

# Exchangeable Rasch Matrices\*

Steffen L. Lauritzen<sup>†</sup>  
University of Oxford

September 17, 2007

## Abstract

This article is concerned with binary random matrices that are exchangeable with the probability of any finite submatrix only depending on its row- and column sums. We describe basic representations of such matrices both in the case of full row- and column exchangeability and the case of weak exchangeability. Finally the results are interpreted in terms of random graphs with exchangeable labels and with a view towards their potential application to social network analysis.

*Key words:* partial exchangeability; random graphs; social network analysis; sufficiency.

*AMS Subject classifications:* 62B99, 60C05.

## 1 Introduction

Let us initially consider the sequential case and recall that a sequence of random variables  $X_1, \dots, X_n, \dots$  is said to be *exchangeable* if for all  $n$  and  $\pi \in S(n)$  it holds that

$$X_1, \dots, X_n \stackrel{\mathcal{D}}{=} X_{\pi(1)}, \dots, X_{\pi(n)}$$

where  $S(n)$  is the group of permutations of  $\{1, \dots, n\}$  and  $\stackrel{\mathcal{D}}{=}$  denotes equality in distribution; i.e. the distribution of an exchangeable sequence of random variables is unchanged whenever the order of any finite number of them is rearranged.

---

\*This paper was presented at the Bruno de Finetti Centenary Conference in Rome, November 2006.

<sup>†</sup>Department of Statistics, University of Oxford, 1 South Parks Road, Oxford OX1 3TG, United Kingdom. email: steffen@stats.ox.ac.uk.

de Finetti's theorem for binary sequences [1] then says that a binary sequence  $X_1, \dots, X_n, \dots$  is exchangeable if and only if there exists a distribution function  $F$  on  $[0, 1]$  such that for all  $n$

$$p(x_1, \dots, x_n) = \int_0^1 \theta^{t_n} (1 - \theta)^{n - t_n} dF(\theta),$$

where  $t_n = \sum_{i=1}^n x_i$ . Further, the limiting frequency  $Y = \lim_{n \rightarrow \infty} \sum_{i=1}^n X_i/n$  exists and  $F$  is the distribution function of this limit. Conditionally on  $Y = \theta$ , the sequence  $X_1, \dots, X_n, \dots$  are independent and identically distributed with expectation  $\theta$ .

An alternative formulation of exchangeability focuses on its relationship to sufficiency [2, 3, 4]. A statistic  $t(x)$  is summarizing for  $p$  [5] if for some  $\phi$

$$p(x) = \phi(t(x)).$$

Note that if  $t$  is summarizing for all  $p$  in a family  $\mathcal{P}$  of distributions, it is sufficient for  $\mathcal{P}$ .

For binary variables,  $X_1, \dots, X_n, \dots$  is exchangeable if and only if for all  $n$

$$P(X_1 = x_1, \dots, X_n = x_n) = \phi_n(\sum_{i=1}^n x_i)$$

i.e., if and only if  $t_n = \sum_{i=1}^n x_i$  is summarizing for its distribution. This is due to the fact that  $t_n = \sum_{i=1}^n x_i$  is the maximal invariant for the action of the permutation group  $S(n)$  on the binary sequences of length  $n$  and  $S(n)$  acts transitively on the sequences with given value of  $t_n$ , so that any two such sequences are permutations of each other.

The present paper is investigating the interplay between these ideas in the case of random binary matrices where the situation is somewhat more complex.

The next section is initially giving an overview of results in [6] and the reader is referred to this paper for details not described here. Some of the considerations of [6] are further extended to the case of weakly exchangeable arrays. The last section touches upon the relation of these to random graphs and social network analysis.

## 2 Exchangeable binary matrices

### 2.1 Random Rasch matrices

The Rasch model [7] was originally developed to analyse data from intelligence tests where  $X_{ij}$  indicates a binary outcome when problem  $i$  was attempted by person  $j$ ,  $X_{ij} = 1$  denoting success and  $X_{ij} = 0$  denoting failure. The model is described by 'easinesses'  $\alpha = (\alpha_i)_{i=1, \dots}$  and 'abilities'

$\beta = (\beta_j)_{j=1,\dots}$  so that binary responses  $X_{ij}$  are conditionally independent given  $(\alpha, \beta)$  and

$$P(X_{ij} = 1 | \alpha, \beta) = 1 - P(X_{ij} = 0 | \alpha, \beta) = \frac{\alpha_i \beta_j}{1 + \alpha_i \beta_j}.$$

The model is a potential model for a large variety of phenomena such as, for example, a batter  $i$  getting a hit against a pitcher  $j$  in baseball matches [8] or the occurrence of species  $i$  on island  $j$  [9]; see for example [10] for a survey.

A *random Rasch matrix* has  $(\alpha_i)$  i.i.d. with distribution  $A$ ,  $(\beta_j)$  i.i.d. with distribution  $B$ , the entire sequences  $\alpha$  and  $\beta$  also being independent of each other. Such a model would be relevant if each of batters and pitchers were *a priori* exchangeable. An example of a random Rasch matrix is displayed in Figure 1.

## 2.2 Exchangeable and summarized matrices

Recall that a doubly infinite matrix  $X = \{X_{ij}\}_{1,1}^{\infty,\infty}$  is *row-column exchangeable* [11] (an RCE-matrix) if for all  $m, n, \pi \in S(m), \rho \in S(n)$

$$\{X_{ij}\}_{1,1}^{m,n} \stackrel{\mathcal{D}}{=} \{X_{\pi(i)\rho(j)}\}_{1,1}^{m,n},$$

i.e. if the distribution is unchanged when rows or columns are permuted. A random Rasch matrix is clearly RCE.

A doubly infinite (binary) matrix  $X = \{X_{ij}\}_{1,1}^{\infty,\infty}$  is said to be *row-column summarized* (RCS-matrix) if for all  $m, n$

$$p(\{x_{ij}\}_{1,1}^{m,n}) = \phi_{m,n}\{r_1, \dots, r_m; c_1, \dots, c_n\},$$

where  $r_i = \sum_j x_{ij}$  and  $c_j = \sum_i x_{ij}$  are the row- and column sums.

In contrast to the sequential case, there is no simple summarizing statistic for an RCE matrix. RCE-matrices are generally not RCS-matrices and *vice versa* because the group  $G_{RC}$  of row- and column permutations does not act transitively on matrices with fixed row- and column sums. To see the latter, consider

$$M_1 = \begin{Bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 1 \end{Bmatrix}, \quad M_2 = \begin{Bmatrix} 0 & 0 & 1 \\ 0 & 0 & 1 \\ 1 & 1 & 0 \end{Bmatrix}.$$

The matrices  $M_1$  and  $M_2$  have identical row- and column sums. However, their determinants are different:  $|\det M_1| = 1$  whereas  $|\det M_2| = 0$ . Since the absolute value of the determinant is invariant under row- and column permutations, it follows that no combination of such permutations will ever modify  $M_1$  to become  $M_2$ .



Figure 1: Two random matrices, each of dimension  $100 \times 100$ . The matrix to the left is a random Rasch matrix and thus RCES whereas the matrix to the right is RCE but not RCS. The two matrices have the same overall mean equal to 0.5.

If a matrix is both RCE and RCS, we say that it is an *RCES-matrix*. Random Rasch matrices are RCES matrices since, conditionally on  $(\alpha, \beta)$ , we have

$$p(\{x_{ij}\}_{1,1}^{m,n} | \alpha, \beta) = \prod_{i=1}^m \prod_{j=1}^n \frac{(\alpha_i \beta_j)^{x_{ij}}}{1 + \alpha_i \beta_j} = \frac{\prod_{i=1}^m \alpha_i^{r_i} \prod_{j=1}^n \beta_j^{c_j}}{\prod_{i=1}^m \prod_{j=1}^n (1 + \alpha_i \beta_j)}, \quad (1)$$

which only depends on row- and column sums, implying that this also holds after taking expectation w.r.t.  $(\alpha, \beta)$ .

The difference between an RCE and RCS matrix is not always immediately visible. Figure 1 displays a random Rasch matrix next to an RCE matrix which is not RCS, yet they are not easily distinguishable.

### 2.3 de Finetti's theorem for RCE matrices

The set of distributions  $\mathcal{P}_{RCE}$  of binary RCE matrices is a convex simplex. In particular, *every*  $P \in \mathcal{P}_{RCE}$  has a unique representation as a mixture of extreme points  $\mathcal{E}_{RCE}$  of  $\mathcal{P}_{RCE}$ , i.e.

$$P(A) = \int_{\mathcal{E}} Q(A) \mu_P(Q).$$

The same holds if RCE is replaced by RCS or RCES. In addition, it can be shown that

$$\mathcal{E}_{RCES} = \mathcal{E}_{RCE} \cap \mathcal{P}_{RCS}.$$

The extreme measures are particularly simple. Aldous [11] shows that *for any*  $P \in \mathcal{P}_{RCE}$  *the following are equivalent:*

- $P \in \mathcal{E}_{RCE}$

- The tail  $\sigma$ -field  $\mathcal{T}$  is trivial
- The corresponding RCE-matrix  $X$  is dissociated.

Here the tail  $\mathcal{T}$  is

$$\mathcal{T} = \bigcap_{n=1}^{\infty} \sigma\{X_{ij}, \min(i, j) \geq n\}$$

and a matrix is *dissociated* if for all  $A_1, A_2, B_1, B_2$  with  $A_1 \cap A_2 = B_1 \cap B_2 = \emptyset$

$$\{X_{ij}\}_{i \in A_1, j \in B_1} \perp\!\!\!\perp \{X_{ij}\}_{i \in A_2, j \in B_2}.$$

Following [12], a binary doubly infinite random matrix  $X$  is a  $\phi$ -matrix if  $X_{ij}$  are independent given  $U = (U_i)_{i=1, \dots}$  and  $V = (V_j)_{j=1, \dots}$  where  $U_i$  and  $V_j$  are independent and uniform on  $(0, 1)$  and

$$P(X_{ij} = 1 | U = u, V = v) = \phi(u_i, v_j),$$

In [11, 12, 13] it is shown that *distributions of  $\phi$ -matrices are the extreme points of  $\mathcal{P}_{RCE}$* , i.e. binary RCE matrices are mixtures of  $\phi$ -matrices. Different  $\phi$  may in general have identical distributions of their  $\phi$ -matrix. Clearly, if  $(g, h)$  is a pair of measure-preserving transformations of the unit interval into itself,  $\tilde{\phi}(u, v) = \phi(g(u), h(v))$  yields the same distribution of  $X$  as  $\phi$ . In fact,  $\phi$  is exactly determined up to such a pair of measure-preserving transformations [14].

## 2.4 Rasch type $\phi$ -matrices

As shown in [6], if a  $\phi$ -matrix is also *RCS*, then

$$P\left(\left\{\begin{array}{cc} 1 & 0 \\ 0 & 1 \end{array}\right\} | U = u, V = v\right) = P\left(\left\{\begin{array}{cc} 0 & 1 \\ 1 & 0 \end{array}\right\} | U = u, V = v\right)$$

which holds if and only if  $\phi$  is of *Rasch type*, i.e. if for all  $u, v, u^*, v^*$ :

$$\phi(u, v)\bar{\phi}(u, v^*)\bar{\phi}(u^*, v)\phi(u^*, v^*) = \bar{\phi}(u, v)\phi(u, v^*)\phi(u^*, v)\bar{\phi}(u^*, v^*), \quad (2)$$

where we have let  $\bar{\phi} = 1 - \phi$ . This is the *Rasch functional equation* [15].

Although RCE matrices have no simple summarizing statistics, RCES-matrices do: they are summarized by the empirical distributions of row- and column sums:

$$t_{mn} = \left(\sum_{i=1}^m \delta_{r_i}, \sum_{j=1}^n \delta_{c_j}\right),$$

where  $\delta_s$  is the measure with unit mass in  $s$ . This is a semigroup statistic, and RCES matrices can also be represented via mixtures of characters on the image semigroup (Ressel 2002, personal communication). General solutions of the Rasch functional equation thus represent characters of the image semigroup of the empirical row- and column sum measures.

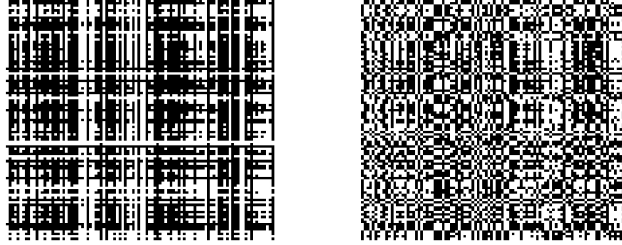


Figure 2: The left-hand matrix is a non-regular RCES  $\phi$ -matrix with columns  $\phi(u, v) = \chi_{\{u \leq v\}}$ . The matrix on the right-hand side is a  $\phi$ -matrix with  $\phi(u, v) = \chi_{\{|u-v| \leq 1-1/\sqrt{2}\}}$ . It is RCE but not RCS. The two matrices have the same overall mean equal to 0.5.

Lauritzen (2003) [6] shows that *any RCES matrix is a mixture of Rasch type  $\phi$ -matrices* and also that *any regular RCES matrix is a mixture of random Rasch matrices*. Here, a random binary matrix is said to be *regular* if

$$0 < P(X_{ij} = 1 | \mathcal{S}) < 1 \text{ for all } i, j,$$

where the *shell  $\sigma$ -algebra*  $\mathcal{S}$  is

$$\mathcal{S} = \bigcap_{n=1}^{\infty} \sigma\{X_{ij}, \max(i, j) \geq n\}.$$

Regular solutions ( $0 < \phi < 1$ ) to the Rasch functional equation are all of the form

$$\phi(u, v) = \frac{a(u)b(v)}{1 + a(u)b(v)},$$

where  $a$  and  $b$  are positive real-valued functions on the unit interval, leading to random Rasch models.

The matrix to the right in Figure 1 is a  $\phi$ -matrix with  $\phi(u, v) = (u+v)/2$ . Since this does not satisfy Rasch's functional equation it is not RCS. The matrix to the left is similarly a Rasch matrix with  $\phi = 6.49186uv/(1 + 6.49186uv)$ . The two matrices have the same overall mean equal to 0.5.

There are interesting non-regular solutions to the Rasch equation, for example

$$\phi(u, v) = \chi_{\{u \leq v\}} = \begin{cases} 1 & \text{if } u \leq v \\ 0 & \text{otherwise.} \end{cases}$$

A corresponding  $\phi$ -matrix is displayed in Figure 2. But there are also solu-

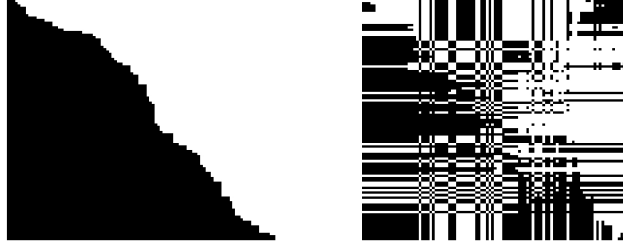


Figure 3: The matrices are obtained from those displayed in Figure 2 by sorting rows and columns according to their sum. The left-hand matrix was a non-regular RCES matrix. The right-hand matrix was RCE but not RCS.

tions such as, for example,

$$\phi(u, v) = \begin{cases} \frac{a(u)b(v)}{1 + a(u)b(v)} & \text{if } 1/3 < u, v < 2/3 \\ \chi_{\{u \leq v\}} & \text{otherwise.} \end{cases}$$

Both of these non-regular solutions imply the existence of incomparable groups, so that some questions are always answered correctly for a subgroup of the persons and some questions never answered by some. More complex variants of the latter example lead to Cantor–Rasch matrices, see [6] for further details.

The difference between RCE and RCES can be striking if the corresponding matrix is manipulated by sorting the rows and columns by their row- and column sums, as shown in Figure 3, where both diagrams have been obtained from Figure 2 in this way.

## 2.5 WE matrices

A random matrix is said to be *weakly exchangeable* [16, 17] (a WE-matrix) if for all  $n$  and  $\pi \in S(n)$

$$\{X_{ij}\}_{1,1}^{n,n} \stackrel{D}{=} \{X_{\pi(i)\pi(j)}\}_{1,1}^{n,n},$$

i.e. if the distribution of  $X$  is unchanged when rows and columns are permuted *using the same* permutation for rows and columns. Similarly we say that a doubly infinite (binary) matrix  $X = \{X_{ij}\}_{1,1}^{\infty,\infty}$  is *weakly summarized* (WS-matrix) if for all  $n$

$$p(\{x_{ij}\}_{1,1}^{n,n}) = \phi_n\{r_1 + c_1, \dots, r_n + c_n\},$$

where  $r_i = \sum_j x_{ij}$  and  $c_j = \sum_i x_{ij}$  are the row- and column sums as before.

Again *WE-matrices are generally not WS-matrices and vice versa*. No joint permutation of rows and columns take  $M_3$  into  $M_4$ , where

$$M_3 = \begin{pmatrix} 0 & 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 & 0 \end{pmatrix}, \quad M_4 = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 0 & 1 & 0 \end{pmatrix}$$

have identical row- and column sums, since  $\det M_3 = 4$  whereas  $\det M_4 = -4$  and simultaneous permutation of rows and columns also preserves the sign of the determinant.

If a matrix is both WE and WS, it is a *WES-matrix*. If in addition,  $\{X_{ij} = X_{ji}\}$ , i.e. the matrix is *symmetric*, we may consider SWE, SWS, SWES matrices, etc. Note that one could also consider the weaker *distributional symmetry* by assuming  $X \stackrel{D}{=} X^\top$ , i.e. that transposition of  $X$  does not alter the distribution of  $X$ . In the following we shall mostly restrict attention to the fully symmetric case and write  $X_{\{ij\}} = X_{ij} = X_{ji}$ .

## 2.6 de Finetti's theorem for SWE matrices

A symmetric binary doubly infinite random matrix  $X$  is a  $\psi$ -matrix if  $X_{\{ij\}}$  are all independent given  $U = (U_i)_{i=1,\dots}$  where  $U_i$  are mutually independent and uniform on  $(0, 1)$  and

$$P(X_{\{ij\}} = 1 | U = u) = \psi(u_i, u_j).$$

Reformulating results in [11] yields that *binary SWE matrices are mixtures of  $\psi$ -matrices*. Exactly as in the case of RCES matrices, it is easy to show that

$$\mathcal{E}_{SWES} = \mathcal{E}_{SWE} \cap \mathcal{P}_{SWS}.$$

implying that SWES matrices are mixtures of  $\psi$ -matrices where  $\psi$  satisfies the Rasch functional equation. The latter follows from the fact that we then must have for all  $y, z \in \{0, 1\}$  that

$$P\left(\left\{\begin{pmatrix} 0 & y & 0 & 1 \\ y & 0 & 1 & 0 \\ 0 & 1 & 0 & z \\ 1 & 0 & z & 0 \end{pmatrix} \mid U = u\right\}\right) = P\left(\left\{\begin{pmatrix} 0 & y & 1 & 0 \\ y & 0 & 0 & 1 \\ 1 & 0 & 0 & z \\ 0 & 1 & z & 0 \end{pmatrix} \mid U = u\right\}\right).$$

Hence *regular SWES  $\psi$ -matrices have the form*

$$\psi(u, v) = \frac{a(u)a(v)}{1 + a(u)a(v)}.$$



There are also non-regular solutions of interest in the symmetric case. For example

$$\psi(u, v) = \begin{cases} 0 & \text{if } u < 1/3 \text{ or } v < 1/3 \\ \frac{a(u)a(v)}{1 + a(u)a(v)} & \text{if } 1/3 < u, v < 2/3 \\ 1 & \text{otherwise.} \end{cases} \quad (3)$$

It seems complex to give a complete description of all symmetric solutions to the Rasch functional equation.

### 3 Random graphs

#### 3.1 Exchangeable matrices as random graphs

The results described in the previous sections become particularly relevant when the binary matrix  $X$  is considered to represent a random graph. This representation can be made in a number of ways. If we consider the rows and columns as labels of two different sets of vertices, an undirected random bipartite graph can be defined from  $X$  by ignoring the diagonal and placing an edge between  $i$  and  $j$  if and only if  $X_{ij} = 1$ .

In this interpretation, an RCE-matrix corresponds to a random bipartite graph with exchangeable labels within each group of graph vertices. Similarly, an RCS-matrix is one where any two bipartite graphs with the same vertex degree for every vertex are equally likely. An RCES matrix represents one where the two distributions of vertex degrees determine the probability of the graph.

If we consider the row-and column numbers to label the *same vertex set*, the matrix  $X$  represents a random (directed) graph by placing a directed edge from  $i$  to  $j$  if and only if  $X_{ij} = 1$ . A WE-matrix then represents a random graph with exchangeable labels.

If the matrix  $X$  is symmetric it naturally represents a random undirected graph, again by placing an edge between  $i$  and  $j$  if and only if  $X_{ij} = 1$ . An SWE-matrix then represents an undirected random graph with exchangeable labels, an SWS-matrix represents a random graph with probability only depending on its vertex degrees, and an SWES matrix one with probability only depending on the distribution of vertex degrees.

Examples of WES and SWES graphs with non-regular  $\psi$ -matrices are displayed in Figure 4.

#### 3.2 Social network analysis

Random graphs with exchangeability properties form natural models for *social networks* [18]. Frank and Strauss [19] consider *Markov graphs* which

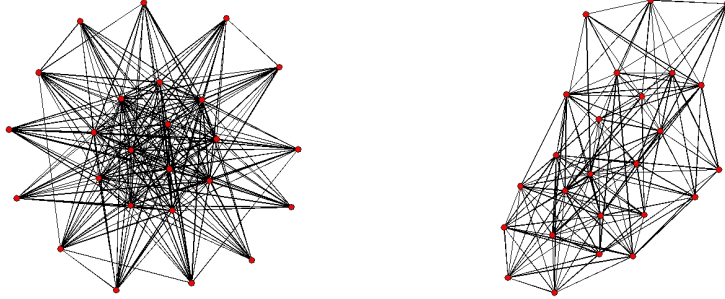


Figure 4: The graph on the left-hand side is a non-regular SWES graph with  $\psi$ -matrix given by (3). The graph on the right-hand side has a  $\psi$ -matrix with  $\phi(u, v) = \chi_{\{|u-v| \leq 1-1/\sqrt{2}\}}$ . It is SWE but not SWES. Both graphs have 25 vertices.

are random graphs with

$$X_{\{i,j\}} \perp\!\!\!\perp X_{\{k,l\}} \mid X_{E \setminus \{\{i,j\}, \{k,l\}\}} \quad (4)$$

whenever all indices  $i, j, k, l$  are different. Here  $E$  denotes the edges in the complete graph on  $\{1, \dots, n\}$ . They show that *weakly exchangeable Markov graphs on  $n$  vertices* all have the form

$$p(\{x_{ij}\}_{1,1}^{n,n}) \propto \exp\{\tau_n t(x) + \sum_{k=1}^{n-1} \eta_{nk} \nu_k(x)\}$$

where  $\tau_n$  and  $\eta_{nk}$  are arbitrary real constants,  $x = \{x_{ij}\}_{1,1}^{n,n}$ ,  $t(x)$  is the number of triangles in  $x$ , and  $\nu_k(x)$  is the number of vertices in  $x$  of degree  $k$ . Such Markov graphs are also SWE, but not extendable as such unless  $\tau_n = 0$  in which case they are also SWES. If  $\tau_n \neq 0$  and  $n > 5$ , they are not SWES.

Typically, exchangeable graphs generated by  $\psi$ -matrices differ from Markov graphs in that they are *dissociated*, hence marginally rather than conditionally independent:

$$X_{\{i,j\}} \perp\!\!\!\perp X_{\{k,l\}} \quad (5)$$

whenever all indices  $i, j, k, l$  are different. In fact *infinite weakly exchangeable Markov graphs are Bernoulli graphs* because the conjunction of (4) and (5) implies complete independence.

It seems unfortunate that random induced subgraph of Markov graphs are not Markov themselves and it could be of interest to develop alternative

models for social networks that preserve their structure when sampling subgraphs based on  $\psi$ -matrix models. This holds, for example, for the *latent space models* [20] and *latent position cluster models* [21], both of which are instances of  $\psi$ -matrix models.

For example, one could consider exchangeable random graphs which for every  $n$  also are summarized by the number of triangles and the empirical distribution of vertex degrees

$$p(\{x_{ij}\}_{1,1}^{n,n}) = f_n\{t(x), \sum_{k=1}^n \delta_{r_k(x)}\}, \quad (6)$$

or similar graphs with summarizing statistics being counts of specific types of subgraph.

Characterizing exchangeable solutions to (6) or functions involving similar statistics in general use in social network analysis, could establish an interesting class of alternatives to the generalizations of Markov graphs known as exponential random graph models [22, 23, 24].

## References

- [1] Bruno de Finetti. Funzione caratteristica di un fenomeno aleatorio. *Atti della R. Accademia Nazionale dei Lincei, Serie 6. Memorie, Classe di Scienze Fisiche, Matematiche e Naturale*, 4:251–299, 1931.
- [2] Steffen L. Lauritzen. Extreme point models in statistics (with discussion). *Scandinavian Journal of Statistics*, 11:65–91, 1984.
- [3] Persi Diaconis and David Freedman. Partial exchangeability and sufficiency. In J. K. Ghosh and J. Roy, editors, *Statistics, Applications and New Directions; Proceedings of the Indian Statistical Institute Golden Jubilee International Conference*, Sankhya A, pages 205–236. Indian Statistical Institute, 1984.
- [4] Steffen L. Lauritzen. *Extremal Families and Systems of Sufficient Statistics*. Springer-Verlag, Heidelberg, 1988. Lecture Notes in Statistics 49.
- [5] David Freedman. Invariants under mixing which generalize de Finetti’s theorem. *Annals of Mathematical Statistics*, 33:916–923, 1962.
- [6] Steffen L. Lauritzen. Rasch models with exchangeable rows and columns. In J. M. Bernardo, M. J. Bayarri, J. O. Berger, A. P. Dawid, D. Heckerman, A. F. M. Smith, and M. West, editors, *Bayesian Statistics 7*, pages 215–232, Oxford, 2003. Clarendon Press.

- [7] Georg Rasch. *Probabilistic Models for Some Intelligence and Attainment Tests*, volume 1 of *Studies in Mathematical Psychology*. Danmarks Pædagogiske Institut, Copenhagen, 1960.
- [8] Sam Gutmann, J. H. B. Kemperman, J. A. Reeds, and L. A. Shepp. Existence of probability measures with given marginals. *Annals of Probability*, 19:1781–1797, 1991.
- [9] J. B. Wilson. Methods for detecting non-randomness in species co-occurrences: A contribution. *Oecologica*, 73:579–582, 1987.
- [10] Gerhard H. Fischer and Iwo W. Molenaar, editors. *Rasch Models. Foundations, Recent Developments, and Applications*. Springer-Verlag, New York, 1995.
- [11] David Aldous. Representations for partially exchangeable random variables. *Journal of Multivariate Analysis*, 11:581–598, 1981.
- [12] Persi Diaconis and David Freedman. On the statistics of vision: the Julesz conjecture. *Journal of Mathematical Psychology*, 24:112–138, 1981.
- [13] D. N. Hoover. Row–column exchangeability and a general model for exchangeability. In G. Koch and F. Spizzichino, editors, *Exchangeability in Probability and Statistics*, pages 281–291. North-Holland, Amsterdam, 1982.
- [14] Olav Kallenberg. On the representation theorem for exchangeable arrays. *Journal of Multivariate Analysis*, 30:137–154, 1989.
- [15] Georg Rasch. Proof that the necessary condition for the validity of the multiplicative dichotomic model is also sufficient. Unpublished manuscript, 1971.
- [16] Bernard Silverman. Limit theorems for dissociated random variables. *Advances of Applied Probability*, 8:806–819, 1976.
- [17] G. K. Eagleson and N. C. Weber. Limit theorems for weakly exchangeable random variables. *Proceedings of the Cambridge Philosophical Society*, 84:73–80, 1979.
- [18] Stanley Wasserman and Katherine Faust. *Social Network Analysis*. Cambridge University Press, Cambridge, UK, 1995.
- [19] Ove Frank and David Strauss. Markov graphs. *Journal of the American Statistical Association*, 81:832–842, 1986.

- [20] Peter D. Hoff, Adrian E. Raftery, and Mark S. Handcock. Latent space approaches to social network analysis. *Journal of the American Statistical Association*, 97:1090–1098, 2002.
- [21] Mark S. Handcock, Adrian E. Raftery, and Jeremy M. Tantrum. Model-based clustering for social networks (with discussion). *Journal of the Royal Statistical Society, Series A*, 170:301–354, 2007.
- [22] Ove Frank. Statistical analysis of change in networks. *Statistica Neerlandica*, 45:283–293, 1991.
- [23] Stanley Wasserman and Philippa E. Pattison. Logit models and logistic regression models for social networks I: An introduction to Markov graphs and  $p^*$ . *Psychometrika*, 61:401–425, 1996.
- [24] Tom A. B. Snijders, Philippa E. Pattison, Garry L. Robins, and Mark S. Handcock. New specifications for exponential random graph models. *Sociological Methodology*, 36:99–153, 2006.