31st Annual International Conference of the IEEE EMBS
Minneapolis, Minnesota, USA, September 2-6, 2009

State-Space Multivariate Autoregressive Models for Estimation of Cortical Connectivity from EEG

B. L. Patrick Cheung, Brady Riedner, Giulio Tononi and Barry D. Van Veen

***Abstract*— We propose using a state-space model to estimate cortical connectivity from scalp-based EEG recordings. A state equation describes the dynamics of the cortical signals and an observation equation describes the manner in which the cortical signals contribute to the scalp measurements. The state equation is based on a multivariate autoregressive (MVAR) process model for the cortical signals. The observation equation describes the physics relating the cortical signals to the scalp EEG measurements and spatially correlated observation noise. An expectation-maximization (EM) algorithm is employed to obtain maximum-likelihood estimates of the MVAR model parameters. The strength of influence between cortical regions is then derived from the MVAR model parameters. Simulation results show that this integrated approach performs significantly better than the two-step approach in which the cortical signals are first estimated from the EEG measurements by attempting to solve the EEG inverse problem and second, an MVAR model is fit to the estimated signals. The method is also applied to data from a subject watching a movie, and confirms that feedforward connections between visual and parietal cortex are generally stronger than feedback connections.**

I. INTRODUCTION

Multivariate autoregressive (MVAR) models have been successfully used for studying causal relationships between various cortical regions using invasive recordings [1]-[3]. They have also been used to model topographic interactions between signals at different scalp EEG measurement locations [4]-[5]. Different metrics of connectivity between regions may be obtained from the MVAR model parameters, e.g., [6]-[7].

Identifying cortical MVAR models from scalp EEG data is a challenging problem because the cortical signals are not directly observed. The conventional approach to this problem proceeds in two stages: first cortical signals are estimated by solving the ill-conditioned inverse problem, and then a MVAR model is fit to the estimated signals. For example, Hui and Leahy [8] use linearly constrained minimum variance beamforming to estimate the cortical signal associated with multiple regions of interest (ROIs) from the measured data. Next the MVAR model parameters are estimated by solving the Yule-Walker equations formed from the estimated cortical signals. Astolfi, et al. [7] take a similar approach in which a minimum-norm inverse method is used to first estimate the cortical signals. Another two-step method is described in [9]. Two-step approaches work well at high signal to noise ratio (SNR). However, the effect of noise on the solution of the inverse problem can cause significant errors in the MVAR model parameter estimates at the SNRs typical of EEG data.

In this paper we take an integrated approach for estimating cortical MVAR models from scalp EEG measurements using a state-space formulation of the problem. This formulation leads to maximum-likelihood (ML) estimates of the MVAR model parameters and results in significant performance improvement relative to two-step approaches at low SNR. The MVAR model between cortical regions is described using a "hidden" state equation. An observation equation uses EEG physics to relate the hidden cortical signals to the measured data and models spatially correlated noise unrelated to the MVAR model. The MVAR model parameters are estimated directly from the observed EEG measurements using a variation of the expectation-maximization (EM) algorithm proposed by Shumway and Stoffer [10]. Our observation equation includes unknown parameters that model the unknown spatial distribution of the cortical signal within each ROI.

The paper is organized as follows: Section II introduces the state-space formulation for the MVAR model and a cortical patch basis [11] representation for the observation equation. Cortical patch bases describe uncertainty in the spatial distribution of the signal within each ROI. The EM algorithm for estimating the MVAR model parameters is also summarized in Section II. Section III presents a simulation-based example that illustrates the benefits of the state-space formulation. Section IV reports the application of the state-space approach to EEG data collected from a subject watching a movie and verifies that feedforward connections between the inferior occipital gyrus and the superior parietal lobule are stronger than backward connections. We conclude in Section V.

II. STATE-SPACE MVAR MODEL FOR EEG

Let $\mathbf{y}_n^j$ and $\mathbf{x}_n^j$ be the j^{th} epoch of the $L \times 1$ EEG measurements at L electrodes and the $M \times 1$ state vector, respectively, at time n. The elements of the state vector $\mathbf{x}_n^j$ are the cortical signals associated with the ROIs in the MVAR model and may represent dipoles, multipoles, or cortical patches. We assume that $\mathbf{y}_n^j, n = 1, 2, \ldots, N, j = 1, 2, \ldots, J$ are generated by the linear time-invariant state-space system

This work was supported in part by the National Institutes of Health under Grant No R21EB005473.

B. L. Patrick Cheung and Barry D. Van Veen are with the Department of Electrical and Computer Engineering, University of Wisconsin-Madison, 1415 Engineering Drive, Madison, WI 53706, USA bcheung@wisc.edu, vanveen@engr.wisc.edu.

B. Riedner and G. Tononi are with the Department of Psychiatry, University of Wisconsin-Madison, WI 53706, USA riedner@wisc.edu, gtononi@wisc.edu.

$$
\begin{aligned}
\mathbf{x}_n^j &= \sum_{r=1}^{p} \mathbf{A}_r \mathbf{x}_{n-r}^j + \mathbf{w}_n^j \\
\mathbf{y}_n^j &= \mathbf{C}\mathbf{x}_n^j + \mathbf{v}_n^j \qquad (1)
\end{aligned}
$$

where $\mathbf{A}_1, \mathbf{A}_2, \ldots, \mathbf{A}_p$ are the $M \times M$ matrices of MVAR model coefficients, p is the model order, $\mathbf{C}$ is the $L \times M$ observation matrix, and $\mathbf{w}_n^j$ and $\mathbf{v}_n^j$ are the j^{th} epoch of $M \times 1$ state noise vector and $L \times 1$ observation noise vector respectively. The M columns of the observation matrix $\mathbf{C}$ consist of the forward solutions that map the cortical signals to the EEG measurements.

The spatial distribution of the cortical signal within each ROI is assumed unknown, so $\mathbf{C}$ contains unknown parameters. We write each column of $\mathbf{C}$ as a linear combination of known basis vectors associated with the corresponding ROI and assume the coefficients are unknown. That is, we write the i^{th} column of $\mathbf{C}$ as $[\mathbf{C}]_i = \mathbf{H}_i \boldsymbol{\phi}_i$ where $\mathbf{H}_i$ is a known $L \times Q$ matrix whose columns are a set of basis vectors defining the space spanned by contributions from the i^{th} cortical ROI and $\boldsymbol{\phi}_i$ is a $Q \times 1$ unknown vector of the corresponding basis coefficients. The unknown spatial attributes associated with the i^{th} ROI, i.e., the i^{th} entry in $\mathbf{x}_n^j$, are represented by $\boldsymbol{\phi}_i$. The forward model $\mathbf{H}_i \boldsymbol{\phi}_i$ is able to represent several different classes of source activity distributions. If an equivalent current dipole is employed, then $Q = 3$ and the columns of $\mathbf{H}_i$ represent the forward model for the 3 components of the moment and $\boldsymbol{\phi}_i$ represents the unknown dipole moment. If an extended source representation is desired, then the columns of $\mathbf{H}_i$ can be chosen as either multipole or cortical patch bases and $\boldsymbol{\phi}_i$ represents the corresponding unknown coefficients.

The state noise $\mathbf{w}_n^j$ and observation noise $\mathbf{v}_n^j$ are assumed to be zero-mean Gaussian vectors with unknown covariance matrices $E[\mathbf{w}_n^j \mathbf{w}_m^{i^T}] = \mathbf{Q}\delta_{n,m}\delta_{j,i}$ and $E[\mathbf{v}_n^j \mathbf{v}_m^{i^T}] = \mathbf{R}\delta_{n,m}\delta_{j,i}$. The $M \times 1$ initial state vector $\mathbf{x}_0^j$ is also assumed to be Gaussian distributed with unknown mean vector $\boldsymbol{\mu}_0$ and covariance matrix $\boldsymbol{\Sigma}_0$.

The EM algorithm of Shumway and Stoffer [10] is extended to compute the ML estimates of the unknown model parameters $\mathbf{A}, \mathbf{Q}, \mathbf{R}, \boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0$ and $\boldsymbol{\phi}_i$. The ML estimates of $\mathbf{A}, \mathbf{Q}, \mathbf{R}, \boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0$ are given in our previous work [12]. The ML estimate of $\boldsymbol{\phi}_i$ at iteration r in the M-step of the EM algorithm satisfies the following equation:

$$
\begin{aligned}
\boldsymbol{\phi}_i^r &= \left(\mathbf{H}_i^T \mathbf{R}^{-1} \mathbf{H}_i\right)^{-1} \mathbf{H}_i^T \mathbf{R}^{-1} \\
&\times \left(\sum_{j=1}^{J} \sum_{n=1}^{N} \mathbf{y}_n^j E[\mathbf{x}_{i,n}^j | \mathbf{Y}_N^J, \boldsymbol{\theta}^r] \right. \\
&- \left. \sum_{k=1, k\neq i}^{M} \mathbf{H}_k \boldsymbol{\phi}_k \sum_{j=1}^{J} \sum_{n=1}^{N} E[(\mathbf{x}_{k,n}^j)^2 | \mathbf{Y}_N^J, \boldsymbol{\theta}^r] \right) \\
&\times \left(\sum_{j=1}^{J} \sum_{n=1}^{N} E[(\mathbf{x}_{i,n}^j)^2 | \mathbf{Y}_N^J, \boldsymbol{\theta}^r] \right)^{-1} \qquad (2)
\end{aligned}
$$

where $\mathbf{Y}_N^J = \{\mathbf{y}_1^1, \ldots, \mathbf{y}_N^1, \ldots, \mathbf{y}_1^J, \ldots, \mathbf{y}_N^J\}$ is the matrix of observed data, $\mathbf{R}$ is the current estimate of the noise covariance matrix, $\mathbf{x}_{i,n}^j$ denotes the i^{th} element of $\mathbf{x}_n^j$, and $\boldsymbol{\theta}^r$ are the ML estimates of the other model parameters at iteration r. The statistics $E[\mathbf{x}_{i,n}^j | \mathbf{Y}_N^J, \boldsymbol{\theta}^r]$ and $E[(\mathbf{x}_{i,n}^j)^2 | \mathbf{Y}_N^J, \boldsymbol{\theta}^r]$ are computed from the fixed interval smoother [13] implemented in the expectation step of the EM algorithm. Algebraic manipulation of (2) can be used to obtain an equation that solves for all $\boldsymbol{\phi}_i^r, i = 1, \ldots, M$ simultaneously. We refer to the approach described in this section as the MLEM algorithm.

III. Simulation Example

A simulation scenario is used to illustrate the performance advantages of the MLEM algorithm relative to a two-step approach. The simulation scenario is inspired by the attributes of the real data study presented in the next section. In particular, we simulate cortical activity from the inferior occipital gyrus (IOG) and the superior parietal lobule (SPL) in each hemisphere. A patch of 10 mm geodesic radius at the center of IOGs and SPLs is defined as the ROI as shown in Fig. 1. Fig. 1 also illustrates the assumed connectivity model used in the simulation. Feedforward connections are dominant and feedback connections are absent. We assumed the state noise variances are the same in all four ROIs. We simulated $L = 56$, $N = 1000$ and $J = 1$. We choose the number of patch basis vectors [11] for each ROI as $Q = 3$ and the model order p as 1, and generated spatially white ($\mathbf{R} = \sigma^2 \mathbf{I}$) observation noise.

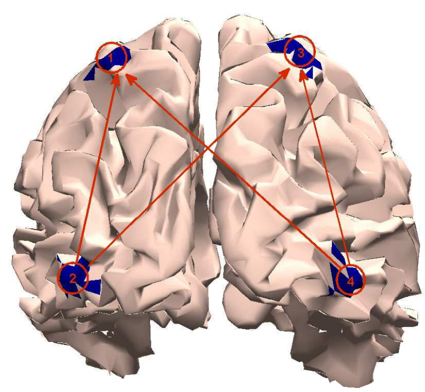

Fig. 1. Locations (in blue) of IOG (region 2 and 4) and SPL (region 1 and 3) and simulated connectivity model. View is from the back.

A two-step approach is implemented based on a least-squares beamformer estimate of the cortical signals. The estimate of $\mathbf{x}_{i,n}^j$ is given by

$$
\hat{\mathbf{x}}_{i,n}^j = \boldsymbol{\phi}_i^{0^T} \mathbf{W}_i^T \mathbf{y}_n^j \qquad (3)
$$

where $\boldsymbol{\phi}_i^0$ is the eigenvector of $\mathbf{W}_i^T \mathbf{S} \mathbf{W}_i$ corresponding to the maximum eigenvalue, $\mathbf{S}$ is the sample covariance matrix of the measured data, $\mathbf{W}_i$ contains columns $(i-1)Q+1$ to iQ of $\mathbf{W} = \mathbf{H}(\mathbf{H}^T\mathbf{H})^{-1}$ and $\mathbf{H} = [\mathbf{H}_1 \mathbf{H}_2 \ldots \mathbf{H}_M]$.

It can be shown that $\mathbf{W}_i$ is orthogonal to $\mathbf{H}_k$ for $k \neq i$ and $\mathbf{W}_i^T\mathbf{H}_i = \mathbf{I}_Q$, so there is no cross-contamination of cortical sources. That is, $\hat{\mathbf{x}}_{i,n}^j$ contains no contributions due to $\mathbf{x}_{k,n}^j, k \neq i$. In the absence of observation noise this least squares beamformer approach perfectly estimates the cortical signals. This beamformer is also equivalent to that of [8] if the observation noise covariance matrix is assumed known. The Yule-Walker equations are then used to obtain the MVAR model parameters from the estimated cortical signals.

Fig. 2 depicts the 16 true and estimated entries of $\mathbf{A}_1$ for 200 independent simulation runs of both the MLEM algorithm and the two-step approach. The SNR, defined as $\mathrm{tr}\{\sum_{j=1}^J \sum_{n=1}^N (\mathbf{C}\mathbf{x}_n^j)(\mathbf{C}\mathbf{x}_n^j)^T\}/\mathrm{tr}\{\sum_{j=1}^J \sum_{n=1}^N (\mathbf{v}_n^j)(\mathbf{v}_n^j)^T\}$, is set to 0 dB. The estimates from the MLEM algorithm are closer to the true values than those from the two-step approach. The true values are within one standard deviation of the mean of the MLEM estimates in every cases, while the two-step method significantly underestimates $\mathbf{A}_1^{11}, \mathbf{A}_1^{12}, \mathbf{A}_1^{14}, \mathbf{A}_1^{22}$ and $\mathbf{A}_1^{44}$ while overestimating $\mathbf{A}_1^{13}, \mathbf{A}_1^{23}$ and $\mathbf{A}_1^{24}$. Here $\mathbf{A}_1^{kl}$ denotes the $(l,k)^{th}$ element of $\mathbf{A}_1$. These errors are a consequence of the observation noise and increase as the SNR decreases.

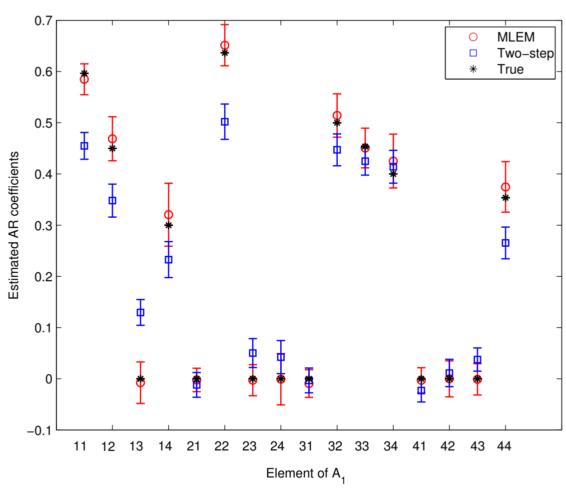

Fig. 2. Mean and standard deviation for the 16 estimated entries in $\mathbf{A}_1$ at SNR of 0 dB calculated over 200 independent simulations. MLEM: red circle, Two-step: blue square, True: black asterisk.

Note that since $p = 1$ in this example, $\mathbf{A}_1^{kl}$ represents the strength of the connection from ROI l to ROI k. In cases where $p > 1$ it is common to use other metrics to measure the strength of influence between ROIs. In this paper we consider the impulse response from one ROI to another and measure the strength of the connection based on the energy of the impulse response. The $M \times M$ impulse response matrix between all ROIs at time i, $\mathbf{\Psi}_i$, is calculated from the estimated MVAR parameter matrices $\mathbf{A}_1, \ldots, \mathbf{A}_p$ as follows [14]:

$$\mathbf{\Psi}_i = \sum_{j=1}^{i} \mathbf{\Psi}_{i-j}\mathbf{A}_j \qquad i = 1, 2, \ldots \tag{4}$$

where $\mathbf{\Psi}_0 = \mathbf{I}_M$ and $\mathbf{A}_j = 0$ for $j > p$.

Fig. 3 depicts the mean and standard deviation of the estimated impulse response energy between each pair of ROIs over 200 independent simulation runs. The improved quality estimates of $\mathbf{A}_1$ obtained with the MLEM algorithm lead to more accurate estimates of the impulse response energy. In particular, the two-step approach significantly underestimates the influence of ROI 2 on ROI 1 and of ROI 2 on ROI 3.

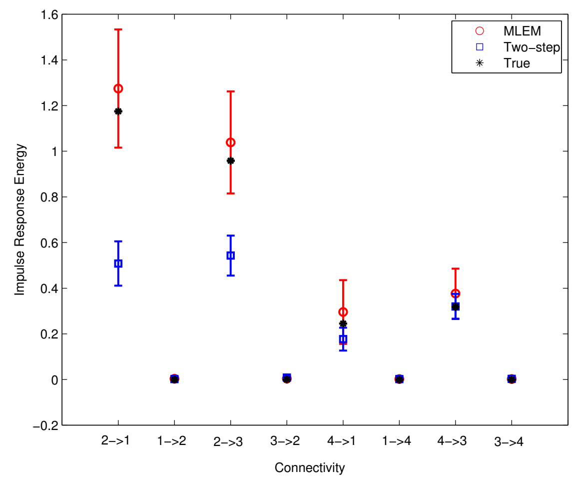

Fig. 3. Mean and standard deviation of the estimated impulse response energy at SNR of 0 dB calculated over 200 independent simulations. MLEM: red circle, Two-step: blue square, True: black asterisk.

IV. Application to Real EEG Data

Data was derived from high-density EEG (256 channel) recordings of a healthy subject passively watching a segment of an engrossing movie (*The Good, The Bad, and The Ugly*). The data were sampled at 1000 Hz and hardware filtered (0.1 Hz - 400 Hz). The data was preprocessed using a low pass Chebyshev Type II filter with 30 Hz cutoff frequency and down sampled to a 62.5 Hz sampling rate. Three artifact free segments of 9.6 sec were selected for processing. Each segment is considered as an epoch in (1). The temporal mean of each trial is subtracted from the data to remove DC offsets. Each of the four ROIs are modeled using three patch bases and an MVAR model of order p = 4 is estimated using the EM algorithm.

Fig. 4 depicts the energy of the impulse response between different connections. "Forward" denotes directed connections from an IOG to an SPL ROI, while "backward" denotes directed connections from an SPL to an IOG ROI. The rightmost set of bars represents the effective forward and backward connectivity calculated as the sum of the impulse response energies across all four connections. The results suggest that feedforward information flow dominates feedback information flow within each hemisphere and across the entire system. This is consistent with the fact that the IOG is a primary sensory area while the SPL is involved in sensory integration and higher order information processing [15] [16].

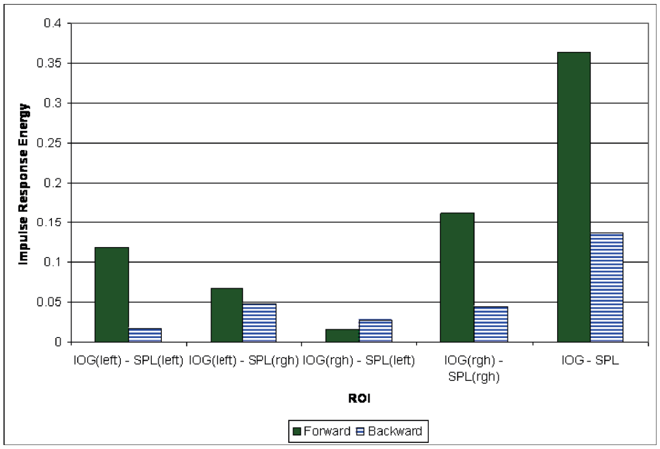

Fig. 4. Impulse Response Energy derived from MLEM estimated MVAR model for a subject passively watching a movie. IOG → SPL: green solid, SPL → IOG: blue stripe.

V. CONCLUSION

We have formulated the problem of estimating cortical connectivity from the scalp EEG measurements as a state-space model and applied the EM algorithm to estimate the model parameters. A patch basis model is used to describe the unknown source activity distribution within each ROI, and the ML estimates of the basis coefficients for each ROI are obtained via the EM algorithm. We demonstrate through simulation that using the state-space model with an EM algorithm gives improved quality MVAR parameter estimates compared to a two-step approach based on beamforming to estimate cortical signals. The result of applying the EM approach to data from a subject watching a movie confirms the dominance of feedforward information flow.

REFERENCES

[1] M. Ding, S. Bressler, W. Yang and H. Liang, "Short-window spectral analysis of cortical event-related potentials by adaptive multivariate autoregressive modeling: data preprocessing, model validation, and variablility assessment", *Biological Cybernetics*, **83**, 2000, pp 35-45.

[2] L. Baccala and K. Sameshima, Partial directed coherence: a new concept in neural structure determination, *Biological Cybernetics*, **84**, 2001, pp 463-474.

[3] A. Brovelli and M. Ding and A. Ledberg and Y. Chen and R. Nakamura and S. Bressler, "Beta osciallations in a large-scale sensorimotor cortical network: Directional influences revealed by Granger causality", *PNAS*, vol. 101, no. 26, 2004, pp 9849-9854

[4] M. Kaminski and M. Ding and W. Truccolo and S. Bressler, "Evaluating causal relations in neural systems: Granger causality, directed transfer function and statistical assessment of significance", *Biological Cybernetics*, **85**, 2001, pp 145-157.

[5] R. Kus, M. Kaminski and K.J. Blinowska, "Determination of EEG activity propagation: pair-wise vs. multichannel estimate", *IEEE Transactions on Biomedical Engeineering*, vol. 51, issue 9, 2004, pp 1501-1510.

[6] M. Winterhalder et al., "Comparison of linear signal processing techniques to infer directed interactions in multivariate neural systems", *Signal Processing*, **85**, 2005, pp 2137-2160

[7] L. Astolfi et al., "Comparison of different cortical connectivity estimators for high-resolution EEG recordings", *Human Brain Mapping*, **28**, 2007, pp 143-157.

[8] H. Hui and R. Leahy, "Linearly Constrained MEG Beamformers for MVAR Modeling of Cortical Interactions", *3rd IEEE International Symposium on Biomedical Imaging: Macro to Nano*, 2006, pp 237-240.

[9] L. Ding, G.A. Worrell, T.D. Lagerlund and B. He, "Ictal Source Analysis: Localization and Imaging of Causal Interactions in Humans", *Neuroimage*, **34**, 2007, pp 575-586.

[10] R.H. Shumway and D.S. Stoffer, "An approach to time series smoothing and forecasting using the EM algorithm", *Journal of Time Series Analysis*, vol. 3, num. 4, 1982, pp 253-264.

[11] T. Limpiti and B.D. Van Veen and R.T. Wakai, "Cortical Patch Basis Model for Spatially Extended Neural Activity", *IEEE Transaction on Biomedical Engineering*, vol. 53, issue 9, 2006, pp. 1740-1754.

[12] B.L. Cheung and B.D. Van Veen, "Estimation of Cortical Multivariate Autoregressive Models for EEG/MEG Using an Expectation-Maximization Algorithm", *5th IEEE International Symposium on Biomedical Imaging: Marco to Nano*, 2008, pp 1235-1238.

[13] H. Rauch and F. Tung and C. Stiebel, "Maximum likelihood estimates of linear dynamic systems", *AIAA Journal*, vol. 3, issue 8, 1965, pp 1445-1450.

[14] H. Lutkepohl, *Introduction to Multiple Time Series Analysis*, Springer-Verlag, Berlin-Heidelberg; 1993.

[15] A. Mechelli, C.J. Price, K.J. Friston and A. Ishai, "Where bottom-up meets top-down: neuronal interactions during perception and imagery", *Cerebral Cortex*, **14**, 2004, pp 1256-1265.

[16] G. Gains, W.L. Thompson and S.M. Kosslyn, "Brain areas underlying visual mental imagery and visual preception: an fMRI study", *Cognitive Brain Research*, vol. 20, issue 2, 2004, pp 226-241.