Ann. Inst. Henri Poincaré Comb. Phys. Interact. 5 (2018), 251–286
DOI 10.4171/AIHPD/54

A symmetry breaking transition in the edge/triangle network model

Charles Radin, Kui Ren, and Lorenzo Sadun

Abstract. Our general subject is the emergence of phases, and phase transitions, in large networks subjected to a few variable constraints. Our main result is the analysis, in the model using edge and triangle subdensities for constraints, of a sharp transition between two phases with different symmetries, analogous to the transition between a fluid and a crystalline solid.

Mathematics Subject Classification (2010). Primary 05C80; Secondary: 82B26.

Keywords. Graph limits, entropy, bipodal structure, phase transitions, symmetry breaking.

Contents

1. Introduction

Our general subject is the emergence of phases, and phase transitions, in large networks subjected to a few variable constraints, which we take as the densities of a few chosen subgraphs. We follow the line of research begun in [21] in which, for k given constraints on a large network one determines a global state on the network, by a variational principle, from which one can compute a wide range of global observables of the network, in particular the densities of all subgraphs. A phase is then a region of constraint values in which all these global observables vary smoothly with the constraint values, while transitions occur when the global state changes abruptly. (See [7, 8, 21, 22, 20, 10] and the survey [19].) Our main focus is on the transition between two particular phases, in the system with the two constraints of edge and triangle density: phases in which the global states have different symmetry.

If one describes networks (graphs) on n nodes by their $n \times n$ adjacency matrices, we will be concerned with *asymptotics* as $n \to \infty$, and in particular limits of these matrices considered as 0-1 valued graphons. (Much of the relevant asymptotics is a recent development [1, 2, 11, 12, 13]; see [14] for an encyclopedic treatment.) Given k subgraph densities as constraints, the asymptotic analysis in [21] leads to the study of one or more k-dimensional manifolds embedded in the infinite dimensional metric space W of (reduced) graphons, the points in the manifold being the emergent (global) states of a phase of the network. That is, there are smooth embeddings, of open connected subsets (called phases) of the phase space $\Gamma \subset [0,1]^k$ of possible constraint values, into W. The embeddings are obtained from the constrained entropy density $s(P)$, a real valued function of the constraints $P \in \Gamma$, through the variational principle [21, 22]: $s(P) = \sup\{\mathcal{S}(g) \,|\, C(g) = P\}$, the constraints being described by $C(g) = P$, for instance edge density $\varepsilon(g) = P_1$ and triangle density $\tau(g) = P_2$, and $\mathcal{S}$ being the negative of the large deviation rate function of Chatterjee–Varadhan [5].

The above framework is modeled on that of statistical physics in which the system is the simultaneous states of many interacting particles, the constraints are the invariants of motion (mass and energy densities for simple materials) and the global states, 'Gibbs states', can be understood in terms of conditional probabilities [27]. The relevant variational principle was proven in [26]. In contrast, for large graphs the constrained entropy optima turn out to be much easier to analyze than Gibbs states. First, it has been found that in all known cases entropy optima are 'multipodal', i.e. for each phase in any graph model they lie in some M dimensional manifold in W corresponding to a decomposition of all the nodes into M equivalence classes [7, 8, 10]. This brings the embedding down into a fixed

finite dimension, at least within each phase, in place of the infinite dimensions of W. In practice, for $k = 2$ this often allows one to find an embedding into 4 dimensions, so it only remains to understand how our 2 dimensional surface sits in those 4 dimensions. The main goal of this paper is to study this near a particular transition for edge/triangle constraints. The phase space Γ of possible values of edge/triangle constraints is the interior of the scalloped triangle of Razborov [24], sketched in the left side of Figure 1.

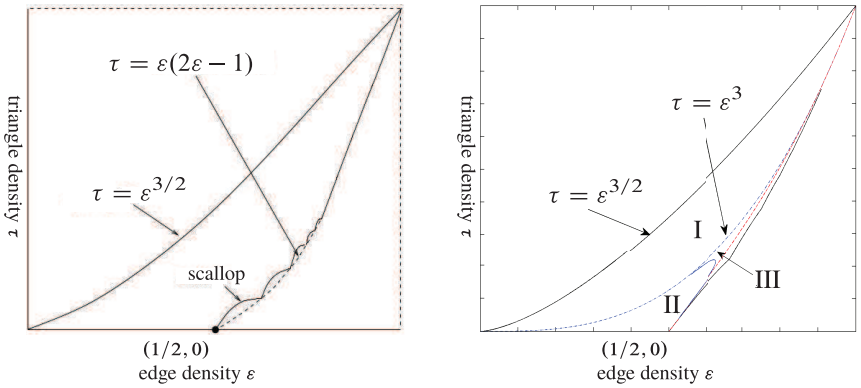

Figure 1. Left: Boundary of the phase space for subgraph constraints edges and triangles. The scalloped region is exaggerated for better visualization; Right: Boundaries of phases I, II and III.

In [22] it was proven that the entropy, viewed as a function of ε and τ, is not differentiable at the Erdös–Rényi curve $\tau = \varepsilon^3$. See the right side of Figure 1. Extensive numerical simulation suggests that the entire region above this curve consists of a single phase, which we call Phase I. Moreover, it was proven in [8] that for (ε, τ) in an open set just above the curve $\tau = \varepsilon^3$, the entropy-optimizing graphons are bipodal (see definition below) and of a form agreeing with the numerical simulations. Numerical simulations also indicate that the bulk of the region below the Erdös–Rényi curve consists of two distinct phases, which we call phases II and III.

The main focus of this paper is the transition between phases II and III. Simulations indicate that all entropy optimal graphons for edge/triangle densities in phases II and III are bipodal, that is given by a piecewise constant function of the form:

$$g(x, y) = \begin{cases} a & x, y < c, \\ d & x < c < y, \\ d & y < c < x, \\ b & x, y > c. \end{cases} \tag{1}$$

Bipodal graphons are generalizations of bipartite graphons, in which $a = b = 0$. Here c, a, d and b are constants taking values between 0 and 1. This bipodal structure was proven [22] for certain points in the interior of Phase II, and is supported by a perturbative analysis [21], but has not been proven in general.

In this paper, we prove the existence of a transition between phases II and III, and examine the structure of this phase transition, under the simplifying assumption that the entropy-maximizing graphons in the relevant regions are indeed bipodal. We do not assume that these entropy optimizers are unique. However, from the bipodal assumption we can prove uniqueness, and also derive, as shown already in [20], the equations determining the transition curve. We further prove that the optimal graphons for edge/triangle constraints (ε, τ) to the left of the II $\leftrightarrow$ III transition have the form:

$$g(x, y) = \begin{cases} \varepsilon - (\varepsilon^3 - \tau)^{1/3} & x, y < 1/2 \text{ or } x, y > 1/2, \\ \varepsilon + (\varepsilon^3 - \tau)^{1/3} & x < \frac{1}{2} < y \text{ or } y < \frac{1}{2} < x, \end{cases} \tag{2}$$

which is highly symmetric in the sense that $c = 1/2$ and $a = b$.

This gives a simple description of what phases I, II and III mean, which we reprise. Phase I is the region above the Erdös–Rényi curve $\tau = \varepsilon^3$. Phase II is the region below the Erdös–Rényi curve where the optimizing graphon is bipodal and symmetric under the interchange $c \leftrightarrow 1 - c$, $a \leftrightarrow b$. Phase III is the region below the Erdös–Rényi curve where the optimizing graphon is bipodal but the symmetry is broken. In addition, there exist three infinite families of smaller phases, clustered near the lower boundary of the profile, where the optimizing graphons are not bipodal; see [9] for investigations of those.

The main result of this paper is the derivation of the lower order terms in ε and τ of the bipodal parameters a, b, c, d as the constraints move to the right of the II $\leftrightarrow$ III transition. That is, we determine how the symmetry is broken at the transition. Before we get into details we should explain the connection between this study of emergent phases and their transitions, and other work on random graphs using similar terminology.

The word phase is used in many ways in the literature, but 'emergent phase' is more specific and refers to a coherent, large scale description of a system of many similar components, in the following sense. Consider large graphs, thought of as systems of many edges on a fixed number of labeled nodes. To work at a large scale means to be primarily concerned with global features or observables, for instance densities of subgraphs.

Finite graphs are fully described (i.e. on a small scale) by their adjacency matrices. In the graphon formalism these have a scale-free description in which each original node is replaceable by a cluster of m nodes, which makes for a convenient analysis connecting large and small scale.

Phases are concerned with large graphs under a small number k of variable global constraints, for instance constrained by the 2 subdensities of edges and triangles. We say one or more phases *emerge* for such systems, corresponding to one or more open connected subsets of parameter values, if: 1) there are *unique* global states (graphons) associated with the sets of constraint values; 2) the correspondence defines a smooth k-dimensional surface in the infinite dimensional space of states.

Note that not all achievable parameter values (the phase space) belong to phases, in particular the boundary of the phase space does not. In fact emergent phases are interesting in large part because they exhibit interesting (singular) boundary behavior in the interior of the phase space. For graphs we see this in at least two ways familiar from statistical mechanics. In the edge/2-star model, studied in detail in [7], there is only one phase but there are achievable parameter values with multiple states associated (as in the liquid/gas transition); see Figure 2. (See also [25] for some related developments.)

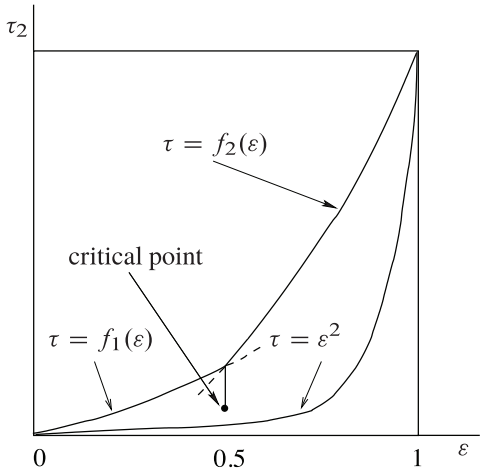

Figure 2. Transition ending in a critical point for edge/2-star model.

In most models constraint values on the Erdös–Rényi curve – parameter values corresponding to iid edges – provide unique states which however do not vary smoothly across the curve, and this provides boundaries of phases [21]. For edge/triangle constraints there is another phase boundary associated with a loss of

symmetry (as in the transition between fluid and crystalline solid). In all models studied so far phases make up all but a lower dimensional subset of the phase space but this may not be true in general.

From an alternative vantage random graphs are commonly used to model some particular network, the idea being to fit parameter values in a parametric family of probability distributions on some space of possible networks, so as to get a manageable probabilistic picture of the one target network of interest; see [4, 16] and references therein. The one parameter Erdös–Rényi (ER) family is often used this way for percolation, as are multiparameter generalizations such as stochastic block models and exponential random graph models (ERGMs). Asymptotics is only a small part of such research, and for ERGMs it poses some difficulties as is well described in the introduction of [4].

Newman et al considered phase transitions for ERGMs; see [17] for a particularly relevant example. Chatterjee–Varadhan introduced in [5] a powerful asymptotic formalism for random graphs with their large deviations theorem, within the previously developed graphon formalism; see [3] for a thorough exposition. This tool was applied to ERMGs in [4], which was then married with Newman's phase analysis in [23] and numerous following works, for instance [15, 28, 6] to cite recent examples.

We now contrast this use of the term phase with the emergent phase analysis discussed above. In the latter a phase corresponds to an embedding of part of the parameter space into the emergent (global) states of the network, so transitions could be interpreted as singular behavior *as those global network states varied*. For ERGMs it was shown in [4] that the analogous map from the parameter space Γ into W, determined by optimization of a free energy rather than entropy, is often many-to-one; for instance *all* states of the 2-parameter edge/2-star ERGM model actually belong to the 1-parameter ER family. Therefore transitions in an ERGM are naturally understood as asymptotic behavior of the *model* under variation of model parameters, rather than singular behavior among naturally varying global states of a network; in keeping with its natural use, a transition in an ERGM says more about the model than about constrained networks.

Acknowledgments. This work was partially supported by NSF grants DMS-1208191, DMS-1509088, DMS-1321018, DMS-1101326 and PHY-1125915.

2. Outline of the calculation

The purpose of this calculation is to understand the transition, in the edge/triangle model, between phases II and III in Figure 1. We take as an assumption that our optimizing graphon at fixed constraint (ε, τ) is bipodal as in (1), and with c being the size of the first cluster of nodes. We denote this graphon by g_{abcd}; see Figure 3. With the notation

$$S_o(z) = -\frac{1}{2}[z\ln(z) + (1-z)\ln(1-z)],\ 0 \le z \le 1 \tag{3}$$

and

$$\mathcal{S}(g) = \int_{[0,1]^2} S_o[g(x,y)]\,dxdy, \tag{4}$$

to obtain $s(P)$ we maximize $\mathcal{S}(g_{abcd})$ by varying the four parameters (a,b,c,d) while holding $P = (\varepsilon, \tau)$ fixed.

It is not hard to check that the symmetric graphon with $a = b = \varepsilon + (\tau - \varepsilon^3)^{1/3}$, $c = \frac{1}{2}$ and $d = \varepsilon - (\tau - \varepsilon^3)^{1/3}$ is always a stationary point of the functional $\mathcal{S}$. But is it a maximum? Near the choices of (ε, τ) where it ceases to be a maximum, what does the actual maximizing graphon look like?

We answer this by doing our constrained optimization in stages. First we fix c and vary (a,b,d) to maximize $\mathcal{S}(g)$ subject to the constraints on ε and τ. Call this maximum value $S(c)$. Since the graphon with parameters $(b,a,1-c,d)$ is equivalent to (a,b,c,d), $S(c)$ is an even function of $c - \frac{1}{2}$. We expand this function in a Taylor series:

$$S(c) = S\Big(\frac{1}{2}\Big) + \frac{1}{2}\ddot{S}\Big(\frac{1}{2}\Big)\Big(c - \frac{1}{2}\Big)^2 + \frac{1}{24}\dddot{S}\Big(\frac{1}{2}\Big)\Big(c - \frac{1}{2}\Big)^4 + \cdots \tag{5}$$

where dots denote derivatives with respect to c. It appears likely that the infinite Taylor series is convergent, thanks to the analytic inputs to the optimization problem, but this is not actually needed for our analysis. All we need are estimates on the errors in the 2nd- and 4th-order Taylor polynomials, which follow from the smoothness of S.

If $\ddot{S}\big(\frac{1}{2}\big)$ is negative, then the symmetric graphon is stable against small changes in c. If $\ddot{S}\big(\frac{1}{2}\big)$ becomes positive (as we vary ε and τ), then the local maximum of $S(c)$ at $c = \frac{1}{2}$ becomes a local minimum. As long as $\dddot{S}\big(\frac{1}{2}\big) < 0$, new local maxima will appear at $c \approx \frac{1}{2} \pm \sqrt{-6\ddot{S}\big(\frac{1}{2}\big)/\dddot{S}\big(\frac{1}{2}\big)}$. In this case, as we pass from Phase II to Phase III (say, along a curve $\gamma(t)$ that crosses the phase transition curve transversally at $t = 0$), we should expect the optimal $\big|c - \frac{1}{2}\big|$ to be exactly zero for $t < 0$, and to vary as $\sqrt{t}$ for $t > 0$.

If $\dddot{S}\left(\frac{1}{2}\right)$ is positive when $\ddot{S}\left(\frac{1}{2}\right)$ passes through zero, something very different happens. Although $c = \frac{1}{2}$ is a *local* maximum whenever $\ddot{S}\left(\frac{1}{2}\right) < 0$, there are local maxima elsewhere, at locations determined largely by the higher order terms in the Taylor expansion (5). When $\ddot{S}\left(\frac{1}{2}\right)$ passes below a certain threshold, one of these local maxima will have a higher entropy than the local maximum at $c = \frac{1}{2}$, and the optimal value of c will change discontinuously.

The calculations below give analytic formulas for $\ddot{S}\left(\frac{1}{2}\right)$ and $\dddot{S}\left(\frac{1}{2}\right)$ as functions of ε and τ. We can then evaluate these formulas numerically to fix the location of the phase transition curve and determine which of the previous two paragraphs more accurately describes the phase transition near a given point on that curve. We then compare these results to numerical sampling that is done *without* the simplifying assumption that optimizing graphons are bipodal.

Our results indicate that

- The assumption of bipodality is justified, and
- $\dddot{S}\left(\frac{1}{2}\right)$ remains negative on the phase transition curve, implying that c does not change discontinuously. Rather, in the asymmetric bipodal phase III, $\left|c - \frac{1}{2}\right|$ goes as the square root of the distance to the phase transition curve. And $c = \frac{1}{2}$ on the curve.

3. Exact formulas for $\ddot{S}\left(\frac{1}{2}\right)$ and $\dddot{S}\left(\frac{1}{2}\right)$

We now present our perturbation calculations. The argument is long and complicated, yet elementary. To improve flow while still providing the necessary details to make a convincing case of the correctness of the results, some of the steps have been placed in the appendix.

3.1. Varying (a, b, d) for fixed c. The first step in the calculation is to derive the variational equations for maximizing $\mathcal{S}(g_{abcd})$ for fixed (ε, τ, c). We first express the edge and triangle densities and $S = \mathcal{S}(g_{abcd})$ as functions of (a, b, c, d):

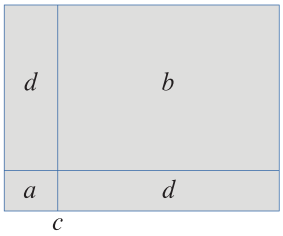

Figure 3. The parameter set (a, b, c, d) of a bipodal graphon.

$$\varepsilon = c^2 a + 2c(1-c)d + (1-c)^2 b, \tag{6a}$$

$$\tau = c^3 a^3 + 3c^2(1-c)ad^2 + 3c(1-c)^2 bd^2 + (1-c)^3 b^3, \tag{6b}$$

$$S = c^2 S_o(a) + (1-c)^2 S_o(b) + 2c(1-c)S_o(d). \tag{6c}$$

Next we compute the gradient of these quantities with respect to (a, b, d), where ∇_3 is an abbreviation for $(\partial_a, \partial_b, \partial_d)$:

$$\nabla_3 \varepsilon = (c^2, (1-c)^2, 2c(1-c)), \tag{7a}$$

$$\begin{aligned}\nabla_3 \tau = (&3c^3 a^2 + 3c^2(1-c)d^2, 3(1-c)^3 b^2 + 3c(1-c)^2 d^2, \\ &6c^2(1-c)ad + 6c(1-c)^2 bd)\end{aligned} \tag{7b}$$

$$\nabla_3 S = (c^2 S_o'(a), (1-c)^2 S_o'(b), 2c(1-c)S_o'(d)). \tag{7c}$$

The maximum of S (for fixed c, ε and τ) must occur at a point where a is strictly between 0 and 1, thanks to the divergence of $S_o'(a)$ to $\pm\infty$ as $a \to 0$ or $a \to 1$. Likewise, b and d remain bounded away from 0 and 1. Thus $\nabla_3 S$ must be a linear combination of $\nabla_3 \varepsilon$ and $\nabla_3 \tau$ so

$$\begin{aligned} 0 &= \det \begin{pmatrix} \nabla_3 \varepsilon \\ \nabla_3 \tau \\ \nabla_3 S \end{pmatrix} \\ &= c^2(1-c)^2 6c(1-c) \\ &\quad \det \begin{pmatrix} 1 & 1 & 1 \\ a^2 c + d^2(1-c) & d^2 c + b^2(1-c) & cad + (1-c)bd \\ S_o'(a) & S_o'(b) & S_o'(d) \end{pmatrix} \end{aligned} \tag{8}$$

Expanding the determinant and dividing by $6c^3(1-c)^3$, we obtain

$$\begin{aligned} 0 = f(a,b,c,d) := &S_o'(a)[cd(a-d) - (1-c)b(b-d)] \\ &+ S_o'(b)[ca(a-d) - (1-c)d(b-d)] \\ &+ S_o'(d)[c(d^2 - a^2) + (1-c)(b^2 - d^2)]. \end{aligned} \tag{9}$$

3.2. Strategy for varying c. We just showed how optimizing the entropy for fixed ε, τ, and c is equivalent to setting $f = 0$. From now on we treat $f = 0$ as an additional constraint. The three constraint equations $\varepsilon = \varepsilon_o$, $\tau = \tau_o$, $f = 0$ define a curve in (a, b, d, c) space, which we can parametrize by c. Since ε, τ

and f are constant, derivatives of these quantities along the curve must be zero. By evaluating these derivatives at $c = 1/2$, we will derive formulas for $\dot{a}(1/2)$, $\dot{b}(1/2)$, etc., where a dot denotes a derivative with respect to c along this curve. These values then determine $\ddot{S}\left(\frac{1}{2}\right)$ and $\dddot{S}\left(\frac{1}{2}\right)$.

We also make use of symmetry. The parameters (a, b, c, d) and $(b, a, 1-c, d)$ describe the same reduced graphon. Thus, if the conditions $\varepsilon = \varepsilon_0$, $\tau = \tau_0$, $f = 0$ trace out a unique curve in (a, b, c, d) space, then we must have

$$\begin{aligned} \dot{a}(1/2) + \dot{b}(1/2) &= \ddot{a}(1/2) - \ddot{b}(1/2) \\ &= \ddot{a}(1/2) + \dddot{b}(1/2) \\ &= \dddot{a}(1/2) - \ddddot{b}(1/2) \\ &= \dot{d}(1/2) \\ &= \ddot{d}(1/2) \\ &= 0, \end{aligned} \tag{10}$$

since all of these quantities are odd under the interchange $(a, b, c, d) \leftrightarrow (b, a, 1-c, d)$.

In fact, it is possible to derive the relations (10), and similar relations for higher derivatives, *without* assuming uniqueness of the curve. We proceed by induction on the degree k of the derivatives involved. At 0-th order we already have that $a(1/2) = b(1/2)$.

The kth derivative of ε and τ, evaluated at $c = 1/2$, are

$$\begin{aligned} \varepsilon^{(k)}(1/2) = &\frac{1}{4}(a^{(k)}(1/2) + b^{(k)}(1/2) + 2d^{(k)}(1/2)) \\ &+ (\text{lower order derivatives of } a, b, d), \end{aligned} \tag{11a}$$

$$\begin{aligned} \tau^{(k)}(1/2) = &\frac{3}{8}((a(1/2)^2 + d(1/2)^2)a^{(k)}(1/2) + (b(1/2)^2 + d(1/2)^2)b^{(k)}(1/2)) \\ &+ \frac{3}{4}(a(1/2) + b(1/2))d(1/2)d^{(k)}(1/2) \\ &+ (\text{lower order derivatives and their products}). \end{aligned} \tag{11b}$$

When k is odd, the lower-order terms vanish by induction, and we are left with

$$0 = a^{(k)}(1/2) + b^{(k)}(1/2) + 2d^{(k)}(1/2), \tag{12a}$$

$$0 = (a(1/2)^2 + d(1/2)^2)(a^{(k)}(1/2) + b^{(k)}(1/2)) + 4a(1/2)d(1/2)d^{(k)}(1/2), \tag{12b}$$

since $b(1/2) = a(1/2)$. The matrix $\begin{pmatrix} 1 & 2 \\ a(1/2)^2+d(1/2)^2 & 4a(1/2)d(1/2) \end{pmatrix}$ is non-singular, having determinant $-2(d(1/2)-a(1/2))^2$, so $a^{(k)}(1/2) + b^{(k)}(1/2) = d^{(k)}(1/2) = 0$.

When k is even, the kth derivative of f, evaluated at $c = 1/2$, is of the form

$$\begin{aligned}&(\text{even function with respect to the interchange}) \times (a^{(k)}(1/2) - b^{(k)}(1/2)) \\ &\quad + \text{lower order derivatives.}\end{aligned} \tag{13}$$

As before, the lower order terms vanish by induction, and the even function is nonzero, so we get $a^{(k)}(1/2) = b^{(k)}(1/2)$.

Henceforth we will freely use the relations in (10) to simplify our expressions. Here are the steps of the calculation:

- setting $\dot{f}(1/2) = 0$ determines $\dot{a}(1/2)$;
- setting $\ddot{\varepsilon}(1/2) = \ddot{\tau}(1/2) = 0$ determines $\ddot{a}(1/2)$ and $\ddot{d}(1/2)$ (in terms of $\dot{a}(1/2)$);
- setting $\dddot{f}(1/2) = 0$ determines $\dddot{a}(1/2)$;
- setting $\dddot{\varepsilon}(1/2) = \dddot{\tau}(1/2) = 0$ determines $\dddot{a}(1/2)$ and $\dddot{d}(1/2)$;
- once all derivatives of (a, b, d) up to 4th order are evaluated at $c = 1/2$, we explicitly compute $\ddot{S}(1/2)$ and $\dddot{S}(1/2)$.

3.3. Derivatives of ε and τ. The edge density and its first four derivatives are

$$\varepsilon = c^2 a + 2(c-c^2)d + (1-c)^2 b, \tag{14a}$$

$$\dot{\varepsilon} = c^2\dot{a} + 2(c-c^2)\dot{d} + (1-c)^2\dot{b} + 2ca - 2(1-c)b + 2(1-2c)d, \tag{14b}$$

$$\begin{aligned}\ddot{\varepsilon} = {} & c^2\ddot{a} + (1-c)^2\ddot{b} + 2(c-c^2)\ddot{d} + 4c\dot{a} - 4(1-c)\dot{b} + 4(1-2c)\dot{d} \\ & + 2a + 2b - 4d,\end{aligned} \tag{14c}$$

$$\begin{aligned}\dddot{\varepsilon} = {} & c^2\dddot{a} + (1-c)^2\dddot{b} + 2c(1-c)\dddot{d} + 6c\ddot{a} - 6(1-c)\ddot{b} + 6(1-2c)\ddot{d} \\ & + 6\dot{a} + 6\dot{b} - 12\dot{d}\end{aligned} \tag{14d}$$

$$\begin{aligned}\ddddot{\varepsilon} = {} & c^2\ddddot{a} + (1-c)^2\ddddot{b} + 2c(1-c)\ddddot{d} + 8c\dddot{a} - 8(1-c)\dddot{b} + 8(1-2c)\dddot{d} \\ & + 12\ddot{a} + 12\ddot{b} - 24\ddot{d}.\end{aligned} \tag{14e}$$

Setting $c = 1/2$ and applying the symmetries (10) gives

$$\dot{\varepsilon}(1/2) = \frac{1}{4}(\dot{a} + \dot{b} + 2\dot{d}) = 0, \tag{15a}$$

$$\begin{aligned}\ddot{\varepsilon}(1/2) &= \frac{1}{4}(\ddot{a} + \ddot{b} + 2\ddot{d}) + 4\dot{a} + 4(a-d)\\ &= \frac{1}{2}(\ddot{a} + \ddot{d}) + 4\dot{a} + 4(a-d),\end{aligned} \tag{15b}$$

$$\dddot{\varepsilon}(1/2) = \frac{1}{4}(\dddot{a} + \dddot{b} + 2\dddot{d}) = 0, \tag{15c}$$

$$\begin{aligned}\ddddot{\varepsilon}(1/2) &= \frac{1}{4}(\ddddot{a} + \ddddot{b} + 2\ddddot{d}) + 8\dddot{a} + 24\ddot{a} - 24\ddot{d}\\ &= \frac{1}{2}(\ddddot{a} + \ddddot{d}) + 8\dddot{a} + 24(\ddot{a} - \ddot{d}),\end{aligned} \tag{15d}$$

where all terms on the right hand side are evaluated at $c = 1/2$.

Similarly, the derivatives of the triangle density at $c = 1/2$ are as follows (see appendix for details):

$$\dot{\tau}(1/2) = 0, \tag{16a}$$

$$\begin{aligned}\ddot{\tau}(1/2) = {} & (3/4)[(a^2 + d^2)\ddot{a} + 2ad\ddot{d}] + (3/2)a\dot{a}^2\\ & + 3(3a^2 + d^2)\dot{a} + 6a(a^2 - d^2),\end{aligned} \tag{16b}$$

$$\dddot{\tau}(1/2) = 0, \tag{16c}$$

$$\begin{aligned}\ddddot{\tau}(1/2) = {} & (1/4)[3a^2\ddddot{a} + 24a\dot{a}\dddot{a} + 18a\ddot{a}^2 + 36\dot{a}^2\ddot{a}\\ & + 3d^2\ddddot{a} + 36d\ddot{d}\ddot{a} + 18a\ddot{d}^2 + 6ad\dddot{d}]\\ & + 6[3a^2\ddot{a} + 18a\dot{a}\ddot{a} + 6\dot{a}^3 + \ddot{a}d^2 + 6\dot{a}d\ddot{d}]\\ & + 36[3a^2\ddot{a} + 6a\dot{a}^2 - d^2\ddot{a} - 2ad\ddot{d}]\\ & + 48[3a^2\dot{a} - 3d^2\dot{a}],\end{aligned} \tag{16d}$$

where once again all quantities on the right hand side of these equations are evaluated at $c = 1/2$.

3.4. Derivatives of $f(a, b, c, d)$. Computing derivatives of f with respect to c involves organizing many terms. We use the notation x_i to denote one of $\{a, b, d\}$, and f_i to denote a partial derivative with respect to i. Since f is linear in c, we need only take one partial derivative in the c direction, but arbitrarily many in the other directions.

We have

$$\dot{f} = f_c + \sum_i f_i \dot{x}_i, \tag{17}$$

$$\ddot{f} = 2\sum_i f_{ci}\dot{x}_i + \sum_i f_i \ddot{x}_i + \sum_{ij} f_{ij}\dot{x}_i\dot{x}_j, \tag{18}$$

since $f_{cc} = 0$, and

$$\begin{aligned}
\dddot{f} =& = 3\sum_i f_{ci}\ddot{x}_i + 3\sum_{ij} f_{cij}\dot{x}_i\dot{x}_j + \sum_i f_i\dddot{x}_i + 3\sum_{i,j} f_{ij}\ddot{x}_i\dot{x}_j + \sum_{ijk} f_{ijk}\dot{x}_i\dot{x}_j\dot{x}_k \\
&= 3f_{ca}\ddot{a} + 3f_{cb}\ddot{b} + 3f_{cd}\ddot{d} + 3f_{caa}\dot{a}^2 + +6f_{cab}\dot{a}\dot{b} \\
&\quad + 6f_{cad}\dot{a}\dot{d} + 3f_{cbb}\dot{b}^2 + 6f_{cbd}\dot{b}\dot{d} + 3f_{cdd}\dot{d}^2 \\
&\quad + f_a\dddot{a} + f_b\dddot{b} + f_d\dddot{d} + 3f_{aa}\dot{a}\ddot{a} + 3f_{ab}\dot{a}\ddot{b} + 3f_{ad}\dot{a}\ddot{d} \\
&\quad + 3f_{ab}\dot{b}\ddot{a} + 3f_{bb}\dot{b}\ddot{b} + 3f_{bd}\dot{b}\ddot{d} + 3f_{ad}\ddot{a}\dot{d} + 3f_{bd}\ddot{b}\dot{d} \\
&\quad + 3f_{dd}\ddot{d}\dot{d} + f_{aaa}\dot{a}^3 + 3f_{aab}\dot{a}^2\dot{b} + 3f_{aad}\dot{a}^2\dot{d} \\
&\quad + 3f_{abb}\dot{a}\dot{b}^2 + 6f_{abd}\dot{a}\dot{b}\dot{d} + 3f_{add}\dot{a}\dot{d}^2 + 3f_{bbd}\dot{b}^2\dot{d} \\
&\quad + 3f_{bdd}\dot{b}\dot{d}^2 + f_{ddd}\dot{d}^3 + f_{bbb}\dot{b}^3.
\end{aligned} \tag{19}$$

When $c = 1/2$ we have $\dot{d} = \ddot{d} = 0$, so many of the terms vanish. We then have

$$\begin{aligned}
\dddot{f}(1/2) &= 3f_{ca}\ddot{a} + 3f_{cb}\ddot{b} + 3f_{cd}\ddot{d} + 3f_{caa}\dot{a}^2 + 6f_{cab}\dot{a}\dot{b} \\
&\quad + 3f_{cbb}\dot{b}^2 + f_a\dddot{a} + f_b\dddot{b} + 3f_{aa}\dot{a}\ddot{a} + 3f_{ab}\dot{a}\ddot{b} \\
&\quad + 3f_{ad}\dot{a}\ddot{d} + 3f_{ab}\dot{b}\ddot{a} + 3f_{bb}\dot{b}\ddot{b} + 3f_{bd}\dot{b}\ddot{d} \\
&\quad + f_{aaa}\dot{a}^3 + 3f_{aab}\dot{a}^2\dot{b} + 3f_{abb}\dot{a}\dot{b}^2 + f_{bbb}\dot{b}^3 \\
&= (3f_{ca} + 3f_{cb})\ddot{a} + 3f_{cd}\ddot{d} + (3f_{caa} - 6f_{cab} + 3f_{cbb})\dot{a}^2 \\
&\quad + (f_a - f_b)\dddot{a} + (3f_{aa} + 3f_{ab} - 3f_{ab} - 3f_{bb})\dot{a}\ddot{a} \\
&\quad + (3f_{ad} - 3f_{bd})\dot{a}\ddot{d} + (f_{aaa} - 3f_{aab} + 3f_{abb} - f_{bbb})\dot{a}^3,
\end{aligned} \tag{20}$$

evaluated at $c = 1/2$, where in the last step we have also used $\dot{b}(1/2) = -\dot{a}(1/2)$, $\ddot{b}(1/2) = \ddot{a}(1/2)$ and $\dddot{b}(1/2) = -\dddot{a}(1/2)$. Solving for $\dddot{a}(1/2)$ then gives

$$\begin{aligned}
\dddot{a}(1/2) = \big[&3(f_{ca} + f_{cb})\ddot{a} + 3f_{cd}\ddot{d} + 3(f_{caa} - 2f_{cab} + f_{cbb})\dot{a}^2 \\
&+ 3(f_{aa} - f_{bb})\dot{a}\ddot{a} + 3(f_{ad} - f_{bd})\dot{a}\ddot{d} \\
&+ (f_{aaa} - 3f_{aab} + 3f_{abb} - f_{bbb})\dot{a}^3\big]/(f_b - f_a),
\end{aligned} \tag{21}$$

evaluated at $c = 1/2$. What remains is to compute the partial derivatives of f that appear in equation (21).

Computing all the partial derivatives f_a, f_b, etc. is a long but straightforward exercise in calculus that can be found in the appendix.

3.5. Evaluating $\dot{a}$, $\ddot{a}$, $\dddot{d}$, $\ddddot{a}$, $\ddddot{a}$, and $\ddddot{d}$. Using the partial derivatives of f computed in the appendix, we have

$$\begin{aligned}\dot{f}(1/2) &= S_o'(a)[a^2 - d^2 + (1/2)d\dot{a} + (1/2)d\dot{b} - b\dot{b}] \\ &\quad + S_o'(b)[a^2 - d^2 + a\dot{a} - (1/2)d\dot{a} - (1/2)d\dot{b}] \\ &\quad + S_o'(d)[2d^2 - 2a^2 - a\dot{a} + b\dot{b}] \\ &\quad + \dot{a}S_o''(a)[-(1/2)(a-d)^2] + \dot{b}S_o''(b)[(1/2)(a-d)^2] \\ &= S_o'(a)[2a^2 - 2d^2 + 2a\dot{a}] + S_o'(d)[2d^2 - 2a^2 - 2a\dot{a}] \\ &\quad - \dot{a}S_o''(a)(a-d)^2,\end{aligned} \tag{22}$$

where all terms on the right are evaluated at $c = 1/2$. Setting this equal to zero then yields

$$0 = 2(a^2 - d^2)[S_o'(a) - S_o'(d)] + \dot{a}[2a(S_o'(a) - S_o'(d)) - S_o''(a)(a-d)^2] \tag{23}$$

and

$$\dot{a}(1/2) = \frac{2(d^2 - a^2)[S_o'(a) - S_o'(d)]}{2a[S_o'(a) - S_o'(d)] - S_o''(a)(a-d)^2}, \tag{24}$$

evaluated at $c = 1/2$. It is convenient to define the ratio

$$\alpha := \frac{\dot{a}(1/2)}{a(1/2) - d(1/2)}. \tag{25}$$

We now use this value of $\dot{a}(1/2)$ to compute $\ddot{a}(1/2)$ and $\ddot{d}(1/2)$:

$$0 = 2\ddot{\varepsilon}(1/2) = \ddot{a} + \ddot{d} + 8\dot{a} + 8(a-d), \tag{26}$$

where all quantities are computed at $c = 1/2$. Likewise,

$$0 = \frac{4}{3}\ddot{\tau} = (a^2 + d^2)\ddot{a} + 2ad\ddot{d} + 2a\dot{a}^2 + 4(3a^2 + d^2)\dot{a} + 8a(a^2 - d^2). \tag{27}$$

However, $2ad$ times equation (26) is

$$0 = 2ad\ddot{a} + 2ad\ddot{d} + 16ad\dot{a} + 16ad(a-d). \tag{28}$$

Subtracting these equations gives

$$\begin{aligned} 0 &= (a-d)^2\ddot{a} + 2a\dot{a}^2 + (12a^2 + 4d^2 - 16ad)\dot{a} \\ &\quad + 8a^3 - 8ad^2 - 16a^2 d + 16ad^2 \\ &= (a-d)^2\ddot{a} + 2a\dot{a}^2 + 4(a-d)(3a-d)\dot{a} + 8a(a-d)^2. \end{aligned} \tag{29}$$

Solving for $\ddot{a}$ then gives

$$\begin{aligned} \ddot{a}(1/2) &= \frac{-2a(\dot{a}^2 - 4(a-d)(3a-d)\dot{a} - 8a(a-d)^2}{(a-d)^2} \\ &= -2a\alpha^2 - 4\alpha(3a-d) - 8a. \end{aligned} \tag{30}$$

We then obtain $\dddot{d}(1/2)$ from equation (26):

$$\begin{aligned} \dddot{d}(1/2) &= -\ddot{a} - 8\dot{a} - 8a + 8d \\ &= 2a\alpha^2 + 4\alpha(a+d) + 8d. \end{aligned} \tag{31}$$

Plugging all of these expressions back into equation (21) yields

$$\begin{aligned} \ddot{a}(1/2) = \big[&\{[6S_o''(a)(a^2-d^2) + 12a(S_o'(a) - S_o'(d))]\ddot{a} \\ &+ [6S_o''(d)(d^2-a^2) + 12d(S_o'(d) - S_o'(a))]\ddot{d} \\ &+ [6S_o'''(a)(a^2-d^2) + 24(d-a)S_o''(a) + 12(S_o'(a) - S_o'(d))]\dot{a}^2 \\ &+ [-3(a-d)^2 S_o'''(a) + 6dS_o''(a) + 6(S_o'(a) - S_o'(d))]\dot{a}\ddot{a} \\ &+ 6[(a-d)S_o''(a) - aS_o''(d)]\dot{a}\ddot{d} \\ &+ [-S_o''''(a)(a-d)^2 + 6aS_o'''(a) - 6S_o''(a)]\dot{a}^3\} \\ &\big[S_o''(a)(a-d)^2 - 2a[S_o'(a) - S_o'(d)\big]^{-1}. \end{aligned} \tag{32}$$

Finally, we need $\dddot{a}(1/2)$ and $\ddddot{d}(1/2)$. From equations (15) and (16),

$$\begin{aligned} 0 &= 2\dddot{\varepsilon}(1/2) \\ &= \dddot{a} + \ddddot{d} + 16\ddot{a} + 48(\ddot{a} - \ddot{d}), \end{aligned}$$

$$\begin{aligned} 0 &= \frac{4}{3}\dddot{\tau}(1/2) \\ &= (a^2+d^2)\dddot{a} + 2ad\ddddot{d} + 8a\dot{a}\ddot{a} + 6a\ddot{a}^2 + 12\dot{a}^2\ddot{a} + 12d\dddot{d}\ddot{a} + 6a\dddot{d}^2 \\ &\quad + 8[3a^2\ddot{a} + 18a\dot{a}\ddot{a} + 6\dot{a}^3 + \ddot{a}d^2 + 6\dot{a}d\dddot{d}] \\ &\quad + 48[3a^2\ddot{a} + 6a\dot{a}^2 - d^2\ddot{a} - 2ad\dddot{d}] + 64[3a^2\dot{a} - 3d^2\dot{a}] \end{aligned} \tag{33}$$

$$
\begin{aligned}
0 &= (4/3)\dddot{\tau}(1/2) - 4ad\dddot{\varepsilon}(1/2) \\
&= (a-d)^2\dddot{a} + 8a\dot{a}\ddot{a} + 6a\ddot{a}^2 + 12\dot{a}^2\ddot{a} \\
&\quad + 12d\ddot{d}\ddot{a} + 6a\ddot{d}^2 \\
&\quad + \dddot{a}(24a^2 + 8d^2 - 32ad) + 144a\dot{a}\ddot{a} + 48\dot{a}^3 + 48\dot{a}d\ddot{d} \\
&\quad + 48[\ddot{a}(3a^2 - 2ad - d^2) + 6a\dot{a}^2] + 192(a^2 - d^2)\dot{a}.
\end{aligned} \tag{34}
$$

Solving for $\dddot{a}$ gives

$$
\begin{aligned}
\dddot{a}(1/2) = -\big[&8a\dot{a}\ddot{a} + 6a\ddot{a}^2 + 12\dot{a}^2\ddot{a} + 12d\ddot{d}\ddot{a} + 6a\ddot{d}^2 \\
&+ \dddot{a}(24a^2 + 8d^2 - 32ad) + 144a\dot{a}\ddot{a} + 48\dot{a}^3 + 48\dot{a}d\ddot{d} \\
&+ 48[\ddot{a}(3a^2 - 2ad - d^2) + 6a\dot{a}^2] \\
&+ 192(a^2 - d^2)\dot{a}\big]/(a-d)^2.
\end{aligned} \tag{35}
$$

We then obtain

$$
\dddot{d}(1/2) = -\dddot{a} - 16\ddot{a} - 48(\ddot{a} - \ddot{d}) \tag{36}
$$

3.6. Derivatives of S. Finally, we compute derivatives of the functional S. We have

$$
S = c^2 S_o(a) + (1-c)^2 S_o(b) + 2c(1-c)S_o(d), \tag{37}
$$

$$
\begin{aligned}
\dot{S} &= 2cS_o(a) - 2(1-c)S_o(b) + 2(1-2c)S_o(d) \\
&\quad + c^2 S_o'(a)\dot{a} + (1-c)^2 S_o'(b)\dot{b} + 2c(1-c)S_o'(d)\dot{d},
\end{aligned} \tag{38}
$$

$$
\begin{aligned}
\ddot{S} &= 2[S_o(a) + S_o(b) - 2S_o(d)] + 4cS_o'(a)\dot{a} - 4(1-c)S_o'(b)\dot{b} \\
&\quad + 4(1-2c)S_o'(d)\dot{d} + c^2[S_o'(a)\ddot{a} + S_o''(a)\dot{a}^2] + (1-c)^2[S_o'(b)\ddot{b} + S_o''(b)\dot{b}^2] \\
&\quad + 2c(1-c)[S_o'(d)\ddot{d} + S_o''(d)\dot{d}^2],
\end{aligned} \tag{39}
$$

$$
\begin{aligned}
\dddot{S} &= 6[S_o'(a)\dot{a} + S_o'(b)\dot{b} - 2S_o'(d)\dot{d}] \\
&\quad + 6c[S_o'(a)\ddot{a} + S_o''(a)\dot{a}^2] - 6(1-c)[S_o'(b)\ddot{b} + S_o''(b)\dot{b}^2] \\
&\quad + 6(1-2c)[S_o'(d)\ddot{d} + S_o''(d)\dot{d}^2] \\
&\quad + c^2[S_o'(a)\dddot{a} + 3S_o''(a)\dot{a}\ddot{a} + S_o'''(a)\dot{a}^3] \\
&\quad + (1-c)^2[S_o'(b)\dddot{b} + 3S_o''(b)\dot{b}\ddot{b} + S_o'''(b)\dot{b}^3] \\
&\quad + 2c(1-c)[S_o'(d)\dddot{d} + 3S_o''(d)\dot{d}\ddot{d} + S_o'''(d)\dot{d}^3],
\end{aligned} \tag{40}
$$

$$\begin{aligned}\dddot{S} &= 12[S_o'(a)\ddot{a} + S_o''(a)\dot{a}^2 + S_o'(b)\ddot{b} + S_o''(b)\dot{b}^2 - 2S_o'(d)\ddot{d} - 2S_o''(d)\dot{d}^2] \\ &\quad + 8c[S_o'(a)\ddot{a} + 3S_o''(a)\dot{a}\ddot{a} + S_o'''(a)\dot{a}^3] \\ &\quad - 8(1-c)[S_o'(b)\ddot{b} + 3S_o''(b)\dot{b}\ddot{b} + S_o'''(b)\dot{b}^3] \\ &\quad + 8(1-2c)[S_o'(d)\ddot{d} + 3S_o''(d)\dot{d}\ddot{d} + S_o'''(d)\dot{d}^3] \\ &\quad + c^2[S_o'(a)\dddot{a} + 4S_o''(a)\dot{a}\ddot{a} + 3S_o''(a)\ddot{a}^2 + 6S_o'''(a)\dot{a}^2\ddot{a} + S_o''''(a)\dot{a}^4] \\ &\quad + (1-c)^2[S_o'(b)\dddot{b} + 4S_o''(b)\dot{b}\ddot{b} + 3S_o''(b)\ddot{b}^2 + 6S_o'''(b)\dot{b}^2\ddot{b} + S_o''''(b)\dddot{b}] \\ &\quad + 2c(1-c)[S_o'(d)\dddot{d} + 4S_o''(d)\dot{d}\ddot{d} + 3S_o''(d)\ddot{d}^2 \\ &\quad + 6S_o'''(d)\dot{d}^2\ddot{d} + S_o''''(d)\dot{d}^4]\end{aligned} \tag{41}$$

Evaluating at $c = 1/2$ and using symmetry gives $\dot{S}(1/2) = \dddot{S}(1/2) = 0$ and

$$\begin{aligned}\ddot{S}(1/2) &= 4[S_o(a) - S_o(d)] + 4S_o'(a)\dot{a} \\ &\quad + (1/2)[S_o'(a)\ddot{a} + S_o''(a)\dot{a}^2 + S_o'(d)\ddot{d}],\end{aligned} \tag{42a}$$

$$\begin{aligned}\ddddot{S}(1/2) &= 24[S_o'(a)\ddot{a} + S_o''(a)\dot{a}^2 - S_o'(d)\ddot{d}] \\ &\quad + 8[S_o'(a)\ddot{a} + 3S_o''(a)\dot{a}\ddot{a} + S_o'''(a)\dot{a}^3]) \\ &\quad + (1/2)[S_o'(a)\dddot{a} + 4S_o''(a)\dot{a}\ddot{a} + 3S_o''(a)\ddot{a}^2 + 6S_o'''(a)\dot{a}^2\ddot{a} + S_o''''(a)\dot{a}^4] \\ &\quad + (1/2)[S_o'(d)\dddot{d} + 3S_o''(d)\ddot{d}^2],\end{aligned} \tag{42b}$$

where the right hand sides of both equations are evaluated at $c = 1/2$, using the equations (24), (30), (31), (32), (35), and (36).

In this and the previous five subsections we have derived relations between certain derivatives, which will be used in the next subsection and in the following section.

3.7. 2-star densities. We illustrate the power of these results by analyzing how the density τ_2 of 2-stars behaves near the phase transition. The parameters of the optimizing graphons are not directly observable, and a skeptic might argue that non-analyticity of these parameters does not prove the existence of a phase transition. However, τ_2 is an observable, and we shall see that its derivatives with respect to (ε, τ) change discontinuously at the transition.

For any graphon g, the 2-star density is

$$\tau_2(g) = \iiint_{[0,1]^3} g(x,y)g(x,z)dx\,dy\,dz. \tag{43}$$

This is neatly expressed in terms of the degree function $D(x) = \int_0^1 g(x, y)dy$:

$$\tau_2(g) = \int_0^1 D(x)^2 dx. \tag{44}$$

For bipodal graphons, this works out to

$$\tau_2 = cD_1^2 + (1-c)D_2^2, \tag{45}$$

where

$$D_1 = ca + (1-c)d \quad \text{and} \quad D_2 = cd + (1-c)b. \tag{46}$$

When $c = 1/2$ and $a = b$, both D_1 and D_2 equal ε, so $\tau_2 - \varepsilon^2$ is identically zero in the symmetric bipodal phase II. To approximate $\tau_2 - \varepsilon^2$ on the other side of the phase transition, we simply take derivatives of τ_2 with respect to c and use a Taylor series. We have

$$\tau_2 = cD_1^2 + (1-c)D_2^2, \tag{47a}$$

$$\dot{\tau}_2 = D_1^2 - D_2^2 + 2cD_1\dot{D}_1 + 2(1-c)D_2\dot{D}_2, \tag{47b}$$

$$\ddot{\tau}_2 = 4D_1\dot{D}_1 - 4D_2\dot{D}_2 + 2c\dot{D}_1^2 + 2cD_1\ddot{D}_1 + 2(1-c)\dot{D}_2^2 + 2(1-c)D_2\ddot{D}_2. \tag{47c}$$

Evaluating at $c = 1/2$, using previously determined values of $\dot{a}$, $\ddot{a}$, $\dot{b}$, $\ddot{b}$, $\dot{d}$, and $\ddot{d}$, gives $\dot{\tau}_2 = 0$ and

$$\ddot{\tau}_2 = (2+\alpha)^2(d-a)^2/2. \tag{48}$$

Thus,

$$\begin{aligned} \tau_2 - \varepsilon^2 &= (2+\alpha)^2(d-a)^2\Big(c - \frac{1}{2}\Big)^2/4 + O((c-\tfrac{1}{2})^4) \\ &= -3(2+\alpha)^2(d-a)^2\ddot{S}/(2\dddot{S}) + O(\ddot{S}^2), \end{aligned} \tag{49}$$

when $\ddot{S} > 0$ and $\dddot{S} < 0$. If we imagine crossing from Phase II to Phase III at time $t = 0$ along a curve that intersects the phase boundary transversally, then $\tau_2 - \varepsilon^2$ will be zero for $t < 0$ and will grow linearly with t for $t > 0$.

4. Expansion near the triple point $(\varepsilon, \tau) = \left(\frac{1}{2}, \frac{1}{8}\right)$

The phase transition between phases II and III occurs where $\ddot{S} = 0$. This curve intersects the ER curve $\tau = \varepsilon^3$ (i.e., the boundary of phase I) at the triple point $(\varepsilon, \tau) = \left(\frac{1}{2}, \frac{1}{8}\right)$. A natural question is how the transition curve approaches that triple point. In this section we prove that it is tangent to second order to the Erdös–Rényi curve.

Theorem 4.1. *Suppose that optimizing graphons in a neighborhood of* $(\varepsilon, \tau) = \left(\frac{1}{2}, \frac{1}{8}\right)$ *are all bipodal, and hence that the boundary between phases* II *and* III *can be expressed by the condition* $\ddot{S} = 0$*. Then, on the restriction of the phase boundary to the neighborhood,* τ *can be expressed as a function of* ε*, and as* $\varepsilon \to \frac{1}{2}^+$, $\tau = \varepsilon^3 - 8\left(\varepsilon - \frac{1}{2}\right)^3 + O\left(\varepsilon - \frac{1}{2}\right)^4$.

Proof. Let

$$A(a,d) = \frac{1}{2}S_o''(a)(a-d)^2 - a(S_o'(a) - S_o'(d)), \tag{50a}$$

$$B(a,d) = -2(a+d)(S_o'(a) - S_o'(d)), \tag{50b}$$

$$C(a,d) = 4(S_o(a) - aS_o'(a) - S_o(d) + dS_o'(d)). \tag{50c}$$

Then

$$\alpha = \frac{-B}{2A}, \tag{51}$$

(evaluated as usual at $c = 1/2$) and our formula for $\ddot{S}$ works out to

$$\ddot{S} = A\alpha^2 + B\alpha + C = \frac{4AC - B^2}{4A}. \tag{52}$$

As long as $A \neq 0$ (i.e. $d \neq a$), the phase transition occurs precisely where the discriminant $\Delta = B^2 - 4AC$ vanishes.

We now do series expansions for A, B and C in powers of $\delta a := a - \frac{1}{2}$ and $\delta d := d - \frac{1}{2}$. We begin with the Taylor Series for $S_o(z)$ and its derivatives:

$$S_o(z) = \frac{\ln(2)}{2} - \sum_{n=0}^{\infty} \frac{4^n\left(z - \frac{1}{2}\right)^{2n+2}}{(n+1)(2n+1)}, \tag{53a}$$

$$S_o'(z) = -2\sum_{n=0}^{\infty} \frac{4^n\left(z - \frac{1}{2}\right)^{2n+1}}{2n+1}, \tag{53b}$$

$$S_o''(z) = -2\sum_{n=0}^{\infty} 4^n\left(z - \frac{1}{2}\right)^{2n}. \tag{53c}$$

(To derive these expressions, expand

$$S_o''(z) = \frac{-1}{2}\left(\frac{1}{z} + \frac{1}{1-z}\right) = \frac{-2}{1 - 4\left(z - \frac{1}{2}\right)^2} \tag{54}$$

as a geometric series, and then integrate term by term to obtain the series for $S_o'(z)$ and $S_o(z)$.) Plugging these expansions into the formula for A yields

$$A = -(\delta a - \delta d)^2 \sum_{n=0}^{\infty} 4^n (\delta a)^{2n} + (1 + 2\delta a) \sum_{n=0}^{\infty} \frac{4^n}{2n+1} (\delta a^{2n+1} - \delta d^{2n+1}) := \sum_{m=1}^{\infty} A_m, \tag{55}$$

where A_m is a homogeneous m-th order polynomial in δa and δd. The first few terms are

$$A_1 = \delta a - \delta d, \tag{56a}$$

$$A_2 = \delta a^2 - \delta d^2, \tag{56b}$$

$$A_3 = \frac{4}{3}(\delta a^3 - \delta d^3), \tag{56c}$$

$$A_4 = \frac{4}{3}(-\delta a^4 + 6\delta a^3 \delta d - 3\delta a^2 \delta d^2 - 2\delta a \delta d^3). \tag{56d}$$

We do similar expansions of B and C:

$$B = 4(a + \delta a + \delta d) \sum_{n=1}^{\infty} \frac{4^n (\delta a^{2n+1} - \delta d^{2n+1})}{2n+1} = \sum_{m=1}^{\infty} B_m, \tag{57a}$$

$$B_1 = 4(\delta a - \delta d), \tag{57b}$$

$$B_2 = 4(\delta a^2 - \delta d^2), \tag{57c}$$

$$B_3 = \frac{16}{3}(\delta a^3 - \delta d^3), \tag{57d}$$

$$B_4 = \frac{16}{3}(\delta a^4 + \delta a^3 \delta d - \delta a \delta d^3 - \delta d^4) \tag{57e}$$

and

$$C = 4 \sum_{n=0}^{\infty} \frac{4^n (\delta a^{2n+1} - \delta d^{2n+1})}{2n+1} + 4 \sum_{n=0}^{\infty} \frac{4^n (\delta a^{2n+2} - \delta d^{2n+2})}{n+1} = \sum_{m=1}^{\infty} C_m, \tag{58a}$$

$$C_1 = 4(\delta a - \delta d), \tag{58b}$$

$$C_2 = 4(\delta a^2 - \delta d^2), \tag{58c}$$

$$C_3 = \frac{16}{3}(\delta a^3 - \delta d^3), \tag{58d}$$

$$C_4 = 8(\delta a^4 - \delta d^4). \tag{58e}$$

Note that $4A_m = B_m = C_m$ when $m = 1$, 2, or 3. This implies that the discriminant Δ vanishes through 4th order, and the leading nonzero term is

$$\begin{aligned}\Delta_5 &= 2B_1B_4 + 2B_2B_3 - 4A_1C_4 - 4A_2C_3 - 4A_3C_2 - 4A_4C_1\\ &= 2B_1B_4 + 2B_2B_3 - B_1C_4 - B_2B_3 - B_3B_2 - 4A_4C_1\\ &= B_1(2B_4 - C_4 - 4A_4)\\ &= \frac{8B_1}{3}\left(3\delta a^4 - 8\delta a^3\delta d + 6\delta a^2\delta d^2 - \delta d^4\right)\\ &= \frac{32}{3}(\delta a - \delta d)^4(3\delta a + \delta d).\end{aligned} \tag{59}$$

In fact, all terms in the expansion of Δ are divisible by $(\delta a - \delta d)^4$, as can be seen by evaluating Δ and its first three derivatives with respect to δd at $\delta d = \delta a$. We can view $3\delta a + \delta d$ as the *leading* term in the expansion of $\frac{3\Delta}{32(\delta a - \delta d)^4}$. Setting $\Delta = 0$ then gives

$$\delta d = -3\delta a + O(\delta a^2). \tag{60}$$

Since

$$\varepsilon = \frac{a+d}{2}; \qquad \tau - \varepsilon^3 = \frac{(a-d)^3}{8}, \tag{61}$$

we have

$$\tau = \varepsilon^3 - 8\left(\varepsilon - \frac{1}{2}\right)^3 + O\left(\left(\varepsilon - \frac{1}{2}\right)^4\right), \tag{62}$$

as required. Note that these expressions only apply when $\delta a < 0$, i.e., when $d > a$, i.e., for $\varepsilon > \frac{1}{2}$. When $\varepsilon < \frac{1}{2}$, bipodal graphons with $\delta d = -3\delta a$ would lie *above* the ER curve. □

5. Numerical values of $\ddot{S}$ and $\dddot{S}$

We now evaluate $\ddot{S}$ and $\dddot{S}$ numerically, following the formulas given in (42), with help from formulas (24), (30), (31), (32), (35), and (36).

Figure 4 shows the boundary of the symmetrical bipodal phase II. For better visualization, we use the coordinate $(\varepsilon, \tau - \varepsilon^3)$ instead of (ε, τ) (in which the phase boundary curve bends too much to see the details). In these new coordinates, the ER $\tau - \varepsilon^3 = 0$ curve becomes horizontal, and is the upper boundary in the plot.

The part of the phase boundary on which $\ddot{S} = 0$, denoted by Σ, is illustrated with thick red line.

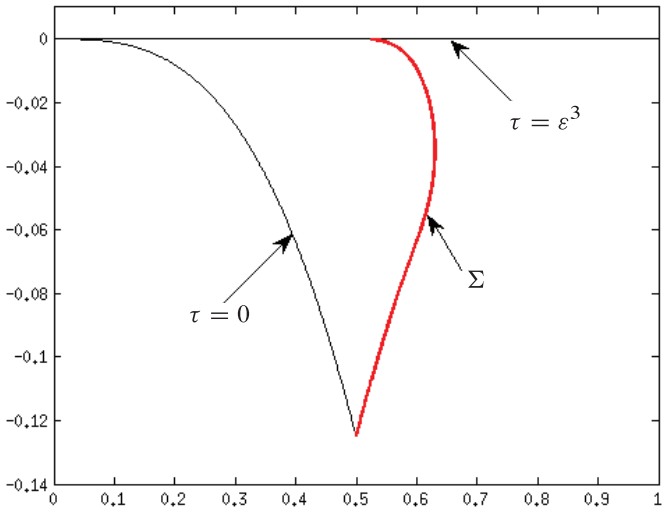

Figure 4. The boundary of the bipodal phase in the $(\varepsilon, \tau - \varepsilon^3)$ coordinate. The part of phase boundary on which $\ddot{S} = 0$ is shown in the thick red line. Note how this line is tangent to the ER curve at the triple point, as required by Theorem 4.1.

In Figure 5, we show the values of $\ddot{S}$ and $\dddot{S}$ on a tube around the phase transition curve Σ. The plot is again in the $(\varepsilon, \tau - \varepsilon^3)$ coordinate. To better visualize the transition, we also plot the functions $\mathrm{sgn}(\ddot{S})$ and $\mathrm{sgn}(\dddot{S})$, where

$$\mathrm{sgn}(x) = \begin{cases} 1, & x \geq 0, \\ -1, & x < 0, \end{cases}$$

is the sign function. Note how $\dddot{S}$ is always negative along Σ. Near the transition, we should thus expect optimizing graphons to have $|c - \frac{1}{2}| \approx \sqrt{-6\ddot{S}\left(\frac{1}{2}\right)/\dddot{S}\left(\frac{1}{2}\right)}$ when $\ddot{S}\left(\frac{1}{2}\right) > 0$, and to have $c = \frac{1}{2}$ when $\ddot{S}\left(\frac{1}{2}\right) < 0$. In the next section we confirm this prediction with direct sampling of graphons.

The function $\dfrac{\partial \ddot{S}}{\partial \varepsilon}$ along Σ, expressed as a function of $\tau - \varepsilon^3$, is shown separately in Figure 6. Since $\left(c - \frac{1}{2}\right)^2$ is proportional to $-\ddot{S}/\dddot{S}$, this gives a measure of how quickly $\left|c - \frac{1}{2}\right|$ grows as we cross Σ from left to right. This sensitivity is greatest near the bottom of Σ, and approaches zero as we approach the triple point.

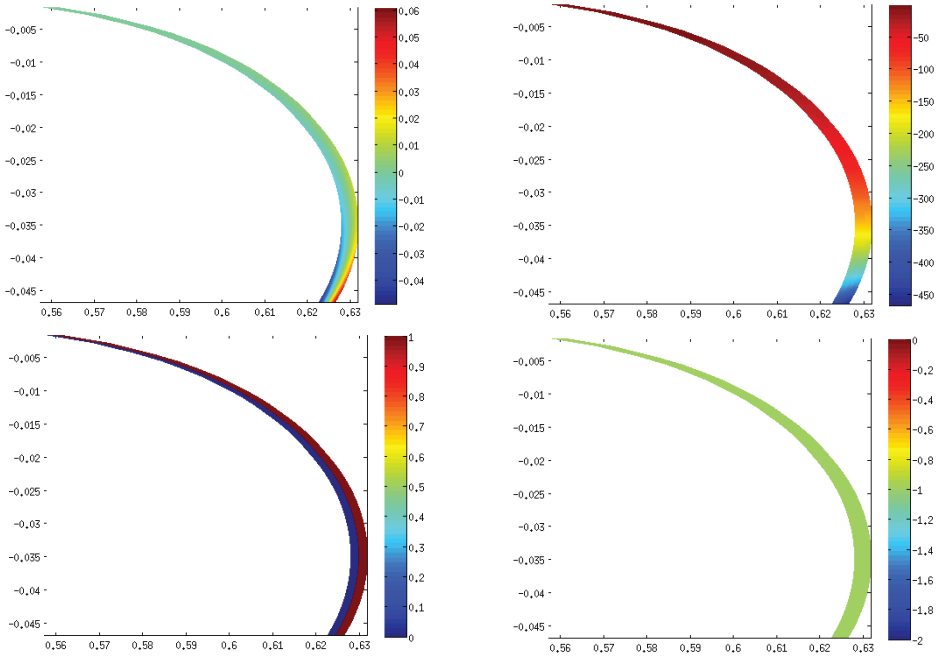

Figure 5. The functions $\ddot{S}$ (top left), $\mathrm{sgn}(\ddot{S})$ (bottom left), $\dddot{S}$ (top right) and $\mathrm{sgn}(\dddot{S})$ (bottom right) in a neighborhood of the phase transition curve Σ.

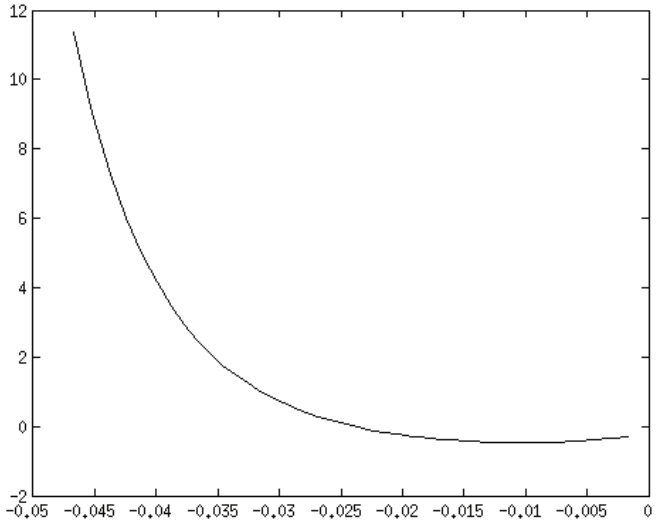

Figure 6. The plot of $\dfrac{\partial \ddot{S}}{\partial \varepsilon}$ on the curve Σ, i.e. as a function of $\tau - \varepsilon^3$.

6. Comparison to numerical sampling

In this section, we perform some numerical simulations using the sampling algorithm we developed in [20], and compare them to the results of the perturbation analysis. The algorithm samples random values of $\mathcal{S}$ in the parameter space of graphons (with constrained values of (ε, τ)), and then takes the maximum of the sampled values. This sampling algorithm is extremely expensive computationally, but when sufficiently large sample sizes are reached, we can achieve desired accuracy; see the discussions in [20]. We emphasize that our sampling algorithm does not assume bipodality of the maximizing graphons. In fact, we always start with the assumption that the graphons are 16-podal. Bipodal structure is deduced when all but two clusters have zero size.

In Figure 7 we show how a density fails to be analytic in (ε, τ) near the transition between phase II and phase III. We plot the sampled values of $\tau_2 - \varepsilon^2$ as a function of ε for two fixed values of $\tau - \varepsilon^3$, where τ_2 is the density of two-stars. Within phase II, $\tau_2 - \varepsilon^2 = 0$ is identically zero. Outside of phase II, we see that $\tau_2 - \varepsilon^2$ grows linearly with distance to Σ, as suggested by equation (49).

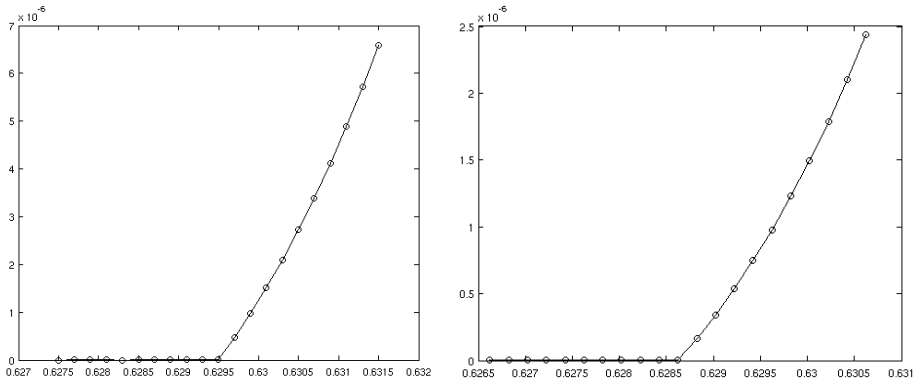

Figure 7. The value of $\tau_2 - \varepsilon^2$ as a function of ε at two different τ values, 0.2184 (left) and 0.2147 (right), cross the boundary of phase II.

In Figure 8, we plot the values of $\left|c - \frac{1}{2}\right|$ for optimizing graphons (found with the numerical sampling algorithm) and the real part of $\sqrt{-6\ddot{S}\left(\frac{1}{2}\right)/\dddot{S}\left(\frac{1}{2}\right)}$ (given by the perturbation analysis) as functions of ε for two different values of $\tau - \varepsilon^3$. As should be expected, the perturbative calculation gives a very good fit when we are reasonably close to the phase transition curve Σ, but starts to be less accurate when we go farther out. A similar plot for the corresponding values of a, b, and d are shown in Figure 9. The parameters a and b show square root singularities,

just like c, but d does not. This is to be expected from the fact that $\dot{d}(1/2) = 0$, while $\dot{a}(1/2)$ and $\dot{b}(1/2)$ are nonzero.

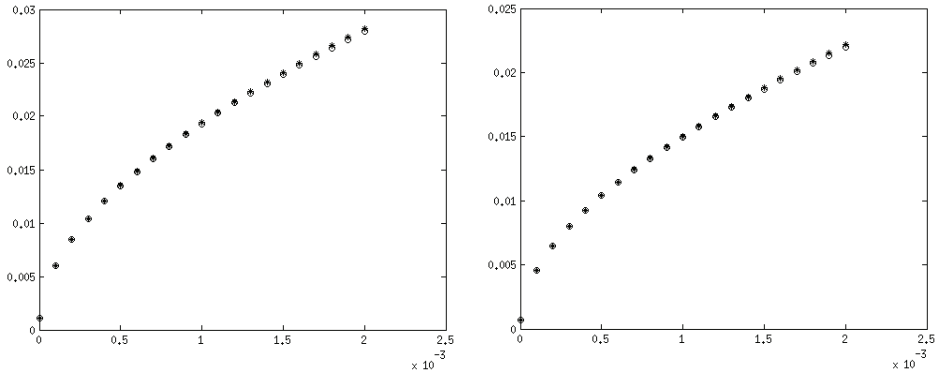

Figure 8. Comparison between the curve $\left(\varepsilon, |c - \frac{1}{2}|\right)$ (in $\circ$, given by the sampling algorithm) and $\left(\varepsilon, \sqrt{-6\ddot{S}\left(\frac{1}{2}\right)/\dddot{S}\left(\frac{1}{2}\right)}\right)$ (in $*$, given by the perturbation calculation) at two different $\tau - \varepsilon^3$ values: -0.0348 (left), -0.0300 (right).

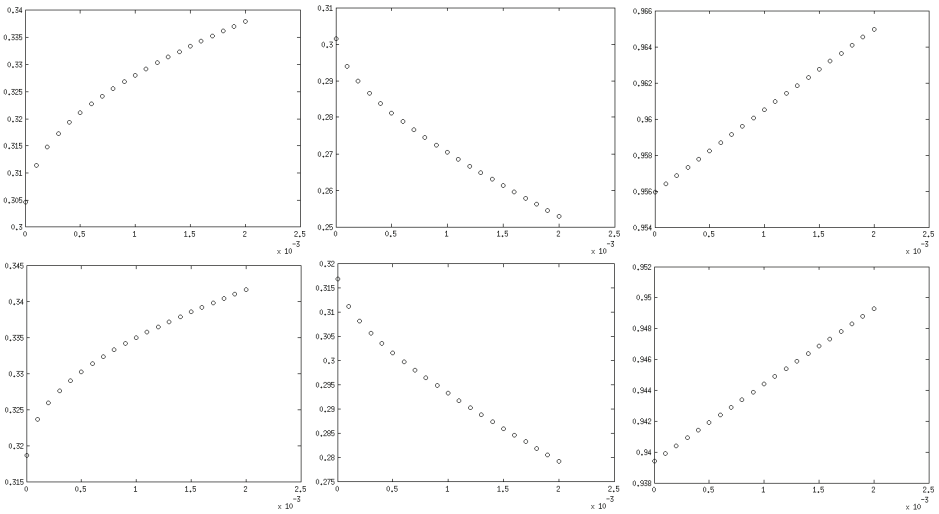

Figure 9. The plots of a (left), b (middle) and d (right) versus $\varepsilon - \varepsilon_0$ at two different values of τ: 0.2147 (top row) and 0.2184 (bottom row).

Finally, we show in Figure 10 some typical graphons that we obtained using the sampling algorithm. We emphasize again that in the sampling algorithm, we did not assume bipodality of the optimizing graphons. The numerics indicate that the optimizing graphons are really bipodal close to (on both sides) the phase

transition curve Σ. These serve as numerical evidence to justify the perturbation calculations in this work.

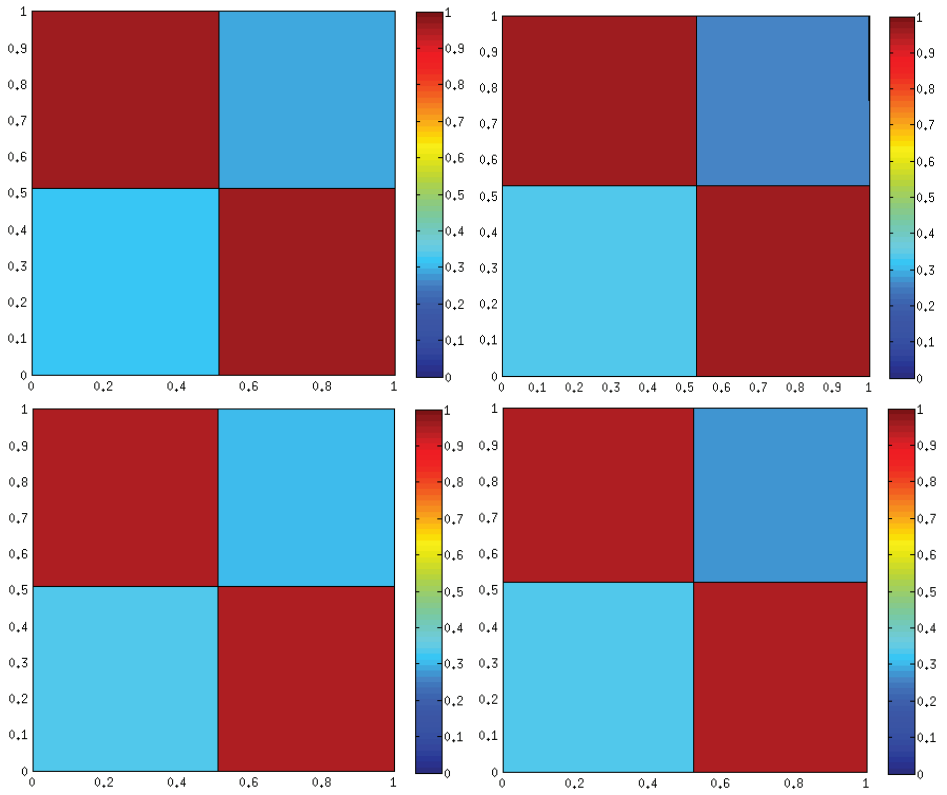

Figure 10. Typical graphons at points close to the phase boundary. The values of (ε, τ) are (from top left to bottom right): $(0.6299, 0.2147)$, $(0.6315, 0.2147)$, $(0.6290, 0.2184)$ and $(0.6306, 0.2184)$.

7. Conclusion

Our analysis began with the assumption that all entropy optimizers for relevant constraint parameters are multipodal; in fact bipodal, but we emphasize the more general feature. Multipodality is a fundamental aspect of the asymptotics of constrained graphs. It is the embodiment of phase emergence: the set of graphs with fixed constraints develops ("emerges into") a well-defined global state as the number of nodes diverges, *by partitioning the set of all nodes into a finite* (*usually small*) *number of equivalent parts,* $\{P_1, P_2, \ldots\}$, of relative sizes c_j, and

uniform (i.e. constant) probability p_{ij} of an edge between nodes in P_i and P_j. This unexpected fact has been seen in all simulations, and was actually proven throughout the phase spaces of edge/k-star models [7] and in particular regions of a wide class of models [8]. In this paper we are taking this as given, for a certain region of the edge/triangle phase space; we are assuming that our large constrained graphs emerge into such global states, and not just multipodal states but bipodal states, in accordance with simulation of the relevant constraint values in this model.

We analyzed the specific constraints of edge density approximately $1/2$ and triangle density less than $1/8$, and drew a variety of conclusions. First we used the facts, proven in [20], that there is a smooth curve in the phase space, indicated in Figure 1, such that: (i) to the left of the curve the (reduced) entropy optimizer is unique and is given by (2); (ii) the entropy optimizer is also unique to the right of the curve but is no longer symmetric. The curve thus represents a boundary between distinct phases of different symmetry. These facts were established in [20]. The analysis in this paper concerns the details of this transition. Here we want to speak to the significance of the results, in terms of symmetry breaking.

We take the viewpoint that the global states for these constraint values are sitting in a 4 dimensional space given by a, b, d and c, quantities which describe the laws of interaction of the elements of the parts P_j into which the node set is partitioned. (The notion of 'vertex type' has been introduced and analyzed in [10] and is useful in more general contexts.)

The symmetry $c_1 = c_2 = 1/2$ and $a = b$ is therefore a symmetry of the *rules* by which the global state is produced. Each part P_j into which the set of nodes is partitioned can be thought of as embodying the symmetry between the nodes it contains, but the symmetry of the global states is a higher-level symmetry, between the way these equivalence classes P_j are connected. In this way the transition studied in this paper is an analogue of the symmetry-breaking fluid/crystal transition studied in the statistical analysis of matter in thermal equilibrium [27].

One consequence is a contribution to the problem concerning the fluid/crystal transition. It has been known experimentally since the work of PW Bridgman in the early twentieth century that no matter how one varies the thermodynamic parameters (say mass and energy density) between fluid and crystal phases one *must* go through a singularity or phase transition. An influential theoretical analysis by L. Landau attributed this basic fact to the difference in symmetry between the crystal and fluid states, but this argument has never been completely convincing [18]. In our model this fact follows for phases II and III from the two-

step process of first simplifying the space of global states to a finite dimensional space (of interaction rules) and then realizing the symmetry as acting in that space; it is then immediate that phases with different symmetry cannot be linked by a smooth (analytic) curve. This is one straightforward consequence of the powerful advantage of understanding our global states as multipodal, when compared say with equilibrium statistical mechanics.

Appendix A. Calculational details

The calculation of our results is very long, with many opportunities for error, so to aid any reader who doubts its correctness we include in these appendices those details which were left out of the main text for readability.

A.1. Derivatives of τ. We compute the triangle density and its derivatives as follows:

$$\tau = c^3a^3 + 3c^2(1-c)ad^2 + 3c(1-c)^2bd^2 + (1-c)^3b^3, \tag{63a}$$

$$\begin{aligned}\dot{\tau} = {} & c^3\frac{d}{dc}(a^3) + 3c^2(1-c)\frac{d}{dc}(ad^2) + 3c(1-c)^2\frac{d}{dc}(bd^2) + (1-c)^3\frac{d}{dc}(b^3)\\ & + 3c^2a^3 + 3(2c-3c^2)ad^2 + 3(1-4c+3c^2)bd^2 - 3(1-c)^2b^3,\end{aligned} \tag{63b}$$

$$\begin{aligned}\ddot{\tau} = {} & c^3\frac{d^2}{dc^2}(a^3) + 3c^2(1-c)\frac{d^2}{dc^2}(ad^2) + 3c(1-c)^2\frac{d^2}{dc^2}(bd^2)\\ & + (1-c)^3\frac{d^2}{dc^2}(b^3) + 6c^2\frac{d}{dc}(a^3) + 6(2c-3c^2)\frac{d}{dc}(ad^2)\\ & + 6(1-4c+3c^2)\frac{d}{dc}(bd^2) - 6(1-c)^2\frac{d}{dc}(b^3)\\ & + 6ca^3 + 6(1-3c)ad^2 + 6(3c-2)bd^2 + 6(1-c)b^3,\end{aligned} \tag{63c}$$

$$\begin{aligned}\dddot{\tau} = {} & c^3\frac{d^3}{dc^3}(a^3) + 3c^2(1-c)\frac{d^3}{dc^3}(ad^2) + 3c(1-c)^2\frac{d^3}{dc^3}(bd^2)\\ & + (1-c)^3\frac{d^3}{dc^3}(b^3) + 9c^2\frac{d^2}{dc^2}(a^3) + 9(2c-3c^2)\frac{d^2}{dc^2}(ad^2)\\ & + 9(1-4c+3c^2)\frac{d^2}{dc^2}b(d^2) - 9(1-c)^2\frac{d^2}{dc^2}(b^3) + 18c\frac{d}{dc}(a^3)\\ & + 18(1-3c)\frac{d}{dc}(ad^2) + 18(3c-2)\frac{d}{dc}(bd^2) + 18(1-c)\frac{d}{dc}(b^3)\\ & + 6a^3 - 18ad^2 + 18bd^2 - 6b^3,\end{aligned} \tag{63d}$$

$$\begin{aligned}\dddot{\tau} = {} & c^3\frac{d^4}{dc^4}(a^3) + 3c^2(1-c)\frac{d^4}{dc^4}(ad^2) + 3c(1-c)^2\frac{d^4}{dc^4}(bd^2) \\ & + (1-c)^3\frac{d^4}{dc^4}(b^3) + 12c^2\frac{d^3}{dc^3}(a^3) + 12(2c-3c^2)\frac{d^3}{dc^3}(ad^2) \\ & + 12(1-4c+3c^2)\frac{d^3}{dc^3}(bd^2) - 12(1-c)^2\frac{d^3}{dc^3}(b^3) \\ & + 36c\frac{d^2}{dc^2}(a^3) + 36(1-3c)\frac{d^2}{dc^2}(ad^2) \\ & + 36(3c-2)\frac{d^2}{dc^2}(bd^2) + 36(1-c)\frac{d^2}{dc^2}(b^3) + 24\frac{d}{dc}(a^3) \\ & - 72\frac{d}{dc}(ad^2) + 72\frac{d}{dc}(bd^2) - 24\frac{d}{dc}(b^3).\end{aligned} \tag{63e}$$

Once again we evaluate at $c = 1/2$, making use of symmetry to simplify terms:

$$\begin{aligned}\dot{\tau}(1/2) &= (1/8)\Big[\frac{d}{dc}(a^3) + 3\frac{d}{dc}(ad^2) + 3\frac{d}{dc}(bd^2) + \frac{d}{dc}(b^3)\Big] \\ &\quad + (3/4)(a^3 + ad^2 - bd^2 - b^3) \\ &= (1/8)\Big[\frac{d}{dc}(a^3) + 3\frac{d}{dc}(ad^2) + 3\frac{d}{dc}(bd^2) + \frac{d}{dc}(b^3)\Big] \\ &= \frac{3}{8}[(a^2+d^2)(\dot{a}+\dot{b}) + 4ad\dot{d}] \\ &= 0\end{aligned} \tag{64}$$

where all quantities on the right hand side are evaluated at $c = 1/2$. Continuing to higher derivatives,

$$\begin{aligned}\ddot{\tau}(1/2) &= (1/8)\Big[\frac{d^2}{dc^2}(a^3) + 3\frac{d^2}{dc^2}(ad^2) + 3\frac{d^2}{dc^2}(bd^2) + \frac{d^2}{dc^2}(b^3)\Big] \\ &\quad + (3/2)\Big[\frac{d}{dc}(a^3) + \frac{d}{dc}(ad^2) - \frac{d}{dc}(bd^2) - \frac{d}{dc}(b^3)\Big] \\ &\quad + 3[a^3 - ad^2 - bd^2 + b^3] \\ &= (1/4)\Big[\frac{d^2}{dc^2}(a^3) + 3\frac{d^2}{dc^2}(ad^2)\Big] + 3\Big[\frac{d}{dc}(a^3) + \frac{d}{dc}(ad^2)\Big] \\ &\quad + 6(a^3 - ad^2),\end{aligned} \tag{65}$$

$$\dddot{\tau} = 0, \tag{66}$$

$$
\begin{aligned}
\dddot{\tau}(1/2) &= (1/8)\Big[\frac{d^4}{dc^4}(a^3) + 3\frac{d^4}{dc^4}(ad^2) + 3\frac{d^4}{dc^4}(bd^2) + \frac{d^4}{dc^4}(b^3)\Big] \\
&\quad + 3\Big[\frac{d^3}{dc^3}(a^3) + \frac{d^3}{dc^3}(ad^2) - \frac{d^3}{dc^3}(bd^2) - \frac{d^3}{dc^3}(b^3)\Big] \\
&\quad + 18\Big[\frac{d^2}{dc^2}(a^3) - \frac{d^2}{dc^2}(ad^2) - \frac{d^2}{dc^2}(bd^2) + \frac{d^2}{dc^2}(b^3)\Big] \\
&\quad + 24\Big[\frac{d}{dc}(a^3) - 3\frac{d}{dc}(ad^2) + 3\frac{d}{dc}(bd^2) - \frac{d}{dc}(b^3)\Big] \\
&= (1/4)\Big[\frac{d^4}{dc^4}(a^3) + 3\frac{d^4}{dc^4}(ad^2)\Big] + 6\Big[\frac{d^3}{dc^3}(a^3) + \frac{d^3}{dc^3}(ad^2)\Big] \\
&\quad + 36\Big[\frac{d^2}{dc^2}(a^3) - \frac{d^2}{dc^2}(ad^2)\Big] + 48\Big[\frac{d}{dc}(a^3) - 3\frac{d}{dc}(ad^2)\Big].
\end{aligned}
\tag{67}
$$

To continue further we must expand the derivatives of a^3 and ad^2:

$$\frac{d}{dc}(a^3) = 3a^2\dot{a}, \tag{68a}$$

$$\frac{d^2}{dc^2}(a^3) = 3a^2\ddot{a} + 6a\dot{a}^2, \tag{68b}$$

$$\frac{d^3}{dc^3}(a^3) = 3a^2\dddot{a} + 18a\dot{a}\ddot{a} + 6\dot{a}^3, \tag{68c}$$

$$\frac{d^4}{dc^4}(a^3) = 3a^2\ddddot{a} + 24a\dot{a}\dddot{a} + 18a\ddot{a}^2 + 36\dot{a}^2\ddot{a}, \tag{68d}$$

and

$$\frac{d}{dc}(ad^2) = \dot{a}d^2 + 2ad\dot{d}, \tag{69a}$$

$$\frac{d^2}{dc^2}(ad^2) = \ddot{a}d^2 + 4\dot{a}d\dot{d} + 2a\dot{d}^2 + 2ad\ddot{d}, \tag{69b}$$

$$\frac{d^3}{dc^3}(ad^2) = \dddot{a}d^2 + 6\ddot{a}d\dot{d} + 6\dot{a}d\ddot{d} + 6\dot{a}\dot{d}^2 + 2ad\dddot{d} + 6a\dot{d}\ddot{d}, \tag{69c}$$

$$
\begin{aligned}
\frac{d^4}{dc^4}(ad^2) &= \ddddot{a}d^2 + 8\dddot{a}d\dot{d} + 12\ddot{a}d\ddot{d} + 12\ddot{a}\dot{d}^2 \\
&\quad + 8\dot{a}d\dddot{d} + 24\dot{a}\dot{d}\ddot{d} + 2ad\ddddot{d} + 8a\dot{d}\dddot{d} + 6a\ddot{d}^2.
\end{aligned}
\tag{69d}
$$

At $c = 1/2$, $\dot{d} = \dddot{d} = 0$, so this simplifies to

$$\frac{d}{dc}(ad^2)(1/2) = \dot{a}d^2, \tag{70a}$$

$$\frac{d^2}{dc^2}(ad^2)(1/2) = \ddot{a}d^2 + 2ad\ddot{d}, \tag{70b}$$

$$\frac{d^3}{dc^3}(ad^2)(1/2) = \dddot{a}d^2 + 6\dot{a}d\ddot{d}, \tag{70c}$$

$$\frac{d^4}{dc^4}(ad^2)(1/2) = \ddddot{a}d^2 + 12\ddot{a}d\ddot{d} + 6a\ddot{d}^2 + 2ad\dddot{d}, \tag{70d}$$

all evaluated at $c = 1/2$.

Plugging back in, this yields

$$\begin{aligned}
\ddot{\tau}(1/2) &= (1/4)\Big[\frac{d^2}{dc^2}(a^3) + 3\frac{d^2}{dc^2}(ad^2)\Big] + 3\Big[\frac{d}{dc}(a^3) + \frac{d}{dc}(ad^2)\Big] \\
&\quad + 6(a^3 - ad^2) \\
&= (1/4)[3a^2\ddot{a} + 6a\dot{a}^2 + 3\ddot{a}d^2 + 6ad\ddot{d}] + 3[3a^2\dot{a} + \dot{a}d^2] \\
&\quad + 6a(a^2 - d^2) \\
&= (3/4)[(a^2 + d^2)\ddot{a} + 2ad\ddot{d}] + (3/2)a\dot{a}^2 + 3(3a^2 + d^2)\dot{a} \\
&\quad + 6a(a^2 - d^2)
\end{aligned} \tag{71}$$

and

$$\begin{aligned}
\dddot{\tau}(1/2) &= (1/4)\Big[\frac{d^4}{dc^4}(a^3) + 3\frac{d^4}{dc^4}(ad^2)\Big] + 6\Big[\frac{d^3}{dc^3}(a^3) + \frac{d^3}{dc^3}(ad^2)\Big] \\
&\quad + 36\Big[\frac{d^2}{dc^2}(a^3) - \frac{d^2}{dc^2}(ad^2)\Big] + 48\Big[\frac{d}{dc}(a^3) - 3\frac{d}{dc}(ad^2)\Big] \\
&= (1/4)[3a^2\dddot{a} + 24a\dot{a}\ddot{a} + 18a\ddot{a}^2 + 36\dot{a}^2\ddot{a} \\
&\quad + 3d^2\dddot{a} + 36d\ddot{d}\ddot{a} + 18a\ddot{d}^2 + 6ad\dddot{d}] \\
&\quad + 6[3a^2\ddot{a} + 18a\dot{a}\ddot{a} + 6\dot{a}^3 + \ddot{a}d^2 + 6\dot{a}d\ddot{d}] \\
&\quad + 36[3a^2\ddot{a} + 6a\dot{a}^2 - d^2\ddot{a} - 2ad\ddot{d}] \\
&\quad + 48[3a^2\dot{a} - 3d^2\dot{a}],
\end{aligned} \tag{72}$$

where all quantities on the right hand side of these equations are evaluated at $c = 1/2$.

A.2. Partial derivatives of $f(a,b,c,d)$. This is a long but straightforward exercise in calculus:

$$\begin{aligned} f &= S_o'(a)[cd(a-d)-(1-c)b(b-d)] \\ &\quad + S_o'(b)[ca(a-d)-(1-c)d(b-d)] \\ &\quad + S_o'(d)[c(d^2-a^2)+(1-c)(b^2-d^2)]; \end{aligned} \tag{73}$$

$$\begin{aligned} f_a &= S_o''(a)[cd(a-d)-(1-c)b(b-d)] \\ &\quad + S_o'(a)(cd)+S_o'(b)(2ac-cd)+S_o'(d)(-2ac), \end{aligned} \tag{74a}$$

$$\begin{aligned} f_b &= S_o''(b)[ca(a-d)-(1-c)d(b-d)] \\ &\quad + S_o'(a)[(1-c)d-2(1-c)b]+S_o'(b)[-(1-c)d]+S_o'(d)[2(1-c)b], \end{aligned} \tag{74b}$$

$$\begin{aligned} f_c &= S_o'(a)[ad-d^2+b^2-bd]+S_o'(b)[a^2-ad+bd-d^2] \\ &\quad + S_o'(d)[2d^2-a^2-b^2] \end{aligned} \tag{74c}$$

$$\begin{aligned} f_d &= S_o''(d)[c(d^2-a^2)+(1-c)(b^2-d^2)] \\ &\quad + S_o'(a)[-2cd+ac+(1-c)b]+S_o'(b)[-ac+2(1-c)d-(1-c)b] \\ &\quad + S_o'(d)[2cd-2(1-c)d]; \end{aligned} \tag{74d}$$

$$f_{ac} = S_o''(a)[ad-d^2+b^2-bd]+dS_o'(a)+(2a-d)S_o'(b)-2aS_o'(d), \tag{75a}$$

$$f_{bc} = S_o''(b)[a(a-d)+d(b-d)]+S_o'(a)(-d+2b)+S_o'(b)d-2bS_o'(d), \tag{75b}$$

$$\begin{aligned} f_{dc} &= S_o''(d)[2d^2-a^2-b^2]+S_o'(a)[a-b-2d] \\ &\quad + S_o'(b)[-a-2d+b]+S_o'(d)4d, \end{aligned} \tag{75c}$$

$$\begin{aligned} f_{aa} &= S_o'''(a)[cd(a-d)-(1-c)b(b-d)]+S_o''(a)(2cd) \\ &\quad + 2cS_o'(b)-2cS_o'(d), \end{aligned} \tag{75d}$$

$$\begin{aligned} f_{bb} &= S_o'''(b)[ca(a-d)-(1-c)d(b-d)] \\ &\quad + S_o''(b)(-2(1-c)d)-2(1-c)S_o'(a)+2(1-c)S_o'(d), \end{aligned} \tag{75e}$$

$$f_{ad} = S_o''(a)[-2cd + ca + (1-c)b] + cS_o'(a) - cS_o'(b) - 2acS_o''(d), \quad (75f)$$

$$f_{bd} = S_o''(b)[-ac + 2(1-c)d - (1-c)b] + (1-c)S_o'(a) - (1-c)S_o'(b) + 2(1-c)bS_o''(d); \quad (75g)$$

$$f_{caa} = S_o'''(a)[d(a-d) + b(b-d)] + 2dS_o''(a) + 2S_o'(b) - 2S_o'(d), \quad (76a)$$

$$f_{cab} = S_o''(a)(2b-d) + (2a-d)S_o''(b), \quad (76b)$$

$$f_{cbb} = S_o'''(b)[a(a-d) + d(b-d)] + 2dS_o''(b) + 2S_o'(a) - 2S_o'(d), \quad (76c)$$

$$f_{aaa} = S_o''''(a)[cd(a-d) - (1-c)b(b-d)] + 3cdS_o'''(a), \quad (76d)$$

$$f_{aab} = S_o'''(a)[-2(1-c)b + (1-c)d] + 2cS_o''(b), \quad (76e)$$

$$f_{abb} = S_o'''(b)(2ac - dc) - 2(1-c)S_o''(a), \quad (76f)$$

$$f_{bbb} = S_o''''(b)[ca(a-d) - (1-c)d(b-d)] - 3(1-c)dS_o'''(b). \quad (76g)$$

We now compute the relevant terms at $c = 1/2$. (The right hand side of equations (77-36) are all intended to be evaluated at $c = 1/2$.) We have

$$\begin{aligned} f_a(1/2) &= S_o''(a)[(1/2)d(a-d) - (1/2)a(a-d)] + (1/2)dS_o'(a) \\ &\quad + (a - d/2)S_o'(a) - aS_o'(d) \\ &= S_o''(a)[-(1/2)(a-d)^2] + a[S_o'(a) - S_o'(d)] \end{aligned} \quad (77a)$$

and

$$\begin{aligned} f_b(1/2) &= S_o''(a)[(1/2)a(a-d) - (1/2)d(a-d)] + S_o'(a)[(d/2) - b] \\ &\quad + S_o'(a)[-(d/2)] + bS_o'(d) \\ &= (1/2)S_o''(a)(a-d)^2 - a[S_o'(a) - S_o'(d)] \\ &= -f_a(1/2), \end{aligned} \quad (77b)$$

so

$$f_b(1/2) - f_a(1/2) = S_o''(a)(a-d)^2 - 2a[S_o'(a) - S_o'(d)]. \quad (78)$$

Note that this is the same as the denominator in the formula (24) for $\dot{a}(1/2)$.

We have

$$f_{ca}(1/2) = f_{cb}(1/2) = S_o''(a)(a^2 - d^2) + 2a[S_o'(a) - S_o'(d)], \tag{79a}$$

$$f_{cd}(1/2) = 2S_o''(d)(d^2 - a^2) + 4d[S_o'(d) - S_o'(a)], \tag{79b}$$

$$\begin{aligned} f_{caa}(1/2) = f_{cbb}(1/2) = S_o'''(a)[a^2 - d^2] + 2dS_o''(a) \\ + 2[S_o'(a) - S_o'(d)], \end{aligned} \tag{79c}$$

$$f_{cab}(1/2) = 2S_o''(a)(2a - d), \tag{79d}$$

$$\begin{aligned} f_{caa}(1/2) - 2f_{cab}(1/2) + f_{cbb}(1/2) = 2S_o'''(a)(a^2 - d^2) + 8(d - a)S_o''(a) \\ + 4[S_o'(a) - S_o'(d)]; \end{aligned} \tag{79e}$$

$$f_{aa}(1/2) = -f_{bb}(1/2) = S_o'''(a)[-(1/2)(a - d)^2] + dS_o''(a) + S_o'(a) - S_o'(d), \tag{80a}$$

$$f_{aa}(1/2) - f_{bb}(1/2) = -(a - d)^2 S_o'''(a) + 2dS_o''(a) + 2[S_o'(a) - S_o'(d)], \tag{80b}$$

$$f_{ad}(1/2) = -f_{bd}(1/2) = (a - d)S_o''(a) - aS_o''(d), \tag{80c}$$

$$f_{ad}(1/2) - f_{bd}(1/2) = 2(a - d)S_o''(a) - 2aS_o''(d); \tag{80d}$$

$$f_{aaa}(1/2) = -f_{bbb}(1/2) = S_o''''(a)[-(1/2)(a - d)^2] + (3/2)dS_o'''(a), \tag{81a}$$

$$f_{aab}(1/2) = -f_{abb}(1/2) = (1/2)S_o'''(a)(d - 2a) + S_o''(a), \tag{81b}$$

$$(f_{aaa} - 3f_{aab} + 3f_{abb} - f_{bbb})(1/2) = -S_o''''(a)(a - d)^2 + 6aS_o'''(a) - 6S_o''(a). \tag{81c}$$

References

[1] C. Borgs, J. Chayes, and L. Lovász, Moments of two-variable functions and the uniqueness of graph limits. *Geom. Funct. Anal.* **19** (2010), no. 6, 1597–1619. MR 2594615 Zbl 1223.05193

[2] C. Borgs, J. Chayes, L. Lovász, V.T. Sós, and K. Vesztergombi, Convergent sequences of dense graphs. I. Subgraph frequencies, metric properties and testing. *Adv. Math.* **219** (2008), no. 6, 1801–1851. MR 2455626 Zbl 1213.05161

[3] S. Chatterjee, *Large deviations for random graphs.* Lecture notes from the 45th Probability Summer School held in Saint-Flour, June 2015. Lecture Notes in Mathematics, 2197. École d'Été de Probabilités de Saint-Flour. Springer, Cham, 2017. MR 3700183 Zbl 1375.60009

[4] S. Chatterjee and P. Diaconis, Estimating and understanding exponential random graph models. *Ann. Statist.* **41** (2013), no. 5, 2428–2461. MR 3127871 Zbl 1293.62046

[5] S. Chatterjee and S. R. S. Varadhan, The large deviation principle for the Erdős–Rényi random graph. *European J. Combin.* **32** (2011), no. 7, 1000–1017. MR 2825532 Zbl 1230.05259

[6] R. Eldan and R. Gross, Exponential random graphs behave like mixtures of stochastic block models. Preprint, 2017. arXiv:1707.01227v2 [math.PR]

[7] R. Kenyon, C. Radin, K. Ren, and L. Sadun, Multipodal structures and phase transitions in large constrained graphs. *J. Stat. Phys.* **168** (2017), no. 2, 233–258. MR 3667360

[8] R. Kenyon, C. Radin, K. Ren, and L. Sadun, Bipodal structure in oversaturated random graphs. *Int. Math. Res. Not. IMRN* **2018**, no. 4, 1009–1044.

[9] R. Kenyon, C. Radin, K. Ren and L. Sadun, The phases of large networks with edge and triangle constraints. *J. Phys. A* **50** (2017), no. 43, 435001, 22 pp. MR 3718855 Zbl 06814247

[10] H. Koch, Vertex order in some large constrained random graphs. *SIAM J. Math. Anal.* **48** (2016), no. 4, 2588–2601. MR 3537007 Zbl 1344.05125

[11] L. Lovász and B. Szegedy, Limits of dense graph sequences. *J. Combin. Theory Ser. B* **96** (2006), no. 6, 933–957. MR 2274085 Zbl 1113.05092

[12] L. Lovász and B. Szegedy, Szemerédi's lemma for the analyst. *Geom. Funct. Anal.* **17** (2007), no. 1, 252–270. MR 2306658 Zbl 1123.46020

[13] L. Lovász and B. Szegedy, Finitely forcible graphons. *J. Combin. Theory Ser. B* **101** (2011), no. 5, 269–301. MR 2802882 Zbl 223.05248

[14] L. Lovász, *Large networks and graph limits.* American Mathematical Society Colloquium Publications, 60. American Mathematical Society, Providence, R.I., 2012. MR 3012035 Zbl 1292.05001

[15] E. Lubetzky and Y. Zhao, On replica symmetry of large deviations in random graphs. *Random Structures Algorithms* **47** (2015), no. 1, 109–146. MR 3366814 Zbl 1348.05195

[16] M. E. J. Newman, *Networks.* An Introduction. Oxford University Press, Oxford, 2010. MR 2676073 Zbl 1195.94003

[17] J. Park and M. E. J. Newman, Solution for the properties of a clustered network. *Phys. Rev. E* **72** (2005), 026136.

[18] A. Pippard, *Elements of classical thermodynamics.* Cambridge University Press, Cambridge, 1979, 122.

[19] C. Radin, Phases in large combinatorial systems. To appear in *Ann. Inst. Henri Poincaré Comb. Phys. Interact.* Preprint 2016. arXiv:1601.04787v2 [math.CO]

[20] C. Radin, K. Ren and L. Sadun, The asymptotics of large constrained graphs. *J. Phys. A* **47** (2014), no. 17, 175001, 20 pp. MR 3197554 Zbl 1290.05139

[21] C. Radin and L. Sadun, Phase transitions in a complex network. *J. Phys. A* **46** (2013), no. 30, 305002, 12 pp. MR 3083277 Zbl 1314.82011

[22] C. Radin and L. Sadun, Singularities in the entropy of asymptotically large simple graphs. *J. Stat. Phys.* **158** (2015), no. 4, 853–865. MR 3311483 Zbl 1320.05064

[23] C. Radin and M. Yin, Phase transitions in exponential random graphs. *Ann. Appl. Probab.* **23** (2013), no. 6, 2458–2471. MR 3127941 Zbl 1278.05225

[24] A. Razborov, On the minimal density of triangles in graphs. *Combin. Probab. Comput.* **17** (2008), no. 4, 603–618. MR 2433944 Zbl 1170.05036

[25] C. Reiher and S. Wagner, Maximum star densities. Preprint 2017. arXiv:1708.01822v2 [math.CO]

[26] D. Ruelle, Correlation functionals. *J. Mathematical Phys.* **6** (1965), 201–220. MR 0201174 Zbl 0127.21501

[27] D. Ruelle, *Statistical mechanics*: *rigorous results.* W. A. Benjamin, New York and Amsterdam, 1969. MR 0289084 Zbl 0177.57301

[28] M. Yin, A. Rinaldo, and S. Fadnavis, Asymptotic quantization of exponential random graphs. *Ann. Appl. Probab.* **26** (2016), no. 6, 3251–3285. MR 3582803 Zbl 1356.05138

Communicated by Christian Borgs

Received May 12, 2016; revised November 9, 2017; accepted February 6, 2018

Charles Radin, Department of Mathematics, The University of Texas at Austin, 2515 Speedway, RLM 8.100, Austin, TX 78712

e-mail: radin@math.utexas.edu

Kui Ren, Department of Mathematics and ICES, The University of Texas at Austin, 2515 Speedway, RLM 8.100, Austin, TX 78712

e-mail: ren@math.utexas.edu

Lorenzo Sadun, Department of Mathematics, The University of Texas at Austin, 2515 Speedway, RLM 8.100, Austin, TX 78712

e-mail: sadun@math.utexas.edu