



First-order limits, an analytical perspective

Jaroslav Nesetril, Patrice Ossona de Mendez

► To cite this version:

Jaroslav Nesetril, Patrice Ossona de Mendez. First-order limits, an analytical perspective . 2014.
hal-01135593

HAL Id: hal-01135593

<https://hal.science/hal-01135593>

Preprint submitted on 25 Mar 2015

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

FIRST-ORDER LIMITS, AN ANALYTICAL PERSPECTIVE

JAROSLAV NEŠETŘIL AND PATRICE OSSONA DE MENDEZ

ABSTRACT. In this paper we present a novel approach to graph (and structural) limits based on model theory and analysis. The role of Stone and Gelfand dualities is displayed prominently and leads to a general theory, which we believe is naturally emerging. This approach covers all the particular examples of structural convergence and it put the whole in new context. As an application, it leads to new intermediate examples of structural convergence and to a “grand conjecture” dealing with sparse graphs. We survey the recent developments.

1. INTRODUCTION

Let G_1, G_2, \dots, G_n be a sequence of graphs with increasing orders. The theory of graph limits allows us to capture some common features of such a sequence by a single structure, the limit object. Such features include property testing and subgraph frequencies. Positive instances of this procedure are related to both dense and sparse graphs. This recently very active area of combinatorics is discussed for instance in [23].

Can one demand that the limit object satisfies some further properties such as degree distribution, or bounded diameter, or non-existence of certain subgraphs?

This we approach by the notion of first-order limits of structures, which was introduced in [30, 33].

This approach is based on the combination of model theory and functional analysis, and proved to be useful in dealing with “intermediate” classes, such as the class of trees and forests [32]. It also generalizes (and as well puts in a new context) several notions of graph convergence studied earlier [6, 24, 5, 3].

In Sections 3–6 we review here some of these results. However we also provide a new setting which generalizes and complements the earlier work of the authors [30]. We now view our approach as a natural approach, which lies midway between functional analysis approach to quantum theory and limits of graphs. This, perhaps even closer to functional analysis, we tried to explain in Section 2.

Date: March 26, 2015.

2010 Mathematics Subject Classification. Primary 03C13 (Finite structures), 03C98 (Applications of model theory), 05C99 (Graph theory), 06E15 (Stone spaces and related structures), Secondary 28C05 (Integration theory via linear functionals).

Key words and phrases. Graph and Relational structure and Functional Analysis and Graph limits and Structural limits and Radon measures and Stone space and Model theory and First-order logic and Measurable graph.

Supported by grant ERCCZ LL-1201 and CE-ITI P202/12/G061, and by the European Associated Laboratory “Structures in Combinatorics” (LEA STRUCO).

Supported by grant ERCCZ LL-1201 and by the European Associated Laboratory “Structures in Combinatorics” (LEA STRUCO).

Here is an outline of what we do: We generalize aspects of the theory of graph limits, moving from a study a homomorphism and subgraph profiles to considering full statistics of first-order formula satisfaction. In our setting the underlying notion of convergence is, in essence, model theoretic, and relies on the following notion [30]:

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 $\phi(\mathbf{A})$ the *satisfying set* of ϕ in \mathbf{A} :

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

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

$$(1) \quad \langle \phi, \mathbf{A} \rangle = \frac{|\phi(\mathbf{A})|}{|A|^p}$$

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

A sequence $(\mathbf{A}_n)_{n \in \mathbb{N}}$ of finite σ -structures is *FO-convergent* if the sequence $(\langle \phi, \mathbf{A}_n \rangle)_{n \in \mathbb{N}}$ converges for every first-order formula ϕ . It is important that one can derive 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 .

We shall see that this abstract generalization of left convergence and local convergence can be seen as a particular (commutative) case of a general framework, which occurs naturally in functional analysis (in its description of quantum physics). This connection not only legitimates some of the constructions introduced to study first-order limits of structures, but it also put the whole framework in perspective, for a possible extension to the non-commutative setting of quantum logic. This connection will be discussed in Section 2, which may be seen as generalization and more uniform treatment of [33].

Encouraged by this general analytic setting, we show that 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. In this setting, there is a one-to-one map $\mathbf{A} \mapsto \mu_{\mathbf{A}}$ from the class of finite σ -structures to the space of probability measures on S , with

$$(2) \quad \int_S \mathbf{1}_{\phi}(T) d\mu_{\mathbf{A}}(T) = \langle \phi, \mathbf{A} \rangle,$$

where $\mathbf{1}_{\phi}$ is the indicator function of the clopen subset of S dual to the formula ϕ in Stone duality. In this setting, a sequence $(\mathbf{A}_n)_{n \in \mathbb{N}}$ of finite σ -structures is first-order convergent if and only if the measures $\mu_{\mathbf{A}_n}$ converge weakly to some measure μ . In such a case, the probability measure μ represents the limit of the sequence $(\mathbf{A}_n)_{n \in \mathbb{N}}$ and, for every first-order formula ϕ it holds

$$(3) \quad \int_S \mathbf{1}_{\phi}(T) d\mu(T) = \lim_{n \rightarrow \infty} \int_S \mathbf{1}_{\phi}(T) d\mu_{\mathbf{A}_n}(T) = \lim_{n \rightarrow \infty} \langle \phi, \mathbf{A}_n \rangle.$$

This answers in a very general form the problem (originally posed by J. Chayes, see [23, Introduction]) whether convergence leads to a limit distribution.

The particular definition (1) of the Stone pairing of a first-order formula and a finite structure can be extended, under certain conditions, to infinite structures. An obvious necessary condition for such an extension is that the domain of the considered infinite structure should be a measurable space, with the property that every first-order definable set is measurable. Note that the difficulty here stands

in the fact that the class of definable sets is closed by projections, what is not the case of measurable sets. Consequently, we introduce in Section 5 the following notions: a *Relational Sample Space* (RSS) is a relational structure, whose domain is a standard Borel space, with the property that every first-order definable subset of a power of the domain is measurable (with respect to Borel product measure). A *modeling* \mathbf{M} [33] is an RSS equipped with a probability measure denoted $\nu_{\mathbf{M}}$. In this setting, we can extend the definition of Stone pairing to modelings: for a modeling \mathbf{M} and a first-order formula ϕ with free variables x_1, \dots, x_p we define

$$(4) \quad \langle \phi, \mathbf{M} \rangle = \int \cdots \int \mathbf{1}_{\phi(\mathbf{M})}(v_1, \dots, v_p) d\nu_{\mathbf{M}}(v_1) \cdots d\nu_{\mathbf{M}}(v_p),$$

where $\mathbf{1}_{\phi(\mathbf{M})}$ is the indicator function of the set

$$(5) \quad \phi(\mathbf{M}) = \{(v_1, \dots, v_p) \in M^p : \mathbf{M} \models \phi(v_1, \dots, v_p)\}.$$

This definition naturally leads to the two following representation problems, which can be seen as analogs of celebrated Aldous-Lyons problem on graphings.

Problem 1. Characterize those probability measures μ on X for which there exists a sequence $(\mathbf{A}_n)_{n \in \mathbb{N}}$ of finite structures such that for every first-order formula $\phi(x_1, \dots, x_p)$ it holds

$$(6) \quad \int_S \mathbf{1}_{\phi}(T) d\mu(T) = \lim_{n \rightarrow \infty} \langle \phi, \mathbf{A}_n \rangle.$$

Problem 2. Characterize those first-order convergent sequences of finite structures $(\mathbf{A}_n)_{n \in \mathbb{N}}$ for which there exists a modeling \mathbf{M} such that for every first-order formula ϕ it holds

$$(7) \quad \langle \phi, \mathbf{M} \rangle = \lim_{n \rightarrow \infty} \langle \phi, \mathbf{A}_n \rangle.$$

As we shall see, such modeling limits do not exist in general (but sometimes they do) and thus we shall be interested in the following problem.

Problem 3. Characterize classes with modeling limits, that is classes \mathcal{C} of finite structures, such that for every first-order convergent sequence $(\mathbf{A}_n)_{n \in \mathbb{N}}$ of structures in \mathcal{C} there is a modeling \mathbf{M} such that $\langle \phi, \mathbf{M} \rangle = \lim_{n \rightarrow \infty} \langle \phi, \mathbf{A}_n \rangle$ for every first-order formula ϕ .

It appears that this last problem is, perhaps surprisingly, related to the sparse-dense dichotomy. The authors introduced in [29, 27] a dichotomy of classes of graphs, between (sparse) *nowhere dense* classes and (dense) *somewhere dense* classes. This dichotomy, which can be expressed in numerous (and non trivially equivalent) ways, appears to be deeply related to first-order related properties of these classes. Based on a characterization of nowhere dense classes by means of Vapnik-Chervonenkis dimension of model-theoretic interpretations [1] and a characterization of random free hereditary classes of graphs [25], we proved in [33] that a monotone class of graphs with modeling limits is necessarily nowhere dense. We believe that nowhere dense-somewhere dense dichotomy actually gives the answer to Problem 3.

Conjecture 1 ([33]). Every monotone nowhere dense class of graphs has modeling limits.

By a combination of model theory and functional analysis methods, we have been able to prove some particular cases: classes of graphs with bounded degree, classes of structures with bounded tree-depth [33], as well as the class of all forests [32] admit modeling limits. In this paper, we sketch a generalization to any class of graphs where the maximum degree is bounded by a function of the girth (or unbounded if the graph is acyclic). But before, in the next section, we outline an even more general approach (pointing to a possible non-commutative version).

2. ABSTRACT ANALYTIC FRAMEWORK

We introduced in [30] the notion of first-order limits of structures as a generalization and unified treatment of various notions of limits. This notions also provided limit objects for new “intermediate classes”. However first-order limits are interesting on their own. Particularly, as we show here, they display a rich spectrum of interconnection to the functional analysis approach to quantum theory [12]. For instance, the connection to C^* -algebras makes more natural the introduction of several spaces and constructions, whose appearance in our framework may have earlier seem a little mysterious.

The formalism we adopt here is indeed close to the one of quantum theory, specially when considering significant objects and spaces. This aspect will be developed in this section, where we consider our approach to structure limits under the lights of statistical and quantum physics, as formalized by functional analysis.

Let us recall some basics: The set of all possible states of a physical system is the *phase space* of the system. In statistical physics, a *statistical ensemble* is the phase space of a physical system together with a method of averaging physical quantities, called *observables*, related to this system. In a classical system with phase space Ω , the observable quantities are real functions defined on Ω , and they are averaged by integration with respect to a certain probability measure μ on Ω . In a quantum system described by vectors in a Hilbert space \mathcal{H} , the observable quantities are defined by self-adjoint operators acting on \mathcal{H} , and are averaged using a certain positive, normalized functional ρ , defined on the algebra $\mathfrak{A}(\mathcal{H})$ of operators on \mathcal{H} (such functionals on \mathcal{H} are called *states*) [26].

2.1. C^* -algebras. A new unifying view consists in considering as a primary object the C^* -algebra \mathfrak{A} of observables. Recall that a C^* -algebra \mathfrak{A} is a Banach algebra with an involution $x \mapsto x^*$, such that the relation $\|x^*x\| = \|x\|^2$ holds for any element $x \in \mathfrak{A}$. The two standard examples of C^* -algebras are the following (see [16] for the standard text):

First example is the space $C_0(X)$ of all continuous complex-valued functions which vanish at infinity on a locally compact Hausdorff space X , equipped with the uniform norm $\|f\| = \sup_{x \in X} |f(x)|$, and involution defined as the complex-conjugate: $f^*(x) = \overline{f(x)}$. Gelfand’s representation theorem states that every commutative C^* -algebra \mathfrak{A} is isometrically $*$ -isomorphic to the algebra $C_0(\Phi_{\mathfrak{A}})$, where $\Phi_{\mathfrak{A}}$ is the topological space of all the *characters* of \mathfrak{A} (that is: of the non-zero homomorphisms $f : \mathfrak{A} \rightarrow \mathbb{C}$) equipped with the relative weak- $*$ topology. Given $a \in \mathfrak{A}$, one defines the *Gelfand transform* \hat{a} of a , that is the function $\hat{a} : \Phi_{\mathfrak{A}} \rightarrow \mathbb{C}$ by $\hat{a}(f) = f(a)$. Then the map $a \mapsto \hat{a}$ defines a norm-decreasing, unit-preserving algebra homomorphism from \mathfrak{A} to $C_0(\Phi_{\mathfrak{A}})$, which is the *Gelfand representation* of \mathfrak{A} .

The second standard example is the algebra $\mathfrak{B}(\mathcal{H})$ of all bounded linear operators on a Hilbert space \mathcal{H} , where involution is defined as the adjoint operator, and norm is defined as the operator norm. Gelfand–Naimark theorem states that any C^* -algebra is isometrically and symmetrically isomorphic to a C^* -subalgebra of some C^* -algebra of the form $\mathfrak{B}(\mathcal{H})$.

In our (discrete) setting, we shall be interested in the specific case where the algebra \mathfrak{A} of observables is *approximately finite* [7], that is in the case where there exists an increasing sequence $\mathfrak{A}_1 \subseteq \mathfrak{A}_2 \subseteq \dots \subseteq \mathfrak{A}$ of finite-dimensional sub- C^* -algebras of \mathfrak{A} such that the union $\bigcup_j \mathfrak{A}_j$ is norm-dense in \mathfrak{A} . Approximately finite C^* -algebras (or *AF C^* -algebras*) have been fully classified by Elliott [11] using K -theory for C^* -algebras. In particular, a C^* -algebra \mathfrak{A} is commutative and approximately finite if and only if it is the algebra of continuous functions on the Stone space of a Boolean algebra (follows from [8, Proposition 3.1] and Stone duality). This very particular case corresponds to the framework we used to study structural limits. Fig. 1 describes the different mathematical objects entering this grand picture.

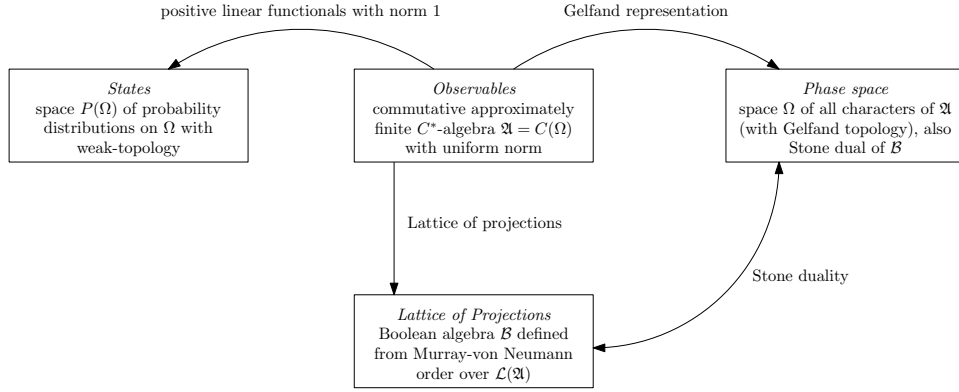


FIGURE 1. The different spaces in the standard case

In this context, let us recall that to each Boolean algebra \mathcal{B} is associated a topological space, denoted $S(\mathcal{B})$, called the *Stone space* of \mathcal{B} , which is a totally disconnected compact Hausdorff space. The points in $S(\mathcal{B})$ are the ultrafilters on \mathcal{B} , or equivalently the homomorphisms from \mathcal{B} to the two-element Boolean algebra. The mapping $K : \mathcal{B} \rightarrow 2^{S(\mathcal{B})}$ defined by

$$K(\phi) = \{T \in S(\mathcal{B}) : \phi \in T\}$$

gives a one-to-one correspondence between elements of \mathcal{B} and clopen subsets of $S(\mathcal{B})$. The topology on $S(\mathcal{B})$ is generated by a basis consisting of all (clopen) sets of the form $\{x \in S(\mathcal{B}) \mid b \in x\}$, where b is an element of \mathcal{B} . For every Boolean algebra \mathcal{B} , $S(\mathcal{B})$ is a compact totally disconnected Hausdorff space, and Stone representation theorem [36] states that every Boolean algebra \mathcal{B} is isomorphic to the algebra of clopen subsets of its Stone space $S(\mathcal{B})$. Note that the Stone dual of a countable Boolean algebra is a compact Polish space. The similarity of Stone duality with Gelfand representation is not surprising, if one considers the motivation that led Stone to this theorem:

“The writer’s interest in the subject, for example, arose in connection with the spectral theory of symmetric transformations in Hilbert space and certain related properties of abstract integrals.”
(M.H. Stone [36])

2.2. Structural Limits. In the context of structural limits, the general picture of Fig. 1 takes more concrete form depicted on Fig. 2. This is a refinement of the approach the authors presented in [30, 33, 32].

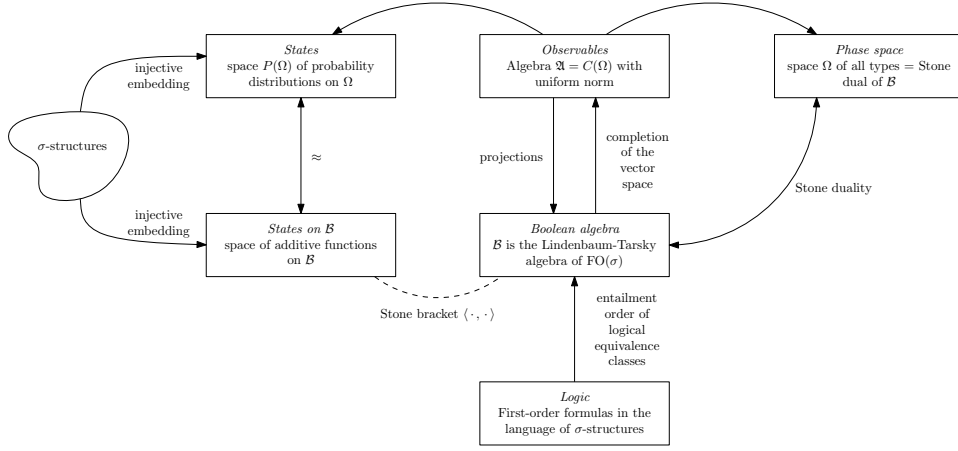


FIGURE 2. The spaces considered in the study of structural limits

Let us explain the role of the different spaces entering this scheme: We consider the class $\text{FO}(\sigma)$ of all first-order formulas constructed with equality and the symbols in the considered signature σ , and the Boolean algebra \mathcal{B} defined as the Lindenbaum-Tarski algebra of $\text{FO}(\sigma)$, that is the Boolean algebra of the classes of these formulas with respect to logical equivalence. For instance, in the case of undirected graphs, formulas are constructed from the symbols of equality and adjacency, as well as standard connectives (conjunctions, disjunction, negation) and quantification over vertices.

The Stone space $S(\mathcal{B})$ associated to \mathcal{B} is a topological space, whose topology is generated by clopen subsets, which correspond exactly to first-order formulas (up to logical equivalence), that is to elements of \mathcal{B} . A point in $S(\mathcal{B})$ corresponds to a maximal set of consistent formulas.

As to each formula ϕ is associated a clopen subset $K(\phi)$ of $S(\mathcal{B})$, the indicator functions $\mathbf{1}_{K(\phi)}$ of clopen subsets of $S(\mathcal{B})$ are continuous functions from $S(\mathcal{B})$ to \mathbb{R} . An algebra \mathfrak{A} can be constructed from \mathcal{B} as the algebra of formal finite linear combinations of elements of \mathcal{B} quotiented by the relation $\phi + \psi = \phi \wedge \psi + \phi \vee \psi$. Each such finite linear combination can be written as a sum $x = \sum_{i=1}^n \alpha_i \phi_i$ where $\phi_i \wedge \phi_j = 0$ whenever $i \neq j$. It turns out that $\|x\| = \max_i |\alpha_i|$ is independent of the chosen decomposition, and defines a norm on \mathfrak{A} (this norm is nothing but the $\|\cdot\|_\infty$ norm on $C(S(\mathcal{B}))$). The completion of \mathfrak{A} for this norm is the Banach algebra $C(S(\mathcal{B}))$ of all continuous functions from $S(\mathcal{B})$ to \mathbb{R} . (Note that if we consider for \mathfrak{A} complex linear combinations, we similarly construct the C^* -algebra of all complex valued functions on $S(\mathcal{B})$.)

Finite σ -structures embed injectively in the space of additive functions on the Boolean algebra \mathcal{B} by $\mathbf{A} \mapsto \langle \cdot, \mathbf{A} \rangle$, where

$$\langle \phi, \mathbf{A} \rangle = \frac{|\{(v_1, \dots, v_p) \in A^p : \mathbf{A} \models \phi(v_1, \dots, v_p)\}|}{|A|^p}.$$

The following fact is worth noticing:

Fact 1. There is a one-to-one correspondence between probability measures μ on $S(\mathcal{B})$ and additive functions $f_\mu : \mathcal{B} \rightarrow [0, 1]$ such that $f(1) = 1$, with the property that for every probability measure μ on $S(\mathcal{B})$ and every $\phi \in \mathcal{B}$ it holds

$$f_\mu(\phi) = \mu(K(\phi)).$$

This fact, which we proved in [30] as a lemma, is actually known in the more general setting of MV-algebras [21, 34]. We give here a short proof for the sake of completeness:

Proof. The vector space $V(\mathcal{B})$ generated by the indicator functions $\mathbf{1}_{K(\phi)}$ for $\phi \in \mathcal{B}$ forms a subalgebra of $C(S(\mathcal{B}))$. The subalgebra $V(\mathcal{B})$ separates points of $S(\mathcal{B})$: if $T_1, T_2 \in S(\mathcal{B})$, then there exists $\phi \in T_1 \setminus T_2$ thus $\mathbf{1}_{K(\phi)}(T_1) \neq \mathbf{1}_{K(\phi)}(T_2)$. As the constant 1 function (corresponding to the maximum of the Boolean algebra) belongs to $V(\mathcal{B})$, the subalgebra $V(\mathcal{B})$ is dense in $C(S(\mathcal{B}))$, according to Stone–Weierstrass theorem. As $S(\mathcal{B})$ is a Radon-space, probability measures μ on $S(\mathcal{B})$ correspond (in a one-to-one correspondence) to positive functionals F_μ on $C(S(\mathcal{B}))$ such that $F_\mu(1) = 1$ (according to Riesz representation theorem). As $V(\mathcal{B})$ is dense in $C(S(\mathcal{B}))$, every positive functional on $C(S(\mathcal{B}))$ is uniquely determined by its restriction on $V(\mathcal{B})$, and conversely every positive functional on $V(\mathcal{B})$ uniquely extend (by continuity) to a positive functional on $C(S(\mathcal{B}))$. Moreover, a norm 1 positive functional on $V(\mathcal{B})$ is uniquely determined by its values on the generating set $\{\mathbf{1}_{K(\phi)} : \phi \in \mathcal{B}\}$, that is by the function $f_\mu : \mathcal{B} \rightarrow [0, 1]$ defined by $f_\mu(\phi) = \mu(K(\phi))$. It is easily checked that f_μ satisfies $f_\mu(1) = 1$ and

$$\phi \wedge \psi = 0 \implies f_\mu(\phi \vee \psi) = f_\mu(\phi) + f_\mu(\psi),$$

and that every function f satisfying these inequalities uniquely defines a norm 1 positive functional on $V(\mathcal{B})$, hence a measure μ on $S(\mathcal{B})$, such that $f = f_\mu$. \square

It follows that finite σ -structures embed injectively in the space $P(S(\mathcal{B}))$ of probability distributions on $S(\mathcal{B})$: to a structure \mathbf{A} with domain A we associate the unique measure $\mu_{\mathbf{A}}$ such that for every $\phi \in \text{FO}(\sigma)$ with free variables x_1, \dots, x_p it holds

$$\int_{S(\mathcal{B})} \mathbf{1}_{K(\phi)} d\mu_{\mathbf{A}}^{\otimes p} = \langle \phi, \mathbf{A} \rangle.$$

Then, a sequence $(\mathbf{A}_n)_{n \in \mathbb{N}}$ is FO-convergent if and only if the probability measures $\mu_{\mathbf{A}_n}$ converge weakly.

More can be said when considering convergence of finite σ -structures. The group S_ω of permutations of \mathbb{N} acts naturally on $\text{FO}(\sigma)$ by permuting the free variables: for a formula ϕ with free variables x_{i_1}, \dots, x_{i_p} and for a permutation τ , the formula $\tau \cdot \phi$ with free variables x_{j_1}, \dots, x_{j_p} where $j_k = \tau(i_k)$ is defined by

$$(\tau \cdot \phi)(x_{j_1}, \dots, x_{j_p}) := \phi(x_{i_1}, \dots, x_{i_p}).$$

It follows that S_ω acts on the Boolean algebra \mathcal{B} , and the action $S_\omega \curvearrowright \mathcal{B}$ defines an action $S_\omega \curvearrowright S(\mathcal{B})$ on the Stone space of \mathcal{B} . As it is clear that for every σ -structure \mathbf{A} , every permutation τ , and every first-order formula ϕ it holds

$$\langle \tau \cdot \phi, \mathbf{A} \rangle = \langle \phi, \mathbf{A} \rangle.$$

Hence if a sequence $(\mathbf{A}_n)_{n \in \mathbb{N}}$ is FO-convergent, then the measures $\mu_{\mathbf{A}_n}$ converge weakly to some S_ω -invariant probability measure on $S(\mathcal{B})$. More generally, every injection $f : \mathbb{N} \rightarrow \mathbb{N}$ defines a transformation $P_f : S(\mathcal{B}) \rightarrow S(\mathcal{B})$, which is continuous, which is a homeomorphism if (and only if) f is bijective (that is a permutation on \mathbb{N}). The set of the transformations P_f has a monoid structure, as $P_f \circ P_g = P_{f \circ g}$. Every measure μ that is obtained either from a finite structure or as weak limit of measures associated to finite structures is invariant under every P_f . Hence $(S(\mathcal{B}), \Sigma_{S(\mathcal{B})}, \mu, P)$ is a measure preserving dynamical system, where $\Sigma_{S(\mathcal{B})}$ is the Borel σ -algebra of $S(\mathcal{B})$ and P is the monoid of transformations indexed by injections $f : \mathbb{N} \rightarrow \mathbb{N}$.

We summarize the situation as follows:

Theorem 2. *Let σ be an at most countable signature, let $X \subseteq \text{FO}(\sigma)$ be a fragment of first-order logic, let \mathcal{B} be the Lindenbaum-Tarski algebra of X , let S be the Stone dual of \mathcal{B} , and let $P(S)$ be the space of probability distributions on S . Then:*

To each σ structure \mathbf{A} corresponds a measure $\mu_{\mathbf{A}}$ such that for every formula $\phi \in X$ with free variables in $\{x_1, \dots, x_p\}$ it holds

$$\int_S \mathbf{1}_{K(\phi)} d\mu_{\mathbf{A}} = \langle \phi, \mathbf{A} \rangle;$$

A sequence $(\mathbf{A}_n)_{n \in \mathbb{N}}$ of σ -structures is X -convergent if and only if the associated measures $\mu_{\mathbf{A}_n}$ converge weakly to some measure $\mu \in P(S)$. In such a case, for every formula $\phi \in X$ with free variables in $\{x_1, \dots, x_p\}$ it holds

$$\lim_{n \rightarrow \infty} \langle \phi, \mathbf{A}_n \rangle = \int_S \mathbf{1}_{K(\phi)} d\mu.$$

Moreover, if Γ is a group of automorphisms of \mathcal{B} (for instance the group of permutations of the free variables), then Γ acts naturally on S ; if all the measures $\mu_{\mathbf{A}}$ associated to finite σ -structures are Γ -invariant, then so are all the measures obtained as weak limits of measures associated to finite σ -structures.

The proof is an easy modification of the proof given in [30], where we proved Theorem 2 without the appendix on group.

2.3. Limits with respect to Parameters. The above abstract setting can be combined with ideas underlying left limits of graphs (as explained in [23]) and, particularly, by property testing. The universal framework, complementing the approach of [30], and using some of the above theory, is to build a notion of limits as follows:

We consider a set \mathcal{O} , whose elements we shall call *objects*, and a set \mathcal{F} of bounded mappings $f : \mathcal{O} \rightarrow \mathbb{R}$, which we call *parameters*. Then we define the notion convergence of a sequence of objects with respect to the family of parameters.

Because we are looking for a notion of limits with respect with the mappings in \mathcal{F} , we consider on \mathcal{O} the *initial topology* with respect to \mathcal{F} , that is the coarsest topology that makes every mapping $f \in \mathcal{F}$ continuous. Note that \mathcal{O} is not Hausdorff in general. Then the space $C_b(\mathcal{O}, \mathbb{R})$ of bounded continuous functions on \mathcal{O}

with sup norm $\|f\| = \sup_{O \in \mathcal{O}} |f(O)|$ is a commutative Banach algebra (that can be extended to a commutative C^* -algebra). According to Gelfand representation theorem, $C_b(\mathcal{O}, \mathbb{R})$ is isometrically isomorphic to the algebra of continuous functions on a compact Hausdorff space, which is known to be homeomorphic to the Stone-Čech compactification $\beta\mathcal{O}$ of \mathcal{O} . Recall that the Stone-Čech compactification $\beta\mathcal{O}$ of \mathcal{O} is the largest compact Hausdorff space generated by \mathcal{O} , in the sense that any map from \mathcal{O} to a compact Hausdorff space factors (in a unique way) through $\beta\mathcal{O}$.

Let $\text{Profile} : \mathcal{O} \rightarrow [0, 1]^{\mathcal{F}}$ denote the mapping defined by

$$\text{Profile}(O) = (f(O))_{f \in \mathcal{F}},$$

and let \sim be the equivalence relation

$$O_1 \sim O_2 \iff \text{Profile}(O_1) = \text{Profile}(O_2).$$

Then $K\mathcal{O} = \mathcal{O} / \sim$ is the Kolmogorov quotient of \mathcal{O} , and $\beta\mathcal{O}$ is homeomorphic to the closure of $\{\text{Profile}(O) : O \in \mathcal{O}\}$ in $[0, 1]^{\mathcal{F}}$.

This leads to the following natural notion of convergence of a sequence of objects of \mathcal{O} with respect to parameters in \mathcal{F} :

Definition 3. A sequence $(O_n)_{n \in \mathbb{N}}$ is \mathcal{F} -convergent if, for every $f \in \mathcal{F}$ the sequence $(f(O_n))_{n \in \mathbb{N}}$ is convergent.

Note that the limit of an \mathcal{F} -convergent sequence of objects in \mathcal{O} is uniquely defined as a point of $\beta\mathcal{O}$.

If $P \in \mathbb{R}[x_1, \dots, x_p]$ is a real polynomial and $f_1, \dots, f_p \in \mathcal{F}$, we denote by $P(f_1, \dots, f_p)$ the mapping from \mathcal{O} to \mathbb{R} defined by

$$P(f_1, \dots, f_p) : O \mapsto P(f_1(O), \dots, f_p(O)).$$

Let $A(\mathcal{O}, \mathcal{F})$ be the sub-algebra of $C_b(\mathcal{O})$ formed by above mappings. Stone-Weierstrass theorem on uniform approximation of continuous functions by polynomials extends in this setting (and this is a folklore):

Lemma 4. The sub-algebra $A(\mathcal{O}, \mathcal{F})$ is dense in $C_b(\mathcal{O}) = C(\beta\mathcal{O})$.

Proof. The algebra $A(\mathcal{O}, \mathcal{F})$ contains a non-zero constant (because of constant polynomials) and $A(\mathcal{O}, \mathcal{F})$ separates the points of $\beta\mathcal{O}$ (as it separates the points of $K\mathcal{O}$). Thus the result follows from Stone-Weierstrass theorem. \square

Note that above results are very general, as we made no assumptions on our objects or parameters. Nevertheless, even that has some direct applications of this point of view to the theory of left-convergence.

Let $\text{hom}(F, G)$ denote the number of homomorphisms from F to G , and let $t(F, G)$ be the *homomorphism density* of F in G defined by

$$t(F, G) = \frac{\text{hom}(F, G)}{|G|^{|F|}},$$

that is the probability that a random map from F to G is a homomorphism.

Proposition 1. Let \mathcal{O} be the set of all finite graphs, and let \mathcal{F} be the set of all the mappings $G \mapsto t(F, G)$ considered as mappings

$$t(F, \cdot) : \mathcal{O} \rightarrow [0, 1].$$

Then the notion of convergence defined above is the left-convergence, and $\beta\mathcal{O}$ is homeomorphic to the space of graphons.

Let \mathcal{SQ} denote the algebra of *standard quantum graphs*, that is the commutative algebra constructed from vector space with basis formed by $\{K_1\}$ and all the finite graphs without isolated vertices, by defining multiplication by linearity from the case of finite graphs, defining the unit of the algebra as K_1 , and the product of two graphs without isolated vertices F_1 and F_2 as their disjoint union.

We shall make use of the following Lemma [23, Corollary 5.45, p. 74]:

Lemma 5. *The graph parameters $\text{hom}(F, \cdot)$ (where F ranges over simple graphs) are linearly independent. Equivalently, the graph parameters $\text{hom}(F, \cdot)$ (where F ranges over connected simple graphs) are algebraically independent.*

Proposition 2. *Let \mathcal{O} be the set of all finite graphs, and let \mathcal{F} be the set of homomorphism densities*

$$t(F, \cdot) : \mathcal{O} \rightarrow [0, 1].$$

Then the mapping

$$T : \mathcal{SQ} \rightarrow A(\mathcal{O}, \mathcal{F}) \quad \sum_{i=1}^{\ell} a_i F_i \mapsto \sum_{i=1}^{\ell} a_i t(F_i, \cdot)$$

is an algebra isomorphism of \mathcal{SQ} and $A(\mathcal{O}, \mathcal{F})$.

Proof. The algebraic independence of the parameters $t(F, \cdot)$ for graphs without isolated vertices will follow from the linear independence of the parameters $t(F, \cdot)$ for simple graphs F either equal to K_1 or without isolated vertices, which we now prove:

Assume that there exist non-isomorphic simple graphs F_1, \dots, F_ℓ (all of them being either K_1 or without isolated vertices) and reals a_1, \dots, a_ℓ such that

$$\sum_{i=1}^{\ell} a_i t(F_i, G) = 0$$

holds for every graph G . Let $N = \max_{i=1}^{\ell} |F_i|$, and let F'_i be the graph obtained from F_i by adding $N - |F_i|$ isolated vertices. Then

$$\text{hom}(F'_i, G) = |G|^{|F'_i|} t(F'_i, G) = |G|^N t(F_i, G).$$

Hence the equation $\sum_{i=1}^{\ell} a_i \text{hom}(F'_i, G) = 0$ holds for every graph G . As no two graphs F'_i are isomorphic, it follows from Lemma 5 that a_1, \dots, a_ℓ are all zero. \square

Let us give two direct consequences of this lemma. The first consequence can be seen as an analog Stone Weierstrass uniform approximation of a continuous function by polynomials.

Say that a graph parameter f is *left continuous* if the sequence $(f(G_n))_{n \in \mathbb{N}}$ converges for every left convergent sequence of graphs $(G_n)_{n \in \mathbb{N}}$.

Proposition 3. *For every left-continuous graph parameter f and every $\epsilon > 0$ there exists a quantum graph $\mathbf{F} = \sum_{i=1}^{\ell} a_i F_i$ such that for every graph G it holds*

$$|t(\mathbf{F}, G) - f(G)| < \epsilon.$$

Proof. This is an immediate consequence of the density of $T(\mathcal{SQ})$ in $C(\beta\mathcal{O})$. \square

Say that a quantum graph \mathbf{F} is *non-negative* (which we denote by $\mathbf{F} \geq 0$) if $t(\mathbf{F}, G) \geq 0$ for every graph G . The non-negativity of quantum graphs was subject to intensive study (see [23]). The second consequence gives yet another characterization of this property.

Proposition 4. *Let \mathbf{F} be a quantum graph. Then $\mathbf{F} \geq 0$ if and only if for every $\epsilon > 0$ there exists a quantum graph \mathbf{H} such that $\mathbf{H} \geq 0$ and $\|\mathbf{F} - \mathbf{H}^2\| < \epsilon$, that is:*

$$\inf_G t(\mathbf{H}, G) \geq 0 \quad \text{and} \quad \sup_G |t(\mathbf{F}, G) - t(\mathbf{H}^2, G)| < \epsilon.$$

Proof. One direction is clear. For the other direction, assume $\mathbf{F} \geq 0$. Let h be the left-continuous graph parameter defined by $h(G) = (t(\mathbf{F}, G) + \epsilon/2)^{1/2}$. Let \mathbf{H} be a quantum graph such that $\|h - T(\mathbf{H})\| < \alpha$, where $\alpha(\|\mathbf{F} + \epsilon/2K_1\|^{1/2} + \alpha) < \epsilon/2$. Then

$$\inf_G t(\mathbf{H}, G) \geq \inf_G t(h, G) - \alpha \geq (\epsilon/2)^{1/2} - \alpha \geq 0.$$

Moreover,

$$\|(\mathbf{F} + \epsilon/2K_1) - \mathbf{H}^2\| = \|h^2 - T(\mathbf{H}^2)\| \leq \|h + T(\mathbf{H})\| \|h - T(\mathbf{H})\| \leq (2\|h\| + \alpha)\alpha < \epsilon/2.$$

Thus

$$\|\mathbf{F} - \mathbf{H}^2\| \leq \|(\mathbf{F} + \epsilon/2K_1) - \mathbf{H}^2\| + \|\epsilon/2K_1\| < \epsilon.$$

□

It is also pleasant to note that the sup norm defined on $C_b(\mathcal{O})$ defines naturally (by identification of \mathcal{SQ} and $A(\mathcal{O}, \mathcal{F})$) defines a norm

$$\left\| \sum_i a_i F_i \right\| = \sup_{G \in \mathcal{O}} \left| \sum_i a_i t(F_i, G) \right|$$

for quantum graphs.

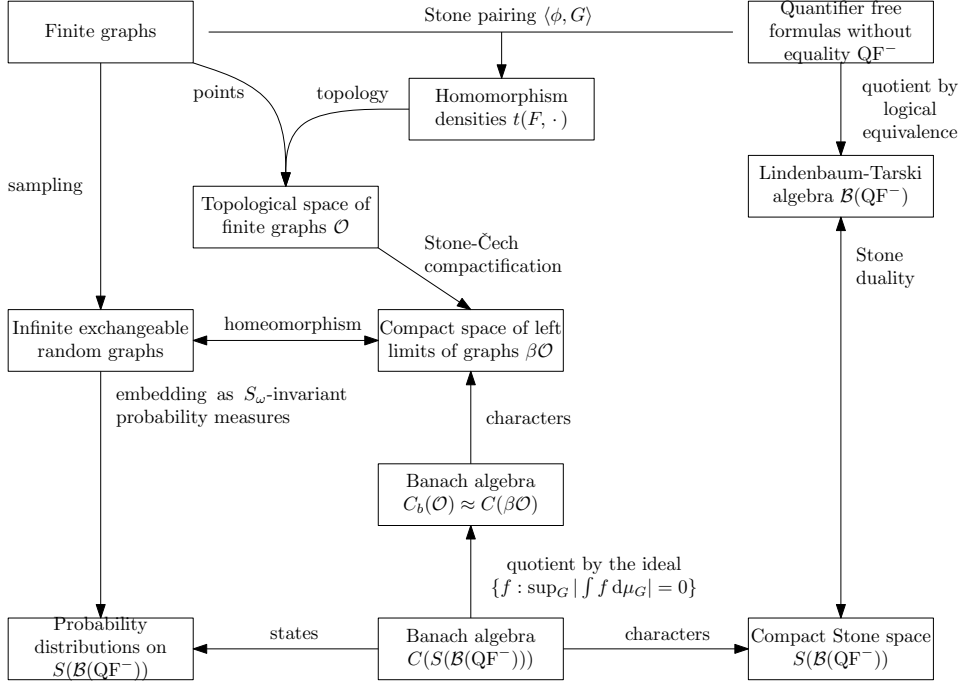
As we shall see later on, left-convergence of graphs is equivalent to (structural) QF^- -convergence, where QF^- is the fragment of quantifier-free formula without equality. The global situation, which we just outlined, is depicted on Fig. 3

3. FRAGMENTS OF INTEREST

We review here a more detailed analysis of the convergence notions based of Lindenbaum-Tarski algebras of different fragments, in the context of Section 2 and of our earlier approach [30]. We consider the following fragments:

- FO_0 : the fragment of sentences;
- FO_p : the fragment of formulas with free variables included in $\{x_1, \dots, x_p\}$;
- QF : the fragment of quantifier-free formulas;
- FO^{local} : the fragment of local formulas;
- $\text{FO}_1^{\text{local}}$: the fragment of local formulas with single free variable.

Let us review some particular important cases of our theory induced by specific fragments of first-order logic.

FIGURE 3. Left-convergence and Structural QF^- -convergence.

3.1. Elementary Limits. Our starting case consists in considering *elementary convergence*: A sequence $(\mathbf{A}_n)_{n \in \mathbb{N}}$ is *elementary convergent* if, for every sentence (that is formulas without any free variables) θ there is an integer N such that either all the \mathbf{A}_n with $n \geq N$ satisfy θ , or no \mathbf{A}_n with $n \geq N$ satisfies θ . In other words, elementary convergence is convergence defined by the fragment $\text{FO}_0(\sigma)$ of first-order sentences. The Stone dual S of the Lindenbaum-Tarski algebra of first-order sentences is the space of complete theories. Recall that a *complete theory* is a maximal consistent set of sentences. Gödel's celebrated completeness theorem of first-order logic asserts that every complete theory has a model. In other words, for every consistent set $T \subset \text{FO}_0(\sigma)$ of sentences there exists a σ -structure \mathbf{A} that is a *model* of T , that is such that \mathbf{A} satisfies every sentence in T . Moreover, according to Löwenheim-Skolem theorem, if the signature σ is at most countable then every consistent theory $T \subset \text{FO}_0(\sigma)$ has a model whose domain is at most countable.

Conversely, to each σ -structure \mathbf{A} one associates the *complete theory* $\text{Th}(\mathbf{A})$ of \mathbf{A} , which is the set of all the sentences satisfied by \mathbf{A} . Although it is easily checked to the mapping $\mathbf{A} \mapsto \text{Th}(\mathbf{A})$ is 1 – 1 on finite structures, this is not the case for infinite structures; two structures \mathbf{A} and \mathbf{B} sharing the same complete theory are said to be *elementary equivalent*.

In such a context, the Stone bracket $\langle \phi, \mathbf{A} \rangle$ gets only values 0 (if $\mathbf{A} \not\models \phi$) or 1 (if $\mathbf{A} \models \phi$), and the measure $\mu_{\mathbf{A}}$ associated to \mathbf{A} is a Dirac measure at the point $\text{Th}(\mathbf{A})$. The class of all finite σ -structures can accordingly be identified with an open subset O of S . The elementary limits of FO_0 -convergent sequences of finite σ -structures correspond to the points of the closure of O , which correspond to the

complete theories T having the *finite model property*, that is the complete theories T such a every finite subset of T has a finite model.

It follows that every FO_0 -convergent sequence of finite σ -structures has a limit that can be represented as a σ -structure with an at most countable domain.

Note that characterizing elementary limits of finite structures is essentially difficult, as finite model property is not decidable in general. However, it is worth noticing that first-order formulas without functional symbols where all existential quantifications appear first in the formula, has the finite model property. Precisely, Ramsey [35] showed that a sentence

$$\exists x_1 \dots \exists x_n \forall y_1 \dots \forall y_m \phi$$

(where ϕ is quantifier-free and has no function symbols) has a model if and only if it has a model of size bounded by n plus the number of constants in ϕ .

3.2. Left limits and the Quantifier Free Fragment. We can precise our intuition by considering the case of undirected graphs with convergence driven by the fragment QF^- of quantifier free formulas without equality (meaning that the equality symbol cannot be used). It is easily checked that the corresponding notion of convergence of a sequence $(G_n)_{n \in \mathbb{N}}$ of graphs is nothing but the *left-convergence* introduced by Lovász *et al.*, which is based on the convergence of the *homomorphism profile* t , which associates to each finite graph F the probability $t(F, G_n)$ that a random map from F to G_n is a homomorphism (that is an adjacency-preserving map). The Stone dual $S(\text{QF}^-)$ of the Lindenbaum-Tarski algebra of QF^- can be described as follows: each point $p \in S(\text{QF}^-)$ is a maximal consistent set of formulas QF^- , hence is uniquely determined by either $x_i \sim x_j$ or $\neg(x_i \sim x_j)$ for each pair $(i, j) \in \mathbb{N}^2$, stating that x_i is adjacent or not to x_j . Each point p can thus be represented by means of a (labeled) countable graph with vertex set \mathbb{N} , and the left limit of a sequence of graphs can be represented as an S_ω -invariant probability distribution on the Stone space, that is as an *infinite exchangeable random graph*. Then it follows from Aldous and Hoover extensions of de Finetti's theorem to exchangeable arrays [2, 19] that left limits of graphs can be represented by means of a *graphon*, that is a symmetric measurable function $W : [0, 1] \times [0, 1] \rightarrow [0, 1]$. However, note that in the case of the fragment QF^- , the mapping $G \mapsto \mu_G$ is not injective, as two vertices that can be uniformly blown up in order to obtain a same graph are mapped to the same probability measure. This special aspect is due to the weakness of expressive power of the QF^- fragment, and it disappears when one considers the fragment QF of all quantifier-free formulas, which essentially defines the same notion of convergence.

3.3. Local limits and the $\text{FO}_1^{\text{local}}$ fragment. We now consider the case of undirected graphs with degree at most D , and the fragment $\text{FO}_1^{\text{local}}$ of local formulas with a single free variable. Recall that a formula ϕ is *t-local* if its satisfaction only depends on a distance t -neighborhood of the free variables, and that ϕ is *local* if it is *t-local* for some $t \in \mathbb{N}$.

Local convergence of graphs with bounded degree has been defined by Benjamini and Schramm [5]. In our setting, local convergence can be defined as follows: A sequence $(G_n)_{n \in \mathbb{N}}$ of graphs with maximum degree D is *locally-convergent* if, for every $r \in \mathbb{N}$, the distribution of the isomorphism types of the distance r -neighborhood of a random vertex of G_n converges as n goes to infinity. Local convergence witnesses

the particular importance of the fragment $\text{FO}_1^{\text{local}}$: As each isomorphism type T of the distance r -neighborhood of a vertex can be characterized by the satisfaction of an r -local formula $\phi_T(x)$ (where x stands for the root vertex), it follows that $\text{FO}_1^{\text{local}}$ -convergence implies local convergence. Conversely, as the satisfaction of an r -local formula with a single free variable only depends on the isomorphism type of the (rooted) distance r -neighborhood of the free variable, $\text{FO}_1^{\text{local}}$ -convergence is equivalent local convergence.

In general, points of the Stone dual of $\text{FO}_1^{\text{local}}$ correspond to the complete theories of rooted connected structures. As we consider graphs with bounded degrees, we know that two rooted connected countable graphs with maximum degree at most D are elementarily equivalent if and only if they are isomorphic. It follows that in this case the Stone dual of $\text{FO}_1^{\text{local}}$ is the space of all (isomorphism classes of) rooted connected countable graphs with maximum degree D .

3.4. Interrelations of Fragments and Reductions. A peculiar feature of the fragments FO_0 and $\text{FO}_1^{\text{local}}$, stated as Theorem 7, is a consequence of Gaifman locality theorem [14]. This we recall now:

Theorem 6. *For every first-order formula $\phi(x_1, \dots, x_n)$ there exist integers t and r such that ϕ is equivalent to a Boolean combination of t -local formulas $\xi_s(x_{i_1}, \dots, x_{i_s})$ and sentences of the form*

$$(8) \quad \exists y_1 \dots \exists y_m \left(\bigwedge_{1 \leq i < j \leq m} \text{dist}(y_i, y_j) > 2r \wedge \bigwedge_{1 \leq i \leq m} \psi(y_i) \right)$$

where ψ is r -local. Furthermore, we can choose

$$r \leq 7^{\text{qrang}(\phi)-1}, \quad t \leq (7^{\text{qrang}(\phi)-1} - 1)/2, \quad m \leq n + \text{qrang}(\phi),$$

and, if ϕ is a sentence, only sentences (8) occur in the Boolean combination. Moreover, these sentences can be chosen with quantifier rank at most $f(\text{qrang}(\phi))$, for some fixed function f .

The proof of the following reduction theorem essentially uses Theorem 6.

Theorem 7 ([30]). *Let (\mathbf{A}_n) be a sequence of finite λ -structures. Then (\mathbf{A}_n) is FO -convergent if and only if it is both $\text{FO}_1^{\text{local}}$ -convergent and FO_0 -convergent. Precisely, (\mathbf{A}_n) is FO_p -convergent if and only if it is both $\text{FO}_p^{\text{local}}$ -convergent and FO_0 -convergent.*

Of course, if we constrain the class of structures on which we study limits, we get stronger reductions. For example, we now prove that in the case of a sequence of vertex-transitive structures in a wide class, most of the hierarchy of X -convergence collapses. Recall that a class of graphs \mathcal{C} is *wide* if, for every integers d, m there is an integer $N(d, m)$ such that every graph $G \in \mathcal{C}$ with order at least $N(d, m)$ contains a subset of m vertices pairwise at distance at least d [4]. (In particular, every class of graphs with bounded degree is wide.)

Lemma 8. *For every sentence θ there exists constants r, m and a formula $\phi \in \text{FO}_1^{\text{local}}$ such that for every vertex-transitive λ -structure \mathbf{A} with at least m vertices pairwise at distance greater than $2r$ it holds*

$$\mathbf{A} \models \theta \iff \mathbf{A} \models \exists x \phi(x) \iff \forall x \phi(x) \iff \langle \phi, \mathbf{A} \rangle = 1.$$

Proof. If \mathbf{A} is vertex-transitive then for every $u, v \in A$ and every $\phi \in \text{FO}_1$ it holds $\mathbf{A} \models \phi(u)$ if and only if $\mathbf{A} \models \phi(v)$. Hence $\mathbf{A} \models \exists x \phi(x)$ is equivalent to $\mathbf{A} \models \forall x \phi(x)$.

According to Theorem 6, the sentence θ is equivalent to a Boolean combination of sentences of the form

$$\begin{aligned} \mathbf{A} \models \exists y_1 \dots \exists y_m & \left(\bigwedge_{1 \leq i < j \leq m} \text{dist}(y_i, y_j) > 2r \wedge \bigwedge_{1 \leq i \leq m} \psi(y_i) \right) \\ \iff \mathbf{A} \models \exists y_1 \dots \exists y_m & \left(\bigwedge_{1 \leq i < j \leq m} \text{dist}(y_i, y_j) > 2r \right) \wedge \exists y \bigwedge_{1 \leq i \leq m} \psi(y) \\ \iff \mathbf{A} \models \exists y & \bigwedge_{1 \leq i \leq m} \psi(y) \end{aligned}$$

(where the first equivalence comes from vertex-transitivity, and the second from the existence of m vertices far apart.) Moreover, by vertex transitivity, for every $\phi_1, \phi_2 \in \text{FO}_1^{\text{local}}$ it holds

$$\mathbf{A} \models (\exists x_1 \phi_1(x_1)) \wedge (\exists x_2 \phi_2(x_2)) \iff \mathbf{A} \models (\exists x \phi_1(x) \wedge \phi_2(x)).$$

It follows that the Boolean combination of the conditions $\mathbf{A} \models \exists y \bigwedge_{1 \leq i \leq m} \psi(y)$ is equivalent to the condition $\mathbf{A} \models \exists x \phi(x)$ for some $\phi \in \text{FO}_1^{\text{local}}$. \square

As a consequence, we obtain the following:

Theorem 9. *Let \mathcal{C} be a wide class of vertex-transitive λ -structures. Then, for every sequence $(\mathbf{A}_n)_{n \in \mathbb{N}}$ with $\mathbf{A}_n \in \mathcal{C}$ and $|A_n| \rightarrow \infty$ the following conditions are equivalent:*

- (1) *the sequence $(\mathbf{A}_n)_{n \in \mathbb{N}}$ is FO-convergent,*
- (2) *the sequence $(\mathbf{A}_n)_{n \in \mathbb{N}}$ is $\text{FO}_1^{\text{local}}$ -convergent,*
- (3) *for every $\phi \in \text{FO}_1^{\text{local}}$, either $(\exists x) \phi(x)$ is satisfied by all but finitely many \mathbf{A}_n or $(\exists x) \phi(x)$ is satisfied by only finitely many \mathbf{A}_n ,*
- (4) *the sequence $(\mathbf{A}_n)_{n \in \mathbb{N}}$ is FO_0 -convergent.*

Proof. That (1) implies (2) is trivial. That (2) is equivalent to (3) follows from vertex-transitivity. That (3) implies (4) follows from Lemma 8 and that (4) implies (3) is trivial. That (2) implies $\text{FO}_1^{\text{local}}$ convergence is a consequence of Gaifman locality theorem and wideness of \mathcal{C} . Finally, that (4) together with $\text{FO}_1^{\text{local}}$ convergence implies (1) follows from Theorem 7. \square

Let us formulate the following consequence for Cayley graphs.

Corollary 1. *Let $d \in \mathbb{N}$, and let $(G_n)_{n \in \mathbb{N}}$ be a sequence of (edge-colored directed) Cayley graphs generated by d elements of groups with orders going to infinity. Then the following conditions are equivalent:*

- (1) *the sequence $(G_n)_{n \in \mathbb{N}}$ is FO-convergent,*
- (2) *the sequence $(G_n)_{n \in \mathbb{N}}$ is $\text{FO}_1^{\text{local}}$ -convergent,*
- (3) *the sequence $(G_n)_{n \in \mathbb{N}}$ is FO_0 -convergent.*

We have already noticed (Theorem 7) that FO-convergence can be reduced to FO_0 -convergence and to $\text{FO}_1^{\text{local}}$ -convergence (thanks to Gaifman locality theorem). Furthermore, for wide classes (like classes with bounded degree), $\text{FO}_1^{\text{local}}$ -convergence is equivalent to $\text{FO}_1^{\text{local}}$ -convergence. This explains the particular importance of the fragments $\text{FO}_1^{\text{local}}$ and $\text{FO}_1^{\text{local}}$.

4. LIMIT CONNECTIVITY

The importance of the fragments reviewed in the previous section also appears in the study of the limit notion of “connected components” for a convergent sequence of λ -structures. Note that connectivity of structures (which fails to be first-order definable) plays an important and non-trivial aspect of convergence of sparse graphs. The relative size and the speed of convergence plays a role here. Part of this problem is expressed by the following notions:

For a σ -structure \mathbf{A} , a vertex $v \in A$, and an integer d , we define $B_d(\mathbf{A}, v)$ as the substructure of \mathbf{A} induced by vertices at distance at most d from v in \mathbf{A} (that is at distance at most d from v in the Gaifman graph of \mathbf{A}).

Definition 10 (Residual sequence). A sequence $(\mathbf{A}_n)_{n \in \mathbb{N}}$ of λ -structures is *residual* if

$$\lim_{d \rightarrow \infty} \limsup_{n \rightarrow \infty} \sup_{v \in A_n} \nu_{\mathbf{A}_n}(B_d(\mathbf{A}_n, v)) = 0.$$

In other words, a sequence is residual if its limit has only zero-measure connected components

Definition 11. A sequence $(\mathbf{A}_n)_{n \in \mathbb{N}}$ of λ -structures is *non-dispersive* if

$$\lim_{d \rightarrow \infty} \liminf_{n \rightarrow \infty} \sup_{v \in A_n} \nu_{\mathbf{A}_n}(B_d(\mathbf{A}_n, v)) = 1.$$

In other words, a non-dispersive sequence is a sequence whose limit is essentially connected.

For a σ -structure \mathbf{A} and a vertex $\rho \in A$, we denote by (\mathbf{A}, ρ) the rooting of the structure \mathbf{A} at ρ , that is the σ^\bullet -structure (where σ^\bullet is obtained by adding a unary symbol R to signature σ) obtained from \mathbf{A} by putting exactly ρ in relation R .

In the case of rooted structures, we usually want a stronger statement that the structures remain concentrated around their roots: a sequence $(\mathbf{A}_n, \rho_n)_{n \in \mathbb{N}}$ of rooted λ -structures is *ρ -non-dispersive* if

$$\lim_{d \rightarrow \infty} \liminf_{n \rightarrow \infty} \nu_{\mathbf{A}_n}(B_d(\mathbf{A}_n, \rho_n)) = 1.$$

The notion of non-dispersiveness is only an approximation of connectiveness. As such it is not supported by a single equivalence relation. However, we can introduce the notion of a component relation system, which approximates component equivalence.

Definition 12. A *component relation system* for a class \mathcal{C} of λ -structures is a sequence ϖ_d of equivalence relations such that for every $d \in \mathbb{N}$ and for every $\mathbf{A} \in \mathcal{C}$ there is a partition of the ϖ_d -equivalence classes of A into two parts $\mathcal{E}_0(\varpi_d, \mathbf{A})$ and $\mathcal{E}_+(\varpi_d, \mathbf{A})$ such that:

- every class in $\mathcal{E}_0(\varpi_d, \mathbf{A})$ is a singleton;
- $\nu_{\mathbf{A}}(\bigcup \mathcal{E}_0(\varpi_d, \mathbf{A})) < \epsilon(d) + \eta(|A|)$ (where $\lim_{d \rightarrow \infty} \epsilon(d) = \lim_{n \rightarrow \infty} \eta(n) = 0$);
- two vertices x, y in $\bigcup \mathcal{E}_+(\varpi_d, \mathbf{A})$ belong to a same connected component of \mathbf{A} if and only if $\mathbf{A} \models \varpi_d(x, y)$ (i.e. iff x and y belong to a same class).

Definition 13 ([33]). A family of sequence $(\mathbf{A}_{i,n})_{n \in \mathbb{N}}$ ($i \in I$) of λ -structures is *uniformly elementarily convergent* if, for every formula $\phi \in \text{FO}_1(\lambda)$ there is an integer N such that it holds

$$\forall i \in I, \forall n' \geq n \geq N, \quad (\mathbf{A}_{i,n} \models (\exists x)\phi(x)) \implies (\mathbf{A}_{i,n'} \models (\exists x)\phi(x)).$$

(Note that if a family $(\mathbf{A}_{i,n})_{n \in \mathbb{N}}$ ($i \in I$) of sequences is uniformly elementarily convergent, then each sequence $(\mathbf{A}_{i,n})_{n \in \mathbb{N}}$ is elementarily convergent.)

Now that we have introduced all the necessary notions, we can state the following fundamental result, which allows to study FO^{local} -convergent sequences (and more generally FO -convergent sequences) of disconnected λ -structures essentially component-wise.

Theorem 14 (Extended comb structure [32]). *Let $(\mathbf{A}_n)_{n \in \mathbb{N}}$ be an FO^{local} -convergent sequence of finite λ -structures with component relation system ϖ_d .*

Then there exist $I \subseteq \mathbb{N} \cup \{0\}$ and, for each $i \in I$, a real α_i and a sequence $(\mathbf{B}_{i,n})_{n \in \mathbb{N}}$ of λ -structures, such that $\mathbf{A}_n = \bigcup_{i \in I} \mathbf{B}_{i,n}$ (for all $n \in \mathbb{N}$), $\sum_{i \in I} \alpha_i = 1$, and for each $i \in I$ it holds

- $\alpha_i = \lim_{n \rightarrow \infty} \frac{|\mathbf{B}_{i,n}|}{|\mathbf{A}_n|}$, and $\alpha_i > 0$ if $i \neq 0$;
- if $i = 0$ and $\alpha_0 > 0$ then $(\mathbf{B}_{i,n})_{n \in \mathbb{N}}$ is FO^{local} -convergent and residual;
- if $i > 0$ then $(\mathbf{B}_{i,n})_{n \in \mathbb{N}}$ is FO^{local} -convergent and non-dispersive.

Moreover, if $(\mathbf{A}_n)_{n \in \mathbb{N}}$ is FO -convergent we can require the family $\{(\mathbf{B}_{i,n})_{n \in \mathbb{N}} : i \in I\}$ to be uniformly elementarily-convergent.

The proof of Theorem 14 is elaborated as it needs a detailed analysis of the speed of growth of components and we refer to [32] for proof. An important consequence of this theorem is to reduce the problem of the construction of limit objects to the basic cases of residual sequences and non-dispersive sequences, see for instance Theorem 19.

5. LIMIT OBJECTS OF SPARSE STRUCTURES

5.1. Borel Structures. Given a signature σ , *Borel σ -structure* is a σ -structure \mathbf{A} , whose domain A is a standard Borel space, such that:

- for every relational symbol $R \in \sigma$ with arity r , the set

$$R^{\mathbf{A}} = \{(v_1, \dots, v_r) \in A^r : \mathbf{A} \models R(v_1, \dots, v_r)\}$$

of R in \mathbf{A} is a Borel subset of A^r ;

- for every functional symbol $f \in \sigma$ with arity r , the corresponding map $f^{\mathbf{A}} : A^r \rightarrow A$ is a Borel map.

Note that Borel structures are a natural generalization of Borel graphs (see [20]).

It is easily checked that if \mathbf{A} is a Borel σ -structure and if ν is a probability measure on A , then for every quantifier-free formula $\phi \in \text{QF}(\sigma)$, the set $\phi(\mathbf{A})$ is Borel. Whence we can extend the definition of Stone bracket to pairs formed by a Borel structure \mathbf{A} equipped with a probability measure ν and a quantifier free formula ϕ with free variables in x_1, \dots, x_p by

$$\langle \phi, (\mathbf{A}, \nu) \rangle = \nu^{\otimes p}(\phi(\mathbf{A})).$$

Then, we shall say that a Borel σ -structure \mathbf{A} equipped with a probability measure ν is a *QF-limit* of a sequence $(\mathbf{F}_n)_{n \in \mathbb{N}}$ of finite σ -structures if for every quantifier free formula ϕ it holds

$$\langle \phi, (\mathbf{A}, \nu) \rangle = \lim_{n \rightarrow \infty} \langle \phi, \mathbf{F}_n \rangle.$$

Example 15. Let $(G_n)_{n \in \mathbb{N}}$ be a left convergent sequence with $|G_n| \rightarrow \infty$.

Then there exists a Borel graph \mathbf{G} with a probability measure ν such that (\mathbf{G}, ν) is a QF-limit of the sequence $(G_n)_{n \in \mathbb{N}}$ if and only if the sequence $(G_n)_{n \in \mathbb{N}}$ converges to a *random-free* graphon, that is a graph which is $\{0, 1\}$ -valued almost everywhere.

Some more exotic examples can be obtained by considering functional symbols:

Example 16 ([9]). A *tree semi-lattice* is a structure \mathbf{A} whose signature only contains a binary functional symbol \wedge , which satisfies the following axioms:

- (1) the operation (defined by) \wedge is associative, commutative, and idempotent;
- (2) for every x, y, z such that $x \wedge z = x$ and $y \wedge z = y$ it holds $x \wedge y \in \{x, y\}$.

Every QF-convergent sequence $(\mathbf{T}_n)_{n \in \mathbb{N}}$ of tree semi-lattices with k -colored domain there exists a Borel tree semilattice \mathbf{L} with k -colored domain and a probability measure ν such that (\mathbf{L}, ν) is a QF-limit of the sequence $(\mathbf{T}_n)_{n \in \mathbb{N}}$.

Example 17. In this example, let $(G_n)_{n \in \mathbb{N}}$ be a sequence of finite nonabelian groups with increasing orders. Assume each G_n is either simple or is a symmetric group. Then the sequence $(G_n)_{n \in \mathbb{N}}$ is QF-convergent and $\text{GL}(2, \mathbb{R})$ (with Haar probability measure) is a QF-limit of $(G_n)_{n \in \mathbb{N}}$.

Indeed, according to [10], the probability that a non-trivial fixed word w equals identity tends to 0 as n goes to infinity, and the same holds for the sequence of symmetric groups, according to [15]. Thus, for every quantifier formula $\phi(x_1, \dots, x_p)$, the probability that $\phi(x_1, \dots, x_p)$ will hold in G_n for n going to infinity for independent uniform random assignments of x_1, \dots, x_p tends either to 1 or 0, depending on the fact that $\phi(x_1, \dots, x_p)$ holds trivially or not. Also, according to [13], the the probability that a non-trivial fixed word w equals identity is 0 for a connected finite-dimensional nonsolvable Lie groups (for instance $\text{GL}(2, \mathbb{R})$). Thus $\text{GL}(2, \mathbb{R})$ is a QF-limit of $(G_n)_{n \in \mathbb{N}}$.

5.2. Modelings. We introduced in [33] — as candidate for a possible limit object of sparse structures — the notion of modeling, which extends the notion of graphing introduced for bounded degree graphs. A *relational sample space* is a relational structure \mathbf{A} (with signature λ) with additional structure: The domain A of \mathbf{A} is a standard Borel space (with Borel σ -algebra $\Sigma_{\mathbf{A}}$) with the property that every subset of A^p that is first-order definable in $\text{FO}(\lambda)$ is measurable (in A^p with respect to the product σ -algebra). A *modeling* is a relational sample space equipped with a probability measure (denoted $\nu_{\mathbf{A}}$). For brevity we shall use the same letter \mathbf{A} for structure, relational sample space, and modeling. The definition of modelings allows us to extend Stone pairing naturally to modelings: the *Stone pairing* $\langle \phi, \mathbf{A} \rangle$ of a first-order formula ϕ (with free variables in $\{x_1, \dots, x_p\}$) and a modeling \mathbf{A} , is defined by

$$\langle \phi, \mathbf{A} \rangle = \nu^{\otimes p}(\phi(\mathbf{A})),$$

where

$$\phi(\mathbf{A}) = \{(v_1, \dots, v_p) \in A^p : \mathbf{A} \models \phi(v_1, \dots, v_p)\}.$$

Note that every finite structure canonically defines a modeling (with same universe, discrete σ -algebra, and uniform probability measure) and that in the definition above matches the definition of Stone pairing of a formula and a finite structure introduced earlier. Also note that every modeling σ -structure is obviously a Borel σ -structure, but that the converse does not have to hold in general, as projections of Borel sets are not Borel in general.

For a fragment X of FO, we shall say that a sequence $(\mathbf{A}_n)_{n \in \mathbb{N}}$ is X -convergent to a modeling \mathbf{L} (or that \mathbf{L} is a *modeling X -limit* of $(\mathbf{A}_n)_{n \in \mathbb{N}}$), and we shall note $\mathbf{A}_n \xrightarrow{X} \mathbf{L}$, if for every $\phi \in X$ it holds $\lim_{n \rightarrow \infty} \langle \phi, \mathbf{A}_n \rangle = \langle \phi, \mathbf{L} \rangle$. We shall say that a class \mathcal{C} of λ -structures *has modeling X -limits* if, every X -convergent sequence $(\mathbf{A}_n)_{n \in \mathbb{N}}$ of λ -structures with $\mathbf{A}_n \in \mathcal{C}$ has a modeling X -limit. For instance, we proved in [33] that classes of graphs with bounded degrees and classes of graphs with bounded tree-depth have modeling FO-limits, and we proved in [32] that the class of rooted (k vertex-colored) trees has modeling FO-limits. However, the notion of modeling limit is, in some sense, only applicable to sparse structures:

Theorem 18 ([33]). *Let \mathcal{C} be a monotone class of graphs.*

If \mathcal{C} has modeling FO-limits then \mathcal{C} is nowhere dense, that is: for every integer p there exists an integer $N = N(\mathcal{C}, p)$ such that no graph in \mathcal{C} contains the p subdivision of K_N as a subgraph.

For more on nowhere dense graphs, we refer the reader to [27, 28, 29, 31]. The importance of nowhere dense classes and the strong relationship of this notion with first-order logic is exemplified by the recent result of Grohe, Kreutzer, and Siebertz [18], which states that (under a reasonable complexity theoretic assumption) deciding first-order properties of graphs in a monotone class \mathcal{C} is fixed-parameter tractable if and only if \mathcal{C} is nowhere dense. Motivated by strong model theoretical properties and manifold characterizations of nowhere dense classes, we conjectured that the above theorem is tight. This may be seen as the most challenging problem of this paper.

Conjecture 1. Every nowhere dense class of graphs has modeling FO-limits.

The problem of the existence of a modeling limit can be reduced to the study of FO^{local} -convergence, and then to two “typical” particular cases:

A modeling \mathbf{A} with universe A satisfies the *Finitary Mass Transport Principle* if, for every $\phi, \psi \in \text{FO}_1(\lambda)$ and every integers a, b such that

$$\begin{cases} \phi \vdash (\exists^{\geq a} y) (x_1 \sim y) \wedge \psi(y) \\ \psi \vdash (\exists^{\leq b} y) (x_1 \sim y) \wedge \phi(y) \end{cases}$$

it holds

$$a \langle \phi, \mathbf{A} \rangle \leq b \langle \psi, \mathbf{A} \rangle.$$

It is clear that every finite structure satisfies the Finitary Mass Transport Principle, hence every modeling FO-limit of finite structures satisfies the Finitary Mass Transport Principle, too.

The following stronger version of this principle, which is also satisfied by every finite structure, does not automatically hold in the limit.

A modeling \mathbf{A} with universe A satisfies the *Strong Finitary Mass Transport Principle* if, for every measurable subsets X, Y of A , and every integers a, b , the following property holds:

If every $x \in X$ has at least a neighbors in Y and every $y \in Y$ has at most b neighbors in X then $a \nu_{\mathbf{A}}(X) \leq b \nu_{\mathbf{A}}(Y)$.

The importance of residual and non-dispersive sequences appears again in the following result, which complements Theorem 14.

Theorem 19 ([32]). *Let \mathcal{C} be a hereditary class of structures.*

Assume that for every $\mathbf{A}_n \in \mathcal{C}$ and every $\rho_n \in A_n$ ($n \in \mathbb{N}$) the following properties hold:

- (1) *if $(\mathbf{A}_n)_{n \in \mathbb{N}}$ is $\text{FO}_1^{\text{local}}$ -convergent and residual, then it has a modeling $\text{FO}_1^{\text{local}}$ -limit;*
- (2) *if $(\mathbf{A}_n, \rho_n)_{n \in \mathbb{N}}$ is FO^{local} -convergent and ρ -non-dispersive then it has a modeling FO^{local} -limit.*

Then \mathcal{C} has modeling FO-limits.

Moreover, if in cases (1) and (2) the modeling limits satisfy the Strong Finitary Mass Transport Principle, then \mathcal{C} has modeling FO-limits that satisfy the Strong Finitary Mass Transport Principle.

6. CONSTRUCTIONS

In this section we add two useful extensions of some of the above results, by means of two basic operations. The first is the notion of (linear) convex combination of modelings. This may be used to give an inverse theorem for the comb structure Theorem 14. The second introduces the powerful tool of interpretation schemes, which gives the possibility to transport results from a class of λ -structures to a class of σ -structures. This also leads to a new characterization of nowhere dense classes.

6.1. Convex Combinations.

Definition 20 (Convex combination of Modelings). Let \mathbf{H}_i be λ -modelings for $i \in I \subseteq \mathbb{N}$ and let $(\alpha_i)_{i \in I}$ be positive real numbers such that $\sum_{i \in I} \alpha_i = 1$.

Let \mathbf{H} be the disjoint union of the \mathbf{H}_i , let $\Sigma_{\mathbf{H}} = \{\bigcup_i X_i : X_i \in \Sigma_{\mathbf{H}_i}\}$ and, for $X \in \Sigma_{\mathbf{H}}$, let $\nu_{\mathbf{H}}(X) = \sum_i \alpha_i \nu_{\mathbf{H}_i}(X \cap H_i)$.

Then \mathbf{H} is the *convex combination* of modelings \mathbf{H}_i with *weights* α_i and we denote it by $\coprod_{i \in I} (\mathbf{H}_i, \alpha_i)$.

The following converse of Theorem 14 is proved in [33]:

Theorem 21. *Assume J is a countable set, α_i ($i \in I$) are reals, and $(\mathbf{B}_{i,n})_{n \in \mathbb{N}}$ ($i \in I$) are sequences of λ -structures such that $\alpha_i = \lim_{n \rightarrow \infty} \frac{|\mathbf{B}_{i,n}|}{|\bigcup_{j \in I} \mathbf{B}_{j,n}|}$ ($\forall i \in I$), $\sum_{i \in I} \alpha_i = 1$, and for each $i \in I$, $(\mathbf{B}_{i,n})_{n \in \mathbb{N}}$ is FO^{local} -convergent. Then $\mathbf{A}_n = \bigcup_{i \in I} \mathbf{B}_{i,n}$ is FO^{local} -convergent. Also, if there exist λ -modelings \mathbf{L}_i ($i \in I$) such that for each $i \in I$, $\mathbf{B}_{i,n} \xrightarrow{\text{FO}^{\text{local}}} \mathbf{L}_i$, then $\mathbf{A}_n \xrightarrow{\text{FO}^{\text{local}}} \coprod_{i \in I} (\mathbf{L}_i, \alpha_i)$.*

Moreover, if the family $\{(\mathbf{B}_{i,n})_{n \in \mathbb{N}} : i \in I\}$ is uniformly elementarily-convergent, then $(\mathbf{A}_n)_{n \in \mathbb{N}}$ is FO-convergent. Also, if there exist λ -modelings \mathbf{L}_i ($i \in I$) such that for each $i \in I$ it holds $\mathbf{B}_{i,n} \xrightarrow{\text{FO}} \mathbf{L}_i$ (for $i > 0$ and $i = 0$ if $\alpha_0 > 0$) and $\mathbf{B}_{0,n} \xrightarrow{\text{FO}_0} \mathbf{L}_0$ (if $\alpha_0 = 0$) then $\mathbf{A}_n \xrightarrow{\text{FO}} \coprod_{i \in I} (\mathbf{L}_i, \alpha_i)$.

6.2. Interpretation Schemes. Interpretation Schemes (introduced in this setting in [33]) generalize to other logics than first-order logic.

Definition 22. Let \mathcal{L} be a logic (here either first-order logic or $\mathcal{L}_{\omega_1\omega}$). For $p \in \mathbb{N}$ and a signature λ , $\mathcal{L}_p(\lambda)$ denotes the set of the formulas in the language of λ in logic \mathcal{L} , with free variables in $\{x_1, \dots, x_p\}$.

Let κ, λ be signatures, where λ has q relational symbols R_1, \dots, R_q with respective arities r_1, \dots, r_q .

An \mathcal{L} -interpretation scheme \mathbf{l} of λ -structures in κ -structures is defined by an integer k — the *exponent* of the \mathcal{L} -interpretation scheme — a formula $E \in \mathcal{L}_{2k}(\kappa)$, a formula $\theta_0 \in \mathcal{L}_k(\kappa)$, and a formula $\theta_i \in \text{FO}_{r_i k}(\kappa)$ for each symbol $R_i \in \lambda$, such that:

- the formula E defines an equivalence relation of k -tuples;
- each formula θ_i is compatible with E , in the sense that for every $0 \leq i \leq q$ it holds

$$\bigwedge_{1 \leq j \leq r_i} E(\mathbf{x}_j, \mathbf{y}_j) \vdash \theta_i(\mathbf{x}_1, \dots, \mathbf{x}_{r_i}) \leftrightarrow \theta_i(\mathbf{y}_1, \dots, \mathbf{y}_{r_i}),$$

where $r_0 = 1$, boldface \mathbf{x}_j and \mathbf{y}_j represent k -tuples of free variables, and where $\theta_i(\mathbf{x}_1, \dots, \mathbf{x}_{r_i})$ stands for $\theta_i(x_{1,1}, \dots, x_{1,k}, \dots, x_{r_i,1}, \dots, x_{r_i,k})$.

For a κ -structure \mathbf{A} , we denote by $\mathbf{l}(\mathbf{A})$ the λ -structure \mathbf{B} defined as follows:

- the domain B of \mathbf{B} is the subset of the E -equivalence classes $[\mathbf{x}] \subseteq A^k$ of the tuples $\mathbf{x} = (x_1, \dots, x_k)$ such that $\mathbf{A} \models \theta_0(\mathbf{x})$;
- for each $1 \leq i \leq q$ and every $\mathbf{v}_1, \dots, \mathbf{v}_{r_i} \in A^{kr_i}$ such that $\mathbf{A} \models \theta_0(\mathbf{v}_j)$ (for every $1 \leq j \leq r_i$) it holds

$$\mathbf{B} \models R_i([\mathbf{v}_1], \dots, [\mathbf{v}_{r_i}]) \iff \mathbf{A} \models \theta_i(\mathbf{v}_1, \dots, \mathbf{v}_{r_i}).$$

From the standard properties of model theoretical interpretations (see, for instance [22] p. 180), we state the following: if \mathbf{l} is an \mathcal{L} -interpretation of λ -structures in κ -structures, then there exists a mapping $\tilde{\mathbf{l}} : \mathcal{L}(\lambda) \rightarrow \mathcal{L}(\kappa)$ (defined by means of the formulas $E, \theta_0, \dots, \theta_q$ above) such that for every $\phi \in \mathcal{L}_p(\lambda)$, and every κ -structure \mathbf{A} , the following property holds (while letting $\mathbf{B} = \mathbf{l}(\mathbf{A})$ and identifying elements of B with the corresponding equivalence classes of A^k):

For every $[\mathbf{v}_1], \dots, [\mathbf{v}_p] \in B^p$ (where $\mathbf{v}_i = (v_{i,1}, \dots, v_{i,k}) \in A^k$) it holds

$$\mathbf{B} \models \phi([\mathbf{v}_1], \dots, [\mathbf{v}_p]) \iff \mathbf{A} \models \tilde{\mathbf{l}}(\phi)(\mathbf{v}_1, \dots, \mathbf{v}_p).$$

It directly follows from the existence of the mapping $\tilde{\mathbf{l}}$ that

- an FO-interpretation scheme \mathbf{l} of λ -structures in κ -structures defines a continuous mapping from $S(\mathcal{B}(\text{FO}(\kappa)))$ to $S(\mathcal{B}(\text{FO}(\lambda)))$;
- an $\mathcal{L}_{\omega_1\omega}$ -interpretation scheme \mathbf{l} of λ -structures in κ -structures defines a measurable mapping from $S(\mathcal{B}(\text{FO}(\kappa)))$ to $S(\mathcal{B}(\text{FO}(\lambda)))$.

Definition 23. Let κ, λ be signatures. A *basic \mathcal{L} -interpretation scheme* \mathbf{l} of λ -structures in κ -structures with *exponent* k is defined by a formula $\theta_i \in \mathcal{L}_{kr_i}(\kappa)$ for each symbol $R_i \in \lambda$ with arity r_i .

For a κ -structure \mathbf{A} , we denote by $\mathbf{l}(\mathbf{A})$ the structure with domain A^k such that, for every $R_i \in \lambda$ with arity r_i and every $\mathbf{v}_1, \dots, \mathbf{v}_{r_i} \in A^k$ it holds

$$\mathbf{l}(\mathbf{A}) \models R_i(\mathbf{v}_1, \dots, \mathbf{v}_{r_i}) \iff \mathbf{A} \models \theta_i(\mathbf{v}_1, \dots, \mathbf{v}_{r_i}).$$

It is immediate that every basic \mathcal{L} -interpretation scheme \mathbf{l} defines a mapping $\tilde{\mathbf{l}} : \mathcal{L}(\lambda) \rightarrow \mathcal{L}(\kappa)$ such that for every κ -structure \mathbf{A} , every $\phi \in \mathcal{L}_p(\lambda)$, and every $\mathbf{v}_1, \dots, \mathbf{v}_p \in A^k$ it holds

$$\mathbf{l}(\mathbf{A}) \models \phi(\mathbf{v}_1, \dots, \mathbf{v}_p) \iff \mathbf{A} \models \tilde{\mathbf{l}}(\phi)(\mathbf{v}_1, \dots, \mathbf{v}_p)$$

We deduce the following general properties:

Proposition 5 ([33]). *Let \mathbf{l} be an FO-interpretation scheme of λ -structures in κ -structures.*

Then, if a sequence $(\mathbf{A}_n)_{n \in \mathbb{N}}$ of finite κ -structures is FO-convergent then the sequence $(\mathbf{l}(\mathbf{A}_n))_{n \in \mathbb{N}}$ of (finite) λ -structures is FO-convergent.

Proposition 6 ([32]). *Let \mathbf{l} be a basic $\mathcal{L}_{\omega_1\omega}$ -interpretation scheme of λ -structures in κ -structures.*

For every κ -modeling \mathbf{A} , $\mathbf{l}(\mathbf{A})$ is a λ -modeling such that for every $\phi \in \mathcal{L}_p(\lambda)$ it holds

$$\langle \phi, \mathbf{l}(\mathbf{A}) \rangle = \langle \tilde{\mathbf{l}}(\phi), \mathbf{A} \rangle.$$

Interpretations are a powerful tool. For a recent application to graph polynomials, see [17].

7. OPEN PROBLEMS

Our main conjecture here is Conjecture 1, which asserts that every monotone nowhere dense class of graphs has modeling limits.

There is a modest progress toward a resolution of this conjecture:

- any monotone class of graphs with bounded degree has modeling limit [30],
- the class of (colored rooted) forests has modeling limits [32].

As an extension of the results in [32], every FO-convergent sequence $(G_n)_{n \in \mathbb{N}}$ with $\text{girth}(G_n) \rightarrow \infty$ has a modeling limit. From this follows the first positive result concerning nowhere dense classes which do not have bounded average degree.

For a graph G , we denote by $b_1(G)$ the *cyclomatic number* of G (i.e. $b_1(G) = \|G\| - |G| + 1$), and by $\text{Diam}(G)$ its diameter.

Proposition 7. *Let $f : \mathbb{N} \rightarrow \mathbb{N}$ be a mapping and let \mathcal{C}_f be the monotone nowhere dense class of all (colored) graphs G such that for every graph $G \in \mathcal{C}$ it holds*

$$b_1(G) \leq f(\text{Diam}(G)).$$

Then the class \mathcal{C}_f has modeling limits.

Note that such classes include in particular all classes with bounded degree, and more generally classes of graphs in which the maximum degree is bounded by a function of the girth. Proposition 7 is currently our most general result toward Conjecture 1.

Problem 4. We consider 2-colored linear order, that is transitive tournaments with vertices colored black or white.

Do every FO-convergent sequence of 2-colored linear order admit a modeling FO-limit?

Note that if the answer to Problem 4 is yes, then it follows (by suitable interpretation) that every class of graphs with bounded pathwidth has modeling FO-limits. Similarly, if the above Example 16 extends to FO-convergence in the sense that every FO-convergent sequence of tree semilattices with k -colored domain has a modeling FO-limit then it would imply that every class of graphs with bounded treewidth has modeling FO-limits.

Quantifier-free fragment corresponds to the best known (and thoroughly investigated) type of convergence, which is left-convergence in the case of graphs and hypergraphs. However, our knowledge of QF-convergence is far from being complete. First, it is not completely clear what form the limit object should have in

case of a relational structures with several non-unary relations. More important, the extension of the theory (notions of (hyper)graphons, cut metric, regularity theorem) in the simple case of a signature reduced to a single binary function is completely open. In general this problem might be too difficult to be directly addressed. It appears that a natural restriction consists in limiting the power of the binary function:

Problem 5. Let σ be a signature reduced to a single binary function $*$ (denoted as an operation), and let T_0 be a finite set of universal formulas (in the first-order language of σ). Assume that there exists a function f such that for every p , every model \mathbf{A} of T_0 , and every $v_1, \dots, v_p \in A$, the σ -structure $\mathbf{A}\langle v_1, \dots, v_p \rangle$ generated by v_1, \dots, v_p — that is the substructure of \mathbf{A} whose domain is the smallest subset of A which contains v_1, \dots, v_p and is closed under the operation $*$ — has a domain of cardinality at most $f(p)$.

Does there exist a graphon-like limit object for QF-convergent sequences of finite models of T_0 ? Does there exist a regularity theorem for finite models of T_0 ?

Note that in some simple cases, we can answer positively. For instance in the case of tree-semilattices [9], which corresponds to associative, commutative, idempotent operation $*$ with the additional property

$$\forall x \forall y \forall z \quad [(x * z = x) \wedge (y * z = y)] \rightarrow [(x * y = x) \vee (x * y = y)].$$

However, this case is too simple in the sense that limit objects are Borel tree-semilattices, which means that tree-semilattices form a “random-free” class.

A possible candidate for a tractable non random-free example is the class of general semilattices, defined by general associative, commutative, idempotent operations $*$.

We can also consider signatures σ only containing unary functional symbols. In the case of a signature σ reduced to a single unary function, it appears that not only a limits of QF-convergent sequences of finite structures can be represented by a Borel σ -structure, but the class of these σ -structures actually admit modeling FO-limits, as an extension of the results in [32]. Note that for a signature σ consisting of at least two unary functional symbols, there is no hope to have modeling FO-limits (as such a class is sufficiently general to allow in it an interpretation of the class of all directed graphs)

REFERENCES

- [1] H. Adler and I. Adler, *Interpreting nowhere dense graph classes as a classical notion of model theory*, European J. Combin. **36** (2014), 322–330.
- [2] D. Aldous, *Exchangeability and related topics*, École d’été de probabilités de Saint-Flour, XIII — 1983, Lecture Notes in Mathematics, vol. 1117, Springer, 1985, pp. 1–198.
- [3] D. Aldous and R. Lyons, *Processes on unimodular random networks*, Electronic Journal of Probability **12** (2007), 1454–1508.
- [4] A. Atserias, A. Dawar, and M. Grohe, *Preservation under extensions on well-behaved finite structures*, 32nd International Colloquium on Automata, Languages and Programming (ICALP) (Springer-Verlag, ed.), Lecture Notes in Computer Science, vol. 3580, 2005, pp. 1437–1449.
- [5] I. Benjamini and O. Schramm, *Recurrence of distributional limits of finite planar graphs*, Electron. J. Probab. **6** (2001), no. 23, 13pp.

- [6] C. Borgs, J.T. Chayes, L. Lovász, V.T. Sós, B. Szegedy, and K. Vesztegombi, *Graph limits and parameter testing*, Proc. 38th Annual ACM Symp. Principles of Dist. Comp., 2005, pp. 51–59.
- [7] O. Bratteli, *Inductive limits of finite dimensional C^* -algebras*, Trans. AMS **171** (1972), 195–234.
- [8] ———, *Structure spaces of approximately finite-dimensional C^* -algebras*, Journal of Functional Analysis **16** (1974), no. 2, 192 – 204.
- [9] P. Charbit, L. Hosseini, and P. Ossona de Mendez, *Random free limits of structures and the example of tree-semilattices.*, in preparation, 2014.
- [10] J. Dixon, L. Pyber, A. Seress, and A. Shalev, *Residual properties of free groups and probabilistic methods*, J. Reine Angew. Math. **556** (2003), 159–172.
- [11] G.A. Elliott, *On the classification of inductive limits of sequences of semisimple finite-dimensional algebras*, Journal of Algebra **38** (1976), no. 1, 29 – 44.
- [12] K. Engesser, D.M. Gabbay, and D. Lehmann (eds.), *Handbook of quantum logic and quantum structures: Quantum logic*, Els, 2009.
- [13] D.B.A. Epstein, *Almost all subgroups of Lie group are free*, J. Algebra **19** (1971), 261–262.
- [14] H. Gaifman, *On local and non-local properties*, Proceedings of the Herbrand Symposium, Logic Colloquium '81, 1982.
- [15] A. Gamburd, S. Hoory, M. Shahshahani, A. Shalev, and B. Virág, *On the girth of random Cayley graphs*, Random Struct. Alg. **35** (2009), no. 1, 100–117.
- [16] I. Gelfand and M.A. Naimark, *On the imbedding of normed rings into the ring of operators in Hilbert space*, Rec. Math. [Mat. Sbornik] N.S. **12(54)** (1943), no. 2, 197–217.
- [17] A.J. Goodall, J. Nešetřil, and P. Ossona de Mendez, *Strongly polynomial sequences as interpretations*, arXiv:1405.2449 [math.CO], May 2014.
- [18] M. Grohe, S. Kreutzer, and S. Siebertz, *Deciding first-order properties of nowhere dense graphs*, arXiv:1311.3899 [cs.LO], November 2013.
- [19] D. Hoover, *Relations on probability spaces and arrays of random variables*, Tech. report, Institute for Advanced Study, Princeton, NJ, 1979.
- [20] A.S. Kechris, S. Solecki, and S. Todorcevic, *Borel chromatic numbers*, Advances in Mathematics **141** (1999), 1–44.
- [21] T. Kroupa, *Every state on semisimple MV-algebra is integral*, Fuzzy Sets and Systems **157** (2006), no. 20, 2771–2782.
- [22] D. Lascar, *La théorie des modèles en peu de maux*, Cassini, 2009.
- [23] L. Lovász, *Large networks and graph limits*, Colloquium Publications, vol. 60, American Mathematical Society, 2012.
- [24] L. Lovász and B. Szegedy, *Limits of dense graph sequences*, J. Combin. Theory Ser. B **96** (2006), 933–957.
- [25] ———, *Regularity partitions and the topology of graphons*, An irregular mind (Szemerédi is 70) (I. Bárány and J. Solymosi, eds.), Bolyai Society Mathematical Studies, vol. 21, Springer, 2010, pp. 415–446.
- [26] R.A. Minlos, *Encyclopedia of mathematics*, ch. Statistical ensemble, 2002.
- [27] J. Nešetřil and P. Ossona de Mendez, *First order properties on nowhere dense structures*, The Journal of Symbolic Logic **75** (2010), no. 3, 868–887.
- [28] ———, *Sparse combinatorial structures: Classification and applications*, Proceedings of the International Congress of Mathematicians 2010 (ICM 2010) (Hyderabad, India) (R. Bhatia and A. Pal, eds.), vol. IV, World Scientific,

- 2010, pp. 2502–2529.
- [29] ———, *On nowhere dense graphs*, European Journal of Combinatorics **32** (2011), no. 4, 600–617.
 - [30] ———, *A model theory approach to structural limits*, Commentationes Mathematicae Universitatis Carolinae **53** (2012), no. 4, 581–603.
 - [31] ———, *Sparsity (graphs, structures, and algorithms)*, Algorithms and Combinatorics, vol. 28, Springer, 2012, 465 pages.
 - [32] ———, *Modeling limits in hereditary classes: Reduction and application to trees*, arXiv:1312.0441 [math.CO], December 2013.
 - [33] ———, *A unified approach to structural limits (with application to the study of limits of graphs with bounded tree-depth)*, arXiv:1303.6471v2 [math.CO], October 2013.
 - [34] G. Panti, *Invariant measures in free MV-algebras*, Communications on Algebra **36** (2008), no. 8, 2849–2861.
 - [35] F.P. Ramsey, *On a problem of formal logic*, Proc. London Math. Soc. series 2, 264–286, vol. 30, 1930.
 - [36] M.H. Stone, *The theory of representations of Boolean algebras*, Transactions of the American Mathematical Society **40** (1936), 37–111.

JAROSLAV NEŠETŘIL, COMPUTER SCIENCE INSTITUTE OF CHARLES UNIVERSITY (IUUK AND ITI), MALOSTRANSKÉ NÁM.25, 11800 PRAHA 1, CZECH REPUBLIC
E-mail address: `nesetril@iuuk.mff.cuni.cz`

PATRICE OSSONA DE MENDEZ, CENTRE D'ANALYSE ET DE MATHÉMATIQUES SOCIALES (CNRS, UMR 8557), 190-198 AVENUE DE FRANCE, 75013 PARIS, FRANCE — AND — COMPUTER SCIENCE INSTITUTE OF CHARLES UNIVERSITY (IUUK), MALOSTRANSKÉ NÁM.25, 11800 PRAHA 1, CZECH REPUBLIC
E-mail address: `pom@ehess.fr`