International Journal of Scientific & Engineering Research, Volume 4, Issue 7, July-2013
ISSN 2229-5518

Type-2 Fuzzy Sets for Modeling and Classifying Non-stationary Systems with Application in Brain-Computer Interfacing

Himadri Nath Moulick[1] & Joyjit Patra[1]
[1]Asst.Prof . C.S.E dept. & [1]Asst.prof. C.S.E.dept.
Aryabhatta Institute Of Engineering And Management,Durgapur.
Durgapur,West Bengal,India.
[1]himadri80@gmail.com &
[1]joyjitpatra_it35@rediffmail.com

Arun Kanti Manna[2]
[2]Asst.prof. C.S.E. dept.
Modern Institute Of Engineering & Technology
Bendel,Hoogly,West Bengal,India
[2]talk2manna@gmail.com

Abstract— A correct representation of uncertainty in measurement is crucial in many applications. Statistical approach sometimes is not the best choice, especially when the knowledge of the measurement process refers only to the support of the values and does not allow a correct assumption on the probability density function (pdf) of the measured variable. In this paper we present an approach that uses the concept of generalized fuzzy numbers, namely Type-2 fuzzy sets, in order to handle the intrinsic dispersion of the possible pdfs associated to a variable. The relation between our representation and the so called Random Fuzzy Variables (RFV) will be also investigated. The use of this representation allows to easily implement the uncertainty propagation, through a functional model, by working directly on the Type-2 fuzzy numbers and by evaluating simultaneously the propagation results for the whole set of confidence levels. Anyway, when a statistical analysis can be performed, the results can be embedded in this generalized representation. Moreover, the new approach allows to assign to the final measurement value a reliable confidence level also in this case, by combining the expanded uncertainty evaluated following IEC-ISO Guide recommendations with the Type-2 fuzzy numbers associated to the output variable. An example of this representation will be also provided .The IT2FLS design methods have been empirically verified in this work in the realm of pattern recognition. In particular, the potential and the suitability ofIT2FLS to the problem of classification of motor imagery (MI) related patterns in electroencephalogram (EEG) recordings has been investigated. The outcome of this study bears direct relevance to the development of EEG-based brain-computer interfaces (BCIs) since the problem under examination poses a major difficulty for the state-of-the-art BCI methods. The IT2FLS classifier is evaluated in this work on multi-session EEG data sets in the framework of an off-line BCI. Its performance is quantified in terms of the classification accuracy (CA) rates and has been found to be favorable to that of analogous systems employing a conventional T1FLS, along with linear discriminant analysis (LDA) and support vector machine (SVM), commonly utilized in MI-based BCI systems.

Keywords-Uncertainty,probability-possibility transformations, Type-2 fuzzy variables. pattern recognition, electroencephalogram, brain-computer interface.

I. INTRODUCTION

The correct representation of the measurement associated to a given variable is a focal point in many applications. An exhaustive description of the principal recommendations about how a reliable expression of the measurement and of its uncertainty has to be performed, is contained in the IEC-ISO "Guide to the expression in measurement" [1] which we address almost totally. Principally, the IEC-ISO Guide states that the measurement cannot be expressed by a single value, but by a distribution of values over an interval within which the measurements lie with a given confidence level. So, detailed rules are provided in order to evaluate this distribution with the highest confidence level associated. The probabilistic approach represents the natural way of computing uncertainty estimation and performing uncertainty propagation through a functional model, but recently many limitations of this approach have been focused. In particular, in order to perform a correct probabilistic representation of the measurement, a set of independent observations is needed. In particular, in order to perform a correct probabilistic representation of the measurement, a set of independent observations is needed. However, in many applications, the value assigned to a certain variable is taken from manuals, calibration reports, handbook, reference values, so that any assumption on the probability density function (pdf) associated to a variable cannot be reliable. Moreover, in order to propagate the uncertainty through a generic function *f*, the joint pdf and the statistical correlation have to be estimated. Again, if a very weak knowledge is available about some of the involved variables, all these estimations can lead to a strong error propagation, thus producing eventually a biased expression of the combined uncertainty.

In all these cases, in particular when a type-B evaluation [1] of uncertainty is needed, alternative methods have to be implemented. Recently [3, 4], a fuzzy approach has been investigated in order to represent uncertainty in measurement when the available information is poor and does not allow a statistical analysis for uncertainty handling. The concepts of fuzzy variables and fuzzy sets have been introduced by Zadeh [5, 6] as an extension of the traditional concept of membership of a variable a to a set A. In crisp set theory this membership is represented by a one $(a \in A)$ or by a zero $(a \notin A)$. whereas in fuzzy set theory it can be modeled by a MF $\mu_A(a)$ such that $0 \leqslant \mu_A(a) \leqslant 1$, with μA(a) convex and normal (i.e., there exists at least one value b such that $\mu_A(b) = 1)$ .The set A is called fuzzy subset and the support of A is the set of points at which μA(a) is positive. The α−level set (or α−cut) of A is a non fuzzy set, denoted by Aα, defined as Aα = {a|μA(a) _ α}. In [7] Zadeh also introduced the concept of possibility theory as a mathematical counterpart of probability theory that deals with uncertainty by means of fuzzy sets, so that a fuzzy/possibility approach is denoted. Moreover, in [3, 4] the authors also underline that the fuzzy/possibility approach is between interval analysis and probability theory. The former is the less expressive because uses only the information of upper and lower bounds of an interval, without any relation with a level of confidence (so with a membership degree). This is not sufficient to use IEC-ISO Guide recommendations in uncertainty expression. Otherwise, the probability approach is somehow too rich for representing relative lack of information coming from human experts or imprecise sensors.

Appropriate modeling of non-stationary responses of realworld systems is a challenging systems engineering problem. Its complexity can be particularly acute as the intrinsic characteristics of real-world systems are often severely nonlinear. In this paper, the emphasis is on robust handling nonstationary effects in pattern recognition problems, where the inference is drawn under uncertain conditions. Classifier systems designed for effective accounting for non-stationary responses often involve a mechanism to monitor or assess the validity of stationary feature distribution hypothesis normally made during a classifier design. The assessment outcome can be used to update the classifier to track the system evolution. The problem of assessing the stationarity hypothesis can be addressed with three main approaches: data-driven, analytical or knowledge-based [1]. A data-driven approach directly inspects data coming from the process and assumes that the available data set is large enough to assess the validity of the stationarity hypothesis with large confidence [2][3]. This solution guarantees a good drift detection ability without requiring any a priori information about the process under investigation. The analytical modality assumes that a mathematical description of the process generating the data is available: only few data are hence required to assess the hypothesis [4][5]. The knowledge-based modality assumes instead that some a priori information about the process (but not the model) is available, e.g., derived from data samples based on causal analysis [4][6]. The identification of the time instant associated with the loss in stationarity allows the designer to take actions, e.g., by updating the classifier network weights to track the process evolution [7] or retraining the classifier [8][9] exactly when needed.In this work, special attention is given to non-stationarity manifestations in neurophysiological data. In particular, the challenging task of discrimination of patterns in the electroencephalogram (EEG) signals recorded from subjects performing motor imagery (MI), e.g. imagining left or right hand movement, as part of an EEG-based brain-computer interface (BCI) system. The underlying EEG data demonstrate a broad spectrum of non-stationary effects at different temporal levels. They mostly arise out of the variability of the brain state dynamics due to changing mental focus, motivation, and biofeedback effects, among others, during BCI experiment. In the BCI community, the issue of long-term changes in the salient EEG characteristics, mainly between experimental sessions, is considered to pose a significant challenge. It has been approached by several methods with varying degree of success. In [9], a neural classifier was retrained every day of the subject's training session and then embedded in the BCI that was operated the following day. Although this approach involving frequent inter-session recalibrating is commonly exploited, it only partially mitigates the effect of long-term non-stationarities and it is rather impractical. Shenoy et al. [10] investigated changes in the EEG feature distribution obtained in training and test BCI sessions with feedback.

II. TYPE-2 FUZZY SETS FOR UNCERTAINTY HANDLING

In [11] Zadeh firstly introduced the concept of generalized fuzzy sets. Suppose that *A* is a fuzzy set and suppose that the MF $\mu A(a)$ associated is allowed to be a fuzzy subset in the interval [0, 1]. In order to differentiate this kind of generalized fuzzy sets from the classical ones Zadeh refers to them as Type-2 fuzzy sets. More generally, he gives a recursive definition of Type-*n* fuzzy sets as follows

International Journal of Scientific & Engineering Research, Volume 4, Issue 7, July-2013
ISSN 2229-5518

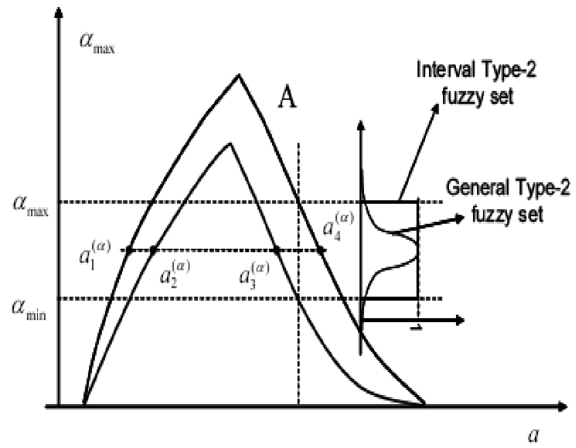

Fig. 1. Example of an interval Type-2 MF

Definition II.1: A fuzzy set is of Type-*n*, $n = 2, 3, \ldots$, if its MF ranges over fuzzy sets of Type-$(n - 1)$. The MF of a fuzzy set of Type-1 ranges over the interval [0, 1].

In a recent literature [13,14], various classes of Type-2 MFs are inspected, but anyway a particular one, including the so called Interval Type-2 fuzzy sets, has been widely investigated and applied in various contexts such as decision making, timeseries forecasting, control of mobile robots [14], etc. Interval Type-2 fuzzy sets are the most widely used Type-2 fuzzy sets because they are simple to use and because it is very difficult to justify the use of any kind of Type-2 fuzzy sets. In this case, the MF $\mu A(a)$ is an Interval Type-2 fuzzy set so that it can be represented only by its lower and upper bounds (i.e. by two Type-1 MFs). This situation is depicted in Fig. 1 and compared with other typologies of non Interval Type-2 MFs (denoted as General Type-2 MFs). In order to identify how to easily operate on this class of more complex fuzzy sets, in [12] the concept of interval of confidence of Type-2 is introduced.

Let us recall now some basic notions. Assume that the lower and upper bounds of an interval of confidence, instead of being ordinary numbers, are fuzzy numbers, that themselves have intervals of confidence. We will denote this kind of Type-2 interval of confidence as

$A = [[a1, a2], [a3, a4]]$

such that $a1$ _ $a2$ _ $a3$ _ $a4$. When $a1 = a2$ and $a3 = a4$ the interval of confidence of Type-2 becomes an interval of Type-1 and if $a1 = a2 = a3 = a4$ the interval becomes of Type-0 (i.e., a number). Consider now a sequence of intervals of confidence of Type-2 that depends on α, that is

$$\forall \alpha \in [0,1], \quad \forall a_1^{(\alpha)},\, a_2^{(\alpha)},\, a_3^{(\alpha)},\, a_4^{(\alpha)}$$

$$\mathcal{A}_\alpha = \left[[a_1^{(\alpha)}, a_2^{(\alpha)}], [a_3^{(\alpha)}, a_4^{(\alpha)}] \right],$$

$$a_1^{(\alpha)} \leqslant a_2^{(\alpha)} \leqslant a_3^{(\alpha)} \leqslant a_4^{(\alpha)}.$$

In order to perform algebraic operations on Type-2 fuzzy sets let us consider now that a fuzzy number of Type-2 can be constructed in two ways.

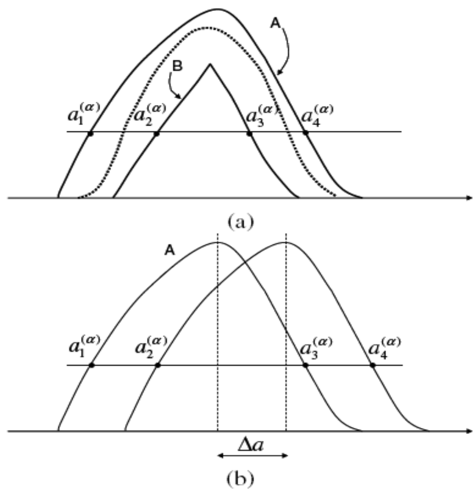

Fig. 2. Two ways of building a Type-2 fuzzy number (a) and (b)

1) Given a Type-1 fuzzy number *A* and a convex fuzzy subset *B* we build a Type-2 fuzzy number as shown in Fig. 2 (a). Note that we can identify a gamma of Type- 1 MFs belonging to the range [*B*,*A*], as for example the dotted MF.

2) The second kind of construction considers a Type-1 fuzzy number *A* and its translation of a certain Δa thus obtaining Fig. 2 (b). The latter interpretation is commonly used in literature [13,14]. It can be seen as a blurring of a Type-1 MF around a central value, thus producing the corresponding Type-2 MF. Otherwise, the former representation is the one we address in this paper, since the fuzzy subset *B* is naturally the inner MF (i.e., a lower bound) and the fuzzy set *A* corresponds to the outer MF (i.e., an upper bound). This point of view will allow us to directly construct the Type-2 MF in the context of uncertainty representation. In order to investigate also the relation among Type-2 MF and RFV let us refer to [12]. Let us consider newly a Type-2 fuzzy number by its α−cuts

$$\left[[a_1^{(\alpha)}, a_2^{(\alpha)}], [a_3^{(\alpha)}, a_4^{(\alpha)}] \right]$$

Now, let us assign to each segment $[a_1^{(\alpha)}, a_2^{(\alpha)}]$ and $[a_3^{(\alpha)}, a_4^{(\alpha)}]$ a pdf $fL(\alpha, x)$ and $fR(\alpha, x)$ respectively. Therefore, in the interval of confidence, the lower and the upper bounds become random variables. Figure 3 shows this concept, with *FL* and *FR* the probability distribution functions associated.

In [12] the authors also show that the envelope of a RFV is a Type-2 fuzzy number, and that, while the operations on the RFV are necessarily performed by sum-product convolution (since the confidence interval is represented by two pdfs), the

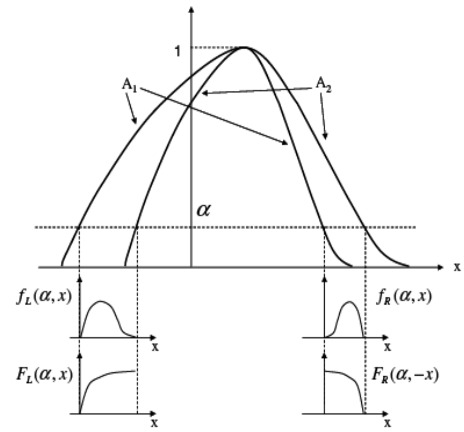

Fig. 3. Random Fuzzy Variables embedded in Type-2 MFs

operations on Type-2 fuzzy numbers can be performed by maxmin convolution that corresponds to the use of the so called Extension Principle (EP) by Zadeh. They also show that the application of EP can be turned into working directly on $\alpha-$cuts, under the assumption of independent variables. Anyway, the use of RFV, is necessary when systematic errors or their corrections are partially unknown, so that standard approaches produce a wrong evaluation of uncertainty. In this case a particular class of rectangular Type-1 MF are embedded into the RFV, in order to model systematic errors or their uncomplete correction. In the following, we will assume that systematic errors are completely corrected, so that only random effects should be considered. So, in order to express uncertainty we consider the use of Interval Type-2 MF.

This paper reports an empirical investigation into systematic data-driven approaches to fuzzy classifier design within the framework of a so-called interval type-2 (IT2) FLS that represents a subcategory of T2FLSs (cf. section III). A computationally effective design methodology is essential to deal with systems involving large data sets, particularly in cases of systems exhibiting acute non-linear and non-stationary characteristics. With the central objective of automating the classifier design process, several innovative methods for fuzzy rule base structure initialization and its parameter optimization are devised and analyzed in this work. The proposed enhancements are incremental and heuristic in nature. A complete design process and implementation of a BCI classifier is discussed in the paper. The primary aim is to examine the effectiveness of a novel IT2FLS-based approach to robust multi-session BCI classification. Therefore, special attention is paid to the classifier's capability to generalize well across a few data sets obtained at different times. The presented instance of brain signal pattern recognition illustrates the challenging nature of a more general problem of reliable analysis and interpretation of EEG in the presence of non-stationary effects. This paper is organized as follows. Section II outlines the specific problem of MI related EEG pattern recognition considered in the paper. Section III elaborates on the T2FL methodology developed and employed in this work. In section IV, the results of the BCI experiments are demonstrated and discussed. Conclusions are then presented in section V.

III. MOTOR IMAGERY-RELATED EEG PATTERN RECOGNITION PROBLEM

A. EEG Data Description The EEG data were recorded in the Intelligent System Research Centre, University of Ulster at Magee, Derry, UK.

The EEG data were obtained from 8 subjects in a timed experimental recording procedure where the subjects were instructed to imagine moving the left or the right hand depending on the horizontal location (left/right) of a target basket displayed at the bottom of a monitor screen (Fig. 1). Each trial was 7 s in length. A ball was displayed at the top of the screen for the first 3 s. In the meantime, at t = 2 s acoustic stimulus signified the beginning of a trial. At t = 3 s the ball started moving to the bottom of the screen. Therefore the segment of the data recorded after t = 3 s of each trial is considered as event related. The horizontal component of the ball movement was continuously controlled in on-line experiments by a subject via the biofeedback mechanism [12] employing an IT2FLS classifier. In this paper however, the fuzzy classifier is evaluated in off-line analysis, after the data have been collected. In addition, it was designed for discrete classification of entire EEG trials resembling the concept of single MI related EEG trial classification, unlike continuous classification at every time point in on-line mode.

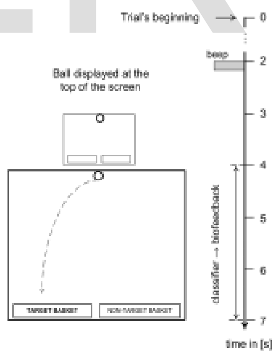

Figure 1. Illustration of a BCI basket paradigm.

The EEG trials were recorded with a g.tec amplifier from two bipolar EEG channels over C3 and C4 locations (10/20 system) [20]. They were then sampled at a frequency of 125 Hz and band-pass filtered in the frequency range 0.5–30 Hz. The data were obtained over 10 sessions, each session consisting of 160 trials. Four consecutive sessions for each subject were arbitrarily selected for further off-line analysis.

B. Feature Extraction

International Journal of Scientific & Engineering Research, Volume 4, Issue 7, July-2013
ISSN 2229-5518

Due to the oscillatory character of the MI induced brain phenomena, reflected in the sensorimotor EEG activity, a method of spectral analysis was employed in the signal quantification. The EEG rhythmical content within μ (8–12 Hz) and (18–25 Hz) ranges was examined in this regard since it reveals lateralized signal power patterns over C3 and C4 locations characteristic of the MI that a subject is performing (left vs. right MI). In particular, when the sensorimotor area of the brain is activated during the imagination of hand movement, the interplay between contralateral attenuation of the μ rhythm and ipsilateral enhancement of the central β oscillations in different phases of MI can usually be observed. These processes occur due to the neurophysiological mechanisms of the so-called event-related desynchronization (ERD) and event-related synchronization (ERS) [20]. The exact EEG manifestations and frequency bands of ERS and ERD vary from subject to subject.The event-related segment of each EEG trial (the last 4 s corresponding to 500 samples) was divided into rectangular windows depending on the settings of two parameters: window length and the amount of overlap. Next, the frequency-related information was extracted from every window with a parametric power spectral density (PSD) method that employs Yule-Walker algorithm for autoregressive (AR) modeling [21]. The EEG features were calculated in each time window as the total energy within the bands of interest (adjusted μ and β).This served as an element C3/C4 i x of the feature vector **x** in (1).

C. Classification Problem

Given a set of feature vectors representing EEG trials, the classification task undertaken in this work is to associate them with classes of mental tasks, more specifically – MI. This instance of brain signal pattern recognition is dichotomous in nature since the differentiation between an imagination of left and right hand movement only is aimed. The problem is challenging due to various non-stationary effects inherent to the on-going electrophysiological brain activity, as discussed briefly in section I. Here, the emphasis is placed on the effective dealing with long-term changes in EEG spectral patterns correlated with MI. In particular, robust inter-session classification performance is the main focus of this work as it represents an urgent need in BCI. A successful method is expected to maintain a satisfactory accuracy rate over a few sessions recorded at distant times (here: once a week) without the need for frequent inter-session adjustments. The major difficulty in this regard lies in the session-to-session variability of the salient EEG features. The next section proposes an IT2FLS-based approach to discrimination of MI induced EEG patterns as an instance of a broad category of non-stationary pattern recognition problems, where no underlying analytical system model or its probabilistic description is available. The emphasis is on design methodology for an IT2FLS classifier to effectively exploit its framework for handling variability in data.

Type-2 fuzzy numbers for type-B uncertainty handling

Let us suppose that the measurement of the variable *X* is provided in the form (*X*0±*UX*)*M*, where *UX* is the expanded uncertainty of *X*, taken from manuals, calibration reports etc., and *M* is the measure unit (in the following omitted for notation simplicity). Then, in the case of incomplete knowledge of the pdf associated to *X* (i.e., in a type-B uncertainty expression), it is possible to build a gamma of pdfs starting from the declaration of *X* and by various assumptions and knowledge of the performed measurement process. Under these considerations, we can have the situation in Fig. 4. Obviously, given the support [*x*1, *x*2] the gaussian probability density function (gpdf) with $\sigma = (x2-x1)/6$ (i.e., containing the 99.73% of the pdf in the support) is the most localized pdf around the central value $xm = (x1 + x2)/2$, whereas the uniform probability density function (updf) with the same support is the least localized pdf. Using the probability-possibility transformations introduced in [8] the Possibility Distributions (PDs) shown in Fig. 5

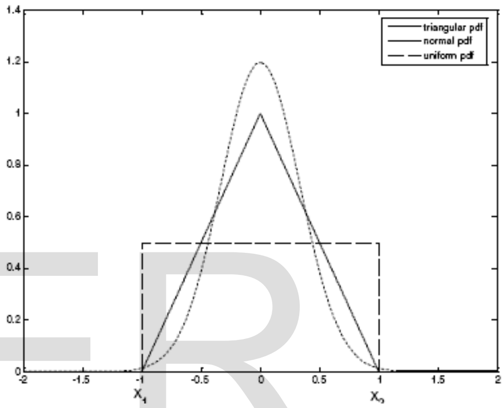

Fig. 4. Various pdfs associated to a type-B uncertainty evaluation

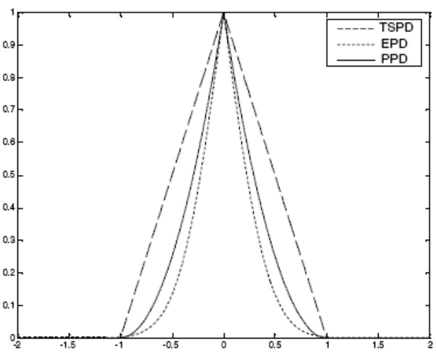

Fig. 5. PDs related to various pdfs associated to a type-B expression of uncertainty are obtained.

Note that there is a set of possible PDs associated with the support [*x*1, *x*2] ranging from the PD related to a gpdf (that we will denote as Error function Possibility Distribution (EPD)) to the PD related to the updf (Triangular Symmetric possibility Distribution (TSPD) in the following), through the internal

Parabolic Possibility Distribution (PPD). If one considers as the lower bound of this set of PDs the dotted one (i.e., the EPD) and as the upper bound the dashed one (i.e., the TSPD) then a Type-2 fuzzy number is obtained, denoted by its Type-2 α−cut, $\mathcal{A}_\alpha = [[x_1^{(\alpha)}, x_2^{(\alpha)}], [x_3^{(\alpha)}, x_4^{(\alpha)}]]$, where $x_1^{(\alpha)}$, $x_2^{(\alpha)}$

$x_3^{(\alpha)}$ and $x_4^{(\alpha)}$ can be easily obtained as

$$x_1^\alpha = x_m - (1-\alpha)(x_m - x_1),$$

$$x_2^\alpha = x_m - \sqrt{2}/6(x_2 - x_1) \cdot \text{erf}^{-1}(1-\alpha),$$

$$x_3^\alpha = x_m + \sqrt{2}/6(x_2 - x_1) \cdot \text{erf}^{-1}(1-\alpha),$$

$$x_4^\alpha = x_m + (1-\alpha)(x_2 - x_m),$$

Now, each variable can be represented by means of this Type-2 MF, whose support is taken from the type-B expression of the measurement value. Otherwise, when a set of repeated observations for a given variable is available, then the IECISO Guide recommends to perform a statistical analysis (i.e., a type-A uncertainty evaluation), so that a reliable estimation of the correct pdf can be extracted. Let us suppose, for example, that a gpdf is the best MLE for a given variable. Then, the use of the probability-possibility transformation involving only the gpdf, thus producing a EPD, leads to the degenerate Type-2 α−cut

$$x_1^\alpha = x_2^\alpha = x_m - \sqrt{2}/6(x_2 - x_1) \cdot \text{erf}^{-1}$$

……….(2)

and

$$x_4^\alpha = x_3^\alpha = x_m + \sqrt{2}/6(x_2 - x_1) \cdot \text{erf}^{-1}($$

…………………(3)

Note that, in this case, the Type-2 MF reduces to a Type-1 MF, embedding the information added by the statistical analysis. So, in order to propagate, through a function *f*, the uncertainty of each variable, the unique representation by means of Type-2 MFs can be adopted, so that the operations involved in *f* can be applied directly on the Type-2 α−cuts [[$x\alpha 1$, $x\alpha 2$], [$x\alpha 3$, $x\alpha 4$]] working as summarized in [12].

IV. A SYSTEMATIC APPROACH TO THE IT2FLS CLASSIFIER DESIGN

A. A Brief Introduction to T2FLS Classification Although the concept of T2FL was introduced in the fuzzy community over three decades ago, it remained in the realm of theoretical studies until recent work by Karnik and Mendel [14]. The introduction of an IT2 fuzzy set (FS), re-definition of T2 fuzzy operators and T2 inference mechanism have encouraged further advancement. Finally, the development of computationally efficient algorithms for T2FLSs has led to a revival of marked interest in their practical applicability to address a broad spectrum of problems where more than static imprecision in data needs to be accounted for. The concept of T2FL can be briefly described as expanded conventional T1FL based on FSs that are themselves fuzzy. In consequence, another dimension of fuzziness is introduced to the definition of a T2FS. The two-dimensional domain of support for additional secondary membership functions, referred to as a foot of uncertainty (FOU) [15], plays an important role in handling inconsistently varying information content. The enhanced flexibility in modeling the associated uncertainty due to the increased number of degrees of freedom underlies the potential of T2FLSs to outperform their T1 counterparts in problems where classification or approximation is to be made under uncertain, variable conditions. On the other hand, special care has to be exercised in T2FLS development in order to appropriately exploit the T2FL apparatus for handling uncertainty without sacrificing its generalization capability. This objective underpins the investigations into optimal design approaches in the domain of IT2FL classification. The IT2FL methodology is targeted in this work due to its computationally efficient implementation, which is an important asset in practical applications considering the complexity overhead of general T2FL tools. An IT2FLS relies on IT2FSs, whose secondary membership functions over the FOU are constant and equal one [15]. This substantially simplifies operations on FSs and facilitates transparent flow of uncertainties through a T2FLS. Here, IT2FSs with uncertain mean are utilized in the framework of the proposed fuzzy classifier. The FOU of such FS is presented in Fig. 2, which illustratively juxtaposes a T1FL and T2FL rule pattern adopted in the reported study.

B. IT2FLS Design

The rule base of the IT2FLS classifier developed in this work is of Mamdani type. Thus, a template of a fuzzy rule, shown in Fig. 2, is the following:

$$\text{IF } X_1 \text{ is } \tilde{A}_1 \text{ AND ... AND } X_{2N_{win}} \text{ is } \tilde{A}_{2N_{win}}$$
$$\text{THEN } class \text{ is } \left[c_{left}, c_{right}\right].$$

FSs *Xi* (*i*=1,..,2*Nwin*) are the T1 fuzzified components (Gaussian T1FSs) of an input feature vector ***x*** (cf. (1) in section IIB) to account for the possibility of stationary uniform noise present in the feature space. *Ãi*’s denote IT2FSs with uncertain mean and *C* is the interval centroid of the consequent T2FS representing the class that the input feature vector is assigned to. Hence, the IT2FLS rule base models uncertainty related to the variability of EEG features and the vagueness of a crisp MI label, i.e. left vs. right. The latter concept accounts for the difficulty in producing unambiguous mental task category.To facilitate gradient-based optimization, all the FSs are Gaussian.The IT2FLS classifier was designed in a two-stage procedure, inspired by general FLS methodology. Firstly, an initial fuzzy rule base was identified and secondly, its parameters were tuned using a global optimization approach. The design was conducted on the so-called calibration data set, split into a validation and a training subset. The final evaluation was performed on an unseen test data set.

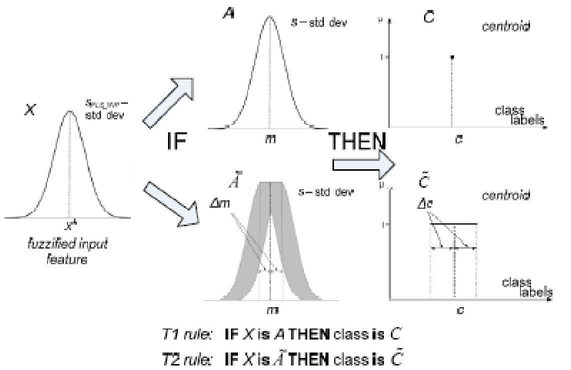

Figure 2. Illustrative comparison of T1FL and T2FL rule patterns.

1) Fuzzy Rule Base Structure Identification

The dimensionality of the EEG feature space, i.e. 2*Nwin* (cf. (1) in section IIB), determines the number of the fuzzy classifier's inputs. An initial fuzzy rule base was found through partitioning of the input space domain since it has been considered as one of the most effective methods of FLS structure identification. The main objective of this design phase was to obtain a compact data representation that reflects the underlying distribution of the features and thus captures their salient characteristics preserving at the same time their class assignments. To this end, a general clustering approach was adopted to devise a conventional prototype T1FLS rule base and in the spirit of partially-dependent initialization [15], it was then extended to serve as an initial T2FLS framework. In the first place, a mapping-constrained agglomerative (MCA) clustering algorithm was employed to reinforce the consistency in the mapping from the input to the output space. The MCA algorithm has been proven to be robust to noise and outliers that can affect the input-output relationship [22]. However, due to the excessive susceptibility of an original single-pass (sp) MCA to variations in the input data ordering, a heuristic modification was found essential. To this end, a multipass (mp) MCA algorithm was developed in this work. Firstly, an original MCA was iterated several times (parameterized) with the core input data appended with the data points representing means of clusters found in the previous iteration. The core data were shuffled at each iteration. Moreover, for every iteration the record of a cluster validity index, based on the classification performance of the prototype T1FLS (see below) reported on a separate validation set, serving as a performance measure of the given cluster structure was kept. The maximum of this measure determined the iteration that resulted in the target cluster structure. The prototype T1FLS rule base was derived without any extra parameters from the clusters identified using the MCAbased scheme with the number of rules equal to the number of clusters. An unquestionable asset of the MCA in this regard stems from the fact that it provides information not only about the cluster position in the input space (the cluster means) but also determines their spread in terms of the standard deviation estimate. Moreover, initialization of the corresponding fuzzy rule consequents is straightforward due to the consistency in the input-output mapping promoted by the algorithm. As shown in Fig. 2, the T1FLS prototype's consequents are expressed in terms of crisp class labels, -1 and 1, associated with left and right MI, respectively. In addition, for a comparative evaluation, the fuzzy c-means (FCM) clustering was employed due to its wide applicability in fuzzy rule base identification [23]. The algorithm requires the prior assumption of the number of clusters, which was selected in this work based on the above-mentioned cluster validity index. The input data space was clustered and the resultant cluster centers projected on each input dimension served as rule prototypes. The widths of the FSs were calculated as the onedimensional standard deviations of the subset of the input data points with the membership degree in the corresponding clusters above a certain threshold (parameterized). Since FCM does not explicitly account for the mapping between the input and the output space, the fuzzy rule consequents were uniformly randomized in the interval [-1,1]. An arbitrary initialization scheme was also verified in this work due to its relative simplicity and thus low computational cost. An algorithm, similar to Wang-Mendel one-pass method [15], consisting of dividing each dimension of the universe of discourse into a given number of intervals (parameterized) associated with arbitrary FSs and then constructing a rule base from the combination of FSs with the highest firing degrees for training data set was used. The rule firing degree was calculated additively over all input data points and a parameterized threshold was applied. The corresponding consequents were initialized randomly as in the FCM-based approach. A clear disadvantage of this method is the uniformity of the shapes (the same widths) and the distribution of FSs, which do not necessarily reflect the original data structure. As mentioned, the cluster validity index was exploited at this design stage to identify an optimal set of parameters for all the initialization schemes considered in this work. This was accomplished in the framework of 5-fold crossvalidation (CV) on a calibration data set. Thus, the initialization methods are assumed to have already been in their optimal setups before the next steps of the fuzzy classifier design are taken. The outcome of their comparative analysis in the configuration with the fully developed IT2FLS classifiers is discussed in section IV. After the prototype T1FLS rule base was initialized, it was extended to serve as a framework for an IT2FLS. Each T1FL rule was described in terms of its antecedent FSs *Ai* (*i*=1,.., 2*Nwin*), parameterized with vector $\boldsymbol{m}$ of their means and vector $\boldsymbol{s}$ of their standard deviations, and a crisp consequent, *c*. As can be noticed in Fig. 2, the uncertainty bounds of the FSs defining the antecedent and the consequent part of an IT2FL rule can be controlled by additional quantities, Δm and Δc, respectively. Therefore, the formulae for IT2FL rule induction from a classical T1FL rule prototype are straightforward:

$$m_1 = m - \Delta m, \quad m_2 = m + \Delta m$$

$$c_{left} = c - \Delta c, \quad c_{right} = c + \Delta c.$$

……………………………..(4)

Vectors ***m***1 and ***m***2 refer to the lower and the upper bound of the uncertain means in IT2FSs, *Ãi* , and *cleft*, *cright* define the consequent centroid (cf. section IIIB). The standard deviations, ***s***, of prototype T1FSs are kept the same for the resultant IT2FSs. The constrained parameterization of Δm with a multiplicative factor *dm* in (5) facilitated parameter selection.

$$\Delta m = dm\, s\,.$$

……………………(5)

Finally, ***sfuzz_inp*** used in the description of the fuzzified inputs was set as a scaled (scalar *a*) vector of the standard deviations of the input features in a training set. The parameters *dm*, Δc and *a*, assumed to be homogeneous for the entire rule base, determine the initial bounds of the uncertainty modeled in the system. Their selection procedure is discussed in subsection IIIB.3.

2) Parameter optimization – learning approaches

The second stage of the IT2FLS classifier design, after setting up an initial rule base, the quantities such as ***m***1, ***m***2, ***s***, *cleft*, *cright* and ***sfuzz_inp*** were tuned for every rule. A global performance optimization approach was adopted in this regard. The proposed learning algorithm is based on the concept of steepest gradient descent with the mean square error loss function since the classifier's output is continuous (with thresholding in the recall phase when a dichotomous class label is needed). The training method consists of three stages and heuristically combines two approaches known in the domain of IT2FLSs, the conventional steepest gradient descent algorithm developed by Liang and Mendel [24] and the method based on the dynamic optimal rate theorem [25][26]. This hybridization led to more robust and effective searching of a multimodal nonlinear space for an optimal configuration of the system parameters than the conventional Liang and Mendel's approach in the given pattern recognition problem. In the first learning stage, the conventional steepest descent was applied with learning rates being reduced by a constant factor every 10 epochs. The selection of their initial values was an important part of the design. An algorithm based on the dynamic optimal rate theorem was applied in the second phase to accelerate the optimization of the parameters of the fuzzy rule consequents. In particular, the combination of sample-by-sample training of the standard deviations s_{fuzz_inp} and the antecedent parameters $m_1, m_2, s, c_{left}, c_{right}$ and s_{fuzz_inp} a batch update of the consequents c_{left} and c_{right} was adopted. With the help of a validation data-set, an early stopping criterion was applied in the first and the second stage for terminating the training process, and more importantly, to enhance generalization capabilities of the classifier. In the third stage, the IT2FLS's parameters were fine tuned using an algorithm similar to that of the first stage. However, the learning rates were significantly reduced and the number of epochs was limited to an arbitrary number of 5. The resulting setup of the system parameters was accepted only if the classifier's performance in terms of the CA rate improved in comparison with the outcome of the second stage. Otherwise, the parameter configuration was rolled back. An analogous learning algorithm was developed for a T1FLS classifier to conduct a fair comparative analysis.

3) Experimental setup

The IT2FLS design process was used in two experimental paradigms adopted in this work. In the first one, the rule base parameters, *dm*, Δc , *a*, and initial learning rates α were selected on one-session data using an extensive grid search in the parameter space. To this end, a multiple-run 5-fold CV procedure was employed with the data split into training (60% of the one-session data set), validation (20%) and test (20%) subsets. The average test classification error (over all 5 test subsets and multiple runs) served as a criterion for the parameter identification.It provided an estimate of the classifier's within-session generalization properties. For greater clarity, it should be re-emphasized that the selection of parameters for rule base initialization, described in subsection IIIB.1, was performed independently at an earlier stage with the cluster validity index as a performance measure. In the second experimental setup, referred to as a singlepass training-test procedure, two-session data sets were involved. One session assumed a role of a calibration set whereas the other one served as unseen test data. Initial parameters were adopted from the earlier CV-based selection process conducted on the same session (calibration) data, which were next used (80% for training and 20% for validation) to design a fuzzy classifier for evaluation on the second session data set. This experimental paradigm allowed for verification of the IT2FLS's capability to deal with ses sionto- session non-stationarities in the EEG features.

V. CONCLUSIONS AND FUTURE PERSPECTIVES

The paper has reported an advanced methodology for the design and implementation of an IT2FL classifier for data intensive non-stationary systems. The initialization scheme involving the modified MCA clustering and the enhanced gradient descent-based learning algorithm were found effective in alleviating the problem of poor initial conditions and slow convergence. It can thus be concluded that there is a need to investigate systematic data-driven design approaches since a classifier's performance can be improved by the appropriate choice of a fuzzy rule base initialization method and a parameter learning scheme. The major aim of the comparative evaluation of several IT2FLS design variants was to identify the optimal approach to development of a robust fuzzy classifier for brain signal pattern recognition with its

potential applicability in other studies. Thus, the design approaches presented in this paper heavily exploited MI related brain signal recordings as a data source.
Their heuristic modifications were motivated by the need to effectively address some of the challenges of BCI classification. Robust handling of non-stationary effects observed in the relevant EEG features was the key objective in this regard. Since the problem is generic in nature, it is envisaged that the proposed methodology should lend itself to tackling its various forms in a range of real-world applications. The proposed IT2FLS has been shown to offer a promising potential in accounting for non-stationary long-term variability in neurophysiological data. The concept of addressing this problem, where no underlying functional model explaining various origins of non-stationarities and their manifestations is available, in a data-driven framework of uncertainty handling apparatus deserves special attention in this regard. It appears particularly beneficial in applications like BCI design due to its suitability for rapid system prototyping and development. Further research is intended to explore the ways that the uncertainty bounds of the classifier's output can be effectively exploited with the aim of further improving the performance of the classifier. A complementary analysis in on-line mode, involving a moving window approach to feature extraction and continuous classification, is also intended for future work. This will facilitate a more extensive assessment of the proposed method and allow for the use of other performance metrics.

References

[1] IEC-ISO, *Guide to the Expression of Uncertainty in Measurement*, October 1993.

[2] A. Ferrero and S. Salicone, "The random-fuzzy variables: A new approach to the expression of uncertainty in measurement," *IEEE Trans. Instrum. Meas.*, vol. 53, no. 5, pp. 1370 – 1377, October 2004.

[3] G. Mauris, L. Berrah, L. Foulloy, and A. Haurat, "Fuzzy handling of measurement errors in instrumentation," *IEEE Trans. Instrum. Meas.*, vol. 49, no. 1, pp. 89 – 93, February 2000.

[4] G. Mauris, V. Lasserre, and L. Foulloy, "A fuzzy approach for the expression of uncertainty in measurement," *Measurement*, vol. 29, pp. 165 – 177, 2001.

[5] L. A. Zadeh, "Fuzzy sets," *Information and Control*, vol. 8, pp. 338 – 353, 1965.

[6] L. A. Zadeh, "Fuzzy logic and approximating reasoning," *Synthese*, vol. 30, pp. 407 – 428, 1975.

[7] L. A. Zadeh, "Fuzzy sets as a basis for a theory of possibility," *Fuzzy Sets and Systems*, vol. 1, no. 1, pp. 9 – 34, 1978.

[8] D. Dubois, L. Foulloy, G. Mauris, and H. Prade, "Probability-possibility transformations, triangular fuzzy sets, and probabilistic inequalities," *Reliable Computing*, vol. 10, no. 4, pp. 273 – 297, August 2004.

[9] A. Ferrero, R. Gamba, and S. Salicone, "A method based on random fuzzy variables for on-line estimation of the measurement uncertainty of dsp-based instruments," *IEEE Trans. Instrum. Meas.*, vol. 53, no. 5, pp. 1362 – 1369, October 2004.

[10] A. Ferrero and S. Salicone, "A comparative analysis of the statistical and random-fuzzy approaches in the expression of uncertainty in measurement," *IEEE Trans. Instrum. Meas.*, vol. 54, no. 4, pp. 1475 – 1481, August 2005.

[11] L. A. Zadeh, "The concept of a linguistic variable and its application to approximate reasoning - Part I-II-III," in *Information Science*, 1975, vol. 8, 8, 9, pp. 199 – 249, 301 – 357, 43 – 80.

[12] A. Kaufmann and M. M. Gupta, *Introduction to fuzzy arithmetic: theory and applications*, Van Nostrand Reinhold Co., New York, NY, USA, 1985.

[13] J. M. Mendel, *Uncertain Rule-Based Fuzzy Logic Systems*, Prentice Hall, Upper Saddle River, NJ, USA, 2000.

[14] J. M. Mendel and R. I. Bob John, "Type-2 fuzzy sets made simple," *IEEE Trans. Fuzzy Syst.*, vol. 10, no. 2, pp. 117 – 127, April 2002.

[15] L.H. Chiang, E.L. Russel, and R. D. Braatz, *Fault Detection and Diagnosis in Industrial Systems*. New York: Springer-Verlag, 2001.

[16] M. Basseville and I.V. Nikiforov, *Detection of Abrupt Changes: Theory and Application*. Englewood Cliffs, NJ: Prentice-Hall, 1993.

[17] R.O. Duda and P.E. Hart, Pattern Classification and Scene Analysis. New York: Wiley, 1973.

[4] V. Venkatasubramanian, R. Rengaswamy, K. Yin, and S. N. Kavuri, "Review of process fault diagnosis—Parts I, II, III," *Comput. Chem. Eng.*, vol. 27, no. 3, pp. 293–346, 2003.

[18] P. Frank, "Fault diagnosis in dynamic systems using analytical knowledge- based redundancy—A survey and some new results," *Automatica*, vol. 26, no. 3, pp. 459–474, 1990.

[19] M.Y. Chow, R.N. Sharpe, and J.C. Hung, "On the application and design

of artificial neural networks for motor fault detection. Parts I, II," *IEEE Trans. Ind. Electron.*, vol. 40, no. 2, pp. 181–196, Apr. 1993.

[20] A. Kuh, "Comparison of tracking algorithms for single layer threshold networks in the presence of random drift," *IEEE Trans. Signal Process.*, vol. 45, no. 3, pp. 640–649, Mar. 1997.

[8] L. Rutkowski, "Adaptive probabilistic neural networks for pattern classification in time-varying environment," *IEEE Trans. Neural Net.*, vol. 15, no. 4, pp. 811–827, Jul. 2004.

[21] J. del R. Millan et al. "A local neural classifier for the recognition of EEG patterns associated to mental tasks," *IEEE Trans. Neural Net.*, vol.13, pp.678–685, 2002.

[22] P. Shenoy et al., "Towards adaptive classification for BCI," *J. Neural Eng.*, vol.3, pp.13–23, 2006.

[23] P. Sykacek et al., "Probabilistic methods in BCI research," *IEEE Trans. Neural Syst. Rehab. Eng.*, vol.11, pp.192–195, 2003.

[24] J.R. Wolpaw et al., "Brain-computer interfaces for communication and control," *Clin. Neurophysiol*, vol.113, pp.767–791, 2002.

[25] S. Mitra,, S.K. Pal, "Fuzzy Sets in Pattern Recognition and Machine Intelligence," *Fuzzy Sets and Systems*, vol.156, pp.381–386, 2005.

[26] N.N. Karnik, J.M. Mendel, "An introduction to type-2 fuzzy logic systems," in *Proc.1998 IEEE FUZZ Conf.*, Anchorage, AK, pp.915–920.

[27] J.M. Mendel, *Uncertain Rule-Based Fuzzy Logic Systems: Introduction and New Directions*, Prentice-Hall, USA, 2001.

[28] J. Zeng and Z.-Q. Liu, "Type-2 Fuzzy Hidden Markov Models and Their Application to Speech Recognition," *IEEE Trans. Fuzzy Sys.*, vol. 14, pp.454–467, 2006.

[29] H. Wu and J. Mendel, "Classification of Battlefield Ground Vehicles Using Acoustic Features and Fuzzy Logic Rule-Based Classifiers," *IEEE Trans. Fuzzy Sys.*, vol. 15, pp.56–71, 2007.

[30] W.W. Tan, C.L. Foo, and T.W. Chua, "Type-2 Fuzzy System for ECG Arrhythmic Classification," in *Proc. FUZZ-IEEE 2007*, London, 2007.

Intelligence, 2007, pp. 193–200.

Principles, Clinical Applications and Related Fields;