



Optimal Noise Density Approximation for Stochastic Resonance

Rujipan Sampanna^{1*} Sanya Mitaim¹

¹*Department of Electrical and Computer Engineering, Faculty of Engineering,
Thammasat University, Thailand*

* Corresponding author's Email: rujipan.s@bu.ac.th

Abstract: A noise benefit or stochastic resonance (SR) occurs when noise enhances the performance of a signal system. The performance is maximum when the noise probability density function (pdf) matches the signal system. Researchers have studied SR using noise with assumed forms of pdf such as Gaussian or uniform or other popular pdfs. Then they studied the SR effect by adjusting the parameters of that noise pdf. This paper proposes a use of finite Gaussian mixture model (GMM) that approximates any noise pdf with finite moments. Thus varying the parameters of a GMM can reshape the pdf to new forms. Then we can optimize the performance of a signal system by using proper GMM parameters. Any optimization algorithm such as particle swarm optimization (PSO) can help search for respective parameters of GMM that lead to the approximation of the (locally) optimal pdf. We show the results using P300 classification systems as test cases. Two-sample t-tests show that GMM noise can improve the performance of the noise-free system as well as the Gaussian-noise systems with p -value < 0.05 .

Keywords: Stochastic resonance, Gaussian mixture model, Particle swarm optimization, P300 classification.

1. Introduction

Noise benefit or stochastic resonance (SR) effect is a phenomenon when noise at appropriate intensity levels can enhance weak input signals. SR effect can occur in engineering systems such as signal processing, image processing, communications, and control [1-4]. Researchers show that noise can improve statistical signal detection for array-based nonlinear correlators in Neyman-Pearson (NP) and maximum-likelihood (ML) signal detection [5, 6]. They show that noise benefit rate improves in terms of the small-quantizer noise limit as the number of array quantizers increases. Noise enhanced hypothesis testing is studied in the restricted NP criterion where the optimal additive noise can be represented by a discrete random variable with a certain number of point masses [7]. Noise can also enhance model in the binary hypothesis testing, where the optimal additive noise is added to increase the detection probability P_D and decrease the false alarm probability P_{FA} [8].

Many types of noise pdf such as uniform, Gaussian, Laplacian noise, and other noise pdfs can enhance the performance of nonlinear systems [1-4]. These studies show that different noise pdfs distinctly affect the system performance. The optimal shape of noise pdf depends on characteristics of a signal system. SR studies often pick a shape or form of pdf in advance. This pre-selected pdf can ease up the task of SR induction but it can lead to inferior performance if the selected choice of pdf is not the optimal one. Gaussian mixture model (GMM) can model any noise pdf with finite statistics. Many applications use GMM to improve performance of signal processing systems such as nucleus segmentation, predicting epileptic seizures, image restoration, blood pressure estimation and brain MR image segmentation [9-13]. Thus we propose a use of GMM to approximate an optimal noise pdf for the SR effect. The GMM can approximate the optimal noise pdf. Then we propose a use of particle swarm optimization (PSO) to search for (local) optimal parameters of Gaussian mixture model.

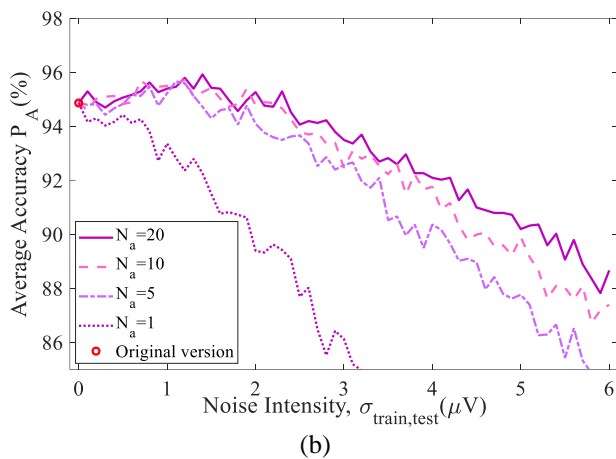
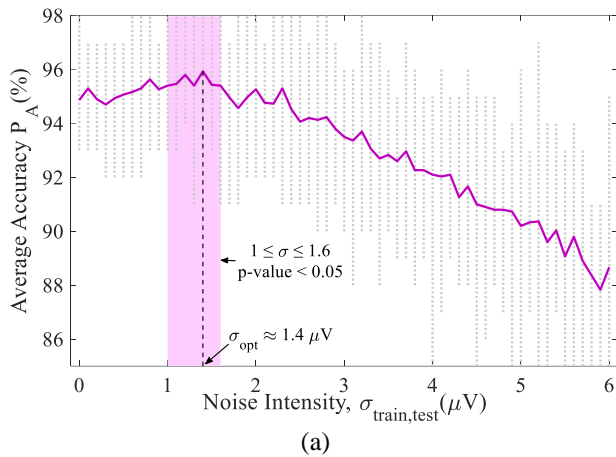


Figure 1. Gaussian noise enhances accuracy of P300 classification: (a) The system uses array of ESVM classifiers with $N_a = 20$ stages. (b) The test shows how classification accuracy P_A increases as the number of stages N_a increases

We use the brain-computer interface (BCI) system as a test case to show how we can induce the SR effect using (locally) optimal GMM noise. BCI is a communication system that allows a person to send commands to an external device through direct measurements of brain activities without using any movement [14-16]. We specifically test our proposed method using P300 applications as it has been one of the most important types of EEG signals in BCI applications [17-20] that include P300 speller paradigm, neurophone, controlling a wheelchair and a robotic arm [19-23].

Earlier results show that noise can benefit the BCI systems in many ways [24-28]. Noise can improve accuracy of the P300 classification system [29] when we compare it with existing methods such as [30-33]. BCI systems in actual use may also have limitations from small amount of signal data for training [14, 15] or the need for reduction of data collection times (the number of signal repetitions) [32, 33]. Noise can improve the accuracy of an array

of P300 classifications systems as shown in Fig.1 and can also reduce the number signal repetitions and thus reduce collection time [27-29]. Noise can also enhance the accuracy of the system using multiple user dataset in training phase [29]. The noise-added system still performs well on the reduced channel and restricted positions of commercial recording systems [29].

Fig. 1 shows Gaussian noise enhances accuracy of P300 classification. The classification system uses array of BCI competition III (A) dataset with $N_r = 15$ signal repetitions and 64 channels. The additive i.i.d. noise is Gaussian with zero mean and with standard deviation or intensity σ . We vary the intensity σ and repeat the test 30 times and show the average. (a) The system uses array of ESVM classifiers with $N_a = 20$ stages. The vertical dash lines show the variation from different noise realizations. The system shows an optimal point at (approximately) $\sigma_{opt} \approx 1.4 \mu\text{V}$ and classification accuracy P_A of 95.93%. Two-sample t-tests show that there is noise benefit (or improvement of the classification accuracy P_A) with p -value < 0.05 for the noise intensity σ in the interval $1 \mu\text{V} \leq \sigma \leq 1.6 \mu\text{V}$. (b) The test shows how classification accuracy P_A increases as the number of stages N_a increases.

Section 2 describes the SR effect with the optimal noise pdf for nonlinear systems. Section 3 discusses the choice of the EEG classification systems that we use to test the SR effect. The proposed system uses arrays of ESVM classification systems [27-29, 32, 33] with additive noise in the training phase and the testing phase (or in actual use). Section 4 shows the extensive experimental results on the BCI competition II dataset Iib [34] and the BCI competition III dataset II [35]. Then, we show that a use of finite Gaussian mixture model (GMM) to approximate any noise pdf with finite moments. The optimal noise pdf can improve the accuracy and reduce data collection time [27-29]. We use the particle swarm optimization (PSO) technique to search for a (locally) optimal GMM noise density for the P300 classification system as shown in Fig. 3. Experimental results show that different noise pdf leads to different system response that also affects the system's performance. Addition of noise from a pdf that matches the system's characteristics can maximally enhance the performance.

2. SR-optimal noise PDF for nonlinear systems

Additive noise can enhance the performance of some nonlinear systems. Most researchers study SR by assuming known forms of noise pdf in advance such as Gaussian noise. Then they vary the intensity level to examine the SR effects [1-8]. They also derive the conditions for optimal noise to use in the EM algorithm [4, 36]. Yet the optimal noise pdf is not fully explored. We propose the use of GMM to approximate the optimal noise pdf. Then we use particle swarm optimization (PSO) to search for the GMM parameters that approximates the optimal pdf.

2.1 Optimal noise PDF approximation

Finite Gaussian mixture model (GMM) approximates any pdf with finite variance. Thus varying the GMM parameters can reshape the pdf to any forms if the model consists of enough number of components.

The Gaussian mixture model (GMM) $p(x)$ is a weighted sum of N_c Gaussian densities:

$$p(x) = \sum_{i=1}^{N_c} w_i p_i(x|\lambda_i) \quad (1)$$

where $\lambda_i = \{w_i, \mu_i, \sigma_i^2\}$ contains the mean μ_i and variance σ_i^2 of the Gaussian pdf and w_i are the mixture weights such that $\sum_{i=1}^{N_c} w_i = 1$. The i^{th} component is a univariate Gaussian density $p_i(x|\lambda_i) \sim \mathcal{N}(\mu_i, \sigma_i^2)$ of the form:

$$p_i(x|\lambda_i) = \frac{1}{\sigma_i \sqrt{2\pi}} \exp\left(-\frac{(x-\mu_i)^2}{2\sigma_i^2}\right) \quad (2)$$

2.2 Optimal GMM noise PDF search using PSO

Finding parameters of the GMM that maximize the performance is an optimization problem. PSO is one of the techniques that can search for the parameter of Gaussian mixture density to obtain the (locally) optimal standard deviation σ_{opt} , mean μ_{opt} and weight w_{opt} of each mixture component. The objective function of classification problem is to maximize the classification accuracy P_A subject to the mixture's intensity σ , mean μ , and mixture weight w between the lower bound σ_l, μ_l, w_l and upper bound σ_u, μ_u, w_u [37-39]:

$$\begin{aligned} &\text{Maximize } P_A(\lambda) \text{ subject to } \sigma_l \leq \sigma \leq \sigma_u \\ &\mu_l \leq \mu \leq \mu_u, w_l \leq w \leq w_u \end{aligned} \quad (3)$$

We can also apply other optimization techniques to solve this optimization problem. Here we show how PSO can search for the parameters of optimal noise pdf for P300 classification system using the following steps [37-39]:

Step 1: Define the number of particles $q = 5$. Note that more particles can give better results but it requires more search time.

Step 2: Randomly generate initial particles $\sigma_i^0, \mu_i^0, w_i^0, i = 1, \dots, q$ in the range $(\sigma_l, \sigma_u) = (0.01, 500)$, $(\mu_l, \mu_u) = (-500, 500)$ and $(w_l, w_u) = (0, 1)$.

Step 3: Compute the ESVM classification accuracy P_A (objective function value) at $\lambda_i^0 = \{\sigma_i^0, \mu_i^0, w_i^0\}$ as $P_A(\lambda_1^0), P_A(\lambda_2^0), \dots, P_A(\lambda_q^0)$.

Step 4: Set the initial velocity of each particle v_i^0 to zero. Set the iteration number as $k = 1$.

Step 5: Obtain the personal best values $\lambda_i^{lbest} = \{\sigma_i^{lbest}, \mu_i^{lbest}, w_i^{lbest}\}$ that give highest value of the objective function $P_A(\lambda_i^j)$ from the i^{th} particle in all previous iterations $j = 1, \dots, k$, and obtain the global best value $\lambda_i^{gbest} = \{\sigma_i^{gbest}, \mu_i^{gbest}, w_i^{gbest}\}$ that gives the highest value of the objective function $P_A(\lambda_i^j)$ from all particles $i = 1, \dots, q$ in all previous iterations $j = 1, \dots, k$.

$$\lambda_i^{lbest} = \underset{1 \leq j \leq k}{\operatorname{argmax}} P_A(\lambda_i^j) \quad (4)$$

$$\lambda_i^{gbest} = \underset{1 \leq j \leq k, 1 \leq i \leq q}{\operatorname{argmax}} P_A(\lambda_i^j) \quad (5)$$

Step 6: Compute the velocity v_i^{k+1}

$$v_i^{k+1} = v_i^k + \alpha_i(\lambda_i^{lbest} - \sigma_i^k) + \beta_i(\lambda_i^{gbest} - \sigma_i^k) \quad (6)$$

where $\lambda_i = \{\sigma_i, \mu_i, w_i\}$, α_i and β_i are uniform (0,1) random numbers.

Step 7: Update the particles $\sigma_i^{k+1}, \mu_i^{k+1}$ and w_i^{k+1}

$$\sigma_i^{k+1} = \sigma_i^k + v_i^{k+1} \quad (7)$$

$$\mu_i^{k+1} = \mu_i^k + v_i^{k+1} \quad (8)$$

$$w_i^{k+1} = w_i^k + v_i^{k+1} \quad (9)$$

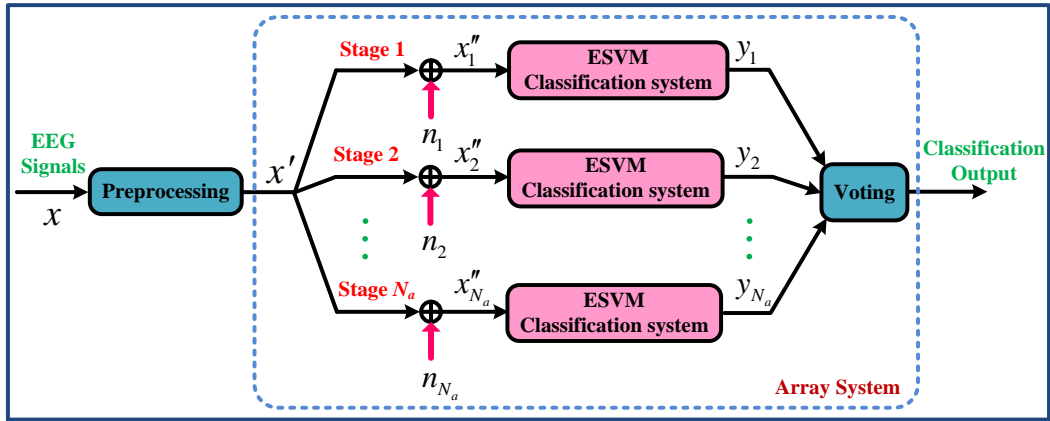


Figure. 2 A noisy array of ESVM classifiers for P300 classification

Step 8: Evaluate the objective function at the current

$$\lambda_i^k = \{\sigma_i^k, \mu_i^k, w_i^k\} \text{ as } P_A(\lambda_1^k), P_A(\lambda_2^k), \dots, P_A(\lambda_q^k) \quad (10)$$

Step 9: Check the convergence of the PSO process. The process converges when the positions of all particles converge to the same solution (the same noise intensity, mean, weight). Thus we obtain the (local) optimal noise intensity σ_{opt} , noise mean μ_{opt} and weight w_{opt} that provides the maximum accuracy $P_A(\sigma_{opt}, \mu_{opt}, w_{opt})$. In case the objective function values P_A does not converge, go to Step 5.

3. SR-Array of P300 classification systems

This section shows how the proposed algorithm can search for locally optimal GMM noise pdf to improve accuracy of nonlinear system. We use “noisy” EEG classification system for P300 as shown in Fig. 2 [29] as a case study. The classification system is an array of N_a identical P300 classification systems with additive i.i.d. noise at each stage.

Each stage processes the same filtered signal x' of the raw EEG data x obtained from the sensors. Then each stage i adds i.i.d. GMM noise n_i to the signal x' to obtain the noise-added signal x_i'' : $x_i'' = x' + n_i$. Our goal is to obtain the (locally) optimal noise pdf. Training the system with known (labelled) data gives us (local) optimal parameters to use in the actual EEG signal classification system (or the testing phase). In practice, the training phase can be performed in advance or before each session.

3.1 ESVM classifier

We use an ESVM as a classifier as ESVM

combines several support vector machines (SVM) classifiers to solve problems associated with signal variations between subjects and over time for P300 classification [33]. An SVM classifier training process needs to find an optimal hyperplane that separates two classes with the largest possible margin in order to increase the performance of the classifier for the unknown data (testing data). In the training phase, we divide the training dataset $\{x\}$ into M clusters. Then we train the k^{th} SVM classifier with the data cluster k^{th} to obtain the respective weights w_{kj} , bias b_k , and Lagrangian multipliers α_{kj} . $K_{kj}(E_{kj}, x) = (E_{kj}^T x + 1)^n$ is the polynomial kernel function where n is the degree of a polynomial. And support vectors $E_{kj}, j = 1, 2, \dots, L_k$, where L_k is the number of support vectors of the k^{th} classifier. The output of an SVM classifier for input data x is as follows:

$$f_k(x) = \text{sgn}(\sum_{j=1}^{L_k} y_{kj} \alpha_{kj} K_{kj}(E_{kj}, x) + b_k) \quad (11)$$

for $k = 1, \dots, M$. In this work we use $n = 3$.

P300 speller paradigm implies a two-class classification problem with class labels $y_i \in \{-1, 1\}$ as “present” and “absent”. Thus, the output of the ESVM classifier is a sign of the sum of M SVM classifiers:

$$\hat{y} = S(x) = \text{sgn}(\sum_{k=1}^M f_k(x)) \quad (12)$$

We test the classifier using different numbers of clusters for each dataset. We use the cluster number M that gives the highest accuracy (17 clusters for A, B and 8 clusters for C), as shown in Table 1 to examine the noise benefits.

Table 1. Datasets of P300 speller paradigm

Dataset (Subject)	Number of Characters	Number of Signals		#Clusters M	#Signals per cluster
		Target	Non-target		
BCI IIIA (A)					
Training:	85	2,550	12,750	17	5
Testing:	100	3,000	15,000		
BCI IIIB (B)					
Training:	85	2,550	12,750	17	5
Testing:	100	3,000	15,000		
BCI II (C)					
Training:	42	1,260	6,300	8	5
Testing:	31	930	4,650		

3.2 P300 signal features

Table 1 gives the descriptions of P300 dataset recorded from 64 electrodes, 240 Hz sampling rate and 0.1-60 Hz filter. We denote the datasets as Subject A for BCI competition III dataset II, Subject B for BCI competition III dataset II, and Subject C for BCI competition II dataset IIb. Each dataset represents a set of characters where each character was repeated 15 times ($N_r = 15$ signal repetitions) to reinforce the P300 responses [34,35]. In our experiments we also test the classification system using fewer numbers of repetitions as well.

We use the EEG signal x in the time window of 0-667 ms after stimulus using a 240 Hz sampling rates [33]. The signal passes through a 0.1-20 Hz bandpass filter to obtain x' . Then we use amplitudes of x' as features. Thus the dimensionality of feature space is 896 samples for 64 channels (667 ms \times 240 Hz = 160 samples per channel and reduces to 14 samples per channel). Then we add Gaussian mixture noise n to x' to explore the SR effect. P300 classification uses samples of EEG signal $x(t)$ from all channels as features. We examine the array of N_a ESVM classifiers. We also vary the signal repetitions N_r from 1 to 15. The character prediction process considers the row (r) and column (c) that have the highest scores from different signal repetitions and sum up the scores of \hat{y} from corresponding rows and columns as follows [33]:

$$c = \operatorname{argmax}_{1 \leq i \leq 6} \sum_{e=1}^{N_r} \hat{y}_{ie} \quad (13)$$

$$r = \operatorname{argmax}_{7 \leq i \leq 12} \sum_{e=1}^{N_r} \hat{y}_{ie} \quad (14)$$

where $c \in \{1, 2, \dots, 6\}$ and $r \in \{7, 8, \dots, 12\}$ are

column and row numbers that have the highest scores, and $N_r = 15$ is the number of signal repetitions.

4. Experimental results

We test the GMM-noise SR for P300 classification using EEG data as shown in Table 1 [34,35]. We use an original classification system as a building block for an array system as shown in Fig. 2. Then we add i.i.d. GMM noise with parameters $\lambda = \{\sigma_i, \mu_i, w_i\}_{i=1}^{N_c}$ to the original signal data x' .

We examine the SR effects using the following steps: **Step 1:** Preprocessing and feature extraction. Use bandpass filtering and other feature extraction techniques. **Step 2:** ESVM classifier training. Divide the training dataset into M clusters and use each cluster to train SVM classifier. **Step 3:** Test the system with noise addition in training and testing phases. We calculate the classification accuracy P_A as a ratio of the number of correct classification outputs N_C and the total number of testing characters N_T in the experiment: $P_A = \frac{N_C}{N_T} \times 100\%$. The classification accuracy varies with noise realizations. Thus we repeat each test 30 times and find the average to determine the system performance P_A .

Fig. 1 shows the SR effect using Gaussian noise $\mathcal{N}(0, \sigma)$ in training and testing phases. The P300 classification system uses an array of ESVM classifiers with number of stages $N_a = 20$ and number of signal repetitions $N_r = 15$. The result also shows that the classification accuracy improves when we increase the number of stages. The accuracy improves when we add more stages and use a suitable level of noise intensity.

Table 2. P300 classification accuracy using array of ESVM classifiers with Gaussian and GMM noise

Dataset (Subject)	Number of signal repetition N_r	Accuracy without noise $P_A(0)$ (%)	Number of stages N_a	Maximum accuracy with noise, P_{A_n} (%)					Accuracy improvement $P_{A_n, \max}(\lambda) - P_A(0)$ (%) (p -value)	Accuracy improvement $P_{A_n, \max}(\lambda) - P_{A_n}(\sigma)$ (%) (p -value)
				Gaussian $P_{A_n}(\sigma)$	GMM $P_{A_n}(\lambda)$					
					Number of mixtures, N_c					
				$\mathcal{N}(0, \sigma_{opt})$	2	3	4	5		
BCI IIIA (A)	14	95.00	1	96.00	96.80	96.80	96.40	95.16	1.80 (0.0057)	0.80 (0.0362)
			10	97.00	96.80	97.20	97.00	96.80	2.20 (5.62×10^{-5})	0.20 (0.0202)
			20	97.20	97.20	96.80	97.40	98.00	3.00 (3.27×10^{-16})	0.80 (0.0148)
BCI IIIB (B)	14	94.00	1	96.60	97.00	96.40	95.40	96.40	3.00 (0.0031)	0.40 (0.0073)
			10	96.40	97.20	97.00	97.20	97.20	3.20 (1.07×10^{-6})	0.80 (0.0002)
			20	97.60	96.80	98.00	96.80	99.00	5.00 (2.21×10^{-9})	1.40 (0.0001)
BCI II (C)	4	96.77	1	98.71	98.49	98.71	98.71	99.35	2.58 (0.0376)	0.64 (0.00005)
			10	100	100	100	100	100	3.23 (9.14×10^{-5})	0
			20	100	100	100	100	100	3.23 (9.14×10^{-5})	0

Table 3. Comparison of the P300 classification accuracy of the proposed methods and existing methods

Methods	BCI IIIA (A)	BCI IIIB (B)	BCI II (C)
Existing methods			
ICA [40]	-	-	100 ($N_r = 5$)
SVM [32]	-	-	100 ($N_r = 5$)
ESVM [33]	97 ($N_r = 15$)	96 ($N_r = 15$)	-
CNN-1 [30]	97 ($N_r = 15$)	92 ($N_r = 15$)	-
MCNN-1 [30]	97 ($N_r = 15$)	94 ($N_r = 15$)	-
Proposed method	98 ($N_r = 14$)	99 ($N_r = 14$)	100 ($N_r = 4$)

Then we examine the use of GMM noise. We search the appropriate optimal GMM noise pdf by varying the noise intensity, noise mean, and mixture weights with specific signal repetition, number of mixtures, and array size. We use PSO to search for locally optimal noise intensity σ from 0.01 μV to 500 μV , noise mean μ varies from -500 μV to 500 μV , and weight of Gaussian mixture w varies between 0 to 1 (weight sum of all Gaussian mixture must be equal to 1) in training phase and testing phase. We test several number of mixtures $N_c = 2, \dots, 5$. We also test the system with different number of stages: $N_a = 1, 10$, and 20. We test the system using $N_r = 1$ to 15 signal repetitions.

Fig. 3 shows samples of GMM approximation of optimal noise pdf's for Subjects A, B, and C for P300 classification systems using ESVM. PSO search for the optimal noise pdf can give different locally optimal pdfs. Table 2 compares the classification accuracy of the array systems that use Gaussian noise and GMM noise. We use the smallest number of signal repetition N_r that already gives the highest accuracy. The accuracy in bold is the maximum accuracy among classifiers that use the same number of stages N_a .

Subject A: Accuracy $P_{A_n} = 98.00\%$ with 5-component GMM $p(n) \sim 0.082\mathcal{N}(-3.20, 3.48) +$

$0.202\mathcal{N}(2.56, 8.45) + 0.259\mathcal{N}(-7.30, 5.71) + 0.218\mathcal{N}(-2.10, 7.75) + 0.239\mathcal{N}(-6.40, 4.21)$ for $N_a = 20$ and $N_r = 14$.

Subject B: Accuracy $P_{A_n} = 99.00\%$ with 5-component GMM $p(n) \sim 0.163\mathcal{N}(-1.03, 0.1) + 0.258\mathcal{N}(1.44, 0.1) + 0.164\mathcal{N}(2.03, 0.1) + 0.178\mathcal{N}(0.81, 0.1) + 0.237\mathcal{N}(0.46, 0.1)$ for $N_a = 20$ and $N_r = 14$.

Subject C: Accuracy $P_{A_n} = 100\%$ with 2-component GMM $p(n) \sim 0.9999\mathcal{N}(47.97, 90.18) + 0.0001\mathcal{N}(17.65, 100)$ for $N_a = 10$ and $N_r = 4$.

Two-sample t-tests confirm that there is an increase of accuracy from GMM noise benefits: $P_{A_n}(\lambda) > P_A(0)$ with p -value < 0.05 . Another two-sample t-tests also confirm that GMM noise can improve the accuracy over the Gaussian noise: $P_{A_n}(\lambda) > P_{A_n}(\sigma)$ with p -value < 0.05 for Subjects A and B. For Subject C, both GMM and Gaussian noise can give perfect classification.

We compare the proposed method with other existing methods [30,32,33,40] as shown in Table 3. The results show that optimal noise density can enhance the accuracy of the BCI systems in general. This implies that we can reduce the number of signal repetitions or the collection time and so can speed up the responses with high accuracy.

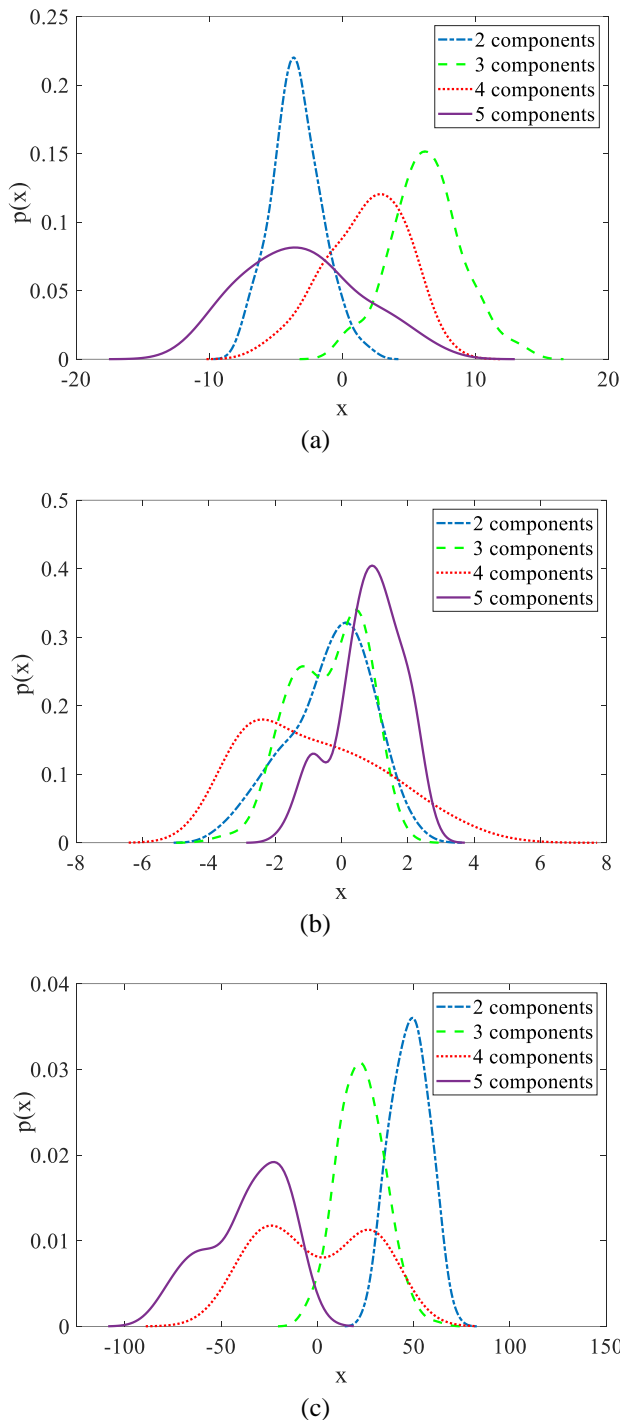


Figure. 3 SR-optimal noise pdf approximation using GMMs with 2, 3, 4 and 5 components for P300 classifications. (a)-(c) Each subject (A, B, or C) tend to use different optimal noise pdfs: (a) subject A, (b) subject B, and (c) subject C

5. Discussion and conclusion

Noise can enhance the performance of a nonlinear system. Experimental results show that different noise pdf's lead to different system responses and that affect its performance. Addition of noise from a pdf that matches the system's

characteristics can maximally enhance the performance. Statistical tests confirm that GMM noise can provide an approximation to any noise pdf that provide another boost to the complex classification problems.

The results lead to a call for investigations to other signal processing applications that can also benefit from noise. A signal processing application may require a more complex setting of noise addition as well as its pdf that may also depend on the system parameters or different stage of the system. Solving this complex optimization problem to obtain an exact solution can be discouraging. A method or an algorithm for approximation of a local solution still remains a preferred choice.

The use of GMM noise pdf offers such a solution to the design of any SR-optimal system for practical use. A nonlinear system that processes GMM noise can be difficult to obtain a closed form analysis. But we can approximate or search for locally optimal solutions using existing optimization techniques. PSO is an example of such algorithm that gives local solution to this complex optimization problem. Other techniques may be more suitable and more efficient to obtain the SR-optimal parameters. The most efficient technique on how to maximize the system performance remains open research problems.

Acknowledgements

R. Sampanna would like to acknowledge the support from Bangkok University, Thailand.

References

- [1] B. Kosko, *Noise*, Viking, 2006.
- [2] S. Mitaim and B. Kosko, "Adaptive Stochastic Resonance", In: *Proc. of the IEEE*, Vol.86, No.11, pp.2152–2183, 1998.
- [3] S. Mitaim and B. Kosko, "Adaptive Stochastic Resonance in Noisy Neurons Based on Mutual Information", *IEEE Transactions on Neural Networks*, Vol.15, pp.1526–1540, 2004.
- [4] O. Osoba, S. Mitaim, and B. Kosko, "The Noisy Expectation-Maximization Algorithm", *Fluctuation and Noise Letters*, Vol.12, No.03, pp.1350012, 2013.
- [5] A. Patel and B. Kosko, "Optimal Mean-Square Noise Benefits in Quantizer-Array Linear Estimation", *IEEE Signal Processing Letters*, Vol.17, No.12, pp.1005–1009, 2010.

- [6] A. Patel and B. Kosko, "Noise Benefits in Quantizer-Array Correlation Detection and Watermark Decoding", *IEEE Transactions on Signal Processing*, Vol.59, pp.488–505, 2011.
- [7] S. Bayram, S. Gultekin, and S. Gezici, "Noise Enhanced Hypothesis-Testing According to Restricted Neyman-Pearson criterion", *Digital Signal Processing*, Vol.25, No.Supplement C, pp.17 – 27, 2014.
- [8] S. Liu, T. Yang, X. Zhang, X. Hu, and L. Xu, "Noise Enhanced Binary Hypothesis-Testing in a New Framework", *Digital Signal Processing*, Vol.41, No.Supplement C, pp.22 – 31, 2015.
- [9] H. G. Lee and S. C. Lee, "Nucleus Segmentation Using Gaussian Mixture based Shape Models", *IEEE Journal of Biomedical and Health Informatics*, Vol.22, No.1, pp.235–243, 2018.
- [10] S. Lee, C. H. Park, and J. H. Chang, "Improved Gaussian Mixture Regression Based on Pseudo Feature Generation Using Bootstrap in Blood Pressure Estimation", *IEEE Transactions on Industrial Informatics*, Vol.12, No.6, pp.2269–2280, 2016.
- [11] M. Niknejad, H. Rabbani, and M. Babaie-Zadeh, "Image Restoration Using Gaussian Mixture Models with Spatially Constrained Patch Clustering", *IEEE Transactions on Image Processing*, Vol.24, No.11, pp.3624–3636, 2015.
- [12] G. Tian, Y. Xia, Y. Zhang, and D. Feng, "Hybrid Genetic and Variational Expectation-Maximization Algorithm for Gaussian-Mixture-Model-Based Brain MR Image Segmentation", *IEEE Transactions on Information Technology in Biomedicine*, Vol.15, No.3, pp.373–380, 2011.
- [13] A. Shahidi Zandi, R. Tafreshi, M. Javidan, and G. A. Dumont, "Predicting Epileptic Seizures in Scalp EEG Based on a Variational Bayesian Gaussian Mixture Model of Zero-Crossing Intervals", *IEEE Transactions on Biomedical Engineering*, Vol.60, No.5, pp.1401–1413, 2013.
- [14] F. Lotte, M. Congedo, A. Lcuyer, F. Lamarche, and B. Arnaldi, "A Review of Classification Algorithms for EEG-Based Brain Computer Interfaces", *Journal of Neural Engineering*, Vol.4, No.2, pp.R1, 2007.
- [15] K. R. Muller, M. Krauledat, G. Dornhege, G. Curio, and B. Blankertz, "Machine Learning Techniques for Brain-Computer Interfaces", *Biomedical Technology*, Vol.49, No.1, pp.11–22, 2004.
- [16] J. R. Wolpaw, N. Birbaumer, D. J. McFarland, G. Pfurtscheller, and T. M. Vaughan, "Brain Computer Interfaces for Communication and Control", *Clinical Neurophysiology*, Vol.113, No.6, pp.767–791, 2002.
- [17] S. Amiri, A. Rabbi, L. Azinfar, and R. Fazel-Rezai, "A Review of P300, SSVEP, and Hybrid P300/SSVEP Brain-Computer Interface Systems", *Brain-Computer Interface Systems, Chapter 10. IntechOpen*, 2013.
- [18] L. Bonnet, F. Lotte, and A. Lcuyer, "Two Brains, One Game: Design and Evaluation of a Multiuser BCI Video Game Based on Motor Imagery", *IEEE Transactions on Computational Intelligence and AI in Games*, Vol.5, No.2, pp.185–198, 2013.
- [19] J. Long, J. Wang, and T. Yu, "An Efficient Framework for EEG Analysis with Application to Hybrid Brain Computer Interfaces Based on Motor Imagery and P300", *Computational Intelligence and Neuroscience*, Vol.2017, pp.1–6, 2017.
- [20] R. A. Ramadan and A. V. Vasilakos, "Brain Computer Interface: Control Signals Review", *Neurocomputing*, Vol.223, pp.26 – 44, 2017.
- [21] S. He, R. Zhang, Q. Wang, Y. Chen, T. Yang, Z. Feng, Y. Zhang, M. Shao, and Y. Li, "A P300-Based Threshold-Free Brain Switch and Its Application in Wheelchair Control", *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, Vol.25, No.6, pp.715–725, 2017.
- [22] M. K. Mukerjee, "NeuroPhone: Brain-Mobile Phone Interface using a Wireless EEG Headset", *Tech. Rep. TR2010-666*, Dartmouth College, Computer Science, Hanover, NH, 2010.
- [23] A. F. Perez Vidal, M. A. Oliver Salazar, and G. Salas Lopez, "Development of a Brain-Computer Interface Based on Visual Stimuli for the Movement of a Robot Joints", *IEEE Latin America Transactions*, Vol.14, No.2, pp.477–484, 2016.
- [24] H. Nakamura, A. Matsubara, and S. Nishifuji, "Noise-assisted auditory brain computer interface", *In: Proc. of International Conf. on Consumer Electronics*, pp. 1–5, 2017.
- [25] S. Nishifuji, A. Matsubara, Y. Sugita, A. Iwata, H. Nakamura, and H. Hirano, "Eyes-closed brain computer interface using modulation of steady state visually evoked potential and auditory steady-state response", *In: Proc. of International Conf. on Instrument and Control Engineers of Japan*, pp.1138–1143, 2017.

- [26] L. F. Rebolledo-Herrera and F. G. Espinosa, "Novel parameter tuned methodology for under-damped stochastic resonance applied to EEG signal enhancement", In: *Proc. of International Conf. on Systems, Man, and Cybernetics*, pp.002128–002132, 2016.
- [27] R. Sampanna and S. Mitaim, "Noise Enhanced Array Signal Detection in P300 Speller Paradigm using ICA-Based Subspace Projections", In: *Proc. of International Conf. on Medicine and Biology Society*, pp.4239–4242, 2013.
- [28] R. Sampanna and S. Mitaim, "Noise Benefits in Motor Imagery Classification using Ensemble Support Vector Machine", In: *Proc. of International Conf. on Biomedical Circuits and Systems Conference*, pp.53–56, 2014.
- [29] R. Sampanna and S. Mitaim, "Noise Benefits in the Array of Brain-Computer Interface Classification Systems", *Informatics in Medicine Unlocked*, Vol.12, pp.88 – 97, 2018.
- [30] H. Cecotti and A. Graser, "Convolutional Neural Networks for P300 Detection with Application to Brain-Computer Interfaces", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol.33, pp.433–445, 2011.
- [31] U. Hoffmann, G. Garcia, J. M. Vesin, K. Diserens, and T. Ebrahimi, "A Boosting Approach to P300 Detection with Application to Brain-Computer Interfaces", In: *Proc. of International Conf. on Neural Engineering*, pp. 97–100, 2005.
- [32] M. Kaper, P. Meinicke, U. Grossekhoefer, T. Lingner, and H. Ritter, "BCI Competition 2003-Data Set IIB: Support Vector Machines for the P300 Speller Paradigm", *IEEE Transactions on Biomedical Engineering*, Vol.51, pp.1073–1076, 2004.
- [33] A. Rakotomamonjy and V. Guigue, "BCI Competition III: Dataset II Ensemble of SVMs for BCI P300 Speller", *IEEE Transactions on Biomedical Engineering*, Vol.55, pp.1147–1154, 2007.
- [34] B. Blankertz, "BCI Competition 2003, [online]. available: <http://www.bbci.de/competition/ii/>," 2003.
- [35] B. Blankertz, "BCI Competition 2004, [online]. available: <http://www.bbci.de/competition/iii/>," 2004.
- [36] O. Osoba and B. Kosko, "The Noisy Expectation-Maximization Algorithm for Multiplicative Noise Injection", *Fluctuation and Noise Letters*, Vol.15, No.01, pp.1650007, 2016.
- [37] E. K. P. Chong and S. H. Zak, *An Introduction to Optimization 4th Edition*, Wiley, 2013.
- [38] S. S. Rao, *Engineering Optimization: Theory and practices*, Wiley, 1996.
- [39] A. Ravindran, K. M. Ragsdell, and G. V. Reklaitis, *Engineering Optimization: Methods and Applications, 2nd Edition*, Wiley, 2006.
- [40] N. Xu, X. Gao, B. Hong, X. Miao, S. Gao and F. Yang, "BCI competition 2003-data set IIB: enhancing P300 wave detection using ICA-based subspace projections for BCI applications", *IEEE Transactions on Biomedical Engineering*, Vol.51, No. 6, pp.1067-1072, 2004.