

ANNALES
HENRI LEBESGUE

PAUL JUNG

JIHO LEE

SAM STATON

HONGSEOK YANG

A GENERALIZATION OF
HIERARCHICAL
EXCHANGEABILITY ON TREES
TO DIRECTED ACYCLIC GRAPHS

UNE GÉNÉRALISATION DE
L'ÉCHANGEABILITÉ HIÉRARCHIQUE SUR
LES ARBRES AUX GRAPHS DIRIGÉS
ACYCLIQUES

Keywords: Bayesian nonparametrics, exchangeability, hierarchical exchangeability, Aldous–Hoover representation, de Finetti representation.

2020 *Mathematics Subject Classification:* 60G09.

DOI: <https://doi.org/10.5802/ahl.74>

(*) Staton was supported by a Royal Society University Research Fellowship. Yang was supported by the Engineering Research Center Program through the National Research Foundation of Korea (NRF) funded by the Korean Government MSIT (NRF-2018R1A5A1059921), and also by Next-Generation Information Computing Development Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Science, ICT (2017M3C4A7068177). Jung and Lee were supported by the National Research Foundation of Korea (NRF-2017R1A2B2001952) and also through a National Research Foundation of Korea grant funded by the Korean Government MSIT (NRF-2019R1A5A1028324).

ABSTRACT. — Motivated by the problem of designing inference-friendly Bayesian non-parametric models in probabilistic programming languages, we introduce a general class of partially exchangeable random arrays which generalizes the notion of hierarchical exchangeability introduced in Austin and Panchenko (2014). We say that our partially exchangeable arrays are DAG-exchangeable since their partially exchangeable structure is governed by a collection of Directed Acyclic Graphs. More specifically, such a random array is indexed by $\mathbb{N}^{|V|}$ for some DAG $G = (V, E)$, and its exchangeability structure is governed by the edge set E . We prove a representation theorem for such arrays which generalizes the Aldous-Hoover and Austin-Panchenko representation theorems.

RÉSUMÉ. — Motivés par le problème de la conception de modèles bayésiens non paramétriques à inférence pour les langages de programmation probabilistes, nous introduisons une classe générale de tableaux aléatoires partiellement échangeables, qui généralise la notion d'échangeabilité hiérarchique introduite par Austin et Panchenko (2014). Nous disons que nos tableaux partiellement échangeables sont DAG-échangeables puisque leur structure partiellement échangeable est gouvernée par une collection de graphes dirigés acycliques (DAG). Plus spécifiquement, un tel tableau aléatoire est indexé par $\mathbb{N}^{|V|}$ pour un certain DAG $G = (V, E)$, et sa structure d'échangeabilité est gouvernée par l'ensemble d'arêtes E . Nous démontrons un théorème de représentation pour de tels tableaux, qui généralise les théorèmes de représentation d'Aldous-Hoover et Austin-Panchenko.

1. Introduction

In [AP14], Austin and Panchenko consider a random array indexed by ℓ -tuples of paths over a collection of ℓ infinitely-branching rooted trees, of finite depths $\{r_1, \dots, r_\ell\}$, where each path in the ℓ -tuple starts from the root of one tree in the collection and ends at a leaf of that tree.⁽¹⁾ If the branches emanating from a given vertex are labelled by \mathbb{N} , then the index set of the random array is $\mathbb{N}^{r_1} \times \dots \times \mathbb{N}^{r_\ell}$. This random array is *hierarchically exchangeable*, defined in [AP14], if its joint distribution remains invariant under rearrangements that preserve the structure of each rooted tree in the collection underlying the index set. In other words, if its joint distribution remains invariant under any map $\tau = (\tau_1, \dots, \tau_\ell)$ on the index set, where τ_i is a rooted-graph isomorphism of the i^{th} rooted tree.

In their work, Austin and Panchenko prove that such arrays have a representation in the spirit of the celebrated Aldous-Hoover representation for exchangeable arrays of random variables [Ald81, Ald85, Hoo79]. In the special case where all trees in the collection have a depth of one, $r_1 = \dots = r_\ell = 1$, i.e., are copies of \mathbb{N} rooted at \emptyset (see Figure 1.1), 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, Chapter 7] (see also [Ald85, Aus12]) for the definition of separate exchangeability, a statement of the Aldous-Hoover theorem, and additional background on what are now classic results in the theory of exchangeable random arrays.

In this work, we generalize hierarchical exchangeability in [AP14] from trees to directed acyclic graphs (DAGs). Our generalization is motivated by issues related to

⁽¹⁾They studied such a random array to address an issue arising in spin glasses, which is to prove the predictions of the Mézard and Parisi ansatz for diluted spin glass models [MP01]. See [Pan15] for details.



Figure 1.1. A collection of rooted trees of depth 1

exchangeable random processes from Bayesian nonparametrics and their implementations in probabilistic programming languages, which we will describe in a separate subsection of this introduction.

Let us describe our generalization of hierarchical exchangeability. Recall 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). For a given DAG, we use the partial order on its vertices: $v \preceq w$ whenever there is a path (possibly of length zero) from v to w . Conversely, we can regard a finite partially ordered set as a DAG having a directed edge (v, w) if $v \prec w$ and there is no v' such that $v \prec v' \prec w$.

In the sequel, we slightly abuse notation by writing G when we refer to the vertex set of G . Conversely, when we refer to a subset or some other vertex set, we will often assume an underlying edge structure induced by the partial order. A subset W of a DAG G is *downward-closed* (with respect to the partial ordering) if $w' \in W$ and $w \preceq w'$ implies that $w \in W$. We will see that, in the context of hierarchical exchangeability, being “higher” in the hierarchy than v corresponds to being smaller than v in the partial order \preceq . We will often say “closed” instead of “downward-closed” for simplicity. For a closed subset C of G , we denote the collection of closed subsets of C by

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

DEFINITION 1.1. — Let $C \in \mathcal{A}_G$. 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 $D \in \mathcal{A}_C$, then every C -type multi-index $\alpha \in \mathbb{N}^C$ can be restricted to a D -type multi-index $\alpha|_D \in \mathbb{N}^D$. A C -type random array in a Borel space \mathcal{X} is a family of random variables,

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

indexed by C -type multi-indices. Here each X_α is a random variable taking values in \mathcal{X} .

Remark 1.2. — We henceforth do not mention the Borel space \mathcal{X} , and just say C -type random array.

DEFINITION 1.3. — Let G be a finite DAG. A G -automorphism of \mathbb{N}^G is a bijection $\tau : \mathbb{N}^G \rightarrow \mathbb{N}^G$ such that

$$(1.1) \quad \alpha|_C = \beta|_C \iff (\tau(\alpha))|_C = (\tau(\beta))|_C, \quad \text{for all } C \in \mathcal{A}_G \text{ and all } \alpha, \beta \in \mathbb{N}^G.$$

For $C \in \mathcal{A}_G$, we say that a C -type random array \mathbf{X}_C is DAG-exchangeable if for every G -automorphism τ ,

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

where $\tau(\mathbf{X}_C)$ is the random array $(\mathbf{X}_{\tau(\alpha)} : \alpha \in \mathbb{N}^C)$.

Remark 1.4. — The index set of a C -type random array, \mathbb{N}^C , also has a natural DAG structure which is infinitely-branching for every nonterminal vertex. Since this structure is somewhat complicated, we will delay its explanation until Example 2.4.

DAG-exchangeability is an instance of *partial exchangeability* which means that it does not require invariance with respect to all bijections of the index set \mathbb{N}^G , but rather, only with respect to some subgroup of these bijections [Ald85]. In the case of DAG-exchangeability, this subgroup is precisely the set of G -automorphisms. In Examples 2.1 and 2.3, we present examples of bijections which are or are-not G -automorphisms.

Our main result is a representation theorem for DAG-exchangeable arrays. The Aldous–Hoover theorem represents an exchangeable array $\mathbf{X} = (X_{ij})$ with a measurable function of collections of independent uniform random variables corresponding to the entries in the array and the symmetries of the array, as well as an additional uniform random variable corresponding to the symmetries of the symmetries. Our representation is similar in that it represents arrays with a measurable function of collections of independent uniform random variables. Again, one collection corresponds to the entries in the array, and the other collections correspond to the symmetries of the array induced by DAG-exchangeability, as well as the symmetries of the symmetries.

THEOREM 1.5. — *If a G -type random array \mathbf{X} is DAG-exchangeable, then there exists a measurable function $g : [0, 1]^{\mathcal{A}_G} \rightarrow \mathcal{X}$ such that*

$$(1.2) \quad \mathbf{X} = (X_\alpha : \alpha \in \mathbb{N}^G) \stackrel{d}{=} \left(g(U_{\alpha|C} : C \in \mathcal{A}_G) : \alpha \in \mathbb{N}^G \right)$$

where the U_β are independent $[0, 1]$ -uniform random variables.

In fact, in this work we prove a slightly more general result than the above, namely that “consistent” representations can be found for each C -type array, simultaneously for all $C \in \mathcal{A}_G$.

DEFINITION 1.6. — *Let \mathcal{C} be a sequence consisting of distinct closed sets of G 's vertices. A \mathcal{C} -type random array collection in \mathcal{X} is a sequence of random variable families*

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

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

In particular, for every DAG-exchangeable random array with indices in \mathbb{N}^G and each downward-closed subset C of G , there is a canonical way of generating a random array with the index set \mathbb{N}^C . Our main representation theorem provides representations for all such induced random arrays, simultaneously. We will refer to such a family of consistent representations, for all induced C -type arrays, as being *fine-grained*.

There are many reasons for considering such generalized random arrays as we do here. Partial exchangeability was considered by de Finetti himself. For instance in [Fin75, Chapter 12], he discusses its role in both parametric and nonparametric

Bayesian statistics. In the 1980s the subject flourished⁽²⁾, and d -dimensional random arrays (i.e., matrices and tensors) emerged as fundamental structures underlying the theory of partial exchangeability. Indeed, in the foundational work of [Hoo79], separately exchangeable arrays together with their joint and weak exchangeable counterparts, are seen to arise quite naturally as mathematical objects. However, even there, the question of when representations arise for other partially exchangeable random arrays is posed in [Hoo79, Section 7]. As already mentioned, [AP14] is one work in this direction using probabilistic arguments in the spirit of [Ald81] and [Kal05]. Also, using ultraproducts and other model theoretic tools in the spirit of [Hoo79], the work of [CT17, CT18] introduce the quite general notion of *relative* exchangeability, from which many forms of partial exchangeability can be extracted.

The main contribution of this work is the extension, in terms of a probabilistic proof, of the concrete framework of hierarchical exchangeability. As noted in the works of Crane and Towsner, hierarchical exchangeability can already be extracted from their more abstract framework; we will show in the appendix that DAG-exchangeability for single arrays also falls under the umbrella of their abstract framework and thus can also be (nontrivially) realized in that framework (we have yet to see whether DAG-exchangeability for families of arrays, which is addressed in our main result, fits into the picture of relative exchangeability, albeit there are some indications that it should). On the level of applications of DAG-exchangeability, we provide a summary of our motivations with respect to Bayesian nonparametric models and probabilistic programming in the next subsection of this introduction.

Let us mention that recently, partial exchangeability was found to have ramifications in the study of random graphs, their limits, and their statistical properties [Aus08, CCB16, CF17, CD18, DJ08, VR15]. We remark that the exchangeability of graphs is not the topic of this paper—instead we use directed graphs as a tool to create and describe our probabilistic symmetries. It is however, not unreasonable to envision that our results can be applied to this line of research in the future.

Following the applications presented in the next subsection, the rest of the paper is organized as follows. In Section 2, we present some examples which motivate our notion of DAG-exchangeable arrays, and illustrate the probabilistic symmetries induced by G -automorphisms. Our examples include, in particular, how hierarchical exchangeability fits into the framework of DAG-exchangeability. In Section 3 we start by extending the notion of DAG-exchangeability to collections of C -type random arrays, and then present our main result. The proof of this result comprises Section 4. 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⁾Two somewhat recent surveys of the early results in this field are given in [Aus12] and [Ald10].

1.1. Applications to Probabilistic Programming

In terms of applications, our motivation comes from studying generative models of array-like structures through probabilistic programming languages⁽³⁾. These are high-level languages for statistical modeling that come equipped with separate Bayesian inference engines, which implement statistical inference algorithms such as the Metropolis–Hastings algorithm and Gibbs sampling.

In that context, one application of exchangeability and Aldous–Hoover type theorems is to identify when an elaborate, hierarchical generative model can be replaced by an equivalent one, with better independence properties, that is more amenable to inference engines. To briefly summarize this, we provide a concrete illustration in the case of a 2-dimensional exchangeable random array. A statistical programmer can implement this as an abstract data type with the following functions:

```
(1.3) newArray: () -> Array
      newRow: Array -> Row
      newColumn: Array -> Column
      entry: (Array, Row, Column) -> real
```

One can build a finite part of the array by writing a program such as:

```
a = newArray();
r1 = newRow(a); ... ; rm = newRow(a);
c1 = newColumn(a); ... ; cn = newColumn(a);
result[1,1] = entry(a,r1,c1); ... ; result[1,n] = entry(a,r1,cn);
result[2,1] = entry(a,r2,c1); ... ; result[2,n] = entry(a,r2,cn);
...
result[m,1] = entry(a,rm,c1); ... ; result[m,n] = entry(a,rm,cn)
```

so that the program will randomly return an $m \times n$ array which is a projection of the ultimate infinite random array.

To be more precise, we suppose that the rows correspond to people, and the columns to movies, and that there is a 1 in a particular cell of an array if that person liked that movie, and a 0 otherwise. In this scenario, the “infinite relational model” [KTG⁺06, OR15, XTYK06] provides a reasonable array. The model implicitly clusters the people into sorts, and clusters movies into genres, and there is an implicit probability that each sort of person will like a movie in each genre. This can be described in a generative way, where we build the array up, as follows [OR15, Example IV.1].

- An **Array** will be represented by a mutable data structure in memory that records how many people are currently grouped into each sort, and how many movies are currently in each genre, as well as how many people of each sort liked how many movies of each genre.
- **newRow** works according to the “Chinese Restaurant Process” [GvdV17] as follows. Suppose that there are currently m sorts of people, and c_i people of

⁽³⁾At this stage, Church [GMR⁺08] and Anglican [MMYW16, WMM14] 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.

each sort. Then the new person will be of sort $i \leq m$ with probability

$$\frac{c_i}{1 + \sum_{j=1}^m c_j}$$

and a new sort with probability

$$\frac{1}{1 + \sum_j c_j}.$$

- When we add a new person (row), we need to also update the new entries in the array – does the new person like the existing movies (columns)? We do this using a Polya urn scheme for each genre j . Suppose that the new person is of sort i . We now run through the movies in genre j , and like each one in turn with probability

$$\frac{a + 1}{a + b + 2}$$

where a is the number of “likes” between people of sort i and movies in genre j , including the movies considered for the new person so far, and b is the number of “dislikes”.

- Similarly, `newColumn` works according to the Chinese Restaurant Process, picking a genre for a new movie, and deciding whether the existing people like the new movie.

This implementation is a natural generative program, but the conditional independence appears to be complicated because each step depends on the last. Because of the exchangeability properties of this infinite relational model, the Aldous–Hoover theorem provides an equivalent implementation that is less generative but with much clearer independence properties. In this particular example we can use a stick-breaking implementation of the Chinese Restaurant Process, and a Beta distribution to implement the Polya urn [OR15, Example IV.6].

- An `Array` is represented by two random countable partitions of $[0, 1]$, say $r, c : [0, 1] \rightarrow \mathbb{N}$, chosen by stick breaking according to the Dirichlet process, together with a random function $e : \mathbb{N}^2 \rightarrow [0, 1]$, i.i.d. uniformly distributed, giving the probability of liking for any given sort of person and genre. The functions r, c , and e are determined by the empirical distribution of the array in question.
- Each person and movie is assigned a uniformly distributed number in $[0, 1]$.
- The probability that a person p likes a movie m is $e(r(p), c(m))$. One way to set this up is to say that each person-movie pair (p, m) is assigned a uniform random $s_{m,p} \in [0, 1]$ i.i.d., and the person p likes the movie m if $s_{m,p} < e(r(p), c(m))$.

The Aldous–Hoover theorem guarantees that the above procedure provides an accurate sample from the array in question. Putting this together, we have a random function $F : [0, 1]^3 \rightarrow \{0, 1\}$ given by $F(p, m, s_{m,p}) = [s_{m,p} < e(r(p), c(m))]$. Picking a parameterization for the random function F gives an ordinary function $f : [0, 1]^4 \rightarrow \{0, 1\}$, which is an Aldous–Hoover representation.

An inference engine can take advantage of the clear independence between people and movies in this implementation.

From the interface (1.3) we can directly read off the DAG for random arrays given in Example 2.1 (b). Furthermore, the interface suggests generalizations that are more complicated and hierarchical, but which are natural to consider in a generative model. For instance, we may extend the interface in (1.3) such that each array cell contains not just a real number but also an array:

```
newArray: () -> Array
newRow: Array -> Row
newColumn: Array -> Column
entry: (Array,Row,Column) -> real

newNestedArray: (Array,Row,Column) -> NestedArray
newNestedRow: NestedArray -> NestedRow
newNestedColumn: NestedArray -> NestedColumn
nestedEntry: (NestedArray,NestedRow,NestedColumn) -> real
```

cdrt Then, we may change the first implementation of the infinite relational model such that a person is represented by a nested row while a row denotes a sort of people; similarly, a movie is represented by a nested column while a column denotes a genre of movies. From this interface we can directly read off the DAG for random block matrices given in Example 2.2(b).

If such a complex generative model is suitably exchangeable (see [SYA⁺17], [SSY⁺18]), then our theorem says that it could just have well been implemented by sampling from uniform distributions, and thus the conditional independence relationships can be properly understood. For instance, for the nested version of the infinite relational model from above, our theorem provides its equivalent variant with better independence properties that provide the definitions of both `entry` and `nestedEntry` operations. Note that the `entry` operation depends only on the row and column indices of the outer matrix, while the `nestedEntry` operation uses all of the row and column indices of the outer and nested matrices. Only a fine-grained representation theorem is able to handle operations depending on different parts of the indices.

Remark 1.7. — In the 2-dimensional situation there are many further things to be said about programming and inference with Aldous–Hoover representations and generative models, for instance:

- in terms of computability, the Aldous–Hoover representation of a generative model need not be computable in general [AAF⁺18, FR12];
- in terms of estimation, by placing some smoothness assumptions on the functions involved, there are optimal algorithms for estimating the Aldous–Hoover representation (e.g. [CGZ15]). In particular, methods such as those in [CGZ15] exemplify the usefulness of having functional representations of the type pursued in this work.

It would be interesting and useful to extend these developments to more complex DAG-exchangeable settings, but we will not explore this further here.

2. Examples

In this section we consider examples of DAG-exchangeable random arrays (Definition 1.1) and a fine-grained generalization (Definition 1.6).

Our first example shows that DAG-exchangeability generalizes several popular notions of exchangeability from the literature, when we choose G appropriately.

Example 2.1. — 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.

- (a) *de Finetti Sequences:* The most common discretely-indexed 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. Every permutation on \mathbb{N} is a G -automorphism. In this case, DAG-exchangeability becomes the standard notion of exchangeability for random sequences in de Finetti’s classic result.
- (b) *Aldous–Hoover Arrays:* 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 multi-index set \mathbb{N}^G is $\mathbb{N}^{\{r,c\}} \simeq \mathbb{N}^2$. Thus, a G -type multi-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. A G -automorphism τ corresponds to a pair of permutations π, π' on \mathbb{N} , with π acting on the row index and π' acting on the column index. In this case, DAG-exchangeability becomes Aldous–Hoover (separate) exchangeability.
- (c) *Hierarchical Exchangeability:* Let r_1, \dots, r_ℓ 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, \dots, r_ℓ , respectively [AP14]. Formally, this process is a family of random variables of the following form:

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

This stochastic process is a G -type array for the DAG G in *Figure 2.1*. So

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

and

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

Thus, $(X_\alpha : \alpha \in \mathbb{N}^{r_1} \times \dots \times \mathbb{N}^{r_\ell})$ is the same thing as a G -type array. In this case, DAG-exchangeability is the same thing as hierarchical exchangeability.

Of course, our framework 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.

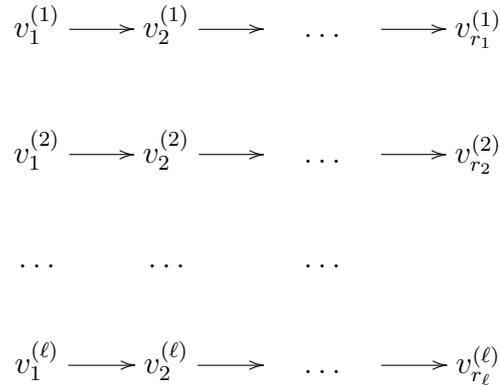


Figure 2.1. The DAG for multi-path-indexed random arrays in [AP14].

Example 2.2. — In order to illustrate the generality of our setting, we present some other instances of DAG-exchangeable arrays that are not mainstream in the exchangeability literature. We mention again that the framework of [CT17] and [CD18] can be used to derive DAG-exchangeability as in our appendix. Thus these examples, except for the last part of (b), fit into their abstract framework as well.

(a) *Sequences of Random Matrices:* The multi-index set for the DAG G in Fig-

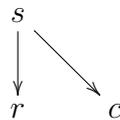


Figure 2.2. The DAG for sequences of random matrices.

ure 2.2 is $\mathbb{N}^G = \mathbb{N}^{\{s,r,c\}}$ which can be thought of as an infinite sequence (arbitrarily labeled) of matrices of the type found in Example 2.1 (b). For $\alpha \in \mathbb{N}^G$, the number $\alpha(s)$ determines which matrix to look at, while $\alpha(r)$ and $\alpha(c)$ are the row and column numbers of the matrix. When these matrices are DAG-exchangeable, the sequence is an exchangeable sequence, and each matrix is a separately exchangeable Aldous-Hoover array. Note that two entries, say $X_{r_0,c_0}^{(s_1)}$ and $X_{r_0,c_0}^{(s_2)}$, of two different arrays in this sequence, which are in the same position (r_0, c_0) , are only related through the exchangeability of the sequence (s_1, s_2, \dots) , and not through their position (r_0, c_0) . Thus this structure has a different partial exchangeability than a three-dimensional separately exchangeable Aldous-Hoover array. For instance, DAG-exchangeability permits the use of different permutations for the rows and columns of $X^{(s_1)}$ and those of $X^{(s_2)}$, while the three dimensional separate exchangeability forbids it.

Sequences of operators naturally arise in mathematical physics— see for instance [BR97, Chapter 6.2]. A specific example is an exchangeable sequence of non-Hermitian random matrices with separately exchangeable entries [BCC11, BC12].

(b) *Random Block Matrices:*

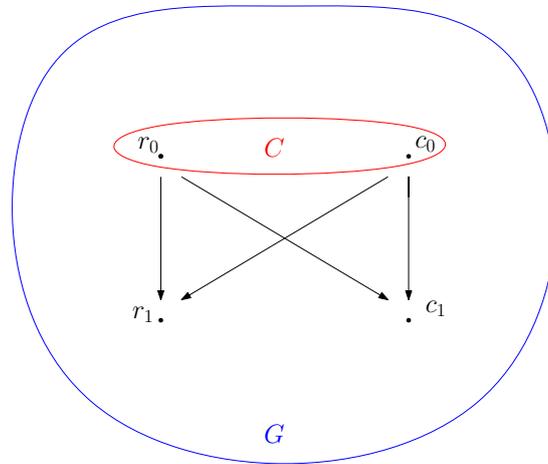


Figure 2.3. The DAG for random block matrices; $\mathcal{C} = (C, G)$.

Consider the DAG G in Figure 2.3. The 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 $(\alpha(r_1), \alpha(c_1))^{\text{th}}$ entry of the nested matrix, which is itself stored at the $(\alpha(r_0), \alpha(c_0))^{\text{th}}$ entry of the outer nesting matrix. We want to point out that if in Example 2.1 (c), one takes $\ell = 2$ and $r_1 = r_2 = 2$, then the multi-index set is the same as in this example. Thus, as a stochastic process, this example is just a special case of Example 2.1 (c). However, as an *exchangeable* stochastic process, this relationship no longer holds. The presence of the additional directed edges here means that a random G -type array should satisfy more symmetries than Example 2.1 (c), particularly those symmetries that are expected to hold for exchangeable random block matrices.

Now, let $C = \{r_0, c_0\}$, a closed subset of G . A small generalization of the random block matrix, similar to our example in Section 1.1, is a random structure that is simultaneously a random matrix (Example 2.1 (b)) and a random block matrix. This can be thought of as a random matrix where each cell contains both a value in \mathcal{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.

(c) *Random Block Matrices and Sequences:* One can use all the previous examples to build new examples. For instance, in Figure 2.4, all three DAGs give the multi-index set $\mathbb{N}^{\{s, r_0, c_0, r_1, c_1\}}$. The edge sets of the three DAGs, however, lead to three different random structures under DAG-exchangeability. The left side can be thought of as a sequence of random block matrices. But if one removes the edge from $s \rightarrow c_0$ (see Figure 2.5), then the DAG-exchangeable array is better thought of as one single random block matrix with rows subject to

two-level hierarchical exchangeability. The right figure can be thought of as a block matrix such that in each entry of each inner matrix, one finds a sequence of random variables (rather than a single random variable), thus it is a block matrix of sequences. If one removes the edge from $c_1 \rightarrow s$ in the middle figure, then one can still view it as a block matrix of sequences, but the distribution of the sequence no longer depends on which “inner” column it is associated with. Finally, the figure on the right has many natural interpretations in terms of DAG-exchangeability. We invite the reader to ponder upon the interesting different interpretations for the associated DAG-exchangeable arrays in this case.

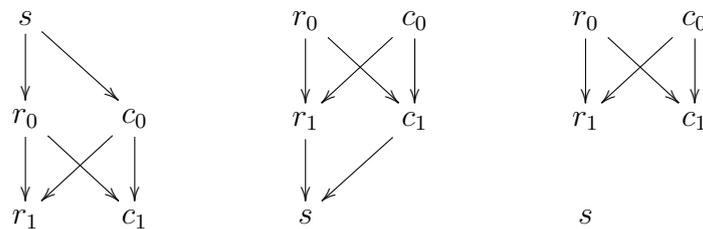


Figure 2.4. The DAGs for different extensions of random block matrices.

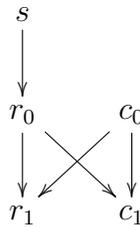


Figure 2.5. The DAG for random block matrices subject to two-level hierarchical exchangeability.

- (d) *Random Walls:* Here is another example of a random array collection with $\mathcal{C} \neq (G)$. Consider the graph G consisting of three vertices x, y, z and no edges. Define \mathcal{C} as follows:

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

A \mathcal{C} -type random array collection consists of three random variable families, namely, $\mathbf{X}_{C_{xy}}$, $\mathbf{X}_{C_{yz}}$ and $\mathbf{X}_{C_{zx}}$ (see Figure 2.6). 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 0 (see Figure 2.7). Viewing the tuple $\mathbb{X} = (\mathbf{X}_{C_{xy}}, \mathbf{X}_{C_{yz}}, \mathbf{X}_{C_{zx}})$ 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.

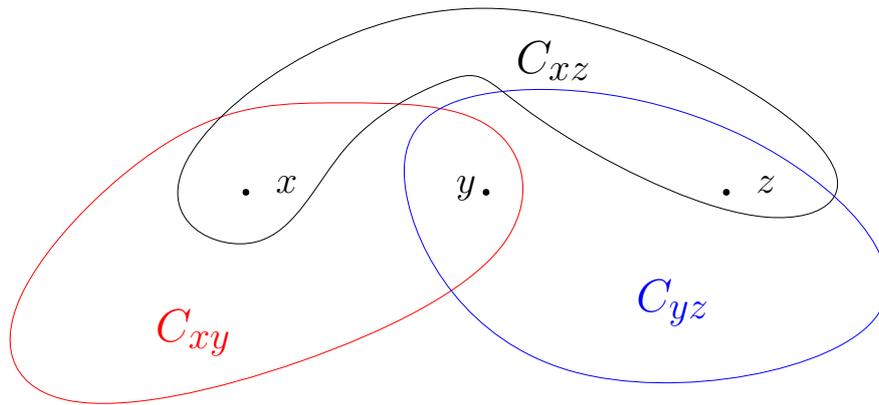


Figure 2.6. The DAG for random walls; $\mathcal{C} = \{C_{xy}, C_{xz}, C_{yz}\}$.

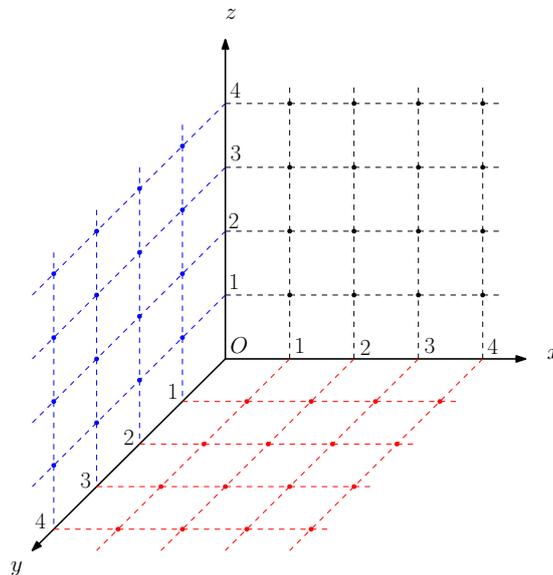


Figure 2.7. The multi-index set for random walls.

Axes are not part of the walls, i.e., the walls have no intersections.

Example 2.3. — This example further illustrates the notion of G -automorphism.

- (a) *Nested Sequences:* When G is just a single edge $(v_1 \rightarrow v_2)$, the random array with multi-indices in \mathbb{N}^G represents a random sequence whose elements are again sequences. A bijection τ on the multi-index set \mathbb{N}^G is a G -automorphism if and only if it is of the form

$$\tau[(v_1, v_2) \mapsto (i, j)] = [(v_1, v_2) \mapsto (\pi(i), \pi'_i(j))]$$

for some permutations π, π'_i on \mathbb{N} . Here $(v_1, v_2) \mapsto (i, j)$ represents a multi-index in \mathbb{N}^G mapping v_1 and v_2 to i and j , respectively. Note the dependency of π'_i on the value i of v_1 . This dependence allows τ to use different permutations for v_2 according to different values of i . However, when the edge $\vec{v_1 v_2}$ is removed from G , this dependency is no longer allowed: in this case, for a

bijection τ to be a G -automorphism, it should have the form

$$\tau[(v_1, v_2) \mapsto (i, j)] = [(v_1, v_2) \mapsto (\pi(i), \pi'(j))]$$

for some permutations π, π' on \mathbb{N} precisely as in Example 2.1 (b). Note the use of a single bijection π' for all values i of v_1 , when the edge $\overrightarrow{v_1 v_2}$ is removed.

- (b) *Sequences of Random Matrices (revisited)*: When G is the DAG in Figure 2.2, the following bijection τ on \mathbb{N}^G is a G -automorphism:

$$\tau([(s, r, c) \mapsto (i, j, k)]) = [(s, r, c) \mapsto (i, \pi_i(j), k)].$$

Here $[(s, r, c) \mapsto (i, j, k)]$ is the multi-index in \mathbb{N}^G mapping vertices (s, r, c) to natural numbers (i, j, k) , and π_n is the permutation on \mathbb{N} that cycles the first n numbers (i.e., $\pi_n(1) = 2, \pi_n(2) = 3, \dots, \pi_n(n) = 1$, and $\pi_n(m) = m$ for $m > n$). As we previously explained, a random array with indices in \mathbb{N}^G is a sequence of random matrices. The G -automorphism τ , in this example, permutes the rows of these matrices, but the way it does so depends on the position of a matrix in the sequence. If we remove the edge \overrightarrow{sr} from G , this dependence is no longer permitted, so that τ stops being a G -automorphism. On the other hand, removing the other edge \overrightarrow{sc} from G is harmless; τ continues to meet the conditions of being a G -automorphism.

- (c) *Random Block Matrices and Sequences (revisited)*: Consider the middle multi-index set in Figure 2.4. The allowable G -automorphisms for this multi-index set are combinations of (a) permuting rows of blocks, (b) permuting columns of blocks, (c) permuting rows within a given block, (d) permutating columns within a given block, and (e) permuting a sequence stored in the entry of a nested matrix.
- (d) *G-automorphism*: More generally, for any finite DAG G , a G -automorphism τ always has the form of applying a permutation to the number associated with each vertex by a given multi-index. The choice of permutation for each vertex is allowed to vary, but only in a way consistent with the structure of G . When $\overrightarrow{w_1 v}, \dots, \overrightarrow{w_p v}$ are all the incoming edges to a vertex v in G (i.e., w_1, \dots, w_p is the set of parents of v), the permutation for v should have the form $\pi_{(n_1, \dots, n_p)}$, where the subscripts are the numbers assigned to w_1, \dots, w_p by a given multi-index α , i.e., $\tau(\alpha)(v) = \pi_{(\alpha(w_1), \dots, \alpha(w_p))}(\alpha(v))$.

As remarked in the introduction, our final Example 2.4 illustrates that the multi-index sets of DAG-type random arrays also have a natural infinitely-branching DAG structure.

Example 2.4. — Ignoring the edges in Figure 2.7 which are merely a visual aid, the vertex set is a natural infinite multi-index set corresponding to the DAG-exchangeable array of Figure 2.6. When DAGs have directed edges, the principle still holds: there is a natural infinitely-branching DAG $G' = (V', E')$ underlying the multi-index set of any DAG-exchangeable array for some finite DAG $G = (V, E)$.

The infinitely-branching DAG G' has a vertex set which replaces each vertex v in G by a countably infinite number of vertices (considered copies of the original v). Its edge set is chosen such that if there is an edge from a copy of w to a copy of v in G' ,

then there is an edge \overrightarrow{wv} in G . The precise definition of G' requires a few notations. For each vertex $v \in V$, let

$$V_v = \{w \in V : \text{there is a path (possibly 0-length) from } w \text{ to } v \text{ in } G\}, \quad \mathbb{I}_v = \mathbb{N}^{V_v}.$$

Here, \mathbb{I}_v represents the set of copies of v , which together replace the vertex v . Thus, $V' = \bigcup_{v \in V} \mathbb{I}_v$. Note that every edge \overrightarrow{wv} in E induces a map from vertices in \mathbb{I}_v to those in \mathbb{I}_w in the graph G' . The map transforms a vertex $\alpha \in \mathbb{I}_v$ to $\alpha|_{V_w} \in \mathbb{I}_w$, the restriction of α to the sub-domain V_w . In G' , there is a directed edge from $\alpha \in \mathbb{I}_w$ to $\beta \in \mathbb{I}_v$ if and only if there is a directed edge from w to v in E and the restriction map induced by this edge maps β to α .

- (a) G' for Hierarchical Exchangeability: The infinite graphs for Austin–Panchenko arrays are collections of infinitely-branching trees. See Figure 2.8. Concretely, consider such an array for the DAG G in Figure 2.1. Let w_1, \dots, w_ℓ be the terminal vertices of ℓ paths in G ; in the figure, they are labeled by $v_{r_1}^{(1)}, \dots, v_{r_\ell}^{(\ell)}$. The graph G' in this case consists of ℓ infinitely-branching trees of depths r_1, \dots, r_ℓ , respectively. The multi-indices for the array are tuples $(\alpha_1, \dots, \alpha_\ell)$ of vertices of G' such that $\alpha_i \in \mathbb{I}_{w_i}$ for $1 \leq i \leq \ell$. They can also be understood as tuples of ℓ paths in G' , where the i^{th} path starts from the root of the i^{th} tree and repeatedly moves toward the leaves by taking the $v_j^{(i)}$ child at step j until the path hits a leaf. Also, \mathbb{I}_{w_i} is isomorphic to \mathbb{N}^{r_i} for all i in this case. Thus, the multi-indices just defined are precisely the elements of $\mathbb{N}^{r_1} \times \dots \times \mathbb{N}^{r_\ell}$, the multi-index set that we have used to describe Austin–Panchenko arrays thus far.

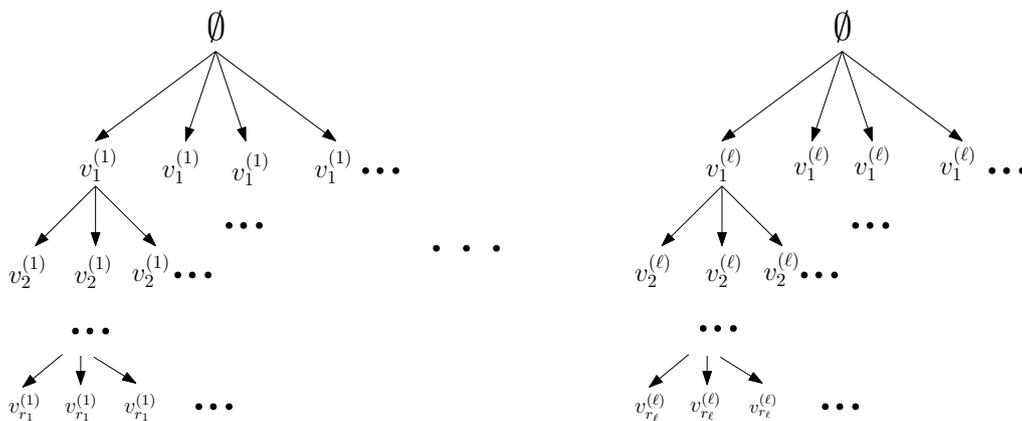


Figure 2.8. An Austin–Panchenko forest with ℓ trees.

- (b) G' for Random Block Matrices: As shown in Figure 2.9, the G' corresponding to random block matrices has infinitely many copies of r_0, c_0, r_1 and c_1 , respectively. The copies of r_0 and c_0 correspond to the rows and columns of the outer matrix, and those of r_1 and c_1 to the rows and columns of the inner nested matrices. The latter copies of r_1 and c_1 are grouped when they belong to the same nested matrix, and the copies in the same group have incoming edges from one copy of r_0 and one copy of c_0 , which express the position of

the block (or inner matrix) within the outer matrix. The multi-indices in this case are pairs of copies of r_1 and c_1 that belong to the same group.

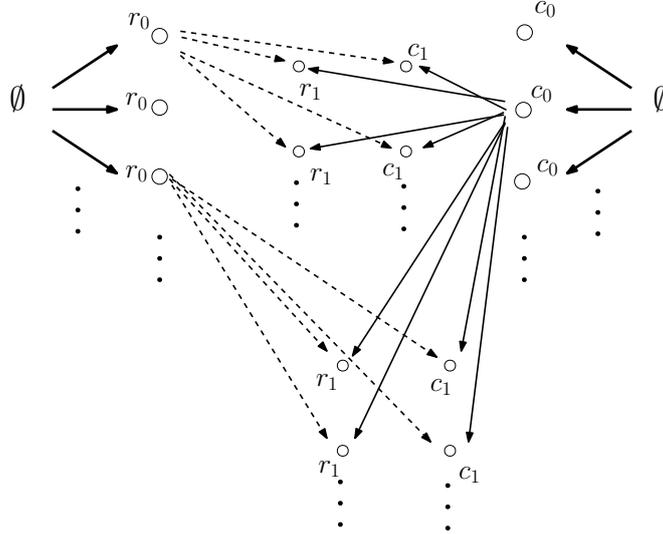


Figure 2.9. Part of the infinitely-branching DAG G' corresponding to random block matrices.

2.1. Overview of Infinitary DAGs and Their Finite Presentations

This section is indeed confusing. The complication comes because we tried to justify the use of DAG in the title of the paper. It may be better just to say that the index set is \mathbb{N}^G for some DAG G in this part of other paper.

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 indices of the random arrays we will consider. *Infinitary* here means each vertex, other than a terminal vertex, has 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

- (a) \mathbb{N}
- (b) $\mathbb{N} \times \mathbb{N}$
- (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

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

All four index sets above are associated with infinitary DAGs, in fact forests or collections of trees, of finite depth. A given index in each of the four index sets

then corresponds to a collection of leaves of the associated DAG (one leaf from each tree in the forest). While it is possible to also view indices of general DAGs as a collection of “terminal vertices”, for complete generality, it is much more convenient to associate a subgraph to this collection of leaves— this subgraph will just be a collection of paths (one path from each tree in the forest). 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. The index set in (c') is associated with a DAG consisting of ℓ infinitary trees of depths r_1, \dots, r_ℓ , respectively. The element

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

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 $v_j^{(i)}$ 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 method for constructing an index set by first building an infinitary DAG of finite depth. The four index sets that we discussed above can all be naturally constructed by this method.

Assume that a *finite* DAG G is given. This G depicts the skeleton of an *infinitary* DAG to be built. 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, \dots, r_ℓ (see *Figure 2.1*).

Next, we 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

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

Here a path may be zero-long, and D_v always contains v . For instance, if G is just a single vertex v without any edges, $D_v = \{v\}$. The vertex set of G' is $\bigcup_v \mathbb{N}^{D_v}$ where $\mathbb{N}^{D_v} \stackrel{\text{def}}{=} \mathbb{N}^{|D_v|}$. That is, each vertex of G' is a function

$$\alpha : D_v \rightarrow \mathbb{N},$$

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 then used to connect the various copied vertices with directed edges. In particular, 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, and they turn out to be forests (see *Figure 2.8*). We will soon see examples where the infinitary DAGs are not forests (see *Figure 2.9*).

Finally, we set the index set to be

$$\{\alpha : \alpha \text{ is a function from the vertex set of } G \text{ to } \mathbb{N}\}.$$

One should view this as a collection of subgraphs of G' in the following way: For each $\alpha : G \rightarrow \mathbb{N}$, the subgraph of G' associated to α has vertex and edge sets

$$V'' = \left\{ \alpha \Big|_{D_v} : v \in G \right\}, \quad E'' = \left\{ (\alpha, \beta) : (\alpha, \beta) \text{ is an edge in } G' \text{ and } \alpha, \beta \in V'' \right\}.$$

3. Main Result

Let G be a finite DAG and recall that \mathcal{A}_G denotes the set of all closed subsets of G . By definition, a G -automorphism τ induces a bijection on the C -type multi-indices $\beta \in \mathbb{N}^C$ for any $C \in \mathcal{A}_G$:

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

Slightly abusing notation, we reuse τ to denote this induced map. Also, a bijection $\tau : \mathbb{N}^C \rightarrow \mathbb{N}^C$ acts on a C -type random array $\mathbf{X}_C = (X_\alpha : \alpha \in \mathbb{N}^C)$ by

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

DEFINITION 3.1. — *Let \mathcal{C} be a sequence of distinct closed subsets of G . A \mathcal{C} -type random array collection $\mathbb{X} = (\mathbf{X}_C : C \in \mathcal{C})$ is DAG-exchangeable if it is equal in distribution to $(\tau(\mathbf{X}_C) : C \in \mathcal{C})$ for every G -automorphism τ , that is,*

$$\mathbb{X} \stackrel{d}{=} (\tau(\mathbf{X}_C) : C \in \mathcal{C}).$$

We denote the set of all multi-indices over some $D \in \mathcal{A}_G$ by

$$I_C \stackrel{\text{def}}{=} \bigcup_{D \in \mathcal{A}_G} \mathbb{N}^D.$$

Also, we introduce the following notations for multi-indices $\alpha \in I_G$:

$$\begin{aligned} \text{Dom} : I_G &\rightarrow \mathcal{A}_G, & \text{Dom}(\alpha) &\stackrel{\text{def}}{=} \text{the set of vertices where } \alpha \text{ is defined,} \\ \text{Restr} : I_G &\rightarrow 2^{\mathcal{A}_G}, & \text{Restr}(\alpha) &\stackrel{\text{def}}{=} \{ \alpha|_C : C \in \mathcal{A}_G \}, \\ \text{Restr}^o : I_G &\rightarrow 2^{\mathcal{A}_G}, & \text{Restr}^o(\alpha) &\stackrel{\text{def}}{=} \text{Restr}(\alpha) \setminus \{ \alpha \}. \end{aligned}$$

The $\text{Dom}(\alpha)$ is the domain of the multi-index α , and the next two are about the restrictions of α : $\text{Restr}(\alpha)$ is the set of all the restrictions, while $\text{Restr}^o(\alpha)$ consists of only the strict restrictions.

Consider the array $(U_\alpha : \alpha \in I_G)$, where the U_α 's are i.i.d. uniform random variables. Let

$$\mathbf{V}_\alpha \stackrel{\text{def}}{=} (U_\beta : \beta \in \text{Restr}(\alpha)).$$

Then, for all G -automorphisms τ ,

$$(\mathbf{V}_\alpha : \alpha \in I_G) \stackrel{d}{=} (\mathbf{V}_{\tau(\alpha)} : \alpha \in I_G)$$

Thus, for any sequence \mathcal{C} of distinct closed subsets of G and any family of measurable functions $(f_C : C \in \mathcal{C})$ with $f_C : [0, 1]^{\mathcal{A}_C} \rightarrow \mathcal{X}$,

$$\left((f_C(\mathbf{V}_\alpha) : \alpha \in \mathbb{N}^C) : C \in \mathcal{C} \right)$$

is DAG-exchangeable, where for the argument of f_C we identify U_β with the $Dom(\beta)$ -coordinate of the product space $[0, 1]^{A_C}$.

As usual, our representation theorem is the converse of the previous statement.

THEOREM 3.2. — *If $(\mathbf{X}_C : C \in \mathcal{C})$ is DAG-exchangeable, then*

$$\left((X_{C,\alpha} : \alpha \in \mathbb{N}^C) : C \in \mathcal{C} \right) \stackrel{d}{=} \left((f_C(U_\beta : \beta \in Restr(\alpha)) : \alpha \in \mathbb{N}^C) : C \in \mathcal{C} \right)$$

for some family of measurable functions $(f_C : C \in \mathcal{C})$ with $f_C : [0, 1]^{A_C} \rightarrow \mathcal{X}$ and independent $[0, 1]$ -uniform random variables $\mathbf{U} = (U_\alpha : \alpha \in I_G)$.

Remark 3.3. — This is the fine-grained generalization of Theorem 1.5 alluded to in the introduction. Indeed, Theorem 1.5 is a simple corollary of the above theorem, since it is just the special case $\mathcal{C} = \{G\}$.

Example 3.4. — Let us illustrate the application of the above theorem with a very simple example. Consider Example 2.2(a), which uses the DAG in Figure 2.2. Let \mathbb{X} be a \mathcal{C} -type random array collection where $\mathcal{C} = (R, C)$ for $R = \{s, r\}$ and $C = \{s, c\}$. Define

$$\mathbf{X}' \stackrel{\text{def}}{=} \left((X_{R,ij}, X_{C,ik}) : i, j, k \in \mathbb{N} \right).$$

This array \mathbf{X}' is just a way of rewriting \mathbb{X} with straightforward adjustment on indices, and it is easy to see that \mathbb{X} is DAG-exchangeable if and only if

$$\mathbf{X}' \stackrel{d}{=} \left((X_{R,\pi(i)\tau_i(j)}, X_{C,\pi(i)\rho_i(k)}) : i, j, k \in \mathbb{N} \right)$$

for all permutations $\pi, \tau_i, \rho_i \in S_{\mathbb{N}}$ with $i \in \mathbb{N}$. If \mathbb{X} is DAG-exchangeable, Theorem 3.2 tells us that \mathbf{X}' has a representation of the following form:

$$\begin{aligned} & \left((X_{R,ij}, X_{C,ik}) : i, j, k \in \mathbb{N} \right) \\ & \stackrel{d}{=} \left((f_R(U_{000}, U_{i00}, U_{ij0}), f_C(U_{000}, U_{i00}, U_{i0k})) : i, j, k \in \mathbb{N} \right) \end{aligned}$$

for some measurable functions f_R and f_C and independent $[0, 1]$ -uniform random variables U_{ijk} for $i, j, k \in \{0\} \cup \mathbb{N}$.

The main result of this paper is a probabilistic proof of Theorem 3.2. Our proof is based on an induction whose inductive step involves reasoning about sophisticated conditional independence, similar to other proofs in the exchangeability literature [AP14, Kal05]. More concretely, in the next subsection 3.1, we provide a different version of Theorem 3.2, from which the theorem follows immediately. Then, we give a detailed proof of this strengthened version of Theorem 3.2 in Section 4.

Remark 3.5. — In Appendix A, we provide an alternative model-theoretic proof of the special case Theorem 1.5 using a result of Crane and Towsner on the representation of relatively exchangeable random structures [CT17]. Crane and Towsner’s result has been formulated and proved in a model-theoretic setting. A large part of our second proof is about translating the graph-theoretic statement of Theorem 1.5 to a model-theoretic one in Crane and Towsner’s representation theorem, and showing

that after translation, the statement satisfies the conditions of Crane and Towsner, and when translated backwards, their conclusion gives the claimed representation of our theorem.

3.1. Representations of Random Arrays Induced by Symmetries

We start with a motivation for this section. Consider de Finetti’s theorem, the most simple and fundamental result on exchangeability. The theorem is commonly stated as, “for any exchangeable random sequence, there is a random measure $\mu(\cdot)$ such that conditional on a realization $\mu(\omega)$, the sequence is (conditionally) i.i.d. with common distribution $\mu(\omega)$.” The theorem was restated by [Ald81] as a representation in terms of uniform random variables:

For any exchangeable sequence $\mathbf{X} = (X_n : n \in \mathbb{N})$ on a Borel space \mathcal{X} , there exists a measurable function $f : [0, 1]^2 \rightarrow \mathcal{X}$ such that for an i.i.d. sequence of uniform random variables $\mathbf{U} = (U_n : n \in \mathbb{N} \cup \{0\})$,

$$(3.1) \quad (X_n : n \in \mathbb{N}) \stackrel{d}{=} \left(f(U_0, U_n) : n \in \mathbb{N} \right).$$

Here, U_0 takes the role of selecting the random measure μ , while each U_n samples X_n under the law μ . If we choose \mathbf{U} so that (3.1) is true almost surely instead of in-distribution by the transfer Theorem [Kal02, Theorem 6.10], one can easily see that U_0 is independent of \mathbf{X} given μ . (See [Kal05, Lemma 7.1 and Theorem 1.1] for example. This type of result is also commonly known as the Hewitt–Savage theorem.)

In this section, we define an array associated to DAG-exchangeable arrays, which we call its *symmetry array*. It generalizes the role that the empirical random measure μ plays for exchangeable sequences. We also propose a variant of Theorem 3.2, which states that the uniform random variables in the representation affect the array only through this symmetry array. This corresponds to the fact that U_0 is independent of \mathbf{X} given μ .

Let \mathcal{C} be a sequence of distinct closed subsets of G , and $\mathbb{X} = (\mathbf{X}_C : C \in \mathcal{C})$ be a DAG-exchangeable \mathcal{C} -type random array collection.

We say that a G -automorphism τ of \mathbb{N}^G fixes $\alpha \in I_G$ if $\tau(\alpha) = \alpha$. We define \mathcal{F}_α to be the sub- σ -field of $\sigma(\mathbb{X})$ consisting of \mathbb{X} -measurable events that are invariant under every α -fixing G -automorphism τ :

$$(3.2) \quad \mathcal{F}_\alpha \stackrel{\text{def}}{=} \sigma \left(\left\{ \mathbb{X}^{-1}(B) : B \text{ is Borel, and if } \left((x_\beta : \beta \in \mathbb{N}^C) : C \in \mathcal{C} \right) \in B \text{ and } \tau \text{ fixes } \alpha, \text{ then } \left((x_{\tau(\beta)} : \beta \in \mathbb{N}^C) : C \in \mathcal{C} \right) \in B \right\} \right)$$

One should think of the \mathcal{F}_α ’s as the σ -fields which contain the information concerning the symmetries in the array. In particular, \mathcal{F}_α contains the information about the symmetries in the array which fix the multi-index α .

For instance, in Example 3.4, consider $\alpha \in \mathbb{N}^R$ defined by $\alpha(s) = 1$ and $\alpha(r) = 3$. Then, $X_{R,13}$ is \mathcal{F}_α -measurable, but in general any of $X_{R,23}$, $X_{R,12}$ and $X_{C,13}$ is not.

Another example is the empirical distribution \mathcal{E} of the sequence $(X_{R,13}, X_{C,1k} : k \in \mathbb{N})$, that is,

$$\mathcal{E} \stackrel{\text{def}}{=} \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k \leq n} \delta_{(X_{R,13}, X_{C,1k})}.$$

By de Finetti–Hewitt–Savage theorem, \mathcal{E} exists almost surely. One can also easily check that it is \mathcal{F}_α -measurable. Let

$$\mathcal{E}' \stackrel{\text{def}}{=} \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k \leq n} \delta_{X_{C,1k}}.$$

This empirical distribution is \mathcal{F}_α -measurable as well. In fact, it is measurable with respect to a smaller σ -field $\mathcal{F}_{\alpha|_{\{s\}}} \subseteq \mathcal{F}_\alpha$.

Remark 3.6. — Restricting α shrinks the σ -field \mathcal{F}_α . That is, for all closed $D \subseteq \text{Dom}(\alpha)$, we have that $\mathcal{F}_{\alpha|_D} \subseteq \mathcal{F}_\alpha$. This is because every α -fixing τ is also an $\alpha|_D$ -fixing G -automorphism and so an event invariant under the latter kind of G -automorphism is also invariant under the former kind.

It will be convenient to encode into random variables, the information contained in the various \mathcal{F}_α 's. These random variables will later serve as a mechanism by which we make different overlapping representations consistent (in a manner later described). Using the facts that the elements of each \mathbf{X}_C take values in a Borel space and that each \mathcal{F}_α is countably generated, we may define:

DEFINITION 3.7. — *Given a DAG-exchangeable array collection \mathbb{X} , we define an associated random symmetry array $\mathbf{S} = (S_\alpha : \alpha \in I_G)$ (after extending the underlying probability space if needed) to be any array \mathbf{S} satisfying*

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

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

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

To see why such an \mathbf{S} exists, fix $C_0 \in \mathcal{C}$ and pick $\alpha_0 \in \mathbb{N}^{C_0}$. There exists a random variable S_{α_0} that generates \mathcal{F}_{α_0} ([Res14, Chapter 3, Exercice 13]). Furthermore, since S_{α_0} is \mathbb{X} -measurable, there exists a measurable function f such that $S_{\alpha_0} = f(\mathbf{X}_C : C \in \mathcal{C})$. Now for each G -automorphism τ , we can define

$$S_{\tau(\alpha_0)} \stackrel{\text{def}}{=} f(\tau(\mathbf{X}_C) : C \in \mathcal{C}).$$

Repeating this procedure for each $C \in \mathcal{C}$ gives $\mathbf{S} = ((S_\alpha : \alpha \in \mathbb{N}^C) : C \in \mathcal{C})$, which satisfies the two required properties.

For all $\beta \in I_G$, if β is a restriction of α , the random variable S_β is \mathcal{F}_α -measurable. This is because S_β is \mathcal{F}_β -measurable but the σ -field \mathcal{F}_β is included in \mathcal{F}_α .

PROPOSITION 3.8. — *Let \mathcal{C} be a sequence of distinct closed subsets of G and let \mathbf{S} be the symmetry array defined as above. If $(\mathbf{X}_C : C \in \mathcal{C})$ is DAG-exchangeable, there exist a family of measurable functions $(h_C : C \in \mathcal{A}_G)$ with $h_C : [0, 1]^{A_C} \rightarrow \mathcal{X}$*

and a collection of independent $[0, 1]$ -uniform random variables $\mathbf{U} = (U_\alpha : \alpha \in I_G)$ such that

$$(3.3) \quad (S_\alpha : \alpha \in I_G) \stackrel{d}{=} (S'_\alpha : \alpha \in I_G)$$

where

$$(3.4) \quad S'_\alpha \stackrel{\text{def}}{=} h_{\text{Dom}(\alpha)} \left((S'_\beta : \beta \in \text{Restr}^o(\alpha)), U_\alpha \right)$$

for $\alpha \in I_G$.

Proposition 3.8 provides a representation for \mathbf{S} that is built out of a collection of independent random variables $(U_\alpha : \alpha \in I_G)$ and appropriate measurable functions. The representation is given in terms of the inductively-defined random variables $(S'_\alpha : \alpha \in I_G)$, with induction being applied to the size of the domain of each multi-index in I_G . Two immediate consequences of the representation are that each S_α depends only on $(U_\beta : \beta \in \text{Restr}(\alpha))$, and that its dependence on $(U_\beta : \beta \in \text{Restr}^o(\alpha))$ is always mediated via $(S'_\beta : \beta \in \text{Restr}^o(\alpha))$.

We now show that Proposition 3.8 implies Theorem 3.2. Note that using induction, we can convert $h_{\text{Dom}(\alpha)}$ to a function $h'_{\text{Dom}(\alpha)}$ for each α such that

$$(3.5) \quad (S_\alpha : \alpha \in I_G) \stackrel{d}{=} \left(h'_{\text{Dom}(\alpha)}(U_\beta : \beta \in \text{Restr}(\alpha)) : \alpha \in I_G \right).$$

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

$$\begin{aligned} & h'_{\text{Dom}(\alpha)}(U_\beta : \beta \in \text{Restr}(\alpha)) \\ &= h_{\text{Dom}(\alpha)} \left(\left(h'_{\text{Dom}(\beta)}(U_\gamma : \gamma \in \text{Restr}(\beta)) : \beta \in \text{Restr}^o(\alpha) \right), U_\alpha \right). \end{aligned}$$

Proof of Theorem 3.2. — $X_{D,\alpha}$ is \mathbb{X} -measurable, and it is fixed under the action of every α -fixing G -automorphism. Thus, $X_{D,\alpha}$ is \mathcal{F}_α -measurable by the definition of the σ -field \mathcal{F}_α . This means that $X_{D,\alpha}$ is also S_α -measurable because $\sigma(S_\alpha) = \mathcal{F}_\alpha$. Furthermore, $X_{D,\alpha}$ takes values in a Borel space. Thus, there exists a measurable function f_α such that $X_{D,\alpha} = f_\alpha(S_\alpha)$ almost surely. By the DAG-exchangeability of the collection of ordered pairs $((\mathbf{X}_C, \mathbf{S}_C) : C \in \mathcal{C})$, we can pick f_α such that it depends only on $\text{Dom}(\alpha)$ and not on the value of α itself. This means that we can write $X_{D,\alpha} = f_D(S_\alpha)$ almost surely, by writing f_α as $f_{\text{Dom}(\alpha)}$. Plugging in (3.5) finishes the proof. \square

Remark 3.9. — Before getting into the proof of Proposition 3.8, we recall a generic property of exchangeable structures. Whenever $\mathbf{X} = (X_n : n \in \mathbb{N})$ is a sequence, by Kolmogorov’s extension theorem, its exchangeability is equivalent to the seemingly weaker condition that the distribution of \mathbf{X} is invariant under the action of *finite* permutations (permutations fixing all but finitely many elements). In particular, if \mathbf{X} is exchangeable, then

$$\mathbf{X} = (X_n : n \in \mathbb{N}) \stackrel{d}{=} (X_{\tau(n)} : n \in \mathbb{N}) = \tau(\mathbf{X})$$

Table 3.1. Notation Guide

Symbol	Object
G	finite DAG or vertex set of a finite DAG
C	downward-closed (w.r.t. partial ordering) subset of G
W, H	arbitrary subsets of G
\mathcal{A}_G	set of all downward-closed subsets of G
\mathcal{C}	sequence of distinct, downward-closed subsets of G (i.e. $\mathcal{C} \subset \mathcal{A}_G$)
$\mathbb{N}^G, \mathbb{N}^C, \mathbb{N}^H$	index sets corresponding to vertex sets of G, C, H
I_C	the multi-graph index set: $\bigcup_{D \in \mathcal{A}_C} \mathbb{N}^D$
α, β, γ	elements of some index set
τ, ρ	G -automorphisms

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 \mathbf{X} .

Let Y be \mathbf{X} -measurable and let K be a subgroup of the infinite permutations. By definition, Y is invariant under the action of K if and only if $(Y, \mathbf{X}) \stackrel{d}{=} (Y, \tau(\mathbf{X}))$ for all $\tau \in K$. By the above paragraph, this is equivalent to having

$$(Y, X_n : n \in F) \stackrel{d}{=} (Y, X_{\tau(n)} : n \in F)$$

for all $\tau \in K$ and all *finite* $F \subseteq \mathbb{N}$. Moreover, we have $(Y, \mathbf{X}) \stackrel{d}{=} (Y, \rho(\mathbf{X}))$ for any injection ρ on \mathbb{N} such that its arbitrary restriction to finite sets can be extended to an element in K .

4. Proof of the Main Result

To prove our main result, it remains to prove Proposition 3.8. Let G, \mathcal{C} and \mathbb{X} be a finite DAG, a sequence of distinct closed subsets of G , and a \mathcal{C} -type DAG-exchangeable random array collection from the proposition. Also let \mathbf{S} be the symmetry array defined as in Section 3.1.

Overview of the proof of Proposition 3.8

As typical for probabilistic proofs in the exchangeability literature, our proof is by induction. Before describing an overview of our proof, it is pedagogical to introduce a natural alternative approach which is also an induction; the difficulty in realizing this alternative approach helps clarify what we believe to be the “crux” of proving a representation theorem for DAG-exchangeable arrays, and is what eventually guides us in how to organize the actual proof. The alternative approach is first to apply the inductive hypothesis to every $C \in \mathcal{A}_G \setminus \{G\}$, i.e. assume a representation (3.3) for $(S_\alpha : \alpha \in I_C)$ for every such C , and then to prove the inductive step by combining these representations to get (3.3) for the entirety of G . Although this seems natural,

this approach is difficult to implement. This is because different C and C' in $\mathcal{A}_G \setminus \{G\}$ may still share vertices (i.e., $C \cap C' \neq \emptyset$) and the representations obtained by applying the inductive hypothesis to C and C' may differ on those vertices; these representations may induce different representations for $(S_\alpha : \alpha \in I_{C \cap C'})$. We will henceforth say that these two representations are *consistent* if they are the same for all $\alpha \in I_{C \cap C'}$.

The architecture of our induction is built to overcome the above-described difficulty, i.e., built to make representations corresponding to C and C' consistent on $C \cap C'$. In fact, there are two levels of induction in our proof. At the top level, there is a rather simple induction on the number of vertices n of G . The top-level inductive assumption allows us to assume representations for all closed subsets with less than n vertices. The more difficult second-level induction is designed to make consistent, in a systematic way, the possibly different representations for all the different closed subsets having less than n vertices. The base step $k = 0$ of our second-level induction, is to choose from any one of the closed subsets of size $n - 1$, a representation for the array $(S_\alpha : \alpha \in I_{G_0})$ for the closed G_0 defined by

$$(4.1) \quad G_0 \stackrel{\text{def}}{=} G \setminus T.$$

Here T is the set of all *terminal* vertices (i.e. vertices with no descendants). Our second-level induction is an induction on the size of sets $A \subset T$ that we will now add to G_0 . More specifically, the inductive step of our second-level induction is to show that, whenever consistent representations for arrays of the form $(S_\alpha : \alpha \in I_{G_0 \cup A})$ exist for A 's such that $|A| = k - 1$, and potentially non-consistent representations exist also for $|A| = k$, then one can appropriately combine the consistent representations at level $k - 1$ to obtain consistent representations at level k .

The second-level induction gives us consistent representations (3.3) for arrays of the form

$$(S_\alpha : \alpha \in I_{G_0 \cup A} \text{ for some } A \subseteq T, |A| = k)$$

for all k such that $|G_0| + k < |G|$. To finish the proof, we must complete the top-level induction by extending the consistent representations at level $n - 1$ (obtained via the second-level induction), which give a joint representation of the array $(S_\alpha : \alpha \in I_G \setminus \mathbb{N}^G)$, to the whole array $(S_\alpha : \alpha \in I_G)$. This final step is easily obtained by an application of an elementary coding lemma (Lemma B.3).

Now, for each $C \in \mathcal{A}_G$, set

$$\mathcal{F}_C \stackrel{\text{def}}{=} \sigma(\mathbf{S}_C).$$

Clearly, we have $\mathcal{F}_D \subseteq \mathcal{F}_C$ whenever $D \in \mathcal{A}_C$. The key to carrying out the above described induction is the following proposition. We will prove this in Section 4.1.

PROPOSITION 4.1. — *Let C, C_1, \dots, C_m be closed subsets of G . Then,*

$$\mathcal{F}_C \perp_{(\mathcal{F}_{C \cap C_i})_{i \leq m}} (\mathcal{F}_{C_i})_{i \leq m}.$$

The proposition will be utilized via the following two immediate Corollaries 4.2 and 4.3. Let the set of all terminal vertices of G be denoted by

$$(4.2) \quad T = \{v_1, \dots, v_t\}, \quad \text{where } |T| = t.$$

Define

$$(4.3) \quad G_0 \stackrel{\text{def}}{=} G \setminus T \quad \text{and} \quad \mathcal{G}_A \stackrel{\text{def}}{=} \mathcal{F}_{G_0 \cup A} \quad \text{for} \quad A \subseteq T.$$

COROLLARY 4.2. — *Let $B, B_1, \dots, B_m \subseteq T$. Then,*

$$\mathcal{G}_B \perp_{(\mathcal{G}_{B \cap B_i})_{i \leq m}} \prod_{i \leq m} (\mathcal{G}_{B_i})_{i \leq m}.$$

Proof. — Apply Proposition 4.1 with $C = G_0 \cup B$ and $C_i = G_0 \cup B_i$. □

COROLLARY 4.3. — *For $k = 0, \dots, t$, let*

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

Then, given \mathcal{G}_{k-1} , the set \mathcal{H}_k is a family of independent σ -fields for all $1 \leq k \leq t$.

Proof. — When $k = t$, the result is immediate because \mathcal{H}_k is a singleton set. Assume that $k < t$. Let $A \subseteq T$ be such that $|A| = k$. Set

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

and

$$\mathcal{G}' \stackrel{\text{def}}{=} \sigma(\{\mathcal{G}_{B \cap A} : B \subseteq T, |B| = k, B \neq A\}).$$

By Corollary 4.2, we have $\mathcal{G}_A \perp_{\mathcal{G}'} \mathcal{G}_{k \setminus A}$. But

$$\mathcal{G}' \subseteq \mathcal{G}_{k-1} \subseteq \sigma(\mathcal{G}' \cup \mathcal{G}_{k \setminus A}).$$

This is because $\mathcal{G}_{B \cap A} \subseteq \mathcal{G}_B$ for every $B \subseteq T$ (which itself follows from the fact that $\mathcal{F}_D \subseteq \mathcal{F}_C$ for closed D, C whenever $D \subseteq C$). Thus, we have that

$$\mathcal{G}_A \perp_{\mathcal{G}_{k-1}} \mathcal{G}_{k \setminus A},$$

from which the result follows. □

For any subset I of I_G and random array $\mathbf{Y} = (Y_\alpha : \alpha \in I_G)$, let us denote the sub-array

$$\mathbf{Y}_I \stackrel{\text{def}}{=} (Y_\alpha : \alpha \in I).$$

Proof of Proposition 3.8. — Without loss of generality we will assume that \mathcal{C} is the set \mathcal{A}_G with some fixed ordering. We use two levels of induction in the proof. The top-level induction is on the number of vertices of G where the $n = 1$ case is simply the de Finetti–Hewitt–Savage theorem. Using the inductive hypothesis for the $n - 1$ case, assume representations exist for

$$(4.4) \quad (S_\alpha : \alpha \in I_{G_0 \cup A} \text{ for some } A \subseteq T, |A| = k)$$

whenever $|G_0| + k < |G|$. Our first objective is to show that such representations can be chosen to be consistent in the sense described at the beginning of this section.

Set $|T| = t$. The case where $|T| = 1$ is obtained directly from the inductive hypothesis and Lemma 4.4, below. In the rest of the proof, we assume that $|T| = t > 1$.

Choose a closed subset of G with $n - 1$ vertices. By the fact that it is closed, it must contain G_0 . By the assumption of (4.4), there exist Borel functions $\{g_C : C \in \mathcal{A}_{G_0}\}$

as well as an array of independent $[0, 1]$ -uniform random variables, \mathbf{U} (which we can assume to be independent from all of the symmetry arrays below), such that

$$(4.5) \quad \begin{aligned} (S_\alpha : \alpha \in I_{G_0}) &\stackrel{d}{=} (S'_\alpha : \alpha \in I_{G_0}), \\ S'_\alpha &\stackrel{\text{def}}{=} g_{\text{Dom}(\alpha)} \left(\mathbf{S}'_{\text{Restr}^o(\alpha)}, U_\alpha \right) \quad \text{for } \alpha \in I_{G_0}. \end{aligned}$$

Similarly, for each $A \subsetneq T$ there exist $\{g_C^A : C \in \mathcal{A}_{G_0 \cup A}\}$ such that

$$(4.6) \quad \begin{aligned} (S_\alpha : \alpha \in I_{G_0 \cup A}) &\stackrel{d}{=} (S_\alpha^A : \alpha \in I_{G_0 \cup A}), \\ S_\alpha^A &\stackrel{\text{def}}{=} g_{\text{Dom}(\alpha)}^A \left(\mathbf{S}_{\text{Restr}^o(\alpha)}^A, U_\alpha \right) \quad \text{for } \alpha \in I_{G_0 \cup A}. \end{aligned}$$

Here, the arrays $\mathbf{S}'_{\text{Restr}^o(\alpha)}$ and $\mathbf{S}_{\text{Restr}^o(\alpha)}^A$ are defined recursively through (4.5) and (4.6). We must next show that the above representations can be chosen to be consistent, to which end we use another (second-level) induction on the sizes of the A 's, say $|A| = k$.

It is pedagogical to go through the easiest step of induction, from $k = 0$ to 1, before dealing with the general inductive step. To simplify notation, let $A = \{v_s\}$ and set

$$(4.7) \quad S_\alpha^s \stackrel{\text{def}}{=} g_{\text{Dom}(\alpha)}^s \left(\mathbf{S}_{\text{Restr}^o(\alpha)}^s, U_\alpha \right) \quad \text{for } \alpha \in I_{G_0 \cup \{v_s\}}$$

and

$$\boldsymbol{\eta}_s \stackrel{\text{def}}{=} \mathbf{S}_{I_{G_0 \cup \{v_s\}}}, \quad \boldsymbol{\theta}_s \stackrel{\text{def}}{=} \mathbf{S}_{I_{G_0 \cup \{v_s\}}}, \quad \mathbf{U}_s \stackrel{\text{def}}{=} \mathbf{U}_{I_{G_0 \cup \{v_s\}} \setminus I_{G_0}}.$$

Using the second equation of (4.7) for each $\alpha \in I_{G_0 \cup \{v_s\}}$, we can express each S_α^s in terms of the S_β^s 's and U_γ 's with $\beta \in I_{G_0}$ and $\gamma \in I_{G_0 \cup \{v_s\}} \setminus I_{G_0}$. The resulting equations can be written as

$$\boldsymbol{\theta}_s = F_s \left(\mathbf{S}_{I_{G_0}}^s, \mathbf{U}_s \right)$$

for an appropriate measurable F_s . By Corollary 4.3,

$$\mathbf{S}_{I_{G_0}}^{\perp\perp} (\boldsymbol{\eta}_s)_s.$$

By construction, $\mathbf{S}_{I_{G_0}} \stackrel{d}{=} \mathbf{S}'_{I_{G_0}} \stackrel{d}{=} \mathbf{S}_{I_{G_0}}^s$ are independent from $(\mathbf{U}_s)_s$, and $(\mathbf{S}_{I_{G_0}}, \boldsymbol{\eta}_s) \stackrel{d}{=} (\mathbf{S}_{I_{G_0}}^s, \boldsymbol{\theta}_s)$. Therefore, by Lemma B.2, we have

$$(\mathbf{S}_{I_{G_0}}, \boldsymbol{\eta}_s)_s \stackrel{d}{=} \left(\mathbf{S}'_{I_{G_0}}, F_s(\mathbf{S}'_{I_{G_0}}, \mathbf{U}_s) \right)_s.$$

Thus we can join the representations given by (4.5) and (4.6) to obtain the following joint distributional equality

$$(4.8) \quad \left(S_\alpha : \alpha \in I_{G_0 \cup \{v_s\}}, s \leq t \right) \stackrel{d}{=} \left(S_\alpha^1 : \alpha \in I_{G_0 \cup \{v_s\}}, s \leq t \right),$$

where $S_\alpha^1 = S'_\alpha$ if $\alpha \in I_{G_0}$ and the rest of the S_α^1 's are defined by the recursive formulae

$$S_\alpha^1 = g_{\text{Dom}(\alpha)}^s \left(\mathbf{S}_{\text{Restr}^o(\alpha)}^1, U_\alpha \right).$$

This is a (consistent) joint representation of

$$(S_\alpha : \alpha \in I_{G_0 \cup A}, A \subseteq T, |A| = k),$$

in the case where $k = 1$.

Let us now generalize the above by carrying out the second-level inductive step on general k to achieve an analogous joint representation at the level $k = t - 1$. Now set $k < t - 1$ to be fixed and assume that we have the following joint representation:

$$(4.9) \quad (S_\alpha : \alpha \in I_{G_0 \cup A}, A \subseteq T, |A| = k) \\ \stackrel{d}{=} (S_\alpha^k : \alpha \in I_{G_0 \cup A}, A \subseteq T, |A| = k) \\ S_\alpha^k \stackrel{\text{def}}{=} g_{\text{Dom}(\alpha)}(\mathbf{S}_{\text{Restr}^\circ(\alpha)}^k, U_\alpha) \quad \text{for } \alpha \in \bigcup_{|A|=k} I_{G_0 \cup A}.$$

Consider the representation in (4.6) for any fixed $B \subseteq T$ with $|B| = k + 1$ (note that $k + 1 \leq t - 1$), and with Borel measurable functions $(g_C^B : C \in \mathcal{A}_{G_0 \cup B})$. We rewrite it here for convenience:

$$(4.10) \quad (S_\alpha : \alpha \in I_{G_0 \cup B}) \stackrel{d}{=} (S_\alpha^B : \alpha \in I_{G_0 \cup B}), \\ S_\alpha^B \stackrel{\text{def}}{=} g_{\text{Dom}(\alpha)}^B(\mathbf{S}_{\text{Restr}^\circ(\alpha)}^B, U_\alpha) \quad \text{for } \alpha \in I_{G_0 \cup B}.$$

Define the following arrays (to ease notation we do not use boldface for these):

$$\eta_B \stackrel{\text{def}}{=} (S_\alpha : \alpha \in I_{G_0 \cup A}, A \subseteq B, |A| = k), \\ \theta'_B \stackrel{\text{def}}{=} (S_\alpha^k : \alpha \in I_{G_0 \cup A}, A \subseteq B, |A| = k), \\ \theta_B^B \stackrel{\text{def}}{=} (S_\alpha^B : \alpha \in I_{G_0 \cup A}, A \subseteq B, |A| = k), \\ U_B \stackrel{\text{def}}{=} (U_\alpha : \alpha \in I_{G_0 \cup A}, A \subseteq B, |A| = k), \\ U_{B^c} \stackrel{\text{def}}{=} (U_\alpha : \alpha \in I_{G_0 \cup A}, A \subseteq T, |A| = k) \setminus U_B, \\ \partial U_B \stackrel{\text{def}}{=} (U_\alpha : \alpha \in I_{G_0 \cup B}) \setminus U_B,$$

$$\eta_T \stackrel{\text{def}}{=} (S_\alpha : \alpha \in I_{G_0 \cup A}, A \subseteq T, |A| = k), \quad \bar{\eta}_B \stackrel{\text{def}}{=} (S_\alpha : \alpha \in I_{G_0 \cup B}), \\ \theta'_T \stackrel{\text{def}}{=} (S_\alpha^k : \alpha \in I_{G_0 \cup A}, A \subseteq T, |A| = k), \quad \bar{\theta}_B^B \stackrel{\text{def}}{=} (S_\alpha^B : \alpha \in I_{G_0 \cup B}).$$

where $\setminus U_B$ denotes deletion of the array U_B . Then, by Corollary 4.2, η_T and $\bar{\eta}_B$ are conditionally independent given η_B , and by construction $\eta_B \stackrel{d}{=} \theta'_B \stackrel{d}{=} \theta_B^B$, all of them independent from U_{B^c} and ∂U_B . Also, by the first lines of (4.9) and (4.10), we have $\eta_T \stackrel{d}{=} \theta'_T$ and $\bar{\eta}_B \stackrel{d}{=} \bar{\theta}_B^B$. Finally, using the second line of (4.10) for each $\alpha \in I_{G_0 \cup B}$, we can express S_α^B in terms of the S_β^B 's in θ_B^B and U_α in ∂U_B . Thus, for an appropriate F_B , we have

$$\bar{\theta}_B^B = F_B(\theta_B^B, \partial U_B).$$

Also, by similar reasoning using the second equation in (4.9), we get

$$\theta'_T = F(\theta'_B, U_{B^c})$$

for some F .

Now, by Lemma B.2,

$$(\eta_B, \eta_T, \bar{\eta}_B) \stackrel{d}{=} (\theta'_B, F(\theta'_B, U_{B^c}), F_B(\theta'_B, \partial U_B))$$

and in particular

$$(\eta_T, \bar{\eta}_B) \stackrel{d}{=} \left(F(\theta'_B, U_{B^c}), F_B(\theta'_B, \partial U_B) \right) = \left(\theta'_T, F_B(\theta'_B, \partial U_B) \right).$$

Moreover, we have $\perp_{\eta_T} (\bar{\eta}_B)_B$ by Corollary 4.3, and also

$$\perp_{\theta'_T} (F_B(\theta'_B, \partial U_B))_B$$

since $F_B(\theta'_B, \partial U_B)$ is a function of $(\theta'_T, \partial U_B \setminus U_{B^c})$, while $(\partial U_B \setminus U_{B^c})_B$ is an independent family which is also independent from θ'_T . Thus, a slight variation of Lemma B.1 shows that

$$(\eta_T, \bar{\eta}_B)_B \stackrel{d}{=} (\theta'_T, F_B(\theta'_B, \partial U_B))_B.$$

Therefore, we have

$$(4.11) \quad (S_\alpha : \alpha \in I_{G_0 \cup B}, B \subseteq T, |B| = k + 1) \\ \stackrel{d}{=} (S_\alpha^{k+1} : \alpha \in I_{G_0 \cup B}, B \subseteq T, |B| = k + 1)$$

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

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

Thus we have built a $(k + 1)$ -version of (4.9). By inducting up to the level $k = t - 1$ (our top-level inductive hypothesis at level $n - 1$ only allows us to go this far), we obtain the following representation, which involves everything except for the S_α 's for $\alpha \in \mathbb{N}^G$.

$$(4.12) \quad (S_\alpha : \alpha \in I_G \setminus \mathbb{N}^G) \stackrel{d}{=} (S'_\alpha : \alpha \in I_G \setminus \mathbb{N}^G), \\ S'_\alpha \stackrel{\text{def}}{=} g_{Dom(\alpha)} (\mathbf{S}'_{Restr^o(\alpha)}, U_\alpha) \quad \text{for } \alpha \in I_G \setminus \mathbb{N}^G.$$

We have joined all the representations on the proper sub-DAGs. Lemma 4.4, below, is the final piece of the puzzle to complete the top-level induction to get a representation for the whole array $(S_\alpha : \alpha \in I_G)$. Using this lemma (defining $\tilde{\mathbf{S}}$ as in the lemma, and defining $\tilde{\mathbf{S}}'$ similarly), we can complete the proof by showing

$$\left(\tilde{\mathbf{S}}, (f_G(\mathbf{S}_{Restr^o(\alpha)}, U_\alpha) : \alpha \in \mathbb{N}^G) \right) \stackrel{d}{=} \left(\tilde{\mathbf{S}}', (f_G(\mathbf{S}'_{Restr^o(\alpha)}, U_\alpha) : \alpha \in \mathbb{N}^G) \right).$$

This distributional equality holds because $(\tilde{\mathbf{S}}, (U_\alpha : \alpha \in \mathbb{N}^G)) \stackrel{d}{=} (\tilde{\mathbf{S}}', (U_\alpha : \alpha \in \mathbb{N}^G))$, which itself follows from Lemma B.1 instantiated with the following data:

$$\mathcal{F} = \text{the trivial } \sigma\text{-field}, \quad I = \{1, 2\}, \\ (T_1, T_2) = (\tilde{\mathbf{S}}, (U_\alpha : \alpha \in \mathbb{N}^G)), \quad (V_1, V_2) = (\tilde{\mathbf{S}}', (U_\alpha : \alpha \in \mathbb{N}^G)). \quad \square$$

LEMMA 4.4. — *Let*

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

Then, there exists a Borel measurable function f_G such that for any array of independent $[0, 1]$ -uniform random variables $(U_\alpha : \alpha \in \mathbb{N}^G)$ which is independent from \mathbf{S} ,

$$(4.13) \quad \left(\tilde{\mathbf{S}}, (S_\alpha : \alpha \in \mathbb{N}^G)\right) \stackrel{d}{=} \left(\tilde{\mathbf{S}}, (f_G(\mathbf{S}_{\text{Restr}^o(\alpha)}, U_\alpha) : \alpha \in \mathbb{N}^G)\right).$$

This Lemma 4.4 is a consequence of Lemma B.3 and will be proved in the following subsection.

4.1. Proofs of Proposition 4.1 and Lemma 4.4

Recall the following three objects: the finite DAG G , the sequence \mathcal{C} of distinct closed subsets of G , and the \mathcal{C} -type DAG-exchangeable random array collection \mathbb{X} . As in the proof of Proposition 3.8, we will assume that \mathcal{C} is the set \mathcal{A}_G with some fixed ordering.

Let us say that a G -automorphism τ is *separated* if for all $v \in G$, there exists a permutation τ_v on \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 separated G -automorphism. Note that every DAG-exchangeable array collection, including our \mathbb{X} , is automatically separately exchangeable. Therefore, for each (not necessarily closed) subset $H \subseteq G$ and multi-index $\alpha \in \mathbb{N}^H$, we may define $\mathcal{F}_\alpha^{\text{sep}}$ to be the σ -field of all events which are invariant under the actions of all separated α -fixing G -automorphisms.

$$\mathcal{F}_\alpha^{\text{sep}} \stackrel{\text{def}}{=} \sigma\left(\left\{\mathbb{X}^{-1}(B) : B \text{ is Borel, and if } \left((x_\beta : \beta \in \mathbb{N}^C) : C \in \mathcal{C}\right) \in B, \right. \right. \\ \left. \left. \text{then } \left((x_{\tau(\beta)} : \beta \in \mathbb{N}^C) : C \in \mathcal{C}\right) \in B \text{ for all separated } \alpha\text{-fixing } \tau\right\}\right).$$

It is important to remember that the domain $H \subset G$ of $\alpha \in \mathbb{N}^H$ here is not necessarily closed. We will use the letter H below to denote such a general subset of G , while continuing our convention that C and D denote closed subsets. Since a Borel set B on the right-hand side above has less G -automorphisms it has to be invariant with respect to, compared to the definition of \mathcal{F}_α in (3.2), it follows that $\mathcal{F}_\alpha \subset \mathcal{F}_\alpha^{\text{sep}}$.

The missing ingredient, common to the proofs of both Proposition 4.1 and Lemma 4.4, is the following conditional independence result which appears as Corollary 5.6 in the celebrated paper of Hoover [Hoo79]:

PROPOSITION 4.5. — *Define $\mathcal{F}_\alpha^{\text{sep}}$ as above. Let $I_1, I_2, I_3 \subseteq \bigcup_{H \subseteq G} \mathbb{N}^H$ be such that, for all $\alpha_1 \in I_1$ and $\alpha_2 \in I_2$, we have $\alpha_1 \cap \alpha_2 \in I_3$. Then, $(\mathcal{F}_\alpha^{\text{sep}} : \alpha \in I_1)$ is conditionally independent from $(\mathcal{F}_\alpha^{\text{sep}} : \alpha \in I_2)$ given $(\mathcal{F}_\alpha^{\text{sep}} : \alpha \in I_3)$.*

In the proposition, $\alpha_1 \cap \alpha_2$ means the restriction of α_1 to the set of vertices that get mapped to the same values by α_1 and α_2 .

4.1.1. Overview of the proof of Proposition 4.1

First note that if we replaced \mathcal{F}_C (respectively for the C_i and $C \cap C_i$) in Proposition 4.1 by

$$(4.14) \quad \sigma(\{\mathcal{F}_\alpha^{sep} : \alpha \in I\}), \quad I = \bigcup_{H \subseteq C} \mathbb{N}^H$$

(respectively for C_i and $C \cap C_i$), then the result would immediately follow from Proposition 4.5. While the σ -fields related to DAG-exchangeability in Proposition 4.1 are not of the type in (4.14), using the structure of I_C and the fact that C is closed, it is possible to express \mathcal{F}_C as

$$(4.15) \quad \mathcal{F}_C = \sigma(\{\mathcal{F}_\alpha^{sep} : \alpha \in I_C\})$$

(similarly for C_i and $C \cap C_i$). This is established in Lemma 4.6 below. A final order of business required to employ Proposition 4.5 is a sort of converse: we also need to express (4.14) in terms of σ -fields related to DAG-exchangeability since these are the σ -fields that one conditions on in Proposition 4.5. This will be done in Lemma 4.7.

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

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

so that henceforth

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

which we view as an array of arrays. Using this notation, we may rewrite (3.2) as

$$(4.16) \quad \mathcal{F}_\alpha = \sigma\left(\left\{\mathbf{X}^{-1}(B) : B \text{ is Borel, and if } (x_\beta : \beta \in \mathbb{N}^G) \in B, \right. \right. \\ \left. \left. \text{then } (x_{\tau(\beta)} : \beta \in \mathbb{N}^G) \in B \text{ for all } \alpha\text{-fixing } \tau \right\}\right)$$

Let us point out that an α -fixing G -automorphism τ , acting on \mathbf{X} , fixes $\alpha|_C$ for $C \in \mathcal{A}_{\text{Dom}(\alpha)}$, but does not necessarily fix $\alpha|_H$ for any arbitrary (non-closed) subset H . We assume that \mathcal{F}_α^{sep} is rewritten similarly as a σ -field defined in terms of \mathbf{X} , instead of \mathbb{X} .

For a subset $H \subseteq G$, not necessarily closed, let H^o denote the largest subset of H which is closed in G . The closed graph H^o is well-defined since a union of closed subsets is again closed. For example, consider the case that the vertex and edge sets of G are $\{v_1, v_2, v_3, v_4\}$ and $\{\overrightarrow{v_1v_2}, \overrightarrow{v_2v_3}, \overrightarrow{v_3v_4}\}$, respectively. If $H_1 = \{v_1, v_3\}$ and $H_2 = \{v_2, v_3, v_4\}$, then $H_1^o = \{v_1\}$ and $H_2^o = \emptyset$.

The *closure* of a subset $H \subseteq G$ is the smallest closed subset containing H .

LEMMA 4.6. — *Let $\alpha \in I_G$. Then,*

$$\mathcal{F}_\alpha = \bigcap_{n \geq 1} \mathcal{F}_\alpha^n = \bigcap_{n \geq 1} \mathcal{F}_\alpha^{sep, n},$$

where

$$\mathcal{F}_\alpha^n \stackrel{\text{def}}{=} \sigma \left(\left\{ \mathbf{X}_\beta : \beta \in \mathbb{N}^G \text{ and} \right. \right. \\ \left. \left. \text{there is } C \in \mathcal{A}_{\text{Dom}(\alpha)} \text{ s.t. } \beta|_C = \alpha|_C \text{ and } \beta(v) > n \text{ for all } v \notin C \right\} \right)$$

and

$$\mathcal{F}_\alpha^{\text{sep},n} \stackrel{\text{def}}{=} \sigma \left(\left\{ \mathbf{X}_\beta : \beta \in \mathbb{N}^G \text{ and} \right. \right. \\ \left. \left. \text{there is } H \subseteq \text{Dom}(\alpha) \text{ s.t. } \beta|_H = \alpha|_H \text{ and } \beta(v) > n \text{ for all } v \notin H \right\} \right).$$

In particular, for all $\alpha \in I_G$,

$$\mathcal{F}_\alpha^{\text{sep}} = \mathcal{F}_\alpha.$$

Proof. — Pick $\alpha \in I_G$. Define an injection $\tau_k : \mathbb{N} \rightarrow \mathbb{N}$, 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 on \mathbb{N}^G such that $\rho_1(\beta)(v) = \tau_{\alpha(v)}(\beta(v))$ if $v \in \text{Dom}(\alpha)$, and $\rho_1(\beta)(v) = \beta(v) + 1$ otherwise.

The injection ρ_1 can be made to act on any \mathbf{X} -measurable Y . This is because such Y is equal to $f(\mathbf{X})$ almost surely for some measurable f and we can define $\rho_1(Y)$ to be $f(\rho_1(\mathbf{X}))$.⁽⁴⁾ The choice of f does not matter here for the following reason. The DAG-exchangeability of \mathbf{X} implies $\rho_1(\mathbf{X}) \stackrel{d}{=} \mathbf{X}$. Thus,

$$\left(\mathbf{X}, f(\mathbf{X}), g(\mathbf{X}) \right) \stackrel{d}{=} \left(\rho_1(\mathbf{X}), f(\rho_1(\mathbf{X})), g(\rho_1(\mathbf{X})) \right)$$

for all measurable f and g . This in turn implies that whenever $f(\mathbf{X}) = g(\mathbf{X})$ almost surely, we also have almost sure equality between $f(\rho_1(\mathbf{X}))$ and $g(\rho_1(\mathbf{X}))$.

We can regard each $E \in \mathcal{F}_\alpha$ as an \mathbf{X} -measurable random variable $\mathbf{1}_E$ and apply ρ_1 to it. The outcome $\rho_1(E)$ of this application is the same as E . Thus, for all $E \in \mathcal{F}_\alpha$, we have that $\rho_1^n(E) = E$ almost surely for all n . Meanwhile, by construction, $\rho_1^n(F) \in \mathcal{F}_\alpha^{\text{sep},n}$ for any event F . Thus, we obtain $\mathcal{F}_\alpha \subseteq \mathcal{F}_\alpha^{\text{sep},n}$, which implies the inclusion $\mathcal{F}_\alpha \subseteq \bigcap_{n \geq 1} \mathcal{F}_\alpha^{\text{sep},n}$.

Now we show the other inclusion. Consider the below conditions on G -automorphisms ρ :

- (1) ρ fixes α ;
- (2) $\rho(\beta)(v) = \beta(v)$ for all β and v having some $u \preceq v$ with $\beta(u) > n$.

In the second condition, we use the partial order $u \preceq v$ introduced earlier, which means that there is a path of length possibly zero from the vertex u to the vertex v in G . Let \mathcal{T}_n be the set of G -automorphisms satisfying these two conditions. Then, $\bigcup_{n \in \mathbb{N}} \mathcal{T}_n$ generates all the finite⁽⁵⁾ α -fixing G -automorphisms. We claim that $\mathcal{F}_\alpha^{\text{sep},n}$ is invariant under the action of any $\rho \in \mathcal{T}_n$.

⁽⁴⁾ $\rho_1(\mathbf{X}) = (\mathbf{X}_{\rho_1(\alpha)})_\alpha$.

⁽⁵⁾ A G -automorphism τ is finite if $\tau(\alpha) = \alpha$ for all but finitely many $\alpha \in \mathbb{N}^G$.

To see the claim, fix $H \subseteq \text{Dom}(\alpha)$ and $\beta \in \mathbb{N}^G$ such that $\beta|_H = \alpha|_H$ and $\beta(v) > n$ for all $v \notin H$. Then, $\rho(\beta)(v) = \beta(v)$ for $v \notin H$ or $v \in H^\circ$; see the remark following (4.16). For $v \in H \setminus H^\circ$, there exists $u \notin H$ such that $u \prec v$. Otherwise, the closure of $H^\circ \cup \{v\}$ is in H , contradicting the maximality of H° . But then $\beta(u) > n$ by the choice of β . Thus $\rho(\beta)(v) = \beta(v)$ holds again in this case.

Combined with the remark following Proposition 3.8, the claim implies that $\bigcap_{n \geq 1} \mathcal{F}_\alpha^{sep, n}$ is invariant under all α -fixing G -automorphisms. Therefore, $\bigcap_{n \geq 1} \mathcal{F}_\alpha^{sep, n} \subseteq \mathcal{F}_\alpha$. Together with our proof for the other inclusion, this gives $\mathcal{F}_\alpha = \bigcap_{n \geq 1} \mathcal{F}_\alpha^{sep, n}$, as desired.

The equality $\mathcal{F}_\alpha^{sep} = \bigcap_{n \geq 1} \mathcal{F}_\alpha^{sep, n}$ is a standard fact. It can also be obtained by repeating our argument for $\mathcal{F}_\alpha = \bigcap_{n \geq 1} \mathcal{F}_\alpha^{sep, n}$ for the DAG which has the same vertex set as G , but has no edges. Thus, if $\text{Dom}(\alpha)$ is a closed subset (equivalently, $\alpha \in I_G$), we have that $\mathcal{F}_\alpha^{sep} = \mathcal{F}_\alpha$.

Since $\mathcal{F}_\alpha^n \subseteq \mathcal{F}_\alpha^{sep, n}$, to complete the proof, we just need to show that $\mathcal{F}_\alpha \subseteq \bigcap_{n \geq 1} \mathcal{F}_\alpha^n$. We point out that proving this inclusion is not needed for what we are trying to show in this subsection, namely, Proposition 4.1 and Lemma 4.4. However, we spell out the proof here since $\mathcal{F}_\alpha \subseteq \bigcap_{n \geq 1} \mathcal{F}_\alpha^n$ is a natural statement which may be useful for future related works.

Let n be a natural number large enough that $n > \max_{v \in \text{Dom}(\alpha)} \alpha(v)$. Define $\rho_n(\beta)(v) = \beta(v)$ if $v \in \text{Dom}(\alpha)$ and $\beta(u) = \alpha(u)$ for all $u \preceq v$, and $\rho_n(\beta)(v) = \beta(v) + n$ otherwise. Then, ρ_n is an injection and satisfies the following claim.

CLAIM. — *Let H be a subset of $\text{Dom}(\alpha)$ that is not necessarily closed. Then, for all $\beta \in \mathbb{N}^G$, if $\beta|_H = \alpha|_H$ and $\beta(v) > \alpha(v)$ for all $v \in \text{Dom}(\alpha) \setminus H$, we have $\rho_n(\beta)(v) = \beta(v)$ for $v \in H^\circ$, and $\rho_n(\beta)(v) = \beta(v) + n$ for $v \notin H^\circ$.*

It is easy to see why the claim holds for $v \in H^\circ$ or $v \notin H$. If $v \in H \setminus H^\circ$, there exists $u \in \text{Dom}(\alpha) \setminus H$ such that $u \prec v$; otherwise, the closure of $H^\circ \cup \{v\}$ is in H , contradicting the maximality of H° . Therefore, $\beta(u) \neq \alpha(u)$. This implies that $\rho_n(\beta)(v) = \beta(v) + n$, proving the claim.

By the claim, $\rho_n(F) \in \mathcal{F}_\alpha^n$ for every $F \in \mathcal{F}_\alpha^{sep, n}$. Meanwhile, $\rho_n(E) = E$ for every $E \in \mathcal{F}_\alpha$. Also, $\mathcal{F}_\alpha \subseteq \mathcal{F}_\alpha^{sep, n}$. Thus, $\mathcal{F}_\alpha \subseteq \mathcal{F}_\alpha^n$. \square

LEMMA 4.7. — *Let $\alpha \in \mathbb{N}^H$ and $H \subseteq G$. Define $\alpha^\circ \stackrel{\text{def}}{=} \alpha|_{\text{Dom}(\alpha)^\circ}$. Then, $\mathcal{F}_\alpha^{sep} = \mathcal{F}_{\alpha^\circ}$.*

Proof. — It is clear that $\mathcal{F}_{\alpha^\circ} = \mathcal{F}_{\alpha^\circ}^{sep} \subseteq \mathcal{F}_\alpha^{sep}$. Similarly to the proof of the previous lemma, one can show that $\mathcal{F}_\alpha^{sep, n}$ is invariant under the action of a G -automorphism τ if

- (1) τ fixes α° , and
- (2) $\tau(\beta)(v) = \beta(v)$ for all β and v having some $u \preceq v$ with $\beta(u) > n$.

This shows that $\bigcap_{n \geq 1} \mathcal{F}_\alpha^{sep, n} \subseteq \mathcal{F}_{\alpha^\circ}$. One can show that $\mathcal{F}_\alpha^{sep} = \bigcap_{n \geq 1} \mathcal{F}_\alpha^{sep, n}$ by applying Lemma 4.6 to \mathbf{X} , noting that every DAG-exchangeable array is also a *separately* exchangeable array. \square

Now we are ready to complete the main task of this subsection, namely, the proofs of Proposition 4.1 and Lemma 4.4.

Proof of Proposition 4.1. — As already noted, by Lemma 4.6, we have

$$\mathcal{F}_C = \sigma(\{\mathcal{F}_\alpha^{sep} : \alpha \in I_C\})$$

and

$$\mathcal{F}_{C_i} = \sigma(\{\mathcal{F}_\alpha^{sep} : \alpha \in I_{C_i}\}).$$

Now for any $\alpha_1 \in I_C$ and $\alpha_2 \in I_{C_i}$, we have $Dom(\alpha_1 \cap \alpha_2) \subseteq C \cap C_i$. Thus, by Proposition 4.5,

$$\mathcal{F}_C \underset{\mathcal{G}'}{\perp\!\!\!\perp} \mathcal{F}_{C_1}, \dots, \mathcal{F}_{C_m}$$

where

$$\mathcal{G}' \stackrel{\text{def}}{=} \sigma(\{\mathcal{F}_\alpha^{sep} : Dom(\alpha) \subseteq C \cap C_i \text{ for some } i \leq m\}).$$

However, Lemma 4.6 and Lemma 4.7 imply that

$$\mathcal{G}' = \sigma(\{\mathcal{F}_\alpha : \alpha \in I_{C \cap C_i} \text{ for some } i \leq m\}). \quad \square$$

Proof of Lemma 4.4. — Fix $\alpha \in \mathbb{N}^G$ and $\beta_1, \dots, \beta_r \in I_G$ such that $\beta_k \neq \alpha$. Similar to the proof of Proposition 4.1, we can show that

$$S_\alpha \underset{\mathbf{S}_{Restr^o(\alpha)}}{\perp\!\!\!\perp} S_{\beta_1}, \dots, S_{\beta_r}.$$

Since β_1, \dots, β_r are arbitrary, we have

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

The joint distributional equality (4.13) can now be obtained by applying Lemma B.3 with $\xi_\alpha = S_\alpha$ and $\eta_\alpha = \mathbf{S}_{Restr^o(\alpha)}$. \square

Appendix 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, CT17, CT18]). In the appendix, we prove Theorem 1.5, a simplified version of our more general Theorem 1.5, using a representation theorem of Crane and Towsner [CT17] which is formulated and proved using model-theoretic tools. We deal with this simplified version only here because we have not yet been able to derive the full version from model theoretic results.

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

We restate Theorem 1.5.

THEOREM A.1 (Theorem 1.5). — *If \mathcal{C} is the singleton sequence (G) and $(\mathbf{X}_C : C \in \mathcal{C})$ is DAG-exchangeable, then there exists a measurable function $f : [0, 1]^{A_G} \rightarrow \mathcal{X}$ such that*

$$(1.2) \quad (X_\alpha : \alpha \in \mathbb{N}^G) \stackrel{d}{=} \left(f(U_{\alpha|_C} : C \in \mathcal{A}_G) : \alpha \in \mathbb{N}^G \right)$$

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

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

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

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

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

$$(A.1) \quad (X_\alpha : \alpha \in I) \stackrel{d}{=} (f(U_b : b \in B(\alpha)) : \alpha \in I)$$

where

- $B(\alpha)$ is the set of all anti-chains in $E(\alpha) \stackrel{\text{def}}{=} \{[\alpha]_{R_k} : k \leq n\} \cup \{\{\alpha\}\}$, 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 A.3. — The original theorem [CT18] has an additional condition that \mathcal{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* \mathcal{M} of type n for some natural number n is a tuple (I, R_1, \dots, R_n) of a set I and binary relations $\{R_k\}$ on I . When another structure $\mathcal{N} = (J, S_1, \dots, S_n)$ 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 \mathcal{M} . A common way of generating a substructure is to restrict \mathcal{M} with a subset J_0 of I :

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

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

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

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

Crane and Towsner used a structure $\mathcal{M} = (I, R_1, \dots, R_n)$ 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 \mathcal{M}* or *\mathcal{M} -exchangeable* if for all finite subsets J of I and embeddings $\tau : \mathcal{M}|_J \rightarrow \mathcal{M}$,

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

where $\tau(\mathbf{X}) \stackrel{\text{def}}{=} (X_{\tau(\alpha)} : \alpha \in I)$. 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 $\mathcal{M} = (I, R_1, \dots, R_n)$. More specifically, they impose requirements on the R_k 's, and in doing so, they gauge the \mathcal{M} -exchangeability condition.

DEFINITION A.4. — *The structure \mathcal{M} is ultrahomogeneous if for all finite substructures*

$$\mathcal{N} = (J, S_1, \dots, S_k)$$

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

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 v on \mathbb{Z} . The lack of any integers strictly between 2 and 3 prevents the construction of such an v . The structure $(\mathbb{Q}, <)$ is dense, and does not suffer from this kind of problem. These examples highlight one intuition behind ultrahomogeneity: that \mathcal{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 \mathcal{M} is *orderly*. Many binary relations on the underlying set I of \mathcal{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 \mathcal{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.5. — *A relation R is basic explicit in R_1, \dots, 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, \dots, r_n for binary relations that are interpreted as R_1, \dots, 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, \dots, R_k if it is a Boolean combination of basic explicit relations in R_1, \dots, R_k .

DEFINITION A.6. — 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 $\{D_1, \dots, D_m\}$ of D made out of equivalence classes D_i of R , and permutations π on $[m]$, there exists an automorphism v on \mathcal{M} such that⁽⁶⁾

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

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

$$(A.2) \quad \begin{aligned} (k_1, k_2)[R_1](k'_1, k'_2) &\iff k_1 = k'_1, \\ (k_1, k_2)[R_2](k'_1, k'_2) &\iff k_1 = k'_1 \wedge k_2 = k'_2. \end{aligned}$$

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 v on \mathbb{N}^2 :

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

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.7. — 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 $D_1 \cap D_2 = \emptyset$.

⁽⁶⁾Here m may be the first countable ordinal, in which case π is a permutation on \mathbb{N} .

DEFINITION A.8. — 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, \dots, R_{k-1} ;
- R'_k freely contains R_k ; and
- if an equivalence relation S is explicit in R_1, \dots, 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 (A.2). 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 three relations:

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

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 1.5

Let $G = (V, E)$ be the DAG in Theorem 1.5. 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 \mathcal{M} that has \mathbb{N}^G as its underlying set and satisfies the conditions of Theorem A.2, especially the orderly condition.

LEMMA A.9. — There exists an enumeration of V , $(v_\ell : 1 \leq \ell \leq n)$, so that $V_\ell \stackrel{\text{def}}{=} \{v_1, \dots, v_\ell\}$ 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 $V_\ell \stackrel{\text{def}}{=} \{v_1, \dots, v_\ell\}$ 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 $\{v_1, \dots, v_{\ell+1}\}$ is closed, as required. \square

From now on, we write $\mathcal{M} = (I, R_{v_1}, \dots, R_{v_n})$, where the v_k are enumerated as in Lemma A.9 and R_v is defined by

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

LEMMA A.10. — \mathcal{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 $n \leq m - 1$. We will prove the claim for the case that $n = m$.

Let $\mathcal{N} = (J, S_1, \dots, S_m)$ be a substructure of \mathcal{M} , and τ an embedding from \mathcal{N} to \mathcal{M} . Because of the way that we constructed the enumeration $(v_\ell : 1 \leq \ell \leq m)$, 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

$$I|_W \stackrel{\text{def}}{=} \mathbb{N}^W, \quad R_{v_\ell}|_W \stackrel{\text{def}}{=} \{(\alpha|_W, \alpha'|_W) : (\alpha, \alpha') \in R_{v_\ell}\},$$

$$\mathcal{M}|_W \stackrel{\text{def}}{=} (I|_W, R_{v_1}|_W, \dots, R_{v_{m-1}}|_W), \quad \mathcal{N}|_W \stackrel{\text{def}}{=} (J|_W, S_1|_W, \dots, S_{m-1}|_W),$$

Then, $\mathcal{N}|_W$ is a finite substructure of $\mathcal{M}|_W$. Furthermore, there exists a function $\tau_0 : J|_W \rightarrow I|_W$ such that $\tau_0(\beta) = \tau(\beta')|_W$ whenever $\beta = \beta'|_W$. In fact, the function τ_0 is an embedding from $\mathcal{N}|_W$ to $\mathcal{M}|_W$. By induction hypothesis, τ_0 can be extended to an automorphism v_0 on $\mathcal{M}|_W$.

We now extend v_0 to an automorphism on \mathcal{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 $v : I \rightarrow I$ as follows:

$$v(\alpha)(v_k) \stackrel{\text{def}}{=} \begin{cases} v_0(\alpha|_W)(v_k) & \text{if } k \leq m - 1, \\ \alpha(v_k) & \text{if } k = m \text{ and } \alpha \notin J, \\ \pi_{\alpha|_W}(\alpha(v_k)) & \text{if } k = m \text{ and } \alpha \in J. \end{cases}$$

Then, v is the desired extension of τ . □

LEMMA A.11. — The sequence R_{v_1}, \dots, R_{v_n} is orderly.

Proof. — For each $k \leq n$, define

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

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

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

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 1.5. — 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} :

$$(A.3) \quad \mathbf{X} = \left(X_\alpha : \alpha \in \mathbb{N}^G \right) \stackrel{d}{=} \left(f(U_b : b \in B(\alpha)) : \alpha \in \mathbb{N}^G \right)$$

where $B(\alpha)$ is the set of all anti-chains in $E(\alpha) \stackrel{\text{def}}{=} \{[\alpha]_{R_k} : k \leq n\} \cup \{\{\alpha\}\}$, 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 (A.3) to the claimed representation of Theorem 1.5. A crucial part of this translation is the following function φ from $B \stackrel{\text{def}}{=} \bigcup \{B(\alpha) : \alpha \in \mathbb{N}^G\}$ to $J \stackrel{\text{def}}{=} \{\alpha|_C : C \in \mathcal{A}_G \text{ and } \alpha \in \mathbb{N}^G\}$:

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

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 $C \in \mathcal{A}_G$ 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 $\{\alpha|_C : C \in \mathcal{A}_G\}$ 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 ,

$$\begin{aligned} \varphi(b) = \varphi(b') &\implies \left(b = b' \text{ or } \{b, b'\} \right. \\ &\quad \left. = \left\{ \{\{\alpha\}\}, \{[\alpha]_{R_v} : v \in M\} \right\} \text{ for some } \alpha \in \mathbb{N}^G \right) \end{aligned}$$

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

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

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

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

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

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

Then,

$$\begin{aligned} &\left((U_b : b \in B \setminus B_0), (U_b, U_{b'} : b = \{\{\alpha\}\} \text{ and} \right. \\ &\quad \left. b' = \{[\alpha]_{R_v} : v \in M\} \text{ for some } \alpha \in \mathbb{N}^G) \right) \\ &\quad \stackrel{d}{=} \left((U'_{\varphi(b)} : b \in B \setminus B_0), (g(U'_\alpha) : \alpha \in \mathbb{N}^G) \right) \end{aligned}$$

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

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

which implies

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

as desired. □

Appendix B. Supplementary Results

We will sometimes write $(\xi_a)_{a \in I}$ to mean a family of random variables, and also refer to such a family as an *array*. Also, we will use the following notation for conditional distribution properties.

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

The first Lemma B.1 is a standard result from probability theory whose proof we omit.

LEMMA B.1. — Let $(T_a, V_a)_{a \in I}$ be a multi-indexed family of random variables, and let \mathcal{F} be a σ -field. Assume the following hold:

- $\mathbf{P}[T_a \in \cdot | \mathcal{F}] = \mathbf{P}[V_a \in \cdot | \mathcal{F}]$ almost surely
- $\perp_{\mathcal{F}} (T_a)_{a \in I}$
- $\perp_{\mathcal{F}} (V_a)_{a \in I}$

Then, $\mathbf{P}[(T_a)_a \in \cdot | \mathcal{F}] = \mathbf{P}[(V_a)_a \in \cdot | \mathcal{F}]$ almost surely, and consequently, $(T_a)_a \stackrel{d}{=} (V_a)_a$.

The next Lemma B.2, which is a simple application of the previous result, is used to synchronize representations using different functions.

LEMMA B.2. — Let $\xi_0, (\xi_a)_{a \in I}$ be random variables such that $\xi_0 \stackrel{d}{=} \xi_a$, and let $\zeta = (\zeta_a)_{a \in I}$ 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 $a \in I$
- $\perp_{\xi_0} (\eta_a)_{a \in I}$

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 ζ .

Proof. — Since $(\xi_0, \zeta_a) \stackrel{d}{=} (\xi_a, \zeta_a)$, we can replace ξ_a by ξ_0 in the first bullet. Let $T_a = (\xi_0, \eta_a)$, $V_a = (\xi_0, \phi_a(\xi_0, \eta_a))$, $\mathcal{F} = \sigma(\xi_0)$. Then, $(T_a, V_a)_a$ and \mathcal{F} satisfy the conditions of Lemma B.1. The desired result immediately follows. \square

The following coding Lemma B.3 can be found in [Kal05, Lemma 7.6]:

LEMMA B.3. — 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 $a, b \in A$.
- $\xi_a \perp_{\eta_a} (\xi_b)_{b \neq a}, \boldsymbol{\eta}$ for all $a \in A$.

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

Acknowledgments

We would like to thank Henry Towsner for correspondence explaining his work with Crane and for giving us insightful feedback. About five years ago, Benjamin Bloem-Reddy suggested to Staton and Yang to study the paper by Austin and Panchenko, which became the starting point of our work. We are grateful to him for his excellent suggestion. We thank Tim Austin, Cameron Freer, Daniel Roy, and Daeseok Lee for helpful discussions and comments on the paper.

BIBLIOGRAPHY

- [AAF⁺18] Nathanael Leedom Ackerman, Jeremy Avigad, Cameron E. Freer, Daniel M. Roy, and Jason M. Rute, *On the computability of graphons*, <https://arxiv.org/abs/1801.10387>, 2018. ↑332
- [Ack15] Nathanael L. Ackerman, *Representations of $\text{Aut}(M)$ -invariant measures: Part I*, <https://arxiv.org/abs/1509.06170>, 2015. ↑357
- [AFP16] Nathanael L. Ackerman, Cameron E. Freer, and Rehana Patel, *Invariant measures concentrated on countable structures*, Forum Math. Sigma **4** (2016), article no. 17. ↑357
- [Ald81] David J. Aldous, *Representations for partially exchangeable arrays of random variables*, J. Multivariate Anal. **11** (1981), no. 4, 581–598. ↑326, 329, 344
- [Ald85] ———, *Exchangeability and related topics*, École d'Été de Probabilités de Saint-Flour XIII 1983, Lecture Notes in Mathematics, Springer, 1985, pp. 1–198. ↑326, 328
- [Ald10] ———, *More uses of exchangeability: representations of complex random structures*, Probability and mathematical genetics. Papers in honour of Sir John Kingman, London Mathematical Society Lecture Note Series, vol. 378, Cambridge University Press, 2010, pp. 35–63. ↑329
- [AP14] Tim D. Austin and Dmitry Panchenko, *A hierarchical version of the de Finetti and Aldous–Hoover representations*, Probab. Theory Relat. Fields **159** (2014), no. 3–4, 809–823. ↑326, 329, 333, 334, 340, 343
- [Aus08] Tim Austin, *On exchangeable random variables and the statistics of large graphs and hypergraphs*, Probab. Surv. **5** (2008), 80–145. ↑329
- [Aus12] ———, *Exchangeable random arrays*, 2012, Notes for IAS workshop, <https://www.math.ucla.edu/~tim/ExchnotesforIISc.pdf>. ↑326, 329
- [BC12] Charles Bordenave and Djalil Chafaï, *Around the circular law*, Probab. Surv. **9** (2012), 1–89. ↑334
- [BCC11] Charles Bordenave, Pietro Caputo, and Djalil Chafaï, *Spectrum of non-hermitian heavy tailed random matrices*, Commun. Math. Phys. **307** (2011), no. 2, 513–560. ↑334
- [BR97] Ola Bratteli and Derek W. Robinson, *Operator Algebras and Quantum Statistical Mechanics. Vol. 2: Equilibrium States Models in Quantum Statistical Mechanics*, 2nd ed., Texts and Monographs in Physics, Springer, 1997. ↑334
- [CCB16] Diana Cai, Trevor Campbell, and Tamara Broderick, *Edge-exchangeable graphs and sparsity*, Advances in Neural Information Processing Systems 29 (NIPS 2016), 2016, pp. 4249–4257. ↑329
- [CD18] Harry Crane and Walter Dempsey, *Edge exchangeable models for interaction networks*, J. Amer. Statist. Assoc. **113** (2018), no. 523, 1311–1326. ↑329, 334
- [CF17] François Caron and Emily B. Fox, *Sparse graphs using exchangeable random measures*, J. R. Stat. Soc. Ser. B. Stat. Methodol. **79** (2017), no. 5, 1295–1366. ↑329
- [CGZ15] Yu Lu Chao Gao and Harrison H. Zhou, *Rate-optimal graphon estimation*, Ann. Stat. **43** (2015), no. 6, 2625–2652. ↑332
- [CT17] Harry Crane and Henry Towsner, *Relative exchangeability with equivalence relations*, Arch. Math. Logic **57** (2017), no. 5–6, 1–24. ↑329, 334, 343, 357, 358
- [CT18] ———, *Relatively exchangeable structures*, J. Symb. Log. **83** (2018), no. 2, 416–442. ↑329, 357, 358
- [DJ08] Persi W. Diaconis and Svante Janson, *Graph limits and exchangeable random graphs*, Rend. Mat. Appl. **28** (2008), no. 1, 33–61. ↑329

- [Fin75] Bruno de Finetti, *Theory of probability. a critical introductory treatment. vols 1 and 2*, Wiley Series in Probability and Mathematical Statistics, John Wiley & Sons, 1974-1975, English translation by Antonio Machi and Adrian Smith. ↑328
- [FR12] Cameron E. Freer and Daniel M. Roy, *Computable de Finetti measures*, Ann. Pure Appl. Logic **163** (2012), no. 5, 530–546. ↑332
- [GMR⁺08] Noah D. Goodman, Vikash Mansinghka, Daniel M. Roy, Keith Bonawitz, and Joshua B. Tenenbaum, *Church: a language for generative models*, Proceedings of the Twenty-Fourth Conference on Uncertainty in Artificial Intelligence (2008) (David McAllester and Petri Myllymaki, eds.), AUAI Press, 2008. ↑330
- [GvdV17] Subhashis Ghosal and Aad van der Vaart, *Fundamentals of nonparametric Bayesian inference*, Cambridge Series in Statistical and Probabilistic Mathematics, vol. 44, Cambridge University Press, 2017. ↑330
- [Hoo79] Douglas N. Hoover, *Relations on probability spaces and arrays of random variables*, 1979, Preprint. ↑326, 329, 353
- [Kal02] Olav Kallenberg, *Foundations of modern probability*, 2nd ed., Probability and Its Applications, Springer, 2002. ↑344
- [Kal05] ———, *Probabilistic Symmetries and Invariance Principles*, Probability and Its Applications, Springer, 2005. ↑326, 329, 343, 344, 365
- [KTG⁺06] Charles Kemp, Joshua B. Tenenbaum, Thomas L. Griffiths, Takeshi Yamada, and Naonori Ueda, *Learning Systems of Concepts with an Infinite Relational Model*, Proceedings of the Twenty-First National Conference on Artificial Intelligence (AAAI-06), AAAI Press, 2006, pp. 381–388. ↑330
- [MMYW16] David T. Meent, Jan-Willem van de Meent, Hongseok Yang, and Frank D. Wood, *Design and Implementation of Probabilistic Programming Language Anglican*, IFL 2016: Proceedings of the 28th Symposium on the Implementation and Application of Functional Programming Languages (Tom Schrijvers, ed.), Association for Computing Machinery, 2016, pp. 1–12. ↑330
- [MP01] Marc Mézard and Giorgio Parisi, *The Bethe lattice spin glass revisited*, Eur. Phys. J. B Condens. Matter Phys. **20** (2001), no. 2, 217–233. ↑326
- [OR15] Peter Orbanz and Daniel M. Roy, *Bayesian Models of Graphs, Arrays and Other Exchangeable Random Structures*, IEEE Trans. Pattern Anal. Mach. Intell. **37** (2015), no. 2, 437–461. ↑330, 331
- [Pan15] Dmitry Panchenko, *Hierarchical exchangeability of pure states in mean field spin glass models*, Probab. Theory Relat. Fields **161** (2015), no. 3-4, 619–650. ↑326
- [Res14] Sidney Ira Resnick, *A probability path*, reprint of the 2005 edition ed., Modern Birkhäuser Classics, Springer, 2014. ↑345
- [SSY⁺18] Sam Staton, Dario Stein, Hongseok Yang, Nathanael L. Ackerman, Cameron Freer, and Daniel M. Roy, *The Beta–Bernoulli Process and Algebraic Effects*, 45th International Colloquium on Automata, Languages, and Programming: ICALP 2018 (Ioannis Chatzigiannakis, Christos Kaklamanis, Marx Dániel, and Donald Sanella, eds.), Leibniz International Proceedings in Informatics, vol. 107, Schloss Dagstuhl, 2018. ↑332
- [SYA⁺17] Sam Staton, Hongseok Yang, Nathanael L. Ackerman, Cameron Freer, and Daniel M. Roy, *Exchangeable random process and data abstraction*, 2017, Workshop on Probabilistic Programming Semantics (PPS 2017), available at <https://pps2017.luddy.indiana.edu/2017/01/02/exchangeable-random-processes-and-data-abstraction/>. ↑330, 332
- [VR15] Victor Veitch and Daniel M. Roy, *The Class of Random Graphs Arising from Exchangeable Random Measures*, <https://arxiv.org/abs/1512.03099>, 2015. ↑329

- [WMM14] Frank D. Wood, Jan Willem van de Meent, and Vikash Mansinghka, *A New Approach to Probabilistic Programming Inference*, Proceedings of the Seventeenth International Conference on Artificial Intelligence and Statistics (Samuel Kaski and Jukka Corander, eds.), Proceedings of Machine Learning Research, vol. 33, PMLR, 2014, pp. 1024–1032. ↑330
- [XTYK06] Zhao Xu, Volker Tresp, Kai Yu, and Hans-Peter Kriegel, *Infinite hidden relational models*, Proceedings of the Twenty-Second Conference on Uncertainty in Artificial Intelligence (2006) (Rina Dechter and Thomas S. Richardson, eds.), AUAI Press, 2006, pp. 544–551. ↑330

Manuscript received on 15th December 2018,
revised on 20th September 2019,
accepted on 15th April 2020.

Recommended by Editor D. Aldous.
Published under license CC BY 4.0.



This journal is a member of Centre Mersenne.



Paul JUNG
Department of Mathematical Sciences, KAIST,
291 Daehak-ro, Yuseong-gu,
Daejeon, (Republic of Korea)
pauljung@kaist.ac.kr

Jiho LEE
Department of Mathematical Sciences, KAIST,
291 Daehak-ro, Yuseong-gu,
Daejeon, (Republic of Korea)
efidiaf@gmail.com

Sam STATON
Department of Computer Science,
University of Oxford,
Parks Road,
Oxford, OX1 3QD, (United Kingdom)
sam.staton@cs.ox.ac.uk

Hongseok YANG
School of Computing, KAIST,
291 Daehak-ro, Yuseong-gu,
Daejeon, (Republic of Korea)
hongseok.yang@kaist.ac.kr