Computing Derivatives in Interval Type-2 Fuzzy Logic Systems

Jerry M. Mendel, *Life Fellow, IEEE*

***Abstract*—This paper makes type-2 fuzzy logic systems much more accessible to fuzzy logic system designers, because it provides mathematical formulas and computational flowcharts for computing the derivatives that are needed to implement steepest-descent parameter tuning algorithms for such systems. It explains why computing such derivatives is much more challenging than it is for a type-1 fuzzy logic system. It provides derivative calculations that are applicable to any kind of type-2 membership functions, since the calculations are performed without prespecifying the nature of those membership functions. Some calculations are then illustrated for specific type-2 membership functions.**

***Index Terms*—Derivations, fuzzy logic system, type-2 fuzzy logic system.**

I. Introduction

A TYPE-2 fuzzy logic system (FLS) lets us directly model (and, subsequently, minimize the effects of) a variety of uncertainties[1] that cannot be directly modeled using a type-1 FLS. The price paid for being able to do this is somewhat greater complexity for a type-2 FLS than for a type-1 FLS; but, if one is unable to achieve satisfactory performance—in the face of uncertainties—using a type-1 FLS, then this may be a small price to pay for the improved performance. Of course, to achieve the improved performance one must first be able to *design* a type-2 FLS. Although there are different approaches to doing this, the most popular one to-date uses steepest descent algorithms (also referred to as *back-propagation algorithms*) for adjusting all design parameters, and such algorithms require the computation of first-derivatives of an objective function with respect to each and every design parameter. The purpose of this paper is to provide such first-derivative formulas since they do not appear in the existing literature. Doing this will make type-2 FLSs much more accessible to FLS designers. To begin, we review the essence of a type-2 FLS.

A type-2 FLS (just as a type-1 FLS) contains four components: rules, fuzzifier, inference engine and output-processor (Fig. 1). During the operation of a type-2 FLS, measurements activate the fuzzifier, inference engine and output processor blocks in that order. The output processor contains two components: *type-reduction* and *defuzzification*. When arbitrary type-2 fuzzy sets[2] are used a type-2 FLS is computationally prohibitive. On the other hand, when all type-2 fuzzy sets are modeled as interval sets, then we obtain an *interval type-2 FLS*, and such FLSs are very practical.[3]

Depending upon the way in which input measurements ($\mathbf{x}$) to the FLS are fuzzified, either as singletons, type-1 fuzzy numbers or type-2 fuzzy numbers, three kinds of interval type-2 FLSs are possible: interval singleton type-2 FLS, interval type-1 nonsingleton type-2 FLS, and interval type-2 nonsingleton type-2 FLS. In the main body of this paper, we focus on an interval singleton type-2 FLS, but provide the extensions of our results to an interval type-2 nonsingleton type-2 FLS (which contains an interval type-1 nonsingleton type-2 FLS as a special case) in Appendix B.

The inference engine first produces a *firing set*, which is then used to produce an output consequent set membership function (MF) for each fired rule, which can then be used to produce a MF for all (combined) fired rules. For an interval singleton type-2 FLS, it is possible to obtain *closed-form formulas* for all of these quantities, and these results are given in Section A.2 of Appendix A; they are obtained by using well-known closed-form formulas for the join and meet of interval sets (e.g., [1] and [6]).

The type-reducer leads to a *type-reduced set* that provides an interval of uncertainty for the output of a type-2 FLS that is analogous to a confidence interval that provides an interval of uncertainty for a probabilistic system. The more uncertainties that occur in a type-2 FLS, which translate into more uncertainties about its MFs, the larger will be the type-reduced set, and *vice versa*. Regardless of which type-reduction method[4] we choose, the type-reduced set for an interval type-2 FLS is an *interval type-1 set*, and its two end-points can be computed using an exact iterative method developed by Karnik and Mendel [2] whose steps are listed in Section A.3 of Appendix A. Because the type-reduced set is an interval set, its defuzzified value is simply the average value of its two end-points. A formula for the defuzzified output is given in Section A.4 of Appendix A. Using the formulas given in Appendix A, it is possible to compute the input-output relation, $y = f_{s2}(\mathbf{x})$, of an interval singleton type-2 FLS, and these formulas are the starting point for the design of such a FLS.

It is well known that a type-1 FLS is characterized by a *fuzzy basis function (FBF) expansion* (e.g., [8]), and that such an expansion is not only useful for computing the output of that FLS but is also used during its design, especially as the starting point

Manuscript received January 15, 2002; revised June 24, 2003.
The author is 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.2003.822681

[1]Because these uncertainties have been enumerated many times before, we refer the reader to [6, p. 68] for a list of them, as well as for discussions about them.

[2]See Section A.1 of Appendix A for some important definitions about type-2 fuzzy sets.

[3]See [7] for many reasons supporting the use of interval type-2 sets.

[4]Type-reduction is briefly reviewed in Section A.3 of Appendix A.

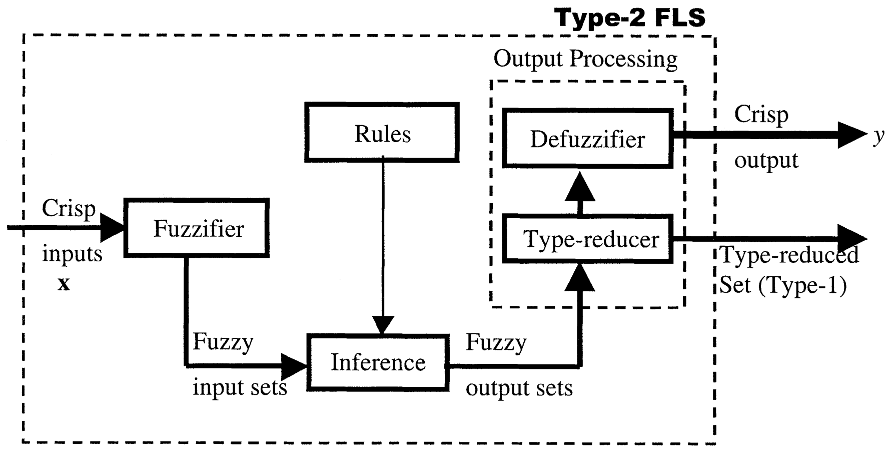

Fig. 1. Type-2 FLS.

for computing derivatives of an objective function with respect to MF parameters. An interval singleton type-2 FLS is characterized by *two* fuzzy basis function (FBF) expansions [5], one associated with the left end-point of the type-reduced set, and the other associated with the right end-point of the type-reduced set. Unlike the FBF expansion for a type-1 FLS, the FBF expansions for an interval singleton type-2 FLS cannot be used to actually compute the left and right end-points of the type-reduced set, because the latter are in terms of two crossover points, $L(\mathbf{x})$ and $R(\mathbf{x})$ (L and R for short), that are computed using the Karnik–Mendel iterative procedures. By the end of those procedures not only are L and R computed but so also are the left and right end-points of the type-reduced set. Interestingly enough, the formulas for the two FBF expansions can however be used during the design of the FLS, as we explain below.

By "design" we mean specify or optimize the parameters that characterize the interval type-2 FLS. A type-2 FLS design method is associated with the following design problem.

> We are given a collection of N input-output numerical data *training* pairs, $(\mathbf{x}^{(1)} : y^{(1)}), (\mathbf{x}^{(2)} : y^{(2)}), \ldots, (\mathbf{x}^{(N)} : y^{(N)})$, where $\mathbf{x}$ is the vector input and y is the scalar output of an interval singleton type-2 FLS. Our goal is to completely specify this type-2 FLS using the training data.

A design method establishes how to specify all the parameters of the antecedent and consequent membership functions using the training pairs $(\mathbf{x}^{(1)} : y^{(1)}), (\mathbf{x}^{(2)} : y^{(2)}), \ldots, (\mathbf{x}^{(N)} : y^{(N)})$. The most popular design method—the back-propagation method—is one in which all MF parameters are tuned using a steepest descent algorithm whose general structure is

$$\theta(n+1) = \theta(n) - \alpha_\theta \left.\frac{\partial e^{(t)}}{\partial \theta}\right|_n \tag{1}$$

where θ denotes any one of the FLS design parameters

$$e^{(t)} = \frac{1}{2}\left[f_{s2}\left(\mathbf{x}^{(t)}\right) - y^{(t)}\right]^2, \qquad t = 1, \ldots, N \tag{2}$$

and $[\,]|_n$ indicates that after taking the derivative of $e^{(t)}$ with respect to a specific θ we must replace all remaining θ values by $\theta(n)$. The challenge to developing easy-to-use steepest descent algorithms is to establish formulas for the derivatives $\partial e^{(t)}/\partial\theta$. Generally, it is much more complicated to compute such derivatives for an interval type-2 FLS than it is for a type-1 FLS, because of the following.

- In an interval singleton type-2 FLS the design parameters appear in[5] *upper* and *lower* MFs, whereas in a singleton type-1 FLS they appear in a single MF.
- In an interval singleton type-2 FLS, type-reduction establishes the two parameters $L(\mathbf{x})$ and $R(\mathbf{x})$, which in turn establish the upper and lower firing-interval MFs that are used to compute the left and right end-points of the type-reduced set [see (A-16) and (A-17)]. There is no type-reduction in a type-1 FLS.

In the rest of this paper we establish mathematical formulas to compute the derivatives $\partial e^{(t)}/\partial\theta$. The derivations of these formulas can be approached in different ways, e.g. choose a type-2 fuzzy set's membership function footprint of uncertainty[6] (FOU) as soon as possible or defer the choice of a FOU for as long as possible. In this paper, we take the latter approach, because by doing so our results are applicable to any kind of FOU.

Section II provides some fundamental assumptions; Section III provides general formulas for the right and left end-points of the type-reduced set; Section IV provides formulas for derivatives of $e^{(t)}$ with respect to antecedent MF parameters; Section V provides formulas for derivatives of $e^{(t)}$ with respect to consequent MF parameters; Section VI provides an example; Section VII provides conclusions. Note that the formulas in Sections III–V are independent of the choices made for the type-2 antecedent and consequent MFs. Finally, Appendix A has important background material about type-2 fuzzy sets, and fuzzy inference engine, type-reduction and defuzzification for interval singleton type-2 FLSs; and, Appendix B presents derivative formulas for interval type-2 nonsingleton type-2 FLSs.

[5]Upper and lower MFs are defined in Definition A-8.
[6]The FOU is defined in Definition A-6.

TABLE I
RULE-ORDERED AND RULE-REORDERED QUANTITIES

Calculations	**Rule-Ordered**	**Rule-Reordered**
Left-end calculations[b] y_l	$\mathbf{z}_l = (z_l^1, \ldots, z_l^M)^T$	$\mathbf{y}_l = (y_l^1, \ldots, y_l^M)^T$
	$\mathbf{y}_l = \mathbf{Q}_l \mathbf{z}_l$ $\mathbf{Q}_l$ is an $M \times M$ permutation matrix[a]	
	$y_l = \dfrac{\sum_{i=1}^{L} \bar{f}^i z_l^i + \sum_{j=L+1}^{M} \underline{f}^j z_l^j}{\sum_{i=1}^{L} \bar{f}^i + \sum_{j=L+1}^{M} \underline{f}^j}$	$y_l = \dfrac{\sum_{i=1}^{L} \bar{g}^i y_l^i + \sum_{j=L+1}^{M} \underline{g}^j y_l^j}{\sum_{i=1}^{L} \bar{g}^i + \sum_{j=L+1}^{M} \underline{g}^j}$
Right-end calculations[b] y_r	$\mathbf{z}_r = (z_r^1, \ldots, z_r^M)^T$	$\mathbf{y}_r = (y_r^1, \ldots, y_r^M)^T$
	$\mathbf{y}_r = \mathbf{Q}_r \mathbf{z}_r$ $\mathbf{Q}_r$ is an $M \times M$ permutation matrix[a]	
	$y_r = \dfrac{\sum_{i=1}^{R} \underline{f}^i z_r^i + \sum_{j=R+1}^{M} \bar{f}^j z_r^j}{\sum_{i=1}^{R} \underline{f}^i + \sum_{j=R+1}^{M} \bar{f}^j}$	$y_r = \dfrac{\sum_{i=1}^{R} \underline{g}^i y_r^i + \sum_{j=R+1}^{M} \bar{g}^j y_r^j}{\sum_{i=1}^{R} \underline{g}^i + \sum_{j=R+1}^{M} \bar{g}^j}$

a. There are many kinds of permutation matrices. The kind we use has columns that are *elementary vectors* (i.e., vectors all of whose elements are zero except one element that is one) and these vectors are arranged—permuted—so as to move elements in the original matrix or vector to new locations in the transformed matrix or vector.

b. Note that in [2] and [5], $\underline{f}^l(\mathbf{x})$ and $\bar{f}^l(\mathbf{x})$ are used to denote both the rule-ordered and rule reordered lower and upper firing intervals. This does not cause any problems, unless—as in the present paper—formulas for y_l and y_r are used for further computations.

II. ASSUMPTIONS

We make four fundamental assumptions.

1) Parameters to be tuned are different for each rule and for each antecedent and consequent, i.e. *no parameters are shared across rules or MFs*.
2) Formulas for antecedent and consequent MFs are not specified ahead of time.
3) Derivatives needed for a steepest descent tuning-algorithm are to be computed by means of mathematical formulas.
4) Center-of-sets type-reduction is used.

If some parameters are shared across rules or MFs then at some point in our analyzes below a detour must be taken. We will indicate precisely where this occurs. By not specifying formulas for antecedent and consequent MFs ahead of time, our results will be as general as possible. If mathematical formulas for derivatives cannot be obtained, it may still be possible to determine derivatives numerically using perturbation techniques. We do not cover such techniques in this paper because the kinds of primary MFs that one usually chooses can be described mathematically, e.g., triangles, trapezoids, Gaussians, etc. We choose center-of-sets type-reduction because there is an explicit appearance of antecedent and consequent MF parameters for it. The same is true for height type-reduction but is not true for centroid or center-of-sums type-reduction (see Table A-I).

III. GENERAL EXPRESSIONS FOR y_l AND y_r

Although, as discussed in Section I, we always compute y_l and y_r using the Karnik–Mendel iterative procedures, we use formulas for y_l and y_r to compute derivatives that are needed in the back-propagation update algorithms. Such formulas can be deduced from step 4 of the iterative procedure (Section A.3), the paragraph just below that procedure, and (A-16) and (A-17), and are

$$y_l = \frac{\sum_{i=1}^{M} g_l^i y_l^i}{\sum_{i=1}^{M} g_l^i} = \frac{\sum_{i=1}^{L} \bar{g}^i y_l^i + \sum_{j=L+1}^{M} \underline{g}^j y_l^j}{\sum_{i=1}^{L} \bar{g}^i + \sum_{j=L+1}^{M} \underline{g}^j} = y_l(\bar{g}^1, \ldots, \bar{g}^L, \underline{g}^{L+1}, \ldots, \underline{g}^M, y_l^1, \ldots, y_l^M) \qquad (3)$$

and

$$y_r = \frac{\sum_{i=1}^{M} g_r^i y_r^i}{\sum_{i=1}^{M} g_r^i} = \frac{\sum_{i=1}^{R} \underline{g}^i y_r^i + \sum_{j=R+1}^{M} \bar{g}^j y_r^j}{\sum_{i=1}^{R} \underline{g}^i + \sum_{j=R+1}^{M} \bar{g}^j} = y_r(\underline{g}^1, \ldots, \underline{g}^R, \bar{g}^{R+1}, \ldots, \bar{g}^M, y_r^1, \ldots, y_r^M). \qquad (4)$$

These formulas cannot be used as is because the f_l^i, y_l^i, f_r^i, and y_r^i have been *reordered* during step-1 of the two iterative procedures used to compute y_l and y_r. In order to compute derivatives of y_l and y_r with respect to MF parameters, we need to know exactly where specific antecedent and consequent MF parameters are located, and this is very difficult to ascertain when y_l and y_r are not in rule-ordered format. So, our first task is to re-express

TABLE II
FORMULAS USED TO CALCULATE y_l AND y_r IN RULE-ORDERED FORMAT[a]

y_l **calculations**	y_r **calculations**
$\underline{\mathbf{f}} = \left(\underline{f}^1, \underline{f}^2, \ldots, \underline{f}^M\right)^T$	
$\bar{\mathbf{f}} = \left(\bar{f}^1, \bar{f}^2, \ldots, \bar{f}^M\right)^T$	
$\underline{\mathbf{g}} = \left(\underline{g}^1, \underline{g}^2, \ldots, \underline{g}^M\right)^T \equiv \mathbf{Q}_l \underline{\mathbf{f}}$	$\underline{\mathbf{g}} = \left(\underline{g}^1, \underline{g}^2, \ldots, \underline{g}^M\right)^T \equiv \mathbf{Q}_r \underline{\mathbf{f}}$
$\bar{\mathbf{g}} = \left(\bar{g}^1, \bar{g}^2, \ldots, \bar{g}^M\right)^T \equiv \mathbf{Q}_l \bar{\mathbf{f}}$	$\bar{\mathbf{g}} = \left(\bar{g}^1, \bar{g}^2, \ldots, \bar{g}^M\right)^T \equiv \mathbf{Q}_r \bar{\mathbf{f}}$
$\mathbf{E}_1 \equiv \left(\mathbf{e}_1 \mid \mathbf{e}_2 \mid \cdots \mid \mathbf{e}_L \mid \mathbf{0} \mid \cdots \mid \mathbf{0}\right)^T \quad L \times M$	$\mathbf{E}_3 \equiv \left(\mathbf{e}_1 \mid \mathbf{e}_2 \mid \cdots \mid \mathbf{e}_R \mid \mathbf{0} \mid \cdots \mid \mathbf{0}\right)^T \quad R \times M$
$\mathbf{e}_i = L \times 1$ elementary vector	$\mathbf{e}_i = R \times 1$ is the ith elementary vector
$\mathbf{E}_2 \equiv \left(\mathbf{0} \mid \cdots \mid \mathbf{0} \mid \varepsilon_1 \mid \varepsilon_2 \mid \cdots \mid \varepsilon_{M-L}\right)^T \quad (M-L) \times 1$	$\mathbf{E}_4 \equiv \left(\mathbf{0} \mid \cdots \mid \mathbf{0} \mid \varepsilon_1 \mid \varepsilon_2 \mid \cdots \mid \varepsilon_{M-R}\right)^T \quad (M-R) \times 1$
$\varepsilon_i = (M-L) \times 1$ elementary vector	$\varepsilon_i = (M-R) \times 1$ is the ith elementary vector
$\mathbf{M}_{l1} \equiv \mathbf{Q}_l^T \mathbf{E}_1^T \mathbf{E}_1 \mathbf{Q}_l \quad M \times M$	$\mathbf{M}_{r1} \equiv \mathbf{Q}_r^T \mathbf{E}_3^T \mathbf{E}_3 \mathbf{Q}_r \quad M \times M$
$\mathbf{M}_{l2} \equiv \mathbf{Q}_l^T \mathbf{E}_2^T \mathbf{E}_2 \mathbf{Q}_l \quad M \times M$	$\mathbf{M}_{r2} \equiv \mathbf{Q}_r^T \mathbf{E}_4^T \mathbf{E}_4 \mathbf{Q}_r \quad M \times M$
$\mathbf{r}_l \equiv (\underbrace{1, 1, \cdots, 1}_{L}, 0, \cdots, 0)^T \quad M \times 1$	$\mathbf{r}_r \equiv (\underbrace{1, 1, \cdots 1}_{R}, 0, \cdots, 0)^T \quad M \times 1$
$\mathbf{s}_l \equiv (0, 0, \cdots 0, \overbrace{1, \cdots, 1}^{M-L})^T \quad M \times 1$	$\mathbf{s}_r \equiv (0, 0, \cdots 0, \overbrace{1, \cdots, 1}^{M-R})^T \quad M \times 1$
$\mathbf{a}_l \equiv \mathbf{M}_{l1} \mathbf{z}_l \quad M \times 1$	$\mathbf{a}_r = \mathbf{M}_{r1} \mathbf{z}_r \quad M \times 1$
$\mathbf{b}_l \equiv \mathbf{M}_{l2} \mathbf{z}_l \quad M \times 1$	$\mathbf{b}_r = \mathbf{M}_{r2} \mathbf{z}_r \quad M \times 1$
$\mathbf{c}_l^T \equiv \mathbf{r}_l^T \mathbf{Q}_l \quad 1 \times M$	$\mathbf{c}_r^T = \mathbf{r}_r^T \mathbf{Q}_r \quad 1 \times M$
$\mathbf{d}_l^T \equiv \mathbf{s}_l^T \mathbf{Q}_l \quad 1 \times M$	$\mathbf{d}_r^T = \mathbf{s}_r^T \mathbf{Q}_r \quad 1 \times M$

a. Using the same symbols, $\underline{\mathbf{g}}$ and $\bar{\mathbf{g}}$, for both y_l and y_r calculations will not cause any problems in later calculations.

(3) and (4) in rule-ordered format. Along the way, we shall also remove the explicit dependence of y_l on L and y_r on R.

Rule-ordered firing intervals are denoted as $F^l(\mathbf{x}')$. We have labeled the *rule reordered firing intervals* $G^l(\mathbf{x}')$, i.e.,

$$F^l(\mathbf{x}') = [\underline{f}^l, \bar{f}^l] \tag{5}$$

$$G^l(\mathbf{x}') = [\underline{g}^l, \bar{g}^l]. \tag{6}$$

In (A-15), observe that $y^i = [y_l^i, y_r^i]$ $(i = 1, \ldots, M)$, and that in the two Karnik–Mendel procedures it is the y_l^i $(i = 1, \ldots, M)$ that are reordered[7] in order to compute y_l, whereas it is the y_r^i $(i = 1, \ldots, M)$ that are reordered to compute y_r. In this paper, we continue to let y_l^i and y_r^i denote *rule-reordered* values; however, we introduce z_l^i and z_r^i to denote their *rule-ordered* counterparts. Table I summarizes the rule-ordered and rule-unordered quantities that we will need in the rest of this paper. The question that we address next is how do we go from the rule-reordered versions of y_l and y_r to the rule-ordered versions?

A. y_l Re-Expressed in Rule-Ordered Format

We want to re-express y_l, given in rule-reordered form (Table I), in terms of a rule-ordered quantities, i.e., in terms of $\bar{f}^i$, $\underline{f}^j$, z_l^i and z_l^j. To begin, we define a collection of vectors and matrices that are summarized in Table II.

[7]For other type-reduction methods, what gets reordered can be deduced from the column in Table A-I labeled "y_l^i and y_r^i defined."

Fact 1: y_l can be re-expressed in terms of rule-ordered quantities as

$$\begin{aligned} y_l &= \frac{\bar{\mathbf{f}}^T \mathbf{M}_{l1} \mathbf{z}_l + \underline{\mathbf{f}}^T \mathbf{M}_{l2} \mathbf{z}_l}{\mathbf{r}_l^T \mathbf{Q}_l \bar{\mathbf{f}} + \mathbf{s}_l^T \mathbf{Q}_l \underline{\mathbf{f}}} \\ &= \frac{\bar{\mathbf{f}}^T \mathbf{a}_l + \underline{\mathbf{f}}^T \mathbf{b}_l}{\bar{\mathbf{f}}^T \mathbf{c}_l + \underline{\mathbf{f}}^T \mathbf{d}_l} \\ &= \frac{\sum_{i=1}^M a_{l,i} \bar{f}^i + \sum_{j=1}^M b_{l,j} \underline{f}^j}{\sum_{i=1}^M c_{l,i} \bar{f}^i + \sum_{j=1}^M d_{l,j} \underline{f}^j}. \end{aligned} \tag{7}$$

Proof: We must re-express the four sums that appear in the rule-reordered version of y_l, given in Table I, i.e.,[8]

$$\begin{aligned} \sum_{i=1}^{L} \bar{g}^i y_l^i &= (\mathbf{E}_1 \bar{\mathbf{g}})^T (\mathbf{E}_1 \mathbf{y}_l) \\ &= (\mathbf{E}_1 \mathbf{Q}_l \bar{\mathbf{f}})^T (\mathbf{E}_1 \mathbf{Q}_l \mathbf{z}_l) \\ &= \bar{\mathbf{f}}^T \mathbf{Q}_l^T \mathbf{E}_1^T \mathbf{E}_1 \mathbf{Q}_l \mathbf{z}_l \\ &= \bar{\mathbf{f}}^T \mathbf{M}_{l1} \mathbf{z}_l \end{aligned} \tag{8}$$

$$\begin{aligned} \sum_{j=L+1}^{M} \underline{g}^j y_l^j &= (\mathbf{E}_2 \underline{\mathbf{g}})^T (\mathbf{E}_2 \mathbf{y}_l) \\ &= (\mathbf{E}_2 \mathbf{Q}_l \underline{\mathbf{f}})^T (\mathbf{E}_2 \mathbf{Q}_l \mathbf{z}_l) \\ &= \underline{\mathbf{f}}^T \mathbf{Q}_l^T \mathbf{E}_2^T \mathbf{E}_2 \mathbf{Q}_l \mathbf{z}_l \\ &= \underline{\mathbf{f}}^T \mathbf{M}_{l2} \mathbf{z}_l \end{aligned} \tag{9}$$

[8]Note, for example, that $(\mathbf{E}_1 \bar{\mathbf{g}})^T = \bar{\mathbf{g}}^T \mathbf{E}_1^T = (\bar{g}^1, \bar{g}^2, \ldots, \bar{g}^M)(\mathbf{e}_1|\mathbf{e}_2|\cdots|\mathbf{e}_L|\mathbf{0}|\cdots|\mathbf{0}) = (\bar{g}^1, \bar{g}^2, \ldots, \bar{g}^L)$ and $\mathbf{E}_1 \mathbf{y}_l = (\mathbf{e}_1|\mathbf{e}_2|\cdots|\mathbf{e}_L|\mathbf{0}|\cdots|\mathbf{0})^T \times (y_l^1, y_l^2, \ldots, y_l^M)^T = (y_l^1, y_l^2, \ldots, y_l^L)^T$, so that $\sum_{i=1}^L \bar{g}^i y_l^i = (\mathbf{E}_1 \bar{\mathbf{g}})^T (\mathbf{E}_1 \mathbf{y}_l)$.

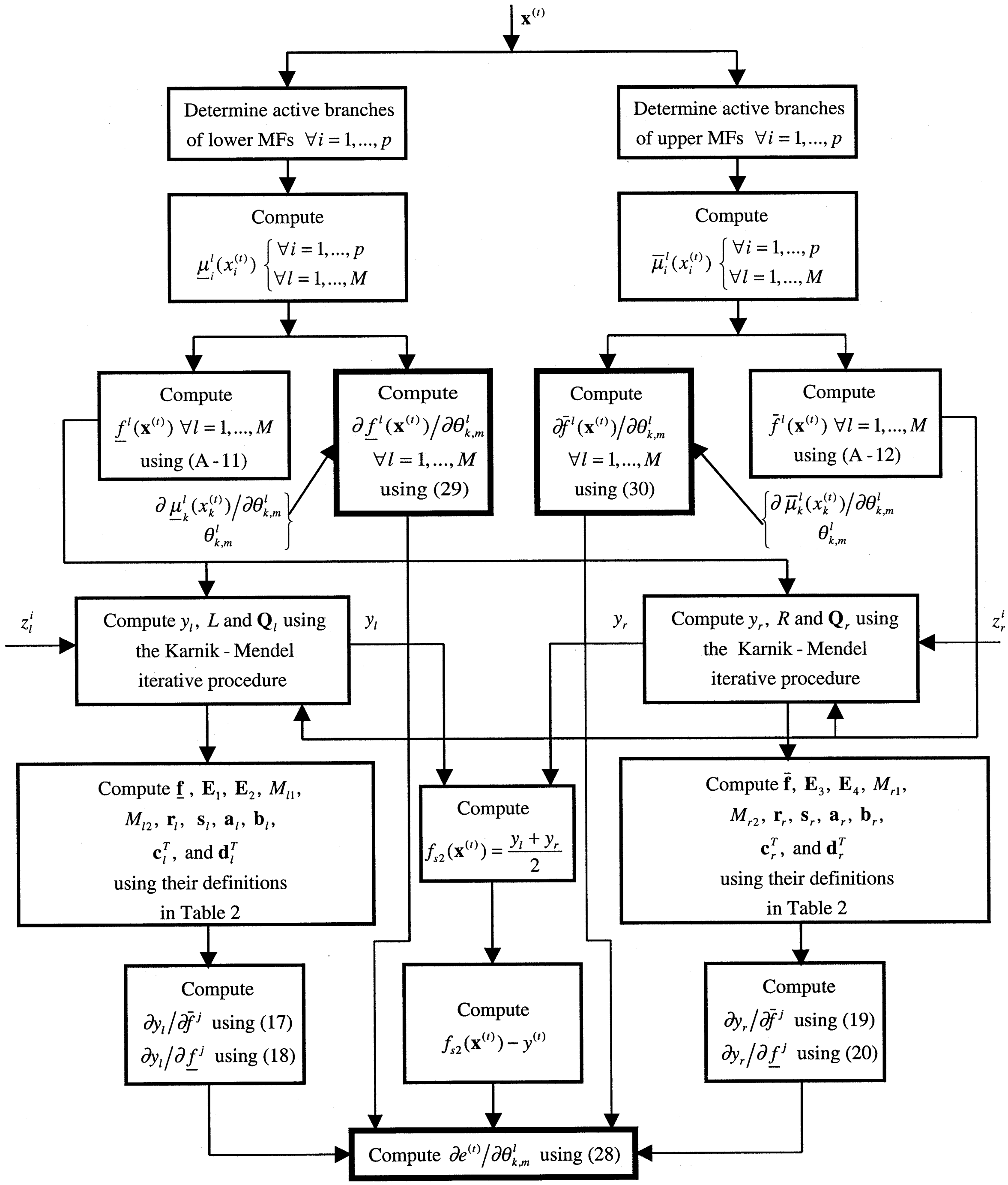

Fig. 2. Computational flow-chart: derivatives for antecedent parameters.

$$\sum_{i=1}^{L} \bar{g}^i = \mathbf{r}_l^T \bar{\mathbf{g}} = \mathbf{r}_l^T \mathbf{Q}_l \bar{\mathbf{f}} \tag{10}$$

$$\sum_{j=L+1}^{M} \underline{g}^j = \mathbf{s}_l^T \underline{\mathbf{g}} = \mathbf{s}_l^T \mathbf{Q}_l \underline{\mathbf{f}}. \tag{11}$$

Substitute (8)–(11) into $y_l = (\sum_{i=1}^{L} \bar{g}^i y_l^i + \sum_{j=L+1}^{M} \underline{g}^j y_l^j)/(\sum_{i=1}^{L} \bar{g}^i + \sum_{j=L+1}^{M} \underline{g}^j)$ to obtain the first term on the right-hand side of (7). The second term appearing on the right-hand side of (7) is obtained by using the quantities $\mathbf{a}_l$, $\mathbf{b}_l$, $\mathbf{c}_l^T$, and $\mathbf{d}_l^T$, that are defined in the last four rows of Table II, in the first term on the right-hand side of (7). The last term on the right-hand side of (7) is just an expanded version of the second term. ■

Observe that (7) involves the *entire* $\bar{\mathbf{f}}$ and $\underline{\mathbf{f}}$ vectors and the entire $\mathbf{z}_l$ vector. This is good because we can then take the derivatives of y_l with respect to any element in $\bar{\mathbf{f}}$ or $\underline{\mathbf{f}}$ without having to worry ahead of time whether or not it actually appears in y_l.

TABLE III
UPPER AND LOWER MF PARAMETER DEPENDENCIES (THE x_k VALUE ESTABLISHES THE *ACTIVE BRANCH*)

x_k	$\bar{\mu}_k^l(x_k)$	$\underline{\mu}_k^l(x_k)$
$x_k < m_{k1}^l$	$m_{k1}^l,\ \sigma_k^l$	$m_{k2}^l,\ \sigma_k^l$
$m_{k1}^l \le x_k \le (m_{k1}^l + m_{k2}^l)/2$	None	$m_{k2}^l,\ \sigma_k^l$
$(m_{k1}^l + m_{k2}^l)/2 \le x_k \le m_{k2}^l$	None	$m_{k1}^l,\ \sigma_k^l$
$x_k > m_{k2}^l$	$m_{k2}^l,\ \sigma_k^l$	$m_{k1}^l,\ \sigma_k^l$

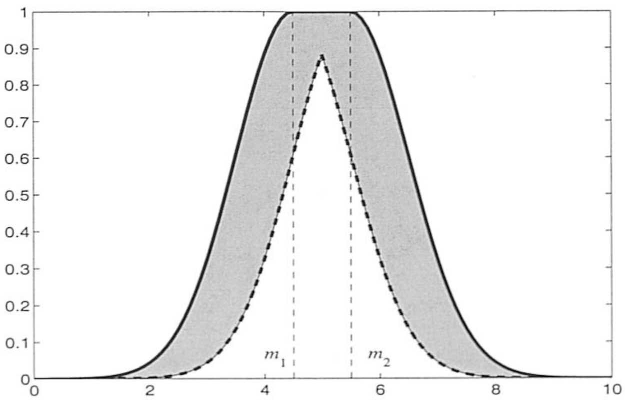

Fig. 3. FOU for Gaussian primary MF with uncertain mean.

Matrices $\mathbf{M}_{l1}$ and $\mathbf{M}_{l2}$ and vectors $\mathbf{r}_l$ and $\mathbf{s}_l$ will automatically dispose of the unnecessary elements of $\bar{\mathbf{f}}$ and $\underline{\mathbf{f}}$, since they depend on L.

B. y_r Re-Expressed in Rule-Ordered Format

Fact 2: y_r can be re-expressed in terms of rule-ordered quantities as

$$\begin{aligned} y_r &= \frac{\underline{\mathbf{f}}^T\mathbf{M}_{r1}\mathbf{z}_r + \bar{\mathbf{f}}^T\mathbf{M}_{r2}\mathbf{z}_r}{\mathbf{r}_r^T\mathbf{Q}_r\underline{\mathbf{f}} + \mathbf{s}_r^T\mathbf{Q}_r\bar{\mathbf{f}}} \\ &= \frac{\underline{\mathbf{f}}^T\mathbf{a}_r + \bar{\mathbf{f}}^T\mathbf{b}_r}{\underline{\mathbf{f}}^T\mathbf{c}_r + \bar{\mathbf{f}}^T\mathbf{d}_r} \\ &= \frac{\sum_{i=1}^M a_{r,i}\underline{f}^i + \sum_{j=1}^M b_{r,j}\bar{f}^j}{\sum_{i=1}^M c_{r,i}\underline{f}^i + \sum_{j=1}^M d_{r,j}\bar{f}^j}. \end{aligned} \tag{12}$$

Proof: Just follow the proof of Fact 1 using quantities that are defined in the right-hand column of Table II. ■

IV. CALCULATION OF $\partial e^{(i)}/\partial\theta_{k,m}^l$ FOR ANTECEDENT PARAMETERS

Antecedent parameters are the parameters that characterize antecedent MFs. For example, a Gaussian primary MF with uncertain mean, as defined in (A-7), is characterized by three parameters, m_1, m_2 and σ. Temporarily, let us denote any one of the antecedent parameters that will be tuned as $\theta_{k,m}^l$ ($k = 1,\ldots,p$ and $l = 1,\ldots,M$). Index m denotes the fact that there can be more than one parameter associated with the MF of each antecedent (k) and rule (l). Here we use the chain rule to compute $\partial e^{(i)}/\partial\theta_{k,m}^l$. Our starting point is (2), in which $f_{s2}(\mathbf{x})$ is given by (A-18), which we restate here as

$$f_{s2}(\mathbf{x}) = \frac{y_l(\mathbf{x}) + y_r(\mathbf{x})}{2}. \tag{13}$$

Hence

$$\begin{aligned} \frac{\partial e^{(i)}}{\partial\theta_{k,m}^l} &= \frac{\partial e^{(i)}}{\partial f_{s2}}\left[\frac{\partial f_{s2}}{\partial y_l}\frac{\partial y_l}{\partial\theta_{k,m}^l} + \frac{\partial f_{s2}}{\partial y_r}\frac{\partial y_r}{\partial\theta_{k,m}^l}\right] \\ &= \frac{1}{2}\left[f_{s2}\left(\mathbf{x}^{(i)}\right) - y^{(i)}\right]\left[\frac{\partial y_l}{\partial\theta_{k,m}^l} + \frac{\partial y_r}{\partial\theta_{k,m}^l}\right] \end{aligned} \tag{14}$$

where we used the facts that $\partial e^{(i)}/\partial f_{s2} = [f_{s2}(\mathbf{x}^{(i)}) - y^{(i)}]$ and $\partial f_{s2}(\mathbf{x})/\partial y_1 = \partial f_{s2}(\mathbf{x})/\partial y_2 = 1/2$. In (14), we now treat y_l and y_r as functions of $\underline{\mathbf{f}}$ and $\bar{\mathbf{f}}$; hence

$$\frac{\partial y_l}{\partial\theta_{k,m}^l} = \sum_{j=1}^M\left(\frac{\partial y_l}{\partial\bar{f}^j}\frac{\partial\bar{f}^j}{\partial\theta_{k,m}^l} + \frac{\partial y_l}{\partial\underline{f}^j}\frac{\partial\underline{f}^j}{\partial\theta_{k,m}^l}\right) \tag{15}$$

$$\frac{\partial y_r}{\partial\theta_{k,m}^l} = \sum_{j=1}^M\left(\frac{\partial y_r}{\partial\bar{f}^j}\frac{\partial\bar{f}^j}{\partial\theta_{k,m}^l} + \frac{\partial y_r}{\partial\underline{f}^j}\frac{\partial\underline{f}^j}{\partial\theta_{k,m}^l}\right). \tag{16}$$

We now need to evaluate all of the derivatives in (15) and (16).

Fact 3: The following are true:

$$\frac{\partial y_l}{\partial\bar{f}^j} = \frac{a_{l,j} - y_l c_{l,j}}{\bar{\mathbf{f}}^T\mathbf{c}_l + \underline{\mathbf{f}}^T\mathbf{d}_l} \tag{17}$$

$$\frac{\partial y_l}{\partial\underline{f}^j} = \frac{b_{l,j} - y_l d_{l,j}}{\bar{\mathbf{f}}^T\mathbf{c}_l + \underline{\mathbf{f}}^T\mathbf{d}_l} \tag{18}$$

$$\frac{\partial y_r}{\partial\bar{f}^j} = \frac{b_{r,j} - y_r d_{r,j}}{\underline{\mathbf{f}}^T\mathbf{c}_r + \bar{\mathbf{f}}^T\mathbf{d}_r} \tag{19}$$

$$\frac{\partial y_r}{\partial\underline{f}^j} = \frac{a_{r,j} - y_r c_{r,j}}{\underline{\mathbf{f}}^T\mathbf{c}_r + \bar{\mathbf{f}}^T\mathbf{d}_r}. \tag{20}$$

Proof: Because all calculations are alike, we only provide the derivation of (18). From the second and third forms of y_l in (7), it follows that

$$\frac{\partial y_l}{\partial\underline{f}^j} = \frac{(\bar{\mathbf{f}}^T\mathbf{c}_l + \underline{\mathbf{f}}^T\mathbf{d}_l)b_{l,j} - (\bar{\mathbf{f}}^T\mathbf{a}_l + \underline{\mathbf{f}}^T\mathbf{b}_l)d_{l,j}}{\left(\bar{\mathbf{f}}^T\mathbf{c}_l + \underline{\mathbf{f}}^T\mathbf{d}_l\right)^2} \tag{21}$$

so that

$$\frac{\partial y_l}{\partial\underline{f}^j} = \frac{b_{l,j}}{\bar{\mathbf{f}}^T\mathbf{c}_l + \underline{\mathbf{f}}^T\mathbf{d}_l} - \frac{y_l d_{l,j}}{\bar{\mathbf{f}}^T\mathbf{c}_l + \underline{\mathbf{f}}^T\mathbf{d}_l} = \frac{b_{l,j} - y_l d_{l,j}}{\bar{\mathbf{f}}^T\mathbf{c}_l + \underline{\mathbf{f}}^T\mathbf{d}_l}. \tag{22}$$

Q.E.D. ■

TABLE IV
DERIVATIVES OF $\bar{\mu}_k^l(x_k)$

x_k	$\partial\bar{\mu}_k^l(x_k)/\partial\theta_{k,1}^l$	$\partial\bar{\mu}_k^l(x_k)/\partial\theta_{k,2}^l$	$\partial\bar{\mu}_k^l(x_k)/\partial\theta_{k,3}^l$
$x_k < m_{k1}^l$	$(x-m_{k1})N(m_{k1},\sigma;x)/\sigma^2$	0	$(x-m_{k1})^2 N(m_{k1},\sigma;x)/\sigma^3$
$m_{k1}^l \le x_k \le (m_{k1}^l + m_{k2}^l)/2$	0	0	0
$(m_{k1}^l + m_{k2}^l)/2 \le x_k \le m_{k2}^l$	0	0	0
$x_k > m_{k2}^l$	0	$(x-m_{k2})N(m_{k2},\sigma;x)/\sigma^2$	$(x-m_{k2})^2 N(m_{k2},\sigma;x)/\sigma^3$

TABLE V
DERIVATIVES OF $\underline{\mu}_k^l(x_k)$

x_k	$\partial\underline{\mu}_k^l(x_k)/\partial\theta_{k,1}^l$	$\partial\underline{\mu}_k^l(x_k)/\partial\theta_{k,2}^l$	$\partial\underline{\mu}_k^l(x_k)/\partial\theta_{k,3}^l$
$x_k \le (m_{k1}^l + m_{k2}^l)/2$	0	$(x-m_{k2})N(m_{k2},\sigma;x)/\sigma^2$	$(x-m_{k2})^2 N(m_{k2},\sigma;x)/\sigma^3$
$x_k > (m_{k1}^l + m_{k2}^l)/2$	$(x-m_{k1})N(m_{k1},\sigma;x)/\sigma^2$	0	$(x-m_{k1})^2 N(m_{k1},\sigma;x)/\sigma^3$

Next, we focus on computing $\partial\underline{f}^j/\theta_{k,m}^l$ and $\partial\bar{f}^j/\theta_{k,m}^l$ that are needed in (15) and (16). Formulas for $\underline{f}^l(\mathbf{x})$ and $\bar{f}^l(\mathbf{x})$ are given in (A-11) and (A-12), respectively. Observe that the former are in terms of the antecedent upper MFs $\bar{\mu}_k^l$ whereas the latter are in terms of the antecedent lower MFs $\underline{\mu}_k^l$.

Fact 4: Parameter $\theta_{k,m}^l$ can *only* appear in $\underline{\mu}_k^l$ or $\bar{\mu}_k^l$ and cannot appear in $\underline{\mu}_j^l$ and $\bar{\mu}_j^l$ for $j \neq k$.

Proof: This is a direct result of Assumption 1. ■

As a direct consequence of Fact 4, we see that[9]

$$\frac{\partial \underline{f}^j}{\partial\theta_{k,m}^l} = \begin{cases} \frac{\partial \underline{f}^l}{\partial\theta_{k,m}^l} \\ 0\ \forall j \neq l \end{cases} \tag{23}$$

$$\frac{\partial \bar{f}^j}{\partial\theta_{k,m}^l} = \begin{cases} \frac{\partial \bar{f}^l}{\partial\theta_{k,m}^l} \\ 0\ \forall j \neq l \end{cases}. \tag{24}$$

Hence, (15) and (16) simplify to

$$\frac{\partial y_l}{\partial\theta_{k,m}^l} = \frac{\partial y_l}{\partial \bar{f}^l}\frac{\partial \bar{f}^l}{\partial\theta_{k,m}^l} + \frac{\partial y_l}{\partial \underline{f}^l}\frac{\partial \underline{f}^l}{\partial\theta_{k,m}^l} \tag{25}$$

$$\frac{\partial y_r}{\partial\theta_{k,m}^l} = \frac{\partial y_r}{\partial \bar{f}^l}\frac{\partial \bar{f}^l}{\partial\theta_{k,m}^l} + \frac{\partial y_r}{\partial \underline{f}^l}\frac{\partial \underline{f}^l}{\partial\theta_{k,m}^l} \tag{26}$$

so that [see (14)]

$$\frac{\partial y_l}{\partial\theta_{k,m}^l} + \frac{\partial y_r}{\partial\theta_{k,m}^l} = \left(\frac{\partial y_l}{\partial \bar{f}^l} + \frac{\partial y_r}{\partial \bar{f}^l}\right)\frac{\partial \bar{f}^l}{\partial\theta_{k,m}^l} + \left(\frac{\partial y_l}{\partial \underline{f}^l} + \frac{\partial y_r}{\partial \underline{f}^l}\right)\frac{\partial \underline{f}^l}{\partial\theta_{k,m}^l}. \tag{27}$$

Fact 5: It is true that

$$\frac{\partial e^{(i)}}{\partial\theta_{k,m}^l} = \frac{1}{2}\left[f_{s2}\left(\mathbf{x}^{(i)}\right) - y^{(i)}\right] \times \left[\left(\frac{\partial y_l}{\partial \bar{f}^l} + \frac{\partial y_r}{\partial \bar{f}^l}\right)\frac{\partial \bar{f}^l}{\partial\theta_{k,m}^l} + \left(\frac{\partial y_l}{\partial \underline{f}^l} + \frac{\partial y_r}{\partial \underline{f}^l}\right)\frac{\partial \underline{f}^l}{\partial\theta_{k,m}^l}\right]. \tag{28}$$

Proof: Substitute (27) into (14). ■

Fact 6: It is also true that for *product* t-norm

$$\frac{\partial \underline{f}^l}{\partial\theta_{k,m}^l} = \left[T_{\substack{i=1\\i\neq k}}^p \underline{\mu}_i^l(x_i)\right] \times \frac{\partial \underline{\mu}_k^l(x_k)}{\partial\theta_{k,m}^l} \tag{29a}$$

$$\frac{\partial \bar{f}^l}{\partial\theta_{k,m}^l} = \left[T_{\substack{i=1\\i\neq k}}^p \bar{\mu}_i^l(x_i)\right] \times \frac{\partial \bar{\mu}_k^l(x_k)}{\partial\theta_{k,m}^l} \tag{30a}$$

and for *minimum* t-norm

$$\frac{\partial \underline{f}^l}{\partial\theta_{k,m}^l} = \frac{\partial \underline{\mu}_k^l(x_k)}{\partial\theta_{k,m}^l} \times \prod_{j=1}^p \xi\left[\underline{\mu}_j^l(x_j) - \underline{\mu}_k^l(x_k)\right] \tag{29b}$$

$$\frac{\partial \bar{f}^l}{\partial\theta_{k,m}^l} = \frac{\partial \bar{\mu}_k^l(x_k)}{\partial\theta_{k,m}^l} \times \prod_{j=1}^p \xi\left[\bar{\mu}_j^l(x_j) - \bar{\mu}_k^l(x_k)\right] \tag{30b}$$

where $\xi(\zeta) = 1$, when $\zeta \ge 0$ and $\xi(\zeta) = 0$, when $\zeta < 0$.

Proof: Apply the chain rule to (A-11) and (A-12) making use of Fact 4.[10] ■

The remaining calculations of $\partial\underline{\mu}_k^l(x_k)/\partial\theta_{k,m}^l$ and $\partial\bar{\mu}_k^l(x_k)/\partial\theta_{k,m}^l$ require specification of antecedent MFs and their associated FOUs.

So, as to see the forest from the trees, we next present a computational flow chart (see Fig. 2) for the calculations of $\partial e^{(i)}/\partial\theta_{k,m}^l$. We show the inherent parallelism in the computations and where additional information is needed before the computations can be completed. Many of the computations only have to be done one time regardless of which antecedent parameters are tuned. The ones in the heavier outlined blocks must be done for each $\theta_{k,m}^l$.

See Appendix B for comparable results for an interval type-2 nonsingleton type-2 FLS.

[9]We hasten to point out that if an MF parameter is *shared* across some or all antecedent MFs, then (23) and (24) are invalid, and a different analysis must be performed from this point on.

[10]For minimum t-norm, note for example, that $\min[a(\theta_1), b(\theta_2), c(\theta_3)] = a(\theta_1)$ if $a(\theta_1) < b(\theta_2)$ and $a(\theta_1) < c(\theta_3)$, and equals λ otherwise, where λ is *not* a function of θ_1. Consequently, $\partial\min[a(\theta_1), b(\theta_2), c(\theta_3)]/\partial\theta_1 = \partial a(\theta_1)/\partial\theta_1$ if $a(\theta_1) < b(\theta_2)$ and $a(\theta_1) < c(\theta_3)$, and equals 0 otherwise. This can be expressed mathematically as: $\partial\min[a(\theta_1), b(\theta_2), c(\theta_3)]/\partial\theta_1 = [\partial a(\theta_1)/\partial\theta_1]\times\xi[b(\theta_2) - a(\theta_1)] \times \xi[c(\theta_3) - a(\theta_1)]$. Note that $\xi(\cdot)$ is a step function.

TABLE VI
MEANINGS OF y_l^i, y_r^i, $\underline{f}^i$, $\overline{f}^i$ AND M IN (A-15) FOR DIFFERENT TYPE-REDUCTION METHODS[a]. IN THIS TABLE, ALL SYMBOLS REFER TO RULE-ORDERED QUANTITIES

Type-reduction method	y_l^i **and** y_r^i **defined**	$\underline{f}^i$ **and** $\overline{f}^i$ **defined**[d]	***M* defined**
center-of sets	left and right end points of the centroid of the consequent of the *i*th rule; treated as consequent parameters	lower and upper firing degrees of the *i*th rule; contains antecedent MF parameters	number of rules
centroid[b]	$y_l^i = y_r^i = y^i$, 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 output of the FLS's output; contains antecedent and consequent MF parameters	number of sampled points
center-of-sums[c]	$y_l^i = y_r^i = y^i$, 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; contains antecedent and consequent MF parameters	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; treated as a consequent parameter[e]	lower and upper firing degrees of the *i*th rule; contains antecedent MF parameters	number of rules

a. Comparable results for *modified height type-reduction* can be found in [6] Section 9.5.4.

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

c. 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).

d. See Section A.2 for formulas for $\underline{f}^i$ and $\overline{f}^i$, as well as for rule outputs and unioned rule outputs.

e. The parameters y^i are shared across both y_l and y_r; so the calculation of derivatives of consequent parameters, as described in Section V, must be modified in this case.

V. CALCULATION OF $\partial e^{(i)}/\partial\theta^j$ FOR CONSEQUENT PARAMETERS

Generally speaking, consequent parameters are the parameters that characterize consequent MFs. When, however, we use center-of-set type-reduction, as we have assumed in Assumption 4 (Section II), then those parameters can be replaced by the two end-points of the centroids of the type-2 consequent sets (see Table A-I in Appendix A). Doing this can reduce the number of design parameters. For example, if the consequent MF is also a Gaussian primary MF with uncertain mean, that is characterized by three design parameters, then using the two end-points of the centroids of this type-2 MF reduces the number of design parameters from three to two.

Note that the consequent parameters do not need the "k" or "m" subscripts in $\theta_{k,m}^j$ (k and m are associated with a specific antecedent). Additionally, (see the formulas for y_l and y_r in Table I) $\theta^j = z_l^j$ or $\theta^j = z_r^j$.[11]

[11]Note that this parameterization is true for center-of-sets and height type-reduction but is not true for centroid or center of sums type-reduction.

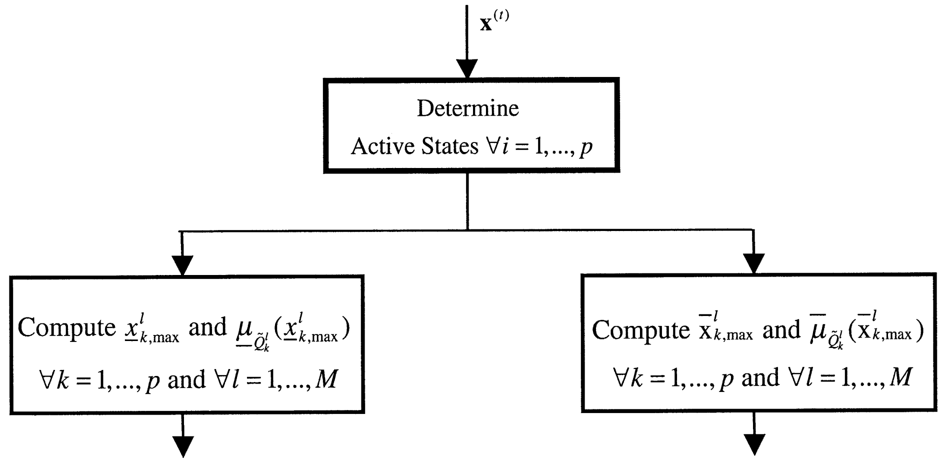

Fig. 4. Modification to top part of Fig. 2.

TABLE VII
$\underline{\mu}_{\tilde{Q}_k^l}(\underline{x}^l_{k,\max})$ AND $\overline{\mu}_{\tilde{Q}_k^l}(\bar{x}^l_{k,\max})$ UNDER PRODUCT T-NORM. FOR MINIMUM T-NORM, REPLACE σ_{k1}^2 BY σ_{k1}, σ_{k2}^2 BY σ_{k2} AND $\sigma_k^{l^2}$ BY σ_k^l. STATES ARE DEFINED IN TABLE VIII

State	$\underline{\mu}_{\tilde{Q}_k^l}(\underline{x}^l_{k,\max})$	$\overline{\mu}_{\tilde{Q}_k^l}(\bar{x}^l_{k,\max})$
(1)	$\exp\left[-\frac{1}{2}\left(m_{k2}^l - x_k'\right)^2 / \left(\sigma_{k1}^2 + \sigma_k^{l^2}\right)\right]$	$\exp\left[-\frac{1}{2}\left(m_{k1}^l - x_k'\right)^2 / \left(\sigma_{k2}^2 + \sigma_k^{l^2}\right)\right]$
(2)	$\exp\left[-\frac{1}{2}\left(m_{k2}^l - x_k'\right)^2 / \left(\sigma_{k1}^2 + \sigma_k^{l^2}\right)\right]$	1
(3)	$\exp\left[-\frac{1}{2}\left(m_{k2}^l + m_{k1}^l - 2x_k'\right)^2 / 4\sigma_{k1}^2 - \frac{1}{2}\left(m_{k2}^l - m_{k1}^l\right)^2 / 4\sigma_k^{l^2}\right]$	1
(4)	$\exp\left[-\frac{1}{2}\left(m_{k1}^l - x_k'\right)^2 / \left(\sigma_{k1}^2 + \sigma_k^{l^2}\right)\right]$	1
(5)	$\exp\left[-\frac{1}{2}\left(m_{k1}^l - x_k'\right)^2 / \left(\sigma_{k1}^2 + \sigma_k^{l^2}\right)\right]$	$\exp\left[-\frac{1}{2}\left(m_{k2}^l - x_k'\right)^2 / \left(\sigma_{k2}^2 + \sigma_k^{l^2}\right)\right]$

From (2) and (13), it is easy to show that

$$\frac{\partial e^{(i)}}{\partial z_l^j} = \frac{\partial e^{(i)}}{\partial f_{s2}\left(\mathbf{x}^{(i)}\right)} \frac{\partial f_{s2}\left(\mathbf{x}^{(i)}\right)}{\partial y_l} \frac{\partial y_l}{\partial z_l^j} = \frac{1}{2}\left[f_{s2}\left(\mathbf{x}^{(i)}\right) - y^{(i)}\right] \frac{\partial y_l}{\partial z_l^i} \tag{31}$$

and

$$\frac{\partial e^{(i)}}{\partial z_r^j} = \frac{1}{2}\left[f_{s2}\left(\mathbf{x}^{(i)}\right) - y^{(i)}\right] \frac{\partial y_r}{\partial z_r^j}. \tag{32}$$

Fact 7: It is true that

$$\frac{\partial y_l}{\partial z_l^j} = \mathbf{e}_j^T \left(\frac{\mathbf{M}_{l1}^T \bar{\mathbf{f}} + \mathbf{M}_{l2}^T \underline{\mathbf{f}}}{\mathbf{r}_l^T \mathbf{Q}_l \bar{\mathbf{f}} + \mathbf{s}_l^T \mathbf{Q}_l \underline{\mathbf{f}}} \right) \tag{33}$$

$$\frac{\partial y_r}{\partial z_r^j} = \mathbf{e}_j^T \left(\frac{\mathbf{M}_{r1}^T \underline{\mathbf{f}} + \mathbf{M}_{r2}^T \bar{\mathbf{f}}}{\mathbf{r}_r^T \mathbf{Q}_r \underline{\mathbf{f}} + \mathbf{s}_r^T \mathbf{Q}_r \bar{\mathbf{f}}} \right) \tag{34}$$

where $\mathbf{e}_j$ is the jth $M \times 1$ unit vector.

Proof: Using the vector calculus fact that $\mathrm{grad}_{\mathbf{z}} \boldsymbol{\alpha}^T \mathbf{z} = \boldsymbol{\alpha}$, it is easy to show, from (7) and (12), that

$$\mathrm{grad}_{\mathbf{z}_l} y_l = \left(\frac{\bar{\mathbf{f}}^T \mathbf{M}_{l1} + \underline{\mathbf{f}}^T \mathbf{M}_{l2}}{\mathbf{r}_l^T \mathbf{Q}_l \bar{\mathbf{f}} + \mathbf{s}_l^T \mathbf{Q}_l \underline{\mathbf{f}}} \right)^T = \left(\frac{\mathbf{M}_{l1}^T \bar{\mathbf{f}} + \mathbf{M}_{l2}^T \underline{\mathbf{f}}}{\mathbf{r}_l^T \mathbf{Q}_l \bar{\mathbf{f}} + \mathbf{s}_l^T \mathbf{Q}_l \underline{\mathbf{f}}} \right) \tag{35}$$

$$\mathrm{grad}_{\mathbf{z}_r} y_r = \left(\frac{\underline{\mathbf{f}}^T \mathbf{M}_{r1} + \bar{\mathbf{f}}^T \mathbf{M}_{r2}}{\mathbf{r}_r^T \mathbf{Q}_r \underline{\mathbf{f}} + \mathbf{s}_r^T \mathbf{Q}_r \bar{\mathbf{f}}} \right)^T = \left(\frac{\mathbf{M}_{r1}^T \underline{\mathbf{f}} + \mathbf{M}_{r2}^T \bar{\mathbf{f}}}{\mathbf{r}_r^T \mathbf{Q}_r \underline{\mathbf{f}} + \mathbf{s}_r^T \mathbf{Q}_r \bar{\mathbf{f}}} \right). \tag{36}$$

Equations (33) and (34) follow directly upon application of $\mathbf{e}_j^T$ to (35) and (36), respectively.[12] ■

[12]Recall that $\mathrm{grad}_{\mathbf{z}_l} y_l = [\partial y_l/\partial z_l^1, \partial y_l/\partial z_l^2, \ldots, \partial y_l/\partial z_l^M]^T$, so that $\mathbf{e}_j^T \mathrm{grad}_{\mathbf{z}_l} y_l = \partial y_l/\partial z_l^j$.

TABLE VIII
DEFINITIONS OF THE FIVE STATES. FOR MINIMUM T-NORM, REPLACE σ_{k1}^2 BY σ_{k1}, σ_{k2}^2 BY σ_{k2} AND $\sigma_k^{l^2}$ BY σ_k^l

State	Product t-norm
(1)	$x_k' \le m_{k1}^l$ and $x_k' < \frac{m_{k1}^l + m_{k2}^l}{2} - \frac{\sigma_{k1}^2 (m_{k2}^l - m_{k1}^l)}{2\sigma_k^{l^2}}$
(2)	$x_k' \in [m_{k1}^l, m_{k2}^l]$ and $x_k' < \frac{m_{k1}^l + m_{k2}^l}{2} - \frac{\sigma_{k1}^2 (m_{k2}^l - m_{k1}^l)}{2\sigma_k^{l^2}}$,
(3)	$x_k' \in [m_{k1}^l, m_{k2}^l]$ and $x_k' \in \left[\frac{m_{k1}^l + m_{k2}^l}{2} - \frac{\sigma_{k1}^2 (m_{k2}^l - m_{k1}^l)}{2\sigma_k^{l^2}}, \frac{m_{k1}^l + m_{k2}^l}{2} + \frac{\sigma_{k1}^2 (m_{k2}^l - m_{k1}^l)}{2\sigma_k^{l^2}} \right]$
(4)	$x_k' \in [m_{k1}^l, m_{k2}^l]$ and $x_k' > \frac{m_{k1}^l + m_{k2}^l}{2} + \frac{\sigma_{k1}^2 (m_{k2}^l - m_{k1}^l)}{2\sigma_k^{l^2}}$
(5)	$x_k' \ge m_{k2}^l$ and $x_k' > \frac{m_{k1}^l + m_{k2}^l}{2} + \frac{\sigma_{k1}^2 (m_{k2}^l - m_{k1}^l)}{2\sigma_k^{l^2}}$

This completes the derivations of general derivative formulas. To proceed further, the FOUs of MFs must be specified. Because calculations of $\partial e^{(i)}/\partial \theta^j$ for consequent parameters are so straightforward, we do not include a flowchart for their implementation. Just implement (31)–(34), and use the vectors and matrices that are defined in Table II.

VI. EXAMPLE

In order to compute $\partial \underline{f}^l / \partial \theta_{k,m}^l$ and $\partial \overline{f}^l / \partial \theta_{k,m}^l$ using (29a) and (30a), we need $\partial \underline{\mu}_k^l(x_k)/\partial \theta_{k,m}^l$ and $\partial \overline{\mu}_k^l(x_k)/\partial \theta_{k,m}^l$. Here, we will compute $\partial \underline{\mu}_k^l(x_k)/\partial \theta_{k,m}^l$ and $\partial \overline{\mu}_k^l(x_k)/\partial \theta_{k,m}^l$ for *Gaussian primary MFs with uncertain means* (see Example A-1). To start off, we restate the results of Example A-1 using the more explicit notations for antecedent MFs ($k = 1, \ldots, p$ and $l = 1, \ldots, M$)

$$\mu_k^l(x_k) = \exp\left\{ -\frac{1}{2}\left(\frac{x_k - m_k^l}{\sigma_k^l} \right)^2 \right\} \quad m_k^l \in [m_{k1}^l, m_{k2}^l] \tag{37}$$

$$\overline{\mu}_k^l(x_k) = \begin{cases} N(m_{k1}^l, \sigma_k^l; x_k) & x_k < m_{k1}^l \\ 1 & m_{k1}^l \le x_k \le m_{k2}^l \\ N(m_{k2}^l, \sigma_k^l; x_k) & x_k > m_{k2}^l \end{cases} \tag{38}$$

$$\underline{\mu}_k^l(x_k) = \begin{cases} N(m_{k2}^l, \sigma_k^l; x_k) & x_k \le \frac{(m_{k1}^l + m_{k2}^l)}{2} \\ N(m_{k1}^l, \sigma_k^l; x_k) & x_k > \frac{(m_{k1}^l + m_{k2}^l)}{2} \end{cases}. \tag{39}$$

We summarize the parameters that $\bar{\mu}_k^l(x_k)$ and $\underline{\mu}_k^l(x_k)$ depend upon, as a function of x_k, in Table III. Its results were obtained by examining (38) and (39) (see, also, Fig. 3). Let $\theta_{k,1}^l \equiv m_{k1}^l$, $\theta_{k,2}^l \equiv m_{k2}^l$, and $\theta_{k,3}^l \equiv \sigma_k^l$. Tables IV and V provide $\partial \overline{\mu}_k^l(x_k)/\partial \theta_{k,m}^l$ and $\partial \underline{\mu}_k^l(x_k)/\partial \theta_{k,m}^l$, respectively.

The results in these Tables IV and V would be used in Fig. 2 as follows. The tests on variable $x_k^{(t)}$, given in Tables IV and V, let us implement the top two blocks in which we have to determine the active branches of the lower and upper MFs. The results in Table V provide $\partial \underline{\mu}_k^l(x_k^{(t)})/\partial \theta_{k,m}^l$ which is needed to compute $\partial \underline{f}^l(x_k^{(t)})/\partial \theta_{k,m}^l$, and the results in Table IV provide $\partial \overline{\mu}_k^l(x_k^{(t)})/\partial \theta_{k,m}^l$ which is needed to compute $\partial \overline{f}^l(x_k^{(t)})/\partial \theta_{k,m}^l$.

TABLE IX
UPPER AND LOWER MF PARAMETER DEPENDENCIES (SEE TABLE VII)

State	$\underline{\mu}_{\tilde{Q}_k^l}(\underline{x}_{k,\max}^l)$	$\overline{\mu}_{\tilde{Q}_k^l}(\overline{x}_{k,\max}^l)$
(1)	m_{k2}^l, σ_k^l, σ_{k1}	m_{k1}^l, σ_k^l, σ_{k2}
(2)	m_{k2}^l, σ_k^l, σ_{k1}	None
(3)	m_{k2}^l, m_{k1}^l, σ_k^l, σ_{k1}	None
(4)	m_{k1}^l, σ_k^l, σ_{k1}	None
(5)	m_{k1}^l, σ_k^l, σ_{k1}	m_{k2}^l, σ_k^l, σ_{k2}

* Includes input MF parameter dependencies.

VII. CONCLUSION

We have made type-2 FLSs much more accessible to FLS designers by providing mathematical formulas and computational flowcharts for computing the derivatives that are needed to implement steepest-descent parameter tuning algorithms for such systems. We have demonstrated why computing such derivatives is much more challenging than it is for a type-1 FLS, and have provided derivative calculations that are applicable to any kind of type-2 MF, since most of the calculations can be performed without prespecifying the nature of those MFs. Eventually, one does have to specify the nature of the type-2 MF in order to complete the calculations. We showed how to complete the calculations for a Gaussian primary MF with uncertain means.

It is important for the reader to remember the four assumptions stated in Section II. If any of them are not obeyed, then some or all of the results of this paper must be modified.

TABLE X
DERIVATIVES OF* $\underline{\mu}_{\tilde{Q}_k^l}(\underline{x}_{k,\max}^l) \equiv \underline{\mu}$

State	$\partial\underline{\mu}/\partial\theta_{k,1}^l$	$\partial\underline{\mu}/\partial\theta_{k,2}^l$	$\partial\underline{\mu}/\partial\theta_{k,3}^l$	$\partial\underline{\mu}/\partial\theta_{k,4}^l$	$\partial\underline{\mu}/\partial\theta_{k,5}^l$
(1)	0	$\partial(1)/\partial\theta_{k,2}^l$	$\partial(1)/\partial\theta_{k,3}^l$	$\partial(1)/\partial\theta_{k,4}$	0
(2)	0	$\partial(1)/\partial\theta_{k,2}^l$	$\partial(1)/\partial\theta_{k,3}^l$	$\partial(1)/\partial\theta_{k,4}$	0
(3)	$\partial(3)/\partial\theta_{k,1}^l$	$\partial(3)/\partial\theta_{k,2}^l$	$\partial(3)/\partial\theta_{k,3}^l$	$\partial(3)/\partial\theta_{k,4}$	0
(4)	$\partial(4)/\partial\theta_{k,1}^l$	0	$\partial(4)/\partial\theta_{k,3}^l$	$\partial(4)/\partial\theta_{k,4}$	0
(5)	$\partial(4)/\partial\theta_{k,1}^l$	0	$\partial(4)/\partial\theta_{k,3}^l$	$\partial(4)/\partial\theta_{k,4}$	0

* e.g., $\partial(4)/\partial\theta_{k,1}^l \equiv \partial\underline{\mu}/\partial\theta_{k,1}^l$ for State (4). Use Table VII to complete this computation.

TABLE XI
DERIVATIVES OF* $\overline{\mu}_{\tilde{Q}_k^l}(\overline{x}_{k,\max}^l) \equiv \overline{\mu}$

State	$\partial\overline{\mu}/\partial\theta_{k,1}^l$	$\partial\overline{\mu}/\partial\theta_{k,2}^l$	$\partial\overline{\mu}/\partial\theta_{k,3}^l$	$\partial\overline{\mu}/\partial\theta_{k,4}^l$	$\partial\overline{\mu}/\partial\theta_{k,5}^l$
(1)	$\partial(1)/\partial\theta_{k,1}^l$	0	$\partial(1)/\partial\theta_{k,3}^l$	0	$\partial(1)/\partial\theta_{k,5}$
(2)	0	0	0	0	0
(3)	0	0	0	0	0
(4)	0	0	0	0	0
(5)	0	$\partial(5)/\partial\theta_{k,2}^l$	$\partial(5)/\partial\theta_{k,3}^l$	0	$\partial(5)/\partial\theta_{k,5}$

* e.g., $\partial(1)/\partial\theta_{k,1}^l \equiv \partial\overline{\mu}/\partial\theta_{k,1}^l$ for State (1). Use Table VII to complete this computation.

APPENDIX A
CALCULATIONS REQUIRED TO IMPLEMENT AN INTERVAL SINGLETON TYPE-2 FUZZY LOGIC SYSTEM

In this appendix we summarize all of the calculations that are needed to implement an interval singleton type-2 FLS. For detailed derivations of all of these results as well as more background on type-2 FLSs, see [6]. To begin we define important terms that are associated with a type-2 fuzzy set and a type-2 FLS.

A. Preliminaries

Consider a type-2 FLS having p inputs $x_1 \in X_1, \ldots, x_p \in X_p$ and one output $y \in Y$. We assume there are M rules where the lth rule has the form

$$R^l : \text{IF } x_1 \text{ is } \tilde{F}_1^l \text{ and} \cdots \text{and } x_p \text{ is } \tilde{F}_p^l, \quad \text{THEN } y \text{ is } \tilde{G}^l, \qquad l = 1, \ldots, M. \tag{A-1}$$

This rule represents a type-2 relation between the input space $X_1 \times \cdots \times X_p$, and the output space, Y, of the type-2 FLS. Associated with the p antecedent type-2 fuzzy sets, $\tilde{F}_i^l$, are the type-2 MFs $\mu_{\tilde{F}_i^l}$ $(i = 1, \ldots, p)$, and associated with the consequent type-2 fuzzy set $\tilde{G}^l$ is its type-2 MF $\mu_{\tilde{G}^l}$. We frequently use the simpler notation μ_i^l for $\mu_{\tilde{F}_i^l}$.

Definition A-1: A *type-2 fuzzy set*, denoted $\tilde{A}$, is characterized by a (three-dimensional) *type-2 membership function* (MF) $\mu_{\tilde{A}}(x, u)$, i.e.,

$$\tilde{A} = \int_{x\in X}\int_{u\in J_x} \mu_{\tilde{A}}(x,u)/(x,u) \quad J_x \subseteq [0,1] \tag{A-2}$$

where $\int\int$ denotes union over all admissible x and u, and $0 \le \mu_{\tilde{A}}(x,u) \le 1$. At each fixed value of $x \in X$, J_x is the *primary membership* of x and x is called the *primary variable*. ■

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

$$\mu_{\tilde{A}}(x = x', u) \equiv \mu_{\tilde{A}}(x') = \int_{u\in J_{x'}} f_{x'}(u)/u \quad J_{x'} \subseteq [0,1] \tag{A-3}$$

in which $0 \le f_{x'}(u) \le 1$. Because $\forall x' \in X$, we drop the prime notation on $\mu_{\tilde{A}}(x')$, and refer to $\mu_{\tilde{A}}(x)$ as a secondary MF; it is a type-1 fuzzy set, which we also refer to as a *secondary set*. ■

Definition A-3: The *domain* of a secondary MF is called the *primary membership* of x. In (A-2) and (A-3), J_x is the primary membership of x, where $J_x \subseteq [0,1]$ for $\forall x \in X$. ■

Definition A-4: The *amplitude* of a secondary MF is called a *secondary grade*. In (A-3), $f_{x'}(u)$ is a secondary grade. ■

Definition A-5: An *interval type-2 fuzzy* set is a type-2 fuzzy set all of whose secondary MFs are type-1 interval sets, i.e., $f_x(u) = 1, \forall u \in J_x \subseteq [0,1], \forall x \in X$. ■

Interval secondary MFs reflect a uniform uncertainty at the primary memberships of x, and are the ones most commonly used in a type-2 FLS. Note that an interval set can be represented just by its domain interval, which can be expressed in terms of its left and right end-points as $[l, r]$, or by its center and spread as $[c - s, c + s]$, where $c = (l + r)/2$ and $s = (r - l)/2$.

Definition A-6: Uncertainty in the primary memberships of a type-2 fuzzy set, $\tilde{A}$, consists of a bounded region that we call the FOU. It is the union of all primary memberships, i.e.,

$$\mathrm{FOU}(\tilde{A}) = \bigcup_{x \in X} J_x. \quad \blacksquare \tag{A-4}$$

The term FOU is very useful, because it not only focuses our attention on the uncertainties inherent in a specific type-2 MF, whose shape is a direct consequence of the nature of these uncertainties, but it also provides a very convenient verbal description of the entire domain of support for all the secondary grades of a type-2 MF.

Definition A-7: Consider a family of type-1 MFs $\mu_A(x|p_1, p_2, \ldots, p_v)$ where $p_1, p_2, \ldots, p_v$ are parameters, some or all of which vary over some range of values, i.e., $p_i \in P_i(i = 1, \ldots, v)$. A *primary MF* is any one of these type-1 MFs, e.g., $\mu_A(x|p_1 = p_{1'}, p_2 = p_{2'}, \ldots, p_v = p_{v'})$. For short, we use $\mu_A(x)$ to denote a primary MF. It will be subject to some restrictions on its parameters. The family of all primary MFs creates a FOU. ■

Two examples of very useful primary MFs are: Gaussian MF with uncertain mean and certain standard deviation, and Gaussian MF with certain mean and uncertain standard deviation.

Definition A-8: An *upper MF* and a *lower MF* are two type-1 MFs that are bounds for the FOU of a type-2 fuzzy set $\tilde{A}$. The upper MF is associated with the upper bound of $FOU(\tilde{A})$, and is denoted $\overline{\mu}_{\tilde{A}}(x), \forall x \in X$. The lower MF is associated with the lower bound of $FOU(\tilde{A})$, and is denoted $\underline{\mu}_{\tilde{A}}(x), \forall x \in X$, i.e.,

$$\overline{\mu}_{\tilde{A}}(x) \equiv \overline{FOU(\tilde{A})} \qquad \forall x \in X \tag{A-5}$$

and

$$\underline{\mu}_{\tilde{A}}(x) \equiv \underline{\mathrm{FOU}(\tilde{A})} \quad \forall x \in X \tag{A-6}$$

Because the domain of a secondary MF has been constrained in Definition A-1 to be contained in [0, 1], lower and upper MFs always exist. ■

Example A-1: Consider the case of a Gaussian primary MF having a fixed standard deviation, σ, and an uncertain mean that takes on values in $[m_1, m_2]$, i.e.,

$$\mu_A(x) = \exp\left[-\frac{1}{2}\left(\frac{x-m}{\sigma}\right)^2\right] \qquad m \in [m_1, m_2]. \tag{A-7}$$

Corresponding to each value of m we will get a different membership curve (Fig. 3). The uniform shading for the FOU denotes interval sets for the secondary MFs and represents the entire interval type-2 fuzzy set $\mu_{\tilde{A}}(x, u)$.

The upper MF, $\overline{\mu}_{\tilde{A}}(x)$, is

$$\overline{\mu}_{\tilde{A}}(x) = \begin{cases} N(m_1, \sigma; x) & x < m_1 \\ 1 & m_1 \le x \le m_2 \\ N(m_2, \sigma; x) & x > m_2 \end{cases} \tag{A-8}$$

where, for example, $N(m_1, \sigma; x) \equiv \exp[-(1/2)(x - m_1/\sigma)^2]$. The thick solid curve in Fig. 3 denotes the upper MF. The lower MF, $\underline{\mu}_{\tilde{A}}(x)$, is

$$\underline{\mu}_{\tilde{A}}(x) = \begin{cases} N(m_2, \sigma; x) & x \le \frac{m_1+m_2}{2} \\ N(m_1, \sigma; x) & x > \frac{m_1+m_2}{2} \end{cases}. \tag{A-9}$$

The thick dashed curve in Fig. 3 denotes the lower MF. ■

From this example we see that sometimes an upper (or a lower) MF cannot be represented by just one mathematical function over its entire x-domain. It may consist of several branches each defined over a different segment of the entire x-domain.[13] When the input, x, is located in a specific x-domain segment, we call its corresponding MF branch an *active branch* ([4], [6]), e.g., in (A-9), when $x > (m_1 + m_2)/2$, the active branch for $\underline{\mu}_{\tilde{A}}(x)$ is $N(m_1, \sigma; x)$.

B. Fuzzy Inference Engine Results

The major result for an interval singleton type-2 FLS is summarized in the following

Theorem A-1: [5], [6] In an interval singleton type-2 FLS using product or minimum t-norm: a) the result of the input and antecedent operations, is an interval type-1 set, called the *firing set*, i.e.,

$$F^l(\mathbf{x}') = \left[\underline{f}^l(\mathbf{x}'), \overline{f}^l(\mathbf{x}')\right] \equiv [\underline{f}^l, \overline{f}^l] \tag{A-10}$$

where

$$\underline{f}^l(\mathbf{x}') = \underline{\mu}_{\tilde{F}_1^l}(x_1') \star \cdots \star \underline{\mu}_{\tilde{F}_p^l}(x_p') \equiv \underline{\mu}_1^l(x_1) \star \cdots \star \underline{\mu}_p^l(x_p) \tag{A-11}$$

and

$$\overline{f}^l(\mathbf{x}') = \overline{\mu}_{\tilde{F}_1^l}(x_1') \star \cdots \star \overline{\mu}_{\tilde{F}_p^l}(x_p') \equiv \overline{\mu}_1^l(x_1) \star \cdots \star \overline{\mu}_p^l(x_p) \tag{A-12}$$

b) the rule R^l fired output consequent set, $\mu_{\tilde{B}^l}(y)$, is the interval type-2 fuzzy set

$$\mu_{\tilde{B}^l}(y) = \int_{b^l \in \left[\underline{f}^l \star \underline{\mu}_{\tilde{G}^l}(y), \overline{f}^l \star \overline{\mu}_{\tilde{G}^l}(y)\right]} 1/b^l, \qquad y \in Y \tag{A-13}$$

where $\underline{\mu}_{\tilde{G}^l}(y)$ and $\overline{\mu}_{\tilde{G}^l}(y)$ are the lower and upper membership grades of $\mu_{\tilde{G}^l}(y)$. (c) suppose that N of the M rules in the FLS fire, where $N \le M$, and the combined output type-1 fuzzy set, $\mu_{\tilde{B}}(y)$, is obtained by combining the fired output consequent sets by taking the union of the rule R^l fired output consequent sets; then (A-14), as shown at the bottom of the next page, holds.

A complete proof of this theorem can be found in [5] and [6]. Generalizations of this theorem to the very important case when the input to the type-2 FLS is a type-2 fuzzy set are also given in those references (see, also Appendix B for some of those results).

C. Type-Reduction

Five different type-reduction methods are described in [6]. Each is inspired by what we do in a type-1 FLS, when we defuzzify the (combined) output of the inference engine using a va-

[13]This is not peculiar to type-2 fuzzy sets. For example, a type-1 triangular MF consists of two branches each defined over a different segment of the domain variable.

riety of defuzzification methods that all do some sort of centroid calculation, and, is based on computing the *centroid of a type-2 fuzzy set*. Using the Extension Principle, Karnik and Mendel [2] defined the centroid of a type-2 fuzzy set; it is a type-1 fuzzy set. Computing the centroid of a general type-2 fuzzy set can be very intensive; however, for an interval type-2 fuzzy set, two exact iterative methods for computing its centroid have been developed by them. This was possible because the centroid of an interval type-2 fuzzy set is an interval type-1 fuzzy set, and such sets are completely characterized by their left- and right-end points; hence, computing the centroid of an interval type-2 fuzzy set only requires computing those two end-points.

The different kinds of type-reduction can all be expressed as

$$Y_{TR}(\mathbf{x}) = [y_l, y_r] = \int_{y^1 \in [y_l^1, y_r^1]} \cdots \int_{y^M \in [y_l^M, y_r^M]} \int_{f^1 \in [\underline{f}^1, \overline{f}^1]} \cdots \int_{f^M \in [\underline{f}^M, \overline{f}^M]} 1 \Big/ \frac{\sum_{i=1}^M f^i y^i}{\sum_{i=1}^M f^i} \tag{A-15}$$

in which y_l^i, y_r^i, $\underline{f}^i$, $\overline{f}^i$ and M have different meanings, as summarized in Table VI. In this paper, we only focus on center-of-sets type-reduction.

In (A-15), all symbols refer to quantities that are rule-ordered. In the Karnik–Mendel iterative procedures, (one for computing y_l and one for computing y_r) that we summarize next, all quantities are reordered according to step 1: The Karnik–Mendel iterative procedure for computing y_r is as follows:[14]

1) Without loss of generality, assume that the precomputed y_r^i are arranged in *ascending order*; i.e., $y_r^1 \le y_r^2 \le \cdots \le y_r^M$. Re-order the f_r^i accordingly and call them g_r^i.
2) Compute y_r as $y_r = \sum_{i=1}^M g_r^i y_r^i / \sum_{i=1}^M g_r^i$ by initially setting $g_r^i = (\underline{g}^i + \overline{g}^i)/2$ for $i = 1, \ldots, M$, where $\underline{g}^i$ and $\overline{g}^i$ have been previously computed using (A-11) and (A-12), respectively, and let $y_r' \equiv y_r$.
3) Find $R (1 \le R \le M-1)$ such that $y_r^R \le y_r' \le y_r^{R+1}$.
4) Compute $y_r = \sum_{i=1}^M g_r^i y_r^i / \sum_{i=1}^M g_r^i$ with $g_r^i = \underline{g}^i$ for $i \le R$ and $g_r^i = \overline{g}^i$ for $i > R$, and let $y_r'' \equiv y_r$.
5) If $y_r'' \ne y_r'$, then go to Step 6). If $y_r'' = y_r'$, then stop and set $y_r'' \equiv y_r$.
6) Set y_r' equal to y_r'', and return to Step 3).

The iterative procedure for computing y_l is very similar to the one just given for y_r. In step 1, it is now the precomputed y_l^i that are arranged in *ascending order*; i.e., $y_l^1 \le y_l^2 \le \cdots \le y_l^M$, and the f_l^i that are reordered accordingly (they are now called g_l^i). In step 2 y_l is computed as $y_l = \sum_{i=1}^M g_l^i y_l^i / \sum_{i=1}^M g_l^i$ by initially setting $g_r^i = (\underline{g}^i + \overline{g}^i)/2$ for $i = 1, \ldots, M$. In step 3, $L (1 \le L \le M-1)$ is found such that $y_l^L \le y_l' \le y_l^{L+1}$; and, in step 4) y_l is computed as $y_l = \sum_{i=1}^M g_l^i y_l^i / \sum_{i=1}^M g_l^i$ with $g_l^i = \overline{g}^i$ for $i \le L$ and $g_l^i = \underline{g}^i$ for $i > L$.

These two four-step iterative-procedures (steps 1) and 2) are initialization steps) have been proven by Karnik and Mendel [2] to converge to the exact solutions in no more than M iterations.

Observe that in these procedures, the computed numbers R and L (called *cross-over points* or *switch-points*) are very important. For $i \le R$, $g_r^i = \underline{g}^i$, whereas for $i > R$ $g_r^i = \overline{g}^i$; hence, y_r can be represented as

$$y_r = y_r \left(\underline{g}^1, \ldots, \underline{g}^R, \overline{g}^{R+1}, \ldots, \overline{g}^M, y_r^1, \ldots, y_r^M\right). \tag{A-16}$$

Additionally, $g_l^i = \overline{g}^i$ for $i \le L$ and $g_l^i = \underline{g}^i$ for $i > L$, so that y_l can be represented as

$$y_l = y_l \left(\overline{g}^1, \ldots, \overline{g}^L, \underline{g}^{L+1}, \ldots, \underline{g}^M, y_l^1, \ldots, y_l^M\right). \tag{A-17}$$

D. Defuzzification

Because Y_{TR} is an interval set, we defuzzify it using the average of y_l and y_r; hence, the defuzzified output of an interval singleton type-2 FLS is

$$y(\mathbf{x}) = f_{s2}(\mathbf{x}) = \frac{y_l + y_r}{2}. \tag{A-18}$$

Appendix B

Results for Interval Type-2 Nonsingleton Type-2 FLSs

In this appendix, we provide results comparable to those given in the main body of the paper, but for an interval type-2 nonsingleton type-2 FLS. For such a FLS not only are the rule antecedents and consequent characterized by interval type-2 fuzzy sets, but the inputs that activate the FLS are also interval type-2 fuzzy sets (a special case of which is a type-1 fuzzy set). We denote the MF for input x_i by $\mu_{\tilde{X}_i}(x_i)$, with lower and upper MFs $\underline{\mu}_{\tilde{X}_i}(x_i)$ and $\bar{\mu}_{\tilde{X}_i}(x_i)$, respectively.

A. Fuzzy Inference Engine Results

The major results for an interval type-2 nonsingleton type-2 FLS are in [5] and are also summarized in of [6, Th. 12–1]. For the purposes of this paper, we only need the following procedure to compute $F^l(\mathbf{x}') = [\underline{f}^l(\mathbf{x}'), \overline{f}^l(\mathbf{x}')] \equiv [\underline{f}^l, \overline{f}^l]$.

1) Choose a t-norm (product or minimum) and create the functions $\underline{\mu}_{\tilde{Q}_k^l}(x_k)$ and $\overline{\mu}_{\tilde{Q}_k^l}(x_k)$, where

$$\underline{\mu}_{\tilde{Q}_k^l}(x_k) = \int_{x_k \in X_k} \left[\underline{\mu}_{\tilde{X}_k}(x_k) \star \underline{\mu}_{\tilde{F}_k^l}(x_k)\right] / x_k \tag{B-1}$$

[14]This procedure is a special case of computing a fractionally linear function [3]; however, it was developed independently of their work.

$$\mu_{\tilde{B}}(y) = \int_{b \in \left[\left[\underline{f}^1 \star \underline{\mu}_{\tilde{G}^1}(y)\right] \vee \cdots \vee \left[\underline{f}^N \star \underline{\mu}_{\tilde{G}^N}(y)\right], \left[\overline{f}^1 \star \overline{\mu}_{\tilde{G}^1}(y)\right] \vee \cdots \vee \left[\overline{f}^N \star \overline{\mu}_{\tilde{G}^N}(y)\right]\right]} 1/b, \qquad y \in Y. \quad \blacksquare \tag{A-14}$$

and

$$\overline{\mu}_{\tilde{Q}_k^l}(x_k) = \int_{x_k \in X_k} \left[\overline{\mu}_{\tilde{X}_k}(x_k) \star \overline{\mu}_{\tilde{F}_k^l}(x_k)\right] / x_k. \quad \text{(B-2)}$$

2) Let $\underline{x}_{k,\max}^l$ and $\bar{x}_{k,\max}^l$ denote the values of x_k that are associated with $\sup_{x_k} \underline{\mu}_{\tilde{Q}_k^l}(x_k)$ and $\sup_{x_k} \overline{\mu}_{\tilde{Q}_k^l}(x_k)$, respectively. Compute $\underline{x}_{k,\max}^l$ and $\bar{x}_{k,\max}^l$.

3) Evaluate $\underline{f}_k^l(x_k')$ and $\overline{f}_k^l(x_k')$ where

$$\underline{f}_k^l(x_k') = \underline{\mu}_{\tilde{Q}_k^l}\left(\underline{x}_{k,\max}^l\right) \quad \text{(B-3)}$$

and

$$\overline{f}_k^l(x_k') = \overline{\mu}_{\tilde{Q}_k^l}\left(\overline{x}_{k,\max}^l\right). \quad \text{(B-4)}$$

Note that $\underline{x}_{k,\max}^l$ and $\bar{x}_{k,\max}^l$ will depend upon measurement x_k'.

4) Compute $\underline{f}^l(\mathbf{x}')$ and $\overline{f}^l(\mathbf{x}')$ as

$$\underline{f}^l(\mathbf{x}') = T_{k=1}^p \underline{f}_k^l(x_k') = T_{k=1}^p \underline{\mu}_{\tilde{Q}_k^l}\left(\underline{x}_{k,\max}^l\right) \quad \text{(B-5)}$$

and

$$\overline{f}^l(\mathbf{x}') = T_{k=1}^p \overline{f}_k^l(x_k') = T_{k=1}^p \overline{\mu}_{\tilde{Q}_k^l}\left(\overline{x}_{k,\max}^l\right). \quad \text{(B-6)}$$

Note that these results are comparable to those in part a) of Theorem A-1. The main difference between computing the firing interval for an interval type-2 nonsingleton type-2 FLS and an interval singleton type-2 FLS is having to compute $\underline{x}_{k,\max}^l$ and $\bar{x}_{k,\max}^l$. Note also that parts b) and c) of Theorem A-1 apply as is to the present nonsingleton case.

B. Changes

The changes in Sections I–V are as follows.

1) In Section I, change $f_{s2}(\mathbf{x}^{(t)})$ in (2) to $f_{ns2-2}(\mathbf{x}^{(t)})$, where "$ns2-2$" is short for *type-2 nonsingleton type-2*.
2) In all sections, wherever the phrase "interval singleton type-2 FLS" is used, replace it with "interval type-2 nonsingleton type-2 FLS."
3) There are no changes to Sections II and III.
4) **Fact 4** is changed to: For specific values of k and l, antecedent parameters $\theta_{k,m}^l$ can only appear in $\underline{\mu}_{\tilde{Q}_k^l}(\underline{x}_{k,\max}^l)$ and $\overline{\mu}_{\tilde{Q}_k^l}(\overline{x}_{k,\max}^l)$ and cannot appear in $\underline{\mu}_{\tilde{Q}_j^l}(\underline{x}_{j,\max}^l)$ and $\overline{\mu}_{\tilde{Q}_j^l}(\overline{x}_{j,\max}^l)$ for $j \neq k$. This is a direct result of Assumption 1 in Section II and the facts that $\underline{x}_{k,\max}^l$ and $\overline{x}_{k,\max}^l$ only depend on one value of k and l.
5) In **Fact 5**, change $f_{s2}(\mathbf{x}^{(t)})$ in (28) to $f_{ns2-2}(\mathbf{x}^{(t)})$.
6) In **Fact 6**, replace $\underline{\mu}_i^l(x_i)$ by $\underline{\mu}_{\tilde{Q}_i^l}(\underline{x}_{i,\max}^l)$ and $\bar{\mu}_i^l(x_i)$ by $\overline{\mu}_{\tilde{Q}_i^l}(\overline{x}_{i,\max}^l)$ for all $i = 1, \ldots, M$. Although we will still refer to the equations for computing $\partial \underline{f}^l(x_k^{(t)})/\partial\theta_{k,m}^l$ and $\partial \bar{f}^l(x_k^{(t)})/\partial\theta_{k,m}^l$ as (29) and (30), respectively, it is to be understood that those equations are now modulo these changes.
7) The flowchart in Fig. 2 is modified as follows.
   a. Replace the top four blocks of Fig. 2 by the blocks in Fig. 4.
   b. Change $f_{s2}(\mathbf{x}^{(t)})$ in the center lower two blocks to $f_{ns2-2}(\mathbf{x}^{(t)})$.
   c. In the block for computing $\underline{f}^l(\mathbf{x}^{(t)})$, replace "(A-11)" by "(B-5), in which $\mathbf{x}' \equiv \mathbf{x}^{(t)}$."
   d. In the block for computing $\bar{f}^l(\mathbf{x}^{(t)})$, replace "(A-12)" by "(B-6), in which $\mathbf{x}' \equiv \mathbf{x}^{(t)}$."
8) In Section V, change $f_{s2}(\mathbf{x}^{(t)})$ in (31) and (32) to $f_{ns2-2}(\mathbf{x}^{(t)})$.

C. An Example

In order to compute $\partial \underline{f}^l/\partial\theta_{k,m}^l$ and $\partial \overline{f}^l/\partial\theta_{k,m}^l$ using (29) and (30), we need $\partial \underline{\mu}_{\tilde{Q}_k^l}(\underline{x}_{k,\max}^l)/\partial\theta_{k,m}^l$ and $\partial \overline{\mu}_{\tilde{Q}_k^l}(\overline{x}_{k,\max}^l)/\partial\theta_{k,m}^l$. Here, we will compute $\partial \underline{\mu}_{\tilde{Q}_k^l}(\underline{x}_{k,\max}^l)/\partial\theta_{k,m}^l$ and $\partial \overline{\mu}_{\tilde{Q}_k^l}(\overline{x}_{k,\max}^l)/\partial\theta_{k,m}^l$ for antecedent *Gaussian primary MFs with uncertain means* and input measurement *Gaussian primary MFs with uncertain standard deviations*. Formulas for antecedent MFs and their upper and lower MFs are given in (37)–(39). Formulas for input measurement MFs and their upper and lower MFs are

$$\mu_k(x_k) = \exp\left\{-\frac{1}{2}\left(\frac{x_k - x_k'}{\sigma_k}\right)^2\right\} \quad \sigma_k \in [\sigma_{k1}, \sigma_{k2}] \quad \text{(B-7)}$$

$$\overline{\mu}_k(x_k) = N(x_k', \sigma_{k2}; x_k) \quad \text{(B-8)}$$

$$\underline{\mu}_k(x_k) = N(x_k', \sigma_{k1}; x_k) \quad \text{(B-9)}$$

Results for $\underline{\mu}_{\tilde{Q}_k^l}(\underline{x}_{k,\max}^l)$ and $\overline{\mu}_{\tilde{Q}_k^l}(\bar{x}_{k,\max}^l)$ are given in Tables VII and VIII. Detailed derivations of these results can be found for product t-norm, in [6, pp. 394–399]. The results for minimum t-norm are derived in exactly the same manner as the ones for product t-norm. We summarize the parameters that $\underline{\mu}_{\tilde{Q}_k^l}(\underline{x}_{k,\max}^l)$ and $\overline{\mu}_{\tilde{Q}_k^l}(\bar{x}_{k,\max}^l)$ depend upon, as a function of x_k, in Table IX. Its results were obtained by examining Table VII.

Let $\theta_{k,1}^l \equiv m_{k1}^l$, $\theta_{k,2}^l \equiv m_{k2}^l$, $\theta_{k,3}^l \equiv \sigma_k^l$, $\theta_{k,4} \equiv \sigma_{k1}$ and $\theta_{k,5} \equiv \sigma_{k2}$. From Table IX observe that $\underline{\mu}$ depends on σ_{k1} in all five of its states, and $\overline{\mu}$ depends on σ_{k2} only in States (1) and (5). No state depends on both σ_{k1} and σ_{k2}. Tables X and XI provide *roadmaps* of nonzero or zero derivatives of $\underline{\mu}_{\tilde{Q}_k^l}(\underline{x}_{k,\max}^l) \equiv \underline{\mu}$ and $\overline{\mu}_{\tilde{Q}_k^l}(\overline{x}_{k,\max}^l) \equiv \overline{\mu}$. You can compute the exact derivatives by using the formulas in Table VII.

The results in Tables VIII, X, and XI would be used in Figs. 2 and 4 as follows. The tests on variable $x_k' \equiv x_k^{(t)}$ in VIII let us implement the top block of Fig. 4 in which we have to determine the active states. The results in Table X provide $\partial \underline{\mu}_{\tilde{Q}_k^l}(\underline{x}_{k,\max}^l)/\partial\theta_{k,m}^l$ which is needed to compute $\partial \underline{f}^l(x_k^{(t)})/\partial\theta_{k,m}^l$, and the results in Table V provide $\partial \overline{\mu}_{\tilde{Q}_k^l}(\overline{x}_{k,\max}^l)/\partial\theta_{k,m}^l$ which is needed to compute $\partial \bar{f}^l(x_k^{(t)})/\partial\theta_{k,m}^l$.

REFERENCES

[1] N. N. Karnik and J. M. Mendel, "Operations on type-2 fuzzy sets," *Fuzzy Sets Syst.*, vol. 122, pp. 327–348, 2001.

[2] ——, "Centroid of a type-2 fuzzy set," *Inform. Sci.*, vol. 132, pp. 195–220, 2001.

[3] V. Kreinovich, A. Lakeyev, J. Rohn, and P. Kahl, *Computational Complexity and Feasibility of Data Processing and Interval Computations*. Dordrecht, The Netherlands: Kluwer, 1997, ch. 10.

[4] Q. Liang and J. M. Mendel, "Interval type-2 fuzzy logic systems," presented at the FUZZ-IEEE '00, San Antonio, TX, May 2000.

[5] ——, "Interval type-2 fuzzy logic systems: theory and design," *IEEE Trans. Fuzzy Syst.*, vol. 8, pp. 535–550, Aug. 2000.

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

[7] ——, "On the importance of interval sets in type-2 fuzzy logic systems," presented at the Joint 9th IFSA World Congr. 20th NAFIPS In. Conf., Vancouver, BC, Canada, July 2001.

[8] L.-X. Wang and J. M. Mendel, "Fuzzy basis functions, universal approximation, and orthogonal least-squares learning," *IEEE Trans. Neural Networks*, vol. 3, pp. 807–813, Sept. 1992.

Jerry M. Mendel (S'59–M'61–SM'72–F'78–LF'04) received the Ph.D. degree in electrical engineering from the Polytechnic Institute of Brooklyn, Brooklyn, NY.

Currently, he is Professor of Electrical Engineering and Associate Director for Education and Outreach of the Integrated Media Systems Center at the University of Southern California, Los Angeles, where he has been since 1974. He has published over 430 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 target classification and computing with words, and spatio–temporal fusion of decisions.

Dr. Mendel is a Distinguished Member of the IEEE Control Systems Society. He was President of the IEEE Control Systems Society in 1986, and is presently Chairman of the Fuzzy Technical Committee and a Member of the Aminstrative Committee of the IEEE Neural Networks Society. 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 Millennium Medal.