A Differential Evolution based Adaptive Neural Type-2Fuzzy Inference System for Classification of Motor Imagery EEG Signals

Debabrota Basu[1], Saugat Bhattacharyya[1], Dwaipayan Sardar[1], Amit Konar[1], D.N. Tibarewala[2]

[1]Dept. Electronics and Telecommunication Engineering,
[2]School of Bioscience and Engineering,
Jadavpur University,
Kolkata, India.
basudebabrota29@gmail.com,
saugatbhattacharyya@live.com, dsnalkel@gmail.com,
konaramit@yahoo.co.in, biomed.ju@gmail.com

Atulya K. Nagar
Department of Math and Computer Science,
Liverpool Hope University,
United Kingdom.
nagara@hope.ac.uk

Abstract— **This paper proposes a new classification algorithm which aims at predicting different states from an incoming non-stationary signal. To overcome the failure of standard classifiers at generalizing the patterns for such signals, we have proposed an Interval Type-2 Fuzzy based Adaptive neural fuzzy Inference System (ANFIS). Through the introduction IT2F system, we have aimed at improving the uncertainty management of the fuzzy inference system. Besides that using DE in forward and backward pass and improving the forward pass function we have improved the parameter update on wide range of nodal functions without any quadratic approximation in forward pass. The proposed algorithm is tested on a standard electroencephalography (EEG) dataset and it is noted that the proposed algorithm performs better than other standard classifiers including the classical ANFIS algorithm.**

Keywords— **Interval Type-2 Fuzzy System, Adaptive Neural Fuzzy Inference, Differential Evolution, Brain-computer Interfacing, Electroencephalography**

I. INTRODUCTION

Brain-computer Interfacing aims to decode the mental brain states to control external devices like a prosthetic robot arm. It has wide application in fields like neuro-rehabilitation, communication [5], military and entertainment. The brain states are identified from signals obtained through devices like Electroencephalography (EEG), functional magnetic resonance imaging (fMRI), magnetoencephalography (MEG), functional near-infrared spectroscopy (fNIRS) electrocorticography (ECoG), and intra- cortical electrodes [2]. Among these, EEG signals are widely used among researchers because it is non-invasive, easily available, portable, and has good temporal resolution [3].

BCI technologies based on EEG recording is made of the following components: i) *Pre-processing*, where the incoming signals are filtered in the required frequency band, ii) *Feature Extraction*, where relevant information (corresponding to the different brain states of the user) are extracted from the filtered signals, and iii) *Classification*, where the features are used to discriminate between the different brain states [4]. In the past, researchers have employed various time-, frequency-, and time-frequency algorithms as feature extractors [16] and standard machine learning algorithms like Support Vector Machines, Hidden Markov Model, Discriminant Analysis, and Neural Networks as classifiers [1].

It is a widely accepted belief that signals acquired from an EEG recording are non-linear, non-stationary and non-Gaussian in nature [16]. These characteristics pose a problem towards the generalization of the training of a classifier, as the classifiers in such cases may yield inaccurate results when an unknown signal is fed to the classifiers. To tackle this problem, the classifier designed must be adaptive in nature and thus, in the present study we propose an Adaptive Neural Type-2 Fuzzy Inference System (ANT2FIS) algorithm to differentiate among the various mental states of a user. As the membership function of a type-1 fuzzy set has no uncertainty associated with it, type-2 fuzzy set comes as a natural replacement of type-1 fuzzy rule base of ANFIS architecture while dealing with EEG signals.

Adaptive Neural Fuzzy Inference System (ANFIS), first proposed by Jang [9] as a fuzzy inference counterpart of Adaptive neural networks, is a strong and standard neural fuzzy inference tool which can be used to predict the chaotic dynamics of a system. But due to its Type-1 fuzzy membership pattern, it fails to handle noise and uncertainty in case of chaotic and Ergodic signals. Also, ANFIS is dependent and sensitive to the parameter sets defined by the user. This inability to handle uncertainty becomes quite

prominent, in case of Ergodic and complex signals like EEG which inspired us to develop associate type-2 fuzzy sets with classical ANFIS for BCI application.

Type-2 fuzzy sets are developed to minimize the effect of uncertainties in rule-base fuzzy logic systems. In our current scenario, the EEG features are classified using our proposed Type-2 fuzzy sets with the fuzzy inference system of ANFIS to minimize the adverse effects of uncertainty. This has made our algorithm a better tool to handle and classify EEG signals. It is more robust, efficient, user independent and handles the uncertainty of EEG signals much better than the previous model. The inefficiency of uncertainty minimization is not the only shortcoming of Type-1 ANFIS. The node functions used in ANFIS have to be piecewise differentiable. This poses a constraint in determination of membership functions for ANFIS model. To make the functions more flexible for real-time uncertain situations like detection of the mental states from an EEG signal, we have to deal with this constraint. Thus we have used an evolutionary algorithm named Differential Evolution (DE) [15, 18] for the adaptation of weights of adaptive network instead of the gradient descent method used in classical ANFIS. This makes our algorithm more robust and a better classifier for a vast range of classification problems.

We have tested the effectiveness of our proposed algorithm by on a standard EEG BCI competition dataset. The datasets consists of EEG data obtained for two mental imagery states: right and left hand motor imagery. We have applied Extreme Energy Ratio (EER) criterion [19] as features in the present study. A comparison with the classical ANFIS and other standard classifiers suggests that our algorithm performs much better in dealing with EEG signals.

The rest of the paper is organized as follows: In Section II, we describe the methodology to our proposed algorithm. In Section III we have given the details of the experiment on which our proposed algorithm is tested. The results and comparisons are discussed in Section IV followed by the concluding remarks in Section V.

II. METHODOLOGY

A. Adaptive Neural Fuzzy Inference Systems (ANFIS)

ANFIS is a very popular and efficient neural fuzzy inference system used for function reconstruction, classification of miscellaneous data and signals, and design of controllers. It was first proposed by Jang [9] as an efficient combination of the adaptive neural network structures and Takagi-Sugeno model based fuzzy inference systems. This hybrid learning system maps the input dataset to the output based on both human reasoning, in the form of fuzzy rule bases and stipulated input-output data pairs.

In ANFIS, a fuzzy decision tree is created to classify the data into one of p^n linear regression models to minimize the sum of squared errors given by
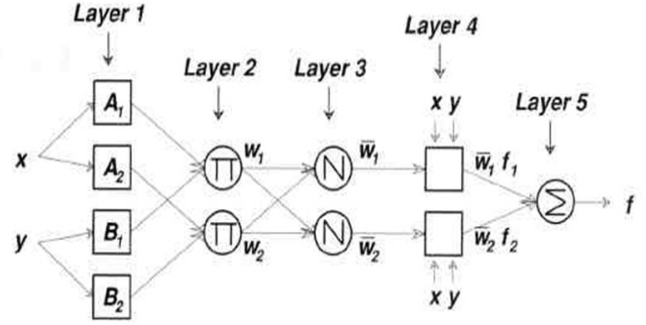
$$SSE = \sum_k E_k^2$$

(1)

Fig. 1. ANFIS network.

where, E_k denotes the error for the kth input vector between the obtained value and the desired value of the output, n is the number of input variables, and, p is the number of fuzzy partitions defined for each input variable.

The ANFIS used here contains five layers as delineated in Fig. 1. The node functions of each layer are fuzzy functions of same type. The nodes in each layer are either adaptive or fixed in nature, which is denoted by a square and circle, respectively, as shown in Fig. 1. The adaptive nodes consist of parameters required for the learning of the network, while the fixed nodes perform only fuzzy theoretic operations.

Due to adaptive approach of ANFIS, it can be used to classify the non-stationary signal like EEG very efficiently, where the fuzzy layers are used to capture the stochastic nature of the EEG signals [24]. But in this case we are bound by two assumptions: i) The node functions of the network have to be piecewise differentiable as it is optimized using gradient descent algorithm, and, ii) Based on quadratic error cost and the polynomial coefficients using least mean square (LMS) algorithm. But as it is based on symmetric error treatment and it is greatly influenced by outliers it may cause instability around points for higher number of partitions. Not only is that, type-1 fuzzy set not sufficient alone to handle all the uncertainties in an Ergodic and complex signal like EEG.

Type-2 fuzzy set, first proposed by Zadeh [20],was improved and exploited by Mendel [13]. It is shown in [24] that type-2 fuzzy logic can handle the uncertainties embedded in systems for classification and pattern recognition problems. As EEG itself is a non-stationary, Ergodic and complex signal, it is hard to classify this signal. Thus, type-2 fuzzy logic systems are used to handle the embedded uncertainty. But introduction of type-2 fuzzy membership functions in every node causes complex nodal functions which are harder than the type-1 membership functions to optimize. Besides that previously mentioned assumptions make the inference system erroneous. To cope up with these situations we have used the differential evolution algorithm to optimize the parameters. As it can work on wider range of functions and can optimize them in less time with better efficiency and shows improved stability.
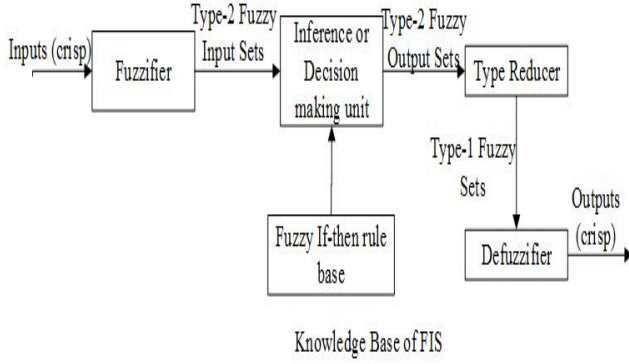
Fig. 2. Block Diagram of Type-2 fuzzy inference system

Fig. 3. The ANT2FIS network

B. Adaptive Neural Type 2 Fuzzy Inference Systems (ANT2FIS)

Basic type-2 fuzzy inference system consists of five blocks as shown in Fig. 2. In the input block the input vectors are fuzzified using type-2 fuzzifier (or fuzzy encoder) block. The type-2 fuzzy membership functions for each of the inputs is converted in corresponding membership functions. Now these membership functions are transferred to the next block, named inference block, where the fuzzy knowledge base i.e. the fuzzy if-then rules are used to calculate the output which is also a type-2 membership function. Now this type-2 function is reduced into a type-1 membership function using a type reducer block. Finally this set is taken as input in the output block to generate the final classification results.

1) ANT2FIS architecture

For simplicity, it is assumed the fuzzy system as depicted in figure 3 has two inputs x and y and one output z. In [21], Mendel et al have shown the basic type-2 fuzzy rules can be taken as extensions of type-1 fuzzy IF-THEN rules. Similarly we can define here two fuzzy rules as follows,

$$R^1 : IF\ x\ is\ \tilde{A}_1\ and\ y\ is\ \tilde{B}_1,\ THEN\ z\ is\ \tilde{c}^1.$$
$$R^2 : IF\ x\ is\ \tilde{A}_2\ and\ y\ is\ \tilde{B}_2,\ THEN\ z\ is\ \tilde{c}^2. \tag{2}$$

The architecture of the corresponding ANT2FIS will be as shown in the Fig. 3. The different layers in this structure and their corresponding fuzzy reasoning is described as follows,

Layer 0: Each and every state variable, say x and y, forms the layer 0 or the *input layer* of the ANT2FIS.

Layer 1: Each node of this layer is a square or adaptive node with a set of premise parameters. In this layer the node functions correspond to the predefined type-2 fuzzy membership functions. Outputs of these nodes are the membership values of the corresponding inputs. Thus, output of the ith node can be given as,
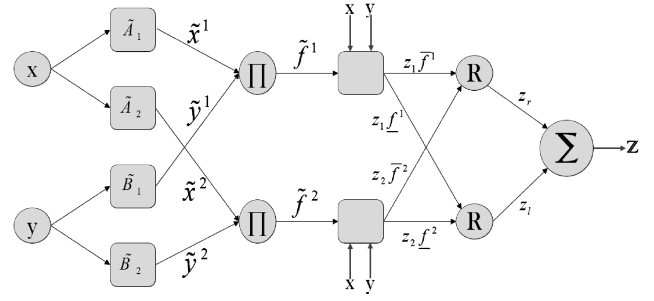
$$O_i^1 \equiv \mu_{\tilde{A}_i}(x) \tag{3}$$

Here, $\mu_{\tilde{A}_i}$ is the type 2 fuzzy membership function, x is the input variable and $\tilde{A}_i$ is the type-2 fuzzy set of a linguistic term for the ith node. Thus, O_i^1 is given by the cross section of the membership function defined on $\tilde{A}_i$ at the input point.

These membership functions can be a bell curve or a Gaussian curve varying between 1 and 0. It was shown in [9], ANFIS will meet the criterion of Stone-Weierstrass theorem if the class of the membership function remains invariant under multiplication. Here, we have chosen the node functions of Gaussian type. As Gaussian MFs always engender another Gaussian MF under product operation, which in turn assures the convergence of the inference system, every node is assigned Gaussian type-2 MFs given by,

$$\mu_{\tilde{A}_i}(x) = \exp\left[-\frac{1}{2}\left(\frac{x-m_i}{\sigma_i}\right)^2\right] \tag{4}$$

where, $\sigma_i \in \left[\sigma_i^1, \sigma_i^2\right]$.

Thus, the upper and lower MFs can be given as (Fig 4),

$$\bar{\mu}_{\tilde{A}_i}(x) = \exp\left[-\frac{1}{2}\left(\frac{x-m_i}{\sigma_i^2}\right)^2\right] \tag{5}$$

$$\underline{\mu}_{\tilde{A}_i}(x) = \exp\left[-\frac{1}{2}\left(\frac{x-m_i}{\sigma_i^1}\right)^2\right] \tag{6}$$

Initially, the total range of the signal is partitioned using p type-2 MFs. In Fig. 3, p=2. But in the real time simulation on EEG where features are taken as the input of ANT2FIS,

$$p=\ dimension\ of\ feature*2 \tag{7}$$

For the robustness of application and ease of calculation, the peak to peak amplitude variation of the EEG is reduced to [-1, 1]. An example of such an initialization for 2 one dimensional inputs is shown in Fig. 5. This uniform partition

of the amplitude interval distributes the uncertainty handling property of the system throughout the interval. Thus,

No. of nodes in layer 1 = no. of inputs (or features) $*p$

The parameters as delineated in Fig. 4 control the nature of the antecedent MFs. These are termed as premise parameters.
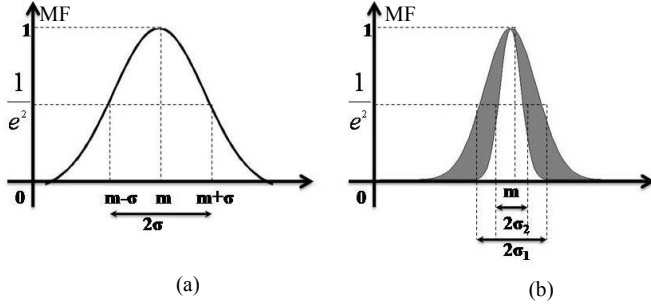

(a)   (b)

Fig. 4. (a) Type-1 Gaussian membership function with mean m and standard deviation σ. (b) Type-2 Gaussian membership function with premise parameters {m, σ_1, σ_2} where m is the mean and standard deviation is varying in the range σ_1 to σ_2.
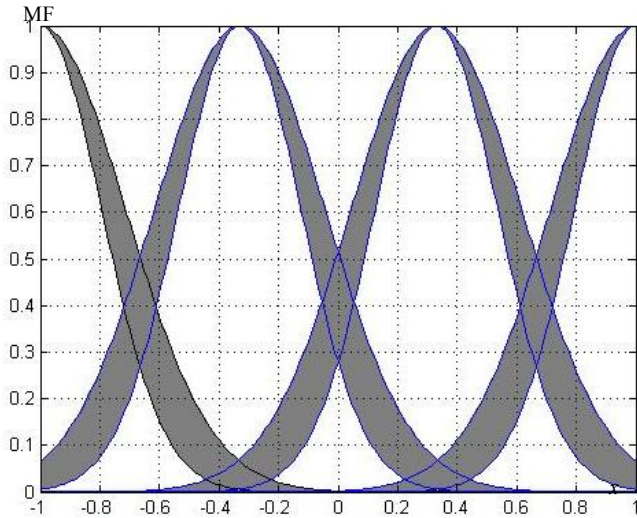

Fig. 5. A typical initialization of type-2 MFs in layer 1 for our simulation (range is set as [-1, 1])

Layer 2: It represents the fuzzy rule base of ANT2FIS. Every node in this layer is a fixed node, labeled as node Π, whose output is the product of all the incoming signals from the previous node and each node output represents the firing strength of a rule. It is given by

$$O_i^2 \equiv \tilde{f}^i(x, y) = \mu_{\tilde{A}_i}(x) * \mu_{\tilde{B}_i}(y)$$
(8)

where, $i = 1, 2, ..2 * D$, D is the dimension of input vector and, * denotes the product t-norm of the type-2 fuzzy memberships which in turn generates another type-2 MF [11]. Here, D=1.

Layer 3: The next part is the type reducer block where the type-2 fuzzy set is reduced to type-1. The final un-

normalized output of the type-2 FLS is obtained by applying the Extension Principle [32] on the type-1 decision making rule. This step is implemented using both layer 3 and layer 4. The output of layer 3 is,

$$O_i^3 \equiv \left\{ z_i . \overline{f}^i, z_i * \underline{f}^i \right\}$$
(9)

where, $z_i = p_i x + q_i y + r_i$

Here, i =1, 2. Each of the nodes contains parameter sets $\{p_i, q_i, r_i\}$ which are called *consequent parameters*.

Layer 4: This layer acts as the final reducer block [10]. Two nodes which belong to this layer, labeled by R, accumulate all the upper membership functions into one node and the lower MFs into another node. They consecutively generate two possible extremities of the output interval z_l and z_r, where,

$$O_1^4 \equiv z_r = \sum_{i=1}^{N} z_i \overline{f}^i$$

$$O_2^4 \equiv z_l = \sum_{i=1}^{N} z_i \underline{f}^i$$
(10)

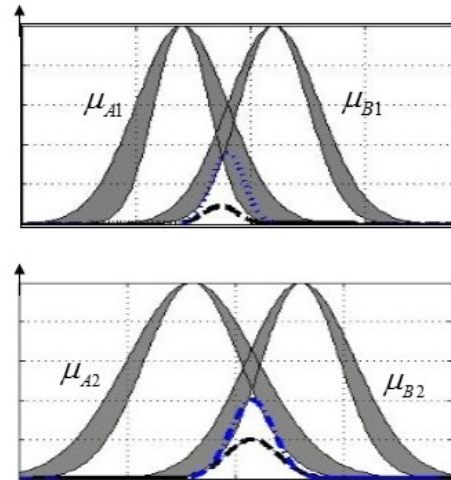Here, N= no. of rule bases= no. of nodes in layer 3. In this example N=2.

Fig. 6. Fuzzy reasoning in layer 2 using product t-norm of non-singleton type-2 fuzzy sets (blue line shows the UMF of output and black line shows the LMF of the output)

Layer 5: This layer contains one accumulator node which accumulates z_l and z_r and divide their sum by 2. Thus, the final output of the ANT2FIS,

$$O_5 \equiv z = \frac{1}{2}\left(z_l + z_r\right)$$
(11)

Thus, ANT2FIS with de-normalized output is a five layer structure as depicted above.

2) Differential Evolution based learning strategy

Like ANFIS, ANT2FIS also uses a two-pass learning cycle. In the forward pass the set of premise parameters, S is fixed and the consequent parameters S' are computed using Least Squared Error (LSE) algorithm during off-line learning. Whereas, in the backward pass, S' is fixed and S is computed using gradient descent algorithm. But due to their inefficiencies and limited performance range, both of them are replaced using the differential evolution algorithm.

The forward pass problem can be stated as, $AX = B$ where, X contains the consequent parameters. If the premise parameters are known and the inputs are given, A and B will be known. But still the problem will be over-determined and generally has no unique solution, as the dimension of B is much larger than that of X. It can be solved this considering as an optimization problem with cost function,

$$C = \det(AX - B) \qquad (12)$$

with the desired minimum value of 0. Now we can solve it using DE. The pseudo-code for basic DE algorithm is mentioned in the appendix.

The backward process is defined by (1) where Ek can be defined as,

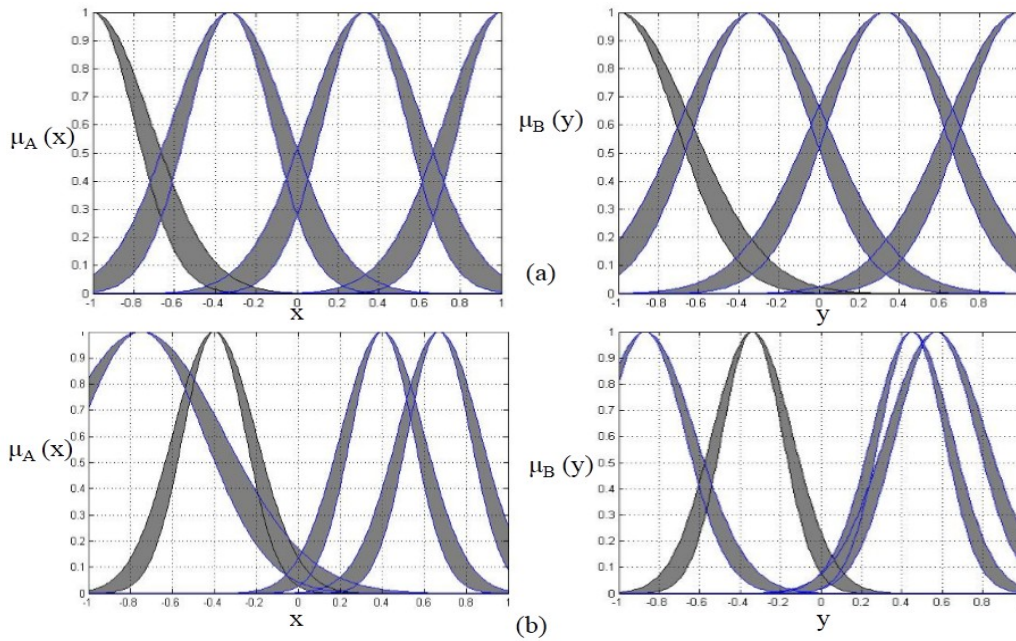$$E_k = \sum_{d=1}^{D} \left(z_{kd} - z_k^* \right)^2 \qquad (14)$$

Fig. 7. Adaptation of premise parameters using DE to get modified premise parameters with minimum error in output: (a) initially generated and uniformly distributed Gaussian MFs (b) the type-2 MFs after minimizing the error using DE based adaptation of premise parameters (# partitions=4, # input variables=2)
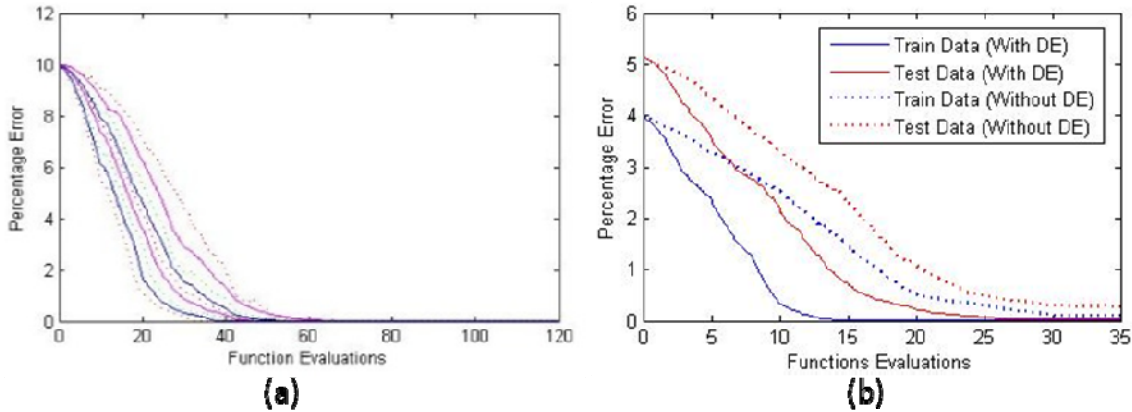

Fig. 8. Variation of average percentage error in output after using DE for forward pass and backward pass: (a) decrease in the error with FEs for different subjects for different subjects, (b) convergence of percentage error with FEs in backward pass for training data (blue line) and test data (red line)

where, z_{kd} is the value of the output vector for the dth dimension of the kth component and z_k^* is the desired output for the kth component. Now, if we consider the premise parameters as known, the cost function as described in (1) manifests into finding optimized premise parameters. The adapted versions of the fuzzy membership functions after backward propagation and the corresponding decrement in error with every function evaluation of DE are shown in Fig. 7 and 8. Fig. 8(b) substantiates improved convergence performance of DE while associated with ANT2FIS.

III. EXPERIMENTS

We have tested our proposed classifier on a standard EEG dataset obtained from the BCI Competition 2008- Graz dataset IIB in this study. This data set consists of EEG data from 9 right handed subjects, and further details on the dataset are published in [12]. The cue based screening paradigm consisted of two classes, namely the motor imagery (MI) of left hand (class 1) and right hand (class 2). Each session consisted of six runs with ten trials each and thus 120 trials (6 runs×10 trials×2 classes) are obtained from each subject at the end of each session.

The EEG signals were recorded using three bipolar recordings (C3, Cz and C4) with a sampling frequency of 250 Hz. The recordings had a dynamic range of ±100 μV for the screening and the ±50 μV for the feedback sessions. They were filtered between 0.5 Hz and 100 Hz using a band pass filter, and a notch filter at 50 Hz was enabled and the electrode position Fz served as ground. In addition to the EEG channels, the Electrooculogram (EOG) was recorded using three mono-polar electrodes with an amplifier setting of dynamic range ±1 mV [12]. From the EEG signals, the eye movement and blink artifacts were removed using the EOG signals.

The features extracted from the filtered EEG signals are determined using Extreme Energy Criterion [19]. At first, each of the EEG observations are rearranged such that we have a matrix X of dimension $N \times T$ corresponding to a single trial. Here, N is the number of electrodes and T is the number of samples in the recording period. The sample covariance, C, when computed by (15) results in a matrix of $N \times N$ dimensions. Averaging over all the samples of a class, we have the covariance of that class. So, C_0 and C_1 are obtained. If ϕ is considered to be a spatial filter, the signal energy is given by $\phi^T XX^T \phi = \phi^T C \phi$. Then, the Extreme Energy Ratio (EER) criterion for distinguishing the classes is given by (16). There can be two filter ϕ_{max} and ϕ_{min} respectively for maximizing and minimizing the ratio in (16). The Eigen vectors corresponding to the maximum and minimum Eigen values of the matrix $C_1^{-1} C_0$ gives the spatial filters ϕ_{max} and ϕ_{min}, respectively. One filter pair suffices if a single source has to be identified. The energy values of the signal filtered by these two filters can be treated as the features of an EEG observation. If m sources are to be identified, we have $2m$ values for each EEG sample. For m sources, ϕ_{max} (ϕ_{min}) is a set of filters given by m generalized Eigen vectors of matrix pair (C_0, C_1) which correspond m maximal (minimal) Eigen values [16].

$$C = XX^T \tag{15}$$

$$R(\phi) = \frac{\phi^T C_0 \phi}{\phi^T C_1 \phi} \tag{16}$$

In this paper, we consider 3 sources (C3, C4 and Cz) and thus, the feature vector for each sample is of dimension 1×6. These are energy values of the filtered signals where the filters are the Eigen vectors corresponding to the Eigen values arranged in ascending order. So, the first (six) energy value in the feature vector is computed using the Eigen vector corresponding to the smallest (largest) Eigen value, which are the features used as inputs in our proposed algorithm.

IV. RESULTS AND DISCUSSION

The EER features extracted in the previous section are fed as inputs to our proposed ANT2IFS classifier. As mentioned in the previous section, the experimental dataset comprises 5 sessions. So, we have selected the first four sessions for training the ANT2IFS classifier and the last dataset for testing the ANT2IFS classifier, from each subject. The performance of the training dataset is determined by applying k-fold validation technique [1] and the best, worst and average classification accuracy (CA) are used as performance metric in our current study. The output of the test dataset is determined from the incoming features on the trained classifier (for each subject) and the classification accuracy determines the performance of the classifier on test condition. Table I gives the results during training and testing phase of the classifier. The mean of the average classification accuracy obtained by the classifier on the nine subjects is 86.41% during training and 84.29% during testing.

TABLE I. PERFORMANCE MEASURE OF ANT2FIS CLASSIFIER DURING TRAINING AND TESTING OF THE EEG DATASET FOR EIGHT SUBJECTS (SUB)

Sub.	Training			Testing
	Best CA	Worst CA	Average CA	
1	97.80	86.10	93.95	85.17
2	96.50	84.70	92.05	83.30
3	97.90	84.20	91.25	83.30
4	96.80	84.30	89.55	83.40
5	97.40	83.90	90.65	83.00
6	96.90	85.40	91.15	84.40
7	95.40	84.30	92.85	83.40
8	98.90	86.70	94.80	85.76
9	**99.50**	**87.80**	**96.65**	**86.86**
Mean	97.45	85.27	92.54	84.29

The performance of our proposed classifier has been compared with its competitors: ANFIS, Support Vector Machine (SVM), Nave Bayesian (NB) and Feed Forward Neural Network (FFNN) using McNemars Test [8] on the test dataset. McNemars test compares the performance of two classifiers, say A and B. The null hypothesis of the test is as follows: algorithms A (here ANT2FIS) and B (competitive algorithms used for comparison) should have the same error rate, i.e., the number of feature vectors misclassified by A but

not by B (n_{01}) must be equal to the number of feature vectors misclassified by B but not by A (n_{10}). The test also employs a comparator statistic Z_j which follows a distribution with one degree of freedom and is given by

$$Z_j = \frac{\left(\left|n_{01} - n_{10}\right|\right)^2}{n_{01} + n_{10}} \tag{1(17)}$$

The hypothesis is rejected if $Z_j \geq 3.841459$, i.e., the confidence level is over 95%. Table-II provides the results of comparative performance of the proposed ANT2FIS classifier with four other classifiers, and the last column indicates that the hypothesis is in favor of only ANFIS. In other words, the ANFIS only is comparable with the proposed ANT2FIS in performance with a 95% confidence level.

TABLE II. STATISTICAL COMPARISON OF ANT2FIS CLASSIFIER USING MCNEMAR'S TEST

Competitors	n_{01}	n_{10}	Z_j	Acceptance/ Rejection of Hypothesis
ANT2FIS	0	3	1.33	Accepted
SVM	19	35	4.17	Rejected
NB	27	46	4.44	Rejected
FFNN	17	46	6.11	Rejected

In Fig. 9, we have compared the average accuracies of type-1 or classical ANFIS and our proposed ANT2FIS over 25 runs for each of them. We have plotted the accuracies of these two algorithms with respect to the test datasets and it is evident that Interval Type-2 ANFIS is more efficient and accurate to classify the EEG signals.
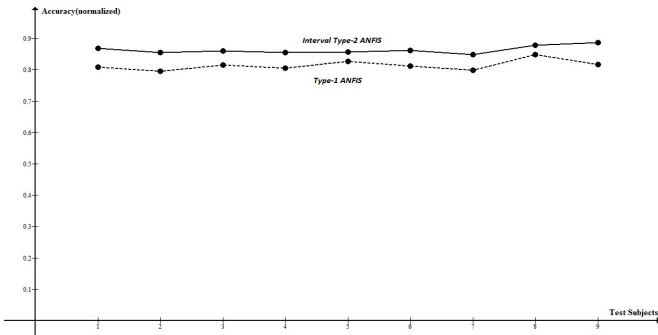

Fig. 9. Comparison of the accuracies of ANFIS and ANT2FIS

V. CONCLUSION

This paper introduces a novel interval type-2 neural fuzzy inference system for classification of EEG signals, required for BCI applications. In this study, we have used the algorithm to decode between two mental states of 9 different subjects. For this purpose, we have employed Extreme Energy Ratio criterion as features which are fed as inputs to our algorithm. From the classification results and statistical tests in Section IV, we have shown that our proposed algorithm ANT2FIS performs better than most standard classifiers. Also, as noted from Section IV, ANIT2FIS performs better than the classical ANFIS algorithm for the non-stationary EEG signals. No Free Lunch Theorem [27] tells us that there are no context-independent or application independent reasons to favor one learning or classification method over another. IT2 fuzzy systems are natural extensions of the original type-1 ANFIS systems, which ensures the performance and IT2 FSs just improves it.The improved performance is because of the following two points: i) by applying interval type-2 fuzzy sets on our algorithm, we have managed to handle uncertainty in the EEG signal, ii) by employing Differential Evolution in place of gradient descent method in backward pass and LSE in forward pass, and, iii) by changing the cost function of forward pass, the deficiencies of Symmetric error treatment & great outliers influence is overcome. That is how we have increased the flexibility of weights selection for the adaptive nodes. The major reason of its inefficiency in a few cases may be that the designed FOU covers too much or too little uncertainty that the system does not have. For our future work, we aim to use the ANT2FIS algorithm as classifier for real-time control of prosthetic devices, which would find applications in fields of rehabilitation engineering and to include other varieties of this algorithm using other kind of fuzzy reasoning like other norms and IF-THEN rules.

APPENDIX

PSEUDO-CODE OF DIFFERENTIAL ALGORITHM

An iteration of the classical DE algorithm consists of the four basic steps—initialization of a population of vectors, mutation, crossover or recombination, and finally selection. The main steps of classical DE are given below:

I. INITIALIZATION

Set the generation number t = 0 and randomly initialize a population of NP individuals

$$\vec{P}_t = \{\overrightarrow{X_1}(t),\ \overrightarrow{X_2}(t),\ \ldots,\overrightarrow{X_{NP}}(t)\}$$

with $\overrightarrow{X_i}(t)\ = \{x_{i_1}(t),\ x_{i_2}(t),....,x_{i_D}(t)\}$

and, each individual uniformly distributed in the range

$[\overrightarrow{X_{min}},\ \overrightarrow{X_{max}}]$, where $\overrightarrow{X_{min}} = \{x_{min_1},\ x_{min_2},\ \ldots,x_{min_D}\}$

and, $\overrightarrow{X_{max}} = \{x_{max_1},\ x_{max_2},\ \ldots,x_{max_D}\}$ with $i = \{1, 2, \ldots ,NP\}$

II. **While** stopping criterion is not reached, **do**

for $i = 1$ to NP

a. Mutation:

Generate a donor vector $\overrightarrow{V_i}(t)\ = \{v_{i_1}(t),\ v_{i_2}(t),....,v_{i_D}(t)\}$ corresponding to the ith target vector $\overrightarrow{X_1}(t)$ by the following scheme

$$\vec{V_1}(t) = \vec{X_{r_1}}(t) + F * (\vec{X_{r_2}}(t) - \vec{X_{r_3}}(t))$$

where r_1, r_2 and r_3 are distinct random integers in [1, NP].

b. Crossover:

Generate trial vector $\vec{U_i}(t) = \{u_{i_1}(t), u_{i_2}(t),....,u_{i_D}(t)\}$ for the ith target vector $\vec{X_i}(t)$ by binomial crossover as

$$u_{ij}(t) = \begin{cases} v_{ij}(t) \, if \, rand(0,1) < Cr \\ x_{ij}(t) \quad\quad\quad otherwise \end{cases}$$

c. Selection:

Evaluate the trial vector $\vec{U_i}(t)$

if $f\left(\vec{U_i}(t)\right) \leq f\left(\vec{X_i}(t)\right)$,

then $\vec{X_i}(t+1) = \vec{U_i}(t)$

$$f\left(\vec{X_i}(t+1)\right) = f\left(\vec{U_i}(t)\right)$$

end if

end for

d. Increase the counter value $t = t + 1$.

end while

The parameters used in the algorithm namely, scaling factor "F" and crossover rate "Cr", should be initialized before calling the "while" loop. The terminate condition can be defined in many ways, a few of which include: 1) fixing the number of iterations N, 2) when best fitness of population does not change appreciably over successive iterations, and 3) either of 1) and 2), whichever occurs earlier.

REFERENCES

[1] E. Alpaydin, Introduction to Machine Learning (Adaptive Computation and Machine Learning). Massachusetts: The MIT Press, 2004.

[2] R. A. Andersen, S. Musallam, and B. Pesaran, "Selecting the signals for a brain-machine interface," Curr. Opinion in Neurobio., vol. 14(6), pp. 720 – 726, 2004.

[3] K. Koka, R. Patwardhan, and W. Besio, "Computer simulation and tank experimental verification of concentric ring electrodes," in 26th Annual International Conference of the IEEE in Engineering in Medicine and Biology Society, 2004, IEMBS '04, pp. 2243–2246, San Francisco, CA, 2004.

[4] S. Bhattacharyya, A. Khasnobish, A. Konar, D.N. Tibarewala, and A.K. Nagar, "Performance analysis of left/right hand movement classification from EEG signal by intelligent algorithms," in 2011 IEEE Symposium on Computational Intelligence, Cognitive Algorithms, Mind, and Brain (CCMB), pp. 1–8, Paris, 2011.

[5] S. Bhattacharyya, A. Sengupta, T. Chakraborti, D. Banerjee, A. Khasnobish, A. Konar, D.N. Tibarewala, and R. Janarthanan, "EEG controlled remote robotic system from motor imagery classification," in Third International Conference on Computing, Communication Networking Technologies (ICCCNT), pp. 1–8, Coimbatore, 2012.

[6] Y. Chae, J. Jeong, and S. Jo, "Toward brain-actuated humanoid robots: Asynchronous direct control using an EEG-based BCI," IEEE Transactions on Robotics, vol. 28(5), pp. 1131–1144, 2012.

[7] J.J. Daly, and J.R. Wolpaw, "Braincomputer interfaces in neurological rehabilitation," The Lancet Neurologyr, vol. 7 (11), pp. 1032–1043, 2008.

[8] T.G. Dietterich, "Approximate statistical tests for comparing supervised classification learning algorithms", Neural Computing, vol. 10(7), pp. 1895–1923, 1998.

[9] J.-S.R. Jang, "ANFIS: Adaptive network based fuzzy inference system," IEEE Transactions on Systems, Man and Cybernetics, vol. 23(3), pp. 665–685, 1993.

[10] N.N. Karnik, J.M. Mendel, "Centroid of a type-2 fuzzy sets," Information Sciences, vol. 132(1–4), pp. 195–220, 2001.

[11] N.N. Karnik, and J.M. Mendel, "Operations on type-2 fuzzy sets," Fuzzy Sets and Systems, vol. 122(2), pp. 327–348, 2001b.

[12] R. Leeb, F. Lee, C. Keinrath, R. Scherer, H. Bischof, and G. Pfurtscheller, "Brain-computer communication: Motivation, aim and impact of exploring a virtual apartment," IEEE Transactions on Neural Systems and Rehabilitation Engineering, vol. 5(4), pp. 473–482, 2007.

[13] J.M. Mendel, R.I. John, and F. Liu, "Interval type-2 fuzzy logic systems made simple", IEEE Transactions on Fuzzy Systems, vol. 14(6), pp. 808–821, 2006.

[14] G. Pfurtscheller, and F.H. Lopes da Silva, "Event-related EEG/MEG synchronization and desynchronization: basic principles," Clinical neurophysiology, vol. 110(11), pp. 1842–1857, 1999.

[15] K. Price, R.M. Storn, and J.A. Lampinen, Differential Evolution: A Practical Approach to Global Optimization. Natural Computing Series, NJ, USA: Springer-Verlag, 2005.

[16] S. Sanei, and J.A. Chambers, EEG Signal Processing. NY: John Wiley, 2008.

[17] M.D. Serruya, N.G. Hatsopoulos, L. Paninski, R. Matthew, and J.P. Donoghue, "Instant neural control of a movement signal," Nature, vol. 416 (6877), pp. 141–142, 2002.

[18] R. Storn, and K. Price, "Differential evolution a simple and efficient heuristic for global optimization over continuous spaces," Journal of Global Optimization, vol. 11 (4), pp. 341–359, 1997.

[19] S. Sun, "The extreme energy ratio criterion for EEG feature extraction" in Proceedings of the 18th international conference on Artificial Neural Networks, Part II, pp. 919–928, Berlin, Heidelberg, 2008.

[20] Q. Liang, and J.M. Mendel, "An introduction to type-2 TSK fuzzy logic systems," in 1999 IEEE international fuzzy systems conference proceedings, vol. 3, pp. 1534–1539, 1999.

[21] Q. Liang, and J.M. Mendel, "MPEG VBR video traffic modeling and classification using fuzzy technique," IEEE Transactions on Fuzzy Systems, vol.9 (1), pp.183-193, Feb 2001.

[22] J. Long, Y. Li, H. Wang, T. Yu, J. Pan, and F. Li, "A hybrid brain computer interface to control the direction and speed of a simulated or real wheelchair". IEEE Transactions on Neural Systems and Rehabilitation Engineering, vol. 20(5), pp. 720–729, 2012.

[23] L. Zadeh, "Fuzzy sets", Information and Control, 1965.

[24] J. Zeng, and Z.Q. Liu, "Type-2 Fuzzy Sets for Pattern Recognition: The State-of-the-Art", Journal of Uncertain Systems, vol.1 (3), pp.163-177, 2007.