Breaking of ensemble equivalence for complex networks

Roccaverde, A.

CHAPTER 6

Breaking of Ensemble Equivalence for Perturbed Erdős-Rényi Random Graphs

This chapter is based on:
F. den Hollander, M. Mandjes, A. Roccaverde, and N. Starreveld. Breaking of ensemble equivalence for perturbed erdős-rényi random graphs. *arXiv:1807.07750*

Abstract

In a previous paper we analysed a simple undirected random graph subject to constraints on the total number of edges and the total number of triangles. We considered the dense regime in which the number of edges per vertex is proportional to the number of vertices. We showed that, as soon as the constraints are *frustrated*, i.e., do not lie on the Erdős-Rényi line, there is breaking of ensemble equivalence, in the sense that the specific relative entropy per edge of the *microcanonical ensemble* with respect to the *canonical ensemble* is strictly positive in the limit as the number of vertices tends to infinity. In the present paper we analyse what happens near the Erdős-Rényi line. It turns out that the way in which the specific relative entropy tends to zero depends on whether the total number of triangles is slightly larger or slightly smaller than typical. We identify what the constrained random graph looks like asymptotically in the microcanonical ensemble.

§6.1 Introduction

In this paper we analyse random graphs that are subject to *constraints*. Statistical physics prescribes what probability distribution on the set of graphs we should choose when we want to model a given type of constraint [53]. Two important choices are:

(1) The *microcanonical ensemble*, where the constraints are *hard* (i.e., are satisfied by each individual graph).

(2) The *canonical ensemble*, where the constraints are *soft* (i.e., hold as ensemble averages, while individual graphs may violate the constraints).

For random graphs that are large but finite, the two ensembles are obviously different and, in fact, represent different empirical situations. Each ensemble represents the unique probability distribution with *maximal entropy* respecting the constraints. In the limit as the size of the graph diverges, the two ensembles are traditionally *assumed* to become equivalent as a result of the expected vanishing of the fluctuations of the soft constraints, i.e., the soft constraints are expected to behave asymptotically like hard constraints. This assumption of *ensemble equivalence* is one of the corner stones of statistical physics, but it does *not* hold in general (see [97] for more background).

In a series of papers the question of possible breaking of ensemble equivalence was investigated for various choices of the constraints, including the degree sequence and the total number of edges, wedges and triangles. Both the *sparse regime* (where the number of edges per vertex remains bounded) and the *dense regime* (where the number of edges per vertex is of the order of the number of vertices) have been considered. The effect of *community structure* on ensemble equivalence has been investigated as well. Relevant references are [48], [50], [37], [93] and [92]. In [37] we considered a random graph subject to constraints on the total number of edges and the total number of triangles, in the dense regime. With the help of *large deviation theory for graphons*, see [31], we derived a variational formula for $s_\infty = \lim_{n\to\infty} n^{-2} s_n$, where n is the number of vertices and s_n is the *relative entropy* of the microcanonical ensemble with respect to the canonical ensemble. We found that $s_\infty > 0$ when the constraints are *frustrated*. In the present paper we analyse the behaviour of s_∞ when the constraints are close to but different from those of the Erdős-Rényi random graph, and we identify what the constrained random graph looks like asymptotically in the microcanonical ensemble. It turns out that the behaviour changes when the total number of triangles is larger, respectively, smaller than that of the Erdős-Rényi random graph with a given total number of edges.

While breaking of ensemble equivalence is a relatively new concept in the theory of random graphs, there are many studies on the asymptotic structure of random graphs. In the pioneering work [31], followed by [70], the large deviation principle for dense Erdős-Rényi random graphs was proven and the asymptotic structure of constrained Erdős-Rényi random graphs was described as the solution of a variational problem. In the past few years significant progress has been made regarding sparse random graphs as well. We refer the reader to [30], [36], [71] and [105]. Two other random graph models that have been extensively studied are the exponential random graph model

and the constrained exponential random graph model. Exponential random graphs, which are related to the canonical ensemble we consider in this paper, were introduced rather early in the physics literature, see [80] and the references therein, and rigorously analysed in detail in [12] and [29]. In [12] the authors investigated the mixing time of the Glauber dymanics and they showed that, in some cases, graphs drawn from the exponential random graph model, behave asymptotically like Erdős-Rényi random graphs with a biased parameter. In [29] the authors verified and generalised this result using the machinery developed in [31]. Their main result was an asymptotic expression for the logarithm of the partition function in terms of a variational problem. Additionally, they showed that, in the edge-triangle model, a phase transition, which is defined as a discontinuity in the derivative of the logarithm of the partition function, occurs for specific values of the parameters. The existence of phase transitions in the exponential random graph model was investigated further in [87] and [101] and for directed graphs in [4]. An analysis of sparse exponential random graphs was carried out in [103]. A second random graph model, which is also related to the random graph models we study in this paper, and has received significant attention in the literature, is the constrained exponential random graph model, we refer the reader to [3], [63], [67] and [102] for a detailed description and analysis. A stream of research that is relevant to our work concerns the asymptotic description of the structure of graphs drawn from the microcanonical ensemble with a constraint on the edge and triangle density. In [86] the authors studied the behavior of random graphs with edge and triangle densities close to the Erdős-Rényi curve. They managed to identify the scaling behavior close to the curve by proving a bound on the entropy function. In one of the results in this paper we rigorously prove the results of [86] and we determine the exact structure of constrained random graphs close to the Erdős-Rényi curve. The same question was investigated in [75] for a constraint on the edge and triangle density close to the lower boundary curve of the admissibility region. In [64] the authors managed to determine, through extensive simulations, curves in the admissibility region where phase transitions occur in the structure of constrained random graphs.

The remainder of this paper is organised as follows. In Section 6.2 we define the two ensembles, give the definition of equivalence of ensembles in the dense regime, some basic facts about graphons and we the *variational representation* of s_∞ derived in [37] when the constraints are on the total numbers of subgraphs drawn from a finite collection of subgraphs. We also recall the analysis of s_∞ in [37] for the special case where the subgraphs are the edges and the triangles. In Section 6.3 we state our main theorems. Proofs are given in Sections 6.4 and 6.5.

§6.2 Definitions and preliminaries

The microcanonical and the canonical ensemble, as well as the relative entropy density have been defined in Section 1.4.1 and 1.4.2. Graphons and their properties have been defined in Section 5.2.2 5.2.3 and 5.3 of Chapter 5. In this section we recall the definition of ensemble equivalence in the dense regime and the main two theorems of Chapter 5.

6.2.1 Definition. Following [37] P_{mic} and P_{can} are said to be equivalent in the dense regime when

$$s_\infty := \lim \frac{1}{n^2} S_n(P_{\text{mic}} \mid P_{\text{can}}) = 0. \tag{6.1}$$

The key result in [37] is the following variational formula for s_∞.

6.2.2 Theorem. [37] *Subject to* (5.30), (5.32) *and* (5.33),

$$\lim_{n\to\infty} n^{-2} S_n(P_{\text{mic}} \mid P_{\text{can}}) =: s_\infty \tag{6.2}$$

with

$$s_\infty = \sup_{\tilde{h}\in\tilde{W}} \left[\vec{\theta}^*_\infty \cdot \vec{T}(\tilde{h}) - I(\tilde{h})\right] - \sup_{\tilde{h}\in\tilde{W}^*} \left[\vec{\theta}^*_\infty \cdot \vec{T}(\tilde{h}) - I(\tilde{h})\right]. \tag{6.3}$$

Theorem 6.2.2 and the compactness of $\tilde{W}^*$ give us a *variational characterisation* of ensemble equivalence: $s_\infty = 0$ if and only if at least one of the maximisers of $\vec{\theta}^*_\infty \cdot \vec{T}(\tilde{h}) - I(\tilde{h})$ in $\tilde{W}$ also lies in $\tilde{W}^* \subset \tilde{W}$. Equivalently, $s_\infty = 0$ when at least one the maximisers of $\vec{\theta}^*_\infty \cdot \vec{T}(\tilde{h}) - I(\tilde{h})$ satisfies the hard constraint. Theorem 6.2.2 allows us to identify cases where ensemble equivalence holds ($s_\infty = 0$) or is broken ($s_\infty > 0$). In [37] a detailed analysis was given for the special case where the constraint is on the total number of edges and the total number of triangles. The analysis in [37] relied on the large deviation principle for dense Erdős-Rényi random graphs established in [31]. The function defined in (5.19) plays a crucial role and is related to the rate function of the large deviation principle.

6.2.3 Theorem. [37] *For the edge-triangle model, $s_\infty = 0$ when*

- $T_2^* = T_1^{*3}$,
- $0 < T_1^* \leq \frac{1}{2}$ *and* $T_2^* = 0$,

while $s_\infty > 0$ when

- $T_2^* \neq T_1^{*3}$ *and* $T_2^* \geq \frac{1}{8}$,
- $T_2^* \neq T_1^{*3}$, $0 < T_1^* \leq \frac{1}{2}$ *and* $0 < T_2^* < \frac{1}{8}$,
- (T_1^*, T_2^*) *lies on the scallopy curve in Figure* 6.1.

Here, T_1^*, T_2^* are in fact the limits $T_{1,\infty}^*, T_{2,\infty}^*$ in (5.32), but in order to keep the notation light we now also suppress the index ∞.
Theorem 6.2.3 is illustrated in Fig. 6.1. The region on and between the blue curves corresponds to the choices of (T_1^*, T_2^*) that are *graphical*, i.e., there exists a graph with edge density T_1^* and triangle density T_2^*. The red curves represent ensemble equivalence, the blue curves and the grey region represent breaking of ensemble equivalence, while in the white region between the red curve and the lower blue curve we do not know what happens. Breaking of ensemble equivalence arises from *frustration* between the values of T_1^* and T_2^*.

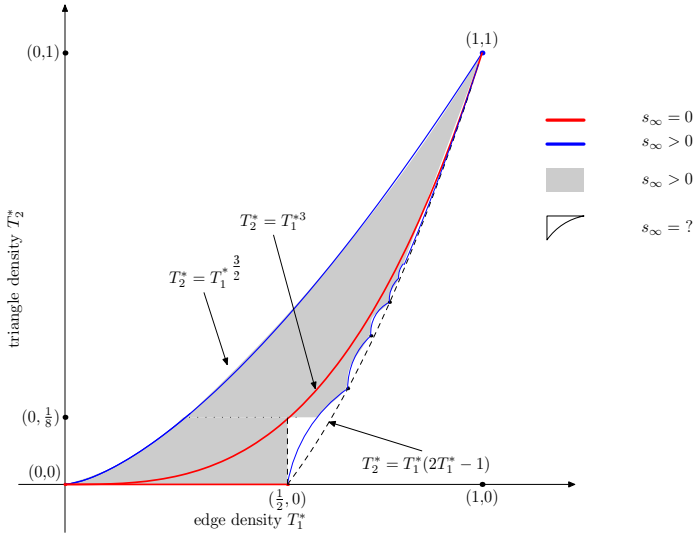

Figure 6.1: The admissible edge-triangle density region is the region on and between the blue curves [86].

The lower blue curve, called the *scallopy curve*, consist of infinitely many pieces labelled by $\ell \in \mathbb{N} \setminus \{1\}$. The ℓ-th piece corresponds to $T_1^* \in (\frac{\ell-1}{\ell}, \frac{\ell}{\ell+1}]$ and a T_2^* that is a function of T_1^* given by

$$T_2^* = \frac{(\ell-1)\left(\ell - 2\sqrt{\ell(\ell - T_1^*(\ell+1))}\right)\left(\ell + \sqrt{\ell(\ell - T_1^*(\ell+1))}\right)^2}{\ell^2(\ell+1)^2}. \tag{6.4}$$

We refer the reader to [83], [85], [86] and [88] for more details.

§6.3 Theorems

In this section we present our results which address the following two issues:

- In Theorems 6.3.1–6.3.3 we identify the scaling behaviour of s_∞ for fixed T_1^* and $T_2^* \downarrow T_1^{*3}$, respectively, $T_2^* \uparrow T_1^{*3}$. It turns out that the way in which s_∞ tends to zero differs in the two cases.
- In Propositions 6.3.5–6.3.7 we characterise some possible asymptotic structures of random graphs drawn from the microcanonical ensemble when the hard constraint is on the edge and triangle density. Our results indicate that the structure of the graphs differs for $T_2^* \downarrow T_1^{*3}$, respectively, $T_2^* \uparrow T_1^{*3}$.

In the sequel we make the following two assumptions:

Assumption 1. Fix the edge density $T_1^* \in (0,1)$ and consider the triangle density $T_1^{*3} + \epsilon$, for some ϵ either positive or negative. For this pair of constraints we consider

the Lagrange multipliers $\vec{\theta}^*_\infty(\epsilon) := (\theta^*_1(\epsilon), \theta^*_2(\epsilon))$ as defined in Section 5.3.1. Then, for ϵ sufficiently small, we have the representation

$$\sup_{\tilde{h}\in\tilde{W}} \Big[\theta^*_1(\epsilon)T_1(\tilde{h}) + \theta^*_2(\epsilon)T_2(\tilde{h}) - I(\tilde{h})\Big] = \theta_1 T^*_1 - I(T^*_1) + (\gamma_1 T^*_1 + \gamma_2 T^{*3}_1)\epsilon + O(\epsilon^2), \tag{6.5}$$

where $\theta_1 := \theta_1(0), \gamma_1 = \theta'_1(0)$ and $\gamma_2 = \theta'_2(0)$.

In Section 6.4.1 we show that Assumption 1 is true when $T^*_1 \in [\frac{1}{2}, 1)$. For $T^*_1 \in (0, \frac{1}{2})$ we can prove (6.8) and (6.9) below, but with $\geq$ replacing the equality. If Assumption 1 is true, then we again obtain (6.8) and (6.9) with equality. If it fails, then we have strict inequality.

Assumption 2. Fix the edge density $T^*_1 \in (0, 1)$ and consider the triangle density $T^{*3}_1 + \epsilon$, for some ϵ either positive or negative. For this pair of constraints we consider the microcanonical entropy

$$-J(\epsilon) := \sup\{-I(\tilde{h}) :\ \tilde{h} \in \tilde{W},\ T_1(\tilde{h}) = T^*_1, T_2(\tilde{h}) = T^{*3}_1 + \epsilon\}. \tag{6.6}$$

Then for ϵ sufficiently small the solution of (6.6), denoted by h^*_ϵ, has the following form

$$h^*_\epsilon = T^*_1 + g_\epsilon, \qquad \text{where} \qquad g_\epsilon = g_{11}1_{I\times I} + g_{12}1_{(I\times J)\cup(J\times I)} + g_{22}1_{J\times J}, \tag{6.7}$$

with $g_{11}, g_{12}, g_{22} \in [-T^*_1, 1 - T^*_1]$ and $I, J \subset [0, 1]$.

Assumption 2 is based on the following intuitive argument. Suppose we want to maximise the microcanonical entropy among all piecewise constant graphons. Then we expect the entropy to decrease when we add more structure, i.e., more steps, in the graphon. A piecewise constant graphon with m steps corresponds to a random graph where the vertices are divided into m groups, and within each group we make an ER random graph with some probability. We expect that the microcanonical entropy will decrease as m increases. This statement is also supported by extensive numerical experiments performed in [65].

The methodology we rely on in order to analyse the variational problem in (6.6) does not always identify the exact optimal graphon. It identifies a candidate optimal graphon, which is sufficient, in some cases for the scaling behaviour of the relative entropy s_∞. We call these graphons *balance optimal*. Roughly speaking, a balance optimal graphon is obtained when solving the optimisation problem in (6.6) in a smaller class of graphons than the whole class of graphons that satisfy the hard constraint. This is the class of graphons satisfying the conditions in Assumption 1 and such that the values g_{11}, g_{12}, g_{22} all correspond to contributions of the same order. The precise definition of a balance optimal graphon is given in Section 6.5. We want to investigate in this chapter whether the global maximizer of (6.6) lies in this smaller class of graphons. We show that balance optimisers have specific structural properties. But, for the case of a perturbation upwards, the unique optimal graphon does not lie in this class, and this happens because $\lambda(I)$ gets very small as $\epsilon \downarrow 0$ while g_{11} stays constant. We refer the reader to [66]. For the case of a perturbation downwards the

exact structure of the unique optimal graphon is still not known: the only results we are aware of come from an extensive numerical study, [63]. From this numerical study it seems that, at least for $T_1^* \in (0, \tilde{T}_1^*)$, with $\tilde{T}_1^* \approx 0.44$, the unique global optimiser is indeed a balance optimal graphon. In this chapter we investigate this question further by identifying the balance optimal graphons and comparing them with the results established numerically in [63].

Balance optimal graphons are candidate optimisers of $J(\epsilon)$. In what follows, because all the graphons we derive are balance optimal graphons, we simply speak of optimal graphons. When at some point a clear distinction is needed we say so. Another important feature is that balance optimal graphons are in general not unique. In the following sections we construct various balance optimal graphons, exhibiting the different structures that can emerge. The variational problem $J(\epsilon)$ in (6.6) has been solved in [66] for the case $T_2^* > T_1^{*3}$, while the case $T_2^* < T_1^{*3}$ still remains unsolved. In this chapter we consider only a small perturbation around the typical values, but the advantage of our method is that it is simpler and yields more intuition about the way the constraint is attained. Moreover, it also applies for the case $\epsilon < 0$, which has not been rigorously analysed before. In [63] the authors identify the maximizers of the microcanonical entropy numerically. The optimal graphons obtained numerically in [63] agree structurally with the balance optimal graphons that we find.

6.3.1 Theorem. *For $T_1^* \in (0,1)$ and $T_1^* \neq \frac{1}{2}$,*

$$\lim_{\epsilon \downarrow 0} \epsilon^{-1} s_\infty(T_1^*, T_1^{*3} + 3T_1^*\epsilon) = \frac{6}{1-2T_1^*} \log \frac{T_1^*}{1-T_1^*} \in (0,\infty). \tag{6.8}$$

6.3.2 Theorem. *For $T_1^* \in (0, \frac{1}{2}]$,*

$$\limsup_{\epsilon \downarrow 0} \epsilon^{-2/3} s_\infty(T_1^*, T_1^{*3} - T_1^{*3}\epsilon) \leq \frac{1}{4} \frac{T_1^*}{1-T_1^*} \in (0,\infty). \tag{6.9}$$

6.3.3 Theorem. *For $T_1^* \in (\frac{1}{2}, 1)$,*

$$\limsup_{\epsilon \downarrow 0} \epsilon^{-2/3} s_\infty(T_1^*, T_1^{*3} - T_1^{*3}\epsilon) \leq f(T_1^*, \bar{T}_1^*) \in (0,\infty), \tag{6.10}$$

where $\bar{T}_1^* \in (-T_1^*, 0)$ is the unique point where the function $x \to f(T_1^*, x)$, defined by

$$f(T_1^*, x) := T_1^{*2} \frac{I(T_1^* + x) - I(T_1^*) - I'(T_1^*)x}{x^2}, \quad x \in (-T_1^*, 0), \tag{6.11}$$

attains its global minimum.

We illustrate these results in Figure 6.2. In the left panel we plot the limits in the right-hand side of (6.9)–(6.10) as a function of T_1^*. In the right panel we plot $s_\infty(T_1^*, T_1^{*3} + \epsilon)$ as a function of ϵ, for ϵ sufficiently small, and for four different values of T_1^*.

6.3.4 Remark. We believe, and there is numerical evidence in [63], that the results in (6.9) and (6.10) hold with equality and that the corresponding limits exist.

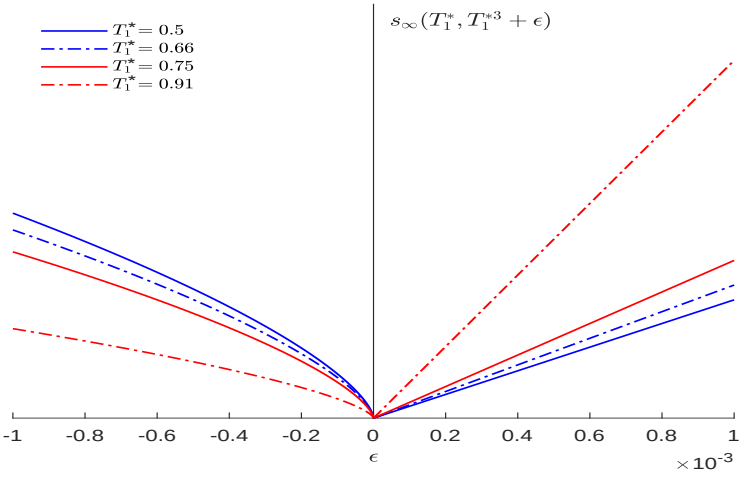

Figure 6.2: Limit of scaled s_∞ as a function of ϵ for ϵ sufficiently small.

In Proposition 6.3.5–6.3.7 below we identify the structure of balance optimal graphons corresponding to the perturbed constraints in the microcanonical ensemble in the limit as $n \to \infty$.

6.3.5 Proposition. *When the ER-line is approached from above, a balance optimal graphon is given by*

$$h = T_1^* + \sqrt{\epsilon}\, g^* + O(\epsilon) \qquad \text{(global perturbation)} \tag{6.12}$$

with g^* given by

$$g^*(x,y) = \begin{cases} 2, & (x,y) \in [0,\frac{1}{2}]^2, \\ 0, & (x,y) \in [0,\frac{1}{2}] \times (\frac{1}{2},1] \cup (\frac{1}{2},1] \times [0,\frac{1}{2}], \\ -2, & (x,y) \in (\frac{1}{2},1]^2. \end{cases} \tag{6.13}$$

It is important to mention that the balance optimal graphon determined in (6.13) is not unique, in the sense that there are multiple graphons that are balance optimal. From Proposition 6.3.5 we also see that it is possible that the class of balance optimisers does not contain the actual unique optimiser of $J(\epsilon)$. For this pair of constraints, and from [66], we have that the actual unique optimiser, denoted by h_ϵ^*, is given by

$$h_\epsilon^*(x,y) = \begin{cases} h_{11}, & (x,y) \in [0,\lambda\epsilon]^2, \\ 1 - T_1^* + h_1\epsilon, & (x,y) \in [0,\lambda\epsilon] \times (\lambda\epsilon,1] \cup (\lambda\epsilon,1] \times [0,\lambda\epsilon], \\ T_1^* + h_2\epsilon, & (x,y) \in (\lambda\epsilon,1]^2, \end{cases} \tag{6.14}$$

where

$$\lambda := \frac{1}{(1-2T_1^*)^2}, \qquad h_1 := \frac{1}{2}h_2, \qquad h_2 := -\frac{2}{1-2T_1^*}. \tag{6.15}$$

The term h_{11} solves the equation $I'(h_{11}) = 3I'(1 - T_1^*)$ and is constant as $\epsilon \downarrow 0$. For details on this issue we refer to [66]. As mentioned above, balance optimal graphons have the structural property that g_{11}, g_{12}, g_{22} all contribute equally to the constraint. This is not the case for the graphon in (6.14) because only g_{12} and g_{22} contribute to the constraint to leading order. The exact computations are provided in Section 6.5.

From (6.13) and (6.14) we observe that balance optimal graphons can have structures very different from the optimal graphons.

6.3.6 Proposition. *When the ER-line is approached from below and $T_1^* \in (0, \frac{1}{2}]$, a balance optimal graphon is given by*

$$h = T_1^* + \epsilon^{1/3} g^* + O(\epsilon^{1/3}) \qquad \text{(global perturbation)} \tag{6.16}$$

with g^* given by

$$g^*(x,y) = \begin{cases} -T_1^*, & (x,y) \in [0, \frac{1}{2}]^2, \\ T_1^*, & (x,y) \in [0, \frac{1}{2}] \times (\frac{1}{2}, 1] \cup (\frac{1}{2}, 1] \times [0, \frac{1}{2}], \\ -T_1^*, & (x,y) \in (\frac{1}{2}, 1]^2. \end{cases} \tag{6.17}$$

This g^* is not unique, in the sense that there are multiple graphons that are balance optimal.

6.3.7 Proposition. *When the ER-line is approached from below and $T_1^* \in (\frac{1}{2}, 1)$, the unique balance optimal graphon is given by*

$$h = T_1^* + g_\epsilon^* \qquad \text{(local perturbation)} \tag{6.18}$$

with g_ϵ^* defined by

$$g_\epsilon^*(x,y) := \begin{cases} \frac{T_1^{*2}}{T_1^*} \epsilon^{2/3}, & (x,y) \in [0, 1 - \frac{T_1^*}{|T_1^*|} \epsilon^{1/3}]^2 \\ T_1^* \epsilon^{1/3}, & (x,y) \in [0, 1 - \frac{T_1^*}{|T_1^*|} \epsilon^{1/3}] \times [1 - \frac{T_1^*}{|T_1^*|} \epsilon^{1/3}, 1] \quad \text{or} \\ & (x,y) \in [1 - \frac{T_1^*}{|T_1^*|} \epsilon^{1/3}, 1] \times [0, 1 - \frac{T_1^*}{|T_1^*|} \epsilon^{1/3}], \\ \bar{T}_1^*, & (x,y) \in [1 - \frac{T_1^*}{|T_1^*|} \epsilon^{1/3}, 1]^2, \end{cases} \tag{6.19}$$

with $\bar{T}_1^* \in (-T_1^*, 0)$ defined in Theorem 6.3.3.

In conclusion, Theorems 6.3.1–6.3.3 say that at a fixed density of the edges it is less costly in terms of relative entropy to increase the density of triangles than to decrease it. The ER-line represents a crossover in the cost (see Figure 6.2, right panel). Above the ER-line the cost is linear in the distance, below the ER-line the cost is proportional to the $\frac{2}{3}$-power of the distance. Propositions 6.3.5–6.3.7 show that the optimal perturbation of the ER-graphon is global above the ER-line and below the ER-line when the edge density is less than $\frac{1}{2}$ and local below the ER-line when the edge density is larger than $\frac{1}{2}$.

§6.4 Proofs of Theorems 6.3.1-6.3.3

In this section we prove Theorems 6.3.1–6.3.3. Along the way we use the results given in Propositions 6.3.5–6.3.7, which we prove in Section 6.5.

§6.4.1 Proof of Theorem 6.3.1

For ease of notation we drop the superscript $*$ from the constraint on the edge density and write T_1 instead of T_1^*. Let

$$T_1(\epsilon) = T_1, \qquad T_2(\epsilon) = T_1^3 + 3T_1\epsilon. \tag{6.20}$$

The factor $3T_1$ appearing in front of the ϵ is put in for convenience. We know that for every pair of graphical constraints $(T_1(\epsilon), T_2(\epsilon))$ there exists a unique pair of Lagrange multipliers $(\theta_1(\epsilon), \theta_2(\epsilon))$ corresponding to these constraints. For an elaborate discussion on this issue we refer the reader to [37]. By considering the Taylor expansion of the Lagrange multipliers $(\theta_1(\epsilon), \theta_2(\epsilon))$ around $\epsilon = 0$, we obtain

$$\theta_1(\epsilon) = \theta_1 + \gamma_1\epsilon + \tfrac{1}{2}\Gamma_1\epsilon^2 + O(\epsilon^3), \qquad \theta_2(\epsilon) = \gamma_2\epsilon + \tfrac{1}{2}\Gamma_2\epsilon^2 + O(\epsilon^3), \tag{6.21}$$

where

$$\theta_1(0) = \theta_1 = I'(T_1),\ \gamma_1 = \theta_1'(0),\ \Gamma_1 = \theta_1''(0),\ \theta_2(0) = 0,\ \gamma_2 = \theta_2'(0),\ \Gamma_2 = \theta_2''(0). \tag{6.22}$$

We denote the two terms in the expression for s_∞ in (6.3) by I_1, I_2, i.e.,

$$s_\infty = \sup_{\tilde{h}\in\tilde{W}} \left[\vec{\theta}_\infty \cdot \vec{T}(\tilde{h}) - I(\tilde{h})\right] - \sup_{\tilde{h}\in\tilde{W}^*} \left[\vec{\theta}_\infty \cdot \vec{T}(\tilde{h}) - I(\tilde{h})\right] = I_1 - I_2, \tag{6.23}$$

and we let $s_\infty(\epsilon)$ denote the relative entropy corresponding to the perturbed constraints. We distinguish between the cases $T_1 \in [\frac{1}{2}, 1)$ and $T_1 \in (0, \frac{1}{2})$.

Case I $T_1 \in [\frac{1}{2}, 1)$**:** From [37, Section 5], if $T_1 \in [\frac{1}{2}, 1)$ and $T_2 \in [\frac{1}{8}, 1)$, then the corresponding Lagrange multipliers (θ_1, θ_2) are both non-negative. Hence from [29, Theorem 4.1] we have that

$$I_1 := \sup_{\tilde{h}\in\tilde{W}} \left[\theta_1(\epsilon)T_1(\tilde{h}) + \theta_2(\epsilon)T_2(\tilde{h}) - I(\tilde{h})\right] = \sup_{0\le u\le 1} \left[\theta_1(\epsilon)u + \theta_2(\epsilon)u^2 - I(u)\right], \tag{6.24}$$

and, consequently,

$$I_1 = \sup_{0<u<1} \left[\theta_1(\epsilon)u + \theta_2(\epsilon)u^3 - I(u)\right] = \theta_1(\epsilon)u^*(\epsilon) + \theta_2(\epsilon)u^*(\epsilon)^3 - I(u^*(\epsilon)). \tag{6.25}$$

The optimiser $u^*(\epsilon)$ corresponding to the perturbed multipliers $\theta_1^*(\epsilon)$ and $\theta_2^*(\epsilon)$ is analytic in ϵ, as shown in [87]. Therefore, a Taylor expansion around $\epsilon = 0$ gives

$$u^*(\epsilon) = T_1 + \delta\epsilon + \tfrac{1}{2}\Delta\epsilon^2 + O(\epsilon^3), \tag{6.26}$$

where $\delta = u^{*\prime}(0)$ and $\Delta = u^{*\prime\prime}(0)$. Hence I_1 can be written as

$$I_1 = \theta_1 T_1 - I(T_1) + (\gamma_1 T_1 + \gamma_2 T_1^3)\epsilon + O(\epsilon^2). \tag{6.27}$$

Moreover,

$$\begin{aligned} I_2 &= \big[\theta_1 + \gamma_1\epsilon + \tfrac{1}{2}\Gamma_1\epsilon^2 + O(\epsilon^3)\big]T_1 + \big[\gamma_2\epsilon + \tfrac{1}{2}\Gamma_2\epsilon^2 + O(\epsilon^3)\big](T_1^3 + 3T_1\epsilon) - \inf_{\tilde{h}\in\tilde{W}_\epsilon^*} I(\tilde{h}) \\ &= \theta_1 T_1 + \gamma_1 T_1\epsilon + \tfrac{1}{2}\Gamma_1 T_1\epsilon^2 + T_1^3\gamma_2\epsilon + \tfrac{1}{2}\Gamma_2 T_1^3\epsilon^2 + 3T_1\gamma_2\epsilon^2 - J^{\downarrow}(\epsilon) + O(\epsilon^3), \end{aligned} \tag{6.28}$$

where

$$J^{\downarrow}(\epsilon) := \inf_{\tilde{h}\in\tilde{W}_\epsilon^*} I(\tilde{h}), \qquad \tilde{W}_\epsilon^* := \{\tilde{h}\in\tilde{W}\colon\, T_1(\tilde{h}) = T_1,\ T_2(\tilde{h}) = T_1^3 + 3T_1\epsilon\}. \tag{6.29}$$

Consequently,

$$s_\infty(T_1^*, T_1^{*3} + 3T_1^*\epsilon) = J^{\downarrow}(\epsilon) - I(T_1) + O(\epsilon^2). \tag{6.30}$$

Denote by $\tilde{h}_\epsilon^{(2)}$ one of the, possibly multiple, balance optimisers of the variational problem $J^{\downarrow}(\epsilon)$. From Proposition 6.3.5 we know that, for ϵ sufficiently small, any graphon in the equivalence class $\tilde{h}_\epsilon^{(2)}$, denoted by $h_\epsilon^{(2)}$, has the form $h_\epsilon^{(2)} = T_1 + \sqrt{\epsilon}g^* + O(\epsilon)$, where the graphon g^* was defined in (6.13). By considering the Taylor expansion of the function I around $\epsilon = 0$, we get

$$\begin{aligned} I(h_\epsilon^{(2)}) &= I(T_1) + I'(T_1)\sqrt{\epsilon}\int_{[0,1]^2} \mathrm{d}x\,\mathrm{d}y\, g^*(x,y) \\ &\quad + \tfrac{1}{2}I''(T_1)\,\epsilon\int_{[0,1]^2} \mathrm{d}x\,\mathrm{d}y\, g^*(x,y)^2 + o(\epsilon) \\ &= I(T_1) + \tfrac{1}{2}I''(T_1)\,\epsilon\int_{[0,1]^2} \mathrm{d}x\,\mathrm{d}y\, g^*(x,y)^2 + o(\epsilon) \\ &= I(T_1) + I''(T_1)\epsilon + o(\epsilon) \\ &= I(T_1) + \frac{1}{2}\frac{1}{T_1(1-T_1)}\epsilon + o(\epsilon). \end{aligned} \tag{6.31}$$

But, from (6.14), a straightforward computation of the entropy of h_ϵ^* shows that

$$J^{\downarrow}(\epsilon) = I(T_1) + \frac{6}{1-2T_1^*}\log\frac{T_1^*}{1-T_1^*}\epsilon + o(\epsilon). \tag{6.32}$$

Hence we obtain that the global optimiser is not a balance optimiser and that

$$s_\infty(T_1^*, T_1^{*3} + 3T_1^*\epsilon) = \frac{6}{1-2T_1^*}\log\frac{T_1^*}{1-T_1^*}\,\epsilon + o(\epsilon). \tag{6.33}$$

Case II $T_1 \in (0, \frac{1}{2})$**:** Consider the term

$$I_1 := \sup_{\tilde{h}\in\tilde{W}} \Big[\theta_1(\epsilon)T_1(\tilde{h}) + \theta_2(\epsilon)T_2(\tilde{h}) - I(\tilde{h})\Big],$$

as above. If Assumption 1 applies, then this case is proved in the same way as Case I. Otherwise, consider the straightforward lower bound

$$\sup_{\tilde{h}\in\tilde{W}}\Big[\theta_1(\epsilon)T_1(\tilde{h})+\theta_2(\epsilon)T_2(\tilde{h})-I(\tilde{h})\Big]\geq\sup_{0\leq u\leq 1}\big[\theta_1(\epsilon)u+\theta_2(\epsilon)u^3-I(u)\big]\,. \tag{6.34}$$

The arguments used in Case I after (6.25) apply, and the result in (6.30) is obtained with an inequality instead of an equality.

§6.4.2 Proof of Theorem 6.3.2

In this section we omit the computations that are similar to those in the proof of Theorem 6.3.1. Let

$$T_1(\epsilon)=T_1,\qquad T_2(\epsilon)=T_1^3-T_1^3\;\epsilon. \tag{6.35}$$

The factor T_1^3 appearing in front of the ϵ is put in for convenience in the computations. The perturbed Lagrange multipliers are

$$\theta_1(\epsilon)=\theta_1+\gamma_1\epsilon+\tfrac{1}{2}\Gamma_1\epsilon^2+O(\epsilon^3),\qquad \theta_2(\epsilon)=\gamma_2\epsilon+\tfrac{1}{2}\Gamma_2\epsilon^2+O(\epsilon^3), \tag{6.36}$$

where

$$\theta_1=I'(T_1),\qquad \gamma_1=\theta_1'(0),\qquad \Gamma_1=\theta_1''(0)\qquad \gamma_2=\theta_2'(0),\qquad \Gamma_2=\theta_2''(0). \tag{6.37}$$

We denote the two terms in the expression for s_∞ in (6.3) by I_1, I_2, i.e., $s_\infty = I_1 - I_2$, and let $s_\infty(\epsilon)$ denote the perturbed relative entropy. The computations for I_1 are similar as before, because the exact form of the constraint does not affect the expansions in (6.26) and (6.27). For I_2, on the other hand, we have

$$\begin{aligned} I_2&=\theta_1T_1+\gamma_1T_1\epsilon+\tfrac{1}{2}\Gamma_1T_1\epsilon^2+T_1^3\gamma_2\epsilon+\tfrac{1}{2}\Gamma_2T_1^3\epsilon^2-T_1^3\gamma_2\epsilon^2-J_1^{\uparrow}(\epsilon)\\ &=\theta_1T_1+\gamma_1T_1\epsilon+T_1^3\gamma_2\epsilon-J_1^{\uparrow}(\epsilon)+O(\epsilon^2),\end{aligned} \tag{6.38}$$

where

$$J_1^{\uparrow}(\epsilon):=\inf_{\tilde{h}\in\tilde{W}_\epsilon^*}I(\tilde{h}),\qquad \tilde{W}_\epsilon^*:=\{\tilde{h}\in\tilde{W}\colon\,T_1(\tilde{h})=T_1,\;T_2(\tilde{h})=T_1^3-T_1^3\;\epsilon\}. \tag{6.39}$$

Consequently,

$$s_\infty(T_1^*,T_1^*-T_1^{*3}\epsilon)=J_1^{\uparrow}(\epsilon)-I(T_1)+O(\epsilon^2). \tag{6.40}$$

Denote by $\tilde{h}_\epsilon^*$ one of the, possibly multiple, optimisers of the variational problem $J_1^{\uparrow}(\epsilon)$. From Proposition 6.3.6 we know that, for $T_1^*\in(0,\frac{1}{2}]$, a balance optimal graphon in the equivalence class $\tilde{h}_\epsilon^*$, denoted by h_ϵ^* for simplicity in the notation, has the form

$$h_\epsilon^*=T_1^*+\epsilon^{1/3}g^*+O(\epsilon^{1/3}) \tag{6.41}$$

with g^* given by

$$g^*(x,y)=\begin{cases}-T_1^*, & (x,y)\in[0,\frac{1}{2}]^2,\\ \;\;T_1^*, & (x,y)\in[0,\frac{1}{2}]\times(\frac{1}{2},1]\cup(\frac{1}{2},1]\times[0,\frac{1}{2}],\\ -T_1^*, & (x,y)\in(\frac{1}{2},1]^2.\end{cases} \tag{6.42}$$

Hence

$$J_1^{\uparrow}(\epsilon) \leq I(T_1) + \frac{1}{2}T_1^{*2}I''(T_1)\epsilon^{2/3} \leq I(T_1) + \frac{1}{4}\frac{T_1^*}{1-T_1^*}\epsilon^{2/3}, \tag{6.43}$$

which gives

$$s_\infty(T_1^*, T_1^* - T_1^{*3}\epsilon) \leq \frac{1}{4}\frac{T_1^*}{1-T_1^*}\epsilon^{2/3} + o(\epsilon^{2/3}). \tag{6.44}$$

§6.4.3 Proof of Theorem 6.3.3

The computations leading to the expression for the relative entropy in the right-hand side of (6.10) are similar as those in Section 6.4.2, and we omit them. Hence we have

$$s_\infty(T_1^*, T_1^* - T_1^{*3}\epsilon) = J_2^{\uparrow}(\epsilon) - I(T_1) + O(\epsilon^2), \tag{6.45}$$

where, for $T_1 \in (\frac{1}{2}, 1)$,

$$J_2^{\uparrow}(\epsilon) := \inf_{\tilde{h}\in\tilde{W}_\epsilon^*} I(\tilde{h}), \qquad \tilde{W}_\epsilon^* := \{\tilde{h} \in \tilde{W} : T_1(\tilde{h}) = T_1,\ T_2(\tilde{h}) = T_1^3 - T_1^3\ \epsilon\}. \tag{6.46}$$

Denote by $\tilde{h}_\epsilon^*$ one of the, possibly multiple, optimisers of the variational problem $J_2^{\uparrow}(\epsilon)$. From Proposition 6.3.7 we know that, for $T_1 \in (\frac{1}{2}, 1)$, a balance optimal graphon, in the equivalence class $\tilde{h}_\epsilon^*$, denoted by h_ϵ^* for simplicity in the notation, has the form

$$h_\epsilon^* = T_1^* + g_\epsilon^* \tag{6.47}$$

with g_ϵ^* given by

$$g_\epsilon^*(x,y) := \begin{cases} \frac{T_1^{*2}}{\bar{T}_1^*}\epsilon^{2/3}, & (x,y) \in [0, 1 - \frac{T_1^*}{|\bar{T}_1^*|}\epsilon^{1/3}]^2 \\ T_1^*\epsilon^{1/3}, & (x,y) \in [0, 1 - \frac{T_1^*}{|\bar{T}_1^*|}\epsilon^{1/3}] \times [1 - \frac{T_1^*}{|\bar{T}_1^*|}\epsilon^{1/3}, 1] \ \text{ or} \\ & (x,y) \in [1 - \frac{T_1^*}{|\bar{T}_1^*|}\epsilon^{1/3}, 1] \times [0, 1 - \frac{T_1^*}{|\bar{T}_1^*|}\epsilon^{1/3}], \\ \bar{T}_1^*, & (x,y) \in [1 - \frac{T_1^*}{|\bar{T}_1^*|}\epsilon^{1/3}, 1]^2. \end{cases} \tag{6.48}$$

The term $\bar{T}_1^* \in (-T_1^*, 0)$ is defined in Theorem 6.3.3. Hence we have

$$s_\infty(T_1^*, T_1^* - T_1^{*3}\epsilon) \leq f(T_1^*, \bar{T}_1^*)\epsilon^{2/3} + o(\epsilon^{2/3}), \tag{6.49}$$

where $\bar{T}_1^* \in (-T_1^*, 0)$ is the unique point where the global minimum of the function $x \to f(T_1^*, x)$ defined by

$$f(T_1^*, x) := T_1^{*2}\frac{I(T_1^* + x) - I(T_1^*) - I'(T_1^*)x}{x^2}, \quad x \in (-T_1, 0). \tag{6.50}$$

We need to show that, for every $T_1^* \in (0,1)$ and for every $x \in (-T_1, 0)$, $f(T_1, x) > 0$ or equivalently that

$$I(T_1^* + x) - I(T_1^*) - I'(T_1^*)x > 0. \tag{6.51}$$

From the mean-value theorem we have that there exists $\xi \in (T_1^* + x, T_1^*)$ such that $I'(T_1^* + x) - I(T_1^*) = I'(\xi)x$. Hence we have that

$$f(T_1^*, x) = (I'(\xi) - I'(T_1^*))x > 0, \tag{6.52}$$

which follows because I' is an increasing function, $x \in (-T_1, 0)$ and $\xi \in (T_1^* + x, T_1^*)$. More detailed arguments are provided in the following section.

§6.5 Proofs of Propositions 6.3.5–6.3.7

In this section we prove Propositions 6.3.5–6.3.7. In Section 6.5.1 we prove Proposition 6.3.5 and in Section 6.5.2 we prove Propositions 6.3.6 - 6.3.7. The proof of Proposition 6.3.7 is similar to the proof of Proposition 6.3.6, only the computations are different. In Section 6.4 the following variational problems were encountered:

(1) For $T_1 \in (0,1)$,

$$J^{\downarrow}(\epsilon) = \inf\big\{I(\tilde{h})\colon\ \tilde{h} \in \tilde{W},\, T_1(\tilde{h}) = T_1,\, T_2(\tilde{h}) = T_1^3 + 3T_1\epsilon\big\}. \tag{6.53}$$

(2) For $T_1 \in (0, \frac{1}{2}]$,

$$J_1^{\uparrow}(\epsilon) = \inf\big\{I(\tilde{h})\colon\ \tilde{h} \in \tilde{W},\, T_1(\tilde{h}) = T_1,\, T_2(\tilde{h}) = T_1^3 - T_1^3\epsilon\big\}. \tag{6.54}$$

(3) For $T_1 \in (\frac{1}{2}, 1)$,

$$J_2^{\uparrow}(\epsilon) = \inf\big\{I(\tilde{h})\colon\ \tilde{h} \in \tilde{W},\, T_1(\tilde{h}) = T_1,\, T_2(\tilde{h}) = T_1^3 - T_1^3\epsilon\big\}. \tag{6.55}$$

In order to prove Propositions 6.3.5–6.3.7, we need to analyse these three variational problems, for ϵ sufficiently small, which is the objective of this section. The variational formula in (6.53) has been rigorously analysed in [66], and hence we study the variational formulas in (6.54) and (6.55), under the assumption that the optimiser lies in the class of balance optimal graphons. We remind the reader that we suppose Assumption 2 to be true. We analyse the variational formulas with the help of a perturbation argument. In particular, we show that the balance optimal perturbations are those given in (6.12), (6.16) and (6.18), respectively. The results in Propositions 6.3.6–6.3.7 follow directly from the following two lemmas.

6.5.1 Lemma. *Let $T_1 \in (0, \frac{1}{2}]$. For $\epsilon > 0$ consider the variational formula for $J_1^{\uparrow}(\epsilon)$ given in* (6.54)*. Then, for ϵ sufficiently small,*

$$J_1^{\uparrow}(\epsilon) \leq I(T_1) + \frac{1}{4}\frac{T_1}{1-T_1}\epsilon^{2/3} + o(\epsilon^{2/3}). \tag{6.56}$$

6.5.2 Lemma. *Let $T_1 \in (\frac{1}{2}, 1)$. For $\epsilon > 0$ consider the variational formula for $J_2^{\uparrow}(\epsilon)$ given in* (6.55)*. Then, for ϵ sufficiently small,*

$$J_2^{\uparrow}(\epsilon) \leq I(T_1) + f(T_1, \bar{T}_1^*)\epsilon^{2/3} + o(\epsilon^{2/3}), \tag{6.57}$$

where $f(T_1, x)$, $x \in (-T_1, 0)$, and $\bar{T}_1^*$ were defined in Theorem 6.3.3.

6.5.3 Remark. As argued in Remark 6.3.4, we believe, and there is numerical evidence in [63], that the results in (6.56) and (6.57) hold with equality.

In what follows we use the notation $f(\epsilon) \asymp g(\epsilon)$, for two functions f, g, when $\frac{f(\epsilon)}{g(\epsilon)}$ converges to a *positive* constant, as $\epsilon \downarrow 0$.

§6.5.1 Proof of Proposition 6.3.5

In this section we prove Proposition 6.3.5 given that Assumption 2 holds. In order to find the optimal perturbation when the ER-line is approached from above, we need to solve $J^{\downarrow}(\epsilon)$ in (6.53). The following construction shows intuitively why balance optimal perturbations have the form given in (6.12). Consider an inhomogeneous ER-random graph on n vertices. We split the vertices of the graph into two parts of equal size, i.e., of size $n/2$. In one part we connect two vertices with probability $T_1+2\sqrt{\epsilon}$, in the other part we connect two vertices with probability $T_1-2\sqrt{\epsilon}$, and we connect vertices lying in different parts with probability T_1. This graph has expected edge density equal to

$$\frac{1}{\binom{n}{2}}\left(T_1\left(\frac{n}{2}\right)^2+(T_1+2\sqrt{\epsilon})\binom{\frac{n}{2}}{2}+(T_1-2\sqrt{\epsilon})\binom{\frac{n}{2}}{2}\right)=T_1. \tag{6.58}$$

Similarly, the expexted triangle density is equal to

$$\begin{aligned}&\frac{1}{\binom{n}{3}}\left(\binom{\frac{n}{2}}{3}(T_1+2\sqrt{\epsilon})^3+\frac{n}{2}\binom{\frac{n}{2}}{2}2T_1^2T_1+\binom{\frac{n}{2}}{3}(T_1-2\sqrt{\epsilon})^3\right)\\&=T_1^3+3T_1\frac{n-4}{n-1}\epsilon\\&\sim T_1^3+3T_1\epsilon,\end{aligned}$$

for n large. Below when we speak of optimal perturbation we mean balance optimal. In the proof below we will see that the optimal perturbation is indeed given by the graphon counterpart of the inhomogeneous ER-random graph described above. We now proceed to the technical details of the proof.

With a slight abuse of notation we write $I(\cdot)$ for both cases of a graphon and a real number. We consider the variational formula $J^{\downarrow}(\epsilon)$, with $\epsilon>0$ given in (6.53). We denote by $\tilde{h}_\epsilon^{*\downarrow}$ one of the, possibly multiple, optimisers of $J^{\downarrow}(\epsilon)$. For simplicity in the notation, in what follows we work with a representative element, denoted by $h_\epsilon^{*\downarrow}$, of the equivalence class $\tilde{h}_\epsilon^{*\downarrow}$. We write the optimiser $h_\epsilon^{*\downarrow}$ in the form $h_\epsilon^{*\downarrow}=T_1+\Delta H_\epsilon$ for some bounded symmetric function ΔH_ϵ defined on the unit square $[0,1]^2$ and taking values in $\mathbb{R}$. This term will be called the *perturbation term*. The optimiser $h_\epsilon^{*\downarrow}$ has to satisfy the conditions on the edge and triangle densities, i.e.,

$$T_1(h_\epsilon^{*\downarrow})=T_1,\qquad T_2(h_\epsilon^{*\downarrow})=T_1^3+3T_1\epsilon. \tag{6.59}$$

Hence the perturbation term ΔH_ϵ needs to satisfy the constraints

$$(G_1)\colon\quad \int_{[0,1]^2}\mathrm{d}x\,\mathrm{d}y\ \Delta H_\epsilon(x,y)=0 \tag{6.60}$$

and

$$\begin{aligned}(G_2)\colon\ 3T_1&\int_{[0,1]^3}\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z\ \ \Delta H_\epsilon(x,y)\Delta H_\epsilon(y,z)\\&+\int_{[0,1]^3}\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z\ \ \Delta H_\epsilon(x,y)\Delta H_\epsilon(y,z)\Delta H_\epsilon(z,x)=3T_1\epsilon.\end{aligned} \tag{6.61}$$

In what follows we prove the result stated in Proposition 6.3.5, i.e., the optimal perturbation is a three-step function and is of order $\sqrt{\epsilon}$.

In Assumption 2 it is stated that it suffices to restrict to graphons that can be written in the form $T_1+\Delta H_\epsilon^{(2)}$, where $\Delta H_\epsilon^{(2)}$ is a bounded symmetric function defined on $[0,1]^2$, taking three non-zero values. In what follows, for simplicity in the computations and without loss of generality, we suppose that the optimal graphon has the form

$$\Delta H_\epsilon^{(2)} = g_{11}1_{I\times I} + g_{12}1_{(I\times J)\cup(J\times I)} + g_{22}1_{J\times J}. \tag{6.62}$$

Then (G_1) above becomes

$$\lambda(I)^2 g_{11} + 2\lambda(I)(1-\lambda(I))g_{12} + (1-\lambda(I))^2 g_{22} = 0 \tag{6.63}$$

and the two integrals in (G_2) become

$$\begin{aligned}\int_{[0,1]^3} \mathrm{d}x\ \mathrm{d}y\ \mathrm{d}z\ \Delta H_\epsilon(x,y)\Delta H_\epsilon(y,z) = {}& \lambda(I)^3 g_{11}^2 + 2\lambda(I)^2(1-\lambda(I))g_{11}g_{12} \\ & + 2\lambda(I)(1-\lambda(I))^2 g_{12}g_{22} + \lambda(I)(1-\lambda(I))g_{12}^2 \\ & + (1-\lambda(I))^2 g_{22}^2,\end{aligned} \tag{6.64}$$

and

$$\begin{aligned}\int_{[0,1]^3} \mathrm{d}x\ \mathrm{d}y\ \mathrm{d}z\ \Delta H_\epsilon(x,y)\Delta H_\epsilon(y,z) = {}& \lambda(I)^3 g_{11}^2 + 2\lambda(I)^2(1-\lambda(I))g_{11}g_{12} \\ & + 2\lambda(I)(1-\lambda(I))^2 g_{12}g_{22} + \lambda(I)(1-\lambda(I))g_{12}^2 \\ & + (1-\lambda(I))^2 g_{22}^2,\end{aligned} \tag{6.65}$$

and a similar expression can be computed for the second integral in (G_2). We now give the formal definition of a balance optimal graphon:

6.5.4 Definition. For $T_1 \in (0,1)$, a graphon $T_1 + \tilde{h}_\epsilon$, $\epsilon > 0$, is called balanced if it has the structure given in (6.62) and the terms $\lambda(I)^2 g_{11}, \lambda(I)(1-\lambda(I))g_{12}$ and $(1-\lambda(I))^2 g_{22}$ are all of the same order when ϵ is sufficiently small.

6.5.5 Definition. For $\epsilon > 0$ a graphon $\tilde{h}_\epsilon$ is called balance optimal if it solves the following optimisation problem:

$$J_{bal}(\epsilon) := \inf\{I(\tilde{h}), \tilde{h} \in \tilde{W}, \tilde{h} \text{ is balanced, } T_1(\tilde{h}) = T_1, T_2(\tilde{h}) = T_1^3 + 3T_1\epsilon\}. \tag{6.66}$$

It is straightforward to observe that, for $\epsilon > 0$,

$$J_{bal}(\epsilon) \geq J(\epsilon). \tag{6.67}$$

In what follows we essentially determine $J_{bal}(\epsilon)$ for ϵ sufficiently small. We distinguish two cases, first $g_{12} = 0$ and then $g_{12} \neq 0$.

Case $g_{12} = 0$: The values of g_+ and g_- are such so that $T_1 + \Delta H_\epsilon^{(2)}$ satisfies the conditions in (6.60) and (6.61). We proceed with the condition in (6.61). A standard computation yields

$$\int_{[0,1]^3} \mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z\ \Delta H_\epsilon^{(2)}(x,y)\Delta H_\epsilon^{(2)}(y,z) = \lambda(I)^3\, g_+^2 + \lambda(J)^3\, g_-^2 \tag{6.68}$$

and

$$\int_{[0,1]^3} \mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z\ \Delta H_\epsilon^{(2)}(x,y)\Delta H_\epsilon^{(2)}(y,z)\Delta H_\epsilon^{(2)}(z,x) = \lambda(I)^3\, g_+^3 + \lambda(J)^3\, g_-^3. \tag{6.69}$$

From (6.60) we obtain the first order condition

$$\lambda(I)^2 g_+ + \lambda(J)^2 g_- = 0. \tag{6.70}$$

Using the condition in (6.70), we get that (6.61) equals

$$g_-^2\, 3T_1 \frac{\lambda(J)^3}{\lambda(I)}(\lambda(J)+\lambda(I)) - g_-^3 \frac{\lambda(J)^3}{\lambda(I)^3}(\lambda(I)^3 - \lambda(J)^3) = 3T_1\epsilon + o(\epsilon). \tag{6.71}$$

There are multiple ways in which the condition in (6.71) can be met. We show that the lowest possible value of I is attained when $g_+ \asymp \sqrt{\epsilon}, g_- \asymp -\sqrt{\epsilon}$ and $\lambda(I), \lambda(J)$ are constant. To that end we distinguish the following cases:

(I)

$$g_-^2\, 3T_1 \frac{\lambda(J)^3}{\lambda(I)}(\lambda(J)+\lambda(I)) \asymp \epsilon, \qquad g_-^3 \frac{\lambda(J)^3}{\lambda(I)^3}(\lambda(I)^3 - \lambda(J)^3) = o(\epsilon), \tag{6.72}$$

which splits into three sub-cases:

(Ia)

$$g_+ \asymp \epsilon^{1/2}, \quad g_- \asymp -\epsilon^{1/2}, \quad \frac{\lambda(J)}{\lambda(I)} \asymp 1. \tag{6.73}$$

(Ib)

$$g_+ \asymp \epsilon^{1/2+\delta/3}, \quad g_- \asymp -\epsilon^{1/2-\delta}, \quad \frac{\lambda(J)^3}{\lambda(I)} \asymp \epsilon^{2\delta}, \quad \delta \in (0, \tfrac{1}{2}). \tag{6.74}$$

(Ic)

$$g_+ \asymp \epsilon^{1/2-3\delta}, \quad g_- \asymp -\epsilon^{1/2+\delta}, \quad \frac{\lambda(J)^3}{\lambda(I)} \asymp \epsilon^{-2\delta}, \quad \delta \in (0, \tfrac{1}{6}). \tag{6.75}$$

(1d)

$$g_+ \asymp \epsilon^{2/3}, \quad g_- = \bar{g} \in (-T_1, 0), \quad \lambda(J) \asymp \epsilon^{1/3}. \tag{6.76}$$

(II)

$$g_-^2\, 3T_1 \frac{\lambda(J)^3}{\lambda(I)}(\lambda(J)+\lambda(I)) \asymp \epsilon^{1+\delta}, \quad -g_- \frac{1}{\lambda(I)^2} \asymp \epsilon^{-\delta}, \quad \delta > 0. \tag{6.77}$$

A simple calculation shows that, in all the cases above, $\lambda(I) + \lambda(J) \asymp 1$ and $\lambda(I)^3 - \lambda(J)^3 \asymp 1$, and hence we can omit these two factors from the analysis below. In what follows we exclude cases (Ib), (Ic) and (II) one by one by comparing them to graphons of the type given in case (Ia).

Case (Ib): We show that, for $\epsilon > 0$ sufficiently small, graphons having the structure indicated in (Ia) yield smaller values of the function I than graphons with the structure in (Ib). We consider two graphons, denoted by $T_1 + g^*$ and $T_1 + \hat{g}^*$, where g^* is as in Case (Ia) and $\hat{g}^*$ is as in Case (Ib). Before giving the technical details of the proof, we present a heuristic argument why $I(T_1 + g^*) < I(T_1 + \hat{g}^*)$. In what follows we will denote by $B(p)$ a Bernoulli random variable with parameter p. The function $-I(x)$, $x \in [0,1]$, defined in (5.19) represents the entropy of a $B(x)$ random variable with parameter x. On the graphon space the function $-I(h)$, $h \in W$, defined in (5.20), can be seen as the expectation of the entropy of a Bernoulli random variable with a random parameter (the expectation is with respect to the random parameter), i.e., $B(h(X,Y))$ with (X,Y) a uniformly distributed random variable on $[0,1]^2$. For $h \in W$ we have

$$-I(h) = \int_{[0,1]^2} \mathrm{d}x\, \mathrm{d}y\, [-I(h(x,y))] = \mathbb{E}[-I(h(X,Y))]. \tag{6.78}$$

Hence we have the following equivalence

$$I(T_1 + g^*) < I(T_1 + \hat{g}^*) \Leftrightarrow \mathbb{E}[-I(T_1 + g^*(X,Y))] > \mathbb{E}[-I(T_1 + \hat{g}^*(X,Y))], \tag{6.79}$$

where (X,Y) is a uniformly distributed random vector on $[0,1]^2$. Instead of working with entropy, it is intuitively simpler to work with the relative entropy with respect to the random variable $B(\frac{1}{2})$. The relative entropy is defined by

$$I_{\frac{1}{2}}(x) := x \log \frac{x}{\frac{1}{2}} + (1-x) \log \frac{1-x}{\frac{1}{2}}, \qquad x \in [0,1]. \tag{6.80}$$

Note that

$$\begin{aligned} &\mathbb{E}[-I(T_1 + g^*(X,Y))] > \mathbb{E}[-I(T_1 + \hat{g}^*(X,Y))] \Leftrightarrow \\ &\mathbb{E}[I_{\frac{1}{2}}(T_1 + g^*(X,Y))] < \mathbb{E}[I_{\frac{1}{2}}(T_1 + \hat{g}^*(X,Y))]. \end{aligned} \tag{6.81}$$

We first give an intuitive argument and afterwards prove that

$$\mathbb{E}[I_{\frac{1}{2}}(T_1 + g^*(X,Y))] < \mathbb{E}[I_{\frac{1}{2}}(T_1 + \hat{g}^*(X,Y))]. \tag{6.82}$$

We distinguish between the cases $T_1 \in (0, \frac{1}{2}]$ and $T_1 \in (\frac{1}{2}, 1)$. The case $T_1 \in (0, \frac{1}{2}]$ follows by using similar arguments as in case $T_1 \in (\frac{1}{2}, 1)$. We treat in detail only the case $T_1 \in (\frac{1}{2}, 1)$.

The relative entropy of a random variable with respect to $B(\frac{1}{2})$ is zero if and only if that random variable is equal to $B(\frac{1}{2})$. So, in order to compare the relative entropies in (6.82), we need to see how close the Bernoulli random variables with random parameters $T_1 + g^*(X,Y)$ and $T_1 + \hat{g}^*(X,Y)$ are to $B(\frac{1}{2})$. We are considering the case $T_1 > \frac{1}{2}$. Hence the random variables $B(T_1 + g^*(X,Y))$ and $B(T_1 + \hat{g}^*(X,Y))$ will be close to $B(\frac{1}{2})$ when the random parameters $T_1 + g^*(X,Y)$ and $T_1 + \hat{g}^*(X,Y)$ are close to $\frac{1}{2}$. This is the case when $g^*(X,Y)$ and $\hat{g}^*(X,Y)$ are negative. These events occur with probabilities

$$\mathbb{P}(T_1 + g^*(X,Y) < T_1) = \mathbb{P}(g^*(X,Y) < 0) = \mathbb{P}(g^*(X,Y) = g_-) = \lambda(J)^2 \asymp 1, \tag{6.83}$$

because of the properties of the graphon in Case (Ia). Similarly, we have that

$$\mathbb{P}(T_1 + \hat{g}^*(X,Y) < T_1) = \mathbb{P}(\hat{g}^*(X,Y) < 0) = \mathbb{P}(\hat{g}^*(X,Y) = g_-) = \lambda(\hat{J})^2 \asymp \epsilon^{4\delta/3}, \tag{6.84}$$

for some $\delta \in (0, \frac{1}{2}]$, because of the properties of the graphon in Case (Ib). Hence we see that the random variable $B(T_1 + g^*(X,Y))$ is closer to the random variable $B(\frac{1}{2})$ with much higher probability than the random variable $B(T_1 + \hat{g}^*(X,Y))$. We can see this by computing the corresponding expectations,

$$\mathbb{E}(g^*(X,Y) \mid g^*(X,Y) = g_-)\, \mathbb{P}(g^*(X,Y) = g_-) = g_-\, \mathbb{P}(g^*(X,Y) = g_-) \asymp \epsilon^{1/2}, \tag{6.85}$$

while

$$\begin{aligned}\mathbb{E}(\hat{g}^*(X,Y) \mid \hat{g}^*(X,Y) = \hat{g}_-)\, \mathbb{P}(\hat{g}^*(X,Y) = \hat{g}_-) &= \hat{g}_-\, \mathbb{P}(\hat{g}^*(X,Y) = \hat{g}_-) \\ &\asymp \epsilon^{1/2-\delta}\epsilon^{4\delta/3} = \epsilon^{1/2+\delta/3}.\end{aligned}$$

In what follows we complete this argument by adding the technical details. We work out the expressions in the left-hand and right-hand sides of (6.82). The expression in the right-hand side of (6.82) can be written as

$$\mathbb{E}[I_{\frac{1}{2}}(T_1 + g^*(X,Y))] = L I_{\frac{1}{2}}(T_1 + g_+) + K I_{\frac{1}{2}}(T_1 + g_-) + (1 - L - K) I_{\frac{1}{2}}(T_1), \tag{6.86}$$

for some constants $L := \mathbb{P}(g^*(X,Y) = g_+)$ and $K = \mathbb{P}(g^*(X,Y) = g_-)$ independent of ϵ. Similarly,

$$\mathbb{E}[I_{\frac{1}{2}}(T_1 + \hat{g}^*(X,Y))] = \lambda(\hat{I})^2 I_{\frac{1}{2}}(T_1 + \hat{g}_+) + \epsilon^{4\delta/3} I_{\frac{1}{2}}(T_1 + \hat{g}_-) + (1 - \lambda(\hat{I})^2 - \epsilon^{4\delta/3}) I_{\frac{1}{2}}(T_1), \tag{6.87}$$

where $\lambda(\hat{I})^2 = \mathbb{P}(\hat{g}^*(X,Y) = \hat{g}_+) \asymp 1$ and $\mathbb{P}(\hat{g}^*(X,Y) = \hat{g}_-) \asymp \epsilon^{4\delta/3}$. Moreover, we recall that from the properties of the graphons in Case (Ia) and Case (Ib) we get

$$g_+ \asymp \sqrt{\epsilon}, \quad g_- \asymp -\sqrt{\epsilon}, \quad \hat{g}_+ \asymp \epsilon^{1/2+\delta/3}, \quad \hat{g}_- \asymp \epsilon^{1/2-\delta}, \quad \delta \in (0, \tfrac{1}{2}]. \tag{6.88}$$

Hence, for $T_1 \in (\frac{1}{2}, 1]$ and ϵ sufficiently small, because of (6.88), we obtain the following inequalities:

$$I_{\frac{1}{2}}(T_1 + g_+) > I_{\frac{1}{2}}(T_1 + \hat{g}_+) > I_{\frac{1}{2}}(T_1 + g_-) > I_{\frac{1}{2}}(T_1 + \hat{g}_-). \tag{6.89}$$

Using a Taylor expansion of the function I around T_1 and the first order conditions

$$L g_+ + K g_- = 0 \quad \text{and} \quad \lambda(\hat{I})^2 \hat{g}_+ + \lambda(\hat{J})^2 \hat{g}_- = 0, \tag{6.90}$$

we observe that (6.86) and (6.87) are read

$$\mathbb{E}[I_{\frac{1}{2}}(T_1 + g^*(X,Y))] = I_{\frac{1}{2}}(T_1) + \tfrac{1}{2} I''_{\frac{1}{2}}(T_1)(L g_+^2 + K g_-^2) + o\left(g_+^2 + g_-^2\right) \tag{6.91}$$

and

$$\begin{aligned}\mathbb{E}[I_{\frac{1}{2}}(T_1 + \hat{g}^*(X,Y))] = I_{\frac{1}{2}}(T_1) &+ \tfrac{1}{2} I''_{\frac{1}{2}}(T_1)(\lambda(\hat{I}^2)\hat{g}_+^2 + \lambda(\hat{J})^2 \hat{g}_-^2) \\ &+ o\left(\lambda(\hat{I}^2)\hat{g}_+^2 + \lambda(\hat{J})^2 \hat{g}_-^2\right).\end{aligned} \tag{6.92}$$

Using (6.88), we observe that $Lg_+^2 + Kg_-^2 \asymp \epsilon$ and

$$\lambda(\hat{I}^2)\hat{g}_+^2 + \lambda(\hat{J})^2\hat{g}_-^2 \asymp \epsilon^{1+2\delta/3} + \epsilon^{4/3\delta}\epsilon^{1-2\delta} \asymp \epsilon^{1-2\delta/3}. \tag{6.93}$$

Hence, for ϵ sufficiently small,

$$\mathbb{E}\left[I_{\frac{1}{2}}(T_1 + g^*(X,Y))\right] < \mathbb{E}\left[I_{\frac{1}{2}}(T_1 + \hat{g}^*(X,Y))\right], \tag{6.94}$$

which proves (6.82).

Similar arguments can be used for the case $T_1 \in (0, \frac{1}{2})$ to show that graphons, as in Case (Ic), yield larger values of $I(\cdot)$ for ϵ sufficiently small. We omit the details.

Case (1d): In this case we have that the optimal graphon is constant on a subset of the unit square with a size tending to zero as $\epsilon \downarrow 0$. Such a graphon yields

$$\begin{aligned} I(T_1 + g^*) &= \lambda(I)^2 I(T_1 + g_+) + 2(1-\lambda(I))(1-\lambda(J))I(T_1) + \lambda(J)^2 I(T_1 + g_-) \\ &= \lambda(I)^2(I(T_1) + I'(T_1)g_+ + o(\epsilon^{2/3})) + 2(1-\lambda(I))(1-\lambda(J))I(T_1) \\ &\quad + \lambda(J)^2 I(T_1 + g_-) \\ &= I(T_1) - \lambda(J)^2 I(T_1) - \lambda(J)^2 \bar{g} I'(T_1) + \lambda(J)^2 I(T_1 + \bar{g}) \\ &= I(T_1) + \epsilon^{2/3}\left(I(T_1 + \bar{g}) - \bar{g}I'(T_1) - I(T_1)\right) + o(\epsilon^{2/3}). \end{aligned} \tag{6.95}$$

The second equality follows by considering a Taylor expansion around $\epsilon = 0$ in the terms that go to zero as $\epsilon \downarrow 0$, i.e, g_+. In the third equality we use (6.70). What remains is to show that

$$I(T_1 + \bar{g}) - I(T_1) - \bar{g}I'(T_1) > 0 \tag{6.96}$$

for $\bar{g} \in (-T_1, 0)$. From the mean-value theorem we have that $I(T_1+\bar{g}) - I(T_1) = I'(\xi)\bar{g}$ for some $\xi \in (T_1+\bar{g}, T_1)$. Since $\bar{g} < 0$ and I is a convex function, i.e., I' is an increasing function, we have that $I'(\xi) < I'(T_1)$. This proves the claim above. From (6.95) we observe that graphons having the form as in Case (1d) yield larger values of I, for ϵ sufficiently small, than graphons as in Case (1a).

Case (II): This case is simpler to exclude than the ones above. Indeed, suppose that (6.77) holds. Then either $\lambda(I)$ should become small or $-g_-$ should become large. But $g_- \asymp -\epsilon^{-\delta}$ is not possible because g_- should stay bounded in $(-T_1, 0)$ as $\epsilon \downarrow 0$. Hence the only possibility is $\lambda(I) \asymp \epsilon^{\eta}$ and $g_- \asymp -\epsilon^{\zeta}$ for some η, ζ such that $\zeta - 2\eta = -\delta$, because of the second condition in (6.77). From the first condition in (6.77) we have that $2\zeta - \eta = 1 + \delta$. Solving these two equations we obtain that $\eta = \frac{1}{3} + \delta$ and $\zeta = \frac{2}{3} + \delta$. From (6.70) we then get that $g_+ \asymp \epsilon^{-\delta}$, which is not possible because g_+ should stay bounded in $(0, 1 - T_1)$ as $\epsilon \downarrow 0$.

At this point we summarise our findings. We considered the variational formula for $J^{\downarrow}(\epsilon)$ given in (6.53) and we assumed that we can restrict ourselves to piece-wise constant graphons (see Assumption 2) subject to the constraints in (6.60) and (6.61).

Afterwards, without loss of generality, we restricted ourselves to an even smaller class of graphons, those of the form

$$g = g_+ 1_{I\times I} + g_- 1_{J\times J} \tag{6.97}$$

for some $g_+ > 0$, $g_- < 0$ and $I, J \subset [0,1]$ with $\lambda(I)^2 + \lambda(J)^2 \leq 1$. At the end of this section we elaborate on the case $g_{12} \neq 0$. More specifically, we have shown that the optimal perturbation satisfies $g_+ \asymp \epsilon^{1/2}$, $g_- \asymp -\epsilon^{1/2}$ and $\lambda(I) \asymp 1$, $\lambda(J) \asymp 1$. Hence the solution to $J^{\downarrow}(\epsilon)$ has the form $T_1 + g^*\sqrt{\epsilon} + o(\epsilon)$, where $g^* = g_+ 1_{L\times L} + g_- 1_{K\times K}$ for some $g_+ > 0, g_- < 0, L, K \in (0,1)$ independent of ϵ, is a symmetric function defined on $[0,1]^2$. From the constraints (6.60) and (6.61) we have that $g_+ L^2 = -g_- K^2$ and $L^3 g_+^2 + K^3 g_-^2 = 1$. A simple calculation shows that

$$\begin{aligned} I(T_1 + g\sqrt{\epsilon}) &= I(T_1) + I'(T_1)(L^2 g_+ + K^2 g_-)\sqrt{\epsilon} + \tfrac{1}{2} I''(T_1)(L^2 g_+^2 + K^2 g_-^2)\epsilon + o(\epsilon) \\ &= I(T_1) + \tfrac{1}{2} I''(T_1)(L^2 g_+^2 + K^2 g_-^2)\epsilon + o(\epsilon). \end{aligned}$$

Hence, in order to find the optimal graphon we need to solve the following optimisation problem:

$$\begin{aligned} &\min \left(L^2 g_+^2 + K^2 g_-^2\right) \\ &\text{such that } L + K \leq 1, \quad g_+ L^2 + g_- K^2 = 0, \quad L^3 g_+^2 + K^3 g_-^2 = 1. \end{aligned} \tag{6.98}$$

This is equivalent to

$$\begin{aligned} &\min \left(\frac{1}{K} + \frac{1}{L} - \frac{2}{K+L}\right) \\ &\text{such that } L + K \leq 1. \end{aligned} \tag{6.99}$$

From a standard computation we find that the optimal K, L should satisfy $K + L = 1$. Hence we need to minimize $\frac{1-2L+L^2}{L(1-L)}$. This function is convex in $L \in (0,1)$ and attains a unique minimum at the point $L = \frac{1}{2}$. Having computed L, K we find $g_+ = -g_- = 2$, and so the optimal solution to $J^{\downarrow}(\epsilon)$, for ϵ sufficiently small, is the graphon

$$h_\epsilon^{*\downarrow}(x,y) = \begin{cases} T_1 + 2\sqrt{\epsilon}, & \text{if } (x,y) \in [0,\tfrac{1}{2}]^2, \\ T_1, & \text{if } (x,y) \in [0,\tfrac{1}{2}] \times (\tfrac{1}{2},1] \text{ or } (\tfrac{1}{2},1] \times [0,\tfrac{1}{2}], \\ T_1 - 2\sqrt{\epsilon}, & \text{if } (x,y) \in (\tfrac{1}{2},1]^2. \end{cases} \tag{6.100}$$

A standard computation shows that $T_1(h_\epsilon^{*\downarrow}) = T_1$ and $T_2(h_\epsilon^{*\downarrow}) = T_1^3 + 3T_1\epsilon$.

Case $g_{12} \neq 0$**:** By following similar arguments as for the case $g_{12} = 0$, we can show that the optimal values of g_{11}, g_{12}, g_{22}, K and L can be retrieved by solving the following optimisation problem:

$$\begin{aligned} &\min \left(L^2 g_{11}^2 + K^2 g_{22}^2 + 2LK g_{12}^2\right) \\ &\text{such that} \\ &\qquad L + K = 1, \\ &\qquad L^2 g_{11} + K^2 g_{22} + 2LK g_{12} = 0, \\ &\qquad L^3 g_{11}^2 + K^3 g_{22}^2 + 2L^2 K g_{12} g_{11} + 2LK^2 g_{12} g_{22} + LK g_{12}^2 = 1. \end{aligned} \tag{6.101}$$

Suppose first that $L = K = \frac{1}{2}$. Then we have the following optimisation problem:

$$\begin{aligned}&\min \tfrac{1}{4}\left(g_{11}^2 + g_{22}^2 + 2g_{12}^2\right)\\&\text{such that}\\&\qquad g_{11} + g_{22} + 2g_{12} = 0,\\&\qquad g_{11}^2 + g_{22}^2 + 2g_{12}g_{11} + 2g_{12}g_{22} + 2g_{12}^2 = 8.\end{aligned}$$

Introducing Lagrange multipliers, we obtain the solution $g_{12} = 0$ and $g_{11} = -g_{22} = 2$. For arbitrary L, K, substituting

$$g_{12} = -\frac{1}{2}\left(\frac{L}{1-L}g_{11} + \frac{1-L}{L}g_{22}\right)$$

into (6.101) and differentiating the Lagrangian with respect to g_{12}, we obtain $g_{12} = 0$. We observe at this point that this argument holds only for the case where g_{11}, g_{12} and g_{22} go to zero as $\epsilon \downarrow 0$. This is not the case for the actual optimal graphon in (6.14).

Case $g_{12} \neq 0$ and $g_{22} = 0$: From (6.96) we observe that $g_{22} = 0$ yields an equality. Hence in this case the microcanonical entropy will of order ϵ instead of $\epsilon^{2/3}$. From the first-order constraint in (6.60) we obtain

$$g_{12} = -\frac{1}{2}\frac{\lambda}{(1-\lambda)}g_{11}, \tag{6.102}$$

where $\lambda := \lambda(I)$. Then the second order constraint reads

$$g_{11}^2\frac{1}{4}\frac{\lambda^2}{(1-\lambda)^2}\lambda(1-\lambda) = \epsilon. \tag{6.103}$$

Following similar arguments as before, we can show that the case $g_{11} \asymp \epsilon^{\delta}$, $\lambda \asymp \epsilon^{1/3-\delta/3}$ $g_{12} \asymp -\epsilon^{2/3+\delta/2}$ is not optimal. The case where g_{11} or g_{12} are constant, independently of ϵ, is also not optimal, since if one of them is constant then the entropy cost will be $\epsilon^{2/3}$ instead of ϵ. A standard computation yields

$$I(T_1 + g^*) = I(T_1) + \frac{1}{2}I''(T_1)\left(2 + 4\frac{1-\lambda}{\lambda}\right)\epsilon + o(\epsilon), \tag{6.104}$$

while for the graphon defined in (6.100) we have

$$I(h_\epsilon^{*\downarrow}) = I(T_1) + I''(T_1)\epsilon + o(\epsilon). \tag{6.105}$$

Hence we see that $I(T_1 + g^*) > I(T_1 + h_\epsilon^{*\downarrow})$ if and only if $1 - \lambda$ is constant and independent of ϵ. If $1 - \lambda \asymp \epsilon^{\delta}$, then further analysis is needed in order to establish the optimal graphon. In any case, the graphon $h_\epsilon^{*\downarrow}$ is balance optimal, as desired.

§6.5.2 Proof of Lemma 6.5.1 and Lemma 6.5.2

In this section we provide the technical details leading to the optimal perturbation of the variational formula in (6.54). We denote one of the, possibly multiple, optimizers

of (6.54) by $\tilde{h}_\epsilon^{*\uparrow}$. In the proof, in order to keep the notation light, we denote a representative element of this class by $h_\epsilon^{*\uparrow}$. We start by writing the optimizer in the form $h_\epsilon^{*\uparrow} = T_1 + \Delta H_\epsilon$ for some perturbation term ΔH_ϵ. The perturbation term has to be a bounded symmetric function defined on the unit square $[0,1]^2$ taking values in $\mathbb{R}$. The optimizer $h_\epsilon^{*\uparrow}$ has to satisfy the constraints

$$T_1(h_\epsilon^{*\uparrow}) = T_1, \qquad T_2(h_\epsilon^{*\uparrow}) = T_1^3 - T_1^3\epsilon, \tag{6.106}$$

and so the perturbation ΔH_ϵ needs to satisfy the two constraints

$$(K_1): \int_{[0,1]^2} \mathrm{d}x\,\mathrm{d}y\ \Delta H_\epsilon(x,y) = 0 \tag{6.107}$$

and

$$\begin{aligned}(K_2):\ 3T_1 &\int_{[0,1]^3} \mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z\ \ \Delta H_\epsilon(x,y)\Delta H_\epsilon(y,z) \\ &+ \int_{[0,1]^3} \mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z\ \ \Delta H_\epsilon(x,y)\Delta H_\epsilon(y,z)\Delta H_\epsilon(z,x) = -T_1^3\epsilon.\end{aligned} \tag{6.108}$$

Again, from Assumption 2, we restrict to graphons having the form $T_1 + \Delta H_\epsilon$ where

$$\Delta H_\epsilon = g_{11}1_{I\times I} + g_{12}1_{(I\times J)\cup(J\times I)} + g_{22}1_{J\times J}, \tag{6.109}$$

$g_{11}, g_{12}, g_{22} \in (-T_1, 1-T_1)$ and $I \subset [0,1]$, $J = I^c$. From (6.107) we get

$$\lambda(I)^2 g_{11} + 2\lambda(I)(1-\lambda(I))g_{12} + (1-\lambda(I))^2 g_{22} = 0, \tag{6.110}$$

which yields

$$g_{12} = -\frac{1}{2}\left(\frac{\lambda(I)}{1-\lambda(I)}g_{11} + \frac{1-\lambda(I)}{\lambda(I)}g_{22}\right). \tag{6.111}$$

A standard computation shows that the second order integral in (6.108) is equal to

$$\lambda(I)^3 g_{11}^2 + (1-\lambda(I))^3 g_{22}^2 + 2\lambda(I)(1-\lambda(I))g_{12}(\lambda(I)g_{11} + (1-\lambda(I))g_{22} + \frac{1}{2}g_{12}. \tag{6.112}$$

By (6.111) this is equal to

$$\frac{1}{4}\lambda(I)(1-\lambda(I))\left(\frac{\lambda(I)}{(1-\lambda)}g_{11} - \frac{1-\lambda(I)}{\lambda(I)}g_{22}\right)^2. \tag{6.113}$$

From (6.108) we observe that, for ϵ sufficiently small, the first integral will dominate the second integral when g_{11}, g_{12} and g_{22} depend on ϵ. Hence, in order to obtain the condition in (6.108), it must be that

$$\int_{[0,1]^3} \mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z\ \Delta H_\epsilon(x,y)\Delta H_\epsilon(y,z) = 0. \tag{6.114}$$

Then (6.113) yields

$$g_{11} = \frac{(1-\lambda(I))^2}{\lambda(I)^2}g_{22} \tag{6.115}$$

and from (6.111) also

$$g_{12} = -\frac{1-\lambda(I)}{\lambda(I)} g_{22}. \tag{6.116}$$

The third order integral in (6.108) then yields

$$g_{22}\frac{1-\lambda(I)}{\lambda(I)} = -T_1\epsilon^{1/3}. \tag{6.117}$$

We distinguish three cases,

(1)

$$g_{11} \asymp -\epsilon^{1/3}, \quad g_{12} \asymp \epsilon^{1/3} \quad g_{22} \asymp -\epsilon^{1/3}, \quad \frac{1-\lambda}{\lambda} \asymp 1, \tag{6.118}$$

(2)

$$g_{11} \asymp -\epsilon^{2/3-\delta} \ , g_{12} \asymp \epsilon^{1/3}, \ g_{22} \asymp -\epsilon^{\delta}, \ \frac{1-\lambda(I)}{\lambda(I)} \asymp \epsilon^{1/3-\delta}, \ \delta \in (0, \tfrac{1}{3}), \tag{6.119}$$

(3)

$$g_{11} \asymp -\epsilon^{2/3} \ , g_{12} \asymp \epsilon^{1/3}, \ g_{22} = \bar{g} \in (-T_1, 0), \ \frac{1-\lambda(I)}{\lambda(I)} \asymp \epsilon^{1/3}. \tag{6.120}$$

For each of the cases above we compute the value of the function I.

Case (1): For graphons as in Case 1, we have

$$\begin{aligned}
I(T_1 + \Delta H_\epsilon) =& \lambda(I)^2 I(T_1 + g_{11}) + 2\lambda(I)(1-\lambda(I))I(T_1 + g_{12}) \\
&+ (1-\lambda(I))^2 I(T_1 + g_{22}) \\
=& I(T_1) + \frac{1}{2}I''(T_1)\big(\lambda(I)^2 g_{11}^2 + 2\lambda(I)(1-\lambda(I))g_{12}^2 \\
&+ (1-\lambda(I))^2 g_{22}^2\big)\epsilon^{2/3} + o(\epsilon^{2/3}) \\
=& I(T_1) + \frac{1}{2}I''(T_1)\left(\frac{(1-\lambda(I))^4}{\lambda(I)^2} + 2\frac{(1-\lambda(I))^3}{\lambda(I)} + (1-\lambda(I))^2\right) g_{22}^2 \\
&+ o(\epsilon^{2/3}) \\
=& I(T_1) + \frac{1}{2}I''(T_1)\frac{(1-\lambda(I))^2}{\lambda(I)^2} g_{22}^2 + o(\epsilon^{2/3}) \\
=& I(T_1) + \frac{1}{2}I''(T_1)T_1^2\epsilon^{2/3} + o(\epsilon^{2/3}) \\
=& I(T_1) + \frac{1}{4}\frac{T_1}{1-T_1}\epsilon^{2/3} + o(\epsilon^{2/3}).
\end{aligned}$$

We observe that there exist multiple graphons that can yield this result. The only constraint we impose is $g_{22}\frac{1-\lambda(I)}{\lambda(I)} = -T_1\epsilon^{1/3}$. For example, the graphon $T_1 + g^*\epsilon^{1/3}$

with

$$g^*(x,y) = \begin{cases} -T_1, & (x,y) \in [0,\frac{1}{2}]^2, \\ T_1, & (x,y) \in [0,\frac{1}{2}] \times (\frac{1}{2},1] \cup (\frac{1}{2},1] \times [0,\frac{1}{2}], \\ -T_1, & (x,y) \in (\frac{1}{2},1]^2, \end{cases} \quad , \tag{6.121}$$

as in (6.17) is balance optimal. In Case (2) below we construct more graphons that are balance optimal.

Case (2): A similar computation as above shows that

$$\begin{aligned} I(T_1 + \Delta H_\epsilon) &= I(T_1) + \frac{1}{2} I''(T_1) T_1^2 \epsilon^{2/3} + o(\epsilon^{2/3}) \\ &= I(T_1) + \frac{1}{4} \frac{T_1}{1-T_1} \epsilon^{2/3} + o(\epsilon^{2/3}). \end{aligned} \tag{6.122}$$

From cases (1) and (2) we observe that various graphons are balance optimal. Hence we need to investigate the higher-order terms in order to determine the optimal graphon.

Case (3): For this case we have

$$\begin{aligned} I(T_1 + \Delta H_\epsilon) &= \lambda(I)^2 \big(I(T_1) + I'(T_1) g_{11}\big) + 2\lambda(I)(1-\lambda(I)) \big(I(T_1) + I'(T_1) g_{12}\big) \\ &\quad + (1-\lambda(I))^2 I(T_1 + \bar{g}) + o(\epsilon^{2/3}) \\ &= I(T_1) + (1-\lambda(I))^2 \left(-I'(T_1)\bar{g} - I(T_1) + I(T_1+\bar{g})\right) + o(\epsilon^{2/3}) \\ &= I(T_1) + T_1^2 \left(\frac{I(T_1+\bar{g}) - I(T_1) - I'(T_1)\bar{g}}{\bar{g}^2} \right) \epsilon^{2/3} + o(\epsilon^{2/3}). \end{aligned} \tag{6.123}$$

Therefore, in order to determine the optimal graphon, we need to find, for a given $T_1 \in (0,1)$, the minimum of the function

$$f(T_1, x) := T_1^2 \frac{I(T_1+x) - I(T_1) - I'(T_1)x}{x^2} \tag{6.124}$$

in $(-T_1, 0)$. We analyze this function for every $T_1 \in (0,1)$ as x varies from $-T_1$ to 0. For $x = -T_1$ we have

$$f(T_1, -T_1) = -I(T_1) + T_1 I'(T_1) = -\frac{1}{2}\log(1-T_1), \tag{6.125}$$

while for $x \uparrow 0$ we have

$$\lim_{x\uparrow 0} f(T_1, x) = T_1^2 \lim_{x\uparrow 0} \frac{I'(T_1+x) - I'(T_1)}{2x} = \frac{1}{2} T_1^2 I''(T_1) = \frac{1}{4}\frac{T_1}{1-T_1}. \tag{6.126}$$

The first derivative is equal to

$$f'(T_1, x) = T_1^2 \frac{\left(I'(T_1+x) - I'(T_1)\right) x^2 - 2x \left(I(T_1+x) - I(T_1) - I'(T_1)x\right)}{x^4} \tag{6.127}$$

and at the boundary points we have

$$\lim_{x\downarrow -T_1} f'(T_1,x) = -\infty, \quad \lim_{x\uparrow 0} f'(T_1,x) = \frac{1}{6} I^{(3)}(T_1) = -\frac{1}{12}\frac{1-2T_1}{(T_1(1-T_1))^2}. \tag{6.128}$$

We observe that $I^{(3)}(T_1) > 0$ if and only if $T_1 > \frac{1}{2}$. Consider first the two endpoints

$$h_1(T_1) = -\frac{1}{2}\log(1-T_1), \quad h_2(T_1) = \frac{1}{4}\frac{T_1}{1-T_1} \tag{6.129}$$

and observe that

$$h_1'(0) = \frac{1}{2} > h_2'(0) = \frac{1}{4}, \quad h_1'(1-\epsilon) = \frac{1}{2\epsilon} < h_2'(\epsilon) = \frac{1}{4\epsilon^2}. \tag{6.130}$$

Both $h_1(\cdot)$ and $h_2(\cdot)$ are increasing function on $[0,1]$. Hence there is a unique $\bar{T}_1$ such that $h_1(T_1) > h_2(T_1)$ for all $T_1 \in (0,\bar{T}_1)$ and $h_1(T_1) \le h_2(T_1)$ for all $T_1 \in [\bar{T}_1, 1)$. Numerically we find $\bar{T}_1 \approx 0.715$ (see also Figure 6.3). We distinguish three cases: $T_1 \in (0, \frac{1}{2}]$, $T_1 \in (\frac{1}{2}, \bar{T}_1]$ and $T_1 \in (\bar{T}_1, 1)$. The results that follow are not rigorously proven, but are derived by using numerical approximations.

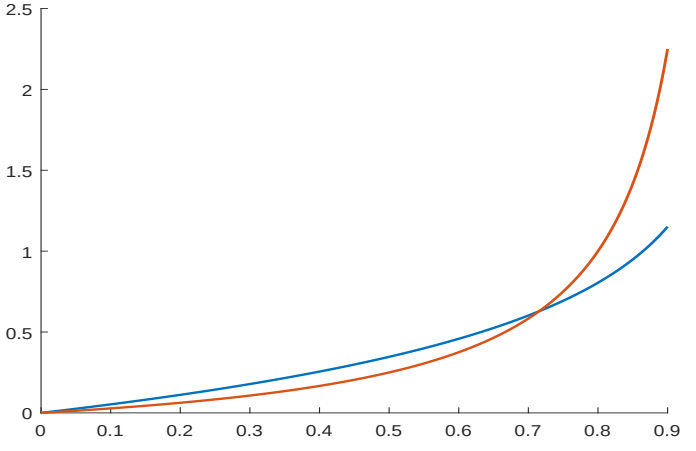

Figure 6.3: Plot of the functions $h_1(\cdot)$ (blue line) and $h_2(\cdot)$ (red line).

Case $T_1 \in (0, \frac{1}{2}]$: We have that $h_1(T_1) = f(T_1, -T_1) > h_2(T_1) = \lim_{x\uparrow 0} f(T_1, x)$. Moreover, $I^{(3)}(T_1) \le 0$, with equality at $T_1 = \frac{1}{2}$. Hence from (6.128) we have that $f(T_1,\cdot)$ decreases away from $f(T_1,-T_1)$ towards $\lim_{x\uparrow 0} f(T_1,x)$. We observe that it is also a decreasing function on $(-T_1, 0)$. Hence we have that

$$f(T_1,x) > \frac{1}{4}\frac{T_1}{1-T_1} \quad \forall x \in (-T_1, 0). \tag{6.131}$$

We illustrate this in Figure 6.4 and 6.5, where we plot $f(T_1,\cdot)$ for $T_1 = 0.1, 0.25, 0.4$ and 0.5.

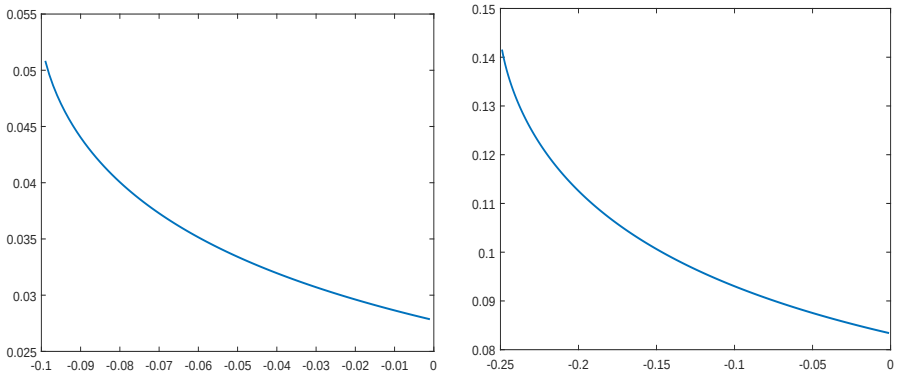

Figure 6.4: Plot of the function $f(T_1, \cdot)$ for $T_1 = 0.1$ (left panel) and $T_1 = 0.25$ (right panel).

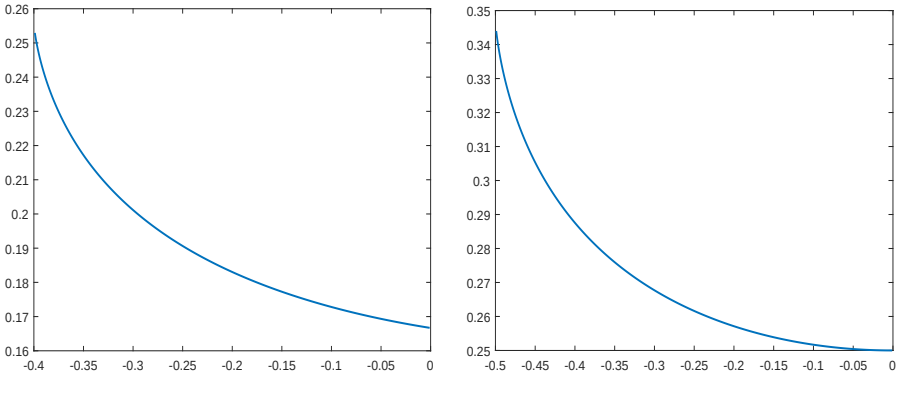

Figure 6.5: Plot of the function $f(T_1, \cdot)$ for $T_1 = 0.4$ (left panel) and $T_1 = 0.5$ (right panel).

Case $T_1 \in (\frac{1}{2}, \bar{T}_1]$**:** In this case we have that $I^{(3)}(T_1) > 0$, and so $f(T_1, x)$ increases into $\lim_{x\uparrow 0} f(T_1, x)$. A similar argument as above in Case 1 shows that there is at least one stationary point $\bar{T}_1^* \in (-T_1, 0)$, which is also a local minimum. Uniqueness of this local minimum is verified numerically, as depicted in Figure 6.6.

Case $T_1 \in (\bar{T}_1, 1)$**:** Using a similar reasoning as in Cases 1 and 2, we get that there is a stationary point $\bar{T}_1^* \in (-T_1, 0)$, which is a local minimum. Uniqueness of this minimum is verified numerically in in Figure 6.7.

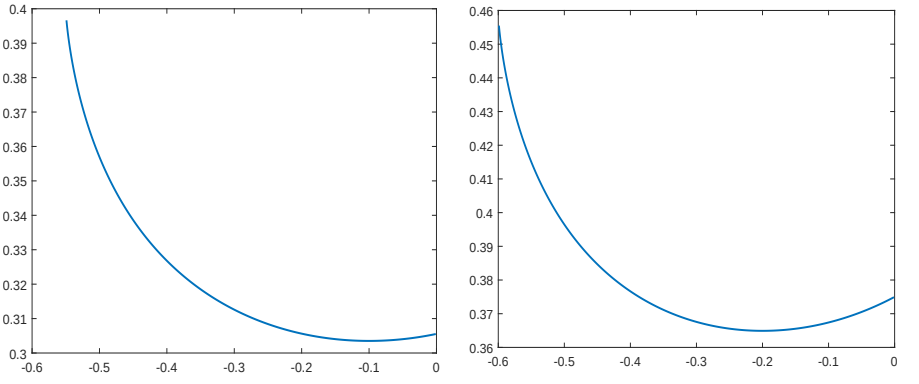

Figure 6.6: Plot of the function $f(T_1, \cdot)$ *for* $T_1 = 0.55$ *(left panel) and* $T_1 = 0.6$ *(right panel).*

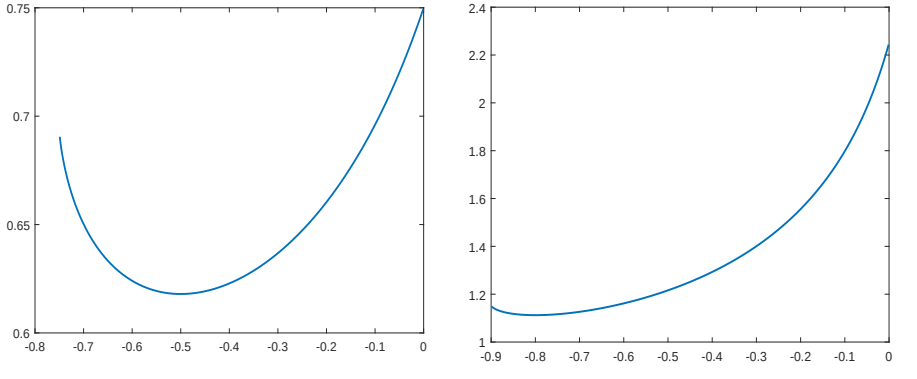

Figure 6.7: Plot of the function $f(T_1, \cdot)$ *for* $T_1 = 0.75$ *(left panel) and* $T_1 = 0.9$ *(right panel).*

From the three cases considered above we observe that if $T_1 \in (0, \frac{1}{2}]$, then

$$f(T_1, x) > \frac{1}{4}\frac{T_1}{1 - T_1}.$$

On the other hand, if $T_1 \in (\frac{1}{2}, 1)$, then the function $f(T_1, \cdot)$ attains a global minimum, denoted by $\bar{T}_1^* \in (-T_1, 0)$, and

$$f(T_1, \bar{T}_1^*) < \frac{1}{4}\frac{T_1}{1 - T_1}.$$

We finally return to (6.123). For $T_1 \in (0, \frac{1}{2}]$ the optimal graphon yields

$$I(T_1 + \Delta H_\epsilon) = I(T_1) + \frac{1}{4}\frac{T_1}{1-T_1}\epsilon^{2/3} + o(\epsilon^{2/3}). \tag{6.132}$$

For $T_1 \in (\frac{1}{2}, 1)$ the optimal graphon yields

$$I(T_1 + \Delta H_\epsilon) = I(T_1) + f(T_1, \bar{T}_1^*)\epsilon^{2/3} + o(\epsilon^{2/3}), \tag{6.133}$$

where $\bar{T}_1^* \in (-T_1, 0)$ is the unique minimizer of the function $f(T_1, \cdot)$ defined in (6.124).