

Approximating the cumulant generating function of triangles in the Erdős-Rényi random graph

C.Giardinà^(a), C. Giberti^(b), E. Magnanini^(a,c)

^(a) University of Modena and Reggio Emilia, via G.Campi 213/b, 41125 Modena, Italy

^(b) University of Modena and Reggio Emilia, via G.Amendola 2, 42122 Reggio Emilia, Italy

^(c) University of Padova, via Trieste 63, 35121 Padova, Italy

July 28, 2020

Abstract

We study the pressure of the ‘edge-triangle model’, which is equivalent to the cumulant generating function of triangles in the Erdős-Rényi random graph. By analyzing finite graphs of increasing volume, as well as the graphon variational problem in the infinite volume limit, we locate a curve in the parameter space where a one-step replica symmetry breaking transition occurs. Sampling a large graph in the broken symmetry phase is well described by a graphon with a structure very close to the one of an equi-bipartite graph.

1 Introduction

Sampling random graphs with prescribed macroscopic properties (such as a given density of certain subgraphs) received considerable attention in recent years. From a statistical physics perspective, one can think of two procedures:

- the *micro-canonical ensemble*, where the sampling is performed with a uniform distribution over the set of all graphs that satisfy the macroscopic constraint exactly;
- the *canonical ensemble*, where the sampling is done with respect to a larger set of graphs that satisfies the macroscopic constraint only on average.

We shall discuss here the simplest non trivial case, i.e. the constraint is on the number of edges and the number of triangles in the graph. In the micro-canonical ensemble these numbers are prescribed exactly. The canonical ensemble is instead provided by the so-called *edge-triangle model*, which is defined by the Boltzmann-Gibbs distribution in which one tunes the average density of edges and the average density of triangles by varying the corresponding conjugate parameters. The edge-triangle model is in turn the simplest example of the more general *exponential random graph* class of models, in which one introduces several parameters to control the density of an arbitrary set of subgraphs.

Whereas the equivalence in the thermodynamic limit of micro-canonical and canonical ensemble is true for several physical systems of interest (often the system is then studied in the canonical ensemble, that is usually more analytically tractable than the micro-canonical one), for random graphs it has been shown that such equivalence can not be taken for granted. In particular [18] identified a region of values for the densities of edges and triangles where there is *ensemble inequivalence*, as measured by a positive relative entropy between the micro-canonical and canonical measures. In this paper we address the following problem:

How does a large graph look like when it is sampled from the edge-triangle model (i.e. imposing some given *average values* for the densities of edges and triangle)? Does the sampling from the edge-triangle model give the same result of the sampling with respect to the microcanonical ensemble (i.e. imposing some given *exact values* of edge and triangle densities)?

1.1 The edge-triangle model and the Erdős-Rényi random graph

To define the setting, let us consider a *graph* with n vertices, that we identify with the elements of the set $[n] = \{1, 2, 3, \dots, n\}$. We shall describe the graph using its adjacency matrix $x = (x_{i,j})_{i,j \in [n]}$, defined as follows: the entry $x_{i,j} = 1$ if the edge connecting vertex i with vertex j is present, and $x_{i,j} = 0$ otherwise. Since the graphs considered in this paper are undirected and without loops, the adjacency matrices will be always symmetric, with 0 or 1 entries and zeros on the diagonal. We will denote by \mathcal{X}_n the set of adjacency matrices of size n , and we also define for later use $\mathcal{X} = \cup_{n \geq 2} \mathcal{X}_n$ the set of adjacency matrices of all sizes. The number of edges and triangles in a graph represented by a matrix $x \in \mathcal{X}_n$ is given by

$$E_n(x) = \sum_{1 \leq i < j \leq n} x_{i,j} \quad T_n(x) = \sum_{1 \leq i < j < k \leq n} x_{i,j} x_{j,k} x_{k,i}. \quad (1)$$

The previous quantities result in random variables if the graph is a *random graph*, i.e. if it is sampled from the set of $2^{\binom{n}{2}}$ undirected simple graphs with n vertices according to some probability distribution. The simplest possible distribution is the so-called Erdős-Rényi model, where each pair of vertices is connected with probability $p > 0$, independently of the other pairs. Thus, in the Erdős-Rényi case the entries of the adjacency matrix, $x = (x_{i,j})_{i,j \in [n]}$, form a set of independent identically distributed (i.i.d.) Bernoulli variables with $\mathbb{P}(x_{i,j} = 1) = p$. In spite of the simple probabilistic set-up the large deviation principle for the Erdős-Rényi graph is far from simple and has been developed only in recent years [14, 16, 13, 25, 12, 32, 17]. In particular, it has been found that the large deviation function may be non-convex.

The exponential random graph model is devised to enhance or decrease the probability of specific geometric structures in the random graph. Here we define the edge-triangle model that involves only triangles and edges [26]. Let $\beta_1, \beta_2 \in \mathbb{R}$, then the probability of a graph with adjacency matrix $x \in \mathcal{X}_n$ in the edge-triangle model is given by:

$$\nu_n(x) = \frac{\exp\left(\frac{6\beta_2}{n} T_n(x) + 2\beta_1 E_n(x)\right)}{Z_n(\beta_1, \beta_2)}, \quad (2)$$

where $Z_n(\beta_1, \beta_2)$ is the *partition function*, i.e. the normalizing factor

$$Z_n(\beta_1, \beta_2) = \sum_{x \in \mathcal{X}_n} \exp\left(\frac{6\beta_2}{n} T_n(x) + 2\beta_1 E_n(x)\right). \quad (3)$$

The factors 6 and 2 are conventional in the definition and account for the permutations of 3 vertices of a triangle and the 2 vertices of an edge. The Erdős-Rényi model with parameter $0 < p < 1$ is embedded in the edge-triangle model, since its distribution:

$$\nu_n^{ER}(x) = \prod_{1 \leq i < j \leq n} p^{x_{i,j}} (1-p)^{1-x_{i,j}} = (1-p)^{\binom{n}{2}} e^{h_p E_n(x)}, \quad \text{where } h_p := \log \frac{p}{1-p}, \quad (4)$$

can be obtained from (2) by setting $\beta_1 = h_p/2$ and $\beta_2 = 0$. If $\beta_2 > 0$ the probability of finding triangles is enhanced with respect to the Erdős-Rényi case, while it is decreased if $\beta_2 < 0$. In the limiting case $\beta_2 \rightarrow -\infty$ the edge-triangle model (2) gives zero probability to graphs containing triangles.

A key quantity in the study of the thermodynamic properties is the *pressure*, that at finite volume is defined as

$$\psi_n(\beta_1, \beta_2) = \frac{1}{n^2} \ln Z_n(\beta_1, \beta_2). \quad (5)$$

By taking derivatives with respect to the model parameters one computes the averages of the edge density $e = 2E_n/n^2$ and of the triangle density $t = 6T_n/n^3$:

$$\langle e \rangle_n = \frac{\partial \psi_n}{\partial \beta_1}, \quad \langle t \rangle_n = \frac{\partial \psi_n}{\partial \beta_2}, \quad (6)$$

where $\langle \cdot \rangle_n$ denotes expectation w.r.t. the measure ν_n defined in (2). We shall be interested in the behavior of very large graphs which mathematically is described by the (thermodynamic) limit $n \rightarrow \infty$. General convexity arguments imply that the thermodynamic limit is well defined, so that the infinite volume pressure exists

$$\psi(\beta_1, \beta_2) := \lim_{n \rightarrow \infty} \frac{1}{n^2} \ln Z_n(\beta_1, \beta_2), \quad (7)$$

and by Lebesgue dominated convergence limits and derivatives can be interchanged so that the relation (6) gives in the thermodynamic limit

$$\langle e \rangle = \frac{\partial \psi}{\partial \beta_1} \quad \langle t \rangle = \frac{\partial \psi}{\partial \beta_2}, \quad (8)$$

where $\langle e \rangle = \lim_{n \rightarrow \infty} \langle e \rangle_n$. We shall work with the parametrization $(\beta_1, \beta_2) = (h_p/2, \alpha/6)$ where we recall $h_p = \ln \frac{p}{1-p}$ and $0 < p < 1$. In this way the pressure of the edge-triangle model can be read as the cumulant generating function of the number of triangles in the Erdős-Rényi random graph. In other words, defining

$$\mu_{n,p}(\alpha) := \frac{1}{\binom{n}{2}} \log \left\langle \exp \left(\frac{\alpha}{n} T_n(X) \right) \right\rangle_n^{ER}, \quad \alpha \in \mathbb{R}, \quad (9)$$

where $\langle \cdot \rangle_n^{ER}$ denotes the expectation w.r.t. the measure ν_n^{ER} , by a simple computation one can show that

$$\mu_{n,p}(\alpha) = \ln(1-p) + 2\psi_n \left(\frac{h_p}{2}, \frac{\alpha}{6} \right). \quad (10)$$

Thus, studying the cumulant generating function of the number of triangles in the Erdős-Rényi random graph is equivalent to studying the pressure of the edge-triangle model.

1.2 Ensemble inequivalence

As the pressure $\psi_n(\beta_1, \beta_2)$ is the crucial quantity in the canonical ensemble, the entropy $s_n(e, t)$ defined as

$$s_n(e, t) = \frac{1}{n^2} \ln \sum_{x \in \mathcal{X}_n} \delta \left(E_n(x) - \frac{n^2 e}{2} \right) \delta \left(T_n(x) - \frac{n^3 t}{6} \right) \quad (11)$$

is the important quantity in the microcanonical ensemble. An heuristic application of the Laplace method would imply

$$\begin{aligned} \psi(\beta_1, \beta_2) &= \lim_{n \rightarrow \infty} \frac{1}{n^2} \ln \int e^{n^2(\beta_1 e + \beta_2 t - s_n(e, t))} de dt \\ &= \sup_{e, t} \left(\beta_1 e + \beta_2 t - \lim_{n \rightarrow \infty} s_n(e, t) \right) \\ &= \sup_{e, t} \left(\beta_1 e + \beta_2 t - s(e, t) \right), \end{aligned} \quad (12)$$

i.e. in the thermodynamic limit the canonical pressure can be obtained from the microcanonical entropy by a Legendre transform. One then says that the two ensembles are *thermodynamic equivalent* if such correspondence also holds in the reversed direction. This is the same as requiring that the microcanonical entropy is strictly convex, i.e. the involution property of the Legendre transform for strictly convex functions.

The problem of *ensemble inequivalence* is, in fact, more general than just thermodynamic inequivalence (see [31] for a recent account). When the correspondence via Legendre transform between pressure and entropy does not hold, the difference between canonical and microcanonical ensemble can then be probed in several ways. In this paper we shall focus on *macrostate inequivalence*. Namely, we ask if sampling a very large graph uniformly at random from the set of all graphs with given densities of edges and triangles (e^*, t^*) is statistically equivalent to sampling a very large graph from the edge-triangle model

with parameter values $\beta_1(e^*, t^*), \beta_2(e^*, t^*)$ that are obtained by inverting the relations (8) with $\langle e \rangle = e^*$ and $\langle t \rangle = t^*$. In the thermodynamic limit graphs are described by the notion of graphon, which is discussed below. Thus the equivalence that we consider amounts to ask if the two sampling procedures produce the same graphon.

The sampling from the micro-canonical ensemble has been investigated for instance in [28, 29, 30, 22, 3]. It has been found that the structure of graphs drawn from the microcanonical ensemble is very rich and may vary a lot as a function of the number of prescribed edges and triangles. For instance, for a choice (e, t) such that $e = \frac{1}{2} + \epsilon$ with $\epsilon \in (\frac{l-2}{2l}, \frac{l-1}{2l+2})$ with $l \in \mathbb{N} \setminus \{1\}$ and t on the scallopy curve, the vertex set of a graph drawn from the microcanonical ensemble can be partitioned into ℓ subsets ($\ell - 1$ of them of the same size and the last of different size). The graph has the form of a complete ℓ -partite graph on these pieces, plus some additional edges in the last piece that create no additional triangles [18].

To the best of our knowledge, sampling from the canonical ensemble has been investigated only in a limited region of the parameters (β_1, β_2) , see the review of know results in Section 2.1. It is the aim of this paper to conduct a systematic exploration of the full parameter space by means of numerical simulations.

1.3 Main results and paper organization

We investigate the sampling from the canonical ensemble by studying the pressure of the edge-triangle model, or equivalently the cumulant generating function $\mu_p(\alpha)$ of triangles in the Erdős-Rényi model with parameter p . We perform numerical simulations for finite graphs and compare them to the variational formulation describing the infinite volume. We shall collect multiple evidences that the structure of graphs in the canonical ensemble has only two possibilities: it is either the constant graphon describing the Erdős-Rényi graph (i.e. independent edges, yet with a modified parameter for the probability of edges accounting for the imposed number of triangles) or it is the graphon describing the 1-step replica symmetric breaking solution (generalizing the bipartite random graphs that is know to be the exact solution for $\alpha \rightarrow -\infty$). By means of different numerical analysis (“cloning” method for a direct measurement and “gradient” method for the solution of the pressure variational problem) we shall identify a curve $\alpha_c(p)$ in the plane (p, α) separating these two regimes called, respectively, the replica symmetric phase and the replica symmetry broken phase. We do not have a proof that replica symmetry broken phase $\alpha < \alpha_c(p)$ is entirely described by the 1-step replica symmetric breaking solution. However, in contrast to the microcanonical sampling, our numerical analysis suggests that in the description of the canonical sampling no higher level of replica symmetry breaking is required. As a consequence of the numerical analysis the value $\alpha_c(p)$ may be identified as a bona-fide critical value for a 1-step replica symmetry breaking transition.

The paper is structured as follows:

- in section 2 we review the variational formulation of the pressure. We recall the results that are known in the literature for the solution of the variational problem and describe the 1-step replica symmetry breaking solution.
- In section 3 we present the numerical analysis of the finite volume pressure based on the ‘cloning method’, which is a population dynamics algorithm.
- In section 4 we solve a discretized version of the variational problem in the infinite volume by a gradient projection method. Here we do not fix a-priori a specific structure for the optimal graphon.
- In section 5 we solve the variational problem restricted to a specific class of graphons, those corresponding to the 1-step replica symmetry breaking solution.

We will argue that the results of sections 4 and 5 coincide (within numeral accuracy) and they are well approximated by the finite volume direct measurements of section 3.

2 Variational formulation

2.1 Review of known results

The theory of graph limits [24, 23, 8, 9, 10] relies on the notion of *graphon*, which describes a random graph in the limit $n \rightarrow \infty$. A graphon is defined as a bounded Borel measurable function $f : [0, 1]^2 \mapsto [0, 1]$ that satisfies $f(x, y) = f(y, x)$ for all $x, y \in [0, 1]$. The idea behind this definition is a mapping of a graph to the unitary square: intuitively the interval $[0, 1]$ represents a continuum of vertices and $f(x, y)$ is associated to the probability of connecting with an edge two vertices x and y . For example, the graphon describing the Erdős-Rényi random graph with parameter p is the constant function f identically equal to p . The set of graphons is denoted by \mathcal{W} . On this set an equivalence relation is introduced according to which $f, g \in \mathcal{W}$ are equivalent if there exists a bijection $\sigma : [0, 1] \rightarrow [0, 1]$ such that σ and σ^{-1} are Borel measurable and preserve the Lebesgue measure, such that $g(x, y) = f(\sigma(x), \sigma(y))$. The set of equivalence classes is denoted by $\widetilde{\mathcal{W}}$, while \tilde{f} is the equivalence class containing $f \in \mathcal{W}$.

In reviewing the known results in the thermodynamic limit, we follow here [15] (for a more general overview of large deviations for random graph see [12]). The thermodynamic limit of the pressure of the edge-triangle model is given by [15, Theorem 3.1]

$$\psi\left(\frac{h_p}{2}, \frac{\alpha}{6}\right) = \frac{1}{2} \sup_{\tilde{f} \in \widetilde{\mathcal{W}}} \left[\alpha \frac{t(\tilde{f})}{3} - I_p(\tilde{f}) \right] - \frac{\ln(1-p)}{2}, \quad (13)$$

where $t(\tilde{f})$, the *density of triangles* in \tilde{f} , is

$$t(\tilde{f}) = \int_0^1 \int_0^1 \int_0^1 f(x, y) f(y, z) f(z, x) dx dy dz, \quad (14)$$

and the entropic term is

$$I_p(\tilde{f}) = \int_0^1 \int_0^1 I_p(f(x, y)) dx dy. \quad (15)$$

Here f is a representative element of the equivalence class \tilde{f} and $I_p(u)$ for $u \in [0, 1]$ denotes the *Bernoulli relative entropy*

$$I_p(u) := u \ln \frac{u}{p} + (1-u) \ln \frac{1-u}{1-p}. \quad (16)$$

Then, from (10) and (13) we get

$$\mu_p(\alpha) := \lim_{n \rightarrow \infty} \mu_{n,p}(\alpha) = \sup_{\tilde{f} \in \widetilde{\mathcal{W}}} H(\tilde{f}), \quad (17)$$

with

$$H(\tilde{f}) = \alpha \frac{t(\tilde{f})}{3} - I_p(\tilde{f}). \quad (18)$$

When the set of maximizers only consists of constant functions we say that we are in the *replica symmetric* phase. Conversely, if the elements of the maximizing set are non-constant functions, then we say that we are in the *replica symmetry breaking* phase.

The infinite-dimensional variational problem (17), involving the non-linear functional H , has been analytically solved only in a region of the parameter values. In particular [15, Theorem 6.2] proves that for all $0 < p < 1$ the system is in the replica symmetric phase for $\alpha > -2$. The variational problem is then reduced to a scalar one, see [15, Theorem 4.1], i.e. for $\alpha > -2$ we have

$$\sup_{0 \leq u \leq 1} \left[\alpha \frac{u^3}{3} - I_p(u) \right] = \alpha \frac{(u^*(\alpha))^3}{3} - I_p(u^*(\alpha)) =: \mu_p^{RS}(\alpha), \quad (19)$$

where $u^*(\alpha)$ is the optimizer that solves the fixed-point equation:

$$\frac{e^{\alpha u^2 + h_p}}{e^{\alpha u^2 + h_p} + 1} = u, \quad u \in [0, 1]. \quad (20)$$

From this, one infers that for $\alpha > -2$ a graph sampled from the edge-triangle model in the limit $n \rightarrow \infty$ will look like an Erdős-Rényi random graph with parameter $u^*(\alpha)$, i.e. edges are independent from each others and present with a probability $u^*(\alpha)$ (we refer to [5, 15] for the precise statement).

As for the region $-\infty < \alpha \leq -2$ the solution of the variational problem (18) is unknown. The Euler-Lagrange equation giving the stationarity condition are given by the following equation, which is the generalization of (20) to the case of non constant functions [15, Theorem 6.1]:

$$f(x, y) = \frac{\exp(\alpha \int_0^1 f(x, z) f(z, y) dz + h_p)}{\exp(\alpha \int_0^1 f(x, z) f(z, y) dz + h_p) + 1}. \quad (21)$$

It has been proved that for α small enough, a graph sampled from the edge-triangle model no longer resembles an Erdős-Rényi graph. In particular [15, Theorem 6.3] show that for α small enough (and for any value of p) the functional $H(f)$ is not maximized at any constant function. This result is based on the fact [15, Theorem 7.1] that one actually proves that in the limit $\alpha \rightarrow -\infty$, the solution of the variational problem (17) is provided by the so-called *equi-bipartite graphon* defined as

$$g(x, y) := \begin{cases} 0 & \text{if } (x, y) \in [0, \frac{1}{2}]^2 \cup [\frac{1}{2}, 1]^2 \\ p & \text{if } (x, y) \in [0, \frac{1}{2}] \times [\frac{1}{2}, 1] \cup [\frac{1}{2}, 1] \times [0, \frac{1}{2}]. \end{cases} \quad (22)$$

and in this limit one has

$$\lim_{\alpha \rightarrow -\infty} \mu_p(\alpha) = H(g) = \frac{1}{2} \ln(1 - p) =: \hat{\mu}_p. \quad (23)$$

For later use, we remark that using the constant graphon f_α defined by $f_\alpha(x, y) = u^*(\alpha)$ for all $x, y \in [0, 1]^2$, one obtains

$$\lim_{\alpha \rightarrow -\infty} \mu_p^{RS}(\alpha) = \lim_{\alpha \rightarrow -\infty} H(f_\alpha) = \ln(1 - p) =: \hat{\mu}_p^{RS}. \quad (24)$$

Thus, in the limit $\alpha \rightarrow -\infty$ the replica symmetric solution is wrong by a factor 2!

2.2 A conjecture

The graphon (22) is the infinite volume correspondent of the *equi-bipartite graph*. In the latter the n vertices are partitioned into two disjoint sets of equal size with no edges connecting two vertices belonging to the same set. Clearly, in such a graph triangles do not exist (more generally bipartite graphs do not contain odd cycles [7, Theorem 4]). Analogously, the density of triangles in the equi-bipartite graphon (22) vanishes since $t(g) = 0$. Thus, as expected, the edge-triangle model in the limit $\alpha \rightarrow -\infty$ is free of triangles.

Guided by the results of our numerical analysis that we illustrate below, we conjecture that for any fixed $0 < p < 1$, there exists a unique finite and negative value $\alpha_c(p)$ such that, crossing $\alpha_c(p)$ from above, the optimizer of H defined in (18) switches from the constant function f_α , identically equal to the solution of the fixed-point equation (20), to the graphon:

$$g_\alpha(x, y) := \begin{cases} p_1(\alpha) & \text{if } (x, y) \in [0, \frac{1}{2}]^2 \cup [\frac{1}{2}, 1]^2 \\ p_2(\alpha) & \text{if } (x, y) \in [0, \frac{1}{2}] \times [\frac{1}{2}, 1] \cup [\frac{1}{2}, 1] \times [0, \frac{1}{2}], \end{cases} \quad (25)$$

where $p_1(\alpha)$ and $p_2(\alpha)$ are functions taking value in $(0, 1)$ that satisfy the following conditions:

$$\lim_{\alpha \rightarrow -\infty} p_1(\alpha) = 0, \quad (26)$$

$$\lim_{\alpha \rightarrow -\infty} p_2(\alpha) = p. \quad (27)$$

We observe that $\lim_{\alpha \rightarrow -\infty} g_\alpha = g$, with g the graphon describing the equi-bipartite graph defined in (22). The rationale behind our conjecture is that the structure of (25) represents the simplest geometry that may emerge from the breaking of the homogeneous graphon. For the resemblance of the structure of the overlap matrix in the 1RSB solution of spin-glasses, we call this graphon the *1-step replica symmetry breaking solution*.

3 Direct measurements for finite graphs

In this section we compute numerically the cumulant generating function of the number of triangles of a finite graph of size n . This expectation is substantially affected by events that, although rare, give a large contribution to the average defining the generating function. A standard tool for estimating the probability of rare events in the Erdős-Rényi random graph is the importance sampling technique, see for instance [6]. Here we follow an approach based on population dynamics, called ‘cloning’ [20, 19, 21, 27, 11, 1, 2]. In this section we adapt the method to a purely geometric problem, by introducing a dynamics for the graph construction.

3.1 Implementing cloning

The cloning algorithm is obtained by tilting a Monte Carlo dynamics that samples from a target distribution. In our case we would like to compute expectations w.r.t. the law of the Erdős-Rényi graph (4). We first describe the dynamical process generating the Erdős-Rényi graph and then we recall how the tilted dynamic arises in the cloning algorithm.

In the Erdős-Rényi graph each edge is present, independently, with probability $p \in (0, 1)$. To generate a graph of size n , we consider the Markov chain $\{X_t, t \in \mathbb{N}\}$, taking values on the set \mathcal{X} of adjacency matrices, defined as follows. We label the n vertices in an arbitrary order. We start by selecting the first two vertices and we connect them with probability p , thereby obtaining a graph of size two. Then we select the third vertex and try to connect it to the first two vertices independently with probability p , thus obtaining a graph of size three. This procedure is repeated until the graph of size n is formed: each time a new vertex is selected, it is connected independently with probability p to each of those already visited. We stipulate that the discrete time step of this process corresponds to the attempt of adding a *single* new edge. Since the evolution from the graph of size i to the graph of size $i + 1$ requires i attempts, then the evolution starting from size two and leading to a graph of size n will require $\mathcal{N}_n = \sum_{i=2}^n i = \binom{n}{2} - 1$ steps.

We now introduce the tilted dynamics. Denoting by P the transition matrix of the Markov chain described above, i.e. $P(x, y) = \mathbb{P}(X_{t+1} = y | X_t = x)$ for $x, y \in \mathcal{X}$, the cumulant generating function of triangles reads

$$\mu_{n,p}(\alpha) = \frac{1}{\mathcal{N}_n + 1} \ln \left(\sum_{x_0, \dots, x_{\mathcal{N}_n} \in \mathcal{X}_n} \nu_0(x_0) P(x_0, x_1) \dots P(x_{\mathcal{N}_n-1}, x_{\mathcal{N}_n}) e^{\frac{\alpha}{n} T(x_{\mathcal{N}_n})} \right), \quad (28)$$

where ν_0 is the initial distribution (i.e. Bern(p)). We rewrite the number of triangles for the graph of size n as

$$T(x_{\mathcal{N}_n}) = \sum_{t=0}^{\mathcal{N}_n} \Delta T(x_t, x_{t+1}), \quad (29)$$

where $\Delta T(x_t, x_{t+1}) = T(x_{t+1}) - T(x_t)$ is the increment of triangles between two consecutive steps. In this way we obtain:

$$\mu_{n,p}(\alpha) = \frac{1}{\mathcal{N}_n + 1} \ln \left(\sum_{x_0, \dots, x_{\mathcal{N}_n} \in \mathcal{X}_n} \nu_0(x_0) P(x_0, x_1) \dots P(x_{\mathcal{N}_n-1}, x_{\mathcal{N}_n}) e^{\frac{\alpha}{n} \Delta T(x_0, x_1)} \dots e^{\frac{\alpha}{n} \Delta T(x_{\mathcal{N}_n-1}, x_{\mathcal{N}_n})} \right). \quad (30)$$

The average in the previous equation can be also computed according to a different dynamics. Indeed, introducing the quantity

$$k_\alpha(x) = \sum_{y \in \mathcal{X}_n} P(x, y) e^{\frac{\alpha}{n} \Delta T(x, y)}, \quad x \in \mathcal{X}_n, \quad (31)$$

and the stochastic matrix p_α with elements

$$P_\alpha(x, y) := P(x, y) e^{\frac{\alpha}{n} \Delta T(x, y)} \frac{1}{k_\alpha(x)}, \quad (32)$$

we can write:

$$\mu_{n,p}(\alpha) = \frac{1}{\mathcal{N}_n + 1} \ln \left(\sum_{x_0, \dots, x_{\mathcal{N}_n} \in \mathcal{X}_n} \nu(x_0) P_\alpha(x_0, x_1) \dots P_\alpha(x_{\mathcal{N}_n-1}, x_{\mathcal{N}_n}) k_\alpha(x_0) \dots k_\alpha(x_{\mathcal{N}_n-1}) \right). \quad (33)$$

As observed in [19] this representation of the average suggests a population dynamics scheme that, starting from a bunch of M initial individuals (clones) with distribution $\nu_0(\cdot)$, make them evolve according to the transition kernel $P_\alpha(\cdot, \cdot)$ and reproduce according to the rate $k_\alpha(\cdot)$. Denoting by $M_{\mathcal{N}_n}$ the size of the corresponding population of clones at the final time \mathcal{N}_n , we get:

$$\mu_{n,p}(\alpha) = \frac{1}{\binom{n}{2}} \ln \left[\frac{M_{\mathcal{N}_n}}{M} \right]. \quad (34)$$

From the computational point of view, to avoid explosion or extinction, at any time step t , the family of M_t of clones is brought back to the original size M by picking uniformly at random M clones out of the M_t available. Recording the ratio $R_t = M_t/M_{t-1}$, the left hand side of the previous display can be written in a telescopic form:

$$\mu_{n,p}(\alpha) = \frac{1}{\binom{n}{2}} \ln \left[\prod_{t=3}^{\mathcal{N}_n} R_t \right] = \frac{1}{\binom{n}{2}} \sum_{t=3}^{\mathcal{N}_n} \ln R_t. \quad (35)$$

Summarizing, formula (33) is implemented, starting from a population of M elements randomly chosen according to ν_0 , by iterating the following steps:

- a) Evolve independently each clone (i.e. graph) with the transition probability P_α .
- b) Replicate each clone with reproduction rate k_α thus obtaining a population of size M' .
- c) The ratio $R = M'/M$ is recorded and then the total number of copies is brought back to M , uniformly choosing M clones among the M' .

The procedure is further illustrated with the example of size $n = 3$ in the Appendix.

3.2 Numerical results

We present here the results of the cloning algorithm for the cumulant generating function of the triangles. We fix a population of $M = 7000$ clones and a value of $p = 0.4$, and we vary the variable α and the graph size n . We denote by $\mu_n^{Cl}(\alpha)$ the cumulant generating function returned by the cloning algorithm for the graph of size n with $p = 0.4$. Similarly, we shorthand $\mu^{RS}(\alpha) := \mu_{0.4}^{RS}(\alpha)$ and we denote the asymptotic values $\hat{\mu} := \hat{\mu}_{0.4}$ and $\hat{\mu}^{RS} := \hat{\mu}_{0.4}^{RS}$. The outcomes of our simulations show that:

- for large values of α , the algorithm reproduces the replica symmetric solution (i.e. $\mu_n^{Cl} \rightarrow \mu^{RS}$ as n increases);
- for small values of α , the algorithm provides a cumulant generating function which is independent of α within statistical fluctuations. This “constant” function approximates better and better the value of the exact solution at $\alpha = -\infty$ (i.e. $\mu_n^{Cl} \rightarrow \hat{\mu}$ as n increases);
- for all values of α and n , the curves furnished by the cloning algorithm are always above the curve that is obtained by taking the maximum between the replica symmetric solution and the exact solution at $\alpha = -\infty$ (i.e. $\mu_n^{Cl} \geq \max\{\mu^{RS}, \hat{\mu}\}$). The discrepancy reduces as n increases.

We observe that the intersection between the replica symmetric solution and the exact solution at $\alpha = -\infty$ occurs at a value $\hat{\alpha} \simeq -110$, thus a substantially small value where the rate of reproduction of triangles in the cloning algorithm is very small. Despite this “extreme” situation, we see strong evidence that the

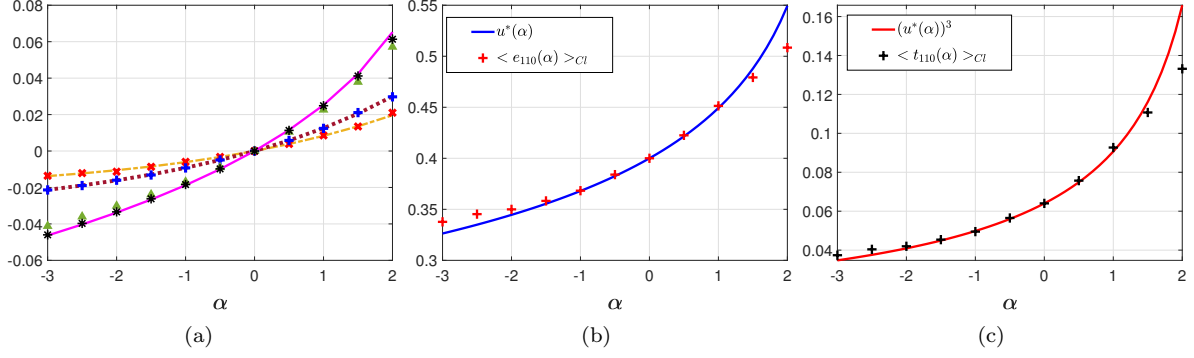


Figure 1: Results of the cloning algorithm for $p = 0.4$ for $\alpha \in [-3, 2]$ (thus in the replica symmetric phase). Left panel: the picture shows the curves $\mu_{n,p}(\alpha)$ for $n = 3$ (dashed-dotted yellow line), $n = 4$ (dotted brown line) and the curve $\mu_p^{RS}(\alpha)$ (magenta continuous line), as well as the output of simulations $\mu_n^{Cl}(\alpha)$ represented with dots: $n = 3$ (x), $n = 4$ (+), $n = 20$ (▲), $n = 110$ (*). Central panel: normalized average number of edges for $n = 110$ (+) in the clones ensemble, together with the solution $u^*(\alpha)$ of the fixed-point equation (20). Right panel: normalized average number of triangles for $n = 110$ (+) with $(u^*(\alpha))^3$ i.e. the expected density of triangles in Erdős-Rényi model with parameter $u^*(\alpha)$.

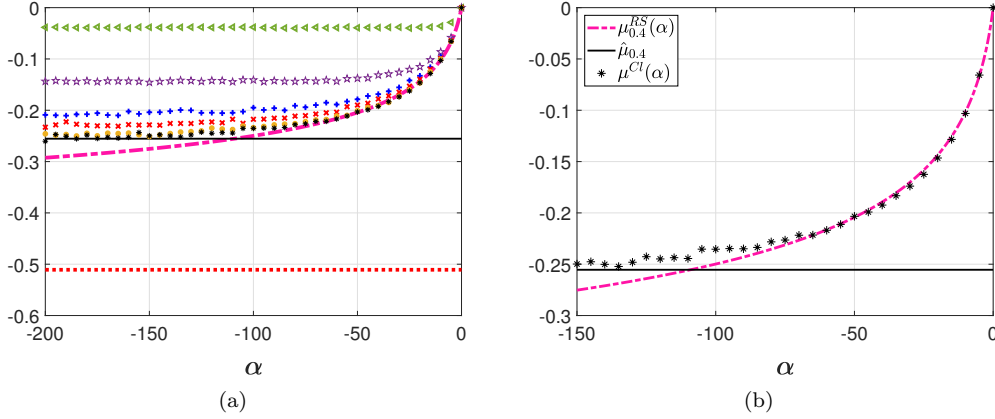


Figure 2: Numerical results of the cloning algorithm for $p = 0.4$ and $\alpha \in [-150, 0]$. Left panel: the picture reports the numerical curves $\mu_n^{Cl}(\alpha)$ with $n = 4$ (◄), $n = 20$ (★), $n = 50$ (+), $n = 70$ (x), $n = 100$ (●), $n = 110$ (*) together with the replica symmetric solution $\mu_p^{RS}(\alpha)$ (dashed-dotted magenta line), its asymptotic value $\hat{\mu}_p^{RS}(\alpha)$ (dotted red horizontal line) and the true asymptotic value $\hat{\mu}_p$ (continuous black horizontal line). Right panel: $\mu_n^{Cl}(\alpha)$ for $n = 110$ (*) together with the replica symmetric solution $\mu_p^{RS}(\alpha)$ (dashed-dotted magenta line) and the asymptotic value $\hat{\mu}_p$ of $\mu_p(\alpha)$.

replica symmetry solution does not hold for small α 's, and for values $\alpha \leq -110$ the solution furnished by the equi-bipartite graphon is a very good approximation for the curve returned by the algorithm.

We discuss our results below. Figure 1 explores the region around $\alpha = 0$. Figure 1(a) shows with dots the results of the cloning algorithm for $n = 3$, $n = 4$, $n = 20$, $n = 110$ and $\alpha \in [-3, 2]$. The picture also reports, with orange dashed-dotted and maroon dotted lines, the exact curves $\mu_{n,0.4}$ for $n = 3$ and $n = 4$, that can be computed explicitly. The curve $\mu_p^{RS}(\alpha)$, obtained by solving numerically the scalar variational problem (19) characterizing the replica symmetric regime for $p = 0.4$, is also displayed (pink continuous line). We observe that for $n = 3$ and $n = 4$ the cloning algorithm perfectly reproduces the

exact result and, as long as n grows, the curves for increasing n 's settle on the curve $\mu^{RS}(\alpha)$.

A further check on the behavior of the algorithm is provided by Fig.1(b). It shows, as a function of α , the average density of edges $\langle e_n(\alpha) \rangle_{Cl}$ in the population of clones, that is $\langle e_n(\alpha) \rangle_{Cl} = \frac{1}{M} \sum_{i=1}^M \frac{E(x^{(i)})}{\binom{n}{2}}$. This quantity is an estimator of the probability p' of finding an edge connecting any two vertices of the graphs produced by the cloning. By effect of the tilting, p' is different from the original value $p = 0.4$ and is a function of α , as expected. Since in replica symmetric regime the edge-triangle model converges (in the limit $n \rightarrow \infty$) to an Erdős-Rényi model with parameter $u^*(\alpha)$, see Section 2, we have reported in Fig.1(b) the optimizer $u^*(\alpha)$. The fair agreement of the values $\langle e_n(\alpha) \rangle_{Cl}$ on the curve of $u^*(\alpha)$, gives a further evidence that the cloning algorithm is well working in this regime of α and reproduces the expected behavior of the edge-triangle model. The same conclusion can be drawn from Fig.1(c) that shows the estimate of the average density of triangles provided by cloning, i.e. $\langle t_n(\alpha) \rangle_{Cl} = \frac{1}{M} \sum_{i=1}^M \frac{T(x^{(i)})}{\binom{n}{3}}$ and the density of triangles in the symmetric replica regime, i.e. $(u^*(\alpha))^3$.

To check if the algorithm is able to reproduce the change occurring around $\alpha \simeq -110$, we have run simulations in a strongly negative region of α . We plot in Fig.2 the results of the cloning method $\mu_n^{Cl}(\alpha)$ (dots) together with the replica symmetric cumulant generating function $\mu^{RS}(\alpha)$. In the same picture we have also reported the asymptotic values $\hat{\mu} = -0.255$ and $\hat{\mu}^{RS} \simeq -0.510$. The left panel of Fig.2 displays the cloning data for several n values, showing the converge towards a limiting profile as n is increased. Figure 2(b) exhibits the results for the largest graph size $n = 110$.

4 Numerical solution of the variational problem

The cloning algorithm implicitly solves the variational problem (17) by producing a population of graphs that approximate the optimizing graphon. Unfortunately, it is hard to scrutinize the structure of the graphon from the adjacency matrix of the graphs, when their size is large. Thus, in order to study the graphon in the replica symmetry broken phase, we solve (17) via a numerical discretization.

4.1 Discretization

The spatial discretization of graphon can be obtained considering the set of $m \times m$ symmetric matrices $\{f_{i,j}\}_{i,j}$ with elements in $[0, 1]$. However, since the graphon that solves (17) does not take the values 0 and 1 [15, Theorem 6.3], we restrict our set of matrices to the following set:

$$\Gamma_{m,\varepsilon} := \{f \in \mathbb{R}^{m \times m} : f_{ij} = f_{ji}, f_{ij} \in [\varepsilon, 1 - \varepsilon], \quad i, j = 1, \dots, m\}, \quad (36)$$

where $\varepsilon > 0$ is a small parameter that bounds the functions away from the singularity of the logarithm in the discretization of $H(\tilde{f})$, that we define as follows:

$$\mathcal{H}_m(f) := \frac{1}{m^2} \left[\frac{\alpha}{3m} \sum_{i,j,k=1}^m f_{ij} f_{jk} f_{ki} - \sum_{i,j=1}^m \left(f_{ij} \ln \frac{f_{ij}}{p} + (1 - f_{ij}) \ln \frac{1 - f_{ij}}{1 - p} \right) \right], \quad f \in \Gamma_{m,\varepsilon}. \quad (37)$$

We remark that, as it happens in the continuous setting, $\mathcal{H}_m(f)$ enjoys a symmetry property. Indeed, given $f, g \in \Gamma_{m,\varepsilon}$ we have $\mathcal{H}_m(g) = \mathcal{H}_m(f)$ if there exists a permutation σ over $\{1, \dots, m\}$ such that $g_{i,j} = f_{\sigma(i), \sigma(j)}$. In this case we call g and f *equivalent*.

We solve numerically the discretization of (17)

$$\max_{f \in \Gamma_{m,\varepsilon}} \mathcal{H}_m(f) \quad (38)$$

by applying the *Gradient Projection* (GP) method with both a constant and variable steplength, the latter chosen according to the Barzilai-Borwein rules [4].

4.2 Numerical results

Being interested in the structure of the optimizer in the replica symmetry breaking region, we have solved (38) for a set of p values ranging from 0.2 to 0.6 and α below -2 , varying it with unitary step. For each value of p and α we have started the iterations of the GP method from a set of 12 initial conditions, using a grid of size $m = 40$ and setting $\varepsilon = 10^{-4}$. Our main findings are:

- a) the presence of only two different geometrical structures of maximizers: the constant ones and chessboard-like one (see Fig.3). The values of the constant maximizers turn out to be equal, within the numerical approximation, to the solution of the fixed-point equation (20);
- b) there exists a critical value $\alpha_c(p)$ such that when $\alpha > \alpha_c(p)$ the optimizer of $\mathcal{H}_m(f)$ is constant whereas when $\alpha < \alpha_c(p)$ it assumes a chessboard-like structure.

We observe that there are different chessboard-like structures that can be reached starting from different initial conditions, as shown in Fig.3. All of them are equivalent to the discretization of the 1-RSB graphon, that by abuse of notation we also denote by g_α :

$$g_\alpha(i, j) = \begin{cases} p_1 & \text{if } (\frac{i}{m}, \frac{j}{m}) \in [0, \frac{1}{2}]^2 \cup [\frac{1}{2}, 1]^2, \\ p_2 & \text{if } (\frac{i}{m}, \frac{j}{m}) \in [0, \frac{1}{2}] \times [\frac{1}{2}, 1] \cup [\frac{1}{2}, 1] \times [0, \frac{1}{2}]. \end{cases} \quad (39)$$

With p fixed, and varying α , we sought the critical value $\alpha_c(p)$ by comparing the values of \mathcal{H}_m computed at the homogeneous solutions f_α and at the chessboard-like ones g_α . We found that the values of f_α coincide, within our approximation and for all α , with $u^*(\alpha)$ (the solution of the replica symmetric fixed-point equation (20)), while the values $p_1(\alpha)$ and $p_2(\alpha)$ occurring in g_α are close to 0 and p , see Tab.1 for large negative α values. This implies that g_α is close to the equi-bipartite graphon g defined in (22). The evidence that the phase transition occurs at the critical value $\alpha_c(p)$ is given by observing that $\mathcal{H}_m(f_\alpha) > \mathcal{H}_m(g_\alpha)$ for $\alpha > \alpha_c(p)$, while $\mathcal{H}_m(f_\alpha) < \mathcal{H}_m(g_\alpha)$ for $\alpha < \alpha_c(p)$. An example is given in Table 1 where, for the case $p = 0.4$, the change of the optimizer shows that the position of the critical value $\alpha_c(p)$ is between -109 and -110 (we recall that α varies with unitary step).

α	$u^*(\alpha)$	f_α	g_α	$\mathcal{H}_{40}(f_\alpha)$	$\mathcal{H}_{40}(g_\alpha)$
-109	0.120486	0.120486	$p_1 \approx 0.000114, p_2 \approx 0.398829$	-0.255336	-0.255356
-110	0.120082	0.120082	$p_1 \approx 0.000100, p_2 \approx 0.399100$	-0.255916	-0.255361

Table 1: Numerical optimizers of (38) for $p = 0.4$ and $\alpha = -109, -110$. We report the solution $u^*(\alpha)$ to equation (20), the constant solution f_α and the two values $p_1(\alpha)$ and $p_2(\alpha)$ taken by the chessboard solution g_α . The last two columns show the values of $\mathcal{H}_{40}(f)$ that give evidence of the phase transition between -109 and -110 .

Denoting by h_α the global optimizer found by the GP method, in Fig.4 we represent $\mathcal{H}_{40}(h_\alpha)$ as a function of α for $p \in \{0.2, 0.4, 0.6\}$. In the replica symmetric region, i.e for $\alpha > \alpha_c(p)$, $\mathcal{H}_{40}(h_\alpha)$ is expected to approximate the solution $\mu^{RS}(\alpha) = \alpha \frac{(u^*(\alpha))^3}{3} - I_p(u^*(\alpha))$. The overlapping of the two curves above $\alpha = -455$, $\alpha = -110$, $\alpha = -69$ and $\alpha = -47$ is shown in Fig.4. Below such thresholds, that we identify as approximations of critical value $\alpha_c(p)$, the optimal solution h_α of \mathcal{H}_{40} switches from the constant one f_α (approximating the fixed-point equation (20)) to a chessboard function g_α . The function $\mathcal{H}_{40}(h_\alpha)$ is nearly flat for $\alpha < \alpha_c(p)$ and very close to the asymptotic value $H(g) = \frac{1}{2} \ln(1-p)$ of $\mu_p(\alpha)$, see (23).

5 Numerical solution of the restricted variational problem: the 1-step RSB optimizer

In this section, in order to study the transition between the replica symmetric region and replica symmetry breaking region, we return to the continuous variational problem (17). In view of the numerical results

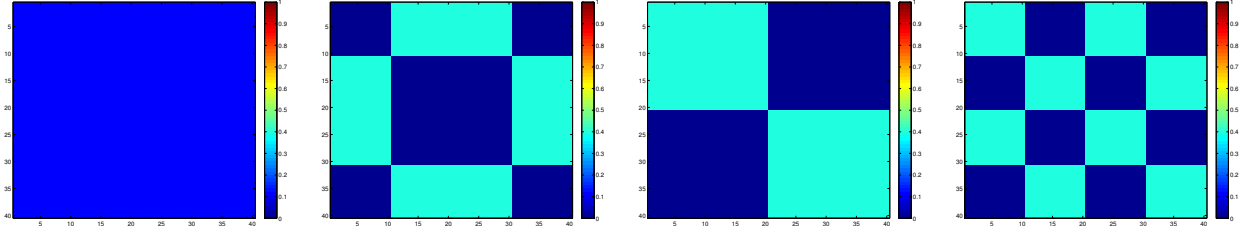


Figure 3: The stationary points of (37) returned by the GP method for $p = 0.4$ and $\alpha = -110$ starting. Different initial conditions, according to the steplength rules used, lead either to the constant function (leftmost panel) and to a chessboard structure (the other panels). The value of the homogeneous solution on the leftmost panel is equal to $u^*(-110)$. The chessboard solutions are equivalent according to our definition.

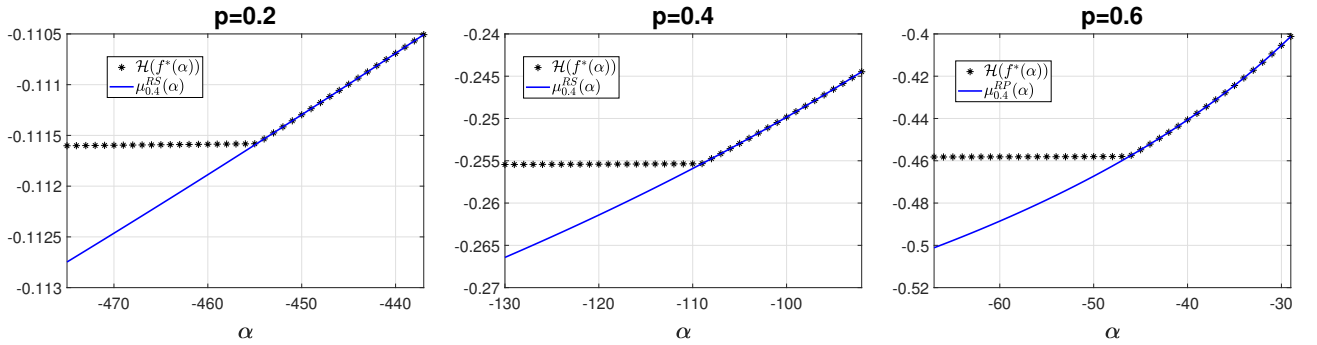


Figure 4: Plot of $\mathcal{H}_{40}(f^*(\alpha))$ as a function of α , where $f^*(\alpha)$ is the numerical solution of (38) (black, dotted line), together with the function $\mu_{0.4}^{RS}(\alpha)$ which solves problem (17) in the replica symmetric regime (continuous, blue line).

of Section 4, here we *assume* that the breaking of the homogeneous graphon gives rise to a solution with the chessboard structure, see Fig.3,

$$g_{p_1, p_2}^{(a)}(x, y) = \begin{cases} p_1 & \text{if } (x, y) \in [0, a]^2 \cup [a, 1]^2 \\ p_2 & \text{if } (x, y) \in [0, a] \times [a, 1] \cup [a, 1] \times [0, a], \end{cases} \quad (40)$$

that we call the *generalized 1-step replica symmetry breaking* solution. We introduce the set $\mathcal{R} = \{g_{p_1, p_2}^{(a)}(x, y) | a, p_1, p_2 \in [0, 1]\} \subset \mathcal{W}$. This set contains the constant graphon (obtained by setting $p_1 = p_2$ for any a or, equivalently, by setting either $a = 0$ or $a = 1$) and, for $a = \frac{1}{2}$, the graphon (25), that we claim to be the solution of (17) for $\alpha < \alpha_c(p)$. Also, the limiting graphon (22) is contained in \mathcal{R} .

We thus restrict the infinite dimensional problem (17) to the finite dimensional one obtained by restricting to the set \mathcal{R} :

$$\sup_{g \in \mathcal{R}} \left[\frac{\alpha t(g)}{3} - I_p(g) \right] \equiv \sup_{p_1, p_2, a \in [0, 1]} \left[\frac{\alpha t(g_{p_1, p_2}^{(a)})}{3} - I_p(g_{p_1, p_2}^{(a)}) \right] =: \mu_p^{1RSB}(\alpha). \quad (41)$$

Being interested in the phase transition, we consider this problem for $\alpha \leq -2$. Using this approach, we aim at locating the critical value $\tilde{\alpha}_c(p)$ denoting the transition that we claim to occur in \mathcal{R} between the homogeneous and the 1-step replica breaking solution. Obviously we can state that $\tilde{\alpha}_c(p) \leq \alpha_c(p)$, but we do not have any argument to assert that the transition in \mathcal{R} is the same occurring in the whole space \mathcal{W} and, as a consequence, that $\tilde{\alpha}_c(p)$ coincides with $\alpha_c(p)$. However, some evidence for the equality of the two critical points will be obtained in this section.

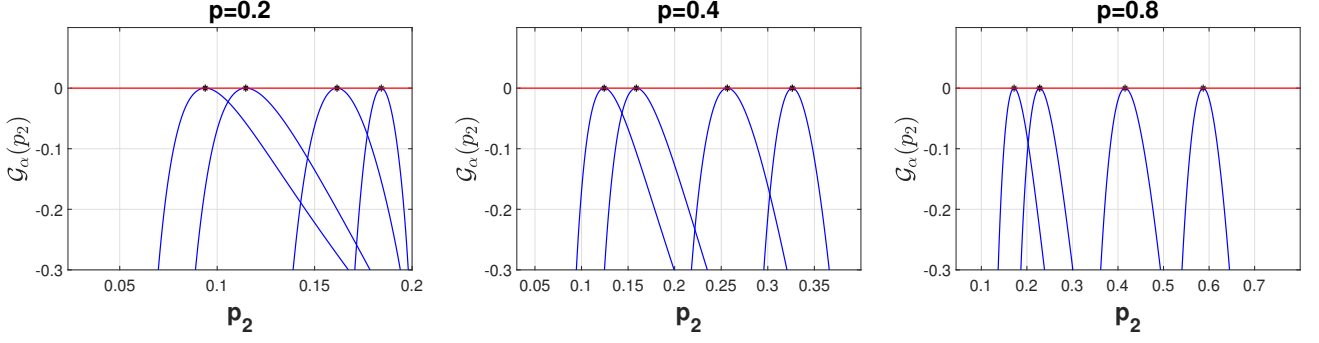


Figure 5: Plot of the function $\mathcal{G}_\alpha(p_2)$ defined in (44) for $p = 0.2, 0.4, 0.8$ (form left to right) for varying α . The black dots on the horizontal line $y = 0$ represent $u^*(\alpha)$. These plots show that $u^*(\alpha)$ is the unique solution to equation (44).

The function to be maximized in (41) can be written as follows:

$$\begin{aligned} \mathcal{F}_\alpha(p_1, p_2, a) &:= \alpha \frac{t(g_{p_1, p_2}^{(a)})}{3} - I_p(g_{p_1, p_2}^{(a)}) \\ &= \frac{\alpha}{3} \{p_1^3[a^3 + (1-a)^3] + 3p_2^2 p_1 a(1-a)\} - 2a(1-a)I_p(p_2) - [a^2 + (1-a)^2]I_p(p_1), \end{aligned}$$

since the density of triangles (14) in the generalized 1-step replica symmetry breaking graphons (40) is

$$t(g_{p_1, p_2}^{(a)}) = p_1^3[a^3 + (1-a)^3] + 3p_2^2 p_1 a(1-a). \quad (42)$$

This function, that satisfies the symmetry $\mathcal{F}_\alpha(p_1, p_2, a) = \mathcal{F}_\alpha(p_1, p_2, 1-a)$, can be defined by continuity up to the boundaries of $[0, 1]^3$ by setting $I_p(0) = I_p(1) = 0$. Therefore $\mathcal{F}_\alpha(p_1, p_2, a)$ attains its maximum on $[0, 1]^3$. Due to the fact that the optimizer of (17) is bounded away from 0 and 1 [15], we assume that the coordinates (p_1, p_2) of the maximum point lie in the interior of $[0, 1]^2$ and thus satisfies stationarity condition $\nabla \mathcal{F}_\alpha(p_1, p_2, a) = 0$, given by the following system of equations:

$$\begin{cases} \alpha [p_1^2(3a(a-1) + 1) + ap_2^2(1-a)] - (a^2 + (1-a)^2) \ln \left(\frac{p_1(1-p)}{p(1-p_1)} \right) = 0, & (a) \\ 2a(1-a) \left[p_1 p_2 \alpha - \ln \left(\frac{p_2(1-p)}{p(1-p_2)} \right) \right] = 0, & (b) \\ (2a-1) [\alpha(p_1^3 - p_2^2 p_1) + 2(I_p(p_2) - I_p(p_1))] = 0, & (c) \end{cases} \quad (43)$$

with $p_1, p_2, a \in [0, 1]$ and $\alpha \leq -2$.

It is simple to check that the point $(p_1^*(\alpha), p_2^*(\alpha), a^*)$ with $p_1^*(\alpha) = p_2^*(\alpha) = u^*(\alpha)$ and any $a^* \in [0, 1]$ is a solution to equation (43) for any α . Indeed, when $p_1 = p_2$ equations (a) and (b) are equal and coincide with the fixed-point equation (20). From the numerical solution of system (43) as α is varied, we got evidence of the existence of a threshold $\tilde{\alpha}_c(p)$ above which $(u^*(\alpha), u^*(\alpha), a^*)$ is the maximum (this holds for any a' , being such parameter meaningless in the homogeneous case). Crossing $\tilde{\alpha}_c(p)$ from above, a non-homogeneous solution with $a^* = 1/2$ and $p_1^*(\alpha) \approx 0$, $p_2^*(\alpha) \approx p$ arises; it turns out to be the maximum of the problem for $\alpha < \tilde{\alpha}_c(p)$.

We give more details on the procedure we followed in order to analyze the solutions of (43). First we observe that there are three special values of a , namely $a = 0, \frac{1}{2}, 1$. As we said above, $a = 0, 1$ correspond to the constant graphon that can be obtained also as a special case (with $p_1 = p_2$) of $a = \frac{1}{2}$. Thus since the cases $a = 0, 1$ can be absorbed in $a = \frac{1}{2}$, we are left with the problem of analyzing (43) for $a = \frac{1}{2}$ and $a \neq \frac{1}{2}$.

- **Case $a \neq \frac{1}{2}$.** From equation (b) in (43) we get $p_1 = \frac{1}{\alpha p_2} \ln \left(\frac{p_2(1-p)}{p(1-p_2)} \right)$ that replaced in (c) gives an

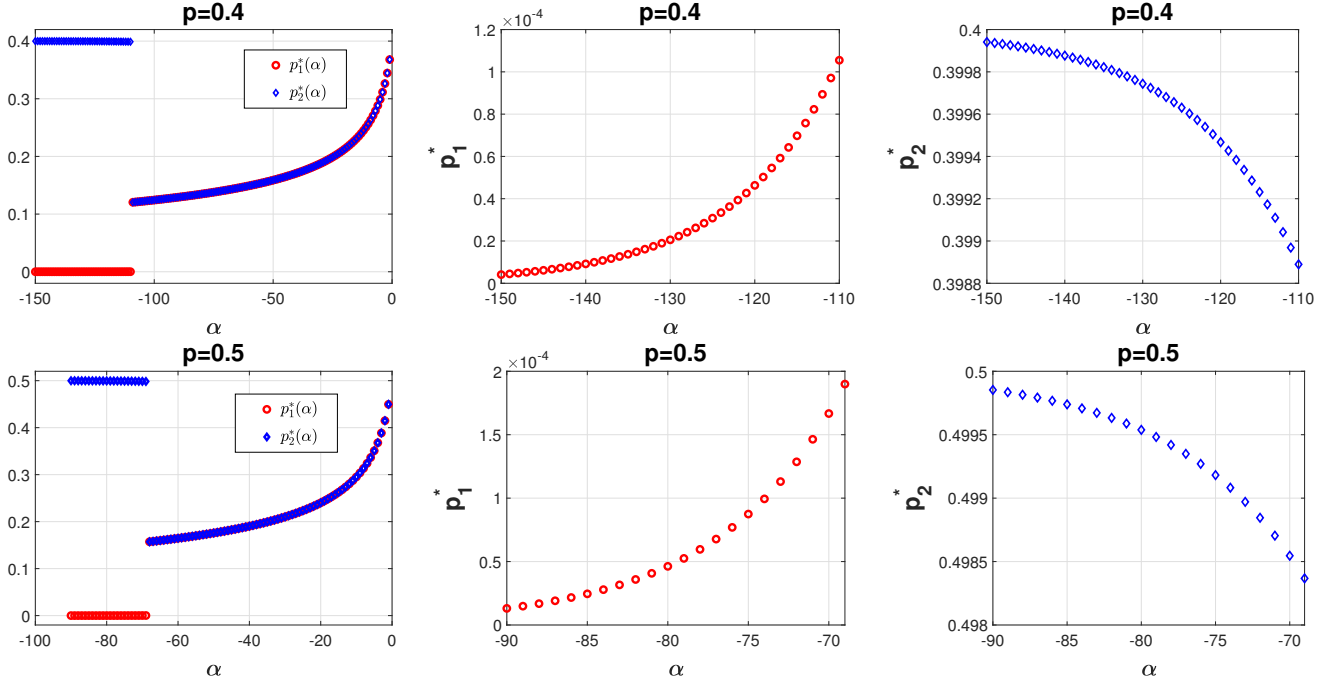


Figure 6: Upper panel: the left picture represents the parameters $p_1^*(\alpha)$ (\circ) and $p_2^*(\alpha)$ (\diamond) of the optimizer of (41), as a function of $\alpha \in [-150, 0]$, where α varies with unitary step, and $p = 0.4$. The pictures in the middle and in the right columns show the behaviour of $p_1^*(\alpha)$ and $p_2^*(\alpha)$ in the zoomed interval $[-150, -110]$. Lower panel: the same as in the upper panel for $p = 0.5$, $\alpha \in [-90, 0]$ and $\alpha \in [-90, -69]$.

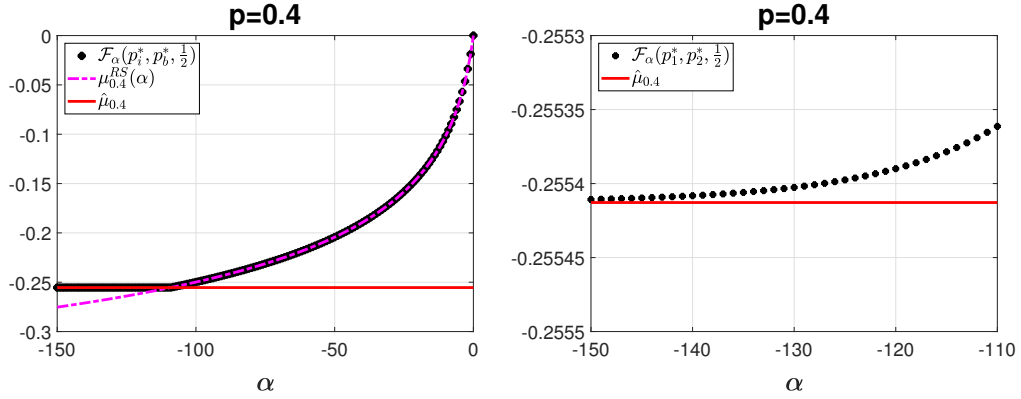


Figure 7: Left column: plot of $\mathcal{F}_\alpha(p_1^*(\alpha), p_2^*(\alpha), \frac{1}{2})$ (black dots) together with the function $\mu_p^{RS}(\alpha)$ (magenta dashed-dotted line) for $p = 0.4$. The horizontal red line represents the asymptotic value $\hat{\mu}_p$, see (23). Right column: zoom on the interval $[-150, -110]$ for the case $p = 0.4$.

equation for p_2 :

$$\mathcal{G}_\alpha(p_2) := \frac{1}{\alpha^2 p_2^3} \left[\ln \left(\frac{p_2(1-p)}{p(1-p_2)} \right) \right]^3 - p_2 \ln \left(\frac{p_2(1-p)}{p(1-p_2)} \right) + 2I_p(p_2) - 2I_p \left(\frac{1}{\alpha p_2} \ln \left(\frac{p_2(1-p)}{p(1-p_2)} \right) \right) = 0. \quad (44)$$

We solved numerically the previous equation for several values of p and α . The function $\mathcal{G}_\alpha(p_2)$ turns out to be a concave function whose maximum, that is also the solution to eq.(44), coincides with the solution

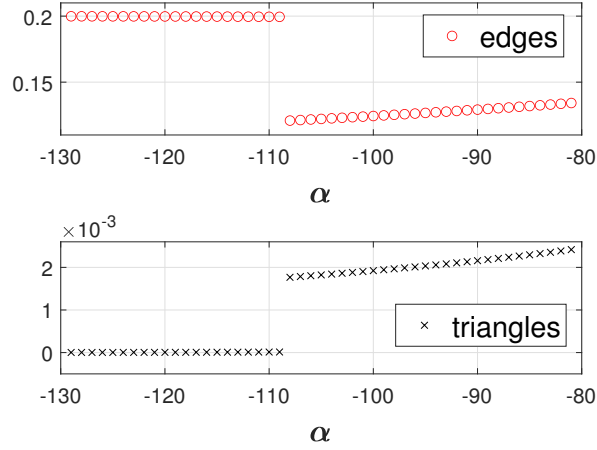


Figure 8: Density of edges and triangles of the optimizing graphon $g_{p_1^*(\alpha), p_2^*(\alpha)}^{(\frac{1}{2})}$ of (41) for varying α and $a = \frac{1}{2}$, $p = 0.4$. The discontinuities make evident the phase transition at the critical point $\tilde{\alpha}_c(p)$.

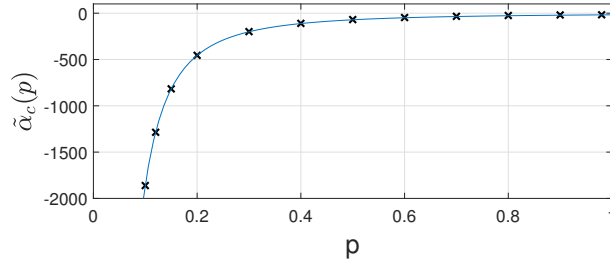


Figure 9: The critical point $\tilde{\alpha}_c(p)$ versus p together with the fitting curve $y = -17.66 x^{-2.027}$.

of eq. (20), i.e. $p_2 = u^*(\alpha)$, see Fig.5. Thus from eq. (b) we get a linear equation for p_1 :

$$\alpha p_1 u^*(\alpha) = \ln \left(\frac{u^*(\alpha)(1-p)}{p(1-u^*(\alpha))} \right),$$

that has the solution $p_1 = u^*(\alpha)$. Indeed, the equation $\alpha x^2 = \ln \left(\frac{x(1-p)}{p(1-x)} \right)$ is equivalent to the fixed-point equation (20) that has a unique solution for $\alpha < 0$.

From this discussion, we conclude that any maximizer with $a \neq \frac{1}{2}$ is necessarily the constant graphon $(u^*(\alpha), u^*(\alpha), a)$.

• **Case $a = \frac{1}{2}$.** In this case the function to be maximized takes the much simpler form $\mathcal{F}_\alpha(p_1, p_2, \frac{1}{2}) = \frac{\alpha}{12}(p_1^3 + 3p_1p_2^2) - \frac{1}{2}(I_p(p_1) + I_p(p_2))$ with the stationarity condition:

$$\begin{cases} \frac{\alpha}{4} [p_1^2 + p_2^2] - \frac{1}{2} \ln \left(\frac{p_1(1-p)}{p(1-p_1)} \right) = 0, & (a) \\ p_1p_2 - \ln \left(\frac{p_2(1-p)}{p(1-p_2)} \right) = 0. & (b) \end{cases} \quad (45)$$

We have computed the numerical solution $(p_1^*(\alpha), p_2^*(\alpha))$ for several values of p , as in Fig.6. The left panels show that crossing a critical value $\tilde{\alpha}_c(p)$ from above the transition between the constant graphon $(p_1^*(\alpha) = p_2^*(\alpha))$ and a 1-step replica symmetry breaking regime takes place. In particular, at $\tilde{\alpha}_c(p)$ the solution jumps towards the point $(0, p)$. The central and right columns of Fig.6 represent the solution below the critical value showing that, in the limit $\alpha \rightarrow -\infty$, the values $p_1^*(\alpha)$ and $p_2^*(\alpha)$ converge to the limits 0 and p , respectively, as conjectured in (26) and (27). A further evidence of this transition is given

in Fig.7 in which the curves $\mathcal{F}_\alpha(p_1^*(\alpha), p_2^*(\alpha), \frac{1}{2})$ and $\mu^{RS}(\alpha)$ are displayed. Above the critical value $\tilde{\alpha}_c(p)$ the two curves overlap thus revealing the replica symmetric phase whereas they separate below (replica breaking regime). For $\alpha < \tilde{\alpha}_c(p)$, the right panel Fig.7 shows that $\mathcal{F}_\alpha(p_1^*(\alpha), p_2^*(\alpha), \frac{1}{2})$ is very close to the asymptotic value (23) (the discrepancy vanishes as $\alpha \rightarrow -\infty$). Figure 8 represents the density of triangles and edges corresponding to the maximizer $g_{p_1^*(\alpha), p_2^*(\alpha)}^{(\frac{1}{2})}$. The former quantity is given in (42) while the latter is

$$e(g_{p_1, p_2}^{(a)}) = \int_{[0,1]^2} g_{p_1, p_2}^{(a)}(x, y) dx dy = p_1[a^2 + (1-a)^2] + 2p_2a(1-a).$$

The jump discontinuity shown in Fig.8, located at $\tilde{\alpha}_c(p)$, makes evident the existence of the transition that separates the replica symmetric phase, in which both quantities decrease for decreasing α , from the replica breaking phase. In this phase the density of edges gets close to the value $\frac{p}{2}$, i.e. the density of the limiting equipartite graphon (22), while the density of triangle jumps towards a value close to zero, being zero the density of triangles of the same graphon. Figure 9 shows the dependence of the critical value $\tilde{\alpha}_c(p)$ from p . The data are in good agreement with a fit $|\tilde{\alpha}_c(p)| \sim p^{-2}$.

We conclude our analysis by Fig.10, which displays the replica symmetric pressure (19), the 1-step replica symmetry breaking pressure (41) and the pressure obtained from the discretized variational problem. We can observe that the three curves perfectly match above the critical threshold $\tilde{\alpha}_c(p)$, whereas in the subcritical phase the pressure obtained from the solution of the discretized problem agrees with the 1RSB pressure and is strictly larger than the replica symmetric pressure. Figure 10 shows that $\tilde{\alpha}_c(p)$ is a singularity of $\alpha \rightarrow \mathcal{F}_\alpha(p_i^*(\alpha), p_b^*(\alpha), \frac{1}{2})$, at which the left and right derivatives are different.

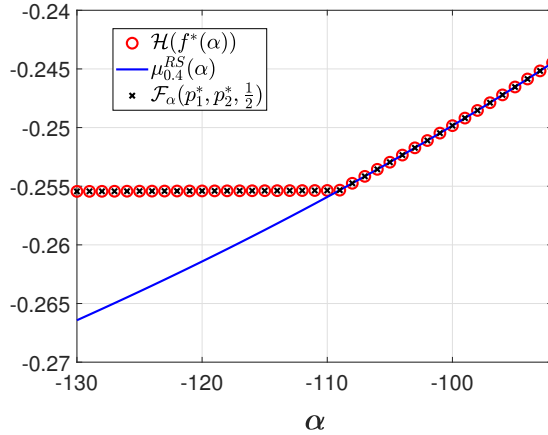


Figure 10: The plot represents $\mu_p^{1RSB}(\alpha) = \mathcal{F}_\alpha(p_1^*(\alpha), p_2^*(\alpha), \frac{1}{2})$ (black crosses), the solution $\mathcal{H}_{40}(f^*(\alpha))$ of the discretized problem (38) (red circles) together with $\mu_p^{RS}(\alpha)$. The picture refers to $p = 0.4$.

Acknowledgments

We acknowledge an enlightening discussion with Remco van der Hofstad. We thank M. Prato and S. Rebegoldi for making available to us their Gradient Projection code.

Appendix: $n = 3$

In order to show how the cloning algorithm works, we explicitly compute the cumulant generating function of triangles in the simplest case, that is the graph of size $n = 3$. Since the probability of the unique triangle

is p^3 , we have

$$\mu_{3,p}(\alpha) = \frac{1}{3} \ln \left\langle \exp \frac{\alpha}{3} T_3(X) \right\rangle_3^{ER} = \frac{1}{3} \ln [e^{\frac{\alpha}{3}} p^3 + 1 - p^3]. \quad (46)$$

We compute again (46) by applying the dynamics described in Sec.3 to a family of M clones. The three steps required to construct the edges of the graph G_3 are represented in Fig.11. The leafs of the tree

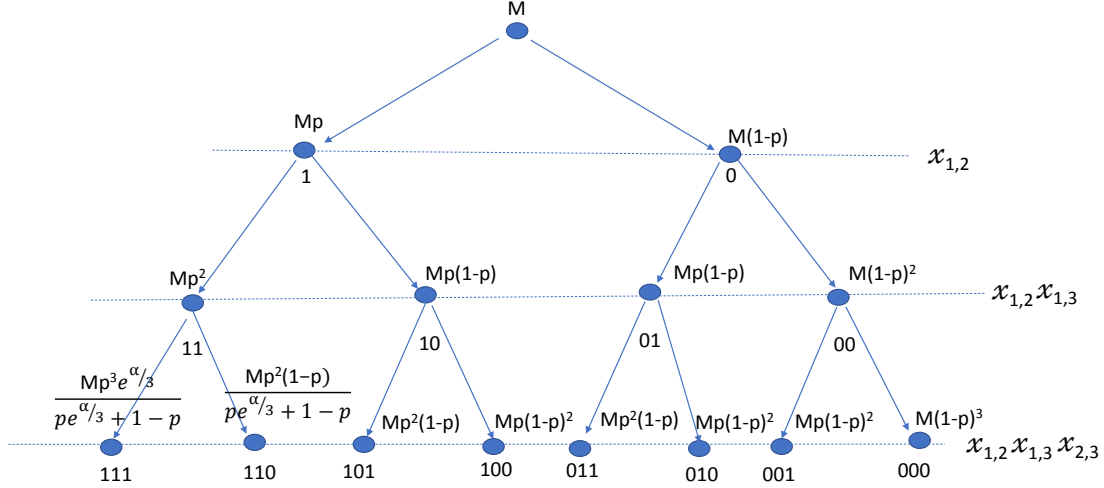


Figure 11: Graphical representation of the evolution step, according to the tilted dynamics $P_\alpha(\cdot, \cdot)$ of cloning scheme. The levels of the tree display the values of the elements of the adjacency matrix $x_{1,2}, x_{1,3}, x_{2,3}$ during the evolution leading to G_3 . M is the initial size of the population. In all levels the expected sizes of the sub-populations with a given configuration of values $x_{1,2}, x_{1,3}, x_{2,3}$ are reported.

represent the occupation variables of the three possible edges, $x_{1,2}, x_{1,3}, x_{2,3}$. In the first step the edge connecting two vertices, say 1 and 2, of each clone is added with probability p . Thus, in the clone population about pM graphs have the edge (1, 2) and $(1-p)M$ graphs have not this edge, see the first level in Fig.11. In the second step the edge (1, 3) is added, still with probability p , leading to four possible values for the pair $(x_{1,2}, x_{1,3})$. Thus, the expected numbers of types are $Mp^2, Mp(1-p), Mp(1-p), M(1-p)^2$, see level 2 in Fig.11. Let us observe that in the first two steps, since $\Delta T(x, y) = 0$, the original and tilted transition probabilities coincide: $P_\alpha(x, y) = P(x, y)$, see (32). The situation changes in the last step. Indeed, the configuration of edges $x = 11$ (which means that $x_{1,2} = 1$ and $x_{1,3} = 1$) may evolve to $y = 111$, with $\Delta T(x, y) = 1$, or to $y = 110$, in which case $\Delta T(x, y) = 0$. For all the other configurations x we have $\Delta T(x, y) = 0$. Thus, see the definition (32) of the tilted probability:

$$P_\alpha(11, 111) = \frac{e^{\frac{\alpha}{3}} p}{e^{\frac{\alpha}{3}} p + 1 - p}, \quad P_\alpha(11, 110) = \frac{1 - p}{e^{\frac{\alpha}{3}} p + 1 - p}, \quad (47)$$

being $P(11, 111) = p$, $P(11, 110) = 1 - p$ and $k_\alpha(11) = pe^{\frac{\alpha}{3}} + 1 - p$. Then, the probability of the paths connecting the root ϕ to the leafs 111, respectively 110, will be, :

$$\mathbb{P}(\phi, 111) = \frac{e^{\frac{\alpha}{3}} p^3}{e^{\frac{\alpha}{3}} p + 1 - p}, \quad \mathbb{P}(\phi, 110) = \frac{p^2(1 - p)}{e^{\frac{\alpha}{3}} p + 1 - p}. \quad (48)$$

The average in (33) can be computed as the sum over the paths from the root to the leafs (equivalently, as the sum over the leafs). Recalling that $k_\alpha(11) = pe^{\frac{\alpha}{3}} + 1 - p$ and observing also that $k_\alpha(x) = 1$ if

$x \neq 11$, from (33) we have:

$$\mu_{3,p}(\alpha) = \frac{1}{3} \ln \left[\mathbb{P}(\phi, 111)k_\alpha(11) + \mathbb{P}(\phi, 110)k_\alpha(11) + \sum_{x \notin \{111, 110\}} \mathbb{P}(\phi, x) \right] \quad (49)$$

$$= \frac{1}{3} \ln \left[\frac{p^3 e^{\frac{\alpha}{3}}}{pe^{\frac{\alpha}{3}} + 1 - p} (pe^{\frac{\alpha}{3}} + 1 - p) + \frac{p^2(1-p)}{pe^{\frac{\alpha}{3}} + 1 - p} (pe^{\frac{\alpha}{3}} + 1 - p) + (1 - p^2) \right] \quad (50)$$

$$= \frac{1}{3} \ln [p^3 e^{\frac{\alpha}{3}} + 1 - p^3], \quad (51)$$

The cloning algorithm simulates (51) by producing a final population of expected size

$$M_3 = M \mathbb{P}(\phi, 111)k_\alpha(11) + M \mathbb{P}(\phi, 110)k_\alpha(11) + \sum_{x \notin \{111, 110\}} M \mathbb{P}(\phi, x).$$

Then, the cumulant generating function is computed from the size of the final population M_3 as

$$\mu_{3,p}(\alpha) = \frac{1}{3} \ln \left[\frac{M_3}{M} \right]. \quad (52)$$

References

- [1] Letizia Angeli, Stefan Grosskinsky, and Adam M Johansen. Limit theorems for cloning algorithms. *arXiv preprint arXiv:1902.00509*, 2019.
- [2] Letizia Angeli, Stefan Grosskinsky, Adam M Johansen, and Andrea Pizzoferrato. Rare event simulation for stochastic dynamics in continuous time. *Journal of Statistical Physics*, 176(5):1185–1210, 2019.
- [3] David Aristoff and Lingjiong Zhu. Asymptotic structure and singularities in constrained directed graphs. *Stochastic Processes and their Applications*, 125(11):4154–4177, 2015.
- [4] Jonathan Barzilai and Jonathan M Borwein. Two-point step size gradient methods. *IMA journal of numerical analysis*, 8(1):141–148, 1988.
- [5] Shankar Bhamidi, Guy Bresler, and Allan Sly. Mixing time of Exponential Random Graphs. In *Foundations of Computer Science, 2008. FOCS'08. IEEE 49th Annual IEEE Symposium on*, pages 803–812. IEEE, 2008.
- [6] Shankar Bhamidi, Jan Hannig, Chia Ying Lee, James Nolen, et al. The importance sampling technique for understanding rare events in Erdős–Rényi random graphs. *Electronic Journal of Probability*, 20, 2015.
- [7] Béla Bollobás. *Modern graph theory*, volume 184. Springer Science & Business Media, 2013.
- [8] Christian Borgs, Jennifer Chayes, László Lovász, Vera T Sós, and Katalin Vesztegombi. Counting graph homomorphisms. In *Topics in discrete mathematics*, pages 315–371. Springer, 2006.
- [9] Christian Borgs, Jennifer T Chayes, László Lovász, Vera T Sós, and Katalin Vesztegombi. Convergent sequences of dense graphs i: subgraph frequencies, metric properties and testing. *Advances in Mathematics*, 219(6):1801–1851, 2008.
- [10] Christian Borgs, Jennifer T Chayes, László Lovász, Vera T Sós, and Katalin Vesztegombi. Convergent sequences of dense graphs ii. Multiway cuts and statistical physics. *Annals of Mathematics*, pages 151–219, 2012.
- [11] Federico Carollo and Carlos Pérez-Espigares. Entanglement statistics in markovian open quantum systems: a matter of mutation and selection. *arXiv preprint arXiv:1910.13940*, 2019.

- [12] Sourav Chatterjee. *Large Deviations for Random Graphs: École D'Été de Probabilités de Saint-Flour XLV-2015*, volume 2197. Springer, 2017.
- [13] Sourav Chatterjee and Amir Dembo. Nonlinear large deviations. *Advances in Mathematics*, 299:396–450, 2016.
- [14] Sourav Chatterjee and Partha S Dey. Applications of stein method for concentration inequalities. *The Annals of Probability*, 38(6):2443–2485, 2010.
- [15] Sourav Chatterjee and Persi Diaconis. Estimating and understanding Exponential Random Graph models. *The Annals of Statistics*, 41(5):2428–2461, 2013.
- [16] Sourav Chatterjee and Srinivasa Varadhan. The large deviation principle for the Erdős-Rényi random graph. *European Journal of Combinatorics*, 32(7):1000–1017, 2011.
- [17] Amir Dembo, Eyal Lubetzky, et al. A large deviation principle for the erdős-rényi uniform random graph. *Electronic Communications in Probability*, 23, 2018.
- [18] F den Hollander, M Mandjes, A Roccaverde, NJ Starreveld, et al. Ensemble equivalence for dense graphs. *Electronic Journal of Probability*, 23, 2018.
- [19] Cristian Giardinà, Jorge Kurchan, Vivien Lecomte, and Julien Tailleur. Simulating rare events in dynamical processes. *Journal of statistical physics*, 145(4):787–811, 2011.
- [20] Cristian Giardinà, Jorge Kurchan, and Luca Peliti. Direct evaluation of large-deviation functions. *Physical review letters*, 96(12):120603, 2006.
- [21] Pablo I Hurtado, Carlos P Espigares, Jesús J del Pozo, and Pedro L Garrido. Thermodynamics of currents in nonequilibrium diffusive systems: theory and simulation. *Journal of Statistical Physics*, 154(1-2):214–264, 2014.
- [22] Richard Kenyon, Charles Radin, Kui Ren, and Lorenzo Sadun. Multipodal structure and phase transitions in large constrained graphs. *Journal of Statistical Physics*, 168(2):233–258, 2017.
- [23] László Lovász. Large networks and graph limits, volume 60 of american mathematical society colloquium publications. *American Mathematical Society, Providence, RI*, 22, 2012.
- [24] László Lovász and Balázs Szegedy. Limits of dense graph sequences. *Journal of Combinatorial Theory, Series B*, 96(6):933–957, 2006.
- [25] Eyal Lubetzky and Yufei Zhao. On replica symmetry of large deviations in random graphs. *Random Structures & Algorithms*, 47(1):109–146, 2015.
- [26] Mark E. J. Newman. *Networks: An Introduction*. Oxford University Press, 2010.
- [27] Carlos Pérez-Espigares and Pablo I Hurtado. Sampling rare events across dynamical phase transitions. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 29(8):083106, 2019.
- [28] Oleg Pikhurko and Alexander Razborov. Asymptotic structure of graphs with the minimum number of triangles. *Combinatorics, Probability and Computing*, 26(1):138–160, 2017.
- [29] Charles Radin and Lorenzo Sadun. Phase transitions in a complex network. *Journal of Physics A: Mathematical and Theoretical*, 46(30):305002, 2013.
- [30] Charles Radin and Lorenzo Sadun. Singularities in the entropy of asymptotically large simple graphs. *Journal of Statistical Physics*, 158(4):853–865, 2015.
- [31] Hugo Touchette. Equivalence and nonequivalence of ensembles: Thermodynamic, macrostate, and measure levels. *Journal of Statistical Physics*, 159(5):987–1016, 2015.
- [32] Yufei Zhao. On the lower tail variational problem for random graphs. *Combinatorics, Probability and Computing*, 26(2):301–320, 2017.