8th International IEEE EMBS Conference on Neural Engineering
Shanghai, China, May 25 - 28 , 2017

Optimal Bandpower Estimation and Tracking via Kalman Filtering for Real-Time Brain-Computer Interfaces*

J. Gruenwald[1,3], C. Kapeller[1,3], K. Kamada[2], J. Scharinger[3] and C. Guger[1,4]

***Abstract*— Brain waves contain fundamental information about cortical activity: signal power within certain frequency bands, which is exploited by a variety of Brain-Computer Interface applications. For real-time systems, these features must be estimated as quickly as possible while maintaining high signal fidelity. Here, we present a statistically optimal signal processing framework for real-time bandpower estimation and tracking. Key components are a spectral shaping stage for increased sensitivity and Kalman filtering of log-transformed bandpower estimates for optimal tracking. The system has one degree of freedom, which allows for adaptive design based on signal dynamics. The overall complexity remains low.**

We evaluated the proposed architecture based on two experiments involving cortical motor functions and receptive-language related cortical areas. First results are promising. Spectral shaping based on a whitening transform increases the sensitivity (z-Score) by up to 60 %. Furthermore, the tracking time lag is substantially reduced relative to conventional approaches.

I. INTRODUCTION

A Brain-Computer Interface (BCI) allows a person to interact with the environment via thoughts alone [1], [2]. BCIs rely on information extracted from brain waves, which are acquired either via the electroencephalogram (EEG, for non-invasive, largely research applications) or the electrocorticogram (ECoG, invasive, largely clinical applications). It has been shown that changes in power within respective frequency bands are directly related to cortical activity and therefore provide essential information about ongoing processes in the brain. This is called *event-related (de-)synchronization* [3] and underlies many BCI applications, such as neuroprosthesis control [4], stroke rehabilitation [5], or communication with severely disabled patients [6]. Furthermore, power changes in higher frequency bands (Gamma and High-Gamma Band, above 32 Hz) are highly location-specific [7]. This allows associating body functions with corresponding cortical areas, a concept termed *functional mapping* which is essential in neurosurgery [8].

An effective and widely used method of bandpower estimation is the computation of the variance of a signal that has been appropriately bandpass-filtered. Each variance estimate is obtained from data within a certain observation window, assuming that the underlying power does not change within this window. Consecutive samples of noisy estimates are commonly smoothed over time via moving-average filters.

The Kalman filter (KF) is widely used in tracking, forecasting, and smoothing signal trajectories embedded in noise [9], [10]. One reason is the low complexity of the filter. More importantly, if signal and noise are jointly Gaussian, it is optimal in the minimum-mean-squared error (MMSE) sense. If not, it approaches the linear MMSE, i.e., it is the best *linear* filter that can be employed to address the tracking problem. One caveat, however, is the assumption that the data follow a certain signal model. Its parameters must be carefully chosen, since this can greatly affect performance.

II. METHODS

Here, we outline the context, define the estimation problem, and present the proposed signal path as shown in Fig. 1.

A. Problem Statement

We assume a discrete-time ECoG recording $y[m]$, sampled at a given frequency f_s. It has been detrended and we ignore power-line interference phenomena for now. We specify the bounds of our frequency band of interest by ω_{lo} and ω_{hi} with $\omega = 2\pi f/f_s$. Within this band, we suppose that an *activation signal* $\check{s}[m] \in \mathbb{R}^+$ modulates the power spectral density (PSD) of $y[m]$, denoted by $P_{yy}(\mathrm{e}^{\jmath\omega})[m]$, as follows:

$$P_{yy}(\mathrm{e}^{\jmath\omega})[m]\big|_{\omega\in[\omega_{\mathrm{lo}},\omega_{\mathrm{hi}}]} = P_{yy}^{(0)}(\mathrm{e}^{\jmath\omega}) + W(\mathrm{e}^{\jmath\omega})\check{s}[m]\,. \quad (1)$$

$P_{yy}^{(0)}(\mathrm{e}^{\jmath\omega})$ refers to a known PSD that does not vary over time, and $W(\mathrm{e}^{\jmath\omega}) \in \mathbb{R}^+$ is a (generally unknown) activation pattern that applies weights to certain frequency components. We neglect the fact that it may be time-varying in reality. The activation signal $\check{s}[m]$ is assumed to be a slow process compared to f_s. Our objective is now to derive a framework for estimating and tracking $\check{s}[m]$ from observed $y[m]$ that satisfies appropriate optimality criteria.

B. Signal Preprocessing

1) Spectral Shaping: As $P_{yy}^{(0)}(\mathrm{e}^{\jmath\omega})$ in (1) approximately follows a $1/f$-curve [11], equalizing it may increase sensitivity. To this end, we model $y[m]$ as an autoregressive (AR) process of order P during time periods where $\check{s}[m]$ vanishes:

$$y[m]\big|_{\check{s}[m]=0} = \sum_{p=1}^{P} a_p y[m-p] + v[m]\,. \quad (2)$$

Here, $v[m] \sim \mathcal{N}\left(0, \sigma_v^2\right)$ are independent and identically distributed (i.i.d.) samples from a zero-mean Gaussian distribution with variance σ_v^2. The coefficients a_p and the variance

*This work was funded by the *Eurostars RapidMaps* Project

[1]The authors are with Guger Technologies OG, Graz, Austria (corresponding author contact: gruenwald@gtec.at)

[2]K. Kamada is with the Department for Neurosurgery, Asahikawa Medical University, Asahikawa, Japan

[3]The authors are with the Department of Computational Perception, Johannes Kepler University, Linz, Austria

[4]C. Guger is with g.tec medical engineering GmbH, Schiedlberg, Austria

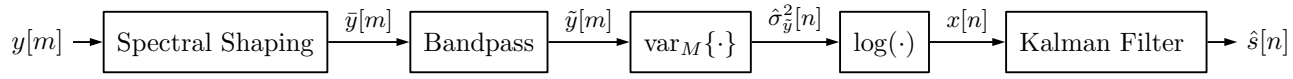

Fig. 1. Proposed signal path.

σ_v^2 can be obtained by appropriate algorithms, e.g., the *Yule-Walker Equations*. The transfer function inverting (2) is then

$$H_{\mathrm{w}}(\mathrm{e}^{\jmath\omega}) = 1 - \sum_{p=1}^{P} a_p \mathrm{e}^{-\jmath\omega p} \tag{3}$$

such that $|H_{\mathrm{w}}(\mathrm{e}^{\jmath\omega})|^2 = \sigma_v^2 / P_{yy}^{(0)}(\mathrm{e}^{\jmath\omega})$ and the filtering result $\bar{y}[m] \approx v[m]$. Hence, for arbitrary $\check{s}[m]$ we get the PSD

$$P_{\bar{y}\bar{y}}(\mathrm{e}^{\jmath\omega})[m] = \left(1 + W(\mathrm{e}^{\jmath\omega}) / P_{yy}^{(0)}(\mathrm{e}^{\jmath\omega}) \check{s}[m]\right) \sigma_v^2 \,. \tag{4}$$

Fig. 2 illustrates the effect of this transform, which is known as *whitening*. Note that with knowledge of $W(\mathrm{e}^{\jmath\omega})$, a transfer function that *truly maximizes* the sensitivity may be found.

2) Bandpass: $\tilde{y}[m]$ results from the ideal bandpass filter

$$H_{\mathrm{bp}}(\mathrm{e}^{\jmath\omega}) = \begin{cases} 1 & \omega_{\mathrm{lo}} \le |\omega| \le \omega_{\mathrm{hi}} \\ 0 & \text{else} \,. \end{cases} \tag{5}$$

3) Variance Estimation: Using Eqs. (4) and (5) and the property $2\pi\sigma_{\tilde{y}}^2[m] = \int_{-\pi}^{+\pi} P_{\tilde{y}\tilde{y}}(\mathrm{e}^{\jmath\omega})[m] \mathrm{d}\omega$, we arrive at

$$\sigma_{\tilde{y}}^2[m] / \sigma_v^2 = (1 + \rho\check{s}[m]) \Delta\omega / \pi \tag{6}$$

with $\rho = \int_{\omega_{\mathrm{lo}}}^{\omega_{\mathrm{hi}}} W(\mathrm{e}^{\jmath\omega}) / P_{yy}^{(0)}(\mathrm{e}^{\jmath\omega}) \mathrm{d}\omega / \Delta\omega$, interpreted as the sensitivity coefficient, and $\Delta\omega = \omega_{\mathrm{hi}} - \omega_{\mathrm{lo}}$. Clearly, $\check{s}[m]$ maps to the signal variance of $\tilde{y}[m]$, which can be estimated.

In practice, one variance estimate is obtained from a set of M consecutive samples. This decimates the sampling rate to f_s/M. The true variance is assumed constant within this window, which can be represented by the *tap vector* $\tilde{\mathbf{y}}[m] = [\tilde{y}[m] \; \tilde{y}[m-1] \; \cdots \; \tilde{y}[m-M+1]]^\top$. Since $\bar{y}[m]$ is regarded as white and Gaussian (cf. Section II-B.1), we infer that $\tilde{\mathbf{y}}[m] \sim \mathcal{N}(\mathbf{0}, \sigma_{\tilde{y}}^2[m]\boldsymbol{\Sigma})$, i.e., it follows a zero-mean multivariate Gaussian distribution, where $\boldsymbol{\Sigma}$ has Toeplitz structure with the autocorrelation coefficients of the bandpass-filter impulse response along its columns.

Since the mean of $\tilde{y}[m]$ is zero and known, the minimum-variance unbiased (MVU) estimator of the variance is the (non-central) second moment, which we can express as

$$\hat{\sigma}_{\tilde{y}}^2[n] = \mathrm{var}_M \left\{\tilde{y}[m]\right\} \big|_{m=nM} = \tilde{\mathbf{y}}[nM]^\top \tilde{\mathbf{y}}[nM] / M \tag{7}$$

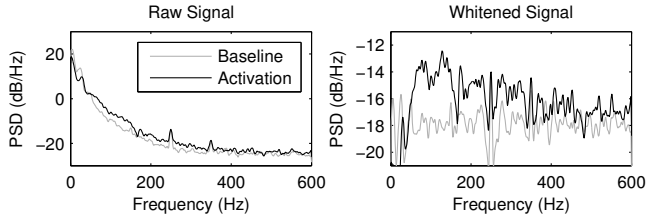

Fig. 2. Exemplary result of the whitening filter based on an AR process of order 30. Spurs in the left picture indicate interfering power-line harmonics. *Baseline*: time periods where $\check{s}[m] = 0$; *Activation*: otherwise.

We are now interested in the distribution of $\hat{\sigma}_{\tilde{y}}^2[n]$. To this end, we express $\tilde{\mathbf{y}}[nM]^\top \tilde{\mathbf{y}}[nM] = \sigma_{\tilde{y}}^2[nM] \boldsymbol{\xi}^\top \boldsymbol{\Sigma} \boldsymbol{\xi}$ with uniform $\boldsymbol{\xi} \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$. As $\boldsymbol{\Sigma}$ is positive (semi-) definite and symmetric, we can decompose it into $\boldsymbol{\Sigma} = \mathbf{P}^\top \boldsymbol{\Lambda} \mathbf{P}$, where $\boldsymbol{\Lambda} = \mathrm{diag}(\lambda_1, \ldots, \lambda_M)$ is a diagonal matrix containing the eigenvalues of $\boldsymbol{\Sigma}$, and $\mathbf{P}$ is a rotation matrix. We thus observe that $\boldsymbol{\nu} = \mathbf{P}\boldsymbol{\xi}$ is uniform as well, so $\boldsymbol{\xi}^\top \boldsymbol{\Sigma} \boldsymbol{\xi} = \boldsymbol{\nu}^\top \boldsymbol{\Lambda} \boldsymbol{\nu}$ with $\boldsymbol{\nu} \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$. Next, we define $\check{\boldsymbol{\nu}} = \left[\nu_1^2 \; \nu_2^2 \; \cdots \; \nu_M^2\right]^\top$, comprising the squared elements of $\boldsymbol{\nu}$, and $\boldsymbol{\lambda} = [\lambda_1 \; \lambda_2 \; \cdots \; \lambda_M]^\top$. Now, (7) translates into

$$\hat{\sigma}_{\tilde{y}}^2[n] = \sigma_{\tilde{y}}^2[nM] \check{\boldsymbol{\nu}}^\top \boldsymbol{\lambda} / M \,. \tag{8}$$

This formulates a mixture distribution of linearly combined, independent random variables following a chi-squared distribution with one degree of freedom, denoted by $\chi^2(1)$. Eq. (8) generalizes the standard result $M\hat{\sigma}_{\tilde{y}}^2[n] / \sigma_{\tilde{y}}^2[nM] \sim \chi^2(M)$ which holds if $\lambda_k = 1 \; \forall k$, i.e., if $\tilde{y}[m]$ were white.

4) Log Transform: From (8) follows for $x[n] = \log \hat{\sigma}_{\tilde{y}}^2[n]$:

$$x[n] = \log \sigma_{\tilde{y}}^2[nM] - \log M + \log \check{\boldsymbol{\nu}}^\top \boldsymbol{\lambda} \,. \tag{9}$$

We highlight the fact that our quantity of interest, $\log \sigma_{\tilde{y}}^2[nM]$, is now embedded in *additive, stationary* noise, which is fully characterized by the bandpass filter impulse response – its moments are thus *known*. Besides, the transform notably reduces the asymmetry of the χ^2 distribution [12]. See Section III-A for a more elaborate discussion.

C. Kalman Filter

In view of (6) and (9) and with $\mu = \mathbb{E}\{\log \check{\boldsymbol{\nu}}^\top \boldsymbol{\lambda}\}$ denoting the expected value of $\log \check{\boldsymbol{\nu}}^\top \boldsymbol{\lambda}$, we can define

$$w[n] = \log \check{\boldsymbol{\nu}}^\top \boldsymbol{\lambda} - \mu \tag{10}$$

$$s[n] = \log(1 + \rho\check{s}[nM]) + \log \frac{\sigma_v^2 \Delta\omega}{\pi M} + \mu \tag{11}$$

and cast $x[n]$ in the form of the KF *observation equation*

$$x[n] = s[n] + w[n] \tag{12}$$

where $x[n]$ is split into a signal term $s[n]$ related to the *log activation* and a noise term $w[n]$ with known moments. For optimality, the KF requires that $w[n] \sim \mathcal{N}(0, \sigma_w^2)$, i.i.d., an issue we will discuss in Section III. It further assumes that $s[n]$ follows a dynamical system model with known parameters, which in general stems from the class of AR processes (cf. Section II-B.1). We argue that the *random walk*

$$s[n] = s[n-1] + u[n] \tag{13}$$

with $u[n] \sim \mathcal{N}(0, \sigma_u^2)$, i.i.d. is appropriate since it provides an architecture that satisfies two key requirements deduced from empirical data: (1) it is a nonstationary process. (2) The parameter σ_u^2 allows for changes in the dynamics of $s[n]$.

TABLE I

RANDOM WALK KALMAN FILTER EQUATIONS [9].

Initialization:	
$s[-1\|-1] = \mathbb{E}\{s[-1]\}$ (can be set to zero)	
$M[-1\|-1] = \sigma_w^2$	
Recursion:	
$s[n\|n-1] = s[n-1\|n-1]$	Prediction
$M[n\|n-1] = M[n-1\|n-1] + \sigma_u^2$	Prediction MSE
$K[n] = M[n\|n-1]/(M[n\|n-1] + \sigma_w^2)$	Kalman Gain
$s[n\|n] = s[n\|n-1] + K[n](x[n] - s[n\|n-1])$	Correction
$M[n\|n] = (1 - K[n])\, M[n\|n-1]$	Minimum MSE
Result:	
$\hat{s}[n] = s[n\|n]$	

Given (12) and (13), the corresponding KF equations are summarized in Table I. The prediction/correction-approach is reflected by twofold indexing. Note that the KF is completely recursive, so storing large sample sets is not needed.

In many scenarios, a short time lag is allowed for better performance. In this case, the *Rauch-Tung-Striebel* (RTS) smoother [10], an extension of the KF and summarized in Table II, can be applied. Upon completion, $\hat{s}_{\text{RTS}}[n]$ holds the smoothed estimate and $M_{\text{RTS}}[n]$ holds the respective MSE.

We emphasize that one outcome of the KF is the MSE itself, which can serve as a valuable measure of realiability.

III. EXPERIMENTS AND RESULTS

Our evaluation is based on two ECoG datasets recorded at Asahikawa Medical University, Japan. The experiments consist of an alternating sequence of baseline periods, where the subject was at rest, and activation periods, during which the subject performed a task. In the *Hand Motor Task* experiment, this task was to solve a Rubik's cube, which caused activation in cortical motor areas. In the *Listening Task* experiment, the subject listened to a story, provoking activation in receptive-language related cortical areas. We show results from channels with high cortical activation.

The data were sampled at 1200 Hz, detrended, and whitened by an AR model of order 30 (trained on baseline data over 10 s). Motivated by Fig. 2 and considering disturbing harmonics of the power line, a bandwidth from 55 Hz to 245 Hz was chosen. A 6$^{\text{th}}$-order infinite impulse response (IIR) butterworth filter served as a bandpass.

TABLE II

RANDOM WALK RAUCH-TUNG-STRIEBEL SMOOTHER (LAG N) [10].

Apply KF up to sample $n+N$.
Initialization:
$M_{\text{RTS}}[n+N] = M[n+N\|n+N]$
$\hat{s}_{\text{RTS}}[n+N] = \hat{s}[n+N]$
Backward Processing for $k = \{n+N-1, n+N-2, \ldots, n\}$
$A[k] = M[k\|k]/M[k+1\|k]$
$M_{\text{RTS}}[k] = M[k\|k] - A^2[k](M[k+1\|k] - M_{\text{RTS}}[k+1])$
$\hat{s}_{\text{RTS}}[k] = s[k\|k] + A[k](\hat{s}_{\text{RTS}}[k+1] - s[k+1\|k])$

TABLE III

COMPARISON OF $\log \breve{\mathbf{v}}^\top \boldsymbol{\lambda}$ AND $\breve{\mathbf{v}}^\top \boldsymbol{\lambda}$ VS. WINDOW SIZE M.

	$\log \breve{\mathbf{v}}^\top \boldsymbol{\lambda}$				$\breve{\mathbf{v}}^\top \boldsymbol{\lambda}$	
M	Mean	Var.	Skew.	Kurt.	Skew.	Kurt.
10	2.08	0.50	−0.47	0.38	1.5	3.6
15	2.55	0.35	−0.42	0.29	1.3	2.5
20	2.87	0.26	−0.38	0.24	1.1	1.9
30	3.32	0.18	−0.33	0.18	0.90	1.3
60	4.05	0.09	−0.24	0.10	0.64	0.63
120	4.77	0.04	−0.18	0.05	0.45	0.31

A. Distribution of the Measurement Noise $w[n]$

In Table III, we provide the most important statistical descriptors of $\breve{\mathbf{v}}^\top \boldsymbol{\lambda}$ and $\log \breve{\mathbf{v}}^\top \boldsymbol{\lambda}$. Since the probability density functions (PDFs) are analytically not tractable, we obtained the values from simulation based on the aforementioned processing parameters. Mean and variance refer to μ (10) and σ_w^2, respectively. Skewness and (excess) kurtosis provide insights into Gaussianity and symmetry. In this respect, the benefit of the log transform is striking. Concluding, we argue that a reasonable choice of M sufficiently well approximates $w[n] \sim \mathcal{N}(0, \sigma_w^2)$, which is required for optimality of the KF.

B. Whiteness of the Measurement Noise $w[n]$

We now want to clarify if the noise $w[n]$ (10) is also white, as required by the KF. Chosing $M = 10$ to allow observations up to $f_s/M/2 = 60$ Hz, Fig. 3 shows two PSDs of experimentally obtained $x[n]$, with and without whitening. They stem from the whole Hand Motor Task experiment over 280 s, where the whitening filter was trained on a baseline segment over 10 s. The third PSD refers to a simulated $x[n]$ without activation (i.e., $\check{s}[m] = 0\ \forall m$) and ideal (thus white) $\bar{y}[m]$. Assumedly, this reduces $x[n]$ to a noise process with mean and standard deviation given by (11) and Table III. We normalized all $x[n]$ by these values and multiplied the obtained PSDs by the Nyquist rate, such that the x-axis sets the benchmark for whiteness and unit variance.

The PSD of the simulated $x[n]$ indeed aligns with the x-axis. This confirms the respective parts of our signal path and justifies the assumption that, under perfect conditions, $w[n]$ is white. Towards lower frequencies, the elevated power in the experimental PSDs accounts for the actual physiological signal, which is of course not present in the simulation.

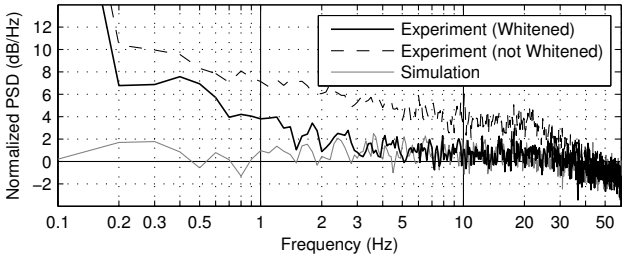

Fig. 3. Comparison of PSDs based on experiments (with and without whitening) and simulation.

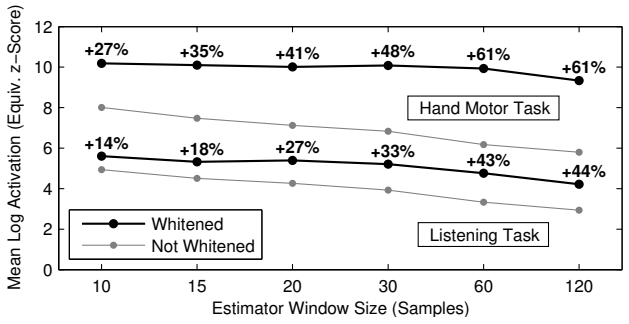

Fig. 4. Sensitivity comparison for different estimator window sizes M.

The benefits of whitening are evident. Furthermore, towards higher frequencies (where the noise $w[n]$ dominates), the experimental PSD with whitening and the PSD based on simulation match perfectly well. This supports our hypothesis that also in the experimental context and with whitening, $w[n]$ is white itself. We can thus finally conclude that the KF architecture is optimal for our tracking problem.

C. Sensitivity Increase due to Whitening

The (mean) amplitude of $x[n]$ during the activation period with respect to the baseline period and noise is a good figure of merit for our purposes. Since we compare values across different sampling rates, we introduce the normalization

$$x[n] \mapsto \sqrt{f_s/M}(x[n] - \mu_{\text{BSL}})/\sigma_{\text{BSL}} \tag{14}$$

where the transformed $x[n]$ is given in a unit we call *Equivalent z-Score*, referring to the z-Score at a fictitious sampling rate of 1 Hz. μ_{BSL} and σ_{BSL} are mean and standard deviation of $x[n]$ during baseline, respectively. In Fig. 4, a consistent improvement due to whitening is apparent.

D. Kalman Filter Performance

Since we allow for a short time lag, we use the RTS smoother to assess the capabilities of the KF. For a window size of $M = 60$, Fig. 5 shows results from two moving-average (MA) filters with time lags of 0.2 s and 0.6 s (half the window size). While the first one properly follows the signal during the activation period, it is overly susceptible to noise during baseline. The second one smooths the baseline signal well, but also removes higher frequency components in the activation phase. Since MA filters with varying window sizes are not practicable, a compromise usually must be made.

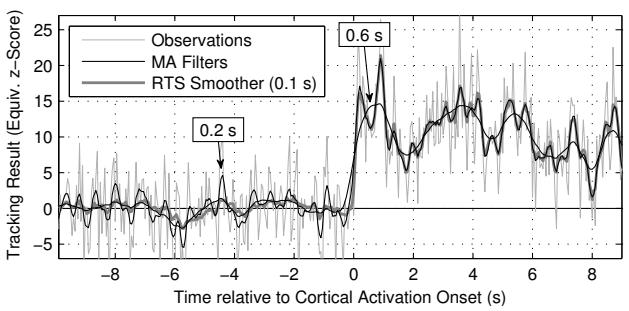

Fig. 5. Comparison of the RTS smoother (with different values of σ_u^2 before and after activation onset) and two MA filters. The values given in seconds are the respective tracking lags.

In contrast to MA approaches, the RTS smoother is capable of following the signal during both baseline and activation period. To illustrate this point, it is parameterized by a small and a large value of σ_u^2 before and after activation onset, respectively. This reveals the potential of adaptive design, which is in fact inevitable for *actual* optimality of the KF. Solutions thereto are available in the literature [13]. We further highlight that, compared to the MA filters, the RTS smoother has a substantially shorter time lag of 0.1 s.

IV. Conclusions and Future Work

We have outlined a signal processing framework that optimizes bandpower tracking for real-time BCIs. The log-transformed MVU bandpower estimates were shown to approximate a white Gaussian noise environment with known variance, where the Kalman filter attains the (linear) MMSE. We thus regard the architecture as *statistically optimal*.

First experiments have shown that the proposed method increases the sensitivity consistently and reduces the tracking lag substantially. The system complexity remains low.

Two matters may be addressed in further research: (1) the spectral shaping transform may use knowledge of the activation pattern to maximize sensitivity; and (2) the Kalman filter must be designed adaptively to *actually* approach optimality.

Since bandpower estimates are key features required by a multitude of BCI applications, the proposed architecture may significantly improve their performance.

References

[1] J. R. Wolpaw, et al., "Brain-Computer Interfaces for Communication and Control," *Clinical Neurophysiology*, vol. 113, no. 6, pp. 767791, Jun. 2002.

[2] J. R. Wolpaw and E. W. Wolpaw, *Brain-Computer Interfaces: Principles and Practice*, New York: Oxford Univ Pr, 2012.

[3] G. Pfurtscheller and F. L. Da Silva, "Event-Related EEG/MEG Synchronization and Desynchronization: Basic Principles," *Clinical Neurophysiology*, vol. 110, no. 11, pp. 1842-1857, 1999.

[4] G. Hotson et al., "Individual Finger Control of a Modular Prosthetic Limb using High-Density Electrocorticography in a Human Subject," *Journal of Neural Engineering*, vol. 13, no. 2, 2016.

[5] N. Sharma et al., "Motor Imagery: A Backdoor to the Motor System after Stroke?," *Stroke*, vol. 37, no. 7, pp. 1941-1952, Jul. 2006.

[6] C. Chatelle et al., "Brain-Computer Interfacing in Disorders of Consciousness," *Brain Injury*, vol. 26, no. 12, pp. 1510-1522, 2012.

[7] K. J. Miller et al., "Spectral Changes in Cortical Surface Potentials during Motor Movement," *Journal of Neuroscience*, vol. 27, no. 9, pp. 2424-2432, Feb. 2007.

[8] N. E. Crone et al., "Functional Mapping of Human Sensorimotor Cortex with Electrocorticographic Spectral Analysis," *Brain*, vol. 121, no. Pt 12, pp. 2271-2299, 1998.

[9] S. M. Kay, *Fundamentals of Statistical Signal Processing, volume I: estimation theory*, Upper Saddle River, NJ: Prentice Hall, 1993.

[10] S. Haykin, ed., "Kalman Filters," in *Kalman Filtering and Neural Networks*, New York: Wiley, 2001, pp. 11-16.

[11] K. J. Miller et al., "Power-Law Scaling in the Brain Surface Electric Potential," *PLoS Computational Biology*, vol. 5, no. 12, Dec. 2009.

[12] M. S. Bartlett and D. G. Kendall, "The Statistical Analysis of Variance-Heterogeneity and the Logarithmic Transformation," *Journal of the Royal Statistical Society* (suppl.), vol. 8, no. 1, pp. 128-138, 1946.

[13] R. Mehra, "On the Identification of Variances and Adaptive Kalman Filtering," *IEEE Transactions on Automatic Control*, vol. 15, no. 2, pp. 175-184, Apr. 1970.