

Idiographic Ising and Divide and Color Models: Encompassing Networks for Heterogeneous Binary Data

M. Marsman^a  and K. Huth^{a,b,c}

^aDepartment of Psychology, University of Amsterdam; ^bDepartment of Psychiatry, Amsterdam University Medical Center; ^cCentre for Urban Mental Health, University of Amsterdam

ABSTRACT

The Ising model is a graphical model that has played an essential role in network psychometrics. It has been used as a theoretical model to conceptualize psychological concepts and as a statistical model to analyze psychological data. Using graphical models such as the Ising model to analyze psychological data has been heavily critiqued since these data often come from cross-sectional applications. An often voiced concern is the inability of the Ising model to express heterogeneity in the population. The idiographic approach has been posed as an alternative and aims to infer individual network structures. While idiographic networks overcome population heterogeneity, it is unclear how they aggregate into established cross-sectional phenomena. This paper establishes a formal bridge between idiographic and cross-sectional network approaches of the Ising model. We ascertain unique topological structures that characterize individuals and aggregate into an Ising model cross-sectionally. This new formulation supports population heterogeneity while being consistent with cross-sectional phenomena. The proposed theory also establishes a new statistical framework for analyzing populations of idiographic networks for binary variables. The Ising model and the divide and color model are special cases of this new framework. We introduce a Gibbs sampling algorithm to estimate models from this new framework.

KEYWORDS

Ising model; random-cluster model; divide and color model; network psychometrics; heterogeneity; idiographic network

Networks have become a popular alternative to the classical latent variable approach of conceptualizing psychological constructs.¹ A prominent example is how we came to think about psychological disorders such as depression (Borsboom, 2008). In classical psychometric theory, symptoms of depression such as weight problems, sleep problems, and a depressed mood, are thought to be caused by an underlying latent variable, depression (Caspi et al., 2014). In contrast, network theory suggests that symptoms form networks of mutually reinforcing variables instead and that there exist causal pathways in which a symptom such as sleep problems can lead to a depressed mood, which in turn then leads to weight problems (e.g., Cramer et al., 2016). This way of conceptualizing psychometric constructs not only had a significant impact on psychopathology research (Borsboom & Cramer, 2013), but it

has also been influential in thinking about what constitutes constructs such as intelligence (van der Maas et al., 2006; van der Maas et al., 2017), personality (Costantini et al., 2019; Cramer et al., 2012), and attitudes (Dalege et al., 2016; Dalege et al., 2019).

In the conceptualization of psychological constructs as networks of mutually reinforcing variables, a graphical model known as the Ising model (Ising, 1925; Lenz, 1920), or Quadratic Exponential model (Cox, 1972), has played an important role. Formally, the Ising model describes the joint distribution of a set of random variables that can be in either one of two states. In psychology, these binary random variables could reflect the presence or absence of a depression-related symptom (Borsboom, 2008; Cramer et al., 2010), the positive or negative evaluation of the trustworthiness of a political candidate (Dalege et al., 2016, 2019), or the correct or

incorrect response to an item in an educational test (Marsman et al., 2015), for example. In the Ising model, the binary variables embody the nodes of a network, and the interactions or associations between nodes then reflect the network structure.

The use of graphical models to conceptualize psychological constructs has also inspired their use to analyze psychological data (Costantini et al., 2015; Dalege et al., 2017; Marsman et al., 2015; Marsman et al., 2019; van Borkulo et al., 2014). But the data that are analyzed often come from cross-sectional applications, and the practice of using graphical models to analyze cross-sectional data has been heavily critiqued (e.g., Bringmann & Eronen, 2018; Forbes et al., 2017, 2019, 2022; Fried & Cramer, 2017). In this paper, we wish to address a prominent critique that the analyzed population needs to be strictly homogeneous regarding the model's characteristics (e.g., E. Bos & Wanders, 2016; F. M. Bos et al., 2017; Brusco et al., 2022). Homogeneity here implies that there can be no individual differences in the associations between the network's variables. Thus the associations at the individual level must mirror associations at the group level. But associations at the group level may be radically different from associations at the individual level (Kievit et al., 2013), a phenomenon that is often referred to as Simpson's paradox (Simpson, 1951) or the ecological fallacy (Robinson, 1950). This objection is particularly problematic for proponents of the Ising model, as the model does not have a natural way to express individual variation.

Because existing graphical models, such as the Ising model, do not naturally accommodate heterogeneity in the population, some researchers have advocated the use of person-centered network strategies instead (e.g., F. M. Bos, et al., 2022; Fisher, 2015; Fisher et al., 2018). The idea behind this idiographic approach is to use repeated measures of an individual's symptoms, say, to infer a person's network structure (Bak et al., 2016; Fisher et al., 2017; de Vos et al., 2017). In this way, unique networks are used to characterize persons, and the idiographic approach can circumvent population heterogeneity. However, whereas cross-sectional applications of graphical models do not accommodate population heterogeneity, the idiographic approach does not generalize to established cross-sectional phenomena. A case in point is the positive manifold (Spearman, 1904), the consistent finding that scores on cognitive tests positively correlate. Where cross-sectional models in psychology are compatible with established group-level phenomena such as the positive manifold (see Savi et al. (2019), for a recent discussion), it is, in general, unclear how models of the individual fit in.

To overcome heterogeneity using an idiographic network approach, we need to find a bridge between networks on the individual and the group levels. This paper aims to introduce that bridge in the case of the Ising model. We formulate an idiographic interpretation of the Ising model that supports population heterogeneity and is consistent with cross-sectional phenomena. We model the idiographic network structures with random graph models and draw a formal connection between the links in these idiographic topologies and cross-sectional correlations. A theoretical framework of Fortuin and Kasteleyn (1972) is the basis of this formulation. Fortuin and Kasteleyn used their theory to unify disparate areas in statistical physics, and Savi et al. (2019) recently used it to formulate an idiographic theory of intelligence that explains how cross-sectional phenomena such as the positive manifold can emerge from models of the individual. We will first use the theory to show how network structures unique to the individual can generate an Ising model in the population and how to traverse the two levels of networks. We will then use Fortuin and Kasteleyn's theory to open up a new statistical framework to analyze a population of idiographic networks. We obtain the Ising model and the divide and color model—a stochastic model for generating networks with correlated variables—as special cases of this new modeling approach. A Gibbs sampling approach allows us to estimate the new models.

This article comprises four sections. In the first section, we briefly review two distinct approaches in network analysis: The graphical model and the random graph model. The two approaches are fundamental to the theory of Fortuin and Kasteleyn and our idiographic formulation of the Ising model. The second section establishes a new network theory that encompasses both random graph models and graphical models. The encompassing network theory builds on the work of Fortuin and Kasteleyn. We use the theory to show how we can obtain the Ising model in cases where the network's topology is a random effect that varies at the individual level and clarify the connection between associations at the individual- and group levels. In the third section, we turn to the statistical analysis of encompassing networks. Here, we will derive the posterior distribution of the idiographic topologies and establish a Gibbs sampling algorithm to estimate the new models. Sample R-code for this can be found at <https://osf.io/beqgr/>. We illustrate the developments using simulated and real-data examples in the fourth section. The paper ends with a discussion.

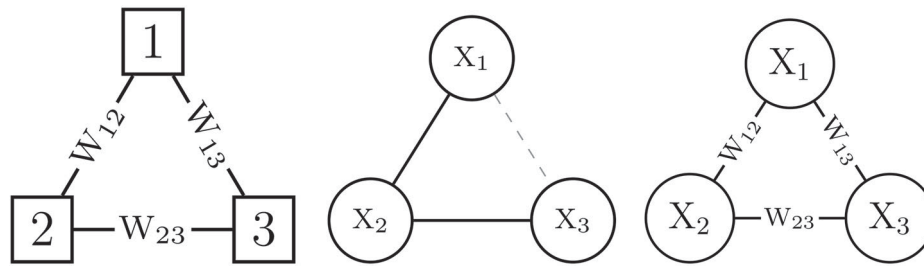


Figure 1. Three distinct approaches to a three variable network. The left figure illustrates a random graph model in which random variables are associated with the edges of the network. The middle figure illustrates a graphical model in which random variables are associated with the nodes of the network. These two approaches combine into the third approach, which is illustrated in the right panel where random variables are associated to both the nodes and the edges of the network.

The network models that underlie Fortuin and Kasteleyn's theory

It was almost fifty years ago that Fortuin and Kasteleyn observed several unexplained relations between the physical phenomena studied in two disparate areas of research. On the one hand, percolation research studies how particles trickle through a porous object, such as how water drips through coffee grounds. On the other hand, in research into magnetism, one studies the way that elementary particles such as electrons may interact to produce a magnet. Fortuin and Kasteleyn set out to discover if the unexplained relations observed between these two research fields were more than just coincidence (Grimmet, 2006) and worked out how the theories of magnetism and percolation relate (Fortuin, 1972a, 1972b; Fortuin & Kasteleyn, 1972). Fortuin and Kasteleyn's approach to consolidate the two scientific theories will be the basis of our idiographic network characterizations.

The theories of percolation and magnetism are formalized using two distinct types of network models. Blending the two scientific theories thus leads to a blending of the two network approaches. Percolation, for example, is conceptualized using a random graph model, and magnetism using the Ising model—a graphical model. Thus, Fortuin and Kasteleyn's theory informs us how the Ising model relates to random graph models, which is key to our idiographic formulation. This paper will use random graph models to characterize the idiographic network structures and use Fortuin and Kasteleyn's theory to relate idiographic networks to the Ising model, cross-sectionally. In this section, we will formally introduce the random graph model and the Ising model.

A random graph model for percolation and idiographic networks

Percolation has traditionally been analyzed using a random-graph model that is now widely known as the

Erdős and Rényi (1960) model (Broadbent & Hammersley, 1957). In general, the Erdős-Rényi model can be used to model the distribution of edges between pairs of nodes in a network. With W_{ij} denoting a binary random variable that, if it is equal to one, indicates the presence of a link between nodes i and j , and if it is equal to zero indicates the absence of that link, the Erdős-Rényi model for an n -variable network is characterized by the following probability distribution

$$p(\mathbf{W} = (w_{12}, \dots, w_{(n-1)n})^T) = \prod_{i=1}^{n-1} \prod_{j=i+1}^n \theta_{ij}^{w_{ij}} (1 - \theta_{ij})^{1-w_{ij}}, \quad (1)$$

where θ_{ij} denotes the probability that the random variable W_{ij} is equal to one. Thus, the Erdős-Rényi model comprises a set of independent Bernoulli variables, one for every pair of nodes in the network. Where the original Erdős-Rényi model has a constant edge inclusion probability θ , our formulation allows for distinct edge inclusion probabilities θ_{ij} for the separate edges w_{ij} .

The left panel in Figure 1 illustrates the Erdős-Rényi model for a network with three nodes. The network nodes are fixed and labeled one to three, and the edges between nodes constitute the random variables W_{ij} . For example, there is an edge between nodes one and two if the random variable W_{12} is equal to one, and there is no edge if W_{12} is equal to zero. The same holds for the other two edges and random variables, W_{13} and W_{23} . The Erdős-Rényi model is thus a model for the topology of the network, as it describes a probability distribution over different realizations of the network's structure. Figure 2 shows two possible realizations from the Erdős-Rényi model for the three-node network. Figure 2 also illustrates the influence of the three model parameters, θ_{12} , θ_{13} and θ_{23} : An edge between nodes two and three, for example, is present ($W_{23} = 1$) with probability θ_{23} . A solid black

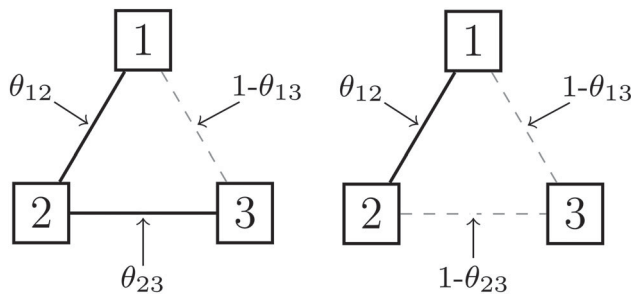


Figure 2. Two realizations of the Erdős-Rényi model for the three node network. Present edges are indicated with a black solid line and absent edges with a gray dashed line.

line indicates this edge in the left panel. An edge between nodes two and three is absent ($W_{23} = 0$) with probability $1 - \theta_{23}$, which is indicated with a dashed gray line in the right panel, for example.

A concept that will become important later on is the clustering of nodes in the network. If nodes are directly or indirectly related, they are said to be in the same cluster or connected component (Fortuin & Kasteleyn, 1972).² Thus, the network on the left of Figure 2 consists of a single cluster since all three of its nodes are either directly or indirectly connected. The network on the right of Figure 2, on the other hand, consists of two distinct clusters: Nodes one and two form one cluster, and Node three forms another.

In percolation research, the presence of an edge between two nodes of the network reflects an open connection between two caveats in a porous object. The absence of an edge then reflects a closed connection. Thus, particles can move between nodes one and three in the network on the left of Figure 2, but they cannot in the network on the right since it has a closed connection between nodes two and three. That is, particles cannot move between the network's clusters. In this paper, however, we use random-graph models to model the distribution of relations between pairs of observables in the idiographic networks. In this way, the presence of an edge in an idiographic network implies a direct association between two symptoms, evaluations, or responses. The absence of an edge then implies that the two observables are not directly related for this person. The left panel of Figure 2 thus corresponds to an idiographic network in which all three variables are related and form a single cluster. The right panel depicts an

idiographic network in which the first two variables are related but are independent of the third. That is, they form two distinct clusters.

A graphical model for magnetism and cross-sectional networks

The Ising model was proposed by Lenz (1920) as a possible model for magnetism, after which his student Ising (1925) wrote his doctoral dissertation about the model (Brush, 1967; Niss, 2005). The Ising model is a graphical model that can be used to model the states of the nodes of the network as a function of the network's structure. Formally, the Ising model is specified by the following probability distribution for n dichotomous random variables X_i that each take values in $\{-1, +1\}$,³

$$p(\mathbf{X} = (x_1, \dots, x_n)^T) = \frac{1}{Z_I} \exp \left(\sum_{i=1}^n \mu_i x_i + \sum_{i=1}^{n-1} \sum_{j=i+1}^n \sigma_{ij} x_i x_j \right), \quad (2)$$

where Z_I is the model's normalizing constant. The Ising model in Equation (2) consists of two sets of parameters: The main effects or thresholds of the network's variables μ_i and their pairwise interactions σ_{ij} . A variable in the network tends to have a positive value ($X_i = +1$) when its main effect is positive ($\mu_i > 0$), and tends to have a negative value ($X_i = -1$) when its main effect is negative ($\mu_i < 0$). In addition, the dyad X_i and X_j tends to align their values when their interaction effect is positive ($\sigma_{ij} > 0$), but tends to be in different states when their interaction effect is negative ($\sigma_{ij} < 0$).

The middle panel of Figure 1 illustrates the graphical model. As opposed to the random graph model that is illustrated in the left panel of Figure 1, the graphical model does not model the relations between nodes but models the states of the nodes instead, which here constitute the random variables X_i . The edges between nodes inform about which variables influence each other. Solid black lines indicate present edges, and dashed gray lines indicated absent edges. Nodes one and two are directly related in the middle panel network of Figure 1, whereas nodes one and three are not. Two possible realizations from the Ising model for the three-node network are shown in Figure 3, together with the influence of the model's parameters. As can be gleaned from Equation (2), the

²In this paper, we use the term "cluster" to refer to a set of nodes that are directly or indirectly connected, i.e., a "connected component" or "maximal clique". This term stems from the percolation literature that forms the basis of the theory in this paper and refers to having an open path between nodes; one can reach all nodes in a cluster by traversing the links between nodes in the cluster. In contemporary literature, clustering is often used to define a group of nodes that are densely connected.

³In practical applications, we often code the dichotomous variables as $\{0, 1\}$ random variables. We use the $\{-1, +1\}$ coding here because it makes the mathematics simpler and the models easier to interpret (see Haslbeck et al., 2022).

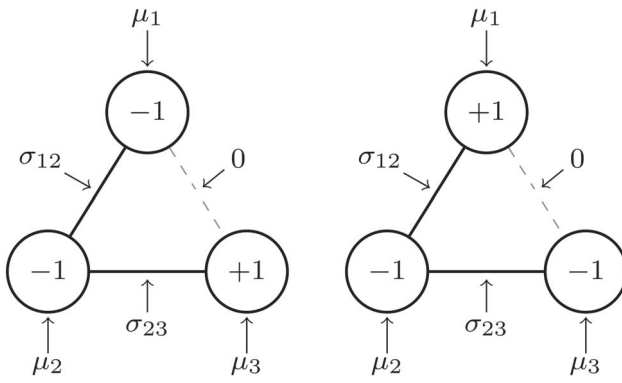


Figure 3. Two realizations of the Ising model for the three node network. Present relations are indicated with black solid line and absent relations with a gray dashed line.

main effects μ_i only affect the individual nodes, while the associations σ_{ij} affect pairs of nodes, i.e., it encodes their relations. In particular, the absence of a relation between two variables, such as the relation between variables X_1 and X_3 , is encoded as an absent association $-\sigma_{13} = 0$ in the Ising model.

In the theory of magnetism, the Ising model is used to model the spin magnetic moments of electrons, which are restricted to either point up ($X = 1$) or point down ($X = -1$). The network's structure then encodes the relations between electrons. If two electrons are close to each other, the electrons are related, and their magnetic moments tend to align. In psychology, the Ising model is used to model the presence or absence of psychological symptoms, the positive or negative evaluation of a political statement, or the correct or incorrect responses to an intelligence test in the population. A positive value ($X_i = +1$) then indicates the presence of a symptom, a positive evaluation of some statement, a correct response to a test question, and a negative value ($X_i = -1$) then indicates the absence of that symptom, the negative evaluation of that statement, or the incorrect response to that test question, respectively.

Observe that the Ising model formulated in Equation (2) does not include a parameter that allows us to express individual differences. Thus, in the Ising model, individual variation can only be summarized in terms of the variation in the configurations of symptoms, attitudes, or responses of persons. Other than that, the model and the structure it describes are invariant across individuals. A natural way to express individual differences with the Ising model will be the topic of the ensuing sections.

The idiographic Ising model: a framework for encompassing networks

In a series of papers, Fortuin and Kasteleyn worked out how the theories of percolation and magnetism

relate (Fortuin, 1972a, 1972b; Fortuin & Kasteleyn, 1972). They linked the two theories by formulating an encompassing network model that simultaneously describes the states of the variables in the network and the relations between them:

$$\begin{aligned} p(\mathbf{X} = \mathbf{x}, \mathbf{W} = \mathbf{w}) &= p(\mathbf{X} = \mathbf{x} | \mathbf{W} = \mathbf{w}) p(\mathbf{W} = \mathbf{w}) \\ &= p(\mathbf{W} = \mathbf{w} | \mathbf{X} = \mathbf{x}) p(\mathbf{X} = \mathbf{x}). \end{aligned}$$

We refer to this model as the encompassing network because it encompasses both types of network models: A random graph model is used to model the network's topology and a graphical model is used to model the network's node states. The random graph model is the edge-marginal of the encompassing network model,

$$p(\mathbf{W} = \mathbf{w}) = \sum_{\mathbf{x}} p(\mathbf{X} = \mathbf{x}, \mathbf{W} = \mathbf{w}),$$

and the graphical model is its node-marginal,

$$p(\mathbf{X} = \mathbf{x}) = \sum_{\mathbf{w}} p(\mathbf{X} = \mathbf{x}, \mathbf{W} = \mathbf{w}),$$

where the sums are with respect to all possible node states and edge states, respectively. The idea of this encompassing network model is thus that both nodes and edges are random variables. The right panel of Figure 1 illustrates this random-node, random-edge model, which combines the ideas from the random-edge model in the left panel and the random-node model in the middle panel.

Fortuin and Kasteleyn used an encompassing network approach to show that the probability of observing a pattern of node states $\mathbf{x} = (x_1, \dots, x_n)^T$ in the Ising model can be characterized as follows (see Appendix A for a proof),

$$\overbrace{p(\mathbf{X} = \mathbf{x})}^{\text{Ising model}} = \sum_{\mathbf{w}} \overbrace{p(\mathbf{X} = \mathbf{x} | \mathbf{W} = \mathbf{w})}^{\text{graph coloring}} \overbrace{p(\mathbf{W} = \mathbf{w})}^{\text{random-cluster model}}, \quad (3)$$

where a random graph model $p(\mathbf{w})$ known as the random-cluster model is used to express the probability of observing a topological structure $\mathbf{w} = (w_{12}, \dots, w_{n(n-1)})^T$, and the conditional distribution $p(\mathbf{x} | \mathbf{w})$ expresses the probability of observing the pattern of variable states $\mathbf{x}$ given the topology $\mathbf{w}$. Since the node states are arbitrarily labeled, the labels are sometimes called colors, and the process of assigning values to the nodes is then called coloring. Thus, $p(\mathbf{x} | \mathbf{w})$ describes the process of coloring the nodes of the graph that is induced by $\mathbf{w}$.

What makes the encompassing network in Equation (3) so special? And why should we care

about its random graph and graph coloring models? Two aspects make Fortuin and Kasteleyn's characterization of the Ising model valuable for modeling psychological phenomena. First, when we assume that the network relations can differ between persons, the persons are characterized by idiographic topologies. Modeling a population of topologies leads us to consider random graph models. Second, we will discuss how the graph coloring model in Equation (3) stipulates a direct connection between the links in the idiographic networks and their correlations in the cross-sectional model. We believe that this connection is essential because it shows us how established psychometric phenomena can emerge from models of the individual. We lose this connection if we abandon the graph coloring model suggested by Fortuin and Kasteleyn in favor of other models (e.g., the coloring process for the random-triangle model that we describe below). Given this graph coloring model, however, only the random-cluster model can generate an Ising model. We will also consider a variant of the encompassing network model that includes a different random graph model, the Erdős-Rényi model. This variant retains the essential connection between idiographic network links and cross-sectional correlations, but no longer reflects the Ising model. We will later see that the encompassing network setup using the Erdős-Rényi model can accommodate data coming from encompassing networks using a random-cluster model for the topologies. We will detail the models and their implications next.

The graph coloring model

The following deterministic rule underlies the encompassing network approach: If two variables are connected, then they must be in the same state:

$$(W_{ij} = 1) \Rightarrow (X_i = X_j).$$

This rule states that an edge or link between variables causes their values to align, and offers a simplified contact process—a model for a disease spreading about the graph's nodes (Grimmett, 2018)—in which two nodes infect each other, and align their values, when they are in direct contact or connected. Van Borkulo et al. (2017) used similar ideas to model the symptom dynamics underlying depression. While the rule does not imply that two aligned variables are connected, we can use it to learn about the graph's topology based on node alignment (i.e., how the disease has spread). Since, by consequence of the above

rule, if two variables are in different states, then they must be disconnected:

$$(X_i \neq X_j) \Rightarrow (W_{ij} = 0).$$

We can thus establish that certain direct links are impossible by tracing which nodes are affected and which nodes are not.⁴

In the contact process, the graph's topology is an essential aspect of a disease's dynamics. In Fortuin and Kasteleyn's encompassing network, this is no different. The deterministic rule underlying the encompassing network identify open paths or clusters as a focal point in modeling the latent topologies. Clusters are groups of directly and indirectly related nodes that, by definition, align their values. This could be a group of co-existing symptoms or attitudes or a set of co-related item responses. The graph coloring model assumes that the network's clusters are independently colored, and is equal to

$$p(\mathbf{X} = \mathbf{x} | \mathbf{w}) = \prod_{c=1}^{\kappa(\mathbf{w})} P(\mathbf{X}_c = \mathbf{1})^{\frac{x_c+1}{2}} P(\mathbf{X}_c = -\mathbf{1})^{1-\frac{x_c+1}{2}} = \frac{1}{2^{\kappa(\mathbf{w})}}, \quad (4)$$

where $\mathbf{X}_c$ denotes the observables in a cluster c and x_c denotes their collective state. The sampling mechanism inspired by this graph coloring model is relatively simple: If the nodes are in the same cluster, then each node in that cluster is set equal to $+1$ with probability $1/2$ or set equal to -1 with probability $1/2$. Fortuin and Kasteleyn showed that this graph coloring model corresponds to an Ising model sans main effects.

Cioletti and Vila (2016) have recently shown how to extend Fortuin and Kasteleyn's formulation to Ising models that include main effects. Their generalization shows that the Ising model's main effects impact the weight assigned to the cluster's values. The coloring probabilities in Cioletti and Vila's generalization become,

$$P(\mathbf{X}_c = \mathbf{1}) = \frac{\exp(\sum_{i \in V_c} \mu_i)}{\exp(\sum_{i \in V_c} \mu_i) + \exp(-\sum_{i \in V_c} \mu_i)},$$

where V_c denotes the set of indices of vertices in cluster c . In this model, there is a greater probability of

⁴Estimating edges from variable alignment (e.g., symptom co-occurrence) is closely related to a recently proposed method for inferring idiographic networks of binary variables by Bodner et al. (2022). Bodner and colleagues proposed to use the Jaccard index, pitting the number of measurements that two variables were both active (i.e., $X_i = 1$ and $X_j = 1$) against the total number of measurements of the two variables excluding the number of measurements in which the variables were both inactive (i.e., $X_i = -1$ and $X_j = -1$). Whereas Bodner et al.'s approach offers a non-parametric approach to estimate the marginal relations in idiographic topologies—considering edges in isolation—the encompassing network model offers a model-based approach for their conditional relations.

generating a positive than generating a negative value for the cluster's nodes if their main effects are positive. In contrast, there is a smaller probability of generating a positive than generating a negative value for the cluster's nodes if their main effects are negative. In psychopathology, for example, the main effects are typically negative to cover the overall absence of symptoms in the population. In education, on the other hand, the main effects are typically positive to cover the overall ability to solve test items.

The graph coloring model introduced by Fortuin and Kasteleyn and generalized by Cioletti and Vila opens up a direct connection between the idiographic topologies and cross-sectional correlations. We discuss this relation below. The focus of this graph coloring model is the network's clusters or open paths. In many applications of random graphs and particularly social networks, the focus is on the network's triangles instead, which leads to the well-known catch-phrase *a friend of my friend is my friend* in friendship networks. Jonasson (1999) proposed a graph coloring model that focuses on triangles instead of clusters or open paths. The random graph model stipulated on the topologies in the triangle-inspired encompassing network was aptly named the random-triangle model. Interestingly, this encompassing network can also have an Ising model as its edge-marginal (Häggström & Jonasson, 1999), albeit constrained to a particular honeycomb lattice topology.

The distribution of idiographic topologies

In their seminal work, Fortuin and Kasteleyn (1972) showed that the edge-marginal in Equation (3) is an Ising model if the topologies are samples from a random-cluster model. The random-cluster model is defined by the following probability distribution over topological structures $\mathbf{w}$:

$$p(\mathbf{W} = \mathbf{w}) = \frac{1}{Z_R} \prod_{i=1}^{n-1} \prod_{j=i+1}^n \overbrace{\theta_{ij}^{w_{ij}} (1 - \theta_{ij})^{1-w_{ij}}}^{\text{Erdős-Rényi model}} \lambda^{\kappa(\mathbf{w})}, \quad (5)$$

where $\kappa(\mathbf{w})$ denotes the number of connected components or clusters in the topological structure that is implied by $\mathbf{w}$, λ is a positive clustering weight, and Z_R is the model's normalizing constant. Observe that the random-cluster model has the same kernel as the Erdős-Rényi model in Equation (1), but differs from the Erdős-Rényi model in terms of the weight it assigns to the network's clusters. The random-cluster

model tends to favor networks with fewer clusters when the weight is smaller than one (i.e., $\lambda < 1$) and tends to favor networks with more clusters when the weight is larger than one (i.e., $\lambda > 1$). The random-cluster model with a unit clustering weight coincides with the Erdős-Rényi model in Equation (1). Fortuin and Kasteleyn showed that the characterization in Equation (3) coincides with an Ising model sans main effects if the topologies are samples from a random-cluster model with a clustering weight of 2.

We use a small simulation to illustrate the influence of the clustering weight on the distribution of topological structures. We simulate topological structures using a clustering weight of 0.5, 1, and 2 and plot the associated cluster distribution in Figure 4. The distributions in Figure 4 confirm that the topological structures, on average, have more distinct clusters with increasing values of the clustering weight. For a clustering weight equal to 0.5, for example, most of the generated networks have two to four components. For a clustering weight of 2, on the other hand, the number of clusters for most of the generated networks increased to seven or more. The cluster distribution for the unit clustering weight resides in between the two other cases, and here most of the generated networks comprise four to seven clusters. In this case, the random-cluster model coincides with the Erdős-Rényi model.

The generalization of Cioletti and Vila (2016) shows that the main effects impact the clustering weights of the random-cluster model, which evolve to

$$\begin{aligned} \lambda_c &= 2 \cosh \left(\sum_{i \in V_c} \mu_i \right) \\ &= \exp \left(\sum_{i \in V_c} \mu_i \right) + \exp \left(- \sum_{i \in V_c} \mu_i \right) \geq 2. \end{aligned} \quad (6)$$

Thus, by including main effects, the random-cluster model assigns different weights to the different clusters,

$$p(\mathbf{W} = \mathbf{w}) = \frac{1}{Z_R} \prod_{i=1}^{n-1} \prod_{j=i+1}^n \theta_{ij}^{w_{ij}} (1 - \theta_{ij})^{1-w_{ij}} \prod_{c=1}^{\kappa(\mathbf{w})} \lambda_c.$$

Since the main effects in empirical applications typically do not sum to zero, their inclusion thus leads to even more fragmentation in the idiographic networks. In psychopathology, for example, the main effects are typically negative to cover the overall absence of symptoms in the population. In education, on the other hand, the main effects are typically positive to cover the overall ability to solve test items. In both

applications, the clustering weights will then be strictly larger than two.

As alluded to at the end of the opening section, there might be reasons that persuade us to consider other population models for the latent topologies. One reason might be the random-cluster model's analytic intractability. The normalizing constant

$$Z_R = \sum_{\mathbf{w}} \prod_{i=1}^{n-1} \prod_{j=i+1}^n \theta_{ij}^{w_{ij}} (1 - \theta_{ij})^{1-w_{ij}} \prod_{c=1}^{\kappa(\mathbf{w})} \lambda_c,$$

involves a sum over all $2^{\frac{1}{2}n(n-1)}$ possible topological states $\mathbf{w}$, and may be difficult to evaluate in practice. Another reason might be the empirical fit of the random-cluster model—and thus also the Ising model. One solution is to use the Erdős-Rényi model rather than the full random-cluster model for the distribution of the latent topologies. The Erdős-Rényi model is analytically tractable and offers great modeling flexibility (see the Encompassing network modeling section in the discussion). The application of an Erdős-Rényi model to describe the distribution of random graphs in Equation (3) is known as the *divide and color model* (Häggström, 2001).

$$\begin{array}{c} \text{Divide and Color model} \qquad \qquad \text{graph coloring} \qquad \text{Erdős-Rényi model} \\ \overbrace{p(\mathbf{X} = \mathbf{x})} = \sum_{\mathbf{w}} \overbrace{p(\mathbf{X} = \mathbf{x} | \mathbf{W} = \mathbf{w})} \overbrace{p(\mathbf{W} = \mathbf{w})}. \end{array} \quad (7)$$

The divide and color model in Equation (7) makes use of the same graph coloring mechanism as Fortuin and Kasteleyn's original encompassing network in Equation (3), but stipulates a different model for the topologies. We will investigate the statistical analysis of the two encompassing network approaches after we have evaluated the relation between idiographic networks and cross-sectional phenomena.

The relation between idiographic topologies and cross-sectional phenomena

The divide and color model does not coincide with an Ising model. However, Häggström (2001) showed that the model does share several features with the particular Ising model in Fortuin and Kasteleyn's formulation. In particular, that the population correlation between any two observables X_i and X_j in the network is non-negative (Fortuin et al., 1971). This observation reveals a significant restriction on both the divide and color model and the particular Ising model that emerges from Fortuin and Kasteleyn's theory. However, the finding that observables non-negatively correlate in the population is consistent with other

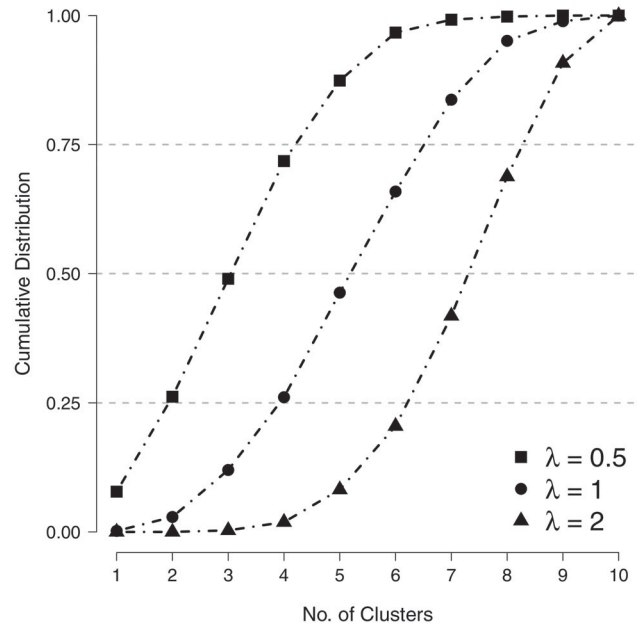


Figure 4. The (cumulative) distribution of clusters in ten-variable networks that were generated from random cluster models with constant edge probabilities $\theta_{ij} = \theta = 0.1$ but with distinct values for the clustering weight.

latent variable approaches (Holland & Rosenbaum, 1986), and with the robust finding of a positive manifold in psychometric research (e.g., Caspi et al., 2014; van der Maas et al., 2006).

The result that observables non-negatively correlate in the divide and color model in Equation (7) and the Fortuin-Kasteleyn model in Equation (3) reveals a deep connection between observables and underlying topologies. This relationship between two observables in the network and their connection in the latent topologies can be made precise (c.f. Grimmet, 2006, Theorem 1.16; Steif & Tykesson, 2017, Eq. 1.1)

$$P(X_i = X_j) = \frac{1}{2} + \frac{1}{2} P(X_i \leftrightarrow X_j) \geq \frac{1}{2}, \quad (8)$$

where $P(X_i \leftrightarrow X_j)$ denotes the probability that the two variables end up in the same cluster. The two-point correlation $\tau(X_i, X_j) = P(X_i = X_j) - \frac{1}{2}$ then equals $\frac{1}{2}P(X_i \leftrightarrow X_j)$, which confirms the non-negative population correlations. Since the states of observables align when they are in the same cluster, their cross-sectional correlation will be high if they end up in the same cluster for many individuals. Thus, the way that the latent topologies cluster has a tremendous impact on observed correlations. The relation in Equation (8) holds for every encompassing network model that uses Fortuin and Kasteleyn's graph coloring model, i.e., for every random graph model $p(\mathbf{w})$ stipulated on the latent topologies. However,

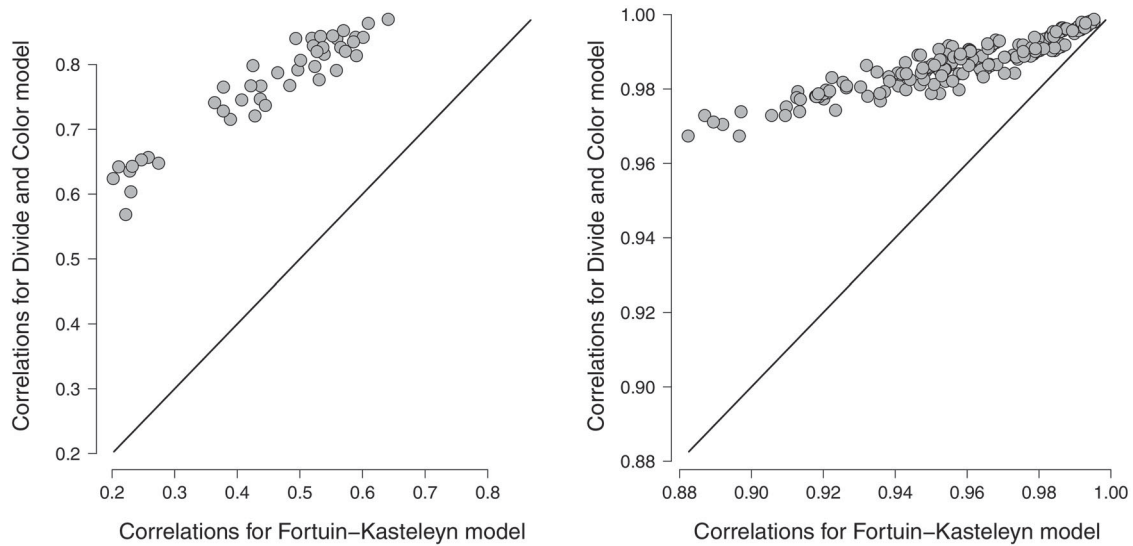


Figure 5. Scatterplots of observed correlations between node states — $\text{Cor}(X_i, X_j)$ — for the divide and color model and the Fortuin and Kasteleyn model for a $n = 10$ variable network in the left panel, and a $n = 20$ variable network in the right panel. The correlations were based on data from $N = 10,000$ cases. The edge probabilities θ_{ij} were sampled uniformly between 0 and 0.5 and the thresholds μ_i were simulated from a standard normal distribution.

Equation (8) only pertains to scenarios that involve no main effects. Equation (1.1) in Steif and Tykesson (2017) and Theorem 2 in Cioletti and Vila (2016) offer the generalizations that are relevant for scenarios that involve main effects.

The topologies that are consistent with the Ising model are heavier fragmented than the topologies that underly the divide and color model. This difference in fragmentation is the result of using a smaller clustering weight in the divide and color model, c.f. Figure 4. Since the edge inclusion probabilities are higher when there is less fragmentation, the two-point population correlations will be higher for the divide and color model than for the Fortuin and Kasteleyn model. In Figure 5 we show that, for models with the same edge inclusion probabilities, the correlations are higher for the divide-and-color than for Fortuin and Kasteleyn's idiographic Ising model. The graph generating mechanisms of the random-cluster model and the Erdős-Rényi model are thus fundamentally different and may lead to divergent predictions and interventions on underlying topologies.

That population correlations in Fortuin and Kasteleyn's theory are a function of the edge inclusion probabilities of the variable pairs hints at a relation between associations in the Ising model and edge probabilities in the random-cluster model in Equation (3). This relationship indeed exists for the Fortuin-Kasteleyn model,

$$\sigma_{ij} = -\frac{1}{2} \log(1 - \theta_{ij}). \quad (9)$$

We illustrate the relation in Figure 6. Observe that the relation is only valid for non-negative associations and is consistent with the restriction to non-negative population correlations. The generalization of the relation to negative associations has been the topic of several investigations that we will not pursue further here (e.g., C. Newman, 1991; Swendsen & Wang, 1987).

The relationship in Equation (9) yields the bridge between group-level associations and links in idiographic networks. On the one hand, it allows us to interpret group-level associations in terms of the proportion of observed edges between variables in the idiographic networks. One substantial result, for example, is that absent associations at the group-level imply the absence of a link between two variables in all idiographic networks. Thus, the idiographic networks mirror the conditional independence property of the Ising model. On the other hand, the relation in Equation (9) also underscores the difference between group-level associations and links at the individual level. At the individual level, two variables will be entirely dependent if they are in the same cluster and independent otherwise. At the group-level, however, this can still generate a wide variety of associations. Moreover, the fact that patterns in observed correlations are reflections of patterns in the underlying edge inclusion probabilities provides us with an alternative view on psychometric models. Savi et al. (2019), for example, characterizes the bifactor structure, which is ubiquitous in psychometrics, in terms of an underlying community or

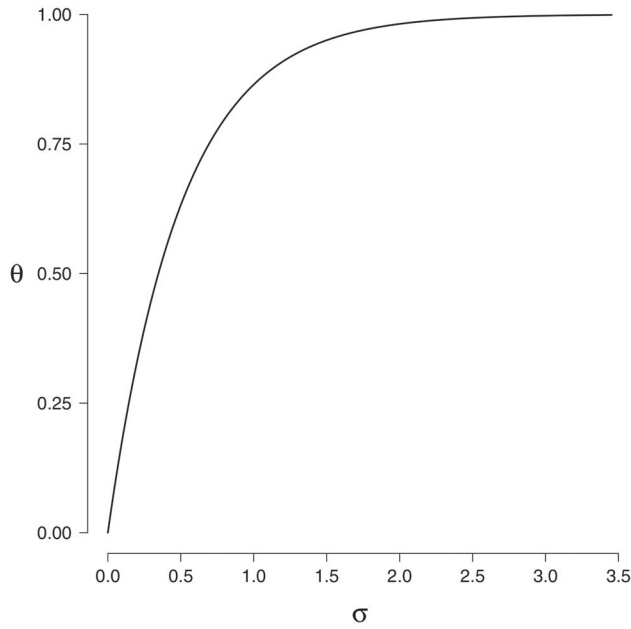


Figure 6. The formal relation between the association parameter σ of the Ising model and the edge probability θ in the random-cluster model.

stochastic block structure. This connection between psychometric models and models from network science offers another exciting avenue for future research (Marsman et al., 2022).

The statistical analysis of encompassing networks

Our theory uses random graphs to accommodate the heterogeneity of associations in the population (i.e., idiographic topologies). However, these random graphs are not directly observed and need to be estimated from the data. We propose to estimate these latent topologies using their posterior distribution. If we assume that k repeated observations $\mathbf{x}_1, \dots, \mathbf{x}_k$ on the network's nodes for a person are independent given the topology $\mathbf{w}$ —the topological equivalence of local independence—the posterior distribution can consistently estimate the clusters of the latent graph. We obtain the Ising model and the divide and color models as special cases of this model setup with $k=1$ measurements and using the random-cluster and Erdős-Rényi priors, respectively. We will establish two important statistical results for this model setup. First, we will show that the random-cluster model is conjugate to the graph coloring model, which allows us to derive a closed-form expression for the posterior distribution of the latent topology. This result holds for scenarios that use a random-cluster model or an Erdős-Rényi model as topological prior. Unfortunately, the closed-form expression does not

make it easy to evaluate the posterior because random-cluster models are intractable. We therefore introduce a Gibbs sampling algorithm to sample from the intractable topological posterior, which allows us to evaluate the posterior via simulations.

Second, the possibility to simulate from the topological posterior opens up the statistical analysis of encompassing networks. Observe that the encompassing network that is consistent with the Ising model (i.e., $k=1$ and a random-cluster model as topological prior) can be estimated using standard software: Equation (9) shows how to translate the estimated edge weights to edge inclusion probabilities, and the proposed posterior simulation method can then be used to estimate the idiographic topologies. The only caveat is that the estimated associations need to be positive. Below, we will use the proposed posterior simulation method to develop a Gibbs sampling algorithm for estimating encompassing networks for $k \geq 1$ using the Erdős-Rényi model as the topological prior. Unfortunately, we could not extend this methodology to encompassing networks that use a random-cluster model as topological prior: The random-cluster model's normalizing constant and its dependency on the network's main effects leads to significant computational challenges. We illustrate the developments using simulated and empirical data in the next section. There, we will also show that encompassing networks using an Erdős-Rényi model as topological prior are high indistinguishable from encompassing networks using a random-cluster model.

The posterior distribution of idiographic networks

If we interpret the random-cluster model to be a prior distribution for the underlying graphs, we can integrate the information from this prior distribution with the coloring process to form posterior distributions of the underlying graphs. This posterior distribution is a specific instance of the Erdős-Rényi model (see Appendix A for a derivation),

$$p(\mathbf{w}|\mathbf{x}) = \prod_{i=1}^{n-1} \prod_{j=i+1}^n \left\{ \begin{array}{l} \text{\textit{x}_i \text{ and } x_j \text{ in the same state}} \\ \left[\theta_{ij}^{w_{ij}} (1 - \theta_{ij})^{1-w_{ij}} \right]^{\delta_{(x_i, x_j)}} \\ \text{\textit{x}_i \text{ and } x_j \text{ in different states}} \\ \times \left[\theta_{ij}^{w_{ij}} 1^{1-w_{ij}} \right]^{1-\delta_{(x_i, x_j)}} \end{array} \right\},$$

where $\delta_{(x_i, x_j)}$ is an indicator function that is equal to one if $x_i = x_j$ and is equal to zero otherwise. The expression of the posterior distribution reveals two

things. First, the probabilities associated to edges between variables that are in different states are zero in the posterior distribution. The adjustment of these edge probabilities follows from the way that graphs are colored, which forces connected nodes to have the same state. As a result, nodes that are in different states could not have connected. Second, the posterior distribution does not consider the network's clusters. In contrast, the prior distribution explicitly does. As a result of these two properties, the edges of the graph are independent Bernoulli variables in the posterior distribution, and the distribution of the edges only depends on its adjacent response variables:

$$\begin{aligned} p(W_{ij} = w_{ij} | x_i = x_j) &= \theta_{ij}^{w_{ij}} (1 - \theta_{ij})^{1-w_{ij}} \\ p(W_{ij} = w_{ij} | x_i \neq x_j) &= 0^{w_{ij}} 1^{1-w_{ij}}, \end{aligned}$$

where the latter expression implies that there can be no edges between variables that are in different states.

Since the posterior distribution comprises independent Bernoulli variables, it is easy to summarize and provide a Bayesian estimate of the latent topology. The posterior mean of a particular edge W_{ij} , for example, is equal to the edge probability θ_{ij} whenever the adjacent variables are aligned or is equal to zero otherwise. Similarly, the posterior variance of this edge is equal to

$$\text{Var}(W_{ij} | x_i, x_j) = \begin{cases} \theta_{ij}(1 - \theta_{ij}) & \text{if } x_i = x_j \\ 0 & \text{if } x_i \neq x_j, \end{cases}$$

which reveals perfect knowledge about edges that lie between variables that are in different states, but there is still much to learn about the remaining edges. Another convenient Bayesian estimate of the idiographic topologies are plausible values (Mislevy, 1991), which we will refer to as plausible networks. A plausible network constitutes a sample from the posterior distribution, and could, in principle, reflect the true underlying network structure.

The posterior distribution that we have derived above concerns a single observation of the idiographic network. It turned out that the posterior distribution is a particular instance of the Erdős-Rényi model, in this case. The posterior distribution based on two or more observations of the idiographic network, however, is a random cluster model (see Appendix A). In particular, suppose that we have k observations $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k$ on the idiographic network of a person, then his or her posterior distribution is equal to a random-cluster model with edge probabilities θ'_{ij} and clustering weights λ'_c . The posterior edge probabilities are specified as

$$\theta'_{ij} = \begin{cases} \theta_{ij} & \text{if } \delta_{ij} = 1 \\ 0 & \text{if } \delta_{ij} = 0 \end{cases}$$

where δ_{ij} is an indicator function that is equal to one if variables x_i and x_j aligned on each of the k occasions, and is equal to zero otherwise. Furthermore, the posterior clustering weight is $\lambda'_c = \lambda_c^{1-k}$, and is equal to one when $k=1$. This posterior distribution reveals three essential properties. First, the random-cluster model is the conjugate prior for the cluster-based coloring process.⁵ Observe that the posterior distribution under the Erdős-Rényi prior—a special case of the random-cluster prior—is also a random-cluster distribution, but with clustering weights $\lambda'_c = \lambda_c^{1-k}$. Second, with more observations, it becomes more likely that the posterior distribution excludes inconsistent edges. Third, the clustering weights of the posterior distribution, i.e., $\lambda'_c = \lambda_c^{1-k}$, are non-increasing functions of k when $\lambda_c \geq 1$. This implies that, given the same value for δ , the posterior distribution tends to favor less fragmented networks if the number of observations k increase. Taken together, the last two observations imply that we (i) accumulate evidence about the clusters of the underlying network and (ii) become more confident that the remaining edges comprise a minimum number of clusters.

Even though conjugacy provides a simple way to update the posterior distribution in light of new observations, the posterior distribution is not easily summarized. This complication is entirely due to the random-cluster model, which does not have simple analytic expressions for the mean and variance of edges in its network, for example. The absent edges—edges w_{ij} for which $\delta_{ij} = 0$ —are an obvious exception, of course, as their posterior expectation and posterior variance are both equal to zero. One way to summarize and analyze the posterior distribution is through simulation. Several methods have been proposed for simulating networks from the random-cluster model, see, for example, Häggström (2002) and Grimmet (2006) for some approaches. However, these methods focus on specific instances of the random-cluster model for which the clustering weights λ' are either integer and larger than one. The posterior clustering weights λ'_c are non-integer when the encompassing network includes main effects and are smaller than or equal to one for $k \geq 1$. We thus need another way to simulate from the random-cluster model. In Appendix B, we derive a straightforward Gibbs sampling approach that works for any non-negative value for the clustering weights.

⁵Similarly, one finds that the random-triangle model is the conjugate prior for the triangle-based coloring process that was analyzed by Jonasson (1999) and Häggström and Jonasson (1999).

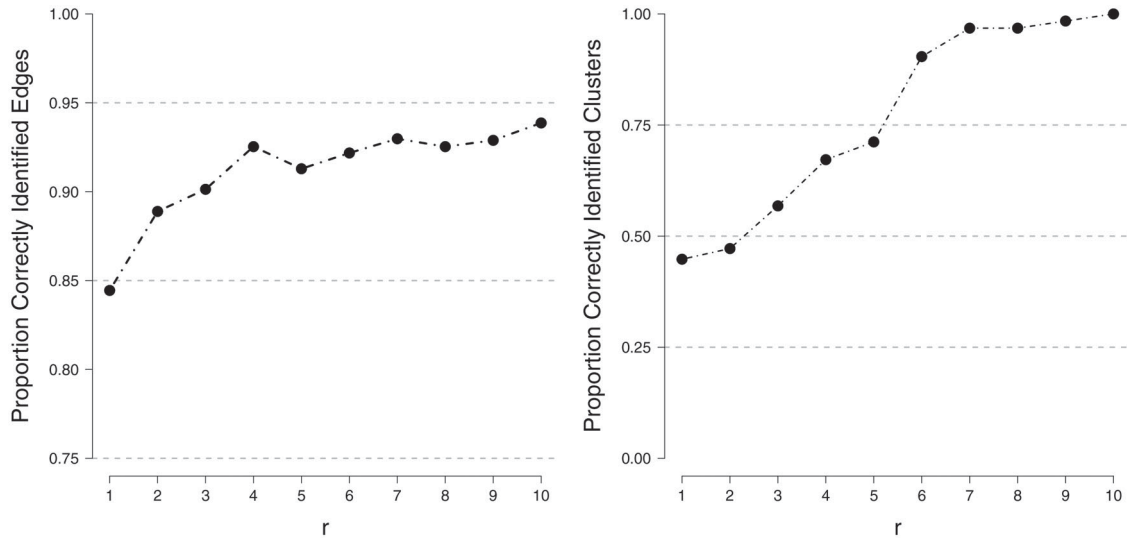


Figure 7. An $n = 10$ variable network was generated for a single person and 25 plausible values were generated based on r observations of the network. The left panel shows the proportion of correctly identified edges and the right panel the proportion of correctly identified clusters (averaged over plausible networks).

We illustrate the results above for estimating a single idiographic network. We generate an idiographic topology for $n = 10$ nodes from a random-cluster model with a constant clustering weight of 2 and constant edge probability $\theta_{ij} = \theta = 0.1$, and generate $k = 10$ observations on the idiographic network. Using the Gibbs sampler from [Appendix B](#), we generate 25 plausible networks from the posterior distribution after $r = 1, 2, \dots, k$ observations. [Figure 7](#) shows the average proportion of correctly identified edges and clusters in the plausible networks after r observations. Here, all clusters seem to be correctly identified after approximately six observations, whereas the proportion of correctly identified edges seems to hit an upper bound.

A Gibbs sampling algorithm to estimate the encompassing network

We set up a Gibbs sampling approach to estimate the encompassing network and use the new sampling procedure to impute the unobserved or missing topologies. We treat the graph coloring model as a likelihood. Assuming that individuals are independent, we have the following likelihood for a person p , $p = 1, \dots, N$,

$$\begin{aligned} L(\mathbf{w}_p, \boldsymbol{\mu} | \mathbf{X}_p) &= \prod_{r=1}^k p(\mathbf{x}_{pr} | \mathbf{w}_p, \boldsymbol{\mu}) \\ &= \frac{\exp\left(\sum_{r=1}^k \sum_{i=1}^n x_{ri} \mu_i\right)}{\left(\prod_{c=1}^{K(\mathbf{w}_p)} 2 \cosh(\sum_{i \in V_c} \mu_i)\right)^k} \mathbf{1}(\mathbf{x}_1, \dots, \mathbf{x}_k, \mathbf{w}_p) \end{aligned}$$

where $\mathbf{1}(\mathbf{x}, \mathbf{w})$ is an indicator function for the set

$$\{\mathbf{x}_1, \dots, \mathbf{x}_k, \mathbf{w} : w_{ij} = 1 \Rightarrow x_{ri} = x_{rj}, \text{ for } i = 1, \dots, k\}.$$

In the encompassing network, a random graph model is used as a population model for the latent topologies, which acts as a topological prior on the idiographic networks. Here, we consider the Erdős-Rényi model as a topological prior,

$$p(\mathbf{w}_p | \boldsymbol{\theta}) = \prod_{i=1}^{n-1} \prod_{j=i+1}^n \theta_{ij}^{w_{ij}} (1 - \theta_{ij})^{1-w_{ij}}.$$

This allows us to formulate a Gibbs sampler to estimate the free parameters of the encompassing network. The main effects $\boldsymbol{\mu}$ and the edge inclusion probabilities $\boldsymbol{\theta}$. We will set up a Gibbs sampler to produce random samples from the multivariate posterior distribution,

$$\begin{aligned} p(\mathbf{w}_1, \dots, \mathbf{w}_N, \boldsymbol{\theta}, \boldsymbol{\mu} | \mathbf{X}) &\propto \\ \prod_{p=1}^N \prod_{r=1}^k p(\mathbf{x}_{pr} | \mathbf{w}_p, \boldsymbol{\mu}) p(\mathbf{w}_p | \boldsymbol{\theta}) p(\boldsymbol{\mu}, \boldsymbol{\theta}), \end{aligned}$$

where $p(\boldsymbol{\mu}, \boldsymbol{\theta})$ denotes a prior distribution for the model parameters. We will assume independence between main effects and edge inclusion probabilities *a priori*, and discuss the individual prior components below. The joint posterior has an intractable form and will be difficult to simulate. We will therefore use a Gibbs sampling algorithm to simulate from the multivariate posterior. Our proposed Gibbs sampler comprises three steps. Each step consists of sampling from the posterior of one of the three parameters $\mathbf{w}$, $\boldsymbol{\mu}$, and

θ conditionally upon all other parameters. We derive the desired full-conditional posterior distributions and discuss how to simulate from them in Appendix C.

Simulated and real-data examples

In this section, we analyze several simulated data sets and an empirical longitudinal data set of symptoms of depression. The simulated data sets will be used to illustrate the recovery of model parameters with the Gibbs sampler, and the analysis of model misspecification. The longitudinal data set will be used to illustrate an encompassing network analysis of panel data.

Numerical example I: parameter recovery in two scenarios

To illustrate parameter recovery, we simulated data for a $n=10$ node encompassing network. The $\frac{1}{2}n(n-1) = 45$ edge inclusion probabilities θ of the Erdős-Rényi model were simulated from a Beta (1, 10) distribution, and the $n=10$ thresholds μ were simulated from a Normal (0, 1) distribution. We simulated data for two scenarios, with $N \times k = 5,000$ observations each. In the first scenario, we sampled $N=500$ topologies from the simulated Erdős-Rényi model and $k=10$ colorings for each topology. In the second scenario, we sampled $N=50$ topologies and $k=100$ colorings for each topology.

With random starting values for the model parameters and topologies, we estimated the parameters for the two scenarios using the Gibbs sampler described in the previous section. We ran the Gibbs samplers for 100,000 iterations each.⁶ The proposed Metropolis algorithm for the thresholds—an independence chain Metropolis algorithm—performed well in this example. Across all 100,000 iterations of the Gibbs sampler, the two scenarios, and the p threshold parameters, all but three proposed values were accepted. The first 5,000 sampled states of μ_1 and θ_{12} —the trace plots—for both scenarios are shown in Figure S1 in the supplementary material. The Gibbs sampler quickly converged for both scenarios.

Figure 8 displays scatter plots of the posterior mean estimates of the thresholds and edge inclusion probabilities for the two scenarios against the data generating parameter values. The horizontal bars represent the 95% highest posterior density intervals—the shortest interval that covers 95% of the posterior mass. All 100,000 iterations of the Gibbs sampling algorithm were used to estimate the posterior means. The scatter plots clearly show that the thresholds were accurately estimated in both scenarios. The trace plots of μ_1 displayed in Figure S1 in the supplementary material suggests that the posterior uncertainty for the thresholds was fairly constant across the two scenarios, and Figure 8 confirms this. The edge inclusion probabilities, however, showed a much lower uncertainty for the $N=500$ topologies scenario than for the $N=50$ topologies scenario, which was also clearly visible in the trace plots of θ_{12} that were displayed in Figure S1 in the supplementary material.

The proportion of correctly identified edges in the latent topology's final state in the Gibbs sampler averaged to 93% for both scenarios. On the other hand, the proportion of correctly identified clusters averaged to approximately 93% in the $k=10$ colorings scenario, and increased to 98% in the $k=100$ colorings scenario. These results confirm our earlier suggestion that there appears to be an upper bound for the proportion of correctly identified edges, but not for the proportion of correctly identified clusters (c.f., Figure 7).

Numerical example II: random-cluster topologies

How does the encompassing network using an Erdős-Rényi model for the latent topologies perform when the topologies are, in fact, sampled from a random-cluster model? We investigate this scenario by using the model from the preceding sections to analyze data generated from a random-cluster distribution for the topologies. We first focus on Scenario 1 of the previous simulation: Data from $N=500$ persons on $k=10$ colorings or measurements were simulated for a $n=10$ variable network. As before, we simulated the $\frac{1}{2}n(n-1) = 45$ edge inclusion probabilities θ from a Beta (1, 10) distribution, and the $p=10$ thresholds μ from a Normal (0, 1) distribution. We generated the $N=500$ topologies from a random-cluster model using the simulated edge inclusion probabilities and thresholds. For a single coloring, this set-up would boil down to an idiographic Ising model.

With random starting values for the model parameters and topologies, we ran the Gibbs sampler for 100,000 iterations. The proposed Metropolis algorithm

⁶We performed all analyses in R (R Core Team, 2019) on a Macbook Pro with 6 Intel Core i9 2.9GHz processors. The Gibbs sampler's 100,000 iterations took about 18 hours for the $N=500$ topologies, $k=10$ colorings scenario, and about 4 hours for the $N=50$ topologies, $k=100$ colorings scenario. We updated the latent topologies and evaluate their clustering in parallel (12 virtual cores on the Macbook). Topological clustering was assessed using the igraph R-package (Csardi & Nepusz, 2019). The R-code that we used for parameter recovery is available at <https://osf.io/beqgr/>.

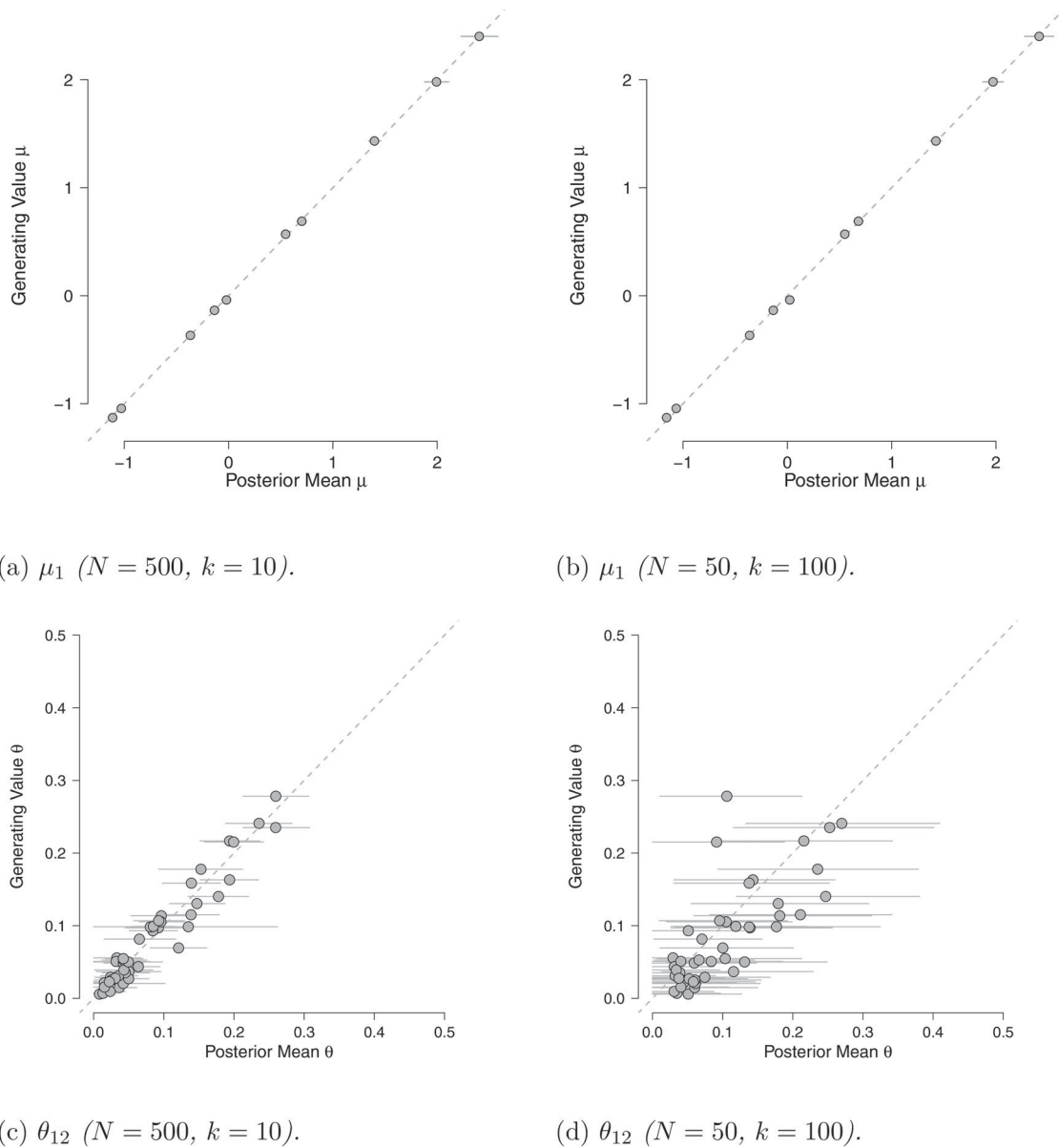


Figure 8. Scatter plots of the posterior mean estimates against the data generating values for the thresholds μ and θ in the two synthetic scenarios. The horizontal gray lines reflect the 95% highest posterior density interval.

for the thresholds—an independence chain Metropolis algorithm—again performed well in this example. Across all 100,000 iterations of the Gibbs sampler and the p threshold parameters, all but three proposed values were accepted. The first 5,000 sampled states of μ_1 and σ_{12} are shown in Figure S2 in the [supplementary material](#). The Gibbs sampler converged quickly.

Figure 9 displays scatter plots of the posterior mean estimates of the thresholds and edge inclusion probabilities against the data generating parameter values. The horizontal bars represent the 95% highest posterior density intervals. All 100,000 iterations of the Gibbs sampling algorithm were used to estimate the posterior means. The scatter plots clearly show

that the thresholds were accurately estimated, and that their estimation was unaffected by the misspecification of the population model. The estimated edge inclusion probabilities, however, clearly differed from their generating values. Given that these parameters belonged to different models, the difference is rather unsurprising.

How does the misspecified population model affect our estimates of the idiographic topologies? The proportion of correctly identified edges in the latent topology's final state in the Gibbs sampler averaged to 96%. This is an even higher percentage than we observed for the correct model set-up in the previous section. The proportion of correctly identified clusters averaged to approximately 94%, which is slightly

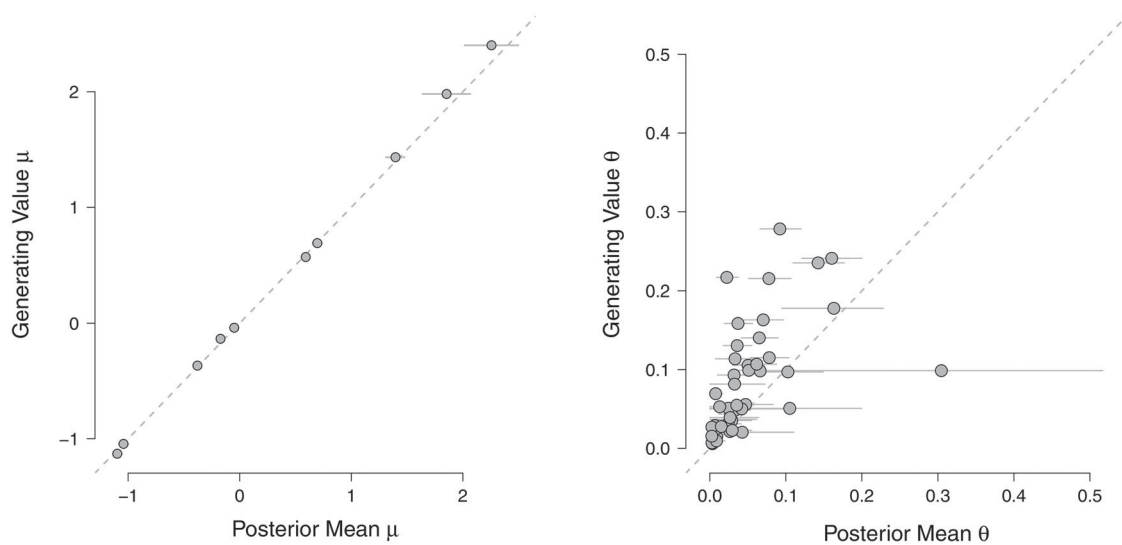


Figure 9. Scatter plots of the posterior mean estimates against the data generating values for the thresholds μ and θ . The horizontal gray lines reflect the 95% highest posterior density interval.

better than in the previous analysis. Thus, our estimates of the idiographic topologies are robust against the misspecification in the scenario that we analyzed here.

Thus far, the misspecified population model only affects estimation of the edge inclusion probabilities. The estimated edge inclusion probabilities of the Erdős-Rényi model appear to be monotonically related to, but systematically smaller than the generating edge inclusion probabilities of the random-cluster model. We expect that this allows the estimated Erdős-Rényi model to accommodate the higher fragmentation in the topologies coming from the random-cluster model (c.f., Figure 5). We next turn to this model mimicry—the ability of a model to accommodate data coming from a competing model.

Assessing the relative fit of the Erdős-Rényi model

Marsman et al. (2016) have recently shown that if a latent variable model is correctly specified, then draws from the posterior distribution of the latent variables can be used to learn about the correct underlying distribution of the latent variables. Thus, assuming that the coloring process $p(\mathbf{X}|\mathbf{W})$ is correct, we can use the distribution of plausible networks to infer if the estimated Erdős-Rényi model fits the observed data. Since the coloring process was correctly specified in this scenario, and its parameter estimates were unaffected by the misspecified population model, we can use Marsman et al.'s theory to assess the fit of the Erdős-Rényi model.

The work of Marsman et al. (2016) comprises two parts. First, they showed that the distribution of plausible values or networks diverges from the postulated

population model—here the Erdős-Rényi model—if it is misspecified. Second, they showed that the plausible network distribution converges toward the correct population distribution of the latent variables. The convergence crucially depends on the consistency of the posterior distribution. Proof of the consistency of the posterior distribution of the latent topologies goes beyond the scope of this paper. However, the causal rule that variables that wire together must be in the same state will help us identify the relevant clusters of the idiographic networks. The results in Figure 7 and the capacity of the Erdős-Rényi model to correctly identify edges and clusters in the latent topologies underline this idea.

What sets the Erdős-Rényi model and the random-cluster model apart is the fragmentation of the topologies they generate. We assess here if we can distinguish the distribution of the number of clusters—the isolated open paths—between a sample from the Erdős-Rényi model and a sample from the random-cluster model (i.e., the plausible networks). Since we do not know the distribution of the number of clusters for the two distributions, we assume here that they are a categorical variable T , say, which ranges between 1 and n . We estimate the parameters of this categorical variable by computing sample averages (i.e., the MLE). Let $\mathbf{w}_p^*$ denote a sample from the posterior distribution of person p (i.e., a plausible network) and $\mathbf{w}_p'$ denote a sample from the estimated population model. If $T_p = T(\mathbf{w}_p)$ denotes the number of clusters in a topology $\mathbf{w}_p$, then T_p^* is the p -th sample of the categorical variable according to the posterior distribution and T_p' the sample from the estimated population model. Under $\mathcal{H}_0$, the two samples come

Table 1. Simulation results for the comparison of cluster distributions of the random-cluster and Erdős-Rényi models in various scenarios. See text for details.

N	n	$\mathcal{H}_0$: Erdős-Rényi				$\mathcal{H}_1$: Random-Cluster			
		$k=1$	$k=5$	$k=10$	Obs.	$k=1$	$k=5$	$k=10$	Obs.
250	10	.06	.04	.06	.06	.05	.08	.08	.09
500	10	.05	.04	.06	.05	.06	.07	.08	.08

from a single underlying categorical distribution. That is,

$$p(\mathbf{T}^*, \mathbf{T}' | \mathcal{H}_0) = p(\mathbf{T}^*, \mathbf{T}' | \hat{\pi}) = \prod_{t=1}^n \hat{\pi}_t^{t^*+t'},$$

where t^* denotes the number of plausible networks with t clusters, t' the number of topologies sampled from the population model that had t clusters, and $\hat{\pi}_t = (t^* + t')/(2N)$. Under $\mathcal{H}_1$, the two samples come from different underlying categorical distribution. That is,

$$p(\mathbf{T}^*, \mathbf{T}' | \mathcal{H}_1) = p(\mathbf{T}^* | \hat{\pi}^*) p(\mathbf{T}' | \hat{\pi}') = \prod_{t=1}^n \hat{\pi}_t^{t^*} \hat{\pi}'_t^{t'},$$

where $\hat{\pi}_t^* = t^*/N$ and $\hat{\pi}'_t = t'/N$. We assess their distinction using a likelihood ratio statistic

$$\Lambda = -2 \log (p(\mathbf{T}^*, \mathbf{T}' | \mathcal{H}_0) / p(\mathbf{T}^*, \mathbf{T}' | \mathcal{H}_1)).$$

In our simulations below, we estimate the sampling distribution of the test statistic by simulating both $\mathbf{T}^*$ and $\mathbf{T}'$ from the estimated population model and compute the corresponding test statistic 250 times.

In Table 1 we show the results from a simulation on a range of scenarios under both $\mathcal{H}_0$ and $\mathcal{H}_1$. For each scenario, we simulated 200 datasets, and ran a Gibbs sampler for 550 iterations, including 50 burnin iterations.⁷ The proportion of p -values below 0.05 was close to its nominal value under $\mathcal{H}_0$. Under $\mathcal{H}_1$ they appear to be above this nominal value, but only slightly. It thus appears that the estimated Erdős-Rényi models can effectively mimic the random-cluster models as they can, to a large extent, account for the data that are generated from the competing random-cluster model.

To further assess the capacity of the Erdős-Rényi model to mimic the random-cluster model, we repeat the simulations leading to Table 1, except now we use the observed (i.e., simulated) topologies. With the topologies observed, the posterior mean of the edge inclusion probabilities in the Erdős-Rényi model can be directly inferred from the corresponding Beta posteriors,

$$E(\theta_{ij} | \mathbf{W}) = \frac{a + W_{+ij}}{a + b + N}.$$

We then simulate N topologies from the estimated Erdős-Rényi model to estimate its cluster distribution. In the column “Obs.,” we show the results from these analyses based on 1,000 simulations. Observe that these results are very much in line with the encompassing network results.

Empirical example: longitudinal data on depression

We next consider an encompassing network analysis of longitudinal depression data from the Addiction Health Evaluation and Disease (AHEAD) management study (Saitz et al., 2013).⁸ The AHEAD study was a randomized controlled trial conducted in Boston between 2006-2010 evaluating chronic disease management for substance dependence in primary care. Participants in the study were at least 18 years of age, diagnosed with alcohol and/or drug dependence, and were currently seeking treatment. Participants were assessed on various measures at baseline and three, six, and twelve months after enrollment in the program. Here, we focus on the network analysis of the depression symptom indicators assessed in the study. The AHEAD study used the PHQ-9 questionnaire (Kroenke et al., 2001) to assess the nine DSM-IV symptoms for Major Depression (American Psychiatric Association, 2000), which we scored as -1 (symptom not present) and $+1$ (symptom present several days to nearly every day). The PHQ-9 questionnaire assesses the following nine depression symptoms: (1) Loss of interest; (2) depressed mood; (3) trouble sleeping or sleeping too much; (4) feeling tired; (5) poor appetite or overeating; (6) low self-esteem; (7) trouble concentrating; (8) more or less energy than usual; (9) self-harm or suicidal thoughts. In the following, we will use their numbers to refer to the symptoms in the plots.

One of the research questions in our analysis is which clusters of depression symptoms we can identify from the idiographic topologies and how these clusters compare against the cross-sectional network estimates. To investigate this research question, we analyzed the symptom level data of all participants that completed all four waves of assessments. In total,

⁷The autocorrelations were higher for the $k=1$ scenarios. For these scenarios, we therefore ran the Gibbs samplers for 1,250 iterations, including 250 burnin iterations.

⁸Data can be requested from <https://www.icpsr.umich.edu/web/ICPSR/studies/33581>.

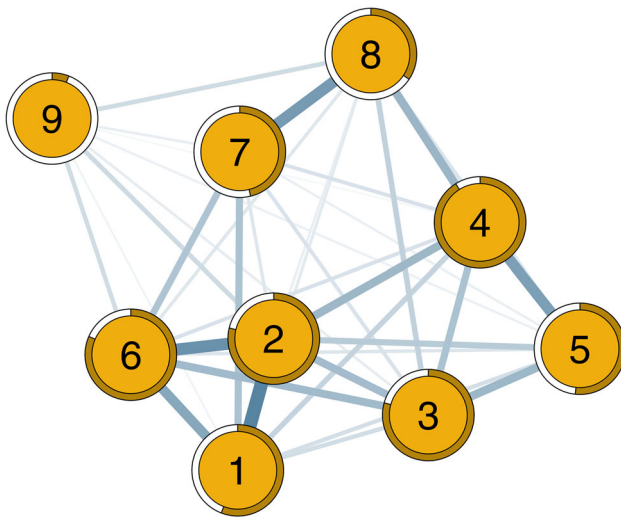


Figure 10. The estimated population network structure for the nine symptoms of depression. Edges reflect the posterior means of the Erdős-Rényi model's edge inclusion probabilities. Pie graphs around each node reflect the node's activation probability in the absence of network relations. The network was visualized using the *qgraph* R-package (Epskamp et al., 2012).

Table 2. Clustering of the idiographic networks of the encompassing network from all four waves split by cluster size. The top row provides the number of possible distinct clusters of a particular size, calculated as $\binom{9}{h}$, where h is the cluster size. The bottom three rows show the number of distinct clusters of a particular size observed in the three samples of plausible networks.

Cluster Size	1	2	3	4	5	6	7	8	9	Total
Possible clusters	9	36	84	126	126	84	36	9	1	511
Sample 1: Observed clusters	9	35	60	52	40	29	12	6	1	244
Sample 2: Observed clusters	9	33	55	43	41	32	12	5	1	231
Sample 3: Observed clusters	9	34	55	49	46	30	11	6	1	241

there were 563 participants in the AHEAD study, of which $N=439$ participants completed all four waves.

With random starting values for the model parameters and topologies, we ran the Gibbs sampler for 100,000 iterations. The proposed Metropolis algorithm for the thresholds again performed well in these analyses, with acceptance rates close to 100%. The first 5,000 sampled states of a threshold parameter and edge inclusion probability are shown in Figure S3 in the supplementary material. The model appeared to converge quickly. We have summarized the parameter values and 95% highest posterior density intervals in Table S1 in the supplementary material.

Figure 10 shows the estimated population network structure. The edges in the network displayed in Figure 10 reflect the posterior means of the Erdős-Rényi model's edge inclusion probabilities. Since we

have used no form of regularization, the network is densely connected. Node two—depressed mood—appears to be the most strongly connected symptom in the network. This finding is in line with previous results, in which depressed mood was found to be the most central and most connected symptom in the network (Langer et al., 2019; McElroy et al., 2018; Santos et al., 2017). Pie graphs around each node reflect the activation probability of that node in the absence of other network relations, i.e.,

$$p(X_i = +1) = \frac{\exp(\mu_i)}{\exp(\mu_i) + \exp(-\mu_i)}.$$

Table S1 in the supplementary material shows that most of the estimated thresholds were positive, which reflects that most persons experienced these symptoms. For these symptoms, $p(X_i = +1) > .5$ and the pie graph was more than half filled. The thresholds that corresponded to the symptoms *trouble concentrating* (node 7), *more or less energy than usual* (node 8), and *self-harm or suicidal thoughts* (node 9) were estimated to be negative, which reflects that most persons did not experience these symptoms. The corresponding pie graphs were less than half filled.

We next analyze the clustering of the idiographic networks that give rise to the cross-sectional networks in Figure 10. Recall that a cluster is a set of nodes that lie on an open path, and the number of nodes in a cluster reflects its size. We observed three interesting findings. First, most individuals showed highly fragmented topologies. On average, individual networks comprised 4.69 clusters. Less than half (i.e., 46.47%) of the participants had one large cluster of at least five nodes in addition to several smaller clusters. Second, in three samples of posterior networks, we found between 231 and 244 distinct clusters. Table 2 shows the number of observed clusters per cluster size for the three samples, together with the number of possible clusters of that size. As can be gleaned from the table, the plausible networks more often comprised smaller (i.e., less than five nodes) than larger clusters (i.e., more than five nodes). For example, for two-variable clusters, we observed between 33 and 35 distinct clusters out of the $\binom{9}{2} = 36$ possible ones. In contrast, we only observed between 11 and 12 distinct seven-variable clusters out of the same $\binom{9}{7} = 36$ possibilities. Third, the most observed clusters consisted of a single node. Here, *self-harm and suicidal thoughts* was most often in an isolated cluster (i.e., in 72.21% of the plausible networks) and *depressed mood* least often (i.e., in 17.08% of the plausible networks). These

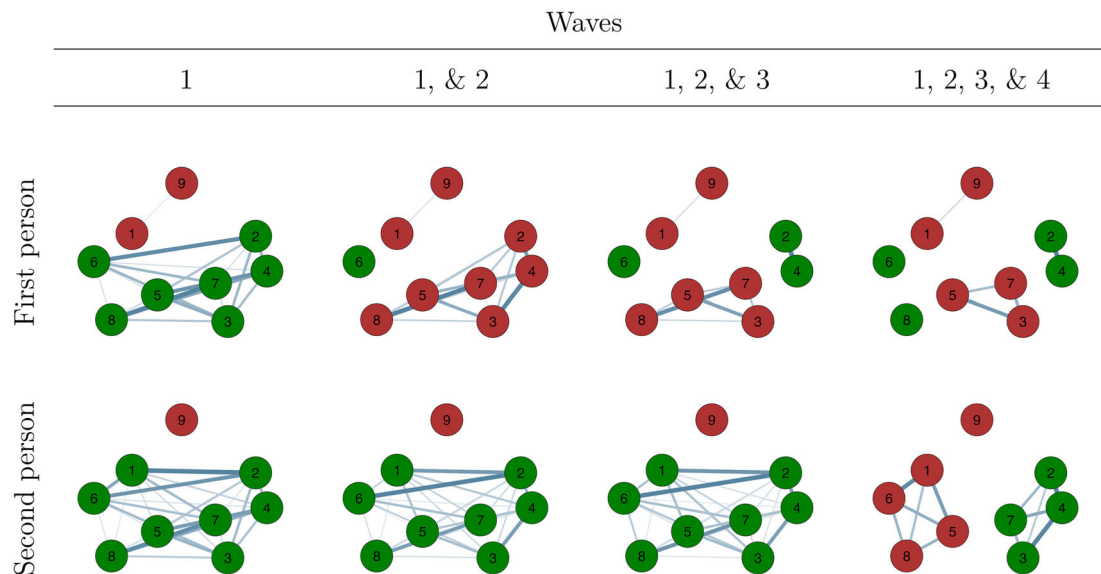


Figure 11. The idiographic networks of two participants across the four waves. Encompassing networks were estimated through consecutively adding data from each wave. After each wave, we ran a Gibbs sampler for 100,000 iterations. The network links depicted here reflect the proportion of links observed in the 100,000 iterations of the Gibbs for these individuals; the thicker the edge, the more often a link was observed. Node coloring represent the symptom states at the particular wave; *green* indicates a present symptom and *red* an absent symptom.

observations imply that *depressed mood* was connected to other nodes in almost 83% of plausible networks, while *self-harm and suicidal thoughts* in less than 18%. To conclude, the dense cross-sectional encompassing network is build up from highly fragmented idiographic networks, i.e., networks consisting of many small clusters.

Next, we zoom in on the idiographic networks of two individuals. Figure 11 shows the idiographic networks obtained in consecutive waves. We added the data from the waves consecutively to estimate the encompassing network and idiographic topologies. Nodes are colored according to the symptom activation at that particular wave (i.e., *green* indicates an active symptom, *red* an inactive symptom). We can identify several interesting findings. First, both individuals depict different trajectories despite similarly densely connected idiographic networks at wave one. While the idiographic topology of the first person quickly became more fragmented, the topology of the second person only showed a difference in the last wave. Second, as we include the data from more waves (i.e., increase k), we start to observe nodes with different values (i.e., $X_i \neq X_j$), and clusters separate. Third, the idiographic network of the first person is more fragmented than that of the second person. The first person's network consists of five clusters with a maximum size of three, while the other person's network consists of three clusters with a maximum size of four. Fourth, we can trace the configuration of the node activity states, as

indicated through the coloring of the nodes in all four plots. States of nodes within one cluster fluctuate jointly across all four measurement occasions. At the first measurement wave, most symptoms are active, while at the last state, inactive symptoms are predominant. The number of active depression symptoms decreased over time for both persons.

In 100,000 iterations of the Gibbs sampler, we observed 32 distinct plausible networks for the first person and 3,445 for the second. We are thus more confident about the first topology than about the second. We show the three most plausible networks (i.e., the most frequently observed ones) for each person in Figure 12. Observe that the plausible networks are even more sparse than the posterior averages in Figure 11. With high probability, the topology of person one contains six or more clusters. Observe that even though we observed many distinct networks for person two, the three most plausible ones had very similar characteristics, comprising only three distinct clusters, with the same three links in the one positive cluster. These topologies need not reflect the correct topology of these individuals, but since they are consistent with everything we know about them, they could be.

Discussion

In this paper, we have introduced an idiographic formulation of the Ising model based on the theory of Fortuin and Kasteleyn. A fundamental aspect of this

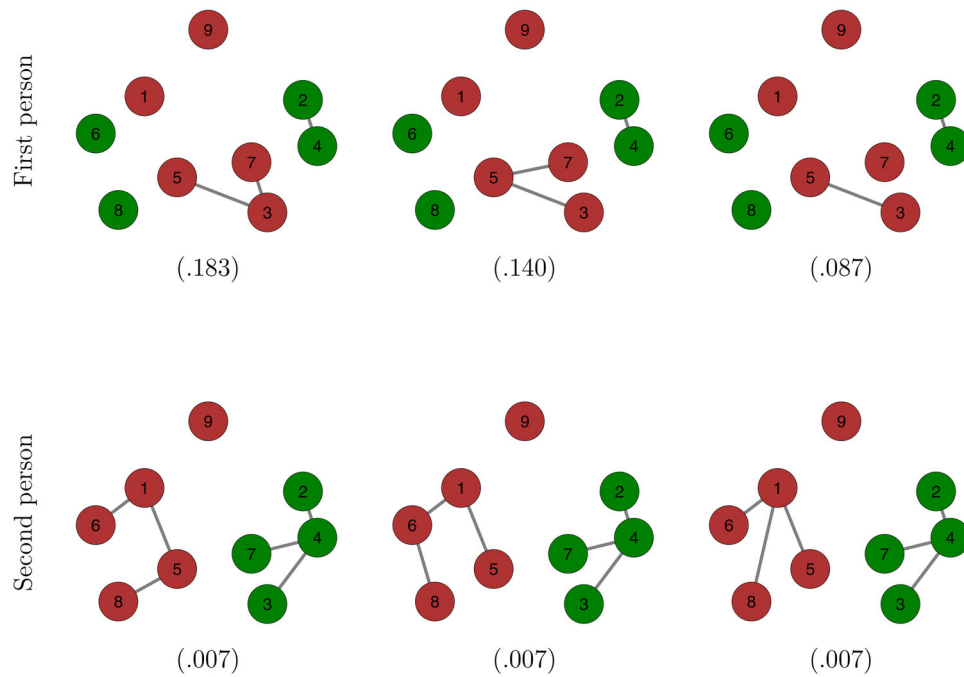


Figure 12. The three most plausible networks for the two persons. Numbers in brackets indicate the proportion of plausible networks with these topologies.

version of the Ising model is that it does not require a homogeneous population for the model to apply. We can thus obtain the Ising model cross-sectionally even if topological structures vary on an individual basis. That is, population homogeneity is a sufficient condition for the Ising model's fit, but not a necessary one. This result suggests a reconciliation of the idiographic and nomothetic approaches in network psychometrics.

Equation (8) describes a formal connection between the two approaches. It expresses observed correlations in the population as a function of the probability that variables link at an individual level. As far as we can tell, such a formal link between correlations at the individual and the group level has not been described before in the psychometric literature. An essential feature of this relation is that it is also valid if the topological structures do not form a random-cluster distribution. The results of Steif and Tykesson (2017), for example, show that the relationship is also valid if the topological structures follow an Erdős-Rényi distribution. The connection between observed correlations and idiographic edge inclusion probabilities thus forms an important bridge between the two levels. It explains how individual networks can generalize to established cross-sectional phenomena and how cross-sectional phenomena can be used to constrain idiographic topologies. In the particular case of the Ising model, where the topological structures follow a random-cluster distribution, Equation (9) provides a parametric link between group-level associations σ_{ij}

and idiographic edge inclusion probabilities θ_{ij} . But, given that the random-cluster model and Erdős-Rényi model are so closely related, one may also expect that the proportion of latent links w_{+ij}/N are closely related to σ , defined as $-\log(1 - \theta)/2$. Figure 13, which depicts these relations for both models, confirms our hunch.⁹

Formulating the Ising model in terms of individual differences in network structure offers a fresh perspective on the conceptualization of psychological constructs and related group-level phenomena. To solidify this approach, we have formulated a framework for encompassing network models, and showed how to estimate models from this new framework and uncover the latent topologies that cause the new perspective. In sum, the results in this paper open up many new lines of research in the field of network psychometrics that we believe will resonate throughout the psychological sciences.

An encompassing network theory

The focus of this article was the idiographic framing of Ising models using the theory of Fortuin and Kasteleyn. Their theory offers a novel network approach that encompasses both graphical models and

⁹Interestingly, this relation appears to be much more stable across conditions for the divide and color model than for Fortuin and Kasteleyn's idiographic Ising model.

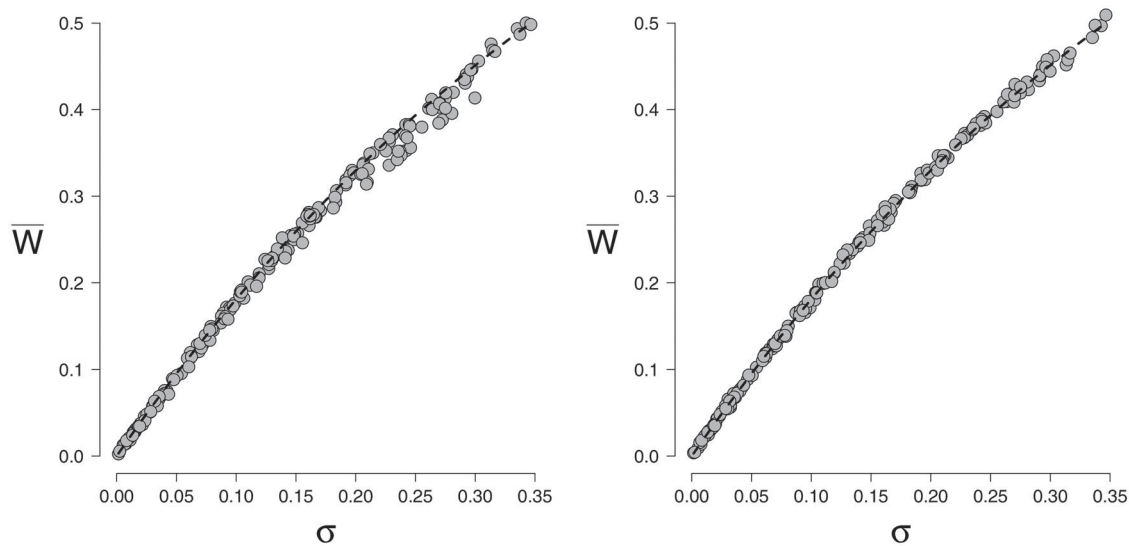


Figure 13. Scatterplots of the proportion of latent links w_{+ij}/N against σ_{ij} for Fortuin and Kasteleyn's idiographic Ising model in the left panel, and the divide and color model in the right panel. The associations were defined as $-\log(1 - \theta_{ij})/2$, where θ_{ij} is the generating edge inclusion probability. Results were based on $N = 10,000$ simulated topologies for a $n = 20$ variable network. The edge inclusion probabilities θ_{ij} were the same for both models and sampled uniformly between 0 and 0.5. The thresholds μ_i were simulated from a standard normal distribution.

random graph models. These two modeling approaches generate streams of literature that are, to a large extent, independent. The fields of Bayesian networks, neural networks, and, more recently, network psychometrics, for example, use graphical models to unveil the relations between observables. On the other hand, the fields of social networks, and, more recently, network science mostly use random graph models to unveil the characteristics of particular entities—nodes in the network—from their relations. Here, the entities could be persons, rail stations, or airports, for example, and their relations could be friendships, and rail or flight connections, respectively.

Because their literature is so isolated from each other, it is often unclear how the two modeling frameworks relate. This obscurity inspired several critiques on the assumptions and methods used in network psychometrics (see, for instance, Brusco et al. 2022; Epskamp et al., 2022; Marsman et al., 2022; Steinley et al., 2017, for recent discussions). The centrality measures used to analyze the relative importance of nodes in the network provide a good example. Initially proposed for (unweighted) random graph models, their extension to weighted graphs offered an opportunity to use them for graphical models such as the Ising model. However, several publications now criticize the application of centrality measures to graphical models, since it is unclear how to interpret them. The connection between random graph models and graphical models examined in this paper offers a way to analyze how centrality measures for

(unweighted) random graphs relate to centrality measures for (weighted) graphical models. And, related to the topic of this paper, how centrality measures for idiographic networks relate to the centrality measures of cross-sectional networks. We believe that the unification of the two modeling approaches will improve our understanding of network models and their methods in psychometrics. Centrality measures are an example of this.

We also think, however, that the relationship between the two types of network models aids our conceptual understanding of psychological constructs, which, according to this paper, resides in the origin of individual topologies. These topologies comprise the unique makeup of psychological disorders, attitudes, personality, or abilities of persons, and it is crucial to establish any regularities in their structure. Recent advances in network science suggest that several topological patterns are common to networks throughout the statistical sciences, such as community structure (Karrer & Newman, 2011; M. Newman, 2012), small worlds (Watts & Strogatz, 1998), and scale-free networks (Barabási & Albert, 1999). It would be interesting to see if these patterns also occur in psychometric topologies, or that they are, in this respect, unique. We have seen, for example, that different random graph models lead to different levels of fragmentation in the idiographic topologies. Furthermore, we also know that the ubiquitous bi-factor structure in psychological measurement can emerge from a community structure in individual topologies (Savi et al.,

2019). But what we do not know, however, is what other topological patterns coincide with the patterns gleaned from observed correlations. Do the same topological patterns that are observed in other areas of network science also occur in these psychometric networks? Are there unique topological patterns that occur in psychometric topologies? Are there plausible psychological processes that generate observed topological patterns? In sum, the unified network theory of Fortuin and Kasteleyn raises intriguing questions about psychological constructs.

Encompassing network modeling

Fortuin and Kasteleyn's theory also inspired us to formulate a class of encompassing network models for populations of idiographic networks. We introduced statistical methods for their analysis, focusing on encompassing networks using an Erdős-Rényi model for the latent topologies. This model proved to be much simpler to estimate than Fortuin and Kasteleyn's original approach using a random-cluster model, since the Erdős-Rényi model does not have an intractable normalizing constant. While being much easier to analyze, our simulations also revealed that an encompassing network using an Erdős-Rényi model for the latent topologies is high indistinguishable from Fortuin and Kasteleyn's original random-cluster formulation. Our version of the encompassing network model thus enjoys great theoretical and practical appeal.

Our formulation of the Erdős-Rényi model—a product of independent Bernoulli distributions—covers the stochastic block model as a special case (Holland et al., 1983). The stochastic block model is a popular variant of the Erdős-Rényi model that groups nodes into blocks or communities (i.e., it differentiates the inclusion probabilities of links between nodes that are part of the same community from the inclusion probabilities of links between nodes from different communities). Savi et al. (2019) revealed that there is a close connection to the bi-factor structure, ubiquitous in psychological data, which makes a stochastic block variant of the encompassing network an interesting topic for future research. It should be relatively straightforward to combine existing Gibbs sampling algorithms for the stochastic block model with the proposed Gibbs sampling algorithm for the encompassing network. The same holds for the degree-corrected stochastic block models proposed by Karrer and Newman (2011).

Another advantage of the Erdős-Rényi model is that it offers great modeling flexibility. Consider, for

example, the following logistic model for the edge inclusion probabilities between variables i and j in the topology of a person p :

$$P(W_{p_{ij}} = 1 | \mathbf{z}_p) = \theta_{p_{ij}} = \frac{\exp(\alpha + \beta^T \mathbf{z}_p)}{1 + \exp(\alpha + \beta^T \mathbf{z}_p)},$$

which extends the Erdős-Rényi model to include person-specific covariates to predict the links in the idiographic topologies. These covariates could be background characteristics, experimental factors, or grouping variables (e.g., school indicators). Even though we would lose conjugacy of the Beta prior on the Bernoulli edge inclusion probabilities, the augmentation scheme of Polson et al. (2013) offers a practical solution in which the normal density is the conjugate prior for the α and β parameters of the logistic model. It is thus relatively straightforward to update our Gibbs sampling algorithm for estimating these regression structures.

The proposed regression set-up also points to an interesting dynamic extension of the model, since $\mathbf{z}_p$ could also be the edge state at a previous time-point. Let $W_{p_{ij}}^{(t)}$ denote the latent link between variables i and j for a person p at a time t . Then, the following logistic model

$$P(W_{p_{ij}}^{(t)} = 1 | W_{p_{ij}}^{(t-1)}) = \theta_{p_{ij}} = \frac{\exp(\alpha + \beta W_{p_{ij}}^{(t-1)})}{1 + \exp(\alpha + \beta W_{p_{ij}}^{(t-1)})},$$

which models the network's dynamics through autoregression of the latent topologies. This points toward hidden-Markov versions of the encompassing network to relax the assumption that the idiographic topologies are constant through time.

Generalizing theory and models beyond binary variables

The Ising model is fundamental to network psychometrics, since many psychological or psychometric variables are binary, such as symptom presence, statement endorsement, and item responses, for example. For this model, we have now formulated a framework that comprises both idiographic and nomothetic approaches. However, the Ising model does not cover all relevant variables in psychological measurement. A comprehensive psychometric theory that connects idiographic and nomothetic approaches would accommodate other variable types. Unfortunately, a broad connection between graphical and random graph models is unavailable, since Fortuin and Kasteleyn's theory only covers Ising and Potts models (Ashkin &

Teller, 1943; Potts, 1952).¹⁰ Generalizing the encompassing network approach to other variable types is an interesting challenge. Cai (2017) recently developed a theory for Gaussian graphical models closely related to the models analyzed in this paper. Spanning the space between the encompassing network approaches of Fortuin and Kasteleyn's and that of Cai offers a fruitful avenue for future research.

Article information

Conflict of Interest Disclosures: Each author signed a form for disclosure of potential conflicts of interest. No authors reported any financial or other conflicts of interest in relation to the work described.

Ethical Principles: The authors affirm having followed professional ethical guidelines in preparing this work. These guidelines include obtaining informed consent from human participants, maintaining ethical treatment and respect for the rights of human or animal participants, and ensuring the privacy of participants and their data, such as ensuring that individual participants cannot be identified in reported results or from publicly available original or archival data.

Funding: This work was supported by Grant 451-17-017 from the Netherlands Organization for Scientific Research.

Role of the Funders/Sponsors: None of the funders or sponsors of this research had any role in the design and conduct of the study; collection, management, analysis, and interpretation of data; preparation, review, or approval of the manuscript; or decision to submit the manuscript for publication.

References

- American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed.). American Psychiatric Association.
- Ashkin, J., & Teller, E. (1943). Statistics of two-dimensional lattices with four components. *Physical Review*, 64(5–6), 178–184. <https://doi.org/10.1103/PhysRev.64.178>
- Bak, M., Drukker, M., Hasmi, L., & van Os, J. (2016). An $n=1$ clinical network analysis of symptoms and treatment in psychosis. *PLoS One*, 11(e0162811), 1–15.
- Barabási, A.-L., & Albert, R. (1999). Emergence of scaling in random networks. *Science*, 286(5439), 509–512.
- Bodner, N., Bringmann, L., Tuerlinckx, F., de Jonge, P., & Ceulemans, E. (2022). ConNECT: A novel network approach for investigating the co-occurrence of binary psychopathological symptoms over time. *Psychometrika*, 87(1), 107–132. <https://doi.org/10.1007/s11336-021-09765-2>
- Borsboom, D. (2008). Psychometric perspectives on diagnostic systems. *Journal of Clinical Psychology*, 64(9), 1089–1108. <https://doi.org/10.1002/jclp.20503>
- Borsboom, D., & Cramer, A. O. J. (2013). Network analysis: An integrative approach to the structure of psychopathology. *Annual Review of Clinical Psychology*, 9(1), 91–121. <https://doi.org/10.1146/annurev-clinpsy-050212-185608>
- Bos, E., & Wanders, R. (2016). Group-level symptom networks in depression. *JAMA Psychiatry*, 73(4), 411. <https://doi.org/10.1001/jamapsychiatry.2015.3103>
- Bos, F. M., Snippe, E., Bruggeman, R., Wichers, M., & van der Krieke, L. (2022). Insights of patients and clinicians on the promise of the experience sampling method for psychiatric care. *Psychiatric Services*, 70(1), 983–991. <https://doi.org/10.1176/appi.ps.201900050>
- Bos, F. M., Snippe, E., de Vos, S., Hartmann, J. A., Simons, C. J. P., van der Krieke, L., de Jonge, P., & Wichers, M. (2017). Can we jump from cross-sectional to dynamic interpretations of networks? Implications for the network perspective in psychiatry. *Psychotherapy and Psychosomatics*, 86(3), 175–177. <https://doi.org/10.1159/000453583>
- Bringmann, L. F., & Eronen, M. I. (2018). Don't blame the model: Reconsidering the network approach to psychopathology. *Psychological Review*, 125(4), 606–615. <https://doi.org/10.1037/rev0000108>
- Broadbent, S., & Hammersley, J. (1957). Percolation processes I. Crystals and mazes. *Mathematical Proceedings of the Cambridge Philosophical Society*, 53(3), 629–641. <https://doi.org/10.1017/S0305004100032680>
- Brusco, M., Steinley, D., Hoffman, M., Davis-Stober, C., & Wasserman, S. (2022). On Ising models and algorithms for the construction of symptom networks in psychopathological research. *Psychological Methods*, 24(6), 735–753. <https://doi.org/10.1037/met0000207>
- Brush, S. G. (1967). History of the Lenz-Ising model. *Reviews of Modern Physics*, 39(4), 883–893. <https://doi.org/10.1103/RevModPhys.39.883>
- Cai, H. (2017). A note on jointly modeling edges and node attributes of a network. *Statistics and Probability Letters*, 121, 54–60. <https://doi.org/10.1016/j.spl.2016.10.014>
- Caspi, A., Houts, R. M., Belsky, D. W., Goldman-Mellor, S. J., Harrington, H., Israel, S., Meier, M. H., Ramrakha, S., Shalev, I., Poulton, R., & Moffitt, T. E. (2014). The p factor: One general psychopathology factor in the structure of psychiatric disorders? *Clinical Psychological Science: A Journal of the Association for Psychological Science*, 2(2), 119–137. <https://doi.org/10.1177/2167702613497473>
- Cioletti, L., & Vila, R. (2016). Graphical representations for Ising and Potts models in general external fields. *Journal of Statistical Physics*, 162(1), 81–122. <https://doi.org/10.1007/s10955-015-1396-5>

¹⁰The Potts model is a generalization of the Ising model that did not attract the interest of psychologists.

- Costantini, G., Epskamp, S., Borsboom, D., Perugini, M., Möttus, R., Waldorp, L. J., & Cramer, A. O. (2015). State of the art personality research: A tutorial on network analysis of personality data in R. *Journal of Research in Personality*, 54, 13–29. <https://doi.org/10.1016/j.jrp.2014.07.003>
- Costantini, G., Richetin, J., Preti, E., Casini, E., Epskamp, S., & Perugini, M. (2019). Stability and variability of personality networks: A tutorial on recent developments in network psychometrics. *Personality and Individual Differences*, 136, 68–78. <https://doi.org/10.1016/j.paid.2017.06.011>
- Cox, D. (1972). The analysis of multivariate binary data. *Journal of the Royal Statistical Society. Series B (Applied Statistics)*, 21(2), 113–120. <https://doi.org/10.2307/2346482>
- Cramer, A. O. J., van Borkulo, C. D., Giltay, E. J., van der Maas, H. L. J., Kendler, K. S., Scheffer, M., & Borsboom, D. (2016). Major depression as a complex dynamic system. *Plos One*, 11(12), e0167490–20. <https://doi.org/10.1371/journal.pone.0167490>
- Cramer, A. O. J., van der Sluis, S., Noordhof, A., Wichers, M., Geschwind, N., Aggen, S. H., Kendler, K. S., & Borsboom, D. (2012). Dimensions of normal personality as networks in search of equilibrium: You can't like parties if you don't like people. *European Journal of Personality*, 26(4), 414–431. <https://doi.org/10.1002/per.1866>
- Cramer, A. O. J., Waldorp, L. J., van der Maas, H. L. J., & Borsboom, D. (2010). Comorbidity: A network perspective. *The Behavioral and Brain Sciences*, 33(2–3), 137–193. <https://doi.org/10.1017/S0140525X09991567>
- Csardi, G., Nepusz, T. (2019). igraph: Network analysis and visualization [Computer software manual]. Retrieved from <https://CRAN.R-project.org/package=igraph> (R-package version 1.2.4.2.)
- Dalege, J., Borsboom, D., van Harreveld, F., van den Berg, H., Conner, M., & van der Maas, H. L. J. (2016). Towards a formalized account of attitudes: The causal attitude network (CAN) model. *Psychological Review*, 123(1), 2–22. <https://doi.org/10.1037/a0039802>
- Dalege, J., Borsboom, D., van Harreveld, F., & van der Maas, H. L. J. (2017). Network analysis on attitudes: A brief tutorial. *Social Psychological and Personality Science*, 8(5), 528–537. <https://doi.org/10.1177/1948550617709827>
- Dalege, J., Borsboom, D., van Harreveld, F., & van der Maas, H. L. J. (2019). A network perspective on political attitudes: Testing the connectivity hypothesis. *Social Psychological and Personality Science*, 10(6), 746–756. <https://doi.org/10.1177/1948550618781062>
- De Vos, S., Wardenaar, K. J., Bos, E. H., Wit, E. C., Bouwmans, M. E. J., & de Jonge, P. (2017). An investigation of emotion dynamics in major depressive disorder patients and healthy persons using sparse longitudinal networks. *PloS One*, 12(6), e0178586–18.
- Epskamp, S., Cramer, A. O. J., Waldorp, L. J., Schmittmann, V. D., & Borsboom, D. (2012). qgraph: Network visualizations of relationships in psychometric data. *Journal of Statistical Software*, 48(4), 1–18. Retrieved from <http://www.jstatsoft.org/v48/i04/> <https://doi.org/10.18637/jss.v048.i04>
- Epskamp, S., Fried, E. I., van Borkulo, C. D., Robinaugh, D. J., Marsman, M., & Dalege, J. (2022) Investigating the utility of fixed-margin sampling in network psychometrics. *Multivariate Behavioral Research*, 56(2), 314–328. <https://doi.org/10.1080/00273171.2018.1489771>
- Epskamp, S., Maris, G., Waldorp, L., & Borsboom, D. (2018). Network psychometrics. In P. Irwing, D. Hughes, & T. Booth (Eds.), *Handbook of psychometrics* (pp. 953–986). Wiley-Blackwell.
- Erdős, P., & Rényi, A. (1960). On the evolution of random graphs. *Publications of the Mathematical Institute of the Hungarian Academy of Sciences*, 5(1), 17–60.
- Fisher, A. J. (2015). Toward a dynamic model of psychological assessment: Implications for personalized care. *Journal of Consulting and Clinical Psychology*, 83(4), 825–836. <https://doi.org/10.1037/ccp0000026>
- Fisher, A. J., Medaglia, J. D., & Jeronimus, B. F. (2018). Lack of group-to-individual generalizability is a threat to human subjects research. *Proceedings of the National Academy of Sciences*, 115(27), E6106–E6115. <https://doi.org/10.1073/pnas.1711978115>
- Fisher, A. J., Reeves, J. W., Glenn, L., Medaglia, J. D., & Rubel, J. A. (2017). Exploring the idiographic dynamics of mood and anxiety via network analysis. *Journal of Abnormal Psychology*, 126(8), 1044–1056.
- Forbes, M. K., Wright, A. G. C., Markon, K. E., & Krueger, R. F. (2017). Evidence that psychopathology symptom networks have limited replicability. *Journal of Abnormal Psychology*, 126(7), 969–988. <https://doi.org/10.1037/abn0000276>
- Forbes, M. K., Wright, A. G. C., Markon, K. E., & Krueger, R. F. (2019). The network approach to psychopathology: Promise versus reality. *World Psychiatry*, 18(3), 272–273. <https://doi.org/10.1002/wps.20659>
- Forbes, M. K., Wright, A. G. C., Markon, K. E., & Krueger, R. F. (2022). Quantifying the reliability and replicability of psychopathology network characteristics. *Multivariate Behavioral Research*, 56(2), 224–242. <https://doi.org/10.1080/00273171.2019.1616526>
- Fortuin, C. (1972a). On the random-cluster model: III. The simple random-cluster model. *Physica*, 59(4), 545–570. [https://doi.org/10.1016/0031-8914\(72\)90087-0](https://doi.org/10.1016/0031-8914(72)90087-0)
- Fortuin, C. (1972b). On the random-cluster model II. The percolation model. *Physica*, 58(3), 393–418. [https://doi.org/10.1016/0031-8914\(72\)90161-9](https://doi.org/10.1016/0031-8914(72)90161-9)
- Fortuin, C., & Kasteleyn, P. (1972). On the random-cluster model: I. Introduction and relation to other models. *Physica*, 57(4), 536–564. [https://doi.org/10.1016/0031-8914\(72\)90045-6](https://doi.org/10.1016/0031-8914(72)90045-6)
- Fortuin, C., Kasteleyn, P., & Ginibre, J. (1971). Correlation inequalities on some partially ordered sets. *Communications in Mathematical Physics*, 22(2), 89–103. <https://doi.org/10.1007/BF01651330>
- Fried, E. I., & Cramer, A. O. J. (2017). Moving forward: Challenges and directions for psychopathological network theory and methodology. *Perspectives on Psychological Science: A Journal of the Association for Psychological Science*, 12(6), 999–1020.
- Geman, S., & Geman, D. (1984). Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 6(6), 721–741.
- Grimmett, G. (2006). *The random-cluster model*. Springer-Verlag.

- Grimmett, G. (2018). *Probability on graphs. Random processes on graphs and lattices* (2nd ed.). Cambridge University Press.
- Häggström, O. (2001). Coloring percolation clusters at random. *Stochastic Processes and Their Applications*, 96(2), 213–242. [https://doi.org/10.1016/S0304-4149\(01\)00115-6](https://doi.org/10.1016/S0304-4149(01)00115-6)
- Häggström, O. (2002). *Finite Markov chains and algorithmic applications*. Cambridge University Press.
- Häggström, O., & Jonasson, J. (1999). Phase transition in the random triangle model. *Journal of Applied Probability*, 36(4), 1101–1115. Retrieved from <https://www.jstor.org/stable/3215581> <https://doi.org/10.1239/jap/1032374758>
- Haslbeck, J., Epskamp, E., Marsman, M., & Waldorp, L. J. (2022). Interpreting the Ising model: The input matters. *Multivariate Behavioral Research*, 56(2), 303–313. <https://doi.org/10.1080/00273171.2020.1730150>
- Holland, P. W., Laskey, K. B., & Leinhardt, S. (1983). Stochastic blockmodels: First steps. *Social Networks*, 5(2), 109–137. [https://doi.org/10.1016/0378-8733\(83\)90021-7](https://doi.org/10.1016/0378-8733(83)90021-7)
- Holland, P. W., & Rosenbaum, P. R. (1986). Conditional association and unidimensionality in monotone latent variable models. *The Annals of Statistics*, 14(4), 1523–1543. Retrieved from <https://www.jstor.org/stable/2241486> <https://doi.org/10.1214/aos/1176350174>
- Ising, E. (1925). Beitrag zur theorie des ferromagnetismus. *Zeitschrift Für Physik*, 31(1), 253–258. <https://doi.org/10.1007/BF02980577>
- Jonasson, J. (1999). The random triangle model. *Journal of Applied Probability*, 36(3), 852–867. Retrieved from <https://www.jstor.org/stable/3215446> <https://doi.org/10.1239/jap/1032374639>
- Karrer, B., & Newman, M. (2011). Stochastic blockmodels and community structure in networks. *Physical Review E*, 83(1), 016107. <https://doi.org/10.1103/PhysRevE.83.016107>
- Kievit, R. A., Frankenhuys, W. E., Waldorp, L. J., & Borsboom, D. (2013). Simpson's paradox in psychological science: A practical guide. *Frontiers in Psychology*, 4(513), 513–514.
- Kroenke, K., Spitzer, R. L., & Williams, J. B. (2001). The PHQ-9: Validity of a brief depression severity measure. *Journal of General Internal Medicine*, 16(9), 606–613.
- Langer, J. K., Tonge, N. A., Piccirillo, M., Rodebaugh, T. L., Thompson, R. J., & Gotlib, I. H. (2019). Symptoms of social anxiety disorder and major depressive disorder: A network perspective. *Journal of Affective Disorders*, 243, 531–538. <https://doi.org/10.1016/j.jad.2018.09.078>
- Lenz, W. (1920). Beiträge zum verständnis der magnetischen eigenschaften in festen körpern. *Physikalische Zeitschrift*, 21, 613–615.
- Maris, G., Bechger, T., & San Martin, E. (2015). A Gibbs sampler for the (extended) marginal Rasch model. *Psychometrika*, 80(4), 859–879.
- Marsman, M., Borsboom, D., Kruis, J., Epskamp, S., van Bork, R., Waldorp, L. J., van der Maas, H. L. J., & Maris, G. (2018). An introduction to network psychometrics: Relating Ising network models to item response theory models. *Multivariate Behavioral Research*, 53(1), 15–35. <https://doi.org/10.1080/00273171.2017.1379379>
- Marsman, M., Maris, G. K. J., Bechger, T. M., & Glas, C. A. W. (2015). Bayesian inference for low-rank Ising networks. *Scientific Reports*, 5, 9050. <https://doi.org/10.1038/srep09050>
- Marsman, M., Maris, G. K. J., Bechger, T. M., & Glas, C. A. W. (2016). What can we learn from plausible values? *Psychometrika*, 81(2), 274–289.
- Marsman, M., Tanis, C. C., Bechger, T. M., & Waldorp, L. J. (2019). Network psychometrics in educational practice. Maximum likelihood estimation of the Curie-Weiss model. In B. P. Veldkamp & C. Sluijter (Eds.), *Theoretical and practical advances in computer-based educational measurement* (pp. 93–120). Springer.
- Marsman, M., Waldorp, L. J., & Borsboom, D. (2022). Towards an encompassing theory of network models: Reply to Brusco, Steinley, Hoffman, Davis-Stober, & Wasserman. *Psychological Methods*, (PsyArXiv: n98qt). <https://doi.org/10.1037/met0000373>
- McElroy, E., Fearon, P., Belsky, J., Fonagy, P., & Patalay, P. (2018). Networks of depression and anxiety symptoms across development. *Journal of the American Academy of Child and Adolescent Psychiatry*, 57(12), 964–973. <https://doi.org/10.1016/j.jaac.2018.05.027>
- Mislevy, R. J. (1991). Randomization-based inference about latent variables from complex samples. *Psychometrika*, 56(2), 177–196. <https://doi.org/10.1007/BF02294457>
- Newman, C. (1991). Disordered Ising systems and random cluster representations. In G. Grimmett (Ed.), *Probability and phase transition* (Vol. 420, pp. 247–260). Springer.
- Newman, M. (2012). Communities, modules and large-scale structure in networks. *Nature Physics*, 8(1), 25–31. <https://doi.org/10.1038/nphys2162>
- Niss, M. (2005). History of the Lenz-Ising model 1920–1950: From ferromagnetic to cooperative phenomena. *Archive for History of Exact Sciences*, 59(3), 267–318. <https://doi.org/10.1007/s00407-004-0088-3>
- Polson, N. G., Scott, J. G., & Windle, J. (2013). Bayesian inference for logistic models using Pólya-Gamma latent variables. *Journal of the American Statistical Association*, 108(504), 1339–1349. <https://doi.org/10.1080/01621459.2013.829001>
- Potts, R. B. (1952). Some generalized order-disorder transformations. *Mathematical Proceedings of the Cambridge Philosophical Society*, 48(1), 106–109. <https://doi.org/10.1017/S0305004100027419>
- R Core Team. (2019). R: A language and environment for statistical computing [Computer software manual]. Vienna, Austria. Retrieved from <https://www.R-project.org/>
- Robinson, W. S. (1950). Ecological correlations and the behavior of individuals. *American Sociological Review*, 15(3), 351–357. <https://doi.org/10.2307/2087176>
- Saitz, R., Cheng, D. M., Winter, M., Kim, T. W., Meli, S. M., Allensworth-Davies, D., Lloyd-Travaglini, C. A., & Samet, J. H. (2013). Chronic care management for dependence on alcohol and other drugs: The AHEAD randomized trial. *JAMA*, 310(11), 1156–1167. <https://doi.org/10.1001/jama.2013.277609>
- Santos, H., Fried, E. I., Asafu-Adjei, J., & Ruiz, R. J. (2017). Network structure of perinatal depressive symptoms in Latinas: Relationship to stress and reproductive biomarkers. *Research in Nursing & Health*, 40(3), 218–228.
- Savi, A. O., Marsman, M., van der Maas, H. L. J., & Maris, G. K. J. (2019). The wiring of intelligence. *Perspectives on Psychological Science*, 16(6), 1034–1061.

- Simpson, E. H. (1951). The interpretation of interaction in contingency tables. *Journal of the Royal Statistical Society. Series B (Methodological)*, 13(2), 238–241. <https://doi.org/10.1111/j.2517-6161.1951.tb00088.x>
- Spearman, C. (1904). “General intelligence,” objectively determined and measured. *The American Journal of Psychology*, 15(2), 201–292. <https://doi.org/10.2307/1412107>
- Steif, J. E., Tykesson, J. (2017). Generalized divide and color models. Retrieved from <https://arxiv.org/abs/1702.04296>. (ArXiv preprint.)
- Steinley, D., Hoffman, M., Brusco, M. J., & Sher, K. J. (2017). A method for making inferences in network analysis: Comment on Forbes, Wright, Markon, and Krueger (2017). *Journal of Abnormal Psychology*, 126(7), 1000–1010. <https://doi.org/10.1037/abn0000308>
- Swendsen, R., & Wang, J. (1987). Nonuniversal critical dynamics in monte carlo simulations. *Physical Review Letters*, 58(2), 86–88.
- van Borkulo, C. D., Borsboom, D., Epskamp, S., Blanken, T. F., Boschloo, L., Schoevers, R. A., & Waldorp, L. J. (2014). A new method for constructing networks from binary data. *Scientific Reports*, 4, 5918. <https://doi.org/10.1038/srep05918>
- van Borkulo, C., Wichers, M., Boschloo, L., Epskamp, S., Schoevers, R., & Kamphuis, J. (2017). *The contact process as a model for predicting network dynamics of psychopathology*.
- van der Maas, H. L. J., Dolan, C. V., Grasman, R. P. P. P., Wicherts, J. M., Huizenga, H. M., & Raijmakers, M. E. J. (2006). A dynamical model of general intelligence: The positive manifold of intelligence by mutualism. *Psychological Review*, 113(4), 842–861.
- van der Maas, H. L. J., Kan, K.-J., Marsman, M., & Stevenson, C. E. (2017). Network models for cognitive development and intelligence. *Journal of Intelligence*, 5(2), 16–17. <https://doi.org/10.3390/jintelligence5020016>
- Watts, D. J., & Strogatz, S. H. (1998). Collective dynamics of ‘small-world’ networks. *Nature*, 393(6684), 440–442.

Appendix A

The Fortuin-Kasteleyn model

The joint distribution of the edge and node variables in the representation of Fortuin and Kasteleyn can be formulated as follows¹¹

$$\begin{aligned} p(\mathbf{w}, \mathbf{x}) &= p(\mathbf{x}|\mathbf{w}) p(\mathbf{w}) \\ &= \prod_{c=1}^{\kappa(\mathbf{w})} \frac{\exp(\sum_{i \in V_c} \mu_i x_i)}{\lambda_c} \frac{1}{Z_R} \prod_{i,j} \theta_{ij}^{w_{ij}} (1 - \theta_{ij})^{1-w_{ij}} \prod_{c=1}^{\kappa(\mathbf{w})} \lambda_c \\ &= \exp\left(\sum_{i=1}^n \mu_i x_i\right) \frac{1}{Z_R} \prod_{i,j} \theta_{ij}^{w_{ij}} (1 - \theta_{ij})^{1-w_{ij}} \mathbf{1}_R(\mathbf{w}, \mathbf{x}), \end{aligned}$$

where $\mathbf{1}_R(\mathbf{w}, \mathbf{x})$ is an indicator function for the set (see, for instance Grimmet, 2006)

$$\{\mathbf{w}, \mathbf{x}: w_{ij} = 1 \Rightarrow x_i = x_j\},$$

which ignores realizations $(\mathbf{w}, \mathbf{x})$ with connections between observables that are in different states. The restriction is the result of the way that the coloring process operates. An other way to formulate this restriction is as follows (Grimmet, 2006)

$$\begin{aligned} &\prod_{i,j} \theta_{ij}^{w_{ij}} (1 - \theta_{ij})^{1-w_{ij}} \mathbf{1}_R(\mathbf{w}, \mathbf{x}) \\ &= \prod_{i,j} \{\theta_{ij} \delta_{(w_{ij}, 1)} \delta_{(x_i, x_j)} + (1 - \theta_{ij}) \delta_{(w_{ij}, 0)}\} \\ &= \prod_{i,j} \left[\theta_{ij}^{w_{ij}} (1 - \theta_{ij})^{1-w_{ij}} \right]^{\delta_{(x_i, x_j)}} \left[0^{w_{ij}} (1 - \theta_{ij})^{1-w_{ij}} \right]^{1-\delta_{(x_i, x_j)}}, \end{aligned}$$

where $\delta_{(a, b)}$ denotes an indicator function that is equal to one if $a = b$ and equal to zero otherwise, and the zero term in the last factor expresses the fact that w_{ij} cannot equal one if $\delta_{(x_i, x_j)} = 0$.

The marginal distribution $p(\mathbf{x})$

The marginal distribution of the node states $\mathbf{x}$ can be obtained from the joint distribution. To wit,

$$\begin{aligned} p(\mathbf{x}) &= \sum_{\mathbf{w}} \frac{\exp\left(\sum_{i=1}^n \mu_i x_i\right)}{Z_R} \prod_{i,j} \left[\theta_{ij}^{w_{ij}} (1 - \theta_{ij})^{1-w_{ij}} \right]^{\delta_{(x_i, x_j)}} \\ &\quad \left[0^{w_{ij}} (1 - \theta_{ij})^{1-w_{ij}} \right]^{1-\delta_{(x_i, x_j)}} = \frac{\exp\left(\sum_{i=1}^n \mu_i x_i\right)}{Z_R} \\ &\quad \prod_{i,j} \left\{ \theta_{ij}^{\delta_{(x_i, x_j)}} 0^{1-\delta_{(x_i, x_j)}} + (1 - \theta_{ij})^{\delta_{(x_i, x_j)}} (1 - \theta_{ij})^{1-\delta_{(x_i, x_j)}} \right\} \\ &= \frac{\exp\left(\sum_{i=1}^n \mu_i x_i\right)}{Z_R} \prod_{i,j} \{\theta_{ij} \delta_{(x_i, x_j)} + 1 - \theta_{ij}\}. \end{aligned}$$

This is a reformulation of the Ising model, since if we express θ_{ij} as $1 - \exp(-2\sigma_{ij})$, we obtain

$$\begin{aligned} p(\mathbf{x}) &= \frac{\exp\left(\sum_{i=1}^n \mu_i x_i\right)}{Z_R} \prod_{i,j} \{(1 - \exp(-2\sigma_{ij})) \delta_{(x_i, x_j)} \\ &\quad + \exp(-2\sigma_{ij})\} = \frac{\exp\left(\sum_{i=1}^n \mu_i x_i\right)}{Z_R} \prod_{i,j} \exp(-2\sigma_{ij}(1 - \delta_{(x_i, x_j)})). \end{aligned}$$

The indicator function may alternatively be expressed as $\delta_{(x_i, x_j)} = \frac{1}{2} + \frac{1}{2} x_i x_j$, such that

$$\begin{aligned} p(\mathbf{x}) &= \frac{\exp\left(\sum_{i=1}^n \mu_i x_i\right)}{Z_R} \prod_{i,j} \exp(-\sigma_{ij} + \sigma_{ij} x_i x_j) \\ &= \frac{1}{\exp\left(\sum_{i,j} \sigma_{ij}\right) Z_R} \left\{ \exp\left(\sum_{i=1}^n \mu_i x_i + \sum_{i,j} \sigma_{ij} x_i x_j\right) \right\}, \end{aligned}$$

which gives the Ising model in Equation (2) upon realizing that the normalizing constant of the Ising model, Z_I , is equal to $Z_R \exp(\sum_{i,j} 2\sigma_{ij})$.

The idiographic posterior

Given the above formulation of the joint distribution of edge and node variables in the Fortuin-Kasteleyn model, we

¹¹We will use $\prod_{i,j}$ as shorthand for $\prod_{i=1}^{n-1} \prod_{j=i+1}^n$.

may express the posterior distribution of the edge variables given the states of the node variables as follows

$$p(\mathbf{w}|\mathbf{x}) = \frac{\exp\left(\sum_{i=1}^n \mu_i x_i\right) \prod_{i,j} [\theta_{ij}^{w_{ij}} (1 - \theta_{ij})^{1-w_{ij}}]^{\delta(x_i, x_j)} [0^{w_{ij}} (1 - \theta_{ij})^{1-w_{ij}}]^{1-\delta(x_i, x_j)}}{\sum_{\mathbf{w}} \frac{\exp\left(\sum_{i=1}^n \mu_i x_i\right) \prod_{i,j} [\theta_{ij}^{w_{ij}} (1 - \theta_{ij})^{1-w_{ij}}]^{\delta(x_i, x_j)} [0^{w_{ij}} (1 - \theta_{ij})^{1-w_{ij}}]^{1-\delta(x_i, x_j)}}{\prod_{i,j} [\theta_{ij}^{w_{ij}} (1 - \theta_{ij})^{1-w_{ij}}]^{\delta(x_i, x_j)} [0^{w_{ij}} (1 - \theta_{ij})^{1-w_{ij}}]^{1-\delta(x_i, x_j)}}}.$$

Next, we characterize the sum in the denominator.

$$p(\mathbf{w}|\mathbf{x}) = \frac{\prod_{i,j} [\theta_{ij}^{w_{ij}} (1 - \theta_{ij})^{1-w_{ij}}]^{\delta(x_i, x_j)} [0^{w_{ij}} (1 - \theta_{ij})^{1-w_{ij}}]^{1-\delta(x_i, x_j)}}{\prod_{i,j} \left\{ \theta_{ij}^{\delta(x_i, x_j)} 0^{1-\delta(x_i, x_j)} + (1 - \theta_{ij})^{\delta(x_i, x_j)} (1 - \theta_{ij})^{1-\delta(x_i, x_j)} \right\}} \\ = \frac{\prod_{i,j} [\theta_{ij}^{w_{ij}} (1 - \theta_{ij})^{1-w_{ij}}]^{\delta(x_i, x_j)} [0^{w_{ij}} (1 - \theta_{ij})^{1-w_{ij}}]^{1-\delta(x_i, x_j)}}{\prod_{i,j} \{ \theta_{ij} \delta(x_i, x_j) + (1 - \theta_{ij}) \}}.$$

Observe that the factors in the denominator are consistent with

$$\theta_{ij} \delta(x_i, x_j) + (1 - \theta_{ij}) = [1 - \theta_{ij}]^{1-\delta(x_i, x_j)},$$

such that the posterior can ultimately be expressed as

$$p(\mathbf{w}|\mathbf{x}) = \prod_{i,j} \left\{ \left[\theta_{ij}^{w_{ij}} (1 - \theta_{ij})^{1-w_{ij}} \right]^{\delta(x_i, x_j)} \left[\frac{0^{w_{ij}} (1 - \theta_{ij})^{1-w_{ij}}}{1 - \theta_{ij}} \right]^{1-\delta(x_i, x_j)} \right\} \\ = \prod_{i,j} \left[\theta_{ij}^{w_{ij}} (1 - \theta_{ij})^{1-w_{ij}} \right]^{\delta(x_i, x_j)} [0^{w_{ij}} 1^{1-w_{ij}}]^{1-\delta(x_i, x_j)}.$$

That is, given the response vector $\mathbf{x}$ the edges in the idiographic networks are independent Bernoulli variables that only depend on the adjacent response variables and the population structure.

Updating the posterior distribution

The posterior distribution after having observed k observations is equal to

$$p(\mathbf{w}|\mathbf{x}_1, \dots, \mathbf{x}_k) = \frac{\prod_{r=1}^k p(\mathbf{x}_r|\mathbf{w}) p(\mathbf{w})}{\sum_{\mathbf{w}} \prod_{r=1}^k p(\mathbf{x}_r|\mathbf{w}) p(\mathbf{w})} \\ = \frac{\prod_{r=1}^k \frac{\exp\left(\sum_{i=1}^n x_{ir} \mu_i\right)}{\prod_{c=1}^{\kappa(\mathbf{w})} \lambda_c} p(\mathbf{w}) \mathbf{1}_R(\mathbf{w}, \mathbf{x})}{\sum_{\mathbf{w}} \prod_{r=1}^k \frac{\exp\left(\sum_{i=1}^n x_{ir} \mu_i\right)}{\prod_{c=1}^{\kappa(\mathbf{w})} \lambda_c} p(\mathbf{w}) \mathbf{1}_R(\mathbf{w}, \mathbf{x})} \\ = \frac{\prod_{c=1}^{\kappa(\mathbf{w})} \lambda_c^{-k} p(\mathbf{w}) \mathbf{1}_R(\mathbf{w}, \mathbf{x})}{\sum_{\mathbf{w}} \prod_{c=1}^{\kappa(\mathbf{w})} \lambda_c^{-k} p(\mathbf{w}) \mathbf{1}_R(\mathbf{w}, \mathbf{x})}$$

where the indicator function $\mathbf{1}_R$ now covers all k observations on the node states. Thus, the numerator of the posterior distribution can be expressed as follows

$$\frac{1}{Z_R} \prod_{i,j} \left[\theta_{ij}^{w_{ij}} (1 - \theta_{ij})^{1-w_{ij}} \right]^{\delta_{ij}} \left[0^{w_{ij}} (1 - \theta_{ij})^{1-w_{ij}} \right]^{1-\delta_{ij}} \prod_{c=1}^{\kappa(\mathbf{w})} \lambda_c^{1-k},$$

where δ_{ij} is an indicator function that is equal to one whenever the states of nodes i and j were the same for all k observations, and is equal to zero otherwise. Thus, the posterior distribution is equal to

$$p(\mathbf{w}|\mathbf{x}_1, \dots, \mathbf{x}_k) =$$

$$\frac{\prod_{i,j} [\theta_{ij}^{w_{ij}} (1 - \theta_{ij})^{1-w_{ij}}]^{\delta_{ij}} [0^{w_{ij}} (1 - \theta_{ij})^{1-w_{ij}}]^{1-\delta_{ij}} \prod_{c=1}^{\kappa(\mathbf{w})} \lambda_c^{1-k}}{\sum_{\mathbf{w}} \prod_{i,j} [\theta_{ij}^{w_{ij}} (1 - \theta_{ij})^{1-w_{ij}}]^{\delta_{ij}} [0^{w_{ij}} (1 - \theta_{ij})^{1-w_{ij}}]^{1-\delta_{ij}} \prod_{c=1}^{\kappa(\mathbf{w})} \lambda_c^{1-k}},$$

which is the expression of a random-cluster model with edge probabilities

$$\theta'_{ij} = \theta_{ij} \times \delta_{ij},$$

and clustering weights

$$\lambda'_c = \lambda_c^{1-k}.$$

Appendix B

A Gibbs sampler for the random-cluster model

The Gibbs sampler (Geman & Geman, 1984) can be used to simulate values from the random-cluster model. To use the Gibbs sampler we set the network to some initial state $\mathbf{w}^{(0)}$ and then update each of the entries in $\mathbf{w}^{(0)}$ given the values on the remaining entries to obtain $\mathbf{w}^{(1)}$. This procedure is the repeated a “sufficient” number of times such that $\mathbf{w}^{(t)}$ at iteration t is a draw from the desired random-cluster model.

We update an entry (a, b) of $\mathbf{w}^{(t)}$ at an iteration t by simulating a value from a so-called full-conditional distributions, i.e., the distribution of W_{ab} given the values of the remaining variables at iteration t —denoted $\mathbf{w}_{\setminus(a,b)}^{(t)}$. The probability that $W_{ab} = 1$ given the values $\mathbf{w}_{\setminus(a,b)}^{(t)}$ on the remaining variables is equal to

$$\frac{\theta_{ab} \prod_{c=1}^{\kappa(\mathbf{w}_{\setminus(a,b)}^{(t)}, 1)} \lambda_c}{\theta_{ab} \prod_{c=1}^{\kappa(\mathbf{w}_{\setminus(a,b)}^{(t)}, 1)} \lambda_c + (1 - \theta_{ab}) \prod_{c=1}^{\kappa(\mathbf{w}_{\setminus(a,b)}^{(t)}, 0)} \lambda_c},$$

where $\kappa(\mathbf{w}_{\setminus(a,b)}^{(t)}, 1)$ denotes the number of clusters of $\mathbf{w}^{(t)}$ when $W_{ab} = 1$ and $\kappa(\mathbf{w}_{\setminus(a,b)}^{(t)}, 0)$ denotes the number of clusters of $\mathbf{w}^{(t)}$ when $W_{ab} = 0$. Thus the full-conditional probabilities depend on whether or not a and b are indirectly connected or not. If a and b are indirectly connected the number of clusters of $\mathbf{w}_{\setminus(a,b)}^{(t)}$ would be the same as the number of clusters of $\mathbf{w}^{(t)}$ and then the probability that $W_{ab} = 1$ is simply θ_{ab} . Let V_c denote the cluster to which both a and b belong, and let V_d and V_e denote the clusters that form when a and b are not directly connected, such that $a \in V_d$ and $b \in V_e$. Then, in the case that $W_{ab} = 0$ would split cluster c into two distinct clusters, d and e , the

probability that $W_{ab} = 1$ is equal to

$$\frac{1}{1 + \frac{(1-\theta_{ab}) \cosh\left(\sum_{i \in V_d} \mu_i\right) \cosh\left(\sum_{i \in V_e} \mu_i\right)}{\theta_{ab} \cosh\left(\sum_{i \in V_c} \mu_i\right)}}. \quad (B1)$$

The above formulation makes use of the clustering weights as in Equation (6). Observe that the full-conditional in Equation (B1) simplifies to $\frac{\theta_{ab}}{2-\theta_{ab}}$ for the standard random-cluster model where the external fields are all equal to zero. In the case that the target random-cluster model uses clustering weights that are equal to λ_c^{1-k} —i.e., after having updated the posterior distribution k times—the full-conditional in Equation (B1) becomes

$$\frac{1}{1 + \frac{1-\theta_{ab}}{\theta_{ab}} \left(\frac{2 \cosh\left(\sum_{i \in V_d} \mu_i\right) \cosh\left(\sum_{i \in V_e} \mu_i\right)}{\cosh\left(\sum_{i \in V_c} \mu_i\right)} \right)^{1-k}}.$$

Appendix C

A Gibbs sampling algorithm to estimate the encompassing network model

The Gibbs sampling algorithm comprises three steps. Each step consists of sampling from the posterior of one of the three parameters $\mathbf{w}$, $\boldsymbol{\mu}$, and $\boldsymbol{\theta}$ conditionally upon all other parameters. Next, we derive the three desired full-conditional posterior distributions and discuss how to simulate from them.

Step 1: Simulating $\mathbf{W}$

The individual topologies are independent given the observed data and the model parameters. The posterior distribution of the topology of a person p is proportional to

$$p(\mathbf{W}_p | \boldsymbol{\mu}, \boldsymbol{\theta}, \mathbf{X}_p) \propto \prod_{r=1}^k p(\mathbf{x}_{pr} | \mathbf{W}_p, \boldsymbol{\mu}) p(\mathbf{W}_p | \boldsymbol{\theta}).$$

Appendix A shows that this characterizes a random-cluster distribution. Let λ_c denote the clustering coefficient for cluster c , and let λ_c^i and λ_c^j denote the clustering coefficients for the clusters that result when $w_{ij} = 0$ splits cluster c into two distinct clusters. That is, $\lambda_c = [2 \cosh(\sum_{h \in V_c} \mu_h)]^{-k}$, $\lambda_c^i = [2 \cosh(\sum_{h \in V_c^i} \mu_h)]^{-k}$, and $\lambda_c^j = [2 \cosh(\sum_{h \in V_c^j} \mu_h)]^{-k}$, where V_c^i denotes the set of nodes in the splitted cluster to which node i belongs, and V_c^j the cluster of nodes to which node j belongs. Then, if $\delta_{pij} = 1$, Appendix B shows that we can sample the links in the idiographic networks using,

$$p(W_{ij} = 1 | \mathbf{w}_{p(ij)}^{\theta}, \boldsymbol{\theta}, \boldsymbol{\mu}, \mathbf{X}_p) = \begin{cases} \frac{\lambda_c \theta_{ij}}{\lambda_c \theta_{ij} + \lambda_c^i \lambda_c^j (1 - \theta_{ij})} & \text{if } w_{pij} = 0 \text{ split clusters } c. \\ \theta_{ij} & \text{if } w_{pij} = 0 \text{ does not split cluster } c, \end{cases}$$

and we set $W_{pij} = 0$ if $\delta_{pij} = 0$.

Step 2: Simulating $\boldsymbol{\mu}$

The likelihood for the threshold of node i , based on the configuration of one person p , can be formulated as

$$L(\mu_i | \mathbf{x}_p, \mathbf{w}_p) = \frac{\exp(2 \sum_{r=1}^k x_{pri} \mu_i) \exp(2 \sum_{r=1}^k \sum_{j \in V_c, j \neq i} x_{prj} \mu_j)}{\left(1 + \exp(2 \mu_i) \exp(2 \sum_{j \in V_c, j \neq i} \mu_j)\right)^k} \\ = \frac{(\beta_i q)^{y_{p+i}}}{(1 + \beta_i q)^k},$$

where $\beta_i = \exp(2 \mu_i)$, $q = \exp(2 \sum_{j \in V_c, j \neq i} \mu_j)$, and $y_{p+i} = \frac{1}{2} + \frac{1}{2} x_{pri}$. If we would only consider the data from a single person, we find that a beta-prime distribution is conjugate to β_i and we would obtain the full-conditional posterior distribution in closed form.¹² The beta-prime distribution is no longer conjugate if more than one topological structure is considered, since then node i can belong to one cluster for one topology and to another cluster for a different cluster. That is, $q = q_p$ depends on the person, and the full likelihood becomes,

$$L(\mu_i | \mathbf{Y}_p, \mathbf{W}) = \prod_{p=1}^n \frac{(\beta_i q_p)^{y_{p+i}}}{(1 + \beta_i q_p)^k}.$$

We stipulate a beta-prime prior distribution on β_i so that its posterior distribution is of the form

$$p(\beta_i | \mathbf{Y}_p, \mathbf{W}, \boldsymbol{\beta}^{(i)}) \propto \prod_{p=1}^n \frac{(\beta_i q_p)^{y_{p+i}}}{(1 + \beta_i q_p)^k} \frac{\beta_i^{a-1}}{(1 + \beta_i)^{a+b}}$$

This posterior distribution is intractable, but in Appendix D we derive an efficient Metropolis algorithm to produce samples from it. We have used $a = b = 1$ in our analyses.

Step 3: Simulating $\boldsymbol{\theta}$

The full-conditional distribution of θ_{ij} is of the form

$$p(\theta_{ij} | w_{+ij}) \propto \theta_{ij}^{w_{+ij}} (1 - \theta_{ij})^{N - w_{+ij}} p(\theta_{ij}),$$

where w_{pij} is the state of the edge between nodes i and j in the topology of a person p . We stipulate a beta (a , b) prior on θ_{ij} , so that the posterior is the beta ($a + w_{+ij}$, $b + N - w_{+ij}$) distribution. We have used $a = b = 1$ in our analyses.

Appendix D

The metropolis algorithm to simulate from $p(\beta_i | \mathbf{Y}, \mathbf{W}, \boldsymbol{\beta}^{(i)})$

We wish to sample from the posterior distribution,

$$p(\beta_i | \mathbf{Y}, \mathbf{W}, \boldsymbol{\beta}^{(i)}) \propto \prod_{p=1}^n \frac{(\beta_i q_p)^{y_{p+i}}}{(1 + \beta_i q_p)^k} \frac{\beta_i^{a-1}}{(1 + \beta_i)^{a+b}}.$$

We consider the Metropolis approach of Maris et al. (2015) to sample from this posterior distribution. After a change of variables $\mu_i = \frac{1}{2} \ln(\beta_i)$, the posterior distribution is proportional to

$$p(\mu_i | \mathbf{Y}, \mathbf{W}, \boldsymbol{\mu}^{(i)}) \propto$$

¹²However, we have to be careful here as thresholds are probably identified at the cluster-level in this case.

$$\prod_{p=1}^N \frac{(\exp(2\mu_i) q_p)^{y_{p+i}}}{(1 + \exp(2\mu_i) q_p)^k} \frac{\exp(2\mu_i)^a}{(1 + \exp(2\mu_i))^{a+b}}.$$

The logarithm of this posterior is concave, has a single mode, and has linear tails:

$$\begin{aligned} \frac{\partial}{\partial \mu_i} \ln(p(\mu_i | \mathbf{Y}, \mathbf{W}, \boldsymbol{\mu}^{(i)})) &\rightarrow \\ \begin{cases} 2(y_{++i} + a) - 2(N \cdot k + a + b) & \text{as } \mu_i \rightarrow \infty, \\ 2(y_{++i} + a) & \text{as } \mu_i \rightarrow -\infty. \end{cases} \end{aligned} \quad (\text{D1})$$

Maris et al. (2015) proposed to use these properties of the focal posterior distribution to create a proposal distribution of the form

$$f(\mu_i) \propto \frac{(d e^{2\mu_i})^{[y_{++i}+a]}}{(1 + d e^{2\mu_i})^{N \cdot k + a + b}}.$$

The logarithm of this proposal distribution—a generalized Beta-Prime distribution—has the same linear tails as the target distribution and has a similar slope. The constant d is used to ensure that the proposal closely matches the target distribution at the Markov chain's current state. It is specified such that the derivatives of the log of the proposal and the log of the target distribution coincide.

If $\hat{\mu}$ is the current state of μ_i in the Markov chain, then it follows that

$$d = \frac{\sum_{p=1}^N k \frac{q_p}{1+\hat{\beta} q_p} + (a+b) \frac{1}{1+\hat{\beta}}}{N \cdot k + a + b - \sum_{p=1}^N k \frac{\hat{\beta} q_p}{1+\hat{\beta} q_p} - (a+b) \frac{\hat{\beta}}{1+\hat{\beta}}},$$

where $\hat{\beta} = \exp(2\hat{\mu})$. The Metropolis algorithm then boils down to the following four steps.

1. Compute d from the above expression.
2. Sample Z from a Beta (a , b) distribution. Then, $\mu' = \log(d^{-1} \frac{Z}{1-Z})/2$.
3. Sample U from a Uniform (0 , 1) distribution, and compute

$$\alpha = \min \left\{ 1, \frac{\prod_{p=1}^N \frac{(\exp(2\mu') q_p)^{y_{p+i}}}{(1 + \exp(2\mu') q_p)^k} \frac{\exp(2\mu')^a}{(1 + \exp(2\mu'))^{a+b}} \times \frac{(d \exp(2\hat{\mu}))^{y_{++i}+a}}{(1+d \exp(2\hat{\mu}))^{N \cdot k + a + b}}}{\prod_{p=1}^N \frac{(\exp(2\hat{\mu}) q_p)^{y_{p+i}}}{(1 + \exp(2\hat{\mu}) q_p)^k} \frac{\exp(2\hat{\mu})^a}{(1 + \exp(2\hat{\mu}))^{a+b}} \times \frac{(d \exp(2\mu'))^{y_{++i}+a}}{(1+d \exp(2\mu'))^{N \cdot k + a + b}}} \right\}$$

4. Set the new state of μ_i to μ' if $U \leq \alpha$ and to $\hat{\mu}$ otherwise.