34th Annual International Conference of the IEEE EMBS
San Diego, California USA, 28 August - 1 September, 2012

A Spatially-Regularized Dynamic Source Localization Algorithm for EEG

E. Pirondini*, B. Babadi, C. Lamus, E. N. Brown, *Fellow, IEEE*, and P. L. Purdon, *Member, IEEE**

***Abstract*— Cortical activity can be estimated from electroencephalogram (EEG) or magnetoencephalogram (MEG) data by solving an ill-conditioned inverse problem that is regularized using neuroanatomical, computational, and dynamic constraints. Recent methods have incorporated spatio-temporal dynamics into the inverse problem framework. In this approach, spatio-temporal interactions between neighboring sources enforce a form of spatial smoothing that enhances source localization quality. However, spatial smoothing could also occur by way of correlations within the state noise process that drives the underlying dynamic model. Estimating the spatial covariance structure of this state noise is challenging, particularly in EEG and MEG data where the number of underlying sources is far greater than the number of sensors. However, the EEG/MEG data are sparse compared to the large number of sources, and thus sparse constraints could be used to simplify the form of the state noise spatial covariance. In this work, we introduce an empirically tailored basis to represent the spatial covariance structure within the state noise processes of a cortical dynamic model for EEG source localization. We augment the method presented in Lamus, et al. (2011) to allow for sparsity enforcing priors on the covariance parameters. Simulation studies as well as analysis of real data reveal significant gains in the source localization performance over existing algorithms.**

I. Introduction

Among the different non-invasive techniques that can be used to study the functionality of the human brain in healthy and diseased subjects, Magnetoencephalography (MEG) and Electroencephalography (EEG) stand out due to their high temporal resolution. They can observe dynamics on a time-scale of milliseconds, reflecting synchronous activity most likely from pyramidal cells within the cerebral cortex. MEG and EEG can therefore provide a dynamic characterization of brain activity that is not possible with slower indirect modalities such as functional magnetic resonance imaging (fMRI) or positron emission tomography (PET) [1]. EEG/MEG source localization requires solution of an ill-conditioned neuromagnetic inverse problem where neuroanatomical, computational, and dynamic constraints are used to arrive at a unique solution. Recent methods have employed spatiotemporal dynamic models based on neurophysiological knowledge, resulting in improved estimation efficiency and spatial localization compared to static methods [2]–[4]. The accuracy of these dynamic methods is strongly influenced by how well the dynamics of the underlying brain current sources is approximated by the state-space model [5].

In this work, we aim to extend the MEG source localization framework developed in [2] to cover EEG. Due to the lower number of sensors typically recorded in EEG compared to MEG, as well as other differences in the inherent characteristics of these two modalities of recording, such an extension presents several challenges. In particular, the characterization of the observation noise model for EEG requires care, whereas in MEG, the observation noise can be well modeled using empty room measurements [2]. Moreover, the assumption of spatially-independent state noise components used in [2] for MEG source localization may be less appropriate for EEG due to its lower spatial resolution.

In order to address these issues, we propose an empirically tailored basis to model the spatial covariance of the state noise processes. The resulting state-space model uses the topology of the source space to allow spatial correlations in the state noise process within local neighborhoods. We incorporate sparsity-enforcing priors into the estimation framework to account for sparsity in the spatial covariance structure. Simulation studies, as well as analysis of real EEG data, confirm significant improvements over traditional methods for EEG source localization.

II. Methods

A. State-Space Model for EEG

EEG recordings provide a measure of the electric field generated by the cortical pyramidal cells at the locations of the sensors distributed on the scalp. Let $y_{i,t}$ denote the signals recorded by sensor i at time t for all $i = 1,2,...,N_y$ and $t = 1,2,...,T$. Let $y_t := [y_{1,t}, y_{2,t}, ..., y_{N_y,t}]'$ be the $N_y \times 1$ vector of observations at time t. The EEG signal arises from the primary and secondary current distributions in the brain, which can be well approximated by a collection of current dipoles distributed perpendicularly on the cortical surface [1]. Let $x_{i,t}$ denote the source amplitude of the dipole i at time t for all $i = 1,2,...,N_x$ and $t = 1,2,...,T$, and let $x_t := [x_{1,t}, x_{2,t}, ..., x_{N_x,t}]'$ be the $N_x \times 1$ vector of dipole sources representing the cortical activity at time t. A typical value of N_y is in the range of 32$\sim$256, whereas N_x can be as high as $\sim$300000. The observation process can be modeled as:

$$y_t = Gx_t + v_t, \tag{1}$$

where G is an $N_y \times N_x$ matrix representing the sensing process and v_t is the $N_y \times 1$ vector of background and

This work was supported by NIH grants DP2-OD006454 (PLP), K25-NS057580 (PLP), DP1-OD003646 (ENB), and R01-EB006385 (ENB).

EP, BB, CL, ENB, and PLP are with the Department of Anesthesia, Critical Care, and Pain Medicine, Massachusetts General Hospital, Boston, MA, and with the Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, Cambridge, MA

ENB is also with Harvard-MIT Division of Health, Science, and Technology, Cambridge, MA

EP is also with SV-EPFL, Lausanne, Switzerland

* Corresponding Authors; email elvira.pirondini@epfl.ch and patrickp@nmr.mgh.harvard.edu

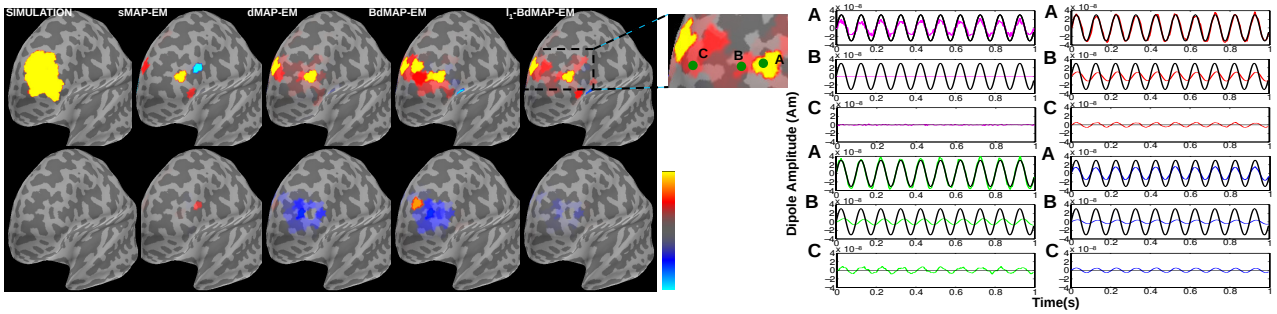

Fig. 1. Pre-frontal simulation results for sMAP-EM (2nd column), dMAP-EM (3rd column), BdMAP-EM (4th column), ℓ_1-BdMAP-EM (5th column) methods at two different time instants: a peak (upper panels) and a zero (bottom panels) of the sinusoidal wave. The first left column shows the simulated activity. The color-bar's maximum (bright yellow) and minimum (bright blue) correspond to ± 0.80 nAm for all methods. The right section of the figure shows a zoomed-in view of the cortical activity obtained in the ℓ_1-BdMAP-EM method, where green dots are representative dipoles labeled A (inside the simulated area), B (inside the simulated area), and C (outside the simulated area). The upper left panels in magenta show the estimated time course of the sMAP-EM method. The bottom left panels in green show the dMAP-EM estimates. The upper right panels in red show the estimated time trace of the BdMAP-EM method and the estimates of the ℓ_1-BdMAP-EM method are reported in the bottom right panels in blue. In all panels the black line represent the time course of the simulated activity.

instrumental noise. The matrix G is a linear map from the neuronal currents to the electric signal recorded by the sensors [6]. The noise vector v_t captures all the uncertainties independent of x_t and can be modeled by a Gaussian random vector of mean zero and covariance matrix C.

The dipole activity can be represented by a spatio-temporal dynamic process based on the underlying neurophysiology [2], as follows:

$$x_{n,t} = \lambda [a_n x_{n,t-1} + (1-a_n) \sum_{i \in N_1(n)} d_{n,i} x_{i,t-1}] + w_{n,t}, \qquad (2)$$

where $N_1(n)$ denotes the set of the nearest neighbors of the nth dipole source, $d_{n,i}$ is the normalized inverse distance between dipoles n and i, $a_n \in [0,1]$ is a weighting factor, λ is a positive scalar ensuring the stability of the autoregressive model, and w_t is the source regeneration process. Equation (2) can be expressed in the following compact form:

$$x_t = F x_{t-1} + w_t, \qquad (3)$$

where the $N_x \times N_x$ matrix F captures the spatial portion of the spatio-temporal autoregressive model. The state noise process modeled by $w_t := [w_{1,t}, w_{2,t}, ..., w_{N_x,t}]'$ represents the temporal innovations of the sources beyond those that can be accounted from local cortical interactions. This process can be approximated by a Gaussian process with covariance matrix Q_t, independent of the measurement noise v_t. In [2] the covariance Q_t was restricted to a diagonal structure. We extend the structure of Q_t by allowing off-diagonal elements in order to capture the spatial dependency in the source regeneration process. Let B an $N_x \times N_x$ matrix with elements:

$$(B)_{n,i} = \begin{cases} 1 & \text{if} \quad n = i \\ \delta_1 d_{n,i} & \text{if} \quad i \in N_1(n) \\ \delta_2 d_{n,i} & \text{if} \quad i \in N_2(n) \\ 0 & \text{otherwise} \end{cases}, \qquad (4)$$

where $N_1(n)$ and $N_2(n)$ denote the set of first and second nearest neighbors of the dipole n, respectively, and δ_1 and δ_2 are positive scaling constant. Let $\{b_1, b_2, ..., b_{N_x}\}$ denote the set of eigen-vectors of B, and let $\{q_1, q_2, ..., q_{N_x}\}$ be the corresponding orthonormal basis spanning the range of B. We consider the class of covariance matrices Q_t that lie in the span of this basis, *i.e.*,

$$Q = \sum_{n=1}^{N_x} \theta_{n,t} q_n q_n', \qquad (5)$$

where $\theta_{n,t}$ are the expansion coefficients.

Suitable priors must be chosen for the parameters $\theta_{n,t}$ in this model. One choice would be to use the conjugate density of the observation likelihood, which in this case is given by the inverse gamma density [2]:

$$p(\theta) = \prod_{n=1}^{N_x} \frac{\beta^\alpha}{\Gamma(\alpha)} \left(\frac{1}{\theta_n}\right)^{a+1} \exp\left(\frac{-\beta}{\theta_n}\right). \qquad (6)$$

The parameters α and β are tuned so that the resulting prior matches the scale of the observation data while remaining non-informative. Alternatively, if we assume that the covariance matrix Q_t has a compact representation in the basis $\{q_1, q_2, ..., q_{N_x}\}$, we can also assume a sparsity enforcing prior on the expansion coefficients $\theta_{n,t}$ [7]. One such prior is given by the Laplacian distribution density:

$$p(\theta) = \gamma^{N_x} \prod_{n=1}^{N_x} \exp(-\gamma \theta_n), \qquad (7)$$

where the parameter γ is tuned in order to maintain the appropriate scaling of the expansion coefficients with the observation data.

B. Inference with Maximum a Posteriori Expectation-Maximization algorithm

Although all the parameters in the state-space and observation model have uncertainties, some of them can be approximated by their fixed a priori estimates based on neurophysiological knowledge. First of all, the lead field matrix G can be computed using a quasi-static approximate solution to the Maxwell's equations [6]. Secondly, the

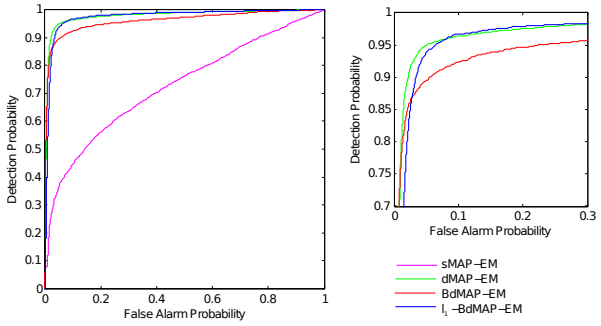

Fig. 2. ROC curves from simulated activity in the pre-frontal area computed by determining the relationship between the detection probability and the false alarm probability at varying thresholds for the null hypothesis. Therefore, the detection probability is given by the fraction of events where the active source was correctly detected for a given threshold. Similarly, the estimate of false alarm probability is given by the proportion of events falsely detected, again for a given threshold (See [2] for details). In the right panel a zoomed-in view of the right corner of the ROC curves, with detection probability 0.7-1 and false alarm probability 0-0.3.

measurement noise covariance C can be reliably estimated from pre-stimulus baseline EEG recordings. The expansion coefficients $\theta_t = [\theta_1, \theta_2, ..., \theta_T]'$ with $\theta_i := [\theta_{1,i}, \theta_{2,i}, ..., \theta_{N_x,i}]'$ for all $i = 1, 2, ..., T$, are the unknown model parameters that can be obtained by a Maximum a Posteriori (MAP) estimator $\hat{\theta}_{t,\mathrm{MAP}} := \arg\max_{\theta_t} p\left(\theta_t | \{y_t\}_{t=1}^T\right)$, where $p(\theta_t | \{y_t\}_{t=1}^T)$ is the posterior density of the parameters θ_t conditioned on the full set of measurements $\{y_t\}_{t=1}^T$. The empirical Bayesian estimate of the source amplitudes, *i.e.,* the conditional mean of the state vector given the full set of measurements $\{y_t\}_{t=1}^T$ and an estimate $\hat{\theta}_{t,\mathrm{MAP}}$ of the model parameters, is denoted by $x_{t|T} := E(x_t | \{y_t\}_{t=1}^T, \hat{\theta}_{t,\mathrm{MAP}})$, where the notation in the subscript of $x_{t|T}$ indicates that the conditioning at time t is on the full set of measurements from time 1 to T. We use the Expectation-Maximization (EM) algorithm to obtain Maximum a Posteriori estimates [8]. The E-step is carried out using the Kalman Filter [9] and Fixed Interval Smoother algorithms [10], which also provide the empirical Bayesian estimate of the source amplitudes at the final EM iteration.

It is easy to show that the M-step of the EM algorithm results in the following update equations for the model parameters,

$$\theta_{n,t}^{(r)} = \frac{\mathrm{tr}\left\{q_n A^{(r)} q_n'\right\} + 2\beta}{T + (2\alpha + 1)}, \quad \text{and} \tag{8}$$

$$\theta_{n,t}^{(r)} = \frac{\sqrt{T^2 + 8\mathrm{tr}\left\{q_n A^{(r)} q_n'\right\}\gamma} - T}{4\gamma}, \tag{9}$$

respectively for the conjugate prior (8) and the Laplacian prior (9), where $A^{(r)}$ is a matrix that can be explicitly calculated using the covariance estimates obtained by the Kalman filter and the Finite Interval Smoother (See [2] for details). The iterations are repeated until the logarithm of the posterior density reaches an asymptote.

C. Simulation Studies and initialization of the parameters

We analyzed simulated data, as well as real data recorded from a human subject, to compare the performance of four different source localization methods: the static MAP-EM (sMAP-EM) [2], the dynamic MAP-EM (dMAP-EM) [2], the dMAP-EM with the spatial covariance basis (BdMAP-EM), and the BdMAP-EM with the Laplacian prior (ℓ_1-BdMAP-EM). For the simulated data, we chose an active region in the pre-frontal area of the left hemisphere. We simulated cortical activity on a density sampled source space of $\sim$300000 dipoles. The time course of the sources in the active region was fixed at a 10 Hz sinusoidal oscillation over a period of 1 second with a sampling frequency of 200 Hz. The lead field matrix was computed with the MNE software [6] using a realistic 3-layer boundary element method (BEM) from the MRI of a human subject. The observation model given by (1) was used to obtain the simulated EEG recordings on 64 sensors. The measurement noise was additive Gaussian with covariance empirically estimated using the MNE software [6] from real EEG data. The simulated dipole amplitudes were scaled uniformly across the active region to achieve a signal-to-noise-ratio (SNR) of 3. This value was chosen to be consistent with empirical estimates of SNR from real data. We employed a source space of $N_x = 1282$ dipoles perpendicular to the cortical surface, with an average spacing of 1.25 cm. However, the model can be easily extended to finer source spaces with unconstrained source orientations. The active region was composed of $\sim$3000 dipoles in the densely-sampled source space. The parameters of the matrix F were set to $a_n = 0.7$ and $\lambda = 0.95$ to improve stability. Finally, the parameters of the covariance basis were set to $\delta_1 = 0.5$ and $\delta_2 = 0.25$.

For the analysis of the real data, we applied the sMAP-EM, dMAP-EM, BdMAP-EM, and ℓ_1-BdMAP-EM to alpha-rhythms elicited during an eyes-open/eyes-closed task recorded using a 64-lead EEG cap from a human subject. The alpha-rhythm is a 10 Hz oscillation originating from the occipital lobe during wakeful relaxation with eyes-closed. The subject was instructed to open and close his or her eyes periodically over a few minutes time. The data were recorded at a sampling frequency of 5 kHz, down-sampled to 200 Hz and filtered in the band of 0.1–50 Hz off-line.

III. Results

Figure 1 shows the spatial distribution of the simulated activity in the pre-frontal area at two different time instants: the peak and the zero of the sinusoidal wave respectively in the upper and in the lower panels. The sMAP-EM (2nd column from the left) produces an estimate that appears more focal than the true active region. The dMAP-EM (3th column) yields a dipole source estimate with a spatial spread comparable to that of the true active region. The BdMAP-EM (4th column) improves over the dMAP-EM in terms of the spatial spread. Finally, the ℓ_1-BdMAP-EM (5th column), which employs the sparsity enforcing prior, provides an estimate with a spatial spread comparable to the dMAP-EM. Interestingly, the ℓ_1-BdMAP-EM is the only method which captures the absence of activity at the zero of the sinusoidal wave (5th column lower panel).

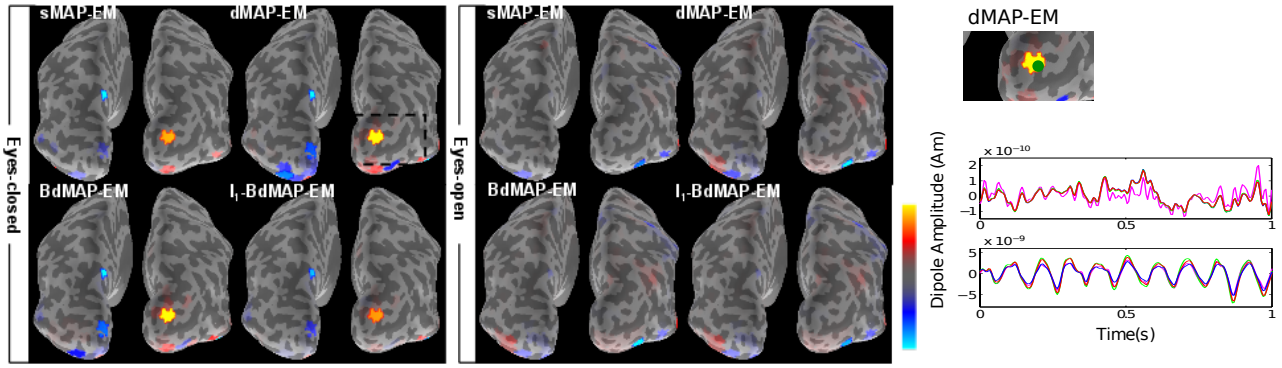

Fig. 3. Analysis of human EEG alpha-rhythm. The color-bar's maximum (bright yellow) and minimum (bright blue) correspond to ± 0.20 nAm for the eyes-closed (left panels) and to ± 0.03 nAm for eyes-open (right panels). The left part of the figure shows a zoomed-in view of the cortical activity obtained in the dMAP-EM method (eyes-closed), where the green dot is a representative dipole of the alpha-rhythm. The time course of the estimations at the green dot is shown for eyes-closed (bottom panel) and eyes-open (upper panel). The sMAP-EM is shown in magenta, the dMAP-EM in green, the BdMAP-EM in red, and the ℓ_1-BdMAP-EM in blue. The units are different for the two panels (1e-10 Am for the upper panel and 1e-09 Am for the bottom one).

The figure also shows the time series estimates obtained by the four methods at a few different dipoles (two inside and one outside the active region). The sMAP-EM provides the best tracking performance outside the true region, but performs poorly inside the area of activity. The other three methods track the time course within the active region with higher precision, but tend to overestimate the activity outside. Among the three dynamic methods, the ℓ_1-BdMAP-EM suppresses the estimate outside the active region the most. The correspondence between the simulated sources and their estimates are evaluated by the root mean square error (RMSE) performance of each method inside and outside the active region. The sMAP-EM exhibits the poorest error performance with an average RMSE of 20.1 nAm inside the active region. The average RMSE of the estimates obtained by the dMAP-EM, BdMAP-EM, and ℓ_1-BdMAP-EM inside the active area are given by 19.5 nAm, 18.2 nAm, and 19.8 nAm respectively. On the other hand, the ℓ_1-BdMAP-EM performs the best outside the active region with an average RMSE 100 pAm, which is consistent with the observed time tracings. The average RMSE for the sMAP-EM, dMAP-EM, and BdMAP-EM are 105 pAm, 148 pAm, and 150 pAm, respectively.

In order to evaluate the sensitivity/specificity tradeoff of the source localization methods, we calculated the receiver operating characteristic (ROC) curve as is shown in Fig.2. The sMAP-EM (magenta) exhibits the lowest initial slope and smallest area-under-the-curve (AUC), requiring a high false alarm rate to achieve a reasonable detection rate. The ROC curves of the dMAP-EM (green), BdMAP-EM (red), and ℓ_1-BdMAP-EM (blue), however, indicate that these three methods outperform the sMAP-EM. In particular, the dMAP-EM and the ℓ_1-BdMAP-EM provides the largest AUC (about 0.97) as well as the highest initial slope.

The performance of the four algorithms on the real data is summarized in Fig.3. The overall performance of the dynamic algorithms in recovering the alpha waves in the occipital lobe during the eyes-closed/eyes-open phase are better than that of the sMAP-EM. In particular the ℓ_1-BdMAP-EM suppressed the amplitude of the estimate the most during the eyes-open phase (*i.e.,* where alpha activity should be absent).

IV. Conclusion

We have proposed a novel state-space model for cortical activity that incorporates the local spatial correlations within an underlying state noise process. A sparsity enforcing prior was used to estimate the model parameters for EEG source localization. Simulation studies as well as the analysis of real data reveal significant performance gains provided by this algorithm in terms of source localization quality over the existing algorithms. Generalizations to broader classes of source spaces and models, as well as improvements in the computational complexity, are currently under study.

References

[1] S. Baillet, J. C. Mosher, and R. M. Leahy, "Electromagnetic brain mapping," *IEEE SIGNAL PROCESSING MAGAZINE*, vol. 18, no. 6, pp. 14–30, 2001.

[2] C. Lamus, M. S. Hämäläinen, S. Temereanca, E. N. Brown, and P. L. Purdon, "A spatiotemporal dynamic distributed solution to the meg inverse problem." *Neuroimage*, Nov 2011.

[3] C. Long, P. Purdon, S. Temereanca, N. Desai, M. Hämäläinen, and E. Brown, "State-space solutions to the dynamic magnetoencephalography inverse problem using high performance computing," *The annals of applied statistics*, vol. 5, no. 2B, p. 1207, 2011.

[4] A. Galka, O. Yamashita, T. Ozaki, R. Biscay, and P. Valdés-Sosa, "A solution to the dynamical inverse problem of eeg generation using spatiotemporal kalman filtering," *NeuroImage*, vol. 23, no. 2, pp. 435–453, 2004.

[5] C. Lamus, C. J. Long, M. S. Hämäläinen, E. N. Brown, and P. L. Purdon, "Parameter estimation and dynamic source localization for the magnetoencephalography (MEG) inverse problem." *Proc IEEE Int Symp Biomed Imaging*, vol. 2007, pp. 1092–1095, May 2007.

[6] M. Hämäläinen and J. Sarvas, "Realistic conductivity geometry model of the human head for interpretation of neuromagnetic data," *IEEE Trans Biomed Eng*, vol. 36, no. 2, pp. 165–171, 1989.

[7] A. Bruckstein, D. Donoho, and M. Elad, "From sparse solutions of systems of equations to sparse modeling of signals and images," *SIAM review*, vol. 51, no. 1, p. 34, 2009.

[8] A. Dempster, N. Laird, and D. Rubin, "Maximum likelihood from incomplete data via the em algorithm," *Journal of the Royal Statistical Society. Series B (Methodological)*, pp. 1–38, 1977.

[9] R. Kalman, "A new approach to linear filtering and prediction problems," *Journal of basic Engineering*, vol. 82, no. 1, pp. 35–45, 1960.

[10] H. Rauch, F. Tung, and C. Striebel, "Maximum likelihood estimates of linear dynamic systems," *AIAA journal*, no. 8, pp. 1445–1450, 1965.