FRACTIONAL ISOMORPHISM OF GRAPHONS

JAN GREBÍK AND ISRAEL ROCHA

Abstract. In this paper we work out the theory of fractional isomorphism of graphons as a generalization to the classical theory of fractional isomorphism of finite graphs. The generalization is given in terms of Markov operators on a Hilbert space and it is characterized in terms of iterated degree distributions, homomorphism density of trees, weak isomorphism of a conditional expectation with respect to invariant sub-$\sigma$-algebras and isomorphism of certain quotients of given graphons. Our proofs use a weak version of the mean ergodic theorem, and correspondences between objects such as Markov projections, sub-$\sigma$-algebras, conditional expectation, etc. That also provides an alternative proof for the characterizations of fractional isomorphism of graphs without the use of Birkhoff–von Neumann Theorem.

1. Introduction

In this paper we introduce the notion fractional isomorphism of graphons, which is the suitable counterpart for fractional isomorphism of graphs. Graphons, introduced by Borgs, Chayes, Lovász, Sós, Szegedy, and Vesztergombi [11], [3], [4], emerged as limit objects in the theory of dense graph limits. We refer the reader to the book [12] for the detailed treatment of the subject.

Our aim in the introduction is to recall all equivalent concepts that describes the notion of fractional isomorphism for finite graphs and introduce the corresponding counterparts for graphons. Here, we follow [16]. If we view the theory of graphons in the intersection of combinatorics and analysis, fruitfully contributing to both, then we must admit that the concepts that we are about to introduce lie more on the analytic side. Therefore, we decided to spend more time in the introduction discussing the motivations and translations between the finite world and measurable world. We hope that the ideas sketched in the introduction would make the concept of fractional isomorphism of graphons understandable for people coming from the combinatorial side without a strong background in analysis. Besides, the combination of the introduction with the classical results that we collect in the appendix should be enough to understand why and how the concepts interact. We will see that many difficulties appear when trying to extend the definitions and proving meaningful characterizations. Before we discuss that in the next sections, let us try to convince the reader of the importance of this concept for the isomorphism problem at the same time that we give some basics of fractional isomorphism of graphs.

We start by recalling that two finite graphs $G$ and $H$ are isomorphic if there is a bijection $\varphi$ between $V(G)$ and $V(H)$ that fully preserves the graph structure, i.e., $\varphi$ sends edges to edges and non-edges to non-edges. Deciding if such a bijection exists, i.e., if two given graphs $G$ and $H$ are isomorphic, is a notorious difficult problem in computer science. It is not known if this problem can be solved in polynomial time nor to be NP-complete. Nevertheless, many relaxations of the isomorphism problem have been investigated, one in particular has theoretical and practical interest; the fractional isomorphism decision problem which can be resolved in polynomial time. We note that $G$ and $H$ are isomorphic if there is a permutation matrix $P$ such that $AP = PB$ where $A$ (resp. $B$) is the adjacency matrix of $G$ (resp. $H$). It is clear that every permutation matrix $P$ is positive and the sum of entries of each row and column are equal to 1, or equivalently $P \geq 0$ and $P1 = PT1 = 1$. Having this in mind we say that finite graphs $G$ and $H$ are fractionally isomorphic if there is a doubly stochastic matrix $S$ such that $AS = SB$. By a doubly stochastic matrix we mean a matrix $S$ such that $S \geq 0$ and $S1 = ST1 = 1$. It follows directly from the definition that fractional isomorphism is a relaxation of isomorphism. Besides, it is easy to see that every doubly stochastic matrix is a square matrix and therefore $G$ and $H$ can be fractionally isomorphic only if $|V(G)| = |V(H)|$.

A general attempt to characterize the isomorphism type of a given graph is to find an invariant that is easy to compute and that would distinguish graphs that are not isomorphic. For example, the assignment $v \mapsto \deg_G(v)$ distinguishes graphs that have different degrees of vertices. Or the assignment...
$v \mapsto \{\deg_G(w) : w \in N(v)\}$ describes the degrees of the neighbors, thus providing a finer information. Iterating this construction yields the so-called **iterated degree sequence** $D(G)$. Formally we define

$$d_1(G) = \{\deg_G(v) : v \in V(G)\} \quad \text{and} \quad d_1(v) = \{\deg_G(w) : w \in N(v)\}$$

and then inductively

$$d_{k+1}(G) = \{d_k(v) : v \in V(G)\} \quad \text{and} \quad d_{k+1}(v) = \{d_k(w) : w \in N(v)\}.$$  

Finally, we put $D(G) = (d_k(G))_{k \in \mathbb{N}}$. It turns out that this sequence is tightly related to fractional isomorphism of graphs. Namely, the iterated degree sequence is used to construct the **color refinement algorithm**, a simple and efficient heuristic to test whether two graphs are isomorphic. The algorithm computes a coloring of the vertices of two graphs based on its iterated degree sequences and compare its colorings. Whenever they are different, we say that color refinement distinguishes the graphs. Whenever they are the same, we do not know whether or not they are isomorphic. Fortunately, the results from Babai, Erdős, and Selkow in the paper [1] imply that color refinement distinguishes almost all non-isomorphic graphs, and in practice this algorithm performs well. Other advanced graph isomorphism algorithms and almost all practical isomorphism softwares uses color refinement underneath. This heuristic goes beyond isomorphism testing and is also useful in a number of other problems (see [2] for further reading). Noticeable, and most relevant for our investigation, is that color refinement does not distinguish $G$ and $H$ if and only if $G$ and $H$ are fractionally isomorphic, which was proved by Tinhofer [18, 19]. That suggests the importance of the concept of fractional isomorphism.

There are additionally other compelling interpretations of this concept which are given by different characterizations. Before we summarize the most important equivalences, we need to introduce equitable partitions and homomorphism vectors. An **equitable partition** of $G$ is a sequence $\mathcal{C} = \{C_j\}_{j < s}$ that is a partition of $V(G)$, i.e., $\bigcup_{j < s} C_j = V(G)$, and $\deg_G(v_0, C_j) = \deg_G(v_1, C_j)$ for every $i, j < s$ such that $v_0, v_1 \in C_i$. That is to say that each induced subgraph $G[C_i]$ must be regular and each of the bipartite graphs $G[C_i, C_j]$ must be biregular. The parameters of $\mathcal{C}$ are given by a pair $(n, \mathcal{C})$, where $n$ is a $s$-dimensional vector and $C$ is $s \times s$ square matrix such that $n(j) = |C_j|$ and $C(i, j) = \deg_G(v, C_j)$, for some $v \in C_i$. That is, the parameters of $\mathcal{C}$ are the numerical information that we can read from $\mathcal{C}$. If $G$ and $H$ admit equitable partitions $\mathcal{C}$ and $\mathcal{D}$ that can be indexed in such a way that the parameters of $\mathcal{C}$ and $\mathcal{D}$ are the same, then we say that $G$ and $H$ have a **common equitable partition**. We say that a partition $\mathcal{C}$ is **coarser** than a partition $\mathcal{D}$ if every element of $\mathcal{D}$ is a subset of some element of $\mathcal{C}$. It is not hard to verify that every finite graph admits a **coarsest equitable partition**, i.e., equitable partition that is coarser than any other equitable partition.

Let $\text{Hom}(F, G)$ be the set of all homomorphism from $F$ to $G$, i.e., number of maps $\varphi : V(F) \rightarrow V(G)$ that sends edges to edges. A result ofLovasz [10] from 1967 states that $G$ and $H$ are isomorphic if and only if

$$|\text{Hom}(F, G)| = |\text{Hom}(F, H)|,$$

where $F$ runs over all finite graphs. This homomorphism vectors will decades later appear in the core of the theory of graph limits, making a bridge between functional analysis and graph theory. Notice that it is natural to ask how the notion of isomorphism changes if we require the homomorphism vectors to be the same only on a fixed subclass of finite graphs. We will see that this is an important idea behind fractional isomorphism.

Having recalled the concepts we are ready to summarize the relevant characterizations for the purposes of this article. We include the references for the corresponding definitions and proofs.

**Theorem 1.1.** Let $G$ and $H$ be finite graphs. Then the following are equivalent:

1. $|\text{Hom}(T, H)| = |\text{Hom}(T, G)|$ for every finite tree $T$ [5].
2. $D(G) = D(H)$, i.e., $G$ and $H$ have the same iterated degree sequences [18, 19].
3. $G$ and $H$ have a common coarsest equitable partition [18].
4. $G$ and $H$ are fractionally isomorphic.
5. $G$ and $H$ have some common equitable partition [19].

The main contribution of this paper is to introduce a counterpart to the notion of fractional isomorphism for graphons and to obtain counterpart to Theorem 1.1. Before we introduce analogous concepts for graphons we recall the basic setting. A graphon is a symmetric measurable function $W : X \times X \rightarrow [0, 1]$ where $(X, \mathcal{B})$ is a standard Borel space endowed with a Borel probability measure $\mu$. We write $W_0$ for the space of all graphons after identifying graphons that are equal almost everywhere. This makes $W_0$ a standard Borel space endowed with a Borel probability measure $\mu$. We write $W_0$ for the space of all graphons after identifying graphons that are equal almost everywhere. This makes $W_0$
a subset of $L^\infty(X \times X, \mu \times \mu)$ or $L^2(X \times X, \mu \times \mu)$ and one may consider the distance inherited on $W_0$ from the corresponding norms. However, the most relevant notion of distance for studying graphons as dense graph limits comes from the cut-norm and is defined as

$$d_2(W, U) = \sup_{A,B \subseteq X} \int_{A \times B} |W - U| \, d\mu \times \mu$$

where the supremum runs over all measurable subsets $A, B$ of $X$. The cut-distance $\delta_2$ is then defined as

$$\delta_2(W, U) = \inf_{\varphi} d_2(W^\varphi, U),$$

where $W^\varphi(x, y) = W(\varphi(x), \varphi(y))$ and the infimum runs over all $\varphi : X \to X$ measurable measure preserving bijections of $X$. Considering $W^\varphi$ and $W$ the same is the measurable analogue of considering two finite graphs the same if they are isomorphic. However, in the qualitative version given by $d_2$ we might get $\delta_2(W, U) = 0$ while there is no single $\varphi$ such that $W^\varphi = U$. Therefore, we say that $W$ and $U$ are isomorphic if we have $\varphi$ such that $W^\varphi = U$ for some measurable measure preserving bijection $\varphi : X \to X$. Besides, we say that $W$ and $U$ are weakly isomorphic if $\delta_2(W, U) = 0$. Notice that it follows that $\delta_2$ is only a pseudometric on $W_0$. We write $\bar{W}$ for the weak equivalence class of the graphon $W$ and $\bar{W}_0$ for the quotient space $W_0$ modulo weak equivalence. It is easy to see that $\bar{\delta}_2$ might be considered as distance on $\bar{W}_0$ and it is one of the main result in the theory of graphons that $(\bar{W}_0, \bar{\delta}_2)$ is a compact metric space (See [11]).

An equivalent description of convergence in the space $\bar{W}_0$ can be obtained via homomorphism densities. Let $F$ and $G$ be finite graphs. The homomorphism density of $F$ in $G$ is defined as

$$t(H, G) = \frac{|\text{Hom}(H, G)|}{|V(G)|^{|V(H)|}}.$$

That is, $t(H, G)$ is the probability that a random map is a homomorphism. Note that the notion is invariant under isomorphism. The analogous notion for graphons is defined as

$$t(F, W) = \int_{g \in \mathcal{X}^{V(F)}} \prod_{v, w \in E(F)} W(g(v), g(w)) \, d\mu^{\otimes |V(F)|}(y)$$

and it is not hard to see that $t(F, W) = t(F, U)$ whenever $W$ and $U$ are weakly equivalent. Remarkably, the authors of [11] prove the equivalence between two types of convergence: a sequence of graphons $W_n$ converges to $W$ in the cut-distance if and only if for every finite graph $F$ we have $t(F, W_n) \to t(F, W)$.

Let $G$ be a finite graph. The natural way to represent $G$ as a graphon is to enumerate $V(G) = \{v_0, \ldots, v_{k-1}\}$, take a partition of $X$ into $k$-many equimeasurable pieces, say $\{A_i\}_{i<k}$, and then define the corresponding graphon $W_G$ that takes values only in the set $\{0, 1\}$ as

$$W_G(x, y) = 1 \iff (\exists i, j < k) \, x \in A_i, y \in A_j, \quad \{v_i, v_j\} \in E(G).$$

We have $t(F, G) = t(F, W_G)$. However, once the quantities are normalized we lose the information about $|V(G)|$. What we mean is that there are finite graphs $G$ and $H$ such that $W_G$ and $W_H$ are weakly isomorphic and $|V(G)| \neq |V(H)|$. That shows a dissimilarity with Lovász’s theorem about homomorphism vectors. A similar phenomena must necessarily occur in our approach to fractional isomorphism of graphons, since it will be a coarser equivalence relation than weak isomorphism. In this case, this contrasts with the fact that all conditions in Theorem 1.1 imply that $|V(G)| = |V(H)|$. However, as in the case of weak equivalence we show that two graphs $G$ and $H$ of the same size are fractionally isomorphic as graphons if and only if they are fractionally isomorphic as graphs.

Now we introduce the counterparts of the concepts that appear in Theorem 1.1. Throughout the text it is convenient to view a graphon $W$ as a self-adjoint integral kernel with the corresponding integral operator $T_W : L^2(X, \mu) \to L^2(X, \mu)$ defined as

$$T_W(f)(x) = \int_X W(x, y)f(y) \, d\mu(y),$$

see for example [12, Section 7.5]. It is a standard fact that $T_W$ is a Hilbert-Schmidt operator, i.e., compact with countable spectrum.

Let $Y$ be a standard Borel space with a probability measure $\iota$. We follow [6] and say that an operator $S : L^2(X, \mu) \to L^2(Y, \iota)$ is a Markov operator if

- $S \geq 0$, i.e., $S(f) \geq 0$ whenever $f \geq 0$,
- $S(1_X) = 1_Y$,
- $S^*(1_Y) = 1_X$. 

The authors of [11, 3] prove the equivalence between two types of convergence: a sequence of graphons $W_n$ converges to $W$ in the cut-distance if and only if for every finite graph $F$ we have $t(F, W_n) \to t(F, W)$.

Now we introduce the counterparts of the concepts that appear in Theorem 1.1. Throughout the text it is convenient to view a graphon $W$ as a self-adjoint integral kernel with the corresponding integral operator $T_W : L^2(X, \mu) \to L^2(X, \mu)$ defined as

$$T_W(f)(x) = \int_X W(x, y)f(y) \, d\mu(y),$$

see for example [12, Section 7.5]. It is a standard fact that $T_W$ is a Hilbert-Schmidt operator, i.e., compact with countable spectrum.

Let $Y$ be a standard Borel space with a probability measure $\iota$. We follow [6] and say that an operator $S : L^2(X, \mu) \to L^2(Y, \iota)$ is a Markov operator if

- $S \geq 0$, i.e., $S(f) \geq 0$ whenever $f \geq 0$,
- $S(1_X) = 1_Y$,
- $S^*(1_Y) = 1_X$. 

The authors of [11, 3] prove the equivalence between two types of convergence: a sequence of graphons $W_n$ converges to $W$ in the cut-distance if and only if for every finite graph $F$ we have $t(F, W_n) \to t(F, W)$.
The class of Markov operators are infinite-dimensional analogues of doubly stochastic matrices.

Each partition \( C \) of \( V(G) \) corresponds uniquely to an equivalence relation on \( V(G) \) and vice-versa. It might be tempting to think that the easiest way how to translate the notion of equitable partition to a measurable setting is to work with so-called smooth equivalence relations on \( X \), i.e., equivalence relations induced by a measurable functions, together with the concept of measure disintegration (see [S Section 17]). Even though it can be verified that the notion that corresponds to equitable partitions can be stated in this language, it turns out that these point-concepts in this context are much more difficult to handle technically. It is much more convenient to work with the corresponding set-concept, i.e., conditional expectations.

Recall that our underlying space is a standard Borel space \((X,B)\) with a Borel probability measure \( \mu \). We say that \( C \) is a relatively complete sub-\( \sigma\)-algebra of \( B \) if it is a sub-\( \sigma\)-algebra and \( Z \in C \) whenever there is \( Z_0 \in C \) such that \( \mu(Z \Delta Z_0) = 0 \). Working with relatively complete sub-\( \sigma\)-algebras rather than sub-\( \sigma\)-algebras helps only to avoid minor technical difficulties. For example the set \( L^2(X,C,\mu) \) of \( C\)-measurable square-integrable functions is naturally a closed subspace of \( L^2(X,\mu) \).

We say that a closed linear subspace \( A \subseteq L^2(X,\mu) \) is \( W\)-invariant if \( T_W(A) \subseteq A \) and we say that relatively complete sub-\( \sigma\)-algebra \( C \) is \( W\)-invariant if \( L^2(X,C,\mu) \) is \( W\)-invariant. Our first observation (which we will prove in the next section) is that \( C \) is \( W\)-invariant if and only if \( T_W \) commutes with the conditional expectation \( \mathbb{E}(\cdot,C,X) \), i.e., \( T_W \circ \mathbb{E}(\cdot,C,X)(f) = \mathbb{E}(\cdot,C,X) \circ T_W(f) \) for every \( f \in L^2(X,\mu) \). In this case we define \( W_C : X \times X \to [0,1] \)

\[
W_C = \mathbb{E}(W(C \times C,X \times X)
\]

and call it a fraction of \( W \). We anticipate that the concept of \( W\)-invariant relatively complete sub-\( \sigma\)-algebras is the smallest relatively complete sub-\( \sigma\)-algebra that is generated by \( C \) and \( \mathbb{E}(\cdot,C,X) \).

The class of Markov operators are infinite-dimensional analogues of doubly stochastic matrices. In the first step define \( i \) be a finite graph and put \( \sigma \).

A suitable way to describe that even though the language is different, the only thing that we actually do is normalizing. Let \( G \), \( V \), \( \mathcal{G} \), \( \mathcal{F} \). Sections 3, 4. First, we restate the concept from the finite case in the language of distributions. Note it might be tempting to think that the easiest way how to translate the notion of equitable partition of a graph is the coarsest equitable partition of a graph is the minimal \( W\)-invariant relatively complete sub-\( \sigma\)-algebra. The analogue of the coarsest equitable partition of a graph is the minimal \( W\)-invariant relatively complete sub-\( \sigma\)-algebra. It is easy to see that the subalgebras that are \( G\)-invariant are exactly those subalgebras that are generated by equitable partitions. As in the finite case, a trivial example of \( W\)-invariant relatively complete sub-\( \sigma\)-algebra is \( B \). Also, if we denote by \( \{\emptyset,X\} \) the smallest relatively complete sub-\( \sigma\)-algebra that contains \( \{\emptyset,X\} \), it is easy to see that \( \{\emptyset,X\} \) is \( W\)-invariant if and only if \( d_{G,W}(x) \) is constant for \( \mu\)-almost every \( x \). In that case \( W_{i(\emptyset,X)} \) is the constant graphon with value \( \int_{X \times X} W(x,y) \, dm \times \mu \).

The analogue of the coarsest equitable partition of a graph is the minimal \( W\)-invariant relatively complete sub-\( \sigma\)-algebra. It is easy to see that such sub-\( \sigma\)-algebras always exists and we denote it as \( C(W) \).

A suitable way to describe \( C(W) \) is to put \( C^W_0 = \{\emptyset,X\} \) and then define inductively \( C^W_i \) to be the smallest relatively complete sub-\( \sigma\)-algebra such that

\[
T_W(L^2(X,C^W_i,\mu)) \subseteq L^2(X,C^W_{i+1},\mu).
\]

The relatively complete sub-\( \sigma\)-algebra that is generated by \( \bigcup_{n \in \mathbb{N}} C^W_i \) is then \( C(W) \).

Next, we describe informally the analogue of iterated degree sequence and we make it precise in Sections 3, 4. First, we restate the concept from the finite case in the language of distributions. Note that even though the language is different, the only thing that we actually do is normalizing. Let \( G \) be a finite graph and put \( n = |V(G)| \). We describe informally a map that encodes the iterated degree sequence. In the first step define \( i_1 = k_1 : V(G) \to [0,1] \) as

\[
v \mapsto \frac{|\text{deg}_G(v)|}{n}.
\]

Let \( A_1 \) be the range of \( k_1 \). Then we can assign to every \( v \in V(G) \) a distribution on \( A_1 \) as

\[
k_2(v)(\{r\}) = \frac{|\{w \in N(v) : i_1(w) = r\}|}{n}
\]

whenever \( r \in A_1 \). In other words, the weight of \( r \in A_1 \) from the point of view of \( v \in V(G) \) is the normalized number of neighbors of \( v \) that have normalized degree \( r \). We define \( i_2(v) = (k_1(v),k_2(v)) \) and let \( A_2 \) be the range of \( i_2 \). We continue iteratively to define \( k_3(v) \) to be the distribution from the point of view of \( v \) on \( A_2 \), and so on. After infinitely many iterations we get a map \( v \mapsto i_G(v) \) where \( i_G(v) \) is an infinite vector of distributions where the range of \( i_G \) is some set \( A_\infty \). The map \( i_G \) naturally encodes an equivalence relation \( E_G \) on \( V(G) \) defined as \( (v,w) \in E_G \) if and only if \( i_G(v) = i_G(w) \). One can verify that the map \( i_G \), the set \( A_\infty \) and the sizes of \( E_G\)-equivalence classes completely determines the iterated degree sequence \( D(G) \). See Figure 1 for an example.
Let $W$ be a graphon on $X$. We describe the first two steps of the construction. We remark that the precise definition is slightly different than the one sketched here (we start with $i_{W,0}$ and omit the auxiliary functions $k_i$). First, we define $i_{W,1} = k_{W,1} : X \to [0,1]$ as

$$x \mapsto \int_X W(x,y) \, d\mu(y).$$

For the second step, we fix $x \in X$ and consider the push-forward of $W(x,\cdot)$ via $i_1$, namely we define

$$k_{W,2}(x)(A) = \int_{i_1^{-1}(A)} W(x,y) \, d\mu(y)$$

whenever $A \subseteq [0,1]$ is a Borel set. This gives a Borel map

$$i_{W,2} : X \to M_{\leq 1}([0,1])$$

from $X$ to the space of all Borel measures of total mass at most 1 (see Appendix $\mathbb{E}$ for notation). Then we define $i_{W,2} = (k_{W,1}(x), k_{W,2}(x))$. Continuing inductively we produce, as in the finite case, a map $x \mapsto i_W(x)$, where $i_W(x)$ is an infinite vector of measures on some spaces. Important thing is that one can inductively define a compact metric space of all such infinite vectors, i.e., there is a compact metric space $M$ such that $i_W : X \to M$ for every graphon $W$. We call $M$ the space of iterated degree distributions. The information about the sizes of the equivalence classes $E_G$ described above are replaced by a distribution $\nu_W$ on $M$ that is a push-forward of $\mu$ via $i_W$. Finally, $\nu_W$ provides the analogue of iterated degree sequence.

As a next step we extract crucial properties of measures of the form $\nu_W$ in order to define abstractly a subspace of Borel probability measures on $M$. That space we call distributions on iterated degree distributions (DIDD). There are two key properties of DIDD. First, we describe a construction that produces from each DIDD $\nu$ an integral kernel $U[\nu]$ with the property that $W_C(W)$ and $U[\nu_W]$ are weakly isomorphic for every graphon $W$. The main observation here is that each vector $i_W(x)$ naturally encodes a measure on $M$ that is absolutely continuous with respect to $\nu_W$. Second, there is a family of real-valued continuous functions $T$ on $M$ such that for every $f \in T$ there is a finite tree $T$ such that

$$t(T,W) = \int f \, d\nu_W = t(T,U[\nu_W])$$

Figure 1. Iterated degree sequence vs. iterated degree distribution of depth 2 on a graph $G$ where $|V(G)| = 1200$. 

| $W(G)$ | $1200$ |
|--------|--------|
| $k_1(v)$ | $\frac{1}{50}$ |
| $k_2(v)$ | $\frac{1}{80}$ |

Degree 40

15 neighbors with degree 24

$E_G$-class

$\nu_W$
whenever \( W \) is a graphon. Moreover, \( \mathcal{T} \) is rich enough so that for every pair \( \nu_0 \neq \nu_1 \) of DIDD there is \( f \in \mathcal{T} \) such that
\[
\int f \, d\nu_0 \neq \int f \, d\nu_1.
\]

Finally, the reason why we work with standard Borel spaces is to be able to create quotient spaces. Namely, for every \( C \) that is a relatively complete sub-\( \sigma \)-algebra of \( \mathcal{B} \) there is a standard Borel space \((Y, \mathcal{D})\) with a probability measure \( \nu \) and a \( C \)-measurable map \( q : X \to Y \) such that \( q^{-1}(\mathcal{D}) \) generates \( C \) and \( \nu \) is the push-forward of \( \mu \) via \( q \). Since this space is unique up to measure isomorphism we denote it as \((X/C, C)\), the probability measure as \( \mu/C \) and the map as \( q_C \). We call it the quotient space of \( X \) by \( C \) and define the graphon \( W/C \) on \( X/C \) as
\[
(W/C)(q_C(x), q_C(y)) = W_C(x, y).
\]
It follows that \( W/C \) is well-defined and that \( W_C \) and \( W/C \) are weakly isomorphic. In fact, the graphon \( W/C(W) \) is the twin-free copy of \( W_{C(W)} \), see \([12, \text{Section 13.1.1}]\). Moreover, we show that \( W/C(W) \) is isomorphic to \( U[q_{W}] \), where \( C(W) \) is the minimal \( W \)-invariant relatively complete sub-\( \sigma \)-algebra of \( \mathcal{B} \).

Now we are ready to state our main result. The conditions are numbered in direct correspondence with their counterparts from Theorem 1.1.

**Theorem 1.2** (Characterizations of fractional isomorphism of graphons). Let \( W, U \in W_0 \). Then the following are equivalent:

1. \( t(T, U) = t(T, W) \) for every finite tree \( T \);
2. \( \nu_W = \nu_U \), where \( \nu_W \) and \( \nu_U \) are the DIDD that we assign to \( W \) and \( U \), respectively;
3. \( W/C(W) \) and \( U/C(U) \) are isomorphic,
4. there is a Markov operator \( S : L^2(X, \mu) \to L^2(X, \mu) \) such that \( T_{U} \circ S = S \circ T_{W} \),
5. there is a \( W \)-invariant relatively complete sub-\( \sigma \)-algebra \( \mathcal{C} \) and a \( U \)-invariant relatively complete sub-\( \sigma \)-algebra \( \mathcal{D} \) such that \( W_{\mathcal{C}} \) and \( U_{\mathcal{D}} \) are weakly isomorphic.

We say that \( U \) and \( W \) are fractionally isomorphic if one of the conditions above is satisfied.

The remaining of the paper is dedicated to make the definitions and statements precise and also to establish the basic tools for the proof of Theorem 1.2. The paper is organized as follows. In Section 2 we prove basic observations about relatively complete sub-\( \sigma \)-algebras, invariant subspaces and the minimal algebra \( C(W) \). In Section 3 we construct the space \( M \), define DIDD, and show the correspondence between integral kernels and DIDD. In Section 4 we prove the main technical result about the collection of tree functions \( \mathcal{T} \) defined on \( M \). Finally, in Section 5 we prove Theorem 1.2. In Appendices A, B, C, D and E we collect several well-known facts about standard Borel spaces, spaces of probability measures, and the connection between sub-\( \sigma \)-algebras, conditional expectations, and Markov operators that we need in our proof. Next, we finish the section with a few observations and open questions.

1.1. **Further remarks and problems.** A direct consequence of (1) in Theorem 1.2 is that the assignment
\[
\overline{W} \mapsto \overline{W}_{C(W)}
\]
is a well defined map from \( \overline{W}_0 \) to \( \overline{W}_0 \) with the property that if there are two sequences of graphons \( W_n \to_{\delta_0} W \) and \( U_n \to_{\delta_0} U \) such that \( W_n \) and \( U_n \) are fractionally isomorphic for every \( n < \omega \), then \( W \) and \( U \) are fractionally isomorphic. That is, the equivalence induced by fractional isomorphism is closed in the product space \( \overline{W}_0 \times \overline{W}_0 \) endowed with the product topology. In particular if \( U \) and \( W \) are weakly isomorphic, then they are fractionally isomorphic. We denote the range of the aforementioned map as \( \mathcal{F} \subseteq \overline{W}_0 \) and call elements of \( \mathcal{F} \) fraction-free graphons. It follows from (3) in Theorem 1.2 that the restriction of the equivalence induced by fractional isomorphism on \( \mathcal{F} \) is equal to the equivalence induced by weak isomorphism. Finally, it follows again from (1) in Theorem 1.2 combined with Theorem 1.1 and Proposition 1.1 that \( \overline{W} \mapsto \nu_W \) is a cut-distance continuous map and therefore the set of those DIDD that correspond to graphons is closed subset of \( M_{\leq 1}(\mathcal{M}) \).

**Question 1.3.** Is \( \overline{W} \mapsto \overline{W}_{C(W)} \) cut-distance continuous? Or equivalently, is \( \mathcal{F} \) closed in \( \overline{W}_0 \)?

**Question 1.4.** Let \( W, U \in W_0 \) be fractionally isomorphic. Is it possible to find sequences \( \{G_n\} \) and \( \{H_n\} \) of finite graphs such that \( G_n \) is fractionally isomorphic to \( H_n \) for each \( n < \omega \) and
\[
G_n \to_{\delta_0} W \quad \text{and} \quad H_n \to_{\delta_0} U?
\]
A positive answer to the second question together with the previous observation would provide a new characterization in Theorem 1.2.
2. Subalgebras

2.1. Kernels. Let \((X, \mathcal{B})\) be a standard Borel space and \(\mu\) be a probability measure on \(X\), i.e., \(\mu \in \mathcal{P}(X)\). In this paper an integral kernel on \(X\) is a \(\mathcal{B} \times \mathcal{B}\)-measurable map
\[
W : X \times X \to [0,1].
\]
The corresponding integral operator \(T_W : L^2(X, \mu) \to L^2(X, \mu)\) defined as
\[
T_W(f)(x) = \int_X W(x,y)f(y) \, d\mu
\]
is a well-defined Hilbert-Schmidt integral operator (see [15, Chapter 4, Exercise 15]). We consider integral kernels \(W\) and \(U\) on \(X\) to be the same if \(T_W = T_U\). This is equivalent with \(U(x,y) = W(x,y)\) for \(\mu\times\mu\)-almost every \((x,y) \in X \times X\) (Theorem A.3). In other words, \(U\) and \(W\) are the same as elements of \(L^\infty(X \times X, \mu \times \mu)\). We say that an integral kernel \(W\) is a graphon (on \(X\)) if \(W(x,y) = W(y,x)\) for \(\mu \times \mu\)-almost every \((x,y) \in X \times X\).

Claim 2.1. Let \(W\) be an integral kernel on \(X\). Then \(T_W\) is self-adjoint if and only if \(W\) is a graphon.

Proposition 2.2. Let \(W\) be a graphon, \(V \subseteq L^2(X,\mu)\) be a closed linear subspace and \(P_V\) be the corresponding orthogonal projection, i.e., \(P_V\) is self-adjoint, \(P_V \circ P_V = P_V\), and \(P_V(L^2(X,\mu)) = V\). Then the following are equivalent
\[
\begin{align*}
(1) & \quad T_W(V) \subseteq V, \\
(2) & \quad \text{there is an orthonormal basis of } V \text{ made of eigenvectors of } T_W, \\
(3) & \quad T_W \text{ commutes with the projection } P_V, \\
(4) & \quad T_W(V^\perp) \subseteq V^\perp, \text{ where } V^\perp = \{f \in L^2(X,\mu) : (\forall g \in V) \langle f, g \rangle = 0\}.
\end{align*}
\]

Proof. Suppose that (1) holds and let \(f \in V^\perp\). Then
\[
\langle T_W(f), g \rangle = \langle f, T_W(g) \rangle = 0
\]
whenever \(g \in V\). This shows that \(T_W(f) \in V^\perp\) and (4) follows. A symmetric argument shows the equivalence \((1) \Leftrightarrow (4)\).

Suppose that (1) and (4) holds. Fix \(f \in L^2(X,\mu)\) together with \(f_0 \in V\) and \(f_1 \in V^\perp\) such that \(f = f_0 + f_1\). Then we have
\[
T_W \circ P_V(f) = T_W(f_0) = P_V(T_W(f_0) + T_W(f_1)) = P_V \circ T_W(f)
\]
and \((3)\) follows.

The implication \((3) \Rightarrow (2)\) follows from the Spectral Theorem for Hilbert-Schmidt operators (see [15]). Finally, the implication \((2) \Rightarrow (1)\) is trivial. \(\Box\)

We say that a closed linear subspace \(V \subseteq L^2(X,\mu)\) is \(W\)-invariant if \(T_W(V) \subseteq V\). It follows that if \(W\) is a graphon, then this notion is equivalent to any of the conditions from Proposition 2.2.

2.2. Conditional Expectation and Invariant Subspaces. Let \((X, \mathcal{B})\) be a standard Borel space and \(\mu \in \mathcal{P}(X)\).

Definition 2.3 (Relative complete sub-\(\sigma\)-algebra). We say that \(\mathcal{C} \subseteq \mathcal{B}\) is a \(\mu\)-relatively complete sub-\(\sigma\)-algebra of \(\mathcal{B}\) if it is a sub-\(\sigma\)-algebra and \(Z \in \mathcal{C}\) whenever there is \(Z_0 \in \mathcal{C}\) such that \(\mu(Z \Delta Z_0) = 0\). We define \(\Theta_\mu\) as the set of all \(\mu\)-relatively complete sub-\(\sigma\)-algebras of \(\mathcal{B}\).

Since the measure \(\mu\) is fixed we will only write relatively complete sub-\(\sigma\)-algebra.

Claim 2.4. Let \(\Phi\) be a family of relatively complete sub-\(\sigma\)-algebra such that \(\emptyset \neq \Phi \subseteq \Theta_\mu\). Then
\[
\bigcap_{\mathcal{C} \in \Phi} \mathcal{C} \in \Theta_\mu.
\]

If \(\mathcal{C} \in \Theta_\mu\), then we define \(L^2(X,\mathcal{C},\mu)\) to be the collection of those functions in \(L^2(X,\mu)\) that are \(\mathcal{C}\)-measurable.

Claim 2.5. Let \(\mathcal{C} \in \Theta_\mu\). Then \(L^2(X,\mathcal{C},\mu)\) is a closed linear subspace and
\[
E(\cdot|\mathcal{C},X) : L^2(X,\mu) \to L^2(X,\mu)
\]
is an orthogonal projection onto \(L^2(X,\mathcal{C},\mu)\).

Proof. This follows directly from Theorem [C.1] parts (1), (2), and (3). \(\Box\)
Let $W$ be an integral kernel and $C \in \Theta_\mu$. In the introduction we assumed that $W$ is a graphon and $C$ is $W$-invariant and then we defined a fraction $W_C$ of $W$ as the conditional expectation with respect to the algebra $C \times C$. Here, we slightly abuse the notation and define $W_C$ as

$$W_C = E(W|B \times C, X \times X)$$

for every integral kernel $W$ and any $C \in \Theta_\mu$. It follows from Theorem C.1 (8) that $W_C$ is an integral kernel on $X$. We show in Claim 2.8 that for graphons and invariant algebras this definition and the definition of a fraction are the same.

**Claim 2.8.** Let $C \in \Theta_\mu$. Then

$$T_{W_C} = T_W \circ E(.|C, X).$$

In particular,

$$T_W \upharpoonright L^2(X, C, \mu) = T_{W_C} \upharpoonright L^2(X, C, \mu).$$

**Proof.** Let $f \in L^2(X, \mu)$ and set $F(x, y) = f(y)$. First we need the following technical observation.

**Subclaim 2.7.** Let $f \in L^2(X, \mu)$ and set $F(x, y) = f(y)$. Then

$$E(F|B \times C, X \times X)(x, \_ \_ | X) = E(f|C, X)$$

for $\mu$-almost every $x \in X$.

**Proof.** First suppose that $f \in L^\infty(X, \mu)$ and put $G(x, y) = E(f|C, X)(y)$. Let $A \in B$ and $B \in C$. Then we have

$$\int_{A \times B} E(F|B \times C, X \times X)(x, y) d\mu(x) d\mu(y) = \int_{A \times B} F(x, y) d\mu(x) d\mu(y) = \int_{A \times B} G(x, y) d\mu(x) d\mu(y)$$

by Theorem C.1 (9) and Fubini’s Theorem [2, Theorem 18.3]. Then we have $E(F|B \times C, X \times X) = G$ by Theorem A.3 and another use of Fubini’s Theorem [2, Theorem 18.3] finishes the proof. The general case follows from the fact that $L^\infty(X, \mu)$ is dense in $L^2(X, \mu)$. □

Subclaim 2.7 together with Theorem C.1 (6) imply that

$$T_{W_C}(f)(x) = \int_X E(W|B \times C, X \times X)(x, y) f(y) d\mu(y) = \int_X E(W|B \times C, X \times X)(x, y) F(x, y) d\mu(y) = \int_X W(x, y) E(F|B \times C, X \times X)(x, y) d\mu(y) = T_W(E(f|C, X))(x)$$

for $\mu$-almost every $x \in X$. The additional part of the claim is an easy consequence. □

We say that $C$ is $W$-invariant if $L^2(X, C, \mu)$ is $W$-invariant, i.e., if $T_W(L^2(X, C, \mu)) \subseteq L^2(X, C, \mu)$. Or equivalently by Claim 2.6

$$T_{W_C} \circ E(.|C, X) = T_W \circ E(.|C, X) = E(.|C, X) \circ T_W \circ E(.|C, X) = E(.|C, X) \circ T_{W_C},$$

i.e., $T_{W_C}$ commutes with $E(.|C, X)$.

**Claim 2.8.** Let $C \in \Theta_\mu$ be $W$-invariant. Then

$$W_C = E(W|B \times C, X \times X) = E(W|C \times C, X \times X).$$

Moreover, if $W$ is a graphon, then so is $W_C$.

**Proof.** Let $U = E(W|C \times C, X \times X)$. We show that $T_W = T_U$. First note that

$$T_{W_C} = T_W \circ E(.|C, X) = E(.|C, X) \circ T_W \circ E(.|C, X)$$

by the comment before this claim. Let $f, g \in L^2(X, \mu)$ and let $F(x, y) = f(y)$ and $G(x, y) = g(x)$. We have

$$\langle T_U(f), g \rangle = \int_{X \times X} G(x, y) E(W|C \times C, X \times X)(x, y) F(x, y) d\mu(x) d\mu(y) = \int_{X \times X} G(x, y) E(W|B \times C, X \times X)(x, y) d\mu(x) d\mu(y) = \int_{X \times X} E(g|C, X)(x) W(x, y) E(F|B \times C, X \times X)(x, y) d\mu(x) d\mu(y) = \langle T_W \circ E(.|C, X)(f), E(.|C, X)(g) \rangle$$

by Subclaim 2.7 together with Theorem C.1 (3), (4), (5), and (6). That implies $T_U = T_{W_C}$ by Proposition A.2.

For the moreover part it is enough to show that $T_{W_C}$ is self-adjoint by Claim 2.1. We have

$$\langle T_{W_C}(f), g \rangle = \langle E(.|C, X) \circ T_W \circ E(.|C, X)(f), g \rangle =$$
\[
\langle f, \mathbb{E}(\mathcal{I} \mathcal{C}, X) \circ T_W \circ \mathbb{E}(\mathcal{I} \mathcal{C}, X) (g) \rangle = \langle f, T_{W_C} (g) \rangle
\]
whenever \( f, g \in L^2(X, \mu) \) and that finishes the proof. \( \Box \)

Taking conditional expectation can be reformulated in the language of quotient spaces. First we recall Theorem E.1 and slightly abuse the notation. For every \( \mathcal{C} \in \Theta_\mu \), there is a standard Borel space \((X/\mathcal{C}, \mathcal{C})\), probability measure \( \mu/\mathcal{C} \in \mathcal{P}(X/\mathcal{C}) \) and a Borel map \( q_\mathcal{C} : X \to X/\mathcal{C} \) such that \( q_\mathcal{C}^* \mu = \mu/\mathcal{C} \). Moreover there is a unique linear isometry

\[ I_\mathcal{C} : L^2(X/\mathcal{C}, \mu/\mathcal{C}) \to L^2(X, \mu) \]

defined as

\[ I_\mathcal{C}(f)(x) = f(q_\mathcal{C}(x)) \]

that is a Markov operator onto \( L^2(X, \mathcal{C}, \mu) \). If we define \( S_\mathcal{C} = I_\mathcal{C}^* \), then \( S_\mathcal{C} \) is a Markov operator, \( S_\mathcal{C} \mid L^2(X, \mathcal{C}, \mu) \) is an isometrical isomorphism and \( S_\mathcal{C} = S_C \circ \mathbb{E}(\mathcal{I} \mathcal{C}, X) \). It follows that \( S_\mathcal{C} \circ I_\mathcal{C} \) is the identity on \( L^2(X, \mathcal{C}, \mu/\mathcal{C}) \) and \( I_\mathcal{C} \circ S_\mathcal{C} \) is equal to \( \mathbb{E}(\mathcal{I} \mathcal{C}, X) \).

**Definition 2.9.** Let \( \mathcal{C} \in \Theta_\mu \) be \( W \)-invariant. We define

\[ W/\mathcal{C} = S_{\mathcal{C} \times \mathcal{C}}(W_\mathcal{C}). \]

**Proposition 2.10.** Let \( W \) be an integral kernel and \( \mathcal{C} \in \Theta_\mu \) be \( W \)-invariant. Then

(i) if \( W \) is a graphon, then \( W/\mathcal{C} \) is a graphon. Furthermore, \( W_\mathcal{C} \) and \( W/\mathcal{C} \) are weakly isomorphic,

(ii) \( T_{W/\mathcal{C}} \circ S_\mathcal{C} = S_\mathcal{C} \circ T_{W_\mathcal{C}} \),

(iii) if \( W \) is a graphon, then we have \( T_{W/\mathcal{C}} \circ S_\mathcal{C} = S_\mathcal{C} \circ T_{W} \).

**Proof.** (i) It follows from Claim 2.8 that if \( W \) is a graphon, then \( W_\mathcal{C} \) is a graphon. By Theorem E.1 we have

\[ W_\mathcal{C}(x, y) = W/\mathcal{C}(q_\mathcal{C}(x), q_\mathcal{C}(y)) \]

for \( \mu \times \mu \)-almost every \( x, y \in X \) and therefore \( W/\mathcal{C}(r, s) = W/\mathcal{C}(s, r) \) for \( \mu/\mathcal{C} \times \mu/\mathcal{C} \)-almost every \( (r, s) \in X/\mathcal{C} \times X/\mathcal{C} \). This gives that \( W/\mathcal{C} \) is a graphon and since \( q_\mathcal{C}^* \mu = \mu/\mathcal{C} \) we get that \( W_\mathcal{C} \) and \( W/\mathcal{C} \) are weakly isomorphic because \( W_\mathcal{C} \) is a pull-back of \( W/\mathcal{C} \).

(ii) First we show that the equality holds for \( h \in L^2(X, \mathcal{C}, \mu) \). By Theorem E.1 there is \( f \in L^2(X/\mathcal{C}, \mu/\mathcal{C}) \) such that \( I_\mathcal{C}(f) = h \). Then we have

\[ \langle T_{W/\mathcal{C}} \circ S_\mathcal{C}(h), g \rangle = \langle T_{W/\mathcal{C}} \circ S_\mathcal{C}(I_\mathcal{C}(f)), g \rangle = \langle T_{W/\mathcal{C}}(f), g \rangle = \int_{X/\mathcal{C}} g(r) W/\mathcal{C}(r, s) f(s) \, d\mu/\mathcal{C} \times d\mu/\mathcal{C} = \int_{X/\mathcal{C} \times X/\mathcal{C}} g(q_\mathcal{C}(x)) W_\mathcal{C}(q_\mathcal{C}(x), q_\mathcal{C}(y)) f(q_\mathcal{C}(y)) \, d\mu \times d\mu = \langle T_{W_\mathcal{C}} \circ I_\mathcal{C}(f), I_\mathcal{C}(g) \rangle = \langle S_\mathcal{C} \circ T_{W_\mathcal{C}}(I_\mathcal{C}(f)), g \rangle = \langle S_\mathcal{C} \circ T_{W_\mathcal{C}}(h), g \rangle \]

for every \( g \in L^2(X/\mathcal{C}, \mu/\mathcal{C}) \). This shows that \( T_{W/\mathcal{C}} \circ S_\mathcal{C}(h) = S_\mathcal{C} \circ T_{W_\mathcal{C}}(h) \) for every \( h \in L^2(X, \mathcal{C}, \mu) \) by Proposition A.2. The general case in (ii) follows from \( S_\mathcal{C} = S_\mathcal{C} \circ \mathbb{E}(\mathcal{I} \mathcal{C}, X) \) together with \( T_{W_\mathcal{C}} = T_{W_\mathcal{C}} \circ \mathbb{E}(\mathcal{I} \mathcal{C}, X) \), where the former holds by Theorem E.1 and the latter by Claim 2.6.

(iii) Suppose that \( W \) is a graphon. Then by Proposition 2.2 we have

\[ \mathbb{E}(\mathcal{I} \mathcal{C}, X) \circ T_{W} = T_{W} \circ \mathbb{E}(\mathcal{I} \mathcal{C}, X). \]

This yields by (ii) and Claim 2.6

\[ T_{W/\mathcal{C}} \circ S_\mathcal{C} = S_\mathcal{C} \circ T_{W_\mathcal{C}} = S_\mathcal{C} \circ T_{W} \circ \mathbb{E}(\mathcal{I} \mathcal{C}, X) = S_\mathcal{C} \circ \mathbb{E}(\mathcal{I} \mathcal{C}, X) \circ T_{W} = S_\mathcal{C} \circ T_{W} \]

and the proof is finished. \( \Box \)

2.3. The Canonical Sequence \( \{C^W_n\}_{n \in \mathbb{N}} \). Let \( W \) be an integral kernel on \( X \).

**Definition 2.11.** We say that \( (\mathcal{D}, \mathcal{E}) \) is a \( W \)-invariant pair where \( \mathcal{D}, \mathcal{E} \in \Theta_\mu \) if

\[ T_W(L^2(X, \mathcal{D}, \mu)) \subseteq L^2(X, \mathcal{E}, \mu). \]

Note that \( \mathcal{C} \in \Theta_\mu \) is \( W \)-invariant if and only if \( (\mathcal{C}, \mathcal{C}) \) is a \( W \)-invariant pair. Suppose that \( \mathcal{C} \in \Theta_\mu \). Then \( (\mathcal{C}, \mathcal{B}) \) is a \( W \)-invariant pair. It follows that the collection \( \Phi(\mathcal{C}) \subseteq \Theta_\mu \) of all \( \mathcal{D} \in \Theta_\mu \) such that \( (\mathcal{C}, \mathcal{D}) \) is a \( W \)-invariant pair is non-empty. We define

\[ m(\mathcal{C}) = \bigcap \Phi(\mathcal{C}). \]

By Claim 2.4 we have \( m(\mathcal{C}) \in \Theta_\mu \).
Claim 2.12. Let $\mathcal{C} \in \Theta_\mu$. Then $(\mathcal{C}, m(\mathcal{C}))$ is a $W$-invariant pair.

Proof. Let $B = T_W(f)^{-1}(A)$ where $A \subseteq \mathcal{C}$ is a Borel set and $f \in L^2(X, \mathcal{C}, \mu)$. Then $B \in \mathcal{D}$ for every $\mathcal{D} \in \Phi(\mathcal{C})$ by the definition of $\Phi(\mathcal{C})$. Therefore $B \in m(\mathcal{C})$. This shows that $T_W(f)$ is $m(\mathcal{C})$-measurable whenever $f \in L^2(X, \mathcal{C}, \mu)$. \hfill $\Box$

Let $X \subseteq B$. Then by Claim 2.4 there is a minimal relatively complete sub-$\sigma$-algebra that contains $X$. We denote it as $\langle X \rangle$.

Definition 2.13 (Canonical sequence $\{C_n^W\}_{n \in \mathbb{N}}$). Define $C_0^W = \{\emptyset, X\}$ and inductively $C_{n+1}^W = m(C_n^W)$. Furthermore, we define

$$C(W) = \left\langle \bigcup_{n \in \mathbb{N}} C_n^W \right\rangle.$$ 

Proposition 2.14. Let $W$ be an integral kernel. Then $C(W)$ is the minimal $W$-invariant relatively complete sub-$\sigma$-algebra of $B$.

Proof. It follows by induction that every $C \in \Theta_\mu$ that is $W$-invariant must contain $C_n^W$ for every $n \in \mathbb{N}$ and therefore $C(W) \subseteq \mathcal{C}$.

Next we show that $C(W)$ is $W$-invariant. Let $X$ denote the closure of $\bigcup_{n \in \mathbb{N}} L^2(X, C_n^W, \mu)$ in $L^2(X, \mu)$. Then we have $X \subseteq L^2(X, C(W), \mu)$ because $L^2(X, C(W), \mu)$ is closed by Claim 2.5 and $\bigcup_{n \in \mathbb{N}} L^2(X, C_n^W, \mu) \subseteq L^2(X, C(W), \mu)$ by the construction of $C(W)$. The continuity of $T_W$ gives $T_W(X) \subseteq L^2(X, C(W), \mu)$ because $T_W(L^2(X, C_n^W, \mu)) \subseteq L^2(X, C(W), \mu)$ for every $n \in \mathbb{N}$ by the construction of $C(W)$.

It is a standard fact that finite linear combinations of functions of the form $1_A$, where $A \subseteq C(W)$, are dense in $L^2(X, C(W), \mu)$. Another standard fact is that $\bigcup_{n \in \mathbb{N}} C_n^W$ is dense in $C(W)$ in the sense that for every $\epsilon > 0$ and $A \subseteq C(W)$ there is $n \in \mathbb{N}$ and $B \in C_n^W$ such that $\mu(A \triangle B) < \epsilon$ (see [8, Exercise 17.43]). We have

$$\|1_A - 1_B\| = \mu(A \triangle B)$$

whenever $A, B \in B$. This implies $X = L^2(X, C(W), \mu)$ and that finishes the proof. \hfill $\Box$

3. Distributions on iterated degree distributions (DIDD)

In this section we assign to a given graphon $W$ a measure on a compact metric space $\mathbb{M}$ that encodes the canonical sequence $\{C_n^W\}_{n \in \mathbb{N}}$ and the graphon $W/C(W)$. Note that by Appendix E describing $C \in \Theta_\mu$ is basically the same as describing a surjective map $q_C : X \to X/C$. In the introduction we have motivated and explained the connection with the iterated degree sequence and here we will get into the technical details.

3.1. The Space $\mathbb{M}$. Let $K$ be a compact metric space. We denote by $M_{\leq 1}(K)$ the space of all Borel measures of total mass $\leq 1$ (see Appendix B for notation.) Let $P^0 = \{\ast\}$ be the one-point space. Define inductively

$$P^{n+1} = M_{\leq 1} \left( \prod_{i \leq n} P^i \right).$$

According to Appendix B $P^n$ is a compact metric space for every $n \in \mathbb{N}$. We put

$$\mathbb{M} = \prod_{n \in \mathbb{N}} P^n$$

and endow it with the product topology. Then by Tychonoff’s Theorem [15, Theorem A3], $\mathbb{M}$ is a compact metric space.

In order to make the presentation more clear we fix the notation as follows. We define $\mathbb{M}_n = \prod_{j \leq n} P^j$ and

$$p_{n,k} : \mathbb{M}_k \to \mathbb{M}_n$$

to be the canonical projections, where $n \leq k \leq \infty$, and we define $\mathbb{M} = \mathbb{M}_\infty$.

Notice that for every $\alpha \in \mathbb{M}$, we have $\alpha(n+1) \in M_{\leq 1}(\mathbb{M}_n)$. This allows to define

$$P = \{\alpha \in \mathbb{M} : (\forall n \in \mathbb{N}) \alpha(n+1) = p^\ast_{n,n+1} \alpha(n+2)\}.$$ 

It follows from Kolmogorov’s Existence Theorem [2, Theorem 36.1] that for every $\alpha \in P$ there is a unique $\mu_\alpha \in M_{\leq 1}(\mathbb{M})$ such that

$$p_{n,\infty}^\ast \mu_\alpha = \alpha(n+1).$$

In fact, we have the following uniform version.
Claim 3.1. The set $\mathcal{P}$ is closed in $\mathcal{M}$ and the map $\alpha \mapsto \mu_\alpha$ that satisfies
\[ p_{n,\infty}^\ast \mu_\alpha = \alpha (n + 1) \]
for every $n \in \mathbb{N}$ is a continuous map from $\mathcal{P}$ to $M_{\leq 1} (\mathcal{M})$.

Proof. Let $\mathbb{P}_n = \{ \alpha \in \mathbb{M} : \alpha (n + 1) = p_{n+1}^\ast \alpha (n + 2) \}$. Then it follows from Theorem [B.4] that $\mathbb{P}_n$ is closed in $\mathbb{M}$ and it is easy to see that $\mathbb{P} = \bigcap_{n \in \mathbb{N}} \mathbb{P}_n$.

It follows from Theorem [B.1] that
\[ A = \bigcup_{n \in \mathbb{N}} C (\mathbb{M}_n, \mathbb{R}) \circ p_{n,\infty} \]
is uniformly dense in $C (\mathbb{M}, \mathbb{R})$. Let $\alpha_k, \alpha \in \mathbb{P}$ for every $k \in \mathbb{N}$ such that $\alpha_k \to \alpha$ in $\mathbb{M}$ (or equivalently in $\mathbb{P}$). This means by definition that
\[ p_{n,\infty}^\ast \mu_{\alpha_k} = \alpha_k (n + 1) \to \alpha (n + 1) = p_{n,\infty}^\ast \mu_\alpha \]
for every $n \in \mathbb{N}$. Then we have
\[ \int_{\mathbb{M}} f \circ p_{n,\infty} (\beta) \, d\mu_{\alpha_k} (\beta) = \int_{\mathbb{M}_n} f (\kappa) \, d\mu_{n,\infty} (\mu_{\alpha_k} (\kappa)) = \int_{\mathbb{M}_n} f (\kappa) \, d\mu_{n,\infty}^\ast (\mu_\alpha (\kappa)) = \int_{\mathbb{M}} f \circ p_{n,\infty} (\beta) \, d\mu_\alpha (\beta) \]
for every $f \in C (\mathbb{M}_n, \mathbb{R})$. It follows from the the uniform density of $A$ that $\mu_{\alpha_k} \to \mu_\alpha$ in $M_{\leq 1} (\mathbb{M})$. \hfill \Box

Finally we are ready to state the main definition of this section.

Definition 3.2. We say that $\nu \in \mathcal{P} (\mathbb{M})$ is a distribution on iterated degree distributions (DIDD) if
\begin{enumerate}
\item $\nu (\mathbb{P}) = 1$,
\item $\mu_\alpha$ is absolutely continuous with respect to $\nu$ for $\nu$-almost every $\alpha \in \mathbb{M}$ that defines $\mu_\alpha$ as in equation (7),
\item the corresponding Radon–Nikodym derivative satisfies $0 \leq \frac{d\mu_\alpha}{d\nu} \leq 1$ for $\nu$-almost every $\alpha \in \mathbb{M}$.
\end{enumerate}

Note that (2) makes sense by (1).

3.2. From Kernels to DIDD. Let $W$ be an integral kernel. We define a measure $\nu_W \in \mathcal{P} (\mathbb{M})$ and show that $\nu_W$ is DIDD. It may be helpful to keep in mind that the measures $\nu_W$ are the analogue of iterated degree sequences of graphs.

Definition 3.3. (Measure $\nu_W$) Let $(X, \mathcal{B})$ be a standard Borel space and $W$ be an integral kernel on $X$. We define $i_{W,0} : X \to \mathbb{M}_0$ to be the constant map (recall that $\mathbb{M}_0 = \{ \ast \}$). Inductively, we define $i_{W,n+1} : X \to \mathbb{M}_{n+1}$ such that
\begin{enumerate}
\item $i_{W,n+1} (x) (j) = i_{W,n} (x) (j)$, for every $j \leq n$ and
\item $i_{W,n+1} (x) (n+1) (A) = \int_{i_{W,n} (A)} W(x, y) \, d\mu(y)$, whenever $A \subseteq \mathbb{M}_n$ is a Borel set.
\end{enumerate}

It follows trivially from the construction that there is a unique map
\[ i_W : X \to \mathcal{M} \]
defined as $i_W (x) (n) = i_{W,n} (x) (n)$. Finally, we define $\nu_W = i_W \mu$.

Claim 3.4. Let $n \in \mathbb{N}$ and suppose that $i_{W,n}$ is a Borel function. Then
\[ \int_{\mathcal{M}_n} g (\kappa) \, d (i_{W,n+1} (x) (n+1)) (\kappa) = \int_X W(x, y) (g \circ i_{W,n}) (y) \, d\mu(y) \]
for every bounded Borel function $g : \mathbb{M}_n \to \mathbb{R}$ and every $x \in X$.

Proof. This a straightforward consequence of (b) from the definition of $i_{W,n+1}$.

Claim 3.5. The maps $i_{W,n}$ and $i_W$ are Borel for every $n \in \mathbb{N}$.

Proof. First note that if $i_{W,n}$ are Borel for each $n \in \mathbb{N}$, then this implies that $i_W$ is Borel by (a). Also it is easy to see that $i_{W,0}$ is Borel. Next, we proceed inductively. Suppose that $i_{W,n}$ is Borel. Then by (a) it is enough to show that
\[ x \mapsto i_{W,n+1} (x) (n+1) \]
is a Borel assignment. To this end recall by [S] Theorem 17.24 that the Borel structure on $M_{\leq 1} (\mathbb{M}_n)$ is generated by the maps
\[ \nu \mapsto \int_X g(\kappa) \, d\nu \]
where \( g : \mathbb{M}_n \to \mathbb{R} \) is a bounded Borel function. Fix such a function \( g : \mathbb{M}_n \to \mathbb{R} \). Then we have by Claim 3.4 that
\[
x \mapsto \int_{\mathbb{M}_n} g(\kappa) d(\nu W_{n+1}(x)(n+1))(\kappa) = \int_X W(x,y)(g \circ i_{W_n})(y) \, d\mu(y)
\]
and it follows from the inductive assumption that this assignment is Borel. This finishes the proof. \( \square \)

Note that \( \nu W \in \mathcal{P}(\mathcal{M}) \) by Claim 3.5. Let \( \mathcal{B}(\mathbb{M}_n) \) denote the Borel \( \sigma \)-algebra of \( \mathbb{M}_n \) whenever \( n \leq \infty \).

**Proposition 3.6.** Let \( W \) be an integral kernel. Then \( \nu W \) is DIDD. Moreover \( i_W(x) \in \mathbb{P} \) for every \( x \in X \).

**Proof.** First we show that \( i_W(x) \in \mathbb{P} \) for every \( x \in X \). This immediately implies (1) in the definition of DIDD. To see this pick \( A \in \mathcal{B}(\mathbb{M}_n) \). We have
\[
i_W(x)(n+1)(A) = i_W(x)(n+1)(A) = \int_{i_W^{-1}(A)} W(x,y) \, d\mu(y) = \int_X W(x,y)1_{i_W^{-1}(A)}(y) \, d\mu(y) = \]
\[
= \int_X W(x,y)1_{i_W^{-1}(p_{n+1}^{-1}(A))}(y) \, d\mu(y) = \int_{i_W^{-1}(p_{n+1}^{-1}(A))} W(x,y) \, d\mu(y) = \]
\[
i_W(x+2)(n+2)(p_{n+2}^{-1}(A)) = i_W(x)(n+2)(p_{n+2}^{-1}(A)) = 0 \]
where the third equality follows from (a). This shows that \( i_W(x) \in \mathbb{P} \) for every \( x \in X \).

Let \( \nu_{n+1} = \mathbb{P}(\mathbb{M}_n) \) be the push-forward of \( \mu \) via \( i_{W_n, \mu} \), i.e., \( \nu_{n+1} = \mathbb{P}(\mathbb{M}_n) \). Then it is easy to see that \( \nu_{n+1} = \mathbb{P}(\mathbb{M}_n) \) and \( \nu_{n+1} = \mathbb{P}(\mathbb{M}_n) \) for every \( n \in \mathbb{N} \). Let \( D_n = \{ i_W^{-1}(A) : A \in \mathcal{B}(\mathbb{M}_n) \} \) and \( x \in X \). Then we have by the definition and Theorem C.1 that
\[
i_W(x)(n+1)(A) = \int_{i_W^{-1}(A)} W(x,y) \, d\mu(y) = \int_{i_W^{-1}(A)} \mathbb{E}(W(x,\cdot)|D_n, X)(y) \, d\mu(y)
\]
for every \( A \in \mathcal{B}(\mathbb{M}_n) \). By Corollary E.3 there is a Borel function \( g_{x,n} : \mathbb{M}_n \to \mathbb{R} \) such that
\[
\mathbb{E}(W(x,\cdot)|D_n, X)(y) = g_{x,n} \circ i_W(y)
\]
for \( \mu \)-almost every \( y \in X \). This gives
\[
\int_{i_W^{-1}(A)} \mathbb{E}(W(x,\cdot)|D_n, X)(y) \, d\mu(y) = \int_A g_{x,n}(\kappa) d\nu_{n+1}(\kappa)
\]
for every \( A \in \mathcal{B}(\mathbb{M}_n) \). This implies that \( i_W(x)(n+1) \) is absolutely continuous with respect to \( \nu_{n+1} \) and \( g_{x,n} \) is the corresponding derivative for every \( x \in X \) and every \( n \in \mathbb{N} \). Moreover by Theorem C.1 we have
\[
0 \leq \frac{dW(x)(n+1)}{d\mu_{n+1}} \leq 1.
\]
Since \( i_W(x) \in \mathbb{P} \), it follows that \( \{ g_{x,n} \circ p_{n+1} \}_{n \in \mathbb{N}} \) is a martingale and we have by the Doob’s Martingale Convergence Theorem [2, Theorem 35.5] and [2, Theorem 35.7] that \( \mu_{i_W(x)} \) is absolutely continuous with respect to \( \nu_{W} \) and
\[
0 \leq \frac{d\nu_{i_W(x)}}{d\mu_{n+1}} \leq 1
\]
for every \( x \in X \). This shows that \( \nu W \) satisfies (2) and (3) from the definition of DIDD and the proof is finished. \( \square \)

As promised, we show that it is possible to reconstruct \( \{ C_W^n \}_{n \in \mathbb{N}} \) (and therefore \( C(W) \)) from \( i_W \) and \( \mathcal{M} \).

**Proposition 3.7.** Let \( W \) be an integral kernel and \( n \in \mathbb{N} \). Then
\[
\left\{ i_W^{-1}(A) : A \in \mathcal{B}(\mathbb{M}_n) \right\} = C_W^n,
\]
i.e., the minimal relatively complete sub-\( \sigma \)-algebra of \( \mathcal{B} \) that makes the map \( i_W \) Borel is \( C_W^n \).

**Proof.** We show inductively that the claim holds for every \( n \in \mathbb{N} \). First let \( n = 0 \), then \( \mathcal{B}(\mathbb{M}_n) = \{ \emptyset, \{ \star \} \} \) and we have by the definition
\[
\left\{ i_W^{-1}(\emptyset), i_W^{-1}(\{ \star \}) \right\} = \{ \emptyset, X \} = C_W^0.
\]
Next suppose that the claim holds for \( n \in \mathbb{N} \) and write
\[
D_{n+1} = \left\{ i_W^{-1}(A) : A \in \mathcal{B}(\mathbb{M}_{n+1}) \right\}.
\]
Note that \( \mathcal{B}(\mathbb{M}_{n+1}) \) is the smallest \( \sigma \)-algebra that contains \( \{ p_{n+1}^{-1}(A) : A \in \mathcal{B}(\mathbb{M}_n) \} \) and makes the maps
\[
\kappa \mapsto \int_{\mathbb{M}_n} f \, d\kappa(n+1)
\]
Borel where \( f : \mathcal{M}_n \to \mathbb{R} \) is a bounded Borel function. This follows from the definition of the product \( \sigma \)-algebra and from \([8\), Theorem 17.24\].

First, by condition (a) of the definition of \( i_{W,n+1} \) and by the inductive hypothesis, we have
\[
i_{W,n+1}(p_{n,n+1}(A)) = i_{W,n}(A) \in C^W_n \subseteq C^W_{n+1},
\]
for every \( A \in \mathcal{B}(\mathcal{M}_n) \). Next, by Claim \[3.4\], the inductive hypothesis, and the definition of \( C^W_{n+1} \), we have that the map
\[
x \mapsto \int_{\mathcal{M}_n} f(k) d(i_{W,n+1}(x)(n+1)) (k) = \int_X W(x,y)(f \circ i_{W,n})(y) \, d\mu(y)
\]
is \( C^W_{n+1} \)-measurable for every bounded Borel function \( f : \mathcal{M}_n \to \mathbb{R} \). Therefore, we get \( \mathcal{D}_{n+1} \subseteq C^W_{n+1} \).

To show the opposite inclusion, suppose that \( f : X \to \mathbb{R} \) is \( C^W_n \)-measurable. Then by the inductive hypothesis and Corollary \[3.2\] there is a Borel function \( g : \mathcal{M}_n \to \mathbb{R} \) such that
\[
f(x) = g \circ i_{W,n}(x)
\]
for \( \mu \)-almost every \( x \in X \). Then, by Claim \[3.4\] the composition of \( i_{W,n+1} \) with
\[
\kappa \mapsto \int_{\mathcal{M}_n} g \, d\kappa(n+1)
\]
is equal to
\[
x \mapsto \int_X W(x,y)f(y) \, d\mu(y)
\]
for \( \mu \)-almost every \( x \in X \). This shows that \( C^W_{n+1} \subseteq \mathcal{D}_{n+1} \) and the proof is finished. \( \square \)

**Corollary 3.8.** Let \( W \) be an integral kernel. Then
\[
\langle \{i_{W}^{-1}(A) : A \in \mathcal{B}(\mathcal{M})\} \rangle = \mathcal{C}(W),
\]
i.e., the minimal relatively complete sub-\( \sigma \)-algebra of \( \mathcal{B} \) that makes the map \( i_W \) Borel is \( \mathcal{C}(W) \).

**Proof.** It is a standard fact that \( \mathcal{B}(\mathcal{M}) \) is generated by \( \bigcup_{n \in \mathbb{N}} \mathcal{B}(\mathcal{M}_n) \) as a \( \sigma \)-algebra (see \([8\), Section 10\]). The rest is an easy consequence of the definition of \( \mathcal{C}(W) \) together with Proposition \[3.7\]. \( \square \)

### 3.3. From DIIDD to Kernels

Let \( \nu \in \mathcal{P}(\mathcal{M}) \) be DIIDD. The properties of DIIDD implies that the assignment \( \alpha \mapsto \mu_\alpha \) is defined \( \nu \)-almost everywhere and that the assumptions of Theorem \[3.3\] are satisfied. That is, there is \( U[\nu] \in L^\infty(\mathbb{M} \times \mathbb{M}, \mu \times \mu) \) such that \( \|U[\nu]\|_\infty \leq 1 \) and
\[
\nu(x)(A) = \int_X U[\nu](x,y)1_A(y) \, d\mu
\]
for every \( A \in \mathcal{B}(\mathcal{M}) \).

**Theorem 3.9.** Let \( W \) be an integral kernel. Then
\[
W_{\mathcal{C}(W)}(x,y) = U[\nu_W](i_W(x),i_W(y))
\]
for \( \mu \times \mu \)-almost every \( (x,y) \in X \times X \).

**Proof.** Define an integral kernel \( U \) on \( X \) as
\[
U(x,y) = U[\nu_W](i_W(x),i_W(y)).
\]
Then, by Theorem \[A.3\] it is enough to show that \( T_{W_{\mathcal{C}(W)}} = T_U \). Moreover since both \( W_{\mathcal{C}(W)} \) and \( U \) are \( \mathcal{C}(W) \times \mathcal{C}(W) \)-measurable (the former by definition and the latter by Corollary \[3.10\]), it is enough to show that for every \( f \in L^2(X,\mathcal{C}(W),\mu) \) we have
\[
\int_X W_{\mathcal{C}(W)}(x,y)f(y) \, d\mu(y) = \int_X U(x,y)f(y) \, d\mu(y)
\]
for \( \mu \)-almost every \( x \in X \). This is equivalent with the fact that the same holds true for every \( f \in \bigcup_{n \in \mathbb{N}} L^2(X,\mathcal{C}^W_n,\mu) \) by the definition of \( \mathcal{C}(W) \).

In order to show this we first recall that for every \( n \in \mathbb{N} \) and for every \( f \in L^2(X,\mathcal{C}^W_n,\mu) \) there is \( g \in L^2(\mathbb{M},\mathcal{B}(\mathcal{M}),\nu_W) \) and a Borel map \( \tilde{g} : \mathcal{M}_n \to \mathbb{C} \) such that
\[
\begin{align*}
(i) & \quad f(y) = g \circ i_{W,n}(y) = \tilde{g} \circ p_{n,\infty} \circ i_W(y) = \tilde{g} \circ i_{W,n}(y) \text{ for } \mu \text{-almost every } y \in X, \\
(ii) & \quad g(\alpha) = \tilde{g} \circ p_{n,\infty}(\alpha) \text{ for every } \alpha \in \mathbb{M}.
\end{align*}
\]
This follows from Proposition 3.7 and Corollary 3.10 together with Theorem E.1.

Then we have

\[ \int_X W_C(W)(x,y) f(y) \, d\mu(y) = \int_X W(x,y)(\tilde{g} \circ i_{W,n})(y) \, d\mu(y) = \]

\[ = \int_{M_n} \tilde{g}(\kappa) \, d(i_{W,n+1}(x)(n+1))(\kappa) = \int_{M_n} \tilde{g}(\kappa) \, d\mu_{n,\infty}(x) = \]

\[ = \int_{M_n} \tilde{g} \circ p_{n, \infty}(\alpha) \, d\mu_{i_{W,n}}(\alpha) = \int_M g(\alpha) \, d\mu_{i_{W,n}}(\alpha) = \]

\[ \int_M U_\nu(W)(x,y) \, d\nu(\alpha) = \int_X U(x,y) f(y) \, d\mu(y) \]

for \( \mu \)-almost every \( x \in X \) where the first equality follows from (i), the second is Claim 3.4, the third and fourth is just the definition of \( \mu_{\alpha} \) combined with Proposition 3.6. The fifth is (ii), the sixth is by definition of \( U_\nu \) and the seventh follows from the definition of \( U \) together with the fact that \( i_{W,\mu} = \nu \). This finishes the proof.

Corollary 3.10. Let \( W \) be a graphon. Then \( W/C(W) \) is isomorphic to \( U_\nu(W) \). In particular \( U_\nu(W) \) is a graphon.

Proof. It follows from Theorems E.1, E.3 together with Corollaries E.2, 3.8 that there is a measurable almost bijection \( j_W : X/C(W) \to M \) such that \( i_W = j_W \circ q_W \) and \( j_W \mu/C(W) = \nu \). A combination of the first part of the proof of Proposition 2.10 (i) and Theorem 3.8 yields

\[ W/C(W)(x,y) = U_\nu(W)(j_W(x), j_W(y)) \]

for \( \mu/C(W) \times \mu/C(W) \)-almost every \( (x,y) \in X/C(W) \times X/C(W) \).

The additional part follows from Proposition 2.10 (i).

4. Tree Functions

This section is the most technical part of the paper. Our aim is to show two things. The first is that if \( W \) is a graphon on \( X \), then

\[ t(T, W) = t(T, W_C) \]

whenever \( C \in \Theta_\mu \) is \( W \)-invariant and \( T \) is a finite tree. The second is that there is a collection \( T \) of continuous functions on \( M \) that corresponds, in a precise sense described in Subsection 4.2 to tree densities of graphons of the form \( U_\nu \) and that separates DIDD.

Since we work with arbitrary integral kernels, not necessarily graphons, the collection \( T \), in fact, corresponds to rooted trees rather than trees.

4.1. Tree Functions and Invariant Subspaces. Let \( (T, v) \) be a rooted tree, i.e., a finite connected acyclic graph \( T = (V, E) \) with a distinguished vertex \( v \in T \). We denote as \( h(T, v) \) the height of \( T \) with respect to \( v \), i.e., number of edges of the longest path in \( T \) that starts in \( v \).

Let \( W \) be an integral kernel on \( X \). We assign inductively to each rooted tree \((T, v)\) a function \( f^W_{(T, v)} : X \to [0,1] \). We start with the unique rooted tree \((T, v)\) such that \( h(T, v) = 0 \) and put \( f^W_{(T, v)}(x) = 1 \) for every \( x \in X \). Let \((T, v)\) be such that \( h(T, v) = n + 1 > 0 \) and suppose that \( \text{deg}_T(v) = k \). We decompose \( T \) into subtrees rooted at neighbors of \( v \) as follows. The sequence of trees \( \{(T_i, v_i)\}_{i<k} \) is such that \( E(T) = \bigcup_{i<k} (v, v_i) \cup \bigcup_{i<k} E(T_i) \). We call this sequence the corresponding decomposition of \((T, v)\).

For any given set \( I \), we denote by \( X^I \) the space of all functions \( y : I \to X \). Suppose that the functions \( f^W_{(T_i, v_i)} \) are defined for every \( i < k \), then we define

\[ f^W_{(T, v)}(x) = \int_{y \in X^{(0,1, \ldots , k-1)}} \prod_{i<k} (f^W_{(T_i, v_i)}(y(i)) W(x, y(i))) \, d\mu_{\oplus k}(y). \]

This allows to define inductively \( f^W_{(T, v)} \) for every integral kernel \( W \) and every rooted tree \((T, v)\).

Proposition 4.1. Let \( W \) be an integral kernel on \( X \) and \((T, v)\) be a finite rooted tree with \( h(T, v) = n \). Then \( f^W_{(T, v)} \) is \( C^W_n \)-measurable and \( f^W_{(T, v)}(x) = f^W_{(T, v)}(x) \) for \( \mu \)-almost every \( x \in X \) whenever \( C \in \Theta_\mu \) is \( W \)-invariant.
Proof. We prove both claims inductively. It is trivial when \( n = 0 \). Suppose that \( h(T, v) = n + 1 > 0 \) and we have the corresponding decomposition \( \{(T_i, v_i)\}_{i < k} \) of \((T, v)\). Then, by Fubini’s Theorem \([2 \text{ Theorem 18.3}]\), we have

\[
\int_{(T,v)} f_W(x) = \int_{x \in X} \prod_{i < k} \left( \int_{y \in X} f_W(x, y) \, d\mu(y) \right) = \prod_{i < k} \left( \int_{y \in X} f_W(x, y) \, d\mu(y) \right).
\]

It follows from the inductive hypothesis and the definition of \( W_n \) that \( f_W(T, v) \) is \( W_n \)-measurable. Moreover, by the definition of \( W_n \), \( W_n \) is \( C \)-invariant, and the fact that \( W_n \subseteq C \), we have

\[
\prod_{i < k} \left( \int_{y \in X} f_W(x, y) \, d\mu(y) \right) = \prod_{i < k} \left( \int_{y \in X} W(x, y) \, d\mu(y) \right) = \prod_{i < k} \left( \int_{y \in X} W(x, y) \, d\mu(y) \right) = f_W(x)
\]

for \( \mu \)-almost every \( x \in X \). Note that the last equality follows from the fact that \( W(x, y) \subseteq C \), \( x \times X \) for \( \mu \)-almost every \( x \in X \). That finishes the proof.

Proposition 4.2. Let \( W \) be a graphon on \( X \) and \((T, v)\) be a rooted tree. Then

\[
t(T, W) = \int_{X} f_W(x) \, d\mu(x).
\]

Proof. We prove by induction on \( h(T, v) \) that

\[
f_W(T, v) = \int_{x \in X} W(x, y) \, d\mu(y)
\]

where \( y(v) \in X \), \((T, v)\) is a rooted tree and \( W \) is a graphon. Once we have this, then the claim immediately follows. Note that we use the assumption that \( W \) is a graphon, i.e., symmetric, implicitly in the equations.

If \( h(T, v) = 0 \), then the conclusion holds. Suppose that \( h(T, v) = n + 1 > 0 \) and take the corresponding decomposition \( \{(T_i, v_i)\}_{i < k} \) of \( T \). Let \( y(v) \in X \). Then we have

\[
f_W(T, v) = \prod_{i < k} \left( \int_{y(v_i) \in X} W(y(v), y(v_i)) f_W(T_i, v_i) \, d\mu \right) = \prod_{i < k} \left( \int_{y(v_i) \in X} W(y(v), y(v_i)) \prod_{r,s \in E(T_i)} W(y(r), y(s)) \, d\mu \right) = \prod_{i < k} \left( \int_{y \in X} W(y, y(v_i)) \prod_{r,s \in E(T_i)} W(y(r), y(s)) \, d\mu \right)
\]

where the first equality is the definition of \( f_W(T, v) \) together with the Fubini’s Theorem \([2 \text{ Theorem 18.3}]\), the second is the inductive hypothesis and the third is the Fubini’s Theorem \([2 \text{ Theorem 18.3}]\) once again.

Corollary 4.3. Let \( W \) be a graphon on \( X \) and \( T \) be a finite tree. Then

\[
t(T, W) = t(T, W_C) = t(T, U|_{V(W)})
\]

whenever \( C \subseteq \Theta_W \) is \( W \)-invariant.

Proof. This is a combination of Propositions \([1, 2, [2.10] \text{ i}]\), and Corollary \([3.10] \text{ i}\).  \( \square \)
4.2. Family $\mathcal{T}$. We define a collection $\mathcal{T} \subseteq C(\mathbb{M}, \mathbb{R})$ that is closed under multiplication and contains $1_{\mathbb{M}}$. Then we show that for every $f \in \mathcal{T}$ there is a rooted tree $(T, v)$ such that for every DIDD $\nu \in \mathcal{P}(\mathbb{M})$ we have

$$\int_{\mathbb{M}} f(\alpha) \, d\nu(\alpha) = \int_{\mathbb{M}} f_{(T, v)}^{(\nu)}(\alpha) \, d\nu(\alpha)$$

and that $\mathcal{T}$ separates points of $\mathbb{M}$.

The collection $\mathcal{T}$ is defined inductively and the construction resembles the inductive construction of the functions $f_{(T, \nu)}^W$. However, we must be more careful because, roughly speaking, $\mathbb{M} \setminus \mathcal{P} \neq \emptyset$. We notice that these problems disappear when computing the integral over DIDD.

During the inductive construction we assign to each function $f$, that we construct, a natural number $h(f) \in \mathbb{N}$ with the property that there is $f_{h(f), \infty} \in C(\mathbb{M}_{h(f)}, \mathbb{R})$ such that

$$f(\alpha) = f_{h(f), \infty} \circ p_{h(f)}(\alpha)$$

for every $\alpha \in \mathbb{M}$. Note that in that case there is for each $k \geq h(f)$ a function that we denote as $f_{k, \infty} \in \mathbb{M}_k$ such that

$$f(\alpha) = f_{k, \infty} \circ p_k(\alpha)$$

for every $\alpha \in \mathbb{M}$. It might happen that we construct $g = f$ such that $h(f) \neq h(g)$, however, as it will be clear later we are only interested in the fact that there is some $h(f) \in \mathbb{N}$.

**Operation I** Let $f \in C(\mathbb{M}, \mathbb{R})$ together with $h(f) \in \mathbb{N}$ with the properties above. Then for any $k \geq h(f)$ we define

$$F(f, k+1)(\alpha) = \int_{\mathbb{M}_k} f_{k, \infty} \, d\alpha(k+1)$$

and put $h(F(f, k+1)) = k + 1$. Note that if $\alpha \in \mathcal{P}$, then $F(f, k+1)(\alpha) = F(f, l+1)(\alpha)$ for every $k, l \geq h(f)$.

**Operation II** Let $\{f_i\}_{i < l} \subseteq C(\mathbb{M}, \mathbb{R})$ together with $\{h(f_i)\}_{i < l}$. Then we define

$$G(f_0, \ldots, f_{i-1})(\alpha) = \prod_{i < l} f_i(\alpha)$$

and put $h(G(f_0, \ldots, f_{i-1})) = \max\{h(f_i) : i \leq l\}$.

We put $\mathcal{T}_0 = \{1_{\mathbb{M}}\}$ and $h(1_{\mathbb{M}}) = 0$. Then define inductively $\mathcal{T}_{j+1}$ to be the union of $\mathcal{T}_j$ together with one iteration of Operation I and II applied on $\mathcal{T}_j$. Finally, we put $\mathcal{T} = \bigcup_{j \in \mathbb{N}} \mathcal{T}_j$. It is clear that $\mathcal{T}$ is closed under Operation I and II. Moreover, for each $f \in \mathcal{T}$ there is a minimal index $j \in \mathbb{N}$ such that $f \in \mathcal{T}_j$. Notice that this implies that the index of $F(f, k+1)$ (resp. $G(f_0, \ldots, f_{i-1})$) is strictly bigger then the index of $f$ (resp. of $f_i$) for every $i < l$.

Let us informally explain how Operations I and II are connected with rooted trees. This is made precise in Proposition 4.5. Let $(T, v)$ be a rooted tree. Then Operation I corresponds to a rooted tree $(T', v)$ that is defined as $V(T') = \{v\} \cup V(T')$ and $E(T') = \{v, w\} \cup E(T')$. Let $(T_i, w_i)_{i < k}$, then the Operation II corresponds to a rooted tree $(T, v)$ where we take a disjoint union of $T_i$ and then glue together all $w_i$ into a new root $v$.

**Proposition 4.4.** The collection $\mathcal{T}$ is closed under multiplication, contains $1_{\mathbb{M}}$ and separate points of $\mathbb{M}$.

**Proof.** It is a direct consequence of the definition of $\mathcal{T}$ that it is closed under multiplication and contains $1_{\mathbb{M}}$.

We define $\mathcal{G}_k = \{f \in \mathcal{T} : h(f) \leq k\}$ and $\mathcal{F}_k = \{f_{k, \infty} : f \in \mathcal{G}_k\}$. Then $\mathcal{F}_k \subseteq C(\mathbb{M}_k, \mathbb{R})$ is closed under multiplication and contains $1_{\mathbb{M}_k}$ for every $k \in \mathbb{N}$. We show inductively that $\mathcal{F}_k$ separate points of $\mathbb{M}_k$. This clearly holds for $k = 0$ because $\mathbb{M} = \{\ast\}$. Let $\kappa \neq \lambda \in \mathbb{M}_{k+1}$. We consider two cases. Either there is $j < k + 1$ such that $\kappa(j) \neq \lambda(j)$. Then by the inductive assumption there is $f \in \mathcal{G}_k \subseteq \mathcal{G}_{k+1}$ such that

$$f_{k+1, \infty}(\kappa) = f_{k, \infty}(p_{k, k+1}(\kappa)) \neq f_{k, \infty}(p_{k, k+1}(\lambda)) = f_{k+1, \infty}(\lambda).$$

Or $\kappa(j) = \lambda(j)$ for every $j < k + 1$ and $\kappa(k+1) \neq \lambda(k+1) \in M_{\leq 1}(\mathbb{M}_k)$. It follows from the inductive assumption that we may apply Corollary 3.1 with $M_k = K$ and $\mathcal{F}_k = \mathcal{F} = \mathcal{A}$ to get $f \in \mathcal{G}_k$ such that

$$\int_{\mathbb{M}_k} f_{k, \infty} \, d\kappa(k+1) \neq \int_{\mathbb{M}_k} f_{k, \infty} \, d\lambda(k+1).$$

Then the function $F(f, k+1) \in \mathcal{G}_{k+1}$ and we have

$$F(f, k+1)_{k+1, \infty}(\kappa) \neq F(f, k+1)_{k+1, \infty}(\lambda).$$

This shows that $\mathcal{F}_{k+1}$ separate points of $\mathbb{M}_{k+1}$.
From this fact we can see that \( \mathcal{T} \) separate points because if \( \nu_0 \neq \nu_1 \in \mathcal{M} \), then there is \( k \in \mathbb{N} \) such that \( \nu_0(k) \neq \nu_1(k) \). Therefore \( p_{k,\infty}(\nu_0) \neq p_{k,\infty}(\nu_1) \) and there is \( f \in \mathcal{G}_k \) such that 
\[
 f(\nu_0) = f_{k,\infty}(p_{k,\infty}(\nu_0)) \neq f_{k,\infty}(p_{k,\infty}(\nu_1)) = f(\nu_1).
\]

This finishes the proof. \( \square \)

**Proposition 4.5.** Let \( f \in \mathcal{T} \). Then there is a rooted tree \( (T, v) \) such that for every DIDD \( \nu \in \mathcal{P}(\mathcal{M}) \) we have 
\[
 f(\alpha) = f^{U[\nu]}_{(T,v)}(\alpha)
\]
for \( \nu \)-almost every \( \alpha \in \mathcal{M} \).

**Proof.** We show this by induction on \( j \in \mathbb{N} \) from the construction \( \mathcal{T} = \bigcup_{j \in \mathbb{N}} \mathcal{T}_j \). If \( j = 0 \) and \( \nu \in \mathcal{P}(\mathcal{M}) \) is DIDD, then we have
\[
 1_{\mathcal{M}}(\alpha) = 1 = f^{U[\nu]}_{(T,v)}(\alpha)
\]where \( (T, v) \) is the trivial rooted tree such that \( |V(T)| = 1 \).

Let \( f \in \mathcal{T}_{j+1} \setminus \mathcal{T}_j \). Suppose first that there is \( \{g_i\}_{i \leq l} \subseteq \mathcal{T}_j \) such that \( f = G(g_0, \ldots, g_{l-1}) \). Then by the inductive assumption there are rooted trees \( \{(T_i, v_i)\}_{i \leq l} \) such that for every DIDD \( \nu \in \mathcal{P}(\mathcal{M}) \) and every \( i < l \) we have
\[
g_i(\alpha) = f^{U[\nu]}_{(T_i,v_i)}(\alpha)
\]for \( \nu \)-almost every \( \alpha \in \mathcal{M} \). Define the rooted tree \( (T, v) \) in two steps, first take a disjoint union \( \bigcup_{i \leq l} T_i \), then glue together all \( \{v_i\}_{i \leq l} \) and call the new vertex \( v \). Suppose that \( deg_T(v) = r \) and write \( \{w_s\}_{s \leq r} \) for the neighbors of \( v \) in \( T \) and \( \{(S_s, v_s)\}_{s \leq r} \) for the corresponding decomposition of \( (T, v) \). Put \( A_i = \{s < r : \{w_s, v_i\} \in E(T_i)\} \). Then \( \{(S_s, v_s)_{s \in A_i}\}, \mathcal{T}_i, v_i \) is the corresponding decomposition of \( (T_i, v_i) \) and we have
\[
f^{U[\nu]}_{(T,v)}(\alpha) = \prod_{\beta \in \mathcal{M}^{A_i}} \left( U[\nu](\alpha, \beta(s)) f^{U[\nu]}_{(S_s,v_s)}(\beta(s)) \right) d\nu_{\mathcal{M}^{A_i}} = = \prod_{i < l} \left( f^{U[\nu]}_{(T_i,v_i)}(\alpha) \right) = \prod_{i < l} g_i(\alpha) = f(\alpha)
\]for \( \nu \)-almost every \( \alpha \in \mathcal{M} \) where the first equality is by the definition of \( f^{U[\nu]}_{(T,v)} \), the second is Fubini’s Theorem, the third is the definition of \( f^{U[\nu]}_{(T,v)} \), the fourth is the inductive assumption, and the fifth is the definition of \( f \).

Suppose that \( f = F(g, k + 1) \) for some \( g \in \mathcal{T}_j \) and \( k \geq h(g) \). Then it follows from the inductive assumption that there is a rooted tree \( (S, w) \) such that for every DIDD \( \nu \in \mathcal{P}(\mathcal{M}) \) we have
\[
g(\alpha) = f^{U[\nu]}_{(S,w)}(\alpha)
\]for \( \nu \)-almost every \( \alpha \in \mathcal{M} \). We define the rooted tree \( (T, v) \) as \( V(T) = V(S) \cup \{v\} \) and \( E(T) = E(S) \cup \{v, w\} \). Then we have \( deg_T(v) = 1 \) and the corresponding decomposition of \( (T, v) \) is \( \{(S, w)\} \).

We have
\[
f^{U[\nu]}_{(T,v)}(\alpha) = \int_M U[\nu](\alpha, \beta) f^{U[\nu]}_{(S,w)}(\beta) \ d\nu = \int_M U[\nu](\alpha, \beta) g(\beta) \ d\nu = = \int_M g(\beta) \ d\mu_\beta = \int_M g_{k,\infty}(p_{k,\infty}(\beta)) \ d\mu_\beta = \int_{M_k} g_{k,\infty}(\kappa) \ d\mu_{k+1,\infty} = \int_{M_k} g_{k,\infty}(\kappa) \ d\mu(k + 1) = F(g, k + 1)(\alpha)
\]for \( \nu \)-almost every \( \alpha \in \mathcal{M} \), where the first equality is by definition, the second is the inductive assumption, the third is by the definition of \( U[\nu] \), the fourth is by definition of \( h(g) \), the fifth is definition of the push-forward measure, the sixth because \( \nu(F) = 1 \), and the seventh is the definition of \( F(g, k + 1) \). \( \square \)

**Corollary 4.6.** Let \( \nu_0, \nu_1 \in \mathcal{P}(\mathcal{M}) \) be DIDD and suppose that \( \nu_0 \neq \nu_1 \). Then there is a rooted tree \( (T, v) \) such that
\[
 \int_M f^{U[\nu_0]}_{(T,v)}(\alpha) \ d\nu_0(\alpha) \neq \int_M f^{U[\nu_1]}_{(T,v)}(\alpha) \ d\nu_1(\alpha).
\]
Moreover if \( U[\nu_0], U[\nu_1] \) are graphons, then there is a finite tree \( T \) such that
\[
t(T, U[\nu_0]) \neq t(T, U[\nu_1]).
\]
Proof. The first assertion follows from a combination of Propositions 4.4 and Corollary B.2. The moreover part is then a consequence of Proposition 4.2. □

5. PROOF OF THEOREM 1.2

We recall the statement.

**Theorem.** Let $W, U \in \mathcal{W}_0$. Then the following are equivalent:

1. $t(T, U) = t(T, W)$ for every finite tree $T$,
2. $\nu_W = \nu_U$ where $\nu_W, \nu_U$ are the DIDD that we assign to $W, U$,
3. $W/C(W)$ and $U/C(U)$ are isomorphic,
4. there is a Markov operator $S : L^2(X, \mu) \to L^2(X, \mu)$ such that $T_W \circ S = S \circ T_U$,
5. there is a $W$-invariant relatively complete sub-$\sigma$-algebra $C$ and a $U$-invariant relatively complete sub-$\sigma$-algebra $D$ such that $W_C$ and $U_D$ are weakly isomorphic.

Proof of Theorem 1.2. To see that (1) implies (2), we apply Corollary 4.3 twice to both $U$ and $W$ to obtain $U[\nu_U] = t(T, U) = t(T, W) = t(T, U[\nu_W])$ for every finite tree $T$. By Corollary 4.6, we must have $\nu_U = \nu_W$.

That (2) implies (3), follows from Corollary 3.10 applied twice to both $U$ and $W$.

Next, we show that (3) implies (4). We let $Y = X/C(W)$, $Z = X/C(U)$, $\mu_Y = \mu/C(W)$, $\mu_Z = \mu/C(U)$, $W_Y = W/C(W)$ and $U_Z = U/C(U)$. By (3) there is a measurable bijection $r : Y \to Z$ such that $\alpha^r(\mu_Y) = \mu_Z$ and

$$W_Y(r, s) = U_Z(\alpha(r), \alpha(s))$$

for $\mu_Y \times \mu_Y$-almost every $r, s \in Y \times Y$. Note that the map

$$S_\alpha : L^2(Y, \mu_Y) \to L^2(Z, \mu_Z)$$

defined as $S_\alpha(f)(r) = f(\alpha^{-1}(r))$ is a Markov isomorphism by Theorem E.3. We have

$$S_\alpha \circ T_{W_Y}(f)(\alpha(r)) = S_{\alpha}(T_{W_Y}(f))(\alpha(r)) = T_{W_Y}(f)(\alpha^{-1}(\alpha(r))) = \int_Y W_Y(r, s)f(s) d\mu_Y = \int_Z U_Z(\alpha(r), \alpha(s)) S_\alpha(f)(\alpha(s)) \mu_Y = T_{U_Z}(S_\alpha(f)(\alpha(r))) = T_{U_Z} \circ S_\alpha(f)(\alpha(r))$$

for $\mu_Y$-almost every $r \in Y$. Therefore

$$S_\alpha \circ T_{W_Y} = T_{U_Z} \circ S_\alpha.$$

Let $f \in L^2(Z, \mu_Z)$ and $g \in L^2(X, \mu)$. Then we have

$$\langle I_{C(U)} \circ T_{U/Z}(f), g \rangle = \langle T_{U/Z}(f), S_C(U)(g) \rangle = \langle f, T_{U/Z} \circ S_C(U)(g) \rangle = \langle f, S_C(U) \circ T_{U/Z}(g) \rangle = \langle I_{C(U)}(f), T_{U/Z}(g) \rangle = \langle T_{U/Z} \circ I_{C(U)}(f), g \rangle$$

by Proposition 2.10 (iii) where $I_{C(U)} : L^2(Z, \mu_Z) \to L^2(X, \mu)$ is the Markov operator given by Theorem E.1. This shows that

$$I_{C(U)} \circ T_{U/C(U)} = I_{C(U)} \circ T_{U/Z} = T_{U/Z} \circ I_{C(U)}$$

by Proposition A.2. Together we get

$$I_{C(U)} \circ S_\alpha \circ S_{C(W)} \circ T_W = I_{C(U)} \circ S_\alpha \circ T_{W/C(W)} = S_{C(W)}$$

$$= I_{C(U)} \circ T_{U/C(U)} \circ S_\alpha \circ S_{C(W)} = T_U \circ I_{C(U)} \circ S_\alpha \circ S_{C(W)}$$

by another use of Proposition 2.10 (iii). Finally it follows from Proposition D.1 that $I_{C(U)} \circ S_\alpha \circ S_{C(W)}$ is a Markov operator and that gives (4).

We proceed to show that (4) $\Rightarrow$ (5).

**Claim 5.1.** There is $C \in \Theta_\mu$ (resp. $D \in \Phi_\mu$) that is $W$-invariant (resp. $U$-invariant) such that the restriction of $S$

$$S_0 : L^2(X, C, \mu) \to L^2(X, D, \mu)$$

is an isometrical isomorphism. Furthermore, $E(\lfloor D \rfloor, X) \circ S \circ E(\lfloor C \rfloor, X)$ is a Markov operator and

$$E(\lfloor D \rfloor, X) \circ S \circ E(\lfloor C \rfloor, X) \circ T_W = T_U \circ E(\lfloor D \rfloor, X) \circ S \circ E(\lfloor C \rfloor, X).$$
Proof. Let $T = S^* \circ S$ and $R = S \circ S^*$. Then by Proposition D.1, $T : L^2(X, \mu) \to L^2(X, \mu)$ and $R : L^2(X, \mu) \to L^2(X, \mu)$ are self-adjoint Markov operators. By Claim 2.1, $T_U$ and $T_W$ are self-adjoint, thus $T \circ T_W = T_W \circ T \quad \text{and} \quad R \circ T_U = T_U \circ R$.

By Theorem D.3, we have

$$\left\| \frac{1}{n} \sum_{k<n} T^k(f) - P(f) \right\|_2 \to 0 \quad \text{and} \quad \left\| \frac{1}{n} \sum_{k<n} R^k(f) - Q(f) \right\|_2 \to 0$$

for every $f \in L^2(X, \mu)$, where $P$ (resp. $Q$) are the projections onto the eigenspaces for eigenvalue 1 of $T$ (resp. $R$). We proceed to show that

$$S \circ P = Q \circ S.$$ To this end let $f \in L^2(X, \mu)$. Then we have $R^k(S(f)) = S(T^k(f))$ and therefore by continuity of $S$, we have

$$Q(S(f)) = \frac{1}{n} \sum_{k<n} R^k(S(f)) = S \left( \frac{1}{n} \sum_{k<n} T^k(f) \right) \to S(P(f)).$$

A similar argument shows that

$$S^* \circ Q = P \circ S^*.$$ It follows from Proposition D.1 that $P$ and $Q$ are Markov operators and

$$P \circ T_W = T_W \circ P \quad \text{and} \quad Q \circ T_U = T_U \circ Q$$

because they are limits in the strong operator topology of elements that have this property. This yields

$$Q \circ S \circ P = T_U \circ Q \circ S \circ P,$$

Let $f \in L^2(X, \mu)$ such that $P(f) = f$. Note that in that case $T(f) = S^* \circ S(f) = P(f) = f$. Then

$$\left\| S(f) \right\|_2^2 = \langle S(f), S(f) \rangle = \langle T(f), f \rangle = \langle f, f \rangle = \left\| f \right\|_2^2$$

and

$$Q \circ S(f) = S \circ P(f) = S(f).$$

Similarly for $g \in L^2(X, \mu)$ such that $Q(g) = g$ we get

$$\left\| S^*(g) \right\|_2 = \left\| g \right\|_2 \quad \text{and} \quad P \circ S^*(g) = S^*(g).$$

This implies that $S_0$, the restriction of $S$ to the range of $P$, is an isometrical isomorphism between the range of $P$ and the range of $Q$. By Theorem D.2, there are $C, D \in \Theta_\mu$ such that

$$P = E(\mathcal{L} C, X) \quad \text{and} \quad Q = E(\mathcal{L} D, X).$$

This finishes the proof. \(\square\)

We claim that

$$R = S_D \circ (E(\mathcal{L} D, X) \circ S_0 \circ E(\mathcal{L} C, X)) \circ I_C : L^2(X, \mu/C) \to L^2(X/D, \mu/D)$$

is a Markov isomorphism such that $R \circ T_{W/C} = T_{U/D} \circ R$. By Proposition D.1, we have that $R$ is a Markov operator. By the same argument as in (3) \(\Rightarrow\) (4) together with Proposition 2.10 (iii) and Claim 5.1 we get

$$R \circ T_{W/C} = S_D \circ E(\mathcal{L} D, X) \circ S_0 \circ E(\mathcal{L} C, X) \circ T_W \circ I_C =$$

$$= S_D \circ T_U \circ E(\mathcal{L} D, X) \circ S_0 \circ E(\mathcal{L} C, X) \circ I_C = T_{U/D} \circ R.$$

Finally note that by Theorem E.1, the map $I_C$ is an isometric isomorphism onto $L^2(X, \mu/C)$ and $S_D$ is isometrical isomorphism when restricted to $L^2(X, \mathcal{D}, \mu)$. Together with Claim 5.1, this shows that $R$ is Markov isomorphism.

Claim 5.2. Let $K$ be a graphon on $(Y, B_Y)$ and $L$ a graphon on $(Z, B_Z)$. Suppose that $R : L^2(Y, \mu_Y) \to L^2(Z, \mu_Z)$ is a Markov isomorphism such that $R \circ T_K = T_L \circ R$. Then there is a measurable bijection $\alpha : Y \to Z$ such that $\alpha^* \mu_Y = \mu_Z$ and

$$K(r, s) = L(\alpha(r), \alpha(s))$$

for $\mu_Y \times \mu_Y$-almost every $r, s \in Y$. 


Proof. It follows from Theorem 3.3 that there is a measurable bijection \( \alpha : Y \to Z \) such that \( \alpha^* \mu_Y = \mu_Z \) and for every \( f \in L^2(Y, \mu_Y) \) we have
\[
R(f)(\alpha(r)) = f(r)
\]
for \( \mu_Y \)-almost every \( r \in Y \). Define an integral kernel \( M : Z \times Z \to [0, 1] \) as
\[
M(\alpha(r), \alpha(s)) = K(r, s).
\]
We proceed to show that
\[
(T_M(1_\alpha(A)), 1_\alpha(B)) = (T_L(1_\alpha(A)), 1_\alpha(B))
\]
for every \( A, B \in \mathcal{B}_Y \). By the fact that \( \alpha \) is a bijection and by Theorem A.3 this is enough to conclude that \( L = M \). Note that we have \( R(1_A) = 1_\alpha(A) \) for every \( A \in \mathcal{B}_Y \). This gives
\[
\langle T_L(1_\alpha(A)), 1_\alpha(B) \rangle = \langle T_K(R(1_A)), 1_\alpha(B) \rangle = \langle (T_L(1_\alpha(A)), 1_\alpha(B)) =
\]
\[
\int_Z R \left( \int_Y K(r, s) 1_A(s) \, d\mu_Y(s) \right) 1_B(t) \, d\mu_Z(t) = \int_Z M(t, u) 1_A(u) \, d\mu_Z(u) 1_B(t) \, d\mu_Z(t) = (T_M(1_\alpha(A)), 1_\alpha(B))
\]
and that finishes the proof. \( \square \)

It follows from Proposition 2.10 (i) that \( W_C \) and \( W/\mathcal{C} \) (resp. \( W_D \) and \( W/\mathcal{D} \)) are weakly isomorphic. Therefore by the Claim 7.2 we have that \( W_C \) and \( U_D \) are weakly isomorphic and this finishes the proof of (4) \( \Rightarrow \) (5).

Finally, we show that (5) \( \Rightarrow \) (1). It follows from Corollary 4.3 that \( t(T, W) = t(T, W_C) \) whenever \( T \) is a tree and \( \mathcal{C} \) is \( W \)-invariant, and similarly \( t(T, U) = t(T, W_D) \). Since \( W_C \) and \( U_D \) are weakly isomorphic, we have \( t(T, W_C) = t(T, W_D) \) and that finishes the proof. \( \square \)

APPENDIX A. STANDARD BOREL SPACES

Let \( X \) be a set and \( \mathcal{B} \) a \( \sigma \)-algebra of subsets of \( X \). We say that \( (X, \mathcal{B}) \) is a standard Borel space if there is a separable completely metrizable topology \( \tau \) on \( X \) such that \( \mathcal{B} \) is equal to the \( \sigma \)-algebra of Borel subsets generated by \( \tau \) (see [3, Section 12]). We denote the space of all Borel measures on \( X \) as \( \mathcal{P}(X) \). Note that the set \( \mathcal{P}(X) \) endowed with the \( \sigma \)-algebra generated by the maps
\[
A \mapsto \mu(A),
\]
where \( A \in \mathcal{B} \), is a standard Borel space (see [3, Section 17]).

Let \( \mu \in \mathcal{P}(X) \). As usual, we consider two \( \mathcal{B} \)-measurable functions \( f, g : X \to \mathbb{C} \) to be the same if
\[
\mu(\{x \in X : f(x) \neq g(x)\}) = 0.
\]
In that case we abuse the notation and write \( f = g \). We write \( L^2(X, \mu) \) (resp. \( L^\infty(X, \mu) \)) for the space of square integrable (resp. bounded) functions on \( X \). We denote the correspondent norm as \( \|f\|_2 \) (resp. \( \|f\|_\infty \)) and we write \( \langle f, g \rangle \) for the scalar product in \( L^2(X, \mu) \). Note that \( L^\infty(X, \mu) \subseteq L^2(X, \mu) \). First we recall several basic facts.

**Proposition A.1.** Let \( (X, \mathcal{B}) \) be a standard Borel space, \( \mu \in \mathcal{P}(X) \) and \( f, g \in L^\infty(X, \mu) \). Then \( f = g \) if and only if
\[
\int_X f(x) 1_A(x) \, d\mu = \int_X g(x) 1_A(x) \, d\mu
\]
for every \( A \in \mathcal{B} \).

**Proposition A.2.** Let \( (X, \mathcal{B}) \) be a standard Borel space, \( \mu \in \mathcal{P}(X) \) and \( S, T : L^2(X, \mu) \to L^2(X, \mu) \) be a bounded linear operators. Then \( S = T \) if and only if
\[
\langle T(f), g \rangle = \langle S(f), g \rangle
\]
for every \( f, g \in L^2(X, \mu) \).

Let \( W \in L^\infty(X \times X, \mu \times \mu) \). We define the corresponding operator \( T_W : L^2(X, \mu) \to L^2(X, \mu) \) as
\[
T_W(f)(x) = \int_X W(x, y)f(y) \, d\mu.
\]
It follows from [15] that \( T_W \) is well-defined bounded Hilbert-Schmidt operator.

**Theorem A.3.** Let \( W, U \in L^\infty(X \times X, \mu \times \mu) \). Then \( T_W = T_U \) if and only if
\[
\mu \times \mu(\{(x, y) \in X \times X : W(x, y) \neq U(x, y)\}) = 0.
\]

**Proof.** Combine Propositions A.1, A.2 \( \square \)
Let \((X, \mathcal{B})\) and \((Y, \mathcal{C})\) be a standard Borel spaces. Suppose that \(\mu \in \mathcal{P}(X)\) and \(f : X \to Y\) is a Borel map. Then we define the push-forward of \(\mu\) via \(f\), in symbols \(f^*\mu\), as
\[f^*\mu(A) = \mu(f^{-1}(A))\]
for every \(A \in \mathcal{C}\). It is a standard fact that \(f^*\mu \in \mathcal{P}(Y)\), see [8 Exercise 17.28].

Appendix B. Compact Spaces

Let \(K\) be a compact metric space. Write \(C(K, \mathbb{R})\) for the vector space of all continuous functions from \(K\) to \(\mathbb{R}\). Then \(C(K, \mathbb{R})\) with the sup-norm and pointwise multiplication is a real Banach algebra. We say that \(A \subseteq C(K, \mathbb{R})\) separates points if for every \(k \neq l \in K\) there is \(f \in A\) such that \(f(k) \neq f(l)\).

Theorem B.1 (Real Stone-Weierstrass). [17] Let \(K\) be a compact metric space and \(A \subseteq C(K, \mathbb{R})\) a subalgebra that contains \(1_K\) and separates points. Then \(A\) is uniformly dense in \(C(K, \mathbb{R})\).

We denote the \(\sigma\)-algebra of Borel sets of \(K\) as \(\mathcal{B}(K)\). Then \((K, \mathcal{B}(K))\) is a standard Borel space and the space of all Radon measures \(M(K)\) corresponds by Riesz representation Theorem [14 Theorem 6.19] to the space of all positive bounded functionals on \(C(K, \mathbb{R})\). We are only interested in those \(\mu \in M(K)\) that are positive real valued and \(\mu(K) \leq 1\), we denote this subspace by \(M_{\leq 1}(K)\). The restriction of the weak* topology on \(M_{\leq 1}(K)\) is compact metrizable [8 Theorem 17.22]. Recall that \(\mu_n \to \mu\) in the weak* topology if
\[
\int_K f \, d\mu_n \to \int_K f \, d\mu.
\]
for every \(f \in C(K, \mathbb{R})\). The \(\sigma\)-algebra of Borel sets generated by the weak* topology on \(M_{\leq 1}(K)\) coincide with the standard Borel structure on \(M_{\leq 1}(K)\) generated by the maps
\[
A \mapsto \mu(A)
\]
where \(A \in \mathcal{B}(K)\) (see [8 Section 17]).

Corollary B.2 (Separating Measures). Let \(K\) be a compact metrizable space and \(A \subseteq C(K, \mathbb{R})\) such that it is closed under multiplication, contains \(1_K\), and separate points. Then for every \(\mu \neq \nu \in M_{\leq 1}(K)\) there is \(f \in A\) such that
\[
\int_K f \, d\mu \neq \int_K f \, d\nu,
\]
i.e., the linear functionals that correspond to \(A\) separate points in \(M_{\leq 1}(K)\).

Proof. It is easy to see that the vector space \(\tilde{A}\) that is generated by \(A\) is actually a subalgebra that satisfies the assumptions of Theorem B.1 and is therefore uniformly dense in \(C(K, \mathbb{R})\). Let \(\mu \neq \nu \in M_{\leq 1}(K)\). There is \(f \in C(K, \mathbb{R})\) such that
\[
\int_K f \, d\mu \neq \int_K f \, d\nu
\]
because functionals that correspond to \(C(K, \mathbb{R})\) separate points of \(M(K)\). Since \(\tilde{A}\) is uniformly dense, we find \(\{g_n\}_{n \in \mathbb{N}}\) such that \(\|g_n - f\|_{sup} \to 0\). This implies that
\[
\int_K g_n \, d\mu \to \int_K f \, d\mu \quad \text{and} \quad \int_K g_n \, d\nu \to \int_K f \, d\nu.
\]
Therefore there is \(n, l \in \mathbb{N}\) and \(\{f_i\}_{i \leq l} \subseteq A\) such that \(g_n = \sum_{i \leq l} f_i\) and
\[
\int_K g_n \, d\mu \neq \int_K g_n \, d\nu.
\]
This implies that there is \(i \leq l\) such that
\[
\int_K f_i \, d\mu \neq \int_K f_i \, d\nu
\]
and we are done. \(\square\)

Let \(K\) be a compact metric space and \(\mu, \nu \in M_{\leq 1}(K)\). Then we say that \(\nu\) is absolutely continuous with respect to \(\mu\), in symbols \(\nu << \mu\), if \(\nu(A) = 0\) whenever \(\mu(A) = 0\). The classical Radon–Nikodym Theorem [14 Theorem 6.10] states that \(\nu << \mu\) if and only if there is a unique \(f \in L^1(K, \mu)\) such that
\[
\nu(A) = \int_A f \, d\mu
\]
for every \(A \in \mathcal{B}(K)\). We call \(f\) the Radon–Nikodym derivative of \(\nu\) with respect to \(\mu\) and denote it as \(d\nu/d\mu\). The following might be considered as the uniform version of Radon–Nikodym Theorem.
Theorem B.3. Let $K$ be a compact metric space, $\mu \in \mathcal{P}(K)$ and $\iota : K \to \mathcal{P}_{\leq 1}(K)$ be a measurable map such that

1. $\iota(x) \ll \mu$,
2. $0 \leq \frac{d(\iota(x))}{d\mu} \leq 1$

for $\mu$-almost every $x \in K$. Then there is $W \in L^\infty(K \times K, \mu \times \mu)$ such that

$$||W||_{\infty} \leq 1$$

and for $\mu$-almost every $x \in K$ we have

$$\iota(x)(A) = \int_X W(x, y)1_A(y) \, d\mu$$

for every $A \in \mathcal{B}(K)$.

Proof. Define a probability measure $\Phi$ on $K \times K$ as the unique extension of

$$\Phi(A \times B) = \int_A \iota(x)(B) \, d\mu.$$  

It is easy to verify that $\Phi \ll \mu \times \mu$ and $W = \frac{d\Phi}{d(\mu \times \mu)}$ works as required. \hfill $\Box$

Theorem B.4. [8] Exercise 17.28 Let $K$ and $L$ be compact metric spaces. Suppose that $f : K \to L$ is a continuous map. Then the map $f^* : \mathcal{M}_{\leq 1}(K) \to \mathcal{M}_{\leq 1}(L)$ defined as $\mu \mapsto f^*\mu$ is continuous.

Proof. Suppose that $\mu_n \to \mu$ in $\mathcal{M}_{\leq 1}(K)$. Then we have

$$\int_L g \, df^*\mu_n = \int_K g \circ f \, d\mu_n \to \int_K g \circ f \, d\mu = \int_L g \, df^*\mu$$

for every $g \in C(L, \mathbb{R})$, where we used the definition of $f^*$ and the fact that $g \circ f \in C(K, \mathbb{R})$. \hfill $\Box$

Appendix C. Conditional Expectation

Let $(X, \mathcal{B})$ be a standard Borel space and $\mu \in \mathcal{P}(X)$. We say that $\mathcal{C} \subseteq \mathcal{B}$ is a relatively complete sub-$\sigma$-algebra if it is a sub-$\sigma$-algebra and $Z \in \mathcal{C}$ whenever there is $Z_0 \in \mathcal{C}$ such that $\mu(Z \triangle Z_0) = 0$. We denote the collection of all relatively complete sub-$\sigma$-algebras as $\Theta_\mu$.

If $\mathcal{C} \in \Theta_\mu$ and $(Y, \mathcal{D})$ is a standard Borel space, then we say that a map $f : X \to Y$ is $\mathcal{C}$-measurable if $f^{-1}(A) \in \mathcal{C}$ for every $A \in \mathcal{D}$.

Theorem C.1. [2] Section 34 Let $(X, \mathcal{B}, \mu)$ be a probability measure space and $\mathcal{C} \in \Theta_\mu$. Then there is a bounded self-adjoint linear operator

$$\mathbb{E}(\cdot | \mathcal{C}, \mu) : L^2(X, \mathcal{B}, \mu) \to L^2(X, \mathcal{B}, \mu)$$

that enjoys the following properties:

1. $\mathbb{E}(\cdot | \mathcal{C}, X)$ is an orthogonal projection,
2. $\mathbb{E}(f | \mathcal{C}, X)$ is $\mathcal{C}$-measurable for every $f \in L^2(X, \mu)$,
3. if $f \in L^2(X, \mu)$ is $\mathcal{C}$-measurable, then $\mathbb{E}(f | \mathcal{C}, X) = f$,
4. $\mathbb{E}(fg | \mathcal{C}, X) = g\mathbb{E}(f | \mathcal{C}, X)$ whenever $f, g \in L^2(X, \mu)$ and $g$ is $\mathcal{C}$-measurable,
5. if $\mathcal{C} \subseteq \mathcal{D}$, then $\mathbb{E}(\mathbb{E}(f | \mathcal{D}, X) | \mathcal{C}, X) = \mathbb{E}(f | \mathcal{C}, X)$ for every $f \in L^2(X, \mu)$,
6. $\int_X f(x)\mathbb{E}(g | \mathcal{C}, X)(x) \, d\mu = \int_X \mathbb{E}(f | \mathcal{C}, X)(x)g(x) \, d\mu$ for every $f, g \in L^2(X, \mu)$,
7. if $f \geq 0$, then $\mathbb{E}(f | \mathcal{C}, X) \geq 0$ for every $f \in L^2(X, \mu)$,
8. $||f||_{\infty} \geq ||\mathbb{E}(f | \mathcal{C}, X)||_{\infty}$ whenever $f \in L^\infty(X, \mu) \subseteq L^2(X, \mu)$,
9. for every $A \in \mathcal{C}$ and $f \in L^2(X, \mu)$ we have

$$\int_A f \, d\mu = \int_A \mathbb{E}(f | \mathcal{C}, X) \, d\mu.$$  

Appendix D. Markov Operators

Our main reference here is [6]. We remark that in [6] all the results are stated for arbitrary measure spaces and that Markov operators are defined on $L^1$ spaces rather than on $L^2$. However, it follows from [6] Chapter 13 that every Markov operator on $L^2$ has a unique extension to a Markov operator on $L^1$ and that the restriction of a Markov operator on $L^1$ to $L^2$ is a Markov operator.

Let $(X, \mathcal{B})$ (resp. $(Y, \mathcal{D})$) be a standard Borel space and $\mu \in \mathcal{P}(X)$ (resp. $\nu \in \mathcal{P}(Y)$). We say that a bounded linear operator $S : L^2(X, \mu) \to L^2(Y, \nu)$ is a Markov operator if

- $S(f) \geq 0$ whenever $f \geq 0$,
\( S(1_X) = 1_Y, \)

\( S^2(1_Y) = 1_X. \)

**Proposition D.1.** [6] Theorem 13.2 and 13.8 The class of Markov operators is closed under adjoints, composition and pointwise limits, in the sense that if \( S_n : L^2(X, \mu) \to L^2(Y, \nu) \) are Markov operators for every \( n \in \mathbb{N} \) and there is \( S : L^2(X, \mu) \to L^2(Y, \nu) \) such that

\[ ||S_n(f) - S(f)||_2 \to 0 \]

for every \( f \in L^2(X, \mu) \), then \( S \) is a Markov operator. Moreover every Markov operator is a contraction, i.e., its norm is bounded by 1.

We say that \( P : L^2(X, \mu) \to L^2(X, \mu) \) is a Markov projection if it is an orthogonal projection and a Markov operator (see [6] Section 13.3). Note that by Theorems C.1, C.7, C.9 we have that \( E(\| C, X) \) is a Markov projection for every \( C \in \Theta_\mu \) (see [6] Remark 13.17).

**Theorem D.2** (Structure of Markov projections). [6] Theorem 13.16 Let \((X, \mathcal{B})\) be a standard Borel space and \( \mu \in \mathcal{P}(X) \). There is a one-to-one correspondence between

1. Markov projections,
2. \( \Theta_\mu \), the relatively complete sub-\( \sigma \)-algebras of \( \mathcal{B} \).

The correspondence is given as

\[ P \mapsto \{ A \in \mathcal{B} : P(1_A) = 1_A \} \quad \text{and} \quad C \mapsto E(\| C, X). \]

**Theorem D.3** (Mean Ergodic Theorem). [6] Theorem 8.6, Example 13.20 Let \((X, \mathcal{B})\) be a standard Borel space, \( \mu \in \mathcal{P}(X) \) and \( S : L^2(X, \mu) \to L^2(X, \mu) \) a Markov operator. Then

\[ \left\| \frac{1}{n} \sum_{k=0}^{n-1} S^k(f) - P(f) \right\|_2 \to 0 \]

for every \( f \in L^2(X, \mu) \), where \( P \) is the orthogonal projection onto the closed subspace \( \{ g \in L^2(X, \mu) : S(g) = g \} \).

Note that it follows from Proposition D.1 that \( P \) is a Markov projection and therefore by Theorem D.2 it is of the form \( E(\| C, X) \) for some \( C \in \Theta_\mu \).

**Appendix E. Quotient Spaces**

Now we use the fact that for standard Borel spaces with probability measure the set maps correspond to point maps.

**Theorem E.1.** Let \((X, \mathcal{B})\) be a standard Borel space, \( \mu \in \mathcal{P}(X) \) and \( C \in \Theta_\mu \). There is a standard Borel space \((Y, \mathcal{D})\) and \( \nu \in \mathcal{P}(Y) \), measurable surjection \( q_C : X \to Y \), and Markov operators

\[ S_C : L^2(X, \mu) \to L^2(Y, \nu) \quad \text{and} \quad I_C : L^2(Y, \nu) \to L^2(X, \mu) \]

such that

1. \( q_C^* \mu = \nu, \)
2. \( S_C^2 = I_C, \)
3. \( S_C \circ E(\| C, X) = S_C, \)
4. \( \| I_C(f) \|_2 = \| f \|_2 \) for every \( f \in L^2(Y, \mathcal{D}), \)
5. \( I_C \circ S_C = E(\| C, X), \)
6. \( S_C \circ I_C \) is the identity on \( L^2(Y, \mathcal{D}), \)
7. \( I_C(f)(x) = f(q_C(x)) \) for every \( f \in L^2(Y, \mathcal{D}). \)

**Proof.** The existence of \((Y, \mathcal{D}), \nu \) and \( q_C \) follows from [8] Exercise 17.43 ii]. Define \( I_C \) by the condition (7). Then it is easy to see that \( I_C \) is a Markov embedding by [6] Section 12.2, Theorem 13.7 and all the other properties follow from [6] Section 13.2, 13.3. \( \square \)

**Corollary E.2.** Let \((X, \mathcal{B})\) and \((Y, \mathcal{D})\) be standard Borel spaces. Suppose that \( \mu \in \mathcal{P}(X) \) and \( f : X \to Y \) is a Borel function. Write \( C \in \Theta_\mu \) for the minimal relatively complete sub-\( \sigma \)-algebra that makes \( f \) Borel. Then for every \( g_0 \in L^2(X, C, \mu) \) there is a Borel map \( g_1 : Y \to C \) such that \( g_0(x) = g_1 \circ f(x) \) for \( \mu \)-almost every \( x \in X \).

**Proof.** Put \( \nu = f^* \mu \in \mathcal{P}(Y) \) and note that by [8] Theorem 21.10 there is a \( Y_0 \in \mathcal{D} \) such that \( Y_0 \subseteq f(X) \) and \( \nu(Y_0) = 1 \). Then use Theorem E.1 \( \square \)
We say that a map $S: L^2(X, \mu) \to L^2(Y, \nu)$ is a Markov isomorphism if it is a Markov operator that is a bijection and $||S(f)||_2 = ||f||_2$ whenever $f \in L^2(X, \mu)$ (see [6, Section 12.2, Theorem 13.7]).

**Theorem E.3.** Let $(X, \mathcal{B})$ (resp. $(Y, \mathcal{D})$) be a standard Borel space and $\mu \in \mathcal{P}(X)$ (resp. $\nu \in \mathcal{P}(Y)$). Then there is a correspondence between

1. Markov isomorphisms $S: L^2(X, \mu) \to L^2(Y, \nu)$,
2. measurable measure preserving bijections $\alpha: X \to Y$.

The correspondence from (2) to (1) is given as $\alpha \mapsto S_\alpha(f)(x) = f(\alpha^{-1}(x))$.

**Proof.** It follows from [6, Theorem 12.11] that there is a correspondence between Markov isomorphisms and measure algebra isomorphisms. It is a standard fact (see [9, Theorem 1.9]) that every measure algebra isomorphism is induced by a measurable measure preserving bijection under the assumption that the spaces are standard Borel. \qed

**References**

[1] L. Babai, P. Erdős, and S. Selkow. Random graph isomorphism. *SIAM Journal on Computing.*, (9):628–635, 1980.
[2] Billingsley, P. Probability and Measure. Third edition. Wiley, 1995.
[3] C. Borgs, J. T. Chayes, L. Lovász, V. T. Sós, and K. Vesztergombi. Convergent sequences of dense graphs. I. Subgraph frequencies, metric properties and testing. *Adv. Math.*, 219(6):1801–1851, 2008.
[4] C. Borgs, J. T. Chayes, L. Lovász, V. T. Sós, and K. Vesztergombi. Convergent sequences of dense graphs. II. Multiway cuts and statistical physics. *Ann. Math.*, 176(1):151–219, 2012.
[5] Holger Dell, Martin Grohe, and Gaurav Rattan. Lovász meets weisfeiler and leman. In *ICALP*, 2018.
[6] T. Eisner, B. Farkas, M. Haase, and R. Nagel. *Operator Theoretic Aspects of Ergodic Theory*. Graduate Texts in Mathematics. Springer International Publishing, 2015.
[7] M. Grohe, K. Kersting, M. Mladenov, and P. Schweitzer. *Color refinement and its applications. In An Introduction to Lifted Probabilistic Inference*. Cambridge University Press, To appear.
[8] Kechris, A.S. Classical Descriptive Set Theory. Graduate texts in mathematics, Springer-Verlag, 1995.
[9] Kerr, D. and Li, H. Independence and Dichotomies. Springer Monographs in Mathematics, Springer International Publishing, 2016.
[10] L. Lovász. Operations with structures. *Acta Mathematica Hungarica*, 18:321–328, 1967.
[11] L. Lovász and B. Szegedy. Limits of dense graph sequences. *J. Combin. Theory Ser. B.*, 96(6):933–957, 2006.
[12] László Lovász. *Large Networks and Graph Limits.*, volume 60 of Colloquium Publications. American Mathematical Society, 2012.
[13] M. Ramana, E. Scheinerman, D. Ullman. Fractional isomorphism of graphs. *Discrete Mathematics.*, 132(1–3):247–265, 1994.
[14] W. Rudin, Real and Complex Analysis, McGraw-Hill, New York 1987.
[15] W. Rudin, Functional Analysis, McGraw-Hill, New York 1973.
[16] E. Scheinerman, D. Ullman. *Fractional Graph Theory, A Rational Approach to the Theory of Graphs.*
[17] M. Stone. The Generalized Weierstrass Approximation Theorem. *Mathematics Magazine.*, 21(4):167–184, 1948.
[18] G. Tinhofer. Graph isomorphism and theorems of birkhoff type. *Computing.*, 36(4):285–300, 1986.
[19] G. Tinhofer. A note on compact graphs. *Discrete Applied Mathematics.*, 30(2–3):253–264, 1991.

Grebík, Rocha: Institute of Computer Science, Czech Academy of Sciences. Pod Vodárenskou věží 2, 182 07, Prague, Czechia. With institutional support RVO:67865807.

E-mail address: greboshrabos@seznam.cz, israelrocha@gmail.com