Testing Hereditary Properties of Ordered Graphs and Matrices

Noga Alon[*], Omri Ben-Eliezer[†], Eldar Fischer[‡]

April 6, 2017

Abstract

We consider properties of edge-colored *vertex-ordered graphs*, i.e., graphs with a totally ordered vertex set and a finite set of possible edge colors. We show that any hereditary property of such graphs is strongly testable, i.e., testable with a constant number of queries. We also explain how the proof can be adapted to show that any hereditary property of 2-dimensional matrices over a finite alphabet (where row and column order is not ignored) is strongly testable. The first result generalizes the result of Alon and Shapira [FOCS'05; SICOMP'08], who showed that any hereditary graph property (without vertex order) is strongly testable. The second result answers and generalizes a conjecture of Alon, Fischer and Newman [SICOMP'07] concerning testing of matrix properties.

The testability is proved by establishing a removal lemma for vertex-ordered graphs. It states that for any finite or infinite family $\mathcal{F}$ of forbidden vertex-ordered graphs, and any $\epsilon > 0$, there exist $\delta > 0$ and k so that any vertex-ordered graph which is ϵ-far from being $\mathcal{F}$-free contains at least $\delta n^{|F|}$ copies of some $F \in \mathcal{F}$ (with the correct vertex order) where $|F| \leq k$. The proof bridges the gap between techniques related to the regularity lemma, used in the long chain of papers investigating graph testing, and string testing techniques. Along the way we develop a Ramsey-type lemma for k-partite graphs with "undesirable" edges, stating that one can find a Ramsey-type structure in such a graph, in which the density of the undesirable edges is not much higher than the density of those edges in the graph.

1 Introduction

Property Testing is dedicated to finding fast algorithms for decision problems of the following type: Given a combinatorial structure S, distinguish quickly between the case where S satisfies a property $\mathcal{P}$ and the case where S is *far* from satisfying the property. Being far means that one needs to modify a significant fraction of the data in S to make it satisfy $\mathcal{P}$. Property Testing was first formally defined by Rubinfeld and Sudan [43], and the investigation in the combinatorial context was initiated by Goldreich, Goldwasser and Ron [31]. This area has been very active over the last twenty years, see, e.g. [30] for various surveys on it.

In this paper we focus on property testing of two-dimensional structures over a finite alphabet, or equivalently, two-variable functions with a fixed finite range. Specifically, we consider graphs and matrices. *Graphs* are functions $G : \binom{V}{2} \to \{0,1\}$ where V is the vertex set; more generally *edge-colored graphs* (with finite color set Σ) are functions $G : \binom{V}{2} \to \Sigma$. *Matrices* over a finite alphabet Σ (or *images*) are functions $M : U \times V \to \Sigma$. In this paper we generally consider edge-colored graphs rather than standard graphs, as the added generality will prove useful later, so the term *graph* usually refers to an edge-colored graph.

For a fixed finite set Σ, a *property* of functions over Σ is simply a collection of functions whose range is Σ. Specifically, an *ordered graph property* is a collection of (edge-colored) graphs $G : \binom{V}{2} \to \Sigma$. An *unordered graph property* is an ordered graph property that is also *invariant under vertex permutations*: If $G \in \mathcal{P}$ and π is any permutation on V_G, then the graph G^π, defined by $G^\pi(\pi(u)\pi(v)) = G(uv)$ for any $u \neq v \in V_G$, satisfies $G^\pi \in \mathcal{P}$. Similarly, an *(ordered) matrix property*, or an *image property*, is a collection of functions $M : [m] \times [n] \to \Sigma$. For simplicity, most definitions given below are only stated for graphs, but they carry over naturally to matrices.

A graph $G : \binom{[n]}{2} \to \Sigma$ is ϵ-far from the property $\mathcal{P}$ if one needs to modify the value $G(ij)$ for at least $\epsilon\binom{n}{2}$ of the edges ij, where ij denotes the (unordered) edge $\{i,j\} \in \binom{[n]}{2}$. A *tester* for the property $\mathcal{P}$ is a randomized algorithm that is given a parameter $\epsilon > 0$ and query access to its input graph G. The tester must distinguish, with error probability at most $1/3$, between the case where G satisfies $\mathcal{P}$ and the case where G is ϵ-far from satisfying $\mathcal{P}$. The tester is said to have *one-sided error* if it always accepts inputs from $\mathcal{P}$, and rejects inputs that are ϵ-far from $\mathcal{P}$ with probability at least $2/3$. It is desirable to obtain testers that are efficient in terms of the *query complexity* (i.e. the maximal possible number of queries made by the tester). A property $\mathcal{P}$ is *strongly testable* if there is a one-sided error tester for $\mathcal{P}$ whose query complexity is bounded by a function $Q(\mathcal{P}, \epsilon)$. In other words, the query complexity of the tester is independent of the size of the input.

From now on, we generally assume (unless it is explicitly stated that we consider unordered graphs) that the vertex set V of a graph G has a total ordering (e.g. the natural one for $V = [n]$), which we denote by $<$. The (induced) *ordered subgraph* of the graph $G : \binom{V}{2} \to \Sigma$ on $U \subseteq V$, where the elements of U are $u_1 < \ldots < u_k$, is the graph $H : [k] \to \Sigma$ which satisfies $H(ij) = G(u_i u_j)$ for any $i < j \in [k]$. For a family $\mathcal{F}$ of "forbidden" graphs, the property $\mathcal{P}_\mathcal{F}$ of $\mathcal{F}$-*freeness* consists of all graphs G for which any ordered subgraph H of G satisfies $H \notin \mathcal{F}$. Finally, a property $\mathcal{P}$ is *hereditary* if it is closed under taking induced subgraphs. That is, for any $G \in \mathcal{P}$ and any ordered subgraph H of G, it holds that $H \in \mathcal{P}$. Note that a property $\mathcal{P}$ is hereditary if and only if $\mathcal{P} = \mathcal{P}_\mathcal{F}$ for some (finite or infinite) family $\mathcal{F}$ of graphs over Σ.

The analogous notions of ordered subgraphs, $\mathcal{F}$-freeness and hereditary properties for matrices are "structure preserving". Here, the *ordered submatrix* of the matrix $M : [m] \times [n] \to \Sigma$ on $A \times B$, where the elements of A and B are $a_1 < \ldots < a_k$ and $b_1 < \ldots < b_l$, is the matrix $N : [k] \times [l] \to \Sigma$ defined by $N(i,j) = M(a_i, b_j)$ for any $i \in [k]$ and $j \in [l]$.

1.1 Previous results on graphs and matrices

Some of the most interesting results in property testing have been those that identify large families of properties that are efficiently testable, and those that show that large families of properties cannot be tested efficiently.

One of the most widely investigated questions in property testing has been that of characterizing the efficiently testable unordered graph properties. In the seminal paper of Goldreich, Goldwasser and Ron [31] it was shown that all unordered graph properties that can be represented by a certain graph partitioning, including properties such as k-colorability and having a large clique, are strongly testable. See also [33]. Alon, Fischer, Krivelevich and Szegedy [4] showed that the property of $\mathcal{F}$-freeness is strongly testable for any finite family $\mathcal{F}$ of forbidden unordered graphs (here the term *unordered graphs* refers to the usual notion of graphs with no order on the vertices). Their main technical result, now known as the *induced graph removal lemma*, is a generalization of the well-known *graph removal lemma* [3, 45].

Theorem 1.1 (Induced graph removal lemma [4]) *For any finite family $\mathcal{F}$ of unordered graphs and $\epsilon > 0$ there exists $\delta = \delta(\mathcal{F}, \epsilon) > 0$, such that any graph G which is ϵ-far from $\mathcal{F}$-freeness contains at least δn^q copies of some $F \in \mathcal{F}$ with q vertices.*

The original proof of Theorem 1.1 uses a strengthening of the celebrated Szemerédi graph regularity lemma [45], known as the *strong graph regularity lemma.*

It is clear that having a removal lemma for a family $\mathcal{F}$ immediately implies that $\mathcal{F}$-freeness is strongly testable: A simple tester which picks a subgraph H whose size depends only on $\mathcal{F}$ and ϵ, and checks whether H contains graphs from $\mathcal{F}$ or not, is a valid one-sided tester for $\mathcal{F}$-freeness. Hence, removal lemmas have a major role in property testing. They also have implications in different areas of mathematics, such as number theory and discrete geometry. For more details, see the survey of Conlon and Fox [21].

By proving a variant of the induced graph removal lemma that also holds for infinite families, Alon and Shapira [8] generalized the results of [4]. The infinite variant is as follows.

Theorem 1.2 (Infinite graph removal lemma [8]) *For any finite or infinite family $\mathcal{F}$ of unordered graphs and $\epsilon > 0$ there exist $\delta = \delta(\mathcal{F}, \epsilon) > 0$ and $q_0 = q_0(\mathcal{F}, \epsilon)$, such that any graph G which is ϵ-far from $\mathcal{F}$-freeness contains at least δn^q copies of some $F \in \mathcal{F}$ on $q \leq q_0$ vertices.*

Theorem 1.2 directly implies that *any* hereditary unordered graph property is strongly testable, exhibiting the remarkable strength of property testing.

Theorem 1.3 (Hereditary graph properties are strongly testable [8]) *Let Σ be a finite set with $|\Sigma| \geq 2$. Any hereditary unordered graph property over Σ is strongly testable.*

Alon, Fischer, Newman and Shapira later presented [6] a complete combinatorial characterization of the graph properties that are testable (with two-sided error) using a constant number of queries, building on results from [26, 33]. Independently, Borgs, Chayes, Lovász, Sós, Szegedy and Vesztergombi obtained an analytic characterization of such properties through the theory of graph limits [15]. See also [38, 37].

An efficient finite induced removal lemma for binary matrices with no row and column order was obtained by Alon, Fischer and Newman [5]. In this case, δ^{-1} is polynomial in ϵ^{-1} (where ϵ, δ play the same roles as in the above removal lemmas). It was later shown by Fischer and Rozenberg [27] that when the alphabet is bigger than binary, the dependence of δ^{-1} on ϵ^{-1} is super-polynomial in general, and in fact testing submatrix-freeness over a non-binary alphabet is at least as hard as testing triangle-freeness in graphs, for which the dependence is also known to be super-polynomial

in general [1], see also [2]. Actually, the main tool in [5] is an efficient conditional regularity lemma for ordered binary matrices, and it was conjectured there that this regularity lemma can be used to obtain a removal lemma for ordered binary matrices.

Conjecture 1.4 (Ordered binary matrix removal lemma [5]) *For any finite family $\mathcal{F}$ of ordered binary matrices and any $\epsilon > 0$ there exists $\delta = \delta(\mathcal{F}, \epsilon)$ such that any $n \times n$ binary matrix which is ϵ-far from $\mathcal{F}$-freeness contains at least δn^{a+b} copies of some $a \times b$ matrix from $\mathcal{F}$.*

In contrast to the abundance of general testing results for two-dimensional structures with an inherent symmetry, such as unordered graphs and matrices, no similar results for *ordered* two-dimensional structures (i.e. structures that do not have any underlying symmetry) have been established. Even seemingly simple special cases, such as F-freeness for a single ordered graph F, or M-freeness for a single 2×2 ordered matrix M, are not known to be strongly testable in general [2]. A good survey on the role of symmetry in property testing is given by Sudan [46], who suggests that the successful characterization of the strongly testable unordered graph properties is attributable to the underlying symmetry of these properties. See also [32].

Despite the lack of general results as above for the ordered case, property testing of multi-dimensional ordered structures has recently been an active area of research. Notable examples of properties that were investigated in the setting of ordered matrices include monotonicity (see, e.g., [18, 19] for some of the recent works in the matrix setting), extensions of monotonicity such as k-monotonicity [17] and more generally poset properties [25], visual and geometric properties of images, such as connectedness, convexity, being a half plane [42, 13] and being a Lipschitz function [10, 14], and local properties, such as consecutive pattern-freeness [12]. Ordered graphs were less investigated in the context of property testing, but are the subject of many works in Combinatorics and other areas. See, e.g., a recent work on Ramsey-type questions in the ordered setting [22], in which it is shown that Ramsey numbers of simple ordered structures might differ significantly from their unordered counterparts.

Finally, we mention a relevant result on *one-dimensional* structures. Alon, Krivelevich, Newman and Szegedy [7] showed that regular languages are strongly testable. One can combine this result with the well-known Higman's lemma in order theory [35] to show that any hereditary property of words (i.e. one dimensional functions) over a finite alphabet is strongly testable.

1.2 Our contributions

We prove generalizations of Theorems 1.3 and 1.2 to the ordered setting, as well as analogous results for matrices. The following result generalizes Theorem 1.3.

Theorem 1.5 (Hereditary properties of ordered graphs are strongly testable) *Fix a finite set Σ with $|\Sigma| \geq 2$. Any hereditary ordered graph property over Σ is strongly testable.*

To prove Theorem 1.5, we establish an *order-preserving* induced graph removal lemma, which holds for finite and infinite families of ordered graphs. This is a generalization of Theorem 1.2.

Theorem 1.6 (Infinite ordered graph removal lemma) *Fix a finite set Σ with $|\Sigma| \geq 2$. For any (finite or infinite) family $\mathcal{F}$ of ordered graphs $F : \binom{[n_F]}{2} \to \Sigma$ and any $\epsilon > 0$ there exist $q_0 = q_0(\mathcal{F}, \epsilon)$ and $\delta = \delta(\mathcal{F}, \epsilon) > 0$, such that any ordered graph $G : \binom{[n]}{2} \to \Sigma$ that is ϵ-far from $\mathcal{F}$-freeness contains at least δn^q induced copies of some graph $F \in \mathcal{F}$ on $q \leq q_0$ vertices.*

An analogue of Theorem 1.5 for matrices is also proved.

Theorem 1.7 (Hereditary properties of ordered matrices are strongly testable) *Fix a finite set Σ with $|\Sigma| \geq 2$. Any hereditary (ordered) matrix property over Σ is strongly testable.*

As in the case of ordered graphs, to prove Theorem 1.7 we establish the following ordered matrix removal lemma, which holds for finite and infinite families of matrices, and settles a generalized form of Conjecture 1.4.

Theorem 1.8 (Infinite ordered matrix removal lemma) *Fix a finite set Σ with $|\Sigma| \geq 2$. For any (finite or infinite) family $\mathcal{F}$ of ordered matrices over Σ and any $\epsilon > 0$ there exist $q_0 = q_0(\mathcal{F}, \epsilon) > 0$ and $\delta = \delta(\mathcal{F}, \epsilon) > 0$, such that any ordered matrix over Σ that is ϵ-far from $\mathcal{F}$-freeness contains at least $\delta n^{q+q'}$ copies of some $q \times q'$ matrix $F \in \mathcal{F}$, where $q, q' \leq q_0$.*

Actually, the proof of Theorem 1.8 is almost identical to that of Theorem 1.6, so we only describe what modifications are needed to make the proof of Theorem 1.6 also work here, for the case of square matrices. However, all proofs can be adapted to the non-square case as well. An outline for the proof of the graph case is given in Section 2, and all of the sections after it are dedicated to the full proof. The needed modifications for the matrix case appear in Subsection 7.3.

To the best of our knowledge, Theorems 1.5 and 1.7 are the first known testing results of this type for ordered two-dimensional structures, and Theorems 1.6 and 1.8 are the first known order-preserving removal lemmas for two-dimensional structures.

It is interesting to note that some of the properties mentioned in Subsection 1.1, such as monotonicity, k-monotonicity, and forbidden-poset type properties in matrices, are hereditary (as all of them can be characterized by a finite set of forbidden submatrices), so Theorem 1.7 gives a new proof that these properties, and many of their natural extensions, are strongly testable. Naturally, our general testers are much less efficient than the testers specifically tailored for each of these properties (in terms of dependence of the underlying constants on the parameters of the problem), but the advantage of our result is its generality, that is, the fact that it applies to any hereditary property. Thus, for example, for any fixed ordered graph H and any integer k, the property that an ordered graph G admits a k-edge coloring with no monochromatic (ordered) induced copy of H is strongly testable. As mentioned above Ramsey properties of this type have been considered in the Combinatorics literature, see [22] and the references therein. Another family of examples includes properties of (integer) intervals on the line. Any interval can be encoded by an edge connecting its two endpoints, where the order on the vertices (the endpoints) is the usual order on the real line. A specific example of a hereditary property is that the given set of intervals is closed under intersection. The forbidden structure is a set of 4 vertices $i < j < k < l$ where ik and jl are edges (representing intervals) whereas jk is a non-edge.

Finally, there are various examples of unordered hereditary graph properties that have simple representations using a small *finite* forbidden family of *ordered* subgraphs, while in the unordered representation, the forbidden family is infinite. Some examples of such properties are bipartiteness, being a chordal graph, and being a strongly chordal graph [23, 16]. For such properties, when the input graph is supplied with the "right" ordering of the vertices, one can derive the strong testability using the version of Theorem 1.6 for *finite* families of forbidden ordered subgraphs – see Theorem 6.1 below – instead of using the infinite unordered version, Theorem 1.2.

1.3 Discussion and open questions

Several possible directions for future research follow from our work.

Dependence between the parameters of the ordered removal lemmas

Our proofs rely heavily on strong variants of the graph regularity lemma. Regularity-based proofs generally have a notoriously bad dependence between the parameters of the problem. In the notation of Theorem 1.6, for a fixed finite family $\mathcal{F}$ of forbidden ordered subgraphs, δ^{-1} is generally very large in terms of ϵ^{-1}, meaning that the number of queries required for the corresponding tester for such properties is very large in terms of ϵ^{-1}. Indeed, the original Szemerédi regularity lemma imposes a tower-type dependence between these parameters [34, 39, 40], while the variant we use is at least as strong (and at least as expensive) as the strong regularity lemma [4], which is known to have a wowzer (tower of towers) type dependence between its parameters [20, 36]. Note that for infinite families $\mathcal{F}$ the dependence between the parameters may be arbitrarily bad [9].

In a breakthrough result of Fox [28], the first known proof for the (unordered) graph removal lemma that does not use the regularity lemma is given. However, the dependence between the parameters there is still of a tower type. In any case, it will be interesting to try to obtain a proof for the ordered case, that does not go through the strong regularity needed in our proof.

Better dependence for specific properties

As discussed in Subsection 1.1, for ordered binary matrices there is an efficient conditional regularity lemma [5], in which the dependence of δ^{-1} on ϵ^{-1} is polynomial. It will be interesting to try to combine the ideas from our proof with this binary matrix regularity lemma, to obtain a removal lemma for finite families of ordered binary matrices with better dependence between the parameters. Ideally, one hopes for a removal lemma with polynomial dependence, but even obtaining such a lemma with, say, exponential dependence will be interesting.

More generally, it will be interesting to find large families of hereditary ordered graph or matrix properties that have more efficient testers than those obtained from our work. See, e.g., [29] for recent results of this type for unordered graph properties.

Characterization of strongly testable ordered properties

For unordered graphs, Alon and Shapira [8] showed that a property is strongly testable using an *oblivious* one-sided tester, which is a tester whose behavior is independent of the size of the input, *if and only if* the property is (almost) hereditary. It will be interesting to obtain similar characterizations in the ordered case.

More generally, in the ordered case there are other general types of properties that may be of interest. Ben-Eliezer, Korman and Reichman [12] recently raised the question of characterizing the efficiently testable *local properties*, i.e., properties that are characterized by a collection of forbidden local substructures. It will also be interesting to identify and investigate large classes of *visual* (or *geometric*) properties. Due to the lack of symmetry, obtaining a complete characterization of the efficiently testable properties of ordered graphs and matrices seems to be very difficult. In fact, considering that all properties whatsoever can be formulated as properties of ordered structures (e.g. strings), any characterization here will have to define and refer to some "graphness" of our setting, even that we do not allow the graph symmetries.

Generalization to ordered hypergraphs and hypermatrices

It will be interesting to obtain similar removal lemmas (and consequently, testing results) for the high-dimensional analogues of ordered graphs and matrices, namely ordered k-uniform hypergraphs and k-dimensional hypermatrices. Such results were proved for unordered hypergraphs [44, 41, 47].

Analytic analogues via graph limits

The theory of graph limits has provided a powerful approach for problems of this type in the unordered case [15, 38, 37]. It will be interesting to define and investigate a limit object for ordered graphs; this may also help with the characterization question above.

2   Paper outline

A proof of a graph removal lemma typically goes along the following lines: First, the vertex set of the graph is partitioned into a "constant" (not depending on the input graph size itself) number of parts, and a corresponding *regularity scheme* is found. The regularity scheme essentially allows that instead of considering the original graph, one can consider a very simplified picture of a constant size structure approximately representing the graph. On one hand, the structure has to approximate the original graph in the sense that we can "clean" the graph, changing only a small fraction of the edges, so that the new graph will not contain anything not already "predicted" by the representing structure. On the other hand, the structure has to be "truthful", in the sense that everything predicted by it in fact already exists in the graph.

In the simplest case, just a regular partition given by the original Szemerédi Lemma would suffice. More complex cases, like [4, 8], require a more elaborate regularity scheme. In our case, Section 5 provides a regularity scheme that addresses both edge configuration and vertex order, combining a graph regularity scheme with a scheme for strings.

Given a regularity scheme, we have to provide the graph cleaning procedure, as well as prove that if the cleaned graph still contains a forbidden subgraph, then the original graph already contains a structure containing many such graphs (this will consist of some vertex sets referenced in the regularity scheme). In Section 6 we show how to use the scheme to prove the removal lemma and the testability theorem for the case of a finite family $\mathcal{F}$ of forbidden subgraphs, while in Section 7 we show how to extend it for the case of a possibly infinite family $\mathcal{F}$. The latter section also contains a formal definition of what it means for the regularity scheme to predict the existence of a forbidden subgraph, while for the finite case it is enough to keep it implicit.

To extract the regularity scheme we need two technical aids. One of which, in Subsection 4.2, is just a rounding lemma that allows us to properly use integer quantities to approximate real ones. While in many works the question of dealing with issues related to the divisibility of the number of vertices is just hand-waved away, the situation here is complex enough to merit a formal explanation of how rounding works.

In Subsection 4.1 we develop a Ramsey-type theorem that we believe to be interesting in its own right. The use of Ramsey-type theorems is prevalent in nearly all works dealing with regularity schemes, as a way to allow us to concentrate only on "well-behaved" structures in the scheme when we are about to clean the graph. Because of the extra complication of dealing with vertex-ordered graphs, we cannot just find Ramsey-type instances separately in different parts of the regularity scheme. Instead, we need to find the well-behaved structure "all at once", and furthermore assure that we avoid enough of the "undesirable" parts where the regularity scheme does not reflect the graph. The fraction of undesirable features, while not large, must not depend on any parameters apart from the original distance parameter ϵ (and in particular must not depend on the size of the regularity scheme), which requires us to develop the new Ramsey-type theorem.

Roughly speaking, the theorem states the following: If we have a k-partite edge-colored graph with sufficiently many vertices in each part, then we can find a subgraph where the edges between every two parts are of a single color (determined by the identity of the two parts). However, we do it in a way that satisfies another requirement: If additionally the original graph is supplied with a

set of "undesirable" edges comprising an α fraction of the total number of edges, then the subgraph we find will include not more than an $(1 + \eta)\alpha$ fraction of the undesirable edges, for an η as small as we would like (in our application $\eta = 1$ will suffice).

2.1 Finding a regularity scheme

To prove the removal lemma we need a *regularity scheme*, that is a sequence of vertex sets whose "interaction" with the graph edges, and in our case also the graph vertex order, allows us to carry a cleaning procedure using combinatorial lemmas.

Historically, in the case of properties like triangle-freeness in ordinary graphs, a regular equipartition served well enough as a regularity scheme. One needs then to just remove all edges that are outside the reach of regularity, such as edges between the sets that do not form regular pairs. When moving on to more general properties of graphs, this is not enough. We need a robust partition (see [26]) instead of just a regular one, and then we can find a subset in each of the partition sets so that these "representative" sets will all form regular pairs. This allows us to decide what to do with problem pairs, e.g. whether they should become complete bipartite graphs or become edgeless (we also need to decide what happens *inside* each partition set, but we skip this issue in the sketch).

For vertex ordered graphs, a single robust partition will not do. The reason is that even if we find induced subgraphs using sets of this partition, there will be no guarantees about the vertex order in these subgraphs. The reason is that the sets of the robust partition could interact in complex ways with regards to the vertex order. Ideally we would like every pair of vertex sets to appear in one of the following two possible ways: Either one is completely before the other, or the two are completely "interwoven".

To interact with the vertex order, we consider the robust partition along with a secondary interval partition. If we consider what happens between two intervals, then all vertices in one of the intervals will be before all vertices in the other one. This suggests that further dividing a robust partition according to intervals is a good idea. However, we also need that inside each interval, the relevant robust partition sets will be completely interwoven. In more explicit terms, we will consider what happens when we intersect them with intervals of a *refinement* of the original interval partition. If these intersections all have the "correct" sizes in relation of the original interval (i.e., a set that intersects an interval also intersects all relevant sub-intervals with sufficient vertex count), then we will have the "every possible order" guarantee.

Section 5 is dedicated to the formulation and existence proof of a regularity scheme suitable for ordered graphs. In Subsection 5.1 we present the concept of *approximating partitions*, showing several useful properties of them. Importantly, the notion of a robust partition is somewhat preserved when moving to a partition approximating it.

In Subsection 5.2 we develop the lemma that gives us the required scheme. Roughly speaking, it follows the following steps.

- We find a base partition P of the graph G, robust enough with regards to the graph edge colors, so as to ensure that it remains robust even after refining it to make it fit into a secondary interval partition.

- We consider an interval partition J of the vertex set V of G, that is robust with respect to P. That is, if we partition each interval of J into a number of smaller intervals (thus obtaining a refinement J'), most of the smaller intervals will contain about the same ratio of members of each set of P as their corresponding bigger intervals.

- Now we consider what happens if we construct a partition resulting from taking the intersections of the members of P with members of J'. In an ideal world, if a set of P intersects an interval of J, then it would intersect "nicely" also the intervals of J' that are contained in that interval. However, this is only mostly true. Also, this "partition by intersections" will usually not be an equipartition.

- We now modify a bit both P and J, to get Q and I that behave like the ideal picture, and are close enough to P and J. Essentially we move vertices around in P to make the intersections with the intervals in J' have about the same size inside each interval of J. We also modify the intersection set sizes (which also affects J a little) so they will all be near multiples of a common value (on the order of n). This is so we can divide them further into an equipartition that refines both the robust graph partition and the interval partition. The rounding Lemma 4.8 helps us here.

The above process generates the following scheme. Q is the modified base equipartition, and its size (i.e. number of parts) is denoted by k. I is the modified "bigger intervals" equipartition, and its size is denoted by m. We are allowed to require in advance that m will be large enough (that is, to have m bigger than a predetermined constant m_0). There is an equipartition Q' of size mt which refines both Q and I. That is, each part of Q' is fully contained in a part of Q and a part of I, and so each part of Q contains exactly t parts of Q'. Moreover, there is the "smaller intervals" equipartition I' which refines I, and has size mb where $b = r(m, t)$ for a two-variable function r that we are allowed to choose in advance (r is eventually chosen according to the Ramsey-type arguments needed in the proof). Each part of I contains exactly b parts of I'. Finally, there is a "perfect" equipartition Q'' which refines Q' and I' and has size mbt, such that inside any bigger interval from I, the intersection of each part of Q' with each smaller interval from I' consists of exactly one part of Q''. Additionally, Q' can be taken to be very robust, where we are allowed to choose the robustness parameters in advance.

We are guaranteed that the numbers m and t are bounded in terms of the above function r, the robustness parameters, and m_0 for which we required that $m \geq m_0$. These bounds do not depend on the size of the input graph. See Lemma 5.8 for more details.

2.2   Proving a finite removal lemma

Consider an ordered colored graph $G : \binom{[n]}{2} \to \Sigma$, and consider a regularity scheme consisting of equipartitions Q, I, Q', I', Q'' for G as described above.

We start by observing that if Q'' is robust enough, then there is a tuple W of "representatives" for Q'', satisfying the following conditions.

- For each part of Q'' there is exactly one representative, which is a subset of this part.

- Each representative is not too small: it is of order n (where the constants here may depend on all other parameters discussed above, but not on the input size n).

- All pairs of representatives are very regular (in the standard Szemerédi regularity sense).

- The densities of the colors from Σ between pairs of representatives are usually similar to the densities of those colors between the pairs of parts of Q'' containing them. Here the *density* of a color $\sigma \in \Sigma$ between vertex sets A and B is the fraction of σ-colored edges in $A \times B$.

Actually, the idea of using representatives, as presented above, was first developed in [4]. Note that each part of Q' contains exactly b representatives (since it contains b parts from Q'') and each small interval of I' contains exactly t representatives.

Now if Q' is robust enough then the above representatives for Q'' also represent Q' in the following sense: Densities of colors between pairs of representatives are usually similar to the densities of those colors between the pairs of parts of Q' containing them.

Consider a colored graph H whose vertices are the small intervals of I', where the "color" of the edge between two vertices (i.e. small intervals) is the $t \times t$ "density matrix" described as follows: For any pair of representatives, one from each small interval, there is an entry in the density matrix. This entry is the set of all colors from Σ that are dense enough between these two representatives, i.e., all colors whose density between these representatives is above some threshold.

An edge between two vertices of H is considered *undesirable* if the density matrix between these intervals differs significantly from a density matrix of the large intervals from I containing them. If Q' is robust enough, then most density matrices for pairs of small intervals are similar to the density matrices of the pairs of large intervals containing them. Therefore, the number of undesirables in H is small in this case.

Consider now H as an m-partite graph, where each part consists of all of the vertices (small intervals) of H that are contained in a certain large interval from I. We apply the undesirability-preserving Ramsey on H, and then a standard multicolored Ramsey within each part, to obtain an induced subgraph D of H with the following properties.

- D has exactly $d_{\mathcal{F}}$ vertices (small intervals) inside each part of H, where $d_{\mathcal{F}}$ is the maximum number of vertices in a graph from the forbidden family $\mathcal{F}$.

- For any pair of parts of H, all D-edges between these parts have the same "color", i.e. the same density matrix.

- For any part of H, all D-edges inside this part have the same "color".

- The fraction of undesirables among the edges of D is small.

Finally we wish to "clean" the original graph G as dictated by D. For any pair Q_1', Q_2' of (not necessarily distinct) parts from Q', let I_1, I_2 be the large intervals from I' containing them, and consider the density matrix that is common to all D-edges between I_1 and I_2. In this matrix there is an entry dedicated to the pair Q_1', Q_2', which we refer to as the set of colors from Σ that are "allowed" for this pair. The cleaning of G is done as follows: For every $u \in Q_1'$ and $v \in Q_2'$, if the original color of uv in G is allowed, then we do not recolor uv. Otherwise, we change the color of uv to one of the allowed colors.

It can be shown that if D does not contain many undesirables, then the cleaning does not change the colors of many edges in G. Therefore, if initially G is ϵ-far from $\mathcal{F}$-freeness, then there exists an induced copy of a graph $F \in \mathcal{F}$ in G with $l \leq d_{\mathcal{F}}$ vertices after the cleaning. Considering our cleaning method, it can then be shown that there exist representatives $R_1, \ldots, R_l$ with the following properties. For any i, all vertices of R_i come before all vertices of R_{i+1} in the ordering of the vertices, and for any $i < j$, the color of $F(ij)$ has high density in $R_i \times R_j$. Recalling that all pairs of representatives are very regular, a well-known lemma implies that the representatives $R_1, \ldots, R_l$ span many copies of F, as desired.

2.3   From finite to infinite removal lemma

After the finite removal lemma is established, adapting the proof to the infinite case is surprisingly not difficult. The only problem of the finite proof is that we required D to have exactly $d_{\mathcal{F}}$ vertices

in each large interval, where $d_{\mathcal{F}}$ is the maximal number of vertices of a graph in $\mathcal{F}$. This requirement does not make sense when $\mathcal{F}$ is infinite. Instead we show that there is a function $d_{\mathcal{F}}(m, t)$ that "plays the role" of $d_{\mathcal{F}}$ in the infinite case.

$d_{\mathcal{F}}(m, t)$ is roughly defined as follows: We consider the (finite) collection $\mathcal{C}(m, t)$ of all colored graphs with loops that have exactly m vertices, where the set of possible colors is the same as that of H (so the number of possible colors depends only on $|\Sigma|$ and t). We take $d_{\mathcal{F}}(m, t)$ to be the smallest number that guarantees the following. If a graph $C \in \mathcal{C}(m, t)$ exhibits (in some sense) a graph from $\mathcal{F}$, then C also exhibits a graph from $\mathcal{F}$ with no more than $d_{\mathcal{F}}(m, t)$ vertices.

The rest of the proof follows as in the finite case, replacing any occurrence of $d_{\mathcal{F}}$ in the proof with $d_{\mathcal{F}}(m, t)$. Here, if G contains a copy of a graph from $\mathcal{F}$ after the cleaning, then there is a set of no more than $d_{\mathcal{F}}(m, t)$ different representatives that are very regular in pairs and have the "right" densities with respect to some $F \in \mathcal{F}$ with at most $d_{\mathcal{F}}(m, t)$ vertices, so we are done as in the finite case.

From ordered graphs to ordered matrices To prove Theorem 1.8 for square matrices, we reduce the problem to a graph setting. Suppose that $M : U \times V \to \Sigma$ is a matrix, and add an additional color σ_0 to Σ. All edges between U and V will have the original colors from Σ, and edges inside U and inside V will have the new color σ_0. Note that we are not allowed to change colors to or from the color σ_0, as it actually signals "no edge". The proof now follows from the proof for graphs: We can ask the partition I into large intervals to "respect the middle", so all parts of I are either fully contained in U or in V. Moreover, colors of edges inside U or inside V are not modified during the cleaning step, and edges between U and V are not recolored to σ_0, since this color does not appear at all between the relevant representatives (and in particular, does not appear with high density).

To adapt the proof of Theorem 1.8 for non-square matrices, we need the divisibility condition to be slightly different than respecting the middle. In the case that $m = o(n)$, we need to construct two separate "large intervals" equipartitions, one for the rows and one for the columns, instead of one such equipartition I as in the graph case. The rest of the proof does not change.

3 Preliminaries and definitions

In general, we may and will assume whenever needed throughout the paper that n is large enough with respect to all relevant parameters. We generally denote "small" parameters and functions (whose values are always positive but can be arbitrarily close to zero) by small Greek letters[1], and "large" parameters and functions (whose values are always finite natural numbers but can be arbitrarily large) by Latin letters. We assume that all parameters in all statements of the lemmas are monotone in the "natural" direction, as in the following examples: $T(\alpha, b) \leq T(\alpha', b')$ for $\alpha \geq \alpha', b \leq b'$, and $\gamma(c, \delta) \leq \gamma(c', \delta')$ for $c \geq c', \delta \leq \delta'$. We also assume that all "small" parameters are smaller than one, and all "large" parameters are larger than one.

We also assume that all functions are "bounded by their parameters", for example $\gamma(\alpha, k) \leq \alpha$ and $T(\alpha, k) \geq k$. These definitions extend naturally to any set of parameters, and are easily seen to be without loss of generality as long as we do not try to optimize bounds.

---

[1] The only exception is λ, which will denote general real numbers, and ℓ which will denote their rounding

Colored graphs and charts

A Σ-*colored graph* $G = (V, c_G)$ is defined by a totally ordered set of vertices V and a function $c_G : \binom{V}{2} \to \Sigma$. That is, G is a complete ordered graph whose edges are colored by elements of Σ. The standard notion of an (ordered) graph is equivalent to a $\{0, 1\}$-colored graph. A Σ-*colored graph with loops* $G' = (V, c_{G'})$ is defined by a totally ordered set V and a function $c_{G'} : \binom{V}{2} \cup V \to \Sigma$. We identify the notation $c_{G'}(vv)$ with $c_{G'}(v)$ for any $v \in V$.

With a slight abuse of notation, we denote by $U_1 \times U_2 = \{\{u_1, u_2\} : u_1 \in U_1, u_2 \in U_2\}$ the set of edges between two disjoint vertex sets U_1 and U_2. A (k, Σ)-*chart* $C = (V_1, \ldots, V_k, c_C)$ is defined by k disjoint vertex sets $V_1, \ldots, V_k$ and a function $c_C : E_C \to \Sigma$, where $E_C = \bigcup_{1 \le i < j \le k} U_i \times U_j$. In other words, it is an edge-colored complete k-partite graph. For C and G as above, we say that C is a *partition* of G if $V = \bigcup_{i=1}^k V_i$ and $c_G(e) = c_C(e)$ for any edge $e \in E_C$. Moreover, C is *equitable* if $||V_i| - |V_j|| \le 1$ for any $1 \le i, j \le k$; an equitable partition is sometimes called an *equipartition*. The *size* $|C|$ of the partition C is the number of parts in it. For a partition C as above, a (k', Σ)-chart C' which is also a partition of G is said to be a *G-refinement* of C if we can write $C' = (V_{11}, \ldots, V_{1j_1}, \ldots, V_{k1}, \ldots, V_{kj_k}, c_{C'})$ where $V_i = \bigcup_{l=1}^{j_i} V_{il}$. Note that $c_G(e) = c_C(e) = c_{C'}(e)$ for any edge $e \in E_C$. We will sometimes omit the coloring from the description of a partition when it is clear from the context (as the coloring is determined by the partition of the vertices and the coloring of the graph).

For two disjoint sets of vertices U, W and a coloring $c : U \times W \to \Sigma$, we say that the *density* of $\sigma \in \Sigma$ in (U, W, c) is $d_\sigma(U, W, c) = |(U \times W) \cap c^{-1}(\sigma)| / |U||W|$. the squared density is denoted by $d_\sigma^2(U, W, c)$. The *index* of (U, W, c) is

$$\text{ind}(U, W, c) = \sum_{\sigma \in \Sigma} d_\sigma^2(U, W, c).$$

Note that $0 \le \text{ind}(U, V, c) \le 1$ always holds. When the coloring c is clear from context, we will usually simply write $d_\sigma(U, V)$ for density and $\text{ind}(U, V)$ for index.

For a chart C as above we define the *index* of C as

$$\text{ind}(C) = \sum_{1 \le i < i' \le k} \frac{|V_i||V_{i'}|}{\binom{|V|}{2}} \text{ind}(V_i, V_{i'}, c \restriction_{V_i \times V_{i'}})$$

where $V = \bigcup_{i=1}^k V_i$. By the Cauchy-Schwarz inequality, for any two partitions C, C' of G where C' is a G-refinement of C we have

$$0 \le \text{ind}(C) \le \text{ind}(C') \le 1. \tag{1}$$

For a function $f : \mathbb{N} \to \mathbb{N}$ and a constant $\gamma > 0$, we say that an equipartition C of size k is *(f, γ)-robust* if there exists no refining equipartition C' of C of size at most $f(k)$ for which $\text{ind}(C') > \text{ind}(C) + \gamma$. The following observation states that for any colored graph G and any equipartition C of G, there exists an (f, γ)-robust equipartition C' refining C. The first explicit definition of robustness was given in [26].

Observation 3.1 (Robust partitioning of colored graphs [26]) *For any integer $k > 0$, function $f : \mathbb{N} \to \mathbb{N}$ and real $\gamma > 0$ there exists $T = T_{3.1}(k, f, \gamma)$ such that for any equipartition C of a colored graph G with $|C| = k$, there exists an (f, γ)-robust equipartition $C' = C'_{3.1}(C, f, \gamma)$ that refines C, where $|C'| \le T$.*

Proof: Initially pick $C' = C$. Now, as long as C' is not (f, γ)-robust, let k' denote the number of parts of C'; we may replace C' by a G-refinement C'' of it with at most $f(k')$ parts and $\text{ind}(C'') > \text{ind}(C') + \gamma$. This process stops after at most $1/\gamma$ iterations, by inequality (1). $\qquad \square$

The definition of robustness immediately implies the following.

Observation 3.2 *Let $P = (V_1, \ldots, V_k)$ be an equipartition of a Σ-colored graph $G = (V, c)$, and suppose that P is $(f \circ g, \gamma)$-robust for two functions $f, g : \mathbb{N} \to \mathbb{N}$ and $\gamma > 0$. Then any equitable refinement of P with no more than $g(k)$ parts is (f, γ)-robust.*

The notion of robustness is stronger than the more commonly used notion of regularity. For a Σ-colored graph $G = (V, c)$ and an equipartition $P = (V_1, \ldots, V_k)$ of G, a pair (V_i, V_j) is ϵ-*regular* if $|d_\sigma(V_i, V_j) - d_\sigma(V_i', V_j')| \leq \epsilon$ for any $\sigma \in \Sigma$ and $V_i' \subseteq V_i, V_j' \subseteq V_j$ that satisfy $|V_i'| \geq \epsilon |V_i|, |V_j'| \geq \epsilon |V_j|$. P is an ϵ-*regular partition* if all but at most $\epsilon \binom{k}{2}$ of the pairs (V_i, V_j) are ϵ-regular. The following lemma states that robust partitions are also regular; a lemma like it is implicit in the ideas of the original proof of [45]. The original was formulated only for non-colored graphs ($\Sigma = \{0, 1\}$), but the extension to colored graphs is not hard (and was also done in prior work).

Lemma 3.3 ([45], see also [26]) *For any $\epsilon > 0$ there exist $f = f_{3.3}^{(\epsilon)} : \mathbb{N} \to \mathbb{N}$ and $\delta = \delta_{3.3}(|\Sigma|, \epsilon) > 0$ such that any (f, δ)-robust equipartition P of a Σ-colored graph G is also ϵ-regular.*

The next lemma was first formulated (with different notation and without the extension to general Σ) in [4], but in a sense the basic idea was already used in implicitly proving Lemma 3.3 in [45]. It will be useful for us later.

Lemma 3.4 ([4], see also [26]) *For any $\epsilon > 0$ there exists $\delta = \delta_{3.4}(|\Sigma|, \epsilon) > 0$, so that for every $f : \mathbb{N} \to \mathbb{N}$, any (f, δ)-robust equipartition $P = (V_1, \ldots, V_k)$ of a Σ-colored graph G, and any equitable refinement $P' = (V_{11}, \ldots, V_{1b}, \ldots, V_{k1}, \ldots, V_{kb})$ of P where $kb \leq f(k)$, choosing the indexes so that $V_i = \bigcup_{r=1}^b V_{ir}$ for any $i \in [k]$, it holds that*

$$\frac{1}{\binom{k}{2} b^2} \sum_{\sigma \in \Sigma} \sum_{i, i' \in [k]} \sum_{j, j' \in [b]} |d_\sigma(V_{ij}, V_{i'j'}) - d_\sigma(V_i, V_{i'})| \leq \epsilon.$$

Another lemma that will be useful later is the following. This is Lemma 3.2 in [4].

Lemma 3.5 ([4]) *For any $\eta > 0$ there exists a function $\beta = \beta_{3.5}^\eta : \mathbb{N} \to \mathbb{N}$, so that for any integer $l > 0$ there exists $\kappa = \kappa_{3.5}(\eta, l)$ with the following property: If $H = ([l], c_H)$ and $V_1, \ldots, V_l$ are disjoint vertex sets of $G = (V, c)$, such that for any $i < j$, (V_i, V_j) is $\beta(l)$-regular and $d_{c_H(ij)}(V_i, V_j) \geq \eta$, then the number of induced H-copies in G with a vertex from V_i playing the role of vertex i of H is at least $\kappa \prod_{i=1}^l |V_i|$.*

Strings and intervals

Consider an ordered set V whose elements are $v_1 < \ldots < v_n$. A *string* $S : V \to \Sigma$ is a mapping from the ordered set V to an alphabet Σ. An *interval partition* $I = (I_1, \ldots, I_k)$ of the string $S : V \to \Sigma$ is a partition $V = I_1 \ldots I_k$ into consecutive substrings of S: That is, there exist $0 = a_0 < \ldots < a_{k-1} < a_k = n$ such that $I_i = S(v_{a_{i-1}+1}) \ldots S(v_{a_i})$ for any $1 \leq i \leq k$. I is *equitable* (or an *interval equipartition*) if $a_i - a_{i-1} \in \{\lfloor n/k \rfloor, \lceil n/k \rceil\}$ for any $1 \leq i \leq k$. An *interval refinement* I' of I is an interval partition of S such that any part of I' is fully contained in a part of I. The *size* $|I|$ of an interval partition I is its number of parts.

Next we define notions of index and robustness that are suitable for strings and interval partitions. Similar notions were established in [11, 24]. For a string $S : V \to \Sigma$, the *density* of $\sigma \in \Sigma$ in S is $d_\sigma(S) = |S^{-1}(\sigma)|/|S|$ where $S^{-1}(\sigma) = \{v \in V : S(v) = \sigma\}$, and the squared density of σ in S is denoted by $d_\sigma^2(S)$. The *index* of S is $\text{ind}(S) = \sum_{\sigma \in \Sigma} d_\sigma^2(S)$. Finally, the *index* of an interval partition $I = (I_1, \ldots, I_k)$ of S is

$$\text{ind}(I) = \sum_{i=1}^{k} \frac{|I_i|}{|V|} \text{ind}(I_i).$$

As in the case of charts, for an interval equipartition I of a string S, we say that I is (f, γ)-robust if any interval equipartition I' of size at most $f(k)$ that refines I satisfies $\text{ind}(I') \leq \text{ind}(I) + \gamma$ (in the other direction, $\text{ind}(I) \leq \text{ind}(I')$ always holds). The following is an analogue of Observation 3.1, and its proof is essentially identical.

Observation 3.6 (Robust partitioning of intervals) *For any integer $k > 0$, function $f : \mathbb{N} \to \mathbb{N}$ and real $\gamma > 0$ there exists $T = T_{3.6}(k, f, \gamma)$ such that any interval equipartition I of a string S where $|I| = k$ has an (f, γ)-robust interval refinement $I' = I'_{3.6}(I, f, \gamma)$ consisting of at most T intervals.*

The next lemma is an analogue of Lemma 3.4 for strings, and its proof is similar.

Lemma 3.7 *For any $\epsilon > 0$ there exists $\delta = \delta_{3.7}(|\Sigma|, \epsilon) > 0$, so that for every $f : \mathbb{N} \to \mathbb{N}$, any (f, δ)-robust interval equipartition $I = (I_1, \ldots, I_k)$ of a string $S : V \to \Sigma$, and any equitable interval refinement $I' = (I_{11}, \ldots, I_{1b}, \ldots, I_{k1}, \ldots, I_{kb})$ of I where $kb \leq f(k)$, choosing the indexes so that $I_i = \bigcup_{r=1}^{b} I_{ir}$ for any $i \in [k]$, it holds that*

$$\frac{1}{kb} \sum_{\sigma \in \Sigma} \sum_{i \in [k]} \sum_{j \in [b]} |d_\sigma(I_{ij}) - d_\sigma(I_i)| \leq \epsilon.$$

We finish by defining the string that a partition of an ordered set induces on that set. The strings that we will consider in this paper are of this type.

Definition 3.8 (String of a partition) *For a partition $P = (V_1, \ldots, V_k)$ of an ordered set V, the P-string $S_P : V \to [k]$ maps any $v \in V$ to the element $i \in [k]$ such that $v \in V_i$.*

With slight abuse of notation, we will also use the notion of an interval partition in the context of ordered graphs; here each interval will simply be a set of consecutive vertices (with no accompanying function, in contrast to the case of strings).

4   Technical aids

We develop here two tools that we will use for our proofs. The first tool is a Ramsey-type theorem that we believe to be interesting in its own right. We will use it to find a "uniform" structure with a global view on our graph.

The second tool is a rounding lemma that allows us to evenly partition graphs also when the number of sets does not divide the number of vertices, without hand-waving away the divisibility issues (which might have been questionable in our context).

4.1 A quantitative Ramsey-type theorem

The multicolored Ramsey number $\text{Ram}(s,k)$ is the minimum integer n so that in any coloring of K_n with s colors there is a monochromatic copy of K_k. It is well known that this number exists (i.e. is finite) for any s and k. For our purposes, we will also need a different Ramsey-type result, that keeps track of "undesirable" edges, as described in the following subsection.

Given a k-partite Σ-chart, we would like to pick a given number of vertices from each partition set, so that all edges between remaining vertices in each pair of sets are of the same color. However, in our situation we also have a "quantitative" requirement: A portion of the edges is marked as "undesirable", and we require that in the chart induced on the picked vertices the ratio of undesirable edges does not increase by much.

Formally, we prove the following, which we state as a theorem because we believe it may have uses beyond the use in this paper.

Theorem 4.1 *There exists a function $R_{4.1} : \mathbb{N} \times \mathbb{N} \times \mathbb{N} \times (0,1] \to \mathbb{N}$, so that if $G = (V_1, \ldots, V_k, c)$ is a k-partite Σ-chart with $n \geq R_{4.1}(|\Sigma|, k, t, \alpha)$ vertices in each class, and $B \subseteq \bigcup_{i<j\in[k]}(V_i \times V_j)$ is a set of "undesirable edges" of size at most $\epsilon\binom{k}{2}n^2$, then G contains an induced subchart $H_{4.1}(G, B, t, \alpha) = (W_1, \ldots, W_k, c \!\restriction_{\bigcup_{1\leq i<j\leq k}(W_i \times W_j)})$ with the following properties.*

- *$|W_i| = t$ for every $1 \leq i \leq k$.*

- *$c \!\restriction_{W_i \times W_j}$ is a constant function for every $1 \leq i < j \leq k$.*

- *The size of $B \cap (\bigcup_{1\leq i<j\leq k}(W_i \times W_j))$ is at most $(1+\alpha)\epsilon\binom{k}{2}t^2$.*

In our use, these "vertices" would actually be themselves sets of a robust partition of the original graph, and "colors" will encode densities; an undesirable pair would have the "wrong" densities. Also, in our use case the undesirability of an edge will be determined solely by its color and the W_i that its end vertices belong to, which means that for each $1 \leq i < i' \leq k$ the edge set $W_i \times W_{i'}$ consist of either only desirable edges or only undesirable edges. When this happens, a later pick of smaller sets $W_i' \subset W_i$ will still preserve the ratio of undesirable edges (we will in fact perform such a pick using the original Ramsey's theorem inside each W_i). The following corollary summarizes our use of the theorem.

Definition 4.2 *Given a k-partite Σ-chart $G = (V_1, \ldots, V_k, c)$ and a set $B \subseteq \bigcup_{i<j\in[k]}(V_i \times V_j)$, we say that B is orderly if for every $1 \leq i < j \leq k$ there are no $e \in (V_i \times V_j) \cap B$ and $e' \in (V_i \times V_j) \setminus B$ for which $c(e) = c(e')$. In other words, the "position" and color of an edge fully determines whether it is in B.*

Corollary 4.3 *There exists a function $R_{4.3} : \mathbb{N} \times \mathbb{N} \times \mathbb{N} \to \mathbb{N}$, so that if $G = (\bigcup_{i=1}^{k} V_i, c)$ is a Σ-colored graph with $|V_i| = n \geq R_{4.3}(|\Sigma|, k, t)$ for any $i \in [k]$ and $V_i \cap V_j = \emptyset$ for any $i \neq j \in [k]$, and $B \subseteq \bigcup_{i<j\in[k]}(V_i \times V_j)$ is an orderly set of "undesirable edges" of size at most $\epsilon\binom{k}{2}n^2$, then G contains an induced subgraph D satisfying the following.*

- *The vertex set of D is $\bigcup_{i=1}^{l} U_i$ where $U_i \subseteq V_i$ and $|U_i| = t$ for any $i \in [k]$.*

- *For any $i \in [k]$, all edges inside U_i have the same color.*

- *For any $i < j \in [k]$, all edges in $U_i \times U_j$ have the same color.*

- *$\sum_{i<j\in[k]} |B \cap (U_i \times U_j)| \leq 2\epsilon\binom{k}{2}t^2$.*

Proof: Take $R_{4.3}(s,k,t) = R_{4.1}(s,k,\mathrm{Ram}(s,t),1)$ (recall that $\mathrm{Ram}(s,t)$ denotes the "traditional" s-colored Ramsey function). By Theorem 4.1, there exists a chart $H = (W_1, \ldots, W_k)$ with the following properties.

- $W_i \subseteq V_i$ and $|W_i| = \mathrm{Ram}(t,|\Sigma|)$ for every $i \in [k]$.

- For any pair $i < j \in [k]$, all edges in $W_i \times W_j$ have the same color.

- $\sum_{i<j\in[k]} |B \cap (W_i \times W_j)| \leq 2\epsilon \binom{k}{2} (\mathrm{Ram}(t,|\Sigma|))^2$.

Observe that for any pair $i < j \in [k]$, either $W_i \times W_j \subseteq B$ or $(W_i \times W_j) \cap B = \emptyset$, since all edges in $W_i \times W_j$ have the same color and B is orderly. Therefore, the number of pairs $i < j$ for which $(W_i \times W_j) \cap B \neq \emptyset$ is at most $2\epsilon \binom{k}{2}$. Now we apply the traditional Ramsey's theorem inside each W_j to obtain a set $U_j \subseteq W_j$ of size t such that all edges inside W_j have the same color. Since $\sum_{i<j} |B \cap (U_i \times U_j)| \leq \sum_{i<j:(W_i \times W_j)\cap B\neq\emptyset} |B \cap (U_i \times U_j)| \leq 2\epsilon \binom{k}{2} t^2$, the proof follows. □

Before moving to the proof of Theorem 4.1 itself, let us quickly note that a quantitative counterpart for the traditional (not k-partite) graph case does not exist (indeed, Corollary 4.3 is a way for us to circumvent such issues).

Proposition 4.4 *For any $\alpha > 0$, m, k, and large enough l, for infinitely many n there is a graph G and a set of undesirable pairs B, so that G has n vertices, B consists of at most $\frac{1}{mk}\binom{n}{2}$ pairs, G has no independent set of size l, and every clique of l vertices in G holds at least $(\frac{1}{m} - \alpha)\binom{l}{2}$ members of B.*

Proof: We construct G for any n that is a multiple of mk larger than lk. The graph G will be the union of k vertex-disjoint cliques, each with n/k vertices. In particular, G contains no independent set with l vertices, and any clique with l vertices must be fully contained in one of the cliques of G.

Now B will be fully contained in the edge-set of G, and will consist of the edge-set of mk vertex-disjoint cliques with n/mk vertices each, so that each of the cliques of G contains m of them. It is now not hard to see that any clique with l vertices in G will contain at least $(\frac{1}{m} - \alpha_l)\binom{l}{2})$ members of B, where $\lim_{l\to\infty} \alpha_l = 0$. □

Moving to the proof, the following is our main lemma. It essentially says that we can have a probability distribution over "Ramsey-configurations" in our chart that has some approximate uniformity properties.

Lemma 4.5 *There exists a function $R_{4.5} : \mathbb{N} \times \mathbb{N} \times \mathbb{N} \times (0,1] \to \mathbb{N}$, so that if $G = (V_1, \ldots, V_k, c)$ is a k-partite Σ-chart with $n \geq R_{4.5}(|\Sigma|, k, t, \delta)$ vertices in each class, then G contains a randomized induced subchart $H_{4.5}(G,t,\delta) = (W_1, \ldots, W_k, c \restriction_{\bigcup_{1\leq i<j\leq k}(W_i\times W_j)})$ satisfying the following properties.*

- *Either $|W_i| = t$ for every $1 \leq i \leq k$, or the chart is empty ($W_i = \emptyset$ for every i).*

- *$c \restriction_{W_i\times W_j}$ is a constant function for every $1 \leq i < j \leq k$ (with probability 1).*

- *For every $1 \leq i \leq k$, every $v \in V_i$ will be in W_i with probability at most t/n.*

- *For every $1 \leq i < j \leq k$, every $v \in V_i$ and every $w \in V_j$, the probability for both $v \in W_i$ and $w \in W_j$ to hold is bounded by $(t/n)^2$.*

15

- *The probability that the chart is empty is at most δ.*

Before we prove this lemma, we show how it implies Theorem 4.1.

Proof of Theorem 4.1 We set $R_{4.1}(a, k, t, \alpha) = R_{4.5}(a, k, t, \alpha/3)$. Given the k-partite Σ-chart G, we take the randomized subchart $H = H_{4.5}(G, t, \alpha/3) = (W_1, \ldots, W_k, c \upharpoonright_{\bigcup_{1 \leq i < j \leq k}(W_i \times W_j)})$, and prove that with positive probability it is the required subchart.

Let $B' = B \cap (\bigcup_{1 \leq i < j \leq k} W_i \times W_j)$ denote the set of undesirable pairs that are contained in H. By the probability bound on pair containment and by the linearity of expectation, $\mathbb{E}[|B'|] \leq (t/n)^2 |B| \leq \epsilon \binom{k}{2} t^2$. Therefore, the probability for $|B'|$ to be larger than $(1 + \alpha)\epsilon \binom{k}{2} t^2$ is bounded by $\frac{1}{1+\alpha} \leq 1 - \alpha/2$. Therefore, with positive probability, both $|B'|$ is not too large and H is not the empty chart. Such an H is the desired subchart. $\qquad\square$

To prove Lemma 4.5 we shall make good use of the following near-trivial observation.

Observation 4.6 *There exists a function $m_{4.6} : \mathbb{N} \times \mathbb{N} \times (0, 1] \to \mathbb{N}$, such that if A is a set of size at least $m_{4.6}(k, t, \delta)$ and $\mathcal{A} = (A_1, \ldots, A_k)$ is a partition of A to k sets, then there exists a randomized subset $B = B_{4.6}(\mathcal{A}, t, \delta)$ satisfying the following properties.*

- *Either $|B| = t$ or $B = \emptyset$.*

- *B is fully contained in a single A_i.*

- *For every $a \in A$, the probability for $a \in B$ is at most $t/|A|$.*

- *The probability for $B = \emptyset$ is at most δ.*

Proof: To choose the randomized subset B, first choose a random index I where $\Pr[I = i] = |A_i|/|A|$ for all $1 \leq i \leq k$. Next, if $|A_I| < t$ then set $B = \emptyset$, and otherwise set B to be a subset of size exactly t of A_I, chosen uniformly at random from all $\binom{|A_I|}{t}$ possibilities. Setting $m_{4.6}(k, t, \delta) = tk/\delta$, it is not hard to see that all properties for the random set B indeed hold. $\quad\square$

Proof of Lemma 4.5 The proof is done by induction over k. The base case $k = 1$ is easy – set $R_{4.5}(|\Sigma|, 1, t, \delta) = t$, and let W_1 be a uniformly random subset of size t of V_1.

For the induction step from $k - 1$ to k, we set $R_{4.5}(|\Sigma|, k, t, \delta) = m_{4.6}(|\Sigma|^s, r, \frac{1}{k+1}\delta)$, where $s = m_{4.6}(|\Sigma|^{k-1}, t, \frac{1}{k+1}\delta)$ and $r = R_{4.5}(|\Sigma|, k - 1, t, \frac{1}{k+1}\delta)$. We set $W_1, \ldots, W_k$ to be the result of the following random process.

First, we set $V_1' \subseteq V_1$ to be a uniformly random subset of size exactly s. Then, for every $2 \leq i \leq k$, we set $V_i' \subseteq V_i$ to be the random set $B_{4.6}(\mathcal{V}_i, r, \frac{1}{k+1}\delta)$, where $\mathcal{V}_i$ is the partition of V_i obtained by classifying every $v \in V_i$ according to the colors $\langle c(w, v)\rangle_{w \in V_1'}$, i.e., two vertices in V_i are in the same partition set if all their pairs with vertices from V_1' have the same colors.

If any of the V_i' came out empty, we set all W_i to $\emptyset$ and terminate the algorithm (this occurs with probability at most $\frac{k-1}{k+1}\delta$), and otherwise we continue. Note now, in particular, that for every $w \in V_1$ and $v \in V_i$ the probability for both $w \in V_1'$ and $v \in V_i'$ to hold is bounded by $(s/n)(r/n)$. This is since the probability guarantees of Observation 4.6 hold for any possible value of V_1'. Also, since each V_i' was independently drawn, for $v \in V_i$ and $w \in V_j$ (where $1 < i < j \leq k$) the probability for both $v \in V_i'$ and $w \in V_j'$ to hold is bounded by $(r/n)^2$.

We now let H' denote the $(k - 1)$-partite Σ-chart induced by $V_2', \ldots, V_k'$, and use the induction hypothesis to (randomly) set $W_2, \ldots, W_k$ as the corresponding vertex sets of $H_{4.5}(H', t, \frac{1}{k+1}\delta)$. As before, if we receive empty sets then we also set $W_1 = \emptyset$ and terminate. Note that for $1 < i \leq k$

and $v \in V_i'$, the probability for v to be in W_i is bounded by t/r. Hence, for $v \in V_i$, the probability for $v \in W_i$ to hold is bounded by $(r/n)(t/r) = t/n$. Similarly, for $1 < i < j \le k$, for every $v \in V_i$ and $w \in V_j$ the probability for both $v \in W_i$ and $w \in W_j$ to hold is bounded by $(t/n)^2$. Also by similar considerations, for $v \in V_1$ and $w \in V_i$, the probability for both $v \in V_1'$ and $w \in W_i$ to hold is bounded by $(s/n)(t/n)$.

Finally, we set W_1 to be the random set $B_{4.6}(\mathcal{V}', t, \frac{1}{k+1}\delta)$, where $\mathcal{V}'$ is the partition of V_1' obtained by classifying each $v \in V_1'$ by the colors $\langle c(v, w) \rangle_{w \in W_i}$. Note that $c(v, w)$ in that expression depends only on v and the index i for which $w \in W_i$, because of how we chose each V_i' above. In particular, after the choice of W_1, the function $c \upharpoonright_{W_1 \times W_i}$ is constant for each $1 < i \le k$. Again, if we got an empty set for W_1, we set all $W_2, \ldots, W_k$ to be empty as well. By similar considerations as in the preceding steps, also here, for any $v \in V_1$ the probability of $v \in W_1$ is bounded by t/n, and for $w \in V_i$ where $1 < i \le k$, the probability of both $v \in W_1$ and $w \in W_i$ is bounded by $(t/n)^2$.

The probability of obtaining empty sets in any of the steps is bounded by δ by a union bound, and all other properties of the random sets $W_1, \ldots, W_k$ have already been proven above. $\qquad\square$

4.2   Multipartitions and rounding

The following is a mechanism to handle "with one stroke" rounding issues throughout the paper.

Definition 4.7 (Multipartitions) *A* multipartition *of a set L is a family M of subsets of L, that in particular includes L and all the singletons $\{i\}$ for $i \in L$, and furthermore every two sets $A, B \in M$ are either disjoint or one is contained in the other.*

To get an idea, an object that can be modeled as a multipartition is a partition of L (the multipartition would contain the partition sets, along with L itself and all singleton sets $\{i\}$), but also other objects, such as a partition and its refinement together, can be modeled as multipartitions. Here is the main lemma.

Lemma 4.8 (rounding feasibility) *If M and N are two multipartitions of the same set L, and $\lambda_i \in \mathbb{R}$ is a real value attached to every $i \in L$, then there exist integer values $\ell_i \in \mathbb{Z}$ attached to $i \in L$, satisfying the following.*

- $\ell_i \in \{\lfloor \lambda_i \rfloor, \lceil \lambda_i \rceil\}$ *for every $i \in L$.*

- $\sum_{i \in A} \ell_i \in \{\lfloor \sum_{i \in A} \lambda_i \rfloor, \lceil \sum_{i \in A} \lambda_i \rceil\}$ *for every $A \in M$ and for every $A \in N$.*

- $\sum_{i \in L} \ell_i \in \{\lfloor \sum_{i \in L} \lambda_i \rfloor, \lceil \sum_{i \in L} \lambda_i \rceil\}$.

Proof: Note that the middle item implies the other two (since M and N in particular include L and all singleton sets $\{i\}$ for $i \in L$). We define the following problem of solving a flow network with minimal and maximal constraints (for an exposition of flow networks see [48]).

- The node set of the flow network is $\{u_A : A \in M\} \cup \{w_A : A \in N\}$.

- The start node is u_L and the target node is w_L.

- For every $A, B \in M$, so that $A \subsetneq B$ and there is no $C \in M$ for which $A \subsetneq C \subsetneq B$, we put an edge from u_B to u_A with minimum flow $\lfloor \sum_{i \in A} \lambda_i \rfloor$ and maximum flow $\lceil \sum_{i \in A} \lambda_i \rceil$.

- For every $A, B \in N$, so that $A \subsetneq B$ and there is no $C \in N$ for which $A \subsetneq C \subsetneq B$, we put an edge from w_A to w_B with minimum flow $\lfloor \sum_{i \in A} \lambda_i \rfloor$ and maximum flow $\lceil \sum_{i \in A} \lambda_i \rceil$.

- For every $i \in L$ we put an edge from u_i to w_i with minimum flow $\lfloor \lambda_i \rfloor$ and maximum flow $\lceil \lambda_i \rceil$.

- We require the total flow of the network to be between $\lfloor \sum_{i \in L} \lambda_i \rfloor$ and $\lceil \sum_{i \in L} \lambda_i \rceil$.

This flow network has a real-valued solution by assigning λ_i flow to each edge of the type u_i, w_i, and then assigning the corresponding sums to all other network edges. Hence (since all constraints are integer-valued), the flow network has an integer-valued solution as well (see, e.g., the analysis of Lawler's algorithm in [48], page 602). Fixing such a solution, and setting ℓ_i to be the flow in the edge u_i, w_i for every $i \in L$, completes the proof. $\qquad \square$

An example of using the lemma is when we want to round the values in a 2-dimensional matrix so that the row sums and column sums are also rounded versions of the original sums (and in particular equal to the original sums if they happen to be integers). In our use the resulting integer values would be set sizes for an equipartition, that in turn refines other partitions with set size requirements.

We also note here that the statement of this lemma is false when we are presented with three multipartitions. Take for example the 3-dimensional matrix of size $2 \times 2 \times 2$, where $\lambda_{111} = \lambda_{100} = \lambda_{010} = \lambda_{001} = \frac{1}{2}$, with all other λ values being zero. Also for each of the three dimensions take the partition into two axes-parallel planes. The values on every set of every partition sum up to 1, and yet there are no corresponding $\ell_{ijk} \in \{\lfloor \lambda_{ijk} \rfloor, \lceil \lambda_{ijk} \rceil\}$ satisfying these constraints.

5 A regularity scheme for ordered graphs

5.1 The approximating partition framework

Definition 5.1 (δ-approximating partitions) *Let $P = (V_1, \ldots, V_k)$ and $Q = (U_1, \ldots, U_l)$ be partitions of a set V of size n. We say that Q is a δ-approximation of P, or equivalently, that P and Q are δ-close, if there exists $T \subseteq V$ with $|T| \leq \delta n$ such that $V_i \setminus T = U_i \setminus T$ for any $1 \leq i \leq \max\{k, l\}$, where for $i > k$ we define $V_i = \phi$, and similarly $U_i = \phi$ for $i > l$.*

Lemma 5.2 *For any $\epsilon > 0$ there exists $\delta = \delta_{5.2}(\epsilon) > 0$ such that any two δ-close partitions P and Q of (the vertex set of) a colored graph $G = (V, c)$ satisfy $|\mathrm{ind}(P) - \mathrm{ind}(Q)| \leq \epsilon$.*

Proof: Let $P = (V_1, \ldots, V_k)$ and $Q = (U_1, \ldots, U_l)$ be δ-close partitions of G, where we assume w.l.o.g. that $k < l$. For any $1 \leq i \leq k$ let $W_i = V_i \cap U_i$, and observe that $\sum_{i=1}^{k} |W_i| \geq (1 - \delta)n$. we say that i is *bad* if $|W_i| \leq (1 - \sqrt{\delta}) \min\{|V_i|, |U_i|\}$ and *good* otherwise. Then $\sum_{i \text{ bad}} |V_i| \leq \sqrt{\delta}n$ and $\sum_{i \text{ bad}} |U_i| \leq \sqrt{\delta}n$. When i and j are both good, we have

$$\left| \mathrm{ind}(V_i, V_j) - \mathrm{ind}(W_i, W_j) \right| \leq \sum_{\sigma \in \Sigma} \left| (d_\sigma^2(V_i, V_j) - d_\sigma^2(W_i, W_j) \right| \leq 2 \sum_{\sigma \in \Sigma} \left| d_\sigma(V_i, V_j) - d_\sigma(W_i, W_j) \right|$$

$$\leq 2 \sum_{\sigma \in \Sigma} \left(d_\sigma(V_i, V_j) \left(\frac{|V_i||V_j|}{|W_i||W_j|} - 1 \right) + \frac{|c^{-1}(\sigma) \cap ((V_i \times V_j) \setminus (W_i \times W_j))|}{|W_i||W_j|} \right) = O(\sqrt{\delta})$$

where the second inequality holds since $d_\sigma(V_i, V_j) + d_\sigma(U_i, U_j) \leq 2$, the third inequality follows from the fact that $|x - y| \leq z - x + z - y$ for $z \geq \max\{x, y\}$ and the last inequality follows from the observation that $|V_i||V_j| = (1 + O(\sqrt{\delta}))|W_i||W_j|$. Similarly, it holds that $|\mathrm{ind}(U_i, U_j) -$

$\text{ind}(W_i, W_j)| = O(\sqrt{\delta})$, so $|\text{ind}(V_i, V_j) - \text{ind}(U_i, U_j)| = O(\sqrt{\delta})$ when i, j are good. We finish by observing that

$$\text{ind}(P) - \text{ind}(Q) \leq \sum_{i < j \text{ good}} \left(\frac{|V_i||V_j|}{\binom{n}{2}}\text{ind}(V_i, V_j) - \frac{|U_i||U_j|}{\binom{n}{2}}\text{ind}(U_i, U_j) \right) + 2\sqrt{\delta} = O(\sqrt{\delta})$$

where the first inequality holds since $\sum_{i \text{ bad}} \sum_{j \neq i} |V_i||V_j|ind(V_i, V_j)/\binom{n}{2} \leq 2\sqrt{\delta}$ and the second inequality is true since $|V_i||V_j| = \left(1 + O(\sqrt{\delta})\right)|U_i||U_j|$ and $\text{ind}(V_i, V_j) = \left(1 + O(\sqrt{\delta})\right)\text{ind}(U_i, U_j)$ when i and j are good, and since $\text{ind}(Q) \leq 1$. Therefore, taking a suitable $\delta = \Theta(\epsilon^2)$ in the statement of the lemma suffices. $\qquad\square$

Lemma 5.3 *Let P, Q be δ-close equipartitions of a colored graph G, where $|P| = |Q|$. Then any equitable refinement Q' of Q is δ-close to an equitable refinement P' of P, with $|P'| = |Q'|$.*

Proof: Write $P = (V_1, \ldots, V_k)$, $Q = (U_1, \ldots, U_k)$, $Q' = (U_{11}, \ldots, U_{1r}, \ldots, U_{k1}, \ldots, U_{kr})$ where $U_i = \bigcup_{j=1}^r U_{ij}$. Also, for any i, j let $W_i = V_i \cap U_i$ and $W_{ij} = V_i \cap U_{ij}$. Then $\sum_{i=1}^k \sum_{r=1}^r |W_{ij}| = \sum_{i=1}^k |W_i| \geq (1 - \delta)n$, so we may take a refinement $P' = (V_{ij})_{1 \leq i \leq k, 1 \leq j \leq r}$ of P as follows: For any i, j we take V_{ij} that contains W_{ij} and ℓ_{ij} arbitrary additional elements from $V_i \setminus W_i$, where ℓ_{ij} is chosen by using Lemma 4.8 in the following manner.

For $1 \leq i \leq k$ and $1 \leq j \leq r$ we set $\lambda_{ij} = |V_i|/r - |W_{ij}|$. We set the multipartition M to consist of all singleton sets $\{ij\}$, the set $[k] \times [r]$, and the sets $\{i\} \times [r]$ for $1 \leq i \leq k$. We set the multipartition N to be the trivial one, just the singleton sets and $[k] \times [r]$. Invoking Lemma 4.8, we claim the following about the resulting ℓ_{ij}: Since $\sum_{j=1}^r \lambda_{ij} = |V_i| - |W_i|$, which is an integer, this will also equal the corresponding sum $\sum_{j=1}^r \ell_{ij} = |V_i| - |W_i|$, so we can get this way a refinement of P. Also, for any integer m (in our case $|V_i|$) it holds that $\lfloor \frac{m}{r} \rfloor = \lceil \frac{m+1}{r} \rceil - 1$, so the resulting V_{ij} would form an equipartition. The last issue that we need to deal with is when we have $\ell_{ij} = -1$ for some i and j, which could in fact happen. We claim however that in such a case we can move to another solution for which $\ell_{ij} = 0$. To see this, we note that $\ell_{ij} = -1$ only if $|V_i|/r$ is not an integer, $|W_{ij}| = |U_{ij}| = \lceil |V_i|/r \rceil$, and $\ell_{ij} = \lfloor \lambda_{ij} \rfloor$. But in this case one can see that there exists $j' \neq j$ so that $\ell_{ij'} = \lceil \lambda_{ij'} \rceil > 0$, and so we can increase ℓ_{ij} by 1 at the expense of $\ell_{ij'}$. $\qquad\square$

Lemma 5.4 *For any $\epsilon > 0$ there exists $\delta = \delta_{5.4}(\epsilon) > 0$ such that the following holds: If P and Q are δ-close equipartitions of a colored graph G, P is (f, δ)-robust and $|P| = |Q|$, then Q is (f, ϵ)-robust.*

Proof: Let P, Q be equipartitions as in the statement and pick $\delta = \delta_{5.2}(\epsilon/3)$. Consider an equitable refinement Q' of Q of size at most $f(|Q|) = f(|P|)$. By Lemma 5.3 there exists some equitable refinement P' of P which δ-approximates Q' where $|P'| = |Q'| \leq f(|P|)$. The robustness of P implies that $\text{ind}(P') - \text{ind}(P) \leq \delta \leq \epsilon/3$. By Lemma 5.2, $|\text{ind}(P) - \text{ind}(Q)| \leq \epsilon/3$ and $|\text{ind}(P') - \text{ind}(Q')| \leq \epsilon/3$. We conclude that $\text{ind}(Q') - \text{ind}(Q) \leq \epsilon$. Thus, Q is (f, ϵ)-robust. $\qquad\square$

The definition of δ-close partitions works exactly the same for interval partitions. We observe that interval equipartitions and their densities are mostly determined by the number of parts.

Observation 5.5 *Any two interval equipartitions I and J of $[n]$ into m parts are $\frac{m^2}{n}$-close to each other. In particular, for any f, m and ϵ, if n is large enough as a function of m and ϵ, then for such I and J the densities satisfy $\frac{1}{m} \sum_{i=1}^m \sum_{\sigma \in \Sigma} |d_\sigma(I_i) - d_\sigma(J_i)| \leq \epsilon$.*

Proof: If I and J are two interval equipartitions of $[n]$ with $|I| = |J| = m$, then we can set $T = \bigcup_{i=1}^{m-1} [i\lfloor \frac{n}{m}\rfloor + 1, i\lceil \frac{m}{m}\rceil]$. Clearly $|T| < m^2$ and $I_i \setminus T = J_i \setminus T$ for every $i \in [m]$. The second part of the observation then follows easily from the first part for n large enough. $\square$

5.2 The core lemmas

Definition 5.6 (Least Common Refinement) *For two partitions $P = (V_1, \ldots, V_k)$ and $Q = (U_1, \ldots, U_l)$ of a colored graph G, the* least common refinement *(LCR) $P \sqcap Q$ of P and Q is the partition $(V_1 \cap U_1, \ldots, V_1 \cap U_l, \ldots, V_k \cap U_1, \ldots, V_k \cap U_l)$ (after removing empty sets from the list).*

Note that even if P and Q are equitable, $P \sqcap Q$ is not necessarily equitable.

The following lemma allows us to combine an "order-respecting" interval partition and a robust graph partition. The last statement in the formulation (about even n and m) is not needed for the rest of the proofs concerning ordered graphs, but we will refer to it when we discuss ordered matrices.

Lemma 5.7 *For any $\delta > 0$ and positive integers k, m and b, there exists $\gamma = \gamma_{5.7}(\delta, k) > 0$ such that the following holds: If P is an equipartition of an n-vertex colored graph G (for $n \geq N_{5.7}(\delta, k, m, b)$) with $|P| = k$, and J is an interval equipartition of S_P of size m which is (f, γ)-robust, where $f(m) \geq mb$, then there exist an interval equipartition $I = I_{5.7}(\delta, P, m, b)$ of size m, an interval equipartition $I' = I'_{5.7}(\delta, P, m, b)$ of size mb which refines I, an equipartition $Q = Q_{5.7}(\delta, P, m, b)$ of size k which δ-approximates P, and an equipartition $Q' = Q'_{5.7}(\delta, P, m, b)$ of size at most $T_{5.7}(\delta, k, m)$ which is a refinement of both I and Q, all satisfying that the LCR $Q'' = I' \sqcap Q'$ is an equipartition of size $|Q''| = b|Q'| = |Q'||I'|/|I|$ (i.e., each set of Q' intersects "nicely" all subintervals of the interval of I that contains it).*

Furthermore, if m and n are even, then I "respects the middle", that is $\sum_{i=1}^{m/2} |I_i| = \frac{n}{2}$.

Proof: We denote $P = (V_1, \ldots, V_k)$, and set $\gamma_{5.7}(\delta, k) = \delta_{3.7}(k, \delta/20)$. The sets of the original interval equipartition J will be denoted by $J_1, \ldots, J_m$.

We denote the eventual intervals of I by $(I_1, \ldots, I_m)$, denote the eventual intervals of I' by $(I_{11}, \ldots, I_{1b}, \ldots, I_{m1}, \ldots, I_{mb})$ where $I_i = \bigcup_{j=1}^b I_{ij}$ for any $i \in [m]$. The eventual sets of Q will be denoted by $(U_1, \ldots, U_k)$, the sets of Q' by $(W_{11}, \ldots, W_{1t}, \ldots, W_{m1}, \ldots, W_{mt})$ where $I_i = \bigcup_{s=1}^t W_{is}$, and the eventual sets of Q'' will be denoted as $W_{ijs} = W_{is} \cap I_{ij}$. We pick $t = k\lceil 20/\delta\rceil$, and correspondingly $T_{5.7}(\delta, k, m) = mt$.

Before choosing the partition intervals and sets themselves, we will choose sizes for the sets, and also choose sets of indexes $K_1, \ldots, K_k$ describing the connection of Q to its refinement Q'. That is, eventually we will have $U_a = \bigcup_{(is) \in K_a} W_{is}$ for every $a \in [k]$. Finally defining the eventual $K_{ia} = \{s : (is) \in K_a\}$, $U_{ia} = I_i \cap U_a$ and $U_{ija} = I_{ij} \cap U_a$, we will also have $U_{ia} = \bigcup_{s \in K_{ia}} W_{is}$ and $U_{ija} = \bigcup_{s \in K_{ia}} W_{ijs}$.

We next determine the sizes $|K_{ia}|$, which will be found through our first use of Lemma 4.8 (and some further processing). We set the following parameters and multipartitions.

- $\lambda_{ia} = t|J_i \cap V_a|/|J_i|$.

- N contains the singleton sets $\{(ia)\}$, the set $[m] \times [k]$, and the set $\{i\} \times [k]$ for every $i \in [m]$. Note that in particular $\sum_{a=1}^k \lambda_{ia} = t$ is an integer, so we also have $\sum_{a=1}^k \ell_{ia} = t$.

20

- M contains the singleton sets $\{(ia)\}$, the set $[m] \times [k]$, and the set $[m] \times \{a\}$ for every $a \in [k]$. Note that since $|J_i| = \frac{n}{m} \pm 1 = (1 \pm \frac{m}{n})\frac{n}{m}$ and $|V_a| = (1 \pm \frac{k}{n})\frac{n}{k}$, we have $\sum_{i=1}^{m} \lambda_{ia} = (1 \pm 2\frac{m+k}{n})\frac{mt}{k}$, which means that for n large enough all the sums $\sum_{i=1}^{m} \ell_{ia}$ will equal $\frac{mt}{k} \pm 1$ (note that $\frac{mt}{k}$ is an integer), and moreover the number of $a \in [k]$ for which $\sum_{i=1}^{m} \ell_{ia} = \frac{mt}{k} + 1$ will equal the number of $a \in [k]$ for which this value is $\frac{mt}{k} - 1$.

After obtaining the values ℓ_{ia} through Lemma 4.8, we obtain ℓ'_{ia} from ℓ_{ia} through the following process: For all a for which $\sum_{i=1}^{m} \ell_{ia} = \frac{mt}{k}$, we set $\ell'_{ia} = \ell_{ia}$ for all $i \in [m]$. Otherwise we each time take an a for which $\sum_{i=1}^{m} \ell_{ia} = \frac{mt}{k} + 1$ and an a' for which $\sum_{i=1}^{m} \ell_{ia'} = \frac{mt}{k} - 1$. We choose i for which $\ell_{ia} > \ell_{ia'}$ set $\ell'_{ia} = \ell_{ia} - 1$, $\ell'_{ia'} = \ell_{ia'} + 1$, and for all other i' we set $\ell'_{i'a} = \ell_{i'a}$ and $\ell'_{i'a'} = \ell_{i'a'}$. The resulting ℓ'_{ia} satisfy $\sum_{a=1}^{k} \ell'_{ia} = t$, $\sum_{i=1}^{m} \ell'_{ia} = \frac{mt}{k}$, and $\ell'_{ia} = \lambda_{ia} \pm 2$.

Now we construct disjoint $K_1, \ldots, K_k \subset [m] \times [t]$ so that for every $i \in [m]$ and every $a \in [k]$ we have $|K_{ia}| = \ell'_{ia}$. By the equations on the sums of ℓ'_{ia} above this is doable, and results in $|K_a| = \frac{mt}{k}$ for every $a \in [k]$.

Next, we determine the sizes of the sets I_{ij} of I' and W_{ijs} of Q'', through a second use of Lemma 4.8. We set the following parameters and multipartitions, for determining $\ell_{ijs} = |W_{ijs}|$.

- We plainly set $\lambda_{ijs} = \lambda = \frac{n}{mbt}$ for all $i \in [m]$, $j \in [b]$ and $s \in [t]$. Since all values are the same, the ℓ_{ijs} will have value differences bounded by 1, as befits the equipartition Q''.

- M consists of the singletons, the set $[m] \times [b] \times [t]$, and the following.

  - The set $\bigcup_{s \in K_{ia}}(\{i\} \times \{j\} \times \{s\})$ for every $i \in [m]$, $j \in [b]$ and $a \in [k]$. This will make U_{ija} have size between $\lfloor \frac{|K_{ia}|}{t}|I_{ij}| \rfloor$ and $\lceil \frac{|K_{ia}|}{t}|I_{ij}| \rceil$ (see about $|I_{ij}|$ below).

  - The set $\{i\} \times \{j\} \times [t]$ for each $i \in [m]$ and $j \in [b]$. Eventually we will have $|I_{ij}| = \sum_{s=1}^{t} \ell_{ijs} \in \{\lfloor \frac{n}{mb} \rfloor, \lceil \frac{n}{mb} \rceil\}$, so I' will be equitable.

  - The set $\{i\} \times [b] \times [t]$ for each $i \in [m]$. Eventually we will have $|I_i| = \sum_{s=1}^{t} \sum_{j=1}^{b} \ell_{ijs} \in \{\lfloor \frac{n}{m} \rfloor, \lceil \frac{n}{m} \rceil\}$, so I will be equitable.

  - If n and m are both even, we also add the sets $[1, m/2] \times [b] \times [t]$ and $[m/2+1, m] \times [b] \times [t]$ to M. Eventually we will have $\sum_{i=1}^{m/2} |I_i| = \sum_{i=m/2+1}^{m} |I_i| = n/2$.

- N consists of the singletons, the set $[m] \times [b] \times [t]$, and the following.

  - The set $\{i\} \times [b] \times \{s\}$ for every $i \in [m]$ and $s \in [t]$. This will ensure that the eventual Q' is equitable.

  - The set $\bigcup_{s \in K_{ia}}(\{i\} \times [b] \times \{s\})$ for every $i \in [m]$ and $a \in [k]$. This will make every U_{ia} have size between $\lfloor \frac{|K_{ia}|}{t}|I_i| \rfloor$ and $\lceil \frac{|K_{ia}|}{t}|I_i| \rceil$.

  - The set $\bigcup_{(is) \in K_a}(\{i\} \times [b] \times \{s\})$ for every $a \in [k]$. This will ensure that the eventual Q is equitable.

After obtaining the values ℓ_{ijs} for the respective set sizes $|W_{ijs}|$, we finally construct the partitions themselves. First we construct I as the only interval partition for which $|I_i| = \sum_{s=1}^{t} \sum_{j=1}^{b} \ell_{ijs}$ for every $i \in [m]$, and I' as the only refinement of I for which $|I_{ij}| = \sum_{s=1}^{t} \ell_{ijs}$ for every $i \in [m]$ and $j \in [b]$. For every $i \in [m]$ and $s \in [t]$ let $b_{is} \in [k]$ be the index such that $(is) \in K_{b_{is}}$. We now go over the indexes $i \in [m]$ and $j \in [b]$, and partition the vertices of I_{ij} into the sets W_{ijs} so that as many members of $V_{b_{is}} \cap I_{ij}$ as possible will go into every W_{ijs}. When we can no longer assign vertices in this manner (because $|I_{ij} \cap V_b|$ will not necessarily equal $\sum_{b_{is}=b} \ell_{ijs}$), we assign the remaining vertices to complete the sets that do not yet have the correct size.

Having defined I, I' and Q'', we define Q' by setting $W_{is} = \bigcup_{j=1}^b W_{ijs}$ for every $i \in [m]$ and $s \in [t]$, and define Q by setting $U_a = \bigcup_{(is) \in K_a} W_{is}$. All properties of I, I', Q, Q' and Q'' immediately follow from the construction, apart from the relationship between Q and P that we still need to prove.

Because of the way we chose the sets W_{ijs} to maximize the number of vertices they contain from the "correct" sets of P, The partitions P and Q will be δ-close if

$$\sum_{i=1}^m \sum_{j=1}^b \sum_{a=1}^k \big||V_a \cap I_{ij}| - |U_a \cap I_{ij}|\big| \leq \delta n.$$

Denote the densities according to the string S_P by $d_{P,a}(I_{ij})$ (where $a \in [k]$), and the densities according to S_Q by $d_{Q,a}(I_{ij})$. For n large enough, because I' is an equipartition (interval sizes differ by not more than 1), we have

$$\frac{1}{n}\sum_{i=1}^m \sum_{j=1}^b \sum_{a=1}^k \big||V_a \cap I_{ij}| - |U_a \cap I_{ij}|\big| \leq \frac{2}{mb}\sum_{i=1}^m \sum_{j=1}^b \sum_{a=1}^k |d_{P,a}(I_{ij}) - d_{Q,a}(I_{ij})|.$$

From now on we bound the sums on the right hand side. Recall that J denotes the original interval equipartition of size m, and let J' be any refinement of J of size mb. By Observation 5.5, for n large enough we have $\frac{1}{mb}\sum_{a \in [k]} \sum_{i \in [m]} \sum_{j \in [b]} |d_{P,a}(I_{ij}) - d_{P,a}(J_{ij})| \leq \delta/20$. By Lemma 3.7, we know that $\frac{1}{mb}\sum_{a \in [k]} \sum_{i \in [m]} \sum_{j \in [b]} |d_{P,a}(J_{ij}) - d_{P,a}(J_i)| \leq \delta/20$. Now, recall that we chose the sets K_a so that $|K_{ia}| = t \cdot d_{P,a}(J_i) \pm 2$. This means that $\frac{1}{m}\sum_{a \in [k]} \sum_{i \in [m]} |d_{P,a}(J_i) - d_{Q,a}(I_i)| \leq \delta/5$ (recalling also how we chose t). Finally, by our construction, for n large enough, $\frac{1}{mb}\sum_{a \in [k]} \sum_{i \in [m]} \sum_{j \in [b]} |d_{Q,a}(I_i) - d_{Q,a}(I_{ij})| \leq \delta/20$. This follows from the size restriction that we ensured for the sets U_{ia} and U_{ija}. Using triangle inequalities with all these bounds on the density differences concludes the proof. $\square$

Lemma 5.8 *For any positive integer k, real value γ, functions $r : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ and $f : \mathbb{N} \to \mathbb{N}$, and any n-vertex ordered colored graph G (for large enough n), there exist an interval equipartition I into m parts where $k \leq m \leq S_{5.8}(\gamma, k, f, r)$, an equipartition Q' of G into mt parts (not necessarily an interval equipartition) which refines I and is additionally (f, γ)-robust, where $mt \leq T_{5.8}(\gamma, k, f, r)$, and an interval equipartition I' into $m \cdot r(m, t)$ parts also refining I, so that the LCR $Q'' = Q' \sqcap I'$ is an equipartition into exactly $mt \cdot r(m, t)$ parts (so each set of Q' intersects "nicely" all relevant intervals in I').*

Moreover, if n is even, then m will be even and I will respect the middle.

Proof: For each $l \in \mathbb{N}$ we define a function $g_l : \mathbb{N} \to \mathbb{N}$ by setting for every $m \in \mathbb{N}$

$$g_l(m) = m \cdot r(m, T_{5.7}(\delta_{5.4}(\gamma), l, m)/m).$$

Then we define a function $h : \mathbb{N} \to \mathbb{N}$ setting for every $l \in \mathbb{N}$

$$h(l) = f(T_{5.7}(\delta_{5.4}(\gamma), l, T_{3.6}(k, g_l, \gamma_{5.7}(\delta_{5.4}(\gamma), l)))).$$

We start with an equipartition P that is $(h, \delta_{5.4}(\gamma))$-robust that we obtain by Observation 3.1, and then with respect to the string S_P we obtain by Observation 3.6 an interval equipartition J that has at least k parts and is $(g_{|P|}, \gamma_{5.7}(\delta_{5.4}(\gamma), |P|))$-robust. Note that $|P| \leq T_{3.1}(1, h, \delta_{5.4}(\gamma))$, and hence $|J| \leq T_{3.6}(k, g_{T_{3.1}(1,h,\delta_{5.4}(\gamma))}, \gamma_{5.7}(\delta_{5.4}(\gamma), T_{3.1}(1, h, \delta_{5.4}(\gamma))))$, which we set as our $S_{5.8}(\gamma, k, f, r)$.

If n is even, then we make sure that k is also even (otherwise we replace it with $k+1$ in all of the above), and then $|J|$ will be even as well (and our subsequent use of Lemma 5.7 will provide an I that respects the middle).

We then invoke Lemma 5.7 to get our partitions $I = I_{5.7}(\delta_{5.4}(\gamma), P, |J|, g_{|P|}(|J|)/|J|)$, $I' = I'_{5.7}(\delta_{5.4}(\gamma), P, |J|, g_{|P|}(|J|)/|J|)$, and $Q' = Q'_{5.7}(\delta_{5.4}(\gamma), P, |J|, g_{|P|}(|J|)/|J|)$. By the size guarantees of Lemma 5.7 we have $|I| = |J|$ (ensuring our size bound for $|I|$), and $|Q'|$ is bounded by $T_{5.7}(\delta_{5.4}(\gamma), T_{3.1}(1, h, \delta_{5.4}(\gamma)), S_{5.8}(\gamma, k, f, r))$, which we set as our $T_{5.8}(\gamma, k, f, r)$.

Lemma 5.7 guarantees all requirements apart from the robustness of Q'. To prove it, we note that Q' is a refinement of the partition $Q = Q_{5.7}(\delta_{5.4}(\gamma), P, |J|, g_{|P|}(|J|)/|J|)$ into at most $T_{5.7}(\delta_{5.4}(\gamma), |P|, T_{3.6}(k, g_{|P|}, \gamma_{5.7}(\delta_{5.4}(\gamma), |P|)))$ parts, where $|Q| = |P|$ and Q and P are $\delta_{5.4}(\gamma)$-close. Hence by invoking Lemma 5.4 (which makes Q (h, γ)-robust), and then Observation 3.2, we get that Q' is indeed (f, γ)-robust. $\qquad\square$

6   The finite case for graphs

This section contains the proof of Theorem 1.6 for the case that the forbidden family $\mathcal{F}$ is finite. This is the ordered generalization of the finite induced graph removal lemma (Theorem 1.1).

Theorem 6.1 (Finite ordered graph removal lemma) *Fix a finite set Σ with $|\Sigma| \geq 2$. For any finite family $\mathcal{F}$ of ordered graphs $F : \binom{[n_F]}{2} \to \Sigma$ and any $\epsilon > 0$ there exists $\delta = \delta(\mathcal{F}, \epsilon) > 0$, such that any ordered graph $G : \binom{V}{2} \to \Sigma$ that is ϵ-far from $\mathcal{F}$-freeness contains at least δn^q induced copies of some graph $F \in \mathcal{F}$.*

The proof of Theorem 1.6 is completed in Section 7, by considering the case where $\mathcal{F}$ is infinite. The proof for the infinite case mostly relies on ideas and tools presented in this section, but requires another step, which is motivated by the ideas of Alon and Shapira [8] for the unordered case.

6.1   Representing subsets

Fix a finite alphabet Σ and a finite family $\mathcal{F}$ over Σ. Let $d_{\mathcal{F}}$ denote the largest number of vertices in a graph from $\mathcal{F}$. Now let $G = (V, c)$ be an n-vertex Σ-colored graph and suppose that I, I', Q', Q'' are equipartitions of G of sizes m, mb, mt, mbt respectively, so that I, I' are interval partitions, I' and Q' refine I, and $Q'' = I' \sqcap Q'$. More specifically, we write $I = (I_1, \ldots, I_m)$, $I' = (I_{11}, \ldots, I_{1b}, \ldots, I_{m1}, \ldots, I_{mb})$, $Q' = (U_{11}, \ldots, U_{1t}, \ldots, U_{m1}, \ldots, U_{mt})$, $Q'' = (U_{111}, \ldots, U_{mbt})$, where $I_j = \bigcup_{r=1}^{b} I_{jr} = \bigcup_{s=1}^{t} U_{js}$ for any $j \in [m]$ and $U_{jrs} = I_{jr} \cap U_{js}$ for any $j \in [m], r \in [b], s \in [t]$. Note that this is the same setting as the one obtained in Lemma 5.8, but we do not apply the lemma at this point; in particular, we currently do not make any assumptions on the equipartitions other than those stated above. We may and will assume whenever needed that n is large enough (as a function of all relevant parameters), and that any tuple of subsets of V considered in this section has at least two parts (i.e., it is not trivial).

Definition 6.2 (Representing subsets) *Let $\alpha, \beta, \mu > 0$ be real numbers and suppose that $A = (A_1, \ldots, A_l)$ is an equipartition of G. We say that $B = (B_1, \ldots, B_l)$ represents A if $B_i \subseteq A_i$ for any $i \in [l]$. Furthermore, we say that B (α, β, μ)-represents A if the following holds.*

- *$B_i \subseteq A_i$ and $|B_i| \geq \alpha n$ for any $i \in [l]$.*

- *All pairs (B_i, B_j) with $i < j \in [l]$ are β-regular.*

- $\frac{1}{\binom{l}{2}} \sum_{i<j\in[l]} \sum_{\sigma\in\Sigma} |d_\sigma(B_i, B_j) - d_\sigma(A_i, A_j)| \le \mu$.

The following lemma is a slight variant of Corollary 3.4 in [4], suggesting that partitions that are robust enough have good representing subsets. The proof follows along the same lines of the proof of Lemma 3.2 in [21], so we omit it.

Lemma 6.3 ([4, 21]) *For any $\mu > 0$ and function $\beta : \mathbb{N} \to (0,1)$ there exist a function $f = f_{6.3}^{(\beta,\mu)} : \mathbb{N} \to \mathbb{N}$ and a real number $\gamma = \gamma_{6.3}(\mu) > 0$, such that for any integer $l > 0$ there is a real number $\alpha = \alpha_{6.3}(\beta,\mu,l) > 0$, all satisfying the following. If $A = (A_1,\ldots,A_l)$ is an (f,γ)-robust equipartition of G, then there exists a tuple $B = (B_1,\ldots,B_l)$ which $(\alpha,\beta(l),\mu)$-represents A.*

The next lemma is not hard to derive from Lemma 6.3 using Lemma 3.4, and is more suitable to our setting.

Lemma 6.4 *For any function $\beta : \mathbb{N} \to (0,1)$, function $g : \mathbb{N} \to \mathbb{N}$, and real number $\mu > 0$, there exist a function $f = f_{6.4}^{\beta,g,\mu} : \mathbb{N} \to \mathbb{N}$ and a real number $\gamma = \gamma_{6.4}(\mu) > 0$, so that for any integer $l > 1$ there exists $\alpha = \alpha_{6.4}(\beta,g,\mu,l) > 0$ satisfying the following: If $A = (A_1,\ldots,A_l)$ is an (f,γ)-robust equipartition of G and $A' = (A_{11},\ldots,A_{1L},\ldots,A_{l1},\ldots A_{lL})$ is an equitable refinement of A, where $lL \le g(l)$ and $A_i = \bigcup_{j=1}^{L} A_{ij}$ for any $i \in [l]$, then there exists $B = (B_{11},\ldots,B_{1L},\ldots B_{l1},\ldots,B_{lL})$ which $(\alpha,\beta(lL),\mu)$-represents A', and satisfies*

$$\frac{1}{\binom{l}{2}L^2} \sum_{i<i'\in[l]} \sum_{j,j'\in[L]} \sum_{\sigma\in\Sigma} |d_\sigma(B_{ij}, B_{i'j'}) - d_\sigma(A_i, A_{i'})| \le 2\mu.$$

Proof: Pick $f = f_{6.4}^{\beta,g,\mu} = f_{6.3}^{(\beta,\mu)} \circ g$ and $\gamma = \gamma_{6.4}(\mu) = \min\{\delta_{3.4}(|\Sigma|,\mu), \gamma_{6.3}(\mu)\}$. Also pick $\alpha = \alpha_{6.4}(\beta,g,\mu,l) = \alpha_{6.3}(\beta,\mu,g(l))$, and suppose that A is (f,γ)-robust. By Observation 3.2 and the fact that $|A'| = lL \le g(l)$, we know that A' is $(f_{6.3}^{(\beta,\mu)}, \gamma_{6.3}(\mu))$-robust, so by Lemma 6.3 there exists a tuple $B = (B_{11},\ldots,B_{1L},\ldots B_{l1},\ldots,B_{lL})$ which $(\alpha_{6.3}(\beta,\mu,lL),\beta(lL),\mu)$-represents A', and by the monotonicity of α, B also $(\alpha,\beta(lL),\mu)$-represents A'. In particular,

$$\frac{1}{\binom{l}{2}L^2} \sum_{i<i'\in[l]} \sum_{j,j'\in[L]} \sum_{\sigma\in\Sigma} |d_\sigma(B_{ij}, B_{i'j'}) - d_\sigma(A'_{ij}, A'_{i'j'})| \le \mu.$$

Now by Lemma 3.4, and since $|A'| \le g(l) \le f(l)$,

$$\frac{1}{\binom{l}{2}L^2} \sum_{i<i'\in[l]} \sum_{j,j'\in[L]} \sum_{\sigma\in\Sigma} |d_\sigma(A'_{ij}, A'_{i'j'}) - d_\sigma(A_i, A_{i'})| \le \mu.$$

Combining the above two inequalities and using the triangle inequality concludes the proof. $\square$

6.2 The graph of the representatives and its coloring

For the next step, let $\Gamma = \Gamma(\Sigma, t)$ denote the collection of all $t \times t$ matrices M of the following form: Each entry of M is a non-empty subset of the color set Σ (where a subset is allowed to appear in multiple entries of M), so $|\Gamma(\Sigma,t)| < 2^{|\Sigma|t^2}$.

Definition 6.5 (Threshold color matrices, threshold graphs, undesirability) *Suppose that $W = (W_{111}, \ldots, W_{mbt})$ represents Q'' and define $W_{jr} = (W_{jr1}, \ldots, W_{jrt})$ and $X_j = (U_{j1}, \ldots, U_{jt})$ for any $j \in [m]$ and $r \in [b]$. Let $0 < \eta < \rho < 1/|\Sigma|$ be real numbers.*

For two t-tuples $A = (A_1, \ldots, A_t)$ and $B = (B_1, \ldots, B_t)$ where $A_s, B_s \subseteq V$ for any $s \in [t]$, the η-threshold matrix $M = M(A, B, \eta) \in \Gamma$ of the pair A, B is the $t \times t$ matrix whose (s, s') entry (for $(s, s') \in [t]^2$) is the set of colors $\sigma \in \Sigma$ that satisfy $d_\sigma(A_s, B_{s'}, c \restriction_{A_s, B_{s'}}) \geq \eta$. Note that this set cannot be empty since $\eta < 1/|\Sigma|$.

The (η, W)-threshold graph H_W^η is an (ordered) Γ-colored graph defined as follows: The vertices of H_W^η are all parts of I', and the color of the edge $I_{jr}I_{j'r'}$ is $M(W_{jr}, W_{j'r'}, \eta)$.

The edge $I_{jr}I_{j'r'}$ of H_W^η is ρ-undesirable if $j' > j$ and at least ρt^2 of the pairs $(s, s') \in [t]^2$ satisfy $M(X_j, X_{j'}, \rho)[s, s'] \nsubseteq M(W_{jr}, W_{j'r'}, \eta)[s, s']$. Finally, H_W^η is ρ-undesirable if at least $\rho\binom{m}{2}b^2$ of the edges $I_{jr}I_{j'r'}$ in it are ρ-undesirable, and ρ-desirable otherwise.

In other words, an edge $I_{jr}I_{j'r'}$ is undesirable if there are many pairs of sets $W_{jrs}, W_{j'r's'}$ in W, for which the density of some original edge color in $W_{jrs} \times W_{j'r's'}$ is significantly smaller than its density in $U_{js} \times U_{j's'}$. H_W^η is undesirable if it contains many undesirable edges. Note that the set of ρ-undesirable edges in H_W^η is orderly: Whether an edge $I_{jr}, I_{j'r'}$ of H_W^η is undesirable or not depends only on its color $M(W_{jr}, W_{j'r'}, \eta)$ and on $M(X_j, X_{j'}, \rho)$.

The following lemma relates the robustness of our partitions to the desirability of the resulting threshold charts.

Lemma 6.6 *For any $0 < \rho < 1/|\Sigma|$ and functions $\beta : \mathbb{N} \to (0, 1/|\Sigma|)$ and $g : \mathbb{N} \to \mathbb{N}$, there exist a function $f = f_{6.6}^{\rho, \beta, g} : \mathbb{N} \to \mathbb{N}$ and positive real numbers $\mu = \mu_{6.6}(\rho) \leq \rho$, $\gamma = \gamma_{6.6}(\rho)$ and $\alpha = \alpha_{6.6}(\rho, \beta, g, m, t)$, such that if Q' is (f, γ)-robust and $|Q''| \leq g(|Q'|)$, then there is a tuple $W = (W_{111}, \ldots, W_{mbt})$ which $(\alpha, \beta(mbt), \mu)$-represents Q'', and furthermore $H_W^{\rho/2}$ is ρ-desirable.*

Proof: Let $0 < \rho < 1/|\Sigma|$ and suppose that $H_W^{\rho/2}$ is ρ-undesirable, where W is any tuple that represents Q''. The definition of undesirability implies that

$$\frac{1}{\binom{m}{2}t^2b^2} \sum_{j < j' \in [m]} \sum_{s, s' \in [t]} \sum_{r, r' \in [b]} \sum_{\sigma \in \Sigma} |d_\sigma(W_{jrs}, W_{j'r's'}) - d_\sigma(U_{js}, U_{j's'})| \geq \frac{\rho\binom{m}{2}b^2\rho t^2\rho/2}{\binom{m}{2}t^2b^2} = \frac{\rho^3}{2}. \quad (2)$$

Indeed, if $M(X_j, X_{j'}, \rho)[s, s'] \nsubseteq M(W_{jr}, W_{j'r'}, \rho/2)[s, s']$ then there exists some $\sigma \in \Sigma$ for which $d_\sigma(U_{js}, U_{j's'}) \geq \rho$ but $d_\sigma(W_{jrs}, U_{j'r's'}) \leq \rho/2$, so each such event contributes $\rho/2$ to the sum in the left hand side.

Therefore, $H_W^{\rho/2}$ is ρ-desirable if the above sum is smaller than $\rho^3/2$. Thus, we pick $\mu(\rho) = \rho^3/5$. Also pick $f_{6.6}^{\rho, \beta, g} = f_{6.4}^{\beta, g, \mu}$, $\gamma_{6.6}(\rho) = \gamma_{6.4}(\mu)$, and $\alpha_{6.6}(\rho, \beta, g, m, t) = \alpha_{6.4}(\beta, g, \mu, mt)$. Since Q' is $\left(f_{6.4}^{\beta, g, \mu}, \gamma_{6.4}(\mu)\right)$-robust, and since $|Q''| \leq g(|Q'|)$, Lemma 6.4 implies that there exists $W = (W_{111}, \ldots, W_{mbt})$ which $(\alpha, \beta(mbt), \mu)$-represents Q', also guaranteeing that the left hand side of (2) is at most $2\mu < \rho^3/2$, so $H_W^{\rho/2}$ is ρ-desirable. $\qquad \square$

Definition 6.7 (Nicely colored subgraph) *Let $W = (W_{111}, \ldots, W_{mbt})$ be a tuple of subsets that represents Q'' and let $\eta > 0$. A subgraph $D = (\bigcup_{j=1}^m D_j, c_D)$ of H_W^η is said to be nicely colored if the following conditions hold.*

- *For any $j \in [m]$, $D_j \subseteq I_j$ and $|D_j| = d_{\mathcal{F}}$.*

- For any fixed $j \in [m]$, all edges inside D_j have the same color from Γ, denoted by $C^{(D)}_{jj}$.

- For any fixed $j < j' \in [m]$, all edges between D_j and $D_{j'}$ have the same color from Γ, denoted by $C^{(D)}_{jj'}$.

The next lemma follows directly from Corollary 4.3.

Lemma 6.8 *For any two integers $m, t > 0$ there exists $R = R_{6.8}(m, t)$ satisfying the following: If $b \geq R_{6.8}(m, t)$, then for any tuple $W = (W_{111}, \ldots, W_{mbt})$ that represents Q'' and any $\eta > 0$ there exists a nicely colored subgraph $D = D_{6.8}(W, \eta)$ of H^{η}_W. Moreover, if H^{η}_W is ρ-desirable for some $\eta < \rho < 1/|\Sigma|$, then the number of ρ-undesirable edges in D is at most $2\rho \binom{m}{2}(d_{\mathcal{F}})^2$.*

Proof: Take $R_{6.8}(m, t) = R_{4.3}(2^{|\Sigma|t^2}, m, d_{\mathcal{F}}) > R_{4.3}(|\Gamma_t|, m, d_{\mathcal{F}})$. Since the set of ρ-undesirable edges in H^{η}_W is orderly, we may apply Corollary 4.3 on H^{η}_W, to get a nicely colored subgraph D of it. If H^{η}_W is ρ-desirable for some $\eta < \rho < 1/|\Sigma|$, then by definition it has at most $\rho \binom{m}{2} b^2$ ρ-undesirable edges, and so the last condition in Corollary 4.3 implies that D has at most $2\rho \binom{m}{2}(d_{\mathcal{F}})^2$ ρ-undesirable edges. $\qquad\square$

6.3 Cleaning the original graph

Definition 6.9 (Cleaned graph) *Let $W = (W_{111}, \ldots, W_{mbt})$ be a tuple of subsets which represents Q'', let $\eta > 0$, and suppose that D is a nicely colored subgraph of H^{η}_W. The cleaned graph $G' = G'(G, D) = (V, c')$ is defined as follows. For any $u < v \in V$ where $u \in I_{js}$ and $v \in I_{j's'}$, we set $c'(uv) = c(uv)$ if $c(uv) \in C^{(D)}_{jj'}[s, s']$, and otherwise we set $c'(uv)$ to an arbitrary color from $C^{(D)}_{jj'}[s, s']$.*

The next lemma states that if D comes from a desirable H^{η}_W, then $G'(G, D)$ is close to G.

Lemma 6.10 *Suppose that D is a nicely colored subgraph of some H^{η}_W with W representing Q'' and $0 < \eta < \rho$, such that at most $2\rho \binom{m}{2} d^2_{\mathcal{F}}$ edges of D are ρ-undesirable. Then $G' = G'(G, D)$ is $(7|\Sigma|\rho + 2/m)$-close to G, where $m = |I|$.*

Proof: Write $G' = (V, c')$ and let $\mathcal{J}$ denote the set of pairs $j < j' \in [m]$ such that $D_j \times D_{j'}$ contains an undesirable edge. An edge $e \in \binom{V}{2}$ may satisfy $c'(e) \neq c(e)$ only if at least one of the following holds (some of the inequalities stated below rely on the assumption that n is large enough).

1. e lies inside some part I_j of I. The number of such edges is $\sum_{j=1}^{m} \binom{|I_j|}{2} \leq m \binom{\lceil n/m \rceil}{2} < \frac{2}{m} \binom{n}{2}$.

2. $e \in I_{j_1} \times I_{j_2}$ where $(j_1, j_2) \in \mathcal{J}$. But $|\mathcal{J}| \leq 2\rho \binom{m}{2}$: The number of ρ-undesirable edges in D is exactly $|\mathcal{J}| d^2_{\mathcal{F}}$, since D is orderly (with respect to the parts $D_1, \ldots, D_m$) and has $d_{\mathcal{F}}$ vertices in each D_i. Thus, $|\mathcal{J}| d^2_{\mathcal{F}} \leq 2\rho \binom{m}{2} d^2_{\mathcal{F}}$, which implies the desired inequality. Therefore, the number of edges e of of this type is less than $3\rho \binom{n}{2}$.

3. $e \in U_{js} \times U_{j's'}$ where $j < j' \in [m]$, $(j, j') \notin \mathcal{J}$, and $M(X_j, X_{j'}, \rho)[s, s'] \nsubseteq C_{jj'}(D)[s, s']$. But since the number of pairs $(s, s') \in [t]^2$ that satisfy this condition for a fixed $(j, j') \notin \mathcal{J}$ is at most ρt^2, only at most $3\rho |I_j||I_{j'}|/2$ of the edges $e \in I_j \times I_{j'}$ belong here, implying that the total number of edges of this type is less than $2\rho \binom{n}{2}$.

26

4. $e \in U_{js} \times U_{j's'}$ where $j < j' \in [m]$, $(j, j') \notin \mathcal{J}$, and $M(X_j, X_{j'}, \rho)[s, s'] \subseteq C_{jj'}(D)[s, s']$, but $d_{c(e)}(U_{js}, U_{j's'}) < \rho$. The number of such edges in $U_{js} \times U_{j's'}$ is at most $|\Sigma| \cdot \rho |U_{js}||U_{j's'}|$, and the total number of such edges is less than $2\rho |\Sigma| \binom{n}{2}$.

Therefore, the total number of edges e with $c(e) \neq c'(e)$ is less than $(7\rho |\Sigma| + 2/m) \binom{n}{2}$. $\qquad \square$

Lemma 6.11 *Let $W = (W_{111}, \ldots, W_{mbt})$ be a tuple that represents Q'' and let $\eta > 0$. If D is a nicely colored subgraph of H_W^{η} and the cleaned $G'(G, D)$ contains a copy of some $F = ([n_F], c_F) \in \mathcal{F}$, then there exist $W_{j_1 r_1 s_1}, \ldots, W_{j_{n_F} r_{n_F} s_{n_F}} \in W$ with the following properties.*

- *For any $i \in [n_F - 1]$, either $j_{i+1} > j_i$, or $j_{i+1} = j_i$ and $r_{i+1} > r_i$.*

- *For any $i < i' \in [n_F]$ it holds that $d_{c_F(ii')}(W_{j_i r_i s_i}, W_{j_{i'} r_{i'} s_{i'}}) \geq \eta$.*

Proof: Suppose that $G'(G, D) = (V, c')$ contains a copy of F whose vertices in V are $v_1 < \ldots < v_{n_F}$. For any $i \in [n_F]$, let $j_i \in [m], s_i \in [t]$ be the indices for which $v_i \in I_{j_i s_i}$ and denote the vertices of D inside I_{j_i} by $D_{j_i} = \{I_{j_i r_{i1}}, \ldots, I_{j_i r_{id_{\mathcal{F}}}}\}$, where $r_{i1} < \ldots < r_{id_{\mathcal{F}}} \in [b]$ for any $i \in [n_F]$. Then for any $i, i' \in [n_F]$ and $l, l' \in [d_{\mathcal{F}}]$, for which either $i < i'$, or $i = i'$ and $l < l'$, it holds that $c_F(ii') = c'(v_i, v_{i'}) \in C_{j_i j_{i'}}^{(D)}[s_i, s_{i'}] = M(W_{j_i r_{il}}, W_{j_{i'} r_{i'l'}}, \eta)[s_i, s_{i'}]$, and so by definition $d_{c_F(ii')}(W_{j_i r_{il} s_i}, W_{j_{i'} r_{i'l'} s_{i'}}) \geq \eta$.

Therefore, the sets $W_{j_1 r_{11} s_1}, \ldots, W_{j_{n_F} r_{n_F n_F} s_{n_F}}$ satisfy the conditions of the lemma: They exist, since $n_F \leq d_{\mathcal{F}}$. The first condition holds since $j_1 \leq \ldots \leq j_{n_F}$, and if $j_i = j_{i+1}$ then $r_{ii} = r_{(i+1)i} < r_{(i+1)(i+1)}$. The second condition holds by the first paragraph of the proof (putting $l = i$ and $l' = i'$). $\qquad \square$

6.4 Proof of Theorem 6.1

Suppose that G is ϵ-far from $\mathcal{F}$-freeness. Take the function $r = R_{6.8}$ (note that r is a two-variable function) and let $g : \mathbb{N} \to \mathbb{N}$ be defined by $g(l) = lr(l, l)$ for any $l \in \mathbb{N}$. Also take $k = \lceil 20/\epsilon \rceil$, $\rho = \epsilon/8|\Sigma|$, and $\beta : \mathbb{N} \to (0, 1/|\Sigma|)$ as a constant function that satisfies $\beta(l) = \beta_{3.5}^{\rho/2}(d_{\mathcal{F}})$ for any $l \in \mathbb{N}$. Also take $f = f_{6.6}^{\rho, \beta, g}$, and $\gamma = \gamma_{6.6}(\rho)$.

Apply Lemma 5.8 with parameters k, γ, r, f, obtaining the equipartitions I, I', Q', Q'' of sizes m, mb, mt, mbt as in the statement of the lemma, where $k \leq m \leq S_{5.8}(\gamma, k, f, r)$, $mt \leq T_{5.8}(\gamma, k, f, r)$, $b = r(m, t) = R_{6.8}(m, t)$, and Q' is (f, γ)-robust. Observe that $|Q''| = mtr(m, t) \leq g(mt) = g(|Q'|)$.

Next, define $\alpha = \alpha_{6.6}(\rho, \beta, g, S_{5.8}(\gamma, k, f, r), T_{5.8}(\gamma, k, f, r))$ and $\mu = \mu_{6.6}(\rho)$. By Lemma 6.6, and since $\beta(l) = \beta_{3.5}^{\rho/2}(d_{\mathcal{F}})$ for any $l \in \mathbb{N}$, there is a tuple W which $(\alpha, \beta_{3.5}^{\rho/2}(d_{\mathcal{F}}), \mu)$-represents Q'', and $H_W^{\rho/2}$ is ρ-desirable. By Lemma 6.8, and since $b = R_{6.8}(m, t)$, there is a nicely colored subgraph $D = D_{6.8}(W, \rho/2)$, containing at most $2\rho \binom{m}{2}(d_{\mathcal{F}})^2$ ρ-undesirable edges.

Lemma 6.10 implies that $G' = G'(G, D)$ is $(7|\Sigma|\rho + 2/m)$-close to G; but $7|\Sigma|\rho + 2/m \leq 7\epsilon/8 + 2/k < \epsilon$, so G' contains a copy of some $F = ([n_F], c_F) \in \mathcal{F}$. Therefore, by Lemma 6.11 (putting $\eta = \rho/2$ in the statement of the lemma), there exist $W_{j_1 r_1 s_1}, \ldots, W_{j_{n_F}, r_{n_F}, s_{n_F}} \in W$ that satisfy the conditions of the lemma. As all pairs of sets from W are $\beta_{3.5}^{\rho/2}(n_F)$-regular (since $n_F \leq d_{\mathcal{F}}$), we can apply Lemma 3.5 to conclude that the number of F-copies in G is at least δn^q for $q = n_F \leq d_{\mathcal{F}}$ and $\delta = \kappa_{3.5}(\rho/2, n_F)\alpha^{n_F} \geq \kappa_{3.5}(\rho/2, d_{\mathcal{F}})\alpha^{d_{\mathcal{F}}}$, concluding the proof.

7 The infinite case

In this section we use the same notation as in Section 6, unless stated otherwise. The proof of Theorem 1.6 follows that of Theorem 6.1 almost word by word, with only one major difference: In the proof of Theorem 6.1 we have picked $d_{\mathcal{F}}$ to be the largest number of vertices in a graph from $\mathcal{F}$, and showed that if G is ϵ-far from $\mathcal{F}$-freeness than there must be a set of at most $d_{\mathcal{F}}$ representatives of parts in Q'', that span a large number of F-copies for some $F \in \mathcal{F}$. However, in the infinite case, such a definition of $d_{\mathcal{F}}$ cannot work. Instead, we take $d_{\mathcal{F}}(m,t)$ to be a parameter that depends on the family $\mathcal{F}$, the size of the alphabet $|\Sigma|$ and the integers m, t (where $m = |I|$, $mt = |Q'|$). It is then shown that with this choice of $d_{\mathcal{F}}$, the proof follows similarly to the finite case, with Lemmas 6.8 and 6.11 being replaced with similar lemmas that are suitable for the infinite case (Lemmas 7.3 and 7.5 below, respectively).

7.1 Embeddability

Definition 7.1 (Embeddability) *For a finite alphabet Σ, integers $m, t > 0$, $\Gamma(\Sigma, t)$-colored graph with loops $H = ([m], c_H)$ and Σ-colored graph $F = ([n_F], c_F)$, we say that F is* embeddable *in H if there exists a mapping $h : [n_F] \to V_H$ with the following properties.*

- *h is weakly order-preserving: $h(1) \leq \ldots \leq h(n_F)$.*

- *There exist integers $s_1, \ldots, s_{n_F} \in [t]$ so that $c_F(ii') \in c_H(h(i), h(i'))[s_i, s_{i'}]$ for any $i < i' \in [n_F]$.*

A family $\mathcal{F}$ of Σ-colored graphs is embeddable *in H if some $F \in \mathcal{F}$ is embeddable in H.*

The next lemma states that the desired $d_{\mathcal{F}}$ is indeed well-defined. It is similar in spirit to the ideas of Alon and Shapira [8] (see Section 4 there).

Lemma 7.2 *Fix a finite alphabet Σ. For any (finite or infinite) family $\mathcal{F}$ of Σ-ordered graphs and integers $m, t > 0$, there exists $d_{\mathcal{F}} = d_{\mathcal{F}}^{(7.2)}(m,t)$ with the following property. If $H = ([m], c_H)$ is a $\Gamma(\Sigma, t)$-colored graph with loops, and if $\mathcal{F}$ is embeddable in H, then there is a graph $F \in \mathcal{F}$ which is embeddable in H and has at most $d_{\mathcal{F}}^{(7.2)}(m,t)$ vertices.*

Proof: Let $\mathcal{H} = \mathcal{H}_{m,t}$ denote the set of all $\Gamma(\Sigma, t)$-colored graphs $H = ([m], c_H)$ with loops, such that $\mathcal{F}$ is embeddable in H. Note that $|\mathcal{H}_{m,t}| \leq |\Gamma(\Sigma, t)|^{m^2} \leq 2^{|\Sigma| t^2 m^2}$. For any $H \in \mathcal{H}$ let $\mathcal{F}_H \subseteq \mathcal{F}$ denote the collection of all graphs in $\mathcal{F}$ that are embeddable in H. Finally define

$$d_{\mathcal{F}}^{(7.2)}(m,t) = \max_{H \in \mathcal{H}_{m,t}} \min_{F \in \mathcal{F}_H} |F|$$

where $|F|$ denotes the number of vertices in F. Since $\mathcal{H}_{m,t}$ is finite, and since the set $\mathcal{F}_H$ is non-empty for any $H \in \mathcal{H}$ (by definition of $\mathcal{H}$), the function $d_{\mathcal{F}}^{(7.2)}(m,t)$ is well defined. Now let H be a graph as in the statement of the lemma and suppose that $\mathcal{F}$ is embeddable in H. Then $H \in \mathcal{H}_{m,t}$, so there exists $F \in \mathcal{F}_H$ of size at most $d_{\mathcal{F}}^{(7.2)}(m,t)$. $\square$

7.2 Adapting the proof for infinite families

For what follows, a *nicely colored (m,t)-subgraph* is defined exactly like a nicely colored subgraph (see Definition 6.7), except that each set D_j is of size $d_{\mathcal{F}}^{(7.2)}(m,t)$. The following is a variant of Lemma 6.8 for the infinite case.

Lemma 7.3 *For any two integers $m,t > 0$ there exists $R = R_{7.3}(m,t)$ satisfying the following: If $b \geq R_{7.3}(m,t)$, then for any tuple $W = (W_{111}, \ldots, W_{mbt})$ that represents Q'' and any $\eta > 0$ there exists a nicely colored (m,t)-subgraph $D = D_{7.3}(W, \eta)$ of H_W^η. Moreover, if H_W^η is ρ-desirable for some $\eta < \rho < 1/|\Sigma|$, then the number of ρ-undesirable edges in D is at most $2\rho\binom{m}{2}(d_{\mathcal{F}}^{(7.2)}(m,t))^2$.*

The proof of Lemma 7.3 is essentially identical to that of Lemma 6.8, with any occurrence of $d_{\mathcal{F}}$ replaced by $d_{\mathcal{F}}^{(7.2)}(m,t)$. In particular we take $R_{7.3}(m,t) = R_{4.3}(2^{|\Sigma|t^2}, m, d_{\mathcal{F}}^{(7.2)}(m,t))$.

Next we state the variant of Lemma 6.10 for the infinite case. The proof is essentially identical.

Lemma 7.4 *Suppose that D is a nicely colored (m,t)-subgraph of some H_W^η with W representing Q'' and $0 < \eta < \rho$, such that at most $2\rho\binom{m}{2}(d_{\mathcal{F}}^{(7.2)}(m,t))^2$ edges of D are ρ-undesirable. Then $G' = G'(G, D)$ is $(7|\Sigma|\rho + 2/m)$-close to G, where $m = |I|$.*

The next lemma is the variant of Lemma 6.11 that we use in the infinite case. In contrast to the previous two lemmas, here the proof is slightly modified, and makes use of Lemma 7.2.

Lemma 7.5 *Let $W = (W_{111}, \ldots, W_{mbt})$ be a tuple that represents Q'' and let $\eta > 0$. If D is a nicely colored (m,t)-subgraph of H_W^η and $G'(G, D)$ contains a copy of a graph from $\mathcal{F}$, then there exist $F = ([n_F], c_F) \in \mathcal{F}$, where $n_F \leq d_{\mathcal{F}}^{(7.2)}(m,t)$, and sets $W_{j_1 r_1 s_1}, \ldots, W_{j_{n_F} r_{n_F} s_{n_F}} \in W$, with the following properties.*

- *For any $i \in [n_F - 1]$, either $j_{i+1} > j_i$, or $j_{i+1} = j_i$ and $r_{i+1} > r_i$.*

- *For any $i < i' \in [n_F]$ it holds that $d_{c_F(ii')}(W_{j_i r_i s_i}, W_{j_{i'} r_{i'} s_{i'}}) \geq \eta$.*

Proof: Consider the Γ-colored graph with loops $D' = ([m], c_{D'})$: For any $j \leq j'$, $c_{D'}(jj') = C_{jj'}^{(D)}$. Suppose that $G'(G, D) = (V, c')$ contains a copy of $A = ([n_A], c_A) \in \mathcal{F}$, whose vertices in V are $v_1 < \ldots < v_{n_A}$. For any $i \in [n_A]$, let $j_i \in [m]$, $s_i \in [t]$ be the indices for which $v_i \in I_{j_i s_i}$. Then for any $i < i' \in [n_A]$ we have $c_A(ii') = c'(v_i v_i') \in C_{j_i j_i'}^{(D)}[s_i, s_{i'}] = c_{D'}(j_i j_i')[s_i, s_{i'}]$, and so A is embeddable in D' (by the mapping $i \mapsto j_i$). By Lemma 7.2, there exists $F = ([n_F], c_F) \in \mathcal{F}$ which is embeddable in D', where $n_F \leq d_{\mathcal{F}}^{(7.2)}(m,t)$. Let $h : [n_F] \to D'$ denote a mapping that satisfies the conditions of Definition 7.1 and let $s_1', \ldots, s_{n_F}' \in [t]$ be the indices satisfying $c_F(ii') \in c_{D'}(h(i), h(i'))[s_i', s_{i'}']$ for any $i < i' \in [n_F]$.

For any $i \in [n_F]$ denote the vertices of D inside $I_{h(i)}$ by $I_{h(i)r_{i1}'}, \ldots, I_{h(i)r_{id_{\mathcal{F}}(m,t)}'}$, where $r_{i1}' < \ldots < r_{id_{\mathcal{F}}(m,t)}' \in [b]$ for any $i \in [n_F]$. The sets $W_{h(1)r_{11}'s_1'}, \ldots, W_{h(n_F)r_{n_F n_F}'s_{n_F}'}$ satisfy the desired conditions: They exist, since $n_F \leq d_{\mathcal{F}}^{(7.2)}(m,t)$, the first condition holds since h is order-preserving, and the second condition holds since $c_F(ii') \in c_{D'}(h(i), h(i'))[s_i', s_{i'}'] = C_{h(i)h(i')}^{(D)}[s_i', s_{i'}']$. $\square$

Proof of Theorem 1.6 The proof goes along the same lines as the proof of Theorem 6.1, but any occurrence of $d_{\mathcal{F}}$ in the proof of Theorem 6.1 and the accompanying lemmas is replaced here by $d_{\mathcal{F}}^{(7.2)}(m,t)$, including in the definitions of the functions β, r, and the term *nicely colored subgraph* is replaced by *nicely colored (m,t)-subgraph*. More specifically, here are the exact changes needed with respect to the proof of Theorem 6.1.

- We take the functions $\beta = \beta_{3.5}^{\rho/2}$ and $r = R_{7.3}$ (in the finite case we took β as a suitable constant function and $r = R_{6.8}$). The function g is defined as $g(l) = lr(l,l)$. Following the application of Lemma 5.8, we have $b = R_{7.3}(m,t)$.

- As in the the proof of Theorem 6.1, there is a tuple W which $(\alpha, \beta(mbt), \mu)$-represents Q'', and $H_W^{\rho/2}$ is ρ-desirable. By Lemma 7.3, and by our new choice of b, there is a nicely colored (m,t)-subgraph D of $H_W^{\rho/2}$, with at most $2\rho\binom{m}{2}\left(d_{\mathcal{F}}^{(7.2)}(m,t)\right)^2$ ρ-undesirable edges.

- Lemma 7.4 implies that G' contains a copy of a graph from $\mathcal{F}$. Now Lemma 7.5 implies the existence of sets $W_{j_1 r_1 s_1}, \ldots, W_{j_{n_F} r_{n_F} s_{n_F}} \in W$ with $n_F \leq d_{\mathcal{F}}^{(7.2)}(m,t)$, that satisfy the conditions of the lemma for $\eta = \rho/2$. Since all pairs of sets from W are $\beta_{3.5}^{\rho/2}(mbt)$-regular, and since $mbt \geq b \geq d_{\mathcal{F}}^{(7.2)}(m,t) \geq n_F$, these pairs are also $\beta_{3.5}^{\rho/2}(n_F)$-regular. We apply Lemma 3.5 to get that the number of F-copies in G is at least δn^q for

$$q = n_F \leq d_{\mathcal{F}}^{(7.2)}(m,t) \leq d_{\mathcal{F}}^{(7.2)}(S_{5.8}(\gamma,k,f,r), T_{5.8}(\gamma,k,f,r)),$$

$$\delta = \kappa_{3.5}\left(\rho/2, n_F\right)\alpha^{n_F} \geq \kappa_{3.5}\left(\rho/2, d_{\mathcal{F}}^{(7.2)}(m,t)\right)\alpha^{d_{\mathcal{F}}^{(7.2)}(m,t)}$$

$$\geq \kappa_{3.5}(\rho/2, d_{\mathcal{F}}^{(7.2)}(S_{5.8}(\gamma,k,f,r), T_{5.8}(\gamma,k,f,r)))\alpha^{d_{\mathcal{F}}^{(7.2)}(S_{5.8}(\gamma,k,f,r), T_{5.8}(\gamma,k,f,r))}.$$

Indeed, the above bounds for q and δ depend only on $|\Sigma|, \epsilon, \mathcal{F}$, and not on n. $\qquad\square$

7.3 Adapting the proof for matrices

Finally we give a sketch of the proof of Theorem 1.8 for square matrices. The proof is very similar to the graph case, so we only describe why the proof for graphs also works here. Finally, we describe shortly how the proof can be adapted to the case of non-square matrices.

Proof sketch for Theorem 1.8 Given a square matrix $M : U \times V \to \Sigma$ where U, V are ordered, and a family $\mathcal{F}$ of forbidden submatrices, consider the Σ'-colored graph $G = (U \cup V, c)$ where $\Sigma' = \Sigma \cup \{\sigma_0\}$ for some $\sigma_0 \notin \Sigma$, and the union $U \cup V$ is ordered as follows: All elements of V come after all elements of U, and the internal orders of U and V remain as before. The edge colors of G satisfy $c(uv) = M(uv)$ for any $u \in U$ and $v \in V$, and $c(uv) = \sigma_0$ otherwise.

The proof now follows as in the graph case. It is important to note that while in the graph case one is allowed to change the color of any edge, here we are not allowed to change the color of an edge from or to the color σ_0. However, the proof still works, by the following observations.

- Since $|U| = |V|$, the number of vertices in G is even, and so the interval partition I obtained here "respects the middle". That is, each part I_j of I will be fully contained in U or in V. Therefore, for every two parts $I_j, I_{j'}$ of I, either all edges in $I_j \times I_{j'}$ are colored by σ_0 or none of them is colored by σ_0.

- It follows that the set of edges of the cleaned graph $G' = G'(G, D)$ that are colored by σ_0 is identical to that of G. In other words, to generate the cleaned graph we do not modify edge colors to or from σ_0. Since G is made $\mathcal{F}$-free only by modifying colors between U and V to other colors in Σ, one needs to modify at least $\epsilon|U||V|$ edge colors, so the proof follows without changing the main arguments.

$\square$

The above proof works for square matrices, but it can be adapted to general $m \times n$ matrices: If $m = \Theta(n)$, then the condition on I needed is slightly different than respecting the middle, but this only slightly changes the structure of the equipartitions that we obtain via Lemma 5.8, without significantly affecting the proof. The proof can also be formulated for matrices with, say, $m = o(n)$ and $m = \omega(1)$, but then Lemma 5.8 needs to be especially adapted to accommodate the two "types" of vertices (row and column). Essentially we will have two interval equipartitions, one of the row vertices and one of the column vertices, along with their corresponding refinements. Finally, the case where $m = \Theta(1)$ is essentially the case of testing one-dimensional strings; strings can be handled as per the discussion in the end of Subsection 1.1.

It is important to note that one cannot use Theorem 1.6 as a black box to prove Theorem 1.8, as the distance of the graph G to $\mathcal{F}$-freeness might (potentially) be significantly smaller than ϵ, considering that the set of σ_0-colored edges in the $\mathcal{F}$-free graph that is closest to G might differ from the set of σ_0-colored edges in G.

References

[1] N. Alon, *Testing subgraphs in large graphs*, Random Structures & Algorithms 21 (2002), 359–370; also FOCS'01: Proc. 42nd Annu. Symp. Foundations of Computer Science (FOCS), IEEE, 2001, 434–441.

[2] N. Alon, O. Ben-Eliezer, *Removal lemmas for matrices*, arXiv:1609.04235, 2016.

[3] N. Alon, R. A. Duke, H. Lefmann, V. Rödl, R. Yuster, *The algorithmic aspects of the regularity lemma*, Journal of Algorithms 16 (1994), 80–109; also FOCS'92: Proc. 33th Annu. Symp. Foundations of Computer Science (FOCS), IEEE, 1992, 473–481.

[4] N. Alon, E. Fischer, M. Krivelevich, M. Szegedy, *Efficient testing of large graphs*, Combinatorica 20 (2000), 451–476; also FOCS'99: Proc. 40th Annu. Symp. Foundations of Computer Science, IEEE, 1999, 656–666.

[5] N. Alon, E. Fischer, I. Newman, *Efficient testing of bipartite graphs for forbidden induced subgraphs*, SIAM J. Comput. 37 (2007), 959–976.

[6] N. Alon, E. Fischer, I. Newman, A. Shapira, *A combinatorial characterization of the testable graph properties: it's all about regularity*, SIAM J. Comput. 39 (2009), 143–167; also STOC'06: Proc. 38th Ann. ACM Symp. Theory of Computing, 251–260, ACM, 2006.

[7] N. Alon, M. Krivelevich, I. Newman, M. Szegedy, *Regular languages are testable with a constant number of queries*, SIAM J. Comput. 30 (2001), 1842–1862; also FOCS'99: Proc. 40th Annu. Symp. Foundations of Computer Science, IEEE, 1999, 645–655.

[8] N. Alon, A. Shapira, *A characterization of the (natural) graph properties testable with one-sided error*, SIAM J. Comput. 37 (2008), 17031727; also FOCS'05: Proc. 46th Annu. Symp. Foundations of Computer Science, IEEE, 2005, 429–438.

[9] N. Alon, A. Shapira, *Every monotone graph property is testable* SIAM J. Comput. 38 (2008), 505–522; also STOC'05: Proc. 37th Ann. ACM Symp. Theory of Computing, 128–137, ACM, 2005.

[10] P. Awasthi, M. Jha, M. Molinaro, S. Raskhodnikova, *Testing Lipschitz functions on hypergrid domains*, Algorithmica 74 (2016), 1055–1081.

[11] M. Axenovich, Y. Person, S. Puzynina, *A regularity lemma and twins in words*, J. Combin. Theory Ser. A 120 (2013), 733–743.

[12] O. Ben-Eliezer, S. Korman, D. Reichman, *Deleting and testing forbidden patterns in multidimensional arrays*, arXiv:1607.03961, 2017.

[13] P. Berman, M. Murzabulatov, S. Raskhodnikova, *Tolerant testers of image properties*, ICALP'16: Proc. 43rd International Colloquium on Automata, Languages, and Programming (ICALP), 90:1–90:14, 2016.

[14] E. Blais, S. Raskhodnikova, G. Yaroslavtsev, *Lower bounds for testing properties of functions over hypergrid domains*, CCC'14: IEEE 29th Conf. Computational Complexity, IEEE, 2014, 309–320.

[15] C. Borgs, J. Chayes, L. Lovász, V. T. Sós, B. Szegedy, K. Vesztergombi, *Graph limits and parameter testing*, STOC'06: Proc. 38th Ann. ACM Symp. Theory of Computing, 261–270, ACM, 2006.

[16] A. Brandstädt, V. B. Le, J. P. Spinrad, *Graph classes: a survey*, SIAM Monographs on Discrete Mathematics and Applications, SIAM, 1999.

[17] C. Canonne, E. Grigorescu, S. Guo, A. Kumar, K. Wimmer, *Testing k-monotonicity*, ITCS'17: Proc. 8th Innovations in Theoretical Computer Science, to appear.

[18] D. Chakrabarty, C. Seshadhri, *Optimal bounds for monotonicity and Lipschitz testing over hypercubes and hypergrids*, STOC'13: Proc. 2013 ACM Symp. Theory of Computing, 419–428, ACM, 2013.

[19] D. Chakrabarty, C. Seshadhri, *An optimal lower bound for monotonicity testing over hypergrids*, Theory Comput. 10 (2014), 453–464.

[20] D. Conlon, J. Fox, *Bounds for graph regularity and removal lemmas*, Geom. Funct. Anal. 22 (2012), 1191–1256.

[21] D. Conlon, J. Fox, *Graph removal lemmas*, Surveys in Combinatorics 2013, 1–49, London Math. Soc. Lecture Note Ser. 409, Cambridge Univ. Press, Cambridge, 2013.

[22] D. Conlon, J. Fox, C. Lee, B. Sudakov, *Ordered Ramsey numbers*, J. Combin. Theory Ser. B 122 (2017), 353–383.

[23] P. Damaschke, *Forbidden ordered subgraphs*, Topics in combinatorics and graph theory, 219–229, Physica, 1990.

[24] U. Feige, T. Koren, M. Tennenholtz, *Chasing ghosts: competing with stateful policies*, SIAM J. Comput. 46 (2017), 190–23; also FOCS'14: Proc. 55th Annu. Symp. Foundations of Computer Science, IEEE, 2014, 100–109.

[25] E. Fischer, I. Newman, *Testing of matrix-poset properties*, Combinatorica 27 (2007), 293–327.

[26] E. Fischer, I. Newman, *Testing versus estimation of graph properties*, SIAM J. Comput. 37 (2007), 482–501; also STOC'05: Proc. 37th Ann. ACM Symp. Theory of Computing, 138–146, ACM, 2005.

[27] E. Fischer, E. Rozenberg, *Lower bounds for testing forbidden induced substructures in bipartite-graph-like combinatorial objects*, Proc. RANDOM 2007, 464–478.

[28] J. Fox, *A new proof of the graph removal lemma*, Ann. of Math. 174 (2011), 561–579.

[29] L. Gishboliner, A. Shapira, *Removal lemmas with polynomial bounds*, STOC'17: Proc. 2017 ACM Symp. Theory of Computing, to appear.

[30] O. Goldreich (ed.), *Property Testing: Current Research and Surveys*, Vol. 6390 of Lecture Notes in Comput. Sci., Springer, 2010.

[31] O. Goldreich, S. Goldwasser, D. Ron, *Property testing and its connection to learning and approximation*, J. ACM 45 (1998), 653–750; also Proc. 37th Annu. Symp. Foundations of Computer Science (FOCS), IEEE, 1996, 339–348.

[32] O. Goldreich, T. Kaufman, *Proximity oblivious testing and the role of invariances*, in Studies in complexity and cryptography, 173–190, Vol. 6650 of Lecture Notes in Comput. Sci., Springer, 2011.

[33] O. Goldreich, L. Trevisan, *Three theorems regarding testing graph properties*, Random Structures Algorithms 23 (2003), 23–57; also FOCS'01: Proc. 42th Annu. Symp. Foundations of Computer Science, IEEE, 2001, 460–469.

[34] W. T. Gowers, *Lower bounds of tower type for Szemerédi's uniformity lemma*, Geom. Funct. Anal. 7 (1997), 322–337.

[35] G. Higman, *Ordering by divisibility in abstract algebras*, Proc. London Math. Soc. (3) 2 (1952), 326–336.

[36] S. Kalyanasundaram, A. Shapira, *A wowzer-type lower bound for the strong regularity lemma*, Proc. Lond. Math. Soc. (3) 106 (2013), 621–649.

[37] L. Lovász, *Large networks and graph limits*, American Mathematical Society, Providence, RI, 2012.

[38] L. Lovász, B. Szegedy, *Szemerédi's Lemma for the Analyst*, Geom. Funct. Anal. 17 (2007), 252–270.

[39] L. M. Lovász, J. Fox, *A tight lower bound for Szemerédi's regularity lemma*, Combinatorica, to appear.

[40] G. Moshkovitz, A. Shapira, *A short proof of Gowers' lower bound for the regularity lemma*, Combinatorica 36 (2016), 187–194.

[41] B. Nagle, V. Rödl, M. Schacht, *The counting lemma for regular k-uniform hypergraphs*, Random Structures & Algorithms 28 (2006), 113–179.

[42] S. Raskhodnikova, *Approximate testing of visual properties*, Approximation, randomization, and combinatorial optimization, 370–381, Lecture Notes in Comput. Sci., 2764, Springer, 2003.

[43] R. Rubinfeld, M. Sudan, *Robust characterizations of polynomials with applications to program testing*, Siam J. Comput. 25 (1996), 252–271.

[44] V. Rödl, J. Skokan, *Regularity lemma for uniform hypergraphs*, Random Structures & Algorithms 25 (2004), 1–42.

[45] E. Szemerédi, *Regular partitions of graphs*, Proc. Colloque Inter. CNRS No. 260 (J. C. Bermond, J. C. Fournier, M. Las Vergnas and D. Sotteau, eds.), 1978, 399–401.

[46] M. Sudan, *Invariance in property testing*, in Property Testing: Current Research and Surveys (O. Goldreich, ed.), Vol. 6390 of Lecture Notes in Computer Science, Springer, 2010, 211–227.

[47] T. Tao, *A variant of the hypergraph removal lemma*, J. Combin. Theory Ser. A 113 (2006), 1257–1280.

[48] J. van Leeuwen, *Graph Algorithms*, Handbook of Theoretical Computer Science (J. van Leeuwen Ed.), vol. A , 525–631, MIT Press, 1990.