The idiographic Ising model

Marsman, M.[1]
1 University of Amsterdam

Correspondence concerning this article should be addressed to:
Maarten Marsman
University of Amsterdam, Psychological Methods
Nieuwe Achtergracht 129B
PO Box 15906, 1001 NK Amsterdam, The Netherlands
E-mail may be sent to m.marsman@uva.nl.

MM was supported by a Veni grant (451-17-017) from the Netherlands Organization for Scientific Research (NWO).

Abstract

The Ising model is a graphical model that has played an essential role in the field of network psychometrics, where it has been used as a theoretical model to conceptualize psychological concepts and as a statistical model for the analysis of psychological data. But in network psychometrics, the psychological data that are analyzed often come from cross-sectional applications, and the practice of using graphical models such as the Ising model to analyze these data has been heavily critiqued. The primary voiced concern centers around the inability of the Ising model to express heterogeneity in the population, and the subsequent necessity to assume that the population is homogeneous w.r.t. the network's structure. The idiographic approach has been posed as an alternative and aims to infer individual network structures to accommodate population heterogeneity. However, it is unclear how models of the individual aggregate into established cross-sectional phenomena that can be described using, for instance, the Ising model. This paper analyzes a formal bridge between idiographic and cross-sectional approaches in the case of the Ising model. We ascertain unique topological structures that characterize individuals and aggregate into an Ising model cross-sectionally. This formulation is consistent with cross-sectional phenomena and accommodates population heterogeneity.

Keywords: Ising model, random-cluster 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 the way that we come 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, Kan, Marsman, & Stevenson, 2017), personality (Constantini et al., 2019; Cramer et al., 2012), and attitudes (Dalege et al., 2016; Dalege, Borsboom, van Harreveld, & van der Maas, 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, Waldorp, van der Maas, & Borsboom, 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, Maris, Bechger, & Glas, 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 (Constantini et al., 2015; Dalege, Borsboom, van Harreveld, & van der Maas, 2017; Marsman et al., 2015; Marsman, Tanis, Bechger, & Waldorp, 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, Wright, Markon, & Krueger, 2017, 2019, in press; Fried & Cramer, 2017). The topic of this paper is a prominent critique on the assumption that the population that is analyzed is strictly homogeneous regarding the model's characteristics (e.g., E. Bos & Wanders, 2016; F. M. Bos et al., 2017; Brusco, Steinley, Hoffman, Davis-Stober, & Wasserman, in press). Homogeneity here implies that there can be no individual differences in the associations between the network's variables, and 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, Frankenhuis, Waldorp, & Borsboom, 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

[1] It is interesting to observe that despite their conceptual differences, the statistical models that underlie the network and latent variables approaches have been shown to be formally related (Epskamp, Maris, Waldorp, & Borsboom, 2018; Marsman et al., 2018).

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, Snippe, Bruggeman, Wichers, & van der Krieke, in press; Fisher, 2015; Fisher, Medaglia, & Jeronimus, 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, Drukker, Hasmi, & van Os, 2016; de Vos et al., 2017; Fisher, Reeves, Glenn, Medaglia, & Rubel, 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 seems to be oblivious 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, it is, in general, unclear how models of the individual fit in.

To overcome population heterogeneity using an idiographic network approach we need to find a bridge between networks on the levels of the individual and the group. This paper aims to introduce that bridge between the idiographic and cross-sectional approaches in the case of the Ising model. An idiographic interpretation of the Ising model is formulated that supports population heterogeneity and is consistent with cross-sectional phenomena. 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, Marsman, van der Maas, and Maris (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 use the theory to show how network structures that are unique to the individual can generate an Ising model cross-sectionally and how to traverse the two levels of networks.

This article comprises three sections. In the first section, we briefly review two distinct approaches in the analysis of networks: 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. In the second section, we use Fortuin and Kasteleyn's theory to formulate our idiographic approach to the Ising model. In this formulation, we obtain the Ising model in cases where the network's topology is a random effect that varies at the individual level, and we clarify the connection between associations at the individual- and group-level. In the third section, we turn to the statistical analysis of idiographic networks. We reveal how we can learn about an idiographic network from its posterior distribution and how we can use the posterior estimates of several idiographic networks to learn about the fit of the Ising model. We illustrate these ideas using an empirical example and some simulations. The paper ends with a discussion.

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 that were studied in two disparate areas of research. Percolation research, on the one hand, studies the way that particles trickle through a porous object, such as the way that water drips through coffee grounds. In research into

magnetism, on the other hand, 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 that they observed between these two fields of research were more than just coincidence (Grimmet, 2006), and worked out how the theories of magnetism and percolation relate (Fortuin, 1972a, 1972b; Fortuin & Kasteleyn, 1972). The approach that Fortuin and Kasteleyn used to consolidate the two scientific theories will be the basis of our idiographic network characterization of the Ising model.

The theories of percolation and magnetism are formalized using two distinct types of network models, and 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 is conceptualized using the Ising model. Thus, the theory that is developed by Fortuin and Kasteleyn informs us how the Ising model relates to random graph models, which is key to our idiographic formulation. In this paper, we 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 a p-variable network is characterized by the following probability distribution

$$p\left(\mathbf{W} = \left(w_{12},\ \ldots,\ w_{(p-1)p}\right)^{\mathsf{T}}\right) = \prod_{i=1}^{p-1} \prod_{j=i+1}^{p} \theta_{ij}^{w_{ij}} \left(1 - \theta_{ij}\right)^{1-w_{ij}}, \tag{1}$$

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.

The left panel in Figure 1 illustrates the Erdős-Rényi model for a network with three nodes. The nodes of the network are fixed and labeled one to three, and the edges between nodes constitute the random variables W_{ij}. 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. Two possible realizations from the Erdős-Rényi model for the three node network are shown in Figure 2. 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}. This is indicated with a black solid line 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 gray dashed line in the right panel, for example.

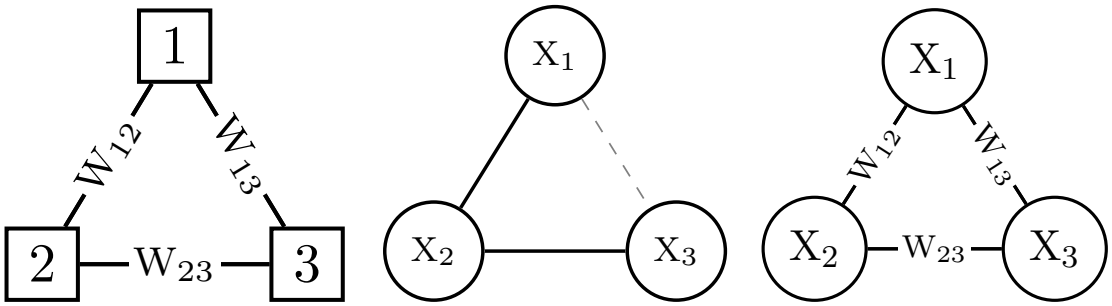

Figure 1. Three distinct approaches to a three variable network. The left figure illustrates a random graph model in which random variables are associated with the edges of the network. The middle figure illustrates a graphical model in which random variables are associated with the nodes of the network. 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.

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. 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.

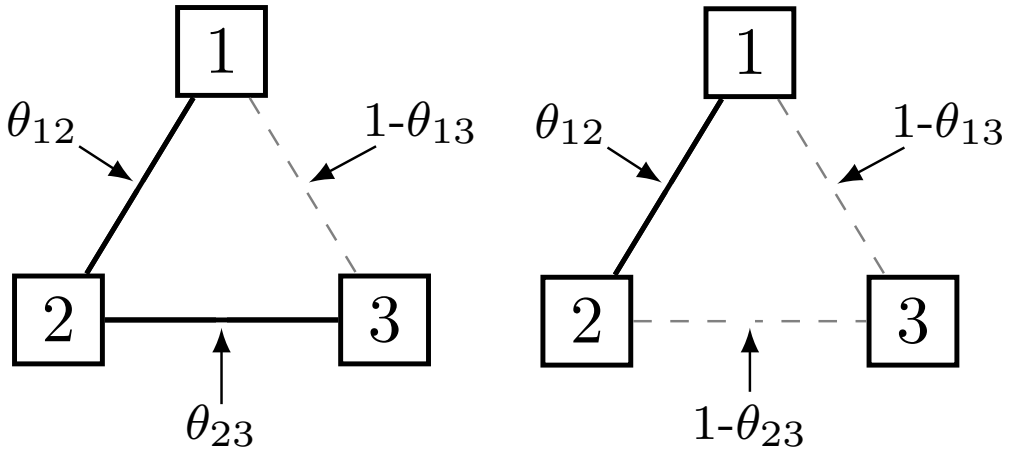

Figure 2. Two realizations of the Erdős-Rényi model for the three node network. Present edges are indicated with a black solid line and absent edges with a gray dashed line.

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, and 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 for this person, the two observables are not directly related. 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 distribution of 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 p dichotomous random variables X_i that each take values in $\{-1,\ +1\}$,[2]

$$p\left(\mathbf{X} = (x_1,\ \ldots,\ x_p)^\mathsf{T}\right) = \frac{1}{\mathrm{Z}_I} \exp\left(\sum_{i=1}^{p} \mu_i x_i + \sum_{i=1}^{p-1} \sum_{j=i+1}^{p} \sigma_{ij} x_i x_j\right), \tag{2}$$

where Z_I is the model's normalizing constant. The Ising model in Eq. (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. Edges that are present are indicated with a black solid line, and absent edges are indicated with a gray dashed line. Nodes one and two are directly related in the middle panel network of Figure 1, for example, 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 Eq. (2), the 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 —$\sigma_{23} = 0$— in the Ising model.

[2] In 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, Epskamp, Marsman, & Waldorp, in press).

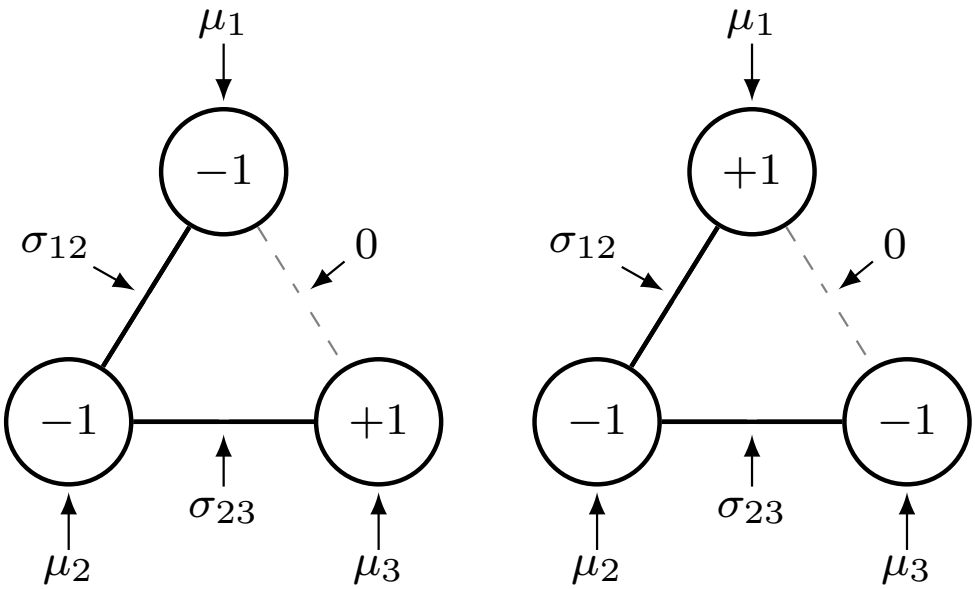

Figure 3. Two realizations of the Ising model for the three node network. Present relations are indicated with black solid line and absent relations with a gray dashed line.

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— 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, or 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 as it is formulated in Eq. (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

To blend the two scientific theories, Fortuin and Kasteleyn showed that the Ising model that is used to conceptualize a magnet is consistent with the view that the network's topology is a random effect. That is, they showed that the probability of observing a pattern

of variable states $\mathbf{x} = (x_1, \ldots)^\mathsf{T}$ in the Ising model can be characterized as follows (see Appendix A for a proof),

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

where a random graph model $p(\mathbf{w})$ is used to express the probability of observing a topological structure $\mathbf{w} = (w_{12}, \ldots)^\mathsf{T}$, and the conditional distribution $p(\mathbf{x} \mid \mathbf{w})$ 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, $p(\mathbf{x} \mid \mathbf{w})$ describes the process of coloring the nodes of the graph that is induced by $\mathbf{w}$.

Fortuin and Kasteleyn's characterization provides an alternative interpretation of the Ising model, one in which the model's unknown topology is an instance of a random graph model. That is, Eq. (3) suggests the following hierarchical scheme for generating a configuration of observables $\mathbf{x}$:[3]

1. Generate a topological structure from the random graph model: $\mathbf{w}^* \sim p(\mathbf{w})$.

2. Color the nodes of the generated graph structure: $\mathbf{x}^* \sim p(\mathbf{x} \mid \mathbf{w}^*)$.

We view the topological structure as a person's constellation of relations between symptoms, responses, or attitudes, and the process of node coloring as the process of configuring the observables in a person's topology. Taken together, the pair $(\mathbf{x}^*, \mathbf{w}^*)$ then constitutes a single person's idiographic network. If we repeat the two-step procedure, we would obtain a new topological structure $\mathbf{w}^{**}$ and configuration of observables $\mathbf{x}^{**}$. The idea is thus that both nodes and edges are random variables and that a particular configuration of nodes and edges then characterizes the individual. 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.

When the generated graphs form a particular distribution, and the graphs' clusters are independently colored, the configurations of observables are consistent with an Ising model in the population. In their seminal work, Fortuin and Kasteleyn (1972) showed that this is the case if the topologies are samples from a random-cluster model. The random-cluster model is defined by the following probability distribution over topological structures $\mathbf{w}$:

$$p(\mathbf{W} = \mathbf{w}) = \frac{1}{\mathrm{Z}_R} \overbrace{\prod_{i=1}^{n-1} \prod_{j=i+1}^{n} \theta_{ij}^{w_{ij}} \, (1 - \theta_{ij})^{1 - w_{ij}}}^{\text{Erdős-Rényi model}} \lambda^{\kappa(\mathbf{w})}, \tag{4}$$

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

[3] An alternative interpretation of Eq. (3) is given by Holland (1990). In his sampling interpretation, an individual can be characterized by a latent topology $\mathbf{w}^*$ that is sampled from a certain population $p(\mathbf{w})$. The proportion of persons with topology $\mathbf{w}^*$ that have a particular configuration $\mathbf{x}^*$ is then equal to $p(\mathbf{x}^* \mid \mathbf{w}^*)$. For the Fortuin and Kasteleyn characterization this will hold at the level of clusters.

normalizing constant. Observe that the random-cluster model has the same kernel as the Erdős-Rényi model in Eq. (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 Eq. (1). Fortuin and Kasteleyn showed that the characterization in Eq. (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 one, two, or three 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.

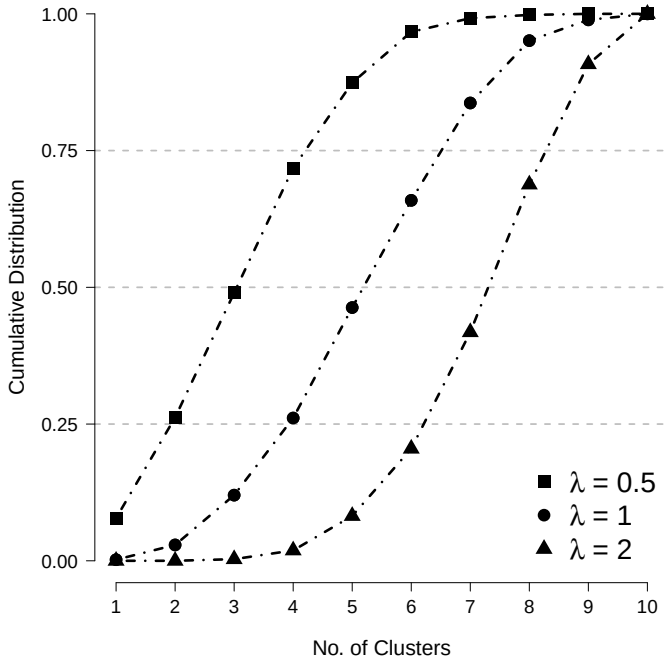

Figure 4. The (cumulative) distribution of clusters in ten-variable networks that were generated from random cluster models with constant edge probabilities $\theta_{ij} = \theta = 0.1$ but with distinct values for the clustering weight.

The application of an Erdős-Rényi model to describe the distribution of random

graphs in Eq. (3) is known as the divide and color model (Häggström, 2001). The divide and color model colors the graph in the same fashion as the Fortuin-Kasteleyn model does. This process colors all nodes in a cluster either $+1$ or -1 with equal probabilities and colors distinct clusters independently,[4]

$$p(\mathbf{X} = \mathbf{x} \mid \mathbf{w}) = \prod_{c=1}^{\kappa(\mathbf{w})} P(\mathbf{X}_c = \mathbf{1})^{\frac{x_c+1}{2}} \, P(\mathbf{X}_c = -\mathbf{1})^{1-\frac{x_c+1}{2}} = \frac{1}{2^{\kappa(\mathbf{w})}}, \tag{5}$$

where $\mathbf{X}_c$ denotes the observables in a cluster c and x_c denotes their collective state. This graph coloring process, in combination with the random-cluster model with a clustering weight of 2, coincides with an Ising model without main effects. 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, Kasteleyn, & Ginibre, 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 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 and the Fortuin-Kasteleyn model 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 and Tykesson, 2017, Eq. 1.1)

$$P(X_i = X_j) = \frac{1}{2} + \frac{1}{2} \, P(X_i \leftrightarrow X_j) \geq \frac{1}{2}, \tag{6}$$

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 values of observables align when they are in the same cluster, their correlation will be high in the population 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 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 wiring 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. The graph generating mechanisms of the random-cluster model and the Erdős-Rényi model are thus fundamentally different and may lead to divergent predictions

[4]The random-triangle model (Häggström & Jonasson, 1999; Jonasson, 1997) illustrates that our exclusive focus on the network's clustering is not strictly necessary for formulating a relation between the Ising model and a random graph model.

and interventions on underlying topologies. The fact that these differences have a noticeable impact on observed correlations makes the fragmentation of idiographic networks an exciting avenue for future research.

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

$$\sigma_{ij} = -\frac{1}{2}\log(1 - \theta_{ij}). \tag{7}$$

We illustrate the relation in Figure 5. 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).

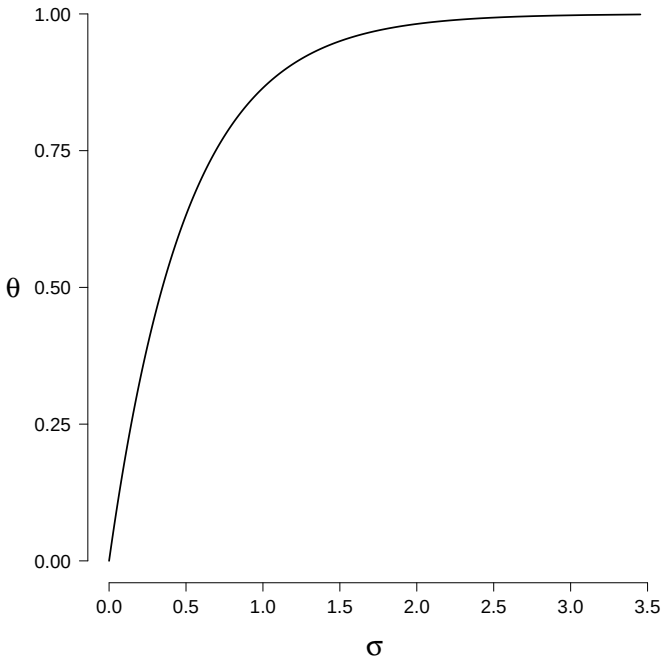

Figure 5. The theoretical relation between the association parameter σ of the Ising model and the edge probability θ in the random-cluster model.

The relationship in Eq. (7) 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 Eq. (7) 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 wiring probabilities provides us with an alternative view on psychometric models. Savi et al. (2019), for example, characterizes the bi-factor structure, which is ubiquitous in psychometrics, in terms of an underlying community or stochastic block structure. This connection between psychometric models and models from network science offers another exciting avenue for future research (Marsman, Waldorp, & Borsboom, 2019).

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 main effects impact the clustering weights of the random-cluster model, the cluster coloring probabilities, and as a result, also influence the link between observables and wiring probabilities in Eq. (6). The clustering weights, for example, evolve to

$$\lambda_c = 2\cosh\left(\sum_{i\in V_c}\mu_i\right) = \exp\left(\sum_{i\in V_c}\mu_i\right) + \exp\left(-\sum_{i\in V_c}\mu_i\right) \geq 2, \tag{8}$$

where V_c denotes the set of indices of the vertices in cluster c. Thus, by including main effects, the random-cluster model now assigns different weights to the different clusters,

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

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. The inclusion of main effects also impacts the coloring probabilities,

$$P(\mathbf{X}_c=\mathbf{1}) = \frac{\exp\left(\sum_{i\in \mathrm{V}_c}\mu_i\right)}{\exp\left(\sum_{i\in \mathrm{V}_c}\mu_i\right)+\exp\left(-\sum_{i\in \mathrm{V}_c}\mu_i\right)} = \frac{1}{\lambda_c}\exp\left(\sum_{i\in \mathrm{V}_c}\mu_i\right).$$

In psychopathology, for example, it becomes more probable that the clusters are colored -1, and in education, on the other hand, it is likely that the clusters are colored $+1$. Theorem 2 of Cioletti and Vila (2016) captures the impact of the main effects on the relation in Eq. (6). The connection between edge probabilities in the random-cluster model and pairwise associations in the Ising model in Eq. (7) is unaffected by the inclusion of main effects, however.

The Statistical Analysis of the Idiographic Ising Model

Our theory uses random graphs to accommodate the heterogeneity of associations in the population. These random graphs, however, are not directly observed and need to

be estimated from observed data. The latent graphs are estimated using their posterior distribution. If we assume that the repeated observations $\mathbf{x}_1$, $\mathbf{x}_2$, ... 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 can then use the posterior distributions of the latent networks to infer about the fit of the model and investigate potential sources of the misfit, in case there is any. We illustrate these methods using some small simulations and an empirical example. Before we illustrate the methods, we start with a cross-sectional analysis of depression symptoms using the Ising model.

A Cross-Sectional Network Analysis of Major Depression

To illustrate the application of an Ising model to psychological data, we analyze the nine aggregate symptoms of major depression in the third Diagnostic and Statistical Manual of Mental Disorders (DSM-III-R) based on data from 8,973 persons in the Virginia Adult Twin Study of Psychiatric and Substance Use Disorders (VATSPSUD; Kendler & Prescott, 2006).[5] We estimate the Ising model using the R-package `IsingFit` (van Borkulo et al., 2014), which combines a l_1-regularized logistic regression (Ravikumar, Wainwright, & Lafferty, 2010; Wainwright, Ravikumar, & Lafferty, 2007) with EBIC edge selection (Barber & Drton, 2015; Chen & Chen, 2008) to estimate the Ising model. This method regularizes small association parameters to exactly zero to prevent overfitting. The R-package `IsingFit` is the most commonly used method to analyze Ising networks in psychology.

Figure 6 shows the estimated network. Observe that the estimated associations are all non-negative. Except for the relation between concentration problems and feelings of worthlessness —there is no edge between these two symptoms— all associations are positive, which is indicated by green edges in the network. The differences in edge thickness indicate differences in association strengths, with thicker edges indicating stronger associations. So, what do these results imply? They imply that (i) it is likely that if we randomly select an individual from the group that happens to have a depressed mood that he or she also experiences a change in appetite. However, it is (ii) more likely that this person also experiences a loss of interest, since there is a larger association between depressed mood and loss of interest than between depressed mood and change of appetite (indicated with a thicker edge in the network). Observe that these two implications or interpretations are about a randomly selected individual from the population and says nothing about what is or is not the case for a particular person. A person that has a depressed mood either experiences a change in appetite or not and this change in appetite is, or is not, influenced by having a depressed mood.

The situation in which an association is absent is somewhat different, however, and implies that the two variables are independent given the rest of the networks' state. There is an absent association between concentration problems and feelings of worthlessness in our example. Thus, for two randomly selected persons from the population that experience the same problems with the seven remaining symptoms in the network, concentration problems and weight loss behave independently. At the individual level, this would imply that concentration problems do not directly influence weight loss or vice versa. This independence

[5] Cramer et al. (2016) provide a detailed network analysis of the 14 disaggregated symptoms.

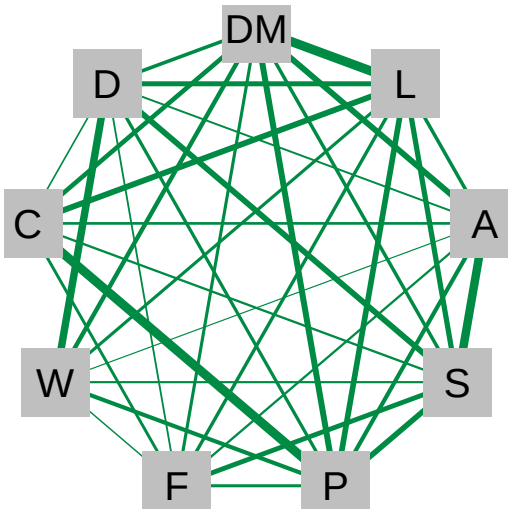

Figure 6. The estimated associations between the nine aggregate symptoms of major depression in the DSM-III-R based on data from 8,973 persons from the VATSPSUD study. Here, DM stands for having a *depressed mood*, L for *loss of interest*, A stands for a *change in appetite or weight* (increased or decreased), S for *sleep problems* (insomnia or hypersomnia), P stands for *psychomotor issues* (agitation or retardation), F stands for fatigue, W for *feelings of worthlessness* (gain or loss), C for *concentration problems* and D for *thoughts of death*. The estimated graph was visualized using the R-package `igraph` (Csardi & Nepusz, 2019).

then applies to every person in the population, which is a powerful implication.

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),

$$p(\mathbf{w} \mid \mathbf{x}) = \prod_{i=1}^{n-1} \prod_{j=i+1}^{n} \left\{ \overbrace{\left[\theta_{ij}^{w_{ij}} \, (1-\theta_{ij})^{1-w_{ij}}\right]^{\delta_{(x_i,\, x_j)}}}^{x_i \text{ and } x_j \text{ in the same state}} \times \overbrace{\left[0^{w_{ij}} \, 1^{1-w_{ij}}\right]^{1-\delta_{(x_i,\, x_j)}}}^{x_i \text{ and } x_j \text{ in different states}} \right\},$$

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:

$$p(W_{ij} = w_{ij} \mid x_i = x_j) = \theta_{ij}^{w_{ij}}(1 - \theta_{ij})^{1-w_{ij}}$$
$$p(W_{ij} = w_{ij} \mid x_i \neq x_j) = 0^{w_{ij}}\, 1^{1-w_{ij}},$$

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

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

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.

We show the posterior expectation of the idiographic topology for two individuals in the VATSPSUD analysis in Figure 7. In the networks of Figure 7, the nodes reflect the symptom states of the individual and are colored red if the symptom is absent and colored green if the symptom is present. The two topological structures in Figure 7 are equal to the topological structure that we showed in Figure 6, with the exception that some edges are now removed. The two persons in Figure 7 have a different configuration of symptoms, which in turn inspires the different topological structures in Figure 7. The idiographic network on the left comprises more potential connections than the one on the right. Observe, however, that the EAP estimate only reflects the population averages and fails to convey the inherent uncertainty of the projected edges and individual topologies. As a result, the EAP estimate does not inform about the plausibility of a particular configuration of links, i.e., what could be a person's topology. Plausible networks, however, capture the uncertainty of edge configurations and make clear what could have been the person's topology. Figure 8 shows a plausible network for the two persons. Observe that, contrary to the EAP estimates in Figure 7, the plausible networks are unweighted and also much sparser than the networks in Figure 8. Of course, neither of the two network structures in Figure 8 needs to be the real underlying network. However, since they are consistent with everything we know about the latent topologies, they could be.

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 $\mathbf{x}_1$, $\mathbf{x}_2$, $\ldots$, $\mathbf{x}_k$ on the

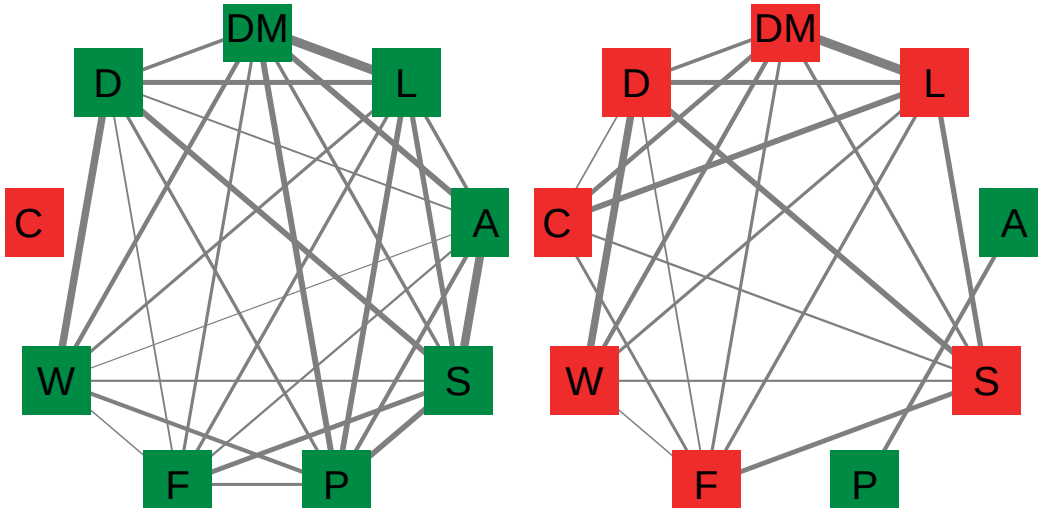

Figure 7. The EAP estimates of the idiographic networks of two persons from the VAT-SPUD study data. Absent symptoms (nodes) are colored red and present symptoms (nodes) are colored green.

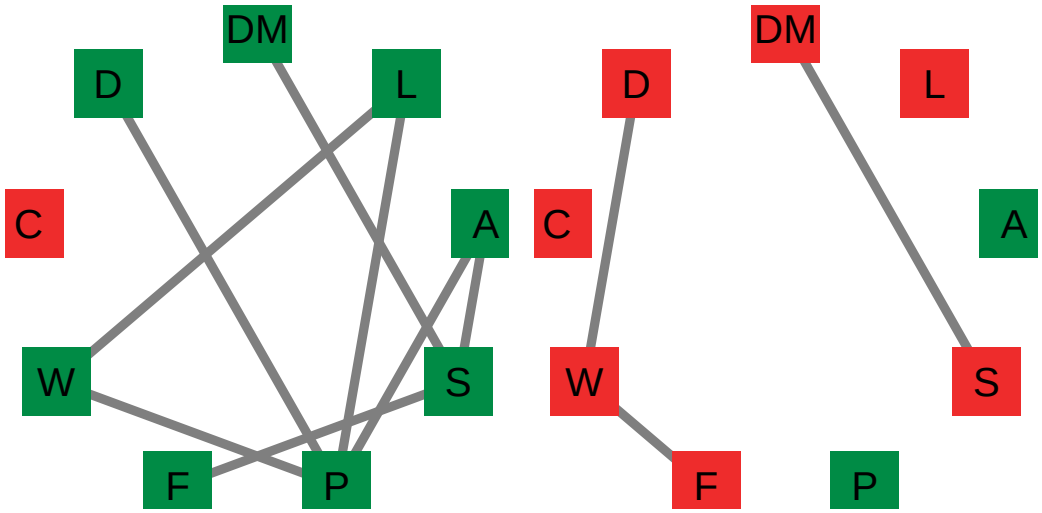

Figure 8. One plausible network for the two persons for which EAP estimates were given in Figure 7.

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 are in the same state for all k observations, 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.[6] Second, with more observations, it becomes more likely that the posterior distribution excludes inconsistent edges. Third, consistent with the posterior distribution after one observation, the posterior distribution after k observations seems to favor less fragmented networks than the prior distribution, since here $\lambda_c > \lambda_c^{1-k}$ for $k \geq 1$. 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. In Appendix B, we derive a Gibbs sampling approach.

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 $r = 1, 2, \ldots, k$ observations. Figure 9 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.

What Can We Learn From Plausible Networks?

Marsman, Maris, Bechger, and Glas (2016) have recently shown that if a latent variable model is correctly specified, then draws from the posterior distribution of the latent variables —plausible values— can be used to learn about the correct underlying distribution of the latent variables. Thus, assuming that the coloring process is correct, we may investigate the distribution of the plausible networks to infer if the estimated random-cluster

[6] Similarly, one finds that the random-triangle model is the conjugate prior for the triangle-based coloring process that was analyzed by Jonasson (1997) and Häggström and Jonasson (1999).

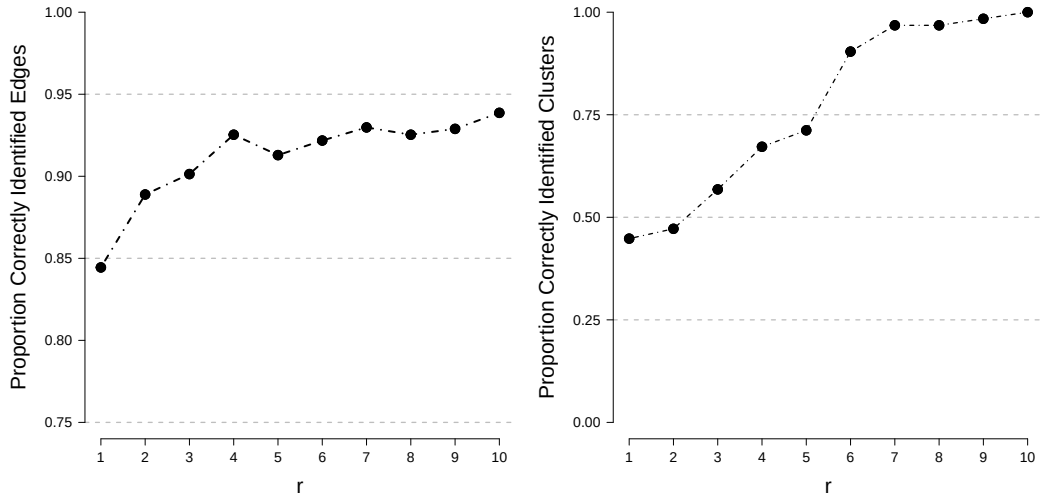

Figure 9. An $n = 10$ variable network was generated for a single person and 25 plausible values were generated based on r observations of the network. The left panel shows the proportion of correctly identified edges and the right panel the proportion of correctly identified clusters (averaged over plausible networks).

model fits the observed data. If the random-cluster model is misspecified, this implies that the Ising model is misspecified. In this case, plausible networks are not only a convenient tool to diagnose the fit of the Ising model, but by studying the (distribution of) plausible networks one can also identify possible modeling directions.

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 random-cluster model— if it is misspecified. Second, they showed that the plausible network distribution converges towards 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 9 underline this idea.

We illustrate the idea first using data simulated from a divide and color model. Idiographic topologies for $n = 10$ variable networks are generated from a Erdős-Rényi model with constant edge probabilities $\theta_{ij} = \theta = 0.15$, for each of $N = 1,000$ cases. Observables were generated using the coupling model $p(\mathbf{x} \mid \mathbf{w})$ including external fields $\boldsymbol{\mu}$. The fields were sampled uniformly between -0.5 and 0.5. The corresponding Ising model is known as the Curie-Weiss model (Kac, 1968; Kochmański, Paszkiewicz, & Wolski, 2013), and the associated random-cluster model was estimated using the methodology of Marsman, Tanis, et al. (2019). One plausible network was generated for each of the N persons based on $k = 1$ and $k = 10$ observations. Figure 1 shows the correct and estimated cluster distributions. Figure 1 reveals that the estimated random-cluster model generates more clusters than the correct random graph model. In line with the results from Marsman et al. (2016), the cluster

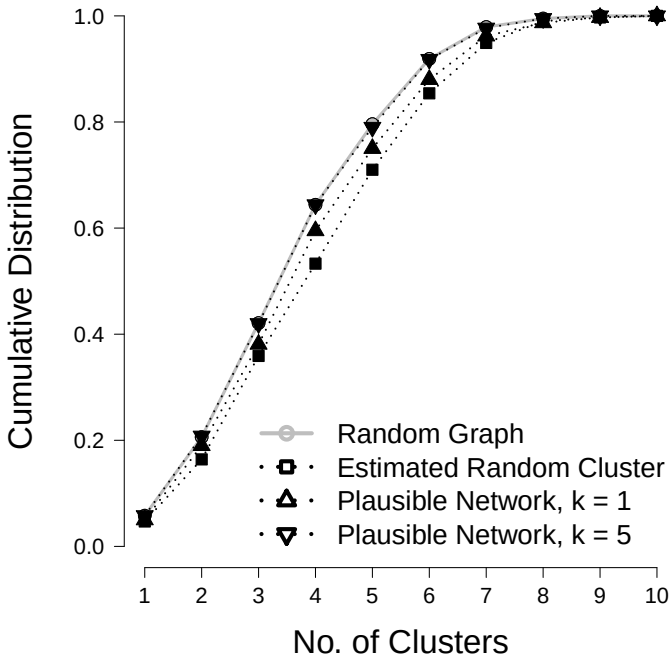

Figure 10. Data were generated from a divide and color model, in which topologies for $n = 10$ variable networks were generated from an Erdős-Rényi model with constant edge probabilities $\theta_{ij} = \theta = 0.15$, for each of $N = 1,000$ persons. Observables were generated using the coloring process of Fortuin and Kasteleyn, including external fields that were sampled uniformly between -0.5 and 0.5. The (cumulative) cluster distributions for the estimated random-cluster model (blue), the plausible networks based on $k = 1$ observations (red), the plausible networks based on $k = 10$ observations (black) and the correct random graph model (gray) are shown here.

distribution of the plausible networks diverges from the postulated population model —the estimated random-cluster model— and converges to the correct random graph model as the number of observations increase. This illustrates that plausible networks can be quite useful to analyze the empirical fit of the random-cluster model, or the Ising model. Observe that even after a single observation the cluster distribution of the plausible networks is distinct from the estimated random-cluster model.

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 version of the Ising model is that it does not require a homogeneous population for the model to apply. The formulation of the model that we have analyzed in this paper shows that we can obtain the Ising model cross-sectionally even if topological structures vary on an individual basis. Thus, homogeneity of the population is a sufficient condition for the Ising model's fit, but not a necessary one, as some authors tend to believe. This result suggests a reconciliation of the idiographic and nomothetic approaches in network psychometrics.

Eq. (6) 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 wiring 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, Eq. (7) provides a parametric link between group-level associations σ_{ij} and idiographic wiring probabilities θ_{ij}.

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. On top of this, we have also shown that we can, in principle, uncover the latent topologies that cause this 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. We will discuss two potential lines of research.

A unified network theory

The focus of this article was the idiographic framing of Ising models using the theory of Fortuin and Kasteleyn. Their theory, however, also offers a statistical framework that comprises both graphical models and random graph models. The 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., in press; Epskamp et al., in press; Marsman, Waldorp, & Borsboom, 2019; Steinley, Hoffman, Brusco, & Sher, 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, 2011), 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, which can be teased apart from observed correlations. The fragmentation of topologies thus seems to be a good place to start. 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.

Generalizing Fortuin and Kasteleyn's theory

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).[7] An idiographic formulation of the popular Gaussian Graphical model (GGM; Lauritzen, 2004), used to analyze networks of continuous variables, does not exist, for example.

One may wonder if the idiographic formulation of the Ising model translates to other models, such as the GGM. To come to such idiographic formulation, we envisage two steps. First, the binary indicator variables $\delta_{(x_i,\, x_j)}$ in the Fortuin-Kasteleyn model can be reformulated as $\delta_{(\text{sign}(x_i),\, \text{sign}(x_j))}$, where $\text{sign}(x)$ is the sign function. The latter formulation of the indicator variable also meaningfully exists for Gaussian variables. Second, the conditional

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

distribution $p(x \mid w)$ and the marginal distribution $p(w)$ in the Fortuin-Kasteleyn model are derived from the posterior $p(w \mid x)$ and marginal $p(x)$. A similar approach would work for the case of Gaussian variables where one could, for example, choose the posterior $p(w \mid x)$ to be an Erdős-Rényi model, as in the case of the Fortuin-Kasteleyn model. It would be interesting to see to what idiographic formulation of the Gaussian graphical model these two choices would lead. The same holds for the categorical variable models of, for example, Hessen (2012).

References

Ashkin, J., & Teller, E. (1943). Statistics of two-dimensional lattices with four components. *Physical Review*, *64*(5), 178–184. doi: 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. doi: 10.1371/journal.pone.0162811

Barabási, A.-L., & Albert, R. (1999). Emergence of scaling in random networks. *Science*, *286*(5439), 509–512. doi: 10.1126/science.286.5439.509

Barber, R. F., & Drton, M. (2015). High dimensional Ising model selection with Bayesian information criteria. *Electronic Journal of Statistics*, *9*(1), 567–607. doi: 10.1214/15-EJS1012

Borsboom, D. (2008). Psychometric perspectives on diagnostic systems. *Journal of Clinical Psychology*, *64*(9), 1089–1108. doi: 10.1002/jclp

Borsboom, D., & Cramer, A. O. J. (2013). Network analysis: An integrative approach to the structure of psychopathology. *Annual Review of Clinical Psychology*, *9*, 91–121. doi: 10.1146/annurev-clinpsy-050212-185608

Bos, E., & Wanders, R. (2016). Group-level symptom networks in depression. *JAMA Psychiatry*, *73*(4), 411. doi: 1

Bos, F. M., Snippe, E., Bruggeman, R., Wichers, M., & van der Krieke, L. (in press). Insights of patients and clinicians on the promise of the experience sampling method for psychiatric care. *Psychiatric Services*. doi: 10.1176/appi.ps.201900050

Bos, F. M., Snippe, E., de Vos, S., Hartmann, J. A., Simons, C. J. P., van der Krieke, L., ... 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. doi: 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. doi: 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. doi: 10.1017/S0305004100032680

Brusco, M., Steinley, D., Hoffman, M., Davis-Stober, C., & Wasserman, S. (in press). On Ising models and algorithms for the construction of symptom networks in psychopathological research. *Psychological Methods*, 1–19. doi: 10.1037/met0000207

Brush, S. G. (1967). History of the Lenz-Ising model. *Reviews of Modern Physics*, *39*(4), 883–893. doi: 10.1103/RevModPhys.39.883

Caspi, A., Houts, R., Belsky, D., Goldman-Mellor, S., Harrington, H., Israel, S., ... Moffit, T. (2014). The p factor: One general psychopathology factor in the structure of psychiatric disorders? *Clinical Psychological Science*, *2*(2), 119–137. doi: 10.1177/2167702613497473

Chen, J., & Chen, Z. (2008). Extended Bayesian information criteria for model selection with large model spaces. *Biometrika*, *95*(3), 759–771. doi: 10.1093/biomet/asn034

Cioletti, L., & Vila, R. (2016). Graphical representations for Ising and Potts models in general external fields. *Journal of Statistical Physics*, *162*(1), 81–122. doi: 10.1007/s10955-015-1396-

5

Constantini, G., Epskamp, S., Borsboom, D., Perugini, M., Mõttus, R., Waldorp, L. J., & Cramer, A. O. J. (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. doi: 10.1016/j.jrp.2014.07.003

Constantini, G., Richetin, J., Preti, E., Casini, E., Epskamp, S., & Perugi, M. (2019). Stability and variability of personality networks: A tutorial on recent developments in network psychometrics. *Personality and Individual Differences*, *136*, 68–78. doi: 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. doi: 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), 1–20. doi: 10.1371/journal.pone.0167490

Cramer, A. O. J., van der Sluis, S., Noordhof, A., Wichers, M., Geschwind, N., Aggen, S. H., ... 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*, 414–431. doi: 10.1002/per.1866

Cramer, A. O. J., Waldorp, L. J., van der Maas, H. L. J., & Borsboom, D. (2010). Comorbidity: A network perspective. *Behavioral and Brain Sciences*, *33*(2–3), 137–193. doi: 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 acount of attitudes: The causal attitude network (CAN) model. *Psychological Review*, *123*(1), 2–22. doi: 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 Psychology and Personality Science*, *8*(5), 528–537. doi: 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. doi: 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*(e0178586), 1–18. doi: 10.1371/journal.pone.0178586

Epskamp, S., Fried, E. I., van Borkulo, C. D., Robinaugh, D. J., Marsman, M., Dalege, J., ... Cramer, A. O. J. (in press). Investigating the utility of fixed-margin sampling in network psychometrics. *Multivariate Behavioral Research*. doi: 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). New York, NY: 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. doi: 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. doi: 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. doi: 10.1037/abn0000311

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. doi: 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*.

Forbes, M. K., Wright, A. G. C., Markon, K. E., & Krueger, R. F. (in press). Quantifying the reliability and replicability of psychopathology network characteristics. *Multivariate Behavioral Research*. doi: 10.1080/00273171.2019.1616526

Fortuin, C. (1972a). On the Random-Cluster model: III. The simple Random-Cluster model. *Physica*, *59*(4), 545-570. doi: 10.1016/0031-8914(72)90087-0

Fortuin, C. (1972b). On the Random-Cluster model II. The percolation model. *Physica*, *58*(3), 393–418. doi: 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. doi: 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. Retrieved from `https://projecteuclid.org/euclid.cmp/1103857443`

Fried, E. I., & Cramer, A. O. J. (2017). Moving forward: Challenges and directions for psychopathological network theory and methodology. *Perspectives on Psychological Science*, *12*(6), 999—1020. doi: 10.1177/1745691617705892

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. doi: 10.1109/TPAMI.1984.4767596

Grimmet, G. (2006). *The random-cluster model*. Heidelberg, Germany: Springer-Verlag.

Häggström, O. (2001). Coloring percolation clusters at random. *Stochastic Processes and their Applications*, *96*(2), 213–242. doi: 10.1016/S0304-4149(01)00115-6

Häggström, O. (2002). *Finite Markov chains and algorithmic applications*. Cambridge, UK: 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`

Haslbeck, J., Epskamp, E., Marsman, M., & Waldorp, L. J. (in press). Interpreting the Ising model: The input matters. *Multivariate Behavioral Research*.

Hessen, D. (2012). Fitting and testing conditional multinormal partial credit models. *Psychometrika*, *77*(4), 693–709. doi: 10.1007/s11336-012-9277-1

Holland, P. W. (1990). On the sampling theory foundations of item response theory models. *Psychometrika*, *55*(4), 577–601. doi: 10.1007/BF02294609

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`

Ising, E. (1925). Beitrag zur theorie des ferromagnetismus. *Zeitschrift für Physik*, *31*(1), 253–258. doi: 10.1007/BF02980577

Jonasson, J. (1997). The random triangle model. *Journal of Applied Probability*, *36*(3), 852–867. Retrieved from `https://www.jstor.org/stable/3215446`

Kac, M. (1968). Mathematical mechanisms of phase transitions. In M. Chrétien, E. Gross, & S. Deser (Eds.), *Statistical physics: Phase transitions and superfluidity, Vol. 1, Brandeis University Summer Institute in Theoretical Physics* (pp. 241–305). New York: Gordon and Breach Science Publishers.

Karrer, B., & Newman, M. (2011). Stochastic blockmodels and community structure in networks. *Physical Review E: Statistical, Nonlinear, and Soft Matter Physics*, *83*(1), 016107.

doi: 10.1103/PhysRevE.83.016107

Kendler, K. S., & Prescott, C. A. (2006). *Genes, environment, and psychopathology. Understanding the causes of psychiatric and substance use disorders.* New York, NY: The Guilford Press.

Kievit, R. A., Frankenhuis, W. E., Waldorp, L. J., & Borsboom, D. (2013). Simpson's paradox in psychological science: A practical guide. *Frontiers in Psychology*, *4*(513), 1–14. doi: 10.3389/fpsyg.2013.00513

Kochmański, M., Paszkiewicz, T., & Wolski, S. (2013). Curie-Weiss magnet–a simple model of phase transition. *European Journal of Physics*, *34*(6), 1555–1573.

Lauritzen, S. (2004). *Graphical models.* Oxford: Oxford University Press.

Lenz, W. (1920). Beiträge zum verständnis der magnetischen eigenschaften in festen körpern. *Physikalische Zeitschrift*, *21*, 613–615.

Marsman, M., Borsboom, D., Kruis, J., Epskamp, S., van Bork, R., Waldorp, L. J., ... Maris, G. K. J. (2018). An introduction to network psychometrics: Relating Ising network models to item response theory models. *Multivariate Behavioral Research*, *53*(1), 15–35. doi: 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). doi: 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. doi: 10.1007/s11336-016-9497-x

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). Cham, Switzerland: Springer.

Marsman, M., Waldorp, L. J., & Borsboom, D. (2019). Towards a grand unified theory of network models: Reply to Brusco, Steinley, Hoffman, Davis-Stober, & Wasserman. *Retrieved from: https://psyarxiv.com/n98qt*. (PsyArXiv preprint.)

Mislevy, R. J. (1991). Randomization-based inference about latent variables from complex samples. *Psychometrika*, *56*(2), 177–196. doi: 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). Dordrecht, the Netherlands: Springer. doi: 10.1007/978-94-015-8326-8_15

Newman, M. (2011). Communities, modules and large-scale structure in networks. *Nature Physics*, *8*, 25–31. doi: 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. doi: 10.1007/s00407-004-0088-3

Potts, R. B. (1952). Some generalized order-disorder transformations. *Mathematical Proceedings of the Cambridge Philosophical Society*, *48*(1), 106–109. doi: 10.1017/S0305004100027419

Ravikumar, P., Wainwright, M. J., & Lafferty, J. D. (2010). High-dimensional Ising model selection using l_1-regularized logistic regression. *Annals of Statistics*, *38*(3), 1287–1319. doi: 10.1214/09-AOS691

Robinson, W. S. (1950). Ecological correlations and the behavior of individuals. *American Sociological Review*, *15*(3), 351–357.

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. doi: 10.1177/1745691619866447

Simpson, E. H. (1951). The interpretation of interaction in contingency tables. *Journal of the Royal Statistical Society. Series B (Methodological)*, *13*(2), 238–241.

Spearman, C. (1904). "General intelligence," objectively determined and measured. *The American Journal of Psychology*, *15*(2), 201–292. doi: 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*.

Swendsen, R., & Wang, J. (1987). Nonuniversal critical dynamics in monte carlo simulations. *Physical Review Letters*, *58*(2), 86–88. doi: 10.1103/PhysRevLett.58.86

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)). doi: 10.1038/srep05918

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. doi: 10.1037/0033-295X.113.4.842

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), 1–17. doi: 10.3390/jintelligence5020016

Wainwright, M. J., Ravikumar, P. K., & Lafferty, J. D. (2007). High-dimensional graphical model selection using l_1-regularized logistic regression. In B. Schölkopf, J. C. Platt, & T. Hoffman (Eds.), *Advances in neural information processing systems 19* (pp. 1465–1472). MIT Press.

Watts, D. J., & Strogatz, S. H. (1998). Collective dynamics of 'small-world' networks. *Nature*, *393*(6684), 440–442. doi: 10.1038/30918

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 follows[8]

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

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

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

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)

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

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

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,

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

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

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

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

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

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

The Idiographic Posterior

Given the above formulation of the joint distribution of edge and node variables in the Fortuin–Kasteleyn model, we may express the posterior distribution of the edge variables given the states of the node variables as follows

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

Next, we characterize the sum in the denominator.

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

Observe that the factors in the denominator are consistent with

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

such that the posterior can ultimately be expressed as

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

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

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

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

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

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

$$
p(\mathbf{w} \mid \mathbf{x}_1,\, \ldots,\, \mathbf{x}_k) = \frac{\prod_{i,\,j} \left[\theta_{ij}^{w_{ij}} (1-\theta_{ij})^{1-w_{ij}}\right]^{\delta_{ij}} \left[0^{w_{ij}} (1-\theta_{ij})^{1-w_{ij}}\right]^{1-\delta_{ij}} \prod_{c=1}^{\kappa(\mathbf{w})} \lambda_c^{1-k}}{\sum_{\mathbf{w}} \prod_{i,\,j} \left[\theta_{ij}^{w_{ij}} (1-\theta_{ij})^{1-w_{ij}}\right]^{\delta_{ij}} \left[0^{w_{ij}} (1-\theta_{ij})^{1-w_{ij}}\right]^{1-\delta_{ij}} \prod_{c=1}^{\kappa(\mathbf{w})} \lambda_c^{1-k}},
$$

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 $\mathbf{w}^{(t)}_{\setminus(a,\, b)}$. The probability that $W_{ab} = 1$ given the values $\mathbf{w}^{(t)}_{\setminus(a,\, b)}$ on the remaining variables is equal to

$$\frac{\theta_{ab} \prod_{c=1}^{\kappa(\mathbf{w}^{(t)}_{\setminus(a,\, b)},\, 1)} \lambda_c}{\theta_{ab} \prod_{c=1}^{\kappa(\mathbf{w}^{(t)}_{\setminus(a,\, b)},\, 1)} \gamma_c + (1-\theta_{ab}) \prod_{c=1}^{\kappa(\mathbf{w}^{(t)}_{\setminus(a,\, b)},\, 10)} \lambda_c},$$

where $\kappa(\mathbf{w}^{(t)}_{\setminus(a,\, b)},\, 1)$ denotes the number of clusters of $\mathbf{w}^{(t)}$ when $W_{ab} = 1$ and $\kappa(\mathbf{w}^{(t)}_{\setminus(a,\, b)},\, 0)$ 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 $\mathbf{w}^{(t)}_{\setminus(a,\, b)}$ 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

$$\frac{1}{1 + \frac{(1-\theta_{ab})\, 2 \cosh\left(\sum_{i \in V_d} \mu_i\right) \cosh\left(\sum_{i \in V_e} \mu_i\right)}{\theta_{ab} \cosh\left(\sum_{i \in V_c} \mu_i\right)}}. \tag{9}$$

The above formulation makes use of the clustering weights as in Eq. (8). Observe that the full-conditional in Eq. (9) 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 Eq. (9) becomes

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