UPPSALA DISSERTATIONS IN MATHEMATICS

105

Degrees in Random Graphs and Tournament Limits

Erik Thörnblad

Department of Mathematics
Uppsala University
UPPSALA 2018

Dissertation presented at Uppsala University to be publicly examined in Polhemssalen, Ångströmlaboratoriet, Lägerhyddsvägen 1, Uppsala, Friday, 9 March 2018 at 13:15 for the degree of Doctor of Philosophy. The examination will be conducted in English. Faculty examiner: Professor Daniel Kráľ (University of Warwick).

Abstract

Thörnblad, E. 2018. Degrees in Random Graphs and Tournament Limits. *Uppsala Dissertations in Mathematics* 105. 26 pp. Uppsala: Department of Mathematics.
ISBN 978-91-506-2677-3.

This thesis consists of an introduction and six papers on the topics of degree distributions in random graphs and tournaments and their limits.

The first two papers deal with a dynamic random graph, evolving in time through duplication and deletion of vertices and edges. In Paper I we study the degree densities of this model. We show that these densities converge almost surely and determine their limiting values exactly as well as asymptotically for large degrees. In Paper II we study the evolution of the maximum degree and provide a precise growth rate thereof.

Paper III deals with a dynamic random tree model known as the vertex-splitting tree model. We show that the degree densities converge almost surely and find an infinite linear system of equations which they must satisfy. Unfortunately we are not able to show that this system has a unique solution except in special cases.

Paper IV is about self-converse generalised tournaments. A self-converse generalised tournament can be seen as a matrix whose entries take values in [0,1] and whose diagonally opposite elements sum to 1. We characterise completely the marginals of such a matrix, and show that such marginals can always be realised by a self-converse generalised tournament.

In Paper V, we define and develop the theory of tournament limits and tournament kernels. We characterise transitive and irreducible tournament limits and kernels, and prove that any tournament limit and kernel has an essentially unique decomposition into irreducible tournament limits or kernels interlaced by a transitive part.

In Paper VI, we study the degree distributions of tournament limits, or equivalently, the marginals of tournament kernels. We describe precisely which distributions on [0,1] which may appear as degree distributions of tournament limits and which functions from [0,1] to [0,1] may appear as the marginals of tournament kernels. Moreover, we show that any distribution or marginal on this form may be realised by a tournament limit or tournament kernel. We also study those distributions and marginals which can be realised by a unique tournament limit or kernel, and find that only the transitive tournament limit/kernel gives rise to a degree distribution or marginal with this property.

Keywords: Random graphs, degree distributions, degree sequences, graph limits, tournaments

Erik Thörnblad, Department of Mathematics, Box 480, Uppsala University, SE-75106 Uppsala, Sweden.

ISSN 1401-2049
ISBN 978-91-506-2677-3
urn:nbn:se:uu:diva-339025 (http://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-339025)

We can forgive a man for making a useful thing as long as he does not admire it. The only excuse for making a useless thing is that one admires it intensely.

— Oscar Wilde

List of papers

This thesis is based on the following papers, which are referred to in the text by their Roman numerals.

I E. Thörnblad. Asymptotic degree distribution of a duplication–deletion random graph model. *Internet Mathematics*, 11(3):289-305, 2015. doi:10.1080/15427951.2015.1009523

II E. Thörnblad. The dominating colour of an infinite Pólya urn. *Journal of Applied Probability*, 53(3):914-924, 2016. doi:10.1017/jpr.2016.49

III S. Ö. Stéfansson and E. Thörnblad. Almost sure degree distribution of the vertex–splitting model. *Stochastic Models*, 32(4):575-592, 2016. doi:10.1080/15326349.2016.1182029

IV E. Thörnblad. Eplett's theorem for self–converse generalised tournaments. To appear in *Australasian Journal of Combinatorics*, 70(3), 2018.

V E. Thörnblad. Decomposition of tournament limits. *European Journal of Combinatorics*, 67:96-125, 2018. doi:10.1016/j.ejc.2017.07.023

VI E. Thörnblad. Tournament limits: Score functions, degree distributions and self–converseness, 2017. *Manuscript.* arXiv:1611.09579

Additional papers

The following papers are not included in this thesis.

VII K. Gabrysch and E. Thörnblad. The greedy walk on an inhomogeneous Poisson process, 2016, *Manuscript.* arXiv:1611.09568

VIII E. Thörnblad. Another proof of Moon's theorem on generalised tournament score sequences, 2016, *Manuscript.* arXiv:1605.06407

IX E. Thörnblad. Tournament limits: degree distributions and score functions. *Conference proceedings BGW 2016.*

X E. Thörnblad and J. Zimmermann. Counting quasi-idempotent irreducible integral matrices, 2017, *Manuscript.* arXiv:1701.03699

Contents

1. Introduction

This thesis consists of an introduction, a short summary of the six papers, and the papers themselves. The papers deal with various topics in graph theory. In this introduction, we introduce graphs, their degree sequences, random graphs and limits of graphs. These are the central objects in the six papers. Although the papers typically deal with specific classes of graphs, for simplicity we focus on so called simple graphs in this introduction.

1.1 Graphs

Throughout this introduction, a *graph* will mean a pair (V, E), where V is a finite set (called the vertex set) and E is a symmetric relation on V (called the edge set), such that $(v, v) \notin E$ for all $v \in V$. Often these are called *simple graphs*. Graphs are typically drawn in the plane as in Figure 1.1, with points representing vertices and lines between them representing edges. Naturally, there are infinitely many ways to embed a graph in the plane. Depending on the problem at hand, the choice of embedding may or may not be relevant.

Each graph also corresponds to a so–called *adjacency matrix*, which is a $0-1$ valued matrix in which the entry in row i and column j is 1 if and only if (i, j) is an edge of the graph. This matrix contains all the information stored in the pair (V, E), and it is well–defined up to simultaneous permutation of rows and columns (this corresponds to relabelling of the vertices). For instance, the graph in Figure 1.1 has adjacency matrix

$$\begin{array}{c} \begin{array}{cccccc} v_1 & v_2 & v_3 & v_4 & v_5 & v_6 \end{array} \\ \begin{pmatrix} 0 & 1 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 0 \end{pmatrix} \begin{array}{c} v_1 \\ v_2 \\ v_3 \\ v_4 \\ v_5 \\ v_6 \end{array} \end{array}.$$

Often one does not care about the vertex labels and instead consider unlabelled graphs. Two labelled graphs induce the same unlabelled graph if their adjacency matrix are identical up to simultaneous row and column permutations.

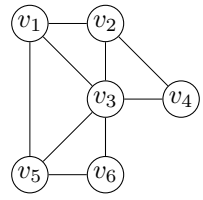

Figure 1.1. A planar drawing of the graph with vertex set $\{v_1, v_2, v_3, v_4, v_5, v_6\}$ and edges $\{(v_1, v_2), (v_1, v_3), (v_1, v_5), (v_2, v_3), (v_2, v_4), (v_3, v_4), (v_3, v_5), (v_3, v_6), (v_5, v_6)\}$.

Properties which are independent of the vertex labels are typically called graph properties. A simple graph property which appears in all papers in this thesis is the distribution of the the degrees of the graph. The *degree* of a vertex is the number of incident edges; equivalently the sum of the corresponding row in the adjacency matrix. For instance, the graph in Figure 1.1 has degree sequence $5,3,3,3,2,2$ when listed in decreasing order. Given a sequence $(d_i)_{i=1}^n$ of non-negative integers, one question is whether there exists a graph with this sequence as its degree sequence. For graphs as defined in this introduction, the answer is provided by the Erdős–Gallai theorem [11], which states that $d_1 \geq d_2 \geq \cdots \geq d_n$ is the degree sequence of some graph if and only if $d_1 + \cdots + d_n$ is even and

$$\sum_{i=1}^{k} d_i \leq k(k-1) + \sum_{i=k+1}^{n} \min(d_i, k)$$

for all $1 \leq k \leq n$. The evenness of the sum of degrees is an immediate consequence of the fact that each edge is incident to precisely two vertices, and there are numerous proofs that the second condition is necessary and sufficient. In Paper IV and VI we deal with similar problems for the classes of directed graphs and their limits.

1.1.1 Random graphs

One direction in graph theory research is to introduce randomness. The classical example of a random graph model is the Erdős-Renyi graph $G(n,p)$. In this model, a graph is drawn from the set of all graphs on n vertices, chosen with probability proportional to $p^k(1-p)^{n-k}$, where k is the number of edges of the graph. Equivalently, a graph sampled from $G(n,p)$ can be generated by flipping a p-biased coin independently for each possible edge, including the edge with probability p and excluding it with probability $1-p$. These graphs have been studied since the late 1950s and produced a wide range of interesting results. The interested reader is refered to [4].

In this thesis we will not deal with the Erdös-Renyi model, but rather with a general extension of it. This is known as the W–random graph $G(n,W)$, where $W : [0,1]^2 \to [0,1]$ is a measurable function satisfying $W(x,y) = W(y,x)$ for

almost all $(x,y) \in [0,1]^2$. To construct a graph from this ensemble, first take an independent sample $X_i \in U[0,1]$, $1 \leq i \leq n$. Then, for all pairs $1 \leq i < j \leq n$, draw an edge between i to j with probability $W(X_i, X_j)$; otherwise no edge is drawn between i and j. By taking $W = p$ everywhere we obtain the Erdös-Renyi model as a special case. These graphs appear in Paper V and VI.

Sometimes one allows for graphs to evolve with time, for instance by (random) addition of vertices and edges as time progresses. The seminal example is the *preferential attachment model*, defined as follows. Start with $\mathscr{T}_0$, a graph consisting of two vertices joined by an edge. For $n \geq 1$, form $\mathscr{T}_n$ by introducing a new vertex and joining it to a vertex in $\mathscr{T}_{n-1}$, chosen with probability proportional to its degree. Although $(\mathscr{T}_n)_{n\geq 1}$ is a stochastic process, the fluctuations created by its random construction tend to smooth out over time, with many of its properties being well–behaved in the limit as $n \to \infty$. For instance, if $d_{k,n}$ is the random variable equal to the proportion of vertices of degree k in $\mathscr{T}_n$, then

$$d_{k,n} \cdot \frac{k(k+1)(k+2)}{4} \to 1 \tag{1.1}$$

almost surely as $n \to \infty$, for each $k \geq 1$. A proof of this appears in [5]. Such a stability of the asymptotic degree distribution has been observed in many dynamic random graph models. We will see some examples of this in Papers I, II and III.

The degree distribution of the preferential attachment model is an example of a *power law*. These frequently occur when considering graphs evolving by rules similar to the preferential attachment model. However, there are many small variations of the preferential attachment model which will break the power law property. In Paper I we consider a preferential attachment–type graph model with deletion of edges. This satisfies a power law in one parameter range, but exhibits another type of degree distribution in other parameter ranges.

1.2 Graph limit theory

As hinted at in the preceding section, interesting questions appear when studying limits of large (random) graph sequences. Of course, there are many ways to "take limits" of graph sequences. One is the *graph limit theory*, which we introduce now. Most claims we make below are highly non–trivial, but proofs to all of them may be found in the comprehensive monograph [14]. Graph limit theory is connected to the Aldous–Hoover theory of exchangeable arrays, a correspondence which is outlined in [9]. In this thesis, we meet the graph limit theory in Papers V and VI.

Graph limit theory deals with large graphs, or rather limits of graph sequences for which the number of vertices tends to infinity. The key concept is

that of "small subgraph densities". Two graphs can be deemed similar if these numbers are similar for the two graphs. This can be encapsulated through the metric defined by

$$d(G,H) = \sum_{n=1}^{\infty} \frac{|t(F_n,G) - t(F_n,H)|}{2^n}$$

for any graphs G and H, where $F_1, F_2, \ldots$, is an enumeration of all finite graphs, and where $t(F,G)$ denotes the probability that a random map $V(F) \to V(G)$ is a graph homomorphism. Clearly this metric is invariant under relabeling of nodes, and the number $t(F,G)$ should be thought of as the density of F inside G. If $d(G,H)$ is small, then G and H have similar subgraph densities for all subgraphs, and G and H can be thought of as similar graphs. The subgraph densities are the main object of study in graph limit theory. Indeed, the numbers $(t(F_i,G))_{i=1}^{\infty}$ and $(t(F_i,H))_{i=1}^{\infty}$ uniquely determine $d(G,H)$.

We are interested in graph sequences where the number of vertices tends to infinity. To define the first type of limit object in the theory, we identify each graph with a point in the Cantor cube $[0,1]^{\infty} \times [0,1]$ via the injective map $G \mapsto (t(F_i,G))_{i=1}^{\infty} \times (|V(G)|^{-1})$. Consider now the closure of the set $\zeta = \{(t(F_i,G))_{i=1}^{\infty} \times (|V(G)|^{-1}) \ : \ G \in \mathscr{G}\} \subseteq [0,1]^{\infty} \times [0,1]$ with respect to the sup norm. It is obvious that the set $\overline{\zeta} \setminus \zeta$ is non-empty. (For instance, the sequence K_n of complete graphs on n vertices, converges to a limit. Since $|V(K_n)| = n \to \infty$, the last coordinate of the limit must be zero, and there is no finite graph with whose last coordinate is zero.)

Formally speaking, the set of graph limits is defined to be the set $\overline{\zeta} \setminus \zeta \subseteq [0,1]^{\infty} \times \{0\}$, so a graph limit is simply a point in $[0,1]^{\infty} \times \{0\}$. These fill in the holes in $\overline{\zeta}$, much like how the irrationals fill in the holes in the real number line. Graph limits are often denoted by Γ. If $(t(F_i,G_n))_{i=1}^{\infty} \times (|V(G_n)|^{-1}) \to \Gamma$ as $n \to \infty$, then we say that $(G_n)_{n=1}^{\infty}$ *converges* to Γ, and we write $G_n \to \Gamma$. Moreoever, we write $t(F,\Gamma) := \lim_{n\to\infty} t(F,G_n)$ and call $t(F,\Gamma)$ the density of F inside Γ. Since the subgraph densities are simply the projections onto all but the last one–dimensional subspaces, a graph limit is uniquely determined by its subgraph densities.

1.2.1 Kernels and graphons

There is another type of limit object, called a *graphon* or a *kernel*. This is a measurable function $W : [0,1]^2 \to [0,1]$ satisfying $W(x,y) = W(y,x)$ almost everywhere. A kernel can be seen as the continuous analogue of the adjacency matrix of a graph. In particular, each graph G on $\{1,2,\ldots,n\}$ induces a kernel defined by

$$W_G(x,y) = \begin{cases} 1 & \text{if } (i-1)/n < x < i/n, (j-1)/n < y < j/n, (i,j) \in E(G) \\ 0 & \text{otherwise.} \end{cases}$$

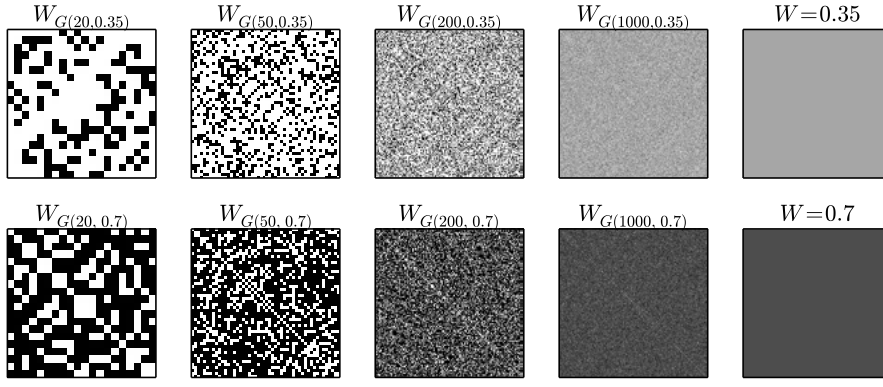

Figure 1.2. Plots of two different kernel sequences and their limiting kernel. The origin is located the top left corner of each subfigure. The values are represented by a grayscale, where black is 1 and white is 0. (Note that all but the two rightmost images consist only of black and white pixels.) Each of the eight non-constant kernels were randomly generated from the $G(n,p)$ ensemble with parameters as specified in the subfigure titles.

(This is simply a way to place the adjacency matrix inside the $[0,1]^2$.) For reasons which will become clear, this needs only be defined up to measure zero.

Parallel to each graph sequence $(G_n)_{n=1}^{\infty}$, we have a corresponding sequence of kernels $(W_{G_n})_{n=1}^{\infty}$. If the former converges to some graph limit Γ, we might hope that the second also converges, this time in the space of functions $[0,1]^2 \to [0,1]$.

In Figure 1.2, we have plotted two different kernel sequences arising from the $G(n,p)$ ensemble with different parameters. Although these are randomly generated, the plots suggest that the randomness will not be visible in the limit. Indeed, a calculation yields that $\int_{A\times A} W_{G(n,p)}(x,y)dxdy \to |A|^2 p$ almost surely for any measurable $A \in [0,1]$, a result which may suggest that the constant kernel $W = p$ is the correct limit kernel of a graph sequence generated from $G(n,p)$ with $n \to \infty$. (This viewpoint is related to the theory of quasirandom graphs [6].) We return to this example later.

The correct norm to consider for kernel convergence is the *cut norm* $\|\cdot\|_\square$, defined by

$$\|W\|_\square := \sup_{S,T} \left| \int_{S\times T} W(x,y)dxdy \right|.$$

where the supremum is over all measurable subsets $S,T \subseteq [0,1]$. The cut metric $d_\square$ is defined by

$$d_\square(W_1,W_2) := \inf_{\phi_1,\phi_2} \|W_1^{\phi_1} - W_2^{\phi_2}\|_\square$$

where the infimum is over all measure-preserving transformations $\phi_1, \phi_2 : [0,1] \to [0,1]$ and $W^{\phi}(x,y) := W(\phi(x), \phi(y))$. Note that the cut metric is only a pseudo-metric, since it identifies elements which differ on a set of measure 0. In the finite case, we mentioned that graph similarity should not depend on the labels of the graphs. Taking the infimum over all measure–preserving transformation has the same effect on kernel similarity.

If $(W_n)_{n=1}^{\infty}$ satisfies $d_{\square}(W_n, W) \to 0$ as $n \to \infty$, we say that W_n converges to W. If $W_n = W_{G_n}$ are the induced kernel from a graph sequences, we typically abuse notation and write $G_n \to W$.

It is also possible to define subgraph densities for kernels. If F is a graph on k vertices and W any kernel, we define

$$t(F,W) := \int_{[0,1]^k} \prod_{(i,j)\in E(F)} W(x_i, x_j) \prod_{i=1}^{k} dx_i.$$

It can be shown that $t(F,G)$ and $t(F,W_G)$ agree up to a small error term which vanishes as the number of vertices of G increases. Therefore the numbers $t(F,G_n)$ converge if and only if the numbers $t(F,W_{G_n})$ converge. This gives a dual notion of convergence, which also implies that graph limits can be seen as equivalence classes of kernels. That is, a kernel represents a limit if and only if $t(F,W) = t(F,\Gamma)$ for all graphs F.

The equivalence between cut norm convergence and subgraph density convergence of kernels implies that each limit is represented by some kernel. However, since limits were defined with reference to kernels, it does not imply that each kernel represents a limit. (It could be that there is a kernel which is not the limit of any graph sequence). This does however turn out to be true. Recall from Section 1.1.1 that each kernel W induces a W-random graph $G(n,W)$. It can be shown that $G(n,W) \to W$ almost surely, implying that each kernel is the limit of some graph sequence, which means that each kernel indeed represents some limit. (As mentioned, this is not an obvious result. By constrast, the fact that each graph limit is the limit of some graph sequence follows by definition.)

Let us return to the sequences in Figure 1.2. For any graph F on k vertices, it holds that

$$t(F, W_{G(n,p)}) \to \int_{[0,1]^k} \prod_{(i,j)\in E(F)} p \prod_{i=1}^{k} dx_i = t(F,p),$$

almost surely, meaning that $G(n,p) \to p$ almost surely. This graph sequence of course also converges to a limit, which may be defined by its subgraph densities

$$t(F,\Gamma) := t(F,p).$$

It should be noted that our version of graph limit theory concerns itself only with dense graph sequences, i.e. sequences for which the number of

edges grows quadratically in the number of vertices. If $(G_n)_{n=1}^{\infty}$ is a graph sequence for which the number of edges grows subquadratically in the number of vertices, then $G_n \to 0$. Recently attempts have been made to extend theory to non–dense graph sequences, but we do not pursue this further here.

2. Summary of papers

2.1 Paper I

Let $0 < p < 1$ and let $\mathscr{G}_0$ be the simple graph consisting of a single vertex. At time n, construct $\mathscr{G}_n$ from $\mathscr{G}_{n-1}$ as follows. With probability p, add a new vertex to the graph, and join it to a vertex chosen uniformly at random and all its neighbours. With probability $1-p$, select a vertex in the graph and remove all its incident edges (so that the vertex becomes isolated).

This random graph model is a version of a model introduced in [2]. In each step, the graph is a union of disjoint complete subgraphs. Although we do not take this viewpoint in this paper, it can be helpful to view these subgraphs as independent birth–death processes. The three cases $0 < p < 1/2$, $p = 1/2$ and $1/2 < p < 1$ turn out to be qualitatively different. In the first case, each complete subgraph will, with probability 1, shrink to just an isolated vertex infinitely many times. In the last case, each complete subgraph has a non–zero probability of "surviving", i.e. its size grows to infinity as $n \to \infty$.

Let d_k be the limiting proportion of the number of vertices of degree k. Our main result is that this exists and satisfies an exact integral formula, which yields the asymptotic results

$$d_k \sim \begin{cases} (-\beta)^{-1}(1-\beta)^{1-\beta}\Gamma(1-\beta)\gamma^{-k}k^{\beta} & \text{if } 0 < p < 1/2, \\ (e\pi)^{1/2}k^{1/4}e^{-2\sqrt{k}} & \text{if } p = 1/2, \\ \gamma\beta^{\beta}\Gamma(\beta+1)k^{-\beta} & \text{if } 1/2 < p < 1, \end{cases}$$

as $k \to \infty$, where $\gamma = (1-p)/p$, $\beta = p/(2p-1)$ and Γ is the Gamma function. That is, in the case $0 < p < 1/2$, the degree distribution decays exponentially, when $p = 1/2$ it lies between a power law and exponential decay, and in the case $1/2 < p < 1$ it satisfies a power law.

2.2 Paper II

We continue to study the dynamic random graph model introduced in Paper I. Móri and Backhausz [3] studied the same model and determined the growth rate of the maximum degree in the cases $0 < p < 1/2$ and $p = 1/2$. We fill in the remaining hole and study the evolution of the maximum degree in the case $1/2 < p < 1$. Our main result is that it grows (almost surely) like $\mu n^{(2p-1)/p}$

as $n \to \infty$, for some random variable $\mu > 0$, where n is the number of vertices in the graph.

Instead of considering individual vertices, we consider the complete subgraphs of the graph. As mentioned, these grow essentially independently. To be able to explot this, we perform a continuous–time embedding of the model. Each vertex is equipped with two exponential clocks – one ringing at rate p and the other ringing at rate $1-p$. If a clock of the first type rings, we add a vertex to the clique of the vertex whose clock rang. If a clock of the second type rings, we remove all edges of vertex whose clock rang. By observing this model at the times where the graph changes, we retrieve the model in Paper I. However, we are now able to fully exploit independence between different cliques.

The idea behind the remainder of the proof is inspired by Galashin [12], who demonstrated a similar result for the preferential–attachment model. We show that there is a unique complete subgraph that remains the largest after some random but finite time, even though all its vertices gets replaced by deletion and duplication. Since the graph consists of disjoint complete subgraphs, the growth rate of the maximum degree must then match the growth rate of the size of the largest complete subgraph. Finally, the growth rate of each individual complete subgraph can be analysed (it evolves independently of everything else).

2.3 Paper III

We study the vertex–splitting tree model introduced in [7, 8]. This is a dynamic random tree model, which is embedded in the plane. Each vertex is given a weight according to its current degree. Each step, a vertex is selected at random with probability proportional to its weight. The selected vertex v is split into two connected vertices v' and v''. The neighbours of v are split into two sets E' and E'' (randomly, according to some other weights), and v' and v'' is made adjacent to all vertices of E' and E'' respectively.

This model generalises some other known models, including the preferential–attachment model, the random plane recurive tree model and the random recursive tree model.

We study the case when the weight of each vertex of degree k is of the form $ak+b$. In [7], it was shown that the expected values of the degree densities converge to a limit described by an infinite linear system of equations. We show the stronger result that the degree densities converge almost surely to the same infinite linear system of equations. The proof method is similar to that of Paper I. Unfortunately, we are not able to show that this system has a unique solution except in special cases. We also consider a two–coloured version of this model as described in [8].

2.4 Paper IV

A generalised tournament G on n vertices is a weighted digraph (without loops) on n vertices, satisfying the condition that the weights of any two edges joining the same two vertices must sum to 1. The converse is obtained by replacing each weight x by a weight $1-x$. If the resulting weighted digraph is isomorphic to the original, then the generalised tournament is said to be self-converse.

In this paper we characterise score sequences of self-converse generalised tournaments. For a generalised tournament on n vertices, we let d_i be the sum of the weights of edges leaving vertex i. The sequence $(d_i)_{i=1}^n$ is known as the score sequence of the generalised tournaments.

Theorem. *A sequence $d_1 \geq d_2 \geq \cdots \geq d_n$ is the score sequence of some generalised tournament if and only if*
(i) $\sum_{i=1}^k d_i \geq \binom{k}{2}$ *for all* $k = 1, \ldots, n$, *with equality if* $k = n$, *and*
(ii) $d_i + d_{n+1-i} = n-1$ *for all* $1 \leq i \leq n$.

This extends an old result about self-converse non-generalised (i.e. where all edge weights are 0 or 1) tournaments due to Eplett [10]. In fact, our proof uses this result to first prove a corresponding result for rational score sequences. A compactness argument allows us to extend this to arbitrary real edge weights.

2.5 Paper V

In this paper we study the limits and kernels of tournaments. (A tournament is a directed complete graph.) These are defined analogously to graph limits and kernels as in Section 1.2. In particular, a tournament kernel is a measurable function $W : [0,1]^2 \to [0,1]$ satisfying $W(x,y) + W(y,x) = 1$ almost everywhere.

The theory of tournament kernels and limits has not been previously written down in the literature, so we do a significant amount of groundwork in this paper. The theory behaves largely as expected, and many results and definitions for finite tournaments extend straightforwardly to the setting of tournament limits and tournament kernels.

Two key notions in tournament theory are those of transitivity and irreducible. A tournament is said to be transitive if it contains no directed cycle (equivalently, if it contains no directed triangle). Analogously we define transitive tournament kernels to be those kernels W for which $t(C_3, W) = 0$. Transitive tournaments can be defined in many equivalent ways, and we show that transitive tournament kernels also have many equivalent definitions.

A tournament is said to be reducible if its vertex set can be decomposed into two non–empty sets A and B, such that all edges between the two sets go

from A to B. Otherwise it is said to be irreducible. Analogously, a tournament kernel W is reducible if there exists non–null measurable $A, B \subseteq [0,1]$ such that $W(x,y) = 1$ for almost all $(x,y) \in A \times B$. Otherwise it is said to be irreducible.

We also define transitive and irreducible tournament limits. One key point is that the definitions for kernels and limits must be sensible, in the sense that if a tournament limit is transitive (irreducible), then any kernel it represents should also be transitive (irreducible). We also define a notion of direct sum. Intuitively speaking, if a tournament kernel or tournament limit is reducible, then it decomposes into a direct sum of at least two subkernels or sublimits. (Irreducibility is therefore equivalent to being indecomposable.)

A classical result [15] for finite tournaments states that that any finite tournament has a unique decomposition into induced irreducible subtournaments $G_1, \dots, G_m$, such that all edges go from G_i to G_j for any $i < j$. Inspired by this result, our main aim is to prove a similar result for tournament kernels and tournament limits. Our main result is that each tournament kernel has an essentially unique decomposition into a decomposition into irreducible components (interlaced by a transitive part). Since each tournament limit is represented by many tournament kernels, this gives potentially many different decompositions of the tournament limit. However, we are able to show that the induced decomposition of the tournament limit is unique up to order isomorphism of the indexing set.

2.6 Paper VI

In this paper we study the degree distributions of tournament limits and kernels, and score functions of tournament kernels. The *degree distribution* of a tournament limit Γ is the unique probability distribution on $[0,1]^2$ with (m,n):th moment equal to $t(S_{m,n}, \Gamma)$, where $S_{m,n}$ denotes the directed graph with a central vertex, joined by outgoing edges to m other vertices, and by incoming edges to n other vertices (none of which are the same). The *score function* of a kernel W is defined as its first marginal $f_W(x) := \int_0^1 W(x,y)dy$.

Our main result is the following.

Theorem. *Let $f : [0,1] \to [0,1]$ be a measurable function. The following statements are equivalent.*

(i) The function f satisfies $\int_B f(x)dx \geq Leb(B)^2/2$ for any measurable $B \subseteq [0,1]$, with equality whenever $Leb(B) = 1$.

(ii) There exists a tournament kernel with score function f.

(iii) There exists a tournament limit Γ for which the first marginal of its degree distribution is distributed like $f(U)$, where $U \sim [0,1]$.

This can be seen as a generalisation of Landau's theorem [13] on admissible score sequences, which states that a a sequence $(d_i)_{i=1}^n$ is the outdegree

sequence of a tournament on n vertices if and only if $\sum_{i\in J} d_i \geq \binom{|J|}{2}$ for any subset $J \subseteq \{1,\dots,n\}$, with equality if $J = \{1,\dots,n\}$.

A natural secondary question is which score functions can be realised by a unique (up to weak isomorphism) tournament kernel. We show that this only happens for the score function $f(x) = x$ (and its relatives), which is realised by the transitive tournament limit $W(x,y) = \mathbb{1}(x \leq y)$. Equivalently, the all degree distributions are realised by many distinct tournament limits except that of the transitive limit. This mirrors a result due to Avery [1], who characterised out-degree sequences of tournaments which are realised by a unique tournament. In particular, these tournaments are close to being transitive (indeed, they converge to the transitive tournament limit). In light of this, our result should not be surprising.

Finally we consider self-converse tournament limits and their score functions, and provide a partial characterisation of these.

3. Summary in Swedish

Denna avhandling består av en inledning och sex artiklar. Gemensamt för samtliga artiklar är att de ligger inom ramen för ämnet grafteori. I de tre första artiklarna studeras två sorters slumpgrafer som växer med tiden enligt en slumpmässig regel. I de tre sista artiklarna studeras turneringar och deras gränsvärden.

Artikel I och II behandlar en slumpgrafsmodell som växer med tiden. Varje tidsteg utför vi följande operation. Med sannolikhet p skapas en kopia av en slumpmässigt vald tidigare existerande nod. Med sannolikhet $1-p$ raderas alla kanter från en tidigare existerande nod. I artikel I fokuserar vi på gradtalsfördelningen då tiden går mot oändligheten. Även om grafen växer slumpmässigt visar det sig att gradtalsdensiteterna konvergerar nästan säkert mot ett gränsvärde som vi kan bestämma exakt. I artikel II studerar vi hur det maximala gradtalet beter sig då tiden går mot oändligheten. Även denna parameter visar sig växa på ett förutsägbart sätt trots att den underliggande parametern är slumpmässig.

I artikel III studerar vi en annan slumpgrafsmodell som växer med tiden. I denna modell är varje graf ett träd, vilket betyder att det inte finns några cykler. Återigen studerar vi den asymptotiska gradtalsfördelningen, och visar att denna existerar nästan säkert. Gränsvärdesfördelningen visar sig vara lösningen till ett oändligt ekvationssystem. Detta bekräftar en tidigare förmodan som fanns i litteraturen, även om vi inte kan visa att ekvationssystemet alltid har en entydig lösning.

I artikel IV, V och VI studerar vi turneringar, det vill säga är kompletta grafer där varje kant fått en riktning. Artikel IV berör så kallade generalisade turneringar. Här tillåter vi två viktade riktade kanter mellan varje par av noder, med villkoret att vikten mellan varje par av kanter mellan två noder ska summera till 1. Vi studerar en speciell typ av generaliserade turnerningar som har egenskapen att de är isomorfa till den generaliserade turnering man får om man ersätter varje kantvikt x med $1-x$.

Artikel V och VI behandlar så kallade grafgränsvärden för turneringar. Gränsvärden av grafer kan studeras på många olika sätt. Angreppssätt vi väljer i de två sista artiklarna är relativt nytt och grundar sig en i teori som utvecklats främst under 2000-talet. I denna teori definierar man ett avståndsmått, där två stora grafer ligger nära varandra om de har liknande densiteter av mindre delgrafer. Detta ger upphov till en typ av konvergens. En annan, i stort sett ekvivalent, typ av konvergens ges av den så kallade $\square$–metriken, som bygger på en slags integralnorm. Gränsvärdesobjekten kallas för grafgränsvärden

eller kärnor, där grafgränsvärdena kan ses som ekvivalensklasser av kärnorna. Dessa är rent algebraiska respektive rent analytiska objekt, vilket innebär att teorin blir en sammanvävning av grafteori, algebra, analys och integrationsteori. Detta gör den extremt flexibel, i den mening att man ofta kan välja hur man vill angripa ett visst problem.

Teorin som vi använder är endast intressant ur ett matematiskt perspektiv så länge man har att göra med grafföljder där antalet kanter växer kvadratiskt med antalet noder. Om antalet kanter växer subkvadratiskt, så blir motsvarande gränsvärdesobjekt helt enkelt det triviala 0–gränsvärdet. Exempelvis går det därmed inte att studera träd i denna teorin, eftersom dessa har ett linjärt antal kanter i antalet noder. Det bör dock sägas att motsvarande teorier har utvecklats för grafföljder där antalet kanter växer subkvadratiskt i antalet noder, men vi går inte närmare in på det i denna avhandling.

Vårt bidrag är att utveckla grafgränsvärdesteorin till gränsvärden av turneringar. Dessa har ett kvadratiskt antal kanter i antalet noder, så vi kan förvänta oss att se icke-triviala gränsvärden. I artikel V formulerar vi i första hand en koherent teori för turneringsgränsvärden. I andra hand visar vi att dessa har en entydig uppdelning i så kallade irreducibla gränsvärden. Detta speglar ett känt uppdelningsresultat för ändliga turneringar. I artikel VI studerar vi gradtalsfördelningen hos turneringsgränsvärden och ger villkor på vilka fördelningar som kan vara gradtalsfördelningar för turneringsgränsvärden.

4. Acknowledgements

Above all I owe my gratitude to my adviser Svante, for his direction and guidance over the past five years, and for spotting endless mistakes in my papers. Thanks also to my coadviser Erik who always kept me entertained at work and to Cecilia who always gave me kind advice and believed in me. I am also grateful to my first collaborator Sigurður, for many interesting discussions, and to Walter, for taking such good care of all the Phd students.

This department would not be what it is without its fantastic administration staff. My thanks go to Elisabeth, Inga-Lena and Lisbeth. I asked you more questions than I thought existed (and all were answered!). In particular, thanks to Fredrik for being so supportive during my first years of teaching.

Thank you to all current and former PhD students, including Anna, Brendan, Kostas, Jonathan, Juozas, Linnéa, Marta, Natasha and Sam, for being my colleagues and becoming my friends. I am particularly grateful to Björn, for being full of good advice and for being my rival in the racquet sports, to Daniah and Ernest, for being such fun to talk with, and to Filipe and Tilo, for always being so helpful.

I am lucky enough to have been collaborating with several PhD students at the department. Andrea, thank you for telling me about the susejd rol. You always have something bizarre to say. Jakob, you made me enjoy research again and "gifted" me a paper. Thanks! Johanna, thank you for being a great office mate. Katja, I enjoyed the three years we spent together in the office. Thank you for taking care of me when I moved to the probability corridor! I enjoyed collaborating with all four of you – I would suggest more of it between the PhD students. Although many of us work in different areas, evidently it is possible to find common denominators.

Thank you to all my friends for sticking with me. Alicia, thank you for moral support – I know you will become a great researcher. Anja, thank you for being my wailing wall and greatest supporter. You taught me more than anyone else. Bing, thank you for always being there for me and for always listening to me talk about my research. Iuliana, thank you for seeing me through university – I know I would not have made it without you. Wenting and Xiao, thank you both for studying, eating and discussing with me. Many thanks also to David, Dylan, Jim, Nico, Shyam and all my friends at UASS for chess lunches, chess fikas and chess discussions. I will admit there were many a day I came to work only because we had scheduled a chess lunch.

Finally I would like to thank my family for all the support they give.

References

[1] P. Avery. Condition for a Tournament Score Sequence to be Simple. *Journal of Graph Theory*, 4:157–164, 1980.

[2] Á. Backhausz and T. Móri. Asymptotic properties of a random graph with duplications. *Journal of Applied Probability*, 52(2):375–390, 2015.

[3] Á. Backhausz and T. Móri. Further properties of a random graph with duplications and deletions. *Stochastic Models*, 32(1):99–120, 2015.

[4] B. Bóllobas. *Random Graphs*. Cambridge University Press, second edition, 2001.

[5] B. Bóllobas, O. Riordan, J. Spencer, and G. Tusnády. The Degree Sequence of a Scale-Free Random Graph Process. *Random Structures and Algorithms*, 18(3):279–290, 2001.

[6] F. R. K. Chung, R. L. Graham, and R. M. Wilson. Quasi-Random Graphs. *Combinatorica*, 9(4):345–362, 1989.

[7] F. David, M. Dukes, T. Jonsson, and S. u. Ö. Stefansson. Random Tree Growth by Vertex Splitting. *Journal of Statistical Mechanics Theory and Experiment*, 04, 2009. doi:10.1088/1742-5468/2009/04/P04009.

[8] F. David, C. Hagendorf, and K. J. Wiese. A growth model for RNA secondary structures. *Journal of Statistical Mechanics Theory and Experiment*, 128, 2008. doi: 10.1088/1742-5468/2008/04/P04008.

[9] P. Diaconis and S. Janson. Graph Limits and Exchangeable Random Graphs. *Rend. Mat. Appl. (7)*, 7(1):33–61, 2008.

[10] W. J. R. Eplett. Self–converse tournaments. *Canadian Mathematical Bulletin*, 22:23–27, 1979.

[11] P. Erdös and T. Gallai. Graphs with Prescribed Degrees of Vertices. *Mat. Lapok.*, 11:264–274, 1960.

[12] P. Galashin. Existence of a persistent hub in the convex preferential attachment model. *Probability and Mathematical Statistics*, 36(1):59–74, 2016.

[13] H. G. Landau. On dominance relations and the structure of animal societies: III The condition for a score structure. *Bulletin of Mathematical Biophysics*, 15(2):143–148, 1953.

[14] L. Lovász. *Large Networks and Graph Limits*. Colloquium Publications. American Mathematical Society, 2012.

[15] J. W. Moon. *Topics on Tournaments*. Holt, Rinehart and Winston, Inc., 1968.

UPPSALA DISSERTATIONS IN MATHEMATICS

Dissertations at the Department of Mathematics
Uppsala University

---

1. Torbjörn Lundh: *Kleinian groups and thin sets at the boundary*. 1995.
2. Jan Rudander: *On the first occurrence of a given pattern in a semi-Markov process*. 1996.
3. Alexander Shumakovitch: *Strangeness and invariants of finite degree*. 1996
4. Stefan Halvarsson: *Duality in convexity theory applied to growth problems in complex analysis*. 1996.
5. Stefan Svanberg: *Random walk in random environment and mixing.* 1997.
6. Jerk Matero: *Nonlinear elliptic problems with boundary blow-up*. 1997.
7. Jens Blanck: *Computability on topological spaces by effective domain representations*. 1997.
8. Jonas Avelin: *Differential calculus for multifunctions and nonsmooth functions*. 1997.
9. Hans Garmo: *Random railways and cycles in random regular graphs*. 1998.
10. Vladimir Tchernov: *Arnold-type invariants of curves and wave fronts on surfaces*. 1998.
11. Warwick Tucker: *The Lorenz attractor exists*. 1998.
12. Tobias Ekholm: *Immersions and their self intersections*. 1998.
13. Håkan Ljung: *Semi Markov chain Monte Carlo*. 1999.
14. Pontus Andersson: *Random tournaments and random circuits*. 1999.
15. Anders Andersson: *Lindström quantifiers and higher-order notions on finite structures*. 1999.
16. Marko Djordjević *Stability theory in finite variable logic*. 2000.
17. Andreas Strömbergsson: *Studies in the analytic and spectral theory of automorphic forms*. 2001.
18. Olof-Petter Östlund: *Invariants of knot diagrams and diagrammatic knot invariants*. 2001.
19. Stefan Israelsson: *Asymptotics of random matrices and matrix valued processes*. 2001.
20. Yacin Ameur: *Interpolation of Hilbert spaces*. 2001.
21. Björn Ivarsson: *Regularity and boundary behavior of solutions to complex Monge-Ampère equations*. 2002.
22. Lars Larsson-Cohn: *Gaussian structures and orthogonal polynomials*. 2002.
23. Sara Maad: *Critical point theory with applications to semilinear problems without compactness*. 2002.
24. Staffan Rodhe: *Matematikens utveckling i Sverige fram till 1731*. 2002.
25. Thomas Ernst: *A new method for q-calculus*. 2002.
26. Leo Larsson: *Carlson type inequalities and their applications*. 2003.

27. Tsehaye K. Araaya: *The symmetric Meixner-Pollaczek polynomials*. 2003.
28. Jörgen Olsén: *Stochastic modeling and simulation of the TCP protocol*. 2003.
29. Gustaf Strandell: *Linear and non-linear deformations of stochastic processes*. 2003.
30. Jonas Eliasson: *Ultrasheaves*. 2003.
31. Magnus Jacobsson: *Khovanov homology and link cobordisms*. 2003.
32. Peter Sunehag: *Interpolation of subcouples, new results and applications*. 2003.
33. Raimundas Gaigalas: *A non-Gaussian limit process with long-range dependence*. 2004.
34. Robert Parviainen: *Connectivity Properties of Archimedean and Laves Lattices*. 2004.
35. Qi Guo: *Minkowski Measure of Asymmetry and Minkowski Distance for Convex Bodies*. 2004.
36. Kibret Negussie Sigstam: *Optimization and Estimation of Solutions of Riccati Equations*. 2004.
37. Maciej Mroczkowski: *Projective Links and Their Invariants*. 2004.
38. Erik Ekström: *Selected Problems in Financial Mathematics*. 2004.
39. Fredrik Strömberg: *Computational Aspects of Maass Waveforms*. 2005.
40. Ingrid Lönnstedt: *Empirical Bayes Methods for DNA Microarray Data*. 2005.
41. Tomas Edlund: *Pluripolar sets and pluripolar hulls*. 2005.
42. Göran Hamrin: *Effective Domains and Admissible Domain Representations*. 2005.
43. Ola Weistrand: *Global Shape Description of Digital Objects*. 2005.
44. Kidane Asrat Ghebreamlak: *Analysis of Algorithms for Combinatorial Auctions and Related Problems*. 2005.
45. Jonatan Eriksson: *On the pricing equations of some path-dependent options*. 2006.
46. Björn Selander: *Arithmetic of three-point covers*. 2007.
47. Anders Pelander: *A Study of Smooth Functions and Differential Equations on Fractals*. 2007.
48. Anders Frisk: *On Stratified Algebras and Lie Superalgebras*. 2007.
49. Johan Prytz: *Speaking of Geometry*. 2007.
50. Fredrik Dahlgren: *Effective Distribution Theory*. 2007.
51. Helen Avelin: *Computations of automorphic functions on Fuchsian groups*. 2007.
52. Alice Lesser: *Optimal and Hereditarily Optimal Realizations of Metric Spaces*. 2007.
53. Johanna Pejlare: *On Axioms and Images in the History of Mathematics*. 2007.
54. Erik Melin: *Digital Geometry and Khalimsky Spaces*. 2008.
55. Bodil Svennblad: *On Estimating Topology and Divergence Times in Phylogenetics*. 2008.

56. Martin Herschend: *On the Clebsch-Gordan problem for quiver representations*. 2008.
57. Pierre Bäcklund: *Studies on boundary values of eigenfunctions on spaces of constant negative curvature*. 2008.
58. Kristi Kuljus: *Rank Estimation in Elliptical Models*. 2008.
59. Johan Kåhrström: *Tensor products on Category O and Kostant's problem*. 2008.
60. Johan G. *Granström: Reference and Computation in Intuitionistic Type Theory*. 2008.
61. Henrik Wanntorp: *Optimal Stopping and Model Robustness in Mathematical Finance*. 2008.
62. Erik Darpö: *Problems in the classification theory of non-associative simple algebras*. 2009.
63. Niclas Petersson: *The Maximum Displacement for Linear Probing Hashing*. 2009.
64. Kajsa Bråting: *Studies in the Conceptual Development of Mathematical Analysis*. 2009.
65. Hania Uscka-Wehlou: *Digital lines, Sturmian words, and continued fractions*. 2009.
66. Tomas Johnson: *Computer-aided computation of Abelian integrals and robust normal forms*. 2009.
67. Cecilia Holmgren: *Split Trees, Cuttings and Explosions*. 2010.
68. Anders Södergren: *Asymptotic Problems on Homogeneous Spaces*. 2010.
69. Henrik Renlund: *Recursive Methods in Urn Models and First-Passage Percolation*. 2011.
70. Mattias Enstedt: *Selected Topics in Partial Differential Equations*. 2011.
71. Anton Hedin: *Contributions to Pointfree Topology and Apartness Spaces*. 2011.
72. Johan Björklund: *Knots and Surfaces in Real Algebraic and Contact Geometry*. 2011.
73. Saeid Amiri: *On the Application of the Bootstrap*. 2011.
74. Olov Wilander: *On Constructive Sets and Partial Structures*. 2011.
75. Fredrik Jonsson: *Self-Normalized Sums and Directional Conclusions*. 2011.
76. Oswald Fogelklou: *Computer-Assisted Proofs and Other Methods for Problems Regarding Nonlinear Differential Equations*. 2012.
77. Georgios Dimitroglou Rizell: *Surgeries on Legendrian Submanifolds*. 2012.
78. Granovskiy, Boris: *Modeling Collective DecisionMaking in Animal Groups*. 2012
79. Bazarganzadeh, Mahmoudreza: *Free Boundary Problems of Obstacle Type, a Numerical and Theoretical Study*. 2012
80. Qi Ma: *Reinforcement in Biology*. 2012.
81. Isac Hedén: *Ga-actions on Complex Affine Threefolds*. 2013.
82. Daniel Strömbom: *Attraction Based Models of Collective Motion*. 2013.
83. Bing Lu: Calibration: *Optimality and Financial Mathematics*. 2013.

84. Jimmy Kungsman: *Resonance of Dirac Operators*. 2014.
85. Måns Thulin: *On Confidence Intervals and Two-Sided Hypothesis Testing*. 2014.
86. Maik Görgens: *Gaussian Bridges – Modeling and Inference*. 2014.
87. Marcus Olofsson: *Optimal Switching Problems and Related Equations*. 2015.
88. Seidon Alsaody: *A Categorical Study of Composition Algebras via Group Actions and Triality*. 2015.
89. Håkan Persson: *Studies of the Boundary Behaviour of Functions Related to Partial Differential Equations and Several Complex Variables*. 2015.
90. Djalal Mirmohades: *N-complexes and Categorification*. 2015.
91. Shyam Ranganathan: *Non-linear dynamic modelling for panel data in the social sciences*. 2015.
92. Cecilia Karlsson: *Orienting Moduli Spaces of Flow Trees for Symplectic Field Theory*. 2016.
93. Olow Sande: *Boundary Estimates for Solutions to Parabolic Equations*. 2016.
94. Jonathan Nilsson: *Simple Modules over Lie Algebras*. 2016.
95. Marta Leniec: *Information and Default Risk in Financial Valuation*. 2016.
96. Arianna Bottinelli: *Modelling collective movement and transport network formation in living systems*. 2016.
97. Katja Gabrysch: *On Directed Random Graphs and Greedy Walks on Point Processes*. 2016.
98. Martin Vannestål: *Optimal timing decisions in financial markets*. 2017.
99. Natalia Zabzina: *Mathematical modelling approach to collective decision-making*. 2017.
100. Hannah Dyrssen: *Valuation and Optimal Strategies in Markets Experiencing Shocks*. 2017.
101. Juozas Vaicenavicius: *Optimal Sequential Decisions in Hidden-State Models*. 2017.
102. Love Forsberg: *Semigroups, multisemigroups and representations*. 2017.
103. Anna Belova: *Computational dynamics – real and complex*. 2017.
104. Ove Ahlman: *Limit Laws, Homogenizable Structures and Their Connections*. 2018.
105. Erik Thörnblad: *Degrees in Random Graphs and Tournament Limits*. 2018.