J Stat Phys (2015) 158:853–865
DOI 10.1007/s10955-014-1151-3

Singularities in the Entropy of Asymptotically Large Simple Graphs

Charles Radin · Lorenzo Sadun

Received: 22 July 2014 / Accepted: 13 November 2014 / Published online: 21 November 2014

Abstract We prove that the asymptotic entropy of large simple graphs, as a function of fixed edge and triangle densities, is nondifferentiable along a certain curve. We also determine the precise modified-bipartite structure of asymptotic graphs with edge density 1/2 and triangle density in the interval $[0, 1/8]$.

Keywords Grapon · Extremal graphs · Phase transition · Random graph · Graph limits

Mathematics Subject Classification 05C35, 05C30

1 Introduction

Large networks or graphs can be modeled with statistical mechanics, with edges playing the role of particles in a large volume, and other sub-graphs playing the role of energy. One may then construct a 'microcanonical ensemble' by constraining the density of edges and of some other sub-graphs, and may thereby hope to see the emergence of phases and phase transitions, as graph size diverges, via properties of the entropy. (See [14] for previous work in this vein, and for references to treatments using grand canonical ensembles for graphs, which will be discussed below.) Just as one can sometimes investigate the structure of phases in a phase diagram from the optimal configurations corresponding to the boundary of the phase space (energy ground states), so too one may hope to investigate the phases of large graphs by the graphs corresponding to the analogous phase space boundary. Extremal graph theory [4] deals with graphs in which conflicting graph invariants are on the verge of contradiction. A classic example due to Mantel from 1907 shows that, among graphs of order n, as the edge number increases beyond $\lfloor n^2/4 \rfloor$ a graph can no longer be bipartite and must contain

C. Radin (✉) · L. Sadun
Department of Mathematics, The University of Texas at Austin, Austin, TX 78712, USA
e-mail: radin@math.utexas.edu

L. Sadun
e-mail: sadun@math.utexas.edu

a triangle. Generalizing slightly, the Mantel problem is to determine those graphs with fixed edge density e which minimize, and those which maximize, the possible values t of triangle density. In this vein extremal graph theory is concerned with qualitative features of graphs with invariants on the boundary ∂S of the space S of possible values of some particular set of invariants which, for the Mantel problem, are the edge and triangle densities, e and t. (The set ∂S for the Mantel problem was finally determined in [13], and the optimizing graphs in [12].) In this paper we are concerned with a natural generalization of extremal graph theory to the interior of S. Borrowing an idea from physics, it is possible that qualitative graph features which are forced in an absolute sense on a subset P of the boundary of S are still retained for typical graphs in some phase, a region of S abutting P. (We define 'phase' below and 'typical' in the next section.) For instance for the Mantel problem there is evidence in [14] that for edge density less than 1/2 there is a region of S abutting the interval $(e, t) \in [0, 1/2] \times \{0\}$ of ∂S, in which now a typical graph is nearly bipartite. (The vertices are divided into two clusters of nearly equal size, with nearly all edges connecting vertices in one cluster to vertices in the other.) One of our main results, Theorem 6.2, proves this for the interior interval $(e, t) \in 1/2 \times [0, 1/8]$.

One objective in such a study is 'phase transitions', boundaries between phases in which the competition between invariants which has traditionally been studied on ∂S is extended into the interior of S, and now concerns typical graphs. We study typical graphs using entropy and the graph limit formalism, which we sketch after the following summary of results.

Consider the set $\hat{G}^n$ of simple graphs G with set $V(G)$ of (labeled) vertices, edge set $E(G)$ and triangle set $T(G)$, where the cardinality $|V(G)|$, of $V(G)$, is n. (A graph is 'simple' if the edges are undirected and there are no multiple edges or loops.) We will be concerned with the asymptotics of $\hat{G}^n$ as n diverges, specifically in the relative number of graphs as a function of the cardinalities $|E(G)|$ and $|T(G)|$.

Let $Z^{n,\alpha}_{e,t}$ be the number of graphs in $\hat{G}^n$ such that the edge and triangle densities, $e(g)$ and $t(g)$, satisfy:

$$e(G) \equiv \frac{|E(G)|}{\binom{n}{2}} \in (e-\alpha, e+\alpha) \quad \text{and} \quad t(G) \equiv \frac{|T(G)|}{\binom{n}{3}} \in (t-\alpha, t+\alpha). \tag{1}$$

Graphs g in $\cup_{n\geq 1}\hat{G}^n$ are known to have edge and triangle densities, $(e(g), t(g))$, whose accumulation points form a compact subset R of the (e, t)-plane bounded by three curves, $c_1 : (e, e^{3/2}),\ \ 0 \leq e \leq 1$, the line segment $l_1 :\ (e, 0),\ \ 0 \leq e \leq 1/2$, and a certain scalloped curve $(e, h(e)),\ \ 1/2 \leq e \leq 1$, lying above the curve $(e, e(2e-1)),\ \ 1/2 \leq e \leq 1$, and meeting it when $e = e_k = k/(k+1),\ \ k \geq 1$; see [12,13] and references therein, and Fig. 1. (Note the minor shift in emphasis from S, as discussed earlier, to the accumulation points R of S.)

We are interested in the relative number of graphs with given numbers of edges and triangles, asymptotically in the number of vertices. More precisely we will analyze the entropy density, the exponential rate of growth of $Z^{n,\alpha}_{e,t}$ as a function of n. First consider

$$s^{n,\alpha}_{e,t} = \frac{\ln(Z^{n,\alpha}_{e,t})}{n^2}, \text{ and } s(e,t) = \lim_{\alpha\downarrow 0}\lim_{n\to\infty} s^{n,\alpha}_{e,t}. \tag{2}$$

The limits defining the entropy density $s(e, t)$ are proven to exist in [14]. The objects of interest for us are the qualitative features of $s(e, t)$ in the interior of R. In particular, a phase is commonly defined as a maximal connected open subset in which the entropy density is analytic [15]. Our other main result is

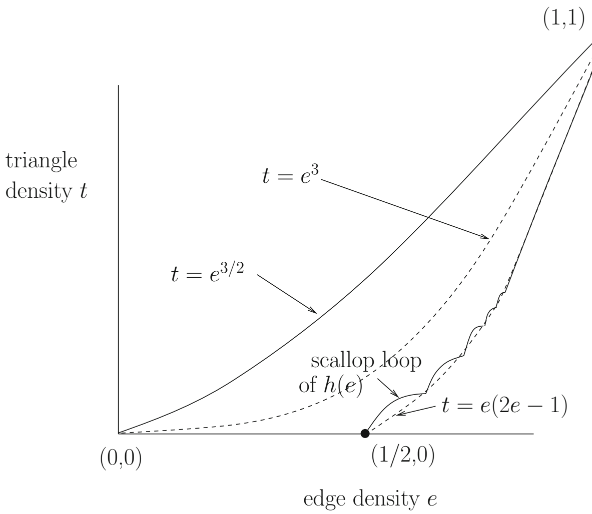

Fig. 1 The phase space R, outlined in solid lines

Theorem 1.1 *In the interior of its domain R the entropy density $s(e,t)$ satisfies:*

$$s(e, e^3) - s(e,t) \geq c|t - e^3| \tag{3}$$

for some $c = c(e) > 0$. Therefore for fixed e, $s(e,t)$ attains its maximum at $t = e^3$ but is not differentiable there. For $t < e^3$ we have the stronger inequality

$$s(e, e^3) - s(e,t) \geq \tilde{c}|t - e^3|^{\frac{2}{3}}. \tag{4}$$

for some $\tilde{c} = \tilde{c}(e) > 0$.

So the graph of $s(e,t)$ has its maxima, varying t for fixed e, on a sharp crease at the curve $t = e^3$, $0 < e < 1$, and is not concave for t just below e^3. The importance of the result lies in the implication from (3) of the lack of differentiability of $s(e,t)$ on the crease, and thus the existence of a phase transition, and the implication from (4) of a lack of concavity of $s(e,t)$, discussed below.

The following bounds on the entropy indicate that the exponents in Theorem 1.1 are sharp:

Theorem 1.2 *For $t > e^3$ the entropy density $s(e,t)$ satisfies:*

$$s(e, e^3) - s(e,t) \leq c'|t - e^3| \tag{5}$$

for some $c' = c'(e) > 0$. For $t < e^3$ we have the weaker inequality

$$s(e, e^3) - s(e,t) \leq \tilde{c}'|t - e^3|^{\frac{2}{3}}. \tag{6}$$

for some $\tilde{c}' = \tilde{c}'(e) > 0$.

We begin with a quick review of the formalism of graph limits, as recently developed in [2,3,8–10]; see also the recent book [7]. The main value of this formalism here is that one can use large deviations on graphs with independent edges [6] to give an optimization formula for $s(e,t)$ [14].

2 Graphons

Consider the set $\mathcal{W}$ of all symmetric, measurable functions

$$g : (x, y) \in [0, 1]^2 \to g(x, y) \in [0, 1]. \tag{7}$$

Think of each axis as a continuous set of vertices of a graph. For a graph $G \in \hat{G}^n$ one associates

$$g^G(x, y) = \begin{cases} 1 & \text{if } (\lceil nx \rceil, \lceil ny \rceil) \text{ is an edge of } G \\ 0 & \text{otherwise,} \end{cases} \tag{8}$$

where $\lceil y \rceil$ denotes the smallest integer greater than or equal to y. For $g \in \mathcal{W}$ and simple graph H we define

$$t(H, g) \equiv \int_{[0,1]^\ell} \prod_{(i,j) \in E(H)} g(x_i, x_j)\, dx_1 \cdots dx_\ell, \tag{9}$$

where $\ell = |V(H)|$, and note that for a graph G, $t(H, g^G)$ is the density of graph homomorphisms $H \to G$:

$$\frac{|\mathrm{hom}(H, G)|}{|V(G)|^{|V(H)|}}. \tag{10}$$

We define an equivalence relation on $\mathcal{W}$ as follows: $f \sim g$ if and only if $t(H, f) = t(H, g)$ for every simple graph H. Elements of $\mathcal{W}$ are called "graphons", elements of the quotient space $\tilde{\mathcal{W}}$ are called "reduced graphons", and the class containing $g \in \mathcal{W}$ is denoted $\tilde{g}$. Equivalent functions in $\mathcal{W}$ differ by a change of variables in the following sense. Let Σ be the space of measure-preserving maps $\sigma : [0, 1] \to [0, 1]$, and for f in $\mathcal{W}$ and $\sigma \in \Sigma$, let $f_\sigma(x, y) \equiv f(\sigma(x), \sigma(y))$. Then $f \sim g$ if and only if there exist σ, σ' in Σ such that $f_\sigma = g_{\sigma'}$ almost everywhere; see Cor. 2.2 in [2]. The space $\mathcal{W}$ is compact with respect to the 'cut metric' defined as follows. First, on $\mathcal{W}$ define:

$$d_\blacksquare(f, g) \equiv \sup_{S,T \subseteq [0,1]} \left| \int_{S \times T} [f(x, y) - g(x, y)]\, dx dy \right|. \tag{11}$$

Then on $\tilde{\mathcal{W}}$ define the cut metric by:

$$\tilde{d}_\blacksquare(\tilde{f}, \tilde{g}) \equiv \inf_{\sigma, \sigma' \in \Sigma} d_\blacksquare(f_\sigma, g_{\sigma'}). \tag{12}$$

We will use the fact, which was shown in [3], that the cut metric is equivalent to the metric

$$\delta_{\mathrm{hom}}(\tilde{f}, \tilde{g}) \equiv \sum_{j \geq 1} \frac{1}{2^j} |t(H_j, f) - t(H_j, g)|, \tag{13}$$

where $\{H_j\}$ is a countable set of simple graphs, one from each graph-equivalence class.

Also note that if each vertex of a finite graph is split into the same number of 'twins', each connected to the same vertices, the result stays in the same equivalence class, so for a convergent sequence $\tilde{g}^{G_j}$ one may assume $|V(G_j)| \to \infty$.

The following was proven in [14], adapting a proof in [5]. These proofs are based on a fundamental large deviation result in [6].

Theorem 2.1 *([14]) For any possible pair* (e, t), $s(e, t) = \max[-I(g)]$, *where the maximum is over all graphons* g *with* $e(g) = e$ *and* $t(g) = t$, *where*

$$e(g) = \int_{[0,1]^2} g(x, y)\,dxdy, \qquad t(g) = \int_{[0,1]^3} g(x, y)g(y, z)g(z, x)\,dxdydz \tag{14}$$

and the rate function is

$$I(g) = \int_{[0,1]^2} I_0(g(x, y))\,dxdy, \text{ where } I_0(u) = \frac{1}{2}\left[u\ln(u) + (1-u)\ln(1-u)\right]. \tag{15}$$

3 Proof of Theorems 1.1 and 1.2

Proof of Theorem 1.1. Fix a graphon g with edge density e. We can always write such a graphon as $g = g_e + \delta g$ where g_a is the constant function on $[0, 1]^2$ with value a. We then compute

$$\delta t(g) := t(g) - e^3 = 3e^2 \int_{[0,1]^2} \delta g(x, y)\,dxdy + 3e \int_{[0,1]^3} \delta g(x, y)\delta g(y, z)\,dxdydz \\ + \int_{[0,1]^3} \delta g(x, y)\delta g(y, z)\delta g(z, x)\,dxdydz. \tag{16}$$

The first term on the right hand side is zero, since $\int_{[0,1]^2} \delta g(x, y)\,dxdy = \delta e = 0$. If we think of δg as the integral kernel of the Hermitian trace class operator $T_{\delta g}$ on $L^2([0, 1])$, then using the inner product $\langle \cdot, \cdot \rangle$ and trace Tr we can rewrite the remaining terms as

$$\delta t = 3e\langle \phi_1, T_{\delta g}^2 \phi_1\rangle + Tr(T_{\delta g}^3), \tag{17}$$

where $\phi_1(x) = 1$ is the constant function on $[0, 1]$. Note that the first term is non-negative. Using again the fact that $\int_{[0,1]^2} \delta g(x, y)\,dxdy = 0$,

$$\begin{aligned} \delta I &= \int_{[0,1]^2} [I_0(e + \delta g(x, y)) - \delta g(x, y)I_0'(e) - I_0(e)]\,dxdy \\ &= \int_{[0,1]^2} \frac{I_0(e + \delta g(x, y)) - \delta g(x, y)I_0'(e) - I_0(e)}{\delta g(x, y)^2} \delta g(x, y)^2\,dxdy \\ &\geq f_-(e) \int_{[0,1]^2} \delta g(x, y)^2\,dxdy, \end{aligned} \tag{18}$$

where

$$f(e, x) = \frac{I_0(e + x) - xI_0'(e) - I_0(e)}{x^2}, \tag{19}$$

and $f_-(e) = \inf_x f(e, x)$. The positivity of $f(e, x)$ follows from the concavity of I_0. Since $\lim_{x\to 0} f(e, x) = \frac{I_0''(e)}{2} = \frac{1}{4e(1-e)}$, $f_-(e)$ is a positive number less than or equal to $\frac{1}{4e(1-e)}$. □

Lemma 3.1 $|Tr(T_{\delta g}^3)| \leq (Tr(T_{\delta g}^2))^{3/2}$, *with equality if and only if* $T_{\delta g}$ *is a rank 1 operator.*

Proof Since $T_{\delta g}$ is an Hermitian trace class operator it has pure discrete spectrum. If $\{\mu_i\}$ are the eigenvalues of $T_{\delta g}$, then

$$|Tr(T^3_{\delta g})| = |\sum_i \mu_i^3| \leq \sum_i |\mu_i^3| \leq \max_j |\mu_j| \sum_i \mu_i^2 \leq (\sum_i \mu_i^2)^{3/2} = [Tr(T^2_{\delta g})]^{3/2}. \tag{20}$$

If $T_{\delta g}$ has rank one, then $Tr(T^3_{\delta g}) = \mu^3 = \pm[Tr(T^2_{\delta g})]^{3/2}$. If $T_{\delta g}$ has rank bigger than 1, then $\max_j(\mu_j)$ is strictly smaller than $\sqrt{\sum_i \mu_i^2}$.

We next give an estimate for $I(g)$ when $t < e^3$. If $\delta t < 0$, then

$$-\delta t = -Tr(T^3_{\delta g}) - 3e\langle\phi_1, T^2_{\delta g}\phi_1\rangle \leq -Tr(T^3_{\delta g}) \leq [Tr(T^2_{\delta g})]^{3/2} \leq \left(\frac{\delta I}{f_-(e)}\right)^{3/2}. \tag{21}$$

This implies that

$$\delta I \geq f_-(e)(-\delta t)^{2/3}. \tag{22}$$

Using $|\delta t| \leq e^3$ this also implies a linear estimate

$$\delta I \geq \frac{f_-(e)}{e}|\delta t| \tag{23}$$

for $\delta t < 0$.

Finally, we estimate $I(g)$ when $t > e^3$. Since $\langle\phi_1, T^2_{\delta g}\phi_1\rangle \leq Tr(T^2_{\delta g})$, and since $Tr(T^2_{\delta g}) \leq 1$, we have

$$\begin{aligned}\delta t &\leq Tr(T^3_{\delta g}) + 3eTr(T^2_{\delta g}) \leq (Tr(T_{\delta g})^2)^{3/2} + 3eTr(T^2_{\delta g}) \leq (3e+1)Tr(T^2_{\delta g}) \\ &\leq \frac{(3e+1)\delta I}{f_-(e)},\end{aligned} \tag{24}$$

so

$$\delta I \geq \frac{f_-(e)\delta t}{3e+1}. \tag{25}$$

□

Proof of Theorem 1.2. To prove Theorem 1.2 we only have to exhibit graphons such that $I(g) - I_0(e) < c'(t - e^3)$ for $t > e^3$ and another graphon such that $I(g) - I_0(e) < \tilde{c}'(t - e^3)^{3/2}$ for $t < e^3$. For the first, consider the graphon

$$g(x, y) = \begin{cases} e - \epsilon & x, y < \frac{1}{2} \\ e & x < \frac{1}{2} < y \text{ or } y < \frac{1}{2} < x \\ e + \epsilon & x, y > \frac{1}{2} \end{cases} \tag{26}$$

We compute

$$I(g) - I_0(e) = \frac{1}{4}I_0''(e)\epsilon^2 + O(\epsilon^4),$$

while

$$t - e^3 = \frac{3}{4}\epsilon^2 e,$$

so

$$I(g) - I_0(e) = \frac{I_0''(e)}{3e}(t - e^3) + O((t - e^3)^2).$$

The graphon (26) is not believed to be optimal, except at $e = 1/2$. However, this graphon gives an upper bound for the optimal $I(g)$, and hence a lower bound for the entropy.

For $t < e^3$, consider graphons of the form

$$\tilde{g}(x, y) = \begin{cases} e - \epsilon & x, y < \frac{1}{2} \text{ or } x, y > \frac{1}{2}, \\ e + \epsilon & x < \frac{1}{2} < y \text{ or } y < \frac{1}{2} < x, \end{cases}$$

with $\epsilon > 0$. Then

$$I(\tilde{g}) - I_0(e) = \frac{1}{2} I_0''(e) + O(\epsilon^4)$$

and

$$t = e^3 - \epsilon^3,$$

so

$$I(\tilde{g}) - I_0(e) = \frac{1}{2} I_0''(e)(e^3 - t)^{2/3} + O((e^3 - t)^{4/3}.$$

The graphon $\tilde{g}$ is provably optimal for $e = 1/2$ (see Theorem 6.2 below), and is believed optimal for all $e < 1/2$. However, as with the $t > e^3$ case, optimality is not needed for the current proof, as we are only seeking lower bounds for the entropy.

Since $s(e, t) - s(e, e^3)$ is bounded below by an appropriate power of $|t - e^3|$ for t close to e^3, and since t takes values in the finite interval $[0, e^{3/2}]$, we can find positive constants c' and $\tilde{c}'$ such that $s(e, t) - s(e, e^3)$ is bounded below by $-c'(t - e^3)$ or $-\tilde{c}'|e^3 - t|^{2/3}$ for all values of t. □

4 Other Graph Models

We now generalize Theorem 1.1 to graph models where we keep track of the number of graph homomorphisms $H \to G$ for some fixed graph H, not necessarily triangles. We can compute the entropy of graphs with $e(g^G)$ within α of e and $t(H, g^G)$ within α of t, and define the entropies $s_{e,t}^{n,\alpha}$ and $s(e, t)$ exactly as in equation (2). The proof of Theorem 2.1 carries over almost word-for-word to show the following.

Theorem 4.1 *For any possible pair (e, t), $s(e, t) = \max[-I(g)]$, where the maximum is over all graphons g with $e(g) = e$ and $t(H, g) = t$.*

Note that if H has k edges the constant graphon g_e satisfies $t(H, g_e) = e^k$.

Theorem 4.2 *For fixed $0 < e < 1$ the entropy density $s(e, t)$ achieves its maximum at $t = e^k$ and is not differentiable with respect to t at that point.*

Proof Following the proof of Theorem 1.1, we write $g = g_e + \delta g$ and expand both $I(g)$ and $t(H, g)$ in terms of δg. The estimate (18) still applies. The only difference is the expansion of $t(H, g)$.

Since $t(H, g)$ is the integral of a polynomial expression in g, we can expand δt as a polynomial in δg. This must take the form

$$\delta t = \int_{[0,1]^2} h_1(x,y)\delta g(x,y)\,dxdy + \int_{[0,1]^4} h_2(w,x,y,z)\delta g(w,x)\delta g(y,z)\,dwdxdydz \\ + \int_{[0,1]^3} h_3(x,y,z)\delta g(x,y)\delta g(y,z)\,dxdydz + \cdots, \tag{27}$$

where the non-negative functions $h_1(x, y)$, $h_2(w, x, y, z)$, $h_3(x, y, z)$, etc., are computed from the graphon from which we are perturbing. However, that graphon is a constant g_e, so each function h_i is also a constant. Thus there are non-negative constants $c_1, c_2, \ldots$, such that

$$\delta t = c_1 \int_{[0,1]^2} \delta g(x,y)\,dxdy + c_2 \int_{[0,1]^4} \delta g(w,x)\delta g(y,z)\,dwdxdydz \\ + c_3 \int_{[0,1]^3} \delta g(x,y)\delta g(y,z)\,dxdydz + \cdots \tag{28}$$

The first two terms integrate to zero. All subsequent terms are bounded by a multiple of $Tr(T^2_{\delta g})$, since we can take the absolute value of the integrand and replace all but two copies of δg with the constant 1. (Note that $Tr(T^2_{|\delta g|} = Tr(T^2_{\delta g}) = \iint g(x,y)^2 dxdy.$)

Since there are only a finite number of terms, $|\delta t|$ is bounded above by a constant multiple of $Tr(T^2_{\delta g})$ while δI is bounded below by a constant multiple of $Tr(T^2_{\delta g})$. Combining these observations yields the analog of Theorem 1.1, and we conclude that $s(e, t)$ cannot have a 2-sided derivative with respect to t at $t = e^k$. □

A more careful analysis of the terms in the sum (28) shows that each term is either positive-definite, is dominated by a positive-definite term, or scales as $Tr(T^2_{\delta g})^{3/2}$ or higher, implying that the concavity of $s(e, t)$ just below the curve $t = e^k$ is the same as for the triangle model. However, this analysis is not needed for the proof of Theorem 4.2 and has been omitted.

There do exist some graphs H, such as "k-stars" with k edges and one vertex on all of them, such that the lowest value of t for fixed e is on the 'Erdős-Rényi curve', $t = e^k,\ 0 < e < 1$. For such graphs the analysis of what happens for $\delta t < 0$ is moot and $s(e, t)$ may have a 1-sided derivative at (e, e^k).

5 Legendre Transform and Exponential Random Graphs

We return temporarily to the special case in which H is a triangle. Note that it has been fundamental to our analysis to use the optimization characterization of $s(e, t)$ of Theorem 2.1 (Theorem 3.1 in [14]). Treating this as an optimization with constraints, one might naturally introduce Lagrange multipliers β_1, β_2 and consider the following optimization system,

$$\max_g[-I(g) + \beta_1 e(g) + \beta_2 t(g)];\ \ e(g) = e;\ \ t(g) = t, \tag{29}$$

namely maximize

$$\Psi_{\beta_1,\beta_2}(g) = -I(g) + \beta_1 e(g) + \beta_2 t(g) \tag{30}$$

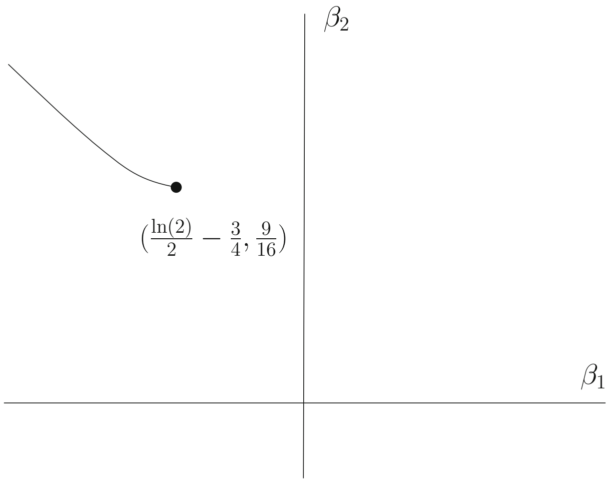

Fig. 2 The curve of all singularities of $\psi(\beta_1, \beta_2)$, for $\beta_2 > -1/2$

for fixed (β_1, β_2) and then adjust (β_1, β_2) to achieve the desired values of $e(g)$ and $t(g)$. The 'free energy density'

$$\psi(\beta_1, \beta_2) = \max_g \Psi_{\beta_1,\beta_2}(g) \tag{31}$$

is directly related to the normalization in exponential random graph models and basic information in such models is simply obtainable from it [1,5,11,15]. It can be considered the Legendre transform of $s(e, t)$, but since the domain of $s(e, t)$ is not convex, the relationship between $s(e, t)$ and $\psi(\beta_1, \beta_2)$ must be more complicated than is common for Legendre transforms. In particular, although it has been proven ([5,15]) that $\psi(\beta_1, \beta_2)$ has singularities as a function of (β_1, β_2) (see Fig. 2) it does not seem straightforward to use this to prove singularities in $s(e, t)$. This is what necessitated the different approach we have taken here. We will try to clarify the relationship between $\psi(\beta_1, \beta_2)$ and $s(e, t)$ through differences in the optimization characterizations of these quantities.

As one crosses the curve in Fig. 2 by increasing β_2 at fixed β_1, the unique graphon maximizing $\Psi_{\beta_1,\beta_2}(g)$ jumps from lower to higher value of $e(g)$, but still $t(g) = e(g)^3$ [5,15]. We emphasize that whenever $\beta_2 > -1/2$, one is on the Erdős-Rényi curve $t = e^3$ indicated in Fig. 1 [5,15]. This is significant in interpreting the singularities of $s(e, t)$ and $\psi(\beta_1, \beta_2)$. The singularities or 'transition' characterized in Theorem 1.1 and associated with *crossing* the Erdős-Rényi curve is presumably between graphs of different character but similar densities; we expect that graphons maximizing $s(e, t)$, for $t > e^3$, are related to those (discussed below) for the upper boundary of its domain R, while for $t < e^3$ they are related to those for the lower boundary of R. (The latter are the subject of [1,14].) On the other hand, the transition in Fig. 2, associated with varying (β_1, β_2), is between graphs of similar character (independent edges) but different densities. This phenomenon is unrelated to the transition of Theorem 1.1, although still associated with the Erdős-Rényi curve, and which we understand as follows.

Assume one optimizes $\Psi_{\beta_1,\beta_2}(g)$ for fixed (β_1, β_2), where (β_1, β_2) is adjusted so that maximizing graphons g satisfy $e(g) = e$ and $t(g) = t$ to match the desired values of (e, t) in which we are interested for $s(e, t)$. It may happen that for special (β_1, β_2) there are also optimizing g with other densities, $(e(g), t(g)) \neq (e, t)$. This degeneracy is what is occurring precisely for the (β_1, β_2) on the singularity curve of Fig. 2. All such g clearly solve the maximization problem for $s[e(g), t(g)]$; they are appearing together when we fix (β_1, β_2) because the value of $\Psi_{\beta_1,\beta_2}(g)$ happens to be the same for all these g, a phenomenon of no particular relevance to the original optimization problem of $s(e, t)$. So in this sense degenerate solutions associated with the Legendre transform can be misleading; they point to a 'transition' which is foreign to the maximization problem for $s(e, t)$. We next consider other consequences from using the Legendre transform.

One issue of importance to those who study exponential random graph models is that for no β_2 is there a maximizer g of the free energy density Ψ satisfying $t(g) > e(g)^3$, though there clearly are such optimizers of the entropy density s as we see for instance from Fig. 1.

Theorem 5.1 *For every β_2 and every maximizer g of $\Psi(g)$, $t(g) \leq e(g)^3$.*

Proof Suppose the graphon g' satisfies $t(g') > [e(g')]^3$ and maximizes the free energy

$$\Psi(g) = -I(g) + \beta_1 e(g) + \beta_2 t(g), \tag{32}$$

for some β_1 and β_2. It follows from Theorem 4.2 in [5] that $\beta_2 < 0$. Let g_e be the constant graphon with the same edge density as G'. Since $t > e^3$, $\beta_2(g_e) > \beta_2 t(g')$. Also, $-I(g_e) > -I(g')$, since for given edge density $-I(g)$ is maximized at g_e. But then $\Psi(g_e) > \Psi(g')$, and g' is not a maximizer, which is a contradition. □

6 Optimizing Graphons

Having established in Theorem 1.1 a phase transition on the Erdős-Rényi curve, we consider the forms of the graphons that maximize $s(e, t)$ on each side of the curve. We previously [14] determined the optimizing graphons on the lower boundary of the region R, including the scalloped curve. We now compute the optimizing graphons on the upper boundary and on the curve $e = 1/2$ below the Erdős-Rényi line.

6.1 The Upper Boundary

On the upper boundary, the graphons are unique and have zero entropy. The following theorem appears to be well known by the experts, and is included for completeness.

Theorem 6.1 *Any graphon with $t = e^{3/2}$ takes the form*

$$g(x, y) = \begin{cases} 1 & x, y < \sqrt{e} \\ 0 & \textit{otherwise.} \end{cases} \tag{33}$$

up to a measure-preserving transformation.

Proof Let T_g be the operator on $L^2[0, 1]$ with integral kernel g. We already know that $t = Tr(T_g^3) \leq Tr(T_g^2)^{3/2}$, with equality if and only if T_g is rank 1. However,

$$Tr(T_g^2) = \iint g(x, y)g(y, x)dxdy = \iint g(x, y)^2 dxdy \leq \iint g(x, y)dxdy = e, \tag{34}$$

with equality if and only if $g(x, y)^2 = g(x, y)$ almost everywhere, i.e. $g(x, y) = 0$ or 1 almost everywhere.

Combining the two results, we have that $t \le e^{3/2}$, with equality if and only if two conditions are met: $g(x, y) = \alpha(x)\alpha(y)$ for some positive function α, (i.e. T_g has rank one), and $g(x, y)$ is a 0–1 function, implying that $\alpha(x)$ is a 0-1 function.

By applying a measure-preserving transformation to $[0, 1]$ we can assume that α is the characteristic function of an interval $[0, s]$. We then compute $e = s^2$ and $t = e^3$. In short, each point on the upper boundary for the allowed region in the e–t plane is achieved by a unique reduced graphon, namely the equivalence class of the graphon (33). □

6.2 The special case of $e = 1/2$

Theorem 6.2 *When $e = 1/2$ and $t \le e^3$, the graphon*

$$\tilde{g}(x, y) = \begin{cases} 1/2 + \epsilon & x < 1/2 < y \text{ or } x > 1/2 > y \\ 1/2 - \epsilon & x, y < 1/2 \text{ or } x, y > 1/2, \end{cases} \tag{35}$$

where $\epsilon = (e^3 - t)^{1/3}$, maximizes $s(e, t)$. Furthermore, every maximizing graphon is of the form $\tilde{g}_\sigma$ for some measure-preserving transformation σ.

Proof We use perturbation theory, writing $g(x, y) = e + \delta g(x, y)$. When $e = 1/2$, the n^{th} derivative $I_0^{(n)}(x)$ is positive for n even and zero for n odd. This means that $[I_0(e + x) - I_0(e)]/x^2$ is a convex function of x^2 (since it is a power series in x^2 with positive coefficients). This allows us to find a formula for δg that simultaneously maximizes $-Tr(T_{\delta g}^3)$ for fixed $Tr(T_{\delta g}^2)$, minimizes the positive-definite quadratic term in δt (to be zero), and minimizes δI for fixed $Tr(T_{\delta g}^2)$. This must therefore be a minimizer of the rate function and a maximizer of the entropy. We assume throughout that $\iint \delta g(x, y) dx\, dy = 0$. □

Lemma 6.3 *Let $T_{\delta g}$ be a rank-one operator: $T_{\delta g} f = c\langle\alpha, f\rangle\, \alpha$ where $\langle\alpha, \alpha\rangle = 1$. Then $\delta t = c^3$.*

Proof Since $c\langle g_1, \alpha\rangle\langle\alpha, g_1\rangle = \int_{[0,1]^2} \delta g(x, y) = 0$, we must have $\langle g_1, \alpha\rangle = 0$. This makes the quadratic term $3e\langle g_1, T_{\delta g}^2 g_1\rangle$ identically zero. Since $T_{\delta g}$ is rank one with unique eigenvalue c, $\delta t = Tr(T_{\delta g}^3) = c^3$. □

In other words, any rank-1 perturbation of the form $\delta g = c|\alpha\rangle\langle\alpha|$ will give the most negative possible value of δt, with the positive-definite quadratic contribution being zero and the cubic term achieving the bounds of Lemma 3.1.

Now we try to minimize $\int_{[0,1]} I_0[e + \delta g(x, y)]\, dxdy$ for fixed $Tr(T_\delta g^2)$. By convexity, this is minimized when $\delta g(x, y)^2$ is constant. This is clearly the case when the graphon is given by (35), showing that this graphon is in fact a minimizer of the rate function.

To see that this minimizer is unique, we can restrict our attention (by Lemma 3.1) to rank-1 perturbations. Among rank-1 perturbations $\delta g(x, y) = c\alpha(x)\alpha(y)$, so having $\delta g(x, y)^2$ constant means that $\alpha(x)^2$ is constant. Since the integral of α is zero, we must have $\alpha(x) = +1$ on a set of measure 1/2 and -1 on a set of measure 1/2. Up to measure-preserving automorphism, we then have

$$\alpha(x) = \begin{cases} 1 & x > 1/2; \\ -1 & x < 1/2. \end{cases} \tag{36}$$

This means that any graphon that minimizes $I(g)$ for fixed $e = 1/2$ and fixed $t \leq e^3$ must be $\tilde{g}$, up to a measure-preserving transformation.

6.3 Lagrange Multipliers on the $e = 1/2$ line

We proved in Theorem 5.1 that maximizing graphons for $s(e, t)$ for $t > e^3$ cannot be found using the Legendre transform. We now show that this also applies to certain values of $t < e^3$, starting with $e = 1/2$ and t close to $1/8$.

When $e = 1/2$, knowing precisely the optimizing graphon $\tilde{g}$ allows us to compute $s(e, t)$:

$$s\left(\frac{1}{2}, t\right) = \frac{-1}{2}\left[I_0\left(\frac{1}{2}+\epsilon\right) + I_0\left(\frac{1}{2}-\epsilon\right)\right] = -I_0\left(\frac{1}{2}+\epsilon\right), \tag{37}$$

for all $t < 1/8$, since $I_0(u) = I_0(1-u)$.

Now consider the optimization using Lagrange multipliers. The Euler-Lagrange equations are:

$$-I_0'[g(x, y)] + \beta_1 + \beta_2 h(x, y) = 0, \tag{38}$$

where

$$h(x, y) = 3\int_{[0,1]} g(x, z)g(y, z)\, dz \tag{39}$$

is the first variation of $t(g)$ with respect to $g(x, y)$. For our $g = \tilde{g}$, this becomes:

$$\begin{aligned}\beta_1 + 3\beta_2\left(\frac{1}{4}-\epsilon^2\right) &= I_0'\left(\frac{1}{2}+\epsilon\right) = \frac{1}{2}\ln\left[\frac{\frac{1}{2}+\epsilon}{\frac{1}{2}-\epsilon}\right]\\ \beta_1 + 3\beta_2\left(\frac{1}{4}+\epsilon^2\right) &= I_0'\left(\frac{1}{2}-\epsilon\right) = \frac{1}{2}\ln\left[\frac{\frac{1}{2}-\epsilon}{\frac{1}{2}+\epsilon}\right],\end{aligned} \tag{40}$$

which are satisfied if and only if

$$\beta_2 = -\frac{4}{3}\beta_1 = \frac{I_0'\left(\frac{1}{2}-\epsilon\right) - I_0'\left(\frac{1}{2}+\epsilon\right)}{6\epsilon^2} = -\frac{1}{6\epsilon^2}\ln\left[\frac{\frac{1}{2}+\epsilon}{\frac{1}{2}-\epsilon}\right]. \tag{41}$$

Notice that β_1 and β_2 diverge as $\epsilon \downarrow 0$ (equivalently, as $t \uparrow 1/8$).

However, solutions to the Euler-Lagrange equations are not necessarily local maxima of Ψ. It is easy to check by differentiation of (37) that there are $0 < c_1 < c_2 < 1/8$ such that $s(1/2, t)$ is strictly concave on $(0, c_1)$ but strictly convex on $(c_2, 1/8]$. Convexity implies that $\tilde{g}$ is not a maximizer for $\Psi(\beta_1, \beta_2)$ for $c_2 < t < 1/8$, but is rather a local *minimizer* with respect to variation of t, and so there are no (β_1, β_2) for which maximizers of $\Psi(\beta_1, \beta_2)$ are maximizers of $s(1/2, t)$ for t just below $1/8$. While the precise calculation was done for $e = 1/2$ using equation (37), this phenomenon is simply due to inequality (4), and actually occurs for all e, not just for $e = 1/2$. In fact from the proof of Theorem 4.2 this phenomenon can be extended to subgraphs H other than triangles. [However, as noted above, for some H the Erdős-Rényi curve is actually the lower boundary of the domain of the entropy, in which case there are no 'missing' points below it.]

Acknowledgments We gratefully acknowledge useful discussions with Mei Yin. This work was partially supported by NSF Grants DMS-1208941 and DMS-1101326.

References

1. Aristoff, D., Radin, C.: Emergent structures in large networks. J. Appl. Probab. **50**, 883–888 (2013)
2. Borgs, C., Chayes, J., Lovász, L.: Moments of two-variable functions and the uniqueness of graph limits. Geom. Funct. Anal. **19**, 1597–1619 (2010)
3. Borgs, C., Chayes, J., Lovász, L., Sós, V.T., Vesztergombi, K.: Convergent graph sequences I: subgraph frequencies, metric properties, and testing. Adv. Math. **219**, 1801–1851 (2008)
4. Bollobás, B.: Extremal Graph Theory. Academic Press, London (1978)
5. Chatterjee, S., Diaconis, P.: Estimating and understanding exponential random graph models. Ann. Stat. **41**, 2428–2461 (2013)
6. Chatterjee, S., Varadhan, S.R.S.: The large deviation principle for the Erdős-Rényi random graph. Eur. J. Comb. **32**, 1000–1017 (2011)
7. Lovász, L.: Large networks and graph limits. American Mathematical Society, Providence (2012)
8. Lovász, L., Szegedy, B.: Limits of dense graph sequences. J. Combin. Theory Ser. B **96**, 933–957 (2006)
9. Lovász, L., Szegedy, B.: Szemerédi's lemma for the analyst. GAFA **17**, 252–270 (2007)
10. Lovász, L., Szegedy, B.: Finitely forcible graphons. J. Combin. Theory Ser. B **101**, 269–301 (2011)
11. Newman, M.E.J.: Networks: an Introduction. Oxford University Press, Oxford (2010)
12. O. Pikhurko and A. Razborov, Asymptotic structure of graphs with the minimum number of triangles, arXiv:1203.4393.
13. Razborov, A.: On the minimal density of triangles in graphs. Combin. Probab. Comput. **17**, 603–618 (2008)
14. Radin, C., Sadun, L.: Phase transitions in a complex network. J. Phys. A: Math. Theor. **46**, 305002 (2013)
15. Radin, C., Yin, M.: Phase transitions in exponential random graphs. Ann. Appl. Probab. **23**, 2458–2471 (2013)