Annals of Mathematics **176** (2012), 151–219
http://dx.doi.org/10.4007/annals.2012.176.1.2

Convergent sequences of dense graphs II. Multiway cuts and statistical physics

By C. Borgs, J. T. Chayes, L. Lovász, V. T. Sós, and K. Vesztergombi

Abstract

We consider sequences of graphs (G_n) and define various notions of convergence related to these sequences including "left-convergence," defined in terms of the densities of homomorphisms from small graphs into G_n, and "right-convergence," defined in terms of the densities of homomorphisms from G_n into small graphs.

We show that right-convergence is equivalent to left-convergence, both for simple graphs G_n, and for graphs G_n with nontrivial nodeweights and edgeweights. Other equivalent conditions for convergence are given in terms of fundamental notions from combinatorics, such as maximum cuts and Szemerédi partitions, and fundamental notions from statistical physics, like energies and free energies. We thereby relate local and global properties of graph sequences. Quantitative forms of these results express the relationships among different measures of similarity of large graphs.

Contents

Research of L. L. and K. V. supported by grant No. CNK 77780 from the National Development Agency of Hungary and by the National Science Foundation under agreement No. DMS-0835373. Research of V. T. S. supported in part by OTKA grants T032236, T038210, T042750.

1. **Introduction**

Growing sequences of graphs arise naturally in many contexts, both fundamental and applied. How do we characterize and classify such sequences? In particular, under what conditions do such sequences converge to something nontrivial and yet sufficiently universal to be conceptually meaningful? A considerable part of graph theory and combinatorics in the past fifty years has been devoted to classifying large, but *finite* graphs. However, surprisingly, until the work here, there was not a general theory for sequences of dense graphs that grow without bound. This paper is the second of two papers in which we develop a theory of *convergent sequences* of dense graphs; see [5] for an announcement of some of these results.

Our theory draws heavily on perspectives and results from both combinatorics and statistical physics. We will therefore explain our results in both languages and provide examples of relevance to both fields.

Consider a dense sequence of simple graphs (G_n) such that the number of nodes in G_n goes to infinity with n (where, as usual, a graph is simple if it has no loops and no multiple edges, and dense means that the average degree grows like the number of vertices in G_n). In this paper we will consider several natural notions of convergence for such a sequence—some motivated by combinatorics and others by statistical physics. Our main result will be a theorem showing that many of these notions of convergence are equivalent. These equivalences allow simple proofs of many of the fundamental results in combinatorics and also provide a framework for addressing some previously unapproachable questions; see, e.g., [2]. These equivalences also help to unify central notions of combinatorics, discrete optimization, and statistical physics.

From the point of view of combinatorics, our theory can be viewed as a substantial generalization of the theory of quasirandom graphs, which are sequences of graphs that "look like" random graphs . Obviously, there are many ways in which one could make this precise, but interestingly, many natural ways in which a sequence of graphs could be defined to be quasirandom turn out to be equivalent [13], [6].

Here we prove similar equivalences for the notion of convergent graph sequences. In fact, most of the equivalences for quasirandom graphs are immediate corollaries of the general theory developed here and in our companion paper [3]. A notable exception is the spectral representation of quasirandom graphs: while it turns out that convergence of the spectrum is implied by our other conditions of convergence, it is not equivalent in our general setting. Indeed, already in the setting of generalized quasirandom graph sequences considered in [8], neither the knowledge of the limiting spectrum of the adjacency matrices nor the knowledge of the limiting spectrum of the Laplacians is enough to characterize the sequences.

From the viewpoint of physics, our results show that convergence of various thermodynamic quantities, notably microcanonical free energies or ground state energies for all so-called "soft-core" models, is equivalent to convergence of apparently more local graph properties, as defined below.

1.1. *Equivalent notions of convergence.* The first notion of convergence for a sequence (G_n) we consider is what we call "left-convergence." It was introduced in the companion of this paper [3] and is a way of characterizing a large graph G in terms of the number of copies of a small graph F that are contained in G. Given two simple graphs F and G, we denote the number homomorphisms from F to G by $\hom(F, G)$. Let $t(F, G)$ be the probability that a random map $\phi : V(F) \to V(G)$ is a homomorphism,

$$t(F, G) = \frac{1}{|V(G)|^{|V(F)|}} \hom(F, G), \tag{1.1}$$

where $V(G)$ and $V(F)$ are the set of vertices in G and F, respectively. We then called a sequence (G_n) of simple graphs *left-convergent* if the "homomorphism densities" $t(F, G_n)$ converge for all simple graphs F.

Instead of testing a graph sequence (G_n) with homomorphisms "from the left," i.e., with homomorphisms from a small graph F into the graphs (G_n), one might want to test (G_n) with homomorphisms "from the right," i.e., one might want to consider the homomorphisms from G_n into some small graph H. For this to be interesting, we have to work with weighted graphs, i.e., graphs H with nodeweights $\alpha_i(H) > 0$ for the nodes $i \in V(H)$ and edgeweights $\beta_{ij}(H) \in \mathbb{R}$ for the edges $ij \in E(H)$. A simple graph can be considered as a weighted graph with all nodeweights and edgeweights equal to 1. The homomorphism number from a simple graph G into a weighted graph H is then defined as

$$\text{(1.2)} \qquad \hom(G,H) = \sum_{\phi:V(G)\to V(H)} \prod_{u\in V(G)} \alpha_{\phi(u)}(H) \prod_{uv\in E(G)} \beta_{\phi(u),\phi(v)}(H),$$

where $E(G)$ denotes the set of edges in G. We will often restrict ourselves to so-called "soft-core" graph, i.e., complete graphs H with all loops present, strictly positive nodeweights $\alpha_i(H) > 0$, and strictly positive edgeweights $\beta_{ij}(H) = \beta_{ji}(H) > 0$.

For soft-core graphs H, these homomorphism numbers "from the right" typically grow or fall exponentially in the number of edges of G. Since the number of edges in a sequence of dense graphs grows like the square of the number of nodes, it seems natural to define a sequence (G_n) of graphs to be right-convergent if $\frac{1}{|V(G_n)|^2} \ln \hom(G_n,H)$ converges for every soft-core graph H. For reasons explained below, we will call such a sequence *naively right-convergent*.

Naive right-convergence turns out to be interesting from both a combinatorics and a statistical physics point of view. Indeed, as we will see below, the convergence of $\frac{1}{|V(G_n)|^2} \ln \hom(G_n,H)$ for a certain graph H on two nodes is equivalent to the convergence of the density of the largest cut in G_n; and right-convergence is equivalent to the convergence of the density of the largest cut in weighted multiway cut problems. From the viewpoint of physics, the homomorphism number $\hom(G,H)$ is just the canonical partition function of a suitable soft-core model on the graph G. One might therefore guess that naive right-convergence corresponds to the convergence of the free energies of these models, but due to our normalization, it actually corresponds to the convergence of ground state energies; see Section 2.3 below.

In contrast to the notion of left-convergence, which corresponds to the convergence of local properties like the density of triangles or the density of 4-cycles, naive right-convergence thus corresponds to convergence of global properties like the density of the largest cut and the ground state energies of suitable soft-core models. This raises the question whether the *a priori* quite different notions of left- and right-convergence are equivalent, the starting point of this paper. While it turns out that left-convergence is not equivalent to naive right-convergence (hence the term naive), a strengthened condition involving homomorphisms for which the number of vertices in G_n that map onto a given $i \in V(H)$ is restricted to be a given fraction of $V(G_n)$ gives equivalence.

In addition to left and right-convergence, we consider several other natural notions of convergence, all of which turn out to be equivalent. Among these notions is that of convergence in a suitably defined metric, a concept already considered in [3]. Another one concerns partitions and the graphs obtained from taking "quotients" with respect to these partitions. More precisely, given a partition $\mathcal{P} = (V_1, \dots, V_q)$ of a graph G, we define the q-quotient $G/\mathcal{P}$ as the weighted graph on $[q]$ with edgeweights β_{ij} given by the edge density between V_i and V_j. (In the theory of Szemerédi partitions, the graph $G/\mathcal{P}$ is often called a cluster graph.) For two graphs G and G' on at least q nodes, we may then want to know how close the sets of q-quotients of these two graphs are. Measuring similarity in terms of Hausdorff distance, this leads to a fourth notion of convergence, convergence of quotients.

In addition to the above four notions, we will be interested in several notions of convergence motivated by statistical physics. We will, in particular, ask under which conditions on a sequence of graphs (G_n) the ground state energies and free energies of finite spin systems defined on G_n are convergent. We also address the same question for the so-called microcanonical ground state energies and free energies. We will show that left-convergence of (G_n) implies convergence of the ground state energies and the free energies of all "soft-core" finite spin systems on (G_n), and we will show that both convergence of the microcanonical ground state energies and convergence of the microcanonical free energies are equivalent to left-convergence.

1.2. *The limit object.* Given the equivalence of the above six notions of convergence, one might want to ask whether a convergent sequence has a natural limit object, in terms of which the limiting homomorphism densities, quotients, free energies, etc. can be expressed.

We start with an example, the random graph sequence $(\mathbb{G}(n,p))$ where, as usual, $\mathbb{G}(n,p)$ is the graph on n nodes in which any two nodes are connected independently with probability p. It is not hard to see that $t(F, \mathbb{G}(n,p))$ converges to $p^{|E(F)|}$ with probability one. Interestingly, this limit can be written as the homomorphism density of a finite weighted graph. Indeed, defining the homomorphism densities of a weighted graph G with nodeweights $\alpha_i(G) > 0$ and $\beta_{ij}(G) \in \mathbb{R}$ by

$$t(F,G) = \frac{\hom(F,G)}{\alpha_G^k}, \tag{1.3}$$

where k is the number of nodes in the simple graph F and $\alpha_G = \sum_{i \in V(G)} \alpha_i(G)$ is the total nodeweight of G, we clearly have that $p^{|E(F)|} = t(F, G_0)$, where G_0 is the graph with one node, a loop at this node, and weight p for the loop. (The node weight is irrelevant in this case, and can, e.g., be set to 1.) This raises the question of which graph sequences have a limit that can be expressed in terms of a finite, weighted graph, which in turn leads to the notion of generalized quasirandom graphs, studied in detail in [8].

For a left-convergent sequence of simple graphs, the limit cannot be expressed in terms of a finite graph in general. Given that one of our equivalences is convergence in metric, one might therefore want to define the limit in the usual abstract way by identifying sequences that are Cauchy. But it turns out that there is a much more natural limit object in terms of measurable, bounded, symmetric 2-variable functions, which we call *graphons*.

It was already observed by Frieze and Kannan [7] that functions of this form are natural generalizations of weighted graphs. (They proved a Regularity Lemma for this generalization.) Of more relevance for us is the work of Lovász and Szegedy [9], who showed that the limit points of left-convergent graph sequences can be identified with graphons, in the sense that given a left-convergent sequence (G_n), there exists a graphon W such that the limit of the homomorphism densities can be expressed in terms of suitably defined homomorphism densities of W.

The notion of a graphon is useful in an even wider setting, and will, in particular, allow us to find simple expressions for the limit objects corresponding to the various notions of convergence considered in this paper. Moreover, most of the statements of our main theorems, Theorems 2.8 and 2.9, have a natural formulation for sequences of uniformly bounded graphons $W_n \in \mathcal{W}$, with proofs that turn out to be much cleaner than the corresponding direct proof of these theorems in terms of graphs. Indeed, many of the technical details of this paper concern rounding techniques that reduce Theorems 2.8 and 2.9 to the corresponding statements for sequence of graphons. It turns out that this approach naturally gives not only the equivalence of the above notions for sequences of *simple* graphs but also for sequences of *weighted* graphs; see Section 2.4 for the precise statements.

The organization of this paper is as follows. In the next section we define our main concepts and state our results; first for sequences of simple graphs, and then for sequences of weighted graphs. The analogues of these concepts and results for graphons are presented in the Section 3 and proved in Section 4. In Section 5 we give the details of the rounding procedures needed to reduce the results of Section 2 to those of Section 3. In our final section, Section 6, we discuss weaker notions of convergence; in particular, convergence of the spectrum of the adjacency matrices, including an example that shows that the convergence of spectra is not sufficient for convergence from the left.

2. Convergent sequences of graphs

2.1. *Definitions.* We start by recalling the definition of left-convergence.

Definition 2.1 ([3]). A sequence (G_n) of simple graphs is called *left-convergent* if the homomorphism densities $t(F, G_n)$ converge for all simple graphs F.

Next we formalize the definition of right-convergence in terms of homomorphism for which the number of vertices in G that map onto a given $i \in V(H)$ is restricted to be a given fraction. To this end, we label the nodes of H as $1, \ldots, q$, and define Pd_q to be the set of vectors $\mathbf{a} \in \mathbb{R}_q$ for which $\sum_i a_i = 1$ and $a_i \geq 0$ for all $i \in [q]$. Given a probability distribution $\mathbf{a} \in \mathsf{Pd}_q$, we set

(2.1)
$$\Omega_{\mathbf{a}}(G) = \Big\{\phi : V(G) \to [q] \colon \Big| |\phi^{-1}(\{i\})| - a_i |V(G)| \Big| \leq 1 \quad \text{for all} \quad i \in [q]\Big\}$$

and define a constrained version of the homomorphism numbers by

$$\hom_{\mathbf{a}}(G, H) = \sum_{\phi \in \Omega_{\mathbf{a}}(G)} \prod_{uv \in E(G)} \beta_{\phi(u)\phi(v)}(H). \tag{2.2}$$

Note the absence of the factors $\alpha_i(H)$ corresponding to the nodeweights. These would be essentially the same for each term and are not carried along. This quantity is natural from the viewpoint of statistical physics: it is the microcanonical partition function on G of a model characterized by the weights in H, at fixed "particle densities" specified by $\mathbf{a}$.

Definition 2.2. A sequence (G_n) of simple graphs (G_n) is called *right-convergent* if

$$\frac{1}{|V(G_n)|^2} \ln \hom_{\mathbf{a}}(G_n, H)$$

converges for every soft-core graph H and every probability distribution $\mathbf{a}$ on $V(H)$, and it is called *naively right-convergent* if

$$\frac{1}{|V(G_n)|^2} \ln \hom(G_n, H)$$

converges for every soft-core graph H.

Example 2.3 (Max-Cut). Let H be the weighted graph on $\{1, 2\}$ with nodeweights $\alpha_1(H) = \alpha_2(H) = 1$ and edgeweights $\beta_{11}(H) = \beta_{22}(H) = 1$ and $\beta_{12}(H) = e$ (where e is the base of the natural logarithm). The leading contributions to $\hom(G, H)$ then come from the maps $\phi :\ V(G) \to \{1, 2\}$ such that the bichromatic edges of ϕ form a maximal cut in G. Using the fact that there are only $2^{|V(G)|}$ mappings, we get that

$$\operatorname{maxcut}(G) \leq \frac{\ln \hom(G, H)}{|V(G)|^2} \leq \operatorname{maxcut}(G) + \frac{\ln 2}{|V(G)|},$$

where $\operatorname{maxcut}(G)$ is the density of the largest cut, i.e., the number of edges in this cut divided by $|V(G)|^2$. This implies, in particular, that for a naively right-convergent sequence (G_n), the density of the largest cut is convergent.

Next we define the metric introduced in [3]. It is derived from the so-called cut-norm and expresses similarity of global structure; graphs with small distance in this metric have cuts of similar size. This is easily made precise for two simple graphs G and G' on the same set V of nodes, where we define

$$d_\square(G,G') = \max_{S,T\subset V}\left|\frac{e_G(S,T)}{|V|^2} - \frac{e_{G'}(S,T)}{|V|^2}\right|,$$

with $e_G(S,T)$ denoting the number of edges in G that have one endpoint in S and one endpoint in T (with edges in $S\cap T$ counted twice).

But some care is needed when G and G' have different nodesets. Here we use the notion of fractional overlays; see [3] for a motivation of our definition. We will give the definition in the more general case where both G and G' are weighted graphs.

Definition 2.4 ([3]). Let G, G' be weighted graphs with nodeset V and V', respectively. For $i\in V$ and $u\in V'$, let $\mu_i = \alpha_i(G)/\alpha_G$ and $\mu'_u = \alpha_u(G')/\alpha_{G'}$. We then define the set of fractional overlays of G and G' as the set $\mathcal{X}(G,G')$ of probability distributions X on $V\times V'$ such that

$$\sum_{u\in V'} X_{iu} = \mu_i \quad \text{for all} \quad i\in V \quad \text{and} \quad \sum_{i\in V} X_{iu} = \mu'_u \quad \text{for all} \quad u\in V',$$

and we set

$$\delta_\square(G,G') = \min_{X\in\mathcal{X}(G,G')}\ \max_{S,T\subset V\times V'}\Bigg|\sum_{\substack{(i,u)\in S\\(j,v)\in T}} X_{iu}X_{jv}\big(\beta_{ij}(G)-\beta_{uv}(G')\big)\Bigg|. \tag{2.3}$$

One of the main results of [3], and one of the main inputs needed for this paper, is the statement that left-convergence is equivalent to convergence in the metric $\delta_\square$.

Another notion of convergence that we will also show to be equivalent is the convergence of "quotients." The quotients of a simple graph G are defined in terms of the partitions $\mathcal{P}=\{V_1,\dots,V_q\}$ of its node set by contracting all nodes in a given group to a new node, leading to a weighted graph $G/\mathcal{P}$ on q nodes. More precisely, we define $G/\mathcal{P}$ as the weighted graph on $[q]$ with weights

$$\alpha_i(G/\mathcal{P}) = \frac{|V_i|}{|V(G)|} \qquad \text{and} \qquad \beta_{ij}(G/\mathcal{P}) = \frac{e_G(V_i,V_j)}{|V_i|\cdot|V_j|}. \tag{2.4}$$

The quotient graph $G/\mathcal{P}$ thus has nodeweights proportional to the sizes of the classes in $\mathcal{P}$, and edgeweights that are equal to the edge densities between the corresponding classes of $\mathcal{P}$. We denote the set of quotients obtained by considering all possible partitions of $V(G)$ into q classes by $\widehat{\mathcal{S}}_q(G)$. Since a quotient $G/\mathcal{P}\in\widehat{\mathcal{S}}_q(G)$ can be characterized by $q+q^2$ real numbers (the node- and edgeweights of $G/\mathcal{P}$), we may consider the set $\widehat{\mathcal{S}}_q(G)$ as a subset of $\mathbb{R}^{q+q^2}$.

It might therefore seem natural to consider two q-quotients as close if their ℓ_1 distance on $\mathbb{R}^{q+q^2}$ is small. But for our purpose, the following distances between two weighted graphs H, H' on q labeled nodes are more useful:

$$d_1(H,H') = \sum_{i,j\in[q]} \Big| \frac{\alpha_i(H)\alpha_j(H)\beta_{ij}(H)}{(\alpha_H)^2} - \frac{\alpha_i(H')\alpha_j(H')\beta_{ij}(H')}{(\alpha_{H'})^2} \Big| + \sum_{i\in[q]} \Big| \frac{\alpha_i(H)}{\alpha_H} - \frac{\alpha_i(H')}{\alpha_{H'}} \Big| \tag{2.5}$$

and

$$d_\square(H,H') = \sup_{S,T\subset[q]} \Big| \sum_{\substack{i\in S\\ j\in T}} \Big(\frac{\alpha_i(H)\alpha_j(H)\beta_{ij}(H)}{(\alpha_H)^2} - \frac{\alpha_i(H')\alpha_j(H')\beta_{ij}(H')}{(\alpha_{H'})^2} \Big) \Big| + \sum_{i\in[q]} \Big| \frac{\alpha_i(H)}{\alpha_H} - \frac{\alpha_i(H')}{\alpha_{H'}} \Big|. \tag{2.6}$$

Let (X,d) be a metric space. As usual, the *Hausdorff metric* d^{Hf} on the set of subsets of X is defined by

$$d^{\mathrm{Hf}}(S,S') = \max\Big\{ \sup_{x\in S} \inf_{y\in S'} d(x,y), \sup_{x\in S'} \inf_{y\in S} d(x,y) \Big\}. \tag{2.7}$$

Definition 2.5. A sequence (G_n) of simple graphs has *convergent quotients* if for all $q\geq 1$, the sequence of sets of quotients $\widehat{\mathcal{S}}_q(G_n)$ is a Cauchy sequence in the Hausdorff distance d_1^{Hf}.

In addition to the four notions of convergence defined above, we will also consider convergence of the free energies and ground state energies of certain models of statistical physics. The models we will be concerned with are so-called soft-core spin systems with finite spin space. They are defined in terms of a finite set $[q]=\{1,\dots,q\}$, a symmetric $q\times q$ matrix J with entries in $\mathbb{R}$ (we denote the set of these matrices by Sym_q), and a vector $h\in\mathbb{R}^q$. A "spin configuration" on a simple graph G is then given by a map $\phi: V(G)\to[q]$, and the energy density of such a spin configurations is defined as

$$\mathcal{E}_\phi(G,J,h) = -\frac{1}{|V(G)|} \sum_{u\in V(G)} h_{\phi(u)} - \frac{2}{|V(G)|^2} \sum_{uv\in E(G)} J_{\phi(u)\phi(v)}. \tag{2.8}$$

Here h_i has the meaning of a generalized magnetic field, describing the preference of the "spin" $\phi(u)$ to be aligned with $i\in[q]$, and J_{ij} represents the strength of the interaction between the spin states $i,j\in[q]$. Note that we divided the second sum by $|V(G)|^2$ to compensate for the fact that, in a dense graph, the number of edges grows like the square of the number of nodes. Our normalization therefore guarantees that the energy density stays bounded uniformly in the size of $V(G)$.

As usual, the partition function on a simple graph G is defined as

$$Z(G,J,h) = \sum_{\phi:V(G)\to[q]} e^{-|V(G)|\mathcal{E}_\phi(G,J,h)}, \tag{2.9}$$

and the free energy and ground state energy per node are defined as

$$\widehat{\mathcal{F}}(G,J,h) = -\frac{1}{|V(G)|}\ln Z(G,J,h) \tag{2.10}$$

and

$$\widehat{\mathcal{E}}(G,J,h) = \min_{\phi:V(G)\to[q]} \mathcal{E}_\phi(G,J,h), \tag{2.11}$$

respectively. We will often leave out the qualifier "per node," and refer to the quantities $\widehat{\mathcal{F}}(G,J,h)$ and $\widehat{\mathcal{E}}(G,J,h)$ as *free energy* and *ground state energy* of the *model* (J,h) *on* G. More specifically, J is called the coupling constant matrix, and h is called the magnetic field, and the model (J,h) will be referred to as the *soft-core model with spin state* $[q]$, *coupling constant matrix* J *and magnetic field* h. We are also interested in the so-called microcanonical versions of these quantities, defined as

$$Z_{\mathbf{a}}(G,J) = \sum_{\phi\in\Omega_{\mathbf{a}}(G)} \exp\Big(-|V(G)|\mathcal{E}_\phi(G,J,0)\Big), \tag{2.12}$$

$$\widehat{\mathcal{F}}_{\mathbf{a}}(G,J) = -\frac{1}{|V(G)|}\ln Z_{\mathbf{a}}(G,J) \tag{2.13}$$

and

$$\widehat{\mathcal{E}}_{\mathbf{a}}(G,J) = \min_{\phi\in\Omega_{\mathbf{a}}(G)} \mathcal{E}_\phi(G,J,0). \tag{2.14}$$

In this microcanonical version, the magnetic field h would only add a constant, and therefore we do not consider it.

Example 2.6 (The Ising Model). The simplest model that fits into our framework is the so-called Ising model. It has spin configurations $\phi : V(G) \to \{-1,+1\}$, and the energy density of a spin configuration ϕ is defined as

$$\mathcal{E}_\phi(G,J,h) = -\frac{1}{|V(G)|^2}\sum_{uv\in E(G)} K\phi_u\phi_v - \frac{1}{|V(G)|}\sum_{u\in V(G)}\mu\phi_u,$$

where K and μ are real parameters. Note that this fits into our scheme by setting $J_{\phi,\phi'} = \frac{K}{2}\phi\phi'$ and $h_\phi = \mu\phi$.

Definition 2.7. Let (G_n) be a sequence of simple graphs. We say that (G_n) has convergent ground state energies and free energies if $\widehat{\mathcal{E}}(G_n,J,h)$ and $\widehat{\mathcal{F}}(G_n,J,h)$ converge for all q, all $h\in\mathbb{R}^q$, and all $J\in\mathsf{Sym}_q$, respectively. Similarly, we say that (G_n) has convergent microcanonical ground state energies and free energies if $\widehat{\mathcal{E}}_{\mathbf{a}}(G_n,J)$ and $\widehat{\mathcal{F}}_{\mathbf{a}}(G_n,J)$ converge for all q, all $\mathbf{a}\in\mathsf{Pd}_q$, and all $J\in\mathsf{Sym}_q$, respectively.

2.2. *Main results for sequences of simple graphs.* The main results of this paper are summarized in the following theorems, except for the results concerning the limiting expression for the ground state energy and free energy, which require some additional notation and are stated in Theorem 3.7 in Section 3.6.

THEOREM 2.8. *Let (G_n) be a sequence of simple graphs such that $|V(G_n)| \to \infty$ as $n \to \infty$. Then the following statements are equivalent*:

(i) *The sequence (G_n) is left-convergent.*
(ii) *The sequence (G_n) is a Cauchy sequence in the metric $\delta_\square$.*
(iii) *The quotients of (G_n) are convergent in the Hausdorff distance d_1^{Hf}.*
(iv) *The sequence (G_n) is right-convergent.*
(v) *The microcanonical ground state energies of (G_n) are convergent.*
(vi) *The microcanonical free energies of (G_n) are convergent.*

Conditions (i) and (ii) were shown to be equivalent in [3]. Extending Example 2.3, it is easy to see that conditions (iv) and (v) are equivalent. (See Lemma 5.7 for a quantitative relation.) Note finally that statements (iii)–(vi) implicitly contain a parameter q, referring to the number of classes in a partition, or the number of nodes in the soft-core graph under consideration. One might therefore ask whether the equivalence of (iii)–(vi) holds separately for each q. While this is true for the equivalence of (iv) and (v), our proofs suggest that this is not the case for the equivalence of (iii) and (v) or (vi).

In contrast to the notions of convergence discussed in Theorem 2.8, convergence of the energies and free energies $\widehat{\mathcal{E}}(G_n, J, h)$ and $\widehat{\mathcal{F}}(G_n, J, h)$ (and naive right-convergence) are not equivalent to left-convergence; see Example 6.3 for a counterexample. But left-convergence does imply convergence of the energies and free energies, as well as naive right-convergence. It also implies convergence of the spectrum. This is the content of our second theorem.

THEOREM 2.9. *Let (G_n) be a left-convergent sequence of simple graphs such that $|V(G_n)| \to \infty$ as $n \to \infty$. Then the following holds*:

(i) *The sequence (G_n) is naively right-convergent.*
(ii) *The ground state energies of (G_n) are convergent.*
(iii) *The free energies of (G_n) are convergent*
(iv) *The spectrum of (G_n) is convergent in the sense that if $\lambda_{n,1} \geq \lambda_{n,2} \geq \cdots \geq \lambda_{n,|V(G_n)|}$ are the eigenvalues of the adjacency matrix of G_n, then $|V(G_n)|^{-1}\lambda_{n,i}$ and $|V(G_n)|^{-1}\lambda_{n,|V(G_n)|+1-i}$ converge for all $i > 0$.*

These theorems, as well their analogues for sequences of weighted graphs, Theorems 2.14 and 2.15, are proved in Section 5, except for the statement about spectra, which is proved in Section 6.

2.3. *Ground state energies, maximum multiway cuts, and quotients.* In this section, we discuss the combinatorial meaning of our results, in particular the relation between ground state energies and generalized max-cut problems on one hand, and the relation between ground state energies and quotients on the other.

We start with the former. To this end, we insert (2.8) into (2.11), leading to

(2.15)
$$-\widehat{\mathcal{E}}(G,J,h) = \max_{\phi:V(G)\to[q]}\Big(\frac{1}{|V(G)|}\sum_{u\in V(G)} h_{\phi(u)} + \frac{2}{|V(G)|^2}\sum_{uv\in E(G)} J_{\phi(u)\phi(v)}\Big).$$

Let us first consider the case of zero magnetic field. For the special case where $q=2$, $J_{ij}=\frac{1}{2}(1-\delta_{ij})$ and $h=0$, the ground state energy of this model can easily be calculated, giving that $-\widehat{\mathcal{E}}(G,J,0)$ is just equal to the density of the largest cut,

$$-\widehat{\mathcal{E}}(G,J,0) = \max_{S\subset V(G)} \frac{e_G(S,V\setminus S)}{|V(G)|^2}.$$

For general q and J, we obtain a natural generalization to weighted multiway cuts. As in Example 2.3, the solution to this weighted multiway cut problem gives a good approximation to $\log\hom(G,H)$ for general soft-core graphs H. More precisely, if $\beta_{ij}(H)=e^{2J_{ij}}$, then

$$\frac{1}{|V(G)|^2}\ln\hom(G,H) = -\widehat{\mathcal{E}}(G,J,0) + O\Big(\frac{1}{|V(G)|}\Big), \tag{2.16}$$

with the implicit constant in the error term depending on the nodeweights of H; see Lemma 5.7. As a consequence, naive right-convergence is equivalent to convergence of the ground state energies for models without magnetic fields.

Turning to nonzero magnetic fields, even the simplest case $q=2$ and $J_{ij}=\frac{1}{2}(1-\delta_{ij})$ leads to a problem that, while quite natural from a combinatorial point of view, to our knowledge has not been studied in the literature. Taking, e.g., $h_i=\mu\delta_{i1}$ with $\mu\in\mathbb{R}$, we get the following generalization of the standard max-cut problem:

$$-\widehat{\mathcal{E}}(G,J,h) = \max_{S\subset V(G)}\Big(\frac{e_G(S,V\setminus S)}{|V(G)|^2} + \mu\frac{|S|}{|V(G)|}\Big).$$

This problem interpolates, to some extent, between the standard max-cut problem (where the size of S is ignored) and the max-bisection problem (where the size of S is prescribed exactly). We will call it the "biased max-cut problem," and the generalization to arbitrary q, J, and h the "biased weighted multiway cut problem."

Considering finally the microcanonical ground state energies,

$$-\widehat{\mathcal{E}}_{\mathbf{a}}(G,J)=\frac{2}{|V(G)|^2}\max_{\phi\in\Omega_{\mathbf{a}}(G)}\sum_{uv\in E(G)}J_{\phi(u)\phi(v)}, \tag{2.17}$$

we are faced with a multiway max-cut problem where the number of vertices in $\phi^{-1}(\{i\})$ is constrained to be approximately equal to $a_i|V(G)|$.

Remark 2.10. If we leave out the convergence of microcanonical free energies, whose combinatorial significance is less clear, the theorems proved in this paper (together with Example 6.3) lead to the following interesting hierarchy of max-cut problems. The weakest form of convergence is that of naive right-convergence, which is equivalent to the convergence of the density of the largest weighted multiway cut (ground state energies with zero magnetic field). The next strongest notion is that of convergence of biased weighted multiway cuts (ground state energies with nonzero magnetic field). The strongest is that of convergence of the weighted multiway cuts with prescribed proportions for the different parts of the cut (microcanonical ground state energies). The remaining notions of convergence (left-convergence, convergence in metric, convergence of quotients, and right-convergence) are equivalent to the convergence of the weighted multiway cuts with arbitrary prescribed proportions.

Turning finally to the relation between quotients and ground state energies, let us note that any map ϕ contributing to the right-hand side of (2.15) defines a partition $\mathcal{P}=(V_1,\dots,V_q)$ of $V(G)$: just set $V_i=\phi^{-1}(\{i\})$. As a consequence, we can rewrite $\widehat{\mathcal{E}}(G,J,h)$ as

$$\widehat{\mathcal{E}}(G,J,h)=-\max_{H\in\widehat{\mathcal{S}}_q(G)}\Big(\sum_{i=1}^{q}\alpha_i(H)h_i+\sum_{i,j=1}^{q}\alpha_i(H)\alpha_j(H)\beta_{ij}(H)J_{ij}\Big). \tag{2.18}$$

This relation shows that the consideration of quotients is quite natural when analyzing weighted multiway cut problems (a.k.a. ground state energies). It also immediately gives that convergence of quotients implies convergence of the ground state energies. The corresponding relation for the microcanonical ground state energies is more complicated due to the fact that a quotient H contributing to $\mathcal{E}_{\mathbf{a}}(G,J)$ has nodeweights that are only approximately equal to the entries of $\mathbf{a}$.

Remark 2.11. Together with the concept of the cut-metric introduced in (2.3), quotients also allow for a very concise formulation of Szemerédi's Regularity Lemma [12], at least in its weak form of Frieze and Kannan [7]. In this formulation, the Weak Regularity Lemma states that given $\varepsilon>0$ and any simple graph G, we can find a $q\le 4^{1/\varepsilon^2}$ and a quotient $H\in\widehat{\mathcal{S}}_q(G)$ such that $\delta_\square(G,H)\le\varepsilon$; see [3] for details.

2.4. *Extension to weighted graphs.* Although Theorems 2.8 and 2.9 are stated for simple graphs, it turns out that the proofs of most of these statements hold more generally, namely for any sequence (G_n) of *weighted* graphs such that (G_n) has uniformly bounded edgeweights and *no dominant nodeweights* in the sense that

$$\frac{\alpha_{\max}(G_n)}{\alpha_{G_n}} \to 0 \qquad \text{as} \qquad n \to \infty, \tag{2.19}$$

where $\alpha_{\max}(G) = \max_{i\in V(G)} \alpha_i(G)$.

We use the symbols $\alpha(G)$ and $\beta(G)$ to denote the vector of nodeweights and the matrix of edgeweights of a weighted graph G. Recall that $\alpha_G = \sum_i \alpha_i(G)$, and set

$$\alpha_{\min}(G) = \min_{i\in V(G)} \alpha_i(G) \quad \text{and} \quad \beta_{\max}(G) = \max_{ij\in E(G)} |\beta_{ij}(G)|.$$

A sequence (G_n) has uniformly bounded edgeweights if $\sup_n \beta_{\max}(G_n) < \infty$.

We generalize the homomorphism numbers $\hom(G, H)$ to the case where both G and H are weighted. Assume thus that H is soft-core, with

$$\alpha_i(H) = e^{h_i} \quad \text{and} \quad \beta_{ij}(H) = e^{2J_{ij}}, \tag{2.20}$$

and that G is a general weighted graph. Setting $\beta_{uv}(G) = 0$ if uv is not an edge in G, we then define

$$\begin{aligned}\hom(G,H) = \sum_{\phi:V(G)\to V(H)} \exp\Big(\sum_{u\in V(G)} \alpha_u(G) h_{\phi(u)} \\ + \sum_{u,v\in V(G)} \alpha_u(G)\alpha_v(G)\beta_{uv}(G) J_{\phi(u)\phi(v)}\Big),\end{aligned} \tag{2.21}$$

an expression that reduces to (1.2) if G is simple.

Remark 2.12. This notation allows us to express partition functions as homomorphism numbers of weighted graphs. For every simple graph G,

$$Z(G, J, h) = \hom(G', H),$$

where G' is obtained from G by weighting its edges by $1/|V(G)|$.

Recall that we defined the metric $\delta_\square$ for general weighted graphs. Let H be a soft-core graph with nodeset $[q]$, and let $\mathbf{a} \in \mathsf{Pd}_q$. For a weighted graph G, we then set

$$\Omega_{\mathbf{a}}(G) = \Big\{ \phi : V(G) \to [q] \colon \Big| \sum_{u\in\phi^{-1}(\{i\})} \alpha_u(G) - a_i \alpha_G \Big| \le \alpha_{\max}(G) \Big\}, \tag{2.22}$$

and we define

$$\hom_{\mathbf{a}}(G,H) = \sum_{\phi\in\Omega_{\mathbf{a}}(G)} \exp\Big(\sum_{u,v\in V(G)} \alpha_u(G)\alpha_v(G)\beta_{uv}(G) J_{\phi(u)\phi(v)}\Big), \tag{2.23}$$

where J is again related to the edgeweights of H by (2.20).

To generalize the notion of quotients to a weighted graph G, let us again consider a partition $\mathcal{P} = (V_1, \dots, V_q)$ of the nodeset of G. We then define the quotient $G/\mathcal{P}$ to be the weighted graph with nodeset $[q]$ and weights

$$\alpha_i(G/\mathcal{P}) = \frac{\alpha_{G[V_i]}}{\alpha_G} \qquad \text{and} \qquad \beta_{ij}(G/\mathcal{P}) = \sum_{\substack{u \in V_i \\ v \in V_j}} \frac{\alpha_u(G)\alpha_v(G)\beta_{uv}(G)}{\alpha_{G[V_i]}\alpha_{G[V_j]}}, \tag{2.24}$$

where $\alpha_{G[V_i]} = \sum_{u \in V_i} \alpha_u(G)$ is the total weight of the partition class V_i. As before, we call $G/\mathcal{P}$ a *q-quotient of* G if $\mathcal{P}$ is a partition of $V(G)$ into q classes, and denote the set of q-quotients of a given graph G by $\widehat{\mathcal{S}}_q(G)$.

To define a soft-core spin model on G, let $[q] = \{1, \dots, q\}$, let $h \in \mathbb{R}^q$, let J be a symmetric $q \times q$ matrix with entries in $\mathbb{R}$, and let $\phi : V(G) \to [q]$. We then generalize the definition (2.8) to

$$\mathcal{E}_\phi(G, J, h) = -\sum_{u \in V(G)} \frac{\alpha_u(G)}{\alpha_G} h_{\phi(u)} - \sum_{u,v \in V(G)} \frac{\alpha_u(G)\alpha_v(G)\beta_{uv}(G)}{\alpha_G^2} J_{\phi(u)\phi(v)}. \tag{2.25}$$

The partition function, free energy, and ground state energy of the model (J, h) on the weighted graph G are then defined in the same way as in the unweighted case; see equations (2.9), (2.10), and (2.11). Similarly, the microcanonical partition functions, free energies, and ground state energies on a weighted graph G are again defined by (2.12), (2.13), and (2.14). Note by definition, the energies (2.25), and hence also the partition functions, free energies, and ground state energies, are invariant under rescaling of the nodeweights of G.

Example 2.13 (The Inhomogeneous Ising Model). Recall the Ising model from Example 2.6, with spin space $\{-1, +1\}$, coupling constants $J_{\phi,\phi'} = \frac{K}{2}\phi\phi'$, and magnetic fields $h_\phi = \mu\phi$. When defined on a simple graph, it is often called a "homogeneous model" because the coupling constants and magnetic fields are constant. But if we take the graph G to be weighted with edgeweights $\beta_{uv}(G)$ (but still unit nodeweights), the model becomes an "inhomogeneous Ising model," with energy density

$$\mathcal{E}_\phi(G, J, H) = -\frac{1}{|V(G)|^2} \sum_{uv \in E(G)} K_{uv}\phi_u\phi_v - \frac{1}{|V(G)|} \sum_{u \in V(G)} \mu\phi_u,$$

where the coupling constants, $K_{uv} = K\beta_{uv}(G)$, represent variations due to inhomogeneities in the underlying crystal structure.

Just as for simple graphs, a sequence (G_n) of weighted graphs with uniformly bounded edgeweights is called *left-convergent* if $t(F, G_n)$ converges for every simple graph F. A sequence (G_n) of weighted graphs is called *right-convergent* if

$$\frac{\ln \hom_{\mathbf{a}}(G_n, H)}{\alpha_{G_n}^2}$$

converges for every soft-core graph H and every probability distribution $\mathbf{a}$ on $V(H)$, and it is called *naively right-convergent* if

$$\frac{\ln \hom(G_n, H)}{\alpha_{G_n}^2}$$

converges for every soft-core graph H.

The following two theorems generalize Theorems 2.8 and 2.9 to weighted graphs.

THEOREM 2.14. *Let (G_n) be a sequence with uniformly bounded edgeweights and no dominant nodeweights. Then the following statements are equivalent*:

(i) *The sequence (G_n) is left-convergent.*
(ii) *The sequence (G_n) is a Cauchy sequence in the metric $\delta_\square$.*
(iii) *The quotients of (G_n) are convergent in the Hausdorff distance d_1^{Hf}.*
(iv) *The microcanonical ground state energies of (G_n) are convergent.*

If, in addition, $\alpha_{G_n}^2/|V(G_n)| \to \infty$, then the following is also equivalent to the statements above:

(v) *The sequence (G_n) is right-convergent.*

If the assumption of no dominant nodeweights is replaced by the stronger assumption that all nodes have weight one and $|V(G_n)| \to \infty$, then the following is also equivalent:

(vi) *The microcanonical free energies of (G_n) are convergent.*

THEOREM 2.15. *Let (G_n) be a left-convergent sequence of weighted graphs with uniformly bounded edgeweights. Then*:

(i) *If (G_n) has no dominant nodeweights and $\alpha_{G_n}^2/|V(G_n)| \to \infty$, then the sequence (G_n) is naively right-convergent.*
(ii) *If (G_n) has no dominant nodeweights, then the ground state energies of (G_n) are convergent.*
(iii) *If all nodes have weight one and $|V(G_n)| \to \infty$, then the free energies of (G_n) are convergent.*
(iv) *The spectrum of (G_n) is convergent in the sense that if $\lambda_{n,1} \ge \lambda_{n,2} \ge \dots \ge \lambda_{n,|V(G_n)|}$ are the eigenvalues of the adjacency matrix of G_n, then $|V(G_n)|^{-1}\lambda_{n,i}$ and $|V(G_n)|^{-1}\lambda_{n,|V(G_n)|+1-i}$ converge for all $i > 0$.*

As pointed out earlier, the equivalence of the first two statements in Theorem 2.14 was proved in the first part of this paper [3]. Here our main focus is on establishing the equivalence of convergence in metric with the other notions of convergence, i.e., the equivalence of (ii) through (vi). Let us note that the additional condition needed for the equivalence of (vi) with the remaining statements is not merely a technical condition. In fact, not all left-convergent sequences of graphs lead to convergent microcanonical free energies if we allow nonconstant nodeweights; see Example 6.4 in Section 6.

Remark 2.16. The reader may notice that none of our theorems assumed that the sequence (G_n) is dense, in the sense that the edge density

$$\frac{1}{\alpha_G^2} \sum_{u,v \in V(G_n)} \alpha_u(G_n)\alpha_v(G_n)\beta_{uv}(G_n)$$

is bounded from below by a constant. That does not mean, however, that our theorems say very much for nondense sequences. Indeed, if the edge density of G_n tends to zero, then most of the statements of the theorem become trivial. The ground state energies and free energies, as well as their microcanonical counterparts, tend to zero, the homomorphism density $t(F, G_n)$ of every simple graph tends to zero, etc.

A similar remark applies to disordered spin systems. While our results for the free energies require that the nodeweights are one, they do not require that $\beta_{uv}(G_n)$ has a definite sign. But if $\frac{1}{|V(G_n)|^2} \sum_{u,v \in V(G_n)} \beta_{uv}(G_n)$ tends to zero (which will happen with probability one if, e.g., β_{uv} is chosen i.i.d. from $\{-1, +1\}$), then the limiting free energies are zero as well. This is due to the fact that we have chosen the ferromagnetic normalization $|V(G_n)|^{-2}$ for the energy $\mathcal{E}_\phi$ per node, rather than the “spin-glass” normalization $|V(G_n)|^{-3/2}$.

Remark 2.17. Let H be a soft-core graph on q nodes, and let $\mathbf{a} \in \mathsf{Pd}_q$. Extending Example 2.3, it is easy to see

$$\frac{\ln \hom_{\mathbf{a}}(G, H)}{\alpha_G^2} = -\widehat{\mathcal{E}}_{\mathbf{a}}(G, J) + O\left(\frac{|V(G_n)|}{\alpha_G^2}\right),$$

with J given by (2.20). (See Lemma 5.7 for a quantitative relation.) This shows why right-convergence is equivalent to the convergence of the microcanonical ground state energies if $\alpha_{G_n}^2/|V(G_n)| \to \infty$.

On the other hand, if we consider sequences (G_n) with $\alpha_{G_n}^2/|V(G_n)| \to c$ for some $c \in (0, \infty)$, then

$$\frac{\ln \hom_{\mathbf{a}}(G, H)}{\alpha_G^2} = -\frac{1}{c}\widehat{\mathcal{F}}_{\mathbf{a}}(G, cJ) + o(1),$$

and right-convergence becomes equivalent to the convergence of the microcanonical free energies.

The least interesting case is the case $\alpha_{G_n}^2/|V(G_n)| \to 0$. In this case,

$$\frac{\ln \hom_{\mathbf{a}}(G,H)}{|V(G_n)|} = \log q + o(1)$$

and the homomorphism numbers $\hom_{\mathbf{a}}(G,H)$ do not contain any interesting information about G_n as $n \to \infty$.

3. **Convergent sequences of graphons**

In this section we discuss the generalization of the concepts and results of the last section to graphons, already mentioned in Section 1.

Definition 3.1. A *graphon* is a bounded measurable function $W : [0,1]^2 \to \mathbb{R}$ which is symmetric; i.e., $W(x,y) = W(y,x)$ for all $(x,y) \in [0,1]^2$.

We denote the subset of graphons with values in some bounded interval I by $\mathcal{W}_I$.

3.1. *Graphons as limits of left-convergent graph sequences.* Let $W \in \mathcal{W}$, and let F be a simple graph with $V(F) = \{1,\dots,k\}$. Following [9], we then define the homomorphism density of W as

$$t(F,W) = \int_{[0,1]^k} \prod_{ij\in E(F)} W(x_i,x_j)\,dx. \tag{3.1}$$

It is not hard to see that this definition extends the definition of homomorphism densities from graphs to graphons. Indeed, let G be a weighted graph on n nodes, and let $I_1,\dots,I_n$ be consecutive intervals in $[0,1]$ of lengths $\alpha_1(G)/\alpha_G$, $\dots$, $\alpha_n(G)/\alpha_G$, respectively. We then define W_G to be the step function that is constant on sets of the form $I_u \times I_v$, with

$$W_G(x,y) = \beta_{uv}(G) \qquad \text{if} \qquad (x,y) \in I_u \times I_v. \tag{3.2}$$

Informally, we consider the adjacency matrix of G and replace each entry (u,v) by a square of size $\alpha_u(G)\alpha_v(G)/\alpha_G^2$ with the constant function β_{uv} on this square. With the above definitions, we have that $t(F,G) = t(F,W_G)$.

Let (G_n) be a sequence of weighted graphs and W be a graphon. We say that $G_n \to W$ if $t(F,G_n) \to t(F,W)$ for every simple graph F. Generalizing the results of [9] to weighted graphs the following was shown in [3].

Theorem 3.2. *For every left-convergent sequence (G_n) of weighted graphs with uniformly bounded edgeweights, there exists a $W \in \mathcal{W}$ such that $G_n \to W$. Conversely, for every $W \in \mathcal{W}$, there exists a sequence (G_n) of weighted graphs with uniformly bounded edgeweights such that $G_n \to W$.*

3.2. *The metric space of graphons.* We will need several norms on the space of graphons. In addition to the standard L_∞, L_1, and L_2 norms of a graphon W (denoted by $\|W\|_\infty$, $\|W\|_1$, and $\|W\|_2$ respectively), we need the cut-norm introduced in [7]. It is defined by

$$\|W\|_\square = \sup_{S,T\subset[0,1]} \left| \int_{S\times T} W(x,y)dxdy \right|,$$

where the supremum goes over measurable subsets of $[0,1]$.

There are several equivalent ways of generalizing the definition of the distance $\delta_\square$ to graphons; see [3]. Here, we define the *cut-distance* of two graphons by

$$\delta_\square(U,W) = \inf_\phi \|U - W^\phi\|_\square, \tag{3.3}$$

where the infimum goes over all invertible maps $\phi: \ [0,1] \to [0,1]$ such that both ϕ and its image are measure preserving, and W^ϕ is defined by $W^\phi(x,y) = W(\phi(x),\phi(y))$. It is not hard to show that this distance indeed extends the distance of weighted graphs, in the sense that $\delta_\square(G,G') = \delta_\square(W_G, W_{G'})$, where W_G is the step function defined in (3.2). We will use the notation $\delta_\square(G,W) = \delta_\square(W_G,W)$ for a weighted graph G and graphon W.

Similar construction can be applied to the L_p norm on $\mathcal{W}$, and we can define distance $\delta_p(U,W) = \inf_\phi \|U - W^\phi\|_p$. (We will need this construction only near the end of the paper for $p = 2$.)

It is not hard to check that $\delta_\square$ satisfies the triangle inequality, so after identifying graphons with distance zero, the space $(\mathcal{W}, \delta_\square)$ becomes a metric space, denoted by $\widetilde{\mathcal{W}}$. The subspace corresponding to the graphons in $\mathcal{W}_I$ will be denoted by $\widetilde{\mathcal{W}}_I$. It was shown in [10] that the space $\widetilde{\mathcal{W}}_{[0,1]}$ is compact. This immediately implies that for any bounded interval I, the metric space $\widetilde{\mathcal{W}}_I$ is compact as well.

One of the main results of our companion paper [3] is the following theorem.

THEOREM 3.3 ([3]). *Let I be a bounded interval, and let (W_n) be a sequence of graphons with values in I.*

(i) *$t(F, W_n)$ is convergent for all simple graphs F if and only if (W_n) is a Cauchy sequence in the metric $\delta_\square$.*

(ii) *Let W be an arbitrary graphon. Then $t(F,W_n) \to t(F,W)$ for all simple graphs F if and only if $\delta_\square(W_n, W) \to 0$.*

In particular, it follows that $G_n \to W$ if and only if $\delta_\square(W_{G_n}, W) \to 0$. We call two graphons W and W' *weakly isomorphic* if $t(F,W) = t(F,W')$ for every simple graph F. It follows from Theorem 3.3 that this is equivalent to $\delta_\square(W,W') = 0$. The results of [4] imply a further equivalent condition: there exists a third graphon U such that $W = U^\phi$ and $W' = U^\psi$ for two measure-preserving functions $\phi, \psi : [0,1] \to [0,1]$.

By the compactness of $\widetilde{\mathcal{W}}_I$, any Cauchy sequence of graphons $W_n \in \mathcal{W}_I$ has a limit $W \in \mathcal{W}_I$, but this does not guarantee uniqueness. Indeed, every graphon weakly isomorphic to W could serve as the limit graphon. It follows from the discussion above that this covers all the nonuniqueness, in other words, the limit is unique as an element of $\widetilde{\mathcal{W}}_I$.

3.3. *Quotients and approximations by step functions.* We call a function $W: [0,1]^2 \to [0,1]$ a *step function*, if $[0,1]$ has a partition $\{S_1, \dots, S_k\}$ into a finite number of measurable sets such that W is constant on every product set $S_i \times S_j$. It can be seen that every step function is at cut-distance zero from W_G for some finite, weighted graph G. Graphons, as limits of finite graphs, can thus be approximated by step functions in the cut-distance. One way to find such an approximation is as follows. Given a graphon $W \in \mathcal{W}$ and a partition $\mathcal{P} = (V_1, \dots, V_q)$ of $[0,1]$ into measurable sets, we define a finite, weighted graph $W/\mathcal{P}$ on $[q]$ by setting

$$\alpha_i(W/\mathcal{P}) = \lambda(V_i) \qquad \text{and} \qquad \beta_{ij}(W/\mathcal{P}) = \frac{1}{\lambda(V_i)\lambda(V_j)} \int_{V_i \times V_j} W(x,y)\,dx\,dy$$

(if $\lambda(V_i)\lambda(V_j) = 0$, we define $\beta_{ij}(W/\mathcal{P}) = 0$) and the corresponding function $W_\mathcal{P}$ by

$$W_\mathcal{P}(x,y) = \sum_{i,j=1}^{q} \beta_{ij}(W/\mathcal{P}) \mathbf{1}_{x \in V_i} \mathbf{1}_{y \in V_j}. \tag{3.4}$$

We call the graph $W/\mathcal{P}$ a *q-quotient* of W and use $\widehat{\mathcal{S}}_q(W)$ to denote the set of all q-quotients of W.

It is not hard to check that the averaging operation $W \mapsto W_\mathcal{P}$ is contractive with respect to the norms $\|\cdot\|_1$, $\|\cdot\|_2$ and $\|\cdot\|_\square$ on $\mathcal{W}$:

$$\|W_\mathcal{P}\|_1 \le \|W\|_1, \qquad \|W_\mathcal{P}\|_2 \le \|W\|_2, \qquad \text{and} \qquad \|W_\mathcal{P}\|_\square \le \|W\|_\square. \tag{3.5}$$

The following theorem is an extension of the Weak Regularity Lemma [7] from graphs to graphons and states that every graphon can be well approximated by a step function.

THEOREM 3.4. *Let $U \in \mathcal{W}$ and $k \ge 1$.*

(i) *There exists a partition $\mathcal{P}$ of $[0,1]$ into at most k measurable parts such that*

$$\|U - U_\mathcal{P}\|_\square < \sqrt{\frac{2}{\log_2 k}} \|U\|_2.$$

(ii) *There exists a $q \le k$ and a quotient $H \in \widehat{\mathcal{S}}_q(U)$ such that*

$$\delta_\square(U, H) < \sqrt{\frac{2}{\log_2 k}} \|U\|_2.$$

The first statement of the theorem gives an approximation of a graphon by step functions and is essentially due to Frieze and Kannan [7]. Indeed, with a slightly worse constant, it follows from Theorem 12 of [7]. In the above form, the first statement of the theorem is proved in Section 4.4.2 below. The second statement gives an approximation by a finite, weighted graph, a factor $U/\mathcal{P} \in \widehat{\mathcal{S}}_q(U)$, and can easily be seen to be equivalent to the first. Stronger versions of the regularity lemma for graphons, in particular a version of the original Szemerédi lemma, can be found in [9], [10].

We will also need a fractional version of q-quotients with which it will be easier to work. First, a *fractional partition* of a set $[0,1]$ into q classes (briefly, a fractional q-partition) is a q-tuple of measurable functions $\rho_1, \dots, \rho_q : [0,1] \to [0,1]$ such that for all $x \in [0,1]$, we have $\rho_1(x) + \cdots + \rho_q(x) = 1$. Given a fractional q-partition $\rho = (\rho_1, \dots, \rho_q)$ of $[0,1]$, we then set

$$\alpha_i(\rho) = \int_0^1 \rho_i(x)dx$$

and define U/ρ to be the weighted graph on $[q]$ with weights

(3.6)
$$\alpha_i(U/\rho) = \alpha_i(\rho) \quad \text{and} \quad \beta_{ij}(U/\rho) = \frac{1}{\alpha_i(\rho)\alpha_j(\rho)} \int_{[0,1]^2} \rho_i(x)\rho_j(y)U(x,y)\,dx\,dy.$$

If $\alpha_i(\rho)\alpha_j(\rho) = 0$, we set $\beta_{ij}(U/\rho) = 0$. We call U/ρ a *fractional q-quotient* of U and denote the set of these fractional q-quotients by $\mathcal{S}_q(U)$.

3.4. *Energy, entropy, and free energy.* Recall the definition (2.8) of the energy density of spin configuration $\phi : V(G) \to [q]$ on a simple graph G. Such a spin configuration defines a partition $\mathcal{P} = (V_1, \dots, V_q)$ of $V(G)$ via $V_i = \phi^{-1}(\{i\})$. In terms of this partition, we can rewrite the energy of the configuration ϕ as

$$\mathcal{E}_\phi(G,J,h) = -\frac{1}{|V(G)|}\sum_i h_i \sum_{u \in V(G)} \mathbf{1}_{u \in V_i} - \frac{1}{|V(G)|^2}\sum_{i,j} J_{ij} \sum_{u,v \in V(G)} \mathbf{1}_{u\in V_i}\mathbf{1}_{v \in V_j}\mathbf{1}_{uv \in E(G)}.$$

Our attempt to generalize this form to graphons leads to the following definitions. Given a graphon W, an integer $q \geq 1$, a matrix $J \in \mathsf{Sym}_q$, and a vector $h \in \mathbb{R}^q$, we define the *energy of a fractional q-partition* ρ of $[0,1]$ as

(3.7)
$$\mathcal{E}_\rho(W,J,h) = -\sum_i h_i \int_{[0,1]} \rho_i(x)\,dx - \sum_{i,j} J_{ij} \int_{[0,1]^2} \rho_i(x)\rho_j(y)W(x,y)\,dx\,dy.$$

The *ground state energy of the model (J,h) on W* is then defined as

$$\mathcal{E}(W,J,h) = \inf_\rho \mathcal{E}_\rho(W,J,h), \tag{3.8}$$

where the infimum runs over all fractional q-partitions of $[0,1]$. The most important energy measure for us will be the *microcanonical ground state energy*, given by

$$\mathcal{E}_{\mathbf{a}}(W,J) = \inf_{\rho:\alpha(\rho)=\mathbf{a}} \mathcal{E}_\rho(W,J,0), \tag{3.9}$$

where the infimum now runs over all fractional q-partitions $[0,1]$ such that $\alpha(\rho)=\mathbf{a}$. Note that

$$\mathcal{E}(W,J,h) = \inf_{\mathbf{a}\in\mathsf{Pd}_q} \Big(\mathcal{E}_{\mathbf{a}}(W,J) - \sum_i a_i h_i\Big). \tag{3.10}$$

As we will see in Theorem 3.7, the definitions (3.8) and (3.9) are not only natural analogues of the corresponding definitions for finite graphs, but they are also the correct limiting expressions of the ground state energies of convergent graphs sequences.

The definition of the free energy of graphs ((2.10) and (2.13)) does not carry over to graphons in a direct way. In fact, there is no natural notion of homomorphism numbers from a graphon W into a finite graph H, which is related to the fact that $\hom(G,H)$ is *not* invariant under blow ups of its first argument (where, as usual, the blow up of a weighted graph G on n nodes is the graph $G[k]$ on kn nodes labeled by pairs iu, $i\in V(G)$, $u=1,\dots,k$, with edgeweights $\beta_{iu,jv}(G[k])=\beta_{ij}(G)$ and nodeweights $\alpha_{iu}(G[k])=\alpha_i(G)$). To circumvent this difficulty, we define the free energy of a graphon W by a variational formula involving the entropy of a fractional q-partition ρ of $[0,1]$,

$$\mathsf{Ent}(\rho) = \int_0^1 \mathsf{Ent}(\rho(x))\,dx \quad\text{with}\quad \mathsf{Ent}(\rho(x)) = -\sum_{i=1}^q \rho_i(x)\ln\rho_i(x). \tag{3.11}$$

In terms of this entropy we define the *free energy* of the model (J,h) on W as

$$\mathcal{F}(W,J,h) = \inf_\rho \Bigg(\mathcal{E}_\rho(W,J,h) - \mathsf{Ent}(\rho)\Bigg), \tag{3.12}$$

where the infimum again runs over all fractional q-partitions of $[0,1]$. The microcanonical free energy is defined analogously:

$$\mathcal{F}_{\mathbf{a}}(W,J) = \inf_{\rho:\alpha(\rho)=\mathbf{a}} \Bigg(\mathcal{E}_\rho(W,J,0) - \mathsf{Ent}(\rho)\Bigg), \tag{3.13}$$

where the infimum again runs over all fractional q-partitions of $[0,1]$ such that $\alpha(\rho)=\mathbf{a}$. Note that again

$$\mathcal{F}(W,J,h) = \inf_{\mathbf{a}\in\mathsf{Pd}_q} \Big(\mathcal{F}_{\mathbf{a}}(W,J) - \sum_i a_i h_i\Big). \tag{3.14}$$

While the definitions (3.12) and (3.13) may seem unintuitive from a mathematical point of view, they are quite natural from a physics point of view. Ultimately, the most convincing justification for these definitions is again given by our results, which prove that the limiting expressions of the free energies of a convergent sequence of graphs are given by (3.12) and (3.13).

3.5. *Equivalent notions of convergence.* Next we state the graphon version of the main result of this paper, Theorem 2.8. It gives several equivalent properties characterizing convergence in the space of graphons.

THEOREM 3.5. *Let I be a bounded interval, and let (W_n) be a sequence of graphons in W_I. Then the following statements are equivalent*:

(i) *For all simple graphs F, the sequence of homomorphism densities $t(F, W_n)$ is convergent.*
(ii) *(W_n) is a Cauchy sequence in the cut-metric $\delta_\square$.*
(iii) *For every $q \geq 1$, the sequence $(\mathcal{S}_q(W_n))$ is Cauchy in the Hausdorff distance d_1^{Hf}.*
(iv) *The sequence $(\mathcal{E}_{\mathbf{a}}(W_n, J))$ is convergent for all $q \geq 1$, all $\mathbf{a} \in \mathsf{Pd}_q$, and all $J \in \mathsf{Sym}_q$.*
(v) *The sequence $(\mathcal{F}_{\mathbf{a}}(W_n, J))$ is convergent for all $q \geq 1$, all $\mathbf{a} \in \mathsf{Pd}_q$, and all $J \in \mathsf{Sym}_q$.*

The reader may notice that the analogue of statement (iv) of Theorem 2.8, i.e., right-convergence of the sequence (W_n), is missing in the above theorem. This is because there is no natural notion of homomorphism numbers from a graphon W into a finite graph H, as explained above. Condition (iv) here corresponds to condition (v) in Theorem 2.8, which (as remarked earlier) is easily seen to be equivalent to condition (iv) in Theorem 2.8.

Finally, taking into account the representations (3.10) and (3.14), we immediately get the following corollary of Theorem 3.5.

COROLLARY 3.6. *Let I be a bounded interval, and let (W_n) be a sequence of graphons in W_I. If $t(F, W_n) \to t(F, W)$ for some $W \in \mathcal{W}$ and all simple graphs F, then $\mathcal{E}(W_n, J, h) \to \mathcal{E}(W, J, h)$ and $\mathcal{F}(W_n, J, h) \to \mathcal{F}(W, J, h)$ for all $q \geq 1$, $h \in \mathbb{R}^q$, and $J \in \mathsf{Sym}_q$.*

By this corollary, the convergence of the energies $\mathcal{E}(W_n, J, h)$ and free energies $\mathcal{F}(W_n, J, h)$ is necessary for the convergence of the homomorphism densities $t(F, W_n)$, but it is not sufficient. In fact, it is not that hard to construct two graphons W and W' that have different homomorphism densities, but for which $\mathcal{E}(W, J, h) = \mathcal{E}(W', J, h)$ and $\mathcal{F}(W, J, h) = \mathcal{F}(W', J, h)$ for all $q \geq 1$, $h \in \mathbb{R}^q$, and $J \in \mathsf{Sym}_q$; see Example 6.1.

3.6. *Limit expressions for convergent sequences of graphs.* Our next theorem states that the limiting quantities referred to in Theorems 2.14 and 2.15 are equal to the corresponding objects defined for graphons.

THEOREM 3.7. *Let $W \in \mathcal{W}$, and let G_n be a sequence of weighted graphs with uniformly bounded edgeweights and no dominant nodeweights. Let F be a simple graph, let $q \geq 1$, $\mathbf{a} \in \mathsf{Pd}_q$, and $J \in \mathsf{Sym}_q$, and let H be a soft-core weighted graph with $\beta_{ij}(H) = e^{2J_{ij}}$. If $\delta_\square(G_n, W) \to 0$, then*

$$\begin{aligned} t(F, G_n) &\to t(F, W), \\ d_1^{\mathrm{Hf}}(\widehat{\mathcal{S}}_q(G_n), \widehat{\mathcal{S}}_q(W)) &\to 0, \\ \widehat{\mathcal{E}}_{\mathbf{a}}(G_n, J) &\to \mathcal{E}_{\mathbf{a}}(W, J), \\ \widehat{\mathcal{E}}(G_n, J, h) &\to \mathcal{E}(W, J, h). \end{aligned}$$

If, in addition, $\alpha_{G_n}^2/|V(G_n)| \to \infty$, then

$$\begin{aligned} -\frac{1}{\alpha_{G_n}^2} \ln \hom_{\mathbf{a}}(G_n, H) &\to \mathcal{E}_{\mathbf{a}}(W, J), \\ -\frac{1}{\alpha_{G_n}^2} \ln \hom(G_n, H) &\to \mathcal{E}(W, J, 0). \end{aligned}$$

If, in addition, all nodes in G_n have weight one, then

$$\begin{aligned} \widehat{\mathcal{F}}_{\mathbf{a}}(G_n, J) &\to \mathcal{F}_{\mathbf{a}}(W, J), \\ \widehat{\mathcal{F}}(G_n, J, h) &\to \mathcal{F}(W, J, h). \end{aligned}$$

We illustrate the last theorem and the expression (3.12) for the limiting free energy in a few simple examples: first the standard ferromagnetic Ising model on a general convergent sequence of simple graphs, next the particularly simple special case in which the convergent sequence is just a sequence of complete graphs, and finally an example of a so-called disordered Ising ferromagnet. We end this section with a general result on the free energy of disordered spin systems.

Example 3.8 (Ising Model on Convergent Graphs Sequences). Consider the inhomogeneous Ising model of Example 2.13 with $K > 0$ (called the ferromagnetic Ising model), and assume that G_n is a sequence of simple graphs such that $G_n \to W$ from the left. By Theorems 3.3 and 3.7, the free energy $\widehat{\mathcal{F}}(G_n, J, h)$ converges to the free energy $\mathcal{F}(W, J, h)$ defined in (3.12). Expressing the fractional partitions $\rho_\pm(x)$ as $\frac{1}{2}(1 \pm m(x))$, we rewrite this expression as

$$\begin{aligned} \mathcal{F}(W, J, h) = \inf_{m:[0,1]\to[-1,1]} \Bigg(&-\frac{K}{2} \int W(x,y) m(x) m(y) dx dy - \mu \int m(x) dx \\ &+ \int \frac{1}{2}(1 + m(x)) \ln\Big(\frac{1}{2}(1 + m(x))\Big) + \int \frac{1}{2}(1 - m(x)) \ln\Big(\frac{1}{2}(1 - m(x))\Big) \Bigg), \end{aligned}$$

where the infimum goes over all measurable functions $m : [0, 1] \to [-1, 1]$.

Example 3.9 (Curie-Weiss Model). Next we specialize to the case where $G_n = K_n$, the complete graph on n nodes. In the physics literature, the Ising model on this graph is known as the mean-field Ising model or as the Curie-Weiss model. For the complete graph, the frequencies $t(\cdot, K_n)$ are easily calculated: $t(F, K_n) = 1 + O(1/n)$, implying that K_n converges to the constant function 1 from the left. By Theorems 3.3 and 3.7, the free energies $\widehat{\mathcal{F}}(K_n, J, h)$ therefore converge to

$$\mathcal{F}(1, J, h) = \inf_{m \in [-1,1]} \Bigg(-\frac{K}{2} m^2 - \mu m \\ + \frac{1+m}{2} \ln(1+m) + \frac{1-m}{2} \ln(1-m) \Bigg) - \ln 2.$$

It is not hard to see that the infimum is in fact a minimum and that the minimizer obeys the equation

$$m = \tanh(Km + \mu),$$

which is the well-known mean-field equation for the "order parameter" m. For $\mu = 0$, this equation has either one or three solutions, depending on whether $K \le 1$ or $K > 1$. The largest solution,

$$M(K) = \max\{m \colon m = \tanh(Km)\},$$

is called the magnetization, and both $m = M(K)$ and $m = -M(K)$ are minimizers for the free energy. It is not hard to see that $M(K) = 0$ for $K \in [0, 1]$ and that for $K > 1$, the function $K \mapsto M(K, 0)$ is a real analytic function that takes values that lie strictly between 0 and 1. As a consequence, the free energy in zero magnetic field, $\mathcal{F}(1, K, 0)$, is an analytic function of K on both $(0, 1)$ and $(1, \infty)$, with a singularity (called a phase transition) at $K = 1$, and

$$\mathcal{F}(1, J, 0) = -\ln 2 \quad \text{if} \quad K \le 1 \qquad \text{and} \qquad \mathcal{F}(1, J, 0) < -\ln 2 \quad \text{if} \quad K > 1.$$

We will use this fact later to give a counterexample showing that not all left-convergent sequences of graphs lead to convergent microcanonical free energies if we allow nonconstant nodeweights.

The function $m(x)$ in Example 3.8 is the inhomogeneous analogue of this order parameter m, and more generally, the fractional partitions $\rho_i(x)$ in (3.12) represent inhomogeneous order parameters for a soft-core spin system with spin space $[q]$.

Example 3.10 (Disordered Ising Ferromagnets). Our next example concerns the Ising model on a simple graph G with nonconstant coupling constants. Writing the varying coupling constants as $K\beta_{uv}$, this can clearly be modeled in our framework by moving from the simple graph G to a weighted graph G' with nodeweights one and edgeweights $\beta_{uv}(G') = \beta_{uv}$. To be specific, let us assume that the weights β_{uv} are chosen i.i.d. from some probability distribution with bounded support and expectation $\bar{\beta}$. It is quite easy to show that whenever the original sequence G_n is left-convergent with $G_n \to W$, then the sequence G'_n is left-convergent with probability one and $G'_n \to \bar{\beta}W$. Thus

$$\widehat{\mathcal{F}}(G'_n, J, h) \to \mathcal{F}(\bar{\beta}W, J, h) = \mathcal{F}(W, \bar{\beta}J, h) \qquad \text{with probability 1.}$$

In order to interpret this result, let us first consider the case where the distribution of β_{uv} is symmetric and $\bar{\beta} = 0$. This represents a so-called spin-glass, and our result only expresses the well-known fact that, with the normalization chosen in equations (2.8) and (2.9), the free energy of a spin glass is zero. For nontrivial results in spin glasses, one would need to scale $J_{\phi(u)\phi(v)}$ by $1/\sqrt{|V(G)|}$ rather than $1/|V(G)|$. If $\bar{\beta}$ is positive, the model describes a so-called disordered ferromagnet, and the above identity expresses the fact that, provided that the coupling asymmetry is strong enough, a disordered ferromagnet on a sequence of dense graphs has the same thermodynamic limit as a homogeneous ferromagnet on the same graph sequence, except for a rescaling of the coupling constant.

As our next proposition shows, the above result holds for arbitrary soft-core spin systems with finite spin space.

Proposition 3.11. *Let (G_n) be a sequence of simple graphs, and let (G'_n) be a sequence of weighted graphs with $V(G'_n) = V(G_n)$, $E(G'_n) = E(G_n)$, nodeweights one, and edgeweights $\beta_{uv}(G'_n) = X^{(n)}_{uv}$, where $X^{(n)}_{uv}$ are real valued i.i.d. random variables with compact support and expectation $\bar{\beta}$. Let $q \geq 1$, $h \in \mathbb{R}^q$, $J \in \mathsf{Sym}_q$, and assume that $\widehat{\mathcal{F}}(G_n, \bar{\beta}J, h)$ converges as $n \to \infty$. Then $\widehat{\mathcal{F}}(G'_n, J, h)$ converges with probability one and*

$$\lim_{n\to\infty} \widehat{\mathcal{F}}(G'_n, J, h) = \lim_{n\to\infty} \widehat{\mathcal{F}}(G_n, \bar{\beta}J, h) \qquad \textit{with probability } 1.$$

Note that the proposition only requires that $\widehat{\mathcal{F}}(G_n, \bar{\beta}J, h)$ is convergent, a condition that is weaker than left-convergence of the original sequence (G_n).

The proof of the proposition gives a similar statement for an arbitrary function from the set of graphs into $\mathbb{R}$ that is invariant under graph isomorphisms and continuous with respect to the cut-metric. As a consequence, an analogue of the above proposition holds, e.g., for the ground state energies $\widehat{\mathcal{E}}(G'_n, J, h)$.

4. **Proof of Theorem 3.5**

The equivalence of (i) and (ii) was proved in [3]. In fact, the following quantitative form is true. (Conclusion (a) was proved in [9] and (b) was proved in [3].)

THEOREM 4.1. *Let* $U, W \in \mathcal{W}$ *and* $C = \max\{1, \|W\|_\infty, \|U\|_\infty\}$.

(a) *Let* F *be a simple graph. Then*

$$|t(F,U) - t(F,W)| \le 4|E(F)|C^{|E(F)|-1}\delta_\square(U,W).$$

(b) *Suppose that for some* $k \ge 1$,

$$|t(F,U) - t(F,W)| \le 3^{-k^2}$$

for every simple graph F *on* k *nodes. Then*

$$\delta_\square(U,W) \le \frac{22C}{\sqrt{\log_2 k}}.$$

This theorem should motivate the rest of the section, where we prove quantitative forms of the main implications among (ii)–(v). We start with some preliminaries.

4.1. *Preliminaries.*

4.1.1. *More on distances for weighted graphs.* Recall that the q-quotients of a graphon U are weighted graphs on q nodes with total nodeweight one. We will often identify these weighted graphs with a point $(\mathbf{a}, X) \in \mathbb{R}^{q+q^2}$, where $\mathbf{a} \in \mathbb{R}^q$ is the vector of nodeweights and $X \in \mathsf{Sym}_q$ is the matrix of edgeweights of the quotient under consideration.

To work with quotients, we will use several different distances on weighted graphs. In addition to the distances d_1 and $d_\square$ introduced in (2.5) and (2.6), we use the ℓ_2-norm of a weighted graph H,

$$\|H\|_2 = \|W_H\|_2 = \left(\sum_{i,j\in V(H)} \frac{\alpha_i(H)\alpha_j(H)}{\alpha_H^2}\beta_{ij}(H)^2\right)^{1/2}$$

and the ℓ_2 distance between two weighted graphs H and H' with the same nodeset and identical nodeweights,

$$d_2(H,H') = \frac{1}{\alpha_H^2}\left(\sum_{i,j\in V} \alpha_i(H)\alpha_j(H)\Big(\beta_{ij}(H) - \beta_{ij}(H')\Big)^2\right)^{1/2}. \tag{4.1}$$

Note that for two weighted graphs with the same nodeset and identical nodeweights, these distances are related to the corresponding norms on graphons by $d_1(H,H') = \|W_H - W_{H'}\|_1$, $d_2(H,H') = \|W_H - W_{H'}\|_2$ and $d_\square(H,H') = \|W_H - W_{H'}\|_\square$.

For a fixed $\mathbf{a} \in \mathsf{Pd}_n$, it will be convenient to introduce on Sym_n the inner product

$$\langle X, Y \rangle_{\mathbf{a}} = \sum_{i,j=1}^{n} a_i a_j X_{ij} Y_{ij} \tag{4.2}$$

and the corresponding norms

$$\|X\|_{\mathbf{a},1} = \sum_{i,j=1}^{n} a_i a_j |X_{ij}|, \quad \|X\|_{\mathbf{a},2} = \langle X, X \rangle_{\mathbf{a}}^{1/2} \tag{4.3}$$

and

$$\|X\|_{\mathbf{a},\square} = \max_{S,T \subseteq [n]} \Big| \sum_{\substack{i \in S \\ j \in T}} a_i a_j X_{ij} \Big|.$$

Note that with these definitions, we have

$$\frac{1}{n^2} \|X\|_{\mathbf{a},1} \leq \|X\|_{\mathbf{a},\square} \leq \|X\|_{\mathbf{a},1} \leq \|X\|_{\mathbf{a},2} \leq \|X\|_{\infty}. \tag{4.4}$$

Note also that for two weighted graphs H, H' with the same nodeweights $\alpha_i(H) = \alpha_i(H')$ and edgeweights $\beta(H) = X$, $\beta(H') = X'$, the above norms allow us to express the distances introduced in (2.5) and (4.1) as

$$d_1(H, H') = \|X - X'\|_{\mathbf{a},1}, \quad d_2(H, H') = \|X - X'\|_{\mathbf{a},2},$$

and

$$d_\square(H, H') = \|X - X'\|_{\mathbf{a},\square},$$

where $\mathbf{a}$ is the vector with components $a_i = \alpha_i(H)/\alpha_H = \alpha_i(H')/\alpha_{H'}$. We will make repeated use of this representation in this paper.

4.1.2. *Fractional and integer quotients.* We start by discussing the relationship between fractional and integer quotients. Let $U \in \mathcal{W}$, let $q \geq 1$, and let $\mathbf{a} \in \mathsf{Pd}_q$. In addition to the sets $\mathcal{S}_q(U)$ and $\widehat{\mathcal{S}}_q(U)$ introduced in Section 3.3, we need the set $\mathcal{S}_{\mathbf{a}}(U)$ of quotients $H \in \mathcal{S}_q(U)$ with $\alpha(H) = \mathbf{a}$, and similarly for $\widehat{\mathcal{S}}_{\mathbf{a}}(U)$, as well as the sets $\mathcal{B}_{\mathbf{a}}(U) = \{X \in \mathsf{Sym}_q : (\mathbf{a}, X) \in \mathcal{S}_q(U)\}$ and $\widehat{\mathcal{B}}_{\mathbf{a}}(U) = \{X \in \mathsf{Sym}_q : (\mathbf{a}, X) \in \widehat{\mathcal{S}}_q(U)\}$.

Note that these sets are invariant under measure preserving bijections $\phi : [0,1] \to [0,1]$. Indeed, for any such ϕ, let $U^\phi(x,y) = U(\phi(x), \phi(y))$ and $\rho^\phi(x) = \rho(\phi(x))$. Then $U/\rho = U^\phi/\rho^\phi$, implying that $\mathcal{S}_q(U) = \mathcal{S}_q(U^\phi)$. In a similar way, one proves that $\widehat{\mathcal{S}}_q(U) = \widehat{\mathcal{S}}_q(U^\phi)$, $\mathcal{B}_{\mathbf{a}}(U) = \mathcal{B}_{\mathbf{a}}(U^\phi)$, and $\widehat{\mathcal{B}}_{\mathbf{a}}(U) = \widehat{\mathcal{B}}_{\mathbf{a}}(U^\phi)$.

The next lemma states that the set of quotients $\mathcal{S}_q(U)$ is compact in the topology induced by the metric d_1 defined in (2.5).

LEMMA 4.2. *Let $U \in \mathcal{W}$ and $q \geq 1$. Then $(\mathcal{S}_q(U), d_1)$ is compact.*

Proof. Let $H_1, H_2, \ldots \in \mathcal{S}_q(U)$. Then there are fractional partitions $\rho^{(i)}$ of $[0,1]$ such that $H_n = U/\rho^{(n)}$. For each $i \in [q]$, let $\mu_i^{(n)}$ be the measure on the Borel sets of $[0,1]$ with density function ρ_i^n. By going to a subsequence, we may assume that the sequence $(\mu_i^{(n)}(D) :\ n = 1, 2, \ldots)$ is convergent for every $i \in [q]$ and rational interval D. Let $\mu_i(D)$ be its limit. From the fact that $\mu_i^{(n)} \le \lambda$ (the Lebesgue measure), it follows that μ_i extends to all Borel sets as a measure and that this measure is absolutely continuous with respect to λ. Hence the function $\rho_i = d\mu_i/d\lambda$ is well defined. It also follows that $0 \le \rho_i \le 1$ almost everywhere and that $\sum_i \rho_i(x) = 1$ for almost all x. So changing the ρ_i on a set of measure 0, we get a fractional partition $\rho = (\rho_1, \ldots, \rho_q)$ of $[0,1]$.

Let $\varepsilon > 0$. Let $\mathcal{P}$ be a partition of $[0,1]$ into rational intervals such that $\|U - U_{\mathcal{P}}\|_1 \le \varepsilon/3$. Then

$$\begin{aligned}\Big|\int_{[0,1]^2} \rho_i^{(n)}(x)\rho_j^{(n)}(y)U(x,y)\,dx\,dy - \int_{[0,1]^2} \rho_i(x)\rho_j(y)U(x,y)\,dx\,dy\Big| \\ \le \Big|\int_{[0,1]^2} \rho_i^{(n)}(x)\rho_j^{n}(y)\Big(U(x,y) - U_{\mathcal{P}}(x,y)\Big)\,dx\,dy\Big| \\ + \Big|\int_{[0,1]^2} \Big(\rho_i^{(n)}(x)\rho_j^{n}(y) - \rho_i(x)\rho_j(y)\Big)U_{\mathcal{P}}(x,y)\,dx\,dy\Big| \\ + \Big|\int_{[0,1]^2} \rho_i(x)\rho_j(y)\Big(U(x,y) - U_{\mathcal{P}}(x,y)\Big)\,dx\,dy\Big|.\end{aligned}$$

The first and third terms on the right-hand side are bounded by $\|U - U_{\mathcal{P}}\|_1/3$; the middle term will be less than $\varepsilon/3$ if n is large enough, since if D is a step of $\mathcal{P}$, then

$$\int_D \rho_i^{(n)}(x)\,dx \to \int_D \rho_i(x)\,dx$$

by the construction of ρ_i. Since $\alpha_i(U/\rho^{(n)}) = \mu_i^n([0,1]) \to \mu_i([0,1]) = \alpha_i(U/\rho)$ for all i, this implies that

$$\beta_{ij}(U/\rho^{(1)}) = \frac{1}{\alpha_i(U/\rho^{(n)})\alpha_j(U/\rho^{(n)})}\int_{[0,1]^2} \rho_i^{(n)}(x)\rho_j^{(n)}(y)U(x,y)\,dx\,dy \to \beta_{ij}(U/\rho)$$

whenever $\alpha_i(U/\rho)\alpha_j(U/\rho) > 0$.

If $\alpha_i(U/\rho)\alpha_j(U/\rho) = 0$, we cannot conclude anything about the limit of $\beta_{ij}(U/\rho^{(n)})$, but fortunately, this is not needed. Indeed, in order to show that $d_1(U/\rho^{(n)}, U/\rho) \to 0$ as $n \to \infty$, we only need to show that $\beta_{ij}(U/\rho^{(n)}) \to \beta_{ij}(U/\rho)$ if $\alpha_i(U/\rho)\alpha_j(U/\rho) > 0$. To see this, we note that the first sum in (2.5) is a sum of terms of the form

$$\Big|\alpha_i(U/\rho)\alpha_j(U/\rho)\beta_{ij}(U/\rho) - \alpha_i(U/\rho^{(n)})\alpha_j(U/\rho^{(n)})\beta_{ij}(U/\rho^{(n)})\Big|.$$

If $\alpha_i(U/\rho)\alpha_j(U/\rho) = 0$, then the first term in this difference is identically zero, while the second tends to zero as $n \to \infty$ due to the facts that $\alpha(U/\rho^{(n)}) \to \alpha(U/\rho)$ and $|\beta_{ij}(U/\rho^{(n)})| \leq \|U\|_\infty$. □

The following lemma is easy to prove along the same lines. Here d_1 is again the distance defined in (2.5), while $d_{\mathbf{a},1}$ is the distance induced by the norm $\|\cdot\|_{\mathbf{a},1}$ defined in (4.3).

LEMMA 4.3. *Let $U \in \mathcal{W}$, let $q \geq 1$, and let $\mathbf{a} \in \mathsf{Pd}_q$. Then $(\mathcal{S}_q(U), d_1)$ is the closure of $(\widehat{\mathcal{S}}_q(U), d_1)$ and $(\mathcal{B}_\mathbf{a}(U), d_{\mathbf{a},1})$ is the closure of $(\widehat{\mathcal{B}}_\mathbf{a}(U), d_{\mathbf{a},1})$.*

While the two sets $\mathcal{S}_q(U)$ and $\widehat{\mathcal{S}}_q(U)$ are equal if U is a step function (see Proposition 5.3), they are not equal in general. This is the content of the following example.

Example 4.4. Let $W \in \mathcal{W}_{[0,1]}$ be positive definite as a kernel. The fractional partition $(\rho, 1-\rho)$ of $[0,1]$ with $\rho \equiv 1/2$ gives a weighted graph $(\mathbf{a}, B)$ on two nodes, with both nodeweights $a_i = 1/2$, and all edgeweights $B_{ij} = \int W(x,y)dxdy$. Using the positive definiteness of W, it is then not hard to see that any fractional partition σ with $W/\sigma = W/\rho$ must actually be equal to ρ almost everywhere. Thus $(\mathbf{a}, B)$ cannot be obtained from any fractional partition other than ρ; in particular, not from any ordinary partition. Hence $\widehat{\mathcal{S}}_q(W) \neq \mathcal{S}_q(W)$.

When analyzing the relationship between ground state energies and quotients, we will naturally be lead to the Hausdorff distance between the subsets of quotients $\mathcal{S}_\mathbf{a}(U)$ and $\mathcal{S}_\mathbf{a}(W)$ for two graphons U and W. The following lemma relates the Hausdorff distance of these two sets to the Hausdorff distance between $\mathcal{S}_q(U)$ and $\mathcal{S}_q(W)$.

LEMMA 4.5. *For any two graphons $U, W \in \mathcal{W}$ and $q \geq 1$,*

$$d_1^{\mathrm{Hf}}(\mathcal{S}_q(U), \mathcal{S}_q(W)) \leq \max_{\mathbf{a}} d_1^{\mathrm{Hf}}(\mathcal{S}_\mathbf{a}(U), \mathcal{S}_\mathbf{a}(W)) \leq (1 + 2\|W\|_\infty) d_1^{\mathrm{Hf}}(\mathcal{S}_q(U), \mathcal{S}_q(W)).$$

Proof. The lower bound is trivial. Let $d = \max_\mathbf{a} d_1^{\mathrm{Hf}}(\mathcal{S}_\mathbf{a}(U), \mathcal{S}_\mathbf{a}(W))$, and let $H \in \mathcal{S}_q(U)$. Then $H \in \mathcal{S}_\mathbf{a}(U)$ for some $\mathbf{a}$, and so by the definition of Hausdorff distance, there is an $H' \in \mathcal{S}_\mathbf{a}(W)$ such that $d_1(H, H') \leq d$. Thus H' is a point in $\mathcal{S}_q(W)$ such that $d_1((\mathbf{a}, B), (\mathbf{a}, B')) = \|B - B'\|_{\mathbf{a},1} \leq d$.

To prove the upper bound, it will be convenient to introduce the distance

$$\tilde{d}_1((\mathbf{a}, B), (\mathbf{b}, C)) = \sum_{i,j} |a_i a_j B_{ij} - b_i b_j C_{ij}|$$

and the Hausdorff distance $\tilde{d}_1^{\mathrm{Hf}}$ inherited from $\tilde{d}_1$. As we will see below, we then have that

$$\tilde{d}_1^{\mathrm{Hf}}(\mathcal{S}_{\mathbf{a}}(W),\mathcal{S}_{\mathbf{b}}(W)) \le 2\|\mathbf{a}-\mathbf{b}\|_1\,\|W\|_\infty \tag{4.5}$$

for all $\mathbf{a},\mathbf{b}\in\mathsf{Pd}_q$.

Before establishing the bound (4.5), we show how it can be used to prove the upper bound of the lemma. Let $(\mathbf{a},B)\in\mathcal{S}_q(U)$, and let $d'=d_1^{\mathrm{Hf}}(\mathcal{S}_q(U),\mathcal{S}_q(W))$. By the definition of Hausdorff distance, there is a weighted graph $(\mathbf{c},D)\in\mathcal{S}_q(W)$ such that

$$d_1((\mathbf{a},B),(\mathbf{c},D))=\tilde{d}_1((\mathbf{a},B),(\mathbf{c},D))+\|\mathbf{a}-\mathbf{c}\|_1\le d',$$

and by the bound (4.5), there is a matrix $B'\in\mathcal{B}_{\mathbf{a}}(W)$ such that

$$\tilde{d}_1((\mathbf{c},D),(\mathbf{a},B'))\le 2\|\mathbf{a}-\mathbf{c}\|_1\|W\|_\infty.$$

Hence

$$\begin{aligned}d_1((\mathbf{a},B),(\mathbf{a},B'))&=\tilde{d}_1((\mathbf{a},B),(\mathbf{a},B'))\le\tilde{d}_1((\mathbf{a},B),(\mathbf{c},D))+\tilde{d}_1((\mathbf{c},D),(\mathbf{a},B'))\\&\le\tilde{d}_1((\mathbf{a},B),(\mathbf{c},D)+2\|\mathbf{a}-\mathbf{c}\|_1\|W\|_\infty\le(1+2\|W\|_\infty)d',\end{aligned}$$

which completes the proof of the upper bound of the lemma.

We are left with the proof of (4.5). Let $H\in\mathcal{S}_{\mathbf{a}}(U)$, so that $H=U/\rho$ for some fractional partition $\rho=(\rho_1,\dots,\rho_q)$ with $\alpha_i(\rho)=a_i$. It is easy to define a fractional partition $\rho'=(\rho_1',\dots\rho_q')$ of $[0,1]$ with $\alpha_i(\rho)=b_i$ and $\sum_i\|\rho_i-\rho_i'\|_1=\|\mathbf{a}-\mathbf{b}\|_1$. In order to prove the bound (4.5), we will show that

$$\tilde{d}_1(U/\rho,U/\rho')\le 2\|\mathbf{a}-\mathbf{b}\|_1\|U\|_\infty.$$

Let $i,j\in[q]$. Then

$$\begin{aligned}&\Big|a_ia_j\beta_{ij}(U/\rho)-b_ib_j\beta_{ij}(U/\rho')\Big|\\&\qquad=\left|\int_{[0,1]^2}U(x,y)\Big(\rho_i(x)\rho_j(y)-\rho_i'(x)\rho_j'(y)\Big)\,dx\,dy\right|\\&\qquad\le\|U\|_\infty\int_{[0,1]^2}\Big|\rho_i(x)\rho_j(y)-\rho_i'(x)\rho_j'(y)\Big|\,dx\,dy\\&\qquad\le\|U\|_\infty\int_{[0,1]^2}\Big|\rho_i(x)\rho_j(y)-\rho_i(x)\rho_j'(y)\Big|\,dx\,dy\\&\qquad\quad+\|U\|_\infty\int_{[0,1]^2}\Big|\rho_i(x)\rho_j'(y)-\rho_i'(x)\rho_j'(y)\Big|\,dx\,dy\\&\qquad=\|U\|_\infty\Big(a_i\|\rho_j-\rho_j'\|_1+\|\rho_i-\rho_i'\|_1b_j\Big).\end{aligned}$$

Summing over i and j this gives the desired bound. □

4.1.3. *Ground state energies and quotients.* We close our section on preliminaries with an expression of the ground state energy and the free energy in a "finite" way in terms of the corresponding quotients. Let $J \in \mathsf{Sym}_q$ and $h \in \mathbb{R}^q$. Using the closedness of $\mathcal{S}_q(W)$ and $\mathcal{B}_{\mathbf{a}}(W)$ and the fact that the map $(\mathbf{a}, X) \mapsto \langle X, J\rangle_{\mathbf{a}} + \langle \mathbf{a}, h\rangle$ is continuous in the d_1-metric, one easily shows that

$$\mathcal{E}(W,J,h) = -\max_{(\mathbf{a},X)\in\mathcal{S}_q(W)} \Big(\langle X, J\rangle_{\mathbf{a}} + \langle \mathbf{a}, h\rangle\Big) \tag{4.6}$$

and

$$\mathcal{E}_{\mathbf{a}}(W,J) = -\max_{X\in\mathcal{B}_{\mathbf{a}}(W)} \sum_{i,j=1}^{q} a_i a_j X_{ij} J_{ij} = -\max_{X\in\mathcal{B}_{\mathbf{a}}(W)} \langle X, J\rangle_{\mathbf{a}}. \tag{4.7}$$

4.2. *From distances to quotients.* The next theorem is a quantitative form of the implication (ii)⇒(iii) in Theorem 3.5.

THEOREM 4.6. *Let $q \geq 1$ and $U, W \in \mathcal{W}$. Then*

$$d_1^{\mathrm{Hf}}(\mathcal{S}_q(U), \mathcal{S}_q(W)) \leq q^2 \delta_\square(U,W).$$

Proof. We first prove that

$$d_\square^{\mathrm{Hf}}(\mathcal{S}_{\mathbf{a}}(U), \mathcal{S}_{\mathbf{a}}(W)) \leq \|U-W\|_\square \tag{4.8}$$

for all $\mathbf{a} \in \mathsf{Pd}_q$. Let $H \in \mathcal{S}_{\mathbf{a}}(U)$. Then there exists a fractional partition $\rho = (\rho_1, \dots, \rho_q)$ of $[0,1]$ such that $H = U/\rho$. Let $H' = (\mathbf{a}, \beta(W/\rho))$. Then for every $S, T \subseteq [q]$, we have

$$\begin{aligned}
&\Big|\sum_{\substack{i\in S\\ j\in T}} a_i a_j (\beta_{ij}(H) - \beta_{ij}(H'))\Big| \\
&\qquad = \Big|\sum_{\substack{i\in S\\ j\in T}} \int_{[0,1]^2} \rho_i(x)\rho_j(y)(U(x,y) - W(x,y))\,dx\,dy\Big| \\
&\qquad = \Big|\int_{[0,1]^2} \Big(\sum_{i\in S}\rho_i(x)\Big)\Big(\sum_{j\in T}\rho_j(y)\Big)(U(x,y)-W(x,y))\,dx\,dy\Big| \\
&\qquad \leq \|U-W\|_\square,
\end{aligned}$$

and hence $d_\square(H,H') \leq \|U-W\|_\square$, which proves the bound (4.8).

Since the sets $\mathcal{S}_{\mathbf{a}}(U)$ and $\mathcal{S}_{\mathbf{a}}(W)$ are invariant under measure preserving bijections, the bound (4.8) implies that $d_\square^{\mathrm{Hf}}(\mathcal{S}_{\mathbf{a}}(U), \mathcal{S}_{\mathbf{a}}(W)) \leq \delta_\square(U,W)$, and taking into account the bound (4.4), this in turn implies that

$$d_1^{\mathrm{Hf}}(\mathcal{S}_{\mathbf{a}}(U), \mathcal{S}_{\mathbf{a}}(W)) \leq q^2 \delta_\square(U,W). \tag{4.9}$$

Together with Lemma 4.5 this gives the desired bound on $d_1^{\mathrm{Hf}}(\mathcal{S}_q(U), \mathcal{S}_q(W))$. □

4.3. *From quotients to energies.* The next theorem is a quantitative version of the implication (iii)⇒(iv) from Theorem 3.5.

THEOREM 4.7. *Let* $q \geq 1$, $\mathbf{a} \in \mathsf{Pd}_q$, $J \in \mathsf{Sym}_q$, *and* $U, W \in \mathcal{W}$. *Then*

$$|\mathcal{E}_{\mathbf{a}}(U,J) - \mathcal{E}_{\mathbf{a}}(W,J)| \leq (1 + 2\|W\|_\infty)\, \|J\|_\infty\, d_1^{\mathrm{Hf}}(\mathcal{S}_q(U), \mathcal{S}_q(W)).$$

Proof. In view of Lemma 4.5, it is enough to prove that

$$|\mathcal{E}_{\mathbf{a}}(U,J) - \mathcal{E}_{\mathbf{a}}(W,J)| \leq \|J\|_\infty d_1^{\mathrm{Hf}}(\mathcal{S}_{\mathbf{a}}(U), \mathcal{S}_{\mathbf{a}}(W)). \tag{4.10}$$

Let $H \in \mathcal{S}_{\mathbf{a}}(U)$ attain the maximum in the representation (4.7) for $\mathcal{E}_{\mathbf{a}}(U,J)$, so that

$$\mathcal{E}_{\mathbf{a}}(U,J) = -\langle J, \beta(H)\rangle_{\mathbf{a}}.$$

By the definition of Hausdorff distance, there is an $H' \in \mathcal{S}_{\mathbf{a}}(W)$ such that $d_1(H,H') \leq d_1^{\mathrm{Hf}}(\mathcal{S}_{\mathbf{a}}(U), \mathcal{S}_{\mathbf{a}}(W))$. Then

$$\begin{aligned}
\mathcal{E}_{\mathbf{a}}(W,J) - \mathcal{E}_{\mathbf{a}}(U,J) &\leq \langle J, \beta(H)\rangle_{\mathbf{a}} - \langle J, \beta(H')\rangle_{\mathbf{a}} \\
&= \langle J, \beta(H) - \beta(H')\rangle_{\mathbf{a}} \leq \|J\|_\infty \sum_{i,j=1}^{q} a_i a_j |\beta_{ij}(H) - \beta_{ij}(H')| \\
&= \|J\|_\infty d_1(H,H') \leq \|J\|_\infty d_1^{\mathrm{Hf}}(\mathcal{S}_{\mathbf{a}}(U), \mathcal{S}_{\mathbf{a}}(W)).
\end{aligned}$$

In a similar way, one proves a lower bound of $-\|J\|_\infty d_1^{\mathrm{Hf}}(\mathcal{S}_{\mathbf{a}}(U), \mathcal{S}_{\mathbf{a}}(W))$, giving (4.10) and hence the statement of the theorem. □

The following theorem is the analogue of Theorem 4.7 for the ground state energies $\mathcal{E}(W,J,h)$ and is a quantitative version of the first statement from Corollary 3.6.

THEOREM 4.8. *Let* $q \geq 1$, $h \in \mathbb{R}^q$, $J \in \mathsf{Sym}_q$, *and* $U, W \in \mathcal{W}$. *Then*

$$|\mathcal{E}(U,J,h) - \mathcal{E}(W,J,h)| \leq \max\{\|J\|_\infty, \|h\|_\infty\}\, d_1^{\mathrm{Hf}}(\mathcal{S}_q(U), \mathcal{S}_q(W)). \tag{4.11}$$

Proof. This bound is proved in the same way as the bound (4.10) and is left to the reader. □

4.4. *From energies back to distances.* Combining the bounds (4.9) and (4.10), we get

$$|\mathcal{E}_{\mathbf{a}}(U,J) - \mathcal{E}_{\mathbf{a}}(W,J)| \leq q^2 \|J\|_\infty \delta_\square(W,U). \tag{4.12}$$

The next theorem, which is one of the main results in this paper, gives a bound in the opposite direction, and thereby provides a quantitative proof of the implication (iv)⇒(ii) in Theorem 3.5.

THEOREM 4.9. *Let* $U, W \in \mathcal{W}$, *and suppose that*

$$|\mathcal{E}_{\mathbf{a}}(U,J) - \mathcal{E}_{\mathbf{a}}(W,J)| \leq \frac{\varepsilon^2}{64q^2} \|J\|_\infty \max\{\|U\|_\infty, \|W\|_\infty\}$$

for all $q \le 4^{9/\varepsilon^2}$, $\mathbf{a} \in \mathsf{Pd}_q$, *and* $J \in \mathsf{Sym}_q$. *Then*

$$\delta_\square(U, W) \le \varepsilon \max\{\|U\|_\infty, \|W\|_\infty\}.$$

The proof of Theorem 4.9, to be given in the next sections, is based on the following idea, which is very similar to the main idea in the proof of the Weak Regularity Lemma. For $q \ge 1$, let

$$\mathcal{L}_q(U) = \max_{(\mathbf{a},B)\in\mathcal{S}_q(U)} (-\mathcal{E}_{\mathbf{a}}(U, B)) \tag{4.13}$$

and

$$\Delta_q(U) = \sqrt{\mathcal{L}_{4q}(U) - \mathcal{L}_q(U)}. \tag{4.14}$$

We will show that

$$\delta_\square(U, H) \le \Delta_q(U) \tag{4.15}$$

whenever $H = (\mathbf{a}, B)$ is such that it attains the maximum in (4.13). Since $0 \le \mathcal{L}_q(U) \le \|U\|_\infty^2$ for all q, $\mathcal{L}_q$ cannot decrease by a substantial amount too many times implying, in particular, that there must be a $q \le 4^{9/\varepsilon^2}$ such that $\Delta_q \le \frac{\varepsilon}{3}\|U\|_\infty$. But this implies that for this q, the maximizer in (4.13) must be a good approximation to U in the $\delta_\square$ distance, $\delta_\square(U, H) \le \frac{\varepsilon}{3}\|U\|_\infty$. Thus a good knowledge of the ground state energies allows us to calculate a good approximation to the graphon U by a finite graph in the $\delta_\square$ distance.

4.4.1. *The geometry of fractional quotients.* In this subsection, we give a different representation for $\mathcal{L}_q(U)$, which will allow us to prove (4.15). To this end, we first prove the following lemma.

LEMMA 4.10. *Given* $q \ge 1$, $\mathbf{a} \in \mathsf{Pd}_q$, *and* $U \in \mathcal{W}$, *let*

$$L_{\mathbf{a}}(U) = \max_{B\in\mathcal{B}_{\mathbf{a}}(U)} \|B\|_{\mathbf{a},2}^2. \tag{4.16}$$

Then

$$L_{\mathbf{a}}(U) = \max_{B\in\mathcal{B}_{\mathbf{a}}(U)} \Big(-\mathcal{E}_{\mathbf{a}}(U, B)\Big), \tag{4.17}$$

where any B *that attains the maximum in the first expression also attains the maximum in the second expression, and vice versa.*

Proof. Since $\langle X, B\rangle_{\mathbf{a}} \le \|X\|_{\mathbf{a},2}\|B\|_{\mathbf{a},2} \le L_{\mathbf{a}}(U)$, we have

$$\|B\|_{\mathbf{a},2}^2 = \langle B, B\rangle_{\mathbf{a}} \le -\mathcal{E}_{\mathbf{a}}(U, B) = \max_{X\in\mathcal{B}_{\mathbf{a}}(U)} \langle X, B\rangle_{\mathbf{a}} \le L_{\mathbf{a}}(U).$$

Taking the maximum over $B \in \mathcal{B}_{\mathbf{a}}(U)$, we obtain the identity (4.17) as well as the statement that any matrix that attains the maximum in (4.16), also attains the maximum in (4.17). To prove the converse statement, we use that $\langle (X-B),(X-B)\rangle_{\mathbf{a}} \geq 0$ for all $X, B \in \mathcal{B}_{\mathbf{a}}(U)$ implying, in particular, that

$$-2\mathcal{E}_{\mathbf{a}}(U,B) \leq \|B\|_{\mathbf{a},2}^2 + L_{\mathbf{a}}(U).$$

If B_0 is such that $-\mathcal{E}_{\mathbf{a}}(U,B)$ attains its maximum for $B = B_0$, we therefore have that

$$2L_{\mathbf{a}}(U) = -2\mathcal{E}_{\mathbf{a}}(U,B_0) \leq \|B_0\|_{\mathbf{a},2}^2 + L_{\mathbf{a}}(U) \leq 2L_{\mathbf{a}}(U),$$

which implies that $\|B_0\|_{\mathbf{a},2}^2 = L_{\mathbf{a}}(U)$, as required. □

4.4.2. *Step function approximation.* As a consequence of Lemma 4.10, we may rewrite $\mathcal{L}_q(U)$ as

$$\mathcal{L}_q(U) = \sup_{\mathcal{P}\in P_q} \|U_{\mathcal{P}}\|_2^2, \tag{4.18}$$

where the supremum goes over all partitions of $[0,1]$ into q classes. Indeed, let $\mathcal{P}$ be a partition of $[0,1]$ into q classes, and let $\mathbf{a} = \alpha(U/\mathcal{P})$. Then $U/\mathcal{P}$ is a quotient of U, and

$$\|U_{\mathcal{P}}\|_2^2 = \sum_{i,j=1}^{q} \alpha_i(U/\mathcal{P})\alpha_j(U/\mathcal{P})\beta_{ij}(U/\mathcal{P}) = \|\beta(U/\mathcal{P})\|_{\mathbf{a},2}^2.$$

Using the fact that $\mathcal{B}_{\mathbf{a}}(U)$ is the closure of $\widehat{\mathcal{B}}_{\mathbf{a}}(U)$, we now rewrite the right-hand side of (4.18) as

$$\sup_{\mathcal{P}\in P_q} \|U_{\mathcal{P}}\|_2^2 = \sup_{\mathbf{a}\in \mathsf{Pd}_q} \sup_{B\in\widehat{\mathcal{B}}_{\mathbf{a}}(U)} \|B\|_{\mathbf{a},2}^2 = \max_{\mathbf{a}\in \mathsf{Pd}_q} \max_{B\in\mathcal{B}_{\mathbf{a}}(U)} \|B\|_{\mathbf{a},2}^2.$$

With the help of Lemma 4.10, this gives (4.18). In particular, it follows that

$$\mathcal{L}_q(U) \leq \|U\|_2^2. \tag{4.19}$$

The next lemma will be important in proving bounds on the approximation by step functions.

LEMMA 4.11. *For every partition $\mathcal{P}$ of $[0,1]$ into q classes, we have*

$$\|U - U_{\mathcal{P}}\|_{\square}^2 \leq \mathcal{L}_{4q}(U) - \|U_{\mathcal{P}}\|_2^2.$$

Proof. Let S and T be arbitrary measurable subsets of $[0,1]$, and let $\mathcal{P}'$ be the partition of $[0,1]$ generated by S, T and $\mathcal{P}$. Clearly $\mathcal{P}'$ has at most $4q$ classes. Since $U_{\mathcal{P}'}$ gives the best L_2-approximation of U among all step functions with steps $\mathcal{P}'$, we conclude that for every real number t, we have

$$\|U - U_{\mathcal{P}'}\|_2^2 \leq \|U - U_{\mathcal{P}} - t\mathbf{1}_{S\times T}\|_2^2,$$

which in turn implies that

$$\|U - U_{\mathcal{P}'}\|_2^2 \leq \|U - U_{\mathcal{P}}\|_2^2 - 2t\langle \mathbf{1}_{S\times T}, U - U_{\mathcal{P}}\rangle + t^2.$$

Choosing $t = \langle \mathbf{1}_{S\times T}, U - U_{\mathcal{P}}\rangle$, this gives

$$\langle \mathbf{1}_{S\times T}, U - U_{\mathcal{P}}\rangle^2 \le \|U-U_{\mathcal{P}}\|_2^2 - \|U-U_{\mathcal{P}'}\|_2^2 = \|U_{\mathcal{P}'}\|_2^2 - \|U_{\mathcal{P}}\|_2^2 \le \mathcal{L}_{4q}(U) - \|U_{\mathcal{P}}\|_2^2.$$

Since the supremum of the left-hand side over all sets S, T is just $\|U - U_{\mathcal{P}}\|_\square^2$, this proves the statement of the lemma. □

It is instructive to show that Lemma 4.11 implies the Weak Regularity Lemma for graphons, Theorem 3.4.

Proof of Theorem 3.4. Set $\varepsilon = \|U\|_2\sqrt{2/\log_2 k}$. If $\|U - U_{\mathcal{P}}\|_\square \ge \varepsilon$ for all partitions $\mathcal{P}$ with at most k classes, then by Lemma 4.11, $\mathcal{L}_{4q}(U) - \|U_{\mathcal{P}}\|_2^2 \ge \varepsilon^2$ for every $1 \le q \le k$ and every $\mathcal{P} \in \mathcal{P}_q$. Hence $\mathcal{L}_{4q}(U) - \mathcal{L}_q \ge \varepsilon^2$ for every $1 \le q \le k$, which in turn implies that

$$\mathcal{L}_{4k} \ge \Big(\Big\lfloor \frac{1}{2}\log_2 k\Big\rfloor + 1\Big)\varepsilon^2 > \Big(\frac{1}{2}\log_2 k\Big)\varepsilon^2 \ge \|U\|_2,$$

which contradicts (4.19). □

The following corollary verifies (4.15).

COROLLARY 4.12. *Let $q \ge 1$, $U \in \mathcal{W}$, and $H \in \mathcal{S}_q(U)$. Then*

$$\delta_\square(U, H) \le \sqrt{\mathcal{L}_{4q}(U) - \|H\|_2^2}. \tag{4.20}$$

If H attains the maximum in (4.13), *then*

$$\delta_\square(U, H) \le \sqrt{\mathcal{L}_{4q}(U) - \mathcal{L}_q(U)}.$$

Proof. By Lemma 4.10, the second bound of the lemma immediately follows from the first. Thus it is enough to prove (4.20). Let $\mathcal{P}$ be a partition of $[0,1]$ into q classes, and let $U/\mathcal{P} = H$ be the corresponding integer quotient of U. By Lemma 4.11, we have that

$$\delta_\square(U, U/\mathcal{P})^2 \le \|U - U_{\mathcal{P}}\|_\square^2 \le \mathcal{L}_{4q} - \|H\|_2^2.$$

Since $\mathcal{S}_{\mathbf{a}}(U)$ is the closure of $\widehat{\mathcal{S}}_{\mathbf{a}}(U)$, this gives (4.20), as desired. □

4.4.3. *Completion of the proof.* Rescaling W and U by a constant factor if necessary, we may assume that $\|U\|_\infty, \|W\|_\infty \le 1$. Let $\tau = \varepsilon^2/(64q^2)$, and let $q_0 = 4^{\lceil 9/\varepsilon^2\rceil - 1}$.

Since $0 \le \mathcal{L}_q(U) \le 1$, there is a $1 \le q \le q_0$ such that $\mathcal{L}_{4q}(U) - \mathcal{L}_q(U) \le \frac{\varepsilon^2}{9}$. Choose $H = (\mathbf{a}, B) \in \mathcal{S}_q(U)$ in such a way that $\mathcal{L}_q(U) = -\mathcal{E}_{\mathbf{a}}(U, B)$. We have

$$\delta_\square(U, W) \le \delta_\square(U, H) + \delta_\square(H, H') + \delta_\square(H', W).$$

Let us estimate the three terms on the right-hand side separately.

By Corollary 4.12, we have

$$\delta_\square(U, H) \le \frac{\varepsilon}{3}, \tag{4.21}$$

and by Lemma 4.10, we have that $-\mathcal{E}_{\mathbf{a}}(U,B) = \|B\|^2_{\mathbf{a},2}$. Due to the assumption that $\|U\|_\infty \le 1$, we also have $\|B\|_\infty \le 1$.

Let $Y \in \mathcal{B}_{\mathbf{a}}(W)$ attain the maximum in the definition of $-\mathcal{E}_{\mathbf{a}}(W,B)$. Then

$$\langle Y,B\rangle_{\mathbf{a}} = -\mathcal{E}_{\mathbf{a}}(W,B) \ge -\mathcal{E}_{\mathbf{a}}(U,B) - \tau = \|B\|^2_{\mathbf{a},2} - \tau, \tag{4.22}$$

and also

$$\begin{aligned}\langle Y,Y\rangle_{\mathbf{a}} &\le -\mathcal{E}_{\mathbf{a}}(W,Y) \le -\mathcal{E}_{\mathbf{a}}(U,Y) + \tau \\ &= \max_{X\in\mathcal{B}_{\mathbf{a}}(U)} \langle X,Y\rangle_{\mathbf{a}} + \tau \le \|B\|_{\mathbf{a},2}\|Y\|_{\mathbf{a},2} + \tau \\ &\le \frac{1}{2}\Big(\|B\|^2_{\mathbf{a},2} + \|Y\|^2_{\mathbf{a},2}\Big) + \tau,\end{aligned}$$

implying that

$$\langle Y,Y\rangle_{\mathbf{a}} \le \|B\|^2_{\mathbf{a},2} + 2\tau.$$

Hence

$$\begin{aligned}\|B-Y\|^2_{\mathbf{a},2} &= \|B\|^2_{\mathbf{a},2} + \|Y\|^2_{\mathbf{a},2} - 2\langle B,Y\rangle_{\mathbf{a}} \\ &\le \|B\|^2_{\mathbf{a},2} + (\|B\|^2_{\mathbf{a},2} + 2\tau) - 2(\|B\|^2_{\mathbf{a},2} - \tau) = 4\tau.\end{aligned}$$

Let $H' = (\mathbf{a}, Y)$. Using Cauchy-Schwarz, we get that

$$\delta_\square(H,H') \le \|B-Y\|_{\mathbf{a},\square} \le q\sqrt{4\tau} \le \frac{\varepsilon}{4}. \tag{4.23}$$

We are left with a bound on $\delta_\square(H',W)$. To this end, we again use Corollary 4.12, this time in the form of the bound (4.20), which gives that

$$\Big(\delta_\square(H',W)\Big)^2 \le \mathcal{L}_{4q}(W) - \|Y\|^2_{\mathbf{a},2}. \tag{4.24}$$

By the definition of $\mathcal{L}_{\mathbf{b}}$ and the conditions of the theorem, we have that $L_{\mathbf{b}}(W) \le L_{\mathbf{b}}(U) + \tau$ for every $\mathbf{b} \in \mathsf{Pd}_{4q}$, and hence

$$\mathcal{L}_{4q}(W) \le \mathcal{L}_{4q}(U) + \tau. \tag{4.25}$$

On the other hand, (4.22) implies that $\|Y\|^2_{a2} + \|B\|^2_{a2} \ge 2\langle Y,B\rangle_{\mathbf{a}} \ge 2\|B\|^2_{a2} - 2\tau$, and so

$$\|Y\|^2_{a2} \ge \|B\|^2_{2a} - 2\tau = \mathcal{L}_q(U) - 2\tau. \tag{4.26}$$

Combining (4.24), (4.25), and (4.26), we get

$$\Big(\delta_\square(H',W)\Big)^2 \le \mathcal{L}_{4q}(U) + \tau - \mathcal{L}_q(U) + 2\tau \le \frac{\varepsilon^2}{9} + 3\tau \le \Big(\frac{5\varepsilon}{12}\Big)^2,$$

and so

$$\delta_\square(H',W) \le \frac{5\varepsilon}{12}. \tag{4.27}$$

To sum up, by (4.21), (4.23), and (4.27), we get

$$\delta_\square(U,W) \le \delta_\square(U,H) + \delta_\square(H,H') + \delta_\square(H',W) \le \frac{\varepsilon}{3} + \frac{\varepsilon}{4} + \frac{5\varepsilon}{12} = \varepsilon,$$

which completes the proof of Theorem 4.9.

4.5. *From distances to free energies and back.* In this section we prove the implications (ii)⇒(v)⇒(iv), which will complete the proof of Theorem 3.5. Again, we prove two (simple) quantitative versions.

THEOREM 4.13. *Let $q \geq 1$, let $\mathbf{a} \in \mathsf{Pd}_q$, let $J \in \mathsf{Sym}_q$, let $h \in \mathbb{R}^q$, and let $U, W \in \mathcal{W}$. Then*

$$\Big|\mathcal{F}_{\mathbf{a}}(U,J) - \mathcal{F}_{\mathbf{a}}(W,J)\Big| \leq \|J\|_1 \delta_\square(U,W) \tag{4.28}$$

and

$$\Big|\mathcal{F}(U,J,h) - \mathcal{F}(W,J,h)\Big| \leq \|J\|_1 \delta_\square(U,W). \tag{4.29}$$

Proof. Since the left-hand side of the above bounds does not change if we replace U by U^ϕ for a measure preserving bijection $\phi : [0,1] \to [0,1]$, it is enough to prove the lemma with a bound in terms of $\|U - W\|_\square$ instead of $\delta_\square(U,W)$. Let $\rho = (\rho_1, \dots, \rho_q)$ be a fractional partition of $[0,1]$. Recall the definition (3.7) of $\mathcal{E}_\rho(W,J,h)$. Using the fact that the cut-norm $\|\cdot\|_\square$ can be rewritten as

$$\|W\|_\square = \sup_{f,g\colon\ [0,1]\to[0,1]} \left| \int W(x,y) f(x) g(y)\, dx\, dy \right|, \tag{4.30}$$

where the suprema go over measurable and functions $f, g :\ [0,1] \to [0,1]$, we then have

$$\Big|\mathcal{E}_\rho(U,J,h) - \mathcal{E}_\rho(W,J,h)\Big| \leq \|J\|_1 \|U - W\|_\square.$$

Recalling the definitions (3.12) and (3.13), this completes the proof. □

THEOREM 4.14. *Let $q \geq 1$, $\mathbf{a} \in \mathsf{Pd}_q$, $J \in \mathsf{Sym}_q$, and let $U, W \in \mathcal{W}$. Let $\varepsilon > 0$ and $c = (2\ln q)/\varepsilon$. Then*

$$\Big|\mathcal{E}_{\mathbf{a}}(W,J) - \mathcal{E}_{\mathbf{a}}(U,J)\Big| \leq \frac{1}{c}\Big|\mathcal{F}_{\mathbf{a}}(W,cJ) - \mathcal{F}_{\mathbf{a}}(U,cJ)\Big| + \varepsilon.$$

Proof. Using the fact that $\mathsf{Ent}(\rho) \leq \ln q$, we get by (3.9) and (3.13) that

$$\Big|\mathcal{E}_{\mathbf{a}}(W,J) - \mathcal{F}_{\mathbf{a}}(W,J)\Big| \leq \ln q,$$

and similarly for U. Hence

$$\begin{aligned}\Big|\mathcal{E}_{\mathbf{a}}(W,J) - \mathcal{E}_{\mathbf{a}}(U,J)\Big| &= \frac{1}{c}\Big|\mathcal{E}_{\mathbf{a}}(W,cJ) - \mathcal{E}_{\mathbf{a}}(U,cJ)\Big| \\ &\leq \frac{1}{c}\Big(\Big|\mathcal{F}_{\mathbf{a}}(W,cJ) - \mathcal{F}_{\mathbf{a}}(U,cJ)\Big| + 2\ln q\Big),\end{aligned}$$

which proves Theorem 4.14 and thereby also completes the proof of Theorem 3.5. □

5. **Graphs vs. graphons**

We will use the results of the last section to prove Theorem 2.14 and Theorem 2.15(i)–(iii). Indeed, if we have a sequence of graphs (G_n), we can consider the sequence of associated graphons W_{G_n} and apply Theorem 3.5 to that sequence. The main technical issue here will be to relate parameters like $t(F, G)$, $\widehat{\mathcal{E}}_{\mathbf{a}}(G, J)$, and $\widehat{\mathcal{F}}_{\mathbf{a}}(G, J)$ to the corresponding parameters $t(F, W_G)$, $\mathcal{E}_{\mathbf{a}}(W_G, J)$, and $\mathcal{F}_{\mathbf{a}}(W_G, J)$ of the associated graphon. In some cases, this relationship is trivial:

$$t(F, G) = t(F, W_G) \tag{5.1}$$

for any two graphs F and G; but the corresponding relations for the ground state energies and free energies hold only asymptotically. A related technical issue will be the relationship between fractional and integral partitions, which will be more complicated than for graphons. (Compare, e.g., Lemma 4.3 and Theorem 5.4.)

5.1. *Fractional partitions and quotients.* Recall the definition of quotient graphs from Section 2.4. We will often consider $\widehat{\mathcal{S}}_q(G)$ as a subset of $\mathbb{R}^{q+q^2}$, denoting its elements H as $(\mathbf{a}, X)$, with $X = \beta(H) \in \mathsf{Sym}_q$ and $\mathbf{a} = \alpha(H) \in \mathsf{Pd}_q$. Given a vector $\mathbf{a} \in \mathsf{Pd}_q$, we finally introduce the set $\widehat{\mathcal{B}}_{\mathbf{a}}(G)$ of all weighted adjacency matrices of all quotients of G with nodeweights $\mathbf{a}$, $\widehat{\mathcal{B}}_{\mathbf{a}}(G) = \{X \in \mathsf{Sym}_q \colon (\mathbf{a}, X) \in \widehat{\mathcal{S}}_q(G)\}$.

For a finite graph G, the set $\widehat{\mathcal{S}}_q(G)$ is typically a very large finite set, which makes it difficult to work with. It will be convenient to introduce a fractional version of quotients. First, a *fractional partition* of a set V into q classes (briefly, a fractional q-partition) is a q-tuple of functions $\rho_1, \dots, \rho_q : V \to [0, 1]$ such that for all $x \in V$, we have $\rho_1(x) + \cdots + \rho_q(x) = 1$.

Let G be a weighted graph. For every fractional partition $\rho = (\rho_1, \dots, \rho_q)$ of $V(G)$, we define the *fractional quotient* G/ρ as the weighted graph with nodeweights

$$\alpha_i(G/\rho) = \sum_{u \in V(G)} \frac{\alpha_u(G)}{\alpha_G} \rho_i(u)$$

and edges weights

$$\beta_{ij}(G/\rho) = \frac{1}{\alpha_i(G/\rho)\alpha_j(G/\rho)} \sum_{u,v \in V(G)} \frac{\alpha_u(G)\alpha_v(G)\beta_{uv}(G)}{\alpha_G^2} \rho_i(u)\rho_j(v);$$

compare to the expressions (3.6) for graphons. To distinguish the fractional quotients from the quotients introduced in Section 2.4, we will often call the latter integer quotients. Note that the above definition reduces to the definition (2.24) if $\rho_i(x)$ is the indicator function of the event that $x \in V_i$. Note also that neither the integer quotients nor the fractional quotients of a graph G change if we rescale all nodeweights of G by a constant factor.

We call a graph H a fractional q-quotient of G if $H = G/\rho$ for some fractional q-partition of $V(G)$, and we denote the set of all fractional q-quotients of G by $\mathcal{S}_q(G)$. Finally, we define the fractional analogue of the set $\widehat{\mathcal{B}}_{\mathbf{a}}(G)$ as $\mathcal{B}_{\mathbf{a}}(G) = \{X \in \mathsf{Sym}_q \colon (\mathbf{a}, X) \in \mathcal{S}_q(G)\}$.

It follows from Lemma 4.2 and Proposition 5.3 that $\mathcal{S}_q(G)$ is a closed set, and it is not hard to see that $\mathcal{S}_q(G)$ is connected, but in general it is not convex (see Example 5.2). Obviously, $\mathcal{S}_q(G)$ contains $\widehat{\mathcal{S}}_q(G)$, but it is not its closure in general (since the latter is a finite set). We will come back to how well $\widehat{\mathcal{S}}_q(G)$ approximates $\mathcal{S}_q(G)$ in Lemma 5.4. Most of the time, we will work with the fractional versions, which are much easier to handle.

We can use fractional partitions to define fractional versions of ground state energy by replacing the partitions in the definition by fractional partitions. For every fractional partition ρ of $V(G)$, define

$$\begin{aligned}(5.2)\qquad \mathcal{E}_\rho(G,J,h) = &- \sum_{u\in V(G)} \sum_i h_i \frac{\alpha_u(G)}{\alpha_G}\rho_i(u) \\ &- \sum_{u,v\in V(G)} \sum_{i,j} \frac{\alpha_u(G)\alpha_v(G)}{\alpha_G^2}\beta_{uv}(G)\rho_i(u)\rho_j(v)J_{ij}.\end{aligned}$$

If ρ is a proper partition corresponding to a map $\phi : V(G) \to [q]$, then $\mathcal{E}_\rho(G,J,h) = \mathcal{E}_\phi(G,J,h)$. Using this notation, we can define

$$(5.3)\qquad \mathcal{E}(G,J,h) = -\max_\rho \mathcal{E}_\rho(G,J,h) = -\max_{(\mathbf{a},X)\in\mathcal{S}_q(G)} \Big(\langle X,J\rangle_{\mathbf{a}} + \langle \mathbf{a},h\rangle\Big)$$

and

$$(5.4)\qquad \mathcal{E}_{\mathbf{a}}(G,J) = -\max_{\rho:\alpha(\rho)=\mathbf{a}} \mathcal{E}_\rho(G,J,0) = -\max_{X\in\mathcal{B}_a(G)} \langle X,J\rangle_{\mathbf{a}}.$$

We will come back to how well these fractional versions approximate the "real" versions in Section 5.3.

We conclude with a couple of examples illustrating the set of quotients and its complexity. In particular, we see that $\mathcal{S}_q$ is not convex in general.

Example 5.1. Let $K_1(p)$ be a single node with a loop with weight p. For every fractional q-quotient H of $K_1(p)$, we have $\beta(H) \equiv p$, and so $\mathcal{B}_{\mathbf{a}}(K_1(p))$ consists of a single $q \times q$ matrix with constant entry p, no matter what value we choose for $\mathbf{a} \in \mathsf{Pd}_q$.

Example 5.2. This example gives a weighted graph G for which $\mathcal{S}_q(G)$ is not convex. Let $L_2(p)$ be the two-node graph with a loop with weight p at each node (and no other edge). Let ρ be a fractional q-partition of $V(L_2(p)) = \{u,v\}$, and let H denote the corresponding quotient. Then

$$\alpha_i(H) = \frac{1}{2}(\rho_i(u)+\rho_i(v)) \text{ and } \beta_{ij}(H)) = \frac{(\frac12)^2\rho_i(u)\rho_j(u)p + (\frac12)^2\rho_i(v)\rho_j(v)p}{\alpha_i(H)\alpha_j(H)}.$$

For $q = 2$, the fractional partition ρ can be expressed by two parameters x, y: $\rho_1(u) = x$, $\rho_2(u) = 1 - x$, $\rho_1(v) = y$, $\rho_2(v) = 1 - y$, which reduces to one parameter, say the parameter x, if we fix $\alpha(H)) = \mathbf{a}$ for some $\mathbf{a} \in \mathsf{Pd}_2$. The edgeweights $\beta(H)$ can then be expressed as a quadratic function in x, giving that $\mathcal{B}_{\mathbf{a}}(L_2(p))$ is a nonconvex function in the parameter x in the space of 2×2 matrices. Then of course $\mathcal{S}_2(L_2(p))$ is not convex either.

5.2. *Quotients of graphs and graphons.* We start by noting the following simple fact.

PROPOSITION 5.3. *For every weighted graph G and every $q \geq 1$,*

$$\widehat{\mathcal{S}}_q(G) \subseteq \mathcal{S}_q(G) = \mathcal{S}_q(W_G) = \widehat{\mathcal{S}}_q(W_G).$$

Proof. It is obvious that $\widehat{\mathcal{S}}_q(G) \subseteq \mathcal{S}_q(G)$ and $\widehat{\mathcal{S}}_q(W_G) \subseteq \mathcal{S}_q(W_G)$, so we only have to show that $\mathcal{S}_q(G) \subseteq \widehat{\mathcal{S}}_q(W_G)$ and $\mathcal{S}_q(W_G) \subseteq \mathcal{S}_q(G)$. Every fractional q-partition ρ of $V(G)$ gives a (nonfractional) q-partition $(S_1, \dots, S_q)$ of $[0, 1]$ as follows. Partition the interval I_v corresponding to $v \in V(G)$ into q intervals $I_{v1}, \dots, I_{vq}$ of lengths $\rho_1(v)\alpha_v(G)/\alpha_G, \dots, \rho_q(v)\alpha_v(G)/\alpha_G$, respectively, and define $S_i = \cup_{v \in V(G)} I_{vi}$. It is straightforward to check that $G/\rho = (W_G)/\mathcal{P}$, and hence $\mathcal{S}_q(G) \subseteq \widehat{\mathcal{S}}_q(W_G)$. Finally, every fractional partition ρ of $[0, 1]$ defines a fractional partition $\bar{\rho}$ of $V(G)$ by

$$\bar{\rho}_i(v) = \int_{I_v} \rho_i(x)\, dx.$$

Again, it is easy to check that $G/\bar{\rho} = W/\rho$. This proves that $\mathcal{S}_q(W_G) \subseteq \mathcal{S}_q(G)$ and completes the proof of the proposition. $\square$

The following technical lemma asserts that by restricting our attention to integral partitions we do not lose too much, provided the graph has no dominating nodeweights.

THEOREM 5.4. *For every weighted graph G and every $q \geq 1$,*

$$d_1^{\mathrm{Hf}}(\widehat{\mathcal{S}}_q(G), \mathcal{S}_q(G)) \leq q\sqrt{\frac{\alpha_{\max}(G)}{\alpha_G}}\big(1 + 4\beta_\infty(G)\big).$$

Proof. Let $c = \alpha_{\max}(G)/\alpha_G$. We have to show that for every $H \in \mathcal{S}_q(G)$ there is an $H' \in \widehat{\mathcal{S}}_q(G)$ such that

$$d_1(H, H') \leq q\sqrt{c}\big(1 + 4\beta_\infty(G)\big).$$

Since quotients and fractional quotients do not change if we rescale the weights of G, we may assume that $\alpha_G = 1$.

Let $a_i = \alpha_i(H)$ and $B_{ij} = a_i a_j \beta_{ij}(H)$, and let $\rho = (\rho_1, \dots, \rho_q)$ be a fractional partition of $V(G)$ such that $H = G/\rho$. In other words, let ρ be such that

$$\sum_{u\in V(G)} \alpha_u(G)\rho_i(u) = a_i \qquad \text{and} \qquad \alpha_u(G)\alpha_v(G)\beta_{uv}(G)\rho_i(u)\rho_j(v) = B_{ij}.$$

Let $\mathcal{P} = (V_1, \dots, V_q)$ be a random partition of $V(G)$ obtained by "rounding" ρ as follows. For every $u \in V(G)$, we draw a random index i from the probability distribution $(\rho_1(u), \dots, \rho_q(u))$, and we put u in V_i. Let $H' = G/\mathcal{P}$, and set $a_i' = \alpha_i(H')$ and $B_{ij}' = a_i'a_j'\beta_{ij}(H')$.

We use a standard (though somewhat lengthy) second moment argument to show that with large probability, a' is close to a and B' is close to B. Let X_{iu} be the indicator variable that we put u in V_i. Clearly $\mathsf{E}(X_{iu}) = \rho_i(u)$. Using that X_{iu} and X_{ju} are independent if $u \neq v$,

$$\begin{aligned}\mathsf{E}((a_i' - a_i)^2) &= \sum_{u\neq v} \alpha_u(G)\alpha_v(G)\mathsf{E}((X_{iu} - \rho_i(u))(X_{iv} - \rho_i(v))) \\ &\quad + \sum_u \alpha_u(G)^2\mathsf{E}((X_{iu} - \rho_i(u))^2) \\ &= \sum_{u\in V(G)} \alpha_u(G)^2(\rho_i(u) - \rho_i(u)^2) \le c \sum_{u\in V(G)} \alpha_u(G)\rho_i(u) = ca_i,\end{aligned}$$

and summing over all i, we get

$$\mathsf{E}(\|\mathbf{a} - \mathbf{a}'\|_2^2) \le c. \tag{5.5}$$

The argument for B is similar but more involved. Let us assume for the moment that $|\beta_{uv}(G)| \le 1$. Writing $B_{ij}' - B_{ij}$ as

$$B_{ij}' - B_{ij} = \sum_{u,v\in V(G)} \alpha_u(G)\alpha_v(G)\beta_{uv}(G)(X_{iu}X_{jv} - \rho_i(u)\rho_j(v))$$

and introducing the shorthand α_u for $\alpha_u(G)$, we bound

(5.6)

$$\begin{aligned}\mathsf{E}((B_{ij}' - B_{ij})^2) &= \sum_{u_1,v_1,u_2,v_2} \alpha_{u_1}\alpha_{u_2}\alpha_{v_1}\alpha_{v_2}\beta_{u_1v_1}(G)\beta_{u_2v_2}(G) \\ &\qquad \times \mathsf{E}\Big(\big(X_{iu_1}X_{jv_1} - \rho_i(u_1)\rho_j(v_1)\big)\big(X_{iu_2}X_{jv_2} - \rho_i(u_2)\rho_j(v_2))\big)\Big) \\ &\le \sum_{u_1,v_1,u_2,v_2} \alpha_{u_1}\alpha_{u_2}\alpha_{v_1}\alpha_{v_2}\Big|\mathsf{E}\Big(\big(X_{iu_1}X_{jv_1} - \rho_i(u_1)\rho_j(v_1)\big) \\ &\qquad\qquad \times \big(X_{iu_2}X_{jv_2} - \rho_i(u_2)\rho_j(v_2))\big)\Big)\Big|,\end{aligned}$$

where the sum goes over nodes $u_1, v_1, u_2, v_2 \in V(G)$. Consider any term above:

$$\mathsf{E}\Big(\big(X_{iu_1}X_{jv_1} - \rho_i(u_1)\rho_j(v_1)\big)\big(X_{iu_2}X_{jv_2} - \rho_i(u_2)\rho_j(v_2))\big)\Big). \tag{5.7}$$

If u_1, u_2, v_1, v_2 are all different, then $X_{iu_1}, X_{jv_1}, X_{iu_2}, X_{jv_2}$ are independent, and hence this expectation is 0.

Next, suppose that there is one coincidence. If this coincidence is $u_1 = v_1$, we can use the independence of the three random variables X_{iu_2}, X_{jv_2}, and $X_{iu_1}X_{jv_1}$ to conclude that this gives again no contribution, and similarly for $u_2 = v_2$. Consider one of the other four coincidences, say $u_1 = u_2$. Then the expectation in (5.7) is $\rho_i(u_1)\rho_j(v_1)\rho_j(v_2)$, and the contribution of these terms to the sum in (5.6) is bounded by

$$\sum_{u_1,v_1,v_2} \alpha_{u_1}^2 \alpha_{v_1}\alpha_{v_2}\rho_i(u_1)\rho_j(v_1)\rho_j(v_2) \le ca_ia_j.$$

There are four similar terms, so the total is bounded by $4ca_ia_j$.

In the case of two coincidences, we have either $u_1 = u_2$ and $v_1 = v_2$ or $u_1 = v_2$ and $v_1 = u_2$ or $v_1 = u_1$ and $v_2 = u_2$. Consider the case $u_1 = u_2 = u$, $v_1 = v_2 = v \neq u$. The expectation in (5.7) is then $\rho_i(u)\rho_j(v)\big(1 - \rho_i(u)\rho_j(v)\big)$. The contribution of these terms to the sum in (5.6) is at most

$$\sum_{u,v} \alpha_u^2\alpha_v^2\rho_i(u)\rho_j(v) \le ca_ia_j.$$

The two other cases are similar, giving a total of at most $3ca_ia_j$.

For three coincidences, there are four cases, which all are similar. Taking, e.g., the case $u_1 = u_2 = v_1 = u$ and $v_2 = v \neq u$, we get $\rho_i(u)\rho_j(v)\delta_{ij}(1 - \rho_i(u))$. The sum of these terms over u and v gives a contribution that is at most

$$\delta_{ij}\sum_{u,v} \alpha_u^3\alpha_v\rho_i(u)\rho_j(v) \le ca_ia_j.$$

The other three terms are similar, giving a total contribution of $4ca_ia_j$.

We are left with the case of four coincidences, $u_1 = u_2 = v_1 = v_2 = u$, which gives an expectation of $\rho_i(u)\delta_{ij} - 2\rho_i(u)^3\delta_{ij} + \rho_i(u)^2\rho_j(u)^2$, and a total contribution of at most

$$\sum_u \alpha_u^4\Big(\rho_i(u)\delta_{ij} + \rho_i(u)\rho_j(u)\Big) \le ca_i\delta_{ij} + ca_ia_j.$$

To sum up, we get that

$$\mathsf{E}((B_{ij} - B'_{ij})^2) \le 12ca_ia_j + ca_i\delta_{ij}$$

whenever $\beta_\infty(G) \le 1$. Rescaling the edgeweights of G to remove the condition $\beta_\infty(G) \le 1$, this gives

$$\mathsf{E}(\|B - B'\|_2^2) = \mathsf{E}\Big(\sum_{i,j}(B_{ij} - B'_{ij})^2\Big) \le 13c(\beta_\infty(G))^2.$$

Combined with (5.5) and Cauchy-Schwarz, this gives

$$\begin{aligned}\mathsf{E}\Big(d_1(H,H')\Big) &= q\mathsf{E}\Big(\frac{1}{q}\sum_{i=1}^{q}|a_i-a_i'|\Big)+q^2\mathsf{E}\Big(\frac{1}{q^2}\sum_{i,j=1}^{q}|B_{ij}-B_{ij}'|\Big)\\ &\le q\Big(\mathsf{E}\Big(\frac{1}{q}\sum_{i=1}^{q}|a_i-a_i'|^2\Big)\Big)^{1/2}+q^2\Big(\mathsf{E}\Big(\frac{1}{q^2}\sum_{i,j=1}^{q}|B_{ij}-B_{ij}'|^2\Big)\Big)^{1/2}\\ &\le \sqrt{qc}+q\sqrt{13c}\beta_\infty(G)\le q\sqrt{c}\big(1+4\beta_\infty(G)\big).\end{aligned}$$

Hence with positive probability, $d_1(H,H')\le q\sqrt{c}\big(1+4\beta_\infty(G)\big)$, as required. □

5.3. *Ground state energies of graphs and graphons.* We start with the remark that Proposition 5.3 and equation (5.4) imply that

$$\mathcal{E}_{\mathbf{a}}(G,J)=\mathcal{E}_{\mathbf{a}}(W_G,J). \tag{5.8}$$

The next theorem relates this common value to the microcanonical ground state energy $\widehat{\mathcal{E}}_{\mathbf{a}}$ introduced in Section 2.4.

THEOREM 5.5. *Let G be a weighted graph, and let $q\ge 1$, $\mathbf{a}\in\mathsf{Pd}_q$ and $J\in\mathsf{Sym}_q$. Then*

$$|\widehat{\mathcal{E}}_{\mathbf{a}}(G,J)-\mathcal{E}_{\mathbf{a}}(G,J)|\le 6q^3\frac{\alpha_{\max}(G)}{\alpha_G}\beta_{\max}(G)\,\|J\|_\infty. \tag{5.9}$$

First we show that the fractional version of $\widehat{\mathcal{E}}(G,J,h)$ does not carry new information, at least if we restrict ourselves to weighted graphs without loops.

PROPOSITION 5.6. *Let $q\ge 1$, $J\in\mathsf{Sym}_q$ and $h\in\mathbb{R}^q$. If G is a weighted graph with $\beta_{xx}(G)=0$ for all $x\in V(G)$, then*

$$\widehat{\mathcal{E}}(G,J,h)=\mathcal{E}(G,J,h). \tag{5.10}$$

In the more general case where $\beta_{uu}(G)$ is arbitrary, we have

$$\Big|\widehat{\mathcal{E}}(G,J,h)-\mathcal{E}(G,J,h)\Big|\le 2\frac{\alpha_{\max}}{\alpha_G}\beta_{\max}(G)\|J\|_\infty. \tag{5.11}$$

To prove these results, we need some preparation.

5.3.1. *Preliminaries.* Let ρ and ρ' be two fractional partitions of $[0,1]$. We define the distance

$$d_1(\rho,\rho')=\frac{1}{q}\sum_{i=1}^{q}\int_{[0,1]}|\rho_i(x)-\rho_i'(x)|dx \tag{5.12}$$

on fractional q-partitions. For a weighted graph G and two fractional q-partitions of $V(G)$, we define

$$d_{1,G}(\rho,\rho') = \frac{1}{q}\sum_{i=1}^{q}\sum_{v\in V(G)}\frac{\alpha_v(G)}{\alpha_G}|\rho_i(v)-\rho_i'(v)|. \tag{5.13}$$

(If G has nodeweights one, we often leave out the subscript of G and denote this distance by $d_1(\rho,\rho')$ as well.)

The following inequalities are immediate consequences of the definitions (3.7), (5.2), and (2.25). Let $W\in\mathcal{W}$, let G be a weighted graph, and let $q\geq 1$, $J\in\mathsf{Sym}_q$, and $h\in\mathbb{R}^q$. If ρ,ρ' are fractional q-partitions of $[0,1]$, then

$$|\mathcal{E}_\rho(W,J,h)-\mathcal{E}_{\rho'}(W,J,h)|\leq q(2\|J\|_\infty\|W\|_\infty+\|h\|_\infty)\,d_1(\rho,\rho'). \tag{5.14}$$

If ρ,ρ' are fractional q-partitions of $V(G)$, then

$$|\mathcal{E}_\rho(G,J,h)-\mathcal{E}_{\rho'}(G,J,h)|\leq q(2\|J\|_\infty\beta_{\max}(G)+\|h\|_\infty)\,d_{1,G}(\rho,\rho'). \tag{5.15}$$

If G' is a weighted graph on the same nodeset as G, then

$$\Big|\mathcal{E}_\phi(G,J,h)-\mathcal{E}_\phi(G',J,h)\Big|\leq\max\{\|h\|_\infty,q^2\|J\|_\infty\}d_\square(G,G'), \tag{5.16}$$

and if G and G' also have the same nodeweights, then

$$\Big|\mathcal{E}_\phi(G,J,h)-\mathcal{E}_\phi(G',J,h)\Big|\leq q^2\|J\|_\infty d_\square(G,G'). \tag{5.17}$$

5.3.2. *Proof of Theorem* 5.5. Without loss of generality, we may assume that $\alpha_G=1$ and $\beta_{\max}(G)=1$. First we prove that

$$\mathcal{E}_{\mathbf{a}}(G,J)\leq\widehat{\mathcal{E}}_{\mathbf{a}}(G,J)+2q\|J\|_\infty\alpha_{\max}(G). \tag{5.18}$$

Rewrite the microcanonical ground state energy as

$$\widehat{\mathcal{E}}_{\mathbf{a}}(G,J)=-\max_{\phi\in\Omega_{\mathbf{a}}(G)}\sum_{u,v\in V(G)}\frac{\alpha_u(G)\alpha_v(G)\beta_{uv}(G)}{\alpha_G^2}J_{\phi(u)\phi(v)}, \tag{5.19}$$

let $\phi: V(G)\to[q]$ be a map attaining the optimum on the right-hand side, and let ρ be the corresponding partition of $V(G)$, considered as a fractional partition. Then $\widehat{\mathcal{E}}_{\mathbf{a}}(G,J)=\mathcal{E}_\rho(G,J,0)$ and $|\alpha_i(\rho)-a_i|\leq\alpha_{\max}(G)$. It is now easy construct another fractional partition ρ' with $\alpha_i(\rho')=a_i$ and $d_{1,G}(\rho,\rho')\leq\alpha_{\max}(G)$. Invoking (5.15), the inequality (5.18) follows.

The main part of the proof is to show that

$$\widehat{\mathcal{E}}_{\mathbf{a}}(G,J)\leq\mathcal{E}_{\mathbf{a}}(G,J)+6q^3\alpha_{\max}\,\|J\|_\infty. \tag{5.20}$$

For a given fractional partition ρ of $V(G)$ with $\alpha(\rho)=\mathbf{a}$, call a node v *bad* if $\rho(v)$ is not a 0-1 vector. Suppose that there are at least $q+1$ bad nodes, and let S be any set of $q+1$ bad nodes. For a bad node v, the vector $\rho(v)$ has at least two fractional entries, so the selected nodes have at least $2q+2$ fractional entries. If we fix the sums $\sum_{i=1}^q\rho_i(v)$ for $v\in S$ and $\sum_{v\in S}\rho_i(v)$ for $i=1,\dots,q$, we have fixed $2q+1$ sums, so there is a family of solutions with dimension at least 1. That is to say, we have an affine family $\rho_{t,i}(v) = \rho_i(v) + tr_i(v)$ of "deformations" of ρ such that ρ_t is a fractional partition of $V(G)$ for every t, $\alpha(\rho_t) = \mathbf{a}$, and $r_i(v) = 0$ unless $v \in S$ and $0 < \rho_i(v) < 1$.

Let X and X_t be such that $G/\rho = (\mathbf{a}, X)$ and $G/\rho_t = (\mathbf{a}, X_t)$. Then

$$\langle J, X_t\rangle_{\mathbf{a}} = \langle J, X\rangle_{\mathbf{a}} + C_1 t + C_2 t^2,$$

where

$$C_1 = 2 \sum_{\substack{u\in S\\ v\in V(G)}} \alpha_u(G)\alpha_v(G)\beta_{uv}(G) \sum_{i,j=1}^{q} J_{ij} r_j(u)\rho_i(v)$$

and

$$C_2 = \sum_{u,v\in S} \alpha_u(G)\alpha_v(G)\beta_{uv}(G) \sum_{i,j=1}^{q} J_{ij} r_i(u) r_j(v).$$

Choosing the sign of t so that $C_1 t \geq 0$, we increase the absolute value of t until there is at least one new pair (v,i) for which $\rho_{t,i}(v)$ is 0 or 1, while we still have $\rho_t \geq 0$. Starting with an optimal fractional partition, we repeat this operation until we are left with a set R of at most q bad nodes. Then we replace the resulting fractional partition $\tilde\rho$ on R by any integer partition $(V_1,\dots,V_q)$ obeying the condition

$$\Big|\sum_{u\in R} \tilde\rho_i(u)\alpha_u(G) - \sum_{u\in V_i} \alpha_u(G)\Big| \leq \alpha_{\max}(G).$$

How much do these operations decrease the value $\langle J, X\rangle_{\mathbf{a}}$? Replacing ρ by ρ_t, we lose at most $C_2 t^2$. Since for every u, $(\rho_1(u)+tr_1(u),\dots,\rho_q(u)+tr_q(u))$ is still a fractional partition, we have $\sum_i r_i(u)t = 0$ and $0 \leq \rho_i(u) + r_i(u)t \leq 1$ implying, in particular, that $\sum_i |r_i(u)t| = 2\sum_i |r_i(u)t| \mathbf{1}_{r_i(u)t<0} \leq 2\sum_i \rho_i(u) \leq 2$. Hence

$$\begin{aligned}|C_2 t^2| &\leq \|J\|_\infty \sum_{u,v\in S} \alpha_u(G)\alpha_v(G)|\beta_{uv}(G)| \sum_{i,j=1}^{q} |r_i(u)t|\cdot|r_j(v)t| \\ &\leq 4\|J\|_\infty \sum_{u,v\in S} \alpha_u(G)\alpha_v(G) = 4\|J\|_\infty \alpha_{G[S]}^2 \leq 4\|J\|_\infty (q+1)\alpha_{\max}\alpha_{G[S]}.\end{aligned}$$

Thus the cost of replacing one fractional entry in S by an integer entry is not more than $4\|J\|_\infty(q+1)\alpha_{\max}\alpha_{G[S]}$. To estimate the total cost of reducing the set of bad nodes to a set R of at most q nodes, we formulate the following game. There are n items of prices $\alpha_1 \geq \cdots \geq \alpha_n$, which sum to 1, and there are $q-1$ copies of each. At each step, you select $q+1$ different items and pay the total price; then your adversary points at q of them, which you have to give back without compensation. The game stops when there are at most q different items left. Your goal is to minimize your total payment. How much do you have to pay, if both you and your adversary play optimally?

Let us follow the simple greedy strategy of selecting the $q+1$ cheapest items each time. It is easy to argue that the best strategy for the adversary is to take away all but the cheapest of these $q+1$ items at each time. Then you pay $(\alpha_n + \alpha_{n-1} + \cdots + \alpha_{n-q+1})$ $q-1$ times, $(\alpha_{n-1} + \alpha_{n-2} + \cdots + \alpha_{n-q})$ $q-1$ times, etc. In total you pay for every item at most $(q-1)(q+1) < q^2-1$ times, and so your total cost is less than q^2-1, leading to a decrease in the value of $\langle J, X\rangle_{\mathbf{a}}$ that is less than $4\|J\|_\infty (q+1)(q^2-1)\alpha_{\max}$.

To estimate the cost to convert the fractional partition $\tilde{\rho}$ on R to an ordinary partition $\mathcal{P} = (V_1, \dots, V_q)$, we bound the difference

$$\sum_{u,v\in V(G)} \alpha_u(G)\alpha_v(G)\beta_{uv}(G) \sum_{i,j=1}^{q} J_{ij}\Big(\mathbf{1}_{u\in V_i}\mathbf{1}_{v\in V_j} - \tilde{\rho}_i(u)\tilde{\rho}_j(v)\Big)$$

$$= \sum_{\substack{u\in V(G)\\ v\in R}} \alpha_u(G)\alpha_v(G)\beta_{uv}(G) \sum_{i,j=1}^{q} J_{ij}\Big(\tilde{\rho}_i(u) + \mathbf{1}_{u\in V_i}\Big)\Big(\mathbf{1}_{v\in V_j} - \tilde{\rho}_j(v)\Big)$$

by $4\|J\|_\infty \alpha_{G[R]} \le 4q\alpha_{\max}\|J\|_\infty$, leading to an overall bound of

$$4\|J\|_\infty\Big((q+1)(q^2-1)+q\Big)\alpha_{\max} \le 6q^3\alpha_{\max}\|J\|_\infty.$$

This concludes the proof of (5.20). □

5.3.3. *Proof of Proposition* 5.6. We first prove the identity (5.10). Rewriting both $\mathcal{E}(G,J,H)$ and $\widehat{\mathcal{E}}(G,J,h)$ in terms of factors, this amounts to showing that

$$\max_{(\mathbf{a},X)\in\widehat{\mathcal{S}}_q(G)}\Big(\langle X,J\rangle_{\mathbf{a}} + \langle \mathbf{a}, h\rangle\Big) = \max_{(\mathbf{a},X)\in\mathcal{S}_q(G)}\Big(\langle X,J\rangle_{\mathbf{a}} + \langle \mathbf{a}, h\rangle\Big).$$

Let ρ be a fractional q-partition of G, and let $G/\rho = (\mathbf{a}, X)$. Assuming without loss of generality that $\alpha_G = 1$, we have the identity

$$\langle X,J\rangle_{\mathbf{a}} + \langle \mathbf{a}, h\rangle = \sum_{u,v\in V(G)} \alpha_u(G)\alpha_v(G)\beta_{uv}(G)\sum_{i,j=1}^{q} J_{ij}\rho_i(u)\rho_j(v) + \sum_{u\in V(G)}\alpha_u(G)\sum_{i=1}^{q} h_i\rho_i(u).$$

For a fixed $u \in V(G)$, this is a linear function of $(\rho_1(u), \dots, \rho_q(u))$ (here we use that $\beta_{uu}(G) = 0$), and so its maximum is attained at a vertex of the simplex Pd_q, i.e., a vector $(\rho_1(u), \dots, \rho_q(u))$ that is integer valued. Repeating this for every $u \in V(G)$, we see that the maximum over fractional partitions is attained for an ordinary partition, as desired.

To obtain the bound (5.11), we note that the error from removing the diagonal terms can be bounded by

$$2\sum_{u\in V(G)} \frac{(\alpha_u(G))^2|\beta_{uu}(G)|}{\alpha_G^2}\|J\|_\infty \le 2\frac{\alpha_{\max}(G)}{\alpha_G}\beta_{\max}(G)\|J\|_\infty. \qquad \square$$

5.4. *Graph homomorphisms and ground state energy of graphons.* The next lemma generalizes the bound in Example 2.3 and gives a quantitative version of the bound (2.16).

LEMMA 5.7. *Let G be an weighted graph on n nodes, and let H be a softcore weighted graph on q nodes, with weights $\alpha_i(H) = e^{h_i}$ and $\beta_{ij}(H) = e^{2J_{ij}}$. Then*

$$-\widehat{\mathcal{E}}(G,J,0)+\frac{\ln\alpha_{\min}(H)}{\alpha_G} \le \frac{\ln \hom(G,H)}{\alpha_G^2} \le -\widehat{\mathcal{E}}(G,J,0)+\frac{\ln\alpha_{\max(H)}}{\alpha_G}+\frac{n\log q}{\alpha_G^2}$$

and for every $\mathbf{a}\in\mathsf{Pd}_q$,

$$-\widehat{\mathcal{E}}_{\mathbf{a}}(G,J)+\frac{\ln\alpha_{\min}(H)}{\alpha_G} \le \frac{\ln \hom_{\mathbf{a}}(G,H)}{\alpha_G^2} \le -\widehat{\mathcal{E}}_{\mathbf{a}}(G,J)+\frac{\ln\alpha_{\max(H)}}{\alpha_G}+\frac{n\log q}{\alpha_G^2}.$$

Proof. We prove the first inequality; the proof of the second is similar. Write $\hom(G,H)$ as

$$\hom(G,H) = \sum_{\phi:V(G)\to V(H)} \alpha_\phi e^{-\alpha_G^2\mathcal{E}_\phi(G,J,0)},$$

where $\alpha_\phi = \prod_{i\in V(G)}\alpha_{\phi(i)}(H)^{\alpha_i(G)}$. Since, by definition, the minimum of $\mathcal{E}_\phi(G,J,0)$ is the ground state energy $\widehat{\mathcal{E}}(G,J,0)$, we have

$$\hom(G,H) \le \sum_\phi \alpha_\phi e^{-\alpha_G^2\widehat{\mathcal{E}}(G,J,0)} \le q^n\alpha_{\max}(H)^{\alpha_G}e^{-\alpha_G^2\widehat{\mathcal{E}}(G,J,0)},$$

and

$$\hom(G,H) \ge \max_\phi \alpha_\phi e^{-\alpha_G^2\mathcal{E}_\phi(G,J,0)} \ge \alpha_{\min}(H)^n e^{-\alpha_G^2\widehat{\mathcal{E}}(G,J)},$$

from which the lemma follows. □

5.5. *Free energies of graphs and graphons.* We now turn to the main theorem of this section, namely that the free energy of G is close to that of W_G.

THEOREM 5.8. *Let $q\ge 1$, $\mathbf{a}\in\mathsf{Pd}_q$, $h\in\mathbb{R}^q$, and $J\in\mathsf{Sym}_q$. Let G be a graph on n nodes with all nodeweights* 1. *Then*

$$\left|\widehat{\mathcal{F}}_{\mathbf{a}}(G,J)-\mathcal{F}_{\mathbf{a}}(W_G,J)\right| \le \frac{12q^2}{n^{1/4}}+\frac{65q^2}{\sqrt{\ln n}}\|J\|_\infty\beta_{\max}(G)$$

and

$$\left|\widehat{\mathcal{F}}(G,J,h)-\mathcal{F}(W_G,J,h)\right| \le \frac{12q^2}{n^{1/4}}+q^2\frac{65}{\sqrt{\ln n}}\|J\|_\infty\beta_{\max}(G)+\frac{5q^2}{n^{1/2}}\|h\|_\infty.$$

The proof of this inequality is more involved than the proof of the corresponding statement for ground state energies. The additional difficulties here are not just technical. They are related to the fact, noted earlier, that there is no natural way to define homomorphism numbers from graphons to finite graphs. Thus, while we could define approximations to the ground state energies $\mathcal{E}(W,J,h)$ and $\mathcal{E}_{\mathbf{a}}(W,J)$ that involved only integer partitions, it is not possible to do the same thing in the case of the free energies $\mathcal{F}(W, J, h)$ and $\mathcal{F}_{\mathbf{a}}(W, J)$ — since the entropy $\mathsf{Ent}(\rho)$ of an integer partition is zero. In other words, we will have to translate the information contained in the discrete sums defining $Z(G, J, h)$ and $Z_{\mathbf{a}}(G, J)$ into entropy information involving fractional partitions.

This is best explained in the case where the graph under consideration is a blow up $G[k]$ of a much smaller graph G. In this situation, there are large classes of configurations that have exactly the same energy density. Indeed, for $u \in V(G)$, let V_u be the set of nodes in $V(G[k])$ that are blow ups of u, and let $k_i(u)$ be the number of nodes in V_u that are mapped onto $i \in [q]$. Then all configurations $\phi : V(G[k]) \to [q]$ with given numbers $\{k_i(u)\}$ have the same energy. Counting how many such configurations we can find, we will get a term that eventually will lead to a term $\mathsf{Ent}(\rho)$ in an optimization problem. In a final step, we will use the Weak Regularity Lemma to approximate the graphs in a convergent sequence (G_n) by blow ups of a suitable sequence of smaller graphs.

5.5.1. *Entropies.* Recall the definition (3.11) of the entropy of a fractional partition ρ. If ρ is a fractional partition of a finite set V, this definition can be modified in the following way:

$$\mathsf{Ent}(\rho) = -\frac{1}{|V|} \sum_{i=1}^{q} \sum_{v \in V} \rho_i(v) \ln \rho_i(v) = \frac{1}{|V|} \sum_{v \in V} \mathsf{Ent}(\rho(v)).$$

Let ρ be a fractional q-partition of $[0, 1]$ and $\mathcal{P} = \{I_1, \dots, I_n\}$ be an equipartition of $[0, 1]$, i.e., a partition such that all classes of $\mathcal{P}$ have the same Lebesgue measure. We define the fractional partition $\rho_{\mathcal{P}}$ of $[0, 1]$ and the fractional partition $\rho/\mathcal{P}$ of $[n]$ as follows:

$$(\rho/\mathcal{P})_i(v) = \frac{1}{n} \int_{I_v} \rho_i(x)\, dx \qquad \text{and} \qquad (\rho_{\mathcal{P}})_i(y) = (\rho/\mathcal{P})_i(v) \qquad \text{if } y \in I_v.$$

PROPOSITION 5.9. *For every fractional q-partition ρ of $[0, 1]$ and every equipartition $\mathcal{P} = \{I_1, \dots, I_n\}$ of $[0, 1]$, we have*

$$\mathsf{Ent}(\rho) \le \mathsf{Ent}(\rho_{\mathcal{P}}) = \mathsf{Ent}(\rho/\mathcal{P}).$$

Proof. The equality of $\mathsf{Ent}(\rho_{\mathcal{P}})$ and $\mathsf{Ent}(\rho/\mathcal{P})$ is straightforward. The function $\mathsf{Ent}(\mathbf{x}) = -\sum_{i=1}^{q} x_i \ln x_i$ is concave for $\mathbf{x} \in \mathsf{Pd}_q$, so the inequality follows by Jensen's inequality. □

As a consequence, we have the following finite formula for the free energy of the graphon W_G associated with a graph G with nodeweights 1:

$$\mathcal{F}_{\mathbf{a}}(W_G) = \inf_{\rho} \Big(\mathcal{E}_\rho(G, J, 0) - \mathsf{Ent}(\rho) \Big), \tag{5.21}$$

where ρ ranges over all fractional partitions of $V(G)$ with $\alpha(\rho) = \mathbf{a}$.

Together with (5.15), the next lemma shows that the quantities on the right-hand side of (5.21) are continuous functions of ρ.

LEMMA 5.10. *Let ρ, ρ' be fractional q-partitions (of a finite set or of $[0,1]$). If $d_1(\rho, \rho') \le 1/e$, then*

$$|\mathsf{Ent}(\rho) - \mathsf{Ent}(\rho')| \le q d_1(\rho, \rho') \ln \frac{1}{d_1(\rho, \rho')}.$$

Proof. We do the proof for fractional partitions of $[0,1]$. Define $f : [0,1] \to \mathbb{R}$ by $f(x) = -x \ln x$. As a consequence of the concavity of f, we have that

$$|f(x) - f(y)| \le \max\{f(|x-y|), f(1-|x-y|\} \le g(|x-y|),$$

where $g(x)$ is the concave hull of $\max\{f(x), f(1-x)\}$,

$$g(x) = \begin{cases} f(x) & \text{if} \quad x \in [0, 1/e], \\ 1/e & \text{if} \quad x \in (1/e, 1-1/e), \\ f(1-x) & \text{if} \quad x \in [1-1/e, 1]. \end{cases}$$

By Hölder's inequality, we thus have

$$\begin{aligned} \frac{1}{q}|\mathsf{Ent}(\rho) - \mathsf{Ent}(\rho')| &= \Big|\frac{1}{q}\sum_{i=1}^{q} \int_{[0,1]} \Big(f(\rho_i(x)) - f(\rho_i'(x))\Big) dx\Big| \\ &\le \frac{1}{q}\sum_{i=1}^{q} \int_{[0,1]} g\Big(|\rho_i(x) - \rho_i'(x)|\Big) dx \\ &\le g\Big(\frac{1}{q}\sum_{i=1}^{q} \int_{[0,1]} |\rho_i(x) - \rho_i'(x)| dx\Big) \\ &= g(d_1(\rho, \rho')) = f(d_1(\rho, \rho')), \end{aligned}$$

where we used the assumption that $d_1(\rho, \rho') \le 1/e$ in the last step. This proves the lemma. □

The following lemma is also easy to prove

LEMMA 5.11. *Let G be an weighted graph on n nodes, and let H be a soft-core weighted graph on q nodes, with weights $\alpha_i(H) = e^{h_i}$ and $\beta_{ij}(H) = e^{2J_{ij}}$. Then*

$$\widehat{\mathcal{F}}(G, J, h) \le \widehat{\mathcal{E}}(G, J, h) \le \widehat{\mathcal{F}}(G, J, h) + \ln q$$

and

$$\widehat{\mathcal{F}}_{\mathbf{a}}(G, J) \le \widehat{\mathcal{E}}_{\mathbf{a}}(G, J) \le \widehat{\mathcal{F}}_{\mathbf{a}}(G, J) + \ln q.$$

5.5.2. *Blowups of a graph.* Instead of directly relating the free energies of G and W_G, we first look at blow ups of G.

LEMMA 5.12. *Let G be a weighted graph with nodeweights one, let $q \geq 1$, and let $\mathbf{a} \in \mathsf{Pd}_q$, $h \in \mathbb{R}^q$, and $J \in \mathsf{Sym}_q$. Denote the k-fold blow up of G by $G[k]$. Then*

$$\left|\widehat{\mathcal{F}}_{\mathbf{a}}(G[k], J) - \mathcal{F}_{\mathbf{a}}(W_G, J)\right| \leq \frac{2q^2}{k}\|J\|_\infty \beta_{\max}(G) + 3q^2 \frac{\ln(k+1)}{k} \tag{5.22}$$

and

$$\left|\widehat{\mathcal{F}}(G[k], J, h) - \mathcal{F}(W_G, J, h)\right| \leq \frac{2q^2}{k}\max\{\|J\|_\infty \beta_{\max}(G), \|h\|_\infty\} + 3q^2 \frac{\ln(k+1)}{k}. \tag{5.23}$$

Proof. Let $V(G) = \{1, \dots, n\}$, and let $I_1, \dots, I_n \subset [0,1]$ be consecutive intervals of lengths $1/n$. Given a configuration $\phi\colon\ V(G[k]) \to [q]$ and a node $u \in V(G)$, let $k_i(u)$ be the number of nodes $u' \in V(G[k])$ such that u' is a copy of u and $\phi(u') = i$, and set $\widehat{\rho}_i(x) = k_i(u)/k$ whenever $x \in I_u$. Let $\mathcal{R}_{\mathbf{a}}$ be the set of all fractional q-partitions ρ of $V(G)$ such that $\alpha(\rho) = \mathbf{a}$, and let $\widehat{\mathcal{R}}_{\mathbf{a}}$ be the set fractional q-partitions τ of $V(G)$ such that $\tau_i(x)$ is an integer multiple of $1/k$, and

$$|\alpha_i(\tau) - a_i\Big| \leq \frac{1}{nk} \qquad \text{for all } i \in [q]. \tag{5.24}$$

Then $\Omega_{\mathbf{a}}(G[k])$ is precisely the set of configuration $\phi\colon\ V(G[k]) \to [q]$ for which $\widehat{\rho} \in \widehat{\mathcal{R}}_{\mathbf{a}}$.

We write the energy density of the configuration ϕ as

$$\mathcal{E}_\phi(G[k], J, 0) = \frac{1}{n^2}\sum_{i,j=1}^{q} J_{ij} \sum_{u,v \in V(G)} \beta_{uv}(G)\widehat{\rho}_i(u)\widehat{\rho}_j(v).$$

The number of configurations ϕ corresponding to a fixed set of numbers $(k_i(u))$ $(i \in [q],\ u \in [n])$ is given by the product of multinomials

$$\prod_{u \in [n]} \frac{k!}{k_1(u)! \dots k_q(u)!}.$$

To continue, we approximate the factorials by the leading term in their asymptotic expansion. Neglecting, for the moment, the error term, we have

$$\prod_{u \in [n]} \frac{k!}{k_1(u)! \dots k_q(u)!} \approx \prod_{u \in [n]} \frac{(k/e)^k}{(k_1(u)/e)^{k_1(u)} \dots (k_q(u)/e)^{k_q(u)}} = \exp\Big(nk\mathsf{Ent}(\widehat{\rho})\Big).$$

To bound the error in the above approximation, we use the following simple inequality, valid for all integers $m \geq 1$:

$$\Big(\frac{m}{e}\Big)^m \leq m! \leq em\Big(\frac{m}{e}\Big)^m.$$

As a consequence, we have that

$$\Big(\frac{1}{ek}\Big)^{qn} e^{nk\mathsf{Ent}(\widehat{\rho})} \le \prod_{u\in[n]} \frac{k!}{k_1(u)!\dots k_q(u)!} \le (ek)^{qn} e^{nk\mathsf{Ent}(\widehat{\rho})}.$$

Bounding finally the number of choices for the qn-tuple $(k_i(u))$ by $(k+1)^{n(q-1)} \le (k+1)^{nq}$, we conclude that

$$(ek)^{-qn} \max_{\widehat{\rho}\in\widehat{\mathcal{R}}_{\mathbf{a}}} e^{nk(\mathsf{Ent}(\widehat{\rho})-\mathcal{E}_{\widehat{\rho}}(W_G,J,0))} \le Z_{\mathbf{a}}(G[k],J)$$

$$\le (ek(k+1))^{qn} \max_{\widehat{\rho}\in\widehat{\mathcal{R}}_{\mathbf{a}}} e^{nk(\mathsf{Ent}(\widehat{\rho})-\mathcal{E}_{\widehat{\rho}}(W_G,J,0))}.$$

The above bound on the partition function implies that

(5.25)
$$\Big|\widehat{\mathcal{F}}_{\mathbf{a}}(G[k],J) - \min_{\widehat{\rho}\in\widehat{\mathcal{R}}_{\mathbf{a}}} \Big(\mathcal{E}_{\widehat{\rho}}(W_G,J,0) - \mathsf{Ent}(\widehat{\rho})\Big)\Big| \le q\frac{\ln(ek(k+1))}{k} \le 3q\frac{\ln(k+1)}{k}.$$

By (5.21), we have

$$\mathcal{F}_{\mathbf{a}}(W_G,J) = \min_{\rho\in\mathcal{R}_{\mathbf{a}}} \Big(\mathcal{E}_{\rho}(W_G,J,0) - \mathsf{Ent}(\rho)\Big). \tag{5.26}$$

To complete the proof of the lemma, we therefore have to compare the fractional partitions in $\widehat{\mathcal{R}}_{\mathbf{a}}$ to those in $\mathcal{R}_{\mathbf{a}}$.

Let $\hat{\rho} \in \widehat{\mathcal{R}}_{\mathbf{a}}$ attain the minimum in the expression on the left-hand side of (5.25). Using the fact that $\hat{\rho}$ obeys the constraint (5.24), it is not hard to show that there exists a fractional q-partition $\rho \in \mathcal{R}_{\mathbf{a}}$ such that $d_1(\rho,\rho') \le \frac{1}{k}$. Inequality (5.14) gives

$$\Big|\mathcal{E}_{\rho}(W_G,J,0) - \mathcal{E}_{\widehat{\rho}}(W_G,J,0)\Big| \le 2q\|J\|_\infty \|W_G\|_\infty d_1(\rho,\widehat{\rho}) \le 2q\|J\|_\infty \beta_{\max}(G)\frac{1}{k},$$

while Lemma 5.10 (together with the fact that $|\mathsf{Ent}(\rho) - \mathsf{Ent}(\widehat{\rho})| \le \ln q \le \frac{q}{k}\ln(k+1)$ if $k \le 2$) implies that

$$|\mathsf{Ent}(\rho) - \mathsf{Ent}(\widehat{\rho})| \le q\frac{\ln(k+1)}{k}.$$

Hence, using also (5.25),

$$\begin{aligned} \hat{\mathcal{F}}_{\mathbf{a}}(G[k],J) &\ge \mathcal{E}_{\widehat{\rho}}(W_G,J,0) - \mathsf{Ent}(\widehat{\rho}) - 3q\frac{\ln(k+1)}{k} \\ &\ge \mathcal{E}_{\rho}(W_G,J,0) - \mathsf{Ent}(\rho) \\ &\quad - 2q\|J\|_\infty\beta_{\max}(G)\frac{1}{k} - q\frac{\ln(k+1)}{k} - 3q\frac{\ln(k+1)}{k} \\ &\ge \mathcal{F}_{\mathbf{a}}(W_G,J) - 2q\|J\|_\infty\beta_{\max}(G)\frac{1}{k} - 4q\frac{\ln(k+1)}{k}. \end{aligned} \tag{5.27}$$

To prove a bound in the opposite direction, consider a fractional q-partition ρ that attains the minimum in (5.26). Given this partition, we will construct a partition $\widehat{\rho} \in \widehat{\mathcal{R}}_{\mathbf{a}}$. Let $b_i(u) = k\rho_i(u)$; then by the Integer Making Lemma [1], there exists integers $k_i(u)$ such that

$$|b_i(u) - k_i(u)| < 1 \qquad (1 \le i \le q,\ 1 \le u \le n), \tag{5.28}$$

$$\Big|\sum_{i=1}^q b_i(u) - \sum_{i=1}^q k_i(u)\Big| < 1 \qquad (1 \le u \le n), \tag{5.29}$$

and

$$\Big|\sum_{u=1}^n b_i(u) - \sum_{i=u}^n k_i(u)\Big| < 1 \qquad (1 \le i \le q). \tag{5.30}$$

Since $\sum_i b_i(u) = k$ is an integer, (5.29) implies $\sum_i k_i(u) = k$, and so $\widehat{\rho}_i(u) = k_i(u)/k$ is a fractional partition. Furthermore (5.30) implies that $|\alpha_i(\widehat{\rho}) - a_i| \le 1/(nk)$, and so $\widehat{\rho} \in \widehat{\mathcal{R}}_{\mathbf{a}}$. Finally, (5.28) gives that

$$d_1(\rho, \widehat{\rho}) \le \frac{q}{k}.$$

Hence, using Lemma 5.10 (this time together with the fact that $|\mathsf{Ent}(\rho) - \mathsf{Ent}(\widehat{\rho})| \le \ln q \le \frac{q^2}{k}\ln(k+1)$ if $1 \le k \le \lfloor qe \rfloor$) and the inequalities (5.14) and (5.25) again, we get

$$\begin{aligned}
\mathcal{F}_{\mathbf{a}}(W_G, J) &= \mathcal{E}_\rho(W_G, J, 0) - \mathsf{Ent}(\rho) \\
&\ge \mathcal{E}_{\widehat{\rho}}(G[k], J, 0) - \mathsf{Ent}(\widehat{\rho}) - 2q^2\|J\|_\infty \beta_{\max}(G)\frac{1}{k} - q^2\frac{\ln(k+1)}{k} \\
&\ge \widehat{\mathcal{F}}_{\mathbf{a}}(G[k], J) - 3q\frac{\ln(k+1)}{k} - 2q^2\|J\|_\infty \beta_{\max}(G)\frac{1}{k} - q^2\frac{\ln(k+1)}{k} \\
&\ge \widehat{\mathcal{F}}_{\mathbf{a}}(G[k], J) - 2q^2\|J\|_\infty \beta_{\max}(G)\frac{1}{k} - 3q^2\frac{\ln(k+1)}{k},
\end{aligned}$$

where in the last step we assume (without loss of generality) that $q \ge 2$. Together with (5.27), this proves the bound (5.22). The bound (5.23) is proved in the same way; in fact, its proof is slightly easier. □

We also need the following lemma of a somewhat similar nature.

LEMMA 5.13. *Let G be a graph with nodeweights 1, and let $q \ge 1$, $\mathbf{a} \in \mathsf{Pd}_q$, $h \in \mathbb{R}^q$, and $J \in \mathsf{Sym}_q$. Let G' be obtained from G by adding k new isolated nodes with nodeweights 1. Then*

(5.31)
$$|\hat{\mathcal{F}}_{\mathbf{a}}(G', J) - \hat{\mathcal{F}}_{\mathbf{a}}(G, J)| \le \frac{k}{|V(G)|}\Big(\frac{q}{2}\ln|V(G)| + (q+2)\Big(\beta_{\max}\|J\|_\infty + \frac{1}{2}\ln q\Big)\Big),$$

$$|\hat{\mathcal{F}}(G', J, h) - \hat{\mathcal{F}}(G, J, h)| \le \frac{k}{|V(G)|}\Big(\beta_{\max}\|J\|_\infty + \|h\|_\infty + \ln q\Big), \tag{5.32}$$

and

$$|\mathcal{F}_{\mathbf{a}}(W_{G'},J)-\mathcal{F}_{\mathbf{a}}(W_G,J)|\le\frac{2q^2k}{|V(G)|}\|J\|_\infty\beta_{\max}. \tag{5.33}$$

Proof. It suffices to prove the case $k=1$. Let $n=|V(G)|$. Let $\phi\in\Omega_{\mathbf{a}}(G)$. We claim that after changing the value of ϕ on at most $\lfloor(q-1)/2\rfloor$ nodes, it can be extended to the new node to get a configuration $\phi'=\phi'(\phi)\in\Omega_{\mathbf{a}}(G')$. Indeed, let $V_i=\phi^{-1}(\{i\})$, and let $\delta_i=(n+1)a_i-|V_i|$. Then $\sum_i\delta_i=1$ and $-1+a_i\le\delta_i\le a_i+1$ (by the assumption that $\phi\in\Omega_{\mathbf{a}}(G)$). Let S_+ be the set of indices for which $\delta_i>1$, and let S_- be the set of indices for which $\delta_i\le 0$. Since $\sum_i\delta_i=1$, we know that $|S_-|\ge|S_+|-1$. Choose $|S_+|-1$ vertices of G in such a way that each has a different image in S_-, and change the images of each of them to a different element of S_+. If we map the new vertex $n+1$ to the remaining element in S_+, we obtain a configuration $\phi'\in\Omega_{\mathbf{a}}(G')$. Since $|S_+|+|S_+|-1\le q$, the number of vertices whose image was changed is at most $\lfloor(q-1)/2\rfloor$, as claimed.

If $\phi'=\phi'(\phi)$ is obtained from ϕ by the above procedure, then

$$|\mathcal{E}_\phi(G,J,0)-\mathcal{E}_{\phi'}(G',J,0)|\le 2\|J\|_\infty\beta_{\max}n\lceil(q-1)/2\rceil\le n(q-1)\beta_{\max}\|J\|_\infty.$$

It is also not hard to check that each configuration ϕ' can arise from at most $n^{\lfloor(q-1)/2\rfloor}q^{\lfloor(q-1)/2\rfloor}\le(nq)^{q/2}$ different configurations ϕ. As a consequence,

$$\begin{aligned}Z_{\mathbf{a}}(G',J)&=\sum_{\psi\in\Omega_{\mathbf{a}}(G')}e^{-\frac{1}{n+1}\mathcal{E}_\psi(G',J,0)}\ge\frac{1}{(nq)^{q/2}}\sum_{\phi\in\Omega_{\mathbf{a}}(G)}e^{-\frac{1}{n+1}\mathcal{E}_{\phi'(\phi)}(G',J,0)}\\&\ge\frac{1}{(nq)^{q/2}}\sum_{\phi\in\Omega_{\mathbf{a}}(G)}e^{-\frac{1}{n}\mathcal{E}_\phi(G',J,0)}e^{-q\beta_{\max}\|J\|_\infty}\\&=\frac{1}{(nq)^{q/2}}\exp(-q\beta_{\max}\|J\|_\infty)Z_{\mathbf{a}}(G,J).\end{aligned}$$

Conversely, from every $\psi\in\Omega_{\mathbf{a}}(G')$ we can construct a $\phi\in\Omega_{\mathbf{a}}(G)$ by deleting the new node and changing the image of at most $\max\{1,\lfloor(q-1)/2\rfloor\}\le q/2$ nodes (where we used that, without loss of generality, $q\ge 2$ since otherwise we do not have to change any nodes). This time, there are at most $q(nq)^{\max\{1,\lfloor(q-1)/2\rfloor\}}\le q(nq)^{q/2}$ different configurations $\psi\in\Omega_{\mathbf{a}}(G')$ that can give rise to the same configuration ϕ. As a consequence, we now have

$$\begin{aligned}Z_{\mathbf{a}}(G',J)&=\sum_{\psi\in\Omega_{\mathbf{a}}(G')}e^{-\frac{1}{n+1}\mathcal{E}_\psi(G',J,0)}\\&\le q(nq)^{q/2}e^{nq\beta_{\max}\|J\|_\infty}\sum_{\phi\in\Omega_{\mathbf{a}}(G)}e^{-\frac{1}{n+1}\mathcal{E}_\phi(G,J,0)}\\&\le q(nq)^{q/2}\exp((q+1)\beta_{\max}\|J\|_\infty)Z_{\mathbf{a}}(G,J).\end{aligned}$$

Combined with the trivial inequality $e^{-n\beta_{\max}\|J\|_\infty} \le Z_{\mathbf{a}}(G,J) \le q^n e^{n\beta_{\max}\|J\|_\infty}$, this gives

$$\begin{aligned}|\hat{\mathcal{F}}_{\mathbf{a}}(G',J) - \hat{\mathcal{F}}_{\mathbf{a}}(G,J)| &= \left|\frac{1}{n+1}\ln Z_{\mathbf{a}}(G',J) - \frac{1}{n}\ln Z_{\mathbf{a}}(G,J)\right| \\ &\le \frac{q+2}{n}\beta_{\max}\|J\|_\infty + \frac{\ln q}{n} + \frac{q\ln(nq)}{2n} \\ &= \frac{q+2}{n}\Big(\beta_{\max}\|J\|_\infty + \frac{1}{2}\ln q\Big) + \frac{q}{2n}\ln n.\end{aligned}$$

This proves (5.31). The inequality (5.32) follows from the observation that

$$Z(G',J,h) = Z(G,J,h)\sum_{i=1}^{q} e^{h_i},$$

and the inequality (5.33) follows easily from Theorem 4.13. □

5.5.3. *Conclusion.* To conclude the proof of Theorem 5.8, we use the following form of the Weak Regularity Lemma due to Frieze and Kannan [7]; see also [3]. We define, for a weighted graph G and a partition $\mathcal{P} = (V_1,\dots,V_k)$ of $V(G)$, the weighted graph $G_{\mathcal{P}}$ on $V(G)$ with nodeweights $\alpha(G_{\mathcal{P}}) = \alpha(G)$ and edgeweights $\beta_{uv}(G_{\mathcal{P}}) = \beta_{ij}(G/\mathcal{P})$ if $(u,v) \in V_i \times V_j$. We call $\mathcal{P}$ *equitable* if

$$\left\lfloor \frac{|V(G)|}{k} \right\rfloor \le |V_i| \le \left\lceil \frac{|V(G)|}{k} \right\rceil$$

for all $i \in [q]$.

Lemma 5.14 ([7]). *For every weighted graph G with all nodeweights 1 and integer $1 \le k \le |V(G)|$, there is an equitable partition $\mathcal{P}$ of $V(G)$ into k classes such that*

$$d_\square(G, G_{\mathcal{P}}) \le \frac{20}{\sqrt{\log_2 k}}\beta_{\max}(G).$$

With the help of this lemma, we now complete the proof of Theorem 5.8 as follows. Let $k = \lceil n^{1/2} \rceil$. It will be convenient to assume that $m = n/k$ is an integer. To this end, add $k' = k\lceil n/k \rceil - n \le n^{1/2}k$ new isolated nodes to G. By Lemma 5.13, the cumulative change to $\widehat{\mathcal{F}}_{\mathbf{a}}(G,J)$ and $\mathcal{F}_{\mathbf{a}}(W_G,J)$ can be bounded by

$$\frac{1}{n^{1/2}}\Big(\frac{q}{2}\ln n + \frac{q+2}{2}\ln q + (2q^2+q+2)\beta_{\max}\|J\|_\infty\Big) \le \frac{q^2}{n^{1/2}}\Big(\frac{1}{2}\ln n + \frac{1}{2} + 5\beta_{\max}\|J\|_\infty\Big).$$

By the Weak Regularity Lemma 5.14, we may now choose an equitable partition $\mathcal{P}$ of $V(G)$ into k classes such that

$$d_\square(G, G_{\mathcal{P}}) \le \frac{20\beta_{\max}(G)}{\sqrt{\log_2 k}} \le \frac{20\sqrt{2}\beta_{\max}(G)}{\sqrt{\ln n}}.$$

To complete the proof, we use the triangle inequality,

$$\begin{aligned}\left|\widehat{\mathcal{F}}_{\mathbf{a}}(G,J)-\widehat{\mathcal{F}}_{\mathbf{a}}(W_G,J)\right| &\le \left|\widehat{\mathcal{F}}_{\mathbf{a}}(G,J)-\widehat{\mathcal{F}}_{\mathbf{a}}(G_{\mathcal{P}},J)\right|\\ &\quad+\left|\widehat{\mathcal{F}}_{\mathbf{a}}(G_{\mathcal{P}},J)-\mathcal{F}_{\mathbf{a}}(W_{G_{\mathcal{P}}},J)\right|\\ &\quad+\left|\mathcal{F}_{\mathbf{a}}(W_{G_{\mathcal{P}}},J)-\mathcal{F}_{\mathbf{a}}(W_G,J)\right|.\end{aligned}$$

Here the first term is bounded by $q^2\|J\|_\infty d_\square(G,G_{\mathcal{P}})$ by (5.16), (2.12) and (2.13), and the last term is bounded by the same quantity by Theorem 4.13.

To estimate the middle term, let $G'=G/\mathcal{P}$. Then $G_{\mathcal{P}}=G'[m]$ and $W_{G_{\mathcal{P}}}=W_{G'}$, and hence by Lemma 5.12,

$$\begin{aligned}\left|\widehat{\mathcal{F}}_{\mathbf{a}}(G_{\mathcal{P}},J)-\mathcal{F}_{\mathbf{a}}(W_{G'},J)\right| &= \left|\widehat{\mathcal{F}}_{\mathbf{a}}(G'[m],J)-\mathcal{F}_{\mathbf{a}}(W_{G'},J)\right|\\ &\le \frac{q^2}{m}\Big(2\|J\|_\infty\beta_{\max}(G)+3(1+\ln m)\Big)\\ &\le \frac{q^2}{n^{1/2}}\Big(4\|J\|_\infty\beta_{\max}(G)+6\Big(1+\frac{1}{2}\ln n\Big)\Big).\end{aligned}$$

Combining the various error terms, we get that

$$\begin{aligned}\left|\widehat{\mathcal{F}}_{\mathbf{a}}(G,J)-\widehat{\mathcal{F}}_{\mathbf{a}}(W_G,J)\right| &\le \frac{q^2}{n^{1/2}}\Big(\frac{7}{2}\ln n+\frac{13}{2}+9\beta_{\max}(G)\|J\|_\infty\Big)\\ &\quad+q^2\frac{40\sqrt{2}}{\sqrt{\ln n}}\|J\|_\infty\beta_{\max}(G)\\ &\le \frac{q^2}{n^{1/2}}\Big(\frac{13}{2}+\frac{14n^{1/4}}{e}\Big)+q^2\frac{40\sqrt{2}+9}{\sqrt{\ln n}}\|J\|_\infty\beta_{\max}(G)\\ &\le \frac{12q^2}{n^{1/4}}+q^2\frac{65}{\sqrt{\ln n}}\|J\|_\infty\beta_{\max}(G).\end{aligned}$$

This proves the first bound of the theorem. The proof of the second bound is completely analogous and is left to the reader.

5.6. *Proof of Theorem* 2.14. Let (G_n) be a sequence of graphs with uniformly bounded edgeweights and no dominating nodeweight.

The equivalence of (i) and (ii) was proved in [3].

Theorem 5.4 and Proposition 5.3 imply that $d_1^{\mathrm{Hf}}(\widehat{\mathcal{S}}_q(G_n),\mathcal{S}_q(W_{G_n}))\to 0$, and hence the sequence $\widehat{\mathcal{S}}_q(G_n)$ is Cauchy in the d_1^{Hf} distance if and only if the sequence $\mathcal{S}_q(W_{G_n})$ is. By Theorem 3.5, this happens if and only if the graphon sequence (W_{G_n}) is convergent, which is equivalent to (i).

Similarly, equation (5.8) and Theorem 5.5 imply that

$$|\widehat{\mathcal{E}}_{\mathbf{a}}(G_n,J)-\mathcal{E}_{\mathbf{a}}(W_{G_n},J)|\to 0,$$

and hence the sequence $\widehat{\mathcal{E}}_{\mathbf{a}}(G_n,J)$ is convergent if and only if the sequence $\mathcal{E}_{\mathbf{a}}(W_{G_n},J)$ is. By Theorem 3.5, this happens for all a and J if and only if the graphon sequence (W_{G_n}) is convergent, which is again equivalent to (i).

Next suppose that $\alpha_{G_n}^2/n \to \infty$. Lemma 5.7 implies that for every weighted graph $H = (\mathbf{a}, J)$,

$$\left| \frac{\hom_{\mathbf{a}}(G_n, H)}{\alpha_{G_n}^2} - \mathcal{E}_{\mathbf{a}}(W_{G_n}, J) \right| \to 0,$$

and hence the sequence $(\hom_{\mathbf{a}}(G_n, H)/\alpha_{G_n}^2)$ is convergent if and only if the sequence $(\mathcal{E}_{\mathbf{a}}(W_{G_n}, J))$ is. As we have seen, this is equivalent to (i).

Now suppose that all nodeweights in the graphs G_n are 1. Then Theorem 5.8 implies that $|\widehat{\mathcal{F}}_{\mathbf{a}}(G_n, J) - \mathcal{F}_{\mathbf{a}}(W_{G_n}, J)| \to 0$, and hence the sequence $\widehat{\mathcal{F}}_{\mathbf{a}}(G_n, J)$ is convergent if and only if the sequence $\mathcal{F}_{\mathbf{a}}(W_{G_n}, J)$ is. We conclude by Theorem 3.5 as before. Similar arguments also prove Theorems 2.15 and 3.7.

5.7. *Proof of Proposition* 3.11. The following lemma is a slight generalization of Lemma 4.3 in [3], and the proof is essentially the same.

Lemma 5.15. *Let $\lambda > 0$, and let H be a weighted graph on n nodes with nodeweights 1. Let X_{ij} $(ij \in E(H))$ be independent random variables such that $\mathsf{E}(X_{ij}) = \beta_{ij}(H)$ and $|X_{ij}| \le C$. Let G be the random graph on $V(H)$ with edgeweights X_{ij}. Then*

$$d_\square(H, G) < C\left(\sqrt{\frac{\lambda + 4\log 4}{n}} + \frac{1}{n}\right)$$

with probability at least $1 - e^{-\lambda n/4}$.

Turning to the proof of Proposition 3.11, recall that we are considering two sequences (G_n) and (G'_n) of graphs. Consider a third sequence, G''_n, with $V(G''_n) = V(G'_n)$, $E(G'_n) = E(G''_n)$ and $\beta_{uv}(G''_n) = \bar\beta$.

Lemma 5.15 implies that with probability one, $d_\square(G'_n, G''_n) \to 0$ as $n \to \infty$. Combined with the easy bound (5.16), this immediately gives the statement of the proposition. Indeed, using the fact that $\widehat{\mathcal{F}}(G_n, \bar\beta J, h) = \widehat{\mathcal{F}}(G_n,'' J, h)$, we may use (5.16) to bound

$$\begin{aligned}\left|\widehat{\mathcal{F}}(G'_n, J, h) - \widehat{\mathcal{F}}(G_n, \bar\beta J, h)\right| &= \left|\widehat{\mathcal{F}}(G'_n, J, h) - \widehat{\mathcal{F}}(G_n,'' J, h)\right| \\ &\le q^2 \max\{\|J\|_\infty, \|h\|_\infty\} d_\square(G,'' G').\end{aligned}$$

6. **Weaker convergence**

6.1. *Counterexamples.* By Theorems 4.13 and 4.14, graphons that are near in the cut-metric have similar free energies, and thus also similar ground state energies. Our first example shows that the converse does not hold. Indeed, it gives a family of distinct graphons that have the same free energies and ground state energies.

Example 6.1 (Block Diagonal Graphons). Given $0 < \alpha < 1$ and $\beta_1, \beta_2 \geq 0$, let W be the block diagonal graphon

$$W(x,y) = \begin{cases} \beta_1 & \text{if} \quad 0 \leq x,y \leq \alpha, \\ \beta_2 & \text{if} \quad \alpha \leq x,y \leq 1, \\ 0 & \text{otherwise.} \end{cases} \tag{6.1}$$

It is easy to express the free energies of W in terms of the free energies of the constant graphons $W_1 \equiv \alpha\beta_1$ and $W_2 \equiv (1-\alpha)\beta_2$. If $q \geq 1$, $h \in \mathbb{R}^q$, and $J \in \mathsf{Sym}_q$, then

$$\begin{aligned} \mathcal{F}(W,J,h) &= \sum_{u=1,2} \min_{\rho^{(u)} \in \mathsf{Pd}_q} \Big(\alpha_u^2 \beta_u \sum_{i,j} J_{ij} \rho_i^{(u)} \rho_j^{(u)} - \alpha_u \sum_i \rho_i^{(u)} h_i - \alpha_u \mathsf{Ent}(\rho^{(u)}) \Big) \Big) \\ &= \sum_{u=1,2} \alpha_u \mathcal{F}(\beta_u \alpha_u, J, h), \end{aligned} \tag{6.2}$$

where $\alpha_1 = \alpha$ and $\alpha_2 = 1-\alpha$. The same calculation also shows that

$$\mathcal{E}(W,J,h) = \sum_{u=1,2} \alpha_u \mathcal{E}(\beta_u \alpha_u, J, h). \tag{6.3}$$

Choosing

$$\beta_1 = 1/\alpha \qquad \text{and} \qquad \beta_2 = 1/(1-\alpha) \qquad \text{with} \qquad \alpha \in (0, 1/2],$$

we obtain a one-parameter family of distinct graphons that cannot be distinguished by their free energies or ground state energies since $\mathcal{F}(W,J,h) = \mathcal{F}(1,J,h)$ and $\mathcal{E}(W,J,h) = \mathcal{E}(1,J,h)$ for all W in the family.

Obviously, two distinct graphons that can be distinguished by their ground state energies without magnetic fields can also be distinguished if we allow magnetic fields. Our next example shows that the converse is not true.

Example 6.2. Consider again the block diagonal graphon defined in (6.1). It is easy to calculate the ground state energy of this graphon for $h = 0$, giving

$$\mathcal{E}(W,J,0) = \Big(\alpha^2\beta_1 + (1-\alpha)^2\beta_2\Big)\mathcal{E}(1,J,0).$$

Choosing

$$\beta_1 = \frac{\lambda}{\alpha} \qquad \text{and} \qquad \beta_2 = \frac{1-\alpha\lambda}{(1-\alpha)^2}$$

with $\alpha \in (0,1/2]$ and $\lambda \in (0, 1/\alpha]$, we obtain a two-parameter family of distinct graphons that cannot be distinguished by the ground state energies without magnetic fields.

But only the subfamily considered in Example 6.1, i.e., the subfamily with $\lambda = 1$, remains indistinguishable if we allow magnetic fields. Indeed, consider the case $q = 2$, $J_{ij} = 1 - \delta_{i,1}$, and $h_i = c\delta_{1,i}$ from the biased max-cut problem discussed in Section 2.3. The biased max-cut for $W \equiv \beta$ can be easily calculated, giving

$$-\mathcal{E}(\beta, J, h) = \max_{a \in [0,1]} \Big(2a(1-a)\beta + ca \Big) = \frac{\beta + c}{2} + \frac{c^2}{8\beta}$$

provided $|c| \le 2\beta$. Taking into account the relation (6.3), we conclude that for all graphons W in the above family, we have

$$-\mathcal{E}(W, J, h) = \frac{c+1}{2} + + \frac{c^2}{8}\Big(\frac{\alpha}{\lambda} + \frac{(1-\alpha)^2}{1-\alpha\lambda} \Big)$$

provided $|c| \le \min\{2\lambda/\alpha, (1-\alpha\lambda/(1-\alpha)^2\}$. Thus two elements of the family can be distinguished by the biased max-cut problem unless $\frac{\alpha}{\lambda} + \frac{(1-\alpha)^2}{1-\alpha\lambda} = 1$, i.e., unless $\lambda = 1$, as claimed.

We have seen in the previous sections that right-convergence implies convergence of the ground state energies, which in turn implies convergence of the ground state energies without magnetic fields, and hence naive right-convergence. Using Examples 6.1 and 6.2, it is not hard to show that right-convergence is in fact strictly stronger than convergence of the ground state energies, which in turn is strictly stronger than naive right-convergence. This is the content of the next example.

Example 6.3. We first give an example of a sequence of simple graphs that has convergent ground state energies, but is not left-convergent, and therefore also not right-convergent. Let $p \le 1/2$, let $G_n = \mathbb{G}(n, p)$, and let G_n' be the disjoint union of two random graphs $\mathbb{G}(n, 2p)$. With probability one, G_n then converges from the left to the constant graphon $W \equiv p$, and G_n' converges from the left to the graphon W' defined by (6.1) with $\alpha = 1/2$ and $\beta_1 = \beta_2 = 2p$. As a consequence, $\widehat{\mathcal{E}}(G_n, J, h) \to \mathcal{E}(W, J, h)$ and $\widehat{\mathcal{E}}(G_n', J, h) \to \mathcal{E}(W', J, h)$. By the identity (6.3), $\mathcal{E}(W, J, h) = \mathcal{E}(W', J, h)$ for all $q \ge 1$, all $J \in \mathsf{Sym}_q$, and all $h \in \mathbb{R}^d$, implying that the ground state energies of G_n and G_n' converge to the same limiting ground state energy. Interleaving the two sequences (G_n) and (G_n'), we get a sequence of simple graphs that is not left-convergent, but has convergent ground state energies. (Taking into account the identity (6.2), we see that this sequence has convergent free energies as well.)

In a similar way, we can use Example 6.2 to construct a sequence of simple graphs that is naively right-convergent but does not have convergent ground state energies. Indeed, let W be the constant graphon $W \equiv p$, and let W' be the graphon defined in (6.1) with $\alpha = 1/2$, $\beta_1 = p$, and $\beta_2 = 3p$. Then $\mathcal{E}(W', J, 0) = \Big(\frac{p}{4} + \frac{3p}{4}\Big)\mathcal{E}(1, J, 0) = \mathcal{E}(W, J, 0)$. Let $G_n = \mathbb{G}(p, n)$ and G_n' be the disjoint union of $\mathbb{G}(p, n)$ and $\mathbb{G}(3p, n)$. If H is a soft-core graph on q nodes with $\beta_{ij}(H) = e^{2J_{ij}}$, then $\frac{1}{n^2}\hom(G_n, H) \to \mathcal{E}(W, J, 0)$ and $\frac{1}{(2n)^2}\log\hom(G_n', H) \to$

$\mathcal{E}(W', J, 0) = \mathcal{E}(W, J, 0)$. Interleaving the two sequences, we thus obtain a sequence which is naively right-convergent, but does not have convergent ground state energies once we allow for nonzero magnetic fields.

We finally give an example showing that the statements of Theorems 2.14, 2.15, and 3.7 concerning the free energy do not hold if we relax the condition that (G_n) has nodeweights one.

Example 6.4. Let G be the weighted graph on $\{1,2\}$ with weights $\beta_{11}(G) = \beta_{22}(G) = 1$, $\alpha_1(G) = 1/3$, and $\alpha_2(G) = 2/3$, and let G_n be obtained from G by blowing up each node n times, $G_n = G[n]$. Then G_n converges to the block-diagonal graphon

$$W(x,y) = \begin{cases} 1 & \text{if} \quad 0 \le x, y \le 1/3 \quad \text{or} \quad 1/3 \le x, y \le 1, \\ 0 & \text{otherwise.} \end{cases}$$

But the free energies and microcanonical free energies of G_n do not converge to those of W. Indeed, let $q \ge 2$, $\mathbf{a} \in \mathsf{Pd}_q$, and $J \in \mathsf{Sym}_q$. Proceeding as in the proof of Lemma 5.12, it is then not hard to show that $\widehat{\mathcal{F}}_{\mathbf{a}}(G[n], J)$ converges to

$$\mathcal{F}_{\mathbf{a}}^{\infty} = \inf_{\rho} \Big(\mathcal{E}_\rho(G, J, 0) + \frac{1}{2} \sum_{x \in V(G)} \sum_{i \in [q]} \rho_i(x) \log \rho_i(x) \Big),$$

where the infimum goes over all fractional partitions ρ of $V(G) = \{1,2\}$ obeying the constraint $\frac{1}{3}\rho_i(1) + \frac{2}{3}\rho_i(2) = a_i$, while $\mathcal{F}(G[n], J, 0)$ converges to $\mathcal{F}^{\infty} = \min_{\mathbf{a} \in \mathsf{Pd}_q} \mathcal{F}_{\mathbf{a}}^{\infty}$. Note that the nodeweights of G enter into the energy term $\mathcal{E}_\rho(G, J, 0)$ and the condition on ρ, but not into the entropy term $\frac{1}{2} \sum_x \sum_i \rho_i(x) \log \rho_i(x)$, in contrast to the corresponding expression for the microcanonical free energies of the limit W,

$$\mathcal{F}_{\mathbf{a}}(W, J) = \inf_{\rho} \Big(\mathcal{E}_\rho(G, J, 0) + \frac{1}{3} \sum_{i \in [q]} \rho_i(1) \log \rho_i(1) + \frac{2}{3} \sum_{i \in [q]} \rho_i(2) \log \rho_i(2) \Big),$$

where the nodeweights enter into both the energy and the entropy term.

Specializing to the Ising model with spin space $\{-1, +1\}$ and coupling constants $J_{\phi,\phi'} = \frac{K}{2}\phi\phi'$, we write the limit $\mathcal{F}_{\mathbf{a}}^{\infty}$ as

$$\mathcal{F}_{\mathbf{a}}^{\infty} = - \max_{\substack{m_1, m_2 \in [-1,1] \\ \frac{1}{3}m_1 + \frac{2}{3}m_2 = m}} \Big(\frac{K}{2} \Big(\frac{1}{9} m_1^2 + \frac{4}{9} m_2^2 \Big) + \frac{1}{2} \mathsf{Ent}(m_1) + \frac{1}{2} \mathsf{Ent}(m_2) \Big)$$

and the free energy of the limit W as

$$\mathcal{F}_{\mathbf{a}}(W, J) = - \max_{\substack{m_1, m_2 \in [-1,1] \\ \frac{1}{3}m_1 + \frac{2}{3}m_2 = m}} \Big(\frac{K}{2} \Big(\frac{1}{9} m_1^2 + \frac{4}{9} m_2^2 \Big) + \frac{1}{3} \mathsf{Ent}(m_1) + \frac{2}{3} \mathsf{Ent}(m_2) \Big).$$

Here $m = m(\mathbf{a}) = a_+ - a_-$ and

$$\mathsf{Ent}(m) = -\frac{1+m}{2}\ln\Big(\frac{1+m}{2}\Big) - \frac{1-m}{2}\ln\Big(\frac{1-m}{2}\Big).$$

Let $K = 3/2$, let $\widehat{m} \geq 0$ be such that

$$\max_{m_2\in[-1,1]}\Big(\frac{m_2^2}{3} + \frac{1}{2}\mathsf{Ent}(m_2)\Big) = \frac{\widehat{m}^2}{3} + \frac{1}{2}\mathsf{Ent}(\widehat{m}),$$

and let $\widehat{\mathbf{a}} \in \mathsf{Pd}_2$ and be such that $\widehat{m} = \widehat{a}_1 - \widehat{a}_2$. Using the fact that

$$\max_{M\in[-1,1]}\Big(\frac{\tilde{K}}{2}M^2 + \mathsf{Ent}(M)\Big) \geq \ln 2 = \mathsf{Ent}(0),$$

with equality if and only if $\tilde{K} \leq 1$ (see Example 3.9), it is then not hard to check that

$$\mathcal{F}^\infty = \mathcal{F}^\infty_{\widehat{\mathbf{a}}} = -\frac{\widehat{m}^2}{3} - \frac{1}{2}\mathsf{Ent}(\widehat{m}) - \frac{1}{2}\mathsf{Ent}(0) < -\ln 2,$$

while

$$\begin{aligned}\mathcal{F}_{\widehat{\mathbf{a}}}(W,J) &\geq \mathcal{F}(W,J,0)\\ &= -\max_{m_1,m_2}\Big(\frac{m_1^2}{12} + \frac{m_2^2}{3} + \frac{1}{3}\mathsf{Ent}(m_1) + \frac{2}{3}\mathsf{Ent}(m_2)\Big) = \mathsf{Ent}(0) = -\ln 2.\end{aligned}$$

This proves that $\lim_{n\to\infty}\widehat{\mathcal{F}}_{\widehat{\mathbf{a}}}(G[n],J) < \mathcal{F}_{\widehat{\mathbf{a}}}(W,J)$ and $\lim_{n\to\infty}\widehat{\mathcal{F}}(G[n],J,0) < \mathcal{F}(W,J,0)$.

Interspersing the sequence $(G[n])$ with an arbitrary sequence of simple graphs that converges to W, this also yields an example of a convergent sequence of weighted graphs whose free energies and microcanonical free energies do not converge.

6.2. *Naive right-convergence with two weights.* We have seen that naive right-convergence is not enough to guarantee left-convergence. But there is a way of saving the equivalence of right-convergence with left-convergence, by considering target graphs H with two edgeweights β_{ij} and γ_{ij}. We call these graphs *doubly weighted*. We say that H is *soft-core* if $\beta_{ij}, \gamma_{ij} > 0$ for all $i, j \in V(H)$. The value $\hom(G,H)$ is defined as

$$\begin{aligned}\hom(G,H) = &\sum_{\phi:V(G)\to V(H)}\prod_{u\in V(G)}\alpha_{\phi(u)}(H)\\ &\times\prod_{uv\in E(G)}\beta_{\phi(u),\phi(v)}(H)\prod_{uv\notin E(G)}\gamma_{\phi(u),\phi(v)}(H).\end{aligned}$$

THEOREM 6.5. *Let (G_n) be a sequence with uniformly bounded edgeweights, and nodeweights* 1. *Then (G_n) is left-convergent if and only if*

$$\frac{\ln\hom(G_n,H)}{|V(G)|^2}$$

has a limit for each doubly weighted soft-core graph H.

Proof. The proof of the "only if" part is analogous to the proof of the first statement in Theorem 2.15 and is left to the reader. The idea of the proof of the "if" part is that one can use the second set of edgeweights to force the dominating partition to have prescribed sizes, and thereby show that the microcanonical ground state energies converge. To be more precise, let $q \geq 1$, $\mathbf{a} \in \mathsf{Pd}_q$, and $J \in \mathsf{Sym}_q$. Define a doubly weighted graph H_C by

$$\alpha_i = 1, \qquad \gamma_{ij} = \exp\Big(\frac{-C}{a_i}\mathbf{1}_{i=j}\Big), \qquad \beta_{ij} = \exp\Big(J_{ij} + \frac{C}{a_i}\mathbf{1}_{i=j}\Big) \qquad (i, j \in [q]).$$

Then for every graph G,

$$\hom(G, H_C)\exp(C|V(G)|^2) = \sum_{\phi} \alpha_\phi \exp(\mathcal{E}_\phi(G, J, 0)) \exp\Big(-\sum_i \frac{C}{a_i}\big(a_i n - |\phi^{-1}(i)|\big)^2\Big).$$

The last factor is maximized when $\phi \in \Omega_{\mathbf{a}}(G)$, from which it is not hard to show that

$$\lim_{C\to\infty}\Big(\frac{\ln \hom(G, H_C)}{|V(G)|^2} + C\Big) = \mathcal{E}_{\mathbf{a}}(G, J),$$

uniformly in G, and hence the theorem follows. □

6.3. *Convergence of spectra.* Every graphon $W \in \mathcal{W}$ defines an operator $T_W:\ L_2[0,1] \to L_2[0,1]$ by

$$T_W f(x) = \int_0^1 W(x, y) f(y)\, dy.$$

It is well known that this operator is self-adjoint and compact, and hence it has a discrete real spectrum $\Lambda(W)$, whose only possible point of accumulation is 0. We consider $\Lambda(W)$ as multiset. For $i \geq 1$, let $\lambda_i(W)$ denote the i^{th} largest element of the spectrum (counting multiplicities), provided the spectrum has at least i positive elements; otherwise, let $\lambda_i(W) = 0$. Similarly, let $\lambda'_i(W)$ denote the i^{th} smallest element of the spectrum, provided the spectrum has at least i negative elements; otherwise, let $\lambda'_i(W) = 0$.

It is known that for $k \geq 2$, the sum $\sum_{\lambda\in\Lambda(W)} \lambda^k$ is absolute convergent. In fact, we have

$$\sum_{\lambda\in\Lambda(W)} \lambda^2 = \|W\|_2^2 \quad \text{and} \quad \sum_{\lambda\in\Lambda(W)} \lambda^k = t(C_k, W) \quad \text{for all } k \geq 3. \tag{6.4}$$

It follows that $\sum_{i=1}^m \lambda_i^3 \leq t(C_3, W)$, and hence

$$\lambda_m \leq \frac{t(C_3, W)^{1/3}}{m^{1/3}}. \tag{6.5}$$

For a graph G with n nodes, we consider its adjacency matrix A_G and its eigenvalues $\mu_1 \geq \mu_2 \geq \cdots \geq \mu_n$. We define its *normalized eigenvalues*

$\lambda_i = \mu_i/n$, $(i = 1, \dots, n)$. Again for $k \geq 3$, we have

$$\sum_{i=1}^{n} \lambda_i^k = t(C_k, G). \tag{6.6}$$

We note that the spectrum of W_G is the normalized spectrum of G, together with infinitely many 0's.

The following is a generalization of Theorem 2.9(ii) to weighted graphs, and it also gives the values of the limiting eigenvalues.

THEOREM 6.6. *Let W be a graphon, and let $(G_m : \; m = 1, 2, \dots)$ be a sequence of weighted graphs with uniformly bounded edgeweights tending to W. Let $|V(G_m)| = n_m$, and let $\lambda_{m,1} \geq \lambda_{m,2} \geq \cdots \geq \lambda_{n,n_m}$ be the normalized spectrum of G_m. Then for every $i \geq 1$,*

$$\lambda_{m,i} \to \lambda_i(W) \quad \textit{and} \quad \lambda_{m,n_m+1-i} \to \lambda'_i(W) \quad \textit{as} \quad n \to \infty.$$

We can prove a bit more.

THEOREM 6.7. *Let $(W_1, W_2, \dots)$ be a sequence of uniformly bounded graphons, converging (in the $\delta_\square$ metric) to a graphon W. Then for every $i \geq 1$,*

$$\lambda_i(W_n) \to \lambda_i(W) \quad \textit{and} \quad \lambda'_i(W_n) \to \lambda'_i(W) \quad \textit{as} \quad n \to \infty. \tag{6.7}$$

Proof. If the conclusion does not hold, then there is an $i_0 \geq 0$ for which (say) $\lambda_i(W_n) \not\to \lambda_i(W)$. Choosing a suitable subsequence, we may assume that for each $j \geq 1$, the limits

$$\mu_j = \lim_{n\to\infty} \lambda_j(W_n) \qquad \text{and} \qquad \mu'_j = \lim_{n\to\infty} \lambda'_j(W_n)$$

exist and that $\mu_{i_0} \neq \lambda_{i_0}(W)$.

We claim that for every $k \geq 4$,

$$\lim_{n\to\infty} \sum_{j=1}^{\infty} \lambda_j^k(W_n) \longrightarrow \sum_{j=1}^{\infty} \mu_j^k, \quad \text{and} \quad \lim_{n\to\infty} \sum_{j=1}^{\infty} {\lambda'_j}^k(W_n) \longrightarrow \sum_{j=1}^{\infty} {\mu'_j}^k. \tag{6.8}$$

Indeed, the sequence $t(C_3, W_n)$ is convergent, and hence it is bounded by some constant c; but then (6.5) tells us that $\lambda_m(W_n) < (c/m)^{1/3}$, and hence the sum $\sum_m \lambda_m^k(W_n)$ is uniformly majorized by the convergent series $\sum_m (c/m)^{4/3}$. Hence we can take the limit term-by-term in the sums on the left-hand side.

Using once more the convergence of $t(C_k, W_n)$ to $t(C_k, W)$, we conclude that for every $k \geq 4$, we have

$$\sum_{j=1}^{\infty} \mu_j^k + \sum_{j=1}^{\infty} {\mu'_j}^k = \sum_{j=1}^{\infty} \lambda_j(W)^k + \sum_{j=1}^{\infty} \lambda'_j(W)^k. \tag{6.9}$$

To conclude, it suffices to prove that the two sums on each side are the same term-by-term:

$$\mu_j = \lambda_j(W) \qquad \text{and} \qquad \mu'_j = \lambda'_j(W) \qquad (j \geq 0). \tag{6.10}$$

Indeed, this can be proved by induction on j. Let λ_j occur a times in the sequence $(\lambda_1, \lambda_2, \dots)$ and b times in the sequence $(\mu_1, \mu_2, \dots)$. Let $-\lambda_j$ occur a' times in the sequence $(\lambda'_1, \lambda'_2, \dots)$ and b' times in the sequence $(\mu'_1, \mu'_2, \dots)$. Assume by induction that $\lambda_i = \mu_i$ for $i < j$ and that $\lambda'_i = \mu'_i$ whenever $|\lambda'_i| > \lambda_j$ or $|\mu'_i| > \lambda_j$. Subtracting the contribution of these terms from both sides of (6.9), and sending $k \to \infty$ through the even numbers, the left-hand side is asymptotically $(b + b')\lambda_j^k$, while the right-hand side is $(a + a')\lambda_j^k$. This implies that $a + a' = b + b'$. Similarly, letting k tend to infinity through the odd numbers, we get that $a - a' = b - b'$. This implies that $a = b$ and $a' = b'$ so, in particular, $\lambda_j = \mu_j$ as claimed. □

Let I be a bounded interval. The previous theorem then states that for all $i \geq 1$, the maps $W \mapsto \lambda_i(W)$ and $W \mapsto \lambda'_i(W)$ are continuous maps from $(\mathcal{W}_I, \delta_\square)$ to $\mathbb{R}$. By the compactness of $(\mathcal{W}_I, \delta_\square)$, these maps are uniformly continuous, implying the following.

Corollary 6.8. *For every bounded interval I, every $\varepsilon > 0$, and every $i \geq 1$, there is a $\delta_i > 0$ such that if $U, W \in \mathcal{W}_I$ and $\delta_\square(U, W) \leq \delta_i$, then*

$$|\lambda_i(U) - \lambda_i(W)| \leq \varepsilon \quad \textit{and} \quad |\lambda'_i(U) - \lambda'_i(W)| \leq \varepsilon.$$

For the special case when the sequence (G_n) is quasirandom with density p, the largest normalized eigenvalue of G_n tends to p, while the others tend to 0. In this special case, this statement has a converse. If (G_n) is a sequence of graphs such that the edge-density on G_n tends to p, the largest normalized eigenvalue of G_n tends to p, and all the other eigenvalues tend to 0, then (G_n) is quasirandom.

This converse, however, does not extend to a characterization of convergent graph sequences in any direct way. Consider two regular nonisomorphic graphs G_1 and G_2 with the same spectrum, say the incidence graphs of two nonisomorphic finite projective planes of the same order n. Consider the blow ups $G_1(n)$ and $G_2(n)$, $n = 1, 2, \dots$, and merge them into a single sequence. This sequence is not convergent, but all graphs in it have the same edge density, and the spectra of all graphs are the same except for the 0's.

7. Quasi-inner product and noneffective arguments

There is an alternative way of expressing ground state energies, which leads to a shorter, but noneffective proof of our main result, the equivalence of left and right-convergence. Moreover, it reduces significantly the set of test graphs $\{(\mathbf{a}, H)\}$ respectively $\{(\mathbf{a}, J)\}$ in Definition 2.2 respectively in Definition 2.7 (see Corollary 7.4).

The best space to work with graphons is the compact space $\widetilde{\mathcal{W}}_I$. (For simplicity, assume that $I = [0, 1]$.) This space has no linear structure, and the sum $U+W$ or inner product $\langle U,W\rangle$ of two graphons cannot be defined in a way that would be invariant under weak isomorphism. However, we can replace the inner product by the following version, which will be very useful:

$$\mathcal{C}(U,W)=\sup_{\phi}\langle U,W^{\phi}\rangle=\sup_{\phi}\int_{[0,1]^2} U(x,y)W(\phi(x),\phi(y))\,dx\,dy,$$

where the supremum is taken over all measure preserving bijections $\phi:\ [0,1]\to[0,1]$.

We can use this quasi-inner product to express ground state energies. Let $\mathbf{a}\in\mathsf{Pd}_q$ and $J\in\mathsf{Sym}_q$, and let H be the weighted graph on $[q]$ with nodeweights $\mathbf{a}$ and edgeweights J. Then for every graphon W,

$$\mathcal{E}_{\mathbf{a}}(W,J)=\mathcal{C}(W,W_H). \tag{7.1}$$

We can also express the cut norm with this functional as

$$\|W\|_\square=\sup_{S,T\subseteq[0,1]}\langle W,\mathbf{1}_{S\times T}\rangle=\sup_{a,b\in[0,1]}\mathcal{C}(W,\mathbf{1}_{[0,a]\times[0,b]}). \tag{7.2}$$

The functional $\mathcal{C}(U,W)$ has many good properties. It follows, just like for the cut norm in [3], that

$$\mathcal{C}(U,W)=\sup_{\phi}\langle U,W^{\phi}\rangle=\sup_{\phi}\langle U^{\phi},W\rangle=\sup_{\sigma,\tau}\langle U^{\sigma},W^{\tau}\rangle, \tag{7.3}$$

where ϕ ranges over all measure preserving bijections $[0,1]\to[0,1]$, and σ,τ range over all measure preserving, but not necessarily bijective maps $[0,1]\to[0,1]$. Hence the overlay functional is invariant under measure preserving transformations of the graphons U and W; i.e., it is a functional on the space $\widetilde{\mathcal{W}}_I\times\widetilde{\mathcal{W}}_I$. It also follows that this quantity has the (somewhat unexpected) symmetry property $\mathcal{C}(U,W)=\mathcal{C}(W,U)$ and satisfies the inequalities

$$\langle U,W\rangle\le\mathcal{C}(U,W)\le\|U\|_2\|W\|_2,\quad \mathcal{C}(U,W)\le\|U\|_\infty\|W\|_1. \tag{7.4}$$

This supports the claim that $\mathcal{C}(.,.)$ behaves like some kind of inner product. This analogy is further supported by the following identity, a kind of "Cosine Theorem," relating it to the distance δ_2 derived from the L_2-norm:

$$2\mathcal{C}(U,W)=\|U\|_2^2+\|W\|_2^2-\delta_2(U,W)^2=\delta_2(U,0)^2+\delta_2(W,0)^2-\delta_2(U,W)^2. \tag{7.5}$$

But we have to be a bit careful; the functional $\mathcal{C}(U,W)$ is not bilinear, only subadditive:

$$\mathcal{C}(U+V,W)\le\mathcal{C}(U,W)+\mathcal{C}(V,W). \tag{7.6}$$

It is homogeneous for positive scalars. If $\lambda>0$, then

$$\mathcal{C}(\lambda U,W)=\mathcal{C}(U,\lambda W)=\lambda\mathcal{C}(U,W) \tag{7.7}$$

and $\mathcal{C}(U,W) = \mathcal{C}(-U,-W)$, but $\mathcal{C}(U,W)$ and $\mathcal{C}(-U,W)$ are not related in general.

A less trivial property of this functional is that it is continuous in each variable with respect to the $\delta_\square$ distance. This does not follow from (7.5), since the distance $\delta_2(U,W)$ is not continuous with respect to $\delta_\square$, only lower semicontinuous.

LEMMA 7.1. *If* $\delta_\square(U_n,U) \to 0$ *as* $n \to \infty$ $(U, U_n \in \mathcal{W}_1)$, *then* $\mathcal{C}(U_n,W) \to \mathcal{C}(U,W)$ *for every* $W \in \mathcal{W}_1$.

Proof. We may assume that $\|U_n - U\|_\square \to 0$. By subadditivity (7.6), we have

$$-\mathcal{C}(U-U_n,W) \le \mathcal{C}(U_n,W) - \mathcal{C}(U,W) \le \mathcal{C}(U_n-U,W),$$

and hence it is enough to prove that $\mathcal{C}(U_n-U,W), \mathcal{C}(U-U_n,W) \to 0$. In other words, it suffices to prove the lemma in the case when $U_n \to U = 0$.

The usual inner product $\langle U,W\rangle$ is continuous in each variable with respect to the cut norm, which was noted, e.g., in [11, Lemma 2.2]. Since $\mathcal{C}(U_n,W) \ge \langle U_n,W\rangle$, it follows that $\liminf_n \mathcal{C}(U_n,W) \ge 0$.

To prove the opposite inequality, we start with the case when W is a stepfunction. Write $W = \sum_{i=1}^m a_i \mathbf{1}_{S_i\times T_i}$. Then using (7.6) and (7.2), we get

$$\begin{aligned}\mathcal{C}(U_n,W) &\le \sum_{i=1}^m \mathcal{C}(U_n, a_i\mathbf{1}_{S_i\times T_i})\\ &= \sum_{i=1}^m \mathcal{C}(a_iU_n, \mathbf{1}_{S_i\times T_i}) \le \sum_{i=1}^m \|a_iU_n\|_\square = \sum_{i=1}^m |a_i|\|U_n\|_\square.\end{aligned}$$

Since every term tends to 0, we get that $\limsup \mathcal{C}(U_n,W) \le 0$.

Now if W is an arbitrary kernel, then for every $\varepsilon > 0$, we can find a stepfunction W' such that $\|W-W'\|_1 \le \varepsilon/2$. Then $\mathcal{C}(U_n,W') \to 0$, and hence $\mathcal{C}(U_n,W') \le \varepsilon/2$ if n is large enough. But then

$$\mathcal{C}(U_n,W) \le \mathcal{C}(U_n,W-W') + \mathcal{C}(U_n,W') \le \|U_n\|_\infty\|W-W'\|_1 + \varepsilon/2 \le \varepsilon.$$

This shows that $\limsup_n \mathcal{C}(U_n,W) \le 0$ and completes the proof. □

Remark 7.2. While the functional $\mathcal{C}(U,W)$ is continuous in each variable, it is not continuous as a function on $\widetilde{\mathcal{W}}_1 \times \widetilde{\mathcal{W}}_1$. Let (G_n) be any quasirandom graph sequence, and let $W_n = U_n = 2W_{G_n} - 1$. Then $U_n, W_n \to 0$ in the cut norm (and so also in $\delta_\square$), but $\mathcal{C}(U_n,W_n) = 1$ for all n.

We use the quasi-inner product to give a proof of what can be considered as the main result in this paper, the equivalence of (ii) and (iv) in Theorem 3.5.

THEOREM 7.3. *A sequence* (W_n) *of graphons in* $\mathcal{W}_I$ *is convergent in the* $\delta_\Box$ *distance if and only if* $(\mathcal{E}_{\mathbf{a}}(W_n, J))$ *is convergent for every* $\mathbf{a} \in \mathsf{Pd}_q$ *and* $J \in \mathsf{Sym}_q$.

Proof. Note that by (7.1), $(\mathcal{E}_{\mathbf{a}}(W_n, J))$ is convergent for every $\mathbf{a} \in \mathsf{Pd}_q$ and $J \in \mathsf{Sym}_q$ if and only if $\mathcal{C}(W_n, W_H)$ is convergent for every weighted graph H.

Suppose that $W_n \to W$ in the cut distance. We may apply measure preserving transformations so that the $W_n \to W$ in the cut norm. Then for every $U \in \mathcal{W}_1$, by (7.6) and Lemma 7.1, we have

$$\mathcal{C}(U, W_n) - \mathcal{C}(U, W) \le \mathcal{C}(U, W_n - W) \to 0,$$

and hence $\limsup_n \mathcal{C}(U, W_n) \le \mathcal{C}(U, W)$. Replacing U by $-U$ shows that $\liminf_n \mathcal{C}(U, W_n) \ge \mathcal{C}(U, W)$, and hence $\lim_n \mathcal{C}(U, W_n) = \mathcal{C}(U, W)$. In particular, $\mathcal{C}(W_n, W_H)$ is convergent for every weighted graph H.

Conversely, let (W_n) be a sequence that is not convergent in the cut distance. By the compactness of the graphon space, it has two subsequences, (W_{n_i}) and (W_{m_i}), converging to different (not weakly isomorphic) graphons W and W'. Then there is a graphon U such that $\mathcal{C}(W, U) \ne \mathcal{C}(W', U)$; in fact, (7.5) implies

$$(\mathcal{C}(W', W') - \mathcal{C}(W', W)) + (\mathcal{C}(W, W) - \mathcal{C}(W', W)) = \delta_2(W', W)^2 \ge \delta_\Box(W', W)^2 > 0,$$

and so either $\mathcal{C}(W', W') \ne \mathcal{C}(W, W')$ or $\mathcal{C}(W, W') \ne \mathcal{C}(W, W)$.

Suppose that $\mathcal{C}(W, U) \ne \mathcal{C}(W', U)$, and let (H_k) be any sequence of simple graphs such that $H_k \to U$ in the $\delta_\Box$ distance. Then by Lemma 7.1, we have $\mathcal{C}(W, W_{H_k}) \ne \mathcal{C}(W', W_{H_k})$ if k is large enough, and for this simple graph H_k, the sequence $\mathcal{C}(W_n, W_{H_k})$ is not convergent. □

The proof above is not effective; it does not provide explicit inequalities between the different distance measures that we considered, like Theorems 4.6, 4.7, or 4.9. However, it has the following corollary. Let F be a simple graph, and let H_F be the soft-core weighted graph with $V(H_F) = V(F) = [q]$, nodeweights 1, and edgeweights

$$\beta_{i,j}(H_F) = \begin{cases} e, & \text{if } (i,j) \in E(F), \\ 1, & \text{otherwise.} \end{cases}$$

Let $\mathbf{u} = (1/q, \dots, 1/q) \in \mathbb{R}^q$.

COROLLARY 7.4. (a) *A sequence* (G_n) *of weighted graphs is right-convergent if and only if*

$$\frac{1}{\alpha_{G_n}^2} \ln \hom_{\mathbf{u}}(G_n, H_F)$$

is convergent for every simple graph F.

(b) *A sequence* (G_n) *of weighted graphs has convergent microcanonical ground state energies if and only if* $\mathcal{E}_{\mathbf{u}}(G_n, J)$ *converges for every symmetric* 0-1 *matrix* J *with* 0*'s in the diagonal.*

References

[1] Zs. Baranyai, On the factorization of the complete uniform hypergraph, in *Infinite and Finite Sets*, *Colloq. Math. Soc. János Bolyai*, **10**, North-Holland, Amsterdam, 1975, pp. 91–108. MR 0416986. Zbl 0306.05137.

[2] B. Bollobás, C. Borgs, J. Chayes, and O. Riordan, Percolation on dense graph sequences, *Ann. Probab.* **38** (2010), 150–183. MR 2599196. Zbl 1190.60090. http://dx.doi.org/10.1214/09-AOP478.

[3] C. Borgs, J. T. Chayes, L. Lovász, V. T. Sós, and K. Vesztergombi, Convergent sequences of dense graphs. I. Subgraph frequencies, metric properties and testing, *Adv. Math.* **219** (2008), 1801–1851. MR 2455626. Zbl 1213.05161. http://dx.doi.org/10.1016/j.aim.2008.07.008.

[4] C. Borgs, J. Chayes, and L. Lovász, Moments of two-variable functions and the uniqueness of graph limits, *Geom. Funct. Anal.* **19** (2010), 1597–1619. MR 2594615. Zbl 1223.05193. http://dx.doi.org/10.1007/s00039-010-0044-0.

[5] C. Borgs, J. Chayes, L. Lovász, V. T. Sós, and K. Vesztergombi, Counting graph homomorphisms, in *Topics in Discrete Mathematics*, *Algorithms Combin.* **26**, Springer-Verlag, New York, 2006, pp. 315–371. MR 2249277. Zbl 1129.05050. http://dx.doi.org/10.1007/3-540-33700-8_18.

[6] F. R. K. Chung, R. L. Graham, and R. M. Wilson, Quasi-random graphs, *Combinatorica* **9** (1989), 345–362. MR 1054011. Zbl 0715.05057. http://dx.doi.org/10.1007/BF02125347.

[7] A. Frieze and R. Kannan, Quick approximation to matrices and applications, *Combinatorica* **19** (1999), 175–220. MR 1723039. Zbl 0933.68061. http://dx.doi.org/10.1007/s004930050052.

[8] L. Lovász and V. T. Sós, Generalized quasirandom graphs, *J. Combin. Theory Ser. B* **98** (2008), 146–163. MR 2368030. Zbl 1127.05094. http://dx.doi.org/10.1016/j.jctb.2007.06.005.

[9] L. Lovász and B. Szegedy, Limits of dense graph sequences, *J. Combin. Theory Ser. B* **96** (2006), 933–957. MR 2274085. Zbl 1113.05092. http://dx.doi.org/10.1016/j.jctb.2006.05.002.

[10] ———, Szemerédi's lemma for the analyst, *Geom. Funct. Anal.* **17** (2007), 252–270. MR 2306658. Zbl 1123.46020. http://dx.doi.org/10.1007/s00039-007-0599-6.

[11] ———, Testing properties of graphs and functions, *Israel J. Math.* **178** (2010), 113–156. MR 2733066. Zbl 05823069. http://dx.doi.org/10.1007/s11856-010-0060-7.

[12] E. Szemerédi, Regular partitions of graphs, in *Problèmes Combinatoires et Théorie des Graphes* (Colloq. Internat. CNRS, Univ. Orsay, Orsay, 1976), *Colloq. Internat. CNRS* **260**, CNRS, Paris, 1978, pp. 399–401. MR 0540024. Zbl 0413.05055.

[13] A. Thomason, Pseudorandom graphs, in *Random Graphs '85* (Poznań, 1985), *North-Holland Math. Stud.* **144**, North-Holland, Amsterdam, 1987, pp. 307–331. MR 0930498. Zbl 0632.05045.

(Received: March 17, 2007)

Microsoft Research, Cambridge, MA
E-mail: borgs@microsoft.com

Microsoft Research, Cambridge, MA
E-mail: jchayes@microsoft.com

Department of Computer Science, Eötvös Loránd University,
Budapest, Hungary
E-mail: lovasz@cs.elte.hu

Alfréd Rényi Institute of Mathematics, Hungarian Academy of Sciences,
Budapest, Hungary
E-mail: sos@renyi.hu

Department of Computer Science, Eötvös Loránd University,
Budapest, Hungary
E-mail: veszterk@cs.elte.hu