Subgrouping with Chain Graphical VAR Models

Jonathan J. Park[1], Sy-Miin Chow[1], Sacha Epskamp[2], and Peter C. M. Molenaar[1]

[1]Department of Human Development and Family Studies, The Pennsylvania State University

[2]Department of Psychology, National University of Singapore

Author Note

Research reported in this publication was supported by the Intensive Longitudinal Health Behavior Cooperative Agreement Program funded by the National Institutes of Health under Award Number U24AA027684; and National Science Foundation grant IGE-1806874. Correspondence concerning this article should be addressed to Jonathan J. Park, the Pennsylvania State University, 029 Health and Human Development Building, University Park, PA, 16802 or by e-mail at jjp409@psu.edu.

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

Word Count: 8131

Subgrouping with Chain Graphical VAR Models

In the past decade, network models have seen a great deal of use in the social and behavioral sciences. These network models have pushed the study of human behavior and psychology to new heights by encouraging researchers to consider psychological constructs in a different light (Borsboom & Cramer, 2013). More recently, networks have been increasingly applied to intensive longitudinal data (i.e., data where $T > 50$ time-points Beck & Jackson, 2020; Park et al., 2021) in what are referred to as dynamic network models. These "dynamic" network models are in-line with recent calls for the social and behavioral sciences to attend to idiographic phenomena (A. J. Fisher, 2015; Molenaar, 2004). However, a criticism of idiographic research pertains to the perceived lack of inference $N = 1$ studies may provide for the general populace and vice versa: the generalizability of nomothetic results to idiographic studies (Hamaker et al., 2005; Molenaar, 2004).

This 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-20^{th} century, a series of debates took place regarding the placement of psychology as an idiographic or nomothetic science (see Runyan, 1983). However, recent years have seen a blurring in this distinction between idiographic and nomothetic views of the behavioral sciences on both the methodological and substantive areas of the field. Relevant to the current work, several methods have been proposed in recent years for bridging the divide between nomothetic and idiographic research.

We refer to the class of methods that blur idiographic and nomothetic lines of inquiry as "idio-thetic" methods (Beck & Jackson, 2020). Idio-thetic methods have seen a wide range of development including mixed-effects modeling, latent class analysis (McCutcheon, 1987), and mixture modeling (Muthen, 2001). In the context for this paper, we propose a novel idio-thetic extension to one dynamic network approach and

compare its performance to a benchmark alternative. Idio-thetic methods in the dynamic network literature have already seen use in various fields ranging from major depressive disorder (MDD; De Vos et al., 2017; Lydon-Staley et al., 2019), emotion dynamics (Park et al., 2021; Wright et al., 2019), and ADHD (Gates et al., 2014). In all of these fields, it is critically important that signal be differentiated from noise. This can typically be accomplished by aggregating or pooling information across individuals to identify trends that exist for a plurality of subjects (Gates & Molenaar, 2012), or by assuming a fixed structure that applies to all individuals (Bulteel et al., 2016). The strengths of the idio-thetic approach resides in its flexibility to identify broad group-level insights from individual-level data or vice versa.

Some examples of idio-thetic dynamic network methods include the multilevel VAR (Bringmann et al., 2013), the multi-var (Z. F. Fisher et al., 2020), the Group Iterative Multiple Model Estimation procedure and its subgrouping extension (GIMME and S-GIMME; Gates et al., 2017; Gates & Molenaar, 2012), and–more recently–the alternating least squares vector autoregressive model (ALS VAR; Bulteel et al., 2016; Takano et al., 2020). These idio-thetic approaches all vary in the degree to which they allow for group- and individual-specific expressions in dynamic networks. For instance, S-GIMME identifies group- and subgroup-specific paths and fixes those structures prior to allowing for the estimation of person-specific effects (Gates et al., 2017). In contrast, the multilevel VAR assumes all individuals share a common dynamic network structure but are allowed to vary to some degree but does not incorporate a means for clustering individuals into subgroups. While this broad class of methods is still undergoing a large boom in development, some recent methodologies are noteworthy.

Most contemporary dynamic network clustering algorithms operate by identifying subgroups of individuals with similar over-time relations among a series of variables, as captured using 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 $t-1$. Clustering approaches motivated by this modeling framework focus on detecting idio-thetic dynamics that remain strong across successive measurement occasions. This is in contrast to alternative clustering approaches, such as S-GIMME, which emphasize similarities in the directionality of contemporaneous relations among variables, the strengths of which would have greatly diminished from one measurement occasion to the next.

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 have relatively accurate recovery of parameters in simulation studies, with several useful properties. (Bulteel et al., 2016; Takano et al., 2020). Figure 1A demonstrates some configurations where the ALS VAR may perform quite well in enumerating subgroups of subjects. In theory, the ALS VAR can differentiate groups which differ by the magnitude of edges; specifically (Figure 1A; Takano et al., 2020). This would be most likely to occur in instances comparing clinical samples to controls where past research has suggested denser connectivity among the clinically depressed relative to healthy controls (CITE) and is feature particularly unique to 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 1B).

Insert Figure 1

The current work introduces an alternative approach to the ALS VAR: the subgrouped chain graphical VAR (scGVAR). Similar to the ALS VAR, the scGVAR 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, some contemporaneous information is incorporated when conducting the community detection.

Vector Autoregression

We begin this section by detailing the standard vector autoregressive (VAR) model which forms the basis of the graphical VAR and the ALS VAR methods. 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, μ 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 graphical VAR (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). Further standardization can be applied to the matrices of the standard VAR model to obtain coefficients of the partial directed correlations (PDCs) and the partial contemporaneous correlations (PCCs) which correspond to the lagged

effects and the contemporaneous covariances in the standard VAR, respectively with the equations below:

$$PCC_{(\eta_i,\eta_j)} = -\frac{k_{ij}}{\sqrt{k_{ii}k_{jj}}} \tag{3}$$

$$PDC_{(\eta_i,\eta_j)} = \frac{\phi_{ij}}{\sqrt{\psi_{ii}k_{jj} + \phi_{ij}^2}} \tag{4}$$

where ϕ_{ij} is the lag(1) regression weight of the j^{th} variable onto the i^{th} variable, and ψ_{ii} is the variance of the i^{th} element of $\zeta_{i,t}$ upon itself.

Estimation of Graphical VAR Models

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" towards 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 $\mathbf{K}$ and $\mathbf{\Phi}$ 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 $\sum_{j,k} |\hat{\phi}_{ijk}^{(m+1)} - \hat{\phi}_{ijk}^{(m)}| < \epsilon \sum_{j,k} |\hat{\phi}_{ijk}^{\text{RIDGE}}|$ 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 towards 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.

Subgrouping Graphical VAR Models

Upon fitting N GVAR models, one may be interested in identifying a group-level structure. This can currently be accomplished with LASSO regularization the `graphicalVAR` package to fit a chain GVAR. A chain GVAR is produced by "chaining" or stacking together the time series of multiple subjects to produce a set of group-based estimates across subjects. However, samples may be comprised of multiple smaller groups whom share similarities with other subjects and form cohesive subgroups. Subgrouping of VAR models is not new and many approaches currently exist for clustering individuals in the VAR or related frameworks (Bulteel et al., 2016; Gates et al., 2017; Takano et al., 2020). Ultimately, the group of subgrouping with VAR models is to derive idio-thetic inference. That is, identify groups of individuals who are relatively homogeneous.

The approach for subgrouping GVARs can be seen in Figure 2 and is henceforth referred to as a subgrouped chain graphical VAR (scGVAR). 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 and the same sign ($\pm$). 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

A by identifying the configuration of **A** which maximizes the conductance of the network in an iterative procedure.

The procedure begins by first estimating a cluster solution with WalkTrap (Pons & Latapy, 2006). The cluster solutions from WalkTrap are then assessed for their conductance. Conductance can be conceptualized as the relative number of edges entirely contained within a cluster relative to edges between different clusters. Thus, subjects are more similar if they share more structural similarities with in-group members and relatively few structural similarities with members outside their group. Formally, we can see this expressed as:

$$\omega(S_k) = \frac{\sum\limits_{i \in S_k j \notin S_k} a_{ij}}{min[A(S_k), A(\bar{S}_k))]} \tag{7}$$

where **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 associated with the k^{th} subgroup–or $\mathbf{A}(\bar{S}_k)$–the total number of edges that are unrelated to the k^{th} subgroup. 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.

Insert Figure 2

The conductance of a full graph can then be expressed as the arithmetic average of all pairwise conductance values:

$$\omega(G) = 1 - \frac{1}{K}\sum_{k}\omega(S_k) \tag{8}$$

where we have also subtracted from 1 so that higher conductance values indicate better delineation between subgroups.

Sparsity is then induced into the **A** matrix to a greater degree by subtracting the next minimum value from all cells in **A** and the process is repeated until an optimal subset of **A** is found which produces a cluster solution with the highest conductance

value. Once the optimal partition is found, a chain GVAR model is estimated for each subgroup.

The results of the scGVAR algorithm include a sample-level chained GVAR, K chained GVAR models for each 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.

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 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.

Estimation of the ALS VAR

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

$$L_K = \sum_{i=1}^{I} \sum_{t=2}^{T_i} (\eta_{it} - \hat{\eta}_{it})^2 \tag{9}$$

where L_K is the sum of squared prediction errors and $\hat{\eta}_{it}$ is the $p \times 1$ vector of predicted scores for the i^{th} subject at time, t.

Bulteel et al. (2016) proposed a four step procedure for estimating ALS VAR models which are presented in Figure 2 below 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 effects (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) is used to identify K clusters of individuals who are closest in their Euclidean distances. The scGVAR in contrast utilizes the WalkTrap algorithm. 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 9).

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 K from $K = 1$ to K_{max}:

$$st_k = \frac{L_{K-1} - L_K}{L_K - L_{K+1}} \tag{10}$$

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., 2020). 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.

Simulation Study

A Monte Carlo simulation study was conducted to evaluate several key questions regarding the performance of these algorithms: 1. How well do the algorithms perform as the effective distance between them increases? 2. What happens as the number of subgroups increases? 3. Does the performance of these algorithms improve with a greater number of time-points? 4. How well do these algorithms perform when the dynamics occur when the sampling rate is (mis)matched to the true processes?

Several design factors were considered to address these questions including:

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

10-variate stationary VAR(1) models were randomly simulated for $N = 52$ subjects across the 500 MC runs. When 2 clusters were present, each subgroup consisted of 26-subjects whereas when 4 clusters were present, each subgroup consisted of 13-subjects. Auto-regressive elements of the VAR models were fixed to 0.70 for all subjects. Following this, 10 cross-regressions were added to every subject's model according to their simulated subgroup membership. Of the 10-total cross-regressions, 3-

or 9-edges would be specific to one subgroup with the remaining edges being common to all subgroups; or common paths. In small-scale tests, this number of unique distances corresponded with significant differences in the Bhattcharaya distance (Kailath, 1967); an extension of the Mahalanobis distance which quantifies the distance between two multivariate distributions. The values of 3- and 9-edges were selected as they were settings which corresponded to relatively weak and large separations between the simulated VAR models.

The number of time-points were meant to reflect our empirical illustration ($T = 100$) as well as provide an asymptotic assessment of the performance of the algorithms ($T = 500$). Finally, we simulated a condition where all dynamics occurred contemporaneously (i.e., faster than the VAR(1) sampling rate) to highlight the performance of the scGVAR and its ability to capture "faster" dynamics.

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., 2020). Further, the ALS VAR should perform quite well when subgroups are highly differentiated and when the number of clusters is 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 over time. Thus, the ALS VAR should perform better when $D = 9$ and when $C = 2$. In contrast, the scGVAR should perform well at C increases. The time parameters were selected to mirror our empirical illustration (e.g., $T = 100$) and provide potentially asymptotic performance measures ($T = 500$). Finally, the ALS VAR does not explicitly estimate contemporaneous effects and thus we tested this condition to highlight the strengths of the scGVAR.

When comparing the contemporaneous networks, the model implied residual covariance matrix was constructed from the results of the ALS VAR output and compared to the true residual covariance matrices; however, due to the relative complexity in additionally computing standard errors from this approach, power and specificity calculations of the contemporaneous networks are calculated similarly to how the scGVAR metrics were calculated.

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.

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 estiamted group membership. The ARI can be

calculated as:

$$ARI_{HA} = \frac{\binom{N}{2}(a+d) - [(a+b)(a+c)+(c+d)(b+d)]}{\binom{N}{2}^2 - [(a+b)(a+c)+(c+d)(b+d)]} \tag{11}$$

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).

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:

$$\text{Bias}(\theta) = \frac{1}{H}\sum_{h=1}^{H}(\hat{\theta}_h - \theta) \tag{12}$$

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.

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 2-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:

$$Spec(\theta) = \frac{TN_\theta}{TN_\theta + FP_\theta} \tag{13}$$

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:

$$RMSE(\theta) = \sqrt{\frac{1}{H}\sum_{h=1}^{H}(\hat{\theta}_h - \theta)^2} \tag{14}$$

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.

Insert Figure 3

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.

Insert Figure 4

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. The scGVAR biases tended to decrease with greater Bhattcharaya distances; however, the biases did not tend to decrease with greater sample-sizes. 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.

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 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 3. 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.

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 2-group solution where 2 of the true 4 groups would be merged together.

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 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.

Contemporaneous. 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 4.

Interestingly, the biases in the ALS VAR were relatively low for the

Table 1
Table of simulation results comparing the scGVAR to the ALS VAR for the 2-subgroup and 4-subgroup conditions across different sample sizes, T and distances D. Isolated rows indicate conditions where only contemporaneous effects we generated.

	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
$T_{500};D_3;G_4$	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

contemporaneous networks with values of -0.0002 in both conditions. Further, the RMSEs for the recovered contemporaneous networks 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 consisent 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.

Simulation Review. Overall, the ALS VAR produced relatively unbiased estimates of both the lagged- and contemporaneous-networks. Further, the algorithm tended to exhibit high power and acceptable levels of specificity in the lagged conditions when only 2-clusters were present.

In the $C = 2$ conditions, the ALS VAR had excellent recovery of the underlying

subgroups; however, it had difficulty in enumerating the proper number of subgroups when $C = 4$. 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 Bhattcharaya distance configurations, all networks still shared non-zero autoregressions and 7- or 1-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 4 communities as the ALS VAR tended to enumerate only 2 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., $C = 4$; and larger Bhattcharaya 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 1A; Takano et al., 2020). 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 1B). In contrast, when a large degree of common edges are expected with a small number of differentiating points, the scGVAR may be preferred (Figure 1C). 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 1D). 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).

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 posess 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 ($n = 27$) or control groups ($n = 27$) who were pair-matched by age, gender, smoking behavior, and BMI. However, due to convergence issues, 7-participants were excluded from the analysis ($N = 48$; $n_{MDD} = 24$; $n_{control} = 23$). Participants were asked to respond to survey measures 3-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.

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 7-point Likert-type scale and had been previously used in another study (De Vos et al., 2017).

Data Preprocessing

All variables were linearly detrended as well as centered using their within-subject means. Missing data were approximately 8.2% and 6.8% in the MDD and control groups, respectively. Missing data were not imputed for the current investigation. Thus, the average length of time for subjects was approximately 83.2-observations ($SD = 7.4$; De Vos et al., 2017).

Analysis

The scGVAR was fitted to the entire sample in two ways: first, using the top-down diagnostic markers provided by the dataset. 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. The selection of the γ-hyperparameter is left to the researcher's discretion. For the current study, 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. Both models and their results are discussed below.

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 5; 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.

Insert Figure 5

Results from the scGVAR analysis did not fit 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 6).

Insert Figure 6

Group- and Subgroup-level Analysis for scGVAR. The group-level network produced by the scGVAR is provided in Figure 7 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.

Insert Figure 7

Overall, the scGVAR delineated 2 major subgroups of $N > 2$ individuals (Figure 8). All other subjects were classified as singletons or simple dyads, $N = 17$. The first recovered subgroup contained 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 3-measurements per day in the current investigation (Park et al., 2021; Wright et al., 2019).

Insert Figure 8

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 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.

Subgroup-level Analysis for ALS VAR. The ALS VAR converged on a 2 subgroup solution. Each subgroup was comprised of $n_1 = 17; n_2 = 33$ participants, respectively and exhibited drastically different affect dynamics. Similar to the findings of the 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 wuth 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 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.

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–in both instances–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–as noted earlier–is in line with prior studies on affect which suggest that these dynamics may take place faster than our sampling window of 3-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 exist 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. Of course, this also depends on the actual timescale of events.

Further, 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 8) 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 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.

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., 2020; Gates et al., 2017). These methods are invaluable in the study of human behavior as we attempt to understand individuals

at treatment or clinical levels as well as garner a stronger understanding of processes in the general populace. In the current work, we introduced the scGVAR as a novel idio-thetic method for estimating chain graphical VAR 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, unlike multilevel VAR and the ALS VAR, 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 similar 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 efects (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 2 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 return similar cluster solutions, the ALS VAR may be preferable as one may infer that there is

agreement between the more conservative ALS VAR and the more liberal scGVAR in the number of subgroups. Thus, the preference would be towards the less biased point-estimates of the ALS VAR. 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 merge similar yet distinct groups if they share a large degree of similarity. 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 CHULL procedure.

Further, 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 different network characteristics than 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.

Limitations

Our study is limited by several factors. Our simulation 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 contemporaneous dynamics (e.g., S-GIMME; Gates et al., 2017). Thus, future investigations should attempt to compare these related methodologies against one another. Recent work by Liu et al. (2020) has also extended the ALS VAR to incorporate random effects in a methodology dubbed the dynamic mixture model (DMM). It would be interesting to see the extent to which solutions from the ALS VAR and the scGVAR differ from solutions derived by 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.

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.

Further, 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 continous-forms to see if cluster solutions differ appreciably by methodology.

References

Abegaz, F., & Wit, E. (2013). Sparse time series chain graphical models for reconstructing genetic networks. *Biostatistics*, *14*(3), 586–599.

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.

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.

Borsboom, D., & Cramer, A. O. (2013). Network analysis: An integrative approach to the structure of psychopathology. *Annual review of clinical psychology*, *9*, 91–121.

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.

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.

Ceulemans, E., & Kiers, H. A. (2006). Selecting among three-mode principal component models of different types and complexities: A numerical convex hull based method. *British journal of mathematical and statistical psychology*, *59*(1), 133–150.

Chen, J., & Chen, Z. (2008). Extended bayesian information criteria for model selection with large model spaces. *Biometrika*, *95*(3), 759–771.

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*. New York: Guilford.

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.

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 (TIST)*, *7*(2), 1–22.

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.

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*, *6*(3), 416–427.

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.

Fisher, A. J. (2015). Toward a dynamic model of psychological assessment: Implications for personalized care. *Journal of consulting and clinical psychology*, *83*(4), 825.

Fisher, Z. F., Kim, Y., Fredrickson, B., & Pipiras, V. (2020). Penalized estimation and forecasting of multiple subject intensive longitudinal data. *arXiv preprint arXiv:2007.05052*.

Foygel, R., & Drton, M. (2010). Extended bayesian information criteria for gaussian graphical models. *arXiv preprint arXiv:1011.6640*.

Friedman, J., Hastie, T., & Tibshirani, R. (2008). Sparse inverse covariance estimation with the graphical lasso. *Biostatistics*, *9*(3), 432–441.

Gates, K. M., Lane, S. T., Varangis, E., Giovanello, K., & Guiskewicz, K. (2017). Unsupervised classification during time-series model building. *Multivariate behavioral research*, *52*(2), 129–148.

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.

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.

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.

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

Hubert, L., & Arabie, P. (1985). Comparing partitions. *Journal of classification*, *2*(1), 193–218.

Kailath, T. (1967). The divergence and bhattacharyya distance measures in signal selection. *IEEE transactions on communication technology*, *15*(1), 52–60.

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., et al. (2017). The hierarchical taxonomy of psychopathology (hitop): A dimensional alternative to traditional nosologies. *Journal of abnormal psychology*, *126*(4), 454.

Kuppens, P., Allen, N. B., & Sheeber, L. B. (2010). Emotional inertia and psychological maladjustment. *Psychological science*, *21*(7), 984–991.

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.

Liu, S., Ou, L., & Ferrer, E. (2020). Dynamic mixture modeling with dynr. *Multivariate Behavioral Research*, 1–15.

Lundh, L.-G. (2015). The person as a focus for research—the contributions of windelband, stern, allport, lamiell, and magnusson. *Journal of Person-Oriented Research*, *1*(1-2), 15–33.

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.

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

Molenaar, P. C. (2004). A manifesto on psychology as idiographic science: Bringing the person back into scientific psychology, this time forever. *Measurement*, *2*(4), 201–218.

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.

Pons, P., & Latapy, M. (2006). Computing communities in large networks using random walks. *J. Graph Algorithms Appl*.

Rothman, A. J., Levina, E., & Zhu, J. (2010). Sparse multivariate regression with covariance estimation. *Journal of Computational and Graphical Statistics*, *19*(4), 947–962.

Runyan, W. M. (1983). Idiographic goals and methods in the study of lives. *Journal of Personality*, *51*(3), 413–437.

Ryan, O., Kuiper, R. M., & Hamaker, E. L. (2018). A continuous-time approach to intensive longitudinal data: What, why, and how? *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.

Steinley, D. (2004). Properties of the hubert-arable adjusted rand index. *Psychological methods*, *9*(3), 386.

Takano, K., Stefanovic, M., Rosenkranz, T., & Ehring, T. (2020). Clustering individuals on limited features of a vector autoregressive model. *Multivariate Behavioral Research*, 1–19.

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society: Series B (Methodological)*, *58*(1), 267–288.

Ward, J. H. (1963). Hierarchical grouping to optimize an objective function. *Journal of the American statistical association*, *58*(301), 236–244.

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

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.

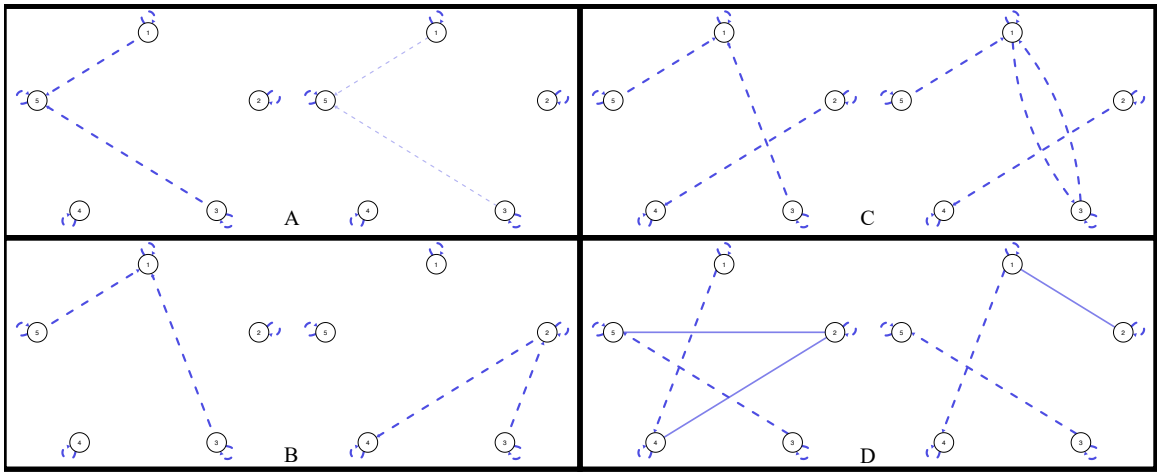

Figure 1. Optimal settings for the ALS VAR (A & B) and the scGVAR (C & D). Dashed edges represent lagged effects while solid edges indicate contemporaneous effects. More vibrant edges are stronger effects while faded edges are weaker. The ALS VAR outperforms the scGVAR in cases where: A) subgroups differ 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

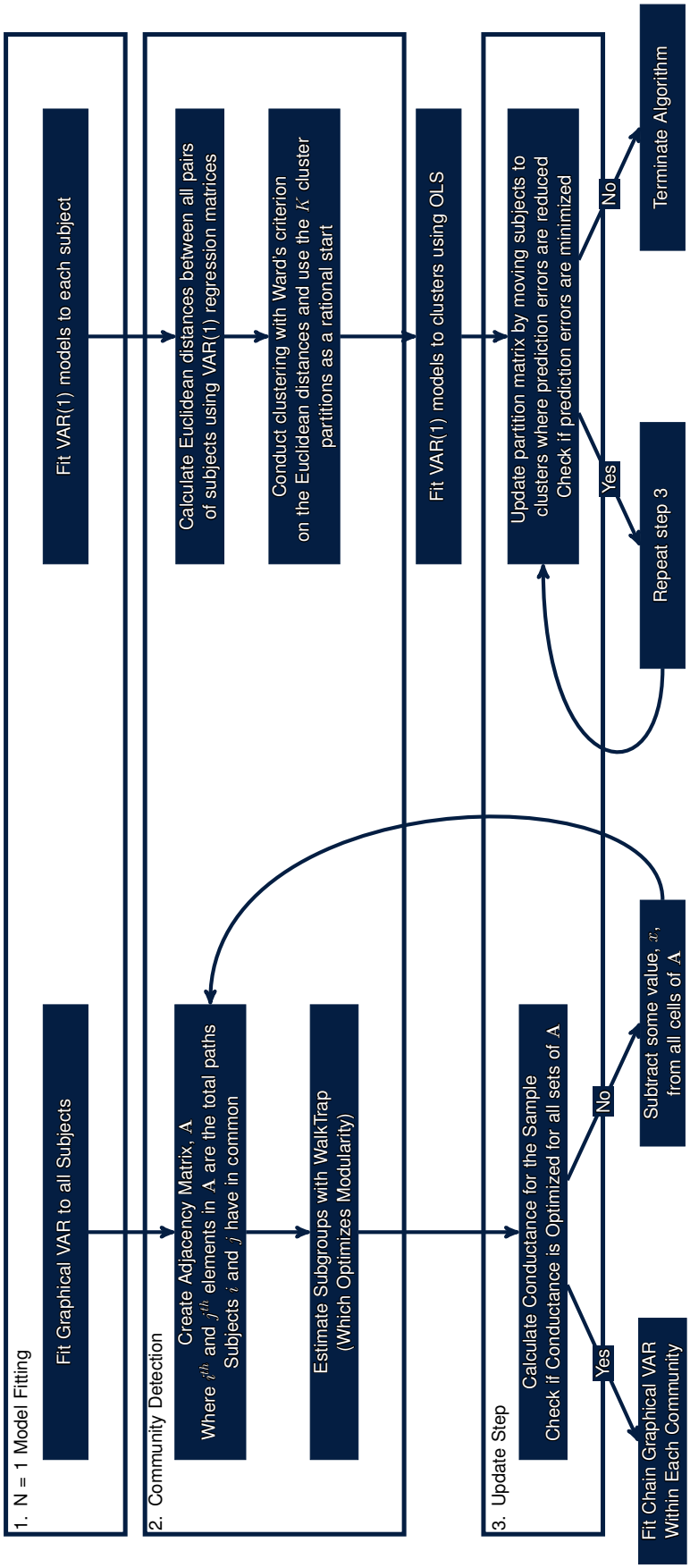

Figure 2. Procedure for scGVAR (Left) and the ALS VAR (Right). BOxes indicate common steps while dark boxes highlight algorithm-specific steps taken by the scgVAR and ALS VAR, respectively.

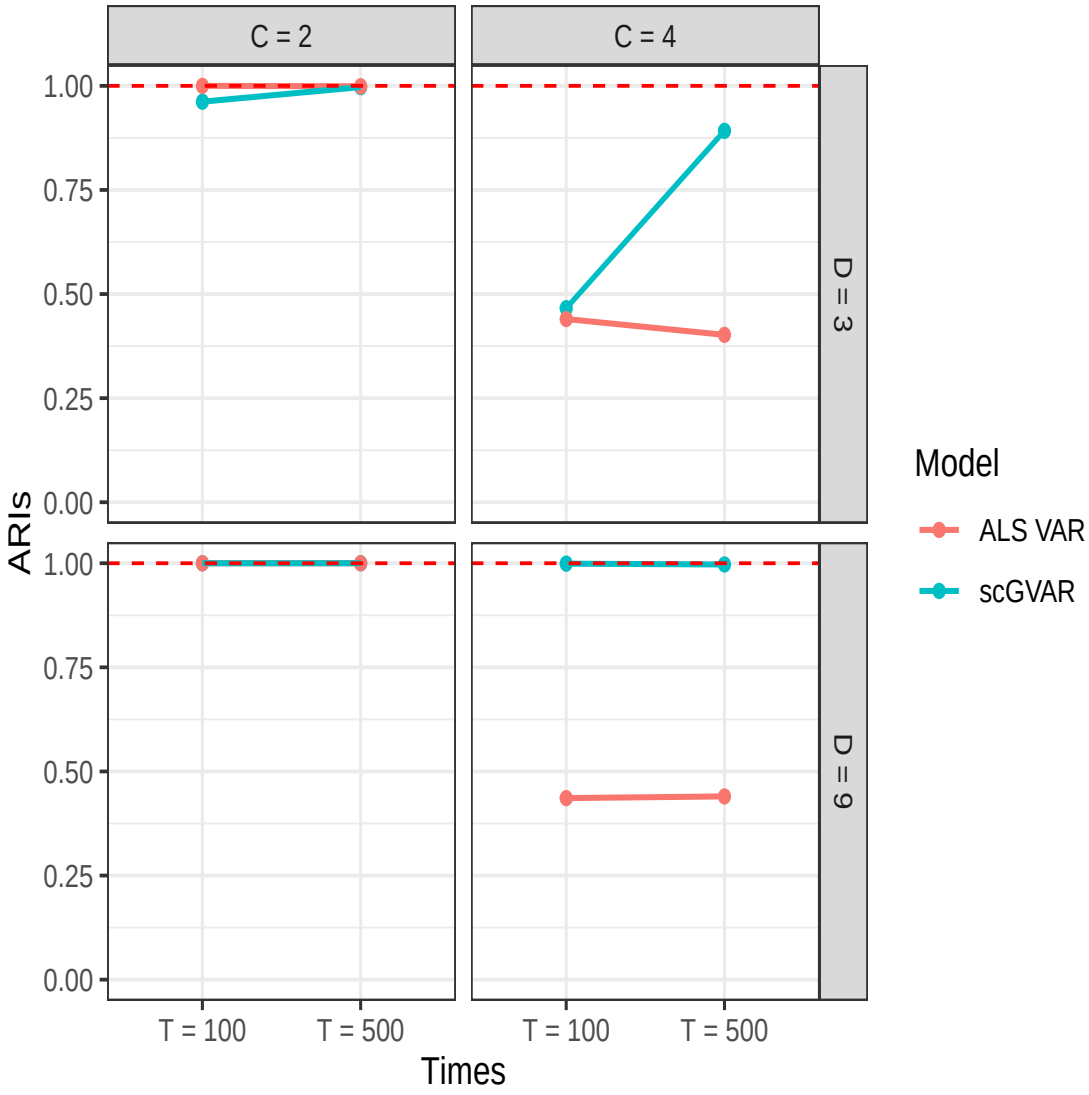

Figure 3. ARI values across simulation conditions. C indicates the number of clusters, D indicates the number of cluster-specific features, and T took on values of 100 or 500. Higher ARI values indicate better accuracy in subgrouping performance. The scGVAR tended to perform better with larger separation between subgroups and larger T.

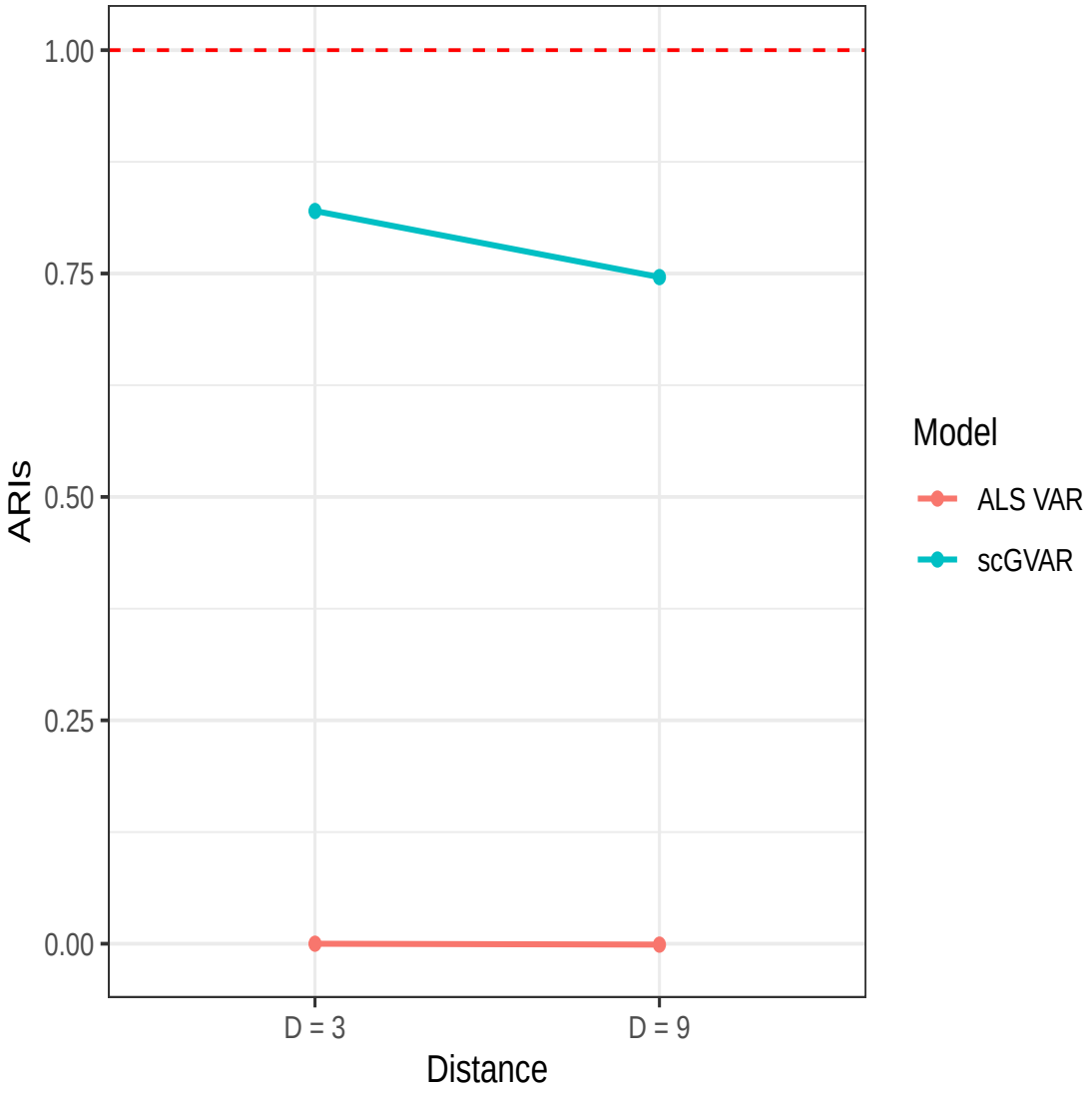

Figure 4. ARI values for the contemporaneous-only simulations. *C* indicates the number of clusters, *D* indicates the number of cluster-specific features, and *T* took on values of 100 or 500. Higher ARI values indicate better accuracy in subgrouping performance. The scGVAR tended to perform better than the ALS VAR in terms of subgrouping accuracy for contemporaneous-only conditions.

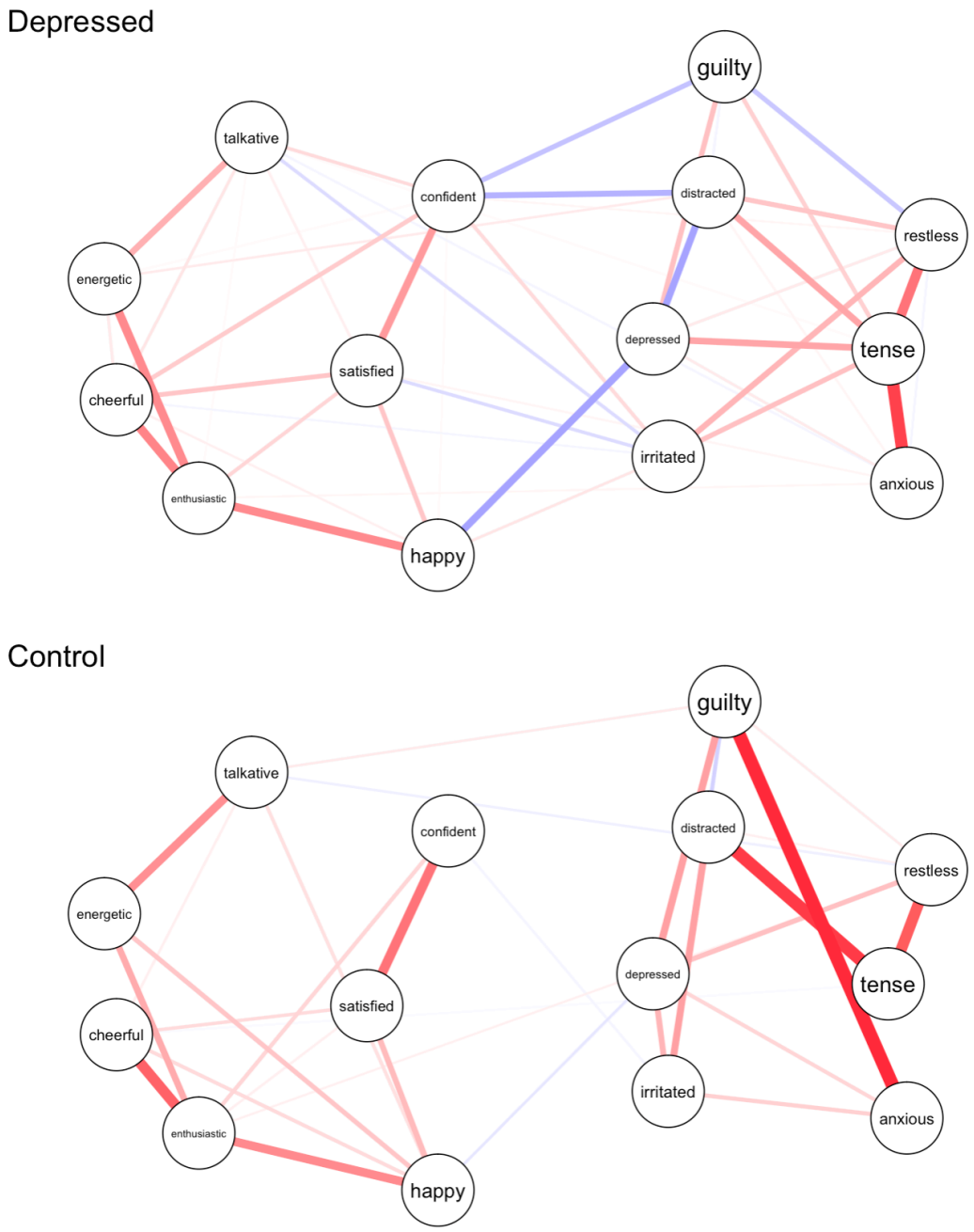

Figure 5. Between-subjects networks from the scGVAR function. Red edges indicate positive associations among variables and blue edges indicate negative associations.

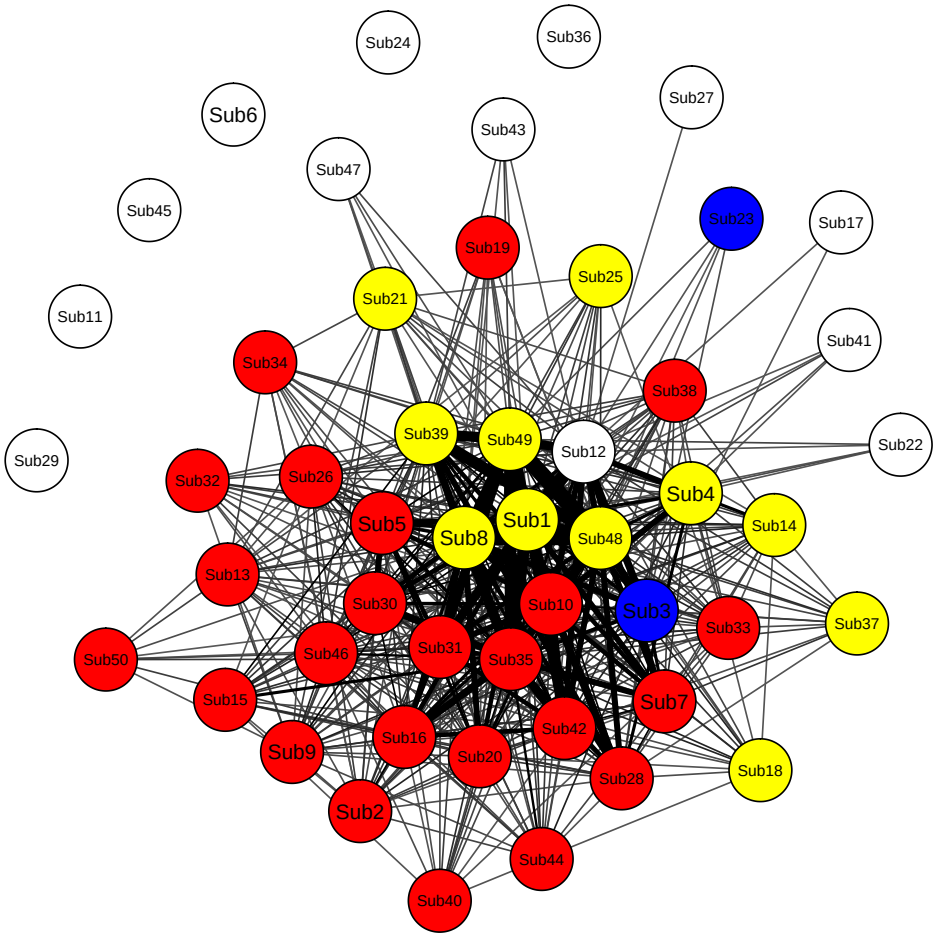

Figure 6. Subgroup membership from the scGVAR applied to the whole sample of controls and depressed subjects. Communities did not align with clinical diagnoses.

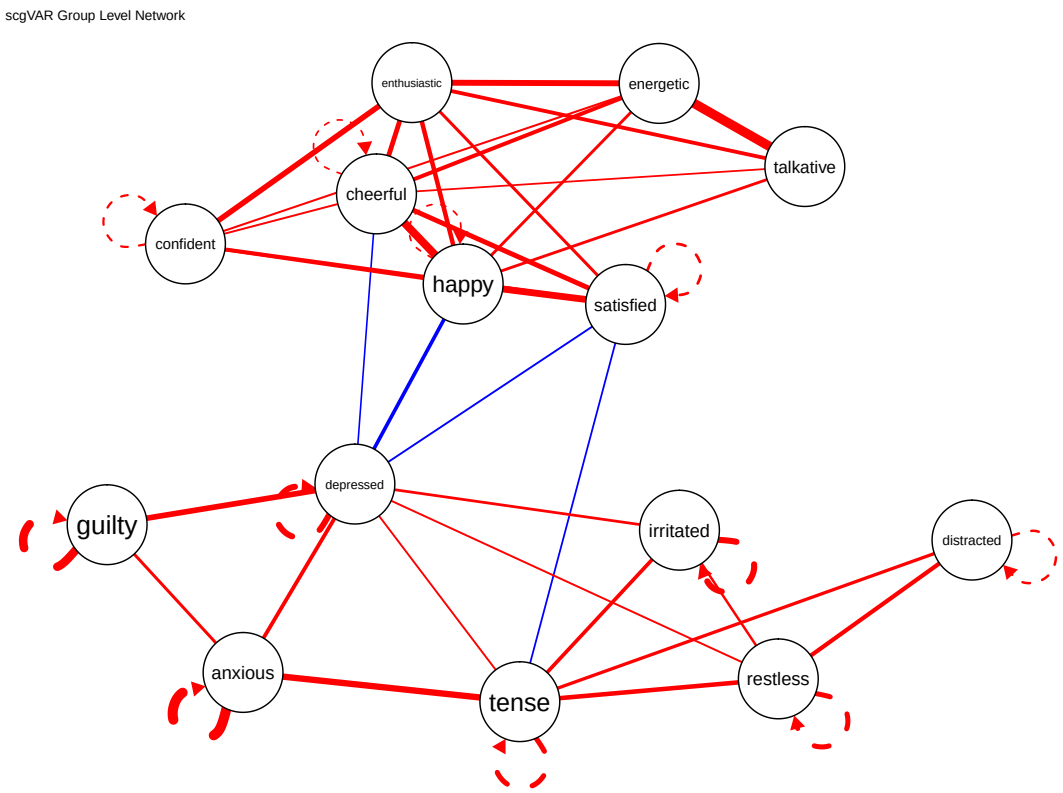

Figure 7. Group-level network of the scGVAR. Red edges indicate positive connections and blue indicate negative connections. Solid lines are indicative of contemporaneous effects and dashed lines are indicative of lagged effects.

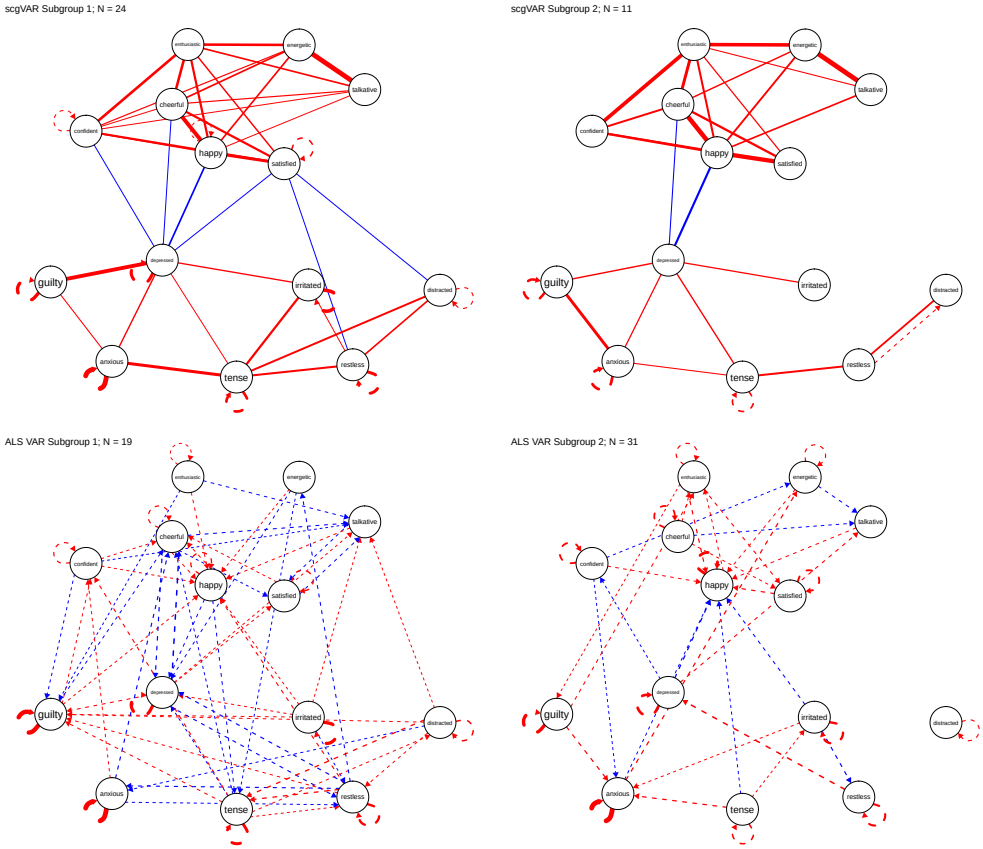

Figure 8. Recovered subgroups by method with scGVAR on top and the ALS VAR on the bottom row. Subgroups comprised of N > 2 participants shown. Red edges indicate positive associations and blue edges are negative associations. Dashed lines are lagged effects whereas bold lines are contemporaneous effects.