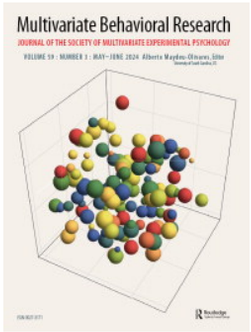

Multivariate Behavioral Research

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/hmbr20

Subgrouping with Chain Graphical VAR Models

Jonathan J. Park, Sy-Miin Chow, Sacha Epskamp & Peter C. M. Molenaar

To cite this article: Jonathan J. Park, Sy-Miin Chow, Sacha Epskamp & Peter C. M. Molenaar (2024) Subgrouping with Chain Graphical VAR Models, Multivariate Behavioral Research, 59:3, 543-565, DOI: 10.1080/00273171.2023.2289058

To link to this article: https://doi.org/10.1080/00273171.2023.2289058

Published online: 13 Feb 2024.

MULTIVARIATE BEHAVIORAL RESEARCH
2024, VOL. 59, NO. 3, 543–565
https://doi.org/10.1080/00273171.2023.2289058

Subgrouping with Chain Graphical VAR Models

Jonathan J. Park[a], Sy-Miin Chow[a], Sacha Epskamp[b], and Peter C. M. Molenaar[a]

[a]Department of Human Development and Family Studies, The Pennsylvania State University; [b]Department of Psychology, National University of Singapore

ABSTRACT

Recent years have seen the emergence of an "idio-thetic" class of methods to bridge the gap between nomothetic and idiographic inference. These methods describe nomothetic trends in idiographic processes by pooling intraindividual information across individuals to inform group-level inference or vice versa. The current work introduces a novel "idio-thetic" model: the subgrouped chain graphical vector autoregression (scGVAR). The scGVAR is unique in its ability to identify subgroups of individuals who share common dynamic network structures in both lag(1) and contemporaneous effects. Results from Monte Carlo simulations indicate that the scGVAR shows promise over similar approaches when clusters of individuals differ in their contemporaneous dynamics and in showing increased sensitivity in detecting nuanced group differences while keeping Type-I error rates low. In contrast, a competing approach—the Alternating Least Squares VAR (ALS VAR) performs well when groups were separated by larger distances. Further considerations are provided regarding applications of the ALS VAR and scGVAR on real data and the strengths and limitations of both methods.

KEYWORDS

Community detection; dynamic network modeling; vector autoregression; psychopathology

1. Introduction

The divide between idiographic and nomothetic inference in the behavioral sciences is far from new and stems from the early days of scientific psychology or even earlier (Hamaker, 2004; Lundh, 2015; Molenaar, 2004). During the mid-20th century, a series of debates took place regarding the placement of psychology as an idiographic or nomothetic science (see Runyan, 1983). Recent years have witnessed a surge of methods, referred to as "idio-thetic" methods (Beck & Jackson, 2020), that integrate individual and group (population)-level information within the same modeling framework. Examples of idio-thetic methods include mixed-effects modeling, latent class analysis (McCutcheon, 1987), and mixture modeling (Liu et al., 2020; Muthen, 2001). This paper introduces a novel idio-thetic extension to the chain graphical vector autoregression model (cGVAR; Epskamp, Waldorp, et al., 2018), termed the subgrouped chain graphical VAR (scGVAR). The scGVAR utilizes a community detection approach to identify subgroup-level models in addition to the group- and individual-level results produced by the cGVAR. As a result, the scGVAR identifies homogeneous subgroups within potentially heterogeneous samples which exhibit distinct dynamics in the GVAR framework. Unlike related work, the scGVAR is novel in how it leverages information to identify subgroups and employs regularization techniques to enhance the power of detecting shared dynamic patterns among subgroups, even in the presence of characteristics that dominate the sample.

Network models have significantly impacted the social and behavioral sciences over the last decade, reshaping the study of human behavior and psychology by encouraging researchers to consider psychological constructs from different perspectives (Borsboom & Cramer, 2013). More recently, dynamic network models have gained traction, particularly in the analysis of intensive longitudinal data (i.e., data where $T > 50$ time-points; Beck & Jackson, 2020; Park et al., 2021). These dynamic network approaches have found diverse applications, ranging from exploring major depressive disorder (MDD; De Vos et al., 2017; Lydon-Staley et al., 2019) and emotion dynamics (Park et al., 2021; Wright et al., 2019) to investigating ADHD (Gates et al., 2014). These approaches align with the call to attend to idiographic phenomena

CONTACT Jonathan J. Park jjp409@psu.edu The Pennsylvania State University, 029 Health and Human Development Building, University Park, PA 16802, USA

(e.g., A. J. Fisher, 2015; Molenaar, 2004). That is, many of these approaches begin with individual model fitting to each individual's data in turn (e.g., Hamaker et al., 2005), followed by integrated procedures to identify and summarize homogeneous group/subgroup ("nomothetic") structures from individual modeling results to circumvent the lack of generalizability of $N = 1$ results to the populations or sub-populations at large (e.g., Bulteel et al., 2016; Epskamp, Waldorp, et al., 2018; Lane et al., 2019).

As distinct from other idio-thetic methods such as mixed, latent class, and mixture modeling, beginning with $N = 1$ analysis allows for more in-depth explorations of each individual's person-specific dynamic network structures prior to pooling information across individuals or subgroups. However, myriad issues remain unresolved. One critical issue resides in many current dynamic network approaches' low power in detecting subgroups when a substantial number of network paths overlapping with each other and are only differentiable from a limited number of network paths. For example, Henry et al. (2019) utilized idio-thetic methods to analyze a sample of individuals with ADHD and ASD and uncovered distinct connections between specific regions of interest (ROIs) of the brain that characterized each disorder. Notably, despite these unique dynamics within each group, a substantial majority of connections were found to be shared across the entire sample, spanning both ADHD and ASD populations.

1.1. Idio-thetic dynamic network clustering methods

Many contemporary dynamic network clustering algorithms operate by identifying subgroups of individuals with similar over-time relations among a series of variables, as captured by a dynamic model of choice. In this article, we propose a dynamic network clustering approach that focuses on the vector autoregressive (VAR) model as the operating model. The VAR model consists of *structural paths* that delineate the lagged relations among a series of manifest variables between times t and $t-1$ when of a lag-order of 1; or VAR(1). This is in contrast to clustering approaches based on alternative (but equivalent) structural VAR formulations (Gates et al., 2017; Kim et al., 2007; Lütkepohl, 2005) that target "contemporaneous" patterns of associations; namely, patterns that span a shorter duration than what may be statistically discerned at the next observed time point. VAR-based clustering approaches offer several advantages that contribute to their widespread adoption across various scientific domains such as relatively straightforward identification constraints in contrast to related models such as the structural VAR (see Gates et al., 2010). Additionally, the availability of tools and techniques to easily implement the VAR process, including the integration of regularization methods and handling high-dimensional data, further enhances their appeal and applicability in various research settings. Indeed, several existing dynamic network models are built on the VAR framework. Such as the multilevel VAR (Bringmann et al., 2013), the multi-var (Z. F. Fisher et al., 2022), S-GIMME (Gates et al., 2017; Gates & Molenaar, 2012), and—more recently—the alternating least squares VAR (ALS VAR; Bulteel et al., 2016; Takano et al., 2021). These idio-thetic approaches all vary in the extent and procedures through which group- and individual-specific paths are identified, evaluated, and retained.

The ALS VAR is one of the benchmark idio-thetic methods that identify subgroups of individuals based on their VAR structures. The ALS VAR has been shown to yield relatively accurate recovery of parameters in simulation studies, with several useful properties such as the ability to separate groups which differ by the magnitude of their parameters (e.g., effect sizes; Bulteel et al., 2016; Takano et al., 2021). A weakness of this approach—however—is that the ALS VAR has some difficulty in separating subgroups which share a large degree of commonalities but differ in a few specific structures (Takano et al., 2021). The scGVAR—discussed in full below—identifies subgroups of individuals who share common dynamic network structures prior to fitting subgroup-level models. However, the scGVAR is designed to provide greater sensitivity in identifying clusters of individuals with minor differences in dynamic network structures, made possible by focusing on the absence/presence of structural paths, with greater sparsity induced through regularization. In addition, contemporaneous information can be incorporated when conducting community detection.

Figure 1(A) demonstrates some configurations where the ALS VAR may perform well in enumerating subgroups of subjects. That is, the ALS VAR can differentiate groups which differ by the magnitude of their edges; specifically (Figure 1(A); Takano et al., 2021). This is well suited, for instance, in applications comparing clinical samples to controls where past research has suggested denser connectivity among the clinically depressed relative to healthy controls (e.g., Lydon-Staley et al., 2019) and is a feature unique to

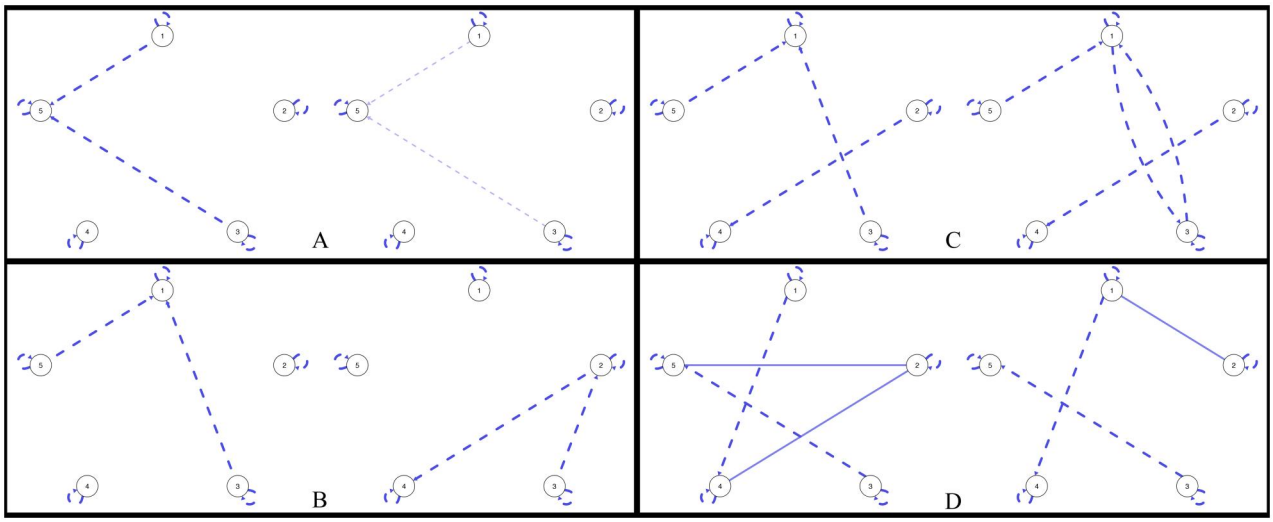

Figure 1. Optimal settings for the ALS VAR (A & B) and the scGVAR (C & D). Networks represent a five-variate VAR(1) model. Dashed edges represent lagged effects while solid edges indicate contemporaneous effects. Edges which return to their node indicate autoregressive effects. Faded edges are weaker connections between nodes. The ALS VAR outperforms the scGVAR in cases where: (A) subgroups differ based on the magnitude of effects; note that the structure is the same but the right network is substantially weaker and (B) subgroups are differentiated by a large degree of differences. The scGVAR performs better than the ALS VAR when (C) subgroups differ by a small number of unique features and (D) subgroups differ only by their contemporaneous effects.

the ALS VAR relative to other idio-thetic methods. Additionally, the ALS VAR performs well when groups vary by a large number of their dynamics over time (Figure 1(B)). However, the ALS VAR may have difficulty separating subgroups which share many common edges but differ in a few, key connections but is a scenario where our scGVAR would be advantageous in Figure 1(C) such as in prior work showing minor structural differences between subgroups (e.g., Henry et al., 2019; Price et al., 2017). Additionally, when subgroups may differ due to some combination of lagged- or contemporaneous-effects, the scGVAR—in its unique structuring of contemporaneous effects—will perform well. In contrast, the ALS VAR does not explicitly model contemporaneous effects and is not expected to perform as well.

The remainder of this article is organized as follows. First, we formally introduce the VAR(1) model as the foundational model behind both the scGVAR and the ALS VAR prior to formally introducing the algorithms in the following sections. Second, we introduce a simulation study designed to highlight the unique strengths and weaknesses of both the scGVAR and the ALS VAR. We then conclude with an empirical example where both algorithms are applied and results are interpreted to help guide substantive researchers in implementing either algorithm.

2. Vector autoregression

We begin this section by detailing the standard vector autoregression or VAR model which forms the basis of the graphical VAR (GVAR), its extensions, and the ALS VAR. In the VAR, past values on measures predict present values at a given lag order. For example, a backward shift by 1-unit of time would be known as a lag-1 VAR or VAR(1) and can be written in matrix notation as:

$$\boldsymbol{\eta}_{i,t} = \boldsymbol{\mu}_i + \boldsymbol{\Phi}_i(\boldsymbol{\eta}_{i,t-1} - \boldsymbol{\mu}_i) + \boldsymbol{\zeta}_{i,t} \tag{1}$$

where $\boldsymbol{\eta}_{i,t}$ represents the p-variate vector of scores on p items at a given time, t for a given individual, i, $\boldsymbol{\mu}$ represents a vector of means on the p variables, $\boldsymbol{\Phi}_i$ is the $p \times p$ matrix of lagged auto- and cross-regressions, and $\boldsymbol{\zeta}$ is the p-variate vector of innovations with mean of zero and a time-invariant $p \times p$ covariance, $\boldsymbol{\Psi}_i$:

$$\boldsymbol{\zeta}_{i,t} \sim N(0, \boldsymbol{\Psi}_i) \tag{2}$$

A GVAR model is then simply the estimation of an additional network of contemporaneous effects upon the residual covariance matrix. Formally, this is accomplished by fitting a Gaussian Graphical Model (GGM) on the variance-covariance matrix of the residuals. This yields the precision matrix, $\mathbf{K} = \boldsymbol{\Psi}^{-1}$, of the VAR model. A diagonal structure for $\mathbf{K}$ implies a diagonal structure for $\boldsymbol{\Psi}$, and thus conditional independence among the nodes (Epskamp, van Borkulo, et al., 2018; Park et al., 2021). Notably, the GGM

assumes multivariate normality of the variables and violations of this assumption can prevent GGMs from properly characterizing the joint likelihood function which may result in biases (Epskamp, Waldorp, et al., 2018). Alternative graphical models may be applied when the form of the data differ from standard continuous variables such as the Ising model for binary data or a mixed graphical model for categorical and Poisson-distributed variables (e.g., Haslbeck & Waldorp, 2015; Van Borkulo et al., 2014). Further standardization can be applied to the matrices of the VAR model to obtain coefficients of the partial directed correlations (PDCs) and the partial contemporaneous correlations (PCCs) which correspond to partialed and standardized lagged effects and the contemporaneous residual covariances in the VAR, respectively with the equations:

$$PCC_{(\eta_m,\eta_n)} = -\frac{k_{mn}}{\sqrt{k_{mm}k_{nn}}} \tag{3}$$

$$PDC_{(\eta_m,\eta_n)} = \frac{\phi_{mn}}{\sqrt{\psi_{mm}k_{nn} + \phi_{mn}^2}} \tag{4}$$

where k_{mn} is the $(m,n)^{th}$ cell and k_{nn} is the n^{th} diagonal cell of the $\mathbf{K}$ matrix, ϕ_{mn} is the lag(1) cross-regression weight of the m^{th} variable onto the n^{th} variable, ψ_{mm} is the variance of the m^{th} element of $\mathbf{\Psi}$.

A cGVAR may then produced by "chaining" or stacking together the time-series of multiple subjects to create a pooled dataset and fitting a GVAR model to the entire chained time-series to produce a set of group-based estimates across all subjects.

2.1. Estimation of graphical VAR models

Given our interest in detecting shared subgroup paths in high-dimensional networks, use of regularization routines to simplify network structure is important to prevent over-fitting. For the current work, we focus on GVAR models estimated with the" least absolute shrinkage and selection operator" (LASSO; Tibshirani, 1996). LASSO is a method for variable selection typically used in instances where the number of parameters, p, is relatively large compared to sample size, N; particularly in cases where p exceeds N.

In these instances, for the same p, smaller N will yield sparser structures with coefficients" pulled" toward zero. LASSO is incorporated into the GVAR using the multivariate regression with covariance estimation (MRCE; Abegaz & Wit, 2013; Rothman et al., 2010). The MRCE algorithm optimizes both $\mathbf{\Phi}_i$ and $\mathbf{K}_i$ matrices using cyclical-coordinate descent and the graphical LASSO, respectively (Abegaz & Wit, 2013; Epskamp, van Borkulo, et al., 2018; Rothman et al., 2010). It has been shown in simulation studies to outperform similar regularization methods when residual covariances are relatively high (Rothman et al., 2010). These qualities make it preferable for estimation of GVAR models where covariances are expected to exist in the form of the partial contemporaneous correlations. The MRCE solves for sparse estimates of $\mathbf{\Phi}_i$ and $\mathbf{K}_i$. This is done by adding two penalties to the negative log-likelihood function: $g(.)$. This allows us to obtain a sparse estimates of $\mathbf{\Phi}_i$ depending on $\mathbf{K}_i$ and vice-versa:

$$(\hat{\mathbf{\Phi}}_i, \hat{\mathbf{K}}_i) = \underset{\mathbf{\Phi}_i,\mathbf{K}_i}{\arg\min} \left\{ g(\mathbf{\Phi}_i, \mathbf{K}_i) + \lambda_1 \sum_{j' \neq j} |k_{j'j}| + \lambda_2 \sum_{j=1}^{p} \sum_{k=1}^{q} |\phi_{jk}| \right\} \tag{5}$$

where λ_1 and λ_2 are tuning parameters for the $\mathbf{K}_i$ and $\mathbf{\Phi}_i$ matrices, respectively. Note, the subscripts j and k denote the row and column elements of $\mathbf{\Phi}$. It is noted by Rothman et al. (2010) that this optimization problem is not convex. However, solving for $\mathbf{\Phi}_i$ or $\mathbf{K}_i$ while the other is fixed renders the problem convex. Thus, they propose a multi-stage algorithm where $\mathbf{\Phi}_i$ is solved for a given $\mathbf{K}_i$ and $\mathbf{K}_i$ is solved for using the $\mathbf{\Phi}_i$ derived in the prior step. While the bulk of the algorithm and its proof can be found in the appendices of Rothman et al. (2010), we discuss briefly how the MRCE algorithm works in the `graphicalVAR` package in R.

The second algorithm detailed by Rothman et al. (2010, p. 952) is used to identify the optimal estimates of $\hat{\mathbf{\Phi}}_i$ and $\hat{\mathbf{K}}_i$ in the `graphicalVAR` package and is completed in a multi-step procedure. For a set of fixed values of both tuning parameters, we may initialize $\hat{\mathbf{\Phi}}_i(0)$ and $\hat{\mathbf{K}}_i(\hat{\mathbf{\Phi}}_i^{(0)})$, then conduct the following steps:

1. Compute $\hat{\mathbf{\Phi}}_i^{(m+1)} \leftarrow \hat{\mathbf{\Phi}}_i(\hat{\mathbf{K}}_i^{m})$ using the first algorithm proposed by Rothman et al. (2010, p. 6)
2. Compute $\hat{\mathbf{K}}_i^{(m+1)} \leftarrow \hat{\mathbf{\Phi}}_i^{(m+1)}$ with the GLASSO algorithm (Friedman et al., 2008)
3. If $\sum_{j,k} \left| \hat{\phi}_{ijk}^{(m+1)} - \hat{\phi}_{ijk}^{(m)} \right| < \epsilon \sum_{j,k} |\hat{\phi}_{ijk}^{\text{RIDGE}}|$ then stop, otherwise return to step 1

Model selection of graphical VAR models in the `graphicalVAR` package is determined by selecting the combination of tuning parameters which minimize the extended Bayesian information criterion (EBIC; Chen & Chen, 2008) for a given level of a set hyperparameter, γ:

$$\text{EBIC} = -2L + E\log(N) + 4\gamma E\log(P) \tag{6}$$

where L is the log-likelihood, N is the sample size, E is the number of non-zero edges in both $\hat{\mathbf{\Phi}}_i$ and the upper triangle of $\hat{\mathbf{K}}_i$, P is the number of nodes in the network, and γ is a hyperparameter set by the user to further penalize model selection toward sparser models (Epskamp & Fried, 2018; Foygel & Drton, 2010). The decision to use EBIC in lieu of BIC is because it allows researchers to specify the degree of sparsity they desire in the obtained network. This hyperparameter value may also be set to 0.00 if one wishes to use BIC as a selection criterion.

3. Subgrouping graphical VAR models

Upon fitting N GVAR models, one may be interested in identifying a group-level structure, or multiple subgroups characterized by similar structures. One way of doing so is to apply a community detection algorithm to a measure of choice that can help quantify the similarities between any two individuals. Our proposed approach, the scGVAR, is summarized in Figure 2. In scGVAR, we utilize WalkTrap as the community detection approach which—broadly—identifies communities which share more connections with one another than would be expected by chance (see Pons & Latapy, 2005, for a complete review). The measure for quantifying similarities across individuals is an $N \times N$ matrix called the adjacency matrix, $\mathbf{A}$, with element $a_{i,j}$ $(i \neq j)$ indicating the number of network coefficients (i.e., PDCs and the PCCs) shared by the i^{th} and j^{th} subjects. In the case of the PDCs, the i^{th} and j^{th} subjects are common if they share—or mutually do not share—a structural edge in the same direction between two nodes as well as the same sign (i.e., ±). In the case of the PCCs," common" is defined solely by common sign between the same two nodes. Estimation of the PDCs and PCCs is still performed at the individual level using LASSO as previously described. In addition, sparsity is further introduced into $\mathbf{A}$ by applying a threshold to" zero-out" elements of $\mathbf{A}$ through optimization of the conductance—the ratio of within-group connections to connections to external-groups—of the resultant network structures (Sinclair & Jerrum, 1989). A complete description of the technical components of the scGVAR may be found in Appendix A.

The results of the scGVAR algorithm include a sample-level cGVAR, K cGVAR models for each identified subgroup, and N GVAR models for each subject. The group- and subgroup-level networks of the scGVAR model are fully parameterized network structures which contain estimated coefficients of the average strength of edges at the group- and subgroup-level. This allows researchers to further distinguish subgroups by not only direction of lagged effects but also their magnitude and polarity at the subgroup-level.

4. The alternating least squares (ALS) VAR

Initially proposed by Bulteel et al. (2016), the ALS VAR is an idio-thetic approach for clustering individuals into groups and estimating a VAR model upon

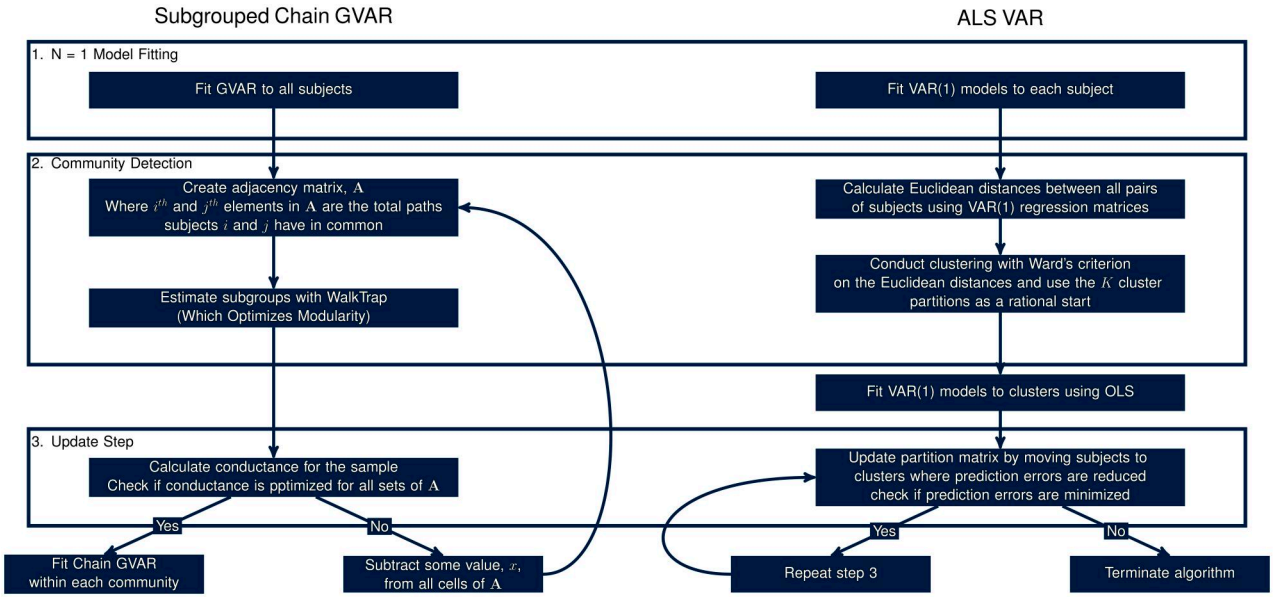

Figure 2. Procedure for scGVAR (left) and the ALS VAR (right). BOxes indicate common steps while dark boxes highlight algorithm-specific steps taken by the scgVAR and ALS VAR, respectively.

the clusters in a similar manner to the scGVAR. The model formula for the ALS VAR is identical to that of a standard VAR in Equation (1). The ALS VAR will estimate a series of individual VAR models and—upon clustering—fit a VAR model to the chained time-series of all subjects within a cluster.

4.1. Estimation of the ALS VAR

For a set of user-specified K clusters, a partition matrix and K regression coefficient matrices are estimated by minimizing the objective following function:

$$L_K = \sum_{i=1}^{I} \sum_{t=2}^{T_i} (\eta_{it} - \hat{\eta}_{it})^2 \tag{7}$$

where L_K is the sum of squared prediction errors, η_{it} are the $p \times 1$ vector of predicted scores for the i^{th} subject at time, t, and $\hat{\eta}_{it}$ are the $p \times 1$ vector of predicted scores for the i^{th} subject at time, t. Thus, by minimizing L_K, the ALS VAR is seeking a partition of K clusters where the prediction error is minimized across all subjects.

Bulteel et al. (2016) proposed a four step procedure for estimating ALS VAR models which are presented in Figure 2 and is similar in many ways to the scGVAR algorithm. For instance, both approaches initialize their search for subgroups by fitting N person-specific models and constructing adjacency matrices from the N VAR coefficient matrices. Notably, the ALS VAR looks solely at the $\mathbf{\Phi}_i$ matrices while the scGVAR assessed both the $\mathbf{\Phi}_i$ and the matrix of contemporaneous relations (i.e., PCCs; Equation (4)). However, the ALS VAR constructs the adjacency matrix using the Euclidean distances between each subject's $\mathbf{\Phi}_i$ coefficients matrices whereas the scGVAR uses a count of common paths. Following this, the ALS VAR uses Ward's criterion (Ward, 1963) to identify K clusters of individuals who are closest in their Euclidean distances where the value of K is designated by the researcher as the total number of clusters they would like to enumerate up to. In contrast, the scGVAR utilizes the WalkTrap algorithm to determine the number of clusters purely based on the data. This initial cluster solution is then used as the rational start for the ALS VAR algorithm. In the second step, a VAR(1) model is fitted to the chained time-series of all subjects within a cluster. The third step updates the adjacency matrix by moving subjects to clusters where their prediction errors are minimized and VAR(1) models are fitted again to the updated clusters. The third step is repeated until L_K is minimized (see Equation (7)).

4.2. Model selection of ALS VARs

Currently, the ALS VAR utilizes the Convex Hull (CHULL; Ceulemans & Kiers, 2006) procedure for model selection. This method is a numerical means of identifying the "elbow" in a scree plot. In this instance, assessing the relationship between the number of clusters and misfit in terms of the sum of squared prediction errors (Bulteel et al., 2016). This procedure is completed by calculating for each number of subgroups from $K = 1$ to K:

$$st_k = \frac{L_{K-1} - L_K}{L_K - L_{K+1}} \tag{8}$$

and selecting the clustering partition where st_k is maximized which is indicative that an additional cluster would not significantly increase model fit—or reduce the sum of squared prediction errors. This can be thought of as identifying the point at which the prediction errors changed significantly in transitioning from L_{K-1} to L_K relative to the change in prediction error going from L_K to L_{K+1}.

In simulation studies, the ALS VAR has performed quite well in distinguishing clusters based on unique structural differences and has been found to be able to distinguish between groups with identical structures with only magnitudinal differences separating them (Bulteel et al., 2016; Takano et al., 2021). However, these same simulation studies report some degree of misclassification when groups lie "between" one another in the Euclidean space due to the nature of the ALS algorithm described above prioritizing the ratio of improvement of fit to the number of enumerated clusters rather than solely minimizing discrepancy of fit.

5. Simulation study

A Monte Carlo (MC) simulation was conducted to evaluate several key questions regarding the performance of these algorithms: (1) How well would the scGVAR and the ALS VAR perform as the distance between subgroups—defined as the number of common cross-regression parameters—increases? (2) What would happen as the number of subgroups increases? (3) Do the scGVAR and ALS VAR perform better as the number of time-points increases? And (4) How well do the scGVAR and ALS VAR perform when dynamics occur significantly faster than the sampled rate (i.e., in the presence of contemporaneous effects)?

5.1. Simulation conditions

Several design factors were manipulated to address these questions including:

1. Distances between subgroups: $D = 3$ or $D = 9$
2. Number of subgroups: $G = 2$ or $G = 4$
3. Number of time-points per subject: $T = 100$ or $T = 500$
4. Lag-only and contemporaneous-only networks

Ten-variate stationary VAR(1) models were used to simulate data for $N = 52$ subjects across 500 MC runs. In the two-group condition, each subgroup consisted of 26-subjects each whereas the four-group condition resulted in 13-subjects per subgroup. The transition matrices of the 10-variate VAR(1) models were specified such that all subjects had auto-regressive elements fixed to 0.70 regardless of group membership. Following this, 10 cross-regressions with a weight of 0.20 were added to each subgroup model. Subgroups would possess either three ornine unique cross-regressions depending on the distance condition, D. Following the generation of each of the G subgroup models, subjects were simulated according the subgroups they were assigned to.

6. Simulation rationale

Distances, D, between subgroups were defined as the number of common cross-regression parameters shared between the subgroups. The decision to use three- and nine-unique cross-regression parameters was motivated by small-scale simulations that revealed that three- and nine-cross-regressive parameters coincided with significant differences between the multivariate distributions of the simulated VAR(1) models. These distances were derived using the Bhattcharaya distance (BD; Kailath, 1967). Formally, the BD is an extension of the Mahalanobis distance which quantifies the distance between two multivariate distributions with larger values corresponding to greater distances between the distributions.

The number of subgroups, G, was set to two and four subgroups. While two subgroups is the minimum for testing the results of the algorithms, four subgroups was considered because it has been used in similar simulation studies for evaluating cluster recovery (e.g., Gates et al., 2017) and reflects or exceeds the number of subgroups found in empirical work (e.g., Lane et al., 2019; Wright et al., 2019).

The number of time-points were meant to closely mirror our empirical illustration ($T = 100$). Historically, the minimum number of recommended time-points in time-series analysis has been at least $T = 50$ (see Hecht & Zitzmann, 2021, for further discussion); however, given the large dimensionality of the VAR(1) models in our simulations and empirical analyses, we chose a larger minimum value. Furthermore, we chose $T = 500$ to provide an asymptotic assessment of the performance of the algorithms.

Finally, we simulated a condition where all dynamics occurred contemporaneously (i.e., faster than the VAR(1) sampling rate). These dynamics were represented as covariances in the residual covariance matrix of the simulated VAR(1) models to highlight the performance of the scGVAR and its ability to capture" faster" dynamics. While the ALS VAR does not explicitly estimate contemporaneous effects, we tested this condition primarily to highlight the ability of the scGVAR to recover and cluster based on this information. When comparing the contemporaneous networks, the model implied residual covariance matrices were constructed from the results of the ALS VAR output and compared to the true residual covariance matrices (see subsection on Performance Measures).

Ultimately, the goal of the simulations were to highlight the strengths of both the approaches in different circumstances. As mentioned in the introduction, both the ALS VAR and the scGVAR have unique attributes that may grant them the edge over the other. For instance, the ALS VAR should be able to identify subgroups and their corresponding dynamic network structures in a lag(1) context with relatively little bias (Bulteel et al., 2016; Takano et al., 2021). Furthermore, the ALS VAR should perform well when subgroups were highly differentiated and when the number of clusters was relatively small with increasing performance over time. In contrast, the scGVAR should perform well when a large degree of differences among groups manifests as contemporaneous effects; which the ALS VAR does not explicitly model. Additionally, the clustering algorithm implemented by the scGVAR should allow it to separate groups that differ by relatively smaller number of edges relative to the ALS VAR.

In line with our expectations above, the simulation parameters highlighted the strengths of both approaches. Specifically, the manipulation of distances highlighted the strengths of the ALS VAR when the separation was relatively high (i.e., $D = 9$). In contrast, smaller degrees of separation should see lower performance for the ALS VAR but potentially higher performance for the scGVAR. Similarly, Bulteel et al. (2016) noted poor performance of the ALS VAR

when communities were relatively close in their dynamics as Time increases. Thus, the ALS VAR should perform better when $D = 9$ and when $G = 2$. In contrast, the scGVAR should perform well at G increases from 2 to 4.

6.1. Performance measures

The following section details the performance measures used in the MC study to evaluate the performance of the two algorithms in terms of quality of cluster assignment as well as in accuracy of point estimates at the subgroup level. We begin by discussing metrics on cluster assignment followed by a discussion of the metrics used for evaluating the recovery of point-estimates at the subgroup level.

6.1.1. Cluster recovery

In past work on assessing cluster accuracy, the Hubert-Arabie Adjusted Rand Index (referred to henceforth as the ARI; Hubert & Arabie, 1985) has been used to evaluate the quality of cluster assignments in simulation designs (Lane et al., 2019). The ARI, typically takes values ranging from 0 to 1 with 1 indicating a perfect match between true and estimated group membership. The ARI can be calculated as:

$$ARI_{HA} = \frac{\binom{N}{2}(a+d) - [(a+b)(a+c)+(c+d)(b+d)]}{\binom{N}{2}^2 - [(a+b)(a+c)+(c+d)(b+d)]} \tag{9}$$

where N is the number of subjects, a represents the number of hits, b indicates the number of false negative classifications (i.e., pairs classified as separate when they actually belonged together), c is the number of false positive classifications, and d is the number of true negative classifications. In evaluation, excellent recovery is considered $ARI_{HA} > 0.90$, good recovery $ARI_{HA} > 0.80$, moderate recovery is $ARI_{HA} > 0.65$, and values below $ARI_{HA} = 0.65$ are considered poor (Lane et al., 2019; Steinley, 2004).

6.1.2. Parameter recovery at subgroup level

Several metrics were used to assess the recovery of parameters at the subgroup level, including: bias, root mean square error (RMSE), power, and specificity and are described below:

Bias is defined as:

$$\text{Bias}(\theta) = \frac{1}{H}\sum_{h=1}^{H}\left(\hat{\theta}_h - \theta\right) \tag{10}$$

where θ and $\hat{\theta}$ represent the parameter of interest and its estimate and H is the total number of MC runs. Bias quantifies the direction (based on sign) and the magnitude of the deviation from the true parameter estimate relative to the actual parameter estimate. For this and all subsequent simulation metrics, these parameters were calculated for each parameter and averaged across all MC runs.

Power is defined as the proportion of MC runs where the parameter of interest was identified as non-zero when it truly is non-zero. Typically, power can be calculated by taking a proportion of the number of MC runs where a truly non-zero parameter is identified as non-zero based on the 95% CI of the estimated parameters. For the ALS VAR, these calculations could be implemented for the lagged simulations; however, in the contemporaneous conditions, this was not computationally feasible. Likewise, this method for calculating statistical power was not viable for the scGVAR due to the regularized estimates not returning standard errors. Prior interpretations of regularized networks indicate that regularized estimates that are non-zero are effectively "selected" for (Epskamp & Fried, 2018). Thus, we considered non-zero regularized edges as true non-zeros when calculating statistical power after rounding to two-decimal places similar to simulation results conducted by Park et al. (2021).

Specificity measures were also calculated for the simulations; however, specificity calculations differed from conventional settings for the scGVAR as standard errors are not returned. Typically, specificity rates are computed in MC studies as the proportion of MC runs in which the 95% CI of a parameter whose true value is zero contained zero. As the regularization procedure is considered, in and of itself, a form of variable selection, we calculated specificity based on whether parameters were set to 0 or not after rounding to the second decimal place for all results. Specificity calculations were carried out in the traditional sense for the ALS VAR when standard errors were accessible using the typical approaches defined as:

$$Spec(\theta) = \frac{TN_\theta}{TN_\theta + FP_\theta} \tag{11}$$

where TN_θ indicates the number of true negatives based on whether the true parameter is captured by the 95% CI of the estimated parameter and FP indicates the number of false positives.

The RMSE quantifies the degree of deviation an estimate $\hat{\theta}$ has from its true value θ and is calculated as:

$$RMSE(\theta) = \sqrt{\frac{1}{H}\sum_{h=1}^{H}\left(\hat{\theta}_h - \theta\right)^2} \tag{12}$$

6.2. Simulation results

The simulations revealed several areas of promise for both algorithms. We begin by discussing the general performance of the ALS VAR and the scGVAR. Following this, we explore how the two algorithms performed across the simulated conditions we specified previously. The full table of results can be found in Table 1.

6.2.1. Distances between subgroups

In line with expectations, the ALS VAR tended to perform better with greater distances between communities when $G = 2$. Notably, the ALS VAR exhibited increased statistical power, lower RMSEs, and generally high performance on ARIs when distances jumped from 3 to 9 edges separating the communities see Table 1 and Figure 3.

Similarly, the scGVAR exhibited a gain in performance as D increased from 3 to 9 in terms of ARIs, Bias, Power, and Specificity with RMSEs remaining stable across all conditions see Figure 3. The scGVAR biases tended to decrease with greater distances; however, the biases did not tend to decrease with greater sample-sizes. These results have been found in other implementations of LASSO regularization where bias-related estimates tended to plateau at higher sample-sizes (Williams & Rast, 2020).

6.2.2. Number of communities

When the number of groups was equal to 2, the ALS VAR and the scGVAR both performed rather well with the ALS VAR exhibiting noticeably lower power and specificity relative to the scGVAR—see Figure 4—but lower biases and a higher cluster accuracy.

The scGVAR performed notably better than the ALS VAR when the number of communities increased from 2 to 4 on all performance measures. With respect to ARI performance, the scGVAR scored ARI values ranging from $[0.46; 0.99]$ while the ALS VAR only ranged from $[0.40; 0.44]$ when $G = 4$ as seen in Figure 5. The high instance of incorrect cluster associations would result in incorrect estimates and thus is confounded with the poor performance on all other metrics. This likely is associated with how the ALS VAR and the scGVAR identify communities in their search algorithms. As noted by Bulteel et al. (2016), the ALS VAR struggles to find the correct number of communities when there is a large degree of overlap or when a community lies "between" two others in the Euclidean space. In our simulated example, this is highlighted when we have multiple groups with many common edges. A supplemental simulation was conducted in the "four-group condition" where the number of subjects within a subgroup was held constant (i.e., $n_G = 26$) relative to the two-group conditions but total N was allowed to vary to ($N = 104$) and are included in Table 1. The results of these simulations showed slight variation in the subgrouping accuracy for the scGVAR and the ALS VAR. Namely, the scGVAR improved slightly whereas the ALS VAR showed a slight decrease in performance.

An important note is that the ALS VAR often" merged" communities together when they shared a large degree of overlap. This explains why the ALS VAR still returned fairly acceptable biases across both configurations of G. That is, the algorithm would enumerate a two-group solution where two of the true four groups would be merged together.

6.2.3. Length of time

In general, both the ALS VAR and the scGVAR exhibited improved performance as the number of measurement occasions increased from $T = 100$ to $T = 500$. However, the biases did not tend to decrease

Table 1. Table of simulation results comparing the scGVAR to the ALS VAR for the two-subgroup and four-subgroup conditions across different sample sizes, T and distances D.

	ARI		Bias		Power		Specificity		RMSE	
	scGVAR	ALS VAR	scGVAR	ALS VAR	scGVAR	ALS VAR	scGVAR	ALS VAR	scGVAR	ALS VAR
$T_{100};D_3;G_2$	0.962	1.000	−0.011	−0.001	0.999	0.540	0.936	0.820	0.154	0.178
$T_{500};D_3;G_2$	0.997	0.999	−0.013	−0.001	1.000	0.563	0.961	0.785	0.154	0.177
$T_{100};D_9;G_2$	1.000	1.000	−0.011	−0.001	1.000	0.550	0.941	0.816	0.153	0.176
$T_{500};D_9;G_2$	1.000	1.000	−0.013	−0.002	1.000	0.573	0.952	0.777	0.152	0.175
$T_{100};D_3;G_4$	0.466	0.440	−0.016	−0.006	0.996	0.611	0.771	0.684	0.152	0.171
$T_{100};D_3;G_4$	0.537	0.402	−0.015	−0.006	0.997	0.626	0.809	0.653	0.153	0.171
$T_{500};D_3;G_4$	0.892	0.402	−0.014	−0.007	0.999	0.678	0.879	0.563	0.155	0.170
$T_{100};D_9;G_4$	0.999	0.436	−0.011	−0.014	1.000	0.760	0.907	0.425	0.152	0.156
$T_{500};D_9;G_4$	0.997	0.440	−0.013	−0.015	1.000	0.835	0.919	0.291	0.153	0.156
$T_{500};D_3;G_2$	0.820	0.000	0.007	0.000	1.000	1.000	0.574	0.367	0.180	0.070
$T_{500};D_9;G_2$	0.746	−0.001	0.007	0.000	0.998	1.000	0.565	0.302	0.185	0.120

Notes: Isolated rows indicate conditions where only contemporaneous effects we generated. Consideration was made to keep all sample sizes the same. The gray row denotes a condition where N was manipulated but all other parameters were held constant to the condition above it. Differences are relatively minor-relative to the condition above-with slight improvements to the ARI for the scGVAR and a slight decrease for the ALS VAR.

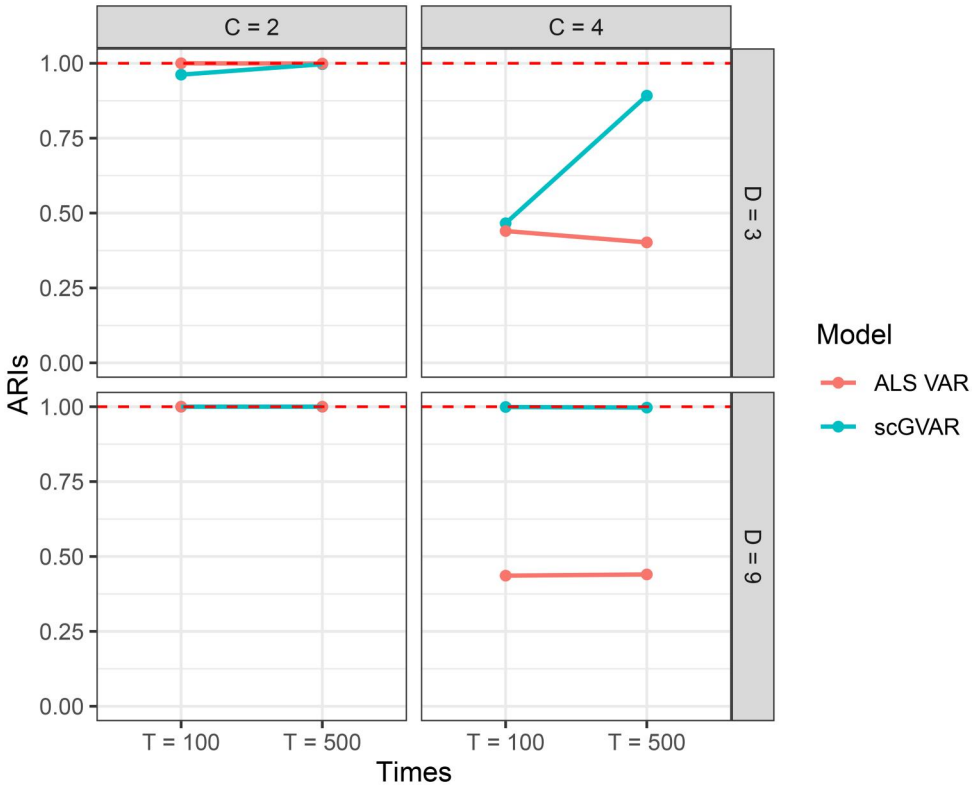

Figure 3. ARI values across simulation conditions. *G* indicates the number of clusters, *D* indicates the distance between subgroups, and *T* took on values of 100 or 500. Higher ARI values indicate better accuracy in subgrouping performance. Dashed line indicates the ideal value of the ARI at 1.00. The scGVAR tended to perform better with larger separation between subgroups and larger *T*.

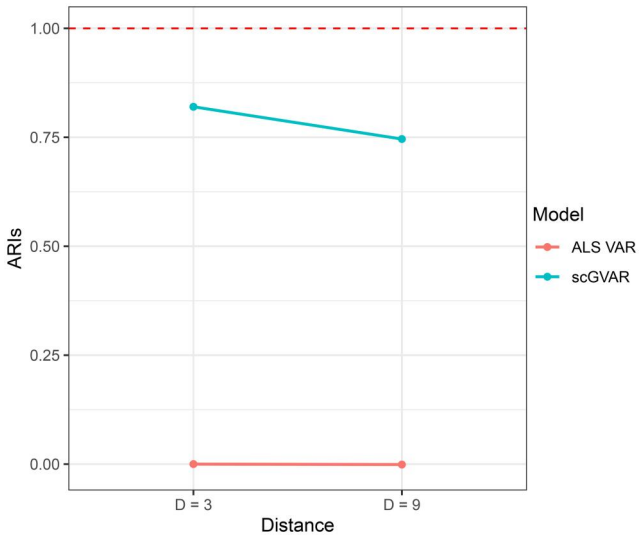

Figure 4. ARI values for the contemporaneous-only simulations. *D* indicates distance between subgroups. Higher ARI values indicate better accuracy in subgrouping performance. Dashed line indicates the ideal value of the ARI at 1.00. The scGVAR tended to perform better than the ALS VAR in terms of subgrouping accuracy for contemporaneous-only conditions.

with greater sample-sizes for the scGVAR. These results have been found in other implementations of LASSO where bias-related estimates tended to plateau at higher sample-sizes (Williams & Rast, 2020). This was confirmed when the biases decreased with increasing sample-size in a follow-up small-scale simulation with the regularization turned off. Similar to the ALS VAR, the scGVAR RMSEs tended to remain relatively consistent across simulated conditions irrespective of sample-size or Bhattcharaya distance.

6.2.4. Contemporaneous effects

When considering a network with only autoregressive effects and contemporaneous edges, the ALS VAR performed notably worse in its classification accuracy; though, this is unsurprising since the residual covariances are not utilized when clustering individuals. The ARIs for the ALS VAR across both Bhattcharaya distances were effectively 0.00. In contrast to the ALS VAR, the scGVAR explicitly models the contemporaneous network of partial correlations and leverages this information during the clustering procedure. As expected, the ALS VAR performed well in recovering the underlying communities with ARIs ranging from 0.75 to 0.82 and can be seen in Figure 6.

Interestingly, the biases in the ALS VAR were relatively low for the contemporaneous networks with values of −0.0002 in both conditions. Furthermore, the RMSEs for the recovered contemporaneous networks

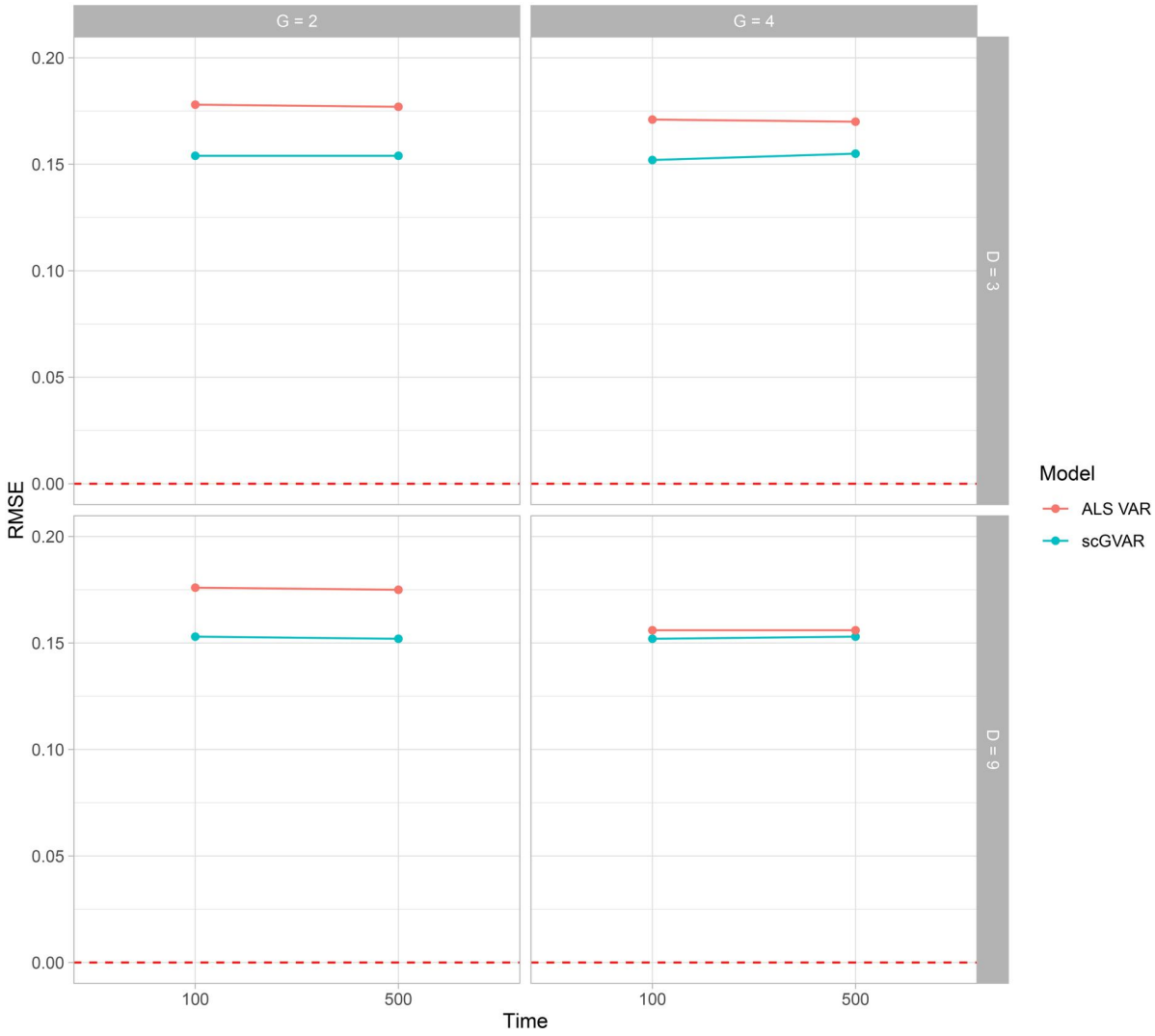

Figure 5. RMSE values across all lagged simulation conditions. G indicates the number of subgroups, D is the distance between subgroups, and time is allowed to vary between 100 and 500 time-points. The red, dashed line indicates an ideal RMSE value of 0.00. RMSEs tended to remain constant for the scGVAR across conditions. In contrast, RMSEs improved for the ALS VAR as the distance between groups grew larger.

tended to increase as Bhattcharaya distances increased from 0.07 to 0.12.

The scGVAR showed better bias profiles when compared to the lagged networks with biases ranging from 0.006 to 0.007 when moving from small to large Bhattcharaya distances; however, unlike the ALS VAR, the scGVAR RMSEs remained relatively consistent regardless of the Bhattcharaya distance, $[0.180, 0.184]$. The ALS VAR had high power with 100% across both distance configurations; however, the specificity was rather low at 0.37 to 0.30 as distance increased between the clusters. Similarly, the scGVAR had high power, 99.9% to 99.8% but had higher specificity than the ALS VAR with values of 0.57 across both conditions.

6.3. Simulation review

Overall, the ALS VAR produced relatively unbiased estimates of both the lagged- and contemporaneous-networks. Furthermore, the algorithm tended to exhibit high power and acceptable levels of specificity in the lagged conditions when only two clusters were present.

In the $G = 2$ conditions, the ALS VAR had excellent recovery of the underlying subgroups; however, it had difficulty in enumerating the proper number of subgroups when $G = 4$. These results are similar to those found by Bulteel et al. (2016) who found that the algorithm had some difficulties in separating groups that shared a large degree of similarity. Across our D configurations, all networks still shared non-

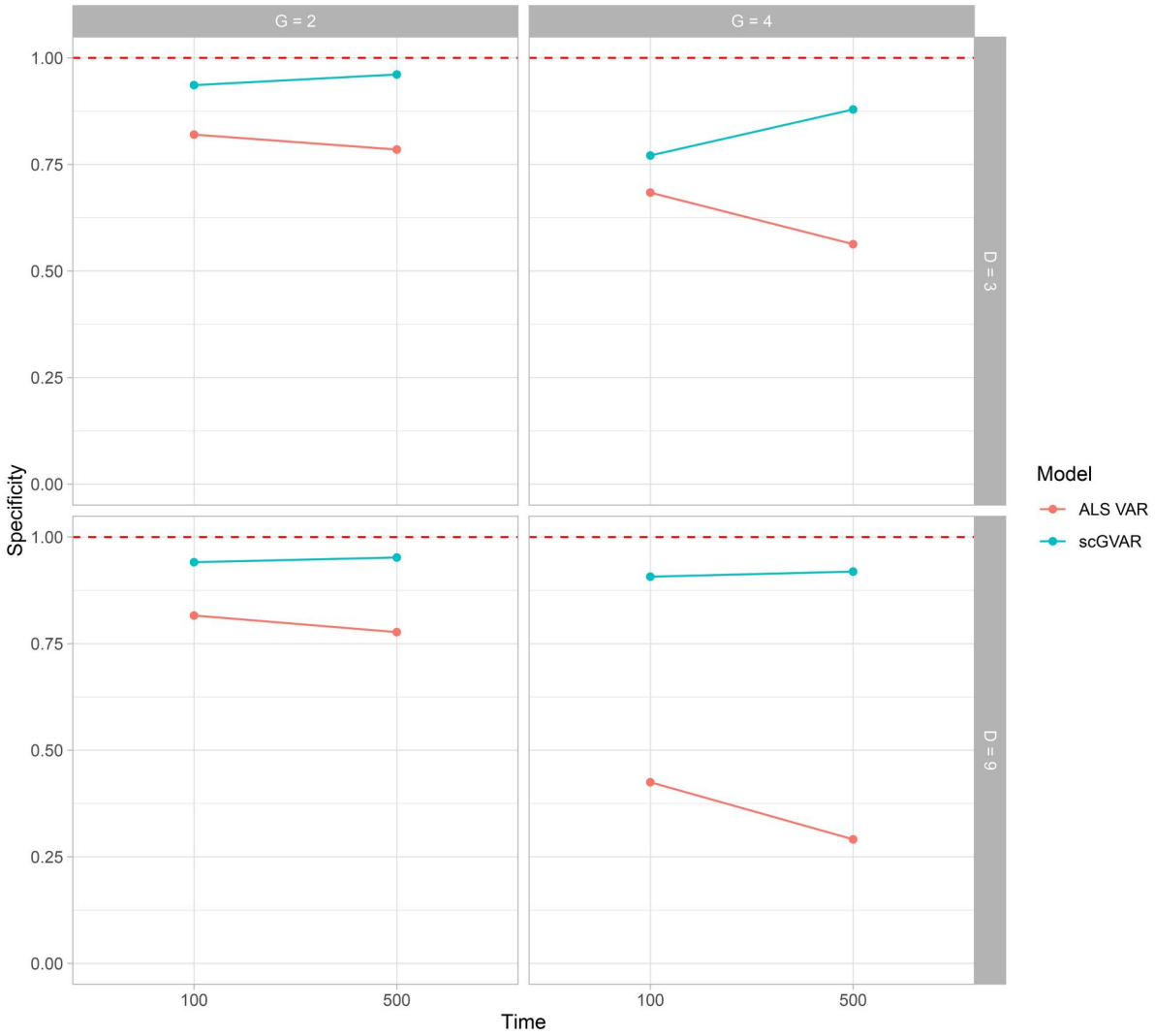

Figure 6. Specificity values across all lagged simulation conditions. G indicates the number of subgroups, D is the distance between subgroups, and time is allowed to vary between 100 and 500 time-points. The red, dashed line indicates an ideal specificity value of 1.00. scGVAR tended to increase in specificity with greater time-points. In contrast, the ALS VAR seemed to exhibit decreased performance with larger *T* in the four-group conditions. Notably, this may be due to it merging subjects from different groups and incorrectly detecting effects that were not present for those subgroups.

zero autoregressions and seven- or one-additional cross-lagged edge. In both of these configurations, both communities shared more edges than they had uniquely to themselves. This was further compounded when there were four communities as the ALS VAR tended to enumerate only two communities a majority of the time. This likely is related to the CHULL procedure implemented by the ALS VAR which prioritizes relative improvements to prediction errors as a ratio to the number of clusters identified; thus, two merged clusters performed 'good enough' to be accepted.

Generally, the scGVAR consistently returned more biased estimates than the ALS VAR across all conditions. Our simulations revealed that the scGVAR seems to outperform the ALS VAR when groups differ on a minority of edges; particularly in relatively noisy conditions (i.e., $G = 4$; and larger distances).

Figure 1 provides a synthesis of our simulation results with respect to conditions where one algorithm may be preferred over the other. The ALS VAR may be more effective than the scGVAR when groups are expected to differ by their magnitude (Figure 1(A); Takano et al., 2021). This could be expected when differentiating groups with psychopathologies which have been shown to differ in the density of their networks rather than in specific connections. Additionally, the ALS VAR may be desirable when groups are expected to differ by a relatively large number of edges with clear separations between the clusters (Figure 1(B)). In contrast, when a large degree of common edges are expected with a small number

of differentiating points, the scGVAR may be preferred (Figure 1(C)). Additionally, the scGVAR seemed to be the clear favorite in terms of recovering a network of contemporaneous effects; particularly if groups are expected to differ based on their contemporaneous effects (Figure 1(D)). As shown in previous works, these results may be expected in a number of studies where the time-scale of the process under study is faster than the sampling rate (Park et al., 2021). Overall, the simulation results highlight unique strengths to both algorithms. Specific to the scGVAR, subgroup recovery was dependent on time, T when the number of subgroups was 4. This is due to how the subgrouping is conducted using the parameter estimates themselves; in instances of low statistical power, the dynamic networks will be more difficult to recover and may lead to erroneous assignments or even spurious clusters being generated.

7. Illustrative example

Idio-thetic methods have seen increased use in clinical samples to explain potential reasons for the great deal of heterogeneity observed in clinical samples (De Vos et al., 2017; Gates et al., 2014). These implementations may yield insight regarding the degree to which individuals differ in their dynamic network structures and could potentially be used to identify common dynamics that meaningfully characterize groups of individuals. For example, past work by Gates et al. (2014) found unique temporal dynamics in the neural relations of ADHD patients which corresponded to different diagnostic sub-types. Thus, we hope to apply both the scGVAR and the ALS VAR to a clinical sample of individuals with Major Depressive Disorder (MDD) to potentially identify unique subgroups of individuals who may possess unique or otherwise informative dynamic network structures.

The following data were taken from work completed by De Vos et al. (2017) as part of the Mood and Movement in Daily Life (MOOVD) study and will be used to demonstrate the scGVAR and the ALS VAR. The data contains a sample of $N = 54$ individuals falling into either clinical Major Depressive Disorder ($n_{MDD} = 27$) or control groups ($n_{CON} = 27$) who were pair-matched by age, gender, smoking behavior, and BMI. However, due to convergence issues, seven participants were excluded from the analysis ($N = 48$; $n_{MDD} = 24$; $n_{CON} = 23$). Specifically, the models for these participants—even following the normalization procedures—failed to return estimates for the scGVAR during optimization. Participants were asked to respond to survey measures three times per day for 30-days resulting in a dataset with a maximum number of 90-responses per participant.

Past work by Wright et al. (2019) has indicated that many studies on emotion dynamics contain contemporaneous effects due to the speed at which emotion dynamics tend to occur relative to the sampling intervals utilized by researchers. Thus, the scGVAR is expected to identify subgroups based largely on contemporaneous effects whereas the ALS VAR may prioritize identifying subgroups which strongly differ from one another.

7.1. Measures

The daily diary measure consisted of 14-affective items (7 positive; 7 negative). The positive affect (PA) items consisted of items such as, *Feeling Talkative*, *Enthusiastic*, and *Energetic*. The negative affect (NA) items were comprised of similar item stems, *Feeling Tense*, *Anxious*, and *Depressed*. The items were scored on a seven-point Likert-type scale and had been previously used in another study (De Vos et al., 2017).

7.2. Data preprocessing

Data were preprocessed using identical procedures as those used in De Vos et al. (2017) using code provided by the authors. Missing data were approximately 8.2% and 6.8% in the MDD and control groups, respectively. The average length of time for subjects was approximately 83.2-observations. These missing data were imputed using the "Amelia" package in R rendering 90 total observations per subject. Following this, data were detrended by first fitting a nonparametric smoothing spline to each univariate time-series for each item and secondly by subtracting the spline curve from the original time-series. Finally, to address the Likert-type responses with skewed distributions, each variable was transformed using a normal quantile transformation to relax the assumption of normality (see Bogner et al., 2012). For the between-subjects networks, within-subjects means were calculated across each variable prior to these data treating procedures to obtain their within-person mean levels and fed directly into the analyses.

7.3. Analysis

The scGVAR was fitted to the entire sample in two ways. First, subjects were pre-partitioned based on their diagnostic categorization (i.e., MDD or Control).

Following this, the scGVAR and the ALS VAR were run without any *a priori* specification as to potential group membership based on the clinical diagnoses using the data to define the subgroups on their own. While the selection of the γ-hyperparameter is left to the researcher's discretion, we set $\gamma = 0.00$ in favor of model discovery of individual networks. Following the scGVAR search, the ALS VAR was fitted to the data with 2 subgroups extracted to match the expected number of potential groups (i.e., an MDD and a Control group). Both models and their results are discussed below.

7.4. Results

A between-subjects partial correlation network was fitted onto the time-averaged means of each subject within their respective diagnostic groups and aligned with established research looking at the affective dynamics of individuals with MDD (Figure 7; Schweren et al., 2018). Specifically, we found more densely connected networks among the MDD individuals relative to the controls; particularly with respect to the cross-affective (i.e., positive–negative) connections. This greater network density and strong cross-affective connectivity can be described as high emotional rigidity (Lydon-Staley et al., 2019) and has been related to the concept of emotional inertia in the modeling of within-subjects data (Kuppens et al., 2010) where individuals tend to become trapped by prior emotional states (e.g., strong lagged terms). We saw high rigidity in the MDD group from the confirmatory approach where subjects with MDD tended to have strong positive associations within their negative affect networks that also suppressed their positive affect networks. While the control group also had some strongly connected edges in their negative affect networks, these variables did not tend to effect their positive affect variables to the same degree.

Results from the scGVAR analysis did not align with clinical diagnoses. That is, participants were not evenly divided into MDD and control groups when using the scGVAR on the data. Instead, our data aligned with observations by De Vos et al. (2017) that participants were highly heterogeneous. Two clear clusters of participants did emerge from the analyses and is discussed in detail below (see Figure 8).

7.5. Group- and subgroup-level analysis for scGVAR

The group-level network produced by the scGVAR is provided in Figure 9 and shows strong connections within affective states. That is, positive and negative affect variables tended to be positively associated with one another. Additionally, we saw strong autoregressions among all of the negative affect variables and several positive affect variables indicating some degree of emotional inertia. Finally, we observed negative associations between the positive and negative affect variables with the strongest connection being a negative edge between depression and happiness.

Overall, the scGVAR delineated two major subgroups (Figure 10). All other subjects were classified as singletons or simple dyads, $n_{singles} = 17$. The presence of such a large degree of smaller, fractured groups may be a reflection of a significant degree of individual differences in the dataset as observed by previous studies (e.g., De Vos et al., 2017). To simplify discussion of the results, only the major subgroups will be discussed. The first recovered subgroup contained $n_{G1} = 21$ participants (44.6%) from the sample. The average network structure for this subgroup was relatively dense with both lagged and contemporaneous effects and participants tended to overlap to a large degree. Specifically, participants in the first recovered subgroup tended to share anywhere between 48 ($\sim 13\%$) and 76 ($\sim 20\%$) common edges in terms of both directionality—when applicable—and polarity. It should be noted that the vast majority of dynamics take place contemporaneously. This is unsurprising as prior studies have found that affective dynamics may take place at faster intervals than the three-measurements per day in the current investigation (Park et al., 2021; Wright et al., 2019).

We saw that the first subgroup was fairly similar to the overall group-level model. It was characterized by several autoregressive components; namely on "Guilt," "Tense," and" Anxiety"; among other negative affect variables. This indicates that—on average—participants in this subgroup tended to experience these emotions across time-intervals. Thus, feelings of guilt, tension, and anxiety tended to spill over and predict higher levels of these emotions in subsequent occasions.

With respect to the contemporaneous effects, we see a high degree of connectivity between nodes indicative of positive and negative affect. This suggests that negative and positive affect items influence one another freely within the time-intervals but not between them. For example, Feelings of depression tend to contemporaneously suppress feelings of happiness and satisfaction in one interval of time. However, these feelings of depression do not necessarily predict those emotions in the next interval. In addition, we

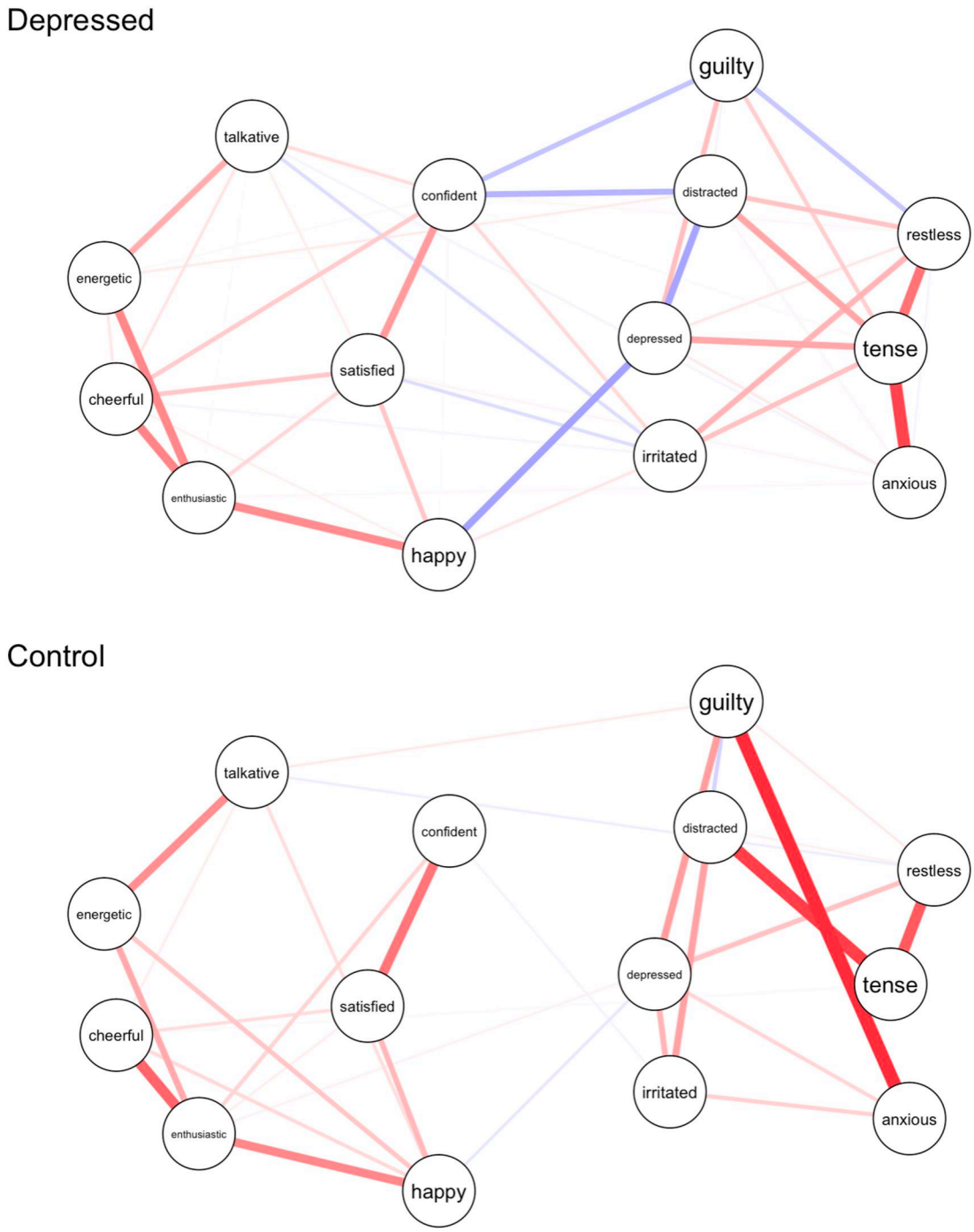

Figure 7. Between-subjects networks from the scGVAR function. Red edges indicate positive associations among variables and blue edges indicate negative associations.

tend to see the strongest connections within positive and negative affect items. For instance, the strongest connections appear between "Cheerful" and "Happiness"; "Energetic" and "Talktative"; "Tense" and "Restless", etc.

The distinguishing difference between the first and second recovered subgroups was that the second recovered subgroup had noticeably weaker connections among the negative affect variables, no autoregressive terms for the positive affect variables, and a lower degree of connectivity between the positive and negative affect variables.

7.5.1. Subgroup-level analysis for ALS VAR

The ALS VAR converged on a two subgroup solution. Each subgroup was comprised of $n_{G1} = 17; n_{G2} = 33$ participants, respectively and exhibited drastically different affect dynamics. Similar to the findings of the

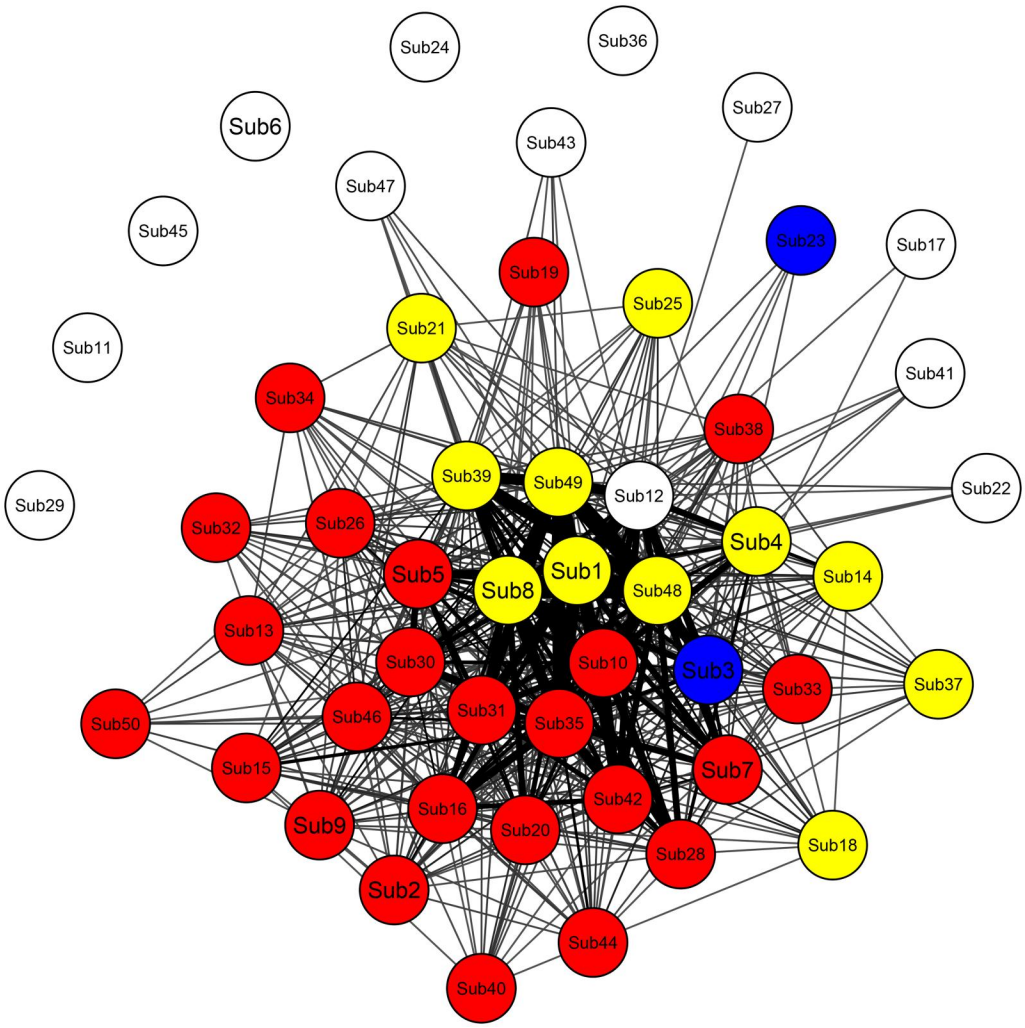

Figure 8. Subgroup membership from the scGVAR applied to the whole sample of controls and depressed subjects. Nodes indicate individual subjects where subjects of the same color belong to the same communities. Subjects whose nodes are white were not placed in any community. Communities did not align with clinical diagnoses.

scGVAR output, we find that the sample did not diverge based on the clinical diagnoses with members of both groups comprised of both MDD and control participants. Additionally, we note that—while the scGVAR output contained networks comprised of almost entirely contemporaneous edges, we note that the ALS VAR output only allows for lagged effects. As well, the scGVAR output was notably less dense than the ALS VAR networks. We discuss the key differences and characteristics as follows.

The first subgroup was characterized by strong autoregressive features in both the positive and negative affect variables indicating a large degree of cross-occasion spillover. Additionally—and similar to the first scGVAR subgroup—the first subgroup had a strong degree of connectivity between the positive and negative affect variables. However, unlike in the scGVAR, the positive and negative affect variables did not seem to cluster together as clearly with many negative connections among variables of the same valence which is not typically expected (e.g., greater anxiety is associated with lower restlessness).

The second subgroup recovered by the ALS VAR was noticeably less densely connected both overall as well as between the positive and negative affect variables when compared to the first subgroup. This second subgroup was also characterized by a clearer degree of item-level clustering within the positive and negative variables (e.g., all negative affect variables were positively associated with one another).

Another interesting finding was that while the two subgroups of the ALS VAR did share some similarities (e.g., strong autoregressions on similar terms), other edges were polar opposites to one another. For instance, in the first subgroup, greater irritation was associated with greater happiness. In contrast, greater irritation was associated with lower happiness. This was the case with many edges across the two

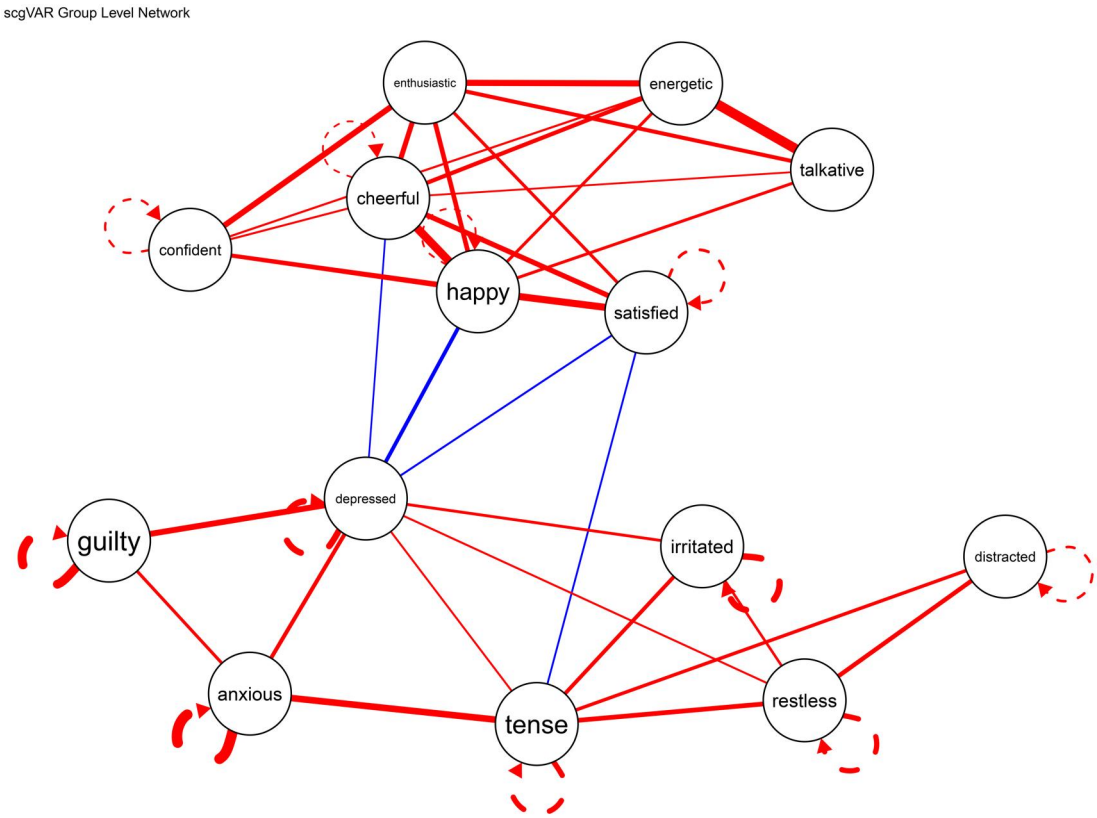

Figure 9. Group-level network of the scGVAR. Red edges indicate positive connections and blue indicate negative connections. Solid lines are indicative of contemporaneous effects and dashed lines are indicative of lagged effects.

subgroups. However, several shared paths emerged as well (e.g., greater confidence associated with greater happiness) which was also a common edge in the scGVAR networks as well.

7.5.2. scGVAR and ALS VAR comparisons

The results of the empirical illustration present two different pictures with some overlap as well. Beginning with the commonalities, we found that the estimated subgroups did not fall along clinical diagnostic lines. Instead, the groups fell along clearly distinguishable and structurally unique classifications. Thus, in line with results presented by De Vos et al. (2017), we see that the sample is fairly heterogeneous in both the ALS VAR and the scGVAR.

The scGVAR subgroups consisted primarily of contemporaneous rather than lagged effects. This finding was in line with prior studies on affect which suggest that these dynamics may take place faster than our sampling window of three-measurements per day (Park et al., 2021; Wright et al., 2019). In a clear demonstration of the estimation differences, the ALS VAR does not contain any contemporaneous information and, instead, all effects existed as lagged effects. This has profound implications for how one may interpret these networks as well as how the time-scale of effects are perceived.

Furthermore, we observed that the scGVAR networks were sparser than the ALS VAR networks. This is also in line with our expectations as the scGVAR is a regularized procedure which shrinks weaker edges to zero whereas the ALS VAR has no means of shrinking edges resulting in denser networks. Overall, this resulted in clearer network of dynamics for interpretation when compared to the ALS VAR's recovered networks which were much more dense. In some instances, edges in the ALS VAR were also somewhat counter-intuitive. For example, Subgroup 1 of the ALS VAR (see bottom left panel in Figure 10) has several positive edges spanning between positive and negative affect variables. Interpretationally, this could be interpreted as "Depression" has a positive effect on "Confidence" in the subsequent time interval. Indeed, many edges between the two subgroups recovered by the ALS VAR are opposite to edges found in the other. That said, the ALS VAR accomplishes the goal of maximizing the differences between the subgroup-level VAR models. In contrast, the scGVAR models share a significant portion of edges with each other and differ by only a few edges. The differences between the scGVAR and ALS VAR are not entirely

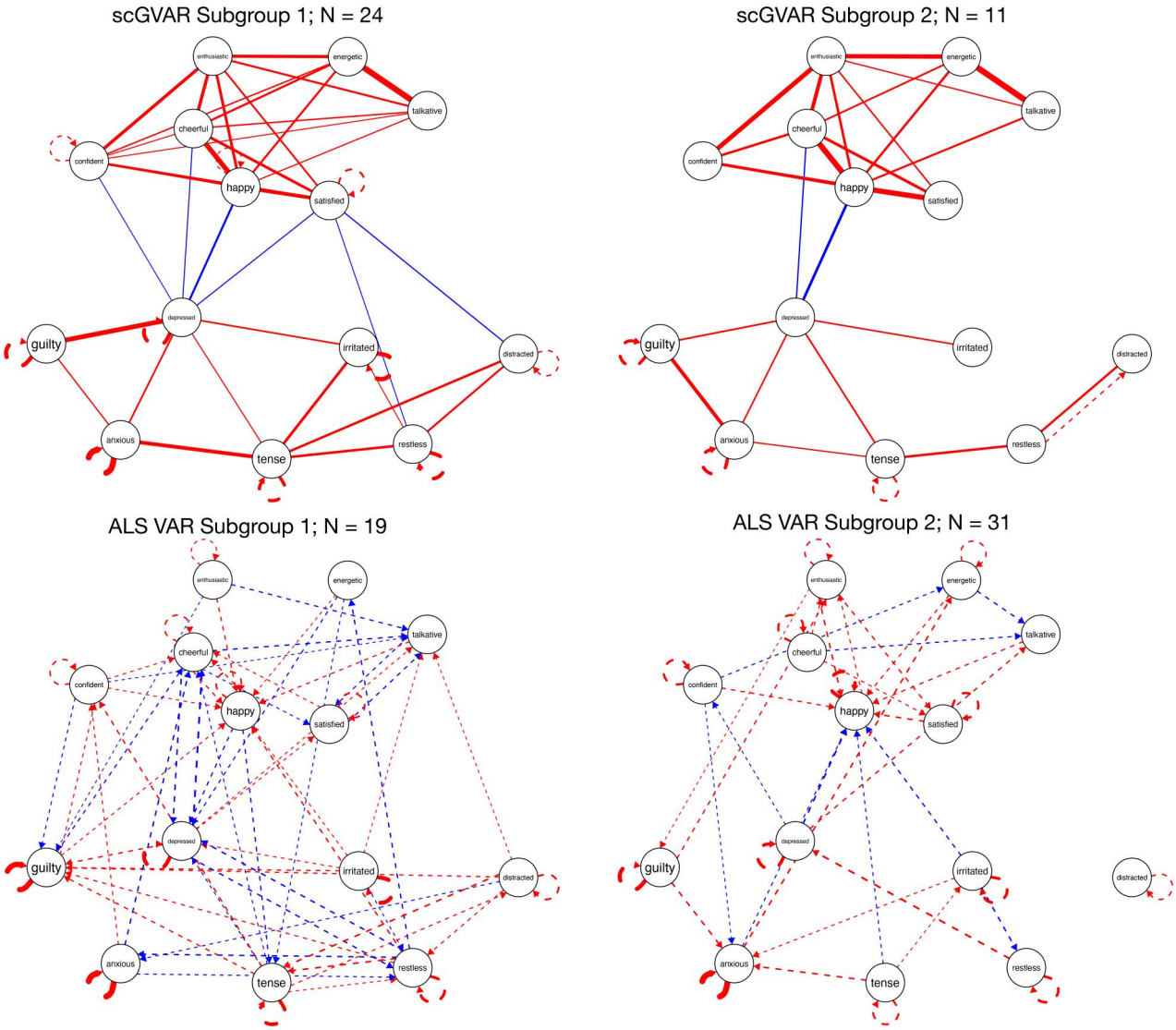

Figure 10. Recovered subgroups by method with scGVAR on top and the ALS VAR on the bottom row. Subgroups comprised of N >2 participants shown. Red edges indicate positive associations and blue edges are negative associations. Dashed lines are lagged effects whereas bold lines are contemporaneous effects.

surprising as they prioritize different information. The ALS VAR—which seeks to minimize the sums of squares predictions errors, L_K, has been shown to perform quite well even when groups are separated by as little as 1-edge.

8. Discussion

The past decade has seen a boom in the use and development of idio-thetic methods (Bulteel et al., 2016; Z. F. Fisher et al., 2022; Gates et al., 2017). These methods are invaluable in the study of human behavior as we strive to understand individuals at treatment or clinical levels as well as garner a stronger understanding of processes in the general populace. We introduced the scGVAR as a novel idio-thetic method for estimating chain GVAR models to the time-series data of multiple subjects. The scGVAR allows researchers to identify homogeneous subgroups in a data-driven manner from subject-level data. The scGVAR model has several strengths over some alternative idio-thetic models. For example, the scGVAR readily comes with built-in regularization tools and is prepared to estimate non-directional contemporaneous effect networks (Epskamp, Waldorp, et al., 2018). Similarly, the scGVAR framework is distinct from other methods such as S-GIMME (Gates et al., 2017)—which is based on the structural VAR framework—in how contemporaneous effects are estimated. In S-GIMME, model building and constraints allow for identified directional contemporaneous relations whereas the contemporaneous effects in the scGVAR are non-directional partial correlations.

Prior work has discussed the ramifications of modeling directional or non-directional effects (Park et al., 2021). In these cases, it is left to researchers to decide which model is most appropriate for their research questions.

The results of our simulation present strengths and weaknesses of both clustering algorithms. Based on these results, we provide some insight as to when one algorithm may be preferred over another. To summarize, the ALS VAR returned consistently less biased results when compared to the scGVAR and outperformed the scGVAR in many respects when only one clusters were present. In contrast, the scGVAR tended to outperform the ALS VAR in the presence of distinguishing contemporaneous effects in addition to conditions when a large number of similar clusters existed. Overall, we believe that the two methods can be used in tandem with one another to leverage both of their strengths when exploring dynamic networks.

For example, in instances where both the scGVAR and the ALS VAR enumerate similar cluster solutions, the ALS VAR may be preferable due to its typically lower biases in the point-estimates. In contrast, when the cluster solutions differ, attention should be paid to *how* the solutions differ. As our simulation results showed, the ALS VAR may combine similar yet distinct groups if their dynamics are very similar to one another. Thus, the scGVAR may be preferable if the scGVAR returns solutions that generally represent subsets of a solution given by the ALS VAR as this would indicate that the ALS VAR may not have separated the groups as much as it could have due to the conservative nature of the CHULL procedure.

Furthermore, the results of our empirical illustration showed distinct structural differences in the estimated subgroups of individuals within the MOOVD sample. Previously, De Vos et al. (2017) conducted confirmatory analyses on participants in the MOOVD sample and found a great deal of within-group heterogeneity in both the MDD and control conditions. The data-driven approach taken by the current investigation found that the participants did not fall into clusters based on their clinical diagnoses in either the ALS VAR or the scGVAR. These findings could suggest that the two clinical groups of participants may share more—or fewer in the case of the scGVAR—similarities than expected by their diagnostic labels. Alternatively, the results may indicate the need for these algorithms to focus on network characteristics other than their structural differences. For example, Lydon-Staley et al. (2019) and Schweren et al. (2018) tested the degree to which individuals with depression differed based on their network density; which was a result replicated in our between-subjects networks. Thus, the methods used here—which are more concerned with differences in network structure—may be insufficient for identifying the true source of differences between MDD and controls.

8.1. Limitations

The study is limited by several factors. Our MC study investigated the strengths of the ALS VAR and the scGVAR when identifying solely lagged or contemporaneous networks; however, one or the other is unlikely to ever truly occur with real data. Future investigations could address the performance of these algorithms when both contemporaneous and cross-lagged effects are present and how those joint effects bias estimation. Relatedly, other idio-thetic approaches exist which explicitly estimate directed contemporaneous dynamics (e.g., S-GIMME; Gates et al., 2017). Thus, future investigations should attempt to compare these related methodologies against one another. Additionally, both the scGVAR and the ALS VAR are limited in their ability to subgroup individuals. Namely, subgroups *must* be assigned to a group or exist as a singleton case. This is primarily due to the "hard" nature of the clustering algorithms implemented by both methods discussed here. Some work has been done with methods that allow for probability-based assignment such as recent work by Liu et al. (2020), which extended the ALS VAR to incorporate random effects in a methodology dubbed the dynamic mixture model (DMM). Relatedly, fuzzy clustering approaches have been explored for use in clustering dynamic networks to data where subjects may belong to multiple groups (Park et al., 2023). Future work could explore the extent to which solutions from the ALS VAR and the scGVAR differ from solutions derived by fuzzy clustering approaches and the DMM approach given the additional random effects components, it would be interesting to see how sensitive DMM remains to groups which differ by magnitude but not structure.

Another limiting factor of our simulation was the number of subgroups tested. In our simulations, we enumerated two- and four-groups. The upper limit of four-subgroups was in-line with other simulation studies that assessed subgrouping performance with similar algorithms (Gates et al., 2017) and also matched or exceeded other empirical works (e.g., Lane et al., 2019; Wright et al., 2019); however, further work should investigate how these algorithms perform under conditions where many subgroups exist.

Similarly, we did not manipulate the number of variables in our dynamic networks and used a fixed network size of 10-variables. For this simulation, our primary focus was on the effect of the distance between the VAR models. In theory, using our definition of distance between VAR models should allow for a clearer idea of how well these algorithms perform regardless of the number of variables; however, there are some considerations. For example, the scGVAR relies on the number of common parameters each pair of subjects shares. Thus, smaller networks could result in a smaller total possible number of similarities and could result in more difficulty separating groups if the proportion of common and unique paths in dynamics are small between groups. Future work should validate these models in the context of smaller dimensional and larger dimensional networks.

Specific to our empirical illustrations, our results were limited by our treatment of the day-to-day and within-day lags as equivalent. Following in the design by De Vos et al. (2017), we treated all time-points as equidistant and kept the lag between the night and the following morning. This may have had the effect of biasing the results of our empirical illustration.

The scGVAR does not return standard error estimates due to its use of regularization and estimates could not be readily obtained for the contemporaneous networks of the ALS VAR. Thus, our power and specificity calculations had to be adjusted using a heuristic rounding rule and may be improved upon with future work. Preliminary simulations revealed issues with the scGVAR where cluster solutions would fail if communities differed based on the presence or absence of an edge. For example, when two groups differ based on whether an effect is present or not. In the instance that these cases are expected to exist within a dataset, we added the functionality, `z.count = TRUE`, which uses the full information of the adjacency matrix when searching for communities. This argument exhibits improved performance when subgroups are differentiated by present or absent edges but comes at the cost of overall performance in other situations and was not included in the simulation discussions. Furthermore, the clustering approach implemented by the scGVAR—and other approaches such as S-GIMME—make use of discrete clustering algorithms when identifying groups. Recent developments in fuzzy clustering may be illuminating avenues for future research given recent calls for psychological research to move from a discretized to a dimensional approaches (Drakopoulos et al., 2016; Kotov et al., 2017). These fuzzy alternatives may be useful in identifying groups of individuals who do not align exactly with larger subgroups and could potentially be used to identify individuals who are exhibiting transitory dynamics from one state to another (Bolin et al., 2014).

Recent works have extended the dynamic network approach to the continuous-time domain (Ryan et al., 2018). These works are further bolstered by work establishing equivalence between discrete-time and continuous-time models (Chow et al., 2021; Demeshko et al., 2015). Future work should attempt to extend these idio-thetic approaches into the continuous-time domain and using the identified equivalences between the discrete- and continuous-forms to see if cluster solutions differ appreciably by methodology.

Article information

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

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

Funding: This research reported in this publication was supported by the National Institutes of Health under grants U24AA027684 and UL1 TR002014; and National Science Foundation grant IGE-1806874 alongside grants UH3 OD023389, 5 R01 HD097189-04, U2C OD023375, and 5 UL1TR002014-06.

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

Acknowledgements: The ideas and opinions expressed herein are those of the authors alone, and endorsement by the authors' institution or the respective funding agencies is not intended and should not be inferred.

ORCID

Sy-Miin Chow http://orcid.org/0000-0003-1938-027X
Sacha Epskamp http://orcid.org/0000-0003-4884-8118

References

Abegaz, F., & Wit, E. (2013). Sparse time series chain graphical models for reconstructing genetic networks. *Biostatistics*, *14*(3), 586–599. https://doi.org/10.1093/biostatistics/kxt005

Beck, E. D., & Jackson, J. J. (2020). Consistency and change in idiographic personality: A longitudinal ESM network study. *Journal of Personality and Social Psychology*, *118*(5), 1080–1100. https://doi.org/10.1037/pspp0000249

Bogner, K., Pappenberger, F., & Cloke, H. L. (2012). The normal quantile transformation and its application in a flood forecasting system. *Hydrology and Earth System Sciences*, *16*(4), 1085–1094. https://doi.org/10.5194/hess-16-1085-2012

Bolin, J. H., Edwards, J. M., Finch, W. H., & Cassady, J. C. (2014). Applications of cluster analysis to the creation of perfectionism profiles: A comparison of two clustering approaches. *Frontiers in Psychology*, *5*, 343. https://doi.org/10.3389/fpsyg.2014.00343

Borsboom, D., & Cramer, A. O. (2013). Network analysis: An integrative approach to the structure of psychopathology. *Annual Review of Clinical Psychology*, *9*(1), 91–121. https://doi.org/10.1146/annurev-clinpsy-050212-185608

Bringmann, L. F., Vissers, N., Wichers, M., Geschwind, N., Kuppens, P., Peeters, F., Borsboom, D., & Tuerlinckx, F. (2013). A network approach to psychopathology: New insights into clinical longitudinal data. *PLoS One*, *8*(4), e60188. https://doi.org/10.1371/journal.pone.0060188

Bulteel, K., Tuerlinckx, F., Brose, A., & Ceulemans, E. (2016). Clustering vector autoregressive models: Capturing qualitative differences in within-person dynamics. *Frontiers in Psychology*, *7*, 1540. https://doi.org/10.3389/fpsyg.2016.01540

Ceulemans, E., & Kiers, H. A. (2006). Selecting among three-mode principal component models of different types and complexities: A numerical convex hull based method. *The British Journal of Mathematical and Statistical Psychology*, *59*(Pt 1), 133–150. https://doi.org/10.1348/000711005X64817

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

Chow, S.-M., Losardo, D., Park, J. J., & Molenaar, P. C. (2021). *Continuous-time dynamic models: Connections to structural equation models and other discrete-time models*. Guilford.

Demeshko, M., Washio, T., Kawahara, Y., & Pepyolyshev, Y. (2015). A novel continuous and structural var modeling approach and its application to reactor noise analysis. *ACM Transactions on Intelligent Systems and Technology*, *7*(2), 1–22. https://doi.org/10.1145/2710025

De Vos, S., Wardenaar, K. J., Bos, E. H., Wit, E. C., Bouwmans, M. E., & De Jonge, P. (2017). An investigation of emotion dynamics in major depressive disorder patients and healthy persons using sparse longitudinal networks. *PLoS One*, *12*(6), e0178586. https://doi.org/10.1371/journal.pone.0178586

Drakopoulos, G., Kanavos, A., Makris, C., & Megalooikonomou, V. (2016). On converting community detection algorithms for fuzzy graphs in neo4j. *arXiv preprint arXiv:1608.02235*.

Epskamp, S., & Fried, E. I. (2018). A tutorial on regularized partial correlation networks. *Psychological Methods*, *23*(4), 617–634. https://doi.org/10.1037/met0000167

Epskamp, S., van Borkulo, C. D., van der Veen, D. C., Servaas, M. N., Isvoranu, A.-M., Riese, H., & Cramer, A. O. (2018). Personalized network modeling in psychopathology: The importance of contemporaneous and temporal connections. *Clinical Psychological Science: A Journal of the Association for Psychological Science*, *6*(3), 416–427. https://doi.org/10.1177/2167702617744325

Epskamp, S., Waldorp, L. J., Mõttus, R., & Borsboom, D. (2018). The gaussian graphical model in cross-sectional and time-series data. *Multivariate Behavioral Research*, *53*(4), 453–480. https://doi.org/10.1080/00273171.2018.1454823

Fisher, A. J. (2015). Toward a dynamic model of psychological assessment: Implications for personalized care. *Journal of Consulting and Clinical Psychology*, *83*(4), 825–836. https://doi.org/10.1037/ccp0000026

Fisher, Z. F., Kim, Y., Fredrickson, B., & Pipiras, V. (2022). Penalized estimation and forecasting of multiple subject intensive longitudinal data. *psychometrika*, *87*(2), 1–29.

Foygel, R., & Drton, M. (2010). Extended Bayesian information criteria for gaussian graphical models. *Advances in neural information processing systems*, *23*.

Friedman, J., Hastie, T., & Tibshirani, R. (2008). Sparse inverse covariance estimation with the graphical lasso. *Biostatistics*, *9*(3), 432–441. https://doi.org/10.1093/biostatistics/kxm045

Gates, K. M., Henry, T., Steinley, D., & Fair, D. A. (2016). A Monte Carlo evaluation of weighted community detection algorithms. *Frontiers in Neuroinformatics*, *10*, 45. https://doi.org/10.3389/fninf.2016.00045

Gates, K. M., Lane, S. T., Varangis, E., Giovanello, K., & Guskiewicz, K. (2017). Unsupervised classification during time-series model building. *Multivariate Behavioral Research*, *52*(2), 129–148. https://doi.org/10.1080/00273171.2016.1256187

Gates, K. M., & Molenaar, P. C. (2012). Group search algorithm recovers effective connectivity maps for individuals in homogeneous and heterogeneous samples. *NeuroImage*, *63*(1), 310–319. https://doi.org/10.1016/j.neuroimage.2012.06.026

Gates, K. M., Molenaar, P. C., Hillary, F. G., Ram, N., & Rovine, M. J. (2010). Automatic search for fMRI connectivity mapping: An alternative to granger causality testing using formal equivalences among SEM path modeling, VAR, and unified SEM. *NeuroImage*, *50*(3), 1118–1125. https://doi.org/10.1016/j.neuroimage.2009.12.117

Gates, K. M., Molenaar, P. C., Iyer, S. P., Nigg, J. T., & Fair, D. A. (2014). Organizing heterogeneous samples using community detection of gimme-derived resting state functional networks. *PLoS One*, *9*(3), e91322. https://doi.org/10.1371/journal.pone.0091322

Hamaker, E. L. (2004). *Time series analysis and the individual as the unit of psychological research* Universiteit van Amsterdam.

Hamaker, E. L., Dolan, C. V., & Molenaar, P. C. (2005). Statistical modeling of the individual: Rationale and application of multivariate stationary time series analysis. *Multivariate Behavioral Research*, *40*(2), 207–233. https://doi.org/10.1207/s15327906mbr4002_3

Haslbeck, J., & Waldorp, L. J. (2015). mgm: Estimating time-varying mixed graphical models in high-dimensional data. *arXiv preprint arXiv:1510.06871*.

Hecht, M., & Zitzmann, S. (2021). Sample size recommendations for continuous-time models: Compensating

shorter time series with larger numbers of persons and vice versa. *Structural Equation Modeling: A Multidisciplinary Journal*, *28*(2), 229–236. https://doi.org/10.1080/10705511.2020.1779069

Henry, T. R., Feczko, E., Cordova, M., Earl, E., Williams, S., Nigg, J. T., Fair, D. A., & Gates, K. M. (2019). Comparing directed functional connectivity between groups with confirmatory subgrouping gimme. *NeuroImage*, *188*, 642–653. https://doi.org/10.1016/j.neuroimage.2018.12.040

Hubert, L., & Arabie, P. (1985). Comparing partitions. *Journal of Classification*, *2*(1), 193–218. https://doi.org/10.1007/BF01908075

Kailath, T. (1967). The divergence and Bhattacharyya distance measures in signal selection. *IEEE Transactions on Communications*, *15*(1), 52–60. https://doi.org/10.1109/TCOM.1967.1089532

Kim, J., Zhu, W., Chang, L., Bentler, P. M., & Ernst, T. (2007). Unified structural equation modeling approach for the analysis of multisubject, multivariate functional MRI data. *Human Brain Mapping*, *28*(2), 85–93. https://doi.org/10.1002/hbm.20259

Kotov, R., Krueger, R. F., Watson, D., Achenbach, T. M., Althoff, R. R., Bagby, R. M., Brown, T. A., Carpenter, W. T., Caspi, A., Clark, L. A., Eaton, N. R., Forbes, M. K., Forbush, K. T., Goldberg, D., Hasin, D., Hyman, S. E., Ivanova, M. Y., Lynam, D. R., Markon, K., … Zimmerman, M. (2017). The hierarchical taxonomy of psychopathology (HiTOP): A dimensional alternative to traditional nosologies. *Journal of Abnormal Psychology*, *126*(4), 454–477. https://doi.org/10.1037/abn0000258

Kuppens, P., Allen, N. B., & Sheeber, L. B. (2010). Emotional inertia and psychological maladjustment. *Psychological Science*, *21*(7), 984–991. https://doi.org/10.1177/0956797610372634

Lane, S. T., Gates, K. M., Pike, H. K., Beltz, A. M., & Wright, A. G. (2019). Uncovering general, shared, and unique temporal patterns in ambulatory assessment data. *Psychological Methods*, *24*(1), 54–69. https://doi.org/10.1037/met0000192

Liu, S., Ou, L., & Ferrer, E. (2020). Dynamic mixture modeling with dynr. *Multivariate Behavioral Research*, *56*(6), 941–955. https://doi.org/10.1080/00273171.2020.1794775

Lundh, L.-G. (2015). The person as a focus for research-the contributions of windelband, stern, allport, lamiell, and magnusson. *Journal for Person-Oriented Research*, *1*(1-2), 15–33. https://doi.org/10.17505/jpor.2015.03

Lütkepohl, H. (2005). *New introduction to multiple time series analysis*. Springer Science & Business Media.

Lydon-Staley, D. M., Xia, M., Mak, H. W., & Fosco, G. (2019). Adolescent emotion network dynamics in daily life and implications for depression. *Journal of Abnormal Child Psychology*, *47*(4), 717–729. https://doi.org/10.1007/s10802-018-0474-y

McCutcheon, A. L. (1987). *Latent class analysis* (Number 64). Sage.

Molenaar, P. C. (2004). A manifesto on psychology as idiographic science: Bringing the person back into scientific psychology, this time forever. *Measurement: Interdisciplinary Research & Perspective*, *2*(4), 201–218. https://doi.org/10.1207/s15366359mea0204_1

Muthen, B. (2001). Latent variable mixture modeling. *New Developments and Techniques in Structural Equation Modeling*, *2*, 1–33.

Park, J. J., Chow, S.-M., Fisher, Z. F., & Molenaar, P. C. (2021). Affect and personality: Ramifications of modeling (non-) directionality in dynamic network models. *European Journal of Psychological Assessment*, *36*(6), 1009–1023. https://doi.org/10.1027/1015-5759/a000612

Park, J. J., Chow, S., & Molenaar, P. (2023). What the fuzz!? leveraging ambiguity in dynamic network models. https://doi.org/10.31234/osf.io/ehtkf

Pons, P., & Latapy, M. (2005). Computing communities in large networks using random walks. In *Computer and Information Sciences-ISCIS 2005: 20th International Symposium, Istanbul, Turkey, October 26–28, 2005. Proceedings 20* (pp. 284–293). Springer Berlin Heidelberg.

Price, R. B., Lane, S., Gates, K., Kraynak, T. E., Horner, M. S., Thase, M. E., & Siegle, G. J. (2017). Parsing heterogeneity in the brain connectivity of depressed and healthy adults during positive mood. *Biological Psychiatry*, *81*(4), 347–357. https://doi.org/10.1016/j.biopsych.2016.06.023

Rothman, A. J., Levina, E., & Zhu, J. (2010). Sparse multivariate regression with covariance estimation. *Journal of Computational and Graphical Statistics: A Joint Publication of American Statistical Association, Institute of Mathematical Statistics, Interface Foundation of North America*, *19*(4), 947–962. https://doi.org/10.1198/jcgs.2010.09188

Runyan, W. M. (1983). Idiographic goals and methods in the study of lives. *Journal of Personality*, *51*(3), 413–437. https://doi.org/10.1111/j.1467-6494.1983.tb00339.x

Ryan, O., Kuiper, R. M., & Hamaker, E. L. (2018). A continuous-time approach to intensive longitudinal data: What, why, and how?. In: Van montfort, K., Oud, J., Voelkle, M. (Eds.), *Continuous time modeling in the behavioral and related sciences* (pp. 27–54). Springer.

Schweren, L., van Borkulo, C. D., Fried, E., & Goodyer, I. M. (2018). Assessment of symptom network density as a prognostic marker of treatment response in adolescent depression. *JAMA Psychiatry*, *75*(1), 98–100. https://doi.org/10.1001/jamapsychiatry.2017.3561

Sinclair, A., & Jerrum, M. (1989). Approximate counting, uniform generation and rapidly mixing Markov chains. *Information and Computation*, *82*(1), 93–133. https://doi.org/10.1016/0890-5401(89)90067-9

Steinley, D. (2004). Properties of the Hubert-arable adjusted rand index. *Psychological Methods*, *9*(3), 386–396. https://doi.org/10.1037/1082-989X.9.3.386

Takano, K., Stefanovic, M., Rosenkranz, T., & Ehring, T. (2021). Clustering individuals on limited features of a vector autoregressive model. *Multivariate Behavioral Research*, *56*(5), 768–786. https://doi.org/10.1080/00273171.2020.1767532

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society: Series B (Methodological)*, *58*(1), 267–288. https://doi.org/10.1111/j.2517-6161.1996.tb02080.x

Van Borkulo, C. D., Borsboom, D., Epskamp, S., Blanken, T. F., Boschloo, L., Schoevers, R. A., & Waldorp, L. J. (2014). A new method for constructing networks from binary data. *Scientific Reports*, *4*(1), 5918. https://doi.org/10.1038/srep05918

Ward, J. H. (1963). Hierarchical grouping to optimize an objective function. *Journal of the American Statistical Association*, *58*(301), 236–244. https://doi.org/10.1080/01621459.1963.10500845

Williams, D. R., & Rast, P. (2020). Back to the basics: Rethinking partial correlation network methodology. *The British Journal of Mathematical and Statistical Psychology*, *73*(2), 187–212. https://doi.org/10.1111/bmsp.12173

Wright, A. G., Gates, K. M., Arizmendi, C., Lane, S. T., Woods, W. C., & Edershile, E. A. (2019). Focusing personality assessment on the person: Modeling general, shared, and person specific processes in personality and psychopathology. *Psychological Assessment*, *31*(4), 502–515. https://doi.org/10.1037/pas0000617

Appendix A
Full description of the scGVAR algorithm

The approach for scGVAR models can be seen in Figure 2. In the first step, GVAR models are fitted to all subjects. Following this, an $N \times N$ adjacency matrix is generated, $\mathbf{A}$. Any element, $a_{i,j}$ where $i \neq j$, indicates the number of structural paths the i^{th} and j^{th} subject share in common. In the case of the scGVAR, "common" refers to both the PDCs and the PCCs. In the case of the PDCs, the i^{th} and j^{th} subjects are common if they share—or mutually do not share—a structural edge in the same direction between two nodes as well as the same sign (±). In the case of the PCCs, "common" is defined solely by common sign between the same two nodes. A similar approach has been implemented by the subgrouping algorithm in the Group Iterative Multiple Model Estimation (S-GIMME) procedure (Gates et al., 2017). Next, sparsity is introduced into $\mathbf{A}$ by identifying the configuration of $\mathbf{A}$ which maximizes the conductance of the network in an iterative procedure which is detailed below.

WalkTrap

The procedure begins by first estimating a cluster solution with WalkTrap (Pons and Latapy, 2005). WalkTrap is a community detection algorithm which generates a matrix of transition probabilities for any pair of nodes in a network. Each element of the transition matrix then corresponds to the probability of going from one node to the other following a random walk of a given length, typically four steps (Gates et al., 2016). A hierarchical clustering method such as Ward's method (Ward, 1963) is then applied to the transition probabilities matrix to derive communities with the lowest within-cluster variance. WalkTrap is an unsupervised community detection method which operates by maximizing the modularity of a given network; however, modularity maximization has frequently been criticized due to issues related to its ability to enumerate smaller communities when networks are comprised of varying sizes (Gates et al., 2016). An adjustment was made to the WalkTrap procedure to ameliorate the influence that modularity maximization has on the selected solution by maximizing conductance.

Conductance

Conductance can be conceptualized as the relative number of edges entirely contained within a cluster relative to edges between a cluster and other clusters. Thus, high conductance would be when subjects in a subgroup are entirely connected to each other and subgroups share nothing in common with one another. Formally, we can see this expressed as:

$$\omega(S_k) = \frac{\sum_{i \in S_k j \notin S_k} a_{ij}}{\min\left[\mathbf{A}(S_k), \mathbf{A}(\bar{S}_k)\right]} \tag{13}$$

where $\mathbf{A}$ is the adjacency matrix, S_k is the k^{th} subgroup, and the denominator is which ever value is lowest between $\mathbf{A}(S_k)$—the total number of edges between the i^{th} subject to all other members of the k^{th} subgroup to which they are assigned—or $\mathbf{A}(\bar{S}_k)$—the total number of edges between the i^{th} subject and to subjects in unrelated subgroups. In this case, we can see that lower pair-wise conductance values would indicate that a pair of subgroups is more orthogonal whereas higher values would indicate a greater degree of overlap between clusters. The conductance of a full graph can then be expressed as the arithmetic average of all pair-wise conductance values:

$$\omega(G) = 1 - \frac{1}{K}\sum_{k} \omega(S_k) \tag{14}$$

where we have also subtracted from 1 so that higher conductance values indicate better delineation between subgroups. Sparsity is induced in the $\mathbf{A}$ matrix by subtracting the minimum value of $\mathbf{A}$ from all cells generating an adjusted adjacency matrix, $\mathbf{A}_{adj}$. The WalkTrap algorithm is then applied to the sparse matrix, $\mathbf{A}_{adj}$, and conductance is calculated. Following this, greater sparsity is induced by taking the adjusted adjacency matrix and subtracting the minimum value from all cells of $\mathbf{A}_{adj}$ again. This process is repeated iteratively until an optimal subset of the adjacency matrix is found which produces a cluster solution with the highest conductance value. Once the optimal partition is found, a cGVAR model is estimated for each subgroup and returned to the user.