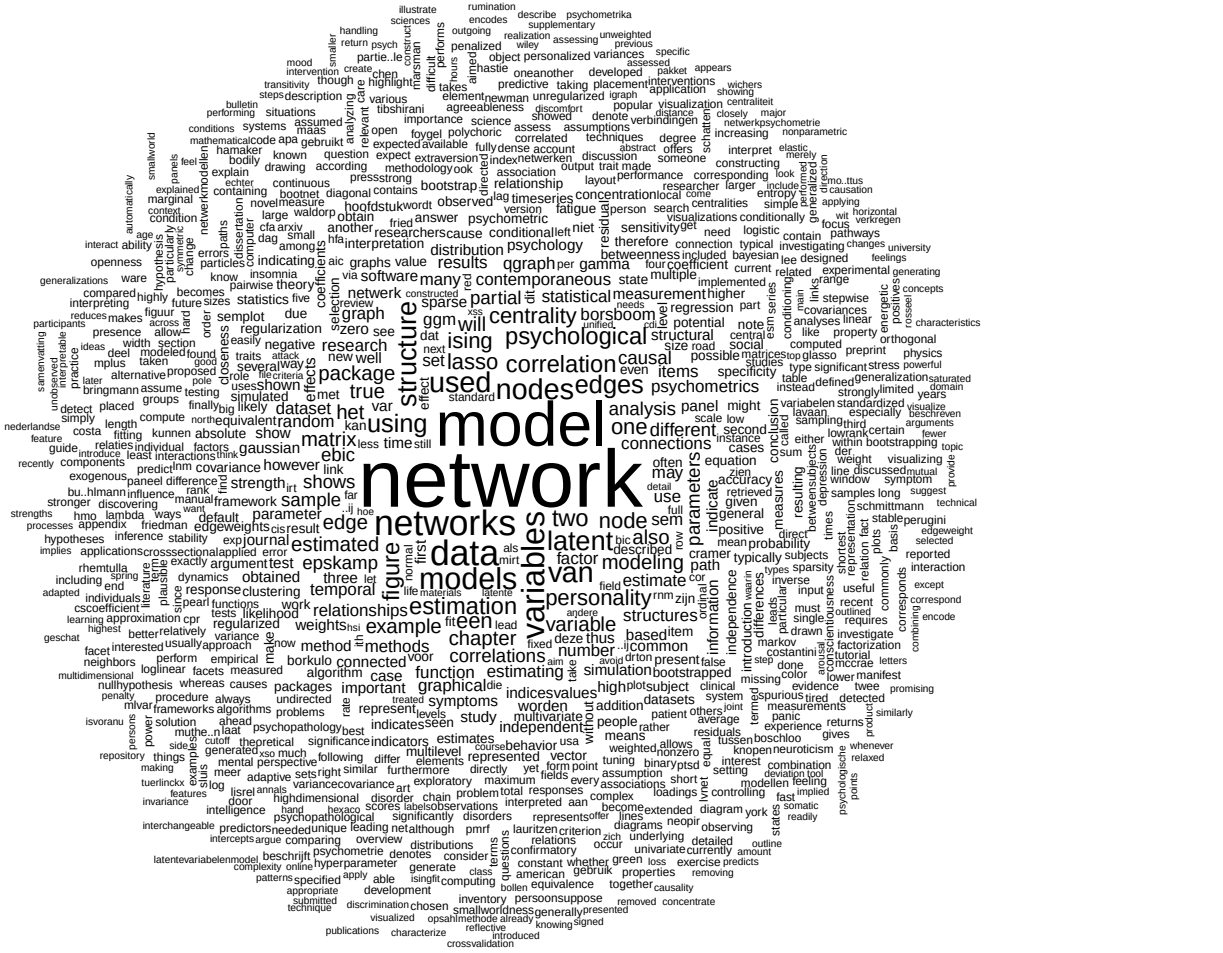
Network Psychometrics

PhD Dissertation

Sacha Epskamp

24-10-2016


network
model
networks
variables
data
models
structure
figure
used
nodes
edges
psychological
centrality
lasso
correlation
estimation
latent
van
personality
variable
chapter
using
one
different
connections
parameters
estimated
epskamp
relationships
modeling
path
factor
two
node
also
partial
statistical
measurement
regression
psychometrics
analysis
panel
might
ising
will
graph
sample
edge
matrix
epic
shows
package
research
regularization
correlations
function
estimating
simulation
bootstrapped
graphical
important
indicies
values
high
symptoms
worden
independent
multivariate
network
psychopathology
mental
perspective
variancecovariance
exploratory
maximum
fields
every
associations
complex
loadings
controlling
psychopathological
disorders
represents
lauritzen
criterion
diagrams
neopir
observing
underlying
detailed
outline
relations
occur
confirmatory
univariate
currently
amount
constant
whether
green
loss
exercise
predicts
removing
gebruik
properties
equivalence
together
causality
bollen
computing
generate
hyperparameter
complexity
online
patterns
specified
apply
able
isingfit
appropriate
development
submitted
technique
discrimination
chosen
smallworldness
generally
presented
inventory
persoonsuppose
removed
concentrate
visualized
opsahlmethode
already
knowing
signed
publications
characterize
reflective
introduced
crossvalidation

Contents

i

ii

iii

iv

Outline

Chapter 1

Introduction: Psychological Networks

This chapter has been adapted from:

- Epskamp, S., Borsboom, D., and Fried, E.I. (2012) Estimating psychological networks and their accuracy: a tutorial paper. *arXiv preprint*, arXiv:1604.08462.

Chapter 1

Introduction: Psychological Networks

1.1 Introduction

There are over 7 billion people in the world, each with a different brain containing 15 to 33 billion neurons. These people are intelligent entities who develop and change over time and who interact with each other in complicated social structures. Consequently, human behavior is likely to be complex. In recent years, research on dynamical systems in psychology has emerged, which is analogous to other fields such as biology and physics. One popular and promising line of research involves the modeling of psychological systems as causal systems or networks of cellular automata (Van Der Maas et al., 2006; Borsboom, 2008; Cramer, Waldorp, van der Maas, & Borsboom, 2010; Borsboom, Cramer, Schmittmann, Epskamp, & Waldorp, 2011). The general hypothesis is that noticeable macroscopic behavior—the co-occurrence of aspects of psychology such as cognitive abilities, psychopathological symptoms, or behavior—is not due to the influence of unobserved common causes, such as general intelligence, psychopathological disorders, or personality traits, but rather to emergent behavior in a network of interacting psychological, sociological, biological, and other components. I will term such networks *psychological networks*

Parts of this chapter have been adapted from: Epskamp, S., Borsboom, D., and Fried, E.I. (2012) Estimating psychological networks and their accuracy: a tutorial paper. *arXiv preprint*, arXiv:1604.08462.

to distinguish these models from other networks used in psychology, such as social networks and neural networks.

Figure 1.1 shows an example of such a psychological network, estimated on the `bfi` dataset from the *psych* R package (Revelle, 2010). This dataset contains the responses of 2,800 people on 25 items designed to measure the Big Five personality traits (McCrae & Costa, 1997). The network shows many meaningful connections, such as "make friends easily" being linked to "make people feel at ease," "don't talk a lot" being linked to "find it difficult to approach others," and "carry the conversation to a higher level" being linked to "know how to captivate people." A more detailed description on how to interpret such networks is included below and is discussed further (with respect to personality) in Chapter 10. Psychological networks are strikingly different to network models typically used in complexity research. These are networks between variables (nodes can take one of multiple states) rather than between concrete entities, such as people, computers, cities, and so forth. Furthermore, we do not know the structure of psychological networks. Due to the nascent status of this field, we do not yet know basic properties of psychological networks such as connectivity and clustering, which forces researchers to estimate network structures from data.

This dissertation addresses the problem of how to estimate network models from psychological data, and how such models should consequently be analyzed and interpreted. This is the first dissertation fully devoted to estimating and interpreting psychological networks. Some of the methods discussed (most notably the *qgraph* package; Chapter 9) have already grown to be commonly used psychological research. As a result, the number of researchers working on the estimation of psychological networks has grown substantively during the course of this PhD project. We can now speak of a new field of research: *network psychometrics*. The goal of this dissertation is to start of researchers in this field on the right foot.

Because most chapters of this dissertation utilize a certain class of network models, pairwise Markov random fields (PMRF; Lauritzen, 1996; Murphy, 2012), this introduction is followed by a general introduction to interpreting such models (adapted from Epskamp, Borsboom, & Fried, 2016, which is further used in Chapter 3). The introduction will conclude with a general outline of the dissertation.

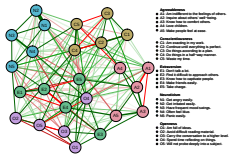


Figure 1.1: Example of a network model estimated the BFI personality dataset from the psych package in R. Nodes represent variables (in this case, personality inventory items) and links between the nodes (also termed *edges*) represent partial correlation coefficients. Green edges indicate positive partial correlations, red edges indicate negative partial correlations, and the width and saturation of an edge corresponds to the absolute value of the partial correlation. Estimation technique as outlined in Chapter 2 was used.

1.2 Psychological Networks

A psychological network is a model in which *nodes* represent observed psychological variables, usually psychometric test items such as responses to questions about whether a person suffered from insomnia or fatigue in past weeks. These nodes are connected by *edges* which indicate some statistical relationship between them. These models are conceptually different from commonly used reflective latent variable models that explain the co-occurrence among symptoms (e.g., the fact that individuals often suffer from sadness, insomnia, fatigue, and concentration problems at the same time) by invoking an underlying unobserved latent trait (e.g., depression) as the common cause of all the symptoms. Psychological networks offer a different conceptual interpretation of the data and explain such co-occurrences via direct relationships between symptoms; for example, someone who sleeps poorly will be tired, and someone who is tired will not concentrate well (Fried et al., 2015; Schmittmann et al., 2013). Such relationships can then be more easily interpreted when drawn as a network structure where edges indicate pathways on which nodes can affect each other. The edges can differ in strength of connection, also termed *edge weight*, indicating if a relationship is strong (commonly visualized with thick edges) or weak (thin, less saturated edges) and positive (green edges) or negative (red edges). After a network structure is estimated, the visualization of the graph itself tells the researcher a detailed story of the multivariate dependencies in the data. Additionally, many inference methods from graph theory can be used to assess which nodes are the most important in the network, termed the *most central nodes*.

Directed and Undirected Networks

In general, there are two types of edges that can be present in a network: an edge can be directed, in which case one head of the edge has an arrowhead indicating a one-way effect, or an edge can be undirected, indicating some mutual relationship. A network that contains only directed edges is termed a *directed network*, whereas a network that contains only undirected edges is termed an *undirected network* (Newman, 2010). Many fields of science consider directed networks interesting because they can be used to encode causal structures (Pearl, 2000). For example, the edge insomnia → fatigue can be taken to indicate that insomnia causes fatigue. The work of Pearl describes that such causal structures can be tested using only observational cross-sectional data and can even be estimated to a certain extent (Kalisch, Mächler, Colombo, Maathuis, & Bühlmann, 2012; Scutari, 2010). However, when temporal information is lacking, there is only limited information present in cross-sectional observational data. Such estimation methods typically only work under two very strict assumptions (a) that all entities which play a causal role have been measured and (b) that the causal chain of cause and effect is not cyclic (i.e., a variable cannot cause itself via any path). Both assumptions are not very plausible in psychological systems. Furthermore, such directed networks suffer from the problem that many equivalent models can exist that feature the same relationships found in the data (MacCallum, Wegener, Uchino, & Fabrigar, 1993); this makes the interpretation of structures difficult. For example, the structure insomnia → fatigue → concentration is statistically equivalent to the structure insomnia ← fatigue → concentration as well as the structure insomnia ← fatigue ← concentration: All three only indicate that insomnia and concentration problems are conditionally independent after controlling for fatigue.

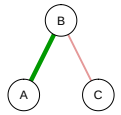


Figure 1.2: Example of a pairwise Markov random field. Node A positively interacts with node B, and node B negatively interacts with node C. Nodes A and C are *conditionally independent* given node B.

For the reasons outlined above, psychological networks estimated on cross-sectional data are typically undirected networks. The current state-of-the-art method for estimating undirected psychological network structures involves the estimation of PMRFs. A PMRF is a network model in which edges indicate the full conditional association between two nodes after conditioning on all other nodes in the network. This means when two nodes are connected, there is a relationship between these two nodes that cannot be explained by any other node in the network. Simplified, it can be understood as a partial correlation controlling for all other connections. The absence of an edge between two nodes indicates that these nodes are conditionally independent of each other given the other nodes in the network. Thus, a completely equivalent undirected structure (compared to the structures described above) would be insomnia — fatigue — concentration, indicating that insomnia and concentration problems are conditionally independent after controlling for fatigue.

Figure 1.2 shows a PMRF similar to the example described above. In

this network, there is a positive relationship between insomnia and fatigue and a negative relationship between fatigue and concentration. The positive edge is thicker and more saturated than the negative edge, indicating that this interaction effect is stronger than that of the negative edge. This network shows that insomnia and concentration do not directly interact with each other in any way other than through their common connection with fatigue. Therefore, fatigue is the most important node in this network—a concept we will later quantify as *centrality*. These edges can be interpreted in several different ways.[1] First, as shown above, the model is in line with causal interpretations of associations among the symptoms. Second, this model implies that insomnia and fatigue predict each other after controlling for concentration; even when we know someone is concentrating poorly, that person is more likely to suffer from insomnia when we observe that person suffering from fatigue. Similarly, fatigue and concentration predict each other after controlling for insomnia. After controlling for fatigue, there is no longer any predictive quality between insomnia and concentration, even though these variables are correlated; fatigue now mediates the prediction between these two symptoms. Finally, these edges can represent genuine symmetric causal interactions between symptoms (e.g., in statistical physics, a PRMF called the Ising model is used to model particles that cause neighboring particles to be aligned).

Network Inference

In the first step of network analysis, the obtained network is typically presented graphically to show the structure of the data. Afterwards, inference methods derived from graph theory can be applied to the network structure. The estimated PRMF is always a weighted network, which means that we not only look at the structure of the network (e.g., are two nodes connected or not) but also at the strength of connection between pairs of nodes. Because of this, many typical inference methods that concern the global structure of the network (e.g., small-worldness, density, and global clustering; Kolaczyk, 2009; Newman, 2010; Watts & Strogatz, 1998) are less useful in the context of psychological networks because they only take into account whether nodes are connected or not and not the strength of association among nodes. Because the global inference methods for weighted networks and PRMFs are still in development and no consensus has been reached, the network inference section focuses on local network properties: How are two nodes related, and what is the influence of a single node?

[1] A more detailed description of the interpretation of such models can be found in Chapter 6 and Chapter 8.

Relationship between two nodes. The relationship between two nodes can be assessed in two ways. First, we can directly assess the edge weight. This is always a number that is nonzero because an edge weight of zero would indicate there is no edge. The sign of the edge weight (positive or negative) indicates the type of interaction, and the absolute value of the edge weight indicates the strength of the effect. For example, a positive edge weight of 0.5 is equal in strength to a negative edge weight of −0.5 and both are stronger than an edge weight of 0.2. Two strongly connected nodes influence each other more easily than two weakly connected nodes. This is similar to how two persons standing closer to each other can communicate more easily (via talking) than two people standing far away from each other (via shouting)—two strongly connected nodes are closer to each other. As such, the length of an edge is defined as the inverse of the edge strength. Finally, the distance between two nodes is equal to the sum of the lengths of all edges on the shortest path between two nodes (Newman, 2010).

Node centrality. The importance of individual nodes in the network can be assessed by investigating the node centrality. A visualization of a network, such as the one shown in Figure 1.2, is an abstract rendition of a high-dimensional space in two dimensions. Although visualizations of network models often aim to place highly connected nodes into the center of the graph, for instance using the Fruchterman-Reingold algorithm (Fruchterman & Reingold, 1991), the two-dimensional visualization cannot properly reflect the true space of the model. Thus, the metric distance between the placement of nodes in the two-dimensional space has no direct interpretation as it has in multidimensional scaling, for instance. Therefore, graph theory has developed several methods to more objectively quantify which node is most central in a network. Three such centrality measures have appropriate weighted generalizations that can be used with psychological networks (Opsahl, Agneessens, & Skvoretz, 2010). First, node strength, also called degree in unweighted networks (Newman, 2010), simply adds up the strength of all connected edges to a node; if the network is made up of partial correlation coefficients, the node strength equals the sum of absolute partial correlation coefficients between a node and all other nodes. Second, closeness takes the inverse of the sum of all the shortest paths between one node and all other nodes in the network. Thus, where node strength investigates how strongly a node is directly connected to other nodes in the network, closeness investigates how strongly a node is indirectly connected to other nodes in the network. Finally, betweenness looks at how many of

the shortest paths between two nodes go through the node in question; the higher the betweenness, the more important a node is in connecting other nodes.

1.3 Outline

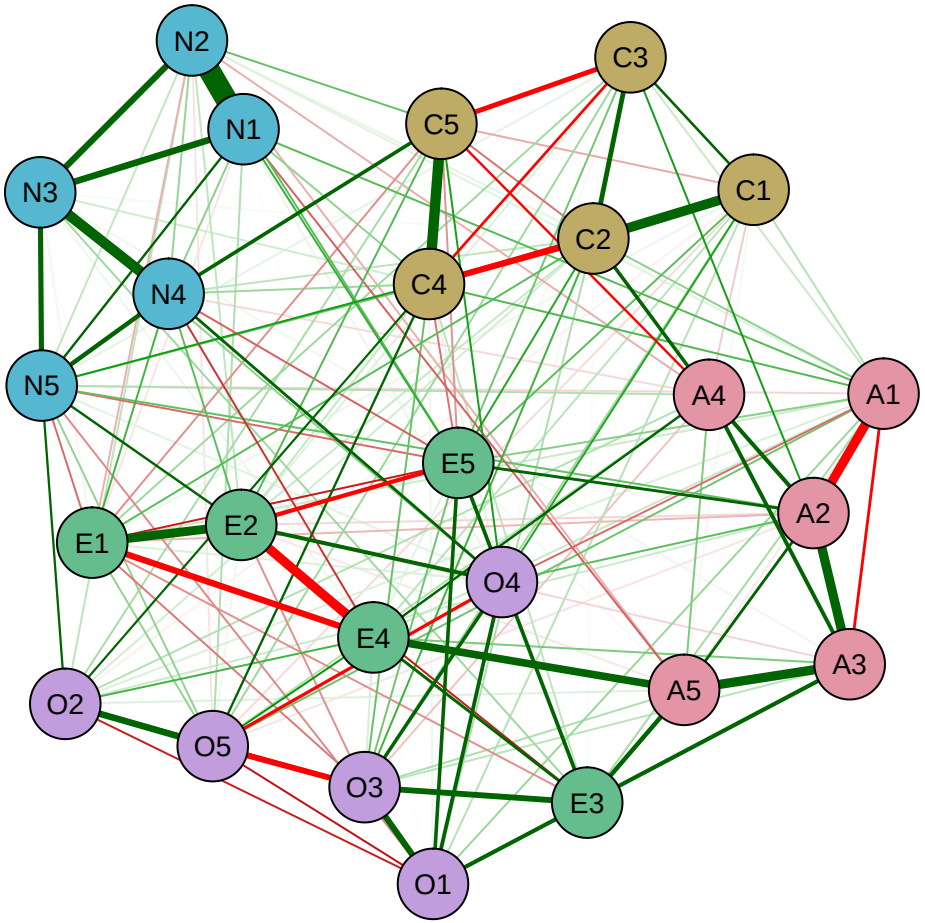
The above describes, in short, the basis of the methodology introduced in this dissertation. Although the methods and software discussed in this dissertation are applicable to many fields both in and outside of psychology[2], it is in three fields where the method has been most applied: clinical psychology, psychometrics, and personality research. As a result of this diversity, the audience of contributions in network psychometrics varies as well between empirical researchers without much programming knowledge, researchers with strong technical skills, and researchers that are familiar with programming in R. To this end, the dissertation is split in three parts: Part I is aimed at empirical researchers with an emphasis on clinical psychology, Part II is aimed at technical researchers with an emphasis on psychometrics, and Part III is aimed at R users with an emphasis on personality research.

Network models have gained extensive footing in clinical psychology for their ability to highlight the dynamics that may lead to someone developing or maintaining mental illness (Cramer et al., 2010; Borsboom & Cramer, 2013; Borsboom, in press). For this reason, Part I of this dissertation is aimed at empirical researchers, with an emphasis on clinical psychology. Chapter 2 will introduce the reader to the network estimation technique we now most often employ on cross-sectional data: model selection on a series of regularized networks. Chapter 3 continues on this topic and presents methods for assessing the accuracy of the network estimated as well as the stability of inferences made on the network structure after observing only subsets of people. Chapter 4 continuous discussing network estimation using regularization and provides a more critical note to argue that results from the methods described in this part are not without problems and certain conclusions cannot be drawn. Finally, Chapter 5 gives a conceptual introduction to time-series modeling on the data of a single patient in clinical practice.

[2] In recent literature, the methodology has been applied to diverse fields, such as attitude formation (Dalege et al., 2016), test validity (Ziegler, Booth, & Bensch, 2013), dog personality (Goold, Vas, Olsen, & Newberry, 2015), plant breeding (da Silva, Cecon, & Puiatti, 2015; Silva, Rigo, Possa, & Rigo, 2016), and ecology (Wilkins, Shiroka, Joseph, Hubbard, & Safran, 2015; Goell, Ozkundakci, Hilbert, & Adrian, 2016).

Part II, on the other hand, is aimed at psychometricians and researchers with strong technical skills. Simply stated, these chapters contain equations. Chapter 6 will further outline the modeling of time-series data. This chapter will give the theoretical justification for the model described in Chapter 5 and extends this modeling framework to situations where time-series data is available from multiple people. Rather than contrasting network modeling to the latent variable modeling of classical psychometrics, Chapter 7 and Chapter 8 outline how these frameworks can be combined. Chapter 7 will introduce the network model for multivariate normal data as a formal psychometric model and will cast this model in the well-known structural equation modeling (SEM) framework. Combining network models into SEM allows for two new modeling frameworks, both of which show promise in psychometrics. Finally, Chapter 8 will outline the relationship between the best known network model for binary data, the Ising model, to well-known psychometric modeling frameworks such as logistic regression, loglinear modeling, and item-response theory.

Part III continues the combination of using networks to augment classical psychometrics in three tutorial papers aimed at empirical researchers familiar with programming in R. Two of these chapters are aimed at the third field in which network modeling has grown prominent: personality research (Cramer, Sluis, et al., 2012). These chapters are not technical and are all tutorials using R. Chapter 9 will outline how network-based visualizations can be used to gain insight into high-dimensional correlational structures as well as large exploratory factor analysis results. Chapter 10 introduces the network perspective to personality researchers. Both Chapter 9 and Chapter 10 will show the use of these methods in personality inventory datasets. Finally, Chapter 11 will introduce an R package that can be used to draw path diagrams and to visualize correlational structures directly from SEM software output.



Agreeableness
- A1: Am indifferent to the feelings of others.
- A2: Inquire about others' well−being.
- A3: Know how to comfort others.
- A4: Love children.
- A5: Make people feel at ease.

Conscientiousness
- C1: Am exacting in my work.
- C2: Continue until everything is perfect.
- C3: Do things according to a plan.
- C4: Do things in a half−way manner.
- C5: Waste my time.

Extraversion
- E1: Don't talk a lot.
- E2: Find it difficult to approach others.
- E3: Know how to captivate people.
- E4: Make friends easily.
- E5: Take charge.

Neuroticism
- N1: Get angry easily.
- N2: Get irritated easily.
- N3: Have frequent mood swings.
- N4: Often feel blue.
- N5: Panic easily.

Openness
- O1: Am full of ideas.
- O2: Avoid difficult reading material.
- O3: Carry the conversation to a higher level.
- O4: Spend time reflecting on things.
- O5: Will not probe deeply into a subject.

Part I

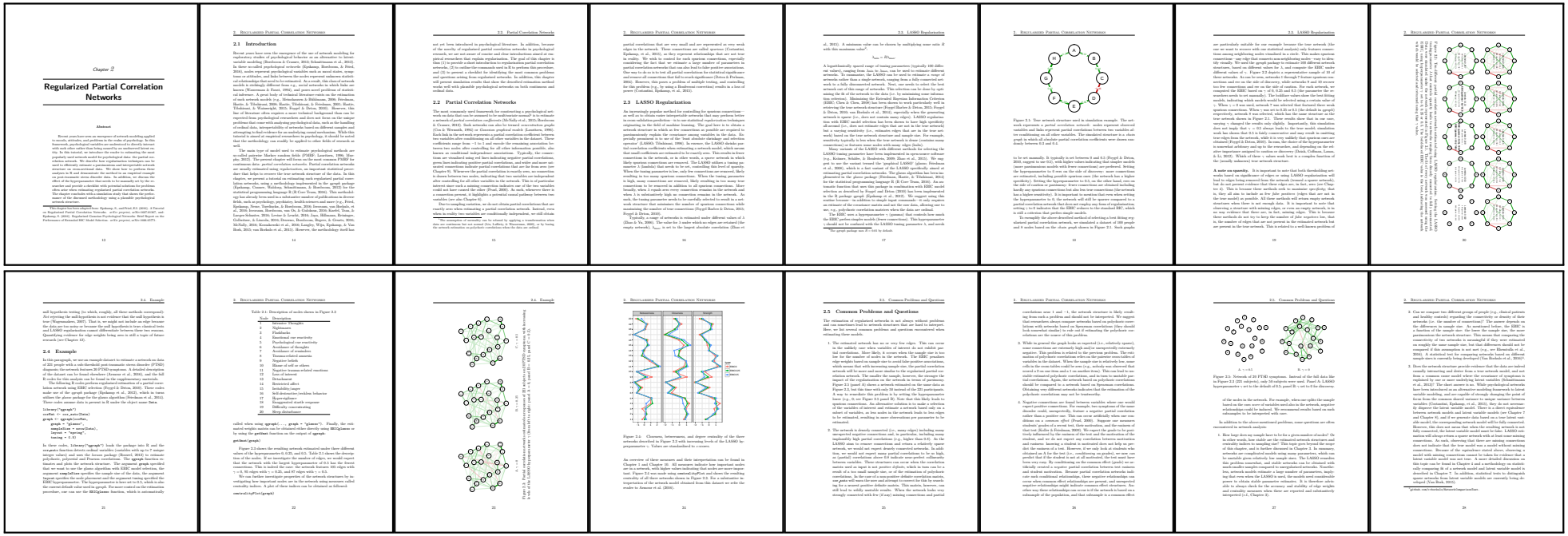
A Practical Guide to Network Psychometrics

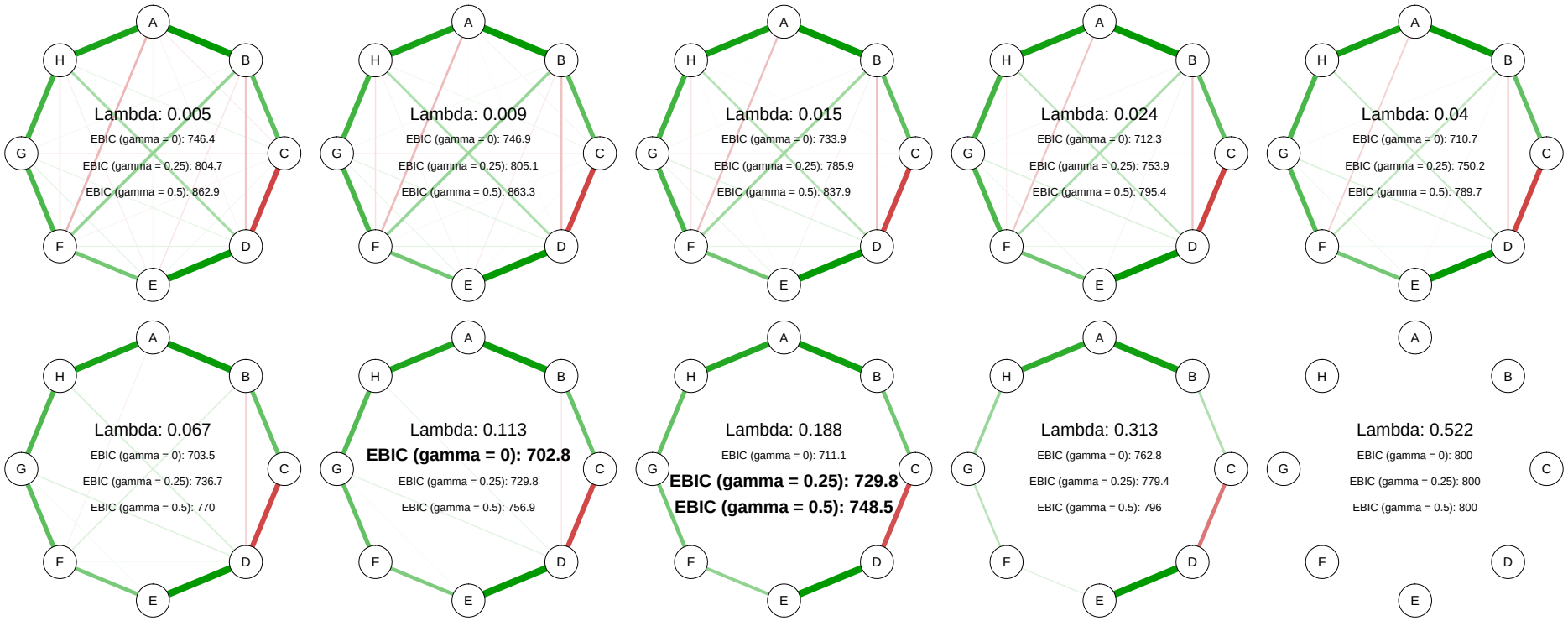
Chapter 2

Regularized Partial Correlation Networks

This chapter has been adapted from:

- Epskamp, S., and Fried, E.I. (2016). A Tutorial on Regularized Partial Correlation Networks. *arXiv preprint*, arXiv:1607.01367.
- Epskamp, S. (2016). Regularized Gaussian Psychological Networks: Brief Report on the Performance of Extended BIC Model Selection. *arXiv preprint*, arXiv:1606.05771.



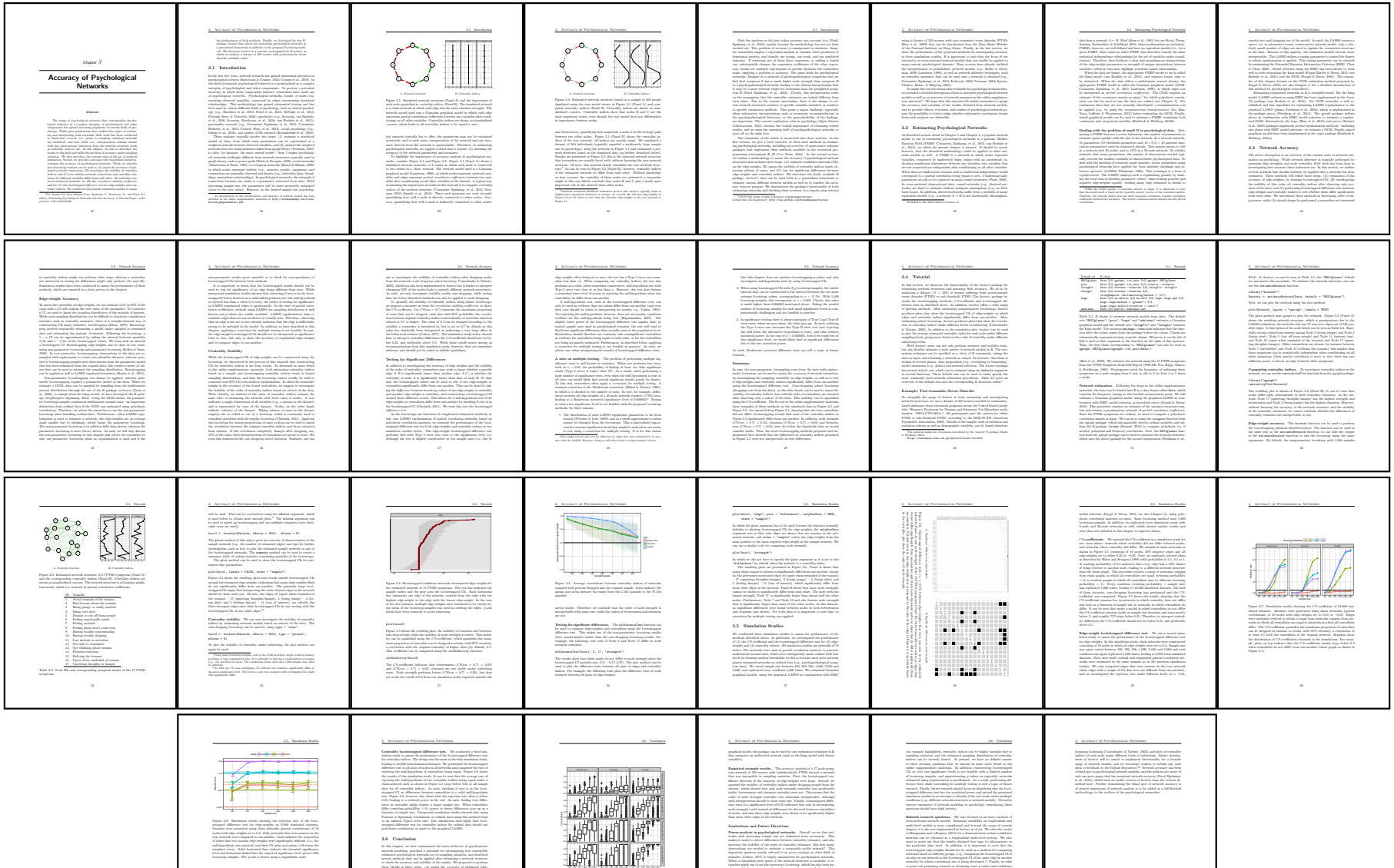

Lambda: 0.005
EBIC (gamma = 0): 746.4
EBIC (gamma = 0.25): 804.7
EBIC (gamma = 0.5): 862.9
Lambda: 0.009
EBIC (gamma = 0): 746.9
EBIC (gamma = 0.25): 805.1
EBIC (gamma = 0.5): 863.3
Lambda: 0.015
EBIC (gamma = 0): 733.9
EBIC (gamma = 0.25): 785.9
EBIC (gamma = 0.5): 837.9
Lambda: 0.024
EBIC (gamma = 0): 712.3
EBIC (gamma = 0.25): 753.9
EBIC (gamma = 0.5): 795.4
Lambda: 0.04
EBIC (gamma = 0): 710.7
EBIC (gamma = 0.25): 750.2
EBIC (gamma = 0.5): 789.7
Lambda: 0.067
EBIC (gamma = 0): 703.5
EBIC (gamma = 0.25): 736.7
EBIC (gamma = 0.5): 770
Lambda: 0.113
EBIC (gamma = 0): 702.8
EBIC (gamma = 0.25): 729.8
EBIC (gamma = 0.5): 756.9
Lambda: 0.188
EBIC (gamma = 0): 711.1
EBIC (gamma = 0.25): 729.8
EBIC (gamma = 0.5): 748.5
Lambda: 0.313
EBIC (gamma = 0): 762.8
EBIC (gamma = 0.25): 779.4
EBIC (gamma = 0.5): 796
Lambda: 0.522
EBIC (gamma = 0): 800
EBIC (gamma = 0.25): 800
EBIC (gamma = 0.5): 800
A
B
C
D
E
F
G
H

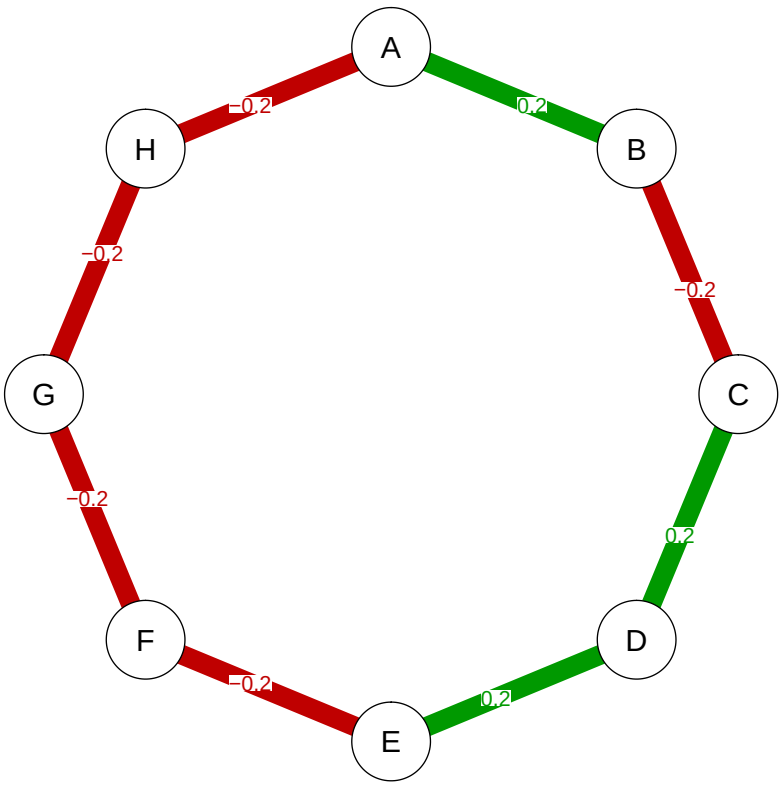
Chapter 3

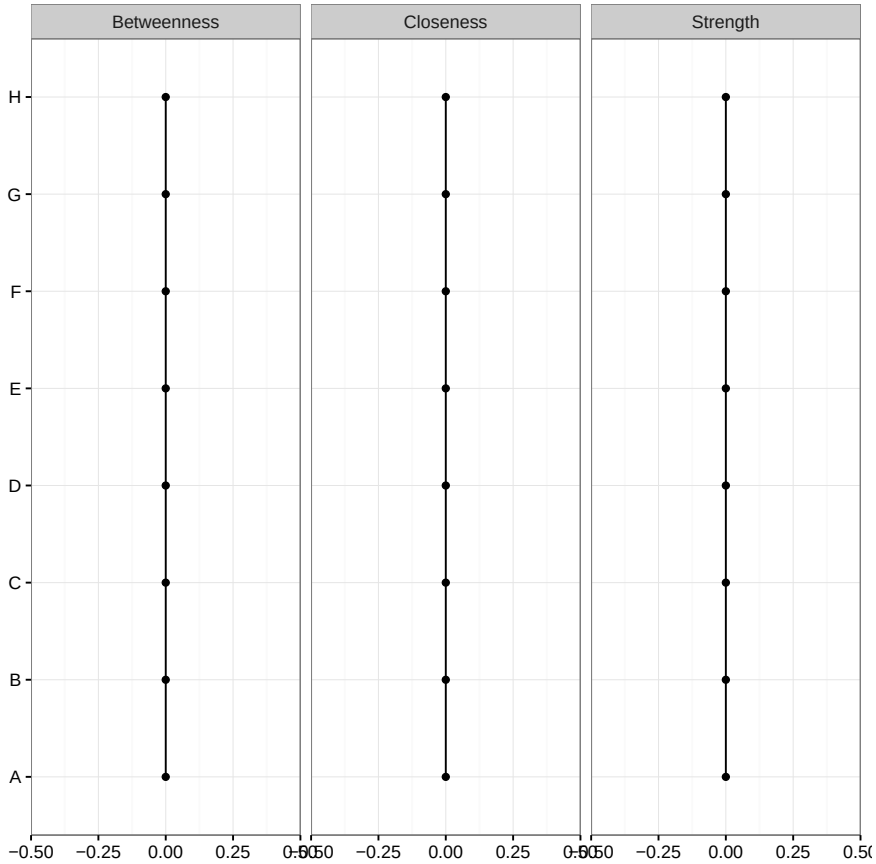
Accuracy of Psychological Networks

This chapter has been adapted from:

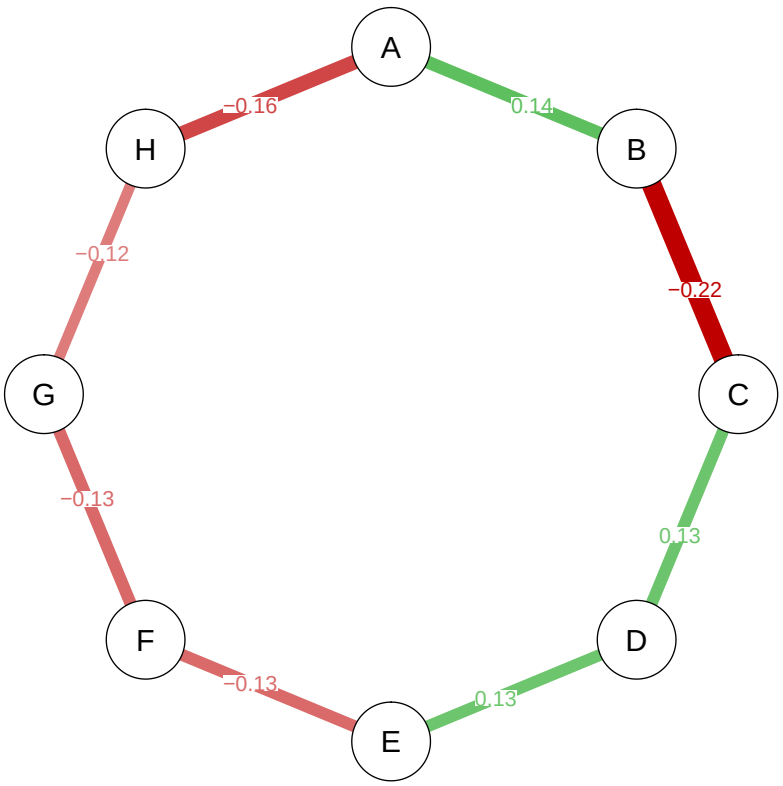
- Epskamp, S., Borsboom, D., and Fried, E.I. (2016). Estimating Psychological Networks and their Accuracy: A Tutorial Paper. *arXiv preprint*, arXiv:1604.08462.

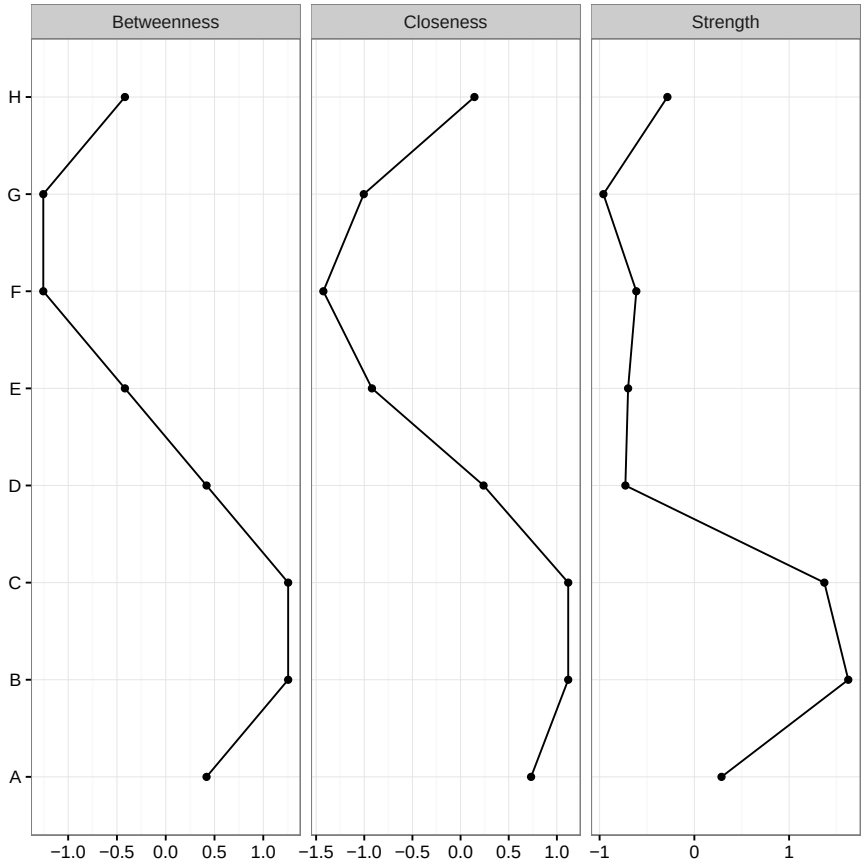






- True network and centrality indices
- All edge-weights and centralities are equal





- Estimated network and centrality indices
- How to judge if these differences are interpretable?

Chapter 4

Network Estimation and Sparsity

This chapter has been adapted from:

- Epskamp, S., Kruis, J., and Marsman, M. (2016). Estimating psychopathological networks: be careful what you wish for. *arXiv preprint*, arXiv:1604.08045.

Chapter 4

Network Estimation and Sparsity

Abstract

Network models, in which psychopathological disorders are conceptualized as a complex interplay of psychological and biological components, have become increasingly popular in the recent psychopathological literature (Borsboom et al., 2011). These network models often contain significant numbers of unknown parameters, yet the sample sizes available in psychological research are limited. As such, general assumptions about the true network are introduced to reduce the number of free parameters. Incorporating these assumptions, however, means that the resulting network will always reflect the particular structure assumed by the estimation method—a crucial and often ignored aspect of psychopathological networks. For example, observing a sparse structure and simultaneously assuming a sparse structure does not imply that the true model is, in fact, sparse. To illustrate this point, we discuss a recently published paper that reveals a high-dimensional network of psychopathological symptoms (Boschloo et al., 2015). Furthermore, we show the effect of the assumption of sparsity in three simulation studies.

This chapter has been adapted from: Epskamp, S., Kruis, J., and Marsman, M. (2016). Estimating psychopathological networks: be careful what you wish for. *arXiv preprint*, arXiv:1604.08845.

4.1 Introduction

Recent psychological literature has focused on a network approach to model many different psychological phenomena (Schmittmann et al., 2013). Such networks can be high-dimensional structures (i.e., the number of unknown parameters is much larger than the available data), which are hard to estimate without making general assumptions about the underlying true model structure. If the true model is assumed to be sparse, thus containing a small number of connections relative to the number of possible connections, a methodology can be applied that will return such a sparse network structure. In other words, assuming a sparse network structure results in estimating a sparse network structure, which means that certain conclusions cannot be drawn from observing such a structure. In this chapter, we argue that care should be taken in interpreting the obtained network structure because the estimation procedure may pollute the results. We will illustrate this by showing examples of networks obtained when sparse networks are estimated even when the true network structure is dense.

4.2 Network Psychometrics

The network approach has been particularly promising in the field of psychopathology. Within this framework, symptoms (e.g., insomnia, fatigue, and concentration problems) are no longer treated as interchangeable indicators of some latent mental disorder (e.g., depression). Instead, symptoms play an active causal role. For example, insomnia leads to fatigue, fatigue leads to concentration problems, and so forth (Borsboom & Cramer, 2013). Psychopathological disorders, then, are not interpreted as the common cause of observed symptoms but rather as emergent behaviors that result from a complex interplay of psychological and biological components. To group such a complex structure, a network model can be used in which variables such as symptoms or moods are represented by nodes. Nodes are connected by edges that indicate associations between nodes. This line of research has led to intuitive new insights about various psychopathological concepts such as comorbidity (Borsboom et al., 2011; Cramer et al., 2010), the impact of life events (Cramer, Borsboom, Aggen, & Kendler, 2012; Fried et al., 2015), and sudden life transitions (e.g., sudden onset of a depressive episode; van de Leemput et al., 2014; Wichers, Groot, Psychosystems, ESM Group, & EWS Group, 2016).

The growing popularity of the network perspective on psychological phenomena has culminated in the emergence of a new branch of psychology

dedicated to the estimation of network structures on psychological data—network psychometrics (Epskamp, Maris, Waldorp, & Borsboom, in press). This field focuses on tackling the problem of estimating network structures involving large numbers of parameters in high-dimensional models. When cross-sectional data are analyzed, the most popular models that are used are the Gaussian Graphical Model (GGM; Lauritzen, 1996) for continuous data and the Ising model (Ising, 1925) for binary data. Both the GGM and the Ising model fall under a general class of models called *Markov Random Fields*. These models represent variables as nodes which are connected by edges but only if the variables are *conditionally independent*. The strength of an edge (i.e., its absolute deviance from zero) demonstrates the strength of the association between two variables after conditioning on all other variables in the network; this is also termed *concentration* (Cox & Wermuth, 1993). In the GGM, edges directly correspond to partial correlation coefficients. The Ising model does not allow for such standardization, but edge weights can be similarly interpreted. A more detailed conceptual introduction of these models can be found in Chapter 1. A technical introduction to the GGM will follow in Chapter 6, and a technical introduction to the Ising Model will follow in Chapter 8.

In both models, we must estimate a weight matrix that contains $P(P-1)/2$ number of parameters, where P is the number of nodes, in order to encode the network structure. These parameters encompass the conditional relationship between two nodes after conditioning on all other nodes in the network and can be shown to be quite unstable with relatively low sample sizes (see Chapter 3). "Relatively low sample sizes," is a loose description and has not yet been well-defined. A general rule would be to have at least as many observations as the number of parameters. But, as will be shown later, this general rule still results in unstable estimates. A common solution to overcome the problem of estimating many parameters is to reduce this number by using some form of regularization or penalization. A particularly promising technique is to apply the 'least absolute shrinkage and selection operator' (LASSO; Tibshirani, 1996) to the edge weights of the network. The LASSO penalizes the sum of absolute parameter values such that the estimated values shrink to zero. That is, the absolute parameter estimates will be small and will often equal exactly zero. Therefore, the resulting model is almost always sparse; only a relatively few number of parameters will be estimated to be nonzero. The use of LASSO typically leads to better performance in cross-validations (i.e., overfitting is prevented) and results in more easily interpretable models compared to nonregularized Ising models. Most important is that if the true network structure is sparse, the LASSO performs well in estimating this network

structure and, more specifically, in estimating fewer edges to be nonzero that are actually zero in the true network (i.e., fewer false positives).

The LASSO uses a tuning parameter that controls the sparsity, which can be chosen to minimize some criterion such as the Extended Bayesian Information Criterion (EBIC; Chen & Chen, 2008). This methodology has been shown to work well for both the GGM (Foygel & Drton, 2010) and the Ising model (Foygel Barber & Drton, 2015; van Borkulo et al., 2014), has been implemented in easy-to-use software (e.g., Epskamp et al., 2012; van Borkulo & Epskamp, 2014), and has been utilized in an increasing number of publications (e.g., Dalege et al., 2016; Isvoranu, van Borkulo, et al., 2016; Kossakowski et al., 2016; Langley et al., 2015; Rhemtulla et al., 2016; van Borkulo et al., 2015). For a more thorough introduction to this methodology, we recommend reading van Borkulo et al. (2014) and Chapter 2. To date, the largest application of a psychological network estimated using the LASSO was carried out by Boschloo et al. (2015), who measured 120 psychiatric symptoms in 34,653 subjects and modeled these with an Ising model. We use their work in this chapter to illustrate our concerns regarding the interpretation of network structures that are the result of applying network methodology to data.

4.3 A Sparse Network Model of Psychopathology

The network of Boschloo et al. (2015) shows a network structure in which symptoms representative of a disorder strongly cluster together. Although Boschloo and her colleagues admitted that the found network structure closely represents the structure that is imposed by the Diagnostic and Statistical Manual of Mental Disorders (DSM; American Psychiatric Association, 2013), they concluded that the found structure indicates that symptoms are not interchangeable, which is presumed in the DSM. Commonly, a DSM diagnosis requires an individual to have X out of Y symptoms, regardless of which specific symptoms. This means that two people with vastly different symptoms can be assigned the same diagnosis. This interchangeability results from an underlying causal notion of unobserved diseases causing symptoms rather than symptoms having an active causal role on each other—a notion more formally known as the common cause model (Schmittmann et al., 2013). Boschloo and her colleagues concluded that the network structure shows that symptoms are not interchangeable, mainly due to apparent differences in the number of connections and the strength of the connections between symptoms (a relative small number

of pathways between disorders and the presence of some negative connections).

Although we do not necessarily disagree with the notion that symptoms play an active causal role in psychopathology, we wish to point out that the conclusion that symptoms are not interchangeable is difficult to ascertain from a sparse approximated network structure alone. This is because the LASSO relies on the assumption that the true network structure is sparse; the LASSO will always search for a model in which relatively few edges and paths explain the co-occurrence of all nodes. As a result, the LASSO can have a low sensitivity (i.e., not all true edges are detected) but always has a high specificity (i.e., fewer false positives; van Borkulo et al., 2014). It is this reason why network analysts prefer the LASSO; edges that are estimated by the LASSO are likely to represent true edges. Moreover, the LASSO returns a parsimonious explanation of the data using only a few connections that can be interpreted as causal pathways (Lauritzen, 1996; Pearl, 2000). That the LASSO yields a possible explanation, however, does not mean that the LASSO provides the only explanation, nor does it indicate that other explanations are false. In the case of Boschloo et al., the sparse explanation found by the LASSO can give great insight regarding a possible way in which psychopathological symptoms interact with each other. However, merely finding a sparse structure does not mean that other explanations (e.g., a latent variable model with interchangeable symptoms) are disproved. Simply stated, using the LASSO always returns a sparse structure, that is what the LASSO does.

4.4 The Bet on Sparsity

The LASSO is capable of retrieving the true underlying structure but only if that true structure is sparse. Any regularization method makes the assumption that the true structure can be simplified in some way (e.g., is sparse) because otherwise too many observations are needed to estimate the network structure. This principle has been termed the *bet on sparsity* (Hastie et al., 2001). But what if the truth is not sparse, but dense?

Such a case would precisely arise if the true model were a latent common cause model in which one or several latent variables contribute to scores on completely interchangeable indicators. This is a feasible alternative because the Ising model can be shown to be mathematically equivalent to a certain type of latent variable model: the multidimensional item response model (MIRT; Reckase, 2009), with posterior normal distributions on the latent traits (Marsman, Maris, Bechger, & Glas, 2015; see also Chapter 8).

The corresponding Ising model is a low-rank network that will often be dense (i.e., all possible edges are present). Intuitively, this makes sense because the Ising model parameterizes conditional dependencies between items after conditioning on all other items, and no two items can be made conditionally independent if the common cause model is true. A low-rank weighted network will show indicators of a latent variable as clusters of nodes that are all strongly connected with each other. Therefore, if a common cause model is the true origin of the co-occurrences in the dataset, the corresponding Ising model should show the indicators to cluster together. Then if LASSO regularization is used, the corresponding network would likely feature sparsity but the nodes would still be clustered together—much like the results shown by Boschloo et al.

It is this relationship between the Ising model and MIRT that has led researchers to estimate the Ising model using a different form of regularization, by estimating a low-rank approximation of the network structure (Marsman et al., 2015). Such a structure is strikingly different than the sparse structure returned by LASSO estimation. Whereas the LASSO will always yield many edge parameters to be exactly zero, a low-rank approximation generally estimates no edge to be exactly zero. Thus a low-rank approximation will typically yield a dense network. On the other hand, this dense network is highly constrained by the eigenvector structure, leading many edge parameters to be roughly equivalent to each other rather than compared to the strongly varying edge parameters LASSO estimation allows. For example, the data can always be recoded such that a Rank 1 approximation only has positive connections. These are key points that cannot be ignored when estimating a network structure. Regardless of the true network structure that underlies the data, the LASSO will always return a sparse network structure. Similarly, a low-rank approximation will always return a dense low-rank network structure. Both methods tackle the bet on sparsity in their own way—sparsity in the number of nonzero parameters or sparsity in the number of nonzero eigenvalues—and both can lose the bet.

4.5 Estimating an Ising Model When the Truth Is Dense

Here we illustrate the effect that the estimation procedure has on the resulting Ising model in two examples. First, we simulated 1,000 observations from the true models shown in Figure 4.1. The first model is called a Curie-Weiss model (Kac, 1966), which is fully connected and in which all edges

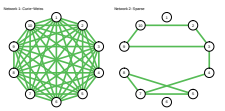


Figure 4.1: True network structures used in simulation study. The first network is a Curie-Weiss network: a fully connected network in which all edges have the same strength. The second network is a random sparse network. All edge weights are 0.2.

have the same strength (here set to 0.2). This network is a true Rank 1 network, which has been shown to be equivalent to a unidimensional Rasch model (Marsman et al., 2015). The Rasch model is a latent variable model in which all indicators are interchangeable. Figure 4.2 shows the results using three different estimation methods—sequential univariate logistic regressions for unregularized estimation (see Chapter 8), LASSO estimation using the *IsingFit*[1] R package (van Borkulo & Epskamp, 2014), and a Rank 2 approximation (Marsman et al., 2015)—on the first n number of rows in the simulated dataset. It can be seen that the unregularized estimation shows many spurious differences in edge strength, including many negative edges. The LASSO performs better but estimates a sparse model in which edge weights vary and in which many edges are estimated to be exactly zero. The Rank 2 approximation works best in capturing the model, which is not surprising because the true model is a Rank 1 network. At high sample sizes, all methods perform well in obtaining the true network structure.

The second model in Figure 4.1 corresponds to a sparse network in which 20% of randomly chosen edge strengths are set to 0.2 and in which the remaining edge strengths are set to 0 (indicating no edge). As Figure 4.3

[1] All LASSO analyses in this chapter made use the default setup of *IsingFit*, using a hyperparameter (γ) value of 0.25 as well as the AND-rule.

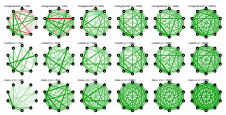


Figure 4.2: Examples of estimated network structures when the true network is a Curie-Weiss network, using different sample sizes and estimation methods. Graphs were drawn using the qgraph package without setting a maximum value (i.e., the strongest edge in each network has full saturation and width).

shows, the LASSO now performs very well in capturing the true underlying structure. Because both the unregularized estimation and the Rank 2 approximation estimate a dense network, they have a very poor specificity (i.e., many false-positive edges). In addition, the Rank 2 approximation retains spurious connections even at high sample sizes (choosing a higher rank will lead to a better estimation). Thus, this example serves to show that the LASSO and low-rank approximations only work well when the assumptions on the true underlying model are met. In particular, using a low-rank approximation when the truth is sparse will result in many false positives, whereas using a LASSO when the truth is dense will result in many false negatives. Even when the true model is one in which every node represents an interchangeable symptom, the LASSO would still return a model in which nodes could be interpreted to not be interchangeable.

For the second example, we simulated data under the latent variable model as shown in Figure 4.4, using an MIRT model (Reckase, 2009). In this model, the symptoms for dysthymia and generalized anxiety disorder (GAD) were taken from the supplementary materials of Boschloo et al. (2015), with the exception of the GAD symptom "sleep disturbance," which we split in two: insomnia and hypersomnia. The item discriminations of each symptom were set to 1, indicating that symptoms are interchangeable,

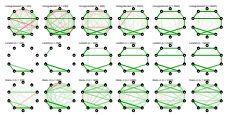


Figure 4.3: Examples of estimated network structures when the true network is sparse, using different sample sizes and estimation methods. Graphs were drawn using the qgraph package without setting a maximum value (i.e., the strongest edge in each network has full saturation and width).

and item difficulties were set to 0. All latent variables were simulated to be normally distributed with a standard deviation of 1, and the correlation between dysthymia and GAD was set to 0.55—similar to the empirically estimated comorbidity (Kessler, Chiu, Demler, & Walters, 2005). Nodes 2 and 3 in dysthymia and nodes 6 and 7 in GAD are mutually exclusive, which we modeled by adding orthogonal factors with slightly higher item discriminations of 1.1 and −1.1. Furthermore, nodes 7, 8, 9, and 10 of dysthymia are identical to nodes 6, 7, 8, and 9 of GAD respectively, which we modeled by adding orthogonal factors with item discriminations of 0.75. These nodes are typically not identical because a sleep structure is imposed in datasets such as the one analyzed by Boschloo et al. That is, if someone does not exhibit the symptom "low mood," that person is never asked about insomnia in the depression scale because he or she is assumed to not have this symptom. We did not impose a sleep structure to keep the simulation study simple. Such shared symptoms are termed *bridge symptoms* in network analysis because they are assumed to connect the clusters of disorders and explain comorbidity (Borsboom et al., 2011; Cramer et al., 2010). In sum, the model shown in Figure 4.4 generates data that are plausible given the latest disease conceptualization of psychopathology.

Figure 4.5 shows the simulated and recovered network structures. First we simulated 10 million observations from this model and estimated the

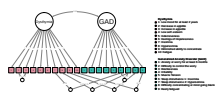


Figure 4.4: A multidimensional IRT model (MIRT) used in simulating data. All latent variables were normally distributed with standard deviation of 1 and all symptoms were binary. The edges in this model correspond to item discrimination parameters.

corresponding Ising model using nonregularized estimation by framing the Ising model as a log-linear model (Agresti, 1990, see Chapter 8) (the estimation was done using the *IsingSampler* package, Epskamp, 2014). Panel A shows the results, which give a good proxy of the true corresponding Ising structure. It can be seen that the true model is dense, meaning that indicators of the disorders cluster together. Two negative connections are formed between the mutually exclusive indicators, and bridging connections are formed between the shared indicators. Next, we simulated 1,000 observations from this model in Figure 4.4 and estimated the Ising model in various ways. Panel B shows unregularized estimation via a log-linear model and shows many spurious strong connections, including many more negative connections than present in the true model. As such, Panel B highlights our need to regularize—even in a sizable dataset of 1,000 observations for a 19-node network. The simulated data has 22.2 observations for every parameter, far more than the 4.9 observations per parameter in Boschloo et al. (2015). Thus, even with a high sample size (*34,653 subjects*) and even when more subjects are measured than there are parameters present, it can still be advisable to use some form of regularization, as is done by Boschloo et al. (Boschloo et al., 2015) in using the LASSO. Panel C shows the result from using the LASSO, using the *IsingFit* package (van Borkulo et al., 2014). In this model, the clustering is generally retrieved—two of the bridging connections are retrieved and one negative connection is retrieved. However, the resulting structure is much more sparse than the true model,

and interpreting this structure could lead to the same conclusions determined by Boschloo and her colleagues: The number of connections differed across symptoms, connection strengths varied considerably across symptoms, and relatively few connections connected the two clusters. Finally, Panel D shows the result of a Rank 2 approximation, which is equivalent to a two-factor model. Here, it can be seen that although a dense structure is retrieved that shows the correct clustering, relationships of the clustering (the negative and bridging edges) are not retrieved. The online supplementary materials[2] show that with a higher sample size ($n = 5{,}000$) the estimation is more accurate and that the unregularized and LASSO estimations result in similar network structures.

Different Estimation Techniques

In light of the examples discussed in this chapter, researchers may wonder when they should and should not use a particular estimation method. For example, low-rank estimation is more suited in the example demonstrated in Figure 4.2, whereas LASSO estimation fits better in the example shown in Figure 4.3. These conclusions, however, depend on knowing the true network structure as shown in Figure 4.1—something a researcher will not know in reality. The choice of estimation method, therefore, is not trivial. Choosing the estimation method depends on three criteria: (1) the prior expectation of the true network structure, (2) the relative importance the researcher attributes to sensitivity (discovery) and specificity (caution), and (3) the practical applicability of an estimation procedure. When a researcher expects the true network to be low rank (e.g., due to latent variables), low-rank estimation should be preferred over LASSO regularization. On the other hand, when a researcher expects the network to be sparse, LASSO regularization should be used. In addition, LASSO regularization should be preferred when a researcher aims to have high specificity (i.e., to refrain from estimating an edge that is missing in the true model). Finally, practical arguments can play a role in choosing an estimation procedure as well. LASSO, particularly in combination with EBIC model selection, is relatively fast even with respect to large datasets. As a result, researchers could apply bootstrapping methods to the estimation procedure to further investigate the accuracy of parameter estimation (see Chapter 3), which may not be feasible for slower estimation procedures.

We focused the argumentation on LASSO regularization and low-rank approximations these are the main methodologies that have been

[2] http://sachaepskamp.com/files/IsingLatentNetworks.pdf

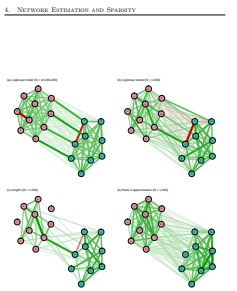


Figure 4.5: Estimated network structures based on data generated by the MIRT model in Figure 4.4.

applied in psychological literature and present two extreme cases of a range of different network structures that can be estimated. Because these methods lie on the extreme ends of sparsity relative to dense networks, they best exemplify the main point of this chapter: In small sample sizes, some assumptions of the true model must be made (e.g., the true model is sparse), and these assumptions influence the resulting network structure (e.g., the obtained network is sparse). This does not mean that LASSO and low-rank approximation are the only methods available. An alternative, for example, is to use *elastic-net* estimation, which uses LASSO regularization with edge regression (penalizing the sum of squared coefficients). The *elasticIsing* package (Epskamp, 2016) can be used to accomplish this; it uses cross-validation in selecting the tuning parameters. The online supplementary materials show an example of elastic-net applied to the data analyzed in Figure 4.5. It is noteworthy that the elastic-net procedure selected a dense network (i.e., ridge regression) over LASSO regularization, indicating that data-driven evidence can be garnered to argue whether or not LASSO regularization should be used. The obtained network, like the unregularized network in Figure 4.5 (Panel B), also shows many connections which were falsely estimated to be negative—the true model being sparse because a low-rank approximation returning a dense model seems to indicate that the true model is dense. These characteristics from the networks we obtain are a consequence of the method used to estimate a network structure (specifically the assumptions made by the employed method about the data-generating network structure) and often pollute the resulting estimated model (Kruis & Maris, 2015).

Recently it has been demonstrated that three, statistically indistinguishable, representations of the Ising model exist that explain observed associations between binary variables either through a common cause (latent variable), through the reciprocal effect between variables (network), or through the conditioning on a common effect (collider variable; Epskamp et al., in press; Marsman et al., 2015; Kruis & Maris, 2016). Consequently, when a model from one of these frameworks can sufficiently describe the associative structure of the measured variables, there exists an alternative representation for other frameworks that can also accurately represent the structure of the data. For example, Boschloo et al.'s sparse network structure (Boschloo et al., 2015), resulting from the LASSO being applied to the data, can also be described by a multidimensional latent variable model (with a single latent variable for each clique in the network) and residual correlations. As such, obtaining sufficient fit for a statistical network model cannot be regarded as evidence for the theoretical model, where a network structure acts as the causal mechanism from which associations between

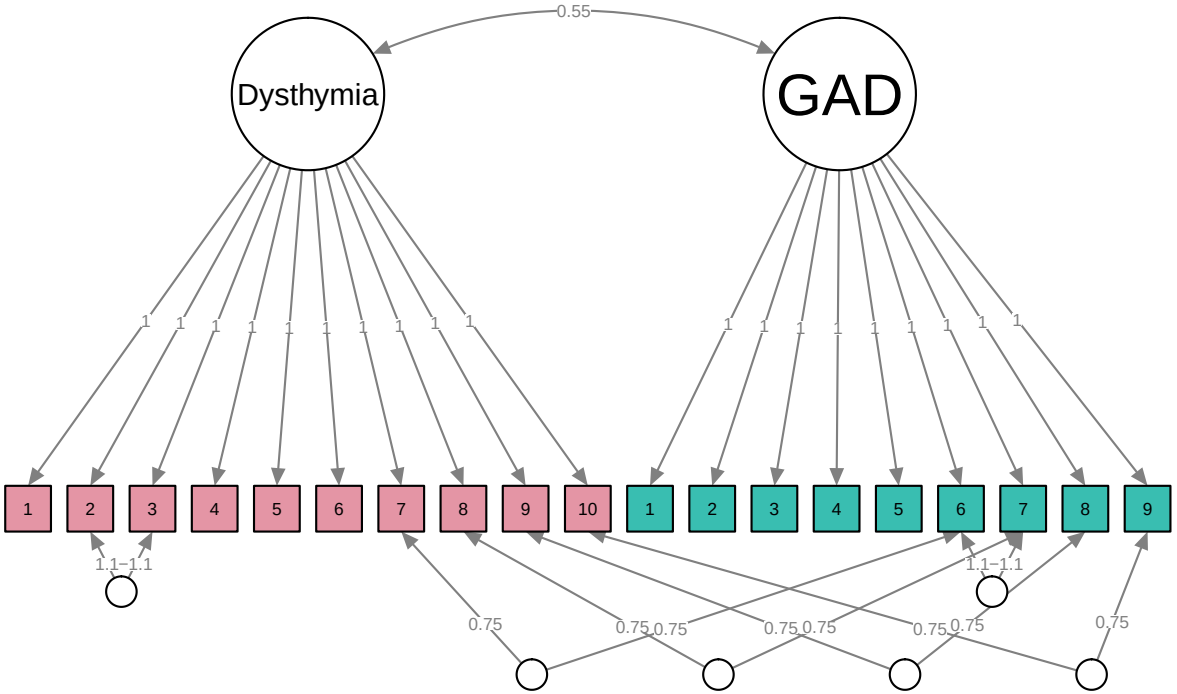
tion methods makes an assumption on the underlying true model structure: The LASSO assumes a sparse structure whereas low-rank approximation assumes a dense but low-rank structure. Investigating the results of the estimation methods cannot validate these assumptions. The LASSO always yields a sparse structure, which does not mean that the true underlying structure could not have been dense. On the other hand, low-rank approximations rarely produce sparse structures, but that does not mean that the true underlying structure could not have been sparse.

Figure 4.2 illustrates this point by showing that LASSO estimation when the true network structure is a Curie-Weiss model still results in a sparse structure. This means that observing any of the sparse structures shown in Figure 4.2 does not mean that the nodes in the network could not represent interchangeable indicators of a single latent trait. Figure 4.5 illustrates this point again in a plausible scenario in psychopathology and also shows that when the true network structure is complicated and neither sparse nor low rank, as is the case here, all regularization methods partly fail even when using a relatively large sample size. As such, interpreting the sparsity of such a structure is questionable: the LASSO resulting in a sparse model gives as little evidence for the true model being sparse because a low-rank approximation returning a dense model seems to indicate that the true model is dense. These characteristics from the networks we obtain are a consequence of the method used to estimate a network structure (specifically the assumptions made by the employed method about the data-generating network structure) and often pollute the resulting estimated model (Kruis & Maris, 2015).

Recently it has been demonstrated that three, statistically indistinguishable, representations of the Ising model exist that explain observed associations between binary variables either through a common cause (latent variable), through the reciprocal effect between variables (network), or through the conditioning on a common effect (collider variable; Epskamp et al., in press; Marsman et al., 2015; Kruis & Maris, 2016). Consequently, when a model from one of these frameworks can sufficiently describe the associative structure of the measured variables, there exists an alternative representation for other frameworks that can also accurately represent the structure of the data. For example, Boschloo et al.'s sparse network structure (Boschloo et al., 2015), resulting from the LASSO being applied to the data, can also be described by a multidimensional latent variable model (with a single latent variable for each clique in the network) and residual correlations. As such, obtaining sufficient fit for a statistical network model cannot be regarded as evidence for the theoretical model, where a network structure acts as the causal mechanism from which associations between

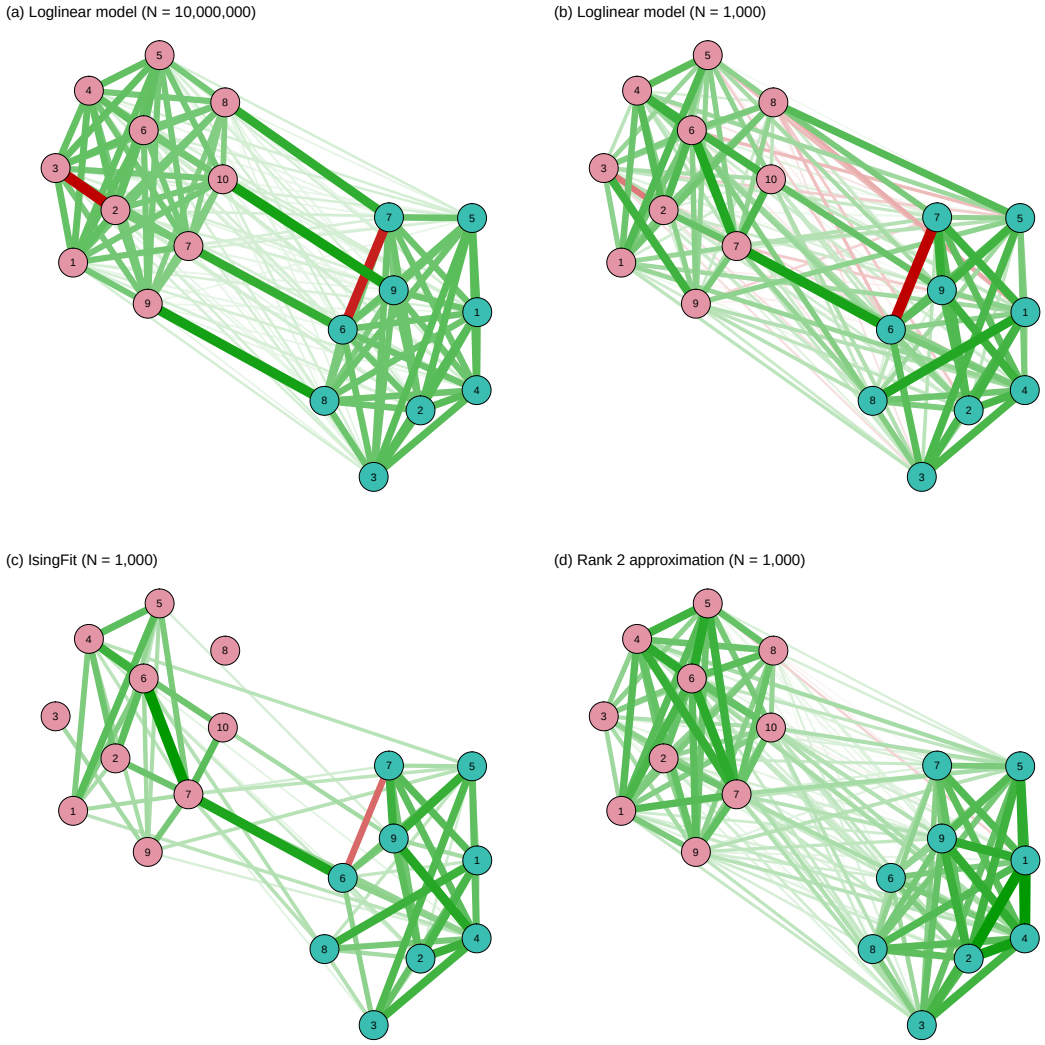
variables emerge. We therefore advise, in general, to tread carefully when drawing inferences about the theoretical causal mechanisms that generate the data from statistical model fit.

Network models show great promise in mapping out and visualizing relationships present in the data and are useful to comprehend high dimensional multivariate relationships. In addition, network models can be powerful tools to estimate the backbones of potential causal relationships—if those relationships are assumed to exist. Using the LASSO to estimate such network structures is a powerful tool in performing fast high-dimensional model selection that results in fewer false positives, and interpreting network structures obtained from the LASSO can illuminate the strong relationships present in the dataset. Important to realize is that using LASSO estimation will result in a sparse structure, and similarly, using a low-rank approximation will result in a dense low-rank result. Our aim here is not to argue against using the LASSO or to argue that estimating network structures is wrong. Our aim is to clarify that choosing the estimation method is not trivial and can greatly impact both the estimated structure as well as any conclusions drawn from that structure.



Dysthymia
- 1: Low mood for at least 2 years
- 2: Decrease in appitite
- 3: Increase in appetite
- 4: Low self−esteem
- 5: Indecisiveness
- 6: Feelings of Hopelessness
- 7: Insomnia
- 8: Hypersomnia
- 9: Diminished ability to concentrate
- 10: Fatigue

Generalized Anxiety Disorder (GAD)
- 1: Anxiety or worry for at least 6 months
- 2: Difficulty to control the worry
- 3: Restlessness
- 4: Irritability
- 5: Muscle Tension
- 6: Sleep disturbance 1: Insomnia
- 7: Sleep disturnance 2: Hypersomnia
- 8: Difficulty concentrating or mind going blank
- 9: Easily fatigued


(a) Loglinear model (N = 10,000,000)
(b) Loglinear model (N = 1,000)
(c) IsingFit (N = 1,000)
(d) Rank 2 approximation (N = 1,000)

Chapter 5

Personalized Network Modeling in Psychopathology

This chapter has been adapted from:

- Epskamp, S., van Borkulo, C.D., van der Veen, D.C., Servaas, M.N., Isvoranu, A.M., Riese, H., and Cramer, A.O.J. Personalized Network Modeling in Psychopathology: The Importance of Contemporaneous and Temporal Connections.

Chapter 5

Personalized Network Modeling in Psychopathology

Abstract

Two methodological trends have become prominent in the fields of psychopathology and psychiatry: (1) technological developments in collecting time-intensive, repeated, intra-individual measurements in order to capture symptom fluctuations and other time varying factors (e.g., emotions/affect) in daily life (i.e., time-series), and (2) an increasing number of statistical tools for for estimating associations between these measurements (i.e., network structures) based on these time-series data. Combining these two trends allows for the estimation of intra-individual network structures. Using vector-autoregression (VAR), two networks can be obtained: a *temporal network*, in which one investigates if symptoms (or other relevant variables) predict one another over time, and a *contemporaneous network*, in which one investigates if symptoms predict one another in the same window of measurement. The network literature using these models has so far mostly focused on the temporal network. Here we argue that temporal relations between psychopathological variables might typically unfold within shorter time intervals (e.g., minutes) than the time intervals commonly and feasibly used in current time-series studies (e.g., hours). As a result, such temporal relations will be captured

This chapter has been adapted from: Epskamp, S., van Borkulo, C.D., van der Veen, D.C., Servaas, M.N., Isvanian, A.M., Riese, H., and Cramer, A.O.J. Personalized Network Modeling in Psychopathology: The Importance of Contemporaneous and Temporal Connections.

in the contemporaneous network, rather than in the temporal network. Both temporal and contemporaneous networks may highlight *potential* causal pathways—they are not definitive proof of causality but may lead to meaningful insights. As such, both types of networks function as hypothesis *generators*. We conclude the chapter with empirical examples of such analyses on symptoms time-series data from clinical cases.

5.1 Introduction

Recent years have witnessed an emergence of two distinct trends in the study of psychopathology. First, technological advances have permitted the gathering of intensive repeated measurements of patients and healthy controls with the Experience Sampling Method (ESM: Aan het Rot, Hogenelst, & Schoevers, 2012; Myin-Germeys et al., 2009; Wichers, Lothmann, Simons, Nicolson, & Peeters, 2012). With ESM, participants are measured repeatedly within short time intervals during daily life. For example, someone is queried five times a day during a period of two weeks on his or her level of insomnia, depressed mood, and fatigue since the previous measurement. We will term the time frame on which one reports the *window of measurement*. The resulting time-series data allow for the investigation of intra-individual processes (Hamaker, 2012). The second trend is the network perspective on psychopathology, in which mental disorders are interpreted as the consequence of a dynamical interplay between symptoms and other variables (Borsboom & Cramer, 2013; Cramer et al., 2016; Cramer & Borsboom, 2015). This literature uses network models in an attempt to understand and predict the dynamics of psychopathology. From this perspective, symptoms are not seen as passive indications of a mental disorder but rather play an active role, making symptoms prime candidates for interventions (Borsboom, in press; Fried, Epskamp, et al., 2016).

Time-series data of a single individual offer a promising gateway into understanding the dynamical processes that may occur within that individual over time (e.g., Bringmann et al., 2013, 2015; Pe et al., 2015; Wigman et al., 2015). Such personalized network structures are typically estimated using a statistical technique called *vector-autoregression* (VAR; van der Krieke et al., 2015). Predominantly, VAR analyses have focused on the estimation of *temporal relationships* (relationships that occur between different windows of measurement). However, as will be outlined in Chapter 6, the residuals of the VAR model can be further used to estimate *contemporaneous relationships* (relationships that occur in the same window of measurement),

which are not yet commonly used in the field. In this chapter, we argue that both network structures provide valuable hypothesis-generating information directly applicable to the study of psychopathology as well as to clinical practice. We focus the majority of the discussion on explaining contemporaneous partial correlation networks, as these are not yet often utilized in the literature of intra-individual analysis. We exemplify this by analyzing two ESM datasets obtained from patients.

5.2 Temporal and Contemporaneous Networks

In time-series data analysis with an average time-interval of a few hours, a typical density statistical assumption is violated: consecutive responses are not likely to be independent (e.g., someone who is tired between 9:00 and 11:00 is likely to still be tired between 11:00 and 13:00). The minimal method of coping with this violation of independence is the lag-1 VAR model (van der Krieke et al., 2015). In this model, a variable in a certain window of measurement is predicted by the same variable in the previous window of measurement (*autoregressive* effects) and all other variables in the previous window of measurement (cross-lagged effects; Selig & Little, 2012)[1]. This model does not assume auto-correlations between larger differences in time (e.g., lag-2) are zero, but merely that such relationships can be fully explained by the lag-1 model. These *autoregressive* and *cross*-lagged effects can be estimated and visualized in a *network* (Bringmann et al., 2013). In this network, measured variables (such as symptoms) are represented by nodes. When one variable *predicts* another in the next window of measurement, we draw a link with an arrowhead pointing from one node to the other. We term this network the *temporal network*.

The predictive effects shown in the temporal network satisfy the assumption that in a causal relationship the cause must precede the effect. Therefore, these are often interpreted to be indicative of causal relationships. Only interpreting temporal coefficients, however, does not utilize VAR to its full potential. The residuals of the temporal VAR model are correlated; correlations in the *same* window of measurement remain that cannot be explained by the temporal effects. These correlations can be used to compute a network of *partial correlations* (Wild et al., 2010). In such a network, each variable is again represented by a node. Links (without arrowhead) between two nodes indicate the partial correlation obtained after controlling for both temporal effects and all other variables in the

[1] VAR can be seen as an ordinary regression where the predictors are lagged variables.

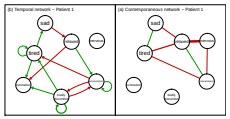


Figure 5.1: Two network structures that can be estimated with time-series data analysis, based on data of a clinical patient ($n = 52$) measured over a period of two weeks. The model was estimated using the *graphicalVAR* package for R. Circles (nodes) represent variables, such as symptoms, and connections (links, both bidirectional drawn as simple lines or directed drawn as an arrow) indicate predictive relationships. Green links indicate positive relationships, red links indicate negative relationships, and the width and saturation of a link indicates the strength of the relationship. The network on the left shows a *temporal network*, in which a link denotes that one variable predicts another in the next window of measurement. The network on the right shows a *contemporaneous network*, in which links indicate partial correlations between variables in the same window of measurement, after controlling for all other variables in the same window of measurement and all variables of the previous window of measurement.

same window of measurement. We term this network the *contemporaneous network*[2].

Figure 5.1 shows an example of the two network structures obtained from a VAR analysis. These networks are estimated using ESM data of a clinical patient, and are further described and interpreted in Section 5.6.

[2] The contemporaneous network should not be confused with a network of lag-0 (partial) correlation. Such a network would (1) not take into account that responses are not independent, and (2) present a mixture of temporal and contemporaneous effects. Thus, we obtain the contemporaneous network from the residuals of the VAR model, since only then relationships between windows of measurement and relationships *within* windows of measurement are separated.

The temporal and contemporaneous networks were estimated at the same time, using the methodology outlined by Abegaz and Wit (2013). The temporal network (a) shows auto-regressions (an arrow of a node pointing at itself) on three variables: 'tired', 'bodily discomfort' and 'concentration'. Thus, when this patient is tired she is likely still tired during the next window of measurement. There are cross-lagged relationships between several variables. For example, this patient being tired predicts her to ruminate more during the next window of measurement... The contemporaneous network (b) shows, among other relationships, a relationship between 'relaxed' and 'nervous': when this patient was tired she was also more likely to relax poorer, so *reported during the same window of measurement*. This can be seen as the direct consequence of a plausible causal relationship: being nervous might lead you to feel less relaxed (or vice-versa). There is no reason why such a causal relationship should take a few hours to occur, which brings us to the main point of this chapter.

5.3 Causation at the Contemporaneous Level

In a typical ESM study, the time between consecutive measurements is a few hours[3]. As such, the temporal network will only contain predictive effects of measured variables on other measured variables about a few hours later. However, it is likely that many causal relationships occur *much faster* than a timeframe of a few hours. Take for example a classical causal model:

Turn on sprinklers → Grass is wet.

Turning on the sprinklers causes the grass to become wet. This causal effect occurs very fast: after turning on the sprinklers it takes perhaps a *few seconds* for the grass to become wet. If we take measures of sprinklers ("on" or "off") and the wetness of the grass every two hours, it would be rather improbable to capture the case in which the sprinklers were turned on *just* before the grass became wet. As such, the temporal network would not contain a connection between turning on the sprinklers and the grass being wet, and likely would only contain a temporal auto-regression of the grass being wet (because it takes time for grass to dry). However, after controlling for this auto-regression, we most likely would find a connection between these variables in the contemporaneous network: in windows of measurement where the sprinklers were on we would likely also find that the grass was wet.

[3] Notable exceptions are sampling designs in which individuals are asked to fill-out questionnaires once a day or week.

We can think in a similar vein about psychopathological causal relationships. For example, a patient suffering from panic disorder might anticipate a panic attack by experiencing somatic arousal (e.g., sweating, increased heart rate):

Somatic arousal → anticipation of panic attack

In this structure, an arrow indicates that whatever is on the left causes whatever is on the right. This patient anticipates a panic attack, because the patient is experiencing somatic arousal. This causal effect would likely occur within minutes, not hours. Someone who experiences somatic arousal between 13:00 and 15:00 might still experience somatic arousal between 15:00 and 17:00. Thus, we can expect to find auto-regressions. However, between somatic arousal and anticipation of panic attack we would likely only find a contemporaneous connection.

In sum, relations between symptoms and other variables can plausibly unfold faster than the time-window of measurement; such relationships will be captured in the contemporaneous network. Figure 5.1 showed, however, that the contemporaneous network has no direction (links have no arrowheads). To understand *how* such undirected networks can still highlight potential causal pathways, we need to delve into the literature on estimation of networks in psychopathology.

5.4 Partial Correlation Networks

As outlined above, the contemporaneous relationships can be interpreted and drawn as a network of partial correlations. In this section, we describe how such partial correlation networks can be interpreted and how links in such a network can be seen as indicative of causal relationships. Partial correlation networks are part of a more general class of undirected (i.e., no arrows) networks (formally called Markov Random Fields; Lauritzen, 1996) that have been introduced to psychopathology in response to the call for conceptualizing psychopathological behavior (such as symptoms) as complex networks (Borsboom et al., 2011; Cramer et al., 2010). After the initial introduction of partial correlation networks to the psychopathological literature (Borsboom & Cramer, 2013), the use of undirected networks in psychopathology gained considerable traction following the introduction of easy-to-use estimation methods and publicly available software packages for both estimation and visualization (Epskamp et al., 2012; van Borkulo et al., 2014). Ever since, such network structures have been extensively applied to research in the fields of psychopathology and psychiatry, such

as comorbidity (Boschloo et al., 2015), autism (Ruzzano, Borsboom, & Geurts, 2015), post-traumatic stress disorder (McNally et al., 2015), psychotic disorders (van Borkulo, et al., 2014; Isvanian, Borsboom, et al., 2016), major depression (Fried et al., 2015; van Borkulo et al., 2015), and clinical care (Kroese et al., 2016).

Partial correlation networks have become so prominent because they present a relatively easy method to estimate and visualize potential causal pathways, while taking into account that observational data (i.e., no experimental interventions) only contains limited information on such causal relationships. In observational data, causality is reflected only in the conditional independence structure (Pearl, 2000). Conditional independence means that two variables are no longer correlated at fixed levels of a third variable. A partial correlation network shows conditional independence, because when the partial correlation between two variables after conditioning on all others equals zero, then that means two variables are conditionally independent. Therefore, two nodes that are not directly connected via a link are conditionally independent.

Taking again the patient described above suffering from a panic disorder, suppose we expand the causal structure to include this patient's pathway related to avoiding feared situations:

Somatic arousal → anticipation of panic attack → avoidance of feared situations.

Anticipating a panic attack might cause this patient to avoid feared situations, such as malls or busy shopping streets[4]. The causal structure indicates that we would expect to be able to predict this patient avoiding feared situations given that he or she is experiencing somatic arousal. However, if we already knew this person is anticipating a panic attack, we already predict that this person will avoid feared situations. Then, observing somatic arousal on top of anticipating the panic attack does not improve this prediction. Thus, we would expect non-zero partial correlations between somatic arousal and anticipation of panic attack, and between anticipation of panic attack and avoidance. We would furthermore expect a partial correlation of *zero* between somatic arousal and avoidance behavior; somatic arousal and avoidance behavior are conditionally independent given the anticipation of a panic attack. Consequently, we would expect the following partial correlation network:

Somatic arousal — anticipation of panic attack — avoidance behavior.

[4] These relationships should be taken as an example. The direction of such effects is still at topic of debate, and likely differs from patient to patient (Pitliya, 1980).

Finding such a partial correlation network often does not allow one to find the true direction of causation. This is due to two technical arguments: (1) equivalent non models explain the same conditional independencies and (2) these models only work when we can assume the causal structure is acyclic (i.e., contains no feedback loops). Concerning the first argument, we can summarize the above causal structure as $A \to B \to C$, in which A and C are conditionally independent given B. This conditional independence, however, also holds for two other models: $A \leftarrow B \leftarrow C$ and $A \leftarrow B \to C$ (Pearl, 2000). In general, we cannot distinguish between these three models using only observational data. Adding more variables only increases this problem of potentially equivalent models, making it difficult to construct such a network only from observational data. Even when such a network can be constructed, we must assume that the structure is not self-enforcing. That is, a variable cannot cause itself via some chain (e.g., $A \to B \to C \to A$). In psychopathology, however, this assumption likely does not hold (in our example above: anticipating a panic attack might cause more somatic arousal). As a result of these problems with directed structures when temporal or experimental information is lacking, undirected networks have been more successful in this emergent research field. In an undirected network, the observation that A and C are conditionally independent given B is represented by only one model: $A - B - C$ (Lauritzen, 1996).

To summarize, the contemporaneous network is an undirected network without arrowheads. This network shows a link when two variables are not conditionally independent given both all responses in the previous window of measurement and responses of all other variables in the current window of measurement. When two variables are conditionally independent, no link is drawn. If a partial correlation network is present, we would expect such a link, and if a causal relationship were not present, we would not expect such a link. Therefore, the links in the contemporaneous network can be indicative of causal relationships. However, as finding the direction of such relationships is hard, we do not attempt to do so and keep the links in the contemporaneous network without direction.

5.5 Generating Causal Hypotheses

The connections in both the temporal and contemporaneous network cannot be interpreted as true causal relationships except under strong assumptions. The pathways shown can only be *indicative* of potential causal relationships. Such a pathway is a necessary condition for causality (we would

expect such relationships when there is a true causal effect), but not sufficient (the relationship can also be spurious and due to, e.g., unobserved causes; Pearl, 2000). Therefore, these networks can be seen as *hypothesis generating*. To test for causality one needs to investigate what happens after experimentally changing one variable. If fatigue causes concentration problems, we would expect concentration levels to change after experimentally making someone fatigued. Experimentally changing concentration levels should, on the other hand, not influence fatigue. Such causal testing can only be done experimentally; it is hard to infer causality from observational data, no matter how often and intensive someone is measured and how intensive the sampling rate is.

In addition to generating hypotheses on causal links, both networks also generate hypotheses on which nodes are important. The importance of nodes in a network can be quantified with descriptive measures called *centrality measures* (Costantini, Epskamp, et al., 2015; Newman, 2010; Opsahl et al., 2010). A node with a high centrality is said to be 'central', indicating the node is well connected in the network. Such a central node may be a prime candidate for intervention, as targeting this node will influence the rest of the system. This is not only the case for central nodes in the temporal network, but also for central nodes in the contemporaneous network. Even when a node has no temporal connections, it can still carry a lot of information on subsequent measurements, purely by being central in the contemporaneous network. For example, if A predicts B and C in the same window of measurement (contemporaneous links), and B and C both predict themselves in the next window of measurement (autoregressions), then as a result A is able to predict B and C in the next window of measurement, even though no cross-lagged relationships might be found in the temporal network.

While experimental intervention is needed to test causal hypotheses, such hypotheses on causal relationships and central nodes might be hard to verify in practice. For example, one cannot wait with forming treatment plans until after lengthy experimental designs have been tested on a clinical patient. In addition, in intensive treatments for example, multiple nodes are likely to be targeted simultaneously; the causal effect of one particular node is hard to test. Furthermore, it might not be known how certain symptoms can be treated at all (e.g., feelings of derealisation when the patient is suffering from a comorbid depressonalisation disorder, a disorder that is often concurrent with a panic disorder). Still, the obtained insights are useful: the personalized networks can be discussed with the patient and, when the patient recognizes the discovered relationships, help to generate hypotheses and choose interventions that target these nodes (Kroese et al.,

2016).

5.6 Clinical Example

To exemplify how the described symptom networks can be utilized in clinical practice, we analyzed ESM data obtained from two patients treated in a tertiary outpatient clinic in Groningen. Patient 1 was a female patient, aged 23, who received cognitive behavioral therapy (CBT) for a severe panic disorder and a depressive disorder secondary to the panic disorder. Her response rate was 74%. Patient 2 was a female patient, aged 53 suffering from major depressive disorder, in early partial remission after having received electroconvulsive therapy (ECT). Her response rate was 93%, and data collection started one day after her last ECT session.

Methods

The patients received an extensive briefing plus written user instructions for the ESM measurements. Digital support was available 24/7. Patient data were gathered during normal daily life with an ESM tool developed for an ongoing epidemiological study. With our secured server system (RoQua; roqua.nl; Sytema & Van der Krieke, 2013), text messages with links to online questionnaires were sent to the patient's smartphone. All items could be answered on a 7-point Likert scale varying from '1=not at all' to '7=very much'. Measurement occasions were scheduled five times a day every three hours for two weeks (maximal number of possible measurements = 70), and took three to five minutes to complete. The timing of the measurements was adjusted to their individual daily rhythm with the last measurement timed 30 minutes before going to bed. Patients were instructed to fill-out the questionnaires as soon as possible after receiving the text message. The patient received a reminder after 30 minutes, and after 60 minutes the link was closed. The protocol used was submitted to the ethical review board of the University Medical Center Groningen, who confirmed that formal assessment was not required. Prior to participation, patients were fully informed about the study after which they gave written informed consent.

To estimate simple network structures with not many nodes, we selected seven of the administered variables that usually should interact with each other: feeling sad, being tired, ruminating, experiencing bodily discomfort, feeling nervous, feeling relaxed and being able to concentrate. Network structures were standardized as described by Wild et al. (2010) to avoid misleading parameter estimates in the network structure (Rubinov, Tserlinckx, Bruse, & Cuelemans, 2016). The networks were estimated using the

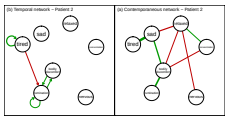


Figure 5.2: Contemporaneous (panel a) and temporal (panel b) network structures of a clinical patient ($n = 65$) measured over a period of two weeks. The model was estimated using the graphicalVAR package for R.

graphicalVAR package for R (Epskamp, 2015), which uses penalized maximum likelihood estimation to estimate model parameters (strength of connections) while simultaneously controlling for parsimony (which links are removed; Abegaz & Wit, 2013; Rothman et al., 2010). The graphicalVAR package estimates 2,500 different models, varying 50 levels of parsimony in the temporal network and 50 levels of parsimony in the contemporaneous network. Bayesian Information Criterion (BIC) model selection was used to select the best fitting model. A more detailed description of the estimation procedure is beyond the scope of this chapter. An introduction to model selection of regularized networks is provided in Chapter 2, and a methodological introduction to this modeling framework is provided in Chapter 6. We refer the reader to (Abegaz & Wit, 2013) for the estimation procedure used.

Results

Figure 5.1 shows the two network structures of Patient 1. The temporal network in Panel (a) shows several connections involving bodily discomfort: whenever she experienced bodily discomfort, she ruminated more, felt more tired and was less able to concentrate three hours later. The contemporaneous network in Panel (b) shows that feeling relaxed plays a central role in the network. Whenever she was relaxed she experienced less sadness,

tiredness, nervousness and was better able to concentrate (and vise versa, e.g., whenever she experienced less sadness she was more relaxed). In the case of Patient 1, therapy sessions revealed that intensively cleaning her house was her way of coping with stress. This lead to bodily discomfort and eventually rumination about her inability to do the things in the way she used to do things, resulting in a sad mood. Teaching her other ways to cope with stress broke this negative pattern.

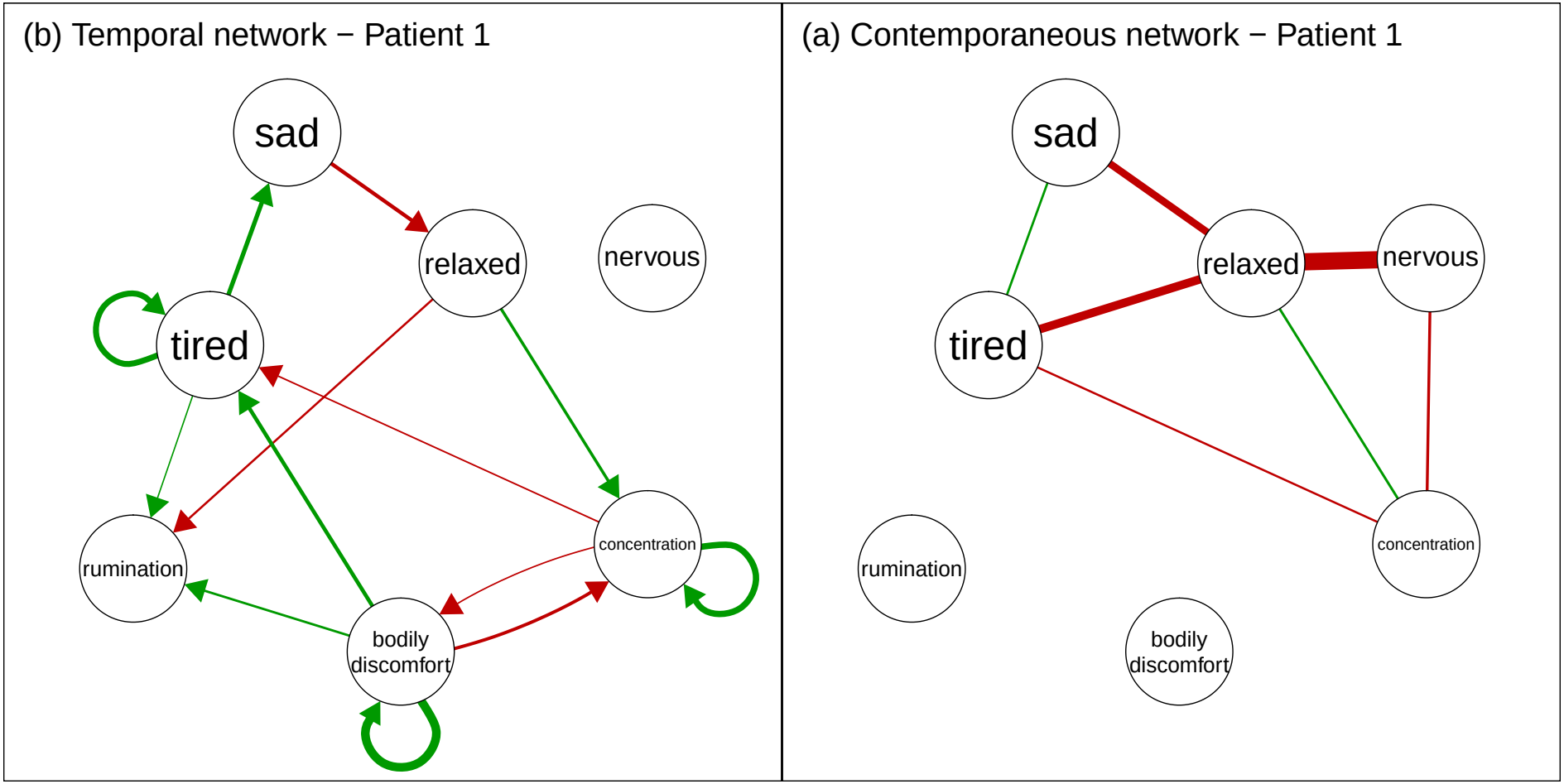
Figure 5.2 shows the two network structures of Patient 2. The contemporaneous network in Panel (b) features more connections than the temporal network in Panel (a). In the contemporaneous network, the node bodily discomfort has a central role. Whenever Patient 2 experienced bodily discomfort (in her case, palpitations), she felt sadder, less relaxed, ruminated more and was less able to concentrate within the same window of measurement. This fits the pathology of a panic disorder where bodily sensations are interpreted catastrophically. The temporal network shows the effects over time and highlights a potential feedback loop, where bodily discomfort rumination (in her case, catastrophic interpretations of the bodily sensations) leads to more attention to bodily discomfort, causing more rumination. Feeling tired seems also to lead to more rumination in time.

5.7 Conclusion

In this chapter we argued that when analyzing intra-individual time-series data in clinical settings, researchers should focus on both temporal and contemporaneous relationships. While temporal networks are commonly estimated and interpreted in the network approach to psychopathology (e.g., Bringmann et al., 2013), contemporaneous networks, especially when drawn as a partial correlation network, are not commonly used. We have argued that both contemporaneous and temporal networks can highlight meaningful relationships, interpretable and useable by patients and clinicians in treatment, as well as present researchers with hypothesis generating exploratory results on potential causal relationships. Such personalized knowledge can be used for intervention selection (e.g., choosing which symptoms to treat), as well as generate testable hypotheses pertaining to the individual patient that can be used to perform experiments. In addition to temporal relationship, contemporaneous relationships are also important in discovering psychological dynamics, as such relationships can also occur at a much faster time scale than the typical lag interval used in ESM studies.

A main limitation of the VAR method is that, even when contemporaneous networks are estimated, the results depend on the lag-interval used. If the lag interval is too long, meaningful relationships might not be retrieved (e.g., some dynamics might occur between days or weeks rather than hours). Conversely, if the relationship is too fast, and dissipates fast, it might also not be retrieved (e.g., if the effect of a relation dissipates after minutes, it might not be captured in a design that measures persons hourly or slower). The optimal lag-interval is often unknown, and can even differ between individuals and different variables. The lag-interval used is typically chosen in part due to practical reasons; it is not feasible for a patient to fill out a questionnaire often during a day (e.g., 20 times a day). The data gathering can also not take too long (e.g., more than two weeks), as the VAR model typically assumes people do not structurally change (Hadlock & Waldorp, 2016a; Wichers et al., 2016). While effects that are slower than the lag-interval could be captured in a second temporal network (e.g., a network between days in addition to a network between minutes; de Haans-Rietdijk, Kuppens, & Hamaker, 2016), such methods require more observations.

The aim of this chapter is not to argue against interpreting temporal coefficients—both temporal and contemporaneous effects contain meaningful information on how the observed variables relate to one-another. Regardless, we strongly argue that the temporal and contemporaneous relationships should not be over-interpreted, as these merely highlight potential causal pathways. So what is the use then of contemporaneous and temporal networks if they do not allow for causal interpretation? We argue that, for an individual patient, it is hardly relevant if relationships in lag-for data are causal or not. What matters is that both types of networks give the clinician as well as the patient a personalized, and visualized window into a patient's daily life. Moreover, this personalized window comes with a host of opportunities to arrive at tailor-made intervention strategies (e.g., treating central symptom of patient), and to monitor progress (e.g., will "deactivating" central symptom result in the deactivation of other symptoms?). Additionally, discussing the idiographic intricacies of networks with the patient offers ample opportunity for the patient to gain insight into his/her strengths and pitfalls and for reinforcing a sense of participation in one's own cure. Personalized care is on everybody's agenda, and rightly so; given its benefits, so should network modeling of psychopathology data at the level of the individual be.


(b) Temporal network − Patient 1
sad
relaxed
nervous
tired
rumination
concentration
bodily discomfort
(a) Contemporaneous network − Patient 1
sad
relaxed
nervous
tired
rumination
concentration
bodily discomfort

Part II

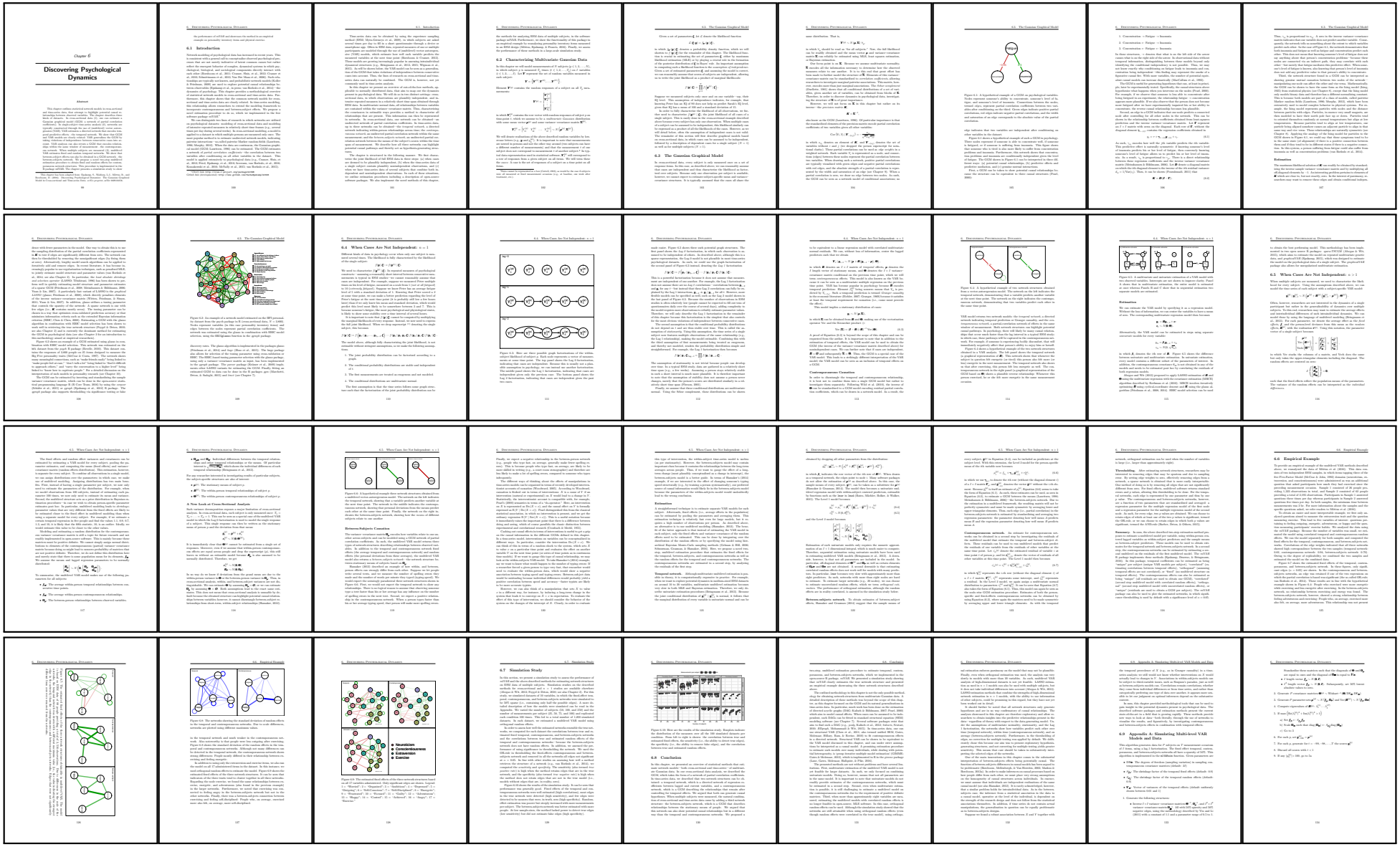
Technical Advances in Network Psychometrics

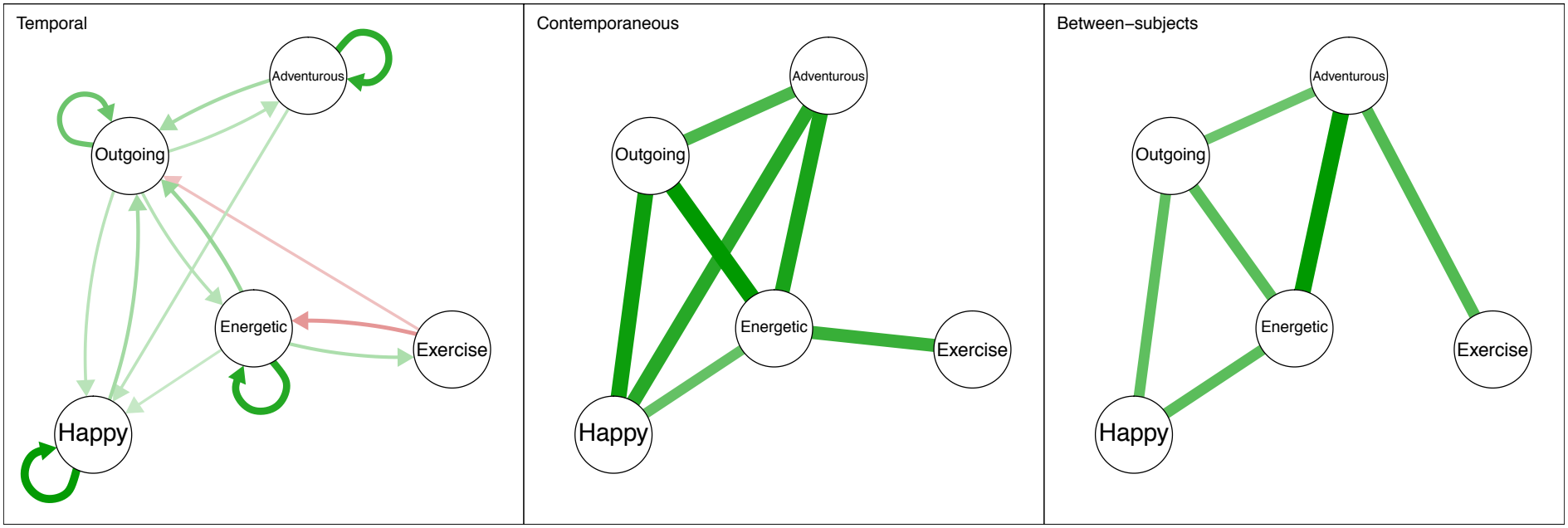
Chapter 6

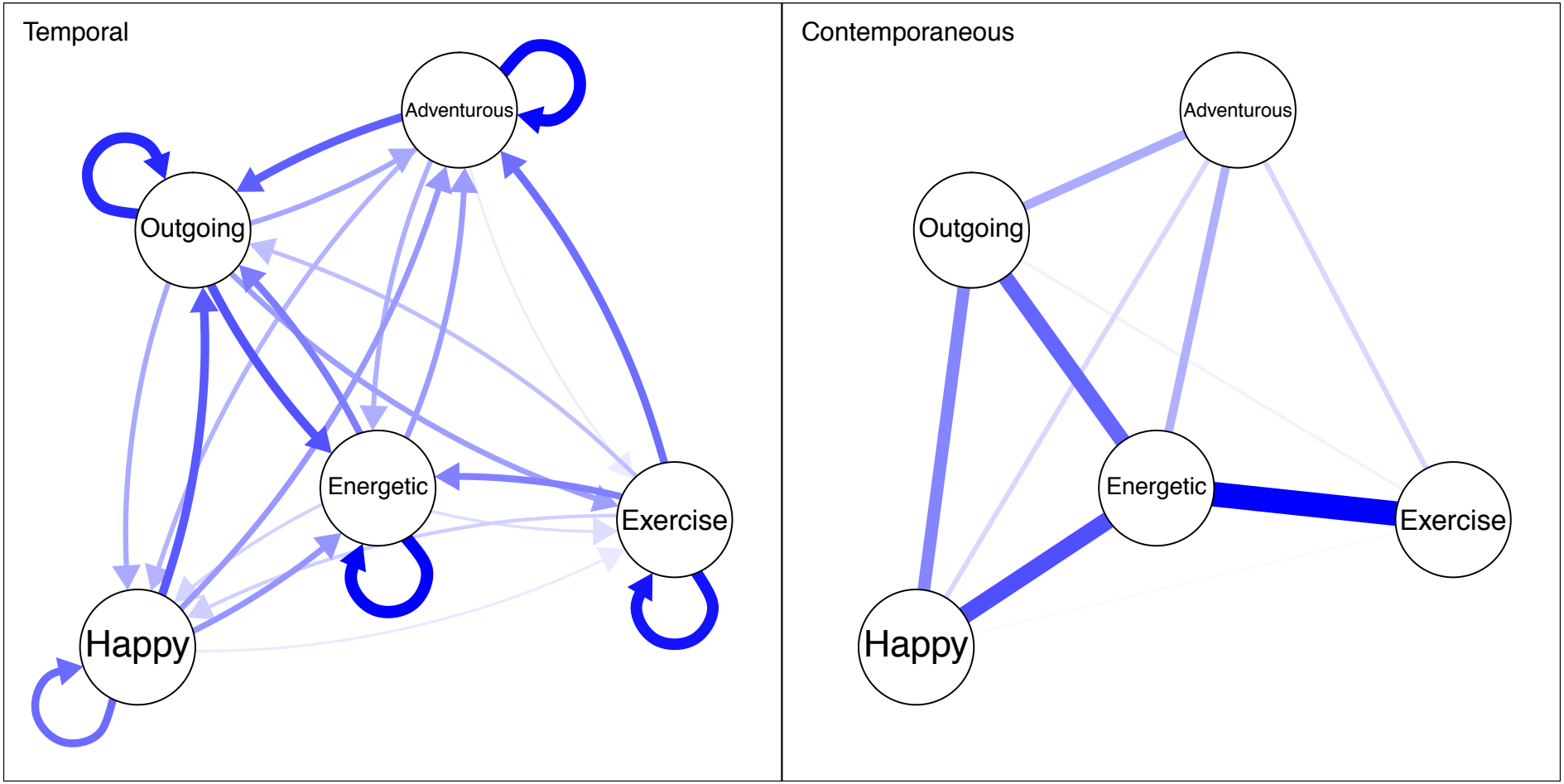
Discovering Psychological Dynamics

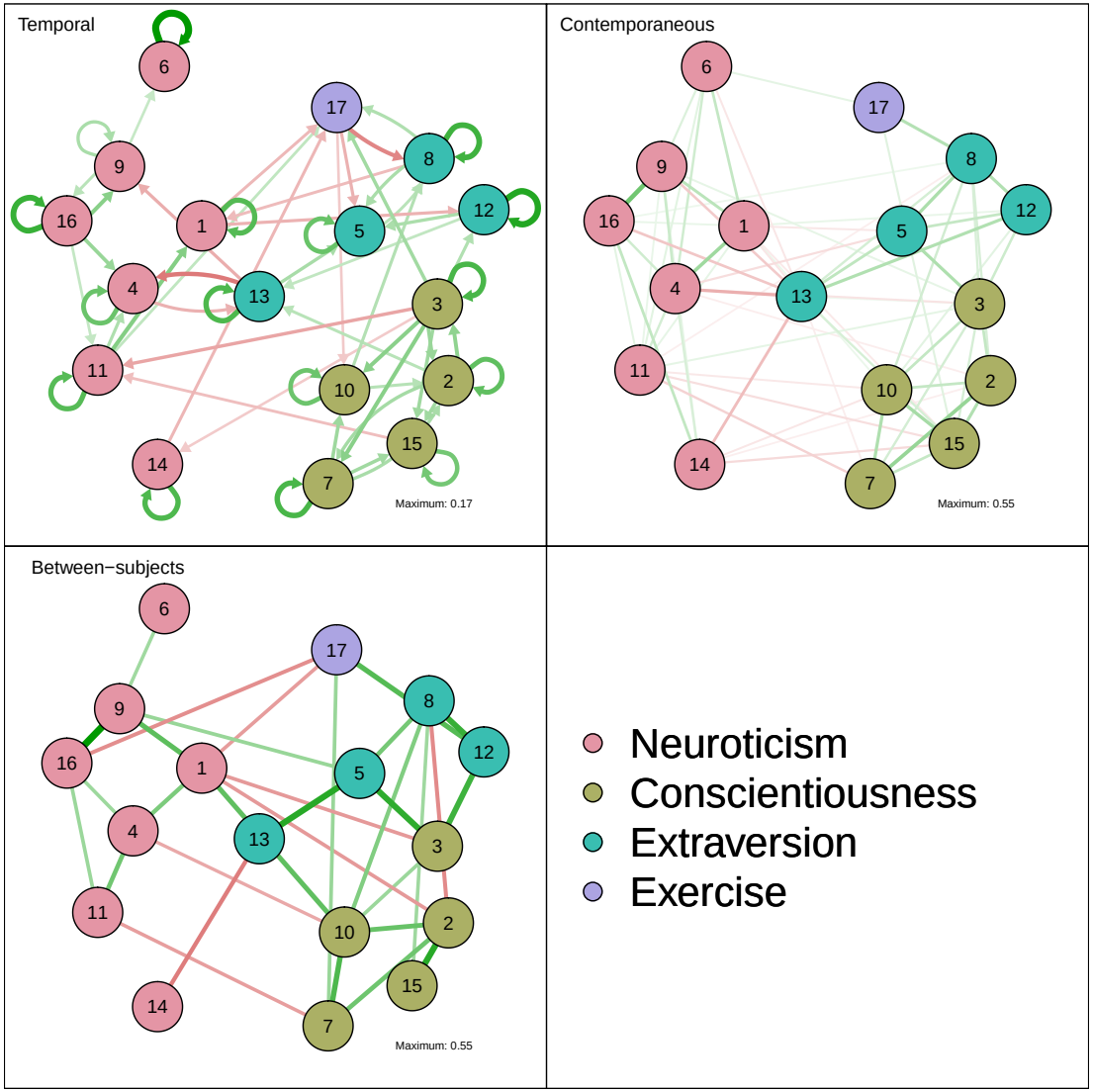
This chapter has been adapted from:

- Epskamp, S., Waldorp, L.J., Mõttus, R., and Borsboom, D. (2016). Discovering Psychological Dynamics: The Gaussian Graphical Model in Cross-sectional and Time-series Data. *arXiv preprint*, arXiv:1609.04156.




Temporal
Adventurous
Outgoing
Energetic
Exercise
Happy
Contemporaneous
Adventurous
Outgoing
Energetic
Exercise
Happy
Between−subjects
Adventurous
Outgoing
Energetic
Exercise
Happy


Temporal
Adventurous
Outgoing
Energetic
Exercise
Happy
Contemporaneous
Adventurous
Outgoing
Energetic
Exercise
Happy

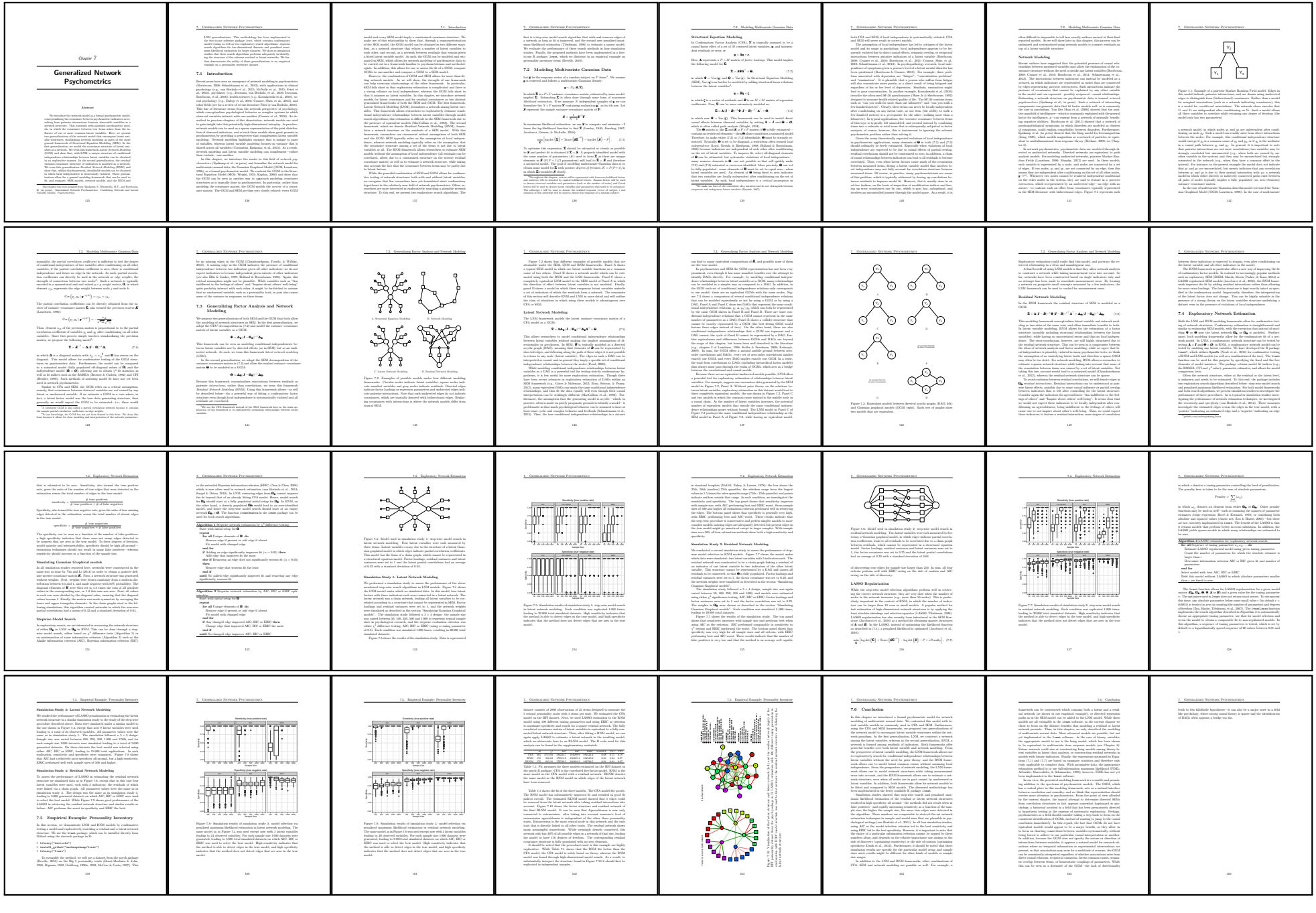

Temporal
Maximum: 0.17
Contemporaneous
Maximum: 0.55
Between−subjects
Maximum: 0.55
Neuroticism
Conscientiousness
Extraversion
Exercise

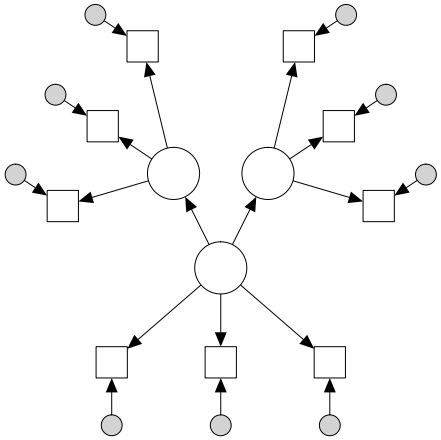
Chapter 7

Generalized Network Psychometrics

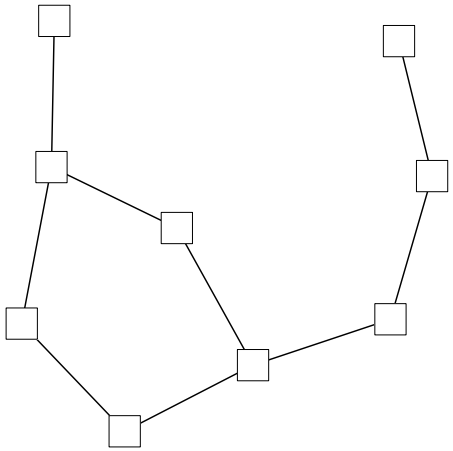
This chapter has been adapted from:

- Epskamp, S., Rhemtulla, M.T., and Borsboom, D. (in press). Generalized Network Psychometrics: Combining Network and Latent Variable Models. *Psychometrika*.

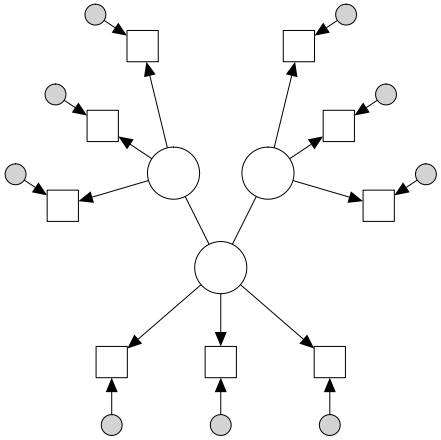




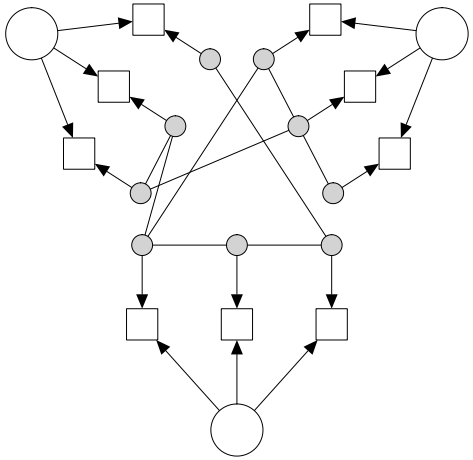
A. Structural Equation Modeling



B. Network Modeling

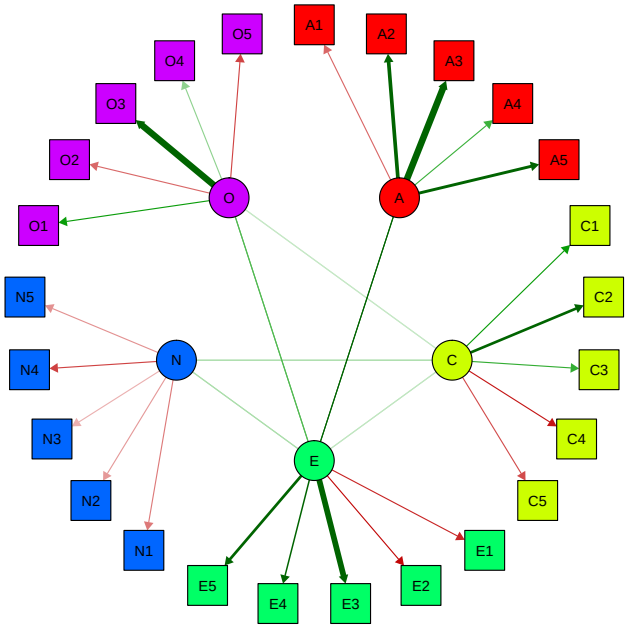


C. Latent Network Modeling

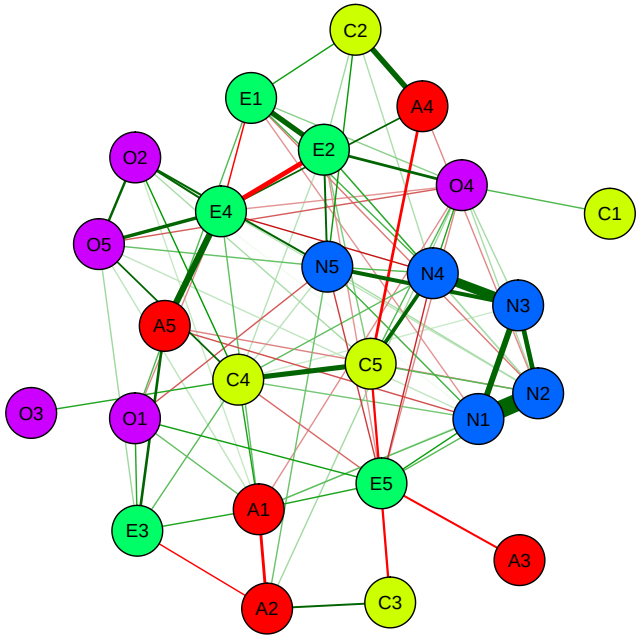


D. Residual Network Modeling

Factor structure & latent network



Residual network



Agreeableness
- A1: Am indifferent to the feelings of others.
- A2: Inquire about others' well−being.
- A3: Know how to comfort others.
- A4: Love children.
- A5: Make people feel at ease.

Conscientiousness
- C1: Am exacting in my work.
- C2: Continue until everything is perfect.
- C3: Do things according to a plan.
- C4: Do things in a half−way manner.
- C5: Waste my time.

Extraversion
- E1: Don't talk a lot.
- E2: Find it difficult to approach others.
- E3: Know how to captivate people.
- E4: Make friends easily.
- E5: Take charge.

Neuroticism
- N1: Get angry easily.
- N2: Get irritated easily.
- N3: Have frequent mood swings.
- N4: Often feel blue.
- N5: Panic easily.

Openness
- O1: Am full of ideas.
- O2: Avoid difficult reading material.
- O3: Carry the conversation to a higher level.
- O4: Spend time reflecting on things.
- O5: Will not probe deeply into a subject.

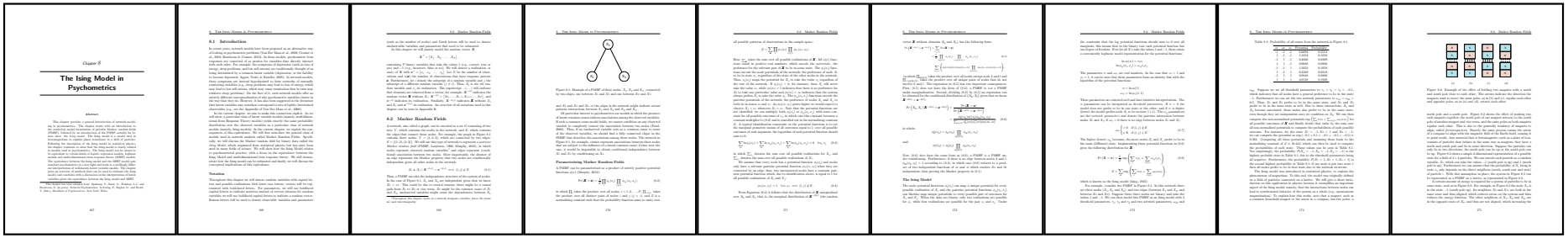
	df	χ^2	AIC	BIC	EBIC	RMSEA	TLI	CFI
CFA	265	4713.94	183233.7	183589.9	184542.4	0.08	0.75	0.78
RNM	172	806.63	179511.0	180419.4	182848.2	0.04	0.94	0.97
RLNM	176	843.18	179539.5	180424.2	182789.5	0.04	0.94	0.97

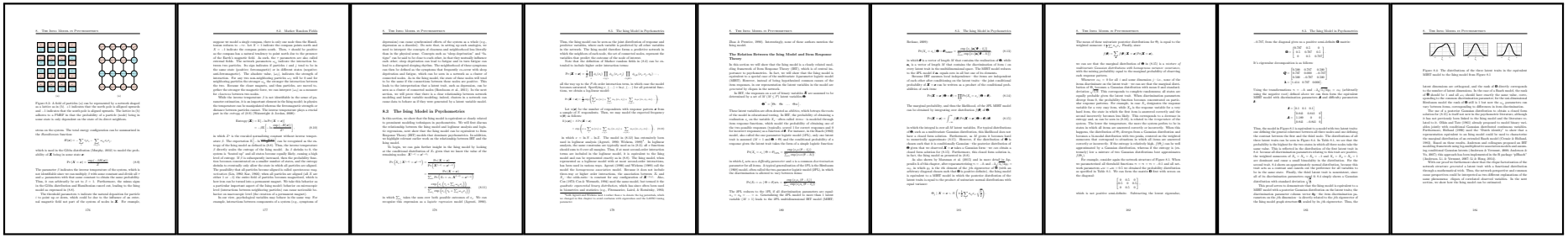
Chapter 8

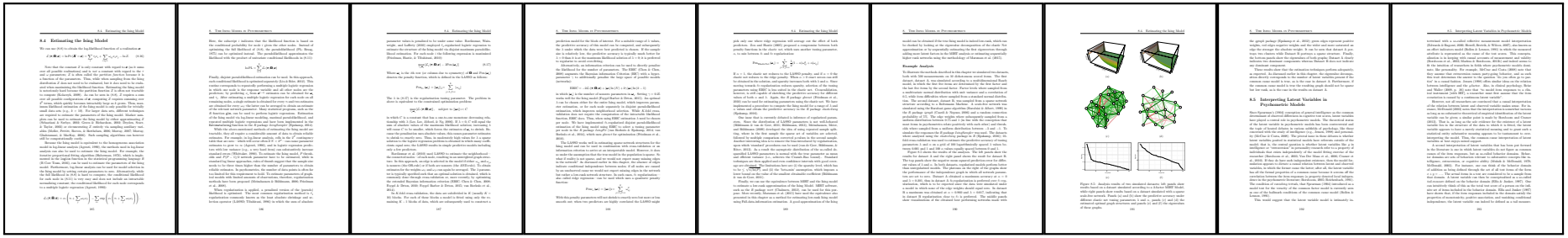
The Ising Model in Psychometrics

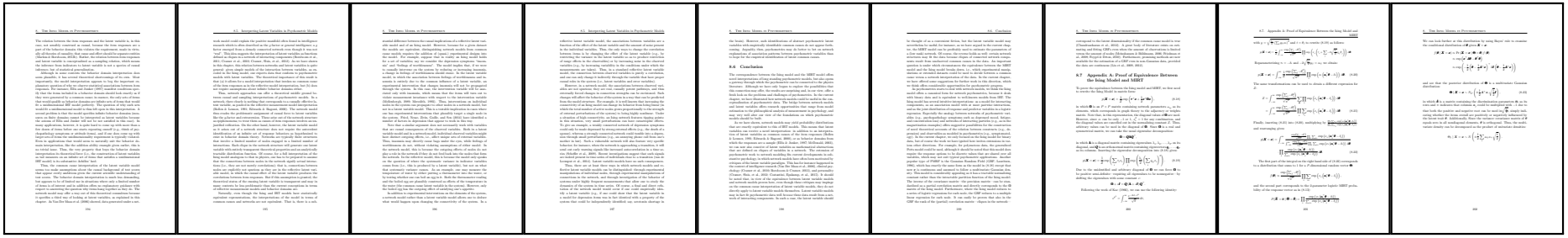
This chapter has been adapted from:

- Epskamp, S., Maris, G., Waldorp, L.J., and Borsboom, D. (in press). Network Psychometrics. In Irwing, P., Hughes, D., and Booth, T. (Eds.), *Handbook of Psychometrics*. New York: Wiley.









- Introduces the Ising model
- Ising model in statistical physics
- Equivalences between the Ising model and psychometric models:
  - Logistic regression
  - Loglinear modeling
  - Multivariate IRT
- Estimation of the Ising model
- Interpreting latent variables

Part III

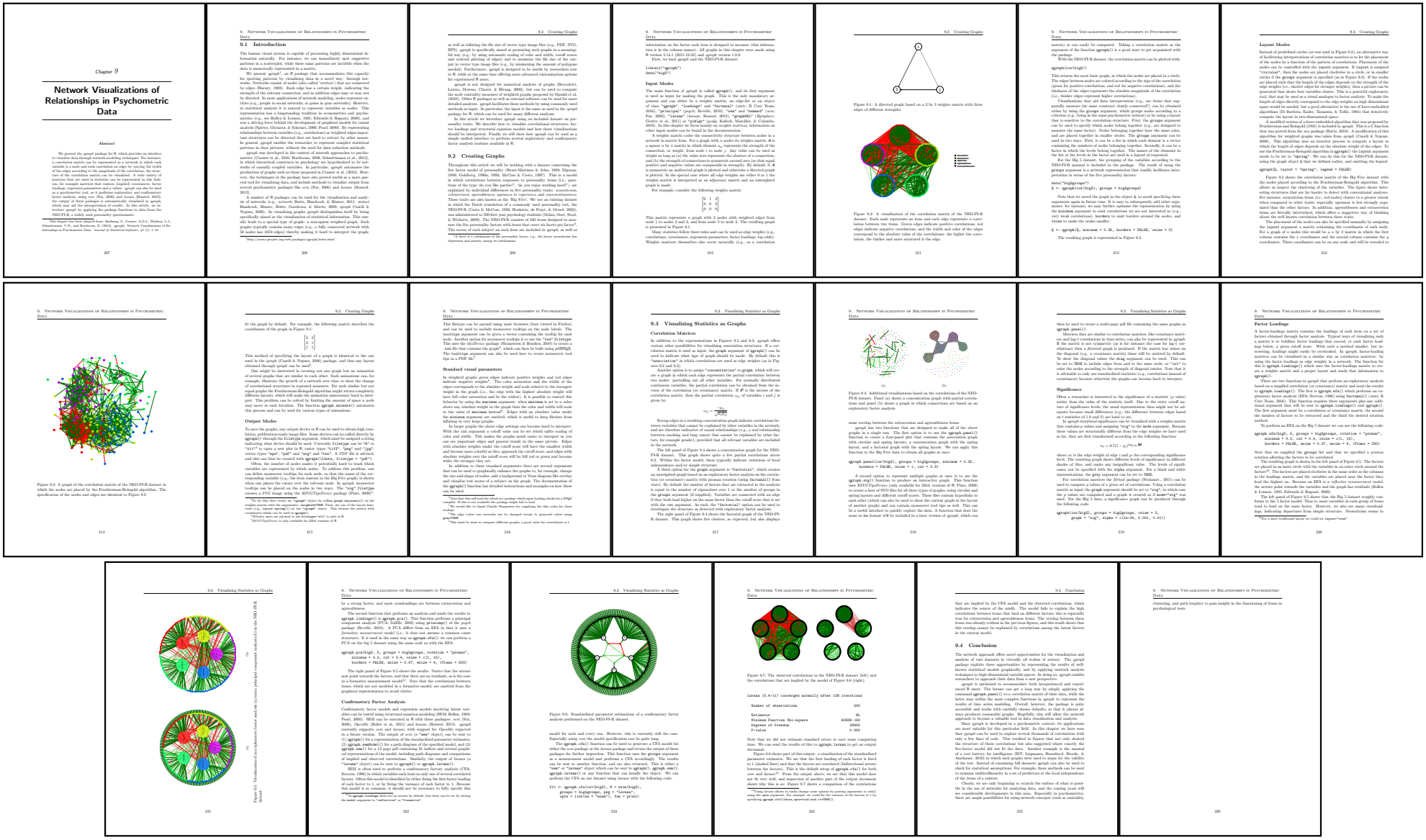
Visualizing Psychometrics and Personality Research

Chapter 9

Network Visualizations of Relationships in Psychometric Data

This chapter has been adapted from:

- Epskamp, S., Cramer, A.O.J., Waldorp, L.J., Schmittmann, V.D., and Borsboom, D. (2012). qgraph: Network Visualizations of Relationships in Psychometric Data. *Journal of Statistical Software, 48* (1), 1–18.



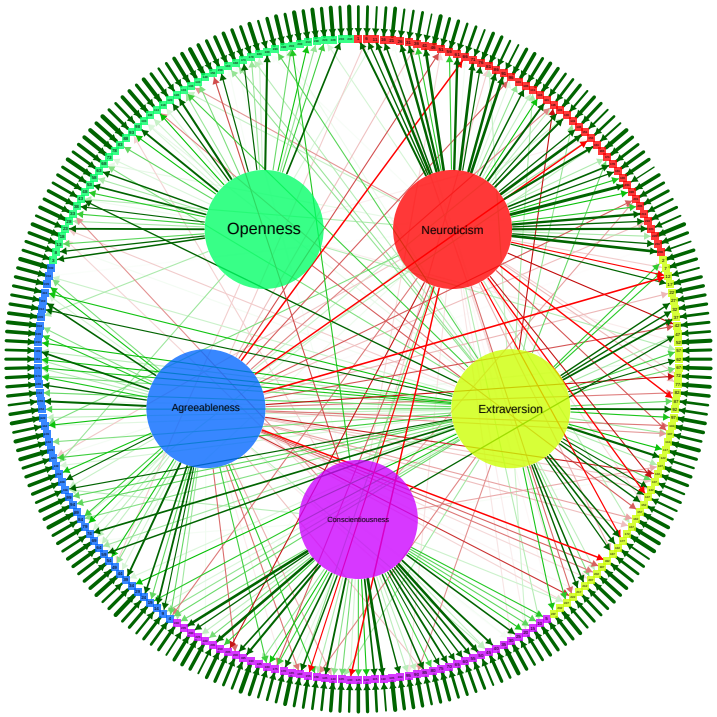
Chapter 9

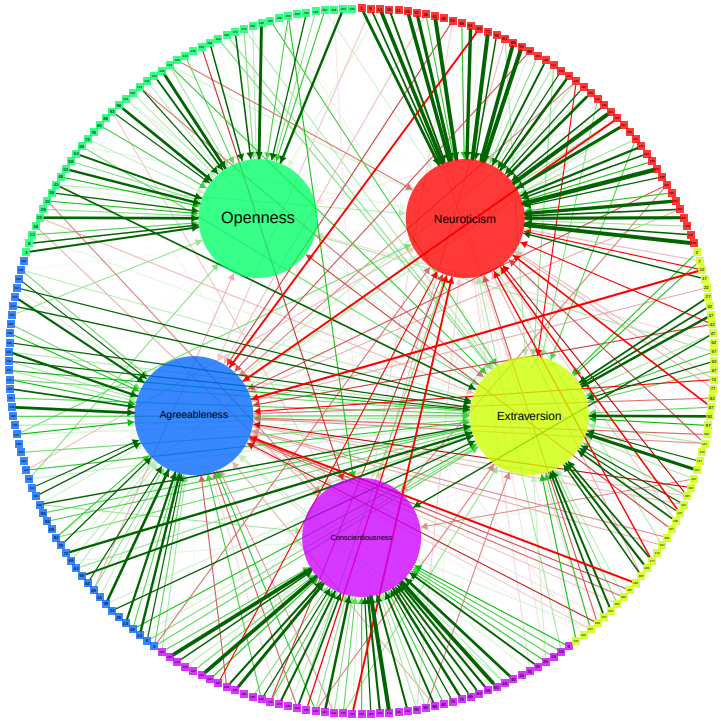
Network Visualizations of Relationships in Psychometric Data

Abstract

We present the qgraph package for R, which provides an interface to visualize data through network modeling techniques. For instance, correlation matrices can be visualized as a network where each node represents a variable or a node and each correlation an edge by varying the width of the edges according to the magnitude of the correlation, the structure of the correlation matrix can be visualized. A wide variety of matrices that are used in statistics can be represented in this fashion, for examples matrices that contain (implied) covariances, factor loadings, regression parameters and p values. qgraph can also be used to visualize results of a PCA, exploratory factor analysis, confirmatory factor analysis, using cvm (Fox, 2006) and lavaan (Rosseel, 2012); the output of these packages is automatically visualized in qgraph, which may aid the interpretation of results. In this article, we introduce qgraph by applying the package functions to data from the NEO-PI-R, a widely used personality questionnaire.

This chapter has been adapted from: Epskamp, S., Cramer, A.O.J., Waldorp, L.J., Schmittmann, V.D., and Borsboom, D. (2012). qgraph: Network Visualizations of Relationships in Psychometric Data. *Journal of Statistical Software*, *48* (1): 1–18.


Openness
Neuroticism
Agreeableness
Extraversion
Conscientiousness


Openness
Neuroticism
Agreeableness
Extraversion
Conscientiousness

Chapter 10

State of the aRt Personality Research

This chapter has been adapted from:

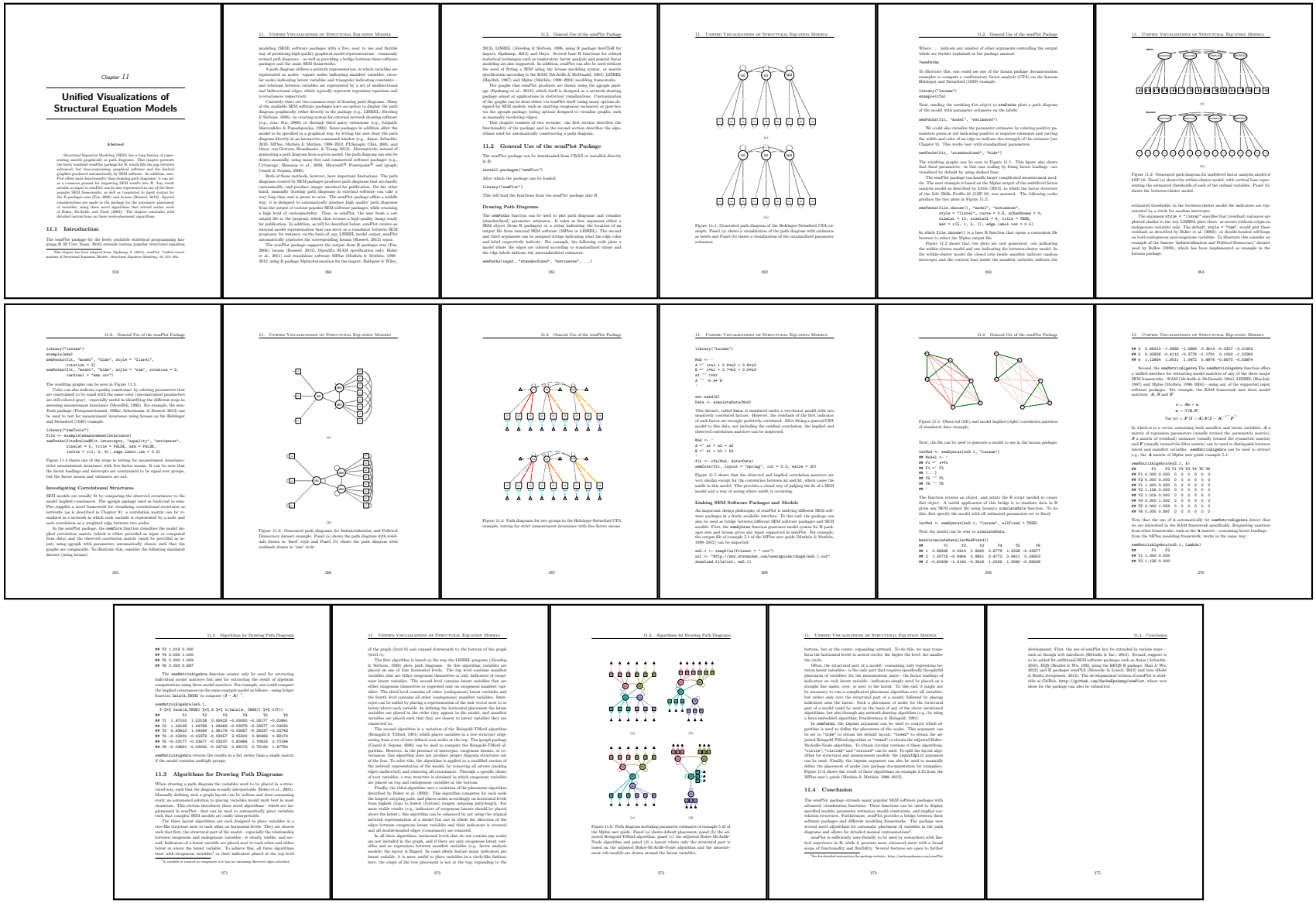
- Costantini, G., Epskamp, S., Borsboom, D., Perugini, M., Mõttus, R., Waldorp, L. J., and Cramer, A. O. J. (2014). State of the aRt personality research: A tutorial on network analysis of personality data in R. *Journal of Research in Personality 54,* 13–29. (The first two authors contributed equally to this work).

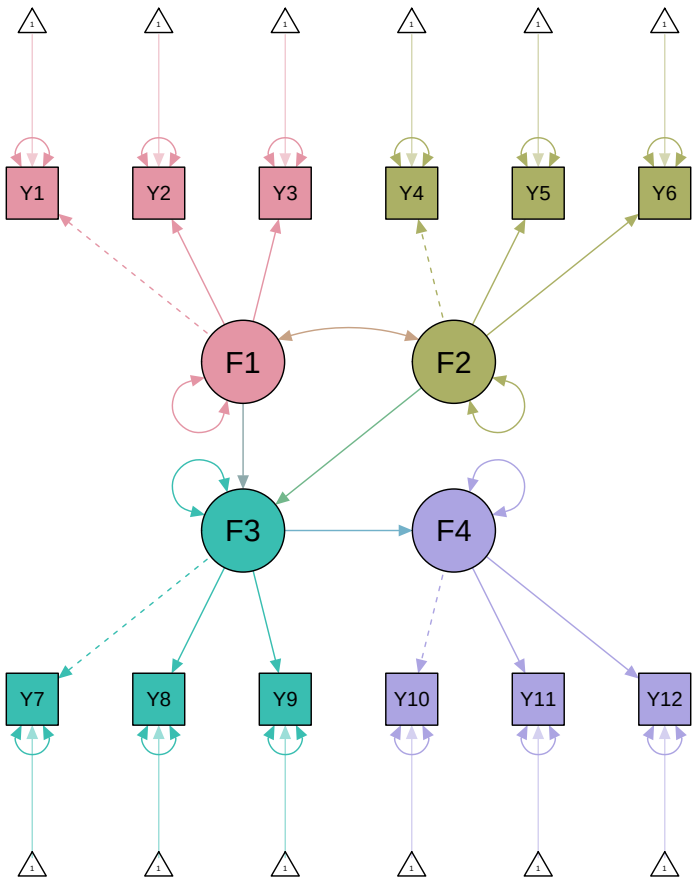
Chapter 11

Unified Visualizations of Structural Equation Models

This chapter has been adapted from:

- Epskamp, S. (2015). semPlot: Unified visualizations of Structural Equation Models. *Structural Equation Modeling, 22*, 474–483.



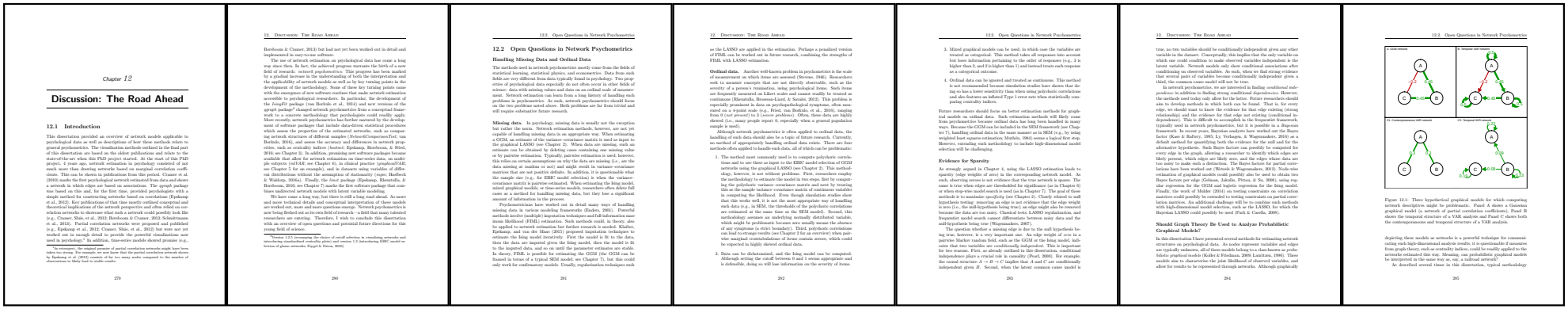

1
Y1
Y2
Y3
Y4
Y5
Y6
F1
F2
F3
F4
Y7
Y8
Y9
Y10
Y11
Y12

Part IV

Conclusion

Chapter 12

Discussion: The Road Ahead



Chapter 12

Discussion: The Road Ahead

12.1 Introduction

This dissertation provided an *overview* of network models applicable to psychological data as well as descriptions of how these methods relate to general psychometrics. The visualization methods outlined in the final part of this dissertation are based on the oldest publications and relate to the state-of-the-art when this PhD project started. At the start of this PhD project, 4 years ago, network estimation in psychology consisted of not much more than drawing networks based on marginal correlation coefficients. This can be shown in publications from this period. Cramer et al. [2010] marks the first psychological network estimated from data and shows a network in which edges are based on associations. The qgraph package was based on this and, for the first time, provided psychologists with a simple method for constructing networks based on correlations (Epskamp et al., 2012). Key publications of that time mostly outlined conceptual and theoretical implications of the network perspective and often relied on correlation networks to showcase what such a network could possibly look like (e.g., Cramer, Sluis, et al., 2012; Borsboom & Cramer, 2013; Schmittmann et al., 2013). Partial correlation networks were proposed and published (e.g., Epskamp et al., 2012; Cramer, Sluis, et al., 2012) but were not yet worked out in enough detail to provide the powerful visualizations now used in psychology.[1] In addition, time-series models showed promise (e.g.,

[1] In retrospect, the original promoter of partial correlation networks might have been taken too strong. For example, we now know that the partial correlation network shown by Epskamp et al. [2012] consists of far too many nodes compared to the number of observations to likely lead to stable results.

279

Borsboom & Cramer, 2013) but had not yet been worked out in detail and implemented in easy-to-use software.

The use of network estimation on psychological data has come a long way since then. In fact, the achieved progress warrants the birth of a *new* field of research: *network psychometrics*. This progress has been marked by a gradual increase in the understanding of both the interpretation and the applicability of network models as well as by key turning points in the development of the methodology. Some of these key turning points came with the emergence of new software routines that made network estimation accessible to psychological researchers. In particular, the development of the *lavaqtl* package (van Borkulo et al., 2014) and new versions of the qgraph package[2] changed network psychometrics from a conceptual framework to a concrete methodology that psychologists could readily apply. More recently, network psychometrics has further matured by the development of software packages that include data-driven estimation procedures which assess the properties of the estimated networks, such as comparing network structures of different samples (NetworkComparisonTest; van Borkulo, 2016), and assess the accuracy and differences in network properties, such as centrality indices (bootnet; Epskamp, Borsboom, & Fried, 2016; see Chapter 3). In addition, promising new software packages became available that allow for network estimation on time-series data, on multiple subjects (mlVAR; see Chapter 6), in clinical practice (graphicalVAR; see Chapter 5 for an example), and in datasets using variables of different distributions without the assumption of stationarity (mgm; Haslbeck & Waldorp, 2016a). Finally, the *bootnet* package (Epskamp, Borsboom, & Fried, 2016; see Chapter 7) marks the first software package that combines unbiased network models with latent variable modeling.

We have come a long way, but there is still a long road ahead. As more and more technical details and conceptual interpretation of these models are worked out, more and more questions emerge. Network psychometrics is now being fleshed out as its own field of research—a field that many talented researchers are entering. Therefore, I wish to conclude this dissertation with an overview of open questions and potential future directions for this young field of science.

[2] Version 1.0 (corresponding the release of cutoff selection in visualizing networks and interchanging standardized centrality plots) and version 1.3 (introducing EBIC model selection of gloses networks; Foygel & Drton, 2010).

280

12. Discussion: The Road Ahead

12.2 Open Questions in Network Psychometrics

Handling Missing Data and Ordinal Data

The methods used in network psychometrics mostly come from the fields of statistical learning, statistical physics, and econometrics. Data from such fields are very different from data typically found in psychology. Two properties of psychological data especially do not often occur in other fields of science: data with missing values and data on an ordinal scale of measurement. Network estimation can learn from a long history of handling such problems in psychometrics. As such, network psychometrics should focus on the two problems noted above. Both problems are far from trivial and will require substantive future research.

Missing data. In psychology, missing data is usually not the exception but rather the norm. Network estimation methods, however, are not yet capable of handling missing data in an appropriate way. When estimating a GGM, an estimate of the variance–covariance matrix is used as input to the graphical LASSO (see Chapter 2). This matrix, however, cannot be estimate can be obtained by deleting cases containing one missing value or by pairwise estimation. Typically, pairwise estimation is used; however, this relies on certain assumptions on why the data are missing (i.e., are the data missing at random or not) and might result in variance–covariance matrices that are not positive definite. In addition, it is questionable what the sample size (e.g., for EBIC model selection) is when the variance–covariance matrix is pairwise estimated. When estimating the Ising model, mixed graphical models, or time-series models, researchers often delete full cases as a method for handling missing data, but they lose a significant amount of information in the process.

Psychometricians have worked out in detail many ways of handling missing data in various modeling frameworks (Enders, 2001). Powerful methods involve (multiple) imputation techniques and full-information maximum likelihood (FIML) estimation. Such methods could, in theory, also be applied to network estimation but further research is needed. Kloiber, Epskamp, and van der Maas (2015) proposed imputation techniques to estimate the Ising model iteratively. First the model is fit to the data, then the data are imputed given the Ising model, then the model is fit to the imputed data, and so on until the parameter estimates are stable. In theory, FIML is possible for estimating the GGM (the GGM can be framed in terms of a typical SEM model; see Chapter 7), but this could only work for confirmatory models. Usually, regularization techniques such

281

as the LASSO are applied in the estimation. Perhaps a penalized version of FIML can be worked out in future research, combining the strengths of FIML with LASSO estimation.

Ordinal data. Another well-known problem in psychometrics is the scale of measurement on which items are answered (Stevens, 1946). Researchers seek to measure concepts that are not directly observable, such as the severity of a person's rumination, using psychological items. Such items are frequently measured on Likert scales and cannot readily be treated as continuous (Rhemtulla, Brosseau-Liard, & Savalei, 2012). This problem is especially prominent in data on psychological symptoms, often measured on a 4-point scale (e.g., Fried, van Borkulo, et al., 2016), ranging from 0 (*not present*) to 3 (*severe problems*). Often, these data are highly skewed (i.e., many people report 0, especially when a general population sample is used).

Although network psychometrics is often applied to ordinal data, the handling of such data should also be a topic of future research. Currently, no method of appropriately handling ordinal data exists. There are four methods often applied to handle such data, all of which can be problematic:

1. The method most commonly used is to compute polychoric correlations and to use these as input to the EBIC model selection of GGM networks using the graphical LASSO (see Chapter 2). This method, however, is not without problems. First, researchers employ the methodology to estimate the model in two steps, first by computing the polychoric variance–covariance matrix and next by treating this as the sample variance–covariance matrix of continuous variables in computing the likelihood. Even though simulation studies show that this works well, it is not the most appropriate way of handling such data (e.g., in SEM, the thresholds of the polychoric correlations are estimated at the same time as the SEM model). Second, this methodology assumes an underlying normally distributed variable, which might be problematic because zero usually means the absence of any symptoms (a strict boundary). Third, polychoric correlations can lead to strange results (see Chapter 2 for an overview) when pairwise marginal contingencies of items contain zeroes, which could be expected in highly skewed ordinal data.
2. Data can be dichotomized, and the Ising model can be computed. Although setting the cutoff between 0 and 1 seems appropriate and is defensible, doing so will lose information on the severity of items.

282

12. Discussion: The Road Ahead

3. Mixed graphical models can be used, in which case the variables are treated as categorical. This method takes all responses into account but loses information pertaining to the order of responses (e.g., 3 is higher than 2, and 2 is higher than 1) and instead treats each response as a categorical outcome.
4. Ordinal data can be ignored and treated as continuous. This method is not recommended because simulation studies have shown that doing so has a lower sensitivity than when using polychoric correlations and also features an inflated Type 1 error rate when statistically computing centrality indices.

Future researchers should focus on better estimation methods for graphical models on ordinal data. Such estimation methods will likely come from psychometrics because ordinal data has long been handled in many ways. Because the GGM can be included in the SEM framework (see Chapter 7), handling ordinal data in the same manner as in SEM (e.g., by using weighted least squares estimation; Muthén, 1984) seems a logical first step. However, extending such methodology to include high-dimensional model selection will be challenging.

Evidence for Sparsity

As strongly argued in Chapter 4, using the LASSO estimation leads to sparsity (edge weights of zero) in the corresponding network model. As such, observing zeroes is not evidence that the true network is sparse. The same is true when edges are thresholded for significance (as in Chapter 6) or when step-wise model search is used (as in Chapter 7). The goal of these methods is to maximize *specificity* (see Chapter 2). Closely related to null hypothesis testing, *removing* an edge is not evidence that the edge weight is zero (i.e., the null-hypothesis being true); an edge might also be removed because the data are too noisy. Classical tests, LASSO regularization, and frequentist model search cannot differentiate between noisy data and the null-hypothesis being true (Wagenmakers, 2007).

The question whether a missing edge is due to the null hypothesis being true, however, is a very important one. An edge weight of zero in a pairwise Markov random field, such as the GGM or the Ising model, indicates that two variables are *conditionally independent*. This is important for two reasons. First, as already outlined in this dissertation, conditional independence plays a crucial role in causality (Pearl, 2000). For example, the causal structure $A \leftrightarrow B \rightarrow C$ implies that A and C are conditionally independent *given* B. Second, when the latent common cause model is

283

12. Discussion: The Road Ahead

true, no two variables should be conditionally independent given any other variable in the dataset. Conceptually, this implies that the only variable on which one could condition to make observed variables independent is the latent variable. Network models only show conditional associations after conditioning on observed variables. As such, when we find strong evidence that several pairs of variables become conditionally independent given a third, the common cause model will not be true.

In network psychometrics, we are interested in finding *conditional independence* in addition to finding strong conditional dependencies. However, the methods used today only allow for the latter. Future researchers should aim to develop methods in which both can be found. That is, for every edge, we should want to know the evidence for that edge existing (strong relationship) and the evidence for that edge not existing (conditional independence). This is difficult to accomplish in the frequentist framework, typically used in network psychometrics, but it is possible in a *Bayesian* framework. In recent years, Bayesian analysts have worked out the Bayes factor (Kass & Raftery, 1995; Ly, Verhagen, & Wagenmakers, 2016) as a default method for quantifying both the evidence for the null and for the alternative hypothesis. Such Bayes factors can possibly be computed for every edge in the graph, allowing a researcher to identify which edges are likely present, which edges are likely zero, and the edges where data are too noisy to make such a distinction. The Bayes factors for partial correlations have been worked out (Wetzels & Wagenmakers, 2012). Node-wise estimation of graphical models could possibly also be used to obtain two Bayes factors per edge (Gelman, Jakulin, Pittau, & Su, 2008), using regular regression for the GGM and logistic regression for the Ising model. Finally, the work of Mulder [2016] on testing constraints on correlation matrices could possibly be extended to testing constraints on partial correlation matrices. An additional challenge will be to combine such methods with high-dimensional model selection, such as the LASSO, for which the Bayesian LASSO could possibly be used (Park & Casella, 2008).

Should Graph Theory Be Used to Analyze Probabilistic Graphical Models?

In this dissertation I have presented several methods for estimating network structures on psychological data. As nodes represent variables and edges are typically unknown, all of these models belong to a class known as *probabilistic graphical models* (Koller & Friedman, 2009; Lauritzen, 1996). These models aim to characterize the joint likelihood of observed variables, and allow for results to be represented through networks. Although graphical

284

12.2. Open Questions in Network Psychometrics

Figure 12.1: Three hypothetical graphical models for which computing network descriptives might be problematic. Panel A shows a Gaussian graphical model (a network of partial correlation coefficients); Panel B shows the temporal structure of a VAR analysis and Panel C shows both the contemporaneous and temporal structure of a VAR analysis.

depicting these models as networks is a powerful technique for communicating each high-dimensional analysis results, it is questionable if measures from graph theory, such as centrality indices, could be readily applied to the networks estimated this way. Meaning, can probabilistic graphical models be interpreted in the same way as, say, a railroad network?

As described several times in this dissertation, typical methodology

285

for analysing weighted networks—such as computing centrality measures (Opsahl et al., 2010)—are often used on models obtained in network psychometrics. In this methodology, each edge is first transformed into a *length* (i.e., the inverse absolute value of an edge weight), then the resulting network is analyzed as, for example, a railroad network would. The distance between two nodes in a network is defined through computing *shortest* paths, an important property in computing both closeness node centrality (i.e., the inverse sum of such distances) and the betweenness node centrality (i.e., measuring how often a node lies on the shortest path). Such methods are insightful when researchers can interpret these shortest paths to be sensible (e.g., passengers using a rail network will probably only travel through the shortest path), but it might make less sense in the context of probabilistic graphical models.

Figure 12.1 illustrates some examples of where interpreting such structures as a railroad network would lead to a strange interpretation of the data. When investigating the distance between nodes A and C in a typical network analysis would indicate that the *shortest path* is the path $A \rightarrow B \rightarrow C$, resulting in B having the highest betweenness and the direct path between A and C being ignored. However, such a result would not take all the information of the model into account. In fact, A and C are marginally independent; the correlation between A and C is exactly zero, indicating that knowing A contains no information on C and vice versa. Such a structure could emerge if B is a common effect of A and C, in which case distributing A can, in no way, have any effect on C. As such, it is questionable what it means for B to have a high betweenness if no causal effect goes through B.

Panel B shows a vector auto-regression (VAR) temporal model in which the contemporaneous structure is ignored. This network would lead to a similar conclusion because of the network in Panel A. The shortest path from A to C goes via B. In this network, however, edges indicate Lag-1 effects. This means that the path $A \rightarrow B \rightarrow C$ indicates a Lag-2 effect, whereas the direct path $A \rightarrow B$ indicates a Lag-1 effect. Such paths are not even comparable because they indicate completely different temporal structures. Finally, Panel C shows both model matrices obtained from a VAR model. Suppose a researcher is interested in identifying which node is best able to predict all nodes at later measurement. In this case, only investigating the temporal structure would lead to the conclusion that the most important node is B. However, such an analysis would not take the contemporaneous network, in which Node A is highly central, into account.

286

12. Discussion: The Road Ahead

Information theory. A potential solution for such problems is to not interpret probabilistic graphical models as networks, but rather, for what they are: full characterizations of the joint likelihood. In this line of thinking, the graphical representation is only useful for visualizing the statistical results but should not be over interpreted. The estimated model, nonetheless, is extremely powerful because it captures the associational structure of a dataset without the need for underlying theory on the causal mechanisms. A possible solution for inference methods then lies in the use of *information theory* (Cover & Thomas, 2012), which has shown to be a promising gateway to understanding the full complexity of such systems (Ugam, Aparicio, & Shoi, 2013; Quax, Kandhai, & Sloot, 2013).

In information theory, we can make use of the Shannon entropy (Cover & Thomas, 2012) of a set of random variables, $\boldsymbol{Y}$, which denotes the average amount of bits of information needed to communicate a discrete outcome. When dealing with continuous variables, as we do in the GGM and VAR models, we can define the *differential entropy*:

$$h(\boldsymbol{Y}) = -\mathbb{E}\log_2 f(\boldsymbol{Y}),$$

in which $f(\boldsymbol{y})$ denotes the density function of $\boldsymbol{Y}$. This measure can be computed for any number of variables and quantifies their volatility—in the case of a single continuous variable, the entropy is directly related to the variance. Now, divide $\boldsymbol{Y}$ in two subsets $\boldsymbol{Y}^{(1)}$ and $\boldsymbol{Y}^{(2)}$. We can then quantify the association in the two subsets using the *mutual information*:

$$I\left(\boldsymbol{Y}^{(1)}; \boldsymbol{Y}^{(2)}\right) = h\left(\boldsymbol{Y}^{(1)}\right) + h\left(\boldsymbol{Y}^{(2)}\right) - h\left(\boldsymbol{Y}^{(1)}, \boldsymbol{Y}^{(2)}\right).$$

This measure can act as a general measure for strength of association between any set of variables to any other set of variables.

In the network perspective, we treat strongly associated variables as densely connected and usually take the analogy of such strongly connected variables to be close to one another. Mutual information can then be seen as a new form of quantifying *closeness* between nodes or sets of nodes—the inverse of distance. Therefore, mutual information is an alternative to the shortest path length. This measure not only takes the shortest paths into account but all other paths as well. For example, the mutual information of two variables or nodes (we often use these terms interchangeably), $I(Y_1; Y_2)$, can be taken as a measure of how close these two nodes are to each other. The mutual information between two sets, $I\left(\boldsymbol{Y}^{(1)}; \boldsymbol{Y}^{(2)}\right)$—for example, in which Set (1) contains the symptoms of depression and Set (2) contains the symptoms of generalized anxiety—can be taken as a measure of closeness

287

12. Discussion: The Road Ahead

between two groups of nodes. Furthermore, the mutual information of one variable, (Y_1), with respect to all other variables, $(Y_{-(1)})$ and $I(Y_1; Y_{-(1)})$, can be used as a centrality measure. Finally, when temporal dependence is present, only computing the information one node has on all nodes at the next measurement, $I(Y_1; Y_{t+1})$, can be taken as a temporal centrality measure, which takes into account both the contemporaneous and temporal network (see Chapter 5 and Chapter 6).

When $\boldsymbol{Y}$ has a multivariate normal distribution with zero $\boldsymbol{\mu}$ and variance–covariance matrix $\boldsymbol{\Sigma}$, as is the case in both the GGM and graphical VAR model, the differential entropy becomes (Cover & Thomas, 2012, p. 250):

$$h(\boldsymbol{Y}) = \frac{1}{2}\log_2\left((2\pi e)^J \cdot |\boldsymbol{\Sigma}|\right).$$

As can be seen, this measure is a direct function of the size of the variance–covariance matrix $\boldsymbol{\Sigma}$. This expression allows us to compute all mutual informations described above. For example, the mutual information between two variables can be computed as:

$$I(Y_1; Y_2) = -\frac{1}{2}\log_2\left(1 - \rho_{12}^2\right).$$

As can be seen, this measure is a direct property of the explained variance ρ_{12}^2 between two variables. The mutual information of one variable with all other variables becomes:

$$I(Y_1; Y_{-(1)}) = \frac{1}{2}\log_2\left(\frac{|\boldsymbol{\Sigma}_{(11)}|}{|\boldsymbol{\Sigma}|}\right),$$

in which $\boldsymbol{\Sigma}_{(11)}$ denotes the variance–covariance matrix without row and column 1. Finally, in a stationary time series of multivariate normal data (e.g., a Lag-1 VAR model), the temporal closeness above becomes:

$$I\left(y_{t,j}; \boldsymbol{y}_{t+1}\right) = \frac{1}{2}\log_2\left(\frac{\det\left(\boldsymbol{\Sigma}_{\mathrm{TP}}^{(1)}\right)}{|\boldsymbol{\Sigma}_{\mathrm{TP}}^{(2)}|}\right),$$

in which $\boldsymbol{\Sigma}_{\mathrm{TP}}^{(1)}$ denotes a subsetted Toeplitz matrix:

$$\boldsymbol{\Sigma}_{\mathrm{TP}}^{(1)} = \mathrm{Var}\left(\boldsymbol{Y}_{t:t}, \boldsymbol{Y}_{t+1}\right),$$

which can be obtained from a VAR analysis.

Applying these metrics to the networks shown in Figure 12.1 leads to strikingly different interpretations. In Panel A, Nodes A and C are now

288

12.2. Open Questions in Network Psychometrics

shown to be independent; thus, Node B does not have a problematic interpretation of having a high betweenness. Now in Panel C,[3] Node A is shown to have slightly more information over the next time point than Node B, even though Node B has more temporal connections. Information theory is a promising gateway to analyzing the network models obtained. However, future researchers must thoroughly test and validated these metrics on psychological data.

The Importance of Intercepts

The network models outlined in this dissertation are all models of second-order moments. That is, they model variances and covariances but not expected values. The parameters that do model the expected value in the GGM, Ising model, and VAR model are the intercepts. When drawing a network, these are ignored. As such, when using the network structure to compute centrality, for example, intercepts are not taken into account. The problem with this approach is that links are formed between variables that may have largely different intercepts. Particularly in models of binary variables, links may be formed between variables that never have high entropy at the same time. For example, whenever a person is in danger of suffering from suicidal ideation (high entropy), we might expect that person to always experience sadness (low entropy). If we apply a virus-spreading analogy (Borsboom et al., 2011), such nodes would never "infect" each other; the link would never be used.

Figure 12.2, Panel A, shows an example of a network we might estimate on educational data. This network is a fully connected Ising model, also called a Curie-Weiss model, which is known to be equivalent to the IRT model shown in Panel B (Marsman et al., 2015; see also Chapter 8). As such, IRT models are often used and work well on educational data, the Ising model of Panel A is not measurable. We can see a link between the items "1 + 1" and "0.07692 + 0.3409." This link represents a very plausible predictive relationship. Knowing someone can answer "0.07692 + 0.3409" tells us that person can also answer "1 + 1" (because the person is able to answer the much harder question as well), and knowing someone cannot answer "1 + 1" also tells us that person cannot answer "0.07692 + 0.3409" (because that person cannot answer the simpler question). These items likely do not have high entropy at the same time in any person's life. This means that when a person is struggling with "1 + 1," and may or may not answer this item correctly (high entropy), it is highly likely that this

[3] Panel B only shows half the information needed to characterize the full likelihood; see Chapter 6.

289

12. Discussion: The Road Ahead

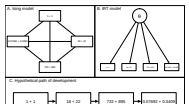


Figure 12.2: Three potential network structures on educational data. Panels A and B are equivalent models that do not show the difficulty of an item in Panel C. Items are ordered according to their difficulty, pointing from the easier item to the more difficult item.

person will never be able to answer "0.07692 + 0.3409" correctly (always an incorrect response; low entropy). Conversely, whenever a person correctly answers the item "0.07692 + 0.3409," that person will likely always answer the item "1 + 1" correctly.

The network perspective would lead to the following interpretation of Panel A, that one could influence the probability of correctly answering "0.07692 + 0.3409" by training someone on the question "1 + 1." However, this sense unlikely. Teaching a person the techniques needed to answer "1 + 1" would not prepare that person to answer "0.07692 + 0.3409," which requires knowledge of decimal points, counting over tens, and so forth. The latent variable model in Panel B would implicate that training someone on one of the items would never help that person answer other items correctly: Although I do not wish to argue against mathematical ability, I do think that such an assumption might also be too strict. Children learn by making items, and learning how to make one item helps a child make another item. This learning, however, does not jump wildly as would be expected from

290

12.2. Open Questions in Network Psychometrics

Panel A, rather it follows a straight path. Learning "1 + 1" helps to answer the item "2 + 2," which helps to answer the item "3 + 3", and so forth.

Panel C shows a network that is based only on intercepts rather than covariances. Here, each item points to the first harder item. This very different network structure shows that people first learn "1 + 1," then "10 + 22," and so forth. I term such a network structure a *path of development*. The network shown in Panel C is merely a hypothetical example of what a network that also takes intercepts into account could look like. In perfect unidimensional cases, it might look like Panel C, whereas in multidimensional cases, one could envision, for example, parallel paths or path splitting. I do not seek to propose a new modeling framework in this section but merely wish to highlight that taking intercepts into account could lead to different ways of investigating the phenomena of interest.

Complexity

The network models, as outlined in this dissertation, are but one of the many consequences that may come from a more general hypothesis of *complexity*. Psychological behavior plausibly is the result of emergent behavior in a complex system of interacting psychological, biological, and sociological components. Simply stated, psychology is complex. People's behavior is dictated by their brains, which consist of billions of neurons, formed by many years of development. As such, every person is a mini universe of complexity. These universes, in turn, interact with one another in complicated social networks. Perhaps, psychology is one of the hardest fields to tackle. It is, in my opinion, only logical that many behaviors have no simple explanation.

The network model is but one attempt at grasping this complexity; we should not get sidetracked by believing it is the only possible attempt. The hypothesis of complexity is not limited to the expectations that data are generated due to an underlying (sparse) network model of, at most, second-order moments. This hypothesis reaches further, with many more implications. This point of view can take psychological research in many different directions—rather than merely the estimation of network models. For example, long-term predictions can be made on the effects of intervention, without understanding the true underlying causal mechanisms. Also of particular importance is the work done by van de Leemput et al. [2014] and Wichers et al. [2016] on identifying early warning signals for phase transitions in psychology, such as the onset of depression. I think that the hypothesis of complexity has much to offer in the years to come and will change psychological research in ways we cannot imagine now.

291

12. Discussion: The Road Ahead

12.3 Conclusion

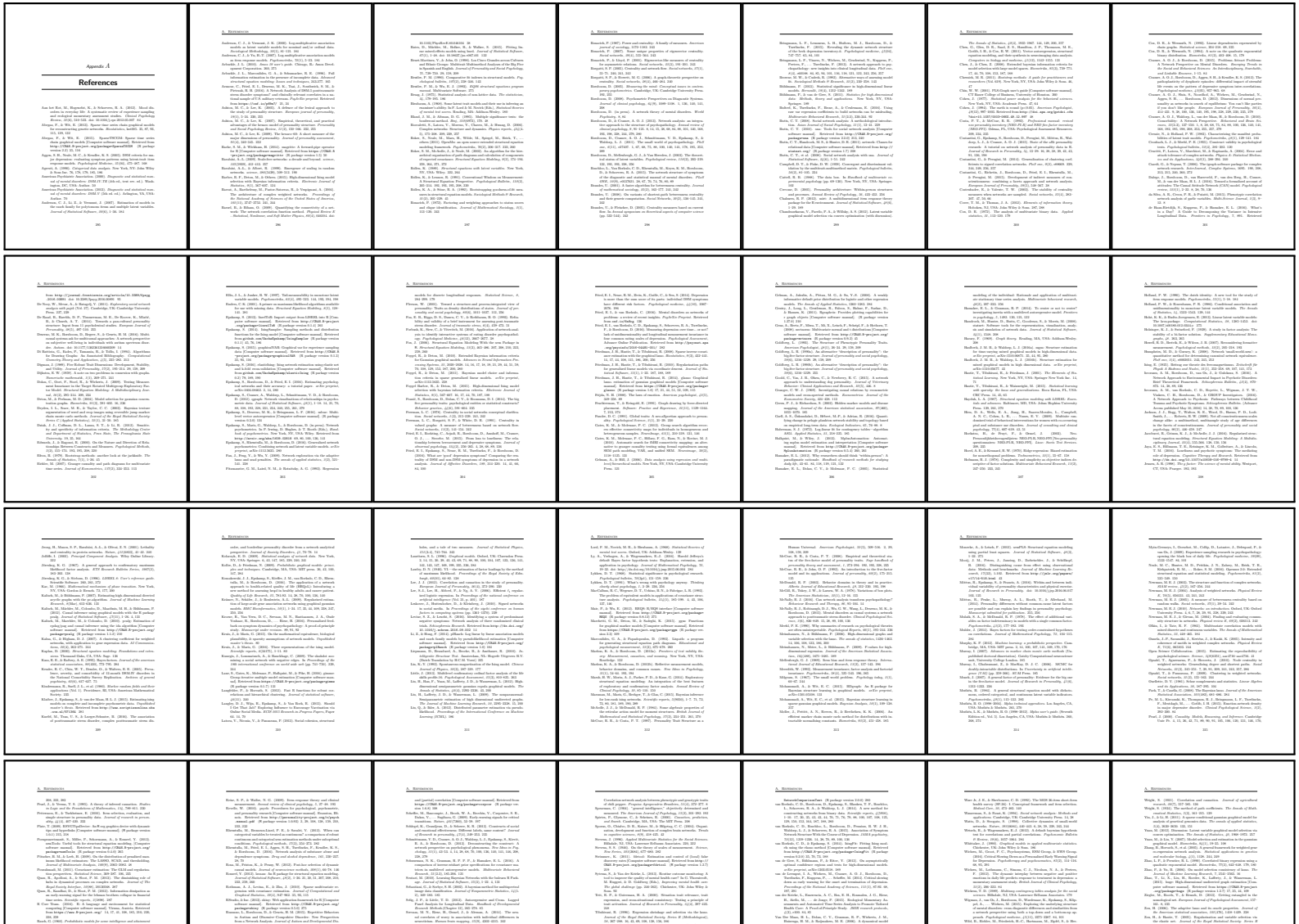
In this discussion, I outlined various topics for future research which can be tackled in network psychometrics: improving centrality measures, handling missing and ordinal data, quantifying evidence for sparsity in the network, and incorporating intercepts in inference on these models. This is just a highlight of several future directions: many more can be conceived, such as tackling heterogeneity, improving multilevel estimation of contemporaneous effects, handling a mixture of observational and experimental data, and extending networks to nonlinear dynamics. Finally, I noted a far more general field of research—complexity in psychology—of which network modeling is merely a small part. The network models proposed in this dissertation add much to the toolbox of psychological and psychometric researchers. Network psychometrics, however, is still a young field of research with many unanswered questions. The full utility of these methods and their place in psychological research and psychometrics will be determined on the road ahead.

292

- Missing data
- Ordinal data
- Evidence for Sparsity
  - Bayesian statistics
- Should probabilistic models be analyzed as networks?
  - Information theory
- Importance of intercepts
- Complexity

Part V

Appendices



Appendix B

Contributed Work

This appendix contains a list of the contributed publications and software packages until the completion of this PhD project. The PDF version of the dissertation contains links to all publications and software repositories, which can also be found on my website at:

`www.sachaepskamp.com`

B.1 Publications

This section lists all manuscripts that are submitted or accepted for publication at the time of submission of this dissertation. As such, some of the titles and journals of yet unpublished work might change before publications as a result of the review process.

Main Author Publications

- **Epskamp, S.**, and Fried, E.I. A Tutorial on Regularized Partial Correlation Networks (pre-print archive link).
  - Chapter 2
  - Submitted to *Psychological Methods*
- **Epskamp, S.** , Regularized Gaussian Psychological Networks: Brief Report on the Performance of Extended BIC Model Selection (pre-print archive link).
  - Section 2.6

  - Under review at *Network Science*
- **Epskamp, S.**, Borsboom, D., and Fried, E.I. Estimating Psychological Networks and their Accuracy: A Tutorial Paper (pre-print archive link).
  - Chapter 3 and Chapter 1
  - Revision under review at *Behavior Research Methods*
- **Epskamp, S.**, Kruis, J., and Marsman, M. Estimating psychopathological networks: be careful what you wish for (pre-print archive link).
  - Chapter 4
  - Revision under review at *PlosOne*
- **Epskamp, S.**, van Borkulo, C.D., van der Veen, D.C., Servaas, M.N., Isvoranu, A.M., Riese, H., and Cramer, A.O.J. Personalized Network Modeling in Psychopathology: The Importance of Contemporaneous and Temporal Connections.
  - Chapter 5
  - Submitted to *Psychological Medicine*
- **Epskamp, S.**, Waldorp, L.J., Möttus, R., and Borsboom, D. Discovering Psychological Dynamics: The Gaussian Graphical Model in Cross-sectional and Time-series Data (pre-print archive link).
  - Chapter 6
  - Under review at *Multivariate Behavioral Research*
- **Epskamp, S.**, Rhemtulla, M., and Borsboom, D. Generalized Network Psychometrics: Combining Network and Latent Variable models (pre-print archive link).
  - Chapter 7
  - Accepted for publication in *Psychometrika*
- **Epskamp, S.**, Maris, G., Waldorp, L. J., and Borsboom, D. (in press). Network Psychometrics. To appear in: Irwing, P., Hughes, D., and Booth, T. (Eds.), *Handbook of Psychometrics*. New York: Wiley (link).
  - Chapter 8

- **Epskamp, S.**, Cramer, A.O.J., Waldorp, L.J., Schmittmann, V.D. and Borsboom, D. (2012) qgraph: Network Visualizations of Relationships in Psychometric Data. *Journal of Statistical Software, 48(4)*, 1-18 (link).
  - Chapter 9
- Costantini, G., **Epskamp, S.** (the first two authors contributed equally to this work), Borsboom, D., Perugini, M., Möttus, R., Waldorp, L. J., and Cramer, A. O. (2014). State of the aflt personality research: A tutorial on network analysis of personality data in R. *Journal of Research in Personality, 54*, 13-29 (link).
  - Chapter 10
- **Epskamp, S.** (2015). semPlot: Unified visualizations of Structural Equation Models. *Structural Equation Modeling, 22*, 747-757 (link).
  - Chapter 11
- Van Bork, R., **Epskamp, S.** (the first two authors contributed equally to this work), Rhemtulla, M., Borsboom, D., and Van der Maas, H.L.J. What is P? Some Risks of General Factor Modeling.
  - Under review at *Theory and Psychology*

Collaborations

- Golino, H., **Epskamp, S.**. Exploratory graph analysis: a new approach for estimating the number of dimensions in psychological research.
  - Under review at *PlosOne*
- Wagenmakers, E-J., Marsman, M., Jamil T., Ly, A., Verhagen, J., Love, J., Selker, R., Gronau, Q.F., Smiras, M., **Epskamp, S.**, Matzke, D., Rouder, J.N., Morey, R.D. Bayesian Inference for Psychology. Part I: Theoretical Advantages and Practical Ramifications
  - Under review at *Psychonomic Bulletin & Review*
- Tio, P., **Epskamp, S.**, Noordhof, A., and Borsboom, D. (2016). Mapping the manuals of madness: Comparing the ICD-10 and DSM-IV-TR using a network approach. International Journal of Methods in Psychiatric Research (link).

- Fried, E. I., van Borkulo, C. D., **Epskamp, S.**, Schoevers, R. A., Tuerlinckx, F., and Borsboom, D. (2016). Measuring Depression Over Time. . or not? Lack of Unidimensionality and Longitudinal Measurement Invariance in Four Common Rating Scales of Depression. *Psychological Assessment*. Advance online publication (link).
- **Open Science Collaboration**. (2015). Estimating the reproducibility of psychological science. *Science, 349 (6251)* (link).
  - I setup and coordinated the "analysis audit" in which a team of students worked on replicating all analyses done in R.
- Wichers, M., Groot, P.C., **Psychosystems**, ESM Group, EWS Group (2016). *Psychotherapy and Psychosomatics, 85*: 114-116 (link).
- Nuijten, M.B., Hartgerink, C. H. J., van Assen, M. A. L. M., **Epskamp, S.**, and Wicherts, J. M. (2015). The prevalence of statistical reporting errors in psychology (1985-2013). *Behavior Research Methods* (link).
- Fried, E. I., **Epskamp, S.**, Nesse, R.M., Tuerlinckx, F., and Borsboom, D. (2016). What are 'good' depression symptoms? Comparing the centrality of DSM and non-DSM symptoms of depression in a network analysis. *Journal of Affective Disorders, 189*: 314-320 (link).
- Kossakowski, J. J., **Epskamp, S.**, Kieffer, J. M., van Borkulo, C. D., Rhemtulla, M., and Borsboom, D. (2015). The application of a network approach to health-related quality of life (HRQoL): Introducing a new method for assessing HRQOL in healthy adults and cancer patient. *Quality of Life Research, 0962-9343*, 1-12 (link).
- Langley, D. J., Wijn, R., **Epskamp, S.**, and Van Bork, R. (2015). Should I Get That Jab? *ECIS 2015 Research-in-Progress Papers*, Paper 64. (link).
- Wigman, J. T. W., Van Os, J., Borsboom, D., Wardenaar, K. J., **Epskamp, S.**, Klippel, A, MERGE, Viechtbauer, W., Myin-Germeys, I., and Wichers, M. (2015). Exploring the underlying structure of mental disorders: cross-diagnostic differences and similarities from a network perspective using both a top-down and a bottom-up approach. *Psychological Medicine, 45*, 2375-2387 (link).
- Fried, E. I., Bockting, C., Arjadi, R., Borsboom, D., Amshoff, M., Cramer, A. O. J., **Epskamp, S.**, Tuerlinckx, F., Carr, D., and

Stroebe, M. (2015). From loss to loneliness: The relationship between depressive symptoms and bereavement. *Journal of Abnormal Psychology, 124*, 256-65 (link).

- van Borkulo, C. D., Borsboom, D., **Epskamp, S.**, Blanken, T. F., Boschloo, L., Schoevers, R. A., and Waldorp, L. J. (2014). A new method for constructing networks from binary data. *Scientific Reports, 4*, 5918. (link).
- Borsboom, D., Cramer, A.O.J., Schmittmann, V.D., **Epskamp, S.** and Waldorp L.J. (2011) The Small World of Psychopathology. *PLoS ONE 6, e27407*. (link).
- Borsboom, D., **Epskamp, S.**, Kievit, R.A., Cramer, A.O.J., and Schmittmann,V.D. (2011). Transdiagnostic networks. *Perspectives on Psychological Science, 6*, 610-614. (link).
- Schmittmann, V.D., Cramer, A.O.J., Waldorp, L.J., **Epskamp, S.**, Kievit, R.A. and Borsboom, D. (2011). Deconstructing the construct: Anetwork perspective on psychological phenomena. *New Ideas in Psychology, 31*, 43-53. (link).

B.2 Software

Stable R packages

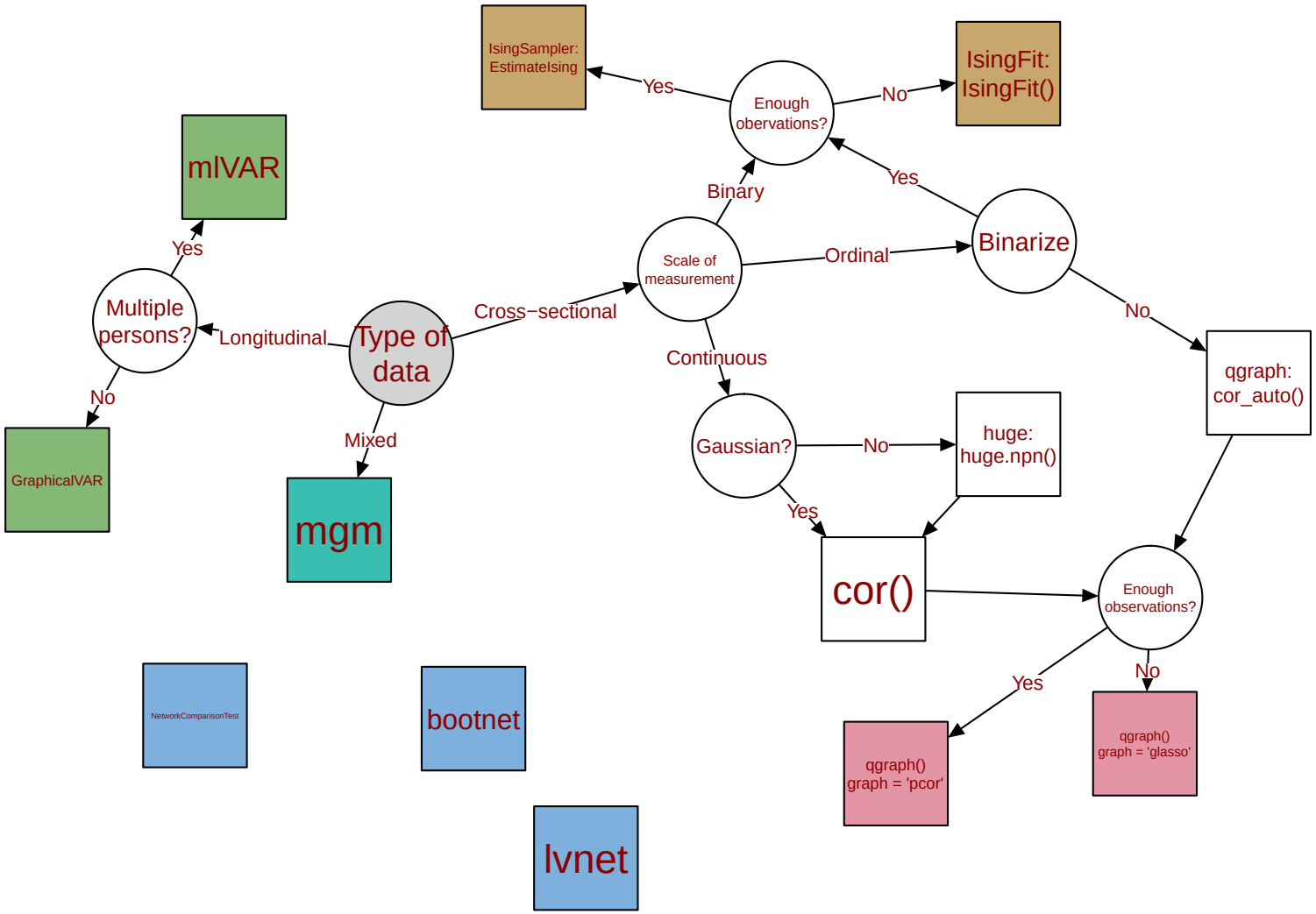
- **qgraph**
  - Network drawing, construction and estimation and network-based data visualization (Link to CRAN repository)
- **semPlot**
  - Path diagrams and visual analysis of various SEM packages' output (Link to CRAN repository)
- **IsingSampler**
  - Sampling methods and distribution functions for the Ising model (Link to CRAN repository)
- **lisrelToR**
  - Import output from LISREL into R (Link to CRAN repository)

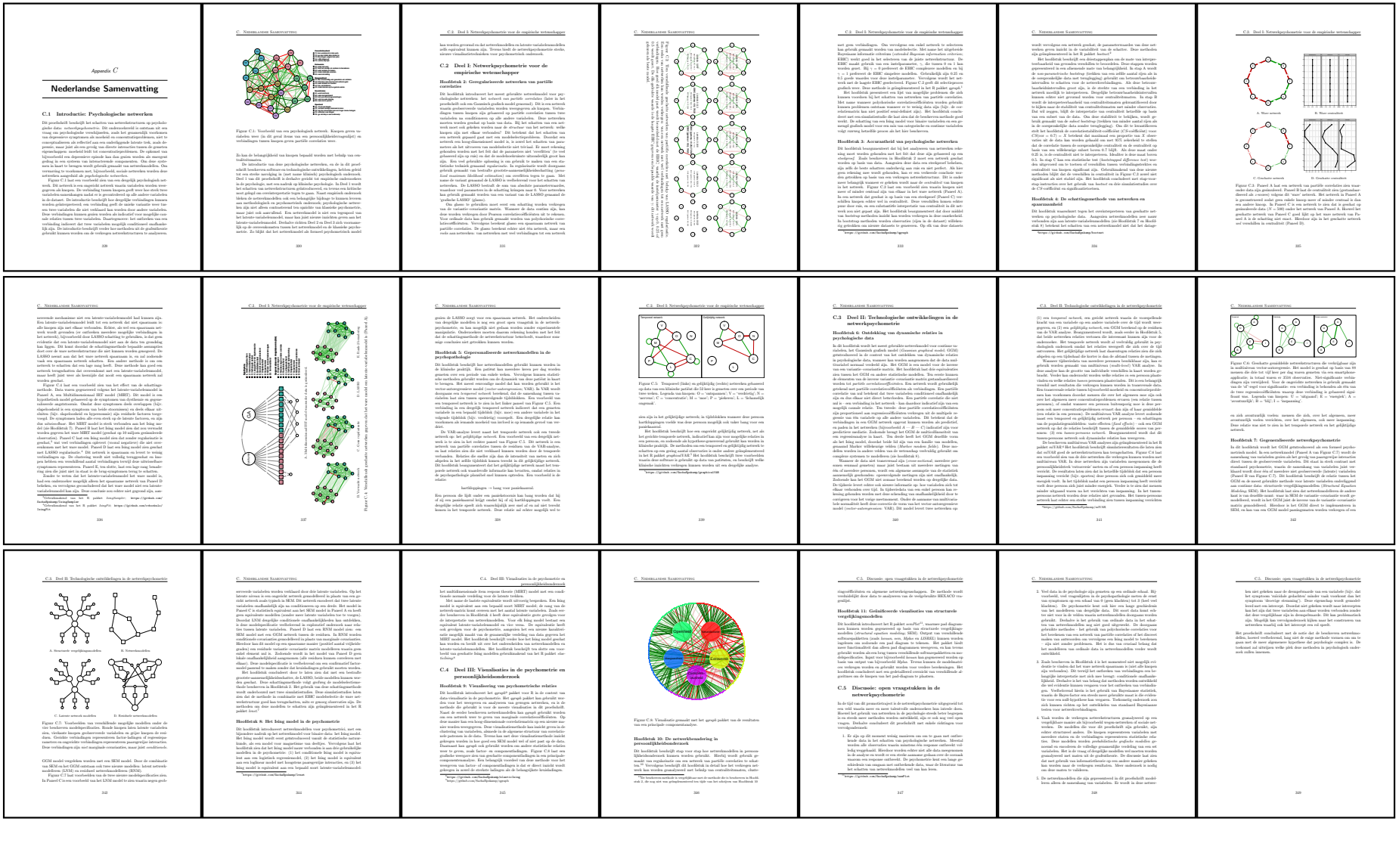
- **graphicalVAR**
  - Estimate temporal and contemporaneous effects on $N = 1$ longitudinal data (Link to CRAN repository)
- **mlVAR**
  - Multi-level vector autoregression (Link to CRAN repository)
- **bootnet**
  - General robustness tests and plots for network models (Link to CRAN repository)
- **elasticIsing**
  - Ising model estimation using elastic net regularization (Link to CRAN repository)
- **lvnet**
  - Latent variable network modeling (Link to CRAN repository)

Collaborations

- **IsingFit**
  - Fitting Ising models using the eLasso method (Link to CRAN repository)
- **statcheck**
  - Fitting Ising models using the eLasso method (Link to CRAN repository)
- **JASP**
  - A low fat alternative to SPSS, a delicious alternative to R. (Link to website)

The Psychosystems Ecosystem




Appendix C
Nederlandse Samenvatting
C.1 Introductie: Psychologische netwerken
C.2 Deel I: Netwerkpsychometrie voor de empirische wetenschapper
Hoofdstuk 2: Geregulariseerde netwerken van partiële correlaties
C.3 Deel II: Technologische ontwikkelingen in de netwerkpsychometrie
Hoofdstuk 6: Ontdekking van dynamische relaties in psychopathologie
C.4 Deel III: Visualisaties in de psychometrie en persoonlijkheidsonderzoek
Hoofdstuk 9: Visualisering van psychometrische relaties
Hoofdstuk 10: De netwerkbenadering in persoonlijkheidsonderzoek
C.5 Discussie: open vraagstukken in de netwerkpsychometrie

Thank you for your attention!