Evolving neuro-fuzzy system based online identification of a bio-inspired flapping wing micro aerial vehicle

Author:
ferdaus, M; Anavatti, SG; Garratt, M

Publication details:
IEEE symposium series on computational intelligence
v. 2018-January
pp. 1 - 8
9781538627259 (ISBN)

Event details:
Ieee SSCI
Honolulu
2017-11-27 - 2017-12-01

Publication Date:
2018-02-08

Publisher DOI:
https://doi.org/10.1109/SSCI.2017.8285226

Evolving Neuro-Fuzzy System based Online Identification of a Bio-inspired Flapping Wing Micro Aerial Vehicle

Md Meftahul Ferdaus*
School of Engineering
and Information Technology
UNSW Canberra@ADFA
Canberra BC 2610, Australia
Email: m.ferdaus@unsw.edu.au

Mahardhika Pratama
School of Computer Science
and Engineering
Nanyang Technological University
Singapore, 639798, Singapore
Email: mpratama@ntu.edu.sg

Sreenatha G. Anavatti, and
Matthew A. Garratt
School of Engineering
and Information Technology
UNSW Canberra@ADFA
Canberra BC 2610, Australia
Email: s.anavatti@adfa.edu.au;
M.Garratt@adfa.edu.au

***Abstract*—Bio-Inspired Flapping Wing Micro Air Vehicles (BIFW MAVs) are highly nonlinear and overactuated system. Besides, they may suffer from various uncertainties and perturbations like wind gust, sensor error etc. Modelling of such complex nonlinear system by considering the uncertainties is very difficult for the conventional first principle methods. However, numerous advantages of BIFW MAVs such as vertical take-off and landing, hovering, quick turn, and enhanced manoeuvrability attract researchers to develop their accurate modelling, and to do so Evolving Intelligent Systems (EISs) is an appropriate candidate since they do not need any information about the system dynamics. In this work, an advanced EIS called Generic Evolving Neuro-Fuzzy Inference System (GENEFIS) is employed to identify a four-wing BIFW MAV Multi Input Multi Output nonlinear model on the fly from the data stream, where an efficient online identification of the BIFW MAV model is observed.**

I. Introduction

Micro Aerial Vehicles (MAV) is a miniature of autonomous Unmanned Aerial Vehicles (UAVs), which is also known as μUAV. Usually, the MAVs have a maximum dimension of 6 inches with a gross take-off weight of nearly 200 g. They are similar to a small bird or insect with a flight velocity ranging between 10 ms^{-1} and 20 ms^{-1}. MAVs are classified into three subdivisions, namely fixed wing, rotary wing and flapping wing MAVs. Recently, MAVs are becoming more practical and affordable due to developments in batteries, microelectronics and sensor technology, and advancement in micro-electromechanical arrangements and micro-manufacturing techniques. Among various MAVs, bio-inspired flapping wing (BIFW) MAVs are becoming popular among researchers due to their capability to perform at lower Reynolds numbers compare to their fixed wings, and rotary wings counterpart. The BIFW MAVs are able to facilitate critical manoeuvres like vertical take-off and landing, gliding, roll banking, backward and sideways flying, which are not possible for similar sized fixed or rotary wing MAVs. Furthermore, BIFW MAVs can generate rapid acceleration during manoeuvres. The biggest challenges, benefits and feasibility of utilizing BIFW as MAV are described in [1], [2]. These huge benefits of BIFW MAVs make worthy to investigate the flight dynamics of various flapping wing creatures of natural world. The investigations on their flight dynamics of last two decades are discussed in [3]–[5]. From their investigation it is observed that Hummingbirds and dragonflies are able to hover and dart deftly and quickly, and bumblebees can fly without any lift coefficient, which are some highly attractive features for BIFW MAVs. Based on the source of mimicking, BIFW MAVs can be divided into two classes, such as 1) Ornithopters (bird-like), and 2) Entemopters (insect-like). Between this two categories, Entemopters are more appropriate as FW MAV due to their hovering and manoeuvring capability in tight space. Besides, some of the insects can create an asymmetric flapping angle and flapping frequencies or can tilt the stroke planes to generate controlled forces and moments [6].

The advantages of various flapping wing creatures described in the previous paragraph inspire researchers to develop appropriate BIFW MAV model and their efficient control methods. Nguyen et al. [7] fabricated an Entemopter with two fixed and two flapping wings. Their Entemopter has the capability of producing a vertical thrust to lift-off a weight of approximately 14 grams at 10 Hz frequency. Their Entemopter demonstrated a successful free flight with an excellent control accuracy. Du et al. [8] developed a first principle FW MAV model, where they have evaluated their model by attaching with a Linear Quadratic Regulator (LQR) mechanism based light controller. Their FW MAV model is linearised to fit with the LQR flight controller. Besides, an iterative learning based tuner to tune the input weighting matrix of their LQR is used to deal with unmodelled parameters. Inspired by the unique feature of wing actuation system and flapping profile of dragonflies, a self-learning wing actuation system around a bearing mechanism for a Dragonfly-inspired MAV (DI-MAV) is fabricated in [9]. All the techniques of modelling the FW MAV discussed in this paragraph based on first principle technique, where the precise mathematical model is compulsory. However, the

978-1-5386-2726-6/17/$31.00 ©2017 IEEE

BIFW MAVs are highly nonlinear and over-actuated system. Therefore, they express a complex mathematical model, where the incorporation of various uncertainties is very difficult. An alternative solution to the problem is the utilization of model-free knowledge-based or data-driven techniques.

A. Related work

Fuzzy logic and Neural networks are the two popular data-driven or knowledge-based techniques employed in modelling various MAVs [10], [11]. A fuzzy neural network system is developed in [12] to achieve a stable hovering of an insect-like FW MAV. Stable flight test with high manoeuvrability is observed from their simulation results. Nonetheless, their controller is not implemented in hardware yet. To deal with the uncertainties exists in the nonlinear complex FW MAV model a radial basis function neural network (RBFNN) is utilized in [13]. The stability of their closed loop control system is verified through Lyapunov stability theorem. By selecting the appropriate control variables, a good trajectory tracking performance is observed from their simulation results. However, their control technique is not tested experimentally yet. A fuzzy controller namely hybrid adaptive fuzzy controller (HAFC) is developed in [14] to control a dragonfly-like FW MAV. Their HAFC can tune its parameters by using the hybrid adaptive law to minimize the tracking error. The simulation results of the HAFC are compared with a direct adaptive (DA) method based fuzzy controller (DAFC). Better tracking performance is observed from the HAFC than the DAFC. However, the controller is not tested experimentally yet. All the fuzzy and neural network techniques discussed until now are based on batch learning process, therefore static in nature. Some of them are adaptive due to their capability of tuning parameters. However, they cannot evolve their structure by adding, or deleting rules or neurons. Therefore, they cannot cope with the sudden changes in FW MAV's flight test. A solution to the problem is the utilization of the evolving intelligent system in fuzzy and neural network techniques. Recently researchers are trying to employ evolving fuzzy and neural networks in FW MAVs. An evolutionary algorithm called ModNet encoding is implemented in a recurrent neural network in [15] to evolve both the structure and weights of the neural network. Their evolutionary neuro controller can adjust the dihedral, the twist and the sweep of the wings at every moment to balance the FW MAV body. A Spiking Neural Network (SNN) based controller is developed for an FW MAV called RoboBee in [16], where they have utilized a reward-modulated Hebbian plasticity mechanism to adopt a leaky integrate-and-fire spiking neural network in flight.

To identify a highly nonlinear overactuated system like FW MAV, EIS is an appropriate candidate since they learn from scratch with no base knowledge and change their structure by adding or deleting rules to cope with changing system dynamics [17]. EIS fully work in a single-pass learning scenario which is scalable for online real-time requirement under limited computational resources such as MAVs platform. The EIS research area grows rapidly after seminal work by Angelov, namely eTS [18], eTS+ [19], Simpl_eTS [20], AnYa [21], [22], [23]. Inspired by various evolving fuzzy and neuro-fuzzy approaches, an advanced EIS based technique called Generic Evolving Neuro-Fuzzy Inference System (GENEFIS) proposed in [24], [25] is employed in our work for online identification of a four wings BIFW MAV. GENEFIS consists of some unique features such as: 1) the utilization of the Datum Significance plus (DS+) method as a rule growing mechanism minimizes the inconsistency problem of Datum Significance (DS) method; 2) integration of the Generalized Adaptive Resonance Theory+ (GART+) with an aim to upgrade the premise parameters with respect to input data distributions; 3) amalgamation of the fuzzily weighted generalized recursive least square (FWGRLS) method to tune the output parameters; 4) application of a fast kernel-based metric approach to capture fuzzy set and rule level redundancy; 5) casting of a new online feature selection technique, which has the ability to handle the curse of dimensionality; and 6) utilization of multivariate Gaussian function. These special characteristics of GENEFIS make it an appropriate candidate to identify a highly non-linear overactuated four-winged BIFW MAV with uncertainties.

II. BIFW MAV Platform

The FW MAV used in our work is a bio-inspired four wings Entemopter. The determination process of the BIFW MAV system dynamics is explained in [26], where they have developed an MAV flight simulator since flight simulators are often utilized due to the high cost and time consumption to develop the set-up for experimental flight test. From this flight simulator various manoeuvrability such as take-off, rolling, pitching, and yawing of an FW MAV is analysed. The flapping angle (ϕ) in the flapping profile of the BIFW MAV is presented as:

$$\phi(t) = \phi_a \cos(\pi f t) \quad (1)$$

where ϕ_a is the flapping amplitude in radian, f is the flapping frequency in Hz, t is the time is second. The angle of attack (α_{aoa}) can be expressed as:

$$\alpha_{aoa} = \alpha_{mn} - \alpha_p \sin(\omega dt + \Psi) \quad (2)$$

where α_{mn} is mean angle of attack in radian, α_p is amplitude of pitching oscillation in radian, dt is time step in seconds, and Ψ is the phase difference between the pitching and plunging motion. All the four wings of the FW MAV follows the same flapping profile.

In the simulator, each wing is controlled by an actuator. Body coordination of a four wings FW MAV is displayed in Fig. 1. Each actuator is controlled by eight (8) flapping parameters. A parametric analysis is accomplished to find the dominant flapping parameters. After a complete parametric analysis, it is found that the flapping amplitude is the dominant one among the eight parameters to control the FW MAV. By tuning the flapping amplitude, the take-off, rolling, and pitching motion is observed. Only for the yawing motion flapping phase needs to be tuned. Tuning of these parameters

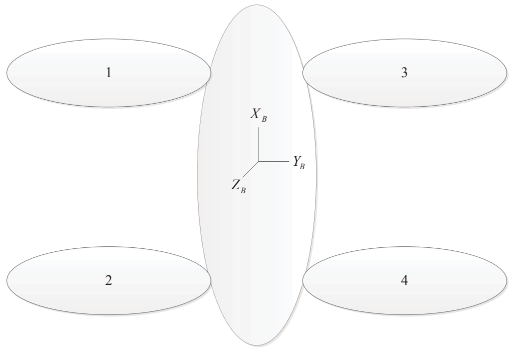


Fig. 1: Body coordination of an FW MAV. Numbers indicate the actuator number

TABLE I: Effects of flapping parameters in different manoeuvring of FW MAV (ϕ_0: Flapping amplitude(degree); Ψ: Phase(degree))

Actuators with corresponding Flapping parameter	Action
Actuator: 1, 2, 3, 4; ϕ_0: 90	Take-off
Actuator: 1, 2; ϕ_0: 90 and Actuator: 3, 4; ϕ_0: 60	Roll-right
Actuator: 1, 2; ϕ_0: 60 and Actuator: 3, 4; ϕ_0: 90	Roll-left
Actuator: 1, 3; ϕ_0: 90 and Actuator: 2, 4; ϕ_0: 60	Pitch-up
Actuator: 1, 3; ϕ_0: 60 and Actuator: 2, 4; ϕ_0: 90	Pitch-down
Actuator: 1, 4; Ψ: 90 and Actuator: 2, 3; Ψ: 60	Yaw-right
Actuator: 1, 4; Ψ: 60 and Actuator: 2, 3; Ψ: 90	Yaw-left

and their effect on various manoeuvring is summarized in in TABLE I .

From the FW MAV flight simulator the input output data is collected to develop the data-driven model, where the four input datasets are the four flapping amplitudes applied to four actuators. The rotational velocities (ω_{bx}, ω_{by}, and ω_{bz}) and translational velocities (v_{bx},v_{by}, and v_{bz}) of the FW MAV body in three dimensional plane are six output datasets. From four inputs and six outputs, a Multi-Input Multi-output (MIMO) nonlinear FW MAV model is identified online using the GENEFIS. The online learning policy of the GENEFIS is descried in the next section.

III. Online learning mechanism of GENEFIS

An advanced evolving fuzzy system called Generic Evolving Neuro-Fuzzy Inference System (GENEFIS) [24] is employed in this work to identify the BIFW MAV. GENEFIS is a TS fuzzy system that consists of multidimensional membership functions in the input space with ellipsoid contours in arbitrary positions. All the estimated one dimensional membership functions represent a portion in the input space partition by assigning the Gaussian functions own center and width. Thereby, generating 1st order first-order polynomials as consequent parts of the fuzzy rules. In GENEFIS, a typical fuzzy rule can be expressed as follows:

$$\text{IF } Z \text{ is } R_i, \text{ then } \eta_i = b_{0i} + b_{1i}\zeta_1 + b_{2i}\zeta_2 + ... + b_{ki}\zeta_k \quad (3)$$

where R_i represents the degree of membership of the $i-th$ rule constructed from a concatenation of fuzzy sets, k denotes the dimension of input feature, Z is an input vector of interest, b_i is the consequent parameter, ζ_k is the $k-th$ input feature. The predicted output of the model can be expressed as:

$$\hat{\eta} = \sum_{i=1}^{j} \Psi(\zeta)\eta_i(\zeta) = \frac{\sum_{i=1}^{j} R_i\eta_i}{\sum_{i=1}^{j} R_i} = \frac{\sum_{i=1}^{j} \exp(-(Z-\Theta_i)\Sigma_i^{-1}(Z-\Theta_i)^T)\eta_i}{\sum_{i=1}^{j} \exp(-(Z-\Theta_i)\Sigma_i^{-1}(Z-\Theta_i)^T)} \quad (4)$$

where Θ_i is the centroid of the $i-th$ fuzzy rule $\Theta_i \in \Re^{1\times j}$, Σ_i is a non-diagonal covariance matrix $\Sigma_i \in \Re^{k\times k}$, and k is the number of fuzzy rules.

A. Rule growing and pruning mechanism

The DS method developed in [27] is utilized in GENEFIS as a rule growing mechanism. The DS method is geared into the multivariate Gaussian membership function and first order consequent of GENEFIS. In this work DS method can be expressed as follows:

$$D_{sg_n} = |e_{rr}| \frac{\det(\Sigma_{j+1})^k}{\sum_{i=1}^{j+1} \det(\Sigma_i)^k} \quad (5)$$

where D_{sg_n} denotes the the significance of the $n-th$ datum, system error $|e_{rr}| = |tr_n - \eta_n|$, tr_n is the target and η_n is perdicted output from GENEFIS at $n-th$ episode. To reduce sensitivity to underfitting or overfitting effect, e_{rr} can be replaced by the global mean and standard deviation of e_{rr} as outlined in the original GENEFIS paper [24]. The condition in expanding the rule base utilizing (5) is $D_{sg} \geq \delta$, where δ is a predefined threshold. On the other hand, the Extended Rule significance (ERS) method is employed in GENEFIS to prune rules. A numerous modifications are made in ERS theory to fit them with GENEFIS. The ERS theory employed in GENEFIS can be expressed as:

$$\mathcal{E}_{inf}^i = \sum_{i=1}^{j+1} \eta_i \frac{\det(\Sigma_i)^k}{\sum_{i=1}^{j} \det(\Sigma_i)^k} \quad (6)$$

where $\mathcal{E}_{inf}^i \leq k_e$, k_e orchestrating a plausible tradeoff between compactness and simplicity of the rule base. In this work, the allocated value for δ is $\delta = [0.0001, 1]$, and $k_e = 10\%$ of δ.

B. Generalized Adaptive Resonance Theory+ (GART+)

Generalized Adaptive Resonance Theory (GART) [28] is used in GENEFIS as another technique of granulating input features and adapting premise parameters. It is observed that GART and its successor improved GART (IGART) suffers from a category growing problem. In GRAT the compatibility measure is done utilizing the maximal membership degree of a new datum to all available rules. In first round if the selected category expresses a higher membership degree than a predefined threshold σ_a, then the match tracking mechanism is committed. However, the first round winning category fails to beat the match tracking threshold σ_b, which deactivates that category and increases the value of threshold σ_a to find a better candidate. A larger width is required in the next selected category to cope with the increased value of σ_a, otherwise it fabricates a new category. Nonetheless, a category with a larger radii may contain more than one distinguishable data clouds and thereby marginalizing the other clusters in every training episode. In incremental learning environment this effect is known as *cluster delamination*. To relieve from the *cluster delamination* effect, in GENEFIS the size of fuzzy rule are constrained, which allows a limited grow or shrink of a category.

1) Category Choice: In GART+ [29], Bayess decision theory is utilized to determine the most compatible or winning category. The Bayesian concept not only consider the matching factor of a category to an injected datum, but also consider the population of the category through the category prior probability. A posterior probability of the $i-th$ category can be expressed as:

$$\hat{P}_r(\Psi_i|Z) = \frac{\hat{p}_r(Z|\Psi_i)\hat{P}_r(\Psi_i)}{\sum_{i=1}^{j}\hat{p}_r(Z|\Psi_i)\hat{P}_r(\Psi_i)} \tag{7}$$

where $\hat{p}_r(\Psi_i|Z)$ and $\hat{P}_r(\Psi_i)$ represents the likelihood and the prior probability correspondingly, which ca also be expressed as:

$$\hat{P}_r(Z|\Psi_i) = \frac{1}{(2\pi V_{Hi})^{1/2}}\exp(-(Z-\Theta_i)\Sigma_i^{-1}(Z-\Theta_i)^T) \tag{8}$$

$$\hat{P}_r(\Psi_i) = \frac{N_i}{\sum_{i=1}^{j} N_i} \tag{9}$$

where N_i denotes the number of times that $i-th$ category wins the competition. V_{Hi} determines the estimated hyper-volume of feature space covered by the $i-th$ category, which can be expressed as:

$$V_{Hi} = \det(\Sigma_i) \tag{10}$$

In short, when a category satisfy the condition $win = \arg\max(\hat{P}_r(\Psi_i|Z))$, then the category is a candidate to undergo a resonance.

2) Vigilance Test: There are two goals to perform the vigilance test. The first goal concerns about the capability of the winning category to accommodate a new datum. The second goal is to reduce the size of the category, where a rule is not allowed to have a volume higher than the threshold V_{Hmax}, that is calculated from $V_{Hmax} \equiv \sigma_b \sum_{i=1}^{j} V_{Hi}$. However, the conditions for attaining these two goals is presented below:

$$\text{Case I: } \phi_{win} \geq \sigma_a,\ V_{Hwin} \leq V_{Hmax} \tag{11}$$

where ϕ_{win} is the membership degree of the winning rule to seize the latest datum. More importantly, the condition in (11) is indicating the capability of the selected category to accommodate the newest datum and emphasizing on the limited size of a category. σ_a is usually set close to 1. Contrarily, the value of σ_b is set as [0.0001, 0.1]. Then the adaptation mechanism of focal point Θ_i, and the dispersion matrix Σ_i is generated by the equations as follows:

$$\Theta_{win}^{new} = \frac{N_{win}^{old}}{N_{win}^{old}+1}\Theta_{win}^{old} + \frac{(Z-\Theta_{win}^{old})}{N_{win}^{old}+1} \tag{12}$$

$$\Sigma_{win}(new)^{-1} = \frac{\Sigma_{win}(old)^{-1}}{1-\alpha} + \frac{\alpha}{1-\alpha} \frac{(\Sigma_{win}(old)^{-1}(Z-\Theta_{win}^{new}))(\Sigma_{win}(old)^{-1}(Z-\Theta_{win}^{new}))^T}{1+\alpha(Z-\Theta_{win}^{old})\Sigma_{win}(old)^{-1}(Z-\Theta_{win}^{old})^T} \tag{13}$$

$$N_{win}^{new} = N_{win}^{old} + 1 \tag{14}$$

where $\alpha = 1/(N_{win}^{old}+1)$, N_{win}^{old} denotes the number of training samples populating the winning cluster. Besides, (13) is able to prompt the training process since a direct adjustment of the dispersion matrix is occurring without the necessity to re-inverse the dispersion matrix. With respect to the conditions in (11), some pertinent likelihoods may emerge in the rehearsal process and they are outlined as follows:

Case II: $\phi_{win} < \sigma_a,\ V_{Hwin} > V_{Hmax}$

In this circumstance, the injected training sample cannot be touched by any rules of GENEFIS. The statistical contribution of the datum needs to be calculated by using (5). When both conditions are satisfied, a new rule is generated and it parameters are assigned as follows:

$$\Theta_{j+1} = Z \tag{15}$$

$$\text{diag}\left(\Sigma_{j+1}\right) = \frac{\max((\Theta_i-\Theta_{i-1}),(\Theta_i-\Theta_{i+1}))}{\sqrt{\frac{1}{\ln(\epsilon)}}} \tag{16}$$

where the value of ϵ is 0.5. Equation (16) ensures a a sufficient coverage of the newly added rule, which is proved in [30]. It helps GENEFIS to explore untouched regions in the feature space fitting a superfluous cluster at whatever point a relatively unexploited region or knowledge is fed, which is a mandatory element to confronting possible non-stationary and evolving qualities of the system being modelled. Note that proper initialization of inverse covariance matrix plays crucial role to the success of multivariate Gaussian fuzzy rule. Although it meets the e-completeness criterion, Equation (16) requires re-inversion phase which sometime leads to instability when the covariance matrix is not full-rank. As an alternative, the inverse covariance matrix can be initialized as:

$$\Sigma^{-1} = k_{fs}I \tag{17}$$

where k_{fs} is a user-defined parameter.

Case III: $\phi_{win} \geq \sigma_a,\ V_{Hwin} > V_{Hmax}$

This situation is indicating the capability of existing rule base to cover the current data easily. However, the width of the chosen cluster is oversized. This datum creates a redundancy when added to the rule base. To mitigate the adverse impact, one of the solutions is to replace the selected cluster merely by this datum. Then the fuzzy region is eased as follows:

$$\Theta_{win} = Z \tag{18}$$

$$\Sigma_{win}(new)^{-1} = \frac{1}{k_{win}}\Sigma_{win}(old)^{-1} \tag{19}$$

where k_{win} is a constant with a value of 1.1, and the width of the cluster is reduced until a desirable fuzzy region is obtained while satisfying $V_{Hwin} \leq V_{Hmax}$.

Case IV: $\phi_{win} < \sigma_a,\ V_{Hwin} \leq V_{Hmax}$

The same action is taken as in Case I, i.e. the adjustment process is executed to stimulate the category to move towards this training example.

C. Mechanism of Merging Rules

Redundancy of fuzzy rules is a major obstacle in the evolving neuro-fuzzy system. It occurs when new data samples fill up the gap between two different clusters, and they become overlapping, sometimes even indistinguishable. In such condition, a similarity measure is needed to quantify the similarity level of those clusters. The clusters are set to move together when they are deemed identical.

The kernel-based metric is a well-known feasible method to deal with the redundancy of fuzzy rules on the fly. In GENEFIS the kernel based metric method is not only refurbished to deal with redundancy of the fuzzy sets, but also to delve cues of fuzzy clusters redundancy. For example, the similarity or level of interaction between two Gaussian membership functions A and B can be obtained by comparing centroids and widths of two fuzzy sets in one joint formula [31].

$$S_{kerr} = e^{-|\Theta_A - \Theta_B| - |\rho_A - \rho_B|} \tag{20}$$

Equation (20) has the following interesting property:

$$\begin{aligned} S_{kerr}(A,B) = 1 &\Leftrightarrow |\Theta_A - \Theta_B| + |\rho_A - \rho_B| = 0 \\ &\Leftrightarrow \Theta_A = \Theta_B \wedge \rho_A = \rho_B \end{aligned} \tag{21}$$

$$S_{kerr}(A,B) < \epsilon \Leftrightarrow |\Theta_A - \Theta_B| > g \vee |\rho_A - \rho_B| > g \tag{22}$$

where g is a positive constant. The first condition is the case of perfect match, when the similarity measure between two fuzzy sets has a maximal degree, which is one. In the second condition it is observed that an embedded set is not coupled with a set covering it having a significantly larger width. It may result in an inaccurate exhibition of the data in one local region. Therefore, two clusters are merged together, only when the similarity of them is higher than the threshold $S_{kerr} \geq \epsilon$. In this work the threshold value of ϵ is set to 0.8. The merging between two similar fuzzy sets is occurred by using the following equation:

$$\Theta_{new} = (\max(U) + \min(U))/2 \tag{23}$$

$$\rho_{new} = (\max(U) - \min(U))/2 \tag{24}$$

where $U = \{\Theta_A \pm \rho_A, Theta_B \pm \rho_B\}$. Eigenvalue, eigenvector of the non-diagonal covariance matrix, or the distance from the centre to cutting point of ellipsoid is used to extract the radii of multivariate Gaussian function. The essence of the first method benefits from the eigenvalue and eigenvector of the non-diagonal covariance matrix and a specific α-cut value. Here the chosen value of α-cut is $e^{-1/2} \approx 0.6$. In this work the fuzzy set is built as follows:

$$\nu_i = \Theta_i \tag{25}$$

$$\rho_i = \max_{k=1,\ldots u}\left(\frac{r}{\sqrt{\lambda_k}}\cos(\varphi(e_k, b_i))\right) \tag{26}$$

where λ_k denotes the eigenvalue for kth dimension, e_k represents eigenvector with respect to kth dimension, ν_i labels the modal value of the fuzzy sets which is usually centre, ρ_i stands for the spread of the set., b_i denotes the vector expressing ith axis $b_i = (0, 0, ..., 1, ..., 0)$, where the value of 1 indicates the ith position. φ shows the angle spanned between b_i and e_k.

$$\varphi(e_k, b_i) = \arccos\left(\left|\frac{e_j^T b_i}{|e_j||b_i|}\right|\right) \tag{27}$$

The first method incurs an accurate approximation of the fuzzy set of multivariate Gaussian kernel. However, it suffers from huge computational efforts due to quantifying eigenvalue and eigenvector in every training episode. Thereby, the importance of another method is that the radii of the fuzzy sets can be bestowed by the distance from the center to the cutting point or the axis-parallel intersection with the ellipsoid which is expressed as:

$$\rho_i = \frac{r}{\sqrt{\Sigma_{ii}}} \tag{28}$$

where Σ_{ii} is the diagonal elements of the inverse covariance matrix. Though the second method is less accurate than the first method, it is computationally faster than the first one.The fuzzy set merging scenario is not a part of the main training process. It is executed after the training process when showing the fuzzy rules to the operator to improve fuzzy rule transparency. To reduce the model's complexity, the fuzzy rule merging strategy is carried out. The similarity of each fuzzy set is encapsulated and blended to a single value by using $minimum$ operator as follows:

$$S_{rule}(A,B) = \min_{h=1,\ldots k}(S_{kerr}(A_h, B_h)) \tag{29}$$

If $S_{rule} > \epsilon_1$, where ϵ_1 is a tolerable similarity degree whose default value is 0.8, the two rules A and B can be merged. The more superfluous rule can be evicted and consequently alleviate rule-base burden. By applying a weighted average of two rules, following rule merging formulas are achieved:

$$\Theta_{dm}^{new} = \frac{\Theta_{dm}^{old} N_{dm}^{old} + \Theta_{i+1} N_{dm+1}}{N_{dm}^{old} + N_{dm+1}^{old}} \tag{30}$$

$$\Sigma_{dm}(new)^{-1} = \frac{\Sigma_{dm}(old)^{-1} N_{dm}^{old} + \Sigma_{dm+1}(old)^{-1} N_{dm+1}^{old}}{N_{dm}^{old} + N_{dm+1}^{old}} \tag{31}$$

$$N_{dm}^{new} = N_{dm}^{old} + N_{dm+1}^{old} \tag{32}$$

where the rule dm is a more dominant rule surrounded by more data points than the rule $dm+1$, i.e. $N_{dm} > N_{dm+1}$. In a TS fuzzy model dissimilar consequent parameters of two fuzzy rules may observed despite of conceiving similarity in the premise part. It may create an inconsistency weakening interpretability of explanatory module. As a solution the consequent parameters of the two fuzzy rules can be merged.

This merging is done according to a combination rule, which has some synergies to the notion of participatory learning [32]:

$$\omega_{dm}^{new} = \omega_{dm}^{old} + \gamma g(\omega_{dm}^{old} - \omega_{dm+1}^{old}) \tag{33}$$

$$\gamma = \frac{N_{dm}^{old}}{N_{dm}^{old} + N_{dm+1}^{old}} \tag{34}$$

where

$$g = \begin{cases} 1 & \text{if } S_{rule} \leq S_{out} \\ 0 & \text{if } S_{rule} > S_{out} \end{cases} \tag{35}$$

labels a degree of consistency, S_{out} represents the similarity of the output parameters. When the consequent parameters are more similar than the antecedents, g becomes 1, and consequent merging are occurred. A feasible way to predict the similarity between two local sub-models is by means of the angle formed between them as expressed below:

$$\varphi = \arccos\left(\left|\frac{\omega_i^T \omega_{i+1}}{|\omega_i||\omega_{i+1}|}\right|\right) \tag{36}$$

where the range of φ is $[0, \pi]$, $w_i = [b_{1i}, b_{2i}, ..., b_{k,i}]$, and $w_{i+1} = [b_{1,i+1}, b_{2,i+1}, ..., b_{k,i+1}]$. Now the similarity between two hyperplanes is defined in [31] as follows:

$$S_{out}(\eta_i, \eta_{i+1}) = \begin{cases} 1 - \frac{2}{\pi}\varphi & \text{where } \varphi \in \left[0, \frac{\pi}{2}\right] \\ \frac{2}{\pi}\left(\varphi - \frac{\pi}{2}\right) & \text{where } \varphi \in \left[\frac{\pi}{2}, \pi\right] \end{cases} \tag{37}$$

The dynamic adaptation of the premise parameters of fuzzy rules are considered as the major reason of redundancy problem. There is no necessity to execute the similarity check in every training observation as it retards the training process. Thereby, in GENEFIS the rule merging procedure is carried out only when the resonance is carried out to minimize the complexity.

D. Mechanism of Online Feature Selection

A mechanism which enables a device to select the most informative input attributes on the fly is known as online feature selection mechanism, which is missing in majority of the evolving neuro-fuzzy systems since they utilize a standalone mechanism. To achieve this online mechanism in GENEFIS, an input selection algorithm called Input Selection (IS) method is utilized which has the capability to pinpoint the most important input features in online during the training process. The IS is embedded in GENEFIS's learning machine to switch off input features those are inactive in their lifespan. How a input variable in the nth training episode is contributing is determined by the significance of the input and output parameters of the fuzzy system as follows:

$$IS_h = \tau_h I_h \tag{38}$$

where τ denotes the sensitivity of the output parameters or local subsystems of hth input feature, whereas I_h stands for the sensitivity of the input parameters of the hth attribute. The final expression of τ_h, I_h, and eventually IS_h is epitomized in (39), (40), and (41) respectively as follows:

$$\tau_h = \frac{\sum_{i=1}^{j} \Omega_{hi}}{\sum_{h=1}^{k} \sum_{i=1}^{j} \Omega_{hi}} \tag{39}$$

$$I_h = \frac{\det(\Sigma_h)^j}{\sum_{h=1}^{k} \det(\Sigma_h)^j} \tag{40}$$

With the incorporation of both the input and output parts, the IS method can finally be expressed as follows:

$$IS_h = \frac{\det(\Sigma_h)^j}{\sum_{h=1}^{k} \det(\Sigma_h)^j} * \frac{\sum_{i=1}^{j} \Omega_{hi}}{\sum_{h=1}^{k} \sum_{i=1}^{j} \Omega_{hi}} \tag{41}$$

where $\Omega_{hi} = [b_{h1}, b_{h2}, ..., b_{hj}]$ is a composition of the hth input parameters in every rule. If $I_h \leq k_{in}$, where k_{in} is a threshold regulating deleting frequency of the inconsequential input features, and $k_{in} = 0.0001$, the observed input attributes are superfluous during the training process. They can be dispossessed by sustaining the accuracy the model, on the other hand they may be maintained for next training episodes. Because GENEFIS starts its learning process from scratch, the online feature selection scenario is only after observing several training samples to prevent the input attributes to be deleted too early during the training phase.

E. Mechanism of Adapting Consequent Parameters

For adapting the rule consequent the Fuzzily Weighted Generalised Recursive Least Square (FWGRLS) method [29] is used in GENEFIS, which can be expressed as follows:

$$K_l(n) = P_i(n-1)\zeta_{en}(\frac{1}{\wedge_i(n)} + \zeta_{en} P_i(n-1)\zeta_{en})^{-1} \tag{42}$$

$$P_i(n) = P_i(n-1) - K_l(n)\zeta_{en} P_i(n-1) \tag{43}$$

$$\begin{aligned} \eta_i(n) &= \eta_i(n-1) - \alpha P_i(n) \bigtriangledown \Psi(\eta_i(n-1)) \\ &\quad + K_l(n)(tr(n) - \zeta_{en}\eta_i(n)) \end{aligned} \tag{44}$$

where $K_l(n)$ is the Kalman gain, and $P_i(n)$ is the covariance matrix, $\zeta_{en} = [\zeta_1^n, \zeta_2^n, ..., \zeta_k^n]^T$ is the extended input vector, $\eta_i(n)$ is the local subsystem of the ith rule, α is the regularization parameter, which is set close to zero $\alpha \approx 10^{-7}$. In GENEFIS, the quadratic weight decay function is utilized and can be expressed as follows:

$$\Psi(\eta_i(n-1)) = \frac{1}{2}(\eta_i(n-1))^2 \tag{45}$$

Its gradient can be obtained as follows:

$$\bigtriangledown \Psi(\eta_i(n-1)) = \eta_i(n-1) \tag{46}$$

The function aims to trigger the weight being adjusted to reduce by a factor proportional to its current value. As a consequence, a small fluctuation is observed in the magnitude of the output parameters which intensifies the generalization capability. In the Bayesian viewpoint, a decay term in the cost function features a prior distribution of the local sub-model. In the case of quadratic decay function, the prior distribution of output parameters is Gaussian which renders the output

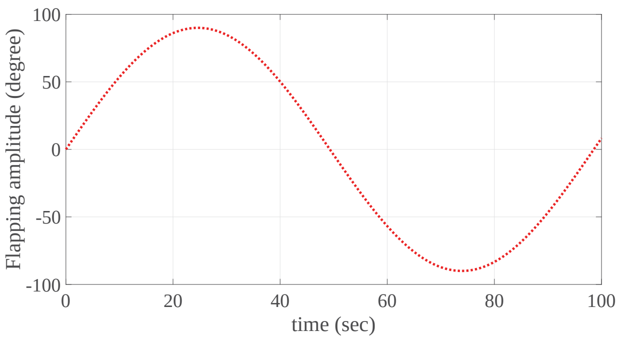


Fig. 2: Input: Flapping amplitude

TABLE II: RMSE in BIFW MAV Online Identification

BIFW MAV properties	RMSE (training)	RMSE (validation)
v_{bx}	0.00193	0.00107
v_{by}	0.00191	0.00327
v_{bz}	0.00174	0.00038
ω_{by}	0.00268	0.00738
ω_{bz}	0.00307	0.00060

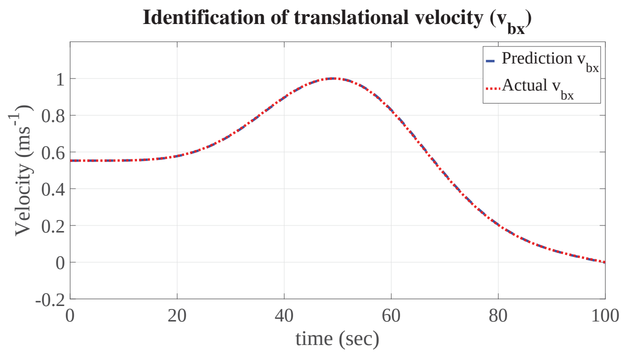
parameters to be cautiously distributed [33]. This method can be seen as a derivation of the fuzzily weighted generalized recursive least square (FWGRLS) method [18].

IV. Online Identification Results of BIFW MAV System

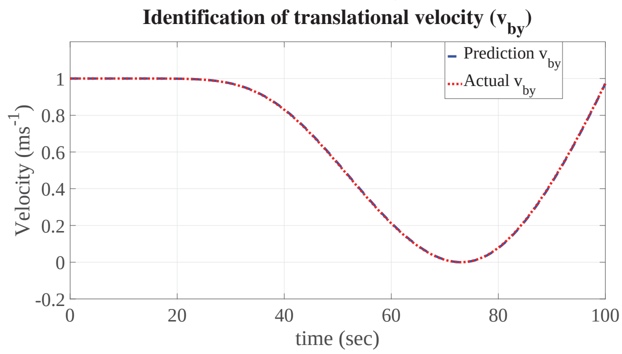
The data used for the MIMO nonlinear BIFW MAV online identification is based on a 100 sec simulation in Simulink with a time step of 0.1 sec. In a physical BIFW MAV model, they can change their flapping amplitude within a range of $-90°$ and $90°$. Therefore, a sinusoidal flapping amplitude varying between $-90°$ and $90°$ is applied to all four actuators of the four wings of BIFW MAV as shown in Fig. 2, which helps the GENEFIS in the online identification of BIFW MAV since the input datasets are obtained within the maximum possible range. Besides, they are smooth in the way of changing, which can easily be identified by GENEFIS. In GENEFIS, only one multivariate Gaussian membership function is utilized. From Fig. 3 and Fig. 4 it is clearly observed that all the translation velocities (v_{bx},v_{by}, and v_{bz}), and the rotational velocities (ω_{by}, and ω_{bz}) are identified precisely. In our identification 60% samples are used for training and remaining 40% samples are for validation. The root mean square error (RMSE) is calculated in both training and validation period and summarized in TABLE II. From all these results it is observed that the highly nonlinear and overactuated BIFW MAV is identified effectively on the fly using GENEFIS.

V. Conclusion

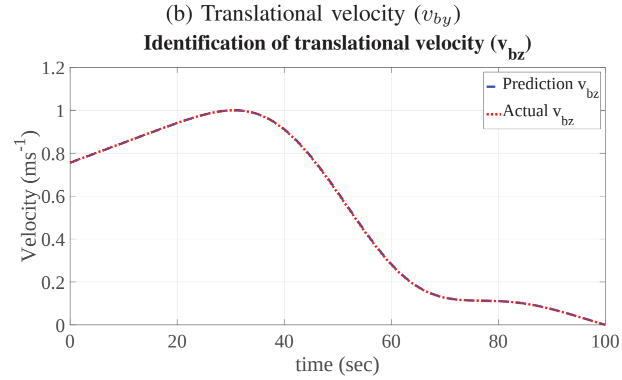
In this work, a data-driven model free identification of a BIFW MAV model on the fly is witnessed through GENEFIS. The amalgamation of Datum Significance plus (DS+) method



(a) Translational velocity (v_{bx})



(b) Translational velocity (v_{by})



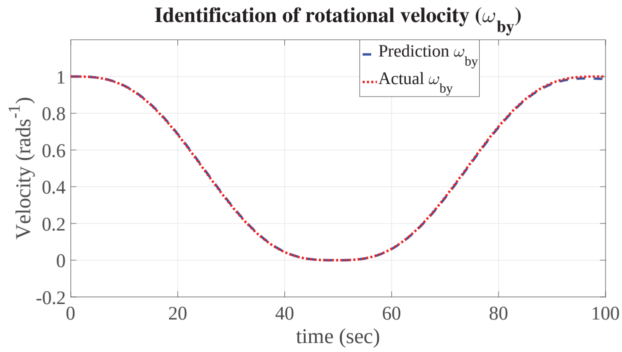
(c) Translational velocity (v_{bz})

Fig. 3: Identification of FW MAVs translational velocity (v_{bx}, v_{by}, v_{bz})

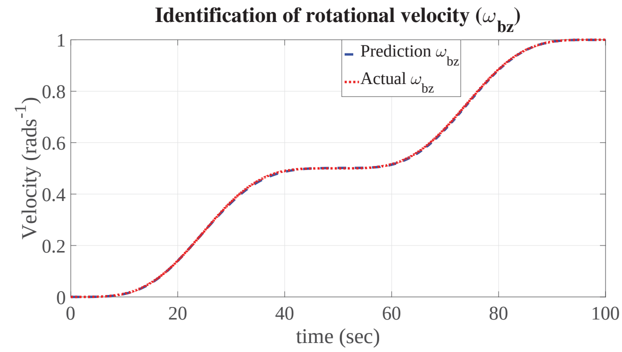
as a rule growing mechanism, integration of the Generalized Adaptive Resonance Theory+ (GART+), integration of the FWGRLS method to tune the output parameters, and application of a new online feature selection technique in GENEFIS make it an appropriate candidate to model a highly nonlinear MIMO BIFW MAV in online. The accurate identification of the translational and rotational body motion of BIFW MAV with a maximum RMSE of only 0.007 evaluates its performance. In future, GENEFIS based flexible controller will be developed for FW MAVS.

References

[1] C. P. Ellington, "The novel aerodynamics of insect flight: applications to micro-air vehicles," *Journal of Experimental Biology*, vol. 202, no. 23, pp. 3439–3448, 1999.



(a) Rotational velocity (ω_{by})



(b) Rotational velocity (ω_{bz})

Fig. 4: Identification of FW MAVs translational velocity ($\omega_{bx}, \omega_{by}, \omega_{bz}$)

[2] D. J. Pines and F. Bohorquez, "Challenges facing future micro-air-vehicle development," *Journal of aircraft*, vol. 43, no. 2, pp. 290–305, 2006.

[3] A. P. Willmott and C. P. Ellington, "The mechanics of flight in the hawkmoth Manduca sexta. I. Kinematics of hovering and forward flight." *Journal of Experimental Biology*, vol. 200, no. 21, pp. 2705–2722, 1997.

[4] Y. Lin, Y. Xu, H. Chen, M. J. Bender, A. L. Abbott, and R. Müller, "Optimal Threshold and LoG Based Feature Identification and Tracking of Bat Flapping Flight," in *Applications of Computer Vision (WACV), 2017 IEEE Winter Conference on*. IEEE, 2017, pp. 418–426.

[5] S. M. Nogar, A. Gogulapati, J. J. McNamara, A. Serrani, M. W. Oppenheimer, and D. B. Doman, "Control-Oriented Modeling of Coupled Electromechanical-Aeroelastic Dynamics for Flapping-Wing Vehicles," *Journal of Guidance, Control, and Dynamics*, 2017.

[6] G. Taylor and A. Thomas, "Animal flight dynamics ii. longitudinal stability in flapping flight," *Journal of theoretical biology*, vol. 214, no. 3, pp. 351–370, 2002.

[7] Q.-V. Nguyen, W. L. Chan, and M. Debiasi, "Design, fabrication, and performance test of a hovering-based flapping-wing micro air vehicle capable of sustained and controlled flight," 2014.

[8] C.-p. Du, J.-x. Xu, and Y. Zheng, "Application of iterative learning tuning to a dragonfly-like flapping wing micro aerial vehicle," in *Control and Decision Conference (CCDC), 2016 Chinese*. IEEE, 2016, pp. 4215–4220.

[9] J. M. Kok and J. Chahl, "Design and manufacture of a self-learning flapping wing-actuation system for a Dragonfly-inspired MAV," in *54th AIAA Aerospace Sciences Meeting*, 2016, p. 1744.

[10] M. M. Ferdaus, S. G. Anavatti, M. A. Garratt, and M. Pratama, "Fuzzy Clustering based Nonlinear System Identification and Controller Development of Pixhawk based Quadcopter," in *Advanced Computational Intelligence (ICACI), 2017 IEEE International Conference on*. IEEE, 2017, pp. 223–230.

[11] M. M. Ferdaus, S. G. Anavatti, M. Pratama, and M. A. Garratt, "Online Identification of a Rotary Wing Unmanned Aerial Vehicle from Data Streams," 2017.

[12] Q. Guo, M. Hu, R. Wei, J. Xu, and H. Song, "Hovering control based on fuzzy neural networks for biomimetic flying robotic," in *Information and Automation, 2008. ICIA 2008. International Conference on*. IEEE, 2008, pp. 504–508.

[13] W. He, Z. Yan, and C. Sun, "Trajectory tracking control of a flapping wing micro aerial vehicle via neural networks," in *Advanced Robotics and Mechatronics (ICARM), International Conference on*. IEEE, 2016, pp. 443–448.

[14] M. S. Couceiro, N. M. Ferreira, and J. T. Machado, "Hybrid adaptive control of a dragonfly model," *Communications in Nonlinear Science and Numerical Simulation*, vol. 17, no. 2, pp. 893–903, 2012.

[15] J.-B. M. S. D. Laurent and M. T. Druot, "Evolution of a neural network for the control of a flapping-wing animat," 2005.

[16] T. S. Clawson, S. Ferrari, S. B. Fuller, and R. J. Wood, "Spiking neural network (SNN) control of a flapping insect-scale robot," in *Decision and Control (CDC), 2016 IEEE 55th Conference on*. IEEE, 2016, pp. 3381–3388.

[17] E. D. Lughofer, "Flexfis: A robust incremental learning approach for evolving takagi–sugeno fuzzy models," *IEEE Transactions on fuzzy systems*, vol. 16, no. 6, pp. 1393–1410, 2008.

[18] P. P. Angelov and D. P. Filev, "An approach to online identification of takagi-sugeno fuzzy models," *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, vol. 34, no. 1, pp. 484–498, 2004.

[19] P. Angelov, "Evolving Takagi-Sugeno Fuzzy Systems from Data Streams (eTS+)," *Evolving Intelligent Systems: Methodology and Applications*, pp. 21–50.

[20] P. Angelov and D. Filev, "Simpl_ets: a simplified method for learning evolving takagi-sugeno fuzzy models," in *Fuzzy Systems, 2005. FUZZ'05. The 14th IEEE International Conference on*. IEEE, 2005, pp. 1068–1073.

[21] P. Angelov and R. Yager, "A new type of simplified fuzzy rule-based system," *International Journal of General Systems*, vol. 41, no. 2, pp. 163–185, 2012.

[22] P. Angelov, E. Lughofer, and X. Zhou, "Evolving fuzzy classifiers using different model architectures," *Fuzzy Sets and Systems*, vol. 159, no. 23, pp. 3160–3182, 2008.

[23] P. Angelov, "Fuzzily connected multimodel systems evolving autonomously from data streams," *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, vol. 41, no. 4, pp. 898–910, 2011.

[24] M. Pratama, S. G. Anavatti, and E. Lughofer, "Genefis: toward an effective localist network," *IEEE Transactions on Fuzzy Systems*, vol. 22, no. 3, pp. 547–562, 2014.

[25] M. Pratama, S. G. Anavatti, M. Garratt, and E. Lughofer, "Online identification of complex multi-input-multi-output system based on generic evolving neuro-fuzzy inference system," in *Evolving and Adaptive Intelligent Systems (EAIS), 2013 IEEE Conference on*. IEEE, 2013, pp. 106–113.

[26] J. Kok and J. Chahl, "A low-cost simulation platform for flapping wing MAVs," in *SPIE Smart Structures and Materials+ Nondestructive Evaluation and Health Monitoring*. International Society for Optics and Photonics, 2015, pp. 94 290L–94 290L.

[27] G.-B. Huang, P. Saratchandran, and N. Sundararajan, "An efficient sequential learning algorithm for growing and pruning rbf (gap-rbf) networks," *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, vol. 34, no. 6, pp. 2284–2292, 2004.

[28] K. S. Yap, C. P. Lim, and I. Z. Abidin, "A Hybrid ART-GRNN Online Learning Neural Network With ϵ Insensitive Loss Function," *IEEE transactions on neural networks*, vol. 19, no. 9, pp. 1641–1646, 2008.

[29] R. J. Oentaryo, M. J. Er, L. San, L. Zhai, and X. Li, "Bayesian art-based fuzzy inference system: A new approach to prognosis of machining processes," in *Prognostics and Health Management (PHM), 2011 IEEE Conference on*. IEEE, 2011, pp. 1–10.

[30] E. Lughofer, "Human-inspired evolving machinesthe next generation of evolving intelligent systems," *IEEE SMC newsletter*, vol. 36, 2011.

[31] E. Lughofer, J.-L. Bouchot, and A. Shaker, "On-line elimination of local redundancies in evolving fuzzy systems," *Evolving Systems*, vol. 2, no. 3, pp. 165–187, 2011.

[32] R. R. Yager, "A model of participatory learning," *IEEE transactions on systems, man, and cybernetics*, vol. 20, no. 5, pp. 1229–1234, 1990.

[33] D. J. MacKay, "A practical Bayesian framework for backpropagation networks," *Neural computation*, vol. 4, no. 3, pp. 448–472, 1992.