LARGE GRAPHS AND THE RELATIONSHIP BETWEEN EDGE AND TRIANGLE DENSITY

Submitted to the Department of Mathematics in partial fulfillment of the requirements for the degree of Bachelor of Arts in Mathematics with Honors

Jack Latta

advised by

Sukhada Fadnavis

March 24, 2014
Harvard University

Acknowledments

First and foremost I would like to thank my advisor Sukhada Fadnavis for stoking my interest in combinatorics and guiding me toward the theory of large graphs. You were invaluable in helping me develop a clearer vision for this beautiful topic. I'd also like to thank the other members of the Harvard community, both professors and students, that have aided me in my growth the past four years. You are many, but I must in particular mention Benedict Gross, Howard Georgi, Joseph Ruescher, Tony Liu, and Maya Jonas Silver. Thank you to my mom for teaching me a love of learning, my dad for teaching me to have faith in myself, and my sister for teaching me to have faith in others. And, finally, thank you to my brother Mason C. Sloan the First- he knows why.

Contents

1 Introduction **3**

2 Algebra of Graph Homomorphisms **5**
2.1 Gluing Algebra 6
2.1.1 Connection Rank 6
2.1.2 Q_S/f and the Idempotent Basis 8
2.1.3 Finite Connection Rank and Computation 9
2.2 Graph Homomorphisms 10
2.2.1 Homomorphism Functions and Connection Rank . . . 11
2.3 Reflection Positivity 14
2.3.1 Kruskal-Katona Bound 15
2.3.2 Goodman's Bound 16

3 Graphons **17**
3.1 Definition of Graphons 18
3.2 Cut Norm 21
3.3 Szemerédi Regularity Lemma 22
3.3.1 Compactness of $(\widetilde{\mathscr{W}_0}, \delta_\square)$ 23
3.4 Convergence of Dense Graph Sequences 24

4 Extremal Graph Theory **27**
4.1 Inequalities and Complete Graphs 27
4.1.1 Generalized Turan's Theorem 29
4.2 Variational Calculus of Graphons 30
4.3 Razborov's Proof 32

Chapter 1

Introduction

In 1907, Mantel proved that all n node graphs with at least $n^2/4$ edges contain a triangle in what was probably the first result in extremal graph theory [10]. Define $t(F, G)$ as the probability that a random map of the nodes of F into G preserves the adjacencies of F. Then Mantel's statement says that if the edge density $t(K_2, G)$ is greater than one half, then the triangle density $t(K_3, G)$ is greater than zero. The natural question arising from this result is: Given that a graph has some edge density, what triangle densities can it have? Call the space of all possible values $(t(K_2, G), t(K_3, G))$ $D_{2,3}$. Finding $D_{2,3}$ proved difficult. In 1976, Lovász and Somonovitis conjectured the bounding curves [14], but the conjecture remained unproven until Razborov solved the problem in 2008 using his flag algebra technique [19]. In this thesis, instead of flag algebras, the modern theory of large dense graphs is developed, ultimately providing the relationship between edge density and triangle density.

However, the theory of large graphs is widely applicable outside of extremal graph theory, as large graphs are ubiquitous not only in mathematics but also in computer science, statistical physics, biology, engineering, and many other fields. Large graphs present unique difficulties. Due to their size, it is generally hard to compute graph parameters on them, as most parameters, at least naively, take time exponential in the size of the graph to compute. Graphs also can be sufficiently large that even looking at all of the nodes and edges takes too much time. In this case, it is important to know what information can be gained from looking at subgraphs of the graph. In order for the sampling of subgraphs to prove useful, there must be some notion

of distance between graphs that describes how different two graphs are with respect to these subgraphs, although it is not clear how this distance can be usefully defined. Finally, just as a metal, actually a large graph of atoms and their interactions, can be approximated as a smooth solid with certain properties, it is convenient to define a notion of the convergence of sequences of graphs so that large graphs can be approximated by integrable functions.

Thus, on the way to Razborov's result these questions associated with large graphs are answered. In Chapter 2 it is shown that a large class of graph parameters define a factor algebra that is a finite dimensional inner product space, culminating in the proof that such parameters can be evaluated on graphs with finite tree-width in polynomial time. Further, with one other requirement, these parameters are equivalent to homomorphism functions, a result that allows for the proof of Goodman's Bound and the Kruskal-Katona Bound, partial progress toward the categorization of $D_{2,3}$. Chapter 3 concerns the theory of convergent graph sequences. Convergence is defined as convergence in homomorphism density for all graphs F. Graphons are proposed as a limit, and it is shown that they have the desired behavior. The cut distance provides a notion of distance between graphs or graphons, and the Szemerédi Regularity Lemma is used to prove that the space defined by the cut distance is compact. The Counting Lemma is then used to show that graphons are, in fact, the appropriate limit object. At this point, the theory of large graphs is well developed, and one could easily branch into the study of spin models of statistical physics, further algorithmic study of large graphs, or the topology of large graphs and graphons, among others, but that is outside of the scope of this thesis. Instead, in Chapter 4 the continuity of graphons is used first to prove the Generalized Turan's Theorem, after which the variational calculus of graphons is developed. With this final tool, Razborov's proof translated into the language of graphons is presented.

Chapter 2

Algebra of Graph Homomorphisms

The primary difficulty of understanding how parameters act on large graphs is the size of the graphs. One avenue for simplifying large graphs is to define a product between graphs to create algebraic relationships, with the idea being that evaluation of a parameter on a graph can be broken into subproblems. As in [11], we define k- labeled graphs with a corresponding operation called the gluing product, which defines a space of graphs Q_k, and, for some graph parameter f, a connection matrix $M(f, k)$. It turns out that a surprising number of connection matrices $M(f, k)$ have finite rank, which allows the definition of an inner product space Q_k/f with finite dimension. In other words, with respect to such a parameter f, any graph can be written as a linear combination of a finite number of basis elements in Q_k/f. This has profound implications, two of which are explored here. First, for parameters with finite connection rank and graphs with bounded tree-width, one can use dynamic programming to compute the parameter in time polynomial in the size of the graph. Second, parameters with finite connection rank are intimately connected with homomorphism functions, which will allow for the proof of Goodman's Bound and the Kruskal-Katona Bound, partial progress toward the characterization of $D_{2,3}$.

2.1 Gluing Algebra

A *S-labeled graph* is a graph with an identification of the labels $x \in S$ with distinct nodes of the graph. If $S = [k]$, the graph is called k-labeled, and if no node has multiple labels, the graph is called simply labeled. The *gluing product* of two k-labeled graphs is computed by taking the disjoint union of the two graphs, and then identifying the nodes with the same label, possibly creating multiple edges between two nodes. Notice that the product is associative and commutative and always yields a k-labeled graph. This product can be extended in the natural way to linear combinations of k-labeled graphs to define the quantum algebra Q_k. Any graph parameter $f : Q_0 \to \mathbb{R}$ defines a Frobenius inner product $\langle , \rangle$ on $x, y \in Q_k : \langle x, y \rangle = f([[xy]])$ where $[[,]]$ is the unlabeling operator, projecting quantum graphs into Q_0.

2.1.1 Connection Rank

Consider any graph parameter f and integer $k \geq 0$. The *connection matrix* $M(f, k)$ is the symmetric matrix whose rows and columns are indexed by the simply k-labeled graphs G_i and the element in the ith row and jth column is $\langle G_1, G_2 \rangle$. A graph parameter is *reflection positive* if all of its connection matrices are positive semidefinite and define its *connection rank* $r(f, k)$ as $\mathrm{rk}(M(f, k))$.

Many familiar graph parameters have finite connection rank. Below are three examples to give intuition for why this is often the case and which parameters one might expect to (and not to) have finite connection rank. We frequently use the fact that $r(f) + r(g) \geq r(f + g)$:

Example 2.1.1: Let $\chi_q(G)$ denote the number of q-colorings of a graph G and $\chi_q(G, f)$ for $f : [k] \to [q]$ denote the number of q-colorings of G that assign each labeled node i the color $f(i)$. Then for k-labeled graphs G_1 and G_2, the q-colorings of $[[G_1 G_2]]$ can be split into cases depending on the function f, giving that

$$\chi_q([[G_1 G_2]]) = \sum_f \chi_q(G_1, f)\chi_q(G_2, f)$$

Thus, $M(\chi_q, k)$ can be written as the sum of q^k rank 1 matrices and so $r(\chi_q, k) \leq q^k$.

Example 2.1.2: Let $Ham(G)$ denote the number of Hamiltonian cycles of G. Consider the form of a Hamiltonian cycle on the k-labeled G_1G_2. Following the cycle will define a cyclic ordering of the labeled nodes $(i_1, \ldots, i_k)$. Further, the path from i_l to i_{l+1} either uses edges in G_1 or G_2. If the edges used are in G_m, let $j_l = m$. Then each Hamiltonian path can be associated with a trace $T = (i_1, \ldots, i_k, j_1, \ldots, j_k)$. Define $Ham(G_m, T)$ as the number of sets of paths in G_m such that every path is from some i_l to i_{l+1} with $j_l = m$, does not cover any of the other labeled nodes, are node-disjoint on the unlabeled edges, and cover every unlabeled node of G_m. Intuitively, $Ham(G_m, T)$ is the number of halves of a Hamiltonian path that satisfy some trace T. Then

$$Ham(G_1G_2) = \sum_T Ham(G_1, T)Ham(G_2, T)$$

and so by the same logic as the above example, $r(Ham, k)$ is bounded by the number of possible traces, $2^k(k-1)!$ [12].

It's worth noting quickly that the above argument does not work for Eulerian paths, as the j_l above would need to be replaced by something like in-degree, which is only bounded by the edge multiplicity of the labeled nodes, which is unbounded. In fact, the parameter counting the number of Eulerian paths in a graph has infinite connection rank.

Example 2.1.3: The chromatic number χ does not have finite connection rank. Consider the submatrix N of $M(\chi, 0)$ indexed by graphs K_i. Then it is clear that $N_{i,j} = max(i, j)$ and so N has infinite rank. Thus, χ has infinite connection rank.

It's clear that this argument can be extended to any *maxing* parameter, ie one in which $f(G_1G_2) = max(f(G_1), f(G_2))$. In particular, the size of the maximum clique also has infinite connection rank.

Many 0-1 valued parameters, called *properties* also have finite connection rank. For example, Robertson and Seymour [21] showed that all properties such that $f(G) = 1$ implies $f(G') = 1$ for all minors of G G' have finite connection rank, and Godlin, Kotex, and Makowski [7] showed that all properties that can described by a monadic second order formula do as well.

The above examples of q-colorings and Hamiltonian paths also suggest that parameters with finite connection rank can be computed more efficiently than those with infinite connection rank through something like dynamic programming. This notion can be formalized, but first we need to know a bit more about the structure of Q_k.

2.1.2 Q_S/f and the Idempotent Basis

Let $N_k(f)$ denote the kernel of the inner product associated with f. In other words, let

$$N_k(f) = x \in Q_k : f(xy) = 0 \ \forall y \in Q_k$$

Let $Q_k/f := Q_k/N_k(f)$.

Theorem 2.1.1: The dimension of Q_k/f is $r(f,k)$.

This follows immediately from the definition of $r(f,k)$. □

Further, if f is reflection positive, then the inner product is positive definite on Q_k, implying that Q_k/f is an inner product space. Consider the linear transformation A_x associated with $x \in Q_k$ such that $A_x y = xy$. Since the inner product is definite, $x \to A_x$ is injective. Since the A_x commute and satisfy the Frobenius identity, they can be simultaneously diagonalized. Thus, Q_k/f is isomorphic to the algebra of diagonal matrices, and so it has an *idempotent basis* B_k. This logic can be extended to any set $S \subset \mathbb{N}$ to define an idempotent basis B_S. For two idempotents in Q/f, p resolves q if $pq = q$. If $p \in B_S$ and $|T| = |S| + 1$, the number of elements in B_T that resolve p is called the *degree* of p and is denoted $deg(p)$.

This algebraic structure has far reaching ramifications. We present without proof two results that will be useful later, as well as one immediate corollary. The omitted proofs consist only of various algebraic manipulations, see [12] for details:

Lemma 2.1.1: Let $S, T \subset \mathbb{N}$ be finite sets, $p \in B_{S \cap T}$, and let $q \in B_S$ resolve p. Then for $x \in Q_T/f$, $f(p)f(qx) = f(q)f(px)$.

Lemma 2.1.2: If two idempotents $q \in B_T$ and $r \in B_S$ resolve the same idempotent $p \in B_{S \cap T}$, then $qr \neq 0$.

Proof: By the above lemma,

$$f(qr) = \frac{f(q)}{f(p)} f(pr) = \frac{f(q)}{f(p)} f(r) > 0$$

and so $qr \neq 0$.

Lemma 2.1.3: If $S \subset T$ and $q \in B_T$ resolves $p \in B_S$, then $\deg(q) \geq \deg(p)$.

2.1.3 Finite Connection Rank and Computation

As suggested by Examples 2.1.1 and 2.1.2, parameters with finite connection rank give rise to a decomposition of the calculation of the parameter into subproblems. We show here that this allows such parameters to be calculated efficiently on graphs with small treewidth [11].

A *tree decomposition* of a graph G is a tree T and a set of subgraphs $\{G_i\}_{i \in V(T)}$ such that $\cup G_i = G$ and, if i is on the path from j to k in T, $V(G_i) \supseteq V(G_j) \cap V(G_k)$. A graph G has *treewidth* k if k is the largest integer such that there is no tree decomposition of G into subgraphs of size at most k.

Theorem 2.1.2: [12] If f is a graph parameter with $r(f,k)$ finite, then f can be computed in polynomial time for graphs with treewidth at most k .

Proof: Essentially, we do a large amount of precomputation, and then use dynamic programming.

It is easy to confirm that $r(f,k)$ is non-decreasing in k, so since $r(f,k)$ is finite, so is $r(f,l)$ for all $l \leq k$. By Theorem 2.1.1, then, each of the algebras Q_l/f have finite dimension. For each l, compute a basis B_l, as well as the product of any two elements in the basis.

Next, for every m-labeled graph H for $m \leq k$ with at most $k+1$ nodes, every ordered subset $S \subseteq V(H)$ with $|S| \leq k$, and every $F_S \in B_{|S|}$, let H' be the m-labeled graph formed from gluing the labels of F_S to the corresponding element of S on the graph H. Rewrite H' as a linear combination of the elements B_m.

Finally, compute $f(G)$ for all $G \in B_0$. Thus ends our precomputation, which depends only on k and $r(f,k)$.

Let G be some graph with treewidth k and decomposition using the notation above. Call any leaf r of T the root, and for $i \neq r$, let i' denote its parent. Then for all $i \neq r$, $S_i = V(G_i) \cap V(G_{i'})$ is a cutset in G with $k_i \leq k$ nodes. Define F_i as the union of all graphs G_j where j is a descendant of i (including i) after which the k_i nodes of S_i are labeled. Our goal is to express every F_i in the basis B_{k_i}.

Assume we have done so for all proper descendants of i. By definition, F_i is obtained from G_i by attaching different branches F_j at the sets S_j. By assumption, we know how to express each F_j as a linear combination of elements B_{k_j}, and so F_i is a linear combination of graphs, each of which

is G_i with some number of basis graphs attached at some $S \subseteq V(G_i)$ with $|S| \leq k$. If multiple bases are attached at the same S, they can be easily replaced by a single graph, as we have precomputed all of the products of bases. This yields a linear combination of k_i labeled graphs in the basis B_{k_i}, as desired. Finally, when we reach the root, consider it 0-labeled, giving an expression for G in the basis B_0, which immediately gives the value of $f(G)$ as desired. □

2.2 Graph Homomorphisms

Consider two simple graphs G and H. A *homomorphism* ψ from G to H is a map from $V(G)$ to $V(H)$ such that if $(i,j) \in E(G)$, $(\psi(i), \psi(j)) \in E(H)$. Let $hom(F,G)$ denote the number of homomorphisms from F to G. A *weighted graph* is a looped simple graph with positive real weights α_i associated with the nodes and real weights β_{ij} associated with the edges. The notion of homomorphisms can be extended to weighted graphs H with nodeweights α_i and edgeweights β_{ij} as follows:

$$\alpha_\psi = \prod_{u \in V(F)} a_{\psi(u)}$$

$$hom_\psi(F,H) = \prod_{(u,v) \in E(F)} \beta_{\psi(u)\psi(v)}$$

$$hom(F,H) = \sum_{\psi: V(F) \to V(H)} \alpha_\psi hom_\psi(F,H)$$

We define the homomorphism densities as

$$t(F,G) = \frac{hom(F,G)}{v(G)^{v(F)}}$$

and

$$t(F,H) = \frac{hom(F,H)}{(\sum_{v \in V(H)} \alpha_v)^{v(F)}}$$

for G simple and H weighted, respectively. $inj(F,G)$ and t_{inj} can be defined analogously where ψ is restricted to being injective. It's clear that for simple graphs F, G

$$hom(F,G) = \sum_P inj(F/P, G)$$

where P ranges over all partitions of $V(F)$ and G/P is the graph that results from merging each element of the partition P into a single node, with the new nodes connected only if they have adjacent pre-images. It follows immediately that we can use the Möbius inversion of the partition lattice to get [24]:

$$inj(F,G) = \sum_P \mu_P hom(F/P, G)$$

We can now prove the following:

Theorem 2.2.1: The simple graph parameter $hom(.,G)$ determines a simple graph G.

Proof: By the above relation between $hom(F,G)$ and $inj(F,G)$, if for all F $hom(F,G) = hom(F,G')$, then $inj(F,G) = inj(F,G')$ for all F. But then $inj(G',G) = inj(G,G') = inj(G,G) > 0$, and so there is an injective map from G into G' and from G' into G, and so $G \cong G'$. □

We will not prove it here, but an analogous result to the above theorem holds for weighted graphs [11]. If weighted graphs H_1 and H_2 have no twin nodes and $hom(F,H_1) = hom(F,H_2)$ for all simple graphs F, then $H_1 \cong H_2$. A weighted graph H has twin nodes if there exist some two nodes u and v such that $\beta_{uj} = \beta_{vj}$ for all j. If such twins exist, H can be reduced to a graph with one less node by removing u and changing v's weight to $\alpha_u + \alpha_v$ without affecting any of its values with respect to homomorphism functions. Thus, the above result could equivalently state that if two weighted graphs have identical homomorphism functions, then after twin reduction they are isomorphic. In the next section, we will complete this identification of graphs with parameters, but first, it's worthwhile to see some examples of homomorphism functions.

Example 2.2.1: $hom(G,K_n)$ is the number of n colorings of G.

Example 2.2.2: $hom(P_k,G)$ is the number of walks of length $k-1$ in G.

Example 2.2.3: The number of stable sets $stab(G)$ is $hom(G,K_2^\circ)$ where K_2° is an edge with a loop added at one of the nodes. This is because for each homomorphism ψ, the set of nodes that ψ sends to the unlooped node in K_2 form an independent set in G.

2.2.1 Homomorphism Functions and Connection Rank

We are now ready to unify our notions of homomorphism functions and finite connection rank.

Theorem 2.2.2: For f a graph parameter defined on multigraphs without loops, $f = hom(., H)$ for some weighted graph H on q nodes if and only if it is reflection positive, $f(K_0) = 1$, and $r(f, k) \leq q^k$ for all $k \geq 0$.

Proof: ($\Rightarrow$) [2] For any two k-labeled graphs F_1 and F_2 and $\psi : [k] \to V(H)$, let $hom^{\psi}(G, H) : Q_k \to \mathbb{R}$ be the number of homomorphisms from G to H that send i to $\psi(i)$ for $i \in [k]$. Then by definition

$$hom^{\psi}([[F_1F_2]], H) = hom^{\psi}(F_1, H)hom^{\psi}(F_2, H)$$

Since

$$hom([[F_1F_2]], H) = \sum_{\psi} \alpha_{\psi} hom^{\psi}([[F_1F_2]], H)$$

$M(hom(., H), k)$ is the sum of positive semidefinte rank 1 matrices. Since there are $v(H)^k$ possible choices for ψ, $hom(., H)$ is reflection positive and has rank at most $v(H)^k$. It is easy to confirm that $hom(K_0, H) = 1$ for all H.

($\Leftarrow$) [13] By Lemma 2.1.3, if some idempotent $p \in B_S$ is resolved by some $q \in B_T$ with $|T| = |S| + 1$, then $\deg(q) \geq \deg(p)$. This process can be repeated indefinitely to show that for arbitrary $|T| > |S|$, the dimension of Q_T is at least $\deg(p)^{|T-S|}$. Since by assumption the dimension is bound by q^k, then, $\deg(p) \leq q$ and, in particular, there is some largest degree $D \leq q$ of some idempotent such that all idempotents that resolve it also have degree D.

For some set S fix an idempotent basis $p \in B_S$ with degree D. Let $q_1^u, \ldots, q_D^u$ be the elements of $B_{S \cup \{u\}}$ for any $u \in \mathbb{N} - S$. Then for any finite $T \supset S$ and $\psi : T - S \to \{1, \ldots, D\}$, let

$$q_{\psi} = \prod_{v \in T-S} q^v_{\psi(v)}$$

It follows from Lemma 2.1.1 that

$$f(q_{\psi}) = f\left(\prod_{v \in T-S} q^v_{\psi(v)}\right) = \left(\prod_{v \in T-S} \frac{f(q_{\psi(v)})}{f(p)}\right) f(p) \neq 0$$

and so $q_{\psi} \neq 0$.

We claim that the basic idempotents of Q_T/f that resolve p are the elements q_{ψ}. We show this by induction. If $|T - S| = 1$, the result follows trivially. Now assume that $|T - S| > 1$. For some $u \in T - S$, let $U = S \cup \{u\}$ and

$W = T - \{u\}$. By the inductive assumption, the idempotent bases resolving p are of the form q_ψ. Call one of these r. By Lemma 2.1.2, $rq_i^u \neq 0$ for all i. Further, both it and the idempotent bases of B_T clearly resolve r. Since $rq_i^u \neq rq_j^u$ for $i \neq j$, and the degree of r is D, it follows that each rq_i^u is an idempotent basis in B_T. We thus have all of the elements in B_T that resolve p, and so Q_T/f behaves as we claimed.

Note that an idempotent q_ψ resolves q_i^u if and only if $\psi(u) = i$. It follows that

$$q_i^u = \sum_{\psi:\psi(u)=i} q_\psi$$

We are ready to define H, a weighted looped complete graph on D nodes with nodeweights α_i and edgeweights β_{ij}. Pick some $u \in \mathbb{N} - S$. Let $\alpha_i = f(q_i^u)/f(p)$ (which is trivially greater than zero).

Choose another distinct $v \in \mathbb{N} - S$ and let $W = S \cup \{u, v\}$. Let K_{uv} denote the graph on W which was just one edge connecting u and v, and let k_{uv} be the corresponding graph in Q_W. pk_{uv} trivially resolves p, and so it can be written as a linear combination of elements of B_W that resolve p. We define β_{ij} as the elements that arise from this linear combination, namely

$$pk_{uv} = \sum_{i,j} \beta_{ij} q_i^u q_j^v$$

It's worth noting that $\beta_{ij} = \beta_{ji}$, as is necessary, because $pk_{uv} = pk_{vu}$.

Using the above expression of q_i^u as a sum of q_ψs, we get the following:

$$\begin{aligned} pk_{uv} &= \sum_{i,j \in V(H)} \beta_{ij} q_i^u q_j^v \\ &= \sum_{i,j \in V(H)} \beta_{ij} \sum_{\psi:\psi(u)=i,\psi(v)=j} q_\psi \\ &= \sum_{\psi} \beta_{\psi(u)\psi(v)} q_\psi \end{aligned}$$

Let G be a Z-labeled graph with $V(G) = Z \subseteq \mathbb{N} - S$ and let g be the

corresponding element of Q/f. Then

$$\begin{aligned} pg &= \prod_{uv \in E(G)} pk_{uv} \\ &= \prod_{uv \in E(G)} \left(\sum_{\psi} \beta_{\psi(u)\psi(v)} q_{\psi} \right) \\ &= \sum_{\psi} \left(\prod_{uv \in E(G)} \beta_{\psi(u)\psi(v)} q_{\psi} \right) \end{aligned}$$

Since $p \in Q_S/f$ and $g \in Q_Z/f$ with $S \cap Z = \emptyset$, $f(p)f(q) = f(pq)$ and so (in the third step using Lemma 2.1.2)

$$\begin{aligned} f(g) &= f(pg)/f(p) \\ &= \sum_{\psi} \left(\prod_{uv \in E(G)} \beta_{\psi(u)\psi(v)} f(q_{\psi}) \right) / f(p) \\ &= \sum_{\psi} \left(\prod_{uv \in E(G)} \beta_{\psi(u)\psi(v)} \left(\prod_{v \in V(G)} \frac{f(q_i^v)}{f(p)} \right) f(p) \right) / f(p) \\ &= \sum_{\psi} \left(\prod_{uv \in E(G)} \beta_{\psi(u)\psi(v)} \prod_{v \in V(G)} \alpha_{\phi(v)} \right) \\ &= hom(G, H) \end{aligned}$$

as desired. □

2.3 Reflection Positivity

We are now ready to do some extremal graph theory. All arguments about $t(F, H)$ will hold for all H, so for notational simplicity denote $t(F, H)$ by just F. Further, denote the number of labels of a graph using superscripted dots. Unless otherwise stated, the labels are applied to graphs from nodes with smallest degree first. When this is not the case, the labeling will be described explicitly.

The reflection positivity of homomorphism numbers gives many free inequalities. For example, for any submatrix M indexed by the graphs $\{F_1, \ldots, F_n\}$

and vector $v \in \mathbb{R}^n$, reflection positivity implies that $v^T M v \geq 0$. Written in terms of the F_i,

$$\begin{aligned} 0 &\leq \sum_{i,j=1}^{n} v_i v_j [[F_i F_j]] \\ &\leq \left[\left[\left(\sum_{i=1}^{n} v_i F_i\right)\right]\right] \end{aligned}$$

Thus, $[[x^2]] \geq 0$ for all $x \in Q_k$.

Another property implied by reflection positivity is that all submatrices of $M(hom(., H), k)$ have non-negative determinant. In particular, for a submatrix indexed by two elements F_1 and F_2,

$$[[F_1^2]][[F_2^2]] \geq [[F_1 F_2]]^2$$

Finally, if F is a subgraph of G, then any homomorphism of G into some H gives a corresponding homomorphism of F into H and so $F - G \geq 0$

2.3.1 Kruskal-Katona Bound

Given some edge density, the Kruskal-Katona Bound gives an upper bound on triangle density.

Theorem 2.3.1: $t(K_3, H) \leq t(K_2, H)^{3/2}$

Proof: [9] Let $F_1 = K_2^{\cdot\cdot}$ and $F_2 = P_3^{\cdot\cdot}$. Then since

$$\begin{aligned} [[F_1^2]][[F_2^2]] &\geq [[F_1 F_2]]^2 \\ \Rightarrow P_4 K_2 &\geq K_3^2 \\ \Rightarrow K_2^3 &\geq K_3^2 \\ \Rightarrow t(K_3, H) &\leq t(K_2, H)^{3/2} \end{aligned}$$

Where the second to last step follows from the fact that K_2^2 is a subgraph of P_4. □

Can this bound be attained? It turns out that one can get arbitrarily close to any point on the curve. Consider a graph G that has subgraph K_k with $n - k$ accompanying isolated nodes. Then $t(K_2, G) = \frac{k(k-1)}{n^2}$ and $t(K_3, G) = \frac{k(k-1)(k-2)}{n^3}$. As n approaches infinity, $t(K_2, G)$ approaches $(\frac{k}{n})^2$

and $t(K_2, G)$ approaches $(\frac{k}{n})^3$. Thus, one can write a sequence of graphs whose homomorphism densities converge to any point on the Kruskal-Katona line. One benefit of defining the limit of a convergent sequence of graphs is that the Kruskal-Katona bound will be saturated in this way for all choices of $t(K_2, G)$.

2.3.2 Goodman's Bound

Goodman's Bound gives a similar lower bound on triangle density.

Theorem 2.3.2: [8] $t(K_3, G) \geq t(K_2, G)(2t(K_2, G) - 1)$ for all G.

Proof: Notice that $x = K_3^{\cdot\cdot\cdot} - P_3^{\cdot\cdot\cdot} - P_3^{\cdot\cdot,\cdot} + K_2^{\cdot\cdot}K_1^{\cdot}$ is idempotent, where the two P_3s correspond to paths of length two with label 3 and 2 on the middle node, respectively. This means that

$$[[x^2]] = K_3 - 2P_3 + K_2K_1 = K_3 - 2P_3 + K_2$$

since $t(K_1, H) = 1$ for all H. Similarly,

$$[[(K_2^{\cdot} - K_2)^2]] = P_3 - K_2^2$$

Since unlabeled squares of quantum graphs are always positive, we thus have that

$$\begin{aligned} 0 &\leq [[x^2]] + 2[[(K_2^{\cdot} - K_2)^2]] \\ &\leq K_3 - 2K_2^2 + K_2 \\ \Rightarrow K_3 &\geq K_2(2K_2 - 1) \end{aligned}$$

as desired. □

It's clear that this bound is not strict for $t(K_2, G) < .5$, as then

$$t(K_2, G)(2t(K_2, G) - 1) < 0 \leq t(K_3, G).$$

For any graph K_n, $t(K_2, K_n) = \frac{n-1}{n}$ and $t(K_3, K_n) = \frac{(n-1)(n-2)}{n^2}$ so K_n saturates the inequality. It turns out that the only time the inequality is saturated is at $t(K_2, G) = \frac{n-1}{n}$ for integer n, but we don't yet have the tools to prove this. In order to optimize our lower bound further, we need to move away from the discrete space of graphs into a space in which we can make continuous deformations. This space will turn out to be the space of graphons, which is described in Chapter 3.

Chapter 3

Graphons

The example of the Kruskal-Katona Bound strongly suggests defining a limit object of a sequence of convergent graphs, the idea being that defining such object will allow us to move from the discrete space of graphs into a continuous one, which will open up many new lines of attack mathematically. First, one must define which sequences of graphs are convergent. We have already seen above that homomorphism functions $hom(., G)$ not only describe many important properties of graphs, but also fully determine simple graphs. This suggests a notion of convergence first defined by Borgs, Chayes, Lovász, Sós, and Vesztergombi in 2006[2]. A sequence of graphs (G_n) is called *convergent* if $t(F, G_n)$ converges for all F.

Next, we need to define a limit object. The definition is motivated by considering the adjacency matrix of a graph G, and then normalizing it, showing that graphs can be represented by symmetric step functions $[0, 1]^2 \to [0, 1]$. Graphons are generalizations of graphs, and defined as the set of all symmetric measurable functions $[0, 1]^2 \to [0, 1]$. In section 1, we go on to prove that graphons can be given appropriate homomorphism functions, and that all graphs have a corresponding graphon with identical homomorphism function.

In section 2, a distance on the space of graphons called the cut distance, first suggested by Frieze and Kannan [6], is motivated and presented. Section 3 introduces the Szemerédi Regularity Lemma, which, when applied to graphons, says that graphons can be approximated up to some error ϵ by a stepfunction with $k(\epsilon)$ steps. In particular, $k(\epsilon)$ does not depend on the specific graphon. These two notions combine to allow for the proof that the

space of graphons such that graphons with cut distance 0 are identified with each other is compact with respect to the cut distance.

With the result that the cut distance defines a compact space, in section 4 we prove that each graphon is the limit of some graph sequence and that each graph sequence has some graphon as its limit. Thus, it is confirmed that graphons are the appropriate limit object.

3.1 Definition of Graphons

A *kernel* is a symmetric, measurable function $W : [0,1]^2 \to \mathbb{R}$. Denote the space of all kernels $\mathscr{W}$. If the range of W is $[0,1]$, it is called a *graphon* [15], and the space of graphons is called $\mathscr{W}_0$. We also let $\mathscr{W}_1$ denote the space of graphons with range $[-1,1]$. For any graph H, the corresponding kernel is defined as follows: if the weights of the edges of H are β_{ij} and the weights of the nodes of H are α_i and $\alpha_H = \sum \alpha_i$, then W_H is a stepfunction partitioned into steps S_i of size α_i/α_H such that $W(x,y) = \beta_{ij}$ for $x \in S_i, y \in S_j$. If all $0 \leq \beta_{ij} \leq 1$, W_H is a graphon. Intuitively, the graphon representation of a graph looks very similar to an adjacency matrix- both can be thought of as look up tables giving the adjacency between nodes of a graph.

In order to have a useful definition, it must be the case that homomorphism densities of graphs extend to kernels. Define

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

It's clear that this is an analogous definition to that of homomorphism density of graphs, and so $t(F,G) = t(F,W_G)$ as we would desire. Crucially, it is also the case that $t(.,W)$ is multiplicative and reflection positive. This means, for example, that Goodman's Bound and the Kruskal-Katona Bound extend to graphons. It is not the case in general, however, that $t(.,W)$ has finite connection rank.

Example 3.1.1: let S_n denote the nth star, namely a tree with a root and

n leaves. Then

$$\begin{aligned} t(S_n, W) &= \int_{[0,1]^{V(S_n)}} \prod_{ij \in E(S_n)} W(x_i, x_j) \prod_{i \in V(S_n)} dx_i \\ &= \int_{[0,1]^{n+1}} \prod_{i=1}^{n} W(x, y_i) \prod_{i=1}^{n} dy_i dx \\ &= \int_0^1 \left(\int_0^1 W(x, y) dy \right)^n dx \\ &= \int_0^1 (d_W(x))^n dx \end{aligned}$$

where $d_W(x)$ is called the normalized degree function and defined as

$$d_W(x) := \int_0^1 W(x, y) dy$$

Example 3.1.2: [12] We already noted that the number of Eulerian paths in a graph F has infinite connection rank, and so cannot be written as a homomorphism function $t(., H)$ for some weighted graph H. Let $Eul(F)$ be the set of orientations $\vec{F}$ of F with equal in degree and out degree on each vertex. We claim that $|Eul(F)| = t(F, W)$ for $W(x, y) = 2cos(2\pi(x - y))$. Below we let $d_{in}^{\vec{F}}(i)$ and $d_{out}^{\vec{F}}(i)$ denote the in and out degrees of some vertex

i under the orientation $\vec{F}$:

$$\begin{aligned}
t(F,W) &= \int_{[0,1]^{V(F)}} \prod_{ij\in E(F)} 2cos(2\pi(x-y)) \prod_{i\in V(F)} dx_i \\
&= \int_{[0,1]^{V(F)}} \prod_{ij\in E(F)} e^{2\pi i(x_i-x_j)} + e^{2\pi i(x_j-x_i)} \prod_{i\in V(F)} dx_i \\
&= \int_{[0,1]^{V(F)}} \left(\sum_{\vec{F}} \prod_{ij\in E(\vec{F})} e^{2\pi i(x_i-x_j)} \right) \prod_{i\in V(F)} dx_i \\
&= \int_{[0,1]^{V(F)}} \left(\sum_{\vec{F}} \prod_{i\in V(\vec{F})} e^{2\pi i(d^{\vec{F}}_{in}-d^{\vec{F}}_{out})x_i} \right) \prod_{i\in V(F)} dx_i \\
&= \int_{[0,1]^{V(F)}} \left(\sum_{\vec{F}\in Eul(F)} 1 + \sum_{\vec{F}\notin Eul(F)} \prod_{i\in V(\vec{F})} e^{2\pi i(d^{\vec{F}}_{in}-d^{\vec{F}}_{out})x_i} \right) \prod_{i\in V(F)} dx_i \\
&= \int_{[0,1]^{V(F)}} \left(\sum_{\vec{F}\in Eul(F)} 1 \right) \prod_{i\in V(F)} dx_i \\
&= |Eul(F)|
\end{aligned}$$

where all cases with $\vec{F} \notin Eul(F)$ cancel because the period of each $e^{2\pi ik}$ for $k \neq 0$ is $1/k$, and so integrating from 0 to 1 cancels out the element.

Thus, the class $t(.,W)$ retains the properties of $t(.,H)$ while expanding the possible uses of homomorphism functions. In many cases, the function $t(.,W)$ is all we will care about with respect to some kernel. For this reason we say that two kernels W_1 and W_2 are *weakly isomorphic* if $t(F,W_1) = t(F,W_2)$ for all simple graphs F and call the space of equivalence classes of weakly isomorphic kernels $\widetilde{\mathscr{W}}$.

As an extension, for a k-labeled graph G let $V = V(G) - [k]$. Then define

$$t_{x_1,\ldots,x_k}(F,W) = \int_{x\in[0,1]^V} \prod_{ij\in E(V)} W(x_i,x_j) \prod_{i\in V} dx_i$$

It's clear that $t_{\mathbf{x}}(F_1F_2,W) = t_{\mathbf{x}}(F_1,W)t_{\mathbf{x}}(F_2,W)$ and, for F' the graph obtained by unlabeling the k node,

$$t_{x_1,\ldots,x_{k-1}}(F',W) = \int_0^1 t_{x_1,\ldots,x_k}(F,W)dx_k$$

3.2 Cut Norm

A fundamental question one might ask about graphs is how different they are from each other. First, consider two graphs F, G with the same number of nodes n. One natural choice is the edit distance

$$d_1(F,G) = \frac{|E(F)\Delta E(G)|}{n^2}$$

which is the distance associated with the L_1 norm. This doesn't really work well with our definition of convergence. For example, consider two graphs A, B with n nodes that are generated as follows: for any two nodes i and j, put an edge between them with probability $1/2$. Then for large n, the two graphs will have nearly identical homomorphism densities, and our notion of distance should say that these two graphs are similar. However, the edit distance between these two graphs has expected value $1/2$, the same expected value as the edit distance between one of these graphs and the complete graph.

The cut distance addresses this problem by instead trying to isolate the portions of the graphs that are most different. Let $e_G(S,T)$ with $S, T \in [n]$ denote the edges in G with first node in S and second node in T. Then we define

$$d_\square(F,G) = max_{S,T} \frac{|e_F(S,T) - e_G(S,T)|}{n^2}$$

In this case, $d_\square(A,B)$ is $O(1/\sqrt{n})$, which seems more reasonable.

However, there are still two more issues. First, labeling shouldn't matter, as this has no bearing on homomorphism densities. Second, we would like to be able to compare graphs with different numbers of nodes. To address the first problem, we will just take the minimum value of our distance across all relabelings. We let $\hat{G}$ denote relabelings of a graph G. To address the second, we introduce the concept of the blowup of a graph. The blowup $G(k)$ is formed from G by creating k copies of each node in G and then connecting nodes in $G(p)$ if and only if their corresponding node was connected in G. Then for F and G with m and n nodes, we define the *cut distance*

$$\delta_\square(F,G) = \min_{\hat{F},\hat{G}} \lim_{k\to\infty} d_\square(\hat{F}(kn), \hat{G}(km))$$

Working with kernels, we analogously define the cut norm [6]

$$||W||_\square = \sup_{S,T\subseteq[0,1]} \left| \int_{S\times T} W(x,y)dxdy \right|$$

and, letting ψ be invertible measure preserving maps from $[0,1]$ to $[0,1]$ and $W^{\psi}(x,y) := W(\psi(x),\psi(y))$, we define the *cut distance* as

$$\delta_{\square}(W_1, W_2) \inf_{\psi} ||W_1 - W_2^{\psi}||_{\square}$$

It's immediate that $||,||_{\square} \leq ||,||_1$. It's easy, although a bit tedious, to confirm that for two graphs F and G, $\delta_{\square}(F,G) = \delta_{\square}(W_F, W_G)$ and, further, that two kernels W_1 and W_2 are weakly isomorphic if and only if $\delta_{\square}(W_1, W_2) = 0$. This second fact allows us to consider $\delta_{\square}$ as a metric over $\widetilde{\mathscr{W}}$, the space of weak isomorphism classes of kernels. We would like to show that $(\widetilde{\mathscr{W}}, \delta_{\square})$ is compact, but, in order to do so, we need the Szemerédi Regularity Lemma.

3.3 Szemerédi Regularity Lemma

In this section a partition P will be defined as a finite set of measurable sets S_i such that $S_i \cap S_j = \emptyset$ and $\cup S_i = [0,1]$. A step function U is associated with some partition P and has the property that $U(x,y)$ depends only on which set S_i x and y are in. A refinement Q of a partition P is a partition such that each S_i is a subset of one of the sets of P. Given some kernel W and partition P, we define

$$W_P(x,y) = \frac{1}{\lambda(S_i)\lambda(S_j)} \int_{S_i \times S_j} W(x,y)dxdy$$

for $x \in S_i$, $y \in S_j$. $_P$ is called the stepping operator, sending kernels to a stepfunction that averages W over its steps $S_i \times S_j$. It is easy to confirm that the stepping operator is contractive in the cut norm.

All graphs have corresponding graphons that are stepfunctions. In order for graphons to be useful for describing graphs, it must be the case that all graphons are "graph-like" in some way. The Szemerédi Regularity Lemma gives this link, showing that graphons can be made arbitrarily close to step functions with respect to the cut distance. We only state the lemmas here without proof, as the proofs are outside of the scope of this thesis. See [23] and [6] for the complete proofs.

Theorem 3.3.1: [6] Weak Regularity Lemma for Kernels. For every $W \in \mathscr{W}$ and $k \geq 1$ there is a step function U with k steps such that

$$||W - U||_{\square} < \frac{2}{\sqrt{\log k}} ||W||_2$$

where $||,||_2$ is the L_2 norm. $\square$

As desired, the above theorem states that each W can be approximated by a stepfunction with k steps, and the error term depends only on $||W||_2$ (which, for graphons, is less than 1) and k. In order to prove compactness, it is convenient to have the ability to successively refine the graphons' corresponding stepfunctions, so we will use a different yet equivalent form of the Weak Regularity Lemma, which can be stated as follows:

Lemma 3.3.1: For $W \in \mathscr{W}_1$, $k > m \geq 1$, and some m-partition P of [0,1], there is a k-partition Q refining P such that

$$||W - W_Q||_\square \leq \frac{2}{\sqrt{\log k/m}}$$

In particular, this lemma says that no matter what partition P one starts with, there is a refinement Q such that W_Q is arbitrarily close to W in the cut distance.

Lemma 3.3.2: For $W \in \mathscr{W}_1$, U a step function, and P the partition of [0, 1] into the steps of U,

$$||W - W_P||_\square \leq 2||W - U||_\square$$

Proof: Because the stepping operator is contractive, $2||W-U||_\square \geq ||W - U||_\square + ||(W-U)_P||_\square$. Further,

$$\begin{aligned} ||W-U||_\square + ||(W-U)_P||_\square &= ||W-U||_\square + ||W_P - U||_\square \\ &\geq ||W - W_P|| \end{aligned}$$

by triangle inequality.

3.3.1 Compactness of $(\widetilde{\mathscr{W}_0}, \delta_\square)$

Theorem 3.3.2: [16] $(\widetilde{\mathscr{W}_0}, \delta_\square)$ is compact.

Proof: $(\widetilde{\mathscr{W}_0}, \delta_\square)$ is compact if and only every sequence of graphons (W_i) has a convergent subsequence. By Lemma 3.3.1, it is possible to construct partitions $P_{n,k}$ of [0, 1] and stepfunctions $W_{n,k} = (W_n)_{P_{n,k}}$ such that $P_{n,k+1}$ refines $P_{n,k}$, $|P_{n,k}| = s_k$ is only a function of k, and $||W_n - W_{n,k}||_\square < 1/k$. Assume without loss of generality that each partition class in $P_{n,k}$ is an interval. If not, there is some measure preserving bijection on W_n to make this the case.

For each k, we can find a subsequence (W_n) such that the length of the ith interval of the $W_{n,k}$ converges for each $i \leq s_k$ and the value of the the product of the ith and jth intervals converge because W_n is bounded and s_k is finite. Thus, each subsequence $W_{n,k}$ converges almost everywhere to some stepfunction U_k with s_k intervals. Let P_k denote the partition of $[0,1]$ into the steps of U_k. It's clear that since each $P_{n,k}$ was a refinement of all $P_{n,l}$ for $l < k$, it is also the case that $U_k = (U_l)_{P_k}$.

Choose a random point $(x, y) \in [0,1]^2$. Since each U_k is a refinement of U_l for $k > l$, $(U_k(x,y))$ is a bounded martingale, and so the sequence (U_k) converges almost everywhere. Let U be the limit of U_k. For arbitrary positive ϵ, since $U_k \to U$, there exists some $k > 1/\epsilon$ such that $||U - U_k||_1 < \epsilon$. Similarly, since $(W_{n,k}) \to U_k$, there exists some n such that $||W_{n,k} - U_k||_1 < \epsilon$. Finally, by the construction of $P_{n,k}, \delta_\square(W_{n,k}, W_n) < 1/k < \epsilon$. Thus by triangle inequality and the fact that $||, ||_1 \geq ||, ||_\square$

$$\begin{aligned}\delta_\square(U, W_n) &\leq \delta_\square(U, U_k) + \delta_\square(U_k, W_{n,k}) + \delta_\square(W_{n,k}, W_n) \\ &\leq ||U - U_k||_1 + ||U_k - W_{n,K}||_1 + \delta_\square(W_{n,k}, W_n) \\ &< 3\epsilon\end{aligned}$$

Thus, the subsequence (W_n) converges to U, and so $(\widetilde{\mathscr{W}_0}, \delta_\square)$ is compact. $\square$

3.4 Convergence of Dense Graph Sequences

We are almost ready to prove that graphons are, in fact, the appropriate limit object for convergent graph sequences. Because we now know that $(\widetilde{\mathscr{W}_0}, \delta_\square)$ is compact, we now just need that nearby graphons have similar homomorphism functions. The Counting Lemma provides exactly that.

Theorem 3.4.1: [15] Counting Lemma: For F simple and $W_1, W_2 \in \mathscr{W}_0$,

$$|t(F, W_1) - t(F, W_2)| \leq e(F)\delta_\square(W_1, W_2)$$

Proof: Label the edges of F as $e_t = (i_t, j_t)$, and let $E_t = \{e_1, \ldots, e_t\}$ and

$$X_t = (W_1(x_{i_t}, x_{j_t}) - W_2(x_{i_t}, x_{j_t})) \prod_{ij \in E_{t-1}} W_1(x_i, x_j) \prod_{ij \in E(F) - E_t} W_2(x_i, x_j)$$

The X_t are constructed explicitly so that their sum will telescope, and so

$$\sum_{t=1}^{e(F)} X_t = \prod_{ij \in E(F)} W_1(x_i, x_j) - \prod_{ij \in E(F)} W_2(x_i, x_j)$$

Further, since $0 \leq W(x,y), U(x,y) \leq 1$

$$\begin{aligned}\left|\int_0^1 \int_0^1 X_t dx_{i_t} dx_{j_t}\right| &\leq \left|\int_0^1 \int_0^1 W_1(x_{i_t}, x_{j_t}) - W_2(x_{i_t}, x_{j_t}) dx_{i_t} dx_{j_t}\right| \\ &\leq \delta_\square(W_1, W_2)\end{aligned}$$

We are now prepared to complete the proof, first using the definition of homomorphism functions on graphons, and then using the above relations.

$$\begin{aligned}|t(F, W_1) - t(F, W_2)| &= \left|\int_{[0,1]^{V(F)}} d\mathbf{x} \prod_{ij \in E(F)} W_1(x_i, x_j) - \prod_{ij \in E(F)} W_2(x_i, x_j)\right| \\ &= \left|\int_{[0,1]^{V(F)}} d\mathbf{x} \sum_{t=1}^{e(F)} X_t\right| \\ &\leq e(F)\delta(W_1, W_2)\end{aligned}$$

□

First, we show that the limit graphon exists for every convergent graph sequence.

Theorem 3.4.2: [15] For any convergent graph sequence (G_n), there exists a graphon W such that $G_n \to W$.

Proof: Because $(\widetilde{\mathscr{W}_0}, \delta_\square)$ is compact, the sequence W_{G_n} has some convergent subsequence (W_{G_i}) and there exists some graphon W such that $\delta_\square(W_{G_i}, W) \to 0$. Then it follows by the Counting Lemma that for every simple F

$$\begin{aligned}|t(F, G_i) - t(F, W)| &= |t(F, W_{G_i}) - t(F, W)| \\ &\leq e(F)\delta_\square(W_{G_i}, W)\end{aligned}$$

and so $t(F, G_i) \to t(F, W)$. Since $(t(F, G_n))$ has a subsequence that converges to $t(F, W)$ it must be the case that $(t(F, G_n))$ converges to $t(F, W)$, and so (G_n) converges to W. □

Next, we show that convergence can also be characterized by the cut distance.

Theorem 3.4.3: [12] $(G_n) \to W$ if and only if $\delta_\square(W_{G_n}, W) \to 0$.

Proof: As above, the Counting Lemma shows that $\delta_\square(W_{G_n}, W) \to 0$ implies $G_n \to W$. Now assume that $G_n \to W$. Then by definition $t(F, G_n) \to t(F, W)$ for all simple F. Consider the map ζ that sends $W \in \widetilde{\mathscr{W}_0}$ to the

family $(t(F, W))$ of all functions with simple F. This map is trivially injective by the definition of $\widetilde{\mathscr{W}_0}$ and clearly continuous by the definition of homomorphism functions on graphons. Thus, since $(\widetilde{\mathscr{W}_0}, \delta_\square)$ is compact, ζ^{-1} is continuous, and so $\delta_\square(W_{G_n}, W) \to 0$. □

To complete our identification of graphons as the limit object of convergent sequences, we would like for every graphon to be the limit of some simple graph sequence. This is the case. Generate a *W-random graph* $G(n, W)$ as follows. Let the n nodes be associated with some x_i chosen uniformly at random from the interval [0, 1]. Then for each set of nodes x_i and x_j, put an edge between them with probability $W(x_i, w_j)$, with the decision about each edge made independently. It would require too much of a diversion in probability theory to prove here, but it is the case that

Theorem 3.4.4: $(G(n, W))$ converges to W with probability 1.

Proof: See [3]. □

Thus, all convergent graph sequences converge to some W, every W has a sequence that converges to it, and convergence is well characterized by the cut distance. Graphons are the appropriate limit object.

Chapter 4

Extremal Graph Theory

We are now ready to characterize the possible relationships between edge density and triangle density. The Kruskal-Katona Bound gave a strict upper bound on triangle density, $t(K_3, G) \leq t(K_2, G)^{3/2}$. However, Goodman's Bound $t(K_3, G) \geq t(K_2, G)(2t(K_2, G) - 1)$ is not strict and is, in fact, only attained when $t(K_2, G) = \frac{k-1}{k}$ for integer k. Graph algebras alone were not able to improve this lower bound, but graphons provide new tools. First, we prove that the complete graphs are the extremal points of the lower bound. Next, graphons can be continuously deformed, and so one can define a variational calculus on the space of graphons. This additional tool will be the key to Razborov's proof.

4.1 Inequalities and Complete Graphs

In this chapter, we are considering homomorphism densities of complete graphs. In characterizing a linear inequality on quantum graphs g that are linear combinations of complete graphs, ideally testing a subset of such graphs and having the inequality being satisfied on this subset would imply that it is satisfied on all graphs. Goodman's Bound is saturated by the complete graphs, and so one might conjecture that the complete graphs are the desired subset. This is the case.

Theorem 4.1.1: [22] For g a quantum graph that is a linear combination of complete graphs, $t(g, .) \geq 0$ if and only if $t(g, K_n) \geq 0$ for all $n \geq 1$.

Proof: If $t(g, .) \geq 0$, then $t(g, K_n) \geq 0$ trivially.

If $t(g, K_n) \geq 0$ for all $n \geq 1$, then we want to show that $t(g, W) \geq 0$ for all W. Because $(\widetilde{\mathscr{W}_0}, \delta_\square)$ is compact, if the above holds for some dense subset of the space, then it holds for the whole space. We choose the dense subset W_H, where H is a simple weighted graph with edgeweights 0 or 1 and nodeweights α_i such that $\sum \alpha_i = 1$. It follows from Theorem 3.4.4 that this is a dense subset. Assume for the sake of contradiction that there is some H such that $t(g, H) < 0$. Choose the H with this property first with the minimum number of nodes, then such that $t(g, H)$ is minimized.

We claim that H is a complete graph. First, notice that since all homomorphisms in $hom(K_n, H)$ are injective and g is a linear combination of complete graphs, $t(g, H)$ is multilinear in the α_i. If two nodes i and j are not adjacent, then $t(g, H)$ has no term with $\alpha_i \alpha_j$, and so $t(g, H) = k + l\alpha_i + m\alpha_j$ for some k, l, m that don't depend on α_i and α_j. While keeping $\alpha_i + \alpha_j$ constant, one can change the values of α_i and α_j without increasing $t(g, H)$ until α_i or α_j is 0, giving a graph with 1 less node than our original H with negative homomorphism function, a contradiction. Thus, H is a complete graph. We also claim that all weights are equal. Since H is complete, we now know that $t(g, H)$ is a symmetric bilinear polynomial. Again only considering two weights α_i and α_j, $t(g, H) = k + l\alpha_i + m\alpha_j + n\alpha_i\alpha_j$ for k, l, m, n not depending on α_i and α_j. Since $t(g, H)$ is symmetric, $l = m$. Since $\alpha_i + \alpha_j$ is constant, then, $t(g, H) = k' + n\alpha_i\alpha_j$ for some k' that doesn't depend on α_i or α_j. If $n \geq 0$, the value is minimized by letting $\alpha_i = 0$, a contradiction as $\alpha_i > 0$. Thus, $n < 0$ and so $t(g, H)$ is minimized when $\alpha_i = \alpha_j$. But then the sign of $t(g, H)$ is the sign of $t(g, K_n)$ for some n, and it is nonnegative. Thus, we have a contradiction, and so $t(g, .) \geq 0$. $\square$

This result can be restated in a form more useful to our study of the relationship between edge density and triangle density. For integer $m \geq 1$ and graphon W, let $\mathbf{t}_W = (t(K_2, W), \ldots, t(K_m, W))$ and T_m the set of all vectors $\mathbf{t}_W$. Then T_m is the closure of the set of all $\mathbf{t}_G$ by Theorem 3.4.2 and 3.4.4. In particular, in this language, the above theorem states that the extreme points of the convex hull of T_m are the vectors $\mathbf{T}_{K_n}$ and $(1, 1, \ldots, 1)$. Applying this to the $m = 3$ case gives Bollobás's result [1], that $D_{2,3}$ lies above the convex hull of the points $(1, 1)$ and $(t(K_2, K_n), t(K_2, K_n)) = \left(\frac{n-1}{n}, \frac{(n-1)(n-2)}{n^2}\right)$, improving our lower bound from the quadratic Goodman's Bound to the polygonal curve formed by the points $\mathbf{T}_{K_n}$.

4.1.1 Generalized Turan's Theorem

Turan's Theorem states that the graph G on n nodes with the maximum number of edges without containing K_k as a subgraph is obtained by equipartitioning the nodes into $k-1$ subsets and then connecting all nodes not in the same subset. These graphs are called Turan graphs and denoted $T(n, k-1)$. The natural extension of this result is the question of how to maximize the number of edges of a graph subject to the restriction that there are no subgraphs $L_1, L_2, \ldots, L_m$. It turns out that, up to a negligible error, the optimizer depends only on the chromatic numbers of the L_i, denoted $\chi(L_i)$. We develop the appropriate result below, which mirrors the results proved by Zykov in 1949 [25].

Lemma 4.1.1: [12] For $r \geq 2$

$$\max\{t(K_2, W) : W \in \widetilde{\mathscr{W}_0}, t(K_r, W) = 0\} = 1 - \frac{1}{r-1}$$

with unique optimizer $W_{K_{r-1}}$.

Proof: Consider the inequality

$$r^r t(K_r, W) - (r-1)t(K_2, W) + r - 2 \geq 0$$

It's clear that it holds for all $W = W_{K_n}$, and so by Theorem 4.1.1 it holds for all W. Further, equality only potentially holds for $W = W_{K_n}$. Checking, it only holds for $W = W_{K_{r-1}}$. Since $t(K_r, K_{r-1}) = 0$, the inequality reduces to the result.

Lemma 4.1.2: For $L_1, \ldots, L_m$ simple graphs and $r = \min \chi(L_i)$,

$$\max\{t(K_2, W) : t(L_1, W) = \cdots = t(L_m, W) = 0\} = 1 - \frac{1}{r-1}$$

with unique optimizer $W_{K_{r-1}}$

Proof: Let $\chi(L_1) = r$. Then $t(L_1, K_r) > 0$ and so if $t(L_1, W) = 0$, $t(K_r, W) = 0$. The result then follows immediately from the above lemma.

Using the above Lemma, it is easy to show, although we do not here, that

Theorem 4.1.2: [4] (Generalized Turan's Theorem) Let $L_1, \ldots, L_m$ be simple graphs with $r = \min \chi(L_i)$ and G a graph that does not contain any L_i as a subgraph. Then

$$e(G) \leq \left(1 - \frac{1}{r-1} + o(1)\right) \binom{v(G)}{2}$$

4.2 Variational Calculus of Graphons

In the previous section, we described one way in which a graphon W_H can be continuously deformed: by decreasing the weight of one node while increasing the weight of another. We can generalize this notion. The below results were first shown by Razborov, although we use the notation of Lovász here [12]. Define a family of weight functions $\alpha_s : [0,1] \to \mathbb{R}^+$ such that $\int_0^1 \alpha_s(x)dx = 1$ and $\dot{\alpha}_s = \frac{d}{ds}a_s$ exists and is bounded. Each a_s gives rise to a graphon W_s with

$$t(F, W_s) = \int_{[0,1]^{V(F)}} \prod_{ij \in E(F)} W(x_i, x_j) \prod_{i \in V(F)} \alpha_s(x_i) \prod_{i \in V(F)} dx_i$$

Let $F^\dagger$ denote the sum of alll possible 1-labelings of F. Then

$$\frac{d}{ds}t(F, W_s) = \langle \dot{\alpha}_s, t_x(F^\dagger, W_s) \rangle$$

Another possible variation is characterized by a family $U_s \in \mathscr{W}_0$ such that $\dot{U}_s(x,y)$ exists and is bounded for all $x, y, 0 \leq s \leq 1$. Let F^{ij} denote the graph obtained from deleting the edge ij from F, $F^\ddagger = \frac{1}{2}\sum_{ij} F^{ij\cdot\cdot}$ with the labels going to the nodes i and j and $F^\sharp = K_2^{\cdot\cdot} F^\ddagger$. Then it follows from the definition of homomorphism functions that

$$\frac{d}{ds}t(F, U_s) = \langle \dot{U}_s, t_{xy}(F^\ddagger, U_s) \rangle$$

For some function f of $\mathscr{W}_0$, we call W a local minimizer on f if there is some $\epsilon > 0$ such that $f(U) \geq f(W)$ for all graphons U with $||U - W||_1 < \epsilon$.

In both of the below lemmas, let $\Phi : \mathbb{R}^m \to \mathbb{R}$ be a differentiable function, $\Phi_i = \frac{\partial}{\partial x_i}\Phi$, $F_1, \ldots F_m$ be simple graphs, W be a local minimizer of $\Phi(t(F_1, W), \ldots, t(F_m, W))$, and $a_i = \Phi_i(t(F_1, W), \ldots, t(F_m, W))$.

Lemma 4.2.1: [18] For almost all $x \in [0,1]$,

$$\sum_{i=1}^{m} a_i(t_x(F_i^\dagger, W) - v(F_i)t(F_i, W)) = 0$$

Proof: Let $\phi : [0,1] \to [-1,1]$ be a measurable function such that $\int \phi = 0$. Then $\alpha_s(x) = 1 + s\phi(x)$ defines a family of weight functions with corresponding graphons W_s. Since W is a local minimizer and $\lim_{s\to 0} W_s =$

W, $\frac{d}{ds}\Phi(t(F_1, W_s), \ldots, t(F_m, W_s))|_{s=0} = 0$. Thus,

$$\begin{aligned} 0 &= \frac{d}{ds}\Phi(t(F_1, W_s), \ldots, t(F_m, W_s))|_{s=0} \\ &= \sum_{i=1}^{m} \alpha_i \int_0^1 \phi(x) t_x(F_i^\dagger, W) dx \\ &= \int_0^1 \phi(x) \sum_{i=1}^{m} a_i t_x(F_i^\dagger, W) dx \end{aligned}$$

In particular, this relation holds for arbitrary ϕ, and so $\sum_{i=1}^m a_i t_x(F_i^\dagger, W)$ is a constant function of x almost everywhere. Integrating over x gives the desired result.

Lemma 4.2.2: [18] For almost all $x, y \in [0, 1]$,

$$\sum_{i=1}^{m} a_i t_{xy}(F_i^\sharp, W) \leq 0$$

Proof: The proof is very similar to the above. Let $U \in \mathscr{W}_1$ be a function such that $U(x, y) \geq 0$ if $W(x, y) = 0$ and $U(x, y) \leq 0$ if $W(x, y) = 1$. Then $U_s = W + sU$ defines a family of graphons for $0 \leq s < \epsilon$ for some $\epsilon > 0$. Further, since W is a local minimizer, $\frac{d}{ds}\Phi(t(F_1, W_s), \ldots, t(F_m, W_s))|_{s=0} \geq 0$. Thus,

$$\begin{aligned} 0 &\leq \frac{d}{ds}\Phi(t(F_1, W_s), \ldots, t(F_m, W_s))|_{s=0} \\ &= \int_{[0,1]^2} U(x, y) \sum_{i=1}^{m} a_i t_{xy}(F_i^\ddagger, W) dx dy \end{aligned}$$

Since this holds for arbitrary U subject to the constraints, it follows that $\sum_{i=1}^m a_i t_{xy}(F_i^\ddagger, W) = 0$ if $0 < W(x, y) < 1$, is less than or equal to 0 if $W(x, y) = 1$ and greater than or equal to 0 if $W(x, y) = 1$. Notice that $W(x, y) = t_{xy}(K_2^{\cdot\cdot})$. Then

$$\begin{aligned} \sum_{i=1}^{m} a_i t_{xy}(F_i^\sharp, W) &= t_{xy}(K_2^{\cdot\cdot}) \left(\sum_{i=1}^{m} a_i t_{xy}(F_i^\ddagger, W) \right) \\ &= W(x, y) \left(\sum_{i=1}^{m} a_i t_{xy}(F_i^\ddagger, W) \right) \end{aligned}$$

Reading off from our results for the various cases for $W(x, y)$, we get that $\sum_{i=1}^m a_i t_{xy}(F_i^\sharp, W) \leq 0$ as desired.

4.3 Razborov's Proof

The variational calculus finally allows for the complete characterization of $D_{2,3}$. We present Razborov's proof in 3 parts following a version of the proof given by Lovász in the language of graphons. First, we consider the conjecture of Lovász and Simonovitis and the bound that it implies. We then prove the bound in the interval $t(K_2, G) \in [1/2, 2/3]$, first proven by Fisher in 1989. Finally, we present a proof of the bound for all edge densities, a result which was first shown by Razborov in 2007.

Lovász and Simonovitis conjectured the following:

Theorem 4.3.1: [14] For $t(K_2, W) = d$, the minimum of $t(K_3, W)$ for $W \in \widehat{\mathscr{W}_0}$ is attained by W_H where H is a weighted complete graph on $k = \lceil \frac{1}{1-d} \rceil$ nodes with edgeweights 1 and nodeweights all equal except for one node which has weight no greater than the others.

Proof: First, it's worth noting that the above does not claim that W_H is unique. In fact, this is not the case.

We now find the bound implied by the above claim. Nikiforov [17] suggested letting the weights of the first $k-1$ nodes be $\alpha_1 = \frac{1+u}{k}$ and the weight of the final node be $a_k = \frac{1-(k-1)u}{k}$. Then

$$\begin{aligned} d &= t(K_2, W_H) \\ &= (k-1)\frac{1+u}{k}\left(1 - \frac{1+u}{k}\right) + \frac{1-(k-1)u}{k}\left(1 - \frac{1-(k-1)u}{k}\right) \\ &= \frac{k-1}{k}\left((1+u)\frac{k-1-u}{k} + \frac{1-ku+u}{k}(1-u)\right) \\ &= \frac{k-1}{k}(1-u^2) \\ \Rightarrow u &= \sqrt{1 - \frac{kd}{k-1}} \end{aligned}$$

Similarly, we can calculate the triangle density $f(d)$

$$\begin{aligned} f(d) &= t(K_3, W_H) \\ &= \frac{(k-1)(k-2)}{k^2}(1+u)^2(1-2u) \end{aligned}$$

Thus, if we can show that $f(d)$ is the lower bound for $D_{2,3}$, then the conjecture will be proved. We set out to do so below. □

The $k = 3$ case is easier and was shown by Fisher in 1989:

Theorem 4.3.2: [5] If $t(K_2, G) = d$, then $t(K_3, G) \geq f(d)$ for

$$f(d) = \frac{(k-1)(k-2)}{k^2}\left(1 + \sqrt{1 - \frac{kd}{k-1}}\right)^2\left(1 - 2\sqrt{1 - \frac{kd}{k-1}}\right)$$

and $k = \lceil \frac{1}{1-d} \rceil = 3$.

Proof: We keep the proof valid for general k as long as possible as the proof for general k is the same at the beginning. Let W be the graphon with $\frac{k-2}{k-1} \leq t(K_2, W) = d \leq \frac{k-1}{k}$ that minimizes $\phi(W) = t(K_3, W) - f(t(K_2, W))$. Assume that this value is negative. By Goodman's Bound, $\phi(W) \geq 0$ for d at the endpoints, and so $\frac{k-2}{k-1} < d < \frac{k-1}{k}$, which means that W is a local minimizer in $\widetilde{\mathscr{W}_0}$. This means we can use the results from the variational calculus, letting $\lambda = f'(d) = 3\frac{k-2}{k}(1+u)$. Notice that λ positive. In all of the below, we, as before, let $t(F, W)$ be denoted by F.

By Lemma 4.2.1,

$$3K_3^{\cdot} - 2\lambda K_2^{\cdot} = 3K_3 - 2\lambda K_2$$

Multiplying by $K_2^{\cdot}$ and then unlabeling gives

$$3K_4^{ij,ik} - 2\lambda K_3^{ij} = 3K_3K_2 - 2\lambda K_2K_2$$

This is our first relevant equality. We develop a relevant inequality as follows: By Lemma 4.2.2, $3K_3^{\cdot\cdot} - \lambda K_2^{\cdot\cdot} \leq 0$. Multiplying each side by the always non-negative $K_3^{\cdot\cdot} - 2K_2^{\cdot} + 1$ and then unlabeling gives

$$3K_4^{ij} - 6K_4^{ij,ik} \leq (\lambda - 3)K_3 + \lambda K_2 - 2\lambda P_3$$

We need one more inequality. Consider the probability that a map of 4 nodes into some graphon W creates a subgraph with a specific node isolated and the rest forming a triangle. The probability of this is clearly at least zero, and it can be expressed using the principle of inclusion exclusion in terms of homomorphism functions as $K_3 - 3K_4^{ij,ik} + 3K_4^{ij} - K_4$. Thus,

$$3K_3^{ij,ik} - 3K_3^{ij} \leq K_3 - K_4$$

Adding our three inequalities together and replacing K_2 with d gives

$$(\lambda + 3d - 2)K_3 \geq \lambda(2d^2 - d) + K_4$$

For $k = 3$, $\lambda + 3d - 2 > 0$ and so removing the always positive K_4 term and solving for K_3 gives

$$\begin{aligned} K_3 &\geq \frac{\lambda(2d^2 - d)}{\lambda + 3d - 2} \\ &= f(d) \end{aligned}$$

after some algebraic manipulation. Thus, $\phi(W)$ is nonnegative and our bound is correct for $k = 3$. □

It makes sense that throwing out the K_4 term works for the $k = 3$ case because the optimizing graphons are in the form W_H where H has only three nodes. However, for $k > 3$, we need an explicit bound for K_4.

Theorem 4.3.3: [12] [19] If $t(K_2, G) = d$, then $t(K_3, G) \geq f(d)$ where

$$f(d) = \frac{(k-1)(k-2)}{k^2}\left(1 + \sqrt{1 - \frac{kd}{k-1}}\right)^2\left(1 - 2\sqrt{1 - \frac{kd}{k-1}}\right)$$

where $k = \lceil \frac{1}{1-d} \rceil$.

Proof: Our plan is to prove the result by induction. We have a base case, $k = 3$. We now just need to show that if the result holds for some k, then it holds for $k + 1$. We use the same proof as above up until the result

$$(\lambda + 3d - 2)K_3 \geq \lambda(2d^2 - d) + K_4.$$

We now derive a bound for K_4. By induction, we have a bound for $t(K_3, U)$ for all graphons U with $t(K_2, U) < \frac{k-2}{k-1}$. We thus look to express K_4 in terms of $t(K_3, U)$ for such a graphon U. Pick any $z \in [0, 1]$. Define a graphon W_z that is the same as W modified by the weight function $W(z, .)/d_W(z)$ where $d_W(z)$ is the degree of z as defined in Example 3.1.1. We aim to show that $t(K_2, W_z) \leq \frac{k-2}{k-1}$.

Let $\mu = 2\lambda d - 3K_3$. Then the equality $3K_3^{\cdot} - 2\lambda K_2^{\cdot} = 3K_3 - 2\lambda K_2$ in the above proof can be restated as $3t_z(K_3^{\cdot}, W) - 2\lambda d_W(z) = -\mu$. By its construction, $t(K_2, W_z) = \frac{t_z(K_3, W)}{d_W(z)^2}$ and so

$$t(K_2, W_z) = \frac{2\lambda d_W(z) - \mu}{3d_W(z)}$$

We thus want to try to find good bounds for $\mu, d_W(z)$, and λ. The explicit description of λ above gives $\lambda \in [3(k-2)/k, 3(k-2)/(k-1)]$. By Goodman's

Bound and our original hypothesis, $d(2d-1) \leq t(K_3, W) < f(d)$ and so, after some algebraic manipulation

$$2d\lambda - 3f(d) < \mu \leq 2d\lambda - 3d(2d-1)$$

It takes a bit of algebraic manipulation, but these two relations imply that

$$\frac{k-1}{k-2} < \frac{3\mu}{\lambda^2} \leq \frac{k-2}{k-3}$$

Since we have already show that $3K_3^{\cdot\cdot} - \lambda K_2^{\cdot\cdot} \leq 0$, it follows by unlabeling one node that $3t_z(K_3^{\cdot}, W) \leq \lambda d_W(z)$, and so using the definition of μ we get

$$\frac{\mu}{2\lambda} \leq d_W(z) \leq \frac{\mu}{\lambda}$$

Returning to our equation, for $t(K_2, W_z)$,

$$\frac{\partial}{\partial d_W(z)} \frac{2\lambda d_W(z) - \mu}{3d_W(z)} = \frac{2(\mu - \lambda d_W(z))}{3d_W(z)^3}$$

and so $t(K_2, W_z)$ is a non-decreasing function in the range of $d_W(z)$. This means we can plug in $d_W(z) = \frac{\mu}{\lambda}$ to get the bound

$$t(K_2, W_z) \leq \frac{\lambda^2}{3\mu} \leq \frac{k-2}{k-1}$$

Thus, the edge density of W_z is sufficiently small that we can use the inductive hypotheis that $t(K_3, W_z) \geq f(t(K_2, W_z))$. Thus,

$$\begin{aligned} t(K_4, W) &= \int_0^1 t_z(K_4^{\cdot}, W) dz \\ &= \int_0^1 t(K_3, W_z) d_W(z)^3 dz \\ &\geq \int_0^1 f\left(\frac{2\lambda d_W(z) - \mu}{3d_W(z)^2}\right) d_W(z)^3 dz \end{aligned}$$

In some sense, this is enough to prove the theorem, but the integral is essentially impossible to use in its current form. Playing around with the function inside the integral, though, shows that it is roughly linear in $d_W(z)$. The idea, then, is to give a linear lower bound of the function inside the integral. Let $g(z) = \frac{2\lambda z - \mu}{3z^2}$ and $h(z) = f(g(z))z^3$. Since g is continuous

and $g(\frac{\mu}{2\lambda}) \leq \frac{k-3}{k-2} \leq g(\frac{\mu}{\lambda})$, there exists some z_0 such that $g(z_0) = \frac{k-3}{k-2}$ for $z_0 \in [\frac{\mu}{2\lambda}, \frac{\mu}{\lambda}]$. Thus,

$$h(z_0) = f\left(\frac{k-3}{k-2}\right) z_0^3 = \frac{(k-3)(k-4)}{(k-2)^2} z_0^3$$

Now that everything is explicit, one can prove using simple yet messy calculus (which we omit here) that

$$h(z) - h(z_0) \geq \frac{1}{3}(2\lambda^2 - 3\mu)(z - z_0)$$

Thus,

$$\begin{aligned} t(K_4, W) &\geq \int_0^1 h(d_W(z))dz \\ &\geq \frac{1}{3}\int_0^1 (2\lambda^2 - 3\mu) d_W(z) dz + h(z_0) - \frac{1}{3}(2\lambda^2 - 3\mu) z_0 \\ &= \frac{(k-3)(k-4)}{(k-2)^2} z_0^3 + \frac{1}{3}(2\lambda^2 - 3\mu)(d - z_0) \end{aligned}$$

Plugging this into the inequality

$$(\lambda + 3d - 2)K_3 \geq \lambda(2d^2 - d) + K_4$$

gives, after some algebraic manipulation, an expression only in terms of k, u, and $y = w_0/(1+u)$:

$$\begin{aligned} &(1+u)^2 \left(y - \frac{k-2}{k}\right) \times \\ &\left(\frac{k-1}{k}(1-u) - \left(\frac{k-3}{k-2}u + \frac{3k-7}{k-2}\right) y + \frac{2(k-1)(k-3)}{k(k-2)}(1+u)y^2\right) \geq 0 \end{aligned}$$

Further, we know that $0 < u < \frac{1}{k-1}$ and can easily show from our previous inequalities that $\frac{k-2}{k} \leq y \leq \frac{(k-2)^2}{k(k-3)}$. After another tedious calculus calculation, one can show that the above value subject to the constraints is always negative, and so we have a contradiction. There is no W such that $f(K_3, W) < f(t(K_2, W))$. □

Thus, not only have we proved a bound, but we have graphs which give values arbitrarily close to the bound, and so $D_{2,3}$ is finally completely characterized. The next question one might ask is how to characterize $D_{2,n}$ for all n. This

is highly non-trivial and is outside of the scope of this thesis, but it turns out that the ideas behind flag algebras or, equivalently, graphons, are the key. In fact, after Razborov's proof in 2007, more general proofs came quickly. In 2011, Nikiforov extended the result to $D_{2,4}$[17], and, in 2012, Reiher completely characterized $D_{2,n}$ for all n [20]. The application of graph limit theory to extremal graph theory continues to be an active field of research to this day.

Bibliography

[1] B. Bollobás. Relations between sets of complete subgraphs. *Combinatorics*, Proc. 5th British Comb. Conf. (ed. C.St.J.A. Nash-Williams, J. Sheehan), Utilitas Math. (1976), 79-84.

[2] C. Borgs, J. Chayes, L. Lovász, V.T. Sós and K. Vesztergombi. Counting graph homomorphism. *Topics in Discrete Mathematics* (ed. M. Klazar, J. Kratochvil, M. Loebl, J. Matoušek, R. Thomas, P. Valtr), Springer (2006), 315-371.

[3] P. Diaconis and S. Janson. Graph limits and exchangeable random graphs. *Rendiconti di Matematica* **28** (2008), 33-61.

[4] P. Erdős and M. Simonovitis. A limit theorem in graph theory. *Studia Sci. Math. Hungar.* **1** (1966), 51-57.

[5] D.C. Fisher. Lower bounds on the number of triangles in a graph. *J. Graph Theory* **13** (1989), 505-512.

[6] A. Frieze and R. Kanaan. Quick approximation to matrices and applications. *Combinatorica* **19** (1999), 175-220.

[7] B. Godlin, T. Kotek and J.A. Makowsky. Evaluations of Graph Polynomials. *34th Int. Workshop on Graph-Theoretic Concepts in Comp. Sci.*, Lecture Notes in Comp. Sic. 5344, Springer, Berlin-Heidelberg (2008), 183-194.

[8] A.W. Goodman. On sets of acquaintances and strangers at any party. *Amer. Math. Monthly* **66** (1959) 778-783.

[9] G.O.H. Katona. Continuous versions of some extremal hypergraph problems. *Combinatorics*, (Keszthely, Hungary, 1976), Coll. Math. Soc. J. Bolyai **18** (J. Bolyai Mat. Soc., Budapest, 1978), 653-678.

[10] W. Mantel. Problem 28, solution by H. Gouwentak, W. Matel, J. Teixeira de Mattes, F. Schuh and W.A. Wythoff. *Wiskundige Opgaven*, 10:60-61, 1907.

[11] L. Lovász. The Rank of connection matrices and the dimension of graph algebras. *Eur. J. Comb.* **27** (2006), 962-970.

[12] L. Lovász. *Large Networks and Graph Limits*. American Mathematical Society, Rhode Island, 2012.

[13] L. Lovász. Operations with structures. *Acta Math. Hung.* **18** (1967), 321-328.

[14] L. Lovász and M. Simonovitis. On the number of complete subgraphs. *Combinatorics*, Proc. 5th British Comb. Conf. (ed. C.St.J.A.Nash-Williams, J.Sheehan), Utilitas Math. (1976), 429-441.

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

[16] L. Lovász and B. Szegedy. Szemerédi's Lemma for the analyst. *Geom. Func. Anal.* **17** (2007), 252-270.

[17] V. Nikiforov. The number of cliques in graphs of given order and size. *Trans. Amer. Math. Soc.* **363** (2011), 1599-1618.

[18] A.A. Razborov. Flag Algebras. *J. Symbolic Logic* **72** (2007), 1239-1282.

[19] A.A. Razborov. On the mimimal density of triangles in graphs. *Combin. Prob. Comput.* **17** (2008), 603-618.

[20] C. Reiher. The clique density theorem. arXiv:1212.2454 [math.CO].

[21] N. Robertson and P. Seymour. Graph Minors. XX. Wagner's Conjecture. *J. Combin. Theory B* **92**, 325-357.

[22] R.H. Schelp and A. Thomason. Remark on the number of complete and empty subgraphs. *Combin. Prob. Comput.* **7** (1998), 217-219.

[23] E. Szemerédi. Regular partitions of graphs. *Colloque Inter. CNRS* (J.-C. Fournier, M. Las Vergnas and D. Sotteau, eds.) (1978) 399-401.

[24] J.H. van Lint and R.M. Wilson. *A Course in Combinatorics*. Cambridge University Press (1992).

[25] A.A. Zykov. On some properties of linear complexes. *Mat. Sbornik* **24** (1949), 163-188. *Amer. Math. Soc. Transl.* **79**, 1-33.