Logical Methods in Computer Science
Volume 20, Issue 4, 2024, pp. 13:1–13:52
https://lmcs.episciences.org/

Submitted May 16, 2022
Published Nov. 12, 2024

ON THE RELATIVE ASYMPTOTIC EXPRESSIVITY OF INFERENCE FRAMEWORKS *

VERA KOPONEN [a] AND FELIX WEITKÄMPER [b]

[a] Department of Mathematics, Uppsala University, Sweden.
e-mail address: vera.koponen@math.uu.se

[b] Institut für Informatik, Ludwig-Maximilians-Universität München, Munich, Germany.
e-mail address: felix.weitkaemper@lmu.de

Abstract. We consider logics with truth values in the unit interval $[0,1]$. Such logics are used to define queries and to define probability distributions. In this context the notion of almost sure equivalence of formulas is generalized to the notion of asymptotic equivalence. We prove two new results about the asymptotic equivalence of formulas where each result has a convergence law as a corollary. These results as well as several older results can be formulated as results about the relative asymptotic expressivity of inference frameworks. An inference framework $\mathbf{F}$ is a class of pairs $(\mathbb{P}, L)$, where $\mathbb{P} = (\mathbb{P}_n : n = 1, 2, 3, \ldots)$, $\mathbb{P}_n$ are probability distributions on the set $\mathbf{W}_n$ of all σ-structures with domain $\{1, \ldots, n\}$ (where σ is a first-order signature) and L is a logic with truth values in the unit interval $[0,1]$. An inference framework $\mathbf{F}'$ is asymptotically at least as expressive as an inference framework $\mathbf{F}$ if for every $(\mathbb{P}, L) \in \mathbf{F}$ there is $(\mathbb{P}', L') \in \mathbf{F}'$ such that $\mathbb{P}$ is asymptotically total variation equivalent to $\mathbb{P}'$ and for every $\varphi(\bar{x}) \in L$ there is $\varphi'(\bar{x}) \in L'$ such that $\varphi'(\bar{x})$ is asymptotically equivalent to $\varphi(\bar{x})$ with respect to $\mathbb{P}$. This relation is a preorder. If, in addition, $\mathbf{F}$ is at least as expressive as $\mathbf{F}'$ then we say that $\mathbf{F}$ and $\mathbf{F}'$ are asymptotically equally expressive. Our third contribution is to systematize the new results of this paper and several previous results in order to get a preorder on a number of inference systems that are of relevance in the context of machine learning and artificial intelligence.

1. Introduction

In modern artificial intelligence, logics with (truth) values in the unit inverval $[0,1]$ are used not only as query languages, but also to define the probability distributions with respect to which queries are evaluated. Such probability distributions can be defined by formalisms called *probabilistic graphical models (PGMs)* which are determined by a finite graph (with a vertex for each relation symbol) and formulas that express (conditional) probabilities of individual relations. We can view a formula $\varphi(x_1 \ldots, x_n)$ in a probability logic $L(\sigma)$ over a

Key words and phrases: probability logic, statistical relational artificial intelligence, probabilistic graphical models, scalable inference, finite model theory, logical convergence laws, logical expressivity.

* ACM Class F.4.1. The research leading to this publication was partially conducted during a visit of the second author to the University of Uppsala, enabled by LMUExcellent and funded by the Federal Ministry of Education and Research (BMBF) and the Free State of Bavaria under the Excellence Strategy of the Federal Government and the Länder.

LOGICAL METHODS IN COMPUTER SCIENCE
DOI:10.46298/LMCS-20(4:13)2024

signature σ as a function taking a σ-structure $\mathcal{A}$ and a tuple $a_1, \ldots, a_n \in A$ as input and returning a real number in $[0, 1]$ as output (see Definition 4.1); this number will be denoted $\mathcal{A}(\varphi(a_1, \ldots, a_n))$ and called the *value* of $\varphi(a_1, \ldots, a_n)$ in $\mathcal{A}$. Then for any finite relational σ we can define an $L(\sigma)$*-network* as a directed acyclic graph with vertex set σ, and for every vertex $R \in \sigma$ of arity n, an $L(\sigma)$-formula $\theta_R(x_1, \ldots x_n)$ using only relation symbols among the parents of R in the directed acyclic graph (see Definition 5.6). For any relation $R \in \sigma$, any finite domain A and tuple $\bar{a}$ of elements from A such that its length matches the arity of R, the conditional probabilitity of $R(a_1, \ldots, a_n)$, given the interpretation of all the parents of R, is specified by the value of $\theta_R(a_1, \ldots, a_n)$. Let σ be a finite and relational signature and let $\mathbf{W}_n$ denote the set of all σ-structures with domain $[n] = \{1, \ldots, n\}$. In the way indicated above each $L(\sigma)$-network now defines a probability distribution on $\mathbf{W}_n$ for each n (see Definition 5.7).

We will mostly work with a logic that we call PLA^+, or $PLA^+(\sigma)$ if we want to indicate the signature used, which uses aggregation functions instead of quantifiers (see Definition 3.2). In this logic we can express all queries (on finite structures) that can be expressed by first-order logic because the aggregation functions max and min can be used to express existential and universal quantification. Of course, there are many more aggregation functions, for example the average, also called arithmetic mean (of a finite sequence of reals from $[0, 1]$). With such aggregation functions PLA^+ can express, for example, each stage of the definitions of PageRank and SimRank (more about this in Example 4.10). We will use PLA^+ both as a query language and as a language to specify probability distributions on $\mathbf{W}_n$ via $PLA^+(\sigma)$-networks.

PLA^+ is a very expressive query language, and PLA^+-networks can define a great variety of probability distributions. For this reason we do not expect to be able to prove general theorems for all PLA^+-queries and all PLA^+-networks. Therefore we look for sublogics, say L and L', of PLA^+ such that (a) these logics are expressive enough to be useful, and (b) we can prove general results for probability distributions defined by L-networks and queries expressed by L'. These sublogics will be defined mainly by restricting PLA^+ to formulas that only use aggregation functions that have certain "nice" properties. The most extreme case is to ban all aggregation functions, so let us call a PLA^+-formula which does not use any aggregation function *aggregation-free*.

Let $\mathbb{G}$ denote a $PLA^+(\sigma)$-network and let $\mathbb{P}_n$ denote the probability distribution on $\mathbf{W}_n$ that is determined by $\mathbb{G}$. Our first main result, Theorem 5.11, is that if every formula associated to $\mathbb{G}$ contains only continuous aggregation functions (in the sense of Definition 3.5), then every $PLA^+(\sigma)$-formula that contains only continuous aggregation functions is asymptotically equivalent to an aggregation-free formula with respect to the sequence of distributions $\mathbb{P} = (\mathbb{P}_n : n =, 1, 2, 3, \ldots)$. The notion of "asymptotic equivalence" (Definition 5.4) is a generalization of the notion of "almost sure equivalence" that makes sense for logics with more than two truth values. Intuitively speaking, two formulas are asymptotically equivalent if, with high probability, their values are almost the same. Moreover, as stated in Corollary 5.13, if $\varphi(\bar{x})$ is a $PLA^+(\sigma)$ formula and $\mathbb{G}$ is as above then $\varphi(\bar{x})$ satisfies a convergence law with respect to $\mathbb{P}$. In the more technical Corollary 7.21 we show how the mentioned results can be used to approximate probabilities of queries without any reference to the domain size.

Our second main result still considers a $PLA^+(\sigma)$-network $\mathbb{G}$ such that all associated formulas use only continuous aggregation functions. However, as query language, we use the two-valued *conditional probability logic (CPL)* introduced in [Kop20]. CPL is an extension of first-order logic which allows constructions that can express statements like "the relative frequency of $\bar{x}$ that satisfy $\varphi_1(\bar{x})$ among $\bar{x}$ that satisfy $\varphi_2(\bar{x})$ is at least (a constant) c, or alternatively, is at least as large as c plus the relative frequency of $\bar{x}$ that satisfy $\psi_1(\bar{x})$ among $\bar{x}$ that satisfy $\psi_2(\bar{x})$" where c is a non-negative real. Here the conditions expressed by $\varphi_1(\bar{x})$ et cetera may themselves be CPL-formulas, so this construction can be nested and viewed as a kind of quantification. The second main result, Theorem 8.6, is that if the $CPL(\sigma)$-formula $\varphi(\bar{x})$ is *safe (with respect to* $\mathbb{G}$*)* then $\varphi(\bar{x})$ is almost surely equivalent to a quantifier-free first-order formula. The condition that $\varphi(\bar{x})$ is *safe* roughly means that in every relative frequency statement as described above that occurs in $\varphi(\bar{x})$, the constant c does not belong to a certain finite set of numbers that is determined by the syntactic structure of $\varphi(\bar{x})$ and by $\mathbb{G}$. We obtain two corollaries. The first is a convergence law for safe CPL-formulas. The second roughly says that there is a $PLA^+(\sigma)$-network $\tilde{\mathbb{G}}$ such that all associated PLA^+- formulas are aggregation-free, and for every safe CPL-formula $\varphi(\bar{x})$ and for sufficiently large n, the probability that a sequence of parameters satisfies $\varphi(\bar{x})$ under the distribution $\tilde{\mathbb{P}}_n$ determined by $\tilde{\mathbb{G}}$ approximates the corresponding probability under $\mathbb{P}_n$ arbitrarily closely. This implies that the first probability can be estimated, with as high accuracy as we like for large enough n, in time which is independent from the domain size.

Our third contribution, Theorem 9.7 illustrated by Figure 1, is to systematize the new results of this article and those in [Kop20, KW23, SS88] by means of the notions of inference framework and relative asymptotic expressivity of inference frameworks. [1] An *inference framework* (for a signature σ) is a class $\mathbf{F}$ of pairs $(\mathbb{P}, L)$ where $\mathbb{P} = (\mathbb{P}_n : n \in \mathbb{N}^+)$, each $\mathbb{P}_n$ is a probability distribution on $\mathbf{W}_n$, and L is a logic (the associated query language) which uses the signature σ (see Definition 9.1). Note that we allow L to depend on $\mathbb{P}$. The reason is that it allows us make finer distinctions between queries which are "easy" to evaluate with respect to a given sequence $\mathbb{P}$ of probability distributions and queries which are "hard" to evaluate with respect to the same $\mathbb{P}$.

The asymptotic expressivity of an inference framework should now be studied on both components of its pairs $(\mathbb{P}, L)$. We call an inference framework $\mathbf{F}'$ *asymptotically at least as expressive* as another inference framework $\mathbf{F}$ if for every $(\mathbb{P}, L) \in \mathbf{F}$ there is $(\mathbb{P}', L') \in \mathbf{F}'$ such that $\mathbb{P}$ is asymptotically total variation equivalent to $\mathbb{P}'$ (see Definition 5.3) and for every $\varphi(\bar{x}) \in L$ there is $\varphi'(\bar{x}) \in L'$ such that $\varphi'(\bar{x})$ is asymptotically equivalent to $\varphi(\bar{x})$ with respect to $\mathbb{P}$ (or equivalently $\mathbb{P}'$, see Definition 9.2). If in addition, $\mathbf{F}$ is asymptotically at least as expressive as $\mathbf{F}'$ then we say that they are *asymptotically equally expressive.*

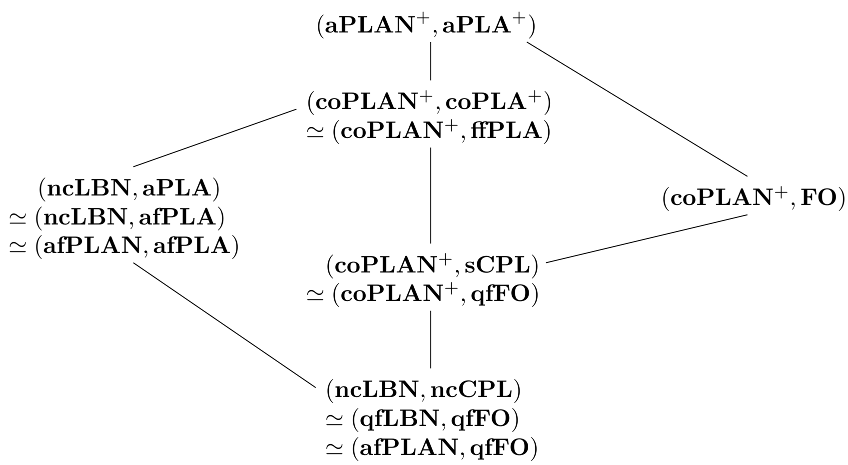
In the discussion above we have already implicitly seen examples of inference frameworks and results stating that two inference frameworks are asymptotically equally expressive. For example, if $coPLA^+$ is the set of all PLA^+-formulas that use only continuous aggregation functions and $\mathbf{F}$ is the set of all pairs $(\mathbb{P}, coPLA^+)$ such that $\mathbb{P}$ is defined by a $coPLA^+$-network, then Theorem 5.11 implies that $\mathbf{F}$ is asymptotically equally expressive as the inference framework $\mathbf{F}'$ consisting of all pairs $(\mathbb{P}', afPLA)$ where $\mathbb{P}'$ is defined by a $coPLA^+$-network and $afPLA$ is set of all aggregation-free formulas in PLA^+.

Naturally one can question whether our notion of "asymptotically at least as expressive" is the most relevant one. In particular, the asumption about asymptotic total variation equivalence may appear to be too strong, since the logic(s) used in a particular inference framework may not be able to "define" all subsets of the probability space $\mathbf{W}_n$. But we

[1]We did not include the inference framework to which the main results in [Jae98] apply because we have not determined where it belongs relative to the inference frameworks of Theorem 9.7 or Figure 1, but we suspect that it belongs to the equivalence class in the bottom of Figure 1.

believe that a useful notion of "asymptotically at least as expressive" should be *transitive* and all other seemingly reasonable candidates that we have considered (except the one mentioned above and in Definition 9.2 below) turned out *not* to be transitive.

Figure 1: The inference frameworks are defined in Definition 9.6. The symbol $\simeq$ denotes asymptotic equal expressivity. A path upwards means that the the upper inference framework is asymptotically more expressive (i.e. $\prec$ holds). The absence of a path "upwards" between two inference framewords means that the inference frameworks are incomparable with respect to asymptotic expressivity.



Related work. Researchers in statistical relational artificial intelligence, a branch of artificial intelligence and machine learning (see e.g. [DRKNP16, GT07, KMG15, GVdBP21] for introductions to the field), have developed several different formalisms to specify probabilistic graphical models on an abstract relational level. They include relational Bayesian networks [Jae97], relational logistic regression [KBK$^+$14], Bayesian logic programs [KDR07], and lifted Bayesian networks [Kop20]. Each of these formalisms can be viewed as a $L(\sigma)$-network for a suitable logic L and signature σ. In fact, each of the mentioned formalisms can be viewed as a $PLA^+(\sigma)$-network. Besides notational differences, the main difference between PLA^+ and the probability logic in [Jae98] is that PLA^+ allows for more general connectives.

As before, let σ be a finite and relational signature, $\mathbf{W}_n$ the set of all σ-structures with domain $[n]$, and $\mathbb{P}_n$ a probability distribution on $\mathbf{W}_n$. Jaeger's main result in [Jae98] can now be formulated as saying that if $ecPLA^+$ is set of PLA^+-formulas which use only *exponentially convergent* aggregation functions and if $\mathbb{P}_n$ is defined by an $ecPLA^+(\sigma)$-network, then the probability of every first-order query converges as $n \to \infty$. The aggregation functions noisy-or, maximum and minimum are exponentially convergent, but not the arithmetic or geometric means. On the other hand the arithmetic and geometric means as well as (for $\alpha > 0$) $length_\alpha$ are continuous, but not exponentially convergent, where for every sequence of reals $\bar{r} = (r_1, \ldots, r_m)$, $length_\alpha(\bar{r}) = m^{-\alpha}$. Therefore Theorem 5.11 and its corollaries apply to different sequences of probability distributions and different queries than the main result in [Jae98].

The main results in [KW23] can be stated as follows: Suppose that $\mathbb{P}_n$ is defined by a lifted Bayesian network for σ (as defined in [Kop20]) and let $aPLA(\sigma)$ be the set of

$PLA^+(\sigma)$-formulas in which only *admissible* aggregation functions are allowed and which also satisfies two other conditions (see Definitions 3.5 and 4.14). Then every $aPLA(\sigma)$-formula is *asymptotically equivalent* to an aggregation-free formula, and every $aPLA(\sigma)$-formula satisfies a convergence law (as $n \to \infty$). Admissibility is a weak form of continuity (so continuity implies admissibility) satisfied by the aggregation functions maximum and minimum. Every sequence of distributions defined by a lifted Bayesian network is asymptotically total variation equivalent to a sequence of distributions defined by a $coPLA^+$-network, but not vice versa. So Theorem 5.11 is, with respect to probability distributions, more general than the main result in [KW23], but not with respect to queries. Moreover, Theorem 5.11 cannot be generalized to hold for $aPLA$-queries, or for first-order queries. The reason is as follows. First, the aggregation function $length_\alpha$ is continuous, so if R is a binary relation symbol and $\alpha \in (0, 1)$, then there is a $coPLA^+$-network such that the formula $R(x, y)$ has probability $n^{-\alpha}$ (for any pair, independently of other pairs) where n is the domain size. By results on random graphs by Shelah and Spencer [SS88], if α is rational then there is a first-order sentence such that its probability does not converge as $n \to \infty$. Since every first-order formula is equivalent to some $aPLA$-formula we cannot have convergence for all $aPLA$-formulas.

From [Kop20, Theorem 3.16] it follows that lifted Bayesian networks only can define probability distributions $\mathbb{P}_n$ where for each $R \in \sigma$ the probability of $R(\bar{x})$ is either constantly 0 or tends to a constant $c > 0$ as $n \to \infty$. It seems like the same is true for relational Bayesian networks as in [Jae98] for the following reason: Suppose for a contradiction that a sequence of probability distributions $\mathbb{P}_n$ is defined by a relational Bayesian network and that the probability that a pair of parameters satisfies an atomic formula $R(x, y)$ is $n^{-\alpha}$ (independently of other pairs) where $\alpha \in (0, 1)$ is rational and n is the size of the domain. By[SS88] again there is a first-order sentence such that its probability does not converge as $n \to \infty$, but this contradicts the main result in [Jae98].

In [Kop20] Koponen studied so-called lifted Bayesian networks, for producing probability distributions, and a fragment of conditional probability logic (CPL) which extends first-order logic and avoids certain "critical" parameters as a query language. Theorem 8.6 considers another fragment of CPL and more general, up to asymptotic total variation equivalence, probability distributions. The fragment of CPL considered in Theorem 8.6 avoids first-order quantifiers and "unsafe" parameters in "conditional proportion statements".

Other works than [Jae98, Kop20, KW23] about the asymptotics of logics with respect to probability distributions defined by probabilistic graphical models are less related to the results presented here. They include the following publications: [CM19, MBGS19, PBK$^+$14, Wei21, Wei24, FWF23].

Organization. This article is organized as follows. Section 2 fixes some basic notation and terminology, and states a result that will be used later. Section 3 defines the notions of *connective* and *aggregation function* that we will use and also defines the notion of *(strongly) admissible* aggregation function. Section 4 defines the general notion of a *logic* that we will use, as well as the particular logics that will be considered later. Section 5 defines the notion of $L(\sigma)$-network and states our first main result, Theorem 5.11, and a corollary. It also recalls the notion of lifted Bayesian network from [Kop20] and the main results from [Kop20] and [KW23] (which are used in Theorem 9.7). Section 6 proves a key technical result about asymptotic elimination of strongly admissible aggregation functions, Proposition 6.4, which is used in Section 7 to prove Theorem 5.11. Section 7 completes the proof of Theorem 5.11 by, roughly speaking, proving that the preconditions of Proposition 6.4 are satisfied. In

Section 8 we introduce the notion of a *safe CPL*-formula and state and prove our second main result, Theorem 8.6. Then Section 9 defines the notions of *inference framework* and *asymptotically at least as expressive (inference framework)* and states and proves our last result, Theorem 9.7. Finally, Section 10 briefly recalls what we have done.

2. Preliminaries

We let $\mathbb{N}$ be the set of all non-negative integers and we let $\mathbb{N}^+$ be the set of positive integers. Finite sequences/tuples of objects are denoted by $\bar{a}, \bar{b}, \bar{r}, \bar{x}$, et cetera. (Typically the objects of the sequence are numbers, elements from the domain of a structure or logical variables.)

For a finite sequence $\bar{a}$, $|\bar{a}|$ denotes the length of the sequence and $\text{rng}(\bar{a})$ denotes the set of elements occurring in the sequence $\bar{a}$. For a set A, $|A|$ denotes its cardinality. If A is a set and $n \in \mathbb{N}^+$, then A^n is the set of all finite sequences of length n of elements from A and $A^{<\omega}$ is the set of all finite nonempty sequences of elements from A, so $A^{<\omega} = \bigcup_{n \in \mathbb{N}^+} A^n$ The set $\{1, \ldots, n\}$ will be denoted by $[n]$.

Structures in the sense of first-order logic will be denoted by calligraphic letters such as $\mathcal{A}, \mathcal{B}, \ldots$. Unless otherwise specified, their domains/universes are denoted by the corresponding non-calligraphic letter $A, B, \ldots$ (see for example [EF99] for basics about structures in the sense of first-order logic and (finite) model theory). We will consider different logics, but the semantics are always based on first-order structures. If σ is a (first-order) *signature*, also called *vocabulary*, then a *σ-structure* is a first-order structure in which all symbols in σ have an interpretation (and, for counting reasons, we assume that no other symbols are interpreted in a σ-structure). A signature σ is *finite and relational* if it is finite and contains only relation symbols. If R is a relation symbol of a signature σ, its intepretation in a σ-structure $\mathcal{A}$ is denoted by $R^{\mathcal{A}}$. If $\sigma' \subset \sigma$ are signatures and $\mathcal{A}$ is a σ-structure then the reduct of $\mathcal{A}$ to σ' is denoted by $\mathcal{A} \restriction \sigma'$. For more about notation concerning logical concepts, see Section 4.

By *directed acyclic graph (DAG)* we mean a directed graph without loops or directed cycles. Let $G = (V, G)$ be a DAG. If $v \in V$ then $\text{par}(v)$ denotes the set of *parents of* v, that is, the set of all vertices $w \in V$ such that $(w, v) \in E$. For $v \in V$ we define the *maximal path rank* of v, denoted $\text{mp}(v)$, to be the maximal integer $n > 0$ such that there is a directed path $v_0, \ldots, v_n \in V$ (meaning that $(v_i, v_{i+1}) \in E$ for all i) with $v_n = v$. We define the *maximal path rank* of $\mathcal{G}$, denoted $\text{mp}(\mathcal{G})$, as $\text{mp}(\mathcal{G}) = \max(\text{mp}(v) : v \in V)$.

We call a random variable *binary* if it can only take the value 0 or 1. The following is a direct consequence of [AS00, Corollary A.1.14] which in turn follows from the Chernoff bound [Che52]:

Lemma 2.1. *Let Z be the sum of n independent binary random variables, each one with probability p of having the value 1. 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 $2e^{-c_\varepsilon pn}$.*

Corollary 2.2. *Let $p \in [0, 1]$ and let $\varepsilon > 0$. Let Z be the sum of n independent binary random variables $Z_1, \ldots, Z_n$, where for each $i = 1, \ldots, n$ the probability that Z_i equals 1 belongs to the interval $[p - \varepsilon, p + \varepsilon]$. Then there is $c_\varepsilon > 0$, depending only on ε, such that the probability that $Z > (1 + \varepsilon)(p + \varepsilon)n$ or $Z < (1 - \varepsilon)(p - \varepsilon)n$ is less than $2e^{-c_\varepsilon pn}$.*

Proof. Let $Z'_1, \ldots, Z'_n$ be independent binary random variables where each Z'_i takes value 1 with probability (exactly) $p + \varepsilon$ (assuming $p + \varepsilon \leq 1$). Let $Z' = Z'_1 + \ldots + Z'_n$. By Lemma 2.1, the probability that $Z' > (p + \varepsilon)n + \varepsilon(p + \varepsilon)n = (1 + \varepsilon)(p + \varepsilon)n$ is less than $2e^{-a_\varepsilon pn}$ where

$a_\varepsilon > 0$ depends only on ε. Since, for each i, the probability that $Z_i = 1$ is less or equal to the probability that $Z'_1 = 1$ it follows that the probability that $Z > (1+\varepsilon)(p+\varepsilon)n$ is not larger than the probability that $Z' > (1+\varepsilon)(p+\varepsilon)n$ which is less than $2e^{-a_\varepsilon pn}$. A similar argument shows that the probability that $Z < (1-\varepsilon)(p-\varepsilon)n$ is less than $2e^{-b_\varepsilon pn}$ for some $b_\varepsilon > 0$ depending only on ε. □

3. Connectives and aggregation functions

The idea behind a k-ary connective is that it assigns a truth value to every k-tuple of truth values. Recall from the previous section that $[0,1]^{<\omega}$ is the set of all finite sequences $\bar{r}$ where each entry of $\bar{r}$ belongs to $[0,1]$. So $\big([0,1]^{<\omega}\big)^k$ is the set of all k-tuples $(\bar{r}_1, \ldots, \bar{r}_k)$ where, for each $i = 1, \ldots, k$, $\bar{r}_i$ is a finite sequence of reals from $[0,1]$. Note that $\bar{r}_i$ and $\bar{r}_j$ are allowed to have different length if $i \neq j$. The role of a k-ary aggregation function is to assign a truth value to every k-tuple $(\bar{r}_1, \ldots, \bar{r}_k) \in \big([0,1]^{<\omega}\big)^k$. In particular a unary (1-ary) aggregation function takes only one sequence $\bar{r} \in [0,1]^{<\omega}$ (of arbitrary finite length) as input. When used in a logic, in Section 4, the role of aggregation functions will be to "aggregate over a domain", similarly to (generalized) quantifiers in the context of 0/1-valued logics. Alternatively (as in e.g. [Jae98]), one can view a k-ary aggregation (or "combination") function as a mapping from k-tuples of finite multisets (of reals in $[0,1]$) into $[0,1]$. However, the "symmetry condition" in Definition 3.2 below implies that our notion of aggregation function is exchangeable, in the context of this article, with the notion of an aggregation function as operating on multisets.

Definition 3.1. A function $\mathsf{C} : [0,1]^k \to [0,1]$ where $k \in \mathbb{N}^+$ will also be called a *connective*.

Definition 3.2. Let $F : \big([0,1]^{<\omega}\big)^k \to [0,1]$, so F takes k sequences (not necessarily of the same length) as input. We call F an *aggregation function* if F is symmetric in the sense that if $\bar{r}_1, \ldots, \bar{r}_k \in [0,1]^{<\omega}$ and for each $i = 1, \ldots, k$, $\bar{\rho}_i$ is an arbitrary reordering of the entries of $\bar{r}_i$, then $F(\bar{\rho}_1, \ldots, \bar{\rho}_k) = F(\bar{r}_1, \ldots, \bar{r}_k)$.

Example 3.3. (a) The aggregation functions listed below are common when analyzing data. For $\bar{r} = (r_1, \ldots, r_n) \in [0,1]^{<\omega}$, define
(1) $\max(\bar{r})$ to be the *maximum* of all r_i,
(2) $\min(\bar{r})$ to be the *minimum* of all r_i,
(3) $\mathrm{am}(\bar{r}) = (r_1 + \ldots + r_n)/n$, so 'am' is the *arithmetic mean*.
(4) $\mathrm{gm}(\bar{r}) = \big(\prod_{i=1}^n r_i\big)^{(1/n)}$, so 'gm' is the *geometric mean*.
(5) $\text{noisy-or}(\bar{r}) = 1 - \prod_{i=1}^n (1 - r_i)$.

(b) Another example of an aggregation function (now with arity 2) is the pseudometric $\mu_1^u : \big([0,1]^{<\omega}\big)^2 \to [0,1]$ of Definition 3.9 which compares how similar (or close) two sequences $\bar{r}, \bar{\rho} \in [0,1]^{<\omega}$ are if we disregard the ordering of the entries in each sequence.

(c) For another example, define $S(x) = (1 + e^{-x})^{-1}$ for all $x \in \mathbb{R}$, let $w_1, \ldots, w_k \in [0,1]$ be weights such that their sum equals 1, and define $G : \big([0,1]^{<\omega}\big)^k \to [0,1]$ by $G(\bar{r}_1, \ldots, \bar{r}_k) = S\big(\sum_{i=1}^k w_i \cdot \mathrm{am}(\bar{r}_i)\big)$. G is used to define Bayesian networks in the context of *Domain-Size-Aware Relational Logistic Regression* models [FWF23].

(d) For $\alpha \in (0,1)$ define $length_\alpha : [0,1]^{<\omega} \to [0,1]$ by $length_\alpha(\bar{r}) = 1/|\bar{r}|^\alpha$ for all $\bar{r} \in [0,1]^{<\omega}$ (so $length_\alpha(\bar{r})$ depends only on the length of $\bar{r}$). $length_\alpha$ can, for example, be used to define a $coPLA^+(\sigma)$-network where the signature σ contains only a binary relation symbol such that this network induces a probability distribution on directed graphs where each directed edge has probability $1/n^\alpha$, independently of other edges, where n is the number of vertices.

(e) More examples (e.g. conditional arithmetic means) are given in Section 5 of [KW23].

Now we will isolate two classes of aggregation functions that are sufficiently benign that we can prove results concerning the asymptotic expressivity of inference frameworks that use only such aggregation functions. Later we will see that the two classes contain useful aggregation functions. Before defining these two classes we need to define the notion of convergence testing sequence of sequences from $[0,1]^{<\omega}$.

Intuitively speaking, an infinite sequence (of finite sequences) $\bar{r}_n \in [0,1]^{<\omega}$, $n \in \mathbb{N}$, is convergence testing if there are k and $c_1, \dots, c_k, \alpha_1, \dots \alpha_k \in [0,1]$ such that, as $n \to \infty$, all entries in $\bar{r}_n$ congregate ever closer to the "convergence points" $c_1, \dots, c_k$ and, for each i, the proportion of entries in $\bar{r}_n$ that are close to c_i tends ever closer to α_i.

Definition 3.4 (Convergence testing sequence)**.** A sequence $\bar{r}_n \in [0,1]^{<\omega}$, $n \in \mathbb{N}$, is called *convergence testing* for parameters $c_1, \dots, c_k \in [0,1]$ and $\alpha_1, \dots \alpha_k \in [0,1]$ if the following hold, where $r_{n,i}$ denotes the ith entry of $\bar{r}_n$:

(1) $|\bar{r}_n| < |\bar{r}_{n+1}|$ for all $n \in \mathbb{N}$.
(2) For every disjoint family of open (with respect to the induced topology on $[0,1]$) intervals $I_1, \dots I_k \subseteq [0,1]$ such that $c_i \in I_i$ for each i, there is an $N \in \mathbb{N}$ such that $\mathrm{rng}(\bar{r}_n) \subseteq \bigcup_{j=1}^{k} I_j$ for all $n \geq N$, and for every $j \in \{1, \dots, k\}$,

$$\lim_{n\to\infty} \frac{|\{i \leq |\bar{r}_n| : r_{n,i} \in I_j\}|}{|\bar{r}_n|} = \alpha_j.$$

More generally, a sequence of m-tuples of sequences $(\bar{r}_{1,n}, \dots, \bar{r}_{m,n}) \in \big([0,1]^{<\omega}\big)^m$, $n \in \mathbb{N}$, is called *convergence testing* for parameters $c_{i,j} \in [0,1]$ and $\alpha_{i,j} \in [0,1]$, where $i \in \{1, \dots, m\}$, $j \in \{1, \dots, k_i\}$ and $k_1, \dots k_m \in \mathbb{N}^+$, if for every fixed $i \in \{1, \dots, m\}$ the sequence $\bar{r}_{i,n}$, $n \in \mathbb{N}$, is convergence testing for $c_{i,1}, \dots, c_{i,k_i}$, and $\alpha_{i,1}, \dots, \alpha_{i,k_i}$.

Next we define the notion of (strongly) admissible aggregation function. The intuition behind admissibility and strong admissibility is that they are "(partial) continuity conditions" suitable for aggregation functions. We would argue that strong admissibility is a more "continuity-like" condition than admissibility because the aggregation functions max and min are admissible but not strongly admissible and one could argue that max and min, as aggregation functions, should not be considered to be continuous: If n is large then it is reasonable to view the sequences $(r_1, \dots, r_n)$ and $(\rho_1, \dots, \rho_n)$ as very similar (or "close" to each other) if $r_i = 0$ for all $i = 1, \dots, n$, $\rho_i = 0$ for $i = 1, \dots, n-1$ and $\rho_n = 1$, but $\max(\bar{r}) = 0$ and $\max(\bar{\rho}) = 1$.

Definition 3.5 (Admissibility and continuity)**.** (i) An aggregation function $F : \big([0,1]^{<\omega}\big)^m \to [0,1]$ is called *strongly admissible*, or *continuous*, if the following two conditions hold:

(1) For all $n_1, \dots, n_m \in \mathbb{N}^+$, F is continuous on the set $[0,1]^{n_1} \times \dots \times [0,1]^{n_m}$.
(2) For all convergence testing sequences of tuples $(\bar{r}_{1,n}, \dots, \bar{r}_{m,n}) \in \big([0,1]^{<\omega}\big)^m$, $n \in \mathbb{N}$, and $(\bar{\rho}_{1,n}, \dots, \bar{\rho}_{m,n}) \in \big([0,1]^{<\omega}\big)^m$, $n \in \mathbb{N}$, with the same parameters $c_{i,j}, \alpha_{i,j} \in [0,1]$, $\lim_{n\to\infty} |F(\bar{r}_{1,n}, \dots, \bar{r}_{m,n}) - F(\bar{\rho}_{1,n}, \dots, \bar{\rho}_{m,n})| = 0$.

(ii) An aggregation function $F : \left([0,1]^{<\omega}\right)^m \to [0,1]$ is called *admissible* if condition (1) above holds and condition (2) above holds whenever the parameters $\alpha_{i,j}$ are *positive* for all i and j.

The next proposition and example show that a number of useful aggregation functions are indeed (strongly) admissible. Proposition 3.6 below was proved in [KW23], because the proof in [KW23] that the arithmetic and geometric means are admissible still works if one allows the parameters $\alpha_{i,j}$ to be 0, thus showing that the functions are strongly admissible.

Proposition 3.6. *(i) The functions* am *(*arithmetic mean*) and* gm *(*geometric mean*) are strongly admissible (in other words, continuous).*
(ii) The functions max *and* min *are admissible.*

Example 3.7 (More (strongly) admissible aggregation functions)**.** The aggregation functions in parts (b), (c) and (d) in Example 3.3 are strongly admissible. In the case of (c) and (d) this follows easily from their definitions and the fact that arithmetic mean is strongly admissible. In the case of (b) it follows from the characterization of strong admissibility given below by Definition 3.11 and Proposition 3.12. In Sections 5 and 6 of [KW23] more examples of admissible aggregation functions are given, for example a "conditional arithmetic mean".

It is not hard to see that max and min are not strongly admissible and that noisy-or is not even admissible.

The above given definition of (strong) admissibility is fairly straightforward and natural, as well as useful for proving that some functions are (strongly) admissible. But we do not see how it can be used directly in the proof of Proposition 6.4 which is used to prove one of our main results, Theorem 5.11. Therefore we give a different characterization of (strong) admissibility (Definition 3.11) below. For this we need to consider functional representations of sequences in $[0,1]^{<\omega}$ and two pseudometrics on $[0,1]^{<\omega}$.

Definition 3.8 (Functional representations of sequences)**.** Let $n \in \mathbb{N}^+$ and let $\bar{r} = (r_1, \ldots, r_n) \in [0,1]^n$. We will associate a function from $[0,1]$ to $[0,1]$ with $\bar{r}$ in two different ways, one way where the order of the entries in $\bar{r}$ matters and one in which the order does not influence the associated function.

(1) Define $\mathfrak{f}_{\bar{r}}$, which we call the *ordered functional representation of* $\bar{r}$, as follows: For every $a \in [0, 1/n)$, let $\mathfrak{f}_{\bar{r}}(a) = r_1$, for every $i = 1, \ldots, n-1$ and every $a \in [i/n, (i+1)/n)$, let $f(a) = r_{i+1}$ and finally let $f(1) = r_n$.
(2) Define $\mathfrak{g}_{\bar{r}}$, which we call the *unordered functional representation of* $\bar{r}$, as follows: Let $\bar{\rho} = (\rho_1, \ldots, \rho_n)$ be a reordering of $\bar{r}$ such that, for all $i = 1, \ldots, n-1$, $\rho_i \leq \rho_{i+1}$ and let $\mathfrak{g}_{\bar{r}} = \mathfrak{f}_{\bar{\rho}}$.

Definition 3.9 (Pseudometrics on sequences)**.** (1) First we recall the L_1 and L_∞ norms: for every (bounded and integrable) $f : [0,1] \to \mathbb{R}$ they are defined as

$$\|f\|_1 = \int_{[0,1]} |f(x)| dx \qquad \text{and} \qquad \|f\|_\infty = \sup\{|f(a)| : a \in [0,1]\}.$$

(2) For $\bar{r}, \bar{\rho} \in [0,1]^{<\omega}$ we define

$$\mu_1^u(\bar{r}, \bar{\rho}) = \|\mathfrak{g}_{\bar{r}} - \mathfrak{g}_{\bar{\rho}}\|_1,$$
$$\mu_\infty^o(\bar{r}, \bar{\rho}) = \|\mathfrak{f}_{\bar{r}} - \mathfrak{f}_{\bar{\rho}}\|_\infty.$$

(3) For arbitrary $k > 1$ we can define a function on $\big([0,1]^{<\omega}\big)^k$, also denoted μ_1^u and μ_∞^o (to avoid making notation more complicated), as follows: For all $(\bar{r}_1, \ldots, \bar{r}_k), (\bar{r}'_1, \ldots, \bar{r}'_k) \in \big([0,1]^{<\omega}\big)^k$ let

$$\mu_1^u\big((\bar{r}_1, \ldots, \bar{r}_k), (\bar{r}'_1, \ldots, \bar{r}'_k)\big) = \max\big(\mu_\infty^u(\bar{r}_1, \bar{r}'_1), \ldots, \mu_1^u(\bar{r}_k, \bar{r}'_k)\big)$$

and similarly for μ_∞^o.

From well-known results in analysis it follows that μ_1^u and μ_∞^o are symmetric and satisfy the triangle inequality so they are pseudometrics on $[0,1]^{<\omega}$. It is easy to see that none of them is a metric since it can happen that $\mu_1^u(\bar{r}, \bar{\rho}) = 0$ and $\bar{r} \neq \bar{\rho}$. For example, if $\bar{r} = (0, 1/2, 1)$ and $\bar{\rho} = (0, 0, 1/2, 1/2, 1, 1)$ then $\mu_1^u(\bar{r}, \bar{\rho}) = 0$. Note that for all $\bar{r}, \bar{\rho} \in [0,1]^{<\omega}$, $\mu_1^u(\bar{r}, \bar{\rho}), \mu_\infty^o(\bar{r}, \bar{\rho}) \leq 1$.

Definition 3.10. Let $F : \big([0,1]^{<\omega}\big)^k \to [0,1]$ be an aggregation function and let μ be any of the the pseudometrics defined in Definition 3.9. Also let $X \subseteq \big([0,1]^{<\omega}\big)^k$. We say that F is *asymptotically uniformly continuous on* X if for every $\varepsilon > 0$ there are n and $\delta > 0$ such that if $(\bar{r}_1, \ldots, \bar{r}_k), (\bar{\rho}_1, \ldots, \bar{\rho}_k) \in X$, $|\bar{r}_i|, |\bar{\rho}_i| \geq n$ for all i and $\mu_1^u\big((\bar{r}_1, \ldots, \bar{r}_k), (\bar{\rho}_1, \ldots, \bar{\rho}_k)\big) < \delta$, then $\big|F(\bar{r}_1, \ldots, \bar{r}_k) - F(\bar{\rho}_1, \ldots, \bar{\rho}_k)\big| < \varepsilon$.

Definition 3.11 (Alternative characterization of (strong) admissibility)**.** An aggregation function $F : \big([0,1]^{<\omega}\big)^m \to [0,1]$ is called *strongly admissible sensu novo* if the following two conditions hold:

(1) For all $k_1, \ldots, k_m \in \mathbb{N}^+$ and all $c_{i,j}, \alpha_{i,j} \in [0,1]$, for $i = 1, \ldots, m$ and $j = 1, \ldots, k_i$, and all sufficiently small $\delta > 0$, F is asymptotically uniformly continuous on $X_1 \times \ldots \times X_m$ where, for each $i = 1, \ldots, m$,

$$\begin{aligned} X_i = \big\{ &\bar{r} \in [0,1]^{<\omega} : \mathrm{rng}(\bar{r}) \subseteq \{c_{i,1}, \ldots, c_{i,k_i}\} \text{ and, for each } j = 1, \ldots, k_i, \\ &\text{there are between } (\alpha_{i,j} - \delta)|\bar{r}| \text{ and } (\alpha_{i,j} + \delta)|\bar{r}| \text{ coordinates in } \bar{r} \\ &\text{which equal } c_{i,j} \big\}. \end{aligned}$$

(2) For all $k_1, \ldots, k_m \in \mathbb{N}^+$ and all $c_{i,j}, \alpha_{i,j} \in [0,1]$, for $i = 1, \ldots, m$ and $j = 1, \ldots, k_i$, and every $\varepsilon > 0$, there is $\delta > 0$ such that if, for $i = 1, \ldots, m$, $\bar{r}_i, \bar{\rho}_i \in [0,1]^{<\omega}$ and
   (a) $|\bar{\rho}_i| = |\bar{r}_i|$,
   (b) $\mu_\infty^o(\bar{r}_i, \bar{\rho}_i) < \delta$,
   (c) $\mathrm{rng}(\bar{r}_i) \subseteq \{c_{i,1}, \ldots, c_{i,k_i}\}$, and
   (d) for each $j = 1, \ldots, k_i$, there are between $(\alpha_{i,j} - \delta)|\bar{r}_i|$ and $(\alpha_{i,j} + \delta)|\bar{r}_i|$ coordinates in $\bar{r}_i$ which equal $c_{i,j}$,

   then $|F(\bar{r}_1, \ldots, \bar{r}_m) - F(\bar{\rho}_1, \ldots, \bar{\rho}_m)| < \varepsilon$.

An aggregation function $F : \big([0,1]^{<\omega}\big)^m \to [0,1]$ is called *admissible sensu novo* if the above conditions hold under the restriction that $\alpha_{i,j}$ is *positive* for all i and j.

Proposition 3.12. *(i) An aggregation function is strongly admissible sensu novo if and only if it is strongly admissible.*
(ii) An aggregation function is admissible sensu novo if and only if it is admissible.

Proposition 3.12 is proved just like Proposition 6.5 in [KW23], because the only difference between strong admissibility (sensu novo) and admissibility (sensu novo) is that in the former notions (but not the latter) we allow the parameters $\alpha_{i,j}$ to be zero. This does not affect the proof of Proposition 6.5 in [KW23].

4. Logics

In this section we define the general notion of logic that we will use as well as the various concrete logics that will be studied. We are pragmatic and minimalistic and will define a logic (with values in the unit interval) to be something that has a few key properties necessary for making sense of definitions and results that follow. (We are not aware of any commonly accepted notion of a many-valued logic in general.) In the context of 0/1-valued logics a commonly used definition of a logic appears in [Ebb85, Definition 1.1.1, p. 27] and this definition is stronger than the one we give below, i.e. every logic in that sense is a logic in our sense. A difference between [Ebb85, Definition 1.1.1, p. 27] and our notion of a logic is (even when restricting to 0/1-valued logics) that we allow formulas of a logic to have free variables. In the rest of this section (as in the whole article) let σ be a finite relational signature.

Definition 4.1. (i) By a *logic (for σ)* we mean a set L of objects, called *formulas*, such that the following hold:

(1) For every $\varphi \in L$ a finite set $Fv(\varphi)$ of so-called *free variables* of φ is associated to φ. If we write $\varphi(\bar{x})$ where $\varphi \in L$ then we mean that $Fv(\varphi) \subseteq \text{rng}(\bar{x})$ and when using this notation we assume that there are no repetitions in the sequence $\bar{x}$ (although we occasionally repeat this assumption).
(2) To every triple $(\varphi(\bar{x}), \mathcal{A}, \bar{a})$ such that $\varphi(\bar{x}) \in L$, $\mathcal{A}$ is a finite σ-structure and $\bar{a} \in A^{|\bar{x}|}$ a number $\alpha \in [0, 1]$ is associated. We write $\mathcal{A}(\varphi(\bar{a})) = \alpha$ to express that α is the number, or *value*, associated to the triple $(\varphi(\bar{x}), \mathcal{A}, \bar{a})$. (We allow $\bar{x}$ and $\bar{a}$ to be empty which is the case if the formula has no free variables.)

(ii) We let the expressions '$\mathcal{A} \models \varphi(\bar{a})$' and '$\mathcal{A} \not\models \varphi(\bar{a})$' mean the same as $\mathcal{A}(\varphi(\bar{a})) = 1$ and $\mathcal{A}(\varphi(\bar{a})) = 0$, respectively.
(iii) Suppose that L is a logic, $\varphi(\bar{x}, \bar{y}) \in L$ (where $\text{rng}(\bar{x}) \cap \text{rng}(\bar{y}) = \emptyset$), $\mathcal{A}$ is a finite σ-structure and $\bar{a} \in A^{|\bar{x}|}$. Then $\varphi(\bar{a}, \mathcal{A}) = \{\bar{b} \in A^{|\bar{y}|} : \mathcal{A}(\varphi(\bar{a}, \bar{b}) = 1\}$.
(iv) Let L and L' be logics. Then let us write $L \leq L'$ if for every $\varphi(\bar{x}) \in L$ there is $\varphi'(\bar{x}) \in L'$ such that for every finite σ-structure $\mathcal{A}$ and every $\bar{a} \in A^{|\bar{x}|}$, $\mathcal{A}(\varphi(\bar{a})) = \mathcal{A}(\varphi'(\bar{a}))$.

Definition 4.2. Suppose that L is a logic for σ. We say that $\varphi(\bar{x}) \in L$ and $\psi(\bar{x}) \in L$ are *equivalent* if, for every finite σ-structure $\mathcal{A}$ and every $\bar{a} \in A^{|\bar{x}|}$, $\mathcal{A}(\varphi(\bar{a})) = \mathcal{A}(\psi(\bar{a}))$.

4.1. Some notation and concepts concerning first-order logic.

Some basic concepts and notation regarding first-order logic will be used in the sequel and are defined below.

Definition 4.3. (i) We let $FO(\sigma)$, respectively $qfFO(\sigma)$, denote the set of all first-order formulas, respectively quantifier-free first-order formulas, that can be constructed by using the signature σ.
(ii) Constructions of the form '$x = y$' and '$R(x_1, \ldots, x_r)$', where $x, y, x_1, \ldots, x_r$ are variables, and $R \in \sigma$ has arity r, are called *atomic first-order formulas (over σ)*. By a *first-order literal (over σ)* we mean a first-order atomic formula (over σ) or a negation of such one.
(iii) If $\varphi(\bar{x}) \in L$, $\mathcal{A}$ is a σ-structure and $\bar{a} \in A^{|\bar{x}|}$, then the notation '$\mathcal{A} \models \varphi(\bar{a})$' has the usual meaning of first-order logic, and we let '$\mathcal{A}(\varphi(\bar{a})) = 1$' have the same meaning as '$\mathcal{A} \models \varphi(\bar{a})$'.

Definition 4.4 (Atomic σ-types)**.** (i) A consistent set p of first-order literals over σ is called an *atomic σ-type*. If an atomic σ-type is denoted by $p(\bar{x})$ it is understood that every variable

that occurs in a formula in $p(\bar{x})$ occurs in the sequence $\bar{x}$.
(ii) An atomic σ-type $p(\bar{x})$ is called *complete* if for every first-order atomic formula $\varphi(\bar{x})$ over σ, either $\varphi(\bar{x})$ or $\neg\varphi(\bar{x})$ belongs to $p(\bar{x})$.
(iii) If $p(\bar{x})$ is an atomic σ-type and $\text{rng}(\bar{y}) \subseteq \text{rng}(\bar{x})$, then $p(\bar{x})\upharpoonright\bar{y}$ (or $p\upharpoonright\bar{y}$) denotes the set of all formulas $\varphi \in p(\bar{x})$ such that every variable of φ occurs in $\bar{y}$.
(iv) If $p(\bar{x})$ is an atomic σ-type and $\sigma' \subset \sigma$, then $p(\bar{x})\upharpoonright\sigma' = p(\bar{x}) \cap FO(\sigma')$.
(v) If $p(\bar{x})$ is an atomic σ-type, where $\bar{x} = (x_1, \ldots, x_k)$, then the *identity fragment* of p is the set of all formulas in $p(\bar{x})$ of the form $x_i = x_j$ or $\neg(x_i = x_j)$ (abbreviated $x_i \neq x_j$).

When convenient we will identify, notationally, an atomic σ-type $p(\bar{x})$ with the formula obtained by taking the conjunction of all formulas in $p(\bar{x})$. With this convention, if $\mathcal{A}$ is a σ-structure and $\bar{a} \in A^{|\bar{x}|}$ the notation $\mathcal{A} \models p(\bar{a})$ makes sense and means, with model theoretic language, that $\bar{a}$ *realizes* $p(\bar{x})$ (in the structure $\mathcal{A}$). Note that if $\sigma = \emptyset$, then an atomic σ-type $p(\bar{x})$ will only contain literals of the form $z = y$ or $z \neq y$ where $z, y \in \text{rng}(\bar{x})$.

4.2. **Probabilistic logics with aggregation functions.** We now define a quite general logic, PLA^+ (where PLA stands for *probability logic with aggregation functions*), with truth values in $[0, 1]$. Most of the logics that we will study more closely will be sublogics of PLA^+. It follows directly from the definitions that the *probability logic* of Jaeger in [Jae98] is a sublogic of PLA^+.

Definition 4.5 (Syntax of $PLA^+(\sigma)$)**.** The formulas of the logic $PLA^+(\sigma)$ are constructed as follows.[2]

(1) For each $c \in [0, 1]$, $c \in PLA^+(\sigma)$ (i.e. c is a formula) and $Fv(c) = \emptyset$. We also let $\bot$ and $\top$ denote 0 and 1, respectively.
(2) For all variables x and y, '$x = y$' belongs to $PLA^+(\sigma)$ and $Fv(x = y) = \{x, y\}$.
(3) For every $R \in \sigma$, say of arity r, and any choice of variables $x_1, \ldots, x_r$, $R(x_1, \ldots, x_r)$ belongs to $PLA^+(\sigma)$ and $Fv(R(x_1, \ldots, x_r)) = \{x_1, \ldots, x_r\}$.
(4) If $n \in \mathbb{N}^+$, $\varphi_1(\bar{x}), \ldots, \varphi_n(\bar{x})_n \in PLA^+(\sigma)$ and $\mathsf{C} : [0, 1]^n \to [0, 1]$ is a continuous connective, then $\mathsf{C}(\varphi_1, \ldots, \varphi_n)$ is a formula of $PLA^+(\sigma)$ and its set of free variables is $Fv(\varphi_1) \cup \ldots \cup Fv(\varphi_n)$.
(5) If $k \in \mathbb{N}^+$, $\varphi_1(\bar{x}, \bar{y}), \ldots, \varphi_k(\bar{x}, \bar{y}) \in PLA^+(\sigma)$, $p^=(\bar{x}, \bar{y})$ is an atomic $\emptyset$-type (i.e. a possibly empty description of the equalities and nonequalities between the variables in the sequence $\bar{x}\bar{y}$), where $\bar{x}$ and $\bar{y}$ are sequences of distinct variables such that $\text{rng}(\bar{x}) \cap \text{rng}(\bar{y}) = \emptyset$ and $F : \left([0, 1]^{<\omega}\right)^k \to [0, 1]$ is an aggregation function, then
$$F(\varphi_1(\bar{x}, \bar{y}), \ldots, \varphi_k(\bar{x}, \bar{y}) : \bar{y} : p^=(\bar{x}, \bar{y}))$$
is a formula of $PLA^+(\sigma)$ and its set of free variables is
$$\Big(\bigcup_{i=1}^{k} Fv(\varphi_i)\Big) \setminus \text{rng}(\bar{y}),$$
so this construction binds the variables in $\bar{y}$.

[2]The '+' in PLA^+ is there because the syntax of $PLA^+(\sigma)$ allows, unlike $PLA(\sigma)$ defined in [KW23], that $p^=$ in item (5) is not complete, and because item (4) is more general than the corresponding part (about connectives) for $PLA(\sigma)$ in [KW23].

Definition 4.6 (Semantics of $PLA^+(\sigma)$)**.** For each σ-structure $\mathcal{A}$, each formula $\varphi(\bar{x}) \in PLA^+(\sigma)$ and every $\bar{a} \in A^{|\bar{x}|}$, we define a real number, denoted $\mathcal{A}(\varphi(\bar{a}))$, in the interval $[0,1]$, called the *value of* $\varphi(\bar{a})$ *in* $\mathcal{A}$, as follows (where if φ has no free variable we just omit $\bar{x}$ and $\bar{a}$):

(1) For every $c \in [0,1]$ and every σ-structure $\mathcal{A}$, $\mathcal{A}(c) = c$.
(2) For every σ-structure $\mathcal{A}$ and all $a, b \in A$, $\mathcal{A}(a = b) = 1$ if $\mathcal{A} \models a = b$ and otherwise $\mathcal{A}(a = b) = 0$.
(3) For every $R \in \sigma$, of arity r say, every finite σ-structure $\mathcal{A}$ and all $\bar{a} \in A^r$, $\mathcal{A}(R(\bar{a})) = 1$ if $\mathcal{A} \models R(\bar{a})$ and otherwise $\mathcal{A}(R(\bar{a})) = 0$.
(4) If $n \in \mathbb{N}^+$, $\varphi_1(\bar{x}), \ldots, \varphi_n(\bar{x}) \in PLA^+(\sigma)$ and $\mathsf{C} : [0,1]^n \to [0,1]$ is a continuous connective, then for every finite σ-structure $\mathcal{A}$ and every $\bar{a} \in A^{|\bar{x}|}$,

$$\mathcal{A}\big(\mathsf{C}(\varphi_1(\bar{a}), \ldots, \varphi_n(\bar{a}))\big) \;=\; \mathsf{C}\big(\mathcal{A}(\varphi_1(\bar{a})), \ldots, \mathcal{A}(\varphi_n(\bar{a}))\big).$$

(5) If $k \in \mathbb{N}^+$, $\bar{x}$ and $\bar{y}$ are sequences of distinct variables such that $\mathrm{rng}(\bar{x}) \cap \mathrm{rng}(\bar{y}) = \emptyset$, $\varphi_1(\bar{x}, \bar{y}), \ldots, \varphi_k(\bar{x}, \bar{y}) \in PLA^+(\sigma)$, $p^=(\bar{x}, \bar{y})$ is an atomic $\emptyset$-type, $F : \big([0,1]^{<\omega}\big)^k \to [0,1]$ is an aggregation function, $\mathcal{A}$ is a finite σ-structure and $\bar{a} \in A^{|\bar{x}|}$, then

$$\mathcal{A}\big(F(\varphi_1(\bar{a}, \bar{y}), \ldots, \varphi_k(\bar{a}, \bar{y}) : \bar{y} : p^=(\bar{a}, \bar{y}))\big) = F(\bar{r}_1, \ldots, \bar{r}_k)$$

if there is some $\bar{b} \in A^{|\bar{y}|}$ such that $p^=(\bar{a}, \bar{b})$ holds and, for $i = 1, \ldots, k$,

$$\bar{r}_i = \big(\mathcal{A}(\varphi_i(\bar{a}, \bar{b})) : \bar{b} \in A^{|\bar{y}|} \text{ and } p^=(\bar{a}, \bar{b}) \text{ holds}\big),$$

and otherwise $\mathcal{A}\big(F(\varphi_1(\bar{a}, \bar{y}), \ldots, \varphi_k(\bar{a}, \bar{y}) : \bar{y} : p^=(\bar{a}, \bar{y}))\big) = 0$.

With Definition 3.1 of a connective we can, by using the semantics of Lukasiewicz logic (see for example [Ber08, Section 11.2]), define continuous connectives which, when restricted to $\{0,1\}$, have the usual meanings of $\neg$, $\wedge$, $\vee$, and $\to$.

Definition 4.7 (Some special continuous connectives)**.** (1) Let $\neg : [0,1] \to [0,1]$ be defined by $\neg(x) = 1 - x$.
(2) Let $\wedge : [0,1]^2 \to [0,1]$ be defined by $\wedge(x, y) = \min(x, y)$.
(3) Let $\vee : [0,1]^2 \to [0,1]$ be defined by $\vee(x, y) = \max(x, y)$.
(4) Let $\to: [0,1]^2 \to [0,1]$ be defined by $\to (x, y) = \min(1,\ 1 - x + y)$.
(5) Let $\mathrm{wm} : [0,1]^3 \to [0,1]$ (where wm stands for *weighted mean*) be defined by $\mathrm{wm}(x, y, z) = x \cdot y + (1 - x) \cdot z$.

We now give examples of the expressivity of $PLA^+(\sigma)$. Note that in Examples 4.8 and 4.9 only strongly admissible aggregation functions are used. In Example 4.10 only admissible aggregation functions are used (but the details are in [KW23]).

Example 4.8 (Similarity measure)**.** Let $E_1, \ldots, E_k \in \sigma$ be binary relation symbols. A measure of the *similarity of two elements* x *and* y, with respect to $E_1, \ldots, E_k$, is given by considering the fraction of elements which have the same connections to x and y. This can be expressed in $PLA(\sigma)$ by the formula

$$\mathrm{am}\bigg(\bigwedge_{i=1}^{k} \big((E_i(z, x) \leftrightarrow E_i(z, y)) \wedge (E_i(x, z) \leftrightarrow E_i(y, z))\big) : z : y \neq z \wedge x \neq z\bigg)$$

which we call $\psi(x, y)$. By nesting aggregation functions we can express other relations, properties and statements. "The similarity to x of the most similar other element" is given by the formula $\max(\psi(x, y) : y : x \neq y)$. "The average similarity of x to other elements" is given by $\mathrm{am}(\psi(x,y) : y : x \neq y)$. "The lowest similarity score between any two elements" is expressed by $\min((\min(\psi(x,y) : y : x \neq y)) : x :)$. In this example we considered atomic reations E_i, but it is possible to replace E_i with an arbitrary $PLA^+(\sigma)$-formula with two free variables.

Example 4.9 (Similarity profile)**.** Suppose that $E_1, \ldots, E_k \in \sigma$, $\psi(x,y)$ is the formula from Example 4.8, $\mathcal{A}$ is a finite σ-structure and $a \in A$. Then $\bar{p} = (\mathcal{A}(\psi(a,c)) : c \in A, c \neq a)$ is a sequence containing the similarity scores of (a,c) as c ranges over all other elements in A. Let us call $\bar{p}$ the "similarity profile" of a to other elements. Let now $b \in A$ and $\bar{q} = (\mathcal{A}(\psi(b,c)) : c \in A)$. Recall the pseudometric μ_1^u on $[0,1]^{<\omega}$ from Definition 3.9, which is a strongly admissible aggregation function $\mu_1^u : \left([0,1]^{<\omega}\right)^2 \to [0,1]$, and $\mu_1^u(\bar{p},\bar{q})$ measures how close the similarity profile of a is to the similarity profile of b; the smaller $\mu_1^u(\bar{p},\bar{q})$ is, the closer are the similarity profiles of a and b. So the formula $\neg\mu_1^u(\psi(x,z),\psi(y,z) : z)$ gives a higher value, or "similarity profile score", if the similarity profiles of x and y are closer. Intuitively speaking, if the similarity profile score is close to 1 then x and y are two possibly quite unrelated entities with "nearly isomorphic" connections (with respect to $E_1, \ldots, E_k$) to the rest of the world.

Example 4.10 (SimRank and PageRank)**.** In [KW23] it is demonstrated that every stage of SimRank [JW02] can be expressed by a $PLA(\sigma)$-formula (defined in Definition 4.14 below) that uses only admissible aggregation functions. One can also show (which is simpler) that every stage of PageRank [BP98] can be expressed by a $PLA(\sigma)$-formula with only admissible aggregation functions.

Definition 4.11 (Aggregation-free and basic probability formulas)**.**
(i) A formula of $PLA^+(\sigma)$ in which no aggregation function appears is called *aggregation-free.*
(ii) If $n \in \mathbb{N}^+$, $\alpha_1, \ldots, \alpha_n \in [0,1]$ and $\psi_1(\bar{x}), \ldots, \psi_n(\bar{x}) \in PLA(\sigma)$ are such that each ψ_i is a conjunction of first-order literals, then the formula $\bigwedge_{i=1}^n \left(\psi_i(\bar{x}) \to \alpha_i\right)$ is called a *basic probability formula.*

Remark 4.12. A basic probability formula which is also a sentence, that is, a formula without free variables, has the form $\bigwedge_{i=1}^n(\top \to c_i)$ where $c_i \in [0,1]$ (and recall that $\top = 1$). The formula $\bigwedge_{i=1}^n(\top \to c_i)$ is equivalent to c where $c = \min\{c_1, \ldots, c_n\}$, so every basic probability sentence is equivalent to a sentence of the form c for some $c \in [0,1]$.

The next lemma is proved in the same way as the corresponding result in [KW23, Lemma 3.10] where $PLA(\sigma)$-formulas are considered.

Lemma 4.13. *If $\varphi(\bar{x}) \in PLA^+(\sigma)$ is aggregation-free then $\varphi(\bar{x})$ is equivalent to a basic probability formula.*

We now define different sublogics of $PLA^+(\sigma)$ which we will study, both as query languages and as languages for defining networks that induce probability distributions. These sublogics will be obtained from $PLA^+(\sigma)$ by restricting the kind of aggregation functions, or connectives, that are allowed in formulas to such ones considered in Section 3.

Definition 4.14 (Sublogics of $PLA^+(\sigma)$)**.** (1) $PLA(\sigma)$ is defined like $PLA^+(\sigma)$ except that in (5) in Definition 4.5 we require that $p^=(\bar{x},\bar{y})$ is a *complete* atomic $\emptyset$-type and (4) in Definition 4.5 is restricted to only apply to the continuous connectives $\neg$, $\wedge$, $\vee$, $\to$ and wm.

(2) $aPLA^+(\sigma)$, respectively $aPLA(\sigma)$, is the subset of $PLA^+(\sigma)$, respectively $PLA(\sigma)$, where only *admissible* aggregation functions F are allowed in part (5) of Definition 4.5.

(3) $coPLA^+(\sigma)$, respectively $coPLA(\sigma)$, is the subset of $PLA^+(\sigma)$, respectively $PLA(\sigma)$, where only *strongly admissible* (or *continuous*) aggregation functions F are allowed in parts (5) of Definition 4.5.
(4) $afPLA(\sigma)$ is the set of all aggregation-free formulas in $PLA^+(\sigma)$.

Note that every aggregation-free formula in $PLA^+(\sigma)$ is, by Lemma 4.13, equivalent to a basic probability formula, which is a $PLA(\sigma)$ formula. Observe also that $afPLA(\sigma) \subseteq coPLA(\sigma) \subseteq coPLA^+(\sigma) \subseteq PLA^+(\sigma)$, $coPLA(\sigma) \subseteq aPLA(\sigma) \subseteq PLA(\sigma)$ and $coPLA(\sigma^+) \subseteq aPLA^+(\sigma) \subseteq PLA^+(\sigma)$.

The following was proved in [KW23, Lemma 3.11] for $PLA(\sigma)$ and the (simple) proof is easily generalized to $PLA^+(\sigma)$.

Lemma 4.15 (Truth value invariance under isomorphisms)**.** *Let $\mathcal{A}$ and $\mathcal{B}$ be isomorphic σ-structures and let f denote an isomorphism from $\mathcal{A}$ to $\mathcal{B}$. If $\varphi(\bar{x}) \in PLA^+(\sigma)$ and $\bar{a} \in A^{|\bar{x}|}$, then $\mathcal{A}(\varphi(\bar{a})) = \mathcal{B}(\varphi(f(\bar{a})))$.*

4.3. **Conditional probability logic.** Besides PLA^+ and its sublogics we will also consider *conditional probability logic (CPL)* [Kop20] as a logic for queries and for defining networks. CPL is a two-valued logic which extends first-order logic and with which one can express that the relative frequency of tuples that satisfy (some condition) $\varphi_1(\bar{x})$, conditioned on tuples satisfying $\varphi_2(\bar{x})$ at least as large as the relative frequency of tuples that satisfy $\varphi_3(\bar{x})$, conditioned on tuples satisfying $\varphi_4(\bar{x})$. If a probability distribution is given on the set of structures with a given finite domain, then we can use CPL to ask what the probability is that the frequency of one event (possibly conditioned on another event) is larger than the frequency of another event (possibly conditioned on yet another event). When CPL is used in the definition of a *lifted Bayesian network* in the sense of [Kop20], it expresses "threshold conditions" when the probability of a relation changes from one value to another. Concrete examples of its expressive power are found in Example 3.5 and remarks 3.4 and 3.6 in [Kop20].

Definition 4.16 (Syntax of $CPL(\sigma)$)**.** Suppose that σ is a finite relational signature. Then the set of *conditional probability formulas over* σ, denoted $CPL(\sigma)$, is defined inductively as follows:

(1) Every atomic σ-formula belongs to $CPL(\sigma)$ (where 'atomic' has the same meaning as in first-order logic with equality).
(2) If $\varphi, \psi \in CPL(\sigma)$ then $(\neg\varphi), (\varphi \wedge \psi), (\varphi \vee \psi), (\varphi \rightarrow \psi), (\varphi \leftrightarrow \psi), (\exists x \varphi) \in CPL(\sigma)$ where x is a variable. (As usual, in practice we do not necessarily write out all parentheses.) We consider $\forall x \varphi$ to be an abbreviation of $\neg\exists x \neg\varphi$.
(3) If $r \geq 0$ is a real number, $\varphi, \psi, \theta, \tau \in CPL(\sigma)$ and $\bar{y}$ is a sequence of distinct variables, then

$$\Big(r + \|\varphi \mid \psi\|_{\bar{y}} \;\geq\; \|\theta \mid \tau\|_{\bar{y}}\Big) \in CPL(\sigma) \quad \text{and}$$
$$\Big(\|\varphi \mid \psi\|_{\bar{y}} \;\geq\; \|\theta \mid \tau\|_{\bar{y}} + r\Big) \in CPL(\sigma).$$

In both these new formulas all variables of φ, ψ, θ and τ that appear in the sequence $\bar{y}$ become *bound*. So this construction can be seen as a sort of quantification, which may become more clear by the provided semantics below.

A formula $\varphi \in CPL(\sigma)$ is called *quantifier-free* if it contains no quantifier, that is, if it is constructed from atomic formulas using only the connectives $\neg, \wedge, \vee, \rightarrow, \leftrightarrow$.

Definition 4.17 (Semantics of $CPL(\sigma)$)**.** (1) The interpretations of $\neg, \wedge, \vee, \rightarrow, \leftrightarrow$ and $\exists$ are as in first-order logic.

(2) Suppose that $\mathcal{A}$ is a *finite* σ-structure and let $\varphi(\bar{x}, \bar{y}), \psi(\bar{x}, \bar{y}), \theta(\bar{x}, \bar{y}), \tau(\bar{x}, \bar{y}) \in CPL(\sigma)$. Let $\bar{a} \in A^{|\bar{x}|}$.

(a) We define $\varphi(\bar{a}, \mathcal{A}) = \big\{\bar{b} \in A^{|\bar{y}|} : \mathcal{A} \models \varphi(\bar{a}, \bar{b})\big\}$.

(b) The expression

$$\mathcal{A} \models \Big(r + \|\varphi(\bar{a}, \bar{y}) \mid \psi(\bar{a}, \bar{y})\|_{\bar{y}} \ \geq \ \|\theta(\bar{a}, \bar{y}) \mid \tau(\bar{a}, \bar{y})\|_{\bar{y}}\Big)$$

means that $\psi(\bar{a}, \mathcal{A}) \neq \emptyset$, $\tau(\bar{a}, \mathcal{A}) \neq \emptyset$ and

$$r + \frac{\big|\varphi(\bar{a}, \mathcal{A}) \cap \psi(\bar{a}, \mathcal{A})\big|}{\big|\psi(\bar{a}, \mathcal{A})\big|} \ \geq \ \frac{\big|\theta(\bar{a}, \mathcal{A}) \cap \tau(\bar{a}, \mathcal{A})\big|}{\big|\tau(\bar{a}, \mathcal{A})\big|}$$

and in this case we say that $\Big(r + \|\varphi(\bar{a}, \bar{y}) \mid \psi(\bar{a}, \bar{y})\|_{\bar{y}} \ \geq \ \|\theta(\bar{a}, \bar{y}) \mid \tau(\bar{a}, \bar{y})\|_{\bar{y}}\Big)$ is true (or holds) in $\mathcal{A}$. If $\psi(\bar{a}, \mathcal{A}) = \emptyset$ or $\tau(\bar{a}, \mathcal{A}) = \emptyset$ or

$$r + \frac{\big|\varphi(\bar{a}, \mathcal{A}) \cap \psi(\bar{a}, \mathcal{A})\big|}{\big|\psi(\bar{a}, \mathcal{A})\big|} \ < \ \frac{\big|\theta(\bar{a}, \mathcal{A}) \cap \tau(\bar{a}, \mathcal{A})\big|}{\big|\tau(\bar{a}, \mathcal{A})\big|}$$

then we write

$$\mathcal{A} \not\models \Big(r + \|\varphi(\bar{a}, \bar{y}) \mid \psi(\bar{a}, \bar{y})\|_{\bar{y}} \ \geq \ \|\theta(\bar{a}, \bar{y}) \mid \tau(\bar{a}, \bar{y})\|_{\bar{y}}\Big)$$

and say that $\Big(r + \|\varphi(\bar{a}, \bar{y}) \mid \psi(\bar{a}, \bar{y})\|_{\bar{y}} \ \geq \ \|\theta(\bar{a}, \bar{y}) \mid \tau(\bar{a}, \bar{y})\|_{\bar{y}}\Big)$ is false in $\mathcal{A}$.

(c) The meaning of

$$\mathcal{A} \models \Big(\|\varphi(\bar{a}, \bar{y}) \mid \psi(\bar{a}, \bar{y})\|_{\bar{y}} \ \geq \ \|\theta(\bar{a}, \bar{y}) \mid \tau(\bar{a}, \bar{y})\|_{\bar{y}} + r\Big)$$

is defined similarly.

5. Probabilistic graphical models, sequences of probability distributions, and asymptotic elimination of aggregation functions

In this section we define the (parametrized) probabilistic graphical models that will be used for defining probability distributions on $\mathbf{W}_n$, the set of σ-strutures, for some finite relational signature σ, with domain $[n]$. In particular we define the notion of $L(\sigma)$-network where $L(\sigma)$ is an arbitrary logic for σ. The notion of $L(\sigma)$-network is general enough to encompass all directed probabilistic graphical models that we are aware of (e.g. relational Bayesian networks, [Jae98], lifted Bayesian networks [Kop20], relational logistic regression models [MBGS19, PBK$^+$14]) by choosing an appropriate logic $L(\sigma)$). In most results mentioned in this article we consider $L(\sigma)$-networks where $L(\sigma)$ is a sublogic of $PLA^+(\sigma)$ obtained by putting restrictions on the kind of aggregation functions that may be used. This is because we want to understand the role of aggregation functions in the interplay between those logics used to define probability distributions and (possibly other) logics used to define queries. It follows immediately from the definitions that every relational Bayesian network (for σ) in [Jae98] is a $PLA^+(\sigma)$-network, modulo some notational and terminological differences.

In Section 5.2 we state our first main result, Theorem 5.11, which is about $PLA^+(\sigma)$-networks and queries that use only strongly admissible (or continuous) aggregation functions, and its corollaries. As discussed in Section 1, Theorem 5.11 and Corollary 5.13 differ from other results with similar general aims in the field of statistical relational AI, e.g. [Jae98, MBGS19, PBK+14, Kop20, KW23], because they apply to *different* sequences of probability distributions or to *different* queries.

Throughout this section (as in the rest of the article) we assume that σ is a finite relational signature and that $\mathbf{W}_n$ denotes the set of all σ-structures with domain $[n]$. We begin by defining some notions that are relevant to all probabilistic graphical models and logics that we consider. Then we define $L(\sigma)$-networks, lifted Bayesian networks and then state the main results of [Kop20, KW23] which will be used later.

5.1. **Sequences of probability distributions and asymptotic equivalence.** Let L be a logic for σ.

Definition 5.1. By a *sequence of probability distributions* we mean a sequence $(\mathbb{P}_n : n \in \mathbb{N}^+)$ such that for every n, $\mathbb{P}_n$ is a probability distribution on $\mathbf{W}_n$.

Definition 5.2. Let $\mathbb{P}_n$ is a probability distribution on $\mathbf{W}_n$, let $\varphi(\bar{x}) \in L$, and let $\bar{a} \in [n]^{|\bar{x}|}$. Then

$$\mathbb{P}_n(\varphi(\bar{a})) = \mathbb{P}_n\big(\{\mathcal{A} \in \mathbf{W}_n : \mathcal{A}(\varphi(\bar{a})) = 1\}\big).$$

If $p(\bar{x})$ is a finite set of formulas (e.g. an atomic σ-type) then we let $\mathbb{P}_n(p(\bar{a})) = \mathbb{P}_n(\varphi(\bar{a}))$ where $\varphi(\bar{x})$ is the conjunction of all formulas in $p(\bar{x})$.

Definition 5.3 (Asymptotic total variation equivalence)**.** Two sequences of probability distributions $(\mathbb{P}_n : n \in \mathbb{N}^+)$ and $(\mathbb{P}'_n : n \in \mathbb{N}^+)$ are called *asymptotically total variation equivalent*, denoted $(\mathbb{P}_n : n \in \mathbb{N}^+) \sim_{tv} (\mathbb{P}'_n : n \in \mathbb{N}^+)$, if there is a function $\delta : \mathbb{N}^+ \to \mathbb{R}$ such that $\lim_{n\to\infty} \delta(n) = 0$ and for all sufficiently large n and every $\mathbf{X} \subseteq \mathbf{W}_n$, $|\mathbb{P}_n(\mathbf{X}) - \mathbb{P}'_n(\mathbf{X})| \leq \delta(n)$.

Definition 5.4 (Asymptotic equivalence of formulas)**.** Let $\varphi(\bar{x}), \psi(\bar{x}) \in L$.
(i) Let $\mathbb{P} = (\mathbb{P}_n : n \in \mathbb{N}^+)$ be a sequence of probability distributions. We say that $\varphi(\bar{x})$ and $\psi(\bar{x})$ are *asymptotically equivalent with respect to* $\mathbb{P}$ if for all $\varepsilon > 0$

$$\mathbb{P}_n\Big(\big\{\mathcal{A} \in \mathbf{W}_n : \text{ there is } \bar{a} \in A^{|\bar{x}|} \text{ such that } |\mathcal{A}(\varphi(\bar{a})) - \mathcal{A}(\psi(\bar{a}))| > \varepsilon\big\}\Big) \to 0$$

as $n \to \infty$.
(ii) Suppose in addition that L is a logic such that for every $\varphi(\bar{x}) \in L$, every finite σ-structure $\mathcal{A}$ and every $\bar{a} \in A^{|\bar{a}|}$, the value $\mathcal{A}(\varphi(\bar{a}))$ is either 0 or 1. Then we say that $\varphi(\bar{x}) \in L$ and $\psi(\bar{x}) \in L$ are *almost surely equivalent with respect to* $\mathbb{P}$ if

$$\lim_{n\to\infty} \mathbb{P}_n\big(\{\mathcal{A} \in \mathbf{W}_n : \text{ for all } \bar{a} \in [n]^{|\bar{x}|}, \mathcal{A}(\varphi(\bar{a})) = \mathcal{A}(\psi(\bar{a}))\}\big) \;=\; 1.$$

For logics with only the truth values 0 and 1, the notions of asymptotic equivalence and almost sure equivalence are equivalent as stated by the following lemma, the straightforward proof of which is left to the reader.

Lemma 5.5. *Let $\mathbb{P} = (\mathbb{P}_n : n \in \mathbb{N}^+)$ be a sequence of probability distributions and suppose that L is a logic such that for every $\varphi(\bar{x}) \in L$, every finite σ-structure $\mathcal{A}$ and every $\bar{a} \in A^{|\bar{a}|}$, the value $\mathcal{A}(\varphi(\bar{a}))$ is either 0 or 1. Let $\varphi(\bar{x}), \psi(\bar{x}) \in L$. Then $\varphi(\bar{x})$ and $\psi(\bar{x})$ are asymptotically equivalent with respect to $\mathbb{P}$ if and only if they are almost surely equivalent with respect to $\mathbb{P}$.*

5.2. **$L(\sigma)$-Networks.** Now we define the notion of $L(\sigma)$-network, where $L(\sigma)$ is a logic for σ, and how it induces a probability distribution on $\mathbf{W}_n$.

Definition 5.6 ($L(\sigma)$-network)**.** Suppose that $L(\sigma)$ is a logic for σ and suppose that, for every $\sigma' \subset \sigma$, $L(\sigma')$ is a logic for σ' and whenever $\sigma'' \subset \sigma' \subset \sigma$, then $L(\sigma'') \subset L(\sigma') \subset L(\sigma)$.
(i) A $L(\sigma)$-*network* is determined by the following two components:

(1) A DAG $\mathbb{G}$ with vertex set σ.
(2) To each relation symbol $R \in \sigma$ a formula $\theta_R(\bar{x}) \in L(\text{par}(R))$ (where $\text{par}(R)$ is the set of parents of R in the DAG) is associated where $|\bar{x}|$ equals the arity of R. We call θ_R the *probability formula associated to* R by the network.

(ii) For technical reasons it will be convenient to consider, for the empty signature $\sigma = \emptyset$, a (unique) $L(\sigma)$-network, denoted $\mathbb{G}^\emptyset$, such that its underlying DAG has empty vertex set and consequently no probability formula.
(iii) Let $\mathbb{G}$ denote an $L(\sigma)$-network, let $\sigma' \subseteq \sigma$, and suppose that for every $R \in \sigma'$, $\text{par}(R) \subseteq \sigma'$. Then the $L(\sigma')$-network specified by the induced subgraph of the underlying DAG of $\mathbb{G}$ with vertex set σ' and the probability formulas θ_R for all $R \in \sigma'$ will be called the $L(\sigma')$-*subnetwork of* $\mathbb{G}$ *induced by* σ'.

We use the convention to denote an $L(\sigma)$-network by the same symbol (e.g. $\mathbb{G}$) as its underlying DAG.

It follows immediately from its definition that $PLA^+(\sigma)$ and all sublogics of it described in Definition 4.14 satisfy the conditions in the beginning of the above definition, so it makes sense to talk about for example a $PLA^+(\sigma)$-network, a $PLA(\sigma)$-network, a $aPLA(\sigma)$-network, a $coPLA^+(\sigma)$-network, and an $afPLA(\sigma)$-network.

Now we define how an $L(\sigma)$-network defines a probability distribution on $\mathbf{W}_n$.

Definition 5.7 (The sequence of probability distributions induced by an $L(\sigma)$-network)**.** Let $L(\sigma)$ be a logic for σ which satisfies the conditions in the beginning of Definition 5.6 and let $\mathbb{G}$ be an $L(\sigma)$-network.
(i) If σ is empty then $\mathbb{P}_n$, the probability distribution on $\mathbf{W}_n$ induced by $\mathbb{G}$, is the unique probability distribution on (the singleton set) $\mathbf{W}_n$.
(ii) Now suppose that σ is nonempty and suppose tat for each $R \in \sigma$, its arity is denoted by k_R and the probability formula corresponding to R is denoted by $\theta_R(\bar{x})$ where $|\bar{x}| = k_R$. Suppose that the underlying DAG of $\mathbb{G}$ has mp-rank ρ. For each $0 \leq r \leq \rho$ let $\mathbb{G}_r$ be the subnetwork which is induced by $\sigma_r = \{R \in \sigma : \text{mp}(R) \leq r\}$ and note that $\mathbb{G}_\rho = \mathbb{G}$. Also let $\mathbb{G}_{-1} = \mathbb{G}^\emptyset$ and let $\mathbb{P}_n^{-1}$ be the unique probability distribution on $\mathbf{W}_n^{-1} = \mathbf{W}_n^\emptyset$. By induction on r we define, for every $r = 0, 1, \ldots, \rho$, a probability distribution $\mathbb{P}_n^r$ on the set $\mathbf{W}_n^r$ of all σ_r-structures with domain $[n]$ as follows: For every $\mathcal{A} \in \mathbf{W}_n^r$, let $\mathcal{A}' = \mathcal{A} \restriction \sigma_{r-1}$ and

$$\mathbb{P}_n^r(\mathcal{A}) \; = \; \mathbb{P}_n^{r-1}(\mathcal{A}') \prod_{R \in \sigma_r \setminus \sigma_{r-1}} \; \prod_{\bar{a} \in R^{\mathcal{A}}} \mathcal{A}'\big(\theta_R(\bar{a})\big) \prod_{\bar{a} \in [n]^{k_R} \ \setminus \ R^{\mathcal{A}}} \big(1 - \mathcal{A}'\big(\theta_R(\bar{a})\big)\big).$$

Finally we let $\mathbb{P}_n = \mathbb{P}_n^\rho$ and note that $\mathbf{W}_n = \mathbf{W}_n^\rho$, so $\mathbb{P}_n$ is a probability distribution on $\mathbf{W}_n$ and we call $(\mathbb{P}_n : n \in \mathbb{N}^+)$ *the sequence of probability distributions induced by* $\mathbb{G}$.

From the above definition it follows immediately that if a probability distribution on $\mathbf{W}_n$ can be induced by a relational Bayesian network (for σ) in the sense of [Jae98] then it can be induced by a $PLA^+(\sigma)$-network. The following lemma is a straightforward consequence of Definition 5.6 and it uses the notation and assumptions of this definition.

Lemma 5.8. *If $0 \leq r \leq \rho$, $R \in \sigma_r$ has arity k, $\theta_R(\bar{x}) \in L(\mathrm{par}(R))$ is the formula associated to R according to Definition 5.6, $n \in \mathbb{N}^+$, $\bar{a} \in [n]^k$ and $\mathcal{A} \in \mathbf{W}_n^{r-1}$, then*

$$\mathbb{P}_n^r\big(\{\mathcal{B} \in \mathbf{W}_n^r : \mathcal{B} \models R(\bar{a})\} \mid \{\mathcal{B} \in \mathbf{W}_n : \mathcal{B}{\restriction}\sigma_{r-1} = \mathcal{A}\}\big) \ = \ \mathcal{A}(\theta_R(\bar{a})).$$

Definition 5.9. Let $\mathbb{G}$ be an $L(\sigma)$-network. We say that $\varphi(\bar{x})$ and $\psi(\bar{x})$ are *asymptotically equivalent with respect to to* $\mathbb{G}$, denoted $\varphi(\bar{x}) \sim_{\mathbb{G}} \psi(\bar{x})$, if $\varphi(\bar{x})$ and $\psi(\bar{x})$ are asymptotically equivalent with respect to the sequence of probability distributions induced by $\mathbb{G}$.

Example 5.10. We give examples of (conditional) probabilities that can be modelled with a $PLA^+(\sigma)$-network. Let σ be a finite relational signature and suppose that $\mathbb{G}$ is a DAG with vertex set σ. Let $R \in \sigma$. We show how the probability of R, conditioned on $\mathrm{par}(R)$, may be expressed. In all examples, we need to define a formula $\theta_R(\bar{x}) \in PLA^+(\sigma)$ as in Definition 5.6 which assigns the probability that $R(\bar{x})$ holds, conditioned on the values of the relations in $\mathrm{par}(R)$.

If R $\mathrm{par}(R) = \emptyset$ (i.e. if R has no parents in the DAG) then we can let $\theta_R(\bar{x})$ be some constant $c \in [0, 1]$, meaning that the probability of $R(\bar{x})$ is c. Or we can (for example) let $\alpha \in (0, 1)$ and then let $\theta_R(\bar{x})$ be $length_\alpha(y = y : y :)$ (where $y \notin \mathrm{rng}(\bar{x})$). This expresses that the probability of $R(\bar{x})$ is $n^{-\alpha}$ where n is the cardinality of the domain.

For all remaining examples suppose that $\mathrm{par}(R) \neq \emptyset$. If we want to express that the probability of $R(\bar{x})$ depends (only) on the complete atomic $\mathrm{par}(R)$-type that $\bar{x}$ satisifies, then we can let $\theta_R(\bar{x})$ be $\bigwedge_i (p_i(\bar{x}) \to \gamma_i)$ where $p_i(\bar{x})$ ranges over all complete atomic $\mathrm{par}(R)$-types in the variables $\bar{x}$. This expresses that if $p_i(\bar{x})$ holds then the probability of $R(\bar{x})$ is the value of the formula γ_i (without free variables). Here γ_i may (for example) be a constant $c_i \in [0, 1]$ expressing that if $p_i(\bar{x})$ holds, then the probability of $R(\bar{x})$ is c_i, or it may (for example) be a $PLA^+(\sigma)$-formula of the form $length_\alpha(y = y : y :)$, expressing that if $p_i(\bar{x})$ holds, then the probability of $R(\bar{x})$ is $n^{-\alpha}$ where n is the cardinality of the domain.

Let $\varphi_1(x, y), \ldots, \varphi_k(x, y) \in PLA^+(\mathrm{par}(R))$. If R is binary then we can let $\theta_R(x, y)$ be the formula $\psi(x, y)$ from Example 4.8 with E_i replaced by φ_i, which expresses that the probability of $R(x, y)$ equals the similarity of x and y with respect to this similarity measure.

If R is unary then we can instead let $\theta_R(x)$ be the formula $\mathrm{am}(\psi(x, y) : y : x \neq y)$ which expresses that the probability of $R(x)$ equals the average similiarity of x to other elements with respect to $\varphi_1(x, y), \ldots, \varphi_k(x, y)$.

Suppose again that R is binary. Then we can let $\theta_R(x, y)$ be

$$1 - \mu_1^u\big(\psi(x, z), \psi(y, z) : z : x \neq z, y \neq z\big)$$

which (recalling Example 4.9) expresses that the probability of $R(x, y)$ equals the "proximity" of the similarity profiles of x and y. Alternatively, if $E \in \mathrm{par}(R)$ then $\theta_R(x, y)$ can, for any $l \in \mathbb{N}$, be the $PLA^+(\mathrm{par}(R))$-formula which expresses the lth stage (or approximation) of SimRank with respect to E. This means that the probability of $R(x, y)$ equals the lth stage of the approximation of SimRank of x and y.

Now we have the concepts that allow us to state our first main result. Recall Definition 4.14 of $coPLA(\sigma)$.

Theorem 5.11 (Asymptotic elimination of continuous aggregation functions)**.** *Let σ be a finite relational signature, let $\mathbb{G}$ be a $coPLA^+(\sigma)$-network and let $(\mathbb{P}_n : n \in \mathbb{N}^+)$ be the sequence of probability distributions induced by $\mathbb{G}$.*

(i) *If $\varphi(\bar{x}) \in coPLA^+(\sigma)$ then $\varphi(\bar{x})$ is asymptotically equivalent to a basic probability formula with respect to $\mathbb{G}$.*

(ii) *For every atomic σ-type $p(\bar{x})$, every $m \in \mathbb{N}^+$ and every $\bar{a} \in [m]^{|\bar{x}|}$, $\lim_{n\to\infty} \mathbb{P}_n(p(\bar{a}))$ exists and depends only on p and $\mathbb{G}$.*

Remark 5.12 (Expressivity of *coPLA*$^+$)**.** Recall that *coPLA*$^+$ may only use continuous aggregation functions. The aggregation functions am and μ_1^u are continuous, as stated before. Therefore the formulas from examples 4.8 and 4.9 that express "the similarity of x and y", respectively "the similarity profile score och x and y" are in *coPLA*$^+$.

But max and min are not continuous and cannot be expressed in some indirect way by *coPLA*$^+$, for otherwise Corollary 5.13 would contradict a nonconvergence result in [SS88] as discussed in the introduction. It follows that the first-order quantifiers cannot be expressed by *coPLA*$^+$, so it does not subsume first-order logic.

However in some contexts one may be more interested in the proportion (rather than existence) of elements that satisfy some (two-valued) formula. For this we can use the continuous aggregation function am which returns this proportion. If we want to express that the proportion of elements that satisfy a formula is at least as large as the proportion that satisfies another formula, then we could, tentatively, use the connective $\leq (x, y)$ which returns 1 if $x \leq y$ and 0 otherwise. But since this connective is not continuous it cannot be used in *PLA*$^+$, so instead we consider a continuous approximation of $\leq$, say $\leq_\varepsilon (x, y) = 1$ if $x \leq y$, 0 if $x \geq y + \varepsilon$, and $(y + \varepsilon - x)/\varepsilon$ if $y < x < y + \varepsilon$, where $\varepsilon > 0$ is small. With $\leq_\varepsilon$ and am we can, within *coPLA*$^+$, express approximations of statements about proportions. The idea of considering continuous approximations of connectives or aggregation functions can of course be used more generally to get *coPLA*$^+$-formulas that estimate discrete/discontinuous properties.

The proof of Theorem 5.11 consists of two parts. One part is to show that if, with high probability, structures in $\mathbf{W}_n$ satisfy certain saturation conditions (defined in Section 6) then strongly admissible aggregation functions can be asymptotically eliminated. This is stated by Proposition 6.4 and proved in Section 6. The other part is to prove that with probability tending to 1 as the domain size tends to infinity, a random structure from $\mathbf{W}_n$ satisfies these saturation conditions. This is proved in Section 7. The following corollary is quickly derived from Theorem 5.11.

Corollary 5.13 (Convergence of probability)**.** *Let σ be a finite relational signature, let $\mathbb{G}$ be a coPLA$^+(\sigma)$-network, and let $(\mathbb{P}_n : n \in \mathbb{N}^+)$ be the sequence of probability distributions induced by $\mathbb{G}$. If $\varphi(\bar{x}) \in coPLA^+(\sigma)$ then there are $c_1, \ldots, c_k \in [0, 1]$, depending only on φ and $\mathbb{G}$, such that for every $m \in \mathbb{N}^+$, every $\bar{a} \in [m]^{|\bar{x}|}$ and every $\varepsilon > 0$,*

$$\lim_{n\to\infty} \mathbb{P}_n\big(\big\{\mathcal{A} \in \mathbf{W}_n : \mathcal{A}(\varphi(\bar{a})) \in \bigcup_{i=1}^{k}[c_i - \varepsilon, c_i + \varepsilon]\big\}\big) \;=\; 1$$

and for all $i = 1, \ldots, k$

$\mathbb{P}_n\big(\big\{\mathcal{A} \in \mathbf{W}_n : |\mathcal{A}(\varphi(\bar{a})) - c_i| < \varepsilon\big\}\big)$ converges as $n \to \infty$

to a number which depends only on φ, c_i and $\mathbb{G}$.

Proof. Let $\mathbb{G}$ and $\varphi(\bar{x})$ be as assumed. By Part (i) of Theorem 5.11, there is a basic probability formula $\psi(\bar{x})$ which is asymptotically equivalent to $\varphi(\bar{x})$ with respect to $\mathbb{G}$. Then $\psi(\bar{x})$ has the form $\bigwedge_{i=1}^{k}(\psi_i(\bar{x}) \to c_i)$ where, for each i, $c_i \in [0, 1]$ and $\psi_i(\bar{x})$ is a conjunction of first-order literals. Without loss of generality we can assume that each ψ_i is the conjunction of all formulas in a complete atomic σ-type and that every complete atomic σ-type in the variables $\bar{x}$ is represented by some ψ_i. Note that for every $\mathcal{A} \in \mathbf{W}_n$ and every $\bar{a} \in [n]^{|\bar{x}|}$ we have $\mathcal{A}(\psi(\bar{a})) \in \{c_1, \dots, c_k\}$ and $\mathcal{A}(\psi(\bar{a})) = c_i$ if $\mathcal{A} \models \psi_i(\bar{a})$. Let $c \in \{c_1, \dots, c_k\}$ and suppose that $i_1, \dots, i_t$ enumerates all i such that $c_i = c$. Then

$$\mathbb{P}_n\big(\{\mathcal{A} \in \mathbf{W}_n : \mathcal{A}(\psi(\bar{a})) = c\}\big) = \mathbb{P}_n\big(\bigvee_{j=1}^{t} \psi_{i_j}(\bar{a})\big) = \sum_{j=1}^{t} \mathbb{P}_n\big(\psi_{i_j}(\bar{a})\big).$$

By Part (ii) of Theorem 5.11, it follows that the above probability converges, as $n \to \infty$, to a number which depends only on ψ and $\mathbb{G}$. Since $\varphi(\bar{x}) \sim_{\mathbb{G}} \psi(\bar{x})$ the conclusions of the corollary follow. □

5.3. **Lifted Bayesian networks.** The notions and results of this subsection are not needed until Section 9, but we introduce them here since Section 5 has introduced the other probabilistic graphical model considered here, the $L(\sigma)$-network (e.g. $PLA^+(\sigma)$-network). In one of our main results, Theorem 9.7, we will consider (sequences of) probability distributions induced by lifted Bayesian networks in the sense of [Kop20]. It is not hard (by combining CPL, constants in $[0, 1]$, and the "weighted mean" connective in Definition 4.7) to construct a logic, say $CPL^*(\sigma)$, such that every lifted Bayesian network for σ corresponds to an $CPL^*(\sigma)$-network. In order to avoid the somewhat cumbersome translation of results in [Kop20] about lifted Bayesian networks to results about $CPL^*(\sigma)$-networks we do not however do this, but instead we formulate the results in this section in their original form which uses the notion of a lifted Bayesian network. The intuition behind the next definition is that if $R \in \sigma$ and the condition expressed by the CPL-formula $\chi_{R,i}(\bar{x})$ holds, then the probability of $R(\bar{x})$ is (the number) $\mu(R \mid \chi_{R,i})$.

Definition 5.14 (Lifted Bayesian network)**.** Let σ be a finite relational signature. A *lifted Bayesian network for* σ is determined by the following components:

(a) An acyclic directed graph (DAG) $\mathbb{G}$ with vertex set σ.
(b) For each $R \in \sigma$, a number $\nu_R \in \mathbb{N}^+$, formulas $\chi_{R,i}(\bar{x}) \in CPL(\mathrm{par}(R))$, for $i = 1, \dots, \nu_R$, where $|\bar{x}|$ equals the arity of R, such that $\forall\bar{x}\big(\bigvee_{i=1}^{\nu_R} \chi_{R,i}(\bar{x})\big)$ is valid (i.e. true in all $\mathrm{par}(R)$-structures) and if $i \neq j$ then $\exists\bar{x}\big(\chi_{R,i}(\bar{x}) \wedge \chi_{R,j}(\bar{x})\big)$ is unsatisfiable. Each $\chi_{R,i}$ will be called an *aggregation formula (of* $\mathbb{G}$*)*.
(c) For each $R \in \sigma$ and each $1 \leq i \leq \nu_R$, a number denoted $\mu(R \mid \chi_{R,i})$ (or $\mu(R(\bar{x}) \mid \chi_{R,i}(\bar{x}))$) in the interval $[0, 1]$.

Observe that Definition 5.14 makes sense if σ is empty. In this case the underlying DAG has empty vertex set (and edge set) and no numbers or formulas as in parts (b) and (c) of the definition need to be specified. A lifted Bayesian network $\mathbb{G}$ for σ induces a probability distribution on $\mathbf{W}_n$ in a way explained in [Kop20, Definition 3.11], but as we will not need the details we omit the definition here. We refer to [KW23, Example 5.3] for an example of a lifted Bayesian network.

Now we state previous results about lifted Bayesian networks, CPL and PLA in [Kop20] and [KW23] that will be used in proving Theorem 9.7 which describes the relative asymptotic expressivity of the inference frameworks considered in this article. These results use the notion of a (non)critical CPL-formula. The intuition behind this notion is that a $CPL(\sigma)$-formula is *critical* with respect to a lifted Bayesian network $\mathbb{G}$ for σ if it contains a subformula of the form

$$\Big(r + \|\varphi \mid \psi\|_{\bar{y}} \ \geq \ \|\theta \mid \tau\|_{\bar{y}}\Big) \quad \text{or} \quad \Big(\|\varphi \mid \psi\|_{\bar{y}} \ \geq \ \|\theta \mid \tau\|_{\bar{y}} + r\Big).$$

where $r = \alpha - \beta$, $\alpha = \lim_{n\to\infty} \frac{\mathbb{P}_n(\varphi_1(\bar{a}))}{\mathbb{P}_n(\varphi_2(\bar{a}))}$, $\beta = \lim_{n\to\infty} \frac{\mathbb{P}_n(\psi_1(\bar{a}))}{\mathbb{P}_n(\psi_2(\bar{a}))}$ and $\varphi(\bar{x})$ and $\psi(\bar{x})$ are quantifier-free first-order formulas, the length of $\bar{x}$ is bounded in terms of the length of the subformula, $\varphi_1(\bar{x})$ implies $\varphi_2(\bar{x})$, and $\psi_1(\bar{x})$ implies $\psi_2(\bar{x})$. Otherwise the $CPL(\sigma)$-formula is *noncritical* with respect to $\mathbb{G}$. Observe that *every first-order formula is noncritical with respect to every lifted Bayesian network*, since a first-order formula does not contain a subformula of the form considered above The numbers α and β above depend only on the underlying DAG of $\mathbb{G}$ and the numbers $\mu(R \mid \chi_{R,i})$ in the definition of a lifted Bayesian network. Therefore it makes sense to say that an aggregation formula of $\mathbb{G}$ is (or is not) noncritical with respect to $\mathbb{G}$. The exact definition of noncritical formula in [Kop20] is quite technical and appears in [Kop20, Definitions 4.29 and 4.30]. A simplified and stronger definition of noncritical $CPL(\sigma)$-formula is given in [KW23, Definition 6.7].

Theorem 5.15 [Kop20, Theorems 3.14 – 3.16]**.** *Let $\mathbb{G}$ be a lifted Bayesian network for σ, suppose that all aggregation functions of $\mathbb{G}$ are noncritical with respect to $\mathbb{G}$, and let $\mathbb{P} = (\mathbb{P}_n : n \in \mathbb{N}^+)$ be the sequence of probability distributions induced by $\mathbb{G}$.*

- (i) *If $\varphi(\bar{x}) \in CPL(\sigma)$ is noncritical with respect to $\mathbb{G}$, then $\varphi(\bar{x})$ is almost surely equivalent to a quantifier-free formula (with respect to $\mathbb{P}$). Moreover, if $m \in \mathbb{N}^+$ and $\bar{a} \in [m]^{|\bar{x}|}$, then $\lim_{n\to\infty} \mathbb{P}_n(\varphi(\bar{a}))$ exists. In particular, if φ has no free variable (i.e. is a sentence), then $\lim_{n\to\infty} \mathbb{P}_n(\varphi)$ is either 0 or 1.*
- (ii) *For every aggregation formula $\chi_{R,i}$ of $\mathbb{G}$ there is a quantifier-free formula $\chi'_{R,i} \in CPL(\sigma)$ such that every relation symbol of $\chi'_{R,i}$ occurs in $\chi_{R,i}$ and such that if $\mathbb{G}'$ is the lifted Bayesian network for σ*
  - (a) *with the same underlying DAG as $\mathbb{G}$, and*
  - (b) *where, for every $R \in \sigma$, the aggregation formula $\chi_{R,i}$ is replaced by $\chi'_{R,i}$ and $\mu(R \mid \chi'_{R,i}) = \mu(R \mid \chi_{R,i})$,*

  then $(\mathbb{P}'_n : n \in \mathbb{N}^+) \sim_{tv} (\mathbb{P}_n : n \in \mathbb{N}^+)$ where $(\mathbb{P}'_n : n \in \mathbb{N}^+)$ is the sequence of probability distributions induced by $\mathbb{G}'$.

Remark 5.16. The conclusion of Theorem 3.16 in [Kop20] is somewhat weaker than part (ii) of Theorem 5.15 above since the former does not refer to asymptotic total variation equivalence, but the proof of [Kop20, Theorem 3.16] shows, using the notation above, that $(\mathbb{P}'_n : n \in \mathbb{N}^+) \sim_{tv} (\mathbb{P}_n : n \in \mathbb{N}^+)$ (as seen from the proof of [Kop20, Corollary 4.42 (c)]).

Theorem 5.17 [KW23, Theorem 6.8]**.** *Let $\mathbb{G}$ be a lifted Bayesian network for σ and suppose that all aggregation functions of $\mathbb{G}$ are noncritical with respect to $\mathbb{G}$. If $\varphi(\bar{x}) \in PLA(\sigma)$ and all aggregation functions in φ are admissible, then $\varphi(\bar{x})$ is asymptotically equivalent to a basic probability formula with respect to $\mathbb{G}$.*

6. Asymptotic elimination of strongly admissible aggregation functions

In this section we prove a result, Proposition 6.4, which is one component of the proof of Theorem 5.11 and of independent interest since it shows that if a saturation condition (given by Definition 6.2 and Assumption 6.3) holds almost surely, then strongly admissible aggregation functions can be eliminated from $PLA^+(\sigma)$-formulas.

Throughout this section, for each $n \in \mathbb{N}^+$, $\mathbb{P}_n$ is a probability distribution on $\mathbf{W}_n$, the set of all σ-structures with domain $[n]$, where σ is a finite relational signature. When saying that two formulas are asymptotically equivalent then it is with respect to $\mathbb{P} = (\mathbb{P}_n : n \in \mathbb{N}^+)$. When denoting an atomic σ-type by notation like $p(\bar{x}, \bar{y})$, or a formula by $\varphi(\bar{x}, \bar{y})$, we assume that $\bar{x}$ and $\bar{y}$ are sequences of different variables and $\text{rng}(\bar{x}) \cap \text{rng}(\bar{y}) = \emptyset$ although this assumption may be repeated.

Let $p(\bar{x}, \bar{y})$ be an atomic $\emptyset$-type, let $\mathcal{A}$ be a finite σ-structure and let $\bar{a} \in A^{|\bar{x}|}$. The equalities and inequalities in $p(\bar{x}, \bar{y})$ specifies how many "degrees of freedom" we have for choosing $\bar{b} \in A^{|\bar{y}|}$ such that $p(\bar{a}, \bar{b})$ holds. Since the "degrees of freedom" for $\bar{y}$ when $\bar{x}$ has been assigned some fixed values will play a role in the proof that will follow we now define the $\bar{y}$-dimension of $p(\bar{x}, \bar{y})$ which captures this idea.

Definition 6.1. Let $p(\bar{x}, \bar{y})$ be an atomic σ-type. The *$\bar{y}$-dimension of* $p(\bar{x}, \bar{y})$, denoted $\dim_{\bar{y}}(p)$, is the maximal $d \in \mathbb{N}$ such that there are a σ-structure $\mathcal{A}$, $\bar{a} \in A^{|\bar{x}|}$ and $\bar{b} \in A^{|\bar{y}|}$ such that $\mathcal{A} \models p(\bar{a}, \bar{b})$ and $|\text{rng}(\bar{b}) \setminus \text{rng}(\bar{a})| \geq d$.

Observe that if $p^=(\bar{x}, \bar{y})$ is an atomic $\emptyset$-type, $d = \dim_{\bar{y}}(p)$, $\mathcal{A} \in \mathbf{W}_n$, and $\bar{a} \in [n]^{|\bar{x}|}$ satisfies $p^= \restriction \bar{x}$, then (for large enough n) $|p^=(\bar{a}, \mathcal{A})| = (n - |\text{rng}(\bar{a})|)^d \sim n^d$. The intuitive content of Assumption 6.3 below is that for every complete atomic σ-type $p(\bar{x}, \bar{y})$ and $q(\bar{x}) = p \restriction \bar{x}$, there is $\alpha \in [0, 1]$ such that with high probability (for large n), if $q(\bar{a})$ holds then the proportion of $\bar{b}$ such that $p(\bar{a}, \bar{b})$ holds is close to α. The next definition will be used to make this idea more precise.

Definition 6.2 (Saturation and unsaturation)**.** Let $\bar{x}$ and $\bar{y}$ be sequences of different variables such that $\text{rng}(\bar{x}) \cap \text{rng}(\bar{y}) = \emptyset$ and let $p(\bar{x}, \bar{y})$ and $q(\bar{x})$ be atomic σ-types such that $q \subseteq p$. Let also $0 \leq \alpha \leq 1$ and $d = \dim_{\bar{y}}(p)$.

(1) A finite σ-structure $\mathcal{A}$ is called (p, q, α)-*saturated* if, whenever $\bar{a} \in A^{|\bar{x}|}$ and $\mathcal{A} \models q(\bar{a})$, then $\left|\{\bar{b} \in A^{|\bar{y}|} : \mathcal{A} \models p(\bar{a}, \bar{b})\}\right| \geq \alpha |A|^d$.
(2) A finite σ-structure $\mathcal{A}$ is called (p, q, α)-*unsaturated* if, whenever $\bar{a} \in A^{|\bar{x}|}$ and $\mathcal{A} \models q(\bar{a})$, then $\left|\{\bar{b} \in A^{|\bar{y}|} : \mathcal{A} \models p(\bar{a}, \bar{b})\}\right| \leq \alpha |A|^d$.

Assumption 6.3. For all $m, n \in \mathbb{N}^+$ and $\delta > 0$ there is $\mathbf{Y}_n^{m,\delta} \subseteq \mathbf{W}_n$ such that

(1) $\lim_{n \to \infty} \mathbb{P}_n(\mathbf{Y}_n^{m,\delta}) = 1$, and
(2) for every complete atomic σ-type $p(\bar{x}, \bar{y})$ such that $|\bar{x}| + |\bar{y}| \leq m$, if $q(\bar{x}) = p \restriction \bar{x}$ and $\dim_{\bar{y}}(p) > 0$, then there is $\alpha_{p,q} \in [0, 1]$ depending only on p, q and $\mathbb{P}$, such that every $\mathcal{A} \in \mathbf{Y}_n^{m,\delta}$ is $(p, q, \alpha_{p,q} - \delta)$-saturated and $(p, q, \alpha_{p,q} + \delta)$-unsaturated.

Proposition 6.4. *If Assumption 6.3 holds and $\varphi(\bar{x}) \in PLA^+(\sigma)$ contains only strongly admissible aggregation functions, then $\varphi(\bar{x})$ is asymptotically equivalent to a basic probability formula.*

The rest of this section is devoted to proving Proposition 6.4 and its proof is concluded by Corollary 6.11. The proofs follow the pattern of the proofs in [KW23, Section 7], but there are subtle differences throughout. The reason is partly that, unlike $PLA(\sigma)$, $PLA^+(\sigma)$ allows constructions as in part (5) of Definition 4.5 where $p^=$ is not necessarity a *complete* atomic σ-type, and partly that we allow the numbers $\alpha_{p,q}$ in Assumption 6.3 to be 0 and we did not need to bother with this "convergence to 0 case" in [KW23].

Remark 6.5 (Eliminating aggregation functions of higher arities)**.** The results below up to Proposition 6.10 are stated and proved only for (unary) admissible aggregations functions $F : [0,1]^{<\omega} \to [0,1]$ but the results hold also for admissible aggregation functions $F : \big([0,1]^{<\omega}\big)^k \to [0,1]$, where $k > 1$, and basic probability formulas $\psi_i(\bar{x},\bar{y})$, $i = 1,\dots,k$ (in place of $\psi(\bar{x},\bar{y})$). The proofs in the general case work out in the same way but the notation becomes messier since the assumptions and notation introduced in Assumption 6.7 below for $\psi(\bar{x},\bar{y})$ need to be considered for all $\psi_i(\bar{x},\bar{y})$.

Lemma 6.6. *Suppose that $\psi(\bar{x},\bar{y}) = \bigwedge_{i=1}^{t}(\psi_i(\bar{x},\bar{y}) \to c_i)$ is a basic formula where each ψ_i is a conjunction of literals. Let be $p^=(\bar{x},\bar{y})$ be an atomic $\emptyset$-type and let $F : [0,1]^{<\omega} \to [0,1]$ be an aggregation function.*

(i) *Suppose that, for some $1 \leq s \leq t$, $\psi_i(\bar{x},\bar{y}) \wedge p^=(\bar{x},\bar{y})$ is consistent if $1 \leq i \leq s$ and inconsistent if $i > s$. Then for every finite σ-structure $\mathcal{A}$ and every $\bar{a} \in A^{|\bar{x}|}$ that satisfies $p^= \restriction \bar{x}$,*

$$\mathcal{A}\big(F\big(\psi(\bar{a},\bar{y}) : \bar{y} : p^=(\bar{a},\bar{y})\big)\big) \;=\; \mathcal{A}\big(F\big(\bigwedge_{i=1}^{s}(\psi_i(\bar{a},\bar{y}) \to c_i) : \bar{y} : p^=(\bar{a},\bar{y})\big)\big).$$

(ii) *If F is admissible and $\psi_i(\bar{x},\bar{y}) \wedge p^=(\bar{x},\bar{y})$ is inconsistent for all $i = 1,\dots,t$, then for all finite σ-structures $\mathcal{A}$ and all $\bar{a} \in A^{|\bar{x}|}$ that satisfy $p^= \restriction \bar{x}$,*

$$\mathcal{A}\big(F\big(\psi(\bar{a},\bar{y}) : \bar{y} : p^=(\bar{a},\bar{y})\big)\big) \;=\; F(\bar{r})$$

where $\bar{r}$ is the sequence of length 1 the unique entry of which is 1.

Proof. (i) Suppose that the assumptions of (i) hold. If $i > s$ then $\psi_i(\bar{x},\bar{y}) \wedge p^=(\bar{x},\bar{y})$ is inconsistent, so for every finite σ-structure $\mathcal{A}$, every $\bar{a} \in A^{|\bar{x}|}$ and every $\bar{b} \in A^{|\bar{y}|}$, if $\mathcal{A} \models p^=(\bar{a},\bar{b})$ then $\mathcal{A}(\psi_i(\bar{a},\bar{b})) = 0$ and hence $\mathcal{A}(\psi_i(\bar{a},\bar{b}) \to c_i) = 1$. Now we get

$$\mathcal{A}\big(\bigwedge_{i=1}^{t}(\psi_i(\bar{a},\bar{b}) \to c_i)\big) = \min\{\mathcal{A}(\psi_i(\bar{a},\bar{b}) \to c_i) : i = 1,\dots,t\} =$$
$$\min\{\mathcal{A}(\psi_i(\bar{a},\bar{b}) \to c_i) : i = 1,\dots,s\} = \mathcal{A}\big(\bigwedge_{i=1}^{s}(\psi_i(\bar{a},\bar{b}) \to c_i)\big).$$

(ii) Suppose that $\psi_i(\bar{x},\bar{y}) \wedge p^=(\bar{x},\bar{y})$ is inconsistent for all $i = 1,\dots,t$. Then, for every finite σ-structure $\mathcal{A}$, every $\bar{a} \in A^{|\bar{x}|}$ and every $\bar{b} \in A^{|\bar{y}|}$, if $\mathcal{A} \models p^=(\bar{a},\bar{b})$ then we have $\mathcal{A}(\psi_i(\bar{a},\bar{b}) \to c_i) = 1$ for all i, so $\mathcal{A}(\psi(\bar{a},\bar{b})) = 1$. Hence, the sequence $\bar{r} = (\mathcal{A}(\psi(\bar{a},\bar{y})) : \bar{y} : p^=(\bar{a},\bar{y}))$ is constantly 1. Let $\bar{r}'$ be the sequence of length 1 the only entry of which is 1. Then $\mu_1^u(\bar{r},\bar{r}') = 0$. Since F is strongly admissible it is strongly admissible sensu novo (according to Proposition 3.12) and by condition (1) in the definition of admissibility sensu novo (Definition 3.11) we get $F(\bar{r}) = F(\bar{r}')$. □

Assumption 6.7. Until Proposition 6.10 we make, without loss of generality, the following assumptions: Let $\kappa \in \mathbb{N}^+$ and let $\bar{x}$ and $\bar{y}$ be sequences of distinct variables such that $\text{rng}(\bar{x}) \cap \text{rng}(\bar{y}) = \emptyset$ and $|\bar{x}| + |\bar{y}| \leq \kappa \in \mathbb{N}^+$. Let $\psi(\bar{x},\bar{y})$ be a basic probability formula. Then $\psi(\bar{x},\bar{y})$ is equivalent to a basic probability formula of the form

$$\bigwedge_{i=1}^{s}\bigwedge_{j=1}^{t_i}\big(p_{i,j}(\bar{x},\bar{y}) \to c_{i,j}\big),$$

where each $p_{i,j}(\bar{x},\bar{y})$ is a complete atomic σ-type and (since $c_{i,j}$ may be zero) every complete atomic σ-type in the variables $\bar{x}\bar{y}$ equals $p_{i,j}$ for some $1 \leq i \leq s$ and $1 \leq j \leq t_i$. Furthermore, we may assume (by reordering if necessary) that for all $i = 1, \dots, s$ and all $1 \le j, j' \le t_i$, $p_{i,j} \restriction \bar{x} = p_{i,j'} \restriction \bar{x}$. Let $q_i(\bar{x}) = p_{i,1} \restriction \bar{x}$ for each i. *Without loss of generality we may therefore assume that $\psi(\bar{x}, \bar{y})$ has the above described form.*

Lemma 6.8. *Suppose that $F : [0,1]^{<\omega} \to [0,1]$ is a strongly admissible aggregation function and let $p^{=}(\bar{x}, \bar{y})$ be an atomic $\emptyset$-type. Fix an index $1 \le i \le s$. There is $d_i \in [0,1]$, depending only on ψ, $p^{=}$ and F, such that for every $\varepsilon > 0$ there is $\delta > 0$ such that for all sufficiently large n, all $\mathcal{A} \in \mathbf{Y}_n^{\kappa,\delta}$, and all $\bar{a} \in [n]^{|\bar{x}|}$, if $\mathcal{A} \models q_i(\bar{a})$, then*

$$\left|\mathcal{A}\big(F\big(\psi(\bar{a}, \bar{y}) : \bar{y} : p^{=}(\bar{a}, \bar{y})\big)\big) - d_i\right| < \varepsilon.$$

Proof. Fix $1 \le i \le s$. If $p_{i,j}(\bar{x}, \bar{y}) \wedge p^{=}(\bar{x}, \bar{y})$ is inconsistent for all $j = 1, \dots, t_i$, then, by Lemma 6.6 (ii), the conclusion is immediate.

So now suppose that there is at least one j such that $p_{i,j}(\bar{x}, \bar{y}) \wedge p^{=}(\bar{x}, \bar{y})$ is consistent. By Lemma 6.6 (i), we may without loss of generality modify ψ and assume that $p_{i,j}(\bar{x}, \bar{y}) \wedge p^{=}(\bar{x}, \bar{y})$ is consistent, hence $p^{=} \subseteq p_{i,j}$, for all $j = 1, \dots, t_i$, and that $p_{i,j}$, $j = 1, \dots, t_i$, enumerates all complete atomic σ-types in the variables $\bar{x}\bar{y}$ which extend q_i and $p^{=}$.

Let $l = \dim_{\bar{y}}(p^{=})$ and, for $j = 1, \dots, t_i$, let $l_j = \dim_{\bar{y}}(p_{i,j})$. Suppose first that $l = 0$. Then $p^{=} \wedge q_i(\bar{x})$ has a unique extension to a complete atomic σ-type with variables $\bar{x}, \bar{y}$, so $t_i = 1$. Also, if $\mathcal{A} \models p_{i,j}(\bar{a}, \bar{b})$, then $\mathrm{rng}(\bar{b}) \subseteq \mathrm{rng}(\bar{a})$. It follows that for every finite σ-structure $\mathcal{A}$ and $\bar{a} \in A^{|\bar{a}|}$ such that $q_i(\bar{a})$ holds there is a unique $\bar{b} \in A^{|\bar{y}|}$ such that $\mathcal{A} \models p_{i,j}(\bar{a}, \bar{b})$ and hence the sequence $(\mathcal{A}(\psi(\bar{a}, \bar{y}) : \bar{y} : p^{=}(\bar{a}, \bar{y}))$ has a single coordinate which is $c_{i,1}$. Therefore we can let $d_i = F(\bar{r})$ where $\bar{r}$ is the sequence of length 1 where the only coordinate is $c_{i,1}$.

Now suppose that $l > 0$. If $l_j < l$ then let $\alpha_j = 0$. If $l_j = l$ then let $\alpha_j \in [0,1]$ be the number associated to $p_{i,j}$ and q_i by Assumption 6.3.

Let $\delta > 0$, $\mathcal{A}_1 \in \mathbf{Y}_{n_1}^{\kappa,\delta}$, $\mathcal{A}_2 \in \mathbf{Y}_{n_2}^{\kappa,\delta}$, $\bar{a}_1 \in [n_1]^{|\bar{x}|}$, $\bar{a}_2 \in [n_2]^{|\bar{x}|}$, $\mathcal{A}_1 \models q_i(\bar{a}_1)$, $\mathcal{A}_2 \models q_i(\bar{a}_2)$, and, for $k = 1, 2$, let

$$\bar{r}_k = \big(\mathcal{A}_k\big(\psi(\bar{a}_k, \bar{b})\big) : \bar{b} \in [n_k]^{|\bar{y}|} \text{ and } \mathcal{A} \models p^{=}(\bar{a}, \bar{b})\big). \tag{6.1}$$

It follows directly from the definition of $\bar{r}_k$ and assumptions about ψ that $\mathrm{rng}(\bar{r}_1), \mathrm{rng}(\bar{r}_2) \subseteq \{c_{i,1}, \dots, c_{i,t_i}\}$. To prove the lemma it now suffices to show that for any $\varepsilon > 0$ and all large enough n_1 and n_2, $|F(\bar{r}_1) - F(\bar{r}_2)| < \varepsilon$. Since we assume that F is strongly admissible it follows from Proposition 3.12 that F is strongly admissible sensu novo. From Condition (1) of the definition of strong admissibility sensu novo (Definition 3.11) it follows that it now suffices to show that there is a constant $C > 0$ which depends only on ψ and $p^{=}$ such that if n_1 and n_2 are sufficiently large, then $\mu_1^u(\bar{r}_1, \bar{r}_2) < \delta C$.

Let $k \in \{1, 2\}$ and note that $|\bar{r}_k| \sim (n_k)^l$. By Assumption 6.3, if n_k is large enough the following holds for each $j = 1, \dots, t_i$:

- If $l_j < l$ then $|p_{i,j}(\bar{a}_k, \mathcal{A}_k)| \le (n_k)^{l_j}$ where $\frac{(n_k)^{l_j}}{(n_k)^l} \to 0$ as $n_k \to \infty$, so $|p_{i,j}(\bar{a}_k, \mathcal{A}_k)| \le \delta(n_k)^l = (\alpha_j + \delta)(n_k)^l$, as $\alpha_j = 0$ in this case.
- If $l_j = l$ then $(\alpha_j - \delta)(n_k)^l \le |p_{i,j}(\bar{a}_k, \mathcal{A}_k)| \le (\alpha_j + \delta)(n_k)^l$.

Let $c \in [0,1]$ and suppose that there are exactly m indices $j = j_1, \dots, j_m$ such that $c_{i,j} = c$. Every $\bar{b} \in p_{i,j}(\bar{a}_k, \mathcal{A}_k)$ contributes to a coordinate $c_{i,j}$ in the sequence $\bar{r}_k$. Therefore the number c will occur between

$$(\alpha_{j_1} + \ldots + \alpha_{j_m} - m\delta)(n_k)^l \quad \text{and} \quad (\alpha_{j_1} + \ldots + \alpha_{j_m} + m\delta)(n_k)^l$$

times in $\bar{r}_k$. This implies that $\mu_1^u(\bar{r}_1, \bar{r}_2) \leq \delta C$ for a constant C that depends only on t_i which in turn depends only on ψ and $p^=$. This concludes the proof. □

Corollary 6.9. *Let $F : [0,1]^{<\omega} \to [0,1]$ be a strongly admissible aggregation function and let $p^=(\bar{x}, \bar{y})$ be an atomic $\emptyset$-type. Then there is a basic probability formula $\theta(\bar{x})$ such that for every $\varepsilon > 0$ there is $\delta > 0$ such that for all sufficiently large n, all $\mathcal{A} \in \mathbf{Y}_n^{\kappa,\delta}$, and all $\bar{a} \in [n]^{|\bar{x}|}$,*

$$\left|\mathcal{A}\big(F\big(\psi(\bar{a}, \bar{y}) : \bar{y} : p^=(\bar{a}, \bar{y})\big)\big) - \mathcal{A}\big(\theta(\bar{a})\big)\right| < \varepsilon.$$

Proof. Recall that from Assumption 6.7 $q_i(\bar{x}) = p_{i,j} \restriction \bar{x}$ for all i (and all $j = 1, \ldots, t_i$). By Lemma 6.6 (i), we may without loss of generality modify ψ and assume that $p_{i,j}(\bar{x}, \bar{y}) \wedge p^=(\bar{x}, \bar{y})$ is consistent for all $j = 1, \ldots, t_i$, and that $p_{i,j}$, $j = 1, \ldots, t_i$, enumerates all complete atomic σ-types in the variables $\bar{x}\bar{y}$ which extend q_i and $p^=$.

For every $i = 1, \ldots, s$, let $d_i \in [0,1]$ be as in Lemma 6.8. Let $q'_1(\bar{x}), \ldots, q'_m(\bar{x})$ enumerate all complete atomic $\emptyset$-types in the variables $\bar{x}$ which are different from $p^= \restriction \bar{x}$. We show that if $\theta(\bar{x})$ is the formula $\bigwedge_{i=1}^{s}(q_i(\bar{x}) \to d_i) \ \wedge \ \bigwedge_{j=1}^{m}(q'_j(\bar{x}) \to 0)$ then the lemma holds. Let $\varepsilon > 0$. Let $\mathcal{A} \in \mathbf{Y}_n^{\kappa,\delta}$ and $\bar{a} \in [n]^{|\bar{x}|}$. If $\bar{a}$ does not satisfy $p^= \restriction \bar{x}$, then it satisfies $q'_j(\bar{x})$ for some j and (no matter what δ is)

$$\mathcal{A}\big(F\big(\psi(\bar{a}, \bar{y}) : \bar{y} : p^=(\bar{a}, \bar{y})\big)\big) \ = \ 0 \ = \ \mathcal{A}\Big(\bigwedge_{i=1}^{s}(q_i(\bar{a}) \to d_i) \ \wedge \ \bigwedge_{j=1}^{m}(q'_j(\bar{a}) \to 0)\Big).$$

Now suppose that $\bar{a}$ satisfies $p^= \restriction \bar{x}$ and hence it satisfies $q_i(\bar{x})$ for some i. Then

$$\mathcal{A}\Big(\bigwedge_{i=1}^{s}(q_i(\bar{a}) \to d_i) \ \wedge \ \bigwedge_{j=1}^{m}(q'_j(\bar{a}) \to 0)\Big) \ = \ d_i. \tag{6.2}$$

From Lemma 6.8 we have that if $\delta > 0$ is small enough, then for every $i = 1, \ldots, s$, all sufficiently large n, all $\mathcal{A} \in \mathbf{Y}_n^{\kappa,\delta}$, and all $\bar{a} \in [n]^{|\bar{x}|}$, if $\mathcal{A} \models q_i(\bar{a})$, then

$$\left|\mathcal{A}\big(F\big(\psi(\bar{a}, \bar{y}) : \bar{y} : p^=(\bar{a}, \bar{y})\big)\big) - d_i\right| < \varepsilon. \tag{6.3}$$

The corollary now follows from (6.2) and (6.3). □

Proposition 6.10. *Suppose that $\varphi(\bar{x}, \bar{y}), \psi(\bar{x}, \bar{y}) \in PLA^+(\sigma)$ are asymptotically equivalent and that $\psi(\bar{x}, \bar{y})$ is a basic probability formula. Let $p^=(\bar{x}, \bar{y})$ be an atomic $\emptyset$-type. If $F : [0,1]^{<\omega} \to [0,1]$ is a strongly admissible aggregation function, then $F\big(\varphi(\bar{x}, \bar{y}) : \bar{y} : p^=(\bar{x}, \bar{y})\big)$ is asymptotically equivalent to a basic probability formula.*

Proof. Suppose that $\varphi(\bar{x}, \bar{y}), \psi(\bar{x}, \bar{y}) \in PLA(\sigma)$ are asymptotically equivalent and that $\psi(\bar{x}, \bar{y})$ is a basic probability formula. Without loss of generality we may assume that ψ has the form described in Assumption 6.7. Let $\kappa = |\bar{x}| + |\bar{y}|$ and $\varepsilon > 0$. By Corollary 6.9 there is a basic probability formula $\theta(\bar{x})$ such that for all small enough $\delta > 0$ and large enough n, if $\mathcal{A} \in \mathbf{Y}_n^{\kappa,\delta}$ and $\bar{a} \in [n]^{|\bar{x}|}$, then

$$\left|\mathcal{A}\big(F\big(\psi(\bar{a}, \bar{y}) : \bar{y} : p^=(\bar{a}, \bar{y})\big)\big) \ - \ \mathcal{A}\big(\theta(\bar{a})\big)\right| < \varepsilon/2. \tag{6.4}$$

For $\delta > 0$ and $n \in \mathbb{N}^+$ let

$$\mathbf{X}_n^{\delta} = \big\{\mathcal{A} \in \mathbf{W}_n : \text{for all } \bar{a} \in [n]^{|\bar{x}|} \text{ and all } \bar{b} \in [n]^{|\bar{y}|} \text{ such that } p^=(\bar{a}, \bar{b}) \text{ holds,}$$
$$\left|\mathcal{A}(\varphi(\bar{a}, \bar{b})) - \mathcal{A}(\psi(\bar{a}, \bar{b}))\right| < \delta\big\}.$$

Since $\varphi(\bar{x}, \bar{y})$ and $\psi(\bar{x}, \bar{y})$ are asymptotically equivalent we have $\lim_{n\to\infty} \mathbb{P}_n(\mathbf{X}_n^\delta) = 1$. By Assumption 6.3 we have $\lim_{n\to\infty} \mathbb{P}_n(\mathbf{Y}_n^{\kappa,\delta}) = 1$ and hence $\lim_{n\to\infty} \mathbb{P}_n(\mathbf{X}_n^\delta \cap \mathbf{Y}_n^{\kappa,\delta}) = 1$.

It now suffices to prove that if $\delta > 0$ is small enough, then for all sufficiently large n, all $\mathcal{A} \in \mathbf{X}_n^\delta \cap \mathbf{Y}_n^{\kappa,\delta}$ and all $\bar{a} \in [n]^{|\bar{x}|}$,

$$\big|\mathcal{A}\big(F\big(\varphi(\bar{a}, \bar{y}) : \bar{y} : p^{=}(\bar{a}, \bar{y})\big)\big) - \mathcal{A}\big(\theta(\bar{a})\big)\big| < \varepsilon. \tag{6.5}$$

The statement (6.5) follows from (6.4) and the following (to be proved)

$$\big|\mathcal{A}\big(F\big(\varphi(\bar{a}, \bar{y}) : \bar{y} : p^{=}(\bar{a}, \bar{y})\big)\big) - \mathcal{A}\big(F\big(\psi(\bar{a}, \bar{y}) : \bar{y} : p^{=}(\bar{a}, \bar{y})\big)\big) < \varepsilon/2. \tag{6.6}$$

Hence it remains to prove that if $\delta > 0$ is small enough then (6.6) holds for all sufficiently large n, all $\mathcal{A} \in \mathbf{X}_n^\delta \cap \mathbf{Y}_n^{\kappa,\delta}$ and all $\bar{a} \in [n]^{|\bar{x}|}$.

Let $\mathcal{A} \in \mathbf{X}_n^\delta \cap \mathbf{Y}_n^{\kappa,\delta}$ and $\bar{a} \in [n]^{|\bar{x}|}$. If $\bar{a}$ does not satisfy $p^{=}\restriction\bar{x}$ then

$$\mathcal{A}\big(F\big(\varphi(\bar{a}, \bar{y}) : \bar{y} : p^{=}(\bar{a}, \bar{y})\big)\big) \; = \; 0 \; = \; \mathcal{A}\big(F\big(\psi(\bar{a}, \bar{y}) : \bar{y} : p^{=}(\bar{a}, \bar{y})\big)\big).$$

Now suppose that $\bar{a}$ satisfies $p^{=}\restriction\bar{x}$ and hence $\bar{a}$ satisfies $q_i(\bar{x})$ (as in Assumption 6.7) for some i. Then the following two sequences are nonempty:

$$\bar{r} = \big(\mathcal{A}\big(\varphi(\bar{a}, \bar{b})\big) : \bar{b} \in [n]^{|\bar{y}|} \text{ and } p^{=}(\bar{a}, \bar{b}) \text{ holds}\big),$$
$$\bar{\rho} = \big(\mathcal{A}\big(\psi(\bar{a}, \bar{b})\big) : \bar{b} \in [n]^{|\bar{y}|} \text{ and } p^{=}(\bar{a}, \bar{b}) \text{ holds}\big).$$

First suppose that every $p_{i,j}(\bar{x}, \bar{y})$ (as in Assumption 6.7) is *in*consistent with $p^{=}$. Then all entries in $\bar{\rho}$ are equal to 1. Since $\mathcal{A} \in \mathbf{X}_n^\delta$ we get $\mu_\infty^o(\bar{r}, \bar{\rho}) < \delta$. Since F is strongly admissible, hence strongly admissible sensu novo, if follows from Condition (2) of the definition of strong admissibility sensu novo, that if δ is small enough, then $|F(\bar{r}) - F(\bar{\rho})| < \varepsilon/2$ and (6.6) follows immediately from this.

Now suppose that at least one $p_{i,j}$ is consistent with $p^{=}$. By Lemma 6.6 we may, without loss of generality, assume that every $p_{i,j}$ is consistent with $p^{=}$. Then we can argue in the same way as we argued in the proof of Lemma 6.8 and conclude that there are $\alpha_j, c_{i,j} \in [0, 1]$, corresponding to $p_{i,j}$, for $j = 1, \dots, t_i$, such that if $c \in [0, 1]$ and $j_1, \dots, j_m$ enumerates all $c_{i,j}$ such that $c_{i,j} = c$, then c appears between

$$(\alpha_{j_1} + \ldots + \alpha_{j_m} - m\delta)n^l \quad \text{and} \quad (\alpha_{j_1} + \ldots + \alpha_{j_m} + m\delta)n^l$$

times in $\bar{\rho}$. As $\mathcal{A} \in \mathbf{X}_n^\delta$ we get $\mu_\infty^o(\bar{r}, \bar{\rho}) < \delta$. Since F is strongly admissible, hence strongly admissible sensu novo, it follows from Condition (2) of the definition of admissibility sensu novo, that if δ is small enough, then $|F(\bar{r}) - F(\bar{\rho})| < \varepsilon/2$ which implies that (6.6) holds. □

Corollary 6.11. *Let $\varphi(\bar{x}) \in PLA^+(\sigma)$ and suppose that all aggregation functions in φ are strongly admissible. Then $\varphi(\bar{x})$ is asymptotically equivalent to a basic probability formula.*

Proof. We use induction on the complexity of formulas. If $\varphi(\bar{x})$ is of one of the forms described in parts (1) and (2) of the definition of $PLA^+(\sigma)$ (Definition 4.5), then $\varphi(\bar{x})$ is aggregation-free and then $\varphi(\bar{x})$ is asymptotically equivalent to a basic probability formula by virtue of Lemma 4.13, since equivalence implies asymptotic equivalence.

Now suppose that $\mathsf{C} : [0, 1]^k \to [0, 1]$ is a continuous connective and $\varphi_1(\bar{x}), \dots, \varphi_k(\bar{x}) \in PLA^+(\sigma)$, so $\mathsf{C}(\varphi_1(\bar{x}), \dots, \varphi_k(\bar{x})) \in PLA^+(\sigma)) \in PLA^+(\sigma)$. If each $\varphi_i(\bar{x})$ is asymptotically equivalent to a basic probability formula $\varphi_i'(\bar{x})$, then it follows from the continuity of C that $\mathsf{C}(\varphi_1(\bar{x}), \dots, \varphi_k(\bar{x}))$ is asymptotically equivalent to $\mathsf{C}(\varphi_1'(\bar{x}), \dots, \varphi_k'(\bar{x}))$. Since $\mathsf{C}(\varphi_1'(\bar{x}), \dots, \varphi_k'(\bar{x}))$ is aggregation-free it follows from Lemma 4.13 that it is equivalent to a basic probability formula $\psi(\bar{x})$. But then $\mathsf{C}(\varphi_1(\bar{x}), \dots, \varphi_k(\bar{x}))$ is asymptotically equivalent to $\psi(\bar{x})$.

Now suppose that $F : \big([0,1]^{<\omega}\big)^k \to [0,1]$ is a strongly admissible aggregation function, $p^=(\bar{x}, \bar{y})$ is an atomic σ-type, $\varphi_1(\bar{x}, \bar{y}), \dots, \varphi_k(\bar{x}, \bar{y}) \in PLA^+(\sigma)$ and that each $\varphi_i(\bar{x}, \bar{y})$ is asymptotically equivalent to a basic probability formula $\varphi'_i(\bar{x}, \bar{y})$. Then Proposition 6.10 combined with Remark 6.5 implies that $F(\varphi_1(\bar{x}, \bar{y}), \dots, \varphi_k(\bar{x}, \bar{y}) : \bar{y} : p^=(\bar{x}, \bar{y}))$ is asymptotically equivalent to a basic probability formula. □

7. Saturation and convergence of atomic types: finishing the proof of Theorem 5.11

Theorem 5.11 follows from Proposition 6.4 and the proofs in this section. We consider a finite relational signature σ and a $PLA^+(\sigma)$-network $\mathbb{G}$ whose every probability formula uses only *strongly admissible* aggregation functions (if it uses any at all). We want to prove that the claims of Theorem 5.11 hold. For this we will use induction on the maximal path rank, or mp-rank, of the underlying DAG of $\mathbb{G}$, also denoted $\mathbb{G}$. In order to make the inductive step work out we have to prove (in the base case and in the inductive step) a few claims none of which explicitly states that the claims of Theorem 5.11 hold. But these few claims (labelled (1)–(5) in Assumption 7.2) in conjunction with Proposition 6.4 imply Theorem 5.11. Perhaps a bit counter-intuitively the base case of the induction will *not* be the case when $\mathbb{G}$ has mp-rank 0 (i.e. when $\mathbb{G}$ has no edges). Instead the base case will be when $\sigma = \emptyset$, in other words when the DAG has no vertices (and hence no edges) and in this case we have the convention that the empty DAG has mp-rank '-1'. The base case, for an empty signature, is stated by Lemma 7.1. To make the notation consistent with the notation used in the inductive step formulated in Assumption 7.2, we denote the empty signature of the base case (Lemma 7.1) by σ'.

Lemma 7.1 (The base case)**.** *Suppose that $\sigma' = \emptyset$ and, for all $n \in \mathbb{N}^+$, let $\mathbf{W}'_n$ be the set of all σ'-structures with domain $[n]$ (so $\mathbf{W}'_n$ is a singleton set), and let $\mathbb{P}'_n$ be the unique probability distribution on $\mathbf{W}'_n$.*

(a) *Let $k \in \mathbb{N}^+$, $\varepsilon' > 0$, and let $\delta' : \mathbb{N}^+ \to \mathbb{R}^{\geq 0}$ be any function such that $\lim_{n\to\infty} \delta'(n) = 0$, Then there are $\mathbf{Y}'_n \subseteq \mathbf{W}'_n$, for $n \in \mathbb{N}^+$, such that the following hold:*
   (1) $\lim_{n\to\infty} \delta'(n) = 0$.
   (2) $\mathbb{P}'_n(\mathbf{Y}'_n) \geq 1 - \delta'(n)$ *for all sufficiently large n.*
   (3) *For every complete atomic σ'-type $p'(\bar{x})$ with $|\bar{x}| \leq k$ there is a number which we denote $\mathsf{P}'(p'(\bar{x}))$, or just $\mathsf{P}'(p')$, such that for all sufficiently large n and all $\bar{a} \in [n]$ which realize the identity fragment of p',*
$$\big|\mathbb{P}'_n\big(\{\mathcal{A}' \in \mathbf{W}'_n : \mathcal{A}' \models p'(\bar{a})\}\big) \;-\; \mathsf{P}'(p'(\bar{x}))\big| \;\leq\; \varepsilon'.$$
   (4) *For every complete atomic σ'-type $p'(\bar{x}, \bar{y})$ with $|\bar{x}\bar{y}| \leq k$ and $0 < \dim_{\bar{y}}(p'(\bar{x}, y)) = d$, if $q'(\bar{x}) = p' \restriction \bar{x}$, then there is $\alpha \in [0,1]$ such that, for all sufficiently large n, every $\mathcal{A}' \in \mathbf{Y}'_n$ is $(p', q', \alpha - \varepsilon')$-saturated and $(p', q', \alpha + \varepsilon')$-unsaturated.*

(b) *If $\varphi(\bar{x}) \in PLA^+(\emptyset)$ and every aggregation function in φ is strongly admissible, then $\varphi(\bar{x})$ is asymptotically equivalent, with respect to $(\mathbb{P}'_n : n \in \mathbb{N}^+)$, to a basic probability formula.*

Proof. (a) Suppose that $\sigma' = \emptyset$ and let $k \in \mathbb{N}^+$ and $\varepsilon' > 0$ be given. Also let $\delta' : \mathbb{N}^+ \to \mathbb{R}^{\geq 0}$ be any function such that $\lim_{n\to\infty} \delta'(n) = 0$, so (1) holds. For every complete atomic σ'-type $p'(\bar{x})$ let $\mathsf{P}'(p'(\bar{x})) = 1$. Observe that, for every n, if $\bar{a} \in [n]$ and $\bar{a}$ realizes the identity fragment of $p'(\bar{x})$, then $\bar{a}$ realizes $p'(\bar{x})$ in the unique $\mathcal{A}'$ of $\mathbf{W}'_n$. Hence, for trivial reasons we have (3).

For every n let $\mathbf{Y}'_n$ be the set of all $\mathcal{A}' \in \mathbf{W}'_n$ such that for every complete atomic σ'-type $p'(\bar{x}, \bar{y})$ with $|\bar{x}\bar{y}| \leq k$ and $0 < \dim_{\bar{y}}(p'(\bar{x}, \bar{y})) = |\bar{y}|$, if $q(\bar{x}) = p{\restriction}\bar{x}$, then for all sufficiently large n, every $\mathcal{A}' \in \mathbf{Y}'_n$ is $(p', q', 1 - \varepsilon')$-saturated and $(p', q', 1 + \varepsilon')$-unsaturated. Suppose that $p'(\bar{x}, \bar{y})$ is a complete atomic σ'-type with $|\bar{x}\bar{y}| \leq k$ and $0 < \dim_{\bar{y}}(p'(\bar{x}, \bar{y})) = |\bar{y}|$. Let $q'(\bar{x}) = p'{\restriction}\bar{x}$ and suppose that $\mathcal{A}' \models q'(\bar{a})$ where $\mathcal{A}' \in \mathbf{W}'_n$. Then $\mathcal{A}' \models p'(\bar{a}, \bar{b})$ for every $\bar{b} \in [n]^{|\bar{y}|}$ consisting of different elements no one of which occurs in $\bar{a}$. There are $n^{|\bar{y}|} - Cn^{|\bar{y}|-1}$ such $\bar{b}$ for some constant C. So if $n^{|\bar{y}|} - Cn^{|\bar{y}|-1} \geq (1 - \varepsilon')n^{|\bar{y}|}$ then $\mathcal{A}'$ is $(p', q', 1 - \varepsilon')$-saturated. For trivial reasons, $\mathcal{A}'$ is also $(p', q', 1 + \varepsilon')$-unsaturated. Hence we have proved (4). The above argument shows that for all large enough n the unique member of $\mathbf{W}'_n$ belongs to $\mathbf{Y}'_n$, so it follows that (2) holds.

(b) Note that we have proved that (1) – (4) hold for all choices of $k \in \mathbb{N}^+$ and $\varepsilon' > 0$. Therefore Assumption 6.3 holds and consequently, by Proposition 6.4, if $\varphi(\bar{x}) \in PLA^+(\emptyset)$ and every aggregation function in φ is strongly admissible, then $\varphi(\bar{x})$ is asymptotically equivalent to a basic probability formula. □

Note that Lemma 7.1 implies the statement of Theorem 5.11 in the case when $\sigma = \emptyset$. For the rest of this section we make the following assumptions: σ is a nonempty finite relational signature. $\mathbb{G}$ is a $PLA^+(\sigma)$-network with mp-rank $\rho \geq 0$. For every $R \in \sigma$, the corresponding probability formula θ_R contains only strongly admissible aggregation functions. σ' is the set of $R \in \sigma$ such that the mp-rank of R is less than ρ and $\mathbb{G}'$ is the $PLA^+(\sigma')$-subnetwork induced by σ'. $\mathbf{W}'_n$ be the set of all σ'-structures with domain $[n]$ and $\mathbf{W}_n$ is the set of σ-structures with domain $[n]$. $\mathbb{P}'_n$ and $\mathbb{P}_n$ are the probability distributions induced by $\mathbb{G}'$ and $\mathbb{G}$ on $\mathbf{W}'_n$ and $\mathbf{W}_n$, respectively. Now we assume the following (which, by Lemma 7.1, holds if $\sigma' = \emptyset$):

Assumption 7.2 (Induction hypothesis)**.** For all $k \in \mathbb{N}^+$ and all $\varepsilon' > 0$ there are $\delta' : \mathbb{N}^+ \to \mathbb{R}^{\geq 0}$ and $\mathbf{Y}'_n \subseteq \mathbf{W}'_n$, for $n \in \mathbb{N}^+$, such that the following hold:
(1) $\lim_{n\to\infty} \delta'(n) = 0$.
(2) $\mathbb{P}'_n(\mathbf{Y}'_n) \geq 1 - \delta'(n)$ for all sufficiently large n.
(3) For every complete atomic σ'-type $p'(\bar{x})$ with $|\bar{x}| \leq k$ there is a number which we denote $\mathsf{P}'(p'(\bar{x}))$, or just $\mathsf{P}'(p')$, such that for all sufficiently large n and all $\bar{a} \in [n]$ which realize the identity fragment of p',

$$\left|\mathbb{P}'_n\big(\{\mathcal{A}' \in \mathbf{W}'_n : \mathcal{A}' \models p'(\bar{a})\}\big) \; - \; \mathsf{P}'(p'(\bar{x}))\right| \; \leq \; \varepsilon'.$$

(4) For every complete atomic σ'-type $p'(\bar{x}, \bar{y})$ with $|\bar{x}\bar{y}| \leq k$ and $0 < \dim_{\bar{y}}(p'(\bar{x}, \bar{y})) = d$, if $q'(\bar{x}) = p'{\restriction}\bar{x}$, then there is $\alpha \in [0, 1]$ such that, for all sufficiently large n, every $\mathcal{A}' \in \mathbf{Y}'_n$ is $(p', q', \alpha - \varepsilon')$-saturated and $(p', q', \alpha + \varepsilon')$-unsaturated.
(5) For every $R \in \sigma \setminus \sigma'$ there is a basic probability formula $\chi_R(\bar{x}) \in PLA^+(\sigma')$ such that $\chi_R(\bar{x}) \sim_{\mathbb{G}'} \theta_R(\bar{x})$, where $\theta_R \in PLA^+(\sigma')$ is the probability formula corresponding to R in $\mathbb{G}$, and for all sufficiently large n, all $\mathcal{A}' \in \mathbf{Y}'_n$ and all $\bar{a} \in [n]^{|\bar{x}|}$,

$$\left|\mathcal{A}'(\theta_R(\bar{a})) - \mathcal{A}'(\chi_R(\bar{a}))\right| \leq \varepsilon'.$$

Outline of the inductive step and how conditions (1) – (5) imply Theorem 5.11. We fix some arbitrary $k \in \mathbb{N}^+$ and $\varepsilon' > 0$. The we carry out an argument, finished in Proposition 7.18, which shows that we can find $\varepsilon > 0$, which can be made as small as we like if ε' is chosen small enough, $\delta : \mathbb{N}^+ \to \mathbb{R}^{\geq 0}$, and $\mathbf{Y}_n \subseteq \mathbf{W}_n$, for $n \in \mathbb{N}^+$, such that if σ',

ε', δ', $\mathbf{W}'_n$, $\mathbf{Y}'_n$ and $\mathbb{P}'_n$ are replaced by σ, ε, δ, $\mathbf{W}_n$, $\mathbf{Y}_n$ and $\mathbb{P}_n$, respectively, then parts (1) – (4) of Assumption 7.2 hold. It follows, as stated in Corollary 7.19, that *for every k and every $\varepsilon > 0$*, there are $\delta : \mathbb{N}^+ \to \mathbb{R}^{\geq 0}$ and $\mathbf{Y}_n \subseteq \mathbf{W}_n$, for $n \in \mathbb{N}^+$, such that if σ', ε', δ'_n, $\mathbf{W}'_n$, $\mathbf{Y}'_n$ and $\mathbb{P}'_n$ are replaced by σ, ε, δ_n, $\mathbf{W}_n$, $\mathbf{Y}_n$ and $\mathbb{P}_n$, respectively, then parts (1) – (4) of Assumption 7.2 hold. Hence Assumption 6.3 holds and therefore Proposition 6.4 implies that if $\varphi(\bar{x}) \in PLA^+(\sigma)$ and all aggregation functions in φ are strongly admissible, then $\varphi(\bar{x})$ is asymptotically equivalent to a basic probability formula. This is stated in Corollary 7.20.

It remains to show that Part (5) of Assumption 7.2 holds for every $R \in \sigma^* \setminus \sigma$ if $\sigma \subset \sigma^*$ and the corresponding probability formula $\theta_R \in PLA^+(\sigma)$ (of a $PLA(\sigma^*)$-network $\mathbb{G}^*$ having $\mathbb{G}$ as a subnetwork) contains only strongly admissible aggregation functions. So suppose that $\sigma^* \supset \sigma$ and $\mathbb{G}^*$ is a $PLA^+(\sigma^*)$-network such that $\mathbb{G}$ is a subnetwork of $\mathbb{G}^*$ and for every $R \in \sigma^* \setminus \sigma$ the corresponding probability formula $\theta_R \in PLA^+(\sigma)$ contains only strongly admissible aggregation functions. Then, by Corollary 7.20, for every $R \in \sigma^* \setminus \sigma$ there is a basic probability formula χ_R such that $\theta_R \sim_{\mathbb{G}} \chi_R$. Since σ^* is finite, it follows that for any $\varepsilon > 0$ there are $\mathbf{Y}^*_n \subseteq \mathbf{W}_n$ for all $n \in \mathbb{N}^+$ such that $\lim_{n\to\infty} \mathbb{P}_n(\mathbf{Y}^*_n) = 1$ and the following holds for all n: If $R \in \sigma^* \setminus \sigma$, $\mathcal{A} \in \mathbf{Y}^*_n$ and $\bar{a} \in [n]^r$, where r is the arity of R, then

$$|\mathcal{A}(\theta_R(\bar{a})) - \mathcal{A}(\chi_R(\bar{a}))| \leq \varepsilon. \tag{7.1}$$

We can now replace $\mathbf{Y}_n$ by $\mathbf{Y}_n \cap \mathbf{Y}^*_n$, but in order to not introduce new notation we still call this new set $\mathbf{Y}_n$. By modifying δ_n slightly if necessary the conditions (1) – (4) still hold if σ', ε', δ'_n, $\mathbf{W}'_n$, $\mathbf{Y}'_n$ and $\mathbb{P}'_n$ are replaced by σ, ε, δ_n, $\mathbf{W}_n$, $\mathbf{Y}_n$ and $\mathbb{P}_n$, respectively. In addition, condition (5) now holds if σ', σ, $\mathbb{G}'$, $\mathbb{G}$ and $\mathbf{Y}'_n$ are replaced by σ, σ^*, $\mathbb{G}$, $\mathbb{G}^*$ and $\mathbf{Y}_n$, respectively.

This reduces the proof of the inductive step to Proposition 7.18, Corollary 7.19 and Corollary 7.20. Thus it remains to prove these results. But before doing this, we explain how Theorem 5.11 follows from the conditions (1) – (5) of Assumption 6.3 with σ', ε', δ'_n, $\mathbf{W}'_n$, $\mathbf{Y}'_n$ and $\mathbb{P}'_n$ replaced by σ, ε, δ_n, $\mathbf{W}_n$, $\mathbf{Y}_n$ and $\mathbb{P}_n$, respectively. Part (ii) of Theorem 5.11 follows from condition (3) with the mentioned replacements as stated by Corollary 7.19. Condition (4) (with the mentioned replacements) implies that Assumption 6.3 holds and therefore Proposition 6.4 implies Part (i) of Theorem 5.11. The statement which is relevant for Part (i) of Theorem 5.11 appears clearly in Corollary 7.20.

From now until Proposition 7.18 we fix $k \in \mathbb{N}^+$ and $\varepsilon' > 0$. In the proofs that follow we will consider relativizations of $\mathbb{P}_n$ to some subsets of $\mathbf{W}_n$ according to the next definition.

Definition 7.3. (i) If $\mathbf{Y}' \subseteq \mathbf{W}'_n$ then we define

$$\mathbf{W}^{\mathbf{Y}'} = \{\mathcal{A} \in \mathbf{W}_n : \mathcal{A}{\restriction}\sigma' \in \mathbf{Y}'\} \quad \text{and if } \mathcal{A} \in \mathbf{W}^{\mathbf{Y}'} \text{ and } \mathcal{A}{\restriction}\sigma' = \mathcal{A}', \text{ then}$$

$$\mathbb{P}^{\mathbf{Y}'}(\mathcal{A}) \;=\; \frac{\mathbb{P}'_n(\mathcal{A}')}{\mathbb{P}'_n(\mathbf{Y}')} \prod_{R \in \sigma \setminus \sigma'} \prod_{\bar{a} \in R^{\mathcal{A}}} \mathcal{A}'\big(\theta_R(\bar{a})\big) \prod_{\bar{a} \in [n]^{k_R} \setminus R^{\mathcal{A}}} \big(1 - \mathcal{A}'\big(\theta_R(\bar{a})\big)\big)$$

where k_R is the arity of R.
(ii) If $\mathcal{A}' \in \mathbf{W}'_n$, then we let

$$\mathbf{W}^{\mathcal{A}'} = \mathbf{W}^{\{\mathcal{A}'\}} \quad \text{and, for every } \mathcal{A} \in \mathbf{W}^{\mathcal{A}'},$$

$$\mathbb{P}^{\mathcal{A}'}(\mathcal{A}) = \mathbb{P}^{\{\mathcal{A}'\}}(\mathcal{A}) = \prod_{R \in \sigma \setminus \sigma'} \prod_{\bar{a} \in R^{\mathcal{A}}} \mathcal{A}'\big(\theta_R(\bar{a})\big) \prod_{\bar{a} \in [n]^{k_R} \setminus R^{\mathcal{A}}} \big(1 - \mathcal{A}'\big(\theta_R(\bar{a})\big)\big)$$

Then $\mathbb{P}^{\mathbf{Y}'}$ and $\mathbb{P}^{\mathcal{A}'}$ are probability distributions on $\mathbf{W}^{\mathbf{Y}'}$ and $\mathbf{W}^{\mathcal{A}'}$, respectively. Note also that if $\mathbf{Y}' \subseteq \mathbf{W}'_n$, $\mathcal{A}' \in \mathbf{Y}'$ and $\mathcal{A} \in \mathbf{W}^{\mathcal{A}'}$, then

$$\mathbb{P}^{\mathbf{Y}'}(\mathcal{A}) = \frac{\mathbb{P}'_n(\mathcal{A}')}{\mathbb{P}'_n(\mathbf{Y}')}\mathbb{P}^{\mathcal{A}'}(\mathcal{A}), \tag{7.2}$$

and in particular, taking $\mathbf{Y}' = \mathbf{W}'_n$, we have, for every $\mathcal{A} \in \mathbf{W}_n$,

$$\mathbb{P}_n(\mathcal{A}) = \mathbb{P}'_n(\mathcal{A}{\restriction}\sigma')\mathbb{P}^{\mathcal{A}{\restriction}\sigma'}(\mathcal{A}). \tag{7.3}$$

We now state a few basic lemmas which will be useful.

Lemma 7.4. *For every n, if $\mathbf{Y}' \subseteq \mathbf{W}'_n$ then $\mathbb{P}_n(\mathbf{W}^{\mathbf{Y}'}) = \mathbb{P}'_n(\mathbf{Y}')$.*

Proof. By using (7.3) in the first line below we get

$$\mathbb{P}_n(\mathbf{W}^{\mathbf{Y}'}) \;=\; \sum_{\mathcal{A}'\in\mathbf{Y}'}\sum_{\mathcal{A}\in\mathbf{W}^{\mathcal{A}'}}\mathbb{P}_n(\mathcal{A}) \;=\; \sum_{\mathcal{A}'\in\mathbf{Y}'}\sum_{\mathcal{A}\in\mathbf{W}^{\mathcal{A}'}}\mathbb{P}'_n(\mathcal{A}')\mathbb{P}^{\mathcal{A}'}(\mathcal{A}) \;=$$
$$\sum_{\mathcal{A}'\in\mathbf{Y}'}\mathbb{P}'_n(\mathcal{A}')\sum_{\mathcal{A}\in\mathbf{W}^{\mathcal{A}'}}\mathbb{P}^{\mathcal{A}'}(\mathcal{A}) \;=\; \sum_{\mathcal{A}'\in\mathbf{Y}'}\mathbb{P}'_n(\mathcal{A}') \;=\; \mathbb{P}'_n(\mathbf{Y}'). \qquad \square$$

Lemma 7.5. *For every n,*
(i) if $\mathbf{X} \subseteq \mathbf{W}_n$ and $\mathcal{A}' \in \mathbf{W}'_n$, then $\mathbb{P}_n(\mathbf{X} \mid \mathbf{W}^{\mathcal{A}'}) = \mathbb{P}^{\mathcal{A}'}(\mathbf{X} \cap \mathbf{W}^{\mathcal{A}'})$, and
(ii) if $\mathbf{X} \subseteq \mathbf{W}_n$ and $\mathbf{Y}' \subseteq \mathbf{W}'_n$, then $\mathbb{P}_n(\mathbf{X} \mid \mathbf{W}^{\mathbf{Y}'}) = \mathbb{P}^{\mathbf{Y}'}(\mathbf{X} \cap \mathbf{W}^{\mathbf{Y}'})$.

Proof. Let $\mathbf{X} \subseteq \mathbf{W}_n$.

(i) Let $\mathcal{A}' \in \mathbf{W}'_n$. Using Lemma 7.4 in the first line below and (7.3) in the second line below, we get

$$\mathbb{P}_n(\mathbf{X} \mid \mathbf{W}^{\mathcal{A}'}) \;=\; \frac{\mathbb{P}_n(\mathbf{X}\cap\mathbf{W}^{\mathcal{A}'})}{\mathbb{P}_n(\mathbf{W}^{\mathcal{A}'})} \;=\; \frac{\mathbb{P}_n(\mathbf{X}\cap\mathbf{W}^{\mathcal{A}'})}{\mathbb{P}'_n(\mathcal{A}')} \;=$$
$$\frac{\mathbb{P}'_n(\mathcal{A}')\sum_{\mathcal{A}\in\mathbf{X}\cap\mathbf{W}^{\mathcal{A}'}}\mathbb{P}^{\mathcal{A}'}(\mathcal{A})}{\mathbb{P}'_n(\mathcal{A}')} \;=\; \mathbb{P}^{\mathcal{A}'}(\mathbf{X}\cap\mathbf{W}^{\mathcal{A}'}).$$

(ii) Let $\mathbf{Y}' \subseteq \mathbf{W}'_n$. Using that $\mathbf{X} \cap \mathbf{W}^{\mathbf{Y}'}$ is the disjoint union of all $\mathbf{X} \cap \mathbf{W}^{\mathcal{A}'}$ such that $\mathcal{A}' \in \mathbf{Y}'$, Lemma 7.4, Part (i) of this lemma and (7.2), we get

$$\mathbb{P}_n(\mathbf{X} \mid \mathbf{W}^{\mathbf{Y}'}) \;=\; \frac{\mathbb{P}_n(\mathbf{X}\cap\mathbf{W}^{\mathbf{Y}'})}{\mathbb{P}_n(\mathbf{W}^{\mathbf{Y}'})} \;=\; \sum_{\mathcal{A}'\in\mathbf{Y}'}\frac{\mathbb{P}_n(\mathbf{X}\cap\mathbf{W}^{\mathcal{A}'})}{\mathbb{P}_n(\mathbf{W}^{\mathbf{Y}'})} \;=$$
$$\sum_{\mathcal{A}'\in\mathbf{Y}'}\frac{\mathbb{P}_n(\mathbf{W}^{\mathcal{A}'})}{\mathbb{P}_n(\mathbf{W}^{\mathbf{Y}'})}\,\mathbb{P}_n(\mathbf{X} \mid \mathbf{W}^{\mathcal{A}'}) \;=\; \sum_{\mathcal{A}'\in\mathbf{Y}'}\frac{\mathbb{P}'_n(\mathcal{A}')}{\mathbb{P}'_n(\mathbf{Y}')}\mathbb{P}^{\mathcal{A}'}(\mathbf{X}\cap\mathbf{W}^{\mathcal{A}'}) \;=$$
$$\sum_{\mathcal{A}'\in\mathbf{Y}'}\mathbb{P}^{\mathbf{Y}'}(\mathbf{X}\cap\mathbf{W}^{\mathcal{A}'}) \;=\; \mathbb{P}^{\mathbf{Y}'}(\mathbf{X}\cap\mathbf{W}^{\mathbf{Y}'}). \qquad \square$$

Remark 7.6 (Properties of $\mathbb{P}^{\mathcal{A}'}$)**.** Fix any n and any $\mathcal{A}' \in \mathbf{W}'_n$.
(i) If $\mathcal{A} \in \mathbf{W}^{\mathcal{A}'}$ then (by the definitions of $\mathbb{P}_n$ and $\mathbb{P}^{\mathcal{A}'}$)

$$\mathbb{P}^{\mathcal{A}'}(\mathcal{A}) = \mathbb{P}_n\big(\mathcal{A} \mid \mathcal{A}{\restriction}\sigma' = \mathcal{A}'\big).$$

It follows that if $p'(\bar{x})$ is a complete atomic σ'-type, $\mathcal{A}' \models p'(\bar{a})$ and $p(\bar{x}) \supset p'(\bar{x})$ is an atomic σ-type, then

$$\mathbb{P}^{\mathcal{A}'}\big(\{\mathcal{A} \in W^{\mathcal{A}'} : \mathcal{A} \models p(\bar{a})\}\big) \;=\; \mathbb{P}_n\big(\{\mathcal{A} \in \mathbf{W}_n : \mathcal{A} \models p(\bar{a}) \mid \{\mathcal{A} \in \mathbf{W}_n : \mathcal{A}{\restriction}\sigma' = \mathcal{A}'\}\big).$$

(ii) For every $\alpha \in \{0,1\}$, every $R \in \sigma \setminus \sigma'$ and every $\bar{a} \in [n]^r$, where r is the arity of R, let $\mathbf{E}^\alpha_{R,\bar{a}} = \{\mathcal{A} \in \mathbf{W}^{\mathcal{A}'} : \mathcal{A} \models R^\alpha(\bar{a})\}$ where R^0 and R^1 denote $\neg R$ and R, respectively. It follows from the definition of $\mathbb{P}^{\mathcal{A}'}$ that

(a) for every $R \in \sigma \setminus \sigma'$ and every $\bar{a} \in [n]^r$, where r is the arity of R, $\mathbb{P}^{\mathcal{A}'}(\mathbf{E}^1_{R,\bar{a}}) = \mathcal{A}'(\theta_R(\bar{a}))$, $\mathbb{P}^{\mathcal{A}'}(\mathbf{E}^0_{R,\bar{a}}) = 1 - \mathcal{A}'(\theta_R(\bar{a}))$, and

(b) if $\alpha_1, \ldots, \alpha_m \in \{0,1\}$, $R_1, \ldots, R_m \in \sigma \setminus \sigma'$ and $\bar{a}_1, \ldots, \bar{a}_m$ are tuples where $|\bar{a}_i|$ equals the arity of R_i for each i, and for all $1 \leq i < j \leq m$, $R_i \neq R_j$ or $\bar{a}_i \neq \bar{a}_j$, then the events $\mathbf{E}^{\alpha_1}_{R_1,\bar{a}_1}, \ldots, \mathbf{E}^{\alpha_m}_{R_m,\bar{a}_m}$ are independent.

Lemma 7.7. *Suppose that $p'(\bar{x})$ is a complete atomic σ'-type and that $p(\bar{x}) \supseteq p'(\bar{x})$ is a (possibly partial) atomic σ-type. There is a number which we denote $\mathsf{P}(p(\bar{x}) \mid p'(\bar{x}))$, or just $\mathsf{P}(p \mid p')$, such that for all sufficiently large n, all $\bar{a} \in [n]^{|\bar{x}|}$ and all $\mathcal{A}' \in \mathbf{Y}'_n$ such that $\mathcal{A}' \models p'(\bar{a})$,*

$$\left|\mathbb{P}^{\mathcal{A}'}\big(\{\mathcal{A} \in \mathbf{W}^{\mathcal{A}'} : \mathcal{A} \models p(\bar{a})\}\big) \ - \ \mathsf{P}(p(\bar{x}) \mid p'(\bar{x}))\right| \ \leq \ \varepsilon_p$$

where $\varepsilon_p > 0$ depends only on ε and on p and ε_p can be made arbitrarily small by taking ε' sufficiently small.

Proof. Let $\mathcal{A}' \in \mathbf{Y}'_n$. Suppose that $\bar{x} = (x_1, \ldots, x_m)$ and $\bar{a} = (a_1, \ldots, a_m) \in [n]^m$. It follows from Remark 7.6 and Lemma 5.8 that

$$\mathbb{P}^{\mathcal{A}'}\big(\{\mathcal{A} \in W^{\mathcal{A}'} : \mathcal{A} \models p(\bar{a})\}\big) \ = $$
$$\mathbb{P}_n\big(\{\mathcal{A} \in \mathbf{W}_n : \mathcal{A} \models p(\bar{a}) \mid \{\mathcal{A} \in \mathbf{W}_n : \mathcal{A}{\restriction}\sigma' = \mathcal{A}'\}\big) \ = $$
$$\prod_{\substack{R \in \sigma \setminus \sigma' \text{ and} \\ R(x_{i_1}, \ldots, x_{i_r}) \in p(\bar{x})}} \mathcal{A}'(\theta_R(a_{i_1}, \ldots, a_{i_r})) \prod_{\substack{R \in \sigma \setminus \sigma' \text{ and} \\ \neg R(x_{i_1}, \ldots, x_{i_r}) \in p(\bar{x})}} \big(1 - \mathcal{A}'(\theta_R(a_{i_1}, \ldots, a_{i_r}))\big).$$

Assumption 7.2 (5) now says that for every $R \in \sigma \setminus \sigma'$ there is a basic probability formula χ_R which is asymptotically equivalent to θ_R, with respect to $\mathbb{G}'$ (where θ_R is the probability formula associated to R by $\mathbb{G}$), and if r is the arity of R then the following holds for all $b_1, \ldots, b_r \in [n]$ (assuming that n is large enough):

$$\big|\mathcal{A}'(\theta_R(b_1, \ldots, b_r)) - \mathcal{A}'(\chi_R(b_1, \ldots, b_r))\big| \leq \varepsilon'.$$

Hence

$$\Big|\mathbb{P}^{\mathcal{A}'}\big(\{\mathcal{A} \in W^{\mathcal{A}'} : \mathcal{A} \models p(\bar{a})\}\big) \ - $$
$$\prod_{\substack{R \in \sigma \setminus \sigma' \\ R(x_{i_1}, \ldots, x_{i_r}) \in p(\bar{x})}} \mathcal{A}'(\chi_R(a_{i_1}, \ldots, a_{i_r})) \prod_{\substack{R \in \sigma \setminus \sigma' \\ \neg R(x_{i_1}, \ldots, x_{i_r}) \in p(\bar{x})}} \big(1 - \mathcal{A}'(\chi_R(a_{i_1}, \ldots, a_{i_r}))\big)\Big| \ \leq \ \varepsilon_p$$

where $\varepsilon_p > 0$ depends only on p and ε'. Now we note that every number $\mathcal{A}'(\chi_R(a_{i_1}, \ldots, a_{i_r}))$ in the above expression depends only on p'. The reason is that since χ_R is a basic probability formula the value $\mathcal{A}'(\chi_R(a_{i_1}, \ldots, a_{i_r}))$ depends only on the complete atomic σ'-type which is realized by $(a_{i_1}, \ldots, a_{i_r})$ in $\mathcal{A}'$ and we are assuming that $\mathcal{A}' \models p'(\bar{a})$. In other words, there is a constant, which we denote by $\mathsf{P}(p(\bar{x}) \mid p'(\bar{x}))$ such that for all sufficiently large n, all $\bar{a} \in [n]$ and all $\mathcal{A}' \in \mathbf{Y}'_n$ such that $\mathcal{A}' \models p'(\bar{a})$,

$$\left|\mathbb{P}^{\mathcal{A}'}\big(\{\mathcal{A} \in \mathbf{W}^{\mathcal{A}'} : \mathcal{A} \models p(\bar{a})\}\big) \ - \ \mathsf{P}(p(\bar{x}) \mid p'(\bar{x}))\right| \ \leq \ \varepsilon_p. \qquad \square$$

Definition 7.8. Let ε^* be the maximum of ε' and of all ε_p as in Lemma 7.7 where the atomic σ-type $p(\bar{x}, \bar{y})$ is subject to the constraint that $|\bar{x}\bar{y}| \leq k$.

Lemma 7.9. *Suppose that $p'(\bar{x})$ is a complete atomic σ'-type and that $p(\bar{x}) \supseteq p'(\bar{x})$ is a (possibly partial) atomic σ-type. Then for all sufficiently large n and all $\bar{a} \in [n]^{|\bar{x}|}$ which realize the identity fragment of $p'(\bar{x})$ (and hence of p) we have*

$$\left|\mathbb{P}_n\big(\{\mathcal{A} \in \mathbf{W}_n : \mathcal{A} \models p(\bar{a})\} \mid \mathbf{W}^{\mathbf{Y}'_n}\big) \; - \; \mathsf{P}(p(\bar{x}) \mid p'(\bar{x})) \cdot \mathsf{P}'(p'(\bar{x}))\right| \; \leq \; 7\varepsilon^*.$$

Proof. Let $\bar{a} \in [n]^{|\bar{x}|}$ realize the identity fragment of $p'(\bar{x})$. Furthermore,
let $\mathbf{X}_n$ be the set of all $\mathcal{A} \in \mathbf{W}_n$ such that $\mathcal{A} \models p(\bar{a})$,
let $\mathbf{X}'_n$ be the set of all $\mathcal{A}' \in \mathbf{W}'_n$ such that $\mathcal{A}' \models p'(\bar{a})$, and
let $\mathbf{Z}'_n$ be the set of all $\mathcal{A}' \in \mathbf{Y}'_n$ such that $\mathcal{A}' \models p'(\bar{a})$.
From parts (2) and (3) of Assumption 7.2 it easily follows that (for large enough n)
$\mathbb{P}'_n(\mathbf{Z}'_n)/\mathbb{P}'_n(\mathbf{Y}'_n)$ differs from $\mathbb{P}'_n(\mathbf{Z}'_n)$ by at most ε^*,
$\mathbb{P}'_n(\mathbf{Z}'_n)$ differs from $\mathbb{P}'_n(\mathbf{X}'_n)$ by at most ε^* and
$\mathbb{P}'_n(\mathbf{X}'_n)$ differs from $\mathsf{P}'(p'(\bar{x}))$ by at most ε^*.
Consequently,

$$\mathsf{P}'(p'(\bar{x})) - 3\varepsilon^* \leq \frac{\mathbb{P}'_n(\mathbf{Z}'_n)}{\mathbb{P}'_n(\mathbf{Y}'_n)} \leq \mathsf{P}'(p'(\bar{x})) + 3\varepsilon^*. \tag{7.4}$$

By Lemma 7.5, $\mathbb{P}_n(\mathbf{X}_n \mid \mathbf{W}^{\mathbf{Y}'_n}) = \mathbb{P}^{\mathbf{Y}'_n}(\mathbf{X} \cap \mathbf{W}^{\mathbf{Y}'_n})$. Then, using (7.2), we have

$$\mathbb{P}^{\mathbf{Y}'_n}\big(\mathbf{X}_n \cap \mathbf{W}^{\mathbf{Y}'_n}\big) = \sum_{\mathcal{A}' \in \mathbf{Y}'_n} \mathbb{P}^{\mathbf{Y}'_n}\big(\mathbf{X}_n \cap \mathbf{W}^{\mathcal{A}'}\big) = \sum_{\mathcal{A}' \in \mathbf{Z}'_n} \mathbb{P}^{\mathbf{Y}'_n}\big(\mathbf{X}_n \cap \mathbf{W}^{\mathcal{A}'}\big) =$$

$$\sum_{\mathcal{A}' \in \mathbf{Z}'_n} \sum_{\mathcal{A} \in \mathbf{X}_n \cap \mathbf{W}^{\mathcal{A}'}} \mathbb{P}^{\mathbf{Y}'_n}(\mathcal{A}) = \sum_{\mathcal{A}' \in \mathbf{Z}'_n} \sum_{\mathcal{A} \in \mathbf{X}_n \cap \mathbf{W}^{\mathcal{A}'}} \frac{\mathbb{P}'_n(\mathcal{A}')}{\mathbb{P}'_n(\mathbf{Y}'_n)} \mathbb{P}^{\mathcal{A}'}(\mathcal{A}) =$$

$$\sum_{\mathcal{A}' \in \mathbf{Z}'_n} \frac{\mathbb{P}'_n(\mathcal{A}')}{\mathbb{P}'_n(\mathbf{Y}'_n)} \sum_{\mathcal{A} \in \mathbf{X}_n \cap \mathbf{W}^{\mathcal{A}'}} \mathbb{P}^{\mathcal{A}'}(\mathcal{A}) = \sum_{\mathcal{A}' \in \mathbf{Z}'_n} \frac{\mathbb{P}'_n(\mathcal{A}')}{\mathbb{P}'_n(\mathbf{Y}'_n)} \mathbb{P}^{\mathcal{A}'}\big(\mathbf{X}_n \cap \mathbf{W}^{\mathcal{A}'}\big).$$

By Lemma 7.7 and (7.4),

$$\sum_{\mathcal{A}' \in \mathbf{Z}'_n} \frac{\mathbb{P}'_n(\mathcal{A}')}{\mathbb{P}'_n(\mathbf{Y}'_n)} \mathbb{P}^{\mathcal{A}'}\big(\mathbf{X}_n \cap \mathbf{W}^{\mathcal{A}'}\big) \leq \sum_{\mathcal{A}' \in \mathbf{Z}'_n} \frac{\mathbb{P}'_n(\mathcal{A}')}{\mathbb{P}'_n(\mathbf{Y}'_n)} \big(\mathsf{P}(p(\bar{x}) \mid p'(\bar{x})) + \varepsilon^*\big) =$$

$$\frac{\mathbb{P}'_n(\mathbf{Z}'_n)}{\mathbb{P}'_n(\mathbf{Y}'_n)} \big(\mathsf{P}(p(\bar{x}) \mid p'(\bar{x})) + \varepsilon^*\big) \leq \big(\mathsf{P}'(p'(\bar{x})) + 3\varepsilon^*\big)\big(\mathsf{P}(p(\bar{x}) \mid p'(\bar{x})) + \varepsilon^*\big) \leq$$

$$\mathsf{P}'(p'(\bar{x})) \cdot \mathsf{P}(p(\bar{x}) \mid p'(\bar{x})) + 7\varepsilon^*$$

and in a similar way

$$\sum_{\mathcal{A}' \in \mathbf{Z}'_n} \frac{\mathbb{P}'_n(\mathcal{A}')}{\mathbb{P}'_n(\mathbf{Y}'_n)} \mathbb{P}^{\mathcal{A}'}\big(\mathbf{X}_n \cap \mathbf{W}^{\mathcal{A}'}\big) \geq \mathsf{P}'(p'(\bar{x})) \cdot \mathsf{P}(p(\bar{x}) \mid p'(\bar{x})) - 7\varepsilon^*.$$ ☐

Lemma 7.10. *Suppose that $p'(\bar{x})$ is a complete atomic σ'-type and that $p(\bar{x}) \supseteq p'(\bar{x})$ is an (possibly partial) atomic σ-type. Then for all sufficiently large n and all $\bar{a} \in [n]^{|\bar{x}|}$ which realize the identity fragment of $p'(\bar{x})$ we have*

$$\left|\mathbb{P}_n\big(\{\mathcal{A} \in \mathbf{W}_n : \mathcal{A} \models p(\bar{a})\}\big) \; - \; \mathsf{P}(p(\bar{x}) \mid p'(\bar{x})) \cdot \mathsf{P}'(p'(\bar{x}))\right| \; < \; 9\varepsilon^*.$$

Proof. Let $\bar{a} \in [n]^{|\bar{x}|}$ realize the identity fragment of $p'(\bar{x})$ and we can assume that n is large enough that $\delta'(n) \leq \varepsilon^*$. Let $\mathbf{X}_n$ be the set of all $\mathcal{A} \in \mathbf{W}_n$ such that $\mathcal{A} \models p(\bar{a})$. We have

$$\mathbb{P}_n\big(\mathbf{X}_n\big) = \mathbb{P}_n\big(\mathbf{X}_n \mid \mathbf{W}^{\mathbf{Y}'_n}\big)\mathbb{P}_n\big(\mathbf{W}^{\mathbf{Y}'_n}\big) + \mathbb{P}_n\big(\mathbf{X}_n \mid \mathbf{W}_n \setminus \mathbf{W}^{\mathbf{Y}'_n}\big)\mathbb{P}_n\big(\mathbf{W}_n \setminus \mathbf{W}^{\mathbf{Y}'_n}\big).$$

By the use of Lemma 7.4 and by Part (2) of Assumption 7.2, we also have

$$\mathbb{P}_n\big(\mathbf{W}_n \setminus \mathbf{W}^{\mathbf{Y}'_n}\big) = 1 - \mathbb{P}_n\big(\mathbf{W}^{\mathbf{Y}'_n}\big) = 1 - \mathbb{P}'_n(\mathbf{Y}'_n) \leq \delta'(n).$$

It follows that $\mathbb{P}_n\big(\mathbf{X}_n \mid \mathbf{W}_n \setminus \mathbf{W}^{\mathbf{Y}'_n}\big)\mathbb{P}_n\big(\mathbf{W}_n \setminus \mathbf{W}^{\mathbf{Y}'_n}\big) \leq \delta'(n)$. By Lemma 7.4 and Part (2) of Assumption 7.2, $\mathbb{P}_n\big(\mathbf{W}^{\mathbf{Y}'_n}\big) = \mathbb{P}'_n\big(\mathbf{Y}'_n\big) \geq 1 - \delta'(n)$. It now follows from Lemma 7.9 that $\mathbb{P}_n\big(\mathbf{X}_n\big)$ differs from $\mathsf{P}(p(\bar{x}) \mid p'(\bar{x})) \cdot \mathsf{P}'(p'(\bar{x}))$ by at most $7\varepsilon^* + 2\delta'(n) \leq 9\varepsilon^*$. □

Definition 7.11. For every (possibly partial) σ-type $p(\bar{x})$ such that $p'(\bar{x}) = p{\restriction}\sigma'$ is a complete atomic σ'-type, we define $\mathsf{P}(p(\bar{x})) = \mathsf{P}'(p'(\bar{x})) \cdot \mathsf{P}(p(\bar{x}) \mid p'(\bar{x}))$.

With this definition we can reformulate Lemma 7.10 as follows:

Corollary 7.12. *If $p(\bar{x})$ is an (possibly partial) atomic σ-type such that $p{\restriction}\sigma'$ is a complete atomic σ'-type, then, for all sufficiently large n and all $\bar{a} \in [n]^{|\bar{x}|}$ which realize the identity fragment of $p(\bar{x})$ we have*

$$\big|\mathbb{P}_n\big(\{\mathcal{A} \in \mathbf{W}_n : \mathcal{A} \models p(\bar{a})\}\big) \ - \ \mathsf{P}(p(\bar{x}))\big| \ < \ 9\varepsilon^*.$$

Lemma 7.13. *Suppose that $p(\bar{x}, y)$ and $q(\bar{x})$ are complete atomic σ-types such that $|\bar{x}y| \leq k$, $\dim_y(p) = 1$ and $q \subseteq p$. Then there are $\gamma \in [0,1]$ and $c > 0$ such that for all sufficiently large n and all $\mathcal{A}' \in \mathbf{Y}'_n$,*

$$\mathbb{P}^{\mathcal{A}'}\big(\{\mathcal{A} \in \mathbf{W}^{\mathcal{A}'} : \mathcal{A} \textit{ is } (p, q, (\gamma - 5\varepsilon^*))\textit{-saturated and } (p, q, (\gamma + 5\varepsilon^*))\textit{-unsaturated}\}\big) \ \geq \ 1 \ - \ e^{-cn}.$$

Proof. Let n be large enough that Part (4) of Assumption 7.2 holds. Suppose that $p(\bar{x}, y)$ and $q(\bar{x})$ are complete atomic σ-types such that $|\bar{x}y| \leq k$, $\dim_y(p) = 1$ and $q \subseteq p$. Let $p' = p{\restriction}\sigma$ and $q' = q{\restriction}\sigma'$. Moreover, let $p^y(\bar{x}, y)$ include $p'(\bar{x}, y)$ and all $(\sigma \setminus \sigma')$-formulas in $p(\bar{x}, y)$ which contain the variable y.

Let $\mathcal{A}' \in \mathbf{Y}'_n$. By Part (4) of Assumption 7.2 there is $\alpha \in [0,1]$ such that $\mathcal{A}'$ is $(p', q', \alpha - \varepsilon')$-saturated and $(p', q', \alpha + \varepsilon')$-unsaturated. By the same assumption, α does not depend on the particular choice of $\mathcal{A}' \in \mathbf{Y}'_n$. Let

$$\beta = \mathsf{P}(p^y(\bar{x}, y) \mid p'(\bar{x}, y)),$$

where $\mathsf{P}(p^y(\bar{x}, y) \mid p'(\bar{x}, y))$ is like in Lemma 7.7, and let

$$\gamma = \alpha\beta.$$

For every $\bar{a} \in [n]^{|\bar{x}|}$ let

$$B'_{\bar{a}} \ = \ \big\{b \in [n] : \mathcal{A}' \models p'(\bar{a}, b)\big\}.$$

Since $\mathcal{A}'$ is $(p', q', \alpha - \varepsilon')$-saturated and $(p', q', \alpha + \varepsilon')$-unsaturated we have

$$(\alpha - \varepsilon')n \leq |B'_{\bar{a}}| \leq (\alpha + \varepsilon')n.$$

For every $\bar{a} \in [n]^{|\bar{x}|}$ and every $\mathcal{A} \in \mathbf{W}^{\mathcal{A}'}$ let

$$B_{\bar{a},\mathcal{A}} \ = \ \big\{b \in [n] : \mathcal{A} \models p^y(\bar{a}, b)\big\}$$

and note that $B_{\bar{a},\mathcal{A}} \subseteq B'_{\bar{a}}$ for every $\bar{a}$ and every $\mathcal{A} \in \mathbf{W}^{\mathcal{A}'}$. It follows that if $\alpha = 0$ then the conclusion of the lemma follows with $\gamma = 0$ because $\varepsilon' \leq \varepsilon^*$. So for the rest of the proof we assume that $\alpha > 0$. By starting with a sufficiently small ε' we can assume that ε^* is small enough so that $\alpha, \beta > \varepsilon^*$. Let

$$\mathbf{X}_{\bar{a}} \;=\; \big\{\mathcal{A} \in \mathbf{W}^{\mathcal{A}'} : \text{ either } \mathcal{A} \not\models q(\bar{a}) \text{ or } \\ (1-\varepsilon^*)(\gamma - 2\varepsilon^*)n \leq |B_{\bar{a},\mathcal{A}}| \leq (1+\varepsilon^*)(\gamma + 2\varepsilon^*)n\big\}.$$

Observe that if $\mathcal{A} \in \mathbf{W}^{\mathcal{A}'}$, $\mathcal{A} \models q(\bar{a})$ and $\mathcal{A} \models p^y(\bar{a}, b)$, then $\mathcal{A} \models p(\bar{a}, b)$. Hence every $\mathcal{A} \in \bigcap_{\bar{a}\in[n]^{|\bar{x}|}} \mathbf{X}_{\bar{a}}$ is $(p, q, (1-\varepsilon^*)(\gamma-2\varepsilon^*))$-saturated and $(p, q, (1+\varepsilon^*)(\gamma+2\varepsilon^*))$-unsaturated.

Fix any $\bar{a}$ such that $\mathcal{A}' \models q'(\bar{a})$ (and note that $\mathcal{A} \models q(\bar{a})$ implies $\mathcal{A}' \models q'(\bar{a})$). By Remark 7.6, for all distinct $b, c \in B'_{\bar{a}}$, the events

$$\mathbf{E}_b = \{\mathcal{A} \in \mathbf{W}^{\mathcal{A}'} : \mathcal{A} \models p^y(\bar{a}, b)\} \;\text{ and }\; \mathbf{E}_c = \{\mathcal{A} \in \mathbf{W}^{\mathcal{A}'} : \mathcal{A} \models p^y(\bar{a}, c)\}$$

are independent. Moreover, by the choice of β and the definition of ε^* (Definition 7.8), for each $b \in B'_{\bar{a}}$,

$$\beta - \varepsilon^* \leq \mathbb{P}_n^{\mathcal{A}'}(\mathbf{E}_b) \leq \beta + \varepsilon^*.$$

Let $Z : \mathbf{W}^{\mathcal{A}'} \to \mathbb{N}$ be the random variable defined by

$$Z(\mathcal{A}) = \big|\{b \in B'_{\bar{a}} : \mathcal{A} \models p^y(\bar{a}, b)\}\big|.$$

It follows from Corollary 2.2 that

$$\mathbb{P}_n^{\mathcal{A}'}\big(Z > (1+\varepsilon^*)(\beta+\varepsilon^*)|B'_{\bar{a}}|\big) \;<\; 2\exp\big(-c_{\varepsilon^*}(\beta+\varepsilon^*)|B'_{\bar{a}}|\big)$$

and

$$\mathbb{P}_n^{\mathcal{A}'}\big(Z < (1-\varepsilon^*)(\beta-\varepsilon^*)|B'_{\bar{a}}|\big) \;<\; 2\exp\big(-c_{\varepsilon^*}(\beta-\varepsilon^*)|B'_{\bar{a}}|\big)$$

where the constant $c_{\varepsilon^*} > 0$ depends only on ε^*. By using that $\alpha n \geq |B'_{\bar{a}}| - \varepsilon' n \geq |B'_{\bar{a}}| - \varepsilon^* n$ we see that

$$(\gamma + 2\varepsilon^*)n \geq (\alpha\beta + 2\varepsilon^*)n \geq \beta\alpha n + 2\varepsilon^* n \geq \beta(|B'_{\bar{a}}| - \varepsilon^* n) + 2\varepsilon^* n \geq \\ \beta|B'_{\bar{a}}| - \beta\varepsilon^* n + 2\varepsilon^* n \geq \beta|B'_{\bar{a}}| + \varepsilon^* n \geq \beta|B'_{\bar{a}}| + \varepsilon^*|B'_{\bar{a}}| = (\beta+\varepsilon^*)|B'_{\bar{a}}|$$

and by similar reasoning we get

$$(\gamma - 2\varepsilon^*)n \leq (\beta - \varepsilon^*)|B'_{\bar{A}}|.$$

It follows that if $Z > (1+\varepsilon^*)(\gamma + 2\varepsilon^*)n$ then $Z > (1+\varepsilon^*)(\beta+\varepsilon^*)|B'_{\bar{a}}|$, and if $Z < (1-\varepsilon^*)(\gamma - 2\varepsilon^*)n$ then $Z < (1-\varepsilon^*)(\beta-\varepsilon^*)|B'_{\bar{a}}|$. Hence we have

$$\mathbb{P}^{\mathcal{A}'}\big(\mathbf{W}^{\mathcal{A}'} \setminus \mathbf{X}_{\bar{a}}\big) \;<\; 2\exp\big(-c_{\varepsilon^*}(\beta+\varepsilon^*)|B'_{\bar{a}}|\big) + 2\exp\big(-c_{\varepsilon^*}(\beta-\varepsilon^*)|B'_{\bar{a}}|\big) \;\leq \\ 2\exp\big(-c_{\varepsilon^*}(\beta+\varepsilon^*)(\alpha-\varepsilon^*)n\big) + 2\exp\big(-c_{\varepsilon^*}(\beta-\varepsilon^*)(\alpha-\varepsilon^*)n\big) \;\leq\; e^{-dn}$$

for some constant $d > 0$ that depends only on ε^*, p and q. Since the argument works for all $\bar{a} \in [n]^{|\bar{x}|}$ it follows that

$$\mathbb{P}^{\mathcal{A}'}\Bigg(\bigcap_{\bar{a}\in[n]^{|\bar{x}|}} \mathbf{X}_{\bar{a}}\Bigg) \;\geq\; 1 - n^{|\bar{x}|}e^{-dn} \geq 1 - e^{-cn}$$

for some constant $c > 0$. Since

$$(1 \pm \varepsilon^*)(\gamma \pm 2\varepsilon^*) = \gamma \pm 2\varepsilon^* \pm \varepsilon^*\gamma \pm 2(\varepsilon^*)^2$$

and (if $0 < \varepsilon^* < 1$) $|2\varepsilon^* \pm \varepsilon^*\gamma \pm 2(\varepsilon^*)^2| \leq 5\varepsilon^*$ we get the conclusion of the lemma. □

The next lemma generalizes the previous one to types $p(\bar{x}, \bar{y})$ where the length of $\bar{y}$ is greater than one.

Lemma 7.14. *Suppose that $p(\bar{x}, \bar{y})$ and $q(\bar{x})$ are complete atomic σ-types such that $|\bar{x}\bar{y}| \leq k$, $\dim_{\bar{y}}(p) = |\bar{y}|$ and $q \subseteq p$. Then there are $\gamma \in [0, 1]$ and $c > 0$ such that for all large enough n and all $\mathcal{A}' \in \mathbf{Y}'_n$,*

$$\mathbb{P}^{\mathcal{A}'}\big(\{\mathcal{A} \in \mathbf{W}^{\mathcal{A}'} : \mathcal{A} \textit{ is } (p, q, \alpha - 11^{|\bar{y}|-1} \cdot 5\varepsilon^*)\textit{-saturated} \\ \textit{and } (p, q, \alpha + 11^{|\bar{y}|+1} \cdot 5\varepsilon^*)\textit{-unsaturated}\}\big) \ \geq \ 1 - e^{-cn}.$$

Proof. We prove the lemma by induction on $m = |\bar{y}|$. The base case $m = 1$ is given by Lemma 7.13 (since then $11^{m-1} \cdot 5\varepsilon^* = 5\varepsilon^*$). Let $p(\bar{x}, \bar{y})$ and $q(\bar{x})$ be as assumed in the lemma where $\bar{y} = (y_1, \ldots, y_{m+1})$. Let $p_m(\bar{x}, y_1, \ldots, y_m)$ be the restriction of p to formulas with variables among $\bar{x}, y_1, \ldots, y_m$. By induction hypothesis, there are $\alpha \in [0, 1]$ and $a > 0$ such that for all sufficiently large n and all $\mathcal{A}' \in \mathbf{Y}'_n$,

$$\mathbb{P}^{\mathcal{A}'}\big(\{\mathcal{A} \in \mathbf{W}^{\mathcal{A}'} : \mathcal{A} \text{ is } (p_m, q, \alpha - 11^{m-1} \cdot 5\varepsilon^*)\text{-saturated} \\ \text{and } (p_m, q, \alpha + 11^{m-1} \cdot 5\varepsilon^*)\text{-unsaturated}\}\big) \ \geq \ 1 - e^{-an}.$$

By Lemma 7.13, there are $\beta \in [0, 1]$ and $b > 0$ such that for all sufficiently large n and all $\mathcal{A}' \in \mathbf{Y}'_n$,

$$\mathbb{P}^{\mathcal{A}'}\big(\{\mathcal{A} \in \mathbf{W}^{\mathcal{A}'} : \mathcal{A} \text{ is } (p, p_m, (\beta - 5\varepsilon^*)\text{-saturated} \\ \text{and } (p, p_m, (\beta + 5\varepsilon^*)\text{-unsaturated}\}\big) \ \geq \ 1 \ - \ e^{-bn}.$$

Suppose that $\mathcal{A} \in \mathbf{W}^{\mathcal{A}'}$ is $(p_m, q, \alpha - 11^{m-1} \cdot 5\varepsilon^*)$-saturated and $(p, p_m, \beta - 5\varepsilon^*)$-saturated. Then one straightforwardly finds that $\mathcal{A}$ is

$$(p, q, (\alpha - 11^{m-1} \cdot 5\varepsilon^*)(\beta - 5\varepsilon^*))\text{-saturated}$$

and by calculations we get the following quite crude estimate

$$(\alpha - 11^{m-1} \cdot 5\varepsilon^*)(\beta - 5\varepsilon^*) \geq \alpha\beta - 11^m \cdot 5\varepsilon^*.$$

Hence $\mathcal{A}$ is $(p, q, \alpha\beta - 11^m \cdot 5\varepsilon^*)$-saturated. In a similar way it follows that if $\mathcal{A} \in \mathbf{W}^{\mathcal{A}'}$ is $(p_m, q, \alpha + 11^{m-1} \cdot 5\varepsilon^*)$-unsaturated and $(p, p_m, \beta + 5\varepsilon^*)$-unsaturated, then $\mathcal{A}$ is $(p, q, \alpha\beta + 11^m \cdot 5\varepsilon^*)$-unsaturated. Let $\gamma = \alpha\beta$. Since there is $c > 0$ such that $e^{-an} + e^{-bn} \leq e^{-cn}$ for all large enough n we get the desired estimate of the probability in the statement of the lemma. □

Corollary 7.15. *Let $p(\bar{x}, \bar{y})$ and $q(\bar{x})$ are complete atomic σ-types such that $|\bar{x}\bar{y}| \leq k$, $d = \dim_{\bar{y}}(p) > 0$, $q \subseteq p$. Then there are $\gamma \in [0, 1]$, depending only on p, q and $\mathbb{G}$, and $c > 0$ such that for all sufficiently large n and all $\mathcal{A}' \in \mathbf{Y}'_n$,*

$$\mathbb{P}^{\mathcal{A}'}\big(\{\mathcal{A} \in \mathbf{W}^{\mathcal{A}'} : \mathcal{A} \textit{ is } (p, q, \alpha - 11^{d-1} \cdot 5\varepsilon^*)\textit{-saturated} \\ \textit{and } (p, q, \alpha + 11^{d+1} \cdot 5\varepsilon^*)\textit{-unsaturated}\}\big) \ \geq \ 1 - e^{-cn}.$$

Proof. Suppose that $p(\bar{x}, \bar{y})$, where $\bar{y} = (y_1, \ldots, y_m)$ is a complete σ-type and let $q(\bar{x}) = p{\restriction}\bar{x}$. Let $d = \dim_{\bar{y}}(p)$. Then there is a subsequence $y_{i_1}, \ldots, y_{i_d}$ (of distinct variables) such that if $j \in [m] \setminus \{i_1, \ldots, i_d\}$ then there is $i \in \{i_1, \ldots, i_d\}$ such that the formula $y_j = y_i$ belongs to p. By reordering the sequence $\bar{y}$ we can assume that $i_1 = 1, \ldots, i_d = d$ (and hence, if $d < j \leq m$ then $y_j = y_i$ belongs to p for some $i \leq d$). Let $p^*(\bar{x}, y_1, \ldots, y_d)$ be the set of all formulas $\varphi \in p(\bar{x}, \bar{y})$ such that all variables in φ belong to $\text{rng}(\bar{x}) \cup \{y_1, \ldots, y_d\}$. It follows that $\dim_{(y_1,\ldots,y_d)}(p^*) = d$ and for every σ-structure $\mathcal{A}$ and every $\bar{a} \in A^{|\bar{x}|}$ then number of $\bar{b} \in A^{|\bar{y}|}$

such that $\mathcal{A} \models p(\bar{a}, \bar{b})$ equals the number of $\bar{b} \in A^d$ such that $\mathcal{A} \models p^*(\bar{a}, \bar{b})$. Lemma 7.14 implies that there are $\gamma \in [0,1]$ and $c > 0$ such that for all large enough n,

$$\mathbb{P}^{\mathcal{A}'}\big(\{\mathcal{A} \in \mathbf{W}^{\mathcal{A}'} : \mathcal{A} \text{ is } (p^*, q, \gamma - 11^{d-1} \cdot 5\varepsilon^*)\text{-saturated} \\ \text{and } (p^*, q, \gamma + 11^{d+1} \cdot 5\varepsilon^*)\text{-unsaturated}\}\big) \;\geq\; 1 - e^{-cn}.$$

By the construction of p^* the above lower bound on the probability also holds if p^* is replaced by p. □

Definition 7.16. For every n, let $\mathbf{Y}_n$ be the set of all $\mathcal{A} \in \mathbf{W}^{\mathbf{Y}'_n}$ such that whenever $p(\bar{x}, \bar{y})$ and $q(\bar{x})$ are complete atomic σ-types with $|\bar{x}\bar{y}| \leq k$, $0 < \dim_{\bar{y}}(p) = d$, $q \subseteq p$ and γ is the number associated to p and q as in Corollary 7.15, then $\mathcal{A}$ is $(p, q, \gamma - 11^{|\bar{y}|-1} \cdot 5\varepsilon^*)$-saturated and $(p, q, \gamma + 11^{|\bar{y}|-1} \cdot 5\varepsilon^*)$-unsaturated.

Lemma 7.17. *There is a constant $c > 0$ such that for all sufficiently large n, $\mathbb{P}_n\big(\mathbf{Y}_n\big) \geq \big(1 - e^{-cn}\big)\big(1 - \delta'(n)\big)$.*

Proof. There are, up to changing variables, only finitely many atomic σ-types $p(\bar{x})$ such that $|\bar{x}| \leq k$. It follows from Corollary 7.15 that there is a constant $c > 0$ such that for all large enough n and all $\mathcal{A}' \in \mathbf{Y}'_n$,

$$\mathbb{P}_n^{\mathcal{A}'}\big(\mathbf{Y}_n \cap \mathbf{W}^{\mathcal{A}'}\big) \geq 1 - e^{-cn}.$$

Since $\mathbf{Y}_n \subseteq \mathbf{W}^{\mathbf{Y}'_n}$ we have $\mathbb{P}_n(\mathbf{Y}_n) = \mathbb{P}_n\big(\mathbf{Y}_n \mid \mathbf{W}^{\mathbf{Y}'_n}\big)\mathbb{P}_n\big(\mathbf{W}^{\mathbf{Y}'_n}\big)$. By Lemma 7.4, $\mathbb{P}_n(\mathbf{W}^{\mathbf{Y}'_n}) = \mathbb{P}'_n(\mathbf{Y}'_n)$ and by Lemma 7.5 we have $\mathbb{P}_n(\mathbf{Y}_n \mid \mathbf{W}^{\mathbf{Y}'_n}) = \mathbb{P}^{\mathbf{Y}'_n}(\mathbf{Y}_n \cap \mathbf{W}^{\mathbf{Y}'_n})$. Hence $\mathbb{P}_n(\mathbf{Y}_n) = \mathbb{P}^{\mathbf{Y}'_n}(\mathbf{Y}_n \cap \mathbf{W}^{\mathbf{Y}'_n})\mathbb{P}'_n\big(\mathbf{Y}'_n\big)$. Then, reasoning similarly as in the proof of Lemma 7.9 (using (7.2)), we get

$$\mathbb{P}^{\mathbf{Y}'_n}\big(\mathbf{Y}_n \cap \mathbf{W}^{\mathbf{Y}'_n}\big) \;=\; \sum_{\mathcal{A}' \in \mathbf{Y}'_n} \mathbb{P}^{\mathbf{Y}'_n}\big(\mathbf{Y}_n \cap \mathbf{W}^{\mathcal{A}'}\big) \;=\; \sum_{\mathcal{A}' \in \mathbf{Y}'_n} \sum_{\mathcal{A} \in \mathbf{Y}_n \cap \mathbf{W}^{\mathcal{A}'}} \mathbb{P}^{\mathbf{Y}'_n}(\mathcal{A}) =$$

$$\sum_{\mathcal{A}' \in \mathbf{Y}'_n} \sum_{\mathcal{A} \in \mathbf{Y}_n \cap \mathbf{W}^{\mathcal{A}'}} \frac{\mathbb{P}'_n(\mathcal{A}')}{\mathbb{P}'_n(\mathbf{Y}'_n)} \mathbb{P}^{\mathcal{A}'}(\mathcal{A}) \;=\; \sum_{\mathcal{A}' \in \mathbf{Y}'_n} \frac{\mathbb{P}'_n(\mathcal{A}')}{\mathbb{P}'_n(\mathbf{Y}'_n)} \sum_{\mathcal{A} \in \mathbf{Y}_n \cap \mathbf{W}^{\mathcal{A}'}} \mathbb{P}^{\mathcal{A}'}(\mathcal{A}) =$$

$$\sum_{\mathcal{A}' \in \mathbf{Y}'_n} \frac{\mathbb{P}'_n(\mathcal{A}')}{\mathbb{P}'_n(\mathbf{Y}'_n)} \mathbb{P}^{\mathcal{A}'}\big(\mathbf{Y}_n \cap \mathbf{W}^{\mathcal{A}'}\big) \;\geq\; \sum_{\mathcal{A}' \in \mathbf{Y}'_n} \frac{\mathbb{P}'_n(\mathcal{A}')}{\mathbb{P}'_n(\mathbf{Y}'_n)} \big(1 - e^{-cn}\big) = \big(1 - e^{-cn}\big).$$

Using Part (2) of Assumption 7.2 we get

$$\mathbb{P}_n(\mathbf{Y}_n) \;=\; \mathbb{P}^{\mathbf{Y}'_n}\big(\mathbf{Y}_n \cap \mathbf{W}^{\mathbf{Y}'_n}\big)\mathbb{P}'_n(\mathbf{Y}'_n) \;\geq\; \big(1 - e^{-cn}\big)\big(1 - \delta'(n)\big).$$ □

Proposition 7.18. *Let $\varepsilon = 11^{k-1} \cdot 5\varepsilon^*$. Then there is a function $\delta : \mathbb{N} \to \mathbb{R}^{\geq 0}$ such that if σ', ε', δ', $\mathbf{W}'_n$, $\mathbf{Y}'_n$ and $\mathbb{P}'_n$ are replaced by σ, ε, δ, $\mathbf{W}_n$, $\mathbf{Y}_n$ and $\mathbb{P}_n$, respectively, then parts (1) – (4) of Assumption 7.2 hold.*

Proof. From Lemma 7.17 it follows that there is a function $\delta : \mathbb{N} \to \mathbb{R}^{\geq 0}$ such that $\lim_{n\to\infty} \delta(n) = 0$ and $\mathbb{P}_n(\mathbf{Y}_n) \geq 1 - \delta(n)$ for all sufficiently large n. Hence parts (1) and (2) of Assumption 7.2 hold if $\mathbb{P}'_n$, $\mathbf{Y}'_n$ and δ' are replaced by $\mathbb{P}_n$, $\mathbf{Y}_n$ and δ, respectively.

Let $\varepsilon = 11^{k-1} \cdot 5\varepsilon^*$. By Corollary 7.12, Part (3) of Assumption 7.2 holds if σ', $\mathbb{P}'_n$, $\mathbf{W}'_n$ and ε' are replaced by σ, $\mathbb{P}_n$, $\mathbf{W}_n$ and ε, respectively. By Corollary 7.15 and Definition 7.16, Part (4) of Assumption 7.2 holds if σ', $\mathbf{Y}'_n$ and ε' are replaced by σ, $\mathbf{Y}_n$ and ε, respectively. □

Corollary 7.19. *For all $k \in \mathbb{N}^+$ and all $\varepsilon > 0$ there are $\delta : \mathbb{N}^+ \to \mathbb{R}^{\geq 0}$ and $\mathbf{Y}_n \subseteq \mathbf{W}_n$, for $n \in \mathbb{N}^+$, such that the following hold:*

(1) $\lim_{n\to\infty} \delta(n) = 0$.
(2) $\mathbb{P}_n(\mathbf{Y}_n) \geq 1 - \delta(n)$ *for all sufficiently large* n.
(3) *For every complete atomic* σ*-type* $p(\bar{x})$ *with* $|\bar{x}| \leq k$ *there is a number which we denote* $\mathsf{P}(p(\bar{x}))$*, or just* $\mathsf{P}(p)$*, such that for all sufficiently large* n *and all* $\bar{a} \in [n]^{|\bar{x}|}$ *which realize the identity fragment of* p*,*

$$\left|\mathbb{P}_n\big(\{\mathcal{A} \in \mathbf{W}_n : \mathcal{A} \models p(\bar{a})\}\big) \ - \ \mathsf{P}(p(\bar{x}))\right| \ \leq \ \varepsilon.$$

(4) *For every complete atomic* σ*-type* $p(\bar{x}, \bar{y})$ *with* $|\bar{x}\bar{y}| \leq k$ *and* $0 < \dim_{\bar{y}}(p(\bar{x}, y)) = d$*, if* $q(\bar{x}) = p{\restriction}\bar{x}$*, then there is* $\alpha \in [0,1]$ *such that, for all sufficiently large* n*, every* $\mathcal{A} \in \mathbf{Y}_n$ *is* $(p, q, \alpha - \varepsilon)$*-saturated and* $(p, q, \alpha + \varepsilon)$*-unsaturated.*

Proof. This follows from Proposition 7.18 because $k \in \mathbb{N}^+$ and $\varepsilon' > 0$ in the argument until Proposition 7.18 are arbitrary and (for any fixed k) the choice of ε in Proposition 7.18 tends to zero as ε' tends to zero. □

Corollary 7.20. *If all aggregation functions in* $\varphi(\bar{x}) \in PLA^+(\sigma)$ *are strongly admissible, then* $\varphi(\bar{x})$ *is asymptotically equivalent (with respect to* $\mathbb{G}$*) to a basic probability formula.*

Proof. According to Corollary 7.19, for every $k \in \mathbb{N}^+$ and every $\varepsilon > 0$ there are $\delta : \mathbb{N}^+ \to \mathbb{R}^{\geq 0}$ and $\mathbf{Y}_n \subseteq \mathbf{W}_n$, for $n \in \mathbb{N}^+$, such that (1), (2) and (4) of that corollary hold. This means that Assumption 6.3 holds. Hence Proposition 6.4 implies that if all aggregation functions in $\varphi(\bar{x}) \in PLA^+(\sigma)$ are strongly admissible, then $\varphi(\bar{x})$ is asymptotically equivalent to a basic probability formula. □

With corollaries 7.19 and 7.20 we have completed the proof of Theorem 5.11. Informally speaking, the next corollary tells that for every $PLA^+(\sigma)$-network $\mathbb{G}$ that uses only strongly admissible aggregation functions there is a $PLA^+(\sigma)$-network $\tilde{\mathbb{G}}$ which uses no aggregation functions at all and is such that every query defined by a $PLA^+(\sigma)$-formula with only strongly admissible aggregation functions has the same asymtotic probability whether we compute it with $\mathbb{G}$ or with $\mathbb{G}'$.

Corollary 7.21 (aggregation-free networks)**.** *Let* σ *be a finite relational signature and let* $\mathbb{G}$ *be a* $PLA^+(\sigma)$*-network such that every probability formula of* $\mathbb{G}$ *contains only strongly admissible aggregation functions. Let* $(\mathbb{P}_n : n \in \mathbb{N}^+)$ *be the sequence of probability distributions induced by* $\mathbb{G}$*.*

By Part (i) of Theorem 5.11, for every $R \in \sigma$*, the probability formula* θ_R *which is associated to* R *by* $\mathbb{G}$ *is asymptotically equivalent to a basic probability formula* χ_R *with respect to* $\mathbb{G}$*. Let* $\tilde{\mathbb{G}}$ *be the* $PLA^+(\sigma)$*-network with the same underlying directed acyclic graph as* $\mathbb{G}$ *and where, for each* $R \in \sigma$*,* χ_R *is the probability formula associated to* R *by* $\tilde{\mathbb{G}}$*. Let* $(\tilde{\mathbb{P}}_n : n \in \mathbb{N}^+)$ *be the sequence of probability distributions induced by* $\tilde{\mathbb{G}}$*. Then the following hold:*

(i) *For every atomic* σ*-type* $p(\bar{x})$*, every* $m \in \mathbb{N}^+$ *and every* $\bar{a} \in [m]^{|\bar{x}|}$*,*

$$\lim_{n\to\infty} \tilde{\mathbb{P}}_n(p(\bar{a})) = \lim_{n\to\infty} \mathbb{P}_n(p(\bar{a})).$$

The common limit depends only on the probability formulas of $\tilde{\mathbb{G}}$ *and the common directed acyclic graph of* $\mathbb{G}$ *and* $\tilde{\mathbb{G}}$*.*

(ii) *Let* $p(\bar{x}, \bar{y})$ *and* $q(\bar{x})$ *be complete atomic* σ*-types such that* $q \subseteq p$*. Then there is* $\alpha \in [0,1]$*, depending only on* p*,* q*, the common directed acyclic graph of* $\mathbb{G}$ *and* $\tilde{\mathbb{G}}$ *and the probability* *formulas of* $\tilde{\mathbb{G}}$ *such that*

$$\lim_{n\to\infty} \tilde{\mathbb{P}}_n\big(\{\mathcal{A} \in \mathbf{W}_n : \mathcal{A} \text{ is } (p,q,\alpha-\varepsilon)\text{-saturated and } (p,q,\alpha+\varepsilon)\text{-unsaturated}\}\big) =$$
$$\lim_{n\to\infty} \mathbb{P}_n\big(\{\mathcal{A} \in \mathbf{W}_n : \mathcal{A} \text{ is } (p,q,\alpha-\varepsilon)\text{-saturated and } (p,q,\alpha+\varepsilon)\text{-unsaturated}\}\big) = 1.$$

(iii) *Let* $\varphi(\bar{x}) \in PLA^+(\sigma)$ *be such that every aggregation function in* φ *is strongly admissible. Then there is a basic probability formula* $\psi(\bar{x})$ *such that* $\varphi(\bar{x})$ *and* $\psi(\bar{x})$ *are asymptotically equivalent with respect to* $\mathbb{G}$ *and with respect to* $\tilde{\mathbb{G}}$. *Without loss of generality we may assume that* $\psi(\bar{x})$ *has the form* $\bigwedge_{i=1}^k(\psi_i(\bar{x}) \to c_i)$ *where each* $\psi_i(\bar{x})$ *is the conjuction of all formulas in a complete atomic* σ*-type* $p_i(\bar{x})$ *and that all complete atomic* σ*-types in the variables* $\bar{x}$ *are enumerated without repetition by* $p_1, \dots, p_k$. *If* $I \subseteq [0,1]$ *is an interval such that no* c_i *is an endpoint of* I, *then for every* $m \in \mathbb{N}^+$ *and every* $\bar{a} \in [m]^{|\bar{x}|}$,

$$\lim_{n\to\infty} \big|\tilde{\mathbb{P}}_n\big(\{\mathcal{A} \in \mathbf{W}_n : \mathcal{A}(\psi(\bar{a})) \in I\}\big) - \mathbb{P}_n\big(\{\mathcal{A} \in \mathbf{W}_n : \mathcal{A}(\varphi(\bar{a})) \in I\}\big)\big| \;=\; 0,$$

where $\tilde{\mathbb{P}}_n\big(\{\mathcal{A} \in \mathbf{W}_n : \mathcal{A}(\psi(\bar{a}))\big)$ *can be computed by only using* $\tilde{\mathbb{G}}$ *and* ψ, *hence in constant time with respect to the domain size.*

Proof sketch. (i) It suffices to consider complete atomic σ-types. Let $p(\bar{x})$ be a complete atomic σ-type. An analysis of the proof leading to Corollary 7.19 shows that the limit $\lim_{n\to\infty} \mathbb{P}_n(p(\bar{a}))$ depends only on the basic probability formulas χ_R first mentioned in Part (5) of Assumption 7.2; the crucial step where χ_R is used is Lemma 7.7. If a probability formula θ_R, as in Part (5) of Assumption 7.2, is a basic probability formula then we can simply let χ_R be the same formula as θ_R to make sure that Part (5) of Assumption 7.2 holds. Let $\mathbb{G}$ and $\tilde{\mathbb{G}}$ be as assumed. By the construction of $\tilde{\mathbb{G}}$ it follows from what has been said that both limits $\lim_{n\to\infty} \mathbb{P}_n(p(\bar{a}))$ and $\lim_{n\to\infty} \tilde{\mathbb{P}}_n(p(\bar{a}))$ depend only on the probability formulas χ_R of $\tilde{\mathbb{G}}$, so the limits are equal.

(ii) If we analyze the proofs above, in particular the proof of Lemma 7.13, we see that the “saturation number” α depends only on limits of the form $\lim_{n\to\infty} \mathbb{P}_n(p(\bar{a}))$ where $p(\bar{x})$ is an atomic σ-type. By Part (i) we get the same limit if $\mathbb{P}_n$ is replaced by $\tilde{\mathbb{P}}_n$ and therefore the conclusion follows.

(iii) Part (ii) implies that the saturation conditions stated by Assumption 6.3 are satisfied by both $(\mathbb{P}_n : n \in \mathbb{N}^+)$ and $(\tilde{\mathbb{P}}_n : n \in \mathbb{N}^+)$. Therefore Proposition 6.4 applies to both $(\mathbb{P}_n : n \in \mathbb{N}^+)$ and $(\tilde{\mathbb{P}}_n : n \in \mathbb{N}^+)$. So if $\varphi(\bar{x}) \in PLA^+(\sigma)$ has only strongly admissible aggregation functions then there are basic probability formulas $\psi(\bar{x})$ and $\tilde{\psi}(\bar{x})$ such that $\varphi(\bar{x}) \sim_{\mathbb{G}} \psi(\bar{x})$ and $\varphi(\bar{x}) \sim_{\tilde{\mathbb{G}}} \tilde{\psi}(\bar{x})$. An analysis of the proof of Proposition 6.4 shows that the constructions of ψ and $\tilde{\psi}$ depend only on the “saturation numbers” α mentioned in Part (ii) of this corollary. Since these saturation numbers are the same (as stated in Part (ii)) for $(\mathbb{P}_n : n \in \mathbb{N}^+)$ and $(\tilde{\mathbb{P}}_n : n \in \mathbb{N}^+)$ it follows that the constructed ψ and $\tilde{\psi}$ are the same. So $\varphi(\bar{x}) \sim_{\mathbb{G}} \psi(\bar{x})$ and $\varphi(\bar{x}) \sim_{\tilde{\mathbb{G}}} \psi(\bar{x})$.

Suppose, without loss of generality, that $\psi(\bar{x})$ is $\bigwedge_{i=1}^k(\psi_i(\bar{x}) \to c_i)$ where each $\psi_i(\bar{x})$ is a conjunction of the formulas in a complete atomic σ-type $p_i(\bar{x})$ and that all complete atomic σ-types are enumerated without repetition by $p_1, \dots, p_k$. By reordering if necessary we may assume that $c_1, \dots, c_l \in I$ and $c_i \notin I$ if $i > l$. Then

$$\mathbb{P}_n\big(\{\mathcal{A} \in \mathbf{W}_n : \mathcal{A}(\psi(\bar{a})) \in I\}\big) = \mathbb{P}_n\bigg(\bigvee_{i=1}^{l} \psi_i(\bar{a})\bigg) = \sum_{i=1}^{l} \mathbb{P}_n\big(\psi_i(\bar{a})\big)$$

and by Part (i) the same holds if '$\mathbb{P}_n$' is replaced by '$\tilde{\mathbb{P}}_n$'. Since $\varphi(\bar{x}) \sim_{\mathbb{G}} \psi(\bar{x})$ we get

$$\lim_{n\to\infty} \left|\tilde{\mathbb{P}}_n\big(\{\mathcal{A} \in \mathbf{W}_n : \mathcal{A}(\psi(\bar{a})) \in I\}\big) - \mathbb{P}_n\big(\{\mathcal{A} \in \mathbf{W}_n : \mathcal{A}(\varphi(\bar{a})) \in I\}\big)\right| \;=\; 0. \qquad \square$$

8. Almost sure elimination of safe CPL-formulas when the distribution is induced by a $coPLA^+$-network

In this section we prove a quantifier elimination result, Theorem 8.6, for the class of "safe" *CPL*-formulas with respect to a sequence of probability distributions $\mathbb{P}$ induced by a $coPLA^+(\sigma)$-network $\mathbb{G}$, that is a $PLA^+(\sigma)$-network that (in its probability formulas) uses only strongly admissible, or continuous, aggregation functions. Unlike the lifted Bayesian networks considered in Section 5.3, such $\mathbb{G}$ can induce $\mathbb{P}$ for which with high likelihood some (or all) relations in σ are sparse. For example, as explained in Example 5.10, if σ contains a binary relation R then, for any fixed $\alpha \in (0, 1)$, a $coPLA^+(\sigma)$-network can express that the probability that $R(x, y)$ holds is $n^{-\alpha}$ (independently of other edges) where n is the cardinality of the domain. Therefore $coPLA^+(\sigma)$-networks can induce the probability distributions studied by Shelah and Spencer in [SS88] in relation to first-order logic.

CPL (see Definition 4.16) is a natural query language since it can express queries about relative frequencies. For example it can express the condition that the relative frequency of $\varphi_1(\bar{x})$ among tuples $\bar{x}$ that satisfy $\varphi_2(\bar{x})$ is at least r, or alternatively, at least as large as the relative frequency of $\psi_1(\bar{x})$ among tuples $\bar{x}$ that satisfy $\psi_2(\bar{x})$.

Theorem 8.6 below provides a quantifier elimination result for such queries. For reasons that become clear below, its scope is restricted to formulas $\varphi(\bar{x}) \in CPL(\sigma)$ which are "safe" with respect to $\mathbb{G}$. We begin with a discussion which motivates the definition of "safe" CPL-formulas further down. As usual let σ be a finite relational signature and $\mathbf{W}_n$ the set of all σ-structures with domain $[n]$. Let $\mathbb{G}$ be a $coPLA^+(\sigma)$-network and let $(\mathbb{P}_n : n \in \mathbb{N}^+)$ be the sequence of probability distributions induced by $\mathbb{G}$. Let $p(\bar{x}, \bar{y})$ and $q(\bar{x})$ be complete atomic σ-types such that $q \subseteq p$. By Corollary 7.19 (4), there is $\alpha \in [0, 1]$, depending only on p, q and $\mathbb{G}$, such that for every $\varepsilon > 0$,

$$\lim_{n\to\infty} \mathbb{P}_n\big(\{\mathcal{A} \in \mathbf{W}_n : \mathcal{A} \text{ is } (p, q, \alpha - \varepsilon)\text{-saturated and } (p, q, \alpha + \varepsilon)\text{-unsaturated}\}\big) \;=\; 1.$$

We can call α the *scaled saturation number* of the pair (p, q), where "scaled" refers to the fact that the $\bar{y}$-dimension of $p(\bar{x}, \bar{y})$ is taken into account in the definition of (un)saturation.

We can extend the idea of scaled saturation numbers to pairs $(\varphi(\bar{x}, \bar{y}), q(\bar{x}))$ where $\varphi(\bar{x}, \bar{y}) \in CPL(\sigma)$ is quantifier-free and $q(\bar{x})$ is a complete atomic σ-type such that $\forall \bar{x}, \bar{y}\big(\varphi(\bar{x}, \bar{y}) \to q(\bar{x})\big)$ is true in every finite σ-structure. Then there are complete atomic σ-types $p_i(\bar{x}, \bar{y})$, $i = 1, \ldots, s$, such that $q \subseteq p_i$ for all i and $\varphi(\bar{x}, \bar{y})$ is equivalent to $\bigvee_{i=1}^{s} p_i(\bar{x}, \bar{y})$ (where p_i is identified with the conjunction of formulas in p_i). Let $d = \max\{\dim_{\bar{y}}(p_i) : i = 1, \ldots, s\}$. Without loss of generality, suppose that $p_1, \ldots, p_t$, $t \le s$, enumerates, without repetition, all p_i such that $\dim_{\bar{y}}(p_i) = d$. Let α_i be the scaled saturation number of (p_i, q) for each i. If $d < |\bar{y}|$ then let $\alpha = 0$. Otherwise (i.e. if $d = |\bar{y}|$) let $\alpha = \alpha_1 + \ldots + \alpha_t$. Call α the *saturation number* of the pair (φ, q). It now follows that if $\varepsilon > 0$ and

$$\mathbf{X}_n^{\varepsilon} = \big\{\mathcal{A} \in \mathbf{W}_n : \text{ if } \bar{a} \in q(\mathcal{A}) \text{ then } (\alpha - \varepsilon)n^{|\bar{y}|} \le |\varphi(\bar{a}, \mathcal{A})| \le (\alpha + \varepsilon)n^{|\bar{y}|}\big\}$$

then $\lim_{n\to\infty} \mathbb{P}_n(\mathbf{X}_n^{\varepsilon}) = 1$.

However, if $\psi(\bar{x}, \bar{y}), \theta(\bar{x}, \bar{y}) \in CPL(\sigma)$ then we can, in general, *not* guarantee that there is $\alpha \in [0,1]$ such that for all $\bar{a} \in [n]^{|\bar{x}|}$ and $\varepsilon > 0$, the following probability converges to 1 as $n \to \infty$:

$$\mathbb{P}_n\big(\{\mathcal{A} \in \mathbf{W}_n : \text{ if } \theta(\bar{a}, \mathcal{A}) \neq \emptyset, \text{ then } \big||\psi(\bar{a}, \mathcal{A}) \cap \theta(\bar{a}, \mathcal{A})|/|\theta(\bar{a}, \mathcal{A})| - \alpha\big| < \varepsilon\}\big).$$

Therefore we cannot in general be sure that $\lim_{n\to\infty} \mathbb{P}_n(\varphi(\bar{a}))$ exists if $\varphi(\bar{x})$ has the form

$$\Big(\|\psi_1(\bar{x}, \bar{y}) \mid \theta_1(\bar{x}, \bar{y})\|_{\bar{y}} \ \geq \ \|\psi_2(\bar{x}, \bar{y}) \mid \theta_2(\bar{x}, \bar{y})\|_{\bar{y}} \ + \ r\Big).$$

This motivates considering CPL-formulas in which the value of terms of the form $\|\psi(\bar{x}, \bar{y}) \mid \theta(\bar{x}, \bar{y})\|_{\bar{y}}$ almost surely converges. This is made precise by the next two definitions.

Definition 8.1. Let $\psi(\bar{x}, \bar{y}), \theta(\bar{x}, \bar{y}) \in CPL(\sigma)$.
(i) We say that $\theta(\bar{x}, \bar{y})$ is *atomically $\bar{x}$-complete* if there is a complete atomic σ-type $q(\bar{x})$ such that $\forall \bar{x}, \bar{y}(\theta(\bar{x}, \bar{y}) \to q(\bar{x}))$ is true in every finite σ-structure.
(ii) We call $\|\psi(\bar{x}, \bar{y}) \mid \theta(\bar{x}, \bar{y})\|_{\bar{y}}$ a *conditional probability term* of $CPL(\sigma)$.
(iii) Suppose that $\theta(\bar{x}, \bar{y})$ is atomically $\bar{x}$-complete and that, for some complete atomic σ-type $q(\bar{x})$, $\forall \bar{x}, \bar{y}(\theta(\bar{x}, \bar{y}) \to q(\bar{x}))$ is true in every finite σ-structure. Let $\alpha \in [0,1]$. The *saturation number of the conditional probability term* $\|\psi(\bar{x}, \bar{y}) \mid \theta(\bar{x}, \bar{y})\|_{\bar{y}}$ *exists and is* α if, for every $\varepsilon > 0$,

$$\lim_{n\to\infty} \mathbb{P}_n\Bigg(\Bigg\{\mathcal{A} \in \mathbf{W}_n : \text{there is } \bar{a} \in [n]^{|\bar{x}|} \text{ such that } \mathcal{A} \models q(\bar{a}) \text{ and } \left|\alpha - \frac{|\psi(\bar{a}, \mathcal{A}) \cap \theta(\bar{a}, \mathcal{A})|}{|\theta(\bar{a}, \mathcal{A})|}\right| > \varepsilon\Bigg\}\Bigg) \ = \ 0.$$

Definition 8.2 (Safe formula)**.** A formula $\varphi(\bar{x}) \in CPL(\sigma)$ is *safe with respect to* $\mathbb{G}$ if the following hold:

(1) $\varphi(\bar{x})$ does not contain $\forall$ or $\exists$.
(2) For every subformula of $\varphi(\bar{x})$ of the form

$$\Big(\|\psi_1(\bar{y}, \bar{z}) \mid \theta_1(\bar{y}, \bar{z})\|_{\bar{z}} \ \geq \ \|\psi_2(\bar{y}, \bar{z}) \mid \theta_2(\bar{y}, \bar{z})\|_{\bar{z}} \ + \ r\Big)$$

or

$$\Big(r \ + \ \|\psi_1(\bar{y}, \bar{z}) \mid \theta_1(\bar{y}, \bar{z})\|_{\bar{z}} \ \geq \ \|\psi_2(\bar{y}, \bar{z}) \mid \theta_2(\bar{y}, \bar{z})\|_{\bar{z}}\Big)$$

and every complete atomic σ-type $q(\bar{y})$ such that $\exists \bar{y}, \bar{z}, \bar{z}'(\theta_1(\bar{y}, \bar{z}) \wedge \theta_2(\bar{y}, \bar{z}') \wedge q(\bar{y}))$ has a finite model, the saturation number of $\|\psi_i(\bar{y}, \bar{z}) \mid \theta_i(\bar{y}, \bar{z}) \wedge q(\bar{y}))\|_{\bar{y}}$ exists for $i = 1, 2$ and if it is denoted by α_i, then $r \neq |\alpha_1 - \alpha_2|$.

Observe that if a formula is safe then every subformula of it is also safe.

Remark 8.3 (Why first-order quantifiers are omitted in safe formulas)**.** As mentioned in Example 3.7 the aggregation function $length_\alpha$, where $\alpha \in (0,1)$, is strongly admissible, or in other words, continuous. Let $\sigma = \{R\}$ where R is binary and let $\mathbf{W}_n$ be the set of all σ-structures with domain $[n]$. We define a $coPLA^+(\sigma)$-network $\mathbb{G}$ by letting the probability formula $\theta_R(x, y)$ of R be $length_\alpha(\psi(x, y, z) : z)$ where $\psi(x, y, z)$ is the formula $z = z$. Let $(\mathbb{P}_n : n \in \mathbb{N}^+)$ be the sequence of probability distributions induced by $\mathbb{G}$. Then for every n, every $\mathcal{A} \in \mathbf{W}_n$ and all $a, b \in [n]$, $\mathbb{P}_n(R(a, b)) = \mathcal{A}(\theta_R(a, b)) = 1/n^\alpha$, where $\mathcal{A}$ is the unique $\emptyset$-structure with domain $[n]$. By a seminal result of Shelah and Spencer [SS88, Theorem 2], if α is rational then there is a first-order sentence $\varphi \in FO(\sigma)$ such that $\lim_{n\to\infty} \mathbb{P}_n(\varphi)$ does not exist (and recall that $FO(\sigma) \subseteq CPL(\sigma)$). So if we want that "safeness" implies convergence we must omit the first-order quantifiers from safe formulas.

Strictly speaking, the result of Shelah and Spencer referred to is about undirected graphs. But each $\mathcal{A} \in \mathbf{W}_n$ gives rise to an undirected graph by, for different $a, b \in [n]$ considering $\{a, b\}$ as an undirected edge of the undirected graph induced by $\mathcal{A}$ if $\mathcal{A} \models a \neq b \wedge R(a,b) \wedge R(b,a)$. In the sentence φ constructed by Shelah and Spencer one then changes every subformula like $R(x,y)$ to '$x \neq y \wedge R(x,y) \wedge R(y,x)$'. Then this modified sentence φ' will have the same probability with respect to $\mathbb{P}_n$ as φ has with respect to the probability distribution on undirected graphs with vertex set $[n]$ which gives every edge probability $1/n^{2\alpha}$ independently of the existence of other edges (and 2α is rational if α is).

The following technical lemma will be used in the proof of Theorem 8.6 in combination with Lemma 8.5 below.

Lemma 8.4. *Let $\psi_i(\bar{x},\bar{y}), \theta_i(\bar{x},\bar{y})) \in CPL(\sigma)$ for $i = 1, 2$. Let $q_1(\bar{x}), \ldots, q_m(\bar{x})$ enumerate all complete atomic σ-types $q(\bar{x})$ such that $\exists \bar{x}, \bar{y}, \bar{z}\big(\theta_1(\bar{x},\bar{y}) \wedge \theta_2(\bar{x},\bar{z}) \wedge q(\bar{x})\big)$ has a finite model. For every finite σ-structure $\mathcal{A}$ and every $\bar{a} \in [n]^{|\bar{x}|}$,*

$$\mathcal{A} \models \Big(\|\psi_1(\bar{a},\bar{y}) \mid \theta_1(\bar{a},\bar{y})\|_{\bar{y}} \ \geq \ \|\psi_2(\bar{a},\bar{y}) \mid \theta_2(\bar{a},\bar{y})\|_{\bar{y}} \ + \ r\Big) \tag{8.1}$$

if and only if

$$\mathcal{A} \models \bigvee_{i=1}^{m} \Big(\|\psi_1(\bar{a},\bar{y}) \mid \theta_1(\bar{a},\bar{y}) \wedge q_i(\bar{a})\|_{\bar{y}} \ \geq \ \|\psi_2(\bar{a},\bar{y}) \mid \theta_2(\bar{a},\bar{y}) \wedge q_i(\bar{a})\|_{\bar{y}} \ + \ r\Big) \tag{8.2}$$

and similarly if '$+\ r$' is moved to the left hand side of '$\geq$'.

Proof. Let $\psi_i(\bar{x},\bar{y}), \theta_i(\bar{x},\bar{y})) \in CPL(\sigma)$ for $i = 1, 2$ and let $q_1(\bar{x}), \ldots, q_m(\bar{x})$ enumerate all complete atomic σ-types $q(\bar{x})$ such that $\exists \bar{x}, \bar{y}, \bar{z}\big(\theta_1(\bar{x},\bar{y}) \wedge \theta_2(\bar{x},\bar{y}) \wedge q(\bar{x})\big)$ has a finite model. Let $\mathcal{A}$ be a finite σ-structure and let $\bar{a} \in A^{|\bar{x}|}$.

First suppose that (8.1) holds. By the semantics of CPL, $\theta_1(\bar{a}, \mathcal{A}) \neq \emptyset$ and $\theta_2(\bar{a}, \mathcal{A}) \neq \emptyset$, so $\mathcal{A} \models \exists \bar{y}, \bar{z}(\theta_1(\bar{a},\bar{y}) \wedge \theta_2(\bar{a},\bar{z}))$ and therefore there is k such that $\mathcal{A} \models q_k(\bar{a})$. By (8.1) and the semantics of CPL we also have

$$\frac{|\psi_1(\bar{a},\mathcal{A}) \cap \theta_1(\bar{a},\mathcal{A})|}{|\theta_1(\bar{a},\mathcal{A})|} \ \geq \ \frac{|\psi_2(\bar{a},\mathcal{A}) \cap \theta_2(\bar{a},\mathcal{A})|}{|\theta_2(\bar{a},\mathcal{A})|} + r.$$

As $\mathcal{A} \models q_k(\bar{a})$ the above implies that

$$\frac{|\psi_1(\bar{a},\mathcal{A}) \cap (\theta_1 \wedge q_k)(\bar{a},\mathcal{A})|}{|(\theta_1 \wedge q_k)(\bar{a},\mathcal{A})|} \ \geq \ \frac{|\psi_2(\bar{a},\mathcal{A}) \cap (\theta_2 \wedge q_k)(\bar{a},\mathcal{A})|}{|(\theta_2 \wedge q_k)(\bar{a},\mathcal{A})|} + r.$$

and hence (by the semantics of CPL) we get

$$\mathcal{A} \ \models \ \Big(\|\psi_1(\bar{a},\bar{y}) \mid \theta_1(\bar{a},\bar{y}) \wedge q_k(\bar{a})\|_{\bar{y}} \ \geq \ \|\psi_2(\bar{a},\bar{y}) \mid \theta_2(\bar{a},\bar{y}) \wedge q_k(\bar{a})\|_{\bar{y}} \ + \ r\Big) \tag{8.3}$$

which implies (8.2).

Now suppose that (8.2) holds. Then there is k such that (8.3) holds. Hence

$$\mathcal{A} \models \exists \bar{y}, \bar{z}((\theta_1 \wedge q_k)(\bar{a},\bar{y}) \wedge (\theta_2 \wedge q_k)(\bar{a},\bar{z}))$$

and

$$\frac{|\psi_1(\bar{a},\mathcal{A}) \cap (\theta_1 \wedge q_k)(\bar{a},\mathcal{A})|}{|(\theta_1 \wedge q_k)(\bar{a},\mathcal{A})|} \ \geq \ \frac{|\psi_2(\bar{a},\mathcal{A}) \cap (\theta_2 \wedge q_k)(\bar{a},\mathcal{A})|}{|(\theta_2 \wedge q_k)(\bar{a},\mathcal{A})|} + r.$$

Since $(\theta_i \wedge q_k)(\bar{a}, \mathcal{A}) = \theta_i(\bar{a}, \mathcal{A})$, for $i = 1, 2$, it follows that

$$\frac{|\psi_1(\bar{a}, \mathcal{A}) \cap \theta_1(\bar{a}, \mathcal{A})|}{|\theta_1(\bar{a}, \mathcal{A})|} \ \geq\ \frac{|\psi_2(\bar{a}, \mathcal{A}) \cap \theta_2(\bar{a}, \mathcal{A})|}{|\theta_2(\bar{a}, \mathcal{A})|} + r.$$

so (8.1) holds.

The case when '$+\ r$' is to the left of '$\geq$' is proved similarly. □

Lemma 8.5. *Suppose that $\varphi(\bar{x})$ has the form*

$$\Big(\|\psi_1(\bar{x}, \bar{y}) \mid \theta_1(\bar{x}, \bar{y})\|_{\bar{y}} \ \geq\ \|\psi_2(\bar{x}, \bar{y}) \mid \theta_2(\bar{x}, \bar{y})\|_{\bar{y}} \ +\ r\Big)$$

Also suppose that, for some complete atomic σ-type $q(\bar{x})$, the sentences $\forall \bar{x}, \bar{y}(\theta_1(\bar{x}, \bar{y}) \to q(\bar{x}))$ and $\forall \bar{x}, \bar{y}(\theta_2(\bar{x}, \bar{y}) \to q(\bar{x}))$ hold in all finite σ-structures. Furthermore, suppose that for $i = 1, 2$ the saturation number of $\|\psi_i(\bar{x}, \bar{y}) \mid \theta_i(\bar{x}, \bar{y})\|_{\bar{y}}$ exists and is α_i. Then:

(1) *If $\alpha_1 < \alpha_2 + r$ then $\varphi(\bar{x})$ is almost surely equivalent to $\perp$.*
(2) *If $\alpha_1 > \alpha_2 + r$ then $\varphi(\bar{x})$ is almost surely equivalent to $q(\bar{x})$.*

If $\varphi(\bar{x})$ is as above except that '$+r$' is moved to the left of '$\geq$', then $\varphi(\bar{x})$ is almost surely equivalent to $\perp$ if $r + \alpha_1 < \alpha_2$, and $\varphi(\bar{x})$ is almost surely equivalent to $q(\bar{x})$ if $r + \alpha_1 > \alpha_2$.

Proof. The numbers α_1 and α_2 depend only on $\theta_1, \theta_2, \psi_1, \psi_2, q$ and $\mathbb{G}$. Thus we can make a case distinction. First suppose that

$$\alpha_1 < \alpha_2 + r.$$

Let $\delta > 0$ be such that if α_1' and α_2' are within distance δ from α_1 and α_2, respectively, then $\alpha_1' < \alpha_2' + r$. Let $\mathbf{X}_n^\delta$ be the set of all $\mathcal{A} \in \mathbf{W}_n$ such that for every $\bar{a} \in [n]^{|\bar{x}|}$, if $\mathcal{A} \models q(\bar{a})$, then, for $i = 1, 2$,

$$\left|\alpha_i - \frac{|\psi_i(\bar{a}, \mathcal{A}) \cap \theta_i(\bar{a}, \mathcal{A})|}{|\theta_i(\bar{a}, \mathcal{A})|}\right| < \delta. \tag{8.4}$$

Then $\lim_{n\to\infty} \mathbb{P}_n(\mathbf{X}_n^\delta) = 1$.

It now suffices to show that for all $\mathcal{A} \in \mathbf{X}_n^\delta$ and all $\bar{a} \in [n]^{|\bar{x}|}$, $\mathcal{A} \not\models \varphi(\bar{a})$, because it then follows that $\varphi(\bar{x})$ is almost surely equivalent to $\perp$. So let $\mathcal{A} \in \mathbf{X}_n^\delta$ and $\bar{a} \in [n]^{|\bar{x}|}$.

If $\mathcal{A} \not\models q(\bar{a})$ then $\theta_1(\bar{a}, \mathcal{A}) = \emptyset$, so by the semantics of CPL, $\mathcal{A} \not\models \varphi(\bar{a})$. Now suppose that $\mathcal{A} \models q(\bar{a})$. By the choice of δ, the definition of $\mathbf{X}_n^\delta$ and (8.4) we get

$$\frac{|\theta_1(\bar{a}, \mathcal{A}) \cap \psi_1(\bar{a}, \mathcal{A})|}{|\theta_1(\bar{a}, \mathcal{A})|} \ <\ \frac{|\theta_2(\bar{a}, \mathcal{A}) \cap \psi_2(\bar{a}, \mathcal{A})|}{|\theta_2(\bar{a}, \mathcal{A})|} \ +\ r.$$

Hence $\mathcal{A} \not\models \varphi(\bar{a})$ according to the semantics of CPL. Now suppose that

$$\alpha_1 > \alpha_2 + r.$$

Let $\delta > 0$ be such that if α_1' and α_2' are within distance δ from α_1 and α_2, respectively, then $\alpha_1' > \alpha_2' + r$. Let $\mathbf{X}_n^\delta$ be defined in the same way as in the previous case, so in particular $\lim_{n\to\infty} \mathbb{P}_n(\mathbf{X}_n^\delta) = 1$. We will show that $\varphi(\bar{x})$ is almost surely equivalent to $q(\bar{x})$. It suffices that prove that for all $\mathcal{A} \in \mathbf{X}_n^\delta$ and $\bar{a} \in [n]^{|\bar{x}|}$, $\mathcal{A} \models \varphi(\bar{a})$ if and only if $\mathcal{A} \models q(\bar{x})$. Let $\mathcal{A} \in \mathbf{X}_n^\delta$ and $\bar{a} \in [n]^{|\bar{x}|}$.

If $\mathcal{A} \models \varphi(\bar{a})$ then, by the semantics of CPL, $\theta_1(\bar{a}, \mathcal{A}) \neq \emptyset$ which (by the assumptions of the lemma) implies that $\mathcal{A} \models q(\bar{a})$. Now suppose that $\mathcal{A} \models q(\bar{a})$. By the choice of δ and definition of $\mathbf{X}_n^\delta$ it follows that, for $i = 1, 2$,

$$\frac{|\theta_1(\bar{a}, \mathcal{A}) \cap \psi_1(\bar{a}, \mathcal{A})|}{|\theta_1(\bar{a}, \mathcal{A})|} \ >\ \frac{|\theta_2(\bar{a}, \mathcal{A}) \cap \psi_2(\bar{a}, \mathcal{A})|}{|\theta_2(\bar{a}, \mathcal{A})|} \ +\ r.$$

Hence $\mathcal{A} \models \varphi(\bar{a})$ according to the semantics of CPL. The last statement, concerning the variant of $\varphi(\bar{x})$ where '$+r$' is to the left of '$\leq$', is proved by straightforward modifications of the given arguments. □

Theorem 8.6. *Let $\mathbb{G}$ be a coPLA$^+(\sigma)$-network. If $\varphi(\bar{x}) \in CPL(\sigma)$ is safe with respect to $\mathbb{G}$, then $\varphi(\bar{x})$ is almost surely equivalent to a quantifier-free formula.*

Proof. Let $\varphi(\bar{x}) \in CPL(\sigma)$ be safe. We use induction on the quantifier rank of formulas. If φ is quantifier-free we are done. So suppose that $\varphi(\bar{x})$ is not quantifier-free and that every safe formula of lower quantifier rank than φ is almost surely equivalent to a quantifier-free formula. If $\varphi(\bar{x})$ has any of the forms $\neg\psi(\bar{x})$, $\psi(\bar{x}) \wedge \theta(\bar{x})$, $\psi(\bar{x}) \vee \theta(\bar{x})$, $\psi(\bar{x}) \to \theta(\bar{x})$ or $\psi(\bar{x}) \leftrightarrow \theta(\bar{x})$ and both $\psi(\bar{x})$ and $\theta(\bar{x})$ are almost surely equivalent to quantifier-free formulas, then it clearly follows that $\varphi(\bar{x})$ is almost surely equivalent to a quantifier-free formula.

Suppose that $\varphi(\bar{x})$ has the form

$$\Big(\|\psi_1(\bar{x},\bar{y}) \mid \theta_1(\bar{x},\bar{y})\|_{\bar{y}} \ \geq \ \|\psi_2(\bar{x},\bar{y}) \mid \theta_2(\bar{x},\bar{y})\|_{\bar{y}} \ + \ r\Big).$$

Let $q_1(\bar{x}), \ldots, q_m(\bar{x})$ enumerate all complete atomic σ-types $q(\bar{x})$ such that

$$\exists \bar{x}, \bar{y}, \bar{z}\big(\theta_1(\bar{x},\bar{y}) \wedge \theta_2(\bar{x},\bar{z}) \wedge q(\bar{x})\big)$$

has a finite model. By Lemma 8.4, $\varphi(\bar{x})$ is equivalent, in all finite σ-structures, to

$$\varphi'(\bar{x}) \ := \ \bigvee_{k=1}^{m} \Big(\|\psi_1(\bar{x},\bar{y}) \mid \theta_1(\bar{x},\bar{y}) \wedge q_k(\bar{x})\|_{\bar{y}} \ \geq \ \|\psi_2(\bar{x},\bar{y}) \mid \theta_2(\bar{x},\bar{y}) \wedge q_k(\bar{x})\|_{\bar{y}} \ + \ r\Big).$$

Hence it suffices to prove that, for every $k = 1, \ldots, m$,

$$\varphi_k(\bar{x}) \ := \ \Big(\|\psi_1(\bar{x},\bar{y}) \mid \theta_1(\bar{x},\bar{y}) \wedge q_k(\bar{x})\|_{\bar{y}} \ \geq \ \|\psi_2(\bar{x},\bar{y}) \mid \theta_2(\bar{x},\bar{y}) \wedge q_k(\bar{x})\|_{\bar{y}} \ + \ r\Big).$$

is almost surely equivalent to a quantifier-free formula. Since $\varphi(\bar{x})$ is assumed to be safe it follows that, for $i = 1, 2$, the saturation number of $\|\psi_1(\bar{x},\bar{y}) \mid \theta_1(\bar{x},\bar{y}) \wedge q_k(\bar{x})\|_{\bar{y}}$ exists and if it is denoted by α_i, then $\alpha_1 \neq \alpha_2 + r$. Now it follows from Lemma 8.5 that $\varphi_k(\bar{x})$ is almost surely equivalent to a quantifier-free formula. The case when $\varphi(\bar{x})$ has the form

$$\Big(r \ + \ \|\psi_1(\bar{x},\bar{y}) \mid \theta_1(\bar{x},\bar{y})\|_{\bar{y}} \ \geq \ \|\psi_2(\bar{x},\bar{y}) \mid \theta_2(\bar{x},\bar{y})\|_{\bar{y}}\Big)$$

is treated similarly. □

Corollary 8.7. *Let $\mathbb{G}$ be a coPLA$^+(\sigma)$-network. If $\varphi(\bar{x}) \in CPL(\sigma)$ is safe with respect to $\mathbb{G}$, then, for every $m \in \mathbb{N}^+$ and $\bar{a} \in [m]^{|\bar{x}|}$, $\lim_{n\to\infty} \mathbb{P}_n(\varphi(\bar{a}))$ exists.*

Proof. Let $\mathbb{G}$ be a *coPLA*$^+(\sigma)$-network and suppose that $\varphi(\bar{x}) \in CPL(\sigma)$ is safe with respect to $\mathbb{G}$. By Theorem 8.6 there is a quantifier-free $\varphi'(\bar{x}) \in CPL(\sigma)$ which is almost surely equivalent to $\varphi(\bar{x})$ with respect to $\mathbb{G}$. But then $\varphi'(\bar{x})$ is also an aggregation-free $PLA^+(\sigma)$-formula which only takes the values 0 or 1. Now Corollary 5.13 says that if $\bar{a} \in [m]^{|\bar{x}|}$ then $\lim_{n\to\infty} \mathbb{P}_n(\varphi'(\bar{a}))$ exists. As φ and φ' are almost surely equivalent also $\lim_{n\to\infty} \mathbb{P}_n(\varphi(\bar{a}))$ exists. □

Corollary 8.8. *Let σ be a finite relational signature and let $\mathbb{G}$ be a $PLA^+(\sigma)$-network such that every probability formula of $\mathbb{G}$ contains only strongly admissible aggregation functions. Let $\tilde{\mathbb{G}}$ be the aggregation-free network from Corollary 7.21 and let $(\tilde{\mathbb{P}}_n : n \in \mathbb{N}^+)$ be the sequence of probability distributions induced by $\tilde{\mathbb{G}}$. Let $\varphi(\bar{x})$ be an arbitrary* CPL*-formula. Then the following holds:*

(i) *$\varphi(\bar{x})$ is safe with respect to $\mathbb{G}$ if and only if it is safe with respect to $\tilde{\mathbb{G}}$.*
(ii) *If $\varphi(\bar{x})$ is safe with respect to $\mathbb{G}$, it is almost surely equivalent to a quantifier-free $\varphi'(\vec{x})$ with respect to $\mathbb{G}$ if and only if it is almost surely equivalent to $\varphi'(\bar{x})$ with respect to $\tilde{\mathbb{G}}$.*
(iii) *If $\varphi(\bar{x})$ is safe with respect to $\mathbb{G}$, then for every $m \in \mathbb{N}^+$ and $\bar{a} \in [m]^{|\bar{x}|}$,*

$$\lim_{n\to\infty} \mathbb{P}_n(\varphi(\bar{a})) = \lim_{n\to\infty} \tilde{\mathbb{P}}_n(\varphi(\bar{a})).$$

Proof. By Corollary 7.21(ii) saturation numbers are the same whether they are calculated with respect to $\mathbb{G}$ or $\tilde{\mathbb{G}}$. Since safety of a *CPL*-formula depends only on the saturation numbers, this immediately implies Point (i). Similarly, the formulas φ' and φ_k and their quantifier-free equivalents in the proof of Theorem 8.6 depend only on φ and the saturation numbers, which coincide for $\mathbb{G}$ and $\tilde{\mathbb{G}}$. Since the limit computed in Corollary 8.7 depends only on the quantifier-free formula, this in turn implies that the limits coincide, whether they are computed in $\mathbb{P}_n$ or in $\tilde{\mathbb{P}}_n$. □

9. Relative asymptotic expressivity of inference frameworks

In this section we tie together results of this article and in [Kop20, KW23, SS88] about sequences of probability distributions defined by different formalisms and queries defined by different logics. We do this by introducing the notion of "inference framework" and by comparing inference frameworks with the notion of "asymptotically at least as expressive". Informally, an inference framework is a set of pairs $(\mathbb{P}, L)$ where $\mathbb{P}$ is a sequence of probability distributions and L is a logic (which may depend on $\mathbb{P}$). In all concrete cases considered here, an inference framework $\mathbf{F}$ will be defined by a particular formalism (and usually with particular restrictions on it) for defining sequences of probability distributions and a logic $\mathcal{L}$ such that for each pair $(\mathbb{P}, L) \in \mathbf{F}$, L is a sublogic of $\mathcal{L}$ determined by $\mathbb{P}$ (although often $L = \mathcal{L}$). Theorem 9.7, illustrated by Figure 1, describes the relationships, with respect to the notion "asymptotically at least as expressive", between a number of inference frameworks which have implicitly been considered in this article and in [Kop20, KW23, SS88].

As usual we let σ be a finite relational signature and, for each $n \in \mathbb{N}^+$, $\mathbf{W}_n$ denotes the set of all σ-structures with domain $[n]$. However, for simplicity, in some notation (like $\mathbf{W}_n$) we do not explicitly show the dependence on σ. *We also assume that σ is nonempty*, which is of course the interesting case. (All previous results in this article hold also for empty σ.)

Definition 9.1. An *inference framework (for σ)* is a class $\mathbf{F}$ of pairs $(\mathbb{P}, L)$ where L is a logic (for σ) and $\mathbb{P} = (\mathbb{P}_n : n \in \mathbb{N}^+)$ where each $\mathbb{P}_n$ is a probability distribution on $\mathbf{W}_n$.

Definition 9.2. Let $\mathbf{F}$ and $\mathbf{F}'$ be inference frameworks for σ.
(i) We write $\mathbf{F} \preccurlyeq \mathbf{F}'$ and say that $\mathbf{F}'$ is *asymptotically at least as expressive as* $\mathbf{F}$ if for every $(\mathbb{P}, L) \in \mathbf{F}$ there is $(\mathbb{P}', L') \in \mathbf{F}'$ such that $\mathbb{P} \sim_{tv} \mathbb{P}'$ and for every $\varphi(\bar{x}) \in L$ there is $\varphi'(\bar{x}) \in L'$ such that $\varphi'(\bar{x})$ is asymptotically equivalent to $\varphi(\bar{x})$ with respect to $\mathbb{P}$.
(ii) By $\mathbf{F} \simeq \mathbf{F}'$ we mean that $\mathbf{F} \preccurlyeq \mathbf{F}'$ and $\mathbf{F}' \preccurlyeq \mathbf{F}$ and if this is the case we may say that $\mathbf{F}$ and $\mathbf{F}'$ are *asymptotically equally expressive.*
(iii) By $\mathbf{F} \prec \mathbf{F}'$ we mean that $\mathbf{F} \preccurlyeq \mathbf{F}'$ and $\mathbf{F}' \not\preccurlyeq \mathbf{F}$

As we discussed in Section 1 one can ask whether our notion of "asymptotically at least as expressive" is the most appropriate one. Especially, one can question why we require that $\mathbb{P}$ and $\mathbb{P}'$ in the definition above are total variation equivalent although the logics L and L' involved may not be able to "define" all subsets of $\mathbf{W}_n$. An important reason for our choice is that we think that "asymptotically at least as expressive" should be a transitive notion (as shown in Lemma 9.4 below) and all other candidates of such a notion that we have considered, for example by weakening the assumption about total variation equivalence, are not transitive.

Remark 9.3. According to the definition of '$\preccurlyeq$', if $\mathbf{F} \preccurlyeq \mathbf{F}'$, $(\mathbb{P}, L) \in \mathbf{F}$, and $\varphi(\bar{x}) \in L$, then there are $(\mathbb{P}', L') \in \mathbf{F}'$ and $\varphi'(\bar{x}) \in L'$ such that $\mathbb{P} \sim_{tv} \mathbb{P}'$ and $\varphi'(\bar{x})$ and $\varphi(\bar{x})$ are asymptotically equivalent *with respect to* $\mathbb{P}$. It follows straightforwardly from the definitions of '$\sim_{tv}$' and 'asymptotic equivalence' that $\varphi'(\bar{x})$ is also asymptotically equivalent to $\varphi(\bar{x})$ with respect to $\mathbb{P}'$, thus establishing a "symmetry" between $\mathbb{P}$ and $\mathbb{P}'$.

Lemma 9.4 (Transitivity of $\preccurlyeq$)**.** *Suppose that* $\mathbf{F}$*,* $\mathbf{F}'$ *and* $\mathbf{F}''$ *are inference frameworks. If* $\mathbf{F} \preccurlyeq \mathbf{F}'$ *and* $\mathbf{F}' \preccurlyeq \mathbf{F}''$*, then* $\mathbf{F} \preccurlyeq \mathbf{F}''$*.*

Proof. Suppose that $\mathbf{F} \preccurlyeq \mathbf{F}'$ and $\mathbf{F}' \preccurlyeq \mathbf{F}''$. Let $(\mathbb{P}, L) \in \mathbf{F}$ and let $\varphi(\bar{x}) \in L$. By assumption there are $(\mathbb{P}', L') \in \mathbf{F}'$ and $\varphi'(\bar{x}) \in L'$ such that $\mathbb{P} \sim_{tv} \mathbb{P}'$ and $\varphi'(\bar{x}) \in L'$ is asymptotically equivalent to $\varphi(\bar{x})$ with respect to $\mathbb{P}$. By Remark 9.3, $\varphi(\bar{x})$ and $\varphi'(\bar{x})$ are asymptotically equivalent with respect to $\mathbb{P}'$. By assumption there are $(\mathbb{P}'', L'') \in \mathbf{F}''$ and $\varphi'' \in L''$ such that $\mathbb{P}' \sim_{tv} \mathbb{P}''$ and $\varphi'(\bar{x}) \in L'$ is asymptotically equivalent to $\varphi'(\bar{x})$ with respect to $\mathbb{P}'$. From $\mathbb{P} \sim_{tv} \mathbb{P}' \sim_{tv} \mathbb{P}''$ it straightforwardly follows (from the defintion of $\sim_{tv}$) that $\mathbb{P} \sim_{tv} \mathbb{P}''$. Since $\varphi(\bar{x})$ and $\varphi'(\bar{x})$ are asymptotically equivalent with respect to $\mathbb{P}'$ and $\varphi'(\bar{x})$ and $\varphi''(\bar{x})$ are asymptotically equivalent with respect to $\mathbb{P}'$ it follows (straightforwardly from the definition of "asymptotic equivalence") that $\varphi(\bar{x})$ and $\varphi''(\bar{x})$ are asymptotically equivalent with respect to $\mathbb{P}'$. Since $\mathbb{P} \sim_{tv} \mathbb{P}'$ it follows that $\varphi(\bar{x})$ and $\varphi''(\bar{x})$ are asymptotically equivalent with respect to $\mathbb{P}$. This proves that $\mathbf{F} \preccurlyeq \mathbf{F}''$. □

The following lemma will be used in the proof of Theorem 9.7:

Lemma 9.5. *(i) There is a sequence of probability distributions* $\mathbb{P}$ *which is induced by a coPLA-network and such that for every sequence of probability distributions* $\mathbb{P}'$ *that is induced by a noncritial lifted Bayesian network* $\mathbb{P} \not\sim_{tv} \mathbb{P}'$*.*
(ii) There are a sentence $\varphi \in FO$*, a sentence* $\psi \in aPLA$ *and a sequence of probability distributions* $(\mathbb{P}_n : n \in \mathbb{N}^+)$ *induced by a coPLA-network such that* $\lim_{n\to\infty} \mathbb{P}_n(\varphi) = \lim_{n\to\infty} \mathbb{P}_n(\{\mathcal{A} \in \mathbf{W}_n : \mathcal{A}(\psi) = 1\})$ *exist and is neither 0 nor 1.*

Proof. Let $R \in \sigma$ have arity k. Let $\mathbb{G}$ be a $PLA(\sigma)$-network such that the underlying DAG has no edges at all and if $Q \in \sigma$ and $Q \neq R$, then θ_Q (the probability formula associated to Q) is '0', so the probability of $Q(\bar{b})$ (for $\bar{b}$ with length matching the arity of Q) is zero. Also let $F : [0,1]^{<\omega} \to [0,1]$ be defined by $F(\bar{r}) = 1/|\bar{r}|^k$ and let $\theta_R(\bar{x})$, where $\bar{x} = (x_1, \ldots, x_k)$ and $\bar{y} = (y_1, \ldots, y_k)$, be the formula

$$F(\bigwedge_{i=1}^{k}(x_i = x_i) \wedge \bigwedge_{i=1}^{k}(y_i = y_i) : \bar{y} : \bigwedge_{i\neq j}(x_i \neq x_j) \bigwedge_{i=1}^{k}\bigwedge_{j=1}^{k}(x_i \neq y_j) \wedge \bigwedge_{i\neq j}(y_i \neq y_j)).$$

It is straightforward to see that F is strongly admissible so $\mathbb{G}$ is a $coPLA$-network. Let $(\mathbb{P}_n : n \in \mathbb{N}^+)$ be the sequence of probability distributions induced by $\mathbb{G}$. By the definition of $\mathbb{P}_n$ it follows that if $\bar{a} \in [n]^k$ is a sequence of different elements, then

$$\mathbb{P}_n(R(\bar{a})) = \frac{1}{(n-k)(n-k-1)\ldots(n-2k+1)} \sim \frac{1}{n^k}.$$

Define the random variable $X_n : \mathbf{W}_n \to \mathbb{R}$ by letting $X_n(\mathcal{A})$ be the number of $\bar{a} \in [n]^k$ such that $\mathcal{A} \models R(\bar{a})$. Then X_n is the sum of $(n-k)(n-k-1)\ldots(n-2k+1)$ independent $\{0,1\}$-valued random variables where each one has probability

$$\frac{1}{(n-k)(n-k-1)\ldots(n-2k+1)}$$

of being 1. By a rough version of the Poisson approximation [Nov19, Equation (1)] we get

$$\left|\mathbb{P}_n(X_n \geq 1) - (1-e^{-1})\right| \;\leq\; C/n^k$$

where $C > 0$ is a constant. It follows that $\lim_{n\to\infty} \mathbb{P}_n(X_n \geq 1) = 1 - e^{-1}$. Let φ denote the first-order sentence $\exists \bar{x}\big(\bigwedge_{i\neq j}(x_i \neq x_j) \wedge R(\bar{x})\big)$ and let ψ denote the sentence

$$\max\big(R(\bar{x}) : \bar{x} : \bigwedge_{i\neq j}(x_i \neq x_j)\big)$$

which belongs to $aPLA$ (because max is admissible). Observe that for every finite structure $\mathcal{A}$, $\mathcal{A}(\psi)$ is either 0 or 1, and $\mathcal{A}(\psi) = 1$ if and only if $\mathcal{A} \models \varphi$. Now let

$$\mathbf{X}_n = \{\mathcal{A} \in \mathbf{W}_n : \mathcal{A} \models \varphi\}.$$

Then $\lim_{n\to\infty} \mathbb{P}_n(\mathbf{X}_n) = 1 - e^{-1}$ so Part (ii) is proved. Let $(\mathbb{P}'_n : n \in \mathbb{N}^+)$ be induced by a noncritical lifted Bayesian network. Since φ is noncritical (because every first-order formula is noncritical), Theorem 5.15 implies that $\mathbb{P}'_n(\mathbf{X}_n)$ converges to either 0 or 1. As $\lim_{n\to\infty} \mathbb{P}_n(\mathbf{X}_n) = 1 - e^{-1}$ we get $(\mathbb{P}_n : n \in \mathbb{N}^+) \not\succeq_{tv} (\mathbb{P}'_n : n \in \mathbb{N}^+)$ so Part (i) is proved. □

In the next definition we use notation for logics that was introduced in Definitions 4.3 and 4.14. Also recall the informal discussion about *(non)critical CPL-formulas* just before Theorem 5.15 and Definition 8.2 of safe CPL-formulas. If $\mathbb{G}$ is a lifted Bayesian network and every aggregation formula of $\mathbb{G}$ is noncritical with respect to $\mathbb{G}$ then we say that $\mathbb{G}$ is a *noncritical lifted Bayesian network*. If all aggregation formulas of $\mathbb{G}$ are quantifier-free then we say that $\mathbb{G}$ is a *quantifier-free lifted Bayesian network*.

Definition 9.6 (Concrete inference frameworks)**.** We define notation for some concrete inference frameworks via the table below which should be understood as follows. The first column gives the name of the inference framework and the second column describes the pairs $(\mathbb{P}, L)$ that belong to the inference framework named on the same row. For notational simplicity the (arbitrary nonempty finite relational) signature σ is suppressed in the notation.

name	contains all $(\mathbb{P}, L)$ such that $\mathbb{P}$ is induced by
$(\mathbf{qfLBN}, \mathbf{qfFO})$	a quantifier-free lifted Bayesian network and $L = qfFO$
$(\mathbf{ncLBN}, \mathbf{FO})$	a noncritical lifted Bayesian network and $L = FO$
$(\mathbf{ncLBN}, \mathbf{CPL})$	a noncritical lifted Bayesian network and $L = CPL$
$(\mathbf{ncLBN}, \mathbf{ncCPL})$	a noncritical lifted Bayesian network $\mathbb{G}$ and L is the set of all noncritical CPL-formulas with respect to $\mathbb{G}$
$(\mathbf{afPLAN}, \mathbf{qfFO})$	a *afPLA*-network and $L = qfFO$
$(\mathbf{ncLBN}, \mathbf{afPLA})$	a noncritical lifted Bayesian network and $L =$ *afPLA*
$(\mathbf{ncLBN}, \mathbf{aPLA})$	a noncritical lifted Bayesian network and $L = aPLA$
$(\mathbf{aPLAN}, \mathbf{aPLA})$	an $aPLA$-network and $L = aPLA$
$(\mathbf{aPLAN}^+, \mathbf{aPLA}^+)$	an $aPLA^+$-network and $L = aPLA^+$
$(\mathbf{coPLAN}^+, \mathbf{qfFO})$	a $coPLA^+$-network and $L = qfFO$
$(\mathbf{coPLAN}^+, \mathbf{FO})$	a $coPLA^+$-network and $L = FO$
$(\mathbf{coPLAN}^+, \mathbf{sCPL})$	a $coPLA^+$-network $\mathbb{G}$ and L is the set of all safe CPL-formulas with respect to $\mathbb{G}$
$(\mathbf{coPLAN}^+, \mathbf{afPLA})$	a $coPLA^+$-network and $L =$ *afPLA*
$(\mathbf{coPLAN}^+, \mathbf{coPLA}^+)$	a $coPLA^+$-network and $L = coPLA^+$

Theorem 9.7 (Relative asymptotic expressivity of inference frameworks)**.**

(1) $(\mathbf{afPLAN}, \mathbf{qfFO}) \simeq (\mathbf{qfLBN}, \mathbf{qfFO}) \simeq (\mathbf{ncLBN}, \mathbf{FO})$
$\simeq (\mathbf{ncLBN}, \mathbf{ncCPL}) \prec (\mathbf{ncLBN}, \mathbf{CPL})$.
(2) $(\mathbf{ncLBN}, \mathbf{ncCPL}) \prec (\mathbf{ncLBN}, \mathbf{afPLA}) \simeq (\mathbf{ncLBN}, \mathbf{aPLA})$.
(3) $(\mathbf{afPLAN}, \mathbf{afPLA}) \simeq (\mathbf{ncLBN}, \mathbf{afPLA}) \simeq (\mathbf{ncLBN}, \mathbf{aPLA})$
$\prec (\mathbf{aPLAN}, \mathbf{aPLA}) \preccurlyeq (\mathbf{aPLAN}^+, \mathbf{aPLA}^+)$.
(4) $(\mathbf{coPLAN}^+, \mathbf{FO}) \prec (\mathbf{aPLAN}^+, \mathbf{aPLA}^+)$.
(5) $(\mathbf{ncLBN}, \mathbf{ncCPL}) \prec (\mathbf{coPLAN}^+, \mathbf{qfFO}) \simeq (\mathbf{coPLAN}^+, \mathbf{sCPL})$
$\prec (\mathbf{coPLAN}^+, \mathbf{FO})$.
(6) $(\mathbf{coPLAN}^+, \mathbf{sCPL}) \prec (\mathbf{coPLAN}^+, \mathbf{afPLA}) \simeq (\mathbf{coPLAN}^+, \mathbf{coPLA}^+)$
$\prec (\mathbf{aPLAN}^+, \mathbf{aPLA}^+)$.
(7) $(\mathbf{afPLAN}, \mathbf{afPLA}) \prec (\mathbf{coPLAN}^+, \mathbf{coPLA}^+)$.
(8) $(\mathbf{ncLBN}, \mathbf{aPLA})$ *and* $(\mathbf{coPLAN}^+, \mathbf{FO})$ *are incomparable with respect to* $\preccurlyeq$.
$(\mathbf{coPLAN}^+, \mathbf{coPLA}^+)$ *and* $(\mathbf{coPLAN}^+, \mathbf{FO})$ *are incomparable with respect to* $\preccurlyeq$.
$(\mathbf{ncLBN}, \mathbf{aPLA})$ *and* $(\mathbf{coPLAN}^+, \mathbf{sCPL})$ *are incomparable with respect to* $\preccurlyeq$.

The main contents of this theorem are illustrated by Figure 1 (in the introduction).

Proof. (1) It follows easily from the definitions that for every *afPLA*-network there is a quantifier-free lifted Bayesian network which induces the same distributions, and vice versa. Therefore $(\mathbf{afPLAN}, \mathbf{qfFO}) \simeq (\mathbf{qfLBN}, \mathbf{qfFO})$. Since every first-order formula is noncritical with respect to every CPL-network we have $(\mathbf{ncLBN}, \mathbf{FO}) \preccurlyeq (\mathbf{ncLBN}, \mathbf{ncCPL})$. Hence, to show the two remaining statements about '$\simeq$' it suffices to show that $(\mathbf{ncLBN}, \mathbf{ncCPL}) \preccurlyeq (\mathbf{qfLBN}, \mathbf{qfFO})$. But this follows from theorem 5.15.

Clearly, $(\mathbf{ncLBN}, \mathbf{ncCPL}) \preccurlyeq (\mathbf{ncLBN}, \mathbf{CPL})$. In [Kop20, Remark 3.18] (which generalizes and example of Keisler and Lotfallah [KL09, Proposition 3.1]) it is shown that a lifted Bayesian network $\mathbb{G}$ and CPL-sentence φ (which is critical with respect to $\mathbb{G}$) exist such that if $(\mathbb{P}_n : n \in \mathbb{N}^+)$ is induced by $\mathbb{G}$, then $\mathbb{P}_n(\varphi)$ does not converge as $n \to \infty$. It follows from Theorem 5.15 that $(\mathbf{ncLBN}, \mathbf{ncCPL}) \prec (\mathbf{ncLBN}, \mathbf{CPL})$.

(2) We clearly have $qfFO \leq afPLA$ so $(\mathbf{ncLBN}, \mathbf{qfFO}) \preccurlyeq (\mathbf{ncLBN}, \mathbf{afPLA})$. Since for example '1/2' is an $afPLA$-formula such that $\mathcal{A}(1/2) = 1/2$ for every finite structure $\mathcal{A}$, it follows that $(\mathbf{ncLBN}, \mathbf{qfFO}) \prec (\mathbf{ncLBN}, \mathbf{afPLA})$. By Part (1) we now get $(\mathbf{ncLBN}, \mathbf{ncCPL}) \prec (\mathbf{ncLBN}, \mathbf{afPLA})$. By Theorem 5.17 we also get $(\mathbf{ncLBN}, \mathbf{afPLA}) \simeq (\mathbf{ncLBN}, \mathbf{aPLA})$.

(3) From Part (2) we have $(\mathbf{ncLBN}, \mathbf{afPLA}) \simeq (\mathbf{ncLBN}, \mathbf{aPLA})$. Theorem 5.15 (ii) states that if $\mathbb{G}$ is a noncritical lifted Bayesian network, then there is a quantifier-free lifted Bayesian network $\mathbb{G}'$ such that the sequences of distributions induced by $\mathbb{G}$ and $\mathbb{G}'$ are asymptotically total variation equivalent. Since for every quantifier-free lifted Bayesian network there is an $afPLA$-network which induces the same sequence of distributions we get $(\mathbf{ncLBN}, \mathbf{afPLA}) \simeq (\mathbf{afPLAN}, \mathbf{afPLA})$ and thus $(\mathbf{afPLAN}, \mathbf{afPLA}) \simeq (\mathbf{ncLBN}, \mathbf{aPLA})$. It follows that in order to show that $(\mathbf{ncLBN}, \mathbf{aPLA}) \prec (\mathbf{aPLAN}, \mathbf{aPLA})$ it suffices to show that $(\mathbf{afPLAN}, \mathbf{afPLA}) \prec (\mathbf{aPLAN}, \mathbf{aPLA})$.

Clearly $(\mathbf{afPLAN}, \mathbf{afPLA}) \preccurlyeq (\mathbf{aPLAN}, \mathbf{aPLA})$. Suppose towards a contradiction that $(\mathbf{aPLAN}, \mathbf{aPLA}) \preccurlyeq (\mathbf{afPLAN}, \mathbf{afPLA})$. By Lemma 9.5, there are a sentence $\psi \in \mathbf{aPLA}$ and $(\mathbb{P}_n : n \in \mathbb{N}^+)$ induced by an $aPLA$-network such that $\lim_{n\to\infty} \mathbb{P}_n(\{\mathcal{A} \in \mathbf{W}_n : \mathcal{A}(\psi) = 1\})$ exists and is neither 0 nor 1. According to the assumption there is a aggregation-free PLA-sentence ψ' which is asymptotically equivalent to ψ with respect to $(\mathbb{P}_n : n \in \mathbb{N}^+)$. As ψ' is aggregation-free it follows that there is $c \in [0, 1]$ such that $\mathcal{A}(\psi') = c$ for *every* finite structure $\mathcal{A}$. But this contradicts that ψ and ψ' are asymptotically equivalent. Finally, we clearly have $(\mathbf{aPLAN}, \mathbf{aPLA}) \preccurlyeq (\mathbf{aPLAN}^+, \mathbf{aPLA}^+)$.

(4) The first-order quantifiers $\exists$ and $\forall$ can be expressed in $aPLA^+$ by using the aggregation functions max and min. Therefore $FO \leq aPLA^+$. We also have $coPLA^+ \leq aPLA^+$, so $(\mathbf{coPLAN}^+, \mathbf{FO}) \preccurlyeq (\mathbf{aPLAN}^+, \mathbf{aPLA}^+)$. It remains to prove that

$$(\mathbf{aPLAN}^+, \mathbf{aPLA}^+) \not\preccurlyeq (\mathbf{coPLAN}^+, \mathbf{FO}).$$

But this follows because, for example, '1/2' is a sentence of $aPLA$ which takes the value 1/2 in every finite structure while first-order sentences can only take the values 0 and 1.

(5) We get $(\mathbf{coPLAN}^+, \mathbf{qfFO}) \simeq (\mathbf{coPLAN}^+, \mathbf{sCPL})$ from Theorem 8.6. From (1) we get $(\mathbf{ncLBN}, \mathbf{ncCPL}) \simeq (\mathbf{qfLBN}, \mathbf{qfFO})$. Since for every quantifier-free lifted Bayesian network there is a $coPLA^+$-network which induces the same sequence of distributions we get $(\mathbf{qfLBN}, \mathbf{qfFO}) \preccurlyeq (\mathbf{coPLAN}^+, \mathbf{qfFO})$. So to prove that $(\mathbf{ncLBN}, \mathbf{ncCPL}) \prec (\mathbf{coPLAN}^+, \mathbf{qfFO})$ it suffices to show that

$$(\mathbf{qfLBN}, \mathbf{qfFO}) \prec (\mathbf{coPLAN}^+, \mathbf{qfFO}).$$

But since every quantifier-free lifted Bayesian network is a noncritical lifted Bayesian network this follows from Lemma 9.5. It remains to prove that $(\mathbf{coPLAN}^+, \mathbf{sCPL}) \prec (\mathbf{coPLAN}^+, \mathbf{FO})$. As $(\mathbf{coPLAN}^+, \mathbf{qfFO}) \simeq (\mathbf{coPLAN}^+, \mathbf{sCPL})$ it suffices to prove that $(\mathbf{coPLAN}^+, \mathbf{qfFO}) \prec (\mathbf{coPLAN}^+, \mathbf{FO})$.
Clearly $(\mathbf{coPLAN}^+, \mathbf{qfFO}) \preccurlyeq (\mathbf{coPLAN}^+, \mathbf{FO})$ so it remains to show that

$$(\mathbf{coPLAN}^+, \mathbf{FO}) \not\preccurlyeq (\mathbf{coPLAN}^+, \mathbf{qfFO}).$$

By Lemma 9.5 there is a first-order sentence φ and $(\mathbb{P}_n : n \in \mathbb{N}^+)$ induced by a *coPLA*-network such that $\lim_{n\to\infty} \mathbb{P}_n(\varphi)$ exists and is neither 0 nor 1. If $(\mathbf{coPLAN^+}, \mathbf{FO}) \preccurlyeq (\mathbf{coPLAN^+}, \mathbf{qfFO})$ then there is a quantifier-free first-order sentence φ' which is asymptotically equivalent to φ with respect to $(\mathbb{P}_n : n \in \mathbb{N}^+)$, which (as the sentences are $\{0,1\}$-valued) means that they are almost surely equivalent with respect to $(\mathbb{P}_n : n \in \mathbb{N}^+)$. But φ' can only be $\perp$ or $\top$ so it means that $\lim_{n\to\infty} \mathbb{P}_n(\varphi)$ is either 0 or 1, a contradiction.

(6) From Theorem 8.6 we get $(\mathbf{coPLAN^+}, \mathbf{sCPL}) \simeq (\mathbf{coPLAN^+}, \mathbf{qfFO})$. Since every quantifier-free first-order formula is equivalent to an aggregation-free *PLA*-formula we get $(\mathbf{coPLAN^+}, \mathbf{qfFO}) \preccurlyeq (\mathbf{coPLAN^+}, \mathbf{afPLA})$. If $\varphi \in afPLA$ is the formula '1/2', then for every $(\mathbb{P}_n : n \in \mathbb{N}^+)$ induced by a *coPLA*-network there is *no* quantifier-free first-order sentence φ' which is asymptotically equivalent to φ, because first-order formulas only take the values 0 or 1. It follows that $(\mathbf{coPLAN^+}, \mathbf{qfFO}) \prec (\mathbf{coPLAN^+}, \mathbf{afPLA})$. This together with $(\mathbf{coPLAN^+}, \mathbf{sCPL}) \simeq (\mathbf{coPLAN^+}, \mathbf{qfFO})$ gives $(\mathbf{coPLAN^+}, \mathbf{sCPL}) \prec (\mathbf{coPLAN^+}, \mathbf{afPLA})$. From Theorem 5.11 we get

$$(\mathbf{coPLAN^+}, \mathbf{afPLA}) \simeq (\mathbf{coPLAN^+}, \mathbf{coPLA^+}).$$

It remains to prove that $(\mathbf{coPLAN^+}, \mathbf{coPLA^+}) \prec (\mathbf{aPLAN^+}, \mathbf{aPLA^+})$. Since

$$(\mathbf{coPLAN^+}, \mathbf{afPLA}) \simeq (\mathbf{coPLAN^+}, \mathbf{coPLA^+})$$

it suffices to prove that $(\mathbf{coPLAN^+}, \mathbf{afPLA}) \prec (\mathbf{aPLAN^+}, \mathbf{aPLA^+})$. But this follows from Part (ii) of Lemma 9.5 because every aggregation-free *PLA*-sentence has the same value in every structure.

(7) We clearly have $(\mathbf{afPLAN}, \mathbf{afPLA}) \preccurlyeq (\mathbf{coPLAN^+}, \mathbf{coPLA^+})$ so we only show that $(\mathbf{coPLAN^+}, \mathbf{coPLA^+}) \not\preccurlyeq (\mathbf{afPLAN}, \mathbf{afPLA})$.
But since $(\mathbf{afPLAN}, \mathbf{afPLA}) \simeq (\mathbf{ncLBN}, \mathbf{afPLA})$ by Part (2) this follows from Lemma 9.5.

(8) The statements

$$\begin{aligned}&(\mathbf{ncLBN}, \mathbf{aPLA}) \not\preccurlyeq (\mathbf{coPLAN^+}, \mathbf{FO}),\\ &(\mathbf{coPLAN^+}, \mathbf{coPLA^+}) \not\preccurlyeq (\mathbf{coPLAN^+}, \mathbf{FO}), \text{ and}\\ &(\mathbf{ncLBN}, \mathbf{aPLA}) \not\preccurlyeq (\mathbf{coPLAN^+}, \mathbf{sCPL})\end{aligned}$$

follow since *aPLA* and *coPLA*$^+$ contain for example the sentence '1/2', the value of which is always 1/2, while sentences in *FO* and in *CPL* can only take the values 0 and 1. The statements

$$\begin{aligned}&(\mathbf{coPLAN^+}, \mathbf{FO}) \not\preccurlyeq (\mathbf{ncLBN}, \mathbf{aPLA}),\\ &(\mathbf{coPLAN^+}, \mathbf{FO}) \not\preccurlyeq (\mathbf{coPLAN^+}, \mathbf{coPLA^+}), \text{ and}\\ &(\mathbf{coPLAN^+}, \mathbf{sCPL}) \not\preccurlyeq (\mathbf{ncLBN}, \mathbf{aPLA})\end{aligned}$$

follow from Lemma 9.5 and the above proved fact that $(\mathbf{coPLAN^+}, \mathbf{afPLA}) \simeq (\mathbf{coPLAN^+}, \mathbf{coPLA^+})$. □

10. Conclusion

We introduced the probability logic *coPLA*$^+$ which allows for probability formulas built using strongly admissible aggregation functions, which satisfy stronger continuity requirements than the admissible aggregation functions used in the *aPLA*-formulas studied by [KW23]. The stricter requirements reduce expressivity, ruling out the classical existential and universal quantifiers and their multivalued counterparts, maximum and minimum. However, $coPLA^{+}$ covers for example the arithmetic and geometric mean as aggregation functions, which can model a dependence of one relation on the relative frequency of another, and can also model (directed versions of) the sparse random graphs studied by Shelah and Spencer [SS88]. We showed that queries expressible in $coPLA^{+}$ are asymptotically equivalent to aggregation-free queries with respect to a given $coPLA^{+}$-network. An analogous quantifier elimination result with respect to $coPLA^{+}$-networks was shown to hold for safe formulas of conditional probability logic, which may include conditional relative frequency quantifiers but not classical universal or existential quantification. As a special case we obtained convergence results for expressive probability logics even over such random graphs where first-order formulas can have divergent probabilities, such as those studied by Shelah and Spencer [SS88]. Finally, we integrated the new results obtained here and previous results in [Kop20, KW23, SS88] by introducing the notion of an inference framework. We classified several inference frameworks related to the present work and to [Kop20, KW23, SS88] by means of their "relative asymptotic expressivity" which is defined using the transitive notion of one inference framework being asymptotically at least as expressive as another.

Acknowledgement

Parts of this work were accomplished during a research visit of the second author to the University of Uppsala, made possible by the support of LMUexcellent, funded by the Federal Ministry of Education and Research (BMBF) and the Free State of Bavaria under the Excellence Strategy of the Federal Government and the Länder.

References

[AS00] N. Alon and J. H. Spencer. *The Probabilistic Method*. John Wiley & Sons, 2 edition, 2000.

[Ber08] M. Bergmann. *An Introduction to Many-Valued and Fuzzy Logic:Semantics, Algebras, and Derivation Systems*. Cambridge University Press, 2008.

[BP98] S. Brin and L. Page. The anatomy of a large-scale hypertextual web search engine. *Computer Networks and ISDN Systems*, 30:107–117, 1998.

[Che52] H. Chernoff. A measure of the asymptotic efficiency for tests of a hypothesis based on the sum of observations. *Annals of Mathemathical Statistics*, 23:493–509, 1952.

[CM19] F. G. Cozman and D. D. Maua. The finite model theory of bayesian network specifications: Descriptive complexity and zero/one laws. *International Journal of Approximate Reasoning*, 110:107–126, 2019.

[DRKNP16] L. De Raedt, K. Kersting, S. Natarajan, and D. Poole. *Statistical Relational Artificial Intelligence: Logic, Probability, and Computation*, volume 32 of *Synthesis Lectures on Artificial Intelligence and Machine Learning*. Morgan & Claypool Publishers, 2016.

[Ebb85] H.-D. Ebbinghaus. Extended logics: The general framework. In J. Barwise and S. Feferman, editors, *Model-Theoretic Logics*, pages 25–76. Springer-Verlag, 1985.

[EF99] H-D. Ebbinghaus and J. Flum. *Finite Model Theory*, volume 2. Springer-Verlag, Berlin Heidelberg New York, 1999. doi:10.1007/3-540-28788-4.

[FWF23] D. Ravdin F. Weitkämper and R. Fabry. Statistical relational learning with scaled weight parameters. In F. Lissi E. Bellodi and R. Zese, editors, *Inductive Logic Programming - 32nd International Conference 2023*, pages 139—153. Springer, 2023.

[GT07] L. Getoor and B. Taskar. *Introduction to Statistical Relational Learning*. The MIT Press, 2007.

[GVdBP21] Sriraam Natarajan Guy Van den Broeck, Kristian Kersting and David Poole, editors. *An Introduction to Lifted Probabilistic Inference*. The MIT Press, 2021.

[Jae97] Manfred Jaeger. Relational Bayesian networks. In Dan Geiger and Prakash P. Shenoy, editors, *UAI 97: Proceedings of the Thirteenth Conference on Uncertainty in Artificial Intelligence, Brown University, Providence, Rhode Island, USA, August 1-3, 1997*, pages 266–273. Morgan Kaufmann, 1997. URL: `https://dslpitt.org/uai/displayArticleDetails.jsp?mmnu=1&smnu=2&article_id=320&proceeding_id=13`.

[Jae98] Manfred Jaeger. Convergence results for relational Bayesian networks. In Dan Geiger and Prakash P. Shenoy, editors, *Proceedings. Thirteenth Annual IEEE Symposium on Logic in Computer Science (LICS 98)*, pages 44–55, 1998. `doi:10.1109/LICS.1998.705642`.

[JW02] G. Jeh and J. Widom. Simrank: A measure of structural-context similarity. In D. Hand, D. A. Keim, and R. NG, editors, *KDD'02: Proceedings of the eighth ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 538–543. ACM Press, 2002.

[KBK+14] Seyed Mehran Kazemi, David Buchman, Kristian Kersting, Sriraam Natarajan, and David Poole. Relational logistic regression. In Chitta Baral, Giuseppe De Giacomo, and Thomas Eiter, editors, *Principles of Knowledge Representation and Reasoning: Proceedings of the Fourteenth International Conference, KR 2014, Vienna, Austria, July 20-24, 2014*. AAAI Press, 2014. URL: `http://www.aaai.org/ocs/index.php/KR/KR14/paper/view/8013`.

[KDR07] K. Kersting and L. De Raedt. Bayesian logic programming. In L. Getoor and B. Taskar, editors, *Introduction to Statistical relational learning*, page 291. MIT Press, Cambridge MA, 2007.

[KL09] H. J. Keisler and W. B. Lotfallah. Almost everywhere elimination of probability quantifiers. *The Journal of Symbolic Logic*, 74:1121–1142, 2009.

[KMG15] A. Kimmig, L. Mihalkova, and L. Getoor. Lifted graphical models: a survey. *Machine Learning*, 99(1):1–45, Apr 2015. `doi:10.1007/s10994-014-5443-2`.

[Kop20] Vera Koponen. Conditional probability logic, lifted Bayesian networks, and almost sure quantifier elimination. *Theoretical Computer Science*, 848:1–27, Dec 2020. `doi:10.1016/j.tcs.2020.08.006`.

[KW23] V. Koponen and F. Weitkämper. Asymptotic elimination of partially continuous aggregation functions in directed graphical models. *Information and Computation*, 293:105061, 2023. `doi:10.1016/j.ic.2023.105061`.

[MBGS19] H. Mittal, A. Bhardwaj, V. Gogate, and P. Singla. Domain-size aware Markov logic networks. In Kamalika Chaudhuri and Masashi Sugiyama, editors, *The 22nd International Conference on Artificial Intelligence and Statistics, AISTATS 2019, 16-18 April 2019, Naha, Okinawa, Japan*, volume 89 of *Proceedings of Machine Learning Research*, pages 3216–3224. PMLR, 2019. URL: `http://proceedings.mlr.press/v89/mittal19a.html`.

[Nov19] S. Y. Novak. On the accuracy of poisson approximation. *Extremes*, 22(4):729–748, Dec 2019. `doi:10.1007/s10687-019-00350-6`.

[PBK+14] D. Poole, D. Buchman, S. M. Kazemi, K. Kersting, and S. Natarajan. Population size extrapolation in relational probabilistic modelling. In Umberto Straccia and Andrea Calì, editors, *Scalable Uncertainty Management - 8th International Conference, SUM 2014, Oxford, UK, September 15-17, 2014. Proceedings*, volume 8720 of *Lecture Notes in Computer Science*, pages 292–305. Springer, 2014. `doi:10.1007/978-3-319-11508-5\_25`.

[SS88] S. Shelah and J. Spencer. Zero-one laws for sparse random graphs. *Journal of the American Mathematical Society*, 1(1):97–115, 2022/05/07/ 1988. Full publication date: Jan., 1988. `doi:10.2307/1990968`.

[Wei21] F. Weitkämper. An asymptotic analysis of probabilistic logic programming, with implications for expressing projective families of distributions. *Theory and Practice of Logic Programming*, 21(6):802–817, 2021. `doi:10.1017/S1471068421000314`.

[Wei24] F. Weitkämper. Functional lifted bayesian networks. In S. Muggleton and A. Tamaddoni-Nezhad, editors, *Proceedings of the 31st International Conference of Inductive Logic Programming 2022*, pages 142–156. Springer, 2024.