Centroid Uncertainty Bounds for Interval Type-2 Fuzzy Sets: Forward and Inverse Problems

Jerry M. Mendel and Hongwei Wu
Signal and Image Processing Institute
Department of Electrical Engineering
University of Southern California
Los Angeles, CA 90089-2564
E-mail: mendel@sipi.usc.edu

***Abstract* - Interval type-2 fuzzy sets (T2 FS) play a central role in fuzzy sets as models for words [6] and in engineering applications of T2 FSs [5]. These fuzzy sets are characterized by their footprints of uncertainty (FOU), which in turn are characterized by their boundaries—upper and lower membership functions (MF). The centroid of an interval T2 FS [3], which is an interval T1 FS, provides a measure of the uncertainty in the interval T2 FS. Intuitively, we anticipate that geometric properties about the FOU, such as its area and the center of gravities (centroids) of its upper and lower MFs, will be associated with the amount of uncertainty in an interval T2 FS. The main purpose of this paper is to demonstrate that our intuition is correct and to quantify the centroid of an interval T2 FS with respect to these geometric properties of its FOU. It is then possible to formulate and solve inverse problems, i.e. going from data to parametric T2 FS models.**

I. INTRODUCTION

Recently, Mendel [6] proposed a fuzzy set (FS) model for words that is based on collecting data from people—*person membership functions (MFs)*—that reflect *intra-* and *inter-levels of uncertainties* about a word, in which a word FS is the union of all such person FSs. The *intra-uncertainty* about a word is modeled using interval type-2 (T2) person FSs, and the *inter-uncertainty* about a word is modeled using an equally weighted union of each person's interval T2 FS. Because an interval T2 FS plays such an important role in this model as well as in engineering applications of T2 FSs (e.g., [5]), we need to understand as much as possible about such sets and how they model uncertainties.

Recall that an *interval T2 FS* $\tilde{A}$ is characterized as [5], [8]:

$$\tilde{A} = \int_{x\in X}\int_{u\in J_x\subseteq[0,1]} 1/(x,u) = \int_{x\in X}\left[\int_{u\in J_x\subseteq[0,1]} 1/u\right]\Big/ x \qquad (1)$$

where x, the *primary variable*, has domain X; u, the *secondary variable*, has domain J_x at each $x \in X$; J_x is called the *primary membership* of x; and, the *secondary grades* of $\tilde{A}$ all equal 1. Uncertainty about $\tilde{A}$ is conveyed by the union of all of the primary memberships, which is called the *footprint of uncertainty* (FOU) of $\tilde{A}$, i.e.

$$FOU(\tilde{A}) = \bigcup_{x\in X} J_x \qquad (2)$$

The *upper membership function* (UMF) and *lower membership function* (LMF) of $\tilde{A}$ are two type-1 MFs that bound the FOU (e.g., see Fig. 5). The UMF is associated with the upper bound of $FOU(\tilde{A})$ and is denoted $\bar{\mu}_{\tilde{A}}(x)$, $\forall x \in X$, and the LMF is associated with the lower bound of $FOU(\tilde{A})$ and is denoted $\underline{\mu}_{\tilde{A}}(x)$, $\forall x \in X$, i. e.

$$\bar{\mu}_{\tilde{A}}(x) \equiv \overline{FOU(\tilde{A})} \quad \forall x \in X \qquad (3)$$

$$\underline{\mu}_{\tilde{A}}(x) \equiv \underline{FOU(\tilde{A})} \quad \forall x \in X \qquad (4)$$

The centroid of an interval T2 FS [2], which is an interval T1 FS, provides a measure of the uncertainty in the interval T2 FS. Intuitively, we anticipate that geometric properties about the FOU, such as its area and the center of gravities (centroids) of its upper and lower MFs, will be associated with the amount of uncertainty in an interval T2 FS.

The main purposes of this paper are to demonstrate that our intuition is correct, to quantify the centroid of an interval T2 FS with respect to these geometric properties of its FOU, and to then formulate and solve inverse problems, i.e. going from data to parametric T2 FS models.

II. CENTROID OF AN INTERVAL TYPE-2 FUZZY SET

Recall that the centroid, $C_{\tilde{A}}$, of the interval T2 FS $\tilde{A}$ is an interval set $[c_l, c_r]$ that is completely specified by its left and right end-points, c_l and c_r, respectively, i.e. [3], [5]

$$C_{\tilde{A}} = [c_l, c_r] = \int_{\theta_1\in J_{x_1}}\cdots\int_{\theta_N\in J_{x_N}} 1 \Big/ \frac{\sum_{i=1}^{N} x_i\theta_i}{\sum_{i=1}^{N}\theta_i} \qquad (5)$$

In this equation, primary variable x has been discretized for computational purposes, such that $x_1 < x_2 < \cdots < x_N$. Unfortunately, no closed-form formulas exist to compute c_l and c_r; however, Karnik and Mendel [3] have developed iterative procedures for computing these end-points, and recently Mendel [7] proved that given a FOU for an interval T2 FS, one that is *symmetrical* about primary variable x at $x = m$, then the centroid of such a T2 FS is also symmetrical about $x = m$. For such a FS it is therefore only necessary to compute either c_l or c_r, resulting in a 50% savings in computation.

Before we summarize the Karnik-Mendel procedures in a form that will be very useful to us, we must first justify the use of the length $c_r - c_l$ as a legitimate measure of the uncertainty of $\tilde{A}$. Wu and Mendel [9] noted that according to Information Theory uncertainty of a random variable is measured by its entropy [2]. 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 MF, $\mu_C(x)$, of an interval FS *C*, where

$$\mu_C(x) = \begin{cases} 1, & x \in [c_l, c_r] \\ 0, & \text{otherwise} \end{cases}, \tag{6}$$

with the probability density function, $p_Y(y)$, of a random variable *Y*, which is uniformly distributed over $[c_l, c_r]$, where

$$p_Y(y) = \begin{cases} 1/(c_r - c_l), & y \in [c_l, c_r] \\ 0, & \text{otherwise} \end{cases}, \tag{7}$$

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 FS *C* to be the same as (or proportional to) that of the random variable *Y*. Since the centroid of a T2 FS is an interval set, its length can therefore be used to measure the extent of the T2 FS's uncertainty.

In the sequel, when we use sampled values of $\underline{\mu}_{\tilde{A}}(x)$ and $\bar{\mu}_{\tilde{A}}(x)$, namely $\underline{\mu}_{\tilde{A}}(x_i)$ and $\bar{\mu}_{\tilde{A}}(x_i)$, where $i = 1, 2, ..., N$, we shall simplify our notation, i.e., without loss of generality

$$\underline{\mu}_{\tilde{A}}(x_i) \equiv \underline{\mu}_i \quad i = 1, ..., N \tag{8a}$$

$$\bar{\mu}_{\tilde{A}}(x_i) \equiv \bar{\mu}_i \quad i = 1, ..., N \tag{8b}$$

The Karnik-Mendel iterative procedures for computing c_l and c_r can be interpreted for the purposes of this paper as follows [9]. Define $c^{(L)}$ and $c^{(R)}$, for $0 \le L, R \le N$, as

$$c^{(L)} \equiv \left[\sum_{i=1}^{L} x_i \bar{\mu}_i + \sum_{j=L+1}^{N} x_j \underline{\mu}_j \right] \Big/ \left[\sum_{i=1}^{L} \bar{\mu}_i + \sum_{j=L+1}^{N} \underline{\mu}_j \right] \tag{9}$$

$$c^{(R)} \equiv \left[\sum_{i=1}^{R} x_i \underline{\mu}_i + \sum_{j=R+1}^{N} x_j \bar{\mu}_j \right] \Big/ \sum_{i=1}^{R} \underline{\mu}_i + \sum_{j=R+1}^{N} \bar{\mu}_j \tag{10}$$

The end-points c_l and c_r for the centroid of an interval T2 FS [given by (5)] are the minimum of all $c^{(L)}$ and the maximum of all $c^{(R)}$, respectively, i.e.[1]

$$c_l = \min_{0 \le L \le N} \left\{ c^{(L)} \right\} = c^{(L^*)} = \left[\sum_{i=1}^{L^*} x_i \bar{\mu}_i + \sum_{j=L^*+1}^{N} x_j \underline{\mu}_j \right] \Big/ \left[\sum_{i=1}^{L^*} \bar{\mu}_i + \sum_{j=L^*+1}^{N} \underline{\mu}_j \right] \tag{11}$$

where

$$L^* = \arg \min_{0 \le L \le N} \left\{ c^{(L)} \right\} \tag{12}$$

and

$$c_r = \max_{0 \le R \le N} \left\{ c^{(R)} \right\} = c^{(R^*)} = \left[\sum_{i=1}^{R^*} x_i \underline{\mu}_i + \sum_{j=R^*+1}^{N} x_j \bar{\mu}_j \right] \Big/ \left[\sum_{i=1}^{R^*} \underline{\mu}_i + \sum_{j=R^*+1}^{N} \bar{\mu}_j \right] \tag{13}$$

where

$$R^* = \arg \max_{0 \le R \le N} \left\{ c^{(R)} \right\} \tag{14}$$

The solutions of (12) and (14), L^* and R^*, are obtained using the Karnik-Mendel iterative procedures, the details of which are not needed in the rest of this paper.

Because closed-form formulas do not exist for c_l and c_r, it is impossible to study how these end-points explicitly depend upon the area of the FOU and the centroids of the upper and lower MFs of the FOU. The approach taken in the rest of this paper is to obtain bounds for both c_l and c_r, and to then examine the explicit dependencies of these bounds on the geometric properties of the FOU.

III. Bounds on c_l and c_r for an Arbitrary FOU

Theorem 1: *The end-points, c_l and c_r, for the centroid of an interval T2 FS are bounded from below and above by (Fig. 1)*

$$\underline{c}_l \le c_l \le \bar{c}_l \tag{15}$$

$$\underline{c}_r \le c_r \le \bar{c}_r \tag{16}$$

where

$$\bar{c}_l = \min\left\{ c_{LMF}, c_{UMF} \right\} \tag{17}$$

$$\underline{c}_r = \max\left\{ c_{LMF}, c_{UMF} \right\} \tag{18}$$

$$c_{LMF} = \sum_{i=1}^{N} x_i \underline{\mu}_i \Big/ \sum_{i=1}^{N} \underline{\mu}_i \tag{19}$$

$$c_{UMF} = \sum_{i=1}^{N} x_i \bar{\mu}_i \Big/ \sum_{i=1}^{N} \bar{\mu}_i \tag{20}$$

$$\underline{c}_l = \bar{c}_l - \frac{\sum_{i=1}^{N} \left(\bar{\mu}_i - \underline{\mu}_i \right)}{\sum_{i=1}^{N} \bar{\mu}_i \sum_{i=1}^{N} \underline{\mu}_i} \times \frac{\sum_{i=1}^{N} \underline{\mu}_i (x_i - x_1) \sum_{i=1}^{N} \bar{\mu}_i (x_N - x_i)}{\sum_{i=1}^{N} \underline{\mu}_i (x_i - x_1) + \sum_{i=1}^{N} \bar{\mu}_i (x_N - x_i)} \tag{21}$$

$$\bar{c}_r = \underline{c}_r + \frac{\sum_{i=1}^{N} \left(\bar{\mu}_i - \underline{\mu}_i \right)}{\sum_{i=1}^{N} \bar{\mu}_i \sum_{i=1}^{N} \underline{\mu}_i} \times \frac{\sum_{i=1}^{N} \bar{\mu}_i (x_i - x_1) \sum_{i=1}^{N} \underline{\mu}_i (x_N - x_i)}{\sum_{i=1}^{N} \bar{\mu}_i (x_i - x_1) + \sum_{i=1}^{N} \underline{\mu}_i (x_N - x_i)} \tag{22}$$

Proof: Provided in the journal version of this paper.

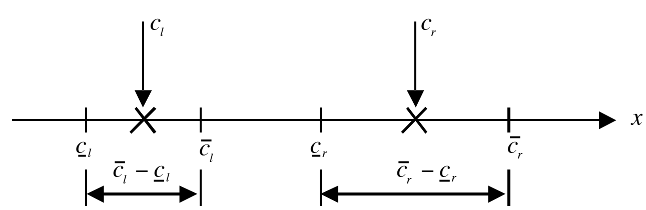

Fig. 1. End-points (X) of the centroid of $\tilde{A}$ and the lower and upper bounds (|) for the two end-points.

---

[1] These theoretical facts are established in [3] and [5].

Next, we re-express the *uncertainty bounds* $\bar{c}_l - \underline{c}_l$ and $\bar{c}_r - \underline{c}_r$, that are obtained from (21) and (22), respectively, in a way that provides enormous insights into these intervals.

Theorem 2: *Let A_{UMF}, A_{LMF}, A_{FOU}, c_{LMF} and c_{UMF} denote the area under the upper MF, the area under the lower MF, the area of the FOU (note that $A_{FOU} = A_{UMF} - A_{LMF}$), the centroid of the lower MF, and the centroid of the upper MF. Then*

$$\bar{c}_l - \underline{c}_l = A_{FOU} \frac{(c_{LMF} - x_1)(x_N - c_{UMF})}{A_{LMF}(c_{LMF} - x_1) + A_{UMF}(x_N - c_{UMF})} \quad (23)$$

$$\bar{c}_r - \underline{c}_r = A_{FOU} \frac{(c_{UMF} - x_1)(x_N - c_{LMF})}{A_{UMF}(c_{UMF} - x_1) + A_{LMF}(x_N - c_{LMF})} \quad (24)$$

Proof: Multiply the numerator and denominators of (21) and (22) each by three Δx terms, and then take the limit as $\Delta x \to 0$. The results in (23) and (24) follow immediately. ■

Comment 1: Theorem 2 demonstrates that the bounding intervals (uncertainty intervals) for the end-points of the centroid of $\tilde{A}$ are indeed expressible in terms of geometric properties of the FOU. It has not made use of any a priori geometric knowledge about the FOU, e.g., the FOU is symmetric; hence its results are most general. Because it has not made use of a priori geometric knowledge, its results may be improved upon by making use of such information. We explore this further in Section V. ■

Theorem 2 lets us obtain many new results about the uncertainty bounds.

Corollary 1: *$\bar{c}_l - \underline{c}_l$ and $\bar{c}_r - \underline{c}_r$ are* shift-invariant.

Proof: The proof for $\bar{c}_l - \underline{c}_l$ focuses on the two factors $(c_{LMF} - x_1)$ and $(x_N - c_{UMF})$ which appear in (23). When the FOU is shifted, $x \to x + m$ in which case $x_1 \to x_1 + m$, $x_N \to x_N + m$, $c_{LMF} \to c_{LMF} + m$ and $c_{UMF} \to c_{UMF} + m$. Consequently, (23) remains unchanged when $x \to x + m$. A similar argument demonstrates that (24) remains unchanged when $x \to x + m$. ■

Comment 2: The results in Corollary 1 mean that we obtain the same *centroid bounds* for a specific FOU regardless of where that FOU is located with respect to its primary variable *(x)*. Of course, we would have hoped/ expected this to be true, and in this corollary our hope/ expectation is mathematically proved. Because of this shift-invariance we can locate the FOU anywhere we choose to on its *x*-axis. ■

Mendel [4], [5] has collected interval end-point data from people about words[2], and has observed that the uncertainty intervals about the left and right-hand end-points are unequal. A non-symmetrical FOU can provide such unequal intervals, whereas (see Corollaries 4 and 5) a symmetrical FOU cannot.

For a *non-symmetrical* FOU, $c_{UMF} \neq c_{LMF}$, and it is useful to express both c_{UMF} and c_{LMF} as functions of how much they each depart from the centroid, $(x_1 + x_N)/2$, of a symmetrical MF. Letting δ_{UMF} and δ_{LMF} denote the departures from symmetry for c_{UMF} and c_{LMF}, respectively, we can express c_{UMF} and c_{LMF} as:

$$c_{UMF} = (x_1 + x_N)/2 + \delta_{UMF} \quad (25)$$

$$c_{LMF} = (x_1 + x_N)/2 + \delta_{LMF} \quad (26)$$

Because $x_1 \le c_{UMF} \le x_N$ and $x_1 \le c_{LMF} \le x_N$, it follows from (25) and (26) that δ_{UMF} and δ_{LMF} are constrained as

$$-(x_N - x_1)/2 \le \delta_{LMF} \le (x_N - x_1)/2 \quad (27)$$

$$-(x_N - x_1)/2 \le \delta_{UMF} \le (x_N - x_1)/2 \quad (28)$$

Corollary 2: *An alternative way to express $\bar{c}_l - \underline{c}_l$ and $\bar{c}_r - \underline{c}_r$ is:*

$$\bar{c}_l - \underline{c}_l = A_{FOU}\left[\frac{2A_{LMF}}{(x_N - x_1) - 2\delta_{UMF}} + \frac{2A_{UMF}}{(x_N - x_1) + 2\delta_{LMF}}\right]^{-1} \quad (29)$$

$$\bar{c}_r - \underline{c}_r = A_{FOU}\left[\frac{2A_{UMF}}{(x_N - x_1) - 2\delta_{LMF}} + \frac{2A_{LMF}}{(x_N - x_1) + 2\delta_{UMF}}\right]^{-1} \quad (30)$$

Proof: Substitute (25) and (26) into (23) and (24).■

Corollary 3: *For an interval T2 FS $\bar{c}_l - \underline{c}_l$ is greater than, equal to, or less than $\bar{c}_r - \underline{c}_r$ if*

$$\frac{\delta_{LMF}}{A_{LMF}\left[(x_N - x_1)^2 - 4\delta_{LMF}^2\right]} \gtreqless \frac{\delta_{UMF}}{A_{UMF}\left[(x_N - x_1)^2 - 4\delta_{UMF}^2\right]} \quad (31)$$

Proof: Eq. (31) follows from (29) and (30) and some simple arithmetic manipulations. ■

Example 1: Special cases of (31) occur when: (1) $\delta_{LMF} > 0 > \delta_{UMF}$ in which case $\bar{c}_l - \underline{c}_l > \bar{c}_r - \underline{c}_r$ (see Fig. 2); and, (b) $\delta_{LMF} < 0 < \delta_{UMF}$, in which case $\bar{c}_l - \underline{c}_l < \bar{c}_r - \underline{c}_r$ (see Fig. 3). ■

It is interesting to study (31) to establish curves above which the > inequality is true and below which the < inequality is true. After a lot of analysis, one can show that $\bar{c}_l - \underline{c}_l > \bar{c}_r - \underline{c}_r$ if: (a) $\Delta_{LMF} > 0$ when $\delta_{UMF} = 0$, or (b)

[2] A group of students were asked the question: "Below are a number of labels that describe an interval or a 'range' that falls somewhere between 0 to 10. For each label, please tell us where this range would start and where it would end." This was done for two collections of 16 and five labels using two different groups of students. See Table 2-2 and Fig. 2-1 in [5] for a summary of results for the 16 labels, and Table 2-3 for a summary of results for the five labels.

$$\Delta_{LMF} > \frac{-(1-4\Delta_{UMF}^2) + \sqrt{(1-4\Delta_{UMF}^2)^2 + 16\left(\frac{A_{LMF}}{A_{UMF}}\right)^2 \Delta_{UMF}^2}}{8\Delta_{UMF}\frac{A_{LMF}}{A_{UMF}}} \quad (32)$$

when $\delta_{UMF} \neq 0$. In (32), $\Delta_{UMF} \equiv \delta_{UMF} / (x_N - x_1)$ and $\Delta_{LMF} \equiv \delta_{LMF} / (x_N - x_1)$. For $\bar{c}_l - \underline{c}_l \leq \bar{c}_r - \underline{c}_r$, change the inequality in (32) from > to ≤. Plots of (32) for $\delta_{UMF} \neq 0$ and five values of A_{LMF} / A_{UMF} are depicted in Fig. 4. Above each curve, $\bar{c}_l - \underline{c}_l > \bar{c}_r - \underline{c}_r$, whereas below each curve $\bar{c}_l - \underline{c}_l < \bar{c}_r - \underline{c}_r$. How to use these general results to design or reconstruct a non-symmetrical FOU from data are topics that are presently under study.

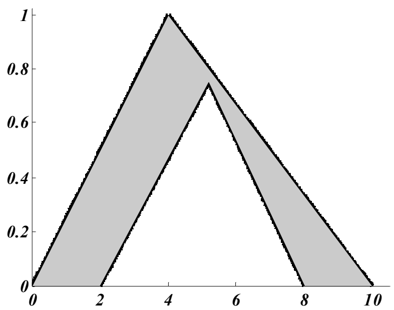

Fig. 2. Non-symmetrical triangular FOU for which $\delta_{UMF} < \delta_{LMF}$

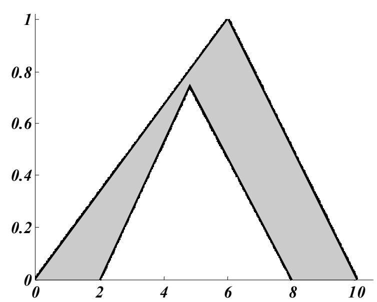

Fig. 3. Non-symmetrical triangular FOU for which $\delta_{UMF} > \delta_{LMF}$.

V. Bounds on c_l and c_r for a Symmetric FOU

Interval T2 FSs with symmetrical FOUs have been very widely used by practitioners of T2 FSs (e.g., [4]). Simplifications to (23) and (24) occur for such FOUs.

Corollary 4: *For a* symmetrical *FOU: (a)* $\bar{c}_l = \underline{c}_r = 0$ *and (b)* $|\bar{c}_l - \underline{c}_l| = |\bar{c}_r - \underline{c}_r| \equiv \Delta c$ *where*

$$\Delta c = |x_1| \left[A_{FOU} / (A_{LMF} + A_{UMF}) \right] \quad (33)$$

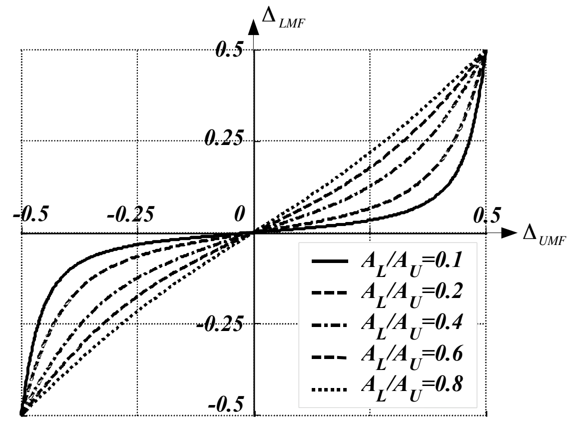

Fig. 4. Universal curves of (38). Note that $A_{LMF} \equiv A_L$ and $A_{UMF} \equiv A_U$

Proof: (a) Shifting the FOU so that it is symmetrical about the origin, it is clear that for a symmetrical FOU, $c_{LMF} = c_{UMF} = 0$; hence, $\bar{c}_l = \underline{c}_r = 0$ follows directly from (17) and (18). (b) Substituting $c_{LMF} = c_{UMF} = 0$ into (23) and (24), we find that:

$$\bar{c}_l - \underline{c}_l = A_{FOU} \left[-x_1 x_N / (-x_1 A_{LMF} + x_N A_{UMF}) \right] \quad (34)$$

$$\bar{c}_r - \underline{c}_r = A_{FOU} \left[-x_1 x_N / (-x_1 A_{UMF} + x_N A_{LMF}) \right] \quad (35)$$

For the shifted FOU, symmetry also means $x_N = -x_1$; hence, (34) and (35) reduce to the same number $-A_{FOU} \left[x_1 / (A_{LMF} + A_{UMF}) \right]$. Because $x_1 < 0$, we take the absolute value of $\bar{c}_l - \underline{c}_l$ and $\bar{c}_r - \underline{c}_r$ to obtain the results in (33). ■

Comment 3: For a symmetrical FOU, because $\bar{c}_l = \underline{c}_r = 0$, the results from our bounding analysis have degenerated into an outer-bound set that bounds the centroid $[c_l, c_r]$, i.e.

$$[c_l, c_r] \subseteq [\underline{c}_l, \bar{c}_r] = [-\Delta c, \Delta c] \quad (36)$$

Such a set may be too conservative. ■

Comment 4: Corollary 4 cannot be used as is for a symmetrical Gaussian FOU, because for such a FOU $|x_1| \to \infty$. This represents yet another shortcoming of trying to use general results for symmetrical FOUs. ■

Recall that the uncertainty bounds in Theorem 1 made no a priori use of the symmetry of a symmetrical FOU. When we do make use of such knowledge, we obtain:

Theorem 3: *Let* $\tilde{A}$ *be an interval T2 FS defined on X whose FOU is symmetrical about* $m \in X$. *Let* c_{HLMF} *and* c_{HUMF} *denote the centroids of half of the (symmetrical) lower and upper MFs, respectively, i.e.*

$$c_{HLMF} = \int_m^\infty x \underline{\mu}(x) dx \Big/ \int_m^\infty \underline{\mu}(x) dx \quad (37)$$

$$c_{HUMF} = \int_m^{\infty} x\overline{\mu}(x)\,dx \Big/ \int_m^{\infty} \overline{\mu}(x)\,dx \quad (38)$$

Then,

$$\underline{c}_r = m + \frac{(c_{HUMF} - m)A_{UMF} - (c_{HLMF} - m)A_{LMF}}{A_{UMF} + A_{LMF}} \quad (39)$$

$$\overline{c}_r = m + \frac{(c_{HUMF} - m)A_{UMF} - (c_{HLMF} - m)A_{LMF}}{2A_{LMF}} \quad (40)$$

and, by symmetry, $\underline{c}_l = -\overline{c}_r$ *and* $\overline{c}_l = -\underline{c}_r$. ■

The proof of this theorem is totally different from the proof of Theorem 1, and will appear in the journal version of this paper. It makes very heavy use of the symmetry of the FOU.

Comment 5: When a symmetrical interval T2 FS $\tilde{A}$ is shifted by $\Delta m = m' - m$ so that $\tilde{A}$ is now symmetrical about m', then in (39) and (40), because A_{LMF} and A_{UMF} remain unchanged, and c_{HUMF}, c_{HLMF} and m are all shifted by Δm, both $\underline{c}_r$ and $\overline{c}_r$ are also shifted by Δm. This again means that $\overline{c}_r - \underline{c}_r$ and $\overline{c}_l - \underline{c}_l$ are shift-invariant (see Corollary 1); hence, in the rest of this section we can focus on a symmetrical interval T2 FS that is symmetrical about the origin.

Corollary 5: *For a symmetrical FOU, let*

$$\Delta c_{new} = \overline{c}_r - \underline{c}_r = \overline{c}_l - \underline{c}_l \quad (41)$$

where $\underline{c}_r$ *and* $\overline{c}_r$ *are in (39) and (40), in which* $m = 0$. *Then*

$$\Delta c_{new} = \overline{c}_r \frac{A_{FOU}}{A_{LMF} + A_{LMF}} \quad (42)$$

Proof: This follows directly from (39) and (40). ■

Comment 6: It is instructive to compare (42) and (33). For the rest of this paper, we shall refer to the results in (33) as Δc_{old}. Clearly $\Delta c_{new} < \Delta c_{old}$ if $\overline{c}_r < |x_1|$. It is possible for $\overline{c}_r > |x_1|$; so, using our two sets of bounds, we are able to conclude that

$$\Delta c = \frac{A_{FOU}}{A_{LMF} + A_{LMF}} \min\left\{|x_1|, \overline{c}_r\right\} = \min\left\{\Delta c_{old}, \Delta c_{new}\right\} \quad (43)$$

Example 2: Here we determine Δc for the symmetrical triangular FOU depicted in Fig. 5. From the simple geometry of this FOU, for which $0 \le h \le 1$, it follows that $|x_1| = b$, $A_{UMF} = b$, $A_{LMF} = ha$, $c_{HUMF} = b/3$, $c_{HLMF} = a/3$ and $A_{FOU} = A_{UMF} - A_{LMF} = b - ha$ so that

$$\Delta c_{old} = |x_1| \frac{A_{FOU}}{A_{LMF} + A_{UMF}} = b\frac{b - ha}{b + ha} \quad (44)$$

$$\Delta c_{new} = \overline{c}_r \frac{b - ha}{b + ha} = \frac{b^2 - ha^2}{6ha} \frac{b - ha}{b + ha} \quad (45)$$

For $\Delta c_{new} \le \Delta c_{old}$, we require

$$h \ge \frac{(b/a)^2}{1 + 6b/a} \quad (46)$$

From (44), (45) and (43), it is straightforward to study the behavior of Δc as a function of both h and b/a. ■

Example 3: Here we determine Δc for the symmetrical Gaussian FOU depicted in Fig. 6, for which

$$\overline{\mu}_{\tilde{A}}(x) = \exp\left(-x^2 / (2\sigma^2)\right) \quad (47)$$

$$\underline{\mu}_{\tilde{A}}(x) = s\exp\left(-x^2 / (2\sigma^2)\right) \quad (48)$$

where $s \in [0,1]$. Note that $\Delta c_{old} = \infty$, so that $\Delta c = \Delta c_{new}$. It is straightforward to show that $A_{UMF} = \sqrt{2\pi}\sigma$, $A_{LMF} = \sqrt{2\pi}s\sigma$, $c_{HUMF} A_{UMF} = 2\sigma^2$ and $c_{HLMF} A_{LMF} = 2s\sigma^2$; consequently, $A_{FOU} = \sqrt{2\pi}\sigma(1-s)$ and

$$\underline{c}_r = \frac{2\sigma}{\sqrt{2\pi}} \frac{(1-s)}{(1+s)} \quad (49)$$

$$\overline{c}_r = \frac{\sigma}{\sqrt{2\pi}} \frac{(1-s)}{s} \quad (50)$$

$$\Delta c = \frac{\sigma}{\sqrt{2\pi}} \frac{(1-s)^2}{s(1+s)} \quad ■ \quad (51)$$

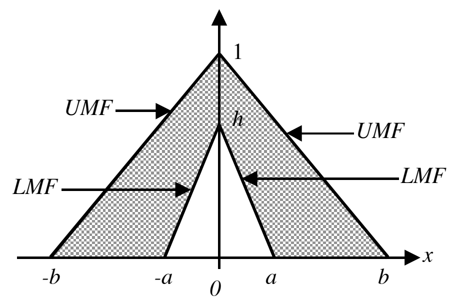

Fig. 5. Symmetrical triangular FOU.

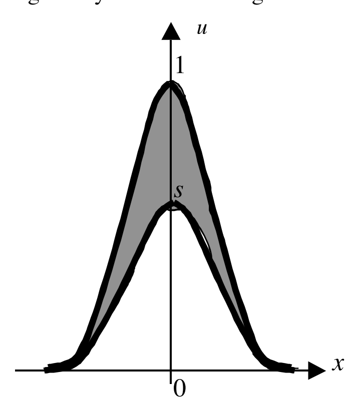

Fig. 6. Symmetrical Gaussian FOU.

Example 4: Using the FOU in Fig. 6, it is possible to solve an interesting *inverse problem*. Suppose that we have collected interval end-point data from a group of n people for a phrase (e.g., *some*), as described in footnote 2. For the purposes of this example, we assume that the uncertainties about the two end-points of this interval-data are the same.

The case when this is not true is currently under investigation. Let $x_1, x_2, ..., x_n$ denote the collected data for one end-point, and x_{avg} and Δx denote the sample average of the n points and the length of the $(1-\alpha)$ confidence interval (which is proportional to the sample standard deviation of the n points). We establish the following two *reasonable* design equations:

$$x_{avg} \equiv (\underline{c}_r + \bar{c}_r)/2 \quad (52)$$

$$\Delta x \equiv \bar{c}_r - \underline{c}_r \quad (53)$$

Next, we determine the parameters of a FOU that satisfy (52) and (53). To that end, we assume the FOU model of Example 3, from which it is possible to solve uniquely for FOU parameters s and σ as:

$$s = \frac{2x_{avg} - \Delta x}{2x_{avg} + 3\Delta x} \quad (54)$$

$$\sigma = \sqrt{2\pi} \frac{(2x_{avg} + \Delta x)(2x_{avg} - \Delta x)}{8\Delta x} \quad (55)$$

What this solution means is: starting with interval data that are collected from a group of people, we can compute the parameters of the scaled Gaussian FOU in Fig. 6, such that the centroid of this interval T2 FS is guaranteed to lie within $\Delta c = \bar{c}_r - \underline{c}_r$. ■

To the best knowledge of the authors, Example 4 represents the first solution of an inverse problem for a T2 FS. It represents a combining of statistics (x_{avg} and Δx) and uncertainty bounds for T2 FSs. In [6], Mendel coined the term *fuzzistics* for the field of experimental fuzzy sets, i.e. the field in which data are collected from people about MFs and related issues are formulated and tested (e.g., [1]). This paper and especially the results in this example illustrate some aspects of type-2 fuzzistics.

VI. CONCLUSIONS

We have demonstrated that the centroid of an interval T2 FS provides a measure of the uncertainty in such a FS. The centroid is a type-1 FS that is completely described by its two end-points. Although it is not possible to obtain closed-form formulas for these end-points, we have established closed-form formulas for upper and lower bounds of the two end-points. Most importantly, these bounds have been expressed in terms of geometric properties of the FOU, namely its area and the center of gravities of its upper and lower MFs. As a result, for the first time it is possible to quantify the uncertainty of an interval T2 FS with respect to these geometric properties of its FOU.

Using the results in this paper, it is possible to examine many "forward" problems, i.e. given a class of FOUs (e.g., triangular, trapezoidal, Gaussian) we can study the bounds on the centroid as a function of the parameter uncertainties that define the FOU.

It is also possible to examine "inverse" problems, i.e. given interval data collected from people about a phrase, and the inherent uncertainties associated with that data which can be described statistically, we can see if it is possible to establish a parametric FOU such that its uncertainty bounds are directly connected to statistical uncertainty bounds. Although we have provided a solution to this problem for one FOU, obtaining solutions for other FOUs is an open issue that is currently under study.

It is quite likely that we will need more quantitative information about a FOU than just its centroid uncertainty bounds if we are to go from uncertain data collected about interval end-points to a unique FOU, because the centroid uncertainty bounds are over-parameterized for some FOUs. This suggests that higher-order moments be established for an interval T2 FS, e.g., dispersion, skewness, and kurtosis. What will be needed for these new uncertainty measures are iterative methods for their computation (analogous to the Karnik-Mendel iterative methods for computing the interval end-points for the centroid of a T2 FS) and quantitative uncertainty bounds for them (analogous to the results presented in this paper for the centroid of a T2 FS). Once these additional results have been developed, then we will be able to establish whether or not it is indeed possible to go from interval end-point data to a unique (non-symmetrical) FOU and if so how to do this.

Connecting data and its uncertainties to a parametric FOU for an interval T2 FS is analogous to estimating parameters in a probability model, and, as is well known, the latter provides a bridge between probability and statistics. We hope that the material in this paper will be the start of much research in providing a bridge between interval T2 FSs and type-2 fuzzistics, something that we believe is needed if computing with words is to become a reality (e.g., [4], [6]).

REFERENCES

[1] T. Bilgic and I. B. Turksen, "Measurement of membership functions: theoretical and empirical work," in *Handbook of Fuzzy Systems, Vol. 1: Foundations*, (D. Dubois and H. Prade, Eds.), pp. 195-228, Kluwer, Boston, MA, 2000.

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

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

[4] J. M. Mendel, "Computing with words, when words can mean different things to different people," in *Proc. of Third International ICSC Symposium on Fuzzy Logic and Applications*, Rochester, NY, June 1999.

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

[6] J. M. Mendel, "Fuzzy sets for words: a new beginning," *Proc. of IEEE Int'l. Conf. on Fuzzy Systems*, St. Louis, MO, pp. 37-42, May 2003.

[7] J. M. Mendel, "On a 50% savings in the computation of the centroid of a *symmetrical* interval type-2 fuzzy set," accepted for publication in *Information Sciences*, 2004.

[8] J. M. Mendel and R. I. Bob John, "Type-2 fuzzy sets made simple," *IEEE Trans. on Fuzzy Systems*, vol. 10, pp. 117-127, April 2002.

[9] H. Wu and J. M. Mendel, "Uncertainty bounds and their use in the design of interval type-2 fuzzy logic systems," *IEEE Trans. on Fuzzy Systems*, vol. 10, pp. 622-639, Oct. 2002.