Latent Causal Modelling of Neuroimaging Data

Morten Mørup[1], **Kristoffer Hougaard Madsen**[1,2], **Lars Kai Hansen**[1] {`mm,khm,lkh`}`@imm.dtu.dk`
1: Informatics and Mathematical Modeling, Technical University of Denmark,
2: Danish Research Centre for Magnetic Resonance

Introduction: To infer the causal relations among sets of measurement variables in neuroimaging data Granger causality [4] and the equivalent directed transfer function (DTF) approach [5] have been invoked. Granger causality is a fundamental tool for the empirical investigation of dynamic interactions in multivariate time-series and causality is here based on the common sense conception that causes always precede their effects. As such, an event taking place in the future cannot cause another event in the past or present. The causal relations are derived by fitting the following autoregressive model to the space-time data $\boldsymbol{X}^{I\times T}$ [5]

$$x_i(t) = \sum_{\tau=1}^{\Upsilon} \sum_{j=1}^{I} q_{i,j}(\tau) x_j(t-\tau) + e_i(t).$$

In the frequency domain these systems of equations correspond to [5]

$$\boldsymbol{x}(f) = \boldsymbol{Q}(f)\boldsymbol{x}(f) + \boldsymbol{e}(f) \Rightarrow \boldsymbol{x}(f) = (\boldsymbol{I} - \boldsymbol{Q}(f))^{-1}\boldsymbol{e}(f) = \boldsymbol{H}(f)\boldsymbol{e}(f)$$

where $\boldsymbol{H}(f) = (\boldsymbol{I} - \boldsymbol{Q}(f))^{-1}$ denotes the transfer function such that $h_{i,j}(f)$ denotes the influence of channel/voxel j to channel/voxel i at frequency f. In the time domain this can equivalently be written as

$$x_i(t) = \sum_{\tau=1}^{T} \sum_{j=1}^{I} h_{i,j}(\tau) e_j(t+1-\tau)$$

where $e_j(t)$ is the input signal to channel j and $h_{i,j}(\tau)$ the influence of channel j to channel i at delay τ.

Methods: Instead of operating with channel specific input functions $\boldsymbol{e}(t)$ derived from the autoregressive model we will operate with latent input functions $\boldsymbol{s}(t)$ based on the following convolutional representation

$$x_i(t) = \sum_{\tau=1}^{\Upsilon} \sum_{d=1}^{D} a_{i,d}(\tau) s_d(t+1-\tau) + \varepsilon_i(t).$$

Where $\varepsilon_i(t)$ is residual noise. The above bilinear convolutive representation assume that the data are caused by underlying latent sources rather than causes stemming from the measurement channels themselves. This we believe to hold for neuroimaging data in general as the voxels/channels record the manifestation of underlying latent neural activity hence it is these underlying sources that drive the measurements rather than voxel/channel specific inputs. We note that this representation corresponds to the convolutive ICA model [8] which was used in [3] for the analysis of spatio-temporal dynamics in EEG. Comparing the above convolutive bilinear decomposition with Granger causal modelling it can be seen that $a_{i,d}(\tau)$ and the transfer tensor $h_{i,j}(\tau)$ in the Granger causality analysis are closely related. The difference being that the transfer tensor $h_{i,j}(\tau)$ is based on channel specific input signals $\boldsymbol{e}(t)$ whereas the transfer tensor in the above convolutive bilinear model is based on latent input signal $\boldsymbol{s}(t)$. Hence, causal relations from latent variables to the measuring channels can be inferred directly from the above convolutive bilinear representation and as such the representation above can potentially perform dimensionality reduction resulting in fewer latent sources than observed measurement voxels/channels. Furthermore, constraints on the causal relations can be directly imposed on $\boldsymbol{A}(\tau)$ such as sparsity or restricting the transfer function to specific delays. These two types of constraints can improve interpretation of the causal relation between the latent sources and the measured voxels/channels. Finally, spatial regions that are caused by the source $s_d(t)$ are automatically grouped in $\boldsymbol{a}_d(\tau)$ while the propagation over time of the source to the voxels/channels is encoded in the temporal mode. As such, the latent modeling above naturally group instantaneous effects in $\boldsymbol{A}(\tau)$ whereas Granger causality is hard to interpret in case of instantaneous propagation between voxels/channels [4] .

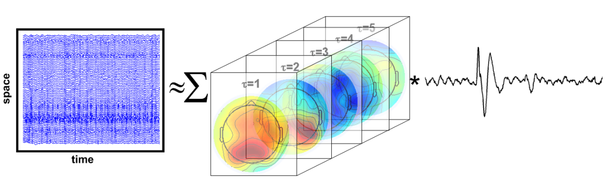

Figure 1: Illustration of the convolutive bilinear model. Each extracted temporal signature has a channel specific delay given by $\boldsymbol{A}(\tau)$. The representation is closely related to Granger causal modelling but contrary to channel specific temporal input functions $\boldsymbol{e}(t)$ the model operates with latent input functions $\boldsymbol{s}(t)$.

Rather than imposing statistical independence on the sources $\boldsymbol{S}$ (as in convolutive ICA) to resolve the model ambiguity of the bilinear convolutive representation we will impose sparsity on the convolutive filter coefficients $\boldsymbol{A}(\tau)$ by regularizing the filter coefficients by the sparsity promoting l_1-norm [2]. We formulate the above model probabilistically and use Bayesian learning for inference. Particularly, we infer the optimal model order D as well as degree of sparsity on $\boldsymbol{A}(\tau)$ using a hierarchical Bayesian approach (automatic relevance determination) [1]. In [7] it has been demonstrated that automatic relevance determination is well suited for tuning the l_1-regularization strength as well as establishing model order. The considered model and corresponding log posterior is given by

$$\begin{array}{lcl} \varepsilon_i(t) & \sim & Normal(0,\sigma^2) \\ \sigma^{-2} & \sim & Gamma(1,\kappa\|\boldsymbol{X}\|_F^2) \\ \boldsymbol{a}_d(\tau) & \sim & Laplace(0,\beta_d) \\ \beta_d & \sim & Gamma(1,\alpha) \\ s_d(t) & \sim & \delta(1-\sum_t s_d(t)^2) \end{array} \qquad \log P(\boldsymbol{X},\mathcal{A},\boldsymbol{S},\sigma^{-2},\boldsymbol{\beta}|\kappa,\alpha) = \left\{ \begin{array}{l} -\frac{\sigma^{-2}}{2}\sum_t^T \|\boldsymbol{x}(t) - \sum_\tau \boldsymbol{A}(\tau)\boldsymbol{s}(t+1-\tau)\|_F^2 \\ -\frac{1}{2}IT\log(\sigma^{-2}) - \kappa\|\boldsymbol{X}\|_F^2\sigma^{-2} \\ +I\Upsilon\log\beta_d - \beta_d(\alpha + \sum_i^I\sum_\tau^\Upsilon |a_{i,d}(\tau)|) \\ +const. \\ \quad s.t. \quad \sum_t s_d(t)^2 = 1 \end{array} \right.$$

We used maximum a posteriori estimation to infer all model parameters (we set $\kappa = 1/10$, $\alpha = 10^{-9}$). In the following we will refer to the above analysis as sparse latent causal modelling (SLCM).

Results: In Figure 2 we analyze a synthetic and real EEG data set. The synthetic EEG data consists of four components. The first two components were based on latent sources observed in multiple channels while the two last components were based on channel specific input functions (as assumed in the Granger DTF approach). The real EEG data set is based on the evoked potential of visual stimulations (for further details see [6]). While the SLCM recover the four components of the synthetically generated data the DTF approach [5] correctly identifies the two channel specific components (3 and 4) but redundantly model the first two components as stemming arbitrarily from some of the multiple channels observing these latent sources. For the real EEG data the SLCM extracts a four component model where the most prominent first component corresponds to visual activation in line with the stimuli. From the extracted transfer tensor $\boldsymbol{A}(\tau)$ the temporal propagation of the sources through the measured channels can be observed.

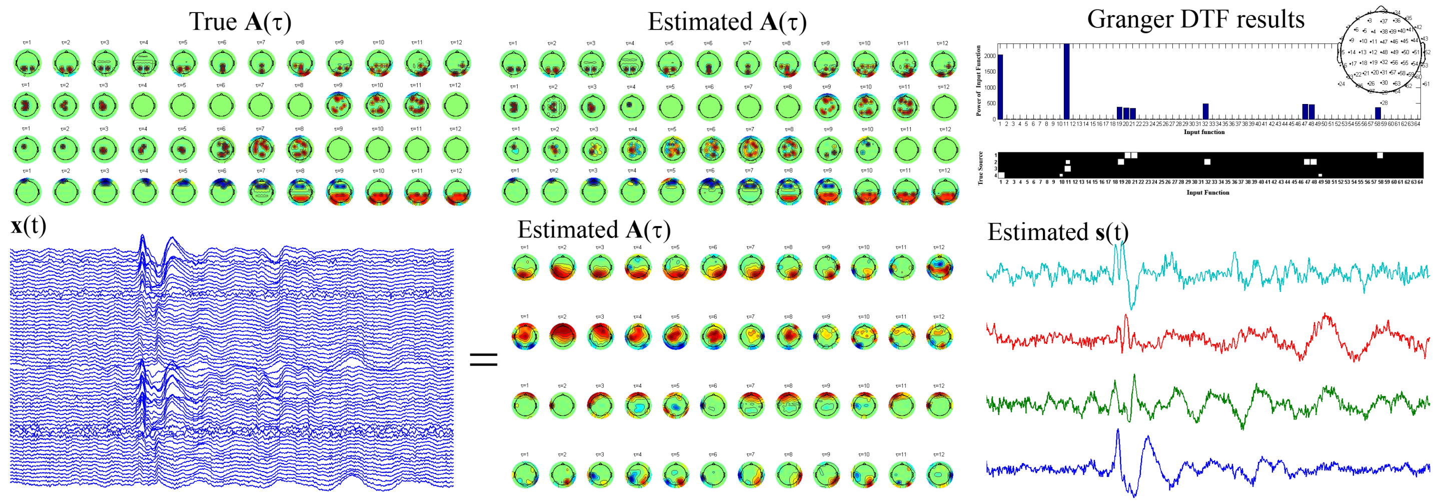

Figure 2: **Top panel:** Analysis of 64 channel synthetic EEG data using the proposed SLCM as well as the Granger DTF approach described in [5]. The data is generated such that the two first components are based on latent sources observed simultaneously in several channels whereas the last two components are generated such that the two latent sources initially are observed in a single channel (channel 11 and 1). As such these two channels Granger causes the remaining channels of the two components. The true $\boldsymbol{A}(\tau)$ show how the simulated sources over time cause the observed activity in the various channels. Next to the true underlying simulated dynamics are given the estimated dynamics from the SLCM model. The model has correctly extracted a four component model and the dynamics are well in accordance with the true underlying dynamics, as such the autocorrelation between the estimated and true latent sources $\boldsymbol{S}^{est}$ and $\boldsymbol{S}^{true}$ are significantly correlated (at an $\alpha = 1\%$ level) only between sources with same component indices (not shown). To the right is given the corresponding Granger analysis based on the DTF approach [5]. From the power of the estimated input functions $\boldsymbol{e}(t)$ of each channel it can be seen that 9 of the 64 input functions are active. The most active input functions are the input functions of channel 11 and 1 corresponding to the two channels of component 3 and 4 that were generated to Granger cause the remaining channels that are activated in the two components respectively. Below are given the significant maximal autocorrelations (on an $\alpha = 1\%$ level) between the input functions and true generated latent sources. Here the estimated inputs functions to channel 11 and 1 are (correctly) significantly correlated to the latent source 3 and 4 respectively. Notice however, that some of the information of component 4 also is significantly coded in the input functions of channel 10 and 49. As the Granger analysis is unable to correctly account for the dynamics of component 1 and 2 the information of these underlying two latent sources are arbitrarily distributed to several of the channels that observe these sources. **Bottom panel:** Analysis of a 64 channel real visual EEG data set. Based on the SLCM approach a four component model was extracted of which the most prominent component (component one) pertains to activation in the occipital areas of the brain whereas the second and third components pertain to frontal activation. The space-time dynamics of the first component reflect a flow from left to right occipital areas and later to more central regions while the frontal-central activation of the second component progress over time to the central areas of the scalp.

Discussion: We have established a close connection between Granger causality and convolutive bilinear models. Contrary to Granger causality that considers causal relations between the measurement variables the convolutive model operate with causal relations to underlying latent sources. We derived a Bayesian approach to estimate the model parameters and demonstrated its success on real and artificial EEG data. The proposed SLCM approach readily generalize to fMRI data where $I \gg T$, this will be the focus in future work.

References

[1] C. M. Bishop. *Pattern Recognition and Machine Learning (Information Science and Statistics).* Springer, 2006.

[2] D. Donoho. For most large underdetermined systems of linear equations the minimal l^1-norm solution is also the sparsest solution. *Communications on Pure and Applied Mathematics*, 59(6):797–829, 2006.

[3] M. Dyrholm, S. Makeig, and L. K. Hansen. Model selection for convolutive ica with an application to spatiotemporal analysis of eeg. *Neural Comput.*, 19(4):934–955, 2007.

[4] C. W. J. Granger. Investigating causal relations by econometric models and cross-spectral methods. *Econometrica*, 37(3):424–438, 1969.

[5] M. Kamiński, M. Ding, W. A. Truccolo, and S. L. Bressler. Evaluating causal relations in neural systems: granger causality, directed transfer function and statistical assessment of significance. *Biol Cybern*, 85(2):145–157, 2001.

[6] M. Mørup, L.K. Hansen, S.M. Arnfred, L.-H. Lim, and K.H. Madsen. Shift-invariant multilinear decomposition of neuroimaging data. *NeuroImage*, 42(4):1439–1450, 2008.

[7] M. Mørup and L. K. Hansen. Automatic relevance determination for multi-way models. *Journal of Chemometrics*, 23:352–363, 2009.

[8] M. S. Pedersen, J. Larsen, U. Kjems, and L. C. Parra. A survey of convolutive blind source separation methods. *Springer Handbook on Speech Processing and Speech Communication*, 2007.