Multipodal Structure and Phase Transitions in Large Constrained Graphs

Richard Kenyon* Charles Radin† Kui Ren ‡ Lorenzo Sadun§

May 3, 2014

Abstract

We study the asymptotics of large, simple, labeled graphs constrained by the densities of k-star subgraphs for two or more k, including edges. We prove that for any set of fixed constraints, such graphs are "multipodal": asymptotically in the number of vertices there is a partition of the vertices into $M < \infty$ subsets $V_1, V_2, \ldots, V_M$, and a set of well-defined probabilities q_{ij} of an edge between any $v_i \in V_i$ and $v_j \in V_j$. We also prove, in the 2-constraint case where the constraints are on edges and 2-stars, the existence of inequivalent optima at certain parameter values. Finally, we give evidence based on simulation, that throughout the space of the constraint parameters of the 2-star model the graphs are not just multipodal but bipodal (M=2), easily understood as extensions of the known optimizers on the boundary of the parameter space, and that the degenerate optima correspond to a non-analyticity in the entropy.

1 Introduction

We study the asymptotics of large, simple, labeled graphs constrained to have certain fixed *subgraph densities* (see definition below). We consider the simplest cases, called *star models*, with $\ell \geq 2$ constraints where the subgraphs are "k-stars": $k \geq 1$ edges with a common vertex, always including edges (1-stars) as one of the constraints.

Using the graphon formalism of Lovász *et al* to frame the asymptotics, we prove that all constrained graphons maximizing the entropy are "multipodal": there is a partition of the vertices into $M < \infty$ subsets $V_1, V_2, \ldots, V_M$, and a set of well-defined probabilities q_{ij} of an edge between any $v_i \in V_i$ and $v_j \in V_j$. In particular the optimizing graphons are piecewise constant, attaining only finitely many values. We also prove, in the 2-constraint

*Department of Mathematics, Brown University, Providence, RI 02912; rkenyon at math.brown.edu
†Department of Mathematics, University of Texas, Austin, TX 78712; radin@math.utexas.edu
‡Department of Mathematics, University of Texas, Austin, TX 78712; ren@math.utexas.edu
§Department of Mathematics, University of Texas, Austin, TX 78712; sadun@math.utexas.edu

case where the constraints are on edges and 2-stars (the "2-star model"), the existence of inequivalent optimizing graphons on a line segment of parameter space. In particular we prove the existence of a phase transition in the sense that the maximizing graphons do not vary continuously with the constraint parameters.

Finally, we give evidence based on simulation that, throughout the space of the constraint parameters of the 2-star model, the optimizers are not just multipodal but bipodal (M=2), easily understood as extensions of the known optimizers on the boundary of the parameter space, and that the inequivalent optimizers correspond to a line of nonanalyticity in the entropy function.

The existence of multipodal optimizers emerged in a series of three papers [RS1, RS2, RRS] on a model with different contraints: edges and triangles, rather than edges and stars. In the triangle model evidence, but not proof, was given that entropy optimizers were M-podal throughout the whole of the parameter space, M growing without bound as edge density approaches 1. Here we *prove* that all optimizers are multipodal, in all star models.

Models related to the above star and triangle models (so-called exponential random graph models (ERGMs)) have been extensively studied and applied: see for instance [N, Lov] and the many references therein. In physics terminology the models in [RS1, RS2, RRS] and this paper are "microcanonical" whereas the ERGMs based on the same subgraph densities are the corresponding "grand canonical" versions or ensembles. In distinction with statistical mechanics with short range forces [Ru, TET], here the microcanonical and grand canonical ensembles are inequivalent [RS1] and in the conclusion below we discuss the extent of the loss of information in ERGMs as compared with microcanonical models. Continuing the analogy with statistical mechanics we also describe the multipodal structure as embodying the emergence of phases in all such parametric families of large graphs, as vertex number grows.

2 Notation and background

Fix distinct positive integers $k_1, \ldots, k_\ell,\ \ell \geq 2$, and consider simple (undirected, with no multiple edges or loops) graphs G with vertex set $V(G)$ of labeled vertices, and for each $k = k_i$, the k-star set $T_k(G)$, the set of graph homomorphisms from a k-star into G. We assume $k_1 = 1$ so the k_1-star is an edge. Let $n = |V(G)|$. The *density* of a subgraph H refers to the relative fraction of maps from $V(H)$ into $V(G)$ which preserve edges: the k-star density is

$$t_k(G) \equiv \frac{|T_k(G)|}{n^{k+1}}. \tag{1}$$

For $\alpha > 0$ and $\tau = (\tau_1, \cdots \tau_\ell)$ define $Z^{n,\alpha}_\tau$ to be the number of graphs with densities

$$t_{k_i}(G) \in (\tau_i - \alpha, \tau_i + \alpha),\ 1 \leq i \leq \ell. \tag{2}$$

We sometimes denote τ_1 by ϵ and $T_1(G)$ by $E(G)$.

Define the *entropy density* s_τ to be the exponential rate of growth of $Z_\tau^{n,\alpha}$ as a function of n:

$$s_\tau = \lim_{\alpha \downarrow 0} \lim_{n\to\infty} \frac{\ln(Z_\tau^{n,\alpha})}{n^2}. \tag{3}$$

The double limit defining the entropy density s_τ is known to exist [RS1]. To analyze it we make use of a variational characterization of s_τ, and for this we need further notation to analyze limits of graphs as $n \to \infty$. (This work was recently developed in [LS1, LS2, BCLSV, BCL, LS3]; see also the recent book [Lov].) The (symmetric) adjacency matrices of graphs on n vertices are replaced, in this formalism, by symmetric, measurable functions $g : [0,1]^2 \to [0,1]$; the former are recovered by using a partition of $[0,1]$ into n consecutive subintervals. The functions g are called graphons.

For a graphon g define the *degree function* $d(x)$ to be $d(x) = \int_0^1 g(x,y)dy$. The k-star density of g, $t_k(g)$, can then be defined as

$$t_k(g) = \int_0^1 d^k(x)\,dx. \tag{4}$$

Finally, the *entropy density* of g is

$$s(g) = \frac{1}{2}\int_{[0,1]^2} S[g(x,y)]\,dxdy, \tag{5}$$

where S is the Shannon entropy function

$$S(w) = -w\log w - (1-w)\log(1-w). \tag{6}$$

The following is a minor variant of a result in [RS1] (itself an adaption of a proof in [CV]):

Theorem 2.1 (The Variational Principle.)**.** *For any feasible set τ of values of the densities $t(g)$ we have $s_\tau = \max[s(g)]$, where the maximum is over all graphons g with $t(g) = \tau$.*

(Some authors use instead the *rate function* $I(g) \equiv -s(g)$, and then minimize I.) The existence of a maximizing graphon $g = g_\tau$ for any constraint $t(g) = \tau$ was proven in [RS1], again adapting a proof in [CV]. We refer to this maximization problem as a *star model*.

We want to consider two graphs *equivalent* if they are obtained from one another by relabeling the vertices. For graphons, the analogous operation is applying a measure-preserving map ψ of $[0,1]$ into itself, replacing $g(x,y)$ with $g(\psi(x),\psi(y))$, see [Lov]. The equivalence classes of graphons under relabeling are called *reduced graphons*, and on this space there is a natural metric, the *cut metric*, with respect to which graphons are equivalent if and only if they have the same subgraph densities for all possible finite subgraphs [Lov].

The graphons which maximize the constrained entropy tell us what 'most' or 'typical' large constrained graphs are like: if g_τ is the only reduced graphon maximizing $s(g)$ with $t(g) = \tau$, then as the number n of vertices diverges and $\alpha_n \to 0$, exponentially most graphs with densities $t_i(G) \in (\tau_i - \alpha_n, \tau_i + \alpha_n)$ will have reduced graphon close to g_τ [RS1].

3 Multipodal Structure

Our main results are the following theorem and the subsequent corollary.

Theorem 3.1. *For any star model, any graphon g which maximizes the entropy $s(g)$, and is constrained by $t(g) = \tau$, is M-podal for some $M < \infty$.*

Proof. The Euler-Lagrange equation for constrained maximization of $s(g)$ is obtained by embedding $g(x, y)$ in a curve $g(x, y) + w\,h(x, y)$ using an arbitrary $h(x, y) = h(y, x)$, and setting equal to 0 the derivative with respect to the real variable w:

$$\frac{d}{dw}\bigg|_{w=0} s(g + wh) + \beta \cdot t(g + wh) = 0, \tag{7}$$

where $\beta = (\beta_1, \ldots, \beta_\ell)$ are Lagrange multipliers. The result, absorbing constants into β, is:

$$2\beta_1 + \sum_{i=2}^{\ell} \beta_i d^{k_i-1}(x) + \sum_{i=2}^{\ell} \beta_i d^{k_i-1}(y) = \ln\left[\frac{1}{g(x, y)} - 1\right]. \tag{8}$$

Solving for $g(x, y)$ gives

$$g(x, y) = \frac{1}{1 + \exp(2\beta_1 + \sum \beta_i d^{k_i-1}(x) + \sum \beta_i d^{k_i-1}(y))}, \tag{9}$$

and integrating with respect to y gives

$$d(x) = \int_0^1 \frac{dy}{1 + \exp(2\beta_1 + \sum \beta_i d^{k_i-1}(x) + \sum \beta_i d^{k_i-1}(y))}. \tag{10}$$

Let $d(x)$ be any solution of (10), let z be a real variable, and consider the function

$$F(z) = z - \int_0^1 \frac{dy}{1 + \exp(2\beta_1 + \sum \beta_i z^{k_i-1} + \sum \beta_i d^{k_i-1}(y))}, \tag{11}$$

where the function $d(y)$ is treated as given. By equation (10), all actual values of $d(x)$ are roots of $F(z)$.

The second term in (11) is an analytic function of z, as follows.

Write $W = \sum_{i=2}^{\ell} \beta_i z^{k_i-1}$ and $Y = \sum_{i=2}^{\ell} \beta_i d(y)^{k_i-1}$ then the integral is

$$\int \frac{d\mu(Y)}{1 + \exp(2\beta_1 + W + Y)}, \tag{12}$$

the convolution of an analytic function of W with an integrable measure $\mu(Y)$. Since the Fourier transform of an analytic function decays exponentially at infinity and the Fourier transform of an integrable measure is bounded, the Fourier transform of the convolution decays exponentially at infinity, so the convolution itself is an analytic function of W. Since W is an analytic function of z, $F(z)$ is an analytic function of z.

Note that $F(z)$ is strictly negative for $z \leq 0$ and strictly positive for $z \geq 1$. Being analytic and not identically zero, $F(z)$ can only have finitely many roots in any compact interval. (By Rolle's Theorem any accumulation point of the roots would have to be an accumulation point of the roots of $F'(z)$, $F''(z)$, etc. So all derivatives of F would have to vanish at the point, making the Taylor series around it identically zero.) In particular $F(z)$ can only have finitely many roots between 0 and 1, implying there are only finitely many values of $d(x)$. Since $d(x)$ and $d(y)$ determine $g(x, y)$ by equation (9), the graphon g is M-podal, where M is the number of distinct values, $d_1, \cdots, d_M$, of $d(x)$; concretely, we can take the partition of the vertex set to be $V_j = d^{-1}(d_j)$. Note that the roots of $F(z)$ are not necessarily values of $d(x)$, so this construction only gives an upper bound to the actual value of M. □

Let c_j be the measure of the set $\{x \in [0, 1] \mid d(x) = d_j\}$. We can apply a measure-preserving transformation so that $d(x) = d_1$ on $[0, c_1]$, $d(x) = d_2$ on $[c_1, c_1 + c_2]$, etc. If x is in the j-th interval and y is in the m-th interval, then

$$g(x, y) = \frac{1}{1 + \exp(2\beta_1 + \sum_{i=2}^{\ell} \beta_i (d_j^{k_i-1} + d_m^{k_i-1}))}. \tag{13}$$

Thus g is piecewise constant. A graph with N vertices corresponding to this graphon will have approximately Nc_j vertices with degree close to Nd_j for each $j \in \{1, 2, \ldots, M\}$. There will be approximately $N^2 c_j c_m / (1 + \exp(2\beta_1 + \sum \beta_i (d_j^{k_i-1} + d_m^{k_i-1})))$ edges between vertices in cluster j and vertices in cluster m (if $j \neq m$, half that if $j = m$), and these edges will be statistically independent of one another.

Corollary 3.2. *The entropy function $s(\tau)$ is piecewise analytic.*

Proof. For fixed M, the entropy and k-star densities of M-podal graphons are analytic functions of the parameters c_i and g_{ij}. We then have an $M + \binom{M+1}{2}$ dimensional variational problem involving analytic functions. The Euler-Lagrange equations for stationary points of the entropy are then finite systems of analytic equatoins, whose solutions must be analytic functions of the parameter τ. Consequently, wherever the maximizing graphon varies continuously with τ, it must vary analytically, as must the entropy. The only places where $s(\tau)$ is not real-analytic is along codimension-1 "phase transition" surfaces where the maximization problem has two (or more) solutions, either with different values of M or within the space of M-podal graphons.

□

4 Phase space

We now simplify to the case $\ell = 2$, and only restrict the number of edges and the number of 2-stars. The phase space (see Fig. 1) is then the set of those $(\epsilon, \tau_2) \subset [0, 1]^2$ which are accumulation points of the values of pairs (edge density, 2-star density) for finite graphs.

The lower boundary (minimum of τ_2 given ϵ) is easily seen to be the Erdös-Rényi curve: $\tau_2 = \epsilon^2$. We now consider the upper boundary, which was determined in [AK].

We call a graphon a *g-clique* if it is bipodal of the form

$$g(x,y) = \begin{cases} 1 & x < c \text{ and } y < c \\ 0 & \text{otherwise} \end{cases} \tag{14}$$

and a *g-anticlique* if it is of the form

$$g(x,y) = \begin{cases} 0 & x > c \text{ and } y > c \\ 1 & \text{otherwise.} \end{cases} \tag{15}$$

Theorem 4.1. *[AK]. For fixed $e(g) \equiv t_1(g) = \epsilon$, any graphon that maximizes the 2-star density $t_2(g)$ is either a g-clique or g-anticlique.*

G-cliques always have $c = \sqrt{\epsilon}$ and 2-star density $\epsilon^{3/2}$. G-anticliques have $c = 1 - \sqrt{1-\epsilon}$ and 2-star density

$$c + c^2 - c^3 = 2\epsilon + [1-\epsilon]^{3/2} - 1. \tag{16}$$

For ϵ small, the g-anticlique has 2-star density $\frac{\epsilon}{2} + O(\epsilon^2)$, which is greater than $\epsilon^{3/2}$. For ϵ close to 1, however, the g-clique has a higher 2-star density than the g-anticlique.

Corollary 4.2. *The upper boundary of the phase space is*

$$\tau_2 = \begin{cases} 2\epsilon + [1-\epsilon]^{3/2} - 1 & \epsilon \le 1/2 \\ \epsilon^{3/2} & \epsilon \ge 1/2. \end{cases} \tag{17}$$

The boundary of the phase space is shown in Fig. 1.

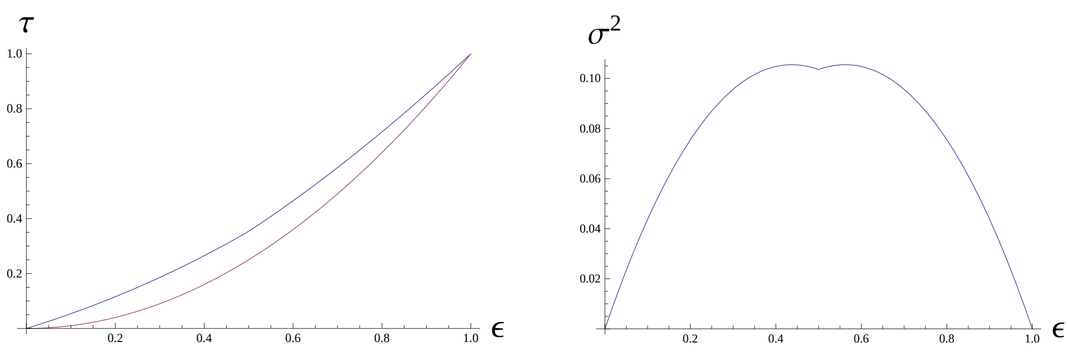

Figure 1: Boundary of the phase space for the 2-star case. Left: true phase boundary; Right: Plot of ϵ versus $\sigma^2 = \tau_2 - \epsilon^2$; in this case the lower boundary becomes the x-axis.

5 Phase Transition for 2-Stars

Theorem 5.1. *For the 2-star model there are inequivalent graphons maximizing the constrained entropy on the line segment* $\{(1/2, \tau_2) \,|\, \tau^* < \tau_2 \leq 2^{-3/2}\}$ *for some* $\tau^* < 2^{-3/2}$*. Moreover, near* $(1/2, 2^{3/2})$*, the maximizing graphons do not vary continuously with the constraint parameters.*

Proof. For any graphon g, consider the graphon $g'(x, y) = 1 - g(x, y)$. The degree functions for g' and g are related by $d'(x) = 1 - d(x)$. If g has edge and 2-star densities ϵ and τ_2, then g' has edge density $1 - \epsilon$ and 2-star density $\int_0^1 (1 - d(x))^2 dx = 1 - 2\epsilon + \tau_2$. Furthermore, $s(g') = s(g)$. This implies that g' maximizes the entropy at $(1 - \epsilon, 1 - 2\epsilon + \tau_2)$ if and only if g maximizes the entropy at (ϵ, τ_2). In particular, if g maximizes the entropy at $(1/2, \tau_2)$, then so does g'. To show that $s(g)$ has a non-unique maximizer along the upper part of the $\epsilon = 1/2$ line, we must only show that a maximizer g is not related to its mirror g' by a measure-preserving transformation of $[0, 1]$.

As noted in Section 4, up to such a transformation there are exactly two graphons corresponding to $(\epsilon, \tau_2) = (1/2, 1/(2\sqrt{2}))$, namely a g-anticlique g_a and a g-clique g_c. These are not related by reordering, since the values of the degree function for the g-clique are $\sqrt{2}/2$ and 0, while those for the g-anticlique are 1 and $1 - \sqrt{2}/2$. Let D be smallest of the following distances in the cut metric: (1) from g_a to g_c, (2) from g_a to the set of symmetric graphons, and (3) from g_c to the set of symmetric graphons.

Lemma 5.2. *There exists* $\delta > 0$ *such that every graphon with* (ϵ, τ_2) *within* δ *of* $(1/2, 2^{-3/2})$ *is within* $D/3$ *of either* g_a *or* g_c*.*

Proof. Suppose otherwise. Then we could find a sequence of graphons with (ϵ, τ_2) converging to $(1/2, 2^{-3/2})$ that have neither g_a nor g_c as an accumulation point. However, the space of reduced graphons is known to be compact [Lov], so there must be some accumulation point g_∞ that is neither g_a nor g_c. Since convergence in the cut metric implies convergence of the density of all subgraphs, $t_1(g_\infty) = 1/2$ and $t_2(g_\infty) = 2^{-3/2}$. But this contradicts the fact that only g_a and g_c have edge and 2-star densities $(1/2, 2^{-3/2})$. □

By the lemma, no graphon with $\epsilon = 1/2$ and $\tau_2 > 1/2\sqrt{2} - \delta$ is invariant (up to reordering) under $g \to 1 - g$. In particular, the entropy maximizers cannot be symmetric, so there must be two (or more) entropy maximizers, one close to g_a and one close to g_c. □

Moreover, on a path in the parameter space from the anticlique on the upper boundary at $\epsilon = \frac{1}{2} - \delta$ to the clique on the upper boundary at $\epsilon = 1/2 + \delta$, there is a discontinuity in the graphon, where it jumps from being close to g_a to being close to g_c. There must be an odd number of such jumps, and if the path is chosen to be symmetric with respect to the transformation $\epsilon \to 1 - \epsilon$, $\tau_2 \to \tau_2 + 1 - 2\epsilon$, the jump points must be arranged symmetrically on the path. In particular, one of the jumps must be at exactly $\epsilon = 1/2$. This shows that the $\epsilon = 1/2$ line forms the boundary between a region where the optimal graphon is close to g_a and another region where the optimal graphon is close to g_c.

6 Simulations

We now show some numerical simulations in the 2-star model ($\ell = 2, k_1 = 1, k_2 = 2$). Our main aim here is to present numerical evidence that the maximizing graphons in this case are in fact *bipodal*, and to clarify the significance of the degeneracy of Theorem 5.1.

To find maximizing K-podal graphons, we partition the interval $[0, 1]$ into K subintervals $\{I_i\}_{i=1,\dots,k}$ with lengths $c_1, c_2, \cdots, c_K$, that is, $I_i = [c_0+\cdots+c_{i-1}, c_0+\cdots+c_i]$ (with $c_0 = 0$). We form a partition of the square $[0, 1]^2$ using the product of this partition with itself. We are interested in functions g that are piecewise constant on the partition:

$$g(x, y) = g_{ij}, \quad (x, y) \in I_i \times I_j, \quad 1 \le i, j \le K, \tag{18}$$

with $g_{ij} = g_{ji}$. We can then verify that the entropy density $s(g)$, the edge density $t_1(g)$ and the 2-star density $t_2(g)$ become respectively

$$s(g) = -\frac{1}{2} \sum_{1\le i,j\le K} [g_{ij} \log g_{ij} + (1 - g_{ij}) \log(1 - g_{ij})]c_i c_j, \tag{19}$$

$$t_1(g) = \sum_{1\le i,j\le K} g_{ij} c_i c_j, \qquad t_2(g) = \sum_{1\le i,j,k\le K} g_{ik} g_{kj} c_i c_j. \tag{20}$$

Our objective is to solve the following maximization problem:

$$\max_{\{c_j\}_{1\le j\le K}, \{g_{i,j}\}_{1\le i,j\le K}} s(g), \quad \text{subject to:} \quad t_1(g) = \epsilon, \quad t_2(g) = \tau_2, \quad \sum_{1\le j\le K} c_j = 1, \quad g_{ij} = g_{ji}. \tag{21}$$

We developed in [RRS] computational algorithms for solving this maximization problem and have benchmarked the algorithms with theoretically known results. For a fixed $\tau \equiv (\epsilon, \tau_2)$, our strategy is to first maximize for a fixed number K, and then maximize over the number K. Let $s^K_{(\epsilon,\tau_2)}$ be the maximum achieved by the graphon g_K, then the maximum of the original problem is $s_{(\epsilon,\tau_2)} = \max_K\{s^K_{(\epsilon,\tau_2)}\}$. Our computational resources allow us to go up to $K = 16$ at this time. See [RRS] for more details on the algorithms and their benchmark with existing results.

The most important numerical finding in this work is that, for every pair (ϵ, τ_2) in the interior of the phase space, the graphons that maximize $s(g)$ are *bipodal*. We need only four parameters (c_1, g_{11}, g_{12} and g_{22}) to describe bipodal graphons (due to the fact that $c_2 = 1-c_1$ and $g_{12} = g_{21}$). For maximizing bipodal graphons, we need only three parameters, since (9) implies that

$$\left(\frac{1}{g_{11}} - 1\right)\left(\frac{1}{g_{22}} - 1\right) = \left(\frac{1}{g_{12}} - 1\right)^2, \tag{22}$$

which was used in our numerical algorithms to simplify the calculations.

We show in Fig. 2 maximizing graphons at some typical points in the phase space. The (ϵ, τ_2) pairs for the plots are respectively: $(0.3, 0.16844286)$ and $(0.3, 0.10339268)$ for the first

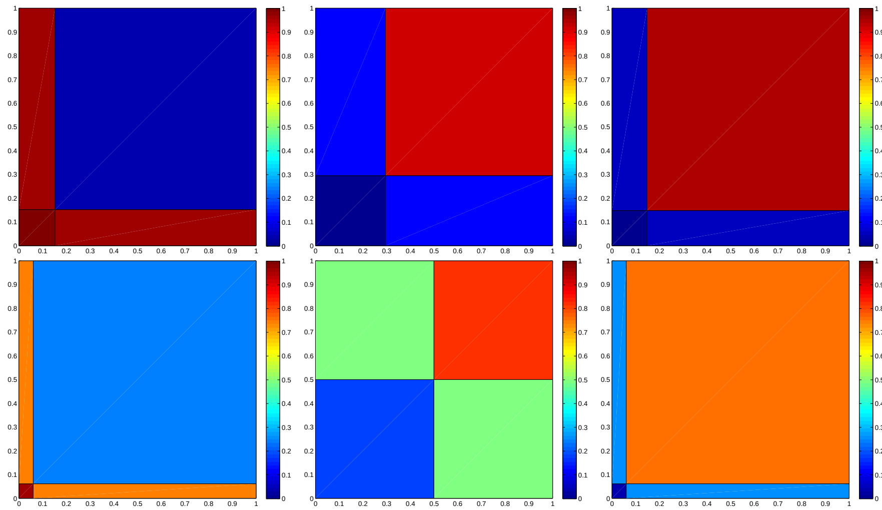

Figure 2: Maximizing graphons at $\epsilon = 0.3$ (left column), $\epsilon = 0.5$ (middle column) and $\epsilon = 0.7$ (right column). For each column τ_2 values decrease from top to bottom.

column (top to bottom), $(0.5, 0.32455844)$ and $(0.5, 0.27485281)$ for the second column, and $(0.7, 0.56270313)$ and $(0.7, 0.50339268)$ for the third column.

The values of s corresponding to the maximizing graphons are shown in the left plot of Fig. 3 for a fine grid of (ϵ, σ^2) (with $\sigma^2 = \tau_2 - \epsilon^2$ as defined in Fig. 1) pairs in the phase space. We first observe that the plot is symmetric with respect to $\epsilon = 1/2$. The symmetry comes from the fact (see the proof of Theorem 5.1) that the map $g \to 1 - g$ takes $\epsilon \to 1 - \epsilon$, $\tau_2 \to 1 - 2\epsilon + \tau_2$ and thus $\sigma^2 \to \sigma^2$. To visualize the landscape of s better in the phase space, we also show the cross-sections of $s_{(\epsilon,\tau_2)}(\epsilon, \sigma^2)$ along the lines $\epsilon_k = 0.05k$, $k = 7, \cdots, 13$, in the right plots of Fig. 3.

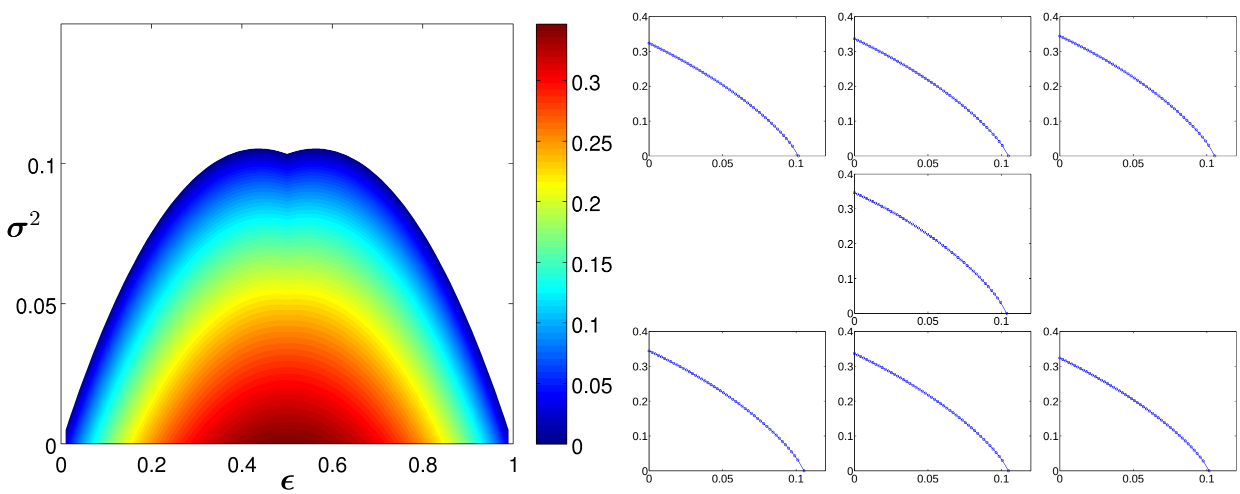

Figure 3: Left: values of $s_{(\epsilon,\tau_2)}$ at different (ϵ, σ^2) pairs; Right: cross-sections of $s_{(\epsilon,\tau_2)}(\epsilon, \sigma^2)$ along lines $\epsilon = \epsilon_k = 0.05k$ $(k = 7, \cdots, 13)$ (from top left to bottom right).

We show in the left plot of Fig. 4 the values of c_1 of the maximizing graphons as a function of the pair (ϵ, σ^2). Here we associate c_1 with the set of vertices among V_1 and V_2 that has the larger probability of an interior edge. This is done to avoid the ambiguity caused by the fact that one can relabel V_1, V_2 and exchange c_1 and c_2 to get an equivalent graph with the same ϵ, τ_2 and s values. We again observe the symmetry with respect to $\epsilon = 1/2$. The cross-sections of $c_1(\epsilon, \sigma^2)$ along the lines of $\epsilon_k = 0.05k$ $(k = 7, \cdots, 13)$ are shown in the right plots of Fig. 4.

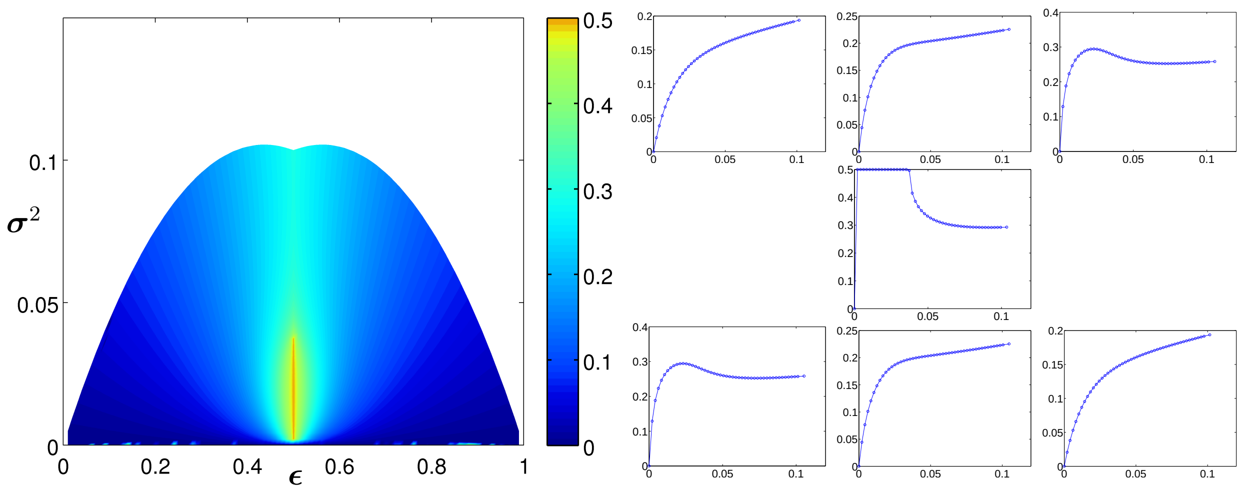

Figure 4: Left: c_1 of the maximizing bipodal graphons as a function of (ϵ, σ^2); Right: cross-sections of $c_1(\epsilon, \sigma^2)$ along lines of $\epsilon_k = 0.05k$ $(k = 7, \cdots, 13)$.

The last set of numerical simulations were devoted to the study of a phase transition in the 2-star model. The existence of this phase transition is suggested by the degeneracy in Theorem 5.1. Our numerical simulations indicate that the functions differ to first order in $\epsilon - 1/2$, and that the actual entropy $s_{(\epsilon,\tau_2)} = \max\{s^L{}_{(\epsilon,\tau_2)}, s^R_{(\epsilon,\tau_2)}\}$ has a discontinuity in $\partial_\epsilon s_{(\epsilon,\tau_2)}$ at $\epsilon = 1/2$ above a critical value τ_2^c. Below τ_2^c, there is a single maximizer, of the form

$$g(x,y) = \begin{cases} \frac{1}{2} + \nu & x, y < \frac{1}{2} \\ \frac{1}{2} - \nu & x, y > \frac{1}{2} \\ \frac{1}{2} & \text{otherwise} \end{cases} \tag{23}$$

Here ν is a parameter related to τ_2 by $\tau_2 = 1/4 + \nu^2/4$. Applying the symmetry $g \to 1 - g$ and reordering the interval $[0, 1]$ by $x \to 1 - x$ sends g to itself.

The critical point τ_2^c is located on the boundary of the region in which the maximizer (23) is stable. The value of τ_2^c can be found by computing the second variation of $s(g)$ within the space of bipodal graphons with fixed values of $(\epsilon = \frac{1}{2}, \tau_2)$, evaluated at the maximizer (23). This second variation is positive-definite for ν small (*i.e.* for τ_2 close to $1/4$) and becomes indefinite for larger values of ν. At the critical value of τ_2^c, $\nu = 2\sqrt{\tau_2^c - 1/4}$ satisfies

$$\left(2S(\frac{1}{2} - \nu) - 2S(\frac{1}{2}) + 3\nu S'(\frac{1}{2} - \nu)\right)(2 - \frac{1}{2}S''(\frac{1}{2} - \nu)) + 8\nu^2 S''(\frac{1}{2} - \nu) = 0 \tag{24}$$

where S' and S'' are respectively the first and second order derivatives of $S(g)$ (defined in (6)) with respect to g. This equation is transcendental, and so cannot be solved in closed form. Solving it numerically for ν leads to the value $\tau_2^c \approx 0.287$, or $\sigma^2 \approx 0.037$. This agrees precisely with what we previously observed in our simulations of optimizing graphons, and corresponds to the point in the left plot Fig. 4 where the $c_1 = 1/2$ region stops.

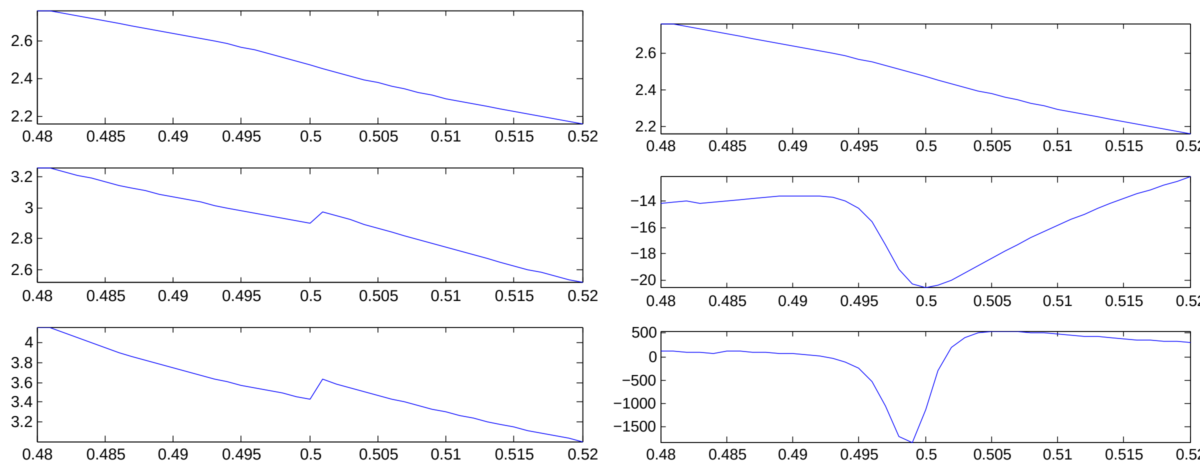

Figure 5: Left: the derivative $\frac{\partial s_{(\epsilon,\tau_2)}}{\partial \epsilon}$ at $\tau_2 = 0.28$ (top), $\tau_2 = 0.30$ (middle) and $\tau_2 = 0.32$ (bottom) in the neighborhood of $\epsilon = 0.5$; Right: the derivatives $\frac{\partial s_{(\epsilon,\tau_2)}}{\partial \epsilon}$ (top), $\frac{\partial^2 s_{(\epsilon,\tau_2)}}{\partial \epsilon^2}$ (middle) and $\frac{\partial^3 s_{(\epsilon,\tau_2)}}{\partial \epsilon^3}$ (bottom) in the neighborhood of $\epsilon = 0.5$ for $\tau_2 = 0.28$.

In the left plot of Fig. 5, we show numerically computed derivatives of $s_{(\epsilon,\tau_2)}$ with respect to ϵ in the neighborhood of $\epsilon = 0.5$ for three different values of τ_2: one below the critical point and two above it. It is clear that discontinuities in the first order derivative of s appears at $\epsilon = 0.5$ for $\tau_2 > \tau_2^c$. When $\tau_2 < \tau_2^c$, we do not observe any discontinuity in the first three derivatives of s.

7 Conclusion

We first compare our results with exponential random graph models (ERGMs) based on the same subgraph densities; see [CD, RY, LZ, Y, YRF, AZ]. For this we focus on the basic optimization problems underlying the two, as follows.

Intuitively the randomness in all such random graph models arises, in modeling large networks, by starting with an assumption that a certain set of subgraphs $H = (H_1, \ldots, H_m)$ are 'significant' for the networks; one can then try to understand a large network as a 'typical' one for certain values $t_H(g) = (t_{H_1}, \ldots, t_{H_m})$ of the densities t_{H_j} of those subgraphs. Large deviations theory can then give probabilistic descriptions of such typical graphs through a variational principle for the constrained entropy, $s_\tau = \sup_{g | t_H(g)=\tau} s(g)$. See [CV].

In this paper, as in [RS1, RS2, RRS], we use such constrained optimization of entropy, and by analogy with statistical mechanics we call such models 'microcanonical'. In contrast, ERGMs are analogues of 'grand canonical' models of statistical mechanics. As noted in Section 2, the microcanonical version consists of maximizing $s(g)$ over graphons g with fixed values $t_H(g) = \tau$. This leads to a constrained-maximum entropy $s_\tau = \sup_{g|t_H(g)=\tau} s(g)$. The optimizing graphons satisfy the Euler-Lagrange variational equation, $\delta[s(g)+\beta\cdot t_H(g)] = 0$, together with the constraints $t_H(g) = \tau$, for some set of Lagrange multipliers $\beta = (\beta_1, \ldots, \beta_m)$.

For the ERGM (grand canonical) approach, instead of fixing $t_H(g)$ one maximizes $F(g) = s(g) + \beta \cdot t_H(g)$ for fixed β, obtaining

$$F_\beta = \sup_g s(g) + \beta \cdot t_H(g). \tag{25}$$

However it is typical for there to be a loss of information in the grand canonical modelling of large graphs. One way to see the loss is by comparing the parameter ("phase") space $\Sigma_{mc} = \{\tau\}$ of the microcanonical model with that for the grand canonical model, $\Sigma_{gc} = \{\beta\}$. For each point β of Σ_{gc} there are optimizing graphons $\tilde{g}_\beta$ such that $F(\tilde{g}_\beta) = F_\beta$, and for each point τ of Σ_{mc} there are optimizing graphons $\tilde{g}_\tau$ such that $t_H(\tilde{g}_\tau) = \tau$ and $s(\tilde{g}_\tau) = s_\tau$. Defining τ' as $t_H(\tilde{g}_\beta)$ it follows that $\tilde{g}_\beta$ maximizes $s(g)$ under some constraint τ, namely $s(\tilde{g}_\beta) = s_{\tau'}$. But the converse fails: there are some τ for which no optimizing $\tilde{g}_\beta$ satisfies $t_H(\tilde{g}_\beta) = \tau$ [CD, RS1].

This asymmetry is particularly acute for star models with $\ell = 2$: it follows from [CD] that all of Σ_{gc} is represented only on the lower boundary curve of Σ_{mc}, $\tau_k = \epsilon^k$: see Fig. 1. If one is interested in the influence of certain subgraph densities in a large network it is therefore preferable to use constrained optimization of entropy rather than to use the ERGM approach.

Next we consider the role of multipodal states in modelling large graphs. In [RS1, RS2, RRS] evidence, but not proof, of multipodal entropy optimizers were found throughout the phase space of the microcanonical triangle model, and here we have proven this to hold in all star models. Consider more general microcanonical graph models with constraints on edge density, $e(g)$, and the densities $t_H(g)$ of a finite number of other subgraphs, H. We make the following conjecture.

Conjecture. The graphons maximizing constrained entropy in such a microcanonical model are always multipodal.

The multipodal structure seems to be an indicator of the emergence of the phases in these networks, as vertex number diverges. Let us explore this suggestion.

In any microcanonical model with constraints on the density of edges and one other subgraph, H, let $k \geq 2$ be the number of edges in H and consider the Erdös-Rényi curve, $\tau_k = \epsilon^k$, in the microcanonical phase space, where the parameter ϵ represents the constraint on edges and τ_k represents the constraint on $t_H(g)$, in the entropy optimization. On this curve there is a general proof [CD] that the optimizing graphons are constant functions on $[0, 1]$, with value ϵ for constraints (ϵ, τ_k). Consider the region in the phase space above the curve, that is, with $\tau_k > \epsilon^k$. There are general proofs [RY, CD] that in the corresponding ERGM models there is a phase transition when the parameter conjugate to the density $t_H(g)$ is positive, adding weight to $t_H(g)$. Theorem 5.1 and the simulations of Section 6 show a transition in the microcanonical 2-star model; it is unclear what, if any, connection it has to that in the corresponding ERGM model.

Question. Is there always a phase transition in a microcanonical model with 2 constraints, above the Erdös-Rényi curve? And if so, does it have any relation to the one in the corresponding ERGM?

By analogy with physics the region above the Erdös-Rényi curve seems to represent a single 'fluid-like' phase, and the question asks if there is always a gas/liquid-like transition in microcanonical models, and if so what is its relation if any to that known in the ERGM.

In the k-star models, *i.e.* star models with two constraints, the Erdös-Rényi curve is the lower boundary of the (microcanonical) phase space, but in the triangle model there is strong evidence of 'solid-like' phases and 'solid/solid transitions' below the curve. Multipodal structure could be a useful tool in understanding the various phases. For instance in the triangle model [RS1, RS2, RRS] even a cursory inspection of the largest values of such a graphon concentrates attention on the conditions under which edges tend to clump together (fluid-like behavior) or push apart into segregated patterns (solid-like behavior). In equilibrium statistical mechanics [Ru] one can rarely understand directly the equilibrium distribution in a useful way, at least away from extreme values of energy or pressure, so one determines the basic characteristics of a model by estimating order parameters or other secondary quantities. In random graph models multipodal structure of the optimizing state gives the hope of more direct understanding of the emergent properties of a model. This would be a significant shift of viewpoint.

Acknowledgments

The authors gratefully acknowledge useful discussions with Mei Yin and references from Miki Simonovits. The computational codes involved in this research were developed and debugged on the computational cluster of the Mathematics Department of UT Austin. The main computational results were obtained on the computational facilities in the Texas Super Computing Center (TACC). We gratefully acknowledge this computational support. This work was partially supported by NSF grants DMS-1208191, DMS-1208941, DMS-1321018 and DMS-1101326.

References

[AK] R. Ahlswede and G.O.H. Katona, Graphs with maximal number of adjacent pairs of edges, Acta Math. Acad. Sci. Hungar. 32 (1978) 97-120

[AZ] D. Aristoff and L. Zhu, On the phase transition curve in a directed exponential random graph model, arxiv:1404.6514

[BCL] C. Borgs, J. Chayes and L. Lovász, Moments of two-variable functions and the uniqueness of graph limits, Geom. Funct. Anal. 19 (2010) 1597-1619.

[BCLSV] C. Borgs, J. Chayes, L. Lovász, V.T. Sós and K. Vesztergombi, Convergent graph sequences I: subgraph frequencies, metric properties, and testing, Adv. Math. 219 (2008) 1801-1851.

[CD] S. Chatterjee and P. Diaconis, Estimating and understanding exponential random graph models, Ann. Statist. 41 (2013) 2428-2461.

[CV] S. Chatterjee and S.R.S. Varadhan, The large deviation principle for the Erdős-Rényi random graph, Eur. J. Comb. 32 (2011) 1000-1017.

[Lov] L. Lovász, Large networks and graph limits, American Mathematical Society, Providence, 2012.

[LS1] L. Lovász and B. Szegedy, Limits of dense graph sequences, J. Combin. Theory Ser. B 98 (2006) 933-957.

[LS2] L. Lovász and B. Szegedy, Szemerédi's lemma for the analyst, GAFA 17 (2007) 252-270.

[LS3] L. Lovász and B. Szegedy, Finitely forcible graphons, J. Combin. Theory Ser. B 101 (2011) 269-301.

[LZ] E. Lubetzky and Y. Zhao, On replica symmetry of large deviations in random graphs, Random Structures and Algorithms (to appear), arXiv:1210.7013 (2012).

[N] M.E.J. Newman, Networks: an Introduction, Oxford University Press, 2010.

[RRS] C. Radin, K. Ren and L. Sadun, The asymptotics of large constrained graphs, J. Phys. A: Math. Theor. 47 (2014) 175001.

[RS1] C. Radin and L. Sadun, Phase transitions in a complex network, J. Phys. A: Math. Theor. 46 (2013) 305002.

[RS2] C. Radin and L. Sadun, Singularities in the entropy of asymptotically large simple graphs, arXiv:1302.3531 (2013).

[Ru] D. Ruelle Statistical Mechanics; Rigorous Results, Benjamin, New York, 1969.

[RY] C. Radin and M. Yin, Phase transitions in exponential random graphs, Ann. Appl. Probab. 23 (2013) 2458-2471.

[TET] H. Touchette, R.S. Ellis and B. Turkington, Physica A 340 (2004) 138-146.

[Y] M. Yin, Critical phenomena in exponential random graphs, J. Stat. Phys. 153 (2013) 1008-1021.

[YRF] M. Yin, A. Rinaldo and S. Fadnavis, Asymptotic quantization of exponential random graphs, arXiv:1311.1738 (2013).