Journal of Algebraic Statistics

On Exchangeability in Network Models

Steffen L. Lauritzen[1], Alessandro Rinaldo[2,*], Kayvan Sadeghi[3]

[1] *Department of Mathematical Sciences, University of Copenhagen, Copenhagen, Denmark*
[2] *Department of Statistics and Data Science, Carnegie Mellon University, Pittsburgh, USA*
[3] *Department of Statistical Science, University College London, London, UK*

Abstract. We derive representation theorems for exchangeable distributions on finite and infinite graphs using elementary arguments based on geometric and graph-theoretic concepts. Our results elucidate some of the key differences, and their implications, between statistical network models that are finitely exchangeable and models that define a consistent sequence of probability distributions on graphs of increasing size.

2000 Mathematics Subject Classifications: 60G09, 05C80, 62H99

Key Words and Phrases: de Finetti's theorem; graphons; Möbius simplex; finite exchangeability; positive semidefinite functions

1. Introduction

Exchangeability is one of most basic forms of probabilistic invariance. When applied to probability distributions on graphs, it is equivalent to requiring that isomorphic graphs have the same probabilities. Indeed, exchangeability provides the probabilistic underpinning to the theory of dense graph limits developed recently in the graph-theoretic literature [see, e.g., 10, 6, 21].

In statistical network modeling, exchangeability is a common simplifying assumption. However, it is typically only enforced for models on graphs of a given size, and not simultaneously over sequences of models on graphs of increasing size. This practice is born out of convenience: it is much easier to formulate probability distributions on finite as opposed to infinite graphs. However, the consequences of relying on this weaker assumption of finite exchangeability can be detrimental to the validity and generalizability of any statistical analysis: the properties of probability distributions on graphs of different sizes that are finitely exchangeable need not be related to each other in any meaningful way (or in any way at all); see [27, 7, 29].

---

*Corresponding author.

Email addresses: `lauritzen@math.ku.dk` (S.L. Lauritzen), `arinaldo@cmu.edu` (A. Rinaldo), `k.sadeghi@ucl.ac.uk` (K. Sadeghi)

In this article, we investigate the relationship between exchangeability of random finite graphs and exchangeability of random infinite graphs using a combination of simple geometric arguments and standard graph-theoretic concepts. Our work can be seen as an extension to the graph setting of the geometric representation of finite exchangeability for random binary sequences developed by [8]. We make the following contributions: (1) we formulate a finite de Finetti's theorem for random graphs that is both elementary and rely on well known graph-theoretic quantities (namely, density homomorphisms) only; (2) we extend this result to obtain a simple derivation of the well-known de Finetti's representation theorem for exchangeable distributions on (infinite) graphs; (3) we provide novel geometric characterizations of all the finite marginals of exchangeable distributions on finite graphs and discuss the implications of our findings.

Related Work. There is a vast literature on exchangeability of random arrays, of which random graphs are a special case; see, e.g., [2, 3], [12], [15], [17], [19], [16] and [28], to name a few. Of particular significance is [10] [but see also 23], which details the connections between exchangeability of random graphs and the notion of graph limits developed in [6] and [21]; see also the book [20]. Similarly, finite exchangeability for sequences and arrays has been thoroughly investigated: see [8], [9], [2] and, in particular, [22]; see also [31].

In the companion paper [18], we rely on tools from the theory of graphical models to study the Markov properties of finitely exchangeable network models. The results derived there complement the ones we obtain in the present paper. We will discuss the connection between the two papers later in Section 5.

The article is organized in the following way. Section 2 describes the geometry of finitely exchangeable distributions on finite graphs and exchangeable distributions on infinite graphs and introduces the Möbius parametrization, which we will use throughout to represent probabilities on graphs. In Section 3 we provide definitions and basic results for homomorphism and isomorphism densities in order to derive a de Finetti's theorem for finitely exchangeable probability distributions on graphs based on the Möbius parametrization in Section 4. In Section 5 we study the manifold of dissociated and exchangeable random graphs and show that there exists dissociated and exchangeable random graphs that are not infinitely extendable.

Notation. For any integer $n \geq 2$ let $\mathcal{L}_n$ and $\mathcal{U}_n$ denote the set of simple labeled graphs and simple unlabeled graphs with node set $[n] := \{1, \ldots, n\}$, respectively, and set $\mathcal{L} = \bigcup_{n=2}^{\infty} \mathcal{L}_n$ and $\mathcal{U} = \bigcup_{n=2}^{\infty} \mathcal{U}_n$. We let $\mathcal{L}_{\infty}$ be the set of infinite simple labeled graphs. For any two graphs G and G' in $\mathcal{L}$, we will write $G \sim G'$ to signify that they are isomorphic and $[G]$ for the equivalence class of all graphs isomorphic to G. With a slight abuse of notation, we will at times identify an unlabeled graph with such an equivalence class $[G]$. We will also identify $\mathcal{L}_n$ with the Boolean algebra of all subsets of the node pairs $\{\{i, j\}, i \neq j\}$ partially ordered by inclusion by identifying each graph in $\mathcal{L}_n$ with the binary vector $\{0, 1\}^{\binom{n}{2}}$ representing its edges. With this identification, $\mathcal{L}_n$ indexes the coordinates of vectors in $\mathbb{R}^{2^{\binom{n}{2}}}$. If G and H are in $\mathcal{L}$, we write $H \subseteq G$ if H is a sub-graph (not necessarily induced) of G. For integers $2 \leq m \leq n \leq \infty$ and a $G \in \mathcal{L}_n$, $G[m]$ is the sub-graph of G induced by the nodes $[m]$.

In our analysis, we will often identify a graph G with its set of edges, hence ignoring isolated nodes. The set of labeled graphs on subsets of $[n]$ without isolated nodes is denoted by $\mathcal{I}_n$ (thus $G \in \mathcal{I}_n$ if and only if it is a subgraph of the complete graph on $[n]$ and has no isolated nodes) and we let $\mathcal{I} = \bigcup_{n=2}^{\infty} \mathcal{I}_n$. Similarly, we let $\mathcal{J}_n$ denote the set of unlabeled graphs without isolated nodes and at most n vertices and $\mathcal{J} = \bigcup_{n=2}^{\infty} \mathcal{J}_n$.

For any finite $n \geq 2$, if $G \in \mathcal{L}_n$ and $\sigma \in \mathcal{S}_n$, the permutation group on $[n]$, we will let G_σ be the graph obtained from G by relabeling its nodes according to σ. Thus nodes i and j are connected in G if and only if $\sigma(i)$ and $\sigma(j)$ are connected in G_σ. Finally, we let $\mathcal{S} = \bigcup_n \mathcal{S}_n$ be the set of all finite permutations.

2. Probabilities on graphs: exchangeability and geometry

We begin by introducing the notions of exchangeable distributions on networks and illustrating their geometry properties.

For finite n, the set of all probability distributions on $\mathcal{L}_n$ can be represented geometrically as the probability simplex in $\mathbb{R}^{\mathcal{L}_n}$, denoted with Δ_n. The coordinates of each vector $p \in \Delta_n$ are indexed by the graphs in $\mathcal{L}_n$, and the $2^{\binom{n}{2}}$ vertices of Δ_n are the unit masses at each $G \in \mathcal{L}_n$.

To formally define exchangeability, we first introduce the notion of **marginal mapping**: for any pair of integers $2 \leq m \leq n$, this mapping is defined to be the function $\Pi_n^m : \Delta_n \to \Delta_m$ mapping any probability distribution $p_n \in \Delta_n$ into the probability distribution $\Pi_m^n(p^n) = p_n^m \in \Delta_m$ given by

$$p_n^m(H) = \sum_{G \in \mathcal{L}_n \,:\, H = G[m]} p_n(G), \quad H \in \mathcal{L}_m. \tag{1}$$

With a slight abuse of notation we shall also think of of each $p_n \in \Delta_n$ as a measure and write $p(G[m])$ for the induced distribution on subgraphs:

$$p(G[m]) = p_n^m(G[m]),$$

where $p_n^m = \Pi_m^n(p)$

Definition 2.1. A probability distribution p on $\mathcal{L}_n$ is m-**exchangeable** when $p(G[m]) = p(G[m]_\sigma)$ for all $\sigma \in \mathcal{S}_m$ and all $G \in \mathcal{L}_n$. Equivalently, p is m-exchangeable when $p(G[m]) = p(G[m]')$ if $G[m] \sim G[m]'$ in $\mathcal{L}_m$. If $m = n$ we say that p is (finitely) **exchangeable**.

We denote with $\mathcal{E}_n \subset \Delta_n$ the set of all exchangeable distributions on $\mathcal{L}_n$. It is easy to show that exchangeable distributions are mixtures of uniform distributions over isomorphic classes. In fact, $\mathcal{E}_n$ is affinely isomorphic to the probability simplex in $\mathbb{R}^{\mathcal{U}_n}$ so that $\mathcal{E}_n$ is a polytope of dimension $|\mathcal{U}_n| - 1$.

Lemma 2.2. $\mathcal{E}_n$ *is a simplex whose vertices are the uniform probability distributions over isomorphic classes of $\mathcal{L}_n$. The dimension of $\mathcal{E}_n$ is equal to $|\mathcal{U}_n| - 1$.*

Proof. For a given class $[H] \in \mathcal{U}_n$ let $p_{[H]}$ be the probability distribution on $\mathcal{L}_n$ corresponding to the uniform distribution over $[H]$. That is, for any $G \in \mathcal{L}_n$,

$$p_{n,[H]}(G) = \begin{cases} \frac{1}{|[H]|} & \text{if } G \in [H], \\ 0 & \text{otherwise.} \end{cases} \tag{2}$$

Then, $p_{[H]} \in \mathcal{E}_n$ for all $[H]$. The vectors $\{p_{[H]}, [H] \in \mathcal{U}_n\}$ are affinely independent, because they are supported on incomparable subsets of $\mathcal{L}_n$, regarded as a poset with respect to the subset inclusion. Thus, their convex hull is a simplex inside $\mathcal{E}_n$. We will show that this simplex in fact coincides with $\mathcal{E}_n$. Let p be any point in $\mathcal{E}_n$. By exchangeability, for any $[H] \in \mathcal{U}_n$, the value of p at each of the coordinates indexed by the graphs in the isomorphism class $[H]$ is the same. Thus,

$$p = \sum_{[H] \in \mathcal{U}_n} v_{[H]} p_{[H]},$$

for some sequence $\{v_{[H]}, [H] \in \mathcal{U}_n\}$ of non-negative numbers. Since $\sum_{\{G \in \mathcal{L}_n\}} p(G) = 1$, it follows that $\sum_{[G]} v_{[G]} = 1$ and, therefore, that p is in the convex hull of $\{p_{[H]}, [H] \in \mathcal{U}_n\}$. Furthermore, since the convex hull of the vectors $\{p_{[H]}, [H] \in \mathcal{U}_n\}$ is a simplex, the sequence $\{v_{[H]}, [H] \in \mathcal{U}_n\}$ is unique.

Consistent models and exchangeability

A highly desirable property of a probability distribution for network data of a given size, say m, is that the distribution be realized as the marginal of probability distributions over networks of larger sizes n for all $n > m$. We refer to this property as probabilistic consistency.

Definition 2.3. A sequence $\{p_n\}_{n=2}^{\infty}$ of probability distributions such that $p_n \in \Delta_n$ for all n is **consistent** if

$$p_m = \Pi_n^m p_n, \quad \forall\, 2 \le m \le n. \tag{3}$$

If a probability distribution over networks of a given size is not part of a consistent sequence, then its properties may not be related in any meaningful way to the properties of any probability distribution over networks of different sizes.

When applied to an exchangeable distribution in $\mathcal{E}_n$, the marginal mapping Π_n^m always yields an exchangeable distribution in $\mathcal{E}_m$. However, Π_n^m is not surjective: there are exchangeable distributions in $\mathcal{E}_m$ that cannot be obtained as marginals of any exchangeable distribution on $\mathcal{E}_n$, for all $n > m > 3$. We formally state this fact in the next result and illustrate it in Theorem 2.6.

Lemma 2.4. *For all integers $4 \le m < n_1 < n_2$, it holds that $\Pi_{n_2}^m(\mathcal{E}_{n_2}) \subsetneq \Pi_{n_1}^m(\mathcal{E}_{n_1}) \subsetneq \mathcal{E}_m$.*

Proof. We will make use of the following graph-theoretic results, due to [1]:

Lemma 2.5. *Let H be a graph with n nodes.*

1. *All the induced subgraphs of H with a fixed but arbitrary number of m nodes, where $2 \leq m \leq n - 2$, are isomorphic if and only if H is a complete or empty graph.*

2. *All the induced subgraphs of H with $n - 1$ nodes are isomorphic if and only if H is a **node-transitive graph**, that is for any two nodes v_1 and v_2, there is some automorphism t such that $t(v_1) = v_2$.*

We will first show that $\Pi_n^m(\mathcal{E}_n) \subsetneq \mathcal{E}_m$ for all $4 \leq m < n$ The linearity of the marginal mapping Π_n^m, $n > m$, implies that $\Pi_n^m(\mathcal{E}_n)$ is a polytope whose vertices are contained in the image under Π_n^m of the vertices of $\mathcal{E}_n$. Thus, consider a vertex of $\mathcal{E}_n$, which, by Lemma 2.2, can be represented by an undirected graph on n nodes, say U. Such a vertex is mapped by Π_n^m into a distribution giving positive probabilities to only induced subgraphs of U of size m. This is a convex combination of uniform distributions over the labeled version of each of the induced subgraphs, which are vertices of $\mathcal{E}_m$. Hence

$$\Pi_n^m(\mathcal{E}_n) \subseteq \mathcal{E}_m. \tag{4}$$

We now show that $\Pi_n^m(\mathcal{E}_n)$ is a strict subset of $\mathcal{E}_m$. To see this, notice that a vertex of $\mathcal{E}_n$ is mapped into a vertex of $\mathcal{E}_m$ if and only if it corresponds to the uniform distribution over isomorphic graphs in $\mathcal{G}_n$ such that all induced subgraphs obtained by removing any set of $n - m$ nodes are isomorphic. By Lemma 2.5, if $n - m = 1$ this condition is satisfied by all node-transitive graphs and if $n - m > 2$ only by the empty and complete graphs. This proves that the inclusion (4) is strict.

We will now prove that $\Pi_{n_2}^m(\mathcal{E}_{n_2}) \subsetneq \Pi_{n_1}^m(\mathcal{E}_{n_1})$, for all integers $4 \leq n_1 < n_2$. Since $\Pi_{n_2}^m(\mathcal{E}_{n_2}) = \Pi_{n_1}^m\left(\Pi_{n_2}^{n_1}(\mathcal{E}_{n_2})\right)$ and, as we just saw, $\Pi_{n_2}^{n_1}(\mathcal{E}_{n_2}) \subsetneq \mathcal{E}_{n_1}$, it holds that $\Pi_{n_2}^m(\mathcal{E}_{n_2}) \subseteq \Pi_{n_1}^m(\mathcal{E}_{n_1})$. Thus, we only need to verify that the previous inclusion is strict. This, in turn, will follow if we exhibit a vertex p of $\mathcal{E}_{n_1}$ that (i) is not in the image under $\Pi_{n_2}^{n_1}$ of $\mathcal{E}_{n_2}$ and (ii) such that $\Pi_{n_1}^m(p)$ is a vertex of $\Pi_{n_1}^m(\mathcal{E}_{n_1})$. We choose p to be the uniform distribution over graphs in $\mathcal{L}_{n_1}$ that are isomorphic to the node-disjoint union of the complete graph on $n_1 - 1$ nodes and one isolated node. By definition, this is a vertex of $\mathcal{E}_{n_1}$ and, by Theorem 2.5, is not in $\Pi_{n_2}^{n_1}(\mathcal{E}_{n_2})$, since it does not belong to the image under $\Pi_{n_2}^{n_1}$ of the vertices of $\mathcal{E}_{n_2}$. Next, $\Pi_{n_1}^m(p)$ obviously belongs to $\mathcal{E}_m$ and can be expressed as the mixture $\frac{m}{n_1}p' + \frac{n_1 - m}{n_1}p''$. Here, p' the uniform distribution over all graphs in $\mathcal{L}_m$ that are isomorphic to the node-disjoint union of a complete graph on $m - 1$ nodes and one isolated node, and p'' is the point mass at the complete graph in $\mathcal{L}_m$. In particular, $\Pi_{n_1}^m(p)$ must be a vertex of $\Pi_{n_1}^m(\mathcal{E}_{n_1})$. To see this, the node-disjoint union of a complete graph on $m - 1$ nodes and one isolated node is not node-transitive, and, by Theorem 2.5, cannot be a vertex of $\Pi_{n_1}^m(\mathcal{E}_{n_1})$. Since $\Pi_{n_1}^m(p)$ is the only point in $\Pi_{n_1}^m(\mathcal{E}_{n_1})$ that has such a mixture representation and is the image of a vertex of $\mathcal{E}_{n_1}$, the claim follows. Thus p satisfies both properties (i) and (ii) and the proof is complete.

Example 2.6. Let p_5 the uniform distribution on $\mathcal{L}_5$ that assigns probability 1/15 over all graphs isomorphic to the union of a four cycle and an isolated node. Then, its image p_5^4 in $\mathcal{E}_4$ under the marginal mapping is the convex combinations of two vertices of $\mathcal{E}_4$:

the uniform distributions over the 3 graphs isomorphic to the 4-cycle and the uniform distribution over the 12 graphs isomorphic to the union of a 2-star and an isolated node. The weights of this mixtures are 3/15 and 12/15, respectively. On the other hand, $P_5^4(\mathcal{E}_5)$ does not contain those two vertices of $\mathcal{E}_4$, verifying that $P_5^4(\mathcal{E}_5)$ is a strict subset of $\mathcal{E}_4$. Furthermore, the point p_5^4 happens to be a vertex of $P_5^4(\mathcal{E}_5)$. To see this, it is enough to observe that, for each of the other 33 unlabeled graphs in $\mathcal{U}_5$, the set of induced subgraphs obtained by removing any one node is different than (and is never contained in) the set consisting of the 4-cycle and the union of a 2-star and an isolated node. Thus, it is not possible to represent p_4^5 as a convex combination of the marginals of uniform distributions over isomorphic graphs on 5 nodes. Since each vertex of $P_5^4(\mathcal{E}_5)$ is the image of some vertex of $\mathcal{E}_5$, the claim follows.

The previous example has led us to conjecture that, for $4 \leq m < n$, each vertex of $\mathcal{E}_n$ is mapped into a vertex of $\Pi_n^m(\mathcal{E}_n)$. When $m = 3$ and $n = 4$ this is clearly not true, since in this case it is easy to see that $\Pi_4^3(\mathcal{E}_4) = \mathcal{E}_3$, which explains the requirement that $m \geq 4$ in Theorem 2.4. For example, the uniform distribution over graphs in $\mathcal{L}_4$ isomorphic to the 3-path is not a vertex of $\Pi_4^3(\mathcal{E}_4)$.

Theorem 2.4 implies that a sequence of finitely exchangeable probability distributions on graphs need not be consistent. That is, if one poses a finitely exchangeable distribution p_n on $\mathcal{L}_n$, while all its marginals will be exchangeable, there is no guarantee *a priori* that p_n can be realized as the marginal of any exchangeable distribution on larger graphs.

In order for a sequence of finitely exchangeable distribution on graphs to be consistent, finite exchangeability needs to be replaced by the stronger notion of exchangeability, which we define next. We remark that, though our definition may appear different from the classic definition of row and column exchangeability of symmetric random binary arrays as in, e.g., [28], [2], [9] and [19], it is in fact equivalent. Below, we will use the symbol "$\overset{d}{=}$" to denote identity in distribution.

To present exchangeability, we first notice that the definition of marginal map can be extended in a straightforward manner to distributions on $\mathcal{L}_\infty$[†]: for any integer $m \geq 2$ and any probability distribution p_∞ over $\mathcal{L}_\infty$, Π_∞^m takes p_∞ into the distribution $\Pi_\infty^m(p_\infty) = p_\infty^m$ in Δ_m given by

$$p_\infty^m(H) = \mathbb{P}(G[m] = H), \quad H \in \mathcal{L}_m,$$

where G is the random graph in $\mathcal{L}_\infty$ with distribution p_∞. By slightly abusing notation again, for any $p \in \mathcal{L}_\infty$ and $G \in \mathcal{L}_m$, we write $p(G)$ for $p_\infty^m(G)$, where $p_\infty^m = \Pi_\infty^m(p)$.

Definition 2.7. A probability distribution p on $\mathcal{L}_\infty$ is exchangeable when $G \overset{d}{=} G_\sigma$, where G denotes the random graph in $\mathcal{L}_\infty$ with distribution p, for all $\sigma \in \mathcal{S}$. Equivalently, p is exchangeable when, for any pair G and G' of isomorphic graphs in $\mathcal{L}$, $p(G) = p(G')$.

It follows that all the finite marginals $\{p_n\}_{n=1,2,\dots}$ of an exchangeable distribution define a consistent sequence of finitely exchangeable distributions. Conversely, by the

---

[†]Unlike the set $\mathcal{L}_n$, which is finite for each n, $\mathcal{L}_\infty$ is uncountable. Viewed as the product set $\{0,1\}^{E(K_\infty)}$, where $E(K_\infty)$ denotes all subsets of edges of an infinite complete graph, $\mathcal{L}_\infty$ is a compact metric space under the product topology. Thus, Borel probability measures are well defined on $\mathcal{L}_\infty$.

Kolmogorov-Bochner extension theorem [see, e.g. 24], the existence of a consistent family of finitely exchangeable distribution will guarantee the existence of an exchangeable distribution on $\mathcal{L}_\infty$.

We let $\mathcal{E}_\infty$ denote the set of all exchangeable distributions on $\mathcal{L}_\infty$. $\mathcal{E}_\infty$ can be identified with a compact subset of $[0,1]^{\mathcal{I}}$ and is a Bauer simplex; see (26) and Section 4.1.1 below. One of our goals in this article is to describe the relationship between finite exchangeability and exchangeability in the present setting. In particular, we seek a geometric characterization of the subset of $\mathcal{E}_m$ given by

$$\lim_{n \to \infty} \Pi_n^m(\mathcal{E}_n) = \bigcap_n \Pi_n^m(\mathcal{E}_n) = \Pi_\infty^m(\mathcal{E}_\infty),$$

which, in light of Theorem 2.4, is a well-defined closed set. We provide a partial solution in Section 5.

The Möbius parametrization

Though canonical, the parametrization corresponding to the set $\mathcal{E}_n$ is not the most convenient. As we will see, exchangeable distributions on graphs are better represented using marginal, as opposed to joint, probabilities. We will refer to this parametrization as the **Möbius parametrization,** which we describe next. We take note that this is not a novel parametrization: it is simply the adaptation to the network setting of the well-known representation of multivariate binary distributions by means of the Möbius inversion formula.

Let M_n be the square matrix of dimension $2^{\binom{n}{2}}$ with entries indexed by graphs in $\mathcal{L}_n$ and given by

$$M_n(F, G) = 1(F \subseteq G), \quad F, G \in \mathcal{L}_n. \tag{5}$$

Then, M_n has full rank [see, e.g., 30] and its inverse has entries

$$M_n^{-1}(F, G) = (-1)^{|G \setminus F|} 1(F \subseteq G), \quad F, G \in \mathcal{L}_n, \tag{6}$$

where, for $F \subseteq G$, $|G \setminus F|$ is the number of edges G has in excess of F. Borrowing the terminology from [11], we define the **Möbius simplex** to be the set

$$\Delta_n^M = \{M_n p, \, p \in \Delta_n\}.$$

As Δ_n^M and Δ_n are in one-to-one correspondence with each other, Δ_n^M is a valid parametrization of all the probability distribution on $\mathcal{L}_n$.

The form and probabilistic interpretation of Δ_n^M are of course quite different from those of Δ_n. Indeed, we will index the coordinates of the point in Δ_n^M by the elements of $\mathcal{I}_n$, which we recall is the set of labeled graphs on subsets of $[n]$ without isolated nodes. For a $p \in \Delta_n$ and a $z \in \Delta_n^M$ with $z = M_n p$, the value of z at any such graph F is just the **marginal probability** that F is a subgraph of a random graph drawn from p. That is,

$$z(F) = \sum_{H \in \mathcal{L}_n \,:\, F \subseteq H} p(H) = \mathbb{P}\left(F \subseteq G\right), \quad F \in \mathcal{I}_n, \tag{7}$$

where G is a random graph in $\mathcal{L}_n$ with distribution p. In particular, the values of z conform to the partial order on $\mathcal{I}_n$: $z(F) \geq z(F')$ if $F \subseteq F'$. Since the Möbius transform (5) is linear and invertible, Δ_n^M is also a polytope (in fact a simplex), whose vertices are the image by M_n of the vertices of Δ_n: $2^{\binom{n}{2}} - 1$ vertices indexed by all non-empty graphs in $\mathcal{I}_n$ and the vector $1 \in \mathbb{R}^{2^{\binom{n}{2}}}$, the Möbius transform of the point mass on the complete graph.

The Möbius parametrization enjoys the following property, referred to as **backward compatibility**. For $m \leq n$, let $p_n \in \Delta_n$ and $p_n^m = \Pi_n^m p_n \in \Delta_m$ be its marginal, and let z_n and z_n^m denote their Möbius transforms. Then,

$$z_n^m(F) = z_n(F), \quad \forall F \in \mathcal{I}_m. \tag{8}$$

Backward compatibility is a direct consequence of the fact that the Möbius parameters are marginal probabilities.

The image of the simplex $\mathcal{E}_n$ of finitely exchangeable probability distributions by the Möbius transform is also a polytope (in fact, a simplex) of the same dimension, denoted by $\mathcal{E}_n^M$, whose vertices are the Möbius transform of the vertices of $\mathcal{E}_n$. Clearly, the vertices of $\mathcal{E}_n^M$ can also be indexed by $\mathcal{I}_n$. By exchangeability and Equation (7), for each $z \in \mathcal{E}_n^M$, $z(F) = z(F')$ whenever $F \sim F'$ in $\mathcal{I}_n$. In fact, using Equation (6), $\mathcal{E}_n^M$ is defined geometrically by these linear constraints, the linear constraint that $z(\emptyset) = 1$ (where $\emptyset$ signifies the empty graph), and the facet defining inequalities

$$\sum_{F' \in \mathcal{I}_n: \, F \subseteq F'} (-1)^{|F' \setminus F|} z(F') \geq 0, \quad \forall F \in \mathcal{I}_n, F \neq \emptyset. \tag{9}$$

3. Homomorphism and isomorphism densities

Next, we recall some graph-theoretic quantities that will play a key role in our derivations. It is not a coincidence that these very same quantities are also used in the theory of graph limits. See, e.g., [20].

Let $G \in \mathcal{L}_n$ and $F \in \mathcal{L}_m$, where $m \leq n$ (this last requirement is not technically necessary; however we will assume it throughout). The **homomorphism density** of H in G is

$$t_{\hom}(F, G) = \frac{\hom(F, G)}{n^m}, \tag{10}$$

and is equal to the fraction of all mappings from $[m]$ into $[n]$ that define a homomorphism (adjacency preserving mapping) between F and G. Density homomorphisms are **multiplicative:**

$$t_{\hom}(F_1 F_2, G) = t_{\hom}(F_1, G) t_{\hom}(F_2, G), \tag{11}$$

where $F_1 F_2$ is the node-disjoint union of F_1 and F_2. As we will see, this is the graph-theoretical counterpart to a fundamental probabilistic property known in the literature on exchangeable arrays as the dissociated property.

A related concept is that of the **injective homomorphism density** of F in G,

$$t_{\text{inj}}(F, G) = \frac{\text{inj}(F, G)}{(n)_m}, \tag{12}$$

where $(n)_m = n!/(n - m)!$, and $\text{inj}(F, G)$ is the number of injective mappings from $[m]$ into $[n]$ that define a homomorphism between F and G.

Remark 3.1. If F has isolated nodes, the values of both $t_{\text{hom}}(F, G)$ and $t_{\text{inj}}(F, G)$ do not change if F is replaced by the smaller sub-graphs induced by the nodes of positive degree. Therefore, there is no loss of generality in assuming that $F \in \mathcal{I}_n$ when dealing with the quantities in (10) and (12).

In a similar manner, for $G \in \mathcal{L}_n$ and $F \in \mathcal{L}_m$, we define the **isomorphism density** of F in G as

$$t_{\text{iso}}(F, G) = \frac{\text{iso}(F, G)}{n^m},$$

where $\text{iso}(F, G)$ is the number of maps from $[m]$ into $[n]$ that preserve both adjacency and non-adjacency, i.e., such that the induced subgraph of G is isomorphic to F. Finally, let

$$t_{\text{ind}}(F, G) = \frac{\text{ind}(F, G)}{(n)_m} \tag{13}$$

be the injective isomorphism density, where $\text{ind}(F, G)$ is the number of injective mappings from $[m]$ into $[n]$ that preserve both adjacency and non-adjacency, i.e. the number of isomorphisms from F into induced subgraphs of G with m nodes.

The next two results, whose proofs are straightforward and therefore omitted, provide a more statistically transparent interpretation of homomorphism and isomorphism densities. The difference between injective and non-injective densities is precisely the difference between sampling with and without replacement.

Lemma 3.2. *Fix a $G \in \mathcal{L}_n$. Let $(U_1, \ldots, U_m)$ be independent random variables uniformly distributed over $[n]$ and H be the random graph on $[m]$ where $i \sim j$ in H if and only if $U_i \sim U_j$ in G. Then, for any graph $F \in \mathcal{L}_m$,*

$$t_{\text{hom}}(F, G) = \mathbb{P}(F \subseteq H) \quad \text{and} \quad t_{\text{iso}}(F, G) = \mathbb{P}(F = H).$$

Let $(U_1', \ldots, U_m')$ be the sequence of random variables describing the outcomes of m draws without replacements of n labelled equiprobable balls. Let H be a random graph on m nodes such that $i \sim j$ if and only if $U_i' \sim U_j'$ in G. Then, for any graph $F \in \mathcal{L}_m$,

$$t_{\text{inj}}(F, G) = \mathbb{P}'(F \subseteq H) \quad \text{and} \quad t_{\text{ind}}(F, G) = \mathbb{P}'(F = H).$$

Using the above representation we immediately obtain the following well-known bound on the difference between subgraph densities arising from injective and non-injective mappings. These bounds will also be used in the proof of Theorem 4.1.

Lemma 3.3. *Let $G \in \mathcal{L}_n$ and $m \leq n$. For any $A \subseteq \mathcal{L}_m$, set*

$$t_{\mathrm{ind}}(A, G) = \sum_{F \in A} t_{\mathrm{ind}}(F, G) \quad and \quad t_{\mathrm{iso}}(A, G) = \sum_{F \in A} t_{\mathrm{iso}}(F, G). \tag{14}$$

Then,

$$\sup_{A \subseteq \mathcal{L}_m} \left| t_{\mathrm{iso}}(A, G) - t_{\mathrm{ind}}(A, G) \right| \leq 1 - \frac{(n)_m}{n^m}. \tag{15}$$

As a result,

$$\sup_{F \in \mathcal{L}_m} \left| t_{\mathrm{hom}}(F, G) - t_{\mathrm{inj}}(F, G) \right| \leq 1 - \frac{(n)_m}{n^m}. \tag{16}$$

Proof. It is enough to prove (15), since (16) clearly follows from it. By Theorem 3.2, (15) can be established by a well known bound on the total variation distance between a sample with and without replacement: see, e.g., [14]. Here we give an alternative proof based on sub-graph densities, and inspired by the arguments used in [22]. Let $\mathrm{notinjiso}(F, G)$ denote the number of non-injective mappings from $[k]$ into $[n]$ that define isomorphisms between F and G, so that $\mathrm{iso}(F, G) = \mathrm{notinjiso}(F, G) + \mathrm{ind}(F, G)$. For any $A \subseteq \mathcal{A}_m$, let $\mathrm{notinjiso}(A, G) = \sum_{F \in A} \mathrm{notinjiso}(F, G)$. By Theorem 3.2, both $t_{\mathrm{ind}}(A, G)$ and $t_{\mathrm{iso}}(A, G)$ are probabilities and, therefore, take values in $[0, 1]$. Thus,

$$
\begin{aligned}
t_{\mathrm{iso}}(A, G) - t_{\mathrm{ind}}(A, G) &= \frac{\mathrm{notinjiso}(A, G) + \mathrm{ind}(A, G)}{n^k} - \frac{\mathrm{ind}(A, G)}{(n)_k} \\
&= \frac{\mathrm{notinjiso}(A, G)}{n^k} - \frac{\mathrm{ind}(A, G)}{(n)_k} \left(1 - \frac{(n)_k}{n^k} \right).
\end{aligned}
$$

Since, trivially, $0 \leq \mathrm{notinjiso}(A, G) \leq n^k - (n)_k$, we obtain that $0 \leq \frac{\mathrm{notinjiso}(A,G)}{n^k} \leq \left(1 - \frac{(n)_k}{n^k} \right)$. Using both bounds in the previous display yields that

$$-t_{\mathrm{ind}}(A, G) \left(1 - \frac{(n)_k}{n^k} \right) \leq t_{\mathrm{iso}}(A, G) - t_{\mathrm{ind}}(A, G) \leq (1 - t_{\mathrm{ind}}(A, G)) \left(1 - \frac{(n)_k}{n^k} \right).$$

The claimed bound follows since $0 \leq t_{\mathrm{ind}}(A, G) \leq 1$.

Remark 3.4. The above bound can be weakened to the simpler bound $\binom{m}{2}/n$. See also Lemma 2.1 in [21].

The value of $t_\circ(F, G)$ remains unchanged if one or both of its arguments F and G are replaced by isomorphic graphs $F' \sim F$ and $G' \sim G$, where $t_\circ$ is any of the densities introduced above. Thus, these graph densities remain well defined if one or both of their arguments belong to $\mathcal{U}$. The next result uses this fact to establish a correspondence between injective densities and the concepts introduced in Section 2. It will be used in the proof of Theorem 4.1.

Lemma 3.5. *Let p_n be the vertex of $\mathcal{E}_n$ corresponding to the uniform distribution over the class $[G]$ and p_n^m its image under the marginal mapping Π_n^m, where $2 \leq m < n$. Let z_n^m be the Möbius transform of p_n^m. Then, for all $F \in \mathcal{I}_m$, $z_n^m(F) = t_{\mathrm{inj}}(F, [G])$ and, for all $F \in \mathcal{L}_m$, $p_n^m(F) = t_{\mathrm{ind}}(F, [G])$.*

Proof. We will give a proof only for the identity involving the injective homomorphism density, since the same arguments apply to the one involving the injective isomorphism density. Let $[H] \in \mathcal{U}_n$ be a given isomorphism class in $\mathcal{L}_n$ and $p_{[H]}$ the point in $\mathcal{E}_n$ corresponding to the uniform distribution over $[H]$. For a given $F \in \mathcal{I}_m$, let $B = \{G \in \mathcal{L}_n \colon F \subseteq G\}$. For any $\sigma \in \mathcal{S}_n$, let $\sigma^{-1}(B) = \{G \in \mathcal{L}_n \colon F \subseteq G_\sigma\}$, that is, the set of G such that $G_\sigma \in B$. With a slight abuse of notation we write $p_{[H]}(B) = \sum_{G \in B} p_{[H]}(G)$. Then,

$$
\begin{aligned}
z_{[H]}^m(F) &= z_{[H]}(F) \\
&= \sum_{G \in \mathcal{L}_n} 1(F \subseteq G) p_{[H]}(G) \\
&= p_{[H]}(B) \\
&= \frac{1}{n!} \sum_{\sigma \in \mathcal{S}_n} p_{[H]}(\sigma^{-1}(B)) \\
&= \frac{1}{n!} \sum_{G \in [H]} \sum_{\sigma \in \mathcal{S}_n} 1(F \subseteq G_\sigma) p_{[H]}(G) \\
&= \sum_{G \in [H]} \frac{|\{\sigma \in \mathcal{S}_n \colon F \subseteq G_\sigma\}|}{n!} p_{[H]}(G) \\
&= t_{\mathrm{inj}}(F, [H]) \sum_{G \in [H]} p_{[H]}(G) \\
&= t_{\mathrm{inj}}(F, [H]).
\end{aligned}
$$

The first identity follows from the backward compatibility of the Möbius transform, the fourth identity follows from exchangeability and the last identity uses the facts that, for any $\sigma \in \mathcal{S}_n$, $G_\sigma \in [H]$ if and only if $G \in [H]$ and that $t_{\mathrm{inj}}(F, \cdot)$ is constant over $[H]$ (with the common value denoted as $t_{\mathrm{inj}}(F, [H])$).

4. de Finetti's theorems for exchangeable distributions on graphs

In this section we will use the sub-graph densities introduced in Section 3 to derive a de Finetti's theorem for finitely exchangeable probability distributions on graphs based on the Möbius parametrization. The results show that the Möbius parameters of a finitely exchangeable distribution on $\mathcal{L}_m$ that extends to a finitely exchangeable distribution on $\mathcal{L}_n$, where $2 \leq m < n$, are the expected values of the injective density homomorphisms. These are approximated uniformly well by the expected values of the density homomorphisms with the approximation error of order $O(m^2/n)$. The proof is a simple application of Theorem 3.5 and Theorem 3.3 is the graph-theoretical counterpart of Theorem 1 in [22].

Theorem 4.1 (de Finetti's theorem for finitely exchangeable distributions on graphs). *Let $p_n \in \mathcal{E}_n$ and $p_n^m = \Pi_n^m p_n$ where $m \leq n$. Let z_n^m be the corresponding Möbius transform of p_n^m. Then, for any subgraph $F \in \mathcal{I}_m$,*

$$z_n^m(F) = \sum_{G \in \mathcal{L}_n} t_{\mathrm{inj}}(F, G) p_n(G) \tag{17}$$

and

$$\max_{F \in \mathcal{I}_m} \left| z_n^m(F) - \sum_{G \in \mathcal{L}_n} t_{\mathrm{hom}}(F, G) p_n(G) \right| \leq 1 - \frac{(n)_m}{n^m}. \tag{18}$$

Proof. The proof relies on Lemma 3.3 and can be regarded as extension to the network setting of the geometric arguments used in [8]. See also [9] and [17], and in particular, [22].

Far any isomorphism class $[H] \in \mathcal{U}_n$, and any $F \in \mathcal{I}_m$,

$$t_{\mathrm{inj}}(F, G) = t_{\mathrm{inj}}(F, G'), \quad \forall G, G' \in [H]. \tag{19}$$

As before, $t_{\mathrm{inj}}(F, [H])$ denotes the common value of $t_{\mathrm{inj}}(F, G)$ for all graphs $G \in [H]$.

Next, let $p_{n,[H]}$ be the finitely exchangeable probability distribution on $\mathcal{L}_n$ corresponding to the uniform distribution over the isomorphic class $[H]$, as described in Equation (2). Since $p_{n,[H]}$ is a vertex of $\mathcal{E}_n$ by Theorem 2.2, its Möbius transform $z_{n,[H]}$ is a vertex of $\mathcal{E}_n^M$. Then, because $\mathcal{E}_n^M$ is a simplex, any point in $z_n \in \mathcal{E}_n^M$ can be written as

$$z_n = \sum_{[H] \in \mathcal{U}_n} w_{[H]} z_{n,[H]},$$

for a unique sequence of non-negative numbers $\{w_{[H]}, [H] \in \mathcal{U}_n\}$ such that $\sum_{[H] \in \mathcal{U}_n} w_{[H]} = 1$.

Let $p_{n,[H]}^m = \Pi_n^m p_{n,[H]}$ be the probability distribution over $\mathcal{L}_m$ obtained by marginalizing over $p_{n,[H]}$ and $z_{n,[H]}^m$ be its Möbius transform. Then, for any $F \in \mathcal{I}_m$,

$$z_{n,[H]}^m(F) = t_{\mathrm{inj}}(F, [H]) = \sum_{G \in [H]} \frac{1}{[H]} t_{\mathrm{inj}}(F, G), \tag{20}$$

where the first identity follows from Theorem 3.5.

As a result, for any $F \in \mathcal{I}_m$,

$$\begin{aligned}
z_n^m(F) &= \sum_{[H] \in \mathcal{U}_n} w_{[H]} z_{n,[H]}^m(F) \\
&= \sum_{[H] \in \mathcal{U}_n} w_{[H]} t_{\mathrm{inj}}(F, [H]) \\
&= \sum_{[H] \in \mathcal{U}_n} w_{[H]} \left(\sum_{G \in [H]} \frac{1}{|[H]|} t_{\mathrm{inj}}(F, G) \right)
\end{aligned}$$

$$= \sum_{[H]\in\mathcal{U}_n} w_{[H]} \left(\sum_{G\in[H]} p_{n,[H]}(G) t_{\mathrm{inj}}(F,G) \right)$$

$$= \sum_{G\in\mathcal{G}_n} t_{\mathrm{inj}}(F,G) \left(\sum_{[H]\in\mathcal{U}_n} w_{[H]} p_{n,[H]}(G) \right)$$

$$= \sum_{G\in\mathcal{G}_n} t_{\mathrm{inj}}(F,G) p_n(G),$$

where the first, second and fourth identities follow from the linearity of marginal operation, Equation (20), and Equation (2), respectively. Thus, Equation (17) follows.

Using the previous identity, for a given $F\in\mathcal{I}_m$ of size, say, k,

$$\left| z_n^m(F) - \sum_{G\in\mathcal{G}_n} t_{\mathrm{hom}}(F,G) p_n(G) \right| = \left| \sum_{G\in\mathcal{G}_n} t_{\mathrm{inj}}(F,G) p_n(G) - \sum_{G\in\mathcal{G}_n} t_{\mathrm{hom}}(F,G) p_n(G) \right|$$

$$\leq \sum_{G\in\mathcal{G}_n} \left| t_{\mathrm{inj}}(F,G) - t_{\mathrm{hom}}(F,G) \right| p_n(G)$$

$$\leq 1 - \frac{(n)_k}{n^k},$$

where the last inequality is due to Theorem 3.3. Equation (18) is established by noting that

$$1 - \frac{(n)_k}{n^k} \leq 1 - \frac{(n)_m}{n^m}, \tag{21}$$

for all integer $k < m$.

The theorem further implies that, for any finitely exchangeable distributions on $\mathcal{L}_n$, the marginal probabilities of all its small sub-graphs are well approximated by a certain mixture of densities homomorphisms of such sub-graphs, with the mixing measure defined over isomorphisms class in $\mathcal{L}_n$. Formally, we have the following:

Corollary 4.2. *Assume $2 \leq m < n$. Let $p_n \in \mathcal{E}_n$ and z_n be its Möbius transform. Then, there exists a probability distribution $\{w_U, U \in \mathcal{U}_n\}$ on $\mathcal{U}_n$, uniquely determined by p_n, such that*

$$\left| z_n(F) - \sum_{U\in\mathcal{U}_n} w_U t_{\mathrm{hom}}(F,U) \right| \leq 1 - \frac{(n)_m}{n^m}, \quad \forall F \in \mathcal{I}_m.$$

Proof. By backward compatibility (8), $z_n^m(F) = z_n(F)$ for all $F \in \mathcal{I}_m$. Furthermore, by Theorem 2.2, each $p_n \in \mathcal{E}_n$ can be written as

$$p_n = \sum_{[G]\in\mathcal{U}_n} p_{n,[G]} w_{[G]},$$

for a unique probability distribution $\{w_{[G]}, [G] \in \mathcal{U}_n\}$ on $\mathcal{U}_n$, where $p_{n,[G]}$ is the uniform distribution over the class $[G] \in \mathcal{U}_n$. The claim follows from the fact that $t_{\mathrm{hom}}(F,\cdot)$ takes on the same value $t_{\mathrm{hom}}(F,[G])$ over $[G]$ and collecting terms.

By Theorem 2.2, p_n being extremal is equivalent to p_n being a uniform distribution over some isomorphism class, say $[G]$ in $\mathcal{L}_n$. If p_n is extremal then $w_{[G]} = 1$. Since homomorphism densities are multiplicative, we can use the fact that $1 - (n)_m/n^m \le \binom{m}{2}/n$ to conclude that, for n of larger order than m^2 and if p_n is an extremal distribution on $\mathcal{L}_n$,

$$z_n(F) \approx z_n(F_1)z_n(F_2), \tag{22}$$

for each $F \in \mathcal{I}_m$ of the form $F = F_1F_2$, where we recall that F_1F_2 is the vertex-disjoint union of F_1 and F_2. As we will show in the next section, the approximation in Equation (22) becomes an equality if p_n is embedded into a sequence of consistent finitely exchangeable distributions that extend to an extremal exchangeable distribution over $\mathcal{L}_\infty$. Furthermore, all such extremal distributions are defined by these identities.

A result analogous to Theorem 4.1 holds also for joint probabilities. We have chosen to focus on marginal probabilities since they are more natural in this context, as they directly lead to the key approximation property of Equation (22).

Corollary 4.3. *Consider the setting of Theorem 4.1. Then, for any $F \in \mathcal{L}_m$,*

$$p_n^m(F) = \sum_{G \in \mathcal{L}_n} t_{\mathrm{ind}}(F, G)p_n(G) \tag{23}$$

and, as a result,

$$\max_{F \in \mathcal{L}_m} \left| p_n^m(F) - \sum_{G \in \mathcal{L}_n} t_{\mathrm{iso}}(F, G)p_n(G) \right| \le 1 - \frac{(n)_m}{n^m}. \tag{24}$$

Furthermore, letting $\tilde{p}_n^m$ the probability distribution on $\mathcal{L}_m$ given by

$$\tilde{p}_n^m(F) = \sum_{G \in \mathcal{L}_n} t_{\mathrm{iso}}(F, G)p_n(G), \quad F \in \mathcal{L}_m,$$

we have

$$d_{\mathrm{TV}}\left(p_n^m, \tilde{p}_n^m\right) \le 1 - \frac{(n)_m}{n^m}, \tag{25}$$

where $d_{\mathrm{TV}}(P, Q)$ denotes the total variation distance between the probability distributions P and Q.

Proof. We omit the proofs of (23) and (24), since they are nearly identical to the proofs of (17) and (18) given above. To prove (25), let $A \subset \mathcal{L}_m$ and recall the quantities defined in eq. (14):

$$t_{\mathrm{ind}}(A, G) = \sum_{F \in A} t_{\mathrm{ind}}(F, G) \quad \text{and} \quad t_{\mathrm{iso}}(A, G) = \sum_{F \in A} t_{\mathrm{iso}}(F, G).$$

Notice that by Theorem 3.3, $|t_{\mathrm{ind}}(A, G) - t_{\mathrm{iso}}(A, G)| \le 1 - \frac{(n)_m}{n^m}$, for any $G \in \mathcal{L}_n$ and $A \subseteq \mathcal{L}_m$. Then,

$$\left| \sum_{F \in A} p_n^m(F) - \sum_{F \in A} \tilde{p}_n^m(F) \right| = \left| \sum_{F \in A} \left(\sum_{G \in \mathcal{G}_n} t_{\mathrm{ind}}(F, G)p_n(G) \right) - \sum_{F \in A} \left(\sum_{G \in \mathcal{G}_n} t_{\mathrm{iso}}(F, G)p_n(G) \right) \right|$$

$$= \left| \sum_{G \in \mathcal{L}_n} (t_{\mathrm{ind}}(A, G) - t_{\mathrm{iso}}(A, G)) \, p_n(G) \right|$$

$$\leq \sum_{G \in \mathcal{L}_n} |t_{\mathrm{ind}}(A, G) - t_{\mathrm{iso}}(A, G)| \, p_n(G)$$

$$\leq 1 - \frac{(n)_m}{n^m}.$$

Inequality (25) now follows since, by definition,

$$d_{\mathrm{TV}}(p_n^m, \tilde{p}_n^m) = \sup_{A \subseteq \mathcal{L}_m} \left| \sum_{F \in A} p_n^m(F) - \sum_{F \in A} \tilde{p}_n^m(F) \right|.$$

This completes the proof.

Just like in Theorem 4.2, we can equivalently express (23) as

$$p_n^m(F) = \sum_{U \in \mathcal{U}_n} t_{\mathrm{ind}}(F, U) w_U,$$

for a probability distribution $\{w_U, U \in \mathcal{U}_n\}$ that is uniquely determined by p_n.

4.1. From finite exchangeability to exchangeability

Below we will strengthen the conclusions of Theorem 4.1 by further assuming that each $p_n \in \mathcal{E}_n$ is an element of a sequence $\{p_n\}_{n=2}^{\infty}$ of finitely exchangeable distributions that are consistent, i.e. satisfy Equation (3). As noted above, each such sequence extends uniquely to one element in the simplex $\mathcal{E}_\infty$ of exchangeable probability distribution on $\mathcal{L}_\infty$. Below, we will establish a de Finetti's-type theorem for exchangeable distributions and, along the way, relate it to the theory of graph limits. This connection is well known and has been elucidated in [10].

We begin by introducing a few concepts that are necessary to represent distributions over infinite graphs and graph sequences. First, it is easy to see that any probability distribution on $\mathcal{E}_\infty$ admits a Möbius parametrization that is completely analogous to the one given for distributions of finite random graphs and are based on marginal probabilities of finite subgraphs without isolated nodes. In detail, for a point $p_\infty \in \mathcal{E}_\infty$ we will write

$$z = z(p_\infty) = (z(F), F \in \mathcal{I}) \in [0, 1]^{\mathcal{I}}, \tag{26}$$

for the sequence of Möbius parameters given by

$$z(F) = \mathbb{P}(F \subset G) \quad F \in \mathcal{I}, \tag{27}$$

where G is an infinite random graph with distribution p_∞ and we recall that $\mathcal{I}$ is the set of all finite graphs without isolated nodes. In particular, the Möbius parametrization

$(z(F), F \in \mathcal{I}) \subset [0,1]^{\mathcal{I}}$ of an exchangeable distribution on $\mathcal{L}_\infty$ satisfies the properties that $z(\emptyset) = 1$, $z(F) = z(F')$ if $F \sim F'$ and Equation (9) holds for all n.

We will also require the notion of graph limits: see [21], [6], [20]. Following [10], we let $\mathcal{U}_\infty$ be the collection of all sequences

$$(x_F, F \in \mathcal{I}) \in [0,1]^{\mathcal{I}} \tag{28}$$

of the form

$$x_F = \lim_n t_{\mathrm{hom}}(F, U_n)$$

for some sequence $\{U_n\}_n$ of unlabeled graphs, with $U_n \in \mathcal{U}_n$ for all n. The set $\mathcal{U}_\infty$ consists of all possible limits of sequences of unlabeled graphs, according to the definition of graph limit of [21] and [6]. Intuitively, one can think of each $U \in \mathcal{U}_\infty$ as an "infinite unlabeled graph". Indeed, notice the similarity between (28) and (26). In order to emphasize the role of density homomorphisms in this definition we will write $t_{\mathrm{hom}}(F, U)$ for the element of the sequence $U \in \mathcal{U}_\infty$ corresponding to $F \in \mathcal{I}$. Notice that if F and H are isomorphic then $t_{\mathrm{hom}}(F, U) = t_{\mathrm{hom}}(H, U)$, for all $U \in \mathcal{U}_\infty$. The set $\mathcal{U}_\infty$ is a compact subset of the compact metric space $[0,1]^{\mathcal{I}}$ when endowed with the metric

$$d(x, y) = \sum_{i=1}^{\infty} \frac{1}{2^i} |x_{F_i} - y_{F_i}|,$$

where $F_1, F_2, \dots$ is an enumeration of all the graphs in $\mathcal{I}$.

The next result provides a representation of the Möbius parameters of the probability distributions in $\mathcal{E}_\infty$ as expected density homomorphism of all the graphs in $\mathcal{I}$. In addition, the Möbius parameters of the extremal distributions in $\mathcal{E}_\infty$ satisfy a defining set by polynomial equations given below in (30).

Theorem 4.4 (de Finetti's theorem for exchangeable random networks). *The Möbius parameters corresponding to the probability distribution $p_\infty \in \mathcal{E}_\infty$ are given by*

$$z(F) = \lim_n \mathbb{E}\left[t_{\mathrm{hom}}(F, G[n])\right] = \mathbb{E}\left[t_{\mathrm{hom}}(F, U)\right], \quad \forall F \in \mathcal{I}. \tag{29}$$

where the expectation is with respect to the distribution of a random variable U taking values in $\mathcal{U}_\infty$. A distribution on $\mathcal{L}_\infty$ is extremal in $\mathcal{E}_\infty$ if and only if its Möbius parameters satisfy the conditions

$$z(F) = z(F_1)z(F_2) \tag{30}$$

for all $F \in \mathcal{I}$ with $F = F_1 F_2$. Furthermore, there exists one deterministic graph limit $U \in \mathcal{U}_\infty$ such that

$$z(F) = t_{\mathrm{hom}}(F, U) = \lim_n t_{\mathrm{hom}}(F, G[n]), \quad \forall F \in \mathcal{I}, \tag{31}$$

where the limit exists almost surely.

Proof. Let $G \in \mathcal{L}_\infty$ be an exchangeable infinite labeled graph. Then, for a fixed $m \geq 2$ and each $n \geq m$, Theorem 4.1 yields that

$$\left| \mathbb{P}\left(G[m] \supset F\right) - \mathbb{E}\left[t_{\text{hom}}(F, G[n])\right] \right| \leq 1 - \frac{(n)_m}{n^m}, \quad \forall F \in \mathcal{I}_m. \tag{32}$$

Let $\{U_n\}_n \subset \mathcal{U}$ be a sequence of random unlabeled graphs such that, for each n, U_n represents the isomorphism class of $G[n]$. Then, for each n, $t_{\text{hom}}(F, G[n]) \overset{d}{=} t_{\text{hom}}(F, U_n)$, where we recall that "$\overset{d}{=}$" denotes identity in distribution, and, as result $\mathbb{E}\left[t_{\text{hom}}(F, G[n])\right] = \mathbb{E}\left[t_{\text{hom}}(F, U_n)\right]$. Taking the limit in n, (32) implies that

$$\mathbb{P}\left(G[m] \supset F\right) = \lim_n \mathbb{E}\left[t_{\text{hom}}(F, U_n)\right],$$

for all $F \in \mathcal{I}_m$ and all $m \in \mathbb{N}$. Thus, by Theorem 3.1 in [10], there exists a random element $U \in \mathcal{U}_\infty$ such that

$$\mathbb{P}\left(G \supset F\right) = \mathbb{E}\left[t_{\text{hom}}(F, U)\right], \quad \forall F \in \mathcal{I},$$

where $\mathbb{E}\left[t_{\text{hom}}(F, U)\right] = \lim_n \mathbb{E}\left[t_{\text{hom}}(F, U_n)\right] = \lim_n \mathbb{E}\left[t_{\text{hom}}(F, G[n])\right]$. Thus, Equation (29) is proved.

Remark 4.5. Few comments are in order.

1. In the notation of [10], U_n converges in distribution to U, both viewed as elements of the space $\overline{\mathcal{U}}$, and, by Theorem 5.3 therein, such a random U is unique.

2. Furthermore, invoking again Theorem 3.1 in [10], we can conclude that $t_{\text{hom}}(F, G[n])$ converges in distribution for each $F \in \mathcal{I}$.

3. Alternatively, we may prove the claim using standard arguments from the theory of weak convergence of probability measures; see, e.g., the proof of Theorem 4 in [8]. Indeed, for each n we let μ_n be the probability distribution of U_n defined over the compact metric space $\overline{\mathcal{U}}$. Then, there exists a subsequence $\{\mu_{n_i}\}_i$ that converges weakly to a probability measure μ over the same space, which we may define to be the distribution of U. Since, for each fixed $F \in \mathcal{I}$, $t_{\text{hom}}(F, \cdot)$ is a bounded and continuous function over $\overline{\mathcal{U}}$ [see, e.g., 20] the result follows.

To show (30), we will rely on the following result of [10].

Lemma 4.6 ([10], Corollary 5.4)**.** *There is a one-to-one correspondence between the extreme points of the set $\mathcal{E}_\infty$ and the set $\mathcal{U}_\infty$, given by*

$$t_{\text{hom}}(F, U) = z(F), \quad \forall F \in \mathcal{I}, \tag{33}$$

where $U \in \mathcal{U}_\infty$ and $z(F)$ is the value of Möbius parameter at F for the corresponding p_∞ (see 27).

We can now prove (30). Assume that p_∞ is extremal in $\mathcal{E}_\infty$ with Möbius parameters $\{z(F), F \in \mathcal{I}\}$. Let $U \in \mathcal{U}_\infty$ its corresponding sequence. Then there exists a sequence $\{U_n\}_n \subset \mathcal{U}$ of unlabeled graphs with $U_n \in \mathcal{U}_n$ for all n such that $t(F, U) = \lim_n t_{\mathrm{hom}}(F, U_n)$, for any $F \in \mathcal{I}$. Consider any pair of node disjoint graphs F_1 and F_2 in $\mathcal{I}$. Without loss of generality, we may take the nodesets of F_1 and F_2 to be $[m_1]$ and $\{m_1 + 1, \ldots, m_1 + m_2\}$, respectively. Using theorem 4.6,

$$z(F_1 F_2) = t(F_1, F_2, U) = \lim_n t_{\mathrm{hom}}(F_1 F_2, U_n) = \lim_n t_{\mathrm{hom}}(F_1, U_n) t_{\mathrm{hom}}(F_2, U_n) = z(F_1) z(F_2),$$

where the third identity follows from the multiplicative property of density homomorphisms, which holds for all $n \geq m_1 + m_2$; see (11). The same argument applies to any pair of node-disjoint graphs F_1 and F_2 in $\mathcal{I}$, and (30) follows. Now suppose that (30) holds. Using Equation (29), for any pair of node-disjoint isomorphic graphs F_1 and F_2 in $\mathcal{I}$,

$$z(F_1 F_2) = \mathbb{E}\left[t_{\mathrm{hom}}(F_1 F_2, U)\right] = \mathbb{E}\left[t_{\mathrm{hom}}^2(F_1, U)\right],$$

where U is the random element in $\mathcal{U}_\infty$ corresponding to the distribution p_∞. Using (30) and the fact that $z(F_1) = z(F_2)$, we have that

$$\mathbb{E}\left[t_{\mathrm{hom}}^2(F_1, U)\right] = z(F_1 F_2) = z(F_1) z(F_2) = z^2(F_1) = \left(\mathbb{E}\left[t_{\mathrm{hom}}(F_1, U)\right]\right)^2,$$

and, therefore, that $t(F_1, U)$ is almost surely constant. Since the choice of F_1 is arbitrary, we conclude that the random variable $t_{\mathrm{hom}}(F, U)$ is almost surely constant for each $F \in \mathcal{I}$ and therefore, by definition, that U is non random. It then follows from Theorem 4.6 that the distribution of G is extremal. Finally, since $t_{\mathrm{hom}}(F, G[n]) \overset{d}{=} t_{\mathrm{hom}}(F, U_n)$ for all n and $F \in \mathcal{I}$ and $\{U_n\}$ is a non-random sequence of graphs in $\mathcal{U}$ with graph limit U, $t_{\mathrm{hom}}(F, G[n]$ converges almost surely to $t_{\mathrm{hom}}(F, U)$.

Theorem 4.4 gives a reformulation of well known results about symmetric binary exchangeable arrays [see, e.g. 2, 3, 19, 28, 12, 15, 16] and can also be directly linked to the theory of graph limits, as shown in particular by [10] [see also 20, Chapter 11]. Our contribution is a relatively simple proof that combines the finite exchangeability bound from Theorem 4.1 with classic arguments from the theory of weak convergence of measure as detailed in [10].

The identity (29) signifies that the Möbius parameters of any $p_\infty \in \mathcal{E}_\infty$ can be expressed as an average of density homomorphisms over graph limits, while Equation (31) expresses the result that there is a one-to-one correspondence between graph limits and extremal distributions in $\mathcal{E}_\infty$ (this formally stated in Theorem 4.6 above).

The proof of the Theorem also reveals that if G is a random graph in $\mathcal{L}_\infty$ with an exchangeable distribution, then, as $n \to \infty$ and for each $F \in \mathcal{I}$, the sequence $t_{\mathrm{hom}}(F, G[n])$ converges in distribution; it also converges almost surely if and only if the distribution of G is extremal in $\mathcal{E}_\infty$. Furthermore, we see that

$$z(F) = \lim_n \mathbb{E}[t_{\mathrm{hom}}(F, G[n])], \quad F \in \mathcal{I}.$$

Of course, a priory, for any infinite (random or deterministic) graph in $\mathcal{L}_\infty$, the limit $\lim_n t_{\mathrm{hom}}(F, G[n])$ needs not exist.

An equivalent version of Theorem 4.4 can also be given for probability parameters as opposed to Möbius parameters. However, the parametrization of extremal distributions in $\mathcal{E}_\infty$ by the induced probabilities of finite graphs does not seem to satisfy any factorization properties, such as the ones expressed in (30) for the Möbius parameters. For this reason, we find the Möbius parametrization more convenient. We refrain from providing the details.

One of the main implications of Theorem 4.4 is that, for any integer $n \geq 2$, if $p_n \in \mathcal{E}_n$ is the marginal of an extremal exchangeable distribution on $\mathcal{L}_\infty$, then, by Equation (30), its Möbius parameters (marginal probabilities) satisfy the identities

$$z_n(F_1 F_2) = z_n(F_1) z_n(F_2), \quad \forall F_1, F_2 \in \mathcal{I}_n, \tag{34}$$

i.e. the approximation Equation (22) holds exactly. This property holding for all n is equivalent to a well-known measure-theoretical property of exchangeable distributions over binary arrays, known as *dissociatedness*; see, e.g., [28]. In fact, dissociatedness is a necessary and sufficient condition for an exchangeable distribution over arrays to be extremal [3]. In the graph limit literature [see, e.g., 20, Chapter 11] an equivalent formulation of Equation (34) for all n is referred to as the *local* property of the associated sequence of distributions.

We will refer to distributions in $\mathcal{E}_n$ satisfying Equation (34) as **dissociated.** Notice that, according to our definition, a dissociated distribution in $\mathcal{E}_n$ needs not be the marginal of any dissociated or even finitely exchangeable distributions over larger graphs. This is in contrast with the classic notion of dissociatedness used in the probabilistic literature, which requires Equation (34) to hold for all n and therefore applies to all the marginals of an exchangeable distribution. Indeed, as we will show below, there exist dissociated distributions in $\mathcal{E}_n$, for all $n \geq 4$, that cannot be extended to to any distributions on larger graphs.

Since dissociated distributions in $\mathcal{E}_n$ contain the Möbius parameters of the marginals of all extremal distributions in $\mathcal{E}_\infty$, in order to understand the subset of $\mathcal{E}_n$ corresponding to image under the marginal mapping of all exchangeable distributions it is crucial to study dissociated distributions, which we do next in the next Section.

4.1.1. Connection with harmonic analysis

There is an interesting connection between Theorem 4.4 and harmonic analysis on semigroups [5, 25]. More precisely, if we consider the semigroup $(\mathcal{J}, +)$ of unlabeled graphs without isolated nodes, where $+$ denotes node disjoint union, the Möbius parameters clearly satisfy

$$z(F) = \phi([F]),$$

for some function $\phi : \mathcal{J} \to \mathbb{R}_+$. A is shown in the lemma below, the function ϕ is **positive definite** on $(\mathcal{J}, +)$, meaning that any matrix of the form

$$m_{ij} = \phi([F_i] + [F_j]), \quad i, j = 1, \ldots, n$$

is positive semidefinite.

Lemma 4.7. *Let G be a random exchangeable graph with Möbius parameters z given as above. Then the function ϕ is bounded and positive definite on $(\mathcal{J}, +)$.*

Proof. Clearly $\phi(\emptyset) = 1$ and ϕ is bounded. Introduce the binary random variables X_{ij} for $i \neq j \in \mathbb{N}$ where $X_{ij} = 1$ if $i \sim j$ in G and $X_{ij} = 0$ otherwise; X is the (random) adjacency matrix of G. Then, clearly

$$z(F) = \mathbb{E}\left(\prod_{ij:i\sim j \in F} X_{ij}\right).$$

So elementary calculations will verify that

$$
\sum_{u,v=1}^{n} c_u c_v \phi([F_u] + [F_v]) = \sum_{u,v=1}^{n} c_u c_v \mathbb{E}\left(\prod_{i\sim j \in F_u} X_{ij} \prod_{i\sim j \in F_v^*} X_{ij}\right)
$$
$$
= \mathbb{E}\left\{\sum_u c_u \prod_{i\sim j \in F_u} X_{ij}\right\}^2 \geq 0
$$

where F_v^* is a copy of F_v which is node-disjoint from F_u. This completes the proof.

We note that the property in Lemma 4.7 is referred to as **reflection positivity** in [21].

Now [4] show that the set of bounded positive definite functions on an Abelian semigroup is a Bauer simplex with the set of **characters** (multiplicative functions) as extreme points; this is essentially equivalent to the statement in Theorem 4.4.

4.1.2. Connection with graphons

The conclusions of Theorem 4.1 can be equivalently expressed using graphons. Indeed, paraphrasing a deep result about exchangeable arrays established by [2] and [15] [see also 16], the Möbius sequence $(z(F), F \in \mathcal{I})$ corresponding to an extremal exchangeable distribution admits the representation

$$z(F) = \int_{[0,1]} \int_{[0,1]^n} \prod_{(j,j) \in E(F)} W(x_i, x_j) dx_1 \ldots dx_n,$$

for all $F \in \mathcal{I}_n$ and all $n \geq 2$, and some symmetric measurable function $W \colon [0,1]^2 \to [0,1]$ (which is not uniquely defined). The same result was also established in the context of graph limits by [6] and [21], who termed the function W a graphon. Furthermore, equation (29) takes the form

$$z(F) = \int_{[0,1]} \left\{\int_{[0,1]^n} \prod_{(j,j) \in E(F)} \phi(\alpha, x_i, x_j) dx_1 \ldots dx_n\right\} d\alpha,$$

for all $F \in \mathcal{I}_n$ and all $n \geq 2$, for some measurable function $\phi \colon [0,1]^3 \to [0,1]$ (not necessarily uniquely defined), symmetric in its last two arguments. See, e.g., Chapter 14 in [2].

5. The manifold of dissociated exchangeable distributions

Let $\mathcal{D}_n \subset \mathcal{E}_n^M$ be the set of Möbius parameters of finitely exchangeable dissociated distributions on $\mathcal{L}_n$. By definition, $\mathcal{D}_n$ is comprised of all the points in $\mathcal{E}_n^M$ that satisfy the system of polynomial equations (34). Therefore, $\mathcal{D}_n$ is the intersection of $\mathcal{E}_n^M$ with a smooth manifold, in fact an affine variety in $(z(F), F \in \mathcal{I}_n)$. For this reason, we will refer to $\mathcal{D}_n$ as the **dissociated manifold**.

Clearly, the image of $\mathcal{D}_n$ under the inverse Möbius transform is a subset of $\mathcal{E}_n$ that can be also defined by a system of polynomial equations in the probability parameters, though these relations are not as simple as the ones in (34).

The next result describes some of the properties of the set $\mathcal{D}_n$. In particular, it shows that if p_n is the marginal of an exchangeable, non-extremal distribution on $\mathcal{L}_\infty$, then its Möbius parameters (marginal probabilities) are mixtures of the Möbius parameters of dissociated distributions in $\mathcal{E}_n$. From this, we obtain a partial geometric characterization of the set $\Pi_\infty^n(\mathcal{E}_\infty)$. It should be apparent now why Möbius parameters (marginal probabilities) are better suited to describe exchangeability in our context. Recall that M_n denotes the M'o bius map defined in (5).

Lemma 5.1. *The dimension of $\mathcal{D}_n$ is the number of unlabeled connected graphs with at most n nodes. If p_∞ is a distribution in $\mathcal{E}_\infty$ and $z = z(p_\infty)$ is as in Equation (27), then*

$$z_n \in \begin{cases} \mathcal{D}_n & \text{if } p_\infty \text{ is extremal} \\ \text{convhull}(\mathcal{D}_n) & \text{otherwise,} \end{cases}$$

where $z_n = (z(F), F \in \mathcal{I}_n)$. As a result, for each $n \geq 2$,

$$P_\infty^n(\mathcal{E}_\infty) \subset \left\{ M_n^{-1} z_n \colon z_n \in \text{convhull}(\mathcal{D}_n) \right\}.$$

Proof. That claim about the dimension of $\mathcal{D}_n$ follows from counting the number of polynomial equations in (34) (see, e.g., [11] and references therein for a similar calculation) and taking into account the fact that $z_n(F) = z_n(F')$ for all $F \sim F'$. To show the second statement, let $p_\infty \in \mathcal{E}_\infty$ and $z = z(p_\infty)$ be as in Equation (27). Then, by (29),

$$z_n(F) = z(F) = \mathbb{E}\left[t_{\text{hom}}(F, U)\right], \quad \forall F \in \mathcal{I}_n,$$

where in the above display U is a random variable taking values in $\mathcal{U}_\infty$. The claim is then established after noting that for any deterministic $U \in \mathcal{U}_\infty$, it holds that $t_{\text{hom}}(F, U) \in \mathcal{D}_n$ for all $F \in \mathcal{I}_n$, by the second part of Theorem 4.4.

Theorem 5.1 should be compared with the results in [13], where it is shown, with a different language, that the set of Möbius parameters in $\mathcal{E}_n^M$ arising from extremal distributions in $\mathcal{E}_\infty$ belongs to $\mathcal{D}_n$ and has a non-empty interior, of dimension equal to the number of unlabeled connected graphs with at most n nodes. The authors further remark that not much else is known about this set for any $n \geq 3$, including its topological properties (though they do show that it is path-connected).

In light of this, one may be led to conjecture that each point in the dissociated manifold $\mathcal{D}_n$ arises as the Möbius parameter of an extremal exchangeable distribution on $\mathcal{E}_\infty$. However, quite surprisingly, this is not the case. The following example provides a family of strictly positive probability distributions in $\mathcal{D}_4$ that are not extendable. Geometrically, this set is a line segment in $\mathcal{D}_4$. Other examples of dissociated distributions with zero entries in $\mathcal{D}_4$ that are not extendable are given in tables 2 and 3, which we discuss in the next section.

Example 5.2. Take $n = 4$. Consider the (strictly positive) probability distribution $p^* = \alpha p_1 + (1 - \alpha)p_2$ on $\mathcal{L}_4$, where $\alpha \in (0, 1)$, p_1 is the Erdös-Renyi distribution on $\mathcal{L}_4$ with $p = 1/2$, and p_2 is the distribution on $\mathcal{L}_4$ corresponding to the mixture of the point mass at the empty graph, the uniform distribution over the 4 graphs isomorphic to the union of a triangle and an isolated node, and the uniform distribution over the 3 graphs isomorphic to the 4-cycle, with weights $1/8$, $1/2$ and $3/8$ respectively. Both p_1 and p_2 are finitely exchangeable, and, therefore, so is p^*. Furthermore, since under both p_1 and p_2 the Möbius parameters corresponding to graphs isomorphic to an edge and to the disjoint union of two edges are $1/2$ and $1/4$ respectively, $p^* \in \mathcal{D}_n$ for each $\alpha \in (0, 1)$. Yet, as p_2 is not contained in the image of the marginal mapping Π_5^4 over $\mathcal{L}_5$, there is no distribution on $\mathcal{L}_n$, $n \geq 5$, whose marginal in $\mathcal{L}_4$ is p^*. Therefore, $z(p^*) \notin \Pi_\infty^5(\mathcal{E}_\infty)$.

It remains unknown whether there exist any simple criteria to determine whether the inverse Möbius transform of any point in $\mathcal{D}_n$ is also in $\Pi_\infty^m(\mathcal{E}_\infty)$.

5.1. Connections with [18].

In [18], we have also investigated exchangeable network models as graphical models on binary data with symmetric restrictions. There we have shown that distributions in $\mathcal{E}_n$ can only be compatible with few Markov properties, and we have identified all the possible conditional independence structures that such distributions may exhibit. Furthermore, we have proved that the only non-trivial conditional independence structure that yields a consistent sequence of finitely exchangeable probability distributions corresponds to a certain bi-directed graphical model for marginal independence. Such a model, which can be thought of as a canonical parametric model encompassing all finitely exchangeable networks of any given size, belongs to the class of marginal models for binary data studied by [11] [see also 26]. In particular, it is obtained from enforcing the dissociatedness constraints (34) in addition to exchangeability. One of the implications of these results is that the image under Π_∞^m of all extremal families in $\mathcal{E}_\infty$ is a strict submodel of a graphical model for marginal independence. Finally, the model can be parameterized as a curved exponential family on $\mathcal{L}_n$ with natural sufficient statistics given by the injective density homomorphisms and dimension equal to the number of connected unlabeled graphs on n nodes.

5.2. Maximum likelihood estimation

In this section we further investigate some of the statistical properties of the models specified by the manifold $\mathcal{D}_n$ of exchangeable and dissociated distributions. We will focus on the basic problem of estimating the Möbius parameters by maximizing the likelihood based on a sample of size one.

If $G \in \mathcal{L}_n$ is the observed network, a maximum likelihood estimator (MLE) of the Möbius parameters under the dissociated model is a point in the set

$$\mathrm{argmax}_{z \in \mathcal{D}_n} \ell(G, z),$$

where $\ell(\cdot, G)$ is the likelihood function, given by

$$z \in \mathcal{D}_n \mapsto \ell(z, G) = \sum_{\{F \in \mathcal{I}_n, :\ G \subseteq F\}} (-1)^{|F \setminus G|} z(F).$$

Using exchangeability, we can rewrite the likelihood function as

$$\ell(z, G) = \sum_{U \in \mathcal{U}_n} (-1)^{E(U) - E(G)} r_U(G) z(U),$$

where for a (labeled or unlabeled) graph G, $E(G)$ is the number of its edges, $r_U(G)$ is the number of graphs in $\mathcal{L}_n$ containing G as a subgraph and belonging to the isomorphism class represented by U, and $z(U)$ is the common value of the coordinates of the Möbius parameters z corresponding to the graphs in the isomorphism class represented by U. See Examples 1 and 2 in [18].

As remarked in the previous Section, points in $\mathcal{D}_n$ correspond to the closure of the mean-value space of a curved exponential family of probability distributions on $\mathcal{L}_n$. The MLE of the Möbius parameters may be on the boundary of $\mathcal{D}_n$ and may not be unique. Both cases are problematic from a statistical standpoint: the former case implies that the probability distribution corresponding to the MLE assigns zero probability to some graphs in $\mathcal{L}_n$ (when in fact all probabilities should be positive) and the latter case renders statistical inference based on such an estimator ill-posed.

In order to study both issues, we have obtained numerically all the possible maximum likelihood estimates under the constraints of exchangeability and dissociatedness for all the realizations of one network on four nodes. We have carried out the calculations in `Mathematica` using the built in optimization method. [11] propose a general algorithm for computing the MLE of the Möbius parameters of marginal models for binary data that could in principle be used for our problem. While such algorithm is more efficient and presumably faster than the brute force optimization, it requires strictly positive counts, a condition that is never satisfied when the data take the form of a single observed network.

When $n = 4$, there are 11 isomorphism classes, shown below in Figure 1 as unlabeled graphs, along with their respective sizes.

Table 2 and 3 show the maximum likelihood estimates of the Möbius parameter and of the actual probabilities, respectively. An empty entry in the table signifies a value of zero.
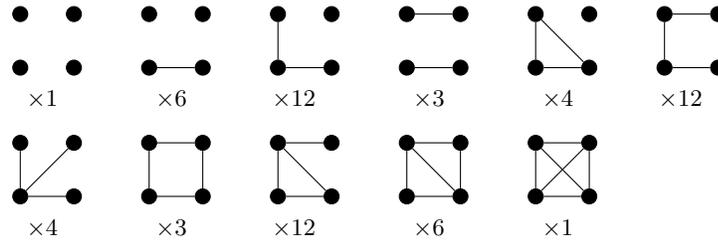
Figure 1: All non-isomorphic graphs $U \in \mathcal{U}_4$, along with the size m of the isomorphism class each represents, denoted by $\times m$.

It is apparent from Table 3 that all the estimates contain zero coordinates, a fact that implies that, with only one observed network, all the maximum likelihood estimates lie on the boundary of the parameter space. Furthermore, the MLE is not unique: it can be seen from Tables 2 and 3 that when the observed network consists of two parallel edges, or is a path or a cycle then the likelihood is maximized along line segments on the boundary of both the simplex and the Möbius simplex. (See also Example 7 in [18]). Finally, direct calculations reveal that, with the exception of the point masses at the empty and complete graphs, none of the maximum likelihood estimates of the probability distributions extend to exchangeable distributions over larger networks.

The fact that there are zeros in the MLEs of Table 3 means that it is relatively easy to check, for each case, that no exchangeable distribution on 5-node graphs can marginalize to that MLE. Consider for example the second row, corresponding to observing a 4-node graph with only one edge. The MLE is a mixture of a point mass at the complete graph and of the uniform distribution over graphs isomorphic to the observed one. In order for the MLE to be the marginal of some exchangeable distribution on 5-node graphs, that distribution must in turn be a mixture of uniform distributions over isomorphic 5-node graphs (and a point mass on the complete graph on 5 nodes) such that the removal of any one node will either be a 4-node graph with one edge or a complete graph. Such distribution does not exist (because there does not exist any 5-node unlabeled graph such that removing any one node will produce as an induced subgraph a 4-node graph with only one edge). Other cases can be checked by similar arguments.

6. Conclusions

It is worth commenting on the difference between Theorem 4.1 and 4.4 and analogous results for finitely exchangeable and exchangeable random binary sequences; see, e.g. [8, 9, 17]. First, the n-dimensional marginal of any exchangeable binary sequence can be described geometrically as a uniquely determined point in the convex hull of the intersection of the one-dimensional variety corresponding to the surface of independence inside the $(2^n - 1)$-dimensional simplex with the n dimensional affine subset of finitely exchangeable distributions. In our setting the manifold $\mathcal{D}_n$ of dissociated distributions (actually, as we saw, a certain non-trivial subset of it) plays an analogous role, though in the Möbius parametrization. However, it is clear that $\mathcal{D}_n$ is much more complex, and, unlike the surface of independence, which has fixed dimension 1, its dimension increases with n. Table 1

Table 1: Dimension of the sets $\mathcal{E}_n$ and $\mathcal{D}_n$ a function of the number of nodes n. The numbers are sourced from OEIS Foundation Inc. (2011), The On-Line Encyclopedia of Integer Sequences, see `http://oeis.org/A000088` and `http://oeis.org/A001349`.

						n			
	3	4	5	6	7	8	9	10	11
$\dim(\mathcal{E}_n)$	3	10	33	155	1,043	12,345	274,667	12,005,167	1,018,997,863
$\dim(\mathcal{D}_n)$	3	9	30	142	995	12,112	273,192	11,989,763	1,018,690,328

provides the dimension of $\mathcal{E}_n$ and of $\mathcal{D}_n$ (which are the number of unlabeled graphs minus 1 and the number of connected unlabeled graphs, respectively) for all nodes of size up to 11. As it can be seen and is also simple to show, the ratio of the dimension of $\mathcal{D}_n$ over that of $\mathcal{E}_n$ converges to 1 very rapidly as n grows. Another striking difference is the fact that not all points on the manifold of dissociated distributions correspond to extremal exchangeable distributions. This is in contrast with the sequence case, in which every point on the surface of independence corresponds to an extremal exchangeable distribution, for each n. Thus, exchangeability in graphs (a special case of exchangeability for binary 2-dimensional array) is considerably more subtle and complicated than the sequence case.

Acknowledgments

Table 2: Maximum likelihood estimates of the Möbius parameters for all possible samples of size one from $\mathcal{L}_4$. Along the rows of the table we report only the isomorphic classes in $\mathcal{G}_4$, each represented as an undirected graph in $\mathcal{G}_4$. Indeed, by exchangeability, isomorphic graphs in $\mathcal{G}_4$ yield the same maximum likelihood estimate. The columns of the table are indexed by all $U \in \mathcal{U}_4$ without isolated nodes. Thus the entry (U', U) represents the maximum likelihood estimate of $z(F)$, where $U = [F]$ if the observed network is a graph in the class represented by U'. An empty cell is equivalent to 0. In addition, $-(3\sqrt{2}/4)+1 \le b$, $d \le (3\sqrt{2}/4)-1$ and $-1/32 \le c \le 1/32$.

$x\backslash U$	U_1	U_2	U_3	U_4	U_5	U_6	U_7	U_8	U_9	U_{10}
x_1	$1/5$	$1/25$	$1/25$	$1/25$	$1/25$	$1/25$	$1/25$	$1/25$	$1/25$	$1/25$
x_2	$1/2$	$5/16$	$1/4$	$1/4$	$1/4$	$1/4$	$1/4$	$1/4$	$1/4$	$1/4$
x_3	$\sqrt{2}-1$	$(3\sqrt{2}/2)-2$	$3-2\sqrt{2}$	$(3\sqrt{2}/4)-1-b$	$1/8$	$(3\sqrt{2}/4)-1+b$	$1/8$			
x_4	$1/2$	$1/4$	$1/4$	$1/8$	$1/16$	$1/32+c$	$1/8$			
x_5	$1/2$	$3/16$	$1/4$	$1/32-c$	$1/8$	$1/8$	$3-2\sqrt{2}$	$1/16$		
x_6	$1/2$	$1/4$	$1/4$	$(3\sqrt{2}/4)-1-d$	$3-2\sqrt{2}$	$(3\sqrt{2}/4)-1+d$				
x_7	$2-\sqrt{2}$	$1-\sqrt{2}/2$	$6-4\sqrt{2}$	$3/16$	$1/8$	$3/16$				
x_8	$1/2$	$5/16$	$1/4$							
x_9	$4/5$	$16/25$	$16/25$	$12/25$	$12/25$	$12/25$	$8/25$	$8/25$	$4/25$	
x_{10}	1	1	1	1	1	1	1	1	1	1

Table 3: Maximum likelihood estimates of the probability parameters for all possible samples of size one from $\mathcal{L}_4$. Along the rows and columns of the Table we report only the isomorphic classes in $\mathcal{G}_4$, each represented as an undirected graph in $\mathcal{U}_4$. Thus, the entry corresponding to the pair (U', U) is the maximum likelihood estimate of the probability of observing a graph in the isomorphism class represented by U if the observed network is a graph in the class represented by U'. An empty cell is equivalent to 0. In addition, $-(3\sqrt{2}/4)+1 \le b, d \le (3\sqrt{2}/4)-1$ and $-1/32 \le c \le 1/32$.

$x\backslash U$	[∅]	[e]	[·L]	[:::]	[◺]	[⊔]	[⋖]	[□]	[△]	[N]	[⋈]
[∅]	1	4/25									
[·L]			1/16	$3-2\sqrt{2}$							
[:::]	1/8				$(3\sqrt{2}/4)-1-b$		$(3\sqrt{2}/1)-1+b$				
[◺]					1/8	1/16	1/32$+c$	1/8			
[⊔]	1/8				1/32$-c$		1/8	1/8			
[⋖]								$3-2\sqrt{2}$			
[□]	1/4				$(3\sqrt{2}/4)-1-d$		$(3\sqrt{2}/4)-1+d$		1/16		1/25
[△]	1/25									4/25	1/4
[⋈]											1

References

[1] Jin Akiyama, Geoffrey Exoo, and Frank Harary. The graphs with all induced subgraphs isomorphic. *The Bulletin of the Malaysian Mathematical Society Series 2*, 2:43–44, 1979.

[2] David Aldous. Representations for partially exchangeable random variables. *Journal of Multivariate Analysis*, 11:581–598, 1981.

[3] David Aldous. Exchangeability and related topics. In P. Hennequin, editor, *École d'Été de Probabilités de Saint–Flour XIII — 1983*, pages 1–198. Springer-Verlag, Heidelberg, 1985. Lecture Notes in Mathematics 1117.

[4] Christian Berg, Jens Peter Reus Christensen, and Paul Ressel. Positive definite functions on Abelian semigroups. *Mathematische Annalen*, 259:253–274, 1976.

[5] Christian Berg, Jens Peter Reus Christensen, and Paul Ressel. *Harmonic Analysis on Semigroups*. Springer-Verlag, New York, 1984.

[6] Christian Borgs, J.T. Chayes, László Lovász, V.T. Sós, and K. Vesztergombi. Convergent sequences of dense graphs I: Subgraph frequencies, metric properties and testing. *Advances in Mathematics*, 219:1801–1851, 2008.

[7] Harry Crane and Walter Dempsey. A framework for statistical network modeling. Available at https://arxiv.org/abs/1509.08185, 2015.

[8] Persi Diaconis. Finite forms of de Finetti's theorem on exchangeability. *Synthese*, 36(2):271–281, 1977.

[9] Persi Diaconis and David Freedman. On the statistics of vision: the Julesz conjecture. *Journal of Mathematical Psychology*, 24:112–138, 1981.

[10] Persi Diaconis and Svante Janson. Graph limits and exchangeable random graphs. *Rendiconti di Matematica, Serie VII*, 28:33–61, 2008.

[11] Mathias Drton and Thomas S. Richardson. Binary models for marginal independence. *Journal of the Royal Statistical Society Series B*, 70(2):287–309, 2008.

[12] Geoffrey K. Eagleson and Neville C. Weber. Limit theorems for weakly exchangeable arrays. *Mathematical Proceedings of the Cambridge Philosophical Society*, 84:123–130, 1978.

[13] Paul Erdös, László Lovász, and Joel Spencer. Strong independence of graphcopy functions. In J.A. Bondy and U.S.R. Murty, editors, *Graph Theory and Related Topics*, pages 165–172, New York, 1979. Academic Press.

[14] David Freedman. A remark on the difference between sampling with and without replacement. *Journal of the American Statistical Association*, 72:681–681, 1977.

[15] Douglas N. Hoover. Relations on probability spaces and arrays of random variables. Preprint, Institute of Advanced Study, Princeton., 1979.

[16] Olav Kallenberg. *Probabilistic Symmetries and Invariance Principles*. Springer–Verlag, New York, 2005.

[17] G. Jay Kerns and Gábor J. Székely. DeFinetti's theorem for abstract finite exchangeable sequences. *Journal of Theoretical Probability*, 19:589–608, 2006.

[18] Steffen Lauritzen, Alessandro Rinaldo, and Kayvan Sadeghi. Random networks, graphical models, and exchangeability. *Journal of the Royal Statistical Society, Series B*, 80:481–508, 2018.

[19] Steffen L. Lauritzen. Exchangeable Rasch matrices. *Rendiconti di Matematica, Serie VII*, 28:83–95, 2008.

[20] László Lovász. *Large Networks and Graph Limits*, volume 60 of *Colloquium Publications*. American Mathematical Society, 2012.

[21] László Lovász and Balász Szegedy. Limits of dense graph sequences. *Journal of Combinatorial Theory, Series B*, 96:933–957, 2006.

[22] František Matúš. Finite partially exchangeable arrays. Technical Report 1856, Institute of Information Theory and Automation, Academy of Sciences of the Czech Republic, Prague, 1995.

[23] Peter Orbantz and Daniel M. Roy. Bayesian models of graphs, arrays, and other exchangeable structures. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 37:437–461, 2015.

[24] Malempati M. Rao. Projective limits of probability spaces. *Journal of Multivariate Analysis*, 1:28–57, 1971.

[25] Paul Ressel. Exchangeability and semigroups. *Rendiconti di Matematica, Serie VII*, 28:63–81, 2008.

[26] Alberto Roverato, Monia Lupparelli, and Luca La Rocca. Log-mean linear models for binary data. *Biometrika*, 100:485–494, 2013.

[27] Cosma Shalizi and Alessandro Rinaldo. Consistency under sampling of exponential random graph models. *Annals of Statistics*, 41:508–535, 2013.

[28] Bernard W. Silverman. Limit theorems for dissociated random variables. *Advances in Applied Probability*, 8:806–819, 1976.

[29] Tom A. B. Snijders. Conditional marginalization for exponential random graph models. *The Journal of Mathematical Sociology*, 34:239–252, 2010.

[30] Richard P. Stanley. *Enumerative Combinatorics*. Cambridge University Press, Cambridge, UK, 2nd edition, 2011.

[31] Alexander Volfovsky and Edoardo M. Airoldi. Sharp total variation bounds for finitely exchangeable arrays. *Statistics and Probability Letters*, 114:54 – 59, 2016.