_k > 0$ it follows that $g([0,1]^s) \cap J_k \neq \emptyset$. From $\alpha_j = 0$ it follows that $g([0,1]^s) \cap J_i$ contains at most one point. But then $g([0,1]^s)$ is not an interval, contradicting that g is continuous. Hence, also condition (2)(d) of Definition 3 holds for $\bar{q}$.

By the assumptions on δ, M and N, it follows that there is an open interval $I_\varepsilon \subseteq [0,1]$ with width (the distance between its endpoints) at most ε and depending only on $\varphi, \bar{r}, M, N$, and δ such that $F(\bar{q}) \in I_\varepsilon$. Furthermore, if $\varepsilon > \varepsilon' > 0$ it follows in the same way that there are $M' \geq M$, $N' \geq N$, and $0 \leq \delta' \leq \delta$ such that if J'_i, α'_i and $I_{\varepsilon'}$ are defined as J_i, α_i, and I_ε, but with M', N' and δ' in place of M, N and δ, respectively, then $I_{\varepsilon'} \subset I_\varepsilon$ and $I_{\varepsilon'}$ has width at most ε'. We now define $D(\bar{r})$ to be the unique member of $\bigcap_{\varepsilon > 0} I_\varepsilon$. It follows that if $\mathcal{A} \in \mathbf{X}^\delta_n$ and $R^{\mathcal{A}}_i(\bar{a}_i) = r_i$ for $i = 1, \dots, t$, then $\mathcal{A}(\psi(\bar{a})) = \mathcal{A}\big(F(\varphi(\bar{a}, y) : y)\big) = \mathcal{A}\big(F(\varphi_{\bar{r}}(\bar{a}, y) : y)\big) = F(\bar{q})$ and $\mathcal{A}(\theta(\bar{a})) = \mathsf{D}(\bar{r})$; hence $|\mathcal{A}(\psi(\bar{a})) - \mathcal{A}(\theta(\bar{a}))| = |F(\bar{q}) - \mathsf{D}(\bar{r})| \leq \varepsilon$. Observe that the definition of $\mathsf{D}(\bar{r})$ does not depend on the choice $\mathcal{A}$ or $\bar{a} \in [n]^k$ as long as $\mathcal{A} \in \mathbf{X}^\delta_n$ and $\bar{a}$ satisfies $ic(\bar{x})$, because (4.3) holds for all $\mathcal{A} \in \mathbf{X}^\delta_n$ and all $\bar{a} \in [n]^k$ that satisfy $ic(\bar{x})$.

For the continuity of D, let $\bar{r}$ be as above and suppose that $\bar{r}' = (r'_1, \dots, r'_t) \in [0,1]^t$ and that $|r_i - r'_i| \leq \delta$ for all $i = 1, \dots, t$. Since C is continuous on the compact space $[0,1]^{t+s}$ it is in fact uniformly continuous. It follows that, for every $\varepsilon' > 0$, if δ is small enough and, for all $i = t+1, \dots, t+s$, the reals $r_i, r'_i \in [0,1]$ are such that $|r_i - r'_i| \leq \delta$, then $|\mathsf{C}_{\bar{r}}(r_{t+1}, \dots, r_{r+s}) - \mathsf{C}_{\bar{r}'}(r'_{t+1}, \dots, r'_{r+s})| \leq \varepsilon'$. Let $\varphi_{\bar{r}'}(\bar{x}, y) := \mathsf{C}_{\bar{r}'}(R_{t+1}(\bar{x}_{t+1}, y), \dots, R_{t+s}(\bar{x}_{t+s}, y))$ For $\mathcal{A}$, $\bar{a}$ and b_i as above we get

$$\Big|\mathcal{A}(\varphi_{\bar{r}}(\bar{a}, b_i)) - \mathcal{A}(\varphi_{\bar{r}'}(\bar{a}, b_i))\Big| \leq \varepsilon'.$$

For $i = 1, \dots, m$ let $q'_i = \mathcal{A}(\varphi_{\bar{r}'}(\bar{a}, b_i))$ so we have $|q_i - q'_i| \leq \varepsilon'$ for all i. Let $\bar{q}' = (q'_1, \dots, q'_m)$. Then, for every $\varepsilon > 0$, if ε' is small enough (which can be arranged by taking δ small enough) we get $|F(\bar{q}) - F(\bar{q}')| \leq \varepsilon$ because F is continuous. Given any $\varepsilon > 0$ we can argue as before (for $\bar{q}$ and $\bar{r}$) to get $|F(\bar{q}') - \mathsf{D}(\bar{r}')| \leq \varepsilon$ (given that δ is small enough). Then we have $|F(\bar{q}) - F(\bar{q}')| \leq \varepsilon$, $|F(\bar{q}') - \mathsf{D}(\bar{r}')| \leq \varepsilon$, and $|F(\bar{q}) - \mathsf{D}(\bar{r})| \leq \varepsilon$, which gives $|\mathsf{D}(\bar{r}) - \mathsf{D}(\bar{r}')| \leq 3\varepsilon$. Hence D is continuous. □

Lemma 13. *Let $\bar{x} = (x_1, \dots, x_k)$ be a sequence of distinct variables and let $y \notin \operatorname{rng}(\bar{x})$ be a variable. Let $ic(\bar{x}, y)$ be an identity constraint for $\bar{x}y$ which implies that $y \neq x_i$ for all $i = 1, \dots, k$, and let $ic'(\bar{x})$ be the restriction of ic to $\bar{x}$. Suppose that $\varphi(\bar{x}, y), \varphi'(\bar{x}, y) \in CLA$ where $\varphi(\bar{x}, y)$ and $\varphi'(\bar{x}, y)$ are asymptotically equivalent with respect to $ic(\bar{x}, y)$ and $F : [0,1]^{<\omega} \to [0,1]$ is a continuous aggregation function. Then $F(\varphi(\bar{x}, y) : y)$ and $F(\varphi'(\bar{x}, y) : y)$ are asymptotically equivalent with respect to $ic'(\bar{x})$.*

Proof. For all $n \in \mathbb{N}^+$ and $\delta > 0$ let $\mathbf{Y}^\delta_n$ be the set of all $\mathcal{A} \in \mathbf{W}_n$ such that for all $\bar{a} \in [n]^k$ that satisfy $ic'(\bar{x})$ and all $b \in [n] \setminus \operatorname{rng}(\bar{a})$,

$$\Big|\mathcal{A}(\varphi(\bar{a}, b)) - \mathcal{A}(\varphi'(\bar{a}, b))\Big| \leq \delta.$$

Since we assume that $\varphi(\bar{x}, y)$ and $\varphi'(\bar{x}, y)$ are asymptotically equivalent with respect to $ic(\bar{x}, y)$ it follows that for all $\delta > 0$, $\lim_{n \to \infty} \mathbb{P}_n\big(\mathbf{Y}^\delta_n\big) = 1$. Let $\varepsilon > 0$. It now suffices to show that if $\delta > 0$ is small enough, n large enough, $\mathcal{A} \in \mathbf{Y}^\delta_n$, and $\bar{a} \in [n]^k$ satisfies $ic'(\bar{x})$, then

$$\Big|\mathcal{A}\big(F(\varphi(\bar{a}, y) : y)\big) - \mathcal{A}\big(F(\varphi'(\bar{a}, y) : y)\big)\Big| \leq \varepsilon. \tag{4.4}$$

So take $\mathcal{A} \in \mathbf{Y}^\delta_n$ and $\bar{a} \in [n]^k$ that satisfies $ic'(\bar{x})$ and let $b_1, \dots, b_m$ enumerate $[n] \setminus \operatorname{rng}(\bar{a})$. For $i = 1, \dots, m$, let $q_i = \mathcal{A}(\varphi(\bar{a}, b_i))$ and $q'_i = \mathcal{A}(\varphi'(\bar{a}, b_i))$, and then let $\bar{q} = (q_1, \dots, q_m)$ and $\bar{q}' = (q'_1, \dots, q'_m)$. Then $|q_i - q'_i| \leq \delta$ for all i. Since we assume that F is continuous it follows from condition (1) of Definition 3 of continuity that if δ is small enough and n large enough then $|F(\bar{q}) - F(\bar{q}')| \leq \varepsilon$, hence (4.4) holds. □

Proof of Theorem 1. We begin with part (a) and prove it by induction on the complexity, or construction, of $\varphi(\bar{x}) \in CLA$. If $\varphi(\bar{x})$ has one of the forms 1, 2, or 3, in Definition 4 then it is already aggregation-free and asymptotically equivalent to itself with respect to $id(\bar{x})$, for every identity constraint $ic(\bar{x})$. Next, suppose that $\varphi(\bar{x})$ has the form $\mathsf{C}(\varphi_1(\bar{x}), \dots, \varphi_k(\bar{x}))$ for some $\varphi_i(\bar{x}) \in CLA$, $i = 1, \dots, k$, and continuous $\mathsf{C} : [0,1]^k \to [0,1]$. Let $ic(\bar{x})$ be an identity constraint for $\bar{x}$. By the induction hypothesis, for $i = 1, \dots, k$, $\varphi_i(\bar{x})$ is asymptotically equivalent to some aggregation-free $\varphi'_i(\bar{x})$ with respect to $ic(\bar{x})$. Then $\varphi'(\bar{x}) := \mathsf{C}(\varphi'_1(\bar{x}), \dots, \varphi'_k(\bar{x}))$ is aggregation-free. Since C is continuous it follows straightforwardly that $\varphi'(\bar{x})$ is asymptotically equivalent to $\varphi(\bar{x})$ with respect to $ic(\bar{x})$.

Now suppose that $\varphi(\bar{x})$ has the form $F(\psi(\bar{x}, y) : y)$ where F is a continuous aggregation function (and $\psi(\bar{x}, y) \in CLA$). Let $ic(\bar{x})$ be an identity constraint on $\bar{x}$ and let $ic'(\bar{x}, y)$ be (the unique up to equivalence) identity constraint that implies $ic(\bar{x})$ and $y \neq x_i$ for all $i = 1, \dots, k$. By the induction hypothesis, $\psi(\bar{x}, y)$ is asymptotically equivalent to some aggregation-free $\psi'(\bar{x}, y)$ with respect to $ic'(\bar{x}, y)$. By Lemma 12, $F(\psi'(\bar{x}, y) : y)$ is asymptotically equivalent to some aggregation-free $\theta(\bar{x})$ with respect to $ic(\bar{x})$. Lemma 13 implies that $F(\psi(\bar{x}, y) : y)$ and $F(\psi'(\bar{x}, y) : y)$ are asymptotically equivalent with respect to $ic(\bar{x})$. Since asymptotic equivalence (with respect to $ic(\bar{x})$) is a transitive relation between formulas (which is easily verified) it follows that $F(\psi(\bar{x}, y) : y)$ is asymptotically equivalent to $\theta(\bar{x})$ with respect to $ic(\bar{x})$.

Finally we prove part (b) of Theorem 1. Let $\varphi(\bar{x}) \in CLA$ and let $ic(\bar{x})$ be an identity constraint for $\bar{x}$. By part (a), $\varphi(\bar{x})$ is asymptotically equivalent to some aggregation-free $\psi(\bar{x})$. By Lemma 3 we may assume that $\psi(\bar{x})$ satisfies the preconditions of the formula in Lemma 8, so that lemma is applicable to $\psi(\bar{x})$.

Let $I = [c, d] \subseteq [0,1]$ be an interval. The cases when $c = d$ or when $c = 0$ and $d = 1$ are trivial, so we assume that $0 < c < d < 1$. (The cases when $c = 0$ or $d = 1$ are treated similarly, with minor modifications of the definitions of I^+_δ and I^-_δ below.) Let $\delta > 0$ be such that $\delta < \min\big((d - c)/2, c/2, (1 - d)/2\big)$ and let , $I^+_\delta = [c - \delta, d + \delta]$ and $I^-_\delta = [c + \delta, d - \delta]$. Lemma 8 implies that there are α, α^+_δ and α^-_δ such that for all n and all $\bar{a} \in [n]^{|\bar{x}|}$ that satisfy $ic(\bar{x})$,

$$\begin{aligned}
&\mathbb{P}_n\big(\{\mathcal{A} \in \mathbf{W}_n : \mathcal{A}(\psi(\bar{a})) \in I\}\big) = \alpha, && (4.5)\\
&\mathbb{P}_n\big(\{\mathcal{A} \in \mathbf{W}_n : \mathcal{A}(\psi(\bar{a})) \in I^+_\delta\}\big) = \alpha^+_\delta, \text{ and} &&\\
&\mathbb{P}_n\big(\{\mathcal{A} \in \mathbf{W}_n : \mathcal{A}(\psi(\bar{a})) \in I^-_\delta\}\big) = \alpha^-_\delta, &&
\end{aligned}$$

where, for some $s \in \mathbb{N}^+$, $R_1, \dots, R_s \in \sigma$, identity constraints $ic_1(\bar{x}_1), \dots, ic_s(\bar{x}_s)$, and continuous $\mathsf{C} : [0,1]^s \to [0,1]$,

$$\alpha_\delta^+ = \int_{\mathsf{C}^{-1}(I_\delta^+)} \mu_{R_1}^{ic_1}(r_1) \cdot \ldots \cdot \mu_{R_s}^{ic_s}(r_s) dr_1 \ldots dr_s, \text{ and}$$

$$\alpha_\delta^- = \int_{\mathsf{C}^{-1}(I_\delta^-)} \mu_{R_1}^{ic_1}(r_1) \cdot \ldots \cdot \mu_{R_s}^{ic_s}(r_s) dr_1 \ldots dr_s.$$

As $I_\delta^- \subseteq I \subseteq I_\delta^+$ we get $\alpha_\delta^- \leq \alpha \leq \alpha_\delta^+$. Let $\mathbf{X}_n^\delta$ be the set of all $\mathcal{A} \in \mathbf{W}_n$ such that for all $\bar{a} \in [n]^{|\bar{x}|}$ that satisfy $ic(\bar{a})$, $|\mathcal{A}(\varphi(\bar{a})) - \mathcal{A}(\psi(\bar{a}))| \leq \delta$. Since for all $n \in \mathbb{N}^+$ and $\bar{a} \in [n]^{|\bar{x}|}$ that satisfy $ic(\bar{a})$,

$$\{\mathcal{A} \in \mathbf{X}_n^\delta : \mathcal{A}(\psi(\bar{a})) \in I_\delta^-\} \subseteq \{\mathcal{A} \in \mathbf{X}_n^\delta : \mathcal{A}(\varphi(\bar{a})) \in I\} \subseteq \{\mathcal{A} \in \mathbf{X}_n^\delta : \mathcal{A}(\psi(\bar{a})) \in I_\delta^+\}$$

we must have

$$\begin{aligned} &\mathbb{P}_n\big(\{\mathcal{A} \in \mathbf{X}_n^\delta : \mathcal{A}(\psi(\bar{a})) \in I_\delta^-\}\big) \leq \\ &\mathbb{P}_n\big(\{\mathcal{A} \in \mathbf{X}_n^\delta : \mathcal{A}(\varphi(\bar{a})) \in I\}\big) \leq \\ &\mathbb{P}_n\big(\{\mathcal{A} \in \mathbf{X}_n^\delta : \mathcal{A}(\psi(\bar{a})) \in I_\delta^+\}\big). \end{aligned} \tag{4.6}$$

Since $\lim_{n\to\infty} \mathbb{P}_n\big(\mathbf{X}_n^\delta\big) = 1$ it follows from (4.5) that, for every $\varepsilon > 0$ and all sufficiently large n and all $\bar{a} \in [n]^{|\bar{x}|}$ that satisfy $ic(\bar{x})$, we have

$$\mathbb{P}_n\big(\{\mathcal{A} \in \mathbf{X}_n^\delta : \mathcal{A}(\psi(\bar{a})) \in I_\delta^+\}\big) = (\alpha_\delta^+ - \varepsilon, \alpha_\delta^+ + \varepsilon), \text{ and}$$

$$\mathbb{P}_n\big(\{\mathcal{A} \in \mathbf{X}_n^\delta : \mathcal{A}(\psi(\bar{a})) \in I_\delta^-\}\big) = (\alpha_\delta^- - \varepsilon, \alpha_\delta^- + \varepsilon).$$

This and (4.6) gives, for sufficiently large n,

$$\alpha_\delta^- - \varepsilon < \mathbb{P}_n\big(\{\mathcal{A} \in \mathbf{X}_n^\delta : \mathcal{A}(\varphi(\bar{a})) \in I\}\big) < \alpha_\delta^+ + \varepsilon.$$

Since $\lim_{n\to\infty} \mathbb{P}_n\big(\mathbf{X}_n^\delta\big) = 1$ it follows that for all sufficiently large n and all $\bar{a} \in [n]^{|\bar{x}|}$ that satisfy $ic(\bar{x})$,

$$\alpha_\delta^- - 2\varepsilon < \mathbb{P}_n\big(\{\mathcal{A} \in \mathbf{W}_n : \mathcal{A}(\varphi(\bar{a})) \in I\}\big) < \alpha_\delta^+ + 2\varepsilon. \tag{4.7}$$

Recall that $\mathsf{C}^{-1}(I_\delta^-) \subseteq \mathsf{C}^{-1}(I_\delta^+)$ and $\alpha_\delta^- \leq \alpha \leq \alpha_\delta^+$. By taking $\delta > 0$ small enough we can make the measure of $\mathsf{C}^{-1}(I_\delta^+) \setminus \mathsf{C}^{-1}(I_\delta^-)$ as small as we like. Due to the expressions above of α_δ^+ and α_δ^- as integrals over $\mathsf{C}^{-1}(I_\delta^+)$ and $\mathsf{C}^{-1}(I_\delta^-)$, respectively, it follows that for all $\varepsilon > 0$, if $\delta > 0$ is sufficiently small then $\alpha_\delta^+ - \alpha_\delta^- < \varepsilon$. So from (4.7) we get

$$\Big|\mathbb{P}_n\big(\{\mathcal{A} \in \mathbf{W}_n : \mathcal{A}(\varphi(\bar{a})) \in I\}\big) - \alpha\Big| \leq 4\varepsilon$$

if $\delta > 0$ is small enough and n is large enough and this completes the proof. □

Final remarks

One can ask if Theorem 1 can be generalized in various ways. For example, can a similar theorem be proved if the CLA is allowed to use more general forms of aggregations (as in [27,32,33]), for example over tuples of elements, or over elements or tuples that satisfy some condition? Another conceivable generalization would be to allow the probability distribution to be generated by some form of probabilistic graphical model [19–21,39], as in e.g. [27,32,33], but adapted to real valued relations.

CRediT authorship contribution statement

Vera Koponen: Writing – review & editing, Writing – original draft, Conceptualization.

Data availability

No data was used for the research described in the article.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

Vera Koponen reports financial support was provided by Swedish Research Council. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

Vera Koponen thanks the anonymous referee for valuable comments including pointing out a couple of gaps in the arguments of the initial manuscript which have now been resolved. Vera Koponen was partially supported by the Swedish Research Council, grant 2023–05238_VR.

References

[1] Y. Glebskii, D. Kogan, M. Liogon'kii, V. Talanov, Range and degree of realizability of formulas in the restricted predicate calculus, Cybernetics 5 (1969) 142–154.
[2] R. Fagin, Probabilities on finite models, J. Symbolic Logic 41 (1976) 50–58.
[3] M. Albert, M. Bouvel, V. Féray, M. Noy, A logical limit law for 231-avoiding permutations, Permutation Patterns 2023 #1 26 (1) (2024).
[4] J.T. Baldwin, Expansions of geometries, J. Symbolic Logic 68 (2003) 803–827.
[5] S.N. Burris, Number Theoretic Density and Logical Limit Laws, Vol. 86, American Mathematical Socitety, 2001.
[6] S. Haber, M. Krivelevich, The logic of random regular graphs, J. Comb. 1 (2010) 389–440.
[7] C.D. Hill, On 0, 1-laws and asymptotics of definable sets in geometric fraïssé classes, Fundam. Math. 239 (2017) 201–219.
[8] K. Risto, On probabilistic elimination of generalized quantifiers, Random Struct. Algorithms 19 (2001) 1–36.
[9] H, J. Keisler, W.B. Lotfallah, Almost everywhere elimination of probability quantifiers, J. Symbolic Logic 74 (2009) 1121–1142.
[10] G. Ph, M.Y. Kolaitis, Vardi, Infinitary logics and 0–1 laws, Inform. Comput. 98 (1992) 258–294.
[11] G. Ph, H.J. Kolaitis, B.L. Prömel, Rotchild, K_{l+1}-Free graphs: asymptotic structure and a 0–1 law, Trans. Am. Math. Soc. 303 (1987) 637–671.
[12] J.F. Lynch, Almost sure theories, Ann. Math. Log. 18 (1980) 91–135.
[13] L. Gregory, Mccolm, MSO zero-one laws on random labelled acyclic graphs, Discrete Math 254 (2002) 331–347.
[14] D. Mubayi, C. Terry, Discrete metric spaces: structure, enumeration, and 0–1 laws, J. Symbolic Log. 84 (2019) 1293–1324.
[15] S. Shelah, J. Spencer, Zero-one laws for sparse random graphs, J. Am. Math. Soc. 1 (1988) 97–115.
[16] C.C. Chang, H.J. Keisler, Continuous Model Theory, Princeton University Press, 1966.
[17] I.B. Yaacov, A. Berenstein, C.W. Henson, A. Usvyatsov, Model theory for metric structures, in: Z. Chatzidakis, D. Machpherson, A. Pillay, A. Wilkie (Eds.) Model Theory With Applications to Algebra and Analysis, London Mathematical Society Lecture Notes Series, Vol. 350, 2008, pp. 315–427.
[18] F. Riguzzi, T. Swift, A survey of probabilistic logic programming, Declarative Logic Program.: Theory Syst. Appl. (2018) 185–228.
[19] G.V.D. Broeck, K. Kerstin, S. Natarajan, D. Poole (Eds.), An Introduction to Lifted Probabilistic Inference, the, MIT Press, 2021.
[20] L. De, Raedt, K. Kersting, S. Natarajan, D. Poole (Eds.), Statistical Relational Artificial Intelligence; Logic, Probability, and Computation, Morgan and Claypool Publishers, 2016.
[21] A. Kimmig, L. Mihalkova, L. Getoor, Lifted graphical models: a survey, Mach Learn 99 (2015) 1–45.
[22] E. Grädel, H. Helal, M. Naaf, R. Wilke, Zero-One laws and almost sure valuations of first-order logic in semiring semantics, Proc. 37th Annu. ACM/IEEE Symp. Logic Comput. Sci. 22 (2022) 1–12.
[23] I. Goldbring, B. Hart, A. Kruckman, The almost sure theory of finite metric spaces, Bull. London Math. Soc. 53 (2021) 1740–1748.
[24] S. Adam-Day, T.-M. Iliant, I. Ilkan, Ceylan, Zero-one laws of graph neural networks, in: Proceedings of 37th Conference on Neural Information Processing Systems, 37th Conference on Neural Information Processing Systems, 2023.
[25] S. Adam-Day, M. Benedikt, Ismail Ilkan Ceylan and Ben Finkelshtein, almost surely asymptotically constant graph neural networks, in: Proceedings of the 38th Conference on Neural Information Processing Systems, the 38th Conference on Neural Information Processing Systems, 2024.
[26] F.G. Cozman, D.D. Mauá, The finite model theory of Bayesian network specifications: descriptive complexity and zero/one laws, Int. J. Approx. Reason. 110 (2019) 107–126.
[27] M. Jaeger, Convergence results for relational Bayesian networks, in: Proceedings of the Thirteenth Annual IEEE Symposium on Logic in Computer Science (LICS 98), the Thirteenth Annual IEEE Symposium on Logic in Computer Science (LICS 98), 1998, pp. 44–55.
[28] M. Jaeger, Reasoning about infinite random structures with relational Bayesian networks, in: Proceedings of the Sixth International Conference on Principles of Knowledge Representation and Reasoning, the Sixth International Conference on Principles of Knowledge Representation and Reasoning, 1998, pp. 570–581.
[29] V. Koponen, Conditional probability logic, lifted Bayesian networks, and almost sure quantifier elimination, Theor. Comput. Sci. 848 (2020) 1–27.
[30] V. Koponen, Random expansions of finite structures with bounded degree, Ann. Pure Appl. Log. 177 (2026) 103665. https://doi.org/10.1016/j.apal.2025.103665
[31] V. Koponen, Y. Tousinejad, Random expansions of trees with bounded height, Theor. Comput. Sci. 1040 (2025) 115201. https://doi.org/10.1016/j.tcs.2025.115201
[32] V. Koponen, F. Weitkämper, Asymptotic elimination of partially continuous aggregation functions in directed graphical models, Inform. Comput. 293 (2023) 105061. https://doi.org/10.1016/j.ic.2023.105061
[33] V. Koponen, F. Weitkämper, On the relative asymptotic expressivity of inference frameworks, Logical Metods Comput. Sci. 20 (2024) 52. https://doi.org/10.46298/lmcs-20(4:13)2024
[34] F. Weitkämper, An asymptotic analysis of probabilistic logic programming, Theory Pract. Logic Program. 21 (2021) 802–817.
[35] F. Weitkämper, Functional lifted Bayesian networks: statistical relational learning and reasoning with relative frequencies, inductive logic programming. ILP 2022, Lect. Notes Comput. Sci. 13779 (2024) 142–156.
[36] J. Lukasiewicz, A. Tarski, Untersuchungen uber den aussagenkalkul, Comptes Rendus S{́e}ances Soci{́e}t{́e} Sci. Lett. Varsovie 23 (1930) 30–50.
[37] N. Alon, J.H. Spencer, The Probabilistic Metod, Second Edition, John Wiley & Sons, 2000.
[38] H. Chernoff, A measure of the asymptotic efficiency for tests of a hypothesis based on the sum of observations, Ann. Math. Stat. 23 (1952) 493–509.
[39] D. Koller, N. Friedman, Probabilistic Graphical Models: Principles and Techniques, MIT Press, 2009.
[40] H.-D. Ebbinghaus, J. Flum, Finite Model Theory, Second Edition, Springer, 1999.