

Robust Recovery of Robinson Property in L^p -Graphons: A Cut-Norm Approach

Mahya Ghandehari^a Teddy Mishura^b

Submitted: Oct 3, 2023; Accepted: Sep 26, 2024; Published: Nov 1, 2024
© The authors. Released under the CC BY-ND license (International 4.0).

Abstract

This paper investigates the Robinson graphon completion/recovery problem within the class of L^p -graphons, focusing on the range $5 < p \leq \infty$. A graphon w is Robinson if it satisfies the Robinson property: if $x \leq y \leq z$, then $w(x, z) \leq \min\{w(x, y), w(y, z)\}$. We demonstrate that if a graphon possesses localized near-Robinson characteristics, it can be effectively approximated by a Robinson graphon in terms of cut-norm. To achieve this recovery result, we introduce a function Λ , defined on the space of L^p -graphons, which quantifies the degree to which a graphon w adheres to the Robinson property. We prove that Λ is a suitable gauge for measuring the Robinson property when proximity of graphons is understood in terms of cut-norm. Namely, we show that (1) $\Lambda(w) = 0$ precisely when w is Robinson; (2) Λ is cut-norm continuous, in the sense that if two graphons are close in the cut-norm, then their Λ values are close; and (3) for $p > 5$, any L^p -graphon w can be approximated by a Robinson graphon, with error of the approximation bounded in terms of $\Lambda(w)$. When viewing w as a noisy version of a Robinson graphon, our method provides a concrete recipe for recovering a cut-norm approximation of a noiseless w . Given that any symmetric matrix is a special type of graphon, our results can be applicable to symmetric matrices of any size. Our work extends and improves previous results, where a similar question for the special case of L^∞ -graphons was answered.

Mathematics Subject Classifications: 05C62, 15A83, 15B52

1 Introduction

Given an incomplete matrix M_I with many missing entries, is it possible to recover the underlying complete matrix M ? An instance of such a scenario is when data is aggregated from customer ratings, where any given customer only rates a few items. In general,

^aDepartment of Mathematical Sciences, Ewing Hall, University of Delaware in Newark, DE, U.S.A. (mahya@udel.edu).

^bDepartment of Mathematics, Victoria Building, Toronto Metropolitan University in Toronto, ON, Canada (tmishura@torontomu.ca).

this problem is ill-posed; indeed, with low dimensional information, it is not possible to uniquely recover high dimensional data. However, in practical applications, it is often the case that additional structure about the original matrix M is known. In the above example regarding customer ratings, it is usually assumed that most people's preferences are based on very few factors, implying that the complete matrix M should be low rank. Leveraging this assumption makes the completion of M_I quite tractable, even if the majority of its entries are missing. In general, the problem of filling in the incomplete entries in a matrix M_I under the condition that the completed matrix satisfies a specified structural property is known as the *matrix completion problem*. Due to its many applications in data science, the completion problem for low rank matrices has gained the attention of many researchers (see e.g. [9, 10, 30, 31, 34, 49, 54, 56]). Furthermore, the matrix completion problem has been explored in the context of various other essential structural properties, including Hankel, Toeplitz, and moment structures (see [16, 17, 29, 57] for some examples). Unfortunately, even known information can be corrupted by noise. In this case, the problem of completion becomes that of *recovery*: Given an observed matrix \hat{M} , find a matrix M satisfying a prescribed structural property such that $\|M - \hat{M}\|$ is small. For examples of recovery of noisy matrices, see [1, 8, 33, 35, 36, 37, 51, 55] for low rank matrices, [18, 43] for monotone matrices, and [7, 12, 52] for covariance matrices.

An important structural property of matrices/graphons is the *Robinson* property. A *Robinson matrix*, also called an *R-matrix*, is a symmetric matrix $A = [a_{ij}]$ such that for $i \leq j \leq k$ we have

$$a_{ik} \leq \min\{a_{ij}, a_{jk}\}.$$

Robinson matrices are well-studied objects in data science, largely because of their connection with the classical seriation problem [19, 40, 45, 48, 53]. The objective of the seriation problem is to use pairwise comparisons of a set of items to recover their linear ordering. The seriation problem is easily transferred to the problem of determining whether a given symmetric matrix can be permuted into a Robinson matrix, and finding the Robinson ordering. This problem can be solved in polynomial time; see [45] for the original algorithm, and [2, 20, 38, 39] for more recent efficient algorithms. When the underlying matrix is perturbed by noise, recovery of a Robinson ordering turns out to be a very challenging problem. For a discussion on the NP-hardness of this problem when ℓ^p -norm approximations (with $p < \infty$) are required, see [4]; and for a polynomial time algorithm when seeking ℓ^∞ -norm approximations, see [13].

To put matrices of various size in one framework, and also to develop a probabilistic Robinson theory that can treat matrices of growing sizes, we focus on *graphons* as a versatile extension of symmetric matrices. Graphons are measurable real-valued functions on $[0, 1]^2$ that are symmetric with respect to the diagonal. A graphon is called an L^p -graphon if its p -norm is finite. L^∞ -graphons are typically referred to as graphons, and were introduced in [42] as the limit objects of converging sequences of dense graphs (i.e. graph sequences $\{G_n\}_{n \in \mathbb{N}}$ in which the number of edges is quadratic in the number of vertices). In [5, 6], Borgs et al. extended this point of view, and proved that L^p -graphons (when $1 < p < \infty$) are limit objects of sparse graph sequences. The mode of convergence in dense/sparse graph limit theory is easily captured using the *cut-norm*. This norm was

introduced in [21] for matrices, and can be extended to graphons in a natural manner:

$$\|w\|_{\square} = \sup_{A, B \subseteq [0,1]} \left| \iint_{A \times B} w(x, y) dx dy \right|,$$

where the supremum is taken over all Lebesgue measurable subsets A, B .

Graph limit theory provides an effective approach for capturing the common large-scale features of networks. This viewpoint is now the base for many network analysis methods. In the limit theory of graphs, graphons represent random processes, called w -random graphs, that generate networks. Networks produced by the same graphon have similar large-scale features. For this reason, we think of graphons as the large-scale blueprint of any graph that they generate. Naturally, to study networks, we only need to study the (underlying) graphon, rather than the individual networks generated by that graphon. We refer to [41] for a comprehensive account of (dense) graph limit theory. To see some applications of graphons in network analysis, we refer to [3, 22, 23, 24, 27, 28, 44].

Robinson L^p -graphons are defined similarly to Robinson matrices (see Subsection 2.2). In this paper, we study the graphon recovery problem for the class of Robinson graphons. The limiting Robinson graphon of a converging graph sequence, if recovered successfully in terms of cut-norm, can be used to seriate the graphs in that sequence (see for example [32] for a randomized algorithm or [47] for a spectral approach). For this reason, we must measure proximity in terms of cut-norm when performing the Robinson graphon recovery. Our objective, therefore, is to obtain a statement of the following general form:

Given a graphon w that is locally almost Robinson, there exists a Robinson graphon u such that $\|w - u\|_{\square}$ is sufficiently small.

To begin, it is essential to formalize the concept of locally almost Robinson graphons. In this paper, we introduce a graphon function denoted as Λ , which serves as a gauge for assessing the Robinson property by quantifying localized (aggregated) deviations from it (Definition 4). In informal terms, we consider a graphon w to be locally almost Robinson if $\Lambda(w)$ is close to 0. Indeed, we prove that Λ serves as a suitable measurement for the Robinson property when $p > 5$ (see Proposition 5 and Theorem 19); that is, we show that Λ is subadditive and satisfies the following three properties:

- **(Recognition)** $\Lambda(w) = 0$ if and only if w is Robinson.
- **(Continuity)** Λ is continuous with respect to the cut-norm.
- **(Recovery)** Given an L^p -graphon w , there exists a Robinson graphon u such that $\|w - u\|_{\square} \leq c_p \Lambda(w)^{\alpha_p}$, where $\alpha_p \in (0, 1)$ and $c_p > 0$ are constants depending only on p .

The final item in the list presented above may be interpreted as follows: if w is locally almost Robinson (i.e., $\Lambda(w)$ is near 0), then a Robinson graphon u can be found so that $\|w - u\|_{\square}$ is also small. We think of u as a *Robinson approximation* for w . Our results

(in particular, Definition 7 and Theorem 19) provide a systematic approach for creating Robinson approximations for graphons that are almost Robinson on a local scale.

The current work extends the results of [14, 26], where a similar problem was addressed within the context of L^∞ -graphons. Specifically, [14] introduced a function Γ defined on the space of L^∞ -graphons. Much like Λ , this function Γ adheres to three crucial properties: (1) recognition: $\Gamma(w) = 0$ precisely when w is Robinson, (2) continuity: when equipped with the cut-norm, the function Γ exhibits continuity over the space of L^∞ -graphons, and (3) recovery: for every L^∞ -graphon w , there exists a Robinson L^∞ -graphon u such that $\|w - u\|_{\square} \leq 14\Lambda(w)^{1/7}$. The approach in the current paper offers a two-fold advantage. Firstly, our newly proposed function Λ possesses the desired properties of recognition, continuity, and recovery for all L^p -graphons with $p > 5$, extending beyond the restricted case of L^∞ -graphons. Secondly, the formula defining the function Λ is less intricate compared to Γ . Leveraging the simplicity of Λ and applying meticulous estimates, we not only extend the results achieved by Γ to encompass L^p -graphons (as detailed in Theorem 19), but also notably improve the outcomes within the domain of L^∞ -graphons (as demonstrated in Corollary 17).

It is worth noting that the proof of continuity and recovery of Γ in [14, 26] cannot be adjusted to work for the setting of L^p -graphons. Indeed, the proof of the $\|\cdot\|_{\square}$ -continuity of Γ in [14] relies upon the continuity of the triangular cut operator when acting on the space of $[-1, 1]$ -valued graphons [14, Lemma 6.1]. This continuity, however, does not hold when the triangular cut is applied to L^p -graphons with $p < \infty$ (see [46]). As a result, the proof in [14] do not work for the unbounded case. Similarly, the proof of stability of Γ in [26, Theorem 3.2] heavily relies on the fact that the graphons at hand do not attain values outside of $[-1, 1]$ (see Remark 3 for more details). Thus, to handle L^p -graphons, the new approach presented here seemed to be necessary.

This paper is organized as follows: In Section 2, we introduce various necessary background involving graphons and the study of the Robinson property. Section 3 discusses previous work that has been done to recover Robinson graphons in the cut-norm; here, we also comparatively summarize our own graphon recovery results. In Section 4, alongside the building of necessary machinery and statement of technical lemmas, we present proofs of our main recovery results. The aforementioned technical lemmas are proven in Appendix 4.3.

2 Notation and background

In this section of the paper, we present necessary background and notation for the reader. We begin with basic notation used throughout the paper; afterwards, we present graph limit theory, stating required definitions alongside notable results, then transition into explanation of the Robinson property for both matrices and graphons.

All functions and sets discussed in this paper are assumed to be measurable. The symmetric difference between two sets A and B is denoted $A\Delta B$, and the measure of A is denoted $|A|$. The notation $\|\cdot\|_p$ is used to represent the standard p -norm; that is, for

a function $f : X \rightarrow \mathbb{R}$ on a measure space X , and $1 \leq p < \infty$,

$$\|f\|_p := \left(\int_X |f|^p \right)^{\frac{1}{p}} \quad \text{and} \quad \|f\|_\infty := \inf_{a \in \mathbb{R}} \{ |f(x)| \leq a \text{ almost everywhere (a.e.)} \}.$$

2.1 Graphons and the cut-norm

We begin by defining the graphon, the fundamental building block of graph limit theory, in several of its myriad forms.

Definition 1 (Graphon space).

- Let \mathcal{W}_0 be the set of all symmetric, measurable functions $w : [0, 1]^2 \rightarrow [0, 1]$. The elements of this set are called *graphons*.
- Let \mathcal{W}^∞ be the set of all bounded, symmetric, measurable functions $w : [0, 1]^2 \rightarrow \mathbb{R}$. The elements of this set are called *kernels*.
- For $p \geq 1$, let \mathcal{W}^p be the set of all symmetric, measurable functions $w : [0, 1]^2 \rightarrow \mathbb{R}$ such that $\|w\|_p < \infty$. The elements of this set are called *L^p -graphons*.

It is clear that $\mathcal{W}_0 \subseteq \mathcal{W}^\infty \subseteq \mathcal{W}^p \subseteq \mathcal{W}^1$ for $p > 1$. We mention as well the space of *step kernels* $\mathcal{S} \subseteq \mathcal{W}^\infty$, which comprises the set of symmetric step functions on $[0, 1]^2$; that is, for $w \in \mathcal{S}$, there exists some partition $\mathcal{P} = \{P_1, \dots, P_m\}$ of $[0, 1]$ into measurable subsets such that w is constant on $P_i \times P_j$ for all $1 \leq i, j \leq m$. We define the space of step graphons $\mathcal{S}_0 \subseteq \mathcal{W}_0$ similarly. The norm of choice for these function spaces is the *cut-norm*, which was introduced in [21] for matrices, and can be naturally extended to the case of graphons as follows.

Definition 2 (Cut-norm). Let $w \in \mathcal{W}^1$. We define the *cut-norm* by

$$\|w\|_\square = \sup_{S, T \subseteq [0, 1]} \left| \iint_{S \times T} w(x, y) \, dx dy \right|$$

where the supremum is taken over all measurable subsets S and T . Moreover, the supremum in the formula for cut-norm is achieved [41, Lemma 8.10].

It is obvious that for any function $w \in \mathcal{W}^1$, and for all $p \geq 1$,

$$\|w\|_\square \leq \|w\|_1 \leq \|w\|_p \leq \|w\|_\infty.$$

Similar to unlabeled graphs, we are interested in ‘unlabeled graphons’; these are defined through an analytic generalization of graph vertex relabelling. Let Φ denote the set of all measure-preserving bijections on $[0, 1]$; that is, if $A \subset [0, 1]$ and $\phi \in \Phi$, then $|\phi(A)| = |\phi^{-1}(A)| = |A|$. The *cut-distance* between two graphons u and w is

$$\delta_\square(u, w) = \inf_{\phi \in \Phi} \|u - w^\phi\|_\square,$$

where $w^\phi(x, y) = w(\phi(x), \phi(y))$. The cut-distance δ_\square is only a pseudometric; we thus identify graphons of cut distance 0 to get the space $\widetilde{\mathcal{W}}_0$ of *unlabeled graphons*. $\widetilde{\mathcal{W}}$ and $\widetilde{\mathcal{W}}^p$ are defined similarly. This has quite the reward: For $p > 1$, the metric space $(\widetilde{\mathcal{W}}^p, \delta_\square)$ has a compact unit ball ([6, Theorem 2.13]). For $p = 1$, additional conditions are required to ensure compactness ([6, Theorem C7]), though we do not consider this case in the paper.

Let $w \in \mathcal{W}^1$ and let $A, B \subseteq [0, 1]$. The *cell average* of w over $A \times B$ is given by

$$\bar{w}(A \times B) = \frac{1}{|A \times B|} \iint_{A \times B} w \, dx dy.$$

For $w \in \mathcal{W}^1$ and a partition $\mathcal{P} = (S_1, \dots, S_k)$ of $[0, 1]$ into measurable sets, we define the function $w_{\mathcal{P}}$ by

$$w_{\mathcal{P}}(s, t) = \frac{1}{|S_i \times S_j|} \iint_{S_i \times S_j} w(x, y) \, dx dy = \bar{w}(S_i \times S_j) \quad \text{if } (s, t) \in S_i \times S_j,$$

The operator $w \mapsto w_{\mathcal{P}}$ is called the *stepping operator*. It is easy to see that the stepping operator is contractive with respect to the cut-norm as well as L^p -norms ; indeed, for $w \in \mathcal{W}^1$, a partition \mathcal{P} of $[0, 1]$, and $p \geq 1$, we have

$$\|w_{\mathcal{P}}\|_p \leq \|w\|_p \quad \text{and} \quad \|w_{\mathcal{P}}\|_\square \leq \|w\|_\square.$$

2.2 The Robinson property

The Robinson property for matrices was first introduced in [50] for the study of the classical seriation problem, whose objective is to order a set of items so that similar items are placed close to one another. A symmetric matrix $A = [a_{ij}]$ is a *Robinson matrix* if

$$i \leq j \leq k \implies a_{ik} \leq \min\{a_{ij}, a_{jk}\}.$$

and is *Robinsonian* if it becomes a Robinson matrix after simultaneous application of a permutation π to its rows and columns. The permutation π is a *Robinson ordering* of A . If the entries a_{ij} of the symmetric matrix A represent similarity of items i and j , then the Robinson ordering of A represents a linear arrangement of the items so that similar items are placed closer together.

The Robinson property for graphons, initially introduced in [14], shares a similar definition. When analyzing large matrices or graphs, particularly in the context of network visualization, it is useful to determine whether these complex structures can be interpreted as samples of a Robinson graphon w . If so, one can harness the Robinson property inherent in the graphon w to obtain quantitative insights concerning the data at hand. Formally, an L^p -graphon $w \in \mathcal{W}^p$ is said to be *Robinson* if

$$x \leq y \leq z \implies w(x, z) \leq \min\{w(x, y), w(y, z)\}.$$

We call an L^p -graphon *Robinson almost everywhere*, or Robinson a.e. for short, if it is equal a.e. to a Robinson L^p -graphon. An L^p -graphon $w \in \mathcal{W}^p$ is called *Robinsonian* if there exists a Robinson L^p -graphon $u \in \mathcal{W}^p$ such that $\delta_\square(w, u) = 0$.

3 Robinson property: measurement and recovery

In order to study the recovery problem for Robinson matrices/graphons, we need to devise a graphon parameter that suitably measures the Robinson property. Such a parameter must *recognize* Robinson graphons, be *continuous* on the space of graphons, and *recover* the Robinson property. These three key features ensure that the graphon parameter recognizes graphs sampled from Robinson graphons as well. Informally speaking, these are graphs whose parameter value is small, and we think of them as ‘almost Robinson graphs’.

We give a brief overview of past results from [14] introducing such a parameter for L^∞ -graphons, before discussing efforts to tackle this problem in the world of L^p -graphons. The parameter $\Gamma : \mathcal{W}_0 \rightarrow [0, 1]$ was introduced in [14] to estimate to what extent a given graphon fails to satisfy the Robinson property. We recall the definition of Γ , noting that it can naturally extend to \mathcal{W}^1 . For $w \in \mathcal{W}^1$ and a measurable subset A of $[0, 1]$, we define $\Gamma(w, A)$ as

$$\begin{aligned} \Gamma(w, A) = & \iint_{y < z} \left[\int_{x \in A \cap [0, y]} (w(x, z) - w(x, y)) dx \right]_+ dy dz \\ & + \iint_{y < z} \left[\int_{x \in A \cap [z, 1]} (w(x, y) - w(x, z)) dx \right]_+ dy dz, \end{aligned}$$

where $[x]_+ = \max(x, 0)$. We then define $\Gamma(w) := \sup_A \Gamma(w, A)$, where the supremum is taken over all measurable subsets of $[0, 1]$. It turns out that Γ suitably measures the Robinson property for L^∞ -graphons:

- (i) (**Recognition** [14, Proposition 4.2]) $w \in \mathcal{W}_0$ is Robinson a.e. if and only if $\Gamma(w) = 0$.
- (ii) (**Continuity** [14, Lemma 6.2]) Γ is continuous on \mathcal{W}_0 with respect to cut-norm.
- (iii) (**Recovery** [26, Theorem 3.2]) For every $w \in \mathcal{W}_0$, there exists a Robinson graphon $u \in \mathcal{W}_0$ satisfying

$$\|u - w\|_\square \leq 14\Gamma(w)^{1/7}.$$

Indeed, Γ recognizes samples of Robinson graphons: These are precisely graph sequences whose Γ -values converge to 0 [26, Theorem 1.2].

Remark 3. It is natural to ask whether these results for Γ can be extended to L^p -graphons; however, in trying to do so, one encounters key issues in proving continuity (ii) and recovery (iii) of Γ . Firstly, if χ denotes the characteristic function of the triangle above the diagonal, then the proof of (ii) is based on the following two facts:

- (a) $|\Gamma(w_1) - \Gamma(w_2)| \leq 2\|w_1 - w_2\|_\square + 2\|(w_1 - w_2)\chi\|_\square$ for $w_1, w_2 \in \mathcal{W}^1$.
- (b) $\|w\chi\|_\square \leq 2\sqrt{\|w\|_\square \|w\|_\infty}$ for $w \in \mathcal{W}^\infty$.

While condition (a) holds in the sparse case, it was shown in [46] that condition (b) cannot be generalized to $w \in \mathcal{W}^p$ with $1 < p < \infty$. In fact, the map $M_\chi : w \mapsto w\chi$ is not $\|\cdot\|_\square$ -continuous on \mathcal{W}^p for $1 < p < \infty$ [46, Corollary 3.4]. So even though the definition of Γ can be extended to L^p -graphons, the proof of continuity in [14] does not naturally extend to the unbounded case. Whether Γ is cut-norm continuous on \mathcal{W}^p remains an open question.

Secondly, the proof of stability of Γ in [27] relies heavily, and at multiple points, on the universal L^∞ -bound on graphons. Most importantly, the key idea of the proof is to locate a large enough cell on which the graphon has a large enough average ([27, Claim 4.11]), assuming that the desired upper bound on the error of estimate fails. The existence of such a cell, which is an integral part of the proof, is not guaranteed when the graphon is not L^∞ -bounded. Indeed, for any $1 \leq p < \infty$, it is easy to construct a graphon w with $\|w\|_p = 1$ for which such a cell does not exist.

4 Robinson recovery of L^p -graphons

In this section, we introduce a new graphon parameter Λ for measuring the Robinson property. We say that $A \leq B$ for sets $A, B \subseteq [0, 1]$ if $a \leq b$ for all $a \in A$ and $b \in B$.

Definition 4 (Robinson parameter). Let $w \in \mathcal{W}^1$. Define

$$\begin{aligned} \Lambda(w) &= \frac{1}{2} \sup_{\substack{A \leq B \leq C, \\ |A|=|B|=|C|}} \left[\iint_{A \times C} w \, dx dy - \iint_{B \times C} w \, dx dy \right] \\ &+ \frac{1}{2} \sup_{\substack{X \leq Y \leq Z, \\ |X|=|Y|=|Z|}} \left[\iint_{X \times Z} w \, dx dy - \iint_{X \times Y} w \, dx dy \right], \end{aligned}$$

where A, B, C and X, Y, Z are measurable subsets of $[0, 1]$.

It is clear that we have $\Lambda(w) \geq 0$ for all $w \in \mathcal{W}^1$, as we can take $A = B = C = X = Y = Z = \emptyset$. Moreover,

$$\Lambda(w) \leq \frac{1}{2} \left(\iint_{(A \cup B) \times C} |w| \, dx dy + \iint_{X \times (Y \cup Z)} |w| \, dx dy \right),$$

implying that $\Lambda(w) \leq \|w\|_p$ for every $p \geq 1$. Note that since the supremum is subadditive, we also have that Λ is subadditive, i.e., for all $u, w \in \mathcal{W}^p$

$$\Lambda(u + w) \leq \Lambda(u) + \Lambda(w). \tag{1}$$

The rest of this section is dedicated to proving that Λ suitably measures the Robinson property for L^p -graphons. We thus show that Λ recognizes the Robinson property, is continuous on L^p -graphons, and can recover the Robinson property. We begin with the proof of recognition, noting that for ease of notation, the upper triangle of the unit square will be written

$$\Delta = \{(x, y) \in [0, 1]^2 : x \leq y\}.$$

Proposition 5 (Recognition). *Let $1 \leq p \leq \infty$ and suppose $w \in \mathcal{W}^p$. Then w is Robinson a.e. if and only if $\Lambda(w) = 0$.*

Proof. Observe that if w is Robinson a.e., then $\Lambda(w) = 0$. To prove the reverse direction, suppose $\Lambda(w) = 0$. For $n \in \mathbb{N}$, let $w_n = w_{\mathcal{P}_n}$, where \mathcal{P}_n is the partition of $[0, 1]$ into n equal-size intervals $I_1 \leq I_2 \leq \dots \leq I_n$. Note that as $\Lambda(w) = 0$, for any choice of measurable subsets $A \leq B \leq C$ of $[0, 1]$ of equal size, we have

$$\iint_{A \times C} w(x, y) \, dx dy \leq \iint_{B \times C} w(x, y) \, dx dy \ \& \ \iint_{A \times C} w(x, y) \, dx dy \leq \iint_{A \times B} w(x, y) \, dx dy.$$

Applying the above inequalities to the sets I_i , we observe that w_n is Robinson. Since $w_n \rightarrow w$ in the L^1 norm, by using the Borel–Cantelli Lemma and going down to a subsequence if necessary, we can assume that $\{w_n\}$ converges to w pointwise a.e. in $[0, 1]^2$. So, there exists a set of measure zero $N \subset [0, 1]^2$ such that $w(x, y) = \lim_n w_n(x, y)$ for $(x, y) \in [0, 1]^2 \setminus N$. As w is symmetric, we can assume that N is symmetric with respect to the main diagonal. Next, we define the symmetric function \tilde{w} as follows:

$$\tilde{w}(x, y) = \begin{cases} w(x, y) & (x, y) \in \Delta \setminus N \\ \sup \{w(u, v) : (u, v) \in ([0, x] \times [y, 1]) \setminus N\} & (x, y) \in \Delta \cap N \end{cases}.$$

Clearly $w = \tilde{w}$ almost everywhere. To show \tilde{w} is Robinson, we observe that for every $(x_1, y_1), (x_2, y_2) \in \Delta$ with $x_2 \leq x_1$ and $y_1 \leq y_2$, we have $\tilde{w}(x_2, y_2) \leq \tilde{w}(x_1, y_1)$. Indeed, the required inequality can be easily verified by considering four cases depending on whether each of (x_1, y_1) and (x_2, y_2) belong to $\Delta \setminus N$ or $\Delta \cap N$; we omit the details of this straightforward verification. \square

We now show that Λ is continuous on L^p -graphons.

Proposition 6 (Continuity). $|\Lambda(w) - \Lambda(u)| \leq 2\|w - u\|_{\square}$.

Proof. Let $w, u \in \mathcal{W}^1$, and fix $\epsilon > 0$. There exist measurable sets $A \leq B \leq C$ with equal size and measurable sets $X \leq Y \leq Z$ with equal size such that

$$\frac{1}{2} \left(\iint_{A \times C} w \, dx dy - \iint_{B \times C} w \, dx dy + \iint_{X \times Z} w \, dx dy - \iint_{X \times Y} w \, dx dy \right) \geq \Lambda(w) - \epsilon. \quad (2)$$

Combining (2) with the definition of $\Lambda(u)$, we get that

$$\begin{aligned} 2(\Lambda(w) - \Lambda(u) - \epsilon) &\leq \iint_{A \times C} w - u + \iint_{B \times C} u - w + \iint_{X \times Z} w - u + \iint_{X \times Y} u - w \\ &\leq 4\|w - u\|_{\square}. \end{aligned}$$

Following the same logic, we can also prove that $\Lambda(u) - \Lambda(w) \leq 2\|w - u\|_{\square} + 2\epsilon$. Taking $\epsilon \rightarrow 0$, we get $|\Lambda(w) - \Lambda(u)| \leq 2\|w - u\|_{\square}$. \square

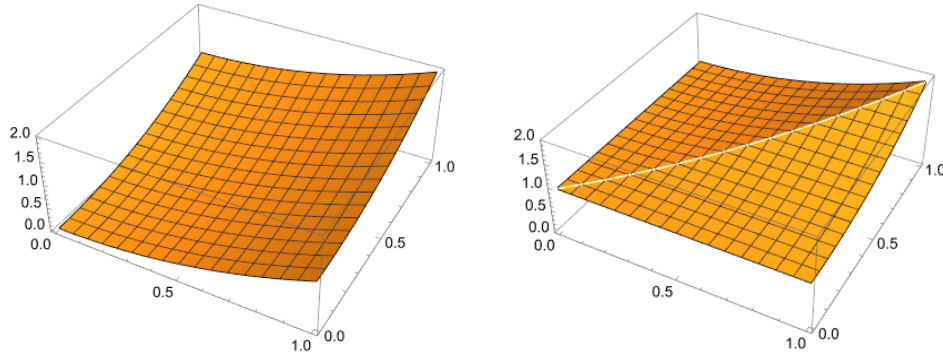


Figure 1: The graphon $w(x, y) = x^2 + y^2$ (left) and its Robinson approximation R_w^α for $\alpha = 0.05$ (right).

Having shown that Λ recognizes the Robinson property and is continuous on L^p -graphons, all that remains is to show it recovers the Robinson property; that is, we endeavour to show that for a given L^p -graphon w , there exists some Robinson L^p -graphon u such that $\|w - u\|_\square$ is bounded above by some power of $\Lambda(w)$. This is a far more challenging task than the previous two properties and must be handled in several parts. We begin by first defining this Robinson graphon u and listing some properties that will be useful in future proofs.

4.1 Robinson approximation of L^p -graphons

Here we introduce the α -Robinson approximation of an L^p -graphon. We borrow the following terminology from [26]. The *upper left* (UL) and *lower right* (LR) regions of a given point $(a, b) \in \Delta$ are given by

$$\begin{aligned} UL(a, b) &= [0, a] \times [b, 1], \\ LR(a, b) &= [a, b] \times [a, b] \cap \Delta. \end{aligned}$$

Definition 7 (α -Robinson approximation for graphons). Let $p \geq 1$ and fix a parameter $0 < \alpha < 1$. Given $w \in \mathcal{W}^p$, the α -Robinson approximation R_w^α of w is the L^p -graphon such that for all $(x, y) \in \Delta$,

$$R_w^\alpha(x, y) = \sup \{ \bar{w}(S \times T) : S \times T \subseteq UL(x, y), |S| = |T| = \alpha \}, \quad (3)$$

where S, T are measurable and $\sup \emptyset = 0$. Moreover, we set $R_w^\alpha = w$ if $\alpha = 0$ and w is Robinson.

For every $w \in \mathcal{W}^p$, the approximation R_w^α is a Robinson graphon. To see this, note that for two points (x_1, y_1) and (x_2, y_2) , if $(x_1, y_1) \in UL(x_2, y_2)$, then $UL(x_1, y_1) \subseteq UL(x_2, y_2)$. So by definition of R_w^α , we get $R_w^\alpha(x_1, y_1) \leq R_w^\alpha(x_2, y_2)$. Therefore, R_w^α satisfies the Robinson property (Subsection 2.2), as for $x \leq y \leq z$, we have $(x, z) \in UL(x, y) \cap UL(y, z)$. See Figure 1 for an example of a graphon w and its Robinson approximation R_w^α .

Remark 8. If w is a Robinson graphon and $0 < \alpha < 1$, then the graphon R_w^α can be different from w , although both graphons are Robinson. This is indeed the case for Robinson step graphons. In general, w does not need to be continuous, but R_w^α satisfies certain continuity properties: Though R_w^α is certainly not always continuous on $[0, 1]^2$ —the outer boundary of thickness α of the unit square is set to 0 by definition—it is continuous on $[\alpha, 1 - \alpha]^2$. We prove this fact and some other useful properties of R_w^α in Proposition 9.

Proposition 9 (Properties of the Robinson approximation). *Let $p \geq 1$, and fix a parameter $0 < \alpha < 1$. Given graphons $w, u \in \mathcal{W}^p$, we have:*

- (i) *If $u \leq w$ pointwise, then $0 \leq R_w^\alpha - R_u^\alpha \leq R_{w-u}^\alpha$.*
- (ii) *R_w^α is continuous on the square $[\alpha, 1 - \alpha]^2$, and $\|R_w^\alpha\|_\infty \leq \alpha^{-\frac{2}{p}} \|w\|_p$.*
- (iii) *For $w, u \in \mathcal{W}^p$, we have $\|R_w^\alpha - R_u^\alpha\|_\infty \leq \|w - u\|_\square$. In particular, $\|R_w^\alpha - R_u^\alpha\|_\square \leq \|w - u\|_\square$.*

We note that for item (iii), only convergence in the L^1 -norm—the weakest L^p -norm—is required. Therefore, convergence in any L^p -norm is enough to guarantee the convergence of the α -Robinson approximations in the cut-norm.

Proof. To prove (i), let $(x, y) \in \Delta$. By definition of the α -Robinson approximation, for every $\epsilon > 0$, there exist sets $A, B \subseteq \text{UL}(x, y)$ such that $R_w^\alpha(x, y) \geq \bar{w}(A \times B) \geq R_w^\alpha(x, y) - \epsilon$ and $C, D \subseteq \text{UL}(x, y)$ such that $R_u^\alpha(x, y) \geq \bar{u}(C \times D) \geq R_u^\alpha(x, y) - \epsilon$. We also note that $|A| = |B| = |C| = |D| = \alpha$. Since $u \leq w$, we have

$$0 \leq \frac{1}{\alpha^2} \iint_{C \times D} (w - u) dx dy = \bar{w}(C \times D) - \bar{u}(C \times D) \leq R_w^\alpha(x, y) - R_u^\alpha(x, y) + \epsilon$$

Letting $\epsilon \rightarrow 0$, we conclude that $0 \leq R_w^\alpha - R_u^\alpha$. Moreover, we note that

$$R_w^\alpha(x, y) - \epsilon - R_u^\alpha(x, y) \leq \frac{1}{\alpha^2} \iint_{A \times B} (w - u) dx dy \leq R_{w-u}^\alpha(x, y),$$

showing that (i) holds true, again by letting $\epsilon \rightarrow 0$.

To prove (ii), let $(x, y) \in [\alpha, 1 - \alpha]^2$, and suppose the sequence $\{(x_n, y_n)\} \subset [\alpha, 1 - \alpha]^2$ converges to (x, y) in the standard Euclidean distance. Then, we have

$$\text{UL}(\min(x_n, x), \max(y_n, y)) \subseteq \text{UL}(x, y) \subseteq \text{UL}(\max(x_n, x), \min(y_n, y)),$$

which implies that

$$R_w^\alpha(\min(x_n, x), \max(y_n, y)) \leq R_w^\alpha(x_n, y_n) \leq R_w^\alpha(\max(x_n, x), \min(y_n, y)). \quad (4)$$

We shall show that both the upper and lower bound in (4) converge to $R_w^\alpha(x, y)$; by the Squeeze Theorem, this will imply our original claim. To do so, we let $\epsilon > 0$ and recall from elementary measure theory that as w is an L^1 function, for every $\epsilon > 0$, there exists

a $\delta > 0$ such that for every $S \subset [\alpha, 1 - \alpha]^2$ satisfying $|S| < \delta$, we have $\iint_S |w| dx dy < \epsilon$. Let $N \in \mathbb{N}$ be such that for all $n \geq N$, we have that $d((x_n, y_n), (x, y)) < \frac{\delta}{2}$. Defining the set $S := \text{UL}(\min(x_n, x), \max(y_n, y)) \Delta \text{UL}(x, y)$, it must then be that

$$\begin{aligned} R_w^\alpha(x, y) &\leq R_w^\alpha(\min(x_n, x), \max(y_n, y)) + \frac{1}{\alpha^2} \iint_S |w| dx dy \\ &\leq R_w^\alpha(\min(x_n, x), \max(y_n, y)) + \frac{\epsilon}{\alpha^2}. \end{aligned}$$

Note that by definition of R_w^α , we have $R_w^\alpha(\min(x_n, x), \max(y_n, y)) \leq R_w^\alpha(x, y)$, so

$$|R_w^\alpha(\min(x_n, x), \max(y_n, y)) - R_w^\alpha(x, y)| < \epsilon.$$

Let $\epsilon \rightarrow 0$, we conclude that the left hand side of (4) converges to $R_w^\alpha(x, y)$. A similar argument shows that the the right hand side of (4) converges to $R_w^\alpha(x, y)$ as well, and we are done.

To prove the norm bound, we note that for every $\epsilon > 0$, there exist sets $A, B \subseteq \text{UL}(x, y)$ where $|A| = |B| = \alpha$ such that $R_w^\alpha(x, y) \geq \bar{w}(A \times B) \geq R_w^\alpha(x, y) - \epsilon$. Then, by Hölder's inequality,

$$R_w^\alpha(x, y) - \epsilon \leq \frac{1}{\alpha^2} \iint_{A \times B} w dx dy \leq \frac{1}{\alpha^2} \|w\|_p \|\mathbb{1}_{A \times B}\|_q = \alpha^{-\frac{2}{p}} \|w\|_p,$$

where we used $\|\mathbb{1}_{A \times B}\|_q = \alpha^{2/q} = \alpha^{2-2/p}$. We let $\epsilon \rightarrow 0$ to finish the proof of (ii).

To prove (iii), fix $(x, y) \in \Delta$, and let $\epsilon > 0$ be arbitrary. By an argument similar to (i), there exist sets A, B of size α such that $A \times B \subseteq \text{UL}(x, y)$, and

$$R_w^\alpha(x, y) - R_u^\alpha(x, y) \leq \bar{w}(A \times B) - \bar{u}(A \times B) + \epsilon \leq \alpha^2 \|w - u\|_\square + \epsilon.$$

Sending ϵ to 0, we get $R_w^\alpha(x, y) - R_u^\alpha(x, y) \leq \alpha^2 \|w - u\|_\square$. A similar argument can be used to show $R_u^\alpha(x, y) - R_w^\alpha(x, y) \leq \alpha^2 \|w - u\|_\square$, so we are done. \square

4.2 Recovering the Robinson property for L^p -graphons

Our proof that Λ recovers the Robinson property requires that the value of R_w^α and \bar{w} both be tightly controlled within sets of specific measure. To accomplish this, we enhance the techniques outlined in [26], and as a result, we need to introduce analogous notation. Let $w \in \mathcal{W}^\infty$ be a graphon with $w \geq 0$ and $\lceil \|w\|_\infty \rceil = M$. Fix a parameter $\alpha \in (0, 1)$ (as in Definition 7), and let m be a fixed integer. For $k \in \{1, \dots, mM - 1\}$, define the k -th *black region* \mathcal{B}_k , the k -th *white region* \mathcal{V}_k and the k -th *grey region* \mathcal{G}_k as follows.

$$\begin{aligned} \mathcal{B}_k &= \left\{ (x, y) \in \Delta : x = y \text{ or } \exists S \times T \subseteq \text{UL}(x, y) \text{ s.t. } |S| = |T| = \alpha \ \& \ \bar{w}(S \times T) > \frac{k}{m} \right\}, \\ \mathcal{V}_k &= \left\{ (x, y) \in \Delta \setminus \mathcal{B}_k : \exists S \times T \subseteq \text{LR}(x, y) \text{ s.t. } |S| = |T| = \alpha \ \& \ \bar{w}(S \times T) \leq \frac{k}{m} \right\}, \\ \mathcal{G}_k &= \Delta \setminus (\mathcal{B}_k \cup \mathcal{V}_k). \end{aligned}$$

We set $\mathcal{B}_0 = \mathcal{V}_{mM} = \Delta$, set $\mathcal{V}_0 = \mathcal{B}_{mM} = \emptyset$, and denote $\mathcal{G} := \bigcup_{k=1}^{mM-1} \mathcal{G}_k$. In addition, we also define the following regions:

$$\mathcal{R}_k := \mathcal{B}_k \cap \mathcal{V}_{k+1}.$$

Black regions provide lower bounds on both R_w^α and \bar{w} while white regions provide upper bounds. So the regions \mathcal{R}_k are used to form tight bounds on the behaviour of both R_w^α and \bar{w} . Grey regions provide no information pertinent to our method of proof. From the definition of the regions, it is easy to observe that

$$\mathcal{B}_0 \supseteq \mathcal{B}_1 \supseteq \cdots \supseteq \mathcal{B}_{mM} \quad \text{and} \quad \mathcal{V}_0 \subseteq \mathcal{V}_1 \subseteq \cdots \subseteq \mathcal{V}_{mM}.$$

Subsequently, we establish that the regions \mathcal{R}_k form a partition of $\Delta \setminus \mathcal{G}$. Our approach mirrors that of [26, Lemma 4.8], with a supplementary sketch of the proof provided for reader's convenience.

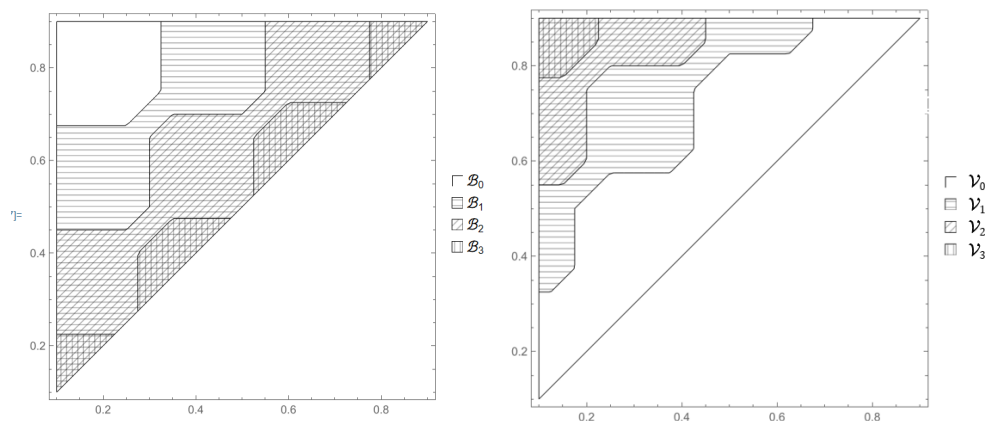


Figure 2: An example of black and white regions for $m = 4$ and $\alpha = 0.1$.

Lemma 10. *Let $w \in \mathcal{W}^\infty$ be such that $w \geq 0$ and $\lceil \|w\|_\infty \rceil = M$, and let \mathcal{B}_k , \mathcal{V}_k , \mathcal{G}_k , and \mathcal{R}_k be as defined above. Then,*

$$\Delta \setminus \left(\bigcup_{k=1}^{mM-1} \mathcal{G}_k \right) = \bigsqcup_{k=0}^{mM-1} \mathcal{R}_k,$$

where \bigsqcup denotes the disjoint union of sets.

Proof. Let $1 \leq i < j \leq mM - 1$, and note that $\mathcal{R}_i \subseteq \mathcal{V}_{i+1} \subseteq \mathcal{V}_j$ and $\mathcal{R}_j \subseteq \mathcal{B}_j$. Since $\mathcal{V}_j \cap \mathcal{B}_j = \emptyset$, we conclude that $\mathcal{R}_i \cap \mathcal{R}_j = \emptyset$ as well. So the regions \mathcal{R}_k are disjoint. Next, observe that

$$\Delta \setminus \left(\bigcup_{k=1}^{mM-1} \mathcal{G}_k \right) = \Delta \setminus \left(\bigcup_{k=1}^{mM-1} \Delta \setminus (\mathcal{B}_k \cup \mathcal{V}_k) \right) = \bigcap_{k=1}^{mM-1} (\mathcal{B}_k \cup \mathcal{V}_k).$$

Now, we consider the expansion of $(\mathcal{B}_1 \cup \mathcal{V}_1) \cap (\mathcal{B}_2 \cup \mathcal{V}_2) \cap \cdots \cap (\mathcal{B}_{mM-1} \cup \mathcal{V}_{mM-1})$ into expressions $X_1 \cap \cdots \cap X_{mM-1}$ with $X_i \in \{\mathcal{B}_i, \mathcal{V}_i\}$. We further note that $X_1 \cap \cdots \cap X_{mM-1} = \emptyset$ whenever $X_i = \mathcal{V}_i$ and $X_j = \mathcal{B}_j$ for some $i < j$; thus, every nonempty term $X_1 \cap \cdots \cap X_{mM-1}$ from the above expansion must be of one of the following forms:

- (i) $X_1 \cap \cdots \cap X_{mM-1} = \mathcal{B}_j \cap \mathcal{V}_{j+1} = \mathcal{R}_j$ with $1 \leq j < mM - 1$ if there is at least one black and one white region amongst the X_i .
- (ii) $X_1 \cap \cdots \cap X_{mM-1} = \mathcal{V}_1 \cap \cdots \cap \mathcal{V}_{mM-1} = \mathcal{V}_1$ if all X_i are white.
- (iii) $X_1 \cap \cdots \cap X_{mM-1} = \mathcal{B}_1 \cap \cdots \cap \mathcal{B}_{mM-1} = \mathcal{B}_{mM-1}$ if all X_i are black.

This completes the proof, as $\mathcal{V}_1 = \mathcal{V}_1 \cap \mathcal{B}_0 = \mathcal{R}_0$ and also $\mathcal{B}_{mM-1} = \mathcal{B}_{mM-1} \cap \mathcal{V}_{mM} = \mathcal{R}_{mM-1}$. \square

By definition, for every $k \in \{1, \dots, mM - 1\}$, the regions \mathcal{B}_k and \mathcal{V}_k have *upper* and *lower boundary* functions $f_k, g_k : [0, 1] \rightarrow [0, 1]$, where f_k is the upper boundary of \mathcal{B}_k and g_k is the lower boundary of \mathcal{V}_k . These are defined in [26] as follows:

$$\begin{aligned} f_k(x) &= \sup\{z \in [x, 1] : (x, z) \in \mathcal{B}_k\}, \\ g_k(x) &= \inf\{z \in [x, 1] : (x, z) \in \mathcal{V}_k\}; \end{aligned}$$

with the convention $\inf \emptyset = 1$. Additionally, we define $f_0(x) = 1$ and $g_{mM}(x) = x$ for all $x \in [0, 1]$ to represent the corresponding boundaries for $\mathcal{B}_0 = \mathcal{V}_{mM} = \Delta$. Finally, since f_k and g_{k+1} are the upper and lower boundaries of \mathcal{B}_k and \mathcal{V}_{k+1} respectively, if the region \mathcal{R}_k is nonempty, then it is bounded from below by g_{k+1} and from above by f_k . We refer to Figure 2 for a visual representation of the regions and boundary functions previously mentioned.

From the definition of the black and white regions, it is easy to see that the functions f_k and g_k are both increasing functions and thus only admit jump discontinuities. We naturally extend the graph of these functions to *boundary curves* by adding vertical line segments connecting any such discontinuities, denoting the resulting curves once again by f_k and g_k respectively.

Let $S, T \subseteq [0, 1]$ be measurable. We say that $S \times T$ *crosses* a boundary curve f_k or g_k if the top-left corner of the cell is strictly above the boundary curve and its bottom-right corner is strictly below the curve. This definition is used as $S \times T$ need not be a connected subset of \mathbb{R}^2 , so a boundary curve can go through the cell without intersecting it. Figure 3 depicts such behavior.

The proof of our recovery theorem is based on the idea that the total area of the grey regions \mathcal{G}_k is small. This allows us to concentrate on the behaviour of w and R_w^α inside the regions \mathcal{R}_k , where their values are strictly controlled. We show that their local average difference in these regions can be controlled by $\Lambda(w)$, leading to the conclusion that $\|w - R_w^\alpha\|_\square$ must be small. We introduce three lemmas here, each necessary in the proof of our main result. We will present the proofs of Lemma 12 and Lemma 13 in Appendix 4.3. The proof of Lemma 11 is very similar to [26, Lemma 4.10], but to keep this article self-contained, we present a sketch of this proof in Appendix 4.3 as well.

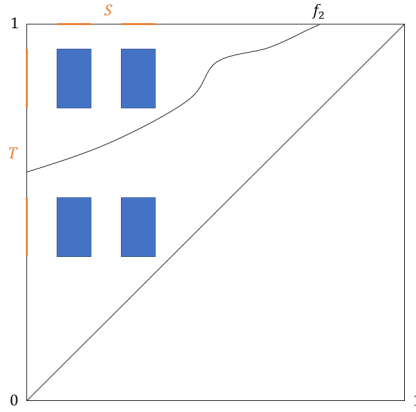


Figure 3: An example of a cell $S \times T$ that crosses a boundary curve f_2 without intersecting it.

Lemma 11. Let $k \in \mathbb{Z}^{\geq 0}$, $w \in \mathcal{W}^p$ with $w \geq 0$, and $\alpha \in (0, 1)$. Then, $\overline{\mathcal{G}_k}$ does not contain any $\beta \times \beta$ square, where $\beta > \alpha$. Here, $\overline{\mathcal{G}_k}$ denotes the closure of \mathcal{G}_k in the Euclidean topology of \mathbb{R}^2 .

Lemma 12. Let $u \in L^\infty[0, 1]$ (not necessarily non-negative), and $P \subseteq [0, 1]$ be a measurable subset such that $\int_P u \, dx \neq 0$. Let $0 < \beta < |P|$ be fixed. Then P can be partitioned into $N := \lceil |P|/\beta \rceil$ subsets P_1, \dots, P_N so that the following conditions are satisfied:

- (i) $P_1 \leq \dots \leq P_{N-1}$.
- (ii) $|P_i| = \beta$ for $1 \leq i \leq N - 1$ and $|P_N| \leq \beta$.
- (iii) $|\int_{P_N} u \, dx| \leq \frac{1}{N} |\int_P u \, dx|$.

Lemma 13. Let $f \in L^1[0, 1]^2$, and let $S, S' \subseteq [0, 1]$ be measurable subsets such that $|S| = |S'|$. Suppose for a constant $C > 0$ we have

$$\iint_{S \times S'} f \, dx dy \geq C.$$

Then, for every $\alpha \in (0, 1)$, there exist measurable sets $T \subset S$ and $T' \subset S'$ such that $|T| = |T'| = \alpha|S|$ and

$$\frac{1}{|T \times T'|} \iint_{T \times T'} f \, dx dy \geq \frac{C}{|S \times S'|}.$$

4.3 Proof of recovering property of Λ

Prefacing our main result, we begin first with a necessary proposition—as the proof of Theorem 19 heavily features “cutting” graphons off at certain values, we must control their behaviour once cut. The following proposition provides such control.

Proposition 14. Let $p > 2$ and $w \in \mathcal{W}^\infty$ with $w \geq 0$. Suppose $\|w\|_p \leq 1$. If R_w^α is the Robinson approximation of w with parameter $\alpha = \|w\|_\infty^{-\frac{p}{3p-2}} \Lambda(w)^{\frac{2p}{5p-2}}$, then

$$\|w - R_w^\alpha\|_\square \leq 4 \left(1 + (8\|w - R_w^\alpha\|_p + 2)\|w\|_\infty^{\frac{2p}{3p-2}} \right) \Lambda(w)^{\frac{p-2}{5p-2}}. \quad (5)$$

The proof of this proposition is inspired by [26, Theorem 3.2], however, the approximation result of [26] would not provide the upper bound needed for Proposition 14. As a corollary of this proposition, we improve upon the bound obtained in [26].

Proof. By [41, Lemma 8.10], there exist measurable $S, T \subseteq [0, 1]$ so that

$$\left| \iint_{S \times T} (w - R_w^\alpha) dx dy \right| = \|w - R_w^\alpha\|_\square.$$

Replacing $S \times T$ with $T \times S$ if necessary, we can assume without loss of generality that

$$\left| \iint_{(S \times T) \cap \Delta} (w - R_w^\alpha) dx dy \right| \geq \frac{1}{2} \|w - R_w^\alpha\|_\square. \quad (6)$$

Fix $\beta \in (\alpha, \frac{8}{7}\alpha)$. Next, we split S into $N_1 := \lceil |S|/\beta \rceil$ subsets S_1, S_2, \dots, S_{N_1} , and T into $N_2 := \lceil |T|/\beta \rceil$ subsets T_1, \dots, T_{N_2} , so that the following conditions are satisfied.

Conditions for S_i :

- (i) $S_1 \leq \dots \leq S_{N_1-1}$.
- (ii) $|S_i| = \beta$ for $1 \leq i \leq N_1 - 1$ and $|S_{N_1}| \leq \beta$.
- (iii) $\left| \iint_{S_{N_1} \times T} (w - R_w^\alpha) dx dy \right| \leq \frac{\|w - R_w^\alpha\|_\square}{N_1}$.

Conditions for T_j :

- (i) $T_1 \leq \dots \leq T_{N_2-1}$.
- (ii) $|T_j| = \beta$ for $1 \leq j \leq N_2 - 1$ and $|T_{N_2}| \leq \beta$.
- (iii) $\left| \iint_{S \times T_{N_2}} (w - R_w^\alpha) dx dy \right| \leq \frac{\|w - R_w^\alpha\|_\square}{N_2}$.

To prove that a partition of S satisfying the above three properties exists, we apply Lemma 12 to the function $u(\cdot) = \int_T (w - R_w^\alpha)(\cdot, y) dy$ and $P = S$; a similar proof shows that such a partition exists for T .

We will assume, without loss of generality, that $N_1, N_2 \geq 8$, as we shall show the proposition holds true if $N_1 \leq 7$ or $N_2 \leq 7$. Assume that $N_1 \leq 7$. Then, $|S| \leq 7\beta$, and so we get

$$\|w - R_w^\alpha\|_\square = \left| \iint (w - R_w^\alpha) \mathbb{1}_{S \times T} \right| \leq \|w - R_w^\alpha\|_\infty (7\beta) \leq \|w\|_\infty (8\alpha) = 8\|w\|_\infty^{-\frac{2p}{3p-2}} \Lambda(w)^{\frac{2p}{5p-2}}.$$

In the above inequalities, we used the fact that since $0 \leq w, R_w^\alpha$ and $\|R_w^\alpha\|_\infty \leq \|w\|_\infty$, we have $\|w - R_w^\alpha\|_\infty \leq \|w\|_\infty$. To verify that the statement of the proposition holds true in this case, we need to show that the right hand side of the previous inequality is bounded above by the right hand side of (5), which reduces to showing that

$$\Lambda(w)^{\frac{p+2}{5p-2}} \leq \|w\|_\infty^{\frac{2}{3p-2}}.$$

However, since $\Lambda(w) \leq 1$ and $p > 2$, we have $\Lambda(w)^{\frac{p+2}{5p-2}} \leq \Lambda(w)^{\frac{2}{3p-2}}$. This, together with the fact that $\Lambda(w) \leq \|w\|_\infty$, proves that Proposition 14 holds if $N_1 \leq 7$. An identical argument shows that the proposition also holds if $N_2 \leq 7$. From this point forward, we make the following assumption:

Assumption ()* $N_1, N_2 \geq 8$.

We now consider the value of $w - R_w^\alpha$ on the sets $S_i \times T_j$. Note that with repeated applications of the triangle inequality, we have

$$\sum_{\substack{1 \leq i < N_1, 1 \leq j < N_2 \\ (S_i \times T_j) \cap \Delta \neq \emptyset}} \left| \iint_{S_i \times T_j} w - R_w^\alpha \right| \geq \left| \iint_{(S \times T) \cap \Delta} w - R_w^\alpha \right| - \left| \iint_{(S_{N_1} \times T) \cup (S \times T_{N_2})} w - R_w^\alpha \right|,$$

which, together with Equation (6) and property (iii) of the partitions $\{S_i\}$ and $\{T_i\}$, implies that

$$\sum_{\substack{1 \leq i < N_1, 1 \leq j < N_2 \\ (S_i \times T_j) \cap \Delta \neq \emptyset}} \left| \iint_{S_i \times T_j} (w - R_w^\alpha) \right| \geq \left(\frac{1}{2} - \frac{1}{N_1} - \frac{1}{N_2} \right) \|w - R_w^\alpha\|_\square.$$

The above inequality, combined with Assumption (*), implies that

$$\sum_{\substack{1 \leq i < N_1, 1 \leq j < N_2 \\ (S_i \times T_j) \cap \Delta \neq \emptyset}} \left| \iint_{S_i \times T_j} (w - R_w^\alpha) \right| \geq \frac{1}{4} \|w - R_w^\alpha\|_\square. \quad (7)$$

Claim 15. *There exists a cell $S' \times T' \subseteq \Delta$ contained in a region $\mathcal{R}_k = \mathcal{B}_k \cap \mathcal{V}_{k+1}$ so that $|S'| = |T'| = \alpha$, and*

$$\left| \iint_{S' \times T'} (w - R_w^\alpha) dx dy \right| \geq \alpha^2 \left[\frac{1}{4} \|w - R_w^\alpha\|_\square - 2(2mM - 1) \|w - R_w^\alpha\|_p \alpha^{1 - \frac{2}{p}} \right].$$

We use the pigeonhole principle to prove the claim. By Lemma 10, $\Delta = \bigcup_{k=0}^{mM-1} \mathcal{R}_k \cup \bigcup_{k=1}^{mM-1} \mathcal{G}_k$, and each of the regions \mathcal{G}_k or \mathcal{R}_k is bounded by boundary curves in $\{f_k, g_l : 1 \leq k \leq mM - 1, 1 \leq l \leq mM\}$. Thus, if a cell $S_i \times T_j$ does not cross the graph of any of these boundary curves, then it must be entirely contained inside one closed region $\overline{\mathcal{R}_k}$

or $\overline{\mathcal{G}_k}$. Next, by Lemma 11, no grey regions $\overline{\mathcal{G}_k}$ can contain any $S_i \times T_j$ with $1 \leq i < N_1$ and $1 \leq j < N_2$, because $|S_i| = |T_j| = \beta$ for such i and j . Thus, these cells must either lie in a single region $\overline{\mathcal{R}_k}$ or cross a boundary curve. Let \mathcal{I} denote the collection of indices (i, j) with $i < N_1$ and $j < N_2$, for which the associated cells $S_i \times T_j$ do not lie in a single region $\overline{\mathcal{R}_k}$. From above,

$$\mathcal{I} = \left\{ (i, j) : \begin{array}{l} 1 \leq i < N_1 \\ 1 \leq j < N_2 \end{array}, \text{ and } \exists 1 \leq k \leq mM - 1 \text{ s.t. } (S_i \times T_j) \text{ crosses } f_k \text{ or } g_k \text{ or } g_{mM} \right\}.$$

Because the lower and upper boundaries f_k, g_k are increasing curves, each f_k and g_k cross at most $2/\beta$ cells from the grid. As there are $2mM - 1$ total f_k and g_k , we have

$$|\mathcal{I}| \leq \frac{2(2mM - 1)}{\beta}.$$

Using Hölder's inequality and that every cell indexed in \mathcal{I} is of size β^2 , we get

$$\sum_{\substack{(i, j) \in \mathcal{I} \\ (S_i \times T_j) \cap \Delta \neq \emptyset}} \left| \iint_{S_i \times T_j} (w - R_w^\alpha) dx dy \right| \leq \frac{2(2mM - 1)}{\beta} \|w - R_w^\alpha\|_p \beta^{2 - \frac{2}{p}}. \quad (8)$$

Putting inequalities (7) and (8) together, it must be true that

$$\sum_{\substack{1 \leq i < N_1, 1 \leq j < N_2 \\ (S_i \times T_j) \cap \Delta \neq \emptyset \\ (i, j) \notin \mathcal{I}}} \left| \iint_{S_i \times T_j} (w - R_w^\alpha) dx dy \right| \geq \frac{1}{4} \|w - R_w^\alpha\|_\square - 2(2mM - 1) \|w - R_w^\alpha\|_p \beta^{1 - \frac{2}{p}}.$$

Since there are at most $(\lfloor \beta^{-1} \rfloor)^2$ cells $S_i \times T_j$ of size $\beta \times \beta$ and $\beta^2 \leq (\lfloor \beta^{-1} \rfloor)^{-2}$, there must exist a cell $S_{i_0} \times T_{j_0} \subseteq \Delta$ so that $|S_{i_0}| = |T_{j_0}| = \beta$, $(i_0, j_0) \notin \mathcal{I}$, and

$$\left| \iint_{S_{i_0} \times T_{j_0}} (w - R_w^\alpha) dx dy \right| \geq \beta^2 \left[\frac{1}{4} \|w - R_w^\alpha\|_\square - 2(2mM - 1) \|w - R_w^\alpha\|_p \beta^{1 - \frac{2}{p}} \right].$$

So $S_{i_0} \times T_{j_0}$ lies entirely in $\overline{\mathcal{R}_k} = \overline{\mathcal{B}_k} \cap \overline{\mathcal{V}_{k+1}}$ for some $0 \leq k \leq mM - 1$. By Lemma 13, we can reduce $S_{i_0} \times T_{j_0}$ to an $\alpha \times \alpha$ cell, called $S' \times T'$, contained in $\mathcal{R}_k = \mathcal{B}_k \cap \mathcal{V}_{k+1}$ satisfying

$$\left| \iint_{S' \times T'} (w - R_w^\alpha) dx dy \right| \geq \alpha^2 \left[\frac{1}{4} \|w - R_w^\alpha\|_\square - 2(2mM - 1) \|w - R_w^\alpha\|_p \beta^{1 - \frac{2}{p}} \right].$$

This inequality holds for all $\beta > \alpha$; taking $\beta \rightarrow \alpha$ proves Claim 15.

The cell $S' \times T'$ is thus contained in some region \mathcal{R}_k . From the definition of R_w^α and the black and white regions, we observe that if $1 \leq k \leq mM - 1$, then $\frac{k}{m} < R_w^\alpha \leq \frac{k+1}{m}$ on $S' \times T'$, and if $k = 0$, then $0 \leq R_w^\alpha \leq \frac{1}{m}$ on $S' \times T'$.

Claim 16. Under the assumptions made so far,

$$\|w - R_w^\alpha\|_\square \leq 4 \left(\frac{2\Lambda(w)}{\alpha^2} + \frac{1}{m} + 2(2mM - 1)\|w - R_w^\alpha\|_p \alpha^{1-\frac{2}{p}} \right).$$

To prove the claim, we consider three cases:

Case 1: Assume that $\iint_{S' \times T'} (w - R_w^\alpha) dx dy > 0$ and $0 \leq k \leq mM - 2$.

Using Claim 15 and the fact that $|S' \times T'| = \alpha^2$, we have

$$\bar{w}(S' \times T') - \frac{k}{m} \geq \overline{w - R_w^\alpha}(S' \times T') \geq \frac{1}{4} \|w - R_w^\alpha\|_\square - 2(2mM - 1) \|w - R_w^\alpha\|_p \alpha^{1-\frac{2}{p}}. \quad (9)$$

Now let (x, y) be the lower right corner of $S' \times T'$. Then $(x, y) \in \mathcal{V}_{k+1}$, implying that $\text{LR}(x, y)$ contains a region $S_a \times T_b$ so that $|S_a| = |T_b| = \alpha$ and that $\bar{w}(S_a \times T_b) \leq \frac{k+1}{m}$. So inequality (9) combined with the definition of Λ implies that

$$\begin{aligned} \Lambda(w) &\geq \frac{\alpha^2}{2} \left(\bar{w}(S' \times T') - \bar{w}(S' \times T_b) + \bar{w}(S' \times T_b) - \bar{w}(S_a \times T_b) \right) \\ &\geq \frac{\alpha^2}{2} \left(\frac{1}{4} \|w - R_w^\alpha\|_\square - 2(2mM - 1) \|w - R_w^\alpha\|_p \alpha^{1-\frac{2}{p}} + \frac{k}{m} - \frac{k+1}{m} \right), \end{aligned}$$

which reduces to

$$\frac{1}{4} \|w - R_w^\alpha\|_\square \leq \frac{2\Lambda(w)}{\alpha^2} + \frac{1}{m} + 2(2mM - 1) \|w - R_w^\alpha\|_p \alpha^{1-\frac{2}{p}}.$$

Case 2: Assume $\iint_{S' \times T'} (w - R_w^\alpha) dx dy \leq 0$ and $1 \leq k \leq mM - 1$.

By a similar argument used to show (9),

$$\frac{k+1}{m} - \bar{w}(S' \times T') \geq \frac{1}{4} \|w - R_w^\alpha\|_\square - 2(2mM - 1) \|w - R_w^\alpha\|_p \alpha^{1-\frac{2}{p}}.$$

Now let (x, y) be the upper left corner of $S' \times T'$. Then $(x, y) \in \mathcal{B}_k$, which means $\text{UL}(x, y)$ contains a region $S_c \times T_d$ such that $|S_c| = |T_d| = \alpha$ and $\bar{w}(S_c \times T_d) > \frac{k}{m}$. Using the definition of Λ , similar to the argument in Case 1, we get

$$\Lambda(w) \geq \frac{\alpha^2}{2} \left(\frac{k}{m} + \frac{1}{4} \|w - R_w^\alpha\|_\square - 2(2mM - 1) \|w - R_w^\alpha\|_p \alpha^{1-\frac{2}{p}} - \frac{k+1}{m} \right),$$

which reduces to

$$\frac{1}{4} \|w - R_w^\alpha\|_\square \leq \frac{2\Lambda(w)}{\alpha^2} + \frac{1}{m} + 2(2mM - 1) \|w - R_w^\alpha\|_p \alpha^{1-\frac{2}{p}}.$$

Case 3: Assume that either $\iint_{S' \times T'} (w - R_w^\alpha) dx dy > 0$ and $k = mM - 1$ or that $\iint_{S' \times T'} (w - R_w^\alpha) dx dy \leq 0$ and $k = 0$.

In the first assumption, we have that $R_w^\alpha > M - \frac{1}{m}$ on $S' \times T'$ whereas $w \leq M$ by definition. Thus, $\frac{\alpha^2}{m} \geq \iint_{S' \times T'} (w - R_w^\alpha) dx dy > 0$. In the second assumption, we have that $R_w^\alpha \leq \frac{1}{m}$ on $S' \times T'$ and by negativity of the integral in the second assumption it must be that $\frac{\alpha^2}{m} \geq \iint_{S' \times T'} (R_w^\alpha - w) \geq 0$. Combining either of these results with Claim 15 gives

$$\frac{\alpha^2}{m} \geq \alpha^2 \left[\frac{1}{4} \|w - R_w^\alpha\|_\square - 2(2mM - 1) \|w - R_w^\alpha\|_p \alpha^{1 - \frac{2}{p}} \right],$$

which can be rearranged to show

$$\frac{1}{4} \|w - R_w^\alpha\|_\square \leq \frac{1}{m} + 2(2mM - 1) \|w - R_w^\alpha\|_p \alpha^{1 - \frac{2}{p}}.$$

So Claim 16 holds in all cases.

To finish the proof, we take $\alpha = \|w\|_\infty^{-\frac{p}{3p-2}} \Lambda(w)^{\frac{2p}{5p-2}}$ and $m = \lceil \Lambda(w)^{-\frac{p-2}{5p-2}} \rceil$ in Claim 16, and we get

$$\frac{1}{4} \|w - R_w^\alpha\|_\square \leq \left(1 + (4\|w - R_w^\alpha\|_p + 2) \|w\|_\infty^{\frac{2p}{3p-2}} \right) \Lambda(w)^{\frac{p-2}{5p-2}} + \left(4\|w - R_w^\alpha\|_p \|w\|_\infty^{\frac{2p}{3p-2}} \right) \Lambda(w)^{\frac{2p-4}{5p-2}}.$$

Since $\Lambda(w) \leq 1$, we get $\Lambda(w)^{\frac{2p-4}{5p-2}} \leq \Lambda(w)^{\frac{p-2}{5p-2}}$; this simplifies the above equation to the desired result. \square

As a corollary to Proposition 14, we obtain an improvement for the earlier results of [26] on Robinson approximation of graphons and kernels.

Corollary 17. *Let $w : [0, 1]^2 \rightarrow [0, 1]$ be a graphon and $u : [0, 1]^2 \rightarrow \mathbb{R}$ be a kernel. Then*

$$\|w - R_w^\alpha\|_\square \leq 44\Lambda(w)^{\frac{1}{5}} \quad \text{and} \quad \|u - R_u^{\alpha^*}\|_\square \leq 44\|u\|_\infty^{\frac{4}{5}} \Lambda(u)^{\frac{1}{5}},$$

where $\alpha = \|w\|_\infty^{-\frac{1}{3}} \Lambda(w)^{\frac{2}{5}}$ and $\alpha^* = \|u\|_\infty^{-\frac{2}{5}} \Lambda(u)^{\frac{2}{5}}$.

Proof. For $w \in \mathcal{W}_0$, let $\alpha_p = \|w\|_\infty^{-\frac{p}{3p-2}} \Lambda(w)^{\frac{2p}{5p-2}}$. Then, by Proposition 14, we have $\|R_w^{\alpha_p} - w\|_\square \leq 44\Lambda(w)^{\frac{p-2}{5p-2}}$, since $\|w\|_\infty \leq 1$ and $\|w - R_w^{\alpha_p}\|_p \leq 1$. However, as $w \in L^p[0, 1]^2$ for every $p \geq 1$, we can allow $p \rightarrow \infty$, showing $\|w - R_w^\alpha\|_\square \leq 44\Lambda(w)^{\frac{1}{5}}$, where $\alpha = \|w\|_\infty^{-\frac{1}{3}} \Lambda(w)^{\frac{2}{5}}$ as desired. For $u \in \mathcal{W}$, we scale by $\|u\|_\infty$ to make a new function $u^* := u/\|u\|_\infty$ to get that

$$\|u^* - R_{u^*}^{\alpha^*}\|_\square \leq 44\Lambda(u^*)^{\frac{1}{5}}, \tag{10}$$

where $\alpha^* = \|u\|_\infty^{-\frac{2}{5}} \Lambda(u)^{\frac{2}{5}}$. We note that by definition of Λ and positivity of $\|u\|_\infty$, it must be the case that $\Lambda(u^*)^{\frac{1}{5}} = \|u\|_\infty^{-\frac{1}{5}} \Lambda(u)^{\frac{1}{5}}$. Furthermore, by definition of $R_{u^*}^{\alpha^*}$, we have $R_{u^*}^{\alpha^*} = \|u\|_\infty^{-1} R_u^{\alpha^*}$. Thus, combining these two observations with (10) yields $\frac{1}{\|u\|_\infty} \|u - R_u^{\alpha^*}\|_\square \leq 44\|u\|_\infty^{-\frac{1}{5}} \Lambda(u)^{\frac{1}{5}}$, which proves the claim. \square

We are now ready to state and prove our recovery result about Λ . We begin with a necessary definition—a key technique in this proof is taking an unbounded graphon and “cutting it off” at a certain threshold.

Definition 18 (*M*-cut-off). Let $w \in \mathcal{W}^1$ and define

$$E_M = \{(x, y) \in [0, 1]^2 : w(x, y) > M\}.$$

The *M*-cut-off of w , denoted by w_M , is defined to be $w_M := (\mathbb{1} - \mathbb{1}_M)w$, where $\mathbb{1}_M$ is the characteristic function of E_M and $\mathbb{1}$ is the characteristic function of $[0, 1]^2$.

Theorem 19. Suppose $w : [0, 1]^2 \rightarrow [0, \infty)$ is an L^p -kernel with $p > 5$ and $\|w\|_p \leq 1$. Then there exists some $\alpha \in [0, \frac{1}{2})$ such that R_w^α , the Robinson approximation of w with parameter α , satisfies

$$\|w - R_w^\alpha\|_\square \leq 78\Lambda(w)^{\frac{p-5}{5p-5}}. \quad (11)$$

The idea of the proof is as follows: we use the triangle inequality to bound $\|w - R_w^\alpha\|_\square$ by the sum of the cut-norms of $w - w_M$, $w_M - R_{w_M}^\alpha$ and $R_{w_M}^\alpha - R_w^\alpha$, and then bound each term using Λ . If w_M is Robinson, then the upper bound is proved without use of Corollary 17. If w_M is not Robinson, then $\|w_M - R_{w_M}^\alpha\|_\square$ must be handled using Corollary 17. Handling these two cases finishes the proof.

Proof. By definition of Robinson approximation, if $\Lambda(w) = 0$, then we set $\alpha = 0$, resulting in $R_w^\alpha = w$. Thus, we assume that $\Lambda(w) > 0$ and define our cut-off value $M = 2\Lambda(w)^{-\frac{1}{p-1}}$. We now consider two cases.

Case 1: Suppose that $\Lambda(w_M) > 0$.

Let $\alpha = \|w_M\|_\infty^{-\frac{2}{5}} \Lambda(w_M)^{\frac{2}{5}}$. The triangle inequality can then be used to show that

$$\|w - R_w^\alpha\|_\square \leq \|w - w_M\|_\square + \|w_M - R_{w_M}^\alpha\|_\square + \|R_{w_M}^\alpha - R_w^\alpha\|_\square.$$

We will proceed by bounding each of the terms on the right hand side one by one, starting with $\|w - w_M\|_\square$. By definition, $\mathbb{1}_M$ is the characteristic function of E_M , the region of $[0, 1]^2$ where $w > M$. Since $M|E_M|^{\frac{1}{p}} \leq \|w\mathbb{1}_M\|_p \leq \|w\|_p \leq 1$, we get $|E_M| \leq (\frac{1}{M})^p$. Therefore, for q satisfying $1/p + 1/q = 1$, it is true that

$$\|\mathbb{1}_M\|_q = |E_M|^{1/q} \leq \left(\frac{1}{M}\right)^{\frac{p}{q}} = M^{1-p} = 2^{1-p}\Lambda(w). \quad (12)$$

It is also true that

$$\|w - w_M\|_\square \leq \|w - w_M\|_1 = \|w\mathbb{1}_M\|_1 \leq \|w\|_p \|\mathbb{1}_M\|_q \leq 2^{1-p}\Lambda(w) \leq 2^{1-p}\Lambda(w)^{\frac{p-5}{5p-5}}, \quad (13)$$

handling the first term. We can further say that $\|w - w_M\|_\square \leq \frac{1}{16}\Lambda(w)^{\frac{p-5}{5p-5}}$ for all $p > 5$. Now we shift focus to $\|w_M - R_{w_M}^\alpha\|_\square$. By Corollary 17, as w_M is bounded,

$$\|w_M - R_{w_M}^\alpha\|_\square \leq 44\|w_M\|_\infty^{\frac{4}{5}} \Lambda(w_M)^{\frac{1}{5}} \leq 44M^{\frac{4}{5}} (\Lambda(w) + 2^{1-p}\Lambda(w))^{\frac{1}{5}},$$

where the second inequality is due to the combination of (1) for $w_M = w - w\mathbb{1}_M$ alongside the fact that $\Lambda(-w\mathbb{1}_M) \leq \|w\mathbb{1}_M\|_1 \leq 2^{1-p}\Lambda(w)$ by (12). This simplifies to

$$\|w_M - R_{w_M}^\alpha\|_\square \leq 44(1 + 2^{1-p})^{\frac{1}{5}} M^{\frac{4}{5}} \Lambda(w)^{\frac{1}{5}} \leq 44(1 + 2^{1-p})^{\frac{1}{5}} 2^{\frac{4}{5}} \Lambda(w)^{\frac{p-5}{5p-5}}.$$

Using the fact that $p > 5$, the above inequality gives us:

$$\|w_M - R_{w_M}^\alpha\|_\square \leq 77.6 \Lambda(w)^{\frac{p-5}{5p-5}}. \tag{14}$$

To bound the third term, we use Proposition 9 (i) and Proposition 9 (ii) to get that

$$\|R_w^\alpha - R_{w_M}^\alpha\|_\square \leq \|R_w^\alpha - R_{w_M}^\alpha\|_\infty \leq \|R_{w-w_M}^\alpha\|_\infty \leq \alpha^{-2} \|w - w_M\|_1.$$

Applying Hölder's inequality with conjugate indices p, q , we conclude

$$\|R_w^\alpha - R_{w_M}^\alpha\|_\square \leq \alpha^{-2} \|w\mathbb{1}_M\|_1 \leq \alpha^{-2} \|w\|_p \|\mathbb{1}_M\|_q.$$

Now, using the upper bound for $\|\mathbb{1}_M\|_q$ provided by (12), and substituting the value of α , we observe that

$$\|R_w^\alpha - R_{w_M}^\alpha\|_\square \leq 2^{1-p} \alpha^{-2} \Lambda(w) = 2^{1-p} \Lambda(w) \left(\|w_M\|_\infty^{-\frac{2}{5}} \Lambda(w_M)^{\frac{2}{5}} \right)^{-2}. \tag{15}$$

Applying (1) to $w = w_M + w\mathbb{1}_M$ gets us

$$\Lambda(w_M) \geq \Lambda(w) - \Lambda(w\mathbb{1}_M) \geq \Lambda(w) - \|w\mathbb{1}_M\|_1 \geq \Lambda(w) - \|w\|_p \|\mathbb{1}_M\|_q \geq (1 - 2^{1-p}) \Lambda(w),$$

and since $p > 5$, this requires that $\Lambda(w_M) \geq \frac{15}{16} \Lambda(w)$. Using this together with $\|w_M\|_\infty \leq M = 2\Lambda(w)^{\frac{-1}{p-1}}$ and inequality (15) implies that

$$\begin{aligned} 2^{p-1} \|R_w^\alpha - R_{w_M}^\alpha\|_\square &\leq \Lambda(w) \|w_M\|_\infty^{\frac{4}{5}} \Lambda(w_M)^{-\frac{4}{5}} \leq \Lambda(w) M^{\frac{4}{5}} (1 - 2^{1-p})^{-\frac{4}{5}} \Lambda(w)^{-\frac{4}{5}} \\ &\leq \frac{2^{\frac{4}{5}}}{(1 - 2^{1-p})^{\frac{4}{5}}} \Lambda(w)^{\frac{p-5}{5p-5}}. \end{aligned}$$

When $p > 5$, we have $2^{\frac{4}{5}} 2^{1-p} (1 - 2^{1-p})^{-\frac{4}{5}} \leq 0.2$, so

$$\|R_w^\alpha - R_{w_M}^\alpha\|_\square \leq 0.2 \Lambda(w)^{\frac{p-5}{5p-5}}. \tag{16}$$

Thus (13), (14), and (16) together imply that $\|w - R_w^\alpha\|_\square \leq 78 \Lambda(w)^{\frac{p-5}{5p-5}}$, proving the statement of the theorem in this case.

Case 2: Suppose that $\Lambda(w_M) = 0$.

Let $\alpha = M^{-\frac{2}{5}} \Lambda(w)^{\frac{2}{5}}$. We proceed similarly to Case 1, bounding each of the three summands in the right hand side of the following inequality:

$$\|w - R_w^\alpha\|_\square \leq \|w - w_M\|_\square + \|w_M - R_{w_M}^\alpha\|_\square + \|R_w^\alpha - R_{w_M}^\alpha\|_\square.$$

The first term on the right side of the inequality can be handled identically to Case 1, yielding $\|w - w_M\|_{\square} \leq \frac{1}{16} \Lambda(w)^{\frac{p-5}{5p-5}}$. For the third term, we can proceed identically to Case 1 and get

$$\|R_w^{\alpha} - R_{w_M}^{\alpha}\|_{\square} \leq 2^{1-p} \alpha^{-2} \Lambda(w).$$

Putting $\alpha = M^{-\frac{2}{5}} \Lambda(w)^{\frac{2}{5}}$ and noting that $p > 5$ results in $\|R_w^{\alpha} - R_{w_M}^{\alpha}\|_{\square} \leq 0.2 \Lambda(w)^{\frac{p-5}{5p-5}}$.

However, the second term $\|w_M - R_{w_M}^{\alpha}\|_{\square}$ must be handled differently. Indeed, when $\Lambda(w_M) = 0$, we have that w_M is Robinson a.e., and we need to approximate $R_{w_M}^{\alpha}$ directly.

Claim 20. *When $w_M : [0, 1]^2 \rightarrow [0, M]$ is Robinson a.e., we have*

$$\|w_M - R_{w_M}^{\alpha}\|_{\square} \leq \iint_{\{|x-y| \leq 2\alpha\}} w_M(x, y) dx dy.$$

To prove the claim, first note that for any point $(x, y) \in \Delta$, we have that $w_M(x, y)$ is an upper bound for every value of w_M over the set $UL(x, y)$; thus, any average over that set (such as in the definition of R_w^{α}) would not exceed $w_M(x, y)$. So $R_{w_M}^{\alpha} \leq w_M$. On the other hand, as w_M is Robinson a.e., we have

$$R_{w_M}^{\alpha}(x, y) = R_{w_M}^{\alpha}(y, x) = \begin{cases} \frac{1}{\alpha^2} \iint_{[x-\alpha, x] \times [y, y+\alpha]} w_M dx dy & \text{for } (x, y) \in [\alpha, 1-\alpha]^2 \cap \Delta \\ 0 & \text{otherwise} \end{cases}. \tag{17}$$

We now introduce an auxiliary function \tilde{w}_M defined as follows:

$$\tilde{w}_M(x, y) = \tilde{w}_M(y, x) = \begin{cases} w_M(x - \alpha, y + \alpha) & (x, y) \in [\alpha, 1 - \alpha]^2 \cap \Delta \\ 0 & \text{otherwise} \end{cases}.$$

Since w_M is Robinson, we have that $R_{w_M}^{\alpha} \geq \tilde{w}_M$. Thus we have $0 \leq w_M - R_{w_M}^{\alpha} \leq w_M - \tilde{w}_M$ pointwise, allowing us to show the following:

$$\begin{aligned} \|w_M - R_{w_M}^{\alpha}\|_{\square} &\leq \|w_M - \tilde{w}_M\|_{\square} = 2 \iint_{[0,1]^2 \cap \Delta} (w_M - \tilde{w}_M) dx dy \\ &= 2 \left(\iint_{[0,1]^2 \cap \Delta} w_M dx dy - \iint_{[\alpha, 1-\alpha]^2 \cap \Delta} \tilde{w}_M dx dy \right) \\ &= 2 \left(\iint_{[0,1]^2 \cap \Delta} w_M dx dy - \iint_{0 \leq x \leq y - 2\alpha \leq 1 - 2\alpha} w_M(x, y) dx dy \right) \\ &\leq \iint_{\{|x-y| \leq 2\alpha\}} w_M(x, y) dx dy. \end{aligned}$$

This proves Claim 20. Finally, note that

$$\iint_{\{|x-y| \leq 2\alpha\}} w_M(x, y) dx dy \leq \|w_M\|_p \|\mathbb{1}_{[0,1]^2 \cap \{|x-y| \leq 2\alpha\}}\|_q \leq (1 - (1 - 2\alpha)^2)^{\frac{1}{q}} \leq (4\alpha)^{1 - \frac{1}{p}},$$

where in the last inequality we used that $0 \leq \alpha \leq 1$. Substituting for α and noting that $p > 5$, we have

$$\|w_M - R_{w_M}^\alpha\|_\square \leq 4^{1-\frac{1}{p}} \left(\frac{\Lambda(w)}{M}\right)^{\left(\frac{2}{5}\right)\left(1-\frac{1}{p}\right)} \leq 4\Lambda(w)^{\frac{2}{5}}.$$

Since $\Lambda(w) \leq 1$ and $\frac{2}{5} \geq (p-5)(5p-5)^{-1}$ for all $p > 5$, we get $\|w_M - R_{w_M}^\alpha\|_\square \leq 4\Lambda(w)^{\frac{2}{5}} \leq 4\Lambda(w)^{\frac{p-5}{5p-5}}$. Therefore, if $\Lambda(w_M) = 0$, we have that $\|w - R_w^\alpha\|_\square \leq 5\Lambda(w)^{\frac{p-5}{5p-5}}$, proving the statement of the theorem in this case. \square

Remark 21 (Concluding remarks). We finish this section with a discussion on the performance of Λ and a comparison with the previously defined function Γ from [14].

- (i) The function Γ provides a continuous and robust mapping on L^p -graphons only for the case $p = \infty$ (see Remark 3), whereas Λ can be used as a suitable gauge of Robinson property for any L^p -graphon with $5 < p \leq \infty$ (see Theorem 19). Moreover, the error (in cut-norm) of the Robinson approximation provided by Γ of a graphon $w : [0, 1]^2 \rightarrow [0, 1]$ is bounded by $14\Gamma(w)^{\frac{1}{7}}$ (see [14, Theorem 3.2]). However, using Λ for Robinson approximation of w leads to a much improved upper bound of $44\Lambda(w)^{\frac{1}{5}}$ (see Corollary 17). This improvement is partly due to the refined definition of Λ (compared to Γ), and partly due to the possibility of using sharper results in Proposition 14 for L^p -graphons to obtain bounds for L^∞ -graphons by letting $p \rightarrow \infty$.
- (ii) Theorem 19 holds only for $p > 5$. This is an artefact of the proof technique used and not an indication that similar results do not hold for $1 \leq p \leq 5$. We anticipate that ideological bottlenecks would appear at $p \leq 2$ and $p = 1$. We conjecture that results analogous to Theorem 19 are valid for $p \geq 2$, though achieving these results would likely require new proof methods and more advanced techniques. Indeed, our proof of Theorem 19 is based on the idea of approximating an L^p -graphon with its M -cut-off, and then applying Robinson approximation results for bounded graphons to the cut-off function. The condition $p > 5$ is essential for obtaining reasonable cut-off approximations (see (13)). This requirement is the primary impediment to adapting the proof of Theorem 19 to $p \leq 5$.

Acknowledgements

We thank Jeannette Janssen for helpful comments which improved this work. We thank the Department of Mathematical Sciences at the University of Delaware for their continued support throughout the process of this research. M. Ghandehari was supported by NSF grants DMS-1902301 and DMS-2408008 while this work was being completed. Finally, we sincerely thank the anonymous reviewer for reading the manuscript and suggesting improvements.

References

- [1] A. Agarwal, S. Negahban, and M.J. Wainwright. Noisy matrix decomposition via convex relaxation: Optimal rates in high dimensions. *The Annals of Statistics*, 40(2):1171–1197, 2012.
- [2] J.E. Atkins, E.G. Boman, and B. Hendrickson. A spectral algorithm for seriation and the consecutive ones problem. *SIAM Journal on Computing*, 28(1):297–310, 1998.
- [3] M. Avella-Medina and F. Parise and M. Schaub and S. Segarra. Centrality measures for graphons: accounting for uncertainty in networks. *IEEE Transactions on Network Science and Engineering*, 7(1):520–537, 2020.
- [4] J.P. Barthélemy and F. Brucker. NP-hard approximation problems in overlapping clustering. *Journal of Classification*, 18(2):159–183, 2001.
- [5] C. Borgs, J. Chayes, H. Cohn, and Y. Zhao. An l^p theory of sparse graph convergence II: L_d convergence, quotients and right convergence. *The Annals of Probability*, 46(1), Jan 2018.
- [6] C. Borgs, J. 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*, 372(5):3019–3062, May 2019.
- [7] T. Cai and W. Liu. Adaptive thresholding for sparse covariance matrix estimation. *Journal of the American Statistical Association*, 106(494):672–684, 2011.
- [8] E.J. Candes and Y. Plan. Matrix completion with noise. *Proceedings of the IEEE*, 98(6):925–936, 2010.
- [9] E.J. Candès and B. Recht. Exact matrix completion via convex optimization. *Foundations of Computational Mathematics*, 9(6):717–772, Dec 2009.
- [10] E.J. Candès and T. Tao. The power of convex relaxation: near-optimal matrix completion. *IEEE Transactions on Information Theory*, 56(5):2053–2080, 2010.
- [11] V. Chandrasekaran, S. Sanghavi, P.A. Parrilo, and A.S. Willsky. Rank-sparsity incoherence for matrix decomposition. *SIAM Journal on Optimization*, 21(2):572–596, 2011.
- [12] Y. Chen, Y. Chi, and A. Goldsmith. Exact and stable covariance estimation from quadratic sampling via convex programming. *IEEE Transactions on Information Theory*, 61(2):4034–4059, 2015.
- [13] V. Chepoi and M. Seston. Seriation in the presence of errors: A factor 16 approximation algorithm for l_∞ -fitting robinson structures to distances. *Algorithmica*, 59:521–568, 2011.
- [14] H. Chuangpishit, M. Ghandehari, M. Hurshman, J. Janssen, and N. Kalyaniwalla. Linear embeddings of graphs and graph limits. *Journal of Combinatorial Theory, Series B*, 113:162–184, Jul 2015.
- [15] Y. de Castro, C. Lacour, and T.M. Pham Ngoc. Adaptive estimation of nonparametric geometric graphs. *Mathematical Statistics and Learning*, 2017.

- [16] A. Eftekhari, D. Yang, and M.B. Wakin. Weighted matrix completion and recovery with prior subspace information. *IEEE Transactions on Information Theory*, 64(6):4044–4071, 2018.
- [17] M. Fazel, T.K. Pong, D. Sun, and P. Tseng. Hankel matrix rank minimization with applications to system identification and realization. *SIAM Journal on Matrix Analysis and Applications*, 34(3):946–977, 2013.
- [18] N. Flammarion, C. Mao, and P. Rigollet. Optimal rates of statistical seriation. *Bernoulli*, 25(1):623–653, 2019.
- [19] F. Fogel, A. d’Aspremont, and M. Vojnovic. Spectral ranking using seriation. *Journal of Machine Learning Research*, 17(1):3013–3057, January 2016.
- [20] D. Fortin. Robinsonian matrices: Recognition challenges. *Journal of Classification*, 34:191–222, 2017.
- [21] A. Frieze and R. Kannan. Quick approximation to matrices and applications. *Combinatorica*, 19:175–220, 02 1999.
- [22] C. Gao, Y. Lu, and H. Zhou. Rate-optimal graphon estimation. *The Annals of Statistics*, 43(1):2624–2652, 2015.
- [23] C. Gao and Z. Ma. Minimax rates in network analysis: graphon estimation, community detection and hypothesis testing. *Statistical Science*, 36(1):16–33, 2021.
- [24] S. Gao and P. Caines. Graphon control of large-scale networks of linear systems. *IEEE Transactions on Automatic Control*, 65(1):4090–4105, 2021.
- [25] M. Ghandehari and J. Janssen. An optimization parameter for seriation of noisy data. *SIAM Journal on Discrete Mathematics*, 33(2):712–730, 2019.
- [26] M. Ghandehari and J. Janssen. Graph sequences sampled from Robinson graphons. *European Journal of Combinatorics*, 116(2024), Paper No. 103859, 31 pp.
- [27] M. Ghandehari, J. Janssen, and N. Kalyaniwalla. A noncommutative approach to the graphon Fourier transform. *Appl. Comput. Harmon. Anal.*, 61:101–131, 2022.
- [28] M. Gkogkas and C. Kuehn and C. Xu. Continuum limits for adaptive network dynamics. *Communications in Mathematical Sciences*, 21(1): 83–106, 2023.
- [29] F. Goyens, C. Cartis, and A. Eftekhari. Nonlinear matrix recovery using optimization on the grassmann manifold. *Applied and Computational Harmonic Analysis*, 62:498–542, 2023.
- [30] D. Gross. Recovering low-rank matrices from few coefficients in any basis. *IEEE Transactions on Information Theory*, 57(3):1548–1566, 2011.
- [31] P. Jain, P. Netrapalli, and S. Sanghavi. Low-rank matrix completion using alternating minimization (extended abstract). In *STOC’13—Proceedings of the 2013 ACM Symposium on Theory of Computing*, pages 665–674. ACM, New York, 2013.
- [32] J. Janssen and A. Smith. Reconstruction of line-embeddings of graphons. *Electronic Journal of Statistics*, 16(1):331–407, 2022.

- [33] R.H. Keshavan, A. Montanari, and S. Oh. Matrix completion from noisy entries. *Journal of Machine Learning Research*, 11:2057–2078, 2010.
- [34] R.H. Keshavan, A. Montanari, and S. Oh. Matrix completion from a few entries. *IEEE Transactions on Information Theory*, 56(6):2980–2998, 2010.
- [35] O. Klopp. Noisy low-rank matrix completion with general sampling distribution. *Bernoulli*, 20(1):282–303, 2014.
- [36] O. Klopp, K. Lounici, and A.B. Tsybakov. Robust matrix completion. *Probability Theory and Related Fields*, 169(1):523–564, Oct 2017.
- [37] V. Koltchinskii, K. Lounici, and A.B. Tsybakov. Nuclear-norm penalization and optimal rates for noisy low-rank matrix completion. *The Annals of Statistics*, 39(5): 2302 – 2329, 2011.
- [38] M. Laurent and M. Seminaroti. A lex-bfs-based recognition algorithm for robinsonian matrices. *Discrete Applied Mathematics*, 222:151–165, 2017.
- [39] M. Laurent, M. Seminaroti, and S. Tanigawa. A structural characterization for certifying robinsonian matrices. *Electronic Journal of Combinatorics*, 24(2), #P2.21, 2017.
- [40] I. Liiv. Seriation and matrix reordering methods: An historical overview. *Statistical Analysis and Data Mining: The ASA Data Science Journal*, 3(2):70–91, 2010.
- [41] L. Lovász. *Large Networks and Graph Limits.*, volume 60 of *Colloquium Publications*. American Mathematical Society, 2012.
- [42] L. Lovász and B. Szegedy. Limits of dense graph sequences. *Journal of Combinatorial Theory, Series B*, 96(6):933–957, 2006.
- [43] R. Ma, T. Cai, and H. Li. Optimal permutation recovery in permuted monotone matrix model. *Journal of the American Statistical Association*, 116(2):1358–1372, 2021.
- [44] G. Medvedev. The nonlinear heat equation on dense graphs and graph limits, *SIAM Journal on Mathematical Analysis*, 46(4):2743–2766, 2014
- [45] B.G. Mirkin, H.L. Beus, and S.N. Rodin. *Graphs and Genes*. Biomathematics. Springer Berlin Heidelberg, 1984.
- [46] T. Mishura. Cut norm discontinuity of triangular truncation of graphons. *Linear Algebra and its Applications*, 650:26–41, 2022.
- [47] A. Natik and A. Smith. Consistency of spectral seriation, [arXiv:2112.04408](https://arxiv.org/abs/2112.04408).
- [48] P. Pr ea and D. Fortin. An optimal algorithm to recognize Robinsonian dissimilarities. *Journal of Classification*, 31(3):351–385, 2014.
- [49] B. Recht. A simpler approach to matrix completion. *Journal of Machine Learning Research*, 12:3413–3430, 2011.
- [50] W. S. Robinson. A method for chronologically ordering archaeological deposits. *American Antiquity*, 16(4):293–301, 1951.

- [51] A. Rohde and A. B. Tsybakov. Estimation of high-dimensional low-rank matrices. *Annals of Statistics*, 39(2):887–930, 2011.
- [52] M. Rudelson and S. Zhou. Reconstruction from anisotropic random measurements. *IEEE Transactions on Information Theory*, 59(6):3434–3447, 2013.
- [53] M. Seston. A simple algorithm for recognize Robinsonian dissimilarities. In *COMP-STAT 2008—Proceedings in Computational Statistics*, pages 241–248, CD-ROM. Physica-Verlag/Springer, Heidelberg, 2008.
- [54] R. Sun and Z.-Q. Luo. Guaranteed matrix completion via non-convex factorization. *IEEE Transactions on Information Theory*, 62(11):6535–6579, 2016.
- [55] M. Tao and X. Yuan. Recovering low-rank and sparse components of matrices from incomplete and noisy observations. *SIAM Journal on Optimization*, 21(1):57–81, 2011.
- [56] B. Vandereycken. Low-rank matrix completion by Riemannian optimization. *SIAM Journal on Optimization*, 23(2):1214–1236, 2013.
- [57] C.-L. Wang, C. Li, and J. Wang. Comparisons of several algorithms for toeplitz matrix recovery. *Computers & Mathematics with Applications*, 71(1):133–146, 2016.

Appendices

Proofs of Lemmas 11, 12 and 13

In this Appendix we present proofs of the lemmas used previously in the paper, recalling their statements for clarity.

Lemma (Lemma 11). Let $k \in \mathbb{Z}^{\geq 0}$, $w \in \mathcal{W}^p$ with $w \geq 0$, and $\alpha \in (0, 1)$. Then, $\overline{\mathcal{G}_k}$ does not contain any $\beta \times \beta$ square, where $\beta > \alpha$. Here, $\overline{\mathcal{G}_k}$ denotes the closure of \mathcal{G}_k in the Euclidean topology of \mathbb{R}^2 .

Proof of Lemma 11. Let $k \in \mathbb{Z}^{\geq 0}$ be fixed and let \mathcal{G}_k be defined with parameter m and α . Towards a contradiction, let $\beta > \alpha$ and suppose there exist measurable subsets $S, T \subseteq [0, 1]$ with $|S| = |T| = \beta$ for which $S \times T \subseteq \overline{\mathcal{G}_k}$. Since $|S| = |T| = \beta > \alpha$, there exist $a_1, a_2 \in S$ and $b_1, b_2 \in T$ such that $a_2 - a_1 > \alpha$ and $b_2 - b_1 > \alpha$. Note that $(a_i, b_j) \in \overline{\mathcal{G}_k}$ for $i, j = 1, 2$. Since every point in $\overline{\mathcal{G}_k}$ that is not on the lower or upper boundary curves must be an interior point of \mathcal{G}_k , by moving the points slightly if necessary, we can assume that

$$a_2 - a_1 > \alpha, \quad b_2 - b_1 > \alpha, \quad \text{and } (a_i, b_j) \in \mathcal{G}_k, i, j = 1, 2.$$

We shall now show the inclusion $(a_1, a_2) \times (b_1, b_2) \subseteq \mathcal{G}_k$. Towards a contradiction, suppose that there exists some point $(z, w) \in \text{UL}(a_2, b_1) \cap \text{LR}(a_1, b_2) \setminus \mathcal{G}_k$, which implies that either $(z, w) \in \text{UL}(a_2, b_1) \cap \text{LR}(a_1, b_2) \cap \mathcal{B}_k$ or that $(z, w) \in \text{UL}(a_2, b_1) \cap \text{LR}(a_1, b_2) \cap \mathcal{V}_k$. The first case implies that $(a_2, b_1) \in \mathcal{B}_k$ while the second case implies that $(a_1, b_2) \in \mathcal{V}_k$, both of which are contradictions with $(a_i, b_j) \in \mathcal{G}_k$. Thus, \mathcal{G}_k contains $(a_1, a_2) \times (b_1, b_2)$.

Clearly, $(a_1, a_2) \times (b_1, b_2)$ contains a closed $\alpha \times \alpha$ rectangle which we denote $[a'_1, a'_2] \times [b'_1, b'_2]$. The two points (a'_1, b'_2) and (a'_2, b'_1) are elements of \mathcal{G}_k , so they fail to satisfy the conditions for both \mathcal{V}_k and \mathcal{B}_k . In particular, we have $\overline{w}([a'_1, a'_2] \times [b'_1, b'_2]) \leq (k - 1)/m$ as $(a'_2, b'_1) \notin \mathcal{B}_k$, as well as $\overline{w}([a'_1, a'_2], [b'_1, b'_2]) > (k - 1)/m$ as $(a'_1, b'_2) \notin \mathcal{V}_k$. These statements form a contradiction, showing the initial claim must be true. \square

Lemma. (Lemma 12) Let $u \in L^\infty([0, 1])$, let $P \subseteq [0, 1]$ be a measurable subset such that $\int_P u \, dx \neq 0$, and let $0 < \beta < |P|$ be fixed. Then P can be partitioned into $N := \lceil |P|/\beta \rceil$ subsets P_1, \dots, P_N so that the following conditions are satisfied:

- (i) $P_1 \leq \dots \leq P_{N-1}$.
- (ii) $|P_i| = \beta$ for $1 \leq i \leq N - 1$ and $|P_N| \leq \beta$.
- (iii) $|\int_{P_N} u \, dx| \leq \frac{1}{N} |\int_P u \, dx|$.

Proof of Lemma 12. For any sets P_1, \dots, P_N satisfying the above properties, we must have

$$|P_N| = \delta := |P| - \beta \left(\left\lceil \frac{|P|}{\beta} \right\rceil - 1 \right).$$

To prove the lemma, it is enough to find a subset $P_N \subseteq P$ satisfying $|P_N| = \delta$ and condition (iii). Indeed, given such a subset P_N , one can form the other sets P_1, \dots, P_{N-1}

by splitting $P \setminus P_N$ into consecutive sets of measure β . If $\delta = \beta$, then the lemma follows trivially from the pigeonhole principle. So, we assume that $\delta < \beta$. Moreover, replacing u by $-u$ if necessary, we can assume that $\int_P u \, dx > 0$. We claim there must exist a set $Q \subset P$ such that $|Q| = |P| - \beta \left(\left\lceil \frac{|P|}{\delta} \right\rceil - 1 \right)$ and $\int_Q u \, dx > 0$. To show this statement, note that

$$\int_P u \, dx = \int_{P^+} u \, dx + \int_{P^-} u \, dx > 0, \quad (18)$$

where $P^+ := \{x \in P : u(x) > 0\}$ and $P^- := \{y \in P : u(y) < 0\}$. If $|P^+| \geq \delta$, then any subset Q of measure δ from P^+ suffices. If $|P^+| < \delta$, then we let $Q = P^+ \cup Q'$ for a subset $Q' \subseteq P^-$ with $|Q'| = \delta - |P^+|$, and note that $\int_Q u \, dx \geq \int_{P^+} u \, dx > 0$ by (18). We now prove the existence of a subset $P_N \subseteq P$ satisfying $|P_N| = \delta$ and condition (iii). Towards a contradiction, assume that for any set $S \subseteq P$ such that $|S| = \delta$, we have

$$\left| \int_S u \, dx \right| > \frac{1}{N} \int_P u \, dx. \quad (19)$$

Consider now $R := P \setminus Q$. For any $p \in P$, let $r_p = \inf\{q : |R \cap [p, q]| = \delta\}$, and let the auxiliary function ϕ be defined as

$$\phi : P \rightarrow \mathbb{R}, \quad \phi(p) := \int_{R \cap [p, r_p]} u \, dx.$$

For ease of writing, let $R_p := R \cap [p, r_p]$, and note that as $|R_p| = \delta$, by (19), we get $\phi(p) \neq 0$ for all $p \in P$. It is easy to see that ϕ is continuous, as for $p, q \in P$, we have

$$|\phi(p) - \phi(q)| = \left| \int_{R_p} u \, dx - \int_{R_q} u \, dx \right| = \left| \int_{R_p \Delta R_q} u \, dx \right| \leq \|u\|_\infty |R_p \Delta R_q| \leq \|u\|_\infty |p - q|.$$

Therefore, as $\phi \neq 0$, it is either strictly positive or strictly negative. Without loss of generality, we assume that $\phi(p) > 0$ for all p , and to avoid violating (19), it must also be that

$$\phi(p) > \frac{1}{N} \int_P u \, dx \quad (20)$$

for all $p \in P$. Consider $\{p_i\}_{i=1}^M \subset R$, where $p_1 = 0$, $p_{i+1} = r_{p_i}$ for $1 < i < M$, and $M = \left\lceil \frac{|P|}{\delta} \right\rceil - 1$. Then, due to the positivity of w and (20), we have that

$$\int_P u \, dx = \sum_{i=1}^M \phi(p_i) + \int_Q u \, dx > \frac{M}{N} \int_P u \, dx. \quad (21)$$

This will lead to a contradiction, as we will show that $M \geq N$, i.e. $\left\lceil \frac{|P|}{\delta} \right\rceil - 1 \geq \left\lceil \frac{|P|}{\beta} \right\rceil$. We note that as $\delta < \beta$, it must be the case that $M \geq N - 1$. Therefore, for (21) to avoid a contradiction, M must be equal to $N - 1$. This forces δ within a tight range of values:

Specifically, $|P|/\left\lceil\frac{|P|}{\beta}\right\rceil \leq \delta < \beta$. However, the exact value of δ is $|P| - \beta\left(\left\lceil\frac{|P|}{\beta}\right\rceil - 1\right)$, and thus if $M = N - 1$, we have

$$\begin{cases} |P| - \beta\left(\left\lceil\frac{|P|}{\beta}\right\rceil - 1\right) \geq \frac{|P|}{\left\lceil\frac{|P|}{\beta}\right\rceil}, \text{ which implies, } |P| \geq \beta\left\lceil\frac{|P|}{\beta}\right\rceil \\ |P| - \beta\left(\left\lceil\frac{|P|}{\beta}\right\rceil - 1\right) < \beta, \text{ or equivalently, } |P| < \beta\left\lceil\frac{|P|}{\beta}\right\rceil \end{cases}$$

This is a contradiction, so $M \geq N$, making (21) a contradiction. Thus the lemma holds true and such a partition of P must exist. \square

Lemma. (Lemma 13) Let $f \in L^1([0, 1]^2)$, and let $S, S' \subseteq [0, 1]$ be measurable subsets such that $|S| = |S'|$. Suppose for a constant $C > 0$ we have

$$\iint_{S \times S'} f \, dx dy \geq C.$$

Then, for every $\alpha \in (0, 1)$, there exist measurable sets $T \subset S$ and $T' \subset S'$ such that $|T| = |T'| = \alpha|S|$ and

$$\frac{1}{|T \times T'|} \iint_{T \times T'} f \, dx dy \geq \frac{C}{|S \times S'|}.$$

Proof of Lemma 13. Suppose there exist integers n, k, l , with $l < k$, so that $|S| = \frac{k}{n}$ and $\alpha|S| = \frac{l}{n}$; the case where one or both of $|S|$ or α is not rational can be done using standard density/approximation arguments. Next, split S into k consecutive sets $S_1 \leq S_2 \leq \dots \leq S_k$ of measure $\frac{1}{n}$; likewise, split S' into k consecutive sets $S'_1 \leq S'_2 \leq \dots \leq S'_k$ also of measure $\frac{1}{n}$ and note that

$$\iint_{S \times S'} f \, dx dy = \sum_{i=1}^k \sum_{j=1}^k \iint_{S_i \times S'_j} f \, dx dy.$$

Let $B_{i,j} := \iint_{S_i \times S'_j} f \, dx dy$ and let $\mathcal{N}_{k,l}$ denote the collection of all l -subsets of the set $\{1, \dots, k\}$. Note that there are $\binom{k-1}{l-1}$ many l -subsets of k elements containing a specific element i_0 . Using this, counting the number of times a specific $B_{i,j}$ appears in the following sum results in

$$\sum_{I \in \mathcal{N}_{k,l}} \sum_{J \in \mathcal{N}_{k,l}} \sum_{\substack{i \in I \\ j \in J}} B_{i,j} = \sum_{i,j=1}^k \binom{k-1}{l-1}^2 B_{i,j} \geq \binom{k-1}{l-1}^2 C.$$

Thus, by the pigeonhole principle, there exist sets $I, J \in \mathcal{N}_{k,l}$ such that

$$\sum_{\substack{i \in I \\ j \in J}} B_{i,j} \geq \frac{\binom{k-1}{l-1}^2}{\binom{k}{l}^2} C = \frac{l^2}{k^2} C.$$

This implies that the sets $T = \cup_{i \in I} S_i$ and $T' = \cup_{i \in J} S'_i$ satisfy both that

$$\iint_{T \times T'} f \, dx dy \geq \frac{l^2}{k^2} C$$

and $|T| = |T'| = \alpha|S|$, showing the lemma holds true. □