A MARGINALISED MARKOV CHAIN MONTE CARLO APPROACH FOR MODEL BASED ANALYSIS OF EEG DATA

Imali Hettiarachchi, Shady Mohamed, Saeid Nahavandi

Centre for Intelligent Systems Research
Deakin University
Australia

ABSTRACT

The work presented in this paper focuses on fitting of a neural mass model to EEG data. Neurophysiology inspired mathematical models were developed for simulating brain's electrical activity imaged through Electroencephalography (EEG) more than three decades ago. At the present well informative models which even describe the functional integration of cortical regions also exists. However, a very limited amount of work is reported in literature on the subject of model fitting to actual EEG data. Here, we present a Bayesian approach for parameter estimation of the EEG model via a marginalized Markov Chain Monte Carlo (MCMC) approach.

Index Terms— Electroencephalography, Nonlinear dynamical systems, Bayesian methods, Parameter Estimation, Particle Filter

1. INTRODUCTION

Electroencephalography (EEG) is a noninvasive functional imaging modality which records the brain's electrical activity at a very high time resolution. Also EEG has the inherent ability of measuring the neural activity more directly compared to other noninvasive modalities such as functional magnetic resonance imaging (fMRI). Conversely, EEG has a lower spatial resolution than fMRI reconstructed images; nevertheless the high temporal resolution is a significant advantage when studying brain dynamics.

The EEG measurements are acquired through multiple surface electrodes placed on the scalp. The recorded electrode potentials are a summed result of the extracellular current flow of sources (brain areas), associated with synchronously activated and vertically oriented large-scale neuronal populations. However, the exact neurophysiological mechanism behind this collective behavior is not totally understood. Several mathematical modeling attempts were taken to explain the physiological bases of the source dynamics, one of them being the neural mass models (NMM).

The NMM's summarizes the behavior of millions of interacting neurons and present them with a few state variables in a nonlinear dynamic system formulation and are capable of producing EEG-like waveforms. The NMM's are originally developed to simulate the alpha rhythms [1] and the activity in the olfactory cortex [2]. Later these models have been extended to simulate spontaneous rhythms in other frequency bands [3], to synthesize epileptic activities [4], to analyse the connectivity and coherence on EEG rhythms [5], to generate the event related responses [6] and data fusion of EEG, magnetoencephalography (MEG), and fMRI [7].

Also the NMM's have been used in effective connectivity analyses [7, 8]. Dynamic Causal Models (DCM) for EEG data analysis [8] uses the NMM presented in referred to as the Jansen's model, to characterize the source dynamics. DCM for EEG basically investigates how the coupling among brain regions is influenced by changes in the experimental context.

The NMM's have been used extensively to reproduce the empirically observed responses and to make mechanistic conclusions of the generation of the signals. However there are only a few instances where model fitting of EEG data is reported in literature. Model fitting let us estimate the values of unknown parameters in a the NMM by fitting the model to real EEG data

The first model fitting to EEG data is reported in [9], where the authors use a nonlinear Kalman filter methodology (Local Linearization filter) for a neural mass model in [10] for the generation of alpha rhythms. The parameter estimation of the model to fit the spontaneous EEG data is done through a maximum likelihood approach through minimizing the negative log likelihood.

The work in [11] uses a NMM fitting approach for the EEG data. They use a genetic algorithm for the parameter estimation of the Jansen's model for the peri-stimulus EEG activity, and then use the estimates to eliminate the alpha related activity with the intention of enhancing the evoked potential through adaptive filtering. The work presented in [12] uses a square-root unscented Kalman filter to estimate the extrinsic input through fitting the actual EEG signal to the Jansen's neural mass model.

Amongst the various available neural mass models [1, 2, 13] our work is focused on the Jansen's model [13], which is originally presented as a lumped parameter model

in [1] to model alpha rhythms. In this paper we present a Bayesian method to estimate the posterior distribution of the parameters of the model. A Bayesian framework is important in constraining the parameter estimation due to the large number of parameters. It also has the advantage of estimating a full posterior of the estimate rather than a single point estimate.

We use a Markov Chain Monte Carlo (MCMC) approach for parameter estimation. Also we adjoin the MCMC approach with another powerful Bayesian technique, the Particle Filter (PF) [14, 15]. The physiologically plausible parameters in the NMM's have an imminent interpretability thus, despite the high computational cost involved when using MCMC strategies; the information carried in the estimation is important for inference.

We will focus on system identification to infer physiological quantities such as intrinsic coupling parameters and synaptic time constants of a single source.We present simulation results for the stimulus-evoked EEG data and validate the proposed method to be used in experimental data analysis.

2. JANSEN'S MODEL

Jansen's model [13] describes a source using three neuronal subpopulations. A column representation of the source is modelled by a population of excitatory pyramidal cells, receiving inhibitory feedback from the local interneuron and excitatory input from the spiny stellate cells.

Each of the neuron populations is modelled at two stages. The first stage transforms the average pulse density of presynaptic input coming to the population of neurons into an average postsynaptic membrane potential (PSP) which can be either excitatory or inhibitory. This is a linear transformation with an impulse response given by, $h(t)_{e,i} = \frac{H_{e,i}}{\tau_{e,i}} t e^{(-t/\tau_{e,i})}$, where, subscripts e, i represent excitatory and inhibitory synapses respectively. The second stage transforms the average membrane potential of a neural population into an average firing rate through a nonlinear transformation,$S(v) = \frac{2e_0}{1+e^{-(rv)}} - e_0$, where $r = 0.56$and $e_0 = 2.5$as given in [6]. Below is the model description for the three neuronal subpopulations,

$$\begin{aligned}
\dot{x}_5 &= \frac{H_e}{\tau_e} c_2 S(x_1) - \frac{2x_5}{\tau_e} - \frac{x_2}{\tau_e^2} \\
\dot{x}_6 &= \frac{H_i}{\tau_i} c_4 S(x_7) - \frac{2x_6}{\tau_i} - \frac{x_3}{\tau_i^2} \\
\dot{x}_4 &= \frac{H_e}{\tau_e} c_1 S(x_0 + c_u u) - \frac{2x_4}{\tau_e} - \frac{x_1}{\tau_e^2} \\
\dot{x}_8 &= \frac{H_e}{\tau_e} c_3 S(x_0) - \frac{2x_8}{\tau_e} - \frac{x_7}{\tau_e^2}
\end{aligned} \tag{1}$$

where c_{1-4} are the intrinsic connections among the sub populations and are given as a fraction of a connectivity constant c_c as reported in [13] with $c_1 = c_c, c_2 = 0.8c_c, c_3 = 0.25c_c$ and $c_4 = 0.25c_c$,c_u is the extrinsic stimulus input strength and u the stimulus input. x_0, x_1 and x_4 are the overall PSP (output) of pyramidal cells, stellate cells and inhibitory cells respectively. Also x_2 and x_3 are the overall excitatory and inhibitory PSP of pyramidal cells. In order to maintain the consistency in the formulation the differential of the pyramidal subpopulation output $\dot{x}_0$included in the equations as $\dot{x}_0 = x_5 - x_6$ together with $\dot{x}_2 = x_5$, $\dot{x}_3 = x_6$, $\dot{x}_1 = x_4$ and $\dot{x}_7 = x_8$, collectively giving 9 ordinary differential equations.

For simplicity at this stage, we consider the measured EEG signal to be the depolarization of the pyramidal cells. Thus the output of the system is given by, $y = x_2 - x_3$.

Taking into account the physiological noise and the measurement noise, the system can be rewritten in a general state space form with stochastic differential equations (SDE) as,

$$dx_t = f(x_t, u_t, \theta)dt + BdW_t \tag{2}$$

$$y_t = gx_t + e_t \qquad e_t \sim N(0, \sigma_e^2) \tag{3}$$

Where,$x = [x_0, x_1, \cdots, x_8]$ the unobserved states of the system, $\theta = [H_e, H_i, \tau_e, \tau_i, c_1, c_2, c_3, c_4, c_u]$ the model parameters, $g = [0\ 0\ 1\ -1\ 0\ 0\ 0\ 0\ 0]^T$, B is a weighting matrix and W_t is a vector of wiener processes. The superscript T denotes the transpose of a vector and $N(\mu, \sigma^2)$ denotes a Normal distribution with mean μ and variance σ^2 all throughout the paper.

3. ESTIMATION METHODS

In Bayesian estimation the estimation of parameters is based on evaluating the posterior distribution $p(\theta|Y_T)$,

$$p(\theta|Y_T) \propto p(\theta)p(Y_T|\theta) \tag{4}$$

where, $p(\theta)$ is the prior on θ and $p(Y_T|\theta)$ the likelihood (marginal) of data. We use a MCMC approach for evaluating this target posterior via a proposal distribution, since drawing samples from the posterior is not straight forward. MCMC basically provides a method of computing high-dimension integrals using Markov chains [16].

In particularly here we use a Metropolis-Hastings (MH) algorithm for the random generation of the samples. MH draws samples from a proposal distribution $\theta^* \sim \pi(\theta^*|\theta)$ and then accepts the value with probability $min\{1, A\}$ where for our application,

$$A = \frac{p(\theta^*)\pi(\theta|\theta^*)p(Y_T|\theta^*)}{p(\theta)\pi(\theta^*|\theta)p(Y_T|\theta)} \tag{5}$$

If it is rejected then current point becomes the next point, $\theta_{t+1} = \theta_t$. This process is repeated $t = 1, 2, \cdots, n$ where n is

the total number of steps the chain. The state of the chain after a large number of steps is then used as a sample of the target distribution. To follow this method we need to evaluate the marginal likelihood $p(Y_T|\theta)$ which is not straight forward with latent variables in the model given in (1).Therefore, we use a joint PF and MH MCMC scheme, referred to as a Pseudo marginal MCMC estimation [17, 18]. Given the physiological parameters, state estimation of the EEG model is carried out by the PF. With this approach we break up the interdependence of the hidden states and the static parameters by marginalizing out the states using a PF.

Among the available estimation techniques for the nonlinear dynamic systems, the PF has proved to be a powerful nonlinear estimation technique in a diverse range of applications [14, 15]. The PF's uses a set of weighted random samples; $\{x_t^i, w_t^i\}$, where $i = 1,2, \ldots, N_s$ and N_s is the no of particles to approximate the posterior density [14]. Detailed discussions on the PF is omitted here and can be found in [14, 15]. For this work, we use a Sequential importance resampling (SIR) algorithm presented in [14].

The proposal distribution $\pi(.|.)$ for the target distribution $p(\theta|Y_T)$ is chosen as a random walk $\theta_k^* = \theta_{k-1} + v_k$ with $v_k \sim N(0, \sigma_v^2)$. The performance accuracy of MH MCMC critically depends on the selection of the variance of the proposal distribution; here, it is selected on observation of the mixing of generated sample chain. The above algorithm is to perform the MCMC in one dimension. In this paper, we use an identical multi-dimensional version for our EEG parameter estimation problem, where the parameter vector is more than one element.

4. SIMULATION DATA

The system of equations (2) and (3) do not have an analytic solution. Also, the algorithms in estimation, work in discrete time. Thus, the equation (2) is implemented with a time discretization. We use an Euler-Maruyama method with a sampling time (Δt) of 0.001s,

$$x_{t+\Delta t} = x_t + f(x_t, u_t, \theta)\Delta t + B\Delta W_t \quad (6)$$

where $\Delta W_t \sim N(0, \Delta t I)$, where in this case I is a 9 X 9 identity matrix. The noise level are set at $\sigma_e^2 = 0.01$and $B = [0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0.1]$. We lump the process (physiological) noise at the pyramidal cell output for simplicity.

The EEG signal consists of two parts: evoked response due to stimulus and ongoing (spontaneous) activity. The work in[6]shows that at low input levels the response components due to stimulus and ongoing activity can be linearly separable. Thus, in the simulations we set the extrinsic input level to $c_u = 100$, and model the input solely with a block input function (Fig.1 (a)), without modeling for the background activity. so the simulated EEG signal (Fig.1 (b)) only reflects the response of the stimuli. c_c is set to 50. Other parameter values used in the simulation are, $H_e = 3.25, H_i = 29.3, \tau_e = 10ms$and $\tau_i = 15ms$.

5. RESULTS AND DISCUSSION

Since there are no ground truth values available for the physiological parameters of the NMM, it is hard to evaluate the performance of an algorithm directly for experimental data. Thus, at this stage we use simulated data to demonstrate the strength of the proposed method for adapting to experimental data analysis.

Fig.1 (c)-(e) shows the PF performance on the simulated noisy EEG signals. Given the parameters to be fixed at values aforementioned the PF gives accurate estimates on the unobserved states. As shown in Fig. 1 (c)-(e), the estimated states (dashed-dotted lines) are very close to the true states shown in solid lines.

To demonstrate the marginalized MH MCMC parameter estimation for the simulated EEG data, here, we limit the parameter space θ only to three parameters of the NMM, as $\theta = [H_e,\ \tau_e,\ c_u]$. All the other parameters are fixed to the values mentioned earlier.

Starting value for the chains is selected as 1 in all cases. Fig. 2 (top) shows the burn-in period (time to start convergence) and mixing of the chain for the parameter H_e, with $\sigma_v^2 = 0.01$. The burn-in period is 500 iterations. Fig. 2 (bottom) shows the mixing of the chain after convergence. The chain mixes well and thus the proposal for H_e is selected as $\theta_{He,k}^* = \theta_{He,k-1} + N(0, 0.1^2)$. Similarly, the variance for parameters τ_e and $c_u$are selected as 0.0001 and 5 respectively.

Fig. 3 shows the histograms of estimated Markov chains. The normal fit to the histograms (dashed lines) are taken as the posterior distribution of the parameters. The fitted Normal distributions for parameters H_e , τ_e and c_u are $N(3.2415, 0.1016^2)$, $N(0.01, 0.00027^2)$ and $N(100.108, 11.007^2)$ respectively. The mean value being very close to the simulated values show that the proposed method of PF marginalised MH MCMC is a robust technique having a great potential in application to EEG parameter estimation.

Above methods can be extended to estimate any number of parameters. However, with correlations among the parameters, it is not guaranteed to give a unique solution. For the NMM presented in the paper if $\theta = [H_e,\ \tau_e,\ c_u, H_i\ ]$, this issue comes in to existence. However, taking $\theta = [A_e,\ c_u,\ A_i\ ]$ where, $A_{e,i} = H_{e,i}/\tau_{e,i}$, a unique solution can be obtained. Thus, precaution should be taken at the stage of model parameterization. With a high dimensional parameter space it is necessary to either constrain the parameters by fixing some terms or using strongly informative priors.

6. CONCLUSION

In this paper we demonstrate the potential of the MH MCMC algorithm to be applied in model fitting of EEG data. The high temporal resolution of the EEG data makes the PF perform excellent in estimating the latent variables.

The work presented here focuses on a single cortical region. A single mini-column representation of a source is not always sufficient to simulate complex activity. Thus, coupled neural masses and multiple source integration through long range connections of the brain will be incorporated in the future studies.

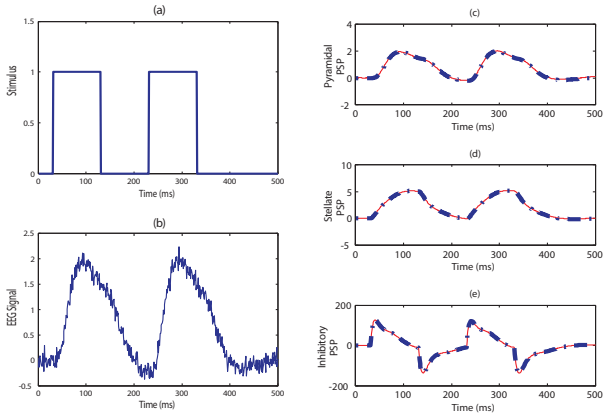


Fig. 1.RIGHT: Stimulus sequence (top) and Simulated EEG signal (bottom) LEFT: State estimation for the simulated EEG signal: True states-thick solid, estimated states-dashed lines

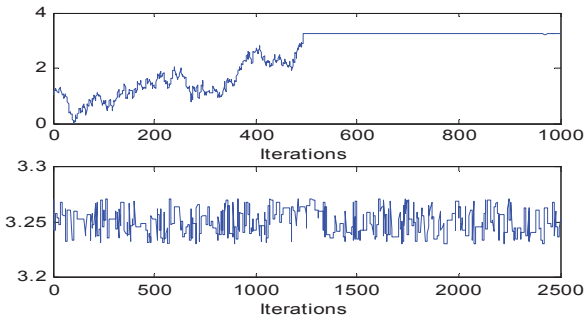


Fig. 2. Markov chain for parameter H_e. TOP: Burn-in period of 500 iterations. BOTTOM: Mixing of the chain after convergence. Y axis shows the value of the sample at each iteration.

7. REFERENCES

[1] F. H. Lopes Da Silva, A. Hoeks, H. Smits and L. H. Zetterberg, "Model of brain rhythmic activity. The alpha rhythm of the thalamus," *KYBERNETIK,* vol. 15, pp. 27-37, 1974.

[2] W. J. Freeman, "Simulation of chaotic EEG patterns with a dynamic model of the olfactory system," *Biological Cybernetics,* vol. 56, pp. 139-150, 1987.

[3] O. David and K. J. Friston, "A neural mass model for MEG/EEG: Coupling and neuronal dynamics," *NeuroImage,* vol. 20, pp. 1743-1755, 2003.

[4] F. Wendling, F. Bartolomei, J. J. Bellanger and P. Chauvel, "Epileptic fast activity can be explained by a model of impaired GABAergic dendritic inhibition," *European Journal of Neuroscience,* vol. 15, pp. 1499-1508, 2002.

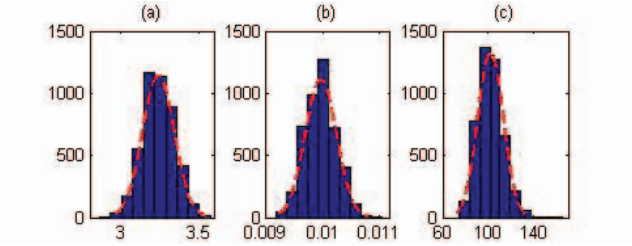


Fig. 3.Histograms of MH generated samples after 10000 iterations (a) H_e (b) τ_e (c) $c_u$and the Normal fit for the posterior density estimates (dashed lines)

[5] M. Zavaglia, L. Astolfi, F. Babiloni and M. Ursino, "The effect of connectivity on EEG rhythms, power spectral density and coherence among coupled neural populations: Analysis with a neural mass model," *IEEE Transactions on Biomedical Engineering,* vol. 55, pp. 69-77, 2008.

[6] O. David, L. Harrison and K. J. Friston, "Modelling event-related responses in the brain," *NeuroImage,* vol. 25, pp. 756-770, 2005.

[7] J. Riera, E. Aubert, K. Iwata, R. Kawashima, X. Wan and T. Ozaki, "Fusing EEG and fMRI based on a bottom-up model: Inferring activation and effective connectivity in neural masses," *Philosophical Transactions of the Royal Society B: Biological Sciences,* vol. 360, pp. 1025-1041, 2005.

[8] S. J. Kiebel, M. I. Garrido, R. J. Moran and K. J. Friston, "Dynamic causal modelling for EEG and MEG," *Cognitive Neurodynamics,* vol. 2, pp. 121-136, 2008.

[9] P. A. Valdes, J. C. Jimenez, J. Riera, R. Biscay and T. Ozaki, "Nonlinear EEG analysis based on a neural mass model," *Biological Cybernetics,* vol. 81, pp. 415-424, 1999.

[10] L. H. Zetterberg, L. Kristiansson and K. Mossberg, "Performance of a model for a local neuron population," *Biological Cybernetics,* vol. 31, pp. 15-26, 1978.

[11] B. H. Jansen, A. Balaji Kavaipatti and O. Markusson, "Evoked potential enhancement using a neurophysiologically-based model," *Methods of Information in Medicine,* vol. 40, pp. 338-345, 2001.

[12] X. Fang, Z. Hu and P. Shi, "Neural mass model driven nonlinear EEG analysis," vol. 6326 LNCS, ed, 2010, pp. 457-466.

[13] B. H. Jansen and V. G. Rit, "Electroencephalogram and visual evoked potential generation in a mathematical model of coupled cortical columns," *Biological Cybernetics,* vol. 73, pp. 357-366, 1995.

[14] M. S. Arulampalam, S. Maskell, N. Gordon and T. Clapp, "A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking," *IEEE Transactions on Signal Processing,* vol. 50, pp. 174-188, 2002.

[15] A. Doucet, N. d. Freitas and N. E. Gordon., Eds., *Sequential Monte Carlo Methods in Practice*. New York: Springer, 2001.

[16] W. R. Gilks, S. Richardson and D. J. Spiegelhalter, Eds., *Markov chain Monte Carlo in practice* Chapman & Hall, 1996.

[17] C. Andrieu and G. O. Roberts, "The pseudo-marginal approach for efficient Monte Carlo computations," *Annals of Statistics,* vol. 37, pp. 697-725, 2009.

[18] E. Jones, J. Parslow and L. Murray, "A bayesian approach to state and parameter estimation in a phytoplankton-zooplankton model," *Australian Meteorological and Oceanographic Journal,* vol. 59, pp. 7-16, 2010.