SINGULARITIES IN THE ENTROPY OF ASYMPTOTICALLY LARGE SIMPLE GRAPHS

CHARLES RADIN AND LORENZO SADUN

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.

1. Introduction

Consider the set $\hat{G}^n$ of simple graphs G with set $V(G)$ of (labelled) vertices, edge set $E(G)$ and triangle set $T(G)$, where the cardinality $|V(G)| = n$. ('Simple' means 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, $(e(g), t(g))$, dense in the 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 [PR] and references therein, and Figure 1.

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 the total entropy $\ln(Z^{n,\alpha}_{e,t})$, as n grows, as follows. 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 [RS]. The objects of interest for us are the qualitative features of $s(e,t)$ in the interior of R. Our main result is:

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

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

2010 *Mathematics Subject Classification.* 05C35, 05C30.

This work was partially supported by NSF grants DMS-1208941 and DMS-1101326.

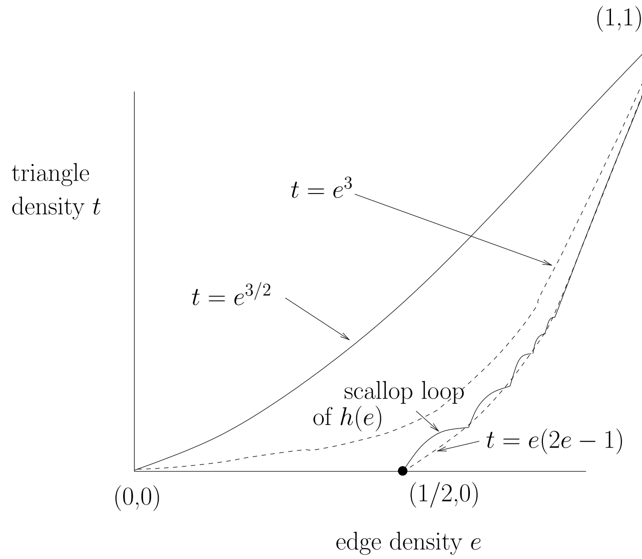

FIGURE 1. The phase space R, outlined in solid lines

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.

We begin with a quick review of the formalism of graph limits; see [BCLSV, LS1, LS2, LS3], and/or the comprehensive treatment in [Lov]. The main value of this formalism here is that one can use large deviations on graphs with independent edges [CV] to give an optimization formula for $s(e,t)$ [RS].

2. Graphons

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

$$g : (x,y) \in [0,1]^2 \to g(x,y) \in [0,1]. \tag{5}$$

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

$$g^G(x,y) = \begin{cases} 1 & \text{if } (\lceil nx \rceil, \lceil ny \rceil) \text{ is an edge of } G \\ 0 & \text{otherwise,} \end{cases} \tag{6}$$

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

$$t(H,g) \equiv \int_{[0,1]^\ell} \prod_{(i,j) \in E(H)} g(x_i,x_j)\, dx_1 \cdots dx_\ell, \tag{7}$$

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

$$\frac{|\mathrm{hom}(H,G)|}{|V(G)|^{|V(H)|}}. \tag{8}$$

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 $\tilde{g}$. The space $\tilde{\mathcal{W}}$ is compact in the metric topology with metric:

$$\delta_{\blacksquare}(\tilde{f}, \tilde{g}) \equiv \sum_{j \geq 1} \frac{1}{2^j} |t(H_j, f) - t(H_j, g)|, \tag{9}$$

where $\{H_j\}$ is a countable set of simple graphs, one from each graph-equivalence class. Equivalent functions in $\mathcal{W}$ differ by a change of variables in the following sense. Let Σ be the space of measure preserving bijections σ of $[0, 1]$, 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 $g = f_\sigma$ for some $\sigma \in \Sigma$. 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 [RS].

Theorem 2.1. *([RS]) 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*

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

and the rate function is

$$I(g) = \int_{[0,1]^2} I_0(g(x, y))\, dxdy, \text{ where } I_0(u) = \frac{1}{2}\left[u \ln(u) + (1 - u) \ln(1 - u)\right]. \tag{10}$$

3. Proof of Theorem 1.1

Proof. 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

$$\begin{aligned} \delta t(g) \quad := \quad & t(g) - e^3 = 3e^2 \int_{[0,1]^2} \delta g(x, y)\, dxdy + 3e \int_{[0,1]^3} \delta g(x, y)\delta g(y, z)\, dxdydz \\ & + \int_{[0,1]^3} \delta g(x, y)\delta g(y, z)\delta g(z, x)\, dxdydz. \end{aligned} \tag{11}$$

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

$$\delta t = 3e\langle\phi_1, T_{\delta g}^2\phi_1\rangle + Tr(T_{\delta g}^3), \tag{12}$$

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 $\int_{[0,1]^2} \delta g(x,y)\,dxdy = 0$,

$$\begin{aligned} \delta I &= \int_{[0,1]^2} [I_0(e+\delta g(x,y)) - \delta g(x,y)I_0'(e) - I_0(e)]\,dxdy \\ &= \int_{[0,1]^2} \frac{I_0(e+\delta g(x,y)) - \delta g(x,y)I_0'(e) - I_0(e)}{\delta g(x,y)^2}\delta g(x,y)^2\,dxdy \\ &\geq f_-(e)\int_{[0,1]^2}\delta g(x,y)^2\,dxdy, \end{aligned} \tag{13}$$

where

$$f(e,x) = \frac{I_0(e+x) - xI_0'(e) - I_0(e)}{x^2}, \tag{14}$$

and $f_-(e) = \inf_x f(e,x)$ is a positive number less than or equal to $\frac{I_0''(e)}{2} = \frac{1}{4e(1-e)}$.

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

$$|Tr(T_{\delta g}^3)| = |\sum_i \mu_i^3| \leq \sum_i |\mu_i^3| \leq \max_j|\mu_j|\sum_i \mu_i^2 \leq (\sum_i \mu_i^2)^{3/2} = [Tr(T_{\delta g}^2)]^{3/2}. \tag{15}$$

If $T_{\delta g}$ has rank one, then $Tr(T_{\delta g}^3) = \mu^3 = \pm[Tr(T_{\delta g}^2)]^{3/2}$. 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

$$-\delta t = -Tr(T_{\delta g}^3) - 3e\langle\phi_1, T_{\delta g}^2\phi_1\rangle \leq -Tr(T_{\delta g}^3) \leq [Tr(T_{\delta g}^2)]^{3/2} \leq \left(\frac{\delta I}{f_-(e)}\right)^{3/2}. \tag{16}$$

This implies that

$$\delta I \geq f_-(e)(-\delta t)^{2/3}. \tag{17}$$

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

$$\delta I \geq \frac{f_-(e)}{e}|\delta t| \tag{18}$$

for $\delta t < 0$.

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

$$\delta t \leq Tr(T_{\delta g}^3) + 3eTr(T_{\delta g}^2) \leq (Tr(T_{\delta g})^2)^{3/2} + 3eTr(T_{\delta g}^2) \leq (3e+1)Tr(T_{\delta g}^2) \leq \frac{(3e+1)\delta I}{f_-(e)}, \tag{19}$$

so

$$\delta I \geq \frac{f_-(e)\delta t}{3e+1}. \tag{20}$$

□

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 (13) 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

$$\begin{aligned} \delta t \;&=\; \int_{[0,1]^2} h_1(x,y)\delta g(x,y)\,dxdy + \int_{[0,1]^4} h_2(w,x,y,z)\delta g(w,x)\delta g(y,z)\,dwdxdydz \\ &\quad + \int_{[0,1]^3} h_3(x,y,z)\delta g(x,y)\delta g(y,z)\,dxdydz + \cdots, \end{aligned} \tag{21}$$

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, …, such that

$$\begin{aligned} \delta t = c_1 \int_{[0,1]^2} \delta g(x,y)\,dxdy \quad &+ \quad c_2 \int_{[0,1]^4} \delta g(w,x)\delta g(y,z)\,dwdxdydz \\ &+ \quad c_3 \int_{[0,1]^3} \delta g(x,y)\delta g(y,z)\,dxdydz + \cdots \end{aligned} \tag{22}$$

The first two terms integrate to zero, the third is positive semi-definite, and all subsequent terms scale as δg^3 or higher.

For $\delta t < 0$, we must have $(-\delta t)$ of order $\left(\int_{[0,1]^2} \delta g(x,y)^2\,dxdy\right)^{3/2}$ or smaller, which again yields the analogue of equation (4), while for $\delta t > 0$ we simply note that, from (13), $\delta I \geq 0$

which implies $s(e, e^k) \geq s(e, t)$. Together these prove that $s(e, t)$ cannot be differentiable in t at $t = e^k$. $\square$

Note: 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 'Erdös-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 [RS]). Treating this as an optimization with constraints, one might naturally introduce Lagrange multipliers β_1, β_2 and consider the following optimization system,

$$\max_g [-I(g) + \beta_1 e(g) + \beta_2 t(g)];\ e(g) = e;\ t(g) = t, \tag{23}$$

namely maximize

$$\Psi_{\beta_1,\beta_2}(g) = -I(g) + \beta_1 e(g) + \beta_2 t(g) \tag{24}$$

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'

$$\psi(\beta_1, \beta_2) = \max_g \Psi_{\beta_1,\beta_2}(g) \tag{25}$$

is directly related to the normalization in exponential random graph models and basic information in such models is simply obtainable from it [N, CD, RY, AR]. 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 which convexity plays a significant role. In particular, although it has been proven ([CD, RY]) that $\psi(\beta_1, \beta_2)$ has singularities as a function of (β_1, β_2) (see Figure 2) it does not seem straightforward to use this to prove singularities in $s(e, t)$, which therefore 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 Figure 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 $t(g) = e(g)^3$ [CD, RY]. We emphasize that whenever $\beta_2 > -1/2$, one is on the Erdös-Rényi curve $t = e^3$ indicated in Figure 1 [CD, RY]. 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 Erdös-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 on the upper boundary of its domain R, while for $t < e^3$ they are related to those on the lower boundary

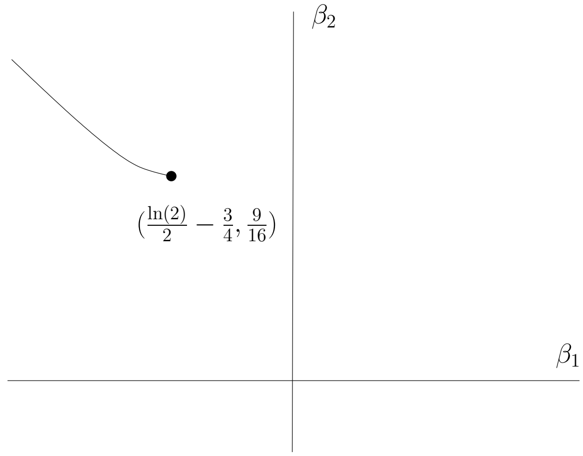

FIGURE 2. The curve of all singularities of $\psi(\beta_1, \beta_2)$, for $\beta_2 > -1/2$

of R. (The latter are the subject of [RS].) On the other hand, the transition in Figure 2, associated with varying (β_1, β_2), is between graphs of similar character (independent edges) but different densities, a phenomenon unrelated to the transition of Theorem 1.1 although still associated with the Erdös-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 Figure 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 in the Lagrange multiplier method can be misleading; they point to a 'transition' which is foreign to the maximization problem for $s(e, t)$. We next consider other features of the Lagrange multiplier method.

One issue of importance to those who study exponential random graph models is that no matter how large β_2 is, one cannot have maximizers g with $t(g) > e(g)^3$, though they certainly exist, as we see for instance from Figure 1. There are however less well known examples of 'missing' values of (e, t) when using the Lagrange multipliers (β_1, β_2), for $t(g) < e(g)^3$, which have a different origin.

5.1. **The special case of $e = 1/2$.** 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.

Lemma 5.1. *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. Assume $c \neq 0$ (or else the lemma is trivial). 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, $\delta t = Tr(T_{\delta g}^3) = c^3$. □

Now we try to minimize $\int_{[0,1]} I[e + c\alpha(x)\alpha(y)]\, dxdy$. By convexity, this is minimized when $[\alpha(x)\alpha(y)]^2$ is constant, which 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 can assume that

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

This means that the graphon that minimizes $I(g)$ for fixed $e = 1/2$ and fixed $t \leq e^3$ is

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

where $\epsilon = (e^3 - t)^{1/3}$, and that the associated reduced graphon is unique. The associated entropy is then

$$s(\frac{1}{2}, t) = \frac{-1}{2}\left[I_0\left(\frac{1}{2} + \epsilon\right) + I_0\left(\frac{1}{2} - \epsilon\right)\right] = -I_0\left(\frac{1}{2} + \epsilon\right), \tag{27}$$

since $I_0(u) = I_0(1-u)$.

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

$$-I_0'[g(x,y)] + \beta_1 + \beta_2 h(x,y) = 0, \tag{28}$$

where

$$h(x,y) = 3\int_{[0,1]} g(x,z)g(y,z)\, dz \tag{29}$$

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

$$\begin{aligned} \beta_1 + 3\beta_2\left(\frac{1}{4} - \epsilon^2\right) &= I_0'\left(\frac{1}{2} + \epsilon\right) = \frac{1}{2}\ln\left[\frac{\frac{1}{2}+\epsilon}{\frac{1}{2}-\epsilon}\right] \\ \beta_1 + 3\beta_2\left(\frac{1}{4} + \epsilon^2\right) &= I_0'\left(\frac{1}{2} - \epsilon\right) = \frac{1}{2}\ln\left[\frac{\frac{1}{2}-\epsilon}{\frac{1}{2}+\epsilon}\right], \end{aligned} \tag{30}$$

which are satisfied if and only if

$$\beta_2 = -\frac{4}{3}\beta_1 = \frac{I_0'\left(\frac{1}{2}-\epsilon\right) - I_0'\left(\frac{1}{2}+\epsilon\right)}{6\epsilon^2} = -\frac{1}{6\epsilon^2}\ln\left[\frac{\frac{1}{2}+\epsilon}{\frac{1}{2}-\epsilon}\right]. \tag{31}$$

Notice that β_1 and β_2 diverge as $\epsilon \downarrow 0$ (equivalently, as $t \uparrow 1/8$). Actually the situation is much worse. $s(1/2, t) = -I(\tilde{g})$ is a smooth function of $t \leq 1/8$. It is easy to check by differentiation of (27) 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 rather a local minimizer with respect to variation of t, and so there are no (β_1, β_2) which can lead to the maximizers of $s(1/2, t)$ for t just below $1/8$. This phenomenon is simply due to inequality (4), and therefore occurs for all e, not just $e = 1/2$. 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 Erdös-Rényi curve is actually the lower boundary of the domain of the entropy, in which case there are no 'missing' points (e, t) below it.]

References

[AR] D. Aristoff and C. Radin, Emergent structures in large networks, J. Appl. Probab. (to appear), arXiv:1110.1912

[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, arXiv: 1102.2650v3.

[CV] S. Chatterjee and S.R.S. Varadhan, The large deviation principle for the Erdös-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.

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

[PR] O. Pikhurko and A. Razborov, Asymptotic structure of graphs with the minimum number of triangles, arXiv:1203.4393

[RS] C. Radin and L. Sadun, arXiv:1301.1256.

[RY] C. Radin and M. Yin, Phase transitions in exponential random graphs, Ann. Appl. Probab. (to appear), arXiv:1108.0649.

Charles Radin, Department of Mathematics, The University of Texas at Austin, Austin, TX 78712

E-mail address: radin@math.utexas.edu

Lorenzo Sadun, Department of Mathematics, The University of Texas at Austin, Austin, TX 78712

E-mail address: sadun@math.utexas.edu