Available online at www.sciencedirect.com

ScienceDirect

Electronic Notes in DISCRETE MATHEMATICS

www.elsevier.com/locate/endm

Counting configuration–free sets in groups

Juanjo Rué [1,4]

Institut für Mathematik und Informatik
Freie Universität Berlin
Arnimallee 3, 14195 Berlin, Germany

Oriol Serra [2,5]

Departament de Matemàtica Aplicada 4
Universitat Politècnica de Catalunya
Jordi Girona 1, E-08034, Barcelona, Spain

Lluis Vena [3,6]

Computer Science Institute (IUUK and ITI)
Charles University
Malostranské námesti 25, 11800 Praha 1, Czech Republic

Abstract

We present a unified framework to asymptotically count the number of sets, with a given cardinality, free of certain configurations. This is done by combining the hypergraph containers methodology joint with arithmetic removal lemmas. Several applications involving linear configurations are described, as well as some applications in the random sparse setting.

Keywords: Arithmetic Removal Lemma, Hypergraph Container, Arithmetic Combinatorics

1 Introduction

The study of sparse (and probabilistic) analogues of well-known results in extremal combinatorics have become a very active area of research in extremal and random combinatorics (see e.g. the survey by Conlon [2].) One starting point is the *Szemerédi Theorem* [23] on the existence of arbitrarily long arithmetic progressions in sets of integers with positive upper density. This seminal result and the tools arising in its many proofs have been enormously influential in the development of modern discrete mathematics. Additionally, nowadays a large proportion of the research in additive combinatorics is inspired on these achievements.

Sparse analogues of Szemerédi Theorem started in Kohayakawa, Rödl and Łuczack [10] by giving the threshold probability for a random set of the integer interval $[1, n]$ whose subsets of given density contain a.a.s. 3–term arithmetic progressions. The extension of the result to k–term arithmetic progressions was a breakthrough obtained independently, and by different methods, by Conlon and Gowers [3] and by Schacht [19].

Recently, the problem has found a new solution using combinatorial arguments independently by Saxton, and Thomason [18] and by Balogh, Morris, and Samotij [1]. The approach in the above two papers is based on a methodology building on the structure of independent sets in hypergraphs. *Hypergraphs containers* (as it is named in [18]) provides a general framework to attack a wide variety of problems, building on seminal results of Kleitman and Winston for graphs without cycles of length four [9]. The philosophy behind this method is that, for a wide variety of uniform hypergraphs which satisfy mild conditions, one can find a small collection of sets of vertices (which are called *containers*) which contain all independent sets of the given hypergraph, thus providing sensible upper bounds on the number of independent sets. By using appropriate models, solutions of systems of equations or configurations are represented by independent sets in hypergraphs leading to arithmetic appli-

[1] J. R. was partially supported by the FP7-PEOPLE-2013-CIG project *CountGraph* (ref. 630749), the DFG within the research training group *Methods for Discrete Structures* (GRK1408), the collateral PROCOPE-DAAD project *RanConGraph* (ref. 57134837) and the *Berlin Mathematical School.*

[2] O. S. was supported by the Spanish Research Council under project MTM2011-28800-C02-01.

[3] L. V. was supported by the ERC-CZ project *CORES* (ref. LL1201).

[4] Email: `jrue@zedat.fu-berlin.de`

[5] Email: `oriol.serra@upc.edu`

[6] Email: `lluis.vena@gmail.com`

cations of the containers approach.

In addition to important applications in combinatorics, the two papers above mentioned also contain arithmetic applications, providing in particular a new proof of the sparse Szemerédi Theorem. One important ingredient of these proofs, explicitely exposed in [1], is the so-called *Varnavides Theorem* [25]. This is the robust counterpart of Szemerédi Theorem: once a set has positive density, it does not only have one but a positive proportion of the total number of k–term arithmetic progressions. Nowadays there is a rich theory dealing with this type of results, which are rephrased under the name of *Arithmetic Removal Lemmas*. The idea behind them can be traced back to the proof of Roth's Theorem by Ruzsa and Szemerédi [17] and was formulated by Green [8] for a linear equation in an abelian group by using methods of Fourier analysis. The picture was complemented independently by Shapira [20] and by Král, Serra and Vena [13] by proving a removal lemma for linear systems in the integers. These results have been extended in several directions, including arithmetic removal lemmas for a single equation in non-abelian groups, for linear systems over finite fields and for integer linear systems over finite abelian groups (see [12,20,13,14]).

These extensions of Green's Arithmetic Removal Lemma provide proofs of the Szemerédi Theorem in general abelian groups (see also Szegedy's [22]), but cannot handle the robust versions of the multidimensional Szemerédi Theorem (see for instance Solymosi's [21] on Furstenberg and Katznelson's [7]) or, more generally, the appearance and enumeration of finite configurations in dense subsets in abelian groups (as seen in Tao [24, Theorem B.1]). As a consequence, the above mentioned arithmetic removal lemmas cannot be used to show the sparse counterparts of these results (see [1,3,19]).

In this work we develop a general framework to analyze these problems by combining the containers methodology together with robust versions of arithmetic removal lemmas that encompass those previously mentioned finite configurations in abelian groups [26, Theorem 2].

The main tool is Theorem 3.1 (Section 3) which provides an abstract counting result for the number of sets which are free of given systems of configurations in groups (see Section 2 for the definitions and terminology.) Section 4 contains the main contribution of the work. We describe applications of the general framework provided by Theorem 3.1 to natural examples of configurations in abelian groups as well as solutions of single equations in nonabelian groups. The configurations studied are novel and a consequence of arithmetic removal lemmas from homomorphisms over groups (see [26]).

2 Definitions: system of configurations

Let us give the main notion we study.

Definition 2.1 [System of configurations] Let k be a positive integer, let G be a finite set and let A be a map from $G^k = G \times \cdots \times G$ to $\{0, 1\}$. Then the pair (A, G) is said to be a *system of configurations* of degree k. We say that $(g_1, \ldots, g_k) \in G^k$ is a *solution of the system* if $A(g_1, \ldots, g_k) = 1$. We write $S(A, G) = A^{-1}(1)$ for the solution set.

We also denote by $S^j(A, G)$ the subset of $S(A, G)$ where the solutions $\mathbf{g} = (g_1, \ldots, g_k)$ have j different values, namely, the set $\{g_1, \ldots, g_k\}$ has cardinality j.

For a given set $S = \{s_1, \ldots, s_m\} \subset [k]$, let π_S denote the projection $\pi_S : G^k \to G^m$ with $(g_1, \ldots, g_k) \mapsto (g_{s_1}, \ldots, g_{s_m})$. In other words, π_S keeps the coordinates indexed by the elements in S. For $i \in [k]$, let us define the *i-th (A, G)-degree of freedom* as the following quantity associated to the configuration system (A, G):

$$\alpha_i = \max_{\substack{B \subset [k] \\ |B| = i}} \max_{(x_1, \ldots, x_i) \in G^i} \left[S(A, G) \cap \pi_B^{-1}(x_1, \ldots, x_i) \right].$$

Additionally, we define the *restricted i-th (A, G)-degree of freedom* as the quantity

$$\alpha_i^k = \max_{\substack{B \subset [k] \\ |B| = i}} \max_{(x_1, \ldots, x_i) \in G^i} \left[S^k(A, G) \cap \pi_B^{-1}(x_1, \ldots, x_i) \right].$$

The following definition is inspired by Varnavides Theorem [25], which gives a robust version of Roth's Theorem [15]:

Definition 2.2 [Varnavides property, V-property] The system of configurations (A, G) of degree k is said to fulfill the Varnavides property, or V-property, with function γ if, for every $\epsilon > 0$, there exists a $\gamma = \gamma(\epsilon, A) > 0$ such that, for any $B \subset G$ with $|B| \geq \epsilon |G|$, then

$$|B^k \cap S(A, G)| \geq \lfloor \gamma |S(A, G)| \rfloor.$$

Given a set of indices I, a family of systems $\{(A_i, G_i)\}_{i \in I}$ is said to satisfy the V-property if γ is the same function for each member of the family and only depends on ϵ, $\gamma = \gamma(\epsilon)$.

3 Main tool

Theorem 3.1 represents a unified approach to count the number of solution–free sets for systems of configurations.

Theorem 3.1 (Counting independent sets for systems of configurations) *Let k be a fixed positive integer and $\delta > 0$. Let (A, G) be a system of configurations of degree k satisfying the V-property with function $\gamma = \gamma(\delta, A)$. Write $n = |G|$. For each $i \in [1, k]$, let α_i^k be the restricted i-th (A, G)-degree of freedom.*

Assume that each subset of G with more than $\frac{\delta n}{2}$ elements contains a configuration in $S^k(A, G)$. Assume that $\xi = (\gamma - 1)|S(A, G)| + |S^k(A, G)| > 0$. Then for each t such that

$$t \geq C\frac{|G|}{\delta} \max_{\ell \in [2,k]} \left\{ \left(\frac{\alpha_\ell^k}{\alpha_1^k} \frac{1}{k} \binom{k}{l} \right)^{\frac{1}{l-1}} \right\} \text{ and } t \leq \frac{\delta n}{2}$$

where $C = C(k, \epsilon, c)$ is the C appearing in [1, Theorem 2.2] evaluated at

$$\epsilon = \frac{\xi}{|S^k(A, G)|} \text{ and } c = \alpha_1^k \frac{(k-1)!\ |G|}{|S^k(A, G)|},$$

there are

$$t \left[\frac{2e}{\delta^2} \right]^{\delta t} \binom{\delta n}{t}$$

sets of size t with no solution in $S^k(A, G)$. If we assume that $\delta = \min\{\beta/2, 1/40\}$, then the bound can be rewritten as

$$\binom{\beta n}{t}.$$

Theorem 3.1 reads as follows: given a system of configurations (A, G) satisfying the V-property and extra mild natural conditions, we can obtain upper bounds for the number of sets of a given cardinality with no configuration in $S^k(A, G)$.

The proof of Theorem 3.1 involves the construction of an appropriate hypergraph in which the independent sets correspond to sets free of configurations of the given system (A, G). The construction is set up to allow for an application of [1, Theorem 2.2].

Most of the system of configurations in the applications have some common features. Each example consists of a family of configuration systems

$\{(A_i,G_i)\}_{i\in\mathbb{N}}$ where each G_i is growing in size with i and $\lim_{i\to\infty}|G_i|/|S(A,G_i)|=0$. We also require that $\lim_{i\to\infty}\frac{|S^k(A,G_i)|}{|S(A,G_i)|}=1$, and

$$\alpha_1^k\frac{(k-1)!\,|G_i|}{|S^k(A,G_i)|} \tag{1}$$

and $\gamma(\delta,A_i)$ are, respectively, upper and lower bounded uniformly for the whole family. Condition (1) is technical and necessary in the proof of the Theorem 3.1. Observe that the condition $\chi=(\gamma-1)|S(A,G)|+|S^k(A,G)|>0$ in the statement assures that the V-property ensures solutions in $S^k(A,G)$.

4 Applications and further research

So far, authors have found results on the existence of configuration in subsets by studying systems of configurations arising from integer linear systems in $[1,n]$ invariant by translations [19,3], integer linear systems over abelian groups [18] or linear systems over finite fields [18]. These generalize the case for k-term arithmetic progressions [1,3,20,18]. One of the key ingredients in the proofs of these results that provides the V-property are arithmetic removal lemmas for the appropriate context, such as the ones found in [4,8,20,14,13].

All these systems of configurations can be seen as prominent particular cases of homomorphisms of finite abelian groups, context in which an arithmetic removal lemma can be found in [26, Theorem 2]; these include linear homothetic-to-a-point configurations in products of finite abelian groups. The framework of homomorphisms also includes the configurations from the multidimensional Szemerédi setting [6,24], some of which have been treated in [1,19].

The following theorem illustrates an application of Theorem 3.1 which can not be directly obtained form the previously existing tools.

Theorem 4.1 (Rectangles in abelian groups) *Let $\{G_i\}_{i\in\mathbb{N}}$ be a sequence of finite abelian groups, H_i,K_i subgroups of G_i and such that $|H_i|,|K_i|,|G_i|\to\infty$ $(i\to\infty)$.*

For each $\delta>0$ with $\delta<1/40$ there exist a $C=C(\delta)$ and an i_0, depending on the family $\{G_i,H_i,K_i\}_{i\in\mathbb{N}}$ and on δ, for which the following holds.

Let

$$S(A,G_i)=\{(x,x+a,x+b,x+a+b)\ \text{with}\ x\in G_i, a\in H_i, b\in K_i\}$$

be the set of configurations. Assume that $\max\{|H_i|,|K_i|\}\le(|S^k(A,G_i)|/|G_i|)^{2/3}$.

For each $i \geq i_0$ the number of sets free of configurations in $S^k(A, G)$ and with cardinality t such that

$$t > \frac{C}{\delta} \left(\frac{|G_i|^4}{|S^k(A, G_i)|} \right)^{1/3}$$

is bounded from above by $\binom{2\delta|G_i|}{t}$.

When $G_i = \mathbb{Z}_i^2$, $H_i = \mathbb{Z}_i \times \{0\}$, $K_i = \{0\} \times \mathbb{Z}_i$, then the bound on the size t in Theorem 4.1 is, for i large enough, $t > C'' i^{4/3}$. The configurations described in Theorem 4.1 correspond to the edge set of a C_4–bipartite graph, the extremal value of which is $\approx i^{3/2}$ as proved in Füredi and Hajnal [5]. This shows that there is a wide range of values of t to which Theorem 4.1 gives a meaningful bound. More generally, we can consider larger configurations whose bounds for the configuration–free sets are connected to the Zarankiewicz problem [5,11].

Other similar configurations that can be treated are the following ones. Let G be a finite abelian group, G_1 a subgroup of G and $\phi : G_1 \to G$ an injective group homomorphism with $a \neq \pm\phi(a)$ for each $a \in G_1$. The configuration set is the "slanted squares" $\{x, x + a, x + \phi(a), x + a + \phi(a) \ : \ x \in G, a \in G_1\}$.

The same framework of Theorem 3.1 applies to count the number of solution–free sets of equations of the form $x_1 \cdot \cdots \cdot x_k = 1$ in finite non-necessarily abelian groups by using the results in [12].

Another important application of Theorem 3.1 combined with the removal lemma for group homomorphisms is the study of threshold functions for analogues of combinatorial theorems holding in random subsets. More precisely, for given $\delta > 0$ and a system of configurations (A, G), we say that a set $B \subseteq G$ has *the* (δ, A, G)-*Property* if for all subset $B' \subset B$, $|B'| \geq \delta|B|$ $B' \cap S(A, G) \neq \emptyset$. We can study the previous property in the following random model: fix a probability p (that may depend on $|G|$), and consider the binomial random set B_p built by choosing independently each element of G with probability p. By means of the bounds given by Theorem 3.1 we can obtain estimates for the threshold value for p with respect to the (δ, A, G)-Property. These results are in the line of [19, Theorem 2.2;2.3;2.4] (see also [2]).

All these applications are described in the forthcoming paper [16].

References

[1] Balogh, J., R. Morris and W. Samotij, *Independent sets in hypergraphs*, to appear at *Journal of the American Mathematical Society.*

[2] Conlon, D., *Combinatorial theorems relative to a random set*, to appear at *Proceedings of the International Congress of Mathematicians 2014.* Available on-line at *arXiv:1404.3324.*

[3] Conlon, D. and W. Gowers, *Combinatorial theorems in sparse random sets*, submitted. Available on-line at *arXiv:1011.4310v1.*

[4] Frankl, P. and V. Rödl, *Extremal problems on set systems*, Random Structures Algorithms **20** (2002), pp. 131–164.

[5] Füredi, Z. and P. Hajnal, *Davenport-Schinzel theory of matrices*, Discrete Math. **103** (1992), pp. 233–251.

[6] Furstenberg, H., *Ergodic behavior of diagonal measures and a theorem of Szemerédi on arithmetic progressions*, J. Analyse Math. **31** (1977), pp. 204–256.

[7] Furstenberg, H. and Y. Katznelson, *An ergodic szemerédi theorem for commuting transformations*, J. Anal. Math. **34** (1978), pp. 275–291 (1979).

[8] Green, B., *A Szemerédi-type regularity lemma in abelian groups, with applications*, Geom. Funct. Anal. **15** (2005), pp. 340–376.

[9] Kleitman, D. J. and K. J. Winston, *On the number of graphs without 4-cycles*, Discrete Mathematics **41** (1982), pp. 167 – 172.

[10] Kohayakawa, Y., T. Luczak and V. Rödl, *Arithmetic progressions of length three in subsets of a random set*, Acta Arith. **75** (1996), pp. 133–163.

[11] Kövari, T., V. T. Sós and P. Turán, *On a problem of K. Zarankiewicz*, Colloquium Math. **3** (1954), pp. 50–57.

[12] Král, D., O. Serra and L. Vena, *A combinatorial proof of the removal lemma for groups*, J. Combin. Theory Ser. A **116** (2009), pp. 971–978.

[13] Král, D., O. Serra and L. Vena, *A removal lemma for systems of linear equations over finite fields*, Israel J. Math. **187** (2012), pp. 193–207.

[14] Král, D., O. Serra and L. Vena, *On the removal lemma for linear systems over abelian groups*, European J. Combin. **34** (2013), pp. 248–259.

[15] Roth, K. F., *On certain sets of integers*, J. London Math. Soc. **28** (1953), pp. 104–109.

[16] Rué, J., O. Serra and L. Vena, *Counting configuration–free sets in groups.*, in preparation.

[17] Ruzsa, I. Z. and E. Szemerédi, *Triple systems with no six points carrying three triangles*, in: *Combinatorics (Proc. Fifth Hungarian Colloq., Keszthely, 1976), Vol. II*, Colloq. Math. Soc. János Bolyai **18**, 1978 pp. 939–945.

[18] Saxton, D. and A. Thomason, *Hypergraph containers*, submitted. Available online at *arXiv:1204.6595v2*.

[19] Schacht, M., *Extremal results for random discrete structures*, submitted. URL http://www.math.uni-hamburg.de/home/schacht

[20] Shapira, A., *A proof of Green's conjecture regarding the removal properties of sets of linear equations*, J. Lond. Math. Soc. (2) **81** (2010), pp. 355–373.

[21] Solymosi, J., *A note on a question of erdos and graham*, Combin. Probab. Comput. **13** (2004), pp. 263–267.

[22] Szegedy, B., *The symmetry preserving removal lemma*, Proc. Amer. Math. Soc. **138** (2010), pp. 405–408.

[23] Szemerédi, E., *On sets of integers containing no k elements in arithmetic progression*, Acta Arith. **27** (1975), pp. 199–245.

[24] Tao, T., *Mixing for progressions in nonabelian groups*, Forum Math. Sigma **1** (2013), pp. e2, 40.

[25] Varnavides, P., *On certain sets of positive density*, J. London Math. Soc. **34** (1959), pp. 358–360.

[26] Vena, L., *On the removal lemma for linear configurations in finite abelian groups*, arXiv:1505.02280v1 [math.CO] .