Informatics in Medicine Unlocked 12 (2018) 88–97

Contents lists available at ScienceDirect

Informatics in Medicine Unlocked

journal homepage: www.elsevier.com/locate/imu

Noise benefits in the array of brain-computer interface classification systems

Rujipan Sampanna*, Sanya Mitaim

The Department of Electrical and Computer Engineering, Faculty of Engineering, Thammasat University, Thailand

ARTICLE INFO

Keywords:
P300
Motor imagery
Stochastic resonance
Array systems
Ensemble support vector machine
Particle swarm optimization

ABSTRACT

This paper shows that noise can improve the accuracy of brain-computer interface (BCI) systems. Additive Gaussian noise can benefit arrays of ensemble support vector machines (ESVMs) that classify P300 or motor imagery (MI) activities in electroencephalogram (EEG) signals. We show these noise benefits in 64-channel EEG signals from the BCI Competitions II dataset IIb and BCI Competitions III dataset II for the P300 speller paradigm and in 3-channel EEG signals from the BCI Competitions II dataset III and BCI Competitions III dataset IIIa for MI classification systems. We also show that noise can improve the accuracy of EEG classifications based on restricted channel positions in commercial recording systems, such as the 14-channel Emotiv Epoc headset. The experimental results show that noise can provide classifiers with higher accuracy and can reduce the data collection time for P300 classification. The results also show that training ESVMs with a concatenated original dataset and noise-added datasets can improve MI classification. Noise can improve the accuracy of P300 classification for both intra-subject and inter-subject classification systems for multiple users. Addition of noise can significantly affect the parameters of polynomial kernel functions and the number of support vectors of the SVM. This leads to an expansion of the margin between two parallel hyperplanes that eventually improve the classification accuracy. Particle swarm optimization (PSO) can be used to search for the optimal noise intensity.

1. Introduction

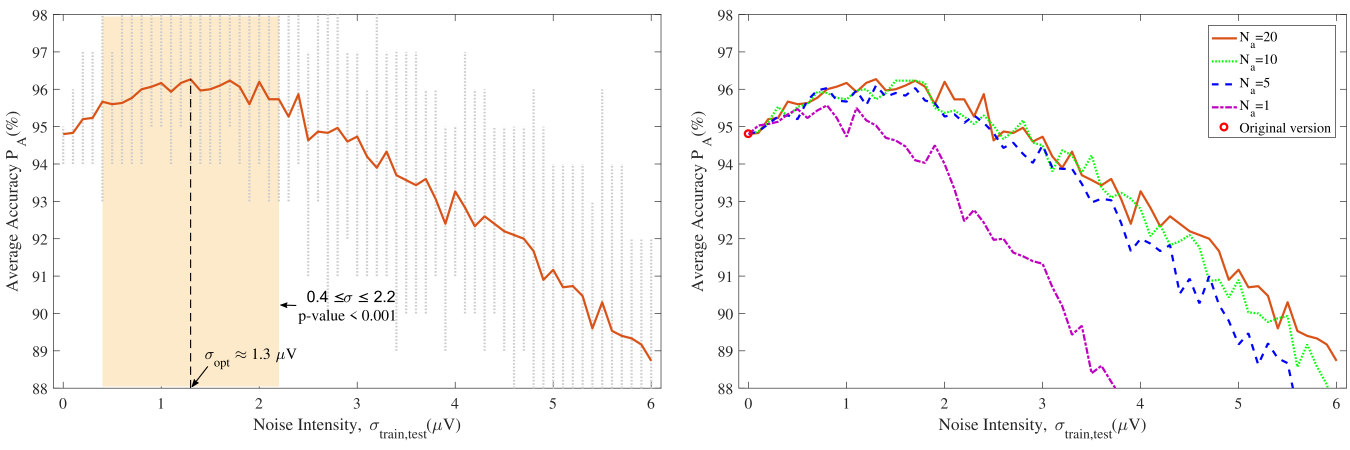
The recently developed brain-computer interfaces (BCIs) use a variety of brain activity phenomenon that have the potential to improve the communication performance of paralyzed people [1–3]. BCIs are a communication system that allows a person to send commands to an external device through direct measurements of brain activities without using any movement. Signal sources include electroencephalograms (EEGs), functional magnetic resonance images (fMRIs), position emission resonance tomography (PET) and magnetoencephalograms (MEGs) [4]. Most BCI researchers focus their attention on EEG-based BCIs because of its non-invasive recordings, low cost, and relative simplicity [4]. There are many types of brain activity patterns in the EEG signals, i.e., slow cortical potential (SCP), P300, steady state visual evoked potential (SSVEP) and motor imagery (MI) [1,5]. A user must produce different brain activity patterns that will be classified and translated into commands.

P300 and MI are important types of EEG signals for BCI applications [5–9]. The applications of P300 include the P300 speller paradigm, neurophones, wheelchair control and robotic arms [8–12]. The purpose of P300 applications is to detect the presence of P300 in an EEG. MI applications include robot control, wheelchair control, car game control, and rehabilitation [6–9]. The purpose of this paradigm is the discrimination of MI tasks such as left hand, right hand, finger, and foot movements.

The performance of a classifier depends on the features and classification algorithm. Popular choices of features include signal amplitudes, wavelet transform coefficients, and autoregressive (AR) model coefficients [1,13]. Independent component analysis (ICA) can remove artifacts and enhance P300 signals [14,15]. Tensor decomposition can find an optimal feature subspace [16]. Classification techniques include support vector machines (SVMs), linear discriminant analysis (LDA), and neural networks [17–19]. Ensemble support vector machines (ESVMs) [17,18] and convolution neural networks (CNNs) [20] better solve the problems of P300 signal variation from different subjects than other techniques [1,18,20]. Other research on EEG classification improvement also includes POMDP approach to optimizing P300 speller BCI paradigm [21], target selection with hybrid feature for BCI-based 2-D cursor control [22], sparse Bayesian classification of EEG for brain-computer interface [23], linked component analysis from matrices to high order tensors with applications to biomedical data [24], multi-kernel extreme learning machine for EEG classification in brain-computer interface [25], feature weighting and regularization of common spatial patterns in EEG-based Motor Imagery BCI [26], and a transform-

* Corresponding author.
E-mail addresses: rujipan.s@bu.ac.th (R. Sampanna), msanya@engr.tu.ac.th (S. Mitaim).

https://doi.org/10.1016/j.imu.2018.07.003
Received 16 April 2018; Received in revised form 19 June 2018; Accepted 8 July 2018
Available online 12 July 2018
2352-9148/



(a) (b)

Fig. 1. Noise enhances the accuracy of P300 classification. The system is tested on the BCI competition III (B) dataset using N_r = 14 signal repetitions and 64 channels. We repeat each test 30 times and show the average. (a) The system uses an array of ESVM classifiers with N_a = 20 stages. The vertical dashed lines show the variation in noise realizations. The system shows an optimal point at (approximately) $\sigma_{opt} \approx 1.3\,\mu V$ and a classification accuracy of 96.27%. Two-sample t-tests show that there is noise benefit with p-value <0.001 for the noise intensity $0.4\,\mu V \le \sigma \le 2.2\,\mu V$. (b) Classification accuracy increases as the number of stages N_a increases.

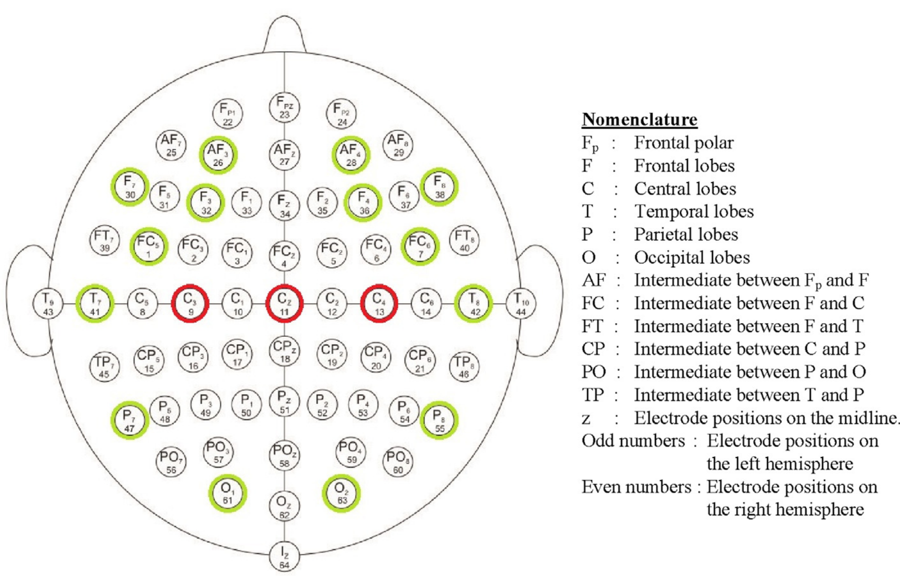


Fig. 2. The channel assignment numbers based on the international 10–20 system (64 channels) and the nomenclature of the positions. The red circles correspond to the 3 channels for recording the MI signals. The green circles correspond to the 14 channels of the Emotiv Epoc headset. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

based feature extraction approach for motor imagery classification [27].

This paper shows how to apply Gaussian noise to improve the accuracy of P300 and MI classification systems. The study uses Gaussian noise although other types of noise such as uniform and Laplace noise can also enhance the performance of BCI systems or other nonlinear systems. For example uniform noise can increase the accuracy of P300 classification for BCI competition III subject A ($\sigma_{opt} = 0.1\,\mu V$ and $N_r = 14$) from 95% to 95.30%. We use Gaussian noise to first show that optimal level of noise can increase the performance of the system. Then we show how optimization techniques such as particle swarm optimization (PSO) can find that optimal level.

Fig. 1 shows an example of the stochastic resonance (SR) effect of a P300 classification system. This SR effect is a phenomenon where noise at appropriate intensity levels can enhance weak input signals [28–31]. The SR effect can also occur in array systems [14,32–35]. Patel and Kosko show that noise can improve statistical signal detection for array-based nonlinear correlators in Neyman-Pearson (NP) and maximum-likelihood (ML) signal detection [33]. They also show that the 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 [36]. Bayram *et al.* show that the optimal additive noise can be represented by a discrete random variable with a certain number of point masses. Liu *et al.* show a noise enhanced 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} [37]. Zhang *et al.* present array-enhanced logical stochastic resonance, where increasing the number of arrays can extend the range of the optimal parameter domain in which a reliable logic output can be obtained [35].

The next section describes the brain-computer interface (BCI) paradigms. Section III 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 [14,17,18] with additive Gaussian noise in the training phase and in the testing phase (or in actual use). The final section shows the extensive experimental results on the BCI competition II dataset IIb (P300) and dataset III (MI) [38,39] and the BCI competition III dataset II (P300) and dataset IIIa (MI) [40]. Then, we show that particle swarm optimization (PSO) [41–43] can

Table 1
Dataset of P300 Speller paradigm.

Dataset	Number of	Number of Signals		
	Characters	Target	Non-target	All Stimulus
BCI IIIA (A)				
Training:	85	2550	12,750	15,300
Testing:	100	3000	15,000	18,000
BCI IIIB (B)				
Training:	85	2550	12,750	15,300
Testing:	100	3000	15,000	18,000
BCI II (C)				
Training:	42	1260	6300	7560
Testing:	31	930	4650	5580

search for the optimal noise intensity.

2. BCI paradigms

BCI systems interpret one or more brain activity patterns in EEG signals [5]: SCP, P300, SSVEP, and MI. Fig. 2 shows the 64 channel assignments based on the international 10–20 system [39,44], the corresponding 3 channels for MI signals, and an example of a commercial EEG acquisition device, such as the 14-channel Emotive Epoc headset [45]. Below we briefly discuss the EEG signals from international databases for P300 speller paradigm and MI paradigm [38,40].

Table 1 gives descriptions of a P300 dataset recorded from 64 electrodes, as shown in Fig. 2, using a 240-Hz sampling rate and 0.1–60-Hz filter. Each dataset represents a set of characters where each character was repeated 15 times ($N_r = 15$ signal repetitions) to reinforce the P300 responses [38–40]. We denote the datasets A, B, and C for the BCI competition III dataset II (Subject A), BCI competition III dataset II (Subject B), and BCI competition II dataset IIb (Subject C).

Table 2 gives descriptions of the dataset of MI paradigm, which consists of 2 classes (right hand movement and left hand movement). There are 346 training trials and 340 testing trials that record from 3 electrodes (C3,Cz,C4) with a 128-Hz sampling rate and 0.5–30-Hz filter for BCI II dataset III [38] and 64 electrodes with a 250-Hz sampling rate and 1–50-Hz filter for BCI III dataset IIIa [40].

3. EEG classification systems

We propose a “noisy” EEG classification system, as shown in Fig. 3, for both P300 classification and MI classification systems. The classification system is an array of conventional classification systems with additive noise. The array system consists of N_a identical systems or stages of P300 or MI classification systems as shown in Fig. 3. Each stage processes the same filtered signal x' of the raw EEG data x that are obtained from the sensors. Then, each stage adds independent noise n_i to the signal x' to obtain the noise-added signal x''_i: $x''_i = x' + n_i$, where n_i is independent Gaussian noise with identical pdf at stage i. Other noise densities can also enhance the BCI systems or other nonlinear systems [28–31]. Training the system with known (labeled) 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.

Table 2
Dataset of MI paradigm.

Dataset	Number of	Number of Signals	
	Trials	Left Hand	Right Hand
		Movement	Movement
BCI II (III)			
Training:	140	70	70
Testing:	140	70	70
BCI III (IIIa):	Subject l1b		
Training:	60	30	30
Testing:	60	30	30
BCI III (IIIa):	Subject k3b		
Training:	90	45	45
Testing:	84	41	43
BCI III (IIIa):	Subject k6b		
Training:	56	28	28
Testing:	56	29	27

3.1. ESVM classifier

We use an ESVM as a classifier since ESVMs combine several support vector machines (SVM) classifiers to solve problems associated with signal variations between subjects and over time [18]. 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, Lagrangian multipliers α_{kj}, $K_{kj}(E_{kj}, x) = (E_{kj}^T x + 1)^n$ is the polynomial kernel function, n is the degree of a polynomial, and support vectors E_{kj}, $j = 1,2, \ldots, L_k$, 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:

$$f_k(x) = sgn\left(\sum_{j=1}^{L_k} y_{kj}\alpha_{kj}K_{kj}(E_{kj}, x) + b_k\right) \tag{1}$$

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

Both the P300 and MI paradigms imply a two-class classification problem with class labels $y_i \in \{-1,1\}$: P300 as “present” and “absent” and MI as “left” or “right”. Thus, the output of the ESVM classifier is a sign of the sum of M SVM classifiers:

$$\hat{y} = S(x) = sgn\left(\sum_{k=1}^{M} f_k(x)\right) \tag{2}$$

We test the classifier with different numbers of clusters and signals, as shown in Table 3. We use the best cluster number M that leads to the highest accuracy to examine the noise benefits.

3.2. P300 signal features

We use the EEG signal x in the time window of 0–667 ms after the stimulus [18]. 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. For SR effect, we add Gaussian noise n to x' before the feature extraction. The noise intensity varies from 1 μV to 10 μV.

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 [18]:

$$c = \arg\max_{1\le i\le 6} \sum_{e=1}^{N_r} \hat{y}_{ie} \tag{3}$$

$$r = \arg\max_{7\le i\le 12} \sum_{e=1}^{N_r} \hat{y}_{ie} \tag{4}$$

where $c \in \{1,2,\ldots,6\}$ and $r \in \{7,8,\ldots,12\}$ are the column and row numbers

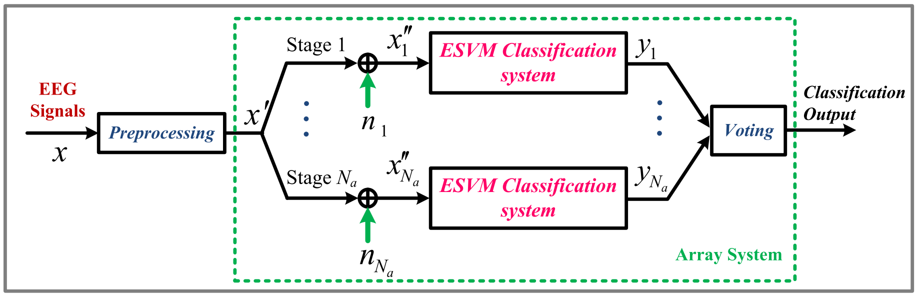


Fig. 3. A diagram of the proposed noisy array of the ESVM classifications systems for P300 and MI classification.

Table 3
Numbers of clusters used in the classification of the P300 and MI datasets.

Dataset	# Clusters (M)	# Signals per cluster	Total number of signals
P300			
BCI III dataset II subject A	17	5	85
BCI III dataset II subject B	17	5	85
BCI II dataset IIb	8	5	40
MI			
BCI II dataset III	5	28	140
BCI III dataset IIIa subject l1b	3	20	60
BCI III dataset IIIa subject k3b	3	30	90
BCI III dataset IIIa subject k6b	4	14	56

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

3.3. Motor imagery signal features

Fig. 4 shows the electrode positions and feature extraction process. We compute the instantaneous powers $P_t = x_t^2$ of the EEG signal x_t during 3–9 s after the order of the left and right cue onset [38,40,46] and use them to obtain the coefficients of the discrete wavelet transform with Daubechies 4 (db4) coefficients [47]. We obtain 18 wavelet coefficients from the 0–1 Hz frequency band [47]. We use the difference in the approximation coefficients of electrode C3 ($A6C3$) and C4 ($A6C4$): $DCA6 = A6C3 - A6C4$. We also obtain the coefficients of the 6^{th} -order AR model [48] from the 3–9 s duration of the EEG signals x'. We use the AR coefficients of two electrodes (C3 and C4).

Thus, the features for each MI signal consist of 18 wavelet coefficients ($DCA6$) and 12 AR coefficients ($ARC3$ = 6 features and $ARC4$ = 6 features). The ESVM uses these 30 features C_x for classification as follows:

$$C_x = [DCA6_1, \cdots, DCA6_{18}, ARC3_1, \cdots, ARC3_6, ARC4_1, \cdots, ARC4_6] \quad (5)$$

3.4. Optimal noise search

Optimal noise intensity depends on the characteristics of signal systems and the classification systems, such as features, classifiers, number of signal repetitions, and number of stages in the array. Determination of the optimal noise level requires solving a complex optimization problem. We can consider selected parameters of the noise pdf and test several realizations to observe the trend and approximate the optimal noise parameter, such as the standard deviation, as shown in Fig. 1.

Many optimization techniques can find a local solution: simulated annealing, genetic algorithm, ant colony optimization, fuzzy optimization and PSO [41–43]. Here, we show how PSO can search for optimal noise intensity or standard deviation σ. Since PSO is the fast and effective optimization technique for resolving complex optimization problem [41–43]. The objective function of the ESVM method is to maximize the classification accuracy P_A subject to the noise intensity σ between the lower bound σ_l and upper bound σ_u [41–43]:

$$\text{Maximize } P_A(\sigma) \quad \text{subject to } \sigma_l \leq \sigma \leq \sigma_u. \quad (6)$$

4. Experimental results

We test the proposed system with EEG data from the BCI competition II dataset IIb (P300) and dataset III (MI) [38,39] and the BCI competition III dataset II (P300) and dataset IIIa (MI) [40]. We use an original classification system as a building block for an array system, as shown in Fig. 3. Then, we add Gaussian noise with the standard deviation σ to the original signal data x'. The noise in each block is independent of that in the others. We also examine the SR effects with three options of noise additions:

Case 1: Adding noise in the training phase.
Case 2: Adding noise in the testing phase.
Case 3: Adding noise in the training and testing phases.

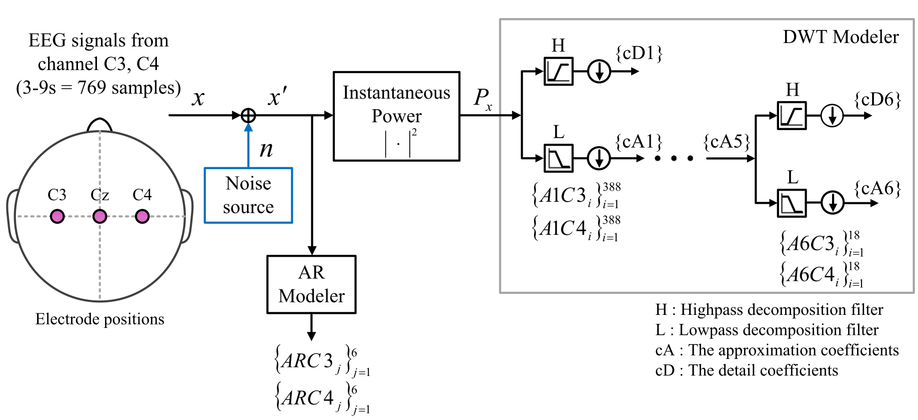
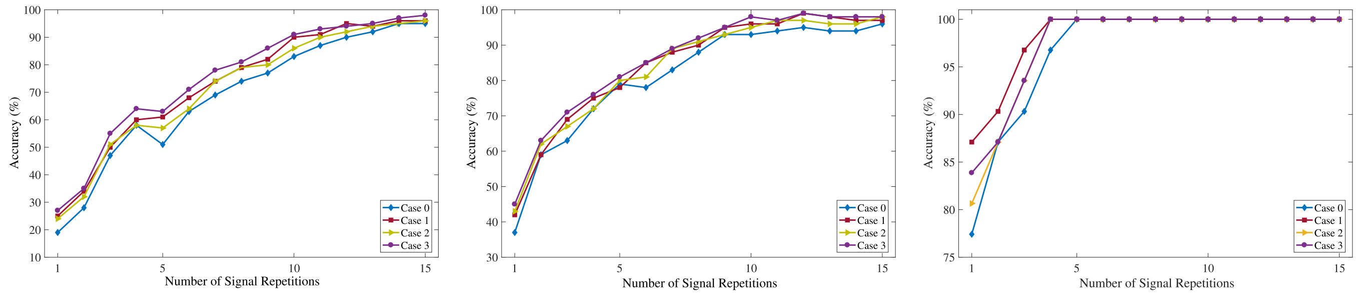


Fig. 4. Electrode positions (C3, Cz, C4) and feature extraction process for MI.



(a) Dataset A (b) Dataset B (c) Dataset C

Fig. 5. P300 classification for datasets A, B, and C using different noise-adding strategies, e.g., Original version (or Case 0): training and testing without noise, Case 1: noise in the training phase, Case 2: noise in the testing phase, and Case 3: noise in the training and testing phases. The classification accuracy of Case 3 is the highest. The experiments use noise intensities from 1 μV to 10 μV and plot the highest value for each repetition number.

Table 4
Classification accuracy of the test datasets for the MI classification system.

Dataset	# Sets in training phase	Train without noise: Test without Noise P_A (%)	Test with noise P_{An} (%) $N_a = 1$	$N_a = 5$	$N_a = 10$	Train with noise: Test without Noise P_{An} (%)	Test with noise P_{An} (%) $N_a = 1$	$N_a = 5$	$N_a = 10$	$\sigma_{opt}(\mu$V)	Increase $P_{A_n} - P_A$ (%)
BCI II (III)	1	86.43	85.29	85.71	85.86	86.43	86.14	86.57	86.57		
	5	91.43	89.71	92.14	91.29	92.14	90.57	**93.57**	92.00		
	10	90.71	90.14	90.71	90.43	92.86	91.29	92.86	92.43		
	Max	92.86	90.14	92.86	92.00	92.86	91.43	**93.57**	92.71	0.003	7.14
BCI III (IIIa) Subject l1b	1	60.00	62.67	61.00	64.67	63.33	65.33	63.33	66.00		
	4	63.33	63.67	65.00	64.33	68.33	67.67	67.67	**70.00**		
	10	61.67	63.33	62.00	63.00	66.67	68.67	69.00	67.00		
	Max	63.33	64.33	65.00	65.00	68.33	68.67	69.33	**70.00**	0.02	10
BCI III (IIIa) Subject k3b	1	55.95	56.19	56.43	56.67	59.52	60.48	59.52	57.38		
	5	54.76	55.71	56.19	56.43	59.52	59.52	58.33	60.24		
	10	54.76	56.19	56.43	56.19	59.52	60.00	**61.90**	60.71		
	Max	57.14	57.14	57.38	57.14	60.71	60.71	**61.90**	60.71	0.08	5.95
BCI III (IIIa) Subject k6b	1	51.79	51.79	51.79	51.79	51.79	51.79	51.79	51.79		
	6	48.21	51.79	51.43	50.00	55.36	58.57	58.93	**61.43**		
	10	51.79	52.50	52.50	52.86	57.14	57.86	60.36	57.14		
	Max	55.36	56.79	56.79	57.86	60.71	60.71	61.07	**61.43**	0.1	9.64

The significance of bold in the table is the intra-subject test and the number of repetitions (Reps N_r) correspond to the maximum accuracy.

Then, 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:

$$P_A = \frac{N_C}{N_T} \times 100\% \tag{7}$$

In general, the classification accuracy may vary with noise realizations. We repeat each test for 30 times and find the average to determine the system performance P_A.

We tested the system using the following steps:

Step 1: Preprocessing and feature extraction. Use bandpass filtering and other feature extraction such as discrete wavelet transform.

Step 2: ESVM classifier training. Divide the training dataset into M clusters and use each cluster to train SVM classifier.

Step 3: Test for accuracy using three options of noise additions mentioned above.

4.1. P300 classification

We vary the noise intensity σ from 0.1 μV to 10 μV in the training phase and testing phase. We test the system based on 1 to 15 signal repetitions. Fig. 1 (a) shows the SR effect (Case 3: Adding noise in training and testing phase) of an array P300 classification system using an ESVM with a number of stages $N_a = 20$ and a number of signal repetitions $N_r = 14$. Fig. 5 shows the classification accuracy of datasets A, B and C using different noise-adding strategies, e.g., Original version: training and testing without noise versus Case 1- Case 3. The accuracy of a noise-added system is higher than the system without noise in all cases.

Fig. 1 (b) shows that the classification accuracy improves with the increasing number of stages $N_a = 1,5,10$, and 20. The accuracy of this system (regular ESVM classification with 14 signal repetitions and a number of stages $N_a = 1$) increases when we add more stages and uses a suitable level of noise intensity. The number of stages $N_a = 20$ gives the maximum accuracy at 96.27% and gives 1.47% improvement over the noiseless classification system at 94.80%. We use two-sample t-tests to confirm that there is an increase of prediction accuracy or noise benefit: $P_A(\sigma) > P_A(0)$ with p-value <0.001 for the noise intensity $0.4\mu \text{ V} \le \sigma \le 2.2\mu$V.

Table 5 shows the classification accuracy of the array systems based on various numbers of signal repetitions N_r and array sizes N_a. The results show the maximum classification accuracy in each case with the corresponding optimal noise intensity (μV). The classification accuracy with noise is higher than that without noise in all cases. The results also show that the classification accuracy tends to increase as the number of stages N_a increases from $N_a = 1$ to $N_a = 10$. We obtain the maximum accuracy values at 97.60% with optimal noise ($\sigma_{opt} = 0.7$ μV) for $N_a = 10$ and $N_r = 15$ signal repetitions, 98.00% with optimal noise ($\sigma_{opt} = 1.3$ μV) for $N_a = 5$ and $N_r = 15$ signal repetitions, and 100% with optimal noise ($\sigma_{opt} = 3.5$ μV) for $N_a = 5$ and $N_r = 4$ signal repetitions for datasets A, B and C, respectively.

Fig. 6 shows the noise benefits in the array system from the study of

Table 5
Classification accuracy of the test datasets for multi-user experiments (64 channels): P300 Classification using an ESVM.

Training set	Testing set	Reps N_r	Train without noise				Train with noise				σ_{opt} (μV)	Increase $P_{An} - P_A$(%)
			Test without noise P_A (%)	Test with noise P_{An} (%)			Test without noise P_{An} (%)	Test with noise P_{An} (%)				
				N_a = 1	N_a = 5	N_a = 10		N_a = 1	N_a = 5	N_a = 10		
A	**A**	14	**95.00**	94.60	95.40	95.80	93.00	94.80	94.80	95.00		
		15	**95.00**	**95.40**	**97.20**	**97.40**	**97.00**	**95.20**	**96.40**	**97.60**		
		Max	95.00	95.40	97.20	97.40	97.00	95.20	96.40	**97.60**	0.7	2.60
B	**B**	12	95.00	96.40	97.20	**97.60**	96.00	96.60	97.20	97.60		
		15	**96.00**	**96.80**	**97.40**	**97.60**	**98.00**	**97.80**	**98.00**	**98.00**		
		Max	96.00	96.80	97.40	97.60	**98.00**	97.80	**98.00**	**98.00**	1.3	2.00
C	**C**	**4**	96.77	98.71	98.06	99.35	**100**	99.35	**100**	**100**		
		15	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**		
		Max	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**	3.5	0
A,B	**A**	10	76.00	78.20	80.20	80.40	76.00	79.00	80.20	77.40		
		15	**90.00**	**91.40**	**92.60**	**92.60**	**91.00**	**90.80**	**93.40**	**93.60**		
		Max	90.00	91.40	92.60	92.60	91.00	90.80	93.40	**93.60**	0.91	3.60
	B	11	**96.00**	96.60	96.40	96.00	95.00	96.00	96.40	96.40		
		12	**96.00**	**97.80**	**98.00**	**98.00**	**97.00**	**97.60**	**98.00**	**98.00**		
		15	95.00	96.20	96.00	95.80	94.00	**97.60**	97.60	97.40		
		Max	96.00	97.80	**98.00**	**98.00**	97.00	97.60	**98.00**	**98.00**	0.91	2.00
	C	**11**	**80.65**	**83.87**	**80.65**	**80.65**	74.19	**82.58**	**83.87**	**83.87**		
		12	77.42	77.42	77.42	77.42	**83.87**	79.35	77.42	77.42		
		15	77.42	77.42	77.42	77.42	74.19	80.65	80.65	80.65		
		Max	80.65	**83.87**	80.65	0.00	**83.87**	82.58	**83.87**	**83.87**	0.91	7.10
B,C	A	**12**	25.00	29.40	28.40	28.00	**36.00**	34.20	**36.20**	**36.60**		
		13	25.00	29.80	29.40	29.20	**36.00**	**34.40**	33.40	33.40		
		15	**34.00**	**32.40**	**35.00**	**35.00**	31.00	31.80	32.40	32.80		
		Max	34.00	32.40	35.00	35.00	36.00	34.40	36.20	**36.60**	0.84	2.60
	B	12	**97.00**	**97.60**	**98.00**	**98.00**	95.00	97.20	98.00	97.80		
		14	96.00	97.20	97.00	96.80	96.00	**98.20**	**98.60**	**98.60**		
		15	96.00	97.40	**98.00**	97.80	**98.00**	97.00	97.40	97.20		
		Max	97.00	97.60	98.00	98.00	98.00	98.20	**98.60**	**98.60**	0.84	1.60
	C	**9**	87.10	87.10	87.10	87.10	**100**	**100**	**100**	**100**		
		10	**96.77**	**96.77**	**96.77**	**96.77**	96.77	96.77	96.77	96.77		
		15	**96.77**	**96.77**	**96.77**	**96.77**	90.32	96.77	96.77	96.77		
		Max	96.77	96.77	96.77	96.77	**100**	**100**	**100**	**100**	0.84	3.23
A,C	**A**	14	**93.00**	91.80	93.60	94.20	87.00	90.80	93.80	93.20		
		15	**93.00**	**93.20**	**96.00**	**96.20**	**94.00**	**92.80**	**96.00**	**96.40**		
		Max	93.00	93.20	96.00	96.20	94.00	92.80	96.00	**96.40**	1.36	3.40
	B	13	31.00	33.20	35.60	36.00	**43.00**	42.20	43.40	44.00		
		14	**42.00**	**41.20**	**42.80**	**42.40**	40.00	45.40	47.20	46.80		
		15	38.00	38.80	41.20	40.00	34.00	**45.60**	**47.80**	**48.20**		
		Max	42.00	41.20	42.80	42.40	43.00	45.60	47.80	**48.20**	0.96	6.20
	C	**8**	**96.77**	96.77	96.77	96.77	93.55	95.48	**100**	94.19		
		13	**96.77**	96.77	96.77	96.77	**100**	**100**	**100**	**100**		
		15	**96.77**	**96.77**	**97.42**	**98.06**	93.55	95.48	**100**	98.06		
		Max	96.77	96.77	97.42	98.06	**100**	**100**	**100**	**100**	2.56	3.23
A,B,C	**A**	**14**	83.00	86.20	88.80	88.20	84.00	87.80	**91.40**	**91.40**		
		15	**86.00**	**86.60**	**90.20**	**90.20**	**88.00**	**88.80**	**91.40**	91.20		
		Max	86.00	86.60	90.20	90.20	88.00	88.80	**91.40**	**91.40**	1.27	5.40
	B	11	**97.00**	97.00	97.00	97.00	96.00	96.40	96.60	96.00		
		12	96.00	**97.60**	**97.60**	**97.80**	97.00	**97.80**	**98.00**	97.40		
		15	94.00	97.00	96.80	97.00	**98.00**	97.00	97.40	**98.20**		
		Max	97.00	97.60	97.60	97.80	98.00	97.80	98.00	**98.20**	0.96	1.20
	C	**9**	87.10	89.03	88.39	87.74	93.55	90.32	90.32	**96.77**		
		10	87.10	88.39	88.39	89.03	**96.77**	94.84	96.13	**96.77**		
		12	**96.77**	**96.77**	**96.77**	**96.77**	**96.77**	**96.77**	**96.77**	**96.77**		
		15	**96.77**	**96.77**	**96.77**	**96.77**	**96.77**	**96.77**	**96.77**	**96.77**		
		Max	**96.77**	**96.77**	**96.77**	**96.77**	**96.77**	**96.77**	**96.77**	**96.77**	1.27	0

The significance of bold in the table is the intra-subject test and the number of repetitions (Reps N_r) correspond to the maximum accuracy.

the four cases (testing dataset of BCI competition III (A)). Fig. 6(a–c) show that the accuracy of the classification in each case increases as the signal repetitions and the number of stages N_a increase. Fig. 6(d–f) show four noise-added cases with 1, 10 and 20 stages, respectively. The accuracy of Case 3 is higher than the accuracy in the other cases. The system shows that the optimal noise can benefit from finding the best classifier in training phase and can achieve the collective noise benefits of the array system in the testing phase.

4.2. Motor imagery classification

We vary the noise intensity σ from $0.0001\,\mu$V to $0.1\,\mu$V in the training phase and testing phase. The number of stages N_a also varies: N_a = 1, 5, and 10.

There are relatively few signals in the datasets. Thus, we concatenate the original dataset with m noise-added datasets in the training phase. Consequently, the total number of signals is $K_d = (m+1)K$, where $K = 140$ is the number of original data signals. We test several quantities ($m = 0,\ \ldots,9$) of noise-added datasets.

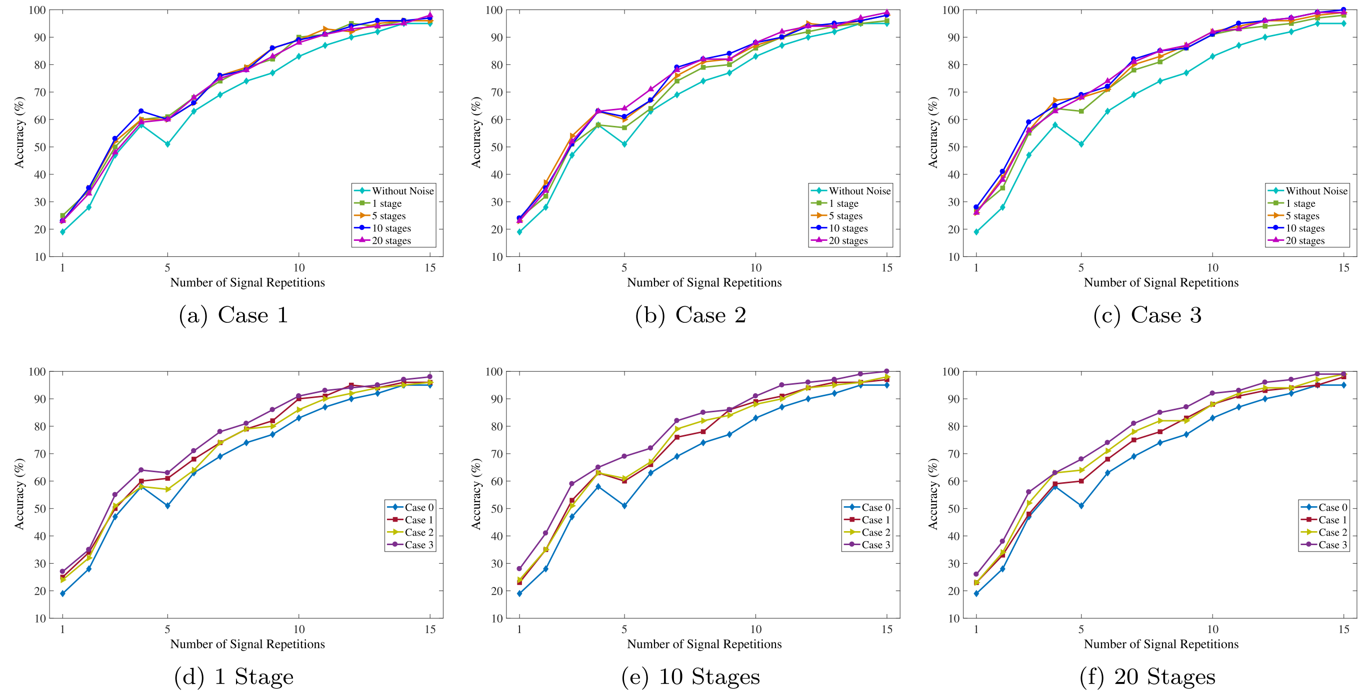


Fig. 6. Noise benefit in an array system of P300 (testing dataset of BCI competition III (A)). (a) Case 1: Adding noise in the training phase. (b) Case 2: Adding noise in the testing phase. (c) Case 3: Adding noise in the training and testing phases. (d–f) Four noise-added cases with N_a = 1, 10 and 20 stages, respectively. The results show that the accuracy of Case 3 is the highest and that the accuracy also increases with the number of stages N_a.

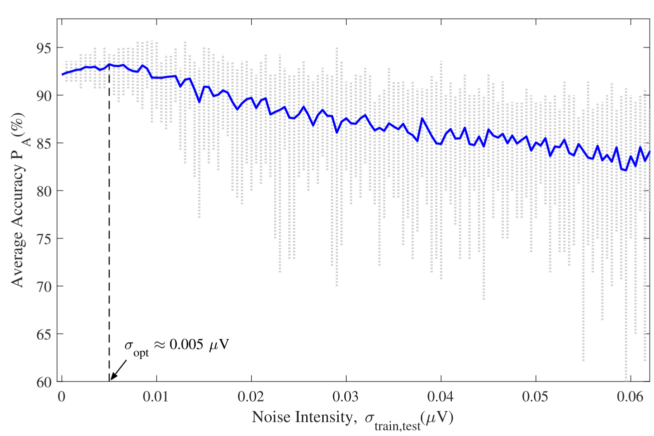


Fig. 7. Noise enhances the accuracy of MI classification. The system is tested on the BCI competition II (III) dataset using 5 sets of noise-added data. The noise intensity (standard deviation σ) varies from 0.0005 μV to 0.1 μV. The system uses an array of ESVM classifiers with N_a = 20 stages.

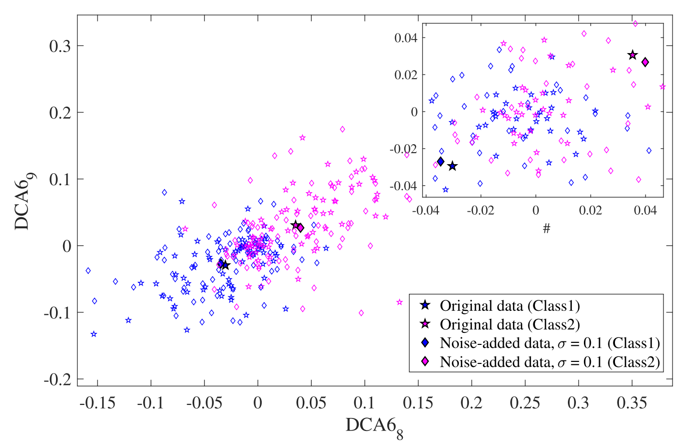


Fig. 8. The scatter plots of two features ($DCA6_8$ and $DCA6_9$) of MI signal features. Blue and magenta markers are features of right and left hand movements respectively. Pentagram markers are features from original data. Diamond markers are features from noise-added data with σ = 0.1. Solid markers are centroids of the respective data clusters. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

Fig. 7 shows an example of the SR effect on the MI classification system. Table 4 shows how array SR can increase the classification accuracy of an MI classification system. The classifier attained maximum accuracy at 93.57% for BCI II (III) (σ_{opt} = 0.003 μV, N_a = 5), 70% for BCI III (IIIa) Subject l1b (σ_{opt} = 0.02 μV, N_a = 10), 61.90% for BCI III (IIIa) Subject k3b (σ_{opt} = 0.08 μV, N_a = 5), and 61.43% for BCI III (IIIa) Subject k6b (σ_{opt} = 0.1 μV, N_a = 10). The results imply that the addition of an appropriate amount of noise in both the training and testing phases can provide the highest accuracy with a proper number of stages N_a.

4.3. Feature analysis

Fig. 8 shows scatter plots of two features ($DCA6_8$ and $DCA6_9$) of MI signal features. Blue and magenta markers are the features of right and left hand movements, respectively. Pentagram markers are features from original data. Diamond markers is the features from noise-added data with σ = 0.1. Solid markers are centroids of the respective data clusters. Centroids of features of noise-added data change from centroid of original data as a result of the original data are added with noise. The distance between centroids of two classes is expanded from 0.089 to 0.092. This possibly leads to improved separation of ESVM classification. The accuracy increases from 92.14% to 93.45%.

As expected, the optimal noise intensity (σ_{opt}) can produce the significantly effect to the parameters of the polynomial kernel functions, which leads to improve the classification accuracy. We analyze with the motor imagery classification (for BCI II (III)) with optimal noise intensity (σ_{opt} = 0.003 μV) The result shows that the number of support

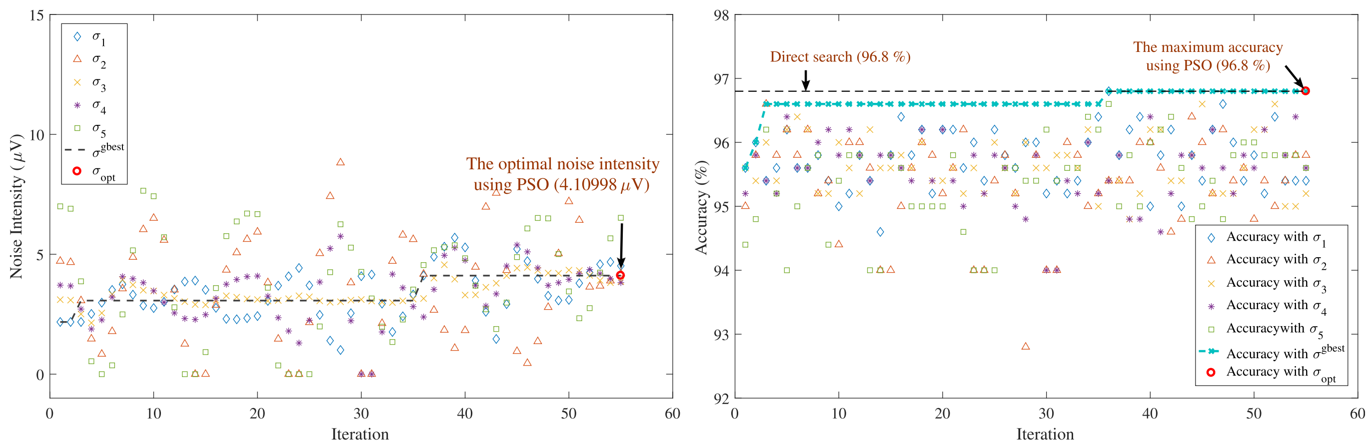


Fig. 9. Optimal noise intensity σ_{opt} search using PSO. The P300 classification system uses a single ESVM classifier ($N_r = 12$ and $N_a = 20$). PSO uses five noise intensity σ values as particles in each iteration. The plots show the paths towards the optimal noise intensity.

Table 6
Classification accuracy of the test datasets (14 channels): P300 classification using ESVM.

Training set	Testing set	Reps N_r	Train without noise				Train with noise				σ_{opt} (μV)	Increase $P_{An} - P_A$ (%)
			Test without noise P_A (%)	Test with noise P_{An} (%) $N_a = 1$	$N_a = 5$	$N_a = 10$	Test without noise P_{An} (%)	Test with noise P_{An} (%) $N_a = 1$	$N_a = 5$	$N_a = 10$		
A	**A**	14	53.00	40.20	53.20	54.60	51.00	**49.80**	58.00	59.40		
		15	**57.00**	**42.00**	**56.80**	**59.40**	**59.00**	49.60	**58.60**	**60.20**		
		Max	57.00	42.00	56.80	59.40	59.00	49.80	58.60	**60.20**	1.8	3.20
B	**B**	14	**76.00**	68.60	74.60	76.20	74.00	71.60	74.60	78.20		
		15	74.00	**70.40**	**77.00**	**79.20**	**77.00**	**74.00**	**77.00**	**80.60**		
		Max	76.00	70.40	77.00	79.20	77.00	74.00	77.00	**80.60**	1.0	4.60
C	**C**	**10**	93.55	93.55	93.55	93.55	93.55	99.35	**100**	**100**		
		13	**100**	99.35	99.35	97.42	100	97.42	97.42	96.77		
		15	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**		
		Max	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**	5.5	0

The significance of bold in the table is the intra-subject test and the number of repetitions (Reps N_r) correspond to the maximum accuracy.

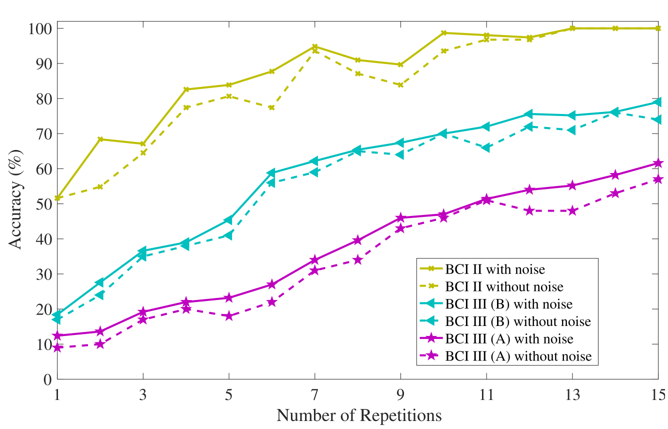


Fig. 10. Accuracy of P300 ESVM-classification using 14-channel (reduced) datasets. The dotted lines and solid lines denote the accuracy of the systems without noise and with noise, respectively. The results show that noise can enhance the classification accuracy for any number of signal repetitions.

vector decrease from 46 to 40 and the margin between two parallel hyperplanes is expanded from 0.3556 to 0.3680. The effect reduces the number of misclassified samples from 47 to 41. We use two-sample t-tests to confirm that there is an increase of margin: Mean of margin with noise is greater than mean of margin without noise with p-value <0.001 for the noise intensity 0.003 μV.

4.4. Optimal noise intensity search using PSO: P300 classification

We can test the system using different noise intensity σ values to approximate the optimal noise level σ_{opt}. PSO can also search for the optimal noise intensity using the following steps [41–43]:

Step 1 Define the number of particles $q = 5$. Note that more particles can give better results, but they require more search time.
Step 2 Randomly generate initial particles σ_i^0, $i = 1, \ldots, q$, in the range $(\sigma_l, \sigma_u) = (0.01, 10)$.
Step 3 Compute the ESVM classification accuracy P_A (objective function value) at σ_i^0 as $P_A(\sigma_1^0), P_A(\sigma_2^0), \ldots, P_A(\sigma_q^0)$.
Step 4 Set the initial velocity of each particle v_i^0 to zero. Set the iteration number to $k = 1$.
Step 5 Obtain the personal best values σ_i^{lbest} that give the highest value of the objective function $P_A(\sigma_i^j)$ from the i^{th} particle in all previous iterations $j = 1, \ldots, k$, and obtain the global best value σ^{gbest} that gives the highest value of the objective function $P_A(\sigma_i^j)$ from all particles $i = 1, \ldots, q$ in all previous iterations $j = 1, \ldots, k$.

$$\sigma_i^{lbest} = \underset{1 \le j \le k}{\arg\max}\, P_A(\sigma_i^j) \tag{8}$$

$$\sigma^{gbest} = \underset{1 \le j \le k, 1 \le j \le k}{\arg\max}\; P_A(\sigma_i^j) \tag{9}$$

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

Table 7
Comparison of the P300 classification accuracy of the proposed methods and existing methods.

Methods	BCI II				BCI III (A)				BCI III (B)			
	Number of signal repetitions				Number of signal repetitions				Number of signal repetitions			
	1	**5**	**10**	**15**	**1**	**5**	**10**	**15**	**1**	**5**	**10**	**15**
Existing methods												
ICA [15] (64 ch)	–	100	100	100	–	–	–	–	–	–	–	–
SVM [17] (10 ch)	64.5	100	100	100	–	–	–	–	–	–	–	–
Gradient boosting [52] (10 ch)	71	100	100	100	–	–	–	–	–	–	–	–
ESVM [18] (64 ch)	–	–	–	–	16	72	83	97	35	75	91	96
CNN-1 [20] [64 ch]	–	–	–	–	16	61	86	97	–	79	91	92
MCNN-1 [20] (64 ch)	–	–	–	–	18	61	82	97	–	77	92	94
RF [53] (64 ch)	–	–	–	–	–	73.80	89.00	97.30	–	80.32	92.30	98.41
Proposed method (PSO)												
ESVM	81.94	**100**	**100**	**100**	28.00	69.00	92.00	**97.60**	45.00	85.00	98.00	**98.60**
SVM	39.68	90.11	96.77	**100**	21.00	57.00	85.00	96.00	44.00	78.00	97.00	97.00

The significance of bold in the table is the intra-subject test and the number of repetitions (Reps N_r) correspond to the maximum accuracy.

$$v_i^{k+1} = v_i^k + \alpha_i(\sigma_i^{lbest} - \sigma_i^k) + \beta_i(\sigma^{gbest} - \sigma_i^k) \tag{10}$$

where α_i and β_i are uniform (0,1) random numbers.

Step 7 Update the particles σ_i^{k+1}

$$\sigma_i^{k+1} = \sigma_i^k + v_i^{k+1} \tag{11}$$

Step 8 Evaluate the objective function at the current σ_i^k as $P_A(\sigma_1^k), P_A(\sigma_2^k), ..., P_A(\sigma_q^k)$.

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). Thus, we obtain the (local) optimal noise intensity σ_{opt} that provides the maximum accuracy $P_A(\sigma_{opt})$. In case the objective function values P_A does not converge, go to Step 5.

Fig. 9 shows how PSO searches for the optimal noise intensity σ_{opt} for P300 classification. PSO creates noise intensity paths towards the red circle, which represents the (local) maximum accuracy at 96.8% with $\sigma_{opt} \approx 4.10998$ μV. In this instance, the accuracy from PSO is equal to that from the direct search method. Both methods can give different approximations of the optimal noise intensity due to the randomness of the signals and PSO's higher resolution of σ as opposed to a pre-determined (fixed) steps of σ in direct search.

The system can find the (local) optimal parameters of classifier for each dataset. We obtain the (local) optimal noise intensity σ_{opt} that provides the maximum accuracy $P_A(\sigma_{opt})$.

4.5. Multiple-user experiments: P300 classification

The proposed method is tested on multiple users where we combine datasets A, B, and C to explore its performance. We conduct both intra-subject tests and inter-subject tests. Intra-subject tests train the system with multiple-user datasets and test it with datasets from the same subjects. Inter-subject tests train the system with multiple-user datasets and test it with datasets from other subjects [49,50].

Xu *et al*. proposed an accuracy improvement of subject-specific P300 classification with the incorporation of inter-subject information [51]. This issue is important for the case of a small amount of training data from the target subject. They propose a classifier calibration strategy named weighted ensemble learning generic information (WELGI) that uses both the intra-subject and inter-subject information for building classifiers based on SVM and SWLDA. This research achieves on the P300 classification and also can be applied to other BCI paradigms. This paper shows how noise can improve the classifiers using data from multiple users. Table 5 shows that the classification accuracy of the intra-subject set and inter-subject set increases at optimal noise (σ_{opt}) level.

4.6. Channel reduction experiment: P300 classification

We also investigate the effectiveness of the algorithm in a simple EEG signal acquisition device such as the Emotiv Epoc. The Emotiv Epoc 14 positions are AF3, F7, F3, FC5, T7, P7, O1, O2, P8, T8, FC6, F4, F8, and AF4, as indicated by the green circles in Fig. 2 [45]. Note that the 14 positions cover only the sides of the scalp, and thus, the signal qualities are not as good as those of the signals obtained from 64 channels.

Table 6 and Fig. 10 show that the systems attain maximum accuracy at 60.20% for dataset A ($\sigma_{opt} = 1.8$ μV, $N_r = 15$, $N_a = 10$), 80.86% for dataset B ($\sigma_{opt} = 1$ μV, $N_r = 15$, $N_a = 10$), and 100% for dataset C ($\sigma_{opt} = 5.5$ μV, $N_r = 10$, $N_a = 5$). The results reveal that noise can enhance the classification accuracy for all 14-channel datasets.

4.7. Comparison with other methods

We compare the proposed method with other existing methods [15,17,18,20,52,53] as shown in Table 7. The ESVM achieves 100% accuracy for BCI competition II ($\sigma_{opt} = 3.5$ μV, $N_a = 5$, and $N_r = 4$), 97.60% for BCI competition III subject A ($\sigma_{opt} = 0.7$ μV, $N_a = 10$, and $N_r = 15$), and 98.60% for BCI competition III subject B ($\sigma_{opt} = 0.84$ μV, $N_a = 5$, and $N_r = 14$). The results show that proper noise can enhance accuracy of the BCI systems in general.

5. Discussions and conclusions

Experimental results show that noise benefits the BCI systems in many ways. Practical BCI systems may have limitations such as having only access to small training data [1,2], the need for reduction of data collection times (the number of signal repetitions) [15,17,18,20,52,53], the restrictions on channel positions in commercial recording systems [45], or accessibility for multiple users [6,7]. The proposed method shows how Gaussian noise can help alleviate these issues. Adding noise to the concatenated datasets can virtually increase the number of training samples and can reduce the chance of over-fitting [2]. The results also show that noise can enhance accuracy of BCI systems for any number of signal repetitions. 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.

Noise can also enhance the accuracy of the system using multiple-user datasets in the training phase. The noise-added system still performs well when another subject uses the system. This indicates that we can apply the system to process signals from multiple users such as BCI-based gaming applications [6,7]. The results on the reduced and restricted positions of commercial recording systems, such as the 14-channel Emotiv Epoc headset, also show the proposed algorithm can improve detections of weak signals.

The use of an array system can increase accuracy in all situations [28,30–35]. Independent noise in each stage collectively benefit the array of BCI systems during the training and testing. The use of particle swarm optimization (PSO) [41–43] shows how optimization algorithms can search for the (local) optimal noise parameter σ_{opt}.

These collective noise benefits suggest that future research should consider the role of noise addition in system design as well as in actual use. The appropriate noise intensity depends on the characteristics of signal and the classification systems, such as features, classifiers, number of signal repetitions, and number of stages in the array. Many types of noise pdf such as uniform, Gaussian, Laplacian noise, and other noise pdfs can enhance the performance of BCI systems or other nonlinear systems [28–31]. Finding the optimal noise pdf in each stage as well as how they are related to maximize the performance of BCI or other nonlinear systems still remains an open research problem.

Acknowledgment

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

References

[1] Lotte F, Congedo M, Lécuyer A, Lamarche F, Arnaldi B. A review of classification algorithms for EEG-based brain–computer interfaces. J Neural Eng 2007;4(2):R1.
[2] Muller K-R, Krauledat M, Dornhege G, Curio G, Blankertz B. Machine learning techniques for brain-computer interfaces. Biomed Technol 2004;49(1):11–22.
[3] Wolpaw JR, Birbaumer N, McFarland DJ, Pfurtscheller G, Vaughan TM. Brain–computer interfaces for communication and control. Clin Neurophysiol 2002;113(6):767–91.
[4] Constantin A, Danyluk A. A Brain-Computer Interface for the Classification of Motor Imagery. Williams College; 2007.
[5] Amiri, S, Rabbi, A, Azinfar, L, Fazel-Rezai, R, A review of P300, SSVEP, and Hybrid P300/SSVEP Brain- Computer Interface Systems, Brain-Computer Interface Systems - recent progress and Future prospects.
[6] Bonnet L, Lotte F, Lécuyer A. Two brains, one game: design and evaluation of a multiuser BCI video game based on motor imagery. IEEE Trans Comp Intell AI Games 2013;5(2):185–98.
[7] Korczowski L, Congedo M, Jutten C. Single-trial classification of multi-user P300-Based brain-computer interface using riemannian geometry. 2015 37th annual international conference of the IEEE engineering in medicine and biology society (EMBC). 2015. p. 1769–72.
[8] Long J, Wang J, Yu T. An efficient framework for EEG analysis with application to hybrid brain computer interfaces based on motor imagery and P300. Comput Intell Neurosci 2017;2017:1–6.
[9] Ramadan RA, Vasilakos AV. Brain computer interface: control signals review. Neurocomputing 2017;223:26–44.
[10] He S, Zhang R, Wang Q, Chen Y, Yang T, Feng Z, Zhang Y, Shao M, Li Y. A P300-Based threshold-free brain switch and its application in wheelchair control. IEEE Trans Neural Syst Rehabil Eng 2016(99):1.
[11] Mukerjee MK. NeuroPhone: brain-mobile phone interface using a wireless EEG headset, Tech. Rep. TR2010–666. Hanover, NH: Dartmouth College, Computer Science; May 2010.
[12] Vidal AFP, Salazar MAO, Lopez GS. Development of a brain-computer interface based on visual stimuli for the movement of a robot joints. IEEE Latin America Trans 2016;14(2):477–84.
[13] Bostanov V. BCI competition 2003-data sets Ib and IIb: feature extraction from event-related brain potentials with the continuous wavelet transform and the t-value scalogram. IEEE (Inst Electr Electron Eng) Trans Biomed Eng 2004;51:1057–61.
[14] Sampanna R, Mitaim S. Noise enhanced array signal detection in P300 speller paradigm using ICA-based subspace projections. 35th annual international conference of the IEEE engineering in medicine and biology society (EMBC). 2013. p. 4239–42.
[15] Xu N, Gao X, Hong B, Miao X, Gao S, Yang F. BCI competition 2003-data set IIb: enhancing P300 wave detection using ICA-based subspace projections for BCI applications. IEEE (Inst Electr Electron Eng) Trans Biomed Eng 2004;51:1067–72.
[16] Onishi A, Phan A-H, Matsuoka K, Cichocki A. Tensor classification for P300-Based brain computer interface. 2012 IEEE international conference on acoustics, speech and signal processing (ICASSP). 2012. p. 581–4.
[17] Kaper M, Meinicke P, Grossekathoefer U, Lingner T, Ritter H. BCI competition 2003-data set IIb: support vector machines for the P300 speller paradigm. IEEE (Inst Electr Electron Eng) Trans Biomed Eng 2004;51:1073–6.
[18] Rakotomamonjy A, Guigue V. BCI competition III: dataset II- ensemble of SVMs for BCI P300 speller. IEEE (Inst Electr Electron Eng) Trans Biomed Eng 2007;55:1147–54.
[19] Selim AE, Wahed MA, Kadah YM. Machine learning methodologies in P300 speller brain-computer interface systems. Proceedings of the radio science conference. 2009. p. 1–9.
[20] Cecotti H, Graser A. Convolutional neural networks for P300 detection with application to brain-computer interfaces. IEEE Trans Pattern Anal Mach Intell 2011;33:433–45.
[21] Park J, Kim KE. A POMDP approach to optimizing P300 speller BCI paradigm. IEEE Trans Neural Syst Rehabil Eng 2012;20(4):584–94.
[22] Long J, Li Y, Yu T, Gu Z. Target selection with hybrid feature for BCI-based 2-d cursor control. IEEE (Inst Electr Electron Eng) Trans Biomed Eng 2012;59(1):132–40.
[23] Zhang Y, Zhou G, Jin J, Zhao Q, Wang X, Cichocki A. Sparse bayesian classification of EEG for brain–computer interface. IEEE Trans Neural Networks Learn System 2016;27(11):2256–67.
[24] Zhou G, Zhao Q, Zhang Y, Adalı T, Xie S, Cichocki A. Linked component analysis from matrices to high-order tensors: applications to biomedical data. Proc IEEE 2016;104(2):310–31.
[25] Zhang Y, Wang Y, Zhou G, Jin J, Wang B, Wang X, Cichocki A. Multi-kernel extreme learning machine for EEG classification in brain-computer interfaces. Expert Syst Appl 2018;96:302–10.
[26] Mishuhina V, Jiang X. Feature weighting and regularization of common spatial patterns in EEG-based motor imagery BCI. IEEE Signal Process Lett 2018;25(6):783–7.
[27] Baali H, Khorshidtalab A, Mesbah M, Salami MJE. A transform-based feature extraction approach for motor imagery tasks classification. IEEE J Transl Eng Health Med 2015;3:1–8.
[28] Kosko B. Noise, viking. 2006.
[29] Mitaim S, Kosko B. Adaptive stochastic resonance. Proc IEEE 1998;86(11):2152–83.
[30] Osoba O, Mitaim S, Kosko B. The noisy expectation-maximization algorithm. Fluctuation Noise Lett 2013;12(03):1350012.
[31] Mitaim S, Kosko B. Adaptive stochastic resonance in noisy neurons based on mutual information. IEEE Trans Neural Network 2004;15:1526–40.
[32] Das A, Stocks NG, Hines EL. Enhanced coding for exponentially distributed signals using suprathreshold stochastic resonance. Commun Nonlinear Sci Numer Simulat 2009;14:223–32.
[33] Patel A, Kosko B. Noise benefits in quantizer-array correlation detection and watermark decoding. IEEE Trans Signal Process 2011;59:488–505.
[34] Stocks N, Information transmission in parallel threshold arrays: suprathreshold stochastic resonance, Phys Rev E 63.
[35] Zhang L, Xiao Y, Wu X. Array enhanced logical stochastic resonance in the presence of delta correlated noise. Proceedgins of the eighth international conference on fuzzy systems and knowledge discovery (FSKD). 2011. p. 2417–21.
[36] Bayram S, Gultekin S, Gezici S. Noise enhanced hypothesis-testing according to restricted neyman–pearson criterion. Digit Signal Process 2014;25(Supplement C):17–27.
[37] Liu S, Yang T, Zhang X, Hu X, Xu L. Noise enhanced binary hypothesis-testing in a new framework. Digit Signal Process 2015;41(Supplement C):22–31.
[38] Blankertz B. BCI competition 2003. 2003 [online]. available. http://www.bbci.de/competition/ii/.
[39] Blankertz B, Muller KR, Curio G, Vaughan TM, Schalk G, Wolpaw JR, Schlogl A, Neuper C, Pfurtscheller G, Hinterberger T, Schroder M, Birbaumer N. The BCI competition 2003: progress and perspectives in detection and discrimination of EEG single trials. IEEE (Inst Electr Electron Eng) Trans Biomed Eng 2004;51:1044–51.
[40] Blankertz B. BCI competition 2004. 2004 [online]. available. http://www.bbci.de/competition/iii/.
[41] Chong EKP, Zak SH. An introduction to optimization 4th Edition. Wiley; 2013.
[42] Rao SS. Engineering optimization: theory and practices. Wiley; 1996.
[43] Ravindran A, Ragsdell KM, Reklaitis GV. Engineering optimization: methods and applications. second ed. Wiley; 2006.
[44] Homan RW, Herman J, Purdy P. Cerebral location of international 10–20 system electrode placement. Electroencephalogr Clin Neurophysiol 1987;66(4):376–82.
[45] Emotiv Systems, [online]. available: http://emotiv.com.
[46] Sampanna R, Mitaim S. Noise benefits in motor imagery classification using ensemble support vector machine. 2014 IEEE biomedical circuits and systems conference (BioCAS). 2014. p. 53–6.
[47] Ming-Ai L, Rui W, Dong-Mei H, Jin-Fu Y. Feature extraction and classification of mental EEG for motor imagery, natural computation, 2009. Icnc '09. Fifth international conference on. 2009. p. 139–43.
[48] Xu B, Song A, Wu J. Algorithm of imagined left-right hand movement classification based on wavelet transform and AR parameter mode. The 1st international conference on bioinformatics and biomedical engineering (ICBBE). 2007. p. 539–42.
[49] Dalebout SD, Robey RR. Comparison of the intersubject and intrasubject variability of exogenous and endogenous auditory evoked potentials. J Am Acad Audiol 1997(8):342–54.
[50] Kang H, Nam Y, Choi S. Composite common spatial pattern for subject-to-subject transfer. IEEE Signal Process Lett 2009;16(8):683–6.
[51] Xu M, Liu J, Chen L, Qi H, He F, Zhou P, Wan B, Ming D. Incorporation of inter-subject information to improve the accuracy of subject-specific p300 classifiers. Int J Neural Syst 2016;26(03):1650010.
[52] Hoffmann U, Garcia G, Vesin JM, Diserens K, Ebrahimi T. A boosting approach to P300 detection with application to brain-computer interfaces, conference proceedings. 2nd international IEEE EMBS conference on neural engineering, vol 2005. 2005. p. 97–100.
[53] Farooq F, Kidmose P. Random forest classification for P300 based brain computer interface applications. 21st european signal processing conference (EUSIPCO 2013). 2013. p. 1–5.