Moments of Two-Variable Functions and the Uniqueness of Graph Limits

Christian Borgs
Jennifer Chayes
László Lovász

Microsoft Research, Redmond, WA 98052

October 2008; revised May 2009

Abstract

For a symmetric bounded measurable function W on $[0,1]^2$ and a simple graph F, the homomorphism density

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

can be thought of as a "moment" of W. We prove that every such function is determined by its moments up to a measure preserving transformation of the variables.

The main motivation for this result comes from the theory of convergent graph sequences. A sequence (G_n) of dense graphs is said to be convergent if the probability, $t(F,G_n)$, that a random map from $V(F)$ into $V(G_n)$ is a homomorphism converges for every simple graph F. The limiting density can be expressed as $t(F,W)$ for a symmetric bounded measurable function W on $[0,1]^2$. Our results imply in particular that the limit of a convergent graph sequence is unique up to measure preserving transformation.

Contents

1 Introduction

Let $\mathcal{W}$ be the set of bounded symmetric measurable functions $W: [0,1]^2 \to \mathbb{R}$, and let $\mathcal{W}_0$ denote the set of functions in $\mathcal{W}$ with values in $[0,1]$. For every $W \in \mathcal{W}$ and every finite graph F, we define the integral

$$t(F,W) = \int\limits_{[0,1]^n} \prod_{ij\in E(F)} W(x_i,x_j) \prod_{i\in V(F)} dx_i \,. \tag{1}$$

Our interest in these integrals stems from graph theory (see next paragraph), but such integrals appear in physics, statistics, and other areas. In many respects, these integrals can be thought of as 2-variable analogues of moments of 1-variable functions, so instead of moment sequences, such 2-variable functions have a "moment graph parameter" (function defined on graphs). Just like moments of a 1-variable function determine the function up to measure preserving transformations, these "moments" determine the 2-variable function up to measure preserving transformations. The exact formulation and proof of this fact is the main goal of this paper.

Our main motivation for this study comes from the theory of convergent graph sequences. Let F and G be two simple graphs (graphs without loops and multiple edges). Let us map the nodes of F randomly into $V(G)$, and let $t(F,G)$ denote the probability that this map preserves adjacency. For example, $t(K_2,G)$ denotes the edge density of G. In general, we call $t(F,G)$ the *homomorphism density* or simply the *density* of F in G.

We call a sequence of simple graphs (G_n) *convergent*, if $t(F,G_n)$ has a limit for every simple graph F. The notion of convergent graph sequences was introduced by Borgs, Chayes, Lovász, Sós and Vesztergombi [2], see also [3], and further studied in [4] and [5]. Lovász and Szegedy [12] proved that every convergent graph sequence has a "limit object" in the form of a function $W \in \mathcal{W}_0$ in the sense that

$$t(F,G_n) \longrightarrow t(F,W) \qquad \text{as} \quad n\to\infty \tag{2}$$

for every simple graph F. In this case we say that G_n *converges to* W. It was also shown in [12] that for every function $W \in \mathcal{W}_0$ there is a convergent sequence (G_n) of simple graphs converging to W. To complete the picture, the results in this paper imply that *the limit object is unique up to measure preserving transformations.*

2 Results

For the precise statement of our results, we need some definitions. Instead of the interval $[0,1]$, we consider two-variable functions on an arbitrary probability space; while this does not add real generality it leads to a cleaner picture. We need a few definitions.

We start by recalling some basic notions from probability theory. Let $(\Omega, \mathcal{A}, \pi)$ be a probability space (where Ω is the underlying set, $\mathcal{A}$ is a σ-algebra on Ω, and π is a probability measure on $\mathcal{A}$). As usual, $(\Omega, \mathcal{A}, \pi)$ is called complete if $\mathcal{A}$ contains all sets of external measure 0, and the completion of $(\Omega, \mathcal{A}, \pi)$ is obtained by replacing $\mathcal{A}$ with the σ-algebra generated by $\mathcal{A}$ and all subsets $N \subset \Omega$ of external measure 0.

Let $(\Omega, \mathcal{A}, \pi)$ and $(\Omega', \mathcal{A}', \pi')$ be probability spaces, and let ϕ be a measure preserving map from Ω to Ω'. The map ϕ is called an *isomorphism* if it is a bijection between Ω and Ω' and both ϕ and ϕ^{-1} are measure preserving, and it is called an *isomorphism mod* 0 if there are null sets $N \in \mathcal{A}$ and $N' \in \mathcal{A}'$ such that the restriction of ϕ to $\Omega \setminus N$ is an isomorphism between $\Omega \setminus N$ and $\Omega' \setminus N'$ (equipped with the suitable restrictions of $(\mathcal{A}, \pi)$ and $(\mathcal{A}', \pi')$, respectively). In the last case $(\Omega, \mathcal{A}, \pi)$ and $(\Omega', \mathcal{A}', \pi')$ are called *isomorphic mod* 0.

It turns out that several of our results require a little bit more structure than that of an arbitrary probability space. In particular, we will consider Lebesgue (or standard) spaces, i.e., complete probability spaces that are isomorphic mod 0 to the disjoint union of a closed interval (equipped with the standard Lebesgue sets and Lebesgue measure) and a countable set of atoms.[1]

2.1 Graphons and Graph Densities

We are now ready to introduce the main objects studied in this paper.

[1]See [14], Section 2.2 for an axiomatic definition of Lebesgue spaces, and Section 2.4 for the proof that a probability space is Lebesgue if and only if it is isomorphic mod 0 to the disjoint union of a closed interval and a countable set of atoms. See also [6].

Starting from an arbitrary probability space $(\Omega, \mathcal{A}, \pi)$, let $W : \Omega \times \Omega \to \mathbb{R}$ be a bounded, symmetric function measurable with respect to the completion of $(\Omega\times\Omega, \mathcal{A}\times\mathcal{A}, \pi\times\pi)$. We call the quadruple $H = (\Omega, \mathcal{A}, \pi, W)$ a *graphon*, and refer to W as *a graphon on the probability space* $(\Omega, \mathcal{A}, \pi)$. (As discussed above, such functions can be thought of as limits of convergent graph sequences, which explains the name).

From our point of view, graphons obtained by changing W on a set of measure 0, or changing the σ-algebra $\mathcal{A}$ so that W remains measurable, do not differ essentially from the original. However, for technical reasons we have to distinguish them. We say that a graphon is *strong*, if W is measurable with respect to $\mathcal{A}\times\mathcal{A}$ (not just the completion of it). We can always change W on a set of measure 0 to make the graphon strong (Theorem 3.2(i)).

We say that H is *complete*, if the underlying probability space is complete, and we say that it is *Lebesguian*, if the underlying probability space is a Lebesgue space. The *completion*, $\overline{H}$, of H is obtained by completing the underlying probability space, i.e., by replacing $\mathcal{A}$ by its completion $\overline{\mathcal{A}}$.

Let $H = (\Omega, \mathcal{A}, \pi, W)$ be a graphon, and let F be a finite graph with $V(F) = \{1, \dots, k\}$. The definition (1) then can be extended as

$$t(F,H) = \int\limits_{\Omega^k} \prod_{ij\in E(F)} W(x_i, x_j) \prod_{i=1}^{k} d\pi(x_i). \tag{3}$$

Let $H = (\Omega, \mathcal{A}, \pi, W)$ and $H' = (\Omega', \mathcal{A}', \pi', W')$ be two graphons. The goal of this paper is to determine necessary and sufficient conditions under which

$$t(F,H) = t(F,H') \tag{4}$$

for all graphs F.

To this end, we will introduce two different notions of isomorphism. Both will be expressed in terms of the following operation: given a graphon $H' = (\Omega', \mathcal{A}', \pi', W')$ and a measure preserving map ϕ from a probability space $(\Omega, \mathcal{A}, W)$ into $(\Omega', \mathcal{A}', \pi')$, let $(W')^\phi$ be "pull-back" of W', defined by $(W')^\phi(x,y) = W(\phi(x), \phi(y))$. If $H = (\Omega, \mathcal{A}, \pi, W)$ and $G = (\Gamma, \mathcal{B}, \rho, U)$ are two graphons and $\phi : \Omega \to \Gamma$ is measure preserving from the completion $\overline{\mathcal{A}}$ into $\mathcal{B}$ such that $W = U^\phi$ almost everywhere, then we call ϕ a *weak isomorphism* from H to G. Note that a weak isomorphism is not necessarily invertible.

We say that H and H' are *isomorphic mod 0* (in notation $H' \cong H'$), if there exists a map $\phi : \Omega \to \Omega'$ such that ϕ is an isomorphism mod 0 and $(W')^\phi = W$ almost everywhere in $\Omega \times \Omega$. For simplicity, we often drop the qualifier mod 0.

We call H and H' *weakly isomorphic* if there is a third graphon G and weak isomorphisms from H and H' into G. It will follow from Theorems 3.2 and 2.1 that we could require here that G is a strong Lebesguian graphon.

The isomorphism relation $\cong$ is clearly an equivalence relation, and it will follow from Theorem 2.1 (ii) below that weak isomorphism is an equivalence relation as well. Every graphon is weakly isomorphic with its completion, and every pair of isomorphic graphons is weakly isomorphic. It is clear that if two graphons H and H' are weakly isomorphic then (4) holds for every graph H. Theorem 2.1 (ii) below will show that the converse also holds.

To state our results, we need one more notion, the notion of *twins*. Let $H = (\Omega, \mathcal{A}, \pi, W)$ be a graphon. Two points $x_1, x_2 \in \Omega$ are called twins if $W(x_1, y) = W(x_2, y)$ for almost all $y \in \Omega$. Note that relation of being twins is an equivalence relation. We call the graphon H *almost twin-free* if all there exists a set N of measure zero such that no two point in $\Omega \setminus N$ are twins.

2.2 Main results

With these definitions, we can state our main result:

Theorem 2.1 (i) *If H and H' are almost twin-free Lebesguian graphons, then* (4) *holds for every simple graph F if and only if $H \cong H'$.*

(ii) *If H and H' are general graphons, then* (4) *holds for every simple graph F if and only if H and H' are weakly isomorphic.*

A natural idea of the proof of Theorem 2.1 is the following: can we bring a graphon $(\Omega, \mathcal{A}, \pi, W)$ to a "canonical form", so that isomorphic or weakly isomorphic graphons would have identical canonical forms? In the case of functions in a single variable, this is possible, through "monotonization": for every bounded real function on $[0,1]$ there is an unique monotone increasing left-continuous function on $[0,1]$ that has the same moments.

In Section 4 we'll construct not quite a canonical form, but a "canonical ensemble", a probability distribution (H_α) of graphons on the same σ-algebra such that $H \cong H_\alpha$ for almost all α, and two graphons are isomorphic if and only if their ensembles can be coupled so that corresponding graphons are identical (up to sets of measure 0).

An important element of the proof is a curious measure-theoretic fact. Consider a 2-variable function for which all 1-variable functions obtained by fixing one of the variables are measurable. This of course does not in general imply that the 2-variable function is measurable, but it does imply it in some circumstances (see e.g. Corollary 4.2).

As we will see, the second statement of Theorem 2.1 can easily be deduced from the first. In fact, we'll show that *every graphon is weakly isomorphic to a twin-free Lebesguian graphon.* (See Theorem 3.2 for more details of this isomorphism.)

We can also transform a Lebesguian graphon into a graphon whose underlying probability space is the unit interval with the Lebesgue measure, by "resolving" the atoms into intervals of the appropriate length. This form is the most elementary and therefore useful in applications; however, it is not so convenient for the purposes of this paper because we loose twin-freeness.

It is easy to see that if H and H' are weakly isomorphic, then (4) holds not only for simple graphs F but also for graphs with multiple edges (which we'll call *multigraphs* if we want to emphasize that multiple edges are allowed; but we don't allow loops). Thus (4) for simple graphs implies this equation for multigraphs. (This fact will be an important step in the proof, see Section 5.2.)

We can formulate our results in a probabilistic way. Recall that a *coupling* between two probability spaces $(\Omega, \mathcal{A}, \pi)$ and $(\Omega', \mathcal{A}', \pi')$ is a probability distribution on $\mathcal{A} \times \mathcal{A}'$ whose marginals are π and π', respectively. A coupling between two graphons means a coupling between their underlying probability spaces. Let $H = (\Omega, \mathcal{A}, \pi, W)$ be a graphon, and let $X_1, \dots, X_n$ be independent random samples from π. Then we have

$$t(F, H) = \mathsf{E}\Big(\prod_{ij \in E(F)} W(X_i, X_j) \Big).$$

Let $H = (\Omega, \mathcal{A}, \pi, W)$ and $H' = (\Omega', \mathcal{A}', \pi', W')$ two graphons, and suppose that there exists a coupling γ between them such that $W(X_1, Y_1) = W'(X_2, Y_2)$ holds with probability 1 for two independent samples (X_1, X_2) and (Y_1, Y_2) from γ. In this case clearly (4) holds for every graph F. As we will see, Theorem 2.1 implies that in the Lebesguian case the converse also holds.

We sum up the results for the most important special case of functions in $\mathcal{W}$, i.e., bounded, symmetric functions $W : [0,1]^2 \to \mathbb{R}$ which are measurable with respect to the Lebesgue sets on $[0,1]^2$ (the Corollary would remain valid for arbitrary Lebesguian graphons, but this would not be essentially more general).

Corollary 2.2 *For two functions $W, W' \in \mathcal{W}$ the following are equivalent.*

(a) *For every simple graph F, $t(F, W') = t(F, W)$.*

(b) *For every multigraph F, $t(F, W') = t(F, W)$.*

(c) *There exists a function $U \in \mathcal{W}$ and two measure preserving maps $\varphi, \psi : [0,1] \to [0,1]$ such that $W = U^{\varphi}$ and $W' = U^{\psi}$ almost everywhere.*

(d) *There exist two measure preserving maps $\varphi, \psi : [0,1] \to [0,1]$ such that $(W')^{\varphi} = W^{\psi}$ almost everywhere.*

(e) *There exists a probability measure γ on $[0,1] \times [0,1]$ such that each marginal of γ is the Lebesgue measure, and if (X, X') and (Y, Y') are two independent samples from γ, then $W(X, Y) = W'(X', Y')$) with probability 1.*

2.3 Examples

The property of being twin-free is crucial for Theorem 2.1 (i).

Example 1 Let $\phi_k : [0,1] \to [0,1]$ be the map $\phi_k(x) = kx \pmod 1$. For any function $W \in \mathcal{W}$, the functions W^{ϕ_2} and W^{ϕ_3} define graphons that are weakly isomorphic but in general not isomorphic. Indeed, for a "generic" W (say $W = xy$), every point has two twins in W^{ϕ_2} and three twins in W^{ϕ_3}. The pair of maps in Corollary 2.2 (c) go from W, while in (d), they go into $(W^{\phi_3})^{\phi_2} = (W^{\phi_2})^{\phi_3} = W^{\phi_6}$.

Our next example shows that the Lebesgue property is also needed.

Example 2 Let Ω be a subset of $[0,1]$ with inner Lebesgue measure 0 and outer Lebesgue measure 1, and let Ω' be its complement. Let $\mathcal{A}$ and $\mathcal{A}'$ consist of the traces of Lebesgue measurable sets on Ω and Ω', respectively. Let W and W' be the restrictions of the function xy to $\Omega \times \Omega$ and $\Omega' \times \Omega'$, respectively. The identical embeddings $\varphi : \Omega \to [0,1]$ and $\varphi' : \Omega' \to [0,1]$ are measure preserving, and hence $H = (\Omega, \mathcal{A}, \pi, W)$ and $H' = (\Omega', \mathcal{A}', \pi', W')$ are weakly isomorphic. But for every $x \in \Omega$, we have

$$2 \int_{\Omega} W(x,y)\, d\pi(y) = x \notin \Omega',$$

which shows that there is no way to "match up" the points in Ω and Ω' to get an isomorphism mod 0. The same example shows that conclusions (d), (e) in Corollary 2.2 could not be extended to the non-Lebesgue case either.

3 Isomorphism

The main goal of this section is to describe how a general graphon can be transformed into a twin-free Lebesguian graphon. To this end, we have to recall some basic notions from measure theory (mostly because their usage does not seem standard), and then discuss different "isomorphism-like" mappings between graphons.

3.1 Preliminaries

For a set $\mathcal{S}$ of subsets of a set Ω, we denote by $\sigma(\mathcal{S})$ the σ-algebra generated by $\mathcal{S}$. We call a σ-algebra $\mathcal{A}$ *countably generated* if there is countable set $S \subseteq \mathcal{A}$ such that $\sigma(\mathcal{S}) = \mathcal{A}$. This is equivalent to the existence of a sequence $\mathcal{A}_1 \subseteq \mathcal{A}_2 \subseteq \dots$ of finite σ-algebras whose union generates $\mathcal{A}$.

We say that a set $\mathcal{S} \subseteq \mathcal{A}$ is a *basis* for the probability space $(\Omega, \mathcal{A}, \pi)$, if $\sigma(S)$ is dense in $\mathcal{A}$, i.e., for every $X \in \mathcal{A}$ there is a $Y \in \sigma(\mathcal{S})$ such that $\pi(X \triangle Y) = 0$.

Given sets $A \subset \Omega$ and two points $x, y \in \Omega$, we say that A *separates* x and y if $|\{x, y\} \cap A| = 1$. We say that a set $\mathcal{S}$ of subsets of Ω *separates* x and y if there exists a set $A \in \mathcal{S}$ that separates x and y. This leads to a partition $\mathcal{P}[\mathcal{S}]$ of Ω by placing two points in the same class if and only if they are not separated by $\mathcal{S}$. We say that $\mathcal{S}$ is *separating* if it separates any two points in Ω. We'll say that a graphon is separating if its underlying σ-algebra is separating.

A probability space $(\Omega', \mathcal{A}', \pi')$ is called a *full subspace* of $(\Omega, \mathcal{A}, \pi)$ if Ω' is a (not necessarily measurable) subset of Ω of external measure 1, $\mathcal{A}' = \{A \cap \Omega' \mid A \in \mathcal{A}\}$, and $\pi'(A \cap \Omega')) = \pi(A)$ for all $A \in \mathcal{A}$.

Consider two probability spaces $(\Omega, \mathcal{A}, \pi)$ and $(\Omega', \mathcal{A}', \pi')$ and a measure preserving map $\phi : \Omega \to \Omega'$. The map ϕ is called an *embedding* of the first space into the second if ϕ is an isomorphism between $(\Omega, \mathcal{A}, \pi)$ and a full subspace of $(\Omega', \mathcal{A}', \pi')$. We call ϕ an *embedding* of a graphon $H = (\Omega, \mathcal{A}, \pi, W)$ into a graphon $H' = (\Omega', \mathcal{A}', \pi', W')$ if ϕ is an embedding of $(\Omega, \mathcal{A}, \pi)$ into $(\Omega', \mathcal{A}', \pi')$ and $(W')^\phi = W$ almost everywhere.

Let $(\Omega, \mathcal{A}, \pi)$ be a probability space and $f : \Omega \to \mathbb{R}$, a bounded $\mathcal{A}$-measurable function. Let $\mathcal{A}_0 \subseteq \mathcal{A}$ be a sub-σ-algebra. The *conditional expectation* $\mathsf{E}(f \mid \mathcal{A}_0)$ is the set of all $\mathcal{A}_0$-measurable function f' such that $\int_{A_0} f \, d\pi = \int_{A_0} f' \, d\pi$ for all $A_0 \in \mathcal{A}_0$. It is well known that such functions exist and any two such functions differ only on a set of π-measure 0. We'll write (somewhat sloppily) $f' = \mathsf{E}(f \mid \mathcal{A}_0)$ instead of $f' \in \mathsf{E}(f \mid \mathcal{A}_0)$. We say that f is *almost $\mathcal{A}_0$-measurable*, if there is an $\mathcal{A}_0$-measurable function f' such that $f = f'$ π-almost everywhere. Clearly we must have $f' \in \mathsf{E}(f \mid \mathcal{A}_0)$, and it does not matter which representative of $\mathsf{E}(f \mid \mathcal{A}_0)$ we choose, so (again somewhat sloppily) we can say that f is almost $\mathcal{A}_0$-measurable if and only if $f = \mathsf{E}(f \mid \mathcal{A}_0)$ almost everywhere.

3.2 Push-Forward and Quotients

Let $(\Omega, \mathcal{A}, \pi)$ and $(\Omega', \mathcal{A}', \pi')$ be probability spaces and let $\phi : \Omega \to \Omega'$ be a measure preserving map. We have described how to "pull back" a graphon on $(\Omega', \mathcal{A}', \pi')$ to a (weakly isomorphic) graphon on $(\Omega, \mathcal{A}, \pi)$. It is also possible to "push-forward" a graphon $H = (\Omega, \mathcal{A}, \pi, W)$ to a graphon $(\Omega', \mathcal{A}', \pi', W_\phi)$. This is defined by the requirement that

$$\int_{A_1' \times A_2'} W_\phi(x', y') d\pi'(x') d\pi'(y') = \int_{\phi^{-1}(A_1') \times \phi^{-1}(A_2')} W(x, y)\, d\pi(x)\, d\pi(y) \quad (5)$$

for all $A_1', A_2' \in \mathcal{A}'$. The next lemma states that the "push-forward" W_ϕ is well defined, and that $(W_\phi)^\phi$ is a certain conditional expectation of W.

Lemma 3.1 *Let $(\Omega, \mathcal{A}, \pi)$ and $(\Omega', \mathcal{A}', \pi')$ be probability spaces, let $\phi : \Omega \to \Omega'$ be a measure preserving map, and let W be a graphon on $(\Omega, \mathcal{A}, \pi)$.*

(i) *There exists a bounded, symmetric function $W_\phi : \Omega' \times \Omega' \to \mathbb{R}$ that is $\mathcal{A}' \times \mathcal{A}'$ measurable and satisfies* (6)*. It is unique up to changes on a set of measure zero in $\Omega' \times \Omega'$.*

(ii) *Let $\mathcal{A}_\phi = \phi^{-1}(\mathcal{A}')$. Then $(W_\phi)^\phi = \mathsf{E}(W \mid \mathcal{A}_\phi \times \mathcal{A}_\phi)$ almost everywhere.*

(iii) *If ϕ is an embedding of $(\Omega, \mathcal{A}, \pi)$ into $(\Omega', \mathcal{A}', \pi')$, then $(W_\phi)^\phi = W$ almost everywhere.*

Proof. (i) By linearity, it is easy to see that we can restrict ourselves to the case where W takes values in $[0, 1]$. Define a measure μ on $\mathcal{A}' \times \mathcal{A}'$ by

$$\mu(A_1' \times A_2') = \int_{\phi^{-1}(A_1') \times \phi^{-1}(A_2')} W(x, y)\, d\pi(x)\, d\pi(y)$$

for $A'_1, A'_2 \in \mathcal{A}$. With this definition, we have that

$$0 \le \mu(A'_1 \times A'_2) \le \pi(\phi^{-1}(A'_1))\pi(\phi^{-1}(A'_2)) = (\pi' \times \pi')(A'_1 \times A'_2),$$

implying in particular that μ is absolutely continuous with respect to $\pi' \times \pi'$. Hence the Radon-Nikodym derivative,

$$W_\phi = \frac{d\mu}{d(\pi' \times \pi')}, \tag{6}$$

is well defined. Using the above bound once more, together with the fact that $\mu(A_1 \times A_2) = \mu(A_2 \times A_1)$, we furthermore have that

$$0 \le W_\phi(x,y) \le 1 \qquad \text{and} \qquad W_\phi(x,y) = W_\phi(y,x) \tag{7}$$

almost everywhere. Changing W_ϕ on a set of measure zero, we may assume that these relations hold everywhere. To define W_ϕ for a general bounded function W, we use linearity.

(ii) Let $A_1, A_2 \in \mathcal{A}_\phi$, i.e., let $A_1 = \phi^{-1}(A'_1)$ and $A_2 = \phi^{-1}(A'_2)$ for some $A'_1, A'_2 \in \mathcal{A}'$. By the definition of W_ϕ, the fact that ϕ is measure preserving, and the definition of $(W_\phi)^\phi$, we have that

$$\begin{aligned}\int_{A_1 \times A_2} W(x,y)\, d\pi(x)\, d\pi(y) &= \int_{A'_1 \times A'_2} W_\phi(x',y')\, d\pi'(x')\, d\pi'(y') \\ &= \int_{A_1 \times A_2} W_\phi(\phi(x),\phi(y))\, d\pi(x)\, d\pi(y) \\ &= \int_{A_1 \times A_2} (W_\phi)^\phi(x,y)\, d\pi(x)\, d\pi(y).\end{aligned}$$

This implies that $(W_\phi)^\phi = \mathsf{E}(W \mid \mathcal{A}_\phi \times \mathcal{A}_\phi)$ almost everywhere.

(iii) Since ϕ is an isomorphism between $(\Omega, \mathcal{A}, \pi)$ and a subspace of $(\Omega', \mathcal{A}', \pi')$, we know that given any $A \in \mathcal{A}$, we can find an $A' \in \mathcal{A}'$ such that $\phi(A) = A' \cap \phi(\Omega)$. But then $\phi^{-1}(A') = \phi^{-1}(\phi(A)) = A$, proving that $A \in \mathcal{A}_\phi$. Thus $\mathcal{A}_\phi = \mathcal{A}$, which implies that $(W_\phi)^\phi = W$ almost everywhere. □

We can use the "push-forward" construction to define quotients of graphons. Let $H = (\Omega, \mathcal{A}, \pi, W)$ be a graphon, let $\mathcal{P}$ be an arbitrary partition of Ω into disjoint sets, and for $x \in \Omega$, let $[x]$ denote the class in $\mathcal{P}$ that contains the point x. We then define a graphon $H/\mathcal{P} = (\Omega/\mathcal{P}, \mathcal{A}/\mathcal{P}, \pi/\mathcal{P}, W/\mathcal{P})$ and a measure preserving map $\phi : \Omega \to \Omega/\mathcal{P}$ as follows: the points in $\Omega/\mathcal{P}$ are the classes of the partition $\mathcal{P}$, ϕ is the map $\phi : x \mapsto [x]$, $\mathcal{A}/\mathcal{P}$ is the σ-algebra consisting of the sets $A' \subset \Omega/\mathcal{P}$ such that $\phi^{-1}(A') \in \mathcal{A}$, and $(\pi/\mathcal{P})(A') := \pi(\phi^{-1}(A'))$. Then ϕ is measure preserving, and the function $W/\mathcal{P} = W_\phi$ is defined by (5).

3.3 Reductions

Now we are able to state the theorem that allows us to reduce every graphon to a twin-free Lebesguian graphon.

Theorem 3.2 (i) *Let $H = (\Omega, \mathcal{A}, \pi, W)$ be a graphon. Then one can change the value of W on a set of $\pi \times \pi$-measure 0 to get a strong graphon.*

(ii) *Let $H = (\Omega, \mathcal{A}, \pi, W)$ be a graphon. Then there exists a countably generated σ-algebra $\mathcal{A}_0 \subset \mathcal{A}$ such that W is $(\mathcal{A}_0 \times \mathcal{A}_0)$-measurable.*

(iii) *Let $H = (\Omega, \mathcal{A}, \pi, W)$ be a graphon. Then the graphon $H/\mathcal{P}[\mathcal{A}]$ is separating. If H is countably generated, then so is $H/\mathcal{P}[\mathcal{A}]$.*

(iv) *Let $H = (\Omega, \mathcal{A}, \pi, W)$ be a separating graphon on a probability space with a countable basis. Then the completion of H can be embedded into a Lebesguian graphon.*

(v) *Let $H = (\Omega, \mathcal{A}, \pi, W)$ be a graphon, and let $\mathcal{P}$ be the partition into the twin-classes of H. Then $H/\mathcal{P}$ is almost twin-free. If H is Lebesguian, then $H/\mathcal{P}$ is Lebesguian as well. Furthermore, the projection $H \to H/\mathcal{P}$ is a weak isomorphism.*

Corollary 3.3 *Every graphon has a weak isomorphism into a strong Lebesguian graphon.*

The proof of this theorem (which is not hard, but technical) will be given in the rest of this section.

3.3.1 Making a graphon strong

Let $H = (\Omega, \mathcal{A}, \pi, W)$ be a graphon, and let $W' = \mathsf{E}(W \mid \mathcal{A} \times \mathcal{A})$. Then W' is $\mathcal{A} \times \mathcal{A}$-measurable, and changing W' on a set of measure 0, we may assume that W' is symmetric and bounded. Moreover, $\int_{A \times A'} (W' - W) = 0$ for all $A, A' \in \mathcal{A}$, which implies that $\int_S (W' - W) = 0$ for all sets S in the completion of $\mathcal{A} \times \mathcal{A}$, so $W = W'$ almost everywhere. These observations prove part (i) of the Theorem.

3.3.2 Countable generation

We prove a simple lemma, which implies Theorem 3.2(ii), and will also be used at several other places (Sections 4.1 and 5.2).

Lemma 3.4 *Let $(\Omega, \mathcal{A})$ and $(\Omega', \mathcal{A}')$ be measurable spaces, and let $W : \Omega \times \Omega' \to \mathbb{R}$ be a bounded, $(\mathcal{A} \times \mathcal{A}')$-measurable function. Then there exist countably generated σ-algebras $\mathcal{A}_0 \subset \mathcal{A}$ and $\mathcal{A}'_0 \subset \mathcal{A}'$ such that W is $(\mathcal{A}_0 \times \mathcal{A}'_0)$-measurable.*

Proof. Let $\mathcal{C}$ be the set of bounded, $(\mathcal{A} \times \mathcal{A}')$-measurable functions W for which the statement of the lemma is true. The set $\mathcal{C}$ is clearly a vector space that contains the constant function 1 as well as the indicator functions of all rectangles $A \times B$ with $A \in \mathcal{A}$ and $B \in \mathcal{A}'$. If is further not hard to show that if (W_k) is a sequence of non-negative functions in $\mathcal{C}$ and $W_k \uparrow W$ for a bounded function W, then the limiting function W is in $\mathcal{C}$ as well. By the monotone class theorem (see, e.g., Theorem 3.14 in [15]), we conclude that $\mathcal{C}$ contains all bounded functions which are measurable with respect to the σ-algebra generated by the rectangles $A \times B$, i.e., the σ-algebra $\mathcal{A} \times \mathcal{A}'$. $\square$

3.3.3 Merging inseparable elements

If we identify elements in the same class of the partition $\mathcal{P}[\mathcal{A}]$, we get a σ-algebra which is isomorphic under the obvious map. This implies (iii) of Theorem 3.2.

3.3.4 Lebesgue property

Consider a separating graphon $H = (\Omega, \mathcal{A}, \pi, W)$, and assume that $\mathcal{A}$ is generated by the countable set $\mathcal{S}$. Then $\mathcal{S}$ is a basis for the completion of $(\Omega, \mathcal{A}, \pi)$. We invoke the fact (see e.g. [14], Section 2.2) that any separating complete probability space with a countable basis can be embedded into a Lebesgue space. Thus there exists an embedding ψ of the completion of $(\Omega, \mathcal{A}, \pi)$ into a Lebesgue space $(\Omega', \mathcal{L}', \lambda')$. Let W' be the push-forward of W, $W' = W_\psi$. By Lemma 3.1, we have that $(W')^\psi = W$ almost everywhere, which shows that ψ is an embedding of the completion of H into $(\Omega', \mathcal{L}', \lambda', W')$. This proves part (iv) of Theorem 3.2.

3.3.5 Partitions into Twin-Classes

We prove (v) in Theorem 3.2. We may assume that $\mathcal{A}$ is countably generated. ndeed, by Lemma 3.4, we can replace $\mathcal{A}$ by a countably generated σ-algebra $\mathcal{A}_0$. This does not change the relation of being twins: Two points $x, x' \in \Omega$ are twins if and only if the set $A_{x,x'} = \{y \in \Omega : W(x,y) = W(x,y')\}$ has measure 1. Since W is measurable with respect to $\mathcal{A}_0 \times \mathcal{A}_0$, the set $A_{x,x'}$ lies in $\mathcal{A}_0 \subset \mathcal{A}$, implying that x and x' are twins with respect to H if and only if they are twins with respect to H_0.

Let $\mathcal{A}_\mathcal{P}$ consists of those sets in $\mathcal{A}$ that do not separate any pair of twin points. Clearly $\mathcal{A}_\mathcal{P}$ is a σ-algebra.

Claim 1 *W is almost $\mathcal{A}_\mathcal{P} \times \mathcal{A}_\mathcal{P}$-measurable.*

Let $\widetilde{W} = \mathsf{E}(W \mid \mathcal{A}_\mathcal{P} \times \mathcal{A}_\mathcal{P})$. We want to prove that

$$\int_{A\times B} W(x,y)d\pi(x)d\pi(y) = \int_{A\times B} \widetilde{W}(x,y)d\pi(x)d\pi(y) \tag{8}$$

for all $A, B \in \mathcal{A}$. Define the functions

$$g_A = \mathsf{E}(1_A \mid \mathcal{A}_\mathcal{P}), \qquad U_A(y) = \int_A W(x,y)d\pi(x),$$

and

$$V_A(x) = \int W(x,y)g_A(y)d\pi(y).$$

Since $U_A(y) = U_A(z)$ if y, z are twins, the function U_A is $\mathcal{A}_\mathcal{P}$-measurable, similarly for V_A, and obviously for g_A. Repeatedly using the fact that $\int fg = \int f\mathsf{E}(g \mid \mathcal{A}_0)$ if f is $\mathcal{A}_0$-measurable, this implies

$$\begin{aligned}
\int_{A\times B} W(x,y)d\pi(x)d\pi(y) &= \int 1_B(y)U_A(y)d\pi(y) = \int g_B(y)U_A(y)d\pi(y) \\
&= \int V_B(x)1_A(x)d\pi(x) = \int V_B(x)g_A(x)d\pi(x) \\
&= \int W(x,y)g_A(x)g_B(y)d\pi(x)d\pi(y) \\
&= \int \widetilde{W}(x,y)g_A(x)g_B(y)d\pi(x)d\pi(y) \\
&= \int_{A\times B} \tilde{W}(x,y)d\pi(x)d\pi(y).
\end{aligned}$$

(where the last equality follows since $\widetilde{W}$ is $\mathcal{A}_\mathcal{P} \times \mathcal{A}_\mathcal{P}$-measurable). This implies (8) and completes the proof of Claim 1.

Let $\widetilde{W} = \mathsf{E}(W \mid \mathcal{A}_\mathcal{P} \times \mathcal{A}_\mathcal{P})$ as before, then $H_\mathcal{P} = (\Omega, \mathcal{A}_\mathcal{P}, \pi, \widetilde{W})$ is a graphon, which is clearly weakly isomorphic to $(\Omega, \mathcal{A}, \pi, W)$. Let N be the set of points $x \in \Omega$ for which $\{y \in \Omega : \widetilde{W}(x,y) \neq W(x,y)\}$ has positive measure. Then clearly N is a null set, and two points $x, x' \in \Omega \setminus N$ are twins in H if and only if they are twins in $H_\mathcal{P}$. The graphon $H/\mathcal{P}$ is obtained from $H_\mathcal{P}$ by identifying indistinguishable elements, which implies that $H/\mathcal{P}$ is twin-free.

To prove that $H/\mathcal{P}$ is Lebesguian if H is Lebesguian, we invoke the fact (established in Section 3.2 of [14]) that $(\Omega/\mathcal{P}, \mathcal{A}/\mathcal{P}, \pi/\mathcal{P})$ is a Lebesgue space provided $(\Omega, \mathcal{A}, \pi)$ is a Lebesgue space and there exists a countable set $\mathcal{S} \subseteq \mathcal{A}$ that separates two points if and only they are in different partition classes.

To construct such a set $\mathcal{S}$, let $\mathcal{T}$ be a countable set generating $\mathcal{A}$, closed under finite intersections. For $A \in \mathcal{A}$ and $x \in \Omega$, let

$$\mu_x(A) = \int_A W(x,y)d\pi(y).$$

Since W is a bounded $\mathcal{A} \times \mathcal{A}$-measurable function, the function $A \mapsto \mu_x(A)$ is a finite measure for all $x \in \Omega$, while the function $x \mapsto \mu_x(A)$ is a $\mathcal{A}$-measurable function on Ω for all $A \in \mathcal{A}$.

By definition, $x, x' \in \Omega$ are twins iff the set $\{y \in \Omega : W(x,y) = W(x,y')\}$ has measure zero. This is equivalent to the condition that $\mu_x(A) = \mu_{x'}(A)$ for all $A \in \mathcal{A}$. Since the measure $\mu_x(\cdot)$ on $\mathcal{A}$ is uniquely determined by the sets in $\mathcal{T}$, we have that x and x' are twins if and only if $\mu_x(T) = \mu_{x'}(T)$ for all $T \in \mathcal{T}$.

For every $T \in \mathcal{T}$ and rational number r, consider the sets $S_{T,r} = \{x \in \Omega : \ \mu_x(T) \geq r\}$. There is a countable number of these. Furthermore, if x and x' are twins, then they belong to exactly the same sets $S_{T,r}$; if they are not twins, then there is a $T \in \mathcal{T}$ such that $\mu_x(T) \neq \mu_{x'}(T)$, and for any rational number between $\mu_x(T)$ and $\mu_{x'}(T)$, the set $S_{T,r}$ separates x and x'.

This completes the proof of Theorem 3.2.

3.4 Isomorphism and Weak Isomorphism

We conclude this section with relating isomorphism and weak isomorphism.

Lemma 3.5 *Let $H_i = (\Omega_i, \mathcal{A}_i, \pi_i, W_i)$ be graphons with the Lebesgue property ($i = 1, 2$), and let $\phi : \ \Omega_1 \to \Omega_2$ be measure-preserving. If H_1 is almost twin-free, and $W_1 = W_2^\phi$ almost everywhere, then ϕ is an isomorphism mod 0, so in particular $H_1 \cong H_2$.*

Proof. Let

$$\Omega_1' = \{x \in \Omega_1 : \ W_2(\phi(x), \phi(y)) = W_1(x,y) \text{ for almost all } y\},$$

and let $N_1 = \Omega_1 \setminus \Omega_1'$. Then $\pi_1(N_1) = 0$ by Fubini and our assumption that $W_1 = W_2^\phi$ almost everywhere.

Let N_1' be a nullset such that all twin-classes of H_1 have at most one point in $\Omega_1 \setminus N_1'$, and let ϕ' to be the restriction of ϕ to $\Omega_1' \setminus N_1'$. Then ϕ' is injective: indeed, if $x_1, x_2 \in \Omega_1' \setminus N_1'$ and $\phi(x_1) = \phi(x_2)$, then $W_1(x_1, y) = W_2(\phi(x_1), \phi(y)) = W_2(\phi(x_2), \phi(y)) = W_1(x_2, y)$ for almost all y by the definition of Ω_1', hence x_1 and x_2 are twins, a contradiction. As shown in [14], Section 2.5, an injective measure preserving map between Lebesgue spaces has a measurable inverse defined almost everywhere. This implies that $\phi' : \Omega_1' \setminus N_1' \to \Omega_2$ is an isomorphism mod 0, which shows that ϕ is an isomorphism mod 0 as well. □

Corollary 3.6 *If two twin-free graphons with the Lebesgue property are weakly isomorphic, then they are isomorphic.*

4 Canonical Ensembles

We could try to construct a "canonical form" of a graphon by assigning "tags" to the points in Ω. For example, we could tag a point x with its marginal $d(x) = \int W(x,y)\,d\pi(y)$, or by the sequence of marginals of higher powers of W. This, however, would not work: for example, there could be a transitive group of measure-preserving permutations of Ω leaving W invariant, and then all points would still have the same tag.

To break the symmetry, we select an infinite sequence $\alpha = (a_1, a_2, \dots)$ of points in Ω, which we call *anchor points.* Now we can tag each point $x \in \Omega$ with the sequence

$$\Phi_\alpha(x) = (W(x, a_1), W(x, a_2), \dots) \in [0,1]^{\mathbb{N}} \tag{9}$$

(where we assume that $0 \le W \le 1$) The map $x \mapsto \Phi_\alpha(x)$ defines a measurable map from Ω into $[0,1]^{\mathbb{N}}$ (with respect to the standard Borel σ-algebra $\mathcal{L}$ on $[0,1]^{\mathbb{N}}$), which in turn defines a measure λ_α on the sets $S \in \mathcal{L}$ by

$$\lambda_\alpha(S) = \pi(\Phi_\alpha^{-1}(S)), \tag{10}$$

and a graphon W_{Φ_α} on $([0,1]^{\mathbb{N}}, \mathcal{L}, \lambda_\alpha)$ by (5). We denote the completion of $([0,1]^{\mathbb{N}}, \mathcal{L}, \lambda_\alpha, W_{\Phi_\alpha})$ by H_α.

We will show that if $\alpha_1, \alpha_2, \dots$ are taken i.i.d. at random with distribution π then with probability one, then H_α is isomorphic mod 0 to the original graphon H (see Section 4.2 for details). So using an infinite sequence of independent random points as anchor points, the tags of the points contain all information about the points.

These tags are almost canonical, except for the choice of the sequence α. So instead of a canonical form, we get a "canonical ensemble", a probability distribution (H_α) of graphons such that $H \cong H_\alpha$ for almost all α, and two graphons are isomorphic if and only if their ensembles can be coupled so that corresponding graphons are isomorphic.

To prove Theorem 2.1 (i), we will therefore have to show that if H and H' satisfy (4), then we can "couple" the choice of anchor points α in H and β in H' so that $H_\alpha \cong H'_\beta$, thus yielding an isomorphism of H and H'. This second step in the proof will be carried out in Section 5.3.

4.1 Measure theoretic preparation

The next technical lemma will be important in the construction of "canonical ensembles".

Lemma 4.1 *Let $(\Omega, \mathcal{A}, \pi)$ and $(\Omega', \mathcal{A}', \pi')$ be probability spaces, and let $W : \Omega \times \Omega' \to \mathbb{R}$ be a bounded $\mathcal{A} \times \mathcal{A}'$-measurable function. Let $Y_1, Y_2, \dots$ be independent random points from Ω'. Let $\mathcal{A}_0 \subseteq \mathcal{A}$ be the (random) σ-algebra generated by the functions $W(\cdot, Y_k)$. Then with probability 1, W is almost $\mathcal{A}_0 \times \mathcal{A}'$-measurable.*

Proof. By Lemma 3.4, we may assume that $\mathcal{A}$ and $\mathcal{A}'$ are countably generated. Let $\mathcal{A}'_1 \subset \mathcal{A}'_2 \subset \dots$ and $\mathcal{A}'_1 \subset \mathcal{A}'_2 \subset \dots$ be a sequence of finite σ-algebras with $\sigma(\cup_n \mathcal{A}_n) = \mathcal{A}$ and $\sigma(\cup_n \mathcal{A}'_n) = \mathcal{A}'$, and let P'_n denote the partition of Ω' into the atoms of $\mathcal{A}'_n$. For $y \in S \in P'_n$ with $\pi'(S) > 0$, define

$$U_{n,m}(x,y) = \frac{1}{m\pi(S)} \sum_{\substack{j\le m\\ Y_j\in S}} W(x,Y_j)$$

We define $U_{n,m}(x,y) = 0$ if $y \in S \in P'_n$ with $\pi'(S) = 0$.

First we prove that for every $n \geq 1$, every $A \in \mathcal{A}$ and $A' \in \mathcal{A}'_n$, we have with probability 1

$$\int\limits_{A\times A'} U_{n,m}\, d\pi\, d\pi' \longrightarrow \int\limits_{A\times A'} W\, d\pi\, d\pi' \qquad (m\to\infty). \tag{11}$$

It suffices to prove this in the case when $A' = S \in P'_n$ and $\pi'(S) > 0$. Then for every $y_0 \in A'$, we have

$$\int\limits_{A} U_{n,m}(x,y_0)\, d\pi(x) = \frac{1}{m\pi(S)} \sum_{\substack{j\le m\\ Y_j\in S}} \int\limits_{A} W(x,Y_j)\, d\pi(x).$$

hence by the Law of Large Numbers,

$$\int\limits_{A} U_{n,m}(x,y_0)\, d\pi(x) \longrightarrow \frac{1}{\pi(S)} \int\limits_{A\times S} W\, d\pi\, d\pi' \qquad (m\to\infty).$$

Since both sides are independent of $y_0 \in S$, integrating over $y_0 \in S$ equation (11) follows.

The number of choices of n, $A \in \cup_k \mathcal{A}_k$ and $A' \in \mathcal{A}'_n$ is countable, and hence it follows that with probability 1, (11) holds for all $n \geq 1$, every $A \in \cup_k \mathcal{A}_k$ and $A' \in \mathcal{A}'_n$. Since $\cup_k \mathcal{A}_k$ is dense in $\mathcal{A}$, this implies that (11) holds for all $n \geq 1$, every $A \in \mathcal{A}$ and $A' \in \mathcal{A}'_n$.

From now on, we suppose that the choice of the Y_i is such that this holds.

For a fixed n, the indices m have a subsequence $m_1 < m_2 < \dots$ such that U_{n,m_j} converges to some function U_n in the weak-$*$-topology of $L_\infty(\mathcal{A}_0 \times \mathcal{A}'_n)$. Hence by (11),

$$\int\limits_{A\times A'} U_n\, d\pi\, d\pi' = \lim_{j\to\infty} \int\limits_{A\times A'} U_{n,m_j}\, d\pi\, d\pi' = \int\limits_{A\times A'} W\, d\pi\, d\pi'$$

for all $n \geq 1$, every $A \in \mathcal{A}$ and $A' \in \mathcal{A}'_n$. Thus U_n is a representative of $\mathsf{E}(W \mid \mathcal{A} \times \mathcal{A}'_n)$. Since U_n is $\mathcal{A}_0 \times \mathcal{A}'_n$ measurable, it is also a representative of $\mathsf{E}(W \mid \mathcal{A}_0 \times \mathcal{A}'_n)$. This shows that for every $n \geq 1$ we have

$$\mathsf{E}(W \mid \mathcal{A} \times \mathcal{A}'_n) = \mathsf{E}(W \mid \mathcal{A}_0 \times \mathcal{A}'_n) \tag{12}$$

almost everywhere.

By Levy's Upward Theorem, the left hand side of (12) tends to $\mathsf{E}(W \mid \mathcal{A} \times \mathcal{A}') = W$ almost everywhere. The right hand side of (12) tends to $\mathsf{E}(W \mid \mathcal{A}_0 \times \mathcal{A}')$ almost everywhere, so $W = \mathsf{E}(W \mid \mathcal{A}_0 \times \mathcal{A}')$ almost everywhere, which proves the Lemma. □

We formulate a couple of corollaries, the first of which is immediate:

Corollary 4.2 *Let $(\Omega, \mathcal{A}, \pi)$ and $(\Omega', \mathcal{A}', \pi')$ be probability spaces, let $W : \Omega \times \Omega' \to \mathbb{R}$ be a bounded function that is measurable with respect to $\mathcal{A} \times \mathcal{A}'$, and let $\mathcal{A}_0 \subset \mathcal{A}$ be a sub-σ-algebra. If $W(\cdot, y)$ is $\mathcal{A}_0$-measurable for almost all $x \in \Omega$, then W is almost $\mathcal{A}_0 \times \mathcal{A}'$-measurable.*

Corollary 4.3 *Let $(\Omega, \mathcal{A}, \pi, W)$ be a graphon, and let $X_1, X_2, \ldots$ be independent random points from Ω. Let $\mathcal{A}_0 \subseteq \mathcal{A}$ be the (random) σ-algebra generated by the functions $W(\cdot, X_k)$. Then with probability 1, W is almost $\mathcal{A}_0 \times \mathcal{A}_0$-measurable.*

Proof. Let $\mathcal{A}_1$ denote the σ-algebras generated by the functions $W(\cdot, X_{2k})$. Clearly $\mathcal{A}_1 \subseteq \mathcal{A}_0$. By Lemma 4.1, W is almost $\mathcal{A}_1 \times \mathcal{A}$ measurable with probability 1, so we can change it on a set of measure 0 to get an $\mathcal{A}_1 \times \mathcal{A}$ measurable function W'. Let $\mathcal{A}'_2$ be the σ-algebras generated by the functions $W'(X_{2k+1}, \cdot)$. Applying the lemma again, we get that W' is almost $\mathcal{A}_1 \times \mathcal{A}'_2$ measurable. With probability 1, each function $W(X_{2k+1}, \cdot)$ differs from $W'(X_{2k+1}, \cdot)$ on a set of measure 0 only (since the X_{2k+1} are independent of $\mathcal{A}_1$), and so $\mathcal{A}'_2 \subseteq \sigma(\mathcal{A}_0)$. So W' is $\mathcal{A}_0 \times \sigma(\mathcal{A}_0)$ measurable, which implies that W', and hence W, are almost $\mathcal{A}_0 \times \mathcal{A}_0$ measurable. □

4.2 Anchor Sequences

Let us consider the σ-algebra $\mathcal{L}$ on $[0,1]^{\mathbb{N}}$ generated by the sets $A_1 \times A_2 \times \dots$, where each A_i is a Borel subset of $[0,1]$ and only a finite number of factors A_i are different from $[0,1]$. Fix a graphon $H = (\Omega, \mathcal{A}, \pi, W)$ with $0 \le W \le 1$. For every $\alpha \in \Omega^{\mathbb{N}}$, the map $\Phi_\alpha : \Omega \to [0,1]^{\mathbb{N}}$ defined by (9) is measurable, and (10) defines a probability measure on $\mathcal{L}$ with respect to which Φ_α is measure preserving. Thus (6) leads to a symmetric, $\mathcal{L} \times \mathcal{L}$-measurable function $W_{\Phi_\alpha} : [0,1]^{\mathbb{N}} \times [0,1]^{\mathbb{N}} \to [0,1]$ which we denote by W_α. We say that $\alpha \in \Omega^{\mathbb{N}}$ is *regular* if $W = W_\alpha^{\Phi_\alpha}$ almost everywhere.

Lemma 4.4 *Almost all $\alpha \in \Omega^{\mathbb{N}}$ are regular.*

Proof. Let $\mathcal{A}_\alpha$ denote the σ-algebra of subsets of Ω of the form $\Phi_\alpha^{-1}(A)$, where $A \in \mathcal{L}$. Note that $\mathcal{A}_\alpha \subseteq \mathcal{A}$ by the fact that Φ_α is measurable. Further, almost by definition, $\mathcal{A}_\alpha$ is the smallest sub-σ-algebra of $\mathcal{A}$ such that all the functions $W(\cdot, \alpha_i)$ are measurable. As a consequence, we may apply Lemma 4.3 to conclude that for almost all α, W is almost $\mathcal{A}_\alpha \times \mathcal{A}_\alpha$)-measurable, which by Lemma 3.1 gives that $W = W_\alpha^{\Phi_\alpha}$ almost everywhere. □

Let $\mathcal{L}_\alpha$ be the completion of $\mathcal{L}$ with respect to λ_α. Then $([0,1]^{\mathbb{N}}, \mathcal{L}_\alpha, \lambda_\alpha)$ is a complete, Polish space and hence Lebesgue, so $H_\alpha = ([0,1]^{\mathbb{N}}, \mathcal{L}_\alpha, \lambda_\alpha, W_\alpha)$ defines a Lebesguian graphon.

Lemma 4.5 *Let H be a twin free graphon with the Lebesgue property. If α is regular, then Φ_α is an isomorphism mod 0 and $H_\alpha \cong H$.*

Proof. By (10), Φ_α is a measure preserving map from $(\Omega, \mathcal{A}, \pi)$ into $([0,1]^{\mathbb{N}}, \mathcal{L}, \lambda_\alpha)$. Since $(\Omega, \mathcal{A}, \pi)$ is complete, Φ_α is measurable (and measure preserving) from $(\Omega, \mathcal{A}, \pi)$ into $([0,1]^{\mathbb{N}}, \mathcal{L}_\alpha, \lambda_\alpha)$ as well. By the definition of a regular α, $H_\alpha^{\phi_\alpha} = H$ almost everywhere, and by Lemma 3.5, Φ_α is an isomorphism mod 0. □

5 Coupling

5.1 Partially Labeled Graphs and Marginals

We recall some notions from [9]. A *partially labeled graph* is a finite graph in which some of the nodes are labeled by different nonnegative integers. Two partially labeled graphs are *isomorphic*, if there is a label-preserving isomorphism between them. A *k-labeled graph* is a partially labeled graph with labels $1, \dots, k$.

Let F_1 and F_2 be two partially labeled graphs. Their *product* F_1F_2 is defined as follows: we take their disjoint union, and then identify nodes with the same label (retaining the labels, and any multiple edges which this might create). For two unlabeled graphs, F_1F_2 is their disjoint union. Clearly this multiplication is associative and commutative.

Let $H = (\Omega, \mathcal{A}, \pi, W)$ be a graphon, and let $\alpha = (a_0, a_1, \dots)$ be an infinite sequence of points in Ω. Let F be a partially labeled graph with nodes $V(F) = \{1, \dots, k\}$, where nodes $1, \dots, r$ are labeled by distinct nonnegative integers $\ell_1, \dots, \ell_r$. Let $X_i = a_{\ell_i}$ for $1 \le i \le r$, and let $X_{r+1}, \dots, X_k \in \Omega$ be independent points from the distribution π. Define

$$t_\alpha(F, H) = \mathsf{E}\Big(\prod_{ij \in E(F)} W(X_j, X_j) \Big).$$

Of course, this value only depends on those elements of α whose subscripts occur as labels, and we'll sometimes omit the tail of α if it contains no labels. For example, if F is a 2-labeled triangle, then

$$\begin{aligned} t_{a_1 a_2}(F, H) = t_\alpha(F, W) &= \mathsf{E}(W(a_1, a_2) W(a_2, X) W(a_1, X)) \\ &= \int\limits_\Omega W(a_1, a_2) W(a_2, x) W(a_1, x) \, d\pi(x). \end{aligned}$$

It is easy to see that if F_1 and F_2 are two k-labeled graphs, then

$$t(F_1F_2, H) = \int\limits_{\Omega^k} t_{x_1 \dots x_k}(F_1, W) t_{x_1 \dots x_k}(F_2, W) \, d\pi(x_1) \dots d\pi(x_k).$$

5.2 Multiple Edges

Lemma 5.1 *Let $H = (\Omega, \mathcal{A}, \pi, W)$ and $H' = (\Omega', \mathcal{A}', \pi', W')$ be two graphons, and assume that $t(F,H) = t(F,H')$ for every simple graph F. Then $t(F,H) = t(F,H')$ for every multigraph F.*

Proof. We use induction on the number of parallel edges in F. Suppose that F has two nodes, say i and j, connected by more than one edge. Let F_k denote the multigraph obtained from F by subdividing one of these edges by $k-1$ new nodes. Let F' denote the multigraph obtained by removing one copy of the edge ij. So $F_1 = F$, but for $k > 1$, F_k has fewer parallel edges than F, and so we may assume that

$$t(F_k, H) = t(F_k, H')$$

holds for every $k \geq 2$. We consider all the multigraphs F_k and F' as 2-labeled graphs, with i and j labeled 1 and 2.

Since F_k can be thought of as the product of F' and a path P_{k+1} with $k+1$ nodes (the endpoints labeled), we can write

$$t(F,H) = \int_{\Omega^2} W(x,y) t_{xy}(F', H) \, d\pi(x) \, d\pi(y),$$

and

$$t(F_k,H) = \int_{\Omega^2} t_{xy}(P_{k+1}, H) t_{xy}(F', H) \, d\pi(x) \, d\pi(y).$$

The first factor inside the integral can be expressed as

$$t_{xy}(P_{k+1}, H) = \int_{\Omega^k} W(x, x_1) \cdots W(x_{k-1} y) \, d\pi(x_1) \ldots d\pi(x_{k-1}),$$

which we can recognize as k-th power of the kernel W as an integral operator.

At this point, it will be useful to assume that H and H' are countably generated graphons (this can be done without loss of generality by Lemma 3.4). As a consequence, W is an integral operator on the separable Hilbert space

$L_2(\Omega, \mathcal{A}, \pi)$, and since W is bounded, this implies that W is Hilbert-Schmidt and thus compact, which in turn implies that W has a spectral representation:

$$W(x,y) \sim \sum_{n=0}^{\infty} \lambda_n \varphi_n(x)\varphi_n(y). \tag{13}$$

It follows that for every $k \geq 2$,

$$t_{xy}(P_{k+1}, H) = \sum_{n=0}^{\infty} \lambda_n^k \varphi_n(x)\varphi_n(y),$$

and hence

$$t(F_k, H) = \sum_{n=0}^{\infty} \lambda_n^k \int_{\Omega^2} \varphi_n(x)\varphi_n(y) t_{xy}(F', H)\, d\pi(x)\, d\pi(y).$$

Similarly, let

$$W'(x,y) \sim \sum_{i=0}^{\infty} \mu_n \psi_n(x)\psi_n(y)$$

be the spectral representation of W', then we get that for every $k \geq 2$,

$$0 = t(F_k, H) - t(F_k, H') = \sum_{n=0}^{\infty} a_n \lambda_n^k - b_n \mu_n^k, \tag{14}$$

where

$$a_n = \int_{\Omega^2} \varphi_n(x)\varphi_n(y) t_{xy}(F', U)\, d\pi(x)\, d\pi(y)$$

and

$$b_n = \int_{(\Omega')^2} \psi_n(x)\psi_n(y) t_{xy}(F', W)\, d\pi(x)\, d\pi(y)$$

are independent of k. (The integrals exist since $t_{x,y}(F', H)$ is a bounded function of x and y.) It follows that in (14) everything must cancel, in other words, for every value c,

$$\sum \{a_n :\ \lambda_n = c\} = \sum \{b_n :\ \mu_n = c\}$$

(it is known that the sums on both sides have a finite number of terms, since the multiplicities of the eigenvalues are finite).

Now while (13) may not be true with equality, the "trace" with any other kernel gives an equation; in particular,

$$t(F,H) = \sum_{n=0}^{\infty} \lambda_n \int_{\Omega^2} W(x,y) t_{xy}(F',H)\, d\pi(x)\, d\pi(y) = \sum_{n=0}^{\infty} a_n \lambda_n,$$

and similarly

$$t(F,H') = \sum_{n=0}^{\infty} b_n \mu_n,$$

which shows that $t(F,H) = t(F,H')$ as claimed. □

It will be convenient to assume that $0 \le W, W' \le 1$. If this does not hold, we can apply a linear transformation to the values of the functions, to get two functions W_0 and W_0' with $0 \le W_0, W_0' \le 1$. Expanding the product in the definition (3), $t(F, W_0)$ can be written as a linear combination of the values $t(F', W)$, where F' is a subgraph of F. Thus $t(F,W) = t(F,W')$ for every graph F if and only if $t(F,W_0) = t(F,W_0')$ for every graph F (where "graph" could mean either simple graph or multigraph). So (4) holds for W_0 and W_0' if and only if it holds for W and W'. If we prove that this implies $(\Omega, \mathcal{A}, \pi, W_0) \cong (\Omega', \mathcal{A}', \pi', W_0')$, then $H \cong H'$ follows trivially.

5.3 Coupling Anchor Sequences

Consider two graphons $H = (\Omega, \mathcal{A}, \pi, W)$ and $H' = (\Omega', \mathcal{A}', \pi', W')$ satisfying the conditions in Theorem 2.1 (i) and $0 \le W, W' \le 1$. Given two "anchor" sequences $\alpha = (a_1, a_2, \dots)$ from Ω and $\beta = (b_1, b_2, \dots)$ from Ω', let $H_\alpha = ([0,1]^{\mathbb{N}}, \mathcal{L}_\alpha, \lambda_\alpha, W_\alpha)$ and $H'_\beta = ([0,1]^{\mathbb{N}}, \mathcal{L}'_\beta, \lambda'_\beta, W'_\beta)$. We would like to select α and β in such a way that $\lambda_\alpha = \lambda'_\beta$ and $W_\alpha = W'_\beta$ almost everywhere. This will complete the proof of the theorem. By Lemma 4.4, we can guarantee that both α and β are regular by selecting $a_1, a_2, \dots$ as well as $b_1, b_2 \dots$ independently and uniformly from π and π', respectively; however, the equality of W_α and W'_β will only be true if we couple α and β carefully.

The condition on the coupling is described in the following lemma.

Lemma 5.2 *Let $H = (\Omega, \mathcal{A}, \pi, W)$ and $H' = (\Omega', \mathcal{A}', \pi', W')$ be two graphons, and let $\alpha = (a_1, a_2, \dots)$ and $\beta = (b_1, b_2, \dots)$ be regular sequences for H and H', respectively. Suppose that for every partially labeled multigraph F,*

$$t_\alpha(F, H) = t_\beta(F, H').$$

Then $\lambda_\alpha = \lambda'_\beta$ and $W_\alpha = W'_\beta$ almost everywhere (with respect to $\lambda_\alpha = \lambda'_\beta$).

Proof. First, we show that $\lambda_\alpha = \lambda'_\beta$. These probability measures are defined on the σ-algebra $\mathcal{L}$ as the distribution measures of the random variables $W(X, a_1), W(X, a_2), \dots)$ and $W'(Y, b_1), W'(Y, b_2), \dots)$, where X and Y are random points from π and π', respectively. By Lemma 6.1 it therefore suffices to prove that these random variables have the same mixed moments.

Let $(k_1, k_2, \dots)$ be a sequence of nonnegative integers, of which only a finite number is nonzero; say $k_i = 0$ for $i > m$. Then

$$\mathsf{E}(\prod_i W(X, a_i)^{k_i}) = t_\alpha(F, H),$$

where F is the star on $m + 1$ nodes, with the endnodes labeled $1, \dots, m$, and the edge between the center and endnode i replaced by k_i parallel edges. Similarly,

$$\mathsf{E}(\prod_i W'(Y, b_i)^{k_i}) = t_\beta(F, H').$$

These numbers are equal by the hypothesis of the Lemma. This proves that $\lambda_\alpha = \lambda'_\beta$.

Second, we show that $W_\alpha(x, y) = W'_\beta(x, y)$ for almost all $x, y \in [0,1]^{\mathbb{N}}$. It suffices to show that the random variables $Z_1 = (X, Y, W_\alpha(X, Y))$ and $Z_2 = (X, Y, W'_\beta(X, Y))$ (with values from $[0,1]^{\mathbb{N}} \times [0,1]^{\mathbb{N}} \times [0,1]$) have the same distribution, where X and Y are independent points in $(\Omega^{\mathbb{N}}, \lambda_\alpha)$.

We can generate Z_1 by choosing independent uniform random points X' and Y' from Ω, and letting $X = \Phi_\alpha(X')$ and $Y = \Phi_\beta(Y')$. Since α is regular, we have that

$$W_\alpha(X,Y) = W(X',Y')$$

with probability one, and hence

$$Z_1 = (W(X',a_1), W(X',a_2), \dots, W(Y',a_1), W(Y',a_2), \dots, W(X',Y')).$$

Similarly, we have

$$Z_2 = (W'(X'',b_1), W'(X'',b_2), \dots, W'(Y'',b_1), W'(Y'',b_2), \dots, W'(X'',Y'')),$$

where X'' and Y'' are independent random points from π'. To prove that Z_1 and Z_2 have the same distribution, it again suffices to prove that they have the same mixed moments.

A particular mixed moment is given by nonnegative integers $(k_1, k_2, \dots)$, $(l_1, l_2, \dots)$ and m (of which only a finite number is nonzero; say $k_i = l_i = 0$ for $i > n$). Let us define the multigraph F as follows. F has two unlabeled nodes v_x and v_y, and n further nodes labeled $1, \dots, n$. We connect v_x to i by k_i edges, v_y to i by l_i edges ($i = 1, \dots, n$), and v_x to v_y by m edges. Then

$$\begin{aligned}\mathsf{E}\big(W(X',a_1)^{k_1} \cdots W(X',a_n)^{k_n} W(Y',a_1)^{k_1} \\ \cdots W(Y',a_n)^{k_n} W(X',Y')\big) = t_\alpha(F,H).\end{aligned}$$

and similarly

$$\begin{aligned}\mathsf{E}\big(W'(X'',b_1)^{k_1} \cdots W'(X'',b_n)^{k_n} W'(Y'',b_1)^{k_1} \\ \cdots W'(Y'',b_n)^{k_n} W'(X,Y)\big) = t_\beta(F,H').\end{aligned}$$

These two numbers are the same by hypothesis. This completes the proof of the Lemma. □

To prove Theorem 2.1, we next show:

Lemma 5.3 *Let $H = (\Omega, \mathcal{A}, \pi, W)$ and $H' = (\Omega', \mathcal{A}', \pi', W')$ be two Lebesguian graphons such that*

$$t(F,H) = t(F,H').$$

for every multigraph F. Then we can couple sequences $\alpha \in \Omega^{\mathbb{N}}$ with sequences $\beta \in \Omega'^{\mathbb{N}}$ so that if $\alpha, \beta)$ is a sequence from this joint distribution, then

$$t_\alpha(F, H) = t_\beta(F, H').$$

holds almost surely for every partially labeled multigraph F.

Proof. Let $\mathcal{F}_k$ be the set of k-labeled multigraphs. We define recursively a coupling of sequences $\alpha \in \Omega^k$ with sequences $\beta \in \Omega'^k$ so that $t_{\alpha'}(F, H) = t_{\beta'}(F, H')$ holds almost surely for every $F \in \mathcal{F}_k$. Let $(a_1, \dots, a_k)$ and $(b_1, \dots, b_k)$ be chosen from this coupled distribution. Consider two random points X from π and Y from π', and the random variables

$$A = (t_{a_1 \dots a_k X}(F, H) : \; F \in \mathcal{F}_{k+1})$$

and

$$B = (t_{b_1 \dots b_k Y}(F, H) : \; F \in \mathcal{F}_{k+1})$$

with values in $[0, 1]^{\mathcal{F}_{k+1}}$. We claim that the variables A and B have the same distribution. It suffices to show that A and B have the same mixed moments. Consider any moment of A; in other words, let $F_1, \dots, F_m \in \mathcal{F}_{k+1}$, let $q_1, \dots, q_m$ be nonnegative integers, and let $F_i^{q_i}$ be obtained from F_i by replacing each edge in F_i by q_i edges. Then the corresponding moment of A is

$$\mathsf{E}\Big(\prod_{i=1}^{m} t_{a_1 \dots a_k X}(F_i, H)^{q_i}\Big) = \mathsf{E}\big(t_{a_1 \dots a_k X}(F_1^{q_1} \dots F_m^{q_m}, H)\big) = t_{a_1 \dots a_k}(F, H),$$

where the multigraph F is obtained by unlabeling the node labeled $k+1$ in the multigraph $F_1^{q_1} \dots F_m^{q_m}$. Expressing the moments of B in a similar way, we see that they are equal by the induction hypothesis. This proves that A and B have the same distribution.

Using Lemma 6.2 it follows that we can couple the variables X and Y so that $A = B$ with probability 1. In other words, we can replace X and Y by a random variable $(X', Y') \in \Omega \times \Omega'$ so that X' has distribution π, Y' has distribution π', and their joint distribution satisfies

$$t_{a_1 \dots a_k X'}(F, H) = t_{b_1 \dots b_k Y'}(F, H')$$

for every $F \in \mathcal{F}_{k+1}$ with probability 1. Thus we have extended the coupling to $\Omega^k \times \Omega'^k$.

It is clear that this sequence of couplings defines a coupling of $\Omega^{\mathbb{N}}$ with $\Omega'^{\mathbb{N}}$ as claimed. □

5.4 Conclusion of proofs

Proof of Theorem 2.1. Part (i) follows easily: if we choose random sequences (α, β) from the coupled distribution given by Lemma 5.3, then these sequences will be regular with probability 1, and so they satisfy the conditions of Lemma 5.2.

To prove (ii), suppose that $H = (\Omega, \mathcal{A}, \pi, W)$ and $H' = (\Omega', \mathcal{A}', \pi', W')$ satisfy (4) for every simple graph F. By Corollary 3.3, we can find twin-free Lebesguian graphons $G = (\Gamma, \mathcal{B}, \rho, U)$ and $G' = (\Gamma', \mathcal{B}', \rho', U')$ and weak isomorphisms ϕ and ϕ' from H and H' to G and G', respectively. It follows by Theorem 2.1(i) that the G and G' are isomorphic mod 0, so in particular $U = (U')^{\psi'}$ almost everywhere for some measure preserving map $\psi' : \Gamma \to \Gamma'$. Defining $\psi : \Omega \to \Gamma'$ by $\psi(x) = \psi'(\phi(x))$, we conclude that $W = (U')^{\psi}$ almost everywhere. The maps ψ and ϕ' are measure preserving from the completions $\overline{H}$ and $\overline{H'}$ into G'. □

Proof of Corollary 2.2. The equivalence of (a), (b) and (c) follows by Theorem 2.1 (ii) and the fact that a function which is measurable with respect to the completion of $\mathcal{L} \times \mathcal{L}$ is almost everywhere equal to a function which is measurable with respect to $\mathcal{L} \times \mathcal{L}$. In the proof of (c), Theorem 2.1 may give a graphon containing atoms, but it is easy to replace these atoms by intervals of appropriate length.

To prove that (c)$\Longrightarrow$(e), assume that φ, ψ and U exist as in (c). Let $X, X' \in [0, 1]$ be independent random points from the uniform distribution

λ on $[0,1]$. Since φ and ψ are measure preserving, $\varphi(X)$ and $\psi(Y)$ have the same distribution, and hence by Lemma 6.2 there is a coupling measure γ on $[0,1]\times[0,1]$ with marginals λ such that if (X,X') is a random sample from γ, then $\varphi(X)=\psi(X')$ with probability 1. So if (X,X') and (Y,Y') are independent random points from γ, then

$$W(X,Y) = U(phi(X), phi(Y)) = U'(psi(X'), psi(Y')) = W'(X',Y').$$

To prove that (e)$\Longrightarrow$(d), consider the projections $\Phi,\Psi:\ [0,1]^2\to[0,1]$ defined by $\Phi(x,x')=x$ and $\Psi(x,x')=x'$. Then

$$W^\Phi((X,X'),(Y,Y')) = W(X,Y)$$

and

$$(W')^\Psi((X,X'),(Y,Y')) = W'(X',Y')$$

Thus, $W^\Phi=(W')^\Psi$ almost everywhere. Furthermore, Φ and Ψ are measure preserving if we consider the coupling measure γ on $[0,1]$.

Since the completion of $([0,1]^2,\mathcal{L}_2,\gamma)$ is a Lebesgue space, we can find a measure preserving map $\rho:\ ([0,1],\lambda)\to([0,1]^2,\gamma)$. Setting $\varphi=\Phi\circ\rho$ and $\psi=\Psi\circ\rho$, we obtain the desired measure preserving maps $\varphi,\psi:[0,1]\to[0,1]$ such that $W^\varphi=(W')^\psi$ almost everywhere.

Finally, (d)$\Rightarrow$(a) is trivial. □

Acknowledgement

We are grateful to Miklós Laczkovich, Ron Peled, Yuval Peres and Oded Schramm for many useful discussions on the topic of this paper, and to Kati Vesztergombi and Svante Janson for carefully reading an earlier version and suggesting several improvements.

References

[1] N. Alon, W. Fernandez de la Vega, R. Kannan and M. Karpinski: Random sampling and approximation of MAX-CSPs, *J. Comput. System Sci.* **67** (2003), 212–243.

[2] C. Borgs, J. Chayes, L. Lovász, V.T. Sós, K. Vesztergombi, unpublished, 2004.

[3] C. Borgs, J. Chayes, L. Lovász, V.T. Sós, K. Vesztergombi: Counting graph homomorphisms, in: *Topic in Discrete Mathematics* (Klazar, Kratochvil, Loebl, Matousek, Thomas, Valtr, eds.) Springer, Berlin–Heidelberg (2006), 315–371.

[4] C. Borgs, J. Chayes, L. Lovász, V.T. Sós, K. Vesztergombi: Convergent Sequences of Dense Graphs I: Subgraph Frequencies, Metric Properties and Testing, preprint (2006), `http://research.microsoft.com/~borgs/Papers/ConvMetric.pdf`

[5] C. Borgs, J.T. Chayes, L. Lovász, V.T. Sós, and K. Vesztergombi: Convergent Graph Sequences II. H-Colorings, Statistical Physics and Quotients, manuscript (2006)

[6] D.L. Cohn: *Measure Theory*, Birkheuser, Boston, 1980.

[7] P. Erdös, L. Lovász, J. Spencer: Strong independence of graphcopy functions, in: *Graph Theory and Related Topics*, Academic Press, 165-172.

[8] Feller, W.: *An Introduction to Probability Theory and its Applications,* Vol. II, Second edition, Wiley, New York, 1971.

[9] M. Freedman, L. Lovász, A. Schrijver: Reflection positivity, rank connectivity, and homomorphism of graphs *Journal of The American Mathematical Society* (to appear)

[10] O. Goldreich, S. Goldwasser and D. Ron: Property testing and its connection to learning and approximation, *J. ACM* **45** (1998), 653–750.

[11] P.R. Halmos: *Measure Theory*, Graduate Texts in Mathematics **18**, Springer, New York, Heidelberg, Berlin (1991).

[12] L. Lovász and B. Szegedy: Limits of dense graph sequences, Microsoft Research Technical Report TR-2004-79.

[13] L. Lovász, B. Szegedy: Szemerédi's Lemma for the analyst, Microsoft Research Technical Report MSR-TR-2005-90,

`ftp://ftp.research.microsoft.com/pub/tr/TR-2005-90.pdf`

[14] V. A. Rohlin: On the fundamental ideas of measure theory, Translations of the American Mathematical Society, Series 1, Vol. 10, 1-54 (1962). The Russian original appeared in Mat. Sb. 25, 107-150 (1949).

[15] D. Williams: *Probability with Martingales*, Cambridge University Press, (1991).

6 Appendix: Moments and coupling of probability distributions

In this section we prove some probability theory lemmas, that are "well known" but not easy to reference. We start with the fact that if two vector valued random variables have the same mixed moments, then they have the same distribution (cf. Feller [8], Problem XV.9.21).

Lemma 6.1 *Let* $(\Omega, \mathcal{A}, \pi)$ *and* $(\Omega', \mathcal{A}', \pi')$ *be probability spaces, and let* $f : \Omega \to [0,1]^{\mathbb{N}}$ *and* $g : \Omega' \to [0,1]^{\mathbb{N}}$ *be measurable functions, with* $f(x) = (f_1(x), f_2(x), \dots)$ *and* $g(y) = (g_1(y), g_2(y), \dots)$*. If*

$$\int f_1(x)^{k_1} \dots f_n(x)^{k_n} \, d\pi(x) = \int g_1(y)^{k_1} \dots g_n(y)^{k_n} \, d\pi'(y)$$

for every finite sequence of nonnegative integers $k_1, \ldots, k_n$*, then* $\pi(f^{-1}(B)) = \pi'(g^{-1}(B))$ *for every Borel set* $B \subseteq [0,1]^{\mathbb{N}}$*.*

Proof. It suffices to prove that $\pi(f^{-1}(B)) = \pi'(g^{-1}(B))$ for every Borel set of the form $B = I_1 \times I_2 \times \ldots I_n \times [0,1] \times \ldots$, where $I_1, \ldots, I_n$ are intervals. Let $p_{j,m}(x)$ be a polynomial that approximates the indicator function $\mathbf{1}_{I_j}$ on $[0,1]$ in L_1 with error less than $1/m$ ($j = 1, \ldots, n$). Then

$$\int_\Omega p_{1,m}(f_1(x)) \cdots p_{n,m}(f_n(x))\, dx \longrightarrow \int_\Omega \mathbf{1}_{I_1}(f_1(x)) \cdots \mathbf{1}_{I_n}(f_n(x))\, dx$$
$$= \int_{f^{-1}(B)} 1\, dx = \pi(f^{-1}(B)) \qquad (m \to \infty).$$

Similarly,

$$\int_\Omega p_{1,m}(g_1(x)) \cdots p_{n,m}(g_n(x))\, dx \longrightarrow \pi(g^{-1}(B)) \qquad (m \to \infty).$$

But the left hand sides of these two relations are equal for all m, which proves the Lemma. □

We need the following natural fact about coupling.

Lemma 6.2 *Assume that* $(\Omega, \mathcal{A}, \pi)$ *and* $(\Omega', \mathcal{A}, \pi')$ *are Lebesgue spaces, and* $(\Gamma, \mathcal{B}, \rho)$*, a countably generated separating space. Let* $f : \Omega \to \Gamma$ *and* $g : \Omega' \to \Gamma$ *be measure preserving maps. Then there exists a coupling* ν *of* $(\Omega, \mathcal{A}, \pi)$ *and* $(\Omega', \mathcal{A}, \pi')$ *such that*

$$\nu\big\{(x,y) :\; f(x) = g(y)\big\} = 1.$$

Proof. For $A \in \mathcal{A}$, consider the measure $\lambda^A(B) = \pi(A \cap f^{-1}(B))$ defined for $B \in \mathcal{B}$, and its Radon-Nikodym derivative $f^A = d\lambda^A/d\rho$. Since $\lambda^A \le \pi(f^{-1}(B)) = \rho(B)$, this derivative exists, and $0 \le f^A \le 1$ almost everywhere. Furthermore, $f^\emptyset = 0$ and $f^\Omega = 1$ almost everywhere.

Similarly, for $C \in \mathcal{A}'$, define $\mu^C(B) = \pi'(C \cap g^{-1}(B))$ and $g^C = d\mu^C/d\rho$. Finally, let

$$\nu(A \times C) = \int f^A g^C \, d\rho. \tag{15}$$

Clearly

$$\nu(A \times C) \le \int f^A \, d\rho = \pi(A),$$

and similarly $\nu(A \times C) \le \pi'(C)$. Hence in particular $\nu(A \times C) = 0$ if either $\pi(A) = 0$ or $\pi'(C) = 0$.

Claim 2 *If $A_i \in \mathcal{A}$, $C_i \in \mathcal{A}'$ $(i \in I)$ and the sets $A_i \times C_i$ form a (finite or countably infinite) partition of $A \times C$ $(A \in \mathcal{A}$, $C \in \mathcal{A}')$, then $\sum_i \nu(A_i \times C_i) = \nu(A \times C)$.*

It is easy to see that if $A_1, A_2 \in \mathcal{A}$ are disjoint sets and $A = A_i \cup A_2$, then $f^{A_1} + f^{A_2} = f^A$ almost everywhere. It follows that for every $C \in \mathcal{A}'$, we have $\nu(A_1 \times C) + \nu(A_2 \times C) = \nu(A \times C)$. This implies by standard arguments that the claim holds if $|I|$ is finite. This in turn implies that ν extends to a finitely additive measure on the algebra $\mathcal{F}$ of sets that can be written as the union of a finite number of product sets $A \times C$ $(A \in \mathcal{A}$, $C \in \mathcal{A}')$.

In the case of infinite $|I|$, it follows that $\sum_i \nu(A_i \times C_i) \le \nu(A \times C)$; in fact, for every finite $J \subseteq I$, we have $\cup_{i \in J} A_i \times C_i \subseteq A \times C$, and hence by the finite additivity of ν, we have

$$\sum_{i \in J} \nu(A_i \times C_i) = \nu\Big(\cup_{i \in J} A_i \times C_i\Big) \le \nu(A \times C).$$

Since this holds for every finite subset J of I, it also holds for I.

Suppose that there is a partition where $\{A_i \times C_i : \; i = 1 \in \mathbb{N}\}$ of $A \times C$ and an $\varepsilon > 0$ for which

$$\sum_i \nu(A_i \times C_i) < \nu(A \times C) - 4\varepsilon$$

on a set B of positive measure. Now we use that $(\Omega, \mathcal{A}, \pi)$ and $(\Omega', \mathcal{A}, \pi')$ are Lebesgue spaces, so we may assume that they are intervals $[0, a]$ and $[0, b]$

respectively, together with a countable set of atoms. Thinking of the atoms as converging to a from above, we have a compact topology on them. For every i, we can find an open sets $U_i \supseteq A_i$ and $V_i \supseteq C_i$ such that $\pi(U_i) \le \pi(A_i) + \varepsilon 2^{-i}$ and $\pi'(V_i) \le \pi'(C_i) + \varepsilon/2^i$. Also, we can find closed sets $U \subseteq A$ and $V \subseteq C$ such that $\pi(U) \ge \pi(A) - \varepsilon$ and $\pi'(V) \ge \pi'(C) - \varepsilon$. Then

$$\begin{aligned}\nu(U_i \times V_i) &\le \nu(A_i \times C_i) + \nu((U_i \setminus A_i) \times C_i) + \nu(U_i \times (V_i \setminus C_i)) \\ &\le \nu(A_i \times C_i) + \pi(U_i \setminus A_i) + \pi'(V_i \setminus C_i) \le \nu(A_i \times C_i) + 2\varepsilon 2^{-i}.\end{aligned}$$

It follows similarly that

$$\nu(U \times V) \ge \nu(A \times C) - 2\varepsilon.$$

Hence

$$\sum_i \nu(U_i \times V_i) \le \sum_i \nu(A_i \times C_i) + 2\varepsilon < \nu(A \times C) - 2\varepsilon \le \nu(U \times V).$$

The open sets $U_i \times V_i$ cover the compact set $U \times V$, and so a finite number of them also covers. But the contradicts the finite additivity of ν which we already established.

Claim 3 *The setfunction ν extends to a measure on $\mathcal{A} \times \mathcal{A}'$.*

We have seen already that ν extends to $\mathcal{F}$; it follows by Claim 2 that this extension is σ-additive. Thus the Claim follows by the Measure Extension Theorem.

Define $\Delta = \{(x, y) \in \Omega \times \Omega' : f(x) = g(y)\}$. To complete the proof of the Lemma, we want to prove that ν is a coupling between $(\Omega, \mathcal{A}, \pi)$ and $(\Omega', \mathcal{A}, \pi')$ (which is trivial), and that $\nu(\Omega \times \Omega' \setminus \Delta) = 0$. Let $\mathcal{S} \subseteq \mathcal{B}$ be a countable family separating the elements of Γ. Then

$$\Omega \times \Omega' \setminus \Delta = \bigcup_{S \in \mathcal{S}} f^{-1}(S) \times g^{-1}(\Gamma \setminus S) \cup \bigcup_{S \in \mathcal{S}} f^{-1}(\Gamma \setminus S) \times g^{-1}(S).$$

Consider any term here, say $f^{-1}(S) \times g^{-1}(\Gamma \setminus S) = A \times C$. Then

$$\nu(A \times C) = \int f^A g^C \, d\rho = \int_S + \int_{\Gamma \setminus S}.$$

Here

$$\int_S f^A g^C \, d\rho \leq \int_S g^C \, d\rho = \mu^C(S) = \pi'(g^{-1}(\Gamma \setminus S) \cap g^{-1}(S)) = 0,$$

and similarly

$$\int_{\Omega \setminus S} f^A g^C \, d\rho = 0.$$

This proves that $\nu(\Omega \times \Omega' \setminus \Delta) = 0$. □