Flexible Tracking of Auditory Attention

Majid Mirbagheri[1], *Bradley Ekin*[3], *Les Atlas*[3], *Adrian KC Lee*[1,2]

[1]Institute for Learning and Brain Sciences
[2]Department of Speech & Hearing Sciences, [3]Department of Electrical Engineering
University of Washington
Seattle, WA, USA
{mbagheri,bekin,atlas,akclee}@uw.edu

Abstract

Auditory selective attention plays a central role in the human capacity to reliably process complex sounds in multi-source environments. The ability to track the attentional state of individuals in such environments is of interest to neuroscientists and engineers due to its importance in the study of attention-related disorders and its potential application in the hearing aid and advertising industries. The underlying neural basis of auditory attention is not well established, however evidence exists to suggest that cortical activity entrainment to the temporal envelope of speech is modulated by attention. Leveraging this finding, we introduce a probabilistic approach based on Hidden Markov Model Regression (HMMR) to decode the attentional state of the listener with respect to a given speech stimulus. Our method is novel in that it uses only the target stream to detect the attended segments, while existing methods require knowledge about the target and distractor. This is a particular advantage in real-world applications where the number of sources are often time-variant and unknown to the decoder. We use synthetic data to evaluate robustness and tracking capability, and real electrophysiological data to demonstrate how the proposed method achieves accuracies commensurate to BCI (Brain Computer Interface) systems deployed in the field.

Index Terms: auditory, attention tracking, brain computer interface, hidden markov model, regression

1. Introduction

Humans are able to "tune in" to speech sources of interest in noisy environments. Known as the cocktail party effect, this capability of the brain has been investigated extensively by neuroscientists in order to find biomarkers of auditory attention and gain better insight into the underlying neural processes. The two fundamental and interconnected questions often arising in this area are: 1) how does the brain segregate and attend to distinct sound sources, and 2) how can researchers decode the cognitive state of brain during auditory tasks? This work focuses on the latter question with the aim of providing improved human-assisted sound segregation methods for future automatic speech recognition (ASR) and speech processing techniques for hearing aid solutions.

Recent studies suggest that the encoding of auditory objects in the brain while listening to competing speakers is affected by selective attention. Invasive studies with electrocorticography (ECoG) have shown that the high-gamma power activity of the brain (70-150 Hz) favors the attended (target) over the unattended source (distractor) [1]. Other studies show that low-frequency cortical activity ($<$8 Hz) recorded non-invasively via electroencephalography (EEG) and magnetoencephalography (MEG) follows the slow temporal envelope of speech [2, 3]. It has been demonstrated that in dichotic listening (one stream to the left ear, another to the right), cortical entrainment to an attended speech envelope is substantially stronger than to an unattended envelope [4]. This contrast has been used as a criterion to decode the attentional state of the listener in a competing-speaker environment using multi-trial MEG [5], and single-trial EEG [6] recordings. In this arrangement, it is possible to ascertain which of two speech sources is the attended one based on similarity between a reconstruction of the attended stimulus envelope from neural recordings and the envelope of each of the two waveforms over long periods (60 s). Low signal-to-noise ratio (SNR) inherent in MEG and EEG signals forces these methods to rely on multiple repetitions or long durations of data recording. In a separate work focused on addressing this poor temporal resolution [7], a state-space model was used to construct a decoder to estimate attentional states as a function of the correlation between the predicted and real neural responses.

All the above methods share the same assumption of the competing stimulus being available for the decoding. In real-world scenarios however, there is typically no prior knowledge about the distractor signal; there may be no distractor present or there may be multiple distractors that are fully or partially unknown to the decoder. In this paper, we consider the more general problem in which only one target speech signal among all sources is given, and the objective is to detect the time periods when the subject is attending to the signal. For this purpose, we use a probabilistic and dynamic framework in the form of a hidden Markov model (HMM) to decode the attentional state of the listener with respect to the target signal using simultaneous MEG and EEG (M-EEG) recordings. By allowing for constrained temporal evolution of hidden states, HMMs have been successfully used for tracking in adverse conditions for a variety of speech-related applications such as pitch estimation [8] and voice activity detection [9].

In our approach, the HMM states consist of two linear mappings of the target speech to the neural activity modeled by an impulse response function and noise variance corresponding to attended and unattended segments. The length of observation windows and transition probabilities in this model are fixed parameters that constrain temporal dynamics of attentional states based on behavioral and neuroimaging studies suggesting that the process of switching auditory attention between competing streams occurs over time scales on the order of 150-300 ms [10, 11]. In the following section, we present a description of our proposed model. In section 3, we provide details about experimental design and results. We finally conclude our paper

with a few discussions and directions for the future research.

2. The Proposed Hidden Markov Model for Attention Tracking

In our approach, we consider a forward-mapping model to formulate the neural response as a linear function of the auditory stimulus envelope. Known as temporal response functions (TRF), these linear mappings provide a window into different stages of auditory processing in the brain and are thought to be modulated by the attentional state of the listener [4]. The approach we pursue here is to decode attentional modulations by segmenting the neural data based on TRF regime changes. Numerous methods have been proposed for modeling time series with regime changes; of these, piecewise regression is perhaps the best known [12]. However these approaches are unsupervised and do not account for prior knowledge about TRFs. To address this problem, we consider an alternative approach based on Hidden Markov Model Regression (HMMR) [13]. Viewed as an extension of standard HMMs, HMMR preserves the Markov process modeling of HMM with the difference that each state is associated with a regression model rather than constant distributions over time. We now present a basic description of HMMRs for attention tracking.

2.1. Hidden Markov Model Regression

Suppose an acoustic signal with envelope e_i is presented to a subject, evoking the auditory neural response r_i. In our model, the response is assumed to be generated by the following regression model:

$$r_i = \mathbf{g}_{z_i}^T \mathbf{e}_i + \sigma_{z_i}\epsilon_i \;;\; \epsilon_i \sim \mathcal{N}(0,1), \quad (i = 1, \ldots, N) \tag{1}$$

where i is the time index, and $\{z_i\}$ is the hidden discrete-valued sequence taking its values from the set $\{0, 1\}$, which corresponds to unattended and attended states. The τ-dimensional vector $\mathbf{g}_{z_i} = (g_{z_i,0}, \ldots, g_{z_i,\tau-1})^T$ represents the TRF coefficients for the regression model z_i, and σ_{z_i} is the corresponding standard deviation. $\mathbf{e}_i = (e_i, e_{i-1}, \ldots, e_{i-\tau+1})^T$ is a window of the input signal sequence, and ϵ_i's are assumed to be independent standard Gaussian variables representing additive noise. To model slow temporal variations of attentional state, we assume z_i changes value every L samples with $k = \lfloor N/L \rfloor$, $z'_j = z_{(j-1)L+1}$, and $\mathbf{z} = (z'_1, \ldots, z'_k)$ being a homogeneous Markov chain of first order parametrized by the initial state distribution π and the transition matrix A. The HMMR model is then fully parametrized by the parameter vector $\boldsymbol{\theta} = (\pi, A, \mathbf{g}_0, \mathbf{g}_1, \sigma_0, \sigma_1)$. While all these parameters can be estimated from the observed neural response in a fully-unsupervised fashion, we employ a semi-supervised approach by fixing A, learning $\mathbf{g}_{z_i}$'s from a training set offline, and estimating π and σ_{z_i}'s in an unsupervised way from the observed test data. By setting the value of self-transition probabilities in A close to 1, we aim to further impose continuity of $\mathbf{z}$ over time. It has been further shown that while TRFs for attended and unattended auditory stimuli are subject-dependent, their templates are more or less consistent over time [5]. However, this consistency cannot be the case for the state priors and the noise variances, particularly for our problem setup in which the distractor source is assumed to be unknown and can vary from session to session.

2.2. Estimation of Temporal Response Functions

Temporal response functions (TRFs) for attended and unattended states are each estimated offline using neural data recorded in a competing-speaker environment with both attended and unattended streams available. Given the stimuli and the response, TRFs can be computed in different ways. We follow the same procedure as [7] to estimate $\mathbf{g}_{z_i}$'s by finding the solution to the following ℓ_1-regularized least squares problem:

$$\mathbf{g}_0^*, \mathbf{g}_1^* = \tag{2}$$
$$\arg\min_{\mathbf{g}_0, \mathbf{g}_1} \sum_{i=1}^{N} \|r_i - \mathbf{g}_0^T \mathbf{e}_i^u - \mathbf{g}_1^T \mathbf{e}_i^a\|_2^2 + \gamma_0 \|\mathbf{g}_0\|_1 + \gamma_1 \|\mathbf{g}_1\|_1$$

with $\mathbf{e}_i^u$ and $\mathbf{e}_i^a$ respectively denoting the broadband envelope of the unattended and attended speech waveforms. The ℓ_1 regularization term encourages the sparsity known to be a characteristic of TRFs and hence results in more generalizable regression models. We used an interior-point method described in [14] to solve the optimization problem. A k-fold cross-validation procedure (k = 5) was also conducted to find the optimum values for the regularization parameters γ_0, γ_1.

2.3. HMMR Parameter Estimation

Having fixed $\mathbf{g}_0, \mathbf{g}_1$, and A, we proceed to estimate the new parameter set $\boldsymbol{\theta} = (\pi, \sigma_0, \sigma_1)$ via a maximum-likelihood (ML) method. In ML estimation, a log-likelihood function is maximized as following:

$$\mathcal{L}(\boldsymbol{\theta}) = \log p(r_1, \ldots, r_n; \boldsymbol{\theta}) \tag{3}$$
$$= \log \sum_{\mathbf{z}} p(z'_1; \pi) \prod_{j=1}^{k} p(z'_j \mid z'_{j-1}; A) \prod_{i=1}^{n} \mathcal{N}(r_i; \mathbf{g}_{z_i}^T \mathbf{e}_i, \sigma_{z_i})$$

This function is not convex and cannot be maximized directly. However a local maximum can be found using the well-known Baum-Welch algorithm which is an instantiation of the more general Expectation-Maximization (EM) algorithm [15]. The procedure seeks to find the MLE of the marginal likelihood by alternating between the two following steps:

1. *E-step*: This step involves the computation of the complete-data log-likelihood given the observed neural data, the target signal, and the current estimation of the parameter set $\boldsymbol{\theta}^{(q)}$ expressed as:

$$Q(\boldsymbol{\theta}, \boldsymbol{\theta}^{(q)}) = \mathbb{E}\Big[\log p(\mathbf{r}, \mathbf{z} \,|\, \mathbf{e}; \boldsymbol{\theta}) \,|\, \mathbf{r}, \mathbf{e}; \boldsymbol{\theta}^{(q)} \Big] \tag{4}$$

It turns out that the expectation can be easily derived by the computation of the posterior probabilities of the state at the j-th time interval given the whole observation data and the current estimation of the parameter set $\boldsymbol{\theta}^{(q)}$ mathematically written as:

$$\zeta_l^{(q)}(j) = p(z'_j = l \mid \mathbf{r}, \mathbf{e}; \boldsymbol{\theta}^{(q)}) \quad (j = 1, \ldots, k),\ l \in \{0, 1\} \tag{5}$$

By applying Bayesian rule these posterior probabilities can be rewritten as:

$$\zeta_l(j) = \frac{\alpha_l(j)\beta_l(j)}{\sum_{j=1}^{k} \alpha_l(j)\beta_l(j)}, \tag{6}$$

in which:

$$\alpha_l(j) = p(r_1, r_2, \ldots, r_{jL}, z'_j = l \mid \mathbf{e}; \boldsymbol{\theta}) \tag{7}$$

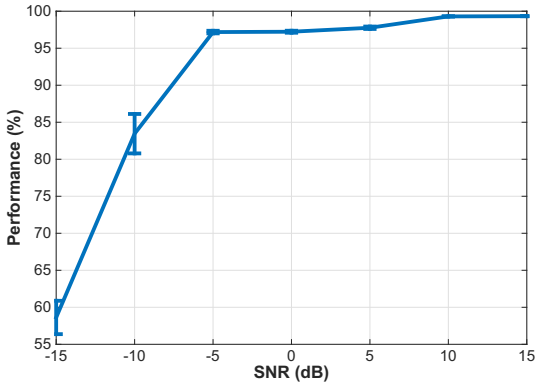

Figure 1: *Robustness against background distortion at different SNR levels tested on synthetic data.*

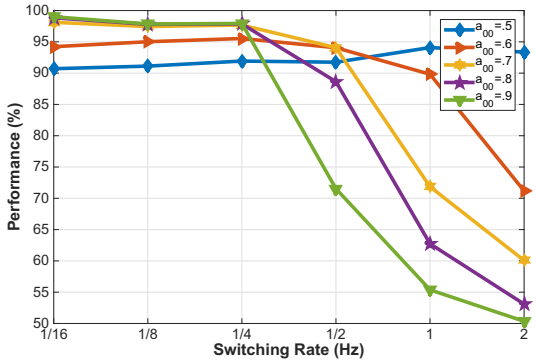

Figure 2: *Simulation results on varying the self-transition probability and its effect on average performance for various state switching rates.*

and:

$$\beta_l(j) = p(r_{jL+1}, \ldots, r_n \mid z'_j = l, \mathbf{e}; \boldsymbol{\theta}) \tag{8}$$

The Baum-Welsh algorithm efficiently computes these posterior probabilities by computing α_{jl}'s and β_{jl}'s through a recursive procedure known as forward-backward algorithm [16], as used in standard HMMs.

2. *M-step*: In this step, the parameters are updated to the values that maximize the conditional expectation in (4) with respect to $\boldsymbol{\theta}$ as in:

$$\boldsymbol{\theta}^{(q+1)} = \arg\min_{\boldsymbol{\theta}} Q(\boldsymbol{\theta}, \boldsymbol{\theta}^{(q)}) \tag{9}$$

It can be shown that the following update rules are derived by solving the above optimization problem for HMMR. Initial state probabilities are updated as following:

$$\pi_l^{(q+1)} = \zeta_l^{(q)}(1) \tag{10}$$

The update rule for the variances can also be derived as a:

$$\sigma_l^{(q+1)} = \frac{1}{L\sum_{j=1}^{k}\zeta_l^{(q)}(j)} \sum_{j=1}^{k} \zeta_l^{(q)}(j) \sum_{i=(j-1)L+1}^{jL} (r_i - \mathbf{g}_l^T \mathbf{e}_i)^2$$

Once the model parameters are all estimated, the most likely sequence of attentional states is found using the Viterbi decoding algorithm [17]. Alternatively, the attended state posterior probabilities, $\zeta_1(j)$, can also be used as a measure of auditory attention with respect to the target.

3. Experiments and Results

We conducted a series of experiments with synthetic and real data to assess the performance of our proposed method. We ran simulations on synthetic data to evaluate the noise robustness and tracking capability of the decoder. We also tested the method with real M-EEG data recorded from 3 subjects. In all of our experiments, we evaluated performance by finding the percentage of correct Viterbi decoder outputs.

3.1. Simulation Results

In the first experiment, we used the forward model in (1) to generate a simulated neural response to two simultaneous speech streams of the same intensity. For this experiment, we chose a sampling frequency F_s of 120 Hz and stimulus duration of 60 s (n = 7200). To generate the response signal, we used two TRFs of length 250 ms (τ = 30) that were estimated from real data corresponding to attended and unattended states. The sequence $\{z_i\}$ periodically alternated between the two states with a total of 2 periods within the whole duration. The response was then synthesized by introducing an i.i.d. Gaussian noise to the sum of the input envelopes filtered by TRFs according to the generated state sequence. Finally, the decoding was performed with respect to the first speech sample with a temporal resolution of 200 ms (L = 24) using a symmetric transition matrix with self-transition probabilities set to 0.8. Figure 1 illustrates the average performance of the proposed method, along with the 95% confidence intervals, for tracking auditory attention in presence of background noise at different SNR levels. It is evident that the method can recover attentional states almost perfectly at SNRs as low as -5 dB. It should be noted that the SNRs here do not account for the interference corresponding to the second speech stream, such that the actual signal-to-interference-plus-noise ratios (SINR) are even lower than the values reported here.

We also evaluated the performance of the tracking capability of our approach by changing the switching rate of the attentional state z_n. For this experiment, the background SNR was set to 10 dB, and the decoder was tested for different switching rates between 1/16–2 Hz, and with self transition probabilities ranging between 0.5–0.9. With random transitions of states (a_{00}=0.5) the proposed method performs consistently well (90% accuracy) across different switching rates as depicted in figure 2. With more conservative transitions (a_{00}>0.6), steady states are tracked more accurately, however the probability of detecting changes is reduced, and therefor the performance drops for increased switching rates.

3.2. M-EEG experiment results

3.2.1. Stimuli, task, and neural response extraction

In the second experiment, we applied our proposed method to data recorded from human subjects. Three healthy, right-handed, normal-hearing subjects (ages 24–27, 1 female, 2 male) participated in the experiment, each giving informed consent according to procedures approved by University of Washington. Subjects reported no history of hearing impairment or neurological disorder. Each subject performed 40 trials, each 60 s in

length. Each trial consisted of excerpts from two audio stories narrated by two different male speakers and dichotically presented to the subject. Each subject attended to the same story throughout all 40 trials. In the first 20 trials, the target streams were presented in the subjects's right ear. The subject was then instructed to switch attention to the left ear, into which the subsequent 20 trial streams were presented. Subjects were required to answer 3 multiple-choice questions on the content of the attended story after each trial. Stimulus intensities in each audio stream within each trial were normalized to have the same root mean square (RMS), and silent gaps longer than 0.5 s in the speech streams were truncated to 0.5 s in duration.

MEG and EEG data were recorded simultaneously while subjects performed the behavioral tasks. A cortical M-EEG source space was constructed resulting in $\sim$3000 dipoles per hemisphere whose currents were estimated from the M-EEG sensor data. More details about data acquisition system, M-EEG preprocessing, and source localization can be found in [18].

To ensure that the neural signals used in the experiment were related to the auditory stimulus, we extracted activity from four different regions of interest (ROIs) in the cortex involved in auditory processing. As a control case, we also included one region from visual cortex in our evaluation. The auditory ROIs consisted of lateral, planum polare, and temporal plane of the superior temporal gyrus, and the anterior transverse temporal gyrus. The calcarine sulcus was chosen for the visual ROI. For all different combination of auditory ROIs, we used the average of corresponding dipole activities in both hemispheres as the neural response in our analyses.

For each subject, 10 trials (5 attending to right, and 5 attending to left) were randomly selected and used to estimate subject-dependent TRFs corresponding to attended and unattended states. The decoding was then performed on the remaining subset of trials separately with respect to both streams. Similar to simulations, we performed the decoding for this M-EEG experiment using a_{00} = 0.8 and L = 24 (200 ms observation windows) considering that the attentional state of the subjects was not expected to switch rapidly during the task. Initialized with different random choices of noise variance and initial state priors, the decoding was repeated 100 times for each subject-trial-target combination.

3.2.2. Results and Discussion

We tested the visual ROI and all different combinations of auditory ROIs. Table 1 shows the best average accuracy of the decoder for combinations consisting of 1, 2, 3, or 4 auditory ROIs along with the corresponding combination and 95%-confidence intervals. As expected, the activity of visual cortex (last column) did not carry any information about auditory attention and resulted in states drawn by chance. For the auditory areas, we achieved decoding accuracies ranging between 58–76% with the highest (75.9%) belonging to the combination of lateral superior and transverse temporal gyri. The degraded accuracy of areas with more ROIs included (3–4) may be indicative of lower SNR in the added regions or simply due to destructive nature of averaging in generating neural responses.

The significance of these results become more evident when compared to accuracies reported for other decoders that take advantage of knowledge about the distractor. For example in [6], the authors have reported a decoding accuracy around 60% for the median subject among 40 when using 10-s long observation windows for EEG signals (as opposed to 200-ms in our experiment). It is likely that part of the gain we achieved is due to the use of source estimation that enables an ROI approach to target relevant cortical anatomy. However, it should be noted that gathering anatomical information for M-EEG source imaging is becoming more economically viable for the general public. Another aspect of the proposed method is that knowledge of any additional source can be easily integrated with the decoder by simply modifying the input vector and the TRFs corresponding to states as following:

Table 1: *Single-trial attentional decoding accuracy using different regions of interest (ROIs) in the brain, lateral (l), planum polare (p), and temporal plane (tp) of the superior temporal gyrus; the anterior transverse temporal gyrus (tr) in auditory cortex, and the calcarine sulcus from visual cortex. Columns 1-4 show the best accuracy achieved by combining dipole activities respectively in 1-4 of the auditory ROIs. Column 5 demonstrates the performance for the control case using one visual ROI.*

1-aROI	2-aROI	3-aROI	4-aROI	vROI
72.7%$\pm$0.97 (l)	**75.9%$\pm$0.88** (l,tr)	72.9%$\pm$0.89 (l,tr,tp)	72.2% $\pm$0.89	50.5 $\pm$0.57

$$\mathbf{g}_{z_i}^{\text{new}} = \left[\mathbf{g}_{z_i}^T \ \mathbf{g}_{1-z_i}^T\right]^T, \text{and} \quad \mathbf{e}_i^{\text{new}} = \left[\mathbf{e}_i^{\text{target}\,T} \ \mathbf{e}_i^{\text{distractor}\,T}\right]^T$$

4. Conclusions

We proposed an HMMR-based approach to track the auditory attentional state of subjects. Using single-trial M-EEG recordings in conditions in which there was no information about the competing stimulus, our proposed method could successfully track attentional modulations with accuracies as high as 75.9%. It is possible that the subjects are attending to the to-be-attended stimuli only 75.9% of the time, but we hypothesize that the errors are partly due to inaccuracies in the estimation of temporal response functions and also their fluctuation over time. A possible solution for this issue is to have a model that accounts for the within-state variability of TRFs by allowing either multiple substates or states parametrized by continuous variables. For future research, a direction to investigate is continuous-space modeling of attentional modulations.

Although we applied our approach to estimated M-EEG source-domain signals, the theory can be directly applied to sensor outputs as well eliminating the need for anatomical information. Another research direction is to treat all sensor outputs together as multivariate observations, and explore how that compares with the current univariate approaches (i.e. optimizing the linear combination of the channels.)

5. Acknowledgements

This work was supported by National Institute on Deafness and Other Communication Disorders grant (R01DC013260), Air Force Office of Scientific Research Young Investigator Program Award (FA9550-12-1-0466), Office of Naval Research Young Investigator Research Program Award (N00014-15-1-2124) and Army Research Office Grant No. W911NF1210277.

6. References

[1] N. Mesgarani and E. F. Chang, "Selective cortical representation of attended speaker in multi-talker speech perception," *Nature*, vol. 485, no. 7397, pp. 233–236, 2012.

[2] H. Luo and D. Poeppel, "Phase patterns of neuronal responses reliably discriminate speech in human auditory cortex," *Neuron*, vol. 54, no. 6, pp. 1001–1010, 2007.

[3] S. J. Aiken and T. W. Picton, "Human cortical responses to the speech envelope," *Ear Hear.*, vol. 29, no. 2, pp. 139–157, 2008.

[4] N. Ding and J. Z. Simon, "Neural coding of continuous speech in auditory cortex during monaural and dichotic listening," *J. Neurophysiol.*, vol. 107, no. 1, pp. 78–89, 2012.

[5] N. Ding and J. Z. Simon, "Emergence of neural encoding of auditory objects while listening to competing speakers," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 109, no. 29, pp. 11 854–11 859, 2012.

[6] J. A. O'Sullivan, A. J. Power, N. Mesgarani, S. Rajaram, J. J. Foxe, B. G. Shinn-Cunningham, M. Slaney, S. A. Shamma, and E. C. Lalor, "Attentional selection in a cocktail party environment can be decoded from single-trial EEG," *Cereb. Cortex*, vol. 25, pp. 1697–1706, 2015.

[7] S. Akram, J. Z. Simon, S. A. Shamma, and B. Babadi, "A state-space model for decoding auditory attentional modulation from meg in a competing-speaker environment," in *Advances in Neural Information Processing Systems*, 2014, pp. 460–468.

[8] M. Wu, D. Wang, and G. J. Brown, "A multipitch tracking algorithm for noisy speech," *Speech and Audio Processing, IEEE Transactions on*, vol. 11, no. 3, pp. 229–241, 2003.

[9] J. Sohn, N. S. Kim, and W. Sung, "A statistical model-based voice activity detection," *IEEE Signal Process. Lett.*, vol. 6, no. 1, pp. 1–3, Jan. 1999.

[10] E. Larson and A. K. C. Lee, "Influence of preparation time and pitch separation in switching of auditory attention between streams," *J. Acoust. Soc. Am.*, vol. 134, no. 2, pp. EL165–71, 2013.

[11] E. Larson and A. K. C. Lee, "Switching auditory attention using spatial and non-spatial features recruits different cortical networks," *Neuroimage*, vol. 84, pp. 681–687, 2014.

[12] V. L. Brailovsky and Y. Kempner, "Application of piece-wise regression to detecting internal structure of signal," *Pattern Recognit.*, vol. 25, no. 11, pp. 1361–1370, 1992.

[13] M. Fridman, "Hidden markov model regression," Ph.D. dissertation, Graduate School of Arts and Sciences, University of Pennsylvania, 1993.

[14] S.-J. Kim, K. Koh, M. Lustig, S. Boyd, and D. Gorinevsky, "An interior-point method for large-scale ℓ_1-regularized least squares," *Selected Topics in Signal Processing, IEEE Journal of*, vol. 1, no. 4, pp. 606–617, 2007.

[15] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *J. R. Stat. Soc. Series B Stat. Methodol.*, vol. 39, no. 1, pp. 1–38, 1977.

[16] L. Rabiner, "A tutorial on hidden markov models and selected applications in speech recognition," *Proc. IEEE*, vol. 77, no. 2, pp. 257–286, 1989.

[17] A. J. Viterbi, "Error bounds for convolutional codes and an asymptotically optimum decoding algorithm," *IEEE Trans. Inf. Theory*, vol. 13, no. 2, pp. 260–269, 1967.

[18] A. K. C. Lee, E. Larson, and R. K. Maddox, "Mapping cortical dynamics using simultaneous meg/eeg and anatomically-constrained minimum-norm estimates: an auditory attention example," *J. Vis. Exp*, vol. 68, pp. e4262–e4262, 2012.