Influence of Parameter Choice on the Detection of High-Dimensional Functional Networks

Britta Pester, Karl-Jürgen Bär, and Lutz Leistritz

Abstract

The detection of directed interactions within networks derived from spatially highly resolved data, such as functional magnetic resonance imaging (fMRI) has been a challenging task for the last years. Commonly this is solved by restricting the analysis to a small number of representative network nodes (e.g. fMRI voxels), to regions of interest (e.g. brain areas) or by using dimension reduction methods like principal or independent component analysis. Recently, these problems have successfully been encountered by combining multivariate autoregressive models and parallel factor analysis. This approach involves a cascade of analysis steps, entailing a number of parameters that have to be chosen carefully. Yet, the question of an appropriate choice of analysis parameters has not been clarified so far—in particular for temporally varying models. In this work we fill this gap. Synthetic data with known underlying ground truth structure are generated to evaluate the correctness of results in dependence on the parameter choice. Resting state fMRI data are used to assess the influence of the involved parameters in the clinical application. We found that model residuals offer a good means for determining appropriate filter algorithm parameters; the model order should be chosen according to two aspects: the well-established information criteria and the fit between Fourier and estimated spectra.

1 Introduction

The human brain is a complex network, exchanging an immense mass of information between remote neuronal areas. Therefore, a thorough investigation of brain processes does not only require the analysis of brain activity but also the consideration brain connectivity [1]. In other words, two regions being active at the same time do not necessarily transfer information between each other. For this reason many approaches have successfully developed in order to quantify the extent of connectivity between different brain regions. Prominent examples are dynamic causal modelling, transfer entropy, Granger causality, directed transfer function and partial directed coherence [1–3].

However, spatially high-resolved data lead to two problems: first, the number of spatial nodes (in this work: fMRI voxels) by far exceed the number of temporal samples (in this work: fMRI volumes). Second, from a practical point of view, the computational capacities are exhausted due to the high network dimensionality: in the fMRI case, the number of network nodes reaches ten thousands up to a hundreds of thousands; in addition, the number of possible connections quadratically rises with the number of network nodes. This makes any conventional connectivity analysis unfeasible. In most cases, the analysis is limited to time series derived from a smaller set of selected or aggregated voxels. Another alternative is the application of dimension reduction methodologies as for example independent or principal component analysis (PCA) [4, 5].

A new approach has been proposed in [6]. Here, a PCA dimensionality reduction from high-dimensional (HD) into low-dimensional (LD) space is combined with a following multivariate autoregressive (MVAR) estimation which is

B. Pester (✉) · L. Leistritz
Institute of Medical Statistics, Computer Sciences and Documentation, Jena, University Hospital, Friedrich Schiller University Jena, Jena, Germany
e-mail: Britta.Pester@med.uni-jena.de
URL: http://www.imsid.uniklinikum-jena.de/IMSID.html

B. Pester
Chair of Clinical Psychology and Behavioural Neuroscience, Technische Universität Dresden, Dresden, Germany

K.-J. Bär
Department of Psychiatry, Jena University Hospital, Friedrich Schiller University Jena, Jena, Germany

transferred back into HD space. This finally enables the calculation of a highly resolved network, i.e. the quantification of directed connectivity from voxel to voxel.

What has been missing so far is an in-depth consideration of the necessary analysis parameters. The proposed methodology requires many analysis configurations, such as settings of the estimation algorithm or the proportion of retained variance after the PCA dimension reduction. Here, we successively vary the involved parameters and show the influence and reciprocal effects of parameter choice on the quality of network identification.

2 Material

In this work, we followed two complementary approaches: first, simulated data with known ground truth structure were generated in order to evaluate the correctness of results in dependence on the parameter choice. Second, resting state fMRI data of 154 healthy subjects were used to assess the influence of the involved parameters in a clinical application.

2.1 Synthetic Data

Simulated time series were realized as time-variant MVAR processes, where the model coefficients were chosen according to pre-defined ground truth networks. These networks were designed in such a way that the network nodes form four non-overlapping clusters, so-called modules [7]. This means that the expected value for an intra-module connection is considerably higher than that for an extra-module connection. The number of network nodes was set to $D = 50$ and the number of temporal samples to $N = 1000$, providing a good balance between network size and temporal resolution [8]. At sample $n = 500$, ground truth changed from one network into another, which enables the generation of time series based on a temporally varying model.

2.2 Resting State fMRI Data

To evaluate the influence of analysis parameters in practice, data from a resting state fMRI experiment conducted by the Department of Psychiatry and Psychotherapy, Jena University Hospital were used [9]. Pseudonymized data of 154 subjects were acquired using the 12 channel head coil at the 3T MRI scanner (MAGNETOM TIM Trio, Siemens). The experiment included a resting state fMRI scan with a subsequent high-resolution, anatomical T1-weighted MR scan. A to-

tal of $N = 240$ volumes were acquired; each consisting of 45 transversal slices covering the whole brain, deliberately including the lower brainstem [9].

3 Methods

3.1 Applied Analysis Steps

The herein applied methodologies are based on time-variant multivariate autoregressive models (tvMVAR) [10]. This tvMVAR approach has been further developed to the *large scale* MVAR model (lsMVAR) that can be used to estimate time-variant approximations of high-dimensional data [8]. Despite the benefit of time variance, this approach offers the possibility to apply any tvMVAR-based connectivity measure in high dimensions, including frequency-selective approaches.

The initial step of the lsMVAR approach is a reduction from HD space comprising D (D large) network nodes to LD space with C (C small) network nodes by means of PCA. Let $\mathbf{x} \in \mathbb{R}^{C \times N}$ be the LD matrix containing the C retained principle components of N temporal samples derived from HD data. Then, consider the LD tvMVAR model of order p for $\mathbf{x}$:

$$\mathrm{x}(n) = \sum_{r=1}^{p} \mathbf{B}^r(n)\mathrm{x}(n-r) + \mathrm{e}(n), \quad n = p+1, \ldots, N, \quad (1)$$

with LD model parameters $\mathbf{B}^r \in \mathbb{R}^{C \times C}$ and LD model residuals $\mathrm{e}(n) \in \mathbb{R}^C$. Then, the whole model can be projected back onto D-dimensional space by a left multiplication of the pseudoinverse of the (truncated) mixing matrix $\mathbf{W} \in \mathbb{R}^{C \times D}$:

$$\mathbf{W}^+ \mathrm{x}(n) = \mathbf{W}^+ \left(\sum_{r=1}^{p} \mathbf{B}^r \mathrm{x}(n-r) + \mathrm{e}(n) \right) \qquad (2)$$

which can be rearranged to

$$\underbrace{\mathbf{W}^+ \mathrm{x}(n)}_{:= \tilde{\mathrm{y}}(n)} = \sum_{r=1}^{p} \underbrace{\mathbf{W}^+ \mathbf{B}^r \mathbf{W}}_{:= \mathbf{A}^r} \underbrace{\mathbf{W}^+ \mathrm{x}(n-r)}_{:= \tilde{\mathrm{y}}(n-1)} + \underbrace{\mathbf{W}^+ \mathrm{e}(n)}_{:= \tilde{\mathrm{e}}(n)} \in \mathbb{R}^D,$$

$$(3)$$

with approximated HD data $\tilde{\mathrm{y}}(n)$, HD model parameters $\mathbf{A}^r$ and HD residuals $\tilde{\mathrm{e}}(n)$. This offers the opportunity for the estimation of *time-variant* MVAR models. In this work, connectivity was assessed by means of time-variant partial directed coherence (PDC) which has the benefit that directed connectivity can be quantified under consideration of various frequencies [1].

3.2 Involved Parameters

The lsMVAR approach requires four parameters that are involved in three different analysis steps:

- TvMVAR parameters were estimated by means of the Kalman Filter [11]. This time-variant approach requires two constants: c_1 regulates the adaption of the covariance matrix; c_2 defines the step-width of the random walk that is used to update the tvMVAR parameters.
- The tvMVAR model p has to be determined. This parameter defines the number of temporal samples in the past that are considered for the estimation of the current value.
- PCA dimension reduction demands an a priori definition of the number of retained PCA components C. This value determines the proportion of variance explanation after PCA, i.e. the higher C, the higher the explanation of variance.

4 Results

4.1 Synthetic Data

Simulations offer the possibility to clearly decide whether the results are correct or not. To assess the goodness of fit between ground truth and computed networks, we used the Cohen's kappa [12]. It quantifies the agreement between two raters; in this case between the derived networks and the known ground truth networks. The results of our simulations can be summarized as follows:

- A quite reasonable possibility for choosing the *Kalman filter parameters* c_1, c_2 is to consider the tvMVAR model residuals. In our simulations, this approach has proven to be useful: synthetic data showed that a high Kappa coefficient—and thus a high accordance between GTNs and PDC networks—corresponds to low mean squared model residuals. Therefore, surveying the model residuals offers a suitable possibility for an adequate choice of c_1, c_2.
- The determination of the *tvMVAR model order* p has proven to be not that clear. Conventional information criteria like Akaike's and Bayesian information criterion [10] provide a first recommendation by establishing a balance between goodness of fit and number of parameters that have to be estimated. However, for frequency-selective approaches like PDC it is important whether the model order is suitable to properly reproduce the frequency spectrum of original data. We found that

the best way is to initially use the information criteria to obtain a first impression of a reasonable region for the choice of p; then, Fourier spectra of real time series should be checked against those of estimated MVAR-based data.

- The successive variation of the *number of retained components* C did not show surprising results for simulated data. Figure 1a shows the performance in dependence on C by means of the area under the receiver operating characteristic curve [13]; clearly, the accordance of PDC networks with GTNs rises with increasing C. Cohen's kappa in dependence on the explained variance is represented in Fig. 1b; again, a higher explanation of variance leads to a deteriorated agreement between GTNs and lsMVAR-driven networks.

4.2 FMRI Data

Individual fMRI connectivity patterns heavily differed between subjects. However, it turned out that in despite of this variation, the influence of parameter choice was similar for the whole group. Therefore, we show the results for one exemplary subject.

First, all parameters were chosen according to the suggestions described in Sect. 4.1, then they were kept fix and successively one parameter has been systematically varied.

- The variation of *Kalman filter parameters* c_1, c_2 showed that the model residuals slightly rise with increasing c_1 while they intensively decrease with increasing c_2. That means: a faster adaption of the covariance matrix and a lower the step width of the random walk lead to a better model fit. As a consequence, it does not appear to be useful to solely consider the model residuals but also whether the adaption of the estimator is satisfying, which of course requires a certain experience of the user in the application of the method.
- According to Akaike's and Bayesian information criterion the *model order* was suggested to be set to $p = 8$. As described in Sect. 4.1, it is important to also consider the spectral properties; we found that for our data, $p \geq 11$ should be preferred with the aim to adequately separate connectivity patterns regarding the frequency domain. The reason is that for $p < 11$, time-frequency-maps of PDC are quite smeared, getting clearer with increasing p, while for $p > 11$, there is hardly any further improvement regarding this point. As an example, Fig. 2a shows the connection from the locus coeruleus complex (LC) to the nucleus raphes dorsalis (DRN), which have proven to be connected during resting state situations [9]. In our
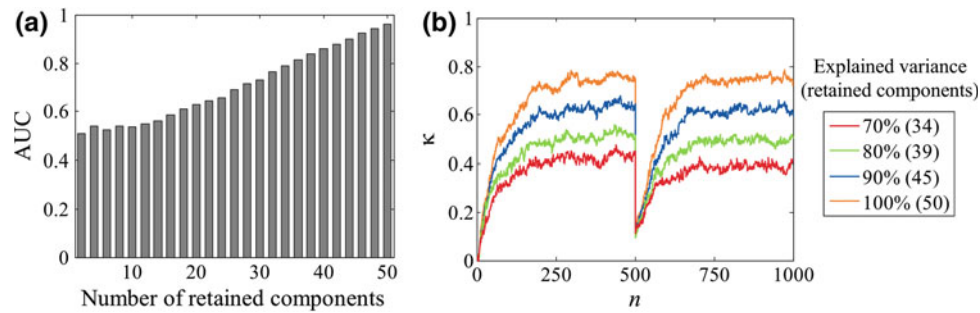
Fig. 1 Performance of lsMVAR-based PDC analysis. Panel **a** shows the temporal mean of the AUC values in dependence on the number of retained components $C = 1, \ldots, 50$. In panel **b**, the temporal dynamics of Cohen's kappa for 70, 80, 90 and 100% explained variance are depicted
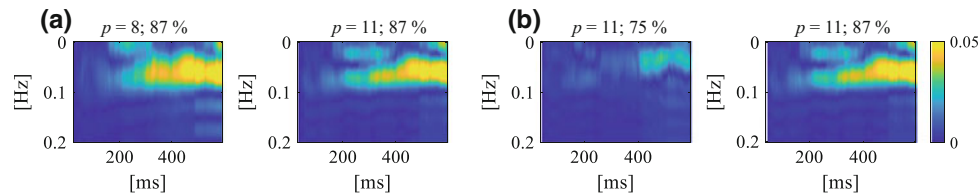

Fig. 2 PDC results of the connection from LC to DRN. Subplot **a** provides a comparison between two different model orders $p = 8$ (suggested choice based on information criteria) and $p = 11$

(data-driven optimum). Analogously, subplot **b** shows the PDC results for two different proportions of explained variance, 75 and 87%

time-variant, frequency-selective analysis approach, the order $p = 8$ suggested by the information criteria is not enough to separate the connection in low frequencies (around 0.06 Hz) emerging during the second half (Fig. 2a, left panel) in a clear manner as compared to $p = 11$ (Fig. 2a, right panel).

- A similar situation is when the number of *retained components* C has to be chosen. Whenever C is increased, the model gets more accurate; on the other hand it has to be considered that a high number of components leads to high computational efforts. Therefore, a good strategy is to inspect the results in dependence on the explained variance and identify a proper balance between explained variance and adequate computational efforts. For our data, we found that an explained variance of around 87% provides a good compromise. Again, the rationale is that this choice represents the point, where for higher values of C the detected networks hardly vary, while for smaller C the derived networks immensely differ when C is changed. This property is by far more pronounced for the choice of C as compared to the choice of p. Figure 2b demonstrates this property: analogous to Fig. 2a, it illustrates the PDC time-frequency maps of the connection from LC to DRN. In the left panel, the map for 75% variance explanation is shown and on the left for 87%. The most striking difference occurs in the lower frequency domain: for 75%, high connections are indicated around 0.03 Hz,

while for 87% it is around 0.06 Hz. Notably, this 0.06 Hz connection remains for higher variance explanation than 87%, this is why that point provides a suitable indicator for a proper choice of C.

5 Discussion and Future Work

Any newly introduced method requires a substantial evaluation to justify the application to real-world data. Nonetheless, in addition it is important to test and understand the influence and mutual effects of analysis configurations in order to avoid misinterpretations due to inappropriate parameter settings. For conventional PDC there has been in-depth work on that aspect based on simulations and EEG data, providing recommendations for the application of this method [14].

However, for the recently proposed lsMVAR approach this point has not been investigated yet. The lsMVAR methodology combines a PCA dimension reduction with tvMVAR modelling and involves four important parameters: two parameters that control the adaption of the estimation algorithm; the tvMVAR model order, defining the number steps in the past that are included for the estimation of the current value; and the number of retained PCA components which corresponds to the proportion of explained variance.

Based on the analysis of synthetic data, we found that model residuals yield a useful indication for a suitable setting of the Kalman filter control parameters. A combination between common information criteria and the consideration of frequency spectra give support in choosing an appropriate tvMVAR model order p. Not surprisingly, a higher number of retained components C leads to a better agreement between GTNs and PDC networks.

For the resting state fMRI data, we found that the choice of Kalman filter parameters based on model residuals is not advisable. Besides surveying the residuals, a sufficient expertise is necessary to find an appropriate compromise between fast adaption and smoothness of the estimated model. The model order p should be chosen in two steps: first, information criteria should be applied to get an appropriate initial value. Second, the results in this range should be inspected with regard to the fit between Fourier and estimated spectra, in order to find out whether this value is adequate. Finally we found that the most impact on the results was made by the explained variance after PCA (i.e. number of retained components C). Similar to p, C should successively be varied to identify the setting when the results of higher C only differ to a small extent.

So far, we inspected the results from a methodological point of view. After finding an appropriate parameter choice, the next step will be to investigate the results in addition to methodological questions—can the lsMVAR approach provide new insights into the default mode network [15]? Furthermore, what has not been done yet is to take advantage of the possibility to explore temporally varying experimental setups [16]. Finally, a comparison between groups is of great interest [17], which however will be a big challenge due to the high number of output data.

References

1. K. Sameshima and L. A. Baccalá, *Methods in brain connectivity inference through multivariate time series analysis*. 1em plus 0.5em minus 0.4em CRC press, 2014.
2. L. Lee, K. Friston, and B. Horwitz, "Large-scale neural models and dynamic causal modelling," *Neuroimage*, vol. 30, no. 4, pp. 1243–1254, 2006.
3. A. K. Seth, "Causal connectivity of evolved neural networks during behavior," *Network-Computation in Neural Systems*, vol. 16, no. 1, pp. 35–54, 2005. [Online]. Available: <Go to ISI > :// WOS:000233575000003 = 0pt.
4. S. M. Smith, A. Hyvärinen, G. Varoquaux, K. L. Miller, and C. F. Beckmann, "Group-pca for very large fMRI datasets," *NeuroImage*, vol. 101, pp. 738–749, 2014.
5. I. Jolliffe, *Principal component analysis*. 1em plus 0.5em minus 0.4em Wiley Online Library, 2002.
6. C. Schmidt, B. Pester, N. Schmid-Hertel, H. Witte, A. Wismüller, and L. Leistritz, "A multivariate granger causality concept towards full brain functional connectivity," *PloS one*, vol. 11, no. 4, p. e0153105, 2016.
7. V. D. Blondel, J. L. Guillaume, R. Lambiotte, and E. Lefebvre, "Fast unfolding of communities in large networks," *Journal of Statistical Mechanics-Theory and Experiment*, 2008.
8. B. Pester, "Novel approaches for exploring highly resolved brain connectivity: development, evaluation and practical application," Thesis, Institute of Medical Statistics, Computer Sciences and Documentation, 2016.
9. K.-J. Baer, F. de la Cruz, A. Schumann, S. Koehler, H. Sauer, H. Critchley, and G. Wagner, "Functional connectivity and network analysis of midbrain and brainstem nuclei," *Neuroimage*, vol. 134, pp. 53–63, 2016.
10. H. Lütkepohl, *New introduction to multiple time series analysis*. 1em plus 0.5em minus 0.4em Springer Science & Business Media, 2005.
11. T. Milde, L. Leistritz, L. Astolfi, W. H. Miltner, T. Weiss, F. Babiloni, and H. Witte, "A new kalman filter approach for the estimation of high-dimensional time-variant multivariate ar models and its application in analysis of laser-evoked brain potentials," *Neuroimage*, vol. 50, no. 3, pp. 960–969, 2010.
12. J. Cohen, "A coefficient of agreement for nominal scales," *Educational and psychological measurement*, vol. 20, no. 1, pp. 37–46, 1960.
13. L. Sachs, *Angewandte Statistik: Anwendung statistischer Methoden*. 1em plus 0.5em minus 0.4em Springer-Verlag, 2013.
14. L. Leistritz, B. Pester, A. Doering, K. Schiecke, F. Babiloni, L. Astolfi, and H. Witte, "Time-variant partial directed coherence for analysing connectivity: a methodological study," *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, vol. 371, no. 1997, p. 20110616, 2013.
15. R. Hindriks, M. H. Adhikari, Y. Murayama, M. Ganzetti, D. Mantini, N. K. Logothetis, and G. Deco, "Can sliding-window correlations reveal dynamic functional connectivity in resting-state fMRI?" *Neuroimage*, vol. 127, pp. 242–256, 2016.
16. J. Liu, B. A. Duffy, D. Bernal-Casas, Z. Fang, and J. H. Lee, "Comparison of fMRI analysis methods for heterogeneous bold responses in block design studies," *NeuroImage*, vol. 147, pp. 390–408, 2017.
17. L. Lim, H. Hart, M. A. Mehta, A. Simmons, K. Mirza, and K. Rubia, "Neural correlates of error processing in young people with a history of severe childhood abuse: an fMRI study," *American Journal of Psychiatry*, vol. 172, no. 9, pp. 892–900, 2015.