Ginzburg–Landau Functionals and Reaction–diffusion Equations in the Large-Graph Limit

Edith Jin Zhang

Submitted in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy
under the Executive Committee
of the Graduate School of Arts and Sciences

COLUMBIA UNIVERSITY

2025

Abstract

Ginzburg–Landau Functionals and Reaction–Diffusion Equations in the Large-Graph Limit

Edith Jin Zhang

Graphons are a notion of infinite graphs that are spatially continuous rather than discrete (as are the nodes of a graph). They have emerged as a mathematical formalism that attempts to preserve the essential properties of graphs as their sizes approach infinity. In this dissertation, we consider two problems on graphs, and we let the size of the graphs approach infinity. We define graphon limits of these two graph problems, and we show that the limiting problems on graphons are consistent with their graph counterparts in the sense that the solutions of the graph problems converge to the solutions of the limiting graphon problems.

In Chapters 3–6, we consider the problem of minimizing graph Ginzburg–Landau (GL) functionals. Graph GL functionals are relaxations of graph-cut functionals on graphs, and they have yielded a variety of insights in image segmentation and graph clustering. For sequences of graphs W_n that converge to a limiting graphon W, we show that the graph GL functional Γ-converges to a continuous and nonlocal functional that we call the *graphon GL functional*. We also show that graphon GL functionals Γ-converge, in a sharp-interface limit, to a graphon total variation (TV) functional. Using the Γ-convergence results, we conclude that the minimizers of the Γ-converging sequences of functionals converge to the minimizers of their limiting functionals. To obtain our convergence results, we require an extension of the underlying function spaces; we discuss the need for a probabilistic interpretation (via Young measures) of the variational problems in the graphon limit.

In Chapters 7–8, we study graph RD equations, which are certain systems of differential equations that are defined on the nodes of a graph. Consider a sequence of graphs that converges to a limiting graphon. We show that the solutions of the sequence of graph RD equations converge to the solution of a limiting graphon RD equation, which we define to be a continuous RD equation that has a nonlocal diffusion term. Furthermore, we show that a sequence of stochastic particle processes (that consist of a random walk and a birth-death process) on the sequence of graphs converges to the solution of the graphon RD equation.

The graphon limits of graph problems can be viewed as continuous and nonlocal counterparts of discrete graph problems, which tend to be large systems of coupled equations. Our results establish precedent and intuition for future work on graphon limits of graph problems, and they build the connection between graphon theory and nonlocal analysis.

Table of Contents

Acknowledgments v

Chapter 1: Introduction 1

1.1 Motivation and background 1

1.2 Contributions 3

1.3 Organization 4

Chapter 2: Graphons 5

2.1 Graphons 5

2.1.1 L^p graphons 6

2.2 Cut norm and convergence 8

2.3 Graphons as nonlocal operators 10

Chapter 3: Ginzburg–Landau and Total-variation Functionals on Graphons 11

3.1 Introduction 11

3.1.1 Related work 14

3.2 Graph functions and functionals 15

3.3 Γ-convergence 19

3.4 Young measures and weak convergence of functions 20

3.5 Graphon functions and functionals 22

Chapter 4: Sequential Limit: ϵ then n (i.e., $\mathrm{GL}_\epsilon^{W_n} \xrightarrow{\Gamma} \mathrm{TV}^{W_n} \xrightarrow{\Gamma} \mathrm{TV}^{W}$) 24

4.1 Limit (1): $\mathrm{GL}_\epsilon^{W_n} \xrightarrow{\Gamma} \mathrm{TV}^{W_n}$ 24

4.2 Limit (2): $\mathrm{TV}^{W_n} \xrightarrow{\Gamma} \mathrm{TV}^{W}$ 25

Chapter 5: Sequential Limit: n then ϵ (i.e., $\mathrm{GL}_\epsilon^{W_n} \xrightarrow{\Gamma} \mathrm{GL}_\epsilon^{W} \xrightarrow{\Gamma} \mathrm{TV}^{W}$) 26

5.1 Limit (3): $\mathrm{GL}_\epsilon^{W_n} \xrightarrow{\Gamma} \mathrm{GL}_\epsilon^{W}$ 26

5.1.1 Well-posedness of GL_ϵ^{W} in $L^\infty((0,1);[-1,1])$ 26

5.1.2 The proof of limit (3) 28

5.2 Limit (4): $\mathrm{GL}_\epsilon^{W} \xrightarrow{\Gamma} TV^{W}$ as $\epsilon \to 0$ for $W \in L^\infty$. 34

5.2.1 The issue of ϵ-scaling 34

Chapter 6: The GL minimizer for several examples 38

6.1 The constant graphon 39

6.2 2×2 stochastic block models (SBMs) 45

6.2.1 Complete bipartite graphon 47

6.2.2 Community-structure graphon 51

Chapter 7: Graphon reaction–diffusion equations 57

7.1 Introduction 57

7.1.1 Relationship to GL equation 58

7.1.2 Related work 58

7.1.3 Contributions 60

7.2 Graphon diffusion equation 61

7.2.1 Graphon diffusion equation 62

7.3 Graphon reaction–diffusion equation 67

7.3.1 Graphon reaction–diffusion equation 67

Chapter 8: Law of large numbers 71

8.0.1 Stochastic term 76

Conclusion 87

References 91

List of Figures

2.1 An example of a graph, its associated adjacency matrix, and its corresponding step graphon. 6

3.1 Four different Γ-convergences of the graph GL functional. 13

6.1 Three types of 2×2 piecewise-constant graphons. 45

7.1 Contributions relating to graph diffusion equations. 59

7.2 Contributions relating to reaction–diffusion equations. 60

Acknowledgements

There are many people who made my past six years at Columbia mostly enjoyable and mostly productive. During my first year at Columbia, I spent a lot of time working on (and comiserating about) problem sets with my friend Peter Wang with whom I shared nearly every single class of both semesters. My second year was completely remote due to the pandemic, and I spent it in Virginia with family. I was able to spend a lot of quality time with my family, and my brother Edward helped me with programming. In the following four years, I spent good time with my friends Vaishnavi Murthy, Ali Zare, William Giovanni Jimenez Senzano IV, Ray Gardella, Aria Halavati, Emily Glazer, Karen Yu, Sean Cohen, Remy Kassem, Madison Ihrig, Jonah Chaban, Mason Cannon, Mark Mazel, Marcus Cooley, Mustafa Ozkoc, Frederico Levy, Esha Trivedi, Jimmy Wang, Stephen Gomez-Peck, Hannah Kirk Nass, Kat Novak, Sofia Barandiaran, James Miller, Claire Chapeau, Lakshman Prabhakar, Rachel Chon, Javed Shaik, Katherine Qian, Katherine Yan, and others.

I am glad that my advisor, Qiang Du, let me explore anything I found interesting and provided his support in all of those pursuits. He supplied ideas on the string method in my “ELBO-MEP" project, suggested the Jordan–Kinderlehrer–Otto connection in my project on variational inference and Fokker–Planck equations, and identified the connection between nonlocal analysis and graphons when I was trying to understand graphon limits. One of the first papers he showed me was a paper by Andrea Bertozzi on graph GL functionals, and funnily enough, after much effort on other projects, graph(on) GL ended up being a main topic of my dissertation.

James Scott was extremely generous with his time and effort in the graphon GL project. He helped identify the importance of Young measures for graphon functionals, and his knowledge in Calculus of Variations were necessary in forming our Γ-convergence proofs. He was happy to get in the weeds of our proofs and explain his ideas. He also gave me two house plants, back in 2021, that are still doing well.

Andrew Blumberg (Columbia Mathematics) helped me become interested in geoemetric data analysis, which occurred when I took his course on that subject in Spring 2021. I enjoyed our conversations while walking around the Columbia campus and Morningside Park in the fall and winter of 2021, and I am especially glad that he helped set up my visit to UCLA in summer of 2022. My collaboration with Mason Porter would not be possible without him. I also learned a lot from my discussions with Chris Wiggins on an eclectic collection of topics, from diffusion bandits (during the pandemic) to graphon spectra (in 2024).

Mason Porter (UCLA Mathematics) took a risk to agree to mentor me for a summer, and he was quite generous with his time during that summer. I am glad that he brought the theory of graphons to my attention. Throughout our (remote) process of developing our paper on graphon GL functionals, he was a careful editor and an extremely useful mentor on academic writing.

Lastly, and most importantly, I am grateful to my family and my partner Ethan for their support and care.

Chapter 1: Introduction

1.1 Motivation and background

The field of applied graph theory is rich and increasingly-studied [1]. Graphs describe systems of individuals that have heterogeneous interactions; examples of real-world graphs include social networks, gene-regulatory networks, and the internet [2, 3, 4, 5, 6, 7]. These graphs can have billions of nodes. Such large graphs are difficult to visualize and understand, and they are intractable when using traditional computational techniques.

In this dissertation, we study two graph-theoretic problems using graphons, which are a type of infinite graph. A graphon is a symmetric and measurable function $W : [0,1]^2 \to \mathbb{R}$ that is a continuum limit of adjacency matrices for increasingly-large graphs [8]. Graphons have become popular in both applications [2, 3, 4, 5, 6, 7] and theoretical studies [9, 10, 11, 12, 8].

There are several reasons to consider graphon limits of graph problems. Graph-based models and computations ought to be robust under small changes in a graph's structure (e.g. the addition or removal of a vertex or edge), and they should work for graphs of different sizes that have similar underlying structure [13]. Furthermore, solving problems on large graphs is often computationally intractable. Graphons provide a way to compare different networks that have different sizes and densities; one consequence is that a large graph may be approximated by a smaller one that has similar structure [7].

Furthermore, because graphons are a familiar object in nonlocal analysis (see Section 2.3), one can use computational techniques from nonlocal analysis (e.g., see [14]) for applications that involve large graphs. Graphons have also been employed in mean-field approximations of dynamical systems on graphs [15, 2]. Mean-field systems are a method that is commonly used in statistical physics to model continuum limits of a large number of agents that each interact with the rest of the population as a whole. Used in this way, graphons can describe the (possibly heterogeneous) interactions between a large number of interacting agents [16].

In this dissertation, we investigate two related problems that graphons are well-suited to due to their continuous nature. In both cases, we show that the problem on large graphs can be approximated by a limiting problem on graphons. First is a class of optimization problems on graphs, and second is a class of partial differential equations (PDEs) on graphs.

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). Because graph optimization problems are discrete, they are combinatorial in nature. In general, combinatorial problems are harder to solve than continuous problems. Approximating optimization problems on graphs by their limiting problems on graphons opens new avenues for approximating the minimizers of the graph problems. Graphons are well-suited to optimization problems because it is possible to continuously deform a graphon and thereby calculate variations of graphon functionals [8, Section 16.2].

One well-known optimization problem on graphs is the minimum-cut problem [17]; it involves minimizing the graph-cut functional. The graph-cut functional is used in applications such as community detection [18, 19], image segmentation [20, 21, 22, 23, 24], among other applications. It is also closely related to the maximum-flow (max-flow) problem on networks [25, 26, 27] and to the total-variation (TV) functional [28]. We study the *graph Ginzburg–Landau (GL) functional*, which is an approximation to the min-cut functional, and we develop the theory behind a *graphon GL functional* and a related *graphon TV functional*. Graphons connect graph functionals to nonlocal functionals [29, 30]. For instance, one can view a graphon TV functional (which we will define in Chapter 3) as a nonlocal perimeter functional [31, 32, 33].

The second set of problems that we consider in the graphon limit relates to *reaction–diffusion (RD) equations* on graphs and graphons. RD equations are partial differential equations that describe the concentration of a substance in which particles diffuse (i.e., travel) and react (i.e., appear and disappear). They have been used in fields such as biology, ecology, chemistry, and physics [34, 35, 36, 37]. RD equations are gradient flows of energy functionals of which the GL functional is a special case. (The gradient flow of the GL functional is called the *Allen–Cahn equation*, which is an RD equation.) In other words, solving RD equations is a time-dependent (i.e., dynamical) problem that is related to the time-independent (i.e., steady-state) GL-minimization problem discussed above.

1.2 Contributions

Our main contributions consist of two results relating to graphon GL functionals, and three results relating to graphon RD equations. For all our results, suppose that a sequence of graphs $\{W_n\}$ converges with respect to cut norm (which is defined in Chapter 2) to a limiting graphon W.

Our contributions that relate to graphon GL functionals are as follows. Corollary 3 establishes that the sequence of graph GL functionals corresponding to the graphs W_n Γ-converges to the graphon GL functional corresponding to the graphon W. Theorem 5.2.1 establishes that, for a fixed graphon W, the graphon GL functional Γ-converges to a graphon TV functional. We prove accompanying compactness properties of the domains of the functionals, which guarantees that the minimizers of the graph GL functionals converge to the minimizers of the graphon GL functionals, which in turn converge to the minimizers of the graphon TV functionals.

Our contributions that relate to graphon RD equations are as follows. Theorem 7.2.1 shows that the solutions of the sequence of graph diffusion equations corresponding to the sequence $\{W_n\}$ converges to the solution of the graphon diffusion equation corresponding to the limiting graphon. The convergence occurs with a linear rate of convergence in L^p norm for $p \geq 1$. Theorem 7.3.1 shows that the solutions of the sequence of graph RD equations corresponding to the sequence $\{W_n\}$ converge to the solution of the limiting graphon RD equation corresponding to the graphon W. The convergence occurs in L^p norm for $p \geq 1$ and is linear. Lastly, Theorem 8.0.1 is a (weak) law of large numbers (LLN) result that shows that birth-death processes on the sequence of graphs $\{W_n\}$ converge in probability to the solution of a graphon RD equation corresponding to the graphon W.

In extending these two related graph problems to their graphon limits, we help develop the theoretical groundwork for graph-to-graphon limits that can be used for other optimization problems and PDEs, among other problems, on graphs.

1.3 Organization

Chapter 2 defines graphons, including the set of graphons that are associated to graphs; the relevant norm of convergence, called cut norm, and its properties; and different function spaces that graphons may reside in. Chapter 2 also discusses how graphons fit into a functional-analytic framework; they are kernels corresponding to nonlocal operators.

In Chapter 3, we introduce and set up the problem of Γ-convergence of graphon GL functionals. We define the relevant functions and functionals on graphs and graphons, the concept of Γ-convergence, and Young measures. We discuss the role of Young measures in our graphon functionals and their importance for compactness results.

In Chapters 4 and 5, we state our Γ-convergence results for the graph GL functional. Chapter 4 is brief, while Chapter 5 is more involved and deals with issues of boundedness and scaling.

Chapter 7 centers on the convergence of the solutions of sequences of graph RD equations to the solution of a limiting graphon RD equation. We define RD equations and the graphon diffusion operator, and state the graph-to-graphon convergence results for the graph diffusion equation and the graph RD equation.

In Chapter 8, we define a stochastic process on graphs and prove a law of large numbers (LLN) result that guarantees that the stochastic process converges in probability to the solution of a limiting graphon RD equation.

Chapter 2: Graphons

2.1 Graphons

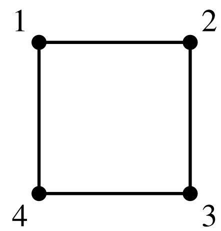
A *graphon*, which is a portmanteau of “graph" and “function", is a bounded, measurable, and symmetric function $W : \Omega^2 \to \mathbb{R}$, where $\Omega \subset \mathbb{R}^d$ is a connected and bounded domain. The set of graphons is $\mathcal{W} = \{W : \Omega^2 \to \mathbb{R}\}$. Throughout this 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 in order to avoid complications relating to isolated points. In the present paper, it makes no difference because W is always integrated. We use the open interval to be consistent with the conventions of functional analysis, which provides the main technical machinery in our paper. More generally, one can choose Ω to be any domain in $\mathbb{R}^d$.

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 $[n] = \{1, \ldots, n\}$. It has an associated adjacency matrix $A^{(n)}$ with entries $A^{(n)}_{ij}$. We associate the graph W_n with a function W_n that takes the constant value $A^{(n)}_{ij}$ on the product interval $I_i \times I_j$, where $I_i = [(i-1)/n,\ i/n)$ for $i = 1, \ldots, n-1$ and $I_0 = (0, 1/n)$. The relationship between the adjacency matrix and the graphon is thus

$$W_n(x, y) = \left\{ A^{(n)}_{ij} \quad \text{for } (x, y) \in I_i \times I_j\,,\ i, j \in \{1, \ldots, n\} \right\}\,. \tag{2.1}$$

We have thereby associated a graph with a step function by associating the ith node with the interval I_i and associating each edge (i, j) with the product interval $I_i \times I_j$. In this way, one can identify each finite graph W_n with a graphon, which we will also denote by W_n (see Remark 2.2.1). We refer to a graphon that corresponds to a finite graph as a *step graphon*.

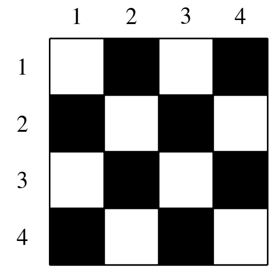
In Figure 2.1, we show an example graph and its corresponding adjacency matrix and step graphon. By convention, the axes of the graphon begin at the upper left.



The 4-cycle graph.

$$\begin{pmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{pmatrix}$$

Its adjacency matrix.



Its corresponding step graphon.

Figure 2.1: An example of a graph, its associated adjacency matrix, and its corresponding step graphon.

One can view any graphon $W : (0,1)^2 \to \mathbb{R}$ 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, where a finite graph is a randomly-drawn sample from a graphon, and (2) limits of sequences of growing graphs. We employ the latter interpretation of graphons.

2.1.1 L^p graphons

Lovasz's original formulation of graphons [8] are known as "L^∞ graphons" because they are functions in the space

$$L^\infty((0,1)^2) = \{W \ : \ \|W\|_\infty := \text{ess sup}_{x,y\in(0,1)^2}|W(x,y)| < \infty\} \,. \tag{2.2}$$

This set of L^∞ graphons [8] arise as limits of dense sequences of graphs[1], which are sequences of graphs whose number of edges increase as $\Theta(n^2)$ in the number n of nodes, which means that the number of edges is bounded above by $c_1 n^2$ and below by $c_2 n^2$ for some constants $c_1, c_2 > 0$. It is common to denote the set of L^∞ graphons by $\mathcal{W}$ 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

[1]Density is a property of sequences of graphs, rather then a property of graphs themselves, because the definition of density is based on the rate at which the number of edges increases in comparison to the number of nodes.

version $W/\|W\|_\infty \in \mathcal{W}_0$.

Sequences of growing graphs in many applications and real-world situations are sparse [18], but applications of these L^∞ graphons typically involve only mean-field models [16, 38]. To combat this issue, researchers have defined so-called L^p *graphons*, which arise as limits of sequences of *sparse graphs*. L^p graphons are obtained by normalizing graphs by their edge density, and they are associated with operators on $L^q((0,1)^2)$.

Sequences of graphs whose number of edges increase as $o(n^2)$ (i.e., grows at a rate strictly less than cn^2 for some constant c) converges to the zero graphon $W \equiv 0$ because the edge set is a set of measure 0 in the $n \to \infty$ 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 yet bounded in the $n \to \infty$ limit only if the number of nonzero terms in the adjacency matrix $A^{(n)} = \{A_{ij}^{(n)}\}_{i,j=1}^n$ is $\Theta(n^2)$.

To extend the theory of graphons to sparse sequences of graphs, researchers introduced L^p graphons [39, 40], 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^∞ graphons to the space

$$L^p((0,1)^2) = \left\{ W \ : \ \|W\|_p := \left(\int_0^1 \int_0^1 |W(x,y)|^p \,dx\,dy \right)^{\frac{1}{p}} < \infty \right\} \tag{2.3}$$

for $p \geq 1$.

L^p graphons are defined in [41, Definition 2.7], and sequences of graphons that converge to an L^p graphon are characterized in [41, Theorem 2.8]. For the purposes of the present paper, we view graphons as functions in $L^p((0,1)^2)$ for $p \in [1,\infty]$, meaning that we assume whatever necessary properties on the underlying graphs that allow them to be L^p functions.

When Ω is a bounded domain, $L^p(\Omega) \subset L^q(\Omega)$ for $p > q \geq 1$. Therefore, the set of L^∞ graphons, which includes all dense graphs and all bounded-degree 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 4.2 and 5.1), we assume that W_n and W are in L^1. In the proof of limit (4) (see Section 5.2, we assume that W_n and W are L^∞.

Researchers have also begun to use the traditional L^∞ graphons as limit objects of sequences of so-called “very sparse" graphs (i.e., sequences of graphs with bounded degree or bounded mean degree as $n \to \infty$ [39, 42]).

In this dissertation, we think of graphons simply as L^p functions, for $p \in \{1, 2, \dots, \infty\}$, rather than as limits of dense or sparse sequences of graphs.

2.2 Cut norm and convergence

The graphon cut norm (which is also known as the “cut norm" and is closely related to the graph-cut functional) is the choice of topology for the space of graphons. Namely, 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 2.2.1 (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*

$$\mathrm{Cut}(S, S^c) = \sum_{i \in S,\, j \in S^c} A_{ij}^{(n)} \,. \tag{2.4}$$

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

$$\mathrm{Cut}(u) = \frac{1}{8} \sum_{i,j=1}^{n} A_{ij}^{(n)} |u_i - u_j|^2 \quad \textit{with} \;\; u : [n] \to \{-1, 1\} \,. \tag{2.5}$$

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 such that $\mathcal{W}$ is a compact metric space [43]. Therefore, any bounded sequence $\{W_n\}_{n \in \mathbb{N}}$ of graphons has a subsequence $W_{n'}$ such that $\|W_{n'} -$

$W\|_\square \to 0$.

Definition 2.2.2. *The* cut norm *of a graphon W is*

$$\|W\|_\square = \sup_{S\subseteq(0,1)} \int_{S\times S^c} W(x,y)\,dx\,dy\,. \tag{2.6}$$

Definition 2.6 highlights the similarity between the cut norm and the graph-cut functional. When a graphon is a step graphon W_n (see equation (2.1)), the cut norm (2.6) 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 (2.6) [44]. A particularly useful one for the present paper is

$$\|W\|_\square = \sup_{\phi,\psi\in L^\infty((0,1);[-1,1])} \int_0^1\int_0^1 W(x,y)\phi(x)\psi(y)\,dx\,dy\,, \tag{2.7}$$

where

$$L^\infty((0,1);[-1,1]) = \{f : (0,1)\to[-1,1]\}\,. \tag{2.8}$$

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

$$\|W\|_\square = \sup_{\phi,\psi\in L^\infty((0,1))} \frac{1}{\|\phi\|_\infty}\frac{1}{\|\psi\|_\infty}\int_0^1\int_0^1 \phi(x)\psi(y)W(x,y)\,dx\,dy\,. \tag{2.9}$$

Remark 2.2.1. *Any finite weighted graph $W_n : [n] \to \mathbb{R}$ is at cut-norm distance 0 from its corresponding step graphon W_n. Therefore, we use the notation W_n for both objects. Similarly, each finite symmetric step function on $(0,1)^2 \to \mathbb{R}$ corresponds to a graph. 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 finite graph. See [45, Section 3.3], [44, Remark 4.6], and [43].*

Remark 2.2.2. *We use the notation "$\overset{\square}{\to}$" to denote convergence in cut norm. Accordingly, $W_n \overset{\square}{\to}$*

W means that $\|W_n - W\|_\square \to 0$.

Remark 2.2.3. *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* $T_W : L^\infty((0,1)) \to L^1(0,1)$*. Here, equivalence of norms* $\|\cdot\|_A$ *and* $\|\cdot\|_B$ *is defined by the property* $c_1\|\cdot\|_A \le \|\cdot\|_B \le c_2\|\cdot\|_A$ *for some constants* c_1, c_2*. In fact, the terms "graphon" and "kernel" are sometimes used interchangeably [44].*

Remark 2.2.4. *It is known that* $\|W\|_\square \le \|W\|_1$ *for any graphon W [44], where* $\|\cdot\|_1$ *is the* $L^1(\Omega^2, \mathbb{R})$ *norm. Additionally, for any step graphon with n steps,* $\|W_n\|_1 \le \sqrt{2n}\|W_n\|_\square$ *[8, 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* $n \to \infty$*.*

2.3 Graphons as nonlocal operators

Graphons are well-suited to the study of PDEs on graphs because they are spatially-continuous objects: they allow us to frame graph problems as functional-analytic problems, which opens up the possibility for more sophisticated problem-solving techniques than those that are available for discrete problems. For example, graphons have been used as kernels that induce nonlocal operators [46, 47, 48, 49]. The class of nonlocal operators that we are interested in are operators of the form

$$\mathcal{L}^W(u)(x) = \int_\Omega W(x,y)\big(u(y) - u(x)\big)dy\,, \tag{2.10}$$

where Ω is a bounded and connected domain in $\mathbb{R}^n$ and $W(x,y) : \Omega^2 \to \mathbb{R}$ is a kernel function that encodes the nontrivial connectivity structure. The operator $\mathcal{L}^W$ is called nonlocal because it depends on the difference $u(y) - u(x)$ for values of x and y that are not necessarily close to each other. In contrast, a local operator is an operator whose value at a single point depends only on the value (or derivatives) of a function at a single point.

If $W \in L^p((0,1)^2)$, then W is associated with an integral operator on $L^q((0,1)^2)$ for $q = \frac{p}{p-1}$. (Similarly, if $W \in L^\infty((0,1)^2$, then it is associated with an integral operator on $L^1((0,1)^2)$.)

Chapter 3: Ginzburg–Landau and Total-variation Functionals on Graphons

3.1 Introduction

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 equivalent to a graph TV functional [28] (see Remark 3.2.1). 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 attempt 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 Γ-convergence of the graph GL functional to the graph TV functional (equivalently, to the graph-cut functional), which was proved in [50], justifies the use of the graph GL functional as a relaxation of the graph TV (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 [51, 52, 53, 54].

In the present paper, we further relax the graph-GL minimization problem to the continuum using a large-graph limit. That is, we evaluate the limiting minimization problem on functions $(0, 1) \to [-1, 1]$ rather than on functions $\{1, \ldots, n\} \to [-1, 1]$. We use the idea of a *graphon* [8], 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.

To consider the convergence of graph GL functionals in the graphon limit, we use a central tool in variational calculus that is known as Γ-convergence. Provided the domain of the underlying function space is compact, the Γ-convergence of a sequence of functionals F_n to a limiting functional F guarantees that the minimizers of F_n converge to the minimizer of F [55].

The original GL theory is a physical model for phase transitions in superconductors [56, 57]. The GL functional, which is sometimes called the Allen–Cahn functional or the Modica–Mortola functional in some applications [58, 59, 60], is

$$\mathrm{GL}_\epsilon(u) = \epsilon \int_\Omega \left|\nabla u^2\right| \, dx + \frac{1}{\epsilon} \int_\Omega \Phi(u(x)) \, dx \,, \quad u \in H^1(\Omega) \,, \tag{3.1}$$

where $\Omega \in \mathbb{R}^d$ is a bounded and connected set, Φ is a double-well potential, and $H^1(\Omega) = W^{1,2}(\Omega) = \{f \in L^p(\Omega) : \left(\|f\|_2^2 + \|f'\|_2^2\right)^{1/2} < \infty\}$ is a Sobolev space [61]. As $\epsilon \to 0$, the functional GL_ϵ Γ-converges to the TV functional (i.e., perimeter function) [62]

$$\mathrm{TV}(u) = \int_\Omega |\nabla u| \, dx \,. \tag{3.2}$$

We take inspiration from van Gennip and Bertozzi [50], who defined the graph GL functional, which is a discrete version of GL_ϵ for functions u on graphs. See Section 3.2 for the definitions of the graph GL and other graph functionals. Van Gennip and Bertozzi proved Γ-convergence of the graph GL functional for a square-lattice graph. They derived both a large-graph (i.e., $n \to \infty$) limit and a sharp-interface (i.e., $\epsilon \to 0$) limit of this functional. As $n \to \infty$, the square-lattice graph of grid size $1/n$ converges to the region $(0, 1)^2 \subset \mathbb{R}^2$.

The growing square-lattice graph is a specific case of a growing graph sequence. One can view it as a mesh approximation of the unit square. In the present paper, we generalize van Gennip and Bertozzi's results to general growing graph sequences, each of which converges to some graphon as $n \to \infty$. We refer to the limit of sequences of graph GL functionals as a *graphon GL functional*. We show that the graphon GL functional Γ-converges to a nonlocal TV functional as $\epsilon \to 0$.

Figure 3.1 illustrates two different sequential Γ-limits of the graph GL functional (3.8). The arrows indicate Γ-convergence. For definitions of the functionals, see Section 3.5. One of the sequential limits is an $n \to \infty$ and then $\epsilon \to 0$ limit; the other is an $\epsilon \to 0$ and then $n \to \infty$ limit. In Sections 3.2 and 3.5, we define the associated limiting functionals. Limit (1) was proven in [50, Theorem 3.1], and limit (2) follows from [11, Theorem 12]. To prove limit (3), we use an approach that is similar to the proof of [11, Theorem 12]. We significantly generalize [50, Theorem 5.2] in the sense that our graph limits are for general sequences of growing graphs, rather than only for a growing square-lattice graph. Limit (4) resembles the classical Modica-Mortola limit [60] (which states that $\mathrm{GL}_\epsilon \xrightarrow{\Gamma} TV$), but concerns the graphon versions of those functionals. We prove limit (4) for $W \in L^\infty((0,1)^2)$ (i.e., graphons that correspond to dense sequences of graphs). Limit (4) is not defined for more general L^p graphons.

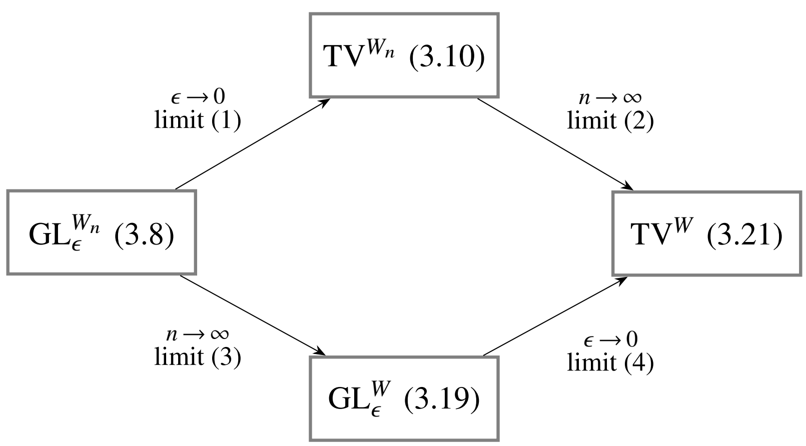


Figure 3.1: Four different Γ-convergences of the graph GL functional.

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

3.1.1 Related work

Our work is inspired by van Gennip and Bertozzi [50], who examined the same four limits as in Figure 3.1 and also the simultaneous limit $\epsilon \to 0$, $n \to \infty$. Their $n \to \infty$ limits hold only in the case of the square-lattice graph. We restate their version of limit (1), which holds for all graphs, and we extend their versions of limits (2)–(4) to a nonlocal case of graph limits in which the limiting GL and TV functionals are the graphon GL and TV functionals (3.19) and (3.21), respectively. Limits (2)–(4) hold for general sequences of graphs, which is why the limiting functionals are nonlocal.

We follow ideas from Braides et al [11], who used Young measures to show that the graph-cut functional (which we define later in (2.5)) Γ-converges to the graphon-cut functional. Our methods and results are related more closely to the results of [11] than to those of [50], but they still differ fundamentally because the domain of the cut functional consists of finite-range functions, whereas the domains of the graph and graphon GL functionals are continuous-range functions.

Trillos and Slepcev [63] also studied a local limit and, analogously to limit (2), obtained Γ-convergence of the perimeter functional for the $n \to \infty$ limit of point clouds. In [64], the same authors obtain a similar Γ-limit in a metric space characterized by optimal transportation maps known as the TL^p metric. A key difference between their paper and ours is that we do not require the limiting functional to be a local quantity, such as Euclidean perimeter or a Dirichlet-type energy. Instead, our limiting functionals are given by nonlocal limits with interaction potential given by graphons. These functionals are a generalized "relaxation," in that local functionals are recovered from them for special cases of singular graphons.

The original theory of graphons [8] considers dense sequences of graphs, and Braides et al. [11] also required the graphs in such sequences to be dense. However, most real-world graphs are sparse [18, 39], 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 [41, 40], and our analysis allows sequences of sparse graphs that converge to L^p graphons. See Section 2.1.1 for more detail.

One recovers different types of TV functionals when taking the $\epsilon \to 0$ limit of the classical GL, graph GL, and graphon GL functionals. In particular, we obtain a nonlocal, continuous TV functional in the $\epsilon \to 0$ limit of the graphon GL. Nonlocal TV functionals have been useful in a variety of applications, especially in image processing [32, 31, 33, 65, 29, 66, 67, 68, 69]. They are also of theoretical interest because they generalize the notion of perimeter from objects with geometric regularity like $\mathbb{R}^2$, on which one can compute gradients of functions, to less regular objects such as metric spaces[70]. The fractional GL functional is a well-studied example of a nonlocal GL functional [71, 72, 73], but thus far it has not been connected to graph theory. A graphon GL is a fractional GL when a graphon is of the form $W(x, y) = \frac{1}{|x-y|^{1+2s}}$, with $s \in (0, 1)$.

3.2 Graph functions and functionals

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 Γ-convergence of those functionals. It is also relevant to analyze the function spaces on which the functionals act.

We consider GL functionals and TV functionals, which are two basic types of 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 (3.1) and the classical TV functional in (3.2). The graph GL and TV functions are discrete and were defined in [74] and [24], respectively. We discuss them in this section. In Section 3.5, we define graph GL and TV functionals, which are continuous and involve Young measures.

We consider two types of function spaces: (1) function spaces of graphons and (2) function spaces of functions on those graphons. As we discussed in Section 2.1, 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

$$\mathcal{V}^n = \{\hat{u} : [n] \to \mathbb{R}\} . \tag{3.3}$$

Each function $\hat{u} \in \mathcal{V}^n$ has an associated step function $u : (0,1) \to \mathbb{R}$. Let $u(x) = \hat{u}(i)$ for $x \in I_i = [(i-1)/n,\ i/n)$. 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

$$\mathcal{V}_b^n = \{u : [n] \to \mathbb{R}\,, u_i \in \{\pm 1\} \text{ for all } i\} \tag{3.4}$$

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

Recall the classical GL functional (3.1), which is defined on functions $u : (0,1) \to \mathbb{R}$ by

$$\mathrm{GL}_\epsilon(u) = \epsilon \int_0^1 |\nabla u|^2 \, dx + \frac{1}{\epsilon} \int_0^1 \Phi(u(x)) \, dx \,, \tag{3.5}$$

where Φ is a double-well potential that has zeros at $s = \pm 1$. The double-well potential Φ can take a general form (see [50, assumptions W_1–W_4]), but we use the standard choice

$$\Phi(s) = (s^2 - 1)^2 \,. \tag{3.6}$$

Note that the classical GL functional 3.1 is defined on functions in $W^{1,2}((0,1))$ (that is, the Sobolev space of functions on $(0,1)$ that are L^2 and whose first derivative is also L^2). The classical TV functionals 3.2 is defined on $W^{1,1}(0,1)$. We extend the domain of both functionals to $L^\infty((0,1))$ by letting them evaluate to $+\infty$ when undefined.

The graph version $\mathrm{GL}_\epsilon^{W_n}$ of the GL functional is analogous to GL_ϵ. It acts on $u \in \mathcal{V}^n$, instead of on $u \in L^\infty((0,1))$. 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 Section 2.2 of [50] for further discussion of graph analogues of calculus operators.) For $u \in \mathcal{V}^n$, we thus obtain

$$\mathrm{GL}_{\epsilon}^{W_n}(u) = \frac{1}{n^2}\sum_{i,j=1}^{n} A_{ij}^{(n)}|u_i - u_j|^2 + \frac{1}{\epsilon n}\sum_{i=1}^{n}\Phi(u_i) \tag{3.7}$$

$$= \int_0^1\int_0^1 W_n(x,y)|u(x)-u(y)|^2 dx dy + \frac{1}{\epsilon}\int_0^1 \Phi(u(x))\,dx\,, \tag{3.8}$$

where (3.7) uses the adjacency matrix and (3.8) uses the definition (2.1) of the step graphon W_n. Similarly, the graph TV functional replaces the term $|\nabla u|$ in (3.2) with a finite difference. It is finite only for binary functions. This functional is

$$\mathrm{TV}^{W_n}(u) = \begin{cases} \sum_{i,j=1}^{n} A_{ij}^{(n)}|u_i - u_j| & \text{if } u \in \mathcal{V}_b^n \\ +\infty & \text{if } u \in \mathcal{V}^n \setminus \mathcal{V}_b^n \end{cases} \tag{3.9}$$

$$= \begin{cases} \int_0^1\int_0^1 W_n(x,y)|u(x)-u(y)|\,dx\,dy & \text{if } u \in \mathcal{V}_b^n \\ +\infty & \text{if } u \in \mathcal{V}^n \setminus \mathcal{V}_b^n\,. \end{cases} \tag{3.10}$$

The Dirichlet energy

$$D(u) = \int_0^1 |\nabla u(x)|^2\,dx \tag{3.11}$$

has a similar form to 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 defining $D(u) = +\infty$ for $u \in L^\infty(0,1)\setminus W^{1,2}(0,1)$. The graph Dirichlet energy

$$D^{W_n}(u) = \begin{cases} \int_0^1\int_0^1 W_n(x,y)|u(x)-u(y)|^2\,dx\,dy & \text{if } u \in \mathcal{V}^n \\ +\infty & \text{if } u \in L^\infty((0,1))\setminus\mathcal{V}^n \end{cases} \tag{3.12}$$

acts on graph functions and replaces the gradient term in (3.11) with a finite difference. The graph Dirichlet energy is a generalization of the graph-cut functional (2.5). If we restrict the graph functions to the range $\{-1,1\}$, then the graph Dirichlet energy is equal to the graph-cut functional; see (3.7) and (3.8).

Remark 3.2.1 (Equivalence of graph-cut, Dirichlet energy, and TV functionals)**.** *The graph-cut functional, which is defined in Equation* (2.5)*, is equal to the graph Dirichlet energy* (3.12) *of the graph function u. The graph-cut is also equal to the graph TV functional* (3.9) *when u has the range* $\{-1, 1\}$.

We define the graphon GL functional

$$\mathrm{GL}_{\epsilon}^{W}(\nu) = \int_0^1 \int_0^1 W(x,y) \int_{\mathbb{R}^2} |\lambda-\mu|^2 \, d\nu_x(\lambda) \, d\nu_y(\mu) \, dx \, dy + \frac{1}{\epsilon} \int_0^1 \int_{\mathbb{R}} \Phi(\lambda) \, d\nu_x(\lambda) \, dx \,. \quad (3.13)$$

The first term of (3.19) is the graphon Dirichlet energy

$$D^W(\nu) = \int_0^1 \int_0^1 W(x,y) \int_{\mathbb{R}^2} |\lambda - \mu|^2 \, d\nu_x(\lambda) \, d\nu_y(\mu) \, dx \, dy \,, \quad (3.14)$$

which is the graphon analogue of the graph Dirichlet energy (3.12). For finite graphs W_n, we use this definition, which is more general than the graph Dirichlet energy (3.12). Equation (3.20) reduces to (3.12) when ν 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 ν_x and ν_y. This is a probabilistic analogue of the term $|u(x) - u(y)|^2$ in the graph GL functional. 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

$$\mathrm{TV}^W(\nu) = \begin{cases} 2\int_0^1 \int_0^1 W(x,y) \int_{\mathbb{R}^2} |\lambda - \mu| \, d\nu_x(\lambda) \, d\nu_y(\mu) \, dx \, dy & \text{if } \nu \in \mathcal{Y}^b \\ +\infty & \text{if } \nu \in \mathcal{Y}((0,1),\mathbb{R}) \setminus \mathcal{Y}^b \,, \end{cases} \quad (3.15)$$

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

To interpret the Young-measure functionals, we think of the "slice" ν_x as analogous to $u_n(x)$, which is the "slice" of the graph function at x. The graphon functionals GL_ϵ^W and TV^W give energies of the state ν, just as the graph functionals $\mathrm{GL}_\epsilon^{W_n}$ and TV^{W_n} give energies of the state u_n of a graph.

3.3 Γ-convergence

The notion of Γ-convergence of functionals is useful in optimization and the calculus of variations [55, 50]. In concert with a certain compactness property, the Γ-convergence of a sequence of functionals guarantees the convergence of corresponding minimizers of the sequence of functionals.

Definition 3.3.1. *Let X be a metric space, and let $F_n : X \to \mathbb{R} \cup \{\pm\infty\}$ be a sequence of functionals. We say that F_n* Γ-converges *to $F : X \to \mathbb{R} \cup \{\pm\infty\}$, which we denote by $F_n \xrightarrow{\Gamma} F$, with respect to $u_n \to u$ if*

1. *for every sequence $\{u_n\}$ such that $u_n \to u$, we have $\liminf_{n\to\infty} F_n(u_n) \geq F(u)$;*

2. *there exists a sequence $\{u_n\}_{n\in\mathbb{N}}$ such that $\limsup_{n\to\infty} F_n(u_n) \leq F(u)$.*

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" [50].

It is useful to be purposeful when choosing the metric under which u_n converges to u. If the convergence metric is stronger, then it is easier to prove the lim inf–inequality for Γ-convergence but harder to construct a recovery sequence for the lim sup–inequality and for compactness. Conversely, if the convergence metric is weaker, then it is harder to prove the lim inf–inequality but easier to construct a recovery sequence. In the present paper, our Γ-convergence results are with respect to narrow convergence of Young measures (see Definition 3.4.4).

3.4 Young measures and weak convergence of functions

In our study of Γ-convergence, we need to use Young measures [75], which extend the feasible set of the GL-minimization and TV-minimization problems to a space of measures. In general, the study of functionals such as $\int_0^1 f(v(x))dx$ – for given continuous bounded f – that act on $v \in L^\infty((0,1))$ present the following challenge: a sequence $\{v_j\}$ may converge to a function v in the weak-* topology on $L^\infty((0,1))$, but it is *not* necessarily the case that $f(v_j) \overset{\star}{\rightharpoonup} f(v)$. Our functionals ($\mathrm{GL}_\epsilon^W$, $\mathrm{GL}_\epsilon^{W_n}$, and so on) act on $L^\infty((0,1))$, and exhibit this kind of structure. To overcome this challenge, we extend these functionals to act on the set of Young measures. Intuitively, the use of Young measures in the present context accommodates the rapidly oscillating minima of the graphon GL functionals. The minima can oscillate arbitrarily rapidly, and Young measures represent an effective limit of those oscillations when composed with continuous bounded functions in the integrals. We discuss this further in Chapter 6.

It is useful to review some definitions and properties that were presented in [11].

Definition 3.4.1 (Young measure)**.** *A* Young measure ν *on* $(0,1)\times\mathbb{R}$ *is a family* $\{\nu_x\}_{x\in(0,1)}$ *of probability measures, which is parametrized by* $x \in (0,1)$*, such that the map*

$$x \mapsto \int_{\mathbb{R}} f(\lambda)\, d\nu_x(\lambda) \tag{3.16}$$

is a Lebesgue-measurable function for every continuous and bounded $f \in C^b(\mathbb{R})$.

Intuitively, ν_x is a "slice" of the Young measure ν at the value 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 3.4.2 (Young measure corresponding to a measurable function)**.** *A* Young measure corresponding to a Lebesgue-measurable function $u : (0,1) \to \mathbb{R}$ *is the family of delta measures*

$$\{\nu_x\}_{x\in(0,1)} = \{\delta_{u(x)}\}_{x\in(0,1)}\,. \tag{3.17}$$

We refer to such measures as *δ-Young measures.* With the definition of ν in equation (3.17), the map (3.16) is the evaluation map $x \mapsto f(u(x))$.This evaluation map is indeed Lebesgue-measurable as required in Definition 3.4.1 because of the Lebesgue-measurability of u and the continuity of f.

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

Definition 3.4.3 (Weak-* topology in $L^\infty((0,1))$ space)**.** *A sequence $f_n \in L^\infty((0,1))$ converges in the weak-* topology to $f \in L^\infty((0,1))$ if*

$$\lim_{n\to\infty} \int_0^1 f_n(x)g(x)\,dx = \int_0^1 f(x)g(x)\,dx$$

for any $g \in L^1((0,1))$. We then say that $f_n \overset{\star}{\rightharpoonup} f$ in L^∞.

Definition 3.4.4 (Narrow convergence of Young measures)**.** *A sequence $\nu^n \in \mathcal{Y}((0,1),\mathbb{R})$ con-*verges narrowly *to $\nu \in \mathcal{Y}((0,1),\mathbb{R})$ if the map* (3.16) *converges in the weak-* topology in $L^\infty((0,1))$ for all continuous and bounded functions $f \in C^b(\mathbb{R})$. That is,*

$$\int_{\mathbb{R}} f(\lambda)\,d\nu_x^n(\lambda) \overset{\star}{\rightharpoonup} \int_{\mathbb{R}} f(\lambda)\,d\nu_x(\lambda)\,. \tag{3.18}$$

Lemma 3.4.1 (Narrow convergence of product Young measures [11, Lemma 8])**.** *Let ν^n be a sequence of Young measures that converges narrowly to $\nu \in \mathcal{Y}((0,1),\mathbb{R})$. Then the sequence of product measures $\nu^n \otimes \nu^n$ 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*

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

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 next lemma is the key compactness property of $\mathcal{Y}((0,1),\mathbb{R})$ that justifies the use of Young measures in our Γ-convergence result.

Lemma 3.4.2 (Prohorov's Theorem, [11, Theorem 9])**.** *Let* $\{u_n\}_{n\in\mathbb{N}}$ *be a bounded sequence in* $L^1((0,1))$*, and let* $\{\nu^n\}_{n\in\mathbb{N}}$ *denote the sequence of corresponding Young measures (see equation* (3.17)*). There then exists a subsequence* $\{u_{n_k}\}_{k\in\mathbb{N}}$ *and a Young measure* ν *such that* ν^{n_k} *converges narrowly to* ν *as* $k \to \infty$*.*

3.5 Graphon functions and functionals

We define the graphon GL functional

$$\mathrm{GL}_\epsilon^W(\nu) = \int_0^1 \int_0^1 W(x,y) \int_{\mathbb{R}^2} |\lambda-\mu|^2 \, d\nu_x(\lambda)\, d\nu_y(\mu)\, dx\, dy + \frac{1}{\epsilon} \int_0^1 \int_{\mathbb{R}} \Phi(\lambda)\, d\nu_x(\lambda)\, dx\,. \quad (3.19)$$

The first term of (3.19) is the graphon Dirichlet energy

$$D^W(\nu) = \int_0^1 \int_0^1 W(x,y) \int_{\mathbb{R}^2} |\lambda-\mu|^2 \, d\nu_x(\lambda)\, d\nu_y(\mu)\, dx\, dy\,, \quad (3.20)$$

which is the graphon analogue of the graph Dirichlet energy (3.12). For finite graphs W_n, we use this definition, which is more general than the graph Dirichlet energy (3.12). Equation (3.20) reduces to (3.12) when ν 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 ν_x and ν_y. This is a probabilistic analogue of the term $|u(x)-u(y)|^2$ in the graph GL functional. 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

$$\mathrm{TV}^W(\nu) = \begin{cases} 2\int_0^1\int_0^1 W(x,y)\int_{\mathbb{R}^2} |\lambda-\mu|\, d\nu_x(\lambda)\, d\nu_y(\mu)\, dx\, dy & \text{if } \nu \in \mathcal{Y}^b \\ +\infty & \text{if } \nu \in \mathcal{Y}((0,1),\mathbb{R}) \setminus \mathcal{Y}^b\,, \end{cases} \tag{3.21}$$

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

To interpret the Young-measure functionals, we think of the “slice" ν_x as analogous to $u_n(x)$, which is the “slice" of the graph function at x. The graphon functionals GL_ϵ^W and TV^W give energies of the state ν, just as the graph functionals $\mathrm{GL}_\epsilon^{W_n}$ and TV^{W_n} give energies of the state u_n of a graph.

Chapter 4: Sequential Limit: ϵ then n (i.e., $\mathrm{GL}_{\epsilon}^{W_n} \xrightarrow{\Gamma} \mathrm{TV}^{W_n} \xrightarrow{\Gamma} \mathrm{TV}^{W}$)

In this section, we prove the limits (1) and (2) from Figure 3.1. Limit (1) was already proven by van Gennip and Bertozzi [50, Theorem 3.1], and we state their result for completeness. Limit (2) was proven on square-lattice graphs in [50, Theorem 4.3] and for point clouds in [64]; we extend this result to general sequences of (weighted, undirected, and simple) graphs. Our proof closely follows the proof of the main theorem of [11].

4.1 Limit (1): $\mathrm{GL}_{\epsilon}^{W_n} \xrightarrow{\Gamma} \mathrm{TV}^{W_n}$

We state key relevant results from [50], which constitute our limit (1). Propositions 4.1.1 and 4.1.2 hold for all finite, undirected, and weighted graphs W_n.

Proposition 4.1.1 (Γ-convergence, [50, Theorem 3.1])**.** *The graph GL functional* (3.8) Γ*-converges to the graph TV functional* (3.10) *as* $\epsilon \to 0$ *with respect to* $u_n \to u$ *in* $\mathcal{V}^n$*. That is,*

$$\mathrm{GL}_{\epsilon}^{W_n} \xrightarrow{\Gamma} \mathrm{TV}^{W_n} \; . \tag{4.1}$$

In concert with Proposition 4.1.1, the following compactness property of the set $\mathcal{V}$ of functions guarantees that the minimizers of the Γ-converging functionals also converge.

Proposition 4.1.2 (Compactness, [50, Theorem 3.2])**.** *Let* $\{\epsilon_n\}_{n=1}^{\infty} \in \mathbb{R}_+$ *be a sequence such that* $\epsilon_n \to 0$ *as* $n \to \infty$*, and let* $\{u_n\}_{n=1}^{\infty} \subset \mathcal{V}$ *be a sequence for which there exists a constant* $C > 0$ *such that* $\mathrm{GL}_{\epsilon_n}^{W_n} < C$ *for all* $n \in \mathbb{N}$*. There then exists a subsequence* $\{u_{n'}\}_{n'=1}^{\infty} \subseteq \{u_n\}_{n=1}^{\infty}$ *and* $u_\infty \in \mathcal{V}_b^n$ *such that* $u_{n'} \to u_\infty$ *as* $n \to \infty$*.*

4.2 Limit (2): $\mathrm{TV}^{W_n} \xrightarrow{\Gamma} \mathrm{TV}^{W}$

We prove limit (2) using [11, Theorem 12], which states that $I_n \xrightarrow{\Gamma} I$ as $W_n \xrightarrow{\square} W$ where

$$I_n(\nu) = \int_0^1 \int_0^1 W_n(x,y) \int_{\mathbb{R}^2} f(\lambda,\mu)\, d\nu_x(\lambda)\, d\nu_y(\mu)\, dx\, dy\,, \tag{4.2}$$

$$I(\nu) = \int_0^1 \int_0^1 W(x,y) \int_{\mathbb{R}^2} f(\lambda,\mu)\, d\nu_x(\lambda)\, d\nu_y(\mu)\, dx\, dy\,. \tag{4.3}$$

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 is used in [11, Theorem 12].

Theorem 4.2.1. *[11, 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 sequence of dense graphs, with $W_n \xrightarrow{\square} W \in \mathcal{W}_0$. We then have that*

$$I_n \xrightarrow{\Gamma} I \text{ as } n \to \infty \tag{4.4}$$

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

Corollary 1. *With the choice $f(s,t) = |s-t|$, we have*

$$\mathrm{TV}^{W_n} \xrightarrow{\Gamma} \mathrm{TV}^{W} \text{ as } n \to \infty \tag{4.5}$$

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

Proposition 4.2.1 (Compactness)**.** *Let $W_n \xrightarrow{\square} W$, and let $u_n \in \mathcal{V}^n$ be a sequence of graph functions such that $\mathrm{TV}^{W_n}(u_n) < M$ for all n and some $M > 0$. Then, there exists a convergence subsequence u_{n_k} that the sequence of corresponding δ-Young measures $\{\delta_{u_{n_k}(x)}\}_{x\in(0,1)}$ converges to a limiting Young measure ν.*

Proof. Because $\{u_n\}$ consists of graph functions, it bounded in $L^1((0,1))$. Then we apply Lemma 3.4.2 (Prohorov's Theorem) to obtain the result. □

Chapter 5: Sequential Limit: n then ϵ (i.e., $\mathrm{GL}_\epsilon^{W_n} \xrightarrow{\Gamma} \mathrm{GL}_\epsilon^{W} \xrightarrow{\Gamma} \mathrm{TV}^{W}$)

In this section, we prove two novel limits. Our main result is limit (3), which extends [11, Theorem 12]. We prove limit (4) only for L^∞ graphons, and we discuss a scaling issue for L^p graphons.

5.1 Limit (3): $\mathrm{GL}_\epsilon^{W_n} \xrightarrow{\Gamma} \mathrm{GL}_\epsilon^{W}$

To prove 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 $\arg\min_{L^\infty((0,1))} \mathrm{GL}_\epsilon^W \subset L^\infty((0,1);[-1,1])$, so that it suffices to consider $u, u_n \in L^\infty((0,1);[-1,1])$.

We then show that the graph Dirichlet energy Γ-converges to the graphon Dirichlet energy. Our proof includes ideas from the proof of [11, Lemma 11], which shows the Γ-convergence of graph-cut functionals, which are similar to Dirichlet energies, except that their domains consist of only functions whose codomain has finite cardinality. Finally, we show that adding a double-well potential does not affect the Γ-convergence. Because each GL functional is a sum of a Dirichlet energy and a double-well potential, this fact yields Γ-convergence of the GL functionals.

5.1.1 Well-posedness of GL_ϵ^W in $L^\infty((0,1);[-1,1])$

We show by contradiction that u is bounded by $M = 1$. So let $M > 1$, and let u^M be the truncation of u at $\pm M$. That is,

$$u^M(x) = \begin{cases} M & \text{if } \ u(x) > M \\ u(x) & \text{if } \ |u(x)| \leq M \\ -M & \text{if } \ u(x) < -M\,. \end{cases} \tag{5.1}$$

We show that $\mathrm{GL}_\epsilon^W(u) \geq \mathrm{GL}_\epsilon^W(u^M)$ when W is any graphon. This implies that the minimizer of GL_ϵ^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. Because the double-well potential $\Phi(s) = (s^2-1)^2$ increases as $s > 1$ increases and decreases as $s < -1$ decreases, we know that

$$\int_0^1 \Phi(u) \geq \int_0^1 \Phi(u^M)\,. \tag{5.2}$$

For the Dirichlet energy, let

$$S_M = \{x : u(x) \geq M\} \tag{5.3}$$

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

$$\begin{aligned} g(x,y) &= |u(x) - u(y)|^2\,,\\ g^M(x,y) &= |u^M(x) - u^M(y)|^2\,. \end{aligned}$$

We want to show that

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

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

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

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 when $s > c$. 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

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

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

5.1.2 The proof of limit (3)

The following lemma appears similar to the result [11, Theorem 12], but it is different in two ways. First, it extends the domain from u_n with finite codomain to $u_n \in L^\infty((0,1))$, and it extends graphons from $W \in L^\infty((0,1)^2)$ to $W \in L^p((0,1)^2)$. Furthermore, Lemma 5.1.1 establishes pointwise convergence rather than Γ-convergence. It provides the foundation for proving the Γ-convergence of the graph Dirichlet energy to the graphon Dirichlet energy, which is stated in Theorem 5.1.1, which in turn is used to prove limit (3) which is stated in Corollary 3.

Lemma 5.1.1. *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 $I_n : \mathcal{Y}((0,1),\mathbb{R}) \to [0,\infty)$ 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 $u_n \in \mathcal{V}^n$ be a sequence of graph functions such that $\sup_n \|u_n\|_\infty < \infty$. We then have that the sequence of corresponding Young measures $\{\nu_x^n = \delta_{u_n(x)}\}_{x\in(0,1)} \subset \mathcal{Y}((0,1),\mathbb{R})$ 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* ν *satisfies*

$$I_n(\nu^{n_k}) \to I(\nu) \tag{5.4}$$

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

$$I(\nu) = \int_{(0,1)^2} W(x,y) \int_{\mathbb{R}^2} f(\lambda,\mu)\, d\nu_x(\lambda)\, d\nu_y(\mu)\, dx\, dy\,.$$

Proof. To prove this result, we use the triangle inequality to break $|I_n(\nu^n) - I(\nu)|$ into two parts. One of the parts converges to 0 due to 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 Lemma 3.4.2, when $\{u_n\}_{n=1}^{\infty}$ is a bounded sequence in $L^\infty((0,1),\mathbb{R})$, we know that a subsequence ν^{n_k} of $\nu^n = \{\nu_x^n = \delta_{u_n(x)}\}$ converges narrowly to a limit point ν. Henceforth, we simplify our notation by using $\{\nu^n\}$ to denote the subsequence $\{\nu^{n_k}\}$. We denote the innermost integrals of the functionals by

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

and we can then 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

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

We show that $(II) \to 0$ using weak convergence of g_n to g. The narrow convergence $\nu^n \xrightarrow{\star} \nu$

implies that $g_n \rightharpoonup^* g$ in $L^\infty((0,1)^2)$ by 3.4.1. That is,

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

Furthermore, $W \in L^p((0,1)^2) \subset L^1((0,1)^2)$, so the definition of weak-L^∞ convergence implies that $(II) \to 0$.

We now show that $(I) \to 0$ because $W_n \overset{\square}{\to} W$. 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 (2.9).

Let $P_k : \mathbb{R}^2 \to \mathbb{R}$ be a sequence of polynomial functions such that

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

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

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

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 bound $\|u_n\|_\infty \le 1$ (which we guaranteed in Section 5.1.1) to obtain

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

In summary,

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

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

Corollary 2. *Let $W_n \overset{\square}{\to} W$, and let $u_n \in \mathcal{V}^n$ be a sequence of graph functions such that* $\sup_n \|u_n\|_\infty < \infty$. *We then have that the sequence of corresponding Young measures $\{\nu_x^n\} \subset \mathcal{Y}((0,1),\mathbb{R})$ is pre-compact in the narrow topology. Moreover, any subsequence $\{\nu_x^n\}$ and any limit point ν satisfies*

$$D^{W_n}(\nu^n) \to D^W(\nu) \tag{5.5}$$

pointwise, where D^{W_n} and D^W are defined in (3.12) *and* (3.20)*, respectively.*

Proof. Choose $f(s,t) = |s - t|^2$ in Theorem 5.1.1. □

Corollary 2 extends [11, Lemma 11] by allowing ν to be any Young measure and allowing $\nu^n = \{\delta_{u_n(x)}\}_{x\in(0,1)}$ to be any δ-Young measure corresponding to some $u_n \in L^\infty((0,1))$. In [11, Lemma 11], ν_n must have support on a finite set of values, with the number of values being 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 ν^n to be any δ-Young measure. We will see that the closure of the set of δ-Young measures in $\mathcal{Y}((0,1),\mathbb{R})$ is the set $\mathcal{Y}((0,1),\mathbb{R})$.

Theorem 5.1.1. *Under the same assumptions as in Corollary 2, we have*

$$D^{W_n} \xrightarrow{\Gamma} D^W , \tag{5.6}$$

where we take Γ*-convergence with respect to narrow convergence of Young measures and cut-norm convergence of graphons.*

Proof. To prove Γ-convergence, it suffices to prove the following two statements.

[(i)]For every $\{\nu^n\}_{n\in\mathbb{N}}$ such that $\nu^n \to \nu$ narrowly, we have $D(\nu) \leq \liminf_n D^{W_n}(\nu^n)$. There exists a sequence $\{\nu^n\}_{n\in\mathbb{N}}$ that converges narrowly to ν with
$D(\nu) \geq \limsup_n D^{W_n}(\nu^n)$.

Statement (i) follows from Corollary 2 because pointwise convergence holds for all $\nu^n \to \nu$ narrowly. To prove statement (ii), it suffices to show that there exists a sequence $\nu^n \in \mathcal{X}_n$ converging narrowly to some ν for any $\nu \in \mathcal{Y}((0,1))$. Once we obtain this sequence, Corollary 2 yields inequality (ii). The existence of such a sequence follows by the denseness of Young measures corresponding to measurable functions in the set of all Young measures [76, Proposition 8]. This proof of statement (ii) is more direct than the analogous proof of [11, Theorem 12], in which ν^n is constructed carefully from a sequence of finite-valued functions u_n. □

Proposition 5.1.1 shows that the sequence of double-well potentials on graph functions converges to the double-well potential on a Young measure. In Corollary 3, we use this result in concert with Theorem 5.1.1 to show that $\mathrm{GL}_\epsilon^{W_n} \xrightarrow{\Gamma} \mathrm{GL}_\epsilon^W$.

Proposition 5.1.1. *Let* $u_n \in L^\infty((0,1))$ *be a sequence of functions that are constant on the intervals* $I_i = \big((i-1)/n,\ i/n\big]$*, and let* $\Phi : \mathbb{R} \to \mathbb{R}$ *be continuous and bounded. We have that*

$$\int_0^1 \Phi(u_n(x))\,dx \to \int_0^1 \int_{\mathbb{R}} \Phi(\lambda)\,d\nu(\lambda)\,dx \tag{5.7}$$

pointwise, where ν is the Young measure that is the narrow limit of the sequence of Young measures ν^n, which have the form $\nu_x^n = \delta_{u_n(x)}$.

1. *Proof.* The limit ν exists by Lemma 3.4.2 (Prohorov's Theorem). By the definition of narrow convergence, the fact that $\nu^n \overset{\star}{\rightharpoonup} \nu$ implies that

$$\Phi(u_n(x)) = \int_{\mathbb{R}} \Phi(\lambda)\, d\nu_x^n(\lambda) \overset{\star}{\rightharpoonup} \int_{\mathbb{R}} \Phi(\lambda)\, d\nu_x(\lambda) \text{ in } L^\infty((0,1))\,. \tag{5.8}$$

The definition of weak-$*$ convergence implies that, for any test function $\psi(x) \in L^1((0,1))$, 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 $n \to \infty$. Using the test function $\psi \equiv 1$ yields the desired result.

□

The Γ-convergence of the graph GL functional (3.8) to the graphon limit as $n \to \infty$ follows from the Γ-convergence of the graph Dirichlet energy (3.12) (see Corollary 2) and the pointwise convergence of the double-well potential (see Proposition 5.1.1). We state this convergence result formally in the next corollary.

Corollary 3. *Under the same assumptions as in Corollary 2, we have*

$$\mathrm{GL}_\epsilon^{W_n} \xrightarrow{\Gamma} \mathrm{GL}_\epsilon^{W} \tag{5.9}$$

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

Proof. Let $\mathrm{GL}_\epsilon^{W_n} = F_n + G_n$ and $\mathrm{GL}_\epsilon^{W} = F + G$, 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 Γ-convergence (see Corollary 2) of the graph Dirichlet energy (3.12) (that is, $F_n \xrightarrow{\Gamma} F$) and the fact that $G_n \to G$ (i.e., the pointwise convergence of the double-well potential that we proved in Proposition 5.1.1), we invoke the fact that Γ-convergence still holds under a continuous perturbation (see [55, Remark 2.2]) in the topology of narrow convergence of Young measures to obtain $F_n + G_n \xrightarrow{\Gamma} F + G$. □

5.2 Limit (4): $\mathrm{GL}_\epsilon^W \xrightarrow{\Gamma} TV^W$ as $\epsilon \to 0$ for $W \in L^\infty$.

5.2.1 The issue of ϵ-scaling

A key property of the classical GL functional (3.1) is its Γ-convergence to the TV functional in the $\epsilon \to 0$ limit [60]. Consider

$$\mathrm{GL}_\epsilon(u) = \epsilon \int_0^1 |\nabla u|^2 \, dx + \frac{1}{\epsilon} \int_0^1 \Phi(u(x)) \, dx \,, \tag{5.10}$$

which is equation (3.1) (but we show it again for convenience). As ϵ 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 ϵ shrinks to zero, the contribution from each jump grows to ∞. The prefactors that ensure that both the Dirichlet energy and double-well potential remain $O(1)$ as $\epsilon \to 0$ are ϵ and $1/\epsilon$, respectively. One can see this by substituting $x \mapsto x/\epsilon$ into the classical GL functional (3.1). Compare (5.10) to the graph GL functional

$$\mathrm{GL}_\epsilon^{W_n} = \int_0^1 \int_0^1 W_n(x, y) |u(x) - u(y)|^2 \, dx \, dy + \frac{1}{\epsilon} \int_0^1 \Phi(u(x)) \, dx \,, \tag{5.11}$$

which is equation (3.8). The graph GL functional (5.11) does not require the prefactor ϵ 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 $\epsilon \to 0$. As in the case of 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 GL_ϵ^W, which is similar to (5.11) 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 ϵ prefactor in L^∞ graphons.

If the graphon Dirichlet energy (3.20) is unbounded, as is the case for L^p graphons, the scaling in ϵ depends on the choice of W. The reason for this is that it is necessary to scale down 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 the Dirichlet energy needs to decay, and vice versa. Formalizing the relationship between the singularity of W and the ϵ-scaling rate is out of the scope of this work.

We prove a version of the classical limit $\mathrm{GL}_\epsilon \xrightarrow{\Gamma} \mathrm{TV}$ for graphon GL and graphon TV functionals. We consider only L^∞ graphons because it is difficult to determine the correct ϵ-scaling for general L^p graphons.

Theorem 5.2.1. *Suppose that $W \in L^\infty((0,1)^2)$. Then, as $\epsilon \to \infty$, we have that*

$$\mathrm{GL}_\epsilon^W \xrightarrow{\Gamma} \mathrm{TV}^W \tag{5.12}$$

with respect to narrow convergence of Young measures.

To prove Theorem 5.2.1, we follow a strategy that resembles the proof of [50, Theorem 3.1]. Suppose that a sequence of Young measures $\nu^n \in \mathcal{Y}((0,1),\mathbb{R})$ converges narrowly to $\nu \in \mathcal{Y}((0,1),\mathbb{R})$. Lemma 5.2.1 shows that the double-well potential Γ-converges either to 0 or to ∞ as $\nu^n \to \nu$ narrowly.

Lemma 5.2.1. *Let $\Phi : \mathbb{R} \to \mathbb{R}$ be a continuous and bounded potential function. Consider the functionals E_ϵ and E_0 on $\mathcal{Y}([0,1],\mathbb{R})$ defined as*

$$E_\epsilon(\nu) = \begin{cases} \dfrac{1}{\epsilon}\displaystyle\int_0^1 \int_{\mathbb{R}} \Phi(\lambda)\, d\nu_x(\lambda)\, dx & \textit{if } \nu \in \mathcal{Y}([0,1],\{-1,1\}), \\ +\infty & \textit{otherwise}, \end{cases} \tag{5.13}$$

and

$$E_0(\nu) = \begin{cases} 0 & \textit{if } \nu \in \mathcal{Y}([0,1],\{-1,1\}), \\ +\infty & \textit{otherwise}, \end{cases} \tag{5.14}$$

where $\mathcal{Y}([0,1],\{-1,1\}) \subset \mathcal{Y}([0,1],\mathbb{R})$ is the set of Young measures with support on $\{-1,1\}$. As $\epsilon \to 0$, we have

$$E_\epsilon \xrightarrow{\Gamma} E_0 \tag{5.15}$$

with respect to narrow convergence of Young measures.

Proof of Lemma 5.2.1. Consider sequences $\{\epsilon_n\}_{n\in\mathbb{N}}$ and $\{\nu^n\}_{n\in\mathbb{N}}$ such that $\epsilon_n \to 0$ and $\nu^n \to \nu$ narrowly for some $\nu \in \mathcal{Y}([0,1],\infty)$. We verify the liminf inequality (i) and the limsup inequality (ii) of Definition 3.3.1.

When $\nu \in \mathcal{Y}([0,1],\{-1,1\})$, it follows that $E_0(\nu) = 0$. The liminf inequality $E_0(\nu) \leq \liminf_{n\to\infty} E_{\epsilon_n}(\nu^n)$ holds because E_{ϵ_n} is nonnegative. When $\nu \in \mathcal{Y}([0,1],\mathbb{R}) \setminus \mathcal{Y}([0,1],\{-1,1\})$, it follows that $\nu^n \notin \mathcal{Y}([0,1],\{-1,1\})$ for sufficiently large n. Therefore, the slice ν^n_x has support in $\mathbb{R} \setminus \{-1,1\}$ for some x, so $\int_{\mathbb{R}} \Phi(\lambda)\, d\nu^n_x(\lambda) > 0$ for that x. Consequently, the liminf inequality holds:

$$\liminf_{n\to\infty} E_{\epsilon_n}(\nu^n) \geq \int_0^1 \liminf_{n\to\infty} \frac{1}{\epsilon_n} \int_{\mathbb{R}} \Phi(\lambda)\, d\nu^n_x(\lambda) dx = \infty = E_0(\nu)\,. \tag{5.16}$$

The first inequality holds because of Fatou's lemma. The first equality holds because $\epsilon_n \to 0$ and the integrand is uniformly bounded away from zero. The second equality holds because $\nu^n \notin \mathcal{Y}([0,1],\{-1,1\})$. We now show the limsup inequality $\limsup_{n\to\infty} E_{\epsilon_n}(\nu^n) \leq E_0(\nu)$. Suppose that $\nu \in \mathcal{Y}([0,1],\{-1,1\})$. We choose $\nu^n = \nu$ for all n, so $E_0(\nu) = 0 = \limsup_{n\to\infty} E_{\epsilon_n}(\nu^n)$. If $\nu \in \mathcal{Y}([0,1],\mathbb{R}) \setminus \mathcal{Y}([0,1],\{-1,1\})$, then $E_0(\nu) = \infty$ and the limsup inequality holds for any sequence ν^n that converges narrowly to ν. □

We now prove Theorem 5.2.1.

Proof of Theorem 5.2.1. To prove the desired result, we use the fact that Γ-convergence is maintained under a continuous perturbation. Note that $\mathrm{GL}^W_\epsilon(\nu) = D^W(\nu) + E_\epsilon(\nu)$, where D^W is the graphon Dirichlet energy (3.20) and E_ϵ is the double-well potential (5.13). The key ingredients in our proof are the facts that the double-well potential Γ-converges (see Lemma 5.2.1) and that the Dirichlet energy (3.20) is continuous in ν. This energy is continuous because ν^n converges narrowly to ν, which implies (by Definition (3.4.4)) for $W \in L^1$ that

$$\int_0^1 \int_0^1 W(x,y) \int_{\mathbb{R}^2} |\lambda - \mu|^2 \, d\nu_x^n(\lambda) \, d\nu_y^n(\mu) \, dx \, dy$$
$$\to \int_0^1 \int_0^1 W(x,y) \int_{\mathbb{R}^2} |\lambda - \mu|^2 \, d\nu_x(\lambda) \, d\nu_y(\mu) \, dx \, dy \, .$$

Therefore,

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

is a continuous perturbation of the double-well potential. Because Γ-convergence is stable under continuous perturbations, equation (5.15) implies that

$$\mathrm{GL}_\epsilon^W \overset{\Gamma}{\to} D^W + \begin{cases} 0 & \text{if } \nu \in \mathcal{Y}^b \\ +\infty & \text{otherwise} \end{cases} \tag{5.17}$$

as $\epsilon \to 0$.

The Γ-limit in (5.17) is equal to $\mathrm{TV}^W(\nu)$. This is true for the following reason. If $\nu \in \mathcal{Y} \setminus \mathcal{Y}^b$, then both the Γ-limit in (5.17) and $\mathrm{TV}^W(\nu)$ are ∞. If $\nu \in \mathcal{Y}^b$, then λ and μ can only take the values ± 1. Therefore, λ and μ can only have support in $\{\pm 1\}$, so $|\lambda - \mu|^2$ equals either $2|\lambda - \mu|$ or 0. □

Chapter 6: The GL minimizer for several examples

It is informative to compute 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 $\epsilon \to 0$ 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 [77, 78] 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 (3.19) 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 functional (3.8).

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 volume constraint, where "volume" refers to the sum of values of a function (or a Young measure) on a graph. For the graphon GL-minimization problem, we thus impose the volume constraint

$$\int_0^1 \int_{\mathbb{R}} \lambda \, d\nu_x(\lambda) \, dx = c \tag{6.1}$$

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

$$\frac{1}{n} \sum_{i=1}^{n} u_i = c \, . \tag{6.2}$$

This constraint entails that the Young measure ν (or the graph function u) has a mean value of c, so ν cannot have all of its mass on either -1 or $+1$ alone. This yields *phase separation*, with the values -1 and 1 corresponding to two different *phases* [59]. We assume that $c \in (-1, 1)$ in Section 6.1 and that $c = 0$ in Section 6.2.

6.1 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 $W \equiv p \in (0, 1)$)[8, Section 10.1], and some growing preferential-attachment graphs (including Barabási–Albert (BA) graphs, for which $W \equiv p \in (0, 1)$) [8, Example 11.44 and Proposition 11.45].

Consider the graph GL functional on the n-node complete graph. Let $A_{ij}^{(n)} = p$ for all i and j in (3.8). This leads to

$$\begin{aligned}
\mathrm{GL}_\epsilon^{W_n}(u) &= \frac{1}{n^2}\sum_{i,j=1}^n p(u_i - u_j)^2 + \frac{1}{\epsilon n}\sum_{i=1}^n \Phi(u_i) \\
&= \frac{p}{n^2}\left(2n\sum_{i=1}^n u_i^2 - 2\sum_{i,j=1}^n u_i u_j\right) + \frac{1}{\epsilon n}\sum_{i=1}^n (u_i^4 - 2u_i^2 + 1) \\
&= \frac{1}{\epsilon n}\sum_{i=1}^n \left(u_i^4 - (2 - 2\epsilon p)u_i^2\right) - \frac{2p}{n^2}\left(\sum_{i=1}^n u_i\right)^2 + \frac{1}{\epsilon} \\
&= \frac{1}{\epsilon n}\sum_{i=1}^n (u_i^2 - (1 - \epsilon p))^2 - \frac{2}{n^2}\left(\sum_{i=1}^n u_i\right)^2 - \frac{1}{\epsilon}(1 - \epsilon p)^2 + \frac{1}{\epsilon}\,.
\end{aligned} \tag{6.3}$$

Fix $\epsilon < 1$. Minimizing (6.3) subject to the volume constraint (6.2) is equivalent to minimizing the energy

$$E_n(v) := \sum_{i=1}^n \left(v_i^2 - 1\right)^2 \,, \tag{6.4}$$

where

$$v_i = \frac{u_i}{\sqrt{1 - \epsilon p}}\,. \tag{6.5}$$

With the change of variables (6.5), the volume constraint (6.2) becomes

$$\frac{1}{n}\sum_{i=1}^{n} v_i = \frac{c}{\sqrt{1-\epsilon p}} \,. \tag{6.6}$$

Minimizing (6.4) subject to the volume constraint (6.6) leads to the Euler–Lagrange (EL) equations

$$v_i^3 - v_i = \tau \quad \text{and} \quad \frac{1}{n}\sum_{i=1}^{n} v_i = \frac{c}{\sqrt{1-\epsilon p}} \,, \tag{6.7}$$

where τ is the Lagrange multiplier associated with the volume constraint.

The derivation of the EL equations(6.7) is as follows. For all suitable perturbations ϕ, we have

$$\begin{aligned}
\frac{d}{dt}\left[E_n(v+t\phi) + \frac{\tau}{n}\sum_{i=1}^{n}(v_i+t\phi)\right]_{t=0} &= \frac{d}{dt}\left[\sum_{i=1}^{n}((v_i+t\phi_i)^2-1)^2 + \frac{d}{dt}\frac{\tau}{n}\sum_{i=1}^{n}(v_i+t\phi)\right]_{t=0} \\
&= \sum_{i=1}^{n} 2((v_i+t\phi_i)^2-1)(2v_i+t\phi_i)\phi_i|_{t=0} + \frac{\tau}{n}\sum_{i=1}^{n}\phi_i \\
&= \sum_{i=1}^{n}\left[4(v_i^3-v_i)\phi_i + \frac{\tau}{n}\phi_i\right] \\
&= 0
\end{aligned}$$

for all ϕ_i implies that for some τ', we have

$$v_i^3 - v_i = \tau' \,.$$

Proposition 6.1.1 (Characterization of GL minimizers for the finite complete graph with constant edge weight)**.** *Let $W_n \equiv p$, and ϵ small enough, let $|c| < \sqrt{1-\epsilon p}$. The minimizers of* $\mathrm{GL}_\epsilon^{W_n}$ *under the volume constraint* (6.6) *are functions u whose range is contained in $\{z_{\pm,n}\}$, 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 $O(1/\sqrt{n})$.*

Remark 6.1.1. *In the $n \to \infty$ limit, the sequence of minimizers approaches* ± 1 *almost everywhere.*

Remark 6.1.2. *The minimizers are not unique, since the energy and the constraints are invariant* *under permutations of the subscripts.*

Proof. With the change of variables (6.5), it suffices to describes minimizers v of the equivalent minimization problem (6.4).

First, we show that the energy functional (6.4) for a minimizer v is bounded above by 1. Consider a candidate minimizer $\tilde{v}$ with $\tilde{v}_i \in \{\pm 1\}$ for all $i \in \{1, \ldots, n\}$ 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; because the two numbers sum to n, one rounds up and one rounds down. (It does not matter which one rounds up and which one rounds down.) There may then 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 volume constraint. 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. Due to equation (6.4), the energy bound $E_n(v) \leq 1$ implies that there must be some 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 that $|\tau| = O(1/\sqrt{n})$. Note that 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 uniformly bounded away from 0, with respect to n. This implies that, whenever $|g(v_i)| = |v_i^2 - 1||v_i| = |\tau| \leq (1/\sqrt{n})(1 + 1/\sqrt{n})^{1/2}$, we have from the inverse function theorem 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 $O(1/\sqrt{n})$, whereas 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 all be a root of $g(x) = \tau$ and satisfies $|v_i^2 - 1| = O(1/\sqrt{n})$. Thus, $v_i \to \pm 1$ as $n \to \infty$.

□

The minimizer v of (6.4) that we constructed has a discontinuous profile; that is, it has jumps between the values $+\sqrt{1 - \epsilon p}$ and $-\sqrt{1 - \epsilon p}$. See the related discussion in [79]. Furthermore, as $n \to \infty$, the limit of the sequence of minimizers v takes the values ± 1 almost everywhere as $n \to \infty$, so (by the change of variables (6.5)) 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 $\mathrm{GL}_\epsilon^{W_n} \xrightarrow{\Gamma} \mathrm{GL}_\epsilon^{W}$ as $W_n \xrightarrow{\square} W$, we know (see Section 5.1) that characterizing the GL minimizers for the constant graph gives us insight into GL minimizers for the constant graphon. We show that (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 helps explain why graphon functionals act on Young measures.

The Young measure $\{\nu_x\}_{x\in(0,1)}$ that is defined by

$$\nu_x = \theta\,\delta_{\sqrt{1-\epsilon p}} + (1-\theta)\delta_{-\sqrt{1-\epsilon p}}\,, \tag{6.8}$$

where $2\theta - 1 = cn$, gives a set of minimizers of the GL functional for the constant graphon with the volume constraint (6.1). The quantities θ and $1-\theta$ are the proportions of the values $+\sqrt{1-\epsilon p}$ and $-\sqrt{1-\epsilon p}$, respectively.

The Young measure (6.8) 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 θ of the points in $[0,1]$ and the value $-\sqrt{1-\epsilon p}$ on a proportion $1-\theta$ of the points. We consider two different Young measures with the volume constraint (6.1) but different amounts of oscillation.

We first consider the Young measure

$$\nu_x^{(1)} = \delta_{u(x)}\,, \ \text{ where } \ u(x) = \begin{cases} -\sqrt{1-\epsilon p}\,, & 0 < x < \frac{1-c}{2} \\ \sqrt{1-\epsilon p}\,, & \frac{1-c}{2} < x < 1\,, \end{cases} \tag{6.9}$$

which represents a non-oscillatory limit. This function $u(x)$ has a single transition that satisfies the volume constraint (6.1). The graphon GL functional (3.19) evaluated on $\nu^{(1)}$ is

$$\mathrm{GL}_\epsilon^W(\nu^{(1)}) = 4(1-\epsilon p)\int_0^{\frac{1-c}{2}}\int_{\frac{1-c}{2}}^1 W(x,y)\,dx\,dy + 4(1-\epsilon p)\int_{\frac{1-c}{2}}^1\int_0^{\frac{1-c}{2}} W(x,y)\,dx\,dy\,. \quad (6.10)$$

We now consider a Young measure $\nu_x^{(2)}$ that satisfies the volume constraint (6.1) but does not correspond to a function u. This Young measure is

$$\nu_x^{(2)} = \frac{1+c}{2}\delta_{\sqrt{1-\epsilon p}} + \frac{1-c}{2}\delta_{-\sqrt{1-\epsilon p}}\,. \quad (6.11)$$

One can think of $\nu_x^{(2)}$ as the limit of increasingly oscillatory functions that satisfies the constraint (6.1). The graphon GL functional (3.19) evaluated on $\nu^{(2)}$ is

$$\mathrm{GL}_\epsilon^W(\nu^{(2)}) = 2\sqrt{1-\epsilon p}(1-c)(1+c)\int_0^1\int_0^1 W(x,y)\,dx\,dy\,. \quad (6.12)$$

We compare the values of $\mathrm{GL}_\epsilon^W(\nu^{(1)})$ and $\mathrm{GL}_\epsilon^W(\nu^{(2)})$ to illustrate the effect of oscillations on the functional value. Suppose that $W \equiv 1 \in (0,1)$ and that the volume constraint (6.1) holds. Equations (6.10) and (6.12) yield

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

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 [59].

The contribution to the constant-graphon GL functional GL_ϵ^W is the same whenever ν_x and ν_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 $\epsilon \to 0$ limit, the Young-measure minimizers (6.8) converge to the limiting Young measure $\nu = \{\nu_x\}$, which is defined by

$$\nu_x = \theta\delta_1 + (1-\theta)\delta_{-1} \tag{6.13}$$

for all $x \in [0,1]$. From the Γ-convergence $\mathrm{GL}_\epsilon^W \xrightarrow{\Gamma} \mathrm{TV}^W$ (see Section 5.2), the Young-measure limit (6.13) of the minimizers (6.8) of GL_ϵ^W is a minimizer of the limiting energy TV^W for $W \equiv p$.

To help with later discussions, also study the (unrealistic) case of "oversaturation", which occurs when the volume constraint imposes that the mean value of v is larger than 1 or is 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 $\sqrt{1-\epsilon p}$ under this condition.

Lemma 6.1.1 (Non-oversaturation of GL minimizers)**.** *When* $|\gamma| := \left|\frac{c}{\sqrt{1-\epsilon p}}\right| \geq 1$ *in the volume constraint* (6.6)*, the minimizer of* (6.4) *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* $\mathrm{GL}_\epsilon^{W_n}(u)$ *with the volume constraint* (6.2) *is* u^* *(which satisfies* $u_i^* = c$ *for all* $i \in [n]$*), and the minimum energy is* $\mathrm{GL}_\epsilon^{W_n}(u^*) = \frac{1}{\epsilon}\Phi(c) = \frac{1}{\epsilon}(c^2-1)^2$*. Additionally, the energies* $E_n(v^*)$ *and* $\mathrm{GL}_\epsilon^{W_n}(u^*)$ *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 volume constraint (6.2) as

$$\sum_{i=1}^{n} v_i = \gamma n\,. \tag{6.14}$$

First, we show that the E_n minimizer 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

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

The function $H_n(v) \leq E_n(v)$ for any v. Furthermore, 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. Consequently, with the volume constraint (6.14), 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 minimizer of E_n entails that $v_i = \gamma$, which yields

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

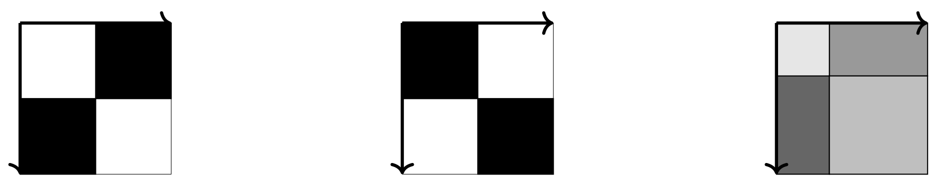
The conclusion about $\mathrm{GL}_\epsilon^{W_n}(u)$ then follows immediately. □

6.2 2×2 stochastic block models (SBMs)

Consider the 2×2 piecewise-constant graphon

$$W(x,y) = \begin{cases} a_{11} & \text{if } (x,y) \in (0,\frac{1}{2}) \times (0,\frac{1}{2}) \\ a_{12} & \text{if } (x,y) \in (0,\frac{1}{2}) \times [\frac{1}{2},1) \\ a_{21} & \text{if } (x,y) \in [\frac{1}{2},1) \times (0,\frac{1}{2}) \\ a_{22} & \text{if } (x,y) \in [\frac{1}{2},1) \times [\frac{1}{2},1)\,, \end{cases} \tag{6.15}$$

where $a_{ij} \in [0,1]$. This graphon is a stochastic block model (SBM) [18]. Researchers use SBMs as generative models of graphs with various types of mesoscale network structures, such as assortative or disassortative block structures [80]. SBMs constitute a relatively general class of graphs and graphons. Indeed, the Szemeredi Lemma implies that one can approximate any L^∞ graphon arbitrarily closely in cut-norm by an SBM [43, Lemma 3.1].



Bipartite graph   Community structure graph   A 2×2 SBM

Figure 6.1: Three types of 2×2 piecewise-constant graphons.

In this subsection, we discuss the 2×2 piecewise-constant graphons in Figure 6.1. If one views the SBM as a generative model, the entries a_{ij} encode the probability that there is an edge between a node in community $i \in \{1,2\}$ and a node in community $j \in \{1,2\}$. However, we view the entries a_{ij} as fixed values that form a block-diagonal adjacency matrix in which adjacency-matrix entries take values W as described in (6.15).

In an assortative SBM, the intra-community edge probabilities a_{ii} and a_{jj} are larger than the inter-community edge probabilities a_{ij} and a_{ji}. Figure 6.1 shows examples of graphons where black represents the value one, white represents the value zero, and grayscale represents values in $(0,1)$. The origin is at the upper-left corner of each graphon. An SBM with planted community structure (see Figure 6.1(b)) is an extreme case of assortative block structure [80]. In a disassortative SBM, the edge probabilities a_{ij} and a_{ji} are larger than a_{ii} and a_{jj}, so inter-community edges are more likely than intra-community edges. An extreme example is a bipartite SBM (see Figure 6.1(a)). Finally, Figure 6.1(c) shows an example of a general 2×2 SBM with different community sizes.

When $a_{21} = a_{12} = 1$ and $a_{11} = a_{22} = 0$ in equation (6.15), the graphon W is a bipartite graphon. When $a_{21} = a_{12} = 0$ and $a_{11} = a_{22} = 1$, the graphon W is a community-structure graphon with two weakly-connected subgraphons.

Seeking a GL minimizer for SBMs is different from seeking a constant-graphon GL minimizer because no volume constraint is needed in the bipartite case. For the constant graphon, the minimizer would be trivial (all 1 or all -1) without a volume constraint, but the bipartite graphon has no such trivial minimizer. Instead, we will see that the community size α acts like a volume constraint.

6.2.1 Complete bipartite graphon

In a bipartite graph (see Figure 6.1(b)), 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

$$A_{ij}^{(n)} = \begin{cases} 1 & \text{if } \ i \in S\,,\ j \in S^c \\ 1 & \text{if } \ i \in S^c\,,\ j \in S \\ 0 & \text{if } \ i \in S\,,\ j \in S \\ 0 & \text{if } \ i \in S\,,\ j \in S^c\,. \end{cases} \tag{6.16}$$

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

$$W(x, y) = \begin{cases} 0 & \text{if } \ (x, y) \in (0, a) \times (0, a) \\ 1 & \text{if } \ (x, y) \in (0, a) \times [a, 1) \\ 1 & \text{if } \ (x, y) \in [a, 1) \times (0, a) \\ 0 & \text{if } \ (x, y) \in [a, 1) \times [a, 1)\,. \end{cases} \tag{6.17}$$

We rewrite the GL functional for the bipartite graph as

$$\begin{aligned} \mathrm{GL}_\epsilon^{W_n}(u) &= \frac{1}{n^2} \sum_{i \in S, j \in S^c} (u_i - u_j)^2 + \frac{1}{\epsilon n} \sum_{i=1}^{n} \Phi(u_i) \qquad (6.18) \\ &= \frac{1}{n^2} \sum_{i \in S, j \in S^c} u_i^2 + \frac{1}{n^2} \sum_{i \in S, j \in S^c} u_j^2 - \frac{2}{n^2} \sum_{i \in S, j \in S^c} u_j + \frac{1}{\epsilon n} \sum_{i=1}^{n} \Phi(u_i) \\ &= \frac{1}{n^2} \sum_{i \in S} |S^c| u_i^2 + \frac{1}{n^2} \sum_{j \in S^c} |S| u_j^2 - \frac{2}{n^2} \sum_{i \in S, j \in S^c} u_i u_j + \frac{1}{\epsilon n} \sum_{i=1}^{n} \Phi(u_i) \\ &= \frac{1}{\epsilon n} \sum_{i \in S} \left(u_i^4 - (2 - \epsilon(1-a)) u_i^2 + 1 \right) + \frac{1}{\epsilon n} \sum_{i \in S^c} \left(u_i^4 - (2 - \epsilon a) u_i^2 + 1 \right) - \frac{2}{n^2} \sum_{i \in S} u_i \sum_{j \in S^c} u_j\,. \end{aligned}$$

By completing the square in the first two sums, we see that (6.18) is equal (up to a constant) to

$$\frac{1}{\epsilon n}\sum_{i\in S}\left(u_i^2-\left(1-\frac{\epsilon(1-a)}{2}\right)\right)^2+\frac{1}{\epsilon n}\sum_{i\in S^c}\left(u_i^2-\left(1-\frac{\epsilon a}{2}\right)\right)^2-\frac{2}{n^2}\sum_{i\in S}u_i\sum_{j\in S^c}u_j\,. \tag{6.19}$$

Define the community size $\sum_{i\in S}u_i = \alpha$, which we combine with the volume constraint (6.2) to obtain

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

We introduce a rescaling of variables that is similar to (6.5). This rescaling is

$$v_i = \begin{cases} \frac{u_i}{\sqrt{1-\epsilon(1-a)/2}} & \text{for } i\in S \\ \frac{u_i}{\sqrt{1-\epsilon a/2}} & \text{for } i\in S^c\,. \end{cases} \tag{6.20}$$

To simplify the notation, we define

$$c_S := \sqrt{1-\epsilon(1-a)/2} \quad \text{and} \quad c_{S^c} := \sqrt{1-\epsilon a/2}\,. \tag{6.21}$$

This yields the equivalent minimization problem

$$E_n(v) := \frac{c_S^4}{\epsilon}\sum_{i\in S}(v_i^2-1)^2+\frac{c_{S^c}^4}{\epsilon}\sum_{i\in S^c}(v_i^2-1)^2-\frac{2}{n}c_Sc_{S^c}\alpha(c-\alpha)\,, \tag{6.22}$$

where, with the change of variables (6.20), the community size $\alpha = \sum_{i\in S}u_i$ gives

$$\alpha = c_S\sum_{i\in S}v_i\,, \quad c-\alpha = c_{S^c}\sum_{i\in S^c}v_i\,. \tag{6.23}$$

Treating v and α as unknowns, the resulting EL equations for minimizing E_n are

$$\begin{aligned}
\frac{4c_S^4}{\epsilon}(v_i^3 - v_i) - \tau_S c_S = 0 \quad & \text{on } S\,, \\
\frac{4c_{S^c}^4}{\epsilon}(v_i^3 - v_i) - \tau_{S^c} c_{S^c} = 0 \quad & \text{on } S^c\,, \\
-\frac{2}{n} c_S c_{S^c}\,\alpha + \tau_S - \tau_{S^c} = 0\,. &
\end{aligned} \tag{6.24}$$

Proposition 6.2.1 (Characterization of the GL minimizers for the bipartite graphon)**.** *Let W_n be the complete bipartite graph with n nodes (see* (6.16)*). The minimizers of* $\mathrm{GL}_\epsilon^{W_n}$ *are functions u that take the values $\{\pm\sqrt{1-\epsilon a/2} + O(1/\sqrt{n})\}$ except for at most a finite number N of nodes, where N 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 ± 1 except for at most N possible nodes, where N is independent of n.

First, we derive the uniform energy bound

$$E_n(\tilde{v}) \leq \frac{3}{\epsilon} \tag{6.25}$$

for a feasible $\tilde{v}$. When both $|S|$ and $|S^c|$ are even, $E_n(\tilde{v}) = 0$ if we take an equal number of 1 and -1 entries in the components of $\tilde{v}$ in S and S^c. This also implies that $\alpha = 0$.

If n is even but both $|S|$ and $|S^c|$ are odd, then we set $\tilde{v}_i = 0$ for one node in S and one node in S^c, and we also take an equal number of 1 and -1 entries in S and likewise in S^c. This again yields $\alpha = 0$. For both of these choices of $\tilde{v}$, we have

$$E_n(\tilde{v}) \leq \frac{c_S^4 + c_{S^c}^4}{\epsilon} = \frac{2}{\epsilon} - 1 + \frac{\epsilon}{4}(1 - 2a + 2a^2) \leq \frac{3}{\epsilon} \tag{6.26}$$

for $a \in (0,1)$ and $\epsilon < 1$. For the remaining case, in which n is odd, 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 $\frac{c_S^4}{\epsilon}$ (if $|S|$ is odd) or $\frac{c_{S^c}^4}{\epsilon}$ (if $|S^c|$ is odd), which again satisfies the uniform energy bound (6.25).

From the energy upper bound (6.25), a minimizer v satisfies

$$\frac{c_S^4}{\epsilon}\sum_{i\in S}(v_i^2-1)^2+\frac{c_{S^c}^4}{\epsilon}\sum_{i\in S^c}(v_i^2-1)^2 \le E_n(v) \le E_n(\tilde{v}) \le \frac{3}{\epsilon}\,. \tag{6.27}$$

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

$$(v_j^2-1)^2 \le \frac{3}{|S|c_S^4} \quad \text{and} \quad (v_{j'}^2-1)^2 \le \frac{3}{|S^c|c_S^4}\,. \tag{6.28}$$

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

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

where the constants m_1 and m_2 do not depend on n. Because the EL equations (6.24) hold for all nodes, including j and j', we then obtain

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

for some constant C that does not depend on n. With these bounds on τ_S and τ_{S^c}, equations (6.24) yield

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

for all nodes in both S and S^c. Similar to the proof of Proposition 6.1.1, for sufficiently large n, the zeros v_i lie within $O(k/\sqrt{n})$ of ±1 and 0. Furthermore, due to the energy bound (6.27), we see that v can be within $O(k/\sqrt{n})$ of 0 at most at a finite number of nodes; that number does not exceed a constant that is independent of n.

From the change of variables (6.20), 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. □

It follows from Proposition 6.2.1 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 $U_1 \subset S$ and $U_2 \subset S^c$ are any sets with the property that $|U_1| = \frac{|S|}{2}$ and $|U_2| = \frac{|S^c|}{2}$.

As $n \to \infty$, the energy bound (6.27) also yields a bound on the limiting graphon energy. Furthermore, this bound 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 $|S^c|$, and one can always choose measures that are binary, which do not contribute to the double-well energy.

6.2.2 Community-structure graphon

Community-structure graphs consist of densely-connected subgraphs (i.e., communities) that are sparsely connected to each other. As in a 2×2 bipartite SBM, a 2×2 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 subgraphs 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

$$A_{ij}^{(n)} = \begin{cases} 1 & \text{if } i\,,\ j \in S \\ 1 & \text{if } i\,,\ j \in S^c \\ 0 & \text{if } i \in S\,,\ j \in S^c \\ 0 & \text{if } i \in S^c\,,\ j \in S\,. \end{cases} \tag{6.29}$$

The corresponding complete community-structure graphon is

$$W(x,y) = \begin{cases} 1 & \text{if } (x,y) \in (0,a) \times (0,a) \\ 0 & \text{if } (x,y) \in (0,a) \times [a,1) \\ 0 & \text{if } (x,y) \in [a,1) \times (0,a) \\ 1 & \text{if } (x,y) \in [a,1) \times [a,1)\,. \end{cases} \tag{6.30}$$

The GL functional for a complete community-structure graph is

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

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

$$E_n(v) = \frac{c_S^4}{\epsilon} \sum_{i \in S} (v_i^2 - 1)^2 + \frac{c_{S^c}^4}{\epsilon} \sum_{i \in S^c} (v_i^2 - 1)^2 - \frac{2}{n} c_S c_{S^c} \alpha^2 - \frac{2}{n} c_S c_{S^c} (c - \alpha)^2\,, \tag{6.31}$$

where α and $c - \alpha$ satisfy the same volume constraints (6.23) as a complete bipartite graph, $c_S v_i = u_i$ for $i \in S$ and $c_{S^c} v_i = u_i$ for $i \in S^c$, and

$$c_S = \sqrt{1 - \epsilon a}\,, \quad c_{S^c} = \sqrt{1 - \epsilon(1 - a)}\,. \tag{6.32}$$

Let $c = 0$. Assuming the community size $\alpha = \sum_{i \in S} u_i$, we have

$$c_S \sum_{i \in S} v_i = \alpha\,, \quad c_{S^c} \sum_{i \in S^c} v_i = -\alpha\,. \tag{6.33}$$

When $|S| = |S^c|$, the energy functional (6.31) reduces to

$$E_n(v) = \frac{2c_S^4}{\epsilon} \sum_{i=1}^{n} (v_i^2 - 1)^2 - \frac{4}{n} c_S^2 \alpha^2\,. \tag{6.34}$$

The first term of the energy (6.34) encourages the values of v to be near ± 1. The second term of (6.34) encourages the $|\alpha|$ 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 tradeoff between the first and the second terms. The volume constraint (6.33) ensures that the sum of the values of v in S is the negative of the sum of the values of v in S^c. We show in Proposition 6.2.2 that the optimal balance in this tradeoff has v near ± 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 of a complete community-structure graph when $|S| = |S^c|$. Let $|S| = |S^c|$ (which implies that $c_S = c_{S^c}$), and let $\alpha := \gamma |S|$. The volume constraint (6.33) then entails that γ / 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 (6.33) as

$$\frac{1}{|S|} \sum_{i \in S} v_i = \frac{\gamma}{c_S}\,, \quad -\frac{1}{|S|} \sum_{i \in S^c} v_i = -\frac{\gamma}{c_S}\,. \tag{6.35}$$

Proposition 6.2.2 (Characterization of the GL minimizers for the complete community-structure graph)**.** *Let W_n be the complete community-structure graph, which has adjacency-matrix elements* (6.29)*, and suppose that $|S| = |S^c|$. The minimizers of $\mathrm{GL}_\epsilon^{W_n}$ 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 γ that minimize E_n are equal to $\pm c_S$, which approach* ± 1 *as $\epsilon \to 0$.*

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, $(h(x) + h(y))/2 \leq h\left((x+y)/2\right)$. 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 means. Consequently, the minimizer v' of H_n is constant on S and constant on S^c but with different values on S and S^c because of the volume constraint (6.33). The minimizer that satisfies the volume constraint (6.35) is the function v' with

$$v'_i = \begin{cases} +\frac{\gamma}{c_S} & \text{if } \; i \in S \\ -\frac{\gamma}{c_S} & \text{if } \; i \in S^c\,. \end{cases} \tag{6.36}$$

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

$$E_n(v') = \frac{nc_S^4}{\epsilon}\left(\frac{\gamma'}{c_S^2} - 1\right)^2 - \frac{4}{n}c_S^2\gamma'|S|^2 \tag{6.37}$$

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

$$\frac{d}{d\gamma'}g_1(\gamma') = \frac{2c_S^2 n}{\epsilon}\left(\frac{\gamma'}{c_S^2} - 1\right) - \frac{4}{n}c_S^2|S|^2\,. \tag{6.38}$$

Setting (6.38) to 0 yields the critical point

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

Therefore, when $|S| = |S^c|$, the critical values of γ 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 (6.36) of H_n is thus

$$v_i' = \begin{cases} +\sqrt{1+\epsilon/2} & \text{if } \ i \in S \\ -\sqrt{1+\epsilon/2} & \text{if } \ i \in S^c . \end{cases} \tag{6.39}$$

Using the change of variables (6.32), we obtain the GL minimizer u, which is defined by

$$u_i = \begin{cases} +\sqrt{1-(\epsilon/2)^2} & \text{if } \ i \in S \\ -\sqrt{1-(\epsilon/2)^2} & \text{if } \ i \in S^c . \end{cases} \tag{6.40}$$

Because $\epsilon \downarrow 0$, we have that $+\sqrt{1-(\epsilon/2)^2} \downarrow 1$ and $-\sqrt{1-(\epsilon/2)^2} \uparrow -1$. Therefore, even when the mean value of v is forced by the volume constraint (6.35) to have an absolute value larger than 1, the optimal u approaches ± 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

$$E_n(\tilde{v}) = -\frac{4}{n}c_S^2|S|^2 . \tag{6.41}$$

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

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

$$E_n(v) = \frac{2c_S^4}{\epsilon}\sum_{i=1}^{n}(v_i^2-1)^2 - \frac{4}{n}c_S^2\gamma^2|S|^2 . \tag{6.42}$$

The second term of (6.42) is larger than $E_n(\tilde{v})$ and the first term of (6.42) 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 γ. The GL minimizer u that corresponds to this candidate minimizer $\tilde{v}$ has components

$$u_i = \begin{cases} +\frac{1}{\sqrt{1-\epsilon/2}} & \text{if } \ i \in S \\ -\frac{1}{\sqrt{1-\epsilon/2}} & \text{if } \ i \in S^c \, . \end{cases} \tag{6.43}$$

As $\epsilon \downarrow 0$, the values of the minimizer u approach ± 1.

□

Chapter 7: Graphon reaction–diffusion equations

7.1 Introduction

Classical reaction–diffusion (RD) equations originated to describe the evolution of a concentration of a substance in chemical reactions in which particles diffuse (i.e., travel) and react (i.e., appear and disappear). They arise in fields such as biology, ecology, chemistry, and physics [34, 35, 36, 37]. RD equations are semi-linear parabolic partial differential equations (PDEs) of the form

$$u_t(x,t) = Lu(x,t) + \Phi(u(x,t)) \,, \tag{7.1}$$

where u is the density or concentration of a substance, L is a diffusion operator, and Φ is a function that incorporates the effect of local reactions.

RD equations that occur on graphs, which we call *graph RD equations*, are used for applications in which there is some heterogeneous connectivity between individuals. Such applications include population ecology on population networks and disease spread on social networks [81, 82, 83, 84].

For a graph with adjacency matrix $\{A_{ij}\}_{i,j=1}^{n}$, a graph RD equation is a system of equations

$$\frac{d}{dt}u_i(t) = \frac{1}{n}\sum_{j=1}^{n} A_{ij}\big(u_j(t) - u_i(t)\big) + \Phi\big(u_i(t)\big) \,, \quad i \in \{1, \ldots, n\} \,, \tag{7.2}$$

where $u_i(t)$ describes the time-varying concentration of a substance on node i of the graph. Hence a graph RD equation on a graph with n nodes is a system of n differential equations, and (7.2) is an analogue of Equation (7.1) in which, in addition to being discrete, has a nonlocal diffusion term that is weighted by the graph's edge-weights $\{A_{ij}\}$.

We consider two types of limits relating to graph RD equations: first, the large-graph limit of graph RD equations for a sequence of growing graphs, and second, the hydrodynamic limit of certain interacting particle systems on graphs. Large-graph limits concern *graphons*, which are a notion of limits of sequences of graphs as the size of the graphs approaches infinity. The hydrodynamic limit of interacting particle systems describes the density of each of the interacting particles as the number of particles approaches infinity; we consider a class of continuous-time interacting particle systems on graphs that involve both a random walk and a birth-death process. The hydrodynamic limit of the random walk gives the diffusion part of the RD equation, whereas the limit of the birth-death process gives the reaction part of the RD equation.

7.1.1 Relationship to GL equation

The EL equation of a variational minimization problem results from equating the first variation $\frac{d}{dt}\left[\mathrm{GL}_\epsilon^W(u+t\varphi)\right]|_{t=0}$ to 0. This amounts to finding stationary points of the GL functional; solving the EL equation is a necessary, but not sufficient, condition that a value is in the range of u^*.

Simplifying $\frac{d}{dt}\left[\mathrm{GL}_\epsilon^W(u+t\varphi)\right]|_{t=0} = 0$ yields

$$4\int_0^1 W(x,y)(u(x)-u(y))dy + \frac{1}{\epsilon}\Phi'(u(x)) = 0\,.$$

7.1.2 Related work

The graph diffusion equation is the system of differential equations (7.2) in the case where $f \equiv 0$. The graphon diffusion equation is a PDE of the form (7.1) in the case where $f \equiv 0$ and L is a nonlocal operator weighted by a kernel W (the graph and graphon diffusion equations will be defined in greater detail in Section 7.2). In Figure 7.1, we denote the graph diffusion equation corresponding to a graph W_n as $\mathrm{Diffusion}^{W_n}$, and the graphon diffusion equation corresponding to a graphon W as $\mathrm{Diffusion}^W$.

Medvedev [48] showed that if the sequence of graphs W_n converges in L^2 norm to W, then the sequence of solutions of $\mathrm{Diffusion}^{W_n}$ converges in L^2 norm to the solution of $\mathrm{Diffusion}^W$. We extend this result to include the case where W_n converges in cut norm to W. The cut norm is a weaker form of convergence than is L^2 convergence, and it is defined in greater detail in Section 2.1.

Angstmann et al [85] showed that Diffusion^{W_n} is the master equation for a node-centric continuous-time random walk (RW) on a graph W_n, which we denote by RW^{W_n}.

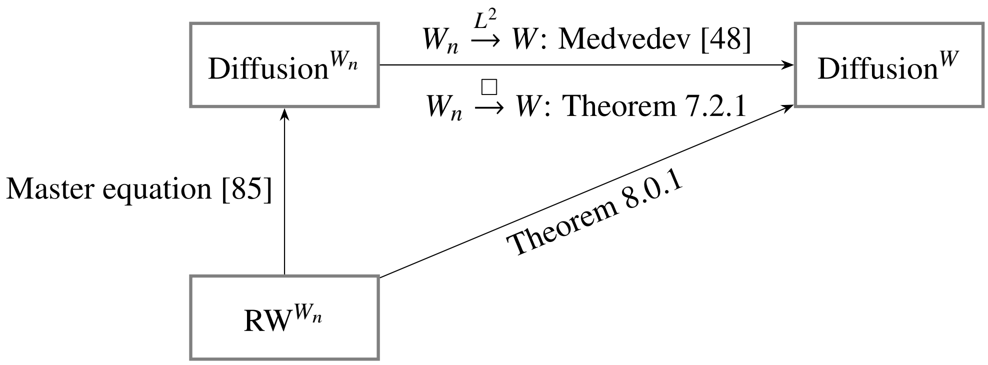


Figure 7.1: Contributions relating to graph diffusion equations.

Figure 7.2 shows more general relations. The diffusion equations are generalized to RD equations, and the random walks are generalized to random walks that are superimposed with birth-death processes (denote these by RWBD^{W_n}). We denote the RW with birth-death processes on a graph W_n by RWBD^{W_n}. Angstmann [85] showed that a RD equation on a graph is the master equation of RWBD^{W_n} on that graph. The reaction function Φ of the limiting RD equation is the birth rate minus the death rate of the stochastic process. Watanabe [86, Theorem 5.1] proves that $\text{RWBD}^{W_n} \to \text{RD}^{W}$ in probability when W_n is a sequence of "quotient graphs", which is a special sequence of graphs that converge in L^2 norm to a given graphon. We extend their result to the case where W_n is any sequence of graphs such that $W_n \xrightarrow{\square} W$.

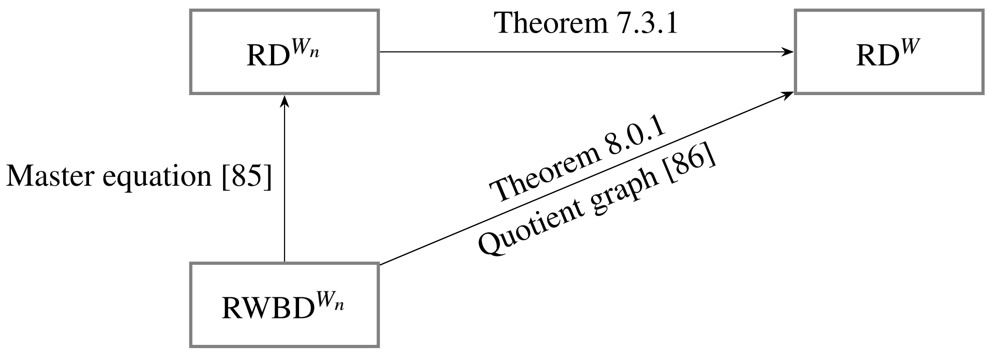


Figure 7.2: Contributions relating to reaction–diffusion equations.

Beyond what is depicted in Figures 7.1 and 7.2, a *graphon random walk* was introduced by Petit et al [87], who prove that the sequence of graph random walks corresponding to W_n converge (in L^2 norm) to the graphon random walk when W_n is a sequence of quotient graphs that converge to W. Furthermore, [87] show that the graphon diffusion equation is the master equation of the graphon random walk. The study of graphon stochastic processes is a topic for future study.

7.1.3 Contributions

In a previous paper [88], we proved that sequences of graph Ginzburg–Landau (GL) functionals Γ-converge to a limiting graphon GL functional. [88] considers a time-independent (i.e., steady-state) problem, whereas this paper considers time-dependent (i.e., dynamical) problems that correspond to gradient flows of energy functionals of which the GL functional is a special case.

The convergence proofs in [88] suggest that graphon GL functionals can serve as proxies for graph GL functionals (as well as minimum-cut functionals), whereas the convergence proofs in this paper suggest that graphon RD equations can serve as proxies for graph RD equations. Effectively, our convergence results show that a large system of nonlocal ODEs on the nodes of a graph can be approximated with a single nonlocal PDE.

Our main results consist of three theorems. For all three theorems, suppose that $W_n \xrightarrow{\square} W$ for some sequence of graphs W_n and a graphon W.

Theorem 7.2.1 shows that the solutions of the sequence of graph diffusion equations corresponding to the sequence $\{W_n\}$ converges to the solution of the graphon diffusion equation corresponding to the limiting graphon. The convergence occurs with a linear rate of convergence in L^p norm for $p \geq 1$. Theorem 7.2.1 extends [48, Theorem 4.1] in two ways: first, we extend the assumption that $W_n \xrightarrow{L^2} W$ to the weaker (see Section 2.1) assumption that $W_n \xrightarrow{\square} W$, and second, we extend convergence of solutions in L^2 norm to convergence in L^p norm for $p \geq 1$.

Theorem 7.3.1 shows that the solutions of the sequence of graph RD equations corresponding to the sequence $\{W_n\}$ converge to the solution of the limiting graphon RD equation corresponding to the graphon W. The convergence occurs in L^p norm for $p \geq 1$ and is linear.

Theorem 8.0.1 is a law of large numbers (LLN) result that generalizes [86, Theorem 5.1]. We define a stochastic process on a graph W_n that consists of a continuous-time random walk and a birth-death process. Theorem 8.0.1 is a weak law LLN that states that as $n \to \infty$, the sequence of graph-based processes converges to the solution of the graphon RD equation corresponding to the graphon W.

7.2 Graphon diffusion equation

Let $W \in L^1((0,1)^2)$ be a graphon. The corresponding graphon diffusion operator (a.k.a. the graphon Laplacian) is defined as

$$\mathcal{L}^W u(x,t) = \int_0^1 W(x,y)\big(u(y,t) - u(x,t)\big)\, dy\,, \tag{7.3}$$

and it is defined for $u \in L^1(0,1)$. Consider a graph on the nodes $[n]$ that has adjacency matrix $A_{ij}^{(n)}$ and let $u_n : [n] \to \mathbb{R}$ be a function on the nodes of the graph. The corresponding graph diffusion operator (a.k.a. the graph Laplacian) is defined as

$$\mathcal{L}^{W_n} u_n(t,i) = \frac{1}{n}\sum_{j=0}^{n-1} A_{ij}^{(n)}\big(u_n(j,t) - u_n(i,t)\big)\,, \tag{7.4}$$

where $u_n(\cdot,t) \in \mathbb{R}^n$ is a n-vector [89]. The graph diffusion operator is a special case of the graphon diffusion operator in the following sense. Define the graphon W_n corresponding to the adjacency matrix $A_{ij}^{(n)}$ as defined in (**??**). Then, the graph diffusion operator (7.4) is equal to the integral operator

$$\mathcal{L}^{W_n} u(x,t) = \int_0^1 W_n(x,y)\big(u(y,t) - u(x,t)\big)\, dy \tag{7.5}$$

when u is a step-function on $(0,1)$ that is constant on the intervals I_k such that u is equal to $u_n(k,t)$ on each I_k for $k = 0, \dots, n-1$. Furthermore, $\mathcal{L}^{W_n}$ can act on any function in $L^1(0,1)$. Hence, (7.5) is a special case of (7.3).

Even though $\mathcal{L}^{W_n}$ is defined on $u \in L^1(0,1)$, we sometimes treat $\mathcal{L}^{W_n}$ as an operator on graph functions (which are stepfunctions on $(0,1)$ that are constant on the intervals I_k). We denote these graph functions by u_n to contrast them from L^1 functions, which are denoted by u.

The graph diffusion equation is defined as

$$\frac{\partial u}{\partial t}(x,t) = \mathcal{L}^{W_n} u(x,t)\,, \tag{7.6}$$

and the graphon diffusion equation is defined as

$$\frac{\partial u}{\partial t}(x,t) = \mathcal{L}^{W} u(x,t)\,. \tag{7.7}$$

Note that $\int_0^1 \mathcal{L}^W u(x,t)dx = 0$ due to the symmetry of W. As a result, $\int_0^1 u_t(x,t)dx = 0$, which means that the total mass is conserved. Thus, we assume zero-flux (Neumann) boundary conditions.

7.2.1 Graphon diffusion equation

For $W \in L^1((0,1)^2)$, let $e^{\mathcal{L}^W t}$ be the semigroup associated with the linear operator $\mathcal{L}^W$. The following lemma shows the stability property of the semigroup. This property is also called non-expansiveness, and in the $p = \infty$ case, it is also called the maximum property.

Lemma 7.2.1. *Let $p \in \mathbb{N} \cup \{+\infty\}$ and let $W \in L^1((0,1)^2)$. For every positive function $g \in L^p(0,1)$*

and for all $t \geq 0$, it holds that

$$\left\|e^{\mathcal{L}^{W}t}g\right\|_{p} \leq \|g\|_{p} .$$

Proof. For $p \in \mathbb{N}$, we have

$$\begin{aligned}
\frac{d}{dt}\|e^{\mathcal{L}^{W}t}g\|_{p}^{p} &= p\left\langle \left|e^{\mathcal{L}^{W}t}g\right|^{p-2} e^{\mathcal{L}^{W}t}g,\, \frac{\partial}{\partial t}e^{\mathcal{L}^{W}t}g \right\rangle \\
&\leq p\int_0^1 \left|e^{\mathcal{L}^{W}t}g(x)\right|^{p-1} \int_0^1 W(x,y)\left[e^{\mathcal{L}^{W}t}g(y) - e^{\mathcal{L}^{W}t}g(x)\right] dy\, dx \\
&= \frac{p}{2}\int_0^1\int_0^1 W(x,y)\left[e^{\mathcal{L}^{W}t}g(y) - e^{\mathcal{L}^{W}t}g(x)\right]\left|e^{\mathcal{L}^{W}t}g(x)\right|^{p-1} dy\, dx \\
&\quad + \frac{p}{2}\int_0^1\int_0^1 W(x,y)\left[e^{\mathcal{L}^{W}t}g(x) - e^{\mathcal{L}^{W}t}g(y)\right] e^{\mathcal{L}^{W}t}\left|e^{\mathcal{L}^{W}t}g(y)\right|^{p-1} dy\, dx \\
&= \frac{p}{2}\int_0^1\int_0^1 W(x,y)\left[e^{\mathcal{L}^{W}t}g(y) - e^{\mathcal{L}^{W}t}g(x)\right]\left[\left|e^{\mathcal{L}^{W}t}g(x)\right|^{p-1} - \left|e^{\mathcal{L}^{W}t}g(y)\right|^{p-1}\right] dy\, dx \\
&\leq 0 .
\end{aligned}$$

The last line is due to the fact that $\left|e^{\mathcal{L}^{W}t}g(\cdot)\right|^{p-1}$ is a monotonic function, hence the two terms in brackets have opposite signs. Gronwall's inequality gives

$$\left\|e^{\mathcal{L}^{W}t}g\right\|_{p} \leq \|g\|_{p} \ . \tag{7.8}$$

Taking the limit $p \to \infty$, we get the inequality (7.8) for $p = \infty$. □

Theorem 7.2.1 extends [48, Theorem 4.1] to the case of cut norm convergence. Specifically, [48, Theorem 4.1] proves the convergence of the solutions of the graph diffusion equations to the solution of a limiting graphon diffusion equation in the case when the graphs converge in L^2 norm to the limiting graphon. We prove a similar result but assume that the graphs converge in cut norm to the limiting graphon.

Theorem 7.2.1. *Let $n \in \mathbb{N}$ and let $p \in \mathbb{N} \cup \{+\infty\}$. Assume that W and W_n are symmetric and measurable functions in $L^1((0,1)^2)$ and that $W_n \xrightarrow{\square} W$ as $n \to \infty$. For $t \geq 0$, let u_n be the solution to the graph diffusion equation (7.6) with initial value $u_n(\cdot, 0) \in L^\infty(0,1)$ and let u be the solution* *to the graphon diffusion equation* (7.7) *with initial value* $u(\cdot, 0) \in L^\infty(0,1)$. *WLOG, assume that* $\|u_n(\cdot,0)\|_\infty = \|u(\cdot,0)\|_\infty = 1$. *Then,*

$$\|u(\cdot,t) - u_n(\cdot,t)\|_p \le \|u(\cdot,0) - u_n(\cdot,0)\|_p + 2t\|W_n - W\|_\square . \tag{7.9}$$

Proof. Define the error $e_n(x,t) = u_n(x,t) - u(x,t)$. We will show that the error decreases in time. Due to the definitions (7.6) and (7.7), have

$$\begin{aligned}\frac{\partial e_n}{\partial t}(x,t) &= \frac{\partial u_n}{\partial t}(x,t) - \frac{\partial u}{\partial t}(x,t)\\ &= \int_0^1 W_n(x,y)\big(u_n(y,t) - u_n(x,t)\big)\,dy - \int_0^1 W(x,y)\big(u(y,t) - u(x,t)\big)\,dy\\ &= \int_0^1 W_n(x,y)\big(e_n(y,t) - e_n(x,t)\big)\,dy + \int_0^1 (W_n - W)(x,y)\big(u(y,t) - u(x,t)\big)\,dy .\end{aligned} \tag{7.10}$$

We can express (7.10) as

$$\frac{\partial e_n}{\partial t} = \mathcal{L}e_n + g_n(u)$$

where

$$\mathcal{L}^{W_n} e_n = \int_0^1 W_n(x,y)\big(e_n(y,t) - e_n(x,t)\big)\,dy$$

and

$$g_n(u) = \int_0^1 (W_n - W)(x,t)(u(y,t) - u(x,t))\,dy .$$

We compute the time-derivative of (7.10). Using the directional derivative of the L^p norm which is the linear functional that is defined, for $h \in L^p$, as

$$(D_f\|f\|_p(h) = \|f\|_p^{1-p} \int_0^1 |f(x)|^{p-2} f(x)h(x)\,dx$$

and the chain rule, we get

$$\frac{1}{p}\frac{d}{dt}\|e_n(\cdot,t)\|_p = \|e_n\|_p^{1-p}\left(\int_0^1 |e_n(x,t)|^{p-2}e_n(x,t)\mathcal{L}^{W_n}e_n\,dx + \int_0^1 |e_n(x,t)|^{p-2}e_n(x,t)g_n(u)\,dx\right)$$
$$=: (I) + (II)\,. \tag{7.11}$$

Term (I) is bounded as

$$\begin{aligned}
(I) &= \|e_n\|_p^{1-p}\int_0^1\int_0^1 |e_n(x,t)|^{p-2}e_n(x,t)W_n(x,y)\big(e_n(y,t)-e_n(x,t)\big)\,dy\,dx\\
&= \|e_n\|_p^{1-p}\bigg(\frac{1}{2}\int_0^1\int_0^1 |e_n(x,t)|^{p-2}e_n(x,t)W_n(x,y)\big(e_n(y,t)-e_n(x,t)\big)\,dy\,dx\\
&\qquad+\frac{1}{2}\int_0^1\int_0^1 |e_n(y,t)|^{p-2}e_n(y,t)W_n(x,y)\big(e_n(x,t)-e_n(y,t)\big)\,dx\,dy\bigg)\\
&= \|e_n\|_p^{1-p}\bigg(\frac{1}{2}\int_0^1\int_0^1 |e_n(x,t)|^{p-2}e_n(x,t)W_n(x,y)\big(e_n(y,t)-e_n(x,t)\big)\,dy\,dx\\
&\qquad-\frac{1}{2}\int_0^1\int_0^1 |e_n(y,t)|^{p-2}e_n(y,t)W_n(x,y)\big(e_n(y,t)-e_n(x,t)\big)\,dx\,dy\bigg)\\
&= \|e_n\|_p^{1-p}\frac{1}{2}\int_0^1\int_0^1 W_n(x,y)\big(e_n(y,t)-e_n(x,t)\big)\Big(|e_n(x,t)|^{p-2}e_n(x,t)-|e_n(y,t)|^{p-2}e_n(y,t)\Big)\,dx\,dy\,.
\end{aligned} \tag{7.12}$$

The function $|e_n(x,t)|^{p-2}e_n(x,t) = |e_n(x,t)|^{p-1}\mathrm{sign}(e_n(x,t))$ is monotonic, and so the quantities $e_n(y,t)-e_n(x,t)$ and $|e_n(x,t)|^{p-2}e_n(x,t)-|e_n(y,t)|^{p-2}e_n(y,t)$ have opposite signs. Hence, $(I) \leq 0$.

Next, we bound (II). Let $q = p/(p-1)$. Using Holder's inequality, the definition of cut norm which is stated in Equation (2.9), we have

$$
\begin{aligned}
(II) &= \|e_n\|_p^{1-p} \int_0^1 |e_n(x,t)|^{p-2} e_n(x,t) g_n(u) dx \\
&\leq \|e_n\|_p^{1-p} \int_0^1 |e_n(x,t)|^{p-1} |g_n(u)| dx \\
&= \|e_n\|_p^{1-p} \left(\int_0^1 |e_n(x,t)|^{(p-1)q} dx \right)^{1/q} \|g_n\|_p \\
&= \|e_n\|_p^{1-p} \|e_n\|_p^{p-1} \left\| \int_0^1 (W_n - W)(x,y)(u(y,t) - u(x,t))\, dy \right\|_p \\
&\leq \left(\int \left(\int (W_n - W)(x,y)(u(y,t) - u(x,t))\, dy \right)^p dx \right)^{1/p} \\
&\leq 2 \|W_n - W\|_\square . \qquad (7.13)
\end{aligned}
$$

where the last line follows from the fact that $\|u_n(\cdot,0)\|_\infty \leq 1$ and $\|u(\cdot,0)\|_\infty \leq 1$ implies that $\|u_n(\cdot,t)\|_\infty \leq 1$ and $\|u(\cdot,t)\|_\infty \leq 1$ due to Lemma 7.2.1, and so $\|e_n(\cdot,t)^{p-1}\|_\infty \leq 2^{p-1}$.

$$
\frac{d}{dt} \|e_n(\cdot,t)\|_p \leq 2 \|W_n - W\|_\square . \qquad (7.14)
$$

Gronwall's lemma applied to (7.14) implies that

$$
\|e_n(t,\cdot)\|_p \leq \|e_n(0,\cdot)\|_p + 2t \|W_n - W\|_\square . \qquad (7.15)
$$

□

Remark 7.2.1. *One can generalize Theorem 7.2.1 to nonlinear diffusion, in which the term $u(y,t) - u(x,t)$ in the diffusion operator is replaced by $D(u(y,t) - u(x,t))$ where $D : \mathbb{R} \to \mathbb{R}$ is a Lipschitz-continuous function. [48, Theorem 4.1] holds for nonlinear diffusion.*

Remark 7.2.2. *One can generalize Theorem 7.2.1 to any convex function of $u_n - u$ (in Theorem 7.2.1, the convex function is $\|\cdot\|_p$). A convex function is needed because, in order to bound term (I), a monotonic multiplying factor is needed.*

7.3 Graphon reaction–diffusion equation

7.3.1 Graphon reaction–diffusion equation

Consider the reaction–diffusion equation

$$u_t = \mathcal{L}u + \Phi(u) \tag{7.16}$$

where $\mathcal{L}$ is a diffusion operator and $f : \mathbb{R} \to \mathbb{R}$ is a nonlinear function. Classically, $\mathcal{L}$ is the Laplacian operator [34]. We define the graph and graphon RD equations by using the graph and graphon diffusion operators in place of $\mathcal{L}$. The graph RD equation, corresponding to a graph W_n, is

$$\frac{\partial u_n}{\partial t}(x,t) = \mathcal{L}^{W_n} u_n(x,t) + \Phi(u_n(x,t)) \tag{7.17}$$

and the graphon RD equation, corresponding to a graphon W, is

$$\frac{\partial u}{\partial t}(x,t) = \mathcal{L}^{W} u(x,t) + \Phi(u(x,t))\,. \tag{7.18}$$

We consider the initial value problems associated with the dynamics and we obtain convergence of the solution of Equation (7.17) to the solution of Equation (7.18) as $n \to \infty$. First, we restate a lemma from [90] that provides a maximum bound property for nonlocal RD equations.

Lemma 7.3.1 (Theorem 2.3 [90]). *Assume that Φ is continuously differentiable on $[-1,1]$ and that $\Phi(1) \le 0 \le \Phi(-1)$. Let u be the solution to the graphon RD equation* (7.18) *for some $W \in L^1((0,1)^2)$. Let $T > 0$. If*

$$\|u(\cdot,0)\|_\infty \le 1\,,$$

then for all $t \in [0,T]$,

$$\|u(\cdot,t)\|_\infty \le 1\,.$$

Note that this lemma holds for both (7.17) and (7.18) because (7.17) is a special case of (7.18).

Also note that Lemma 7.3.1 implies that Φ is globally Lipschitz continuous because we assume that Φ is continuously differentiable on $[-1, 1]$.

Theorem 7.3.1 (Convergence of solutions of graph RD to graphon RD equations)**.** *Let $p \geq 1$. Let W_n be a sequence of graphs such that $W_n \xrightarrow{\square} W$ for some limiting graphon W. For each W_n, let u_n be the solution to the graph RD Equation* (7.17) *with graph W_n and let u be the solution to the graphon RD Equation* (7.18) *with graphon W. Assume that $\|u_n(\cdot, 0) - u(\cdot, 0)\|_p \to 0$ as $n \to \infty$. Then,*

$$\|e_n(\cdot, t)\|_p \leq \|e_n(\cdot, 0)\|_p \, e^{2Kt} + 2\|W_n - W\|_\square \frac{\left(e^{2Kt} - 1\right)}{K} .$$

Proof. Let $e_n(x,t) = u_n(x,t) - u(x,t)$.

$$\begin{aligned}
\frac{\partial e_n}{\partial t}(x,t) &= \frac{\partial u_n}{\partial t}(x,t) - \frac{\partial u}{\partial t}(x,t) \\
&= \int_0^1 W_n(x,y)\big(u_n(y,t) - u_n(x,t)\big)\, dy - \int_0^1 W(x,y)\big(u(y,t) - u(x,t)\big)\, dy \\
&\qquad + \Phi(u_n(x,t)) - \Phi(u(x,t)) \\
&= \int_0^1 W_n(x,y)\big(e_n(y,t) - e_n(x,t)\big)\, dy + \int_0^1 (W_n - W)(x,y)\big(u(y,t) - u(x,t)\big)\, dy \\
&\qquad + \Phi(u_n(x,t)) - \Phi(u(x,t)) . \qquad (7.19)
\end{aligned}$$

Then, e_n satisfies the equation

$$\frac{\partial e_n}{\partial t}(x,t) = Le_n + g_n(u)$$

where

$$\mathcal{L}^{W_n} e_n = \int_0^1 W_n(x,y)\big(e_n(y,t) - e_n(x,t)\big)\, dy ,$$

$$g_n(u) = \int_0^1 (W_n - W)(x,y)\big(u(y,t) - u(x,t)\big)\, dy + \Phi(u_n(x,t)) - \Phi(u(x,t)) .$$

Using the chain rule and the directional derivative (7.2.1), we have

$$
\begin{aligned}
\frac{d}{dt}\|e_n(\cdot,t)\|_p^p &= \|e_n\|_p^{1-p}\int_0^1 |e_n(x,t)|^{p-2}e_n(x,t)\big(\mathcal{L}^{W_n}e_n + g_n(u)\big)dx \\
&= \|e_n\|_p^{1-p}\bigg(\int_0^1 |e_n(x,t)|^{p-2}e_n(x,t)\int_0^1 W_n(x,y)\big(e_n(y,t)-e_n(x,t)\big)\,dy\,dx \\
&\quad + \int_0^1\int_0^1 |e_n(x,t)|^{p-2}e_n(x,t)(W_n-W)(x,y)\big(u(y,t)-u(x,t)\big)\,dy\,dx \\
&\quad + \int_0^1 |e_n(x,t)|^{p-2}e_n(x,t)\big(\Phi(u_n(x,t))-\Phi(u(x,t))\big)\,dx\bigg) \\
&= \|e_n\|_p^{1-p}\left((I)+(II)+(III)\right) .
\end{aligned}
$$

The calculations in Section 7.2.1 (see Equations (7.12) and (7.13)) give

$$
\|e_n\|_p^{1-p}\left((I)+(II)\right) \leq 2\|W_n - W\|_\square . \tag{7.20}
$$

To bound (III), let $q = p/(p-1)$. Hölder's inequality and the fact that Φ is K-Lipschitz give

$$
\begin{aligned}
\|e_n\|_p^{1-p}(III) &= \|e_n\|_p^{1-p}\int_0^1 |e_n(x,t)|^{p-2}e_n(x,t)\big(\Phi(u_n(x,t))-\Phi(u(x,t))\big)\,dx \\
&\leq \|e_n\|_p^{1-p}\|e_n^{p-1}\|_q\|\Phi(u_n(x,t))-\Phi(u(x,t))\|_p \\
&= \|e_n\|_p^{1-p}\|e_n\|_p^{p-1}\|\Phi(u_n(x,t))-\Phi(u(x,t))\|_p \\
&\leq K\|e_n(\cdot,t)\|_p .
\end{aligned} \tag{7.21}
$$

Combining the bounds (7.20) and (7.21), we have that

$$
\frac{1}{2}\frac{d}{dt}\|e_n(\cdot,t)\|_p \leq 2\|W_n - W\|_\square + K\|e_n(\cdot,t)\|_p . \tag{7.22}
$$

Gronwall's inequality applied to (7.22) gives

$$
\|e_n(\cdot,t)\|_p \leq \|e_n(\cdot,0)\|_p\, e^{2Kt} + 2\|W_n - W\|_\square \frac{\left(e^{2Kt}-1\right)}{K} .
$$

□

Remark 7.3.1. *Theorem 7.3.1 reduces to Theorem 7.2.1 in the case* $K = 0$ *because* $\frac{(e^{2Kt}-1)}{K} \to t$ *as* $K \to 0$*. Note that if* Φ *is a monotone function, so that the resulting energy is convex, then the corresponding dynamic system has a long time asymptotic stability. In such a case, one may eliminate the exponential factor* e^{2Kt} *or even prove the conclusion with a negative exponent* K*.*

Remark 7.3.2. *Just like in the diffusion equation case (see Remark 7.2.2), one can generalize Theorem 7.3.1 to show that any convex function of* $u_n - u$ *converges to zero as* $n \to \infty$*.*

Chapter 8: Law of large numbers

In this section, we describe an interacting particle process on graphs that consists of a continuous-time random walk superimposed with a birth-death process. For a sequence of graphons W_n converging in cut norm to a graphon W, we show that the sequence of particle processes on W_n converges in probability to the solution of a nonlocal reaction–diffusion (RD) equation, where the nonlocal diffusion is weighted by the graphon W. In this section, we assume that W_n and W are in $L^\infty((0,1)^2)$. This assumption is used when proving that the stochastic part of the error goes to zero (section 8.0.1).

Let $T_n(t)u$ be the solution of the graph diffusion equation

$$\frac{d}{dt}u(x,t) = \mathcal{L}^{W_n}u(x,t) \tag{8.1}$$

where $\mathcal{L}^{W_n}$ is the graphon diffusion operator defined in (7.3). By solving the linear differential equation (8.1) with an integrating factor, one can express the corresponding semigroup operator using the implicit equation

$$T_n(t)(u_0(x)) = \int_0^1 W_n(x,y)T_n(t)(u_0(y))\,dy\left(1-e^{-t}\right) + u_0(x)e^{-t}\,. \tag{8.2}$$

We now define a stochastic process X^n whose hydrodynamic limit is the solution of (8.1). Consider a graph W_n with nodes $[n]$. Suppose that at time $t = 0$, there is some number of particles on each node. Let $m_k(t)$ be the number of particles on node k at time t, and let $\vec{m}(t) = (m_1(t), m_2(t), \ldots, m_n(t))$. Let $e_i \in \mathbb{R}^n$ be the ith unit vector, and let $\vec{m}$ undergo the stochastic transitions

$$\begin{cases} \vec{m} \to \vec{m} + e_i - e_k & \text{at rate } m_k W_n\left(k/n, i/n\right) n^{-1}, \\ \vec{m} \to \vec{m} + e_k & \text{at rate } lb\left(m_k/l\right), \\ \vec{m} \to \vec{m} - e_k & \text{at rate } ld\left(m_k/l\right). \end{cases} \tag{8.3}$$

Define the stochastic process

$$X^n(x,t) = \frac{m_k(t)}{l} \quad \text{for } x \in I_k\,, \tag{8.4}$$

which is the density of particles if l is a parameter that is proportional to the initial total number of particles. We assume that $l \to \infty$, which is needed in order to keep X^n bounded as $n \to \infty$. This way, the reaction term that acts on X^n will effectively be Lipschitz [91].

The quantity

$$Z^n(t) = X^n(t) - X^n(0) - \int_0^t \mathcal{L}^{W_n}(X^n(s))\, ds - \int_0^t \Phi(X^n(s))\, ds \tag{8.5}$$

is a martingale due to [91, Lemma 2.2]. One can prove that (8.5) is a martingale by showing that $\mathcal{L}^{W_n} + \Phi$ is the infinitesimal generator of X^n, and then using Dynkin's semigroup formula. Hence, we can also write X^n as

$$X^n(t) = X^n(0) + \int_0^t \mathcal{L}^{W_n}(X^n(s))\, ds + \int_0^t \Phi(X^n(s))ds + Z^n(t)\,, \tag{8.6}$$

where $\Phi(x) = b(x) - d(x)$ is a reaction function that describes the birth-death process. We assume that b and d are Lipschitz continuous.

We apply the variation of constants method to (8.6)—that is, we multiply both sides by $e^{-\mathcal{L}^{W_n}t}$, integrate both sides, and rearrange—to get

$$X^n(t) = e^{\mathcal{L}^{W_n}t}X^n(0) + \int_0^t e^{\mathcal{L}^{W_n}(t-s)}\Phi(X^n(s))\, ds + \int_0^t e^{\mathcal{L}^{W_n}(t-s)}\, dZ^n(s)\,. \tag{8.7}$$

Similarly, the solution u_n of the graph RD equation (7.17) is given by

$$u_n(t) = e^{\mathcal{L}^{W_n}(t)} u_n(0) + \int_0^t e^{\mathcal{L}^{W_n}(t-s)} \Phi(u_n(s))\, ds\,. \tag{8.8}$$

Observe that (8.7) and (8.8) look similar, but (8.7) has an additional stochastic term. Our LLN theorem is as follows.

Theorem 8.0.1. *Let W_n and W be L^∞ graphons, and assume that $W_n \overset{\square}{\to} W$. Assume that the reaction term Φ is continuously differentiable on $[-1, 1]$. Let u be the solution to the graphon RD equation 7.18. Suppose that*

[(i)]$\|X^n(0) - u(0)\|_p \to 0$ in probability as $n \to \infty$, and $l = l(n) \to \infty$ as $n \to \infty$.

Then, for $T > 0$,

$$sup_{0 \le t \le T} \|X^n(t) - u(t)\|_p \to 0 \quad \text{in probability as } n \to \infty\,. \tag{8.9}$$

In order to prove Theorem 8.0.1 we use the convergence of the diffusion and reaction terms as shown in Sections 7.2.1 and 7.3.1. We use four lemmas: Lemma 7.2.1 states that $e^{\mathcal{L}^{W_n} t}$ is nonexpansive, Lemma 8.0.1 states that $e^{\mathcal{L}^{W_n}}$ converges to $e^{\mathcal{L}^{W}}$, and Lemma 7.3.1 ensures that the solution of the graphon RD equation is bounded in L^∞. We also prove Lemma 8.0.6, which states that a remaining stochastic term, which is the last term on the right-hand side of (8.7), converges to zero as $n \to \infty$. The proof of Lemma 8.0.6 is more involved; it is in section 8.0.1.

Lemma 8.0.1. *Let $e^{\mathcal{L}^{W_n}}$ and $e^{\mathcal{L}^{W}}$ be the semigroups generated by the operators $\mathcal{L}^{W_n}$ and $\mathcal{L}^{W}$, respectively. Let $g \in L^p(0, 1)$ and let $g_n \in L^p(0, 1)$ be the piecewise-constant function that is constant on the intervals I_k that is obtained by averaging g on each I_k.*

$$\|e^{\mathcal{L}^{W_n} t} g_n - e^{\mathcal{L}^{W} t} g\|_p \to 0 \qquad \text{for all } 0 \le t \le T\,.$$

In particular, as $n \to \infty$,

$$\|e^{\mathcal{L}^{W_n} t} g_n - e^{\mathcal{L}^{W} t} g\|_p \leq \|g_n - g\|_p + 2^p\, t\|W_n - W\|_\square .$$

2. *Proof.* The result follows from Theorem 7.2.1. □

Proof. **Proof of Theorem 8.0.1** Let u_n be the solution to the graph RD equation (7.17). By the triangle inequality,

$$\|X^n(t) - u(t)\|_p \leq \|X^n(t) - u_n(t)\|_p + \|u_n(t) - u(t)\|_p .$$

Theorem 7.3.1 implies that for sufficiently large n and $p \geq 1$, we have

$$\|u_n(t) - u(t)\|_p \leq \|u_n(0) - u(0)\|_p e^{2Kt} + 2\|W_n - W\|_\square \frac{e^{2Kt-1}}{K} .$$

Define the stopping time

$$\tau = \sup\left\{t : \|X^n(t) - u_n(t)\|_p \leq \epsilon\right\}$$

for some $\epsilon > 0$. [90, Theorem 2.3] implies that for all $t \in [0, T]$ and for all $x \in [0, 1]$, we have $\|u_n\| \leq \rho$ for some $\rho > 0$. As a result, for $t < \tau$, we have

$$\|X^n(t)\|_p < \rho + \epsilon . \tag{8.10}$$

Furthermore, because the jump size is 1, we have that $\|X^n(t \wedge \tau)\|_\infty < \rho + \epsilon + 1$.

Let

$$\overline{X}^n(t) = X^n(t \wedge \tau) + \int_{t\wedge\tau}^{t} \mathcal{L}^{W_n} \overline{X}^n(s) ds + \int_{t\wedge\tau}^{t} \Phi\left(\overline{X}^n(s)\right) ds \tag{8.11}$$

be the stochastic process that is equal to $X(t)$ up until the stopping time τ, at which time $\overline{X}$ begins to move according to the same integral equation as u_n.

Note that

$$P\left(\sup_{t\in[0,T]} \|X^n(t)-u(t)\|_\infty \geq \epsilon\right) \leq P\left(\sup_{t\in[0,T]} \left\|\overline{X}^n(t)-u(t)\right\|_\infty \geq \epsilon\right). \tag{8.12}$$

Consequently, in order to prove Theorem 8.0.1 it suffices to bound the right-hand side of (8.12). We proceed using the fact that $\overline{X}^n$ and u_n satisfy the same equation starting at time τ and that at time τ they differ by ϵ.

We can also write $\overline{X}^n(t)$ as

$$\overline{X}^n(t) = \overline{X}^n(0) + \int_0^t \mathcal{L}^{W_n}\overline{X}^n(s)ds + \int_0^t \Phi(\overline{X}^n(s))ds + Z^n(t\wedge\tau). \tag{8.13}$$

Using the variation of constants method, we get

$$\overline{X}^n(t) = e^{\mathcal{L}^{W_n}t}\overline{X}^n(0) + \int_0^t e^{\mathcal{L}^{W_n}(t-s)}\Phi(\overline{X}^n(s))ds + \int_0^t e^{\mathcal{L}^{W_n}(t-s)}dZ^n(s\wedge\tau). \tag{8.14}$$

Due to equations (8.8) and (8.14), we have

$$\begin{aligned}\|\overline{X}^n(t)-u_n(t)\|_p &\leq \left\|e^{\mathcal{L}^{W_n}t}\big(\overline{X}^n(0)-u_n(0)\big)\right\|_p + \left\|\int_0^t e^{\mathcal{L}^{W_n}(s-z)}\big(\Phi(\overline{X}^n(z))-\Phi(u(z))\big)\,dz\right\|_p \\ &\quad + \left\|\int_0^t e^{\mathcal{L}^{W_n}(s-z)}\,dZ^n(s\wedge\tau)\right\|_p \\ &\leq (I)+(II)+(III)\end{aligned}$$

Lemma 7.2.1 implies that

$$(I) \leq \|\overline{X}^n(0)-u_n(0)\|_p. \tag{8.15}$$

Lemmas 7.2.1 and 8.0.1 imply that

$$\begin{aligned}(II) &\le \int_0^t \left\| e^{\mathcal{L}^{W_n}(s-z)}\big(\Phi(\overline{X}^n(z)) - \Phi(u_n(z))\big)\right\|_p \, dz \\ &\le \int_0^t \left\|\big(\Phi(\overline{X}^n(z)) - \Phi(u_n(z))\big)\right\|_p \, dz \\ &\le \int_0^t K\|\overline{X}^n(z) - u_n(z)\|_p \, dz \,. \end{aligned} \tag{8.16}$$

Due to equations (8.15) and (8.16), we have

$$\|\overline{X}^n(t) - u_n(t)\|_p \le \|\overline{X}^n(0) - u_n(0)\|_p + K\int_0^t \|\overline{X}^n(z) - u_n(z)\|_p \, dz + (III) \,.$$

Gronwall's inequality for integral inequalities implies that

$$\|\overline{X}^n(t) - u_n(t)\|_p \le \|\overline{X}^n(0) - u(0)\|_p + Ke^{Kt}\int_0^t c_n(z)e^{-Kz}\, dz + (III) \,. \tag{8.17}$$

In Subsection 8.0.1, we prove that the stochastic term (III) converges to zero as $n \to \infty$. □

8.0.1 Stochastic term

We emulate the proof of [92, Theorem 4.1] while citing lemmas from [86]. We rewrite the stochastic term as

$$Y^n(t) = \int_0^t T_n(t-s)\, dZ^n(s \wedge \tau) \tag{8.18}$$

where

$$T_n(t) = e^{\mathcal{L}^{W_n} t} \tag{8.19}$$

is the semigroup generated by the graphon diffusion operator $\mathcal{L}^{W_n}$. After stating several lemmas, we prove the main result (Lemma 8.0.6) which states that the L^p norm of (8.18) converges to zero in probability as $n \to \infty$ and for $p \ge 1$.

Let $\Lambda X^n(t) = X^n(t) - X^n(t^-)$ be the jump of a process X^n at time t.

Lemma 8.0.2. *[86, Lemma 4.2] Let ξ be a step-function that is constant on the intervals I_k, for $k = 0, 2, \ldots, n-1$. Then,*

$$\sum_{s \le t} \left(\Lambda \langle Z^n(s \wedge \tau), \xi \rangle\right)^2 - \frac{1}{nl} \int_0^{t \wedge \tau} \left\{ \left\langle X^n(s, \cdot), \int_0^1 W_n(\cdot, y)(\xi(y) - \xi(x))^2 \, dy \right\rangle + \left\langle |\Phi|(X^n(s, \cdot)), \xi^2(\cdot) \right\rangle \right\} ds \tag{8.20}$$

is a mean-zero martingale.

The following lemma is similar to [93, Lemma 4.2]), except that [93, Lemma 4.2]) holds for the Laplacian operator corresponding to fully-connected graphs $W_n \equiv 1$, whereas we generalize it to all graphons.

Lemma 8.0.3. *Let W_n be a graph and let $f_k = n\mathbf{1}_{I_k}$. We have that*

$$\left\langle (T_n(t)f)^2 - T_n(t)f_k \mathcal{L}^{W_n} \left(T_n(t)f_k\right), 1 \right\rangle \le 2ne^{-2t} + 1 \,. \tag{8.21}$$

Proof. The proof is similar to the proof of [93, Lemma 4.2]. We separately bound the first term $(I) := \langle \left(T_n(t)f_k\right)^2, 1\rangle$ and the second term $(II) := -\langle T_n(t)f_k \mathcal{L}^W \left(T_n(t)f_k\right), 1\rangle$. Because T_n is self-adjoint, we have

$$\langle (T_n(t)f_k)^2, 1\rangle = \langle T_n(2t)f_k, f_k\rangle \,.$$

Using the definition of T_n (Equation (8.2)) and Lemma 7.2.1 (which implies that $\|T_n(t)f\|_\infty \le \|f\|_\infty$ for any $f \in L^\infty(0,1)$) we have

$$\begin{aligned} T_n(2t)f_k(x) &= \int_0^1 W_n(x,y)T_n(2t)(f_k(y))\, dy\, (1 - e^{-2t}) + f_k(x)e^{-2t} \\ &\le \|W_n\|_\infty \left(1 - e^{-2t}\right) + f_k(x)e^{-2t} \,. \end{aligned} \tag{8.22}$$

Assuming that $\|W_n\|_\infty = 1$, we have that (I) satisfies

$$\langle \big(T_n(t) f_k\big)^2, 1\rangle \le \Big\langle \big(1 - e^{-2t}\big) + f_k e^{-2t},\ f_k \Big\rangle \le n e^{-2t} + 1\,. \tag{8.23}$$

(II) satisfies

$$\begin{aligned}
-\langle T_n(t) f_k \mathcal{L}^W \left(T_n(t) f_k\right), 1\rangle &= -\int T_n(t) f_k(x) \int W(x,y)\big(T_n(t) f_k(y) - T_n(t) f_k(x)\big)\, dx \\
&= -\int \int W(x,y)\big(T_n(2t) f_k(x) f_k(y) - T_n(2t) f_k(x)\big)\, dx \\
&= -\int f_k(x) T_n(2t) \int W(x,y)\big(f_k(y) - f_k(x)\big)\, dx \\
&= -\langle T_n(2t)\mathcal{L}^W f_k,\ f_k\rangle\,.
\end{aligned} \tag{8.24}$$

Next, using (8.2) and the fact that $f_k = n\mathbf{1}_{I_k}$, we have

$$\begin{aligned}
T_n(2t)\mathcal{L}^W f_k(x) &= \int_0^1 W(x,y) T_n(2t)(\mathcal{L}^W f_k(y))\, dy \left(1 - e^{-2t}\right) + \mathcal{L}^W f_k(x) e^{-2t} \\
&\le \left(\int_0^1 W(x,y) \int_{k/n}^{(k+1)/n} n W(y,z)\, dz - f_k(y)\, dy\right)(1 - e^{-2t}) \\
&\quad + \left(\int_{k/n}^{(k+1)/n} n W(x,y) - f_k(x)\, dy\right) e^{-2t} \\
&= \left(d_k - \int_0^1 W(x,y) f_k(y)\, dy\right)(1 - e^{-2t}) + (d_k - f_k(x))\, e^{-2t} \\
&= (d_k - f_k(x)) e^{-2t}
\end{aligned}$$

where $d_k = \int_{I_k} W(x,y) dx \le \int_0^1 W(x,y) dx \le \|W\|_\infty$ is the degree of the kth node. Hence, due to

Equation (8.24), we have that (II) satisfies

$$
\begin{aligned}
-\left\langle T_n(t) f_k \mathcal{L}^W \left(T_n(t) f_k\right), 1\right\rangle &= -\langle (d_k - f_k) e^{-2t}, f_k\rangle \\
&= -e^{-2t} \int_{k/n}^{(k+1)/n} (d_k - n) n \, dx \\
&= e^{-2t}(n - d_k) \\
&\leq n e^{-2t} \,. \qquad (8.25)
\end{aligned}
$$

Combining (8.23) and (8.25), we conclude that (I) + (II) $\leq 2ne^{-2t} + 1$. □

The following lemma states that the integrand in (8.20) is bounded by an integrable function.

Lemma 8.0.4. *Denote* $\xi = T_n(t-s) f_k$ *where* $f_k = \mathbf{1}_{I_k}$.

$$
\left\langle X^n(s), \int_0^1 W_n(\cdot, y)\Big(\xi(\cdot) - \xi(y)\Big)^2 dy \right\rangle + \left\langle |\Phi|(X^n(s)), \xi^2 \right\rangle \;\leq\; c(\rho)(ne^{-2t} + 1) \,. \qquad (8.26)
$$

where $c(\rho)$ *is a constant that depends linearly on* ρ*. Recall that* ρ *is the upper bound as seen in* (8.10)*. Note that the integral of this bound is* $\int_0^t c(\rho)(ne^{-2s} + 1) ds \leq c(\rho)(n + t)$*.*

Proof. We bound the two terms on the left-hand side of Equation (8.26) separately.

$$
\begin{aligned}
\Big\langle X^n(s), \int_0^1 W_n(\cdot, y)\Big(\xi(\cdot) - \xi(y)\Big)^2\Big\rangle \\
&= \int\int W_n(x,y) X^n(s,x)\,(\xi(x) - \xi(y))^2\ dx \\
&\le \|X^n(s)\|_\infty \int\int W_n(x,y)\,(\xi(x) - \xi(y))^2\ dx \\
&= \|X^n(s)\|_\infty \Big[\int\int W_n(x,y)(\xi(x) - \xi(y))\xi(x)\,dx \\
&\qquad - \int\int W_n(x,y)(\xi(x) - \xi(y))\xi(y)\,dx\Big] \\
&= \|X^n(s)\|_\infty \Big[\int\int W_n(x,y)(\xi(x) - \xi(y))\xi(x)\,dx \\
&\qquad + \int\int W_n(x,y)(\xi(x) - \xi(y))\xi(x)\,dx\,dy\Big] \\
&= 2\|X^n(s)\|_\infty \int\int W_n(x,y)(\xi(x) - \xi(y))\xi(x)\,dx\,dy \\
&= 2\|X^n(s)\|_\infty \langle -\xi\mathcal{L}^{W_n}\xi,\ 1\rangle \\
&\le 2(\rho+1)\langle -\xi\mathcal{L}^{W_n}\xi,\ 1\rangle\,, \qquad (8.27)
\end{aligned}
$$

where the last line follows from the fact that for times up to the stopping time τ, X^n is bounded as in (8.10). Furthermore,

$$
\begin{aligned}
\langle|\Phi|(X^n(s)), \xi^2\rangle &= \int_0^1 |\Phi|(X^n(s,x))\xi^2(x)\,dx \\
&\le \big\||\Phi|(X^n)\big\|_\infty \langle \xi^2, 1\rangle \\
&\le c(\rho)\,\langle \xi^2, 1\rangle\,,
\end{aligned}
$$

where c is a linear function of ρ. The last inequality holds because $\|X^n(t)\|_\infty < \rho + 2$ for $t \le \tau$ and because the reaction function Φ is Lipschitz. Thus, the left-hand side of Equation (8.26) is bounded above by $c(\rho)\langle\xi^2 - \xi\mathcal{L}^{W_n}\xi,\ 1\rangle$ where c is a linear function of ρ. For $\xi = T_n(t-s)f_k$, Lemma 8.0.3 gives the bound $\langle\xi^2 - \xi\mathcal{L}^{W_n}\xi,\ 1\rangle \le 2ne^{-2t} + 1$. Thus, the proof is finished. □

The following lemma is restated from [92, Lemma 4.4] (but with minor changes in notation).

Lemma 8.0.5. *([92, Lemma 4.4]) : Let $F_t^{n,l}$ be the completion of the σ-algebra $\sigma(n(s) : s \leq t)$. Let $m(t)$ be a bounded martingale of finite variation defined on $[0,T]$ with $m(0) = 0$ and satisfying*

[(i)]m is right-continuous with left limits. $|\Lambda m(t)| \leq 1$ for $0 \leq t \leq T$. $\sum_{0\leq s\leq t}(\Lambda m(s))^2 - \int_0^t g(s)\,ds$ is a mean-zero martingale with $0 \leq g(s) \leq h(s)$, where $h(s)$ is a bounded deterministic function and $g(s)$ is $F_t^{n,l}$ adapted.

Then, $\mathbb{E}[\exp(m(T))] \leq \exp\left(\frac{3}{2}\int_0^T h(s)\,ds\right)$.

Lemma 8.0.6. *Let Y^n be defined as in* (8.18)*. Assume the corresponding graphon W_n is L^∞. Then,* $\sup_{[0,T]}\|Y^n(t)\|_p \to 0$ *in probability as $n \to \infty$.*

1. *Proof.* The proof first bounds Y^n restricted to a spatial interval I_k and a temporal interval $[0,\bar{t}]$; this "slice" of Y^n is called $\bar{m}$. Then, we estimate the variance of the jumps of $\bar{m}$ and show that that estimate is bounded; with this information, Lemma 8.0.5 gives a bound on $\bar{m}$. We obtain a bound on the whole spatial domain by adding up the bounds on all sub-intervals I_k. To obtain a bound on the whole time interval, we divide the time interval $[0,T]$ into n intervals. Note that Y^n is only nonzero up to time τ, so it would suffice to prove that $\sup_{[0,\tau]}\|Y^n(t)\|_p \to 0$ in probability. However, we show the convergence for the whole interval to avoid confusion.

Fix $\bar{t} \in (0,T]$ and $k \in \{0,1,\ldots,n-1\}$. Let $f_k = n\mathbf{1}_{I_k}$. For $s \in [0,\bar{t}]$, let

$$\bar{m}(s) = \left\langle \int_0^s T_n(\bar{t}-s')\,dZ^n(s'\wedge\tau),\ f_k \right\rangle. \tag{8.28}$$

Note that $\bar{m}$ is a mean-zero martingale for $s \in [0,\bar{t}]$ because Z^n is a martingale. Also note that $\bar{m}(\bar{t}) = Y^n(\bar{t}, k/n)$ and that $\bar{m}(s) = \langle Z^n(s\wedge\tau), \xi\rangle$ for $\xi(x) = \int_0^s T_n(\bar{t}-s')\,ds' f_k(x)$.

For $t \in [0,T]$, Lemma 8.0.2 implies that

$$\sum_{s\leq t}(\Lambda\bar{m}(s))^2 - \frac{1}{nl}\int_0^{t\wedge\tau}\left\{\int_0^1 X^n(s,x)\int_0^1 W_n(x,y)(\xi(x)-\xi(y))^2dx + \left\langle|\Phi|(X^n(s)),\,\xi^2\right\rangle\right\}ds \tag{8.29}$$

is a mean-zero martingale. Let $m(s) = \theta l \bar{m}(s)$. Equation (8.30) multiplied by $(\theta l)^2$, i.e.

$$\sum_{s \le t} (\Lambda m(s))^2 - \frac{\theta^2 l}{n} \int_0^{t \wedge \tau} \left\{ \int_0^1 X^n(s,x) \int_0^1 W_n(x,y)(\xi(x) - \xi(y))^2 dx + \left\langle |\Phi|(X^n(s)), \xi^2 \right\rangle \right\} ds \,, \tag{8.30}$$

is also a mean-zero martingale. Requirements (i) and (ii) of Lemma 8.0.5 hold for m. Lemma 8.0.4 implies that requirement (iii) also holds, in which $h(s) = c(\rho)(ne^{-2t} + 1)$. Hence, Lemma 8.0.5 implies that

$$\mathbb{E}[\exp(\theta l m(\bar{t})] \le \exp\left(\theta^2 l c(\rho)(1 + \bar{t}/n)\right) \le \exp(\theta^2 l c(\rho)) \,, \tag{8.31}$$

where the second inequality follows from the fact that $\bar{t}/n \le 1$. Hence, we have that

$$\begin{aligned} P(Y^n(\bar{t}, k/n) > \epsilon) &= P(\exp(\theta l Y^n(\bar{t}, k/n)) > \exp(\theta l \epsilon)) \\ &\le \mathbb{E}[\exp(\theta l Y^n(\bar{t}, k/n))]\exp(-\theta l \epsilon) \\ &= \mathbb{E}[\exp(m(\bar{t}))]\exp(-\theta l \epsilon) \\ &\le \exp(\theta l (c(\rho)\theta - \epsilon)) \end{aligned}$$

where the second line is due to Markov's inequality and the last line follows from (8.31). We choose $\theta \in [0, 1]$ that depends only on ρ such that $\theta l(c(\rho)\theta - \epsilon) = -a\epsilon^2 l$ for some $a = a(\rho) > 0$. This gives

$$P(Y^n(\bar{t}, k/n) > \epsilon) \le \exp(-a\epsilon^2 l) \,. \tag{8.32}$$

We obtain a bound on $Y^n(\bar{t})$ in $L^2(0,1)$ by using Equation (8.32) for all k; that is,

$$\begin{aligned}
P\left(\|Y^n(\bar{t})\|_p > \epsilon\right) &= P\left(\sum_{k=1}^{n} (Y^n(\bar{t}, k/n))^p > \epsilon^p\right) \\
&\leq \sum_{k=1}^{n} P\left(|Y^n(\bar{t}, k/n)|^p > \epsilon^p\right) \\
&\leq 2\sum_{k=1}^{n} P\left(Y^n(\bar{t}, k/n) > \epsilon\right) \\
&\leq 2n\exp(-a\epsilon^2 l)\,. \qquad (8.33)
\end{aligned}$$

Next, we show a bound for the whole interval $[0,T]$. We split the time interval $[0,T]$ into n intervals $I_0, I_1, \ldots, I_{n-1}$. We use the same interval-size and notation as for the spatial intervals. Differentiating and integrating (8.18) gives

$$Y^n(t) = \int_0^t \mathcal{L}^{W_n} Y^n(s)\,ds + Z^n(t \wedge \tau) \qquad (8.34)$$

Hence, for $t \in I_m$,

$$Y^n(t) = Y^n(mT/n) + \int_{mT/n}^t \mathcal{L}^{W_n} Y^n(s)\,ds + \tilde{\omega}(t)\,, \qquad (8.35)$$

where

$$\tilde{\omega}(t) = Z^n(t \wedge \tau) - Z^n(mT/n \wedge \tau)\,. \qquad (8.36)$$

Taking the L^p norm of (8.35) gives

$$\begin{aligned}
\|Y^n(t)\|_p &\leq \|Y^n(mT/n)\|_p + \int_{mT/n}^t \|\mathcal{L}^{W_n} Y^n(s)\|_p\,ds + \|\tilde{\omega}(t)\|_p \\
&\leq \|Y^n(mT/n)\|_p + 2\|W_n\|_\infty \int_{mT/n}^t \|Y^n(s)\|_p\,ds + \|\tilde{\omega}(t)\|_p\,, \qquad (8.37)
\end{aligned}$$

where (8.37) holds because

$$
\begin{aligned}
\|\mathcal{L}^{W_n}Y^n(s)\|_p^p &= \int_0^1 \left(\int_0^1 W_n(x,y)(Y^n(s,y) - Y^n(s,x)dy\right)^p dx \\
&\le \|W_n\|_\infty^p \int_0^1 \int_0^1 (Y^n(s,y) - Y^n(s,x))^p dy\, dx \\
&\le \|W_n\|_\infty^p \sum_{i=0}^{p} \binom{p}{k} \int_0^1 \int_0^1 Y^n(s,y)^{p-k} \left(-Y^n(s,x)\right)^k dy\, dx \\
&\le \|W_n\|_\infty^p 2^p \|Y^n(s)\|_p^p \,,
\end{aligned}
$$

where the second line holds because of Jensen's inequality.

We assume WLOG that $\|W_n\|_\infty = 1$. Then, Gronwall's inequality applied to (8.37) gives

$$
\|Y^n(t)\|_p \le \left\{\|Y^n(mT/n)\|_p + \|\tilde{\omega}(t)\|_p\right\} \exp(2(t - mT/n))\,. \tag{8.38}
$$

Taking the supremum of (8.38) over the time interval $[mT/n, (m+1)T/n]$ gives

$$
\sup_{t\in[mT/n,(m+1)T/n]} \|Y^n(t)\|_p \quad \le \left\{\|Y^n\,(mT/n)\|_p + \sup_{t\in[mT/n,(m+1)T/]} \|\tilde{\omega}(t)\|_p\right\} e^{2T/n}\,. \tag{8.39}
$$

We then fix $k \in \{0, 1, \ldots, n-1\}$ and $\theta \in [0, 1]$, and let $\omega(t) = \theta l \tilde{\omega}(t)$. Then, $|\Lambda\omega(t)| \le 1$ because $|\Lambda X^n(t)| \le 1$. [86, Lemma 4.1b] implies that

$$
\begin{aligned}
\sum_{mT/n\le s\le t} (\Lambda\omega(s))^2 - \theta^2 l \int_{mT/n\wedge\tau}^{t\wedge\tau} \frac{1}{n} \sum_{\substack{i=0\\ i\neq k}}^{n-1} W_n(i/n, k/n)(X^n(s, i/n) + X^n(s, k/n))\, ds \\
- \theta^2 l \int_{mT/n\wedge\tau}^{t\wedge\tau} |\Phi|(X^n(s, k/n))\, ds
\end{aligned} \tag{8.40}
$$

is a mean-zero martingale for $t \in I_m$. The integrand of (8.40) is bounded by a constant function $c(\rho)$ of ρ because $\|X^n(s,\cdot)\|_\infty \le \rho + 2$ for $s \in [0, \tau]$ and the fact that $|\Phi|$ is K-Lipschitz. Thus, applying Lemma 8.0.5, we obtain

$$\mathbb{E}[\exp(\omega((m+1)T/n)] \leq \exp\left(\int_{mT/n}^{(m+1)T/n} c(\rho)\theta^2 l\, ds\right) \leq \exp(c(\rho)\theta^2 lT/n)\,. \tag{8.41}$$

Using Doob's submartingale inequality and a similar argument as in (8.32), we can choose $\theta \in [0,1]$ and $a = a(\rho) > 0$ so that

$$\begin{aligned} P\left(\sup_{t\in[mT/n,(m+1)T/n]} \tilde{\omega}\,(t, k/n) \geq \epsilon\right) &\leq \exp(\theta l(c(\rho)\theta T/n - \epsilon)) \\ &\leq \exp\left(-a\epsilon^2 l\right). \end{aligned}$$

We abbreviate $\sup_{t\in[mT/n,(m+1)T/n]}$ as $\sup_{t\in I_m}$. The bound of the L^p norm is

$$\begin{aligned} P\left(\sup_{t\in I_m} \|\omega(t)\|_p \geq \epsilon\right) &= P\left(\sup_{t\in I_m} \sum_{k=1}^{n} (\tilde{\omega}(t,k/n))^p \geq \epsilon^p\right) \\ &\leq \sum_{k=1}^{n} P\left(\sup_{t\in I_m} (\tilde{\omega}(t,k/n))^2 \geq \epsilon^2\right) \\ &\leq 2\sum_{k=1}^{n} P\left(\sup_{t\in I_m} \tilde{\omega}(t,k/n) \geq \epsilon\right) \\ &\leq 2n\exp(-a\epsilon^2 l)\,. \end{aligned} \tag{8.42}$$

Combining the bounds (8.33), (8.39), and (8.42), we obtain

$$\begin{aligned} P\left(e^{-2T/n} \sup_{t\in I_m} \|Y^n(t)\|_p > \epsilon\right) &= P\left(\|Y^n\,(mT/n)\|_p + \sup_{t\in[mT/n,(m+1)T/]} \|\tilde{\omega}(t)\|_p > \epsilon\right) \\ &\leq P\big(\,\|Y^n\,(mT/n)\|_p > \epsilon\big) + P\left(\sup_{t\in[mT/n,(m+1)T/]} \|\tilde{\omega}(t)\|_p > \epsilon\right) \\ &\leq 4n\exp(-a\epsilon^2 l)\,. \end{aligned}$$

We obtain a bound on the full time interval $[0, T]$ as follows:

$$
\begin{aligned}
P\left(e^{-2T} \sup_{t\in[0,T]} \|Y^n(t)\|_p > \epsilon\right) &\le \sum_{m=0}^{n-1} P\left(e^{-2T} \sup_{t\in I_m} \|Y^n(t)\|_p > \epsilon\right) \\
&\le 4n^2 \exp(-a\epsilon^2 l)\,. \qquad (8.43)
\end{aligned}
$$

Due to assumption (ii) of Theorem 8.0.1, the quantity (8.43) converges to zero as $n \to \infty$.

□

Conclusion

Graphon limits of graph-theoretic problems offer new analytical tools and mathematical insight into those problems [13, 4, 44]. In this dissertation, we consider the graphon limits of graph GL functionals and graph RD equations. In both cases, we show that the solutions of the graph-based problems converge to the solutions of the limiting graphon-based problems.

In Chapters 3 through 5 of this dissertation, we defined graphon GL and TV functionals, and we showed that their minimizers are consistent with minimizers of the graph GL and TV functionals in the following sense. Given a sequence of graphs W_n that converge in cut norm to a limiting graphon W, the sequence of graph GL functionals (3.8) Γ-converges to the graphon GL functional (3.19). The sequence of graph TV functionals (3.10) also Γ-converges to the graphon TV functional (3.21) as $W_n \overset{\square}{\to} W$. Additionally, we showed that $\mathrm{GL}_\epsilon^W \overset{\Gamma}{\to} \mathrm{TV}^W$ as $\epsilon \to 0$, which resembles the known Γ-limits $GL_\epsilon \overset{\Gamma}{\to} TV$ in the classical case and $GL_\epsilon^{W_n} \overset{\Gamma}{\to} TV^{W_n}$ in the graph case. All of these Γ-convergence results (see our summary in Figure 3.1), in concert with compactness properties of Young measures, which we obtain via Prohorov's Theorem, imply that the minimizers of the Γ-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 the graphon functionals are formulated using Young measures, rather than using functions. This difference highlights that the limiting minimizers are Young measures, which constitute families of functions that can have arbitrary amounts of oscillation, while taking the same values in the same proportions. Chapter 6 contains the derivation of these Young measure minimizers for several examples of graphons.

Another difference between the graph GL and graphon GL functionals, which is highlighted by our limit (4) (see Section 5.2 for the proof of this limit), is that the ϵ-scaling of the graphon GL functional is somewhere between the ϵ-scalings of the classical GL functional and the graph GL functional. The classical GL functional has scalings ϵ 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 ϵ-scalings as the graph GL functional when the graphon W is bounded. We did not determine a scaling for the more general $W \in L^p((0,1)^2)$.

Our limit (3), which we proved in Section 5.1, extends results of Braides et al. [11, Lemma 11, Theorem 12] in two key ways. First, Braides et al. proved Γ-convergence of the graph-cut functional (which acts on finite-range functions) as $n \to \infty$, whereas we proved Γ-convergence of the graph Dirichlet energy, which extends the graph-cut functional to act on L^∞ functions. Second, Braides et al worked with dense sequences of graphs that converge to L^∞ graphons, whereas we proved Γ-convergence for more general sequences of graphs that converge to L^p graphons.

In Chapter 7, we defined graphon RD equations as nonlocal RD equations, where the diffusion term is a nonlocal integral operator whose kernel is the accompanying graphon. Suppose that a sequence of graphs (or graphons) W_n converges in cut norm to a limiting graphon W. Our first convergence result states that the solutions to the sequence of graph RD equations corresponding to the graphs W_n converge to the solution of the graphon RD equation corresponding to W. In other words, graph RD equations, which, for a given graph with n nodes, are a coupled system of n ordinary differential equations, converge in the large-n limit to a single nonlocal partial differential equation that is parameterized by a graphon.

Our second main convergence result that relates to graphon RD equations is discussed in Chapter 8. There, we define a stochastic random-walk birth-death particle system on graphs and show that it converges in probability to the solution of a limiting graphon RD equation. This result is a graph-to-graphon hydrodynamic limit that suggests that certain agent-based models on graphs can be approximated by a single nonlocal PDE.

Our results focused on theory, but one can examine applications of graphon GL functionals and graphon RD equations. For example, graph GL and TV functionals have been employed in image processing [32, 31, 33, 65, 29, 66, 67, 68, 69], and it is worthwhile to pursue analogous applications of graphon GL and TV functionals. Our convergence result on graph RD equations suggests that graphon RD equations can be used to approximate RD equations on large graphs, which include models of population ecology, disease spread, and chemical reactions [35, 82, 34]. Our LLN result in Chapter 8 suggests that microscopic (a.k.a. agent-based) models can be approximated by the macroscopic RD equation. Such agent-based models include models of ecology, information spread, disease, and other social dynamics [94, 95, 96, 97]. We can also extend the convergence results of RD equations to systems of RD equations. For example, SIR-type models with diffusion consist of a system of three RD equations [35].

Another direction for future work is to investigate stochastic particle systems on graphons: that is, one could find a graphon limit of the random-walk birth-death process on graphs. [87] defines a random walk on graphons and shows formally that it is a limit of random walks on an approximating sequence of graphs. They show that the graph random walk converges to the graphon diffusion equation (in a special case of so-called quotient graphs). One could extend their results to include a birth-death process on graphons.

Another question for future research is how to apply our results in practical computations. While the convergence results in Chapters 7 and 8 include convergence rates, the Γ-convergence results do not. 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 [51, 52, 54], 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 (see [77, 78]).

The mathematics of graphon-based problems are somewhat different from the mathematics of graph-based problems. Because graphons are nonlocal kernels and induce nonlocal operators, they introduce numerous theoretical connections between graph theory and nonlocal analysis. In this dissertation, we have developed connections between graphs, graphons, and nonlocal analysis. Connecting these rich mathematical fields to each other suggests many possibilities for future directions in the theory of infinite graphs.

References

[1] A. Barabási, "Network science," *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 371, no. 1987, p. 20 120 375, 2013.

[2] G. S. Medvedev, "Small-world networks of Kuramoto oscillators," *Physica D: Nonlinear Phenomena*, vol. 266, pp. 13–22, 2014.

[3] F Parise and A Ozdaglar, "Graphon games," in *Proceedings of the 2019 ACM Conference on Economics and Computation*, 2019, pp. 457–458.

[4] L Ruiz, L. F. Chamon, and A Ribeiro, "Graphon signal processing," *IEEE Transactions on Signal Processing*, vol. 69, pp. 4961–4976, 2021.

[5] S. Gao and P. E. Caines, "Graphon control of large-scale networks of linear systems," *IEEE Transactions on Automatic Control*, vol. 65, no. 10, pp. 4090–4105, 2019.

[6] R. Vizuete, P. Frasca, and F. Garin, "Graphon-based sensitivity analysis of SIS epidemics," *IEEE Control Systems Letters*, vol. 4, no. 3, pp. 542–547, 2020.

[7] P. J. Wolfe and S. C. Olhede, "Nonparametric graphon estimation," *arXiv preprint arXiv:1309.5936*, 2013.

[8] L Lovász, *Large Networks and Graph Limits*. Providence: American Mathematical Soc., 2012.

[9] C. Borgs, J. T. Chayes, H. Cohn, and N. Holden, "Sparse exchangeable graphs and their limits via graphon processes," *Journal of Machine Learning Research*, vol. 18, no. 210, pp. 1–71, 2018.

[10] C. Borgs, J. T. Chayes, L. Lovász, V. T. Sós, and K. Vesztergombi, "Convergent sequences of dense graphs I: Subgraph frequencies, metric properties and testing," *Advances in Mathematics*, vol. 219, no. 6, pp. 1801–1851, 2008.

[11] A. Braides, P. Cermelli, and S. Dovetta, "Γ-limit of the cut functional on dense graph sequences," *ESAIM: Control, Optimisation and Calculus of Variations*, vol. 26, p. 26, 2020.

[12] P. Diaconis and S. Janson, "Graph limits and exchangeable random graphs," *arXiv preprint arXiv:0712.2749*, 2007.

[13] J. J. Bramburger, “The emerging utility of graphons in applied math,” *SIAM News*, 2023, Available at `https://sinews.siam.org/Details-Page/the-emerging-utility-of-graphons-in-applied-math`.

[14] Q Du, *Nonlocal Modeling, Analysis, and Computation*. Philadelphia: SIAM, 2019.

[15] G. S. Medvedev, “The nonlinear heat equation on dense graphs and graph limits,” *SIAM Journal on Mathematical Analysis*, vol. 46, no. 4, pp. 2743–2766, 2014.

[16] E. Bayraktar, S. Chakraborty, and R. Wu, “Graphon mean field systems,” *The Annals of Applied Probability*, vol. 33, no. 5, pp. 3587–3619, 2023.

[17] J Picard and M Queyranne, “Selected applications of minimum cuts in networks,” *INFOR: Information Systems and Operational Research*, vol. 20, no. 4, pp. 394–422, 1982.

[18] M. E. J. Newman, *Networks*, Second. Oxford, UK: Oxford University Press, 2018.

[19] Y. van Gennip *et al.*, “Community detection using spectral clustering on sparse geosocial data,” *SIAM Journal on Applied Mathematics*, vol. 73, no. 1, pp. 67–83, 2013.

[20] L. I. Rudin, S. J. Osher, and E. Fatemi, “Nonlinear total variation based noise removal algorithms,” *Physica D: Nonlinear Phenomena*, vol. 60, no. 1–4, pp. 259–268, 1992.

[21] Z. Wu and R. Leahy, “An optimal graph theoretic approach to data clustering: Theory and its application to image segmentation,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 15, no. 11, pp. 1101–1113, 1993.

[22] F. Yi and I. Moon, “Image segmentation: A survey of graph-cut methods,” in *2012 International Conference on Systems and Informatics (ICSAI 2012)*, Institute of Electrical and Electronics Engineers, Piscataway, NJ, USA, 2012, pp. 1936–1941.

[23] J. Shi and J. Malik, “Normalized cuts and image segmentation,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 8, pp. 888–905, 2000.

[24] 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.

[25] G Dantzig and D. Fulkerson, “On the max flow min cut theorem of networks,” *Linear Inequalities and Related Systems*, vol. 38, pp. 225–231, 2003.

[26] 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*, Institute of Electrical and Electronics Engineers, Piscataway, NJ, USA, 2010, pp. 2217–2224.

[27] L. R. Ford and D. R. Fulkerson, “Maximal flow through a network,” *Canadian Journal of Mathematics*, vol. 8, pp. 399–404, 1956.

[28] E. Merkurjev, E. Bae, A. L. Bertozzi, and X.-C. Tai, “Global binary optimization on graphs for classification of high-dimensional data,” *Journal of Mathematical Imaging and Vision*, vol. 52, pp. 414–435, 2015.

[29] Y. Lou, X. Zhang, S. J. Osher, and A. L. Bertozzi, “Image recovery via nonlocal operators,” *Journal of Scientific Computing*, vol. 42, no. 2, pp. 185–197, 2010.

[30] X. Zhang and T. F. Chan, “Wavelet inpainting by nonlocal total variation,” *Inverse Problems and Imaging*, vol. 4, no. 1, pp. 191–210, 2010.

[31] I. El Bouchairi, A. Elmoataz, and J. Fadili, “Nonlocal perimeters and curvature flows on graphs with applications in image processing and high-dimensional data classification,” *SIAM Journal on Imaging Sciences*, vol. 16, no. 1, pp. 368–392, 2023.

[32] L. A. Caffarelli, J.-M. Roquejoffre, and O. Savin, “Nonlocal minimal surfaces,” *Communications on Pure and Applied Mathematics*, vol. 63, no. 9, pp. 1111–1144, 2010.

[33] G. Gilboa and S. J. Osher, “Nonlocal operators with applications to image processing,” *Multiscale Modeling & Simulation: A SIAM Interdisciplinary Journal*, vol. 7, no. 3, pp. 1005–1028, 2009.

[34] N. Britton, *Reaction-diffusion Equations and Their Applications to Biology.* London: Elsevier, 1986.

[35] R. Cantrell and C Cosner, *Spatial Ecology via Reaction-Diffusion Equations*. West Sussex, UK: John Wiley & Sons, 2004.

[36] V. K. Vanag and I. R. Epstein, “Cross-diffusion and pattern formation in reaction–diffusion systems,” *Physical Chemistry Chemical Physics*, vol. 11, no. 6, pp. 897–912, 2009.

[37] A De Masi, P. Ferrari, and J. Lebowitz, “Reaction-diffusion equations for interacting particle systems,” *Journal of Statistical Physics*, vol. 44, no. 3, pp. 589–644, 1986.

[38] P. E. Caines and M. Huang, “Graphon mean field games and their equations,” *SIAM Journal on Control and Optimization*, vol. 59, no. 6, pp. 4373–4399, 2021.

[39] B. Bollobás and O. Riordan, “Metrics for sparse graphs,” in *Surveys in Combinatorics 2009*, ser. London Mathematical Society Lecture Note Series, S. Huczynska, J. D. Mitchell, and C. M. Roney-Dougal, Eds., Cambridge, UK: Cambridge University Press, 2009, pp. 211–288.

[40] 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," *The Annals of Probability*, vol. 46, no. 1, pp. 337–396, 2018.

[41] 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," *Transactions of the American Mathematical Society*, vol. 372, no. 5, pp. 3019–3062, 2019.

[42] B. Bollobás and O. Riordan, "Sparse graphs: Metrics and random models," *Random Structures & Algorithms*, vol. 39, no. 1, pp. 1–38, 2011.

[43] L. Lovász and B. Szegedy, "Szemerédi's lemma for the analyst," *GAFA Geometric And Functional Analysis*, vol. 17, pp. 252–270, 2007.

[44] S Janson, "Graphons, cut norm and distance, couplings and rearrangements," *arXiv preprint arXiv:1009.2376*, 2010.

[45] C. Borgs, J. T. Chayes, L. Lovász, V. T. Sós, and K. Vesztergombi, "Convergent sequences of dense graphs II. Multiway cuts and statistical physics," *Annals of Mathematics*, vol. 176, no. 1, pp. 151–219, 2012.

[46] I. El Bouchairi, J. Fadili, Y. Hafiene, and A. Elmoataz, *Nonlocal Continuum Limits of P-Laplacian Problems on Graphs*. Cambridge University Press, 2023.

[47] J. Bramburger and M. Holzer, "Pattern formation in random networks using graphons," *SIAM Journal on Mathematical Analysis*, vol. 55, no. 3, pp. 2150–2185, 2023.

[48] G. Medvedev, "The nonlinear heat equation on dense graphs and graph limits," *SIAM Journal on Mathematical Analysis*, vol. 46, no. 4, pp. 2743–2766, 2014.

[49] G. Heinze, "Graph-based nonlocal gradient systems and their local limits," 2024.

[50] Y van Gennip and A. Bertozzi, "Γ-convergence of graph Ginzburg–Landau functionals," *Advances in Differential Equations*, vol. 17, no. 11–12, pp. 1115–1180, 2012.

[51] J. Budd and Y. van Gennip, "Graph Merriman–Bence–Osher as a semidiscrete implicit Euler scheme for graph Allen–Cahn flow," *SIAM Journal on Mathematical Analysis*, vol. 52, no. 5, pp. 4101–4139, 2020.

[52] E. Merkurjev, T. Kostic, and A. L. Bertozzi, "An MBO scheme on graphs for classification and image processing," *SIAM Journal on Imaging Sciences*, vol. 6, no. 4, pp. 1903–1930, 2013.

[53] Y. van Gennip, N. Guillen, B. Osting, and A. L. Bertozzi, "Mean curvature, threshold dynamics, and phase field theory on finite graphs," *Milan Journal of Mathematics*, vol. 82, pp. 3–65, 2014.

[54] A. L. Bertozzi and A. Flenner, "Diffuse interface models on graphs for classification of high dimensional data," *SIAM Review*, vol. 58, no. 2, pp. 293–328, 2016.

[55] A. Braides, "A handbook of Γ-convergence," in *Handbook of Differential Equations: Stationary Partial Differential Equations. Volume 3*, M. Chipot and P. Quittner, Eds., Amsterdam, The Netherlands: Elsevier, 2006, pp. 101–213.

[56] Q. Du, M. D. Gunzburger, and J. S. Peterson, "Analysis and approximation of the ginzburg–landau model of superconductivity," *Siam Review*, vol. 34, no. 1, pp. 54–81, 1992.

[57] M. Cyrot, "Ginzburg–Landau theory for superconductors," *Reports on Progress in Physics*, vol. 36, no. 2, pp. 103–158, 1973.

[58] S. M. Allen and J. W. Cahn, "A microscopic theory for antiphase boundary motion and its application to antiphase domain coarsening," *Acta Metallurgica*, vol. 27, no. 6, pp. 1085–1095, 1979.

[59] Q. Du and X. Feng, "The phase field method for geometric moving interfaces and their numerical approximations," in *Handbook of Numerical Analysis. Volume 21. Geometric Partial Differential Equations — Part I*, A. Bonito and R. H. Nochetto, Eds., Chapter 5, Amsterdam, The Netherlands: Elsevier, 2020, pp. 425–508.

[60] L. Modica and S. Mortola, "Un esempio di Γ^{-}-convergenza," *Bollettino dell'Unione Matematica Italiana B*, vol. 14, pp. 285–299, 1977.

[61] L. C. Evans, *Partial Differential Equations*, Second. Providence, RI, USA: American Mathematical Society, 2022.

[62] L. Bronsard and R. V. Kohn, "Motion by mean curvature as the singular limit of Ginzburg–Landau dynamics," *Journal of Differential Equations*, vol. 90, no. 2, pp. 211–237, 1991.

[63] N. G. Trillos and D. Slepčev, "Γ-convergence of perimeter on random geometric graphs," Carnegie Mellon University, Tech. Rep., 2013, Available at `https://www.math.cmu.edu/PIRE/pub/pire66/13-CNA-024.pdf`.

[64] N. García Trillos and D. Slepčev, "Continuum limit of total variation on point clouds," *Archive for Rational Mechanics and Analysis*, vol. 220, pp. 193–241, 2016.

[65] X. Liu and L. Huang, "A new nonlocal total variation regularization algorithm for image denoising," *Mathematics and Computers in Simulation*, vol. 97, pp. 224–233, 2014.

[66] P. Arias, V. Caselles, G. Facciolo, V. Lazcano, and R. Sadek, “Nonlocal variational models for inpainting and interpolation,” *Mathematical Models and Methods in Applied Sciences*, vol. 22, no. supp 02, p. 1 230 003, 2012.

[67] 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*, Springer-Verlag, Heidelberg, Germany, 2015, pp. 66–77.

[68] A. El Chakik, A. Elmoataz, and X. Desquesnes, “Mean curvature flow on graphs for image and manifold restoration and enhancement,” *Signal Processing*, vol. 105, pp. 449–463, 2014.

[69] Y. Hafiene, J. Fadili, and A. Elmoataz, “Nonlocal p-laplacian evolution problems on graphs,” *SIAM Journal on Numerical Analysis*, vol. 56, no. 2, pp. 1064–1090, 2018.

[70] J. Mazón, J. Rossi, and J Toledo, “Nonlocal perimeter, curvature and minimal surfaces for measurable sets,” *Journal d'Analyse Mathématique*, vol. 138, pp. 235–279, 2019.

[71] O. Savin and E. Valdinoci, “Γ-convergence for nonlocal phase transitions,” *Annales de l'Institut Henri Poincaré C: Analyse Non Linéaire*, vol. 29, no. 4, pp. 479–500, 2012.

[72] V. E. Tarasov and G. M. Zaslavsky, “Fractional Ginzburg–Landau equation for fractal media,” *Physica A: Statistical Mechanics and its Applications*, vol. 354, pp. 249–261, 2005.

[73] X. Pu and B. Guo, “Well-posedness and dynamics for the fractional Ginzburg–Landau equation,” *Applicable Analysis*, vol. 92, no. 2, pp. 318–334, 2013.

[74] A. L. Bertozzi and A. Flenner, “Diffuse interface models on graphs for classification of high dimensional data,” *Multiscale Modeling & Simulation*, vol. 10, no. 3, pp. 1090–1118, 2012.

[75] B. Dacorogna, *Weak Continuity and Weak Lower Semicontinuity of Non-Linear Functionals*. Heidelberg, Germany: Springer-Verlag, 1982.

[76] M. Valadier, “Young measures,” *Methods of Nonconvex Analysis*, vol. 1446, pp. 152–188, 1990.

[77] R. A. Nicolaides and N. J. Walkington, “Computation of microstructure utilizing Young measure representations,” *Journal of Intelligent Material Systems and Structures*, vol. 4, no. 4, pp. 457–462, 1993.

[78] C. Carstensen and T. Roubíček, “Numerical approximation of young measures in non-convex variational problems,” *Numerische Mathematik*, vol. 84, pp. 395–415, 2000.

[79] Q. Du and J. Yang, “Asymptotically compatible fourier spectral approximations of nonlocal allen–cahn equations,” *SIAM Journal on Numerical Analysis*, vol. 54, no. 3, pp. 1899–1919, 2016.

[80] T. P. Peixoto, “Bayesian stochastic blockmodeling,” in *Advances in Network Clustering and Blockmodeling*, P. Doreian, V. Batagelj, and A. Ferligoj, Eds., Hoboken, NJ, USA: John Wiley & Sons, Inc., 2019, pp. 289–332.

[81] H. Yuasa and M. Ito, “Internal observation systems and a theory of reaction-diffusion equation on a graph,” in *1998 IEEE International Conference on Systems, Man, and Cybernetics*, IEEE, vol. 4, 1998, pp. 3669–3673.

[82] V Colizza, R Pastor-Satorras, and A Vespignani, “Reaction-diffusion processes and metapopulation models in heterogeneous networks,” *Nature Physics*, vol. 3, no. 4, pp. 276–282, 2007.

[83] R. Van Gorder, “A theory of pattern formation for reaction-diffusion systems on temporal networks,” *Proceedings of the Royal Society A*, vol. 477, no. 2247, p. 20 200 753, 2021.

[84] D. Scalise and R. Schulman, “Emulating cellular automata in chemical reaction–diffusion networks,” *Natural Computing*, vol. 15, pp. 197–214, 2016.

[85] C. N. Angstmann, I. C. Donnelly, and B. I. Henry, “Pattern formation on networks with reactions: A continuous-time random-walk approach,” *Physical Review E—Statistical, Nonlinear, and Soft Matter Physics*, vol. 87, no. 3, p. 032 804, 2013.

[86] I Watanabe, “Continuum limit of nonlocal diffusion on inhomogeneous networks,” *Journal of Dynamics and Differential Equations*, pp. 1–20, 2022.

[87] J. Petit, R. Lambiotte, and T. Carletti, “Random walks on dense graphs and graphons,” *SIAM Journal on Applied Mathematics*, vol. 81, no. 6, pp. 2323–2345, 2021.

[88] E Zhang, J Scott, Q Du, and M. Porter, “Ginzburg–Landau functionals in the large-graph limit,” *arXiv preprint arXiv:2408.00422*, 2024.

[89] R. Merris, “Laplacian matrices of graphs: A survey,” *Linear Algebra and its Applications*, vol. 197, pp. 143–176, 1994.

[90] Q. Du, L. Ju, X. Li, and Z. Qiao, “Maximum bound principles for a class of semilinear parabolic equations and exponential time-differencing schemes,” *SIAM Review*, vol. 63, no. 2, pp. 317–359, 2021.

[91] D. J. Blount, *Comparison of a stochastic model of a chemical reaction with diffusion and the deterministic model*. The University of Wisconsin–Madison, 1987.

[92] D. Blount, “Law of large numbers in the supremum norm for a chemical reaction with diffusion,” *The annals of applied probability*, vol. 2, no. 1, pp. 131–141, 1992.

[93] I. Watanabe and H. Toyoizumi, "Comparison between the deterministic and stochastic models of nonlocal diffusion," *Journal of Dynamics and Differential Equations*, vol. 36, no. 1, pp. 231–250, 2024.

[94] E. Lieberman, C. Hauert, and M. A. Nowak, "Evolutionary dynamics on graphs," *Nature*, vol. 433, no. 7023, pp. 312–316, 2005.

[95] F. Altarelli, A. Braunstein, L. Dall'Asta, and R. Zecchina, "Optimizing spread dynamics on graphs by message passing," *Journal of Statistical Mechanics: Theory and Experiment*, vol. 2013, no. 09, P09011, 2013.

[96] D Aldous, "Interacting particle systems as stochastic social dynamics," *Bernoulli*, vol. 19, no. 4, pp. 1122 –1149, 2013.

[97] J Zhang and J. M. Moura, "Diffusion in social networks as SIS epidemics: Beyond full mixing and complete graphs," *IEEE Journal of Selected Topics in Signal Processing*, vol. 8, no. 4, pp. 537–551, 2014.