arXiv:1812.06282v1 [math.PR] 15 Dec 2018

A Generalization of Hierarchical Exchangeability on Trees to Directed Acyclic Graphs

Paul Jung[1], Jiho Lee[1], Sam Staton[2], and Hongseok Yang[3]

[1]Department of Mathematical Sciences, KAIST
[2]Department of Computer Science, University of Oxford
[3]School of Computing, KAIST

December 18, 2018

Abstract

Motivated by problems in Bayesian nonparametrics and probabilistic programming discussed in Staton et al. (2018), we present a new kind of partial exchangeability for random arrays which we call DAG-exchangeability. In our setting, a given random array is indexed by certain subgraphs of a directed acyclic graph (DAG) of finite depth, where each nonterminal vertex has infinitely many outgoing edges. We prove a representation theorem for such arrays which generalizes the Aldous-Hoover representation theorem.

In the case that the DAGs are finite collections of certain rooted trees, our arrays are hierarchically exchangeable in the sense of Austin and Panchenko (2014), and we recover the representation theorem proved by them. Additionally, our representation is fine-grained in the sense that representations at lower levels of the hierarchy are also available. This latter feature is important in applications to probabilistic programming, thus offering an improvement over the Austin-Panchenko representation even for hierarchical exchangeability.

Key words: Bayesian nonparametrics, Exchangeability, Hierarchical exchangeability, Aldous-Hoover representation

1 Introduction

Motivated by issues arising in the study of spin glasses, in [AP14], Austin and Panchenko consider a random array indexed by the paths of a collection of infinitary rooted trees of finite depth, where each path starts from the root of a tree and ends at a leaf of the tree. A random array is *hierarchically exchangeable*, defined in [AP14], if its joint distribution remains invariant under rearrangements that preserve the structure of each tree in the collection of trees underlying the index set (see Example 3.3(c) below for a more precise description). In their work, Austin and Panchenko prove that such tree-indexed arrays have a representation in the spirit of the celebrated Aldous-Hoover representation for exchangeable arrays of random variables. In the special case where all trees in the collection have

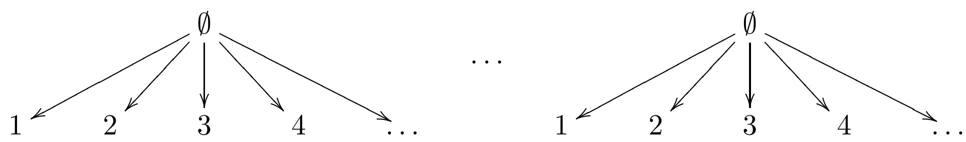

Figure 1: A collection of rooted trees

a depth of one, i.e., are copies of $\mathbb{N}$ rooted at $\emptyset$ (see Figure 1), then hierarchical exchangeability reduces to separate exchangeability also known as row-column exchangeability. The number of trees in the collection corresponds to the dimension of the random array. We refer to [Kal05, Ch. 7] (see also [Ald85, Aus12]) for the definition of separate exchangeability and for additional background on what are now classic results in the theory of exchangeable random arrays.

In this work, motivated by exchangeable random processes used in Bayesian nonparametrics for modeling data and by problems related to probabilistic programming, we consider a certain type of partial exchangeability on random arrays indexed by certain subgraphs of a directed acyclic graph (DAG) of finite depth, where each nonterminal vertex has infinitely many outgoing edges. In analogy to [AP14] we say that a random array, indexed by such subgraphs of a DAG, is DAG-exchangeable if its joint distribution remains invariant under rearrangements that preserve the structure of the DAG; a precise description is given later.

Our main result proves a representation theorem for such arrays which generalizes the Aldous-Hoover representation theorem. In the case that the DAG is a collection of certain rooted trees, our arrays are hierarchically exchangeable, and we recover the representation theorem proved in [AP14].

There are many reasons for considering such generalized random arrays as we do here. On a purely mathematical level, the study of various forms of partial exchangeability has been vibrant since the 1980's, and two somewhat recent surveys of the early results in this field are given in [Aus08] and [Ald09]. Indeed, in the foundational work [Hoo79], the question of when representations arise for partially exchangeable random arrays is posed in Section 7. On the level of applications, we provide a summary of our motivations with respect to Bayesian nonparametric models and probabilistic programming in the next section; a forthcoming companion paper explains these applications in much more detail.

Following the applications presented in the next section, the rest of the paper is organized as follows. In Section 3, we introduce the precise setup by which we work, present some examples from the theory which motivate our main result, and finally define our notion of partial exchangeability. The main result is presented in Section 4, while its proof comprises Section **??**. In the appendix we indicate an alternate route of proving our representation, without the fine-graining discussed above. This alternate method is model-theoretic and is based on the work of [CT17].

2 Applications to Probabilistic Programming

In terms of applications, our motivation comes from studying generative models of array-like structures through probabilistic programming languages[1]. These are high-level languages for statistical modeling that come equipped with a separate Bayesian inference engine (such as MCMC simulation).

[1] At this stage, Church [GMR$^+$08] and Anglican [WvdMM14, TvdMYW16] have some support for advanced Bayesian nonparametric models, through the XRP feature and the `produce`/`absorb` constructs for random processes, but we are really thinking of a next generation of probabilistic programming languages, e.g. [SYA$^+$17], with proper module and library functionality.

In that context, the first attempt at modeling using infinite exchangeable arrays would be simply to use the Aldous-Hoover representation. For a 2-dimensional array, this kind of programming involves the following basic operations:

- find a fresh array (formally, draw a sample from a uniform distribution);
- pick a fresh row in a given array (formally, draw a sample from a uniform distribution);
- pick a fresh column in a given array (formally, draw a sample from a uniform distribution);
- enquire as to the contents of an array in a given row and column (use the representing function together with another uniform sample).

However, an Aldous-Hoover representation might not be efficient or even computable [AAF$^+$, FR12]. In many cases it is preferable to use an implementation that appears to have the same interface but which, internally, builds up the array lazily (on-the-fly). This is analogous to using a Polỳa urn to simulate the Beta-Bernoulli relationship without ever directly sampling from a Beta distribution. Sampling from a distribution satisfies a property known in computer science as 'dataflow': operations can be freely reordered as long as the flow of data is respected. Any computer implementation, however lazy, and whether based on urns, stick-breaking or so forth, will still satisfy the dataflow property. For random arrays, the dataflow property coincides with the statistical property of exchangeability [SYA$^+$17, SSY$^+$18].

From this programming perspective it is natural and easy to design more elaborate generative models. For a simple extension, consider an array where each cell in the array itself contains an infinite array. This would amount to the following additional functionality:

- pick a fresh subrow in a given cell (a given row and column) in a given array;
- pick a fresh subcolumn in a given cell in a given array;
- enquire as to the contents of an array in a given subrow and subcolumn of a given cell in a given array.

If we also retain the idea of each cell containing a value, as well as containing another array, then this functionality is what we call 'fine-grained': although we can enquire based on all the indices (the array, row, column, subrow, and subcolumn) we may also enquire as to the value based on a subset of the indices (the array, row and column) alone.

If such a complex generative model still has the dataflow property, i.e. is suitably exchangeable, then our theorem says that, apart from the computability issues, it could just have well been implemented by sampling from uniform distributions. Thus dataflow in programming and exchangeability in statistics are intimately connected more generally.

3 Setup

In this section, we describe a formal setting for specifying random arrays indexed by certain subgraphs of an acyclic directed graph (DAG), and also for expressing their probabilistic symmetries that arise from the structure of the DAG.

3.1 Overview of Infinitary DAGs and Their Finite Presentations

To begin, we recall the index sets of four well-known exchangeable random arrays, and explain how their elements correspond to subgraphs of infinitary DAGs of finite depth; these subgraphs will serve as the index sets of the random arrays we will consider. *Infinitary* here means each vertex other than terminal ones have infinitely many outgoing edges. For (a) an exchangeable sequence of de Finetti type, (b) a row-column exchangeable array and (c) a general ℓ-dimensional exchangeable array of Aldous-Hoover type, the index sets are respectively given by

$$\text{(a) } \mathbb{N} \qquad \text{(b) } \mathbb{N}\times\mathbb{N} \qquad \text{(c) } \underbrace{\mathbb{N}\times\cdots\times\mathbb{N}}_{\ell \text{ times}}.$$

In the hierarchical-exchangeability setting of [AP14], case (c) is generalized so that each separately exchangeable "dimension" of the array is indexed, not by numbers in $\mathbb{N}$, but rather by sequences of numbers in $\mathbb{N}^r$. The index set in this case is

$$\text{(c') } \mathbb{N}^{r_1}\times\cdots\times\mathbb{N}^{r_\ell}.$$

All four index sets are associated with infinitary DAGs of finite depth. The indices in each set then correspond to certain subgraphs of the associated DAG. For instance, the set $\mathbb{N}$ in (a) is associated with a tree of depth 1 whose root has infinitely many children, i.e. a single tree in Figure 1, and each index $\alpha \in \mathbb{N}$ corresponds to a path from the root to the α-th leaf. We remind the reader that the paths are subgraphs of a tree. The index set in (c') is associated with a DAG consisting of ℓ infinitary trees of depths $r_1, \ldots, r_\ell$, respectively. The element

$$\alpha = \left(\left(v_1^{(1)}, \ldots, v_{r_1}^{(1)}\right), \ldots, \left(v_1^{(\ell)}, \ldots, v_{r_\ell}^{(\ell)}\right) \right) \in \mathbb{N}^{r_1}\times\cdots\times\mathbb{N}^{r_\ell}$$

corresponds to a subgraph made out of ℓ paths, where the i-th path starts from the root of the i-th tree and repeatedly moves toward the leaves by taking the $v_j^{(i)}$-th child at step j until the path hits a leaf.

The presence of such an associated DAG and the correspondence between indices and subgraphs is not accidental. There is a general three-step recipe for constructing an index set by first building an infinitary DAG of finite depth and selecting certain subgraphs of the DAG as indices. The four index sets that we discussed can be constructed by this recipe.

The first step of the recipe is to choose a *finite* DAG G, which depicts the skeleton of an *infinitary* DAG to be built later. In case (a), the finite DAG G is just a single vertex v and no edges; in case (b), G is the DAG with two vertices, r (row) and c (column), and no edges; in case (c'), it is the DAG consisting of ℓ disjoint paths of length $r_1, \ldots, r_\ell$ (see the figure in Example 3.3 below).

The second step is to generate an infinitary DAG G' from G by making infinitely many copies of vertices of G and connecting these copies by edges in an appropriate manner which we now describe. For every vertex v of G, define the *downset* of v to be

$$D_v \stackrel{\text{def}}{=} \{w \ : \ \text{there is a path from } w \text{ to } v \text{ in } G\}.$$

The vertex set of G' is $\bigcup_v \mathbb{N}^{D_v}$. That is, each vertex of G' is a function $\alpha : D_v \to \mathbb{N}$, for some v. The right way to understand such a vertex α is as one of the infinitely many copies of v of G that is assigned an identifier α. This identifier is used to connect the various copied vertices with directed edges. In other words, there is a directed edge from $\alpha \in \mathbb{N}^{D_v}$ to $\beta \in \mathbb{N}^{D_w}$ in G' if (v, w) is a directed edge of G and $\beta|_{D_v} = \alpha$. For each of (a), (b), (c) and (c') from above, the associated infinitary DAGs are constructed this way.

The third and final step of our recipe is to pick subgraphs of G'. Which specific subgraphs to choose for indices depends on the kind of random array one would like to use. But the recipe demands a certain closure property: if a subgraph (V'', E'') of G' is chosen, it is generated from a subset of vertices W of G (which is the whole vertex set in all basic examples) and a function $\alpha : G \to \mathbb{N}$ as follows:

$$V'' = \Big\{\alpha|_C \ : \ C \subseteq \bigcup_{v\in W} D_v \ \text{ such that } \ D_w \subseteq C \text{ for every } w \in C\Big\},$$

$$E'' = \Big\{(\alpha, \beta) \ : \ (\alpha, \beta) \ \text{ is an edge in } G' \ \text{ and } \ \alpha, \beta \in V''\Big\}.$$

The first equality says that V'' is obtained by first taking a vertex α in G' maximal in the sense that there are no outgoing edges from α, and then restricting α with appropriate subsets of G. The next equality says that we do not impose any restriction on the edge set E''. In each of (a), (b), (c) and (c'), the subgraphs used for indices meet this condition; the required W is the set of all vertices of G, and $\alpha : W \to \mathbb{N}$.

3.2 Formal Setting and Examples

Throughout the rest of the paper, we will focus on index sets generated by the recipe just described. This means that each such set comes with a finite DAG G. We will compress the recipe and directly characterize the generated index set from G. We will often abuse notation by writing G when we are referring to the vertex set of G.

We will use the fact that a finite DAG is the same thing as a finite partially ordered set (that is, a set with a binary relation $\preceq$ that is reflexive, transitive, and anti-symmetric). In particular, for a given DAG, we use the partial order on its vertices: $v \preceq w$ whenever there is a path from v to w. This is the reflexive-transitive closure of its anti-symmetric directed edge relation. Conversely, we can regard a finite partially ordered set as a DAG with a directed edge $v \to w$ if $v \prec w$ and there is no v' such that $v \prec v' \prec w$ (here, $\prec$ denotes that reflexivity does not hold).

We say that a subset W of a DAG G is (ancestrally) **closed**[2] (with respect to the partial ordering) if whenever $w' \in W$ and $w \preceq w'$ we have $w \in W$. The **closure** $\bar{W}$ of a set W is the set of all ancestors of W:

$$\bar{W} = \{w' \ : \ \text{there exists } w \in W \text{ such that } w' \preceq w\}.$$

Definition 3.1. Let C be a closed subset of G. A C**-type multi-index** is a function α from the vertices in C to $\mathbb{N}$. We write $\mathbb{N}^C$ for the set of all C-type multi-indices. If $C \subseteq D$ for some closed D, then every D-type multi-index $\alpha \in \mathbb{N}^D$ can be restricted to a C-type multi-index $\alpha|_C \in \mathbb{N}^C$.

Definition 3.2. Let $\mathscr{X}$ be a Borel space, $C \subseteq G$ be closed, and $\mathscr{C}$ be a sequence of distinct closed sets. A C**-type random array** in $\mathscr{X}$ is a family of random variables,

$$\mathbf{X} = (X_\alpha : \alpha \in \mathbb{N}^C)$$

indexed by C-type multi-indices. Here each X_α is a random variable valued in $\mathscr{X}$. A $\mathscr{C}$**-type random array collection** in $\mathscr{X}$ is a sequence of random variable families

$$(\mathbf{X}_C) = (\mathbf{X}_C : C \in \mathscr{C})$$

where each $\mathbf{X}_C$ is a C-type random array.

Remarks.

1. We have generalized to *collections* of random arrays in order to obtain later, in our main result, a fine-grained representation which also allows for representations on sub-DAGs. As mentioned in the abstract, in the special case of hierarchical exchangeability, this allows for an improvement over the Austin-Panchenko representation which is important for applications.
2. We often do not mention the Borel space $\mathscr{X}$, and just say C-type random array and/or $\mathscr{C}$-type random array collection.

Example 3.3. Most discretely indexed stochastic processes can be viewed as G-type random arrays for some DAG G. We illustrate this perspective with basic examples from the literature matching those described in Section 3.1.

[2] Closed sets are sometimes called downsets.

(a) The most common discrete-time stochastic processes are G-type random arrays where G is the graph with only a single vertex v and no edges. The index set $\mathbb{N}^G$ in this case is $\mathbb{N}^{\{v\}} \simeq \mathbb{N}$. Thus, these G-type random arrays are $\mathbb{N}$-indexed families of random variables.

(b) The graph with two vertices r, c and no edges corresponds to infinite random matrices or random arrays indexed by $\mathbb{N}^2$. In other words, the index set $\mathbb{N}^G$ is $\mathbb{N}^{\{r,c\}} \simeq \mathbb{N}^2$. Thus, a G-type index is a pair of two numbers, one denoting a row index and the other denoting a column index. These G-type random arrays have the form $(X_{n,m} : n, m \in \mathbb{N})$ and are random matrices with countably many rows and columns.

(c') Let $r_1, \ldots, r_\ell$ be nonnegative integers. Austin and Panchenko studied a stochastic process indexed by a tuple of paths over ℓ countably-branching trees that have heights $r_1, \ldots, r_\ell$, respectively [AP14]. Formally, this process is a family of random variables of the following form:

$$(X_\alpha : \alpha \in \mathbb{N}^{r_1} \times \ldots \times \mathbb{N}^{r_\ell}).$$

This stochastic process is a G-type array for the following G:

$$\begin{tikzcd}
v_1^{(1)} \arrow[r] & v_2^{(1)} \arrow[r] & \ldots \arrow[r] & v_{r_1}^{(1)} \\
v_1^{(2)} \arrow[r] & v_2^{(2)} \arrow[r] & \ldots \arrow[r] & v_{r_2}^{(2)} \\
\ldots & \ldots & \ldots & \\
v_1^{(\ell)} \arrow[r] & v_2^{(\ell)} \arrow[r] & \ldots \arrow[r] & v_{r_\ell}^{(\ell)}
\end{tikzcd}$$

So $G = \{v_j^{(i)} \mid 1 \leq i \leq \ell, 1 \leq j \leq r_i\}$ and

$$\mathbb{N}^G \simeq \mathbb{N}^{r_1} \times \ldots \times \mathbb{N}^{r_\ell}.$$

Thus, $(X_\alpha : \alpha \in \mathbb{N}^{r_1} \times \ldots \times \mathbb{N}^{r_\ell})$ is the same thing as a G-type array.

Of course, our formal setting is not limited to just recasting well-known exchangeable stochastic processes. Its recipe for defining multi-indices via a DAG makes it easy to define a random array with unusual multi-indices. Furthermore, by moving from random arrays to random array collections, we can express multiple random-variable families whose multi-index sets are related.

Example 3.4. In order to illustrate the generality of our setting, we present two instance of DAG-exchangeable arrays that we have not seen discussed previously in the exchangeability literature.

(a) *Random Block Matrices:* Consider the DAG G:

$$\begin{tikzcd}
r_0 \arrow[r] \arrow[dr] & r_1 \\
c_0 \arrow[r] \arrow[ur] & c_1
\end{tikzcd}$$

The corresponding multi-index set is $\mathbb{N}^G = \mathbb{N}^{\{r_0,c_0,r_1,c_1\}}$, which can be understood as indices of an infinite matrix each of whose entry is again an infinite matrix. For $\alpha \in \mathbb{N}^G$, the pair

$(\alpha(r_0), \alpha(c_0))$ specifies the row and column of the outer matrix, and $(\alpha(r_1), \alpha(c_1))$ those of the nested matrix. Thus, in a random G-type array, each random variable X_α stores the value of the $(\alpha(r_1), \alpha(c_1))$-th entry of the nested matrix, which is itself stored at the $(\alpha(r_0), \alpha(c_0))$-th entry of the outer nesting matrix.

Let $C = \{r_0, c_0\}$, a closed subset of G. A small generalization of the random block matrix, alluded to in the Section 2, is a random structure that is simultaneously a random matrix (Ex. 3.3(b)) and a random block matrix. This can be thought of as a random matrix where each cell contains both a value in $\mathscr{X}$ and another random matrix. It comprises both a C-type random array and a G-type random array. In other words, it is a $\{C, G\}$-type random array collection.

(b) *Random Walls:* Here is another example of a random array collection where $\mathscr{C}$ is different from $\{G\}$. Consider the graph G consisting of three vertices x, y, z and no edges. Let $\mathscr{C}$ be the following closed sets:

$$\mathscr{C} = (C_{xy}, C_{yz}, C_{zx}), \qquad C_{xy} = \{x, y\}, \qquad C_{yz} = \{y, z\}, \qquad C_{zx} = \{z, x\}.$$

A $\mathscr{C}$-type random array collection consists of three random variable families, namely, $\mathbf{X}_{C_{xy}}$, $\mathbf{X}_{C_{yz}}$ and $\mathbf{X}_{C_{zx}}$. These families use different yet related multi-index sets, $\mathbb{N}^{\{x,y\}}$, $\mathbb{N}^{\{y,z\}}$, and $\mathbb{N}^{\{z,x\}}$, respectively. A good way to understand this array collection is to imagine a 3-dimensional grid at points in $\mathbb{N}^{\{x,y,z\}}$. The collection associates a random variable for each point in the xy, yz and zx planes with the respective missing coordinate set to 1. Viewing the tuple $(\mathbf{X}_{C_{xy}}, \mathbf{X}_{C_{yz}}, \mathbf{X}_{C_{zx}})$ in this way, rather than just as three 2-dimensional random arrays, makes it easy to state and study symmetries which involve all three families, as we explain soon.

Motivated by the above examples, we are now ready to introduce the notion of partial exchangeability which we call DAG-exchangeability. We start by defining the symmetries that are displayed by such exchangeability.

Definition 3.5. Let G be a finite DAG. A G**-automorphism** of $\mathbb{N}^G$ is a bijection $\tau \colon \mathbb{N}^G \to \mathbb{N}^G$ such that

$$\alpha|_C = \beta|_C \implies (\tau(\alpha))|_C = (\tau(\beta))|_C, \quad \text{for all } \alpha, \beta \in \mathbb{N}^G \text{ and closed } C \subseteq G. \tag{1}$$

By definition, a G-automorphism τ induces a bijection on the C-type multi-indices $\beta \in \mathbb{N}^C$ for any closed $C \subseteq G$:

$$\tau(\beta) \stackrel{\text{def}}{=} \tau(\alpha)|_C, \qquad \text{for some/any } \alpha \in \mathbb{N}^G \text{ such that } \alpha|_C = \beta.$$

Slightly abusing notation, we reuse τ to denote this induced map.

Also, a bijection $\tau \colon \mathbb{N}^C \to \mathbb{N}^C$ acts on a C-type random array $\mathbf{X}_C$ by

$$\tau(\mathbf{X}_C) \stackrel{\text{def}}{=} (X_{C,\tau(\alpha)} : \alpha \in \mathbb{N}^C).$$

Definition 3.6. A $\mathscr{C}$-type random array collection $(\mathbf{X}_C : C \in \mathscr{C})$ is **DAG-exchangeable** if it is equal in distribution to $(\tau(\mathbf{X}_C) : C \in \mathscr{C})$ for every G-automorphism τ. That is,

$$(\mathbf{X}_C) \stackrel{d}{=} (\tau(\mathbf{X}_C)).$$

We say that a C-type random array is **DAG-exchangeable** if it is so, when viewed as a $\{C\}$-type random array collection.

DAG-exchangeability generalizes several popular notions of exchangeability from the literature, if we choose G appropriately and set $\mathscr{C}$ to be the singleton set consisting of $\{G\}$. For instance, when the graph G consists of only one vertex and does not have any edges, DAG-exchangeability becomes the standard notion of exchangeability for random sequences in de Finetti's classic result. When G is the graph for the infinite random matrix of Example 3.3(b), DAG-exchangeability becomes Aldous-Hoover (separate) exchangeability. When G is the graph in Example 3.3(c), DAG-exchangeability becomes Austin and Panchenko's hierarchical exchangeability. Note that DAG-exchangeability goes further and specifies new kinds of symmetries as displayed by the DAGs in Example 3.4.

DAG-exchangeability is quite strong and is only satisfied by some $\mathscr{C}$-type random array collections. It is natural to ask what such DAG-exchangeable random array collections look like. Answering this question by means of a representation theorem forms the rest of this paper.

4 Main Result

In this section we present our main result. Let G be a finite DAG, and recall that for each closed set C, $\mathbb{N}^C$ is the set of C-type multi-indices. To state our result, we denote the set of (ancestrally) closed subsets of C by

$$\mathscr{A}_C \stackrel{\text{def}}{=} \{D \,:\, D \text{ closed and } D \subseteq C\}.$$

Also, set

$$I_C \stackrel{\text{def}}{=} \bigcup\{\mathbb{N}^D \,:\, D \in \mathscr{A}_C\} \tag{2}$$

and

$$Dom\colon I_C \to \mathscr{A}_C,$$

where $Dom(\alpha)$ is the domain of α for each multi-index $\alpha \in I_G$, i.e. the set of vertices where α is defined.

We say that a G-automorphism τ of $\mathbb{N}^G$ **fixes** $\alpha \in I_G$ if $\tau(\alpha) = \alpha$. Let $\mathscr{C}$ be a sequence of distinct closed sets of vertices. Let $(\mathbf{X}_C : C \in \mathscr{C})$ be a $\mathscr{C}$-type random array collection, for which we often write simply $(\mathbf{X}_C)$. We can define $\mathscr{F}_\alpha$ to be the sub-σ-field of $\sigma\big((\mathbf{X}_C)\big)$ consisting of $(\mathbf{X}_C)$-measurable events that are invariant under every α-fixing G-automorphism τ:

$$\mathscr{F}_\alpha \stackrel{\text{def}}{=} \sigma\Big(\Big\{(\mathbf{X}_C)^{-1}(B) \,:\, B \text{ is Borel, and if } \big((x_\beta : \beta \in \mathbb{N}^C) : C \in \mathscr{C}\big) \in B \text{ and } \tau \text{ fixes } \alpha, \\ \text{then } \big((x_{\tau(\beta)} : \beta \in \mathbb{N}^C) : C \in \mathscr{C}\big) \in B\Big\}\Big)$$

One can think of $\mathscr{F}_\alpha$ as containing the information of the symmetries in the arrays which fix the multi-index α. The above definition takes a little time to digest. We give another representation of this σ-field in (17) below, which may help the reader by giving an alternative perspective.

Since the elements of each $\mathbf{X}_C$ take values in a Borel space and since each $\mathscr{F}_\alpha$ is countably generated, we can choose some random array $\mathbf{S} = (S_\alpha : \alpha \in I_G)$ (after extending the underlying probability space if needed) so that

1. $\sigma(S_\alpha) = \mathscr{F}_\alpha$ for all α, and
2. the random array collection $(\mathbf{S}_C : C \in \mathscr{C})$ with $\mathbf{S}_C = (S_\alpha : \alpha \in \mathbb{N}^C)$, satisfies

$$\Big((\mathbf{X}_C, \mathbf{S}_C) : C \in \mathscr{C}\Big) \stackrel{d}{=} \Big((\tau(\mathbf{X}_C), \tau(\mathbf{S}_C)) : C \in \mathscr{C}\Big)$$

for all G-automorphisms τ, i.e. the collection of ordered pairs $((\mathbf{X}_C, \mathbf{S}_C) : C \in \mathscr{C})$ is DAG-exchangeable.

To actually construct such $\mathbf{S}$, for each $C \in \mathscr{C}$, choose one $\alpha \in \mathbb{N}^C$ and find f such that $S_\alpha \stackrel{\text{def}}{=} f(\mathbf{X}_C : C \in \mathscr{C})$ generates $\mathscr{F}_\alpha$. Then for each G-automorphism τ let $S_{\tau(\alpha)} \stackrel{\text{def}}{=} f(\tau(\mathbf{X}_C) : C \in \mathscr{C})$. The definition is independent of the choice of τ, and $\mathbf{S}$ constructed this way satisfies the desired properties.

The array $\mathbf{S}$ contains all the information of all DAG symmetries in the arrays $(\mathbf{X}_C)$. Note that if $\beta \in I_G$ is a restriction of α, the random variable

$$S_\beta \in \mathscr{F}_\alpha. \tag{3}$$

For $\alpha \in \mathbb{N}^G$, we define

$$\mathit{Restr}(\alpha) \stackrel{\text{def}}{=} \left\{\alpha|_C \;:\; C \in \mathscr{A}_G\right\} \qquad \text{and} \qquad \mathit{Restr}'(\alpha) \stackrel{\text{def}}{=} \mathit{Restr}(\alpha) \setminus \{\alpha\}. \tag{4}$$

The former consists of all restrictions of the multi-index α, while the latter takes only the strict restrictions of α.

Theorem 4.1. *Let $\mathscr{C}$ be a sequence of distinct closed sets, and let $(\mathbf{X}_C : C \in \mathscr{C})$ be a $\mathscr{C}$-type random array collection. If $(\mathbf{X}_C : C \in \mathscr{C})$ is DAG-exchangeable, there exist a family of measurable functions $(g_C : C \in \mathscr{A}_G)$ and a collection of independent $[0,1]$-uniform random variables $\mathbf{U} = (U_\alpha : \alpha \in I_G)$ such that*

$$\Big(S_\alpha : \alpha \in I_G\Big) \stackrel{d}{=} \Big(S'_\alpha : \alpha \in I_G\Big) \tag{5}$$

where

$$S'_\alpha \stackrel{\text{def}}{=} g_{Dom(\alpha)}\left(\Big(S'_\beta : \beta \in \mathit{Restr}'(\alpha)\Big), U_\alpha\right) \tag{6}$$

for $\alpha \in I_G$.

The representation in the theorem gives rise to a representation of $(\mathbf{X}_C)$ in terms of $\mathbf{U}$, which is similar to the one found in standard representation theorems for an exchangeable family of random variables. More concretely, we can use induction and convert $g_{Dom(\alpha)}$ to a function $f_{Dom(\alpha)}$ for each α such that

$$\Big(S_\alpha : \alpha \in I_G\Big) \stackrel{d}{=} \left(f_{Dom(\alpha)}\Big(U_\beta : \beta \in \mathit{Restr}(\alpha)\Big) : \alpha \in I_G\right). \tag{7}$$

The key part of this inductive conversion is to set $f_{Dom(\alpha)}$ using the following equation:

$$f_{Dom(\alpha)}\Big(U_\beta : \beta \in \mathit{Restr}(\alpha)\Big) = \\ g_{Dom(\alpha)}\left(\Big(f_{Dom(\beta)}\big(U_\gamma : \gamma \in \mathit{Restr}(\beta)\big) : \beta \in \mathit{Restr}'(\alpha)\Big), U_\alpha\right).$$

Now note that $X_{C,\alpha}$ is measurable with respect to S_α by the choice of S_α. Furthermore, $X_{C,\alpha}$ takes values in a Borel space. Thus, there exists a measurable h_α such that $X_{C,\alpha} = h_\alpha(S_\alpha)$ almost surely. By the DAG-exchangeability of the collection of ordered pairs $\big((\mathbf{X}_C, \mathbf{S}_C)\big)$, we can pick h_α such that it depends only on $Dom(\alpha)$ and not on the value of α itself. This means that we can write $X_{C,\alpha} = h_{Dom(\alpha)}(S_\alpha)$ almost surely. By combining this observation and the representation for $(S_\alpha : \alpha \in I_G)$ in (7), we obtain:

Corollary 4.2. *If* $(\mathbf{X}_C : C \in \mathscr{C})$ *is DAG-exchangeable, then*

$$\Big(\big(X_{C,\alpha} : \alpha \in \mathbb{N}^C\big) : C \in \mathscr{C}\Big) \overset{d}{=} \left(\Big(h_{Dom(\alpha)}\Big(U_\beta : \beta \in Restr(\alpha)\Big) : \alpha \in \mathbb{N}^C\Big) : C \in \mathscr{C}\right)$$

for some family of measurable functions $(h_C : C \in \mathscr{A}_G)$ *and independent* $[0,1]$*-uniform random variables* $\mathbf{U} = (U_\alpha : \alpha \in I_G)$.

One key benefit of the representation in our theorem is that it tells us how the global information encoded in $\mathbf{U}$ gets used by the array $\mathbf{S}$, and it also tells us what information is shared by two random variables S_α and S_β. As noted before, the array $\mathbf{S}$ captures the various restricted versions of partial exchangeability of $(\mathbf{X}_C)$. This fine-grained representation eventually allows us to prove Theorem 4.1 inductively.

Remark. Before getting into the proof of the main result, we recall a generic property of exchangeable structures. Whenever $\xi = (\xi_n)_{n\in\mathbb{N}}$ is exchangeable, by Kolmogorov's extension theorem, this is equivalent to the seemingly weaker condition that the distribution of ξ is invariant under the action of *finite* permutations (permutations fixing all but finitely many elements). In particular, if ξ is exchangeable, then $(\xi_n)_{n\in\mathbb{N}} \overset{d}{=} (\xi_{\tau(n)})_{n\in\mathbb{N}}$ for any *injection* τ (in fact, Ryll-Nardzewski's theorem tells us the converse is also true). We can extend this sort of argument to other random variables associated to the symmetries of ξ.

Let η be ξ-measurable. Then, η is invariant under the action of a subgroup K of the infinite permutations $S_{\mathbb{N}}$ if and only if $(\eta, \xi) \overset{d}{=} (\eta, \tau\xi)$ for all $\tau \in K$. By the above paragraph, this is equivalent to having $(\eta, \xi) \overset{d}{=} (\eta, \tau\xi)$ for all finite permutations in K. Moreover, we have $(\eta, \xi) \overset{d}{=} (\eta, \rho\xi)$ for any injection $\rho \in K$ such that its arbitrary restriction to finite sets can be extended to an element in K. Thus, as long as any automorphism between finite substructures can be extended to the whole (infinite) structure, the distribution of the array is invariant under the action of injections. This property is called **ultrahomogeneity**. This notion also appears in model-theoretic results on exchangeability (see for example [CT17]). We discuss this in the appendix, along with the proof that the structure $\mathbb{N}^G$ under G-automorphisms is ultrahomogeneous.

5 Proof of the Main Theorem

We will use the following notation for conditional distribution properties.

- $\xi \underset{\mathscr{F}}{\perp\!\!\!\perp} \eta$ (ξ and η are conditionally independent given $\mathscr{F}$.)
- $\underset{\mathscr{F}}{\perp\!\!\!\perp} (\xi_a)_{a\in I}$ (the family $(\xi_a)_{a\in I}$ is conditionally independent given $\mathscr{F}$.)

The first lemma is a standard result from probability theory, of which the proof we omit.

Lemma 5.1. *Let* $(\xi_a, \eta_a)_{a\in I}$ *be a multi-indexed family of random variables, and let* $\mathscr{F}$ *be a* σ*-field. Assume the following hold:*

- $\mathbf{P}[\xi_a \in \cdot|\mathscr{F}] = \mathbf{P}[\eta_a \in \cdot|\mathscr{F}]$ *almost surely*
- $\underset{\mathscr{F}}{\perp\!\!\!\perp} (\xi_a)_{a\in I}$
- $\underset{\mathscr{F}}{\perp\!\!\!\perp} (\eta_a)_{a\in I}$

Then, $\mathbf{P}[(\xi_a)_a \in \cdot|\mathscr{F}] = \mathbf{P}[(\eta_a)_a \in \cdot|\mathscr{F}]$ *almost surely.*

The next lemma, which is a simple application of the previous result, is used to synchronize representations using different functions.

Lemma 5.2. *Let $\xi_0, (\xi_a)_{a\in I}$ be random variables such that $\xi_0 \stackrel{d}{=} \xi_a$, and let $\boldsymbol{\zeta} = (\zeta_a)_{a\in I}$ be a family of independent random variables, which are also independent from $\xi_0, (\xi_a)_{a\in I}$. Let $(\eta_a)_{a\in I}$ be random variables such that for some Borel measurable functions ϕ_a, the following hold:*

- $(\xi_0, \eta_a) \stackrel{d}{=} (\xi_a, \phi_a(\xi_a, \zeta_a))$ *for each* $a \in I$
- $\perp\!\!\!\perp_{\xi_0} (\eta_a)_{a\in I}$

Then, $(\xi_0, \eta_a)_{a\in I} \stackrel{d}{=} (\xi_0, \phi_a(\xi_0, \zeta_a))_{a\in I}$. *In particular, we have* $(\xi_0, \eta_a)_{a\in I} \stackrel{d}{=} (\xi', \phi_a(\xi', \zeta_a))_{a\in I}$ *for any* $\xi' \stackrel{d}{=} \xi_0$ *independent from* $\boldsymbol{\zeta}$.

Proof. Since $(\xi_0, \zeta_a) \stackrel{d}{=} (\xi_a, \zeta_a)$, we can replace ξ_a by ξ_0 in the first bullet. The result then follows by an application of Lemma 5.1. □

Our overall plan to prove the main result is to (i) obtain representations for closed proper subgraphs of G by using induction, and then (ii) glue everything into a joint distributional equality using conditional independence arguments for the S_α's. The above lemmas allow for basic gluing. More complicated gluing will follow from more sophisticated conditional independence arguments for which our next key proposition is one example. We first need more terminology.

Let the set of all terminal vertices (vertices with no descendants) of G be denoted by

$$T = \{v_1, \ldots, v_t\}, \quad \text{where } |T| = t.$$

Also, define

$$G_0 \stackrel{\text{def}}{=} G \setminus T,$$

and for each $C \in \mathscr{A}_G$, set

$$\mathscr{F}_C \stackrel{\text{def}}{=} \sigma(S_\alpha : Dom(\alpha) = C).$$

Clearly, we have $\mathscr{F}_D \subseteq \mathscr{F}_C$ whenever $D \in \mathscr{A}_C$. Finally, for $A \subseteq T$, set

$$\mathscr{G}_A \stackrel{\text{def}}{=} \mathscr{F}_{G_0 \cup A}. \tag{8}$$

Proposition 5.3. *Let $A \subseteq T$, $B = T \setminus A$. Also, let $A_1, \ldots, A_m$ be distinct subsets of A. Then,*

$$\mathscr{G}_A \underset{(\mathscr{G}_{A_i})_{i\le m}}{\perp\!\!\!\perp} (\mathscr{G}_{B\cup A_i})_{i\le m}.$$

The result may seem obvious at first glance, and it is indeed easy to prove for the case $m = 1$. However for $m > 1$, the joint conditional independence seems to be a rather subtle issue. We postpone the proof of this proposition to the next subsection. Using this result, we obtain the following corollary.

Corollary 5.4. *Let $A, B_1, \ldots, B_m \subseteq T$. Then,*

$$\mathscr{G}_A \underset{(\mathscr{G}_{A\cap B_i})_{i\le m}}{\perp\!\!\!\perp} (\mathscr{G}_{B_i})_{i\le m}.$$

Proof. For each $D \subsetneq A$, set $B_D \stackrel{\text{def}}{=} \{B_i : B_i \cap A = D\}$. We have

$$\sigma(\mathscr{G}_{B_i} : B_i \in B_D) \subseteq \mathscr{G}_{D\cup(T\setminus A)},$$

and the result follows by applying Proposition 5.3. □

We immediately also obtain the following result.

Proposition 5.5. *For $k = 0, \dots, t$, let*

$$\mathscr{H}_k \stackrel{\text{def}}{=} \{\mathscr{G}_A : A \subseteq T, |A| = k\} \quad \textit{and} \quad \mathscr{G}_k \stackrel{\text{def}}{=} \sigma(\mathscr{H}_k).$$

Then, given $\mathscr{G}_k$, the set $\mathscr{H}_{k+1}$ is an independent family of σ-fields for all $k < t$.

Proof. Let $A \subseteq T$ be such that $|A| = k$. Set

$$\mathscr{G}_{k\setminus A} \stackrel{\text{def}}{=} \sigma(\mathscr{G}_B : B \subseteq T, |B| = k, B \neq A)$$

and

$$\mathscr{G}' \stackrel{\text{def}}{=} \sigma(\mathscr{G}_B : B \subsetneq A, |B| = k-1).$$

By Corollary 5.4, $\mathscr{G}_A \underset{\mathscr{G}'}{\perp\!\!\!\perp} \mathscr{G}_{k\setminus A}$. But since $\mathscr{G}' \subseteq \mathscr{G}_{k-1}$, we have that $\mathscr{G}_A \underset{\mathscr{G}_{k-1}}{\perp\!\!\!\perp} \mathscr{G}_{k\setminus A}$, from which the result follows. □

The following lemma is the final piece of the puzzle joining all the representations obtained from inductive hypothesis. For each $J \subseteq I_G$ and generic array $(\xi_\alpha : \alpha \in I_G)$, we will write

$$\boldsymbol{\xi}_J \stackrel{\text{def}}{=} (\xi_\alpha : \alpha \in J).$$

Lemma 5.6. *Set*

$$\tilde{\mathbf{S}} \stackrel{\text{def}}{=} (S_\alpha : \alpha \in I_G \setminus \mathbb{N}^G).$$

There exists a Borel measurable function f_G such that for any i.i.d. array of uniform random variables $(U_\alpha : \alpha \in \mathbb{N}^G)$ which is independent from $\mathscr{S}$,

$$\Big(\tilde{\mathbf{S}}, (S_\alpha : \alpha \in \mathbb{N}^G)\Big) \stackrel{d}{=} \Big(\tilde{\mathbf{S}}, (f_G(\mathbf{S}_{Restr'(\alpha)}, U_\alpha) : \alpha \in \mathbb{N}^G)\Big). \tag{9}$$

The proof of this lemma is similar to that of Proposition 5.3. We postpone it for now in order to present the proof of the main theorem.

Proof of Theorem 4.1. Without loss of generality we will assume that $\mathscr{C}$ is the set $\mathscr{A}_G$ with some fixed ordering. We use induction on the number of vertices of G. The $n = 1$ case is simply the de Finetti-Hewitt-Savage theorem. (See Lemma 7.1, [Kal05] for example.) Now assume that the theorem holds for all graphs where the number of vertices is less than n, and let $|G| = n$. The case where $|T| = 1$ is obtained directly from Lemma 5.6 and the inductive hypothesis, so we assume that $|T| \equiv t > 1$.

By the inductive hypothesis, there exist Borel functions $(g_C : C \in \mathscr{A}_{G_0})$ and $(g^s_C : C \in \mathscr{A}_{G_0 \cup \{v_s\}})$ for $s = 1, \dots, t$, as well as an array of i.i.d. uniform random variables, $\mathbf{U}$, which we can assume to be independent from $(S_\alpha : \alpha \in I_G)$, such that

$$\begin{aligned} (S_\alpha : \alpha \in I_{G_0}) &\stackrel{d}{=} (S'_\alpha : \alpha \in I_{G_0}), \\ S'_\alpha &\stackrel{\text{def}}{=} g_{Dom(\alpha)}(\mathbf{S}'_{Restr'(\alpha)}, U_\alpha) \qquad \text{for } \alpha \in I_{G_0} \end{aligned} \tag{10}$$

and

$$\begin{aligned} (S_\alpha : \alpha \in I_{G_0 \cup v_s}) &\stackrel{d}{=} (S^s_\alpha : \alpha \in I_{G_0 \cup v_s}), \\ S^s_\alpha &\stackrel{\text{def}}{=} g^s_{Dom(\alpha)}(\mathbf{S}^s_{Restr'(\alpha)}, U_\alpha) \qquad \text{for } \alpha \in I_{G_0 \cup v_s}. \end{aligned} \tag{11}$$

Now, set
$$\boldsymbol{\xi} \stackrel{\text{def}}{=} \mathbf{S}_{I_{G_0}}, \qquad \boldsymbol{\xi}' \stackrel{\text{def}}{=} \mathbf{S}'_{I_{G_0}}, \qquad \boldsymbol{\xi}_s \stackrel{\text{def}}{=} \mathbf{S}^s_{I_{G_0}},$$
and
$$\boldsymbol{\eta}_s \stackrel{\text{def}}{=} \mathbf{S}_{I_{G_0\cup\{v_s\}}}, \quad \boldsymbol{\zeta}_s \stackrel{\text{def}}{=} \mathbf{U}_{I_{G_0\cup\{v_s\}}\setminus I_{G_0}}, \quad \boldsymbol{\theta}_s \stackrel{\text{def}}{=} \mathbf{S}^s_{I_{G_0\cup\{v_s\}}}.$$
Then, the second equation of (11) can be written as $\boldsymbol{\theta}_s = F_s(\boldsymbol{\xi}_s, \boldsymbol{\zeta}_s)$ for an appropriate measurable F_s. By Proposition 5.5, $\perp\!\!\!\perp_{\boldsymbol{\xi}} (\boldsymbol{\eta}_s)_s$. By construction, $\boldsymbol{\xi} \stackrel{d}{=} \boldsymbol{\xi}' \stackrel{d}{=} \boldsymbol{\xi}_s$ are independent from $(\boldsymbol{\zeta}_s)_s$, and $(\boldsymbol{\xi}, \boldsymbol{\eta}_s) \stackrel{d}{=} (\boldsymbol{\xi}_s, \boldsymbol{\theta}_s)$. Therefore, by Lemma 5.2, we have $(\boldsymbol{\xi}, \boldsymbol{\eta}_s)_s \stackrel{d}{=} (\boldsymbol{\xi}', F_s(\boldsymbol{\xi}', \boldsymbol{\zeta}_s))_s$. Thus we can join the representations given by (10) and (11) to obtain the following joint distributional equality
$$(S_\alpha : \alpha \in I_{G_0\cup\{v_s\}}, s \le t) \stackrel{d}{=} (S^1_\alpha : \alpha \in I_{G_0\cup\{v_s\}}, s \le t), \tag{12}$$
where $S^1_\alpha = S'_\alpha$ if $\alpha \in I_{G_0}$ and the rest of the S^1_α's are defined by the recursive formulae
$$S^1_\alpha = g^s_{Dom(\alpha)}(\mathbf{S}^1_{Restr'(\alpha)}, U_\alpha).$$

This is a joint representation of $(S_\alpha : Dom(\alpha) \in \mathscr{A}_{G_0\cup A}, A \subseteq T, |A| = k)$, where $k = 1$. We will next induct on k to achieve an analogous joint representation for $k = t-1$. Let $k < t-1$ be fixed and assume that we have the following joint representation:
$$\begin{aligned}(S_\alpha : \alpha \in I_{G_0\cup A}, A \subseteq T, |A| = k) &\stackrel{d}{=} (S^k_\alpha : \alpha \in I_{G_0\cup A}, A \subseteq T, |A| = k)\\ S^k_\alpha &\stackrel{\text{def}}{=} g_{Dom(\alpha)}(\mathbf{S}^k_{Restr'(\alpha)}, U_\alpha) \qquad \text{for } \alpha \in \bigcup_{|A|=k} I_{G_0\cup A}.\end{aligned} \tag{13}$$

Since we have assumed that we have a representation for all the proper closed subgraphs of G, then for any fixed $B \subseteq T$ with $|B| = k+1$ (note that $k+1 \le t-1$), we have some Borel measurable functions $(g^B_C : C \in \mathscr{A}_{G_0\cup B})$ such that
$$\begin{aligned}(S_\alpha : \alpha \in I_{G_0\cup B}) &\stackrel{d}{=} (S^B_\alpha : \alpha \in I_{G_0\cup B}),\\ S^B_\alpha &\stackrel{\text{def}}{=} g^B_{Dom(\alpha)}(\mathbf{S}^B_{Restr'(\alpha)}, U_\alpha)) \qquad \text{for } \alpha \in I_{G_0\cup B}.\end{aligned} \tag{14}$$

Define the following arrays (to ease notation we do not use boldface for these):
$$\begin{aligned}
\varphi^0_B &\stackrel{\text{def}}{=} (S_\alpha : \alpha \in I_{G_0\cup A}, A \subseteq B, |A| = k) & \varphi'_B &\stackrel{\text{def}}{=} (S^k_\alpha : \alpha \in I_{G_0\cup A}, A \subseteq B, |A| = k)\\
\varphi_B &\stackrel{\text{def}}{=} (S^B_\alpha : \alpha \in I_{G_0\cup A}, A \subseteq B, |A| = k) & \Upsilon_B &\stackrel{\text{def}}{=} (U_\alpha : \alpha \in I_{G_0\cup A}, A \subseteq B, |A| = k)\\
\psi &\stackrel{\text{def}}{=} (U_\alpha : \alpha \in I_{G_0\cup A}, A \subseteq T, |A| = k) \setminus \Upsilon_B & \psi_B &\stackrel{\text{def}}{=} (U_\alpha : \alpha \in I_{G_0\cup B}) \setminus \Upsilon_B\\
\Gamma &\stackrel{\text{def}}{=} (S_\alpha : \alpha \in I_{G_0\cup A}, A \subseteq T, |A| = k) & \Gamma_B &\stackrel{\text{def}}{=} (S_\alpha : \alpha \in I_{G_0\cup B}),\\
\vartheta &\stackrel{\text{def}}{=} (S^k_\alpha : \alpha \in I_{G_0\cup A}, A \subseteq T, |A| = k), & \vartheta_B &\stackrel{\text{def}}{=} (S^B_\alpha : \alpha \in I_{G_0\cup B}).
\end{aligned}$$

where $\setminus\Upsilon_B$ denotes deletion of the array Υ_B. Then, by Corollary 5.4, Γ and Γ_B are conditionally independent given φ^0_B, and by construction $\varphi^0_B \stackrel{d}{=} \varphi'_B \stackrel{d}{=} \varphi_B$, all of them independent from ψ and ψ_B. Also, by the first lines of (13) and (14) we have $\Gamma \stackrel{d}{=} \vartheta$ and $\Gamma_B \stackrel{d}{=} \vartheta_B$. Finally, for an appropriate F_B, we can write the second line of (14) as $\vartheta_B = F_B(\varphi_B, \psi_B)$. Also note that $\vartheta = G(\Upsilon_B, \psi)$ for an appropriate G, but we can replace Υ_B with φ'_B by recursively applying (13). So we have $\vartheta = F(\varphi'_B, \psi)$ for some F.

Now, by Lemma 5.2,

$$(\varphi_B^0, \Gamma, \Gamma_B) \stackrel{d}{=} (\varphi_B', F(\varphi_B', \psi), F_B(\varphi_B', \psi_B))$$

and in particular

$$(\Gamma, \Gamma_B) \stackrel{d}{=} (F(\varphi_B', \psi), F_B(\varphi_B', \psi_B)) = (\vartheta, F_B(\varphi_B', \psi_B)).$$

Moreover, we have $\underset{\Gamma}{\perp\!\!\!\perp} (\Gamma_B)_B$ by Proposition 5.5 and $\underset{\vartheta}{\perp\!\!\!\perp} (F_B(\varphi_B', \psi_B))_B$ since $F_B(\varphi_B', \psi_B)$ is a function of $(\vartheta, \psi_B \setminus \psi)$, while $(\psi_B \setminus \psi)_B$ is an independent family which is also independent from ϑ. Thus, a slight variation of Lemma 5.1 shows that $(\Gamma, \Gamma_B)_B \stackrel{d}{=} (\vartheta, F_B(\varphi_B', \psi_B))_B$. Therefore, we have

$$(S_\alpha : \alpha \in I_{G_0 \cup B}, B \subseteq T, |B| = k+1) \stackrel{d}{=} (S_\alpha^{k+1} : \alpha \in I_{G_0 \cup B}, B \subseteq T, |B| = k+1) \qquad (15)$$

where $S_\alpha^{k+1} = S_\alpha^k$ for $\alpha \in I_{G_0 \cup A}, A \subseteq T, |A| = k$ and for other $\alpha \in \mathbb{N}^{G_0 \cup B}$, the random variable S_α^{k+1} is defined through the recursive formulae

$$S_\alpha^{k+1} = g_{Dom(\alpha)}^B(\mathbf{S}_{Restr'(\alpha)}^{k+1}, U_\alpha).$$

Thus we have built a $(k+1)$-version of (13). By induction on k, we obtain the following representation, which involves everything except for the S_α's for $\alpha \in \mathbb{N}^G$.

$$\begin{aligned}(S_\alpha : \alpha \in I_G \setminus \mathbb{N}^G) &\stackrel{d}{=} (S'_\alpha : \alpha \in I_G \setminus \mathbb{N}^G),\\ S'_\alpha &\stackrel{\text{def}}{=} g_{Dom(\alpha)}(\mathbf{S}'_{Restr'(\alpha)}, U_\alpha) \qquad \text{for } \alpha \in I_G \setminus \mathbb{N}^G. \end{aligned} \qquad (16)$$

Plugging (16) into (9) and applying Lemma 5.1 completes the proof. ☐

5.1 Proof of Proposition 5.3 and Lemma 5.6

Let us say that an automorphism τ is **separate** if for all $v \in G$, there exists $\tau_v \in S_{\mathbb{N}}$ such that $\tau(\beta)(v) = \tau_v(\beta(v))$ for all $\beta \in \mathbb{N}^G$. The term comes from the fact that an array is separately exchangeable if and only if its distribution is invariant under the action of every separate automorphism. Note that every DAG-exchangeable array $\mathbf{X}$ is automatically separately exchangeable. Therefore, for $\mathbf{X}$, we may define $\mathscr{F}_\alpha^{sep}$ to be the σ-field of all events which are invariant under the action of any separate automorphism that fixes some multi-index $\alpha \in \mathbb{N}^H$. We emphasize that $\mathscr{F}_\alpha^{sep}$ is defined for any $\alpha \in \mathbb{N}^H$ for any arbitrary subgraph $H \subset G$ which is not necessarily closed. We will use the letter H below, to denote arbitrary subgraphs of G.

The missing ingredient, common to the proofs of both Proposition 5.3 and Lemma 5.6, at least in the setting of separate exchangeability, is the following conditional independence result which appears as Corollary 5.6 in the celebrated paper of Hoover [Hoo79]:

Proposition 5.7. *Define $\mathscr{F}_\alpha^{sep}$ as above. Let $I_1, I_2, I_3 \subseteq \bigcup_{H \subseteq G} \mathbb{N}^H$ be such that, for all $\alpha_1 \in I_1$, $\alpha_2 \in I_2$, we have $\alpha_1 \cap \alpha_2 \in I_3$.*

Then, $(\mathscr{F}_\alpha^{sep} : \alpha \in I_1)$ is independent from $(\mathscr{F}_\alpha^{sep} : \alpha \in I_2)$ given $(\mathscr{F}_\alpha^{sep} : \alpha \in I_3)$.

Since the above result is essentially for separately exchangeable arrays, in order to apply this result, we must first establish a relationship between σ-fields related to separate exchangeability and DAG-exchangeability, respectively. The next two lemmas will serve to establish that relationship.

Fix a sequence $\mathscr{C}$ of distinct elements of $\mathscr{A}_G$ and for $\alpha \in \mathbb{N}^G$, define

$$\mathbf{X}_\alpha \stackrel{\text{def}}{=} (X_{C, \alpha|_C} : C \in \mathscr{C})$$

so that henceforth

$$\mathbf{X} \stackrel{\text{def}}{=} (\mathbf{X}_\alpha : \alpha \in \mathbb{N}^G).$$

We can re-write the definition of $\mathscr{F}_\alpha$ for $\alpha \in I_G$ as

$$\mathscr{F}_\alpha \stackrel{\text{def}}{=} \sigma\Big(\Big\{\mathbf{X}^{-1}(B) \ : \ B \text{ is Borel, and if } \big(x_\beta : \beta \in \mathbb{N}^G\big) \in B \text{ and } \tau \text{ fixes } \alpha, \quad \text{then } \big(x_{\tau(\beta)} : \beta \in \mathbb{N}^G\big) \in B\Big\}\Big). \tag{17}$$

Let us point out that an α-fixing automorphism τ fixes $\alpha|_C$ for $C \in \mathscr{A}_{Dom(\beta)}$, but not for every arbitrary subgraph H.

For a subgraph $H \subseteq G$, not necessarily closed, let H^o denote the largest subgraph of H which is closed in G. The closed graph H^o is well-defined since a union of closed subgraphs is again closed. For example, if $V(G) = \{v_1, v_2, v_3, v_4\}$, $E(G) = \{\overrightarrow{v_1v_2}, \overrightarrow{v_2v_3}, \overrightarrow{v_3v_4}\}$ and $H = \{v_1, v_3\}$, $K = \{v_2, v_3, v_4\}$, then we have $H^o = \{v_1\}$, $K^o = \emptyset$.

Lemma 5.8. *Let $\alpha \in I_G$. Then, $\mathscr{F}_\alpha = \bigcap_{n\geq 1} \mathscr{F}_\alpha^n = \bigcap_{n\geq 1} \mathscr{G}_\alpha^n$, where*

$$\mathscr{F}_\alpha^n \stackrel{\text{def}}{=} \sigma(\mathbf{X}_\beta : \beta \in \mathbb{N}^G, \textit{ For some } C \in \mathscr{A}_{Dom(\alpha)}, \beta|_C = \alpha|_C, \beta(v) > n \textit{ for all } v \notin C)$$

and

$$\mathscr{G}_\alpha^n \stackrel{\text{def}}{=} \sigma(\mathbf{X}_\beta : \beta \in \mathbb{N}^G, \textit{ For some } H \subseteq Dom(\alpha), \beta|_H = \alpha|_H, \beta(v) > n \textit{ for all } v \notin H).$$

Proof. Define an injection $\tau_k : \mathbb{N} \to \mathbb{N}$, where

$$\tau_k(k) = k, \quad \tau_k(k-1) = k+1, \quad \text{and} \quad \tau_k(m) = m+1 \ \text{ otherwise.}$$

Let ρ_1 be the injection such that $\rho_1(\beta)(v) = \tau_{\alpha(v)}(\beta(v))$ if $v \in Dom(\alpha)$ and $\rho_1(\beta)(v) = \beta(v) + 1$ otherwise. Since ρ_1 fixes α, we have $\rho_1(E) = E$ for all $E \in \mathscr{F}_\alpha$ and we can easily check that $\rho_1^n(F) \in \mathscr{G}_\alpha^n$ for any event F. Thus, by acting ρ_1^n on $\mathbf{X}$, we obtain $\mathscr{F}_\alpha \subseteq \mathscr{G}_\alpha^n$.

Now we show the other direction. Consider an automorphism ρ that fixes α and also such that $\rho(\beta)(v) = \beta(v)$ whenever there is $u \preceq v$ such that $\beta(u) > n$. We claim that $\mathscr{G}_\alpha^n$ is invariant under the action of ρ. Note that as we take $n \to \infty$, such ρ's generate all the finite automorphisms fixing α.

For a fixed $H \subseteq Dom(\alpha)$ and $\beta \in \mathbb{N}^G$ such that $\beta|_H = \alpha|_H$ and $\beta(v) > n$ otherwise, it is clear that $\rho(\beta)(v) = \beta(v)$ for $v \notin H$ or $v \in H^o$. For $v \in H \setminus H^o$, there exists $u \notin H$ such that $u \prec v$ and $\beta(u) > n$. (Otherwise, the closure of $H^o \cup \{v\}$ is in H, contradicting the maximality of H^o.) Thus $\rho(\beta)(v) = \beta(v)$. Therefore, combined with the remark at the end of Section 4, this shows that $\bigcap_{n\geq 1} \mathscr{G}_\alpha^n$ is invariant under any automorphism fixing α. Therefore, we have $\bigcap_{n\geq 1} \mathscr{G}_\alpha^n \subseteq \mathscr{F}_\alpha$ and thus also $\mathscr{F}_\alpha = \bigcap_{n\geq 1} \mathscr{G}_\alpha^n$.

Since it is obvious that $\mathscr{F}_\alpha^n \subseteq \mathscr{G}_\alpha^n$, the proof is complete if we show that $\mathscr{F}_\alpha \subseteq \bigcap_{n\geq 1} \mathscr{F}_\alpha^n$. Let n be large enough so that $n > \max_{v \in Dom(\alpha)} \alpha(v)$. Now define $\rho_2(\beta)(v) = \tau_{\alpha(v)}(\beta(v))$ if $v \in Dom(\alpha)$ and $\beta(u) = \alpha(u)$ for all $u \preceq v$, and $\rho_2(\beta)(v) = \beta(v) + 1$ otherwise. Then, ρ_2 is an injection fixing α.

Claim. If $H \subseteq Dom(\alpha)$ and H is not a closed subgraph, then for all β such that $\beta|_H = \alpha|_H$ and $\beta(v) > \alpha(v)$ for all $v \in Dom(\alpha) \setminus H$, we have $\rho_2(\beta)(v) = \beta(v)$ if $v \in H^o$ and $\rho_2(\beta)(v) = \beta(v) + 1$ otherwise.

The claim is obvious for $v \in H^o$ or $v \notin H$. If $v \in H \setminus H^o$, there exists $u \in Dom(\alpha) \setminus H$ such that $u \prec v$, because otherwise the closure of $H^o \cup \{v\}$ is in H, contradicting the maximality of H^o. Therefore, $\beta(u) > \alpha(u)$. This implies that $\beta(v) = \beta(v) + 1$, proving the claim.

We have $\rho_2^n(E) = E$ for all $E \in \mathscr{F}_\alpha$. By the claim, we have $\rho_2^n(F) \in \mathscr{F}_\alpha^n$ for all $F \in \mathscr{G}_\alpha^n$. Since we have $\mathscr{F}_\alpha \subseteq \mathscr{G}_\alpha^n$, by acting ρ_2^n on $\mathbf{X}$ we obtain $\mathscr{F}_\alpha \subseteq \mathscr{F}_\alpha^n$. $\square$

By Lemma 5.8, we immediately obtain that $\mathscr{F}_\alpha^{sep} = \bigcap_{n\geq 1} \mathscr{G}_\alpha^n$. Thus, if $Dom(\alpha)$ is a closed subgraph, we obtain $\mathscr{F}_\alpha^{sep} = \mathscr{F}_\alpha$. The next lemma finds a generalization of this fact for arbitrary $\alpha \in \mathbb{N}^H$, $H \subseteq G$.

Lemma 5.9. *Let $\alpha \in \mathbb{N}^H$, $H \subseteq G$. Define $\alpha^o \stackrel{\text{def}}{=} \alpha|_{Dom(\alpha)^o}$. Then, $\mathscr{F}_\alpha^{sep} = \mathscr{F}_{\alpha^o}$.*

Proof. It is clear that $\mathscr{F}_{\alpha^o} = \mathscr{F}_{\alpha^o}^{sep} \subseteq \mathscr{F}_\alpha^{sep}$. Similarly to the proof of the previous lemma, one can show that $\mathscr{G}_{\alpha^o}^n$ is fixed by any automorphism τ that fixes α^o and $\tau(\beta)(v) = \beta(v)$ whenever $\beta(u) > n$ for some $u \preceq v$. This shows that $\mathscr{F}_\alpha^{sep} \subseteq \mathscr{F}_{\alpha^o}$. □

Now we are ready to complete the proof of our main results.

Proof of Proposition 5.3. By Lemma 5.8, we have

$$\mathscr{G}_A = \sigma(\mathscr{F}_\alpha : \alpha \in I_{H_0\cup A}) = \sigma(\mathscr{F}_\alpha^{sep} : \alpha \in I_{H_0\cup A})$$

and

$$\mathscr{G}_{B\cup A_i} = \sigma(\mathscr{F}_\alpha : \alpha \in I_{H_0\cup B\cup A_i}) = \sigma(\mathscr{F}_\alpha^{sep} : \alpha \in I_{H_0\cup B\cup A_i}).$$

Now let $\alpha_1 \in I_{H_0\cup A}$ and $\alpha_2 \in I_{H_0\cup B\cup A_i}$. Then, $Dom(\alpha_1 \cap \alpha_2) \subseteq H_0 \cup A_i$. Thus, we have

$$\mathscr{G}_A \underset{\mathscr{G}'}{\perp\!\!\!\perp} \mathscr{G}_{B\cup A_1}, \ldots, \mathscr{G}_{B\cup A_m}$$

where

$$\mathscr{G}' \stackrel{\text{def}}{=} \sigma(\mathscr{F}_\alpha^{sep} : Dom(\alpha) \subseteq H_0 \cup A_i \text{ for some } i \leq m).$$

However, Lemmas 5.8 and 5.9 imply that

$$\begin{aligned} \mathscr{G}' &= \sigma(\mathscr{F}_\alpha : \alpha \in I_{H_0\cup A_i} \text{ for some } i \leq m) \\ &= \sigma(\mathscr{G}_{A_i}, i \leq m). \end{aligned}$$

□

To prove Lemma 5.6, we deploy Lemma 7.8, [Kal05].

Lemma 5.10 (Coding Lemma)**.** *Let $(\boldsymbol{\xi}, \boldsymbol{\eta}) = ((\xi_a, \eta_a) : a \in A)$ be an array with any multi-index set A. Assume that*

- $(\xi_a, \eta_a) \stackrel{d}{=} (\xi_b, \eta_b)$ *for all $a, b \in A$.*
- $\xi_a \underset{\eta_a}{\perp\!\!\!\perp} (\xi_b)_{b\neq a}, \boldsymbol{\eta}$ *for all $a \in A$.*

Then, at the cost of changing the probability space, there exists a Borel function f and an i.i.d. array of uniform random variables $\boldsymbol{\zeta} = (\zeta_a : a \in A)$ such that $\boldsymbol{\zeta} \perp\!\!\!\perp \boldsymbol{\eta}$ and $\xi_a = f(\eta_a, \zeta_a)$ almost surely for all $a \in A$.

Proof of Lemma 5.6. Fix $\alpha \in \mathbb{N}^G$ and $\beta_1, \ldots, \beta_r \in I_G$ such that $\beta_k \neq \alpha$. Similarly to the proof of Proposition 5.3, we can show that

$$S_\alpha \underset{\mathbf{S}_{Restr'(\alpha)}}{\perp\!\!\!\perp} S_{\beta_1}, \ldots, S_{\beta_r}.$$

Since $\beta_1, \ldots, \beta_r$ is arbitrary, we have

$$S_\alpha \underset{\mathbf{S}_{Restr'(\alpha)}}{\perp\!\!\!\perp} (S_\beta : \beta \in I_G, \beta \neq \alpha).$$

The joint distributional equality (9) can now be obtained by applying Lemma 5.10 for $\xi_\alpha = S_\alpha$ and $\eta_\alpha = \mathbf{S}_{Restr'(\alpha)}$. □

A Model-Theoretic Proof of a Simpler Representation Theorem

Several authors have recently used model-theoretic tools to prove representation theorems for a broad class of exchangeable random structures (e.g. [Ack15, AFP16, CT18, CT17]). In the appendix, we derive a simplified version of Theorem 4.1 from a representation theorem of Crane and Towsner [CT17]. This simplified version appears here because we have not yet been able to derive the full version from model theoretic results.

Let G be a finite DAG, $\mathscr{C}$ a family of closed sets, and $(\mathbf{X}_C : C \in \mathscr{C})$ a $\mathscr{C}$-type random array collection in a Borel space $\mathscr{X}$. Recall that for each closed set C, $\mathbb{N}^C$ is the set of C-type indices. Here is the simplified version of Theorem 4.1 that we mentioned.

Theorem A.1. *If $\mathscr{C}$ is the singleton sequence (G) and $(\mathbf{X}_C : C \in \mathscr{C})$ is DAG-exchangeable, then there exists a measurable function f such that*

$$\big(X_\alpha : \alpha \in \mathbb{N}^G\big) \overset{d}{=} \Big(f\big(U_{\alpha|_C} : C \in \mathscr{A}_G\big) : \alpha \in \mathbb{N}^G\Big) \tag{18}$$

where $\alpha|_C$ is the restriction of α to the vertices in C, and the U_β are independent $[0,1]$-uniform random variables.

In the statement of the theorem, we removed one level of indirection in the indices of random variables, and wrote X_α instead of the technically accurate $X_{G,\alpha}$.

Our proof of this theorem uses a result of Crane and Towsner on the representation of relatively exchangeable random structures [CT17]. Their result has been formulated and proved in a model-theoretic setting. A large part of our proof is about translating the graph-theoretic statement of Theorem A.1 to a model-theoretic one in Crane and Towsner's representation theorem, and showing that after translated, the statement satisfies the conditions of Crane and Towsner, and when translated backwards, their conclusion gives the claimed representation of our theorem.

A.1 Review of Crane and Towsner's Representation Theorem

The following theorem is a minor variant of Crane and Towsner's result in [CT18]. In the theorem, we highlight unexplained terminologies with boldface font, to emphasize that we do not expect a reader to understand them at this point.

Theorem A.2 (Crane, Towsner)**.** *Let $\mathscr{M} = (I, R_1, \dots, R_n)$ be a countably infinite set I with equivalence relations R_k on it, and $\mathscr{X}$ a Borel space. Assume that*

- *$\mathscr{M}$ is an* ***ultrahomogeneous structure****, and*
- *$(R_k : k \le n)$ is an* ***orderly sequence*** *of equivalence relations.*

Then, given a family of $\mathscr{X}$-valued random variables $\mathbf{X} = (X_\alpha : \alpha \in I)$, if the family is ***relatively exchangeable with respect to $\mathscr{M}$*** *(in short,* ***$\mathscr{M}$-exchangeable****), there exists a measurable function f such that*

$$\Big(X_\alpha : \alpha \in I\Big) \overset{d}{=} \Big(f\big(U_b : b \in B(\alpha)\big) : \alpha \in I\Big) \tag{19}$$

where

- *$B(\alpha)$ is the set of all* ***anti-chains*** *in $E(\alpha) \overset{\text{def}}{=} \{[\alpha]_{R_k} : k \le n\} \cup \{\{\alpha\}\}$, the collection of all equivalence classes $[\alpha]_{R_k}$ of α with respect to R_k's, partially-ordered by set inclusion, and*
- *$(U_b : b \in \bigcup_{\alpha \in I} B(\alpha))$ is a collection of independent $[0,1]$-uniform random variables.*

Remark. The original theorem [CT18] has an additional condition that $\mathscr{M}$ satisfies the so-called ω-DAP condition up to the R_k's. In this paper, we consider only a special case of the theorem, and in that case, this condition always holds. It is thus omitted in our presentation of the theorem.

Most of the boldfaced terms are concepts from model theory. In the rest of this subsection, we explain slightly simplified versions of their definitions. For official definitions and detailed backgrounds of these terminologies, see Crane and Towsner's papers [CT17, CT18].

A **structure** $\mathscr{M}$ of type n for some natural number n is a tuple $(I, R_1, \ldots, R_n)$ of a set I and binary relations $\{R_k\}$ on I. When another structure $\mathscr{N} = (J, S_1, \ldots, S_n)$ of the same type satisfies $J \subseteq I$ and $S_k \subseteq R_k$ for all k, we say that it is a **substructure** of $\mathscr{M}$. A common way of generating a substructure is to restrict $\mathscr{M}$ with a subset J_0 of I:

$$\mathscr{M}|_{J_0} \stackrel{\text{def}}{=} (J_0,\, R_1 \cap (J_0 \times J_0),\, \ldots,\, R_n \cap (J_0 \times J_0)).$$

An **embedding** τ from a structure $\mathscr{N} = (J, S_1, \ldots, S_n)$ to a structure $\mathscr{M} = (I, R_1, \ldots, R_n)$ is a function $\tau : J \to I$ such that τ is injective and satisfies

$$\alpha\,[J_k]\,\beta \iff \tau(\alpha)\,[R_k]\,\tau(\beta) \quad \text{for all } \alpha, \beta \in J \text{ and all } k \in [n].$$

Here $[n] = \{m \in \mathbb{N} \,:\, 1 \leq m \leq n\}$. Note that an embedding from $\mathscr{N}$ to $\mathscr{M}$ implies that $\mathscr{N}$ is essentially the same as $\mathscr{M}|_{\tau(J)}$, and provides a sense that $\mathscr{N}$ is a substructure of $\mathscr{M}$ modulo renaming of elements of $\mathscr{N}$. When the embedding is surjective and $\mathscr{M} = \mathscr{N}$, we call τ an **automorphism**.

Crane and Towsner used a structure $\mathscr{M} = (I, R_1, \ldots, R_n)$ with a countably infinite I, to specify an index set for a random-variable family and also a symmetry property of that family. The index set is I itself. They say that a family of random variables $\mathbf{X} = (X_\alpha : \alpha \in I)$ with this index set is **relatively exchangeable with respect to** $\mathscr{M}$ or $\mathscr{M}$**-exchangeable** if for all finite subsets J of I and embeddings $\tau : \mathscr{M}|_J \to \mathscr{M}$,

$$\tau(\mathbf{X}) \stackrel{d}{=} \mathbf{X}$$

where $\tau(\mathbf{X}) \stackrel{\text{def}}{=} \big(X_{\tau(\alpha)} : \alpha \in I\big)$. Embeddings play the role of finite permutations on $\mathbb{N}$ in the standard notion of exchangeability for random sequences.

Nearly all of the remaining terminology in Theorem A.2 describe properties on a structure $\mathscr{M} = (I, R_1, \ldots, R_n)$. More specifically, they impose requirements on the R_k's, and in doing so, they gauge the $\mathscr{M}$-exchangeability condition.

Definition A.3. The structure $\mathscr{M}$ is **ultrahomogeneous** if for all finite substructures

$$\mathscr{N} = (J, S_1, \ldots, S_k)$$

of $\mathscr{M}$ and embeddings τ from $\mathscr{N}$ to $\mathscr{M}$, there exists an automorphism υ on I that extends τ, i.e., $\upsilon|_J = \tau$.

A representative example of an ultrahomogeneous structure is $(\mathbb{Q}, <)$, the set of rational numbers with the usual less-than relation, while a representative counterexample is $(\mathbb{Z}, <)$, the set of integers with the less-than relation. The latter is not ultrahomogeneous because the function τ mapping 2 to 2 and 3 to 4 is an embedding from $(\{2, 3\}, <)$ to $(\mathbb{Z}, <)$, but cannot be extended to the required global function υ on $\mathbb{Z}$. The lack of any integers strictly between 2 and 3 prevents the construction of such an υ. The structure $(\mathbb{Q}, <)$ is dense, and does not suffer from this kind of problem. These examples highlight one intuition behind ultrahomogeneity: that $\mathscr{M}$ does not add any further constraint nor information to that which is present already in an embeddable finite structure.

Our next task is to explain when a sequence $(R_k : k \leq n)$ of equivalence relations of the structure $\mathscr{M}$ is **orderly**. Many binary relations on the underlying set I of $\mathscr{M}$ will appear in our explanation. We call such binary relations simply relations, without mentioning that they are on the set I. Also, R_k refers to the R_k of $\mathscr{M}$. Finally, we remind the reader that I is a countable set and so an equivalence relation on I has only a countable number of equivalence classes.

Definition A.4. A relation R is **basic explicit** in $R_1, \ldots, R_m$ if R has one of the following three forms:

- $R = R_k$ for some k;
- $R = \{(\alpha, \alpha) \,:\, \alpha \in I\}$;
- $R = I_0 \times I$ or $R = I \times I_0$ for some subset I_0 of I that can be defined by a first-order logic formula φ. The formula φ here has one free variable, say x, and may use n symbols $r_1, \ldots, r_n$ for binary relations that are interpreted as $R_1, \ldots, R_n$, in addition to the usual quantifiers and logical connectives from first-order logic. This means $I_0 = \{\alpha \,:\, \varphi(x) \text{ holds when } x = \alpha\}$.

A relation is **explicit** in $R_1, \ldots, R_k$ if it is a Boolean combination of basic explicit relations in $R_1, \ldots, R_k$.

Definition A.5. An equivalence relation S **contains** an equivalence relation R if

$$x\,[R]\,y \implies x\,[S]\,y,$$

or equivalently every equivalence class of R is contained in one of the equivalence classes of S. If, in addition, every equivalence class of S contains the same number (possibly countably infinite) of equivalence classes of R, we say that S **evenly contains** R, and write $\#_R(S)$ for that number. The relation S is said to **freely contain** R if S not only evenly contains R but also satisfies the following condition: for all equivalence classes D of S, partitions $\{D_1, \ldots, D_m\}$ of D made out of equivalence classes D_i of R, and permutations π on $[m]$, there exists an automorphism υ on $\mathscr{M}$ such that[3]

- $\upsilon(D_k) = D_{\pi(k)}$ for all $k \in [m]$; and
- $\upsilon(D') = D'$ for all the other equivalence classes D' of R.

To gain intuition, consider the special case that the structure $\mathscr{M}$ is $(\mathbb{N}^2,\ R_1,\ R_2)$ with the following equivalence relations R_1 and R_2:

$$(k_1, k_2)\,[R_1]\,(k_1', k_2') \iff k_1 = k_1', \qquad (k_1, k_2)\,[R_2]\,(k_1', k_2') \iff k_1 = k_1' \wedge k_2 = k_2'. \quad (20)$$

Note that R_2 is just the equality relation. The relation R_1 freely contains R_2. It contains R_2 because it is a coarser equivalence relation than R_2, the equality relation. This containment is even because each equivalence class of R_1 contains a countable number of equivalence classes of R_2. Checking the remaining condition of free containment is less immediate, but only slightly. Let D be an equivalence class of R and let $\{D_i \,:\, i \in \mathbb{N}\}$ be a partition of D that consists of equivalence classes of S. Then, D has the form $D = \{(k_0, k) \,:\, k \in \mathbb{N}\}$ for some fixed k_0, and each D_i is a singleton set of the form $\{(k_0, k_i)\}$ for some k_i. Given a permutation τ on $\mathbb{N}$, we may fulfill the condition of free containment using the following automorphism υ on $\mathbb{N}^2$:

$$\upsilon(k, k') = \begin{cases} (k, \tau(k')) & \text{if } k = k_0, \\ (k, k') & \text{if } k \neq k_0. \end{cases}$$

When $k = k_0$ and so (k, k') is in the equivalence class D, this function permutes the second component k' according to τ, thus meeting the first bullet point of the condition. Otherwise, (k, k') is not in D, and the function acts as the identity, as required by the second bullet point.

Definition A.6. Let R_1, R_2 be equivalence relations that are contained in an equivalence relation R. Then, R_1 and R_2 are said to be **orthogonal** within R if for any equivalence classes D_1, D_2, D of R_1, R_2, R, respectively, with $D_1, D_2 \subseteq D$, we have $D_1 \cap D_2 \neq \emptyset$.

[3] Here m may be the first countable ordinal, in which case π is a permutation on $\mathbb{N}$.

Definition A.7. The sequence $(R_k : k \leq n)$ of equivalence relations is **orderly** if for each $1 \leq k \leq n$, there exists an equivalence relation R'_k such that

- R'_k is explicit in $R_1, \ldots, R_{k-1}$;
- R'_k freely contains R_k; and
- if an equivalence relation S is explicit in $R_1, \ldots, R_{k-1}$ and strictly contained in R'_k but it is different from R_k, S is either orthogonal to R_k within R'_k or evenly contained in R_k with $\#_{R_k}(S) = \infty$.

A good example of an orderly sequence is (R_1, R_2) made out of relations R_i in (20). The required relations R'_1 and R'_2 are the complete relation $\mathbb{N}^2 \times \mathbb{N}^2$ and the relation R_1, respectively. We focus on R'_2. We have already shown that R_1 freely contains R_2. It is also explicit in R_1, simply because it is R_1. To check the third condition, consider an equivalence relation S explicit in R_1 and strictly contained in R'_2. Although we do not present a detailed calculation, it is possible to show that being explicit implies that S has to be one of the following relations:

$$=, \qquad R_1, \qquad \mathbb{N}^2 \times \mathbb{N}^2.$$

But only the equality relation is strictly contained in R'_2. Thus, S should be the equality relation. That is, $S = R_1$. Our argument so far shows that no S meets the assumptions in the third condition and so the condition holds vacuously.

The remaining concept is **anti-chain**. In a set A with a partial order $\preceq$, an **anti-chain** is a subset A_0 of A such that no two distinct elements of A_0 can be compared by $\preceq$, that is, for all $a, b \in A_0$, if $a \neq b$, then neither $a \preceq b$ nor $b \preceq a$. In Theorem A.2, A_0 is a set of certain subsets of I that are equivalence classes of some equivalence relations, and it is ordered by the subset relation.

A.2 Proof of Theorem A.1

Let $G = (V, E)$ be the DAG in Theorem A.1. Set n to the cardinality of V. The first step is to enumerate the vertices of G such that the order in the enumeration respects the directed edges in E. We use this enumeration to build a structure $\mathscr{M}$ that has $\mathbb{N}^G$ as its underlying set and satisfies the conditions of Theorem A.2, especially the orderly condition.

Lemma A.8. *There exists an enumeration of V, $(v_\ell : 1 \leq \ell \leq n)$, so that $V_\ell \stackrel{\text{def}}{=} \{v_1, \ldots, v_\ell\}$ is closed for every $\ell \leq n$.*

Proof. This is a well-known simple result. A process for enumerating V is called topological sort in combinatorics and computer science. For completeness, we explain the construction of the sequence $(v_\ell : 1 \leq \ell \leq n)$ in the lemma. We construct the sequence inductively. Since V is finite and G is acyclic, there exists a minimal vertex v_1. Our inductive construction starts with the sequence (v_1). Assume that we have enumerated ℓ elements such that $V_\ell \stackrel{\text{def}}{=} \{v_1, \ldots, v_\ell\}$ is closed. Now consider $V \setminus V_\ell$. Since V is finite and partially ordered, so is $V \setminus V_\ell$ and there exists a maximal element $v' \in V \setminus V_\ell$. We set $v_{\ell+1}$ to be this v'. Then, by the maximality of v' in $V \setminus V_\ell$, the set $\{v_1, \ldots, v_{\ell+1}\}$ is closed, as required. □

From now on, we write $\mathscr{M} = (I, R_{v_1}, \ldots, R_{v_n})$, where the v_k are enumerated as in Lemma A.8 and R_v is defined by

$$\alpha \, [R_v] \, \beta \iff \text{for all } w \preceq v, \alpha(w) = \beta(w).$$

Lemma A.9. *$\mathscr{M}$ is ultrahomogeneous.*

Proof. We use induction on n, the cardinality of the vertex set of G. For $n = 1$, the claim is equivalent to the existence of an extension of a bijection between finite subsets of $\mathbb{N}$ to a permutation of $\mathbb{N}$. So, it is obviously true. Now assume that the claim holds if $n \le m-1$. We will prove the claim for the case that $n = m$.

Let $\mathscr{N} = (J, S_1, \dots, S_m)$ be a substructure of $\mathscr{M}$, and τ an embedding from $\mathscr{N}$ to $\mathscr{M}$. Because of the way that we constructed the enumeration $(v_\ell : 1 \le \ell \le m)$, the last vertex v_m is maximal according to the partial order induced by G. That is, v_m is a terminal vertex. Let G' be the subgraph of G with the vertex set $W = \{v_1, \dots, v_{m-1}\}$. Let

$$I|_W \stackrel{\text{def}}{=} \mathbb{N}^W, \qquad R_{v_\ell}|_W \stackrel{\text{def}}{=} \{(\alpha|_W, \alpha'|_W) \,:\, (\alpha, \alpha') \in R_{v_\ell}\},$$
$$\mathscr{M}|_W \stackrel{\text{def}}{=} (I|_W, R_{v_1}|_W, \dots, R_{v_{m-1}}|_W), \qquad \mathscr{N}|_W \stackrel{\text{def}}{=} (J|_W, S_1|_W, \dots, S_{m-1}|_W),$$

Then, $\mathscr{N}|_W$ is a finite substructure of $\mathscr{M}|_W$. Furthermore, there exists a function $\tau_0 : J|_W \to I_W$ such that $\tau_0(\beta) = \tau(\beta')|_W$ whenever $\beta = \beta'|_W$. In fact, the function τ_0 is an embedding from $\mathscr{N}|_W$ to $\mathscr{M}|_W$. By induction hypothesis, τ_0 can be extended to an automorphism υ_0 on $\mathscr{M}|_W$.

We now extend υ_0 to an automorphism on $\mathscr{M}$. Fix $\beta \in J|_W$. Define

$$J_\beta \stackrel{\text{def}}{=} \{\beta' \in J \,:\, \beta'|_W = \beta\}.$$

Construct a permutation of $\mathbb{N}$, say π_β, so that $\pi_\beta(\beta'(v_m)) = \tau(\beta')(v_m)$ for all $\beta' \in J_\beta$. This is possible because J_β is finite. Define $\upsilon : I \to I$ as follows:

$$\upsilon(\alpha)(v_k) \stackrel{\text{def}}{=} \begin{cases} \upsilon_0(\alpha|_W)(v_k) & \text{if } k \le m-1, \\ \alpha(v_k) & \text{if } k = m \text{ and } \alpha \notin J, \\ \pi_{\alpha|_W}(\alpha(v_k)) & \text{if } k = m \text{ and } \alpha \in J. \end{cases}$$

Then, υ is the desired extension of τ. □

Lemma A.10. *The sequence $R_{v_1}, \dots, R_{v_n}$ is orderly.*

Proof. For each $k \le n$, define

$$R'_k \stackrel{\text{def}}{=} \bigcap \bigl\{R_{v_j} \,:\, j < n,\ v_j \preceq v_k \text{ and } v_j \ne v_k\bigr\}.$$

Clearly, R'_k is explicit in $R_{v_1}, \dots, R_{v_{k-1}}$, and R'_k freely contains R_{v_k}. Now consider S such that

1. S is an equivalence relation explicit in $R_{v_1}, \dots, R_{v_{k-1}}$;
2. S is strictly contained in R'_k; and
3. it is not the case that S is evenly contained in R_k with $\#_{R_k}(S) = \infty$.

A more careful analysis of the equivalence relations explicit in $R_{v_1}, \dots, R_{v_{k-1}}$ for this particular model reveals that they are exactly the equivalence relations that are of the form $\bigcap_{i \in I} R_{v_i}$ for $I \subseteq \{1 \dots k-1\}$. Firstly, the third clause in the notion of basic explicit is redundant on this occasion, for I_0 there must be either empty or I: these are the only two definable sets. Secondly, in this circumstance, if a Boolean combination of relations in $R_{v_1}, \dots, R_{v_{k-1}}$ is an equivalence relation then it must actually be an intersection of such relations; we showed this by considering the disjunctive normal forms that a transitive relation may have in this particular model.

From this we can conclude that S is an intersection of R'_k with some R_{v_j}'s where v_j is not an ancestor of v_k. Since $\{v_1, \dots, v_{k-1}\}$ is closed, v_k is not an ancestor of v_j either. Thus, v_j and v_k are incomparable. We use this to show that S is orthogonal to R_{v_k} in R'_k. To this end, consider equivalence classes D, D_1, D_2 of R'_k, S, R_{v_k}, respectively, with $D_1, D_2 \subseteq D$, Pick $\alpha_1 \in D_1$, $\alpha_2 \in D_2$, so that $\alpha_1 \; [R'_k] \; \alpha_2$, and let $\beta \in D$ be given by

$$\beta(v) = \begin{cases} \alpha_1(v) = \alpha_2(v) & \text{if } v \prec v_k, v \neq v_k \text{ ;} \\ \alpha_1(v) & \text{if } v = v_j \text{ where } R_{v_j} \subseteq S \text{ and } v \not\preceq v_k \\ \alpha_2(v) & \text{if } v = v_k \\ \text{anything} & \text{otherwise} \end{cases}$$

so that $\alpha_1 \; [S] \; \beta$ and $\beta \; [R_{v_k}] \; \alpha_2$, i.e. $\beta \in D_1 \cap D_2$. □

Proof of Theorem A.1. The previous lemmas imply that the conditions of Theorem A.2 hold. Thus, we can apply the theorem, and get the following representation of $\mathbf{X}$:

$$\mathbf{X} = \Big(X_\alpha : \alpha \in \mathbb{N}^G\Big) \overset{d}{=} \Big(f\big(U_b : b \in B(\alpha)\big) : \alpha \in \mathbb{N}^G\Big) \tag{21}$$

where $B(\alpha)$ is the set of all anti-chains in $E(\alpha) \overset{\text{def}}{=} \{[\alpha]_{R_k} \; : \; k \leq n\} \cup \{\{\alpha\}\}$, the collection of all equivalence classes $[\alpha]_{R_k}$ of i with respect to the R_k's, partially-ordered by set inclusion, and $(U_b : b \in \bigcup_{\alpha \in \mathbb{N}^G} B(\alpha))$ is a collection of independent $[0,1]$-uniform random variables.

The rest of the proof is about translating the representation in (21) to the claimed representation of Theorem A.1. A crucial part of this translation is the following function φ from $B \overset{\text{def}}{=} \bigcup\{B(\alpha) \; : \; \alpha \in \mathbb{N}^G\}$ to $J \overset{\text{def}}{=} \{\alpha|_C \; : \; C \in \mathscr{A}_G \text{ and } \alpha \in \mathbb{N}^G\}$:

$$\varphi(b) \overset{\text{def}}{=} \begin{cases} \alpha & \text{if } b \in B(\alpha) \text{ for some } \alpha \text{ and } b = \{\{\alpha\}\} \\ \alpha|_{\{w \,:\, w \preceq v_i \text{ for some } i\}} & \text{if } b \in B(\alpha) \text{ for some/any } \alpha \text{ and } b = \{[\alpha]_{R_{v_1}}, \ldots, [\alpha]_{R_{v_k}}\} \end{cases}$$

The function φ is well-defined. In the first case of the above definition, there is only one α. In the second case, there may be multiple choices of α, but they all give rise to the same element in J. Furthermore, φ satisfies three important properties. Firstly, it is surjective, because for any $C \in \mathscr{A}_G$ and $\alpha \in \mathbb{N}^G$, we have

$$\varphi(\{[\alpha]_{R_v} \; : \; v \text{ is } \preceq\text{-maximal in } C\}) = \alpha|_C.$$

Secondly, φ can be restricted to a surjective function from $B(\alpha)$ to $\{\alpha|_C \; : \; C \in \mathscr{A}_G\}$ for all $\alpha \in \mathbb{N}^G$. Finally, it is almost injective in the following sense: when M is the set of $\preceq$-maximal vertices of G,

$$\varphi(b) = \varphi(b') \implies \Big(b = b' \text{ or } \big\{b, b'\big\} = \big\{\{\{\alpha\}\}, \{[\alpha]_{R_v} \; : \; v \in M\}\big\} \text{ for some } \alpha \in \mathbb{N}^G\Big)$$

Let g be a measurable function from $[0,1]$ to $[0,1] \times [0,1]$ such that for any $[0,1]$-uniform U,

$$g(U) \overset{d}{=} (U_1, U_2)$$

for some independent $[0,1]$-uniform random variables U_1 and U_2. Pick a collection of independent $[0,1]$-uniform random variables

$$\mathbf{U}' \overset{\text{def}}{=} (U'_{\alpha|_C} : C \in \mathscr{A}_G \text{ and } \alpha \in \mathbb{N}^G).$$

Recall that M is the set of $\preceq$-maximal vertices. Let

$$B_0 = \{b \in B \; : \; b = \{\{\alpha\}\} \text{ or } b = \{[\alpha]_{R_v} : v \in M\} \text{ for some } \alpha \in \mathbb{N}^G\}.$$

Then,

$$\Big(\big(U_b \;:\; b \in B \setminus B_0\big),\, \big(U_b, U_{b'} \;:\; b = \{\{\alpha\}\} \text{ and } b' = \{[\alpha]_{R_v} : v \in M\} \text{ for some } \alpha \in \mathbb{N}^G\big)\Big)$$
$$\stackrel{d}{=} \Big(\big(U'_{\varphi(b)} \;:\; b \in B \setminus B_0\big),\, \big(g(U'_\alpha) \;:\; \alpha \in \mathbb{N}^G\big)\Big)$$

This and the second property of φ mentioned above imply the existence of a measurable function h such that

$$\Big(\big(U_b : b \in B(\alpha)\big) : \alpha \in \mathbb{N}^G\Big) \stackrel{d}{=} \Big(h\big(U'_{\alpha|_C} : C \in \mathscr{A}_G\big) : \alpha \in \mathbb{N}^G\Big),$$

which implies

$$\big(X_\alpha : \alpha \in \mathbb{N}^G\big) \stackrel{d}{=} \Big((f \circ h)\big(U'_{\alpha|_C} : C \in \mathscr{A}_G\big) : \alpha \in \mathbb{N}^G\Big),$$

as desired. □

References

[AAF$^+$] Nathanael Leedom Ackerman, Jeremy Avigad, Cameron E. Freer, Daniel M. Roy, and Jason M. Rute. On the computability of graphons. arxiv:1802.09598.

[Ack15] Nathanael Ackerman. Representations of Aut(M)-invariant measures: Part I. arXiv:1509.06170, 2015.

[AFP16] Nathanael Ackerman, Cameron Freer, and Rehana Patel. Invariant measures concentrated on countable structures. *Forum of Mathematics, Sigma*, 4, 2016.

[Ald85] David J Aldous. Exchangeability and related topics. In *École d'Été de Probabilités de Saint-Flour XIII 1983*, pages 1–198. Springer, 1985.

[Ald09] David J Aldous. More uses of exchangeability: representations of complex random structures. *arXiv preprint arXiv:0909.4339*, 2009.

[AP14] Tim Austin and Dmitry Panchenko. A hierarchical version of the de Finetti and Aldous-Hoover representations. *Probability Theory and Related Fields*, 159(3-4):809–823, 2014.

[Aus08] Tim Austin. On exchangeable random variables and the statistics of large graphs and hypergraphs. *Probability Surveys*, 5:80–145, 2008.

[Aus12] Tim Austin. Exchangeable random arrays. In *Notes for IAS workshop*, 2012.

[CT17] Harry Crane and Henry Towsner. Relative exchangeability with equivalence relations. *Archive for Mathematical Logic*, pages 1–24, 2017.

[CT18] Harry Crane and Henry Towsner. Relatively exchangeable structures. *The Journal of Symbolic Logic*, 83(2):416–442, 2018.

[FR12] Cameron E. Freer and Daniel M. Roy. Computable de Finetti measures. *Ann. Pure Appl. Logic*, 163(5):530–546, 2012.

[GMR$^+$08] Noah Goodman, Vikash Mansinghka, Daniel Roy, Keith Bonawitz, and Joshua Tenenbaum. Church: a language for generative models. In *Proc. UAI 2008*, 2008.

[Hoo79] Douglas N Hoover. Relations on probability spaces and arrays of random variables. *Preprint, Institute for Advanced Study, Princeton, NJ*, 2, 1979.

[Kal05] Olav Kallenberg. *Probabilistic Symmetries and Invariance Principles*. Springer, 2005.

[SSY+18] Sam Staton, Dario Stein, Hongseok Yang, Nathanael L. Ackerman, Cameron Freer, and Daniel M Roy. The beta-bernoulli process and algebraic effects. In *In Proceedings of 45th International Colloquium on Automata, Languages and Programming*, 2018.

[SYA+17] Sam Staton, Hongseok Yang, Nathanael L. Ackerman, Cameron Freer, and Daniel M Roy. Exchangeable random process and data abstraction. In *Workshop on Probabilistic Programming Semantics (PPS 2017)*, 2017.

[TvdMYW16] David Tolpin, Jan-Willem van de Meent, Hongseok Yang, and Frank D. Wood. Design and implementation of probabilistic programming language anglican. In *Proceedings of the 28th Symposium on the Implementation and Application of Functional Programming Languages, IFL 2016, Leuven, Belgium, August 31 - September 2, 2016*, pages 6:1–6:12, 2016.

[WvdMM14] Frank Wood, Jan Willem van de Meent, and Vikash Mansinghka. A new approach to probabilistic programming inference. In *Proceedings of the 17th International conference on Artificial Intelligence and Statistics*, pages 1024–1032, 2014.