MASARYK UNIVERSITY

Quasirandomness in combinatorics

Filip Kučerák

Master's Thesis

Faculty of Informatics

Brno, Spring 2025

Supervisor: prof. RNDr. Daniel Kráľ, Ph.D., DSc.

Declaration

Hereby I declare that this thesis is my original authorial work, which I have worked out on my own. All sources, references, and literature used or excerpted during elaboration of this work are properly cited and listed in complete reference to the due source.

During the preparation of this thesis, I used the following AI tools: Grammarly for grammar check and ChatGPT for short snippets of code. I declare that I used these tools in accordance with the principles of academic integrity. I checked the content and take full responsibility for it.

Filip Kučerák

Acknowledgments

I would like to thank my supervisor, Dan Kráľ, for all of the invaluable advice and opportunities I was given throughout my studies. I would also like to thank my partner and family for their support during my time at Brno. Finally, I would like to thank my colleagues from the faculty for our many (un)productive conversations.

Abstract

An object is said to be quasirandom if it resembles a truly random object from the same structure class. In this thesis, we focus on determining the boundary of sufficient conditions for the property in the combinatorial setting of permutations. It is known that a permutation is quasirandom if and only if the density of all four-element permutations does not significantly deviate from their expected density of 1/24. Subsets of permutations that can be similarly used to capture the quasirandomness are called quasirandom-forcing. Two natural question arise: determining the size of the smallest such set and classifying the inclusion-wise minimal forcing sets. The aim of the thesis is to contribute towards these questions.

It was known that collections of at most three permutations and collections of four permutations of size four are not quasirandom-forcing. There also exists a quasirandom-forcing set of size six; this is conjectured to be the best possible in terms of the size of a quasirandom-forcing set. We employ analytical methods to show that there is no quasirandom-forcing set of four permutations of size at most five.

Keywords

permutations, permutation limits, quasirandomness, quasirandom-forcing

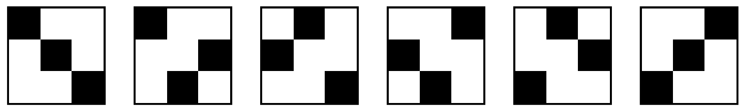
Contents

Notation and Conventions

We use $\mathbb{N}$ to denote non-negative integers including zero. We use $[n]$ to denote the set of first n positive integers and use $[n]_0$ to denote the first $n+1$ non-negative integers. We use $\subset$ to denote not-necessarily proper subsets, i.e., we would use $\subsetneq$ to denote proper subset had there been a chance. As such, we have $\mathbb{N} \subset \mathbb{N}$. Given a set A, we adopt the notation $\sum_a^A f(a)$ for sums and use this notation whenever possible, e.g., for sums, products, minimums, maximums, and probabilities. We use the notation $(x_i)_i^I$ for tuple (possibly infinite) of values indexed by set I and use $\{x_i\}_i^I$ for set whose elements are indexed by I. Given functions $f : A \to B$ and $g : B \to C$, we use gf to always mean composition $g \circ f : A \to C$. We use notation $(x \mapsto f(x))$ to denote function literals, i.e., symbol $(x \mapsto x^2)$ is the same to the set of functions $\mathbb{R} \to \mathbb{R}$ as the symbol 5 is to the set of natural numbers $\mathbb{N}$. Given a proposition V, we write $\mathbf{1}(V)$ for one if V is true and zero if V is false, i.e., $\mathbf{1}$ serves as a indicator function. Given a set A and natural number $k \in \mathbb{N}$, we use $A^{(k)}$ for set of all k-element subsets of A and use binomial coefficient $\binom{n}{k}$ to denote the cardinality of $[n]^{(k)}$; note that if $n < k$, we have $\binom{n}{k} = 0$ and $\binom{0}{0} = 1$. Given a map $f : A \to B$, we define the *kernel* of f to be the partition of domain A into classes which map to the same value under f.

Given vector spaces U and V, we use $\mathcal{L}(U, V)$ to denote the space of linear maps between U and V. Given $n, m \in \mathbb{N}$, we denote by $\mathbb{R}^{m \times n}$ the space of real matrices with m rows and n columns and use standard notation $(M)_{i,j}$ to denote the element of the matrix in i-th row and j-th column and so $(M)_{i,j} = e_i^T M e_j$, where we use e_i for the i-th standard basis vector. Given $n, m \in \mathbb{N}$, we denote by $\mathbb{R}^{n \to m}$ the space of linear maps $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ and identify this set with $\mathbb{R}^{m \times n}$ under the representation in the canonical basis. Whenever we discuss a matrix which is associated to a linear map, we use $\mathbb{R}^{- \to -}$ notation and use $\mathbb{R}^{- \times -}$ otherwise, e.g., when discussing representations of bilinear forms. Given a set A, we use $\mathbb{R}A$ to mean vector space of formal linear combinations of finitely many elements from A with real coefficients, i.e., the free real vector space on A. Given a subset $U \subset V$ of vectors, we use $\mathbb{R}U$ to denote the subspace spanned by the vectors from U and use $\mathbb{R}u$ as a shorthand for $\mathbb{R}\{u\}$.

Let us for set A denote by $\mathrm{Aut}(A)$ the set of functions of the form $A \to A$ that are bijections. Let $\mathbb{S}_n$ denote the group of permutations on n elements, i.e., the set $\mathrm{Aut}([n])$ with multiplication defined by function composition. Let



123 132 213 231 312 321

Figure 1: Elements of symmetric group $\mathbb{S}_3$ and their cover matrices.

$\mathbb{S} = \cup_n^{\mathbb{N}} \mathbb{S}_n$ be the set of all permutations without further structure and define $|\sigma|$ to be the natural number n such that σ is an element of $\mathbb{S}_n$. Throughout the text, we represent permutation $\sigma \in \mathbb{S}_k$ by a tuple $(\sigma(1), \sigma(2), ..., \sigma(k))$ and when k is small, we omit both the parentheses and commas, e.g., we write 132 for permutation $(1, 3, 2)$. To permutation $\sigma \in \mathbb{S}_k$, we associate linear bijection $A_\sigma \in \mathbb{R}^{k \to k}$ with $A_\sigma e_i = e_{\sigma(i)}$ whose matrix we call the *cover matrix* of σ. It is easy to see that mapping sending permutation to its cover matrix is a group homomorphism $\mathbb{S}_k \to \mathrm{SL}(k, \mathbb{R})$, where $\mathrm{SL}(k, \mathbb{R})$ denotes the special linear group of matrices of the form $\mathbb{R}^{k \times k}$ with determinant one. Note that there is another standard representation $B_\sigma e_i = e_{\sigma^{-1}(i)} = A_{\sigma^{-1}} e_i = A_\sigma^T e_i$ which we do not use here.

Chapter 1

Introduction

An object of some structure class is said to be *quasirandom* if it resembles a random object of the same class. The theory was classicaly built in 1980s in the setting of graphs by Rödl [22], Thomason [26] and Chung, Graham and Wilson [4] although the intuitive motive as stated is quite general and has indeed been studied in variety of combinatiorial settings, e.g., in the settings of digraphs, tournaments [3, 12, 1, 8], Latin squares [6], permutations [18, 2, 19, 9, 17], in algebraic settings, e.g., for groups [11], and even in the more general model theoretic settings [8].

Before investigating permutations which are the main subject of the thesis, let us briefly discuss quasirandomness in the setting of graphs. To this end, we choose Erdős-Réryi graph $\mathbb{G}(n,p)$ to be our random graph model. Graph $G \sim \mathbb{G}(n,p)$ is a graph on vertex set $[n]$, where we declare vertex pair $ij \in [n]^{(2)}$ to be an edge independently with probability p. We say that $G(n,p)$ has property $\mathcal{P}$ if the probability that a graph sampled according to $G(n,p)$ has the property $\mathcal{P}$ with high probability, formally

$$\lim_{n\to\infty} \mathbb{P}_G^{\mathbb{G}(n,p)}(G \in \mathcal{P}) = 1.$$

An easy argument using Chernoff bounds can show that $G(n,p)$ has $p+o(1)$ proportion of edges and one can—using slightly more nuanced concentraction results—generalize the argument to all subobjects. Formally, let us denote by $t(H,G)$ for graphs H and G the probability that if we sample random set injection $\psi : V(H) \to V(G)$, the map ψ is a graph homomorphism, i.e., we have that if $uv \in V(H)^{(2)}$ is an edge of H, then $\psi(u)\psi(v)$ is an edge in G. Using linearity, we can observe that given $n \geq |H|$, the expected value of the quantity $t(H,G)$ for $G \sim \mathbb{G}(n,p)$ is $t_{\mathbb{G}(n,p)} := p^{|E(H)|}$ and moreover, given graph H we have that for random graph sequence $G_\bullet := (G_n)_n^{\mathbb{N}}$ with $G_n \sim G(n,p)$ the density sequence $(t(H,G_n))_n^{\mathbb{N}}$ converges to this expectation almost surely. This motivates the following definition. We call graph sequence $G_\bullet = (G_n)_n^{\mathbb{N}}$ *convergent* if $\sup_n^{\mathbb{N}} |G_n| = \infty$ and the limit $t(H,G_\bullet) := \lim_{n\to\infty} t(H,G_n)$ exists, and call such sequence *quasirandom* if $t(H,G_\bullet) = t_{\mathbb{G}(n,p)}(H)$ for every graph H; notice that a random sequence is quasirandom almost surely. A similar definition can be used for any class of structures with good notion of a random model—playing the role of $\mathbb{G}(n, p)$—and good notion of an subobject or structure preserving mapping—playing the role of subgraphs and graph homomorphisms [1].

The definition as stated is *a priori* infinitary in the sense that for sequence $G_\bullet$, one needs to check behaviour of infititely many density sequences $(t(H, G_n))_n^{\mathbb{N}}$, one for each graph H. An important question in the broad area of quasirandomness is determining equivalent, necessary, and sufficient conditions for structure sequence to be quasirandom. Chung, Graham and Wilson, building on an earler work of Rödl [22] and Thomason [26], give a number of equivalent conditions of which we present only few.

Theorem 1 (Chung-Graham-Wilson [4])**.** Given a growing graph sequence $G_\bullet$, the following are equivalent

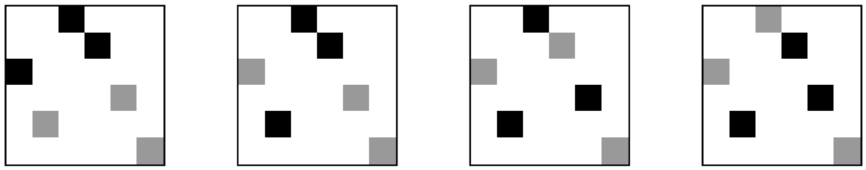
1. the sequence is p-quasirandom,
2. for $X \subset V(G_n)$ we have $e(X) = p|X|^{(2)} + o(n^2)$,
3. $t(K_2, G_n) = p$ and $t(C_4, G_\bullet) = p^4$,
4. if $\lambda_1 \geq ... \geq \lambda_n$ are the eigenvalues of the adjacency matrix of G, then $\lambda_1 = pn + o(n)$ and $\max_i^n |\lambda_i| = o(n)$.

We focus on condition 3 that is quite surprising when seen for the first time. It states that not only do we need to check only finitely many density sequences, if we know that the edge density in $G_\bullet$ converges to p, we only need to check a single graph, the 4-cycle. We call a set of graphs S *quasirandom-forcing* if for convergent sequence $G_\bullet$ we have that if there exists edge density $p \in [0, 1]$ and for every H in set S it holds that $t(H, G_\bullet) = t_{\mathbb{G}(n,p)}(H)$, then $G_\bullet$ is p-quasirandom. An important area of study in quasirandom graph theory is to determine which sets S are quasirandom-forcing. Chung, Graham and Wilson have shown that one can replace C_4 by any even cycle and any complete bipartite graph $K_{2,t}$ for $t \geq 2$. A result of Skokan and Thoma [25] showed that one can genarize this to $K_{a,b}$, i.e., that $\{K_2, K_{a,b}\}$ is forcing and the Forcing Conjecture of Conlon, Fox and Sudakov [5] states that the set $\{K_2, H\}$ is quasirandom forcing if and only if H is bipartite and contains a cycle.

Similar results can be found in the study of other structures. For tournaments (orientations of complete graphs) it is known that singleton set $\{H\}$ is quasirandom-forcing if and only if it is a transitive tournament on at least four vertices or an exceptional tournament on five vertices [8, 1, 8, 13]; for Latin squares it is known that patterns of size 2×3 are quasirandom forcing [6].

In this thesis, we focus on the study of forcing quasirandomness in the setting of permutations. We start by formally defining the required notions and preliminary results, starting with defining the notion of a subobject.

[1] A reader may ask if the choice of subgraphs as our subobjects instead of induced subgraphs and counting labeled occurrences instead of unlabeled is arbitrary, i.e., it has no influence on which convergent sequences are quasirandom, or the choice is deliberate. The theory tells us that the former is the case.



$P = \{1, 3, 4\}$ $P = \{2, 3, 4\}$ $P = \{2, 3, 5\}$ $P = \{2, 4, 5\}$

Figure 1.1: All four sets $P \subset [6]$ for which $(351246)|_P = 312$.

Given a permutation $\pi \in \mathbb{S}_n$ and a k-element subset $P \subset [n]$, we define $\pi|_P \in \mathbb{S}_k$ to be the unique permutation such that for $i, j \in [k]$ with $i < j$, we have that $\pi(P_i) < \pi(P_j)$ if and only if $\pi|_P(i) < \pi|_P(j)$, where $P_1 < ... < P_k$ are ordered elements of P. As an example, given permutation $\pi = (3, 1, 4, 2, 5) \in \mathbb{S}_5$, and set $P = \{1, 3, 4\}$ we have $\pi|_P = (2, 3, 1)$. Given permutations $\sigma \in \mathbb{S}_k$ and $\pi \in \mathbb{S}_n$ we use $d(\sigma, \pi)$ for *density* of σ in π, which we define to be the probability that for uniformly chosen $P \in [n]^{(k)}$ it holds that $\pi|_P = \sigma$. We denote by $d_\rho(\sigma)$ the expected density of σ in permutation uniformly chosen from $\mathbb{S}_n$ for $n \geq |\sigma|$; note that this is well-defined as the expectation is the same for all $n \geq |\sigma|$, namely it is $1/|\sigma|!$. Given a permutation sequence $\pi_\bullet = (\pi_n)_n^{\mathbb{N}}$, we say that the sequence is *convergent* if $\sup_n^{\mathbb{N}} |\pi_n| = \infty$ and $d(\sigma, \pi_\bullet) = \lim_{n\to\infty} d(\sigma, \pi_n)$ converges for all $\sigma \in \mathbb{S}$ and say that convergent sequence is *quasirandom* if $d(\sigma, \pi_\bullet) = d_\rho(\sigma)$ for all $\sigma \in \mathbb{S}$. We denote the space of convergent permutation sequences by $\mathcal{S}$. Given a map $F : \mathcal{S} \to \mathbb{R}^S$ for some set S, we say that the map is *quasirandom-forcing* (from now on forcing) if for convergent permutation sequence $\pi_\bullet \in \mathcal{S}$, we have that if $F(\pi_\bullet) = 0$ holds then π is quasirandom, i.e., the zero set of F contains only quasirandom sequences.

Given a set of permutations $S \subset \mathbb{S}$ a natural map of the sort is the map $G_S : \mathcal{S} \to \mathbb{R}^S$ which sends $\pi_\bullet$ to $(d(\sigma, \pi_\bullet) - d_\rho(\sigma))_{\sigma\in S}$ which we call the *density discrepancy* function. We say that the set of permutations is *quasirandom-forcing* (from now on also forcing) if the associated map G_S is. A result of Kráľ and Pikhurko in [18] states that $\mathbb{S}_4$ is forcing, while the sets $\mathbb{S}_l$ are known not to be forcing for $l < 4$ [7, 18]; this means that—similarly as for graphs, tournaments and Latin squares —quasirandomness in permutations is also captured by only finitely many pattern densities. A natural goal is to classify inclusion-wise minimal forcing sets with first step being understanding the forcing set with minimal size which we denote ω. For the latter, only bounds are known. For upper bounds, the aforementioned result of Kráľ and Pikhurko in [18] shows that $\omega \leq 24$ which was later improved first by Chan et al. in [2] to $\omega \leq 8$ and later by Crudele et al. in [9] to $\omega \leq 6$. To discuss these works we first need some definitions.

We define the linear space of *superpermutations*[2] $\mathbb{R}\mathbb{S}$ to be the free vector

[2]Here, we take inspiration from terminology of quantum graphs taken from [20] by adapting

space generated by the set $\mathbb{S}$, i.e., elements of $\mathbb{RS}$ are formal linear combinations of finitely many permutations. We extend definitions of $d(-,\pi)$ and $d_\rho(-)$ linearly to $\mathbb{RS}$ so that for $\nu = \sum_\sigma^S a_\sigma \sigma$ we have $d(\nu,\pi) = \sum_\sigma^S a_\sigma d(\sigma,\pi)$ and $d_\rho(\nu) = \sum_\sigma^S a_\sigma d_\rho(\sigma)$. In [2] and [9], authors study forcing of a map $G_\nu : \mathcal{S} \to \mathbb{R}$ associated with $\nu \in \mathbb{RS}$ which sends $\pi_\bullet$ to $d(\nu,\pi_\bullet) - d_\rho(\nu)$, that is, they consider discrepancy of the mixed densities. Let us define supp $\nu \subset \mathbb{S}$ to be the set of those permutations σ which have non-zero coefficients a_σ in ν and let us say that ν is of *order k* if the largest permutation in ν is of order k. Notice that if G_ν is forcing and $S = \text{supp } \nu$, then so is G_S: let $a \in \mathbb{R}^S$ be the real vector so that $\nu = \sum_\sigma^S a_\sigma \sigma$, then $G_\nu = a^T G_S$ meaning if $G_S(\pi_\bullet) = 0$ then $G_\nu(\pi_\bullet) = 0$ and $\pi_\bullet$ is quasirandom by forcing of G_ν. The main result of [2] is full description of quasirandom forcing superpermutations $\nu \in \mathbb{RS}_4$, producing ν with $|\text{supp } \nu| = 8$ in the process, thus showing $\omega \le 8$. In [9], authors determine superpermutation $\nu \in \mathbb{RS}_4 \oplus \mathbb{RS}_3$ with support of size 6 which is forcing, giving us the currently best-known upper bound $\omega \le 6$ that is believed to be tight. Moreover, [9] shows that if the coefficients ν_σ of ν are positive and ν has support of size at most 5, then G_ν is not forcing. We discuss the result and tools required in Chapter 3.

For lower bound, the best known bound is $\omega > 3$ by Kurečka shown in [19]. We discuss the result and tools used in Chapter 3. The main aim of this thesis is to contribute to establishing bound $\omega > 4$. In this direction, result of [17] shows that given a 4-element set of permutations $S \subset \mathbb{S}_4$, the set S cannot be forcing; we review the tools used in Chapter 4. Our main result is an extension of this result to sets of four small permutations.

Theorem 2. Let S be a set of four permutations of order at most five, then S is not forcing.

We briefly review the structure of the thesis. In Chapter 2, we survey results regarding permutation limits which give us a workable description of function G_S and serve as a main source of convergent permutation sequences in this thesis. In Chapter 3, we survey gradient-based tools from [19] and their consequences found in [9]. In Chapter 4, we discuss Hessian-based methods found in [17] and describe a non-forcing certificate which we obtain computationally. Finally, in Chapter 5, we use all the aforementioned methods and results to prove the main claim of this thesis. The main body of the text is followed by a series of appendices, the most important of which is Appendix D that describes our computationally found certificates that play a crucial role in various proofs.

the term superposition. There is a different notion of a superpermutation, namely a string of n symbols which contains every permutation on n elements as a substring. Although the clash in names is somewhat unfortunate, I believe it should cause no confusion.

Chapter 2

Permutation Limits

In recent years, study of large combinatorial structures and study of growing combinatorial sequences—e.g., in the study of random graphs, quasirandom graphs and extremal combinatorics—gave rise to a theory of combinatorial limits. The area asks to, for a sequence of combinatorial objects, to create a single object capturing the properties of the elements in the sequence. The exact properties one has in mind heavily depend on the intended use. One viable approach is a very general "syntactic" approach which views combinatorial objects as models of first-order theories and leverages tools coming from the area of model theory such as ultralimit constructions to create "logical" limits—limit objects which satisfy the same first-order formulae as a "large" portion of elements of the sequence—and even "density" limits—limit objects which capture behaviour of the subobject density functionals—using the flag algebra approach originally defined by Razborov [21] which has been an important tool in the area of extremal combinatorics.

The aim of this section is to review one such density limit of convergent permutations sequences (for definition of convergent sequence see introduction), i.e., for convergent sequence $\pi_\bullet$ we aim to define an object μ together with a functional $d(-,\mu) : \mathbb{RS} \to \mathbb{R}$ with the property that $d(\nu,\mu) = d(\nu,\pi_\bullet)$ for every superpermutation $\nu \in \mathbb{RS}$. A simple yet illustrative way is to define the sequence itself to be the limit. This approach represents the most skeletal approach to limit building. The approach of flag-algebras—slightly more concrete—is to proclaim an appropriately chosen subset of functionals $\varphi : \mathbb{RS} \to \mathbb{R}$ to be the limit. Even though flag algebras are both crucial, fascinating and main tool used in the proof $\omega \leq 6$, its survey is outside the scope of this thesis.

In this thesis, we instead focus on a concrete analytical limit, namely measure theoretic object studied in [16, 15] [1]. We define *permuton* to be a Borel measure μ on $[0,1]^2$ with uniform marginals, i.e., we require $\mu([a,b] \times [0,1]) = \mu([0,1] \times [a,b]) = b - a$ for all $0 \leq a \leq b \leq 1$. To define subpermutation density $d(-,\mu)$:

[1] The analytical object studied in the cited works slightly differs from the object presented but Lemma 2.2 and Definition 2.3 in the former citation can be used to translate one to the other.

$\mathbb{RS} \to \mathbb{R}$, we first need to establish some definitions. Given k distinct members $x = (x_i)_i^{[k]}$ of some linear order L (reader should think $[n]$ for some $n \in \mathbb{N}$), we say that collection x induces permutation π if $x_{\pi^{-1}(1)} < x_{\pi^{-1}(2)} < ... < x_{\pi^{-1}(k)}$ (note that for any collection x there is a unique permutation π with this property). As an example, vector $(0.5, 3.14, 10, 2) \in \mathbb{R}^4$ induces permutation $\pi = (1, 3, 4, 2) \in \mathbb{S}_3$ as $\pi^{-1} = (1, 4, 2, 3)$ and $0.5 < 2 < 3.14 < 10$. Given two collections x^1 and x^2 of k-distinct elements in some linear order L, we say that they induce permutation π if x^1 induces π_1, x^2 induces π_2, and $\pi = \pi_2\pi_1^{-1}$. As an example, take sequence $((1, 5), (3, 2), (4, 4), (2, 10)) \in \mathbb{Z}^{2\times 4}$ which induces $(3, 4, 1, 2) \in \mathbb{S}_3$; projection of the sequence to the first coordinate induces permutation $(1, 3, 4, 2) \in \mathbb{S}_3$, projection of the sequence to the second coordinate induces permutation $(3, 1, 2, 4) \in \mathbb{S}_3$ and $(3, 1, 2, 4)(1, 3, 4, 2)^{-1} = (3, 4, 1, 2)$. We define *density* of π in permuton μ to be the probability $d(\pi, \mu)$ that k independently μ-sampled points $(x_i, y_i)_i^{[k]}$ induce permutation π. Although the notion of a permutation induced by a set of points is only meaningful as long as all coordinates are distinct, due to the uniform marginals condition, the probability that we sample distinct coordinates is one.

An important class of permutons are *step permutons* $\mu[A]$ associated to doubly-stochastic matrix $A \in \mathbb{R}^{n\times n}$ which we define for any Borel $X \subset [0, 1]^2$ by

$$\mu[A](X) = n\sum_{i,j}^{[n]^2} A_{i,j}\lambda(X \cap (I_n^i \times I_n^j)) = \int_{x,y}^{X} n\sum_{i,j}^{[n]^2} A_{i,j}1_{I_n^i\times I_n^j}(x, y)d\lambda,$$

where λ is the Lebesgue measure and $I_n^i = [(i-1)/n, i/n] \subset [0, 1]$. Intuitively, permuton μ corresponds to graphical representation of the matrix A in the unit square, each cell of the matrix viewed as a constant square with side of length $1/n$.

The following lemma gives us a simple formula to compute permutation densities in step permutons. Given two posets A and B, let $\mathbf{Ord}(A, B)$ be the set of order preserving mappings from A to B, i.e., mappings f such that $a \leq_A a' \implies f(a) \leq_B f(a')$ for all $a, a' \in A$.

Lemma 3 ([2] Lemma 10)**.** Given permutation $\sigma \in \mathbb{S}_k$ and permuton $\mu[A]$ associated with $A \in \mathbb{R}^{n\times n}$, it holds that

$$d(\sigma, \mu[A]) = \frac{k!}{n^k}\sum_{f,g:\mathbf{Ord}([k],[n])} \frac{\prod_i^{[k]} A_{f(i),g(\sigma(i))}}{\prod_i^{[n]} |f^{-1}(i)||g^{-1}(i)|}.$$

The main result of [16] and [15] shows that permutons are a "good" definition of a limit in the sense that for every convergent sequence of permutations $\pi_\bullet$ there is a unique permuton μ associated to $\pi_\bullet$ such that $d(-, \pi_\bullet) = d(-, \mu)$ (completeness) and moreover that for every permuton μ there is a convergent sequence $\pi_\bullet$ such that the permuton associated to $\pi_\bullet$ is μ almost everywhere (analogue to permutations being "dense" in the space of permutons[2]).

Let us now relate the study of quasirandom sequences with the area of permutation limits. First, by the completeness, we have that each quasirandom sequence corresponds to a single permuton, namely the uniform permuton ρ defined to be the Lebesgue measure on the unit square; ρ as defined is a permuton and moreover, the probability that k points $(x_i, y_i)_i^{[k]}$ induce any permutation $\sigma \in \mathbb{S}_k$ is exactly $d_\rho(\sigma) = 1/k!$. Given a set of permutations $S \subset \mathbb{S}$, define $\tilde{G}_S : \mathcal{P} \to \mathbb{R}^S$ to be a mapping which to each permuton $\mu \in \mathcal{P}$ and $\sigma \in \mathbb{S}$ assigns the density $d(\sigma, \mu) - d(\sigma, \rho)$. As the permuton associated to any quasirandom sequence $\pi_\bullet$ is ρ and for every permuton there is a convergent sequence to which it is associated, to ask whether set S is forcing is to ask, whether there is non-uniform μ such that $\tilde{G}_S(\mu) = 0$; the remaining of the presented work aims to produce for sets S such permuton μ. To do this, we show that the discussed sets S cannot force ρ even on subset $\mathcal{P}' \subset \mathcal{P}$, i.e., even restrictions of $\tilde{G}_S|_{\mathcal{P}'}$ contain non-uniform zeros; in such cases we say that S is not $\mathcal{P}'$-forcing.

As G_S is forcing if and only if $\tilde{G}_S$ is forcing, and there is a clear correspondence between sets $\mathcal{S}$ and $\mathcal{P}$, we abuse notation and denote both functions by G_S.

The theory of permutation limits plays an important role in the proof of the fact that the set $\mathbb{S}_4$ is forcing.

Theorem 4 ([18] Theorem 3)**.** The set $\mathbb{S}_4$ is forcing.

A result of [10] generalizes the above to show that permutons which are structurally simple can be also forced by a finite set of permutations, i.e., if $\mu = \sum_i^{[k]} \alpha_i A_i$ for some non-negative reals α_i and non-trivial polygons A_i, then there exists a set S such that the zero set of the map $H_S(\mu') = (d(\pi, \mu') - d(\pi, \mu))_\pi^S : \mathcal{P} \to \mathbb{R}$ contains only μ.

We conclude this section by reviewing a standard parametrization of a neighbourhood of ρ with hope to finding the desired non-uniform permutons μ in said neighbourhood. Let n be a natural number, then for $i, j \in [n-1]$ we define $B_{i,j} = (e_i - e_{i+1})(e_j - e_{j+1})^T \in \mathbb{R}^{n \times n}$ to be the *n-perturbation* at (i, j). Given formal variables $x = (x_{i,j})_{i,j}^{[n-1]^2}$, we define matrix $\mathbb{J}(x) : \mathbb{R}^{n \times n}$ by equation

$$\mathbb{J}(x) = \mathbb{J} + \sum_{i,j}^{[n-1]^2} x_{i,j} B_{i,j}.$$

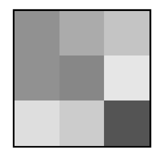
For given $n \in \mathbb{N}$, we define the set of perturbed permutons $\mathcal{P}_n$ to be the set of permutons which are $\zeta_n(x) := \mu\mathbb{J}(x)$ for some $x \in \mathbb{R}^{[n-1]^2}$. Our aim will be to show that any set $S \subset \mathbb{S}_{\leq 5}$ of order four cannot be $\mathcal{P}_n$-forcing for some $n \in \mathbb{N}$ and thus cannot be forcing. Let $\zeta_n(x) \in \mathcal{P}_n$ be a perturbed permuton, then using Lemma 3 we can observe that for any $\sigma \in \mathbb{S}_k$, the density $d(\sigma, \zeta_n(x))$ is

[2] A reader may ask whether the space of permutations and permutons admits a metric under which convergent sequences correspond to Cauchy sequences. The answer is yes but the metric $d_\square$ called cut distance is not trivial to set up and we do not need it here.

$$\begin{pmatrix} x_{1,1}+\frac{1}{3} & -x_{1,1}+x_{1,2}+\frac{1}{3} & -x_{1,2}+\frac{1}{3} \\ -x_{1,1}+x_{2,1}+\frac{1}{3} & x_{1,1}-x_{1,2}-x_{2,1}+x_{2,2}+\frac{1}{3} & x_{1,2}-x_{2,2}+\frac{1}{3} \\ -x_{2,1}+\frac{1}{3} & x_{2,1}-x_{2,2}+\frac{1}{3} & x_{2,2}+\frac{1}{3} \end{pmatrix}$$

$\mathbb{J}(x)$

$$\begin{pmatrix} 0 & 0 & 0 \\ 1 & -1 & 0 \\ -1 & 1 & 0 \end{pmatrix} \qquad \begin{pmatrix} \frac{13}{30} & \frac{1}{3} & \frac{7}{30} \\ \frac{13}{30} & \frac{7}{15} & \frac{1}{10} \\ \frac{2}{15} & \frac{1}{5} & \frac{2}{3} \end{pmatrix}$$



$B_{2,1}$ $\qquad$ $\mathbb{J}(\frac{1}{10}, \frac{1}{10}, \frac{2}{10}, \frac{1}{3})$ $\qquad$ $\zeta_3(\frac{1}{10}, \frac{1}{10}, \frac{2}{10}, \frac{1}{3})$

Figure 2.1: Construction of perturbed permuton $\zeta_3(x) \in \mathcal{P}_3$ for $x \in \mathbb{R}^{2^2}$ defined by $x_{1,1} = 1/10$, $x_{1,2} = 1/10$, $x_{2,1} = 2/10$, and $x_{2,2} = 1/3$.

polynomial in $\mathbb{R}[x]$, we call this polynomial $f_{n,\sigma} \in \mathbb{R}[x]$ the *density polynomial of* σ. Further, we define polynomial $g_{n,\sigma} = f_{n,\sigma} - d(\sigma, \rho)$ which we call the *discrepancy polynomial.* The above states that $G_S|_{\mathcal{P}_n} \circ \zeta_n : \mathbb{R}^{[n-1]^2} \to \mathbb{R}^S$ is a polynomial in each coordinate; our aim is to find a small nontrivial common root of these polynomials.

Chapter 3

First-order Method

The first method—method which will allow us to exclude majority of the studied sets—will be based on the properties of the Jacobian of $G_{n,S} := G_S|_{\mathcal{P}_n} \circ \zeta_n$ and use of Implicit Function Theorem. Informally, our aim is to show that majority of permutation 4-sets are "independent", which gives us enough degrees of freedom with respect to perturbing the random permuton that we can indeed construct permuton which contradicts $\mathcal{P}_n$-forcing. This method is the main tool used in [19] and serves as the uniform reason why 3-sets with sufficiently large permutations are not forcing.

We start by reviewing elements of real analysis. We denote by C^n the set of function with continuous partial derivatives of order at most n and denote by C^∞ the set of *smooth* functions, i.e., functions with continuous partial derivatives of all orders. For function $F : \mathbb{R}^n \to \mathbb{R}^m$ differentiable at point p, we denote by $D_{x,p}F : \mathbb{R}^{n \to m}$ the Jacobian of F with respect to variables x evaluated at point p and use $D_p F : \mathbb{R}^{n \to m}$ to denote the Jacobian of F at point p. Let us recall the Implicit Function Theorem which gives a sufficient condition which allows us to locally express the level set of a smooth function $F : \mathbb{R}^{r+c} \to \mathbb{R}^c$ around a point $(x_0, y_0) \in \mathbb{R}^{r+c}$ as a graph of a smooth function $\phi : \mathbb{R}^r \to \mathbb{R}^c$. The statement, its proof and further discussion can be found in any standard text on multivariable calculus, e.g., [24].

Theorem 5 (Implicit Function Theorem)**.** *Let $r, c \in \mathbb{N}$ and let $F : \mathbb{R}^{r+c} \to \mathbb{R}^c$ be a smooth function with $F(x_0, y_0) = 0$, where $x_0 \in \mathbb{R}^r$ and $y_0 \in \mathbb{R}^c$. Suppose the matrix $D_{y,(x_0,y_0)}F$ is invertible. Then there are open sets $x_0 \in U \subset \mathbb{R}^r$ and $y_0 \in V \subset \mathbb{R}^c$ and a smooth function $\phi : U \to V$ so that for $x \in U$ and $y \in V$, we have that $F(x, y) = 0$ if and only if $\phi(x) = y$.*

Given a permutation set S, we say that the set is n-independent, if $D_0 G_{n,S}$ does not have deficient rank. The following lemma tells us that such sets cannot be forcing. We repeat the proof of [19] for completeness as Chapter 4 builds upon the argument which follows.

Lemma 6 ([19] Lemma 5)**.** *Let S be a finite set of permutations and n a positive integer with $(n-1)^2 > |S|$. If S is n-independent, then S is not $\mathcal{P}_n$-forcing.*

Proof. Let S be a finite set of permutations with $D_0G_{n,S} : \mathbb{R}^{(n-1)^2 \to |S|}$ of rank $|S|$, i.e., $D_0G_{n,S}$ does not have deficient rank and is a linear surjection. As linear surjections split, let $L : \mathbb{R}^{|S| \to (n-1)^2}$ be a map so that $D_0G_{n,S}L$ is identity and let $v \in \text{Im } L^{\perp}$ be a non-zero vector which exists by $(n-1)^2 > |S|$. Define $F : \mathbb{R}^{1+|S|} \to \mathbb{R}^{|S|}$ by sending $(x, y) \in \mathbb{R}^{1+|S|}$ to $G_{n,S}(xv + Ly)$. As $D_{y,0}F = D_0G_{n,S}L$ is invertible, it follows from the Implicit Function Theorem that there exist open sets $U \subset \mathbb{R}^1$ and $V \subset \mathbb{R}^{|S|}$ and a smooth function $\phi : U \to V$ such that for $x \in U$ we have $F(x, \phi(x)) = G_{n,S}(xv + L\phi(x)) = 0$. By choosing $x \neq 0$ and noticing that $L\phi(x) + xv \neq 0$ as v is a non-zero vector orthogonal to image of L, $\zeta_n(L\phi(x) + xv)$ is a non-uniform permuton with correct densities for permutations $\sigma \in S$, proving S is not $\mathcal{P}_n$-forcing. $\square$

3.1 Gradient Polynomial

In this section, our aim is to find a sufficient condition for set S to be n-dependent for *some* sufficiently large natural number $n \in \mathbb{N}$ without determining the rank of individual Jacobians $D_0G_{n,S}$. To this end, the gradient $D_0g_{n,\sigma} \in \mathbb{R}^{(n-1)^2 \to 1}$ can be viewed as a matrix $M_{n,\sigma} \in \mathbb{R}^{(n-1)\times(n-1)}$ defined by $(M_{n,\sigma})_{i,j} = (D_0g_{n,\sigma})_{1,(i,j)}$. Let $s_{n,\sigma} : (0,1)^2 \to \mathbb{R}$ be a step function associated to $M_{n,\sigma}$ which maps $(\alpha, \beta) \in (0,1)^2$ to $(M_{n,\sigma})_{\lceil \alpha n \rceil, \lceil \beta n \rceil}$. The pointwise limit $s_\sigma(\alpha, \beta)$ of the sequence $(s_{n,\sigma}(\alpha, \beta))_n^{\mathbb{N}}$ now represents a infinitesimal perturbation in ρ at (α, β) with radius going to zero and thus the limit s_σ is identically zero.

To avoid vanishing, author in [19] normalizes the sequence and introduces for permutation $\sigma \in \mathbb{S}_k$ the notion of a *gradient polynomial* defined by

$$P_\sigma(\alpha, \beta) = \lim_{n\to\infty} n^3 s_{n,\sigma}(\alpha, \beta).$$

Result of Kurečka is that not only is this function well defined, it is in fact non-trivial and a polynomial in α and β. The aim of this section is to describe the polynomial and survey the consequences of its existence.

Given $k \in \mathbb{N}$, let $B_k : \mathbb{R}^{(k-1) \to k}$ and $D_k : \mathbb{R}^{(k-1) \to (k-1)}$ be matrices defined by

$$(B_k)_{i,l} = (-1)^i \binom{l}{i-1} \text{ and } (D_k)_{i,l} = \mathbf{1}(i = l)(-1)^{l+1} \binom{k-2}{l-1},$$

where recall that we define $\binom{n}{m}$ to be zero whenever $n < m$. Note that D_k is an invertible diagonal matrix and that B_k is an injective linear map with cokernel equal to the span of the all-ones vector $j \in \mathbb{R}^k$; namely this implies that $B_k^T j = 0$. Let us define $\langle \alpha^i \beta^j \mid f \rangle$ to be the coefficient of the monomial $\alpha^i \beta^j$ in the polynomial $f \in \mathbb{R}[\alpha, \beta]$. We define the k-th coefficient matrix of the polynomial $f \in \mathbb{R}[\alpha, \beta]$ to be the matrix $T^{(k)} f : \mathbb{R}^{(k+1) \to (k+1)}$ with $(T^{(k)} f)_{i,j} = \langle \alpha^{(i-1)} \beta^{(j-1)} \mid f \rangle$ for $i, j \in [k+1]$. We are now ready to describe the polynomial P_σ.

Theorem 7 ([19])**.** Let $k \geq 2$ be a natural number and $\sigma \in \mathbb{S}_k$ with cover matrix $A_\sigma \in \mathbb{R}^{k \to k}$, then P_σ is a real polynomial in α, β with $\langle \alpha^i \beta^j \mid f \rangle = 0$ if $i > k-2$ or $j > k-2$. Moreover,

$$T_\sigma := T^{(k-2)} P_\sigma = c_k D_k^T B_k^T A_\sigma B_k D_k$$

where $c_k = k(k-1)/(k-2)!$ is a positive real number. We call T_σ the *gradient matrix* of permutation σ.

Notice that due to the fact that $\langle \alpha^i \beta^j \mid P_\sigma \rangle = 0$ holds for i or j greater than $k-2$, it holds that T_σ fully describes the polynomial P_σ. Given a superpermutation $\nu = \sum_\sigma^S \nu_\sigma \sigma$ of order k, define $P_\nu = \sum_\sigma^S \nu_\sigma P_\sigma$ and define $T_\nu := T^{(k-2)} P_\nu = \sum_\sigma^S \nu_\sigma T^{(k-2)} P_\sigma$; note that matrices $T^{(k-2)} P_\sigma : \mathbb{R}^{(k-1) \to (k-1)}$ are not necessarily matrices $T_\sigma : \mathbb{R}^{(|\sigma|-1) \to (|\sigma|-1)}$ but the former can be a "zero-padded" version of the latter. Let us at this point extend the definition of a cover matrix to the spaces $\mathbb{RS}_k$, by defining it for $\nu = \sum_\sigma^S \nu_\sigma \sigma$ with finite set $S \subset \mathbb{S}_k$ to be the matrix $A_\nu = \sum_\sigma^S \nu_\sigma A_\sigma$. Note that due to distributivity of matrix multiplication, it holds that $T_\nu^{(k-2)} = c_k D_k^T B_k A_\nu B_k D_k$ for $\nu \in \mathbb{RS}_k$.

Given a set S, we say that it is independent if the polynomials $(P_\sigma)_\sigma^S$ are independent as elements of the vector space $\mathbb{R}[\alpha, \beta]$. If they are dependent, there exists $\nu \in \mathbb{RS}$ whose support is subset of S with $P_\nu = 0$; we call such superpermutation the *dependence witness*. The following lemma relates the two notions of dependence presented so far and its proof can be found in [19] in the proof of Lemma 8 which we state as Proposition 9. As the argument is short and the lemma is not directly present in the source as stated, we include it for completeness.

Lemma 8. Let S be a finite set of permutations and let $N \in \mathbb{N}$. If S is independent, then it is n-independent for some $n \geq N$.

Proof. Let $|S| = m$ and assume that there exists $N \in \mathbb{N}$ such that S independent but is n-dependent for all $n \geq N$. Let $(t_n)_n^{\mathbb{N}_{\geq N}}$ be a sequence of normed vectors $t_n \in [-1,1]^S$ with $t_n^T G_{n,S} = 0$ for every $n \geq N$. As $[-1,1]^S$ is sequentially compact and norm is a continuous function, there is an injective $h \in \mathbf{Ord}(\mathbb{N}, \mathbb{N})$ such that $(t_{h(n)})_n^{\mathbb{N}_{\geq N}}$ converges to some $t \in [-1,1]^S$ of norm one. As $P_\sigma = \lim_{n\to\infty} n^3 s_{n,S} = \lim_{n\to\infty} h(n)^3 s_{h(n),S}$ and product of limits is limit of products given the limits converge, we have

$$\sum_\sigma^S t^\sigma P_\sigma = \sum_\sigma^S (\lim_{n\to\infty} t_{h(n)}^\sigma)(\lim_{n\to\infty} h(n)^3 s_{h(n),\sigma}) = \lim_{n\to\infty} h(n)^3 \sum_\sigma^S t_{h(n)}^\sigma s_{h(n),\sigma} = 0$$

as $t_n^T G_{n,S} = 0$ implies $\sum_\sigma^S t_n^\sigma s_{n,\sigma} = 0$. The just presented equality contradicts the assumption that S is independent. □

A direct consequence of Lemma 8 and Lemma 6 is the following.

Proposition 9 ([19] Lemma 8)**.** Let S be a finite set of permutations. If S is independent, then S is not forcing.

This theorem gives us a simple algorithm by which we can computationally eliminate majority of the cases: iterate sets S systematically—we do so in Chapter 5—construct $T_S := (T_\sigma^{(k-2)})_\sigma^S$ where k is the order of the largest permutation in the set and use Gaussian Elimination to determine if matrices in T_S are independent as elements of linear space $\mathbb{R}^{(k-1)\times(k-1)}$. If yes, the set S cannot be forcing. If no, we add set S to the set of possibly forcing sets for which we need stronger tools.

Before discussing higher-order tools, we briefly review important consequences of the theorem above. Note that the theorem claims that any forcing set S has a dependence witness ν. We can use this information with the very explicit description of polynomial P_ν through T_ν to obtain the following corollary used in [19, 9, 17].

Corollary 10. Let $\nu = \sum_\sigma^S \nu_\sigma \sigma$ for $S \subset \mathbb{S}_k$. If $P_\nu = 0$, then $A_\nu = c\mathbb{J}$ for some $c \in \mathbb{R}$. Namely, if $S \subset \mathbb{S}_k$ is forcing with dependence witness $\nu \in \mathbb{R}\mathbb{S}$, then $A_\nu = c\mathbb{J}$.

Proof. Let us denote the matrix A_ν by A. As all permutations in S are of order k, it holds that $T_\nu = c_k D_k^T B_k^T A B_k D_k$ by distributivity. Now as $T_\nu = 0$, we have $B_k^T A B_k = 0$ as $c_k \neq 0$ and both D_k and D_k^T are full-rank diagonal matrices. Let us split $\mathbb{R}^k = \mathbb{R}j \oplus j^\perp$ and note that both subspaces are invariant with respect to A_ν: as A_ν is linear combination of doubly-stochastic matrices, we have that $A^T j = Aj = \lambda j$ for some $\lambda \in \mathbb{R}$ which for $u \in j^\perp$ implies $\langle j, Au\rangle = \lambda\langle j, u\rangle = 0$. We thus split A to $A_j \oplus A_{j^\perp}$, where $A_j : \mathbb{R}^k \to \mathbb{R}j$ has the property that $\ker A_j = j^\perp$ and $A_{j^\perp} : \mathbb{R}^k \to j^\perp$ with property that $\ker A_{j^\perp} = \mathbb{R}j$. Then note that $0 = B_k^T(A_j \oplus A_{j^\perp})B_k = B_k^T A_j B_k + B_k^T A_{j^\perp} B_k = B_k^T A_{j^\perp} B_k = 0$ as $A_j B_k = 0$. The last equality and the fact that B_k is an isomorphism of $\mathbb{R}^k$ onto $j^\perp$ implies $A_{j^\perp} = 0$. As $A_j^T j = A_j j = \lambda j$ and A_j has rank at most one, we have $A = A_j \oplus A_{j^\perp} = A_j = \frac{\lambda}{n} j j^T$ as desired. $\square$

3.2 Subsampling and Fuzzy Covers

We now turn to the notion of a "fuzzy" permutation covers used in [9]. For this, we need to review an important property of the space of superpermutation, namely the *subsampling identity* sometimes called the *chain rule*—as we reserve the latter name for the standard analytical chain rule, we stick to the former. Given permutations $\sigma \in \mathbb{S}_k$, $\pi \in \mathbb{S}_n$ and an integer l such that $k \leq l \leq n$, we have

$$d(\sigma, \pi) = \sum_\tau^{\mathbb{S}_l} d(\sigma, \tau) d(\tau, \pi) = d\left(\sum_\tau^{\mathbb{S}_l} d(\sigma, \tau)\tau, \pi\right). \tag{3.1}$$

To see (3.1), consider the following experiment. Let us uniformly at random sample $U \in [n]^{(l)}$ after which we uniformly at random sample $P \in U^{(k)}$. It is clear that this induces uniform distribution on $P \in [n]^{(k)}$ and as such, we have $d(\sigma, \pi) = \mathbb{P}(\pi|_P = \sigma)$. By law of total probability, we can express this probability as

$$\sum_{\tau\colon \mathbb{P}(\pi|_U=\tau)>0}^{\mathbb{S}_l} \mathbb{P}(\pi|_P = \sigma \mid \pi|_U = \tau)\mathbb{P}(\pi|_U = \tau)$$

and as P is uniformly sampled from $U^{(k)}$, we see that $\mathbb{P}(\pi|_P = \sigma \mid \pi|_U = \tau) = d(\sigma, \tau)$ and $\mathbb{P}(\pi|_U = \tau) = d(\tau, \pi)$ for τ with $\mathbb{P}(\pi|_U = \tau) \neq 0$. The claim follows as for τ with $\mathbb{P}(\pi|_U = \tau) = 0$ we have $d(\tau, \pi) = 0$.

We define $L^{(l)}\sigma = \sum_{\tau}^{\mathbb{S}_l} d(\sigma, \tau)\tau \in \mathbb{R}\mathbb{S}_l$ to be the l-th lift of σ and define the *l-th lift cover* $A_\sigma^{(l)}$ to be the cover matrix of the superpermutation $L^{(l)}\sigma$. The subsampling equality motivates the definition of the quotient space $\mathcal{A}$, given by factoring out the subspace generated by the set $\{\sigma - L^{(l)}\sigma \mid \sigma \in \mathbb{S}_k, l \geq k\}$ from space $\mathbb{R}\mathbb{S}$. Due to (3.1), we see that $d(-, \mu) : \mathbb{R}\mathbb{S} \to \mathbb{R}$ for every permuton μ factors through $d(-, \mu) : \mathcal{A} \to \mathbb{R}$. From this, we conclude the following property of gradient polynomial $P : \mathbb{R}\mathbb{S} \to \mathbb{R}[\alpha, \beta]$.

Lemma 11. The linear map $P : \mathbb{R}\mathbb{S} \to \mathbb{R}[\alpha, \beta]$ factors through $P : \mathcal{A} \to \mathbb{R}[\alpha, \beta]$.

Proof. We need to show that $P_{\sigma - L^{(l)}\sigma} = 0$ for any $\sigma \in \mathbb{S}_k$ and any $l \geq k$. It holds that

$$\begin{aligned} P_{\sigma - L^{(n)}\sigma} &= P_\sigma - \sum_{\tau}^{\mathbb{S}_l} d(\sigma, \tau) P_\tau \\ &= \lim_{n\to\infty} n^3 s_{n,\sigma} - \sum_{\tau}^{\mathbb{S}_l} d(\sigma, \tau) \lim_{n\to\infty} n^3 s_{n,\tau} \\ &= \lim_{n\to\infty} n^3 (s_{n,\sigma} - \sum_{\tau}^{\mathbb{S}_l} d(\sigma, \tau) s_{n,\tau}) \\ &= \lim_{n\to\infty} n^3 \cdot 0 = 0, \end{aligned}$$

where we use that P is linear on $\mathbb{R}\mathbb{S}$, $\lim_{n\to\infty}$ is linear on convergent sequences and $D_0 g_{n,\sigma} = D_0 g_{n,L^{(n)}\sigma} = \sum_{\pi}^{\mathbb{S}_n} d(\sigma, \pi) D_0 g_{n,\pi}$ by subsampling and linearity of D_0. □

As in [9], we for $l \geq k$ define *l-th fuzzy cover* of $\sigma \in \mathbb{S}_k$ to be the matrix $F_\sigma^{(l)} : \mathbb{R}^{l \to l}$ defined by

$$(F_\sigma^{(l)})_{u,v} = \frac{(l-k)!}{\binom{l}{k}} \sum_{j}^{[k]} f_{k,\sigma(j)}^{(l)}(u) f_{k,j}^{(l)}(v),$$

where

$$f_{k,j}^{(l)}(u) = \binom{u-1}{j-1}\binom{l-u}{k-j}.$$

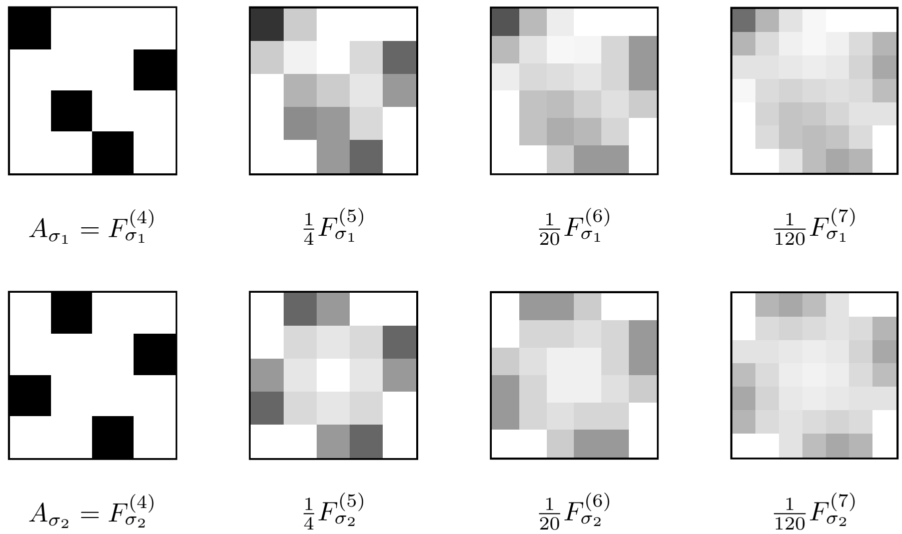


Figure 3.1: The 4th, 5th, 6th, and 7th lift of permutation $\sigma_1 = 1342$ and $\sigma_2 = 3142$

Note that if $l = k$, it holds that $f^{(l)}_{k,j}(u) = \mathbf{1}(u = j)$ from which it follows that $F^{(k)}_\sigma = A_\sigma$. We again extend the definition of lift cover and fuzzy cover linearly to space $\mathbb{RS}$; for $\sigma \in \mathbb{S}_k$ and $l < k$, define $A^{(l)}_\sigma = F^{(l)}_\sigma = 0$ for completeness, at no point of what follows is this pathology important. The following lemma establishes important connection between the notion of a lift cover $A^{(l)}\sigma$ and fuzzy cover $F^{(l)}\sigma$.

Lemma 12 ([9] Lemma 4.5)**.** Given $\sigma \in \mathbb{S}_k$ and $l \geq k$, it holds that

$$A^{(l)}_\sigma - F^{(l)}_\sigma = \frac{(l-1)!}{(k-1)!}\Big(\frac{1}{k} - \frac{1}{l}\Big)\mathbb{J}.$$

By the just stated lemma, it is immediate that $F^{(l)}_\sigma$ has eigenvector j for all $\sigma \in \mathbb{S}_k$ and $l \geq k$.We repeat a simple direct computation from [9] which shows that eigenvalue corresponding to eigenvector is $(l-1)!/(k-1)!$.

To determine the corresponding eigenvalue, we first note that

$$\sum_u^{[l]} f^{(l)}_{k,j}(u) = \sum_u^{[l]} \binom{u-1}{j-1}\binom{l-u}{k-j} = \sum_{u=j}^{l-k+j} \binom{u-1}{j-1}\binom{l-u}{k-j} = \binom{l}{k}$$

where the last equality is witnessed by a bijection which sends $A \in [l]^{(k)}$ to $(A_j, A \cap \{1, ..., A_j - 1\}, A \cap \{A_j + 1, ..., l\})$, where A_j is the j-th element in the set A according to the natural order on $A \subset [l]$. We conclude

$$\sum_{v}^{[l]}(F_\sigma^{(l)})_{u,v} = \frac{(l-k)!}{\binom{l}{k}}\sum_{j}^{[k]} f_{k,\sigma(j)}^{(l)}(u)\Big(\sum_{v}^{[l]} f_{k,j}^{(l)}(v)\Big) = (l-k)!\sum_{j}^{[k]} f_{k,j}^{(l)}(u)$$

that together with the equality

$$f_{k,j}^{(l)}(1) = \binom{0}{j-1}\binom{l-1}{k-j} = \mathbf{1}(j=1)\binom{l-1}{k-j} = \mathbf{1}(j=1)\binom{l-1}{k-1}$$

implies

$$\sum_{v}^{[l]}(F_\sigma^{(l)})_{1,v} = (l-k)!\sum_{j}^{[k]} f_{k,j}^{(l)}(1) = (l-k)!\binom{l-1}{k-1} = \frac{(l-1)!}{(k-1)!}.$$

We say that superpermutation $\nu \in \mathbb{RS}$ of order k has constant cover, if $F_\nu^{(k)}$ (or equivalently $A_\nu^{(k)}$) is a constant matrix, i.e., it is a scalar multiple of $\mathbb{J}$, and say that ν has non-vanishing constant fuzzy cover, if $F_\nu^{(k)} = c\mathbb{J}$ for $c \neq 0$. The next lemma shows that any dependency-witness has a constant cover.

Lemma 13. Let $\nu \in \mathbb{RS}$ be a superpermutation of order k with $P_\nu = 0$ and let $l \geq k$. Then $F_\nu^{(l)} \in \mathbb{RJ}$.

Proof. As ν and $\nu^{(l)} = L^{(l)}\nu \in \mathbb{RS}_l$ map to the same element in $\mathcal{A}$, we have that $P_{\nu^{(l)}} = 0$ which in turn by Corollary 10 implies that $A_{\nu^{(l)}} = A_\nu^{(l)} \in \mathbb{RJ}$. This together with linear extension of Lemma 12 concludes the proof. □

Superpermutations $\nu \in \mathbb{RS}$ with support of size four and non-vanishing constant fuzzy covers were fully described in [9].

Theorem 14 ([9] Proposition 4.22)**.** Let $\nu = \sum_i^{[4]} \nu_{\sigma_i}\sigma_i$ be a superpermutation with non-vanishing constant fuzzy cover, then one of the following cases occurs:

1. all permutations in the support are of order at most three,
2. all permutations in the support are of order four and form a Latin square, i.e., their cover matrices have disjoint supports as functions $[4]^2 \to \{0,1\}$,
3. up to action of symmetries of a square[1] and scaling by a scalar, it belongs

[1]We postpone the formal definition of the action to Chapter 5 page 36.

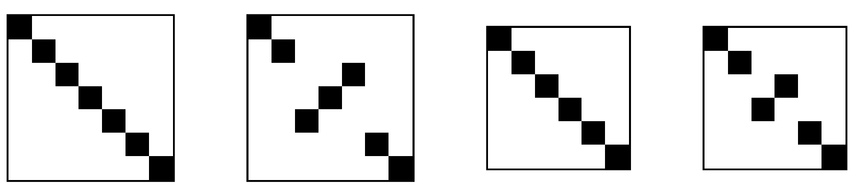
to the finite set

$$\begin{gathered}
3(1234)+3(4321)-4(123)+3(12),\\
3(1234)+3(4321)-4(123)-3(21),\\
3(1324)+3(4231)-4(123)+3(12),\\
3(1324)+3(4231)-4(123)-3(21),\\
3(2143)+3(3412)+4(123)+3(12),\\
3(2413)+3(3142)+4(123)+3(12),\\
3(1234)+3(4321)-2(123)-2(321),\\
3(1324)+3(4231)-2(123)-2(321),\\
3(2143)+3(3412)+2(123)+2(321),\\
3(2413)+3(3142)+2(123)+2(321),\\
36(12345)-36(52341)+15(2143)+10(321),\\
36(12345)-36(52341)-15(3412)-10(123),\\
36(12435)-36(52431)+15(2143)+10(321),\\
36(12435)-36(52431)-15(3412)-10(123).
\end{gathered}$$

The theorem above has an important consequence. Let S be a forcing set of size four whose largest permutation has order k and let ν be its dependency-witness. Then as $F_\nu^{(k)}$ is constant, either ν is one of only finitely many described in Theorem 14, or $F_\nu^{(k)}=0$. From now on, we denote the set of supports of permutations in Theorem 14 by $\mathcal{D}_+$. Using tools from next chapter, we establish that none of the sets in $\mathcal{D}_+$ is forcing, strengthening the statement that the superpermutations mentioned are not forcing.

Before continuing, let us discuss limitations of this method. It is clear that one cannot hope to exclude all sets using only the first-order information as there are examples—e.g., all sets contained in $\mathcal{D}_+$—that are dependent but one may hope that there is only finitely many such sets. The following infinite family of dependent sets shows that this is not the case.

Example 15. Let $l \geq 2$ be a positive integer and let S_l be a set containing permutations $\mathrm{id}_{2l+1} \in \mathbb{S}_{2l+1}$, $\tau_{l,l+2} \in \mathbb{S}_{2l+1}$, $\mathrm{id}_{2l} \in \mathbb{S}_{2l}$ and $\tau_{l,l+1} \in \mathbb{S}_{2l}$, where id_k denotes the identity element of symmetric group $\mathbb{S}_k$ and $\tau_{i,j} \in \mathbb{S}_l$ denotes transposition of $i \in [l]$ and $j \in [l]$. The set S_l is dependent.



id_7 $\tau_{3,5}$ id_6 $\tau_{3,4}$

Figure 3.2: Dependent set S_3.

Proof. We show that for $\nu = \mathrm{id}_{2l+1} - \tau_{l,l+2} - (2l+1)(l+1)^{-2}(\mathrm{id}_{2l} - \tau_{l,l+1})$, we have $P_\nu = 0$. Let $\nu_1 = \mathrm{id}_{2l+1} - \tau_{l,l+2}$ and $\nu_2 = \mathrm{id}_{2l} - \tau_{l,l+1}$. First, it is easy to see that $A_{\nu_1} = (e_l - e_{l+2})(e_l - e_{l+2})^T$ and similarly $A_{\nu_2} = (e_l - e_{l+1})(e_l - e_{l+1})^T$; for intuition see Figure 3.2. Our aim is to show that for vectors $v = D_{2l+1}^T B_{2l+1}^T (e_l - e_{l+2})$ and $u = D_{2l}^T B_{2l}^T (e_l - e_{l+1})$ we have that $v_i = (2l-1)(l+1)^{-1} u_i$ for $i \in [2l-1]$ and $v_i = u_i = 0$ for $i = 2l$. This shows the claim, for $T_{\nu_1} = c_{2l+1} vv^T$, $T_{\nu_2} = c_{2l} uu^T$, and identity

$$c_{2l+1} \frac{(2l-1)^2}{(l+1)^2} = \frac{(2l+1)2l}{(2l-1)!} \frac{(2l-1)^2}{(l+1)^2} = \frac{(2l-1)2l}{(2l-2)!} \frac{(2l+1)}{(l+1)^2} = c_{2l} \frac{2l+1}{(l+1)^2}$$

implies $T_{\nu_1} = (2l+1)(l+1)^{-2}\, T_{\nu_2}^{(2l-1)}$. Expanding the definition reveals for $k \geq 2$ and $a, b \in [k]$ the identity

$$(D_k^T B_k^T (e_a - e_b))_i = \binom{k-2}{i-1} (-1)^{i+1} \left((-1)^a \binom{i}{a-1} - (-1)^b \binom{i}{b-1} \right)$$

which in turn implies

$$v_i = (-1)^{i+1+l} \binom{2l-1}{i-1} \left(\binom{i}{l-1} - \binom{i}{l+1} \right), \text{ and}$$
$$u_i = (-1)^{i+1+l} \binom{2l-2}{i-1} \left(\binom{i}{l-1} + \binom{i}{l} \right) = (-1)^{i+1+l} \binom{2l-2}{i-1} \binom{i+1}{l}$$

for every $i \in [2l-1]$ and $v_i = 0$ for $i = 2l$. Now note that $v_i = u_i = 0$ for $i < l-1$ and otherwise a direct computation expanding the binomial coefficients reveals that

$$\left(\binom{i}{l-1} + \binom{i}{l+1} \right) \binom{i+1}{l}^{-1} = \frac{2l-i}{l+1}$$

and

$$\binom{2l-1}{i-1} \binom{2l-2}{i-1}^{-1} = \frac{2l-1}{2l-i}$$

by which we conclude that for $l-1 \leq i < 2l$—and consequently for $i < 2l$—it holds that $v_i = (2l-1)(l+1)^{-1} u_i$. $\square$

Chapter 4

Second-order Method

In this section, we review the main method used in [17], providing superficially different but morally equivalent proof of Theorem 5 and Corollary 11 which is the main mathematical result needed for our work. We then move to lemmas which allow us to, in a computer assisted fashion, certify that a dependent set is not forcing. We finish the section by describing certificate for dependent set S and illustrate the method on a simple example.

4.1 Walking in the Zero Set

We start by continuing the review of multivariable calculus. Given a smooth function $F : \mathbb{R}^n \to \mathbb{R}^m$, we define $H_x F$ to be the bilinear map induced by $D_x(y \mapsto D_y F)$: note that $D_x(y \mapsto D_y F)$ is element of the set $\mathcal{L}(\mathbb{R}^n, \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m))$ *a priori*, but there is a canonical isomorphism with the set of bilinear maps $\mathcal{L}(\mathbb{R}^n \otimes \mathbb{R}^n, \mathbb{R}^m)$ which maps φ to $(e_i \otimes e_j \mapsto \varphi(e_i)e_j)$. As such we have

$$\begin{aligned}(H_x F(e_i, e_j))_k &= H_x F_k(e_i, e_j) = D_x(y \to D_y F_k)(e_i)e_j \\ &= D_x(y \to D_y F_k e_j)(e_i) = \frac{\partial^2 F_k}{\partial x_i \partial x_j},\end{aligned}$$

where the second to last equality follows from a version of Leibniz rule which can be found in Appendix A. In what follows, we need to determine Hessian of a composition of two maps $F : \mathbb{R}^n \to \mathbb{R}^m$ and $G : \mathbb{R}^m \to \mathbb{R}^s$ which we do using the following equality

$$H_x(GF)(u, v) = H_{F(x)}G(D_x F u, D_x F v) + (D_{F(x)}G)H_x F(u, v)$$

which is derived in Appendix A. In what follows, we primarily study functions of the form $f : \mathbb{R}^n \to \mathbb{R}$ in which case $H_x f$ is a bilinear map of the form $\mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ and as such we identify it with a matrix $M \in R^{n \times n}$ for which it holds that $H_x f(u, v) = u^T M v$ which is standardly called the *Hessian matrix* of function f at x. It is a standard result of theory of smooth maps that such matrix is for a smooth map $f : \mathbb{R}^n \to \mathbb{R}$ a real symmetric matrix.

Let us start by stating a lemma from the area of differential topology that states function $f : \mathbb{R}^n \to \mathbb{R}$ without linear behaviour and with quadratic behaviour which is not degenerate behaves like $c + \sum_i^{[n]} (-1)^{l_i} x_i$ for $c \in \mathbb{R}$ and $l_i \in \mathbb{N}$ for all $i \in [m]$ up to smooth change of coordinates.

Lemma 16 (Morse Lemma)**.** ([14] Lemma 1.1) Let $f : \mathbb{R}^n \to \mathbb{R}$ be a smooth function with $D_0 f = 0$ (zero is a critical point), and with $H_0 f$ being invertible (the critical point is nondegenerate). Then there are neighbourhoods $U, V \subset \mathbb{R}^n$ of zero and a diffeomorphism $u : U \to V$ such that $u(0) = 0$ and for all $x \in U$ it holds that

$$f(x) = f(0) - \sum_{i=1}^{\lambda} u_i(x)^2 + \sum_{i=\lambda+1}^{n} u_i(x)^2,$$

where $\lambda = \lambda_0(f)$ is the *Morse index* of the critical point 0, i.e., λ is the number of negative eigenvalues of the Hessian $H_0 f$.

Let us for natural number $n \in \mathbb{N}$, smooth function $f : \mathbb{R}^n \to \mathbb{R}$, and a vector $x \in \mathbb{R}^n$ define d^+/d^- to be the dimension of the sum of positive/negative eigenspaces of $H_x f$. We define $\kappa_x(f) = \min\{d^+, d^-\}$.

Lemma 17. Let $f : \mathbb{R}^n \to \mathbb{R}$ be a smooth function with $f(0) = 0$, zero being a critical point, and $\kappa_0(f) \geq k$. Then there is an injective smooth map $p : [-1, 1]^k \to \mathbb{R}^n$ with $p(0) = 0$ such that $fp = f(0) = 0$.

Proof. As $\kappa_0(f) \geq k$, there is a linear injection $J : \mathbb{R}^{2k \to n}$ with the property that $J^T H_0 f J$ is an invertible matrix with k positive and k negative eigenvalues: consider map J which sends the standard basis to the system of $2k$ orthonormal eigenvectors that are guaranteed by the Spectral Theorem applied to $H_0 f$ and assumption $\kappa_0(f) \geq k$. For $g = fJ : \mathbb{R}^{2k} \to \mathbb{R}$, we have $D_0 g = (D_0 f)J = 0$ as zero is a critical point of f. Moreover, as J is a linear map, we have $H_0 J = 0$ from which it holds that

$$H_0 g(x, y) = H_{J(0)} f(Jx, Jy) + (D_{J(0)} f) H_0 J(x, y) = (J^T (H_0 f) J)(x, y)$$

and so as $H_0 g = J^T (H_0 f) J$, we have $\kappa_0(g) \geq k$ with zero being a nondegenerate critical point of g. By Morse Lemma, we have open sets $U, V \subset \mathbb{R}^{2k}$ and a diffeomorphism $u : U \to V$ such that for $x \in U$, it holds that

$$g(x) = g(0) - \sum_{i=1}^{k} u_i(x)^2 + \sum_{i=k+1}^{2k} u_i(x)^2 = g(0) + \sum_{i=1}^{k} (u_{k+i}(x)^2 - u_i(x)^2).$$

Let $\rho > 0$ be a real number such that the ball at origin of radius ρ is contained in V, formally $B_\rho(0) = \{x \in \mathbb{R}^{2k} \mid |x| < \rho\} \subset V$, and let $\delta > 0$ be a real number such that $\delta \leq \rho(4k)^{-1/2}$. Define a mapping $r : [-1, 1]^k \to \mathbb{R}^{2k}$ by

$$r(t) = \delta \sum_{i=1}^{k} t_i (e_i + e_{k+i}),$$

where recall that e_i is the i-th standard basis vector of $\mathbb{R}^{2k}$. Due to the inequality

$$\left\|\delta \sum_{i=1}^{k} t_i(e_i + e_{k+i})\right\|^2 = 2\delta^2 \sum_{i=1}^{k} t_i^2 \leq 2\delta^2 k \leq \frac{1}{2}\rho^2 < \rho^2,$$

we conclude that the image of r is contained in V. Define $q = u^{-1}r : [-1,1]^k \to U$, where the composition is well defined by the previous comment. Map q is injective as r is clearly injective by definition and u is a diffeomorphism. We have that

$$u_j(q(t)) = \left\langle e_j, uu^{-1}\Big(\delta \sum_{i=1}^{k} t_i(e_i + e_{k+i})\Big)\right\rangle = \delta \sum_{i=1}^{k} t_i(\mathbf{1}(j=i) + \mathbf{1}(j=k+i)).$$

As such $u_i(q(t)) = u_{k+i}(q(t)) = \delta t_i$ for every $i \in [k]$. At last, we have

$$g(q(t)) = g(0) + \sum_{i=1}^{k}(u_{k+i}(q(t))^2 - u_i(q(t))^2) = g(0).$$

Finally, setting p to be the composition of two injections Jq finishes the proof as $fp = fJq = gq = g(0) = f(0)$. ☐

Lemma 18. Let $F : \mathbb{R}^n \to \mathbb{R}^m$ be a smooth map with $F(0) = 0$ and with Jacobian D_0F of rank at least $m-1$. Let $w \in \mathbb{R}^m$ be a vector of norm one such that $w^T(D_0F) = 0$ and $\kappa_0(w^TF) \geq m$. Then there exists a smooth injective map $p : [-1,1] \to \mathbb{R}^n$ such that $Fp = 0$.

Proof. Let us define $P = (1 - ww^T)$ to be the linear map which projects $\mathbb{R}^m$ onto $w^\perp$. Let $F_1 : \mathbb{R}^n \to w^\perp \subset \mathbb{R}^m$ be defined as $F_1 = PF = (1 - ww^T)F$. As $w^\perp$ is $(m-1)$-dimensional and $w^TD_0F = 0$, we have that D_0F_1 is surjective onto $(m-1)$-dimensional subspace $w^\perp$. It follows from the Implicit Function Theorem that there exist open sets $U \subset \mathbb{R}^{n-m+1}$, $V \subset \mathbb{R}^{m-1}$, and a smooth map $\phi : U \to V$ with the property that for every $x \in U$ and $y \in V$ it holds that $F_1(x,y) = 0$ if and only if $y = \phi(x)$. Define $T : U \to \mathbb{R}^n$ to be a map which sends x to $(x, \phi(x))$ so that $F_1T = 0$ with $T(0) = 0$ by $F_1(0) = 0$. It is immediate that the Jacobian D_0T is an injective linear map.

Next, we examine the map $F_2 = w^TFT : U \to \mathbb{R}$; here, reader should view F_2 as a map $F - F_1 = ww^TF : \mathbb{R}^n \to w\mathbb{R}$ restricted to the $(n-m+1)$-dimensional manifold of the level set of PF around zero which we obtained by the Implicit Function Theorem. First, let us note that by chain rule, we have $D_0F_2 = w^T(D_0F)(D_0T) = 0$, where we use the fact that $T(0) = 0$ and $w^T(D_0F) = 0$. Moreover, we have that for all $x, y \in \mathbb{R}^n$ it holds that

$$\begin{aligned} H_0F_2(x,y) &= H_{T(0)}(w^TF)((D_0T)x, (D_0T)y) + D_0(w^TF)H_0T(x,y) \\ &= H_0(w^TF)((D_0T)x, (D_0T)y) \end{aligned}$$

where the last equality is due to the assumption $D_0(w^TF) = 0$. By assumption we have that $\kappa_0(w^TF) \geq m$ and so both the space of positive eigenvectors and the space of negative eigenvectors of $H_0(w^T F)$ are of dimension at least m. As D_0T is a linear injection onto a subspace $\mathrm{Im}\, T \subset \mathbb{R}^m$ of dimension $n-m+1$, i.e., of codimension $m-1$, it follows that $\mathrm{Im}\, T$ has nontrivial intersection with the space of positive and the space of negative eigenvectors of $H_0(w^T F)$. Thus, we conclude that matrix $(D_0T)^T H_0(w^T F)(D_0T)$ has both a positive and a negative eigenvalue, i.e., $\kappa_0(F_2) \geq 1$. From Lemma 17 applied to F_2, we obtain a smooth injective map $q : [-1,1] \to \mathbb{R}^{n-m+1}$ such that $F_2 q = F_2(0) = 0$. Finally, take $p = Tq : [-1,1] \to \mathbb{R}^n$ then

$$Fp = (1 - ww^T + ww^T)FTq = PFTq + ww^T FTq = F_1Tq + wF_2q = 0.$$

$\square$

The above lemma serves as the main mathematical tool of this chapter. Before stating the problem-specific corollary, we first remark that $\kappa_0(w^T G_{n,S})$ can be determined from eigenvalues of $\sum_\pi^S w_\pi (H_0 g_{n,\pi}) \in \mathbb{R}^{n\times n}$ as due to repeated application of linearity of derivatives, we get

$$\begin{aligned} H_0(w^T G_{n,S}) &= D_0(x \mapsto D_x(w^T G_{n,S})) = D_0(x \mapsto w^T D_x G_{n,S}) \\ &= D_0(x \mapsto \sum_\pi^S w_\pi (D_x G_{n,S})_\pi) = \sum_\pi^S w_\pi D_0\Big(x \mapsto D_x g_{n,\pi}\Big) \\ &= \sum_\pi^S w_\pi (H_0 g_{n,\pi}). \end{aligned}$$

Corollary 19 ([17] Corollary 11)**.** Let $Q \subset S \subset \mathbb{S}$ be a set of permutations with $|Q| + 1 = |S| = m$, where Q is n-independent and S is n-dependent with $w^T \in \mathbb{R}^{m\to 1}$ being the witness to this fact, i.e., $w^T D_0 G_{n,S} = 0$. If the matrix

$$H_0^w G_{n,S} := \sum_\pi^S w_\pi H_0 g_{n,\pi} \in \mathbb{R}^{n\to n}$$

has at least m positive and m negative eigenvalues, then S is not $\mathcal{P}_n$-forcing.

Proof. Due to Q being n-independent, we have that the rank of $D_0 G_{n,S}$ is $m-1$. Let $w \in \mathbb{R}^m$ be the non-zero element of the kernel of $(D_0 G_{n,S})^T$ of norm one. By Lemma 18, there exists an injective map $p : [-1,1] \to \mathbb{R}^{(n-1)^2}$ such that $G_{n,S} p = G_S|_{\mathcal{P}_n} \zeta_n p = 0$. As such permuton $\mu = \zeta_n(p(\epsilon))$ for $\epsilon > 0$ is a non-uniform permuton with $G_S|_{\mathcal{P}_n} \mu = 0$ contradicting $\mathcal{P}_n$-forcing of S. $\square$

As our approach is computational, we need to be aware of numerical precision and problems associated with doing numerical computations. For this, we introduce simple lemmas which allow us to safely certify lower bound on the count of positive resp. negative eigenvalues of a matrix.

Given a set of vectors $(v_i)_i^{[k]}$ in $\mathbb{R}^n$, we define its associated linear map to be map $V : \mathbb{R}^{k\to n}$ sending e_i to v_i. We say that collection of vectors $(v_i)_i^{[k]}$ is orthonormal if it holds that $V^T V = I_k \in \mathbb{R}^{k \to k}$, where I_k is the identity map on $\mathbb{R}^k$. Given a symmetric matrix $M : \mathbb{R}^{n\times n}$, we define its associated symmetric bilinear form $\langle x, y\rangle_M$[1] to be $x^T M y$.

Lemma 20. Let $\delta > 0$, $M \in \mathbb{R}^{n\times n}$ be a real symmetric matrix, and $(v_i)_i^{[m]} \in (\mathbb{R}^n)^{[m]}$ be an orthogonal collection of vectors with the property that for all $i, j \in [m]$ either $i = j$ and $\langle v_i, v_j\rangle_M \geq \delta$ or $i \neq j$ and $|\langle v_i, v_j\rangle_M| < \delta/(m^2 - m)$. Then for all $u \in \mathbb{R}\{v_i\}_i^{[m]}$ we have that $\langle u, u\rangle_M > 0$.

Proof. Let $V : \mathbb{R}^{m\to n}$ be the linear map associated to the collection $(v_i)_i^{[m]}$. We want to show that for $u \in \text{Im } V \subset \mathbb{R}^n$ it holds that $\langle u, u\rangle_M > 0$. It is enough to consider $u = Vw = \sum_i^{[m]} w_i v_i$ with $\|w\| = 1$. After expanding the product and applying the triangle inequality, we obtain

$$\begin{aligned}
\langle u, u\rangle_M &= \sum_{i,j}^{[m]^2} w_i w_j \langle v_i, v_j\rangle_M = \sum_i^{[m]} w_i^2 \langle v_i, v_i\rangle_M + \sum_{i\neq j}^{[m]^2} w_i w_j \langle v_i, v_j\rangle_M \\
&\geq \delta \sum_i^{[m]} w_i^2 - \Big|\sum_{i\neq j}^{[m]^2} w_i w_j \langle v_i, v_j\rangle_M\Big| \geq \delta - \sum_{i\neq j}^{[m]^2} |w_i||w_j||\langle v_i, v_j\rangle_M| \\
&> \delta - \sum_{i\neq j}^{[m]^2} \frac{\delta}{m(m-1)} = \delta - \delta = 0,
\end{aligned}$$

where note that for $\|w\| = 1$, we have $|w_i| \leq 1$ for all $i \in [m]$. $\square$

Lemma 21. Let $\delta > 0$, $M : \mathbb{R}^{n\times n}$ be a real symmetric matrix, and $(v_i)_i^{[m]} \in (\mathbb{R}^n)^{[m]}$ be an orthogonal collection of vectors as in Lemma 20. Then M has at least m positive eigenvalues.

Proof. As M is a real symmetric matrix, it holds by the Spectral Theorem that there exists a real orthogonal matrix V and a real diagonal matrix Λ such that $M = W\Lambda W^T$, W is associated with an orthonormal system of eigenvectors of M, and $(\Lambda_{i,i})_i^{[n]} = (\lambda_i)_i^{[n]}$ is a nonincreasing sequence of eigenvalues of M.

For contradiction, assume that $\lambda_m \leq 0$ so that M has at most $m-1$ positive eigenvalues. Let N be the subspace spanned by eigenvectors $(w_i)_{i=m}^n$ with non-positive eigenvalues. Let $V \subset \mathbb{R}^n$ be the subspace spanned by $(v_i)_i^{[m]}$. As $\dim N + \dim V = n - m + 1 + m = n + 1$, we conclude that $V \cap N \neq 0$. Let $v \in V \cap N$ be a non-zero vector. As v is contained in the subspace N, it can be written as a linear combination $v = \sum_{i=m}^n v_i w_i$ meaning $Mv = \sum_{i=m}^n v_i \lambda_i w_i$

[1] Note that the product may not be positive-definite, i.e., it is not an inner product on the space $\mathbb{R}^n$.

and so at last we obtain inequality

$$\begin{aligned}\langle v, v\rangle_M = \langle v, Mv\rangle &= \Big\langle \sum_{i=m}^{n} v_i w_i, \sum_{i=m}^{n} v_i \lambda_i w_i \Big\rangle \\ &= \sum_{i=m,j=m}^{n} v_i v_j \lambda_j \langle w_i, w_j\rangle = \sum_{i=m}^{n} v_i^2 \lambda_i \leq 0\end{aligned}$$

where the last inequality is due to λ_i being non-positive for all $i \geq m$. The just made conclusion contradicts Lemma 20 as $v \in V$. ☐

We finish this section by stating a problem-specific corollary which we call *second-order certificate lemma* as it describes the main sufficient condition of this section. Given a matrix $A : \mathbb{R}^{n\times n}$, let us define the *diagonal gap* $\gamma : \mathbb{R}^{n\times n} \to [0, +\infty]$ to be

$$\gamma(A) = \frac{\min_i^{[n]} |A_{i,i}|}{\max_{i\neq j}^{[n]^2} |A_{i,j}|}$$

if there is a non-zero element off the diagonal, otherwise set $\gamma(A)$ to be the positive infinity.

Definition 1. Let $M : \mathbb{R}^{n\times n}$ be a real symmetric matrix. We call the linear map $V : \mathbb{R}^{m\to n}$ associated to a collection of vectors $(v_i)_i^{[m]} \subset \mathbb{R}^n$ a *positive/negative m-witness* for M, if

(i) $V^T V$ is diagonal, i.e., $(v_i)_i^{[m]}$ is an orthogonal set,

(ii) $V^T M V$ has positive/negative diagonal, i.e., all elements on diagonal are greater/less than zero, and

(iii) $\gamma(V^T M V) > m(m-1)$.

Lemma 22. If V is a positive m-witness for a matrix $M : \mathbb{R}^{n\times n}$, then V is a negative m-witness for matrix $(-M) : \mathbb{R}^{n\to n}$

Proof. Property *(i)* is independent of M. For property *(ii)* note that $\gamma(-A) = \gamma(A)$. Finally, $(V^T(-M)V)_{i,i} = -(V^T M V^T)_{i,i} < 0$ by V being a positive m-witness. ☐

Lemma 23. If V is a positive m-witness for a real symmetric matrix $M : \mathbb{R}^{n\times n}$, then M has at least m positive eigenvalues.

Proof. As $V^T V$ is a diagonal matrix with positive diagonal, the set $(v_i)_i^{[m]}$ associated with V is orthogonal. Let $\delta = \min_i^{[m]} \langle v_i, v_i\rangle_M > 0$ so that for all $i \in [m]$ the inequality $\langle v_i, v_i\rangle_M \geq \delta$ holds. By $\gamma(V^T M V) > m(m-1)$ for all $i, j \in [m]$ with $i \neq j$, it holds that

$$|\langle v_i, v_j\rangle_M| \leq \max_{i\neq j}^{[m]^2} |\langle v_i, v_j\rangle_M| \leq \gamma(V^T M V)^{-1} \min_i^{[m]} |\langle v_i, v_i\rangle_M| < \frac{\delta}{m(m-1)}.$$

By Lemma 21 with vectors $(v_i)_i^{[m]}$, we have that matrix M has at least m positive eigenvalues. $\square$

Lemma 24 (Certification Lemma)**.** Let $n \in \mathbb{N}$ and let $S \subset \mathbb{S}$ be a set of four permutations with $D_0 G_{n,S}$ of rank three together with n-dependence witness $w \in \mathbb{R}^S$, i.e., we have $w^T D_0 G_{n,S} = 0$. If there is a positive 4-witness and a negative 4-witness for matrix $H_0^w G_{n,S}$, then S is not $\mathcal{P}_n$-forcing.

Proof. Let V_+, V_- be the positive and negative m-witnesses respectively. By Lemma 23 with V_+ and $H := H_0^w G_{n,S}$, the hessian H has at least m positive eigenvalues. By Lemma 23 and Lemma 22 with V_-, the hessian H has at least m negative eigenvalues.

As $D_0 G_{n,S}$ has rank three and $\kappa_0(w^T G_{n,S}) \geq m$, we conclude by Lemma 18 that there exists an injective smooth path $p : [-1,1] \to \mathbb{R}^n$ such that $G_{n,S} p = 0$. Let ϵ be a real positive number, then for any $\sigma \in S$ we have $(G_{n,S}(p(\epsilon)))_\sigma = d(\sigma, \zeta(p(\epsilon))) - d(\sigma, \rho) = 0$ even though $\zeta(p(\epsilon))$ is not the uniform measure. $\square$

4.2 Example

Throughout what follows, we heavily rely on a large number of computationally obtained certificates as described in Lemma 24. Due to the number of certificates and their size, we do not include them in the main body of the text and instead point reader to Appendix D whenever it is appropriate. We devote this section to a concrete example use of the method, namely we use it to show that the set $S = (1234, 1324, 2143, 123)$ is not forcing.

Let $n = 4$ be the order of the perturbed permuton being considered, i.e., our goal is to prove that S is not $\mathcal{P}_4$-forcing. We display $D = D_0 G_{4,S}$ on Figure 4.1 and claim that $w = (1, 4/5, 3/5, -4/5)^T$ generates kernel of D^T, which implies that set S is 4-dependent and moreover that $D_0 G_{4,S}$ is of rank 3. Our aim is to show that Hessian $H_0^w G_{4,S}$ displayed on Figure 4.2 has four positive and four negative eigenvalues. To this end, we use Lemma 24 with positive 4-witness V_+ displayed on Figure 4.3 and negative 4-witness V_- displayed on Figure 4.4. To see that V_+ is indeed a positive witness, direct computation whose result is displayed on Figure 4.3 shows that the matrix $V_+ V_+^T$ is a diagonal matrix, values on diagonal of matrix $V_+^T H_0^w G_{n,S} V_+$ are indeed positive, and finally that we have the inequality $\gamma(V_+^T H_0^w G_{4,S} V_+) = 8638640/311781 \approx 27.7074 > 12$. Similarly, we verify that V_- is a negative 4-witness with $\gamma(V_-^T H_0^w G_{4,S} V_-) = \infty > 12$. Due to Lemma 24 the set S cannot be forcing.

$$D^T = \frac{1}{24576}\begin{pmatrix} 1061 & 1052 & -483 & 2016 \\ 419 & 356 & 363 & 1152 \\ -91 & -100 & 669 & 288 \\ 419 & 356 & 363 & 1152 \\ 533 & 92 & 525 & 1152 \\ 419 & 356 & 363 & 1152 \\ -91 & -100 & 669 & 288 \\ 419 & 356 & 363 & 1152 \\ 1061 & 1052 & -483 & 2016 \end{pmatrix}$$

Figure 4.1: Jacobian $D = D_0 G_{4,S} : \mathbb{R}^{9\to 4}$ for $S = \{1234, 1324, 2143, 123\}$. The rank of the Jacobian is 3 and vector $w = (1, 4/5, 3/5, -4/5)^T$ spans the kernel of D^T.

$$H = \frac{1}{640}\begin{pmatrix} 7 & -5 & -9 & -5 & -20 & 21 & -9 & 21 & 27 \\ -5 & 7 & -5 & 31 & -5 & -17 & -3 & -9 & 21 \\ -9 & -5 & 7 & -3 & 28 & -5 & 3 & -3 & -9 \\ -5 & 31 & -3 & 7 & -5 & -9 & -5 & -17 & 21 \\ -20 & -5 & 28 & -5 & 7 & -5 & 28 & -5 & -20 \\ 21 & -17 & -5 & -9 & -5 & 7 & -3 & 31 & -5 \\ -9 & -3 & 3 & -5 & 28 & -3 & 7 & -5 & -9 \\ 21 & -9 & -3 & -17 & -5 & 31 & -5 & 7 & -5 \\ 27 & 21 & -9 & 21 & -20 & -5 & -9 & -5 & 7 \end{pmatrix}$$

Figure 4.2: Hessian $H = H_0^w G_{4,S} : \mathbb{R}^{9\times 9}$ for $S = \{1234, 1324, 2143, 123\}$.

$$V_+^T = c_1 \begin{pmatrix} 0 & -29050 & -20750 & -36485 \\ -3320 & -21580 & 39010 & -14084 \\ 58100 & -32370 & 0 & 28259 \\ 3320 & -21580 & 39010 & -14084 \\ 0 & -34860 & 0 & 43202 \\ -3320 & -21580 & -39010 & -14084 \\ -58100 & -32370 & 0 & 28259 \\ 3320 & -21580 & -39010 & -14084 \\ 0 & -29050 & 20750 & -36485 \end{pmatrix}$$

$$V_+V_+^T = c_2 \begin{pmatrix} 2046780 & 0 & 0 & 0 \\ 0 & 2066700 & 0 & 0 \\ 0 & 0 & 2092845 & 0 \\ 0 & 0 & 0 & 2084127 \end{pmatrix}$$

$$V_+HV_+^T = c_3 \begin{pmatrix} 7170071200 & 0 & 0 & 0 \\ 0 & 23077461100 & 0 & 258778230 \\ 0 & 0 & 135180091400 & 0 \\ 0 & 258778230 & 0 & 145431644139 \end{pmatrix}$$

Figure 4.3: The positive 4-witness V_+ for $S = \{1234, 1324, 2143, 123\}$ with $n = 4$, diagonal dap $\gamma_+ = 8638640/311781 \approx 27.7074$, $c_1 = 83000^{-1}$, $c_2 = 2075000^{-1}$, and $c_3 = 1102240000000^{-1}$.

$$V_-^T = c_1 \begin{pmatrix} -66 & 12 & 0 & 0 \\ -18 & -4 & 50 & 50 \\ 0 & -45 & 6 & 0 \\ -18 & -4 & -50 & -50 \\ 0 & 74 & 0 & 0 \\ 18 & -4 & 50 & -50 \\ 0 & -45 & -6 & 0 \\ 18 & -4 & -50 & 50 \\ 66 & 12 & 0 & 0 \end{pmatrix}$$

$$V_- V_-^T = c_2 \begin{pmatrix} 5004 & 0 & 0 & 0 \\ 0 & 4939 & 0 & 0 \\ 0 & 0 & 5036 & 0 \\ 0 & 0 & 0 & 5000 \end{pmatrix}$$

$$V_- H V_-^T = c_3 \begin{pmatrix} -21150 & 0 & 0 & 0 \\ 0 & -20097 & 0 & 0 \\ 0 & 0 & -20282 & 0 \\ 0 & 0 & 0 & -10000 \end{pmatrix}$$

Figure 4.4: The negative 4-witness V_- for $S = \{1234, 1324, 2143, 123\}$ with $n = 4$, diagonal dap $\gamma_- = +\infty$, $c_1 = 100^{-1}$, $c_2 = 5000^{-1}$, and $c_3 = 400000^{-1}$.

Chapter 5

Profiles of Forcing Sets

In this chapter, we systematically investigate all possible four element permutation sets, based on the order profile of the permutations in the quadruple. Formally, let *profile* of a set $S \subset \mathbb{S}$ be the vector $(|\pi|)_\pi^S$ where the permutations in S are ordered so that $\pi < \sigma$ if $|\sigma| < |\pi|$ or π and σ are of the same size and π is before σ in the lexicographical ordering of $\mathbb{S}_{|\sigma|}$; the profile vector is thus non-increasing. Throughout this section, we for $\nu \in \mathbb{RS}$ with support of size m use enumeration $(\sigma_i)_i^{[m]}$ of the set supp ν according to the just defined ordering, i.e., $\sigma_i < \sigma_j$ if and only if $i < j$, and use $\nu_i := \nu_{\sigma_i}$ for every $i \in [m]$.

We start by describing necessary condition on the shape of the profile of a forcing set, namely we discuss result shown in [19] that implies that any profile of a forcing set $S \in \mathbb{S}^{(4)}$ needs to be of the form (k_1, k_1, k_2, k_3), where $k_2 = k_1$ or $k_2 = k_1 - 1$. We review a rephrasing of the argument, obtaining lemmas that are needed in what follows.

Lemma 25. Let $\nu \in \mathbb{RS}_k$ be a superpermutation, where all permutations in the support are of size k and let $u \in \mathbb{R}^{k-1}$. If $T_\nu u = 0$, then $A_\nu B_k D_k u = 0$. Moreover, if $u = e_i$ is a basis vector, then $A_\nu B_k e_i = 0$.

Proof. For the proof, let $D = D_k, B = B_k$, and $c = c_k$ as defined in Section 3.1. Let $u \in \mathbb{R}^{k-1}$ be a non-zero vector (for otherwise the claim follows immediately) such that $T_\nu u = cD^T B^T A_\nu BDu = 0$. Then $B^T A_\nu BDu = 0$ as $c \neq 0$ and $D = D^T$ is a linear isomorphism. It follows that $A_\nu BDu \in \ker B^T = \mathbb{R}j$ and thus $A_\nu BDu = aj$ for some scalar $a \in \mathbb{R}$. As $j \in \mathbb{R}^k$ is an eigenvector of both A_ν and A_ν^T, we have $j^T A_\nu BDu = (\sum_\sigma^{\text{supp}\ \nu} \nu_\sigma) j^T BDu = 0$ as $j^T B = 0$, and $j^T A_\nu BDu = j^T aj = ak$. Thus $a = 0$ holds and consequently $A_\nu BDu = 0$ for $u \neq 0$. If $u = e_i$, simply note that $De_i = de_i$ for some non-zero scalar d by definition of D. $\square$

Let Ψ_k be an operator on matrices $\mathbb{R}^{k \times k}$ which "horizontally flips" their matrix representation in the standard basis, i.e., the operator Ψ_k can be seen as precomposition by the permumation matrix associated to $\psi_k = (k, k-1, ..., 1) \in$

$\mathbb{S}_k$ so that $(\Psi_k M)e_i = MA_{\psi_k}e_i = Me_{k-i+1}$ for every $i \in [k]$ and every matrix $M \in \mathbb{R}^{k\times k}$.

Lemma 26. Let $\nu = \nu_1\sigma_1 + \nu_2\sigma_2$ with permutations $\sigma_1, \sigma_2 \in \mathbb{S}_k$, non-zero scalars $\nu_1, \nu_2 \in \mathbb{R}\backslash\{0\}$, and T_ν singular. Then $|\nu_1| = |\nu_2|$. Moreover, if $T_\nu e_{k-1} = 0$, then $\nu_2 = (-1)^k\nu_1$, $\Psi_k A_\rho = (-1)^k A_\rho$ and if k is even, $\sigma_2 = \sigma_1\psi_k$.

Proof. For the proof, let $D = D_k, B = B_k$, and $c = c_k$ with $k = |\sigma_1| = |\sigma_2|$ as defined in Section 3.1. If T_ν is singular, then there is non-zero vector $u \in \mathbb{R}^{k-1}$ such that $T_\nu u = 0$. By Lemma 25, there exists a non-zero vector w such that $A_\nu B w = 0$. As B is injective, this implies that there exists a non-zero vector v such that $0 = A_\nu v = \nu_1 A_{\sigma_1} v + \nu_2 A_{\sigma_2} v$ from which it follows that $\|\nu_1 A_{\sigma_1} v\| = \|\nu_2 A_{\sigma_2} v\|$ and consequently that $|\nu_1|\,\|v\| = |\nu_2|\,\|v\|$ as permuting coordinates of v does not change its norm. The fact that $v \neq 0$ shows $|\nu_1| = |\nu_2|$.

Now assume that $T_\nu e_{k-1} = 0$. By Lemma 25, we have that $A_\nu B_k e_{k-1} = 0$, i.e., we obtain the equality $\nu_1 A_{\sigma_1} B_k e_{k-1} = -\nu_2 A_{\sigma_2} B_k e_{k-1}$. Viewing the equality coordinate-wise, we obtain that it is true for every $l \in [k]$ that

$$\nu_1(-1)^{\sigma_1^{-1}(l)}\binom{k-1}{\sigma_1^{-1}(l)-1} = \nu_2(-1)^{1+\sigma_2^{-1}(l)}\binom{k-1}{\sigma_2^{-1}(l)-1} \tag{5.1}$$

which together with $|\nu_1| = |\nu_2|$ implies that

$$\binom{k-1}{\sigma_1^{-1}(l)-1} = \binom{k-1}{\sigma_2^{-1}(l)-1}$$

and thus either $\sigma_1^{-1}(l) = \sigma_2^{-1}(l)$ or $\sigma_2^{-1}(l) = 1 + k - \sigma_1^{-1}(l)$. As $\sigma_1 \neq \sigma_2$, we have that there exists l such that $\sigma_2^{-1}(l) \neq \sigma_1^{-1}(l)$ and consequently $\sigma_2^{-1}(l) = 1 + k - \sigma_1^{-1}(l)$. Returing to 5.1 with this choice of l, we see that

$$\operatorname{sgn}(\nu_1)\operatorname{sgn}(\nu_2)(-1)^{1+\sigma_1^{-1}(l)+\sigma_2^{-1}(l)} = \operatorname{sgn}(\nu_1)\operatorname{sgn}(\nu_2)(-1)^k = 1$$

from which we conclude that $\nu_1 = (-1)^k\nu_2$. Finally, our aim is to show that $\Psi_k A_\nu = (-1)^k A_\nu$ for every k and that $\sigma_2 = \sigma_1\psi_k$ if k is even. We start with the latter. If k is even, we see that there is no $l \in [k]$ such that $\sigma_1^{(-1)}(l) = \sigma_2^{(-1)}(l)$: otherwise, returing to 5.1 with such l shows that $\operatorname{sgn}(\nu_1)\operatorname{sgn}(\nu_2)(-1) = 1$, contradicting $\nu_1 = (-1)^k\nu_2 = \nu_2$. As such, we have $\sigma_2^{-1}(l) = 1 + k - \sigma_1^{-1}(l)$ for every l and so $\sigma_2^{-1} = \psi_k\sigma_1^{-1}$ which in turn implies $\sigma_2 = \sigma_1\psi_k$. For k even, we have

$$\Psi_k A_\nu = \Psi_k(\nu_1 A_{\sigma_1} + \nu_1 A_{\sigma_1} A_{\psi_k}) = \nu_1 A_{\sigma_1} A_{\psi_k} + \nu_1 A_{\sigma_1} = A_\nu.$$

Let k be odd. If $l \in [k]$ is such that $\sigma_1^{-1}(l) = \sigma_2^{-1}(l)$, then

$$e_l^T(\Psi_k A_\nu) = (\nu_1 e^T_{\sigma_1^{-1}(l)} - \nu_1 e^T_{\sigma_2^{-1}(l)})A_{\psi_k} = 0A_{\psi_k} = 0 = (-1)e_l^T A_\nu.$$

If l is such that $\sigma_1^{-1}(l) = 1 + k - \sigma_2^{-1}(l)$, then

$$\begin{aligned} e_l^T(\Psi_k A_\nu) &= (\nu_1 e^T_{\sigma_1^{-1}(l)} - \nu_1 e^T_{1+k-\sigma_1^{-1}(l)})A_{\psi_k} = \nu_1 e^T_{1+k-\sigma_1^{-1}(l)} - \nu_1 e^T_{\sigma_1^{-1}(l)} \\ &= (-1)e_l^T A_\nu. \end{aligned}$$

As such, we in both cases have that $\Psi_k A_\nu = (-1)^k A_\nu$. $\square$

The following lemma can be deduced from Lemma 13 and proof of Lemma 15 in [19].

Lemma 27 ([19])**.** Let $\nu \in \mathbb{RS}$ with supp $\nu = \{\sigma_1, ..., \sigma_m\}$ for $m \geq 3$ such that $|\sigma_1| \geq |\sigma_2| \geq ... \geq |\sigma_m|$ and $P_\nu = 0$. Then $|\sigma_1| = |\sigma_2|$ and $|\sigma_3| \in \{|\sigma_1|, |\sigma_1| - 1\}$.

Proof. Let $|\sigma_1| = k$. Note that T_{σ_1} is a matrix of rank $k-1$: let $u \in \mathbb{R}^{k-1}$ be a vector such that $T_{\sigma_1} u = 0$, then by Lemma 25 it holds that $A_{\sigma_1} B_k D_k u = 0$ and so $u = 0$ as $A_{\sigma_1} B_k D_k$ is an injective linear map. As T_{σ_1} has full rank, the first row contains non-zero value, i.e., there exists $j \in [k-2]_0$ such that $\langle \alpha^{(k-2)} \beta^j \mid P_{\sigma_1} \rangle \neq 0$. Assuming $|\sigma_2| < |\sigma_1|$, we get that polynomial $P_\nu - \nu_1 P_{\sigma_1}$ is in each variable of degree at most $k-3$ which in turn implies that the polynomial $P_\nu = P_\nu - \nu_1 P_{\sigma_1} + \nu_1 P_{\sigma_1}$ cannot be zero, contradicting the assumption of the statement.

For contradiction, assume $|\sigma_3| < |\sigma_1| - 1 = |\sigma_2| - 1$. Let $\rho = \nu_1 \sigma_1 + \nu_2 \sigma_2$. Due to argument similar to the one above, we have that $P_\rho = \nu_1 P_{\sigma_1} + \nu_2 P_{\sigma_2}$ has no monomials with variable of degree at least $k-3$, i.e. we have that both $T_\rho e_{k-1} = 0$ and $T_\rho e_{k-2} = 0$. Lemma 26 implies that $|\nu_1| = |\nu_2|$ and Lemma 25 implies that $A_\rho B_k e_{k-1} = A_\rho B_k e_{k-2} = 0$. Let $i \in [k]$ be an integer for which $a = \sigma_1^{-1}(i) \neq \sigma_2^{-1}(i) = b$, then $e_i^T A_\rho = \nu_1 e_a^T + \nu_2 e_b^T$. As $|\nu_1| = |\nu_2|$ we have that $|e_a^T B_k e_{k-1}| = |e_b^T B_k e_{k-1}|$ and $|e_a^T B_k e_{k-2}| = |e_b^T B_k e_{k-2}|$ which by definition of B_k implies $\binom{k-1}{a-1} = \binom{k-1}{b-1}$ and $\binom{k-2}{a-1} = \binom{k-2}{b-1}$ which together with $a \neq b$ leads to a contradiction. $\square$

Throughout what follows, we need to enumerate all possible sets with given profile. The following lemma shows that we need only to consider sets modulo action of the dihedral group D_4. Let D_4 be the group of symmetries of a square, formally given by presentation $\langle \alpha, \beta \mid \alpha^4 = \beta^2 = 1, \alpha\beta = \beta^{-1}\alpha^{-1} \rangle$. There is a natural action $\rho_n : D_4 \to \mathrm{Aut}(\mathbb{R}^{n \times n})$ on the space of real square matrices which "horizontally flips" the matrix under β, i.e., $\rho_n(\beta) := \Psi_n$, and rotates the matrix anti-clockwise by 90 degrees under α, i.e., $\rho_n(\alpha)(M) := (\Psi_n M)^T$. Note that the subset of permutation matrices is invariant under this action. To define action $\zeta : D_4 \to \mathrm{Aut}(\mathbb{S})$, we send $g \in D_4$ to a map which sends π to the permutation associated with the matrix $\rho_{|\pi|}(g)(A_\pi)$. We extend this action to m-element permutation sets $\mathbb{S}^{(m)}$ and the space of superpermutations $\mathbb{RS}$ naturally. From now on, we use $g\pi$, gS, and $g\nu$ to denote $\zeta(g)(\pi)$, $\zeta(g)(S)$, and $\zeta(g)(\nu)$.

Lemma 28. Let S be a finite set of permutations and let $g \in D_4$ be a dihedral group element. Then S is forcing if and only if gS is.

Proof. Let us show that $d(\sigma, \pi) = d(g\sigma, g\pi)$. Let $k = |\sigma|$ and $n = |\pi|$. To show $d(\sigma, \pi) \leq d(g\sigma, g\pi)$ it is enough to construct an injection from the set $A = \{P \in [n]^{(k)} \mid \pi|_P = \sigma\}$ to the set $B = \{Q \in [n]^{(k)} \mid (g\pi)|_Q = g\sigma\}$, where the other inequality holds by considering $d(g\sigma, g\pi) \leq d(g^{-1}g\sigma, g^{-1}g\pi) = d(\sigma, \pi)$.

Let $g_1 = \beta = \beta^{-1}$, then map $P \in A$ to $Q = 1 + n - P = \{1 + n - p\}_p^P$: the map is clearly an injection and $(\beta\pi)|_Q = \beta(\pi|_P) = \beta\sigma$ as desired. Let $g_2 =$

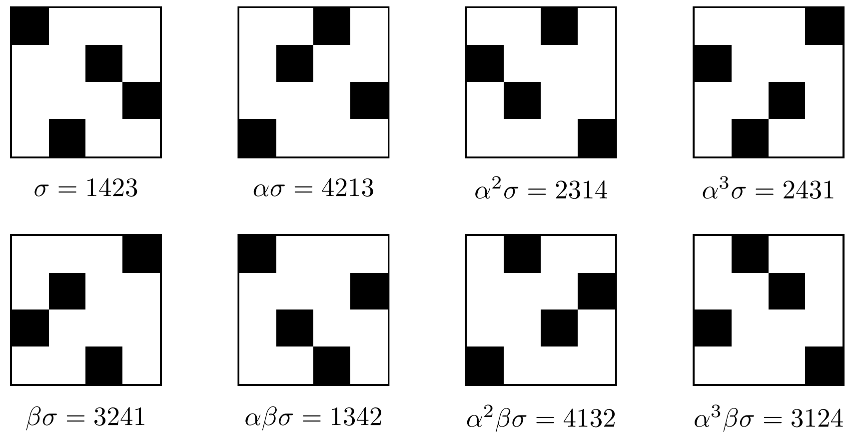


Figure 5.1: Orbit of permutation $\sigma = 1423$ under the action of dihedral group D_4.

$\alpha\beta = \beta^{-1}\alpha^{-1}$, i.e., g_2 is the group element sending π to π^{-1}. To construct the injection, map $P \in A$ to $Q = \pi(P) = \{\pi(p)\}_p^P$. We need to show that $Q \in B$. As $\pi|_P = \sigma$ it holds that $\pi(P_i) = \pi(P)_{\sigma(i)}$ with $(P_i)_i^{[k]}$ being the enumeration of the set $P \subseteq [n]$ under the natural order and similarly for enumeration $(\pi(P)_i)_i^{[k]}$ of the set $\pi(P) \subseteq [n]$. Consequently, it holds that $\pi^{-1}|_{\pi(P)} = \sigma^{-1}$ as $\pi^{-1}(\pi(P)_i) < \pi^{-1}(\pi(P)_j)$ holds if and only if $\pi^{-1}(\pi(P_{\sigma^{-1}(i)})) < \pi^{-1}(\pi(P_{\sigma^{-1}(j)}))$ which holds if and only if $P_{\sigma^{-1}(i)} < P_{\sigma^{-1}(j)}$ which holds if and only if $\sigma^{-1}(i) < \sigma^{-1}(j)$. As any $g \in D_4$ can be expressed as string of g_1 and g_2, e.g., $\alpha = \alpha^{-1} = g_2 g_1$ and $\beta = \beta^{-1} = g_1$, the general claim follows by transitivity of $\leq$.

Assume that S is forcing. To show that gS is forcing, let $\pi_\bullet$ be a convergent sequence so that for every $g\sigma \in gS$ it holds that

$$0 = d(g\sigma, \pi_\bullet) - d_\rho(g\sigma) = \lim_{n\to\infty} d(g\sigma, \pi_n) - d_\rho(g\sigma) = \lim_{n\to\infty} d(\sigma, g^{-1}\pi_n) - d_\rho(\sigma),$$

where we use that $d_\rho(g\sigma) = 1/|g\sigma|! = 1/|\sigma|! = d_\rho(\sigma)$. The equality together with the fact that S is forcing implies that $g^{-1}\pi_\bullet$ is quasirandom, which in turn implies that for every $\tau \in \mathbb{S}_4$, we have $d(\tau, g^{-1}\pi_\bullet) = 1/24$. As $\mathbb{S}_4$ is invariant under action of D_4, we have that for every $\tau \in \mathbb{S}_4$ it holds that $d(g^{(-1)}\tau, g^{(-1)}\pi_\bullet) = d(\tau, \pi_\bullet) = 1/24$. Thus due to Theorem 4, the sequence $\pi_\bullet$ is forcing as desired. □

Thus from now on, we work with sets of orbits instead of the individual sets. As such, any time a concrete set S is mentioned as dependent and or possibly forcing, reader should think of its whole orbit.

Finally, let us make a simple observation here so that we do not repeat it in proofs which follow. If S is a minimal forcing set, S does not contain the set $\mathbb{S}_k$ for any $k \in \mathbb{N}$: due to law of total probability, we have $\sum_\sigma^{\mathbb{S}_k} d(\sigma, \mu) = 1$ which implies that for any $\sigma \in \mathbb{S}_k$ it holds that $d(\sigma, \mu) = 1 - \sum_{\sigma}^{\mathbb{S}_k \setminus \{\sigma\}}$ and consequently it holds that if $G_{\mathbb{S}_k \setminus \{\sigma\}} \mu = 0$, then $G_{\mathbb{S}_k} \mu = 0$ holds; this contradicts the minimality of S. As no collection of three permutations is forcing, this immediately implies that any four-element forcing set S does not contain permutation of order one or two permutations of order two.

We are ready to show that none of the sets in $\mathcal{D}_+$ from Theorem 14 is forcing.

Theorem 29. Let S be a set of four permutations that are support of a superpermutation ν with non-vanishing constant fuzzy cover, i.e., S is an element of the set $\mathcal{D}_+$. Then S is not forcing.

Proof. Let $S \in \mathcal{D}_+$. By Theorem 14, we know that $S \in \mathbb{S}_{\leq 3}^{(4)}$, $S \in \mathbb{S}_4^{(4)}$, or S is (up to action of the group D_4) one of

$$
\begin{array}{cc}
(1234, 4321, 123, 12), & (1234, 4321, 123, 21), \\
(1324, 4231, 123, 12), & (1324, 4231, 123, 21), \\
(2143, 3412, 123, 12), & (2413, 3142, 123, 12), \\
(1234, 4321, 123, 321), & (1324, 4231, 123, 321), \\
(2143, 3412, 123, 321), & (2413, 3142, 123, 321), \\
(12345, 52341, 2143, 321), & (12345, 52341, 3412, 123), \\
(12435, 52431, 2143, 321), & (12435, 52431, 3412, 123),
\end{array}
$$

We denote the last set $\mathcal{E}$. As sets $S \in \mathbb{S}_4^{(4)}$ are not forcing by Theorem 4, we only need to consider the first and the third case.

Let $S \in \mathbb{S}_{\leq 3}^{(4)}$ with dependence witness ν. By Lemma 27, we know that S contains two permutations of size three and as $\mathbb{S}_3$ is not forcing, we know that $|S \cap \mathbb{S}_3| \leq 3$. As S cannot contain two permutations of size two or permutation of size one, profile of S is $(3, 3, 3, 2)$. Enumerating all such sets using computer, we determined that if S is dependent, it is (up to action of D_4) one of the following

$$
\begin{array}{ccc}
(123, 132, 213, 12), & (123, 132, 213, 21), & (123, 132, 321, 12), \\
(123, 132, 321, 21), & (123, 231, 312, 12), & (123, 231, 312, 21).
\end{array}
$$

For the sets above and set $\mathcal{E}$ we use the second-order method to conclude that none of the sets mentioned are in fact forcing; the certificates can be found in Appendix D. $\square$

We conclude this section by stating corollary, which allows us to assume that forcing sets have vanishing fuzzy covers.

Corollary 30. Let S be a forcing set of size four with dependency witness $\nu \in \mathbb{R}\mathbb{S}$ of order k. Then $F_\nu^{(k)} = 0$.

Proof. By Lemma 13, we know that if ν is of order k, it holds that $F_\nu^{(k)} = c\mathbb{J}$ for some real number $c \in \mathbb{R}$. Assume $c > 0$, then S belongs to set $\mathcal{D}_+$ and thus it is not forcing by Theorem 29. We conclude $F_\nu = 0$. $\square$

5.1 Homogeneous Sets

We first look at sets where all permutations are of order $k \in \mathbb{N}$. As the set of all permutations of size k for $k \leq 3$ is known not to be forcing and sets $S \in \mathbb{S}_4^{(4)}$ are known not to be forcing by Theorem 4, we focus on regime $k \geq 5$. Let $S \in \mathbb{S}_k^{(4)}$ be a forcing set for $k \geq 5$, then as $A_\nu = c\mathbb{J}$ for some $c \in \mathbb{R}$ and the number of non-zero elements of any row in A_ν is bounded from above by the size of the support of ν, we conclude that $F_\nu^{(k)} = A_\nu = 0$. The following lemma serves as a necessary condition on set $S \in \mathbb{S}_k^{(4)}$ being forcing which we use to restrict the space of possible forcing sets so that their enumeration becomes computationally feasible even for larger values of k. Before doing so, we need the following definition.

Given a permutation $\tau : \mathbb{S}_k$, it defines a map $\tau- : \mathbb{S}_k \to \mathbb{S}_k$ defined by $\tau(\sigma) = \tau\sigma$ which we extend to map $\tau- : \mathbb{S}_k^{(l)} \to \mathbb{S}_k^{(l)}$ and $\tau- : \mathbb{R}\mathbb{S}_k \to \mathbb{R}\mathbb{S}_k$. Note that if $A_\nu = 0$ for some $\nu \in \mathbb{R}\mathbb{S}_k$, then for any $\tau \in \mathbb{S}_k$ we have that $A_{\tau\nu} = 0$.

Lemma 31. Let $\nu \in \mathbb{S}_k$ be an S-supported superpermutation for some $S \in \mathbb{S}_k^{(4)}$ such that $A_\nu = 0$. Then there exists $\tau \in \mathbb{S}_k$ and ordering $s : [4] \to S$ such that the set $S' = \tau S$, ordering $s' = \tau s$ of set S', and superpermutation $\nu' = \tau\nu$ have the following properties:

(i) $A_{\nu'} = 0$,

(ii) $s'(1) = \mathrm{id}_k$, $s'(2)(1) = 1$, and

(iii) $\nu'_{s'(1)} = -\nu'_{s'(2)} = \nu'_{s'(3)} = -\nu'_{s'(4)}$.

Moreover, set S' with ordering s' and properties *(i)*, *(ii)*, and *(iii)* is determined by $s'(2)$ and $s'(3)$.

Proof. Let $\nu \in \mathbb{S}_k$ be an S-supported superpermutation with $A_\nu = 0$. Given $i \in [k]$ let $i_* : S \to [k]$ be a map sending permutation $\sigma \in S$ to $\sigma(i) \in [k]$ so that $\ker i_*$ represents the partitioning of set S into sets C_l for $l \in [k]$ with the property that for $\sigma_1, \sigma_2 \in C_l$ it holds that $\sigma_1(i) = \sigma_2(i) = l$.

Notice that as $A_\nu = 0$, we have that for every $i \in [k]$ the partition $\ker i_*$ does not contain singleton sets: assume it does so that there exists $i \in [k]$ and $\sigma \in S$ such that $\sigma(i) \neq \sigma'(i)$ for all $\sigma' \in S' \setminus \{\sigma\}$ which together with $\nu_\sigma \neq 0$ implies that $(A_\nu)_{\sigma(i),i} \neq 0$ contradicting $A_\nu = 0$. There are only two types of partitions of four elements without singletons, namely matching with two classes of size two and the trivial partition with single class containing all four elements.

Let G be a graph with vertex set S and let $\sigma_1 \sim \sigma_2$ be an edge if and only if $\sigma_1 \neq \sigma_2$ and there is i such that $\{\sigma_1, \sigma_2\} \in \ker i_*$ (note that this implies not only that $\sigma_1(i) = \sigma_2(i)$ but also that $\sigma_3(i) \neq \sigma_1(i)$ and $\sigma_4(i) \neq \sigma_1(i)$). We show that $G \cong C_4$. First, let us show that from this the claim follows.

First, an edge $\sigma_1 \sim \sigma_2$ implies that $\nu_{\sigma_1} = -\nu_{\sigma_2}$ which implies that there is ordering of the cycle $\hat{s} : [4] \to S$ such that $\nu_{\hat{s}(1)} = -\nu_{\hat{s}(2)} = \nu_{\hat{s}(3)} = -\nu_{\hat{s}(4)}$. By discussion above, it holds that $\ker 1_*$ is either a matching in which case we choose $\sigma' \in \mathbb{S}_k$ so that $\{\hat{s}(1), \sigma'\} \in \ker 1_*$, or $\ker 1_*$ is trivial in which case we choose $\sigma' = \hat{s}(2)$; in both cases we have $\hat{s}(1) \sim \sigma'$ and $\hat{s}(1)(1) = \sigma'(1)$. Now choose one of the two orderings of the cycle s so that $s(1) = \hat{s}(1)$ and $s(2) = \sigma'$ by which $s(1)(1) = s(2)(1)$.

Let $\tau = s(1)^{-1}$, let $S' = \tau S$ be the transformed set with ordering $s' = \tau s$ for which it holds that $s'(1) = 1_k$ and $s'(2)(1) = s'(1)(1) = 1$, and set $\nu' = \tau\nu$ so that $\nu'_{s'(1)} = -\nu'_{s'(2)} = \nu'_{s'(3)} = -\nu'_{s'(4)}$. Let us show that given $s'(2)$ and $s'(3)$, we can reconstruct $s'(4)$. This is immediate as by property *(iii)* and property *(i)* we have that $A_{s'(4)} = A_{s'(1)} - A_{s'(2)} + A_{s'(3)} = A_{\mathrm{id}_k} - A_{s'(2)} + A_{s'(3)}$.

Finally, we prove that $G \cong C_4$. The graph G cannot contain odd cycle, for otherwise the permutations on the odd cycle need to have zero coefficient in ν which contradicts the fact that $S = \operatorname{supp} \nu$. Let us show that for degrees we have $d(\sigma) \geq 2$ for all $\sigma \in S$; this concludes the proof as there is a single bipartite graph on four vertices with minimum degree 2, namely $K_{2,2} \cong C_4$. Assume that permutation $\sigma \in S$ is isolated: this implies that $\ker i_*$ is trivial for all $i \in [k]$ which implies that all the permumations are the same permutation which is clearly a contradiction. Assume $d(\sigma) = 1$ and let σ' be a neighbour of σ in G. Observe that this implies $\sigma = \sigma'$ which is a contradiction: let $i \in [k]$, then $\ker i_*$ is either a matching in which case $\{\sigma, \sigma'\} \in \ker i_*$ and $\sigma(i) = \sigma'(i)$ or the kernel is trivial in which case we have the same conculsion. □

Proposition 32. Let $k \leq 5$ let $S \in \mathbb{S}_k^{(4)}$. Then S is not forcing.

Proof. As $\mathbb{S}_l$ is not forcing for $l < 4$ and any set $S \in \mathbb{S}_k^{(4)}$ is not forcing by Theorem 4, let $k \geq 5$. Given two distinct permutations $\sigma_2, \sigma_3 \in \mathbb{S}_k$ of which neither is the identity permutation, let $S(\sigma_2, \sigma_3)$ be the set $\{\mathrm{id}_k, \sigma_2, \sigma_3, \sigma_4\}$ if matrix $M = A_{\mathrm{id}_k} - A_{\sigma_2} + A_{\sigma_3}$ is a permutation matrix associated with permutation $\sigma_4 \in \mathbb{S}_5$ which is different from id_k, σ_2 and σ_3, and let $S(\sigma_2, \sigma_3)$ be the empty set otherwise. Let $\mathcal{T}(\sigma_2, \sigma_3) = \{\tau S(\sigma_2, \sigma_3) \mid \tau \in \mathbb{S}_k\}$. I claim that

$$\{S \in \mathbb{S}_k^{(4)} \mid S \text{ is forcing}\} \subset \bigcup_{\sigma_2, \sigma_3}^{\mathbb{S}_k^{(2)}} \mathcal{T}(\sigma_2, \sigma_3). \tag{5.2}$$

Assume that $S \in \mathbb{S}_k^{(4)}$ is a forcing and let ν be an S-supported superpermutation with $A_\nu = 0$. Then by Lemma 31, there exists τ such that $S' = \tau S = \{\mathrm{id}_k, \sigma_2, \sigma_3, \sigma_4\}$ and $A_{\mathrm{id}_k} - A_{\sigma_2} + A_{\sigma_3} - A_{\sigma_4} = 0$. This implies that $S' = S(\sigma_2, \sigma_3)$ and consequently $S = \tau^{-1} S' \in \mathcal{T}(\sigma_2, \sigma_3)$.

Let $k = 5$. We computationally enumerate all elements of the right-hand side of (5.2) which produces 159 sets in total. For these we use the second-order method; the list of the 159 permutation sets and their certificates can be found in Appendix D. □

5.2 Sets with One Exceptional Permutation

Next, we show that sets with profile $p(k,l) = (k,k,k,l)$ for $k \leq 6$ and $l < k$ are not forcing. We start by considering regime $4 \leq k \leq 6$ in which we show that all $p(k,l)$-sets are independent and thus cannot be forcing by Proposition 9.

Lemma 33. Let $\sigma \in \mathbb{S}_l$ for $l \geq 3$, then $F_\sigma^{(l+1)}$ has at least three non-zero values in every column but the first and the last, i.e., the vectors $F_\sigma^{(l+1)} e_i$ have for $2 \leq i \leq l$ support of size at least three.

Proof. Fix $2 \leq v \leq l$. Expanding the definition of $F_\sigma^{(l+1)}$ gives

$$(F_\sigma^{(l+1)})_{u,v} = (l+1)^{-1} \sum_{j}^{[l]} \binom{u-1}{\sigma(j)-1} \binom{l+1-u}{l-\sigma(j)} \binom{v-1}{j-1} \binom{l+1-v}{l-j}.$$

It follows that $(F_\sigma^{(l+1)})_{u,v}$ is non-zero if and only if there exists $j \in [l]$ for which all binomial coefficients in the j-the summand are non-zero. After considering supports of the individual binomial coefficients, we get that $(F_\sigma^{(l+1)})_{u,v}$ is non-zero if and only if there exists $j \in [l]$ such that $v \in \{j, j+1\}$ and $u \in \{\sigma(j), \sigma(j)+1\}$ or differently, if $j = v$ and $u \in \{\sigma(v), \sigma(v)+1\}$ or $j = v-1$ and $y \in \{\sigma(v-1), \sigma(v-1)+1\}$; note that as $2 \leq v \leq l$ both cases do occur. Consequently, $(F_\sigma^{(l+1)})_{u,v}$ is non-zero if u is an element of the set $Y_v = \{\sigma(v), \sigma(v)+1, \sigma(v-1), \sigma(v-1)+1\}$. As Y_v is a subset of $[l+1]$, it remains to show that the set Y_v contains at least three elements, but this is immediate as $\sigma(v) \neq \sigma(v-1)$. □

Lemma 34. Let $\sigma \in \mathbb{S}_l$ for $l \geq 2$, then $F_\sigma^{(k)}$ for $k > l+1$ contains a column with support of size at least four.

Proof. Due to similar analysis as in Lemma 33, we see that $(F_\sigma^{(k)})_{u,v} > 0$ if and only if there exists $j \in [l]$ such that $v \in \{j, ..., j+k-l\}$ and $u \in \{\sigma(j), ..., \sigma(j)+k-l\}$. Notice that if $(F_\sigma^{(k)})_{u,v}$ is non-zero for some pair $u, v \in [k]$, then $(F_\sigma^{(k+1)})_{u,v}$ is also non-zero. Thus, it is enough to consider the case $k = l+2$. Choose $v = 2$ so that $(F_\sigma^{(l+2)})_{u,2} > 0$ if $u \in \{\sigma(j), \sigma(j)+1, \sigma(j)+2\}$ for some $j \in \{1,2\}$. This in turn implies that there are at least four distinct values of u such that $(F_\sigma^{(l+2)})_{u,2}$ is non-zero as $\sigma(1) \neq \sigma(2)$. □

Lemma 35. Let $\sigma \in \mathbb{S}_l$ for $4 \leq l \leq 5$, the submatrix spanned by all columns except the first and the last contains four distinct non-zero values, i.e., for linear map $U : \mathbb{R}^{(l-1) \to (l+1)}$ which sends e_i to e_{i+1}, the matrix of $F_\sigma^{(l+1)} U$ contains at least four distinct non-zero values.

Proof. Here, we unfortunately have no better method than to enumerate all permutations of $\mathbb{S}_4 \cup \mathbb{S}_5$ up to action of group D_4 and find the values manually. We list the values in Appendix B. □

Proposition 36. Let $k \leq 6$ and let S be a (k,k,k,l)-set of permutations for $l < k$. Then S is not forcing.

Proof. As minimal forcing set cannot contain a permutation of order one, we can assume $l \geq 2$. By Corollary 30, we assume $F_\nu^{(k)} = 0$. Assume $k \geq l+2$, then as every column of $F^{(k)}_{\nu-\nu_4\sigma_4} = A_{\nu-\nu_4\sigma_4}$ has support of size at most three and $F^{(k)}_{\nu_4\sigma_4}$ has a column whose support has size at least four by Lemma 34, their sum cannot be the zero matrix, a contradiction.

As such, assume $k = l+1$. For $k \in \{3,4\}$, there are only seven dependent sets (up to action of D_4) with profile $(4,4,4,3)$, or $(3,3,3,2)$, namely

$$\begin{array}{ccc}
(123,132,213,12) & (123,132,213,21) & (123,132,321,12) \\
(123,132,321,21) & (123,231,312,12) & (123,231,312,21) \\
 & (1234,1324,2143,123) &
\end{array}$$

which we resolve using the second-order method; their certificates can be found in Appendix D.

Let $k \in \{5,6\}$. We have $F^{(k)}_{\sigma_4} = \nu_4^{-1}(\nu_1 A_{\sigma_1} + \nu_2 A_{\sigma_2} + \nu_3 A_{\sigma_3})$. By Lemma 33, it holds that for every natural number $2 \leq i \leq k$, vector $F^{(k)}_{\sigma_4} e_i$ contains three nonzero values which implies that $\sigma_1(i)$, $\sigma_2(i)$, and $\sigma_3(i)$ are all distinct and so $F^{(k)}_{\sigma} e_i$ contains three non-zero values, namely $\nu_{\sigma_1}, \nu_{\sigma_2}$, and ν_{σ_3}. This in turn implies that the submatrix of $F^{(k)}_{\sigma}$ spanned by all columns but the first and the last contains only three named values. This contradicts the conclusion of Lemma 35 as $l = k-1$ implies $4 \leq l \leq 5$. □

5.3 Matched Profile

In this Section, we consider permutation sets with profile $m(k,l) = (k,k,l,l)$ for $l < k$. By Lemma 27, we know that any set S with profile $m(k,l)$ for $l < k-1$ cannot be forcing and thus focus on case $l = k-1$.

Lemma 37. Let $k \geq 5$ and let S be a forcing set with profile $(k,k,k-1,k-1)$, together with a dependence witness ν. Then k is odd, $\nu_2 = -\nu_1$, and $\nu_4 = -\nu_3$.

Proof. Let $\nu = \xi + \eta$ with $\xi \in \mathbb{RS}_k$ and $\eta \in \mathbb{RS}_{k-1}$. As $P_\nu = 0$, and consequently $P_\xi = -P_\eta$, it holds that $T_\xi e_{k-1} = -T^{(k-2)}_\eta e_{k-1} = 0$ and T_ξ is singular with kernel element e_{k-1}. By Lemma 26, it holds that coefficients in ξ are related by $\nu_1 = (-1)^k \nu_2$ and $\Psi_k A_\xi = (-1)^k A_\xi$. As such, matrix A_ξ has rank at most $\lfloor k/2 \rfloor \leq k-3$ for $k \geq 5$: if k is even, the claim is immediate by horizontal symmetry, and for k odd, the claim follows from horizontal antisymmetry and

$$A_\xi e_{\lceil k/2 \rceil} = \Psi_k A_\xi e_{\lceil k/2 \rceil} = -A_\xi e_{\lceil k/2 \rceil} = 0.$$

As such T_ξ and $T^{(k-2)}_\eta$ both have rank of at most $k-3$ and thus T_η is singular. By Lemma 26, we have $|\nu_3| = |\nu_4|$, i.e., $\nu_4 = (-1)^x \nu_3$ for some $x \in \mathbb{N}$.

By Corollary 30, it follows that $F_\nu^{(k)} = 0$, from which together with the discussion above, we conclude

$$\nu_1\big(A_{\sigma_1} + (-1)^k A_{\sigma_2}\big) = -\nu_3(F_{\sigma_3}^{(k)} + (-1)^x F_{\sigma_4}^{(k)}). \tag{5.3}$$

As j is an eigenvector of A_σ with eigenvalue one for any $\sigma \in \mathbb{S}_k$, and recall that j is an eigenvector of $F_\sigma^{(k)}$ with eigenvalue $(k-1)!/(k-2)! = k-1$ for $\sigma \in \mathbb{S}_{k-1}$, by applying both sides of 5.3 to j, we obtain the equality

$$\nu_1\big(1 + (-1)^k\big)j = -\nu_3(k-1)\big(1 + (-1)^x\big)j. \tag{5.4}$$

If k is even, then the equality implies $2\nu_1 = -\nu_3(k-1)(1+(-1)^x)$ which implies that x is even as both ν_1 and ν_3 are non-zero. Thus we conclude $\nu_4 = \nu_3$. We show that this conclusion leads to a contradiction. Assume it to be true, then

$$\nu_1(A_{\sigma_1} + A_{\sigma_2}) = -\nu_3(F_{\sigma_3}^{(k)} + F_{\sigma_4}^{(k)})$$

holds but the left-hand side is a matrix where each column has support of size at most two, whereas the right-hand side has a column with support of size at least three by Lemma 33 and the fact $(F_\sigma^{(k)})_{i,j} \geq 0$ for all $i, j \in [k]$, i.e., the size of the support of a column in $F_{\sigma_3}^{(k)} + F_{\sigma_4}^{(k)}$ is at least the size of the support of the column in the summands.

Let k be odd. Then by 5.4 it holds that x is odd as ν_1, ν_3, and $k-1$ are non-zero. As such $\nu_4 = -\nu_3$ which concludes the proof. $\square$

The above gives us a computationally checkable necessery condition for $(k, k, k-1, k-1)$-set S with $k \geq 5$ to be forcing. Namely, if ν is the dependency witness of S, it has to hold that matrices $T_{\sigma_1-\sigma_2}^{(k-2)}$ and $T_{\sigma_3-\sigma_4}^{(k-2)}$ have to be linearly dependent as $\nu_1 T_{\sigma_1-\sigma_2}^{(k-2)} + \nu_3 T_{\sigma_3-\sigma_4}^{(k-2)} = T^{(k-2)} P_\nu = 0$. Before moving to the main proposition of this section, we note that due to Example 15, this is the best result possible using only dependence condition.

Proposition 38. Let $k \leq 6$ and let S be a (k, k, l, l)-set of permutations for $l < k$. Then S is not forcing.

Proof. As S does not contain sets $\mathbb{S}_1$ and $\mathbb{S}_2$, we can assume $l \geq 3$ and by Lemma 27, we can assume $k = l + 1$. For $k \geq 5$, it holds that k has to be odd by Lemma 37 and so it suffices to solve cases $k \in \{4, 5\}$. Let S be a forcing set with profile $(k, k, k-1, k-1)$ and dependency witness ν whose fuzzy cover is the zero matrix by Corollary 30. Then it is (up to action of D_4) one of the following:

$$\begin{array}{cc}
(1234, 4321, 123, 321) & (1324, 4231, 123, 321) \\
(2143, 3412, 123, 321) & (2413, 3142, 123, 321) \\
(12345, 14325, 1234, 1324) & (12345, 14325, 4231, 4321) \\
(12543, 14523, 1234, 1324) & (12543, 14523, 4231, 4321)
\end{array}$$
$$(25314, 41352, 2413, 3142).$$

For $k = 4$, we used computer to enumerated all sets with the appropriate profile (up to action of D_4) and checked dependence of the system T_S as described in Chapter 3 to exclude all but the four mentioned sets. For $k = 5$, we used computer to enumerate all sets with appropriate profile (up to action of D_4) and checked whether the matrices $A = T^{(k-2)}_{\sigma_1-\sigma_2}$ and $B = T^{(k-2)}_{\sigma_3-\sigma_4}$ are linearly dependent. For the sets above, we use the second-order method and the certificates can be found in Appendix D. $\square$

5.4 And the Rest

Unfortunately, for the most general case, I found no better approach than to computationally enumerate all of the sets.

Proposition 39. Let $k_1 \leq 5$ and let S be a (k_1, k_1, k_2, k_3)-set of permutations for $k_3 < k_2 < k_1$. Then S is not forcing.

Proof. As $\mathbb{S}_1$ cannot be a subset of a minimal forcing set, we assume $k_3 > 1$ and as a consequence $k_1 \geq 4$. Due to Lemma 27, we can assume $k_2 = k_1 - 1$. We used computer to enumerate all sets with profiles $(5, 5, 4, 3)$, $(5, 5, 4, 2)$, and $(4, 4, 3, 2)$ and checked dependence of system the T_S as described in Chapter 3 to conclude that if set S with one of the mentioned profiles is dependent, it is (up to action of D_4) one of

$$\begin{array}{cc}
(12345, 52341, 2143, 321), & (12345, 52341, 3412, 123), \\
(12435, 52431, 2143, 321), & (12435, 52431, 3412, 123), \\
(1234, 4321, 123, 12), & (1234, 4321, 123, 21), \\
(1324, 4231, 123, 12), & (1324, 4231, 123, 21), \\
(2143, 3412, 123, 12), & (2143, 3412, 123, 21), \\
(2413, 3142, 123, 12), & (2413, 3142, 123, 21).
\end{array}$$

We use the second-order method to conclude that none of the sets mentioned are in fact forcing; the certificates can be found in Appendix D. $\square$

Proof of Theorem 2. Let S be a forcing set with profile (k_1, k_2, k_3, k_4) with $k_4 \leq k_3 \leq k_2 \leq k_1 \leq 5$. By Lemma 27 we have that $k_2 = k_1$ which together with Proposition 32, Proposition 36, Proposition 38, and Proposition 39 shows the statement. $\square$

Chapter 6

Conclusion

The prime objective of this thesis was to improve our understanding of the boundary between systems of pattern densities that do and those that do not force quasirandomness in convergent permutation sequences. In [9], authors show that any forcing superpermutation with positive coefficients needs to have a support of size at least six. In the same work, the authors conjecture that this is true even without constraints on the coefficients, and authors in [17] strengthen the conjecture to say that any forcing set needs to contain at least six permutations; note that a set attaining the bound is known and can be found in [9].

Conjecture 40. There exists no quasirandom-forcing set of size at most five.

In this direction, we contribute in the form of Theorem 2, by showing that sets of four small permutations cannot force quasirandomness. In this work, our proofs heavily rely on computer assistance, both in the form of enumerating sets with linearly dependent gradient polynomials and computing the second-order witnesses as described in Chapter 4. It would be of great value to replace our proofs with "human" proofs; those not relying on any computer-obtained witnesses.

The first goal of future work would be to classify sets with dependent gradients, the same way Theorem 14 classifies superpermutations with non-vanishing constant covers. This goal extends the question posed by Crudele et al. of classifying superpermutations whose constant fuzzy covers are vanishing. In this direction, one may not hope to get classification in terms of a finite set due to an infinite class of dependent sets constructed in Example 15; the same was observed for the Crudele et al. question in their work. To achieve this, a better understanding of fuzzy covers is of interest and more lemmas similar to those found in Chapter 5 are needed, e.g., controlling the shape of the coefficients, structure of the support of fuzzy covers and sets of their distinct non-zero values.

Throughout the text, we study Hessians of discrepancy functions for perturbations of various orders, with the largest perturbation being of order eight. At this value, it becomes computationally difficult to produce even the symbolic expression of the discrepancy polynomial and it is probably not practical to go much further beyond this value. As such, a human approach is needed, with the aforementioned infinite class of Example 15 being a good first case to tackle. A plausible course is to study Hessian sequences instead of the individual Hessians, the same way the work of [19] studies the Jacobian sequence instead of the individual Jacobians. To this end we have only very limited preliminary results.

Bibliography

[1] M. Bucić, E. Long, A. Shapira, and B. Sudakov. "Tournament quasirandomness from local counting". In: *Combinatorica* 41.2 (2021), pp. 175–208.

[2] T. F. N. Chan, D. Král', J. A. Noel, Y. Pehova, M. Sharifzadeh, and J. Volec. "Characterization of quasirandom permutations by a pattern sum". In: *Random Structures & Algorithms* 57.4 (2020), pp. 920–939.

[3] F. Chung and R. L. Graham. "Quasi-random tournaments". In: *Journal of Graph Theory* 15.2 (1991), pp. 173–198.

[4] F. Chung, R. L. Graham, and R. M. Wilson. "Quasi-random graphs". In: *Combinatorica* 9.4 (1989), pp. 345–362.

[5] D. Conlon, J. Fox, and B. Sudakov. "An approximate version of Sidorenko's conjecture". In: *Geometric and Functional Analysis* 20 (2010), pp. 1354–1366.

[6] J. W. Cooper, D. Král', A. Lamaison, and S. Mohr. "Quasirandom latin squares". In: *Random Structures & Algorithms* 61.2 (2022), pp. 298–308.

[7] J. Cooper and A. Petrarca. "Symmetric and asymptotically symmetric permutations". In: *arXiv preprint arXiv:0801.4181* (2008).

[8] L. N. Coregliano and A. A. Razborov. "On the density of transitive tournaments". In: *Journal of Graph Theory* 85.1 (2017), pp. 12–21.

[9] G. Crudele, P. J. Dukes, and J. A. Noel. "Six permutation patterns force quasirandomness". In: *Discrete Analysis* (2024).

[10] R. Glebov, A. Grzesik, T. Klimošová, and D. Král'. "Finitely forcible graphons and permutons". In: *Journal of Combinatorial Theory, Series B* 110 (2015), pp. 112–135.

[11] W. T. Gowers. "Quasirandom groups". In: *Combinatorics, Probability and Computing* 17.3 (2008), pp. 363–387.

[12] R. Hancock, A. Kabela, T. Martins, R. Parente, F. Skerman, J. Volec, et al. "No additional tournaments are quasirandom-forcing". In: *European Journal of Combinatorics* 108 (2023).

[13] R. Hancock, A. Kadela, D. Kráľ, T. Martins, R. Parente, F. Skerman, and J. Volec. "No additional tournaments are quasirandom-forcing". In: *European Journal of Combinatorics* 108 (2023).

[14] M. W. Hirsch. *Differential topology*. Vol. 33. Springer Science & Business Media, 2012.

[15] C. Hoppen, Y. Kohayakawa, C. G. Moreira, B. Ráth, and R. M. Sampaio. "Limits of permutation sequences". In: *Journal of Combinatorial Theory, Series B* 103.1 (2013), pp. 93–113.

[16] C. Hoppen, Y. Kohayakawa, C. G. T. d. A. Moreira, and R. M. Sampaio. "Limits of permutation sequences through permutation regularity". In: *arXiv preprint arXiv:1106.1663* (2011).

[17] D. Král', J.-b. Lee, and J. A. Noel. "Forcing quasirandomness with 4-point permutations". In: *arXiv preprint arXiv:2407.06869* (2024).

[18] D. Král' and O. Pikhurko. "Quasirandom permutations are characterized by 4-point densities". In: *Geometric and Functional Analysis* 23.2 (2013), pp. 570–579.

[19] M. Kurečka. "Lower bound on the size of a quasirandom forcing set of permutations". In: *Combinatorics, Probability and Computing* 31.2 (2022), pp. 304–319.

[20] L. Lovász. *Large networks and graph limits*. Vol. 60. 2012.

[21] A. A. Razborov. "Flag algebras". In: *The Journal of Symbolic Logic* 72.4 (2007), pp. 1239–1282.

[22] V. Rödl. "On universality of graphs with uniformly distributed edges". In: *Discrete Mathematics* 59.1-2 (1986), pp. 125–134.

[23] The Sage Developers. *SageMath, the Sage Mathematics Software System (Version 10.6)*. `https://www.sagemath.org`. 2025.

[24] T. Shifrin. *Multivariable mathematics: linear algebra, multivariable calculus, and manifolds*. John Wiley & Sons, 2004.

[25] J. Skokan and L. Thoma. "Bipartite subgraphs and quasi-randomness". In: *Graphs and Combinatorics* 20 (2004), pp. 255–262.

[26] A. Thomason. "Pseudo-random graphs". In: *North-Holland Mathematics Studies*. Vol. 144. 1987, pp. 307–331.

Appendix A

Elements of multilinear calculus

Let $F : \mathbb{R}^n \to \mathbb{R}^m$ be a smooth function, recall that we define $D_x F$ to be the Jacobian matrix of F at point $x \in \mathbb{R}^n$ and we define $H_x F$ to be the Hessian bilinear form of F at point $x \in \mathbb{R}^n$ which we define as the bilinear form associated to $D_x(y \to D_y F)$. The main goal of this section is to derive an equation for the Hessian of compositions of two smooth maps. I believe the equation presented to be standard but I failed to find a suitable reference.

Throughout this section, for matrix $M : \mathbb{R}^{n\times m}$ and $i \in [m]$ and $j \in [n]$, we write $M^i{}_j$ for $M_{i,j}$, e.g., we have $D_x F^i{}_k = (\partial F_i / \partial x_k)(x)$. We use $\cdot$ to denote any other multiplication than direct function composition or direct matrix multiplication. That is for $F : \mathbb{R}^n \to \mathbb{R}^{m\to o}$ and $G : \mathbb{R}^n \to \mathbb{R}^{o\to p}$ we use $G \cdot F$ to denote the function which sends $x \in \mathbb{R}^n$ to $G(x)F(x) \in \mathbb{R}^{m\to p}$.

Given a differentiable map $F : \mathbb{R}^n \to \mathbb{R}^{m\to o}$ and natural numbers $k \in [n]$, $i \in [o]$, and $j \in [m]$, we use $(D_x F)^{ki}{}_j$ to denote $(\partial F^i{}_j / \partial x_k)(x)$. Let $F : \mathbb{R}^n \to \mathbb{R}^{m\to o}$ and $G : \mathbb{R}^n \to \mathbb{R}^{o\to p}$ be differentiable maps, then the following holds:

$$\frac{\partial (G\cdot F)^i{}_j}{\partial x_k} = \sum_l^{[o]} \frac{\partial (G^i{}_l \cdot F^l{}_j)}{\partial x_k} = \sum_l^{[o]} \frac{\partial G^i{}_l}{\partial x_k} \cdot F^l{}_j + G^i{}_l \cdot \frac{\partial F^l{}_j}{\partial x_k}$$

which for $D_x(G\cdot F) \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^{m\to p})$ and $u \in \mathbb{R}^n$ in turn implies

$$D_x(G\cdot F)(u) = (D_x G(u))F(x) + G(x)(D_x F(u)). \tag{A.1}$$

which can be seen as a version of a Leibniz rule. Using A.1, we obtain the main result of this section. Given smooth functions $F : \mathbb{R}^n \to \mathbb{R}^m$ and $G : \mathbb{R}^m \to \mathbb{R}^o$

as above and $u, v \in \mathbb{R}^n$, we have

$$\begin{aligned}
H_x(GF)(u,v) &= D_x(y \mapsto D_y(GF))(u,v) \\
&= D_x(y \mapsto (D_{F(y)}G)(D_yF))(u,v) \\
&= D_x\big((y \mapsto D_{F(y)}G) \cdot (y \mapsto D_yF)\big)(u,v) \\
&= \big[D_x(y \mapsto D_{F(y)}G)(u)\big](D_xF)v + (D_{F(x)}G)\big[D_x(y \mapsto D_yF)(u)\big]v \\
&= \Big[D_x\big((y \mapsto D_yG)F\big)(u)\Big](D_xF)v + D_{F(x)}G\ H_xF(u,v) \\
&= \Big[\big(D_{F(x)}(y \mapsto D_yG)\big)(D_xFu)\Big](D_xF)v + (D_{F(x)}G)H_xF(u,v) \\
&= H_{F(x)}G(D_xFu, D_xFv) + (D_{F(x)}G)H_xF(u,v),
\end{aligned}$$

where the second equality is due to use of chain rule, fourth is due to (A.1), sixth is due to chain rule and the remaining equations are definitional.

Appendix B

Certificates for Four-Value Lemma

This Appendix accompanies the proof of Lemma 35. Given a permutation $\sigma \in \mathbb{S}_k$, we use $(u, v) \mapsto x$ to denote the fact that $(F_\sigma^{(k+1)})_{u,v} = x$. What follows is enumeration of all $\sigma \in \mathbb{S}_4 \cup \mathbb{S}_5$ (up to action of the group D_4) together with the four non-zero values of $F_\sigma^{(k+1)}$ as described in the statement of the lemma:

$$
\begin{array}{rllll}
\sigma = 1234: & (1,2) \mapsto 4/5, & (2,2) \mapsto 2, & (3,2) \mapsto 6/5, & (3,3) \mapsto 8/5; \\
\sigma = 1243: & (1,2) \mapsto 4/5, & (2,2) \mapsto 2, & (3,2) \mapsto 6/5, & (4,3) \mapsto 2/5; \\
\sigma = 1324: & (1,2) \mapsto 4/5, & (2,2) \mapsto 1/5, & (3,2) \mapsto 6/5, & (4,2) \mapsto 9/5; \\
\sigma = 1342: & (1,2) \mapsto 4/5, & (2,2) \mapsto 1/5, & (3,2) \mapsto 6/5, & (4,2) \mapsto 9/5; \\
\sigma = 1432: & (1,2) \mapsto 4/5, & (2,2) \mapsto 1/5, & (4,2) \mapsto 3/5, & (5,2) \mapsto 12/5; \\
\sigma = 2143: & (1,2) \mapsto 12/5, & (2,2) \mapsto 6/5, & (3,2) \mapsto 2/5, & (1,3) \mapsto 8/5; \\
\sigma = 2413: & (2,2) \mapsto 3/5, & (3,2) \mapsto 2/5, & (5,2) \mapsto 12/5, & (1,3) \mapsto 8/5; \\
\sigma = 12345: & (1,2) \mapsto 5/6, & (2,2) \mapsto 17/6, & (3,2) \mapsto 4/3, & (3,3) \mapsto 13/6; \\
\sigma = 12354: & (1,2) \mapsto 5/6, & (2,2) \mapsto 17/6, & (3,2) \mapsto 4/3, & (3,3) \mapsto 13/6; \\
\sigma = 12435: & (1,2) \mapsto 5/6, & (2,2) \mapsto 17/6, & (3,2) \mapsto 4/3, & (3,3) \mapsto 2/3; \\
\sigma = 12453: & (1,2) \mapsto 5/6, & (2,2) \mapsto 17/6, & (3,2) \mapsto 4/3, & (3,3) \mapsto 2/3; \\
\sigma = 12543: & (1,2) \mapsto 5/6, & (2,2) \mapsto 17/6, & (3,2) \mapsto 4/3, & (3,3) \mapsto 2/3; \\
\sigma = 13254: & (1,2) \mapsto 5/6, & (2,2) \mapsto 1/6, & (3,2) \mapsto 2, & (4,3) \mapsto 1; \\
\sigma = 13425: & (1,2) \mapsto 5/6, & (2,2) \mapsto 1/6, & (3,2) \mapsto 2, & (3,3) \mapsto 1; \\
\sigma = 13452: & (1,2) \mapsto 5/6, & (2,2) \mapsto 1/6, & (3,2) \mapsto 2, & (3,3) \mapsto 1; \\
\sigma = 13524: & (1,2) \mapsto 5/6, & (2,2) \mapsto 1/6, & (3,2) \mapsto 2, & (3,3) \mapsto 1; \\
\sigma = 13542: & (1,2) \mapsto 5/6, & (2,2) \mapsto 1/6, & (3,2) \mapsto 2, & (3,3) \mapsto 1; \\
\sigma = 14325: & (1,2) \mapsto 5/6, & (2,2) \mapsto 1/6, & (4,2) \mapsto 4/3, & (5,2) \mapsto 8/3; \\
\sigma = 14352: & (1,2) \mapsto 5/6, & (2,2) \mapsto 1/6, & (4,2) \mapsto 4/3, & (5,2) \mapsto 8/3;
\end{array}
$$

$\sigma = 14523$: $(1,2) \mapsto 5/6$, $(2,2) \mapsto 1/6$, $(4,2) \mapsto 4/3$, $(5,2) \mapsto 8/3$;
$\sigma = 14532$: $(1,2) \mapsto 5/6$, $(2,2) \mapsto 1/6$, $(4,2) \mapsto 4/3$, $(5,2) \mapsto 8/3$;
$\sigma = 15342$: $(1,2) \mapsto 5/6$, $(2,2) \mapsto 1/6$, $(5,2) \mapsto 2/3$, $(6,2) \mapsto 10/3$;
$\sigma = 15432$: $(1,2) \mapsto 5/6$, $(2,2) \mapsto 1/6$, $(5,2) \mapsto 2/3$, $(6,2) \mapsto 10/3$;
$\sigma = 21354$: $(1,2) \mapsto 10/3$, $(2,2) \mapsto 4/3$, $(3,2) \mapsto 1/3$, $(1,3) \mapsto 5/3$;
$\sigma = 21453$: $(1,2) \mapsto 10/3$, $(2,2) \mapsto 4/3$, $(3,2) \mapsto 1/3$, $(1,3) \mapsto 5/3$;
$\sigma = 21543$: $(1,2) \mapsto 10/3$, $(2,2) \mapsto 4/3$, $(3,2) \mapsto 1/3$, $(1,3) \mapsto 5/3$;
$\sigma = 23514$: $(2,2) \mapsto 2/3$, $(3,2) \mapsto 7/3$, $(4,2) \mapsto 2$, $(3,3) \mapsto 1$;
$\sigma = 24153$: $(2,2) \mapsto 2/3$, $(3,2) \mapsto 1/3$, $(4,2) \mapsto 4/3$, $(5,2) \mapsto 8/3$;
$\sigma = 24513$: $(2,2) \mapsto 2/3$, $(3,2) \mapsto 1/3$, $(4,2) \mapsto 4/3$, $(5,2) \mapsto 8/3$;
$\sigma = 25314$: $(2,2) \mapsto 2/3$, $(3,2) \mapsto 1/3$, $(6,2) \mapsto 10/3$, $(3,3) \mapsto 3/2$;

Appendix C

Implementation of the First-order Method

Throughout the main body of this thesis, we often point to a computational enumeration of sets $S \subset \mathbb{S}$ and computer-assisted determination of the rank of the system T_S as described in Chapter 3. This was achieved using mathematics software system SageMath[23] and the source code in the form of a Jupyter Notebook can be found as an attachment `first-order.ipynb`. For installation of the software, we point reader to Installation Guide[1], and for guide to launching Jupyter Notebook with SageMath, we point reader to Lauching Guide[2]. The code was tested on version 10.6.

[1] `https://doc.sagemath.org/html/en/installation/index.html`
[2] `https://doc.sagemath.org/html/en/installation/launching.html`

Appendix D

Second-order Certificates

There are 187 needed certificates. As such, we include them only in computer readable form. This is done in the form of an attachment `certificates.zip` that can be found attached to the thesis. The archive contains computer readable files

> `dependent.txt`: file contains enumeration of all dependent sets S whose profile is in the set $\{(3,3,3,2)$, $(4,4,3,2)$, $(4,4,3,3)$, $(4,4,4,3)$, $(5,5,4,2)$, $(5,5,4,3)$, $(5,5,4,4)$, $(5,5,5,5)\}$;
>
> $k_1k_2k_3k_4$`.txt`: file contains enumeration of all dependent sets S whose profile is (k_1, k_2, k_3, k_4);
>
> `crudeleetal.txt`: file contains enumeration of all sets S from Theorem 29 whose resolution is needed to prove the claim.

The list contained in `dependent.txt` is union of the lists in the other files; other files are provided only for convenience. For each file in the list above, there is an accompanying certification file whose suffix is `_cert.txt`, i.e., for `dependent.txt` there is a file `dependent_cert.txt`. Each line of the certification file contains the following list of values delimited by semicolon:

1. the permutation set S for which this line is a second-order certificate,
2. natural number n which determines the function $G_{n,S}$ whose Hessian is being studied,
3. a diagonal gap $\gamma_+ \in \mathbb{Q} \cup \{+\infty\}$ for matrix $M_+ = V_+^T H V_+$ (see below),
4. a diagonal gap $\gamma_- \in \mathbb{Q} \cup \{+\infty\}$ for matrix $M_- = V_-^T H V_-$ (see below),
5. the Jacobian $D_0 G_{n,S} \in \mathbb{Q}^{(n-1)^2 \to 4}$ written as a list of four gradient vectors $(D_0 g_{n,\sigma})^S_\sigma$,
6. a vector $w \in \mathbb{Q}^4$ in kernel of $(D_0 G_{n,S})^T$,

7. a Hessian matrix $H = H_0 D_{n,S} \in \mathbb{Q}^{(n-1)^2 \times (n-1)^2}$,

8. a positive 4-witness $V_+ \in \mathbb{Q}^{4 \to (n-1)^2}$ written as a list of vectors $(v_i)_i^{[4]}$ to which V_+ is associated, where v_i is an element of $\mathbb{Q}^{(n-1)^2}$ for every $i \in [4]$,

9. a matrix $M_+ = V_+^T H V_+ \in \mathbb{Q}^{4 \times 4}$ used for verifying the fact that V_+ is a positive witness written as a list of rows,

10. a negative 4-witness $V_- \in \mathbb{Q}^{4 \to (n-1)^2}$ written as a list of vectors $(v_i)_i^{[4]}$ to which V_- is associated, where v_i is an element of $\mathbb{Q}^{(n-1)^2}$ for every $i \in [4]$,

11. a matrix $M_- = V_-^T H V_- \in \mathbb{Q}^{4 \times 4}$ used for verifying the fact that V_- is a negative witness written as a list of rows.

Symbolic variables $(x_{i,j})_{i,j}^{[n-1]^2}$ are always (for computation of gradients and Hessian matrices) ordered lexicographically, e.g., for $n = 3$ we have $x_{1,1} < x_{1,2} < x_{2,1} < x_{2,2}$. All the objects are computed in field of rational numbers $\mathbb{Q}$, meaning we can run the computations exactly. The theory of the certificates can be found in Chapter 4 and example application can be found in Section 4.2.