Combinatorics, Probability and Computing (2016) 25, 213–221.
doi:10.1017/S0963548315000048

First-Order Convergence and Roots†

DEMETRES CHRISTOFIDES[1] and DANIEL KRÁL'[2§]

[1]School of Sciences, UCLan Cyprus, 12–14 University Avenue, 7080 Pyla, Cyprus
(e-mail: dchristofides@uclan.ac.uk)
[2]Mathematics Institute, DIMAP and Department of Computer Science, University of Warwick, Coventry CV4 7AL, UK
(e-mail: d.kral@warwick.ac.uk)

Received 14 March 2014; revised 27 October 2014; first published online 24 February 2015

Nešetřil and Ossona de Mendez introduced the notion of first-order convergence, which unifies the notions of convergence for sparse and dense graphs. They asked whether, if $(G_i)_{i\in\mathbb{N}}$ is a sequence of graphs with M being their first-order limit and v is a vertex of M, then there exists a sequence $(v_i)_{i\in\mathbb{N}}$ of vertices such that the graphs G_i rooted at v_i converge to M rooted at v. We show that this holds for almost all vertices v of M, and we give an example showing that the statement need not hold for all vertices.

2010 *Mathematics subject classification*: Primary 03C98
Secondary 05C63, 28E15

1. Introduction

The theory of limits for combinatorial objects continues to attract ever more attention, and its applications continue to grow in various areas such as extremal combinatorics, computer science and many others. The best-understood case is that of dense graph convergence, which originated in the series of papers by Borgs, Chayes, Lovász, Sós, Szegedy and Vesztergombi [3–5, 12, 13]. This development is also reflected in a recent monograph by Lovász [11]. Another line of research concentrates on the convergence of sparse graphs (such as those with bounded maximum degree), known as Benjamini–Schramm convergence [1,2,8,9]. Nešetřil and Ossona de Mendez [14,15] proposed a notion of first-order convergence to unify the two notions for the dense and sparse settings.

† This work was done during a visit to the Institut Mittag-Leffler (Djursholm, Sweden).
§ The work leading to this creation has received funding from the European Research Council under the European Union's Seventh Framework Programme (FP7/2007-2013)/ERC grant agreement no. 259385.

First-order convergence is a notion of convergence for all relational structures. For simplicity, we limit our exposition to graphs and rooted graphs only, but all our arguments extend naturally to the general setting. If ψ is a first-order formula with k free variables and G is a finite graph, then the *Stone pairing* $\langle \psi, G \rangle$ is the probability that a uniformly chosen k-tuple of vertices of G satisfies ψ. A sequence $(G_n)_{n \in \mathbb{N}}$ of graphs is called *first-order convergent* if the limit $\lim_{n\to\infty} \langle \psi, G_n \rangle$ exists for every first-order formula ψ. A *modelling* M is a (finite or infinite) graph whose vertex set is equipped with a probability measure such that the set of all k-tuples of vertices of M satisfying a formula ψ is measurable in the product measure for every first-order formula ψ with k free variables. In analogy to the graph case, the *Stone pairing* $\langle \psi, M \rangle$ is the probability that a randomly chosen k-tuple of vertices satisfies ψ. If a finite graph is viewed as a modelling with a uniform discrete probability measure on its vertex set, then the stone pairings for the graph and the modelling obtained in this way coincide.

A modelling M is a limit of a first-order convergent sequence $(G_n)_{n \in \mathbb{N}}$ if

$$\lim_{n\to\infty} \langle \psi, G_n \rangle = \langle \psi, M \rangle$$

for every first-order formula ψ. It is not true that every first-order convergent sequence of graphs has a limit modelling [15], but it can be shown, for example, that first-order convergent sequences of trees do [10, 16].

Nešetřil and Ossona de Mendez [15, Problem 1] posed the following problem, which we formulate here for graphs alone.

Problem 1.1. *Let M be a modelling, i.e., a limit of a first-order convergent sequence $(G_n)_{n \in \mathbb{N}}$, and let v be a vertex of M. Does there exist a sequence $(v_n)_{n \in \mathbb{N}}$ of vertices of the graphs $(G_n)_{n \in \mathbb{N}}$ such that the modelling M rooted at v is a limit of the sequence $(G'_n)_{n \in \mathbb{N}}$, where G'_n is obtained from G_n by rooting at v_n?*

We prove that the statement from Problem 1.1 is true for almost every vertex v of M.

Theorem 1.2. *Let M be a modelling, i.e., a limit of a first-order convergent sequence $(G_n)_{n \in \mathbb{N}}$. With probability one, if M' is a modelling obtained from M by rooting at a random vertex v of M, then there exists a sequence $(v_n)_{n \in \mathbb{N}}$ of vertices of $(G_n)_{n \in \mathbb{N}}$ such that M' is a limit of the sequence $(G'_n)_{n \in \mathbb{N}}$, where G'_n is obtained from G_n by rooting at v_n.*

Theorem 1.2 follows from a more general Theorem 3.4, which we prove in Section 3. In Section 4 we present an example that the statement of Theorem 1.2 cannot be strengthened to all vertices v of M, that is, the answer to Problem 1.1 is negative. This also answers a more general problem [15, Problem 2] in the negative.

2. Notation

We assume that the reader is familiar with standard graph theory and logic terminology, as can be found in [6, 7], for example. We briefly review less standard terminology and notation only. Throughout the paper we write $[k]$ for the set of positive integers $\{1, 2, \ldots, k\}$.

There is a close connection between the first-order logic and the so-called Ehrenfeucht–Fraïssé games. The *p-round Ehrenfeucht–Fraïssé game* is played by two players, the spoiler and the duplicator, on two relational structures. We explain the game when played on two graphs G and H. At the beginning of each round, the spoiler chooses a vertex in any one of the two graphs and the duplicator responds by choosing a vertex in the other. One vertex can be chosen several times in different rounds of the game. Let v_i be the vertex chosen in the ith round in G and w_i the vertex chosen in the ith round in H. The duplicator wins the game if the subgraph of G induced by $v_1,\dots,v_p$ and the subgraph of H induced by $w_1,\dots,w_p$ are isomorphic through the isomorphism mapping v_i to w_i.

It can be shown that the duplicator has a winning strategy for the p-round Ehrenfeucht–Fraïssé game played on G and H if and only if G and H satisfy the same first-order sentences with quantifier depth at most p. More generally, suppose that $\psi(x_1,\dots,x_k)$ is a first-order formula with k free variables and with quantifier depth d, G and H are two graphs, and $v_1,\dots,v_k$ and $w_1,\dots,w_k$ are (not necessarily distinct) vertices of G and H, respectively. If the duplicator has a winning strategy for the $(k+d)$-round Ehrenfeucht–Fraïssé game when played on G and H with the vertices $v_1,\dots,v_k$ and $w_1,\dots,w_k$ played in the first k rounds (so it remains to play d rounds of the game), then G satisfies $\psi(v_1,\dots,v_k)$ if and only if H satisfies $\psi(w_1,\dots,w_k)$. This correspondence can be used to show [7] that the set $F^m_{p,q}$ of all non-equivalent first-order formulas with p free variables and quantifier depth at most q for m-rooted graphs is finite for all positive integers m, p and q (the language for m-rooted graphs consists of a single binary relation representing the adjacency and m constants representing the roots).

3. Almost every rooting is good

In this section we prove our main result, which provides a positive answer to Problem 1.1 in the 'almost every' sense. As preparation for the proof, we need to establish several technical lemmas.

Lemma 3.1. *Let ψ be a first-order formula for m-rooted graphs and let $[a,b] \subseteq [0,1]$ be a non-empty interval. For every $\varepsilon > 0$, there exists a first-order formula ψ' such that the following holds for every m-rooted modelling M:*

- *if $\langle \psi, M\rangle \in [a,b]$, then $\langle \psi', M\rangle > 1-\varepsilon$, and*
- *if $\langle \psi, M\rangle \notin (a-\varepsilon, b+\varepsilon)$, then $\langle \psi', M\rangle < \varepsilon$.*

Proof. If ψ is a sentence, that is, it has no free variables, then the statement is trivial. Suppose that ψ has k free variables. Let ψ_n be the first-order formula with nk free variables grouped in n k-tuples, such that ψ_n is true if and only if at least $an - n^{2/3}$ and at most $bn + n^{2/3}$ of these k-tuples satisfy ψ. Formally,

$$\psi_n(x^1_1,\dots,x^1_k,\dots,x^n_1,\dots,x^n_k) = \bigvee_{i=\lceil an-n^{2/3}\rceil}^{\lfloor bn+n^{2/3}\rfloor} \bigvee_{A\in\binom{[n]}{i}} \left(\bigwedge_{j\in A} \psi(x^j_1,\dots,x^j_k) \wedge \bigwedge_{j\notin A} \neg\psi(x^j_1,\dots,x^j_k)\right).$$

https://doi.org/10.1017/S0963548315000048 Published online by Cambridge University Press

The Chernoff bound implies that the formula ψ' can be chosen to be the formula ψ_n for sufficiently large n. □

An interval is a *dyadic interval of order* $k \in \mathbb{N}$ if it is of the form $[a2^{-k}, (a+1)2^{-k}]$ for some integer a. A point x is ε*-far* from an interval J if $|x - y| \geqslant \varepsilon$ for every $y \in J$. Otherwise, we say that x is ε*-close* to J. A *multidimensional interval* is a subset of $[0,1]^d$ that is the product of d intervals; if J is a multidimensional interval, then J_i denotes the ith term in the product. A multidimensional interval J is *dyadic of order* $k \in \mathbb{N}$ if every J_i is dyadic of order k.

The next lemma is a direct consequence of Lemma 3.1. Recall that $F^m_{p,q}$ is the set of all non-equivalent first-order formulas with p free variables and quantifier depth at most q, and the set $F^m_{p,q}$ is finite for all m, p and q.

Lemma 3.2. *Let m, p and q be integers and let $J \subseteq [0,1]^{F^m_{p,q}}$ be a multidimensional interval. For every $\varepsilon > 0$ there exists a first-order formula $\psi^{m,J,\varepsilon}_{p,q}$ such that the following holds for every m-rooted modelling M:*

- *if $\langle \psi, M \rangle \in J_\psi$ for every $\psi \in F^m_{p,q}$, then $\langle \psi^{m,J,\varepsilon}_{p,q}, M \rangle > 1 - \varepsilon$, and*
- *if $\langle \psi, M \rangle$ is ε-far from J_ψ for at least one $\psi \in F^m_{p,q}$, then $\langle \psi^{m,J,\varepsilon}_{p,q}, M \rangle < \varepsilon$.*

If $\psi^{m,J,\varepsilon}_{p,q}$ is the formula from Lemma 3.2, then $\widehat{\psi}^{m,J,\varepsilon}_{p,q}$ is the formula obtained from $\psi^{m,J,\varepsilon}_{p,q}$ by adding m new free variables, such that $\widehat{\psi}^{m,J,\varepsilon}_{p,q}$ is satisfied if and only if $\psi^{m,J,\varepsilon}_{p,q}$ is satisfied for the modelling obtained from M by rooting at the m-tuple specified by the new free variables, that is, the m constants in $\psi^{m,J,\varepsilon}_{p,q}$ are replaced with the m new free variables of $\psi^{m,J,\varepsilon}_{p,q}$. An m-tuple of vertices $v_1, \dots, v_m$ of a modelling M is *negligible* if there exist integers p and q and a dyadic multidimensional interval $J \subseteq [0,1]^{F^m_{p,q}}$ such that

- $\langle \psi, M' \rangle_{\psi \in F^m_{p,q}} \in J$ where M' is the m-rooted modelling obtained from M by rooting at $v_1, \dots, v_m$, and
- there exists $\varepsilon_0 > 0$ such that $\langle \widehat{\psi}^{m,J,\varepsilon}_{p,q}, M \rangle \leqslant \varepsilon$ for every $0 < \varepsilon < \varepsilon_0$.

The next lemma asserts that very few m-tuples can be negligible.

Lemma 3.3. *If M is a modelling and m is an integer, then the set of negligible m-tuples of M is a subset of a set of measure zero.*

Proof. Note that there are countably many triples p, q and $J \subseteq [0,1]^{F^m_{p,q}}$, where J is dyadic. Therefore, it is sufficient to show for every p, q and J that if there exists $\varepsilon_0 > 0$ such that $\langle \widehat{\psi}^{m,J,\varepsilon}_{p,q}, M \rangle < \varepsilon$ for every $0 < \varepsilon < \varepsilon_0$, then there exists a set of measure zero containing all m-tuples $v_1, \dots, v_m$ such that $\langle \psi, M' \rangle_{\psi \in F^m_{p,q}} \in J$, where M' is obtained from M by rooting at $v_1, \dots, v_m$. Fix p, q and J for the rest of the proof. Let X be the set of all such m-tuples, and let k_0 be an integer such that $2^{-k_0} < \varepsilon_0$.

Let $F_k(v_1, \dots, v_m)$ for $k \in \mathbb{N}$ be the function from M^m to $[0,1]$ defined to be $\langle \psi^{m,J,2^{-k}}_{p,q}, M' \rangle$, where M' is the modelling obtained from M by rooting at $v_1, \dots, v_m$. Since the set of tuples satisfying $\widehat{\psi}^{m,J,2^{-k}}_{p,q}$ is measurable, the function F_k is measurable in the corresponding

https://doi.org/10.1017/S0963548315000048 Published online by Cambridge University Press

product space. Moreover, we have

$$\int F_k(v_1,\dots,v_m)\,\mathrm{d}v_1\cdots v_m = \langle \widehat{\psi}^{m,J,2^{-k}}_{p,q}, M\rangle < 2^{-k}$$

for every $k \geqslant k_0$. Observe that Lemma 3.2 yields that

$$X \subseteq \bigcap_{k=k_0}^{\infty} F_k^{-1}([1-2^{-k},1]).$$

Since the function F_k takes values between 0 and 1 (inclusive), the measure of $F_k^{-1}([1-2^{-k},1])$ is less than $2^{-k}/(1-2^{-k})$. It follows that X is a subset of a set of measure zero. □

We are now ready to prove our main theorem.

Theorem 3.4. *Let M be a modelling, i.e., a limit of a first-order convergent sequence $(G_n)_{n\in\mathbb{N}}$, and let m be a positive integer. With probability one, if M' is a modelling obtained from M by rooting at a random m-tuple of vertices of M, then there exists a sequence $(v_{n,1},\dots,v_{n,m})_{n\in\mathbb{N}}$ of m-tuples such that the graphs $(G_n)_{n\in\mathbb{N}}$ rooted at these m-tuples first-order converge to M'.*

Proof. By Lemma 3.3, we can assume that the randomly chosen m-tuple $w_1,\dots,w_m$ of the vertices of M is not negligible. It is enough to show for every p, q and $\delta > 0$ that there exists n_0 such that every graph G_n, $n \geqslant n_0$, contains an m-tuple $v_{n,1},\dots,v_{n,m}$ of vertices such that the graph G'_n obtained from G_n by rooting at the vertices $v_{n,1},\dots,v_{n,m}$ satisfies

$$|\langle\psi, G'_n\rangle - \langle\psi, M'\rangle| \leqslant \delta \tag{3.1}$$

for every $\psi \in F^m_{p,q}$. Note that (3.1) implies that

$$|\langle\psi, G'_n\rangle - \langle\psi, M'\rangle| \leqslant \delta$$

for every $\psi \in F^m_{p',q'}$, for $p' \in [p]$ and $q' \in [q]$.

Fix the integers p and q and the real $\delta > 0$ for the rest of the proof. Choose an integer k such that $2^{-k} < \delta$ and a real $\varepsilon > 0$ such that $2^{-k}+\varepsilon < \delta$. Further, let $J \subseteq [0,1]^{F^m_{p,q}}$ be the dyadic multidimensional interval of order k containing the point $\langle\psi, M'\rangle_{\psi\in F^m_{p,q}}$. Since the m-tuple $v_{n,1},\dots,v_{n,m}$ is not negligible, there exists $\varepsilon' < \varepsilon$ such that

$$\langle \widehat{\psi}^{m,J,\varepsilon'}_{p,q}, M\rangle > \varepsilon'.$$

Since the sequence $(G_n)_{n\in\mathbb{N}}$ converges to M, there exists n_0 such that

$$\langle \widehat{\psi}^{m,J,\varepsilon'}_{p,q}, G_n\rangle > \varepsilon' \tag{3.2}$$

for every $n \geqslant n_0$. By the definition of the formula $\widehat{\psi}^{m,J,\varepsilon'}_{p,q}$, the inequality (3.2) implies that every graph G_n, $n \geqslant n_0$, contains an m-tuple $v_{n,1},\dots,v_{n,m}$ of vertices such that

$$\langle \psi^{m,J,\varepsilon'}_{p,q}, G'_n\rangle > \varepsilon', \tag{3.3}$$

where G'_n is obtained from G_n by rooting at $v_{n,1},\dots,v_{n,m}$. By Lemma 3.2, the Stone pairing $\langle\psi, G'_n\rangle$ is ε'-close to J_ψ for every $\psi \in F^m_{p,q}$. It follows that

$$|\langle\psi, G'_n\rangle - \langle\psi, M\rangle| < 2^{-k} + \varepsilon' < \delta$$

https://doi.org/10.1017/S0963548315000048 Published online by Cambridge University Press

for every $\psi \in F^m_{p,q}$. The proof of the theorem is now finished. □

4. Counterexample

We now show that the statement of Theorem 1.2 cannot be strengthened to all vertices. Before doing so, we need to introduce some additional notation.

If a (finite or infinite) graph G is bipartite, we write $G(A,B)$, where A and B are the two parts of G. The adjacency matrix M of G is the matrix with rows indexed by A and columns indexed by B such that M_{ab} is equal to 1 if the vertices $a \in A$ and $b \in B$ are adjacent, and equal to zero otherwise. If $G(A,B)$ is a bipartite graph and W is a subset of its vertices, then W_A is $A \cap W$ and W_B is $B \cap W$. The adjacency matrix of G restricted to W is the submatrix with rows and columns indexed by W_A and W_B, respectively. Suppose that W is a subset of vertices of $G(A,B)$, W' is a subset of vertices of $G'(A',B')$ and there is a one-to-one correspondence between the vertices of W and W'. When we say that the adjacency matrices of G and G' restricted to W and W' are the same, we mean that they are the same in the stronger sense that the rows/columns for the corresponding vertices are the same.

A bipartite graph $G(A,B)$ is ℓ-universal if every vector from $\{0,1\}^B$ appears at least ℓ times among the rows of the adjacency matrix of G. If W is a subset of vertices of $G(A,B)$, then the ℓ-shadow of W is the multiset S such that each of the vectors $u \in \{0,1\}^{W_A}$ is included in S exactly $\min\{k,\ell\}$ times, where k is the number of times u appears among the columns of the adjacency matrix of G restricted to $W_A \times (B \setminus W_B)$. If $W_A = \emptyset$, then the ℓ-shadow of W consists of $\min\{|B|,\ell\}$ null vectors (*i.e.*, vectors of dimension zero).

The following is the key lemma in our construction.

Lemma 4.1. *Let p and q be two non-negative integers. Let $G(A,B)$ and $G'(A',B')$ be two $(p+q)$-universal graphs, and let $w_1,\dots,w_q$ and $w'_1,\dots,w'_q$ be two sequences of the vertices of $G(A,B)$ and $G'(A',B')$, respectively. Let $W = \{w_1,\dots,w_q\}$ and $W' = \{w'_1,\dots,w'_q\}$. If the adjacency matrices of $G(A,B)$ and $G'(A',B')$ restricted to $W_A \times W_B$ and to $W'_{A'} \times W'_{B'}$, respectively, are the same (with the row/column corresponding to w_i being the same as that of w'_i), and the 2^p-shadows of W and W' are also the same, then the duplicator has a winning strategy for the $(p+q)$-round Ehrenfeucht–Fraïssé game, where the vertices chosen in the first q rounds are $w_1,\dots,w_q$ and $w'_1,\dots,w'_q$.*

Proof. We proceed by induction on p. If $p = 0$ then the graphs induced by the vertices of W and W' are isomorphic, since the adjacency matrices of G and G' restricted to W and W' are the same.

Suppose that $p > 0$. By symmetry, we can assume that the spoiler chooses a vertex of G in the next round. Let w_{q+1} be the chosen vertex. If $w_{q+1} = w_i$ for some $1 \leqslant i \leqslant q$, the duplicator responds with w'_i. So we can now assume that w_{q+1} is different from all the vertices $w_1,\dots,w_q$ and we distinguish two cases based on whether w_{q+1} belongs to A or B.

Let us start with the analysis of the case when $w_{q+1} \in A \setminus A_W$. Let x be the row of the adjacency matrix of G corresponding to w_{q+1}. We will construct a vector $x' \in \{0,1\}^{B'}$ which will determine the response of the duplicator.

Set $x'_{w'_i} = x_{w_i}$ for $w'_i \in W'_{B'}$. Fix a vector $u \in \{0,1\}^{W_A}$. Let $u_0, u_1 \in \{0,1\}^{W_A \cup \{w_{q+1}\}}$ be the two extensions of u, and let m_0 and m_1 be the multiplicities of u_0 and u_1, respectively, in the 2^{p-1}-shadow of $W \cup \{w_{q+1}\}$. Finally, let W_u be the set of the vertices v of $B' \setminus W'_{B'}$ such that the column of v restricted to $W'_{A'}$ is u. If $m_0 + m_1 < 2^p$, then the 2^p-shadow of W' contains the vector u exactly $m_0 + m_1$ times. Set x'_v to 0 for m_0 of the vertices $v \in W_u$ and to 1 for m_1 of such vertices. If $m_0 + m_1 \geqslant 2^p$, at least one of the numbers m_0 or m_1 is at least 2^{p-1}. If $m_0 \geqslant 2^{p-1}$, set x'_v to 1 for $\min\{m_1, 2^{p-1}\}$ of vertices $v \in W_u$ and to 0 for all other $v \in W_u$. If $m_0 < 2^{p-1}$ and $m_1 \geqslant 2^{p-1}$, set x'_v to 0 for m_0 of vertices $v \in W_u$ and to 1 for all other $v \in W_u$. Performing this for every vector $u \in \{0,1\}^{W_A}$, the entire vector $x \in \{0,1\}^{B'}$ is defined.

Since the graph G' is $(p+q)$-universal, there exists a vertex $w'_{q+1} \in A'$ different from the vertices $w'_1, \ldots, w'_q$ such that the row of w'_{q+1} in the adjacency matrix of G' is equal to x'. The duplicator responds with the vertex w'_{q+1}. Observe that the choice of x' implies that the adjacency matrices of G and G' restricted to $W \cup \{w_{q+1}\}$ and $W' \cup \{w'_{q+1}\}$, respectively, are the same and that the 2^{p-1}-shadows of $W \cup \{w_{q+1}\}$ and $W' \cup \{w'_{q+1}\}$ are also the same. The existence of the winning strategy for the duplicator now follows by induction.

It remains to consider the case that $w_{q+1} \in B \setminus B_W$. Let u be the column of w_{q+1} in the adjacency matrix of G restricted to W_A. Clearly, u is contained in the 2^p-shadow of W. Consequently, there is a vertex $w'_{q+1} \in B' \setminus B'_{W'}$ such that the column of w'_{q+1} in the adjacency matrix of G' restricted to $W'_{A'}$ is u. The duplicator responds with the vertex w'_{q+1}. The adjacency matrices of G and G' restricted to $W \cup \{w_{q+1}\}$ and $W' \cup \{w'_{q+1}\}$, respectively, are the same. The 2^{p-1}-shadow of $W \cup \{w_{q+1}\}$ in G is obtained from the 2^p-shadow of W by removing u from the shadow and restricting the multiplicity of each vector to be at most 2^{p-1}. Likewise, the 2^{p-1}-shadow of $W' \cup \{w'_{q+1}\}$ in G' is obtained from the 2^p-shadow of W' by removing u from the shadow and restricting the multiplicity of each vector to be at most 2^{p-1}. Note that if $A = A' = \emptyset$, each of the 2^{p-1}-shadows consists of 2^{p-1} null vectors. Since the 2^p-shadows of W and W' are the same, the 2^{p-1}-shadows of $W \cup \{w_{q+1}\}$ and $W' \cup \{w'_{q+1}\}$ are also the same. The existence of the winning strategy for the duplicator now follows by induction. □

Let $\boldsymbol{s} = (s_n)_{n\in\mathbb{N}}$ be a sequence of integers such that $s_n \geqslant 2$ for every $n \in \mathbb{N}$. For each $x \in [0,1]$, there exists a unique sequence $(x_n)_{n\in\mathbb{N}}$ of integers such that

$$x = \sum_{n=1}^{\infty} \frac{x_n}{\prod_{k=1}^{n} s_k},$$

$0 \leqslant x_n < s_n$ for every n, and there is no $n \in \mathbb{N}$ such that $x_n \neq s_n$ and $x_{n'} = s_{n'}$ for every $n' \geqslant n$. We define $M_{\boldsymbol{s}}$ to be the following modelling. The vertex set of $M_{\boldsymbol{s}}$ is the unit square $[0,1]^2$ with the uniform measure on its Borel subsets. Fix a sequence $(z_n)_{n\in\mathbb{N}}$ of distinct vertices, say $z = (2^{-n}, 0)$, and let $Z = \{z_n, n \in \mathbb{N}\}$. The modelling $M_{\boldsymbol{s}}(A,B)$ is the bipartite graph with $A = [0,1]^2 \setminus Z$ and $B = Z$ such that a vertex $(x,y) \in A = [0,1]^2 \setminus Z$ is adjacent to a vertex $z_n \in B = Z$ if and only if $x_n \neq 0$.

We next verify that every first-order definable subset of $M_{\boldsymbol{s}}^k$ is Borel. A subset X of $M_{\boldsymbol{s}}^\ell$ is *basic* if there exist $v_1, \ldots, v_p \in B$ (we allow $p = 0$), a matrix $M \in \{0,1\}^{\ell \times p}$, an integer q

and a multiset $S \subseteq \{0,1\}^p$ such that the set X is formed by all ℓ-tuples $w_1,\dots,w_\ell \in A$ for which the adjacency matrix restricted to $\{v_1,\dots,v_p,w_1,\dots,w_\ell\}$ is M and the 2^q-shadow of $\{v_1,\dots,v_p,w_1,\dots,w_\ell\}$ is S. In particular, if X is basic, then $X \subseteq A^\ell$.

For a non-negative integer ℓ, a finite subset $B' \subseteq B$, a matrix $M \in \{0,1\}^{\ell\times B'}$ and a subset $T \subseteq \{0,1\}^\ell$, let $X(\ell,B',M,T)$ be the set of ℓ-tuples $w_1,\dots,w_\ell \in A$ such that the the adjacency matrix of M_s restricted to $\{w_1,\dots,w_\ell\} \times B'$ is M and all the columns of the adjacency matrix not associated with vertices of B' belong to T after restricting to $w_1,\dots,w_\ell$. Observe that the set $X(\ell,B',M,T) \subseteq A^\ell$ is Borel for all ℓ, B', M and T. Since every basic set is a countable union of sets $X(\ell,B',M,T)$, every basic set is Borel.

Fix a first-order formula ψ with k free variables and quantifier depth d. By Lemma 4.1, the set of k-tuples of M_s^k satisfying ψ can be partitioned into countably many subsets such that each of them, after a suitable permutation of coordinates, is either a basic set or a product of a basic set and one or more single element subsets of B. Consequently, every first-order definable subset of M_s^k is Borel.

The next lemma directly follows from the definition of a modelling M_s.

Lemma 4.2. *Let $s = (s_n)_{n\in\mathbb{N}}$ be a sequence of integers such that $s_n \geqslant 2$ for every $n \in \mathbb{N}$. The modelling $M_s(A,B)$ is ℓ-universal for every $\ell \in \mathbb{N}$. For all integers p and ℓ, it holds with probability one that a random p-tuple of vertices of M_s contains p different vertices from A and the ℓ-shadow of the p-tuple is the multiset containing each vector $\{0,1\}^p$ with multiplicity ℓ.*

Observe that Lemmas 4.1 and 4.2 yield that $\langle\psi, M_s\rangle = \langle\psi, M_{s'}\rangle$ for every first-order formula ψ and any two sequences s and s'.

We now define the graph $H_n(A,B)$ to be the graph with $A = [2^n] \times [n]$ and $B = [n]$ such that $(a,a') \in A$ is adjacent to $b \in B$ if and only if the bth bit of a when written in binary is 1. We summarize the properties of the graphs H_n in the next lemma.

Lemma 4.3. *Let p and ℓ be two integers. The graph $H_n(A,B)$ is ℓ-universal if $n \geqslant \ell$, and the probability that a random p-tuple of vertices of H_n contains p different vertices from A and the ℓ-shadow of them is the multiset containing each vector $\{0,1\}^p$ with multiplicity ℓ, tends to one as n tends to infinity.*

The next theorem follows directly from Lemmas 4.1–4.3.

Theorem 4.4. *Let $s = (s_n)_{n\in\mathbb{N}}$ be a sequence of integers such that $s_n \geqslant 2$ for every $n \in \mathbb{N}$. For every first-order formula ψ,*

$$\lim_{n\to\infty} \langle\psi, H_n\rangle = \langle\psi, M_s\rangle.$$

In particular, the modelling M_s is a limit of $(H_n)_{n\in\mathbb{N}}$.

Let $\psi_0(x)$ be the first-order formula that is true if x is adjacent to the root. If $s_n = 3$ for every $n \in \mathbb{N}$, then the set of neighbours of every vertex of B in the modelling $M_s(A,B)$ has measure $2/3$. Hence, if M_s' is the modelling obtained from M_s by rooting at an arbitrary

https://doi.org/10.1017/S0963548315000048 Published online by Cambridge University Press

vertex of B, then $\langle \psi_0, M'_s \rangle = 2/3$. Since no vertex of H_n is adjacent to more than $2^{n-1}n$ vertices (the vertices of A are adjacent to at most n vertices each, and each vertex of B is adjacent to $2^{n-1}n$), it holds that

$$\limsup_{n\to\infty} \langle \psi_0, H'_n \rangle \leqslant \frac{1}{2}$$

for any sequence $(H'_n)_{n\in\mathbb{N}}$ of rooted graphs obtained from H_n. We conclude that the sequence $(H_n)_{n\in\mathbb{N}}$, the modelling $M_s(A,B)$ with $\mathbf{s} = (3)_{n\in\mathbb{N}}$ and rooting M_s at any vertex of B provide a counterexample to Problem 1.1.

References

[1] Aldous, D. and Lyons, R. (2007) Processes on unimodular random networks. *Electron. J. Probab.* **12** 1454–1508.

[2] Benjamini, I. and Schramm, O. (2001) Recurrence of distributional limits of finite planar graphs. *Electron. J. Probab.* **6** 1–13.

[3] Borgs, C., Chayes, J., Lovász, L., Sós, V. T., Szegedy, B. and Vesztergombi, K. (2006) Graph limits and parameter testing. In *Proc. 38rd Annual ACM Symposium on the Theory of Computing: STOC*, ACM, pp. 261–270.

[4] Borgs, C., Chayes, J. T., Lovász, L., Sós, V. T. and Vesztergombi, K. (2008) Convergent sequences of dense graphs I: Subgraph frequencies, metric properties and testing. *Adv. Math.* **219** 1801–1851.

[5] Borgs, C., Chayes, J. T., Lovász, L., Sós, V. T. and Vesztergombi, K. (2012) Convergent sequences of dense graphs II: Multiway cuts and statistical physics. *Ann. of Math.* **176** 151–219.

[6] Diestel, R. (2010) *Graph Theory*, Springer.

[7] Ebbinghaus, H.-D. and Flum, J. (2005) *Finite Model Theory*, Springer.

[8] Elek, G. (2007) Note on limits of finite graphs. *Combinatorica* **27** 503–507.

[9] Hatami, H., Lovász, L. and Szegedy, B. Limits of locally-globally convergent graph sequences. *Geom. Funct. Anal.*, **24** (2014), 269–296.

[10] Král', D., Kupec, M. and Tůma, V. Modelings of trees. In preparation.

[11] Lovász, L. (2012) *Large Networks and Graph Limits*, AMS.

[12] Lovász, L. and Szegedy, B. (2006) Limits of dense graph sequences. *J. Combin. Theory Ser. B* **96** 933–957.

[13] Lovász, L. and Szegedy, B. (2010) Testing properties of graphs and functions. *Israel J. Math.* **178** 113–156.

[14] Nešetřil, J. and Ossona de Mendez, P. A model theory approach to structural limits, Comment. *Math. Univ. Carolin.* **53** (2012), 581–603.

[15] Nešetřil, J. and Ossona de Mendez, P. A unified approach to structural limits, and limits of graphs with bounded tree-depth. `arXiv:1303.6471`

[16] Nešetřil, J. and Ossona de Mendez, P. Modeling limits in hereditary classes: Reduction and application to trees. `arXiv:1312.0441`