Local-global convergence, an analytic and structural approach

Jaroslav Nešetřil, Patrice Ossona de Mendez

In memory of Bohuslav Balcar

Abstract. Based on methods of structural convergence we provide a unifying view of local-global convergence, fitting to model theory and analysis. The general approach outlined here provides a possibility to extend the theory of local-global convergence to graphs with unbounded degrees. As an application, we extend previous results on continuous clustering of local convergent sequences and prove the existence of modeling quasi-limits for local-global convergent sequences of nowhere dense graphs.

Keywords: structural limit; Borel structure; modeling; local-global convergence

Classification: 03C13, 03C98, 05C99

The true logic of the world is in the calculus of probabilities.

James Clerk Maxwell

1. Introduction

The study of graph limits recently gained a strong interest, motivated both by the study of large networks and the emerging studies of real evolving networks. The different notions of graph limit (*left convergence* see [7], [8], [9], [10], [19], *local convergence* see [5]), and the basic notions of graph similarity on which they are based, opens a vast panorama. These frameworks have in common to be built on statistics of locally defined motives when the vertices of the graphs in the sequence are sampled uniformly and independently. A unified framework for the study of convergence of structures has been introduced by the authors in [20].

In this setting the notion of convergence is, in essence, model theoretic, and relies on the the following notions:

For a σ-structure $\mathbf{A}$ and a first-order formula φ (in the language of σ, with free variables $x_1, \dots, x_p$), we denote by $\varphi(\mathbf{A})$ the *satisfying set* of φ in $\mathbf{A}$:

$$\varphi(\mathbf{A}) = \{(v_1, \dots, v_p) \in A^p \colon \ \mathbf{A} \models \varphi(v_1, \dots, v_p)\},$$

DOI 10.14712/1213-7243.2015.280

Supported by grant ERCCZ LL-1201 and CE-ITI P202/12/G061, and by the European Associated Laboratory "Structures in Combinatorics" (LEA STRUCO) and partially supported by ANR project Stint under reference ANR-13-BS02-0007.

and we define the *Stone pairing* of φ and $\mathbf{A}$ as the probability

$$\langle \varphi, \mathbf{A} \rangle = \frac{|\varphi(\mathbf{A})|}{|A|^p}$$

that $\mathbf{A}$ satisfies φ for a (uniform independent) random interpretation of the random variables.

A sequence $\mathbf{A} = (\mathbf{A}_n)_{n\in\mathbb{N}}$ of finite σ-structures is FO-*convergent* if the sequence $\langle \varphi, \mathbf{A} \rangle = (\langle \varphi, \mathbf{A}_n \rangle)_{n\in\mathbb{N}}$ converges for every first-order formula φ. One similarly defines a weakened notion of X-*convergence* (for a fragment X of first-order logic) by restricting the range of the test formulas φ to X. In particular, for a sequence of graphs with growing orders, the QF-*convergence* (that is of convergence driven by the fragment QF of quantifier-free formulas) is equivalent to left convergence, and the $\mathrm{FO}^{\mathrm{local}}$-*convergence* (that is of convergence driven by the fragment $\mathrm{FO}^{\mathrm{local}}$ of the so-called local formulas) is equivalent to local convergence when restricting to sequences of graphs with bounded degrees, see [20]. The study of X-convergence of structures and the related problems is called shortly "structural limits". It extends all previously considered types of nongeometric convergence of combinatorial structures (two above and also, e.g. permutations) to general relational structures. A survey of structural limits can be found in [22], [21].

In order to strengthen the notion of local convergence of sequences of connected bounded degree graphs, the notion of *local-global convergence* has been introduced by H. Hatami, L. Lovász and B. Szegedy in [12], based on the framework introduced in [6]. The original definition of local-global convergence, which is based on the total variation distance between the distributions of colored neighborhoods of radius r in the graphs in the sequence, is admittedly quite technical (see Definitions 3.3.1 and 3.4.1). Roughly speaking, the strengthening of local convergence into local-global convergence relates to the extension of the considered properties from first-order logic to existential monadic second order logic.

In this paper we introduce the notion of *lift-Hausdorff convergence*, which is defined generally for graphs and relational structures, based on a simple subsequence completion property (see Definition 3.4.3). The basic notions underlying this definition are the model theoretic notions of *lift* of a structure (that one can see as an augmentation of a structure by colors, new relations, etc.) and the dual notion of *shadow* consisting in forgetting all the additional relations of a lift (see Figure 1). Note that the notions of lift and shadow are close to model theoretic notions of expansion and reduct.

A lift/shadow pair is determined by the data of two signatures $\sigma \subset \sigma^+$, the signature of the shadows and the signature of the lifts. Given a lift/shadow pair, the corresponding notion of lift-Hausdorff convergence is (roughly) defined as follows: A sequence $\mathbf{A} = (\mathbf{A}_n)_{n\in\mathbb{N}}$ is *lift-Hausdorff* convergent if for every convergent subsequence $\mathbf{A}_f^+ = (\mathbf{A}^+_{f(n)})_{n\in\mathbb{N}}$ of lifts there exists a (full) convergent sequence $\mathbf{A}^*$ of lifts extending $\mathbf{A}_f^+$. (In this definition, the condition that $\mathbf{A}_f$ is a subsequence of lifts means that $\mathbf{A}^+_{f(n)}$ is a lift of $\mathbf{A}_{f(n)}$ and $f\colon \mathbb{N} \to \mathbb{N}$ is

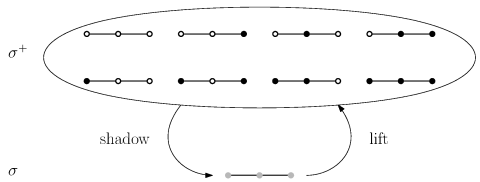


FIGURE 1. The *shadow* operation consists in forgetting some of the relations, while the *lift* operation consists in adding some relations in all possible ways.

an increasing function defining a subsequence of indices; the condition that $\mathbf{A}^*$ extends $\mathbf{A}^+_f$ means that $\mathbf{A}^*_{f(n)} = \mathbf{A}^+_{f(n)}$ for every integer n.)

We prove that this notion indeed extends the usual notion of local-global convergence in the case of monadic lifts and gives a characterization of lift-Hausdorff convergent sequences as Cauchy sequences for an appropriate metric (Theorem 3.3.2). Note that the case of monadic lifts can be interpreted by means of first-order formulas with set parameters.

We give a basic representation theorem for corresponding limits, which is based on the representation theorem for structural limits, see [20]: every limit of a lift-Hausdorff convergent sequence maybe represented by nonempty closed subset of the space of probability measures on some Stone space (Theorem 3.3.3). In Section 4.2 we discuss the possible notion of a limit object for lift-Hausdorff convergent sequences. Two possible notions can be considered:

- a strong notion of limit $\mathbf{L}$ for a sequence $\mathbf{A}$, such that for every convergent sequence $\mathbf{A}^+$ of lifts there exists an admissible lift $\mathbf{L}^+$ of $\mathbf{L}$ that is the limit of $\mathbf{A}^+$ and, conversely, for every admissible lift $\mathbf{L}^+$ of $\mathbf{L}$ there exists a convergent sequence $\mathbf{A}^+$ of lifts with limit $\mathbf{L}^+$;
- a weaker notion of limit $\mathbf{L}$ (in the spirit of the representation of local-global limits by graphings, see [12]) where we ask that for every $\varepsilon > 0$ and every convergent sequence $\mathbf{A}^+$ of lifts there exists an admissible lift $\mathbf{L}^+$ of $\mathbf{L}$ such that for every formula φ we have

$$\left|\langle \varphi, \mathbf{L}^+\rangle - \lim_{n\to\infty} \langle \varphi, \mathbf{A}^+_n\rangle\right| < c_\varphi \varepsilon \tag{1}$$

(where c_φ is a positive constant depending only on φ) and, conversely, for every admissible lift $\mathbf{L}^+$ of $\mathbf{L}$ there exists a convergent sequence $\mathbf{A}^+$ of lifts such that (1) holds.

In these definitions, an essential difficulty lies in the definition of admissible lifts. In the case of local-global convergent sequences of bounded degree graphs with a graphing limit, the notion of admissible lift corresponds to Borel colorings. In a more general setting, an admissible lift of a modeling should (at least) be a modeling itself. However, Borel colorings of a fixed graphing does not induce a closed subset of unimodular probability measures, as noticed in [12] and, more

generally, probability measures associated to modeling lifts of a fixed modeling does not form a closed subset, what is problematic for the stronger notion of limit.

When considering the weaker (standard) definition of a limit, the reverse direction appears to be quite difficult to handle in general. Indeed, in the bounded degree case, the r-neighborhood of a set with measure at most ε has measure at most $d^r\varepsilon$, what allows to easily approximates Borel colorings using a countable base. However such a property does not hold in the general unbounded degree case.

For this reason, we only consider one direction in the definition of limit: a *quasi-limit* $\mathbf{A}$ is a modeling such that for every $\varepsilon > 0$, every formula φ and every convergent sequence $\mathbf{A}^+$ of lifts there exists a modeling lift $\mathbf{L}^+$ of $\mathbf{L}$ such that (1) holds. In this setting, we prove that modelings (introduced in [25]) are not only limit objects for FO-convergent nowhere dense sequences (as proved in [26]) but also quasi-limits for (monadic) lift-Hausdorff convergent nowhere dense sequences (see Theorem 4.2.4).

Let us end this introduction by few remarks. The lifts involved in our local-global structural convergence are all monadic (and can be seen as coloring of vertices). It follows that the expressive power of such lifts is restricted to embeddings and classes which are hereditary. If we would consider more general lifts, such as coloring of the edges[1] then we could represent monomorphisms (not induced substructures) which in turn leads to monotone classes. Monotone classes of graphs which have modeling limits (of which graphing limits are a particular case) were characterized in [26] and coincides with nowhere dense classes (of graphs). This also coincides (in the case of monotone classes of graphs) with the notion of NIP and stable classes, see [1] (see also [23]). For hereditary classes the structure theory and the existence of modeling limits is more complicated (see [22]) and local-global convergence seems to provide a useful framework.

2. Preliminaries

2.1 Relational structures and first-order logic. A *signature* σ is a set of relation symbols with associated arities. In this paper we will consider countable signatures. A *σ-structure* $\mathbf{A}$ is defined by its *domain* A, which is a set, and by interpreting each relation symbol $R \in \sigma$ of arity k as a subset of A^k. We denote by $\mathrm{Rel}(\sigma)$ the set of all finite σ-structures and by $\mathscr{R}\mathrm{el}(\sigma)$ the class of all σ-structures.

A *first-order formula* φ in the language of σ-structures is a formula constructed using disjunction, conjunction, negation and quantification over elements, using the relations in σ and the equality symbol. A variable used in a formula φ is *free* if it is not bound by a quantifier. We always assume that free variables are named $x_1, \ldots, x_n, \ldots$ and we consider formulas obtained by renaming the free variables as distinct. For instance, $x_1 = x_2$ and $x_2 = x_3$ are distinct formulas. We also

[1]This would require to add some first-order restrictions on the lifts, what would not fundamentally change the framework presented in this paper.

Symbol	Fragment
$\mathrm{FO}(\sigma)$ or FO	All first order formulas
$\mathrm{FO}_p(\sigma)$ or FO_p	All first order formulas with free variables within $\{x_1, \ldots, x_p\}$
$\mathrm{FO}_0(\sigma)$ or FO_0	Sentences
$\mathrm{FO}^{\mathrm{local}}(\sigma)$ or $\mathrm{FO}^{\mathrm{local}}$	Local formulas
$\mathrm{FO}_p^{\mathrm{local}}(\sigma)$ or $\mathrm{FO}_p^{\mathrm{local}}$	Local formulas with free variables within $\{x_1, \ldots, x_p\}$
$\mathrm{QF}(\sigma)$ or QF	Quantifier free formulas

TABLE 1. Principal fragments of first-order logic considered in this paper.

consider two constants, 0 and 1 to denote the *false* and *true* statements. We denote by $\mathrm{FO}(\sigma)$ the (countable) set of all first-order formulas in the language of σ-structures. The conjunction and disjunction of formulas φ and ψ are denoted by $\varphi \wedge \psi$ and $\varphi \vee \psi$, and the negation of φ is denoted by $\neg\varphi$. We say that two formulas φ and ψ are *logically equivalent*, which we denote by $\varphi \iff \psi$, if one can infer one from the other (i.e. $\varphi \vdash \psi$ and $\psi \vdash \varphi$). Note that in first-order logic the notions of syntactic and semantic equivalence coincides. In this context we denote by $[\varphi]$ the equivalence class of φ with respect to logical equivalence. It is easily checked that $\mathcal{B}_\sigma = \mathrm{FO}(\sigma)/\!\iff$ is a countable Boolean algebra with minimum 0 and maximum 1, which is called the *Lindenbaum–Tarski algebra* of $\mathrm{FO}(\sigma)$.

In this paper we consider special fragments of first-order logic (see Table 1).

The Lindenbaum–Tarski algebra of a fragment $X \subseteq \mathrm{FO}(\sigma)$ will be denoted by $\mathcal{B}_\sigma^X$. For instance, $\mathcal{B}_\sigma^{\mathrm{QF}} = \mathrm{QF}(\sigma)/\!\iff$.

2.2 Functional analysis. Basic facts from functional analysis, which will be used in this paper, are recalled now.

A *standard Borel space* is a Borel space associated to a Polish space, i.e. a measurable space (X, Σ) such that there exists a metric on X making it a separable complete metric space with Σ as its Borel σ-algebra. Typical examples of standard Borel spaces are $\mathbb{R}$ and the Cantor space. Note that according to Maharam's theorem, all uncountable standard Borel spaces are (Borel) isomorphic. (The authors cannot resist the temptation to mention Balcar's award-winning work [4] in this context.)

In this paper, we shall mainly consider compact separable metric spaces. Note that if (M, d) is a compact separable metric space, then linear functionals on the space of real continuous functions on M can be represented, thanks to Riesz–Markov–Kakutani representation theorem, by Borel measures on M. We denote by $P(M)$ the space of all probability measures on M. A sequence of probability measures $(\mu_n)_{n\in\mathbb{N}}$ is *weakly convergent* if $\int f \,\mathrm{d}\mu_n$ converges for every (real-valued)

continuous function[2] f. Weak convergence defines the *weak topology* of $P(M)$, and (as we assumed that (M,d) is a compact separable metric space) this space is compact, separable, and metrizable by the Lévy–Prokhorov metric (based on the metric d):

$$\mathrm{d}^{\mathrm{LP}}(\mu_1,\mu_2) = \inf\{\varepsilon > 0\colon \mu_1(A) \le \mu_2(A^\varepsilon) + \varepsilon \ \text{ and } \ \mu_2(A) \le \mu_1(A^\varepsilon) + \varepsilon \text{ for all Borel } A\},$$

where $A^\varepsilon = \{x \in M\colon \ \exists\, y \in M \ \ \mathrm{d}(x,y) < \varepsilon\}$.

The Hausdorff metric is defined on the space of nonempty closed bounded subsets of a metric space. Consider a compact metric space (M,d), and let $\mathscr{C}_M$ be the space of nonempty closed subsets of M endowed with the *Hausdorff metric* defined by

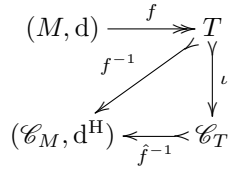
$$\mathrm{d}^{\mathrm{H}}(X,Y) = \max\bigl(\max_{x\in X}\min_{y\in Y}\mathrm{d}(x,y), \max_{y\in Y}\min_{x\in X}\mathrm{d}(x,y)\bigr).$$

One of the most important properties of Hausdorff metric is that the space of nonempty closed subsets of a compact set is also compact (see [13], and [27] for an independent proof). Hence the space $(\mathscr{C}_M,\mathrm{d}^{\mathrm{H}})$ is compact.

We can use the inverse function of a surjective continuous function from a compact metric space (M,d) to a (thus compact) Hausdorff space T to isometrically embed the space $\mathscr{C}_T$ (of nonempty closed subsets of T) into $(\mathscr{C}_M,\mathrm{d}^{\mathrm{H}})$. Then, using the natural injection $\iota\colon T \to \mathscr{C}_T$ (defined by $\iota(x) = \{x\}$) we pull back the Hausdorff distance on $\mathscr{C}_M$ into T:

$$\mathrm{d}_T(x,y) = \mathrm{d}^{\mathrm{H}}(f^{-1}(x), f^{-1}(y)).$$

The situation is summarized in the following diagram.



In this diagram f^{-1} denotes the mapping from T to $\mathscr{C}_M$ and $\hat{f}^{-1}$ the corresponding mapping from $\mathscr{C}_T$ to $\mathscr{C}_M$ defined by $\hat{f}^{-1}(X) = \{y\colon f(y) \in X\}$. Also remark that the metric d_T defined on T is usually not compatible with the original topology of T.

For the topology defined by the metric d_T, one can define the compactification of T, which may be identified with the closure of the image of T (by $f^{-1} \circ \iota$) in $\mathscr{C}_M$.

We shall make use of the following folklore result, which we prove here for completeness.

[2]We do not have to assume that f has compact support as we assumed that (M,d) is compact.

Lemma 2.2.1. *Let X, Y be compact standard Borel spaces and let $f\colon X \to Y$ be continuous. Let $P(X)$ and $P(Y)$ denote the metric space of probability measures on X and Y (with Lévy–Prokhorov metric).*

Then the pushforward by f, that is the mapping $f_*\colon P(X) \to P(Y)$ defined by $f_*(\mu) = \mu \circ f^{-1}$, is continuous.

Proof: Assume $\mu_n \Rightarrow \mu$ is a weakly convergent sequence of measures in $P(X)$. Then for every continuous function $g\colon Y \to \mathbb{R}$ it holds

$$\begin{aligned}\int_Y g(y)\,\mathrm{d}f_*(\mu)(y) &= \int_X g\circ f(x)\,\mathrm{d}\mu(x) = \lim_{n\to\infty}\int_X g\circ f(x)\,\mathrm{d}\mu_n(x)\\ &= \lim_{n\to\infty}\int_Y g(y)\,\mathrm{d}f_*(\mu_n)(y).\end{aligned}$$

Hence $f_*(\mu_n) \Rightarrow f_*(\mu)$. □

2.3 Sequences. In this paper we denoted sequences by sans serif letters. In particular, we denote by A a sequence of structures $\mathsf{A} = (\mathbf{A}_n)_{n\in\mathbb{N}}$, and by $\mathsf{X} = (X_n)_{n\in\mathbb{N}}$ a sequence of sets X_n, where X_n is a subset of the domain A_n of $\mathbf{A}_n$.

Subsequences will be denoted by A_f and X_f, where f is meant to be a strictly increasing function $f\colon \mathbb{N} \to \mathbb{N}$, and represent the sequences $(\mathbf{A}_{f(n)})_{n\in\mathbb{N}}$ and $(X_{f(n)})_{n\in\mathbb{N}}$. Note that $(\mathsf{A}_f)_g = \mathsf{A}_{g\circ f}$.

In order to simplify the notations, we extend binary relations and standard constructions to sequences by applying them component-wise. For instance $\mathsf{X} \subseteq \mathsf{Y}$ means $(\forall\, n \in \mathbb{N})\ X_n \subseteq Y_n$, $\mathsf{X} \cap \mathsf{Y}$ represents the sequence $(X_n \cap Y_n)_{n\in\mathbb{N}}$, and if $f\colon \mathrm{Rel}(\sigma) \to \mathbb{R}$ is a mapping then $f(\mathsf{A})$ represents the sequence $(f(\mathbf{A}_n))_{n\in\mathbb{N}}$.

We find these notations extremely helpful for our purposes.

2.4 Basics of structural convergence. Let σ be a countable signature, let X be a fragment of $\mathrm{FO}(\sigma)$. For $\varphi \in X$ with free variables within $x_1, \ldots, x_p$ and $\mathbf{A}$, we denote by $\langle \varphi, \mathbf{A}\rangle$ the probability that φ is satisfied in $\mathbf{A}$ for a random assignment of elements of A to the free variables of φ (for an independent and uniform random choice of the assigned elements), that is:

$$\langle \varphi, \mathbf{A}\rangle = \frac{|\{\overline{v} \in A^p\colon \mathbf{A} \models \varphi(\overline{v})\}|}{|A|^p}.$$

(Note that the presence of unused free variables does not change the value in the next equation.) In the special case where φ is a sentence, we get

$$\langle \varphi, \mathbf{A}\rangle = \begin{cases} 1 & \text{if } \mathbf{A} \models \varphi,\\ 0 & \text{otherwise.}\end{cases}$$

Two σ-structures $\mathbf{A}$ and $\mathbf{B}$ are *X-equivalent*, what we denote by $\mathbf{A} \equiv_X \mathbf{B}$, if we have $\langle \varphi, \mathbf{A}\rangle = \langle \varphi, \mathbf{B}\rangle$ for every $\varphi \in X$.

Example 2.4.1. Let QF^- be the fragment of quantifier-free formulas that do not use equality.
If a graph G is obtained from a graph H by blowing each vertex into k vertices (i.e. if G is the lexicographic product of H by an edgeless graph of order k) then $G \equiv_{\mathrm{QF}^-} H$.

If X is a fragment including QF or FO_0, then $\equiv_X$ is trivial on finite relational structures.

The case of $\mathrm{FO}_1^{\mathrm{local}}$-equivalence is settled by the next proposition.

Proposition 2.4.2. *For any two finite σ-structures $\mathbf{A}$ and $\mathbf{B}$ we have that $\mathbf{A} \equiv_{\mathrm{FO}_1^{\mathrm{local}}} \mathbf{B}$ if and only if there exists a finite σ-structure $\mathbf{C}$ and two positive integers a and b such that $\mathbf{A}$ is isomorphic to a copies of $\mathbf{C}$ and $\mathbf{B}$ is isomorphic to b copies of $\mathbf{C}$.*

Proof: Let $\mathbf{F}_1, \dots, \mathbf{F}_n, \dots$ be an enumeration of the finite σ-structures (up to isomorphism), and let $\varphi_i(x)$ be a local formula expressing that the connected component of x is isomorphic to $\mathbf{F}_i$ (i.e. that the ball of radius $|F_i| + 1$ around x is isomorphic to F_i). Then $\langle \varphi_i, \mathbf{A} \rangle$ is equal to the product of $|\mathbf{F}_i|/|\mathbf{A}|$ by the number of connected components of $\mathbf{A}$ isomorphic to $\mathbf{F}_i$. Thus there exists a positive integer q and nonnegative integers $p_1, \dots, p_n, \dots$ such that $\langle \varphi_i, \mathbf{A} \rangle = p_i/q$ and the set of all positive p_i values is setwise coprime. Then if $\mathbf{C}$ consists in thus union (over i) of p_i copies of $\mathbf{F}_i$, it is immediate that $\mathbf{A}$ and $\mathbf{B}$ consist in a positive number of copies of $\mathbf{C}$. □

A sequence $\mathbf{A}$ of σ-structures is *X-convergent* if $\langle \varphi, \mathbf{A} \rangle = (\langle \varphi, \mathbf{A}_n \rangle)_{n \in \mathbb{N}}$ converges for each $\varphi \in X$. This provides a unifying to left and local convergence, as mentioned in the introduction: left convergence coincides with QF^--convergence and local convergence with $\mathrm{FO}^{\mathrm{local}}$-convergence (when restricted to graphs with bounded degrees). The term of *structural convergence* covers the general notions of X-convergence.

The basic result of [20], which is going to provide us a guideline for a proper generalization of local-global convergence is the representation theorem for structural limits in terms of probability measures. We adopt [20] to the setting of this paper.

2.5 The representation theorem for structural limits. For a countable signature σ and a fragment X of $\mathrm{FO}(\sigma)$ we denote by S_X^σ the Stone dual of the Lindenbaum–Tarski $\mathcal{B}_\sigma^X$ of X, which is a compact Polish space. Recall that the points of $\mathcal{B}_\sigma^X$ are the maximal consistent subsets of X (or equivalently the ultrafilters on $\mathcal{B}_\sigma^X$). The topology of S_X^σ is generated by the base of the clopen subsets of S_X^σ, which are in bijection with the formulas in X by

$$\varphi \mapsto K(\varphi) = \{t \in S_X^\sigma : \ \varphi \in t\}.$$

In the setting of this paper we work with metric (and, notably, pseudo-metric) spaces. First note that the topology of S_X^σ is metrizable by the several metrics, including the metrics we introduce now.

A *chain covering* of X is an increasing sequence $\mathfrak{X} = (X_1, X_2, \dots)$ of finite sets (i.e. $X_1 \subset X_2 \subset \cdots \subset X_n \subset \dots$) such that every formula in X is logically equivalent to a formula in $\bigcup_{i\geq 1} X_i$. The metric $\delta_\mathfrak{X}$ induced by $\mathfrak{X}$ on S_σ^X is defined by

$$\delta_\mathfrak{X}(t_1, t_2) = \inf\Big\{\frac{1}{n}\colon\ (t_1 \vartriangle t_2) \cap X_n = \emptyset\Big\}, \tag{2}$$

where $t_1 \vartriangle t_2$ stands for the symmetric difference of the sets t_1 and t_2.

First-order limits (shortly FO-limits) and, more generally, X-limits can be uniquely represented by a probability measure μ on the Stone space S dual to the Lindenbaum–Tarski algebra of the formulas. This can be formulated as follows.

Theorem 2.5.1 ([20])**.** *Let σ be a countable signature, let X be a fragment of $\mathrm{FO}(\sigma)$ closed under disjunction, conjunction and negation, let $\mathcal{B}_\sigma^X$ be the Lindenbaum–Tarski algebra of X, and let S_σ^X be the Stone dual of $\mathcal{B}_\sigma^X$.*

Then there is a map T_σ^X from the space $\mathrm{Rel}(\sigma)$ of finite σ-structures to the space of $P(S_\sigma^X)$ of probability measures on the Stone space S_σ^X, such that for every $\mathbf{A} \in \mathrm{Rel}(\sigma)$ and every $\varphi \in X$ we have

$$\langle \varphi, \mathbf{A}\rangle = \int_{S_\sigma^X} \mathbf{1}_\varphi(t)\,\mathrm{d}\mu_\mathbf{A}(t), \tag{3}$$

where $\mu_\mathbf{A} = T_\sigma^X(\mathbf{A})$ and $\mathbf{1}_\varphi$ is the indicator function of the clopen subset $K(\varphi)$ of S_σ^X dual to the formula $\varphi \in X$ in Stone duality, i.e.

$$\mathbf{1}_\varphi(t) = \begin{cases} 1 & \text{if } \varphi \in t, \\ 0 & \text{otherwise.} \end{cases}$$

Additionally, if the fragment X includes FO_0 or QF then the mapping T_σ^X is one-to-one[3].

In this setting, a sequence $\mathbf{A}$ of finite σ-structures is X-convergent if and only if the measures $T_\sigma^X(\mathbf{A}_n)$ converge weakly to some measure μ. Then for every first-order formula $\varphi \in X$ we have

$$\int_{S_\sigma^X} \mathbf{1}_\varphi(t)\,\mathrm{d}\mu(t) = \lim_{n\to\infty} \int_{S_\sigma^X} \mathbf{1}_\varphi(t)\,\mathrm{d}\mu_{\mathbf{A}_n}(t) = \lim_{n\to\infty} \langle \varphi, \mathbf{A}_n\rangle, \tag{4}$$

where $\mu_{\mathbf{A}_n} = T_\sigma^X(\mathbf{A}_n)$.

Assume that a subgroup Γ of the group $\mathfrak{S}_\omega$ of permutations of $\mathbb{N}$ acts on the first-order formulas in X by permuting the free variables. Then this action induces an action on S_σ^X, and the probability measure $T_\sigma^X(\mathbf{A})$ associated with a finite

[3]Note that in [20] the condition on X was erroneously omitted.

structure $\mathbf{A}$ *is obviously* Γ*-invariant, thus so is the weak limit* μ *of a sequence* $(T_\sigma^X(\mathbf{A}_n))_{n\in\mathbb{N}}$ *of probability measures associated with the finite structures of an* X*-convergent sequence. It follows that the measure* μ *appearing in* (4) *has the property to be* Γ*-invariant.*

This theorem generalizes the representation of the limit of a left-convergent sequence of graphs by an infinite exchangeable random graph [2], [16] and the representation of the limit of a local-convergent sequence of bounded degree graphs by a unimodular distribution, see [5]. Figure 2 schematically depicts some of the notions related to the representation theorem.

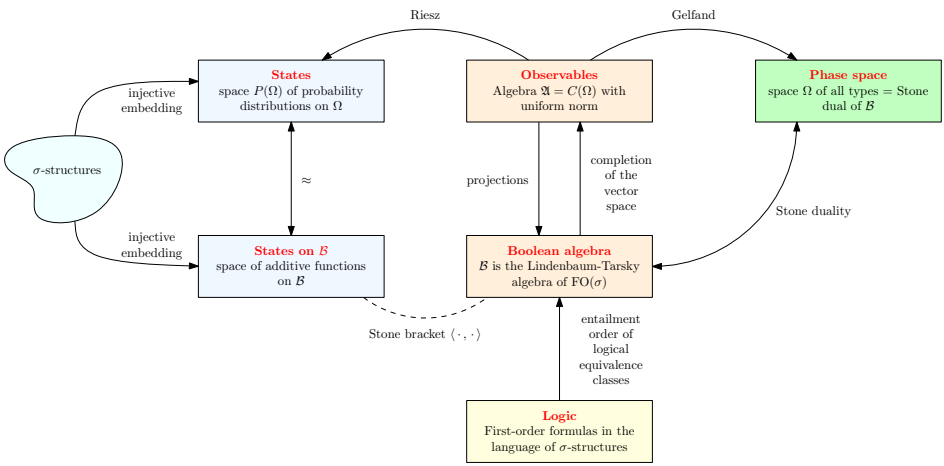


FIGURE 2. Sketch of the spaces involved in the distributional representation of structural limits of σ-structures.

The weak topology of $P(S_\sigma^X)$ is metrizable by using the Lévy–Prokhorov metrics based on the metrics $\delta_{\mathfrak{X}}$ (where $\mathfrak{X}$ is a fixed chain covering of X). Using the fact that $\delta_{\mathfrak{X}}$ is an ultrametric, we obtain the following more practical expression for the Lévy–Prokhorov metric $\mathrm{d}_{\mathfrak{X}}^{\mathrm{LP}}$ associated to $\delta_{\mathfrak{X}}$:

$$\mathrm{d}_{\mathfrak{X}}^{\mathrm{LP}}(\mu_1,\mu_2) = \inf_{n\in\mathbb{N}}\left\{\max\left(\frac{1}{n}, \max_{\varphi\in X_n}\left|\int_{S_\sigma^X} \mathbf{1}_\varphi(t)\,\mathrm{d}(\mu_1-\mu_2)(t)\right|\right)\right\}. \tag{5}$$

This metric in turn uniquely defines a pseudometric $\mathrm{dist}_{\mathfrak{X}}$ on $\mathrm{Rel}(\sigma)$ such that the mapping T_σ^X induces an isometric embedding of $\mathrm{Rel}(\sigma)/{\equiv_X}$ into $P(S_\sigma^X)$:

$$\mathrm{dist}_{\mathfrak{X}}(\mathbf{A},\mathbf{B}) = \mathrm{d}_{\mathfrak{X}}^{\mathrm{LP}}(T_\sigma^X(\mathbf{A}), T_\sigma^X(\mathbf{B})).$$

Note that we have the following expression for $\mathrm{dist}_{\mathfrak{X}}$:

$$\mathrm{dist}_{\mathfrak{X}}(\mathbf{A},\mathbf{B}) = \inf_{n\in\mathbb{N}}\left\{\max\left(\frac{1}{n}, \max_{\varphi\in X_n}|\langle\varphi,\mathbf{A}\rangle - \langle\varphi,\mathbf{B}\rangle|\right)\right\}. \tag{6}$$

It is easily checked that, as expected, a sequence $\mathbf{A}$ is X-convergent if and only if it is Cauchy for $\mathrm{dist}_{\mathfrak{X}}$.

We denote by $\mathscr{M}_\sigma^X \subseteq P(S_\sigma^X)$ the space of probability measures on S_σ^X associated to finite σ-structures:

$$\mathscr{M}_\sigma^X = \{T_\sigma^X(\mathbf{A}) \colon \ \mathbf{A} \in \mathrm{Rel}(\sigma)\}.$$

We denote by $\overline{\mathscr{M}_\sigma^X}$ the weak closure of $\mathscr{M}_\sigma^X$ in $P(S_\sigma^X)$ and by $(\overline{\mathrm{Rel}(\sigma)}^X, \mathrm{dist}_{\mathfrak{X}})$ the completion of the pseudometric space $(\mathrm{Rel}(\sigma), \mathrm{dist}_{\mathfrak{X}})$. Note that $(\overline{\mathrm{Rel}(\sigma)}^X, \mathrm{dist}_{\mathfrak{X}})$ has a dense subspace naturally identified with $(\mathrm{Rel}(\sigma)/\equiv_X, \mathrm{dist}_{\mathfrak{X}})$, and T_σ^X induces an isometric isomorphism of $(\overline{\mathrm{Rel}(\sigma)}^X, \mathrm{dist}_{\mathfrak{X}})$ and $(\overline{\mathscr{M}_\sigma^X}, \mathrm{d}_{\mathfrak{X}}^{\mathrm{LP}})$. Consequently both spaces are separable compact metric spaces.

3. From interpretation to lift convergence

Our basic approach to local-global convergence is by means of lifts of structures, which demands a change of signature. By doing so we still have to preserve some functorial properties and this is done by means of interpretations.

3.1 A categorical approach to interpretations. Interpretations of classes of relational structures in other classes of relational structures are a useful and powerful technique to transfer properties from one class of structures to another (with possibly a different signature).

First, we define interpretations syntactically (in the spirit of [18]), which allows us to organize them as a category. This functorial view will be particularly useful in our setting.

Let τ, σ be countable relational signatures. An *interpretation* I of σ-structures in τ-structures is a triple $(\nu, \eta, (\varrho_R)_{R\in\sigma})$, where:

- $\nu(\overline{x}) \in \mathrm{FO}(\tau)$ is a formula defined on p tuples of variables $\overline{x}$;
- $\eta(\overline{x}, \overline{y}) \in \mathrm{FO}(\tau)$ is a formula defining an equivalence relation on p-tuples (satisfying ν);
- for each relation $R \in \sigma$ of arity k, the formula $\varrho_R(\overline{x}_1, \dots, \overline{x}_k) \in \mathrm{FO}(\tau)$ (with $|\overline{x}_1| = \cdots = |\overline{x}_k| = p$) is compatible with η, meaning

$$\bigwedge_{i=1}^{k} \eta(\overline{x}_i, \overline{y}_i) \vdash \varrho_R(\overline{x}_1, \dots, \overline{x}_k) \ \leftrightarrow \ \varrho_R(\overline{y}_1, \dots, \overline{y}_k).$$

By replacing equality by η, relation R by ϱ_R and by conditioning quantifications using ν one easily checks that the interpretation I allows to associate to each formula $\varphi(x_1, \dots, x_k) \in \mathrm{FO}(\sigma)$ a formula $\widehat{\varphi}(\overline{x}_1, \dots, \overline{x}_k) \in \mathrm{FO}(\tau)$. We define

$$L(\mathsf{I}) \colon \ \mathrm{FO}(\sigma) \ \to \ \mathrm{FO}(\tau)$$

as the mapping $\varphi \mapsto \widehat{\varphi}$.

Note that we have $L(\mathsf{I})(1) = \nu$, $L(\mathsf{I})(x_1 = x_2) = \eta$, and $L(\mathsf{I})(R) = \varrho_R$ for every $R \in \sigma$. Hence $L(\mathsf{I})$ fully determines I.

This definition allows us to consider interpretations $\mathsf{I}\colon \tau \to \sigma$ as morphisms in a *category of interpretations*. The objects of this category are all countable relational signatures (here denoted by $\sigma, \tau, \dots$) and morphisms $\mathsf{I}\colon \tau \to \sigma$ are triples $(\nu, \eta, (\varrho_R)_{R\in\sigma})$ forming an interpretation as above. Morphisms compose as if $\mathsf{I}\colon \tau \to \sigma$ and $\mathsf{J}\colon \sigma \to \kappa$ are interpretations then we can define

$$\mathsf{J} \circ \mathsf{I} = (L(\mathsf{J} \circ \mathsf{I})(x_1 = x_1), L(\mathsf{J} \circ \mathsf{I})(x_1 = x_2), (L(\mathsf{J} \circ \mathsf{I})(S))_{S\in\kappa}).$$

The identity (for σ) is provided by the morphism $(x_1 = x_1, x_1 = x_2, (R)_{R\in\sigma})$. Thus we indeed have a category of interpretations.

A *basic interpretation* in [22] is an interpretation $(\nu, \eta, (\varrho_R)_{R\in\sigma})$ such that $\nu(\overline{x}) := (\overline{x} = \overline{x})$ and $\eta(\overline{x}_1, \overline{x}_2) := (\overline{x}_1 = \overline{x}_2)$. (For instance the identity interpretation defined above is a basic interpretation.)

Note that every basic interpretation $\mathsf{I}\colon \tau \to \sigma$ induces a homomorphism

$$H(\mathsf{I})\colon\ \mathcal{B}_\sigma \to \mathcal{B}_\tau \qquad \text{defined by } H(\mathsf{I})([\varphi]) = [L(\mathsf{I})(\varphi)],$$

where $[\varphi]$ denotes the class of φ for logical equivalence. The mapping H is actually a contravariant functor from the category of interpretations to the category of Boolean algebras.

By Stone duality theorem, the interpretation I also defines a continuous function

$$F(\mathsf{I})\colon\ S_\tau \to S_\sigma \qquad \text{defined by } F(\mathsf{I})(t) = \{\varphi\colon\ L(\mathsf{I})(\varphi) \in t\}.$$

Note that F is a covariant functor from the category of interpretations to the category of Stone spaces.

Finally, the interpretation I also defines a mapping

$$P(\mathsf{I})\colon\ \mathscr{R}\mathrm{el}(\tau) \to \mathscr{R}\mathrm{el}(\sigma)$$

as follows:

- The domain of $P(\mathsf{I})(\mathbf{A})$ is $\nu(\mathbf{A})/\eta$, that is all the η-equivalence classes of p-tuples in $\nu(\mathbf{A})$.
- For every relational symbol $R \in \sigma$ with arity k (and associated formula ϱ_R) we have

  $$P(\mathsf{I})(\mathbf{A}) \models R([\overline{v}_1], \dots, [\overline{v}_k]) \qquad \text{if and only if } \ \mathbf{A} \models \varrho_R(\overline{v}_1, \dots, \overline{v}_k).$$

  (Note that this does not depend on the choice of the representatives $\overline{v}_1, \dots, \overline{v}_k$ of the η-equivalence classes $[\overline{v}_1], \dots, [\overline{v}_k]$.)

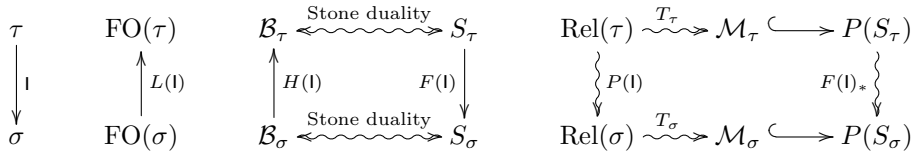
This mapping $P(\mathsf{I})\colon \mathrm{Rel}(\tau) \to \mathrm{Rel}(\sigma)$ is what is usually meant by an interpretation (of σ-structures in τ-structures, see [14]). It is easily checked that the mapping $P(\mathsf{I})$ has the property that for every formula $\varphi \in \mathrm{FO}(\sigma)$ with k free variables and every $[\overline{v}_1], \dots, [\overline{v}_k]$ in $P(\mathsf{I})(\mathbf{A})$ we have

$$P(\mathsf{I})(\mathbf{A}) \models \varphi([\overline{v}_1], \dots, [\overline{v}_k]) \qquad \text{if and only if } \ \mathbf{A} \models L(\mathsf{I})(\varphi)(\overline{v}_1, \dots, \overline{v}_k).$$

The interpretations, which we shall the most frequently consider, belong to the following types of basic interpretations (which are easily checked to be basic interpretations):

- *forgetful interpretations* that simply forget some of the relations;
- *renaming interpretations* that bijectively map a signature to another, mapping a relational symbol to a relational symbol with same arity;
- *projecting interpretations* that forget some symbols and rename others.

Our categorical approach allows us to obtain a more functorial point of view:



In this diagram the mapping $F(\mathsf{I})_*$ is the pushforward defined by $F(\mathsf{I})$ (see Lemma 3.2.1 bellow).

One can also consider the case where we do not consider all first-order formulas. Let X be a fragment of $\mathrm{FO}(\tau)$ and let $Y = L(\mathsf{I})^{-1}(X)$. (Note that if X is closed by disjunction, conjunction and negation, so is $L(\mathsf{I})^{-1}(X)$.) The basic interpretation I then defines a homomorphism

$$H(\mathsf{I})\colon\ \mathcal{B}_\sigma^Y \to \mathcal{B}_\tau^X,$$

which is the restriction of $H(\mathsf{I})$ to $\mathcal{B}_\sigma^Y$. By duality, this homomorphism defines a continuous mapping

$$\widehat{F}(\mathsf{I})\colon\ S_\tau^X \to S_\sigma^Y.$$

In particular, if K_φ is the clopen subset of S_σ^Y defined by $\varphi \in Y$ then $\widehat{F}(\mathsf{I})^{-1}(K_\varphi)$ is the clopen subset $K_{L(\mathsf{I})(\varphi)}$ of S_τ^X. Note that we have $F(\mathsf{I}) = \widehat{F}(\mathsf{I}) \circ \Pi_\tau^X$, where Π_τ^X is the natural projection from S_τ to S_τ^X.

3.2 Metric properties of interpretations. We have seen in the previous section that interpretations define continuous functions between Stone spaces. This property can be used to transfer convergence from one signature to another. This is done in a very general setting we introduce now.

Let $\mathsf{I}\colon \tau \to \sigma$ be a basic interpretation, let X be a fragment of $\mathrm{FO}(\tau)$, let $Y = L(\mathsf{I})^{-1}(X)$, let $\mathfrak{X}$ be a chain covering of X, and let $\mathfrak{Y}$ be a chain covering of Y such that every formula in $L(\mathsf{I})(Y_n)$ is logically equivalent to a formula in X_n.

Let us explain this choice of Y.

In $\operatorname{Im} P(\mathsf{I})$ we should not distinguish two finite structures $P(\mathsf{I})(\mathbf{A})$ and $P(\mathsf{I})(\mathbf{B})$ if there exist a chain $\mathbf{C}_1, \dots, \mathbf{C}_{2n+1}$ of finite structures such that $\mathbf{C}_1 = \mathbf{A}$, $\mathbf{C}_{2n+1} = \mathbf{B}$, $C_{2i-1} \equiv_X \mathbf{C}_{2i}$, and $P(\mathsf{I})(\mathbf{C}_{2i}) = P(\mathsf{I})(\mathbf{C}_{2i+1})$ for $i = 1, \dots, n$. But $P(\mathsf{I})(\mathbf{A}') = P(\mathsf{I})(\mathbf{B}')$ holds if and only if $\langle \varphi, \mathbf{A}' \rangle = \langle \varphi, \mathbf{B}' \rangle$ for every $\varphi \in$

$L(\mathsf{I})(\mathrm{FO}(\sigma))$. Hence the conditions can be rewritten as

$$\begin{aligned} &\forall\, \varphi \in X && \langle \varphi, \mathbf{C}_{2i-1}\rangle = \langle \varphi, \mathbf{C}_{2i}\rangle,\\ &\forall\, \varphi \in L(\mathsf{I})(\mathrm{FO}(\sigma)) && \langle \varphi, \mathbf{C}_{2i}\rangle = \langle \varphi, \mathbf{C}_{2i+1}\rangle. \end{aligned}$$

A necessary (but maybe not sufficient) condition is obviously that

$$\forall\, \varphi \in X \cap L(\mathsf{I})(\mathrm{FO}(\sigma)) \qquad \langle \varphi, \mathbf{A}\rangle = \langle \varphi, \mathbf{B}\rangle,$$

that is:

$$\forall\, \varphi \in L(\mathsf{I})^{-1}(X) \qquad \langle \varphi, P(\mathsf{I})(\mathbf{A})\rangle = \langle \varphi, P(\mathsf{I})(\mathbf{B})\rangle,$$

which we can rewrite as $P(\mathsf{I})(\mathbf{A}) \equiv_Y P(\mathsf{I})(\mathbf{B})$. This shows that the fragment Y is sufficiently small to ensure the continuity of $P(\mathsf{I})$. By our choice of the chain covering $\mathfrak{Y}$ we further get that $P(\mathsf{I})$ induces a *short map* (that is a 1-Lipschitz function). We summarize this in the following lemma.

Lemma 3.2.1. *In the above setting and notation we have:*

$$\mathrm{dist}_{\mathfrak{Y}}(P(\mathsf{I})(\mathbf{A}), P(\mathsf{I})(\mathbf{B})) \le \mathrm{dist}_{\mathfrak{X}}(\mathbf{A}, \mathbf{B}). \tag{7}$$

This can be restated as follows: Let the quotients $(\mathrm{Rel}(\tau)/\equiv_X, \mathrm{dist}_{\mathfrak{X}})$ *and* $(\mathrm{Rel}(\sigma)/\equiv_Y, \mathrm{dist}_{\mathfrak{Y}})$ *be the quotient metric spaces induced by the pseudometric spaces* $(\mathrm{Rel}(\tau), \mathrm{dist}_{\mathfrak{X}})$ *and* $(\mathrm{Rel}(\sigma), \mathrm{dist}_{\mathfrak{Y}})$*. Then the unique continuous function* $\widehat{P}(\mathsf{I})\colon (\mathrm{Rel}(\tau)/\equiv_X, \mathrm{dist}_{\mathfrak{X}}) \to (\mathrm{Rel}(\sigma)/\equiv_Y, \mathrm{dist}_{\mathfrak{Y}})$ *such that* $\widehat{P}(\mathsf{I})([\mathbf{A}]_X) = [P(\mathsf{I})(\mathbf{A})]_Y$ *is a short map.*

Proof: For every pair $\mathbf{A}, \mathbf{B}$ of τ-structures we have

$$\begin{aligned} \mathrm{dist}_{\mathfrak{Y}}(P(\mathsf{I})(\mathbf{A}), P(\mathsf{I})(\mathbf{B})) &= \inf_{n\in\mathbb{N}}\Big\{\max\Big(\frac{1}{n}, \max_{\varphi\in Y_n} |\langle \varphi, P(\mathsf{I})(\mathbf{A})\rangle - \langle \varphi, P(\mathsf{I})(\mathbf{A})\rangle|\Big)\Big\}\\ &= \inf_{n\in\mathbb{N}}\Big\{\max\Big(\frac{1}{n}, \max_{\psi\in L(\mathsf{I})(Y_n)} |\langle \varphi, \mathbf{A}\rangle - \langle \varphi, \mathbf{B}\rangle|\Big)\Big\}\\ &\le \inf_{n\in\mathbb{N}}\Big\{\max\Big(\frac{1}{n}, \max_{\psi\in X_n} |\langle \varphi, \mathbf{A}\rangle - \langle \varphi, \mathbf{B}\rangle|\Big)\Big\}\\ &= \mathrm{dist}_{\mathfrak{X}}(\mathbf{A}, \mathbf{B}). \end{aligned}$$

(In particular $\mathbf{A} \equiv_X \mathbf{B}$ imply that $\mathrm{dist}_{\mathfrak{Y}}(P(\mathsf{I})(\mathbf{A}), P(\mathsf{I})(\mathbf{B})) = 0$ thus $P(\mathsf{I})(\mathbf{A}) \equiv_Y P(\mathsf{I})(\mathbf{B})$ hence $P(\mathsf{I})$ descends to the quotient and there exists a unique map

$$\widehat{P}(\mathsf{I})\colon\ (\mathrm{Rel}(\tau)/\equiv_X, \mathrm{dist}_{\mathfrak{X}}) \to (\mathrm{Rel}(\sigma)/\equiv_Y, \mathrm{dist}_{\mathfrak{Y}})$$

such that $\widehat{P}(\mathsf{I})([\mathbf{A}]_X) = [P(\mathsf{I})(\mathbf{A})]_Y$.) □

Let $\mathscr{M}_{X,\mathsf{I}} = F(\mathsf{I})_*(\mathscr{M}_\tau^X)$ and let $\overline{\mathscr{M}_{X,\mathsf{I}}} = F(\mathsf{I})_*(\overline{\mathscr{M}_\tau^X})$ be the closure of $\mathscr{M}_{X,\mathsf{I}}$ in $(P(S_\sigma^Y), \mathrm{d}_{\mathfrak{Y}}^{\mathrm{LP}})$. We tried to summarize in Figure 3 the relations between the different (pseudo)metric spaces defined from signatures, fragments, and interpretations.

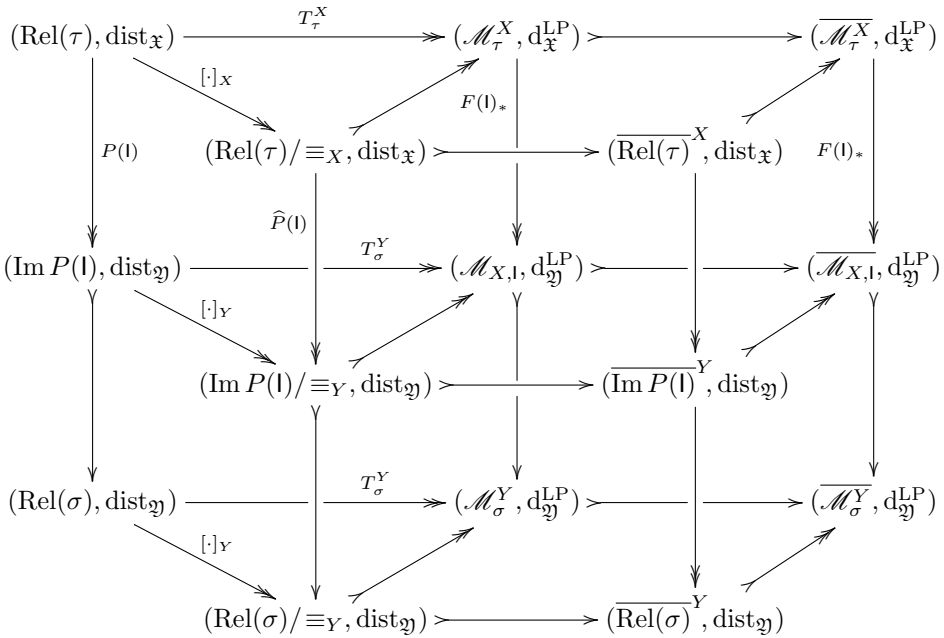


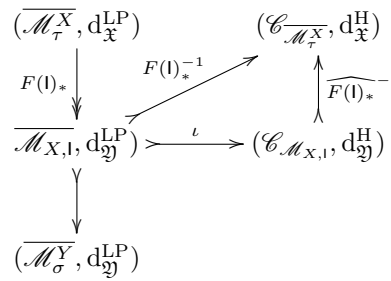
FIGURE 3. The considered (pseudo)metric spaces and their relations. Unlabeled arrows correspond to inclusions $\rightarrowtail$ or isometric embeddings $\twoheadrightarrow$. In this diagram the space $(\overline{\mathrm{Im}\, P(\mathsf{I})}^Y, \mathrm{dist}_{\mathfrak{Y}})$ is the completion of the pseudometric space $(\mathrm{Im}\, P(\mathsf{I}), \mathrm{dist}_{\mathfrak{Y}})$.

3.3 Lift-Hausdorff convergence. We now show how all the above constructions will nicely fit in Definition 3.3.1 of the lift-Hausdorff convergence. We first show how the definition derives from the preceding notions dealing with general basic interpretations.

Let $\mathsf{I}\colon \tau \to \sigma$ be a fixed interpretation, let X be a fixed fragment of $\mathrm{FO}(\tau)$, let $\mathfrak{X}$ be a fixed cover chain of X, and let $\mathfrak{Y}$ be any chain covering of Y such that every formula in $L(\mathsf{I})(Y_n)$ is logically equivalent to a formula in X_n.

According to Lemma 2.2.1 the pushforward mapping $F(\mathsf{I})_*$ is a continuous function from $P(S^X_\tau)$ to $P(S^Y_\sigma)$. Then the Lévy–Prokhorov distance $\mathrm{d}^{\mathrm{LP}}_{\mathfrak{X}}$ on $\overline{\mathscr{M}^X_\tau}$ defines a Hausdorff distance $\mathrm{d}^{\mathrm{H}}_{\mathfrak{X}}$ on the space $\mathscr{C}_{\overline{\mathscr{M}^X_\tau}}$ of nonempty closed subsets of $\overline{\mathscr{M}^X_\tau}$:

$$\mathrm{d}^{\mathrm{H}}_{\mathfrak{X}}(M_1, M_2) = \max\Big(\sup_{\mu_1 \in M_1} \inf_{\mu_2 \in M_2} \mathrm{d}^{\mathrm{LP}}_{\mathfrak{X}}(\mu_1, \mu_2), \sup_{\mu_2 \in M_2} \inf_{\mu_1 \in M_1} \mathrm{d}^{\mathrm{LP}}_{\mathfrak{X}}(\mu_1, \mu_2)\Big). \tag{8}$$



Also the pseudometric $\mathrm{dist}_{\mathfrak{X}}$ on $\mathrm{Rel}(\tau)$ defines a Hausdorff pseudometric $\mathrm{dist}_{\mathfrak{X}}^{\mathrm{H}}$ on the space of nonempty closed subsets of $\mathrm{Rel}(\tau)$ (for the topology induced by the pseudometric $\mathrm{dist}_{\mathfrak{X}}$):

(9)
$$\mathrm{dist}_{\mathfrak{X}}^{\mathrm{H}}(F_1, F_2) = \max\Big(\sup_{\mathbf{A}_1 \in F_1} \inf_{\mathbf{A}_2 \in F_2} \mathrm{dist}_{\mathfrak{X}}(\mathbf{A}_1, \mathbf{A}_2), \sup_{\mathbf{A}_2 \in F_2} \inf_{\mathbf{A}_1 \in F_1} \mathrm{dist}_{\mathfrak{X}}(\mathbf{A}_1, \mathbf{A}_2) \Big).$$

These (pseudo)metrics are related by the following equation (where F_1 and F_2 denote nonempty closed subsets of $\mathrm{Rel}(\tau)$:

$$\mathrm{dist}_{\mathfrak{X}}^{\mathrm{H}}(F_1, F_2) = \mathrm{d}_{\mathfrak{X}}^{\mathrm{H}}(T_\tau^X(F_1), T_\tau^X(F_2)).$$

Using the injective mapping $F(\mathsf{I})_*^{-1}$ we can transfer to $\mathscr{M}_{X,\mathsf{I}}$ the Hausdorff distance $\mathrm{d}_{\mathfrak{X}}^{\mathrm{H}}$ defined on $\mathscr{C}_{\overline{\mathscr{M}_\tau^X}}$, thus defining a distance $\mathrm{d}_{\mathfrak{X},\mathsf{I}}^{\mathrm{H}}$ on $\mathscr{M}_{X,\mathsf{I}}$:

$$\mathrm{d}_{\mathfrak{X},\mathsf{I}}^{\mathrm{H}}(\mu_1, \mu_2) = \mathrm{d}_{\mathfrak{X}}^{\mathrm{H}}(F(\mathsf{I})_*^{-1}(\mu_1), F(\mathsf{I})_*^{-1}(\mu_2)).$$

(Note that this metric usually does not define the same topology as $d_{\mathfrak{Y}}^{\mathrm{LP}}$.)

Using the mapping T_σ^Y we can transfer to $\mathrm{Im}\, P(\mathsf{I})$ the metric $\mathrm{dist}_{\mathfrak{X},\mathsf{I}}^{\mathrm{H}}$ just defined on $\mathscr{M}_{X,\mathsf{I}}$. As T_σ^Y is not injective in general we get this way a pseudometric $\mathrm{dist}_{\mathfrak{X},\mathsf{I}}^{\mathrm{H}}$ on $\mathrm{Im}\, P(\mathsf{I})$:

$$\begin{aligned}\mathrm{dist}_{\mathfrak{X},\mathsf{I}}^{\mathrm{H}}(\mathbf{A}, \mathbf{B}) &= \mathrm{d}_{\mathfrak{X}}^{\mathrm{H}}(F(\mathsf{I})_*^{-1}(T_\sigma^Y(\mathbf{A})), F(\mathsf{I})_*^{-1}(T_\sigma^Y(\mathbf{B}))) \\ &= \mathrm{d}_{\mathfrak{X}}^{\mathrm{H}}(T_\tau^X(P(\mathsf{I})^{-1}(\mathbf{A})), T_\tau^X(P(\mathsf{I})^{-1}(\mathbf{B}))).\end{aligned}$$

Hence we have

$$\mathrm{dist}_{\mathfrak{X},\mathsf{I}}^{\mathrm{H}}(\mathbf{A}, \mathbf{B}) = \mathrm{dist}_{\mathfrak{X}}^{\mathrm{H}}(P(\mathsf{I})^{-1}(\mathbf{A}), P(\mathsf{I})^{-1}(\mathbf{B})). \tag{10}$$

The situation is summarized in the following diagram:

$$\begin{array}{ccc} (\mathscr{C}_{\overline{\mathscr{M}^X_\tau}}, \mathrm{d}^{\mathrm{H}}_{\mathfrak{X}}) & \xleftarrow{T^X_\tau} & (\mathscr{C}_{\overline{\mathrm{Rel}(\tau)}}, \mathrm{dist}^{\mathrm{H}}_{\mathfrak{X}}) \\ \uparrow{\scriptstyle F(\mathsf{I})^{-1}_*} & & \uparrow{\scriptstyle P(\mathsf{I})^{-1}} \\ (\overline{\mathscr{M}_{X,\mathsf{I}}}, \mathrm{d}^{\mathrm{H}}_{\mathfrak{X},\mathsf{I}}) & \xleftarrow{T^Y_\sigma} & (\operatorname{Im} P(\mathsf{I}), \mathrm{dist}^{\mathrm{H}}_{\mathfrak{X},\mathsf{I}}) \end{array}$$

It follows from Lemma 3.2.1 that for every $\mathbf{A}, \mathbf{B} \in \operatorname{Im} P(\mathsf{I})$ we have

$$\mathrm{dist}^{\mathrm{H}}_{\mathfrak{X},\mathsf{I}}(\mathbf{A}, \mathbf{B}) \geq \mathrm{dist}_{\mathfrak{Y}}(\mathbf{A}, \mathbf{B}). \tag{11}$$

(In particular the topology defined by the pseudometric $\mathrm{dist}_{\mathfrak{X},\mathsf{I}}$ is finer that the topology defined by the pseudometric $\mathrm{dist}_{\mathfrak{Y}}$.)

Our basic notion of convergence with respect to an interpretation is the following (which we sketched in the introduction):

Definition 3.3.1 (Lift–Hausdorff convergence)**.** Let $\mathsf{I}\colon \tau \to \sigma$ be a basic interpretation and let X be a fragment of $\mathrm{FO}(\tau)$. A sequence $\mathbf{A}$ of finite σ-structures in $\operatorname{Im} P(\mathsf{I})$ is $\mathsf{I}_*(X)$*-convergent* if for every X-convergent subsequence $\mathbf{B}_f$ of lifts of $\mathbf{A}$ (meaning $P(\mathsf{I})(\mathbf{B}_f) = \mathbf{A}_f$) there exists an X-convergent sequence $\mathbf{C}$ of lifts of $\mathbf{A}$ extending $\mathbf{B}_f$ (i.e. such that $P(\mathsf{I})(\mathbf{C}) = \mathbf{A}$ and $\mathbf{C}_f = \mathbf{B}_f$).

We refer to the general notion of $\mathsf{I}_*(X)$-convergence as *lift-Hausdorff convergence* (or simply *lift convergence*).

This convergence admits an alternative equivalent definition, which justifies the term of "lift-Hausdorff convergence":

Theorem 3.3.2 (Metrization)**.** *Let $\mathfrak{X}$ be an arbitrary cover chain of X, and let $\mathfrak{Y}$ be an arbitrary chain covering of $Y = L(\mathsf{I})^{-1}(X)$ such that every formula in $L(\mathsf{I})(Y_n)$ is logically equivalent to a formula in X_n.*

Then a sequence $\mathbf{A}$ of σ-structures in $\operatorname{Im} P(\mathsf{I})$ is $\mathsf{I}_*(X)$-convergent if and only if it is Cauchy for $\mathrm{dist}^{\mathrm{H}}_{\mathfrak{X},\mathsf{I}}$.

Proof: We consider the two implications.

First assume that the sequence $\mathbf{A}$ of σ-structures in $\operatorname{Im} P(\mathsf{I})$ is $\mathsf{I}_*(X)$-convergent and let $\mathbf{B}_f$ be an X-convergent subsequence such that $P(\mathsf{I})(\mathbf{B}_f) = \mathbf{A}_f$. For every positive integer m, let $N(m)$ be minimum integer such that $f(N(m)) \geq m$. Let $\mathbf{C}_m$ be a τ-structure in $P(\mathsf{I})^{-1}(\mathbf{A}_{N(m)})$ such that $\mathrm{dist}_{\mathfrak{X}}(\mathbf{C}_m, \mathbf{B}_{f(N(m))})$ is minimum. Note that the minimum is attained as $P(\mathsf{I})^{-1}(\mathbf{A}_{N(m)})$ is compact. By definition we have

$$\mathrm{dist}_{\mathfrak{X}}(\mathbf{C}_m, \mathbf{B}_{f(N(m))}) \leq \mathrm{dist}_{\mathfrak{X},\mathsf{I}}(\mathbf{A}_m, \mathbf{A}_{f(N(m))}).$$

As $\mathbf{A}$ is Cauchy for $\mathrm{dist}_{\mathfrak{X},\mathsf{I}}$ and $\mathbf{B}_f$ is Cauchy for $\mathrm{dist}_{\mathfrak{X}}$ it directly follows that $\mathbf{C}$ is Cauchy for $\mathrm{dist}_{\mathfrak{X}}$, i.e. that $\mathbf{C}$ is X-convergent.

We now consider the other direction. Assume that for every X-convergent subsequence $\mathbf{B}_f$ such that $P(\mathsf{I})(\mathbf{B}_f) = \mathbf{A}_f$ there exists a sequence $\mathbf{C}$ such that $\mathbf{C}_f = \mathbf{B}_f$ and $P(\mathsf{I})(\mathbf{C}) = \mathbf{A}$, and assume for contradiction that the sequence $\mathbf{A}$ is not $\mathsf{I}_*(X)$-convergent. Then there exists $\alpha > 0$, such that for every integer N there exist integers $n, m > N$ and $\mathbf{B}_m \in P(\mathsf{I})^{-1}(\mathbf{A}_m)$ such that for every $\mathbf{C}_n \in P(\mathsf{I})^{-1}(\mathbf{A}_n)$ we have $\mathrm{dist}_{\mathfrak{X}}(\mathbf{B}_m, \mathbf{C}_n) > \alpha$. This allows to construct subsequence $\mathbf{B}_f$ and $\mathbf{C}_g$ (where $(f(i), g(i))$ correspond to a pair of admissible values of m and n with $\min(m,n) > \max(f(i-1), g(i-1))$. Moreover, we can assume that $\mathbf{B}_f$ is X-convergent. By assumption the subsequence $\mathbf{B}_f$ can be extended into a full X-convergent sequence, which we (still) denote by $\mathbf{B}$ such that $P(\mathsf{I})(\mathbf{B}) = \mathbf{A}$. In particular, there exist some N such that for every $n, m > N$ we have $\mathrm{dist}_{\mathfrak{X}}(\mathbf{B}_n, \mathbf{B}_m) < \alpha$. In particular, $\mathrm{dist}_{\mathfrak{X}}(\mathbf{B}_{f(n)}\mathbf{B}_{g(n)}) < \alpha$, what contradicts the minimality hypothesis on $\mathrm{dist}_{\mathfrak{X}}(\mathbf{B}_{f(n)}, \mathbf{C}_{g(n)})$. □

Note that Theorem 3.3.2 clearly states that the property of a sequence to be Cauchy for $\mathrm{dist}^{\mathrm{H}}_{\mathfrak{X},\mathsf{I}}$ is independent of the particular choice of the chain coverings $\mathfrak{X}$ and $\mathfrak{Y}$.

Theorem 3.3.3 (Representation)**.** *The $\mathsf{I}_*(X)$-limit of a sequence of an $\mathsf{I}_*(X)$-convergent sequence can be uniquely represented by means of a nonempty compact subset of $\overline{\mathscr{M}_\tau^X}$.*

Proof: Let $\mathbf{A}$ be an $\mathsf{I}_*(X)$-convergent sequence. Let $Y = L(\mathsf{I})^{-1}(X)$. We fix a cover chain $\mathfrak{X}$ of X and a cover chain $\mathfrak{Y}$ of Y such that every formula in $L(\mathsf{I})(Y_n)$ is logically equivalent to a formula in X_n.

According to Lemma 2.2.1 the pushforward mapping $F(\mathsf{I})_*$ is a continuous function from $P(S_\tau^X)$ to $P(S_\sigma^Y)$ hence for every $\mathbf{A}$ the set

$$\mathscr{L}(\mathbf{A}) = \{T_\tau^X(\mathbf{A}^+)\colon\ P(\mathsf{I})(\mathbf{A}^+) = \mathbf{A}\}$$

is a nonempty closed (hence compact) subset of $\overline{\mathscr{M}_\tau^X}$. According to Theorem 3.3.2, the $\mathsf{I}_*(X)$-convergence of $\mathbf{A}$ is equivalent to the convergence of $\mathbf{A}$ according to $\mathrm{dist}^{\mathrm{H}}_{\mathfrak{X},\mathsf{I}}$ metric. As noticed in beginning of Section 3.3 we have

$$\mathrm{dist}^{\mathrm{H}}_{\mathfrak{X},\mathsf{I}}(\mathbf{A},\mathbf{B}) = \mathrm{d}^{\mathrm{H}}_{\mathfrak{X}}(T_\tau^X(P(\mathsf{I})^{-1}(\mathbf{A})), T_\tau^X(P(\mathsf{I})^{-1}(\mathbf{B})) = \mathrm{d}^{\mathrm{H}}_{\mathfrak{X}}(\mathscr{L}(\mathbf{A}), \mathscr{L}(\mathbf{B})).$$

Thus, as $\mathrm{d}^{\mathrm{H}}_{\mathfrak{X}}$ is the Hausdorff distance on the space $\mathscr{C}_{\overline{\mathscr{M}_\tau^X}}$ of nonempty closed subsets of $\overline{\mathscr{M}_\tau^X}$ defined by the Lévy–Prokhorov distance $\mathrm{d}^{\mathrm{LP}}_{\mathfrak{X}}$ on $\overline{\mathscr{M}_\tau^X}$, the $\mathsf{I}_*(X)$-convergence of $\mathbf{A}$ is equivalent to the convergence of the sequence $\mathscr{L}(\mathbf{A}) = (\mathscr{L}(\mathbf{A}_n))_{n\in\mathbb{N}}$ in the Hausdorff sense.

It follows that the limit of $\mathbf{A}$ can be represented uniquely by the Hausdorff limit of $\mathscr{L}(\mathbf{A})$, which is a nonempty compact subset of $\overline{\mathscr{M}_\tau^X}$. □

This lemma gives an easy proof of the following result.

Proposition 3.3.4. *Let $\mathcal{C}$ be a class of σ structures, let $\mathsf{I}\colon \tau \to \sigma$ be an interpretation, and let X, Y be fragments of* $\mathrm{FO}(\tau)$.

If X-convergence implies Y-convergence in the class $\mathcal{D} = \{\mathbf{B}\colon P(\mathsf{I})(\mathbf{B}) \in \mathcal{C}\}$ of τ-structures then $\mathsf{I}_*(X)$-convergence implies $\mathsf{I}_*(Y)$-convergence in the class $\mathcal{C}$.

Proof: Let $\mathbf{A}$ be an $\mathsf{I}_*(X)$-convergent sequence of τ-structures in $\mathcal{C}$ and let $\mathbf{B}_f$ be a Y-converging subsequence of τ-structures (in $\mathcal{D}$) such that $P(\mathsf{I})(\mathbf{B}_{f(n)}) = \mathbf{A}_{f(n)}$. Let $\mathbf{B}_{g\circ f}$ be an X-converging subsequence of $\mathbf{B}_f$. By Definition 3.3.1 there exists an X-convergent sequence $\mathbf{C}$ such that $\mathbf{C}_{g\circ f} = \mathbf{B}_{g\circ f}$ and $P(\mathsf{I})(\mathbf{C}) = \mathbf{A}$ (hence $\mathbf{C}$ is in $\mathcal{D}$). As X-convergence implies Y-convergence on $\mathcal{D}$ the sequence $\mathbf{D}$ is Y convergent, and has the same Y-limit as the Y-convergent sequence $\mathbf{B_f}$ as they share infinitely many elements. It follows that the sequence $\mathbf{D}$ defined by

$$\mathbf{D}_n = \begin{cases} \mathbf{B}_n & \text{if } (\exists\, i)\ n = f(i), \\ \mathbf{C}_n & \text{otherwise,} \end{cases}$$

has the property that $\mathbf{D}_f = \mathbf{B}_f$ and $P(\mathsf{I})(\mathbf{D}) = \mathbf{A}$. We deduce that $\mathbf{A}$ is $\mathsf{I}_*(Y)$-convergent. □

Here are some more remarks indicating convenient properties of $\mathsf{I}_*(X)$-convergence.

First note that if $\mathsf{I}\colon \sigma \to \sigma$ is the identity interpretation, then $\mathrm{dist}_{\mathfrak{X},\mathsf{I}} = \mathrm{dist}_{\mathfrak{X}}$ and $\mathsf{I}_*(X)$-convergence is the same as X-convergence. Also, we have that every sequence $\mathbf{A}$ in $\operatorname{Im} P(\mathsf{I})$ has an $\mathsf{I}_*(X)$-convergent subsequence. Finally, let us remark that for every $\mathsf{I}\colon \tau \to \sigma$, $\mathsf{I}_*(\mathrm{FO})$-convergence implies FO-convergence.

Let $\widehat{\tau}$ be the signature obtained from τ by duplicating each relation symbol countably many times, which we denote by $\widehat{\tau} = \mathbb{N}\tau$. To each symbol $R \in \tau$ correspond the symbols R^i in $\widehat{\tau}$ for $i \in \mathbb{N}$. We define the interpretation I_i obtained from I by replacing relations R by R^i (I_i is a clone of I based on the relations R^i).

Proposition 3.3.5 (Almost $\mathsf{I}_*(X)$-limit probability measure)**.** *Let $\mathbf{A}$ be an $\mathsf{I}_*(X)$-convergent sequence of finite σ-structures.*

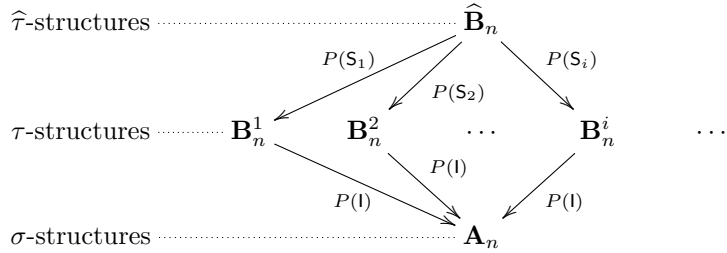
There exists a probability measure $\widehat{\mu} \in \overline{\mathscr{M}_{\widehat{\tau}}^X}$ such that for every $\varepsilon > 0$ and for every $\mathbf{C}$ such that $P(\mathsf{I})(\mathbf{C}) = \mathbf{A}$ there exists $i \in \mathbb{N}$ such that

$$\mathrm{d}_{\mathfrak{X}}^{LP}\big(F(\mathsf{I}_i)_*(\widehat{\mu}), \lim_n T_\tau^X(\mathbf{C}_n)\big) < \varepsilon,$$

where $\lim_n$ stands for the weak limit of probability measures.

Proof: For $i \in \mathbb{N}$ we choose $\mathbf{B}_n^i$ such that $P(\mathsf{I})(\mathbf{B}_n^i) = \mathbf{A}_n$. We construct the $\widehat{\tau}$-structure $\widehat{\mathbf{B}}_n$ by amalgamating all the relations of all the $\mathbf{B}_n^i$. We denote by S_i

the interpreting projection $\widehat{\mathbf{B}}_n \mapsto \mathbf{B}_n^i$. Note that $\mathsf{I}_i = \mathsf{I} \circ \mathsf{S}_i$. Then we have



Then we consider an X-convergent subsequence $\widehat{\mathbf{B}}_f$ of $\widehat{\mathbf{B}}$, the limit of which is represented by the probability measure $\widehat{\mu} \in \mathscr{M}_{\widehat{\tau}}^X$. The measure $\widehat{\mu}$ has obviously the claimed property. □

3.4 Local-global convergence. In this section we show how the abstract framework of Section 3.3 provides a proper setting for local-global convergence.

The notion of local-global convergence of graphs with bounded degrees has been introduced by B. Bollobás and O. Riordan in [6] based on a colored neighborhood metric. In [12], H. Hatami, L. Lovász, and B. Szegedy gave the following equivalent definition:

Definition 3.4.1 ([12])**.** A graph sequence $\mathbf{G}$ of graphs with maximum degree D is *local-global convergent* if for every $r, k \in \mathbb{N}$ and $\varepsilon > 0$ there is an index l such that if $n, m > l$, then for every coloring of the vertices of G_n with k colors, there is a coloring of the vertices of G_m with k colors such that the total variation distance between the distributions of colored neighborhoods of radius r in G_n and G_m is at most $\varepsilon > 0$.

The following is the principal result which relates local-global convergence to a lift-Hausdorff convergence.

Let us consider a fixed countable signature σ and the signature τ obtained from σ by adding countably many unary symbols. Thus $\sigma \subset \tau$. Let

$$\mathsf{Sh}\colon\ \sigma \to \tau$$

be the forgetful interpretation (Sh for "Shadow"). This means $\mathsf{Sh} = (\nu, \eta, (\varrho_R)_{R\in\sigma})$, where $\nu(x_1) := (x_1 = x_1)$, $\eta(x_1, x_2) := (x_1 = x_2)$, and $\varrho_R(x_1, \ldots, x_p) := R(x_1, \ldots, x_p)$ for $R \in \sigma$ with arity p. Then, for instance:

- for a τ-structure $\mathbf{A}$, the σ-structure $P(\mathsf{Sh})(\mathbf{A})$ is obtained from $\mathbf{A}$ by forgetting all unary relations in $\tau \setminus \sigma$;
- for a formula $\varphi \in \mathrm{FO}(\sigma)$, we have $L(\mathsf{Sh})(\varphi) = \varphi$;
- for $t \in S_\tau$ we have $F(\mathsf{Sh})(t) = t \cap \mathrm{FO}(\sigma)$.

By [20] we know that $\mathrm{FO}^{\mathrm{local}}$-convergence coincides with $\mathrm{FO}_1^{\mathrm{local}}$-convergence for graphs with bounded degree. By Proposition 3.3.4 the notions of $\mathsf{Sh}_*(\mathrm{FO}^{\mathrm{local}})$-convergence and $\mathsf{Sh}_*(\mathrm{FO}_1^{\mathrm{local}})$-convergence will also coincide for graphs with bounded degrees. These notions actually coincide with the notion of local-global convergence of graphs with bounded degrees:

Proposition 3.4.2. *Let* $\mathbf{G}$ *be a sequence of graphs with maximum degree* D*. Then the following are equivalent:*

(1) $\mathbf{G}$ *is local-global convergent;*
(2) $\mathbf{G}$ *is* $\mathsf{Sh}_*(\mathrm{FO}_1^{\mathrm{local}})$*-convergent;*
(3) $\mathbf{G}$ *is* $\mathsf{Sh}_*(\mathrm{FO}^{\mathrm{local}})$*-convergent.*

Proof: For classes of colored graphs with degree at most D, $\mathrm{FO}_1^{\mathrm{local}}$-convergence is equivalent to $\mathrm{FO}^{\mathrm{local}}$-convergence (see [20]). It follows from Proposition 3.3.4 that for these graphs $\mathsf{Sh}_*(\mathrm{FO}_1^{\mathrm{local}})$-convergence is equivalent to $\mathsf{Sh}_*(\mathrm{FO}^{\mathrm{local}})$-convergence. Thus we only have to prove the equivalence of local-global convergence and $\mathsf{Sh}_*(\mathrm{FO}_1^{\mathrm{local}})$-convergence.

We consider the fragment $X \subset \mathrm{FO}_1^{\mathrm{local}}$ of formulas consistent with the property of having maximum degree D. Consider a cover chain $\mathfrak{X} = (X_r)_{r\in\mathbb{N}}$ of X where X_r contains (one representative of the equivalence class of) each formula in X that is r-local and use only the r first unary predicates. (Note that $|X_r|$ is finite.)

It is easily checked that every r-local formula $\varphi \in X_r$ is equivalent (on graphs with maximum degree D) to a formula of the form $\bigvee_{B\in\mathcal{F}_\varphi} \zeta_{B,r}(x)$ where $\zeta_{B,r}(x)$ expresses that the ball of radius r rooted at x is isomorphic to the rooted graph B, and $\mathcal{F}_\varphi$ is a finite set of rooted graphs of radius at most r. It easily follows that the maximum of $|\langle\varphi, G_1\rangle - \langle\varphi, G_2\rangle|$ over $\varphi \in X_r$ equals the total variation distance of the distributions of r-balls in G_1 and G_2 where we consider only the r first colors, which we denote by $\mathrm{d}_{TV}^{(r)}(G_1, G_2)$. Then we have

$$\mathrm{dist}_{\mathfrak{X}}(G_1, G_2) = \inf_{r\in\mathbb{N}} \Big\{ \max\Big(\frac{1}{r}, \mathrm{d}_{TV}^{(r)}(G_1, G_2)\Big)\Big\}. \tag{12}$$

As one easily checks that $\mathrm{d}_{TV}(r')(G_1, G_2) \geq \mathrm{d}_{TV}(r)(G_1, G_2)$ if $r' \geq r$ we have that for every fixed integer r we have

$$\min\Big(\frac{1}{r}, \mathrm{d}_{TV}^{(r)}(G_1, G_2)\Big) \leq \mathrm{dist}_{\mathfrak{X}}(G_1, G_2) \leq \max\Big(\frac{1}{r}, \mathrm{d}_{TV}^{(r)}(G_1, G_2)\Big). \tag{13}$$

Now assume $\mathbf{G}$ is $\mathsf{Sh}_*(\mathrm{FO}_1^{\mathrm{local}})$-convergent. Let k, r be fixed integers. Then $\mathsf{Sh}_*(\mathrm{FO}_1^{\mathrm{local}})$-convergence of $\mathbf{G}$ easily implies the convergence of the lifts of G_n by k colors, which means that for every $\varepsilon > 0$ there is an index l such that if $n, m > l$, then for every coloring G_n^+ of G_n with k colors, there is a coloring G_m^+ of G_m with k colors such that $\mathrm{dist}_{\mathfrak{X}}(G_n^+, G_m^+) < \varepsilon$ hence by (13) the total variation distance between the distributions of colored neighborhoods of radius r in G_n and G_m is at most $\varepsilon > 0$, provided that $\varepsilon < 1/r$. Hence $\mathbf{G}$ is local-global convergent.

Assume $\mathbf{G}$ is local-global convergent. Then for every $\varepsilon > 0$, letting $r = \lceil 1/\varepsilon \rceil$, there exists an integer l such that if $n, m > l$, then for every coloring G_n^+ of G_n with r colors, there is a coloring G_m^+ of G_m with r colors such that the total variation distance between the distributions of colored neighborhoods of radius r in G_n and G_m is at most ε. Hence by (13) we have $\mathrm{dist}_{\mathfrak{X}}(G_n^+, G_m^+) < \max(\varepsilon, \mathrm{d}_{TV}^{(r)}(G_n, G_m)) \leq \varepsilon$. (Note that we do not need to use any of the colors with index greater than r.) It follows that $\mathbf{G}$ is $\mathsf{Sh}_*(\mathrm{FO}_1^{\mathrm{local}})$-convergent. □

Motivated by this theorem we can extend the definition of local-global convergence to general graphs and relational structures:

Definition 3.4.3 (Local-global convergence)**.** A sequence $\mathbf{A}$ is *local-global convergent* if it is $\mathsf{Sh}_*(\mathrm{FO}^{\mathrm{local}})$-convergent.

The weaker notion of $\mathsf{Sh}_*(\mathrm{FO}_1^{\mathrm{local}})$-convergence already implies convergence of some graph invariants in an interesting way. This is, for instance, the case of the Hall ratio $\varrho(G) = \alpha(G)/|G|$.

Proposition 3.4.4. *Let* $\mathbf{G}$ *be an* $\mathsf{Sh}_*(\mathrm{FO}_1^{\mathrm{local}})$*-convergent sequence of graphs. The Hall ratio* $\varrho(G_n) = \alpha(G_n)/|G_n|$ *converges.*

Proof: Let $a = \limsup \varrho(G_n)$. Let G_n^+ be obtained by marking (by M) a maximum independent set in G_n. (Thus $G_n = P(\mathsf{Sh})(G_n^+)$.) We extract a subsequence of $\mathbf{G^+}$ with limit measure of $M(G_n^+)$ equal to a, then an $\mathrm{FO}_1^{\mathrm{local}}$-convergent subsequence. According to the lifting property, this subsequence can be extended into a full sequence $\mathbf{G^*}$. Consider the formula

$$\psi(x) := M(x) \wedge (\exists\, y)\ M(y) \wedge \mathrm{Adj}(x, y).$$

Then $\psi(G_n^*)$ is the set of all marked vertices of G_n with a marked vertex in their neighborhood. Hence $\lim_{n\to\infty} \langle \psi, G_n^* \rangle = 0$ (as it converges to 0 on the subsequence where M marks an independent set). Moreover, $M(G_n^*) \setminus \psi(G_n^*)$ is independent. It follows that

$$a = \lim_{n\to\infty} \langle M, G_n^* \rangle = \lim_{n\to\infty} \frac{|M(G_n^*) \setminus \psi(G_n^*)|}{|G_n^*|} \leq \liminf \varrho(G_n) \leq \limsup \varrho(G_n) = a.$$

Hence $\varrho(G_n)$ converges. □

Let us add the following remarks: In such a context it is not possible to distinguish (at the limit) a maximal independent set from a near maximal independent set. Of course this does not change the property that $\alpha(G_n)/|G_n|$ converges nor the measure of the (near) maximal independent set found in the limit.

For the chromatic number, local-global convergence is clearly not strong enough, as witnessed by a local-global convergent sequence $\mathbf{G}$ of bipartite graphs modified by replacing G_n by the disjoint union of G_n and K_{100} for (say) half of the values of n. The obtained sequence is still local-global convergent but the chromatic numbers of G_n oscillate between 2 and 100. To ensure the convergence

of the chromatic number one needs at least $\mathsf{Sh}(\mathrm{FO}_0)$-convergence. However, with $\mathsf{Sh}(\mathrm{FO}_1^{\mathrm{local}})$-convergence it is possible to get the convergence of the minimum integer c such that the graphs G_n can be made c-colorable by removing $o(|G_n|)$ vertices.

We end this section by giving an example showing that not every graphing is a strong local-global limit of a sequence of finite graphs (for a proof that not every graphing is a weak local-global limit of a sequence of finite graphs, that is an answer to the problem posed in [12], see [17]).

Example 3.4.5. Consider the graphing $\mathbf{G}$ with domain $(\mathbb{R}/\mathbb{Z}) \times \{1, 2, \ldots, 6\}$, and edge set

$$\begin{aligned} E = &\{\{(x,1),(y,1)\},\{(x,6),(y,6)\}\colon\ x \in [0,1],\ y = x + \alpha \bmod 1\} \\ &\cup \{\{(x,i),(x,i+1)\}\colon\ x \in [0,1],\ 1 \le i < 5\} \\ &\cup \{\{(x,2),(x,4)\},\{(x,3),(x,5)\}\colon\ x \in [0,1]\} \end{aligned}$$

represented on Figure 4.

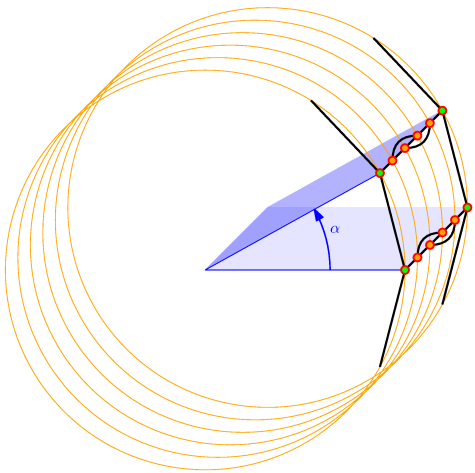
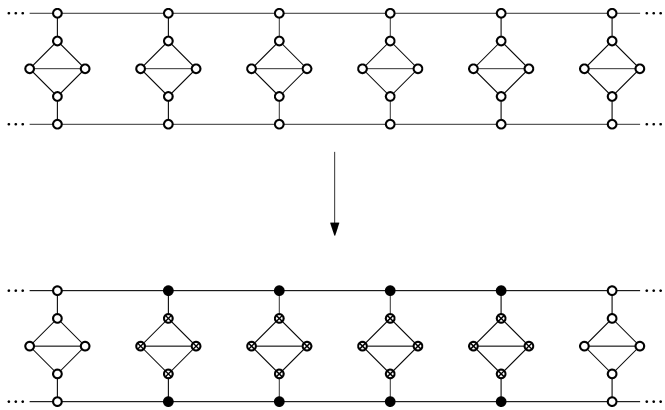


FIGURE 4. Example of a 3-regular graphing that is not a strong local-global limit.

Assume $\mathbf{G}$ is the strong local-global limit of a sequence $\mathbf{G} = (G_n)_{n\in\mathbb{N}}$ of graphs. Almost all neighborhoods in G_n (for n large) look the same: color crossed circle all vertices with two adjacent neighbors and color black all the vertices having exactly two non crossed circle neighbors.



Then, apart from a negligible set of vertices all the vertices are colored either as crossed circle or black. Moreover, almost all black vertices belong either to a long black cycle or a long black path. Recolor the black vertices in blue, green, and purple by dividing all these paths and cycles in almost equal parts (and taking care of globally balancing these colors).

Now consider a local convergent subsequence of the colored G_n. By definition this local subsequence can be extended into a local convergent sequence G^+ of colorings of the graphs in G. By local convergence, every connected component of the sub-graphing induced by green, blue, and violet vertices are monochromatic (apart from a 0 measure set), and these components are invariant by the transformation $(x, y) \mapsto (x + \alpha \bmod 1, y)$ (hence $y \in \{1, 6\}$). However, this sub-graphing has two ergodic components, each of measure 1/6 although each of the blue, green, and violet color contains asymptotically 1/9 of the vertices, a contradiction.

4. Applications

4.1 Clustering. Monadic lifts (i.e. lifts by unary relations) were considered in [24] in the context of continuous clustering of the structures in an $\mathrm{FO}^{\mathrm{local}}$-convergent sequence. One of the main results (see Theorem 4.1.1 bellow) expresses that every $\mathrm{FO}^{\mathrm{local}}$-convergent sequence has monadic lift tracing components while preserving $\mathrm{FO}^{\mathrm{local}}$-convergence. This will be refined in this section under the stronger assumption of $\mathsf{Sh}_*(\mathrm{FO}^{\mathrm{local}})$-convergence (see Theorem 4.1.6).

The analysis in [24] leads to interesting notions: *globular cluster* (corresponding to a limit nonzero measure connected component), *residual cluster* (corresponding to all the zero-measure connected components taken as a whole), and *negligible cluster* (corresponding to the stretched part connecting the other clusters, which eventually disappears at the limit).

Negligible sets intuitively correspond to parts of the graph one can remove, without a great modification of the statistics of the graph: A sequence $\mathsf{X} \subseteq \mathsf{A}$ is

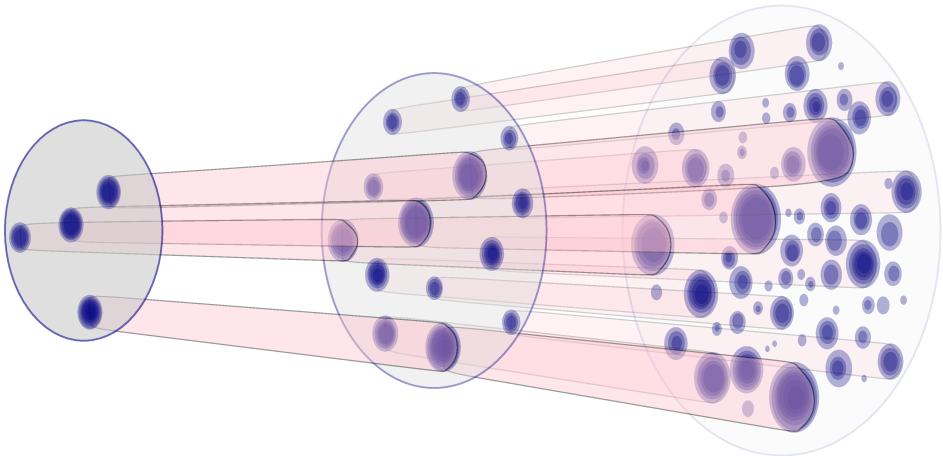


FIGURE 5. Typical shape of a structure continuously segmented by a clustering: dense spots correspond to globular clusters, and the background to the residual cluster. Biggest globular clusters appear first and then move apart from each other, while new (smaller) globular clusters appear and residual cluster becomes sparser and sparser.

negligible in a local-convergent sequence $\mathbf{A}$ if

$$\forall\, d \in \mathbb{N} \qquad \limsup_{n\to\infty} \nu_{\mathbf{A}_n}(\mathrm{N}^d_{\mathbf{A}_n}(X_n)) = 0.$$

This we simply formulate as $\forall\, d \in \mathbb{N}\ \limsup \nu_{\mathbf{A}}(\mathrm{N}^d_{\mathbf{A}}(\mathsf{X})) = 0$.

Two sequences X and Y of subsets are *equivalent* in $\mathbf{A}$ if the sequence $\mathsf{X} \,\Delta\, \mathsf{Y} = (X_n \,\Delta\, Y_n)_{n\in\mathbb{N}}$ is negligible in $\mathbf{A}$. This will be denoted by $\mathsf{X} \approx \mathsf{Y}$. We denote by $\mathsf{0}$ the sequence of empty subsets. Hence $\mathsf{X} \approx \mathsf{0}$ is equivalent to the property that X is negligible. We further define a partial order on sequences of subsets by $\mathsf{X} \preceq \mathsf{Y}$ if the sequence $\mathsf{X} \setminus \mathsf{Y} = (X_n \setminus Y_n)_{n\in\mathbb{N}}$ is negligible in $\mathbf{A}$. Hence "$\preceq$" has $\mathsf{0}$ for its minimum and $\mathsf{X} \approx \mathsf{Y}$ if and only if $\mathsf{X} \preceq \mathsf{Y}$ and $\mathsf{Y} \preceq \mathsf{X}$.

The notion of cluster of a local-convergent sequence is a weak analog of the notion of union of connected components, or more precisely of the topological notion of "clopen subset": A sequence X of subsets of a local-convergent sequence $\mathbf{A}$ is a *cluster* of $\mathbf{A}$ if the following conditions hold:

(1) the lifted sequence $\mathrm{L}_{\mathsf{X}}(\mathbf{A})$ obtained by marking set X_n in $\mathbf{A}_n$ by a new unary relation M_{X} is local-convergent;
(2) the sequence $\partial_{\mathbf{A}}\mathsf{X}$ is negligible in $\mathbf{A}$.

Condition (1) can be seen as a continuity requirement for the subset selection. Condition (2) is stronger than the usual requirement that there are not too many

connections leaving the cluster. We intuitively require that the (asymptotically arbitrarily large) ring around a cluster is a very sparse zone.

A cluster X is *atomic* if for every cluster Y of $\mathbf{A}$ such that $\mathsf{Y} \preceq \mathsf{X}$ either $\mathsf{Y} \approx \mathsf{0}$ or $\mathsf{Y} \approx \mathsf{X}$; the cluster X is *strongly atomic* if X_f is an atomic cluster of $\mathbf{A}_f$ for every increasing function $f\colon \mathbb{N} \to \mathbb{N}$. To the opposite, the cluster X is a *nebula* if for every increasing function $f\colon \mathbb{N} \to \mathbb{N}$, every atomic cluster Y_f of $\mathbf{A}_f$ with $\mathsf{Y}_f \subseteq \mathsf{X}_f$ is trivial (i.e. $\mathsf{Y}_f \approx \mathsf{0}$). Finally, a cluster X is *universal* for $\mathbf{A}$ if X is a cluster of every conservative lift of $\mathbf{A}$.

Two clusters X and Y of a local-convergent sequence $\mathbf{A}$ are *interweaving*, and we note $\mathsf{X} \between \mathsf{Y}$ if every sequence Z with $Z_n \in \{X_n, Y_n\}$ is a cluster of $\mathbf{A}$.

We say that two clusters C_1 and C_2 are

- *weakly disjoint* if $\mathsf{C}_1 \cap \mathsf{C}_2 \approx \mathsf{0}$;
- *disjoint* if $\mathsf{C}_1 \cap \mathsf{C}_2 = \mathsf{0}$;
- *strongly disjoint* if $(\mathrm{N}_{\mathbf{A}}(\mathsf{C}_1) \cap \mathsf{C}_2) \cup (\mathsf{C}_1 \cap \mathrm{N}_{\mathbf{A}}(\mathsf{C}_2)) = \mathsf{0}$.

A cluster C of a local-convergent sequence $\mathbf{A}$ is *globular* if for every $\varepsilon > 0$ there exists $d \in \mathbb{N}$ such that

$$\liminf_{n\to\infty} \sup_{v_n \in C_n} \nu_{\mathbf{A}_n}(\mathrm{N}^d_{\mathbf{A}_n}(v_n)) > (1-\varepsilon) \lim_{n\to\infty} \nu_{\mathbf{A}_n}(C_n).$$

In other words, a cluster C is globular if for every $\varepsilon > 0$ and sufficiently large n, ε-almost all elements of C_n are included in some ball of radius at most d in C_n for some fixed d. (Note that for a cluster C and $v_n \in C_n$, considering $\nu_{\mathbf{A}_n}(\mathrm{N}^d_{\mathbf{A}_n}(v_n))$ or $\nu_{\mathbf{A}_n}(\mathrm{N}^d_{\mathbf{A}_n[C_n]}(v_n))$ makes asymptotically no difference.) Every globular cluster is clearly strongly atomic, but the converse does not hold as witnessed, for instance, by sequence of expanders. The strongly atomic clusters that are not globular are called *open clusters*. Opposite to globular clusters are residual clusters: A cluster X of $\mathbf{A}$ is *residual* if for every $d \in \mathbb{N}$ it holds

$$\limsup_{n\to\infty} \sup_{v_n \in A_n} \nu_{\mathbf{A}_n}(\mathrm{N}^d_{\mathbf{A}_n}(v_n)) = 0.$$

Theorem 4.1.1 ([24])**.** *Let $\mathbf{A}$ be a local convergent sequence of σ-structures. Then there exists a signature σ^+ obtained from σ by the addition of countably many unary symbols M_R and $M_{i,j}$, $i \in \mathbb{N}$, $1 \le j \le N_i$, and a clustering $\mathbf{A}^+$ of $\mathbf{A}$ with the following properties:*

- *For every $i \in \mathbb{N}$, $\bigcup_{j=1}^{N_i} M_{i,j}(\mathbf{A})$ is a universal cluster.*
- *For every $i \in \mathbb{N}$ and every $1 \le j \le N_i$, $M_{i,j}(\mathbf{A})$ is a globular cluster.*
- *Two clusters $M_{i,j}(\mathbf{A})$ and $M_{i',j'}(\mathbf{A})$ are interweaving if and only if $i = i'$.*
- *$M_R(\mathbf{A})$ is a residual cluster.*

This structural theorem is assuming the local convergence of the sequence. If we assume local-global convergence we get stronger results (Theorem 4.1.6 bellow) involving expanding properties which we will define now. This is pleasing as the decomposition into expanders was one of the motivating examples [6] and [12].

The following is a sequential version of expansion property: A structure $\mathbf{A}$ is (d,ε,δ)-*expanding* if for every $X \subset A$ it holds

$$\varepsilon < \nu_{\mathbf{A}}(X) < 1-\varepsilon \quad \Longrightarrow \quad \nu_{\mathbf{A}}(\mathrm{N}^d_{\mathbf{A}}(X)) > (1+\delta)\nu_{\mathbf{A}}(X).$$

This condition may be reformulated as:

$$\inf\Big\{\frac{\nu_{\mathbf{A}}(\mathrm{N}^d_{\mathbf{A}}(X)\setminus X)}{\nu_{\mathbf{A}}(X)}\colon\ \varepsilon<\nu_{\mathbf{A}}(X)<1-\varepsilon\Big\}>\delta.$$

Note that the left hand size of the above inequality is similar to the *magnification* introduced in [3], which is the isoperimetric constant h_{out} defined by

$$h_{\mathrm{out}} = \inf\Big\{\frac{|\mathrm{N}_{\mathbf{A}}(X)\setminus X|}{|X|}\colon\ 0<\frac{|X|}{|A|}<\frac{1}{2}\Big\}.$$

A local-convergent sequence $\mathbf{A}$ is *expanding* if for every $\varepsilon>0$ there exist $d,t\in\mathbb{N}$ and $\delta>0$ such that every $\mathbf{A}_n$ with $n\geq t$ is (d,ε,δ)-expanding. A nontrivial cluster X of $\mathbf{A}$ is *expanding* of $\mathbf{A}$ if $\mathbf{A}[\mathsf{X}]$ is expanding. We have the following equivalent formulations of this concept:

Lemma 4.1.2 ([24])**.** *Let $\mathsf{X}\not\approx\mathsf{0}$ be a cluster of a local convergent sequence $\mathbf{A}$. The following conditions are equivalent:*

(1) *X is an expanding cluster of $\mathbf{A}$;*
(2) *for every $\varepsilon>0$ there exist $d,t\in\mathbb{N}$ such that for every $\mathsf{Z}\subseteq\mathsf{X}$ with $\nu_{\mathbf{A}}(\mathsf{Z})>\varepsilon\nu_{\mathbf{A}}(\mathsf{X})$ it holds*

$$\nu_{\mathbf{A}}(\mathrm{N}^d_{\mathbf{A}}(\mathsf{Z}))>(1-\varepsilon)\nu_{\mathbf{A}}(\mathsf{X});$$

(3) *the sequence X is a strongly atomic cluster of $\mathbf{A}$;*
(4) *for every $\varepsilon>0$ there exists no $\mathsf{Y}\subseteq\mathsf{X}$ such that $\partial_{\mathbf{A}}\mathsf{Y}\approx\mathsf{0}$ and*

$$\varepsilon<\liminf\nu_{\mathbf{A}}(\mathsf{Y})<\limsup\nu_{\mathbf{A}}(\mathsf{X})-\varepsilon.$$

Note that for local-global convergent sequences, the notions of atomic, strongly atomic, and expanding clusters are equivalent.

The case of bounded degree graphs is particularly interesting and our definitions capture this as well. Recall that a sequence $\mathbf{G}$ of graphs is a *vertex expander* if there exists $\alpha>0$ such that $\liminf h_{\mathrm{out}}(G_n)\geq\alpha$. (For more information on expanders we refer the reader to [15].)

Lemma 4.1.3 ([24])**.** *Let $\mathbf{G}$ be a sequence of graphs with maximum degree at most Δ and let $\mathsf{C}\not\approx\mathsf{0}$ be a cluster of $\mathbf{G}$. The following are equivalent:*

- *C is an expanding cluster;*
- *for every $\varepsilon>0$ there exists $\mathsf{X}\subseteq\mathsf{C}$ such that for every $n\in\mathbb{N}$ it holds $|X_n|<\varepsilon|C_n|$ and $\mathbf{G}[\mathsf{C}\setminus\mathsf{X}]$ is a vertex expander.*

We consider a fixed enumeration $\varphi_1, \varphi_2, \dots$ of $\mathrm{FO}^{\mathrm{local}}$. The *profile* $\mathrm{Prof}(\mathsf{C})$ of a cluster C is the sequence formed by $\lim \nu_{\mathbf{A}}(\mathsf{C})$ followed by the values $\lim\langle \varphi_i, \mathbf{A}[\mathsf{C}]\rangle$ for $i \in \mathbb{N}$. The lexicographic order on the profiles is denoted by "$\leq$".

In [24] it was proved that two expanding clusters are either weakly disjoint or interweaving. We now prove a lemma with similar flavor.

Lemma 4.1.4. *Let C_1 be an expanding cluster of a local-convergent sequence $\mathbf{A}$ and let C_2 be a cluster of $\mathbf{A}$.*

Then the limit set of $\nu_{\mathbf{A}}(\mathsf{C}_1 \cap \mathsf{C}_2)$ is included in $\{0, \lim \nu_{\mathbf{A}}(\mathsf{C}_1)\}$.

Proof: Let $\mathsf{X} = \mathsf{C}_1 \cap \mathsf{C}_2$. Assume for contradiction that there exists $0 < \alpha < \lim \nu_{\mathbf{A}}(\mathsf{C}_1)$ and a subsequence $\mathbf{A}_f$ such that $\mathrm{L}_{\mathsf{X}_f}(\mathbf{A}_f)$ is local convergent and $\lim \nu_{\mathbf{A}_f}(\mathsf{X}_f) = \alpha$. As $\delta_{\mathbf{A}}\mathsf{X} \subseteq \delta_{\mathbf{A}}\mathsf{C}_1 \cup \delta_{\mathbf{A}}\mathsf{C}_2$ we deduce that X_f is a cluster of $\mathbf{A}_f$. But $\mathsf{X}_f \subseteq (\mathsf{C}_1)_f$, $\mathsf{X}_f \not\approx \mathsf{0}$, and $\mathsf{X}_f \not\approx \mathsf{C}_1$ (as $\lim \nu_{\mathbf{A}_f}(\mathsf{X}_f) \notin \{0, \lim \nu_{\mathbf{A}}(\mathsf{C}_1)\}$), what contradicts the hypothesis that C_1 is expanding hence strongly atomic (see Lemma 4.1.2). □

The following lemma is a restated version of a Lemma proved in [24].

Lemma 4.1.5. *Two nonnegligible clusters C_1 and C_2 are interweaving if and only if* $\mathrm{Profile}(\mathsf{C}_1) = \mathrm{Profile}(\mathsf{C}_2)$.

Our main result in this section reveals the expanding structure of local-global convergent sequences.

Theorem 4.1.6. *Let σ be a countable relational signature, let σ^+ be the extension of σ by countably many unary symbols N and U_i, $i \in \mathbb{N}$, and let σ^* be the extension of σ by countably many unary symbols N and $M_{i,j}$. Let $\mathsf{I}\colon \sigma^* \to \sigma^+$ be the basic interpretation defined by $\varrho_{U_i}(x) := \bigvee M_{i,j}(x)$ (and all other relations unchanged), and let $\mathsf{Sh}^*\colon \sigma^* \to \sigma$ and $\mathsf{Sh}^+\colon \sigma^+ \to \sigma$ be the natural forgetful interpretations.*

Then for every local-global convergent sequence $\mathbf{A}$ of σ-structure there exists a local-convergent sequence $\mathbf{A}^*$ such that

- *$P(\mathsf{Sh}^*)(\mathbf{A}^*) = \mathbf{A}$;*
- *for every $i, j \in \mathbb{N}$, $M_{i,j}(\mathbf{A}^*)$ is either null or an atomic cluster of $\mathbf{A}^*$, which is interweaving with $M_{i',j'}(\mathbf{A}^*)$ if and only if $i = i'$;*
- *$N(\mathbf{A}^*)$ is a nebula cluster of $\mathbf{A}^*$;*

and such that $\mathbf{A}^+ = P(\mathsf{I})(\mathbf{A}^*)$ has the following properties:

- *$\mathbf{A}^+$ is a local-global convergent sequence such that $P(\mathsf{Sh}^+)(\mathbf{A}^+) = \mathbf{A}$;*
- *for every $i \in \mathbb{N}$, $U_i(\mathbf{A}^+)$ is either null or a cluster of $\mathbf{A}^+$, which can be covered by (finitely many) interweaving atomic clusters;*
- *$N(\mathbf{A}^+)$ is a nebula cluster of $\mathbf{A}^+$.*

Note that this result is in agreement with the intuition: The σ^*-lift is "finer" than the σ^+-lift and thus less likely to be local-global convergent. For instance, we can refer to the subsequence extension property stated in the definition of lift-Hausdorff convergence (Definition 3.3.1).

PROOF: Let $\mathbf{A}$ be a local-global convergent sequence. We select inductively expanding clusters $\mathsf{C}^{i,j}$ of $\mathbf{A}$ as follows: We start with $\mathsf{Z} = \mathsf{0}$, $i = 1$, $j = 1$ and let P be the maximum profile of an expanding cluster of $\mathbf{A}$. Then we repeat the following procedure as long as there exists an expanding cluster of $\mathbf{A}$ that is weakly disjoint from Z.

- If there exists an expanding cluster in $\mathbf{A}$ with profile P that is weakly disjoint from Z we select one as $\mathsf{C}^{i,j}$, we let $\mathsf{Z} \leftarrow \mathsf{Z} \cup \mathsf{C}^{i,j}$, and we increase j by 1.
- Otherwise, we select one with maximum profile as $\mathsf{C}^{i+1,1}$, we let $\mathsf{Z} \leftarrow \mathsf{Z} \cup \mathsf{C}^{i+1,1}$, we let P be the profile of $\mathsf{C}^{i+1,1}$, we increase i by 1, and let $j = 1$.

It is easily checked that by modifying marginally the clusters $\mathsf{C}^{i,j}$ we can make them disjoint and such that $\mathsf{N} = \mathsf{A} \setminus \bigcup_{i,j} \mathsf{C}^{i,j}$ is a nebula cluster. Then by [24, Corollary 5] lifting $\mathbf{A}$ by marking $M_{i,j}$ the cluster $\mathsf{C}^{i,j}$ and N the cluster N we get a local-convergent sequence $\mathbf{A}^*$, which obviously satisfies the conditions stated in the Theorem 4.1.6.

Let $\mathbf{A}^+ = P(\mathsf{I})(\mathbf{A}^*)$. The only property we still have to prove is that $\mathbf{A}^+$ is local-global convergent. According to Definition 3.3.1 this boils down to proving that every local-convergent subsequence $\mathbf{B}^+_f$ of lifts of $\mathbf{A}^+$ can be extended into a full local-convergent sequence of lifts of $\mathbf{A}^+$. We can transfer the relations $M_{i,j}$ from $\mathbf{A}^*$ to $\mathbf{B}^+_f$. This way we obtain a subsequence $\mathbf{B}^*_f$ of lifts of $\mathbf{A}^*$ (which does not need to be local convergent), such that $P(\mathsf{I})(\mathbf{B}^*_f) = \mathbf{B}^+_f$. Let $\mathbf{B}^*_{g\circ f}$ be a local-convergent subsequence of $\mathbf{B}^*_f$. As $P(\mathsf{Sh}^*)(\mathbf{B}^*_{g\circ f}) = \mathbf{A}_{g\circ f}$ and $\mathbf{A}$ is local-global convergent there exists a local-convergent sequence $\mathbf{D}^*$ of lifts of $\mathbf{A}$ extending $\mathbf{B}^*_{g\circ f}$, that is: $P(\mathsf{Sh}^*)(\mathbf{D}^*) = \mathbf{A}$ and $\mathbf{D}^*_{g\circ f} = \mathbf{B}^*_{g\circ f}$. Let $\widehat{\mathsf{C}}^{i,j} = M_{i,j}(\mathbf{D}^*)$. As $(\widehat{\mathsf{C}}^{i,j})_{g\circ f} = (\mathsf{C}^{i,j})_{g\circ f}$ we get that $\widehat{\mathsf{C}}^{i,j}$ is a cluster of $\mathbf{A}$ with same profile as $\mathsf{C}^{i,j}$. According to Lemma 4.1.4, the limit set of $\nu_{\mathbf{A}}(\widehat{\mathsf{C}}^{i,j} \cap \mathsf{C}^{i,j'})$ is included in $\{0, m\}$, where $m = \lim \nu_{\mathbf{A}}(\widehat{\mathsf{C}}^{i,j}) = \lim \nu_{\mathbf{A}}(\mathsf{C}^{i,j'})$. It follows that either $\widehat{\mathsf{C}}^{i,j} \preceq \bigcup_{j'} \mathsf{C}^{i,j'}$ or there exists a subsequence $\mathbf{A}_h$ of $\mathbf{A}$ such that $(\widehat{\mathsf{C}}^{i,j})_h$ is weakly disjoint from the cluster $(\bigcup_{j'} \mathsf{C}^{i,j'})_g$. Marking all the clusters $\mathsf{C}^{i,j}$ and $\widehat{\mathsf{C}}^{i,j}$ in $\mathbf{A}_h$ we get a local-convergent subsequence of lifts, which can be extended into full local-convergent sequence of lifts of $\mathbf{A}$. In this sequence, the marks corresponding to the extension of $(\widehat{\mathsf{C}}^{i,j})_g$ will correspond to a cluster of $\mathbf{A}$ disjoint from all the clusters $\mathsf{C}^{i,j'}$ but with the same profile, which contradicts the construction procedure of the clusters $\mathsf{C}^{i,j}$. Thus $\widehat{\mathsf{C}}^{i,j} \preceq \bigcup_{j'} \mathsf{C}^{i,j'}$, and $\bigcup_j \widehat{\mathsf{C}}^{i,j} \preceq \bigcup_{j'} \mathsf{C}^{i,j'}$. As these two clusters have same limit measure we have $\bigcup_j \widehat{\mathsf{C}}^{i,j} \approx \bigcup_{j'} \mathsf{C}^{i,j'}$. This means that $P(\mathsf{I})(\mathbf{D}^*)$ and $P(\mathsf{I})(\mathbf{B}^*)$ are sufficiently close, so that if we consider the lifts of $\mathbf{A}^+$ defined by $\mathbf{B}^+$ for indices of the form $f(n)$ for some $n \in \mathbb{N}$ and by $P(\mathsf{I})(\mathbf{D}^*)$ for the other indices, we get a local-convergent sequence of lifts of $\mathbf{A}^+$ which extends $\mathbf{B}^+_f$. It follows that $\mathbf{A}^+$ is local-global convergent, what concludes our proof. □

4.2 Local-global quasi-limits. Let us finish this paper in an ambitious way. In [25], [22] we defined the notion of modeling as a limit object from structural convergence.

Modeling limits generalize graphing limits and thus it follows from [20] that FO-convergent sequences of graphs with bounded degrees have modeling limits. In [25] we constructed modeling limits for FO-convergent sequences of graphs with bounded tree-depth, and extended the construction to FO-convergent sequences of trees in [22]. Then existence of modelings for FO-convergent sequences has been proved for graphs with bounded path-width in [11], and eventually for sequences of graphs in an arbitrary nowhere dense class, see [26], which is best possible when considering monotone classes of graphs, see [22]. In fact this provides us with a high level analytic characterization of nowhere dense classes.

Definition 4.2.1. Let $\mathbf{A}$ be a local-global convergent sequence. A modeling $\mathbf{L}$ is a *local-global quasi-limit* of $\mathbf{A}$ if for every local convergent sequence $\mathbf{A}^+$ of lifts of $\mathbf{A}$ and every $\varepsilon > 0$ there exists an admissible lift $\mathbf{L}^+$ of $\mathbf{L}$ (that is a lift $\mathbf{L}^+$ of $\mathbf{L}$ that is a modeling), such that for every local formula φ we have

$$\left|\langle \varphi, \mathbf{L}^+\rangle - \lim_{n\to\infty} \langle \varphi, \mathbf{A}_n^+\rangle\right| < c_\varphi \varepsilon,$$

where c_φ is a positive constant, which depends only on φ.

In other words, the closure of the measures associated to admissible lifts of $\mathbf{L}$ includes the limit Hausdorff limit of the sets of measures associated to lifts of the sequence.

For local-global convergence, it was proved in [12] that graphings still suffice as limit objects. We don't know, however, if every local-global convergent sequence of graphs in a nowhere dense class has a modeling local-global limit. We close this paper by proving that this is almost the case, in the sense that every local-global convergent sequence of graphs in a nowhere dense class has a modeling local-global quasi-limit.

We consider a fixed countable signature σ and the signature τ obtained by adding countably many unary symbols $M_1, M_2, \ldots, M_n, \ldots$ to σ, and the forgetful interpretation $\mathsf{Sh}\colon \tau \to \sigma$. As before we understand local-global convergence as $\mathsf{Sh}_*(\mathrm{FO}^{\mathrm{local}})$-convergence. We fix a chain covering $\mathfrak{X}$ of $\mathrm{FO}^{\mathrm{local}}(\tau)$ (see Section 2.5), from which we derive metrics and pseudo-metrics as in Sections 3.2 and 3.3. We also fix a bijection $\beta\colon \mathbb{N}\times\mathbb{N} \to \mathbb{N}$ (for Hilbert hotel argument) and let Z_c be the renaming interpretation which renames $M_{\beta(c,i)}$ as M_i and forget all the marks not being renamed.

Lemma 4.2.2. *There exists a function $h\colon (0,1) \to \mathbb{N}$ with the following property:*

For every local-global convergent sequence $\mathbf{A}$ of σ-structures there exists a local convergent sequence $\mathbf{B}$ of τ-structures with $P(\mathsf{Sh})(\mathbf{B}) = \mathbf{A}$, such that for every $\varepsilon > 0$ there exists some integer n_0 such that

for every $n \geq n_0$ and every $\mathbf{C} \in P(\mathsf{Sh})^{-1}(\mathbf{A}_n)$ there exists $1 \leq c \leq h(\varepsilon)$ with

$$\mathrm{dist}_{\mathfrak{X}}(\mathbf{C}, P(\mathsf{Z}_c)(\mathbf{B}_n)) < \varepsilon.$$

PROOF: As the space $(\mathrm{Rel}(\tau), \mathrm{dist}_{\mathfrak{X}})$ is totally bounded there exists a mapping $g\colon (0,1) \to \mathbb{N}$ such that for each $\varepsilon > 0$ and each σ-structure $\mathbf{A}$ there is a subset $\mathcal{B}_{\mathbf{A},\varepsilon}$ of $P(\mathsf{Sh})^{-1}(\mathbf{A})$ of cardinality at most $g(\varepsilon)$ with the property that every $\mathbf{C} \in P(\mathsf{Sh})^{-1}(\mathbf{A})$ is at $\mathrm{dist}_{\mathfrak{X}}$-distance at most ε from a τ-structure in $\mathcal{B}_{\mathbf{A},\varepsilon}$. (Such a set may be called an ε-covering.) We construct an infinite sequence $(\mathbf{A}^{(i)})_{i\in\mathbb{N}}$ of τ-structures by listing all the structures in $\mathcal{B}_{\mathbf{A},1/2}$ then all the structures in $\mathcal{B}_{\mathbf{A},1/4}$, etc.

We now construct a τ-structure $\mathbf{A}^+ \in P(\mathsf{Sh})^{-1}(\mathbf{A})$ by letting $M_{\beta(i,j)}(\mathbf{A}^+) = M_j(\mathbf{A}^{(i)})$. Hence $\mathbf{A}^{(i)} = P(\mathsf{Z}_i)(\mathbf{A}^+)$. We say that $\mathbf{A}^+$ is a *universal lift* of $\mathbf{A}$.

Define the function $h\colon (0,1) \to \mathbb{N}$ by

$$h(x) = \sum_{i=1}^{\lceil -\log x \rceil + 1} g(2^{-i}).$$

Then for every $\varepsilon > 0$ and every $\mathbf{B} \in P(\mathsf{Sh})^{-1}(\mathbf{A})$ there is an index $c \leq h(\varepsilon)$ such that $\mathrm{dist}_{\mathfrak{X}}(\mathbf{C}, \mathbf{A}^{(c)}) < \varepsilon/2$, that is such that $\mathrm{dist}_{\mathfrak{X}}(\mathbf{C}, P(\mathsf{Z}_c)(\mathbf{A}^+)) < \varepsilon/2$.

Now consider the local-global convergent sequence A and a sequence A^+ where $\mathbf{A}_n^+$ is a universal lift of $\mathbf{A}_n$. This last sequence has a local convergent subsequence A_f^+, which we extend into a sequence B lifting A.

Let $\varepsilon > 0$. According to local-global convergence of A and local convergence of B there exists n_0 such that for every $n, m \geq n_0$ we have $\mathrm{dist}^{\mathrm{H}}_{\mathfrak{X},\mathsf{Sh}}(\mathbf{A}_n, \mathbf{A}_m) < \varepsilon/4$ and $\mathrm{dist}_{\mathfrak{X}}(\mathbf{B}_n, \mathbf{B}_m) < \alpha$, where α is such that for every $i \leq h(\varepsilon/2)$ we have

$$\mathrm{dist}_{\mathfrak{X}}(\mathbf{X}, \mathbf{Y}) < \alpha \Rightarrow \mathrm{dist}_{\mathfrak{X}}(P(\mathsf{Z}_i)(\mathbf{X}), P(\mathsf{Z}_i)(\mathbf{Y})) < \frac{\varepsilon}{4}.$$

Let $n \geq n_0$ (hence $f(n) \geq n_0$). Let $\mathbf{C} \in P(\mathsf{Sh})^{-1}(\mathbf{A}_n)$. Then there exists $\mathbf{C}' \in P(\mathsf{Sh})^{-1}(\mathbf{A}_{f(n)})$ such that $\mathrm{dist}_{\mathfrak{X}}(\mathbf{C}, \mathbf{C}') < \varepsilon/4$. As $\mathbf{B}_{f(n)} = \mathbf{A}^+_{f(n)}$ is a universal lift of $\mathbf{A}_{f(n)}$ there exists $c \leq h(\varepsilon)$ such that $\mathrm{dist}_{\mathfrak{X}}(P(\mathsf{Z}_c)(\mathbf{B}_{f(n)}), \mathbf{C}') < \varepsilon/2$. As $\mathrm{dist}_{\mathfrak{X}}(\mathbf{B}_n, \mathbf{B}_{f(n)}) < \alpha$ we have $\mathrm{dist}_{\mathfrak{X}}(P(\mathsf{Z}_c)(\mathbf{B}_{f(n)}), P(\mathsf{Z}_c)(\mathbf{B}_n)) < \varepsilon/4$. Altogether, we get $\mathrm{dist}_{\mathfrak{X}}(P(\mathsf{Z}_c)(\mathbf{B}_n), \mathbf{C}) < \varepsilon$ as wanted. □

Definition 4.2.3. A σ-modeling $\mathbf{L}$ is a *quasi-limit* of a local-global convergent sequence A of σ-structures if for every local convergent sequence A^+ of τ-structures with $P(\mathsf{Sh})(\mathsf{A}^+) = \mathsf{A}$ and for every $\varepsilon > 0$ there exists a τ-modeling $\mathbf{L}^+$ with $P(\mathsf{Sh})(\mathbf{L}^+) = \mathbf{L}$ such that $\limsup \mathrm{dist}_{\mathfrak{X}}(\mathbf{L}^+, \mathbf{A}^+) < \varepsilon$.

In other words, for every local-global convergent sequence A there is a modeling $\mathbf{L}$ such that any local convergent sequence lifting A has a limit which is ε-close to an admissible lifting of $\mathbf{L}$. (By admissible, we mean that the lift of $\mathbf{L}$ is itself a modeling.)

Theorem 4.2.4. *Every local-global convergent sequence of graphs in a nowhere dense class has a modeling quasi-limit.*

PROOF: Let A be a local-global convergent of graphs in a nowhere dense class. According to Lemma 4.2.2 there exists a local convergent sequence B of marked

graphs with $P(\mathsf{Sh})(\mathbf{B}) = \mathbf{A}$, such that for every $\varepsilon > 0$ there exists some integer n_0 such that for every $n \geq n_0$ and every $\mathbf{C} \in P(\mathsf{Sh})^{-1}(\mathbf{A}_n)$ there exists $1 \leq c \leq h(\varepsilon)$ with

$$\mathrm{dist}_{\mathfrak{X}}(\mathbf{C}, P(\mathsf{Z}_c)(\mathbf{B}_n)) < \varepsilon.$$

According to [26] the sequence $\mathbf{B}$ has a modeling limit $\mathbf{L}^+$. Then $\mathbf{L} = P(\mathsf{Sh}(\mathbf{L}^+)$ is a modeling quasi-limit of $\mathbf{A}$. □

We conjecture that it is possible to refine the notion of admissible lift and get the reverse direction.

Conjecture 4.2.5. *Every local-global convergent sequence of graphs in a nowhere dense class has a modeling limit.*

References

[1] Adler H., Adler I., *Interpreting nowhere dense graph classes as a classical notion of model theory*, European J. Combin. **36** (2014), 322–330.

[2] Aldous D. J., *Representations for partially exchangeable arrays of random variables*, J. Multivariate Anal. **11** (1981), no. 4, 581–598.

[3] Alon N., *Eigenvalues and expanders*, Theory of computing, Combinatorica **6** (1986), no. 2, 83–96.

[4] Balcar B., Jech T., Pazák T., *Complete CCC Boolean algebras, the order sequential topology, and a problem of von Neumann*, Bull. London Math. Soc. **37** (2005), no. 6, 885–898.

[5] Benjamini I., Schramm O., *Recurrence of distributional limits of finite planar graphs*, Electron. J. Probab. **6** (2001), no. 23, 13 pages.

[6] Bollobás B., Riordan O., *Sparse graphs: metrics and random models*, Random Structures Algorithms **39** (2011), no. 1, 1–38.

[7] Borgs Ch., Chayes J., Lovász L., *Moments of two-variable functions and the uniqueness of graph limits*, Geom. Funct. Anal. **19** (2010), no. 6, 1597–1619.

[8] Borgs Ch., Chayes J., Lovász L., Sós V. T., Szegedy B., Vesztergombi K., *Graph limits and parameter testing*, STOC'06, Proc. of the 38th Annual ACM Symposium on Theory of Computing, 2006, pages 261–270.

[9] Borgs C., Chayes J. T., Lovász L., Sós V. T., Vesztergombi K., *Convergent sequences of dense graphs I. Subgraph frequencies, metric properties and testing*, Adv. Math. **219** (2008), no. 6, 1801–1851.

[10] Borgs C., Chayes J. T., Lovász L., Sós V. T., Vesztergombi K., *Convergent sequences of dense graphs II. Multiway cuts and statistical physics*, Ann. of Math. (2) **176** (2012), no. 1, 151–219.

[11] Gajarský J., Hliněný P., Kaiser T., Král' D., Kupec M., Obdržálek J., Ordyniak S., Tůma V., *First order limits of sparse graphs: plane trees and path-width*, Random Structures Algorithms **50** (2017), no. 4, 612–635.

[12] Hatami H., Lovász L., Szegedy B., *Limits of locally-globally convergent graph sequences*, Geom. Funct. Anal. **24** (2014), no. 1, 269–296.

[13] Hausdorff F., *Set Theory*, Chelsea Publishing, New York, 1962.

[14] Hodges W., *A Shorter Model Theory*, Cambridge University Press, Cambridge, 1997.

[15] Hoory S., Linial N., Wigderson A., *Expander graphs and their applications*, Bull. Amer. Math. Soc. (N.S.) **43** (2006), no. 4, 439–561.

[16] Hoover D., *Relations on Probability Spaces and arrays of Random Variables*, Tech. report, Institute for Advanced Study, Princeton, 1979.

[17] Kun G., Thom A., *Inapproximability of actions and Kazhdan's property (T)*, available at arXiv:1901.03963 [math.GR] (2019), 9 pages.
[18] Lascar D., *La théorie des modèles en peu de maux*, Nouvelle Bibliothèque Mathématique, 10, Cassini, Paris, 2009 (French).
[19] Lovász L., Szegedy B., *Limits of dense graph sequences*, J. Comb. Theory Ser. B **96** (2006), no. 6, 933–957.
[20] Nešetřil J., Ossona de Mendez P., *A model theory approach to structural limits*, Comment. Math. Univ. Carolin. **53** (2012), no. 4, 581–603.
[21] Nešetřil J., Ossona de Mendez P., *First-order limits, an analytical perspective*, European J. Combin. **52** (2016), part B, 368–388.
[22] Nešetřil J., Ossona de Mendez P., *Modeling limits in hereditary classes: reduction and application to trees*, Electron. J. Combin. **23** (2016), no. 2, paper 2.52, 33 pages.
[23] Nešetřil J., Ossona de Mendez P., *Structural sparsity*, Uspekhi Mat. Nauk 71 (2016), no. 1(427), 85–116; translation in Russian Math. Surveys 71 (2016), no. 1, 79–107.
[24] Nešetřil J., Ossona de Mendez P., *Cluster analysis of local convergent sequences of structures*, Random Structures Algorithms **51** (2017), no. 4, 674–728.
[25] Nešetřil J., Ossona de Mendez P., *A unified approach to structural limits (with application to the study of limits of graphs with bounded tree-depth)*, to appear in Mem. Amer. Math. Soc. (2017), 117 pages.
[26] Nešetřil J., Ossona de Mendez P., *Existence of modeling limits for sequences of sparse structures*, published online in J. Symb. Log. (2019), 22 pages. doi:10.1017/jsl.2018.32
[27] Urysohn P., *Works on Topology and Other Areas of Mathematics 1, 2*, State Publ. of Technical and Theoretical Literature, Moskva, 1951, (Russian).

Jaroslav Nešetřil:
Charles University, Faculty of Mathematics and Physics,
Computer Science Institute of Charles University, Malostranské nám. 25, 118 00 Praha 1, Czech Republic

E-mail: nesetril@iuuk.mff.cuni.cz

Patrice Ossona de Mendez:
Centre d'Analyse et de Mathématiques Sociales,
L'Ecole des Hautes Etudes en Sciences Sociales, 190–198 avenue de France, 75013 Paris, France
and
Charles University, Faculty of Mathematics and Physics,
Computer Science Institute of Charles University, Malostranské nám. 25, 118 00 Praha 1, Czech Republic

E-mail: pom@ehess.fr

(*Received* April 27, 2018*, revised* October 16, 2018)