Department of Mathematics and Systems Analysis

Algebraic Statistics

Patrik Norén

DOCTORAL DISSERTATIONS

Aalto University publication series
DOCTORAL DISSERTATIONS 38/2013

Algebraic Statistics

Patrik Norén

A doctoral dissertation completed for the degree of Doctor of Science in Technology to be defended, with the permission of the Aalto University School of Science, at a public examination held in Auditorium G at Otsvängen 1, Otnäs, on 15 April 2013 at 13:00.

Aalto University
School of Science
Department of Mathematics and Systems Analysis

Supervising professor
Prof. Alexander Engström

Thesis advisor
Prof. Alexander Engström

Preliminary examiners
Prof. Volkmar Welker
Philipps-Universität Marburg
Germany

Dr. Carsten Schultz
Freie Universität Berlin
Germany

Opponents
Prof. Volkmar Welker
Philipps-Universität Marburg
Germany

Dr. Axel Hultman
Linköpings universitet
Sweden

Aalto University publication series
DOCTORAL DISSERTATIONS 38/2013

ISBN 978-952-60-5047-8 (printed)
ISBN 978-952-60-5048-5 (pdf)
ISSN-L 1799-4934
ISSN 1799-4934 (printed)
ISSN 1799-4942 (pdf)
http://urn.fi/URN:ISBN:978-952-60-5048-5

Unigrafia Oy
Helsinki 2013

Finland

441 697
Printed matter

A! Aalto University

Aalto University, P.O. Box 11000, FI-00076 Aalto www.aalto.fi

Abstract

Author
Patrik Norén

Name of the doctoral dissertation
Algebraic Statistics

Publisher School of Science

Unit Department of Mathematics and Systems Analysis

Series Aalto University publication series DOCTORAL DISSERTATIONS 38/2013

Field of research Mathematics

Manuscript submitted 11 December 2012 **Date of the defence** 15 April 2013

Permission to publish granted (date) 31 January 2013 **Language** English

☐ **Monograph** ☒ **Article dissertation (summary + original articles)**

Abstract

This thesis on algebraic statistics contains five papers.

In paper I we define ideals of graph homomorphisms. These ideals generalize many of the toric ideals defined in terms of graphs that are important in algebraic statistics and commutative algebra.

In paper II we study polytopes from subgraph statistics. Polytopes from subgraph statistics are important for statistical models for large graphs and many problems in extremal graph theory can be stated in terms of them. We find easily described semi-algebraic sets that are contained in these polytopes, and using them we compute dimensions and get volume bounds for the polytopes.

In paper III we study the topological Tverberg theorem and its generalizations. We develop a toolbox for complexes from graphs using vertex decomposability to bound the connectivity.

In paper IV we prove a conjecture by Haws, Martin del Campo, Takemura and Yoshida. It states that the three-state toric homogenous Markov chain model has Markov degree two. In algebraic terminology this means that a certain class of toric ideals are generated by quadratic binomials.

In paper V we produce cellular resolutions for a large class of edge ideals and their powers. Using algebraic discrete Morse theory it is then possible to make many of these resolutions minimal, for example explicit minimal resolutions for powers of edge ideals of paths are constructed this way.

Keywords Algebra and Statistics.

ISBN (printed) 978-952-60-5047-8 **ISBN (pdf)** 978-952-60-5048-5

ISSN-L 1799-4934 **ISSN (printed)** 1799-4934 **ISSN (pdf)** 1799-4942

Location of publisher Espoo **Location of printing** Helsinki **Year** 2013

Pages 140 **urn** http://urn.fi/URN:ISBN:978-952-60-5048-5

A? Aalto-universitetet

Aalto-universitetet, PB 11000, 00076 Aalto www.aalto.fi

Sammandrag

Författare
Patrik Norén

Doktorsavhandlingens titel
Algebraisk statistik

Utgivare Högskolan för teknikvetenskaper

Enhet Institutionen för matematik och systemanalys

Seriens namn Aalto University publication series DOCTORAL DISSERTATIONS 38/2013

Forskningsområde Matematik

Inlämningsdatum för manuskript 11.12.2012 **Datum för disputation** 15.04.2013

Beviljande av publiceringstillstånd (datum) 31.01.2013 **Språk** Engelska

☐ **Monografi** ☒ **Sammanläggningsavhandling (sammandrag plus separata artiklar)**

Sammandrag

Denna avhandling om algebraisk statistik innehåller fem artiklar.

I artikel I definieras ideal av grafhomomorfier. Dessa ideal generaliserar ett flertal konstruktioner av ideal från grafer som är viktiga i algebraisk statistik samt kommutativ algebra.

I artikel II behandlas polytoper från delgrafsstatistik. Dessa är viktiga för att förstå statistiska modeller som beskriver stora grafer och många problem om ytterlighetsgrafer kan formuleras med dem. Bland verktygen som används är att beskriva semi-algebraiska mängder i polytoperna och genom detta bestämma deras dimension samt begränsa volymen.

I artikel III behandlas den topologiska tverbergssatsen med generaliseringar. Grafkomplexen förstås genom att begränsa sammanhängandegraden medelst hörnnedbrytbarhet.

I artikel IV bevisas att ideal tillhörande markovkedjor med tre tillstånd är genererade i grad två, vilket förmodats av Haws, Martin del Campo, Takemura och Yoshida.

I artikel V skapas cellulära upplösningar för en stor klass av kantideal samt deras potenser. Med algebraisk diskret morseteori görs dessa upplösningar minimala för kantideal från stigar.

Nyckelord Algebra och statistik.

ISBN (tryckt) 978-952-60-5047-8 **ISBN (pdf)** 978-952-60-5048-5

ISSN-L 1799-4934 **ISSN (tryckt)** 1799-4934 **ISSN (pdf)** 1799-4942

Utgivningsort Esbo **Tryckort** Helsingfors **År** 2013

Sidantal 140 **urn** http://urn.fi/URN:ISBN:978-952-60-5048-5

Preface

First of all I thank my advisor Alex Engström. He has thought me a lot of mathematics and almost everything I know about academic life. He has been a great source of ideas, insights and enthusiasm. He is a good friend and co-author too.

I thank Svante Linusson for being my advisor during my time at KTH and I thank Anders Björner for financing my position there.

As a PhD student one gets many opportunities to travel and meet new people, I'm very grateful for this. I thank Bernd Sturmfels for his hospitality during my visits to Berkeley and I thank the Miller foundation for paying for my flights there.

Jonathan Browder, Ragnar Freij, Erik Sjöland, Matthew Stamps and Wouter van Heijst are all awesome in sublime ways and they make our office the best office at Aalto. Matthew Stamps deserves special thanks for helping me make this thesis readable.

Without the support of my family and friends nothing of this would have been possible.

Helsinki, February 12, 2013,

Patrik Norén

Contents

List of Publications

This thesis consists of an overview and of the following publications which are referred to in the text by their Roman numerals.

I Alexander Engström and Patrik Norén. Ideals of graph homomorphisms. *Annals of Combinatorics*, **17**, 2013.

II Alexander Engström and Patrik Norén. Polytopes from subgraph statistics. *arxiv:1011.3552*, 2010.

III Alexander Engström and Patrik Norén. Tverberg's theorem and graph coloring. *arxiv:1105.1455*, 2011.

IV Patrik Norén. The three-state torics homogenous Markov chain model has Markov degree two. *arxiv:1207.0077*, 2012.

V Alexander Engström and Patrik Norén. Cellular resolutions of powers of monomial ideals. *arxiv:1212.2146*, 2012.

Author's Contribution

Publication I: "Ideals of graph homomorphisms"

Norén found most of the theorems and proved them. Engström suggested the question and wrote parts of the introduction.

Publication II: "Polytopes from subgraph statistics"

Norén found most of the theorems and proved them. Engström suggested the question and wrote parts of the introduction.

Publication III: "Tverberg's theorem and graph coloring"

Norén and Engström did equal parts.

Publication IV: "The three-state torics homogenous Markov chain model has Markov degree two"

Norén did everything.

Publication V: "Cellular resolutions of powers of monomial ideals"

Norén found most of the theorems and proved them. Engström suggested the question and wrote parts of the introduction.

1. Introduction

1.1 Introduction to algebraic statistics

The goal of algebraic statistics is to solve statistical problems using tools and techniques from algebra. There is a wide array of problems for which this approach has been highly successful, from improving statistical testing methodology to determining whether a particular model is identifiable or not.

In practice, the algebraic questions arising from statistics typically involve monomial or toric ideals – two well-studied classes of algerbaic objects. These ideals are very well-behaved, but understanding them requires a substantial amount of combinatorial theory.

Another collection of well-studied objects that appear frequently in algebraic statistics are graphs, which provide a convenient way to encode dependencies among random variables. Graphs are also important in statistics since understanding the behavior of large random networks is becoming increasingly important for real-life applications.

This thesis consists of five papers that investigate various combinatorial problems arising from algebraic statistics and its related topics. The document begins with introductions, including literature suggestions, to the methods and themes that are most essential for understanding these articles, e.g., monomial and toric ideals, graph homomorphisms, polytopes from graphs, and connectivity in polytopal complexes, followed by summaries of the articles, and then the articles themselves.

The book by Drton, Sturmfels, and Sullivant [17] provides an excellent introduction to algebraic statistics.

1.2 Methods

1.2.1 Monomial and toric ideals

Toric and monomial ideals are two classes of polynomial ideals that are important in commutative algebra and algebraic statistics.

The central objects in papers I and II are toric ideals. There are many equivalent definitions of toric ideals, but for our purposes, we use the following: A *toric ideal* is a polynomial ideal generated by all binomials $\mathbf{x}^u - \mathbf{x}^v$ such that $u - v$ is in the kernel of an integer matrix A.

In statistics, it is often convenient to collect data into contingency tables. For instance, when performing certain statistical tests, it is useful to have a way to generate tables with the same row and column sums uniformly. Algebraic statistics provides a systematic way of doing this: The idea is to encode tables as monomials and then study the toric ideal generated by the binomials corresponding to tables with the same row and column sums. Given a set of row and columns sums, along with a generating set for the ideal, one can produce a connected graph with vertices corresponding to the contingency tables with the given row and column sums and edges corresponding to binomials that are divisible by an element in the generating set.

The primary aim of algebraic statistics is to understand toric ideals like these and to interpret what various algebraic properties of these ideals mean for the statistical methods. For example, the important property of *normality* is implied by having a square free generating set.

A common approach for understanding complicated toric ideals is decomposing them into simpler ones. The main method for doing this is the *toric fiber product*, introduced by Sullivant [56] and developed further by Sullivant, Engström and Khale [24]. For two toric ideals with sufficiently compatible structures, their toric fiber product is a new toric ideal that inherits many of their individual properties. In this thesis, there are a number of toric ideals defined in terms of graphs, where the toric fiber product often corresponds to gluing the underlying graphs together over a common subgraph or some other simple process of decomposing graphs.

Monomial ideals are polynomial ideals generated by monomials. In algebraic statistics, they have been used to encode systems in reliability theory. They are, in some respect, easier to understand than toric ideals and they provide good examples for different algebraic techniques. The main objects appearing in paper V are square free monomial ideals and their powers.

To understand a polynomial ideal I in a polynomial ring S, and more generally any module over a ring, it is very useful to consider a free resolution. A *free resolution* is an exact sequence of the form

$$0 \leftarrow S/I \leftarrow I \leftarrow S^{b_0} \leftarrow S^{b_1} \leftarrow \cdots.$$

In general, it is a challenging problem to produce minimal resolutions like this; however, sometimes the maps in the sequence can be interpreted as boundary maps in a cell complex. For monomial ideals, this gives rise to the method of cellular resolutions introduced by Sturmfels and Bayer [2].

Literature

- Some general literature on toric ideals:
  [20, 24, 28, 34, 55, 56]

- Some examples of toric ideals in algebraic statistics:
  [10, 12, 17, 29, 33, 35, 36]

- Some examples where monomial ideals are studied:
  [2, 15, 16]

1.2.2 Graph homomorphisms

A *graph homomorphism* is a map from the vertex set of a graph G to the vertex set of a graph H that induces a map from the edge set of G to the edge set of H. Graph homomorphisms give the class of graphs the structure of a category. In algebraic statistics, there are many objects defined in terms of graphs, but often in a somewhat ad hoc way. The main example of this is the toric ideals related to graphical models. Graph homomorphisms can be used to unify these concepts – one of the main themes in paper I.

Graph homomorphisms generalize many of the concepts that are central in graph theory, for example graph colorings and independent sets. A *coloring* of a graph G is a coloring of the vertices of the graph so that no two adjacent vertices have the same color. A coloring of G with n colors can then be encoded as a graph homomorphism from G to the complete graph with n vertices. An independent set in a graph is a set of vertices so that no pair of vertices in the set are adjacent. An *independent set* can be encoded as a graph homomorphism into the graph consisting of an edge between two vertices with a loop attached to one of the vertices.

In paper IV, we use graph colorings to give new versions of the topological Tverberg theorem and, in paper I, one of our results gives a new algebraic method to obstruct the existence of certain graph colorings.

An important problem in graph theory is to count the number of graph homomorphisms between two graphs, especially between large random graphs. In paper III, the main object is a class of polytopes that are obtained by counting graph homomorphisms. These polytopes parameterize exponential random graph models that are used to understand very large networks.

Another application is that outcomes of Markov chains can be interpreted as graph homomorphisms from a directed path to some graph depending on the model. In paper IV, we employ this approach to prove a conjecture by Haws, Martín del Campo, Takemura and Yoshida [36] about toric ideals related to Markov chains.

In paper V, polyhedral complexes obtained from generalizations of graph homomorphisms are used to find cellular resolutions of monomial ideals.

Literature

- Good general texts on graph homomorphisms are:
  [7, 13]

- Different concepts of graph colorings are studied in:
  [5, 6, 7, 18, 38, 45, 57]

- The counting of subgraphs and how it is related to statistics:
  [8, 11, 14, 46, 49, 50, 51]

- Here are some examples of papers studying independent sets:
  [19, 22, 44, 47]

1.2.3 Polytopes from graphs

A *convex polytope* is the convex hull of a finite set of points. If the points are lattice points, the polytope is a lattice polytope.

To every lattice polytope, we associate a toric ideal as follows: The coordinates of the vertices of a lattice polytope can be collected into a matrix A that define a toric ideal. Reversing this process yields a polytope for every toric ideal. The properties of this ideal can often be expressed in terms of the corresponding polytope. In this respect, all the toric ideals that are defined in terms of graphs give polytopes that have a graph theoretic interpretation. A number of the proofs in article I make use of this correspondence.

Polytopes are important in optimization theory. A well known combinatorial optimization problem is to find a largest independent set in a graph. This is encoded by the *independent set polytope*, whose vertices are the incidence vectors of independent sets in a graph. In paper I, these polytopes appear as the polytopes related to toric ideals in algebraic statistics.

An important class of statistical models for large random networks are the exponential random graph models. These models are parameterized by the polytopes of subgraph statistics studied in paper II. These polytopes are obtained by taking the convex hull of all possible vectors of subgraph counts of graphs with a given number of vertices.

A lattice polytope is *normal* if the lattice points in integer dilations of the polytope can be obtained as a sum of the lattice points in the original polytope. This property is important for the toric ideals associated to polytopes. For example, normality can be used to bound the degree of a generating set of a toric ideal. The normality of polytopes related to graphs is used in articles I and IV, where the polytope appearing in paper IV comes from studying the graph homomorphisms from a directed path into a complete directed graph.

Literature

- Some general texts on the importance of polytopes:
  [25, 32, 40, 41, 43, 54, 59]

- Some examples of where polytopes are important in statistics:
  [26, 52]

1.2.4 Connectivity of polytopal complexes

Topological combinatorics has produced many different ways to solve combinatorial problems by translating them to questions about topological spaces. Many problems are translated into questions about the connectivity a particular topological space. The spaces that most often occur in this way are polyhedral complexes. A *polytopal complex* is a set of polytopes that have been glued together over faces. A special example of a polyhedral complex is a *simplicial complex*, where every polytope in the complex is a simplex. An important example of a simplicial complex is *independence complex* of a graph G, whose vertex set is the vertex set of G and whose simplices are the independent sets in G. This kind of complex is studied in papers III and V.

If a complex is contractible, then it is as connected as it can be and proving that a complex is contractible is often the easiest way to prove that a complex is very connected. If a complex can be embedded in such a way that it is convex, then we get a certificate that it is contractible. Another way to show that polyhedral complexes are highly connected is to show that they are shellable. A complex is *shellable* if its facets can be removed sequentially in a well-controlled manner. Shellability implies that the complex topologically is a wedge spheres of some fixed dimension. To show that a simplicial complex is shellable, it is sometimes easier to prove a stronger condition called *vertex decomposablility*. In paper III, the concept of vertex decompasabilitity of the independence complex of graph is translated to a graph theoretic property and used to prove a variant of Tverbergs theorem.

Discrete Morse theory provides a way to reduce the number of cells in a complex without changing its topology. In paper V, discrete Morse theory is used to prove that some cell complexes support minimal cellular resolutions of powers of edge ideals of graphs.

Literature

- Some texts explaining and using topological combinatorics:
  [1, 4, 9, 30, 31]

- The independence complexes are studied in:
  [19, 42, 44]

- The concept of shellability is important in:

  [58]

1.3 Summary

1.3.1 Paper I: Ideals of graph homomorphisms

In this paper, we introduce a new class of toric ideals called ideals of graph homomorphisms. Given two graphs G and H we define an ideal $I_{G\to H}$ in the polynomial ring

$$\mathbf{k}[r_\phi \mid \phi \text{ is a graph homomorphism from } G \text{ to } H].$$

These ideals generalize toric ideals studied in algebraic statistics, where H is taken to be a complete graph with loops. A good introduction to algebraic statistics is the book by Drton, Sturmfels and Sullivant [17]. Polytopes and toric ideals are the two main classes of objects studied here with examples coming from graphs. We prove a number of basic results about these ideals, using methods from Sections 1.2.1, 1.2.2, and 1.2.3, and we show that some of them have very nice algebraic properties, such as being Cohen-Macaulay.

In algebraic statistics, it is important to find generating sets and Gröbner bases of ideals. We use toric fiber products by Sullivant [56] to create generating sets for larger graphs by gluing together the generating sets of smaller graphs. When H is an edge with a loop on one of the vertices, the graph homomorphisms correspond to independent sets of the graph G. In this setting, we provide Gröbner bases for the ideals corresponding to all bipartite graphs, yielding an alternative proof of a classic theorem by Hibi [39]. Explicit Gröbner bases are also constructed for all graphs that are bipartite if one vertex is removed: Such Gröbner bases are quadratic and square-free, which proves that the corresponding rings are Cohen-Macaulay.

1.3.2 Paper II: Polytopes from subgraph statistics

In this paper, we study the convex polytopes obtained by taking the convex hull of all possible vectors of subgraph counts of graphs with a given number of vertices. These polytopes are interesting from an extremal graph theory point of view and they are important in statistics. Polytopes play a key role and, since counting subgraphs is close to counting homomorphisms, we use methods from Section, 1.2.2 and 1.2.3.

In extremal combinatorics, there are many questions about the possible number of certain subgraphs if one fixes the number of another subgraph. These can be translated into questions about what happens close to the boundary of a polytope of subgraph statistics. In statistics, the geometry of these polytopes determine the behavior of maximum likelihood estimations for exponential random graph models, as described by Rinaldo, Fienberg and Zhou [26].

Many questions from extremal graph theory expressed in terms of these polytopes are very hard to understand. Instead of studying the polytopes directly we approximate them by using random graph models. The expectation values of the subgraph counts from random graph models cut out subsets of the polytope. In many cases, these subsets are easily parameterized semi-algebraic sets. We find a family of such sets, called curvy zonotopes, that, in the limit of very large graphs, fill the entire polytope. Using this family of sets, we bound the volume of the polytopes and detemine their dimensions. The volume bound involves the evaluation of a new class of integrals generalizing the Selberg integral [53].

The random graph models we use are exchangeable random graph models. Their usefulness can be explained by their central role in the theory of graph limits developed by Lovász and Szegedy [46]. With this, we prove that the polytopes can be approximated arbitrarily well by curvy zonotopes.

1.3.3 Paper III: Tverberg's theorem and graph coloring

The topological Tverberg theorem states that for any prime power q and any continuous map from the $(d+1)(q-1)$-simplex to $\mathbf{R}^d$, there are q disjoint faces of the simplex whose images intersect. This holds even under some conditions on which vertices can be in the same face. One way to encode these conditions is to say that the faces have to be color classes in a coloring of a certain graph. If the theorem holds, then the graph is called a Tverberg graph. Engström [18] showed that a graph of maximal degree D is a Tverberg graph if $D(D+1) < q$. However, this condition is far from optimal: It is conjectured that there is a constant K such that a graph is Tverberg if $KD < q$.

In this paper, we use methods from section 1.2.4 to prove a fixed parameter version of the conjecture. That is, for every $\epsilon > 0$, there is a constant K_ϵ such that if the graph has $((d+1)(q-1)+1)(1+\epsilon)$ vertices and $q \geq K_\epsilon \Delta$, then there is a q-coloring of the graph with the desired non-empty intersection property.

When proving that a graph is Tverberg, one important object is the independence complex. It is required that the independence complexes are sufficiently (topologically) connected, but Engström [18] conjectured that the complexes are, in fact, shellable. We prove this conjecture by proving that the complexes are vertex decomposable. For a good introduction to topological combinatorics, see [4]. The main graph-theoretic tool for proving both results are algorithms for removing a special types of subgraphs called squids.

1.3.4 Paper IV: The three-state torics homogenous Markov chain model has Markov degree two

The Markov degree of the toric homogenous Markov chain model is the minimal degree in which a particular toric ideal is generated. The ideals studied here were first considered by Takemura and Hara [33] and were further studied by Haws, Martín del Campo, Takemura, and Yoshida [36] who, based on computer calculations, made many conjectures about these ideals. This paper is devoted to proving the conjecture that the Markov degree of the three state Markov chain model is two.

These ideals describe some of the equations satisfied by the probability distribution of a general time homogenous Markov chain. The methods used are very combinatorial, but are inspired by the toric fiber products introduced in Section 1.2.1. The variables in the ring are indexed by graph homomorphisms from a directed path into a complete directed graph and the description in terms of the homomorphisms in Section 1.2.2 is used. When the lengths of paths increase, no new types of moves are needed and thus, the ideals from long paths are constructed from the ideals of shorter paths.

1.3.5 Paper V: Cellular resolutions of powers of monomial ideals

In this paper, we study monomial ideals and powers of monomial ideals, with a focus on edge ideals of graphs and their powers. The edge ideal of a graph G is the ideal $\langle x_u x_v \mid uv \in E(G)\rangle$ in $\mathbf{k}[x_v \mid v \in V(G)]$.

Our goal is to find minimal free resolutions of these ideals and simultaneously find resolutions of all their powers. The main method for doing this uses cellular resolutions. First we construct a non-minimal resolution from a subdivided polytope. Then we use convexity to prove that an acyclicity condition is satisfied in order to support a resolution. From there, we use discrete Morse theory to make the resulting cell complex and resolution smaller –this approach was developed by Betzies and Welker. The discrete Morse theory is applied in two steps, the first step is to reduce the complex to a cell complex isomorphic to a Hom complex that has a nice description. The final step uses optimal Morse matchings from independence complexes of graphs to make the resolution minimal.

Our main example is the edge ideal of a path and all its powers. The primary objects and techniques in this paper are monomial ideals and the machinery introduced in Section 1.2.1. Cell complexes related to the graph homomorphisms introduced Section 1.2.2 support the cellular resolutions of interest and discrete Morse theory, introduced in Section 1.2.4, is the main tool to make the resolutions smaller.

Bibliography

[1] Eric Babson and Dmitry N. Kozlov. Proof of the Lovász conjecture. *Ann. of Math. (2)* **165** (2007), no. 3, 965–1007.

[2] Dave Bayer and Bernd Sturmfels. Cellular resolutions of monomial modules. *J. Reine Angew. Math.* **502** (1998), 123–140.

[3] Bruno Benedetti. Discrete Morse theory for manifolds with boundary. *Trans. Amer. Math. Soc.* **364** (2012), no. 12, 6631–6670.

[4] Anders Björner. Topological Methods. *Handbook of combinatorics, Vol. 1, 2*, 1819–1872, *Elsevier, Amsterdam*, 1995.

[5] Pavle V. M. Blagojević, Benjamin Matschke, and Günter M. Ziegler. Optimal bounds for a colorful Tverberg-Vrećica type problem. *Adv. Math.* **226** (2011), no. 6, 5198–5215.

[6] Pavle V. M. Blagojević, Benjamin Matschke, and Günter M. Ziegler. Optimal bounds for the colored Tverberg problem. `arXiv:0910.4987`, 2009. 11 pp.

[7] Pavle V. M. Blagojević, Benjamin Matschke, and Günter M. Ziegler. A tight colored Tverberg theorem for maps to manifolds. *Topology Appl.* **158** (2011), no. 12, 1445-1452.

[8] Béla Bollobás. Relations between sets of complete subgraphs. *Proceedings of the Fifth British Combinatorial Conference (Univ. Aberdeen, Aberdeen, 1975)*, pp. 79–84. *Congressus Numerantium, No. XV, Utilitas Math., Winnipeg, Man.*, 1976.

[9] Jesús A. De Loera, Jörg Rambau, and Francisco Santos. *Triangulations. Structures for algorithms and applications.* Algorithms and Computation in Mathematics, 25. *Springer-Verlag, Berlin*, 2010. 535 pp.

[10] Mike Develin and Seth Sullivant. Markov bases of binary graph models. *Ann. Comb.* **7** (2003), no. 4, 441–466.

[11] Persi Diaconis and Svante Janson. Graph limits and exchangeable random graphs. *Rend. Mat. Appl. (7)* **28** (2008), no. 1, 33–61.

[12] Persi Diaconis and Bernd Sturmfels. Algebraic algorithms for sampling from conditional distributions. *Ann. Statist.* **26** (1998), no. 1, 363–397.

[13] Reinhard Diestel. *Graph theory.* Fourth edition. Graduate Texts in Mathematics, 173. *Springer, Heidelberg,* 2010. 437 pp.

[14] Adrian Dobra, Stephen E. Fienberg, Alessandro Rinaldo, Aleksandra Slavkovic, and Yi Zhou. Algebraic statistics and contingency table problems: log-linear models, likelihood estimation, and disclosure limitation. *Emerging applications of algebraic geometry*, 63–88, IMA Vol. Math. Appl., 149, *Springer, New York*, 2009.

[15] Anton Dochtermann and Alexander Engström. Algebraic properties of edge ideals via combinatorial topology. *Electron. J. Combin.* **16** (2009), no. 2, 24 pp.

[16] Anton Dochtermann and Alexander Engström. Cellular resolutions of cointerval ideals. *Math. Z.* **270** (2012), no. 1-2, 145–163.

[17] Mathias Drton, Bernd Sturmfels, and Seth Sullivant. *Lectures on algebraic statistics.* Oberwolfach Seminars, 39. *Birkhäuser Verlag, Basel,* 2009. 171 pp.

[18] Alexander Engström. A local criterion for Tverberg graphs. *Combinatorica* **31** (2011), no. 3, 321–332.

[19] Alexander Engström. Complexes of directed trees and independence complexes. *Discrete Math.* **309** (2009), no. 10, 3299–3309.

[20] Alexander Engström. Cut ideals of K_4-minor free graphs are generated by quadrics. *Michigan Math. J.* **60** (2011), no. 3, 705–714.

[21] Alexander Engström. Discrete Morse functions from Fourier transforms. *Experiment. Math.* **18** (2009), no. 1, 45–53.

[22] Alexander Engström. Independence complexes of claw-free graphs. *European J. Combin.* **29** (2008), no. 1, 234–241.

[23] Alexander Engström. Upper bounds on the Witten index for supersymmetric lattice models by discrete Morse theory. *European J. Combin.* **30** (2009), no. 2, 429–438.

[24] Alexander Engström, Thomas Kahle, and Seth Sullivant. Multigraded commutative algebra of graph decompositions, `arxiv:1102.2601`, 2011. 38 pp.

[25] Nicholas Eriksson, Stephen E. Fienberg, Alessandro Rinaldo, and Seth Sullivant. Polyhedral conditions for the nonexistence of the MLE for hierarchical log-linear models. *J. Symbolic Comput.* **41** (2006), no. 2, 222–233.

[26] Stephen E. Fienberg, Alessandro Rinaldo, and Yi Zhou. On the geometry of discrete exponential families with application to exponential random graph models. *Electron. J. Stat.* **3** (2009), 446–484.

[27] Robin Forman. Morse theory for cell complexes. *Adv. Math.* **134** (1998), no. 1, 90–145.

[28] William Fulton. *Introduction to toric varieties.* Annals of Mathematics Studies, 131. *Princeton University Press, Princeton, NJ,* 1993. 157 pp.

[29] Dan Geiger, Christopher Meek, and Bernd Sturmfels. On the toric algebra of graphical models. *Ann. Statist.* **34** (2006), no. 3, 1463–1492.

[30] Mikhail Gromov. Singularities, expanders and topology of maps. I. Homology versus volume in the spaces of cycles. *Geom. Funct. Anal.* **19** (2009), no. 3, 743–841.

[31] Mikhail Gromov. Singularities, expanders and topology of maps. Part 2: From combinatorics to topology via algebraic isoperimetry. *Geom. Funct. Anal.* **20** (2010), no. 2, 416–526.

[32] Martin Grötschel, László Lovász, and Alexander Schrijver. *Geometric algorithms and combinatorial optimization.* Second edition. Algorithms and Combinatorics, 2. *Springer-Verlag, Berlin,* 1993. 362 pp.

[33] Hisayuki Hara and Akimichi Takemura. A Markov basis for two-state toric homogeneous Markov chain model without initial paramaters. *J. Japan Statist. Soc.* **41** (2011), no. 1, 33–49.

[34] Hisayuki Hara and Akimichi Takemura. Markov chain Monte Carlo test of toric homogeneous Markov chains. *Stat. Methodol.* **9** (2012), no. 3, 392–406.

[35] David C. Haws, Abraham Martín del Campo, and Ruriko Yoshida. Degree bounds for a minimal Markov basis for the three-state toric homogeneous Markov chain model. Proceedings of the Second CREST–SBM International Conference "Harmony of Gröbner Bases and the Modern Industrial Society." 2012. 63–98.

[36] David C. Haws, Abraham Martín del Campo, Akimichi Takemura, and Ruriko Yoshida. Markov degree of the three-state toric homogeneous Markov chain model. `arXiv:1204.3070`, 2012. 20 pp.

[37] Pavol Hell and Jaroslav Nešetřil. *Graphs and homomorphisms.* Oxford Lecture Series in Mathematics and its Applications, 28. *Oxford University Press, Oxford,* 2004. 244 pp.

[38] Stephan Hell. Tverberg's theorem with constraints. *J. Combin. Theory, Ser. A* **115** (2008), no. 8, 1402–1416.

[39] Takayuki Hibi. Distributive lattices, affine semigroup rings and algebras with straightening laws. *Commutative algebra and combinatorics (Kyoto, 1985),* 93–109, Adv. Stud. Pure Math., **11**, *North-Holland, Amsterdam,* 1987.

[40] Takayuki Hibi and Hidefumi Ohsugi. Normal polytopes arising from finite graphs. *J. Algebra* **207** (1998), no. 2, 409–426.

[41] Serkan Hoşten and Seth Sullivant. Gröbner bases and polyhedral geometry of reducible and cyclic models. *J. Combin. Theory Ser. A* **100** (2002), no. 2, 277–301.

[42] Jakob Jonsson *Simplicial complexes of graphs.* Lecture Notes in Mathematics, 1928. *Springer-Verlag, Berlin,* 2008. 378 pp.

[43] Thomas Kahle. Neighborliness of marginal polytopes. *Beiträge Algebra Geom.* **51** (2010), no. 1, 45–56.

[44] Kazuhiro Kawamura. Independence complexes and edge covering complexes via Alexander duality. *Electron. J. Combin.* **18** (2011), no. 1, Paper 39, 6 pp.

[45] László Lovász. Kneser's conjecture, chromatic number, and homotopy. *J. Combin. Theory Ser. A* **25** (1978), no. 3, 319–324.

[46] László Lovász and Balázs Szegedy. Limits of dense graph sequences. *J. Combin. Theory Ser. B* **96** (2006), no. 6, 933–957.

[47] Ali Ridha Mahjoub. On the stable set polytope of a series-parallel graph. *Math. Programming* **40** (1988), no. 1, (Ser. A), 53–57.

[48] Ezra Miller and Bernd Sturmfels. *Combinatorial commutative algebra.* Graduate Texts in Mathematics, 227. *Springer-Verlag, New York,* 2005. 417 pp.

[49] Vladimir Nikiforov. The number of cliques in graphs of given order and size. *Trans. Amer. Math. Soc.* **363** (2011), no. 3,1599–1618.

[50] Alexander A. Razborov. On the minimal density of triangles in graphs. *Combin. Probab. Comput.* **17** (2008), no. 4, 603–618.

[51] Alexander A. Razborov. Flag algebras. *J. Symbolic Logic* **72** (2007), no. 4, 1239–1282.

[52] Garry Robins, Pip Pattison, Yuval Kalish, Dean Lusher. An Introduction to Exponential Random Graph (p^*) Models for Social Networks. *Social Networks* **29** (2007), no. 2, 173–191.

[53] Atle Selberg. Bemerkinger om et multipelt integral. *Norsk Mat. Tidsskr.* **26** (1944), 71–78.

[54] Bernd Sturmfels. *Gröbner Bases and Convex Polytopes*. University Lecture Series 8, *American Mathematical Society, Providence, RI,* 1996. 162 pp.

[55] Bernd Sturmfels and Seth Sullivant. Toric geometry of cuts and splits. *Michigan Math. J.* **57** (2008), 689–709.

[56] Seth Sullivant. Toric fiber products. *J. Algebra* **316** (2007), no. 2, 560–577.

[57] Siniša T. Vrećica and Rade T. Živaljević. The colored Tverberg's problem and complexes of injective functions. *J. Combin. Theory, Ser. A* **61** (1992), no. 2, 309–318.

[58] Günter M. Ziegler. Shellability of chessboard complexes. *Israel J. Math.* **87** (1994), no. 1-3, 97–110.

[59] Günter M. Ziegler. *Lectures on polytopes.* Graduate Texts in Mathematics, 152. *Springer-Verlag, New York,* 1995. 370 pp.

Errata

Publication I

Publication II

Publication III

Publication IV

Publication V

Aalto-DD 38/2013

ISBN 978-952-60-5047-8
ISBN 978-952-60-5048-5 (pdf)
ISSN-L 1799-4934
ISSN 1799-4934
ISSN 1799-4942 (pdf)

Aalto University
School of Science
Department of Mathematics and Systems Analysis
www.aalto.fi

BUSINESS +
ECONOMY

ART +
DESIGN +
ARCHITECTURE

SCIENCE +
TECHNOLOGY

CROSSOVER

DOCTORAL
DISSERTATIONS