Proceedings of the 12th European Conference on Combinatorics, Graph Theory and Applications
EUROCOMB'23
Prague, August 28 - September 1, 2023

The dimension of the feasible region of pattern densities

(Extended abstract)

Frederik Garbe* Daniel Král'* Alexandru Malekshahian† Raul Penaguiao‡

Abstract

A classical result of Erdős, Lovász and Spencer from the late 1970s asserts that the dimension of the feasible region of homomorphic densities of graphs with at most k vertices in large graphs is equal to the number of connected graphs with at most k vertices. Glebov et al. showed that pattern densities of indecomposable permutations are independent, i.e., the dimension of the feasible region of densities of k-patterns is at least the number of non-trivial indecomposable permutations of size at most k. We identify a larger set of permutations, which are called Lyndon permutations, whose pattern densities are independent, and show that the dimension of the feasible region of densities of k-patterns is equal to the number of non-trivial Lyndon permutations of size at most k.

DOI: `https://doi.org/10.5817/CZ.MUNI.EUROCOMB23-065`

1 Introduction

A classical result of Erdős, Lovász and Spencer [8] describes the independence of homomorphic densities of graphs in large graphs. Informally speaking, they showed that homomorphic densities of connected graphs are independent and actually determine the densities of

*Faculty of Informatics, Masaryk University, Botanická 68A, 602 00 Brno, Czech Republic. E-mail: `{garbe,dkral}@fi.muni.cz`. Supported by the MUNI Award in Science and Humanities (MUNI/I/1677/2018) of the Grant Agency of Masaryk University.

†Department of Mathematics, King's College London. E-mail: `alexandru.malekshahian@kcl.ac.uk`.

‡Max Planck Institute for the Sciences, Inselstraße 22, 04103 Leipzig, Germany. E-mail: `raul.penaguiao@mis.mpg.de`.

all graphs. We now state their result formally using the language of the theory of graph limits (referring to Section 2 for definitions). Let $\mathcal{G}_k$ be the set of all graphs with at most k vertices and $\mathcal{G}_k^c$ be the set of all connected graphs with at most k vertices; $t(G,W)$ denotes the homomorphism density of a graph G in a graphon W. The aforementioned result of Erdős, Lovász and Spencer [8] asserts that for every $k \in \mathbb{N}$, there exist $x_0 \in [0,1]^{\mathcal{G}_k^c}$ and $\varepsilon > 0$ such that for every $x \in B_\varepsilon(x_0) \subseteq [0,1]^{\mathcal{G}_k^c}$, there exists a graphon W such that $t(G,W)_{G\in\mathcal{G}_k^c} = x$. In addition, there exists a function $f : [0,1]^{\mathcal{G}_k^c} \to [0,1]^{\mathcal{G}_k}$, independent of W, and such that $f(t(G,W)_{G\in\mathcal{G}_k^c}) = t(G,W)_{G\in\mathcal{G}_k}$. In other words, the dimension of the feasible region of homomorphic densities of graphs with at most k vertices in graphons (large graphs) is equal to the number of connected graphs with at most k vertices.

We determine the dimension of the feasible region of densities of k-patterns in permutations; again we refer to Section 2 for definitions. Glebov et al. [10] showed that this dimension is at least the number of non-trivial indecomposable permutations of size at most k. Borga and the last author [2] observed utilizing a result of Vargas [20] that this dimension is at most the number of non-trivial Lyndon permutations of size at most k, and conjectured [2, Conjecture 1.3] that this bound is tight. Our main result asserts that this is indeed the case. Similarly to [10], our argument is based on perturbing a permuton comprised of blow-ups of indecomposable permutations. However, to be able to control the densities of the larger set of all Lyndon permutations, we choose a suitable order of the blow ups of indecomposable permutations and analyze the interplay between the blow-ups using unique decomposition properties into Lyndon words [19].

2 Combinatorial limits

We now introduce notation used throughout this extended abstract. In addition to the monograph by Lovász [16], which provides a comprehensive introduction to the theory of graph limits, we refer the reader to [3–5, 17, 18] for basic results concerning graph limits and to [1, 6, 9, 11–15] for results developing and concerning permutation limits.

2.1 Graph limits

If H and G are two graphs, the *homomorphism density* of H in G, denoted by $t(H,G)$, is the probability that a uniformly random function $f : V(H) \to V(G)$, is a *homomorphism* of H to G. A sequence $(G_n)_{n\in\mathbb{N}}$ of graphs is *convergent* if the number of vertices of G_n tends to infinity and the values of $t(H,G_n)$ converge for every H.

A *graphon* is a symmetric measurable function $W : [0,1]^2 \to [0,1]$, i.e., $W(x,y) = W(y,x)$ for $(x,y) \in [0,1]^2$. The *homomorphism density* of a graph H in a graphon W is defined by

$$t(H,W) = \int_{[0,1]^{V(H)}} \prod_{uv\in E(H)} W(x_u,x_v)\mathrm{d}x_{V(H)}.$$

A graphon W is a *limit* of a convergent sequence $(G_n)_{n\in\mathbb{N}}$ of graphs if $t(H,W)$ is the limit of $t(H,G_n)$ for every graph H. Every convergent sequence of graphs has a limit graphon and every graphon is a limit of a convergent sequence of graphs as shown by Lovász and Szegedy [17]; also see [7] for a relation to exchangeable arrays.

2.2 Permutations

A *permutation* of *size* n is a bijective function π from $[n]$ to $[n]$ (we use $[n]$ to denote the set of the first n positive integers). The permutation π is often viewed as a word $\pi(1)\pi(2)\cdots\pi(n)$ and its size is denoted by $|\pi|$. The *pattern* induced by elements $1 \le k_1 < \cdots < k_m \le n$ is the unique permutation $\sigma : [m] \to [m]$ such that $\sigma(i) < \sigma(i')$ if and only if $\pi(k_i) < \pi(k_{i'})$ for all $i, i' \in [m]$. The *density* of a permutation σ in a permutation π, denoted by $d(\sigma, \pi)$, is the probability that the pattern induced by $|\sigma|$ elements chosen uniformly at random is equal to σ. Similarly to the graph case, we say that a sequence $(\pi_n)_{n\in\mathbb{N}}$ of permutations is *convergent* if the sizes of π_n tend to infinity and the sequence of densities $d(\sigma, \pi_n)$ converges for every permutation σ.

We say that a permutation is *non-trivial* if its size is at least two. The *direct sum* of two permutations π_1 and π_2 is the permutation π of size $|\pi_1| + |\pi_2|$ such that $\pi(k) = \pi_1(k)$ for $k \in [|\pi_1|]$ and $\pi(|\pi_1| + k) = |\pi_1| + \pi_2(k)$ for $k \in [|\pi_2|]$; the permutation π is denoted by $\pi_1 \oplus \pi_2$. A permutation is *indecomposable* if it is not a direct sum of two permutations; note that every permutation is a (possibly iterated) direct sum of indecomposable permutations.

A word $w_1 \cdots w_n$ is *Lyndon* if no proper suffix of the word $w_1 \cdots w_n$ is smaller (in the lexicographic order) than the word $w_1 \cdots w_n$ itself. For example, the word *aab* is Lyndon but the word *aba* is not. We want to use indecomposable permutations as the alphabet to form Lyndon words. For this we introduce an order $\prec$ on the set of indecomposable permutations such that indecomposable permutations of smaller size precede those of larger size. Indecomposable permutations of the same size are ordered lexicographically. Hence, the first five letters are associated with the following five (indecomposable) permutations: $1 \prec 21 \prec 231 \prec 312 \prec 321$. As mentioned above every permutation can be uniquely decomposed into a direct sum of indecomposable permutations and therefore corresponds to a word over the alphabet consisting of indecomposable permutations. A permutation π is *Lyndon* if the word corresponding to the decomposition of π into indecomposable permutations is Lyndon. For example, the permutation $21 \oplus 231 = 21453$ is Lyndon but the permutations $21 \oplus 1 = 213$ and $21 \oplus 21 = 2143$ are not. Note that all indecomposable permutations are Lyndon.

2.3 Permutation limits

A *permuton* is a probability measure Π on the σ-algebra of Borel subsets from $[0,1]^2$ that has uniform marginals, i.e.,

$$\Pi([a,b] \times [0,1]) = \Pi([0,1] \times [a,b]) = b - a$$

for all $0 \le a \le b \le 1$. A Π-*random permutation* of size n is the permutation σ obtained by sampling n points according to the measure Π, sorting them according to their x-coordinates, say $(x_1, y_1), \ldots, (x_n, y_n)$ for $x_1 < \cdots < x_n$ (note that the x-coordinates are pairwise distinct with probability 1), and defining σ so that $\sigma(i) < \sigma(j)$ if and only if $y_i < y_j$ for $i, j \in [n]$. Finally, the *density* of a permutation σ in a permuton Π, which is denoted by $d(\sigma, \Pi)$, is the probability that the Π-random permutation of size $|\sigma|$ is σ.

A permuton Π is a *limit* of a convergent sequence $(\pi_n)_{n\in\mathbb{N}}$ of permutations if, for every permutation σ, $d(\sigma, \Pi)$ is the limit of $d(\sigma, \pi_n)$. Every permuton is a limit of a convergent sequence of permutations and every convergent sequence of permutations has a limit permuton [11, 12].

3 Main result

Let $\mathcal{P}_k$ be the set of all permutations of size at most k, $\mathcal{P}_k^L$ the set of all non-trivial Lyndon permutations of size at most k. Our main result is the following.

Theorem 1. *For every $k \in \mathbb{N}$, there exists $x_0 \in [0,1]^{\mathcal{P}_k^L}$ and $\varepsilon > 0$ such that for every $x \in B_\varepsilon(x_0) \subseteq [0,1]^{\mathcal{P}_k^L}$ there exists a permuton Π such that*

$$d(\sigma, \Pi)_{\sigma\in\mathcal{P}_k^L} = x.$$

In addition, there exists a function $f : [0,1]^{\mathcal{P}_k^L} \to [0,1]^{\mathcal{P}_k}$ such that

$$f\left(d(\sigma, \Pi)_{\sigma\in\mathcal{P}_k^L}\right) = d(\sigma, \Pi)_{\sigma\in\mathcal{P}_k}$$

for every permuton Π.

We next sketch the proof of Theorem 1. We start with the existence of the function f; we remark that the existence of the function f follows from the results presented in the extended abstract [20], and we outline the argument here. Let π be a permutation and let $\pi = \pi_1 \oplus \cdots \oplus \pi_k$ be the (unique) direct sum formed by indecomposable permutations. Further, let $w_1 \cdots w_k$ be the word corresponding to $\pi_1 \oplus \cdots \oplus \pi_k$; it is well-known that the word $w_1 \cdots w_k$ can be *uniquely* expressed as a concatenation of Lyndon words in non-increasing lexicographic order, and let $\pi'_1, \ldots, \pi'_\ell$ be the permutations corresponding to these Lyndon words. For example, if $\pi = 1324576 = 1 \oplus 21 \oplus 1 \oplus 1 \oplus 21$, then π'_1 is $1 \oplus 21 = 132$ and π'_2 is $1 \oplus 1 \oplus 21 = 1243$ which are both Lyndon. It can be shown using [19, Theorem 3.1.1(a)] that the constituents of the product of $\pi'_1 \times \ldots \times \pi'_\ell$ (in the flag algebra sense) are only permutations that either are direct sums of fewer than k indecomposable permutations or are direct sums of k indecomposable permutations but are lexicographically at least as large as π. It follows that every permutation σ that is not Lyndon can be expressed as a polynomial of Lyndon permutations of size at most $|\sigma|$ (in the flag algebra sense), which implies the existence of the function f; in fact, the function f is polynomial.

We next sketch the proof of the main part of Theorem 1, which yields the (matching) lower bound on the dimension on the feasible region of pattern densities. For the lower bound, we use a different mapping of indecomposable permutations to letters; note that this

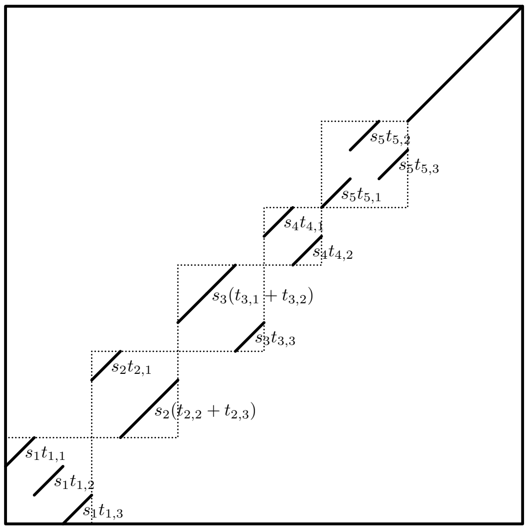

Figure 1: The permuton Π comprised of the "blow-up permutons" of the permutations 321, 312, 231, 21 and 132; the scaling factors s_i and $t_{i,j}$ are placed near their associated parts.

changes which permutations are Lyndon. The *compression* of a permutation π, which is denoted by $\widehat{\pi}$, is the permutation obtained by (iteratively) "merging" consecutive elements that increase by one; for example $\widehat{231} = 21$, $\widehat{3412} = 21$, $\widehat{2341} = 21$, and $\widehat{1342} = 132$. The new order $<$ on indecomposable permutations is defined using $\prec$ on their compressions, and if two different indecomposable permutations have the same compression, then $\prec$ is used directly. For example, $3412 < 321$, and so the letter associated with 3412 precedes the letter associated with 321. Note that while the permutation $321 \oplus 3412 = 3216745$ is Lyndon with respect to $\prec$ it is not with respect to $<$. However, it can be shown that the number of Lyndon permutations of size k is the same with respect to $\prec$ and to $<$.

Fix k and let $\pi_1, \dots, \pi_N$ be all non-trivial Lyndon permutations of size at most k listed in the decreasing (lexicographic) order of the words corresponding to their indecomposable blocks; we emphasize that the modified order $<$ is used both to define which permutations are Lyndon and to order the Lyndon permutations. For $s_1, \dots, s_N \in [0,1]$ and $t_{i,j} \in [0,1]$, $i \in [N]$ and $j \in [|\pi_i|]$ such that the sum of $t_{i,j}$'s is at most one, we define a permuton Π to be the permuton comprised of the "blow-up permutons" of the permutations $\pi_1, \dots, \pi_N$. For each $i \in [N]$ the "blow-up permuton" uses a segment of horizontal length $t_{i,j}$ corresponding to the j'th point of the permutation π_i, $j \in [|\pi_i|]$. The "blow-up permutons" then get scaled by $s_1, \dots, s_N$, respectively; see Figure 1 for illustration. We next consider the Jacobian matrix of the densities $d(\pi_1, \Pi), \dots, d(\pi_N, \Pi)$ viewed as functions of $s_1, \dots, s_N$ and observe that its determinant is a polynomial in the variables s_i and $t_{i,j}$ and the coefficient of the monomial formed by the product of all $t_{i,j}$ is non-zero; the latter is argued by making use of [19, Theorem 3.1.1(a)]. Hence, the Jacobian determinant is not identically zero and so there exists a choice of s_i and $t_{i,j}$ such that the determinant is non-zero, which implies the existence of the point $x_0 \in [0,1]^{\mathcal{P}_k^L}$ and the real $\varepsilon > 0$ from the statement of Theorem 1.

References

[1] J. Balogh, P. Hu, B. Lidický, O. Pikhurko, B. Udvari and J. Volec: *Minimum number of monotone subsequences of length 4 in permutations*, Combin. Probab. Comput. **24** (2015), 658–679.

[2] J. Borga and R. Penaguiao: *The feasible regions for consecutive patterns of pattern-avoiding permutations*, Discrete Math. **346** (2023), 113219.

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

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

[5] C. Borgs, J. T. Chayes, L. Lovász, V. T. Sós and K. Vesztergombi: *Convergent sequences of dense graphs II. Multiway cuts and statistical physics*, Ann. of Math. (2) (2012), 151–219.

[6] T. Chan, D. Král', J. A. Noel, Y. Pehova, M. Sharifzadeh and J. Volec: *Characterization of quasirandom permutations by a pattern sum*, Random Struct. Algor. **57** (2020), 920–939.

[7] P. Diaconis and S. Janson: *Graph limits and exchangeable random graphs*, Rend. Mat. Appl. **28** (2008), 33–61.

[8] P. Erdős, L. Lovász and J. Spencer: *Strong independence of graphcopy functions*, Graph theory and related topics (1979), 165–172.

[9] R. Glebov, A. Grzesik, T. Klimošová and D. Král': *Finitely forcible graphons and permutons*, J. Combin. Theory Ser. B **110** (2015), 112–135.

[10] R. Glebov, C. Hoppen, T. Klimošová, Y. Kohayakawa, D. Král' and H. Liu: *Densities in large permutations and parameter testing*, European J. Combin. **60** (2017), 89–99.

[11] C. Hoppen, Y. Kohayakawa, C. G. T. de A. Moreira, B. Ráth and R. M. Sampaio: *Limits of permutation sequences*, J. Combin. Theory Ser. B **103** (2013), 93–113.

[12] C. Hoppen, Y. Kohayakawa, C. G. T. de A. Moreira and R. M. Sampaio: *Testing permutation properties through subpermutations*, Theor. Comput. Sci. **412** (2011), 3555–3567.

[13] R. Kenyon, D. Král', C. Radin and P. Winkler: *Permutations with fixed pattern densities*, Random Struct. Algor. **56** (2020), 220–250.

[14] D. Král' and O. Pikhurko: *Quasirandom permutations are characterized by 4-point densities*, Geom. Funct. Anal. **23** (2013), 570–579.

[15] M. Kurečka: *Lower bound on the size of a quasirandom forcing set of permutations*, Combin. Probab. Comput. **31** (2022), 304–319.

[16] L. Lovász: Large Networks and Graph Limits, *Colloquium Publications*, volume 60, 2012.

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

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

[19] D. A. Radford: *A natural ring basis for the shuffle algebra and an application to group schemes*, J. Algebra **58** (1979), 432–454.

[20] Y. Vargas: *Hopf algebra of permutation pattern functions*, DMTCS Proceedings vol. AT, 26th International Conference on Formal Power Series and Algebraic Combinatorics (FPSAC 2014) (2014), 839–850.