Random l-colourable structures with a pregeometry

Ove Ahlman and Vera Koponen

July 20, 2012

Abstract

We study finite l-colourable structures with an underlying pregeometry. The probability measure that is used corresponds to a process of generating such structures (with a given underlying pregeometry) by which colours are first randomly assigned to all 1-dimensional subspaces and then relationships are assigned in such a way that the colouring conditions are satisfied but apart from this in a random way. We can then ask what the probability is that the resulting structure, where we now forget the specific colouring of the generating process, has a given property. With this measure we get the following results:

1. A zero-one law.
2. The set of sentences with asymptotic probability 1 has an explicit axiomatisation which is presented.
3. There is a formula $\xi(x, y)$ (not directly speaking about colours) such that, with asymptotic probability 1, the relation "there is an l-colouring which assigns the same colour to x and y" is defined by $\xi(x, y)$.
4. With asymptotic probability 1, an l-colourable structure has a unique l-colouring (up to permutation of the colours).

Keywords: model theory, finite structure, zero-one law, colouring, pregeometry.

1 Introduction

We begin with some background. Let $l \geq 2$ be an integer. Random l-colourable (undirected) graphs were studied by Kolaitis, Prömel and Rothschild in [7] as part of proving a zero-one law for $(l+1)$-clique-free graphs. They proved that random l-colorable graphs satisfies a (labelled) zero-one law, when the uniform probability measure is used. In other words, if $\mathbf{C}_n$ denotes the set of undirected l-colourable graphs with vertices $1, \ldots, n$, then, for every sentence φ in a language with only a binary relation symbol (besides the identity symbol), the proportion of graphs in $\mathbf{C}_n$ which satisfy φ approaches either 0 or 1. They also showed that the proportion of graphs in $\mathbf{C}_n$ which have a unique l-colouring (up to permuting the colours) approaches 1 as $n \to \infty$. In [7] its authors also proved the other statements labelled 1–4 in this paper's abstract, when using the uniform probability measure on $\mathbf{C}_n$, although in case of 3 it is not made explicit. This work was preceeded, and probably stimulated, by an article of Erdös, Kleitman and Rothschild [4] in which it was proved that proportion of triangle-free graphs with vertices $1, \ldots, n$ which are bipartite (2-colourable) approaches 1 as $n \to \infty$.

One can generalise l-colourings from structures with only binary relations to structures with relations of any arity $r \geq 2$ by saying that a structure $\mathcal{M}$ is l-coloured if the elements of $\mathcal{M}$ can be assigned colours from the set of colours $\{1, \ldots, l\}$ in such that if $\mathcal{M} \models R(a_1, \ldots, a_r)$ for some relation symbol R, then $\{a_1, \ldots, a_r\}$ contains at least two elements with different colour. Another way of generalising the notion of l-colouring for graphs, giving the notion of *strong* l-colouring, is to require that if $\mathcal{M} \models R(a_1, \ldots, a_r)$ then $i \neq j$ implies that a_i and a_j have different colours. (If the language has only binary relation symbols then there is no difference between the two notions of l-colouring.)

In [8], Koponen proved that, for every finite relational langauge, if $\mathbf{C}_n$ is the set of l-colourable structures with universe $\{1, \ldots, n\}$, then the statements 1–4 from the abstract hold, for the dimension conditional probability measure, as well as for the uniform probability measure, on $\mathbf{C}_n$. The same results hold if we instead consider *strongly* l-colourable structures. Moreover, the results still hold, for both types of colourings and both probability measures, if we insist that some of the relation symbols are always interpreted as irreflexive and symmetric relations. A consequence of the zero-one law for (strongly) l-colourable structures is that if, for some finite relational vocabulary and for each positive $n \in \mathbb{N}$, $\mathbf{K}_n$ is a set of structures with universe $\{1, \ldots, n\}$ containing every l-colourable structure with that universe, and the probability, using either the dimension conditional measure or the uniform measure, that a random member of $\mathbf{K}_n$ is (strongly) l-colourable approaches 1 as $n \to \infty$, then $\mathbf{K}_n$ has a zero-one law for the corresponding measure; i.e. for every sentence φ, the probability that φ is true in a random $\mathcal{M} \in \mathbf{K}_n$ approaches either 0 or 1 as $n \to \infty$. In [11], Person and Schacht proved that if $\mathcal{F}$ denotes the Fano plane as a 3-hypergraph (so $\mathcal{F}$ has seven elements and seven 3-hyperedges such that every pair of distinct elements are contained in a unique 3-hyperedge) and $\mathbf{K}_n$ is the set of $\mathcal{F}$-free 3-hypergraphs with universe (vertex set) $\{1, \ldots, n\}$, then the proportion of hypergraphs in $\mathbf{K}_n$ which are 2-colourable approaches 1 as $n \to \infty$. Since every 2-colourable 3-hypergraph is $\mathcal{F}$-free, it follows the $\mathcal{F}$-free 3-hypergraphs satisfy a zero-one law if we use the uniform probability measure. As another example, Balogh and Mubayi [1] have proved that if $\mathcal{H}$ denotes the hypergraph with vertices $1, 2, 3, 4, 5$ and 3-hyperedges $\{1, 2, 3\}$, $\{1, 2, 4\}$ and $\{3, 4, 5\}$ and if $\mathbf{K}_n$ denotes the set of $\mathcal{H}$-free 3-hypergraphs with universe $\{1, \ldots, n\}$, then the proportion of hypergraphs in $\mathbf{K}_n$ which are strongly 3-colourable approaches 1 as $n \to \infty$. Since every strongly 3-colourable 3-hypergraph is $\mathcal{H}$-free it follows that $\mathcal{H}$-free 3-hypergraphs satisfy a zero-one law.

In the present article we generalise the work in [8] to the context of (strongly) l-colourable structures with an underlying (combinatorial) pregeometry, also called matroid. Roughly speaking, a structure $\mathcal{M}$ with a pregeometry will be called l-colourable if its 1-dimensional subspaces (i.e. closed subsets of M) can be assigned colours from l given colours in such a way that if R is a relation symbol and $\mathcal{M} \models R(a_1, \ldots, a_r)$, then there are i and j such that the subspaces spanned by a_i and by a_j, respectively, have different colours. A structure $\mathcal{M}$ will be called *strongly* l-colourable if its 1-dimensional subspaces can be assigned colours from l given colours in such a way that if $\mathcal{M} \models R(a_1, \ldots, a_r)$, then any two distinct 1-dimensional subspaces that are included in the closure of $\{a_1, \ldots, a_r\}$ have different colours. The main motivation for this generalisation is to understand how the combinatorics of colourings work out if the elements of a structure are related to each other in a "geometrical way", where in particular, the role of cardinality is taken over by dimension. The main examples of pregeometries for which the results of this article apply are vector spaces, projective spaces and affine spaces over some fixed finite field. Another motivation is the fact that pregeometries have played an important role in the study of infinite models and one may ask to what extent the notion of pregeometry can be combined with the study of asymptotic properties of finite structures. In [8] a framework for studying asymptotic properties of finite structures with an underlying pregeometry was presented. Here we work within that framework, but since we only consider (strongly) l-colourable structures some notions from [8] become simpler here.

We now give rough explanations of the notions that will be involved and the main results. Precise definitions are given in Section 2. We fix an integer $l \geq 2$. L_{pre} denotes a first-order language and for every $n \in \mathbb{N}$, $\mathcal{G}_n$ is an L_{pre}-structure such that $(G_n, \mathrm{cl}_{\mathcal{G}_n})$ is a pregeometry (Definition 2.1) where the closure operator $\mathrm{cl}_{\mathcal{G}_n}$ is definable by L_{pre}-formulas (in a sense given by Definition 2.3 and Assumption 2.12). We will consider the property 'polynomial k-saturation' (Definition 2.10) of the enumerated set $\mathbf{G} = \{\mathcal{G}_n : n \in \mathbb{N}\}$. From Assumption 2.12 it follows that the dimension of G_n approaches infinity as n tends to infinity. The language L_{rel} (from Assumption 2.12) includes L_{pre} and has, in addition, finitely many new relation symbols, all of arity at least 2. By $\mathbf{C}_n$ we denote the set of all L_{rel}-structures $\mathcal{M}$ such that $\mathcal{M} \restriction L_{pre} = \mathcal{G}_n$ and $\mathcal{M}$ is l-colourable (Definition 2.14). By $\mathbf{S}_n$ we denote the set of all L_{rel}-structures $\mathcal{M}$ such that $\mathcal{M} \restriction L_{pre} = \mathcal{G}_n$ and $\mathcal{M}$ is strongly l-colourable (Definition 2.14). For every n, $\delta_n^{\mathbf{C}}$ denotes the probability measure given by Definition 2.17, which means, roughly speaking, that if $\mathbf{X} \subseteq \mathbf{C}_n$, then $\delta_n^{\mathbf{C}}(\mathbf{X})$ is the probability that $\mathcal{M} \in \mathbf{C}_n$ belongs to $\mathbf{X}$ if $\mathcal{M}$ generated by the following procedure: first randomly assign l colours to the 1-dimensional subspaces of M, then, for every relation symbol R that belongs to the vocabulary of L_{rel} but not to the vocabulary of L_{pre}, choose an interpretation of R randomly from all possibilities of interpretations $R^{\mathcal{M}}$ such that the previous assignment of colours is an l-colouring of the resulting structure, and finally forget the colour assignment, leaving us with an L_{rel}-structure. The probability measure $\delta_n^{\mathbf{S}}$ on $\mathbf{S}_n$ is defined similarly (Definition 2.17). If φ is an L_{rel}-sentence then $\delta_n^{\mathbf{C}}(\varphi) = \delta_n^{\mathbf{C}}\big(\{\mathcal{M} \in \mathbf{C}_n : \mathcal{M} \models \varphi\}\big)$ and similarly for $\delta_n^{\mathbf{S}}(\varphi)$.

Theorem 1.1. *Suppose that Assumption 2.12 holds and that $\mathbf{G} = \{\mathcal{G}_n : n \in \mathbb{N}\}$ is polynomially k-saturated for every $k \in \mathbb{N}$. Then, for every L_{rel}-sentence φ, $\delta_n^{\mathbf{C}}(\varphi)$ approaches either 0 or 1, and $\delta_n^{\mathbf{S}}(\varphi)$ approaches either 0 or 1, as $n \to \infty$.*

If F is a field and G is the set of vectors of a vector space or of an affine space over F, or if G is the set of lines of a projective space over F, then (G, cl) where cl is the linear closure operator, affine closure operator, or projective closure operator, respectively, forms a pregeometry (see for example [10] or [9]).

Theorem 1.2. *Suppose that the conditions of Assumption 2.12 hold and that for some finite field F one of the following three cases holds for every $n \in \mathbb{N}$: G_n is an (a) n-dimensional vector space, or (b) n-dimensional affine space, or (c) n-dimensional projective space, over F, and $\mathrm{cl}_{\mathcal{G}_n}$ is the linear, affine or projective closure operator on G_n, respectively. Moreover, assume that L_{pre} is the generic language L_{gen} from Example 2.4, with the intepretations of symbols given in that example.*
(i) There is an L_{rel}-formula $\xi(x, y)$ such that the $\delta_n^{\mathbf{C}}$-probability that the following holds for $\mathcal{M} \in \mathbf{C}_n$ approaches 1 as $n \to \infty$:

> *For all $a, b \in M - \mathrm{cl}_{\mathcal{M}}(\emptyset)$, $\mathcal{M} \models \xi(a, b)$ if and only if every l-colouring of $\mathcal{M}$ gives a and b the same colour.*

(ii) $\lim_{n\to\infty} \delta_n^{\mathbf{C}}\big(\{\mathcal{M} \in \mathbf{C}_n : \ \mathcal{M}$ *has a unique l-colouring*$\}\big) = 1$.
(iii) $\lim_{n\to\infty} \delta_n^{\mathbf{C}}\big(\{\mathcal{M} \in \mathbf{C}_n : \ \mathcal{M}$ *is not l'-colourable if $l' < l$*$\}\big) = 1$.
*(iv) The set $\{\varphi \in L_{rel} : \lim_{n\to\infty} \delta_n^{\mathbf{C}}(\varphi) = 1\}$ forms a countably categorical theory which can be explicitly axiomatised (as in Section 5) by L_{rel}-sentences of the form $\forall \bar{x} \exists \bar{y} \psi(\bar{x}, \bar{y})$ where ψ is quantifier-free, mainly in terms of what we call l-*colour compatible extension axioms, *which involve the formula $\xi(x, y)$ from part (i).*
(v) The statements (i)–(iv) hold if we assume that, for each $n \in \mathbb{N}$, G_n is an n-dimensional vector space over F, $\mathrm{cl}_{\mathcal{G}_n}$ is the linear closure operator on G_n and that

L_{pre} is the language L_F from Example 2.6, with the interpretation of symbols from that example.

The assumptions of Theorem 1.2 imply the assumptions of Theorem 1.3, which is explained in Example 2.11. That is, when dealing with strongly l-colourable structures, the assumptions on the underlying pregeometries can be weaker. By a subspace of a pregeometry we mean a closed set with respect to the given closure operator (Definition 2.3).

Theorem 1.3. *Suppose that the conditions of Assumption 2.12 hold and that $\mathbf{G}$ is polynomially k-saturated for every $k \in \mathbb{N}$. Also assume that for every $n \in \mathbb{N}$, every 2-dimensional subspace of $\mathcal{G}_n$ has at most l different 1-dimensional subspaces.*
(i) There is an L_{rel}-formula $\xi(x, y)$ such that the $\delta_n^{\mathbf{S}}$-probability that the following holds for $\mathcal{M} \in \mathbf{S}_n$ approaches 1 as $n \to \infty$:

> *For all $a, b \in M - \mathrm{cl}_{\mathcal{M}}(\emptyset)$, $\mathcal{M} \models \xi(a, b)$ if and only if every l-colouring of $\mathcal{M}$ gives a and b the same colour.*

(ii) $\lim_{n\to\infty} \delta_n^{\mathbf{S}}\big(\{\mathcal{M} \in \mathbf{S}_n : \ \mathcal{M}$ has a unique strong l-colouring$\}\big) = 1$.
(iii) $\lim_{n\to\infty} \delta_n^{\mathbf{S}}\big(\{\mathcal{M} \in \mathbf{S}_n : \ \mathcal{M}$ is not strongly l'-colourable if $l' < l\}\big) = 1$.
(iv) Suppose, moreover, that the formulas of L_{pre} which, according to Assumption 2.12, define the pregeometry $\mathbf{G} = \{\mathcal{G}_n : n \in \mathbb{N}\}$ are quantifier-free. Then the set $\{\varphi \in L_{rel} : \lim_{n\to\infty} \delta_n^{\mathbf{S}}(\varphi) = 1\}$ forms a countably categorical theory which can be explicitly axiomatised (as in Section 5) by L_{rel}-sentences of the form $\forall \bar{x} \exists \bar{y} \psi(\bar{x}, \bar{y})$ where ψ is quantifier-free, mainly in terms of what we call l-colour compatible extension axioms*, which involve the formula $\xi(x, y)$ from part (i).*

It turns out that Theorem 1.1 follows rather straightforwardly from Theorem 7.32 in [8] when we have proved Lemma 2.21 below. However, Theorem 1.1 in itself does not give information about which sentences have asymptotic probability 1 (or 0), or about properties of the theory consisting of those sentences which have asymptotic probability 1. Neither does it tell us anything about typical properties of large (strongly) l-colourable structures. In order to prove part (i) of Theorems 1.2 and 1.3, which give information of this kind, we treat l-colourable structures and strongly l-colourable structures separately and need to add some assumption(s). The case of strong l-colourings is the easier one and is treated in Section 3; that is, most of the argument leading to part (i) of Theorem 1.3 is carried out in Section 3. The main part of the proof of (i) of Theorem 1.2, dealing with (not necessarily strong) l-colourings, is carried out in Section 4 where we use a theorem from structural Ramsey theory by Graham, Leeb and Rothschild [5].

Once we have established part (i) of Theorems 1.2 and 1.3, which, as said above, is done separately, parts (ii)–(iv) (and (v) of Theorem 1.2) can be proved in a uniform way, that is, it is no longer necessary to distinguish between l-colourable structures and strongly l-colourable structures. This is done in Section 5. It is possible to read Section 5 directly after Section 2 and then consider the details of definability of colorings in Sections 3 and 4, which are independent of each other.

The theorems above generalise the results of Section 9 of [8] to the situation when a nontrivial pregeometry (subject to certain conditions) is present. In other words, if the closure of a set A is always A (so every set is closed) and we let L_{pre} be the language whose vocabulary contains only the identity symbol '=', and, for every $n \in \mathbb{N}$, $\mathcal{G}_n$ is the unique (under these assumtions) L_{pre}-structure with universe $\{1, \ldots, n+1\}$, then (i)–(iv) of Theorems 1.2 and 1.3 hold by results in Section 9 of [8]. Theorem 1.1 includes this case, without reformulation.

Remark 1.4. One may want to consider only L_{rel}-structures in which certain relation symbols from the vocabulary of L_{rel} are always interpreted as irreflexive and symmetric relations (see beginning of Section 2). Theorems 1.1 and 1.2 hold with exactly the same proofs also in this situation. This claim uses that all results of [8] (see Remark 2.1 of that article) hold whether or not one assumes that certain relation symbols are always interpreted as irreflexive and symmetric relations. If a technical assumption is added, explained in Remark 3.7, then Theorem 1.3 also holds in the context when some relation symbols are always interpreted as irreflexive and symmetric relations.

Remark 1.5. In [8], results corresponding to Theorems 1.1–1.3, in the case of trivial pregeometries (i.e. when every set is closed), where proved also for the *uniform probability measure*. The proof used the fact, proved in Section 10 of [8], that, when the pregeometries considered are trivial, then the probability, with the uniform probability measure, that a random (strongly) l-colourable structure with n elements has an l-colouring with relatively even distribution of colours, approaches 1 as $n \to \infty$. We believe that the same is true in the context of Theorems 1.2 and 1.3 above, by proofs analogous to those in Section 10 of [8]. But when the underlying pregeometries are no longer assumed to be trivial, then this condition alone seems to be insufficient for proving analogoues of Theorems 1.1–1.3 if $\delta_n^{\mathbf{C}}$ is replaced by the uniform probability measure on $\mathbf{C}_n$ and $\delta_n^{\mathbf{S}}$ is replaced by the uniform probability measure on $\mathbf{S}_n$. In other words, it appears to be a more difficult task to transfer the results of this article to the uniform probability measure (if possible at all) than was the case in [8].

This article ends with a small errata to [8], which makes explicit some assumptions, used implicity in Section 8 of [8], but not stated explicitly in the places in Sections 7–8 of [8] where they are relevant.

2 Pregeometries and (strongly) l-colourable structures

The notation used here is more or less standard; see [3, 9] for example. The formal languages considered are always first-order and denoted L, often with a subscript. Such L denotes the set of first-order formulas over some vocabulary, also called signature, consisting of constant-, function- and/or relation symbols. First-order structures are denoted with calligraphic letters $\mathcal{A}, \mathcal{B}, \ldots, \mathcal{M}, \mathcal{N}, \ldots$, and their universes with the corresponding non-calligraphic letters $A, B, \ldots, M, N, \ldots$. If the vocabulary of a language L has no constant or function symbols, then we allow an L-structure to have an empty universe. Finite sequences/tuples of objects, usually elements from structures or variables, are denoted with $\bar{a}$, $\bar{x}$, etc. By $\bar{a} \in A$ we mean that every element of the sequence $\bar{a}$ belongs to the set A, and $|A|$ denotes the cardinality of A. A function $f : M \to N$ is called an ***embedding*** of $\mathcal{M}$ into $\mathcal{N}$ if, for every constant symbol c, $f(c^{\mathcal{M}}) = c^{\mathcal{N}}$, for every function symbol g and tuple $(a_1, \ldots, a_r) \in M^r$ where r is the arity of g, $g^{\mathcal{N}}(f(a_1), \ldots, f(a_r)) = f(g^{\mathcal{M}}(a_1, \ldots, a_r))$, and for every relation symbol R and tuple $(a_1, \ldots, a_r) \in M^r$ where r is the arity of R, $\mathcal{M} \models R(a_1, \ldots, a_r) \iff \mathcal{N} \models R(f(a_1), \ldots, f(a_r))$. It follows that an ***isomorphism*** from $\mathcal{M}$ to $\mathcal{N}$ is the same as a surjective embedding from $\mathcal{M}$ to $\mathcal{N}$. Suppose that L' is a language whose vocabulary is included in the vocabulary of L. For any L-structure $\mathcal{M}$, by $\mathcal{M} \upharpoonright L'$ we denote the reduct of $\mathcal{M}$ to L'. If $\mathcal{M}$ is an L-structure and $A \subseteq M$, then $\mathcal{M} \upharpoonright A$ denotes the substructure of $\mathcal{M}$ which is generated by the set A, that is, $\mathcal{M} \upharpoonright A$ is the unique substructure $\mathcal{N}$ of $\mathcal{M}$ such that $A \subseteq N \subseteq M$ and if $\mathcal{N}' \subseteq \mathcal{M}$ and $A \subseteq N' \subseteq M$, then $\mathcal{N} \subseteq \mathcal{N}'$. A third meaning of the symbol '$\upharpoonright$' with respect to structures is given by Definition 2.16. Suppose that A is a set, $n \geq 2$ and $R \subseteq A^n$ an n-ary relation on A. We call R ***irreflexive*** if $(a_1, \dots, a_n) \in R$ implies that $a_i \neq a_j$ whenever $i \neq j$. We call R ***symmetric*** if $(a_1, \dots, a_n) \in R$ implies that $(\pi(a_1), \dots, \pi(a_r)) \in R$ for every permutation π of $\{a_1, \dots, a_n\}$. For any set A, $\mathcal{P}(A)$ denotes the power set of A. A usual, we call a formula ***existential*** if it has the form

$$\exists y_1, \dots, y_m \varphi(x_1, \dots, x_k, y_1, \dots, y_m)$$

where φ is quantifier free.

Definition 2.1. We say that (A, cl), with $\mathrm{cl} : \mathcal{P}(A) \to \mathcal{P}(A)$ is a ***pregeometry*** (also called *matroid*) if it satisfies the following for all $X, Y \subseteq A$:

1. (Reflexivity) $X \subseteq \mathrm{cl}(X)$.
2. (Monotonicity) $Y \subseteq \mathrm{cl}(X) \Rightarrow \mathrm{cl}(Y) \subseteq \mathrm{cl}(X)$.
3. (Exchange property) If $a, b \in A$ then $a \in \mathrm{cl}(X \cup \{b\}) - \mathrm{cl}(X) \Rightarrow b \in \mathrm{cl}(X \cup \{a\})$.
4. (Finite Character) $\mathrm{cl}(X) = \bigcup\{\mathrm{cl}(X_0) : X_0 \subseteq X \text{ and } |X_0| \text{ is finite}\}$.

If $X, Y \subseteq A$ then we say that X is ***independent from*** Y if $\mathrm{cl}(X) \cap \mathrm{cl}(Y) = \mathrm{cl}(\emptyset)$. From the exchange property it follows that X is indepependent from Y if and only if Y is independent from X (symmetry of independence). We will often write $\mathrm{cl}(a_1, \dots, a_n)$ instead of $\mathrm{cl}(\{a_1, \dots, a_n\})$ and say 'a is independent from b' instead of '$\{a\}$ is independent from $\{b\}$ over $\emptyset$'. We say that a set X is ***independent*** if for, each $a \in X$, we have that $\{a\}$ is independent from $X - \{a\}$. We say that a set $X \subseteq A$ is ***closed*** (in (A, cl)) if $\mathrm{cl}(X) = X$. For $X \subseteq A$, the dimension of X is defined as $\dim(X) = \inf\{|Y| : Y \subseteq X \text{ and } X \subseteq \mathrm{cl}(Y)\}$. For more about pregeometries the reader is refered to [9, 10] for example. We will use the following lemma, which has probably been proved somewhere, but for the sake of completeness we give a proof of it here.

Lemma 2.2. *Let $\mathcal{A} = (A, \mathrm{cl})$ be a pregeometry. If $\{a, v_1, ..., v_m, w_1, ..., w_n\} \subseteq A$ is an independent set then* $\mathrm{cl}(a, v_1, ..., v_m) \cap \mathrm{cl}(a, w_1, ..., w_n) = \mathrm{cl}(a)$

Proof. Suppose that $\{a, v_1, ..., v_m, w_1, ..., w_n\} \subseteq A$ is an independent set. By reflexivity $a \in \mathrm{cl}(a, v_1, ..., v_m) \cap \mathrm{cl}(a, w_1, ..., w_n)$ and so by monotonicity

$$\mathrm{cl}(a) \ \subseteq \ \mathrm{cl}(a, v_1, ..., v_m) \cap \mathrm{cl}(a, w_1, ..., w_n).$$

For the opposite direction we assume that $x \in \mathrm{cl}(a, v_1, ..., v_m) \cap \mathrm{cl}(a, w_1, ..., w_n)$ and use induction over n to prove that $x \in \mathrm{cl}(a)$.

Base case: If $n = 0$ then $\mathrm{cl}(a, w_1, ..., w_n) = \mathrm{cl}(a)$ so, as $x \in \mathrm{cl}(a)$, we are done.

Induction step: Suppose that $x \in \mathrm{cl}(a, v_1, ..., v_m) \cap \mathrm{cl}(a, w_1, ..., w_{n+1})$, so we have two cases to consider:

$$\begin{array}{llrl} \text{either} & & x \in & \big(\mathrm{cl}(a, v_1, ..., v_m) \ \cap \ \mathrm{cl}(a, w_1, ..., w_{n+1})\big) \\ & & - & \big(\mathrm{cl}(a, v_1, ..., v_m) \ \cap \ \mathrm{cl}(a, w_1, ..., w_n)\big), \\ \text{or} & & x \in & \mathrm{cl}(a, v_1, ..., v_m) \ \cap \ \mathrm{cl}(a, w_1, ..., w_n). \end{array}$$

In the first case we get the consequence that $x \in \mathrm{cl}(a, w_1, ..., w_{n+1}) - \mathrm{cl}(a, w_1, ..., w_n)$ and hence by the exchange property we get that $w_{n+1} \in \mathrm{cl}(a, w_1, ..., w_n, x)$. We already know that $x \in \mathrm{cl}(a, v_1, ..., v_m)$ and by also using the assumption that $\{a, v_1, ..., v_m, w_1, ..., w_n\}$ is independent we get that

$$1 + m + n = \dim(a, v_1, ..., v_m, w_1, ..., w_n) = \dim(a, v_1, ..., v_m, w_1, ..., w_n, x) =$$

$$\dim(a, v_1, ..., v_m, w_1, ..., w_n, w_{n+1}, x) = \dim(a, v_1, ..., v_m, w_1, ..., w_{n+1}) = 1 + m + n + 1,$$

so $1 + m + n = 1 + m + n + 1$, a contradiction. Hence,

$$x \in \operatorname{cl}(a, v_1, ..., v_m) \cap \operatorname{cl}(a, w_1, ..., w_n),$$

so by the induction hypothesis we get that $x \in \operatorname{cl}(a)$.

By induction we conclude that $\operatorname{cl}(a, v_1, ..., v_m) \cap \operatorname{cl}(a, w_1, ..., w_n) \subseteq \operatorname{cl}(a)$ holds for all n, which finishes the proof. $\square$

We will consider first-order structures $\mathcal{M}$ for which there is a closure operator cl on M such that (M, cl) is a pregeometry and, for each n, the relation $x_{n+1} \in \operatorname{cl}(x_1, \ldots, x_n)$ is definable by a first-order formula without parameters. More precisely, we have the following definition.

Definition 2.3. (i) We say that an L-structure $\mathcal{A}$ is a ***pregeometry*** if there are L-formulas $\theta_n(x_1, \ldots, x_{n+1})$, for all $n \in \mathbb{N}$, such that if the operator $\operatorname{cl}_{\mathcal{A}} : \mathcal{P}(A) \to \mathcal{P}(A)$ is defined by (a) and (b) below, then $(A, \operatorname{cl}_{\mathcal{A}})$ is a pregeometry:

(a) For every $n \in \mathbb{N}$, every sequence $b_1, \ldots, b_n \in A$ and every $a \in A$,

$$a \in \operatorname{cl}_{\mathcal{A}}(b_1, \ldots, b_n) \iff \mathcal{A} \models \theta_n(b_1, \ldots, b_n, a).$$

(b) For every $B \subseteq A$ and every $a \in A$, $a \in \operatorname{cl}_{\mathcal{A}}(B)$ if and only if $a \in \operatorname{cl}_{\mathcal{A}}(b_1, \ldots, b_n)$ for some $b_1, \ldots, b_n \in B$.

(ii) Suppose that $\mathcal{A}$ is a pregeometry in the sense of the above definition. Then, for every $B \subseteq A$, $\dim_{\mathcal{A}}(B)$ denotes the ***dimension*** of B with respect to the closure operator $\operatorname{cl}_{\mathcal{A}}$. In other words, $\dim_{\mathcal{A}}(B) = \min\{|B'| : B' \subseteq B \text{ and } \operatorname{cl}_{\mathcal{A}}(B') \supseteq B\}$. We sometimes abbreviate $\dim_{\mathcal{B}}(B)$ with $\dim(\mathcal{B})$. A closed subset of A is also called a ***subspace*** of $\mathcal{A}$. A substructure $\mathcal{B} \subseteq \mathcal{A}$ is called ***closed*** if its universe B is closed in $(A, \operatorname{cl}_{\mathcal{A}})$.
(iii) Suppose that $\mathbf{G}$ is a set of L-structures. We say that $\mathbf{G}$ is a ***pregeometry*** if there are L-formulas $\theta_n(x_1, \ldots, x_{n+1})$, for all $n \in \mathbb{N}$, such that for each $\mathcal{A} \in \mathbf{G}$, $(A, \operatorname{cl}_{\mathcal{A}})$ is a pregeometry if $\operatorname{cl}_{\mathcal{A}}$ is defined by (a) and (b).

It may happen that for an L-structure $\mathcal{A}$ there are L-formulas θ_n and θ'_n, for $n \in \mathbb{N}$, such that the sequence θ_n, $n \in \mathbb{N}$, defines a different pregeometry on A (according to Definition 2.3 (i)) than does the sequence θ'_n, $n \in \mathbb{N}$. When we use these notions it will, however, be clear that we fix a sequence of formulas θ_n, $n \in \mathbb{N}$, and the pregeometry that they define on each structure from a given set, which will be denoted $\mathbf{G}$.

Example 2.4. (**Generic example**) Every pregeometry (A, cl) can be viewed as a first-order structure $\mathcal{A}$ in the following way. For every $n \in \mathbb{N}$, let P_n be an $(n+1)$-ary relation symbol and let the vocabulary of L_{gen} be $\{P_n : n \in \mathbb{N}\}$. For every $n \in \mathbb{N}$ and every $(a_1, \ldots, a_{n+1}) \in A^{n+1}$, let $(a_1, \ldots, a_{n+1}) \in (P_n)^{\mathcal{A}}$ if and only if $a_{n+1} \in \operatorname{cl}(a_1, \ldots, a_n)$. Then $\mathcal{A}$ is a pregeometry in the sense of Definition 2.3 (i) and $\operatorname{cl} = \operatorname{cl}_{\mathcal{A}}$. It follows that every set of pregeometries $\mathbf{G}$, viewed as L_{gen}-structures is a pregeometry in the sense of Definition 2.3 (iii).

Example 2.5. (**Trivial pregeometries**) If A is a set and $\operatorname{cl}(B) = B$ for every $B \subseteq A$, then (A, cl) is a pregeometry, called a ***trivial preometry***. Let $L_\emptyset$ be the language with vocabulary $\emptyset$, so $L_\emptyset$ can only express whether elements are identical or not. If, for $n > 0$, $\theta_n(x_1, \ldots, x_{n+1})$ denotes a formula which expresses that "x_{n+1} is identical to one of $x_1, \ldots, x_n$", and $\theta_0(x_1)$ is some formula which can never be satisfied, then every $L_\emptyset$-structure is a pregeometry in the sense of Definition 2.3 (i). Moreover, every set $\mathbf{G}$ of $L_\emptyset$-structures is a pregeometry in the sense of Definition 2.3 (iii).

Example 2.6. (**Vector spaces over a finite field**) Let F be a field. Let L_F be the language with vocabulary $\{0, +\} \cup \{f : f \in F\}$, where 0 is a constant symbol, $+$ a binary function symbol and each $f \in F$ represents a unary function symbol. Every vector space over F can be viewed as an L_F-structure by interpreting 0 as the zero vector, $+$ as vector addition and each $f \in F$ as scalar multiplication by f. Now add the assumption that F is *finite*. If, for every $n \in \mathbb{N}$, $\theta_n(x_1, \ldots, x_{n+1})$ is an L_F-formula that expresses that "x_{n+1} belongs to the linear span of $x_1, \ldots, x_n$", then every F-vector space $\mathcal{V}$, viewed as an L_F-structure, is a pregeometry accordning to Definition 2.3 (i). In particular, every set $\mathbf{G}$ of vector spaces over a finite field F, viewed as L_F-structures, is a pregeometry according to Definition 2.3 (iii).

Definition 2.7. We say that the pregeometry $\mathbf{G} = \{\mathcal{G}_n : n \in \mathbb{N}\}$ is ***uniformly bounded*** if there is a function $f : \mathbb{N} \to \mathbb{N}$ such that for every $n \in \mathbb{N}$ and every $X \subseteq G_n$, $\big|\mathrm{cl}_{\mathcal{G}_n}(X)\big| \leq f\big(\dim_{\mathcal{G}_n}(X)\big)$.

Example 2.8. (**Vector space pregeometries**) Let $\mathbf{G} = \{\mathcal{G}_n : n \in \mathbb{N}\}$ is a pregeometry. Suppose that, for every $n \in \mathbb{N}$, $(G_n, \mathrm{cl}_{\mathcal{G}})$ is isomorphic (as a pregeometry) with $(V_n, \mathrm{cl}_{\mathcal{V}_n})$ where each $\mathcal{V}_n$ is a vector space of dimension n over a (fixed) finite field F and $\mathrm{cl}_{\mathcal{V}_n}$ is linear span in $\mathcal{V}_n$. Then $\mathbf{G} = \{\mathcal{G}_n : n \in \mathbb{N}\}$ is uniformly bounded. We get the same conclusion if, instead, each $\mathcal{V}_n$ is a projective space over F with dimension n, or if each $\mathcal{V}_n$ is an affine space over F with dimension n.

Example 2.9. (**Sub-pregeometries of $\mathbb{R}^n$**) Let cl_n denote the linear closure operator in $\mathbb{R}^n$. It is straightforward to verify that whenever $X_n \subseteq \mathbb{R}^n$ and cl'_n is defined by $\mathrm{cl}'_n(A) = \mathrm{cl}_n(A) \cap X_n$ for every $A \subseteq X_n$, then (X_n, cl'_n) is a pregeometry. For every positive integer n choose *finite* $X_n \subseteq \mathbb{R}^n$ and, for all $n \in \mathbb{N}$, let $\mathcal{G}_n = (X_{n+1}, \mathrm{cl}'_{n+1})$. Let L_{gen} be the language from Example 2.4. Then each $\mathcal{G}_n$ can be viewed as a first-order structure in the way explained in that example. It follows that $\mathbf{G} = \{\mathcal{G}_n : n \in \mathbb{N}\}$ is a pregeometry in the sense of Definition 2.3 (iii). Suppose that, in addition, the choice of each X_n is made in such a way that for every $k > 0$ there is m_k such that if $n > 0$ and $a_1, \ldots, a_k \in X_n$, then $|\mathrm{cl}(a_1, \ldots, a_k) \cap X_n| \leq m_k$. Then $\mathbf{G}$ is uniformly bounded.

Definition 2.10. Let $k \in \mathbb{N}$. We say that the pregeometry $\mathbf{G} = \{\mathcal{G}_n : n \in \mathbb{N}\}$ is ***polynomially k-saturated*** if there are a sequence of natural numbers $(\lambda_n : n \in \mathbb{N})$ with $\lim_{n\to\infty} \lambda_n = \infty$ and a polynomial $P(x)$ such that for every $n \in \mathbb{N}$:

(1) $\lambda_n \leq |G_n| \leq P(\lambda_n)$, and

(2) whenever $\mathcal{A}$ is a closed substructure of $\mathcal{G}_n$ and there are $\mathcal{G}$ and $\mathcal{B} \supset \mathcal{A}$ such that $\mathcal{A}$ and $\mathcal{B}$ are closed substructures of $\mathcal{G}$, $\mathcal{G}$ is isomorphic with some member of $\mathbf{G}$ and $\dim_{\mathcal{G}}(A) + 1 = \dim_{\mathcal{G}}(B) \leq k$, then there are closed substructures $\mathcal{B}_i \subseteq \mathcal{M}$, for $i = 1, \ldots, \lambda_n$, such that $B_i \cap B_j = A$ if $i \neq j$, and each $\mathcal{B}_i$ is isomorphic with $\mathcal{B}$ via an isomorphism that fixes A pointwise.

Example 2.11. (i) Let $L_\emptyset$ be the "empty" language from Example 2.5. It is straightforward to verify that if for every $n \in \mathbb{N}$, $\mathcal{G}_n$ is the unique $L_\emptyset$-structure with universe $\{1, \ldots, n+1\}$, then $\mathbf{G}$ is polynomially k-saturated for every $k \in \mathbb{N}$.
(ii) Let F be a finite field and let $L = L_{gen}$ as in Example 2.4 or $L = L_F$ as in Example 2.6. For $n \in \mathbb{N}$ let $\mathcal{V}_n$ be a vector space over F of dimension n. Each $\mathcal{V}_n$ gives rise to a pregeometry (V_n, cl_n) where cl_n is linear span, and each $\mathcal{V}_n$ can be viewed as an L-structure, call it $\mathcal{G}_n$, as in any one of the mentioned examples (depending on whether we take $L = L_{gen}$ or $L = L_F$). Then the pregeometry $\mathbf{G} = \{\mathcal{G}_n : n \in \mathbb{N}\}$ is polynomially

k-saturated for every $k \in \mathbb{N}$. This is explained in some more detail in [8] and the proofs in Section 3.2 of [2] translate to the present context.
(iii) Let F be a finite field. If $\mathcal{G}_n$, for $n \in \mathbb{N}$, is instead the pregeometry obtained from a projective space over F with dimension n, viewed as an L_{gen}-structure as in Example 2.4, then $\mathbf{G} = \{\mathcal{G}_n : n \in \mathbb{N}\}$ is polynomially k-saturated for every $k \in \mathbb{N}$. The same holds if 'projective space' is replaced with 'affine space'. These facts are proved are proved in a slightly different context Section 3.2 of [2], but the proofs there translate straightforwardly to the present context.

Assumption 2.12. We now fix some notation and assumptions for the rest of the paper.

(1) Let $l \geq 2$ be an integer, $P_1, \ldots, P_l$ unary relation symbols and let $V_{col} = \{P_1, \ldots, P_l\}$. The symbols P_i represent colours. Let V_{rel} be a finite nonempty set of relation symbols all of which have arity at least 2. Let ρ be the maximal arity among the relation symbols in V_{rel}.

(2) Let L_{pre} be a language with vocabulary V_{pre}, which is disjoint from both V_{col} and V_{rel}. Suppose that $\mathbf{G} = \{\mathcal{G}_n : n \in \mathbb{N}\}$ is a set of finite L_{pre}-structures where G_n is the universe of $\mathcal{G}_n$ and $\mathbf{G}$ is a pregeometry in the sense of Definition 2.3 (iii). Also, assume that the L_{pre}-formulas $\theta_n(x_1, \ldots, x_{n+1})$, $n \in \mathbb{N}$, define the pregeometry according to Definition 2.3.

(3) Let L_{col} be the language with vocabulary $V_{pre} \cup V_{col}$, let L_{rel} be the language with vocuabulary $V_{pre} \cup V_{rel}$ and let L be the language with vocabulary $V_{pre} \cup V_{col} \cup V_{rel}$.

(4) $\mathbf{G}$ is uniformly bounded and, for every $n \in \mathbb{N}$, if $A \subseteq G_n$ is closed (with respect to $\mathrm{cl}_{\mathcal{G}_n}$) then A is the universe of a substructure of $\mathcal{G}_n$ (or equivalently, A contains all interpretations of constant symbols and is closed under interpretations of function symbols, if such occur in the language).

(5) For every $n \in \mathbb{N}$, if $\mathcal{A}$ is a closed substructure of $\mathcal{G}_n$ and $a_1, \ldots, a_{n+1} \in A$, then $a_{n+1} \in \mathrm{cl}_{\mathcal{G}_n}(a_1, \ldots, a_n) \iff \mathcal{A} \models \theta_n(a_1, \ldots, a_{n+1})$. In other words, the restriction of $\mathrm{cl}_{\mathcal{G}_n}$ to A is definable in $\mathcal{A}$ by the same formulas θ_n.

(6) For every $n \in \mathbb{N}$, if $\mathcal{A}$ is a closed substructure of $\mathcal{G}_n$, then there is m such that $\mathcal{A} \cong \mathcal{G}_m$. Also assume that $\lim_{n\to\infty} \dim(\mathcal{G}_n) = \infty$ and, for every $n \in \mathbb{N}$, $\mathcal{G}_n \restriction \mathrm{cl}_{\mathcal{G}_n}(\emptyset) \cong \mathcal{G}_0$.

(7) For every $n \in \mathbb{N}$, there is a "characteristic" quantifier-free L_{pre}-formula $\chi_{\mathcal{G}_n}(x_1, \ldots, x_{m_n})$ of $\mathcal{G}_n$, where $m_n = |G_n|$, such that if $\mathcal{A}$ is an L_{pre}-structure in which the formulas θ_n define a pregeometry (according to Definition 2.3) and $\mathcal{A} \models \chi_{\mathcal{G}_n}(a_1, \ldots, a_s)$ for some enumeration $a_1, \ldots, a_s$ of A, then $\mathcal{A} \cong \mathcal{G}_n$.

Remark 2.13. (i) If θ_n is quantifier free for every $n \in \mathbb{N}$, then (5) holds. Note that in all examples above, it is possible to let θ_n be quantifier free for every $n \in \mathbb{N}$, either by using using the "generic" language L_{gen} from Example 2.4, or by using some of the other languages mentioned in the examples.
(ii) Observe that by (5), if $\mathcal{A}$ is a closed substructure of $\mathcal{G}_n$ then the formulas θ_n define a pregeometry $(A, \mathrm{cl}_{\mathcal{A}})$, according to Definition 2.3, and for all $X \subseteq A$, $\mathrm{cl}_{\mathcal{A}}(X) = \mathrm{cl}_{\mathcal{G}_n}(X)$. By (5)–(6), for every $k \in \mathbb{N}$, there are only finitely many L_{pre}-structures $\mathcal{A}$, up to isomorphism, such that for some n, $\mathcal{A} \subseteq \mathcal{G}_n$ and $\dim_{\mathcal{G}_n}(A) \leq k$.
(iii) Condition (7) obviously holds if the vocabulary V_{pre} is finite. But we want to be

able to consider languages with infinite vocabularies, such as the langauge L_{gen} from Example 2.4. If we take $L_{pre} = L_{gen}$ with the same interpretations as in Example 2.4 and (1)–(6) hold, then also (7) holds.

Definition 2.14. (i) We say that an L-structure $\mathcal{N}$ is ***l*-coloured** if there is an L-structure $\mathcal{M}$ such that $\mathcal{M} \cong \mathcal{N}$, $\mathcal{M} \upharpoonright L_{pre} = \mathcal{G}_n$ for some $n \in \mathbb{N}$ and $\mathcal{M}$ satisfies the following four conditions:

(1) For all $a \in M$, $\mathcal{M} \models P_1(a) \vee ... \vee P_l(a)$ if and only if $a \notin \mathrm{cl}_{\mathcal{G}_n}(\emptyset)$, in other words, an element has a colour if and only if it does not belong to the closure of $\emptyset$.

(2) If $R \in V_{rel}$ has arity $m \geq 2$ and $a_1, \ldots, a_m \in \mathrm{cl}_{\mathcal{G}_n}(\emptyset)$, then $\mathcal{M} \models \neg R(a_1, \ldots, a_m)$.

(3) For all $i, j \in \{1, ..., l\}$ such that $i \neq j$ and all $a, b \in M - \mathrm{cl}_{\mathcal{G}_n}(\emptyset)$ such that $a \in \mathrm{cl}_{\mathcal{G}_n}(b)$ we have that $\mathcal{M} \models \neg(P_i(a) \wedge P_j(b))$, i.e. dependent elements not belonging to the closure of $\emptyset$ have the same colour.

(4) If $R \in V_{rel}$ has arity $m \geq 2$ and $\mathcal{M} \models R(a_1, ..., a_m)$ then there are $b, c \in \mathrm{cl}_{\mathcal{G}_n}(a_1, ..., a_m) - \mathrm{cl}_{\mathcal{G}_n}(\emptyset)$ such that for every $k \in \{1, ..., l\}$ we have $\mathcal{M} \models \neg(P_k(b) \wedge P_k(c))$.

(ii) We say that $\mathcal{N}$ is ***strongly l*-coloured** if there is an L-structure $\mathcal{M}$ such that $\mathcal{M} \cong \mathcal{N}$, $\mathcal{M} \upharpoonright L_{pre} = \mathcal{G}_n$ for some $n \in \mathbb{N}$ and $\mathcal{M}$ satisfies (1)–(4) above and (5) below:

(5) If $R \in V_{rel}$ has arity $m \geq 2$ and $\mathcal{M} \models R(a_1, ..., a_m)$, then for all $b, c \in \mathrm{cl}_{\mathcal{G}_n}(a_1, ..., a_m) - \mathrm{cl}(\emptyset)$ that are linearly independent ($b \notin \mathrm{cl}_{\mathcal{G}_n}(c)$) and every $k \in \{1, ..., l\}$, $\mathcal{M} \models \neg(P_k(b) \wedge P_k(c))$.

(iii) An L_{rel}-structure is called ***(strongly) l*-colourable** if it can be expanded to an L-structure that is (strongly) l-coloured.
(iv) For $n \in \mathbb{N}$, let $\mathbf{K}_n$ denote the set of all l-coloured structures $\mathcal{M}$ such that $\mathcal{M} \upharpoonright L_{pre} = \mathcal{G}_n$ and let $\mathbf{SK}_n$ denote the set of all strongly l-coloured structures $\mathcal{M}$ such that $\mathcal{M} \upharpoonright L_{pre} = \mathcal{G}_n$. Similarly, let $\mathbf{C}_n$ and $\mathbf{S}_n$ denote the set of l-colourable, respectively, strongly l-colourable structures $\mathcal{M}$ such that $\mathcal{M} \upharpoonright L_{pre} = \mathcal{G}_n$. Finally, let $\mathbf{K} = \bigcup_{n \in \mathbb{N}} \mathbf{K}_n$, $\mathbf{SK} = \bigcup_{n \in \mathbb{N}} \mathbf{SK}_n$ $\mathbf{C} = \bigcup_{n \in \mathbb{N}} \mathbf{C}_n$ and $\mathbf{S} = \bigcup_{n \in \mathbb{N}} \mathbf{S}_n$

It follows that if $\mathcal{M}$ is (strongly) l-colourable (or l-coloured) and all $a_1, \ldots, a_r \in M$ belong to the same 0- or 1-dimensional subspace, then $\mathcal{M} \not\models R(a_1, \ldots, a_r)$.

Remark 2.15. (i) If we say that $\mathcal{M}$ is (strongly) l-coloured then it is presupposed that $\mathcal{N}$ is an is an L-structure. If we say that $\mathcal{M}$ is (strongly) l-colourable then it is presupposed that $\mathcal{M}$ is an L_{rel}-structure.
(ii) From Definition 2.14 it follows that if $\mathcal{M}$ is (strongly) l-coloured or (strongly) l-colourable, then the formulas $\theta_n(x_1, \ldots, x_{n+1})$ from Assumption 2.12 define a pregeometry on M according to Definition 2.3. We always have this pregeometry in mind when speaking of the pregeometry of an (strongly) l-coloured or (strongly) l-colourable structure.
(iii) From the definition of (strongly) l-coloured and (strongly) l-colourable structures and Assumption 2.12 it follows that if $\mathcal{M}$ is a (strongly) l-coloured, or (strongly) l-colourable, structure, and $\mathcal{A}$ is a closed substructure of $\mathcal{M}$, then $\mathrm{cl}_{\mathcal{A}}(X) = \mathrm{cl}_{\mathcal{M}}(X)$ for every $X \subseteq A$. For this reason we will usually omit the subscripts '$\mathcal{A}$' and '$\mathcal{M}$' and just write 'cl'. Also note that from Assumption 2.12 it follows that there is a unique (strongly) l-coloured/colourable structure of dimension 0.

Definition 2.16. Suppose that $\mathcal{M}$ is an L-structure. Let $d \in \mathbb{N}$. The ***d-dimensional reduct*** of $\mathcal{M}$, denoted $\mathcal{M}{\restriction}d$, is the unique L-structure satisfying the following three conditions:

(1) $\mathcal{M}{\restriction}d$ has the same universe as $\mathcal{M}$.

(2) Every symbol in V_{pre} has the same interpretation in $\mathcal{M}{\restriction}d$ as in $\mathcal{M}$.

(3) For each relation symbol $R \in V_{col} \cup V_{rel}$ and tuple $\bar{a} \in M$ of the corresponding arity,

$$\bar{a} \in R^{\mathcal{M}\restriction d} \Leftrightarrow \dim_{\mathcal{M}}(\bar{a}) \leq d \text{ and } \bar{a} \in R^{\mathcal{M}}.$$

Let $\mathbf{K}_n{\restriction}d = \{\mathcal{M}{\restriction}d : \mathcal{M} \in \mathbf{K}_n\}$ and $\mathbf{SK}_n{\restriction}d = \{\mathcal{M}{\restriction}d : \mathcal{M} \in \mathbf{SK}_n\}$.

Notice that if $\mathcal{M}$ is a (strongly) l-colourable structure and d is an integer such that no relation symbol in V_{rel} has higher arity than d, then $\mathcal{M}{\restriction}d = \mathcal{M}$. We also have $\mathbf{K}_n{\restriction}0 = \{\mathcal{G}_n\} = \mathbf{SK}_n{\restriction}0$ for every n. By the ***uniform probability measure*** on a finite set X we mean the probability measure which gives every member of X the same probability $1/|X|$. Recall from Assumption 2.12 that ρ is the highest arity that occurs among the relation symbols of V_{rel}, so $\rho \geq 2$.

Definition 2.17. (i) For every $n \in \mathbb{N}$ and every integer $0 \leq r \leq \rho$ we define a probability measure $\mathbb{P}_{n,r}$ on $\mathbf{K}_n{\restriction}r$ by induction on r as follows. $\mathbb{P}_{n,0}$ is the uniform probability measure on $\mathbf{K}_n{\restriction}0$. For each $1 \leq r \leq \rho$ and $\mathcal{M} \in \mathbf{K}_n{\restriction}r$ we define

$$\mathbb{P}_{n,r}(\mathcal{M}) = \frac{1}{|\{\mathcal{M}' \in \mathbf{K}_n{\restriction}r : \mathcal{M}' \restriction r-1 = \mathcal{M}{\restriction}r-1\}|} \cdot \mathbb{P}_{n,r-1}(\mathcal{M}{\restriction}r-1).$$

(ii) We then define $\delta_n^{\mathbf{K}} = \mathbb{P}_{n,\rho}$ which we call the ***dimension conditional measure on*** $\mathbf{K}_n = \mathbf{K}_n{\restriction}\rho$.
(iii) The ***dimension conditional measure on*** $\mathbf{SK}_n$, denoted $\delta_n^{\mathbf{SK}}$, is defined in the same way, by replacing $\mathbf{K}_n$ with $\mathbf{SK}_n$ in part (i) and then letting $\delta_n^{\mathbf{SK}} = \mathbb{P}_{n,\rho}$.

Example 2.18. Let $L_{pre} = L_F$ as in Example 2.6 and let $F = \mathbb{Z}_2$. Suppose that $l = 2$, so $V_{col} = \{P_1, P_2\}$, and suppose that $V_{rel} = \{R\}$ where R is binary. Let $\mathcal{G}_2 = \mathbb{Z}_2 \times \mathbb{Z}_2$, that is, $\mathcal{G}_2$ is a 2-dimensional vector space over $\mathbb{Z}_2$. From the assumptions that have been made it follows that $\mathbf{K}_2$ is the set of all 2-coloured structures $\mathcal{M}$ such that $\mathcal{M}{\restriction}L_{pre} = \mathcal{G}_2 = \mathbb{Z}_2 \times \mathbb{Z}_2$. We have $|\mathbf{K}_2| = 26$, so if $\mathcal{M} \in \mathbf{K}_2$ is the structure in which all non-zero vectors have colour P_1 and consequently $R^{\mathcal{M}} = \emptyset$, then with the uniform probability measure the probability of $\mathcal{M}$ is $1/26$.

If we want to calculate $\delta_2^{\mathbf{K}}(\mathcal{M})$, where $\mathcal{M}$ is still the same structure, we first need to calculate $\mathbb{P}_{2,0}(\mathcal{M}{\restriction}0)$ which equals 1, because $\mathbb{P}_{2,0}$ is the uniform probability on $\mathbf{K}_2{\restriction}0$ which contains exactly one structure, namely $\mathcal{G}_2 = \mathcal{M}{\restriction}0$. When we consider $\mathbb{P}_{2,1}(\mathcal{M}{\restriction}1)$ we look at structures in $\mathbf{K}_2{\restriction}1$, that is, $\mathcal{G}_2$ with colours added. Since $|\mathbf{K}_2{\restriction}1| = 8$ and the 0-dimensional reduct of every member of $\mathbf{K}_2{\restriction}1$ is $\mathcal{G}_2$ it follows that

$$\mathbb{P}_{2,1}(\mathcal{M}{\restriction}1) \ = \ \frac{1}{|\{\mathcal{M}' \in \mathbf{K}_2 \restriction 1 : \mathcal{M}' \restriction 0 = \mathcal{M} \restriction 0\}|} \cdot \mathbb{P}_{2,0}(\mathcal{M}{\restriction}0) = \frac{1}{8} \cdot 1 = \frac{1}{8}.$$

The last step, to calculate $\delta_2^{\mathbf{K}}(\mathcal{M}) = \mathbb{P}_{2,2}(\mathcal{M})$ is easy, since the only structure in $\mathbf{K}_2{\restriction}2 = \mathbf{K}_2$ which has the same colouring as $\mathcal{M}$ is $\mathcal{M}$ itself. Hence

$$\delta_2^{\mathbf{K}}(\mathcal{M}) \ = \ \mathbb{P}_{2,2}(\mathcal{M}) \ = \ \frac{1}{|\{\mathcal{M}' \in \mathbf{K}_2 \restriction 2 : \mathcal{M}' \restriction 1 = \mathcal{M} \restriction 1\}|} \cdot \mathbb{P}_{2,1}(\mathcal{M}{\restriction}1) = \frac{1}{1} \cdot \frac{1}{8} = \frac{1}{8}.$$

Remark 2.19. We defined $\delta_n^{\mathbf{K}}$ and $\delta_n^{\mathbf{SK}}$ as we did in Definition 2.17 because we are going to use results from [8]. But in the present (more specialised) context, $\delta_n^{\mathbf{K}}$ can be more simply characterised as follows. For every $\mathcal{M} \in \mathbf{K}_n$ we have

$$\delta_n^{\mathbf{K}}(\mathcal{M}) \;=\; \frac{1}{\big|\mathbf{K}_n{\restriction}1\big| \cdot \big|\{\mathcal{M}' \in \mathbf{K}_n : \mathcal{M}'{\restriction}1 = \mathcal{M}{\restriction}1\}\big|},$$

and similarly for $\delta_n^{\mathbf{SK}}$. This is not difficult to prove, by the use of the definitions of l-coloured, and strongly l-coloured, structures. Note that any given colouring of an l-coloured structure $\mathcal{M} \in \mathbf{K}_n$ has probability $1/|\mathbf{K}_n{\restriction}1|$ with this measure.

Definition 2.20. Let $\mathcal{M}$ be an (strongly) l-coloured structure.
(i) Suppose that $\mathcal{B}$ is an (strongly) l-coloured structure and that $\mathcal{A}$ is a closed substructure of $\mathcal{B}$, so $\mathcal{A}$ is also (strongly) l-coloured. We say that $\mathcal{M}$ has the ***$\mathcal{B}/\mathcal{A}$-extension property*** if whenever $\mathcal{A}'$ is a closed substructure of $\mathcal{M}$ and $\sigma_{\mathcal{A}} : \mathcal{A}' \to \mathcal{A}$ is an isomorphism, then there are a closed substructure $\mathcal{B}'$ of $\mathcal{M}$ such that $\mathcal{A}' \subset \mathcal{B}'$ and an isomorphism $\sigma_{\mathcal{B}} : \mathcal{B}' \to \mathcal{B}$ which extends $\sigma_{\mathcal{A}}$.
(ii) Let $k \in \mathbb{N}$. We say $\mathcal{M}$ has the ***k-extension property*** if it has the $\mathcal{B}/\mathcal{A}$-extension property whenever $\mathcal{B}$ is an (strongly) l-coloured structure, $\mathcal{A}$ is a closed substructure of $\mathcal{B}$ and $\dim_{\mathcal{M}}(B) \leq k$.

When saying that two l-coloured structures $\mathcal{A}$ and $\mathcal{A}'$ ***agree on L_{pre} and on closed proper substructures*** we mean that $\mathcal{A}{\restriction}L_{pre} = \mathcal{A}'{\restriction}L_{pre}$ (so in particular, $\mathrm{cl}_{\mathcal{A}} = \mathrm{cl}_{\mathcal{A}'}$) and whenever $\mathcal{U}$ is a closed substructure of $\mathcal{A}$ and $\dim_{\mathcal{A}}(U) < \dim_{\mathcal{A}}(A)$, then $\mathcal{A}{\restriction}U = \mathcal{A}'{\restriction}U$.

Lemma 2.21. *Whenever $\mathcal{M}$ is (strongly) l-coloured, $\mathcal{A}$ is a closed substructure of $\mathcal{M}$ and $\mathcal{A}'$ is an (strongly) l-coloured structure which agrees with $\mathcal{A}$ on L_{pre} and on closed proper substructures, then there is an (strongly) l-coloured structure $\mathcal{N}$ such that $\mathcal{N}{\restriction}L_{pre} = \mathcal{M}{\restriction}L_{pre}$, $\mathcal{N}{\restriction}A = \mathcal{A}'$ and if $\mathcal{U}$ is a closed substructure of $\mathcal{N}$, $\dim_{\mathcal{N}}(U) \leq \dim_{\mathcal{N}}(A')$ and $\mathcal{U} \neq \mathcal{A}'$, then $\mathcal{N}{\restriction}U = \mathcal{M}{\restriction}U$.*

Proof. We only prove the lemma in the case of l-coloured structures. The proof for *strongly* l-coloured structures is a straightforward modification. Suppose that $\mathcal{M}$ is l-coloured, that $\mathcal{A}$ is a closed substructure of $\mathcal{M}$, and therefore l-coloured. Also assume that $\mathcal{A}'$ is l-coloured and agrees with $\mathcal{A}$ on L_{pre} and on closed proper substructures. Observe that by these assumptions and Assumption 2.12, for every $X \subseteq A$ we have $\mathrm{cl}_{\mathcal{M}}(X) = \mathrm{cl}_{\mathcal{A}}(X) = \mathrm{cl}_{\mathcal{A}'}(X)$ and $\dim_{\mathcal{M}}(X) = \dim_{\mathcal{A}}(X) = \dim_{\mathcal{A}'}(X)$, so we can omit the subscripts. The proof splits into three cases.

First suppose that $\dim(A) = 0$. By parts (1) and (2) of the definition of l-coloured structure we have $\mathcal{A} = \mathcal{A}'$ so if $\mathcal{N} = \mathcal{M}$ then the conclusion of the lemma is satisfied.

Now suppose that $\dim(A) = 1$, so $\mathcal{A}$ is a one dimensional structure and therefore all $a \in A - \mathrm{cl}(\emptyset)$ have the same colour in $\mathcal{A}$, say i (that is, $\mathcal{A} \models P_i(a)$). Similarly, $\mathcal{A}'$ is a one dimensional structure so all $a \in A' - \mathrm{cl}(\emptyset)$ have the same colour in $\mathcal{A}'$, say j. Let $\mathcal{N}$ be the structure which satisfies the following conditions:

- $\mathcal{N}{\restriction}L_{pre} = \mathcal{M}{\restriction}L_{pre}$, so in particular $N = M$.
- For every $R \in V_{rel}$, $R^{\mathcal{N}} = \emptyset$.
- For every $a \in M - A$ and every $m \in \{1, \ldots, l\}$, $\mathcal{N} \models P_m(a) \iff \mathcal{M} \models P_m(a)$.
- For every $a \in A - \mathrm{cl}(\emptyset)$, $\mathcal{N} \models P_j(a)$.

Then $\mathcal{N}$ is l-coloured, for trivial reasons, and has the required properties which is easily checked.

Finally suppose that $\dim(A) = k + 1$ where $k \geq 1$. Define $\mathcal{N}$ as follows:

- $\mathcal{N}\restriction k = \mathcal{M}\restriction k$, so in particular $\mathcal{N}\restriction L_{pre} = \mathcal{M}\restriction L_{pre}$.
- Whenever U is a closed subset of $M = N$, $\dim(U) = k + 1$ and $U \neq A$, then $\mathcal{N}\restriction U = \mathcal{M}\restriction U$.
- $\mathcal{N}\restriction A = \mathcal{A}'$.
- Whenever $\bar{a} \in M$, $\dim(\bar{a}) > k + 1$ and $R \in V_{rel}$, then $\bar{a} \notin R^{\mathcal{N}}$.

It remains to prove that $\mathcal{N}$ is l-coloured. Since $\mathcal{N}\restriction k = \mathcal{M}\restriction k$, where $k \geq 1$, it follows that $\mathcal{N}\restriction L_{col} = \mathcal{M}\restriction L_{col}$ and hence conditions (1)–(3) in the definition of l-coloured structure are satisfied.

Now we consider condition (4). Suppose that $\mathcal{N} \models R(\bar{a})$ where $R \in V_{rel}$. We need to show that there are $b, c \in \mathrm{cl}(\bar{a}) - \mathrm{cl}(\emptyset)$ such that b and c have different colours. By the last part of the definition of $\mathcal{N}$ we may assume that $\dim(\bar{a}) \leq k+1$. If $\dim(\bar{a}) \leq k$ then, by the first part of the definition of $\mathcal{N}$, and the assumption that $\mathcal{M}$ is l-coloured it follows that $b, c \in \mathrm{cl}(\bar{a}) - \mathrm{cl}(\emptyset)$ with different colours exist. Now suppose that $\dim(\bar{a}) = k+1$. If $\mathrm{cl}(\bar{a}) \neq A$, then, by the second part of the definition of $\mathcal{N}$ and the assumption that $\mathcal{M}$ is l-coloured, there are $b, c \in \mathrm{cl}(\bar{a}) - \mathrm{cl}(\emptyset)$ with different colours. Finally suppose that $\mathrm{cl}(\bar{a}) = A$. By the third part of the definition of $\mathcal{N}$, $\mathcal{N}\restriction A = \mathcal{A}'$, so $\mathcal{A}' \models R(\bar{a})$ and since $\mathcal{A}'$ is l-colored there are $b, c \in \mathrm{cl}(\bar{a}) - \mathrm{cl}(\emptyset)$ with different colours in $\mathcal{A}'$, and hence (by the third part of the definition of $\mathcal{N}$ again) they have different colours in $\mathcal{N}$. □

In the terminology of [8] (Definition 7.20), Lemma 2.21 says that, for every $k \in \mathbb{N}$, $\mathbf{K}$ and $\mathbf{SK}$ *accept k-substitutions over L_{pre}*. Therefore, Assumption 2.12 and Theorems 7.31 and 7.32 in [8] imply the following:

Theorem 2.22. *Suppose that, for every $k \in \mathbb{N}$, $\mathbf{G}$ is polynomially k-saturated.*
(i) For every $k \in \mathbb{N}$,

$$\begin{aligned} \lim_{n\to\infty} \delta_n^{\mathbf{K}}\big(\{\mathcal{M} \in \mathbf{K}_n : \mathcal{M} \textit{ has the } k\textit{-extension property}\}\big) &= 1 \quad \textit{and} \\ \lim_{n\to\infty} \delta_n^{\mathbf{SK}}\big(\{\mathcal{M} \in \mathbf{SK}_n : \mathcal{M} \textit{ has the } k\textit{-extension property}\}\big) &= 1. \end{aligned}$$

(ii) For every L-sentence φ, $\delta_n^{\mathbf{K}}\big(\{\mathcal{M} \in \mathbf{K}_n : \mathcal{M} \models \varphi\}\big)$ approaches either 0 or 1, and $\delta_n^{\mathbf{SK}}\big(\{\mathcal{M} \in \mathbf{SK}_n : \mathcal{M} \models \varphi\}\big)$ approaches either 0 or 1, as n tends to infinity.

Now we have a zero-one law for (strongly) l-colour*ed* structures, with the dimension conditional probability measure. Next, we look att (strongly) l-colour*able* structures, with a probability measure that is derived from the dimension conditional measure

Definition 2.23. For each n and all $X \subseteq \mathbf{C}_n$ and $Y \subseteq \mathbf{S}_n$ let

$$\begin{aligned} \delta_n^{\mathbf{C}}(X) &= \delta_n^{\mathbf{K}}\big(\{\mathcal{M} \in \mathbf{K}_n : \mathcal{M}\restriction L_{rel} \in X\}\big), \text{ and} \\ \delta_n^{\mathbf{S}}(X) &= \delta_n^{\mathbf{SK}}\big(\{\mathcal{M} \in \mathbf{SK}_n : \mathcal{M}\restriction L_{rel} \in X\}\big). \end{aligned}$$

Intuitively, for $X \subseteq \mathbf{C}_n$, we can think of $\delta_n^{\mathbf{C}}(X)$ as the probability that $\mathcal{M} \in \mathbf{C}_n$ will belong to X if $\mathcal{M}$ is generated by the following procedure: start with $\mathcal{G}_n$ and randomly add colours to the 1-dimensional subspaces of $\mathcal{G}_n$, then add R-relations for each $R \in V_{rel}$

in such a way that the colouring conditions (1)–(4) of Definition 2.14 are respected but apart from this in a random fashion, and finally, forget about the specific colouring, that is, consider the reduct to L_{rel}. The probability measure $\delta_n^{\mathbf{S}}$ can be interpreted analogously. The corollary below states tells that a zero-one law holds for (strongly) l-colourable structures when probability measure $\delta_n^{\mathbf{C}}$ ($\delta_n^{\mathbf{S}}$) is used, in other words, it states the same thing as Theorem 1.1.

Corollary 2.24. *Suppose that, for every $k \in \mathbb{N}$, $\mathbf{G}$ is polynomially k-saturated. For every L_{rel}-sentence φ, $\delta_n^{\mathbf{C}}\big(\{\mathcal{M} \in \mathbf{C}_n : \mathcal{M} \models \varphi\}\big)$ approaches either 0 or 1, and $\delta_n^{\mathbf{S}}\big(\{\mathcal{M} \in \mathbf{S}_n : \mathcal{M} \models \varphi\}\big)$ approaches either 0 or 1, as n tends to infinity.*

Proof. Let φ be an L_{rel}-sentence, so in particular it is an L-sentence. Then

$$\begin{aligned}
&\delta_n^{\mathbf{C}}\big(\{\mathcal{M} \in \mathbf{C}_n : \mathcal{M} \models \varphi\}\big) \\
&= \delta_n^{\mathbf{K}}\big(\{\mathcal{M} \in \mathbf{K}_n : \mathcal{M}{\restriction}L_{rel} \models \varphi\}\big) \quad \text{(by the definition of } \delta_n^{\mathbf{C}}) \\
&= \delta_n^{\mathbf{K}}\big(\{\mathcal{M} \in \mathbf{K}_n : \mathcal{M} \models \varphi\}\big) \quad (\text{since } \mathcal{M} \models \varphi \ \Leftrightarrow \mathcal{M}{\restriction}L_{rel} \models \varphi).
\end{aligned}$$

Since φ is also an L-sentence, Theorem 2.22 implies that $\delta_n^{\mathbf{C}}\big(\{\mathcal{M} \in \mathbf{C}_n : \mathcal{M} \models \varphi\}\big)$ approaches either 0 or 1 as $n \to \infty$. The proof that $\delta_n^{\mathbf{S}}\big(\{\mathcal{M} \in \mathbf{S}_n : \mathcal{M} \models \varphi\}\big)$ approaches either 0 or 1 as $n \to \infty$ is exactly the same; just replace $\mathbf{C}_n$ by $\mathbf{S}_n$ and $\mathbf{K}_n$ by $\mathbf{SK}_n$. □

However, neither the theorem nor its proof gives information about for which L_{rel}-sentences φ we have $\lim_{n\to\infty} \delta_n^{\mathbf{C}}\big(\{\mathcal{M} \in \mathbf{C}_n : \mathcal{M} \models \varphi\}\big) = 1$, nor do we get information about structural properties of (strongly) l-colourable structures. The remaining sections deal with these issues. In hindsight it seems silly that the second author of this article did not notice, in [8], this easy way of proving the zero-one law of (strongly) l-colourable structures with trivial pregeometry, when the measures $\delta_n^{\mathbf{C}}$ (or $\delta_n^{\mathbf{S}}$) are used. But in [8] emphasis was put on extension axioms, which may explain why the above "short cut" to Corollary 2.24 in the case when the underlying pregeomeries are trivial was not noticed.

It will sometimes be convenient to think of l-colourings as functions that assign colours to elements, as done in combinatorics, so we introduce the following terminology.

Definition 2.25. Let $\mathcal{A}$ be an L_{rel}-structure and let $\gamma : A - \mathrm{cl}(\emptyset) \to \{1, ..., l\}$. Let B be a closed subset of A. We say that B is γ-***monochromatic*** if for all $a, b \in B - \mathrm{cl}(\emptyset)$, $\gamma(a) = \gamma(b)$. If B is not γ-monochromatic then it is called γ-***multichromatic***. If $\gamma(a) \neq \gamma(b)$ whenever $a \in B$ and $b \in B$ are independent, then we call B ***strongly*** γ-***multichromatic***. If there is no risk of confusion we may just say monochromatic, multichromatic or strongly multichromatic. We say that γ is a ***(strong)*** l-***colouring*** of $\mathcal{A}$ if the following conditions hold:

1. For every $a \in A - \mathrm{cl}(\emptyset)$, $\mathrm{cl}(a)$ is γ-monochromatic.
2. If $R \in V_{rel}$ and $\mathcal{A} \models R(\bar{a})$ then $\mathrm{cl}(\bar{a})$ is (strongly) γ-multichromatic.

Observe that an L_{rel}-structure $\mathcal{A}$ is (strongly) l-colourable, according to Definition 2.14, if and only if there is an (strong) l-colouring $\gamma : A - \mathrm{cl}(\emptyset) \to \{1, \ldots, l\}$ of $\mathcal{A}$. We will often want to describe the isomorphism type of some particular structure with a sentence, which motivates the following definition.

Definition 2.26. Let $\mathcal{A}$ be an (strongly) l-colourable structure and let $A = \{a_1, \ldots, a_m\}$ By a ***characteristic formula of*** $\mathcal{A}$, with respect to the given enumeration of A, we mean a quantifier-free L_{rel}-formula $\chi_{\mathcal{A}}(x_1, \ldots, x_m)$ such that if $\mathcal{M}$ is an L_{rel}-structure such that the formulas θ_n define a pregeometry $(M, \mathrm{cl}_{\mathcal{M}})$ and $\mathcal{M} \models \chi_{\mathcal{A}}(b_1, \ldots, b_m)$, then the map $a_i \mapsto b_i$, for $i = 1, \ldots, m$, is an embedding of $\mathcal{A}$ into $\mathcal{M}$. Similarly we define a characteristic formula of an (strongly) l-coloured structure. Note that such formulas exist because of the definition of (strongly) l-colourable (or l-coloured) structures and Assumption 2.12 (7) (see also Remark 2.13 (iii)).

3 Definability of strong l-colourings

In this section we study strongly l-coloured structures, where $l \geq 2$ (as always). If a and b are elements of a strongly l-coloured structure and $\mathcal{M} \models P_i(a) \wedge P_i(b)$ for some $i \in \{1, \ldots, l\}$, then we say that a and b have the same colour. The main result of this section, which is essential for the proof of Theorem 1.3, which is finished in Section 5, is the following: there are $k_0 \in \mathbb{N}$ and an L_{rel}-sentence $\xi(x, y)$ such that

- if $\mathcal{M}$ is strongly l-coloured, $a, b \in M - \mathrm{cl}(\emptyset)$ and $\mathcal{M} \models \xi(a, b)$, then a and b have the same colour, and
- if $\mathcal{M}$ is strongly l-coloured and has the k_0-extension property and $a, b \in M - \mathrm{cl}(\emptyset)$, then $\mathcal{M} \models \xi(a, b)$ if and only if a and b have the same colour.

The definition of strongly l-colourable structures implies that if $\mathcal{M}$ is strongly l-colourable, $R \in V_{rel}$ is an r-ary relation symbol (so $r \geq 2$), $\mathcal{M} \models R(a_1, \ldots, a_r)$ and $b, c \in \mathrm{cl}(a_1, \ldots, a_r) - \mathrm{cl}(\emptyset)$ are independent, then a and b must have different colours. It follows that if $a_1, \ldots, a_r \in M$ and the number of 1-dimensional subspaces (i.e. closed subsets) of $\mathrm{cl}(a_1, \ldots, a_r)$ is larger than l, then $\mathcal{M} \not\models R(a_1, \ldots, a_r)$.

Example 3.1. Suppose that $F = \mathbb{Z}_2$ is the 2-element field and, for every $n \in \mathbb{N}$, $\mathcal{G}_n$ is an n-dimensional vector space over F, as in Example 2.6. Let $l = 2$. For every 2-dimensional subspace V of $\mathcal{G}_n$ ($n \geq 2$), the number of 1-dimensional subspaces of V is $2^2 - 1 = 3 > l$. So, with these assumptions, if $\mathcal{M}$ is strongly 2-coloured then $R^{\mathcal{M}} = \emptyset$ for every $R \in V_{rel}$. But if, instead, $l > 2$ then it is possible that $R^{\mathcal{M}} \neq \emptyset$ for every $R \in V_{rel}$.

Since strongly l-coloured structures in which R is interpreted as the empty set for every $R \in V_{rel}$ are not so interesting, the above example motivates the following definition and assumption. Observe that by Assumption 2.12 (6), if $n \in \mathbb{N}$ and $\mathcal{G}'$ is a closed substructure of $\mathcal{G}_n$, then $\mathcal{G}' \cong \mathcal{G}_m$ for some $m \in \mathbb{N}$.

Definition 3.2. (i) If A is a closed subset of G_n, for some n, then let $D(A)$ be the number of 1-dimensional subspaces of A.
(ii) For every $d \in \mathbb{N}$, let $t(d)$ be the maximum of $D(A)$ where A is a subspace of $\mathcal{G}_n$ for some n and $\dim_{\mathcal{G}_n}(A) \leq d$.
(iii) Let $t = \max\{d \in \mathbb{N} : t(d) \leq l\}$.

Note that if $\dim_{\mathcal{G}_n}(A) > l$ then $D(A) > l$, so $t \leq l$. In Example 3.1 we have $t(0) = 0, t(1) = 1, t(2) = 3$ and $t(3) = 8$, so if $l = 2$ then $t = 1$. If, in the same example, $l \in \{3, \ldots, 7\}$, then $t = 2$; if $l = 8$, then $t = 3$, and so on. In order that the arguments that follow work out *we assume that* $t \geq 2$. This is equivalent with the condition, in Theorem 1.3, that for every $n \in \mathbb{N}$, every 2-dimensional subspace of $\mathcal{G}_n$ has at most l different 1-dimensional subspaces.

Let the relation symbols of V_{rel} be $R_1, ..., R_\tau$ with arities $r_1, ..., r_\tau \geq 2$. Without loss of generality we assume that r_1 is the smallest among these arities. By Assumption 2.12

there are L_{pre}-formulas θ_0 and θ_1 such that if $\mathcal{M}$ is strongly l-coloured (or strongly l-colourable), then $\mathcal{M} \models \theta_0(a) \iff a \in \mathrm{cl}(\emptyset)$, and $\mathcal{M} \models \theta_1(a,b) \iff b \in \mathrm{cl}(a)$. Since $L_{pre} \subseteq L_{rel}$, this justifies the use of notation like '$x \in \mathrm{cl}(y)$' when specifying L_{rel}-formulas.

The idea of the formula $\xi(x,y)$ defined below is that whenever a and b do not belong to the closure of $\emptyset$ and $\xi(a,b)$ holds, then a and b must have the same colour (and the converse implication holds if the structure that a and b come from has the k-extension property for large enough k). This is achieved by saying that if a and b are independent then there are $c_2, \dots c_l$ such that every pair of distinct elements from $\{a, c_2, \dots, c_l\}$ is independent and belongs to an R_1-relationship, thus forcing them to have different colours. The same is said about pairs of distinct elements from $\{b, c_2, \dots, c_l\}$, thus forcing the elements of every such pair to have different colours. As $c_2, \dots, c_l$ use up $l-1$ colours and there are only l colours, this forces a and b to have the same colour. In the following definition we will use notation like $\overset{l}{\underset{i=1}{\exists}}\overset{r_1}{\underset{j=1}{\exists}} x_{i,j}$ which is the same as saying $\exists x_{1,1} \dots x_{1,r} \exists x_{2,1} \dots x_{2,r} \exists x_{l,1} \dots x_{l,r_1}$. Moreover, we use triples $(\bullet, \bullet, \bullet)$ to index different variables $z_{(\bullet,\bullet,\bullet)}$.

Definition 3.3. Let $\xi(x,y)$ denote the following L_{rel}-formula:

$$x \in \mathrm{cl}(y) \ \vee \ y \in \mathrm{cl}(x) \ \vee \ \exists y_2, \dots, y_l \overset{l}{\underset{i=2}{\exists}} \overset{r_1-2}{\underset{j=1}{\exists}} z_{(x,i,j)} \overset{l}{\underset{i=2}{\exists}} \overset{r_1-2}{\underset{j=1}{\exists}} z_{(y,i,j)} \overset{l}{\underset{k=2}{\exists}} \overset{k-1}{\underset{i=2}{\exists}} \overset{r_1-2}{\underset{j=1}{\exists}} z_{(k,i,j)}$$
$$\Bigg[\bigwedge_{i=2}^{l} \Big(R_1(x, y_i, z_{(x,i,1)}, \dots, z_{(x,i,r_1-2)}) \ \wedge \ y_i \notin \mathrm{cl}(x) \ \wedge \ R_1(y, y_i, z_{(y,i,1)}, \dots, z_{(y,i,r_1-2)})$$
$$\wedge \ y_i \notin \mathrm{cl}(y) \Big) \ \wedge \ \bigwedge_{j=2}^{i-1} \Big(R_1(y_i, y_j, z_{(i,j,1)}, \dots, z_{(i,j,r_1-2)}) \ \wedge \ y_i \notin \mathrm{cl}(y_j) \Big) \Bigg].$$

The variables $z_{(k,i,j)}, z_{(x,i,j)}$ and $z_{(y,i,j)}$ have the function of "fillers" to get the the right length, r_1, of the tuples. In the case $r_1 = 2$ they are not needed and ξ will look like this:

$$x \in \mathrm{cl}(y) \ \vee \ y \in \mathrm{cl}(x) \ \vee$$
$$\exists y_2, \dots, y_l \Bigg[\bigwedge_{i=2}^{l} \Big(R_1(x, y_i) \ \wedge \ y_i \notin \mathrm{cl}(x) \ \wedge \ R_1(y, y_i) \ \wedge y_i \notin \mathrm{cl}(y) \Big) \ \wedge$$
$$\bigwedge_{j=2}^{i-1} \Big(R_1(y_i, y_j) \ \wedge \ y_i \notin \mathrm{cl}(y_j) \Big) \Bigg].$$

Lemma 3.4. *If $\mathcal{M}$ is strongly l-coloured, $a, b \in M - \mathrm{cl}(\emptyset)$ and $\mathcal{M} \models \xi(a,b)$ then a and b have the same colour in $\mathcal{M}$, i.e. for some $i = 1, \dots, l$ we have $\mathcal{M} \models P_i(a) \wedge P_i(b)$.*

Proof. Let $\mathcal{M}$ be strongly l-coloured. We assume that $\mathcal{M} \models \xi(a,b)$ and $a, b \notin \mathrm{cl}(\emptyset)$. If $a \in \mathrm{cl}(b)$ then we obviously are done by the definition of a colouring, hence assume that a and b are independent. Each y_i that witness the truth of $\xi(a,b)$ must have a different colour from a since they are independent and included in a tuple $(a, y_i, z_{(a,i,1)}, \dots, z_{(a,i,r_1-2)}) \in R_1^{\mathcal{M}}$. In the same way each y_i must have different colour from b. In the same way as for a and b, looking at the definition of ξ, we get that y_i and y_j must have different colour in $\mathcal{M}$ if $i \neq j$. Hence we can conclude that all the elements $a, y_2, \dots, y_l$ have different colours and all the elements $b, y_2, \dots, y_l$ have different colours. But since $\mathcal{M}$ is coloured by only l different colours this implies, by the pigeon hole principle, that a and b must have the same colour. □

For the rest of this section, let

$$k_0 = t(l+1)l.$$

We will now prove that if $\mathcal{M}$ is strongly l-colourable with the k_0-extension property and $a, b \in M - \mathrm{cl}(\emptyset)$ have the same colour in $\mathcal{M}$, then $\mathcal{M} \models \xi(a,b)$. This will be done by defining a structure $\mathcal{B}$ which has the same relations as described by ξ, and showing that this structure is strongly l-colourable. Then we show that if a and b have the same colour in a structure with the k_0-extension property, then they are included in a copy of $\mathcal{B}$ in such a way that, by construction of $\mathcal{B}$, $\xi(a,b)$ holds.

Lemma 3.5. *Let $\mathcal{M}$ be strongly l-coloured and assume that $\mathcal{M}$ has the k_0-extension property. If $a, b \in M - \mathrm{cl}(\emptyset)$ and $\mathcal{M} \models P_i(a) \wedge P_i(b)$ for some $i \in \{1, ..., l\}$, then $\mathcal{M} \models \xi(a,b)$.*

Proof. If we would not be interested in being able to easily adapt the follwing argument to the context where R_1 is always interpreted as an irreflexive and symmetric relation, then some parts of the argument could be simplified (see Remark 3.7). Without loss of generality we may assume that $\mathcal{M} \models P_1(a) \wedge P_1(b)$ where $a, b \in M - \mathrm{cl}(\emptyset)$. If $a \in \mathrm{cl}(b)$ then $\mathcal{M} \models \xi(a,b)$ by definition, hence assume that $a \notin \mathrm{cl}(b)$. Let $\mathcal{A} = \mathcal{M} \restriction \mathrm{cl}(a,b)$ and choose elements $v_2, ..., v_l \in M$ and elements $u_{(a,i,j)}, u_{(b,i,j)}, u_{(k,i,j)} \in M$ for each $2 \leq i \leq l, 1 \leq j \leq t-2$ and $2 \leq k \leq l-1, k \neq i$ such that the set S containing exactly the elements a, b, $v_2, \ldots, v_l$ and $u_{(a,i,j)}, u_{(b,i,j)}, u_{(k,i,j)}$, for i, j, k as indicated above, is an independent set. Such a choice of elements from M is possible because we assume that $\mathcal{M}$ has the k_0-extension property where $k_0 = t(l+1)l$.[1] Let $\mathcal{B}_0$ be the substructure of $\mathcal{M} \restriction L_{pre}$ with universe $\mathrm{cl}(S)$, or equivalently, $\mathcal{B}_0 = \big(\mathcal{M} \restriction \mathrm{cl}(S)\big) \restriction L_{pre}$. Note that $\mathcal{A} \restriction L_{pre} \subseteq \mathcal{B}_0$. Define $\mathcal{B}$ to be the L-structure which is created by expanding $\mathcal{B}_0$ to an L-structure in the following way. We know already that $\mathcal{A} \restriction L_{pre} \subseteq \mathcal{B} \restriction L_{pre}$, so for each $i \in \{1, ..., \rho\}$, every $R_i \in V - V_{pre}$, and every $\bar{a} \in A^{r_i}$, we let $\bar{a} \in R_i^{\mathcal{B}} \Leftrightarrow \bar{a} \in R_i^{\mathcal{A}}$, and for each $j \in \{1, ..., l\}$ and $a \in A$ we let $a \in P_j^{\mathcal{B}} \Leftrightarrow a \in P_j^{\mathcal{A}}$. In this way we obviously get that $\mathcal{A} \subseteq \mathcal{B}$ as L-structures, no matter how we define, in $\mathcal{B}$, interpretations on tuples whose range are not included in A. For every relation symbol $R_i \in V_{rel} - \{R_1\}$ and $\bar{c} \in B^{r_i} - A^{r_i}$ let $\mathcal{B} \not\models R(\bar{c})$. For each $i \in \{2, ..., l\}$ and $i < j \leq l$ fix arbitrary elements

$$\begin{aligned} w_{(a,i,1)}, ..., w_{(a,i,r_1-2)} &\in \mathrm{cl}\big(a, v_i, u_{(a,i,1)}, ..., u_{(a,i,t-2)}\big), \\ w_{(b,i,1)}, ..., w_{(b,i,r_1-2)} &\in \mathrm{cl}\big(b, v_i, u_{(b,i,1)}, ..., u_{(b,i,t-2)}\big) \quad \text{and} \\ w_{(j,i,1)}, ..., w_{(j,i,r_1-2)} &\in \mathrm{cl}\big(v_j, v_i, u_{(j,i,1)}, ..., u_{(j,i,t-2)}\big). \end{aligned}$$

Define $R_1^{\mathcal{B}}$ on $B^{r_1} - A^{r_1}$ in such a way that, for each $2 \leq i \leq l$,

$$\mathcal{B} \models R_1\big(a, v_i, w_{(a,i,1)}, ..., w_{(a,i,r_1-2)}\big) \wedge R_1\big(b, v_i, w_{(b,i,1)}, ..., w_{(b,i,r_1-2)}\big)$$

$$\bigwedge_{k=i+1}^{l} R_1\big(v_k, v_i, w_{(k,i,1)}, ..., w_{(k,i,r_1-2)}\big),$$

and such that $R_1^{\mathcal{B}}$ holds for no other tuples than those indicated in the argument above.

[1]By Assumption 2.12 there is $\mathcal{G}_n$ with dimension k_0 and hence there is a strongly l-coloured structure $\mathcal{B}$ with dimension k_0. By Assumption 2.12 and the definition of strongly l-coloured structures it follows that $\mathcal{B} \restriction \mathrm{cl}_{\mathcal{B}}(\emptyset) \cong \mathcal{M} \restriction \mathrm{cl}_{\mathcal{M}}(\emptyset)$ and since, letting $\mathcal{A} = \mathcal{B} \restriction \mathrm{cl}_{\mathcal{B}}(\emptyset)$, $\mathcal{M}$ has the $\mathcal{B}/\mathcal{A}$-extension property it follows that $\mathcal{M}$ contains an isomorphic copy $\mathcal{B}'$ of $\mathcal{B}$ and B' contains an independent set of cardinality k_0.

In order to complete the definition of $\mathcal{B}$ as an L-structure we need to define the interpretations $P_1^{\mathcal{B}}, \ldots, P_l^{\mathcal{B}}$ on elements in $B - A$. When saying that a 1-dimensional subspace (closed subset) Q gets the colour i we mean that for all $a \in Q - \mathrm{cl}(\emptyset)$, $a \in P_i^{\mathcal{B}}$. Now we define an l-colouring, in $\mathcal{B}$, on $B - A$ according to the following five steps, where we recall that $\mathrm{cl}(a)$ and $\mathrm{cl}(b)$ have colour 1 since, by assumption, $\mathcal{M} \models P_1(a) \wedge P_1(b)$:

(1) For $i = 2, \ldots, l$, $\mathrm{cl}(v_i)$ get the colour i.

(2) By the definition of t and the assumption that $t \geq 2$ it is, for every $i = 2, \ldots, l$, possible to colour all 1-dimensional subspaces of $\mathrm{cl}(a, v_i, u_{(a,i,1)}, ..., u_{(a,i,t-2)})$ which have not yet been assigned colours with the colours $1, \ldots, l$ in such a way that (1) and (2) hold and any two different 1-dimensional subspaces of this space get different colours.

(3) As in (3) it is possible, for every $i = 1, \ldots, l$, to colour all 1-dimensional subspaces of $\mathrm{cl}(b, v_i, u_{(b,i,1)}, ..., u_{(b,i,t-2)})$ which have not yet been assigned colours with the colours $1, \ldots, l$ in such a way that (1) and (2) hold and any two different 1-dimensional subspaces of this space get different colours.

(4) As in (3) and (4) it is possible, for every $i = 1, \ldots, l$, to colour all 1-dimensional subspaces of $\mathrm{cl}(v_j, v_i, u_{(j,i,1)}, ..., u_{(j,i,t-2)})$ with the colours $1, \ldots, l$ in such a way that (1) and (2) hold and any two different 1-dimensional subspaces of this space get different colours.

(5) For every 1-dimensional subspace $Q \subseteq B$ that has not yet been assigned a colour, give Q the colour 1.

Claim. The L-structure $\mathcal{B}$ is a strongly l-coloured structure.

Proof of claim. By the last part of the definition of the colouring of $\mathcal{B}$ and since $\mathcal{A} \subseteq \mathcal{B}$ where $\mathcal{A}$ is a substructure of $\mathcal{M}$, we know that each element has attained at least one colour, so colouring condition (1) of Definition 2.14 is satisfied. The second colouring condition is also satisfied because $\mathcal{A} \subseteq \mathcal{B}$ and $\mathcal{A} \subseteq \mathcal{M}$. If we apply Lemma 2.2 we get the following, for all i, j, k under consideration,

$$\begin{aligned}
\mathrm{cl}(v_k, v_i, u_{(k,i,1)}, ..., u_{(k,i,t-2)}) \cap \mathrm{cl}(b, v_i, u_{(b,i,1)}, ..., u_{(b,i,t-2)}) &= \mathrm{cl}(v_i) \text{ if } k \neq i, \\
\mathrm{cl}(b, v_i, u_{(b,i,1)}, ..., u_{(b,i,t-2)}) \cap \mathrm{cl}(a, v_i, u_{(a,i,1)}, ..., u_{(a,i,t-2)}) &= \mathrm{cl}(v_i), \\
\mathrm{cl}(a, v_i, u_{(a,i,1)}, ..., u_{(a,i,t-2)}) \cap \mathrm{cl}(a, v_j, u_{(a,j,1)}, ..., u_{(a,j,t-2)}) &= \mathrm{cl}(a) \text{ if } i \neq j, \\
\mathrm{cl}(b, v_i, u_{(b,i,1)}, ..., u_{(b,i,t-2)}) \cap \mathrm{cl}(b, v_j, u_{(b,j,1)}, ..., u_{(b,j,t-2)}) &= \mathrm{cl}(b) \text{ if } i \neq j, \\
\mathrm{cl}(a, v_i, u_{(a,i,1)}, ..., u_{(a,i,t-2)}) \cap \mathrm{cl}(v_k, v_i, u_{(k,i,1)}, ..., u_{(k,i,t-2)}) &= \mathrm{cl}(v_i) \text{ if } i \neq k \text{ and} \\
\mathrm{cl}(v_k, v_i, u_{(k,i,1)}, ..., u_{(k,i,t-2)}) \cap \mathrm{cl}(v_j, v_i, u_{(j,i,1)}, ..., u_{(j,i,t-2)}) &= \mathrm{cl}(v_i) \text{ if } k \neq j.
\end{aligned}$$

This shows that the steps (1)–(6) did not give more than one colour to any element of B, and, from the construction it is also clear that dependent elements that do not belong to the closure of $\emptyset$ have obtained the same colour. The colouring restricted to $\mathcal{A} \subseteq \mathcal{B}$ does, since $\mathcal{A} \subseteq \mathcal{M}$ and $\mathcal{M}$ is an l-coloured structure, satisfy all the colouring conditions. Hence the third colouring condition is satisfied for $\mathcal{B}$. If $\mathcal{B} \models R_p(\bar{a})$ for some $R_p \in V_{rel}$, then either $\bar{a} \subset A^{r_p}$ in which case the colouring conditions (4) and (5) are satisfied since

$\mathcal{A} \subseteq \mathcal{M}$ is l-coloured, or $R_p = R_1$ and $\bar{a}$ is identical to one of the following tuples

$$\begin{aligned}&(a, v_i, w_{(a,i,1)}, ..., w_{(a,i,r_1-2)}),\\ &(b, v_i, w_{(b,i,1)}, ..., w_{(b,i,r_1-2)}), \quad \text{or}\\ &(v_k, v_i, w_{(k,i,1)}, ..., w_{(k,i,r_1-2)}),\end{aligned}$$

for some i, k. By the choice of these tuples and the steps (1)–(6) above, it follows that whenever $a, b \in \mathrm{cl}(\bar{a}) - \mathrm{cl}(\emptyset)$ and a is independent from b, then a and b have different colours. Hence colour conditions (4) and (5) are satisfied and we have proved that $\mathcal{B}$ is strongly l-coloured. □

Continuing the proof of Lemma 3.5. By the claim, $\mathcal{B}$ is a strongly l-coloured L-structure and, by the definition of $\mathcal{B}$, $\mathcal{A}$ is a closed substructure of $\mathcal{B}$. Since $B = \mathrm{cl}(S)$ we know that $\dim(B) \leq t(l+1)l = k_0$. As $\mathcal{M}$ has the k_0-extension property and $\mathcal{A}$ is a closed substructure of $\mathcal{M}$, there are a closed substructure $\mathcal{B}' \subseteq \mathcal{M}$ and an isomorphism $f : \mathcal{B}' \to \mathcal{B}$ with which extends the identity function on $\mathcal{A}$, so $\mathcal{A} \subseteq \mathcal{B}'$. From the definition of $\mathcal{B}$ we get that $\mathcal{M} \models \xi(a, b)$. □

Using Lemmas 3.4 and 3.5 we directly get the following:

Corollary 3.6. *If $\mathcal{M}$ is a strongly l-coloured structure with the k_0-extension property and $a, b \in M - \mathrm{cl}(\emptyset)$ then*

$$\mathcal{M} \models \xi(a, b) \qquad \Longleftrightarrow \qquad \mathcal{M} \models P_i(a) \wedge P_i(b) \textit{ for some } i \in \{1, ..., l\}.$$

Remark 3.7. Suppose that we only consider L-structures in which R_1 (a symbol of V_{rel} with minimal arity) is interpreted as an irreflexive and symmetric relation. Then, for every $i = 2, \ldots, l$, the elements $a, v_i, w_{(a,i,1)}, ..., w_{(a,i,r_1-2)}$ from the proof of Lemma 3.5 must be different from each other, where $w_{(a,i,1)}, ..., w_{(a,i,r_1-2)} \in \mathrm{cl}_{\mathcal{M}}(a, v_i, u_{(a,i,1)}, ..., u_{(a,i,t-2)})$, and similarly for the sequences $b, v_i, w_{(b,i,1)}, ..., w_{(b,i,r_1-2)}$ and $v_k, v_i, w_{(k,i,1)}, ..., w_{(k,i,r_1-2)}$. Since $a, b, v_2, \ldots, v_l$ are different, by construction, this can be achieved if, for every n, every closed t-dimensional subset of G_n has cardinality at least r_1, where r_1 is the arity of R_1. Moreover, in the construction of $\mathcal{B}$ we must enlarge $R_1^{\mathcal{B}}$ so that whenever $\mathcal{B} \models R_1(c_1, \ldots, c_r)$ then $\mathcal{B} \models R_1(c_{\pi(c_1)}, \ldots, c_{\pi(c_r)})$ for every permutation π of $\{1, \ldots, r\}$. These changes do not affect the way in which $\mathcal{B}$ is coloured in steps (1)–(6).

4 Definability of l-colourings

Recall Assumptions 2.12. In this section we assume throughout that for some finite field F one of the following three cases hold for every $n \in \mathbb{N}$: (a) G_n is an n-dimensional vector space over F and $\mathrm{cl}_{\mathcal{G}_n}$ is the linear closure operator, or (b) G_n is an n-dimensional affine space over F and $\mathrm{cl}_{\mathcal{G}_n}$ is the affine closure operator, or (c) G_n is an n-dimensional projective space over F and $\mathrm{cl}_{\mathcal{G}_n}$ is the projective closure operator. Moreover, we assume that the language L_{pre} with which $\mathrm{cl}_{\mathcal{G}_n}$ is defined, according to Definition 2.3, is either L_{gen} from Example 2.4 with the same interpretations of symbols as explained in that example, or, provided we are in case (a) above, we have $L_{pre} = L_F$ where L_F is like in Example 2.6 with the same interpretations of symbols as explained there.

The assumption about the language L_{pre} guarantees that there is no other structure on $\mathcal{G}_n$ than that which is needed for defining the pregeometry. Therefore the following result, essentially of basic linear algebra, applies in the present context.

Lemma 4.1. *Let $n, m \in \mathbb{N}$. If $\{a_1, \dots, a_k\} \subseteq G_n$ and $\{b_1, \dots, b_k\} \subseteq G_m$ are independent sets, then there is an L_{pre}-isomorphism from $\mathrm{cl}_{\mathcal{G}_n}(a_1, \dots, a_k)$ to $\mathrm{cl}_{\mathcal{G}_m}(b_1, \dots, b_k)$ which maps a_i to b_i for all $i = 1, \dots, k$.*

In this section we will prove the same kind of result for l-colourable structures with underlying pregeometry $\mathcal{G}_n$ for some n as we did for strongly l-colourable structures in Section 3 (where the assumptions on $\mathcal{G}_n$ made here were not needed). More precisely, we will show that there are $k_0 \in \mathbb{N}$ and an L_{rel}-sentence $\xi(x, y)$ such that

- if $\mathcal{M}$ is l-coloured, $a, b \in M - \mathrm{cl}(\emptyset)$ and $\mathcal{M} \models \xi(a, b)$, then a and b have the same colour, and
- if $\mathcal{M}$ is l-coloured and has the k_0-extension property and $a, b \in M - \mathrm{cl}(\emptyset)$, then $\mathcal{M} \models \xi(a, b)$ if and only if a and b have the same colour.

We will define a certain l-colourable L_{rel}-structure $\mathcal{B}$ which will be used to define the sought after formula $\xi(x, y)$. In order to define such $\mathcal{B}$ we will use a theorem from structural Ramsey theory about colourings of vector spaces, projective spaces and affine spaces over a finite field.

Definition 4.2. Suppose that (V, cl) is a pregeometry.
(i) We call a function $c : V - \mathrm{cl}(\emptyset) \to \{1, \dots, l\}$ an ***l-colouring*** of (V, cl) if whenever $a, b \in V - \mathrm{cl}(\emptyset)$ and $a \in \mathrm{cl}(b)$, then $c(a) = c(b)$.
(ii) Suppose that $c : V - \mathrm{cl}(\emptyset) \to \{1, \dots, l\}$ is an l-colouring of (V, cl) and that W is a subspace (i.e. a closed subset) of V. If all $a \in W - \mathrm{cl}(\emptyset)$ are assigned the same colour by c, then we call W ***c-monochromatic***. If, in addition, there is no closed $U \subseteq V$ such that W is a proper subset of U, then we call W ***maximal c-monochromatic***.

The following theorem was proved by Graham, Leeb and Rothschild [5] and can also be found (in perhaps more accessible form) in [6] (Theorem 9 and Corollary 10 of Section 2.4). Recall that we have fixed a finite vector space F.

Theorem 4.3. *[5] For all $d, l \in \mathbb{N}$ there is a number $N(d, l) \in \mathbb{N}$ such that if $n \geq N(d, l)$, and the pregeometry (V, cl) is isomorphic with an n-dimensional vector space, projective space, or affine space over F and c is an l-colouring of (V, cl), then there exists at least one c-monochromatic subspace of (V, cl) with dimension at least d.*

Let $n = N(2, l)$ for $N(d, l)$ in the above theorem, let $\mathcal{V} = \mathcal{G}_n$ and let c be an l-colouring of $\mathcal{V}$. By our choice of n and Theorem 4.3 there exists at least one c-monochromatic subspace of $\mathcal{V}$ of dimension at least two and hence there also exists at least one maximal c-monochromatic subspace of $\mathcal{V}$ of dimension two. Let $W_1^c, ..., W_{t(c)}^c$ enumerate all the maximal c-monochromatic subspaces of $\mathcal{V}$ of dimension at least two, where $t(c)$ depends on the l-colouring c. (This '$t(c)$' has nothing to do with the '$t(d)$' used in the previous section.) Let C be the set of all l-colourings of $\mathcal{V}$. For each $c \in C$, choose a basis $\{d_1, ..., d_{e_c}\} \subseteq \bigcup_{i=1}^{t(c)} W_i$ for the closure of $\bigcup_{i=1}^{t(c)} W_i$, so in particular, $\bigcup_{i=1}^{t(c)} W_i$ has dimension e_c. Then let $e = \min\{e_c : c \in C\}$. Choose $c_0 \in C$ such that $e_{c_0} = e$ and for every other l-colouring $c \in C$ with $e_c = e$ we have that $t(c) \leq t(c_0)$. For this colouring c_0, let $m = t(c_0)$ and let $W_1 = W_1^{c_0}, ..., W_m = W_m^{c_0}$.

Assume that the relation symbol $R \in V_{rel}$ has minimal arity r among the relation symbols in V_{rel}, so $r \geq 2$. Let $\mathcal{B}$ be the expansion of $\mathcal{V} = \mathcal{G}_n$ to the language L_{rel} defined by, for each relation symbol $Q \in V_{rel} - \{R\}$, letting $Q^{\mathcal{B}} = \emptyset$ and defining $R^{\mathcal{B}}$ in the following way:

- If $v_1, v_2, ..., v_r \in W_i$ for some $i \in \{1, ..., m\}$ or if $\mathrm{cl}(v_1, ..., v_r) = \mathrm{cl}(v_j)$ for some $j \in \{1, \ldots, r\}$, then $\mathcal{B} \models \neg R(v_1, ..., v_r)$.
- If $\{v_1, ..., v_r\} \not\subseteq W_i$ for all $i = 1, ..., m$ and $\mathrm{cl}_{\mathcal{V}}\mathcal{V}(v_1, ..., v_r) \neq \mathrm{cl}(v_j)$ for all $j = 1, \ldots, r$, then $\mathcal{B} \models R(v_1, ..., v_r)$.

Notice that the second case holds if and only if the first case does not hold, so $\mathcal{B}$ is unambiguously defined. Let $b_1, b_2 \in W_1$ be independent (notice that they exist because of the choice of W_1) and let $\mathcal{A} = \mathcal{B} \restriction \mathrm{cl}(\{b_1, b_2\})$. Observe that since $A \subseteq W_1$ it follows that for every $Q \in V_{rel}$, $Q^{\mathcal{A}} = \emptyset$ Let $B = \{b_1, b_2, \ldots, b_\beta\}$ and let $\chi_{\mathcal{B}}(x_1, \ldots, x_\beta)$ be the characteristic formula of $\mathcal{B}$ with respect to the ordering $b_1, b_2, \ldots, b_\beta$ of B. So for every L_{rel}-structure $\mathcal{M}$ we have $\mathcal{M} \models \chi_{\mathcal{B}}(a_1, \ldots, a_\beta)$ if and only if the map $b_i \mapsto a_i$ is an embedding of $\mathcal{B}$ into $\mathcal{M}$.

Definition 4.4. Let $\xi_0(x, y)$ denote the L_{rel}-formula

$$\exists z_3, ..., z_\beta \chi_{\mathcal{B}}(x, y, z_3, ..., z_\beta).$$

We will use $\xi_0(x, y)$ to define the formula $\xi(x, y)$ with the properties that we are looking for, explained in the beginning of this section. Before defining $\xi(x, y)$ we need to assure that $\xi_0(x, y)$ has certain properties which are given by Lemmas 4.5–4.9. Notice that, by construction, $\mathcal{B} \restriction L_{pre} = \mathcal{V}$ so $\mathcal{B}$ and $\mathcal{V}$ have the same universe $B = V$ and $\mathrm{cl}_{\mathcal{B}}$ is the same as $\mathrm{cl}_{\mathcal{V}}$ (which is why we skip the subscripts of 'cl').

Lemma 4.5. *The function $c_0 : B - \mathrm{cl}(\emptyset) \to \{1, \ldots, l\}$ is an l-colouring of $\mathcal{B}$ (according to Definition 2.25). Consequently there exists an l-coloured structure $\mathcal{B}_0$ such that $\mathcal{B}_0 \restriction L_{rel} = \mathcal{B}$ and for every $b \in B - \mathrm{cl}(\emptyset)$ and every $i \in \{1, \ldots, l\}$, $\mathcal{B}_0 \models P_i(b)$ if and only if $c_0(b) = i$.*

Proof. We define a L-structure $\mathcal{B}_0$ by putting colour on $\mathcal{B}$ through the l-colouring c_0. In other words, we let $\mathcal{B}_0$ be the expansion of $\mathcal{B}$ to L such that for every $b \in B - \mathrm{cl}(\emptyset)$, $\mathcal{B}_0 \models P_{c_0(b)} \wedge \bigwedge_{j \neq c_0(b)} \neg P_j(b)$. Then $\mathcal{B}_0 \restriction L_{rel} \cong \mathcal{B}$ so we just need to prove the following:
Claim. $\mathcal{B}_0$ is l-coloured.
We need to check that (1)–(4) of Definition 2.14 are satisfied. Conditions (1) and (3) are satisfied since c_0 is an l-colouring of the underlying pregeometry $\mathcal{V}$ of $\mathcal{B}$. Let $R \in V_{rel}$ be as in the definition of $\mathcal{B}$. Let $Q \in V_{rel}$. If $Q \neq R$ then, by definition of $\mathcal{B}$ and $\mathcal{B}_0$, $Q^{\mathcal{B}_0} = \emptyset$ so (2) and (4) are satisfied for such Q. Now we consider the case $Q = R$. Suppose that $\mathcal{B}_0 \models R(a_1, \ldots, a_r)$. By the definition of $\mathcal{B}$ and $\mathcal{B}_0$ we have

- $\{a_1, \ldots, a_r\} \not\subseteq W_i$ for all $i = 1, \ldots, m$, and
- $\mathrm{cl}(v_1, ..., v_r) \neq \mathrm{cl}(v_j)$ for all $j = 1, \ldots, r$.

In particular, $\{a_1, \ldots, a_r\} \not\subseteq \mathrm{cl}(\emptyset)$ so (2) is satisfied. As $\mathrm{cl}(v_1, ..., v_r) \neq \mathrm{cl}(v_j)$ for all $j = 1, \ldots, r$, it follows that $\mathrm{cl}_{\mathcal{V}}(a_1, \ldots, a_r)$ has dimension at least 2. If $\mathrm{cl}(a_1, \ldots, a_r)$ would be c_0-monochromatic then it would be included in a maximal c_0-monochromatic subspace and, by the first point above, this would contradict the assumption (in the construction of $\mathcal{B}$) that $W_1, \ldots, W_m$ enumerate all maximal c_0-monochromatic subspaces of $\mathcal{V}$ of dimension at least 2. Hence $\mathrm{cl}(a_1, \ldots, a_r)$ is not monochromatic, so (4) is satisfied. Now the claim, and hence the lemma, is proved. □

The structure $\mathcal{B}_0$ from the previous lemma will be used further on. Recall the definition of the L_{rel}-formula $\xi_0(x, y)$ (Definition 4.4).

Lemma 4.6. *If $\mathcal{M}$ is an l-coloured structure, $v, w \in M - \mathrm{cl}(\emptyset)$ and $\mathcal{M} \models \xi_0(v, w)$ then v and w have the same colour, i.e. $\mathcal{M} \models P_i(v) \wedge P_i(w)$ for some $i \in \{1, ..., l\}$.*

Proof. Suppose that $\mathcal{M}$ is an l-coloured structure. Observe that if $\mathcal{B}'$ is a substructure of $\mathcal{M}{\restriction}L_{rel}$, then $\mathcal{M}$ induces a function $c : B' - \mathrm{cl}(\emptyset) \to \{1, \dots, l\}$ by letting, for every $b' \in B' - \mathrm{cl}(\emptyset)$, $c(b') = i$ if and only if $\mathcal{M} \models P_i(b')$. We call such a function c an l-colouring of $\mathcal{B}'$ although, strictly speaking, we can only be sure that it is a colouring of $\mathcal{B}'$ (in the sense of Definition 2.25) if $\mathcal{B}'$ is a closed substructure of $\mathcal{M}$. Recall the definition of $\mathcal{B}$ before Definition 4.4 and that c_0 is, by Lemma 4.5, an l-colouring of $\mathcal{B}$. The lemma will be proved with the help of the following claim.

Claim. Any isomorphism f (if such exists) from the L_{rel}-structure $\mathcal{B}$ to a substructure $\mathcal{B}' \subseteq \mathcal{M}{\restriction}L_{rel}$ induces a bijection between the maximal c_0-monochromatic subspaces of $\mathcal{B}$ and the maximal c-monochromatic subspaces of $\mathcal{B}'$, where c is the l-colouring of $\mathcal{B}'$ induced by $\mathcal{M}{\restriction}L_{rel}$.

Proof of the claim. Suppose that f is an isomorphism from $\mathcal{B}$ to a substructure $\mathcal{B}'$ of $\mathcal{M}{\restriction}L_{rel}$. Assume that c is the l-colouring of $\mathcal{B}'$ induced by $\mathcal{M}{\restriction}L_{rel}$, that is, for all $b \in B' - \mathrm{cl}(\emptyset)$ and $i \in \{1, \dots, l\}$, $\mathcal{M} \models P_i(b)$ if and only if $c(b) = i$. Let $W_1, ..., W_m$ enumerate, without repetition, the maximal c_0-monochromatic subspaces with dimension at least 2 of $\mathcal{V}$, and hence of $\mathcal{B}$, which where chosen when $\mathcal{B}$ was defined and let $b_1, b_2 \in W_1$ be the two independent elements which where chosen in the paragraph before Definition 4.4. Let $W'_1, ..., W'_p$ enumerate, without repetition, the maximal c-monochromatic subspaces of $\mathcal{B}'{\restriction}L_{pre}$ of dimension at least 2. By Theorem 4.3 this sequence is non-empty. We must show that $p = m$ and that there is a permutation π of $\{1, \dots, m\}$ such that $W'_i = f(W_{\pi(i)})$ for all $i = 1, \dots, m$.

Let $i \in \{1, \dots, p\}$. Let $v'_1 \in W'_i$ be arbitrary and, as the dimension of W'_i is at least 2, we can choose $v'_2, \dots, v'_r \in W'_i$ such that $\mathrm{cl}(v'_1, \dots, v'_r) \neq \mathrm{cl}(v'_j)$ for all $j = 1, \dots, r$. We know that W'_i is monochromatic, hence we must have that $\mathcal{B}' \models \neg R(v'_1, ..., v'_r)$. Choose $v_1, \dots, v_r \in B$ such that $f(v_j) = v'_j$ for $j = 1, \dots, r$. Since f is an isomorphism we have that $\mathcal{B} \models \neg R(v_1, ..., v_r)$ and $\mathrm{cl}(v_1, \dots, v_r) \neq \mathrm{cl}(v_j)$ for all $j = 1, \dots, r$. By the definition of $\mathcal{B}$, this implies that $v_1, ..., v_r \in W_{\pi(i)}$ for some $\pi(i) \in \{1, ..., m\}$ (as otherwise we would have $\mathcal{B} \models R(v_1, \dots, v_r)$), contradicting what we have concluded so far).

We have already proved that for each $i \in \{1, ..., p\}$ there is $\pi(i) \in \{1, ..., m\}$ so that $W'_i \subseteq f(\mathcal{W}_{\pi(i)})$. As f is an isomorphism, and therefore preserves dimension of sets, it follows that

$$\dim\Big(\bigcup_{i=1}^{p} W'_i\Big) \leq \dim\Big(\bigcup_{i=1}^{m} W_i\Big).$$

Observe that the l-colouring c of $\mathcal{B}'$ induces an l-colouring c_f of $\mathcal{B}$ by letting $c_f(b) = i$ if and only if $c(f(b)) = i$, for every $b \in B - \mathrm{cl}(\emptyset)$ and every $i \in \{1, \dots, l\}$. Therefore, $f^{-1}(W'_1), \dots, f^{-1}(W'_p)$ is an enumeration of maximal c_f-monochromatic subspaces of $\mathcal{V}$. It follows that if the above inequality would be strict, then $\dim\Big(\bigcup_{i=1}^{m} W_i\Big)$ would not be minimal among all possible choices of l-colourings of $\mathcal{V}$ and corresponding enumeration of maximal monochromatic subspaces, and this would contradict the choice of c_0. Hence we conclude that

$$\dim\Big(\bigcup_{i=1}^{p} W'_i\Big) = \dim\Big(\bigcup_{i=1}^{m} W_i\Big).$$

Recall that we have showed that for every $i \in \{1, \dots, p\}$ there is $\pi(i) \in \{1, \dots, m\}$ such that $W'_i \subseteq f(W_{\pi(i)})$. Suppose, for a contradiction, that this map $\pi : \{1, \dots, p\} \to$

$\{1, \ldots, m\}$ is *not* surjective. Then, as f preserves the dimension of sets, $\bigcup_{i=1}^{p} f(W_{\pi(i)})$ has strictly smaller dimension than $\bigcup_{i=1}^{m} W_i$. Since $\bigcup_{i=1}^{p} W_i' \subseteq \bigcup_{i=1}^{p} f(W_{\pi(i)})$ it follows that $\bigcup_{i=1}^{p} W_i'$ has strictly smaller dimension than $\bigcup_{i=1}^{m} W_i$ which contradicts what we have already proved. Therefore we conclude that $\pi : \{1, \ldots, p\} \to \{1, \ldots, m\}$ is surjective, from which it follows that $p \geq m$.

Recall the notation '$t(c)$' used in the definition of $\mathcal{B}$. By the choice of the colouring c_0 of $\mathcal{V}$ (and of $\mathcal{B}$) we have $p = t(c_f) \leq t(c_0) = m$. As also $p \geq m$ we get $p = m$ and since π is surjective (and p finite) it must be bijective. □

Now we continue with the proof of Lemma 4.6. Assume that $\mathcal{M} \models \xi_0(v, w)$. Then there is a $\mathcal{B}' \subseteq \mathcal{M}$ with $B' = \{v, w, b_3', ..., b_\beta'\}$ and $\mathcal{M} \models \chi_{\mathcal{B}}(v, w, b_3', ..., b_\beta')$. Recall the choice of maximal c_0-monochromatic subspaces $W_1, \ldots, W_m \subseteq V = B$ and independent $b_1, b_2 \in W_1$ in the construction of $\mathcal{B}$ before Definition 4.4. As $\chi_{\mathcal{B}}$ is the characteristic formula of $\mathcal{B}$ with respect to an enumeration of B starting with $b_1, b_2, \ldots$, there is an isomorphism $f : \mathcal{B} \to \mathcal{B}'$ such that $f(b_1) = v$ and $f(b_2) = w$, where $b_1, b_2 \in W_1$. By the claim we have that $f(W_1)$ is a monochromatic subset of $\mathcal{B}'$, with respect to the l-colouring c induced by $\mathcal{M}$, and since $v, w \in f(W_1)$ it follows that v and w must have the same colour in $\mathcal{M}$. □

Remember, from before Definition 4.4, that $\mathcal{A} = \mathcal{B}{\restriction}\mathrm{cl}(b_1, b_2)$, where b_1 and b_2 are independent elements of W_1. Let

$$k_0 = \max\big(\dim(\mathcal{B}), 3\big)$$

Lemma 4.7. *Let $\mathcal{M}$ be an l-coloured structure with the k_0-extension property, suppose that $v, w \in M$ and that $\mathcal{A}'$ is a substructure of $\mathcal{M}$ with universe $cl(v, w)$. If all elements in $A' - \mathrm{cl}(\emptyset)$ have the same colour and there is an isomorphism $f_0 : \mathcal{A}'{\restriction}L_{rel} \to \mathcal{A}$ such that $f_0(v) = b_1$ and $f_0(w) = b_2$ then $\mathcal{M} \models \xi_0(v, w)$.*

Proof. Let $\mathcal{M}$, $v, w \in M - \mathrm{cl}(\emptyset)$ and $\mathcal{A}'$ satisfy the assumptions of the lemma, from which it follows in particular that $\mathcal{A}'$ is a closed substructure of $\mathcal{M}$. Suppose that $f_0 : \mathcal{A}'{\restriction}L_{rel} \to \mathcal{A}$ is an isomorphism such that $f_0(v) = b_1$ and $f_0(w) = b_2$. Let $\mathcal{B}_0$ be the L-expansion of $\mathcal{B}$ from Lemma 4.5 and let $\mathcal{A}_0 = \mathcal{B}_0{\restriction}A$, so $\mathcal{A}_0$ is a closed substructure of $\mathcal{B}_0$. Since, by assumption, all elements of $A' - \mathrm{cl}(\emptyset)$ have the same colour we get $Q^{\mathcal{A}'} = \emptyset$ for all $Q \in V_{rel}$. By the definition of $\mathcal{B}_0$, all elements of $A_0 - \mathrm{cl}(\emptyset) = cl(b_1, b_2) - \mathrm{cl}(\emptyset)$ have the same colour, so $Q^{\mathcal{A}_0} = \emptyset$ for all $Q \in V_{rel}$. Let i be the colour of all elements in $A' - \mathrm{cl}(\emptyset)$. By permuting the colours if necessary we get an l-coloured structure $\mathcal{B}_0'$ such that if $\mathcal{A}_0' = \mathcal{B}_0'{\restriction}A$, then $\mathcal{A}_0'{\restriction}L_{rel} = \mathcal{A}_0{\restriction}L_{rel}$ and f_0 is an L-isomorphism from $\mathcal{A}'$ to $\mathcal{A}_0'$. Since $\mathcal{M}$ satisfies the k_0-extension property and $\dim(\mathcal{B}_0') = k_0$, there is an embedding $f : \mathcal{B}_0' \to \mathcal{M}$ which extends f_0^{-1}. Let $\mathcal{B}'$ be the L_{rel}-reduct of $\mathcal{M}{\restriction}\mathrm{im}(f)$, so $\mathcal{B} \cong \mathcal{B}'$. Since f is an L-isomorphism (where $L_{rel} \subseteq L$) which extends f_0^{-1} we have that $v, w \in B'$ and B' can be enumerated in such a way $v, w, b_3', ..., b_\beta'$ that $\mathcal{M} \models \chi_{\mathcal{B}}(v, w, b_3', ..., b_\beta')$. Hence $\mathcal{M} \models \xi_0(v, w)$. □

Lemma 4.8. *Assume that $\mathcal{M}$ is an l-coloured structure with the k_0-extension property. If $v, w \in \mathcal{M}$ are independent, and have the same colour, then there exists $u \in M - \mathrm{cl}(v, w)$ such that the following holds:*

> *Let $\mathcal{A}_{v,u} = \mathcal{M} \restriction \mathrm{cl}(v, u)$ and let $\mathcal{A}_{w,u} = \mathcal{M} \restriction \mathrm{cl}(w, u)$. Then $\mathcal{A}_{v,u}$ and $\mathcal{A}_{w,u}$ are monochromatic and there exist isomorphisms $f_{v,u} : \mathcal{A}_{v,u} \restriction L_{rel} \to \mathcal{A}$ and $f_{w,u} : \mathcal{A}_{w,u} \restriction L_{rel} \to \mathcal{A}$ such that $f_{v,u}(v) = b_1$, $f_{v,u}(u) = b_2$, $f_{w,u}(w) = b_1$ and $f_{w,u}(u) = b_2$.*

Proof. Suppose that $\mathcal{M}$ is l-coloured with the k_0-extension property and assume that $v, w \in M$ are independent from each other and have the same colour, say 1, without loss of generality. Let $\mathcal{S}$ be any l-coloured structure with dimension 3 and let $\mathcal{S}' = \mathrm{cl}_{\mathcal{S}}(\emptyset)$. By the definition of l-coloured structures and Assumption 2.12, $\mathcal{S}'$ is isomorphic with $\mathcal{M}{\restriction}\mathrm{cl}_{\mathcal{M}}(\emptyset)$. Since $k_0 \geq 3$ and $\mathcal{M}$ has the k_0-extension property it follows that $\mathcal{M}$ has a closed substructure which is isomorphic with $\mathcal{S}$. Therefore $\dim(\mathcal{M}) \geq 3$ and hence there is $u_0 \in M$ such that $\{v, w, u_0\}$ is an independent set. Let $C = \mathrm{cl}(v, w)$ and $\mathcal{C} = \mathcal{M}{\restriction}C$. We will now construct an l-coloured structure $\mathcal{D}$ and show that $\mathcal{C}$ is contained within an isomorphic copy $\mathcal{D}'$ of $\mathcal{D}$. The conclusions of the lemma will then follow easily because of the definition of $\mathcal{D}$.

Let $\mathcal{D}$ have universe $D = \mathrm{cl}(v, w, u_0)$. Interpret the symbols in V_{pre} so that $\mathcal{D}{\restriction}L_{pre}$ is the substructure of $\mathcal{M}{\restriction}L_{pre}$ with universe D. Interpret the symbols of $V_{col} \cup V_{rel}$ so that $\mathcal{D}{\restriction}C = \mathcal{M}{\restriction}C$. For all $d \in D - C$ let $\mathcal{D} \models P_1(d)$ and $\mathcal{D} \not\models P_i(d)$ if $i \neq 1$. For every $Q \in V_{rel}$, of arity q say, and every $\bar{d} \in D^q - C^q$, let $\mathcal{D} \not\models Q(\bar{d})$. From the definition it is clear that $\mathcal{D}$ is l-coloured.

Now we show that if $Q \in V_{rel}$ and $\bar{d} \in \mathrm{cl}(v, u_0)$, then $\mathcal{D} \not\models Q(\bar{d})$. Suppose for a contradiction that $Q \in V_{rel}$, $\bar{d} \in \mathrm{cl}(v, u_0)$ and $\mathcal{D} \models Q(\bar{d})$. By definition of $\mathcal{D}$ we have $\bar{d} \in C$ and, by Lemma 2.2, $C \cap \mathrm{cl}(v, u_0) = \mathrm{cl}(v)$, so $\bar{d} \in \mathrm{cl}(v) \subseteq C$. By the definition of $\mathcal{D}$ we get $\mathcal{M} \models Q(\bar{d})$, which contradicts that $\mathcal{M}$ is l-coloured (as all members of $\bar{d}$ belong to the same 1-dimensional subspace). By a similar proof (replace v by w) it follows that if $Q \in V_{rel}$ and $\bar{d} \in \mathrm{cl}(w, u_0)$, then $\mathcal{D} \not\models Q(\bar{d})$.

Now let $\mathcal{A}_{v,u_0}$ be the substructure of $\mathcal{D}$ with universe $A_{v,u_0} = \mathrm{cl}(v, u_0)$. From the definition of $\mathcal{D}$ it follows that all elements of $\mathcal{A}_{v,u_0}$ have colour 1. Recall the definition of the L_{rel}-structure $\mathcal{A}$ with universe $A = \mathrm{cl}(b_1, b_2)$ before Definition 4.4. As v is independent from u_0, it follows from Lemma 4.1 that there is an L_{pre}-isomorphism $f_{v,u_0} : \mathcal{A}_{v,u_0}{\restriction}L_{pre} \to \mathcal{A}{\restriction}L_{pre}$ such that $f_{v,u_0}(v) = b_1$ and $f_{w,u_0}(u_0) = b_2$. From the definition of $\mathcal{A}$ we have $Q^{\mathcal{A}} = \emptyset$ for every $Q \in V_{rel}$. Since $\mathcal{A}_{v,u_0}$ has universe $\mathrm{cl}(v, u_0)$ it follows from what we proved above and the definition of $\mathcal{A}_{v,u_0}$ that $Q^{\mathcal{A}_{v,u_0}} = \emptyset$ for every $Q \in V_{rel}$ Therefore, f is also an L_{rel}-isomorphism from $\mathcal{A}_{v,u_0}{\restriction}L_{rel}$ to $\mathcal{A}$. In the same way we can show that if $\mathcal{A}_{w,u_0}$ be the substructure of $\mathcal{D}$ with universe $A_{w,u_0} = \mathrm{cl}(w, u_0)$, then all elements of $\mathcal{A}_{w,u_0}$ have colour 1 and there is an isomorphism $f_{w,u_0} : \mathcal{A}_{w,u_0}{\restriction}L_{rel} \to \mathcal{A}$ such that $f_{w,u_0}(w) = b_1$ and $f_{w,u_0}(u_0) = b_2$.

Because of what has been proved above it now suffices to show that there are a substructure $\mathcal{D}' \subseteq \mathcal{M}$ such that $\mathcal{C} \subseteq \mathcal{D}'$ and an isomorphism $f : \mathcal{D} \to \mathcal{D}'$ such that f is the identity on C. Then $u = f(u_0)$ has the desired property. Since $\dim(\mathcal{D}) = 3 \leq k_0$ and $\mathcal{M}$ has the k_0-extension property it follows that, in particular, $\mathcal{M}$ has the $\mathcal{D}/\mathcal{C}$-extension property. Therefore such $\mathcal{D}'$ and f exist. $\square$

Now we put together the previous two lemmas to get the following.

Lemma 4.9. *Assume that $\mathcal{M}$ is l-coloured with the k_0-extension property. If $v, w \in \mathcal{M}$ are independent and have the same colour then there exists $u \in M - \mathrm{cl}(v, w)$ such that $\mathcal{M} \models \xi_0(v, u) \wedge \xi_0(w, u)$.*

Proof. By Lemma 4.8, there is $u \in M - \mathrm{cl}(v, w)$ and monochromatic structures $\mathcal{A}_{v,u}, \mathcal{A}_{w,u} \subseteq \mathcal{M}$ with $A_{v,u} = \mathrm{cl}(v, u)$ and $A_{w,u} = \mathrm{cl}(w, u)$, isomorphisms $f_{v,u} : \mathcal{A}_{v,u} \restriction L_{rel} \to \mathcal{A}$ and $f_{w,u} : \mathcal{A}_{w,u} \restriction L_{rel} \to \mathcal{A}$ with $f_{v,u}(v) = f_{w,u}(w) = a$ and $f_{v,u}(u) = f_{w,u}(u) = b$. So by Lemma 4.7 and $f_{v,u}$ we get $\mathcal{M} \models \xi_0(v, u)$ and then, using $f_{w,u}$ and by Lemma 4.7, we get $\mathcal{M} \models \xi_0(w, u)$. Hence $\mathcal{M} \models \xi_0(v, u) \wedge \xi_0(w, u)$. $\square$

We can finally define the desired L_{rel}-formula $\xi(x, y)$ and prove, in Corollary 4.12, that it has the property of telling whether elements have the same colour or not.

Definition 4.10. Let $\xi(x, y)$ be the L_{rel}-formula

$$x \in \mathrm{cl}(y) \ \vee \ \exists z(\xi_0(x, z) \wedge \xi_0(y, z)).$$

Observe that since $\xi_0(x, y)$ is an existential formula, that is, $\xi_0(x, y)$ has the form $\exists \bar{z}\psi(x, y, \bar{z})$ where ψ is quantifier free, it follows, from the assumptions in the beginning of this section, that $\xi(x, y)$ is logically equivalent to an existential formula. This will be used in Section 5.

Corollary 4.11. *If $\mathcal{M}$ is l-coloured, $v, w \in M - \mathrm{cl}(\emptyset)$, $v \notin \mathrm{cl}(w)$ and $\mathcal{M} \models \xi(v, w)$, then v and w have the same colour, i.e. $\mathcal{M} \models P_i(v) \wedge P_i(w)$ for some $i \in \{1, ..., l\}$.*

Proof. If $\mathcal{M}$ is l-coloured, $v, w \in M - \mathrm{cl}(\emptyset)$, $v \notin \mathrm{cl}(w)$ and $\mathcal{M} \models \xi(v, w)$, then $\mathcal{M} \models \xi_0(v, u) \wedge \xi_0(w, u)$ for some $u \in M$. By Lemma 4.6, v has the same colour as u and u has the same colour as w. Hence, v and w have the same colour. □

Corollary 4.12. *Let $\mathcal{M}$ be l-coloured with the k_0-extension property. If $v, w \in M - \mathrm{cl}(\emptyset)$ then*

$$\mathcal{M} \models \xi(v, w) \iff v \text{ and } w \text{ have the same colour.}$$

Proof. Suppose that $v, w \in M - \mathrm{cl}(\emptyset)$ have the same colour. If v and w are dependent then $v \in \mathrm{cl}(w)$ so $\mathcal{M} \models \xi(v, w)$. Assume that $v \notin \mathrm{cl}(w)$. By Lemma 4.9, there is $u \in M$ such that $\mathcal{M} \models \xi_0(v, u) \wedge \xi_0(w, u)$, so by the definition of ξ we get that $\mathcal{M} \models \xi(v, w)$. The opposite direction is proved like Corollary 4.11 □

5 Almost sure properties and an axiomatisation of the limit theory

In this section we show that if an L_{rel}-formula $\xi(x, y)$ exists which defines the l-colouring of an (strongly) l-coloured structure in the sense of (1) and (2) of Theorem 5.1 below, then we can draw some conclusions about the asymptotic structure of (strongly) l-colurable structures and, if $\xi(x, y)$ is existential then we get an explicit axiomatisation of the set of sentences with limit probability 1. Theorem 5.1 together with the results in Sections 2–4 imply the main results stated in Section 1.

We recall the notation from Definitions 2.14, 2.17 and 2.23. So in particular, $\mathbf{K}_n$ denotes the set of l-coloured structures $\mathcal{M}$ such that $\mathcal{M} \restriction L_{pre} = \mathcal{G}_n$ and $\delta_n^{\mathbf{K}}$ denotes the dimension conditional measure on $\mathbf{K}_n$. $\mathbf{C}_n$ denotes the set of l-colourable structures $\mathcal{M}$ such that $\mathcal{M} \restriction L_{pre} = \mathcal{G}_n$ and $\delta_n^{\mathbf{C}}$ is the probability measure on $\mathbf{C}_n$ derived from $\delta_n^{\mathbf{K}}$. Similarly, $\mathbf{SK}_n$ denotes the set of strongly l-coloured structures $\mathcal{M}$ such that $\mathcal{M} \restriction L_{pre} = \mathcal{G}_n$ and $\delta_n^{\mathbf{SK}}$ denotes the dimension conditional measure on $\mathbf{SK}_n$. $\mathbf{S}_n$ denotes the set of strongly l-colourable structures $\mathcal{M}$ such that $\mathcal{M} \restriction L_{pre} = \mathcal{G}_n$ and $\delta_n^{\mathbf{S}}$ is the probability measure on $\mathbf{S}_n$ derived from $\delta_n^{\mathbf{SK}_n}$. For any L_{rel}-sentence φ, let

$$\delta_n^{\mathbf{C}}(\varphi) \ = \ \delta_n^{\mathbf{C}}\big(\{\mathcal{M} \in \mathbf{C}_n : \mathcal{M} \models \varphi\}\big),$$

and similarly for $\delta_n^{\mathbf{S}}(\varphi)$. In this section we will prove the following result.

Theorem 5.1. *Suppose that the conditions of Assumption 2.12 hold, that $\mathbf{G} = \{\mathcal{G}_n : n \in \mathbb{N}\}$ is polynomially k-saturated for every $k \in \mathbb{N}$ and that there exists an L_{rel}-formula $\xi(x, y)$ and natural number k_0 with the following properties.*

(1) If $\mathcal{M}$ is an l-coloured structure, $a, b \in M - \mathrm{cl}_{\mathcal{M}}(\emptyset)$ and $\mathcal{M} \models \xi(a, b)$, then a and b have the same colour (i.e. $\mathcal{M} \models P_i(a) \wedge P_i(b)$ for some $i \in \{1, \ldots, l\}$).

(2) If $\mathcal{M}$ is an l-coloured structure that has the k_0-extension property and $a, b \in M - \mathrm{cl}_{\mathcal{M}}(\emptyset)$, then $\mathcal{M} \models \xi(a, b)$ if and only if a and b have the same colour.

Then the following hold:

(i) The $\delta_n^{\mathbf{C}}$-probability that the following holds for $\mathcal{M} \in \mathbf{C}_n$ approaches 1 as $n \to \infty$:

For all $a, b \in M - \mathrm{cl}_{\mathcal{M}}(\emptyset)$, $\mathcal{M} \models \xi(a, b)$ if and only if every *l-colouring of $\mathcal{M}$ gives a and b the same colour.*

(ii) $\lim_{n\to\infty} \delta_n^{\mathbf{C}}\big(\{\mathcal{M} \in \mathbf{C}_n : \ \mathcal{M}$ *has a unique l-colouring*$\}\big) = 1$.

(iii) $\lim_{n\to\infty} \delta_n^{\mathbf{C}}\big(\{\mathcal{M} \in \mathbf{C}_n : \ \mathcal{M}$ *is not l'-colourable if $l' < l$*$\}\big) = 1$.

(iv) Suppose, in addition, that $\xi(x, y)$ is an existential formula. Then the set of L_{rel}-sentences φ such that $\lim_{n\to\infty} \delta_n^{\mathbf{C}}(\varphi) = 1$ *forms a countably categorical theory which can be given an explicit axiomatization where every axiom is logically equivalent to a sentence of the form $\forall \bar{x} \exists \bar{y} \psi(\bar{x}, \bar{y})$ where ψ is quantifier-free.*

If the assumptions hold for strongly *l-coloured structures, then (i)–(iv) hold if every occurence of $\mathbf{C}$ is replaced by $\mathbf{S}$.*

Observe that in Sections 3 and 4 we have proved, under the assumptions of Theorems 1.3 and 1.2, respectively, that there are a number k_0 and an L_{rel}-formula $\xi(x, y)$ such that (1) and (2) of Theorem 5.1 are satisfied. Moreover, if all L_{pre}-formulas θ_n which define the pregeometry are quantifier free, then the formula $\xi(x, y)$ obtained in Section 3 and in Section 4 is logically equivalent to an existential formula. Therefore Theorems 1.3 and 1.2 follow from Theorem 5.1 and the results in Sections 3 and 4. So it remains to prove Theorem 5.1.

Proof of Theorem 5.1

The proof is exactly the same in the case of l-colourable structures as in the case of strongly l-colourable structures. Therefore we will speak only of 'l-colourable (or coloured) structures' and use the notations $\mathbf{K}_n$, $\mathbf{C}_n$, $\delta_n^{\mathbf{K}}$ and $\delta_n^{\mathbf{C}}$. (If we replace the mentioned terminology and notation with 'strongly l-colourable (or coloured) structures', $\mathbf{SK}_n$, $\mathbf{S}_n$, $\delta_n^{\mathbf{SK}}$ and $\delta_n^{\mathbf{S}}$, then we have a proof for strongly l-colourable structures.) The general idea of the proof is to first define an L_{rel}-theory $T_{\mathbf{C}}$ such that for every $\varphi \in T_{\mathbf{C}}$, $\lim_{n\to\infty} \delta_n^{\mathbf{C}}(\varphi) = 1$. Then it will follow from compactness that $T_{\mathbf{C}}$ is consistent. The next step is to prove that $T_{\mathbf{C}}$ is complete, which will be done by proving that it is countably categorical and applying Vaught's theorem. When these steps have been carried out it follows easily, since (by compactness) $T_{\mathbf{C}} \models \varphi$ implies $\Delta \models \varphi$ for some finite $\Delta \subseteq T_{\mathbf{C}}$, that for every L_{rel}-sentence φ, either $\lim_{n\to\infty} \delta_n^{\mathbf{C}}(\varphi) = 0$ or $\lim_{n\to\infty} \delta_n^{\mathbf{C}}(\varphi) = 1$.

We assume that the conditions of Assumption 2.12 hold and that $\mathbf{G} = \{\mathcal{G}_n : n \in \mathbb{N}\}$ is polynomially k-saturated for every $k \in \mathbb{N}$. Let k_0 be a natural number and $\xi(x, y)$ an L_{rel}-formula such that

(1) If $\mathcal{M}$ is an l-coloured structure, $a, b \in M - \mathrm{cl}(\emptyset)$ and $\mathcal{M} \models \xi(a,b)$, then a and b have the same colour (i.e. $\mathcal{M} \models P_i(a) \wedge P_i(b)$ for some $i \in \{1, \ldots, l\}$).

(2) If $\mathcal{M}$ is an l-coloured structure that has the k_0-extension property and $a, b \in M - \mathrm{cl}(\emptyset)$, then $\mathcal{M} \models \xi(a,b)$ if and only if a and b have the same colour.

Without loss of generality we may assume that $k_0 \geq 1$.

Lemma 5.2. *Suppose that $\mathcal{M}$ is an l-coloured structure that has the k_0-extension property. Then the following hold:*

(i) For every $i \in \{1, \ldots, l\}$ there is $a \in M$ with colour i (i.e. $\mathcal{M} \models P_i(a)$).

(ii) The formula $\xi(x,y)$ defines an equivalence relation on $M - \mathrm{cl}(\emptyset)$ such that for all $a, b \in M - \mathrm{cl}(\emptyset)$ we have $\mathcal{M} \models \xi(a,b)$ if and only if a and b have the same colour.

(iii) The set $M - \mathrm{cl}(\emptyset)$ is partitioned into exactly l (nonempty) equivalence classes by the equivalence relation defined by $\xi(x,y)$.

Proof. Suppose that $\mathcal{M}$ is an l-coloured structure that has the k_0-extension property.

(i) For every l-coloured $\mathcal{N}$ and 1-dimensional closed substructure $\mathcal{A} \subseteq \mathcal{N}$, all $a \in A - \mathrm{cl}(\emptyset)$ have the same colour, say j. If we change the colour of all $a \in A - \mathrm{cl}(\emptyset)$ to i, say, then the resulting structures is still l-coloured. As we assume that $\mathcal{M}$ has the k_0-extension property (and $\dim(\mathcal{A}) = 1 \leq k_0$) it follows (since $\mathcal{A} \restriction \mathrm{cl}_{\mathcal{A}}(\emptyset) \cong \mathcal{M} \restriction \mathrm{cl}_{\mathcal{M}}(\emptyset)$) that $\mathcal{M}$ has a substructure that is isomorphic with $\mathcal{A}$ and therefore some element of $\mathcal{M}$ has colour i (where i is an arbitrary colour).

(ii) Follows directly from (2).

(iii) By part (i), for every $i \in \{1, \ldots, l\}$, there is some $a \in M - \mathrm{cl}(\emptyset)$ with colour i. Hence, it follows from part (ii) that $M - \mathrm{cl}(\emptyset)$ is partitioned into exactly l different equivalence classes by the relation defined by $\xi(x,y)$. $\square$

The next lemma proves part (i) of Theorem 5.1.

Lemma 5.3. *With $\delta_n^{\mathbf{C}}$-probability approaching 1 as $n \to \infty$ a structure $\mathcal{M} \in \mathbf{C}_n$ has the following property:*

> *For all $a, b \in M - \mathrm{cl}(\emptyset)$ we have $\mathcal{M} \models \xi(a,b)$ if and only if every l-colouring of $\mathcal{M}$ gives a and b the same colour. (In other words, whenever $\mathcal{M}' \in \mathbf{K}_n$ and $\mathcal{M}' \restriction L_{rel} = \mathcal{M}$, then for all $a, b \in M - \mathrm{cl}(\emptyset)$, $\mathcal{M} \models \xi(a,b) \iff \mathcal{M}' \models P_i(a) \wedge P_i(b)$ for some i.)*

Proof. For every $n \in \mathbb{N}$, let

$$\begin{aligned} \mathbf{X}_n^{\mathbf{K}} &= \big\{\mathcal{M} \in \mathbf{K}_n : \mathcal{M} \text{ has the } k_0\text{-extension property}\big\}, \text{ and} \\ \mathbf{X}_n^{\mathbf{C}} &= \big\{\mathcal{M} \in \mathbf{C}_n : \mathcal{M} = \mathcal{N} \restriction L_{rel} \text{ for some } \mathcal{N} \in \mathbf{X}_n^{\mathbf{K}}\big\}. \end{aligned}$$

By the definition of $\delta_n^{\mathbf{C}}$ we have

$$\begin{aligned} \delta_n^{\mathbf{C}}\big(\mathbf{X}_n^{\mathbf{C}}\big) &= \delta_n^{\mathbf{K}}\big(\mathcal{M} \in \mathbf{K}_n : \mathcal{M} \restriction L_{rel} \in \mathbf{X}_n^{\mathbf{C}}\}\big) \\ &= \delta_n^{\mathbf{K}}\big(\{\mathcal{M} \in \mathbf{K}_n : \mathcal{M} \restriction L_{rel} = \mathcal{N} \restriction L_{rel} \text{ for some } \mathcal{N} \in \mathbf{X}_n^{\mathbf{K}}\}\big) \ \geq \ \delta_n^{\mathbf{K}}\big(\mathbf{X}_n^{\mathbf{K}}\big). \end{aligned}$$

By Theorem 2.22 we have $\lim_{n\to\infty} \delta_n^{\mathbf{K}}\big(\mathbf{X}_n^{\mathbf{K}}\big) = 1$ and hence $\lim_{n\to\infty} \delta_n^{\mathbf{C}}\big(\mathbf{X}_n^{\mathbf{C}}\big) = 1$. Therefore it suffices to prove that if $\mathcal{M} \in \mathbf{X}_n^{\mathbf{C}}$ then

$(*)$ for every $\mathcal{M}' \in \mathbf{K}_n$ such that $\mathcal{M}' \restriction L_{rel} = \mathcal{M}$ and all $a, b \in M - \mathrm{cl}(\emptyset)$, we have $\mathcal{M} \models \xi(a, b)$ if and only if $\mathcal{M}' \models P_i(a) \wedge P_i(b)$ for some $i \in \{1, \ldots, l\}$.

So suppose that $\mathcal{M} \in \mathbf{X}_n^{\mathbf{C}}$, $\mathcal{M}' \in \mathbf{K}_n$ and $\mathcal{M}' \restriction L_{rel} = \mathcal{M}$. As $\mathcal{M} \in \mathbf{X}_n^{\mathbf{C}}$ there is $\mathcal{N} \in \mathbf{X}_n^{\mathbf{K}}$ such that $\mathcal{N} \restriction L_{rel} = \mathcal{M}$. By definition of $\mathbf{X}_n^{\mathbf{K}}$, $\mathcal{N}$ has the k_0-extension property, so by Lemma 5.2, $\xi(x, y)$ defines, in $\mathcal{N}$, an equivalence relation on $N - \mathrm{cl}(\emptyset)$ with exactly l equivalence classes. Since $\xi(x, y) \in L_{rel}$ and $\mathcal{N} \restriction L_{rel} = \mathcal{M}$ it follows that $\xi(x, y)$ defines, in $\mathcal{M}$, an equivalence relation on $M - \mathrm{cl}(\emptyset)$ with exactly l equivalence classes.

Note that if $a, b \in M - \mathrm{cl}(\emptyset)$ and $\mathcal{M} \models \xi(a, b)$, then, by (1), we have $\mathcal{M}' \models P_i(a) \wedge P_i(b)$ for some i. It follows that the equivalence relation defined by $\xi(x, y)$ on $M - \mathrm{cl}(\emptyset)$ refines the equivalence relation induced on $M - \mathrm{cl}(\emptyset)$ by the colouring of $\mathcal{M}'$. Since both equivalence relations have exactly l equivalence classes it follows that they are the same relation. In other words, for all $a, b \in M - \mathrm{cl}(\emptyset)$, $\mathcal{M} \models \xi(a, b)$ if and only if $\mathcal{M}' \models P_i(a) \wedge P_i(b)$ for some i. Hence we have proved $(*)$ and the proof of the lemma is finished. $\square$

Observe that Lemma 5.3 immediately implies the following which proves part (ii) of Theorem 5.1:

Corollary 5.4. $\delta_n^{\mathbf{C}}\big(\{\mathcal{M} \in \mathbf{C}_n : \ \mathcal{M} \textit{ has a unique } l\textit{-colouring}\}\big) = 1$

The next corollary proves part (iii) of Theorem 5.1. It implies that there exists an l-colourable structure which cannot be l'-coloured if $l' < l$. This may seem obvious, but if the reader tries to explicitly construct such a structure it may become apparent that it is, on the level of generality considered here, not a trivial problem.

Corollary 5.5. *If* $1 \leq l' < l$, *then*

$$\lim_{n \to \infty} \delta_n^{\mathbf{C}}\big(\{\mathcal{M} \in \mathbf{C}_n : \mathcal{M} \textit{ is not } l'\textit{-colourable}\}\big) \ = \ 1.$$

Proof. Let $1 \leq l' < l$. Note that every l'-colouring of a structure (using only the colours $1, \ldots, l'$) is also an l-colouring. Suppose that $\mathcal{M} \in \mathbf{C}_n$ has an l'-colouring, that is, there is $\mathcal{M}' \in \mathbf{K}_n$ such that $\mathcal{M}' \restriction L_{rel} = \mathcal{M}$ and $(P_i)^{\mathcal{M}'} = \emptyset$ for all $i = (l'+1), \ldots, l$. If n is sufficiently large then $\mathcal{M}$ also has an l-colouring in which all colours $1, \ldots, l$ are used, that is, there is $\mathcal{M}'' \in \mathbf{K}_n$ such that $\mathcal{M}'' \restriction L_{rel} = \mathcal{M}$ and $(P_i)^{\mathcal{M}''} \neq \emptyset$ for all $i = 1, \ldots, l$. Clearly the two colourings of $\mathcal{M}$ are not permutations of each other, that is, there is no permutation π of $\{1, \ldots, l\}$ such that for every $i \in \{1, \ldots, l\}$ and every $a \in M - \mathrm{cl}(\emptyset)$ we have $\mathcal{M}' \models P_i(a)$ if and only if $\mathcal{M}'' \models P_{\pi(i)}(a)$. Hence, for large enough n, if $\mathcal{M} \in \mathbf{C}_n$ has a unique l-colouring and $l' < l$, then $\mathcal{M}$ is not l'-colourable. Therefore Corollary 5.5 follows from Corollary 5.4. $\square$

Now it remains to prove part (iv) of Theorem 5.1. *So for the rest of this section we add the assumption that* $\xi(x, y)$ *is an existential formula.* We will give an explicit axiomatisation of the set of L_{rel}-sentences with asymptotic probability 1 and show that the given axioms form a countably categorical theory. The axiomatisation of the limit theory

$$\Big\{\varphi \in L_{rel} : \lim_{n \to \infty} \delta_n^{\mathbf{C}}(\varphi) = 1\Big\}$$

will be denoted $T_{\mathbf{C}}$ and will consist of four disjoint parts, denoted T_ξ, T_{pre}, T_{iso} and T_{ext}. Note that since we know, by Corollary 2.24, that $\mathbf{C}_n$ satisfies a zero-one law when the measure $\delta_n^{\mathbf{C}}$ is used, it follows that the limit theory is consistent (by compactness) and complete. We will show that whenever $\varphi \in T_{\mathbf{C}}$, then $\lim_{n \to \infty} \delta_n^{\mathbf{C}}(\varphi) = 1$ and that $T_{\mathbf{C}}$ is countably categorical, hence complete. It will then follow that, for every L_{rel}-sentence φ, $T_{\mathbf{C}} \models \varphi$ if and only if $\lim_{n\to\infty} \delta_n^{\mathbf{C}}(\varphi) = 1$. The part of the axiomatisation $T_{\mathbf{C}}$ which we denote T_ξ consists of only one sentence $\varphi_1 \wedge \varphi_2$, where φ_1 and φ_2 are defined below.

Recall that, by Assumption 2.12, in every l-coloured, or l-colourable, structure, the property "x belongs to the closure of $\emptyset$" is defined by the formula $\theta_0(x)$ and the property "y belongs to the closure of $\{x\}$" is defined by the formula $\theta_1(x, y)$.

Definition 5.6. (i) Let $\mathcal{U}$ be an l-colourable structure which is not l'-colourable if $l' < l$, and let $p = |U|$. Such $\mathcal{U}$ exists by Corollary 5.5
(ii) Let φ_1 be an L_{rel}-sentence which expresses that $\xi(x, y)$ defines an equivalence relation on the set of elements *not* satisfying $\theta_0(x)$.
(iii) Let φ_2 be the following L_{rel}-sentence:

$$\exists x_1, \ldots, x_p \Bigg(\chi_{\mathcal{U}}(x_1, \ldots, x_p) \wedge \bigvee_{\substack{I \subseteq \{1,\ldots,p\} \\ |I| = l}} \Bigg[\bigwedge_{i\in I} \neg\theta_0(x_i) \ \wedge \ \bigwedge_{\substack{i,j\in I \\ i\neq j}} \neg\xi(x_i, x_j) \ \wedge \ \forall y \Big(\theta_0(y) \ \vee \ \bigvee_{i\in I} \xi(y, x_i) \Big) \Bigg] \Bigg),$$

where $\chi_{\mathcal{U}}(x_1, \ldots, x_p)$ is the characteristic formula of $\mathcal{U}$ for some enumeration of U.

Lemma 5.7. $\lim_{n\to\infty} \delta_n^{\mathbf{C}}(\varphi_1 \wedge \varphi_2) \ = \ 1.$

Proof. Let $\mathcal{U}^+$ be an l-coloured structure such that $\mathcal{U}^+ \restriction L_{rel} = \mathcal{U}$. By Assumption 2.12 (6) and the definition of l-coloured structures there is a unique, up to isomorphism, l-coloured structure of dimension 0. So if $\mathcal{V} = \mathrm{cl}_{\mathcal{U}^+}(\emptyset)$ then every l-coloured structure has a substructure which is isomorphic to $\mathcal{V}$. It follows that if $\mathcal{M}$ is an l-coloured structure which has the $\mathcal{U}^+/\mathcal{V}$-extension property, then $\mathcal{M}$ has a substructure which is isomorphic to $\mathcal{U}^+$ and therefore $\mathcal{M}\restriction L_{rel}$ has a substructure which is isomorphic to $\mathcal{U}$. Let $k = \max(k_0, \dim(\mathcal{U}^+))$. Note that, by (1), (2) and Lemma 5.2, every l-coloured structure $\mathcal{M}$ with the k-extension property has the following properties:

- $\mathcal{M}$ has a substructure which is isomorphic with $\mathcal{U}^+$.
- $\xi(x, y)$ defines an equivalence relation on $M - \mathrm{cl}(\emptyset)$ such that, for all $a, b \in M - \mathrm{cl}(\emptyset)$, $\mathcal{M} \models \xi(a, b)$ if and only if a and b have the same colour.

We will now prove that if $\mathcal{M}$ is l-coloured and has the k-extension property, then $\mathcal{M} \models \varphi_1 \wedge \varphi_2$.

So suppose that $\mathcal{M}$ is l-coloured and has the k-extension property. Then, as mentioned above, $\mathcal{M}$ has a substructure which is isomorphic to $\mathcal{U}^+$ and ξ defines an equivalence relation on $M - \mathrm{cl}(\emptyset)$, so $\mathcal{M} \models \varphi_1$. It remains to show that $\mathcal{M} \models \varphi_2$. For notational simplicity we assume $\mathcal{U}^+ \subseteq \mathcal{M}$. Let $U = \{a_1, \ldots, a_p\}$ be an enumeration of U such that $\mathcal{M} \models \chi_{\mathcal{U}}(a_1, \ldots, a_p)$. As at least l different colours are needed to colour $\mathcal{U}$, there are $a_{i_1}, \ldots, a_{i_l} \in U - \mathrm{cl}(\emptyset)$ such that if $j \neq j'$ then a_{i_j} has a different colour than $a_{i_{j'}}$, so $\mathcal{M} \models \neg\xi(a_{i_j}, a_{i_{j'}})$. Let $I = \{i_1, \ldots, i_l\}$. Since there are only l colours and all $a_{i_1}, \ldots, a_{i_l}$ have different colours, it follows that every $b \in M - \mathrm{cl}(\emptyset)$ must have the same colour as some a_{i_j} which implies $\mathcal{M} \models \xi(b, a_{i_j})$. Hence $\mathcal{M} \models \varphi_2$.

We have proved that if $\mathcal{M}$ is l-coloured and has the k-extension property, then $\mathcal{M} \models \varphi_1 \wedge \varphi_2$. Consequently,

$$\delta_n^{\mathbf{K}}\big(\{\mathcal{M} \in \mathbf{K}_n : \mathcal{M} \text{ has the } k\text{-extension property}\}\big) \ \leq \ \delta_n^{\mathbf{K}}(\varphi_1 \wedge \varphi_2),$$

so by Theorem 2.22, $\lim_{n\to\infty} \delta_n^{\mathbf{K}}(\varphi_1 \wedge \varphi_2) = 1$. Since $\varphi_1 \wedge \varphi_2$ is an L_{rel}-sentence we have $\mathcal{M} \models \varphi_1 \wedge \varphi_2$ if and only if $\mathcal{M}{\restriction}L_{rel} \models \varphi_1 \wedge \varphi_2$, for every l-coloured structure $\mathcal{M}$. By the definition of $\delta_n^{\mathbf{C}}$ we get $\delta_n^{\mathbf{K}}(\varphi_1 \wedge \varphi_2) = \delta_n^{\mathbf{C}}(\varphi_1 \wedge \varphi_2)$ for every n and hence $\lim_{n\to\infty} \delta_n^{\mathbf{C}}(\varphi_1 \wedge \varphi_2) = 1$. $\square$

Recall Definition 2.25 about l-colourings viewed as functions. The next step is to define, given an l-colourable structure $\mathcal{A}$ and l-colouring $\gamma : A - \mathrm{cl}(\emptyset) \to \{1, \dots, l\}$, a formula which describes which elements have the same colour (with respect to γ).

Definition 5.8. Suppose that $\mathcal{A}$ is an l-colourable L_{rel}-structure with universe $A = \{a_1, \dots, a_\alpha\}$ and that $\gamma : A - \mathrm{cl}(\emptyset) \to \{1, \dots, l\}$ is an l-colouring of $\mathcal{A}$. Then $\zeta_\gamma(x_1, \dots, x_\alpha)$ denotes the formula

$$\bigwedge_{\substack{1\le i\le\alpha\\ a_i\in\mathrm{cl}(\emptyset)}} \theta_0(x_i) \;\wedge \bigwedge_{\substack{1\le i,j\le\alpha\\ a_i,a_j\notin\mathrm{cl}(\emptyset)\\ \gamma(a_i)=\gamma(a_j)}} \xi(x_i, x_j) \;\wedge \bigwedge_{\substack{1\le i,j\le\alpha\\ a_i,a_j\notin\mathrm{cl}(\emptyset)\\ \gamma(a_i)\neq\gamma(a_j)}} \neg\xi(x_i, x_j).$$

Remark 5.9. Notice that if $\gamma : A - \mathrm{cl}(\emptyset)$ and $\gamma' : A - \mathrm{cl}(\emptyset)$ are such that for all $a, b \in A - \mathrm{cl}(\emptyset)$, $\gamma(a) = \gamma(b) \iff \gamma'(a) = \gamma'(b)$, then $\zeta_\gamma = \zeta_{\gamma'}$. This is because ξ only discerns which elements have the same colour and not which colour they have. We will use this in Lemma 5.11.

Recall Definition 2.26 of the characteristic formula of an l-coloured or l-colourable structure, with respect to an ordering of its universe.

Definition 5.10. (i) For every $n \in \mathbb{N}$, let $\eta_n(x_1, \dots, x_n)$ denote the L_{pre}-formula

$$\forall y \Big(\theta_n(x_1, \dots, x_n, y) \;\rightarrow\; \bigvee_{i=1}^{n} y = x_i \Big),$$

and note that, in any l-coloured, or l-colourable, structure, $\eta(x_1, \dots, x_n)$ expresses that $\{x_1, \dots, x_n\}$ is a closed set.
(ii) Suppose that $\mathcal{B}$ is an l-colourable L_{rel}-structure and that $\mathcal{A} \subsetneq \mathcal{B}$ is a closed substructure of $\mathcal{B}$. Let $A = \{a_1, \dots, a_\alpha\}$ and $B = \{a_1, \dots, a_\beta\}$, where $\beta > \alpha$. Moreover, suppose that γ' is an l-colouring of $\mathcal{B}$ which extends an l-colouring γ of $\mathcal{A}$, so $\gamma'{\restriction}A = \gamma$. We call the following sentence an ***instance of the l-colour compatible $\mathcal{B}/\mathcal{A}$-extension axiom***:

$$\begin{aligned}\forall x_1, \dots, x_\alpha \exists x_{\alpha+1}, \dots, x_\beta \Big(&\big[\chi_{\mathcal{A}}(x_1, \dots, x_\alpha) \;\wedge\; \zeta_\gamma(x_1, \dots, x_\alpha) \;\wedge\; \eta_\alpha(x_1, \dots, x_\alpha)\big] \\ &\longrightarrow \big[\chi_{\mathcal{B}}(x_1, \dots, x_\beta) \;\wedge\; \zeta_{\gamma'}(x_1, \dots, x_\beta) \;\wedge\; \eta_\beta(x_1, \dots, x_\beta)\big]\Big).\end{aligned}$$

There are only finitely many l-colourings of $\mathcal{B}$ and therefore there are only finitely many instances of the l-colour compatible $\mathcal{B}/\mathcal{A}$-extension axiom. We define the ***l-colour compatible $\mathcal{B}/\mathcal{A}$-extension axiom*** to be the conjunction of all these instances. A sentence φ is called an ***l-colour compatible extension axiom*** if it is the l-colour compatible $\mathcal{B}/\mathcal{A}$-extension axiom for some closed substructure $\mathcal{A} \subset \mathcal{B}$ where $\mathcal{B}$ is l-colourable.

Observe that every l-colour compatible extension axiom is an L_{rel}-sentence (so none of the symbols $P_1, \dots, P_l$ occurs in it). The next lemma shows that whenever $\mathcal{A} \subset \mathcal{B}$ are l-colour*able* structures and A is closed in $\mathcal{B}$, then there is k such that if $\mathcal{M}$ is an l-colour*ed* structure and has the k-extension property, then $\mathcal{M}$ satisfies the l-colour compatible $\mathcal{B}/\mathcal{A}$-extension axiom. As a corollary we will then get that, with $\delta_n^{\mathbf{C}}$-probability approaching 1 as n tends to infinity, a random $\mathcal{M} \in \mathbf{C}_n$ satisfies the l-colour compatible $\mathcal{B}/\mathcal{A}$-extension axiom.

Lemma 5.11. *Assume that $\mathcal{B}$ is an l-colourable L_{rel}-structure and $\mathcal{A} \subsetneq \mathcal{B}$ is a closed substructure of $\mathcal{B}$. Let $k = \max(k_0, \dim(\mathcal{B}))$. If $\mathcal{M}$ is an l-coloured structure and has the k-extension property, then $\mathcal{M}$ satisfies the l-colour compatible $\mathcal{B}/\mathcal{A}$-extension axiom.*

Proof. Let $\mathcal{B}$ be l-colourable L_{rel}-structures and let $\mathcal{A}$ is a closed substructure of $\mathcal{B}$. Assume that $A = \{a_1, \dots, a_\alpha\}$ and $B = \{a_1, \dots, a_\beta\}$ where $\beta > \alpha$ and let k be as in the lemma. It is enough to prove that if $\mathcal{M}$ is an l-coloured structure with the k-extension property, then $\mathcal{M}$ satisfies each instance of the l-colour compatible $\mathcal{B}/\mathcal{A}$-extension axiom. So assume that $\mathcal{M}$ is an l-coloured structure with the k-extension property and choose an arbitrary instance of the l-colour compatible $\mathcal{B}/A$-extension axiom, which uses a colouring $\gamma' : B - \mathrm{cl}(\emptyset) \to \{1, \dots, l\}$ and its restriction to a A, $\gamma = \gamma' \restriction (A - \mathrm{cl}(\emptyset))$. Assume that

$$\mathcal{M} \models \chi_{\mathcal{A}}(a'_1, \dots, a'_\alpha) \ \wedge \ \zeta_\gamma(a'_1, \dots, a'_\alpha) \ \wedge \ \eta_\alpha(a'_1, \dots, a'_\alpha)$$

for some $a'_1, \dots, a'_\alpha \in M$. Let $\mathcal{A}' = \mathcal{M} \restriction \{a'_1, \dots, a'_\alpha\}$. Then by the definition of $\chi_{\mathcal{A}}$ there is an isomorphism $f : \mathcal{A}' \restriction L_{rel} \to \mathcal{A}$ such that $f(a'_i) = a_i$ for $i = 1, \dots, \alpha$. For all $a'_i \in \{a'_1, \dots, a'_\alpha\} - \mathrm{cl}(\emptyset)$ and $j \in \{1, \dots, l\}$ let

$$\gamma_0(a'_i) = j \quad \iff \quad \mathcal{M} \models P_j(a'_i),$$

so γ_0 is an l-colouring of $\mathcal{A}' \restriction L_{rel}$. By the choice of k and since $\mathcal{M}$ has the k-extension property and $\mathcal{M} \models \zeta_\gamma(a'_1, \dots, a'_\alpha)$ it follows (using (2)) that, for all $a'_i, a'_j \in \{a'_1, \dots, a'_\alpha\} - cl(\emptyset)$,

$$\gamma_0(a'_i) = \gamma_0(a'_j) \quad \iff \quad \mathcal{M} \models \xi(a'_i, a'_j) \quad \iff \quad \gamma(a'_i) = \gamma(a'_j).$$

From this it follows that (by permuting the colours assigned by γ' if necessary) we can find a l-colouring γ'_1 of $\mathcal{B}$ such that if γ_1 is the restriction of γ'_1 to A, then for all $a_i \in \{a_1, \dots, a_\alpha\} - \mathrm{cl}(\emptyset)$ we have $\gamma_1(a_i) = \gamma_1(f(a'_i)) = \gamma_0(a'_i)$ and for all $a, b \in B - \mathrm{cl}(\emptyset)$,

$$\gamma'_1(a) = \gamma'_1(b) \quad \iff \quad \gamma'(a) = \gamma'(b).$$

Now expand $\mathcal{B}$ into an L-structure $\mathcal{B}^+$ by adding colours to it according to γ'_1, that is, if $b \in B - \mathrm{cl}(\emptyset)$ and $\gamma'_1(b) = i$ then let $\mathcal{B}^+ \models P_i(b)$. By the definition of $\mathcal{B}^+$ it follows that f^{-1} is an L-isomorphism from $\mathcal{B}^+ \restriction A$ onto $\mathcal{M} \restriction \{a'_1, \dots, a'_\alpha\}$. Since $\mathcal{M}$ has the k-extension property, we may extend f^{-1} into an embedding $g : \mathcal{B}^+ \to \mathcal{M}$, where the image of g is a closed subset of M. For every $i \in \{\alpha_1, \dots, \beta\}$, let $a'_i = g(a_i)$. Then

$$\mathcal{M} \models \chi_{\mathcal{B}}(a'_1, \dots, a'_\beta) \ \wedge \ \eta_\beta(a'_1, \dots, a'_\beta).$$

Moreover, as g is an L-embedding we have, for every $j \in \{1, \dots, l\}$ and every $a \in B - \mathrm{cl}(\emptyset)$, $\mathcal{M} \models P_j(g(a)) \iff \mathcal{B}^+ \models P_j(a)$. By the choice of γ'_1 it follows that, for all $i, j \in \{1, \dots, \beta\}$ such that $a'_i, a'_j \notin \mathrm{cl}(\emptyset)$, $\gamma'_1(a'_i) = \gamma'_1(a'_j) \iff \gamma'(a_i) = \gamma'(a_j)$. From the definition of $\zeta_{\gamma'}$ and the choice of k it follows that $\mathcal{M} \models \zeta_{\gamma'}(a'_1, \dots, a'_\beta)$. The chosen instance of the l-colour compatible $\mathcal{B}/\mathcal{A}$-extension axiom is hence satisfied by $\mathcal{M}$, and since it was an arbitrary instance, $\mathcal{M}$ has to satisfy all of the instances. $\square$

Now we can prove that every l-colour compatible extension axiom will almost surely be satisfied in an l-colour*able* structure.

Corollary 5.12. *For every l-colour compatible extension axiom φ, $\lim_{n\to\infty} \delta_n^{\mathbf{C}}(\varphi) = 1$.*

Proof. Let φ be an l-colour compatible extension axiom, so for some $\mathcal{A} \subseteq \mathcal{B}$ it is the l-colour compatible $\mathcal{B}/\mathcal{A}$-extension axiom. Let $k = \max(k_0, \dim(\mathcal{B}))$. By Lemma 5.11, for every n, if $\mathcal{M} \in \mathbf{K}_n$ has the k-extension property then $\mathcal{M} \models \varphi$. Hence, for every n,

$$\delta_n^{\mathbf{K}}\big(\{\mathcal{M} \in \mathbf{K}_n : \mathcal{M} \text{ has the } k\text{-extension property}\}\big) \ \leq \ \delta_n^{\mathbf{K}}(\varphi).$$

By Corollary 2.22 we get $\lim_{n\to\infty} \delta_n^{\mathbf{K}}(\varphi) = 1$, so it suffices to show that $\delta_n^{\mathbf{C}}(\varphi) = \delta_n^{\mathbf{K}}(\varphi)$ for all n. But, as in the proof of Lemma 5.7, this follows from the definition of $\delta_n^{\mathbf{C}}$ and the fact that φ is an L_{rel}-sentence. $\square$

The part T_{ext} of the axiomatisation $T_{\mathbf{C}}$ consists, by definition, of all l-colour compatible extension axioms. The axiomatisation $T_{\mathbf{C}}$ also needs to express that the formulas θ_n, $n \in \mathbb{N}$, from Assumption 2.12 define a pregeometry in every model of $T_{\mathbf{C}}$. This is the purpose of the part of $T_{\mathbf{C}}$ which we denote T_{pre}. More specifically, by using the formulas from Assumption 2.12 (2), we can express, with an infinite set T_{pre}, of L_{pre}-sentences, properties (1)–(3) of pregeometries in Definition 2.1 for *finite* sets. In particular, since the closure of a set A should not depend on how we order A, T_{pre} contains, for each $n \geq 1$ and each permutation π of $\{1, \ldots, n\}$, the sentence

$$\forall x_1, \ldots, x_{n+1}\big(\theta_n(x_1, \ldots, x_n, x_{n+1}) \ \longleftrightarrow \ \theta_n(x_{\pi(1)}, \ldots, x_{\pi(n)}, x_{n+1})\big).$$

Lemma 5.13. *Let $\mathcal{M}$ be an L_{rel}-structure such that $\mathcal{M} \models T_{pre}$. Define a closure operator $\mathrm{cl}_{\mathcal{M}}$ as follows:*

(a) For every $n \in \mathbb{N}$ and all $a_1, \ldots, a_{n+1} \in M$, $a_{n+1} \in \mathrm{cl}_{\mathcal{M}}(a_1, \ldots, a_n)$ if and only if $\mathcal{M} \models \theta_n(a_1, \ldots, a_{n+1})$.

(b) For every $A \subseteq M$ and every $a \in M$, $a \in \mathrm{cl}_{\mathcal{M}}(A)$ if and only if there is a finite $A' \subseteq A$ such that $a \in \mathrm{cl}_{\mathcal{M}}(A')$.

Then $(M, \mathrm{cl}_{\mathcal{M}})$ is a pregeometry.

Proof. Suppose that $\mathcal{M} \models T_{pre}$ and let $\mathrm{cl}_{\mathcal{M}}$ be defined by (a) and (b). From (b) it follows that $\mathrm{cl}_{\mathcal{M}}$ has the finiteness property (4) of Definition 2.1 of a pregeometry. From the definition of T_{pre} and (a) it follows that $\mathrm{cl}_{\mathcal{M}}$ has properties (1)–(3) of Definition 2.1 of a pregeometry. Hence, $(M, \mathrm{cl}_{\mathcal{M}})$ is a pregeometry. $\square$

The fourth part of the aximatisation $T_{\mathbf{C}}$, denoted T_{iso}, will express that "every closed finite substructure is l-colourable". For each $n \in \mathbb{N}$ let $\mathcal{M}_{n,1}, \ldots, \mathcal{M}_{n,m_n}$ be an enumeration of all (finitely many) members of $\mathbf{C}_n$, and recall that $\chi_{\mathcal{M}_{n,i}}$ denotes the characteristic formula of $\mathcal{M}_{n,i}$ (see Definition 2.26). Recall that for every n, all structures in $\mathbf{C}_n$ have the same universe (in fact their reduct to L_{pre} is the same). Let $s(n)$ be the cardinality of (the universe of) a structure in $\mathbf{C}_n$. For every $n \in \mathbb{N}$, there is an L_{rel}-sentence ψ_n which expresses that every closed substructure of cardinality $s(n)$ is isomorphic to one of $\mathcal{M}_{n,1}, \ldots, \mathcal{M}_{n,m_n}$. More precisely, we let ψ_n be the L_{rel}-sentence

$$\forall x_1, \ldots, x_{s(n)}\Bigg(\Bigg[\bigwedge_{i\neq j} x_i \neq x_j \ \wedge \ \forall y\Bigg(\theta_{s(n)}(x_1, \ldots, x_{s(n)}, y) \ \rightarrow \ \bigvee_{i=1}^{s(n)} y = x_i\Bigg)\Bigg] \\ \longrightarrow \ \bigvee_{i=1}^{s(n)} \bigvee_{\pi} \chi_{M_{n,i}}(x_{\pi(1)}, \ldots, x_{\pi(s(n))})\Bigg),$$

where the disjunction '$\bigvee_\pi$' ranges over all permutations π of $\{1, \ldots, s(n)\}$. Let $T_{iso} = \{\psi_n : n \in \mathbb{N}\}$. Recall that $T_\xi = \{\varphi_1 \wedge \varphi_2\}$, where φ_1 and φ_2 were defined in Definition 5.6, that T_{ext} is the set of all l-colour compatible extension axioms and that T_{pre} was defined in the paragraph before Lemma 5.13. Now we let

$$T_{\mathbf{C}} \;=\; T_{pre} \cup T_\xi \cup T_{ext} \cup T_{iso}.$$

Notice that $T_{\mathbf{C}}$ contains only L_{rel}-sentences.

Lemma 5.14. *$T_{\mathbf{C}}$ is consistent and countably categorical, hence complete.*

Proof. From the definitions of T_{pre} and T_{iso} it follows that every l-colourable structure is a model of $T_{pre} \cup T_{iso}$. By compactness, Corollary 5.12 and Lemma 5.7, it follows that $T_{\mathbf{C}}$ is consistent.

We now prove that $T_{\mathbf{C}}$ is countably categorical. Assume that $\mathcal{M}$ and $\mathcal{M}'$ are L_{rel}-structures such that $\mathcal{M} \models T_{\mathbf{C}}$, $\mathcal{M}' \models T_{\mathbf{C}}$ and $|M| = |M'| = \aleph_0$. Since $\mathcal{M}, \mathcal{M}' \models T_{pre}$ it follows from Lemma 5.13 that if $\mathrm{cl}_{\mathcal{M}}$ is defined by saying that $a \in \mathrm{cl}_{\mathcal{M}}(A)$ if and only if there are m and $b_1, \ldots, b_m \in A$ such that $\mathcal{M} \models \theta_m(b_1, \ldots, b_m, a)$, then $(M, \mathrm{cl}_{\mathcal{M}})$ is a pregeometry; and similarly for $\mathcal{M}'$. By a back and forth argument we will build partial isomorphisms between $\mathcal{M}$ and $\mathcal{M}'$ such that each new one extends the former ones. The union of these partial isomorphisms shows that $\mathcal{M} \cong \mathcal{M}'$. The main part of the argument is to prove the following:

Claim. Let $\mathcal{A} \subseteq \mathcal{M}$, $\mathcal{A}' \subseteq \mathcal{M}'$ be finite closed substructures (so $\mathrm{cl}_{\mathcal{M}}(\emptyset) \subseteq A$ and $\mathrm{cl}_{\mathcal{M}'}(\emptyset) \subseteq A'$) and suppose that there is an isomorphism $f : \mathcal{A} \to \mathcal{A}'$ such that for all $a, b \in A - \mathrm{cl}_{\mathcal{M}}(\emptyset)$ we have $\mathcal{M} \models \xi(a, b) \iff \mathcal{M}' \models \xi(f(a), f(b))$. Then for every $c \in M - A$ ($d \in M' - A'$) there exist a closed subtructure $\mathcal{B}' \subseteq \mathcal{M}'$ ($\mathcal{B} \subseteq \mathcal{M}$) and an isomorphism $g : \mathcal{M} \restriction \mathrm{cl}_{\mathcal{M}}(A \cup \{c\}) \to \mathcal{B}'$ ($g : \mathcal{B} \to \mathcal{M}' \restriction \mathrm{cl}_{\mathcal{M}'}(A \cup \{d\})$) such that $\mathcal{A}' \subseteq \mathcal{B}'$ ($\mathcal{A} \subseteq \mathcal{B}$), g extends f and for all $a, b \in \mathrm{cl}_{\mathcal{M}}(A \cup \{c\}) - \mathrm{cl}_{\mathcal{M}}(\emptyset)$ ($a, b \in B - \mathrm{cl}_{\mathcal{M}}(\emptyset)$) we have $\mathcal{M} \models \xi(a, b) \iff \mathcal{M}' \models \xi(g(a), g(b))$

Proof of the claim. Let $\mathcal{A} \subseteq \mathcal{M}$, $\mathcal{A}' \subseteq \mathcal{M}'$ be finite closed substructures and $f : \mathcal{A} \to \mathcal{A}'$ an isomorphism such that for all $a, b \in A - \mathrm{cl}_{\mathcal{M}}(\emptyset)$ we have $\mathcal{M} \models \xi(a, b) \iff \mathcal{M}' \models \xi(f(a), f(b))$. Let $A = \{a_1, \ldots, a_\alpha\}$ and $A' = \{f(a_1), \ldots, f(a_\alpha)\}$. Suppose that $c \in M - A$ and let $B = \mathrm{cl}_{\mathcal{M}}(A \cup \{c\})$. (The case when $d \in M' - A'$ is proved in the same way.) Enumerate B as $B = \{b_1, \ldots, b_\beta\}$ so that $b_i = a_i$ for $1 \leq i \leq \alpha$. Note that we must have $\beta > \alpha$. Let $\mathcal{B} = \mathcal{M} \restriction B$.

Since $\mathcal{M} \models T_{\mathbf{C}}$ we have $\mathcal{M} \models \varphi_1 \wedge \varphi_2$, which implies that

(i) $\xi(x, y)$ defines an equivalence relation on $M - \mathrm{cl}_{\mathcal{M}}(\emptyset)$ which we denote $\sim_\xi$,

(ii) there is a substructure of $\mathcal{M}$ which is isomorphic to $\mathcal{U}$ (from Definition 5.6), and for notational simplicity we denote it by $\mathcal{U}$, so $\mathcal{U} \subseteq \mathcal{M}$, and there are $u_1, \ldots, u_l \in U - \mathrm{cl}_{\mathcal{M}}(\emptyset)$ such that if $1 \leq i < j \leq l$ then $u_i \not\sim_\xi u_j$, and

(iii) for all $a \in M - \mathrm{cl}_{\mathcal{M}}(\emptyset)$ there is $j \in \{1, \ldots, l\}$ such that $a \sim_\xi u_j$.

It follows that the restriction of $\sim_\xi$ to $(U \cup B) - \mathrm{cl}_{\mathcal{M}}(\emptyset)$ has exactly l equivalence classes. By assumption, $\xi(x, y)$ is an existential formula. As $U \cup B$ is finite, it follows that there is a *finite* closed substructure $\mathcal{N} \subseteq \mathcal{M}$ such that $U \cup B \subseteq N$ and for all $a, b \in U \cup B$ we have $\mathcal{M} \models \xi(a, b)$ if and only if $\mathcal{N} \models \xi(a, b)$. Consequently:

$$\text{for all } a, b \in (U \cup B) - \mathrm{cl}_{\mathcal{M}}(\emptyset), \;\; a \sim_\xi b \iff \mathcal{N} \models \xi(a, b).$$

Since $\mathcal{M} \models T_{iso}$ we know that $\mathcal{N}$ is an l-colourable structure. Let $\gamma : N - \mathrm{cl}_{\mathcal{N}}(\emptyset) \to \{1, \ldots, l\}$ be an l-colouring of $\mathcal{B}$. Define an equivalence relation $\sim_\gamma$ on $N - \mathrm{cl}_{\mathcal{N}}(\emptyset)$ by

$$a \sim_\gamma b \iff \gamma(a) = \gamma(b) \text{ for all } a, b \in N - \mathrm{cl}_{\mathcal{N}}(\emptyset).$$

Since $\mathcal{U}$ cannot (by its definition) be l'-coloured if $l' < l$, it follows that $\sim_\gamma$ has exactly l equivalence classes. In fact, the restriction of $\sim_\gamma$ to $(U \cup B) - \mathrm{cl}_{\mathcal{N}}(\emptyset)$ has exactly l equivalence classes.

Let $a, b \in (U \cup B) - \mathrm{cl}_{\mathcal{M}}(\emptyset)$ and suppose that $a \sim_\xi b$. Then $\mathcal{M} \models \xi(a, b)$ and by the choice of $\mathcal{N}$ we also get $\mathcal{N} \models \xi(a, b)$. By (1) we must have $\gamma(a) = \gamma(b)$, so $a \sim_\gamma b$. Hence, the restriction of $\sim_\xi$ to $(U \cup B) - \mathrm{cl}_{\mathcal{M}}(\emptyset)$ is a refinement of the restriction of $\sim_\gamma$ to $(U \cup B) - \mathrm{cl}_{\mathcal{M}}(\emptyset)$. As both equivalence relations restricted to $(U \cup B) - \mathrm{cl}_{\mathcal{M}}(\emptyset)$ have exactly l-equivalence classes, it follows that the must coincide. In other words, for all $a, b \in (U \cup B) - \mathrm{cl}_{\mathcal{M}}(\emptyset)$ we have $a \sim_\xi b$ if and only if $a \sim_\gamma b$, so

(iv) for all $a, b \in B - \mathrm{cl}_{\mathcal{M}}(\emptyset)$, $\gamma(a) = \gamma(b) \iff \mathcal{M} \models \xi(a, b)$.

Let $\gamma_B = \gamma\restriction B$ and $\gamma_A = \gamma\restriction A$. Since $\chi_{\mathcal{A}}(x_1, \ldots, x_\alpha)$ is the characteristic formula of $\mathcal{A}$ we have $\mathcal{M} \models \chi_{\mathcal{A}}(a_1, \ldots, a_\alpha)$. From (iv) it follows that $\mathcal{M} \models \zeta_{\gamma_A}(a_1, \ldots, a_\alpha)$ (see Definition 5.8). The assumptions of the claim now imply that

$$\mathcal{M}' \models \chi_{\mathcal{A}}(f(a_1), \ldots, f(a_\alpha)) \ \wedge \ \zeta_{\gamma_A}(f(a_1), \ldots, f(a_\alpha)) \ \wedge \ \eta_\alpha(f(a_1), \ldots, f(a_\alpha)).$$

Since $\mathcal{M}' \models T_{ext}$ it follows, in particular, that $\mathcal{M}'$ satisfies the following instance of the l-colour compatible $\mathcal{B}/\mathcal{A}$-extension axiom:

$$\begin{aligned} \forall x_1, \ldots, x_\alpha \exists y_1, \ldots, y_\beta \Big(&\big[\chi_{\mathcal{A}}(x_1, \ldots, x_\alpha) \ \wedge \ \zeta_{\gamma_A}(x_1, \ldots, x_\alpha) \ \wedge \ \eta_\alpha(x_1, \ldots, x_\alpha)\big] \\ &\longrightarrow \big[\chi_{\mathcal{B}}(x_1, \ldots, x_\beta) \ \wedge \ \zeta_{\gamma_B}(x_1, \ldots, x_\beta) \ \wedge \ \eta_\beta(x_1, \ldots, x_\beta)\big]\Big). \end{aligned}$$

It follows that there are closed substructure $\mathcal{B}' \subseteq \mathcal{M}'$ such that $\mathcal{A}' \subseteq B'$ and an isomorphism $g : \mathcal{B} \to \mathcal{B}'$ which extends f with the property that for all $a, b \in B - \mathrm{cl}(\emptyset) = \mathrm{cl}_{\mathcal{M}}(A \cup \{c\}) - \mathrm{cl}_{\mathcal{M}}(\emptyset)$, $\mathcal{M} \models \xi(a, b)$ if and only if $\mathcal{M}' \models \xi(g(a), g(b))$. □

Continuation of the proof of the lemma. By the claim, it now suffices to prove that the assumptions of the claim hold for at least one pair, $\mathcal{A}, \mathcal{A}'$, such that $\mathcal{A} \subseteq \mathcal{M}$ and $\mathcal{A}' \subseteq \mathcal{M}'$ are closed substructures of the respective superstructure. We claim that this holds for $\mathcal{A} = \mathcal{M}\restriction\mathrm{cl}_{\mathcal{M}}(\emptyset)$ and $\mathcal{A}' = \mathcal{M}'\restriction\mathrm{cl}_{\mathcal{M}'}(\emptyset)$. Indeed, by the definition of l-colourable structure, Assumption 2.12 and $\mathcal{M}, \mathcal{M}' \models T_{iso}$, it follows that with $\mathcal{A}$ and $\mathcal{A}'$ defined in this way we have $\mathcal{A} \cong \mathcal{A}'$. Since $A - \mathrm{cl}_{\mathcal{M}}(\emptyset) = \emptyset$ the preservation of ξ for $a, b \in A - \mathrm{cl}_{\mathcal{M}}(\emptyset)$ is trivially satisfied. This concludes the proof of Lemma 5.14, and hence also of Theorem 5.1. □

Errata for [8]. Here we mention some missed assumptions that should be added to some results of [8].

(i) Theorems 7.31, 7.32 and 7.34 in [8] needs part (5) from Assumption 2.12 in this article. (This assumption could most conveniently be added to Assumption 7.10 in [8].) This assumption is implicitly used in the proof of Lemma 8.5 in [8] and guaratees that if $\mathcal{A}$ is a closed substructure of a represented structure (see Definition 7.4 [8]) $\mathcal{M}$ and $B \subseteq A$, then the closure of B is the same whether computed in $\mathcal{A}$ or in $\mathcal{M}$.

(ii) Theorem 7.32 in [8] needs the following two additional assumptions (implicitly made in the proof of that theorem):

(a) There is, up to isomorphism, a unique represented structure (Definition 7.4 in [8]) with dimension 0. (Note that by the definitions in this article, there is a unique, up to isomorphism, (strongly) l-coloured structure with dimension 0.) This assumption is needed in the proof of Lemma 8.12 in [8]. Without it, one only gets a limit law (convergence, but not necessarily to 0 or 1), since one gets a distinct "limit theory" for every represented isomorphism type of dimension 0.

(b) For every $n \in \mathbb{N}$, there is a "characteristic" L_{pre}-formula $\chi_{\mathcal{G}_n}(x_1, \ldots, x_{m_n})$ of $\mathcal{G}_n$, where $m_n = |G_n|$, such that if $\mathcal{A}$ is an L_{pre}-structure in which the formulas θ_n define a pregeometry and $\mathcal{A} \models \chi_{\mathcal{G}_n}(a_1, \ldots, a_s)$ for some enumeration $a_1, \ldots, a_s$ of A, then $\mathcal{A} \cong \mathcal{G}_n$. This is (7) of Assumption 2.12 in this article (except for omitting here the requirement that $\chi_{\mathcal{G}_n}$ is quantifier-free), and it holds for the examples of pregeometries (and corresponding languages) considered in [8] and in this article. But it is not a consequence of the other assumptions made (in Assumption 7.3 and Assumption 7.10 in [8]), so it needs to be added.

These remarks affect *only* Sections 7–8 in [8], because the other sections do *not* consider (nontrivial) pregemetries.

References

[1] J. Balogh, D. Mubayi, Almost all triangle-free triple systems are tripartite, *Combinatorica*, Vol. 32 (2012) 143–169.

[2] M. Djordjević, The finite submodel property and ω-categorical expansions of pregeometries, *Annals of Pure and Applied Logic*, Vol. 139 (2006), 201-229.

[3] H-D. Ebbinghaus, J. Flum, *Finite Model Theory*, Second Edition, Springer-Verlag (1999).

[4] P. Erdös, D. J. Kleitman, B. L. Rothschild, Asymptotic enumeration of K_n-free graphs, *International Colloquium on Combinatorial Theory. Atti dei Convegni Lincei* 17, Vol. 2, Rome (1976) 19–27.

[5] R. L. Graham, K. Leeb, B. L. Rothschild, Ramsey's theorem for a class of categories, *Advances in Mathematics*, Vol. 8 (1972) 417-433.

[6] R. L. Graham, B. L. Rothschild, J. H. Spencer, *Ramsey Theory*, John Wiley & Sons Inc. (1980).

[7] Ph. G. Kolaitis, H. J. Prömel, B. L. Rothschild, K_{l+1}-free graphs: asymptotic structure and a 0-1 law, *Transactions of The American Mathematical Society*, Vol. 303 (1987) 637–671.

[8] V. Koponen, Asymptotic probabilities of extension properties and random l-colourable structures, *Annals of Pure and Applied Logic*, Vol. 163 (2012) 391-438.

[9] D. Marker, *Model Theory: An Introduction*, Springer-Verlag (2002).

[10] J. Oxley, *Matroid Theory*, Second Edition, Oxford University Press (2011).

[11] Y. Person, M. Schacht, Almost all hypergraphs without Fano planes are bipartite, in Claire Mathieu (editor), *Proceedings of the Twentieth Annual ACM-SIAM Symposium on Discrete Algorithms (SODA 09)*, 217–226, ACM Press (2009).