



## GINZBURG–LANDAU FUNCTIONALS IN THE LARGE-GRAPH LIMIT

EDITH J. ZHANG, JAMES SCOTT, QIANG DU, AND MASON A. PORTER

**ABSTRACT.** Ginzburg–Landau (GL) functionals on graphs, which are relaxations of graph-cut functionals on graphs, have yielded a variety of insights in image segmentation and graph clustering. In this paper, we study large-graph limits of GL functionals by taking a functional-analytic view of graphs as nonlocal kernels. For a graph  $W_n$  with  $n$  nodes, the corresponding graph GL functional  $GL_\epsilon^{W_n}$  is an energy for functions on  $W_n$ . We minimize GL functionals on sequences of growing graphs that converge to functions called graphons. For such sequences of graphs, we show that the graph GL functional  $\Gamma$ -converges to a continuous and nonlocal functional that we call the *graphon GL functional*. We investigate the sharp-interface limits of the graph GL and graphon GL functionals, and we relate these limits to a nonlocal total-variation (TV) functional. We express the limiting GL functional in terms of Young measures and thereby obtain a probabilistic interpretation of the minimization problem in the large-graph limit. Finally, to develop intuition about graphon GL functionals, we determine the GL minimizer for several example families of graphons.

### 1. INTRODUCTION

The study of large graphs has become increasingly common in network analysis, with applications in sociology, systems biology, communications, epidemiology, and many other areas [17]. Because of their nice mathematical properties, graphons have become popular in both applications [35, 49, 55, 59, 69, 70] and theoretical studies [10, 13, 16, 26, 45]. In particular, graphons are well-suited to variational problems on graphs because it is possible to continuously deform a graphon and thereby calculate variations of graphon functionals [45, Section 16.2]. Graphons also connect graph functionals to nonlocal functionals [44, 74]. For instance, one can view a graphon total-variation (TV) functional (which we formulate in the present paper) as a nonlocal perimeter functional [20, 31, 40]. Graphons also have been employed in the derivation of mathematically rigorous mean-field approximations of dynamical systems on graphs [48, 49].

Optimization problems on graphs involve minimizing an energy functional that is defined on functions on graphs (i.e., functions that assign a value to each node of a graph). Optimization problems on large graphs are enormous combinatorial

---

2020 *Mathematics Subject Classification.* 05C63, 49J45, 82B24, 47J30.

*Key words and phrases.* Graphon, graph limit, cut norm, Ginzburg–Landau functional, Allen–Cahn functional,  $\Gamma$ -convergence, Young measure, stochastic block model.

optimization problems, and they often are intractable computationally. One example is the minimum-cut (i.e., “min-cut”) problem, which arises in applications such as community detection [53, 68] and image segmentation [58, 62, 63, 71, 72] and is related to the maximum-flow (i.e., “max-flow”) problem on networks [25, 34, 73].

The classical min-cut problem entails partitioning the set of nodes of a graph into two subsets,  $S$  and  $S^c$ , while minimizing the number of edges that one “cuts” to separate  $S$  and  $S^c$ . The min-cut problem involves minimizing a *graph-cut functional*, which is a special case of a graph Dirichlet energy [50] (see Remark 4.3). As the size of available data increases, increasingly large graphs (with millions of nodes or more) occur in applications. Analyzing large graphs is computationally expensive. For an  $n$ -node graph (i.e., a graph of “size”  $n$ ), the min-cut problem involves optimizing over  $2^n$  possible indicator functions, which each correspond to a possible partition  $\{S, S^c\}$ . This computational cost is a major obstacle in many applications.

One approach to simplify computations in the min-cut problem is to use the graph Ginzburg–Landau (GL) functional. The GL functional is a relaxation of a graph-cut functional that is defined on  $[-1, 1]$ -valued functions instead of on  $\{-1, 1\}$ -valued functions. The  $\Gamma$ -convergence of the graph GL functional to the graph TV functional (equivalently, to the graph-cut functional), which was proved in [66], justifies the use of the graph GL functional as a relaxation of the graph TV functional (i.e., graph-cut functional). However, although the graph GL functional is easier to minimize than the graph-cut functional, the minimization process still relies on approximate algorithms [7, 19, 51, 61, 67].

In the present paper, we further relax the graph-GL minimization problem to a continuum setting using a large-graph limit. That is, we evaluate the limiting minimization problem on functions that map  $((0, 1)) \rightarrow [-1, 1]$  rather than on functions that map  $\{1, \dots, n\} \rightarrow [-1, 1]$ . We use the idea of a *graphon* [45], which generalizes a graph’s adjacency matrix (which is a linear operator that acts on vectors in  $\mathbb{R}^n$ ) to a linear operator that acts on functions in  $L^q((0, 1))$ . This viewpoint allows us to treat problems that involve large graphs as functional-analytic problems.

Lovász’s original formulation of graphons [45] defined them as functions in  $L^\infty((0, 1)^2)$  that arise as limits of *dense sequences of graphs*<sup>1</sup> (i.e., sequences of graphs whose number of edges scales as  $\Theta(n^2)$  in the number  $n$  of nodes, which entails that the number of edges is bounded above by  $c_1 n^2$  and bounded below by  $c_2 n^2$  for some positive constants  $c_1$  and  $c_2$ ). Sequences of growing graphs in many applications and real-world situations are sparse [53], but applications of  $L^\infty$  graphons typically involve only mean-field models [5, 21]. To tackle this issue, researchers have defined  $L^p$  *graphons*, which arise as limits of sparse sequences of graphs [8, 11] (i.e., sequences of graphs whose number of edges increases as  $o(n^2)$ , which entails that the number of edges grows at a rate that is strictly less than  $n^2$ ). One obtains  $L^p$  graphons, which are associated with operators on  $L^q((0, 1)^2)$ , by normalizing graphs by their edge density. Researchers have also used traditional  $L^\infty$  graphons as limit objects of “very sparse” sequences of graphs [8, 9] (i.e., sequences of graphs with

---

<sup>1</sup>Density is a property of a sequence of graphs, rather than a property of a graph itself, because the definition of density is based on the rate at which the number of edges increases in comparison to the number of nodes.

bounded degree or bounded mean degree as  $n \rightarrow \infty$ ). See Section 3.2 for further discussion of sparse graphons.

To consider the convergence of graph GL functionals in the graphon limit, we use a central tool in variational calculus that is known as  $\Gamma$ -convergence. When the domain of the underlying function space is compact, the  $\Gamma$ -convergence of a sequence  $\{F_n\}$  of functionals  $F_n$  to a limiting functional  $F$  guarantees that the minimizers of  $F_n$  converge to the minimizer of  $F$  [15].

The original GL theory is a physical model for phase transitions in superconductors [23, 29]. The GL functional, which is sometimes called the Allen–Cahn functional or the Modica–Mortola functional in some applications [2, 28, 52], is

$$(1.1) \quad \text{GL}_\epsilon(u) = \epsilon \int_\Omega |\nabla u|^2 \, dx + \frac{1}{\epsilon} \int_\Omega \Phi(u(x)) \, dx, \quad u \in H^1(\Omega),$$

where  $\Omega \subset \mathbb{R}^d$  is a bounded and connected set,  $\Phi$  is a double-well potential, and  $H^1(\Omega) = W^{1,2}(\Omega) = \{f \in L^2(\Omega) : (\|f\|_2^2 + \|\nabla f\|_2^2)^{1/2} < \infty\}$  is a Sobolev space [33]. As  $\epsilon \rightarrow 0$ , the functional  $\text{GL}_\epsilon$   $\Gamma$ -converges to the TV functional (i.e., perimeter functional) [18]

$$(1.2) \quad \text{TV}(u) = \int_\Omega |\nabla u| \, dx.$$

We take inspiration from van Gennip and Bertozzi [66], who defined the graph GL functional, which is a discrete version of  $\text{GL}_\epsilon$  for functions  $u$  on graphs. See Section 4.3 for the definitions of the graph GL functional and other graph functionals. Van Gennip and Bertozzi proved  $\Gamma$ -convergence of the graph GL functional for a sequence of growing square-lattice graphs. They derived both a large-graph (i.e.,  $n \rightarrow \infty$ ) limit and a sharp-interface (i.e.,  $\epsilon \rightarrow 0$ ) limit of this functional. As  $n \rightarrow \infty$ , the square-lattice graph with grid size  $1/n$  converges to the region  $(0, 1)^2 \subset \mathbb{R}^2$ .

One can view each graph in a sequence of growing square-lattice graphs as a mesh approximation of the unit square. In the present paper, we generalize van Gennip and Bertozzi’s results to general sequences of growing graphs that converge to a graphon as  $n \rightarrow \infty$ . This convergence is guaranteed because the set of  $L^p$  graphons is (under mild conditions) compact with respect to the cut metric [12, Theorem 2.8].

We refer to the limit of sequences of graph GL functionals as a *graphon GL functional*. We show that the graphon GL functional  $\Gamma$ -converges to a nonlocal TV functional as  $\epsilon \rightarrow 0$ .

In Figure 1, we illustrate two different sequential  $\Gamma$ -limits of the graph GL functional (4.6). One of the sequential limits is an  $n \rightarrow \infty$  and then  $\epsilon \rightarrow 0$  limit; the other is an  $\epsilon \rightarrow 0$  and then  $n \rightarrow \infty$  limit. In Sections 4.3 and 4.5, we define the associated limiting functionals. Limit (1) was proven in [66, Theorem 3.1], and limit (2) follows from [16, Theorem 12]. To prove limit (3), we use an approach that is similar to the proof of [16, Theorem 12]. We significantly generalize [66, Theorem 5.2] in the sense that our graph limits are valid for general sequences of growing graphs, rather than only for a growing sequence of square-lattice graphs. Limit (4) resembles the classical Modica–Mortola limit [52] (which states that  $\text{GL}_\epsilon \xrightarrow{\Gamma} \text{TV}$ ), but it uses graphon versions of GL and TV functionals. We prove limit (4) for

graphons  $W \in L^\infty((0,1)^2)$ , which arise from dense sequences of graphs. Limit (4) is not defined for more general  $L^p$  graphons.

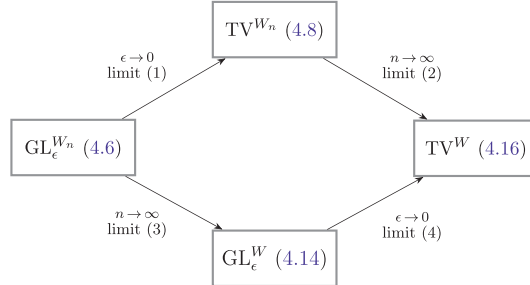


FIGURE 1. Four  $\Gamma$ -convergences of the graph GL functional with different orderings of the  $n \rightarrow \infty$  and  $\epsilon \rightarrow 0$  limits. The arrows indicate  $\Gamma$ -convergence. For definitions of the functionals, see Section 4.

Our results extend the study of graph GL functionals and graph TV functionals to graphon limits. We refer to these limiting functionals as “graphon GL functionals” and “graphon TV functionals”, respectively. We show that the graph GL-minimization and graph TV-minimization problems are consistent with the limiting graphon GL-minimization and graphon TV-minimization problems in the sense of  $\Gamma$ -convergence. That is, the minimizers of the functionals converge to the minimizers of the limiting functionals. We also show that the classical limit  $GL_\epsilon \xrightarrow{\Gamma} TV$  holds for graphons (i.e.,  $GL_\epsilon^W \xrightarrow{\Gamma} TV^W$ ).

Our paper proceeds as follows. In Section 2, we discuss a few key results from related work. In Section 3, we review graphons and formalize the notion of a large-graph limit. In Section 4, we review the GL and TV functionals on graphs and graphons. We also collect some results about Young measures, discuss our function spaces, and define the relevant types of convergence. In Section 5, we prove the sequential  $\Gamma$ -limit with  $\epsilon \rightarrow 0$  and then  $n \rightarrow \infty$ . In Section 6, we prove the sequential  $\Gamma$ -limit with  $n \rightarrow \infty$  and then  $\epsilon \rightarrow 0$ . In Section 7, we determine GL minimizers for a few example families of graphons. Finally, in Section 8, we summarize our results and discuss future research directions.

## 2. RELATED WORK

Our work is inspired by van Gennip and Bertozzi [66], who examined the same four limits as in Figure 1 and also the simultaneous limit  $\epsilon \rightarrow 0$ ,  $n \rightarrow \infty$ . They proved their  $n \rightarrow \infty$  limits specifically for sequences of growing square-lattice graphs. We restate their version of limit (1), which holds for all graphs, and we extend their versions of the limits (2)–(4) to a nonlocal graph-limit scenario in which the limiting GL and TV functionals are the graphon GL and TV functionals (4.14) and (4.16), respectively. Limits (2)–(4) hold for general sequences of graphs. As we discuss in Section 4.5, the limiting graphon functionals are nonlocal.

We adapt ideas from Braides et al. [16], who used Young measures to show that the graph-cut functional (which we define later in equation (3.4))  $\Gamma$ -converges to the

graphon-cut functional. We also use Young measures, which are generalizations of functions on graphs that assign a distribution of possible states (rather than a single fixed value) to each node of a graph. Young measures, which we define and discuss in Section 4.4, play a central role in our paper. Our methods and results are related more closely to the results of [16] than to those of [66]. A key difference from [16] arises from the fact that the domain of the graph-cut functional consists of finite-range functions, whereas the domains of the graph and graphon GL functionals consist of continuous-range functions.

García Trillos and Slepčev [38] also studied a local limit and, analogously to limit (2), obtained  $\Gamma$ -convergence of the perimeter functional for the  $n \rightarrow \infty$  limit of point clouds. The convergence of their graph functions, which are defined on point clouds, occurs in a metric space (which is called the  $TL^p$  space) that is characterized by optimal-transport maps. The  $TL^p$  space also has been employed to obtain results for other large-graph limits of point clouds (see, e.g., [1, 36, 37, 39]). A key difference between our paper and research on  $TL^p$  limits is that we do not require the limiting functional to be a local quantity, such as Euclidean perimeter or the usual Dirichlet energy. Instead, our limiting functionals are nonlocal limits with interaction potentials that are given by graphons. These functionals are “relaxations” in the sense that one can recover local functionals from them as special cases for certain graphons that have singularities on the line  $y = x$ .

The original theory of graphons [45] is concerned with dense sequences of graphs, and Braides et al. [16] also required the graphs in such sequences to be dense. However, most real-world graphs are sparse [8, 53], so this density requirement is a major limitation of much research on graphons. Thankfully, the theory of graphons has been extended to sparse sequences of graphs [11, 12], and our analysis allows sparse sequences of graphs that converge to  $L^p$  graphons. See Section 3.2 for more detail.

One recovers different types of TV functionals when taking the  $\epsilon \rightarrow 0$  limit of the classical GL, graph GL, and graphon GL functionals. In particular, we obtain a nonlocal, continuous TV functional in the  $\epsilon \rightarrow 0$  limit of the graphon GL functional. Nonlocal TV functionals have been useful in a variety of applications, especially in image processing [3, 4, 20, 31, 32, 40, 41, 43, 44]. They also are of theoretical interest because they generalize the notion of perimeter from objects (e.g., ones in  $\mathbb{R}^2$ ) with geometric regularity, on which one can compute gradients of functions, to less regular objects (such as objects in metric spaces) [47]. Fractional GL functionals are one well-studied type of nonlocal functional [57, 60, 64]. One broad class of fractional GL functionals consists of graphons of the form  $W(x, y) = \frac{1}{|x-y|^{1+2s}}$ , with  $s \in (0, 1)$ .

### 3. GRAPHONS, NORMS, AND CONVERGENCE

A “graphon”, which is a portmanteau of “graph” and “function”, is a bounded, measurable, and symmetric function  $W : \Omega^2 \rightarrow \mathbb{R}$ , where  $\Omega \subset \mathbb{R}^d$  is a connected and bounded domain. The set of graphons is  $\mathcal{W} = \{W : \Omega^2 \rightarrow \mathbb{R}\}$ . Throughout this paper, we take  $\Omega = (0, 1)$ . The closed interval  $[0, 1]$  is typically used in the graphon literature, whereas the open interval  $(0, 1)$  is typically used in functional analysis to avoid complications that relate to isolated points. In the present paper, it makes no difference because we always integrate  $W$ . We use the open interval to

be consistent with the conventions of functional analysis, which provides the main technical machinery in our paper.

We consider graphs that are weighted, undirected, and simple (i.e., there are no self-edges or multi-edges). Let  $W_n$  denote a graph with the node set  $[n] = \{1, \dots, n\}$ . The graph  $W_n$  has an associated adjacency matrix  $A^{(n)}$  with entries  $A_{ij}^{(n)}$ . We also associate the graph  $W_n$  with a function  $W_n$  that takes the constant value  $A_{ij}^{(n)}$  on the interval product  $I_i \times I_j$ , where

$$(3.1) \quad I_i = [(i-1)/n, i/n) \text{ for } i \in \{2, \dots, n\} \text{ and } I_1 = (0, 1/n).$$

The relationship between the adjacency matrix and the graphon is thus

$$(3.2) \quad W_n(x, y) = \left\{ A_{ij}^{(n)} \text{ for } (x, y) \in I_i \times I_j, i, j \in [n] \right\}.$$

We have thereby associated a graph with a step function by associating the  $i$ th node with the interval  $I_i$  and associating each edge  $(i, j)$  with the interval product  $I_i \times I_j$ . In this way, one can identify each graph  $W_n$  with a graphon, which we will also denote by  $W_n$  (see Remark 3.3). We refer to a graphon that corresponds to a graph as a *step graphon*.

In Figure 2, we show an example graph and its corresponding adjacency matrix and step graphon. By convention, the axes of the graphon begin at the upper left.

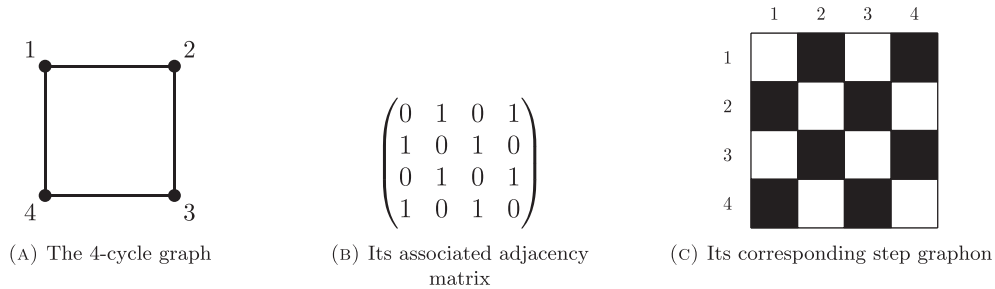


FIGURE 2. An example of a graph, its associated adjacency matrix, and its corresponding step graphon.

One can view any graphon  $W : (0, 1)^2 \rightarrow \mathbb{R}$  as a large-graph limit by thinking of  $(0, 1)$  as a set with infinitely-many nodes and taking  $W(x, y)$  to be the weight of the edge between nodes  $x \in (0, 1)$  and  $y \in (0, 1)$ . One can use graphons to represent both (1) families of graphs (by using a graphon as a probability distribution to generate graphs) and (2) limits of sequences of growing graphs. We employ the latter interpretation of graphons.

**3.1. Cut norm and cut convergence.** The graphon-cut norm, which is also known as the “cut norm” and is closely related to the graph-cut functional, specifies the topology of the space of graphons. Accordingly, both graphs and graphons converge to graphons with respect to cut norm. We introduce the graph-cut functional (which is sometimes called simply a “cut functional”) before introducing the cut norm.

**Definition 3.1** (graph cut). For a partition  $\{S, S^c\}$  of the nodes  $[n]$  of a graph with adjacency matrix  $A^{(n)}$ , the *graph cut* is the functional

$$(3.3) \quad \text{Cut}(S, S^c) = \sum_{i \in S, j \in S^c} A_{ij}^{(n)}.$$

Equivalently, we express the graph-cut functional in terms of a graph function  $u$  that takes values in the set  $\{-1, 1\}$  by using the expression

$$(3.4) \quad \text{Cut}(u) = \frac{1}{8} \sum_{i,j=1}^n A_{ij}^{(n)} |u_i - u_j|^2 \quad \text{with } u : [n] \rightarrow \{-1, 1\}.$$

We define the cut norm  $\|\cdot\|_{\square}$  on the space of graphons. The cut norm is closely related to the graph-cut functional, and it induces a metric that ensures that  $\mathcal{W}$  is a compact metric space [46]. Therefore, any bounded sequence  $\{W_n\}_{n=1}^{\infty}$  of graphons has a subsequence  $\{W_{n'}\}$  such that  $\|W_{n'} - W\|_{\square} \rightarrow 0$ .

**Definition 3.2.** The *cut norm* of a graphon  $W$  is

$$(3.5) \quad \|W\|_{\square} = \sup_{S \subseteq (0,1)} \int_{S \times S^c} W(x, y) \, dx \, dy.$$

Definition 3.2 highlights the similarity between the cut norm and the graph-cut functional. When a graphon is a step graphon  $W_n$  (see equation (3.2)), the cut norm (3.5) becomes the finite sum

$$\|W_n\|_{\square} = \frac{1}{n^2} \sup_{S \subseteq [n]} \sum_{i \in S, j \in S^c} A_{ij}^{(n)},$$

which is the maximum graph cut, normalized by  $1/n^2$ , over all partitions  $\{S, S^c\}$  of the nodes of  $W_n$ . There are other equivalent definitions of the cut norm (3.5) [42]. A particularly useful one for the present paper is

$$(3.6) \quad \|W\|_{\square} = \sup_{\phi, \psi \in L^{\infty}((0,1); [-1,1])} \int_0^1 \int_0^1 W(x, y) \phi(x) \psi(y) \, dx \, dy,$$

where

$$(3.7) \quad L^{\infty}((0, 1); [-1, 1]) = \{f : (0, 1) \rightarrow [-1, 1]\}.$$

By normalizing  $\phi \in L^{\infty}((0, 1))$  to  $\phi/\|\phi\|_{\infty} \in L^{\infty}((0, 1); [-1, 1])$ , we obtain the equivalent definition

$$(3.8) \quad \|W\|_{\square} = \sup_{\phi, \psi \in L^{\infty}((0,1))} \frac{1}{\|\phi\|_{\infty}} \frac{1}{\|\psi\|_{\infty}} \int_0^1 \int_0^1 \phi(x) \psi(y) W(x, y) \, dx \, dy.$$

**Remark 3.3.** One can identify any weighted graph with a step function that maps  $(0, 1)^2 \rightarrow \mathbb{R}$ , and vice versa. We use the notation  $W_n$  for both objects. In concert with the fact that step functions are dense in  $L^p$ , one can approximate any graphon arbitrarily closely in cut norm by a graph. See [14, Section 3.3], [42, Remark 4.6], and [46].

**Remark 3.4.** We use the notation “ $\xrightarrow{\square}$ ” to denote convergence in cut norm. Accordingly,  $W_n \xrightarrow{\square} W$  signifies that  $\|W_n - W\|_{\square} \rightarrow 0$ .

**Remark 3.5.** The cut norm is equivalent to the operator norm of the kernel operator that is induced by the graphon  $T_W(f) = \int_{\Omega} W(x, y)f(y) dy$ , which is a linear operator  $T_W : L^{\infty}((0, 1)) \rightarrow L^1((0, 1))$ . (The norms  $\|\cdot\|_A$  and  $\|\cdot\|_B$  are equivalent when  $c_1\|\cdot\|_A \leq \|\cdot\|_B \leq c_2\|\cdot\|_A$  for some constants  $c_1$  and  $c_2$ .) Additionally, people sometimes use the terms “graphon” and “kernel” interchangeably [42].

**Remark 3.6.** It is known that  $\|W\|_{\square} \leq \|W\|_1$ , where  $\|\cdot\|_1$  is the  $L^1(\Omega^2, \mathbb{R})$  norm, for any graphon  $W$  [42]. Additionally, for any step graphon with  $n$  steps,  $\|W_n\|_1 \leq \sqrt{2n}\|W_n\|_{\square}$  [45, Equation 8.15]. Consequently, for each step graphon, the  $L^1$  norm and cut norm are equivalent for finite  $n$ . However, this is not true when  $n \rightarrow \infty$ .

**3.2.  $L^p$  graphons.** The classical graphons that we have discussed thus far are known as “ $L^{\infty}$  graphons”. They are functions in the space

$$(3.9) \quad L^{\infty}((0, 1)^2) = \{W : \|W\|_{\infty} := \text{ess sup}_{x, y \in (0, 1)^2} |W(x, y)| < \infty\},$$

where the essential supremum  $\text{ess sup}$  equals the supremum up to a measure-0 set.

The set of  $L^{\infty}$  graphons [45] was built for dense sequences of graphs, which are sequences of graphs whose numbers of edges scale with the number  $n$  of nodes as  $\Theta(n^2)$  (i.e., they scale asymptotically in proportion to  $n^2$ ). Any sequence of graphs with  $o(n^2)$  (i.e., asymptotically strictly less than  $cn^2$  for some constant  $c > 0$ ) edges converges to the zero graphon  $W \equiv 0$  because the edge set is a set of measure 0 in the  $n \rightarrow \infty$  limit. One can see this because the Riemann sum

$$\int_0^1 \int_0^1 W_n(x, y) dx dy = \frac{1}{n^2} \sum_{i, j=1}^n A_{ij}^{(n)}$$

is nonzero in the  $n \rightarrow \infty$  limit only if the number of nonzero terms in the adjacency matrix  $A^{(n)} = \{A_{ij}^{(n)}\}_{i, j=1}^n$  is  $\Theta(n^2)$ .

It is common to use  $\mathcal{W}$  to denote the set of  $L^{\infty}$  graphons and to use  $\mathcal{W}_0 \subset \mathcal{W}$  to denote the set of graphons that take values in  $[0, 1]$ . However, one can identify any  $W \in \mathcal{W}$  with its normalized version  $W/\|W\|_{\infty} \in \mathcal{W}_0$ . Therefore, we refer to both  $\mathcal{W}$  and  $\mathcal{W}_0$  as  $L^{\infty}$  graphons to contrast them with  $L^p$  graphons, which we will define shortly.

Although most real-world graphs are sparse, they also often have some nodes with degrees that one expects to grow linearly with  $n$  [53]. For example, perhaps the total number of edges of a graph scales as  $\Omega(n)$  (and perhaps the graph has a heavy-tailed degree distribution). If we use the relationship (3.2) to define sequences of step graphons that correspond to sparse graphs, we obtain limiting graphons that are nonzero only on sets of measure 0. (This occurs because the number of edges is of order  $n$ , rather than of order  $n^2$ .) To extend the theory of graphons to such sparse sequences of graphs, Borgs et al. [11, 12] introduced  $L^p$  graphons (see [12, Definition 2.7]), which allow graphon theory to encompass a much wider variety of sparse-graph sequences.

The set of  $L^p$  graphons extends the set of  $L^{\infty}$  graphons to the space

$$(3.10) \quad L^p((0, 1)^2) = \left\{ W : \|W\|_p := \left( \int_0^1 \int_0^1 |W(x, y)|^p dx dy \right)^{\frac{1}{p}} < \infty \right\}$$

for  $p \geq 1$ . See [12, Theorem 2.8] for a characterization of the sequences of graphons that converge to an  $L^p$  graphon. In the present paper, we view graphons as functions in  $L^p((0, 1)^2)$  for  $p \in [1, \infty]$ , so we assume any necessary properties on step graphons that allow them to be  $L^p$  functions.

When  $\Omega$  is a bounded domain,  $L^p(\Omega) \subset L^q(\Omega)$  for  $p > q \geq 1$ . Therefore, the set of  $L^\infty$  graphons, which includes all dense graphs, is contained in the set of  $L^p$  graphons. Similarly, every  $L^p$  graphon is also an  $L^1$  graphon. In the proofs of limits (2) and (3) (see Sections 5.2 and 6.1, respectively), we assume that  $W_n$  and  $W$  are in  $L^1((0, 1)^2)$ . In the proof of limit (4) (see Section 6.3), we assume that  $W_n$  and  $W$  are in  $L^\infty((0, 1)^2)$ .

#### 4. FUNCTIONS AND FUNCTIONALS ON GRAPHS AND GRAPHONS

Because graphons are functions, it is natural to study them using ideas from functional analysis. Relevant notions include convergence, compactness, functionals on graphs and graphons, and  $\Gamma$ -convergence of those functionals. It is also appropriate to analyze the function spaces on which the functionals act.

We consider GL functionals and TV functionals, which act on “graph functions” (i.e., functions on graphs) and “graphon functions” (i.e., functions on graphons), respectively. Both the GL and TV functionals have classical, graph, and graphon versions. We defined the classical GL functional in (1.1) and the classical TV functional in (1.2). The graph GL and TV functionals are discrete and were defined in [6] and [63], respectively. We discuss them in Section 4.3. In Section 4.5, we define graph GL and TV functionals, which are continuous and involve Young measures.

The classical GL functional (1.1) is defined on the Sobolev space  $W^{1,2}((0, 1))$ , which is the space of functions in  $L^2((0, 1))$  with first derivatives that are also in  $L^2((0, 1))$  [33, Chapter 5]. The classical TV functional (1.2) is defined on  $W^{1,1}((0, 1))$ , which is defined analogously to  $W^{1,2}((0, 1))$  but uses the space  $L^1((0, 1))$ . We extend the domains of both functionals to  $L^\infty((0, 1))$  by imposing the value  $+\infty$  whenever they otherwise would be undefined.

**4.1. Function spaces.** We consider two types of function spaces: (1) spaces in which graphons live (i.e., function spaces of graphons) and (2) spaces of functions that are defined on graphons (i.e., function spaces of functions on graphons). As we discussed in Section 3.2, graphons are always  $L^p$  functions for  $p \in [1, \infty]$ . The functions on graphons that we consider are always functions in  $L^\infty((0, 1))$ .

The space of functions on  $n$ -node graphs is

$$(4.1) \quad \mathcal{V}^n = \{\hat{u} : [n] \rightarrow \mathbb{R}\}.$$

Each function  $\hat{u} \in \mathcal{V}^n$  has an associated step function  $u : (0, 1) \rightarrow \mathbb{R}$ . Let  $u(x) = \hat{u}(i)$  for  $x \in I_i$  for the intervals  $I_i$  that we defined in (3.1). This identification embeds  $\mathcal{V}^n$  into the space  $L^\infty((0, 1))$  of bounded functions. Henceforth, we identify both  $\hat{u}$  and  $u$  as  $u$ . We also consider the subset

$$(4.2) \quad \mathcal{V}_b^n = \{u : [n] \rightarrow \mathbb{R}, u_i \in \{\pm 1\} \text{ for all } i\}$$

of  $\mathcal{V}^n$  that consists of binary graph functions.

**4.2.  $\Gamma$ -convergence.** The notion of  $\Gamma$ -convergence of functionals is useful in optimization and the calculus of variations [15, 66]. In concert with a certain compactness property (which we will specify shortly), the  $\Gamma$ -convergence of a sequence of functionals guarantees the convergence of corresponding minimizers of the functionals.

**Definition 4.1.** Let  $X$  be a metric space, and consider a sequence of functionals  $F_n : X \rightarrow \mathbb{R} \cup \{\pm\infty\}$  for  $n \in \{1, 2, \dots\}$ . We say that  $F_n$   $\Gamma$ -converges to  $F : X \rightarrow \mathbb{R} \cup \{\pm\infty\}$ , which we denote by  $F_n \xrightarrow{\Gamma} F$ , with respect to  $u_n \rightarrow u$  if

- (i)  $\liminf_{n \rightarrow \infty} F_n(u_n) \geq F(u)$  for every sequence  $\{u_n\}$  such that  $u_n \rightarrow u$ ;
- (ii) there exists a sequence  $\{u_n\}_{n=1}^{\infty}$  such that  $\limsup_{n \rightarrow \infty} F_n(u_n) \leq F(u)$ .

**Remark 4.2.** The definition of  $\Gamma$ -convergence of functionals requires a choice of norm for the convergence  $u_n \rightarrow u$ . Therefore,  $\Gamma$ -convergence occurs *with respect to* the convergence  $u_n \rightarrow u$ . Additionally, throughout this paper, we abuse notation and use  $n$  both for indices and for the number of nodes in a graph.

If it is also true that any sequence  $\{u_n\}_{n=1}^{\infty}$  for which  $\{F_n(u_n)\}_{n=1}^{\infty}$  is uniformly bounded has a convergent subsequence  $\{u_{n_k}\}$ , then the corresponding minimizers of  $F_n$  converge to the minimizer(s) of  $F$ . This criterion, which we call the “compactness property”, is sometimes called the “equicoerciveness property” [66].

It is useful to be purposeful when choosing the metric under which  $u_n$  converges to  $u$ . When the convergence metric is stronger (where convergence in a stronger metric implies convergence in a weaker metric), it is easier to prove the lim inf inequality (i.e., the inequality in (i)) for  $\Gamma$ -convergence but harder to construct subsequences for the lim sup inequality (the inequality in (ii)) and compactness. Conversely, when the convergence metric is weaker, it is harder to prove the lim inf inequality but easier to construct a subsequence for the lim sup inequality. In the present paper, we prove  $\Gamma$ -convergence results with respect to narrow convergence of Young measures (see Definition 4.7).

**4.3. Graph functions and functionals.** Recall the classical GL functional (1.1), which is defined on functions  $u : (0, 1) \rightarrow \mathbb{R}$  by

$$(4.3) \quad \text{GL}_\epsilon(u) = \epsilon \int_0^1 |\nabla u|^2 dx + \frac{1}{\epsilon} \int_0^1 \Phi(u(x)) dx,$$

where  $\Phi$  is a double-well potential that has zeros at  $s = \pm 1$ . The double-well potential  $\Phi$  can take a general form (see [66, assumptions  $(W_1)$ – $(W_4)$ ]), but we use the standard choice

$$(4.4) \quad \Phi(s) = (s^2 - 1)^2.$$

The graph version  $\text{GL}_\epsilon^{W_n}$  of the GL functional is analogous to  $\text{GL}_\epsilon$ . It acts on  $u \in \mathcal{V}^n$ , which we treat as piecewise-constant functions  $u \in L^\infty((0, 1))$ . We replace the gradient term  $|\nabla u|^2$  by a finite-difference term that is weighted by the adjacency matrix, and we replace the double-well integral by a finite sum. (See [66, Section 2.2] for further discussion of graph analogues of calculus operators.) For  $u \in \mathcal{V}^n$ ,

we thus obtain

$$(4.5) \quad \text{GL}_\epsilon^{W_n}(u) = \frac{1}{n^2} \sum_{i,j=1}^n A_{ij}^{(n)} |u_i - u_j|^2 + \frac{1}{\epsilon n} \sum_{i=1}^n \Phi(u_i)$$

$$(4.6) \quad = \int_0^1 \int_0^1 W_n(x, y) |u(x) - u(y)|^2 dx dy + \frac{1}{\epsilon} \int_0^1 \Phi(u(x)) dx,$$

where (4.5) uses the adjacency matrix and (4.6) uses the definition (3.2) of the step graphon  $W_n$ . Similarly, the graph TV functional replaces the term  $|\nabla u|$  in (1.2) with a finite difference. It is finite only for binary functions. The graph TV functional is

$$(4.7) \quad \text{TV}^{W_n}(u) = \begin{cases} \sum_{i,j=1}^n A_{ij}^{(n)} |u_i - u_j| & \text{if } u \in \mathcal{V}_b^n \\ +\infty & \text{if } u \in \mathcal{V}^n \setminus \mathcal{V}_b^n \end{cases}$$

$$(4.8) \quad = \begin{cases} \int_0^1 \int_0^1 W_n(x, y) |u(x) - u(y)| dx dy & \text{if } u \in \mathcal{V}_b^n \\ +\infty & \text{if } u \in \mathcal{V}^n \setminus \mathcal{V}_b^n. \end{cases}$$

The Dirichlet energy

$$(4.9) \quad D(u) = \int_0^1 |\nabla u(x)|^2 dx$$

has a similar form as the graph-cut functional, but it acts on  $W^{1,2}((0, 1))$  functions. The Sobolev embedding of  $W^{1,2}((0, 1))$  into  $L^\infty((0, 1))$  allows us to define the Dirichlet energy  $D$  on all  $L^\infty((0, 1))$  functions by setting  $D(u) = +\infty$  for  $u \in L^\infty((0, 1)) \setminus W^{1,2}((0, 1))$ . The graph Dirichlet energy

$$(4.10) \quad D^{W_n}(u) = \begin{cases} \int_0^1 \int_0^1 W_n(x, y) |u(x) - u(y)|^2 dx dy & \text{if } u \in \mathcal{V}^n \\ +\infty & \text{if } u \in L^\infty((0, 1)) \setminus \mathcal{V}^n \end{cases}$$

acts on graph functions and replaces the gradient term in (4.9) with a finite difference.

**Remark 4.3.** The graph Dirichlet energy is a generalization of the graph-cut functional (3.4). If we restrict graph functions to the range  $\{-1, 1\}$ , then the graph Dirichlet energy (4.10) is equal to the graph-cut functional (3.4) (i.e.,  $D^{W_n}(u) = \text{Cut}(u)$  for all graph functions  $u$ ).

**4.4. Young measures and weak-\* convergence of functions.** In our study of  $\Gamma$ -convergence, we need to use Young measures [24], which extend the feasible set of the GL-minimization and TV-minimization problems to a space of measures. In general, the study of the convergence of functionals such as  $\int_0^1 f(v(x)) dx$  (for a given continuous and bounded functional  $f$ ) that act on  $v \in L^\infty((0, 1))$  presents a significant challenge. Because  $L^\infty((0, 1))$  is not a compact space, a sequence<sup>2</sup>  $\{v_n\} \in L^\infty((0, 1))$  may not have a limit in  $L^\infty((0, 1))$ . Therefore, a sequence  $\{f_n(v_n)\}$  may not have a limit  $f(v)$  in which  $v$  is a limit of  $v_n$ . Our functionals ( $\text{GL}_\epsilon^W$ ,  $\text{GL}_\epsilon^{W_n}$ , and so on) act on  $L^\infty((0, 1))$  and have the form  $\int_0^1 f(v(x)) dx$ . To

<sup>2</sup>In this sentence and many other times in the present paper, we abuse the notation “ $\in$ ” to signify that each element of a sequence (rather than the sequence as a single object) is an element of a set.

overcome this challenge, we extend these functionals to act on Young measures (rather than on  $L^\infty$  functions). Intuitively, our use of Young measures accommodates the rapidly oscillating minima of the graphon GL functionals. The minima can oscillate arbitrarily rapidly, and Young measures provide an effective limit of such oscillating functions when they are acted on by continuous bounded functions in the integrals. In Section 7, we further discuss the need for Young measures.

It is useful to review some definitions and properties that were presented in [16].

**Definition 4.4** (Young measure). A *Young measure*  $\nu$  on  $(0, 1) \times \mathbb{R}$  is a family  $\{\nu_x\}_{x \in (0, 1)}$  of probability measures, which are parametrized by  $x \in (0, 1)$ , such that the map

$$(4.11) \quad x \mapsto \int_{\mathbb{R}} f(\lambda) d\nu_x(\lambda)$$

is a Lebesgue-measurable function for every function  $f \in C^b(\mathbb{R})$ , where we use the notation  $C^b$  to signify continuous and bounded functions.

Intuitively,  $\nu_x$  is a “slice” of the Young measure  $\nu$  at the value  $x$ . One can think of  $\nu_x$  as analogous to the value of a function  $u(x)$ . (The value of a function at the point  $x$  is the “slice” of the function at  $x$ .) However, instead of assigning a value (as a function  $u$  does) to each point (i.e., node)  $x$  to designate the “state” of  $x$ , a Young measure  $\nu$  assigns a probability distribution  $\nu_x$  to the point  $x$ . For Young measures on graphons, one can interpret the distribution  $\nu_x$  as incorporating uncertainty into the state of node  $x$ .

Let  $\mathcal{Y}((0, 1), \mathbb{R})$  denote the set of all Young measures on  $(0, 1) \times \mathbb{R}$ . With the next definition, we see how  $\mathcal{Y}((0, 1), \mathbb{R})$  extends the set of  $L^\infty((0, 1))$  functions.

**Definition 4.5** (Young measure corresponding to a measurable function). A *Young measure corresponding to a Lebesgue-measurable function*  $u : (0, 1) \rightarrow \mathbb{R}$  is the family of delta measures

$$(4.12) \quad \{\nu_x\}_{x \in (0, 1)} = \{\delta_{u(x)}\}_{x \in (0, 1)}.$$

We refer to such measures as  $\delta$ -*Young measures*. Because graph functions are special cases of  $L^\infty$  functions, one can similarly define  $\delta$ -Young measures that correspond to graph functions.

For Young measures of the form (4.12), the map (4.11) is the evaluation map  $x \mapsto f(u(x))$ . This evaluation map is Lebesgue-measurable, as required for Definition 4.4, because of the Lebesgue-measurability of  $u$  and the continuity of  $f$ .

We now define the weak-\* topology (which is also known as the “weak-star topology” and the “ultraweak topology”) on the space  $L^\infty$ .

**Definition 4.6** (Weak-\* topology on the space  $L^\infty((0, 1))$ ). A sequence  $\{v_n\} \in L^\infty((0, 1))$  converges in the *weak-\* topology* to  $v \in L^\infty((0, 1))$  if

$$\lim_{n \rightarrow \infty} \int_0^1 v_n(x)g(x) dx = \int_0^1 v(x)g(x) dx$$

for any  $g \in L^1((0, 1))$ . We write that  $v_n \xrightarrow{*} v$  in  $L^\infty((0, 1))$ .

**Definition 4.7** (Narrow convergence of Young measures). A sequence  $\{\nu^n\} \in \mathcal{Y}((0,1), \mathbb{R})$  converges narrowly to  $\nu \in \mathcal{Y}((0,1), \mathbb{R})$  if the right-hand side of the map (4.11) converges in the weak-\* topology in  $L^\infty((0,1))$  for all continuous and bounded functions  $f \in C^b(\mathbb{R})$ . That is, for all  $x \in (0,1)$ , we have

$$(4.13) \quad \int_{\mathbb{R}} f(\lambda) d\nu_x^n(\lambda) \xrightarrow{*} \int_{\mathbb{R}} f(\lambda) d\nu_x(\lambda).$$

**Lemma 4.8** (Narrow convergence of product Young measures [16, Lemma 8]). Let  $\{\nu^n\}$  be a sequence of Young measures that converges narrowly to  $\nu \in \mathcal{Y}((0,1), \mathbb{R})$ . It is then guaranteed that the sequence  $\{\nu^n \otimes \nu^n\}$  of product measures on  $(0,1)^2 \times \mathbb{R}^2$  converges narrowly to  $\nu \otimes \nu \in \mathcal{Y}((0,1)^2, \mathbb{R}^2)$ . That is,  $d(\nu^n \otimes \nu^n)_{(x,y)}(\lambda, \mu) = d\nu_x^n(\lambda) d\nu_y^n(\mu)$  and

$$\iint_{\mathbb{R}^2} f(\lambda, \mu) d(\nu^n \otimes \nu^n)_{(x,y)}(\lambda, \mu) \xrightarrow{*} \iint_{\mathbb{R}^2} f(\lambda, \mu) d(\nu \otimes \nu)_{(x,y)}(\lambda, \mu)$$

for all  $f \in C^b(\mathbb{R}^2)$ , where  $d(\nu \otimes \nu)_{(x,y)}(\lambda, \mu) = d\nu_x(\lambda) d\nu_y(\mu)$ .

The following lemma states the key compactness property of  $\mathcal{Y}((0,1), \mathbb{R})$  that justifies the use of Young measures in our  $\Gamma$ -convergence results that involve graphon limits.

**Lemma 4.9** (Prohorov's Theorem [16, Theorem 9]). Let  $\{u_n\}_{n=1}^\infty$  be a bounded sequence in  $L^1((0,1))$ , and let  $\{\nu^n\}_{n=1}^\infty$  denote the sequence of corresponding Young measures (see equation (4.12)). There then exists a subsequence  $\{u_{n_k}\}_{k=1}^\infty$  and a Young measure  $\nu$  such that  $\nu^{n_k}$  converges narrowly to  $\nu$  as  $k \rightarrow \infty$ .

**4.5. Graphon functions and functionals.** We define the graphon GL functional

$$(4.14) \quad \begin{aligned} \text{GL}_\epsilon^W(\nu) &= \int_0^1 \int_0^1 W(x,y) \int_{\mathbb{R}^2} |\lambda - \mu|^2 d\nu_x(\lambda) d\nu_y(\mu) dx dy \\ &\quad + \frac{1}{\epsilon} \int_0^1 \int_{\mathbb{R}} \Phi(\lambda) d\nu_x(\lambda) dx. \end{aligned}$$

The first term of (4.14) is the graphon Dirichlet energy

$$(4.15) \quad D^W(\nu) = \int_0^1 \int_0^1 W(x,y) \int_{\mathbb{R}^2} |\lambda - \mu|^2 d\nu_x(\lambda) d\nu_y(\mu) dx dy,$$

which is the graphon analogue of the graph Dirichlet energy (4.10). For graphs  $W_n$ , we also use (4.15), which is more general than the graph Dirichlet energy (4.10). Equation (4.15) reduces to (4.10) when  $\nu$  is a Young measure that corresponds to a measurable function  $u$ .

The inner integral  $\int_{\mathbb{R}^2} |\lambda - \mu|^2 d\nu_x(\lambda) d\nu_y(\mu)$  is an expectation of  $|\lambda - \mu|^2$  with respect to the probability measures  $\nu_x$  and  $\nu_y$ . This is a probabilistic analogue of the term  $|u(x) - u(y)|^2$  in the graph GL functional (4.6). When the Young measures are  $\nu_x = \delta_{u(x)}$  and  $\nu_y = \delta_{u(y)}$ , we recover  $|u(x) - u(y)|^2$ .

The graphon TV functional, which also acts on  $\nu \in \mathcal{Y}((0,1), \mathbb{R})$ , is analogous to the graph TV functional, just as the graphon GL functional is analogous to the

graph GL functional. The graphon TV functional is

(4.16)

$$\text{TV}^W(\nu) = \begin{cases} 2 \int_0^1 \int_0^1 W(x, y) \int_{\mathbb{R}^2} |\lambda - \mu| d\nu_x(\lambda) d\nu_y(\mu) dx dy & \text{if } \nu \in \mathcal{Y}^b \\ +\infty & \text{if } \nu \in \mathcal{Y}((0, 1), \mathbb{R}) \setminus \mathcal{Y}^b, \end{cases}$$

where  $\mathcal{Y}^b$  denotes the set of Young measures  $\nu$  with support on  $\{-1, 1\}$  (i.e., the union of the supports of  $\{\nu_x\}_{x \in (0, 1)}$  is  $\{-1, 1\}$ ).

5. SEQUENTIAL LIMIT:  $\epsilon$  THEN  $n$  (I.E.,  $\text{GL}_\epsilon^{W_n} \xrightarrow{\Gamma} \text{TV}^{W_n} \xrightarrow{\Gamma} \text{TV}^W$ )

In this section, we discuss limits (1) and (2) from Figure 1. Limit (1) was already proven by van Gennip and Bertozzi [66, Theorem 3.1], and we state their associated results for completeness. Limit (2) was proven for sequences of growing square-lattice graphs in [66, Theorem 4.3] and for sequences of point clouds in [39]. We prove limit (2) for general sequences of weighted, undirected, and simple graphs. Our proof closely follows the proof of the main theorem of Braides et al. [16].

**5.1. Limit (1):**  $\text{GL}_\epsilon^{W_n} \xrightarrow{\Gamma} \text{TV}^{W_n}$ . We state two key results from [66] that pertain to limit (1). The cruxes of these results are that (1)  $\text{GL}_\epsilon^{W_n} \xrightarrow{\Gamma} \text{TV}^{W_n}$  and that (2) the functionals  $\text{GL}_\epsilon^{W_n}$  and  $\text{TV}^{W_n}$  are defined on a compact set of graph functions. Consequently, the minimizers of  $\text{GL}_\epsilon^{W_n}$  converge to the minimizers of  $\text{TV}^{W_n}$ . Propositions 5.1 and 5.2 hold for all undirected and weighted graphs  $W_n$ .

**Proposition 5.1** ( $\Gamma$ -convergence [66, Theorem 3.1]). *The graph GL functional (4.6)  $\Gamma$ -converges to the graph TV functional (4.8) as  $\epsilon \rightarrow 0$  with respect to  $u_n \rightarrow u$  in  $\mathcal{V}^n$ . That is,*

$$(5.1) \quad \text{GL}_\epsilon^{W_n} \xrightarrow{\Gamma} \text{TV}^{W_n} .$$

In concert with Proposition 5.1, the following compactness property of the set  $\mathcal{V}$  of functions guarantees that the minimizers of the  $\Gamma$ -converging functionals also converge.

**Proposition 5.2** (Compactness [66, Theorem 3.2]). *Let  $\{\epsilon_n\}_{n=1}^\infty \in \mathbb{R}_+$  be a sequence with  $\epsilon_n \rightarrow 0$  as  $n \rightarrow \infty$ , and let  $\{u_n\}_{n=1}^\infty \in \mathcal{V}^n$  be a sequence of graph functions for which there exists a constant  $C > 0$  such that  $\text{GL}_{\epsilon_n}^{W_n}(u_n) < C$  for all  $n \in \{1, \dots, \infty\}$ . There then exists a subsequence  $\{u_{n'}\}_{n'=1}^\infty \subseteq \{u_n\}_{n=1}^\infty$  and binary graph function  $u_\infty \in \mathcal{V}_b^n$  such that  $u_{n'} \rightarrow u_\infty$  as  $n' \rightarrow \infty$ .*

**5.2. Limit (2):**  $\text{TV}^{W_n} \xrightarrow{\Gamma} \text{TV}^W$ . We prove limit (2) as a corollary of [16, Theorem 12], which states that  $I_n \xrightarrow{\Gamma} I$  as  $W_n \xrightarrow{\square} W$ , where

$$(5.2) \quad I_n(\nu) = \int_0^1 \int_0^1 W_n(x, y) \int_{\mathbb{R}^2} f(\lambda, \mu) d\nu_x(\lambda) d\nu_y(\mu) dx dy ,$$

$$(5.3) \quad I(\nu) = \int_0^1 \int_0^1 W(x, y) \int_{\mathbb{R}^2} f(\lambda, \mu) d\nu_x(\lambda) d\nu_y(\mu) dx dy .$$

Limit (2) follows directly by using the integrand  $f(s, t) = |s - t|$  instead of  $f(s, t) = |s - t|^2$ , which is the integrand that was used in [16, Theorem 12].

**Theorem 5.3** ([16, Theorem 12]). *Let the functions  $f \in C^b((0, 1)^2)$  be bounded and continuous, and let  $u \in L^\infty((0, 1))$  and  $\nu \in \mathcal{Y}((0, 1), \mathbb{R})$ . Finally, let  $\{W_n\}_{n=1}^\infty$  be a dense sequence of graphs, with  $W_n \xrightarrow{\square} W \in \mathcal{W}_0$ . We then have that*

$$(5.4) \quad I_n \xrightarrow{\Gamma} I \text{ as } n \rightarrow \infty$$

with respect to the narrow convergence of measures in  $\mathcal{Y}((0, 1), \mathbb{R})$ .

**Corollary 5.4.** *With the choice  $f(s, t) = |s - t|$ , we have*

$$(5.5) \quad \text{TV}^{W_n} \xrightarrow{\Gamma} \text{TV}^W \text{ as } n \rightarrow \infty$$

with respect to narrow convergence of  $\nu^n$  to  $\nu$  in  $\mathcal{Y}((0, 1), \mathbb{R})$ .

**Proposition 5.5** (Compactness). *Let  $W_n \xrightarrow{\square} W$ , and let  $\{u_n\}_{n=1}^\infty \in \mathcal{V}^n$  be a sequence of graph functions such that  $\text{TV}^{W_n}(u_n) < M$  for all  $n \in \{1, \dots, \infty\}$  and some constant  $M > 0$ . There then exists a convergence subsequence  $\{u_{n'}\}$  such that the sequence  $\{\delta_{u_{n'}(x)}\}_{x \in (0, 1)}$  of corresponding  $\delta$ -Young measures converges to a limiting Young measure  $\nu$ .*

*Proof.* Because the sequence  $\{u_n\}$  consists of graph functions, it is bounded in  $L^1((0, 1))$ . We then apply Lemma 4.9 (i.e., Prohorov’s Theorem) to obtain the desired result.  $\square$

6. SEQUENTIAL LIMIT:  $n$  THEN  $\epsilon$  (I.E.,  $\text{GL}_\epsilon^{W_n} \xrightarrow{\Gamma} \text{GL}_\epsilon^W \xrightarrow{\Gamma} \text{TV}^W$ )

In this section, we prove two novel limits. Our main result is limit (3), which extends [16, Theorem 12]. We prove limit (4) only for  $L^\infty$  graphons, and we discuss a scaling issue for  $L^p$  graphons.

**6.1. Limit (3):**  $\text{GL}_\epsilon^{W_n} \xrightarrow{\Gamma} \text{GL}_\epsilon^W$ . There are three key steps in our proof of limit (3).

We first show that the GL-minimization problem is well-posed in the space  $L^\infty((0, 1); [-1, 1])$ . That is, we show that  $\arg \min_{L^\infty((0, 1))} \text{GL}_\epsilon^W \subset L^\infty((0, 1); [-1, 1])$ . It then suffices to consider functions  $u, u_n \in L^\infty((0, 1); [-1, 1])$ .

We then show that the graph Dirichlet energy  $\Gamma$ -converges to the graphon Dirichlet energy. Our proof includes ideas from the proof of [16, Lemma 11] for the  $\Gamma$ -convergence of graph-cut functionals, which are similar to Dirichlet energies, except that their domains consist of functions with finite-cardinality codomains.

Finally, we show that adding a double-well potential does not affect the  $\Gamma$ -convergence. Because each GL functional is a sum of a Dirichlet energy and a double-well potential, this fact yields  $\Gamma$ -convergence of the GL functionals.

**6.1.1. Well-posedness of  $\text{GL}_\epsilon^W$  in  $L^\infty((0, 1); [-1, 1])$ .** We show that the function  $u$  is bounded in magnitude by  $M = 1$ . Let  $u^M$  be the truncation

$$(6.1) \quad u^M(x) = \begin{cases} M & \text{if } u(x) > M \\ u(x) & \text{if } |u(x)| \leq M \\ -M & \text{if } u(x) < -M. \end{cases}$$

We show that  $\text{GL}_\epsilon^W(u) \geq \text{GL}_\epsilon^W(u^M)$  for any  $M > 1$  when  $W$  is any graphon. This implies that the minimizer of  $\text{GL}_\epsilon^W$  is in  $L^\infty((0, 1); [-1, 1])$ .

We separately show the well-posedness in  $L^\infty((0, 1); [-1, 1])$  of the Dirichlet energy and the double-well potential. The double-well potential  $\Phi(s) = (s^2 - 1)^2$  increases as  $s > 1$  increases and decreases as  $s < -1$  decreases, so

$$(6.2) \quad \int_0^1 \Phi(u) \geq \int_0^1 \Phi(u^M)$$

for  $M > 1$ . For the Dirichlet energy, let

$$(6.3) \quad S_M = \{x : u(x) \geq M\}$$

be the set of points  $x \in (0, 1)$  where  $u^M$  and  $u$  differ. To simplify our notation in this discussion, let

$$g(x, y) = |u(x) - u(y)|^2, \\ g^M(x, y) = |u^M(x) - u^M(y)|^2.$$

We want to show that

$$D^W(u) - D^W(u^M) = \int_0^1 \int_0^1 W(x, y) (g(x, y) - g^M(x, y)) \, dx \, dy \\ = \left( \int_{S_M} \int_{S_M^c} + \int_{S_M^c} \int_{S_M} + \int_{S_M} \int_{S_M} \right) W(x, y) (g(x, y) - g^M(x, y)) \, dx \, dy$$

is nonnegative. The integrals over  $S_M \times S_M^c$  and  $S_M^c \times S_M$  are equal because the integrand is symmetric. Both of these integrals are equal to

$$\int_{S_M} \int_{S_M^c} W(x, y) (|u(x) - u(y)|^2 - |M - u(y)|^2) \, dx \, dy.$$

Note that  $|u(x) - u(y)|^2 - |M - u(y)|^2 \geq 0$  because the function  $f(s) = |s - c|^2$  increases as  $s$  increases for  $s > c$ . Consequently, the integral over  $S_M \times S_M^c$  (and hence also the integral over  $S_M^c \times S_M$ ) is nonnegative. The integral over  $S_M \times S_M$  is

$$\int_{S_M} \int_{S_M} W(x, y) (g(x, y) - g^M(x, y)) \, dx \, dy \\ = \int_{S_M} \int_{S_M} W(x, y) (|u(x) - u(y)|^2 - |M - M|^2) \, dx \, dy \\ = \int_{S_M} \int_{S_M} W(x, y) |u(x) - u(y)|^2 \, dx \, dy \geq 0,$$

where we use the nonnegativity of the integrand in the last step to obtain nonnegativity of the integral.

We conclude that  $\int_0^1 \Phi(u(x)) \, dx \geq \int_0^1 \Phi(u^M(x)) \, dx$  and  $D^W(u(x)) \geq D^W(u^M(x))$ . Consequently,  $\text{GL}_\epsilon^W(u(x)) \geq \text{GL}_\epsilon^W(u^M(x))$ . Because  $M > 1$  is arbitrary, it follows that the GL minimizer takes values in  $[-1, 1]$ .

6.1.2. *Proof of limit (3).* The proof of limit (3) requires two key steps. First, we state and prove Theorem 6.3, which establishes  $\Gamma$ -convergence of the graph Dirichlet energy to the graphon Dirichlet energy in the large-graph limit. Second, we show that  $\Gamma$ -convergence holds even after adding a double-well potential to the graph and graphon Dirichlet energies. We thereby obtain  $\Gamma$ -convergence of the graph GL functional to the graphon GL functional in the large-graph limit.

Theorem 6.3, which guarantees  $\Gamma$ -convergence of the graph Dirichlet energy (4.10) to the graphon Dirichlet energy (4.15), resembles [16, Theorem 12], but it extends it in two important ways. First, it extends the domain from finite-codomain  $u_n$  to  $u_n \in L^\infty((0, 1))$ . Second, it extends graphons from  $W \in L^\infty((0, 1)^2)$  to  $W \in L^p((0, 1)^2)$ .

Lemma 6.1 provides the foundation for the proof of Theorem 6.3, which we use to prove limit (3) (see Corollary 6.6).

**Lemma 6.1.** *Let  $W_n \in L^p((0, 1)^2)$  for  $p \geq 1$ , and suppose that  $W_n \xrightarrow{\square} W$ . Let  $f \in C^b(\mathbb{R}^2)$  be a continuous and bounded function, and define the functional  $I_n : \mathcal{Y}((0, 1), \mathbb{R}) \rightarrow [0, \infty)$  as*

$$I_n(\nu^n) = \int_0^1 \int_0^1 W_n(x, y) \int_{\mathbb{R}^2} f(\lambda, \mu) d\nu_x^n(\lambda) d\nu_y^n(\mu) dx dy.$$

Let  $\{u_n\}_{n=1}^\infty \in \mathcal{V}^n$  be a sequence of graph functions such that  $\sup_n \|u_n\|_\infty < \infty$ . It is then the case that the sequence  $\{\nu_x^n\}_{x \in (0, 1)} = \{\delta_{u_n(x)}\}_{x \in (0, 1)} \in \mathcal{Y}((0, 1), \mathbb{R})$  of corresponding Young measures is precompact (i.e., its closure is compact) in the narrow topology. Moreover, any subsequence  $\{\nu_x^{n_k}\}$  of  $\{\nu_x^n\}$  with a corresponding limit point  $\nu$  satisfies

$$(6.4) \quad I_n(\nu^{n_k}) \rightarrow I(\nu)$$

pointwise, where the functional  $I : \mathcal{Y}((0, 1), \mathbb{R}) \rightarrow [0, \infty)$  is

$$I(\nu) = \int_0^1 \int_0^1 W(x, y) \int_{\mathbb{R}^2} f(\lambda, \mu) d\nu_x(\lambda) d\nu_y(\mu) dx dy.$$

*Proof.* To prove this result, we use the triangle inequality to break  $|I_n(\nu^n) - I(\nu)|$  into two parts. One part converges to 0 due to the weak convergence of  $g_n$  to  $g$ , and the other part converges to 0 due to the cut convergence  $W_n \xrightarrow{\square} W$ .

By Prohorov's Theorem (see Lemma 4.9), any sequence  $\{\nu^n\}$  of  $\delta$ -Young measures that corresponds to a sequence  $\{u_n\} \in L^\infty((0, 1))$  has a subsequence  $\{\nu^{n_k}\}$  that converges narrowly to a Young measure  $\nu$ . For the remainder of our paper, we simplify our notation by using  $\{\nu^n\}$  to denote the subsequence  $\{\nu^{n_k}\}$ . We denote the innermost integrals of the functionals by

$$g_n(x, y) = \int_{\mathbb{R}^2} f(\lambda, \mu) d\nu_x^n(\lambda) d\nu_y^n(\mu) = f(u_n(x), u_n(y)),$$

$$g(x, y) = \int_{\mathbb{R}^2} f(\lambda, \mu) d\nu_x(\lambda) d\nu_y(\mu).$$

We write  $I_n(\nu^n) = \int_{(0,1)^2} W_n(x,y)g_n(x,y) dx dy$  and  $I(\nu) = \int_{(0,1)^2} W(x,y)g(x,y) dx dy$ . The triangle inequality gives

$$\begin{aligned} |I_n(\nu^n) - I(\nu)| &= \left| \int_0^1 \int_0^1 (W_n g_n - W g) dx dy \right| \\ &\leq \left| \int_0^1 \int_0^1 (W_n - W) g_n dx dy \right| + \left| \int_0^1 \int_0^1 W (g_n - g) dx dy \right| \\ &\equiv \text{(I)} + \text{(II)}. \end{aligned}$$

We show that (II)  $\rightarrow 0$  using weak convergence of  $g_n$  to  $g$ . The narrow convergence of  $\nu^n$  to  $\nu$  implies that  $g_n$  converges in the weak-\* topology to  $g$  in  $L^\infty((0,1)^2)$  by Lemma 4.8. That is,

$$\int_{\mathbb{R}^2} f(\mu, \lambda) d\nu_x^n(\mu) d\nu_y^n(\lambda) \xrightarrow{*} \int_{\mathbb{R}^2} f(\mu, \lambda) d\nu_x(\mu) d\nu_y(\lambda) \text{ in } L^\infty((0,1)^2, \mathbb{R}).$$

Furthermore,  $W \in L^p((0,1)^2) \subseteq L^1((0,1)^2)$  for  $p \geq 1$ , so the definition of weak convergence in  $L^\infty$  implies that (II)  $\rightarrow 0$ .

We now show that (I)  $\rightarrow 0$  because  $W_n \xrightarrow{\square} W$ . To do this, we approximate  $g_n$  by polynomials to obtain a sum of terms that resembles the definition of cut convergence. This yields an expression that has the same form as the right-hand side of equation (3.8).

Let  $\{P_k\}_{k=1}^\infty : \mathbb{R}^2 \rightarrow \mathbb{R}$  be a sequence of polynomial functions such that

$$|f(a,b) - P_k(a,b)| \rightarrow 0 \text{ as } k \rightarrow \infty \text{ uniformly on } [-1,1]^2.$$

Such a polynomial exists because the set of polynomials is dense in  $L^\infty([-1,1])$ . Using the triangle inequality, we obtain

$$\begin{aligned} \text{(I)} &\leq \left| \int_0^1 \int_0^1 (W_n(x,y) - W(x,y))(f(u_n(x), u_n(y)) - P_k(u_n(x), u_n(y))) dy dx \right| \\ &\quad + \left| \int_0^1 \int_0^1 (W_n(x,y) - W(x,y))P_k(u_n(x), u_n(y)) dy dx \right| \\ &\leq \sup_{(a,b) \in [-1,1]^2} |f(a,b) - P_k(a,b)| (\|W_n\|_{L^1((0,1)^2)} + \|W\|_{L^1((0,1)^2)}) \\ &\quad + \left| \int_0^1 \int_0^1 (W_n(x,y) - W(x,y))P_k(u_n(x), u_n(y)) dy dx \right| \\ &\leq C \sup_{(a,b) \in [-1,1]^2} |f(a,b) - P_k(a,b)| \\ &\quad + \left| \int_0^1 \int_0^1 (W_n(x,y) - W(x,y))P_k(u_n(x), u_n(y)) dy dx \right| \end{aligned}$$

for some constant  $C$  because both the graphon  $W$  and the sequence  $\{W_n\}$  are bounded in  $L^1((0,1)^2)$ .

For any polynomial  $P_k(a, b) = \sum_{i,j=1}^k \alpha_{ij} a^i b^j$ , we use the boundedness of  $\|u_n\|_\infty$  (specifically, we use  $\|u_n\|_\infty \leq 1$ , which we showed in Section 6.1.1) to obtain

$$\begin{aligned} & \left| \int_0^1 \int_0^1 (W_n(x, y) - W(x, y)) P_k(u_n(x), u_n(y)) dx dy \right| \\ &= \left| \int_0^1 \int_0^1 (W_n(x, y) - W(x, y)) \sum_{i,j=1}^k \alpha_{ij} u_n^i(x) u_n^j(y) dx dy \right| \\ &= \left| \sum_{i,j=1}^k \alpha_{ij} \int_0^1 \int_0^1 (W_n(x, y) - W(x, y)) u_n^i(x) u_n^j(y) dx dy \right| \\ &\leq k \max_{i,j} \left\{ \alpha_{ij} \cdot \left| \int_0^1 \int_0^1 (W_n(x, y) - W(x, y)) u_n^i(x) u_n^j(y) dx dy \right| \right\} \\ &\leq C(P_k) \|W_n - W\|_\square \end{aligned}$$

for some constant  $C(P_k)$ .

In summary,

$$(I) \leq C \left( \sup_{(a,b) \in [-1,1]^2} |f(a, b) - P_k(a, b)| + \|W_n - W\|_\square \right),$$

where the constant  $C$  is independent of  $n$ . Choosing  $k$  sufficiently large and then letting  $n \rightarrow \infty$  implies that (I)  $\rightarrow 0$ .  $\square$

**Corollary 6.2.** *Let  $W_n \xrightarrow{\square} W$ , and let  $\{u_n\}_{n=1}^\infty \in \mathcal{V}^n$  be a sequence of graph functions such that  $\sup_n \|u_n\|_\infty < \infty$ . We then have that the sequence  $\{\nu_x^n\} \subset \mathcal{Y}((0, 1), \mathbb{R})$  of corresponding Young measures is precompact in the narrow topology. Moreover, any subsequence  $\{\nu_x^n\}$  and any limit point  $\nu$  satisfy*

$$(6.5) \quad D^{W_n}(\nu^n) \rightarrow D^W(\nu)$$

pointwise, where  $D^{W_n}$  and  $D^W$  are defined in (4.10) and (4.15), respectively.

*Proof.* Choose  $f(s, t) = |s - t|^2$  in Theorem 6.1.  $\square$

Corollary 6.2 extends [16, Lemma 11] by allowing  $\nu$  to be any Young measure and allowing each  $\nu^n = \{\delta_{u_n(x)}\}_{x \in (0,1)}$  to be any  $\delta$ -Young measure that corresponds to some function  $u_n \in L^\infty((0, 1))$ . In [16, Lemma 11],  $\nu^n$  must have support on a finite set of values, where the number of values is independent of  $n$ . The closure of the set of such Young measures is a strict subset of  $\mathcal{Y}((0, 1), \mathbb{R})$ . On the contrary, we allow  $\nu^n$  to be any  $\delta$ -Young measure. As we note below in the proof of Theorem 6.3, the closure of the set of  $\delta$ -Young measures in  $\mathcal{Y}((0, 1), \mathbb{R})$  is the set  $\mathcal{Y}((0, 1), \mathbb{R})$ . See [65, Proposition 8] for a proof of this fact.

**Theorem 6.3.** *Under the same assumptions as in Corollary 6.2, we have*

$$(6.6) \quad D^{W_n} \xrightarrow{\Gamma} D^W,$$

where we take  $\Gamma$ -convergence with respect to narrow convergence of Young measures.

*Proof.* To prove  $\Gamma$ -convergence, it suffices to prove the following two statements.

- (i) For every sequence  $\{\nu^n\}_{n=1}^\infty$  that converges narrowly to  $\nu$ , we have  $D(\nu) \leq \liminf_{n \rightarrow \infty} D^{W_n}(\nu^n)$ .
- (ii) There exists a sequence  $\{\nu^n\}_{n=1}^\infty$  that converges narrowly to  $\nu$  with  $D(\nu) \geq \limsup_{n \rightarrow \infty} D^{W_n}(\nu^n)$ .

Statement (i) follows from Corollary 6.2 because pointwise convergence holds for all sequences  $\{\nu^n\}$  that converge narrowly to  $\nu$ . To prove statement (ii), it suffices to show that there exists a sequence  $\{\nu^n\} = \{\delta_{u_n(x)}\}$  that converges narrowly to some  $\nu$  for any  $\nu \in \mathcal{Y}((0, 1))$ . Once we obtain this sequence, Corollary 6.2 yields inequality (ii). The existence of such a sequence is called the Fundamental Theorem of Young Measures, and it follows from the denseness (with respect to narrow convergence) of Young measures that correspond to measurable functions in the set of all Young measures [65, Proposition 8].  $\square$

**Remark 6.4.** Our proof of statement (ii) is more direct than the analogous proof by Braides et al. [16, Theorem 12], who constructed  $\nu^n$  carefully from a sequence of finite-valued functions  $u_n$ .

Proposition 6.5 guarantees that a sequence of double-well potentials on graph functions converges to a double-well potential on a Young measure. In Corollary 6.6, we use this result in concert with Theorem 6.3 to show that  $\text{GL}_\epsilon^{W_n} \xrightarrow{\Gamma} \text{GL}_\epsilon^W$ .

**Proposition 6.5.** *Let  $\{u_n\} \in L^\infty((0, 1))$  be a sequence of functions that are constant on the intervals  $I_i$  for  $i \in [n]$ , and let  $\Phi : \mathbb{R} \rightarrow \mathbb{R}$  be continuous and bounded. Let  $\nu$  be the Young measure that is the narrow limit of the sequence  $\{\nu^n\}$  of Young measures of the form  $\nu_x^n = \delta_{u_n(x)}$ . It then follows that*

$$(6.7) \quad \int_0^1 \Phi(u_n(x)) dx \rightarrow \int_0^1 \int_{\mathbb{R}} \Phi(\lambda) d\nu(\lambda) dx$$

*pointwise.*

*Proof.* The limit  $\nu$  exists by Lemma 4.9 (i.e., Prohorov’s Theorem). By the definition of narrow convergence, the narrow convergence of  $\nu^n$  to  $\nu$  implies that

$$(6.8) \quad \Phi(u_n(x)) = \int_{\mathbb{R}} \Phi(\lambda) d\nu_x^n(\lambda) \xrightarrow{*} \int_{\mathbb{R}} \Phi(\lambda) d\nu_x(\lambda) \text{ in } L^\infty((0, 1)).$$

For any test function  $\psi(x) \in L^1((0, 1))$ , the definition of weak-\* convergence implies that the sequence of integrals  $\int_0^1 \Phi(u_n(x))\psi(x) dx = \int_0^1 \int_{\mathbb{R}} \Phi(\lambda) d\nu_x^n(\lambda)\psi(x) dx$  converges to  $\int_0^1 \int_{\mathbb{R}} \Phi(\lambda)d\nu_x(\lambda)\psi(x) dx$  as  $n \rightarrow \infty$ . Using the test function  $\psi \equiv 1$  yields the desired result.  $\square$

The  $\Gamma$ -convergence of the graph GL functional (4.6) to the graphon limit as  $n \rightarrow \infty$  follows from the  $\Gamma$ -convergence of the graph Dirichlet energy (4.10) (see Theorem 6.3) and the pointwise convergence of the double-well potential (by Proposition 6.5). We state this convergence result formally in the next corollary.

**Corollary 6.6.** *Under the same assumptions as in Corollary 6.2, we have*

$$(6.9) \quad \text{GL}_\epsilon^{W_n} \xrightarrow{\Gamma} \text{GL}_\epsilon^W$$

*as  $n \rightarrow \infty$  with respect to narrow convergence of Young measures.*

*Proof.* Let  $\text{GL}_\epsilon^{W_n} = F_n + G_n$  and  $\text{GL}_\epsilon^W = F + G$ , where  $F_n$  and  $F$  denote the graph and graphon Dirichlet energies, respectively, and  $G_n$  and  $G$  correspond to the graph and graphon double-well energies, respectively. Given the  $\Gamma$ -convergence (see Theorem 6.3) of the graph Dirichlet energy (4.10) (that is,  $F_n \xrightarrow{\Gamma} F$ ) and the fact that  $G_n \rightarrow G$  (i.e., the pointwise convergence of the double-well potential that follows from Proposition 6.5), we invoke the fact that  $\Gamma$ -convergence still holds under a continuous perturbation (see [15, Remark 2.2]). Consequently, we obtain  $F_n + G_n \xrightarrow{\Gamma} F + G$  with respect to narrow convergence of Young measures.  $\square$

**6.2. The issue of  $\epsilon$ -scaling.** A key property of the classical GL functional (1.1) is its  $\Gamma$ -convergence to the TV functional in the  $\epsilon \rightarrow 0$  limit [52]. Consider

$$(6.10) \quad \text{GL}_\epsilon(u) = \epsilon \int_0^1 |\nabla u|^2 dx + \frac{1}{\epsilon} \int_0^1 \Phi(u(x)) dx,$$

which is equation (1.1) (but we show it again for convenience). As  $\epsilon$  shrinks, the double-well term becomes larger due to its prefactor  $1/\epsilon$ . This larger contribution from the double-well term encourages narrower regions for  $u$  to jump between  $-1$  and  $1$ . However, steeper jumps of  $u$  contribute more to the Dirichlet energy  $\int_0^1 |\nabla u|^2 dx$ . As the interface size  $\epsilon$  shrinks to  $0$ , the contribution from each jump grows to  $\infty$ . The prefactors that ensure that the Dirichlet energy and double-well potential remain  $O(1)$  as  $\epsilon \rightarrow 0$  are  $\epsilon$  and  $1/\epsilon$ , respectively. One can see the associated scaling limits by substituting  $x \mapsto x/\epsilon$  into the classical GL functional (1.1). Compare (6.10) to the graph GL functional

$$(6.11) \quad \text{GL}_\epsilon^{W_n} = \int_0^1 \int_0^1 W_n(x, y) |u(x) - u(y)|^2 dx dy + \frac{1}{\epsilon} \int_0^1 \Phi(u(x)) dx,$$

which is equation (4.6). The graph GL functional (6.11) does not require the prefactor  $\epsilon$  in the graph Dirichlet energy because a graph inherently has no infinitesimal spatial limit. Each jump of  $u$  from  $-1$  to  $1$  (and vice versa) contributes a finite amount to the graph Dirichlet energy, so both terms of the graph GL functional remain  $O(1)$  even when  $\epsilon \rightarrow 0$ . As with the classical GL functional, the double-well potential is a penalty term that enforces  $u$  to be binary.

It is natural to ask how one should scale the graphon GL functional  $\text{GL}_\epsilon^W$ , which is similar to (6.11) in the sense that one computes differences of  $u$  (rather than derivatives of  $u$ ). Note that  $W(x, y) |u(x) - u(y)|^2 = 4W(x, y)$  when  $u(x) = 1$  and  $u(y) = -1$  (or vice versa). Because  $4W(x, y)$  is finite when  $W \in L^\infty((0, 1)^2)$ , one does not need an  $\epsilon$  prefactor for  $L^\infty$  graphons.

If the graphon Dirichlet energy (4.15) is unbounded, as is the case for  $L^p$  graphons, the scaling rate in  $\epsilon$  depends on the choice of  $W$ . The reason for this is that it is necessary to scale the unbounded graphon Dirichlet energy  $D^W$  at an appropriate rate to ensure that both the Dirichlet energy and the double-well potential remain  $O(1)$ . The more singular  $W$  is, the faster one needs to scale down the graphon Dirichlet energy (and vice versa). Formalizing the relationship between the singularity strength of  $W$  and the  $\epsilon$ -scaling is an interesting topic for future work.

**6.3. Limit (4):**  $\text{GL}_\epsilon^W \xrightarrow{\Gamma} \text{TV}^W$  as  $\epsilon \rightarrow 0$  for  $W \in L^\infty$ . We prove a version of the classical limit  $\text{GL}_\epsilon \xrightarrow{\Gamma} \text{TV}$  for graphon GL and graphon TV functionals. We consider only  $L^\infty$  graphons because it is difficult to determine the correct  $\epsilon$ -scaling for general  $L^p$  graphons.

**Theorem 6.7.** *Suppose that  $W \in L^\infty((0, 1)^2)$ . As  $\epsilon \rightarrow \infty$ , we then have*

$$(6.12) \quad \text{GL}_\epsilon^W \xrightarrow{\Gamma} \text{TV}^W$$

*with respect to narrow convergence of Young measures.*

To prove Theorem 6.7, we follow a strategy that resembles the proof of [66, Theorem 3.1]. Suppose that a sequence  $\{\nu^n\} \in \mathcal{Y}((0, 1), \mathbb{R})$  of Young measures converges narrowly to  $\nu \in \mathcal{Y}((0, 1), \mathbb{R})$ . Lemma 6.8 shows that the double-well potential  $\Gamma$ -converges either to 0 or to  $\infty$  as  $\nu^n \rightarrow \nu$  narrowly.

**Lemma 6.8.** *Let  $\Phi : \mathbb{R} \rightarrow \mathbb{R}$  be a continuous and bounded potential function. Consider the functionals  $E_\epsilon$  and  $E_0$  on  $\mathcal{Y}((0, 1), \mathbb{R})$  that are given by*

$$(6.13) \quad E_\epsilon(\nu) = \begin{cases} \frac{1}{\epsilon} \int_0^1 \int_{\mathbb{R}} \Phi(\lambda) d\nu_x(\lambda) dx & \text{if } \nu \in \mathcal{Y}^b \\ +\infty & \text{otherwise} \end{cases}$$

and

$$(6.14) \quad E_0(\nu) = \begin{cases} 0 & \text{if } \nu \in \mathcal{Y}^b \\ +\infty & \text{otherwise,} \end{cases}$$

where we recall that  $\mathcal{Y}^b$  denotes the set of Young measures  $\nu$  with support on  $\{-1, 1\}$ . As  $\epsilon \rightarrow 0$ , we have

$$(6.15) \quad E_\epsilon \xrightarrow{\Gamma} E_0$$

*with respect to narrow convergence of Young measures.*

*Proof of Lemma 6.8.* Consider sequences  $\{\epsilon_n\}_{n=1}^\infty$  and  $\{\nu^n\}_{n=1}^\infty$  such that  $\epsilon_n \rightarrow 0$  and  $\nu^n \rightarrow \nu$  narrowly for some  $\nu \in \mathcal{Y}((0, 1), \mathbb{R})$ . We verify the lim inf inequality (i) and the lim sup inequality (ii) of Definition 4.1.

For  $\nu \in \mathcal{Y}^b$ , we have  $E_0(\nu) = 0$ . Therefore, the lim inf inequality  $E_0(\nu) \leq \liminf_{n \rightarrow \infty} E_{\epsilon_n}(\nu^n)$  holds because  $E_{\epsilon_n}$  is nonnegative. For  $\nu \in \mathcal{Y}((0, 1), \mathbb{R}) \setminus \mathcal{Y}^b$ , we will show that  $\nu^n \notin \mathcal{Y}^b$  for sufficiently large  $n$ . To do this, we prove the contrapositive statement. Suppose that  $\nu^n \in \mathcal{Y}^b$  for arbitrarily large  $n$ . The limiting Young measure  $\nu$  then cannot have support on any point  $y \in \mathbb{R} \setminus \{-1, 1\}$ . Consider a continuous and bounded function  $f$  with support on a small interval around the point  $y$  that does not include the points 1 or  $-1$ . The narrow convergence of  $\nu^n$  to  $\nu$  implies that  $\int_{\mathbb{R}} f(\lambda) d\nu_x^n(\lambda) \xrightarrow{*} \int_{\mathbb{R}} f(\lambda) d\nu_x(\lambda)$ . However,  $\int_{\mathbb{R}} f(\lambda) d\nu_x(\lambda) \geq 0$  for all  $x$  and  $\int_{\mathbb{R}} f(\lambda) d\nu_x^n(\lambda) = 0$  for all  $x$ . Therefore, for  $\nu \in \mathcal{Y}((0, 1), \mathbb{R}) \setminus \mathcal{Y}^b$ , there must be some  $x$  such that  $\nu_x^n$  has support in  $\mathbb{R} \setminus \{-1, 1\}$ . Therefore,  $\int_{\mathbb{R}} \Phi(\lambda) d\nu_x^n(\lambda) > 0$  for that  $x$ , so the lim inf inequality holds. Consequently,

$$(6.16) \quad \liminf_{n \rightarrow \infty} E_{\epsilon_n}(\nu^n) \geq \int_0^1 \liminf_{n \rightarrow \infty} \frac{1}{\epsilon_n} \int_{\mathbb{R}} \Phi(\lambda) d\nu_x^n(\lambda) dx = \infty = E_0(\nu).$$

The first inequality holds by Fatou's Lemma. The first equality holds because  $\epsilon_n \rightarrow 0$  and the integrand is uniformly bounded away from 0. The second equality holds because  $\nu^n \notin \mathcal{Y}^b$ . We now verify the lim sup inequality  $\limsup_{n \rightarrow \infty} E_{\epsilon_n}(\nu^n) \leq E_0(\nu)$ . Suppose that  $\nu \in \mathcal{Y}^b$ . We choose  $\nu^n = \nu$  for all  $n$ , so  $E_0(\nu) = 0 = \limsup_{n \rightarrow \infty} E_{\epsilon_n}(\nu^n)$ . With  $\nu \in \mathcal{Y}((0, 1), \mathbb{R}) \setminus \mathcal{Y}^b$ , it follows that  $E_0(\nu) = \infty$  and thus that the lim sup inequality  $\limsup_{n \rightarrow \infty} E_{\epsilon_n}(\nu^n) \leq E_0(\nu)$  holds for any sequence  $\{\nu^n\}$  that converges narrowly to  $\nu$ .  $\square$

We now prove Theorem 6.7.

*Proof of Theorem 6.7.* To prove the desired result, we use the fact that  $\Gamma$ -convergence persists under a continuous perturbation. Note that  $\text{GL}_\epsilon^W(\nu) = D^W(\nu) + E_\epsilon(\nu)$ , where  $D^W$  is the graphon Dirichlet energy (4.15) and  $E_\epsilon$  is the double-well potential (6.13). The key ingredients in our proof are the guarantee that the double-well potential  $\Gamma$ -converges (by Lemma 6.8) and the fact that the Dirichlet energy (4.15) is continuous in  $\nu$ . The Dirichlet energy is continuous because  $\nu^n$  converges narrowly to  $\nu$ , which implies (by Definition (4.7)) for  $W \in L^1$  that

$$\begin{aligned} & \int_0^1 \int_0^1 W(x, y) \int_{\mathbb{R}^2} |\lambda - \mu|^2 d\nu_x^n(\lambda) d\nu_y^n(\mu) dx dy \\ & \rightarrow \int_0^1 \int_0^1 W(x, y) \int_{\mathbb{R}^2} |\lambda - \mu|^2 d\nu_x(\lambda) d\nu_y(\mu) dx dy \end{aligned}$$

pointwise. Therefore,

$$\text{GL}_\epsilon^W = D^W + \frac{1}{\epsilon} \int_0^1 \int_{\mathbb{R}} \Phi(\lambda) d\nu_x(\lambda) dx$$

is a continuous perturbation of the double-well potential. Because  $\Gamma$ -convergence is stable under continuous perturbations, equation (6.15) implies that

$$(6.17) \quad \text{GL}_\epsilon^W \xrightarrow{\Gamma} D^W + \begin{cases} 0 & \text{if } \nu \in \mathcal{Y}^b \\ +\infty & \text{otherwise} \end{cases}$$

as  $\epsilon \rightarrow 0$ .

The right-hand side of (6.17) is equal to  $\text{TV}^W(\nu)$ . We quickly verify this statement. If  $\nu \in \mathcal{Y}((0, 1), \mathbb{R}) \setminus \mathcal{Y}^b$ , then both the  $\Gamma$ -limit in (6.17) and  $\text{TV}^W(\nu)$  are  $\infty$ . If  $\nu \in \mathcal{Y}^b$ , then  $\lambda$  and  $\mu$  can only take the values  $\pm 1$ . Therefore,  $\lambda$  and  $\mu$  can only have support in  $\{\pm 1\}$ , so  $|\lambda - \mu|^2$  equals either  $2|\lambda - \mu|$  or 0.  $\square$

## 7. GL MINIMIZERS FOR SEVERAL EXAMPLES

It is informative to determine the GL minimizers for some families of graphons. We characterize the minimizers of several graph GL functionals, and we illustrate the resulting Young measure in the graphon limit. Taking the  $\epsilon \rightarrow 0$  limit of a GL minimizer then allows us to infer the associated TV minimizer.

For more general situations than our simple examples, it is difficult to obtain analytical characterizations of graph and graphon GL minimizers. One can seek more general Young-measure minimizers using numerical approximations; see [22, 54] for possible approaches. We leave such endeavors to future work.

The Young measure  $\nu_x = \delta_1$  for all  $x$  (i.e., the constant function  $u(x) = 1$ ) is a trivial minimizer of the graphon GL functional (4.14) for all graphons  $W$ . Similarly,  $\nu_x = \delta_{-1}$  is also a minimizer for all  $W$ . These are minimizers because the double-well potential  $\Phi(s)$  is 0 for  $s = \pm 1$  and the Dirichlet energy term is 0 for constant functions. The constant function  $u \equiv 1$  (which corresponds to the Young measure  $\nu_x \equiv \delta_1$ ) is also a trivial minimizer of the graph GL functional (4.6).

The trivial minimizer is the only minimizer of the GL functional because the GL functional is nonnegative and equals 0 only for the trivial minimizer. To make the minimization problem nontrivial, we use a mean-value constraint. For the graphon GL-minimization problem, we thus impose the mean value

$$(7.1) \quad \int_0^1 \int_{\mathbb{R}} \lambda d\nu_x(\lambda) dx = c$$

for a specified constant  $c \in (-1, 1)$ . The analogous constraint for the graph GL-minimization problem is

$$(7.2) \quad \frac{1}{n} \sum_{i=1}^n u_i = c.$$

This constraint entails that the graph function  $u$  has a mean value of  $c$ . The double-well potential  $\Phi$  forces  $u$  to take values near  $+1$  and  $-1$ , and the mean-value constraint prevents  $u$  from settling at only one of the values  $+1$  or  $-1$ . Consequently,  $u$  exhibits a *phase separation*, where the values  $+1$  and  $-1$  correspond to two different “phases” [28]. We consider  $c \in (-1, 1)$  in Section 7.1 and consider  $c = 0$  in Section 7.2.

Throughout this section, we denote graph functions by  $u$  (rather than by  $u_n$ , as we did previously). Additionally, we denote both functions and Young measures on graphons by  $\nu$ .

**7.1. The constant graphon.** The constant graphon is the large-graph limit of the complete graph (for which  $W \equiv 1$ ), Erdős–Rényi (ER) graphs (for which  $W \equiv p \in (0, 1)$ ) [45, Section 10.1], and some growing preferential-attachment graphs (including Barabási–Albert (BA) graphs, for which  $W \equiv p \in (0, 1)$ ) [45, Example 11.44 and Proposition 11.45].

Consider the graph GL functional (4.6) on the  $n$ -node complete graph, and let  $A_{ij}^{(n)} = p$  for all  $i$  and  $j$ . We thus have

$$(7.3) \quad \begin{aligned} \text{GL}_\epsilon^{W_n}(u) &= \frac{1}{n^2} \sum_{i,j=1}^n p(u_i - u_j)^2 + \frac{1}{\epsilon n} \sum_{i=1}^n \Phi(u_i) \\ &= \frac{p}{n^2} \left( 2n \sum_{i=1}^n u_i^2 - 2 \sum_{i,j=1}^n u_i u_j \right) + \frac{1}{\epsilon n} \sum_{i=1}^n (u_i^4 - 2u_i^2 + 1) \\ &= \frac{1}{\epsilon n} \sum_{i=1}^n (u_i^4 - (2 - 2\epsilon p)u_i^2) - \frac{2p}{n^2} \left( \sum_{i=1}^n u_i \right)^2 + \frac{1}{\epsilon} \\ &= \frac{1}{\epsilon n} \sum_{i=1}^n (u_i^2 - (1 - \epsilon p))^2 - \frac{2p}{n^2} \left( \sum_{i=1}^n u_i \right)^2 - \frac{1}{\epsilon} (1 - \epsilon p)^2 + \frac{1}{\epsilon}. \end{aligned}$$

Fix  $\epsilon < 1$ . By removing constant terms, we see that minimizing (7.3) subject to the mean-value constraint (7.2) is equivalent to minimizing the energy

$$(7.4) \quad E_n(v) := \sum_{i=1}^n (v_i^2 - 1)^2,$$

where

$$(7.5) \quad v_i = \frac{u_i}{\sqrt{1 - \epsilon p}}.$$

With the change of variables (7.5), the mean-value constraint (7.2) becomes

$$(7.6) \quad \frac{1}{n} \sum_{i=1}^n v_i = \frac{c}{\sqrt{1 - \epsilon p}}.$$

Minimizing (7.4) subject to the mean-value constraint (7.6) leads to the Euler–Lagrange equations

$$(7.7) \quad v_i^3 - v_i = \tau \quad \text{and} \quad \frac{1}{n} \sum_{i=1}^n v_i = \frac{c}{\sqrt{1 - \epsilon p}},$$

where  $\tau$  is a Lagrange multiplier that is associated with the mean-value constraint.

**Proposition 7.1** (Characterization of GL minimizers for the complete graph with constant edge weight). *Let  $W_n \equiv p$ , and suppose that  $|c| < \sqrt{1 - \epsilon p}$  for sufficiently small  $\epsilon$ . The minimizers of  $\text{GL}_\epsilon^{W_n}$  under the mean-value constraint (7.2) are functions  $u$  with at most three different values in their codomains. These values are in the set  $\{z_{\pm,n}\}$ , where  $z_{\pm,n} = \pm\sqrt{1 - \epsilon p} + O(1/\sqrt{n})$ , except for at most one node on which  $u$  takes a value that is  $O(1/\sqrt{n})$ .*

As  $n \rightarrow \infty$  and  $\epsilon \rightarrow 0$ , the values  $\{\pm\sqrt{1 - \epsilon p} + O(1/\sqrt{n})\}$  of the minimizers approach  $\{\pm 1\}$ . The minimizers are not unique because the energy (7.4) and the constraint (7.6) (and hence also (7.2)) are invariant with respect to permutations of the subscripts.

We now prove Proposition 7.1.

*Proof.* With the change of variables (7.5), it suffices to describe minimizers  $v$  of the energy (7.4).

We first show that the energy functional (7.4) for a minimizer  $v$  is bounded above by 1. Consider a candidate minimizer  $\tilde{v}$  with  $\tilde{v}_i \in \{\pm 1\}$  for all  $i \in [n]$  that takes the value  $+1$  on  $\frac{n}{2}(1 + \frac{c}{\sqrt{1 - \epsilon p}})$  of the nodes and the value  $-1$  on  $\frac{n}{2}(1 - \frac{c}{\sqrt{1 - \epsilon p}})$  of the nodes. If these numbers of positive and negative values are not integers, we round the numbers; the two numbers sum to  $n$ , so one rounds up and one rounds down. (It does not matter which one rounds up and which one rounds down.) There may be one node  $j$  that does not have an assigned value. On node  $j$ , the candidate minimizer  $\tilde{v}$  takes whichever value allows it to satisfy the mean-value constraint (7.6). Namely,  $\tilde{v}_j = \frac{nc}{\sqrt{1 - \epsilon p}} - \text{round}(\frac{nc}{\sqrt{1 - \epsilon p}})$ , where “round” maps a real number to the nearest integer and 0.5 rounds up. Because  $|\tilde{v}_j| < 1$ , the energy is  $E_n(\tilde{v}) \leq 1$  for this choice of  $\tilde{v}$ . For a minimizer  $v$ , we have  $E_n(v) \leq E_n(\tilde{v}) \leq 1$ .

We now show that the candidate minimizer is an actual minimizer. By equation (7.4), the energy bound  $E_n(v) \leq 1$  implies that there must be some node  $j$  such

that  $|v_j^2 - 1| \leq 1/\sqrt{n}$ . Because  $|\tau| = |v_j^2 - 1||v_j| \leq (1/\sqrt{n})(1 + 1/\sqrt{n})^{1/2}$ , we have  $|\tau| = O(1/\sqrt{n})$ . For large  $n$ , the cubic polynomial  $g(x) = x^3 - x$  is a monotone function around its three simple roots and the derivatives at the roots are constants that are bounded away from 0 uniformly with respect to  $n$ . Therefore, whenever  $|g(v_i)| = |v_i^2 - 1||v_i| = |\tau| \leq (1/\sqrt{n})(1 + 1/\sqrt{n})^{1/2}$ , the Inverse Function Theorem guarantees that either  $|v_i^2 - 1| = O(1/\sqrt{n})$  or  $|v_i| = O(1/\sqrt{n})$ . The contribution to  $E_n$  of any  $v_i$  that satisfies  $|v_i^2 - 1| = O(1/\sqrt{n})$  is  $O(1/\sqrt{n})$ , and the contribution to  $E_n$  of any  $v_i$  that satisfies  $|v_i| = O(\sqrt{1/n})$  is at least  $(1 - O(1/n))^2$ . Therefore, with the bound  $E_n(v) \leq 1$ , we see that there is at most one node  $i$  with  $|v_i| = O(1/\sqrt{n})$  for sufficiently large  $n$ . Every other value of  $v$  must be a root of  $g(x) = \tau$  and satisfies  $|v_i^2 - 1| = O(1/\sqrt{n})$ . Therefore,  $v_i \rightarrow \pm 1$  as  $n \rightarrow \infty$ .  $\square$

The minimizer  $v$  of (7.4) that we constructed has a discontinuous profile, as it has jumps between the values  $+\sqrt{1 - \epsilon p}$  and  $-\sqrt{1 - \epsilon p}$ . See the related discussion in [30]. Furthermore, as  $n \rightarrow \infty$ , the limit  $v$  of the sequence of minimizers takes the values  $\pm 1$  almost everywhere as  $\epsilon \rightarrow 0$ , so (by the change of variables (7.5))  $u$  takes values  $\pm\sqrt{1 - \epsilon p}$  almost everywhere.

We now examine where the optimal  $u$  takes the values  $\pm\sqrt{1 - \epsilon p}$ . In doing so, we illustrate the properties of the GL minimizer in the graphon limit. Because  $\text{GL}_\epsilon^{W_n} \xrightarrow{\Gamma} \text{GL}_\epsilon^W$  as  $W_n \xrightarrow{\square} W$ , we know (see Section 6.1) that characterizing the GL minimizers for the constant graph gives insight into GL minimizers for the constant graphon. We show that the (arbitrarily many) oscillations between the two values  $+\sqrt{1 - \epsilon p}$  and  $-\sqrt{1 - \epsilon p}$  do not affect the optimality of  $u$ . The following discussion also helps explain why graphon functionals act on Young measures. We consider two different Young measures that minimize (7.4) (and thus minimize the constant-graphon GL functional) and that satisfy the mean-value constraint (7.1) but have different amounts of oscillation.

First, consider the Young measure  $\{\nu_x^{(1)}\}_{x \in (0,1)}$  that is defined by

$$(7.8) \quad \nu_x^{(1)} = \frac{1+c}{2} \delta_{\sqrt{1-\epsilon p}} + \frac{1-c}{2} \delta_{-\sqrt{1-\epsilon p}}.$$

The Young measure  $\nu^{(1)}$  is the limit, with respect to narrow convergence in the space of Young measures, of sequences of increasingly oscillatory functions  $u$  that take the value  $\sqrt{1 - \epsilon p}$  on a proportion  $\frac{1+c}{2}$  of the points in  $[0, 1]$  and the value  $-\sqrt{1 - \epsilon p}$  on a proportion  $\frac{1-c}{2}$  of the points in  $[0, 1]$ . Although  $\nu^{(1)}$  is not a function, one can think of it as the limit of increasingly oscillatory functions that satisfy the constraint (7.1). The graphon GL functional (4.14) evaluated at  $\nu^{(1)}$  is

$$(7.9) \quad \text{GL}_\epsilon^W(\nu^{(1)}) = 2(1 - \epsilon p)(1 - c)(1 + c) \int_0^1 \int_0^1 W(x, y) dx dy + \epsilon p^2.$$

The first term on the right-hand side of (7.9) is the Dirichlet-energy term, and the second term (i.e.,  $\epsilon p^2$ ) is the double-well term.

Second, consider the Young measure  $\{\nu_x^{(2)}\}_{x \in (0,1)}$  that is defined by

$$(7.10) \quad \nu_x^{(2)} = \delta_{u(x)}, \quad \text{where } u(x) = \begin{cases} -\sqrt{1 - \epsilon p}, & 0 < x < \frac{1-c}{2} \\ \sqrt{1 - \epsilon p}, & \frac{1-c}{2} < x < 1, \end{cases}$$

which represents a non-oscillatory limit because (1)  $u$  has a single transition and (2)  $u$  has a fixed value on each side of the transition. The location of this transition ensures that  $\nu^{(2)}$  satisfies the mean-value constraint (7.1). The graphon GL functional (4.14) evaluated at  $\nu^{(2)}$  is

$$(7.11) \quad \text{GL}_\epsilon^W(\nu^{(2)}) = 8(1 - \epsilon p) \int_0^{\frac{1-c}{2}} \int_{\frac{1-c}{2}}^1 W(x, y) dx dy + \epsilon p^2.$$

As in equation (7.9), the first term on the right-hand side of (7.11) is the Dirichlet-term and the second term is the double-well term.

To illustrate the effect of oscillations on the value of the graphon GL functional (4.14) for the constant graphon  $W \equiv p$ , we compare the values of  $\text{GL}_\epsilon^W(\nu^{(1)})$  and  $\text{GL}_\epsilon^W(\nu^{(2)})$ . Consider the case  $W \equiv 1$  and suppose that the mean-value constraint (7.1) holds. Equations (7.11) and (7.9) yield

$$\text{GL}_\epsilon^W(\nu^{(1)}) = \text{GL}_\epsilon^W(\nu^{(2)}) = 2(1 - \epsilon p)(1 - c)(1 + c) + \epsilon p^2,$$

which illustrates that oscillations do not increase the value of the graphon GL for the constant graphon. This contrasts with what occurs for classical GL functionals, which are concerned with interfaces and smoothness [28].

The contribution to the constant-graphon GL functional  $\text{GL}_\epsilon^W$  is the same whenever  $\nu_x$  and  $\nu_y$  differ from each other, regardless of how “far”  $x$  and  $y$  are from each other. This situation contrasts sharply with classical GL functionals, where the contribution from variations in the value of  $u$  is  $|\nabla u(x)|^2$  (which penalizes local changes in the value of  $u$  near  $x$ ), rather than  $W(x, y)|u(x) - u(y)|^2$  (which penalizes nonlocal variations in values).

In the  $\epsilon \rightarrow 0$  limit, the Young-measure minimizers (7.8) converge to the limiting Young measure  $\nu = \{\nu_x\}$ , which is defined by

$$(7.12) \quad \nu_x = \theta \delta_1 + (1 - \theta) \delta_{-1}$$

for all  $x \in [0, 1]$ . Because of the  $\Gamma$ -convergence  $\text{GL}_\epsilon^W \xrightarrow{\Gamma} \text{TV}^W$  (see Section 6.3), the Young-measure limit (7.12) of the minimizers (7.8) of  $\text{GL}_\epsilon^W$  is a minimizer of the limiting energy  $\text{TV}^W$  for  $W \equiv p$ .

To help with later discussions, we also study the (unrealistic) case of “oversaturation”, which occurs when the mean-value constraint (7.6) imposes that the mean value of  $v$  is larger than 1 or smaller than  $-1$ . Oversaturation occurs when  $|c| > \sqrt{1 - \epsilon p}$ . In the following lemma, we characterize the minimizers when  $|c| \geq \sqrt{1 - \epsilon p}$  and show that the optimal value of  $|c|$  is  $\sqrt{1 - \epsilon p}$  when there is oversaturation.

**Lemma 7.2** (Non-oversaturation of GL minimizers). *When  $|\gamma| := \left| \frac{c}{\sqrt{1 - \epsilon p}} \right| \geq 1$  in the mean-value constraint (7.6), the minimizer of (7.4) is the constant function  $v^*$ , which satisfies  $v_i^* = \gamma$  for all  $i \in [n]$ ; the minimum energy is  $E_n(v^*) = n(\gamma^2 - 1)^2$ . Equivalently, the minimizer of the constant-graph GL functional  $\text{GL}_\epsilon^{W_n}(u)$  with the mean-value constraint (7.2) is  $u^*$  (which satisfies  $u_i^* = c$  for all  $i \in [n]$ ), and the minimum energy is  $\text{GL}_\epsilon^{W_n}(u^*) = \frac{1}{\epsilon} \Phi(c) = \frac{1}{\epsilon} (c^2 - 1)^2$ . Additionally, the energies  $E_n(v^*)$  and  $\text{GL}_\epsilon^{W_n}(u^*)$  have smaller minima when  $|c| = \sqrt{1 - \epsilon p}$  than when  $|c| > \sqrt{1 - \epsilon p}$ .*

*Proof.* Let  $\gamma = \frac{c}{\sqrt{1-\epsilon p}}$ , and rewrite the mean-value constraint (7.6) as

$$(7.13) \quad \sum_{i=1}^n v_i = \gamma n.$$

We begin by showing that the minimizer of  $E_n$  is the constant function. Let  $h$  be the convex hull of  $f(s) = (s^2 - 1)^2$ . Therefore,  $h = 0$  on  $[-1, 1]$  and  $h = f$  everywhere else.

Define

$$H_n(v) := \sum_{i=1}^n h(v_i),$$

which satisfies  $H_n(v) \leq E_n(v)$  for any  $v$ . Because  $h$  is convex, it follows that  $(h(x) + h(y))/2 \leq h((x + y)/2)$ . Therefore, if  $v_i \neq v_j$  for any  $i \neq j$ , the function  $H_n$  decreases if  $v_i$  and  $v_j$  are both replaced by their mean  $(v_i + v_j)/2$ . Consequently, with the mean-value constraint (7.13), the minimizer of  $H_n$  is the constant function  $v' \equiv \gamma = \frac{c}{\sqrt{1-\epsilon p}}$ . Because  $\gamma \geq 1$ , we have  $H_n(v') = E_n(v')$ , so the entries of the minimizer of  $E_n$  are  $v_i = \gamma$  for all  $i$ , which yields

$$E_n(v') = H_n(v') = n(\gamma^2 - 1)^2.$$

The conclusion about  $\text{GL}_\epsilon^{W_n}(u)$  then follows immediately.  $\square$

**7.2.  $2 \times 2$  stochastic block models (SBMs).** Consider the  $2 \times 2$  piecewise-constant graphon

$$(7.14) \quad W(x, y) = \begin{cases} w_{11} & \text{if } x \in S, y \in S \\ w_{12} & \text{if } x \in S, y \in S^c \\ w_{21} & \text{if } x \in S^c, y \in S \\ w_{22} & \text{if } x \in S^c, y \in S^c, \end{cases}$$

where  $w_{k\ell} \in [0, 1]$  are edge weights and the intervals  $S = (0, a)$  and  $S^c = [a, 1)$ , for some  $a \in (0, 1)$ , are ‘‘communities’’ in  $(0, 1)$ . This graphon is a stochastic block model (SBM) [53]. Researchers use SBMs as generative models of graphs with various types of mesoscale network structures, such as assortative or disassortative block structures [56]. SBMs are a relatively general class of graphs and graphons. Indeed, the Szemerédi Lemma implies that one can approximate any  $L^\infty$  graphon arbitrarily closely in cut norm by an SBM [46, Lemma 3.1].

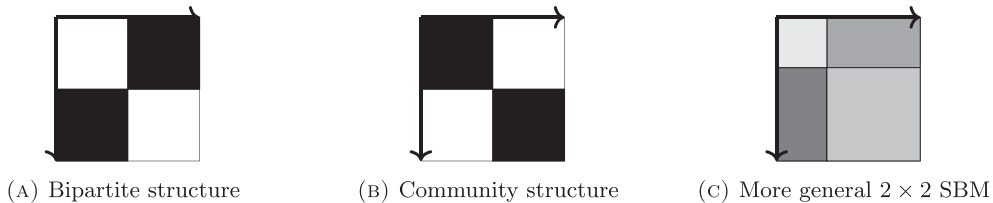


FIGURE 3. Three types of  $2 \times 2$  piecewise-constant graphons.

Consider the  $2 \times 2$  piecewise-constant graphons in Figure 3. We show examples with (A) bipartite structure, (B) community structure, and (C) a more general

block structure. If one views an SBM as a generative model, the entries  $w_{k\ell}$  encode the probability that there is an edge between a node in community  $k \in \{1, 2\}$  and a node in community  $\ell \in \{1, 2\}$ . However, we view the entries  $w_{k\ell}$  as fixed values in a block-diagonal adjacency matrix in which adjacency-matrix entries take the values of  $W(x, y)$  in (7.14).

In the example SBMs in Figure 3, black blocks have the value 1, white blocks have the value 0, and gray blocks have values in the interval  $(0, 1)$ . The origin is at the upper-left corner of each graphon. In an assortative SBM, the intra-community edge probabilities  $w_{kk}$  are larger than the inter-community edge probabilities  $w_{k\ell}$  (with  $k \neq \ell$ ). An SBM with planted community structure (see Figure 3(B)) is an extreme case of an assortative block structure [56]. In a disassortative SBM, the inter-community edge probabilities  $w_{k\ell}$  (with  $k \neq \ell$ ) are larger than the intra-community edge probabilities  $w_{kk}$ , so inter-community edges are more likely than intra-community edges. (In this case, the term ‘‘community’’ is technically a misnomer, but we use it for simplicity.) An extreme example is a bipartite SBM (see Figure 3(A)). Finally, Figure 3(C) shows an example of a general  $2 \times 2$  SBM with two distinct community sizes.

Seeking a GL minimizer for SBMs is different from seeking a constant-graphon GL minimizer because we do not need a mean-value constraint in the SBM case. For the constant graphon, we need a mean-value constraint to ensure that minimizers are not trivial (i.e., not all 1 or all  $-1$ ), but a bipartite graphon does not possess such a trivial minimizer. Trivial minimizers are prevented by the block structure that is encoded in the GL functional, rather than by an explicit constraint. Nontrivial block structures (e.g., ones that correspond to bipartite structure or community structure) incentivize a GL minimizer to distribute positive and negative values to different sets of nodes.

**7.2.1. Complete bipartite graphon.** In a bipartite graph (see Figure 3(A)), each node is in one of two sets,  $S$  and  $S^c$ , with edges only between nodes in different sets. Consider the complete bipartite graph

$$(7.15) \quad A_{ij}^{(n)} = \begin{cases} 0 & \text{if } i \in S, j \in S \\ 1 & \text{if } i \in S, j \in S^c \\ 1 & \text{if } i \in S^c, j \in S \\ 0 & \text{if } i \in S^c, j \in S^c. \end{cases}$$

Suppose that the graph has  $n$  nodes. Let  $a = |S|/n$  and  $1 - a = |S^c|/n$ , so  $a \in (0, 1)$  is the proportion of nodes in the set  $S$ . The corresponding complete bipartite graphon is

$$(7.16) \quad W(x, y) = \begin{cases} 0 & \text{if } (x, y) \in (0, a) \times (0, a) \\ 1 & \text{if } (x, y) \in (0, a) \times [a, 1) \\ 1 & \text{if } (x, y) \in [a, 1) \times (0, a) \\ 0 & \text{if } (x, y) \in [a, 1) \times [a, 1). \end{cases}$$

We rewrite the GL functional for the bipartite graph as

$$\begin{aligned}
(7.17) \quad \text{GL}_\epsilon^{W_n}(u) &= \frac{1}{n^2} \sum_{i \in S, j \in S^c} (u_i - u_j)^2 + \frac{1}{\epsilon n} \sum_{i=1}^n \Phi(u_i) \\
&= \frac{1}{n^2} \sum_{i \in S, j \in S^c} u_i^2 + \frac{1}{n^2} \sum_{i \in S, j \in S^c} u_j^2 - \frac{2}{n^2} \sum_{i \in S, j \in S^c} u_i u_j + \frac{1}{\epsilon n} \sum_{i=1}^n \Phi(u_i) \\
&= \frac{1}{n^2} \sum_{i \in S} |S^c| u_i^2 + \frac{1}{n^2} \sum_{j \in S^c} |S| u_j^2 - \frac{2}{n^2} \sum_{i \in S, j \in S^c} u_i u_j + \frac{1}{\epsilon n} \sum_{i=1}^n \Phi(u_i) \\
&= \frac{1}{\epsilon n} \sum_{i \in S} (u_i^4 - (2 - \epsilon(1-a))u_i^2 + 1) \\
&\quad + \frac{1}{\epsilon n} \sum_{i \in S^c} (u_i^4 - (2 - \epsilon a)u_i^2 + 1) - \frac{2}{n^2} \sum_{i \in S} u_i \sum_{j \in S^c} u_j.
\end{aligned}$$

By completing the squares in the first two sums, we see that the right-hand side of (7.17) is equal (up to a constant) to

$$\begin{aligned}
(7.18) \quad &\frac{1}{\epsilon n} \sum_{i \in S} \left( u_i^2 - \left( 1 - \frac{\epsilon(1-a)}{2} \right) \right)^2 \\
&+ \frac{1}{\epsilon n} \sum_{i \in S^c} \left( u_i^2 - \left( 1 - \frac{\epsilon a}{2} \right) \right)^2 - \frac{2}{n^2} \sum_{i \in S} u_i \sum_{j \in S^c} u_j.
\end{aligned}$$

Define  $\alpha = \sum_{i \in S} u_i$ , which we combine with equation (7.2) and insert into the last term of (7.18) to obtain

$$\sum_{i \in S} u_i \sum_{j \in S^c} u_j = \alpha(c - \alpha).$$

We are treating equation (7.2) as a shorthand notation for the mean value of  $u_i$ , rather than as a genuine constraint.

We introduce the rescaling

$$(7.19) \quad v_i = \begin{cases} \frac{u_i}{\sqrt{1 - \epsilon(1-a)/2}} & \text{for } i \in S \\ \frac{u_i}{\sqrt{1 - \epsilon a/2}} & \text{for } i \in S^c, \end{cases}$$

which is similar to the rescaling (7.5). To simplify notation, we define

$$(7.20) \quad c_S := \sqrt{1 - \epsilon(1-a)/2} \quad \text{and} \quad c_{S^c} := \sqrt{1 - \epsilon a/2}.$$

Minimizing (7.17) is then equivalent to minimizing the energy functional

$$(7.21) \quad E_n(v) := \frac{c_S^4}{\epsilon} \sum_{i \in S} (v_i^2 - 1)^2 + \frac{c_{S^c}^4}{\epsilon} \sum_{i \in S^c} (v_i^2 - 1)^2 - \frac{2}{n} c_S c_{S^c} \alpha(c - \alpha).$$

With the change of variables (7.19), we then obtain

$$(7.22) \quad \alpha = c_S \sum_{i \in S} v_i, \quad c - \alpha = c_{S^c} \sum_{i \in S^c} v_i.$$

Treating  $v$  and  $\alpha$  as unknowns, the resulting Euler–Lagrange equations for minimizing  $E_n$  are

$$(7.23) \quad \begin{aligned} \frac{4c_S^4}{\epsilon}(v_i^3 - v_i) - \tau_S c_S &= 0 \quad \text{on } S, \\ \frac{4c_{S^c}^4}{\epsilon}(v_i^3 - v_i) - \tau_{S^c} c_{S^c} &= 0 \quad \text{on } S^c, \\ -\frac{2}{n} c_S c_{S^c} \alpha + \tau_S - \tau_{S^c} &= 0. \end{aligned}$$

The first two equations in (7.23) are for minimizing  $v$  on  $S$  and on  $S^c$ , respectively, and the third equation is for minimizing  $\alpha$ . There is no constraint equation.

**Proposition 7.3** (Characterization of the GL minimizers for the complete bipartite graphon). *Let  $W_n$  be the complete bipartite graph with  $n$  nodes (see (7.15)). The minimizers of  $\text{GL}_\epsilon^{W_n}$  are functions  $u$  that take the values  $u_i = \pm\sqrt{1 - \epsilon a/2} + O(1/\sqrt{n})$  for  $i \in S^c$  and  $u_i = \pm\sqrt{1 - \epsilon(1-a)/2} + O(1/\sqrt{n})$  for  $i \in S$ , except for at most a finite number  $m$  of nodes, where  $m$  is independent of  $n$ .*

*Proof.* We use a similar argument as we did for the constant graphon to show that the values  $v_i$  of the global minimizer are close to  $\pm 1$  except for at most  $m$  nodes, where  $m$  is independent of  $n$ .

We show that the minimum value of  $E_n$  is bounded above by  $3/\epsilon$  for all  $n$ . To do this, we define a function  $\tilde{v}$  such that

$$(7.24) \quad E_n(\tilde{v}) \leq \frac{3}{\epsilon}.$$

We need to consider different parities of  $n$ ,  $|S|$ , and  $|S^c|$ . We first suppose that  $n$ ,  $|S|$ , and  $|S^c|$  are all even. In this case,  $E_n(\tilde{v}) = 0$  if  $\tilde{v}$  has the same number of 1 values on  $S$  as the number of  $-1$  values on  $S^c$ . We thus obtain  $\alpha = 0$ . Now suppose that  $n$  is even but both  $|S|$  and  $|S^c|$  are odd. We set  $\tilde{v}_i = 0$  for one node in  $S$  and one node in  $S^c$ , and we let  $\tilde{v}$  have the values 1 and  $-1$  on an equal number of entries in  $S$  and  $S^c$ . We then again obtain  $\alpha = 0$ . For both of these choices of  $\tilde{v}$ , we have

$$(7.25) \quad E_n(\tilde{v}) \leq \frac{c_S^4 + c_{S^c}^4}{\epsilon} = \frac{2}{\epsilon} - 1 + \frac{\epsilon}{4}(1 - 2a + 2a^2) \leq \frac{3}{\epsilon}$$

for  $a \in (0, 1)$  and  $\epsilon < 1$ . In the final case,  $n$  is odd, so one of  $|S|$  and  $|S^c|$  is even and the other is odd. In this case, one node has  $\tilde{v}_i = 0$  and the remaining  $\tilde{v}_i$  have an equal number of 1 and  $-1$  entries in  $S$  and an equal number of 1 and  $-1$  entries in  $S^c$ . This again yields  $\alpha = 0$ , and  $E_n(\tilde{v})$  is either  $c_S^4/\epsilon$  (if  $|S|$  is odd) or  $c_{S^c}^4/\epsilon$  (if  $|S^c|$  is odd), which again satisfies the uniform energy bound (7.24).

From the energy bound (7.24), a minimizer  $v$  satisfies

$$(7.26) \quad \frac{c_S^4}{\epsilon} \sum_{i \in S} (v_i^2 - 1)^2 + \frac{c_{S^c}^4}{\epsilon} \sum_{i \in S^c} (v_i^2 - 1)^2 \leq E_n(v) \leq E_n(\tilde{v}) \leq \frac{3}{\epsilon}.$$

Therefore, there exist nodes  $j \in S$  and  $j' \in S^c$  such that

$$(7.27) \quad (v_j^2 - 1)^2 \leq \frac{3}{|S|c_S^4} \quad \text{and} \quad (v_{j'}^2 - 1)^2 \leq \frac{3}{|S^c|c_{S^c}^4}.$$

Because  $|S| = an$  and  $|S^c| = (1-a)n$ , we have

$$|v_j^2 - 1| \leq \frac{m_1}{\sqrt{n}} \quad \text{and} \quad |v_{j'}^2 - 1| \leq \frac{m_2}{\sqrt{n}},$$

where the constants  $m_1$  and  $m_2$  do not depend on  $n$ . The Euler–Lagrange equations (7.23) hold for all nodes, and they thus hold for nodes  $j$  and  $j'$ . Therefore,

$$|\tau_S| \leq \frac{C}{\sqrt{n}} \quad \text{and} \quad |\tau_{S^c}| \leq \frac{C}{\sqrt{n}}$$

for some constant  $C$  that does not depend on  $n$ . With these bounds on  $\tau_S$  and  $\tau_{S^c}$ , the Euler–Lagrange equations (7.23) yield

$$v_i^3 - v_i = O(1/\sqrt{n})$$

for all nodes in both  $S$  and  $S^c$ . Similarly to the proof of Proposition 7.1, for sufficiently large  $n$ , the zeros  $v_i$  lie within  $O(k/\sqrt{n})$  of  $\pm 1$  and 0. Furthermore, due to the energy bound (7.26), we see that  $v$  can be within  $O(k/\sqrt{n})$  of 0 for at most a finite number of nodes. This number of nodes does not exceed a constant that is independent of  $n$ .

From the change of variables (7.19), the GL minimizer  $u$  takes the values  $u_i = \pm\sqrt{1 - \epsilon a/2} + O(1/\sqrt{n})$  for  $i \in S^c$  and  $u_i = \pm\sqrt{1 - \epsilon(1-a)/2} + O(1/\sqrt{n})$  for  $i \in S$ , except for at most a finite number of nodes  $i$ .  $\square$

It follows from Proposition 7.3 that the limiting Young measures for the complete bipartite graphon GL minimizer are

$$\begin{aligned} \nu_x = & \mathbf{1}_{U_1}(x) \delta_{\sqrt{1-\epsilon a/2}} + \mathbf{1}_{S \setminus U_1}(x) \delta_{-\sqrt{1-\epsilon a/2}} \\ & + \mathbf{1}_{U_2}(x) \delta_{\sqrt{1-\epsilon(1-a)/2}} + \mathbf{1}_{S^c \setminus U_2}(x) \delta_{-\sqrt{1-\epsilon(1-a)/2}}, \end{aligned}$$

where  $\mathbf{1}_Z$  denotes the indicator function on the set  $Z$  and  $U_1 \subset S$  and  $U_2 \subset S^c$  are any sets of sizes  $|U_1| = |S|/2$  and  $|U_2| = |S^c|/2$ .

As  $n \rightarrow \infty$ , the energy bound (7.24) also yields a bound on the limiting graphon energy. This bound on the limiting graphon energy shrinks as  $n$  increases. Because one loses the discrete character of graphs in the continuum limit, it no longer makes sense to discuss “even” or “odd”  $|S|$  and  $|S^c|$ .

**7.2.2. Community-structure graphon.** Graphs with community structure have densely-connected subgraphs (i.e., communities) that are sparsely connected to each other. As in a  $2 \times 2$  bipartite SBM, a  $2 \times 2$  community-structure graph involves a partition  $\{S, S^c\}$  of the set of nodes of a graph into two communities. Edges occur frequently between nodes in the same community (either  $S$  or  $S^c$ ), and they occur sparsely between nodes in different communities. We suppose that the communities are complete, and we refer to this example as the “complete community-structure graph”. The complete community-structure graph with communities  $S$  and  $S^c$  has adjacency-matrix elements

$$(7.28) \quad A_{ij}^{(n)} = \begin{cases} 1 & \text{if } i \in S, j \in S \\ 0 & \text{if } i \in S, j \in S^c \\ 0 & \text{if } i \in S^c, j \in S \\ 1 & \text{if } i \in S^c, j \in S^c. \end{cases}$$

The corresponding complete community-structure graphon is

$$(7.29) \quad W(x, y) = \begin{cases} 1 & \text{if } (x, y) \in (0, a) \times (0, a) \\ 0 & \text{if } (x, y) \in (0, a) \times [a, 1) \\ 0 & \text{if } (x, y) \in [a, 1) \times (0, a) \\ 1 & \text{if } (x, y) \in [a, 1) \times [a, 1). \end{cases}$$

The GL functional for a complete community-structure graph is

$$\text{GL}_\epsilon^{W_n}(u) = \frac{1}{n^2} \sum_{i \in S, j \in S} (u_i - u_j)^2 + \frac{1}{n^2} \sum_{i \in S^c, j \in S^c} (u_i - u_j)^2 + \frac{1}{\epsilon n} \sum_{i=1}^n \Phi(u_i),$$

which, with a similar computation to that for a complete bipartite graph, is equivalent up to a constant to the functional

$$(7.30) \quad E_n(v) = \frac{c_S^4}{\epsilon} \sum_{i \in S} (v_i^2 - 1)^2 + \frac{c_{S^c}^4}{\epsilon} \sum_{i \in S^c} (v_i^2 - 1)^2 - \frac{2}{n} c_S c_{S^c} \alpha^2 - \frac{2}{n} c_S c_{S^c} (c - \alpha)^2,$$

where  $\alpha$  and  $c - \alpha$  are the same (see (7.22)) as for a complete bipartite graphon,  $c_S v_i = u_i$  for  $i \in S$  and  $c_{S^c} v_i = u_i$  for  $i \in S^c$ , and

$$(7.31) \quad c_S = \sqrt{1 - \epsilon a}, \quad c_{S^c} = \sqrt{1 - \epsilon(1 - a)}.$$

As for a complete bipartite graphon,  $\alpha$  does not act as a constraint. Instead, it provides a concise notation for the quantity  $\alpha = \sum_{i \in S} u_i$ . We let  $c = 0$  and have the mean-value expressions

$$(7.32) \quad c_S \sum_{i \in S} v_i = \alpha, \quad c_{S^c} \sum_{i \in S^c} v_i = -\alpha.$$

When  $|S| = |S^c|$ , the energy functional (7.30) reduces to

$$(7.33) \quad E_n(v) = \frac{2c_S^4}{\epsilon} \sum_{i=1}^n (v_i^2 - 1)^2 - \frac{4}{n} c_S^2 \alpha^2.$$

The first term of the energy (7.33) encourages the values of  $v$  to be near  $\pm 1$ . The second term of (7.33) encourages  $|\alpha|$  to be as large as possible. In other words, the second term encourages the values of  $v_i$  within a community to either all be very large or all be very small. Therefore, there is a trade-off between the first and the second terms. We show in Proposition 7.4 that the optimal balance in this trade-off has  $v$  near  $\pm 1$  and  $\alpha = |S|$ . This implies that  $v_i = +1$  for all  $i \in S$  and  $v_i = -1$  for all  $i \in S^c$ .

In the following proposition, we characterize the GL minimizers for a complete community-structure graphon with  $|S| = |S^c|$ . Let  $|S| = |S^c|$  (which implies that  $c_S = c_{S^c}$ ) and  $\alpha = \gamma|S|$ , where we recall that  $\gamma = \frac{c}{\sqrt{1 - \epsilon c}}$  (see Lemma 7.2). The expressions in (7.32) then entail that  $\gamma/c_S$  is the mean value of  $v$  on  $S$  and that  $-\gamma/c_S$  is the mean value of  $v$  on  $S^c$ . Therefore, we can rewrite (7.32) as

$$(7.34) \quad \frac{1}{|S|} \sum_{i \in S} v_i = \frac{\gamma}{c_S}, \quad \frac{1}{|S|} \sum_{i \in S^c} v_i = -\frac{\gamma}{c_S}.$$

**Proposition 7.4** (Characterization of the GL minimizers for the complete community-structure graphon). *Let  $W_n$  be the complete community-structure graph, which has the form (7.29), and also suppose that  $|S| = |S^c|$ . The minimizers of  $\text{GL}_\epsilon^{W_n}$  are functions  $u$  that, on  $S$ , take a constant value that approaches  $+1$  as  $\epsilon \downarrow 0$  and, on  $S^c$ , take a constant value that approaches  $-1$  as  $\epsilon \downarrow 0$ . Furthermore, the values of  $\gamma$  that minimize  $E_n$  are equal to  $\pm c_S$ , which approach  $\pm 1$  as  $\epsilon \rightarrow 0$ .*

*Proof.* We separately consider the two cases  $|\gamma/c_S| > 1$  and  $|\gamma/c_S| \leq 1$ .

First suppose that  $\left|\frac{\gamma}{c_S}\right| > 1$ . Let  $h$  be the convex hull of the double-well potential  $f(s) = (s^2 - 1)^2$ . Therefore,  $h = 0$  on  $[-1, 1]$  and  $h = f$  everywhere else. Define

$$H_n(v) := \frac{2c_S^4}{\epsilon} \sum_{i=1}^n h(v_i) - \frac{4}{n} c_S^2 \alpha^2,$$

which is a lower bound of the energy  $E_n$  for all  $v$ . Because  $h$  is convex,  $(h(x) + h(y))/2 \leq h((x+y)/2)$ . Therefore, if  $v_i \neq v_j$  for any  $i \neq j$ , the function  $H_n$  decreases if one replaces both  $v_i$  and  $v_j$  by their mean  $(v_i + v_j)/2$ . Consequently, the minimizer  $v'$  of  $H_n$  is constant on  $S$  and constant on  $S^c$ . However, because of the mean values in (7.34),  $v'$  has different values on  $S$  and  $S^c$ . The minimizer that satisfies (7.34) is the function  $v'$  with

$$(7.35) \quad v'_i = \begin{cases} +\frac{\gamma}{c_S} & \text{if } i \in S \\ -\frac{\gamma}{c_S} & \text{if } i \in S^c. \end{cases}$$

Because  $|\gamma/c_S| > 1$ , we have  $H_n(v') = E_n(v')$ . Let  $\gamma' = \gamma^2$  and consider the energy

$$(7.36) \quad E_n(v') = \frac{nc_S^4}{\epsilon} \left( \frac{\gamma'}{c_S^2} - 1 \right)^2 - \frac{4}{n} c_S^2 \gamma' |S|^2$$

as a function of  $\gamma'$ . We denote this function by  $g_1(\gamma')$ . Its derivative is

$$(7.37) \quad \frac{d}{d\gamma'} g_1(\gamma') = \frac{2c_S^2 n}{\epsilon} \left( \frac{\gamma'}{c_S^2} - 1 \right) - \frac{4}{n} c_S^2 |S|^2.$$

Setting the right-hand side of (7.37) to 0 yields the critical point

$$\gamma' = c_S^2 (1 + \epsilon a).$$

Therefore, when  $|S| = |S^c|$ , the critical values of  $\gamma$  are

$$\gamma \in \left\{ \pm c_S \sqrt{1 + \epsilon/2} \right\} = \left\{ \pm \sqrt{1 - (\epsilon/2)^2} \right\}.$$

The second derivative  $\frac{d^2}{d\gamma'^2} g_1(\gamma') = \frac{4|S|}{\epsilon} > 0$  everywhere, so the critical points are minima of  $g_1$ . The minimizer (7.35) of  $H_n$  is thus

$$(7.38) \quad v'_i = \begin{cases} +\sqrt{1 + \epsilon/2} & \text{if } i \in S \\ -\sqrt{1 + \epsilon/2} & \text{if } i \in S^c. \end{cases}$$

Using the change of variables (7.31), we obtain the GL minimizer  $u$ , which is defined by

$$(7.39) \quad u_i = \begin{cases} +\sqrt{1 - (\epsilon/2)^2} & \text{if } i \in S \\ -\sqrt{1 - (\epsilon/2)^2} & \text{if } i \in S^c. \end{cases}$$

As  $\epsilon \downarrow 0$ , we have  $+\sqrt{1 - (\epsilon/2)^2} \downarrow 1$  and  $-\sqrt{1 - (\epsilon/2)^2} \uparrow -1$ . That is, the positive values of  $u$  converge downward to 1 and the negative values of  $u$  converge upward to  $-1$ . Therefore, even when the mean of  $v$  has an absolute value larger than 1, the optimal  $u$  approaches  $\pm 1$ .

Now suppose that  $|\gamma/c_S| \leq 1$ . We construct a candidate minimizer  $\tilde{v}$  that satisfies  $\gamma = \pm c_S$ , and we show that the energy is larger for  $|\gamma/c_S| < 1$  than for  $|\gamma/c_S| = 1$ . Let  $\tilde{v}_i = +1$  for all nodes  $i \in S$ , and let  $\tilde{v}_i = -1$  for all nodes  $i \in S^c$ . This yields  $\gamma = c_S$  and an energy of

$$(7.40) \quad E_n(\tilde{v}) = -\frac{4}{n} c_S^2 |S|^2.$$

Swapping the signs of  $\tilde{v}$  for  $S$  and  $S^c$  yields  $\gamma = -c_S$  and the same energy (7.40).

Now consider  $v$  with  $\gamma \in (0, c_S)$ , which implies that  $|\gamma/c_S| < 1$ . In this case, the energy (7.30) is

$$(7.41) \quad E_n(v) = \frac{2c_S^4}{\epsilon} \sum_{i=1}^n (v_i^2 - 1)^2 - \frac{4}{n} c_S^2 \gamma^2 |S|^2.$$

The second term of (7.41) is larger than  $E_n(\tilde{v})$  and the first term of (7.41) is nonnegative, so  $E_n(v) \geq E_n(\tilde{v})$ . We can use the same argument for the energy  $E_n(v)$  for  $\gamma \in (-c_S, 0)$ .

We conclude that  $\gamma/c_S = \pm 1$  is the optimal mean value of  $v$  on  $S$  and  $S^c$  when  $|\gamma/c_S| \leq 1$ . Furthermore, the candidate minimizer  $\tilde{v}$  is a minimizer for this value of  $\gamma$ . The GL minimizer  $u$  that corresponds to this candidate minimizer  $\tilde{v}$  has components

$$(7.42) \quad u_i = \begin{cases} +\frac{1}{\sqrt{1-\epsilon/2}} & \text{if } i \in S \\ -\frac{1}{\sqrt{1-\epsilon/2}} & \text{if } i \in S^c. \end{cases}$$

As  $\epsilon \downarrow 0$ , the values of the GL minimizer  $u$  approach  $\pm 1$ . □

It follows from Proposition 7.4 that the limiting Young measure for the complete community-structure graphon GL minimizer with  $|S| = |S^c|$  is

$$\nu_x = \mathbf{1}_S(x) \delta_{\sqrt{1-\epsilon/2}} + \mathbf{1}_{S^c}(x) \delta_{-\sqrt{1-\epsilon/2}}.$$

The assumption that  $|S| = |S^c|$  is a significant simplification and is the reason that the minimizer in this case is simpler than the minimizer for complete bipartite graphons.

## 8. CONCLUSIONS AND DISCUSSION

We studied large-graph limits of Ginzburg–Landau (GL) functionals by taking a functional-analytic view of graphs as nonlocal kernels. We defined graphon GL and total-variation (TV) functionals, and we showed that their minimizers are consistent with minimizers of associated graph GL and TV functionals. Given a sequence  $\{W_n\}$  of graphs that converges in cut norm to a limiting graphon  $W$  (i.e.,  $W_n \xrightarrow{\square} W$ ), the sequence of graph GL functionals (4.6)  $\Gamma$ -converges to the graphon GL functional (4.14), implying that the sequence of minimizers of the GL functionals that correspond to the graphs  $W_n$  converges to a minimizer of the GL functional that corresponds to the graphon  $W$  as  $W_n \xrightarrow{\square} W$ . We also proved that the sequence of graph TV functionals (4.8)  $\Gamma$ -converges to the graphon TV functional (4.16) as  $W_n \xrightarrow{\square} W$ . Additionally, we showed that the graphon GL and TV functionals satisfy the classical  $\Gamma$ -limits  $\text{GL}_\epsilon \xrightarrow{\Gamma} \text{TV}$  and  $\text{GL}_\epsilon^{W_n} \xrightarrow{\Gamma} \text{TV}^{W_n}$  as  $\epsilon \rightarrow 0$ . These  $\Gamma$ -convergence results (see our summary in Figure 1), in concert with compactness properties, imply that the minimizers of the  $\Gamma$ -converging functionals also converge.

The limiting functionals highlight several fundamental differences between the graphon GL functional and both the graph GL functional and the classical GL functional. One difference is that graphon functionals are formulated using Young measures, rather than using functions. This difference highlights the fact that the limiting minimizers are Young measures, which can have arbitrary amounts of oscillation, that take the same values in the same proportions as the minimizing functions. Another difference, which is highlighted by our limit (4) (see Section 6.3 for the proof of this limit), is that the  $\epsilon$ -scaling of the graphon GL functional is somewhere between the  $\epsilon$ -scalings of the classical GL functional and the graph GL functional. The classical GL functional has scalings  $\epsilon$  and  $1/\epsilon$  for the Dirichlet energy and double-well potential, respectively. By contrast, the graph GL functional has the scalings 1 and  $1/\epsilon$ , respectively. The graphon GL functional has the same  $\epsilon$ -scalings as the graph GL functional when the graphon  $W$  is bounded. We did not determine a scaling for more general graphons  $W \in L^p((0,1)^2)$ , but this is worth examining in future efforts.

Our limit (3), which we proved in Section 6.1, extends results of Braides et al. [16, Lemma 11 and Theorem 12] in two key ways. First, Braides et al. proved  $\Gamma$ -convergence of the graph-cut functional (which acts on finite-range functions) as  $n \rightarrow \infty$ , whereas we proved  $\Gamma$ -convergence of the graph Dirichlet energy, which extends the graph-cut functional to act on  $L^\infty$  functions. Second, Braides et al. worked with dense sequences of graphs that converge to  $L^\infty$  graphons, whereas we proved  $\Gamma$ -convergence for more general sequences of graphs that converge to  $L^p$  graphons.

In the present paper, we have been concerned with theoretical ideas, but one can also examine applications of graphon GL functionals that extend existing applications of graph GL functionals. In particular, graph GL and TV functionals have been employed in image processing [3, 4, 20, 31, 32, 40, 41, 43, 44], and it is worthwhile to pursue analogous applications of graphon GL and TV functionals. One benefit of

graphon theory is that graphons provide a way to compare networks that have different sizes and densities [70]. Furthermore, because graphons are a familiar object in nonlocal analysis, one can use computational techniques from nonlocal analysis (e.g., see [27]) for applications that involve large graphs.

Another question for future research is how to apply our results in practical computations. For example, it is important to investigate convergence rates. For approximately what graph sizes  $n$  are the minimizers of a graphon GL functional close enough to the minimizers of an associated graph GL functional? Furthermore, minimizing a graph GL functional requires approximation algorithms [7, 19, 51], and similar algorithms may be useful for minimizing a graphon GL functional. Seeking graphon GL minimizers will also involve seeking Young-measure minimizers, and this in turn will require numerical approximations [22, 54].

#### ACKNOWLEDGEMENTS

Our research was supported in part by the National Science Foundation through grants DMS-2309245 (QD), DMS-1937254 (QD, JS, and EJZ), DMS-1916439 (which is a grant to the American Mathematical Society), and DGE-2036197 (EJZ). Some of the research was conducted as part of EJZ’s visits to UCLA. We thank Andrew Blumberg for useful discussions and two anonymous referees for helpful comments. MAP thanks Leonid Bunimovich for his mentorship when he was a postdoc at Georgia Tech and is honored by the invitation to contribute to this Festschrift.

#### REFERENCES

- [1] A. K. Akash, S. Li and N. García Trillos, *Wasserstein barycenter-based model fusion and linear mode connectivity of neural networks*, arXiv preprint arXiv:2210.06671, 2022.
- [2] S. M. Allen and J. W. Cahn, *A microscopic theory for antiphase boundary motion and its application to antiphase domain coarsening*, Acta Metallurgica **27** (1979), 1085–1095.
- [3] P. Arias, V. Caselles, G. Facciolo, V. Lazzcano and R. Sadek, *Nonlocal variational models for inpainting and interpolation*, Math. Models Methods Appl. Sci. **22** (2012): 1230003.
- [4] J.-F. Aujol, G. Gilboa and N. Papadakis, *Fundamentals of non-local total variation spectral theory*, in: International Conference on Scale Space and Variational Methods in Computer Vision, J.-F. Aujol, M. Nikolova, N. Papadakis (eds), Springer-Verlag, Heidelberg, Germany, 2015, pp. 66–77.
- [5] E. Bayraktar, S. Chakraborty and R. Wu, *Graphon mean field systems*, Ann. Appl. Probab. **33** (2023), 3587–3619.
- [6] A. L. Bertozzi and A. Flenner, *Diffuse interface models on graphs for classification of high dimensional data*, Multiscale Model. Simul. **10** (2012), 1090–1118.
- [7] A. L. Bertozzi and A. Flenner, *Diffuse interface models on graphs for classification of high dimensional data*, SIAM Rev. **58** (2016), 293–328.
- [8] B. Bollobás and O. Riordan, *Metrics for sparse graphs*, in: Surveys in Combinatorics 2009, London Mathematical Society Lecture Note Series, S. Huczynska, J. D. Mitchell, C. M. Roney-Dougal (eds), Cambridge University Press, Cambridge, UK, 2009, pp. 211–288.
- [9] B. Bollobás and O. Riordan, *Sparse graphs: Metrics and random models*, Random Structures Algorithms **39** (2011), 1–38.
- [10] C. Borgs, J. T. Chayes, H. Cohn and N. Holden, *Sparse exchangeable graphs and their limits via graphon processes*, J. Mach. Learn. Res. **18**(210) (2018), 1–71.
- [11] C. Borgs, J. T. Chayes, H. Cohn and Y. Zhao, *An  $L^p$  theory of sparse graph convergence II: LD convergence, quotients and right convergence*, Ann. Probab. **46** (2018), 337–396.

- [12] C. Borgs, J. T. Chayes, H. Cohn and Y. Zhao, *An  $L^p$  theory of sparse graph convergence I: Limits, sparse random graph models, and power law distributions*, Trans. Amer. Math. Soc **372** (2019), 3019–3062.
- [13] C. Borgs, J. T. Chayes, L. Lovász, V. T. Sós and K. Vesztegombi, *Convergent sequences of dense graphs I: Subgraph frequencies, metric properties and testing*, Adv. Math. **219** (2008), 1801–1851.
- [14] C. Borgs, J. T. Chayes, L. Lovász, V. T. Sós and K. Vesztegombi, *Convergent sequences of dense graphs II. Multiway cuts and statistical physics*, Ann. Math. **176** (2012), 151–219.
- [15] A. Braides, *A handbook of  $\Gamma$ -convergence*, in: Handbook of Differential Equations: Stationary Partial Differential Equations, M. Chipot, P. Quittner (eds), vol. 3, Elsevier, Amsterdam, The Netherlands, 2006, pp. 101–213.
- [16] A. Braides, P. Cermelli and Si. Dovetta,  *$\Gamma$ -limit of the cut functional on dense graph sequences*, ESAIM Control Optim. Calc. Var. **26** (2020): 26.
- [17] J. J. Bramburger, *The emerging utility of graphons in applied math*, SIAM News **56**(8) (2023), available at <https://www.siam.org/publications/siam-news/articles/the-emerging-utility-of-graphons-in-applied-math>.
- [18] L. Bronsard and R. V. Kohn, *Motion by mean curvature as the singular limit of Ginzburg–Landau dynamics*, J. Differential Equations **90** (1991), 211–237.
- [19] J. Budd and Y. van Gennip, *Graph Merriman–Bence–Osher as a semidiscrete implicit Euler scheme for graph Allen–Cahn flow*, SIAM J. Math. Anal. **52** (2020), 4101–4139.
- [20] L. A. Caffarelli, J.-M. Roquejoffre and O. Savin, *Nonlocal minimal surfaces*, Comm. Pure Appl. Math. **63** (2010), 1111–1144.
- [21] P. E. Caines and M. Huang, *Graphon mean field games and their equations*, SIAM. J. Control Optim. **59** (2021), 4373–4399.
- [22] C. Carstensen and T. Roubíček, *Numerical approximation of Young measures in non-convex variational problems*, Numer. Math. **84** (2000), 395–415.
- [23] M. Cyrot, *Ginzburg–Landau theory for superconductors*, Rep. Prog. Phys. **36** (1973), 103–158.
- [24] B. Dacorogna, *Weak Continuity and Weak Lower Semicontinuity of Non-Linear Functionals*. Springer-Verlag, Heidelberg, Germany, 1982.
- [25] G. B. Dantzig and D. R. Fulkerson, *12. On the max-flow min-cut theorem of networks*, in: Linear Inequalities and Related Systems, H. W. Kuhn, A. W. Tucker (eds), Annals of Mathematical Studies, vol. 38, Princeton University Press, Princeton, NJ, USA, 1956, pp. 225–231.
- [26] P. Diaconis and S. Janson, *Graph limits and exchangeable random graphs*, arXiv preprint arXiv:0712.2749, 2007.
- [27] Q. Du, *Nonlocal Modeling, Analysis, and Computation*, SIAM, Philadelphia, PA, USA, 2019.
- [28] Q. Du and X. Feng, *The phase field method for geometric moving interfaces and their numerical approximations*, in: Handbook of Numerical Analysis, A. Bonito, R. H. Nochetto (eds), vol. 21. Geometric Partial Differential Equations — Part I, Elsevier, Amsterdam, The Netherlands, 2020, pp. 425–508.
- [29] Q. Du, M. D. Gunzburger and J. S. Peterson, *Analysis and approximation of the Ginzburg–Landau model of superconductivity*, SIAM Rev. **34** (1992), 54–81.
- [30] Q. Du and J. Yang, *Asymptotically compatible Fourier spectral approximations of nonlocal Allen–Cahn equations*, SIAM J. Numer. Anal. **54** (2016), 1899–1919.
- [31] I. E. Bouchairi, A. Elmoataz and J. Fadili, *Nonlocal perimeters and curvature flows on graphs with applications in image processing and high-dimensional data classification*, SIAM J. Imaging Sci. **16** (2023), 368–392.
- [32] A. E. Chakik, A. Elmoataz and X. Desquesnes, *Mean curvature flow on graphs for image and manifold restoration and enhancement*, Signal Processing **105** (2014), 449–463.
- [33] L. C. Evans, *Partial Differential Equations*, American Mathematical Society, Providence, RI, USA, second edition, 2022.
- [34] L. R. Ford and D. R. Fulkerson, *Maximal flow through a network*, Canad. J. Math. **8** (1956), 399–404.
- [35] S. Gao and P. E. Caines, *Graphon control of large-scale networks of linear systems*, IEEE Trans. Automat. Control **65** (2019), 4090–4105.

- [36] N. García Trillos, *Variational Limits of Graph Cuts on Point Clouds*. PhD thesis, Carnegie Mellon University, 2015.
- [37] N. García Trillos and R. W. Murray, *A maximum principle argument for the uniform convergence of graph Laplacian regressors*, SIAM J. Math. Data Sci. **2** (2020), 705–739.
- [38] N. García Trillos and D. Slepčev,  *$\Gamma$ -convergence of perimeter on random geometric graphs*, Technical report, Carnegie Mellon University, 2013, available at <https://www.math.cmu.edu/PIRE/pub/pire66/13-CNA-024.pdf>.
- [39] N. García Trillos and D. Slepčev, *Continuum limit of total variation on point clouds*, Arch. Ration. Mech. Anal. **220** (2016), 193–241.
- [40] G. Gilboa and S. J. Osher, *Nonlocal operators with applications to image processing*, Multiscale Model. Simul. **7** (2009), 1005–1028.
- [41] Y. Hafiene, J. Fadili and A. Elmoataz, *Nonlocal  $p$ -Laplacian evolution problems on graphs*, SIAM J. Numer. Anal. **56** (2018), 1064–1090.
- [42] S. Janson, *Graphons, Cut Norm and Distance, Couplings and Rearrangements*, vol. 4, NYJM Monographs, SUNY at Albany, 2013, available at arXiv:1009.2376.
- [43] X. Liu and L. Huang, *A new nonlocal total variation regularization algorithm for image denoising*, Math. Comput. Simul. **97** (2014), 224–233.
- [44] Y. Lou, X. Zhang, S. J. Osher and A. L. Bertozzi, *Image recovery via nonlocal operators*, J. Sci. Comput. **42** (2010), 185–197.
- [45] L. Lovász, *Large Networks and Graph Limits*, American Mathematical Society, Providence, RI, USA, 2012.
- [46] L. Lovász and B. Szegedy, *Szemerédi’s Lemma for the analyst*, Geom. Funct. Anal. **17** (2007), 252–270.
- [47] J. M. Mazón, J. D. Rossi and J. Toledo, *Nonlocal perimeter, curvature and minimal surfaces for measurable sets*, J. Anal. Math. **138** (2019), 235–279.
- [48] G. S. Medvedev, *The nonlinear heat equation on dense graphs and graph limits*, SIAM J. Math. Anal. **46** (2014), 2743–2766.
- [49] G. S. Medvedev, *Small-world networks of Kuramoto oscillators*, Phys. D **266** (2014), 13–22.
- [50] E. Merkurjev, E. Bae, A. L. Bertozzi and X.-C. Tai, *Global binary optimization on graphs for classification of high-dimensional data*, J. Math. Imaging Vision **52** (2015), 414–435.
- [51] E. Merkurjev, T. Kostić and A. L. Bertozzi, *An MBO scheme on graphs for classification and image processing*, SIAM J. Imaging Sci. **6** (2013), 1903–1930.
- [52] L. Modica and S. Mortola, *Un esempio di  $\Gamma^-$ -convergenza*, Boll. Unione Mat. Ital. **14** (1977), 285–299.
- [53] M. Newman, *Networks*, Oxford University Press, Oxford, UK, second edition, 2018.
- [54] R. A. Nicolaides and N. J. Walkington, *Computation of microstructure utilizing Young measure representations*, J. Intell. Mater. Syst. Struct. **4** (1993), 457–462.
- [55] F. Parise and A. Ozdaglar, *Graphon games*, in: Proceedings of the 2019 ACM Conference on Economics and Computation, New York City, NY, USA, 2019, pp. 457–458.
- [56] T. P. Peixoto, *Bayesian stochastic blockmodeling*, in: Advances in Network Clustering and Blockmodeling, P. Doreian, V. Batagelj, A. Ferligoj (eds), John Wiley & Sons, Inc., Hoboken, NJ, USA, 2019, pp. 289–332.
- [57] X. Pu and B. Guo, *Well-posedness and dynamics for the fractional Ginzburg–Landau equation*, Appl. Anal. **92** (2013), 318–334.
- [58] L. I. Rudin, S. J. Osher and E. Fatemi, *Nonlinear total variation based noise removal algorithms*, Phys. D **60** (1992), 259–268.
- [59] L. Ruiz, L. F. O. Chamon and A. Ribeiro, *Graphon signal processing*, IEEE Trans. Signal Process., **69** (2021), 4961–4976.
- [60] O. Savin and E. Valdinoci,  *$\Gamma$ -convergence for nonlocal phase transitions*, Ann. Inst. H. Poincaré C Anal. Non Linéaire **29** (2012), 479–500.
- [61] X. Shen, H. Sun and X. Tai, *Preconditioned algorithm for difference of convex functions with applications to graph Ginzburg–Landau model*, Multiscale Model. Simul. **21** (2023), 1667–1689.
- [62] J. Shi and J. Malik, *Normalized cuts and image segmentation*, IEEE Trans. Pattern Anal. Mach. Intell. **22** (2000), 888–905.

- [63] A. Szlam and X. Bresson, *Total variation, Cheeger cuts*, in: Proceedings of the 27th International Conference on Machine Learning, vol. 10, 2010, pp. 1039–1046.
- [64] V. E. Tarasov and G. M. Zaslavsky, *Fractional Ginzburg–Landau equation for fractal media*, Phys. A **354** (2005), 249–261.
- [65] M. Valadier, *Young measures*, Methods of nonconvex analysis, **1446** (1990), 152–188.
- [66] Y. van Gennip and A. L. Bertozzi,  $\Gamma$ -convergence of graph Ginzburg–Landau functionals, Adv. Differential Equations **17** (2012), 1115–1180.
- [67] Y. van Gennip, N. Guillen, B. Osting and A. L. Bertozzi, *Mean curvature, threshold dynamics, and phase field theory on finite graphs*, Milan J. Math. **82** (2014), 3–65.
- [68] Y. van Gennip, B. Hunter, R. Ahn, P. Elliott, K. Luh, M. Halvorson, S. Reid, M. Valasik, J. Wo, G. E. Tita, A. L. Bertozzi and P. J. Brantingham, *Community detection using spectral clustering on sparse geosocial data*, SIAM J. Appl. Math. **73** (2013), 67–83.
- [69] R. Vizueté, P. Frasca and F. Garin, *Graphon-based sensitivity analysis of SIS epidemics*, IEEE Control Syst. Lett. **4** (2020), 542–547.
- [70] P. J. Wolfe and S. C. Olhede, Nonparametric graphon estimation, arXiv preprint arXiv: 1309.5936, 2013.
- [71] Z. Wu and R. Leahy, *An optimal graph theoretic approach to data clustering: Theory and its application to image segmentation*, IEEE Trans. Pattern Anal. Mach. Intell. **15** (1993), 1101–1113.
- [72] F. Yi and I. Moon, *Image segmentation: A survey of graph-cut methods*, in: 2012 International Conference on Systems and Informatics (ICSAI 2012), Piscataway, NJ, USA, 2012, pp. 1936–1941.
- [73] J. Yuan, E. Bae and X.-C. Tai, *A study on continuous max-flow and min-cut approaches*, in: 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Piscataway, NJ, USA, 2010, pp. 2217–2224.
- [74] X. Zhang and T. F. Chan, *Wavelet inpainting by nonlocal total variation*, Inverse Probl. Imaging **4** (2010), 191–210.

E. J. ZHANG

Department of Applied Physics and Applied Mathematics, Columbia University, New York, NY, 10027, United States of America

Now affiliated with the Department of Mathematics, University of California, Los Angeles (UCLA), Los Angeles, CA, 90095, United States of America

*E-mail address:* `edith@math.ucla.edu`

J. M. SCOTT

Department of Mathematics and Statistics, Auburn University, Auburn, Alabama, 36849, United States of America

*E-mail address:* `james.m.scott@auburn.edu`

Q. DU

Department of Applied Physics and Applied Mathematics, Columbia University, New York, NY, 10027, United States of America; Data Science Institute, Columbia University, New York, NY, 10027, United States of America

*E-mail address:* `qd2125@columbia.edu`

M. A. PORTER

Department of Mathematics, University of California, Los Angeles (UCLA), Los Angeles, CA, 90095, United States of America; Department of Sociology, University of California, Los Angeles (UCLA), Los Angeles, CA, 90095, United States of America; Santa Fe Institute, Santa Fe, NM, 87501, United States of America

*E-mail address:* `mason@math.ucla.edu`