A de Finetti-type representation of joint hierarchically exchangeable arrays on directed acyclic graphs

Jiho Lee

Department of Mathematical Sciences, KAIST

June 24, 2020

Abstract

We extend the notion of DAG (directed acyclic graph) exchangeability introduced by Jung, L., Staton, Yang (2018) to a wider class of exchangeable structures, which includes jointly exchangeable arrays. We prove a canonical de Finetti-type representation of such arrays. Similar to that work, the inductive proof constructs an associated symmetry array to a given DAG-exchangeable array which encodes the dependence structure corresponding to the hierarchy structure of the DAG. Our proof adapts Kallenberg’s methods to provide a completely probabilistic proof, whereas the proof of the previous result relied on a result of Hoover which uses nonstandard analysis and symbolic logic.

Key words: Exchangeability, hierarchical exchangeability, DAG exchangeability, de Finetti-type representation, joint exchangeability

1 Introduction

DAG exchangeability is a notion of exchangeability on a family of indexed random elements

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

on a Borel space $\mathcal{X}$, where $G = (V, E)$ is a directed acyclic graphs (DAGs). DAG exchangeability was introduced by [JLSY18] as a generalization of hierarchical exchangeability in [AP14]. The main purpose of this paper is to extend [JLSY18] to a wider class of exchangeable structures including jointly exchangeable arrays using probabilistic methods. These methods were first deployed by David Aldous in [Ald81]. Later, Olav Kallenberg applied this method in a systematic way for more general results (see [Kal89] or [Kal92] for example). All of these results are organized in his textbook [Kal05] which is a standard reference for fundamental results in exchangeability.
Our work is motivated by studies on Bayesian inference modeling, probabilistic programming, and neural networks as discussed in the introduction to [JLSY18]. In fact, the original motivation and hope in that work was to obtain a representation for jointly DAG exchangeable arrays as opposed to the representation obtained there for separately DAG exchangeable arrays (precise definitions are given later). In this work, we close this gap by providing such a representation. Briefly, the idea is that de Finetti-type representations of hierarchically exchangeable structures can identify when a hierarchical generative model can be replaced by an equivalent one but with more explicit independence structure (see [SYA+17]). One can also find in [ORT14] a recent survey on various applications of exchangeability theory to Bayesian inference models including [Hof08], [FP+12], and [LOGR12]. Structure theorems on exchangeable processes also provide canonical representations of neural networks with hierarchical symmetries. Readers can consult, for example, [BRT19], [CW16], and [BZSL13] for applications in this direction.

Let $G = (V, E)$ be a DAG. We assume for the rest of the paper that $G$ is finite and simple. Also, when we write $G$ as a set, we refer to the set of vertices $V$. We write $v \prec w$ if there exists a directed nonempty path from $v$ to $w$. Note that $\prec$ defines a partial order in $G$. Conversely, given a finite partially ordered set $(G, \prec)$, we can build a corresponding set of directed edges $E$ by adding the edge $\overline{vw}$ if and only if $v \prec w$ and there is no $v' \in G$ such that $v \prec v' \prec w$. To make this correspondence bijective, we assume that $G$ always have the minimal set of edges under its induced partial order: that is, whenever there is a directed path from $v$ to $w$ that passes other vertices than $v$ and $w$, we have no edge from $v$ to $w$. \footnote{This restriction is necessary in Definition 2.3 (see its preceding paragraph).}

We say that a subgraph $C$ of a DAG $G$ is \textbf{downward-closed} (or just closed) if it is downward-closed under the induced partial order, that is, $v \in C$ whenever there exists $w \in C$ such that $v \prec w$. We write $\mathcal{A}_C$ for the collection of all closed subsets of $C$. For $\alpha \in \mathbb{N}^G$ and $C \in \mathcal{A}_G$, let $\alpha|_C$ denote the restriction of $\alpha$ to $C$ when viewing $\alpha$ as a function from $G$ to $\mathbb{N}$.

**Definition 1.1.** Let $G$ be a DAG. Then, a $G$-\textbf{permutation} is a bijection $\tau : \mathbb{N}^G \to \mathbb{N}^G$ such that

$$\alpha|_C = \beta|_C \iff \tau(\alpha)|_C = \tau(\beta)|_C$$

for all $\alpha, \beta \in \mathbb{N}^G$, $C \in \mathcal{A}_G$. We write $S_N^G$ for the collection of all $G$-permutations. A random array $X = (X_\alpha : \alpha \in \mathbb{N}^G)$ is \textbf{DAG-exchangeable} if for all $\tau \in S_N^G$, we have

$$(X_\alpha : \alpha \in \mathbb{N}^G) \overset{d}{=} (X_{\tau(\alpha)} : \alpha \in \mathbb{N}^G).$$

\footnote{Although in [JLSY18] we used the word “automorphism,” we change the terminology in order to distinguish them with automorphisms of the DAG itself. (See Section 2.)}
1 INTRODUCTION

For any DAG-exchangeable array, we have a canonical representation using independent uniform random variables, as long as the underlying probability space is rich enough, and we will assume this for the rest of the paper.

**Theorem 1.2** ([JLSY18]). Let $G$ be a DAG. Let $X = (X_\alpha : \alpha \in \mathbb{N}^G)$ be a DAG-exchangeable array taking values in a Borel space $\mathcal{X}$. Then, there exist a measurable function $f : [0, 1]^{|G|} \to \mathcal{X}$ and an i.i.d. array $U = (U_\beta : \beta \in I_G)$ of uniform random variables such that

$$X_\alpha = f\left(U_\beta : \beta \in \text{Restr}(\alpha)\right)$$

almost surely for all $\alpha \in \mathbb{N}^G$.

**Example 1.3.** This setup covers the following past results on the representations of exchangeable structures by independent uniform random variables.

(a) Exchangeable sequences: Let $G$ be a graph with a single vertex. Then, $S_{\mathbb{N}}^G$ is simply the group of all bijections from $\mathbb{N}$ to itself. So, a DAG-exchangeable array is merely an exchangeable sequence. Theorem 1.2 implies that for an exchangeable sequence $X = (X_n : n \in \mathbb{N})$, there exist an i.i.d. sequence of uniform random variables $(U_0, U_1, U_2, \ldots)$ and a measurable function $f : [0, 1]^2 \to \mathcal{X}$ such that

$$X_n = f(U_0, U_n)$$

almost surely for all $n \in \mathbb{N}$. This is a variant of de Finetti’s theorem ([DF29], [DF37], [HS55]) proposed by [Ald81].

(b) Separately exchangeable arrays: Let $G = (\{1, 2\}, \emptyset)$. Then, $\mathbb{N}^G = \mathbb{N}^2$ and $S_{\mathbb{N}}^G$ is isomorphic to $(S_{\mathbb{N}})^2$, acting naturally on $\mathbb{N}^2$. Thus, a DAG-exchangeable array is a separately exchangeable array of dimension 2, that is, it satisfies the distributional equation

$$(X_{ij} : i, j \in \mathbb{N}) \overset{d}{=} (X_{\tau(i)\rho(j)} : i, j \in \mathbb{N}).$$

It is guaranteed by either Theorem 1.2 or the Aldous-Hoover theorem ([Ald81], [Hoo79]) that there exist an i.i.d. array of uniform random variables $U = (U_{ij} : i, j \geq 0)$ and a measurable function $f : [0, 1]^4 \to \mathcal{X}$ such that

$$X_{ij} = f(U_{00}, U_{i0}, U_{0j}, U_{ij})$$

almost surely for all $i, j \in \mathbb{N}$. The result can be extended to arrays of higher dimensions. (See [Kal05] for a deep analysis on exchangeable arrays of high dimensions.)

(c) Hierarchically exchangeable arrays: Let $G = (\{v_1, v_2, u_1, u_2\}, E)$ where $E = \{v_1v_2, u_1u_2\}$. Then, a DAG-exchangeable array is an example of hierarchical exchangeability introduced by [AP14], which can be written in the form $(X_{ij, k\ell} : i, j, k, \ell \in \mathbb{N})$ where $i, j, k, \ell$ are the coordinates on $v_1, v_2, u_1, u_2$, respectively. DAG exchangeability allows that for $\tau, \rho \in S_{\mathbb{N}}$ and $\tau_i, \rho_k \in S_{\mathbb{N}}$ for each $i, k \in \mathbb{N}$, we have

$$(X_{ij, k\ell} : i, j, k, \ell \in \mathbb{N}) \overset{d}{=} (X_{\tau(i)\rho(j), \rho(k)\rho(\ell)} : i, j, k, \ell \in \mathbb{N}).$$
The representation theorem by [AP14] allows us to have

$$X_{ij,k\ell} = f(U_{00,00}, U_{i0,00}, U_{ij,00}, U_{ij,k0}, U_{ij,k\ell}, U_{i0,k\ell}, U_{ij,k\ell})$$

almost surely for some measurable function $f$ and some i.i.d. array $U$ of uniform random variables.

---

**Figure 1:** The DAG for Example 1.4

**Example 1.4.** We introduce random block matrices from Example 2.2 of [JLSY18]. Let $V = \{u, v, r, c\}$, $E = \{\overrightarrow{ur}, \overrightarrow{uc}, \overrightarrow{vr}, \overrightarrow{vc}\}$. (See Figure 1.) An array $(X_{ij,k\ell} : i, j, k, \ell \in \mathbb{N})$ is DAG-exchangeable (regarding $i, j, k, \ell$ as coordinates on $u, v, r, c$, respectively) if for all $\tau, \rho \in S_\mathbb{N}$ and $\tau_{ij}, \rho_{ij} \in S_\mathbb{N}$ with $i, j \in \mathbb{N}$, we have

$$(X_{ij,k\ell} : i, j, k, \ell \in \mathbb{N}) \overset{d}{=} (X_{\tau_{ij}(i)\rho_{ij}(j)\tau_{ij}(k)\rho_{ij}(\ell)} : i, j, k, \ell \in \mathbb{N}). \tag{6}$$

By Theorem 1.2, there exist a measurable function $f$ and an i.i.d. array of uniform random variables $U$ such that for all $i, j, k, \ell \in \mathbb{N}$, we have

$$X_{ij,k\ell} = f(U_{00,00}, U_{i0,00}, U_{ij,00}, U_{ij,k0}, U_{ij,k\ell}, U_{i0,k\ell}, U_{ij,k\ell}) \tag{7}$$

almost surely.

To motivate the main objective of this paper, let us revisit (b) of Example 1.3. In the case where the array is jointly exchangeable, that is,

$$(X_{ij} : i, j \in \mathbb{N}) \overset{d}{=} (X_{\tau_{ij}(i)\tau_{ij}(j)} : i, j \in \mathbb{N}) \tag{8}$$

for all $\tau \in S_\mathbb{N}$, we have a representation of the form

$$X_{ij} = f(U_{0}, U_{i}, U_{j}, U_{\{i,j\}}) \tag{9}$$
almost surely for \( i \neq j \) ([Hoo79]). One can see that, compared to (5), the indices on the rows are merged with those on the columns. We can naturally ask if the similar merging occurs on joint versions of DAG-exchangeable arrays. That is, if a random array \( X = (X_{ij,k\ell} : i,j,k,\ell \in \mathbb{N}) \), for instance, satisfies the distributional equation
\[
(X_{ij,k\ell} : i,j,k,\ell \in \mathbb{N}) \overset{d}{=} (X_{\tau(i)\tau(j),\tau(i,j)(k)\rho(i,j)(\ell)} : i,j,k,\ell \in \mathbb{N})
\]
for all \( \tau, \tau(i,j), \rho(i,j) \in S_N \), we can ask whether we have a representation of the form
\[
X_{ij,k\ell} = f(U_{0,00}, U_{i,00}, U_{j,00}, U_{i,j}, k0, U_{i,j}, 0\ell, U_{i,j}, k\ell)
\]
almost surely for \( i \neq j \).

The main objective of this paper is to extend the representation given by Theorem 1.2 to a wider class of exchangeable structures. This new model includes Hoover’s joint exchangeable arrays, the representation (11), and exchangeable arrays associated to arbitrary DAGs with merging of the vertices. We will rigorously define the model in the next section with more examples.

2 Settings and Main Results

Let \( K \) be a subgroup of the automorphism group \( K \) of \( G \). Let \( Aut(G) \) denote the directed graph automorphism group of \( G \). Define a left group action for \( Aut(G) \) acting on \( I_G \) by
\[
\kappa \beta(v) = \beta(\kappa^{-1}(v)), \kappa \in Aut(G), \beta \in I_G.
\]
We can also define the actions of the objects like \( \tau \kappa \) or \( \kappa \tau \) for \( \tau \in S^G \) as composition of functions. For example, let \( G = \{v,u\} \) with no edges and \( K = \{e,\kappa\} \) with \( \kappa \) the nonidentity element, and let \( \beta(v) = 2, \beta(u) = 3 \), and \( \tau = (1 2 3) (3 2 1) \) \( \in S_N \times S_N \cong S^G_N \) acting naturally on \( \mathbb{N}^G \). Then, we have \( (\tau \kappa)(\beta)(v) = 1, (\tau \kappa)(\beta)(u) = 1 \), while \( (\kappa \tau)(\beta)(v) = \tau \beta(u) = 2, (\kappa \tau)(\beta)(u) = \tau \beta(v) = 3 \).

As we can see in Example 1.3, in the case of a separately exchangeable array of dimension \( d \), we can regard \( G \) as a graph of order \( d \) with no edges. An array \( X = (X_\alpha : \alpha \in \mathbb{N}^d) \) is separately exchangeable if and only if
\[
(X_\alpha : \alpha \in \mathbb{N}^d) \overset{d}{=} (X_{\tau \alpha} : \alpha \in \mathbb{N}^d)
\]
for all \( \tau \in S_N^d \). We can describe this in terms of commutativity with the automorphism group of the graph \( G \), which is isomorphic to \( S_d \). That is, \( X \) is jointly exchangeable if and only if (12) holds for all \( \tau \in S_N^d \) that commutes with \( S_d \).

If we require that \( \tau \) commutes with \( \kappa \in K \) for subgroups \( K \) of \( S_d \) instead of the whole \( S_d \), we obtain a different notion of exchangeability between separate and joint exchangeability. For example, if we let

\[\text{Diagonal indices are excluded because we cannot map diagonals to non-diagonals via the action of } S_d. \text{ Likewise, we will exclude some indices in the setting of joint exchangeability on DAGs to guarantee that the group action on the indices is transitive.}\]
2 SETTINGS AND MAIN RESULTS

Figure 2: A diagram associated to permutations commuting with graph automorphisms. \( \tau \in S_N^{\times G} \) commutes with all \( \kappa \in K \subseteq Aut(G) \).

\[ d = 3 \text{ and } K \text{ be the subgroup of } S_3 \text{ generated by } (2, 3), \]

\[ (X_{ijk} : i, j, k \in \mathbb{N}^3^d) = (X_{\tau(i)\sigma(j)\sigma(k)} : i, j, k \in \mathbb{N}^3) \]

for all \( \tau, \sigma \in S_N \).

Considering the above observations, it is tempting to define joint exchangeability on random arrays defined on DAGs by assigning a subgroup \( K \) of \( Aut(G) \) and allowing law-invariance for permutations which commute with \( K \) (see Figure 2). However, there are a few issues we have to handle. One is that the exchangeability structure does not uniquely determine the group \( K \).

**Example 2.1.** Let \( G = \{v_1, v_2, v_3\} \) be a graph with no edges, and let \( K = \mathbb{Z}/3\mathbb{Z} \) acting naturally on \( G \). Then, since \( K \) acts transitively on \( G \), a permutation \( \tau = (\tau_1, \tau_2, \tau_3) \in S_3^3 \) of \( \mathbb{N}^G \) commutes with \( K \) if and only if \( \tau_1 = \tau_2 = \tau_3 \). So we obtain the same exchangeability structure in this setting if we choose either \( K = \mathbb{Z}/3\mathbb{Z} \) or \( K = S_3 \).

The other issue is more serious. Many of the proofs of representation theorems on exchangeable arrays use induction on the dimension of the arrays, and we will also follow this strategy. However, by restricting \( K \) to be a subgroup of \( Aut(G) \), we encounter an issue when deploying this type of induction, as the following example shows.

**Example 2.2.** Let \( V = \{u_1, u_2, v_1, v_2\} \), \( E = \{u_1u_2, v_1v_2\} \). Then, \( Aut(G) \) is a group of order two, where the nonidentity element exchanges \( u_i \) and \( v_i \) for \( i = 1, 2 \) respectively. However, the closed subgraph \( C = \{u_1, u_2, v_1\} \) has a trivial automorphism group. If we assign joint exchangeability on a random array \( X = (X_\alpha : \alpha \in \mathbb{N}^G) \) associated to \( K = Aut(G) \), the permutations in consideration should act identically on vertices \( u_1 \) and \( v_1 \). However, there is no way to assign such a class of exchangeability on random arrays defined on the subgraph \( C \) via its automorphism group, since it has no nontrivial graph automorphism at all.

Both of these issues arise from the nature that the class of permutations that commutes with \( K \) is determined only by the local behavior of \( K \) in the following sense. Let \( Z_K(S_N^{\times G}) \) denote the group of
$G$-permutations that commute with $K$. Let $C_v$ denote the closure of $\{v\}$, i.e. the smallest $C \in \omega_G$ containing $v$. Then, $Z_K(S^N_G)$ consists of all $\tau \in S^N_G$ such that $\tau \kappa(\alpha) = \kappa \tau(\alpha)$ for all $v \in G, \kappa \in K, \alpha \in N^{C_v}$. In other words, the commuting property of a permutation is determined completely by its behavior on individual vertices.

Now, in order to handle the above issues from the examples, instead of a subgroup of $Aut(G)$, we will use a collection of mappings which takes into account the local nature required by joint exchangeability. We will continue to use $K$ to denote such a collection. These mappings are not necessarily defined on the whole of $G$, but only on a specific vertex and its closure. We use the word **isomorphism** to describe a bijective function from a closed subgraph of $G$ to another closed subgraph that preserves directed edges. (Note that we have assumed that $G$ has no redundant edges.)

**Definition 2.3.** Let $G$ be a DAG. A **local isomorphism** of $G$ is a sub-DAG isomorphism on $C_v$,

$$\kappa : C_v \rightarrow C_w,$$

for some $v, w \in G$. A collection $K$ of local isomorphisms is called a **consistent local isomorphism class (CLIC)** of $G$ if

- $K$ contains all the identity mappings and is closed under inversion, composition, and restrictions to subgraphs of the form $C_u$.
- $K$ contains all the trivial extensions. That is, if $\kappa \in K$, then we have $\kappa' \in K$ whenever $\kappa$ is a restriction of $\kappa'$ and $\kappa'(u) = u$ for all $u$ where $\kappa$ is undefined.

**Example 2.4.** Let us go back to the random block matrices in Example 1.4. There are following local isomorphisms along with their inverses and identities:

- $\kappa_{11} : \{u, v, r\} \rightarrow \{u, v, c\}$ where $\kappa_{11}(r) = c$ and $\kappa_{11}(u) = v$.
- $\kappa_{01} : \{u, v, r\} \rightarrow \{u, v, c\}$ where $\kappa_{01}(r) = c$ and $\kappa_{01}(u) = u$.
- $\kappa_{10} : \{u, v, r\} \rightarrow \{u, v, r\}$ where $\kappa_{10}(r) = r$ and $\kappa_{10}(u) = v$.
- $\kappa_1 := \kappa_{10}|_{\{u\}}$.

Note that $\kappa_{10}$ is a trivial extension of $\kappa_1$. We can always identify a local isomorphism with its trivial extensions. A nontrivial CLIC $K_S$ may be generated by any of the following sets $S$:

1. $S = \{\kappa_1\}$ (or equivalently, $S = \{\kappa_1, \kappa_{10}\}$)
2. $S = \{\kappa_{10}\}$
3. $S = \{\kappa_1, \kappa_{11}, \kappa_{01}\}$ (or equivalently, $S = \{\kappa_1, \kappa_{11}, \kappa_{01}, \kappa_{10}\}$)
For another example, let us consider \( G = (V, E) \) with \( V = \{ u_1, u_2, v_1, v_2, v_3 \} \) with edges \( E = \{ u_1 u_2, \overrightarrow{v_1 v_2}, \overrightarrow{v_2 v_3} \} \). It corresponds to Austin and Panchenko’s setting with \( r = 2, d_1 = 2, d_2 = 3 \) (See [AP14]). Although there is no nontrivial automorphism of \( G \), we have the following nontrivial local isomorphisms along with their inverses:

- \( \rho_1 : \{ u_1 \} \rightarrow \{ v_1 \} \).
- \( \rho_2 : \{ u_1, u_2 \} \rightarrow \{ v_1, v_2 \} \) where \( \rho_2(u_i) = v_i \).

A nontrivial CLIC \( K_S \) may be generated by any of the following sets \( S \):

4. \( S = \{ \rho_1 \} \)
5. \( S = \{ \rho_1, \rho_2 \} \)

**Remark.** A local isomorphism need not be extendable to an automorphism. For instance, in the case of Example 2.2, the cause of the described issue is that the local isomorphism \( \kappa : \{ u_1 \} \rightarrow \{ v_1 \} \) cannot be extended to an automorphism of \( C \). Let \( K = \{ id, \kappa, \kappa^{-1}, \rho, \rho^{-1} \} \), where \( \rho : \{ u_1, u_2 \} \rightarrow \{ v_1, v_2 \} \) with \( \rho(u_i) = v_i \) for \( i = 1, 2 \). Then, the (jointly) exchangeable random array associated to the automorphism group of \( G \) is law-invariant under the permutations that commute “locally” with \( K \). Unlike the case using automorphisms, the induced symmetry on the subgraph \( C \) is well-described by just taking the elements in \( K \) which is defined inside \( C \), which are \( id, \kappa, \) and \( \kappa^{-1} \).

Let \( K \) be a CLIC. Given \( v \in G \), let \( K_v \) be the collection of \( \kappa \in K \) defined on \( C_v \) (thought of as functions). We say that two vertices \( v, w \in G \) are **equivalent** under \( K \) if there exists \( \kappa \in K_v \) such that \( \kappa(v) = w \), and denote this relation by \( v \overset{K}{\sim} w \).

We can define a similar equivalence in \( I_G \) as well. Given a local isomorphism \( \kappa : C_v \rightarrow C_w \) in \( K \) and \( \alpha \in \mathbb{N}^{C_v} \), define \( \kappa \alpha \in \mathbb{N}^{C_w} \), where

\[
\kappa \alpha(u) := \alpha(\kappa^{-1} u).
\]

We say that two indices \( \alpha, \beta \in I_G \) are **equivalent** under \( K \) if there exists a bijection \( \phi : \text{Dom}(\alpha) \rightarrow \text{Dom}(\beta) \) such that for each \( v \in \text{Dom}(\alpha) \), there exists \( \kappa \in K_v \) such that \( \kappa \alpha|_{C_v} = \beta|_{C_{\phi(v)}} \). We also write \( \alpha \overset{K}{\sim} \beta \) for this relation. It is easy to check that this defines equivalence relations on both \( G \) and \( I_G \).

As mentioned earlier in the footnote in the paragraph after Example 1.4, to obtain a representation which is consistent for all indices it is convenient to ensure that our index set is transitive under the group action. Thus, instead of \( \mathbb{N}^G \), we restrict our index set to

\[
\mathbb{N}_K^G := \{ \alpha \in \mathbb{N}^G : \alpha|_{C_v} \neq \alpha|_{C_w} \text{ whenever } v \overset{K}{\sim} w \}.
\]

Let us also define

\[
I_K^G := \bigcup_{C \in \mathcal{A}_G} \mathbb{N}_K^G.
\]
2 SETTINGS AND MAIN RESULTS

Figure 3: A commutative diagram associated to Definition 2.5

Roughly speaking, a consistent isomorphism class $K$ is an indicator that restricts the permutations of interest to act identically on vertices that are equivalent under $\sim^K$. The inclusion of identities and taking closure under inversion, composition and restriction, has ensured that $\sim^K$ is an equivalence relation in both $V$ and $I_G^K$. Also, $K$ is uniquely determined by the permutations under which the law of the array is invariant, as we identify each local isomorphism with its trivial extensions.

Now we are ready to define joint DAG-exchangeability and state the main theorem.

**Definition 2.5.** Let $G$ be a DAG and $K$ a CLIC of $G$. A permutation $\tau \in S_N^G$ is said to be $K$-commuting if for all $\beta \in N^G_K$ with $v \in G$ and $\kappa \in K_v$, we have

$$\tau \kappa (\beta) = \kappa \tau (\beta).$$

An array $X = (X_\alpha : \alpha \in N^G_K)$ (or $I^G_K$) is $(G, K)$-exchangeable if

$$(X_\alpha : \alpha \in N^G_K) \overset{d}{=} (X_{\tau \alpha} : \alpha \in N^G_K)$$

for all $K$-commuting $\tau$.

We will keep using the notation $Z_K(S_N^G)$ for the collection of all $K$-commuting permutations. Note that Definition 1.1 is a special case of Definition 2.5 where $K$ consists only of identity mappings.

**Theorem 2.6.** Let $G$ be a finite DAG, $K$ a CLIC of $G$. Then, an array $X = (X_\alpha : \alpha \in N^G_K)$ is $(G, K)$-exchangeable if and only if there exists a measurable function $f : [0, 1]^{|G|} \rightarrow \mathcal{Z}$ such that for all $\alpha \in N^G_K$,

$$X_\alpha \stackrel{a.s.}{=} f(U_{[\alpha|C]_K} : C \in \mathcal{G})$$

for some array $U$ of i.i.d. uniform random variables indexed by $([\beta]_K : \beta \in I^G_K)$, where $[\beta]_K$ denotes the equivalence class of $\beta$ with respect to $\sim^K$. 
Example 2.7. Let us inspect the classes of permutations associated to the CLIC’s introduced in Example 2.4 and their representations. For the first case (random block matrices), each of the cases allows permutations of the following forms, respectively, where \(i, j, k, \ell\) are the index values at \(u, v, r, c\), respectively:

1. \(X_{ij,k\ell} \rightarrow X_{r(i)\tau(j),\rho_{(ij)}(k)\lambda_{ij}(\ell)}\)
2. \(X_{ij,k\ell} \rightarrow X_{r(i)\theta(j),\rho_{ij}(k)\rho_{ij}(\ell)}\)
3. \(X_{ij,k\ell} \rightarrow X_{r(i)\tau(j),\rho_{(ij)}(k)\rho_{(ij)}(\ell)}\)

For each of the three cases, Theorem 2.6 provides a representation of the following forms:

1. \(X_{ij,k\ell} = f(U_{0,0,0}, U_{i,0,0}, U_{j,0,0}, U_{ij,0,0}, U_{ij,k,0}, U_{ij,0,\ell}, U_{ij,k,\ell})\)
2. \(X_{ij,k\ell} = f(U_{0,0,0}, U_{i,0,0}, U_{ij,0,0}, U_{ij,k,\ell})\)
3. \(X_{ij,k\ell} = f(U_{0,0,0}, U_{i,0,0}, U_{ij,0,0}, U_{ij,k,\ell}, U_{ij,\{k,\ell\},\{k,\ell\}})\)

For the second example, each of the cases allows permutations of the following forms, respectively, where \(i, j, k, \ell, m\) are the index values at \(u_1, v_1, u_2, v_2, v_3\), respectively:

4. \(X_{ij,k\ell,m} \rightarrow X_{r(i)\tau(j),\rho_{ij}(k)\lambda_{ij}(\ell)\theta_{ij}(m)}\)
5. \(X_{ij,k\ell,m} \rightarrow X_{r(i)\tau(j),\rho_{ij}(k)\rho_{ij}(\ell)\theta_{ij}(m)}\)

For each of the two cases, Theorem 2.6 provides a representation of the following forms:

4. \(X_{ij,k\ell,m} = f(U_{0,0,0,0}, U_{i,0,0,0}, U_{i,k,0,0}, U_{j,0,0,0}, U_{ij,0,0,0}, U_{ij,k,0,0}, U_{ij,0,\ell,0}, U_{ij,0,\ell,m}, U_{ij,\ell,0,0}, U_{ij,\ell,0,\ell,m}, U_{ij,\ell,0,\ell,m}, U_{ij,\ell,0,\ell,m}, U_{ij,\ell,0,\ell,m}, U_{ij,\ell,0,\ell,m}, U_{ij,\ell,0,\ell,m})\)
5. \(X_{ij,k\ell,m} = f(U_{0,0,0,0}, U_{i,0,0,0}, U_{i,k,0,0}, U_{j,0,0,0}, U_{ij,0,0,0}, U_{ij,k,0,0}, U_{ij,0,\ell,0}, U_{ij,0,\ell,m}, U_{ij,\ell,0,0}, U_{ij,\ell,0,\ell,m}, U_{ij,\ell,0,\ell,m}, U_{ij,\ell,0,\ell,m}, U_{ij,\ell,0,\ell,m}, U_{ij,\ell,0,\ell,m}, U_{ij,\ell,0,\ell,m})\)

2.1 Symmetry random variables associated to jointly DAG-exchangeable arrays

The overall plan of the proof of Theorem 2.6 is similar to that of Theorem 1.2. We deploy induction on the number of vertices of \(G\). To do this, we have to construct random variables which encode the intermediate information associated to \(X\), which we call a symmetry random variables associated to \(X\). We will see that the randomness of the uniform random variables affect \(X\) only through symmetry random variables. A typical example of this phenomenon is the role of the empirical distribution in an exchangeable sequence (see Lemma 7.1 of [Kal05]).
The key property we need to show in this strategy is conditional independence among the involved \( \sigma \)-fields, and that is Proposition 2.10 in our case. It is a parallel of Proposition 4.1 of [JLSY18], of which the proof is based on results of [Hoo79]. One aspect of Hoover’s proof is that it depends heavily on nonstandard analysis and symbolic logic. In this paper, we provide a probabilistic proof of Proposition 2.10 independent of Hoover’s. As mentioned at the beginning of the paper, our strategy resembles that of [Kal05] (especially Chapter 7) in the proof of the Aldous-Hoover representation theorem in a sense that we use systematic tools to prove conditional independence between involved random variables to deploy coding lemmas that provides representations using independent uniform random variables. (See the appendix for the lemmas that we use in the proof of the main result.)

Let \( X \) be a \((G, K)\)-exchangeable array. Let us write \( C_{K, \alpha}(S^G_N) \) for the collection of \( K \)-commuting permutations \( \tau \) such that \( \tau \alpha = \alpha \) for all \( \alpha \in I^G_K \).

The following are basic properties of \( C_{K, \alpha}(S^G_N) \).

(a) If \( \alpha \in \text{Restr}(\beta) \), then \( C_{K, \beta}(S^G_N) \subseteq C_{K, \alpha}(S^G_N) \).

(b) If \( \alpha \overset{K}{\sim} \beta \), then \( C_{K, \alpha}(S^G_N) = C_{K, \beta}(S^G_N) \).

The property (a) is obvious. The new property (b) follows from the fact that \( \tau \in Z^K_K(S^G_N) \) commutes with the elements of \( K \).

Let \( \mathcal{F}_\alpha \) denote the invariant \( \sigma \)-field of permutations fixing \( \alpha \). (we have \( \mathcal{F}_\alpha = \mathcal{F}_\beta \) for \( \alpha \overset{K}{\sim} \beta \) by (b).) We want to construct a Borel-valued random array \( S := (S_\alpha : \alpha \in I^G_K) \) satisfying the following properties, and call it a \textit{random symmetry array} associated to \( X \):

1. \( \mathcal{F}_\alpha = \sigma(S_\alpha) \).
2. The array \((X, S)\) is \((G, K)\)-exchangeable.
3. For \( \alpha \overset{K}{\sim} \beta \), \( S_\alpha = S_\beta \).

Assign a well-ordering on \( \mathcal{A}_G \). For each equivalence class of \( I^G_K \) under \( \overset{K}{\sim} \), choose a representative whose domain is the smallest under this well-ordering. From now on, let us assume that we have a fixed collection of such representatives, and denote this collection by \( \Gamma^G_K \).

The existence of symmetry arrays is a straightforward exercise.

\textbf{Proposition 2.8.} \textit{For any \((G, K)\)-exchangeable array \( X \) taking values in a Borel space, a random symmetry array exists.}\footnote{For those who are concerned with using the axiom of choice in this procedure, we note that it is not the case. Since \( \mathcal{A}_G \) is a finite set, we do not need the well-ordering principle when we choose a well-ordering. When choosing the representatives, for each subgraphs we can assign a well-ordering on the vertices and select the smallest element in the lexicographical order.}
Proof. For $\alpha \in I_K^G$, let us write $\gamma_\alpha$ for the representative of $[\alpha]_K$.

For each $C \in \mathcal{G}$, choose $\gamma_C \in \Gamma$ with $\text{Dom}(\gamma_C) = C$, if there is any. Since $\mathcal{F}_\gamma$ is countably generated, there exists a Borel-valued random variable $S_{\gamma_C}$ such that $\sigma(S_{\gamma_C}) = \mathcal{F}_\gamma$. \footnote{Any countably generated $\sigma$-field can be generated by a random variable taking values on a Borel space. (Exercise 3.13, [Res13])} Since $S_{\gamma_C}$ is $\sigma(X)$-measurable, there exists a measurable function $f_C$ such that $f_C(X) = S_{\gamma_C}$ almost surely. For any other $\gamma \in \Gamma$ with $\text{Dom}(\gamma_C) = C$, choose $\tau \in Z_K(S_{\gamma_C}^G)$ such that $\tau \gamma = \gamma_C$ and let

$$S_\gamma := f_C(\tau X).$$

Note that the choice of $\tau$ is irrelevant. For $\alpha \in I_K^G$, we let

$$S_\alpha := S_{\gamma_\alpha}.$$

Let $\alpha, \beta \in I_K^G$, $\tau \in Z_K(S_{\gamma_\alpha}^G)$ with $\alpha = \tau \beta$. We claim that

$$\phi(\tau X) = S_\beta$$

whenever $\phi(X) = S_\alpha$. The left hand side represents the action of $\tau$ as we regard $S_\alpha$ as an $\sigma(X)$-measurable random element, while on the right hand side $\tau$ acts on $\sigma(S)$-measurable random elements. Once we have that these actions are identical, Lemma A.5 implies that the array $S$ constructed this way satisfies the desired properties.

Let $\text{Dom}(\gamma_\beta) = D$. Then, we have $S_{\gamma_\beta} = f_D(\lambda X)$, for $\lambda \in Z_K(S_{\gamma_\beta}^G)$ such that $\gamma_D = \lambda \gamma_\beta$. Considering the way we have chosen the representatives, $\gamma_\alpha$ and $\gamma_\beta$ are defined on the same domain, and hence for some $\rho$ we have $\gamma_D = \rho \gamma_\alpha$ and thus $S_{\gamma_\alpha} = f_D(\rho X)$.

For $v \in C$, let $\kappa_1, \kappa_2 \in K$ be local isomorphisms such that

$$\kappa_1(\beta|_{C_{s_1(v)}}) = \gamma_\beta|_{C_v}, \kappa_2(\gamma_\alpha|_{C_{s_2\alpha(v)}}) = \alpha|_{C_{s_1(v)}}.$$

Then, $\rho \tau \lambda^{-1}$ fixes $\gamma_D$ at $v \in D$ since

$$\rho \tau \lambda^{-1}(\gamma_D|_{C_v}) = \rho \tau (\gamma_\beta|_{C_v}) = \rho \tau \kappa_1(\beta|_{C_{s_1(v)}})$$

$$= \rho \kappa_1 \tau(\beta|_{C_{s_1(v)}}) = \rho \kappa_1(\alpha|_{C_{s_1(v)}})$$

$$= \rho \kappa_1 \kappa_2(\gamma_\alpha|_{C_{s_2\alpha(v)}}) = \kappa_1 \kappa_2 \rho(\gamma_\alpha|_{C_{s_2\alpha(v)}})$$

$$= \kappa_1 \kappa_2(\gamma_D|_{C_{s_2\alpha(v)}}) = \gamma_D|_{C_v}.$$

Since $v \in D$ is arbitrary, $\rho \tau \lambda^{-1}$ fixes $\gamma_D$. Thus, by definition of $S$, we have

$$f_D(X) = f_D(\rho \tau \lambda^{-1} X)$$

almost surely. By exchangeability the equation holds almost surely if we replace $X$ by $\lambda X$, and hence

$$S_{\gamma_\beta} = f_D(\lambda X) = f_D(\rho \tau X)$$

almost surely. \qed
Once we have an associated symmetry array, we can modify Theorem 2.6 into the following variant. For a generic array $Y = (Y_i : i \in I)$ and $J \subseteq I$, we write $Y_J := (Y_i : i \in J)$.

**Theorem 2.9.** Let $G, K, X$ be as in Theorem 2.6 and let $S$ be a symmetry array of $X$. Then, there exist measurable functions $f_C : [0, 1]^{|G|} \to X$ such that for all $\alpha \in \Gamma^G_K$,

$$S_{\alpha} \overset{a.s.}{=} f_{\text{Dom}(\alpha)}(S_{\text{Restr}^r(\gamma_{\alpha})}, U_{\alpha})$$

for some array $U$ of i.i.d. uniform random variables indexed by $\Gamma^G_K$.

Note that by recursively replacing $S_{\beta}$ with $f_{\text{Dom}(\beta)}(S_{\text{Restr}^r(\gamma_{\beta})}, U_{\beta})$, we obtain the alternate representation of (16) of the form

$$S_{\alpha} = g_{\text{Dom}(\alpha)}(U_{\beta} : \beta \in \text{Restr}(\gamma_{\alpha}))$$

for some measurable functions $g_C$, where $U_{\beta} = U_{\gamma_{\beta}}$ for $\beta \in I^G_K$.

As mentioned earlier, the basic strategy of our proof is using induction on $|G|$, the number of vertices of $G$. We first build representations on the proper subgraphs of $G$, and tie them all together into a representation in the whole $G$. Proposition 2.10 is a key result which makes this “tying” possible.

Let $\alpha, \beta \in I^G_K$, and define $\text{Restr}(\alpha, K)$ to be the collection of $\alpha' \in I^G_K$ such that $\alpha' \overset{K}{\sim} \alpha|_C$ for some $C \in \mathcal{A}_{\text{Dom}(\alpha)}$. Let

$$D_{\alpha, \beta} := \{v \in \text{Dom}(\alpha) : \alpha|_{C_v} \in \text{Restr}(\beta, K)\},$$

and define

$$\alpha \land \beta := \alpha|_{D_{\alpha, \beta}}.$$

These are the joint-exchangeability counterparts of restrictions and intersections of two indices for separate DAG-exchangeability. One can easily check that $\alpha \land \beta \overset{K}{\sim} \beta \land \alpha$ for all $\alpha, \beta \in I^G_K$.

**Proposition 2.10.** Let $\alpha, \ldots, \alpha_n \in I^G_K$. Then, $(S_{\alpha_k} : k \leq n)$ are independent given $(S_{\alpha_k \land \alpha_j} : k \neq j)$.

The next corollary follows from Proposition 2.10. Let us first define some notation.

- $\text{Restr}^r(\alpha, K) := \text{Restr}(\alpha, K) \setminus [\alpha]_K$
- $J_k := (S_{\alpha} : \alpha \in I^G_K, |\text{Dom}(\alpha)| \leq k)$
- $N_k := J_k \setminus J_{k-1}$

Note that $\sigma(S_{\text{Restr}^r(\alpha, K)}) = \sigma(S_{\text{Restr}(\alpha)})$.

**Corollary 2.11.** Let $S_k := (S_{\alpha} : \alpha \in J_k)$. Then, for $1 \leq k \leq |G|$, $S_k$ is a conditionally independent family given $S_{k-1}$. In particular, for $\alpha \in N^G_K$, we have

$$S_{\alpha} \perp \perp S_{\text{Restr}^r(\alpha, K)} \setminus S_{[\alpha]_K}.$$
Proof. Fix $\alpha \in J_k$, and let $A$ be an arbitrary finite subset of $J_k \setminus \{\alpha\}$. By Proposition 2.10, we have

$$S_\alpha \perp \perp (S_\beta : \beta \in A),$$

and since $\sigma(S_{\alpha \land \beta} : \beta \in A) \subseteq \sigma(S_{\text{Restr}'(\alpha,K)}) \subseteq \sigma(S_\alpha)$, we have

$$S_\alpha \perp \perp (S_\beta : \beta \in A).$$

Since $A$ is arbitrary, we have

$$S_\alpha \perp \perp S_{\text{Restr}'(\alpha,K)} (S_\beta : \beta \in A),$$

and since $S_{\text{Restr}'(\alpha,K)} \in \sigma(S_{k-1}) \subseteq \sigma(S_k \setminus S_\alpha)$, we have

$$S_\alpha \perp \perp S_{k-1} \setminus S_\alpha.$$  

(18)

Since $\alpha$ is arbitrary, the proof is complete. $\square$

Proof of Theorem 2.9. We build an induction to show that for all $k \leq |G|$, there exists an i.i.d. array of uniform random variables $(U_\alpha : \alpha \in J_k)$ such that (16) holds for all $\alpha \in J_k$. (The case $k = 0$ is obvious.) Let us assume that there exists an i.i.d. array $W_{k-1} := (W_\alpha : \alpha \in J_{k-1})$ and a family of measurable functions $(f_C : C \in \mathcal{A}_G, |C| \leq k-1)$ such that almost surely,

$$S_\alpha = f_{\text{Dom}(\alpha)}(S_{\text{Restr}'(\alpha)}, W_\alpha), \alpha \in J_{k-1}.$$  

(20)

Fix $\alpha \in N_k$. By (18) from Corollary 2.11, we have

$$S_\alpha \perp \perp S_{\text{Restr}'(\alpha,K)} \setminus S_\alpha.$$  

Thus, by Lemma A.4, there exists a uniform random variable $V_\alpha$ independent of $S_{k-1}$ such that

$$S_\alpha = f_\alpha(S_{\text{Restr}'(\alpha)}, V_\alpha)$$  

(21)

almost surely. By exchangeability, there exists an array of uniform random variables $\partial V_k := (V_\alpha : \alpha \in N_k)$, which are not necessarily independent, such that (21) holds for every $\alpha \in N_k$, with the choice of $f_\alpha$ identical for all $\alpha$ defined on the same domain, which we denote as $f_{\text{Dom}(\alpha)}$.

Now consider an array $\partial W_k := (W_\alpha : \alpha \in N_k)$ of i.i.d. uniform random variables, which are also independent of $W_{k-1}$, and for $\alpha \in N_k$ define

$$S'_\alpha := f_\alpha(S_{\text{Restr}'(\alpha)}, W_\alpha).$$  

(22)

Since we can replace (20) and (22) into equations of the form (17), we can combine them into a one-line expression of the form

$$S'_k = F(W_k),$$  

(23)

for some function $F$ where $S'_k = (S'_{k-1}, (S'_\alpha : \alpha \in N_k))$ and $W_k = (W_\alpha : \alpha \in J_k)$. Note that $W_k = (W_{k-1}, \partial W_k)$ is an i.i.d. array. On the other hand, we have the following properties:
2 SETTINGS AND MAIN RESULTS

- \( P[S_\alpha \in \cdot | S_{k-1}] = P[S'_\alpha \in \cdot | S_{k-1}] \) almost surely for all \( \alpha \in N_k \) since both \( W_\alpha \) and \( V_\alpha \) are independent of \( S_{k-1} \).
- \((S_\alpha : \alpha \in N_k)\) is a conditionally independent family given \( S_{k-1} \) by Corollary 2.11.
- \((S'_\alpha : \alpha \in N_k)\) is a conditionally independent family given \( S_{k-1} \) by construction.

By Lemma \( \ref{A.6} \) we have

\[
S_k = (S_{k-1}, (S_\alpha : \alpha \in N_k)) \overset{d}{=} (S_{k-1}, (S'_\alpha : \alpha \in N_k)).
\]

Thus, with \( \ref{23} \), we can apply Lemma \( \ref{A.2} \) to obtain an array of i.i.d. uniform random variables \( U_k := (U_\alpha : \alpha \in J_k) \) such that

\[
S_k = F(U_k) \tag{24}
\]

almost surely. Splitting (24) back to individual equations of the form (22), we obtain the desired representation for dimension \( k \). Since \( k \leq |G| \) is arbitrary, we have (16) for all \( \alpha \in \Gamma_K^G \) by induction.

**Proof of Theorem 2.6**

Let \( \alpha \in N_K^G \cap \Gamma_K^G \). Since \( X_\alpha \) is \( \sigma(X) \)-measurable and is invariant under permutations fixing \( \alpha \), we have \( X_\alpha \in \mathcal{F}_\alpha \) and hence \( X_\alpha = h(S_\alpha) \) for some measurable function \( h \). By inserting (17), we obtain (15) by identifying \( I_K^G \) modulo \( \Gamma_K^G \) with \( \Gamma_K^G \).

Let \( U_\alpha := U_{\gamma_\alpha} \) and \( U = (U_\alpha : \alpha \in I_K^G) \). To obtain (15) for all \( \alpha \in N_K^G \), it suffices to show that we can choose \( U \) in a way such that \((X, U)\), or equivalently \((S, U)\), is \((G, K)\)-exchangeable. Indeed, if \((X, U)\) is exchangeable, then by transitivity, for \( \beta \in N_K^G \), there exists \( \alpha \in N_K^G \cap \Gamma_K^G \) such that \( \tau \alpha = \beta \), and by exchangeability (15) holds if we replace \( \alpha \) by \( \beta = \tau \alpha \).

By Transfer Lemma \( \ref{A.1} \) there exists a family of measurable functions \( (\phi_C : C \in \mathcal{A}_G) \) such that for any \( \alpha \in \Gamma_K^G \) and any uniform random variable \( V \) independent of \( S_\alpha \), we have \( (S_\alpha, U_\alpha) \overset{d}{=} (S_\alpha, \phi_{Dom(\alpha)}(S_\alpha, V)) \). \( \Box \) Let \( V = (V_\alpha : \alpha \in I_K^G) \) be an array of uniform random variables independent of \( S \), where \( V_\alpha = V_{\gamma_\alpha} \) for all \( \alpha \in I_K^G \) and different components are all independent. Then, by the preceding arguments, we have

\[
S_\alpha = f_G(S_{\text{Restr}^\alpha(\alpha)}, U'_\alpha)
\]

almost surely for all \( \alpha \in \Gamma_K^G \) where \( U'_\alpha = \phi_{Dom(\alpha)}(S_\alpha, V_\alpha) \). (Note that \( S_{\text{Restr}^\alpha(\alpha)} \in \sigma(S_\alpha) \).)

Since \( V \) is an i.i.d. array independent of \( S \), we have

\[
U'_\alpha \overset{S}{=} (U'_\beta : \beta \in \Gamma_K^G \setminus \{\alpha\}) \tag{25}
\]

for all \( \alpha \in \Gamma_K^G \). Also, since \( (U_\alpha : \alpha \in \Gamma_K^G) \) is an independent family and \( S \setminus S_\alpha \in \sigma(U_\beta : \beta \in \Gamma_K^G \setminus \{\alpha\}) \), we have

\[
U_\alpha \overset{S \setminus S_\alpha}{=} (U_\beta : \beta \in \Gamma_K^G \setminus \{\alpha\}) \tag{26}
\]

\(^6\)We can choose \( \phi_C \) to depend only on the domain by exchangeability of \( S \).
Since \( S_\alpha \in \sigma(S \setminus S_{\alpha}, U_{\alpha}) \), we obtain
\[
U_\alpha \perp \perp (U_\beta : \beta \in \Gamma^G_K \setminus \{\alpha\}).
\] (27)

Therefore by (25) and (27), both \((S_{\alpha}, U_{\alpha}) : \alpha \in \Gamma^G_K\) and \((S_{\alpha}, U'_{\alpha}) : \alpha \in \Gamma^G_K\) are conditionally independent family given \( S \). Since \((S_{\alpha}, U_{\alpha}) \overset{d}{=} (S_{\alpha}, U'_{\alpha})\) and both are conditionally independent of \( S \) given \( S_\alpha \), we have
\[
(S, U_{\alpha}) \overset{d}{=} (S, U'_{\alpha})
\] (28)
for all \( \alpha \in \Gamma^G_K \). Therefore, by Lemma A.6, we have
\[
(S, (U_{\alpha} : \alpha \in I^G_K)) \overset{d}{=} (S, (U'_{\alpha} : \alpha \in I^G_K))
\] (29)
where \( U'_{\alpha} := U'_{\gamma_{\alpha}} \). Thus, the relations (16) still hold even if we replace \( U \) by \( U' := (U'_{\alpha} : \alpha \in I^G_K) \).

Since \( S \) and \( V \) are exchangeable and independent of each other, \((S, V)\) is exchangeable. Thus, by Lemma A.5, \((S, U')\) is exchangeable.

\[\square\]

2.2 Proof of Proposition 2.10

Let \( \mathcal{F}_\alpha = \sigma(S_\alpha) \), as in Section 2.1.

Lemma 2.12. For \( \alpha \in I^G_K \), \( n \in \mathbb{N} \), let
\[
\mathcal{F}_\alpha^n := \sigma(X_\beta : \beta \in \mathbb{N}^G_K \text{ there exists } C \in \mathcal{A}_G \text{ such that } \beta|_C \in \text{Restr}(\alpha, K), \beta(v) \geq n \text{ for all } v \not\in C).
\]
Then, \( \mathcal{F}_\alpha = \bigcap_{n \geq 1} \mathcal{F}_\alpha^n \).

Let us introduce some notations to be used in the proof. For \( \alpha, \beta \in I^G_K \) and \( v \in G \), let
\[
A_\alpha(\beta, v) := \{\alpha(w) : w \overset{K}{\sim} v, \text{ there exists some } \kappa \in K_v \text{ such that } \kappa|_{C_v}(u) = \alpha|_{C_w}(u) \text{ for all } u \prec v\},
\]
and
\[
\tau_{\beta, v}(n) := \begin{cases} 
\min\{k > n : k \not\in A_\alpha(\beta, v)\}, & n \not\in A_\alpha(\beta, v), \\
n, & n \in A_\alpha(\beta, v).
\end{cases}
\]
We define \( \rho_\alpha \) to be an injective homomorphism of \( I^G_K \) defined as
\[
\rho_\alpha(\beta)(v) = \tau_{A_\alpha(\beta, v)}(\beta(v)).
\]

To explain in detail, there are three cases:

\footnote{Since any injective homomorphism restricted to a finite subset can be extended to an element of \( \mathbb{Z}_K(S_N^G) \), the law of the array is invariant under injective homomorphisms by Kolmogorov extension theorem.}
1. If $\kappa_\beta|_{C_v} = \alpha|_{C_w}$ for some $w \in \text{Dom}(\alpha)$ and $\kappa \in K$ (or equivalently, $\beta|_{C_v} \in \text{Restr}(\alpha, K)$), then $\rho_\alpha(\beta)(v) = \beta(v)$.

2. If for some $w \in \text{Dom}(\alpha)$ and $\kappa \in K$, we have $\kappa_\beta|_{C_v}(u) = \alpha|_{C_w}(u)$ for all $u \in C_w\setminus\{w\}$ but none of such $w$ and $\kappa$ satisfies $\kappa_\beta|_{C_v}(w) = \alpha(w)$, we let $\rho_\alpha(\beta)(v) = \beta(v) + \ell$, where $\ell$ is the smallest positive integer such that $\beta_{\ell,v}|_{C_v} \notin \text{Restr}(\alpha, K)$, where

$$\beta_{\ell,v}(u) := \beta(u) + \ell 1_{\{u=v\}}.$$

3. Otherwise, we have $\rho_\alpha(\beta)(v) = \beta(v) + 1$.

**Proof.** One can easily see that $\rho_\alpha$ fixes $\alpha$ and commutes with $K$. For any $\beta \in \mathbb{N}_K^G$, $D_{\beta,\alpha}$ is a closed subgraph of $G$, and also $D_{\rho_\alpha(\beta),\alpha} = D_{\beta,\alpha}$. On the other hand, for $v \notin D_{\beta,\alpha}$, one can see that $\rho_\alpha(\beta)(v) \geq \beta(v) + 1$. Thus, we have $\rho_\alpha^n(\alpha) \in \mathcal{F}_\alpha^{n+1}$.

So, for arbitrary $E \in \mathcal{F}_\alpha$, acting $\rho_\alpha^n$ on the inclusion $E \in \sigma(\alpha)$ we obtain $E \in \sigma(\rho_\alpha^n(\alpha)) \subseteq \mathcal{F}_\alpha^{n+1}$.

This shows that $\mathcal{F}_\alpha \subseteq \bigcap_{n \leq 1} \mathcal{F}_\alpha^n$.

To prove the converse, consider the collection $\mathcal{T}_n$ of all finite permutations $\tau \in Z_K(S_{\mathbb{N}}^{nG})$ such that

1. $\tau$ fixes $\alpha$.
2. for all $\beta$, $\tau(\beta)(v) = \beta(v)$ whenever there exists $u \leq v$ such that $\beta(u) > n$.

Then, as $n \to \infty$, the collection $\mathcal{T}_n$ eventually contains all finite permutations in $Z_K(S_{\mathbb{N}}^{nG})$ fixing $\alpha$, and $\mathcal{F}_\alpha^n$ is invariant under the action of $\mathcal{T}_n$. Therefore, we have $\bigcap_{n \geq 1} \mathcal{F}_\alpha^n \subseteq \mathcal{F}_\alpha$.

**Proof of Proposition 2.10.** We use induction on the number of indices $n$. The case $n = 1$ is obvious.

Consider $\rho := \rho_{\alpha_1}$ which is defined as in the proof of Lemma 2.12. As we have seen in the proof of Lemma 2.12 acting $\rho$ on $\beta$ fixes the values on $D_{\beta,\alpha_1}$, which is by definition equal to $\beta \land \alpha_1$, and shifts all the values outside $D_{\beta,\alpha_1}$ by at least $+1$. By exchangeability, we have

$$(S_{\alpha_1}, ..., S_{\alpha_n}) \overset{d}{=} (S_{\alpha_1}, \rho^k(S_{\alpha_2}), ..., \rho^k(S_{\alpha_n})).$$

By Lemma A.3 we have

$$S_{\alpha_1} \overset{\perp}{\rho^k(S_{\alpha_2}), ..., \rho^k(S_{\alpha_n})} S_{\alpha_2}, ..., S_{\alpha_n}$$

for all $k \in \mathbb{N}$. Under $\rho^k$, one can easily see that the indices generating $\mathcal{F}_{\alpha_1}$ falls to the indices generating $\mathcal{F}_{\alpha_1 \land \alpha_j}^{k+1}$. This shows that for all $k \geq 2$,

$$S_{\alpha_1} \overset{\perp}{\mathcal{F}_{\alpha_1 \land \alpha_j}^{k+1} ; 2 \leq j \leq n} S_{\alpha_2}, ..., S_{\alpha_n}.$$
A SUPPLEMENTARY LEMMAS

By the inductive hypothesis, \( S_{a_2}, \ldots, S_{a_n} \) are conditionally independent given \((S_{a_i} : 2 \leq i \neq j \leq n)\). Thus, by Lemma 2.12 and Lemma A.7, the \( \sigma \)-field generated by \( \mathcal{F}_{a_1}^{k} \), \ldots, \( \mathcal{F}_{a_1}^{k} \) decreases to some \( \mathcal{G} \subseteq \sigma(\mathcal{F}_{a_1}^{k} \cup \cdots \cup \mathcal{F}_{a_1}^{k}) \). Applying the backward martingale convergence theorem to (31) as \( k \to \infty \), we obtain

\[
S_{a_1} \perp \mathcal{G} S_{a_2}, \ldots, S_{a_n},
\]

and since each \( \mathcal{F}_{a_1}^{k} \cup \cdots \cup \mathcal{F}_{a_1}^{k} \) is a sub-\( \sigma \)-field of \( S_{a_1} \), we obtain

\[
S_{a_1} \perp \left( \mathcal{F}_{a_1}^{k} \cup \cdots \cup \mathcal{F}_{a_1}^{k} \right) S_{a_2}, \ldots, S_{a_n}. \tag{32}
\]

Since \( S_{a_2}, \ldots, S_{a_n} \) are conditionally independent given \((\mathcal{F}_{a_1}^{k} \cup \cdots \cup \mathcal{F}_{a_1}^{k}) \) and for each \( j \geq 2 \), \( \mathcal{F}_{a_1}^{k} \) is a sub-\( \sigma \)-field of \( \mathcal{F}_{a_j}^{k} \), we have that \( S_{a_2}, \ldots, S_{a_n} \) are conditionally independent given \((\mathcal{F}_{a_1}^{k} \cup \cdots \cup \mathcal{F}_{a_1}^{k}) \). Combining this with (32), we obtain the desired result. \( \square \)

A Supplementary Lemmas

Elementary results that we use in the main text are introduced in this section. All are standard results and frequently used in exchangeability theory. For those results without proofs we have added references where one can find the proofs. We note again that the richness of the probability space is always assumed.

Lemma A.1 (Transfer Lemma: Theorem 6.10, [Kal02]). Let \( X, Y \) be random elements in a Borel space. Then,

1. For all \( X' \overset{d}{=} X \), there exists a measurable function \( f \) such that whenever \( W \) is a uniform random variable independent of \( X' \), then \( Y' := f(X', W) \) satisfies \((X, Y) \overset{d}{=} (X', Y')\).

2. There exist measurable functions \( h \) and \( g \) such that whenever \( W \) is a uniform random variable independent of \( X \) and \( Y', V := h(X, Y, W) \) is a uniform random variable independent of \( X \) satisfying \( Y = g(X, V) \) almost surely.

Lemma A.2 (Corollary 6.11, [Kal02]). Let \( X, Y \) be Borel-valued random variables such that \( X \overset{d}{=} f(Y) \) for some measurable function \( f \). Then, there exists a random variable \( Y' \overset{d}{=} Y \) such that \( X = f(Y') \) almost surely.

Lemma A.3 (Lemma 1.3, [Kal05]). Let \( X, Y, Z \) be random variables such that \((X, Y) \overset{d}{=} (X, Z)\) and \( \sigma(Y) \subseteq \sigma(Z) \). Then, \( X \perp_{Y} Z \).

Lemma A.4 (Proposition 5.13, [Kal02]). Let \( X, Y, Z \) be random elements, where \( X \) lies in a Borel space. Then, \( X \) is conditionally independent of \( Z \) given \( Y \) if and only if there exists a measurable function \( f \) and a uniform random variable \( U \) independent of \( Y, Z \) such that \( X = f(Y, U) \) almost surely.

Lemma A.5. Let \( H \) be a group acting measurably on Borel spaces \( \mathcal{X} \) and \( \mathcal{Y} \), and let \( \mu \) be an \( H \)-invariant probability measure on \( \mathcal{X} \), that is, \( x \) is \( H \)-exchangeable under \( \mu \). Let \( \phi : \mathcal{X} \to \mathcal{Y} \) be a measurable function. If \( \phi(\tau x) = \tau \phi(x) \) \( \mu \)-almost surely for all \( \tau \in H \), then \((x, \phi(x))\) is \( H \)-exchangeable under \( \mu \).
Proof. \( \mu[x \in A, \phi(x) \in B] = \mu[\tau x \in A, \phi(\tau x) \in B] = \mu[\tau x \in A, \tau \phi(x)] \).

\[ \square \]

**Lemma A.6.** Let \( (X_i : i \in I), (Y_i : i \in I) \) be a family of random variables with a countable index set \( I \). For a random variable \( S \), assume that the following are true:

- \( (S, X_i) \overset{d}{=} (S, Y_i) \). Equivalently, \( P[X_i \in |S] = P[Y_i \in |S] \) almost surely.
- Given \( S \), both \( (X_i : i \in I) \) and \( (Y_i : i \in I) \) are conditionally independent families.

Then, we have \( (S, X_i : i \in I) \overset{d}{=} (S, Y_i : i \in I) \).

**Proof.** Without loss of generality, let \( I = \mathbb{N} \). Then for \( n \in \mathbb{N} \) and bounded measurable functions \( f_1, \ldots, f_n \),

\[
E[f_1(X_1) \cdots f_n(X_n)|S] = E[f_1(X_1)|S] \cdots E[f_n(X_n)|S] = E[f_1(Y_1)|S] \cdots E[f_n(Y_n)|S] = E[f_1(Y_1) \cdots f_n(Y_n)|S].
\]

\[ \square \]

**Lemma A.7.** For each \( n \in \mathbb{N} \), let \( (\mathcal{F}^n_k : k \in \mathbb{N}) \) be a sequence of decreasing \( \sigma \)-fields with \( \mathcal{F}^n := \bigcap_{k \geq 1} \mathcal{F}^n_k \). Assume that given \( \mathcal{G} \), the family \( (\mathcal{F}^n_1 : n \in \mathbb{N}) \) is independent.

Then, \( \mathcal{F} := \bigcap_{k \in \mathbb{N}} \bigvee_{n \in \mathbb{N}} \mathcal{F}^n_k \) is a sub-\( \sigma \)-field of \( \mathcal{G} \bigvee \bigvee_{n \in \mathbb{N}} \mathcal{F}^n \). In particular, if \( (\mathcal{F}^n_1 : n \in \mathbb{N}) \) are unconditionally independent, then \( \mathcal{F} = \bigvee_{n \in \mathbb{N}} \mathcal{F}^n \).

**Proof.** For each \( n \in \mathbb{N} \) choose an event \( A_n \in \mathcal{F}^n_1 \), and let \( A \in \mathcal{G} \). Then,

\[
P[A \bigcap_{j \leq N} A_j|\mathcal{G}, \mathcal{F}^n_k : n \in \mathbb{N}] = \mathbf{1}_A \prod_{j \leq N} P[A_j|\mathcal{F}^n_k, \mathcal{G}]
\]

by conditional independence. By backward martingale convergence the right hand side converges to a \( \mathcal{G} \bigvee \bigvee_{n \in \mathbb{N}} \mathcal{F}^n \)-measurable random variable as \( k \to \infty \). Since the collection of all the events of the form \( A \bigcap_{j \leq N} A_j \) is a \( \pi \)-system generating \( \mathcal{G} \bigvee \bigvee_{n \in \mathbb{N}} \mathcal{F}^n \), we can use \( \pi \)-\( \lambda \) arguments to show that for any \( \mathcal{G} \bigvee \bigvee_{n \in \mathbb{N}} \mathcal{F}^n \)-measurable event \( B \), we have

\[
\lim_{k \to \infty} P[B|\mathcal{G}, \mathcal{F}^n_k : n \in \mathbb{N}] \in \mathcal{G} \bigvee \bigvee_{n \in \mathbb{N}} \mathcal{F}^n.
\]

\[ \square \]

Without the conditional independence, we cannot guarantee the result. Consider two sequences of random variables \( X = (X_n : n \in \mathbb{N}) \) and \( Y = (Y_n : n \in \mathbb{N}) \), and let \( P \) be a uniform random variable. Suppose that given \( P \), \( X \) and \( Y \) are independent i.i.d. sequences, where \( P[X_1 = 1|P] = 1 - P[X_1 = -1|P] = P \) and \( P[Y_1 = 1|P] = 1 - P[Y_1 = -1|P] = 1/2 \). Let \( Z_n := X_nY_n \). Then, \( Y \) and \( Z \) are i.i.d. sequences of random variables independent of \( P \), and hence their tail \( \sigma \)-fields are trivial. However, since the tail \( \sigma \)-field of the joint sequence \( (Y, Z) \) recovers \( P \), and hence it is not equal to the join of the tail \( \sigma \)-fields of the components.
This shows that $\mathcal{G} \vee \mathcal{F}^n_k$ converges to $\mathcal{G} \vee \mathcal{F}^n$ as $k \to \infty$. Since $\mathcal{G} \vee \mathcal{F}^n_k \subseteq \mathcal{G} \vee \mathcal{F}^n_k$, we can conclude that

$$\bigcap_{k \in \mathbb{N}} (\mathcal{G} \vee \mathcal{F}^n_k) \subseteq \bigcap_{k \in \mathbb{N}} (\mathcal{G} \vee \mathcal{F}^n_k) = \mathcal{G} \vee \mathcal{F}^n.$$

The last statement is obvious since we always have $\mathcal{G} \vee \mathcal{F}^n_k \subseteq \mathcal{G} \vee \mathcal{F}^n_k$ for each $k \in \mathbb{N}$, which implies that $\mathcal{G} \vee \mathcal{F}^n \subseteq \bigcap_{k \in \mathbb{N}} (\mathcal{G} \vee \mathcal{F}^n_k) = \mathcal{F}$. 

Acknowledgments

The author was 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).

References

[Ald81] David J Aldous. Representations for partially exchangeable arrays of random variables. *Journal of Multivariate Analysis*, 11(4):581–598, 1981.

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

[BRT19] Benjamin Bloem-Reddy and Yee Whye Teh. Probabilistic symmetry and invariant neural networks. *arXiv preprint arXiv:1901.06082*, 2019.

[BZSL13] Joan Bruna, Wojciech Zaremba, Arthur Szlam, and Yann LeCun. Spectral networks and locally connected networks on graphs. *arXiv preprint arXiv:1312.6203*, 2013.

[CW16] Taco Cohen and Max Welling. Group equivariant convolutional networks. In *International conference on machine learning*, pages 2990–2999, 2016.

[DF29] Bruno De Finetti. Funzione caratteristica di un fenomeno aleatorio. In *Atti del Congresso Internazionale dei Matematici: Bologna del 3 al 10 de settembre di 1928*, pages 179–190, 1929.

[DF37] Bruno De Finetti. La prévision: ses lois logiques, ses sources subjectives. In *Annales de l’institut Henri Poincaré*, volume 7, pages 1–68, 1937.

[FP+12] Sandra Fortini, Sonia Petrone, et al. Predictive construction of priors in bayesian nonparametrics. *Brazilian Journal of Probability and Statistics*, 26(4):423–449, 2012.

[Hof08] Peter Hoff. Modeling homophily and stochastic equivalence in symmetric relational data. In *Advances in neural information processing systems*, pages 657–664, 2008.
[Hoo79] Douglas N Hoover. Relations on probability spaces and arrays of random variables. Preprint, 
Institute for Advanced Study, Princeton, NJ, 2, 1979.

[HS55] Edwin Hewitt and Leonard J Savage. Symmetric measures on cartesian products. Transactions of the American Mathematical Society, 80(2):470–501, 1955.

[JLSY18] Paul Jung, Jiho Lee, Sam Staton, and Hongseok Yang. A generalization of hierarchical 
exchangeability on trees to directed acyclic graphs. arXiv preprint arXiv:1812.06282 (to be 
published), 2018.

[Kal89] Olav Kallenberg. On the representation theorem for exchangeable arrays. Journal of Multivariate Analysis, 30(1):137–154, 1989.

[Kal92] Olav Kallenberg. Symmetries on random arrays and set-indexed processes. Journal of Theoretical Probability, 5(4):727–765, 1992.

[Kal02] Olav Kallenberg. Foundations of modern probability. Springer, 2002.

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

[LOGR12] James Lloyd, Peter Orbanz, Zoubin Ghahramani, and Daniel M Roy. Random function 
priors for exchangeable arrays with applications to graphs and relational data. In Advances 
in Neural Information Processing Systems, pages 998–1006, 2012.

[OR14] Peter Orbanz and Daniel M Roy. Bayesian models of graphs, arrays and other exchange-
able random structures. IEEE transactions on pattern analysis and machine intelligence, 
37(2):437–461, 2014.

[Res13] Sidney I Resnick. A probability path. Springer Science & Business Media, 2013.

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