統計数理研究所
The Institute of Statistical Mathematics

HBM 2006
170 T-PM

Decomposition of EEG Time Series into Independent Components by State-Space Modelling

ANDREAS GALKA[1,2,*], KEVIN WONG[3], TOHRU OZAKI[1,3]

[1] Institute of Statistical Mathematics (ISM), Minami-Azabu 4-6-7, Tokyo 106-8569, Japan
[2] Institute of Experimental and Applied Physics, University of Kiel, 24098 Kiel, Germany
[3] Department of Statistical Science, The Graduate University for Advanced Studies, Minami-Azabu 4-6-7, Tokyo 106-8569, Japan

[*] andreas@ism.ac.jp

1. Introduction

Contemporary research in neuroscience produces multi-channel time series representing various aspects of the activity of the human brain, such as EEG, MEG, fMRI and NIRS. For the quantitative analysis of such data new time series analysis tools are needed. A task of high relevance is to separate artefacts (such as the power supply signal or eye movements) from the data and furthermore to decompose the data itself into neurophysiologically meaningful component processes.
Numerous methods have been proposed for such decomposition tasks, many of which belong to the field of *Independent Component Analysis* (ICA) [1]. Here we present an alternative approach to the same task, based on classical stochastic autoregressive time series modelling [2].

2. Methodology

Let the data be given as

$$\mathbf{x}(t) \equiv \big(x_1(t), \ldots, x_N(t)\big)\ ,\ t = 1, \ldots, T$$

where N denotes the number of channels and T the number of sampling times. A *linear state space* (linSS) *model* for $\mathbf{x}(t)$ can be defined by an observation equation

$$\mathbf{x}(t) = \mathsf{C}\,\mathbf{s}(t) + \boldsymbol{\epsilon}(t)$$

where C denotes the *observation matrix*, $\mathbf{s}(t) \equiv \big(s_1(t), \ldots, s_M(t)\big)$ the *source components* (or *factors*), M the dimension of $\mathbf{s}(t)$ and $\boldsymbol{\epsilon}(t)$ measurement noise; and by a state dynamics equation

$$\mathbf{s}(t) = \mathsf{A}\,\mathbf{s}(t-1) + \boldsymbol{\eta}(t)$$

where A denotes the *state transition matrix* and $\boldsymbol{\eta}(t)$ dynamical noise. The covariance matrix of $\boldsymbol{\eta}(t)$ is denoted by $\boldsymbol{\Sigma}_\eta$.
The main model parameters A, C and $\boldsymbol{\Sigma}_\eta$ can be obtained by maximisation of the (log-)likelihood

$$\log \mathcal{L} = -\frac{1}{2}\left(T \log|\boldsymbol{\Sigma}_x| + \sum_{t=1}^{T} \big(\mathbf{x}(t) - \hat{\mathbf{x}}(t|t-1)\big)^{\dagger} \boldsymbol{\Sigma}_x^{-1} \big(\mathbf{x}(t) - \hat{\mathbf{x}}(t|t-1)\big)\right)$$

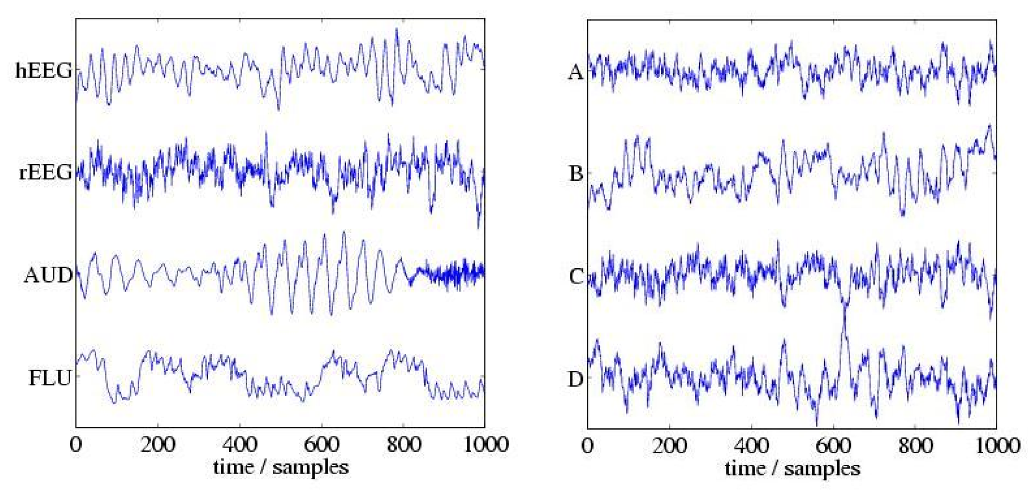
where $\hat{\mathbf{x}}(t|t-1)$ and $\boldsymbol{\Sigma}_x$ denote the observation prediction and the observation prediction covariance matrix, respectively, as provided by a Kalman Filter.
Initial values for A, C and $\boldsymbol{\Sigma}_\eta$ can be obtained by a *multivariate autoregressive* (MAR) *model* [2]

$$\mathbf{x}(t) = \sum_{\tau=1}^{p} \boldsymbol{\Phi}(\tau)\,\mathbf{x}(t-\tau) + \tilde{\boldsymbol{\eta}}(t)$$

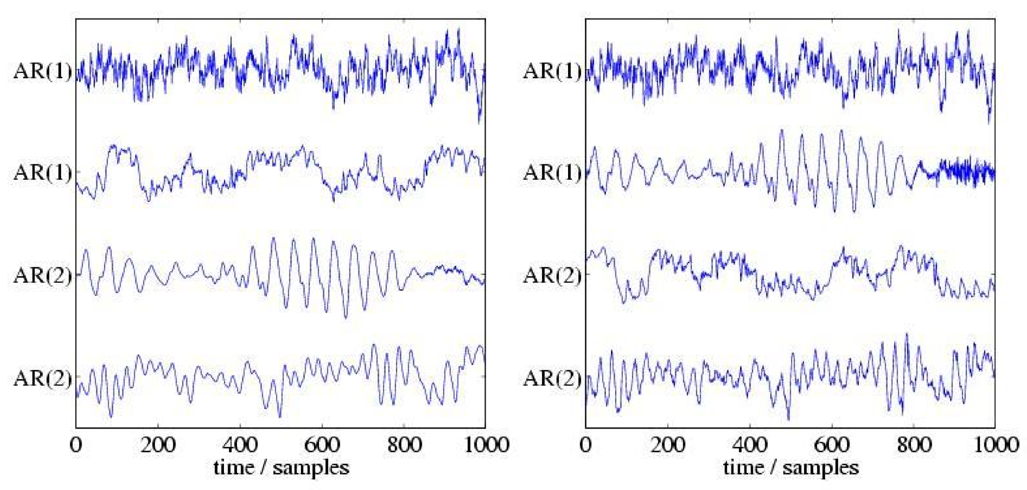
where $\boldsymbol{\Phi}(\tau)$, $\tau = 1, \ldots p$ denotes a set of autoregressive parameter matrices, p an integer model order and $\tilde{\boldsymbol{\eta}}(t)$ dynamical noise. By computing the set of characteristic roots of this model the state transition matrix A can be estimated, such that each real root and each pair of conjugated complex roots gives rise to one source component $s_i(t)$, $i = 1, \ldots, M$. These source components provide an alternative decomposition of the data, instead of the "independent components" of ICA. Even without additional maximisation of likelihood, MAR modelling provides useful decompositions of the data.

3. Simulation study

We demonstrate the feasibility of time series decomposition by MAR and linSS modelling by mixing $N = 4$ completely uncorrelated real-world signals (as "true" sources) by multiplication with a 4×4-matrix:

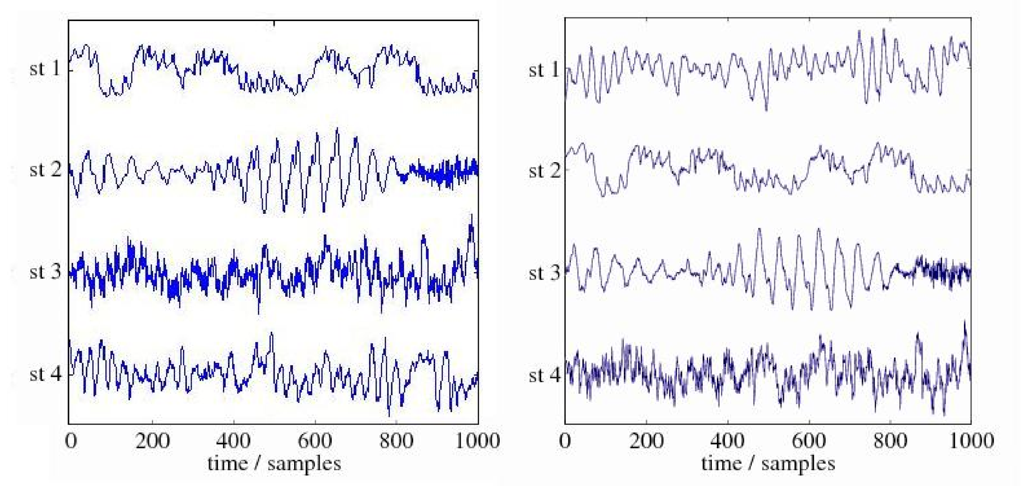


Left: true sources (hEEG = human EEG, rEEG = rat EEG, AUD = voice signal, FLU = turbulent fluid velocity signal); right: simulated data.



Left: decomposition by MAR modelling (model order $p = 7$, keeping only those 4 components with highest Fourier entropy); right: decomposition by linSS modelling, i.e. same components as left, improved by maximisation of likelihood.

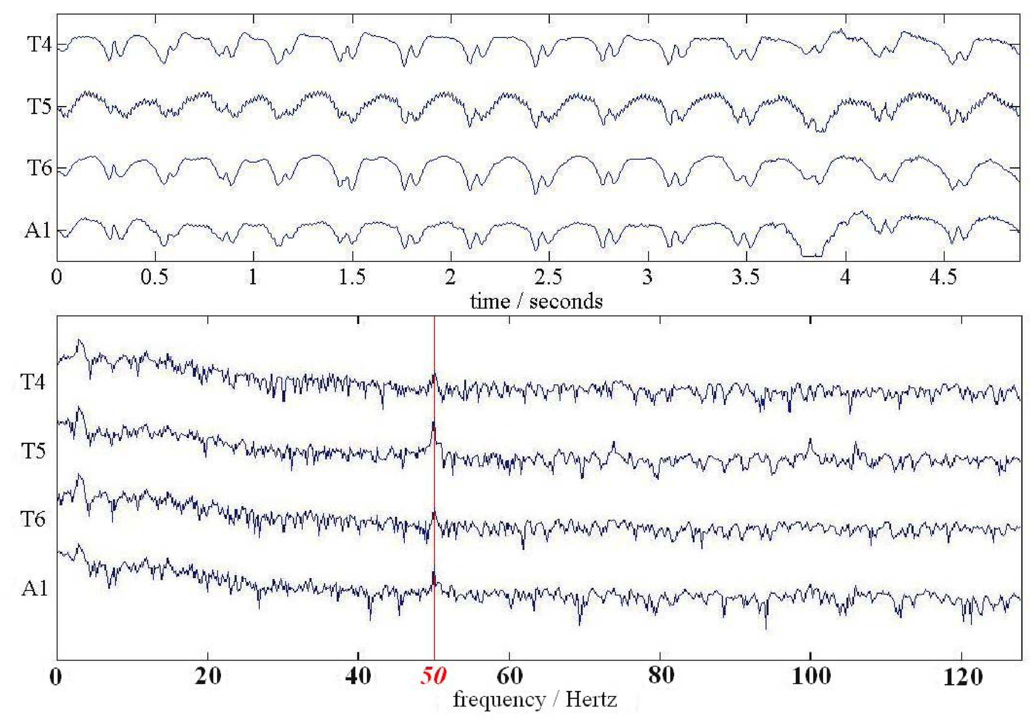
In this simulation also standard ICA algorithms provide very good decompositions:



Left: decomposition by the FastICA algorithm [1]; right: decomposition by the MILCA algorithm [3]. Note that the original ordering of components cannot be reproduced and that components may be inverted.

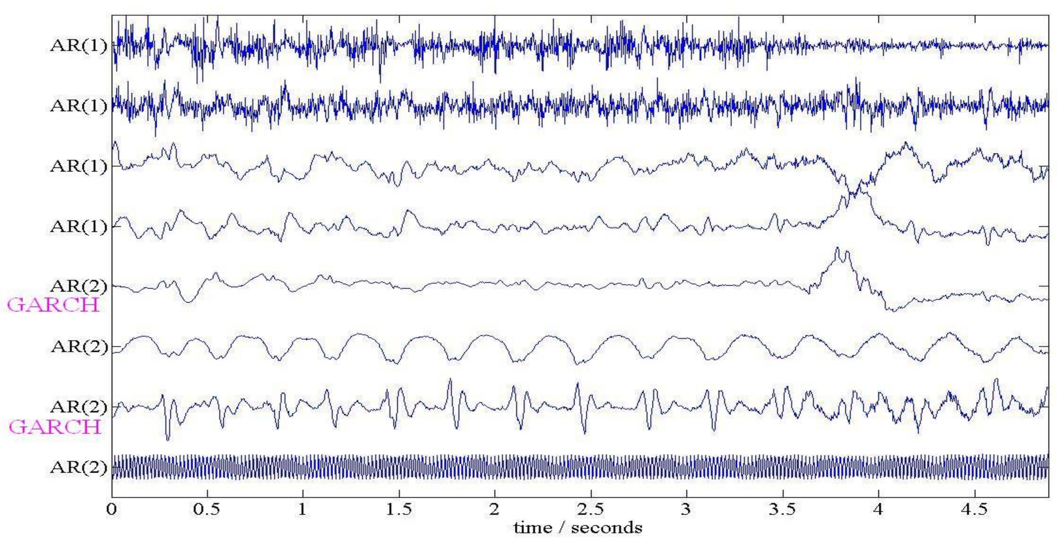
4. Analysis of real EEG data: absence seizure

First we show the example of an EEG time series of an epileptic absence seizure which is contaminated with 50 Hz power supply noise:



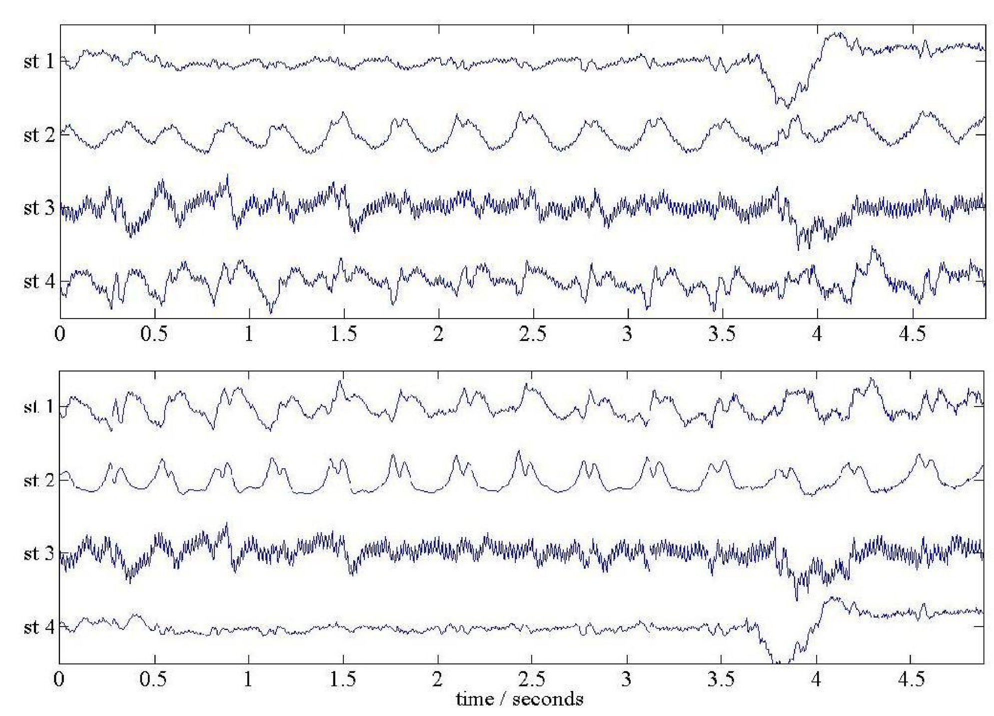
Top: seizure EEG data (sample rate 256 Hz); bottom: corresponding power spectrum (FFT; vertical axis is logarithmic).

The linSS decomposition (with GARCH modelling of covariance for two components, see Section 5) clearly identifies the power supply signal, a low-frequency artefact (presumably eye movement) and seizure-related components (the decomposition based on pure MAR looks similar):



Decomposition by linSS modelling (obtained from MAR with model order $p = 5$).

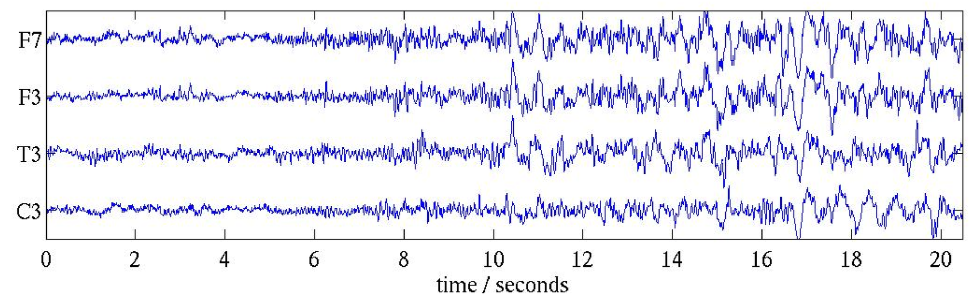
The standard ICA algorithms are found to be unable to provide the same clear distinction of components:



Top: decomposition by the FastICA algorithm [1]; bottom: decomposition by the MILCA algorithm [3].

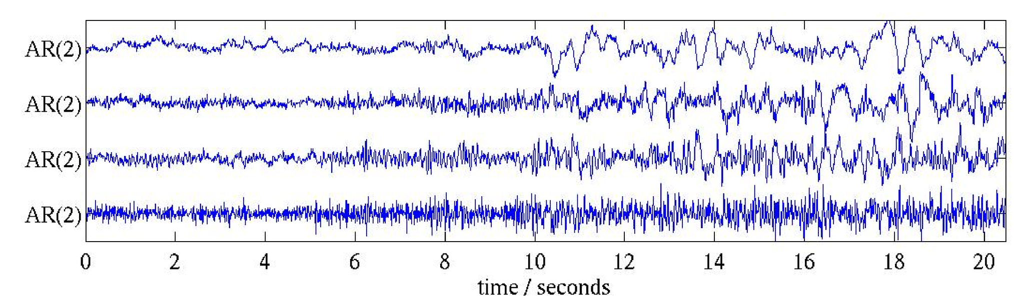
5. Analysis of real EEG data: transition to anaesthesia

Here we show $N = 4$ channels from an EEG time series from a patient experiencing the onset of clinical anaesthesia:



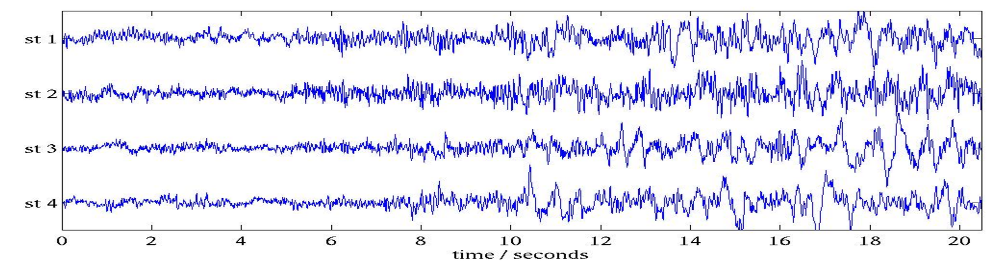
EEG data displaying onset of anaesthesia (sample rate 100 Hz).

By state space modelling we can decompose this data set into sources, some of which represent low-frequency components, representing the gradual loss of consciousness:



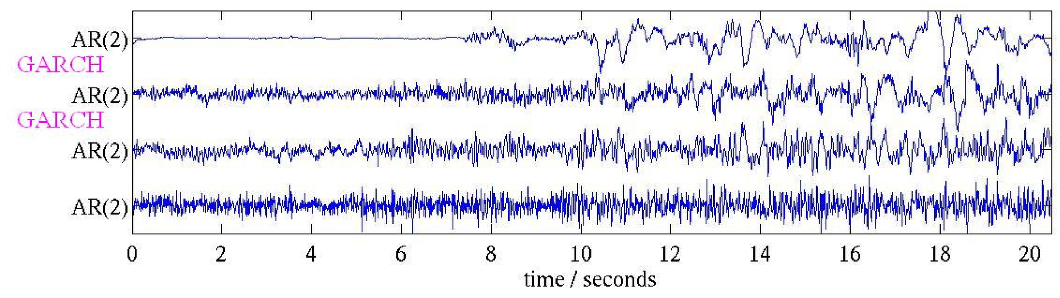
Decomposition by linSS modelling (based on a MAR model with model order $p = 16$, keeping $M = 4$ components). AIC for this model: 6302.21.*
**Akaike Information Criterion*

In contrast, standard ICA algorithms provide essentially uninformative decompositions, although they formally represent decompositions with minimum value of the sample mutual information [3]:



Decomposition by the MILCA algorithm [3].

The apparent non-stationarity of the dynamics can be modelled better by introducing dynamics of the covariances of the dynamical noise terms $\boldsymbol{\eta}(t)$, using a GARCH type model [4,5] (*see also poster* 244 T-PM):



Decomposition by linSS modelling with GARCH covariance dynamics for the first two components. Note that the first component, containing high-amplitude delta activity, remains inactive for the first 8 seconds, thereby roughly locating the time of loss of consciousness. AIC for this model: 5946.15.

6. Summary

- ⋆ ICA can be implemented, and generalised, by classical time series methods.
- ⋆ Likelihood (or, preferably, the *Akaike Information Criterion*, AIC) provides a measure for comparing the quality of competing models: linSS represents a more general model class than MAR, therefore it will achieve better AIC (i.e., better description of data).
- ⋆ Likelihood/AIC shows that by taking dynamics into account (which is not done in most ICA algorithms) the models can be greatly improved.
- ⋆ In simulations ICA algorithms and time series methods achieve about same performance.
- ⋆ With real data time series methods provide more meaningful components.
- ⋆ Time series methods can cope with $M > N$ ("overcomplete bases") and observation noise (hard for ICA methods).
- ⋆ Time series methods can abandon the independence constraint (i.e. model *correlated* components).
- ⋆ Full optimisation for linSS modelling is time consuming; MAR and some ICA methods provide fast alternatives.
- ⋆ The Kalman Filter is able to identify also nonstationary components; a further improvement of the model can be achieved by introducing dynamics of the covariances of the dynamical noise terms $\boldsymbol{\eta}(t)$ by GARCH-type models [4,5].

7. Acknowledgements

This work was supported by the Japanese Society for the Promotion of Science (JSPS) through grants KIBAN 13654075, KIBAN 15500193 and P 03059, and by the Deutsche Forschungsgemeinschaft (DFG) through project GA 673/1-1. EEG data were kindly provided by *U. Stephani* and *H. Muhle* from the Neuropediatric Hospital of the University of Kiel, Germany (epilepsy data), and by *E.R. John* and *L. Prichep* from New York University and Bellevue Hospital, New York, USA (anaesthesia data).

8. References

[1] *A. Hyvärinen, J. Karhunen* and *E. Oja* (2001). *Independent Component Analysis*. Wiley, New York.
[2] G.E.P. Box and G.M. Jenkins (1976). *Time series analysis, forecasting and control* (2nd ed.). Holden-Day, San Francisco.
[3] *H. Stögbauer, A. Kraskov, S.A. Astakhov* and *P. Grassberger* (2004). Least-dependent-component analysis based on mutual information. *Phys. Rev. E*, **70**, 066123.
[4] *A. Galka, O. Yamashita* and *T. Ozaki* (2004). GARCH modelling of covariance in dynamical estimation of inverse solutions. *Phys. Lett. A*, **333**, 261–268.
[5] *K.F.K. Wong, A. Galka, O. Yamashita* and *T. Ozaki* (2006). Modelling non-stationary variance in EEG time series by state space GARCH model. *Computers in Biology and Medicine*, in press.