Abstract—In this research, we have proposed a new scheme to detect and extract the activity of an unknown smooth template in presence of white Gaussian noise with unknown variance. In this regard, the problem is considered a binary hypothesis test, and it is solved employing the generalized likelihood ratio (GLR) method. GLR test uses the maximum likelihood (ML) estimation of unknown parameters under each hypothesis. The ML estimation of the desired signal yields an optimization problem with smoothness constraint which is in the form of a conventional least square error estimation problem and can be solved optimally. The proposed detection scheme is studied for P300 elicitiation from the background electroencephalography signal. In addition, to assume the P300 smoothness, two prior knowledge are considered in terms of positivity and approximate occurrence time of P300. The performance of the method is assessed on both real and synthetic datasets in different noise levels and compared to a conventional signal detection scheme without considering smoothness priors, as well as state-of-the-art linear and quadratic discriminant analysis. The results are illustrated in terms of detection probability, false alarm rate, and accuracy. The proposed method outperforms the counterparts in low-signal-to-noise ratio situations.

Index Terms—Signal activity detection, hypothesis test, smooth signal, white Gaussian noise.

I. INTRODUCTION

Detection theory is a prominent branch of statistical signal processing, which is a well established means to detect signal activity from background noise [1]–[4]. Signal activity detection (SAD) problem can be modeled by a binary hypothesis test, where in the first hypothesis, observations are considered as pure noise, and in the second one, as a mixture of noise and desired signal. Based on available information on signal and noise structure, approaches to solve the problem can be different. The optimal solution of a binary hypothesis test can be obtained by Neyman-Pearson theorem [1]–[4], if either the problem parameters are known or the probability density function (PDF) of observations has monotone likelihood ratio (MLR) property [2], [4]. If none of these two conditions is satisfied, the generalized likelihood ratio (GLR) test is usually employed to solve the problem [1]. The GLR test is a Neyman-Pearson test, for which the unknown PDF parameters under each hypothesis are found using maximum likelihood (ML) estimation. This test is a very powerful tool, which is proved to be asymptotically uniformly most powerful (UMP) among all invariant tests [5], [6].

Herein, a GLR test is proposed to detect the presence of an unknown yet smooth signal in zero-mean white Gaussian noise with unknown variance. The mixing model of signal and noise is assumed to be additive, widely used in the literature [7]. In this test, the ML estimation of the desired signal yields an optimization problem with smoothness constraint. This constraint can be modeled by a quadratic function. Therefore, the problem is reduced to a quadratic optimization problem which is also in the form of a conventional least square error (LSE) estimation. The solution can be seen as the output of a smoothing operator which filters out high-frequency components. The conceived detector also has a constant false alarm rate (CFAR) with respect to the noise variance. The proposed method is compared to a conventional signal detection scheme without considering smoothness priors, and the results demonstrate its significant performance in low-signal-to-noise ratio (SNR) situations. This method can be applied to several applications, dealing with the detection of signal activity in wide-band noise, especially in the absence of reliable priors regarding the signal and noise structures.

Detection of the P300 component from the background electroencephalogram (EEG) signals is used as a case study for the proposed method. It is believed that P300 is a smooth template which its most power lies in the Delta and Theta frequency bands (i.e. between 0.5 to 5Hz) [8]; however, the behavior of background EEG tends to a Gaussian noise which covers the entire frequency band of study (0 to 60Hz) [9]. In addition to the P300 smoothness, two essential prior knowledge are adopted in terms of positivity and approximate occurrence time of this component. In general, P300 detection in the background EEG signals is an important subject in the field of biomedical signal processing. P300 has several applications especially in the context of brain-computer interface (BCI) mind speller [10]–[13]. The P300 component can be elicited by an oddball setting. In such a setting, subjects are exposed to a series of frequent (non-target) and some infrequent (target) visual or audio stimuli. Research findings revealed when a subject is exposed to an expected (target) stimulus, about 300ms post-stimulus, there will be a positive deflection in

Smooth Signal Activity Detection in White Gaussian Noise: Application to P300 Detection

Ali Mobaien
Department of Electrical and Computer Engineering
Shiraz University
Shiraz, Iran
Email: a.mobaien@shirazu.ac.ir

Reza Boostani
Department of Electrical and Computer Engineering
Shiraz University
Shiraz, Iran
Email: boostani@shirazu.ac.ir

Negar Kheirandish
Department of Electrical and Computer Engineering
Shiraz University
Shiraz, Iran
Email: negarkheirandish@yahoo.com
his/her brain signals. This component is known as P300 [14] (Fig. 1b). Due to the very low SNR of P300 in the background EEG, P300 detection is challenging and usually needs several trials to detect this component with acceptable confidence.

There are a massive number of studies that address this issue and various preprocessing and classification strategies are assessed regarding this problem. The most frequent employed classification methods are discriminant analysis (DA), support vector machine, distance-based classifiers, artificial neural networks and ensemble classifiers [15–19]. According to a review article by Philip et al. [20], most of the studies on visual P300 spellers (about 43%) use discriminant analysis in their classification stage due to providing significant results. Hence, the proposed method is compared to linear and quadratic versions of discriminant analysis (LDA and QDA) [21]. These methods are first assessed on a synthetic dataset, which satisfies the problem assumptions; Then, in the second experiment, they are evaluated over the dataset-II from BCI competition III (a P300 speller paradigm). The results are illustrated in terms of detection rate, false alarm rate, and classification accuracy.

II. METHOD

Here, a GLR test is developed to detect the activity of an unknown smooth signal in white Gaussian noise with unknown variance. As mentioned before, P300 is known to be a positive smooth deflection that occurs about 300ms after infrequent stimuli. This gives other prior knowledge about P300 in terms of positivity and approximate occurrence time.

For the ML estimation of noise variance under hypothesis $H_0$ we have $\hat{\sigma}_0^2 = \frac{1}{N}\sum_{n=1}^{N} x_n^2$. Therefore,
\[
f (\bar{x}; \sigma_0^2, H_0) = \left( \frac{1}{\sqrt{2\pi \sigma_0^2}} \right)^N \exp \left( -\frac{1}{2\sigma_0^2} \sum_{n=1}^{N} x_n^2 \right)
\]

For the second hypothesis we have:
\[
f (\bar{x}; \sigma_1^2, m_1, m_2, \hat{\rho}, H_1) = \left( \frac{1}{\sqrt{2\pi \sigma_1^2}} \right)^N \exp \left( -\frac{N}{2} \right)
\]

From (3), it can be seen that without positivity and smoothness constraints, for $m_1 \leq n \leq m_2$, $\hat{\rho}_n = \max\{0, x_n\}$. How to impose smoothness constraint is discussed in the following subsection.

For the ML estimation of $m_1$ and $m_2$, it is clear that without positivity and smoothness constraints, $\hat{m}_1 = k_1$ and $\hat{m}_2 = k_2$. In case of imposing these constraints, $\hat{m}_1$ and $\hat{m}_2$ can be found by a simple search. Now the GLR detector is obtained by likelihood ratio as:
\[
L_{GLR}(\bar{x}) = \frac{f (\bar{x}; \sigma_1^2, \hat{m}_1, \hat{m}_2, \hat{\rho}, H_1)}{f (\bar{x}; \sigma_0^2, H_0)} = \left( \frac{\hat{\sigma}_0^2}{\sigma_1^2} \right)^{N/2}
\]

\[
\left( \frac{\sum_{n=1}^{N} x_n^2}{\sum_{n=m_1}^{m_2} x_n^2 + \sum_{n=m_1}^{m_2} (x_n - \hat{\rho}_n)^2 + \sum_{n=m_2+1}^{N} x_n^2} \right)^{N/2}
\]
Equation (6) can be simplified as \( L_{\text{GLR}}(\mathbf{z}) = (1/(1-t))^N/2 \) where \( t \triangleq \sum_{n=1}^{N} (2\hat{p}_n x_n - \hat{p}_n^2)/\sum_{n=1}^{N} x_n^2. \) Note that \( 0 \leq t < 1 \) and \( L_{\text{GLR}} \) is a monotonic increasing function with respect to \( t. \) Therefore, the GLR test can be further simplified as,

\[
L_{\text{GLR}}(\mathbf{z}) = \frac{1}{\sum_{n=1}^{N} \frac{N}{2}} (2\hat{p}_n x_n - \hat{p}_n^2) H_1, H_0, \eta_{\text{GLR}} \tag{7}
\]

\( \eta_{\text{GLR}} \) can be obtained for a specified false alarm rate either by the use of equation “\( \sup P\{L_{\text{GLR}}(\mathbf{z}) > \eta_{\text{GLR}}; H_0\} = P_{fa}\) (\( P\{\} \) indicates probability), or by means of Monte Carlo simulation [1]. The mentioned equation requires the PDF of \( L_{\text{GLR}} \) which may be challenging to be obtained; Hence, simulation is more feasible. One may concern that the proposed GLR detector is CFAR with respect to noise variance (i.e. noise variance variation would not change \( \eta_{\text{GLR}} \)).

3) Imposing the Smoothness Constraint: The ML estimation of non-zero elements of \( \mathbf{p} \) can be presented as the following optimization problem,

\[
\hat{\mathbf{p}} = \arg \min_{\mathbf{p}} \| \mathbf{x} - \mathbf{p} \|^2 \text{ s.t. } \mathbf{p} \succeq 0 \text{ and smooth }
\]

where \( \mathbf{x} \) and \( \mathbf{p} \) are column vectors formed by elements of \( \mathbf{x} \) and \( \mathbf{p} \) from \( m_1 \) to \( m_2 \), respectively. This presentation is the reduced form of (5). To model the smoothness constraint, the fact that the \( j \)th derivative of a smooth sequence is smaller than a threshold, is used [22]–[24]. Here, a matrix form of discrete approximation of smoothing operator is presented. Let \( \mathbf{d}_j \) be the finite impulse response approximation of length \( M \), for the \( j \)th order derivative operator. For \( j > 2 \), this operator can be achieved iteratively by \( \mathbf{d}_j = \mathbf{d}_{j-1} * \mathbf{d}_1 \) where \( \mathbf{d}_1 = [1, -1, 1] \) and \( * \) sub-index indicates the convolution operator. For a signal length \( L > M \), \( \mathbf{D}_j \in \mathbb{R}^{(L+M-1) \times L} \) is defined as a Toeplitz matrix form of \( \mathbf{d}_j \). For instance, \( \mathbf{d}_2 = [1, -2, 1] \) and

\[
\mathbf{D}_2 = \begin{bmatrix}
1 & 0 & \cdots & 0 \\
-2 & 1 & \cdots & \vdots \\
1 & -2 & \cdots & 0 \\
0 & 1 & \cdots & 1 \\
\vdots & \ddots & \ddots & \vdots \\
0 & \cdots & 0 & 1
\end{bmatrix}_{(L+M-1) \times L}
\]

(Note that the derivative operator \( \mathbf{d}_2 \) is shifted along rows as well as columns of \( \mathbf{D}_2 \).) The left multiplication of \( \mathbf{D}_j \) in any column-wise signal of length \( L \) is equivalent to the \( j \)th order derivative approximation of the signal [24]. Now the problem can be modeled as a convex optimization problem with quadratic constraints,

\[
\hat{\mathbf{p}} = \arg \min_{\mathbf{p}} \| \mathbf{x} - \mathbf{p} \|^2 \text{ s.t. } \mathbf{p} \succeq 0 \text{ and } \| \mathbf{D}_j \mathbf{p} \|^2 \leq \delta^2 \tag{8}
\]

Here \( \delta^2 \) indicates the smoothness bound. For simplicity, the positivity constraint is relaxed and the problem is solved. At last, each element of \( \mathbf{p} \) that has negative value is set to zero. The Lagrangian form of relaxed problem is

\[
\hat{\mathbf{p}} = \arg \min_{\mathbf{p}} \left\{ \| \mathbf{x} - \mathbf{p} \|^2 + \lambda \| \mathbf{D}_j \mathbf{p} \|^2 \right\}, \text{ s.t. } \lambda \geq 0 \tag{9}
\]

where \( \lambda \) is the Lagrange coefficient. This problem is in the form of a classic LSE problem with a quadratic constraint ([22], [25], [26]) and its solution is obtained by:

\[
\hat{\mathbf{p}} = (\mathbf{I} + \lambda \mathbf{D}_j^T \mathbf{D}_j)^{-1} \mathbf{x}, \tag{10}
\]

(\( I \) + \( \lambda \mathbf{D}_j^T \mathbf{D}_j \))^{-1} is a smoothing operator which maps \( \mathbf{x} \) into a smooth subspace. By increasing \( \lambda \) from 0, the smoothness of estimated signal will increase. Fig. 1a shows the estimated signal \( \hat{\mathbf{p}} \) from \( \mathbf{x} \) for different values of \( \lambda \) and \( j = 2 \). By replacing (10) in (7) and letting \( A = (\mathbf{I} + \lambda \mathbf{D}_j^T \mathbf{D}_j)^{-1} \), the matrix form of the proposed hypothesis test is achieved as:

\[
L_{\text{GLR}}(\mathbf{z}) = \frac{\mathbf{x}^T (2\lambda - A) \mathbf{x}}{\mathbf{x}^T \mathbf{x}} H_1, H_0, \eta_{\text{GLR}} \tag{11}
\]

III. METHOD EVALUATION

A. Evaluation Scheme and Parameter Selection

For a detection problem, the detector threshold is chosen in a way that a desired false alarm rate (\( P_{fa} \)) is achieved. A good detector is the one which has a large detection rate (\( P_d \), also known as detector power) for a specified \( P_{fa} \). In general, there are two ways to compare and evaluate different detectors. 1- Calculating \( P_d \) for a fixed \( P_{fa} \) in different SNRs. 2- Calculating \( P_d \) in different values of \( P_{fa} \) for a fixed SNR, known as receiver operating characteristic (ROC) curve. Afterwards, these evaluation metrics are achieved for four detectors driven from subsection II-B, named as \( L, L_2, L_3 \) and \( L_{sp} \), using a synthetic dataset which satisfies the aforementioned assumptions. These four are GLR tests with or without the positivity and smoothness constraints, where “\( p \)” and “\( s \)” sub-indices indicate which constraint is imposed. Also, the proposed method (\( L_{sp} \)) is compared to LDA and QDA in terms of classification accuracy. In this regard, a 10-fold cross validation approach is applied to two equal size groups of

![Fig. 1](image-url)
sequences, one contains and the other does not contain the P300 component. This is done on the mentioned synthetic dataset as well as a real dataset of visual P300 paradigm from BCI competition III.

Note that for the presented results, the smoothing order is set to 2, $\lambda = 50$ for synthetic data, and $\lambda = 5$ for real data. Also, for the case of synthetic data, $m_1$ and $m_2$ are fixed to the sample numbers corresponding to 200 and 400ms respectively, as we know these are the best values based on synthetically generated data. But for the case of real data, the exact position of P300 is unknown, so it is assumed that $200\text{ms} < m_1/f_s < 300\text{ms}$ and $350\text{ms} < m_2/f_s < 500\text{ms}$ ($f_s$ is sampling frequency), then the best $m_1$ and $m_2$ are found by searching through each sequence under test.

B. Data

1) Synthetic Data: To generate each synthetic EEG sequence, an 800ms (with $f_s = 240\text{Hz}$) column vector of zero-mean white Gaussian noise samples is generated as the background activity. Then a synthetic P300 is made by a Gaussian function. The latency, amplitude, and the width of this Gaussian function and also a term for shifting are drawn from normal distributions based on prior knowledge of P300 morphology [14]. Background activity confirms the hypothesis $H_0$ and the sum of background activity and P300 confirms the hypothesis $H_1$.

2) Real Data: The presented method is investigated on dataset II from BCI competition III (P300 speller paradigm) [27]. Data are recorded from two subjects while trying to spell 185 characters. To spell each character, subjects are presented with 12 stimuli, where two of them are target stimuli and are believed to contain P300. Each stimulus is repeated 15 times. EEG signals are collected from 64 channels, bandpass filtered from 0.1 to 60Hz, and digitized at 240Hz.

C. Results

Fig. 2 shows the performance of four mentioned detectors using synthetic data. Fig. 2a presents $P_d$ in different SNRs and $P_{fa} = 0.01$. Also, ROC curves are presented in Fig. 2b. The superiority of the detector which uses both positive and smoothness constraints is clear. It is interesting that this detector is very powerful and gives almost $P_d = .9$ in SNR -10db. The other three detectors are also very powerful since synthetic data fully satisfy the assumptions of the defined detection problem. Note that for values near to zero $L_s$ tends to $L$ and $L_{sp}$ tends to $L_p$ (it can be seen from (10)).

The performance of $L_{sp}$ is assessed and compared with LDA and QDA on both real and synthetic data (Fig. 3). To do this, $L_{sp}$ is used as a classifier and its threshold is set to maximize the classification accuracy of train data. The performance of the methods is assessed using a 10-fold cross-validation approach. For the case of synthetic data, a set of 1000 sequences that confirm $H_0$ and 1000 sequences that confirm $H_1$ are used in different SNRs. The proposed method provides higher accuracy especially in low SNRs (Fig. 3a). For the case of real data, to increase P300 SNR, the average of the first $k$ trials is calculated on 800ms windows of post stimuli signals. Here, we have 185 characters and for each 10 averaged trials containing noise (i.e. 1850 trials) and 2 other averaged trials containing noise plus P300 (i.e. 370 trials). Here, the results for subject A and channel $C_2$ are presented. Fig. 3b shows the performance of $L_{sp}$, LDA and QDA in different SNRs (i.e. different $k$ values). The performance of the methods is assessed using 10-fold cross-validation on 370 target sequences and 370 randomly chosen non-target sequences (five random non-overlapping groups of non-target sequences are generated and the final results are presented as the average results for these groups). The proposed method provides higher accuracy, especially in low SNRs.

IV. Conclusion

In this research, a novel and powerful approach towards smooth signal activity detection in white Gaussian noise is proposed. The method is derived from a GLR test. This test uses the ML estimation of unknown parameters under each hypothesis. The smoothness constraint on the desired signal is led to an optimization problem which is modeled by an LSE estimation problem with quadratic constraints. A smoothing operator is derived from the solution which can estimate the signal with a tunable level of smoothness. The proposed method can be used for P300 detection. This method is applied to synthetic and real datasets of the P300 paradigm and also compared to state-of-the-art LDA and QDA in different SNR levels. The method shows superior results in terms of classification accuracy, especially in low SNRs. This outcome
is very promising since the performance of signal processing methods is mainly diminished in low SNR environments.

REFERENCES

[1] S. M. Kay, Fundamentals of Statistical Signal Processing: Detection Theory. Upper Saddle River, NJ: Prentice-Hall, 1998.
[2] L. L. Scharf, Statistical Signal Processing. Reading, MA: Addison-Wesley, 1991.
[3] E. L. Lehmann and J. P. Romano, Testing Statistical Hypotheses, 3rd ed. New York: Springer Science+Business Media, 2005.
[4] H. V. Poor, An Introduction to Signal Detection and Estimation, 2nd ed. New York: Springer, 1994.
[5] M. Kendall and A. Stuart, The Advanced Theory of Statistics, 4th ed. London: Charles Griffin & Company, 1979, vol. 2.
[6] J. R. Gabriel and S. M. Kay, “On the relationship between the GLRT and UMPI tests for the detection of signals with unknown parameters,” IEEE Trans. Signal Process., vol. 53, no. 11, pp. 4194–4203, nov 2005.
[7] J. Candy, Model-Based Signal Processing. Wiley-IEEE Press, 2005.
[8] J. Polich, “Updating P300: An integrative theory of p3a and p3b,” Clin. Neurophysiol., vol. 118, no. 10, pp. 2128–2148, 2007.
[9] R. Sameni and E. Seraj, “A robust statistical framework for instantaneous electroencephalogram phase and frequency estimation and analysis,” Physiol. Meas., vol. 38, no. 12, p. 2141, 2017.
[10] A. Haider and R. Fazeli-Rezai, “Application of P300 event-related potential in brain-computer interface,” Intech, pp. 19–38, 2017.
[11] L. A. Farwell and E. Donchin, “Talking off the top of your head: Toward a mental prosthesis utilizing event-related brain potentials,” Electroen. Clin. Neuro., vol. 70, no. 6, pp. 510–523, 1988.
[12] A. Mobaien and R. Boostani, “ACSP: Adaptive CSP filter for BCI applications,” in 2016 24th Iranian Conference on Electrical Engineering (ICEE). IEEE, 2016, pp. 466–471.
[13] M. Dehghan, A. Mobaien, and R. Boostani, “A deep neural network-based transfer learning to enhance the performance and learning speed of BCI systems,” Brain-Computer Interfaces, vol. 8, no. 1-2, pp. 14–25, 2021.
[14] S. H. Patel and P. N. Azzam, “Characterization of N200 and P300: Selected studies of the event-related potential,” Int. J. Medical Sci., vol. 2, no. 4, p. 147, 2005.
[15] J. Qu, F. Wang, Z. Xia, T. Yu, J. Xiao, Z. Yu, Z. Gu, and Y. Li, “A novel three-dimensional P300 classifier based on stereo visual stimuli,” IEEE Trans. Human-Mach. Syst., vol. 48, no. 4, pp. 392–399, 2018.
[16] M. Kaper, P. Meinicke, U. Grosskathofer, T. Lingner, and H. Ritter, “BCI competition 2003-data set IIb: Support vector machines for the P300 speller paradigm,” IEEE Trans. Biomed. Eng., vol. 51, no. 6, pp. 1073–1076, 2004.
[17] M. Fira, “Detection of P300 in a BCI Speller,” in International Conference on Hybrid Information Technology. Springer, 2011, pp. 481–487.
[18] H. Cecotti and A. Graser, “Convolutional neural networks for P300 detection with application to brain-computer interfaces,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 33, no. 3, pp. 433–445, 2010.
[19] S. Kundu and S. An, “P300 detection with brain–computer interface application using PCA and ensemble of weighted SVMs,” IETE J. Res., vol. 64, no. 3, pp. 406–414, 2018.
[20] J. T. Philip and S. T. George, “Visual P300 Mind-Speller Brain-Computer Interfaces: A Walk Through the Recent Developments With Special Focus on Classification Algorithms,” Clin. EEG Neurosci., 2019.
[21] G. J. McLachlan, Discriminant Analysis and Statistical Pattern Recognition. John Wiley & Sons, 2004, vol. 544.
[22] W. Gersch, “Smoothness priors,” in New Directions in Time Series Analysis. New York: Springer, 1993, pp. 113–146.
[23] M. P. Tarvainen, P. O. Ranta-Aho, and P. A. Karjalainen, “An advanced detrending method with application to HRV analysis,” IEEE Trans. Biomed. Eng., vol. 49, no. 2, pp. 172–175, 2002.
[24] R. Sameni, “Online filtering using piecewise smoothness priors: Application to normal and abnormal electrocardiogram denoising,” Signal Process., vol. 133, pp. 52–63, 2017.
[25] S. P. Boyd and L. Vandenberghe, Convex Optimization. Cambridge university press, 2004.
[26] W. Gander, “Least squares with a quadratic constraint,” Numer. Math. (Heidelberg), vol. 36, no. 3, pp. 291–307, 1980.
[27] B. Blankertz et al. (2005) BCI competition III webpage. [Online]. Available: http://www.bbci.de/competition/ii/