Uncertainty Bounds and Their Use in the Design of Interval Type-2 Fuzzy Logic Systems

Hongwei Wu and Jerry M. Mendel, *Fellow, IEEE*

***Abstract*—In this paper, we derive inner- and outer-bound sets for the type-reduced set of an interval type-2 fuzzy logic system (FLS), based on a new mathematical interpretation of the Karnik–Mendel iterative procedure for computing the type-reduced set. The bound sets can not only provide estimates about the uncertainty contained in the output of an interval type-2 FLS, but can also be used to design an interval type-2 FLS. We demonstrate, by means of a simulation experiment, that the resulting system can operate without type-reduction and can achieve similar performance to one that uses type-reduction. Therefore, our new design method, based on the bound sets, can relieve the computation burden of an interval type-2 FLS during its operation, which makes an interval type-2 FLS useful for real-time applications.**

***Index Terms*—Interval type-2 fuzzy logic system (FLS), time-series forecasting, type reduction, uncertainty bound.**

I. Introduction

The knowledge used to construct a fuzzy logic system (FLS) is often uncertain. The uncertainties may arise from the following sources: 1) the words used in the antecedents and the consequents of rules can mean different things to different people, 2) consequents obtained by polling a group of experts may differ, 3) the training data are noisy, and 4) the measurements that activate the FLS are noisy [5], [9], [10]. It has been demonstrated that type-2 FLSs are capable of dealing with all such uncertainties [4]–[8].

The most appropriate situations for applying type-2 FLSs are summarized in [10] as follows:

- "Measurement noise is nonstationary, but the nature of the nonstationarity cannot be expressed mathematically ahead of time," e.g., time-series forecasting under variable SNR measurements [5].
- "A data-generating mechanism is time-varying, but the nature of the time variations cannot be expressed mathematically ahead of time," e.g., equalization and co-channel interference reduction for nonlinear and time-varying digital communication channels [6], [7].
- "Feature are described by statistical attributes that are nonstationary, but the nature of the nonstationarity cannot be expressed mathematically ahead of time," e.g., rule-based classification of video traffic [8].
- "Knowledge is mined from experts using IF-THEN questionnaires," e.g., connection admission control for ATM networks [4].

The four kinds of uncertainties mentioned above flow through a type-2 FLS and produce uncertainties at its output. For an interval type-2 FLS (the only kind that is practical to date), the output is uncertain within an interval which is obtained through some kind of *type-reduction* method [2], [3], [10].

Type-reduction is an extension of type-1 defuzzification, obtained by applying the Extension Principle [12] to a specific defuzzification method. It represents a mapping of a type-2 fuzzy set into a type-1 fuzzy set. There exist many kinds of type-reduction methods (e.g., centroid, center-of-sets, center-of-sums, and height type-reduction); but, for an interval type-2 FLS, regardless of the type-reduction method and how its input $\mathbf{x}$ is modeled (e.g., as a singleton, type-1 fuzzy set, or type-2 fuzzy set), the type-reduced set is always an interval set and is determined by its two end points $y_l(\mathbf{x})$ and $y_r(\mathbf{x})$.

In information theory, the uncertainty of a random variable is measured by its entropy [1]. Recall that a one-dimensional random variable that is uniformly distributed over a region has entropy equal to the logarithm of the length of the region. Comparing the membership function (MF), $\mu_Y(y)$, of an interval fuzzy set Y, where

$$\mu_Y(y) = \begin{cases} 1, & \text{y} \in [y_l, y_r] \\ 0, & \text{otherwise} \end{cases} \tag{1}$$

with the probability density function $p_{Y'}(y)$ of a random variable Y', which is uniformly distributed over $[y_l, y_r]$, where

$$p_{Y'}(y) = \begin{cases} \frac{1}{(y_r - y_l)}, & y \in [y_l, y_r] \\ 0, & \text{otherwise} \end{cases} \tag{2}$$

we find that they are almost the same except for their amplitudes. Therefore, it is reasonable to consider the *extent of the uncertainty* of the fuzzy set Y to be the same as (or proportional to) that of the random variable Y'. Since the output of an interval type-2 FLS is uncertain within the type-reduced set, which is an interval type-1 fuzzy set, the *length* of the type-reduced set can therefore be used to measure the extent of the output's uncertainty.

In an interval type-2 FLS, the result of the input and antecedent operations is the firing set $F^i(\mathbf{x})$, which is an interval type-1 fuzzy set, i.e.,

$$F^i(\mathbf{x}) = \left[\underline{f}^i(\mathbf{x}), \bar{f}^i(\mathbf{x})\right] \equiv \left[\underline{f}^i, \bar{f}^i\right], \qquad i = 1, 2, \ldots, M \tag{3}$$

where $\underline{f}^i(\mathbf{x})$ and $\bar{f}^i(\mathbf{x})$ represent the lower and upper firing degrees of the ith rule (formulas for which are given in Section II)

Manuscript received July 11, 2001; revised January 11, 2002 and February 26, 2002.

The authors are with the Signal and Image Processing Institute, Department of Electrical Engineering, University of Southern California, Los Angeles, CA 90089-2564 USA (e-mail: mendel@sipi.usc.edu).

Digital Object Identifier 10.1109/TFUZZ.2002.803496.

and M is the number of rules. When, for example, center-of-sets type-reduction (described in Section II) is used, $y_l(\mathbf{x})$ and $y_r(\mathbf{x})$ can be represented as

$$\begin{aligned} y_l(\mathbf{x}) =& y_l\Big[\bar{f}^1(\mathbf{x}),\dots,\bar{f}^{L^\star(\mathbf{x})}(\mathbf{x}) \\ & \underline{f}^{L^\star(\mathbf{x})+1}(\mathbf{x}),\dots,\underline{f}^{M}(\mathbf{x}), y_l^1,\dots,y_l^M\Big] \quad (4) \\ y_r(\mathbf{x}) =& y_r\Big[\underline{f}^1(\mathbf{x}),\dots,\underline{f}^{R^\star(\mathbf{x})}(\mathbf{x}) \\ & \bar{f}^{R^\star(\mathbf{x})+1}(\mathbf{x}),\dots,\bar{f}^{M}(\mathbf{x}), y_r^1,\dots,y_r^M\Big]. \quad (5) \end{aligned}$$

where y_l^i and y_r^i ($i = 1,\dots,M$) are the end points of the centroid of the consequent type-2 fuzzy sets and $L^\star(\mathbf{x})$ and $R^\star(\mathbf{x})$ are very important switching numbers which depend on the input $\mathbf{x}$ (how to compute them is also described in Section II). Only after $L^\star(\mathbf{x})$ and $R^\star(\mathbf{x})$ are determined are the end points of the type-reduced set determined. Unfortunately, although $L^\star(\mathbf{x})$ and $R^\star(\mathbf{x})$ are related to the input data and the MF parameters of an interval type-2 FLS, they cannot be predetermined as explicit functions of these quantities. To compute $L^\star(\mathbf{x})$ and $R^\star(\mathbf{x})$ we need to implement two iterative procedures, developed by Karnik and Mendel [2], [10], for each given value of $\mathbf{x}$, one for $L^\star(\mathbf{x})$ and a similar one for $R^\star(\mathbf{x})$. The computation of $L^\star(\mathbf{x})$ and $R^\star(\mathbf{x})$ represents a *bottleneck* for interval type-2 FLSs. The main result of this paper is a method for eliminating this bottleneck so that type-2 FLSs are then feasible for real-time applications.

Karnik, Mendel, and Liang [3] have observed that an interval type-2 FLS can be interpreted as a collection of embedded type-1 FLSs (Appendix I provides some background materials about type-2 fuzzy sets, including a definition of embedded type-1 fuzzy sets). We have found that embedded type-1 FLSs play very important roles in understanding uncertainty in a type-2 FLS. Two of them let us compute $y_l(\mathbf{x})$ and $y_r(\mathbf{x})$, whereas some of the others let us compute the *inner-bound set* $[\bar{y}_l(\mathbf{x}), \underline{y}_r(\mathbf{x})]$ and the *outer-bound set* $[\underline{y}_l(\mathbf{x}), \bar{y}_r(\mathbf{x})]$ for the type-reduced set. Fig. 1 shows the type-reduced set and its inner- and outer-bound sets, where $y_l(\mathbf{x}) \in [\underline{y}_l(\mathbf{x}), \bar{y}_l(\mathbf{x})]$ and $y_r(\mathbf{x}) \in [\underline{y}_r(\mathbf{x}), \bar{y}_r(\mathbf{x})]$. In this paper, we show how to compute the inner-bound and the outer-bound sets, explain why they are useful and important and demonstrate that they can be computed without the computation of $L^\star(\mathbf{x})$ or $R^\star(\mathbf{x})$.

In Section II, we provide a new mathematical interpretation to the procedure for computing a type-reduced set. In Section III, we derive the inner- and outer-bound sets for the type-reduced set. In Section IV, we propose a design method for an interval type-2 FLS based on the inner- and outer-bound sets. In Section V, we apply the new design method to the problem of predicting the Mackey–Glass time series. Finally, in Section VI, we draw conclusions.

II. Type-Reduced Fuzzy Set for an Interval Type-2 FLS

A type-reduced fuzzy set for an interval type-2 FLS is a generalized centroid, which can be expressed as [2], [10]

$$Y_{TR}(\mathbf{x}) = [y_l(\mathbf{x}), y_r(\mathbf{x})]$$

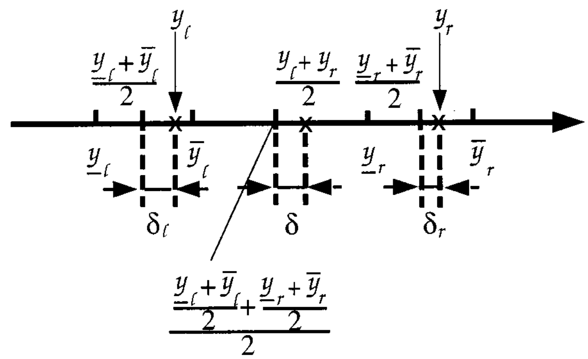


Fig. 1. The type-reduced set $[y_l(\mathbf{x}), y_r(\mathbf{x})]$, its inner- and outer-bound sets $[\bar{y}_l(\mathbf{x}), \underline{y}_r(\mathbf{x})]$ and $[\underline{y}_l(\mathbf{x}), \bar{y}_r(\mathbf{x})]$, and the defuzzified output of the type-reduced set and of its approximation. δ_l is the difference between y_l and the average of $\underline{y}_l$ and $\bar{y}_l$. δ_r is the difference between y_r and the average of $\bar{y}_r$ and $\underline{y}_r$. δ is the difference between the two defuzzified values.

$$= \int_{y^i\in[y_l^i, y_r^i]} \int_{f^i\in[\underline{f}^i, \bar{f}^i]} 1 \Bigg/ \frac{\sum_{i=1}^{M} f^i y^i}{\sum_{i=1}^{M} f^i} \quad (6)$$

where $Y_{TR}(\mathbf{x})$ is an interval type-1 fuzzy set determined by its two end points $y_l(\mathbf{x})$ and $y_r(\mathbf{x})$ and M is the number of rules. For center-of-sets type-reduction, $[\underline{f}^i, \bar{f}^i]$ is the firing interval and $[y_l^i, y_r^i]$ is the centroid of the consequent set of the ith rule. The meanings of $[\underline{f}^i, \bar{f}^i]$ and $[y_l^i, y_r^i]$ in other type-reduction methods are explained in Appendix II.

The firing interval $[\underline{f}^i, \bar{f}^i]$ is determined by [5], [10]

$$\underline{f}^i \equiv \underline{f}^i(\mathbf{x}) = T_{k=1}^{p} \underline{f}_k^i(x_k) \quad (7)$$

and

$$\bar{f}^i \equiv \bar{f}^i(\mathbf{x}) = T_{k=1}^{p} \bar{f}_k^i(x_k) \quad (8)$$

where

$$\underline{f}_k^i(x_k) = \sup_{x_k} \int_{x_k\in X_k} \left[\underline{\mu}_{\widetilde{X}_k}(x_k) \star \underline{\mu}_{\widetilde{F}_k^i}(x_k)\right]/x_k \quad (9)$$

$$\bar{f}_k^i(x_k) = \sup_{x_k} \int_{x_k\in X_k} \left[\bar{\mu}_{\widetilde{X}_k}(x_k) \star \bar{\mu}_{\widetilde{F}_k^i}(x_k)\right]/x_k \quad (10)$$

In these equations, the input $\mathbf{x}$ is a p-dimensional vector, i.e., $\mathbf{x} = [x_1, x_2, \dots, x_p]^T$, $\widetilde{X}_k$ is the type-2 (which includes type-1 and type-0 as special cases) fuzzy model for the kth input, $\widetilde{F}_k^i$ is the type-2 (which includes type-1 as a special case) antecedent set of the ith rule for the kth input, $\underline{\mu}(\cdot)$ and $\bar{\mu}(\cdot)$ are lower and upper membership functions (LMFs, UMFs) and, T and $\star$ represent t-norm operations.

The end points of the type-reduced set, $y_l(\mathbf{x})$ and $y_r(\mathbf{x})$, can be computed using an iterative method, developed by Karnik and Mendel [2], [10], which we reinterpret for the purposes of this paper in the following:

Theorem 1: Assume y_l^i and y_r^i ($i = 1,\dots,M$) are reordered (as required in [2] and [10]) such that

$$y_l^1 \le y_l^2 \le \cdots \le y_l^M \quad (11)$$

$$y_r^1 \le y_r^2 \le \cdots \le y_r^M \quad (12)$$

and define $y_l^{(L)}(\mathbf{x})$ and $y_r^{(R)}(\mathbf{x})$, for $0 \le L, R \le M$, as

$$y_l^{(L)}(\mathbf{x}) \equiv \frac{\sum_{i=1}^{L} \bar{f}^i y_l^i + \sum_{i=L+1}^{M} \underline{f}^i y_l^i}{\sum_{i=1}^{L} \bar{f}^i + \sum_{i=L+1}^{M} \underline{f}^i} \tag{13}$$

$$y_r^{(R)}(\mathbf{x}) \equiv \frac{\sum_{i=1}^{R} \underline{f}^i y_r^i + \sum_{i=R+1}^{M} \bar{f}^i y_r^i}{\sum_{i=1}^{R} \underline{f}^i + \sum_{i=R+1}^{M} \bar{f}^i}. \tag{14}$$

The end points $y_l(\mathbf{x})$ and $y_r(\mathbf{x})$ of the type-reduced fuzzy set of an interval type-2 FLS, given by (6), are the minimum of all $y_l^{(L)}(\mathbf{x})$ and the maximum of all $y_r^{(R)}(\mathbf{x})$, respectively, i.e.,

$$\begin{aligned} y_l(\mathbf{x}) &= \min_{0\le L\le M} \left\{ y_l^{(L)}(\mathbf{x}) \right\} = y_l^{(L^\star(\mathbf{x}))}(\mathbf{x}) \\ &= \frac{\sum_{i=1}^{L^\star(\mathbf{x})} \bar{f}^i y_l^i + \sum_{i=L^\star(\mathbf{x})+1}^{M} \underline{f}^i y_l^i}{\sum_{i=1}^{L^\star(\mathbf{x})} \bar{f}^i + \sum_{i=L^\star(\mathbf{x})+1}^{M} \underline{f}^i} \end{aligned} \tag{15}$$

where

$$L^\star(\mathbf{x}) = \arg \min_{0\le L\le M} \left\{ y_l^{(L)}(\mathbf{x}) \right\} \tag{16}$$

and

$$\begin{aligned} y_r(\mathbf{x}) &= \max_{0\le R\le M} \left\{ y_r^{(R)}(\mathbf{x}) \right\} = y_r^{(R^\star(\mathbf{x}))}(\mathbf{x}) \\ &= \frac{\sum_{i=1}^{R^\star(\mathbf{x})} \underline{f}^i y_r^i + \sum_{i=R^\star(\mathbf{x})+1}^{M} \bar{f}^i y_r^i}{\sum_{i=1}^{R^\star(\mathbf{x})} \underline{f}^i + \sum_{i=R^\star(\mathbf{x})+1}^{M} \bar{f}^i} \end{aligned} \tag{17}$$

where

$$R^\star(\mathbf{x}) = \arg \max_{0\le R\le M} \left\{ y_r^{(R)}(\mathbf{x}) \right\}. \tag{18}$$

The solutions of (16) and (18), $L^\star(\mathbf{x})$ and $R^\star(\mathbf{x})$, are obtained using the Karnik–Mendel iterative procedure [2], [10].

Equation (11) requires a reordering of the M rules for the calculation of $y_l^{(L)}(\mathbf{x})$ and (12) requires another reordering of the M rules for the calculation of $y_r^{(R)}(\mathbf{x})$. In general, these two reorderings are different. In (15) and (17), the $\underline{f}^i$ and $\bar{f}^i$ are associated with the respective reordered rules for $y_l(\mathbf{x})$ and $y_r(\mathbf{x})$.

The Karnik–Mendel iterative procedure for determining $L^\star(\mathbf{x})$ and $R^\star(\mathbf{x})$ is easy to implement and, is given in Appendix III. Karnik and Mendel have shown that *at most* M iterations (M is the number of rules) are needed to determine $L^\star(\mathbf{x})$ and M iterations are needed to determine $R^\star(\mathbf{x})$. In [11], we have shown that *on average* $(M+2)/4$ iterations are needed to determine $L^\star(\mathbf{x})$ and $(M+2)/4$ iterations are needed to determine $R^\star(\mathbf{x})$. Although $L^\star(\mathbf{x})$ and $R^\star(\mathbf{x})$ can be computed in parallel, we see that type-reduction represents a major bottleneck to the use of an interval type-2 FLS in real-time applications, especially when the rule base (i.e., M) of the FLS and the number of input data are large.

III. Inner- and Outer-Bound Sets for a Type-Reduced Set

A type-reduced set is not only associated with the uncertainty of the output of an interval type-2 FLS, but is also crucial to defuzzification. Unfortunately, the time-consuming Karnik–Mendel iterative procedure must be used to obtain the type-reduced set. In this section we provide inner- and outer-bound sets for the type-reduced set, both of which can be calculated without type-reduction. These two sets can not only be used to estimate the uncertainty contained in the output of an interval type-2 FLS, but can also be used to directly derive the defuzzified output under certain conditions. Consequently the inner- and outer-bound sets have the potential to eliminate the computional bottleneck of an interval type-2 FLS.

An interval type-2 FLS can be interpreted as a collection of embedded type-1 FLSs [5], [10]. The following embedded type-1 FLSs only use the LMFs (or UMFs) of the input and antecedent fuzzy sets, together with the left (or right) end points of the centroids of the consequents:

$$\{\text{LMFs, left}\}: y_l^{(0)}(\mathbf{x}) = \frac{\underline{f}^1 y_l^1 + \cdots + \underline{f}^M y_l^M}{\underline{f}^1 + \cdots + \underline{f}^M} \tag{19}$$

$$\{\text{UMFs, left}\}: y_l^{(M)}(\mathbf{x}) = \frac{\bar{f}^1 y_l^1 + \cdots + \bar{f}^M y_l^M}{\bar{f}^1 + \cdots + \bar{f}^M} \tag{20}$$

$$\{\text{LMFs, right}\}: y_r^{(M)}(\mathbf{x}) = \frac{\underline{f}^1 y_r^1 + \cdots + \underline{f}^M y_r^M}{\underline{f}^1 + \cdots + \underline{f}^M} \tag{21}$$

$$\{\text{UMFs, right}\}: y_l^{(0)}(\mathbf{x}) = \frac{\bar{f}^1 y_r^1 + \cdots + \bar{f}^M y_r^M}{\bar{f}^1 + \cdots + \bar{f}^M}. \tag{22}$$

We refer to them as **boundary type-1 FLSs** for an interval type-2 FLS. We have found that boundary type-1 FLSs are very important in deriving the inner- and outer-bound sets of a type-reduced set.

Theorem 2: The end points $y_l(\mathbf{x})$ and $y_r(\mathbf{x})$ of the type-reduced set of an interval type-2 FLS for the input $\mathbf{x}$, are bounded from below and above by (Fig. 1)

$$\underline{y}_l(\mathbf{x}) \le y_l(\mathbf{x}) \le \bar{y}_l(\mathbf{x}) \tag{23}$$

$$\underline{y}_r(\mathbf{x}) \le y_r(\mathbf{x}) \le \bar{y}_r(\mathbf{x}) \tag{24}$$

where

$$\bar{y}_l(\mathbf{x}) = \min \left\{ y_l^{(0)}(\mathbf{x}), y_l^{(M)}(\mathbf{x}) \right\} \tag{25}$$

$$\underline{y}_r(\mathbf{x}) = \max \left\{ y_r^{(0)}(\mathbf{x}), y_r^{(M)}(\mathbf{x}) \right\} \tag{26}$$

and

$$\begin{aligned} \underline{y}_l(\mathbf{x}) = \bar{y}_l(\mathbf{x}) - &\left[\frac{\sum_{i=1}^{M} \left(\bar{f}^i - \underline{f}^i\right)}{\sum_{i=1}^{M} \bar{f}^i \sum_{i=1}^{M} \underline{f}^i} \right. \\ &\left. \times \frac{\sum_{i=1}^{M} \underline{f}^i \left(y_l^i - y_l^1\right) \sum_{i=1}^{M} \bar{f}^i \left(y_l^M - y_l^i\right)}{\sum_{i=1}^{M} \underline{f}^i \left(y_l^i - y_l^1\right) + \sum_{i=1}^{M} \bar{f}^i \left(y_l^M - y_l^i\right)} \right] \end{aligned} \tag{27}$$

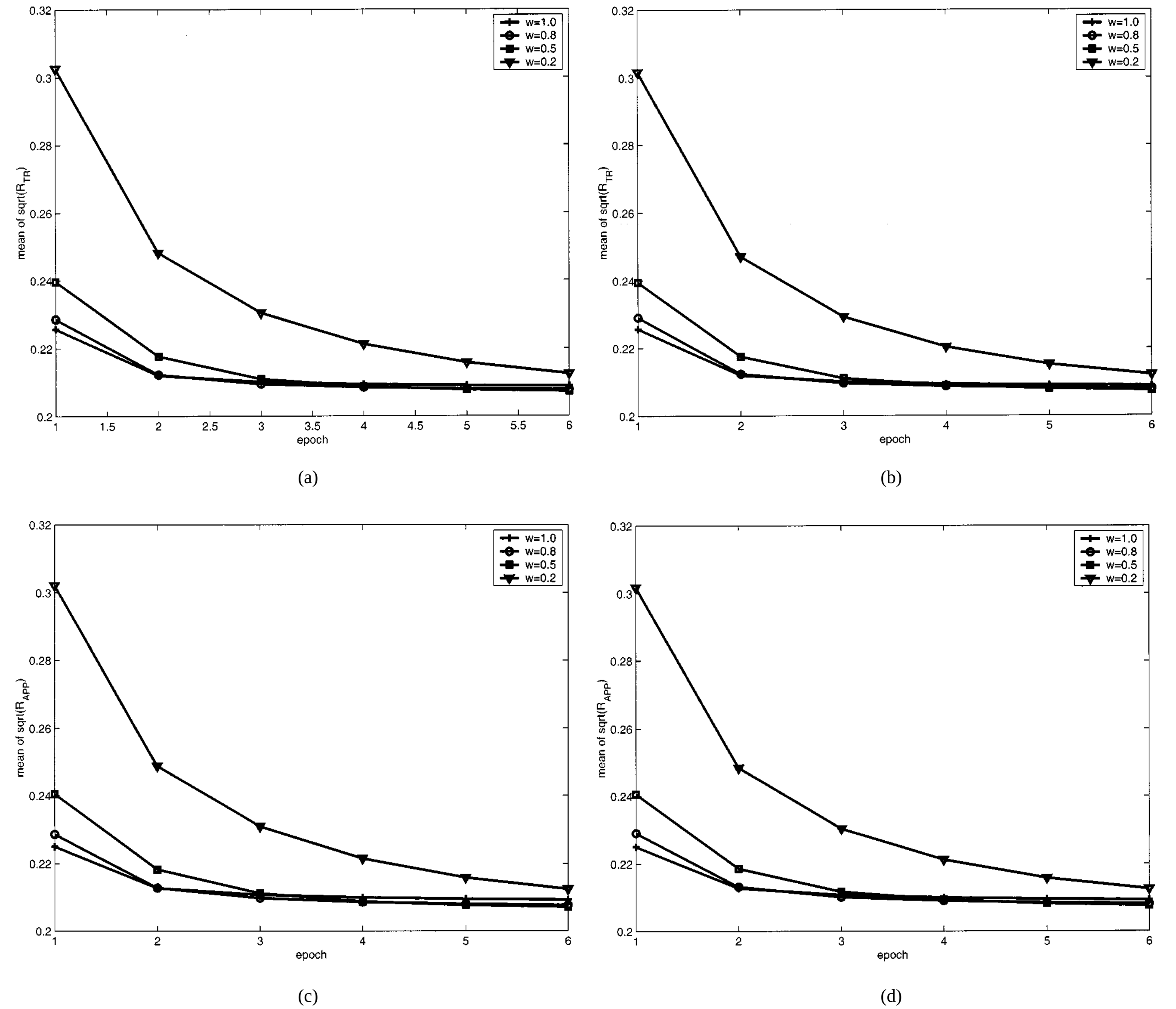


Fig. 2. The mean values of $\sqrt{R_{TR}}$ and $\sqrt{R_{APP}}$ for the testing data, averaged over 50 Monte Carlo iterations (obtained from Tables I and II). (a) Means of $\sqrt{R_{TR}}$ for the designs based on $R_{TR-\bar{\Delta}}$. (b) Means of $\sqrt{R_{TR}}$ for the designs based on $R_{TR-\Delta}$. (c) Means of $\sqrt{R_{APP}}$ for the designs based on $R_{TR-\bar{\Delta}}$. (d) Means of $\sqrt{R_{APP}}$ for the designs based on $R_{TR-\Delta}$.

$$\bar{y}_r(\mathbf{x}) = \underline{y}_r(\mathbf{x}) + \left[\frac{\sum_{i=1}^{M} \left(\bar{f}^i - \underline{f}^i\right)}{\sum_{i=1}^{M} \bar{f}^i \sum_{i=1}^{M} \underline{f}^i} \times \frac{\sum_{i=1}^{M} \bar{f}^i \left(y_r^i - y_r^1\right) \sum_{i=1}^{M} \underline{f}^i \left(y_r^M - y_r^i\right)}{\sum_{i=1}^{M} \bar{f}^i \left(y_r^i - y_r^1\right) + \sum_{i=1}^{M} \underline{f}^i \left(y_r^M - y_r^i\right)} \right]. \tag{28}$$

A proof of Theorem 2 is given in Appendix IV.

We refer to $[\bar{y}_l(\mathbf{x}), \underline{y}_r(\mathbf{x})]$ and $[\underline{y}_l(\mathbf{x}), \bar{y}_r(\mathbf{x})]$ as the inner- and outer-bound sets for the type-reduced set $[y_l(\mathbf{x}), y_r(\mathbf{x})]$ of an interval type-2 FLS. From (27) and (28), we see that the lengths of the intervals $|\bar{y}_l(\mathbf{x}) - \underline{y}_l(\mathbf{x})|$ and $|\bar{y}_r(\mathbf{x}) - \underline{y}_r(\mathbf{x})|$ are determined by how different the lower and upper firing degrees are and how the consequents are distributed. When $\sum_{i=1}^{M}(\bar{f}^i - \underline{f}^i)$ is small (i.e., the uncertainties contained in the firing intervals are small) and/or the difference among y_l^i and the difference among y_r^i $(i = 1, \ldots, M)$ are small[1] (i.e., the consequents are distributed close to each other), then $|\bar{y}_l(\mathbf{x}) - \underline{y}_l(\mathbf{x})|$ and $|\bar{y}_r(\mathbf{x}) - \underline{y}_r(\mathbf{x})|$ are small and consequently the differences of $|y_l(\mathbf{x}) - \underline{y}_l(\mathbf{x})|$, $|y_l(\mathbf{x}) - \bar{y}_l(\mathbf{x})|$, $|y_r(\mathbf{x}) - \underline{y}_r(\mathbf{x})|$ and $|y_r(\mathbf{x}) - \bar{y}_r(\mathbf{x})|$ are small. These observations are consistent with our intuition.

Theorem 2 is true for all type-reduction methods; however, because $[\underline{f}^i, \bar{f}^i]$ and $[y_l^i, y_r^i]$ have different meanings for different type-reduction methods (Appendix II), (25)–(28) may take different values for different type-reduction methods. These values are given in Appendix II.

[1]Although (27) and (28) would then appear to have a 0/0 term, they actually have a $0 \times 0/0$ term, hence, a careful analysis of this case reveals that the second term in (27) and (28) goes to zero.

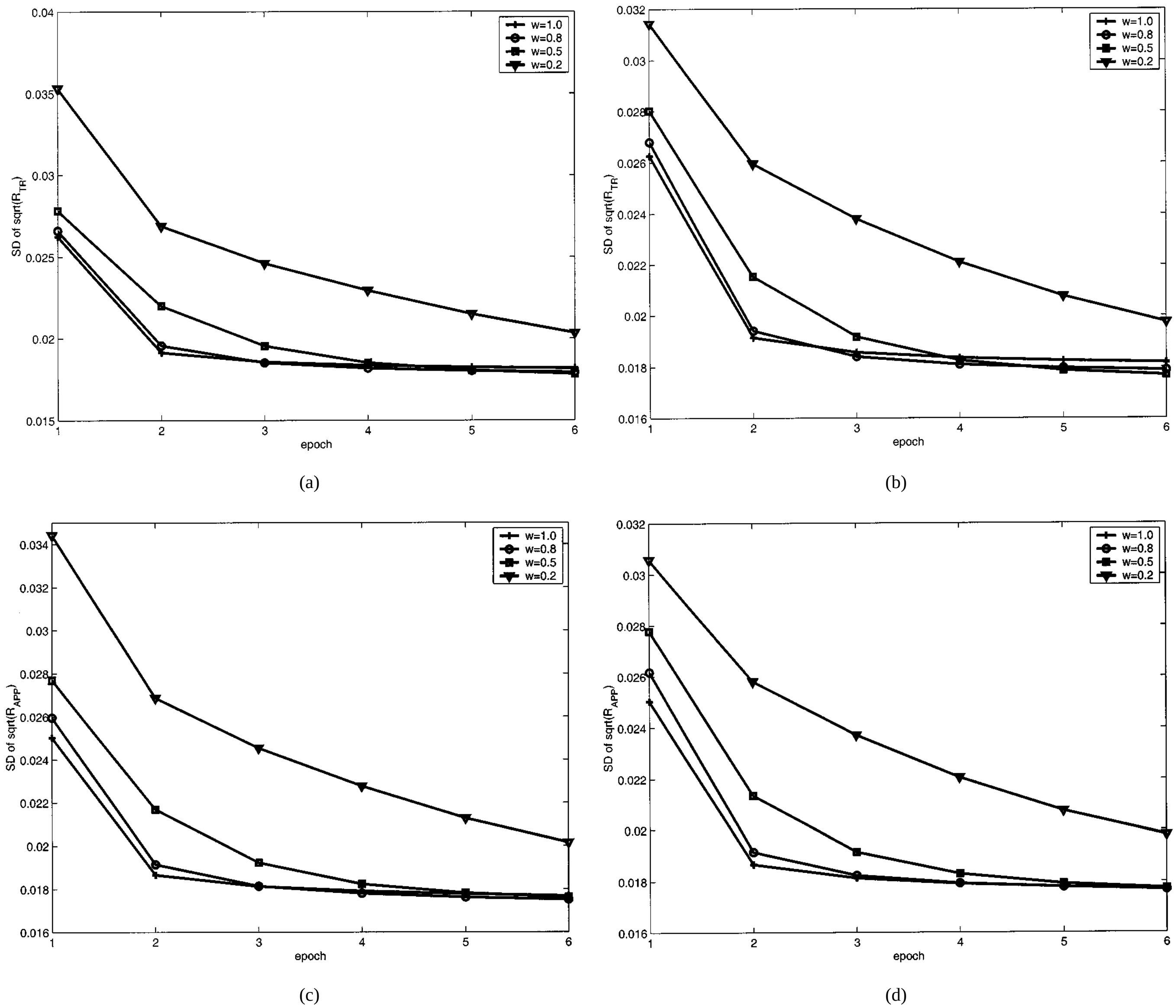


Fig. 3. The standard deviations of $\sqrt{R_{TR}}$ and $\sqrt{R_{APP}}$ for the testing data, averaged over 50 Monte Carlo iterations (obtained from Tables I and II). (a) SDs of $\sqrt{R_{TR}}$ for the designs based on $R_{TR-\bar{\Delta}}$. (b) SDs of $\sqrt{R_{TR}}$ for the designs based on $R_{TR-\Delta}$. (c) SDs of $\sqrt{R_{APP}}$ for the designs based on $R_{TR-\bar{\Delta}}$. (d) SDs of $\sqrt{R_{APP}}$ for the designs based on $R_{TR-\Delta}$.

Our main goal is to *not* perform type-reduction during the real-time operation of a type-2 FLS. We propose, therefore, to approximate the type-reduced set by its inner- and outer-bound sets, i.e., to approximate $[y_l(\mathbf{x}), y_r(\mathbf{x})]$ by $[(\underline{y}_l(\mathbf{x})+\bar{y}_l(\mathbf{x}))/2, (\underline{y}_r(\mathbf{x})+\bar{y}_r(\mathbf{x}))/2]$ and to compute the output of the FLS as $[(\underline{y}_l(\mathbf{x})+\bar{y}_l(\mathbf{x}))/2 + (\underline{y}_r(\mathbf{x})+\bar{y}_r(\mathbf{x}))/2]/2$ (Fig. 1). If this is going to be acceptable, then the difference between $[(\underline{y}_l(\mathbf{x})+\bar{y}_l(\mathbf{x}))/2 + (\underline{y}_r(\mathbf{x})+\bar{y}_r(\mathbf{x}))/2]/2$ and the usual defuzzified output, $[y_l(\mathbf{x})+y_r(\mathbf{x})]/2$, must be small.

Corollary 1: The difference, $\delta(\mathbf{x})$, between the defuzzified outputs of the type-reduced set and its approximation set for the input $\mathbf{x}$, which is defined as

$$\delta(\mathbf{x}) \equiv \left| \frac{y_l(\mathbf{x})+y_r(\mathbf{x})}{2} - \frac{1}{2}\left[\frac{\underline{y}_l(\mathbf{x})+\bar{y}_l(\mathbf{x})}{2} + \frac{\underline{y}_r(\mathbf{x})+\bar{y}_r(\mathbf{x})}{2}\right]\right| \quad (29)$$

is bounded from above as

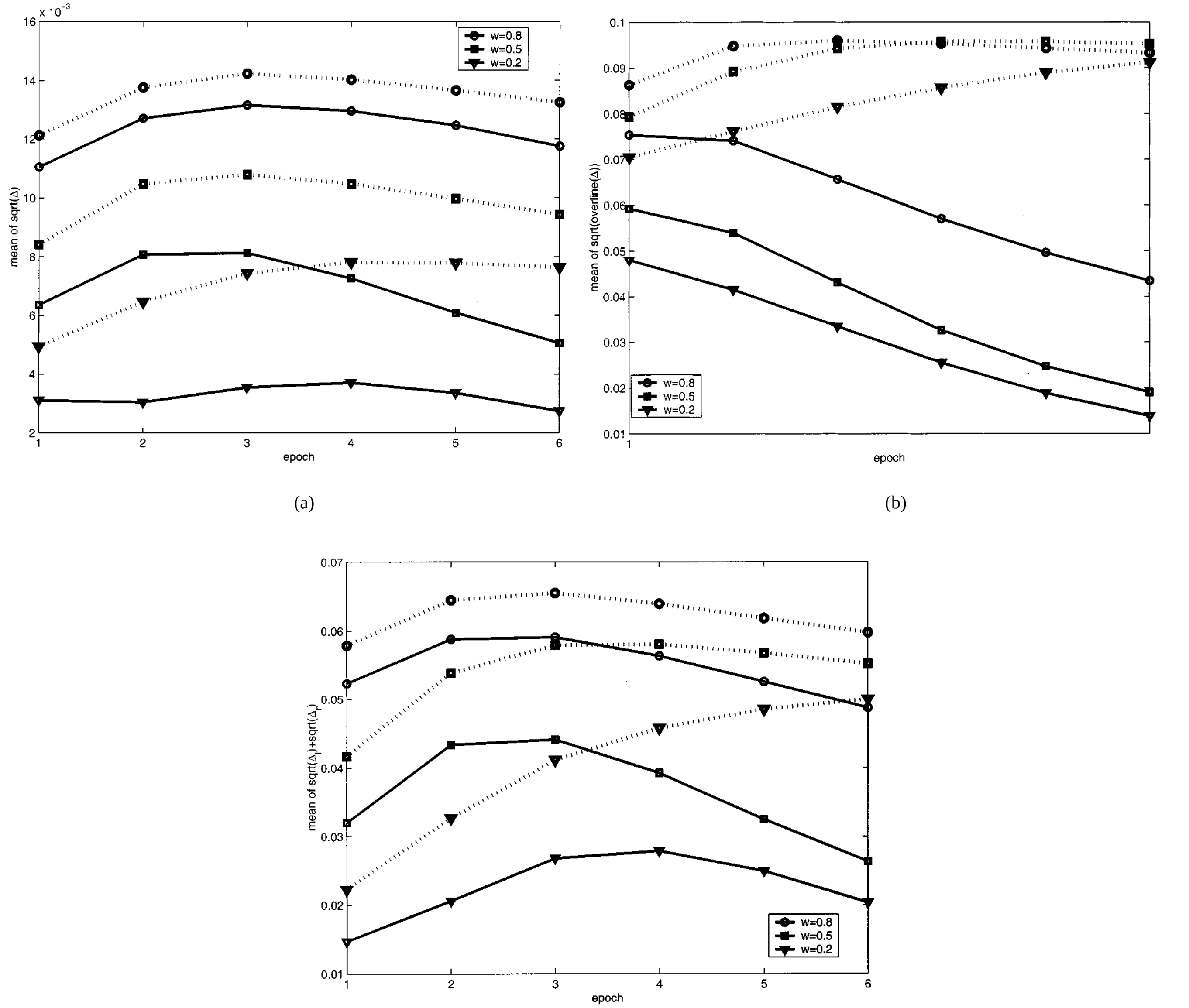
$$\delta(\mathbf{x}) \le \bar{\delta}(\mathbf{x}) \equiv \frac{1}{4}\left[\left(\bar{y}_l(\mathbf{x})-\underline{y}_l(\mathbf{x})\right)+\left(\bar{y}_r(\mathbf{x})-\underline{y}_r(\mathbf{x})\right)\right]. \quad (30)$$

A proof of Corollary 1 appears in Appendix IV.

In Section IV, we shall propose two new risk functions for the design of an interval type-2 FLS, one including $\delta(\mathbf{x})$ and another including $\bar{\delta}(\mathbf{x})$. Here, we demonstrate that using $\bar{\delta}(\mathbf{x})$ has advantages over using $\delta(\mathbf{x})$. Let (see Fig. 1)

$$\delta_l(\mathbf{x}) \equiv \left| y_l(\mathbf{x}) - \frac{\underline{y}_l(\mathbf{x})+\bar{y}_l(\mathbf{x})}{2}\right| \quad (31)$$

$$\delta_r(\mathbf{x}) \equiv \left| y_r(\mathbf{x}) - \frac{\underline{y}_r(\mathbf{x})+\bar{y}_r(\mathbf{x})}{2}\right|. \quad (32)$$



(a) (b) (c)

Fig. 4. The mean values of $\sqrt{\Delta}$, $\sqrt{\bar{\Delta}}$ and $(\sqrt{\Delta_l}+\sqrt{\Delta_r})$ for the testing data, averaged over 50 Monte Carlo iterations (obtained from Tables III and IV). (a) Means of $\sqrt{\Delta}$ for the designs based on $R_{TR-\bar{\Delta}}$ (solid lines) and for the designs based on $R_{TR-\Delta}$ (dotted lines). (b) Means of $\sqrt{\bar{\Delta}}$ for the designs based on $R_{TR-\bar{\Delta}}$ (solid lines) and for the designs based on $R_{TR-\Delta}$ (dotted lines). (c) Means of $(\sqrt{\Delta_l}+\sqrt{\Delta_r})$ for the designs based on $R_{TR-\bar{\Delta}}$ (solid lines) and for the designs based on $R_{TR-\Delta}$ (dotted lines).

Then, $\delta(\mathbf{x})$ can be rewritten as the sum or difference of $\delta_l(\mathbf{x})$ and $\delta_r(\mathbf{x})$, as follows.

- When $[y_l(\mathbf{x})-(\underline{y}_l(\mathbf{x})+\bar{y}_l(\mathbf{x}))/2]\times[y_r(\mathbf{x})-(\underline{y}_r(\mathbf{x})+\bar{y}_r(\mathbf{x}))/2]\geq 0$

$$\delta(\mathbf{x})=\frac{1}{2}\left[\delta_l(\mathbf{x})+\delta_r(\mathbf{x})\right]. \tag{33}$$

- When $[y_l(\mathbf{x})-(\underline{y}_l(\mathbf{x})+\bar{y}_l(\mathbf{x}))/2]\times[y_r(\mathbf{x})-(\underline{y}_r(\mathbf{x})+\bar{y}_r(\mathbf{x}))/2]< 0$

$$\delta(\mathbf{x})=\frac{1}{2}\left|\delta_l(\mathbf{x})-\delta_r(\mathbf{x})\right|. \tag{34}$$

Whereas, for $\bar{\delta}(\mathbf{x})$, since

$$\frac{\bar{y}_l(\mathbf{x})-\underline{y}_l(\mathbf{x})}{2}-\left[\frac{\underline{y}_l(\mathbf{x})+\bar{y}_l(\mathbf{x})}{2}-y_l(\mathbf{x})\right]=y_l(\mathbf{x})-\underline{y}_l(\mathbf{x})\geq 0 \tag{35}$$

and

$$\frac{\bar{y}_l(\mathbf{x})-\underline{y}_l(\mathbf{x})}{2}-\left[y_l(\mathbf{x})-\frac{\underline{y}_l(\mathbf{x})+\bar{y}_l(\mathbf{x})}{2}\right]=\bar{y}_l(\mathbf{x})-y_l(\mathbf{x})\geq 0 \tag{36}$$

which means $[\bar{y}_l(\mathbf{x})-\underline{y}_l(\mathbf{x})]/2\geq\delta_l(\mathbf{x})$ and $[\bar{y}_r(\mathbf{x})-\underline{y}_r(\mathbf{x})]/2\geq\delta_r(\mathbf{x})$ in a similar way, we then have

$$\bar{\delta}(\mathbf{x})\geq\frac{\delta_l(\mathbf{x})+\delta_r(\mathbf{x})}{2} \tag{37}$$

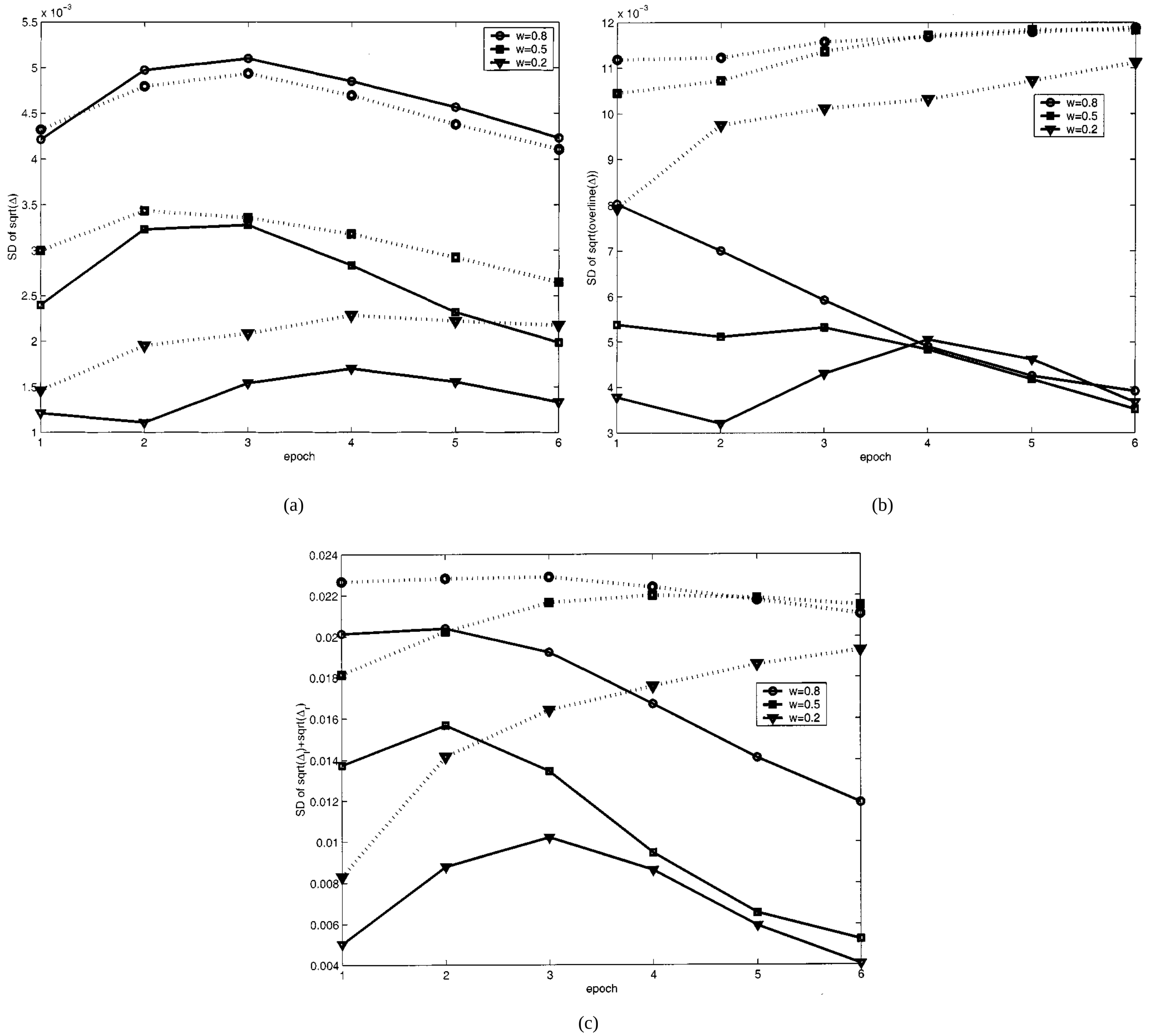


Fig. 5. The standard deviations of $\sqrt{\Delta}$, $\sqrt{\overline{\Delta}}$ and $(\sqrt{\Delta_l} + \sqrt{\Delta_r})$ for the testing data, averaged over 50 Monte Carlo iterations (obtained from Tables III and IV). (a) SDs of $\sqrt{\Delta}$ for the designs based on $R_{TR-\overline{\Delta}}$ (solid lines) and for the designs based on $R_{TR-\Delta}$ (dotted lines). (b) SDs of $\sqrt{\overline{\Delta}}$ for the designs based on $R_{TR-\overline{\Delta}}$ (solid lines) and for the designs based on $R_{TR-\Delta}$ (dotted lines). (c) SDs of $(\sqrt{\Delta_l} + \sqrt{\Delta_r})$ for the designs based on $R_{TR-\overline{\Delta}}$ (solid lines) and for the designs based on $R_{TR-\Delta}$ (dotted lines).

From (34), we see that $\delta(\mathbf{x})$ being small does not necessarily imply that both $\delta_l(\mathbf{x})$ and $\delta_r(\mathbf{x})$ are small; whereas, from (37), we see that $\overline{\delta}(\mathbf{x})$ being small is sufficient for both $\delta_l(\mathbf{x})$ and $\delta_r(\mathbf{x})$ [and, therefore, [from (33) and (34)] $\delta(\mathbf{x})$] to be small, in which case the approximation set $[(\underline{y}_l(\mathbf{x}) + \bar{y}_l(\mathbf{x}))/2, (\underline{y}_r(\mathbf{x}) + \bar{y}_r(\mathbf{x}))/2]$ is close to the type-reduced set $[y_l(\mathbf{x}), y_r(\mathbf{x})]$, *and* their defuzzified outputs are also close.

Why is this important? When using an interval type-2 FLS, we must be concerned about its uncertainty range (i.e., the type-reduced set) as well as its defuzzified output. Therefore, it is important to make both the approximation set and its defuzzified output approach the type-reduced set and its defuzzified output, respectively. From this point of view, using $\overline{\delta}(\mathbf{x})$ is preferred to using $\delta(\mathbf{x})$ during the design.

IV. Application of the Inner- and Outer-Bound Sets to Designing an Interval Type-2 FLS

The major advantage of the inner- and outer-bound sets is they can be calculated without having to use the Karnik–Mendel iterative procedure. If the type-reduced set could be approximated by its inner- and outer-bound sets, then type-reduction could be eliminated and an interval type-2 FLS could lend itself to real-time applications.

Theorem 3: For a group of input–output data $\{\mathbf{x}_i, y_i\}_{i=1}^N$ and an interval type-2 FLS, let the **risk function** (i.e., the sample mean of the squared error), R_{TR}, associated with the type-reduced set $[y_l(\mathbf{x}), y_r(\mathbf{x})]$, be given by

$$R_{TR} \equiv R_{TR}(\mathbf{x}_1, y_1, \ldots, \mathbf{x}_N, y_N)$$

TABLE I
$\sqrt{R_{TR}}$ AND $\sqrt{R_{APP}}$ FOR THE TIME-SERIES FORECASTING EXPERIMENTS BASED ON $R_{TR-\underline{\Delta}}$

			Testing						Validating
			epoch 1	epoch 2	epoch 3	epoch 4	epoch 5	epoch 6	
$w=1$	$\sqrt{R_{TR}}$	mean	0.2256	0.2120	0.2101	0.2094	0.2090	0.2088	0.2089
		SD	0.0263	0.0191	0.0186	0.0184	0.0182	0.0182	0.0144
	$\sqrt{R_{APP}}$	mean	0.2250	0.2127	0.2107	0.2100	0.2095	0.2093	0.2094
		SD	0.0250	0.0187	0.0181	0.0179	0.0178	0.0177	0.0138
$w=0.8$	$\sqrt{R_{TR}}$	mean	0.2285	0.2122	0.2095	0.2085	0.2080	0.2078	0.2079
		SD	0.0266	0.0196	0.0185	0.0182	0.0180	0.0179	0.0139
	$\sqrt{R_{APP}}$	mean	0.2286	0.2128	0.2098	0.2086	0.2080	0.2077	0.2077
		SD	0.0260	0.0191	0.0181	0.0178	0.0176	0.0175	0.0134
$w=0.5$	$\sqrt{R_{TR}}$	mean	0.2396	0.2176	0.2109	0.2087	0.2078	0.2073	0.2074
		SD	0.0278	0.0220	0.0195	0.0185	0.0181	0.0178	0.0136
	$\sqrt{R_{APP}}$	mean	0.2405	0.2183	0.2112	0.2088	0.2077	0.2072	0.2073
		SD	0.0277	0.0217	0.0192	0.0182	0.0178	0.0176	0.0134
$w=0.2$	$\sqrt{R_{TR}}$	mean	0.3026	0.2482	0.2304	0.2213	0.2158	0.2124	0.2128
		SD	0.0353	0.0269	0.0246	0.0229	0.0215	0.0203	0.0156
	$\sqrt{R_{APP}}$	mean	0.3022	0.2487	0.2308	0.2215	0.2158	0.2124	0.2128
		SD	0.0344	0.0269	0.0246	0.0228	0.0213	0.0201	0.0154
$w=0$	$\sqrt{R_{TR}}$	mean	0.4832	0.4835	0.4838	0.4841	0.4844	0.4846	0.4852
		SD	0.0762	0.0808	0.0841	0.0865	0.0883	0.0896	0.0902
	$\sqrt{R_{APP}}$	mean	0.4831	0.4834	0.4837	0.4840	0.4843	0.4846	0.4852
		SD	0.0755	0.0803	0.0837	0.0862	0.0881	0.0894	0.0901

$$=\frac{1}{N}\sum_{i=1}^{N}\left[y_i-\frac{y_l(\mathbf{x}_i)+y_r(\mathbf{x}_i)}{2}\right]^2 \tag{38}$$

and the risk function R_{APP}, associated with its approximation set $[(\underline{y}_l(\mathbf{x})+\bar{y}_l(\mathbf{x}))/2, (\underline{y}_r(\mathbf{x})+\bar{y}_r(\mathbf{x}))/2]$, be given by

$$\begin{aligned} R_{APP} \equiv & R_{APP}(\mathbf{x}_1, y_1, \ldots, \mathbf{x}_N, y_N) \\ = & \frac{1}{N}\sum_{i=1}^{N}\left\{y_i-\frac{1}{2}\left[\frac{\underline{y}_l(\mathbf{x}_i)+\bar{y}_l(\mathbf{x}_i)}{2}\right.\right. \\ & \left.\left.+\frac{\underline{y}_r(\mathbf{x}_i)+\bar{y}_r(\mathbf{x}_i)}{2}\right]\right\}^2 \end{aligned} \tag{39}$$

where $[y_l(\mathbf{x})+y_r(\mathbf{x})]/2$ and $[(\underline{y}_l(\mathbf{x})+\bar{y}_l(\mathbf{x}))/2+(\underline{y}_r(\mathbf{x})+\bar{y}_r(\mathbf{x}))/2]/2$ are the defuzzified outputs of the type-reduced set and its approximation set, respectively. Then

$$\left|\sqrt{R_{TR}}-\sqrt{R_{APP}}\right| \le \sqrt{\frac{1}{N}\sum_{i=1}^{N}\delta^2(\mathbf{x}_i)} \le \sqrt{\frac{1}{N}\sum_{i=1}^{N}\bar{\delta}^2(\mathbf{x}_i)} \tag{40}$$

where $\delta(\mathbf{x})$ and $\bar{\delta}(\mathbf{x})$ are defined in (29) and (30).

Proof: Rewrite (38) and (39) as follows:

$$\begin{aligned} R_{TR} = & \frac{1}{N}\left(\mathbf{y}-\frac{\mathbf{y}_l+\mathbf{y}_r}{2}\right)^T\left(\mathbf{y}-\frac{\mathbf{y}_l+\mathbf{y}_r}{2}\right) \\ = & \frac{1}{N}\left\|\mathbf{y}-\frac{\mathbf{y}_l+\mathbf{y}_r}{2}\right\|^2 \end{aligned} \tag{41}$$

$$\begin{aligned} R_{APP} = & \frac{1}{N}\left[\mathbf{y}-\frac{1}{2}\left(\frac{\underline{\mathbf{y}}_l+\bar{\mathbf{y}}_l}{2}+\frac{\underline{\mathbf{y}}_r+\bar{\mathbf{y}}_r}{2}\right)\right]^T \\ & \times\left[\mathbf{y}-\frac{1}{2}\left(\frac{\underline{\mathbf{y}}_l+\bar{\mathbf{y}}_l}{2}+\frac{\underline{\mathbf{y}}_r+\bar{\mathbf{y}}_r}{2}\right)\right] \\ = & \frac{1}{N}\left\|\mathbf{y}-\frac{1}{2}\left(\frac{\underline{\mathbf{y}}_l+\bar{\mathbf{y}}_l}{2}+\frac{\underline{\mathbf{y}}_r+\bar{\mathbf{y}}_r}{2}\right)\right\|^2 \end{aligned} \tag{42}$$

where $\mathbf{y}, \mathbf{y}_l, \mathbf{y}_r, \underline{\mathbf{y}}_l, \bar{\mathbf{y}}_l, \underline{\mathbf{y}}_r$ and $\bar{\mathbf{y}}_r$ are $N\times 1$ vectors consisting of y_i, $y_l(\mathbf{x}_i)$, $y_r(\mathbf{x}_i)$, $\underline{y}_l(\mathbf{x}_i)$, $\bar{y}_l(\mathbf{x}_i)$, $\underline{y}_r(\mathbf{x}_i)$ and $\bar{y}_r(\mathbf{x}_i)$ $(i=1,\ldots,N)$, respectively. Observe that vectors $[\mathbf{y}-(\mathbf{y}_l+\mathbf{y}_r)/2]$, $[\mathbf{y}-((\underline{\mathbf{y}}_l+\bar{\mathbf{y}}_l)/2+(\underline{\mathbf{y}}_r+\bar{\mathbf{y}}_r)/2)/2]$ and $[(\mathbf{y}_l+\mathbf{y}_r)/2-((\underline{\mathbf{y}}_l+\bar{\mathbf{y}}_l)/2+(\underline{\mathbf{y}}_r+\bar{\mathbf{y}}_r)/2)/2]$ form a triangle in an N-dimensional vector space. Then according to the triangle inequality[2] and Corollary 1

$$\begin{aligned} & \left|\sqrt{R_{TR}}-\sqrt{R_{APP}}\right| \\ & = \frac{1}{\sqrt{N}}\left|\left\|\mathbf{y}-\frac{\mathbf{y}_l+\mathbf{y}_r}{2}\right\|\right. \\ & \qquad \left.-\left\|\mathbf{y}-\frac{1}{2}\left(\frac{\underline{\mathbf{y}}_l+\bar{\mathbf{y}}_l}{2}+\frac{\underline{\mathbf{y}}_r+\bar{\mathbf{y}}_r}{2}\right)\right\|\right| \\ & \le \frac{1}{\sqrt{N}}\left\|\frac{\mathbf{y}_l+\mathbf{y}_r}{2}\right. \\ & \qquad \left.-\frac{1}{2}\left(\frac{\underline{\mathbf{y}}_l+\bar{\mathbf{y}}_l}{2}+\frac{\underline{\mathbf{y}}_r+\bar{\mathbf{y}}_r}{2}\right)\right\| \\ & = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\delta^2(\mathbf{x}_i)} \le \sqrt{\frac{1}{N}\sum_{i=1}^{N}\bar{\delta}^2(\mathbf{x}_i)}. \end{aligned} \tag{43}$$

■

[2]The triangle inequality is $\|\mathbf{y}_1+\mathbf{y}_2\| \le \|\mathbf{y}_1\|+\|\mathbf{y}_2\|$, where $\mathbf{y}_1$ and $\mathbf{y}_2$ are two vectors in an N-dimensional space and $\|\cdot\|$ is the norm defined on the space. From the basic inequality, we can derive $\|\mathbf{y}_1\| = \|(\mathbf{y}_1+\mathbf{y}_2)-\mathbf{y}_2\| \le \|\mathbf{y}_1+\mathbf{y}_2\|+\|\mathbf{y}_2\|$, which means that $\|\mathbf{y}_1\|-\|\mathbf{y}_2\| \le \|\mathbf{y}_1+\mathbf{y}_2\| \le \|\mathbf{y}_1\|+\|\mathbf{y}_2\|$, i.e., $|\|\mathbf{y}_1+\mathbf{y}_2\|-\|\mathbf{y}_1\|| \le \|\mathbf{y}_2\|$. When $\mathbf{y}_1$ represents $((\underline{\mathbf{y}}_l+\bar{\mathbf{y}}_l)/2+(\underline{\mathbf{y}}_r+\bar{\mathbf{y}}_r)/2)/(2-\mathbf{y})$ and $\mathbf{y}_2$ represents $(\mathbf{y}_l+\mathbf{y}_r)/2-((\underline{\mathbf{y}}_l+\bar{\mathbf{y}}_l)/2+(\underline{\mathbf{y}}_r+\bar{\mathbf{y}}_r)/2)/2$, we obtain the inequality in (43).

TABLE II
$\sqrt{R_{TR}}$ AND $\sqrt{R_{APP}}$ FOR THE TIME-SERIES FORECASTING EXPERIMENTS BASED ON $R_{TR-\Delta}$

			Testing						Validating
			epoch 1	epoch 2	epoch 3	epoch 4	epoch 5	epoch 6	
$w=1$	$\sqrt{R_{TR}}$	mean	0.2256	0.2120	0.2101	0.2094	0.2090	0.2088	0.2089
		SD	0.0263	0.0191	0.0186	0.0184	0.0182	0.0182	0.0144
	$\sqrt{R_{APP}}$	mean	0.2250	0.2127	0.2107	0.2100	0.2095	0.2093	0.2094
		SD	0.0250	0.0187	0.0181	0.0179	0.0178	0.0177	0.0138
$w=0.8$	$\sqrt{R_{TR}}$	mean	0.2290	0.2123	0.2097	0.2087	0.2082	0.2080	0.2081
		SD	0.0268	0.0194	0.0184	0.0181	0.0180	0.0179	0.0139
	$\sqrt{R_{APP}}$	mean	0.2290	0.2131	0.2101	0.2090	0.2085	0.2081	0.2082
		SD	0.0262	0.0191	0.0182	0.0179	0.0178	0.0177	0.0136
$w=0.5$	$\sqrt{R_{TR}}$	mean	0.2394	0.2175	0.2111	0.2090	0.2080	0.2075	0.2077
		SD	0.0280	0.0215	0.0192	0.0183	0.0179	0.0177	0.0135
	$\sqrt{R_{APP}}$	mean	0.2405	0.2185	0.2117	0.2093	0.2082	0.2076	0.2078
		SD	0.0278	0.0213	0.0191	0.0183	0.0179	0.0177	0.0135
$w=0.2$	$\sqrt{R_{TR}}$	mean	0.3017	0.2471	0.2292	0.2203	0.2152	0.2121	0.2125
		SD	0.0314	0.0259	0.0238	0.0221	0.0208	0.0198	0.0152
	$\sqrt{R_{APP}}$	mean	0.3018	0.2483	0.2303	0.2211	0.2158	0.2125	0.2130
		SD	0.0306	0.0258	0.0237	0.0220	0.0207	0.0198	0.0152
$w=0$	$\sqrt{R_{TR}}$	mean	0.4832	0.4832	0.4833	0.4833	0.4834	0.4834	0.4835
		SD	0.0613	0.0614	0.0615	0.0617	0.0618	0.0620	0.0633
	$\sqrt{R_{APP}}$	mean	0.4829	0.4830	0.4830	0.4831	0.4831	0.4832	0.4833
		SD	0.0604	0.0605	0.0606	0.0607	0.0608	0.0609	0.0623

In the sequel, we use Δ and $\bar{\Delta}$ to denote the sampled mean-squared δ and $\bar{\delta}$ respectively, i.e.,

$$\Delta \equiv \Delta(\mathbf{x}_1, \ldots, \mathbf{x}_N) = \frac{1}{N}\sum_{i=1}^{N} \delta^2(\mathbf{x}_i) \tag{44}$$

$$\bar{\Delta} \equiv \bar{\Delta}(\mathbf{x}_1, \ldots, \mathbf{x}_N) = \frac{1}{N}\sum_{i=1}^{N} \bar{\delta}^2(\mathbf{x}_i). \tag{45}$$

We usually choose the parameters (e.g., the number of rules and the shapes and parameters of the input, antecedent and consequent MFs) of an interval type-2 FLS to minimize R_{TR}. If we could incorporate the difference between the defuzzified outputs of the type-reduced set and its approximation set, such as $\delta(\mathbf{x})$ or $\bar{\delta}(\mathbf{x})$, during the design procedure, then it should be possible to approximate the type-reduced set by its inner- and outer-bound sets and eliminate type-reduction during the real-time operation period of the interval type-2 FLS.

We propose the following two risk functions for the design of an interval type-2 FLS:

$$R_{TR-\bar{\Delta}}(w) \equiv R_{TR-\bar{\Delta}}(w, \mathbf{x}_1, \ldots, \mathbf{x}_N) = wR_{TR} + (1-w)\bar{\Delta} \tag{46}$$

and

$$R_{TR-\Delta}(w) \equiv R_{TR-\Delta}(w, \mathbf{x}_1, \ldots, \mathbf{x}_N) = wR_{TR} + (1-w)\Delta \tag{47}$$

where $w \in [0,1]$ is a weight. When $w = 1$, $R_{TR-\bar{\Delta}}$ and $R_{TR-\Delta}$ are identical to R_{TR} and both the new design methods reduce to the usual one.

V. Designing an Interval Type-2 FL Predictor for the Mackey–Glass Time Series, Based on $R_{TR-\bar{\Delta}}$ and $R_{TR-\Delta}$

An interval type-2 FLS has been used to predict the chaotic Mackey-Glass time-series in [5]. Liang and Mendel have shown that when the chaotic signal is corrupted by *nonstationary* noise, an interval type-2 FLS achieves much better performance than a type-1 FLS. In this section, we shall design two groups of interval type-2 FLSs, one group based on $R_{TR-\bar{\Delta}}$ and the other group based on $R_{TR-\Delta}$, to predict the Mackey-Glass time-series. We let w in (46) and (47) be 1, 0.8, 0.5, 0.2, and, 0 respectively. We then choose the parameters of the FLS to minimize the corresponding $R_{TR-\bar{\Delta}}(w)$ or $R_{TR-\Delta}(w)$. Since our goal is to demonstrate that the two newly-proposed design methods will permit an interval type-2 FLS to operate without type-reduction, we then compare the operating performance of the interval type-2 FLS with and without type-reduction for all the above values of w and for both design methods.

The Mackey–Glass time series $s(t)$ is modeled as

$$\frac{ds(t)}{dt} = \frac{0.2s(t-\tau)}{1+s^{10}(t-\tau)} - 0.1s(t). \tag{48}$$

When $\tau > 17$, this series exhibits chaotic behavior. After discretization, (48) can be rewritten as

$$f(s,n) = \frac{0.2s(n-\tau)}{1+s^{10}(n-\tau)} - 0.1s(n) \tag{49}$$

and

$$s(n+1) = s(n) + hf(s,n) \tag{50}$$

TABLE III
$\sqrt{\Delta}$, $\sqrt{\bar{\Delta}}$, AND $\sqrt{\Delta_l}+\sqrt{\Delta_r}$ FOR THE TIME-SERIES FORECASTING EXPERIMENTS BASED ON $R_{TR-\bar{\Delta}}$

			Testing						Validating
			epoch 1	epoch 2	epoch 3	epoch 4	epoch 5	epoch 6	
$w=1$	$\sqrt{\Delta}$	mean	0.0145	0.0168	0.0182	0.0191	0.0197	0.0200	0.0201
		SD	0.0053	0.0064	0.0071	0.0075	0.0078	0.0080	0.0079
	$\sqrt{\bar{\Delta}}$	mean	0.0901	0.0958	0.0956	0.0945	0.0937	0.0932	0.0932
		SD	0.0114	0.0112	0.0113	0.0115	0.0116	0.0117	0.0112
	$\sqrt{\Delta_l}+\sqrt{\Delta_r}$	mean	0.0665	0.0707	0.0720	0.0718	0.0715	0.0713	0.0711
		SD	0.0248	0.0241	0.0246	0.0247	0.0245	0.0241	0.0239
$w=0.8$	$\sqrt{\Delta}$	mean	0.0110	0.0127	0.0132	0.0130	0.0125	0.0118	0.0118
		SD	0.0042	0.0050	0.0051	0.0048	0.0046	0.0042	0.0042
	$\sqrt{\bar{\Delta}}$	mean	0.0754	0.0741	0.0657	0.0570	0.0496	0.0435	0.0434
		SD	0.0080	0.0070	0.0059	0.0049	0.0043	0.0039	0.0038
	$\sqrt{\Delta_l}+\sqrt{\Delta_r}$	mean	0.0523	0.0588	0.0591	0.0563	0.0526	0.0487	0.0485
		SD	0.0201	0.0204	0.0192	0.0167	0.0141	0.0119	0.0117
$w=0.5$	$\sqrt{\Delta}$	mean	0.0064	0.0081	0.0081	0.0073	0.0061	0.0050	0.0050
		SD	0.0024	0.0032	0.0033	0.0028	0.0023	0.0020	0.0020
	$\sqrt{\bar{\Delta}}$	mean	0.0592	0.0540	0.0431	0.0326	0.0247	0.0190	0.0190
		SD	0.0054	0.0051	0.0053	0.0048	0.0042	0.0035	0.0035
	$\sqrt{\Delta_l}+\sqrt{\Delta_r}$	mean	0.0320	0.0433	0.0441	0.0392	0.0325	0.0263	0.0262
		SD	0.0137	0.0157	0.0135	0.0095	0.0065	0.0053	0.0052
$w=0.2$	$\sqrt{\Delta}$	mean	0.0031	0.0030	0.0035	0.0037	0.0034	0.0027	0.0027
		SD	0.0012	0.0011	0.0015	0.0017	0.0016	0.0013	0.0013
	$\sqrt{\bar{\Delta}}$	mean	0.0481	0.0415	0.0335	0.0255	0.0188	0.0138	0.0138
		SD	0.0038	0.0032	0.0043	0.0051	0.0046	0.0037	0.0037
	$\sqrt{\Delta_l}+\sqrt{\Delta_r}$	mean	0.0147	0.0206	0.0268	0.0279	0.0249	0.0203	0.0202
		SD	0.0050	0.0088	0.0102	0.0086	0.0059	0.0041	0.0039
$w=0$	$\sqrt{\Delta}$	mean	0.0019	0.0014	0.0010	0.0007	0.0005	0.0004	0.0004
		SD	0.0007	0.0005	0.0004	0.0003	0.0002	0.0002	0.0002
	$\sqrt{\bar{\Delta}}$	mean	0.0435	0.0360	0.0296	0.0244	0.0202	0.0168	0.0168
		SD	0.0021	0.0016	0.0019	0.0024	0.0028	0.0030	0.0030
	$\sqrt{\Delta_l}+\sqrt{\Delta_r}$	mean	0.0159	0.0133	0.0109	0.0090	0.0074	0.0061	0.0061
		SD	0.0075	0.0064	0.0054	0.0046	0.0040	0.0035	0.0035

where h is a small number and the initial values of $s(n)$ for $n \le \tau$ are set randomly.

In our simulations, we chose $h=1$ and $\tau=30$. We assumed $s(k)$ was corrupted by uniformly distributed, zero mean, nonstationary additive noise $n(k)$, so that the available measurements were

$$x(k)=s(k)+n(k), \qquad k=1001,1002,\ldots \tag{51}$$

where -10 dB $\le$ SNR $\le$ 10 dB. Let $\sigma_{n_{-10\ \mathrm{dB}}}$ and $\sigma_{n_{10\ \mathrm{dB}}}$ denote the standard deviations of the noise corresponding to -10 dB and 10 dB SNR, respectively. Then, at each value of k, $\sigma_n(k)$ was assumed to be uniformly distributed in the interval $[\sigma_{n_{-10\ \mathrm{dB}}}, \sigma_{n_{10\ \mathrm{dB}}}]$ that was broken into 100 levels.

In each interval type-2 FLS designed as follows, four antecedents were used for forecasting; namely, $x(k-4)$, $x(k-3)$, $x(k-2)$, and $x(k-1)$ were used to predict $x(k)$ and, two fuzzy sets were used for each antecedent; hence, there were a total of 16 rules. Gaussian primary membership functions of uncertain means ($m \in [m_1, m_2]$) were chosen for the antecedents, Gaussian primary membership functions with uncertain standard deviations ($\sigma \in [\sigma_1, \sigma_2]$) were chosen for the input measurements (i.e., the measurements were modeled as type-2 fuzzy numbers) and center-of-sets type-reduction was used. The parameters to be tuned in each rule included: 1) two mean values for each antecedent ($m^i_{1,k}$, $m^i_{2,k}$, $k=1,\ldots,4$, $i=1,\ldots,16$); 2) two standard deviation values for each input measurement ($\sigma_{1,k}, \sigma_{2,k}$, $k=1,\ldots,4$); and 3) two end points of the centroid of each consequent set (y^i_l, y^i_r, $i=1,\ldots,16$). Hence, there were $2\times4\times16+2\times4+2\times16=168$ parameters for each interval type-2 FLS to be determined during the tuning procedure.

Each interval type-2 FLS was designed based on the first 1000 noisy data: $x(1001),\ldots,x(2000)$. The first 504 design data $[x(1001),\ldots,x(1504)]$ were used to tune the parameters using a steepest descent algorithm so as to minimize the associated risk function $[R_{TR-\bar{\Delta}}(w)$ or $R_{TR-\Delta}(w)]$; whereas, the remaining 496 design data $[x(1505),\ldots,x(2000)]$ were used for testing. The design procedure (training + testing) was implemented for 6 epochs and in each epoch we computed the following quantities for the design testing data:

- the root-mean-squared error between the desired output and the output of the interval type-2 FLS with type-reduction, $\sqrt{R_{TR}}$;

TABLE IV
$\sqrt{\Delta}$, $\sqrt{\bar{\Delta}}$, AND $\sqrt{\Delta_l} + \sqrt{\Delta_r}$ FOR THE TIME-SERIES FORECASTING EXPERIMENTS BASED ON $R_{TR-\Delta}$

			Testing						Validating
			epoch 1	epoch 2	epoch 3	epoch 4	epoch 5	epoch 6	
$w = 1$	$\sqrt{\Delta}$	mean	0.0145	0.0168	0.0182	0.0191	0.0197	0.0200	0.0201
		SD	0.0053	0.0064	0.0071	0.0075	0.0078	0.0080	0.0079
	$\sqrt{\bar{\Delta}}$	mean	0.0901	0.0958	0.0956	0.0945	0.0937	0.0932	0.0932
		SD	0.0114	0.0112	0.0113	0.0115	0.0116	0.0117	0.0112
	$\sqrt{\Delta_l} + \sqrt{\Delta_r}$	mean	0.0665	0.0707	0.0720	0.0718	0.0715	0.0713	0.0711
		SD	0.0248	0.0241	0.0246	0.0247	0.0245	0.0241	0.0239
$w = 0.8$	$\sqrt{\Delta}$	mean	0.0121	0.0138	0.0142	0.0140	0.0137	0.0132	0.0133
		SD	0.0043	0.0048	0.0049	0.0047	0.0044	0.0041	0.0041
	$\sqrt{\bar{\Delta}}$	mean	0.0863	0.0948	0.0960	0.0953	0.0943	0.0933	0.0932
		SD	0.0112	0.0112	0.0116	0.0117	0.0118	0.0119	0.0115
	$\sqrt{\Delta_l} + \sqrt{\Delta_r}$	mean	0.0579	0.0644	0.0655	0.0639	0.0618	0.0597	0.0596
		SD	0.0227	0.0228	0.0229	0.0224	0.0218	0.0211	0.0209
$w = 0.5$	$\sqrt{\Delta}$	mean	0.0084	0.0105	0.0108	0.0105	0.0100	0.0094	0.0094
		SD	0.0030	0.0034	0.0034	0.0032	0.0029	0.0026	0.0028
	$\sqrt{\bar{\Delta}}$	mean	0.0794	0.0892	0.0942	0.0958	0.0957	0.0952	0.0951
		SD	0.0105	0.0107	0.0114	0.0117	0.0118	0.0118	0.0116
	$\sqrt{\Delta_l} + \sqrt{\Delta_r}$	mean	0.0416	0.0539	0.0579	0.0580	0.0568	0.0552	0.0551
		SD	0.0181	0.0202	0.0217	0.0220	0.0219	0.0215	0.0213
$w = 0.2$	$\sqrt{\Delta}$	mean	0.0049	0.0065	0.0074	0.0078	0.0078	0.0076	0.0076
		SD	0.0015	0.0020	0.0021	0.0023	0.0022	0.0022	0.0023
	$\sqrt{\bar{\Delta}}$	mean	0.0705	0.0761	0.0815	0.0857	0.0890	0.0912	0.0911
		SD	0.0079	0.0098	0.0101	0.0103	0.0107	0.0111	0.0111
	$\sqrt{\Delta_l} + \sqrt{\Delta_r}$	mean	0.0222	0.0327	0.0411	0.0457	0.0485	0.0500	0.0499
		SD	0.0083	0.0142	0.0164	0.0176	0.0186	0.0193	0.0192
$w = 0$	$\sqrt{\Delta}$	mean	0.0034	0.0035	0.0035	0.0036	0.0037	0.0037	0.0037
		SD	0.0014	0.0014	0.0014	0.0015	0.0015	0.0015	0.0015
	$\sqrt{\bar{\Delta}}$	mean	0.0666	0.0666	0.0666	0.0665	0.0665	0.0664	0.0664
		SD	0.0066	0.0065	0.0065	0.0065	0.0065	0.0065	0.0064
	$\sqrt{\Delta_l} + \sqrt{\Delta_r}$	mean	0.0213	0.0213	0.0213	0.0213	0.0212	0.0212	0.0211
		SD	0.0087	0.0086	0.0086	0.0085	0.0085	0.0084	0.0084

- the root-mean-squared error between the desired output and the output of the interval type-2 FLS without type-reduction, $\sqrt{R_{APP}}$;
- the root-mean-squared difference between the outputs of the interval type-2 FLS with and without type-reduction, $\sqrt{\Delta}$ and its upper bound $\sqrt{\bar{\Delta}}$;
- the sum of the root-mean-squared difference between the left end points and the root-mean-squared difference between the right end points of the type-reduced set and its approximation set, $\sqrt{\Delta_l} + \sqrt{\Delta_r}$, defined as follows (Fig. 1):

$$\Delta_l \equiv \Delta_l(\mathbf{x}_1, \ldots, \mathbf{x}_N) = \frac{1}{N}\sum_{i=1}^{N} \delta_l^2(\mathbf{x}_i) \tag{52}$$

$$\Delta_r \equiv \Delta_r(\mathbf{x}_1, \ldots, \mathbf{x}_N) = \frac{1}{N}\sum_{i=1}^{N} \delta_r^2(\mathbf{x}_i) \tag{53}$$

where N is the number of testing or validating data ($N = 496$ in our simulation). This quantity reveals the difference between the type-reduced set and its approximation set.

Afterwards, each interval type-2 FLS was *validated* using another set of 496 noisy data $[x(4505), \ldots, x(5000)]$, for which we also computed the above 5 quantities. Validation was done to test our hypothesis that type-reduction is not needed during the real-time operation of an interval type-2 FLS when it is properly designed.

Because the design and validation data are random, we repeated this entire procedure 50 times (50 Monte-Carlo realizations). The means and standard deviations (SDs) of the five quantities $\sqrt{R_{TR}}$, $\sqrt{R_{APP}}$, $\sqrt{\Delta}$, $\sqrt{\bar{\Delta}}$, and $\sqrt{\Delta_l} + \sqrt{\Delta_r}$ for 50 realizations are summarized in Tables I–IV and Figs. 2–5. From these results, we make the following observations.

1) From Tables I and II, we see that for both groups of experiments, based on $R_{TR-\bar{\Delta}}$ and $R_{TR-\Delta}$, respectively, the mean values and the standard deviations of $\sqrt{R_{TR}}$ and $\sqrt{R_{APP}}$, corresponding to the $w = 0$ designs are much worse than those of the other designs, both for the testing data and the validating data (so we only plot the mean values and the standard deviations of $\sqrt{R_{TR}}$ and $\sqrt{R_{APP}}$ for the $w \neq 0$ designs in Figs. 2 and 3; and we only discuss the results for the $w \neq 0$ designs in the following observations). This shows that we should not use

TABLE V
MEANINGS OF y_l^i, y_r^i, $\underline{f}^i$, $\bar{f}^i$, AND M IN (63) FOR DIFFERENT TYPE-REDUCTION METHODS

Type-reduction method	y_l^i and y_r^i defined	$\underline{f}^i$ and $\overline{f}^i$ defined	M defined
center-of-sets	left and right end-points of the centroid of the consequent of the i-th rule	lower and upper firing degrees of the i-th rule	number of rules
centroid[1]	$y_l^i = y_r^i = y^i$, the i-th point in the sampled universe of discourse of the FLS's output	lower and upper membership grades of the i-th sampled point of the FLS's output	number of sampled points
center-of-sums[2]	$y_l^i = y_r^i = y^i$, the i-th point in the sampled universe of discourse of the FLS's output	*sums* of lower and upper membership grades for the i-th sampled point of all rule outputs	number of sampled points
height	$y_l^i = y_r^i = y^i$, a single point in the consequent domain of the i-th rule, usually chosen to be the point having the highest primary membership in the principal MF of the output set	lower and upper firing degrees of the i-th rule	number of rules

1. Prior to calculating the centroid type-reduced set, the fired type-2 fuzzy sets are unioned.

2. Prior to calculating the center-of-sums type-reduced set, the membership functions of the fired type-2 fuzzy sets are added (or a linear combination of them is formed).

$\bar{\Delta}$ or Δ by themselves to tune the parameters of an interval type-2 FLS.

2) From Tables I–IV, we see that the mean values and the SDs of $\sqrt{R_{TR}}$, $\sqrt{R_{APP}}$, $\sqrt{\Delta}$, $\sqrt{\bar{\Delta}}$ and $\sqrt{\Delta_l} + \sqrt{\Delta_r}$ for the testing data and the validating data are identical for the $w = 1$ designs based on $R_{TR-\bar{\Delta}}$ and $R_{TR-\Delta}$. This is because the designs based on $R_{TR-\bar{\Delta}}$ and $R_{TR-\Delta}$ reduce to the usual one when $w = 1$ (so we only plot the mean values and the standard deviations of $\sqrt{\Delta}$, $\sqrt{\bar{\Delta}}$ and $\sqrt{\Delta_l} + \sqrt{\Delta_r}$ for the $w \neq 1$ designs in Figs. 4 and 5).

3) From Figs. 2 and 3, we see that for both groups of experiments, based on $R_{TR-\bar{\Delta}}$ and $R_{TR-\Delta}$, respectively, the mean values and the standard deviations of $\sqrt{R_{TR}}$ and $\sqrt{R_{APP}}$ for the testing data decrease and seem to approach the same limiting values as the training epochs increase. However, when $\bar{\Delta}$ or Δ is weighted too much, e.g., when $w = 0.2$, it takes more epochs of training to let the interval type-2 FLS perform well. From Tables I and II, we see that the results for the validating data show that the mean values and the standard deviations of $\sqrt{R_{TR}}$ and $\sqrt{R_{APP}}$ for all the $w \neq 0$ designs, based on $R_{TR-\bar{\Delta}}$ and $R_{TR-\Delta}$, are close; but the quantities for the $w = 0.2$ designs are a little worse than for the others. If we are only concerned about the performance of an interval type-2 FLS with respect to R_{TR} and R_{APP}, then the designs based on $R_{TR-\bar{\Delta}}$ and $R_{TR-\Delta}$ can achieve similar results.

4) From Tables I and II, it appears that the best results are achieved for the $w = 0.5$ design.

5) From Figs. 4 and 5, we see that the mean values and the standard deviations of $\sqrt{\Delta}$, $\sqrt{\bar{\Delta}}$ and $\sqrt{\Delta_l} + \sqrt{\Delta_r}$ for the designs based on $R_{TR-\bar{\Delta}}$ are always smaller than those based on $R_{TR-\Delta}$ (the solid lines lie below the dotted lines). This shows that the designs based on $R_{TR-\bar{\Delta}}$ can reduce the difference between the type-reduced set and its approximation set, as well as the difference between their defuzzified outputs. If we are also concerned about the approximation of the type-reduced, as well as the defuzzified output, then the designs based on $R_{TR-\bar{\Delta}}$ are preferred to those based on $R_{TR-\Delta}$.

VI. CONCLUSION

Based on a new mathematical interpretation of the Karnik–Mendel iterative procedure for computing the type-reduced set of an interval type-2 FLS, we have derived an inner-bound set and an outer-bound set for the type-reduced set. Our bounds provide estimates of the uncertainty contained in the output of an interval type-2 FLS without having to perform the costly computations of type-reduction.

We have also shown how to incorporate the difference, $\delta(\mathbf{x})$, between the defuzzified outputs of the type-reduced set and its approximation set and its upper bound $\bar{\delta}(\mathbf{x})$ into the design of an interval type-2 FLS, so that the resulting FLS can be used during real-time applications. Our simulation experiments have demonstrated that an interval type-2 FLS designed based on $R_{TR-\Delta}$ or $R_{TR-\bar{\Delta}}$ can operate without type-reduction and can achieve similar performance, in the defuzzified output level, to one that uses type-reduction. Our new method therefore looks very promising to relieve the computation burden of an interval type-2 FLS during operation, which will make an interval type-2 FLS very useful for real-time applications.

We prefer $R_{TR-\bar{\Delta}}$ in the design, because an interval type-2 FLS designed based on it can operate without type-reduction and can achieve similar performance, in both the type-reduced and the defuzzified output levels, to one that uses type-reduction. Before the design, an appropriate weight $w \in [0, 1]$ should be determined in terms of the tradeoff between the convergence speed and the approximation accuracy (a larger value of w results in a faster convergence of the FLS, but a larger difference between the type-reduced set and its approximation set). Our example suggests choosing $w = 0.5$. In summary, after the FLS is designed, its operational equations are: $\underline{y}_l(\mathbf{x})$, $\bar{y}_l(\mathbf{x})$, $\underline{y}_r(\mathbf{x})$ and $\bar{y}_r(\mathbf{x})$, using (25)–(28), $[\underline{y}_l(\mathbf{x}) + \bar{y}_l(\mathbf{x})]/2$ and $[\underline{y}_r(\mathbf{x}) + \bar{y}_r(\mathbf{x})]/2$ (to approximate the type-reduced set) and finally $[(\underline{y}_l(\mathbf{x}) + \bar{y}_l(\mathbf{x}))/2 + (\underline{y}_r(\mathbf{x}) + \bar{y}_r(\mathbf{x})/2)]/2$ (to obtain the defuzzified output from the approximation set).

Appendix I
Background Knowledge About Type-2 Fuzzy Sets and FLSs

In this appendix, we collect some important definitions about type-2 fuzzy sets ([2], [5], [10]).

A. Definition 1

A **type-2 fuzzy set**, denoted $\widetilde{A}$, is characterized by a type-2 membership function $\mu_{\widetilde{A}}(x, u)$, i.e.,

$$\widetilde{A} = \int_{x \in X} \int_{u \in J_x} \mu_{\widetilde{A}}(x, u)/(x, u), \qquad J_x \subseteq [0, 1] \tag{54}$$

where $\int\int$ denotes union over all admissible x and u. At each fixed value of $x \in X$, J_x is the **primary membership** of x.

B. Definition 2

At each value of x, say $x = x'$, the 2D plane whose axes are u and $\mu_{\widetilde{A}}(x', u)$ is called a vertical slice of $\mu_{\widetilde{A}}(x, u)$. A **secondary membership function** is a vertical slice of $\mu_{\widetilde{A}}(x, u)$. It is $\mu_{\widetilde{A}}(x = x', u)$ for $x' \in X$ and $\forall u \in J_{x'} \subseteq [0, 1]$, i.e.,

$$\mu_{\widetilde{A}}(x = x', u) \equiv \mu_{\widetilde{A}}(x') = \int_{u \in J_{x'}} f_{x'}(u)/u, \qquad J_{x'} \subset [0, 1] \tag{55}$$

in which $0 \leq f_{x'}(u) \leq 1$. The type-2 fuzzy set $\widetilde{A}$ can be re-expressed as

$$\begin{aligned}\widetilde{A} &= \int_{x \in X} \mu_{\widetilde{A}}(x)/x \\ &= \int_{x \in X} \left[\int_{u \in J_x} f_x(u)/u\right]/x, \qquad J_x \subseteq [0, 1].\end{aligned} \tag{56}$$

The amplitude, $f_x(u)$, of a secondary membership function is called a **secondary grade**.

C. Definition 3

Assume that each of the secondary membership functions of a type-2 fuzzy set has only one secondary grade that equals one. A **principal membership function** is the union of all such points at which this occurs, i.e.,

$$\mu_{\text{principal}}(x) = \int_{x \in X} u/x \text{ where } f_x(u) = 1 \tag{57}$$

and is associated with a type-1 fuzzy set.

D. Definition 4

Uncertainty in the primary memberships of a type-2 fuzzy set $\widetilde{A}$ consists of a bounded region which is called the footprint of uncertainty (FOU) of $\widetilde{A}$, i.e.,

$$FOU\left(\widetilde{A}\right) = \bigcup_{x \in X} J_x \tag{58}$$

E. Definition 5

The UMFs and LMFs and of a type-2 fuzzy set $\widetilde{A}$ are two type-1 membership functions that are bounds for its FOU. The upper membership function is associated with the upper bound of FOU($\widetilde{A}$) and is denoted $\bar{\mu}_{\widetilde{A}}(x)$, $\forall x \in X$. The lower membership function is associated with the lower bound of FOU($\widetilde{A}$) and is denoted $\underline{\mu}_{\widetilde{A}}(x)$, $\forall x \in X$, i.e.,

$$\bar{\mu}_{\widetilde{A}}(x) \equiv \overline{\text{FOU}\left(\widetilde{A}\right)} \quad \forall x \in X \tag{59}$$

and

$$\underline{\mu}_{\widetilde{A}}(x) \equiv \underline{\text{FOU}\left(\widetilde{A}\right)} \quad \forall x \in X \tag{60}$$

F. Definition 6

For continuous X and J_x, an **embedded type-1 set** A_e of a type-2 fuzzy set $\widetilde{A}$, is

$$A_e = \int_{x \in X} \theta/x, \qquad \theta \in J_x \subseteq [0, 1] \tag{61}$$

For discrete X (with N points) and J_{x_i} (with M_i points)

$$A_e = \sum_{i=1}^{N} \theta_i/x_i, \qquad \theta_i \in J_{x_i} \subseteq [0, 1], \qquad x_i \in X \tag{62}$$

In the continuous case, the number of A_e is uncountable, whereas for the discrete case, there are $\prod_{i=1}^{N} M_i$ of the A_e.

G. Definition 7

An FLS (which contains rules, fuzzifier, inference engine and output processor) is a **type-2 FLS** when either its inputs, antecedents, or consequents are type-2 fuzzy sets. The output processor of a type-2 FLS consists of type-reduction followed by defuzzification. For an **interval type-2 FLS**, the secondary membership functions of the inputs, antecedents and consequent sets are all intervals sets.

H. Definition 8

An **embedded type-1 FLS** for a type-2 FLS is associated with the embedded type-1 fuzzy sets of the inputs, antecedents and consequents. A type-2 FLS can be interpreted as a collection of embedded type-1 FLSs.

Appendix II
Brief Comparison of Different Type-Reduction Methods for Interval Type-2 FLSs

Center-of-sets, centroid, center-of-sums, and height type-reduction can all be expressed as in (6) [2], [10], which we repeat here for the convenience of the readers

$$Y_{TR}(\mathbf{x}) = \int_{y^i \in [y_l^i, y_r^i]} \int_{f^i \in [\underline{f}^i, \bar{f}^i]} 1 \Bigg/ \frac{\sum_{i=1}^{M} f^i y^i}{\sum_{i=1}^{M} f^i}. \tag{63}$$

For the different type-reduction methods, y_l^i, y_r^i, $\underline{f}^i$, $\bar{f}^i$ and M have different meanings, as summarized in TABLE V.

Due to the differences among these type-reduction methods, the boundary type-1 FLSs are different depending on which type-reduction method is implemented. When centroid, center-of-sums and height type-reduction are used, since $y_l^i = y_r^i = y^i$, an interval type-2 FLS has only *two* associated boundary type-1 FLSs, namely

$$\{LMFs\} : y_l^{(0)}(\mathbf{x}) = y_r^{(M)}(\mathbf{x}) = \frac{\sum_{i=1}^{M} \underline{f}^i y^i}{\sum_{i=1}^{M} \underline{f}^i} \tag{64}$$

$$\{UMFs\} : y_l^{(M)}(\mathbf{x}) = y_r^{(0)}(\mathbf{x}) = \frac{\sum_{i=1}^{M} \bar{f}^i y^i}{\sum_{i=1}^{M} \bar{f}^i}. \tag{65}$$

For these type-reduction methods, the inner- and outer-bound sets of the type-reduced set are

$$\begin{aligned}\bar{y}_l(\mathbf{x}) &= \min\left\{y_l^{(0)}(\mathbf{x}), y_l^{(M)}(\mathbf{x})\right\}\\ &= \min\left\{\frac{\sum_{i=1}^{M} \underline{f}^i y^i}{\sum_{i=1}^{M} \underline{f}^i}, \frac{\sum_{i=1}^{M} \bar{f}^i y^i}{\sum_{i=1}^{M} \bar{f}^i}\right\}\end{aligned} \tag{66}$$

$$\begin{aligned}\underline{y}_r(\mathbf{x}) &= \max\left\{y_r^{(0)}(\mathbf{x}), y_r^{(M)}(\mathbf{x})\right\}\\ &= \max\left\{\frac{\sum_{i=1}^{M} \bar{f}^i y^i}{\sum_{i=1}^{M} \bar{f}^i}, \frac{\sum_{i=1}^{M} \underline{f}^i y^i}{\sum_{i=1}^{M} \underline{f}^i}\right\}\end{aligned} \tag{67}$$

and

$$\begin{aligned}\underline{y}_l(\mathbf{x}) = \bar{y}_l(\mathbf{x}) - &\left[\frac{\sum_{i=1}^{M} \left(\bar{f}^i - \underline{f}^i\right)}{\sum_{i=1}^{M} \bar{f}^i \sum_{i=1}^{M} \underline{f}^i}\right.\\ &\left.\times \frac{\sum_{i=1}^{M} \underline{f}^i \left(y^i - y^1\right) \sum_{i=1}^{M} \bar{f}^i \left(y^M - y^i\right)}{\sum_{i=1}^{M} \underline{f}^i \left(y^i - y^1\right) + \sum_{i=1}^{M} \bar{f}^i \left(y^M - y^i\right)}\right]\end{aligned} \tag{68}$$

$$\begin{aligned}\bar{y}_r(\mathbf{x}) = \underline{y}_r(\mathbf{x}) + &\left[\frac{\sum_{i=1}^{M} \left(\bar{f}^i - \underline{f}^i\right)}{\sum_{i=1}^{M} \bar{f}^i \sum_{i=1}^{M} \underline{f}^i}\right.\\ &\left.\times \frac{\sum_{i=1}^{M} \bar{f}^i \left(y^i - y^1\right) \sum_{i=1}^{M} \underline{f}^i \left(y^M - y^i\right)}{\sum_{i=1}^{M} \bar{f}^i \left(y^i - y^1\right) + \sum_{i=1}^{M} \underline{f}^i \left(y^M - y^i\right)}\right]\end{aligned} \tag{69}$$

APPENDIX III
THE KARNIK-MENDEL ITERATIVE PROCEDURE

$L^\star(\mathbf{x})$ can be obtained by the following iterative procedure.

1) Initialize f^i by setting $f^i = (\bar{f}^i + \underline{f}^i)/2$ for $i = 1, \ldots, M$ and compute $y' = \sum_{i=1}^{M} f^i y_l^i / \sum_{i=1}^{M} f^i$.
2) Find L $(1 \le L \le M-1)$ so that $y_l^L \le y' \le y_l^{L+1}$.
3) Set $f^i = \bar{f}^i$ for $1 \le i \le L$ and $f^i = \underline{f}^i$ for $L+1 \le i \le M$ and then compute $y'' = \sum_{i=1}^{M} f^i y_l^i / \sum_{i=1}^{M} f^i$.
4) Check if $y'' = y'$. If yes, stop. Otherwise, go to Step 5).
5) Set y' equal to y''. Go to Step 2).

One pass through Steps 2)–5) is called one iteration. Step 1) is an initialization. $R^\star(\mathbf{x})$ can be obtained using a procedure similar to the above. Only two changes need to be made:

1) y_l^i is replaced by y_r^i;
2) in Step 3), set $f^i = \underline{f}^i$ for $1 \le i \le L$ and $f^i = \bar{f}^i$ for $L+1 \le i \le M$, and then compute $y'' = \sum_{i=1}^{M} f^i y_r^i / \sum_{i=1}^{M} f^i$.

APPENDIX IV
PROOFS OF THEOREM 2 AND COROLLARY 1

A. Proof of Theorem 2

We prove Theorem 2 in two steps. First, we show $[\bar{y}_l(\mathbf{x}), \underline{y}_r(\mathbf{x})]$ is an inner bound for the type-reduced set (Fig. 1). Then, we derive the outer-bound set, $[\underline{y}_l(\mathbf{x}), \bar{y}_r(\mathbf{x})]$, based on the distance between the type-reduced set and its inner-bound set.

1) Inner-Bound Set $[\bar{y}_l(\mathbf{X}), \underline{y}_r(\mathbf{X})]$*:* It follows from (15) that

$$y_l(\mathbf{x}) \le y_l^{(0)}(\mathbf{x}) \text{ and} y_l(\mathbf{x}) \le y_l^{(M)}(\mathbf{x}) \tag{70}$$

and from (17) that

$$y_r(\mathbf{x}) \ge y_r^{(0)}(\mathbf{x}) \text{ and } y_r(\mathbf{x}) \ge y_r^{(M)}(\mathbf{x}) \tag{71}$$

Hence

$$y_l(\mathbf{x}) \le \min\left\{y_l^{(0)}(\mathbf{x}), y_l^{(M)}(\mathbf{x})\right\} \equiv \bar{y}_l(\mathbf{x}) \tag{72}$$

and

$$y_r(\mathbf{x}) \ge \max\left\{y_r^{(0)}(\mathbf{x}), y_r^{(M)}(\mathbf{x})\right\} \equiv \underline{y}_r(\mathbf{x}) \tag{73}$$

Next, we show that $\bar{y}_l(\mathbf{x}) \le \underline{y}_r(\mathbf{x})$ so that $[\bar{y}_l(\mathbf{x}), \underline{y}_r(\mathbf{x})]$ is a valid set. Observe from (19) and (21) that $y_l^{(0)}$ and $y_r^{(M)}(\mathbf{x})$ share the same parameters $\{\underline{f}^1, \ldots, \underline{f}^M\}$, each modified by different variables—$\{y_l^1, \ldots, y_l^M\}$ versus $\{y_r^1, \ldots, y_r^M\}$. Because $(\sum_{i=1}^{M} y^i f^i)/(\sum_{i=1}^{M} f^i)$ is a monotonically increasing function of the variables y^i, it follows that

$$y_l^{(0)}(\mathbf{x}) = \frac{\sum_{i=1}^{M} \underline{f}^i y_l^i}{\sum_{i=1}^{M} \underline{f}^i} \le \frac{\sum_{i=1}^{M} \underline{f}^i y_r^i}{\sum_{i=1}^{M} \underline{f}^i} = y_r^{(M)}(\mathbf{x}).. \tag{74}$$

In a similar manner for $y_l^{(M)}(\mathbf{x})$ and $y_r^{(0)}(\mathbf{x})$, it follows that

$$y_l^{(M)}(\mathbf{x}) = \frac{\sum_{i=1}^{M} \bar{f}^i y_l^i}{\sum_{i=1}^{M} \bar{f}^i} \le \frac{\sum_{i=1}^{M} \bar{f}^i y_r^i}{\sum_{i=1}^{M} \bar{f}^i} = y_r^{(0)}(\mathbf{x}). \tag{75}$$

Consequently

$$\begin{aligned}\bar{y}_l(\mathbf{x}) &\equiv \min\left\{y_l^{(0)}(\mathbf{x}), y_l^{(M)}(\mathbf{x})\right\}\\ &\le \max\left\{y_r^{(0)}(\mathbf{x}), y_r^{(M)}(\mathbf{x})\right\} \equiv \underline{y}_r(\mathbf{x})\end{aligned} \tag{76}$$

which means $[\bar{y}_l(\mathbf{x}), \underline{y}_r(\mathbf{x})]$ is a valid inner-bound set for $[y_l(\mathbf{x}), y_r(\mathbf{x})]$.

2) Outer-Bound Set $[\underline{y}_l(\mathbf{X}), \bar{y}_r(\mathbf{X})]$: We shall analyze the difference between the type-reduced set and its inner-bound set to derive its outer-bound set. For notational simplicity, $L^\star(\mathbf{x}) \equiv L^\star$ and $R^\star(\mathbf{x}) \equiv R^\star$.

1) In (72), we have shown $\bar{y}_l(\mathbf{x}) - y_l(\mathbf{x}) \geq 0$. We next show that $\bar{y}_l(\mathbf{x}) - y_l(\mathbf{x})$ is bounded from above, i.e., $\bar{y}_l(\mathbf{x}) - y_l(\mathbf{x}) \leq c$, from which it follows that $y_l(\mathbf{x}) \geq \bar{y}_l(\mathbf{x}) - c \equiv \underline{y}_l(\mathbf{x})$. To determine c, we begin by using the following inequality:

$$\min[A, B] \leq \alpha A + (1-\alpha) B \tag{77}$$

for $0 \leq \alpha \leq 1$. To understand this inequality we can, without loss of generality, assume $A \geq B$, in which case $\min[A, B] = B$. Thus, we need only show that $B \leq \alpha a + (1-\alpha)B = \alpha(A - B) + B$. But this inequality is valid, since $A \geq B$ and $\alpha \geq 0$. Although (77) is in terms of a free parameter α, in the following we determine an optimal value for α, α^*.

From (25), observe that $\bar{y}_l(\mathbf{x}) - y_l(\mathbf{x}) = \min\left\{y_l^{(0)}(\mathbf{x}) - y_l(\mathbf{x}), y_l^{(M)}(\mathbf{x}) - y_l(\mathbf{x})\right\}$, hence, applying (77) to this equation, we find

$$\bar{y}_l(\mathbf{x}) - y_l(\mathbf{x}) \leq \alpha\left[y_l^{(0)}(\mathbf{x}) - y_l(\mathbf{x})\right] + (1-\alpha)\left[y_l^{(M)}(\mathbf{x}) - y_l(\mathbf{x})\right]. \tag{78}$$

Before finding α^*, we obtain expressions for $y_l^{(0)}(\mathbf{x}) - y_l(\mathbf{x})$ and $y_l^{(M)}(\mathbf{x}) - y_l(\mathbf{x})$.

Expressing $y_l^{(0)}(\mathbf{x})$ and $y_l^{(M)}(\mathbf{x})$ using (13) and using (15) for $y_l(\mathbf{x})$, we find that

$$y_l^{(0)}(\mathbf{x}) - y_l(\mathbf{x}) = \frac{\sum_{i=1}^{M} \underline{f}^i y_l^i}{\sum_{i=1}^{M} \underline{f}^i} - \frac{\sum_{i=1}^{L^\star} \bar{f}^i y_l^i + \sum_{i=L^\star+1}^{M} \underline{f}^i y_l^i}{\sum_{i=1}^{L^\star} \bar{f}^i + \sum_{i=L^\star+1}^{M} \underline{f}^i} \geq 0 \tag{79}$$

where the latter inequality follows from (70), hence,

$$\begin{aligned} y_l^{(0)}(\mathbf{x}) - y_l(\mathbf{x}) &= \frac{\overbrace{\sum_{i=1}^{M} \underline{f}^i y_l^i}^{a}}{\underbrace{\sum_{i=1}^{M} \underline{f}^i}_{c}} - \frac{\overbrace{\sum_{i=1}^{M} \underline{f}^i y_l^i}^{a} + \overbrace{\sum_{i=1}^{L^\star} \left(\bar{f}^i - \underline{f}^i\right) y_l^i}^{b}}{\underbrace{\sum_{i=1}^{M} \underline{f}^i}_{c} + \underbrace{\sum_{i=1}^{L^\star} \left(\bar{f}^i - \underline{f}^i\right)}_{d}} \\ &= \frac{ad - bc}{c(c+d)} = \frac{a}{c} \cdot \frac{d}{c+d} - \frac{b}{c+d} \\ &= y_l^{(0)}(\mathbf{x}) \cdot \frac{\sum_{i=1}^{L^\star} \left(\bar{f}^i - \underline{f}^i\right)}{\sum_{i=1}^{L^\star} \bar{f}^i + \sum_{i=L^\star+1}^{M} \underline{f}^i} \\ &\quad - \frac{\sum_{i=1}^{L^\star} \left(\bar{f}^i - \underline{f}^i\right) y_l^i}{\sum_{i=1}^{L^\star} \bar{f}^i + \sum_{i=L^\star+1}^{M} \underline{f}^i} \\ &\leq y_l^{(0)}(\mathbf{x}) \frac{\sum_{i=1}^{L^\star} \left(\bar{f}^i - \underline{f}^i\right)}{\sum_{i=1}^{L^\star} \bar{f}^i + \sum_{i=L^\star+1}^{M} \underline{f}^i} \\ &\quad - y_l^1 \frac{\sum_{i=1}^{L^\star} \left(\bar{f}^i - \underline{f}^i\right)}{\sum_{i=1}^{L^\star} \bar{f}^i + \sum_{i=L^\star+1}^{M} \underline{f}^i} \\ &= \frac{\left[y_l^{(0)}(\mathbf{x}) - y_l^1\right] \sum_{i=1}^{L^\star} \left(\bar{f}^i - \underline{f}^i\right)}{\sum_{i=1}^{L^\star} \bar{f}^i + \sum_{i=L^\star+1}^{M} \underline{f}^i} \end{aligned} \tag{80}$$

since $y_l^1 \leq y_l^2 \leq \cdots \leq y_l^{L^\star}$. In a similar manner, we find that

$$\begin{aligned} y_l^{(M)}(\mathbf{x}) - y_l(\mathbf{x}) &= \frac{\sum_{i=1}^{M} \bar{f}^i y_l^i}{\sum_{i=1}^{M} \bar{f}^i} - \frac{\sum_{i=1}^{L^\star} \bar{f}^i y_l^i + \sum_{i=L^\star+1}^{M} \underline{f}^i y_l^i}{\sum_{i=1}^{L^\star} \bar{f}^i + \sum_{i=L^\star+1}^{M} \underline{f}^i} \\ &\geq 0 \end{aligned} \tag{81}$$

where the latter inequality also follows from (70), hence

$$\begin{aligned} y_l^{(M)}(\mathbf{x}) - y_l(\mathbf{x}) &= \frac{\sum_{i=L^\star+1}^{M} \left(\bar{f}^i - \underline{f}^i\right) y_l^i}{\sum_{i=1}^{L^\star} \bar{f}^i + \sum_{i=L^\star+1}^{M} \underline{f}^i} \\ &\quad - y_l^{(M)}(\mathbf{x}) \frac{\sum_{i=L^\star+1}^{M} \left(\bar{f}^i - \underline{f}^i\right)}{\sum_{i=1}^{L^\star} \bar{f}^i + \sum_{i=L^\star+1}^{M} \underline{f}^i} \\ &\leq y_l^M \frac{\sum_{i=L^\star+1}^{M} \left(\bar{f}^i - \underline{f}^i\right)}{\sum_{i=1}^{L^\star} \bar{f}^i + \sum_{i=L^\star+1}^{M} \underline{f}^i} \\ &\quad - y_l^{(M)}(\mathbf{x}) \frac{\sum_{i=L^\star+1}^{M} \left(\bar{f}^i - \underline{f}^i\right)}{\sum_{i=1}^{L^\star} \bar{f}^i + \sum_{i=L^\star+1}^{M} \underline{f}^i} \\ &= \frac{\left[y_l^M - y_l^{(M)}(\mathbf{x})\right] \sum_{i=L^\star+1}^{M} \left(\bar{f}^i - \underline{f}^i\right)}{\sum_{i=1}^{L^\star} \bar{f}^i + \sum_{i=L^\star+1}^{M} \underline{f}^i}. \end{aligned} \tag{82}$$

Note that it is always possible to express $\sum_{i=1}^{L^\star}(\bar{f}^i-\underline{f}^i)$ and $\sum_{i=L^\star+1}^{M}(\bar{f}^i-\underline{f}^i)$ in terms of $\sum_{i=1}^{M}(\bar{f}^i-\underline{f}^i)$, as

$$\sum_{i=1}^{L^\star}\left(\bar{f}^i-\underline{f}^i\right)=t\sum_{i=1}^{M}\left(\bar{f}^i-\underline{f}^i\right) \tag{83}$$

and

$$\sum_{i=L^\star+1}^{M}\left(\bar{f}^i-\underline{f}^i\right)=(1-t)\sum_{i=1}^{M}\left(\bar{f}^i-\underline{f}^i\right) \tag{84}$$

where $t\equiv t(\mathbf{x})\in[0,1]$ is determined by $L^\star(\mathbf{x})$ (just solve (83) for t). Substituting (80)–(84) into (78), we find that

$$\begin{aligned}
&\bar{y}_l(\mathbf{x})-y_l(\mathbf{x})\\
&\le\alpha\frac{\left[y_l^{(0)}(\mathbf{x})-y_l^1\right]\sum_{i=1}^{L^\star}\left(\bar{f}^i-\underline{f}^i\right)}{\sum_{i=1}^{L^\star}\bar{f}^i+\sum_{i=L^\star+1}^{M}\underline{f}^i}\\
&\quad+(1-\alpha)\frac{\left[y_l^M-y_l^{(M)}(\mathbf{x})\right]\sum_{i=L^\star+1}^{M}\left(\bar{f}^i-\underline{f}^i\right)}{\sum_{i=1}^{L^\star}\bar{f}^i+\sum_{i=L^\star+1}^{M}\underline{f}^i}\\
&=\frac{\alpha t\left[y_l^{(0)}(\mathbf{x})-y_l^1\right]+(1-\alpha)(1-t)\left[y_l^M-y_l^{(M)}(\mathbf{x})\right]}{t\sum_{i=1}^{M}\bar{f}^i+(1-t)\sum_{i=1}^{M}\underline{f}^i}\\
&\quad\times\sum_{i=1}^{M}(\bar{f}^i-\underline{f}^i)\\
&\equiv g(\alpha,t)\sum_{i=1}^{M}\left(\bar{f}^i-\underline{f}^i\right)\\
&\le\max_{t\in[0,1]}g(\alpha,t)\sum_{i=1}^{M}\left(\bar{f}^i-\underline{f}^i\right)
\end{aligned}\tag{85}$$

where

$$g(\alpha,t)\equiv\frac{\alpha t\left[y_l^{(0)}(\mathbf{x})-y_l^1\right]+(1-\alpha)(1-t)\left[y_l^M-y_l^{(M)}(\mathbf{x})\right]}{t\sum_{i=1}^{M}\bar{f}^i+(1-t)\sum_{i=1}^{M}\underline{f}^i}. \tag{86}$$

Next, we treat α and t as independent variables and use the min–max method to find α^*, i.e.,

$$\alpha^*=\arg\min_{\alpha\in[0,1]}\left[\max_{t\in[0,1]}g(\alpha,t)\right]=\arg\min_{\alpha\in[0,1]}g\left(\alpha,t^*\right). \tag{87}$$

Notice that $g(\alpha,t^*)$ is the maximum of $g(\alpha,t)$ with respect to t and $g(\alpha^*,t^*)$ is the minimum of $g(\alpha,t^*)$ with respect to α. Since $t\in[0,1]$ is indirectly determined by the input x, we cannot arbitrarily choose its value; instead, we consider the worst case (i.e., the maximum value) of $g(\alpha,t)$ with respect to t, to find the upper bound for $\bar{y}_l(\mathbf{x})-y_l(\mathbf{x})$. On the other hand, since $\alpha\in[0,1]$ is a free parameter, we can choose its value arbitrarily to find a tight upper bound for $\bar{y}_l(\mathbf{x})-y_l(\mathbf{x})$.

To find t^* and $g(\alpha,t^*)$, we calculate the partial derivative of $g(\alpha,t)$ in (86) with respect to t as follows:

$$\frac{\partial g(\alpha,t)}{\partial t}=\frac{\alpha\left[y_l^{(0)}-y_l^1\right]\sum_{i=1}^{M}\underline{f}^i-(1-\alpha)\left[y_l^M-y_l^{(M)}(\mathbf{x})\right]\sum_{i=1}^{M}\bar{f}^i}{\left[t\sum_{i=1}^{M}\bar{f}^i+(1-t)\sum_{i=1}^{M}\underline{f}^i\right]^2}. \tag{88}$$

Because the numerator of (88) is not a function of t, we cannot determine t^* by setting $\partial g(\alpha,t)/\partial t=0$. Instead, we must analyze (88) in order to determine t^*. We observe the following from (88).

- When α is chosen so that

$$0\le\alpha<\alpha^* \tag{89}$$

with α^* defined as

$$\alpha^*\equiv\frac{\left[y_l^M-y_l^{(M)}(\mathbf{x})\right]\sum_{i=1}^{M}\bar{f}^i}{\left[y_L^{(0)}(\mathbf{x})-y_l^1\right]\sum_{i=1}^{M}\underline{f}^i+\left[y_l^M-y_l^{(M)}(\mathbf{x})\right]\sum_{i=1}^{M}\bar{f}^i} \tag{90}$$

then $\partial g(\alpha,t)/\partial t<0$, which means that $g(\alpha,t)$ is a monotonically decreasing function with respect to t and, therefore, its maximum value with respect to t occurs at $t^*=0$, i.e.,

$$\begin{aligned}g(\alpha,t^\star)&=\max_{t\in[0,1]}g(\alpha,t)=g(\alpha,0)\\&=\frac{(1-\alpha)\left[y_l^M-y_l^{(M)}(\mathbf{x})\right]}{\sum_{i=1}^{M}\underline{f}^i}\end{aligned} \tag{91}$$

In this case, $g(\alpha,t^*)$ is a monotonically decreasing function with respect to α and, therefore

$$\begin{aligned}&\inf_{0\le\alpha<\alpha^*}g\left(\alpha,t^*\right)\\&\quad=\lim_{\alpha\to\alpha^*}g\left(\alpha,t^*\right)\\&\quad=\frac{\left[y_l^{(0)}(\mathbf{x})-y_l^1\right]\left[y_l^M-y_l^{(M)}(\mathbf{x})\right]}{\left[y_l^{(0)}(\mathbf{x})-y_l^1\right]\sum_{i=1}^{M}\underline{f}^i+\left[y_l^M-y_l^{(M)}(\mathbf{x})\right]\sum_{i=1}^{M}\bar{f}^i}.\end{aligned} \tag{92}$$

- When α is chosen so that

$$\alpha=\alpha^* \tag{93}$$

then $\partial g(\alpha,t)/\partial t=0$, which means $g(\alpha,t)$ is independent of t, i.e.,

$$\begin{aligned}&g\left(\alpha^*,t^*\right)\\&\quad=\max_{t\in[0,1]}g\left(\alpha^*,t\right)=g\left(\alpha^*\right)\end{aligned}$$

$$= \frac{\left[y_l^{(0)}(\mathbf{x}) - y_l^1\right]\left[y_l^M - y_l^{(M)}(\mathbf{x})\right]}{\left[y_l^{(0)}(\mathbf{x}) - y_l^1\right]\sum_{i=1}^{M}\underline{f}^i + \left[y_l^M - y_l^{(M)}(\mathbf{x})\right]\sum_{i=1}^{M}\bar{f}^i}. \tag{94}$$

- When α is chosen so that

$$\alpha^* < \alpha \le 1 \tag{95}$$

then $\partial g(\alpha, t)/\partial t > 0$, which means that $g(\alpha, t)$ is a monotonically increasing function with respect to t and, therefore, its maximum value with respect to t occurs at $t^* = 1$, i.e.,

$$g(\alpha, t^*) = \max_{t\in[0,1]} g(\alpha, t) = g(\alpha, 1) = \frac{\alpha\left[y_l^{(0)}(\mathbf{x}) - y_l^1\right]}{\sum_{i=1}^{M}\bar{f}^i}. \tag{96}$$

In this case, $g(\alpha, t^*)$ is a monotonically increasing function with respect to α and, therefore

$$\begin{aligned}\inf_{\alpha^*<\alpha\le 1} & g(\alpha, t^*) \\ &= \lim_{\alpha\to\alpha^*} g(\alpha, t^*) \\ &= \frac{\left[y_l^{(0)}(\mathbf{x}) - y_l^1\right]\left[y_l^M - y_l^{(M)}(\mathbf{x})\right]}{\left[y_l^{(0)}(\mathbf{x}) - y_l^1\right]\sum_{i=1}^{M}\underline{f}^i + \left[y_l^M - y_l^{(M)}(\mathbf{x})\right]\sum_{i=1}^{M}\bar{f}^i}.\end{aligned} \tag{97}$$

Equations (92), (94), and (97) show that min–max of $g(\alpha, t)$ is achieved when $\alpha = \alpha^*$ and is independent of t, i.e.,

$$\begin{aligned}\min_{\alpha\in[0,1]}\max_{t\le[0,1]} & g(\alpha, t) \\ &= g(\alpha^*) \\ &= \frac{\left[y_l^{(0)}(\mathbf{x}) - y_l^1\right]\left[y_l^M - y_l^{(M)}(\mathbf{x})\right]}{\left[y_l^{(0)}(\mathbf{x}) - y_l^1\right]\sum_{i=1}^{M}\underline{f}^i + \left[y_l^M - y_l^{(M)}(\mathbf{x})\right]\sum_{i=1}^{M}\bar{f}^i}.\end{aligned} \tag{98}$$

Therefore, we substitute $g(\alpha^*)$ into (85). Doing this, we find that

$$\begin{aligned}\bar{y}_l(\mathbf{x}) & - y_l(\mathbf{x}) \\ &\le g(\alpha^*)\sum_{i=1}^{M}\left(\bar{f}^i - \underline{f}^i\right) \\ &= \frac{\left[y_l^{(0)}(\mathbf{x}) - y_l^1\right]\left[y_l^M - y_l^{(M)}(\mathbf{x})\right]\sum_{i=1}^{M}\left(\bar{f}^i - \underline{f}^i\right)}{\left[y_l^{(0)}(\mathbf{x}) - y_l^1\right]\sum_{i=1}^{M}\underline{f}^i + \left[y_l^M - y_l^{(M)}(\mathbf{x})\right]\sum_{i=1}^{M}\bar{f}^i} \\ &= \frac{\sum_{i=1}^{M}\underline{f}^i\left(y_l^i - y_l^1\right)\sum_{i=1}^{M}\bar{f}^i\left(y_l^M - y_l^i\right)}{\sum_{i=1}^{M}\underline{f}^i\left(y_l^i - y_l^1\right) + \sum_{i=1}^{M}\bar{f}^i\left(y_l^M - y_l^i\right)} \\ &\quad\times\frac{\sum_{i=1}^{M}\left(\bar{f}^i - \underline{f}^i\right)}{\sum_{i=1}^{M}\bar{f}^i\sum_{i=1}^{M}\underline{f}^i}\end{aligned} \tag{99}$$

where we have used the definitions of $y_l^{(0)}(\mathbf{x})$ and $y_l^{(M)}(\mathbf{x})$ in (13) to get the last line from the second line.

2) Proceeding in the same way, we get a similar result for $y_r(\theta) - \underline{y}_r(\mathbf{x})$, i.e.,

$$y_r(\mathbf{x}) - \underline{y}_r(\mathbf{x}) \le \frac{\sum_{i=1}^{M}\bar{f}^i\left(y_r^i - y_r^1\right)\sum_{i=1}^{M}\underline{f}^i\left(y_r^M - y_r^i\right)}{\sum_{i=1}^{M}\bar{f}^i\left(y_r^i - y_r^1\right) + \sum_{i=1}^{M}\underline{f}^i\left(y_r^M - y_r^i\right)} \times\frac{\sum_{i=1}^{M}\left(\bar{f}^i - \underline{f}^i\right)}{\sum_{i=1}^{M}\bar{f}^i\sum_{i=1}^{M}\underline{f}^i}. \tag{100}$$

3) From (99) and (100), we obtain the following lower bound $\underline{y}_l(\mathbf{x})$ for $y_l(\mathbf{x})$ and upper bound, $\bar{y}_r(\mathbf{x})$ for $y_r(\mathbf{x})$:

$$y_l(\mathbf{x}) \ge \bar{y}_l(\mathbf{x}) - \left[\frac{\sum_{i=1}^{M}\underline{f}^i\left(y_l^i - y_l^1\right)\sum_{i=1}^{M}\bar{f}^i\left(y_l^M - y_l^i\right)}{\sum_{i=1}^{M}\underline{f}^i\left(y_l^i - y_l^1\right) + \sum_{i=1}^{M}\bar{f}^i\left(y_l^M - y_l^i\right)} \times\frac{\sum_{i=1}^{M}\left(\bar{f}^i - \underline{f}^i\right)}{\sum_{i=1}^{M}\bar{f}^i\sum_{i=1}^{M}\underline{f}^i}\right] \tag{101}$$

with the right-hand side of the inequality defined as $\underline{y}_l(\mathbf{x})$ and

$$y_r(\mathbf{x}) \le \underline{y}_r(\mathbf{x}) + \left[\frac{\sum_{i=1}^{M}\bar{f}^i\left(y_r^i - y_r^1\right)\sum_{i=1}^{M}\underline{f}^i\left(y_r^M - y_r^i\right)}{\sum_{i=1}^{M}\bar{f}^i\left(y_r^i - y_r^1\right) + \sum_{i=1}^{M}\underline{f}^i\left(y_r^M - y_r^i\right)} \times\frac{\sum_{i=1}^{M}\left(\bar{f}^i - \underline{f}^i\right)}{\sum_{i=1}^{M}\bar{f}^i\sum_{i=1}^{M}\underline{f}^i}\right] \tag{102}$$

with the right-hand side of the inequality defined as $\bar{y}_r(\mathbf{x})$. Because $y_l(\mathbf{x}) \le y_r(\mathbf{x})$, $[\underline{y}_l(\mathbf{x}), \bar{y}_r(\mathbf{x})]$ is a valid outer-bound set for $[y_l(\mathbf{x}), y_r(\mathbf{x})]$.

B. Proof of Corollary 1

$\delta(\mathbf{x})$ is the difference between the defuzzified outputs of the type-reduced set $[y_l(\mathbf{x}), y_r(\mathbf{x})]$ and its approximation set $[\underline{y}_l(\mathbf{x}) + \bar{y}_l(\mathbf{x})/2, \underline{y}_r(\mathbf{x}) + \bar{y}_r(\mathbf{x})/2]$ and it can be written as follows:

$$\delta(\mathbf{x}) = \left|\frac{1}{2}\left[\frac{\underline{y}_l(\mathbf{x}) + \bar{y}_l(\mathbf{x})}{2} + \frac{\underline{y}_r(\mathbf{x}) + \bar{y}_r(\mathbf{x})}{2}\right] - \frac{y_l(\mathbf{x}) + y_r(\mathbf{x})}{2}\right|$$

$$
\begin{aligned}
&= \frac{1}{2}\left|\frac{\bar{y}_l(\mathbf{x})+\underline{y}_l(\mathbf{x})}{2} - y_l(\mathbf{x}) + \frac{\bar{y}_r(\mathbf{x})+\underline{y}_r(\mathbf{x})}{2} - y_r(\mathbf{x})\right| \\
&= \frac{1}{2}\left|\left[\frac{\bar{y}_l(\mathbf{x})-y_l(\mathbf{x})}{2} - \frac{y_l(\mathbf{x})-\underline{y}_l(\mathbf{x})}{2}\right]\right. \\
&\qquad \left. + \left[\frac{\bar{y}_r(\mathbf{x})-y_r(\mathbf{x})}{2} - \frac{y_r(\mathbf{x})-\underline{y}_r(\mathbf{x})}{2}\right]\right|. \qquad (103)
\end{aligned}
$$

This equation is the starting point for the derivation of $\bar{\delta}(\mathbf{x})$.

1) From Fig. 1, it is clear that

$$0 \le \frac{\bar{y}_l(\mathbf{x})-y_l(\mathbf{x})}{2} \le \frac{\bar{y}_l(\mathbf{x})-\underline{y}_l(\mathbf{x})}{2} \qquad (104)$$

and

$$0 \le \frac{y_l(\mathbf{x})-\underline{y}_l(\mathbf{x})}{2} \le \frac{\bar{y}_l(\mathbf{x})-\underline{y}_l(\mathbf{x})}{2} \qquad (105)$$

i.e.,

$$-\frac{\bar{y}_l(\mathbf{x})-\underline{y}_l(\mathbf{x})}{2} \le -\frac{y_l(\mathbf{x})-\underline{y}_l(\mathbf{x})}{2} \le 0. \qquad (106)$$

Adding (104) and (106) together, we get

$$
\begin{aligned}
-\frac{\bar{y}_l(\mathbf{x})-\underline{y}_l(\mathbf{x})}{2} &\le \frac{\bar{y}_l(\mathbf{x})-y_l(\mathbf{x})}{2} - \frac{y_l(\mathbf{x})-\underline{y}_l(\mathbf{x})}{2} \\
&\le \frac{\bar{y}_l(\mathbf{x})-\underline{y}_l(\mathbf{x})}{2}. \qquad (107)
\end{aligned}
$$

Hence (see Fig. 1)

$$
\begin{aligned}
\delta_l(\mathbf{x}) &\equiv \left|y_l(\mathbf{x}) - \frac{\bar{y}_l(\mathbf{x})+\underline{y}_l(\mathbf{x})}{2}\right| \\
&= \left|\frac{\bar{y}_l(\mathbf{x})-y_l(\mathbf{x})}{2} - \frac{y_l(\mathbf{x})-\underline{y}_l(\mathbf{x})}{2}\right| \\
&\le \frac{\bar{y}_l(\mathbf{x})-\underline{y}_l(\mathbf{x})}{2}. \qquad (108)
\end{aligned}
$$

2) Similarly (see Fig. 1)

$$\delta_r(\mathbf{x}) \equiv \left|y_r(\mathbf{x}) - \frac{\bar{y}_r(\mathbf{x})+\underline{y}_r(\mathbf{x})}{2}\right| \le \frac{\bar{y}_r(\mathbf{x})-\underline{y}_r(\mathbf{x})}{2}. \qquad (109)$$

3) Combining the aforementioned results, we find that

$$
\begin{aligned}
\delta(\mathbf{x}) &= \frac{1}{2}\left|\frac{\bar{y}_l(\mathbf{x})+\underline{y}_l(\mathbf{x})}{2} - y_l(\mathbf{x}) + \frac{\bar{y}_r(\mathbf{x})+\underline{y}_r(\mathbf{x})}{2} - y_r(\mathbf{x})\right| \\
&\le \frac{1}{2}\left[\left|\frac{\bar{y}_l(\mathbf{x})+\underline{y}_l(\mathbf{x})}{2} - y_l(\mathbf{x})\right| \right. \\
&\qquad \left. + \left|\frac{\bar{y}_r(\mathbf{x})+\underline{y}_r(\mathbf{x})}{2} - y_r(\mathbf{x})\right|\right] \\
&= \frac{1}{2}\left[\delta_l(\mathbf{x}) + \delta_r(\mathbf{x})\right] \\
&\le \frac{1}{4}\left[\left(\bar{y}_l(\mathbf{x}) - \underline{y}_l(\mathbf{x})\right) + \left(\bar{y}_r(\mathbf{x}) - \underline{y}_r(\mathbf{x})\right)\right]. \qquad (110)
\end{aligned}
$$

REFERENCES

[1] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York: Wiley, 1991.

[2] N. N. Karnik and J. M. Mendel, "Centroid of a type-2 fuzzy set," *Inform. Sci.*, vol. 132, pp. 195–220, 2001.

[3] N. N. Karnik, J. M. Mendel, and Q. Liang, "Type-2 fuzzy logic systems," *IEEE Trans. Fuzzy Syst.*, vol. 7, pp. 643–658, Dec. 1999.

[4] Q. Liang, N. N. Karnik, and J. M. Mendel, "Connection admission control in ATM networks using survey-based type-2 fuzzy logic systems," *IEEE Trans. Syst., Man, Cybern. C*, vol. 30, pp. 329–339, Aug. 2000.

[5] Q. Liang and J. M. Mendel, "Interval type-2 fuzzy logic systems: Theory and design," *IEEE Trans. Fuzzy Syst.*, vol. 8, pp. 535–550, Oct. 2000.

[6] ——, "Equalization of nonlinear time-varying channels using type-2 fuzzy adaptive filters," *IEEE Trans. Fuzzy Syst.*, vol. 8, no. 5, pp. 551–563, Oct. 2000.

[7] ——, "Overcoming time-varying co-channel interference using type-2 fuzzy adaptive filters," *IEEE Trans. Circuits Syst. II*, vol. 47, pp. 1419–1429, Dec. 2000.

[8] ——, "MPEG VBR video traffic modeling and classification using fuzzy techniques," *IEEE Trans. Fuzzy Syst.*, vol. 9, no. 1, pp. 183–193, Feb. 2001.

[9] J. M. Mendel, "Computing with words when words can mean different things to different people," presented at the Int. ICSC Congr. Computation Intelligence: Methods Application, 3rd Annual Symp. Fuzzy Logic Application, Rochester, NY, June 1999.

[10] ——, *Uncertain Rule-Based Fuzzy Logic Systems: Introduction and New Directions*. Upper Saddle River, NJ: Prentice-Hall, 2001.

[11] H. Wu and J. M. Mendel, Comput. Complex. Type-Reduction Interval Type-2 Fuzzy Logic Syst., submitted for publication.

[12] L. A. Zadeh, "The concept of a linguistic variable and its application to approximate reasoning-1," *Inform. Sci.*, vol. 8, pp. 199–249, 1975.

Hongwei Wu received the B.Eng. degree in automatic control and the M.Eng. degree in pattern recognition and intelligent systems, from Tsinghua University, Beijing, China, in 1997 and 1999, respectivley. She is currently working toward the Ph.D. degree at the Signal and Image Processing Institute, Department of Electrical Engineering, University of Southern California, Los Angeles

Her research interests include fuzzy logic systems, acoustic signal processing, and pattern recognition.

Jerry M. Mendel (S'59–M'61–SM'72–F'78) received the Ph.D. degree in electrical engineering from the Polytechnic Institute of Brooklyn, Brooklyn, NY, in 1963.

Currently, he is Professor of Electrical Engineering and Associate Director for Education of the Integrated Media Systems Center at the University of Southern California, Los Angeles, where he has been since 1974. He has published over 400 technical papers and is the author and/or editor of eight books, including *Uncertain Rule-based Fuzzy Logic Systems: Introduction and New Directions* (Upper Saddle River, NJ: Prentice-Hall, 2001). His present research interests include type-2 fuzzy logic systems and their applications to a wide range of problems, including computing with words.

Dr. Mendel is a Distinguished Member of the IEEE Control Systems Society, and was was President of that society in 1986. Among his awards are the 1983 Best Transactions Paper Award of the IEEE Geoscience and Remote Sensing Society, the 1992 Signal Processing Society Paper Award, the 2002 TRANSACTIONS ON FUZZY SYSTEMS Outstanding Paper Award, a 1984 IEEE Centennial Medal, and an IEEE Third Millenium Medal.