

MOZHGAN POURMORADNASSERI

Some Problems Related to
Extensions of Polytopes



DISSERTATIONES MATHEMATICAE UNIVERSITATIS TARTUENSIS

112

MOZHGAN POURMORADNASSERI

Some Problems Related to
Extensions of Polytopes



UNIVERSITY OF TARTU
Press

Institute of Computer Science, Faculty of Science and Technology, University of Tartu, Estonia.

Dissertation has been accepted for the commencement of the degree of Doctor of Philosophy (PhD) in informatics on April 20, 2017 by the Council of the Institute of Computer Science, University of Tartu.

Supervisor

Dr. Dirk Oliver Theis
University of Tartu
Tartu, Estonia

Opponents

Dr. Ali Taherkhani
Institute for Advanced Studies in Basic Sciences
Zanjan, Iran

Dr. Kanstantsin Pashkovich
University of Waterloo
Waterloo, Canada

The public defense will take place on June 2nd, 2017 at 16:15 in Liivi 2-405.

The publication of this dissertation was financed by the Institute of Computer Science, University of Tartu.

ISSN 1024-4212
ISBN 978-9949-77-421-0 (print)
ISBN 978-9949-77-422-7 (pdf)

Copyright: Mozghan Pourmoradnasseri, 2017

University of Tartu Press
www.tyk.ee

*To my parents
For their endless love, support and encouragement*

CONTENTS

List of original publications	8
1. Introduction	9
2. Polytopes and Extensions	13
2.1. Preliminaries	13
2.1.1. Linear programming and sizes of linear programming formulations	13
2.1.2. Polytopes and their facets	14
2.1.3. Examples of polytopes	16
2.2. Extensions of polytopes	18
2.2.1. Extension	18
2.2.2. Slack matrix	19
2.2.3. Non-negative factorization	19
2.2.4. Rectangle covering and non-negative rank	20
2.2.5. Lattice embedding	20
2.2.6. Extended formulation of some polytopes	21
2.3. Graph of a polytope	23
3. On the Graph of the Pedigree Polytope	26
3.1. Motivation and previous works	26
3.2. Our result	27
4. From Communication Complexity to Extended Formulation	29
4.1. Basic Model and definitions	29
4.1.1. Deterministic communication complexity	29
4.1.2. Nondeterministic communication complexity	30
4.2. From nondeterministic communication complexity to extended formulation	31
4.2.1. Lower bounds on extension complexity of TSP polytopes	32
4.3. Connection to rectangle graph	33
5. Nondeterministic Communication Complexity of Random Boolean Functions	35
5.1. Motivation and previous works	35
5.2. Our results	36
6. The (Minimum) Rank of Typical Fooling set Matrices	41
6.1. Motivation and previous works	41
6.2. Our results	42
7. Conclusion	45

Bibliography	47
Appendix A. Nondeterministic Communication Complexity of Random Boolean Functions	55
Appendix B. On the Graph of the Pedigree Polytope	87
Acknowledgement	110
Summary in Estonian	111
Publications	113
Nondeterministic Communication Complexity of Random Boolean Functions (Extended Abstract)	115
The Graph of the Pedigree Polytope is Asymptotically Almost Complete (Extended Abstract)	131
The (minimum) rank of typical fooling-set matrices	147
Curriculum Vitae	162
Elulookirjeldus (Curriculum Vitae in Estonian)	163

LIST OF ORIGINAL PUBLICATIONS

Publications included in the thesis

1. Pourmoradnasseri, M. and Theis, D.O. “Nondeterministic Communication Complexity of Random Boolean Functions (Extended Abstract)”. In *Proceedings of Theory and Applications of Models of Computation, TAMC 2016*.
 - Reprinted in *Publications* part of this thesis.
2. Pourmoradnasseri, M. and Theis, D.O. “Nondeterministic Communication Complexity of Random Boolean Functions”, preprint (arXiv:1611.08400).
 - Reprinted as Appendix A of this thesis.
3. Makkeh, A.; Pourmoradnasseri, M. and Theis, D.O. “The Graph of the Pedigree Polytope is Asymptotically Almost Complete (Extended Abstract)”. In *Proceedings of The International Conference on Algorithms and Discrete Applied Mathematics, CALDAM 2017*.
 - Reprinted in *Publications* part of this thesis.
4. Makkeh, A.; Pourmoradnasseri, M. and Theis, D.O. “On the Graph of the Pedigree Polytope” (preprint, arXiv:1611.08431).
 - Reprinted as Appendix B of this thesis.
5. Pourmoradnasseri, M. and Theis, D.O. “The (Minimum) Rank of Typical Fooling Set Matrices”. In *Proceedings of International Computer Science Symposium in Russia, CSR 2017*.
 - Reprinted in *Publications* part of this thesis.

1. INTRODUCTION

Why polytopes?

The goal of linear programming is optimizing a linear function, known as objective function, over a subset of Euclidean space defined by a system of linear equations and inequalities, usually represented by $A\mathbf{x} \leq \mathbf{b}$. Feasible solutions to the problem are the ones that satisfy all the constraints. The set of all feasible solutions can be represented in the Euclidean space with the dimension equal to the number of variables and it is most likely in a high dimension. Each inequality constraint is a half-space and the intersection of these half-spaces gives us a geometric object. Imagine a two-dimensional polygon surrounded by a set of lines (see figure 1). It is possible to end up with an unbounded or empty feasible region, but this is not of our interest here.

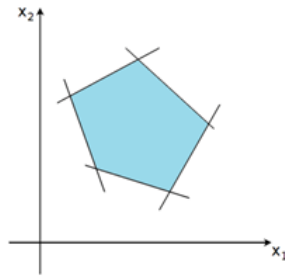


Figure 1. Inequality constraints making a polygon in dimension 2 [34].

There are already very good techniques for solving a linear program problem such as the simplex, the ellipsoid method and the interior point method. To efficiently apply these methods, it might be of advantageous to have a description of the feasible set in the form of $\{\mathbf{x} \in \mathbb{R}^n \mid A\mathbf{x} \leq \mathbf{b}\}$. If this happens and the number of inequalities which are defining the region is not that high, then everything is fine. But most of the times this is not the case!

When applying the linear programming methods, two difficulties may arise. In some cases, the linear description for the problem is known, but the number of constraints are exponential in terms of variables, which gives an exponential running time ¹. In some other cases, particularly in combinatorial optimization the feasible set, aka polytope (higher dimensional polygon), is given as a convex hull of points and the main task is finding a system of linear inequalities which defines the polytope.

Imagine that the vertices of the polygon in figure 1 were given as a set \mathcal{T} . The convex hull of the vertices, the blue area, is denoted by $\text{conv}(\mathcal{T})$. Finding the inequalities corresponding to the lines or planes which are defining the area is not

¹This can sometimes be avoided through the use of Separation Algorithms. This thesis does not discuss the pros and cons of extension vs separation, and their connections.

difficult in dimension 2 or 3, but it is usually challenging in higher dimensions [68].

Let us have a closer look at one of the best-known problems in combinatorial optimization, the *Traveling Salesman Problem*, or TSP for short. A salesman wants to visit every one of n cities and return to the first point. Given the cost of the travel between all pairs of cities, he wants to find the cheapest *tour* for his travel.

There are a total number of $\binom{n}{2}$ roads between each pair of cities and each tour can be specified as a Boolean vector of size $\binom{n}{2}$. An entry of the vector is 1 if the corresponding road is a part of the tour and 0 otherwise. The incidence vector of each tour can be regarded as a point in the space $\mathbb{R}^{\binom{n}{2}}$ and the convex hull of all the points gives a geometric object known as TSP polytope.

Let \mathcal{T} be the set of all the tours. If c is the cost vector, showing the cost of travel between cities, then the linear program formulation the TSP is

$$\text{minimize } c\mathbf{x} \text{ subject to } \mathbf{x} \in \mathcal{T}.$$

There are $(n-1)!/2$ tours for n cities. The set \mathcal{T} is so huge that optimizing over it looks impossible. For $n = 49$, there are already

$$|\mathcal{T}| = 6.20695779626803633543114452368668751926074317733888 \times 10^{60}$$

tours! Dontzig, Fulkerson, and Johnson in their breakthrough paper [13] attacked the problem by linear programming and illustrated the efficiency of their method for solving TSP for $n = 49$ cities — an enormous task for that time.

This is the intuition of their method, which is actually the basis of the current developments on solving TSP [2]. The first step is to replace \mathcal{T} with the TSP polytope, because linear programming always returns one of the points which defines the convex hull (“vertices”). Since the TSP polytope itself is too complicated² to be represented by a linear discription, Dontzig, Fulkerson, and Johnson had the idea to use a relaxation polytope.

An overview of their method is as follows: First, a suitable system of linear inequalities $A\mathbf{x} \leq \mathbf{b}$ is found, such that it is satisfied by all the $\mathbf{x} \in \text{conv}(\mathcal{T})$, and also by some other $\mathbf{x} \notin \text{conv}(\mathcal{T})$. It gives another polytope containing the TSP polytope (see figure 2). Then the algorithm detects an optimum solution \mathbf{x}^* there. However, \mathbf{x}^* is most likely not one of the points in \mathcal{T} , since the the space defined by the linear description is looser than TSP polytope. Keep in mind that the optimal solution of a linear program can be always found in a vertex of the underlying polytope. Therefore $c\mathbf{x}^*$ gives the first lower bound to the actual problem, and it is one of the vertices of the “outer” polytope.

The next step is separating \mathbf{x}^* from the space of feasible solutions. This is indeed the most challenging part of the method. A new linear inequality (hyper-plane) has to be found which is satisfied by all $\mathbf{x} \in \mathcal{T}$ but not by \mathbf{x}^* . By adding

² There are 15,379 types of inequalities in linear description of TSP for $n = 10$ cities, and it is not even certain that the list is complete.

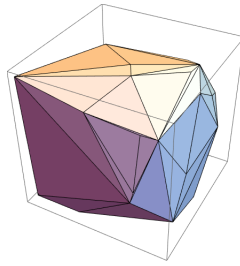


Figure 2. A complicated polytope contained in a cube in dimension 3.

the new hyperplane to the constraints, the algorithm gets a tighter feasible region and now a new \mathbf{x}^* is gained, which is hopefully closer to the actual optimum, by repeating the procedure. It is like cleaving the space repeatedly to reach the the “inner” polytope. This method is called the *cutting plane* and it applies to many other combinatorial optimization problems. Finding the cutting plane requires establishing a common combinatorial property among the feasible solutions and describing it with linear inequalities. This approach led to bilateral developments of combinatorial polyhedral theory and linear programming.

The other idea for speeding up the process is the *branch and cut* algorithm. It splits the feasible solutions into branches usually by assigning a value to some variables and optimizing the linear program in the nodes which are sub-problems actually. Whenever \mathbf{x}^* is found — the lower bound for the solution obtained by the relaxation of the linear program — larger than the optimal solution in that branch, the node is pruned.

The method can be used for tackling other combinatorial optimization problems and it is the base idea behind the **Concord TSP Solver**³, written by David Applegate, Robert E. Bixby, Vašek Chvátal and William J. Cook. The software is currently the fastest TSP solver and can solve tremendously large instances. In April 2006 an instance with 85,900 points was solved using Concorde TSP Solver, taking over 136 CPU-years, a major breakthrough within almost 50 years of solving TSP from 49 cities to 85,900 cities.

All this rapid development of applications of linear programming would not be possible without taking the advantage of “Polyhedral Theory”.

Outline and contributions

In **Chapter 2**, we provide preliminary definitions and results on linear programming and polytopes. In particular we introduce extended formulations in linear programming, combinatorial methods for lower bounding the extension complexity of polytopes and the graph a polytope. We discuss briefly some important examples of polytopes and the known results about their extension complexity.

In **Chapter 3**, we discuss a brief overview of the the following paper:

³<http://www.math.uwaterloo.ca/tsp/concorde/>

- Makkeh, A.; Pourmoradnasseri, M. and Theis, D.O. “The Graph of the Pedigree Polytope is Asymptotically Almost Complete (Extended Abstract)”. In *Proceedings of The International Conference on Algorithms and Discrete Applied Mathematics, CALDAM 2017*.

The author’s contribution among the others, is introducing the stochastic process (s, t) , describing the number of common edges and connected components and analyzing its return to $s = 0$.

In **Chapter 4**, we present an introduction on communication complexity and discuss the connection of nondeterministic communication complexity for lower bounding the extension complexity of polytopes.

In **Chapter 5**, we discuss a brief overview of the the following paper:

- Pourmoradnasseri, M. and Theis, D.O. “Nondeterministic Communication Complexity of Random Boolean Functions (Extended Abstract)”. In *Proceedings of Theory and Applications of Models of Computation, TAMC 2017*.

The paper abounds in application of chernoff bounds, estimates and delicate inequalities which the author worked out for the most part. Also the author’s contribution among the others, is designing the “conditioning on matching” approach to linking fooling set size to independence number in Theorem 3.1(a).

In **Chapter 6**, we discuss a brief overview of the the following paper:

- Pourmoradnasseri, M. and Theis, D.O. “The (Minimum) Rank of Typical Fooling Set Matrices”. In *Proceedings of International Computer Science Symposium in Russia, CSR 2017*

The author’s contribution among the others, is applying the theorem by Ronyai, Babai and Ganapathy to the fooling set case also flushing out the details of the counting arguments involving sparse tee-matrices.

2. POLYTOPES AND EXTENSIONS

2.1. Preliminaries

In this thesis, we take the pre-knowledge of Discrete Mathematic as granted. For terminology, definitions and basic results we refer to the text book by Matoušek and Nešetřil [76]. Also vectors are always represented by boldface characters.

2.1.1. Linear programming and sizes of linear programming formulations

The linear programming problem, LP for short, seeks an optimal solution — such as minimum cost or maximum profit— of a linear function subject to linear constraints. A linear program is a special case of mathematical optimization. LP has found a numerous application in many real life problems [15, 24, 66].

There are several equivalent forms of representing a linear programming problem. One of the most common forms is

$$\begin{aligned} & \text{minimize}_{\mathbf{x}} \mathbf{c}^T \mathbf{x} \\ & \text{subject to } A\mathbf{x} \geq \mathbf{b} \\ & D\mathbf{x} = \mathbf{e} \end{aligned}$$

where $A \in \mathbb{R}^{m \times n}$, $D \in \mathbb{R}^{k \times n}$, $\mathbf{c} \in \mathbb{R}^n$ and \mathbf{b} and $\mathbf{e} \in \mathbb{R}^m$.

The first general method for solving LP was proposed and developed by Kantorovich in 1939 during World War II to optimize the cost of armies. At approximately the same time, Koopmans independently used the linear program formulation in classical economic problems. Later, in 1975, Kantorovich and Koopmans shared the Nobel prize in economics.

In 1947 Dantzig published a method for solving LPs called the *simplex method* [14] which is still noteworthy for its efficiency in practice. Soon after that, in 1948, von Neumann conjectured the so-called *theory of duality*, immediately after Dantzig presented his simplex method, realizing the connection with the problem he had been working on in game theory.

To state it roughly, the simplex method starts from an initial feasible solution, a vertex of the polytope, and moves to another vertex of the polytope P , which is the representation of the feasible solutions, along the edges. The procedure is finding an edge of a polytope P whose direction decreases the value of objective function. The algorithm achieves an optimal solution as soon as it encounters a vertex where no such edge exists. The average case complexity of the simplex method is polynomial [10] and so it is efficient in practice, but its worst case complexity is exponential [56].

For more than half a century, there have been extensive attempts to theoretically explain the good performance of the simplex method. A popular approach in this area was proving that there is always a short walk from every vertex to

the optimal vertex. The Hirsch conjecture is an example of attempts for lower bounding the steps of simplex method. It was posed in 1957 in a question from Hirsch to Dantzig and states that the edge-vertex graph of an n -facet polytope in d -dimensional space has diameter no more than $n - d$ (see section 2.3 for definition of diameter of a polytope.). Despite being one of the most fundamental, basic and old problems in polytope theory, Hirsch's conjecture was disproved in general ¹ more than 50 years later in 2010 by Santos [91]. In 1992, Kalai and Kleitman [50] proved that there always exists a walk of length at most $n^{\log_2 d + 2}$ between every two vertices of a polytope. However the existence of a short walk in the polytope does not guarantee that it can be found by simplex method.

A natural question that could arise was: Is LP solvable in polynomial time, in terms of $n = \dim(\mathbf{x})$ and L , where L is the bitlength of the input? The question was answered affirmatively in 1979 by Khachiyan [52] by introducing the *ellipsoid method*. However, although the method was theoretically the first polynomial-time algorithm for solving LP, it was not efficient in practice.

A few years later, in 1984, the *interior point method* was introduced by Karmarkar [51]. The method has proven to have good result in theory and practice and has been investigated extensively in several variations. It uses the *standard* form of the linear program as

$$\begin{aligned} & \text{minimize}_{\mathbf{x}} \mathbf{c}^T \mathbf{x} \\ & \text{subject to } \mathbf{A}\mathbf{x} = \mathbf{b} \\ & \mathbf{x} \geq 0 \end{aligned}$$

where $A \in \mathbb{R}^{m \times n}$, $\mathbf{c} \in \mathbb{R}^n$, $\mathbf{b} \in \mathbb{R}^m$. However, all the other forms of the linear programs can be converted to the standard form by adding slack variables to inequalities.

Intuitively, given a polytope P and an interior point $\mathbf{a} \in P$, the algorithm produces a sequence of pairs of solutions for primal and dual problem which converge to the optimal solution using a sequence of projections. The computational complexity of the algorithm is $O(n^{3.5}L)$ in general, with L be the size of input [101]. The running time of this method is polynomial in the number of variables and inequalities [98]. Unfortunately, in most of the combinatorial optimization problems, the number of inequality constraints is exponential in terms of variables. Therefore, finding the methods which can give smaller linear description for the problems is of the interest.

2.1.2. Polytopes and their facets

Polytopes are the generalization of polygons in higher dimensions. Convex polytopes are fundamental geometric objects in optimization since they describe the

¹ The Hirsch conjecture is still true for (0,1) polytopes [73]

feasible solution space of the linear programs. In particular, combinatorial optimization searches for an optimum object in a finite set of objects. The objects are represented by vectors and construct the vertices of a convex polytope. The number of vertices is usually exponential in the size of the problem (e.g. all the Hamiltonian cycles or all the spanning trees of a complete graph). In combinatorial optimization, the challenge is not only to reduce the running time of the LP algorithms but also to find an appropriate linear description of the feasible solution space.

In this section we give the basic definitions and fundamental properties of polytopes which will be used in following chapters. For all notions and results from polytope theory mentioned in the presented work, we refer to Ziegler [104].

A \mathcal{V} -polytope is the convex hull of a finite set of points $K = \{\mathbf{v}_1, \dots, \mathbf{v}_n\} \subset \mathbb{R}^d$ with $n \geq 1$:

$$P = \text{conv}(K) = \{\lambda_1 \mathbf{v}_1 + \dots + \lambda_n \mathbf{v}_n : \lambda_i \geq 0, \sum_{i=1}^n \lambda_i = 1\}.$$

An \mathcal{H} -polytope is a bounded intersection of a finite number of half-spaces in some \mathbb{R}^d , which can be presented in the form:

$$P = P(A, \mathbf{z}) = \{\mathbf{x} \in \mathbb{R}^d : A\mathbf{x} \leq \mathbf{z}\} \text{ for some } A \in \mathbb{R}^{m \times d} \text{ and } \mathbf{z} \in \mathbb{R}^m.$$

An \mathcal{H} -polyhedron is the intersection of finitely many half-spaces in \mathbb{R}^d . An \mathcal{H} -polyhedron can be unbounded. In this thesis we are only concerned with polytopes.

A polytope is a point set $P \subset \mathbb{R}^d$ which can be presented either as a \mathcal{V} -polytope or \mathcal{H} -polytope. The *Minkowski-Weyl* theorem states that these two representations are equivalent.

Theorem 1. [70] *A subset $P \subset \mathbb{R}^d$ is the convex hull of a finite set of points (a \mathcal{V} -polytope) if and only if it is a bounded intersection of a finite number of half-spaces (an \mathcal{H} -polytope).*

The importance of the theorem 1 comes from the fact that it ensures that every polytope has both \mathcal{V} -polytope representation and \mathcal{H} -polytope representation and either of them can be referred whenever needed.

A *face* F of a polytope P is defined as an intersection $P \cap H$ where H is an affine hyperplane for which the polytope is contained entirely in one of the two halfspaces determined by the hyperplane. Equivalently, a face F is a subset of the polytope P such that there exists an inequality $\mathbf{a}^T \mathbf{x} \leq b$ which is satisfied by all $\mathbf{x} \in P$ and $F = \{\mathbf{x} | \mathbf{a}^T \mathbf{x} = b\}$.

For every polytope, the empty set and the polytope itself are considered as the (non-proper) faces. *Vertices* of a polytope are dimension zero faces. Line segments, known as *edges*, are faces of polytope with dimension one. The maximal proper faces of a polytope are *facets*. In the other word, facets are faces with dimension $\dim(P) - 1$. For example, the proper faces of a 3-dimensional polytope

are its vertices, edges and the boundary polygons. We do not give the exact definition of the “dimension” here and ask the reader to rely on his/her intuition to understand it.

Having a linear description of a polytope with the minimum number of inequalities, each inequality corresponds to a facet, such as giving shape to the polytope with cutting the space with hyperplanes.

2.1.3. Examples of polytopes

Combinatorial optimization is used to find an optimal value among the set of feasible solutions of a problem. Feasible solutions, can be considered as vectors in some \mathbb{R}^n and the convex hull of these vectors constitutes a polytope.

In this section, we present three well-studied examples of combinatorial polytopes.

Spanning tree polytope. Spanning tree polytope P_{ST} is the convex hull of the characteristic vectors of all spanning trees of the complete graph $K_n = (V, E)$. Letting $\mathcal{T}(n)$ be the set of all spanning trees of the complete graph K_n , then

$$P_{ST} = \text{conv}\{\chi(T) \in \mathbb{R}^E : T \in \mathcal{T}(n)\}.$$

$\chi(T)$ denotes the characteristic vector of the spanning tree, that is $\chi(T) \in \{0, 1\}^E$. An entry of the vector is equal to 1 if and only if its corresponding edge belongs to the spanning tree T .

Edmonds showed that the spanning tree polytope admits the following linear description. $E(S)$ stands for the set of edges induced by the vertex set S .

$$\begin{aligned} \sum_{e \in E} x_e &= n - 1 \\ \sum_{e \in E(S)} x_e &\leq |S| - 1 && \text{for all nonempty } S \subsetneq V \\ x_e &\geq 0 && e \in E. \end{aligned}$$

In this linear formulation, there are exponentially many inequalities, respectively facets. Also, none of the inequality constraints is redundant [22].

Perfect matching polytope. The perfect matching polytope P_{PM} is the convex hull of all characteristic vectors of the perfect matchings of the complete graph $K_n = (V, E)$. If $\mathcal{M}(n)$ is the set of all the perfect matching of K_n ,

$$P_{PM} = \text{conv}\{\chi(M) \in \mathbb{R}^E : M \in \mathcal{M}(n)\}.$$

There is also a linear description by Edmonds [21] for the perfect matching polytope.

$$\begin{aligned}
\sum_{e \in \delta(v)} x_e &= 1 && \text{for all } v \in V \\
\sum_{e \in \delta(U)} x_e &\geq 1 && \text{for all } U \subseteq V \text{ with } |U| \text{ odd} \\
x_e &\geq 0 && e \in E.
\end{aligned}$$

Here $\delta(v)$ is the set of all incident edges to the vertex v and $\delta(U)$ is the set of all edges with exactly one end point in U .

In the linear description of P_{PM} there are n equality constraints in total, one for each vertex and $O(n^2)$ non-negativity constraints, but exponentially many odd-set constraints. Perfect matching polytope is an interesting polytope for its lower bound on extended formulation size which will be discussed later.

Traveling salesman polytope. Another example we go through here, is the *traveling salesman polytope*, P_{TSP} associated with the traveling salesman problem. It is probably one of the most intensively studied problems in combinatorics and computer science. It was first defined in the 1800s by the Irish mathematician W. R. Hamilton and by the British mathematician Thomas Kirkman. Assume a salesman who wants to visit n cities and come back to the first city again. All cities are connected and traveling between every two cities has a cost (flight ticket, time, etc). The goal is choosing an order of cities to travel to that keeps the total cost of the travel as low as possible.

The traveling salesman problem is a typical example of the class of NP-hard problems in mathematics and it has vast applications in science and industry.

More formally, the traveling salesman problem (TSP for short) is this: given a complete graph $K_n = (V, E)$ along with cost c_{ij} for the edge $\{i, j\}$, find a cycle (Hamiltonian cycle) with the minimum cost. Here we consider only the symmetric TSP which means edges of the K_n are not directed ($c_{ij} = c_{ji}$).

Every feasible solution of TSP is a cycle and each cycle is expressed by its Boolean characteristic vector of size $\binom{n}{2}$. The convex hull of all these vectors makes the TSP polytope. In other words, letting $\mathcal{C}(n)$ be the set of all Hamiltonian cycles of the complete graph K_n , then

$$P_{TSP} = \text{conv}\{\chi(C) \in \mathbb{R}^E \mid C \in \mathcal{C}(n)\}.$$

The dimension of P_{TSP} is known by Grötschel and Padberg to be $n(n-3)/2$ [38].

There have been many attempts to understand the TSP polytope and illuminate the structure of its faces and facets ² [35, 95] but only a few are known. Even for the case $n = 10$, it is an open problem whether the current linear description is the complete one or not.

²See <http://www.iwr.uni-heidelberg.de/groups/comopt/software/SMAP0/tsp/tsp.html> for the library of known linear descriptions and clasification of faces of TSP up to 10 cities.

2.2. Extensions of polytopes

The idea of lifting a polytope for finding a more efficient representation of it — mainly for optimizing more efficiently over the polytope — is a well-known topic in linear programming. Extension of a polytopes means basically lifting the polytope into a higher dimensional space by adding more variables.

2.2.1. Extension

An *extended formulation* of a polytope $P \subseteq \mathbb{R}^d$ is the polytope $Q \subseteq \mathbb{R}^e$ with an affine map³ $\pi : \mathbb{R}^e \rightarrow \mathbb{R}^d$ such that $\pi(Q) = P$. Then Q is the *extension* of P . The number of facets of Q is known as the *size* of the extension.

The *extension complexity* of a polytope P , $xc(P)$ is the minimum number of facets among all possible extensions of P . As mentioned earlier, in the \mathcal{H} -polytope representation with the minimum number of inequalities⁴, every inequality corresponds to a facet of the polytope. So extension complexity is the minimum number of inequalities (not equalities) that can describe an extension of the polytope.

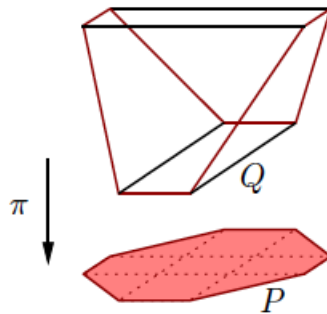


Figure 3. Q with 6 facets is an extension of P with 8 facets [27].

The idea behind the extension of polytopes is as follows: sometimes a polynomial increase in dimension of a polytope makes an exponential decrease in the number of inequalities describing the polytope. Viewed differently a projection of a polytope to a lower dimension, may have larger number of facets. Decreasing the number of facets (inequalities), yields a considerable improvement in the running time of optimizing over the polytope using methods like interior point method since the complexity of the interior point method is polynomial in the size of inequalities and variables.

Among the polytopes associated with combinatorial optimization problems, some of them like the spanning tree polytope of the complete graph or permutahedron have surprisingly small extended formulation. For some like the traveling

³The affine map can be considered safely as a linear map here such that it does not necessarily preserve 0.

⁴We only talk about inequalities and not equalities

salesman problem or matching polytope, it is proven that no polynomial extended formulation exists and there are still many unresolved problems regarding upper and lower bounds on extension of polytopes.

2.2.2. Slack matrix

Let the polytope P be the convex hull of $V = \{\mathbf{v}_1, \dots, \mathbf{v}_n\} \subset \mathbb{R}^d$ and also represented by $\{\mathbf{x} \in \mathbb{R}^d : A\mathbf{x} \leq \mathbf{z}\}$ for some $A \in \mathbb{R}^{m \times d}$ and $\mathbf{z} \in \mathbb{R}^m$. The *slack matrix* of P (with respect to V , A , and \mathbf{z}) is the $m \times n$ matrix S whose ij -th entry is $s_{ij} := z_i - A_i \mathbf{v}_j$, the slack of the j -th element of V with respect to the i -th inequality.

In almost all techniques that provide a lower bound on the size of extended formulation, the slack matrix plays the main role. The non-negative rank and rectangle covering of the slack matrix will be discussed in the following sections.

2.2.3. Non-negative factorization

Yannakakis in his seminal paper [102] showed the equivalence of the geometric parameter, *extension complexity* and the algebraic parameter, *non-negative rank* of the slack matrix associated with the polytope. In this section we study the relationship among these parameters.

The *non-negative rank* of a matrix M , $\text{rank}_+(M)$, is the smallest $r \in \mathbb{N}$ such that M can be expressed as $M = TU$ where $T \in \mathbb{R}_+^{m \times r}$, $U \in \mathbb{R}_+^{r \times n}$ are non-negative matrices. Equivalently $\text{rank}_+(M)$ can be defined as the minimum number r such that M can be decomposed to the sum of r non-negative rank-1 matrices, $M^{(i)}$, $M = \sum_{i=1}^r M^{(i)}$. In this thesis the latter one is taken as the definition.

Theorem 2. [102] *The extension complexity of a polytope P of a dimension greater than zero, is equal to the non-negative rank of its slack matrix.*

Theorem 2 states that finding the lower bound on extension complexity of polytopes is equivalent to finding the lower bound on non-negative rank of the slack matrix. Although determining a non-negative rank of a matrix is a difficult problem in itself [12], matrices are more familiar objects for study. Clearly the normal rank of a matrix is always a lower bound for its non-negative rank but deciding on whether the non-negative rank of a matrix is equal to its normal rank is NP-hard [100]. Finding a reasonable lower bound for non-negative rank is a topic of interest not only for linear programming, but also for other areas such as analyzing data, image and clustering [19, 71]. There are still many unknown problems regarding the non-negative rank and its complexity [55]. A combinatorial method for lower bounding the non-negative rank is determining the rectangle covering number of the matrix, which will be discussed in the next section.

Remark 1. The slack matrix of a polytope can be defined more generally. In section 2.2.2, only the slack of vertices and facets are considered, but one may define the slack matrix containing the slack of vertices and some other additional

faces or even all the faces. In this case, the dimension of the slack matrix is larger than the one defined in 2.2.2, but the positive rank of the matrix stays unchanged [27, 102].

2.2.4. Rectangle covering and non-negative rank

For a positive integer n , a *rectangle* is a product of $R = K \times L \subset [n] \times [m]$ (with $[n] := \{1, \dots, n\}$).

Given a Boolean $n \times m$ matrix M , a *1-rectangle* is rectangle R with $M_{k,\ell} = 1$ for all $(k, \ell) \in R$. A *rectangle covering* of M is a collection of 1-rectangles R_1, \dots, R_r such that $\{(k, \ell) \mid M_{k,\ell} = 1\} = \bigcup_i R_i$, or, informally, every 1-entry of M is contained in one of the 1-rectangles chosen. The *rectangle covering number* [61, 62], $C(M)$, of M is the smallest number of 1-rectangles in a rectangle covering of M .

According to the definition, the non-negative rank of a matrix M is equal to the minimum number r , such that $M = \sum_{i=1}^r M^{(i)}$. Each $M^{(i)}$ is a positive rank-1 matrix, thus its non-zero entries give a rectangle. Obviously, if we discard the values of the non-zero entries in M and just look at the zero-non-zero pattern, the non-zero entries of each $M^{(i)}$ induces a rectangle R_i . So the set of rectangles R_1, \dots, R_r will cover all the non-zero entries of M .

The *support* of matrix M is the matrix obtained by keeping the zero entries and replacing the non-zero entries with 1, denoted by $\text{supp}(M)$. So rectangles of R_1, \dots, R_r give a rectangle covering for the support of the matrix M . Put differently, the $C(\text{supp}(M))$ is a lower bound for $\text{rank}_+(M)$.

2.2.5. Lattice embedding

The set of all faces of a polytope can be regarded as a *lattice*. A lattice is a partially ordered set such that every two elements have a unique supremum and infimum. The *face lattice* of a polytope P , noted by $\mathcal{L}(P)$, is the set of all faces of the polytope, including the trivial faces \emptyset and P , partially ordered by inclusion. In $\mathcal{L}(P)$ facets are the proper maximal faces of $\mathcal{L}(P)$.

Let $\pi : \mathbb{R}^e \rightarrow \mathbb{R}^d$ be an affine map and let $Q \subset \mathbb{R}^e$ be a polytope. Then $\pi(Q)$ is the *projection* of Q under π .

A map f is an *embedding* of a partially ordered set (O, \leq) into (S, \sqsubseteq) if it preserves the order. It means for $u, v \in O$, $u \leq v$ if and only if $f(u) \sqsubseteq f(v)$.

Proposition 1. [27] Let $Q \subset \mathbb{R}^e$ along with the affine map $\pi : \mathbb{R}^e \rightarrow \mathbb{R}^d$ with $\pi(Q) = P$ be the extended formulation of $P \subset \mathbb{R}^d$. Then the map $h : \mathcal{L}(P) \rightarrow \mathcal{L}(Q)$ which assigns $h(F) := Q \cap \pi^{-1}(F)$ to each face F of P is an embedding.

Proof. Obviously $h(\emptyset) = \emptyset$ and $h(P) = Q$. If F is a face of P , then according to the definition there exists an inequality $\mathbf{a}^T \mathbf{x} \leq b$ which represents F . It is also satisfied by all $y \in Q$ with $\mathbf{a}^T \pi(\mathbf{y}) \leq b$ and equality holds if and only if $\pi(\mathbf{y}) \in F$. Hence $h(F)$ is a face of Q represented by the inequality $\mathbf{a}^T \pi(\mathbf{y}) \leq b$. The map

h preserves the order and $\pi(h(F)) = F$, for every face $F \in \mathcal{L}(P)$, therefore it is injective and so, embedding. \square

Some facts can be concluded from the proof of proposition 1. The image $h(F)$ of a face F is a face. Moreover every extension Q of a polytope P , induces an embedding from $\mathcal{L}(P)$ into $\mathcal{L}(Q)$. Hence, the minimum number of facets of a polytope Q , such that the face lattice of P can be embedded into the face lattice of Q , gives a lower bound on the extension complexity of the polytope P .

The embedding $h : \mathcal{L}(P) \rightarrow \mathcal{L}(Q)$ induces a rectangle covering for the support of the slack matrix of P , S , of the size of number of facets in Q [27]. Let $\{F_1, \dots, F_k\}$ be the set of facets of Q . Define the rectangles $R_i = I_i \times J_i$ for $i = \{1, \dots, k\}$ as following. Let I_i be the set of all rows of the slack matrix indexed by faces U of P such that $h(U) \subseteq F_i$ and let J_i be the set of all columns indexed by the vertices v such that $h(\{v\}) \not\subseteq F_i$. The set $\{R_1, \dots, R_k\}$ indeed gives a rectangle covering for the support of S because every non-zero entry of S corresponds to a face U and a vertex v of P , such that $\{v\} \not\subseteq U$. Since the embedding h is order preserving, $h(\{v\}) \not\subseteq h(U)$. By lattice properties, there exists a facet F_i such that it contains $h(U)$ but not $h(\{v\})$.

2.2.6. Extended formulation of some polytopes

After being known that LP is solvable in polynomial time, there had been a sequence of attempts to prove P=NP via finding a polynomial size linear description for known hard problems– in particular TSP [97]. Due to the large size and complicated formulation, it was “hard to tell what they do or do not express” [102]. In his seminal paper, Yannakakis [102] ruled out all the attempts in this direction conveniently by proving that every symmetric⁵ linear formulation of TSP must have an exponential size.

A natural question that may arise is which problems admit polynomial size extended formulation and which do not. Many open problems remain in this area. A few known bounds are mentioned briefly in this section and in the next chapter.

Extended formulation of spanning tree polytope. In this section, we give the polynomial size extended formulation of spanning tree polytope which is due to Martin [67]. It is one of the well-known and simple examples of extended formulation. As it was mentioned earlier, the spanning tree polytope is defined as follows:

$$P_{ST} = \left\{ \mathbf{x} \in \mathbb{R}^E : \begin{aligned} &\sum_{e \in E} x_e = n - 1 \\ &\sum_{e \in E(S)} x_e \leq |S| - 1 && \text{for all nonempty } S \subsetneq V \\ &x_e \geq 0 && e \in E \end{aligned} \right\}.$$

⁵“an LP is called symmetric if every permutation of the cities can be extended to a permutation of all the variables of the LP that preserves the constraints of the LP” [28].

The size of the spanning tree polytope is exponential and the following formulation gives a polynomial size extension of P_{ST} :

$$\begin{aligned}
Q_{ST} = \{(\mathbf{x}, \mathbf{y}) \in \mathbb{R}^E \times \mathbb{R}^{n(n-1)(n-2)} \\
& x_{\{v,w\}} - y_{v,w,u} - y_{w,v,u} = 0 && u, v, w \text{ are distinct} \\
& x_{\{v,w\}} + \sum_{u \in V \setminus \{v,w\}} y_{v,u,w} = 1 && v, w \text{ are distinct} \\
& \sum_{e \in E} x_e = n - 1 \\
& \mathbf{x}, \mathbf{y} \geq 0 && e \in E\}.
\end{aligned}$$

To see how a spanning tree T satisfies the new formulation, it is sufficient to assign the value 1 to $y_{v,w,u}$ if $(v, w) \in T$ and u is on w 's side of the edge (v, w) in T and 0 otherwise. It gives the inclusion $P \subseteq \pi(Q)$. For the proof of the reverse inclusion we refer to [67].

From this formulation, $\text{xc}(P_{ST}(n)) = O(n^3)$. It is an open problem whether $\text{xc}(P_{ST}(n)) = \Theta(n^3)$.

Lower bound on the extension complexity of the perfect matching polytope. The perfect matching polytope is a distinguished example by its extension complexity. The linear program description of this problem as given in 2.1.3, has an exponential size. But there are polynomial algorithms for optimizing a linear function over the perfect matching polytope [21]. The question about the existence of polynomial size extended formulation for matching polytope remained unsolved since Yannakakis' paper [102]. The problem was settled by Rothvoß in 2013 [88] and he showed that surprisingly every extension of the matching polytope has super polynomial size.

Theorem 3. [88] *For all even n , $\text{xc}(P_{PM}(n)) = 2^{\Omega(n)}$.*

This result is particularly interesting because all the other exponential lower bounds on extension complexity of the polytopes are among the polytopes associated with NP-hard problems.

The best previous known lower bound for $\text{xc}(P_{PM}(n))$ was $\Omega(n^2)$ [102]. Also, upper bound $O(n^4)$ on the rectangle covering number of the slack matrix S [27], ensured the rectangle cover by itself can not give any super polynomial lower bound for the extension complexity of the perfect matching polytope. Here we depict the rough idea on the upper bound on the rectangle covering number of S .

There are three types of constraints in 2.1.3. The number of degree constraints and non-negativity constraints is $\Theta(n^2)$ and only the number of odd set constraints is exponential. So it is sufficient to only look at the odd set constraints. If U is an odd set and M is a matching, then it is not difficult to observe that $S_{UM} = |\{\delta(U) \cap M\}| - 1$, where S_{UM} is the entry of the slack matrix corresponding to the odd set U and the matching M . For every pair of edges e_1, e_2 , the rectangle $R_{e_1 e_2} := \{U | e_1, e_2 \in \delta(U)\} \times \{M | e_1, e_2 \in M\}$ induces a 1-rectangle in the slack

matrix. There are $O(n^4)$ many such rectangles and it deduces an upper bound of $O(n^4)$ on the rectangle covering number.

Rothvoß showed [88] that an entry with $S_{UM} = k$ is covered by $\theta(k^2)$ many rectangles. Therefore every polynomial size rectangle cover of the slack matrix, over-covers the non-zero entries of the slack matrix and there is a large gap between $\text{rank}_+(S)$ and $C(S)$ in this problem, hence the rectangle cover bound is not useful in this problem. Then using *hyperplane separation lower bound* suggested by Fiorini, he concluded that the extension complexity of the matching polytope is exponential.

2.3. Graph of a polytope

The k -skeleton of a d dimensional polytope is the set of all faces of the polytope with dimension less than or equal to k . The 1-skeleton or the graph of a polytope P , $G(P)$, is the set of vertices and edges of the polytope. Two vertices are adjacent if they are end-points of an edge (1-dimensional face) of the polytope. A fundamental theorem in polyhedral theory by Balinski states that the graph of a d -dimensional polytope is d -connected [6] (see theorem 3.14 of [104]). As a consequence, the minimum degree of a graph of a polytope is at least the dimension of the polytope.

The graph of a polytope can be regarded as an abstract graph and investigating the graph theoretical properties of it reveals meaningful information about properties of the polytope. For instance, the theorem by Blind and Mani [9] states that so-called simple polytopes⁶ are determined by their graphs. The famous theorem by Steinitz [8] characterizes exactly the 3-dimensional polytopes as the 3-connected planar graphs.

Understanding the graph of polytopes of higher dimensions and in more general form has been more challenging, however [6, 25]. The concept is of interest not only in combinatorial polyhedral theory, but also in combinatorial optimization and theoretical computer science [1, 7, 65].

In combinatorial optimization, particularly after developments of linear programming, polytopes received a considerable amount of attention. Some of the motivations were understanding the running time of simplex method, improving the linear programming techniques and even attacking P vs. NP [102]!

For instance, learning the diameter of a polytope (= diameter of graph of the polytope) gives a lower bound for the number of iterations in simplex and randomized simplex method [47, 48]. The famous example was the Hirsch conjecture which was answered by Santos [91] after fifty years. Although the Hirsch conjecture was disproved in general, attempts for proving an upper bound for diameter of polytopes which is polynomial in the number of facets of polytope are still undergoing [23, 53, 92].

⁶A d -polytope is simple if each of its vertices is adjacent to exactly d edges and also d facets.

Among all the attempts to understand the graph of combinatorial polytopes, TSP polytope and also TSP related polytopes have gotten a considerable amount of attention (e.g., [96]; cf.[35, 74, 77] and their references). The presence of long cycles has been studied ([95], see also [72, 75]), as has the graph density and vertex degrees (e.g., [93], see also [41, 45]).

In theoretical computer science, finding an algorithm which can verify a polytope from its k -skeleton is of interest [29, 46]. Deciding whether a given lattice is a face lattice of a polytope is known to be NP-hard [85]. There are several interesting algorithmic problems regarding the graph of a polytope and unknown facts about computational complexity of them (see [44] for a collection of problems).

For more results on the skeleton of polytopes we refer to the book chapter by Kalai [49].

The natural question that arises is about the connection between graphs of two polytopes P and Q , when Q is an extension of P .

In general, the projection $\pi : Q \rightarrow P$ may project a vertex of Q somewhere in middle of the polytope P and not onto vertices. This type of vertex is called a *hidden vertex* [81]. If a projection projects all the vertices of Q onto vertices of P , then the extension is *without hidden vertices*.

Definition 1. A graph H is a minor of a graph G if it can be obtained from G by any sequence of contracting edges, deleting edges, and deleting isolated vertices.

Definition 2. $G(V, E)$ contains $H(V', E')$ as a model, if G consists $|V'|$ vertex-disjoint connected subgraphs $B_1, \dots, B_{|V'|}$ such that for all distinct i and j some vertex in B_i is adjacent to some vertex in B_j if $\{i, j\} \in E'$ in H .

Clearly, H is a minor of G if and only if G contains H as a model.

The following proposition shows the connection between the graph of two polytopes, while one is an extension of the other. Since a polytope and its graph are essentially the same object, we may refer to the polytope P as its graph $G(P)$.

Proposition 2. If Q is an extension of P with $\pi : Q \rightarrow P$, then

- (a) $G(P)$ is a minor of $G(Q)$.
- (b) If Q is an extension of P without hidden vertices, then $G(P)$ is a model of $G(Q)$.
- (c) If $V(Q)$ and $V(P)$ are in bijection then $G(P)$ is a spanning subgraph of $G(Q)$.

Proof. We start with the general case (a). For every v in vertices of P , the set $F_v := \pi^{-1}(\{v\}) \cap Q$ is a face of Q and hence a polytope itself. Balinski theorem [6] states that the graph of every polytope of dimension k is k -connected. So, every vertex in $G(P)$ corresponds to a connected subgraph of $G(Q)$ and this subgraphs are disjoint.

The preimage of every edge $e = \{v_1, v_2\}$, $F_e := \pi^{-1}(\{e\}) \cap Q$ is also a face of Q and by proposition 1, it contains F_{v_1} and F_{v_2} . Hence, there is a path between F_{v_1} and F_{v_2} in $G(F_e)$. It remains to prove that for every two disjoint edges $\{v_1, v_2\}$

and $\{u_1, u_2\}$ in $G(P)$ there exist two vertex disjoint paths in $G(Q)$, one between F_{v_1} and F_{v_2} and the other between F_{u_1} and F_{u_2} .

Let $S_{v_1v_2}$ and $S_{u_1u_2}$ be the shortest paths between F_{v_1} and F_{v_2} and between F_{u_1} and F_{u_2} . Assume $w \in S_{v_1v_2} \cap S_{u_1u_2}$. It implies $w \in F_{\{v_1, v_2\}} \cap F_{\{u_1, u_2\}}$ and so $w \in \pi^{-1}(\{v_1, v_2\} \cap \{u_1, u_2\}) \cap Q$. It means $\{v_1, v_2\} \cap \{u_1, u_2\} \neq \emptyset$ which is contradiction.

In the case (b), there is no hidden vertex in the extended formulation. So, the vertex set of the face F_e , $e = \{v_1, v_2\}$, can be partitioned to two subsets F_{v_1} and F_{v_2} . Since F_e is connected, there must be an edge between F_{v_1} and F_{v_2} . Therefore $G(Q)$ is a model of $G(P)$.

In the case (c), π induces a bijection between the vertex sets of P and Q . Using part (b), there has to be an edge between F_{v_1} and F_{v_2} in $G(Q)$ for $e = \{v_1, v_2\}$ in $G(P)$. Hence $G(P)$ is a spanning subgraph of $G(Q)$. \square

3. ON THE GRAPH OF THE PEDIGREE POLYTOPE

In this chapter, we discuss the motivation and results of the paper [64], reprinted as the Appendix B of this thesis.

3.1. Motivation and previous works

In this paper, we studied the graph of the pedigree polytope. Our original motivation was the thirty year old conjecture by Grötschel and Padberg [35] stating that the diameter of the graph of TSP polytope is 2. Grötschel and Padberg also extended their question to the family of TSP-related polytopes [35] since there are quite few known facts about the structure of TSP polytope. The conjecture was already proven for the asymmetric TSP [78], but for the symmetric case only the upper bound 4 has been obtained [86].

Padberg and Rao in [78], proved the upper bound 2 on diameter of a class of “algorithmically well-solved” combinatorial problems containing assignment problem, the edge-matching problem on complete graphs, the multi-dimensional assignment problem and many other set partitioning problems. They proposed the diameter of a polytope associated with a combinatorial problem as a measure of complexity of the problem and surprisingly they proved the asymmetric TSP polytope also falls in the same category of the polytopes with small diameter.

The theorem by Papadimitriou [79] states that the non-adjacency of vertices of (Symmetric) Traveling Salesman Problem (TSP) polytopes is NP-complete. The question about non-adjacency of vertices in other families of polytopes has also been studied (cf. [1, 65]).

Pedigree polytopes are a family of TSP-related polytopes introduced by Arthanari [5]. The graph of the pedigree polytope has a nice combinatorial structure and adjacency of vertices can be decided in polynomial time [3].

As in the case of TSP, the vertices of the pedigree polytope, correspond to the Hamiltonian cycles of the complete graph K_n . Therefore the number of vertices is equal to $(n-1)!/2$. In the Arthanari’s idea of the pedigree, cycles evolve over the time. The initial cycle is $\{1, 2, 3\}$ at time 3 and at time $n \geq 4$, the vertex n is inserted into an existing edge of the the cycle with the vertex set $[n - 1]$ and subdivides it into two new edges.

Lemma 1 ([4]). *The pedigree polytope is an extension of TSP polytope, without “hidden” vertices.*

In fact, the vertex set of TSP polytope and pedigree polytope are in bijection. Hence, by the proposition 2, the graph of the TSP polytope is a spanning sub-graph of the pedigree. Unlike TSP polytope graphs, pedigree polytope graphs are not vertex transitive and not even regular. Arthanari’s construction removes the symmetry from the graphs of the polytopes.

3.2. Our result

In our paper we proved the following about the graph of the pedigree polytope.

Theorem 4. *The minimum degree of a vertex on the Pedigree polytope for n cities is $(1 - o(1)) \cdot (n - 1)!/2$ (for $n \rightarrow \infty$).*

Theorem 4 means that the graph of the pedigree polytope is “asymptotically almost complete”. However, in the numerical simulations we observed even for large n (≈ 100) the graph of the pedigree polytope is not complete.

Arthanari defines some combinatorial conditions on the adjacency of the vertices in the pedigree polytope. We do not discuss the conditions here because of being too technical. But fortunately, the proof idea can be understood without the technical details. We refer to the paper [64] for the precise statements.

Consider two cycles A and B , say Alice’s cycle and Bob’s cycle, on vertex set $[n]$. The adjacency of A and B can be seen as a process over the time too by the pedigree graph G_n^{AB} . At time $n + 1$, Alice and Bob insert the vertex $n + 1$ to their cycles. The pedigree graph G_{n+1}^{AB} either stays the same as G_n^{AB} or it arises from G_n^{AB} by adding the vertex $n + 1$ together with edges between $n + 1$ and vertices in $[n]$.

Adjacency of two cycles in the pedigree polytope is determined by the following condition:

Theorem 5 ([3]). *At all times $n \geq 4$, the two vertices of the Pedigree polytope for n cities corresponding to the cycles A and B with node set $[n]$ are adjacent in the Pedigree polytope, if and only if the graph G_n^{AB} is connected.*

For a fixed cycle A and a random cycle B , the pedigree graphs G_{\square}^{AB} will make a sequence of random graphs. At time n , whether the vertex n is added or not or with how many edges is attached to G_{n-1}^{AB} are random events. The necessary condition for having disconnected pedigree graph G_{\square}^{AB} is that an isolated vertex has been created in some step of this random process.

Isolated vertices . For deterministic cycle A and random cycle B , let the random variable Y count the total number of times that an isolated vertex is created in the pedigree graph G_{\square}^{AB} . In other words, $Y = \sum_{n=4}^{\infty} \mathbf{1}_{I_n}$, where I_n denotes the event that, at time n , n is added as an isolated vertex to $G_n^{A,B}$ (and $\mathbf{1}_{\square}$ is the indicator random variable of the event).

Lemma 2. *Whatever Alice does, $\mathbf{E}Y = 2$.*

To understand the importance of the lemma, consider a pedigree graph G_{n-1}^{AB} , just before Alice and Bob make their choices of cycle edges for inserting their new nodes n into it. If they make their choices in a way that n is not a vertex of the new pedigree graph G_n^{AB} , the number of connected components of it doesn’t change. If n is a vertex with incident edges, then the number of connected components can only decrease. The only way that the number of connected components of G_n can increase is if n is an isolated vertex in the new pedigree graph. Hence, Lemma 2

gives an upper bound on the expected number of connected components as well.

From Lemma 2, it is unlikely that the pedigree graph will have many components. The expected number of 2 for the number of components gives this intuition that in the random process of creating G_{\square}^{AB} , most of the time either nothing happens (no new vertex) or edges are created, ultimately reducing the number of components, so the pedigree graph is connected at the end.

This intuition is basically correct, but looking at the process more carefully shows that Alice can pick a strategy that she reduces the chance of merging components in G_{\square}^{AB} .

Theorem 4 actually states, for a random cycle B , chosen uniformly at random from all cycles on the vertex set $[n]$,

$$\min_A \mathbf{P}(\{ \text{the pedigree graph is connected} \}) = 1 - o(1), \quad (3.1)$$

where the minimum is over all cycles on $[n]$.

Adjacency game. We proved the lower bound (3.1), by describing the ‘‘adjacency game’’ between Alice and Bob. Alice’s goal is to make the graph G disconnected using a sophisticated strategy; whereas Bob makes uniformly random choices all the time, blindfolded. We proved that Alice loses with probability $1 - o(1)$. To analyze the game, we study a Markov-like Decision Process with state space $\mathbb{Z}_+ \times \mathbb{Z}_+$. The states are pairs (s, t) , where s is the number of common edges in Alice’s and Bob’s cycles, and t is the number of connected components of the current pedigree graph. We proved that ultimately this process will reach to a state with $t = 1$, i.e. connected pedigree graph, and stays there forever.

The ‘‘adjacency game’’ is as follows: At each time, Alice moves first. She determines her cycle A_n , by choosing an edge of A_{n-1} at each time n and inserting her node n into it. Then Bob moves. He determines B_n in the same way, but he will draw the edge of B_{n-1} into which his new node n is inserted uniformly at random from all edges of B_{n-1} , and his choice is independent of his earlier choices.

We say that Bob wins, if there exists an n_0 such that for all $n \geq n_0$, the pedigree graph G_n^{AB} is connected. We need Bob to win ‘‘uniformly’’, i.e., n_0 must not depend on Alice’s moves.

Using super-martingales, we proved that for large enough n_0 , Alice has to insert her new vertices to A , in a way that decreases the probability of creating isolated vertex in G_{\square}^{AB} between n_0 and $2n_0$. We proved for large enough n'_0 the number of components will not increase anymore and drops to 1 with high probability hence, G_{\square}^{AB} stays connected ever after. Therefore Bob wins.

4. FROM COMMUNICATION COMPLEXITY TO EXTENDED FORMULATION

In the first chapter we studied the basic ideas about extended formulation and how these may help to rule out an exponential number of inequalities. Having a polynomial size linear description of a polytope, it is possible to optimize over it in a polynomial time. But, what are the limitations of this method? Does every polytope admit a polynomial size extended formulation?

Communication complexity is a strong tool for proving lower bounds on different areas of computer science. Usually the basic idea is: if we have some properties of interest — say polynomial size linear description of a polytope, sub-linear space complexity for an algorithm, small query time for a data structure, etc.— then it implies small communication complexity for some known problem, using reduction. Therefore, linear bounds in communication complexity drive lower bounds to the other problems.

In this chapter we will study the connection of communication complexity and extended formulation. Also we briefly mention the exponential lower bound on the extension complexity of the TSP polytope.

4.1. Basic Model and definitions

The important sub-area of complexity theory, *communication complexity*, studies the amount of communication needed to learn or calculate a function. The concept of communication complexity was introduced by Yao [103] in 1979. In general the problem is a system has to do a task, but the information needed for doing the task is distributed among different parties. One obvious solution to the problem is to let all parties reveal their information. But if the communication is expensive, we need to minimize the amount of communication.

In fact, communication complexity is a measure for hardness of a problem when the whole input is not available and it focuses only on the exchanged information and not on the computational ability of the parties for calculation.

Here we introduce the basic model of communication which contains only two parties. We refer to [57] for exact definitions and more details on communication complexity and to [90] for general applications of communication complexity for proving lower bounds in different areas of computer science.

4.1.1. Deterministic communication complexity

In the simplified communication complexity model, the function $f : X \times Y \rightarrow Z$ is given with X , Y and Z being arbitrary finite sets. There are two players, Alice and Bob and the task is evaluating $f(x, y)$, for $x \in X$ and $y \in Y$. Alice only knows x and Bob only knows y , they communicate according to a protocol to verify $f(x, y)$. A naive protocol which always works can be: Alice sends all her input x to Bob,

then Bob determines the value of $f(x, y)$. Although sometimes it is not possible to do better, a protocol with the least communication is of interest.

The so-called deterministic communication complexity of f , $D(f)$, is the minimum number of bits communicated according to a best protocol P on the worst case input (x, y) .

Let $f : X \times Y \rightarrow Z$ be a function, then M_f is the corresponding matrix of f , such that rows are indexed by different values of X and columns are indexed by values of Y and for $(x, y) \in X \times Y$, $M_f(x, y) := f(x, y)$. In this work, we only consider the Boolean functions ($f : X \times Y \rightarrow Z = \{0, 1\}$). So the matrix M_f is Boolean.

Let us look at an easy example. Assume Alice and Bob have their inputs x and y both in $\{0, 1\}^n$ and they are asked to compute the function $\text{NEQ}(x, y)$. They have to output 1, if $x \neq y$ and 0 otherwise. It is not difficult to observe that they can not compute $\text{NEQ}(x, y)$ with communication fewer than n bits (one has to send the whole input) and therefore $D(\text{NEQ}) = n$.

4.1.2. Nondeterministic communication complexity

In the nondeterministic model, there is a third party, the prover, who sees x and y and tries to convince Alice and Bob that $f(x, y) = z$ by sending them certificates. The minimum number of bits communicated by Alice, Bob and the prover according to the best protocol and the worst input, is known as nondeterministic communication complexity.

Looking at the previous example in the nondeterministic case, assume the prover wants to convince Alice and Bob that $\text{NEQ}(x, y) = 1$. In this case, the prover can send the index of the bit in which x and y differ as the certificate. Thus $N(\text{NEQ}) = \lceil \log_2 n \rceil$.

Nondeterministic communication complexity also can be regarded as a two-party model, without the prover, when Alice and Bob can make nondeterministic decisions.

The application of communication complexity in the extended formulation of polytopes is via nondeterministic communication complexity so we are particularly interested in this model here.

Rectangle cover and nondeterministic communication complexity. As we defined in section 2.2.4, a *rectangle* is a product of $R = K \times L \subset [n] \times [n]$ (with $[n] = \{1, \dots, n\}$). Given an $n \times n$ Boolean matrix f , a rectangle R is a *1-rectangle* if $f(k, \ell) = 1$ for all $(k, \ell) \in R$ and it is a *0-rectangle* if $f(k, \ell) = 0$ for all $(k, \ell) \in R$.

Definition 3. The *nondeterministic communication complexity* of a Boolean function f is $\lceil \log_2 C(M_f) \rceil$.

The intuition behind this definition is: Let R_1, R_2, \dots, R_c be the rectangles that cover all the 1 values of f . Assume that Alice and Bob want to verify whether $f(x, y) = 1$ via communication among each other and with the prover. If $f(x, y) = 1$ there exists at least one rectangle which covers $f(x, y)$. The prover should only send the index of that rectangle as the certificate. So $N(f) = \lceil \log_2(c) \rceil$.

Remark 2. Let us look back one more time to proposition ???. In the deterministic model of communication, using the small partitioning of the matrix, M_f , we can conclude $\log_2 \text{rank}(M_f)$ lower bound. We should point out here that it can be a large gap between the “covering number” and the “partitioning number”. For example in the non-equality example, since $N(\text{NEQ}) = \lceil \log_2 n \rceil$ there exists 1-rectangle covering for M_{NEQ} of size $\lceil \log_2 n \rceil$ (it is not hard to find it), whereas the 1-rectangle partition is of size n . So the rank’s lower bound is not relevant to the nondeterministic communication complexity.

Fooling set. Calculating the nondeterministic communication complexity (equivalently the rectangle cover) is not always easy and sometimes methods for lower bounding the quantity is of interest. One of these methods is finding a large fooling set in the communication complexity function.

A *fooling set* is a subset of the domain of the function f , such that no two elements of it can lie in the same 1-rectangle. The size of the largest fooling set of matrix M is denoted by $F(M)$.

Definition 4. A fooling set of the function $f : X \times Y \rightarrow \{0, 1\}$ is the subset $F \subseteq X \times Y$ such that $f(x_i, y_i) = 1$ for all $(x_i, y_i) \in F$ and for each distinct pair of inputs (x_i, y_i) and (x_j, y_j) in F , either $(x_i, y_j) \neq 1$ or $(x_j, y_i) \neq 1$ or both.

Proposition 3. If a Boolean function f has a fooling set of size k , then $C(M_f) \geq k$. In particular $\log_2 k \leq N(f)$.

For the proof of the proposition 3, the main point is: no 1-rectangle can cover two elements of a fooling set at the same time.

The fractional cover number. We briefly review the definition of the fractional cover number which is also used as a bound for rectangle covering number of the matrix of a Boolean function. Let f be a fixed Boolean function, and let R be a random 1-rectangle of f , drawn according to a distribution π . Define

$$\gamma(\pi) := \min_{R \sim \pi} \left\{ \mathbf{P}((x, y) \in R) \mid (x, y) \in \text{supp } f \right\}.$$

The *fractional cover number* is $C^*(M_f) := \min_{\pi} 1/\gamma(\pi)$, where the minimum is taken over all distributions π on the set of 1-rectangles of f .

The following inequalities are well-known [57]. $R^1(M_f)$ denotes for the number of 1s in the largest 1-rectangle.

$$\left. \begin{array}{l} \frac{|\text{supp } M_f|}{R^1(M_f)} \\ F(M_f) \end{array} \right\} \leq C^*(M_f) \leq C(M_f) \leq (1 + \ln R^1(M_f)) C^*(M_f).$$

4.2. From nondeterministic communication complexity to extended formulation

Communication complexity is a powerful tool for proving lower bounds on different computational problems. One surprising connection is between nondetermin-

istic communication complexity and the extension complexity of polytopes. The following evolved communication problem explains this connection (see [90] for more details).

Let P to be a polytope and $FV(f, v)$ to be the non-incidence face-vertex function. It means $FV(f, v) = 1$ if and only if $v \notin f$. It is easily seen that the $M(FV(f, v))$ is the support of a slack matrix of the polytope P . The equivalent representation of Yannakakis' theorem (theorem 2) in the communication complexity terminology is as follows:

Theorem 6. [102] *If the polytope P has an extension Q of size r , then the nondeterministic communication complexity of $FV(f, v)$ of P is at most $\log_2 r$.*

Sketch of the proof. Let S be the face-vertex slack matrix of P , rows are indexed by all the faces and columns are indexed by the vertices of P . Assume that S admits a non-negative factorization of size r (as we mentioned in the remark 1, all the slack matrices of a polytope have the same non-negative rank). Let F be the number of faces of the polytope P and V be the number of vertices of it and the factorization $S = XY$, such that $X_{F \times r}$ and $Y_{r \times V}$ are positive matrices known to Alice and Bob.

If $S_{fv} > 0$, it means $\sum_{j=1}^r X_{fj} \cdot Y_{jv} > 0$. Since all the entries are non-negative, there exists at least one index $k \in \{1, \dots, r\}$ such that $X_{fk} \cdot Y_{kv} > 0$. It is sufficient if the prover sends this index k to Alice and Bob as the certificate. Alice and Bob accept under the condition that their corresponding entries are positive and the communication complexity is at most $\log_2 r$. □

Theorem 6 reduces the problem of proving a lower bound for extension complexity of the polytope P to proving a lower bound on nondeterministic communication complexity of $FV(f, v)$.

4.2.1. Lower bounds on extension complexity of TSP polytopes

As it was explained earlier in 2.2.6, the first lower bound on the size of extended formulation of TSP was given by Yannakakis [102]. The exponential lower bound was given only for symmetric formulation (see also [42, 43]). The problem was left open in general case for 20 years. In 2012, Fiorini, et al. [28] showed that the exponential lower bound holds also in the non-symmetric case. Their technique is to use a new connection between semidefinite programming reformulations of LPs and a special case of communication complexity, in addition to reducing set disjointness. They first proved the lower bound of $2^{\Omega(\sqrt{n})}$ and then after Rothvoß result on the lower bound of the extension complexity of the perfect matching polytope, their lower bound was improved as below. We will not provide the proof here and refer to the paper [28] for the details and more results on other polytopes.

Theorem 7. [28] For every n , $\text{xc}(P_{TSP}) = 2^{\Omega(n)}$.

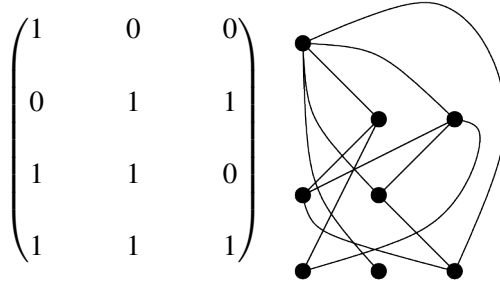
4.3. Connection to rectangle graph

The (Lovász-Saks) rectangle graph, [61] $G_{\boxtimes}(M)$ of the matrix M has as its vertices the 1-entries of M , with two 1-entries being adjacent, if they span a 2×2 rectangle containing a 0-entry of M . More precisely:

$$V(G_{\boxtimes}(M)) = \left\{ (k, \ell) \in [n] \times [n] \mid M_{k, \ell} = 1 \right\}$$

$$E(G_{\boxtimes}(M)) = \left\{ \{(k, \ell), (k', \ell')\} \mid M_{k, \ell} M_{k', \ell'} = 1 \ \& \ M_{k, \ell'} M_{k', \ell} = 0 \right\}.$$

Example 1. Here is an example of a matrix and its corresponding rectangle graph.



The rectangle graph connects communication complexity to graph theory in a satisfactory way.

The inclusion-wise maximal 1-rectangles of M are precisely the vertex sets of the inclusion-wise maximal independent sets of $G_{\boxtimes}(M)$ because if we look at a maximal 1-rectangle and corresponding vertices in the rectangle graph, no two vertices can be adjacent. In particular, $\alpha(G_{\boxtimes}(M))$ —the size of largest independent set in the graph— is equal to the size of the largest 1-rectangle of M . Also, as a consequence, $\chi(G_{\boxtimes}(M)) = C(M)$. $\chi(G)$ stands for the chromatic number of G , the minimum number of colors needed for coloring the vertices of G such that no two adjacent vertices are in the same color.

Each maximal clique of the rectangle graph, also coincides with the maximal fooling set in the matrix and thus $\omega(G_{\boxtimes}(M)) = F(M)$, where $\omega(G)$ stands for the size of the maximum clique in G .

This relationship is mainly important because it eventuates the equivalence of two main conjectures, one in communication complexity and the other in graph coloring.

Conjecture 1. [61] There exists a constant c such that for every Boolean function f ,

$$D(f) \leq (\log \text{rank}(M_f))^c.$$

Conjecture 2. [26, 99] For every graph G ,

$$\chi(G) \leq \exp \log^{O(1)}(\text{rank}(M_f)).$$

The equivalence of these two conjectures was expressed by Lovász and Saks [61] who showed that every graph G is the induced subgraph of $G_{\mathbb{J}}(\mathbb{J} - \text{Adj}(G))$ where \mathbb{J} is the all one matrix. We are not concerned with these conjectures and they are only mentioned here for their importance to computer science.

5. NONDETERMINISTIC COMMUNICATION COMPLEXITY OF RANDOM BOOLEAN FUNCTIONS

In this chapter, we discuss the motivation and results of the paper [84], reprinted as the Appendix A of this thesis.

5.1. Motivation and previous works

The application of communication complexity grows every day in different fields of computer science. The communication complexity model is so easy and clean that can be applied to very diverse areas of computer science by providing lower bounds on algorithms. In addition to combinatorial optimization which is discussed in this chapter, it has been used for finding the lower bounds for sublinear algorithms, compressive sending [20], space-time trade-offs in data structures [69, 82], algorithmic game theory etc. [28, 89, 90].

Studying the communication complexity of random functions is particularly interesting because it tells about the existence of hard functions in a model.

By random function we mean $f : X \times Y \rightarrow \{0, 1\}$, with $|X| = |Y| = n$. We take $f(x, y)$ to be independent Bernoulli random variables with $f(x, y) = 1$, with probability p and $f(x, y) = 0$, with probability $1 - p$ when $p = p(n)$.

Most research on the communication complexity of random functions and their properties, focus on the case that the probability p is constant and especially with $p = 1/2$ [18]. Whereas in the application, most of the time we are dealing with the functions with high density. For instance in combinatorial optimization, when we consider the slack matrix of a polytope, the 0's in a row correspond to the vertices incident to the face and the number of them is very often a polylog of n . It justifies the study of random functions with high density ($p = p(n) \rightarrow 1$ when $n \rightarrow \infty$). Similar to the random graphs field, this probability function makes the study more challenging and more interesting.

Properties of random matrices with Bernoulli distribution of entries has been considered previously. Izhakian, Janson and Rhodes [39] have studied the asymptotic behavior of the triangular rank of random Boolean matrices. The triangular rank is itself important in communication complexity, and is a lower bound to the size of a fooling set. But the case $p \rightarrow 0, 1$ is left as an open question in their paper.

The size of the largest monochromatic rectangle in a random Bernoulli matrix was determined in [80] when p is bounded away from 0 and 1, but their technique fails for $p \rightarrow 1$.

The nondeterministic communication complexity of the clique-vs-stable set problem on random graphs was studied in [11].

Apart from the communication complexity, the studied parameters in the papers coincide to other parameters in other areas. In combinatorics, the rectangle

covering number is equivalent to the strong isometric dimension of graphs [33], and has connections to extremal set theory and coding theory [36, 37]. The size of the largest monochromatic rectangle has application in the analysis of gene expression data [80], and formal concept analysis [17].

The other important connection, as we described in section 4.3, is the rectangle graph with Lovász and Saks construction [61]. The 1-rectangles, covers and fooling sets of a function f correspond to independent sets, colorings and cliques, resp., in a graph constructed from the function. The random rectangle graph obtained from the random function f has considerable differences with the other studied random graph models. For example, in the usual random graph models (Erdős-Renyi, uniform regular), the chromatic number is within a constant factor of the independence ratio (i.e., the quotient independence number over the number of vertices), and, in particular, of the so-called fractional chromatic number (which lies between the two). However, the corresponding statement does not hold in the random graph model deduced by the random Boolean function.

5.2. Our results

We give tight upper and lower bounds for the nondeterministic communication complexity and its most important lower bounds: the fooling set bound; the ratio number of 1-entries over largest 1-rectangle; the fractional cover number. In this chapter the function f often refers to the matrix M_F .

We study the case when probability of a family of events E_n , $n \in \mathbb{N}$ tends to 1 as $n \rightarrow \infty$, i.e. $\lim_{n \rightarrow \infty} \Pr(E_n) = 1$. As is customary, we use the terminology “asymptotically almost surely, a.a.s.” to stand for “with probability tending to 1 as n tends to infinity”.

Largest 1-rectangle

The size of the largest monochromatic rectangle in a matrix with independent (Bernoulli) entries, has been studied due to its application in bioinformatics [58, 59] and the shape of the 1-rectangles was conjectured. The conjecture was proven by Park and Szpankowski [80].

For the random Boolean function $f: X \times Y \rightarrow \{0, 1\}$ with parameter p , they proved if $\Omega(1) = p \leq 1/e$, then, a.a.s. the largest 1-rectangle consists of the 1-entries in a single row or column and if $p \geq 1/e$ but not close to 1, then with $a := \operatorname{argmax}_{b \in \{1, 2, 3, \dots\}} bp^b$, the largest 1-rectangle has a rows and $p^a n$ columns, or vice-versa.

We extended the theorem in [80] for the case that p tends to 0 or 1 quickly.

For $K \subseteq X$, we say the *1-rectangle of f generated by K* is $R := K \times L$ with $L := \left\{ y \in Y \mid \forall x \in K: f(x, y) = 1 \right\}$. The 1-rectangle generated by a subset L of Y is defined similarly.

Theorem 8. Let $f: X \times Y \rightarrow \{0, 1\}$ be a random Boolean function with parameter $p = p(n)$.

(a) If $5/n \leq p \leq 1/e$, then a.a.s., the largest 1-rectangle is generated by a single row or column, and if $p \gg (\ln n)/n$, its size is $(1 + o(1))pn$.

(b) Define

$$\begin{aligned} a_- &:= \lfloor \log_{1/p} e \rfloor, \\ a_+ &:= \lceil \log_{1/p} e \rceil, \text{ and} \\ a &:= \operatorname{argmax}_{b \in \{a_-, a_+\}} bp^b = \operatorname{argmax}_{b \in \{1, 2, 3, \dots\}} bp^b. \end{aligned} \tag{5.1}$$

There exists a constant λ_0 , such that if $1/e \leq p \leq 1 - \lambda_0/n$, then, a.a.s., a largest 1-rectangle is generated by a rows or columns and its size is $(1 + o(1))ap^n$.

The existence of 1-rectangles is rather easy, but proving no larger one exists is fairly difficult. We proved the upper bounds using Chernoff concentration for different shapes of rectangles in the p range, particularly when p tends to 1 quickly.

Proof idea of case (a). We consider three types of rectangles in the matrix. First, rectangles consisting exactly one row or column. Second, rectangles extend over at least 2 rows and 2 columns and they are square. Third, rectangles extend over at least 2 rows and 2 columns and they are not square.

First we proved that there exist a row (or a column) with pn 1s, using median of binomial distribution and independence of rows. Then using Chernoff bound we proved the probability that a row exists with at least $(1 + \varepsilon)pn$ is $o(1)$. Finally by counting argument and union bound we conclude the probability of existing a 1-rectangle of type two and three tends to 0.

Proof idea of case (b). For this part, we look at the rectangles of dimension $k \times \ell$ with $k \leq \ell$. Using union bound, we prove every 1-rectangle must have $k \leq n/\lambda^{2/3}$ when λ is defined through $p = 1 - \lambda/n$. Again using union bound and Chernoff concentration, we establish the upper bound.

Fooling sets

An obvious lower bound to the fooling set size is the *triangular rank*, i.e., the size of the largest triangular submatrix, after permuting rows and columns. Triangular rank has been studied in [39] for random matrices with independent Bernoulli entries with constant parameter p . But the case when $p \rightarrow 0$ or 1 is left as open problem.

In our paper, we studied upper and lower bounds on the size of fooling set pattern contained in a random matrix for different range of p . In particular, we studied the cases $p \rightarrow 0$ or 1.

We obtained the upper bounds by first and second moment calculations on random variable $X_{n,p,r}$ which is the number of fooling set patterns of size r contained in a random Boolean matrix of size n and parameter p .

For the lower bounds, we used some results from random graph theory. First of all, having a random Boolean function $f : X \times Y \rightarrow \{0, 1\}$, consider the bipartite graph H_f whose vertex set is the disjoint union of X and Y , and with $E(H_f) = \text{supp } f$. For random f , this graph is an *Erdős-Renyi random bipartite graph*: each edge is picked independently with probability p .

If $F \subseteq X \times Y$ is a fooling set, then F is a matching in the corresponding bipartite graph H . i.e., $F \subseteq E(H)$. Also F is a *cross-free matching*, i.e., for all $(x, y), (x', y') \in F$, if $(x, y) \neq (x', y')$ then $(x, y') \notin E$ or $(x', y) \notin E$.

Denote by $\nu(H)$ the size of the largest matching in a bipartite graph H and let $\nu^\times(\cdot)$ denote the size largest cross-free matching of a bipartite graph.

Let H be a bipartite graph, and $m = \{e_1, \dots, e_r\} \subseteq E(H)$ a matching in H . Define the graph $G' = G'(H, m)$ with vertex set $V(G') = \{1, \dots, r\}$ and $\{k, \ell\} \in E(G')$ if e_k, e_ℓ induce a $K_{2,2}$ in H . Then $\nu^\times(H) \geq \alpha(G')$ holds: for any stable set A of G' , the set $\{e_j \mid j \in A\}$ is a cross-free matching in H .

Our strategy for obtaining a large cross-free matching will be this: fix a large matching m in H_f , then find a large stable set in the corresponding random graph $G'_{n,p}(m) := G'(H_f, m)$. This random graph behaves similarly to an Erdős-Renyi random graph with $|m|$ vertices and edge-probability p^2 .

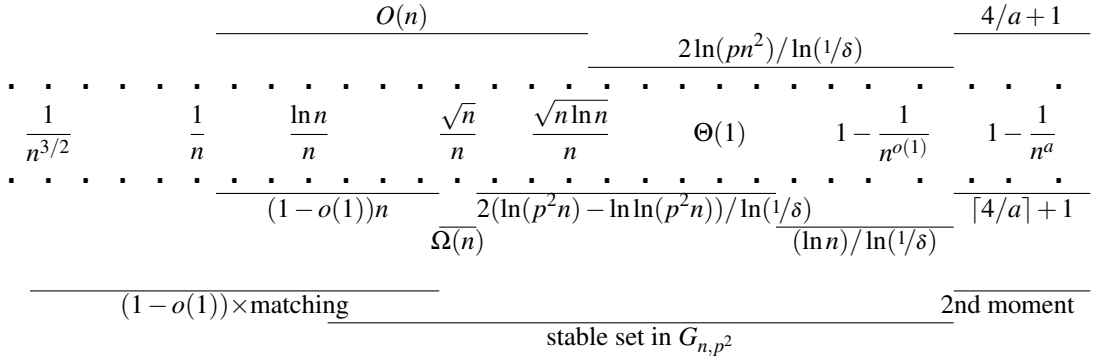


Figure 4. Upper and lower bounds on fooling set sizes. ($\delta := 1 - p^2$)

Theorem 9. Let $f : X \times Y \rightarrow \{0, 1\}$ be a random Boolean function with parameter $p = p(n)$. Define $\bar{p} := 1 - p$ and $\delta := 1 - p^2$.

(a) For $n^{-3/2} \leq p = o(1/\sqrt{n})$, a.a.s., we have

$$F(f) = (1 - o(1))\nu(H_f).$$

(b) If $pn - \ln n \rightarrow \infty$, then, a.a.s., $F(f) \geq a(p^2)$.

(c) If $p \gg \sqrt{(\ln n)/n}$ and $\bar{p} \geq n^{-o(1)}$, then, a.a.s.,

$$F(f) \leq 2 \log_{1/\delta}(pn^2).$$

(d) If $a \in]0, 4[$ is a constant and $\bar{p} = n^{-a}$, then $F(f) \leq 4/a + 1$. If, in addition, $a < 1$, then $F(f) = \lfloor 4/a \rfloor + 1$

Figure 4 summarizes the upper and lower bounds that are obtained from theorem 9 and the methods which are used.

Fractional cover number and cover number

With the results obtained from the size of the fooling set in the random function and the size of largest 1-rectangles, we can now bound the rectangle covering number and the fractional covering number.

The case $p \leq 1/2$ is easy. Let f be a random Boolean function $X \times Y \rightarrow \{0, 1\}$ with parameter p .

If $1/n \ll p \leq 1/2$, we have $C(f) = (1 - o(1))n$. For $p = o(1/\sqrt{n})$, Theorem 5.2(a) gives the lower bound based on the fooling set lower bound and for $1/e \geq p \gg (\ln n)/n$, Theorem 8(a) provides $R^1(f) = (1 + o(1))pn$, a.a.s. and for $1/e \leq p \leq 1/2$, the value of a in Theorem 8(b) is 1, so that $R^1(f) = (1 + o(1))pn$ there, too. We conclude that, a.a.s.,

$$C(f) \geq \frac{|\text{supp } f|}{R^1(f)} = \frac{(1 - o(1))pn^2}{(1 - o(1))pn} = (1 - o(1))n.$$

The case $p > 1/2$ is more challenging in techniques and more interesting in the applications. Define $\bar{p} := 1 - p$, and $\lambda := \bar{p}n$.

Recall again the following inequalities:

$$\left. \begin{array}{l} \frac{|\text{supp } f|}{R^1(f)} \\ F(f) \end{array} \right\} \leq C^*(f) \leq C(f) \stackrel{(*)}{\leq} (1 + \ln R^1(f)) C^*(f). \quad (5.3)$$

Lower bound. Using theorem 8(b), we can establish a lower bound on $C^*(f)$. With $\lambda/n = \bar{p} = 1 - p$, we have a.a.s.,

$$\frac{|\text{supp } f|}{R^1(f)} \geq \frac{(1 + o(1))pn^2}{(1 + o(1))n/e \ln(1/p)} = (1 + o(1)) e p \ln(1/p) n \geq (1 - o(1)) e p \lambda \quad (5.4)$$

For $\bar{p} = o(1)$, this is asymptotic to $e\lambda$.

Upper bound. Then we gave upper bounds on $C^*(f)$ on the fractional covering number by defining suitable distribution on 1-rectangles of f .

Theorem 10. Let $1/2 > p = 1 - \bar{p} = 1 - \lambda/n$.

- (a) If $\ln n \ll \lambda < n/2$, then, a.a.s., $(1 - o(1)) e \lambda \leq C^*(f) \leq (1 + o(1)) e \lambda$
- (b) If $\lambda = \Theta(\ln n)$, then, a.a.s., $C^*(f) = \Theta(\ln n)$.
- (c) If $1 \ll \lambda = o(\ln n)$, then, a.a.s.,

$$(1 - o(1)) \lambda \leq C^*(f) \leq (1 + o(1)) e \max\left(2\lambda, \frac{\ln n}{\ln((\ln n)/\lambda)}\right)$$

Finally, using inequality (*) in (5.3), we prove the following upper bounds on the rectangle covering number.

Corollary 1. *We have $(1 - o(1))\lambda \leq C(f)$, and:*

(a) *if $\ln n \ll \lambda = O(n/\ln n)$, then, a.a.s., $C(f) = O(\lambda \ln n)$;*

(b) *if $\lambda = \Theta(\ln n)$, then, a.a.s., $C(f) = O(\ln^2 n)$;*

(c) *if $1 \ll \lambda = o(\ln n)$, then, a.a.s., $C(f) = O\left(\max\left(\lambda \ln n, \frac{\ln^2 n}{\ln((\ln n)/\lambda)}\right)\right)$.*

□

6. THE (MINIMUM) RANK OF TYPICAL FOOLING SET MATRICES

In this chapter, we discuss the motivation and results of the paper [83], reprinted in the Publication part of this thesis.

6.1. Motivation and previous works

In the section 5.2, we introduced fooling set as a lower bound for the rectangle covering number of a matrix. So, one may be interested to know whether a large fooling set is contained in a given matrix. However, identifying the large fooling sets in a matrix, according to Lovász-Saks rectangle graph construction, is equivalent to detecting large clicks in a graph. Deciding whether a matrix contains a fooling set of a certain size is known to be NP-hard [94]. Therefore, finding an upper bound on the size of the fooling set in terms of some computable properties of the underlying matrix would be interesting.

A fooling set is known under different names in other fields of computer science and mathematics. It is usually used to provide the lower bound on some desired factors. Due to the diverse application of fooling set type lower bounds in different areas, knowing a priori upper bound on the size of the fooling set can show the usefulness of the method in advance.

In polytope theory and combinatorial optimization, the fooling set gives a lower bound on the extension complexity of a polytope and the minimum size of a linear program [27, 102].

In computational complexity, as we discussed before, the logarithm of the size of a fooling set induces a lower bound on the communication complexity of a function [54, 57, 60].

In graph theory, considering the matrix as the adjacency matrix of a bipartite graph, the fooling set corresponds to a cross-free matching, which provides a lower bound on the size of the biclique covering of a graph. A matching is cross-free if no two matching edges induce a C_4 (cycle of size 4) subgraph [16, 40].

Let $f : X \times Y \rightarrow \{0, 1\}$ be a function with the fooling set $(x_1, y_1), \dots, (x_n, y_n) \in X \times Y$. A straightforward upper bound on the size of fooling sets due to Dietzfelinger et al. [18], states that no fooling set in a function f is larger than the square of $\min_A \text{rank}_{\mathbb{F}} A$, when the minimum ranges over all $X \times Y$ matrices A over the field \mathbb{F} with $A_{x,y} = 0$ if and only if $f(x, y) = 0$.

The proof follows from the fact that if $B := A \circ A^T$ then

$$n \leq \text{rank } B \leq \text{rank } A^2,$$

where \circ is the entry-wise (Hadamard) product.

In general the fooling set can be defined on the matrices with entries in the arbitrary field \mathbb{F} (not Boolean entries) and actually in some applications it is important

to consider the matrix in a general field [54].

For fields \mathbb{F} with non-zero characteristics, the upper bound is asymptotically tight [31], and for general fields, it is attained up to a multiplicative constant [30].

6.2. Our results

A natural question would be is the rank of a typical fooling set of size n , closer to its trivial upper bound n or its lower bound \sqrt{n} ? In this paper we tried to answer this question. The question turns out to be surprisingly difficult.

A *fooling-set pattern of size n* is a matrix R with entries in $\{0, 1\} \subseteq \mathbb{F}$ with $R_{k,k} = 1$ for all k and $R_{k,\ell}R_{\ell,k} = 0$ whenever $k \neq \ell$. We say that a fooling-set pattern of size n has *density* $p \in]0, 1]$, if it has exactly $\lceil p \binom{n}{2} \rceil$ off-diagonal 1-entries.

In our paper, we gave partial results for a fooling-set pattern chosen at random according to a sensible distribution. We studied the following distributions:

$Q(n)$ denotes a fooling set pattern drawn uniformly at random from all fooling set patterns of size n ;

$R(n, p)$ denotes a fooling set pattern drawn uniformly at random from all fooling set patterns of size n with density p .

We allow that the density depends on the size of the matrix: $p = p(n)$. As is customary, we use the terminology “asymptotically almost surely, a.a.s.,” to stand for “with probability tending to 1 as n tends to infinity”. $\sigma(M)$ is the *zero-nonzero pattern* of M , that is $\sigma(M(i, j)) = 0$ if and only if $M(i, j) = 0$ and $\sigma(M(i, j)) = 1$ otherwise.

Theorem 11. (a) For every field \mathbb{F} , if $p = O(1/n)$, then, a.a.s., the minimum rank of a matrix with zero-nonzero pattern $R(n, p)$ is $\Omega(n)$.

(b) Let \mathbb{F} be a finite field and $F := |\mathbb{F}|$. (We allow F to grow with n .) If $100 \max(1, \ln \ln F)/n \leq p \leq 1$, then the minimum rank of a matrix over \mathbb{F} with zero-nonzero pattern $R(n, p)$ is

$$\Omega\left(\frac{\log(1/p)}{\log(1/p) + \log(F)} n\right) = \Omega(n/\log(F)).$$

(c) For every field \mathbb{F} , if $p \in]0, 1]$ is a constant, then the minimum rank of a matrix with zero-nonzero pattern $R(n, p)$ is $\Omega(n)$. (The same is true for zero-nonzero pattern $Q(n)$.)

We used different techniques in the proofs of the three parts of the theorem and here we give a brief sketch for each part. For the details we refer to the paper.

In part (a), we study the case that p tends to 0 quickly enough. The idea of the proof comes from the results in random graph theory. We construct a graph G with vertex set $[n]$ from $R(n, p)$ and there is an edge between vertices k and ℓ with $k > \ell$, if and only if $M_{k,\ell} \neq 0$. This construction gives the random graph $G_{n,m,1/2}$. Using Theorem 1.4 in [32], this random graph behaves similarly to the

Erdős-Rényi graph with $q := p/2$. Since $G_{n,p/2}$ has an independent set of size $\Omega(n)$, $G_{n,m,1/2}$ will also have. Independent sets in G are just the lower-triangular submatrices of $R(n, p)$. Therefore in this case rank of the fooling set pattern is lower bounded by $\Omega(n)$.

In part (b), we study the rank of a fooling set pattern $R(n, p)$ when p tends to 0 slowly, and $|\mathbb{F}| = O(1)$. Consider the event:

“There is a matrix M over \mathbb{F} with $\sigma(M) = R(n, p)$, and $\text{rk } M \leq r := n/(2000 \ln F)$.”

We calculate the probability of this event in the case:

1. M contains a *dense* sub-structure (*tee-matrix*) of rank $\text{rk } M$.
2. M contains a *sparse* sub-structure (*tee-matrix*) of rank $\text{rk } M$.

The sub-structure, tee-matrix, is dense if the Hamming weight of its support is of size at least $15pr(n - r)$ and it is sparse otherwise.

In the first case, using Chernoff-like bound, the probability of the event is at most $e^{-\Omega(r)}$ and in the second case, with counting argument in some steps, we prove that the probability of this event tends to 0. Ultimately the rank of every fooling set pattern with the density in the defined range lower bounds by $\Omega(n/\log(F))$.

In part (c), we study the rank of $R(n, p)$ when $p \in]0, 1]$ is a constant. In this case, we apply a theorem of Ronyai, Babai, and Ganapathy [87] on the maximum number of zero-patterns of a family of polynomials for upper bounding the number of all fooling set pattern matrices of rank at most $r := \rho n$, when $\rho < 1/2$.

Theorem 12 (Ronyai-Babai-Ganapathy-2001). *If f is an k -tuple of polynomials in n variables over a field \mathbb{F} with $k \geq n$ and each f_j has degree at most d then, for all m*

$$\left| \left\{ y \in \{0, 1\}^k \mid |y| \leq m \text{ and } y = \sigma(f(u)) \text{ for some } u \in \mathbb{F}^n \right\} \right| \leq \binom{n + md}{n}.$$

In other words, the number of zero-nonzero patterns with Hamming weight at most m is at most $\binom{n + md}{n}$.

Let M be a fooling set of size n and rank at most r . We factorize M to $M = XY$ with $X \in \mathbb{R}^{n \times r}$ and $Y \in \mathbb{R}^{r \times n}$ in a way that for Y :

There are three types of entries in Y : 0, 1 or * (no restriction). In every column there exists at most one 1 and in a column with 1, there is no *. First nonzero in every row is 1 and there is at most one 1 in a row. All-zero rows, at the bottom.

This factorization is always possible using Gaussian elimination row operations. For details we refer to the paper.

Having this special factorization for M , next, we assign the variables $X_{i,j}$ to all entries of X and assign $Y_{i,j}$, to all entries of Y which are *-type. In the next step, using theorem 12, we upper-bound the number of zero-nonzero patterns of the polynomials obtained in M . The number of variables are the total number of $X_{i,j}$'s and $Y_{i,j}$'s. Every polynomial assigned to the entries of M has degree at

most 2 and the Hamming weight of M is at most $m := p \binom{n}{2}$. Finally, with upper bounding the number of possible different Y 's in the factorization, we conclude the probability that a fooling set matrix with zero-nonzero pattern $R(n, p)$ has rank at most r is at most

$$\frac{\binom{n}{r} \binom{2rn - r^2/2 + 2m}{2rn - r^2/2}}{\binom{\binom{n}{2}}{m} 2^m} \rightarrow 0 \text{ (With } \rho < 1/2 \text{ and } r := \rho n). \quad (6.1)$$

The proof for the distribution $Q(n)$ is easier. In fact for the uniform distribution, the total number of fooling set patterns is $3^{\binom{n}{2}}$ and it will be replaced in the denominator of 6.1.

The bound in (b) does not give an $\Omega(n)$ lower bound for infinite fields, or for large finite fields, e.g., $\text{GF}(2^n)$. We conjecture that the bound is still true:

Conjecture 3. For every field \mathbb{F} and for all $p = p(n)$, the minimum rank of a fooling-set matrix with random zero-nonzero pattern $R(n, p)$ is $\Omega(n)$.

7. CONCLUSION

The linear program method plays an important role in solving optimization problems. In this setting, studying the methods for improving the running time of the LP is crucial and in each problem studying the properties of the feasible space in geometric point of view, polyhedral theory, seems inevitable.

Extended formulation is a method for describing the connection between geometric representation of the feasible space and the amount of the information it contains. In other words, extension of a polytope is, in fact, a compressed way of representing a polytope. That is, it provides the connection between linear optimization and communication complexity.

Specifically, nondeterministic communication complexity is a powerful tool for proving lower bounds on the extension complexity of polytopes. Finding a suitable communication complexity problem corresponding to a polytope P and proving a linear lower bound for the nondeterministic communication complexity of it will rule out all the attempts for finding sub-exponential size extension Q of P .

Studying the nondeterministic communication complexity and the parameters related to it, as a rectangle covering number and the fooling set size, is particularly important for showing the limitations of finding polynomial size extension for a polytope associated to an optimization problem. Of these, the communication complexity of random functions is more interesting because it tells about the existence of hard functions in a model. Most studies on the communication complexity of random functions and their properties, focus on the case that the probability p is constant [18]. Whereas in the application, very often we are dealing with functions with high density. For instance in combinatorial optimization, when we consider the slack matrix of a polytope, the 0's in a row correspond to the vertices incident to the face and the number of them is very often the polylog of n . It justifies the study of random functions with high density ($p = p(n) \rightarrow 1$ when $n \rightarrow \infty$). Similar to the random graphs field, this probability function makes the study more challenging and more interesting.

In our paper [84], we have focused on the random Boolean functions $f : X \times Y \rightarrow \{0, 1\}$, with $|X| = |Y| = n$ and density $p = p(n)$. We gave tight upper and lower bounds for the nondeterministic communication complexity and its important lower bounds: the fooling set bound, the ratio number of 1-entries over largest 1-rectangle and the fractional cover number. The parameters we study are of importance beyond Communication Complexity and its direct applications.

The fooling set is an important lower bound for the rectangle covering number of a matrix. While, one may be interested to know whether a large fooling set is contained in a given matrix, identifying the large fooling sets in a matrix is a hard problem in itself. Therefore finding a priori upper bound on the size of the fooling set in terms of some computable properties of the underlying matrix, such as rank of the matrix, would be valuable. Regarding the *typical* minimum

rank of a fooling-set matrix, we asked: Is the minimum rank of a matrix with that zero-nonzero pattern over a field \mathbb{F} closer to its lower bound \sqrt{n} or to its upper bound n ? We studied random patterns with a given density p , and proved an $\Omega(n)$ bound for some. We have to leave open the case when $p \rightarrow 0$ slowly and \mathbb{F} is a large or infinite field. We conjecture that the minimum rank of a fooling set matrix with random zero-nonzero pattern drawn uniformly at random from all fooling set patterns of size n with density p is $\Omega(n)$.

Finally, we investigated the graph of the *pedigree* polytope. The pedigree polytope is an extension of TSP (traveling salesman problem; the most extensively studied problem in combinatorial optimization) polytopes with a nice combinatorial structure. The graph of a polytope can be regarded as an abstract graph and it reveals meaningful information about the properties of the polytope. We proved the minimum degree of a vertex on the Pedigree polytope for n cities is $(1 - o(1)) \cdot (n - 1)!/2$ (for $n \rightarrow \infty$).

BIBLIOGRAPHY

- [1] N Aguilera, R Katz, and P Tolomei. Vertex adjacencies in the set covering polyhedron. *arXiv preprint arXiv:1406.6015*, 2014.
- [2] D. Applegate, R. Bixby, V. Chvátal, and W. Cook. *The Traveling Salesman Problem – A Computational Study*. Princeton Series in Applied Mathematics. Princeton, 2006.
- [3] T. S. Arthanari. On pedigree polytopes and hamiltonian cycles. *Discrete Math.*, 306:1474–1792, 2006.
- [4] Tiru S. Arthanari. Study of the pedigree polytope and a sufficiency condition for nonadjacency in the tour polytope. *Discrete Optimization*, 10(3):224–232, 2013.
- [5] Tiru S Arthanari and M Usha. An alternate formulation of the symmetric traveling salesman problem and its properties. *Discrete Applied Mathematics*, 98(3):173–190, 2000.
- [6] Michel L Balinski et al. On the graph structure of convex polyhedra in n -space. *Pacific J. Math*, 11(2):431–434, 1961.
- [7] Francisco Barahona and Ali Ridha Mahjoub. On the cut polytope. *Mathematical programming*, 36(2):157–173, 1986.
- [8] David W Barnette. On steinitz’s theorem concerning convex 3-polytopes and on some properties of planar graphs. In *The many facets of graph theory*, pages 27–40. Springer, 1969.
- [9] Roswitha Blind and Peter Mani-Levitska. Puzzles and polytope isomorphisms. *Aequationes Mathematicae*, 34(2-3):287–297, 1987.
- [10] K-H Borgwardt. The average number of pivot steps required by the simplex-method is polynomial. *Mathematical Methods of Operations Research*, 26(1):157–177, 1982.
- [11] Gábor Braun, Samuel Fiorini, and Sebastian Pokutta. Average case polyhedral complexity of the maximum stable set problem. In *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques, APPROX/RANDOM 2014, September 4-6, 2014, Barcelona, Spain*, pages 515–530, 2014.
- [12] Joel E. Cohen and Uriel G. Rothblum. Nonnegative ranks, decompositions, and factorizations of nonnegative matrices. *Linear Algebra Appl.*, 190:149–168, 1993.
- [13] George Dantzig, Ray Fulkerson, and Selmer Johnson. Solution of a large-scale traveling-salesman problem. *Journal of the operations research society of America*, 2(4):393–410, 1954.
- [14] George B Dantzig. Maximization of a linear function of variables subject to linear inequalities, in activity analysis of production and allocation. 1951.

- [15] George Bernard Dantzig. *Linear programming and extensions*. Princeton university press, 1998.
- [16] Milind Dawande. A notion of cross-perfect bipartite graphs. *Inform. Process. Lett.*, 88(4):143–147, 2003.
- [17] Milind Dawande, Pinar Keskinocak, Jayashankar M. Swaminathan, and Sridhar Tayur. On bipartite and multipartite clique problems. *J. Algorithms*, 41(2):388–403, November 2001.
- [18] Martin Dietzfelbinger, Juraj Hromkovič, and Georg Schnitger. A comparison of two lower-bound methods for communication complexity. *Theoret. Comput. Sci.*, 168(1):39–51, 1996. 19th International Symposium on Mathematical Foundations of Computer Science (Košice, 1994).
- [19] Chris HQ Ding, Xiaofeng He, and Horst D Simon. On the equivalence of nonnegative matrix factorization and spectral clustering. In *SDM*, volume 5, pages 606–610. SIAM, 2005.
- [20] David L Donoho. Compressed sensing. *IEEE Transactions on information theory*, 52(4):1289–1306, 2006.
- [21] J. Edmonds. Maximum matching and a polyhedron with 0-1 vertices. *J. Res. Nat. Bur. Standards*, 69B:125–130, 1965.
- [22] Jack Edmonds. Matroids and the greedy algorithm. *Mathematical programming*, 1(1):127–136, 1971.
- [23] Friedrich Eisenbrand, Nicolai Hähnle, and Thomas Rothvoß. Diameter of polyhedra: limits of abstraction. In *Proceedings of the twenty-fifth annual symposium on Computational geometry*, pages 386–392. ACM, 2009.
- [24] Andreas T Ernst, Houyuan Jiang, Mohan Krishnamoorthy, and David Sier. Staff scheduling and rostering: A review of applications, methods and models. *European journal of operational research*, 153(1):3–27, 2004.
- [25] William Joshua Espenschied. Graphs of polytopes. 2014.
- [26] Siemion Fajtlowicz. On conjectures of graffiti. *Discrete mathematics*, 72(1-3):113–118, 1988.
- [27] Samuel Fiorini, Volker Kaibel, Kanstantin Pashkovich, and Dirk Oliver Theis. Combinatorial bounds on nonnegative rank and extended formulations. *Discrete Math.*, 313(1):67–83, 2013.
- [28] Samuel Fiorini, Serge Massar, Sebastian Pokutta, Hans Raj Tiwary, and Ronald De Wolf. Exponential lower bounds for polytopes in combinatorial optimization. *Journal of the ACM (JACM)*, 62(2):17, 2015.
- [29] Eric J Friedman. Finding a simple polytope from its graph in polynomial time. *Discrete & Computational Geometry*, 41(2):249–256, 2009.
- [30] Mirjam Friesen, Aya Hamed, Troy Lee, and Dirk Oliver Theis. Fooling-sets and rank. *European Journal of Combinatorics*, 48:143–153, 2015.
- [31] Mirjam Friesen and Dirk Oliver Theis. Fooling-sets and rank in nonzero characteristic. In Jaroslav Nešetřil and Marco Pellegrini, editors, *The Sev-*

- enth European Conference on Combinatorics, Graph Theory and Applications, volume 16 of CRM series, pages 383–390. CRM, 2013.
- [32] Alan Frieze and Michał Karoński. *Introduction to random graphs*. Cambridge University Press, 2015.
- [33] Dalibor Fronček, Janja Jerebic, Sandi Klavzar, and Petr Kovár. Strong isometric dimension, biclique coverings, and sperner’s theorem. *Combinatorics, Probability & Computing*, 16(2):271–275, 2007.
- [34] Carlo Iapige De Gaetani, Noemi Emanuela Cazzaniga, Riccardo Barzagli, Mirko Reguzzoni, and Barbara Betti. Covariance function modelling in local geodetic applications using the simplex method. *Boletim de Ciências Geodésicas*, 22(2):342–357, 2016.
- [35] Martin Grötschel and Manfred W. Padberg. Polyhedral theory. In Eugene L. Lawler, Jan Karel Lenstra, A. H. G. Rinnooy Kan, and David B. Shmoys, editors, *The Traveling Salesman Problem. A Guided Tour of Combinatorial Optimization*, chapter 8, pages 251–306. Wiley, 1985.
- [36] Hossein Hajiabolhassan and Farokhlagha Moazami. Secure frameproof code through biclique cover. *Discrete Mathematics & Theoretical Computer Science*, 14(2):261–270, 2012.
- [37] Hossein Hajiabolhassan and Farokhlagha Moazami. Some new bounds for cover-free families through biclique covers. *Discrete Mathematics*, 312(24):3626–3635, 2012.
- [38] AJ Hoffman, J Wolfe, RS Garfinkel, DS Johnson, CH Papadimitriou, PC Gilmore, EL Lawler, DB Shmoys, RM Karp, JM Steele, et al. *The traveling salesman problem: a guided tour of combinatorial optimization*. J. Wiley & Sons, 1986.
- [39] Zur Izhakian, Svante Janson, and John Rhodes. Superboolean rank and the size of the largest triangular submatrix of a random matrix. *Proceedings of the American Mathematical Society*, 143(1):407–418, 2015.
- [40] S. Jukna and A. S. Kulikov. On covering graphs by complete bipartite subgraphs. *Discrete Math.*, 309(10):3399–3403, 2009.
- [41] Volker Kaibel. Low-dimensional faces of random 0/1-polytopes. In *International Conference on Integer Programming and Combinatorial Optimization*, pages 401–415. Springer, 2004.
- [42] Volker Kaibel, Kanstantsin Pashkovich, and Dirk Oliver Theis. Symmetry matters for the sizes of extended formulations. In *Integer programming and combinatorial optimization*, volume 6080 of *Lecture Notes in Comput. Sci.*, pages 135–148. Springer, Berlin, 2010.
- [43] Volker Kaibel, Kanstantsin Pashkovich, and Dirk Oliver Theis. Symmetry matters for the sizes of extended formulations. *SIAM J. Discrete Math.*, to appear.
- [44] Volker Kaibel and Marc E Pfetsch. Some algorithmic problems in polytope

- theory. In *Algebra, geometry and software systems*, pages 23–47. Springer, 2003.
- [45] Volker Kaibel and Anja Remshagen. On the graph-density of random 0/1-polytopes. In *Approximation, Randomization, and Combinatorial Optimization: Algorithms and Techniques, 6th International Workshop on Approximation Algorithms for Combinatorial Optimization Problems, APPROX 2003 and 7th International Workshop on Randomization and Approximation Techniques in Computer Science, RANDOM 2003, Princeton, NJ, USA, August 24–26, 2003, Proceedings*, pages 318–328, 2003.
- [46] Gil Kalai. A simple way to tell a simple polytope from its graph. *Journal of combinatorial theory, Series A*, 49(2):381–383, 1988.
- [47] Gil Kalai. A subexponential randomized simplex algorithm. In *Proceedings of the twenty-fourth annual ACM symposium on Theory of computing*, pages 475–482. ACM, 1992.
- [48] Gil Kalai. Linear programming, the simplex algorithm and simple polytopes. *Mathematical Programming*, 79(1-3):217–233, 1997.
- [49] Gil Kalai. 20 poly tope skeletons and paths. *Handbook of discrete and computational geometry*, page 455, 2004.
- [50] Gil Kalai and Daniel J Kleitman. A quasi-polynomial bound for diameter of graphs of polyhedra. *Bulletin of the American Mathematical Society*, 1992.
- [51] Narendra Karmarkar. A new polynomial-time algorithm for linear programming. In *Proceedings of the sixteenth annual ACM symposium on Theory of computing*, pages 302–311. ACM, 1984.
- [52] Leonid G Khachiyan. Polynomial algorithms in linear programming. *USSR Computational Mathematics and Mathematical Physics*, 20(1):53–72, 1980.
- [53] Edward D Kim. Polyhedral graph abstractions and an approach to the linear hirsch conjecture. *Mathematical Programming*, 143(1-2):357–370, 2014.
- [54] Hartmut Klauck and Ronald de Wolf. Fooling one-sided quantum protocols. arXiv:1204.4619, 2012.
- [55] Hartmut Klauck, Troy Lee, Dirk Oliver Theis, and Rekha R Thomas. Limitations of convex programming: lower bounds on extended formulations and factorization ranks (dagstuhl seminar 15082). *Dagstuhl Reports*, 5(2):109–127, 2015.
- [56] Victor Klee and George J Minty. How good is the simplex algorithm. Technical report, DTIC Document, 1970.
- [57] Eyal Kushilevitz and Noam Nisan. *Communication complexity*. Cambridge University Press, Cambridge, 1997.
- [58] Stefano Lonardi, Wojciech Szpankowski, and Qiaofeng Yang. Finding bi-

- clusters by random projections. In *Combinatorial Pattern Matching*, pages 102–116. Springer, 2004.
- [59] Stefano Lonardi, Wojciech Szpankowski, and Qiaofeng Yang. Finding biclusters by random projections. *Theoretical Computer Science*, 368(3):217–230, 2006.
- [60] L. Lovász and M. Saks. Möbius functions and communication complexity. In *Proc. 29th IEEE FOCS*, pages 81–90. IEEE, 1988.
- [61] L. Lovász and M. Saks. Communication complexity and combinatorial lattice theory. *Journal of Computer and System Sciences*, 47:322–349, 1993.
- [62] Laszlo Lovász. Communication complexity: a survey. In B. Korte, L. Lovász, H.J. Prömel, and A. Schrijver, editors, *Paths, flows, and VLSI-Layout*, pages 235–265. Springer, 1990.
- [63] Abdullah Makkeh, Mozhgan Pourmoradnasseri, and Dirk Oliver Theis. On the graph of the pedigree polytope. *arXiv preprint arXiv:1611.08431*, 2016.
- [64] Abdullah Makkeh, Mozhgan Pourmoradnasseri, and Dirk Oliver Theis. *The Graph of the Pedigree Polytope is Asymptotically Almost Complete (Extended Abstract)*, pages 294–307. Springer International Publishing, Cham, 2017.
- [65] A Maksimenko. The common face of some 0/1-polytopes with np-complete nonadjacency relation. *Journal of Mathematical Sciences*, 203(6):823–832, 2014.
- [66] Olvi L Mangasarian, R Setiono, and WH Wolberg. Pattern recognition via linear programming: Theory and application to medical diagnosis. *Large-scale numerical optimization*, pages 22–31, 1990.
- [67] R. K. Martin. Using separation algorithms to generate mixed integer model reformulations. *Operations Research Letters*, 10(3):119 – 128, 1991.
- [68] Jiří Matoušek. *Lectures on discrete geometry*, volume 108. Springer New York, 2002.
- [69] Peter Bro Miltersen. Cell probe complexity-a survey. In *Proceedings of the 19th Conference on the Foundations of Software Technology and Theoretical Computer Science, Advances in Data Structures Workshop*, page 2, 1999.
- [70] Hermann Minkowski. Volumen und oberfläche. In *Ausgewählte Arbeiten zur Zahlentheorie und zur Geometrie*, pages 146–192. Springer, 1989.
- [71] Vishal Monga and Mehmet Kivanç Mihçak. Robust and secure image hashing via non-negative matrix factorizations. *IEEE Transactions on Information Forensics and Security*, 2(3-1):376–390, 2007.
- [72] Denis Naddef. Pancyclic properties of the graph of some 0–1 polyhedra. *Journal of Combinatorial Theory, Series B*, 37(1):10–26, 1984.
- [73] Denis Naddef. The hirsch conjecture is true for (0, 1)-polytopes. *Mathematical Programming*, 45(1):109–110, 1989.

- [74] Denis Naddef and Giovanni Rinaldi. The graphical relaxation: a new framework for the symmetric traveling salesman polytope. *Math. Programming*, 58(1, Ser. A):53–88, 1993.
- [75] Dennin J Naddef and William R Pulleyblank. Hamiltonicity in (0–1)-polyhedra. *Journal of Combinatorial Theory, Series B*, 37(1):41–52, 1984.
- [76] Jaroslav Ne et al. *Invitation to discrete mathematics*. Oxford University Press, 2009.
- [77] Marcus Oswald, Gerhard Reinelt, and Dirk Oliver Theis. On the graphical relaxation of the symmetric traveling salesman polytope. *Math. Program.*, 110(1, Ser. B):175–193, 2007.
- [78] Manfred W Padberg and Mendu R Rao. The travelling salesman problem and a class of polyhedra of diameter two. *Mathematical Programming*, 7(1):32–45, 1974.
- [79] Christos H Papadimitriou. The adjacency relation on the traveling salesman polytope is np-complete. *Mathematical Programming*, 14(1):312–324, 1978.
- [80] Gahyun Park and Wojciech Szpankowski. Analysis of biclusters with applications to gene expression data. In *International Conference on Analysis of Algorithms DMTCS proc. AD*, volume 267, page 274, 2005.
- [81] Kanstantsin Pashkovich and Stefan Weltge. Hidden vertices in extensions of polytopes. *Operations Research Letters*, 43(2):161–164, 2015.
- [82] Mihai Patrascu. *Lower bound techniques for data structures*. PhD thesis, Massachusetts Institute of Technology, 2008.
- [83] Mozghan Pourmoradnasseri. The (minimum) rank of typical fooling set matrices. *arXiv preprint arXiv:1608.07038*, 2016.
- [84] Mozghan Pourmoradnasseri and Dirk Oliver Theis. Nondeterministic communication complexity of random boolean functions. *arXiv preprint arXiv:1611.08400*, 2016.
- [85] Jürgen Richter-Gebert. *Realization spaces of polytopes*. Citeseer, 1996.
- [86] F. J. Rispoli and S. Cosares. A bound of 4 for the diameter of the Symmetric Traveling Salesman Polytope. *SIAM J. Discrete Math.*, 11:343–380, 1998.
- [87] Lajos Rónyai, László Babai, and Murali Ganapathy. On the number of zero-patterns of a sequence of polynomials. *Journal of the American Mathematical Society*, 14(3):717–735, 2001.
- [88] Thomas Rothvoss. The matching polytope has exponential extension complexity. In *Proceedings of the 46th Annual ACM Symposium on Theory of Computing*, STOC '14, pages 263–272, New York, NY, USA, 2014. ACM.
- [89] Tim Roughgarden. Barriers to near-optimal equilibria. In *Foundations of Computer Science (FOCS), 2014 IEEE 55th Annual Symposium on*, pages 71–80. IEEE, 2014.

- [90] Tim Roughgarden. Communication complexity (for algorithm designers). *arXiv preprint*, page arXiv:1509.06257, 2015.
- [91] Francisco Santos. A counterexample to the hirsch conjecture. *Annals of mathematics*, 176(1):383–412, 2012.
- [92] Francisco Santos. Recent progress on the combinatorial diameter of polytopes and simplicial complexes. *Top*, 21(3):426–460, 2013.
- [93] A Sarangarajan. A lower bound for adjacencies on the traveling salesman polytope. *SIAM Journal on Discrete Mathematics*, 10(3):431–435, 1997.
- [94] Yaroslav Shitov. On the complexity of Boolean matrix ranks. *Linear Algebra and Its Applications*, 439:2500–2502, 2013.
- [95] Gerard Sierksma. The skeleton of the symmetric traveling salesman polytope. *Discrete applied mathematics*, 43(1):63–74, 1993.
- [96] Gerard Sierksma and Ruud H Teunter. Partial monotoneizations of hamiltonian cycle polytopes: dimensions and diameters. *Discrete applied mathematics*, 105(1):173–182, 2000.
- [97] ER Swart. $P = NP$. Report No. CIS86-02, Department of Computer and Information Science, University of Guelph, Ontario, Canada, 1986.
- [98] Michael J Todd. The many facets of linear programming. *Mathematical Programming*, 91(3):417–436, 2002.
- [99] Cyriel van Nuffelen. A bound for the chromatic number of a graph. *The American Mathematical Monthly*, 83(4):265–266, 1976.
- [100] Stephen A. Vavasis. On the complexity of nonnegative matrix factorization. *SIAM J. Optim.*, 20(3):1364–1377, 2009.
- [101] Stephen J Wright. *Primal-dual interior-point methods*. Siam, 1997.
- [102] Mihalis Yannakakis. Expressing combinatorial optimization problems by linear programs. *J. Comput. System Sci.*, 43(3):441–466, 1991.
- [103] Andrew Chi-Chih Yao. Some complexity questions related to distributive computing (preliminary report). In *Proceedings of the eleventh annual ACM symposium on Theory of computing*, pages 209–213. ACM, 1979.
- [104] Günter M Ziegler. Lectures on polytopes. 1993.

Appendix A. NONDETERMINISTIC COMMUNICATION COMPLEXITY OF RANDOM BOOLEAN FUNCTIONS

The following is copied from [84].

Nondeterministic Communication Complexity of random Boolean functions

Mozhgan Pourmoradnasseri¹ and Dirk Oliver Theis¹

University of Tartu
Institute of Computer Science
Ülikooli 17
51014 Tartu, Estonia,
{mozhgan,dotheis}@ut.ee
WWW: <http://ac.cs.ut.ee/>

Abstract. We study nondeterministic communication complexity and related concepts (fooling sets, fractional covering number) of random functions $f: X \times Y \rightarrow \{0, 1\}$ where each value is chosen to be 1 independently with probability $p = p(n)$, $n := |X| = |Y|$.

Keywords: Communication Complexity, Random Structures

1 Introduction

Communication Complexity lower bounds have found applications in areas as diverse as sublinear algorithms, space-time trade-offs in data structures, compressive sensing, and combinatorial optimization (cf., e.g., [30,11]). In combinatorial optimization especially, there is a need to lower bound *nondeterministic* communication complexity [33,20].

Let X, Y be sets and $f: X \times Y \rightarrow \{0, 1\}$ a function. In nondeterministic communication, Alice gets an $x \in X$, Bob gets a $y \in Y$, and they both have access to a bit string supplied by a prover. In a protocol, Alice sends one bit to Bob; the decision whether to send 0 or 1 is based on her input x and the bit string z given by the prover. Then Bob decides based on his input y , the bit string z given by the prover, and the bit sent by Alice, whether to accept (output 1) or reject (output 0). The protocol is successful, if, (1) regardless of what the prover says, Bob never accepts if $f(x, y) = 0$, but (2) for every (x, y) with $f(x, y) = 1$, there is a proof z with which Bob accepts. The nondeterministic communication complexity is the smallest number ℓ of bits for which there is a successful protocol with ℓ -bit proofs.

Formally, the following basic definitions are common:

- The *support* is the set of all 1-entries: $\text{supp } f := \{(x, y) \mid f(x, y) = 1\}$;
- a *1-rectangle* is a cartesian product of sets of inputs $R = A \times B \subseteq X \times Y$ all of which are 1-entries: $A \times B \subseteq \text{supp } f$;
- a *cover* (or *1-cover*) is a set of 1-rectangles $\{R_1 = A_1 \times B_1, \dots, R_k = A_k \times B_k\}$ which together cover all 1-entries of f , i.e., $\bigcup_{j=1}^k R_j = \text{supp } f$;

- the cover number $C(f)$ of f is the smallest size of a 1-cover.

One can then define the *nondeterministic communication complexity* simply as $N(f) := \log_2 C(f)$ [23].

In combinatorial optimization, one wants to lower bound the nondeterministic communication complexity of functions which are defined based on relations between feasible points and inequality constraints of the optimization problem at hand: Alice has an inequality constraint, Bob has a feasible point, and they should reject (answer 0) if the point satisfies the inequality with equality.

Consider, the following example (it describes the so-called *permutohedron*). Let $k \geq 3$ be a positive integer.

- Let Y denote the permutations π of $[k]$ —the feasible points.
- Let X denote the set of non-empty subsets $U \subsetneq [k]$; such an U corresponds to an inequality constraint $\sum_{u \in U} \pi(u) \geq |U|(|U| + 1)/2$.

Goemans [15] gave an $\Omega(\log k)$ lower bound for the nondeterministic communication complexity of the corresponding function:

$$f(\pi, U) = \begin{cases} 0, & \text{if } \sum_{u \in U} \pi(u) = |U|(|U| + 1)/2; \\ 1, & \text{otherwise, i.e., } \sum_{u \in U} \pi(u) > |U|(|U| + 1)/2. \end{cases}$$

For $k = 3$, see the following table. The rows are indexed by the set X , the columns by the set Y .

	123	132	213	231	312	321
{1}	0	0	1	1	1	1
{2}	1	1	0	1	0	1
{3}	1	1	1	0	1	0
{1, 2}	0	1	0	1	1	1
{1, 3}	1	0	1	0	1	1
{2, 3}	1	1	1	1	0	0

In this situation, the nondeterministic communication complexity lower bounds the logarithm of the so-called *extension complexity*: the smallest number of linear inequalities which is needed to formulate the optimization problem. This relationship goes back to Yannakakis' 1991 paper [33], and has recently been the focus of renewed attention [2,22] and a source of some breakthrough results [10,9]. Other questions remain infamously open, e.g., the nondeterministic communication complexity of the minimum-spanning-tree function: For a fixed number k , Bob has a tree with vertex set $[k]$, Alice has one of a set of inequality constraints (see [31] for the details), and they are supposed to answer 1, if the tree does not satisfy the inequality constraint with equality.

In this paper, we focus on random functions, and we give tight upper and lower bounds for the nondeterministic communication complexity and its most important lower bounds: the fooling set bound; the ratio number of 1-entries over largest 1-rectangle; the fractional cover number. For that, we fix $|X| = |Y| = n$, and, we take $f(x, y)$, $(x, y) \in X \times Y$, to be independent Bernoulli

random variables with parameter $p = p(n)$, i.e., $f(x, y) = 1$ with probability p and $f(x, y) = 0$ with probability $1 - p$.

In Communication Complexity, it is customary to determine these parameters up to within a constant factor of the number of bits, but in applications, this is often not accurate enough. E.g., the above question about the extension complexity of the minimum-spanning-tree polytope asks where in the range between $(1 + o(1))2 \log n$ bits and $(1 + o(1))3 \log n$ bits the nondeterministic communication complexity lies. (Here n should be taken as $|Y| = 2^k - 2$.) Therefore, in our analyses, we focus on the constant factors in our communication complexity bounds.

1.1 Relationship to related work

In core (Communication) Complexity Theory, random functions are usually used for establishing that hard functions exist in the given model of computation. In this spirit, some easy results about the (nondeterministic) communication complexity of random functions and related parameters exist, with p a constant, mostly $p = 1/2$ (e.g., the fooling set bound is determined in this setting in [8]).

In contrast to this, in applications, the density of the matrices is typically close to 1, e.g., in combinatorial optimization, the number of 0s in a “row” $\{y \in Y \mid f(x, y) = 0\}$, is very often polylog of n . This makes necessary to look at these parameters in the spirit of the study of properties of random graph where $p = p(n) \rightarrow 1$ with $n \rightarrow \infty$. In an analogy to the fields of random graphs, the results become both considerably more interesting and also more difficult that way.

The random parameters we analyze have been studied in other fields beside Communication Complexity. Recently, Izhakian, Janson, and Rhodes [18] have determined asymptotically the triangular rank of random Boolean matrices with independent Bernoulli entries. The triangular rank is itself important in Communication Complexity [27] (and its applications [24]), and it is a lower bound to the size of a fooling set. In that paper, determining the behavior for $p \rightarrow 0, 1$ is posed as an open problem.

The size of the largest monochromatic rectangle in a random Bernoulli matrix was determined in [29] when p is bounded away from 0 and 1, but their technique fails for $p \rightarrow 1$.

The nondeterministic communication complexity of a the clique-vs-stable set problem on random graphs was studied in [4].

The parameters we study in this paper are of importance beyond Communication Complexity and its direct applications. In combinatorics, e.g., the cover number coincides with strong isometric dimension of graphs [14], and has connections to extremal set theory and Coding Theory [16,17].

The size of the largest monochromatic rectangle is of interest in the analysis of gene expression data [29], and formal concept analysis [6].

Via a construction of Lovász and Saks [27], the 1-rectangles, covers, and fooling sets of a function f correspond to stable sets, colorings, and cliques, resp.,

in a graph constructed from the function. Consequently, determining these parameters could be thought of as analyzing a certain type of random graphs. This approach does not seem to be fruitful, as the probability distribution on the set of graphs seems to have little in common with those studied in random graph theory. Here is an important example for that. In the usual random graph models (Erdős-Renyi, uniform regular), the chromatic number is within a constant factor of the independence ratio (i.e., the quotient independence number over number of vertices), and, in particular, of the fractional chromatic number (which lies between the two). The corresponding statement (replace “chromatic number” by “cover number”; “independence ratio” by “Hamming weight of f divided by the size of the largest 1-rectangle”; “fractional chromatic number” by “fractional cover number”) is false for random Boolean functions, as we will see in Section 4.

This paper is organized as follows. We determine the size of the largest monochromatic rectangle in Section 2. Section 3 is dedicated to fooling sets: we give tight upper and lower bounds. Finally, in Section 4 we give bounds for both the covering number and the fractional covering number.

1.2 Definitions

A Boolean function $f: X \times Y \rightarrow \{0, 1\}$ can be viewed as a matrix whose rows are indexed by X and the columns are indexed by Y . We will use the two concepts interchangeably. In particular, for convenience, we speak of “row” x and “column” y . We will always take $n = |X| = |Y|$ without mentioning it. Clearly, a *random Boolean function* $f: X \times Y \rightarrow \{0, 1\}$ with parameter p is the same thing as a random $n \times n$ matrix with independent Bernoulli entries with parameter p .

We use the usual conventions for asymptotics: $g \ll h$ and $g = o(h)$ is the same thing. As usual, $g = \Omega(1)$ means that g is bounded away from 0. We are interested in asymptotic statements, usually for $n \rightarrow \infty$. A statement (i.e., a family of events E_n , $n \in \mathbb{N}$) holds *asymptotically almost surely, a.a.s.*, if its probability tends to 1 as $n \rightarrow \infty$ (more precisely, $\lim_{n \rightarrow \infty} \mathbf{P}(E_n) = 1$).

2 Largest 1-rectangle

As mentioned in the introduction, driven by applications in bioinformatics, the size of the largest monochromatic rectangle in a matrix with independent (Bernoulli) entries, has been studied longer than one might expect. Analyzing computational data, Lonardi, Szpankowski, and Yang [25,26] conjectured the shape of the 1-rectangles. The conjecture was proven by Park and Szpankowski [29]. Their proof can be formulated as follows: Let $f: X \times Y \rightarrow \{0, 1\}$ be a random Boolean function with parameter p .

- If $\Omega(1) = p \leq 1/e$, then, a.a.s., the largest 1-rectangle consists of the 1-entries in a single row or column, and $R^1(f) = (1 + o(1))pn$.

- If $p \geq 1/e$ but bounded away from 1, then with $a := \operatorname{argmax}_{b \in \{1,2,3,\dots\}} bp^b$, a.a.s. the largest 1-rectangle has a rows and $p^a n$ columns, or vice-versa.

The existence of these rectangles is fairly obvious. Proving that no larger ones exist requires some work. The problem with the union-bound based proof in [29] is that it breaks down if p tends to 1 moderately quickly. In our proofs, we work with strong tail bounds instead.

Our result extends the theorem in [29] for the case that p tends to 0 or 1 quickly.

For $K \subseteq X$, the 1-rectangle of f generated by K is $R := K \times L$ with

$$L := \left\{ y \in Y \mid \forall x \in K: f(x, y) = 1 \right\}.$$

The 1-rectangle generated by a subset L of Y is defined similarly.

Theorem 2.1. *Let $f: X \times Y \rightarrow \{0, 1\}$ be a random Boolean function with parameter $p = p(n)$.*

(a) *If $5/n \leq p \leq 1/e$, then a.a.s., the largest 1-rectangle is generated by a single row or column, and if $p \gg (\ln n)/n$, its size is $(1 + o(1))pn$.*

(b) *Define*

$$\begin{aligned} a_- &:= \lfloor \log_{1/p} e \rfloor, \\ a_+ &:= \lceil \log_{1/p} e \rceil, \text{ and} \\ a &:= \operatorname{argmax}_{b \in \{a_-, a_+\}} bp^b = \operatorname{argmax}_{b \in \{1,2,3,\dots\}} bp^b. \end{aligned} \tag{1}$$

There exists a constant λ_0 , such that if $1/e \leq p \leq 1 - \lambda_0/n$, then, a.a.s., a largest 1-rectangle is generated by a rows or columns and its size is $(1 + o(1))ap^a n$.

The proof requires us to upper bound the sizes of square 1-rectangles, i.e., $R = K \times L$ with $|K| = |L|$. Sizes of square 1-rectangles have been studied, too. Building on work in [7,6,29], it was settled in [32], for constant p . We need results for $p \rightarrow 0, 1$, but, fortunately, for our theorem, we only require weak upper bounds.

For the proof of (a), we say that a 1-rectangle is *bulky*, if it extends over at least 2 rows and also over at least 2 columns. We then proceed by considering three types of rectangles:

1. those consisting of exactly one row or column (they give the bound in the theorem);
2. square bulky rectangles;
3. bulky rectangles which are not square.

For the proof of (b), we also require an appropriate notion of “bulky”: here, we say that a rectangle of dimensions $k \times \ell$ is bulky if $k \leq \ell$. By again considering square rectangles, we prove that a bulky rectangle must have $k < n/\lambda^{2/3}$. (We always define λ through $p = 1 - \lambda/n$.) By exchanging the roles of rows and

columns, and multiplying the final probability estimate by 2, we only need to consider 1-rectangles with at least as many columns as rows (i.e., bulky ones). Following that strategy yields the statement of the theorem.

The complete proof is in Appendix A.

Remark 1. (a) If $p \geq 1/e$, then

$$1/e^2 \leq \frac{p}{e} \leq p \cdot p^{\log_{1/p} e} \leq p^a \leq \frac{1}{p} \cdot p^{\log_{1/p} e} \leq \frac{1}{pe} \leq 1/e, \quad (2)$$

i.e., $p^a \approx 1/e$, more accurately $p^a = (1 - o_{p \rightarrow 1}(1))/e$.

(b) With $p = 1 - \bar{p} = 1 - \lambda/n$, the following makes the range of $R^1(f)$ clearer:

Since $\bar{p} \leq \ln(1/(1-\bar{p})) \leq \bar{p} + \bar{p}^2$ holds when $\bar{p} \leq 1 - 1/e$, we have

$$\frac{1}{ep} = \frac{n}{e\lambda} \leq p \frac{n}{\lambda} = \frac{p}{\bar{p}} \leq \frac{1}{1+\bar{p}} \cdot \frac{1}{\bar{p}} \leq \log_{1/p} e \leq \frac{1}{\bar{p}} = \frac{n}{\lambda} \quad (3)$$

Corollary 1. For $p = 1 - \frac{\lambda}{n}$ with $\lambda_0 \leq \lambda = o(n)$, we have $R^1(f) = \frac{n^2}{e\lambda} + O(n)$.

See Appendix A for the proof.

3 Fooling sets

A *fooling set* is a subset $F \subseteq X \times Y$ with the following two properties: (1) for all $(x, y) \in F$, $f(x, y) = 1$; and (2) and for all $(x, y), (x', y') \in F$, if $(x, y) \neq (x', y')$ then $f(x, y')f(x', y) = 0$. When f is viewed as a matrix, this means that, after permuting rows and columns, F identifies the diagonal entries of a submatrix which is 1 on the diagonal, and in every pair of opposite off-diagonal entries, at least one is 0. We denote by $F(f)$ the size of the largest fooling set of f . The maximum size of a fooling set of a random Boolean function with $p = 1/2$ is easy to determine (e.g., [8]).

An obvious lower bound to the fooling set size is the *triangular rank*, i.e., the size of the largest triangular submatrix, again after permuting rows and columns. (There is also an upper bound for the fooling set size in terms of the linear-algebraic rank, cf. [8,13], but since our random matrices have high rank, we cannot use that here.) In a recent Proc. AMS paper, Izhakian, Janson, and Rhodes [18] determined the triangular rank of a random matrix with independent Bernoulli entries with constant parameter p . They left as an open problem to determine the triangular rank in the case when $p \rightarrow 0$ or 1, which is our setting.

Our constructions of fooling sets of random Boolean functions make use of ingredients from random graph theory. First of all, consider the bipartite H_f whose vertex set is the disjoint union of X and Y , and with $E(H_f) = \text{supp } f \subseteq X \times Y$. For random f , this graph is an *Erdős-Renyi random bipartite graph*: each edge is picked independently with probability p . Based on the following obvious fact, we will use results about matchings in Erdős-Renyi random bipartite graphs:

Remark 2. Let $F \subseteq X \times Y$. The following are equivalent.

- (a) F is a *fooling set*.
- (b) F satisfies the following two conditions:
 - F is a matching, i.e., $F \subseteq E(H)$;
 - F is *cross-free*, i.e., for all $(x, y), (x', y') \in F$, if $(x, y) \neq (x', y')$ then $(x, y') \notin E$ or $(x', y) \notin E$.

Secondly, fooling sets can be obtained from stable sets in an auxiliary graph: For a random Boolean function f , this graph is an *Erdős-Renyi random graphs*, for which results are available yielding good lower bounds.

Fig. 1 summarizes our upper and lower bounds: Upper bounds are above the dotted lines; lower bounds are below the dotted lines; the range for p is between the dotted lines. All upper bounds are by the 1st moment method.

We emphasize that the upper and lower bounds differ by at most a constant factor. If $p \rightarrow 1$ quickly enough, i.e., $\bar{p} = 1 - p = n^{-a}$ for a constant a , then the upper bounds and lower bounds are even the same except for rounding.

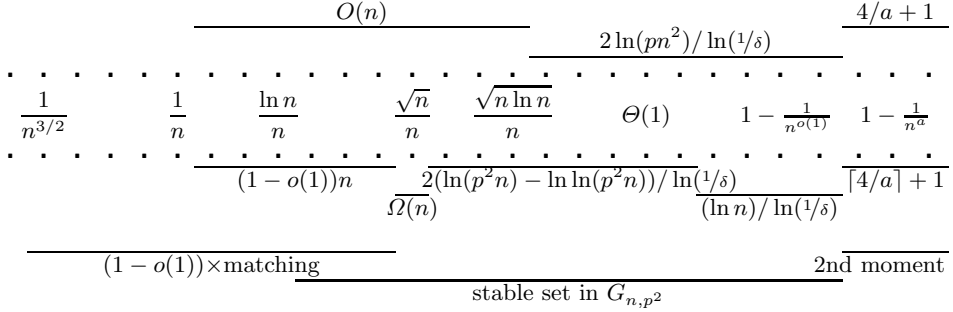


Fig. 1. Upper and lower bounds on fooling set sizes. ($\delta := 1 - p^2$)

3.1 Statement of the theorem, and a glimpse of the proof

Denote by $\nu(H)$ the size of the largest matching in a bipartite graph H . For $q = q(m)$, denote by $\mathbf{G}_{m,q}$ the graph with vertex set $\{1, \dots, m\}$ in which each of the $\binom{m}{2}$ possible edges is chosen (independently) with probability q . Let $a(q) = a_m(q)$ be a function with the property that, a.a.s., every Erdős-Renyi random graph on m vertices with edge-probability q has an independent set of size at least $a_m(q)$.

Theorem 3.1. *Let $f: X \times Y \rightarrow \{0, 1\}$ be a random Boolean function with parameter $p = p(n)$. Define $\bar{p} := 1 - p$ and $\delta := 1 - p^2$.*

- (a) For $n^{-3/2} \leq p = o(1/\sqrt{n})$, a.a.s., we have

$$F(f) = (1 - o(1))\nu(H_f).$$

- (b) If $pn - \ln n \rightarrow \infty$, then, a.a.s., $F(f) \geq a(p^2)$.
(c) If $p \gg \sqrt{(\ln n)/n}$ and $\bar{p} \geq n^{-o(1)}$, then, a.a.s.,

$$F(f) \leq 2 \log_{1/\delta}(pn^2).$$

- (d) If $a \in]0, 4[$ is a constant and $\bar{p} = n^{-a}$, then $F(f) \leq 4/a + 1$. If, in addition, $a < 1$, then $F(f) = \lfloor 4/a \rfloor + 1$

The proof is in Appendix B.

To obtain the bounds in Fig. 1, the following facts from random graph theory are needed.

Theorem 3.2 (Matchings in Erdős-Renyi random bipartite graphs, cf., e.g., [19]). Let $H = (X, Y, E)$ be a random bipartite graph with $|X| = |Y| = n$, and edge probability p .

- (a) If $p \gg 1/n$, then, a.a.s., H has a matching of size $(1 - o(1))n$.
(b) If $p = (\omega(n) + \ln n)/n$ for an ω which tends to ∞ arbitrarily slowly, then, a.a.s., H has a matching of size n .

Theorem 3.3 (Stable sets in Erdős-Renyi random graphs). Let $G = ([m], E)$ be a random graph with $\{u, v\} \in E$ with edge probability $q = q(m)$.

- (a) E.g., [19]: Let $\omega = \omega(m)$ tend to ∞ arbitrarily slowly. If $\omega/m \leq q = 1 - \Omega(1)$, then a.a.s., G has a stable set of size at least

$$2 \frac{\ln(qm) - \ln \ln(qm)}{\ln(1 - q)}.$$

- (b) Greedy stable set: If $q = \Omega(1)$, then, a.a.s., G has a stable set of size at least

$$\frac{\ln(m)}{\ln(1 - q)}.$$

For the region $p = \Theta(1/\sqrt{n})$, there is a corresponding theorem (e.g., [5]). We give here an argument about the expectation based on Turán's theorem. Turán's theorem in the version for stable sets [1] states that in a graph with vertex set V , there exists a stable set of size at least

$$\sum_{v \in V} \frac{1}{\deg(v) + 1},$$

where $\deg(v)$ denotes the degree of vertex v . For random graphs on vertex set $V = [m]$ with edge probability $q = c/m$ for a constant c , using Jensen's inequality, we find that there expected size of the largest stable set is at least

$$\begin{aligned} \mathbf{E} \left(\sum_{v \in V} \frac{1}{\deg(v) + 1} \right) &= \sum_{v \in V} \mathbf{E} \left(\frac{1}{\deg(v) + 1} \right) \\ &\geq \sum_{v \in V} \frac{1}{\mathbf{E} \deg(v) + 1} = \frac{2m}{q(m-1) + 1} \geq \frac{2m}{c+1} = \Theta(m). \end{aligned}$$

4 Fractional cover number and cover number

Armed with the fooling set and 1-rectangle-size lower bounds, we can now bound the cover number and the fractional cover number. We start with the easy case $p \leq 1/2$.

Let f be a random Boolean function $X \times Y \rightarrow \{0, 1\}$ with parameter p , as usual. If $1/n \ll p \leq 1/2$, we have $C(f) = (1 - o(1))n$. Indeed, for $p = o(1/\sqrt{n})$, Theorem 3(a) gives the lower bound based on the fooling set lower bound. For $1/e \geq p \gg (\ln n)/n$, Theorem 2.1(a) yields $R^1(f) = (1 + o(1))pn$, a.a.s., and for $1/e \leq p \leq 1/2$, the value of a in eqn. (1) of Theorem 2.1(b) is 1, so that $R^1(f) = (1 + o(1))pn$ there, too. We conclude that, a.a.s.,

$$C(f) \geq \frac{|\text{supp } f|}{R^1(f)} = \frac{(1 - o(1))pn^2}{(1 - o(1))pn} = (1 - o(1))n.$$

As indicated in the introduction, the case $p > 1/2$ is more interesting, both from the application point of view and from the point of view of the proof techniques.

For the remainder of this section, we assume that $p > 1/2$. Define $\bar{p} := 1 - p$, and $\lambda := \bar{p}n$.

4.1 The fractional cover number

We briefly review the definition of the fractional cover number. Let f be a fixed Boolean function, and let R be a random 1-rectangle of f , drawn according to a distribution π . Define

$$\gamma(\pi) := \min \left\{ \mathbf{P}_{R \sim \pi} ((x, y) \in R) \mid (x, y) \in \text{supp } f \right\}.$$

The *fractional cover number* is $C^*(f) := \min_{\pi} 1/\gamma(\pi)$, where the minimum is taken over all distributions π on the set of 1-rectangles of f .

The following inequalities are well-known [23].

$$\left. \begin{array}{l} |\text{supp } f| \\ R^1(f) \\ F(f) \end{array} \right\} \leq C^*(f) \leq C(f) \stackrel{(*)}{\leq} (1 + \ln R^1(f)) C^*(f). \quad (5)$$

Lower bound Theorem 2.1(b) allows us to lower bound $C^*(f)$. Let f be a random Boolean function $X \times Y \rightarrow \{0, 1\}$ with parameter $p > 1/2$. With $\lambda/n = \bar{p} = 1 - p$, we have a.a.s.,

$$\frac{|\text{supp } f|}{R^1(f)} \geq \frac{(1 + o(1))pn^2}{(1 + o(1))n/e \ln(1/p)} = (1 + o(1)) ep \ln(1/p)n \geq (1 - o(1)) ep\lambda \quad (6)$$

where the last inequality follows from $\bar{p} \leq \bar{p} + \bar{p}^2/2 + \bar{p}^3/3 + \dots = \ln(1/(1 - \bar{p}))$. For $\bar{p} = o(1)$, this is asymptotic to $e\lambda$. It is worth noting that the first inequality in (6) becomes an asymptotic equality if $\bar{p} = o(1)$.

Upper bound We now give upper bounds on $C^*(f)$. To prove an upper bound b on the fractional cover number for a fixed function f , we have to give a distribution π on the 1-rectangles of f such that, if R is sampled according to π , we have, for all (x, y) with $f(x, y) = 1$,

$$\mathbf{P}((x, y) \in R) \geq 1/b.$$

To prove an ‘‘a.a.s.’’ upper bound for a random f , we have to show that

$$\mathbf{P}\left(\exists(x, y): \mathbf{P}((x, y) \in R \mid f \ \& \ f(x, y) = 1) < 1/b\right) = o(1). \quad (7)$$

Our random 1-rectangle R within the random Boolean function f is sampled as follows. Let K be a random subset of X , by taking each x into K independently, with probability q . Then let $R := K \times L$ be the 1-rectangle generated (see p. 5) by the row-set K , i.e., $L := \{y \mid \forall x \in K: f(x, y) = 1\}$.

For $y \in Y$, let the random variable Z_y count the number of $x \in X$ with $f(x, y) = 0$ —in other words, the number of zeros in column y —and set $Z := \max_{y \in Y} Z$. For $(x, y) \in X \times Y$, conditioned on f and $f(x, y) = 1$, the probability that $(x, y) \in R$ equals

$$q(1 - q)^{Z_y} \geq q(1 - q)^Z,$$

so that for every positive integer z , using $1/b = q(1 - q)^z$ in (7),

$$\mathbf{P}\left(\exists(x, y): \mathbf{P}((x, y) \in R \mid f \ \& \ f(x, y) = 1) < q(1 - q)^z\right) = \mathbf{P}(Z > z). \quad (8)$$

To obtain upper bounds on the fractional cover number, we give a.a.s. upper bounds on Z , and choose q accordingly.

Theorem 4.1. *Let $1/2 > p = 1 - \bar{p} = 1 - \lambda/n$.*

- (a) *If $\ln n \ll \lambda < n/2$, then, a.a.s., $(1 - o(1)) pe\lambda \leq C^*(f) \leq (1 + o(1)) e\lambda$*
- (b) *If $\lambda = \Theta(\ln n)$, then, a.a.s., $C^*(f) = \Theta(\ln n)$.*
- (c) *If $1 \ll \lambda = o(\ln n)$, then, a.a.s.,*

$$(1 - o(1)) \lambda \leq C^*(f) \leq (1 + o(1)) e \max\left(2\lambda, \frac{\ln n}{\ln((\ln n)/\lambda)}\right)$$

To summarize, we can determine the fractional cover number accurately in the region $\ln n \ll \lambda \ll n$. For $\lambda = \Theta(\ln n)$ and for $\lambda = \Theta(n)$, we can determine $C^*(f)$ up to a constant. However, for $\lambda = o(\ln n)$, there is a large gap between our upper and lower bounds.

Proof. The lower bounds follow from the discussion above.

Proof of the upper bound in (a). For every constant $t > 0$, let

$$\psi(t) := 1/((1 + t) \ln(1 + t) - t).$$

With

$$h(t) = h(t, n) := \frac{\lambda}{\psi(t) \ln n},$$

using the a standard Chernoff estimate (Theorem 2.1, Eqn.(2.5) in [19]) we find that

$$\mathbf{P}(Z_1 \geq (1+t)\lambda) \leq e^{-\lambda/\psi(t)} \leq e^{-h(t)n},$$

so that, by the union bound,

$$\mathbf{P}(Z \geq (1+t)\lambda) \leq e^{-h(t)}. \quad (10)$$

For every fixed $t > 0$, $h(t)$ tends to infinity with n , so that the RHS in (10) is $o(1)$. Using that in (8), we obtain

$$\mathbf{P}\left(\exists(x, y): \mathbf{P}((x, y) \in R \mid f \& f(x, y) = 1) < q(1-q)^{(1+t)\lambda}\right) = \mathbf{P}(Z > (1+t)\lambda) = o(1),$$

and, taking $q := \frac{1}{(1+t)\lambda}$, we obtain, a.a.s.,

$$\mathbf{C}^*(f) \leq \frac{1}{q(1-q)^{(1+t)\lambda}} \leq \frac{1+t}{1 + \frac{1}{(1+t)\lambda}} e\lambda,$$

where we used $(1-\varepsilon)^k \geq (1-k\varepsilon^2)e^{-k\varepsilon}$ for $\varepsilon < 1$. Since this is true for every $t > 0$, we conclude that, a.a.s., $\mathbf{C}^*(f) \leq (1-o(1))e\lambda$.

Proof of the upper bounds in (b), (c). Here we use a slightly different Chernoff bound (Lemma 13 in the appendix).

For (b), suppose that $\lambda \leq C \ln n$ for a constant $C > 1$. Using Lemma 13 with $\alpha = e^2 C \ln n$, we obtain

$$\mathbf{P}(Z_1 \geq e^2 C \ln n) = O(1/\sqrt{\ln n}) e^{-\lambda} \left(\frac{eC \ln n}{e^2 C \ln n} \right)^\alpha = O(1/\sqrt{\ln n}) e^{-\ln n}.$$

and thus

$$\mathbf{P}(Z \geq e^2 C \ln n) = o(1).$$

We conclude similarly as above: with $q := \frac{1}{e^2 C \ln n}$ we obtain, a.a.s., $\mathbf{C}^*(f) \leq e^3 C \ln n$.

Finally, for (c), if $\lambda = o(\ln n)$, let $\varepsilon > 0$ be a constant, and use Lemma 13 again, with

$$\alpha := \max\left(2\lambda, \frac{(1+\varepsilon) \ln n}{\ln\left(\frac{\ln n}{e\lambda}\right)}\right).$$

We find that

$$\mathbf{P}(Z_1 \geq \alpha) = o(e^{-\alpha \ln(\alpha/e\lambda)}),$$

and the usual calculation (Appendix C.1) shows that $\alpha \ln(\alpha/e\lambda) \geq \ln n$, which implies

$$\mathbf{P}(Z \geq \alpha) = o(1).$$

Conclude similarly as above, with $q := \frac{1}{\alpha}$, we obtain, a.a.s.,

$$\mathbf{C}^*(f) \leq e\alpha = e \max \left(2\lambda, (1 + \varepsilon) \frac{\ln n}{\ln\left(\frac{\ln n}{e\lambda}\right)} \right).$$

One obtains the statement in the theorem by letting ε tend to 0; the e -factor in the denominator of the \ln of the denominator in α is irrelevant as $n \rightarrow \infty$.

The cover number Inequality (*) in (5) gives us corresponding upper bounds on the cover number.

Corollary 2. *We have $(1 - o(1))\lambda \leq \mathbf{C}(f)$, and:*

(a) *if $\ln n \ll \lambda = O(n/\ln n)$, then, a.a.s., $\mathbf{C}(f) = O(\lambda \ln n)$;*

(b) *if $\lambda = \Theta(\ln n)$, then, a.a.s., $\mathbf{C}(f) = O(\ln^2 n)$;*

(c) *if $1 \ll \lambda = o(\ln n)$, then, a.a.s., $\mathbf{C}(f) = O\left(\max\left(\lambda \ln n, \frac{\ln^2 n}{\ln((\ln n)/\lambda)}\right)\right)$.* □

4.2 Binary-Logarithm of the number of distinct rows, and the ratio \mathbf{C}/\mathbf{C}^*

When we view f as a matrix, the binary logarithm of the number of distinct rows is a lower bound on the cover number of f [23]. We have the following.

Proposition 1.

(a) *If $1/2 \geq \bar{p} = \Omega(1/n)$, then, a.a.s., the 2-Log lower bound on $\mathbf{C}(f)$ is $(1 - o(1))\log_2 n$.*

(b) *If $\bar{p} = n^{-\gamma}$ for $1 < \gamma \leq 3/2$, then a.a.s., the 2-Log lower bound on $\mathbf{C}(f)$ is $(1 - o(1))(2 - \gamma)\log_2 n$.*

Proof. Directly from the following Lemma 1 about the number of distinct rows, with $\lambda = n^{1-\gamma}$.

Lemma 1.

(a) *If $1/2 \geq \bar{p} = \Omega(1/n)$, then, a.a.s., f has $\Theta(n)$ distinct non-zero rows.*

(b) *With $\bar{p} = \lambda/n$, if $n^{-1/2} \leq \lambda \leq 1/2$, then, a.a.s., f has $\Omega(\lambda n)$ distinct non-zero rows.*

(The constants in the big-Omegas are absolute.)

For the sake of completeness, we sketch the proof in Appendix C.3.

Erdős-Renyi random graphs have the property that the chromatic number is within a small constant factor from the lower bound one obtains from the independence ratio. For the cover number of Boolean functions, this is not the case. Indeed, Theorem 4.1(c), together with Proposition 1, shows that, a.a.s.,

$$\frac{\mathbf{C}(f)}{\mathbf{C}^*(f)} \geq (1 + o(1)) \frac{\log_2 n}{\frac{\ln n}{\ln\left(\frac{\ln n}{\lambda}\right)}} = \Omega\left(\ln\left(\frac{\ln n}{\lambda}\right)\right),$$

which is $\Omega(\ln \ln n)$ if $\lambda = \ln^{o(1)} n$.

This gap is more pronounced in the (not quite as interesting) situation when $\lambda = o(1)$. Consider, e.g., $\lambda = n^{-\varepsilon}$, for some $\varepsilon = \varepsilon(n) = o(1/\ln \ln n)$, say. Similarly to the proofs of Theorem 4.1, we obtain that $C^*(f) \leq e \max(10, 2/\varepsilon)$. (The max-term comes from the somewhat arbitrary upper bound $Z \leq \max(10, 2/\varepsilon)$.) For the Log-2 lower bound on the cover number, we have $(1 - \varepsilon) \log_2 n$, by Proposition 1, and thus

$$\frac{C(f)}{C^*(f)} = \Omega(\varepsilon \ln n).$$

5 Acknowledgments

The authors would like to thank the anonymous referees for their valuable comments.

Dirk Oliver Theis is supported by Estonian Research Council, ETAG (*Eesti Teadusagentuur*), through PUT Exploratory Grant #620. Mozghan Pourmoradnasseri is recipient of the Estonian IT Academy Scholarship. This research is supported by the European Regional Fund through the Estonian Center of Excellence in Computer Science, EXCS.

References

1. Alon, N., Spencer, J.H.: The Probabilistic Method. Wiley (2008)
2. Beasley, L.B., Klauck, H., Lee, T., Theis, D.O.: Communication complexity, linear optimization, and lower bounds for the nonnegative rank of matrices (dagstuhl seminar 13082). Dagstuhl Reports 3(2), 127–143 (2013)
3. Bollobás, B.: Random graphs, Cambridge Studies in Advanced Mathematics, vol. 73. Cambridge University Press, Cambridge, second edn. (2001)
4. Braun, G., Fiorini, S., Pokutta, S.: Average case polyhedral complexity of the maximum stable set problem. In: Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques, AP-PROX/RANDOM 2014, September 4-6, 2014, Barcelona, Spain. pp. 515–530 (2014), <http://dx.doi.org/10.4230/LIPIcs.APPROX-RANDOM.2014.515>
5. Dani, V., Moore, C.: Independent sets in random graphs from the weighted second moment method. In: Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques, pp. 472–482. Springer (2011)
6. Dawande, M., Keskinocak, P., Swaminathan, J.M., Tayur, S.: On bipartite and multipartite clique problems. J. Algorithms 41(2), 388–403 (Nov 2001), <http://dx.doi.org/10.1006/jagm.2001.1199>
7. Dawande, M., Keskinocak, P., Tayur, S.: On the biclique problem in bipartite graphs. Carnegie Mellon University (1996), gsia Working Paper
8. Dietzfelbinger, M., Hromkovič, J., Schnitger, G.: A comparison of two lower-bound methods for communication complexity. Theoret. Comput. Sci. 168(1), 39–51 (1996), [http://dx.doi.org/10.1016/S0304-3975\(96\)00062-X](http://dx.doi.org/10.1016/S0304-3975(96)00062-X), 19th International Symposium on Mathematical Foundations of Computer Science (Košice, 1994)

9. Fiorini, S., Kaibel, V., Pashkovich, K., Theis, D.O.: Combinatorial bounds on nonnegative rank and extended formulations. *Discrete Math.* 313(1), 67–83 (2013)
10. Fiorini, S., Massar, S., Pokutta, S., Tiwary, H.R., Wolf, R.: Linear vs. semidefinite extended formulations: Exponential separation and strong lower bounds. In: *STOC* (2012)
11. Fiorini, S., Massar, S., Pokutta, S., Tiwary, H.R., Wolf, R.D.: Exponential lower bounds for polytopes in combinatorial optimization. *Journal of the ACM (JACM)* 62(2), 17 (2015)
12. Friesen, M., Hamed, A., Lee, T., Theis, D.O.: Fooling-sets and rank. *European Journal of Combinatorics* 48, 143–153 (2015)
13. Friesen, M., Theis, D.: Fooling-sets and rank in nonzero characteristic. In: Nešetřil, J., Pellegrini, M. (eds.) *The Seventh European Conference on Combinatorics, Graph Theory and Applications*. CRM series, vol. 16, pp. 383–390. CRM (2013)
14. Froncek, D., Jerebic, J., Klavzar, S., Kovár, P.: Strong isometric dimension, biclique coverings, and sperner’s theorem. *Combinatorics, Probability & Computing* 16(2), 271–275 (2007), <http://dx.doi.org/10.1017/S0963548306007711>
15. Goemans, M.X.: Smallest compact formulation for the permutahedron. *Mathematical Programming* 153(1), 5–11 (2015)
16. Hajiabolhassan, H., Moazami, F.: Secure frameproof code through biclique cover. *Discrete Mathematics & Theoretical Computer Science* 14(2), 261–270 (2012), <http://www.dmtcs.org/dmtcs-ojs/index.php/dmtcs/article/view/2131/4075>
17. Hajiabolhassan, H., Moazami, F.: Some new bounds for cover-free families through biclique covers. *Discrete Mathematics* 312(24), 3626–3635 (2012)
18. Izhakian, Z., Janson, S., Rhodes, J.: Superboolean rank and the size of the largest triangular submatrix of a random matrix. *Proceedings of the American Mathematical Society* 143(1), 407–418 (2015)
19. Janson, S., Łuczak, T., Ruciński, A.: *Random graphs*. Wiley-Interscience Series in Discrete Mathematics and Optimization, Wiley-Interscience, New York (2000)
20. Kaibel, V.: Extended formulations in Combinatorial Optimization. *Optima – Mathematical Optimization Society Newsletter* 85, 2–7 (04 2011), www.mathopt.org/Optima-Issues/optima85.pdf
21. Karp, R.M., Sipser, M.: Maximum matchings in sparse random graphs. In: *FOCS*. pp. 364–375 (1981)
22. Klauck, H., Lee, T., Theis, D.O., Thomas, R.R.: Limitations of convex programming: lower bounds on extended formulations and factorization ranks (dagstuhl seminar 15082). *Dagstuhl Reports* 5(2), 109–127 (2015)
23. Kushilevitz, E., Nisan, N.: *Communication complexity*. Cambridge University Press, Cambridge (1997)
24. Lee, T., Theis, D.: Support based bounds for positive semidefinite rank. Tech. Rep. arXiv:1203.3961, arXiv (2012)
25. Lonardi, S., Szpankowski, W., Yang, Q.: Finding biclusters by random projections. In: *Combinatorial Pattern Matching*, pp. 102–116. Springer (2004)
26. Lonardi, S., Szpankowski, W., Yang, Q.: Finding biclusters by random projections. *Theoretical Computer Science* 368(3), 217–230 (2006)
27. Lovás, L., Saks, M.: Communication complexity and combinatorial lattice theory. *Journal of Computer and System Sciences* 47, 322–349 (1993)
28. Mitzenmacher, M., Upfal, E.: *Probability and Computing — Randomized Algorithms and Probabilistic Analysis*. Cambridge (2006)
29. Park, G., Szpankowski, W.: Analysis of biclusters with applications to gene expression data. In: *International Conference on Analysis of Algorithms DMTCS proc. AD*. vol. 267, p. 274 (2005)

30. Roughgarden, T.: Communication complexity (for algorithm designers). arXiv preprint p. arXiv:1509.06257 (2015)
31. Schrijver, A.: Combinatorial optimization. Polyhedra and efficiency., Algorithms and Combinatorics, vol. 24. Springer-Verlag, Berlin (2003)
32. Sun, X., Nobel, A.B.: On the size and recovery of submatrices of ones in a random binary matrix. J. Mach. Learn. Res 9, 2431–2453 (2008)
33. Yannakakis, M.: Expressing combinatorial optimization problems by linear programs. J. Comput. System Sci. 43(3), 441–466 (1991), [http://dx.doi.org/10.1016/0022-0000\(91\)90024-Y](http://dx.doi.org/10.1016/0022-0000(91)90024-Y)

A Proof of Theorem 2.1

We will assume, for simplicity, that $X = Y = [n]$.

A.1 Small p : Proof of Theorem 2.1 (a)

We say that a rectangle is *bulky*, if it extends over at least 2 rows and also over at least 2 columns. The proof of Theorem 2.1 proceeds by considering three types of rectangles:

1. those consisting of exactly one row or column (they give the bound in the theorem);
2. square bulky rectangles;
3. bulky rectangles which are not square.

Let us start with the easiest type (1). The size of such a rectangle is the number of 1s in the chosen row.

Lemma 2. *For all p, n , a.a.s., there exists a row in f containing at least pn 1s. If $p \gg (\ln n)/n$, for every constant $\varepsilon \in]0, 1]$, a.a.s., no row or column has more than $(1 + \varepsilon)pn$ 1s.*

Proof. For the first statement, note that the probability that number of 1s in a fixed row is less than pn is at most $1/2$ (median of a binomial distribution). Since the rows are independent, the probability that all rows have fewer than pn 1s is at most 2^{-n} .

For the second statement, we use an easy Chernoff-type bound (Theorem 4.4(2) in [28]). Denote by X the number of 1s in a fixed row of f . Then

$$\mathbf{P}(X \geq (1 + \varepsilon)pn) \leq e^{-\varepsilon^2 pn/3} \leq e^{-2 \ln n} = n^{-2},$$

where the last inequality holds for large enough n , because $pn \gg \ln n$ implies $pn > 6\varepsilon^{-2} \ln n$ for n sufficiently large. Hence, the probability that a row (or a column) exists which has at least $(1 + \varepsilon)pn$ 1s is $o(1)$.

We now deal with rectangles of type (2).

Lemma 3. *If $p \geq 5/n$, then, a.a.s., there is no square 1-rectangle of size $\sqrt{pn} \times \sqrt{pn}$.*

Proof. We abbreviate $\kappa := pn$. By the union bound, for the probability $q = q(n)$ that there exists a 1-rectangle of size $\sqrt{\kappa} \times \sqrt{\kappa}$, we have

$$q \leq \binom{n}{\sqrt{\kappa}}^2 p^\kappa \leq \left(\frac{e^2 n}{p}\right)^{\sqrt{\kappa}} p^\kappa$$

Applying \ln , we find

$$\ln q \leq \sqrt{\kappa} \ln n + 2\sqrt{\kappa} + \sqrt{\kappa} \ln(1/p) - \kappa \ln(1/p) = \sqrt{\kappa} \left(\ln n + 2 - (\sqrt{\kappa} - 1) \ln(1/p) \right). \quad (11)$$

Now we distinguish cases. If $(2 \ln n)^2/n \leq p \leq 1/e$, then $\sqrt{\kappa} \geq 2 \ln n$, and hence we can bound the expression in the parentheses in (11) as follows:

$$\ln n + 2 - (\sqrt{\kappa} - 1) \ln(1/p) \leq \ln n + 2 - 2 \ln n + 1 \leq -\ln n,$$

for all large enough n . Hence, $q \rightarrow 0$ in this region. If, on the other hand, $5/n \leq p \leq (2 \ln n)^2/n$, then

$$\begin{aligned} \ln q &\leq \sqrt{5} \left(\ln n + 2 - (\sqrt{5} - 1) (\ln n - 2 \ln(2 \ln n)) \right) \\ &\leq -\sqrt{5} (\sqrt{5} - 2) \ln n + O(\ln \ln n) = -\Omega(\ln n). \end{aligned}$$

Hence, $q \rightarrow 0$ in this region, too, which completes the proof of the lemma.

Finally, we come to rectangles of type (3). Consider the probability, ϱ , that f contains a bulky 1-rectangle of size s . By Lemma 3, if such a 1-rectangle has dimensions $a \times b$, we must have $a < \sqrt{pn}$ or $b < \sqrt{pn}$, or else $\varrho = o(1)$. We have $\varrho \leq 2\varrho'$, where ϱ' is the probability that f contains a 1-rectangle of size s consisting of at least as many columns than rows. For ϱ' , we need to consider only 1-rectangles with $a < \sqrt{pn}$. Moreover, increasing b if necessary, w.l.o.g., we may restrict to rectangles generated by a row-set of size a , with $2 \leq a \leq n$ (the LB 2 comes from the condition that the rectangle be bulky).

Lemma 4. *With $\kappa = pn$, if $5 \leq \kappa = O(\text{polylog } n)$, then, a.a.s., there is no bulky rectangle of size at least κ .*

Proof. By the remarks above, we have to bound the probability that there exists a row-set of size $a \in \{2, \dots, \sqrt{\kappa}\}$ which generates a 1-rectangle of size at least κ/a .

Firstly, for a given set K of a rows, we bound the probability that the rectangle it generates has size at least pn . Denote by S the number of columns in the rectangle generated by K . This is a $\text{Bin}(n, p^a)$ r.v. and we find that

$$\mathbf{P}(a \cdot b \geq \kappa) = \mathbf{P}(S \geq \kappa/a) \leq \binom{n}{\kappa/a} p^\kappa = \binom{n}{\kappa/a} \left(\frac{\kappa}{n}\right)^\kappa.$$

Secondly, we sum over all sets K of cardinality a , and compute

$$\binom{n}{a} \binom{n}{\kappa/a} p^\kappa (1 - p^a)^{n - \kappa/a} \leq n^{a + \kappa/a - \kappa + \kappa \log_n \kappa} = n^{-(\kappa(1 - 1/a) - a - o(\kappa))},$$

where $\kappa \log_n \kappa = o(\kappa)$ follows from $\kappa = O(\text{polylog } n)$.

Now, because $a < \sqrt{\kappa}$, we have that the exponent on $1/n$ is $\kappa(1-1/a) - o(\kappa) \geq \kappa/3$, as $a \geq 2$. Finally, summing over all a , we obtain, as an upper bound for the probability that one of these rectangles has size κ or larger, the expression $n^{-(\kappa/3-1)}$ which is $o(1)$, as $\kappa \geq 5$.

For the remaining case, we will need the following numerical fact, whose proof we leave to the reader.

Lemma 5. *There exists an $\varepsilon > 0$ such that, for all $p \in]1/8, 1/e]$ and $a \in \{2, 3\}$,*

$$\left(ap^{a-2} \right)^{p/a} \left(\frac{1-p^a}{1-p/a} \right)^{1-p/a} \leq 1 - \varepsilon. \quad \square$$

Now we deal with bulky rectangles.

Lemma 6. *With $\kappa := pn$, if $\ln^4 n \leq \kappa \leq n/e$, then, a.a.s., there is no bulky rectangle of size at least κ .*

Proof. By the remarks above Lemma 4, we have to bound the probability that there exists a row-set of size $a \in \{2, \dots, \sqrt{\kappa}\}$ which generates a 1-rectangle of size at least κ/a .

For $2 \leq a < \sqrt{\kappa}$, let X_a count the number of columns y with $f(x, y) = 1$ for $x = 1, \dots, a$. We are going to show that

$$P := \sum_{a=2}^{\sqrt{\kappa}} \binom{n}{a} \mathbf{P}(X_a \geq \kappa/a) = o(1).$$

The r.v. X_a has $\text{Bin}(n, p^a)$ distribution. We compute

$$\mathbf{P}(X_a \geq \kappa/a) \leq \binom{n}{\kappa/a} (p^a)^{\kappa/a} \leq \left(\frac{en}{\kappa/a} \right)^{\kappa/a} (p^a)^{\kappa/a} = \left((ea)^{1/(a-1)} p \right)^{\kappa/a} \quad (12)$$

Now, there exists an constant $\varrho < 1$ such that

$$(ea)^{1/(a-1)} \leq \begin{cases} 8\varrho, & \text{for all } a \geq 2, \text{ and} \\ e\varrho, & \text{for } a \geq 4. \end{cases} \quad (13)$$

Consequently, we distinguish two cases:

- (i) $p \leq 1/8$ and
- (ii) $1/8 < p \leq 1/e$.

Case (i): $p \leq 1/8$. In this case, we compute

$$\begin{aligned}
P &= \sum_{a=2}^{\sqrt{\kappa}} \binom{n}{a} \mathbf{P}(X_a \geq \kappa/a) \\
&\leq \sum_{a=2}^{\sqrt{\kappa}} \binom{n}{a} \left(\left((ea)^{1/(a-1)} p \right)^{\kappa/a} \right)^{a-1} && \text{[by (12)]} \\
&\leq \sum_{a=2}^{\sqrt{\kappa}} \binom{n}{a} \left(\varrho^{\kappa/a} \right)^{a-1} && \text{[by (13)]} \\
&\leq \sum_{a=2}^{\sqrt{\kappa}} \binom{n}{a} \left(\varrho^{\sqrt{\kappa}} \right)^{a-1} && \text{[since } a \leq \sqrt{\kappa} \text{ and } \varrho < 1] \\
&= \varrho^{\sqrt{\kappa}} \sum_{a=2}^{\sqrt{\kappa}} \binom{n}{a} \left(\varrho^{\sqrt{\kappa}} \right)^{a-2} \\
&\leq \varrho^{\sqrt{\kappa}} n^2 \sum_{a=0}^{\sqrt{\kappa}-2} \binom{n-2}{a} \left(\varrho^{\sqrt{\kappa}} \right)^a && \text{[replacing } a \rightsquigarrow a-2] \\
&\leq \varrho^{\sqrt{\kappa}} n^2 (1 + \varrho^{\sqrt{\kappa}})^{n-2} && \text{[Binomial theorem]} \\
&\leq \varrho^{\sqrt{\kappa}} n^2 e^{n\varrho^{\sqrt{\kappa}}} \\
&\leq \varrho^{\sqrt{\kappa}} n^2 e^{n\varrho^{\ln^2 n}} && \text{[because } p \geq (\ln^4 n)/n \text{ and } \varrho < 1] \\
&= o(1) && \text{[because } \varrho = 1 - \Omega(1)\text{].}
\end{aligned}$$

Case (ii): $1/8 < p \leq 1/e$. In this case, by (13), the same calculation as in the $p < 1/8$ -case works if the sum is started with $a = 4$. For the first two terms of the sum, $a = 2, 3$, we use a Chernoff bound on X_a , which gives us (e.g., Eqn. (2.4) in [19])

$$\mathbf{P}(X_a \geq \kappa/a) \leq \left(\left(ap^{a-2} \right)^{p/a} \left(\frac{1-p^a}{1-p/a} \right)^{1-p/a} \right)^n. \quad (14)$$

Using Lemma 5, we conclude

$$\begin{aligned}
P &= \sum_{a=2}^{\sqrt{\kappa}} \binom{n}{a} \mathbf{P}(X_a \geq \kappa/a) \\
&= \binom{n}{2} \mathbf{P}(X_2 \geq \kappa/2) + \binom{n}{3} \mathbf{P}(X_3 \geq \kappa/3) + \sum_{a=4}^{\sqrt{\kappa}} \binom{n}{a} \mathbf{P}(X_a \geq \kappa/a) \\
&= o(1) + o(1) + o(1),
\end{aligned}$$

where the first two “ $o(1)$ ”s follow from (14) and Lemma 5, and the third is the same calculation as in the previous case.

This concludes the proof of Theorem 2.1(a).

A.2 Large p : Proof of Theorem 2.1(b)

Now we prove the part of Theorem 2.1 about $p \geq 1/e$. Again, we first prove a statement about square rectangles.

Lemma 7. *For every $\varepsilon > 0$ there exists a constant λ_0 such that, if $n \geq \bar{p}n = \lambda \geq \lambda_0$, then, a.a.s., there is no square 1-rectangle of size*

$$\frac{n}{\lambda^{1-\varepsilon}} \times \frac{n}{\lambda^{1-\varepsilon}}$$

Proof. This is a direct union bound computation. With $b := \frac{n}{\lambda^{1-\varepsilon}}$, the probability that such a 1-rectangle exists is at most

$$\binom{n}{b}^2 p^{b^2} = \binom{n}{b}^2 (1 - \bar{p})^{b^2} \leq e^{b(2 \ln(en/b) - \lambda b/n)} = e^{b \cdot A_b},$$

where

$$\begin{aligned} A_b &= 2 \ln(en/b) - \lambda b/n \\ &= 2 \ln(e \lambda^{1-\varepsilon}) - \lambda^\varepsilon \\ &\leq -1, \end{aligned}$$

where the last inequality holds if $\lambda \geq \lambda_0$ and λ_0 is large enough. The claim follows because $b \rightarrow \infty$.

As above, we need the notion of a “bulky” rectangle: Here, we say that a rectangle of dimensions $k \times \ell$ is *bulky*, if $k \leq \ell$. By Lemma 7, in particular, a.a.s., a bulky rectangle must have $k < n/\lambda^{2/3}$. Again, by exchanging the roles of rows and columns, and multiplying the final probability estimate by 2, we only need to consider 1-rectangles with at least as many columns as rows (i.e., bulky ones).

Proof (Proof of Theorem 2.1(b)). For every $b \in [n]$, denote by X_b the number of columns of the 1-rectangle generated by the row set $\{1, \dots, b\}$ —a random variable with $\text{Bin}(n, p^b)$ distribution. We prove that, for every $1 < u < 2$,

$$\sum_{b=1}^{n/\lambda^{2/3}} \binom{n}{b} \mathbf{P}(bX_b \geq u a p^a n) = o(1), \quad (15)$$

which, together with Lemma 7, proves Theorem 2.1(b).

We split the proof into two lemmas, dealing with the cases $b \leq \log_{1/p} e$ and $b \geq \log_{1/p} e$, resp., stated below. Establishing these lemmas completes the proof of Theorem 2.1(b).

Lemma 8. For every $u \in]1, 2[$ there exists a constant $\lambda_0 \geq 1$ such that, for every $\bar{p} \geq \lambda_0/n$, and for every $1 \leq b \leq \log_{1/\bar{p}} e$, we have

$$\binom{n}{b} \mathbf{P}\left(X_b \geq u \frac{a}{b} p^a n\right) = o_u(1/n).$$

Lemma 9. For every $u \in]1, 2[$ there exists a constant λ_0 such that, if $\bar{p}n = \lambda \geq \lambda_0$, and $\log_{1/\bar{p}} e \leq b \leq n/\lambda^{3/2}$, then

$$\binom{n}{b} \mathbf{P}\left(X_b \geq u \frac{a}{b} p^a n\right) = o_u(1/n).$$

Proof (Proof of Lemma 8). Define

$$\delta := \min\left(u \frac{ap^a}{bp^b} - 1, 1\right).$$

Note that $\delta \geq u - 1 > 0$ by the definition of a in (1). The “1” on the RHS of the minimum is somewhat arbitrary: the particular version of the Chernoff inequality which we refer to, [28, Thm 4.4-2], requires $\delta \leq 1$. Using this Chernoff bound in

$$\mathbf{P}(X_b \geq u ap^a n/b) \leq \mathbf{P}(X_b \geq (1 + \delta) \mathbf{E} X_b) \leq e^{-\delta^2 \mathbf{E} X_b/3},$$

and the inequality $\binom{n}{b} \leq (en/b)^b$, we estimate

$$\begin{aligned} \ln\left(\binom{n}{b} \mathbf{P}\left(X_b \geq u \frac{a}{b} p^a n\right)\right) &\leq b \ln\left(\frac{en}{b}\right) - \delta^2 p^b n/3 \\ &\leq n \left(\frac{b}{n} \ln\left(\frac{en}{b}\right) - \frac{\delta^2}{3e}\right) \quad [\text{since } b \leq \log_{1/\bar{p}} e]. \quad (*) \end{aligned}$$

For any real $b \in [1, \log_{1/\bar{p}} e]$, denote by A_b the term inside the parentheses in (*).

Since $b \mapsto A_b$ is nondecreasing on $[1, n]$, we have, for every $b \in [1, \log_{1/\bar{p}} e]$,

$$\begin{aligned} A_b &\leq A_{\log_{1/\bar{p}} e} \\ &\leq A_{1/\bar{p}} \quad [A. \text{ nondecreasing and } \log_{1/\bar{p}} e \leq \frac{1}{\bar{p}} \leq n, \text{ by (3)}] \\ &\leq A_{n/\lambda_0} \quad [A. \text{ nondecreasing and } 1/\bar{p} \leq n/\lambda_0 \leq n] \\ &= \frac{\ln(e\lambda_0)}{\lambda_0} - \frac{\delta^2}{3e} \\ &\leq \frac{\ln(e^2\lambda_0)}{\lambda_0} - \frac{(u-1)^2}{3e}. \quad [\text{as } \delta \geq u-1.] \end{aligned}$$

Hence, for sufficiently large λ_0 , depending only on u , we have, for all $b \in [1, \log_{1/\bar{p}} e]$,

$$A_b = -\Omega_u(1),$$

so that

$$\mathbf{P}(X_b \geq ap^a n/b) \leq e^{-nA_b} = e^{-\Omega_u(n)} = o_u(1/n)$$

which concludes the proof of the lemma.

Proof (Proof of Lemma 9). By Lemma 7, we already know that, if a bulky 1-rectangles generated by b rows exists with non- $o(1)$ probability, we must have $b < n/\lambda^{2/3}$.

Define δ as follows, $0 < u - 1 \leq \delta := (u - 1) \frac{ap^a}{bp^b} \leq u \frac{ap^a}{bp^b} - 1$, and let

$$\varepsilon := \begin{cases} (u - 1)^2/3, & \text{if } \delta \leq 3/2; \\ \ln^{5/2} - 1 + 2/5, & \text{otherwise.} \end{cases}$$

We do the case distinction because we use two slightly different versions of Chernoff in our estimate of

$$\varrho := \mathbf{P}\left(X_b \geq u \frac{a}{b} p^a n\right).$$

If $\delta \leq 3/2$, then

$$\begin{aligned} \varrho &\leq \mathbf{P}\left(X_b \geq (1 + \delta) \mathbf{E} X_b\right) \\ &\leq e^{-\delta^2 p^b n/3} && \text{[Chernoff, e.g., [19, Cor. 2.3]]} \\ &\leq e^{-(u-1)(u-1) \frac{a}{b} p^a n/3} && \text{[definition of } \delta, \text{ and } \delta \geq u - 1] \\ &= e^{-\varepsilon \frac{a}{b} p^a n}. \end{aligned}$$

If, on the other hand, $\delta > 3/2$, then

$$u \frac{a}{b} p^a n = u \frac{ap^a}{bp^b} \cdot \mathbf{E} X_b \geq (\delta + 1) \cdot \mathbf{E} X_b \geq \frac{5}{2} \cdot \mathbf{E} X_b,$$

and we have, by Eqn. (2.10) in [19, Cor. 2.4],

$$\varrho \leq e^{-\varepsilon \frac{a}{b} p^a n}.$$

In both cases, we conclude

$$\begin{aligned} &\ln \left(\binom{n}{b} \mathbf{P}\left(X_b \geq u \frac{a}{b} p^a n\right) \right) \\ &\leq b \ln(en/b) - \varepsilon \frac{a}{b} p^a n \\ &\leq b \ln(en/b) - \varepsilon \frac{an}{e^2 b} && [p^a \geq 1/e^2 \text{ by (2)}] \\ &\leq b \ln(en/b) - \varepsilon \frac{n^2}{2e^3 \lambda b} && [a \geq \lceil n/e\lambda \rceil \text{ by (3), \& } \lceil n/e\lambda \rceil \geq n/2e\lambda \text{ as } \lambda \leq n/e] \\ &\leq b \ln(e^2 \lambda) - \varepsilon \frac{n^2}{2e^3 \lambda b} && [\text{as } b \geq \log_{1/p} e \geq n/e\lambda, \text{ by (3)}] \\ &\leq \frac{n}{\lambda^{2/3}} \ln(e^2 \lambda) - \frac{\varepsilon}{2e^3} \frac{n}{\lambda^{1/3}} && [\text{as } b \leq n/(\sqrt{\lambda} \ln \lambda)] \\ &= \frac{n}{\lambda^{1/3}} \left(-\frac{\varepsilon}{2e^3} + o_{\lambda \rightarrow \infty}(1) \right). \end{aligned}$$

Hence, if λ is at least a large enough constant, λ_0 , then

$$\binom{n}{b} \mathbf{P}\left(X_b \geq u \frac{a}{b} p^a n\right) = e^{-\Omega_u(n^{2/3})} = o(1/n),$$

and the lemma is proven.

A.3 Proof of Corollary 1

Proof (Proof of the corollary from Theorem 2.1). For the given $p = 1 - \bar{p}$, if $1/e = p^a$, we have

$$\begin{aligned} ap^a &= (1 + O(\bar{p})) \frac{\log_{1/p} e}{e} = \frac{1 + O(\bar{p})}{e \ln \frac{1}{1-\bar{p}}} \\ &= \frac{1 + O(\bar{p})}{e(\bar{p} + \bar{p}^2/2 + \bar{p}^3/3 + \dots)} \stackrel{(*)}{=} \frac{1 + O(\bar{p})}{e\bar{p}} = \frac{1}{e\bar{p}} + O(1) = \frac{n}{e\lambda} + O(1), \end{aligned}$$

where equation $(*)$ uses $\bar{p} = o(1)$. Multiplying by n and invoking Theorem 2.1(b), we obtain the desired bound.

B Proof of Theorem 3.1

The proof of Theorem 3.1 extends over the following three subsections. We first treat upper bounds based on the 1st moment method, then we make the 2nd moment calculation (for the case when $p \rightarrow 1$ quickly), and finally we show how to obtain fooling sets by combining a matching in random bipartite graphs and a stable set in a random (not bipartite) graph.

B.1 Upper bounds: The number of fooling sets of size r

Let the random variable $X = X_r = X_{r,n,p}$ count the number of fooling sets of size r in f . For a set $F \subseteq [n] \times [n]$, denote by A_F the event that F is a fooling set of f . We have

$$X_r = \sum_F \mathbf{I}[A_F], \tag{16}$$

where the sum ranges over all F of the form $F = \{(k_1, \ell_1), \dots, (k_r, \ell_r)\}$, with all the k_j 's distinct, and all the ℓ_j 's distinct. There are $r!$ $\binom{n}{r}^2$ of these sets F , and hence

$$\mathbf{E} X_r = r! \binom{n}{r}^2 p^r \delta^{\binom{r}{2}}.$$

Elementary calculus shows that, for fixed $r \geq 2$, $p \mapsto r! \binom{n}{r}^2 p^r \delta^{\binom{r}{2}}$ is increasing on $[0, 1/\sqrt{r}]$ and decreasing on $[1/\sqrt{r}, 1]$ (see the proof of the (a)-part of Lemma 10). The following lemma describes for which values of r the expectation $\mathbf{E} X_r$ tends to 0 or infinity, resp., in the relevant range of p .

- Lemma 10.** (a) If $e/n \leq p \leq n^{-1/2} \sqrt{\ln n}$, then $\mathbf{E} X_n \rightarrow \infty$.
(b) For constants $c > 1$, $\varepsilon > 0$ if $p = cn^{-1/2} \sqrt{\ln n}$, with $r := (1 + \varepsilon) \frac{n}{c^2}$ we have $\mathbf{E} X_r \rightarrow 0$.
(c) If $p \gg n^{-1/2} \sqrt{\ln n}$ and $1 - p = \bar{p} \geq n^{-o(1)}$, letting

$$r_- := 2 \log_{1/\delta}(pn^2) - 2 \log_{1/\delta} \log_{1/\delta}(pn^2) \text{ and} \\ r_+ := 2 \log_{1/\delta}(pn^2)$$

- we have $\mathbf{E} X_{r_-} \rightarrow \infty$, and $\mathbf{E} X_{r_+} \rightarrow 0$.
(d) If $a \in]0, 4[$ is a constant and $1 - p = \bar{p} = n^{-a}$, then $\mathbf{E} X_r \rightarrow 0$ if $r > 4/a + 1$, and $\mathbf{E} X_r \rightarrow \infty$ if $r < 4/a + 1$.

Proof.

- (a). First of all, we prove that for $r \geq 2$, the function $p \mapsto \mathbf{E} X_{r,p}$ is non-decreasing $]0, r^{-1/2}]$ and non-increasing on $[r^{-1/2}, 1[$.

Clearly, only the function

$$f: p \mapsto p(1 - p^2)^{(r-1)/2}$$

is of interest. Taking the derivative, we obtain

$$f'(p) = (1 - p^2)^{(r-1)/2} - (r-1)p^2(1 - p^2)^{(r-3)/2}.$$

If $0 < p < 1$, then $f'(p) = 0$ and is equivalent to

$$0 = 1 - p^2 - (r-1)p^2 = 1 - rp^2.$$

For $p < 1/\sqrt{r}$, we have $f'(p) > 0$ and $p > 1/\sqrt{r}$, we have $f'(p) < 0$.

Now, for $p = e/n$, using Stirling's formula, we have

$$\mathbf{E} X_n = n! \left(\frac{e}{n}\right)^n \left(1 - \frac{e^2}{n^2}\right)^{\binom{n}{2}} = \Theta(\sqrt{n}),$$

so $\mathbf{E} X_n$ tends to infinity with $n \rightarrow \infty$.

Finally, let $p = n^{-1/2} \sqrt{\ln n}$. We have

$$\frac{\ln \mathbf{E} X_n}{n} = \frac{\ln(n! p^n \delta^{\binom{n}{2}})}{n} = -1 + o(1) + \ln n - \ln(1/p) - \frac{n-1}{2} \ln(1/\delta) \\ \geq -1 + o(1) + \ln n - \ln(1/p) - \frac{n-1}{2} p^2,$$

where we used $\ln(1/\delta) = \ln(1/(1 - p^2)) \leq p^2 + O(p^4)$ and $np^4 = o(1)$ in the last inequality. Replacing p , we get

$$\frac{\ln \mathbf{E} X_n}{n} \geq \frac{1}{2} \ln \ln n + O(1),$$

which proves the claim in (a) for this particular value of p .

(b). First of all, note that, for $4 \leq r < n$, using the estimates

$$\sqrt{r} \left(\frac{r}{e}\right)^r \leq r! \leq r \left(\frac{r}{e}\right)^r, \text{ and}$$

$$\frac{1}{3\sqrt{r}} e^{r-r^2/(n-r)} \left(\frac{n}{r}\right)^r \leq \binom{n}{r} \leq e^r \left(\frac{n}{r}\right)^r,$$

we have

$$1 - \frac{r}{n-r} - O\left(\frac{\ln r}{r}\right) \leq \frac{\ln(r! \binom{n}{r}^2 p^r \delta^{\binom{r}{2}})}{r} - \left(\ln(n^2) - \ln(1/p) - \frac{r-1}{2} \ln(1/\delta) - \ln r \right) \leq 1 + \frac{\ln r}{r}. \quad (*)$$

(We will use this for (c), too.)

Now, with $c > 1$, $p = cn^{-1/2} \sqrt{\ln n}$ and $r = (1 + \varepsilon)n/c^2 = (1 - \Omega(1))n$, we get

$$\begin{aligned} \frac{\ln \mathbf{E} X_r}{r} &= \ln(n^2) - \ln(1/p) - \frac{r-1}{2} \ln(1/\delta) - \ln r + O(1) \\ &= \ln(n^2) - \frac{1}{2} \ln(n/\ln n) - \frac{r-1}{2} \ln(1/\delta) - \ln n + O(1) \\ &= \frac{1}{2} \ln n - \frac{r-1}{2} \ln(1/\delta) + O(\ln \ln n) \\ &= \frac{1}{2} \ln n - \frac{r-1}{2} (p^2 + O(p^4)) + O(\ln \ln n) \\ &= -\frac{\varepsilon}{2} \ln n + O(\ln \ln \ln n), \end{aligned}$$

which proves $\mathbf{E} X_r \rightarrow 0$.

(c). With $r := r_+ = 2 \ln(pn^2)/\ln(1/\delta)$, using the upper bound from (*), we get

$$\begin{aligned} \frac{\ln \mathbf{E} X_r}{r} &\leq \ln(pn^2) - \frac{r-1}{2} \ln(1/\delta) - \ln r + 1 + \frac{\ln r}{r} \\ &= \frac{1}{2} \ln(1/\delta) - \ln r + 1 + \frac{\ln r}{r} \\ &= -\Omega(1), \end{aligned}$$

where the last equation follows from $r \rightarrow \infty$ (due to $\bar{p} \geq n^{-o(1)}$), which also implies $\mathbf{E} X_r \rightarrow 0$.

On the other hand, with $r := r_- = 2 \log_{1/\delta}(pn^2) - 2 \log_{1/\delta} \log_{1/\delta}(pn^2)$, using the upper bound from (*), we get

$$\begin{aligned} \frac{\ln \mathbf{E} X_r}{r} &\geq \ln(pn^2) - \frac{r-1}{2} \ln(1/\delta) - \ln r + 1 - O\left(\frac{\ln r}{r}\right) \\ &\geq (\log_{1/\delta} \log_{1/\delta}(pn^2)) \ln(1/\delta) - \ln r + 1 + O\left(\frac{\ln r}{r}\right) \\ &= \ln \log_{1/\delta}(pn^2) - \ln r + 1 + O\left(\frac{\ln r}{r}\right) \\ &\geq -\ln 2 + 1 + O\left(\frac{\ln r}{r}\right) \\ &= \Omega(1). \end{aligned}$$

Again, the last equation and the conclusion $\mathbf{E} X_r \rightarrow \infty$ follows from $r \rightarrow \infty$.

(d). Finally, let $0 < a < 4$ be a constant and $1 - p = \bar{p} = n^{-a}$. Noting that $\delta = (1 + p)\bar{p} = \Theta(\bar{p})$, if $r = O(1)$, we have

$$\left(\mathbf{E} X_r\right)^{1/r} = \Theta\left(n^2 \bar{p}^{(r-1)/2}\right) = \Theta\left(n^{2-a(r-1)/2}\right),$$

which implies $\mathbf{E} X_r \rightarrow \infty$ if $r > 4/a + 1$, and $\mathbf{E} X_r \rightarrow 0$ if $r < 4/a + 1$.

From this lemma, we immediately get the upper bound on $F(f)$ in Theorem 3.1(c).

Proof (Proof of Theorem 3.1(c)). Follows from (c).

Item (a) of the lemma suggests the question, for which p the value of $F(f)$ drops from $(1 - o(1))n$ to $(1 - \Omega(1))n$. If the expectation is “right”, this happens crossing from $p = \sqrt{(\ln n)/n}$ to $p = (1 + \varepsilon)\sqrt{(\ln n)/n}$. This is supported by the fact that our lower bounds in this region—in the next subsection—appear to be quite simple, in that they only consider one fixed maximal matching in H_f , and delete edges from it until it becomes cross free.

B.2 Second moment calculation

Lemma 11. *If $r = O(1)$ and $p\delta \gg 1/n$, then $\mathbf{Var}(X_r) = o\left(\left(\mathbf{E} X_r\right)^2\right)$.*

Proof. With the notations as in equation (16), let $F_0 := \{(1, 1), \dots, (r, r)\}$, and abbreviate $A_0 := A_{F_0}$. We have

$$\mathbf{E}(X^2) = \mathbf{E} X \cdot \sum_F \mathbf{P}(A_F | A_0)$$

where the sum ranges over all F of the form $F = \{(k_1, \ell_1), \dots, (k_r, \ell_r)\}$, with all the k_j 's distinct, and all the ℓ_j 's distinct, as in (16).

If $F \subset \{r + 1, \dots, n\} \times \{1, \dots, n\}$, then the events A_F and A_0 are clearly independent, so that, with the following sum ranging over these F , we have

$$\sum_F \mathbf{P}(A_F | A_0) = \frac{(n - r)_r}{(n)_r} \mathbf{E} X.$$

Consequently, we have

$$\mathbf{E}(X^2) = \frac{(n - r)_r}{(n)_r} (\mathbf{E} X)^2 + \mathbf{E} X \cdot \sum_F \mathbf{P}(A_F | A_0),$$

where the last sum ranges over all F with $F \cap \{1, \dots, r\} \times \{1, \dots, n\} \neq \emptyset$. For each such F ,

$$\mathbf{P}(A_F | A_0) = O\left(\frac{1}{p\delta n}\right)^{O(r^2)} \mathbf{P}(A_F),$$

with absolute constants in the big- O s.

Hence, if $r = O(1)$ and $p\delta \gg 1/n$,

$$\mathbf{E}(X^2) = \frac{\binom{n}{r}}{(n-r)_r} (\mathbf{E}X)^2 + O\left(\frac{1}{p\delta n}\right)^{O(r^2)} (\mathbf{E}X)^2 = (1 + o(1))(\mathbf{E}X).$$

This proves the statement of the lemma.

Proof (Proof of Theorem 3.1(d)). The upper bound, for general a is in Lemma 10(d). The lower bound when $a < 1$ follows from Lemma 10(d) and Lemma 11.

B.3 Lower bounds: Cross-free sub-matchings

Let $\nu^\times(\cdot)$ denote the size largest cross-free matching of a bipartite graph.

Let H be a bipartite graph, and $m = \{e_1, \dots, e_r\} \subseteq E(H)$ a matching in H . Define the graph $G' = G'(H, m)$ with vertex set $V(G') = \{1, \dots, r\}$ and $\{k, \ell\} \in E(G')$ if e_k, e_ℓ induce a $K_{2,2}$ in H . Then $\nu^\times(H) \geq \alpha(G')$ holds: for any stable set A of G' , the set $\{e_j \mid j \in A\}$ is a cross-free matching in H .

Our strategy for obtaining a large cross-free matching will be this: fix a large matching m in H_f , then find a large stable set in the corresponding random graph $G'_{n,p}(m) := G'(H_f, m)$. This random graph behaves similarly to an Erdős-Renyi random graph with $|m|$ vertices and edge-probability p^2 . The following technical lemma takes care of the dependency issues which arise.

Let $\mathbf{G}_{r,q}$ denote the Erdős-Renyi random graph with r vertices and edge probability q .

Lemma 12. *For all positive integers n, r, a , and $p \in [0, 1]$, we have*

$$\mathbf{P}\left(\nu^\times(H_f) < a \quad \& \quad \nu(H_f) \geq r\right) \leq \mathbf{P}(\alpha(\mathbf{G}_{r,p^2}) < a).$$

Proof. Let \mathcal{M} be the set of matchings of size r of $K_{n,n}$, and for each $m \in \mathcal{M}$ denote by C_m the event that H_f contains m . Fix a matching $m \in \mathcal{M}$. For every edge $e \in E(K_{n,n})$, we have

$$\mathbf{P}(e \in H_f \mid C_m) = p,$$

and these events are jointly independent. Hence, for each potential edge e' of $G'_{n,p}(m)$,

$$\mathbf{P}(e' \in G'_{n,p}(m) \mid C_m) = p^2,$$

again with joint independence of the events.

Now, denote by A the event that there does not exist a cross-free matching of size larger than a in H_f . By the discussion above, A and C_m together imply $\alpha(G'_{n,p}(m)) < a$, so that

$$\mathbf{P}(A \mid C_m) \leq \mathbf{P}(\alpha(G'_{n,p}(m)) < a \mid C_m) = \mathbf{P}(\alpha(\mathbf{G}_{r,p^2}) < a).$$

It follows that

$$\begin{aligned} \mathbf{P}\left(\nu^\times(H_f) < a \quad \& \quad \nu(H_f) \geq r\right) &= \mathbf{P}\left(A \cap \bigcup_m C_m\right) \leq \sum_m \mathbf{P}(A \cap C_m) \\ &= \sum_m \mathbf{P}(A \mid C_m) \mathbf{P}(C_m) \leq \mathbf{P}(\alpha(\mathbf{G}_{r,p^2}) < a), \end{aligned}$$

which concludes the proof of the lemma.

Remark 3. We will use Lemma 12 in the following way: If p , r_- , r_+ are such that both

$$\begin{aligned} \mathbf{P}(\nu(H_f) < r_+) &= o(1), \quad \text{and} \\ \mathbf{P}(\alpha(\mathbf{G}_{r_+,p^2}) < r_-) &= o(1), \end{aligned} \tag{17}$$

then, a.a.s., f has a fooling set of size r_- . Indeed,

$$\begin{aligned} &\mathbf{P}(F(f) < r_-) \\ &\leq \mathbf{P}\left(\nu^\times(H_f) < r_- \quad \& \quad \nu(H_f) \geq r_+\right) + \mathbf{P}(\nu(H_f) < r_+) \\ &\leq \mathbf{P}(\alpha(\mathbf{G}_{r_+,p^2}) < r_-) + \mathbf{P}(\nu(H_f) < r_+) && \text{[Lemma 12]} \\ &= o(1) + o(1) && \text{[by (17)].} \end{aligned}$$

We are now ready to prove the first two items of Theorem 3.1. We start with the easiest part.

Proof (Proof of Theorem 3.1(b)). This is a direct consequence of the remark with $r_- := a(p^2)$ and $r := n$, since, if $pn - \ln n \rightarrow \infty$, then $\nu(H_f) = n$, a.a.s. (e.g., [19, Thm 4.1]).

Proof (Proof of Theorem 3.1(a)). Let $\varepsilon > 0$ be a constant. Proceeding as in Remark 3, with $r_- := r$ and $r_+ := (1 + \varepsilon)r$, if both a.a.s. $\nu(H_f) \geq r$ and a.a.s. $\alpha(\mathbf{G}_{r,p^2}) \geq (1 - \varepsilon)r$, then, a.a.s.,

$$(1 - \varepsilon)\nu(H_f) \leq F(f) \leq \nu(H_f).$$

Letting ε tend to 0 then gives the desired result.

For $n^{-3/2} \leq p = o(n)$, a.a.s., the number of edges of \mathbf{G}_{n,p^2} is $o(1)$, and hence $\alpha(\mathbf{G}_{n,p^2}) = (1 - o(1))n$, while easy arguments show that a.a.s. $\nu(H_f) = \Omega(n)$ with concentration in a window of size $O(\sqrt{n})$. Hence the conditions (17) are satisfied.

For $p = \Omega(1/n)$, a classical result by Karp & Sipser [21] states that there is a function $h:]0, \infty[\rightarrow [0, 1]$ with $\lim_{c \rightarrow \infty} h(c) = 1$ such that if $p = c/n$, then, a.a.s., $\nu(H_f) = (1 - o(1))h(c)/n$. Since $p = o(1/\sqrt{n})$, a.a.s., the number of edges of \mathbf{G}_{n,p^2} is $o(n)$, and hence $\alpha(\mathbf{G}_{n,p^2}) = (1 - o(1))n$. It follows that $F(f) = (1 - o(1))\nu(H_f)$. In particular, if $p \gg 1/n$, then, a.a.s., $\nu(H_f) = (1 - o(1))n$.

C Proofs for Section 4

C.1 The “usual calculation”

With

$$\alpha := \max\left(2\lambda, \frac{(1 + \varepsilon) \ln n}{\ln\left(\frac{\ln n}{e\lambda}\right)}\right),$$

we have to show that

$$\alpha \ln(\alpha/e\lambda) \geq \ln n.$$

We write it down informally. In the following list of inequalities, the each one is implied by the next one:

$$\begin{aligned} \alpha \ln(\alpha/e\lambda) &\geq \ln n && \text{[replace } \alpha \text{ by the 2nd term in the max]} \\ (1 + \varepsilon) \frac{\ln\left(\frac{\alpha}{e\lambda}\right)}{\ln\left(\frac{\ln n}{e\lambda}\right)} &\geq 1 \\ \alpha &\geq \ln^{1/(1+\varepsilon)} n \\ \frac{(1 + \varepsilon) \ln n}{\ln\left(\frac{\ln n}{e\lambda}\right)} &\geq \ln^{1/(1+\varepsilon)} n && \text{[is true.]} \end{aligned}$$

C.2 Chernoff

We have no good reference for the following simple Chernoff estimate (it is almost exactly Theorem 5.4 in [28], except that we allow $\lambda \rightarrow \infty$ slowly). For the sake of completeness, we include it here.

Lemma 13. *Let $\bar{p} = \lambda/n$ with $1 < \lambda = o(n)$, and $2\lambda \leq \alpha \leq n/2$. The probability that a $\text{Bin}(n, \bar{p})$ random variable is at least α is at most*

$$O(1/\sqrt{\alpha}) \cdot e^{-\lambda} \left(\frac{e\lambda}{\alpha}\right)^\alpha. \quad (18)$$

Proof (Proof of Lemma 13). Using Thm 1.1 in [3] (here we need the $\alpha \geq 2\lambda$), and the usual estimates for binomial coefficients, we find that said probability (for n sufficiently large) is at most an absolute constant times

$$\begin{aligned} \mathbf{P}\left(\text{Bin}(n, \bar{p}) = \alpha\right) &\leq \frac{1.1}{\sqrt{2\pi\alpha(n-\alpha)/n}} \left(\frac{\lambda}{\alpha}\right)^\alpha \left(\frac{n-\lambda}{n-\alpha}\right)^{n-\alpha} \\ &\leq \frac{1}{\sqrt{\alpha}} \left(\frac{\lambda}{\alpha}\right)^\alpha \left(1 - \frac{\alpha-\lambda}{n-\alpha}\right)^{n-\alpha} \leq \frac{1}{\sqrt{\alpha}} \left(\frac{\lambda}{\alpha}\right)^\alpha e^{\alpha-\lambda}, \end{aligned}$$

as promised.

C.3 Number of distinct rows

Proof (Proof of Lemma 1). For notational convenience, for $k = 1, \dots, n$, let

$$S_k := \{\ell \mid M_{k,\ell} = 0\}$$

The S_k are random sets, where the events $\ell \in S_k$ are all independent and have probability \bar{p} . For $m \geq 0$, with $\mathbf{0} := \{1, \dots, n\}$, denoting by

$$X_m := |\{S_1, \dots, S_m\} \setminus \{\mathbf{0}\}|,$$

the number of distinct non-zero rows among the first m rows of f , we need to show that $X_n = \Omega(n)$. This is quite easy for $\bar{p} = \Omega(1/n)$, i.e., Item (a). Here, we just prove it in the case that $\bar{p} \leq 1/2n$, i.e., Item (b).

Denote by A_{m+1} the event that the $(m+1)$ st row is zero or a duplicate of the first m rows, i.e., that

$$S_{m+1} \in \{\mathbf{0}, S_1, \dots, S_m\}.$$

We enumerate the distinct sets: $\{S_1, \dots, S_m\} = \{S_{k_1}, \dots, S_{k_{X_m}}\}$. Now, for $m \geq 1$, we have

$$\begin{aligned} \mathbf{P}(A_{m+1} \mid |S_1|, \dots, |S_m|, X_m) &= \mathbf{P}(S_{m+1} \in \{\mathbf{0}, S_1, \dots, S_m\} \mid |S_1|, \dots, |S_m|, X_m) \\ &= \mathbf{P}(S_{m+1} = \mathbf{0}) + \sum_{j=1}^{X_m} \mathbf{P}(S_{m+1} = S_{k_j} \mid |S_1|, \dots, |S_m|, X_m) \\ &= \bar{p}^n + \sum_{j=1}^{X_m} \bar{p}^{|S_{k_j}|} p^{n-|S_{k_j}|} \leq \bar{p}^n + p^n + \max(0, X_m - 1) \bar{p} p^{n-1}, \end{aligned}$$

where the last inequality comes from the fact that, since the S_{k_j} are all distinct, at most one of them has cardinality 0. Hence, for $m \geq 2$,

$$\begin{aligned} \mathbf{P}(A_{m+1} \mid X_m, X_1 = 1) &\leq \bar{p}^n + p^n + (X_m - 1) \bar{p} p^{n-1} \\ &\leq \bar{p}^n + p^n - \bar{p} p^{n-1} + \bar{p} p^{n-1} X_m. \end{aligned}$$

Now, for $m \geq 1$,

$$\begin{aligned} \mathbf{E}(X_{m+1} \mid X_m, X_1 = 1) &= X_m + 1 - \mathbf{P}(A_{m+1} \mid X_m, X_1 = 1), \\ &\geq X_m + 1 - \bar{p}^n - p^n + \bar{p} p^{n-1} - \bar{p} p^{n-1} X_m \\ &= 1 + \bar{p} p^{n-1} - \bar{p}^n - p^n + (1 - \bar{p} p^{n-1}) X_m. \end{aligned}$$

Using the law of total probability and solving the recursion¹, we find that

$$\mathbf{E}(X_m \mid X_1 = 1) \geq (1 + \bar{p} p^{n-1} - \bar{p}^n - p^n) \frac{1 - (1 - \bar{p} p^{n-1})^{m-2}}{\bar{p} p^{n-1}} + (1 - \bar{p} p^{n-1})^{m-1}$$

¹ The recursion: $\mu_{m+1} = \alpha + \beta \mu_m = \dots = \alpha \sum_{j=0}^{m-1} \beta^j + \beta^m \mu_1 = \alpha \frac{1 - \beta^m}{1 - \beta} + \beta^m \mu_1$.

With $\lambda := \bar{p}n$, again, note that, since, by our assumption above, $\lambda \leq 1/2$, using the Bernoulli inequalities $1 - tn \leq (1 - t)^n \leq 1 - tn + t^2 \binom{n}{2}$ for $t < 1$, we have

$$\frac{1}{2} \leq 1 - \lambda \leq p^n \leq p^{n-1} \leq 1 - \lambda \left(\frac{n-1}{n} + \lambda \frac{n-1}{n} \right) \leq 1,$$

so that

$$(1 - \bar{p}p^{n-1})^{m-2} \leq (1 - \bar{p}/2)^{m-2} \leq 1 - \frac{\lambda}{2} \left(\frac{m-2}{n} + \frac{\lambda}{2} \frac{m-2}{n} \right).$$

We conclude that, for $m = n$,

$$\begin{aligned} \mathbf{E}(X_m \mid X_1 = 1) &\geq (1 - p^n) \frac{1 - (1 - \bar{p}p^{n-1})^{m-2}}{\bar{p}p^{n-1}} \\ &\geq \lambda \left(\frac{n-1}{n} + \lambda \frac{n-1}{n} \right) \cdot \frac{\frac{\lambda}{2} \left(\frac{m-2}{n} + \frac{\lambda}{2} \frac{m-2}{n} \right)}{\lambda/n} \geq (1 + o(1)) \frac{\lambda n}{2}. \end{aligned}$$

Since $\mathbf{P}(X_1 = 1) = \mathbf{P}(S_1 = \mathbf{0}) = (1 - \bar{p}^n) = 1 - o(1)$, this implies $\mathbf{E}X_n \geq \mathbf{E}(X_n \mid X_1 = 1) \mathbf{P}(X_1 = 1) \geq (1 - o(1)) \lambda n / 2$.

To obtain the a.a.s. statement from the one about the expectation, we use the usual Martingale-based concentration bound (Corollary 2.27 in [19]): as changing one row can affect X_n by at most 1, we get

$$\mathbf{P}(X_n \leq \lambda n / 4) \leq \mathbf{P}(X_n \leq \mathbf{E}X_n - \lambda n / 4) \leq e^{-(\lambda n)^2 / 32n} = e^{-\Omega(\lambda^2 n)} = o(1),$$

where the last equation follows from the condition $n^{-3/2} = o(\bar{p})$.

Appendix B. ON THE GRAPH OF THE PEDIGREE POLYTOPE

The following is copied from [63].

On the Graph of the Pedigree Polytope

Abdullah Makkeh, Mozghan Pourmoradnasseri, Dirk Oliver Theis*

Institute of Computer Science of the University of Tartu
Ülikooli 17, 51014 Tartu, Estonia
{mozghan, dotheis}@ut.ee

November 29, 2016

Abstract

Pedigree polytopes are extensions of the classical Symmetric Traveling Salesman Problem polytopes whose graphs (1-skeletons) contain the TSP polytope graphs as spanning subgraphs.

While deciding adjacency of vertices in TSP polytopes is coNP-complete, Arthanari has given a combinatorial (polynomially decidable) characterization of adjacency in Pedigree polytopes. Based on this characterization, we study the graphs of Pedigree polytopes asymptotically, for large numbers of cities.

Unlike TSP polytope graphs, which are vertex transitive, Pedigree graphs are not even regular. Using an “adjacency game” to handle Arthanari’s intricate inductive characterization of adjacency, we prove that the minimum degree is asymptotically equal to the number of vertices, i.e., the graph is “asymptotically almost complete”.

Keywords: Traveling Salesman Polytopes, Probabilistic Combinatorics, Extensions of Polytopes, 1-Skeletons/Graphs of Polytopes.

1 Introduction

The graph (1-skeleton) of a polytope has as its vertices (edges) the vertices (edges) of the polytope. The most venerable result on graphs on polytopes: Steinitz’s Theorem states that 3-connected planar graphs are precisely the graphs of 3-dimensional polytopes.

Properties of graphs of polytopes of higher dimension are of interest not only in the combinatorial study of polytopes, but also in Combinatorial Optimization, and Theoretical Computer Science.

For example, the famous Hirsch conjecture in the combinatorial study of polytopes, settled by Santos [15], concerned the diameter of graphs of polytopes.

In Combinatorial Optimization, the study of the graphs of polytopes associated with combinatorial optimization problems was initially motivated by the search for algorithms for these problems.

In Theoretical Computer Science, the theorem by Papadimitriou [13] that Non-Adjacency of vertices of (Symmetric) Traveling Salesman Problem (TSP) polytopes is NP-complete, gave rise to similar results about other families of polytopes (cf. [1, 9] and the references therein, for recent examples).

There have been particularly many attempts to understand the graph of TSP polytopes, and, where this turned out to be infeasible, of TSP-related polytopes (e.g., [19];

*Supported by the Estonian Research Council, ETAG (*Eesti Teadusagentuur*), through PUT Exploratory Grant #620, and by the European Regional Development Fund through the Estonian Center of Excellence in Computer Science, EXCS.

cf. [6, 11, 20]). The presence of long cycles has been studied ([18], see also [12, 10]), as has the graph density / vertex degrees (e.g., [16], see also [8, 7]).

The original motivation for the research in this paper was a 1985 conjecture by Grötschel and Padberg [6] — well-known in polyhedral combinatorial optimization (Problem # 36 on Schrijver’s list [17]) — stating that the graph of TSP polytopes has diameter 2. Already in [6], Grötschel and Padberg extend the question for the diameter to a family of TSP-related polytopes which seemed easier to understand at the time.

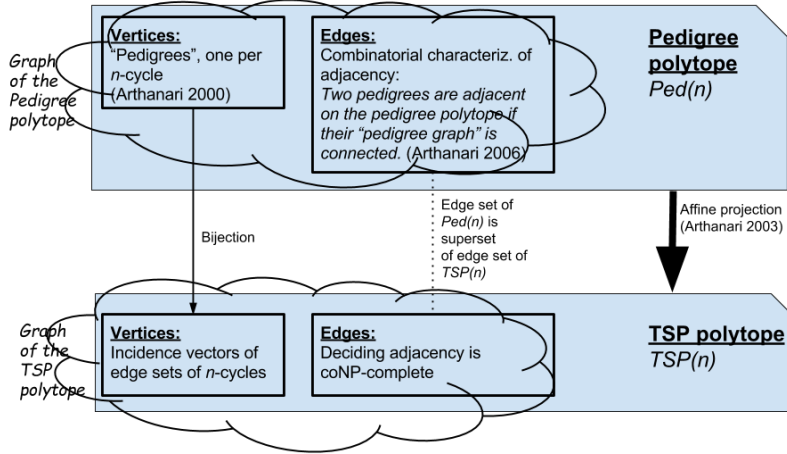


Figure 1: Polytopes and graphs

A more recent family of TSP-related polytopes are the *Pedigree polytopes* of Arthanari [4]. For this family of polytopes, adjacency of vertices can be decided in polynomial time [2]. Moreover, the graphs of the TSP polytopes are spanning subgraphs of the graphs of the Pedigree polytopes [3]. This follows as the Pedigree polytope for n cities is an *extension*, without “hidden” [14] vertices, of the TSP polytope [5]. The vertex set of the TSP polytope for n cities is in a natural bijection with the set of all cycles with node set $[n] := \{1, \dots, n\}$; the same is true for the vertex set of the Pedigree polytope for n cities. See Fig. 1.

The main result of our paper, is the following fact about graphs of Pedigree polytopes. Recall that the number of vertices of either the TSP polytope or the Pedigree polytope for n cities is the number of n -cycles, which is $(n - 1)!/2$.

Theorem 1. *The minimum degree of a vertex on the Pedigree polytope for n cities is $(1 - o(1)) \cdot (n - 1)!/2$ (for $n \rightarrow \infty$).*

In particular, the density graph of Pedigree polytopes is asymptotically equal to 1. Note, though, that while for TSP polytopes, these two statements are equivalent, this is not the case of Pedigree polytopes. The reason is that Pedigree polytopes are not as “symmetric” as TSP polytopes: For every two vertices u, v of the TSP polytope for n cities, there is an affine automorphism of the polytope mapping u to v . (Similar statements are true for monotone-TSP and graphical-TSP polytopes.) This is not true for Pedigree polytopes: Arthanari’s construction removes the symmetry to a large extent.

Numerical simulations show that, even for relatively large n (say, ≈ 100), the graph of the Pedigree polytope is not complete. We have made no attempt, however, to find a non-trivial upper bound for the minimum degree.

We now give a non-technical description of the proof of Theorem 1.

The adjacency game. Alice and Bob play a game on a graph Γ (which both of them see): Alice picks a vertex, then Bob picks a vertex; if the two vertices are adjacent, Bob wins, otherwise Alice wins. The game is silly, Bob always wins (unless there is an isolated vertex). Let us blindfold Bob — he can no longer see the graph or Alice’s move, so the best he can do is pick a random vertex. Alice will win the game with probability $\text{mindeg}(\Gamma)/|V(\Gamma)|$. The game is still trivial, but note that Alice’s chance of winning is linked to the minimum degree.

The game becomes interesting if the graph evolves over time. At each time n , each vertex of the current graph Γ_{n-1} will be replaced by a set of child vertices. Complex rules (known to Alice¹) govern whether or not a particular child vertex inherits the adjacency relation to a particular child of a neighbor of his parent. Also at each time, both Alice and Bob have to update their choices for vertices: Alice can pick any child of her current vertex; Bob, being blindfolded, picks a child of his current vertex uniformly at random.

Now, analyzing the game as a random process (asymptotically for $n \rightarrow \infty$) becomes a non-trivial task, which reveals the minimum degree of Γ_n , as $n \rightarrow \infty$. (Every vertex of Γ_n must have the same number of children for Bob’s random decisions to form a uniformly random vertex.)

We haven’t made clear what exactly Alice’s goal is, or how payout takes place, but these questions will fall in place naturally in our situation (for $n \rightarrow \infty$, with probability tending to 1, Alice will lose in every round).

Our graph Γ_n is the graph of the pedigree polytope for n cities, whose vertices are “pedigrees”. Strictly speaking, pedigrees are combinatorial-geometric objects defined by Arthanari, but to reduce technical overhead, we will work with cycles *only*. So, at time n , Alice and Bob each holds a cycle with node set $[n-1]$. A child-vertex is formed from a parent-vertex by inserting the new node n into the cycle. Alice picking one of these children amounts to inserting the new node n into her cycle; Bob inserts the new node n at a uniformly random position into his cycle.

As for the complex rules governing inheritance of adjacency: these are given by Arthanari’s characterization of adjacency on Pedigree polytopes [2], which are best understood as a process, updating a combinatorial structure at each time.

Pedigrees and how they are adjacent. Arthanari’s beautiful idea of a pedigree is that of a cycle “evolving” over time: Starting from the unique cycle with node set $\{1, 2, 3\}$ at time 3, at time $n \geq 4$, the node n is added to the cycle by subdividing one of its cycle-edges. We say that n is *inserted into* that cycle-edge.

Arthanari’s combinatorial condition for adjacency on the Pedigree polytope has process character, too, with a combinatorial structure, the *pedigree graph* G , evolving over time. Suppose we have two evolving cycles. Let us refer to A as Alice’s cycle, and to B as Bob’s cycle. At time n , Alice chooses a cycle-edge of her current cycle A (with node set $[n-1]$) and inserts her new node n into that cycle-edge to form her new cycle (with node set $[n]$). Then Bob chooses a cycle-edge of his current cycle B , and inserts his new node n into that cycle-edge to form his new cycle.

The pedigree graph G may also change at time n . The new pedigree graph is either equal to the current one, or arises from the current one by adding the new vertex² n with incident edges. The choices of Alice and Bob determine: whether the new vertex is added or not; the number of edges incident to the new vertex n ; the end vertices of these edges.

Arthanari’s combinatorial characterization of adjacency on the Pedigree polytope is now this.

¹Bob doesn’t need them since he plays randomly.

²Trying to reduce confusion, in our terminology, the polytopes have *vertices* which correspond to cycles; the cycles consist of *nodes* and *cycle-edges*, and the the pedigree graph of two cycles has *vertices*.

Theorem 2 ([2]). *At all times $n \geq 4$, the two vertices of the Pedigree polytope for n cities corresponding to the (new) cycles A and B with node set $[n]$ are adjacent in the Pedigree polytope, if, and only if, the (new) graph G is connected.*

Theorem 1 states that, if B is a cycle chosen uniformly at random from all cycles on $[n]$, then

$$\min_A \mathbb{P}(\{\text{the pedigree graph is connected}\}) = 1 - o(1),$$

where the minimum ranges over all cycles on $[n]$. Lower bounding this quantity amounts to studying the following adjacency game: Alice’s goal is to make the graph G disconnected; whereas Bob makes uniformly random choices all the time. We prove that Alice loses with probability $1 - o(1)$. To analyze the game, we study a kind of a Markov Decision Process with state space $\mathbb{Z}_+ \times \mathbb{Z}_+$. The states are pairs (s, t) , where s is the number of common cycle-edges in Alice’s and Bob’s cycles, and t is the number of connected components of the current pedigree graph.

In the next section, we will give rigorous statements corresponding to the hand-waving explanations above. In Section 3, we discuss some basic facts about Bob playing randomly, and discuss the intuition of the proof of the main. That section is followed by a more technical section containing the proofs of the basic properties of pedigree graphs, random or deterministic. In Section 5, we introduce the Markov-Decision-Problem-ish situation that Alice finds herself in. The proof of the main theorem is completed in Section 6. We conclude with a couple of questions for future research which we find compelling.

2 Exact Statements of Definitions, Facts, and Results

2.1 Cycles, One Node at a Time

Our cycles are undirected (so, e.g., there is only one cycle on 3 nodes). For ease of notation, let us say that the *positive direction* on a cycle with node set $[n]$, $n \geq 3$, is the one in which, when starting from the node 1, the node 2 comes before the node 3; the other direction the *negative direction*. When referring to the k th cycle-edge of a cycle, we count the cycle-edges in the positive direction; the 1st one being the one incident on node 1. E.g., in the unique cycle with node set $\{1, 2, 3\}$, the 1st cycle-edge is $\{1, 2\}$, the 2nd is $\{2, 3\}$, and the 3rd is $\{3, 1\}$.

As mentioned in the introduction, Arthanari’s Pedigree is a combinatorial object representing the “evolution” of a cycle “over time”, and the combinatorial definition of adjacency of pedigrees makes use of that step-by-step development. The set of Pedigrees is in bijection with the set of cycles. In our context (we do not have to associate points in space with Pedigrees), defining Pedigrees and then explaining the bijection with cycles is more cumbersome than necessary. For convenience, we use the following more convenient definitions, which mirror the definition of Pedigrees, but they use cycles only. Let us say that an *infinite cycle*³ is a sequence $A = c_n \in \prod_{n=3}^{\infty} [n]$. An infinite cycle A gives rise to an infinite sequence A_n of finite cycles (in the usual graph theory sense), defined inductively as follows:

- A_3 is the unique cycle with node set $\{1, 2, 3\}$;

³The reason why we use this notion of “infinite cycle” is pure convenience. It does not add complexity, but it makes many of statements and proofs less cumbersome. Indeed, instead of an infinite cycle, it is ok to just use a cycle whose length M is longer than all the lengths occurring in the particular argument. So instead of “let A be an infinite cycle, and consider A_k, A_ℓ, A_n ” you have to say “let M be a large enough integer, A_M a cycle of length M , and A_k, A_ℓ, A_n sub-cycles of A_M ”. All the little arguments (e.g., Fact 8 below) have to be done in the same way.

- for $n \geq 3$, A_n is the cycle with node set $[n]$ which arises from adding the node n to A_{n-1} by inserting it into (i.e., subdividing) the c_{n-1} th cycle-edge.

We think of A_\square as a cycle developing over time: At time n , the node n is added.

We will need to access the neighbors of node n in A_n , i.e., the ends of the cycle-edge into which n is inserted (i.e., which is subdivided) when n is added to A_\square . We write $\nu_A^+(n)$ for the neighbor of n in A_n following n in the positive direction, and $\nu_A^-(n)$ for the neighbor of n in A_n following n in the negative direction. The unordered pair $\nu_A(n) = \{\nu_A^+(n), \nu_A^-(n)\}$ is the c_{n-1} th cycle-edge of A_{n-1} , the one into which n was inserted.

These definitions are for $n \geq 4$ but extend naturally for $n = 1, 2, 3$: for $n = 3$ we let $\nu_A^+(3) = 1$, and $\nu_A^-(3) = 2$; for $n = 2$, we let $\nu_A^+(2) = \nu_A^-(2) = 1$. The equation $\nu_A(n) = \{\nu_A^+(n), \nu_A^-(n)\}$ holds for $n \geq 2$ (so $|\nu_A(2)| = 1$); for $n = 1$ we have $\nu_A(1) := \emptyset$.

Remark 3 (*Finding $\nu(k)$ for “old” nodes k*). It is readily verified directly from the definition, that, for $k \geq 2$, $\nu_A^\pm(k)$ can be found as follows: start from node k and walk in positive direction. The first node smaller than k which you encounter is $\nu_A^+(k)$. Similarly, if you walk in negative direction starting from k , the first node smaller than k which you hit, is $\nu_A^-(k)$.

A pair of nodes i, j split each cycle A_n , $n > i, j$ into two (open) segments (i, j do not belong to either segment). We say that the *segment between i and j* is the one which does *not* contain the node $\min(\{1, 2, 3\} \setminus \{i, j\})$ (i.e., 1, unless $1 \in \{i, j\}$, in that case, 2, unless $\{1, 2\} = \{i, j\}$, in that case 3). Note that this does not depend on the choice of $n > i, j$, which justifies to say “the segment of A_\square between i and j ”.

Remark 4 (*Testing/finding n with $\nu(n) = \{i, j\}$*). Given a pair of nodes $\{i, j\}$ and $n' > i, j$, there exists an $n \leq n'$ with $\nu_A(n) = \{i, j\}$ if, and only if, the segment between i and j on $A_{n'}$ is non empty and every node in it is larger than both i and j . In that case, the smallest node, n , in the segment between i and j on A_\square is the one with $\nu(n) = \{i, j\}$.

2.2 The Pedigree Graph

Two infinite cycles A, B give rise to a sequence of graphs G_\square^{AB} which we call the *pedigree graphs*. We omit the superscripted A, B when possible. We speak of *vertices* of the pedigree graphs (rather than nodes). We do this to avoid confusion between the nodes of the cycles A_\square, B_\square and the vertices of G_\square^{AB} , because the vertex set of G_n is a subset of $\{4, \dots, n\}$, and hence of the node set of A_n and B_n . So a node $k \in [n]$ may or may not be a vertex of G_n .

The pedigree graph G_{n-1} is the subgraph of G_n induced by the vertices in $[n-1]$. In other words, G_n is either equal to G_{n-1} (if n is not a vertex), or it arises from G_{n-1} by adding the vertex n together with edges between n and vertices in $[n-1]$.

Example 5. G_1, G_2, G_3 are graphs without vertices. G_4 may be a graph without vertices, or it may consist of a single isolated vertex 4. G_5 could be a graph without vertices; a graph with a single vertex 5; a graph with two isolated vertices 4, 5, or a graph with two vertices 4, 5, linked by an edge. Check figure 2 for possible G_4 and G_5 .

According to Arthanari [2, 3] the condition for the existence of vertices is the following:

- (1) A node $n \in [n]$ is a vertex of G_n , iff $\nu_A(n) \neq \nu_B(n)$.

There are several conditions for the presence of edges between the vertex n and earlier vertices. To make it easier to distinguish these, we speak of edge “types” and give the edges implicit “directions:” from A to B or from B to A . Here are the conditions for edges from n to earlier vertices.

- (2) There is a *type-1 edge* “from A to B ” between n and $k \in [n-1]$, if $\nu_A(n) = \nu_B(k)$. (Note that the condition implies that k is a vertex.)
- (3) There is a *type-1 edge* “from B to A ” Ditto, with A and B exchanged.
- (4) There is a *type-2 edge* “from A to B ” between n and $\ell := \max \nu_A(n)$, unless $\nu_B(\ell) \cap \nu_A(n) \neq \emptyset$. In other words, suppose the node n was inserted into the cycle-edge $\{k, \ell\}$ in A , with $k < \ell$. Now look up the end-nodes of the cycle-edge $\nu_B(\ell)$ into which ℓ was inserted when it was added to B . Unless k coincides with one of these end nodes, there is an edge between n and ℓ .
- (5) *Type-2 edge* “from B to A ” Ditto, with A and B exchanged.

Arthanari’s theorem [2] (Theorem 2) states that, if $n \geq 4$, and A_n, B_n are two cycles with node set $[n]$, then the two vertices of the Pedigree polytope (for n cities) corresponding to A_n and B_n are adjacent, if, and only if, G_n^{AB} is connected.

We will always think of A as “Alice’s cycle” and B as “Bob’s cycle”.

Example 6. Going through an example will help understand the definition of a pedigree graph. Figure 2 shows two cycles A and B evolving over time $n = 3, \dots, 10$, together with the evolving pedigree graph G_n^{AB} .

$n = 3$: As mentioned above, G_3^{AB} is a graph without vertices.

$n = 4$: Alice inserts her new node 4 between into the cycle-edge $\{1, 2\}$ of her cycle A_3 ; Bob inserts his new node 4 into the cycle-edge $\{1, 3\}$ of his cycle B_3 . Hence, $\{1, 2\} = \nu_A(4) \neq \nu_B(4) = \{1, 3\}$, so vertex 4 is added to G_3^{AB} .

$n = 5$: Alice inserts her new node 5 into the cycle-edge $\{2, 4\}$ of her cycle A_4 ; Bob inserts his new node 5 into the cycle-edge $\{1, 2\}$ of his cycle B_4 . Since $\{2, 4\} = \nu_A(5) \neq \nu_B(5) = \{1, 2\}$, vertex 5 is added to G_4^{AB} . Let us check the edges:

- In B_4 , the segment between 2 and 4 contains the node 3 which is smaller than 4. By Remark 4, there is no k with $\nu_A(5) = \nu_B(k)$, and hence no type-1 edge from A to B at this time.
- As $\nu_B(5) = \nu_A(4)$, there is a type-1 edge between 4 and 5 from B to A .
- Since $\max \nu_A(5) = 4$ and $\nu_B(4) = \{1, 3\} \not\ni 2$, there is also a type-2 edge between 5 and 4 from A to B .
- Since $\max \nu_B(5) = 2$ and $\nu_A(2) = \{1\} \ni 1$, there is no type-2 edge incident to 5 from B to A .

$n = 6$: Alice inserts her new node 6 into the cycle-edge $\{2, 3\}$ of her cycle, Bob inserts his new node 6 into the cycle-edge $\{2, 3\}$ of his cycle. Since $\{2, 3\} = \nu_A(6) = \nu_B(6) = \{2, 3\}$, we don’t have a vertex 6 in G_5^{AB} .

$n = 7$: Alice throws into $\{4, 5\}$, Bob throws into $\{3, 4\}$. Since $\{4, 5\} = \nu_A(7) \neq \nu_B(7) = \{3, 4\}$, the vertex 7 is added to G_6^{AB} .

- In B_6 , the segment between 4 and 5 contain nodes 3 and 2 which are smaller than 5. By Remark 4, there is no k with $\nu_A(7) = \nu_B(k)$, and hence no type-1 edge from A to B at this time.
- In A_6 , the segment between 3 and 4 contains the node 2 which is smaller than 4. Again by Remark 4, there is no k with $\nu_B(7) = \nu_A(k)$, and thus no type-1 edge from B to A .
- As $\max \nu_A(7) = 5$ and $\nu_B(5) = \{1, 2\} \not\ni 4$, we have a type-2 edge from A to B between 7 and 5.
- As $\max \nu_B(7) = 4$ and $\nu_A(4) = \{1, 2\} \not\ni 3$, there is also a type-2 edge from B to A between 7 and 4.






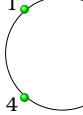
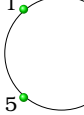
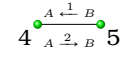


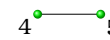


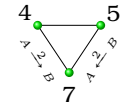


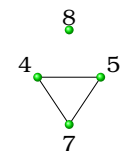


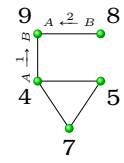


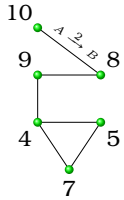
n	A_n	B_n	G_n^{AB}
3			$G_3^{AB} = \emptyset$
4			
5			
6			
7			
8			
9			
10			

Figure 2: Cycles A_n and B_n and corresponding G_n^{AB}

$n = 8$: Alice plays $\{3, 6\}$, Bob chooses $\{1, 4\}$. Since $\{3, 6\} = \nu_A(8) \neq \nu_B(8) = \{1, 4\}$, the vertex 8 is added to $G_7^{A,B}$.

- In B_7 , the segment between 3 and 6 is empty (just the cycle-edge). By Remark 4, there is no k with $\nu_A(8) = \nu_B(k)$, and hence no type-1 edge from A to B .
- For the same reason (segment between 1 and 4 empty), there is no type-1 edge from B to A incident to the vertex 8.
- $\max \nu_A(8) = 6$ and $\nu_B(6) = \{2, 3\} \ni 3$. So there is no type-2 edge from A to B between 8 and a smaller vertex.
- $\max \nu_B(8) = 4$ and $\nu_A(4) = \{1, 2\} \ni 1$. So there is no type-2 edge from B to A between 8 and a smaller vertex.

Hence, vertex 8 is isolated in G_8 .

$n = 9$: Alice chooses $\{1, 3\}$, Bob chooses $\{1, 8\}$. As $\{1, 3\} = \nu_A(9) \neq \nu_B(9) = \{1, 8\}$, the vertex 9 is added to $G_8^{A,B}$.

- As $\{1, 3\} = \nu_A(9) = \nu_B(4) = \{1, 3\}$, there is a type-1 edge from A to B between 9 and 4.
- The segment between 1 and 8, contains the node 3 which is smaller than 8 so there is no type-1 edge from B to A incident to the vertex 9.
- As $\max \nu_A(9) = 3$ and $\nu_B(3) = \{1, 2\} \ni 1$, there is no type-2 edge from A to B between 9 and 3.
- As $\max \nu_B(9) = 8$ and $\nu_A(8) = \{3, 6\} \not\ni 1$, there is a type-2 edge from B to A between 9 and 8.

$n = 10$: Alice chooses $\{3, 9\}$, Bob chooses $\{2, 6\}$. Since $\{3, 9\} = \nu_A(10) \neq \nu_B(10) = \{2, 6\}$, the vertex 10 is added to $G_9^{A,B}$.

- The segment between 3 and 9 in B , contains the node 4 which is smaller than 9 so there is no type-1 edge from A to B incident to the vertex 10.
- Again, in B , the segment between 3 and 9 has a vertex (4) smaller than 9. Remark 4 gives us that there is no k $\nu_B(k) = \nu_A(10)$, so no type-1 edge from B to A is created.
- As $\max \nu_A(10) = 9$ and $\nu_B(9) = \{1, 8\} \not\ni 3$, there is a type-2 edge from A to B between 10 and 9.
- As $\max \nu_B(10) = 6$ and $\nu_A(6) = \{2, 3\} \ni 2$, no type-2 edge from B to A is created.

2.3 Rephrasing Theorem 1

We now rephrase Theorem 1, in terms of the pedigree graph. We also unravel the little-o, and move to the ‘‘Alice-and-Bob’’ letters for the cycles.

Theorem 7 (Theorem 1, rephrased). *For every $\varepsilon > 0$ there is an integer N such that for all $n \geq N$ and all cycles A_n with node set $[n]$, if B_n is drawn uniformly at random from all cycles with node set $[n]$, then*

$$\mathbb{P}(G_n^{AB} \text{ is connected}) \geq 1 - \varepsilon.$$

In symbols, and using infinite cycles, this reads:

$$\forall \varepsilon > 0 \exists N: \forall A \forall n \geq N: \mathbb{P}(G_n^{AB} \text{ is connected}) \geq 1 - \varepsilon,$$

where the probability is taken over all infinite cycles, see the next section. A close look at our proof shows that we are actually proving the following stronger statement (we don't have any use for it, though):

$$\forall \varepsilon > 0 \exists N: \forall A: \mathbb{P}(\forall n \geq N: G_n^{AB} \text{ is connected}) \geq 1 - \varepsilon.$$

3 Pedigree Graphs of Random Cycles

We have to reconcile uniformly random cycles with the “evolution over time” concept of pedigrees. The definition of an infinite cycle makes that very convenient, just do the same as with infinite sequences of coin tosses: Take, as probability measure on the sample space $\prod_{n=1}^{\infty} [n]$ of all infinite cycles the product of the uniform probability measures on each of the sets $[n]$, $n \geq 3$. We refer to the atoms in this probability space as *random infinite cycles*. The following is a basic property of product probability spaces. We will use it mostly without mentioning it.

Fact 8. If B is a random infinite cycle, then, for each $n \geq 3$, the cycle B_n is uniformly random in the set of all cycles with node set $[n]$.

Creating isolated vertices. The first substantial result about the connectedness of the pedigree graph, concerns the creation of isolated vertices.

As outlined in the introduction, we study the situation in which Alice chooses her cycle-edge of A_{n-1} according to a sophisticated strategy, whereas Bob always chooses a uniformly random cycle-edge of B_{n-1} to insert his node n into (which amounts to his cycle B_n being uniformly random in the set of all cycles on $[n]$, by Fact 8). In this section, we adopt a purely “random graph” perspective. For fixed A and random B , the pedigree graphs G_{\circ}^{AB} are a sequence of random graphs, with some weirdo distribution: At time n , whether the new vertex n is added or not, and if it is, how many incident edges it has, and what their end vertices are — these are all random events/variables.

For deterministic A and random B , let the random variable Y count the total number of times that an isolated vertex of the pedigree graph is created. In other words, $Y = \sum_{n=4}^{\infty} \mathbf{1}_{I_n}$, where I_n denotes the event that, at time n , n is added as an isolated vertex to $G_n^{A,B}$ (and $\mathbf{1}_{\circ}$ is the indicator random variable of the event).

Lemma 9. *Whatever Alice does, $\mathbb{E}Y = 2$.*

Moreover, for every $\varepsilon > 0$, if $n_0 \geq 4/\varepsilon + 2$, then, whatever Alice does

$$\mathbb{P}\left(\bigcup_{n \geq n_0} I_n\right) \leq \varepsilon.$$

(We have collected the proofs for this section in the following section, in order not to interrupt the motivating explanations.)

To understand why the lemma is important, consider a pedigree graph at time n , just before Alice and Bob make their choices of cycle-edges into which their respective new nodes n are inserted. If n is not a vertex of the new pedigree graph G_n , the number of connected components of G_{\circ} doesn't change. If n is a vertex, and it does have incident edges, then the number of connected components can only decrease. The only way that the number of connected components of G_n can increase is if n is an isolated vertex in the new pedigree graph. Hence, Lemma 9 provides an upper bound on the expected number of connected components, uniform over n .

The Intuition. From Lemma 9, it is unlikely that the pedigree graph will have many components. Indeed, intuitively, if only 2 isolated vertices are ever created, that means

that most of the time either nothing happens (no new vertex) or edges are created, ultimately reducing the number of components, so the pedigree graph is connected.

While this basic intuition is essentially correct, a closer look reveals some subtleties. First of all, Alice has a big sway in choosing the end vertices of new edges: she can pick the end vertices of type-2 edges from A to B ; and she can influence the end vertices of type-1 edges (both directions).

Secondly, Bob's choices are reduced by the low degrees of the vertices. (A stronger version of (a) is proved as Lemma 12 in Section 4; we do not prove the rest of the lemma because we do not need it for our proof.)

Lemma 10. *The maximum degree of a vertex in a pedigree graph is at most 6:*

- (a) *up to 2 to vertices created in the past; and*
- (b) *up to 6 to future vertices.*

Hence, if a vertex n_0 was created as an isolated vertex or landed in a small connected component, Bob has only 4–6 shots at connecting it to another connected component. The good news is that Alice can never “shut down” a connected component completely: Bob can always extend it by one more vertex.

Lemma 11. *Let C be a connected component of the pedigree graph G_{n-1}^{AB} . There exists a $k \in C$ such that, no matter what Alice's move is at time n , Bob has a move which creates the vertex n and makes it adjacent to k .*

However, for Bob to make a disconnected pedigree graph connected, at some time, he will have to manage to insert his new node in such a way that it has two incident edges, linking two connected components at the same time.

There is no difficulty in realizing that Alice wouldn't stand a chance against a strategically playing Bob. But we claim that the game between a clever Alice and a blindfolded Bob will turn in Bob's favour almost all of the time.

Computer simulations give another indication that some care has to be taken implementing the basic intuition: Even for n as large as 100, even if Alice's cycle is chosen uniformly at random instead of adversarial, the frequency (in 100000 samples) with which we saw a connected pedigree graph was only about 84%. In the remaining 16% of cases, the typical situation is that of one giant connected component containing almost every vertex, and one tiny component growing only very slowly. This indicates that even a *disinterested* Alice can do some damage.

4 Basic Properties of the Pedigree Graph

Lemma 12. *Let A, B be two infinite cycles, and $n \geq 4$.*

- a. *If $\nu_A(n)$ is a cycle-edge of B_{n-1} , then, if n is a vertex, there is no edge in G_n^{AB} “from A to B ” incident on n .*
- b. *If $\nu_A(n)$ is not a cycle-edge of B_{n-1} , then n is a vertex, and the pedigree graph G_n^{AB} has an edge “from A to B ” incident on n :*
 - b.1. *There is a type-1 edge, if and only if every node in the segment on B_\square between $\nu_A^+(n)$ and $\nu_A^-(n)$ is larger than these two⁴. In this case, the other end of the type-1 edge is smallest node in the segment on B between $\nu_A^+(n)$ and $\nu_A^-(n)$.*
 - b.2. *There is a type-2 edge, if and only if there is a node in the segment on B_\square between $\nu_A^+(n)$ and $\nu_A^-(n)$ which is less than at least one of the two nodes. In this case, the other end of the type-2 edge is $\max(\nu_A^+(n), \nu_A^-(n))$.*

In particular, there can be at most one edge “from A to B ” between n and a vertex $k < n$.

⁴Remember that segments are “open”: they don't include the end nodes.

Proof. First of all, that n is a vertex follows immediately, as $\nu_A(n) = \nu_B(n)$ is not possible.

The statements about the edges follow immediately from the definitions of the edges of the pedigree graph: For the first item, use Remark 4; for the second item, by the definition of “segment between”, the two nodes $\nu_A^+(n)$ and $\nu_A^-(n)$ are separated on B by nodes smaller than themselves, so that none of them can be in the $\nu_B(\cdot)$ -set of the other, by Remark 3. \square

Lemma 13. *Let A, B be two infinite cycles, and $n \geq 4$. Then n is an isolated vertex in G_n if, and only if, both*

- (1) $\nu_A(n)$ is a cycle-edge of B_n , and
- (2) $\nu_B(n)$ is a cycle-edge of A_n .

Proof. That the conditions (1) and (2) are sufficient for n to be an isolated vertex is readily verified: Since $\nu_A(n)$ is a cycle-edge of B_n , it must still have been a cycle-edge of B_{n-1} ; similarly $\nu_B(n)$ was a cycle-edge of A_{n-1} . Type-1 edges are immediately excluded. As for type-2 edges, the condition $\nu_A(\max \nu_B(n)) \cap \nu_B(n)$ is trivially satisfied.

For the necessity, suppose (by symmetry) that $\nu_A(n)$ is not a cycle-edge of B_n . Then either it was a cycle-edge of A_\square at time $n-1$ and got destroyed, or it wasn't a cycle-edge of A_\square at time $n-1$ in the first place.

In the former case, the cycle-edge must have been destroyed by through $\nu_B(n) = \nu_A(n)$. But this means that n is not a vertex of the pedigree graph.

In the latter case, Lemma 12 applies, and n is a vertex, but it is not isolated. \square

4.1 Proof of Lemma 9

Recall from page 9 that I_n denotes the event that, at time n , the vertex n is added as an isolated vertex to the pedigree graph.

Lemma 14. *For $n \geq 4$, $\mathbb{P}(I_n) = \frac{4}{(n-1)(n-2)}$.*

Proof. For $x = 1, 2$, denote by E_x the event of condition (x) of Lemma 13. We need to compute $\mathbb{P}(E_1 \cap E_2) = \mathbb{P}(E_1)\mathbb{P}(E_2 | E_1)$. As for E_1 , we have computed this probability in the proof of Lemma 21: it is $2/(n-1)$.

As for $\mathbb{P}(E_2 | E_1)$, conditioning on $\nu_A(n)$ being a cycle-edge amounts to

1. identifying these two nodes to a (super-)node, leaving $n-1$ nodes to consider, and taking a uniformly random cycle on these $n-1$ nodes; and then
2. replacing the super-node by the cycle-edge $\nu_A(n)$, which requires deciding which of the two nodes comes first (in positive orientation).

To calculate the conditional probability $\mathbb{P}(E_2 | E_1)$, we go over all cycle-edges e of A_n , and find the conditional probability of the event that $\nu_B(n) = e$. Since these events are mutually exclusive, we then just add up the probabilities.

Of the $n-1$ nodes after forming the super-node, $n-2$ involve nodes different from n . Each one of the $\binom{n-2}{2}$ possible cycle-edges between these nodes has equal chance of ending up being the one chosen by Bob at time $n-1$. Let us start with the cycle-edges e of A_n with $\nu_A(n) \cap e = \emptyset$, i.e., they are not incident to the super-node. There are $n-4$ of them. For each one of them, we have

$$\mathbb{P}(e = \nu_B(n) | E_1) = 1 / \binom{n-2}{2}.$$

The probability that Bob's chosen cycle-edge is one of the two cycle-edges incident on the super-node is $2/\binom{n-2}{2}$. Let e be one of the two cycle-edges with $|\nu_A(n) \cap e| = 1$. The probability that the orientation of the cycle-edge $\nu_A(n)$ of B_n chosen in step (2) above is

so that Bob's chosen cycle-edge is equal to e is $1/2$. In total, we have, for each of these 2 edges e ,

$$\mathbb{P}(e = \nu_B(n) \mid E_1) = \frac{1/2}{\binom{n-2}{2}}.$$

We add up:

$$\mathbb{P}(E_2 \mid E_1) = \frac{n-4}{\binom{n-2}{2}} + \frac{1/2}{\binom{n-2}{2}} + \frac{1/2}{\binom{n-2}{2}} = \frac{n-3}{\binom{n-2}{2}} = \frac{2}{n-2}.$$

In total, we have

$$\mathbb{P}(I_n) = \mathbb{P}(E_1) \mathbb{P}(E_2 \mid E_1) = \frac{2}{n-1} \cdot \frac{2}{n-2} = \frac{4}{(n-1)(n-2)}.$$

□

We can now complete the proof of Lemma 9.

of Lemma 9. For the expectation, we calculate

$$\begin{aligned} \mathbb{E}Y &= \sum_{n=4}^{\infty} \mathbb{P}(I_n) \\ &= \sum_{n=4}^{\infty} \frac{4}{(n-1)(n-2)} && \text{[by Lemma 14]} \\ &= 4 \sum_{n=4}^{\infty} \left(\frac{1}{n-2} - \frac{1}{n-1} \right) \\ &= 4 \lim_{n \rightarrow \infty} (1/2 - 1/(n-1)) \\ &= 2. \end{aligned}$$

Similarly, for the statement about I_n , we find that

$$\mathbb{P}\left(\bigcup_{n \geq n_0} I_n\right) \leq \sum_{n=n_0}^{\infty} \mathbb{P}(I_n) = 4 \sum_{n=n_0}^{\infty} \left(\frac{1}{n-2} - \frac{1}{n-1} \right) = 4/(n_0 - 2) \leq \varepsilon.$$

This completes the proof of Lemma 9. □

4.2 Proof of Lemma 11

Proof of Lemma 11. Take $k := \max V(C)$. Since k is a vertex, we have $|\nu_B(k) \cap \nu_A(k)| \leq 1$. Suppose that $\nu_B^+(k) \notin \nu_A(k)$ (the other case is symmetric). Then, the first time Bob inserts a node, say n' , into the cycle-edge on the positive side of k , this will create a type-2 edge “from B to A ” between n' and k . Since k is the newest vertex in its component, Bob has not yet inserted a node there, so he can insert n there, now. □

5 The Adjacency Game

At each time, Alice moves first. As already explained, she determines the cycle A by choosing, at each time n , the cycle-edge of A_{n-1} into which her new node n will be inserted. Then Bob moves. He determines B in the same way, but (using Fact 8), he will draw the cycle-edge of B_{n-1} into which his new node n is inserted uniformly at random from all cycle-edges of B_{n-1} , and his choice is independent of his earlier choices.

We say that Bob wins, if there exists an n_0 such that for all $n \geq n_0$, the pedigree graph G_n^{AB} is connected. We need Bob to win “uniformly”, i.e., n_0 must not depend on Alice’s moves.

Let the random variable T_n denote the number of connected components in the pedigree graph G_n^{AB} . To analyze the development of the random process T_n , it turns out to be useful to consider a second random process, S_n . Denote by E_n^\cap the set of cycle-edges that Alice’s cycle and Bob’s cycles have in common,

$$E_n^\cap := E(A_n) \cap E(B_n),$$

and let

$$S_n := |E_n^\cap|$$

count the number of cycle-edges that Alice and Bob have in common. We will distinguish Alice’s moves by whether or not she chooses a common cycle-edge to place her new cycle node. The set $E_n^{\cap*}$ holds those common cycle-edges which are not incident on the cycle-edge which Alice chooses for her new node:

$$E_n^{\cap*} := \{e \in E_n^\cap \mid e \cap \nu_A(n+1) = \emptyset\};$$

we let S^* count the cycle-edges in $E^{\cap*}$:

$$S_n^* := |E_n^{\cap*}|.$$

Finally, denote by E_n^r the set of cycle-edges in Bob’s cycle which are neither common nor incident on Alice’s chosen cycle-edge:

$$E_n^r := \{e \in E(B_n) \setminus E_n^\cap \mid e \cap \nu_A(n+1) = \emptyset\}; \text{ and}$$

$$R_n := |E_n^r|.$$

We are now ready to state and prove the transition probabilities. They depend on whether Alice chooses, for her new node, a common cycle-edge — we refer to that as a c-move by Alice — or a cycle-edge which is in the difference $E(A_n) \setminus E_n^\cap$ — we call that a d-move.

Lemma 15. *The conditional probabilities*

$$\mathbb{P}(S_{n+1} = S_n + \Delta_S \wedge T_{n+1} = T_n + \Delta_T \mid B_n)$$

satisfy these bounds (entries not shown are “= 0”):

Δ_T				
+1	$= \frac{S_n^*}{n}$	$\leq \frac{2}{n}$		
0		$= \frac{R_n}{n}$	$\leq \frac{2}{n}$	$= \frac{1}{n}$
-1				
	-2	-1	0	+1
				Δ_S
	<i>c-move</i> (Alice chooses common cycle-edge)			

Δ_T				
+1				
0		$= \frac{S_n^*}{n}$	$\leq \frac{R_n - T_n + 1}{n}$	$\leq \frac{4}{n}$
-1			$\geq \frac{T_n - 1}{n}$	
	-2	-1	0	+1
				Δ_S
	<i>d-move</i> (Alice chooses cycle-edge in $E(A_n) \setminus E_n^\cap$)			

Proof. Let us start with the case that Alice makes a **c-move**, i.e., she chooses a common cycle-edge to insert her new node $n+1$ into. In symbols: $\nu_A(n+1) \in E_n^\cap$. In this case, there can be no edges “from A to B” incident to the vertex $n+1$ of G_{n+1}^{AB} . We split into disjoint events based on

$$H := |\nu_A(n+1) \cap \nu_B(n+1)|$$

Event $H = 2$. This happens if Bob's random choice of a cycle-edge for his new node $n + 1$ is the same as Alice's, i.e., $\nu_A(n + 1)$. The probability of his happening is $1/n$. In that case, both S and T are unchanged.

Event $H = 1$. We split into two sub-events, " $H = 1$ and $\nu_B(n + 1) \in E_n^\cap$ " and " $H = 1$ and $\nu_B(n + 1) \notin E_n^\cap$ ". Each has probability at most $2/n$.

- $H = 1$ and $\nu_B(n + 1) \in E_n^\cap$ implies that both $\nu_A(n + 1)$ and $\nu_B(n + 1)$ are still common cycle-edges after Alice and Bob have added their new nodes $n + 1$. By Lemma 13, this implies that $n + 1$ is an isolated vertex of G_{n+1} . We have $T_{n+1} = T_n + 1$, and $S_{n+1} = S_n - 1$ (two common cycle-edges destroyed, one new created).
- Suppose $H = 1$ and $\nu_B(n + 1) \notin E_n^\cap$. Simply from $\nu_B(n + 1) \notin E_n^\cap$, applying Lemma 12b. (with A/B exchanged), the existence of an edge incident on $n + 1$ "from B to A " follows. Moreover, that lemma also states that there is only one such edge. Hence, we have $T_{n+1} = T_n$, and $S_{n+1} = S_n$ (one common cycle-edge destroyed, one new created).

Event $H = 0$. We distinguish the same two sub-events as for $H = 0$: " $H = 0$ and $\nu_B(n + 1) \in E_n^\cap$ " and " $H = 0$ and $\nu_B(n + 1) \notin E_n^\cap$ ".

- The sub-event $H = 0$ and $\nu_B(n + 1) \in E_n^\cap$ occurs if Bob hits a cycle-edge in $E_n^{\cap*}$. The probability of that happening is S_n^*/n . By Lemma 12a., we do not have an edge in the pedigree graph incident on $n + 1$, so an isolated vertex is created. Hence $T_{n+1} = T_n + 1$, and $S_{n+1} = S_n - 2$ (two common cycle-edges destroyed, none new created).
- That $H = 0$ and $\nu_B(n + 1) \notin E_n^\cap$ means that that Bob hits a cycle-edge in E_n^r ; the probability of this happening is R_n/n . As in the sub-event above, applying Lemma 12b. gives the existence of exactly one edge "from B to A " incident on $n + 1$. Hence, we have $T_{n+1} = T_n$, and $S_{n+1} = S_n - 1$ (one common cycle cycle-edge destroyed, none new created).

This completely partitions the probability space, with probabilities corresponding to the bounds in the table.

Now let us consider the case that Alice makes a **d-move**, i.e., she chooses a cycle-edge which is not common to both cycles, to insert her new node $n + 1$ into. In symbols: $\nu_A(n + 1) \notin E_n^\cap$. By Lemma 12, this condition implies that $n + 1$ is a vertex of the pedigree graph G_{n+1}^{AB} and there is an edge "from A to B " incident on $n + 1$. This implies that $T_{n+1} \leq T_n$.

We distinguish cases based on the event $\nu_B(n + 1) \in E_n^\cap$.

Event $\nu_B(n + 1) \in E_n^\cap$. We split into three subevents, depending on H (as defined above):

- $\nu_B(n + 1) \in E_n^\cap$ and $H = 0$: This means that Bob hits a cycle-edge in $E_n^{\cap*}$. The probability of that happening is S_n^*/n . In that case, by Lemma 12a., there is no edge "from B to A ", and we have $T_{n+1} = T_n$. One common cycle-edge was destroyed, but none was created, so $S_{n+1} = S_n - 1$.
- $\nu_B(n + 1) \in E_n^\cap$ and $H = 1$: There are at most 2 cycle-edges in B_n which are both common with A_n and have a node in common with $\nu_A(n + 1)$. Hence, the probability that Bob hits one of them is $2/n$. Again, $T_{n+1} = T_n$. One common cycle-edge was destroyed, but another one is created, so $S_{n+1} = S_n$.

Note that $H = 2$ is impossible in a d-move.

Event $\nu_B(n + 1) \notin E_n^\cap$. Before we can proceed, we need to extract the following fact from the proof of Lemma 11:

Fact 16. If G_n^{AB} is not connected and C is a connected component, then, no matter what Alice's move at time $n + 1$, there is a cycle-edge $\{k, k'\}$ of B_n with the properties

- (a) $k' < k$ and $k \in C$;
- (b) if Bob inserts his new node $n + 1$ into that cycle-edge, there will be a type-2 edge "from B to A " between $n + 1$ and k .

We already know that the cycle-edge $\{k, k'\}$ is not in E_n^\cap (Lemma 12a.).

Of the T_n connected components at time n , one contains the end vertex of the edge "from A to B " incident on $n + 1$. For each of the remaining $T_n - 1$ components, take one cycle-edge as described in Fact 16. By property (a), all these cycle-edges are distinct. Hence, the chance that Bob hits one of them is exactly $(T_n - 1)/n$. There may be more possibilities for Bob to reduce the number of connected components, but, by the last sentence in Lemma 12, the number of components which can be joined at time $n + 1$ is at most 2.

Now we can proceed with the event $\nu_B(n + 1) \notin E_n^\cap$. We split into sub-events:

- $\nu_B(n + 1) \notin E_n^\cap$ and $T_{n+1} = T_n - 1$: By what we have discussed, the probability of this happening is at least $(T_n - 1)/n$. As for the change in S_n , no common cycle-edge is destroyed (Lemma 12a.), but it is possible that one is created, so $S_{n+1} \in \{S_n, S_n + 1\}$.
- $\nu_B(n + 1) \notin E_n^\cap$ and $T_{n+1} = T_n$: We split into sub-sub-events:
 - The above conditions and $H = 1$: The probability of that happening is at most $4/n$. Since no common cycle-edge is destroyed, but a new one is created, we have $S_{n+1} = S_n + 1$.
 - The above conditions and $H = 0$: For that to happen, Bob has to hit a cycle-edge in E_n^r . At least $T_n - 1$ of these cycle-edges lead to the situation $T_{n+1} = T_n - 1$, so the probability for this sub-sub-event is at least $\frac{R_n - T_n + 1}{n}$. No common cycle-edge is destroyed, and none is created, so $S_{n+1} = S_n$.

This completely partitions the probability space, with probabilities corresponding to the bounds in the table. \square

The proof of the main theorem now follows the following idea. From the tables in Lemma 15, you see that d-moves have chance of reducing the number of connected components — albeit a small one. Moreover, Alice cannot take a c-move only when $S_n > 0$, but c-moves have a strong tendency to reduce S_n . We prove that the number of d-moves that Alice has to take are frequent enough to lead to a decrease in the number of connected components. This suffices to prove Theorem 6, along the lines sketched on page 9.

The next section gives more details of the proof.

6 Proof of Theorem 7

Using the Azuma-Hoeffding super-martingale tail bound, we prove that, for large enough n_0 , Alice has to take many d-moves between times n_0 and $2n_0$.

Lemma 17. *For every $\varepsilon \in]0, 1[$, if $n_0 \geq \max(900, 8 \ln(1/\varepsilon))$, and $n_1 := 2n_0$ then, whatever Alice does, the probability, conditioned on B_{n_0} and $S_{n_0} \leq \ln^2 n_0$, that among her moves at times $n = n_0 + 1, \dots, n_1$, there are fewer than $n_0/3$ d-moves, is at most ε .*

Proof. Denote by $C \subseteq \{n_0 + 1, \dots, n_1\}$ the set of times in which Alice takes c-move; so Alice takes d-moves at every time $D := \{n_0 + 1, \dots, n_1\} \setminus C$.

Denote by C_n the event that Alice chooses a c-move at time $n + 1$. The reason for the index n instead of $n + 1$ is that C_n and D_n are B_n measurable⁵; we need that for the Martingale argument below.

Now, for every $n \in C$, let

$$L_n := \mathbf{1}_{\{S_n < S_{n-1}\}} - \mathbf{1}_{\{S_n > S_{n-1}\}}.$$

Since, by Lemma 15, when Alice decides for a c-move, $S_n > S_{n-1}$ is equivalent to $S_n = S_{n-1} + 1$, we have

$$S_n \leq S_{n-1} - L_n, \tag{1}$$

and

$$\mathbb{P}(S_{n+1} < S_n \mid B_n, C_n) \geq \frac{S_n^*}{n} + \frac{R_n}{n} \geq \frac{S_n - 4}{n} + \frac{n - S_n - 4}{n} = 1 - 8/n;$$

note that the inequalities remain valid if $S_n < 4$. As for increasing S_n , we have

$$\mathbb{P}(S_{n+1} > S_n \mid B_n, C_n) = 1/n.$$

Hence

$$\mathbb{E}(L_{n+1} \mid B_n, C_n) \geq 1 - 9/n \geq .99, \tag{2}$$

where the last inequality follows since $n_0 \geq 900$.

For the $n \in D$, i.e., when Alice decides for a d-move, we upper-bound by 1 the probability that Alices increases S_n , so,

$$S_{n+1} \leq S_n + 1 - \mathbf{1}_{C_n}, \tag{3}$$

since, by Lemma 15, $S_{n+1} > S_n$ is equivalent to $S_{n+1} = S_n + 1$ for d-moves, too.

Combining inequalities (1) and (3), we have

$$S_{n+1} - S_n \leq 1 - \mathbf{1}_{C_n} + \mathbf{1}_{C_n} \cdot L_{n+1}. \tag{4}$$

We define a super-martingale as follows. Let, $X_{n_0} := S_{n_0}$ and for each $n = n_0, \dots, n_1 - 1$,

$$X_{n+1} = X_n + \mathbf{1}_{C_n} \cdot (.99 - L_{n+1}).$$

(We define $L_n := 0$ for $n \notin C$.)

By induction, X_{n+1} is determined by B_{n+1} , so the measurability property of a super-martingale is given. Moreover, by (2), we have

$$\mathbb{E}(X_{n+1} \mid B_n) = X_n + \mathbf{1}_{C_n} \cdot (.99 - \mathbb{E}(L_{n+1} \mid B_n)) \leq X_n;$$

hence X_n is in fact a super-martingale.

Claim 18. For $n \geq n_0$, we have $X_n \geq S_n + 1.99 \sum_{j=n_0}^{n-1} C_j - (n - n_0)$.

Proof of the claim, by induction. The claim holds for $n = n_0$. Assume that it holds

⁵This is the case because Alice's decision is based on A_n and B_n in a deterministic way, and A_n is deterministically determined from B_{n-1} which is determined by B_n , and so on (induction).

for $n \geq n_0$. We have

$$\begin{aligned} X_{n+1} &= X_n + \mathbf{1}_{C_n} \cdot (.99 - L_{n+1}) \\ &\geq S_n + 1.99 \sum_{j=n_0}^{n-1} C_j - (n - n_0) + \mathbf{1}_{C_n} \cdot (.99 - L_{n+1}) \end{aligned} \quad \text{[I.H.]}$$

$$\begin{aligned} &= S_n + 1.99 \sum_{j=n_0}^{n-1} C_j - (n - n_0) + 1.99 \cdot \mathbf{1}_{C_n} - 1 + (1 - \mathbf{1}_{C_n}) - \mathbf{1}_{C_n} L_{n+1} \\ &\geq S_n + 1.99 \sum_{j=n_0}^{n-1} C_j - (n - n_0) + 1.99 \cdot \mathbf{1}_{C_n} - 1 + S_{n+1} - S_n \end{aligned} \quad \text{[(4)]}$$

$$= S_{n+1} + 1.99 \sum_{j=n_0}^n C_j - (n + 1 - n_0),$$

which completes the proof of the claim. \blacklozenge

From the claim, since $S_n \geq 0$ always, we have

$$X_{n_1} \geq 1.99 \sum_{j=n_0}^{n_1-1} C_j - (n_1 - n_0) = 1.99|C| - (n_1 - n_0) = 1.99|C| - n_0.$$

The event $|D| \leq n_0/3$ implies the event $|C| \geq 2n_0/3$, and hence the event $X_{n_1} \geq 1.2n_0$.

To apply the Azuma-Hoeffding inequality for super-martingales (e.g., Lemma 4.2 in [21]), we first note that $|X_{n+1} - X_n| \leq 1.99$ deterministically. We conclude that the probability of the event $|D| \leq n_0/3$ is at most

$$\exp\left(-\frac{(1.2n_0 - S_{n_0})^2}{2 \cdot 1.99^2 \cdot n_0}\right) \leq \exp\left(-\frac{(1.2n_0 - S_{n_0})^2}{8n_0}\right) \leq \exp\left(-\frac{n_0^2}{8n_0}\right),$$

where the last inequality follows since $S_{n_0} \leq \ln^2 n_0 \leq .2n_0$, whenever $n_0 \geq 300$. We continue, using $n_0 \geq 8 \ln(1/\varepsilon)$,

$$\mathbb{P}(|D| \leq n_0/3) \leq e^{-n_0/8} \leq \varepsilon,$$

which concludes the proof of the lemma. \square

From this, we deduce that must T_\square decrease, but some sophistication is needed, because of the slow divergence of $\sum 1/n$: Indeed, between n_0 and $2n_0$, T_\square decreases only with a constant probability:

Lemma 19. Fix $\delta := 1/42$. If $n_0 \geq \max(900, 8 \ln(1/\delta))$, and $n_1 := 2n_0$ then, whatever Alice does,

$$\mathbb{P}\left(\exists n \in \{n_0 + 1, \dots, n_1\}: T_{n+1} < T_n \mid B_{n_0}, T_{n_0} \geq 2, S_{n_0} \leq \ln^2 n_0\right) \geq 1/7.$$

Proof. From Lemma 17, with probability at least $1 - \delta$, there are at least $n_0/3$ d-moves among Alices moves at times $n = n_0 + 1, \dots, n_1 := 2n_0$.

Denote by D the set of d-moves by Alice in time $\{n_0 + 1, \dots, n_1\}$. By Lemma 15, the probability p that none of these d-moves decreases T_\square is at most

$$p \leq \prod_{n \in D} \mathbb{P}(T_n < T_{n-1} \mid T_{n-1} \geq 2) \leq \prod_{n \in D} \left(1 - \frac{1}{n}\right).$$

Hence

$$\ln p \leq \sum_{n \in D} \ln(1 - 1/n) \leq - \sum_{n \in D} 1/n.$$

The subset D of $\{n_0 + 1, \dots, n_1\}$ of size at least $n_0/3$ which maximizes the last expression is $D := \{5n_0/3, \dots, n_1\}$, so we get

$$\ln p \leq - \sum_{n=5n_0/3}^{n_1} 1/n \leq - \ln\left(\frac{n_1}{5n_0/3}\right) = -\ln(6/5).$$

We conclude that $p \leq 5/6$. In total, the probability that T_\square never decreases is at most $5/6 + \delta = 5/6 + 1/42 = 6/7$. \square

We can boost the probability to $1 - \varepsilon$, for arbitrary $\varepsilon > 0$, by iterating the argument $\Omega(\ln(1/\varepsilon))$ times.

Lemma 20. Fix $\delta := 1/42$. For every $\varepsilon \in]0, 1/56[$, with $a := 10 \ln(2/\varepsilon)$, if

$$n_0 \geq \max(900, 8 \ln(1/\delta), (2a)^{4/\varepsilon}, e^{6/\varepsilon}),$$

and $n_1 := 2an_0$ then, whatever Alice does,

$$\mathbb{P}\left(\exists n \in \{n_0, \dots, n_1\} : T_{n+1} < T_n \mid B_{n_0}, T_{n_0} \geq 2\right) \geq 1 - \varepsilon$$

Before we can prove Lemma 20, we need to control the size of S_\square through the following two lemmas.

Lemma 21. If $n_0 \geq 4$, then, whatever Alice does,

$$\mathbb{P}(S_{n_0} > \ln n_0) \leq 3/\ln n_0$$

Proof. For $e = \{k, \ell\} \in \binom{[n_0]}{2}$, denote by $E_{n_0}(e)$ the event that e is a cycle-edge of B_{n_0} . For $n_0 \geq 4$, by Fact 8,

$$\mathbb{P}(E_{n_0}(e)) = 2/(n_0 - 1).$$

Since

$$S_{n_0} = \sum_{e \in A_{n_0}} \mathbf{1}(E_{n_0}(e)),$$

we find that $\mathbb{E} S_{n_0} = \frac{2n}{n_0-1} \leq 3$ (the last inequality follows from $n_0 \geq 3$). By Markov's inequality, we have

$$\mathbb{P}(S_{n_0} \geq \ln n_0) \leq \frac{3}{\ln n_0}.$$

\square

Lemma 22. For all $\varepsilon \in]0, 1]$ and $b > 1$, if $n_0 \geq \max(10, b^{4/\varepsilon})$ then with $n_1 := bn_0$, whatever Alice does,

$$\mathbb{P}(\exists n \in \{n_0, \dots, n_1 - 1\} : S_n \geq \ln^2 n \mid B_{n_0}, S_{n_0} \leq \ln^2 n_0) \leq \varepsilon.$$

Proof. Whatever Alice does, by Lemma 15, we have $S_{n+1} \leq S_n + 1$ always and the probability that $S_{n+1} = S_n + 1$ is at most $4/n$. If $S_n \geq \ln^2 n$ for an $n \in \{n_0, \dots, n_1 - 1\}$, then either $S_{n_0} \geq \ln n_0$, or the number of times $n \in \{n_0, \dots, n_1 - 1\}$ that $S_{n+1} = S_n + 1$ is at least $\ln^2 n - \ln n_0$. But

$$\ln^2 n - \ln n_0 \geq \ln^2 n_0 - \ln n_0$$

The expected number of times $n \in \{n_0, \dots, n_1 - 1\}$ that $S_{n+1} = S_n + 1$ is

$$\sum_{n=n_0}^{n_1-1} \frac{4}{n} \leq 4(\ln(n_1/n_0) + 1/n_0 - 1/n_1) \leq 4 \ln(n_1) - \ln(n_0).$$

(We have used the well-known bound $\sum_{\ell=m}^n \frac{1}{\ell} \leq \ln(n/(m-1))$.)

Hence, the probability that we have $S_n \geq \ln^2 n$ for some $n \in \{n_0, \dots, n_1 - 1\}$ is at most

$$\frac{4 \ln(n_1) - \ln(n_0)}{\ln^2 n_0 - \ln n_0} = \frac{4 \ln(n_1/n_0) - 1}{\ln n_0 - 1} = \frac{4 \ln b - 1}{\ln n_0 - 1} \leq \varepsilon.$$

The last inequality follows from $4 \ln b \leq \varepsilon \ln n_0 \leq \varepsilon \ln n_0 + 1 - \varepsilon$ (as $\varepsilon \leq 1$). \square

We are now ready to prove Lemma 20.

of Lemma 20. For $j = 0, 1, 2, \dots$, denote by U_j the event that

$$\forall n \in \{(2j+1)n_0, \dots, (2j+3)n_0\}: T_{n+1} \geq T_n.$$

By Lemma 19, we have, for each j ,

$$\mathbb{P}\left(U_j \mid B_{(2j+1)n_0}, T_{(2j+3)n_0} \geq 2, S_{(2j+1)n_0} \leq \ln^2((2j+1)n_0)\right) \leq 6/7,$$

so

$$\mathbb{P}\left(\forall j = 0, \dots, a-1: U_j \mid B_{n_0}, T_{n_0} \geq 2, S_{n_0} \leq \ln^2(n_0)\right) \tag{5}$$

$$= \prod_{j=0}^{a-1} \mathbb{P}\left(U_j \mid B_{n_0}, T_{n_0} \geq 2, S_{n_0} \leq \ln^2(n_0), \forall i = 0, \dots, j-1: U_i\right)$$

$$\leq \prod_{j=0}^{a-1} (\varepsilon + 6/7) \tag{*}$$

$$\leq (7/8)^a \tag{**}$$

$$\leq e^{\ln(2/\varepsilon)} = \varepsilon/2.$$

The last inequality follows from $10 > 1/\ln(8/7)$; inequality (**) follows from $\varepsilon \leq 1/56$.

As for inequality (*), we note that $T_{2jn_0} = 1$ implies the event $\mathbb{C}U_{j-1}$ (\mathbb{C} is the complement), and, by Lemma 22 the probability that there is exists an $n = n_0, \dots, n_1$, with $S_n > \ln^2 n$ is at most ε .

An application of Lemma 21 gets rid of the conditioning on $S_{n_0} \leq \ln^2(n_0)$: Indeed, since from $n_0 \geq e^{6/\varepsilon}$ we have $3/\ln(n_0) \leq \varepsilon/2$, this adds another $\varepsilon/2$ to the final probability. The bound, in total, is ε . \square

Note that Lemma 20 also gets rid of the conditioning on $S_n \leq \ln^2 n$.

We are now ready to complete the proof of the main theorem.

of Theorem 7. Let $\varepsilon' \in]0, 1/2[$ be given. Set $t := 6/\varepsilon'$. Since T_\square can only increase when an isolated vertex is created, we have $T_n \leq Y$, for all $n \geq 4$, where Y is the number of isolated vertices. Hence, by Lemma 9 and Markov's inequality, we have

$$\mathbb{P}\left(\exists n \geq 4: T_n \geq t+1\right) \leq \mathbb{P}(Y \geq t) \leq \mathbb{E}(Y)/t = \varepsilon'/3.$$

Now take $n'_0 \geq 12/\varepsilon' + 2$, and large enough to apply Lemma 20 $n_0 := n'_0$ and to $\varepsilon := \varepsilon'/3t$ (note that this is less than $1/56$). Denote by a be the number defined in that lemma. Applying the lemma t times, for n_0 ranging over $n'_0 + j2an'_0$, $j = 0, \dots, t-1$, the probability that we fail at least once to obtain a decrease in the number of connected components, T_\square , is at most $\varepsilon'/3$. So, with probability at least $1 - 2\varepsilon'/3$, we must have $T_{n_0} = 1$ for one of these n_0 's or for an n between $n'_0 + (t-1)2an'_0$ and $n'_0 + t2an'_0$.

Finally, since $n'_0 \geq 12/\varepsilon' + 2$, by Lemma 9, with probability $1 - \varepsilon'/3$, T_\square will not increase after n'_0 , and hence, with probability $1 - \varepsilon'$, will drop to 1 and stay there for all eternity. Bob wins. \square

7 Some Open Questions

There are two questions which we believe should be asked in the context of our result.

Firstly, are there other polytopes whose graphs are not complete, but the minimum degree is asymptotically that of a complete graph? Could that even be the case for the Traveling Salesman Problem polytope itself?

Secondly, in view of the Traveling Salesman Problem polytope, it would be interesting to find other combinatorial conditions on cycles which are implied by the adjacency of the corresponding vertices on the TSP polytope. The pedigree graph connectedness condition is derived from an extension of the TSP polytope, but maybe there are other combinatorial conditions without that geometric context. The graph resulting from such a condition might be “closer” to the actual TSP polytope graph.

Acknowledgments

The authors would like to thank Kaveh Khoshkhah for pointing us to the idea of analyzing the pair (S, T) of random variables.

References

- [1] N Aguilera, R Katz, and P Tolomei. Vertex adjacencies in the set covering polyhedron. *arXiv preprint arXiv:1406.6015*, 2014.
- [2] T. S. Arthanari. On pedigree polytopes and hamiltonian cycles. *Discrete Math.*, 306:1474–1792, 2006.
- [3] Tiru S. Arthanari. Study of the pedigree polytope and a sufficiency condition for nonadjacency in the tour polytope. *Discrete Optimization*, 10(3):224–232, 2013.
- [4] Tiru S Arthanari and M Usha. An alternate formulation of the symmetric traveling salesman problem and its properties. *Discrete Applied Mathematics*, 98(3):173–190, 2000.
- [5] Samuel Fiorini, Volker Kaibel, Kanstantin Pashkovich, and Dirk Oliver Theis. Combinatorial bounds on nonnegative rank and extended formulations. *Discrete Math.*, 313(1):67–83, 2013.
- [6] Martin Grötschel and Manfred W. Padberg. Polyhedral theory. In Eugene L. Lawler, Jan Karel Lenstra, A. H. G. Rinnooy Kan, and David B. Shmoys, editors, *The Traveling Salesman Problem. A Guided Tour of Combinatorial Optimization*, chapter 8, pages 251–306. Wiley, 1985.
- [7] Volker Kaibel. Low-dimensional faces of random 0/1-polytopes. In *International Conference on Integer Programming and Combinatorial Optimization*, pages 401–415. Springer, 2004.
- [8] Volker Kaibel and Anja Remshagen. On the graph-density of random 0/1-polytopes. In *Approximation, Randomization, and Combinatorial Optimization: Algorithms and Techniques, 6th International Workshop on Approximation Algorithms for Combinatorial Optimization Problems, APPROX 2003 and 7th International Workshop on Randomization and Approximation Techniques in Computer Science, RANDOM 2003, Princeton, NJ, USA, August 24–26, 2003, Proceedings*, pages 318–328, 2003.

- [9] A Maksimenko. The common face of some 0/1-polytopes with np-complete nonadjacency relation. *Journal of Mathematical Sciences*, 203(6):823–832, 2014.
- [10] Denis Naddef. Pancyclic properties of the graph of some 0–1 polyhedra. *Journal of Combinatorial Theory, Series B*, 37(1):10–26, 1984.
- [11] Denis Naddef and Giovanni Rinaldi. The graphical relaxation: a new framework for the symmetric traveling salesman polytope. *Math. Programming*, 58(1, Ser. A):53–88, 1993.
- [12] Dennin J Naddef and William R Pulleyblank. Hamiltonicity in (0–1)-polyhedra. *Journal of Combinatorial Theory, Series B*, 37(1):41–52, 1984.
- [13] Christos H Papadimitriou. The adjacency relation on the traveling salesman polytope is np-complete. *Mathematical Programming*, 14(1):312–324, 1978.
- [14] Kanstantsin Pashkovich and Stefan Weltge. Hidden vertices in extensions of polytopes. *Operations Research Letters*, 43(2):161–164, 2015.
- [15] Francisco Santos. A counterexample to the hirsch conjecture. *Annals of mathematics*, 176(1):383–412, 2012.
- [16] A Sarangarajan. A lower bound for adjacencies on the traveling salesman polytope. *SIAM Journal on Discrete Mathematics*, 10(3):431–435, 1997.
- [17] Alexander Schrijver. *Combinatorial optimization. Polyhedra and efficiency.*, volume 24 of *Algorithms and Combinatorics*. Springer-Verlag, Berlin, 2003.
- [18] Gerard Sierksma. The skeleton of the symmetric traveling salesman polytope. *Discrete applied mathematics*, 43(1):63–74, 1993.
- [19] Gerard Sierksma and Ruud H Teunter. Partial monotoneizations of hamiltonian cycle polytopes: dimensions and diameters. *Discrete applied mathematics*, 105(1):173–182, 2000.
- [20] Dirk Oliver Theis. On the facial structure of symmetric and graphical traveling salesman polyhedra. *Discrete Optimization*, 12:10–25, 2014.
- [21] Nicholas C. Wormald. The differential equation method for random graph processes and greedy algorithms. In M. Karonski and H. J. Proemel, editors, *Lectures on Approximation and Randomized Algorithms*, pages 73–155. PWN, Warsaw, 1999.

ACKNOWLEDGEMENT

First of all, I would like to thank my advisor, Dirk Oliver Theis, for his continued support of my study. I deeply appreciate all his contribution of time, idea and funding that made my PhD possible. He thought me a lot consciously and unconsciously, by his patience and commitment to his students and providing challenging opportunities and step by step assistance. I learned from him how to work under crucial deadlines and be a stronger researcher.

I am very grateful to my reviewers and opponents, Babak Farzad, Härmel Nes-tra, Ali Taherkhani and Kanstantsin Pashkovich for their time and helpful comments and suggestions which considerably improved this dissertation.

I would like to thank Estonian Ministry of Education and Research and Estonian IT academy for financially supporting my PhD study.

I wish to express my sincere appreciation to all my family, especially my parents, my brother and my sister. No matter how far away we are from each other geographically, your love always warms my hurt and gives me courage. All that I am or ever hope to be, I owe to you.

Finally and most importantly, I would like to express my deepest appreciation and love to my husband, Kaveh, for being the best mate, friend, colleague and my son, Bamdad -the sunrise of my life-, for the patience he showed during my thesis writing. I would not be able to pass through challenges without your constant support, understanding and encouragement. You were always around and accompanied me across the moments of hardship and frustration. You are my source of energy and strength. I am truly thankful for having you in my life.

SUMMARY IN ESTONIAN

Polütoopide laienditega seotud ülesanded

Lineaarplaneerimine on optimeerimine matemaatilise mudeliga, mille sihifunktsioon ja kitsendused on esitatud lineaarsete seostega. Paljusid igapäeva elu väljakutseid võime vaadelda lineaarplaneerimise vormis, näiteks miinimumhinna või maksimaalse tulu leidmist. Sisepunkti meetod saavutab häid tulemusi nii teoorias kui ka praktikas ning lahendite leidmise tööaeg ja lineaarsete seoste arv on polünomiaalses seoses. Sellest tulenevalt eksponentsiaalne arv lineaarseid seoseid väljendub ka ekponentsiaalses tööajas.

Iga vajalik lineaarne seos vastab ühele polütoobi P tahule, mis omakorda tähistab lahendite hulka. Üks võimalus tööaja vähendamiseks on suurendada dimensiooni, mille tulemusel väheneks ka polütoobi tahkude arv. Saadud polütoopi Q nimetatakse polütoobi P laiendiks kõrgemas dimensioonis ning polütoobi Q minimaalset tahkude arvu nimetakakse polütoobi P laiendi keerukuseks, sellisel juhul optimaalsete lahendite hulk ei muutu. Tekib küsimus, millisel juhul on võimalik leida laiend Q , mille korral tahkude arv on polünomiaalne.

Mittedeterministlik suhtluskeerukus mängib olulist rolli tõestamaks polütoopide laiendite keerukuse alampiiri. Polütoobile P vastava suhtluskeerukuse leidmine ning alamtõkke tõestamine väistavad võimalused leida laiend Q , mis ei oleks eksponentsiaalne.

Juhuslike funktsioonide suhtluskeerukuse uurimine on tunduvalt huvitavam, kuna selle kaudu on võimalik saada teadmisi ka raskete funktsioonide leidumise ning nende omaduste kohta. Suurem osa uurimistöid keskendub olukorrale, kus tõenäosus p on konstantne. Rakendustes on aga tihti tegemist funktsioonidega, mille tihedus on suur. Seega vaatleme me juhuslikke funktsioone, mille tihedusfunktsioon on $p = p(n)$, kui $n \rightarrow \infty$.

Käesolevas töös keskendume me juhuslikele Boole'i funktsioonidele f , mille tihedusfunktsioon on $p = p(n)$. Me pakume välja vähima ülemtõkke ning suurima alamtõkke mittedeterministliku suhtluskeerukuse jaoks. Lisaks uurime me ka *pedigree* polütoobi graafi. *Pedigree* polütoop on rändkaupmehe ülesande polütoobi laiend, millel on kombinatoorne struktuur. Polütoobi graafi võib vaadelda kui abstraktset graafi ning see annab informatsiooni polütoobi omaduste kohta. Me näitame, et n linna korral on *pedigree* polütoobi minimaalne tipu aste $(1 - o(1)) \cdot (n - 1)!/2$ ($n \rightarrow \infty$). Hetkel ei ole lahendust leidnud küsimus, kas leidub polütoop, mille graaf ei ole täisgraaf, kuid mille tihedusfunktsioon koondub üheks. Täpsemalt, kas selline polütoop võib olla rändkaupmehe ülesande polütoop?

PUBLICATIONS

CURRICULUM VITAE

Personal data

Name: Mozhgan Pourmoradnasseri
Birth: September 8th, 1981
Citizenship: Iranian
Languages: Persian, English
Address: Ülikooli 17, 209 Tartu
51005 Tartumaa, Estonia
Contact: mozhgan@ut.ee

Education

2013–2017: University of Tartu, Ph.D. candidate in Computer Science
2006–2008: University of Zanjan, Zanjan, Iran
M.Sc. in Mathematics
2000–2005: Sharif University of Technology, Tehran, Iran
B.Sc. in Mathematics
1996–2000: Farzanegan High School, Tehran, Iran, secondary education
1993–1996: Somayeh Middle School, Isfahan, Iran, primary education
1988–1993: Malek Primary School, Isfahan, Iran, primary education

Employment

2008–2013 PNU University, Zanjan, Iran,
Lecturer
2005–2006: Farzanegan High School, Tehran, Iran,
Mathematics Teacher

ELULOOKIRJELDUS

Isikuandmed

Nimi: Mozhgan Pourmoradnasseri
Sünniaeg ja -koht: 8. September 1981, Iraan
Kodakondsus: iraanlane
Keelteoskus: pärsia, inglise
Aadress: Ülikooli 17, 209 Tartu
51005 Tartumaa, Eesti
Kontaktandmed: mozhgan@ut.ee

Haridus

2013–2017: Tartu Ülikool, informaatika doktorant
2006–2008: Zānjani Ülikool, Zānjan, Iraan
M.Sc. matemaatika
2000–2005: Šārifi Tehnikaülikooli, Tehran, Iraan
B.Sc. matemaatika
1996–2000: Fārzanegani Gümnaasium, Tehran, Iraan, keskharidus
1993–1996: Somāyehi Gümnaasium, Isfāhan, Iraan, põhiharidus
1988–1993: Maleki Põhikool, Isfāhan, Iraan, põhiharidus

Teenistuskäik

2008–2013: PNU Ülikool, Zānjan, Iraan,
lektor
2005–2006: Fārzanegani Gümnaasium, Tehran, Iraan,
matemaatika õpetaja

DISSERTATIONES MATHEMATICAE UNIVERSITATIS TARTUENSIS

1. **Mati Heinloo.** The design of nonhomogeneous spherical vessels, cylindrical tubes and circular discs. Tartu, 1991, 23 p.
2. **Boris Komrakov.** Primitive actions and the Sophus Lie problem. Tartu, 1991, 14 p.
3. **Jaak Heinloo.** Phenomenological (continuum) theory of turbulence. Tartu, 1992, 47 p.
4. **Ants Tauts.** Infinite formulae in intuitionistic logic of higher order. Tartu, 1992, 15 p.
5. **Tarmo Soomere.** Kinetic theory of Rossby waves. Tartu, 1992, 32 p.
6. **Jüri Majak.** Optimization of plastic axisymmetric plates and shells in the case of Von Mises yield condition. Tartu, 1992, 32 p.
7. **Ants Aasma.** Matrix transformations of summability and absolute summability fields of matrix methods. Tartu, 1993, 32 p.
8. **Helle Hein.** Optimization of plastic axisymmetric plates and shells with piece-wise constant thickness. Tartu, 1993, 28 p.
9. **Toomas Kiho.** Study of optimality of iterated Lavrentiev method and its generalizations. Tartu, 1994, 23 p.
10. **Arne Kokk.** Joint spectral theory and extension of non-trivial multiplicative linear functionals. Tartu, 1995, 165 p.
11. **Toomas Lepikult.** Automated calculation of dynamically loaded rigid-plastic structures. Tartu, 1995, 93 p, (in Russian).
12. **Sander Hannus.** Parametrical optimization of the plastic cylindrical shells by taking into account geometrical and physical nonlinearities. Tartu, 1995, 74 p, (in Russian).
13. **Sergei Tupailo.** Hilbert's epsilon-symbol in predicative subsystems of analysis. Tartu, 1996, 134 p.
14. **Enno Saks.** Analysis and optimization of elastic-plastic shafts in torsion. Tartu, 1996, 96 p.
15. **Valdis Laan.** Pullbacks and flatness properties of acts. Tartu, 1999, 90 p.
16. **Märt Põldvere.** Subspaces of Banach spaces having Phelps' uniqueness property. Tartu, 1999, 74 p.
17. **Jelena Ausekle.** Compactness of operators in Lorentz and Orlicz sequence spaces. Tartu, 1999, 72 p.
18. **Krista Fischer.** Structural mean models for analyzing the effect of compliance in clinical trials. Tartu, 1999, 124 p.
19. **Helger Lipmaa.** Secure and efficient time-stamping systems. Tartu, 1999, 56 p.
20. **Jüri Lember.** Consistency of empirical k-centres. Tartu, 1999, 148 p.
21. **Ella Puman.** Optimization of plastic conical shells. Tartu, 2000, 102 p.
22. **Kaili Müürisep.** Eesti keele arvutigrammatika: süntaks. Tartu, 2000, 107 lk.

23. **Varmo Vene.** Categorical programming with inductive and coinductive types. Tartu, 2000, 116 p.
24. **Olga Sokratova.** Ω -rings, their flat and projective acts with some applications. Tartu, 2000, 120 p.
25. **Maria Zeltser.** Investigation of double sequence spaces by soft and hard analytical methods. Tartu, 2001, 154 p.
26. **Ernst Tungel.** Optimization of plastic spherical shells. Tartu, 2001, 90 p.
27. **Tiina Puolakainen.** Eesti keele arvutigrammatika: morfoloogiline ühestamine. Tartu, 2001, 138 p.
28. **Rainis Haller.** $M(r,s)$ -inequalities. Tartu, 2002, 78 p.
29. **Jan Villemson.** Size-efficient interval time stamps. Tartu, 2002, 82 p.
30. Töö kaitsmata.
31. **Mart Abel.** Structure of Gelfand-Mazur algebras. Tartu, 2003. 94 p.
32. **Vladimir Kuchmei.** Affine completeness of some ockham algebras. Tartu, 2003. 100 p.
33. **Olga Dunajeva.** Asymptotic matrix methods in statistical inference problems. Tartu 2003. 78 p.
34. **Mare Tarang.** Stability of the spline collocation method for volterra integro-differential equations. Tartu 2004. 90 p.
35. **Tatjana Nahtman.** Permutation invariance and reparameterizations in linear models. Tartu 2004. 91 p.
36. **Märt Möls.** Linear mixed models with equivalent predictors. Tartu 2004. 70 p.
37. **Kristiina Hakk.** Approximation methods for weakly singular integral equations with discontinuous coefficients. Tartu 2004, 137 p.
38. **Meelis Käärrik.** Fitting sets to probability distributions. Tartu 2005, 90 p.
39. **Inga Parts.** Piecewise polynomial collocation methods for solving weakly singular integro-differential equations. Tartu 2005, 140 p.
40. **Natalia Saealle.** Convergence and summability with speed of functional series. Tartu 2005, 91 p.
41. **Tanel Kaart.** The reliability of linear mixed models in genetic studies. Tartu 2006, 124 p.
42. **Kadre Torn.** Shear and bending response of inelastic structures to dynamic load. Tartu 2006, 142 p.
43. **Kristel Mikkor.** Uniform factorisation for compact subsets of Banach spaces of operators. Tartu 2006, 72 p.
44. **Darja Saveljeva.** Quadratic and cubic spline collocation for Volterra integral equations. Tartu 2006, 117 p.
45. **Kristo Heero.** Path planning and learning strategies for mobile robots in dynamic partially unknown environments. Tartu 2006, 123 p.
46. **Annely Mürk.** Optimization of inelastic plates with cracks. Tartu 2006. 137 p.
47. **Annemai Raidjõe.** Sequence spaces defined by modulus functions and superposition operators. Tartu 2006, 97 p.
48. **Olga Panova.** Real Gelfand-Mazur algebras. Tartu 2006, 82 p.

49. **Härmel Nestra.** Iteratively defined transfinite trace semantics and program slicing with respect to them. Tartu 2006, 116 p.
50. **Margus Pihlak.** Approximation of multivariate distribution functions. Tartu 2007, 82 p.
51. **Ene Käärik.** Handling dropouts in repeated measurements using copulas. Tartu 2007, 99 p.
52. **Artur Sepp.** Affine models in mathematical finance: an analytical approach. Tartu 2007, 147 p.
53. **Marina Issakova.** Solving of linear equations, linear inequalities and systems of linear equations in interactive learning environment. Tartu 2007, 170 p.
54. **Kaja Sõstra.** Restriction estimator for domains. Tartu 2007, 104 p.
55. **Kaarel Kaljurand.** Attempto controlled English as a Semantic Web language. Tartu 2007, 162 p.
56. **Mart Anton.** Mechanical modeling of IPMC actuators at large deformations. Tartu 2008, 123 p.
57. **Evely Leetma.** Solution of smoothing problems with obstacles. Tartu 2009, 81 p.
58. **Ants Kaasik.** Estimating ruin probabilities in the Cramér-Lundberg model with heavy-tailed claims. Tartu 2009, 139 p.
59. **Reimo Palm.** Numerical Comparison of Regularization Algorithms for Solving Ill-Posed Problems. Tartu 2010, 105 p.
60. **Indrek Zolk.** The commuting bounded approximation property of Banach spaces. Tartu 2010, 107 p.
61. **Jüri Reimand.** Functional analysis of gene lists, networks and regulatory systems. Tartu 2010, 153 p.
62. **Ahti Peder.** Superpositional Graphs and Finding the Description of Structure by Counting Method. Tartu 2010, 87 p.
63. **Marek Kolk.** Piecewise Polynomial Collocation for Volterra Integral Equations with Singularities. Tartu 2010, 134 p.
64. **Vesal Vojdani.** Static Data Race Analysis of Heap-Manipulating C Programs. Tartu 2010, 137 p.
65. **Larissa Roots.** Free vibrations of stepped cylindrical shells containing cracks. Tartu 2010, 94 p.
66. **Mark Fišel.** Optimizing Statistical Machine Translation via Input Modification. Tartu 2011, 104 p.
67. **Margus Niitsoo.** Black-box Oracle Separation Techniques with Applications in Time-stamping. Tartu 2011, 174 p.
68. **Olga Liivapuu.** Graded q -differential algebras and algebraic models in noncommutative geometry. Tartu 2011, 112 p.
69. **Aleksei Lissitsin.** Convex approximation properties of Banach spaces. Tartu 2011, 107 p.
70. **Lauri Tart.** Morita equivalence of partially ordered semigroups. Tartu 2011, 101 p.
71. **Siim Karus.** Maintainability of XML Transformations. Tartu 2011, 142 p.

72. **Margus Treumuth.** A Framework for Asynchronous Dialogue Systems: Concepts, Issues and Design Aspects. Tartu 2011, 95 p.
73. **Dmitri Lepp.** Solving simplification problems in the domain of exponents, monomials and polynomials in interactive learning environment T-algebra. Tartu 2011, 202 p.
74. **Meelis Kull.** Statistical enrichment analysis in algorithms for studying gene regulation. Tartu 2011, 151 p.
75. **Nadežda Bazunova.** Differential calculus $d^3 = 0$ on binary and ternary associative algebras. Tartu 2011, 99 p.
76. **Natalja Lepik.** Estimation of domains under restrictions built upon generalized regression and synthetic estimators. Tartu 2011, 133 p.
77. **Bingsheng Zhang.** Efficient cryptographic protocols for secure and private remote databases. Tartu 2011, 206 p.
78. **Reina Uba.** Merging business process models. Tartu 2011, 166 p.
79. **Uuno Puus.** Structural performance as a success factor in software development projects – Estonian experience. Tartu 2012, 106 p.
80. **Marje Johanson.** $M(r, s)$ -ideals of compact operators. Tartu 2012, 103 p.
81. **Georg Singer.** Web search engines and complex information needs. Tartu 2012, 218 p.
82. **Vitali Retšnoi.** Vector fields and Lie group representations. Tartu 2012, 108 p.
83. **Dan Bogdanov.** Sharemind: programmable secure computations with practical applications. Tartu 2013, 191 p.
84. **Jevgeni Kabanov.** Towards a more productive Java EE ecosystem. Tartu 2013, 151 p.
85. **Erge Ideon.** Rational spline collocation for boundary value problems. Tartu, 2013, 111 p.
86. **Esta Kägo.** Natural vibrations of elastic stepped plates with cracks. Tartu, 2013, 114 p.
87. **Margus Freudenthal.** Simpl: A toolkit for Domain-Specific Language development in enterprise information systems. Tartu, 2013, 151 p.
88. **Boriss Vlassov.** Optimization of stepped plates in the case of smooth yield surfaces. Tartu, 2013, 104 p.
89. **Elina Safiulina.** Parallel and semiparallel space-like submanifolds of low dimension in pseudo-Euclidean space. Tartu, 2013, 85 p.
90. **Raivo Kolde.** Methods for re-using public gene expression data. Tartu, 2014, 121 p.
91. **Vladimir Šor.** Statistical Approach for Memory Leak Detection in Java Applications. Tartu, 2014, 155 p.
92. **Naved Ahmed.** Deriving Security Requirements from Business Process Models. Tartu, 2014, 171 p.
93. **Kerli Orav-Puurand.** Central Part Interpolation Schemes for Weakly Singular Integral Equations. Tartu, 2014, 109 p.
94. **Liina Kamm.** Privacy-preserving statistical analysis using secure multi-party computation. Tartu, 2015, 201 p.

95. **Kaido Lätt.** Singular fractional differential equations and cordial Volterra integral operators. Tartu, 2015, 93 p.
96. **Oleg Košik.** Categorical equivalence in algebra. Tartu, 2015, 84 p.
97. **Kati Ain.** Compactness and null sequences defined by ℓ_p spaces. Tartu, 2015, 90 p.
98. **Helle Hallik.** Rational spline histopolation. Tartu, 2015, 100 p.
99. **Johann Langemets.** Geometrical structure in diameter 2 Banach spaces. Tartu, 2015, 132 p.
100. **Abel Armas Cervantes.** Diagnosing Behavioral Differences between Business Process Models. Tartu, 2015, 193 p.
101. **Fredrik Milani.** On Sub-Processes, Process Variation and their Interplay: An Integrated Divide-and-Conquer Method for Modeling Business Processes with Variation. Tartu, 2015, 164 p.
102. **Huber Raul Flores Macario.** Service-Oriented and Evidence-aware Mobile Cloud Computing. Tartu, 2015, 163 p.
103. **Tauno Metsalu.** Statistical analysis of multivariate data in bioinformatics. Tartu, 2016, 197 p.
104. **Riivo Talviste.** Applying Secure Multi-party Computation in Practice. Tartu, 2016, 144 p.
105. **Md Raknuzzaman.** Noncommutative Galois Extension Approach to Ternary Grassmann Algebra and Graded q-Differential Algebra. Tartu, 2016, 110 p.
106. **Alexander Liyvapuu.** Natural vibrations of elastic stepped arches with cracks. Tartu, 2016, 110 p.
107. **Julia Polikarpus.** Elastic plastic analysis and optimization of axisymmetric plates. Tartu, 2016, 114 p.
108. **Siim Orasmaa.** Explorations of the Problem of Broad-coverage and General Domain Event Analysis: The Estonian Experience. Tartu, 2016, 186 p.
109. **Prastudy Mungkas Fauzi.** Efficient Non-interactive Zero-knowledge Protocols in the CRS Model. Tartu, 2017, 193 p.
110. **Pelle Jakovits.** Adapting Scientific Computing Algorithms to Distributed Computing Frameworks. Tartu, 2017, 168 p.
111. **Anna Leontjeva.** Using Generative Models to Combine Static and Sequential Features for Classification. Tartu, 2017, 167 p.