

# The Arithmetic Recursive Average as an Instance of the Recursive Weighted Power Mean

Christian Wagner

Inst. for Computing & Cybersystems  
Michigan Technological University, USA  
and LUCID, School of Computer Science  
University of Nottingham, UK  
Email: christian.wagner@nottingham.ac.uk

Timothy C. Havens

Dept. of Electrical and Computer Eng. and Computer Science  
Michigan Technological University  
Houghton, MI, USA  
Email: thavens@mtu.edu

Derek T. Anderson

Dept. of Electrical and Computer Eng.  
Mississippi State University  
Mississippi State, MS, USA  
Email: anderson@ece.msstate.edu

**Abstract**—The aggregation of multiple information sources has a long history and ranges from sensor fusion to the aggregation of individual algorithm outputs and human knowledge. A popular approach to achieve such aggregation is the *fuzzy integral* (FI) which is defined with respect to a *fuzzy measure* (FM) (i.e. a normal, monotone capacity). In practice, the discrete FI aggregates information contributed by a discrete number of sources through a weighted aggregation (post-sorting), where the weights are captured by a FM that models the typically subjective ‘worth’ of subsets of the overall set of sources. While the combination of FI and FM has been very successful, challenges remain both in regards to the behavior of the resulting aggregation operators—which for example do not produce symmetrically mirrored outputs for symmetrically mirrored inputs—and also in a manifest difference between the intuitive interpretation of a stand-alone FM and its actual role and impact when used as part of information fusion with a FI. This paper elucidates these challenges and introduces a novel family of *recursive average* (RAV) operators as an alternative to the FI in aggregation with respect to a FM; focusing specifically on the arithmetic recursive average. The RAV is designed to address the above challenges, while also facilitating fine-grained analysis of the resulting aggregation of different combinations of sources. We provide the mathematical foundations of the RAV and include initial experiments and comparisons to the FI for both numeric and interval-valued data.

## I. INTRODUCTION

The aggregation of multiple information sources has a long history and ranges from sensor fusion to the aggregation of individual algorithm outputs and human knowledge. A popular approach to achieve such aggregation is the *fuzzy integral* (FI) which is defined with respect to a *fuzzy measure* (FM) (i.e. a normal, monotone capacity). In practice, the discrete FI aggregates typically objective information contributed by a discrete number of sources through a weighted aggregation (post-sorting), where the weights are captured by a FM which models the typically subjective worth of subsets of the overall set of sources.

In many applications, the FI is used simply as a parametric aggregation function. In this respect, the FM does not semantically mean anything but is just used as a set of parameters to be optimized. However, in of itself, the FM is viewed as a hierarchical weight for a set of sources which does not only capture the individual weight—or more commonly, the

worth—of each source, but also the worth of all combinations of sources within the set (e.g., see Fig. 2).

This *meaning* of the FM is particularly relevant in cases where the FM is specified externally—for example by human experts, and, in cases where an attempt is made at a later stage to interpret the actual values within the FM—for example to develop a better understanding of the aggregation.

A strong source of motivation for this paper is that while the semantic interpretation and explanation of the FM is highly intuitive, current FI based data aggregation applications do not follow this interpretation and do not use the FM to weight all combinations of sources as would intuitively be expected. For example, FI based aggregation of symmetrically mirrored inputs does not result in symmetrically mirrored outputs - as would intuitively be expected. This and similar aspects of FI based data aggregation are discussed further in Section IV.

Importantly, this disconnect between semantic meaning and actual fusion behavior of the FI creates three challenges:

- 1) **External Specification** - it prevents the *a priori* specification of the FM based on external knowledge as an approach to systematically fuse multi-source data. In other words, approaches which generate the FM based on its role of capturing the worth of sources and their combinations do not produce expected results when said FMs are then used in combination with the FI.
- 2) **Validation** - it prevents human interpretation and validation of a FM, thus affecting the capacity for trusting the fused outputs post-aggregation. Essentially, a FM which has been optimized in conjunction with a FI—e.g., through an evolutionary algorithm [1]—may result in optimal FI based fusion for the data considered, but the actual nodes within the FM may not reflect the actual worth of the specific combinations of sources.
- 3) **Knowledge generation** - this is directly related to the Validation challenge. By not maintaining the semantics of the FM, optimized FMs lose the capacity to deliver knowledge on the worths of combinations of sources. In other words, the optimization process does not generate externally useful knowledge on the worth of combinations of sources; we cannot ask the question “What is the worth of the combination of sources  $A$ ,  $B$ , and  $C$ ?”.

Beyond the challenges above, traditional FI based data aggregation is designed to generate one single output. While highly efficient, this approach does not facilitate more fine-grained aggregation where, for example, the aggregated result of each possible source-combination is incrementally known and a decision can be made on whether the fusion should proceed. For example, it may not be meaningful to aggregate two sources providing opposite evidence or overly redundant sources.

In order to address the above challenges, this paper presents a new family of *Recursive Average* (RAV) operators as a novel alternative to the FI, thus providing a new pathway for employing the FM in a data-aggregation context. First, Section II provides key background material on the FM and FI, then Section III introduces the RAV. Section IV highlights the behaviour of the resulting FM-RAV aggregation approach in contrast to traditional FM-FI based aggregation, focusing specifically on the arithmetic RAV. Section V discusses the computational complexity of the RAV, while Section VI provides a brief discussion, conclusions, and further work.

## II. BACKGROUND

### A. Fuzzy Measure

1) *Overview*: A FM captures the possibly subjective worth of every subset in the power set of information sources. In the context of multi-source data aggregation, a FM can be interpreted as a complex weighting structure (a lattice) which assigns weights to sources in a similar fashion to the weights in a weighted average; however, FMs enable the capture of not only the weights of individual sources—referred to as *densities*—but also the weights of all possible combinations of sub-sources. This lattice of weights is shown for an example FM  $g$  with three sources in Fig. 2.

Formally, let  $X = \{x_1, \dots, x_n\}$  be a non-empty finite set of information sources and  $g : 2^X \rightarrow [0, 1]$  be a FM with the following properties [2]:

P1)  $g(\emptyset) = 0$ ,  $g(X) = 1$ ;

P2) If  $A \subseteq B \subseteq X$  then  $g(A) \leq g(B) \leq 1$  ( $g$  is monotonic and non-decreasing).

Note that there is a third property for continuous FMs which is not applicable to discrete FMs such as those in this paper. The measure  $g$  is the confidence or worth of each subset of  $X$ ; hence, P1 tells us that the worth of no sources, or the empty set  $\emptyset$ , is 0 and the worth of *all* sources, or the universal set  $X$ , is 1. P2 tells us that if  $A$  is a subset of  $B$ , then  $B$  is worth at least as much as  $A$ . In other words, the monotonicity constraint of P2 implies the assumption that more sources cannot result in a decrease in worth.

While the majority of applications of the FM is in data aggregation, where, in conjunction with a FI, the FM is used effectively as a complex, constrained parameter set (which can be optimized), a clear attraction of the FM per se is its lattice structure which enables it to capture the intuitive worths of different combinations of input sources. The latter in turn makes the FM suitable for expert-led specification (see below), but also for introspection and validation.

2) *Constructing Fuzzy Measures*: A major challenge facing the application of FMs is the population of the actual FM lattice (see example in Fig. 2 for three sources), i.e., how to determine the values of the variables in the FM. Several approaches exist, including the following.

- *Specification by experts*. While for a small number of sources, the specification of the FM by experts is viable, manual specification for larger numbers of sources rapidly becomes unfeasible. For example, there are already 31 non-empty subsets of a set of 5 sources.
- *Algorithmic specification*. Examples include the Sugeno  $\lambda$  FM [3], the decomposable FM [4], and the more recently introduced data-driven FMs [5, 6]. It is worth noting that some of these algorithms, including the Sugeno  $\lambda$  FM and the decomposable FM, derive the entire FM lattice from the densities alone, with the sole additional constraint being the mathematical correctness of the FM in respect to P1 and P2 above, while others derive the entire FM lattice directly based on some criteria such as the level of agreement over given subsets of sources.
- *Specification through optimization*. Several techniques, including evolutionary algorithms [1] and quadratic programming [7] have been employed to generate FMs while relying on training data and a criterion such as the quality of the fused output. Here, it is important to note that these approaches do not attempt to preserve semantic meaning of the FM, but rather treat it as a set of free weights which are optimized to deliver minimum output error with the specific FI employed.
- *Combinations of the above*. While not a main focus for current application or research, the above approaches can of course be combined. For example, the worth of a number of subsets is known by experts, while the rest is learned from data (e.g., [8])

### B. Fuzzy Integral

There are many types of the FI; see [2] for a detailed discussion. FIs are mostly used for evidence fusion [2, 9–11]. They combine sources of information by accounting for both the support of the question (the evidence  $h$ ) and the expected worth of each subset of sources (as supplied by a FM  $g$ ). Here, we focus on the discrete Sugeno (SFI) and Choquet (CFI) FIs, proposed by Murofushi and Sugeno [12, 13]. Let  $h : X \rightarrow [0, \infty)$  be a real-valued function that represents the evidence or support of a particular hypothesis.<sup>1</sup> The discrete SFI and CFI are defined respectively as:

$$\int_{SFI} h \circ g = SFI_g(h) = \bigvee_{i=1}^n (h(\{x_{\pi(i)}\}) \wedge g(A_i)), \quad (1a)$$

$$\int_{CFI} h \circ g = CFI_g(h) = \sum_{i=1}^n h(\{x_{\pi(i)}\}) [g(A_i) - g(A_{i-1})], \quad (1b)$$

<sup>1</sup>Generally, in data fusion, it is convenient to have  $h : X \rightarrow [0, 1]$ , where each source is normalized to the unit-interval. This is also the appropriate space of  $h$  for use with the Sugeno FI because, by definition,  $g \in [0, 1]$ .

where  $\pi$  is a permutation of  $X$ , such that  $h(\{x_{\pi(1)}\}) \geq h(\{x_{\pi(2)}\}) \geq \dots \geq h(\{x_{\pi(n)}\})$ ,  $A_i = \{x_{\pi(1)}, \dots, x_{\pi(i)}\}$ , and  $g(A_0) = 0$  [3, 14]. The *max* and *min* definition of the Sugeno FI has been extended to other t-conorms and t-norms. Detailed treatments of the properties of FIs can be found in [3, 14, 15].

In some cases, the evidence  $h$  cannot, or should not, be represented simply by numbers;  $h$  would be better represented as an interval-valued or fuzzy number-valued function. An example is the survey question, “How many bottles of wine should I purchase for the reception?” Many people would answer with an interval, e.g., “between 20 and 30,” or a fuzzy number, e.g., “about 25.” Such extensions of both the fuzzy Sugeno and Choquet FIs have been proposed for both interval-valued and also fuzzy number-valued integrands [3, 16–19].

### III. THE RECURSIVE AVERAGE

While the FI provides a powerful and highly successful means of fusing information, several challenges remain—see Sections I and IV. In this section we introduce the family of RAVs—i.e., the recursive weighted power mean—over a set of sources  $S$ , defined as follows.

**Definition 1.** Recursive Weighted Power Mean

$$\text{RAV}_p(S) = \begin{cases} h(S), & \text{if } |S| = 1, \\ \left( \frac{\sum g(B_j) \text{RAV}_p(B_j)^p}{\sum g(B_j)} \right)^{1/p}, & \text{otherwise,} \end{cases} \quad (2)$$

where  $|p| > 0$ ,  $B_j = S \setminus x_j$ ,  $x_j \in S$ , and  $g$  is a FM where at most one of the densities is zero.

Note that for specific values of  $p$ , the RAV follows the weighted power mean and adopts well known average behaviors, such as the harmonic average ( $p = -1$ ), the arithmetic average ( $p = 1$ ), the quadratic average ( $p = 2$ ), etc. For  $p = -\infty$ , the RAV approaches the minimum, and for  $p = +\infty$ , the maximum operator; while for the case where  $p = 0$ , the RAV adopts the geometric mean:

$$\text{RAV}_0(S) = \begin{cases} h(S), & \text{if } |S| = 1, \\ \left( \prod \text{RAV}_0(B_j)^{g(B_j)} \right)^{1/\sum g(B_j)}, & \text{otherwise.} \end{cases} \quad (3)$$

Thus,  $\forall p$ , the RAV for a set of sources is recursively defined as the weighted average of the RAVs of its sub-sources, where the weight at each node is captured by a FM. Note that for conciseness, this paper focuses solely on the instance of the RAV where  $p = 1$ , i.e., the arithmetic RAV. We will further explore the other forms of the RAV in a future publication. For simplicity, we will refer to the arithmetic RAV as RAV throughout this paper.

Figure 1 illustrates the RAV for three sources, highlighting its recursive nature (from top to bottom). Note that the RAV of singletons is the evidence of the given singleton source.

**Remark 1.** The RAV enables the fusion of evidence from multiple sources in conjunction with a FM, similar to the FI. That is, the RAV also accounts for both the support of

the question—the evidence  $h$ —and the expected worth of each subset of sources as supplied by a FM  $g$ . However, the output of the RAV, unlike for the FI, is not a single output; it is a lattice which is identical in structure to that of the FM  $g$ , reflecting the output of the RAV for all possible combinations of sources, including the overall output over all sources:  $\text{RAV}(X)$ .

**Proposition 1.** The RAV is not monotonic non-decreasing.

While the output of the RAV adopts the lattice structure of the original FM, it itself is not a FM as it is not monotonically non-decreasing, i.e., even if it was normalized, it would not satisfy property P2 of a FM as described in Section II-A.

*Proof.* Let  $X = \{x_1, x_2\}$ , with evidence  $h(\{x_1\}) = 0.6$  and  $h(\{x_2\}) = 0.1$ , and worth of  $g(\{x_1\}) = 0.1$  and  $g(\{x_2\}) = 0.9$ . We have  $\text{RAV}(x_1) = 0.6$ ,  $\text{RAV}(x_2) = 0.1$ , and  $\text{RAV}(X) = (0.1 * 0.6 + 0.9 * 0.1)/(0.1 + 0.9) = 0.15$ . Hence,  $\text{RAV}(x_1) > \text{RAV}(X)$ .  $\square$

**Proposition 2.** The output of RAV is bounded by  $\min(h(X)) \leq \text{RAV}(X) \leq \max(h(X))$ .

*Proof.* The proof is trivial as the RAV is a recursive series of arithmetic averages, each bounded by its smallest and largest input. It is also easily extended to the family  $\text{RAV}_p$  which approaches a *min* and *max* operator as  $p$  approaches  $-\infty$  and  $+\infty$  respectively.  $\square$

**Example 1.** Consider a set of three sources  $X = \{x_1, x_2, x_3\}$ , and a FM  $g$  capturing the worth of each source and their combinations. The RAV over all sources is computed as follows.

$$\begin{aligned} \text{RAV}(X) = & (g(\{x_1, x_2\}) \text{RAV}(\{x_1, x_2\}) \\ & + g(\{x_1, x_3\}) \text{RAV}(\{x_1, x_3\}) \\ & + g(\{x_2, x_3\}) \text{RAV}(\{x_2, x_3\})) \\ & / (g(\{x_1, x_2\}) + g(\{x_1, x_3\}) + g(\{x_2, x_3\})), \end{aligned} \quad (4)$$

where

$$\begin{aligned} \text{RAV}(\{x_1, x_2\}) &= \frac{g(\{x_1\})\text{RAV}(\{x_1\}) + g(\{x_2\})\text{RAV}(\{x_2\})}{g^{AR}(\{x_1\}) + \text{RAV}(\{x_2\})}, \\ \text{RAV}(\{x_1, x_3\}) &= \frac{g(\{x_1\})\text{RAV}(\{x_1\}) + g(\{x_3\})\text{RAV}(\{x_3\})}{g^{AR}(\{x_1\}) + \text{RAV}(\{x_3\})}, \\ \text{RAV}(\{x_2, x_3\}) &= \frac{g(\{x_2\})\text{RAV}(\{x_2\}) + g(\{x_3\})\text{RAV}(\{x_3\})}{g^{AR}(\{x_2\}) + \text{RAV}(\{x_3\})}, \end{aligned} \quad (5)$$

and  $\text{RAV}(\{x_1\}) = h(\{x_1\})$ ,  $\text{RAV}(\{x_2\}) = h(\{x_2\})$ ,  $\text{RAV}(\{x_3\}) = h(\{x_3\})$ .

While the RAV does not require a reordering of sources as the FI does, its recursive nature makes it computationally expensive in comparison to the SFI and CFI. We review its computational complexity in detail in Section V. We now proceed by introducing a number of synthetic examples to illustrate the behavior of the RAV, in particular in contrast to the CFI and SFI when used in a data fusion context.

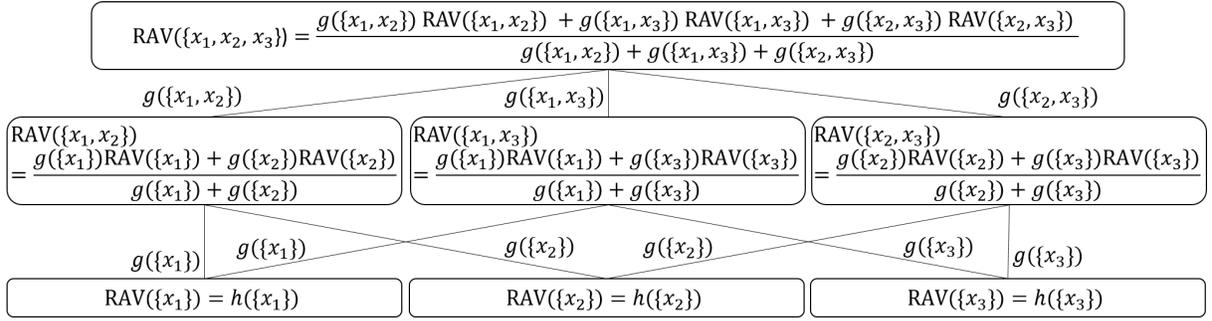


Fig. 1: Arithmetic recursive average for three sources. Note how a FM  $g$  (such as the example in Fig. 2) is used as weights.

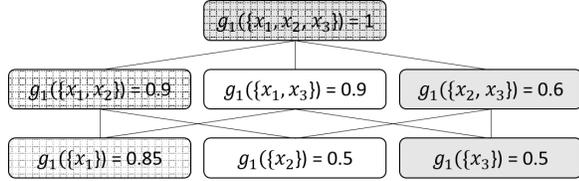


Fig. 2: The FM  $g_1$ .

#### IV. EXPERIMENTS AND ILLUSTRATIVE EXAMPLES

This section provides a series of data fusion examples which serve both to illustrate challenges in FI-based fusion (see Section I) and demonstrate the behavior of the RAV in respect to the SFI and CFI.

##### A. Examples for Synthetic Numeric Data

In order to illustrate the properties of the RAV, while also underpinning the motivation for its design, we provide a series of numeric examples, comparing its application and output to that of the SFI and CFI, where applicable.

1) *Mirrored inputs*: Consider three sources  $X = \{x_1, x_2, x_3\}$  and an associated FM  $g_1$  as depicted in Fig. 2. Furthermore, consider the following two sets of evidence provided by each of the sources:  $H_a = \{h_a(\{x_1\}) = 0.1, h_a(\{x_2\}) = 0.3, h_a(\{x_3\}) = 0.35\}$  and  $H_b = \{h_b(\{x_1\}) = 0.9, h_b(\{x_2\}) = 0.7, h_b(\{x_3\}) = 0.65\}$ . Note that the evidence in  $H_b$  is a mirror image of that in  $H_a$ , viz.,  $h_b(\{x_1\}) = 1 - h_a(\{x_1\})$ . Thus, it is intuitive to expect that the fused outcome across all three sources is also mirrored in relation to the different sets of evidence, i.e. to expect that  $SFI_{g_1}(H_b) = 1 - SFI_{g_1}(H_a)$  and the same for the CFI and RAV. The numeric results for fusing the evidence sets  $H_a$  and  $H_b$  is provided for in Table I for the SFI, the CFI, and the RAV.

Considering Table I, it is clear that the SFI and CFI do not produce outputs that are mirror images, i.e.  $SFI_{g_1}(H_b) \neq 1 - SFI_{g_1}(H_a)$  and  $CFI_{g_1}(H_b) \neq 1 - CFI_{g_1}(H_a)$ . The reason for this non-intuitive result is the way that both FIs focus on the sources providing the largest evidence, as detailed in Section II-B. Specifically, the ordering for evidence set  $H_a$  will result in the permutation:  $\pi(1) = 3, \pi(2) = 2, \pi(3) = 1$ , i.e. the shaded parts of the FM  $g_1$  in Fig. 2 are used. However, for the evidence set  $H_b$ , the permutation will be  $\pi(1) = 1,$

TABLE I: Numeric results (in respect to FM  $g_1$ , see Fig. 2)

	SFI	CFI	RAV
$H_a$	0.35	0.245	0.219
$H_b$	0.85	0.865	0.781
$H_c$	0.9	0.9	0.401

$\pi(2) = 2, \pi(3) = 3$ , resulting in the patterned parts of the FM  $g_1$  in Fig. 2 being used. The RAV however, uses the complete lattice of the FM in both cases, thus returning the intuitive, mirrored results, i.e.  $RAV_{g_1}(H_b) = 1 - RAV_{g_1}(H_a)$ .

2) *Exploitation of the complete FM*: A noteworthy aspect of the example in Section IV-A1 is that for both the SFI and CFI only a part of the FM is used during the fusion step. While this makes them highly efficient at combining the complex worth encoded in the FM with the set of evidence, it also leads to two potential pitfalls.

First, in an optimization context where the FM is optimized/learned from data [1, 20], this may lead to certain parts of the FM not being optimized if training data does not result in all permutations of the FM being used—i.e., ‘visited’ during training.

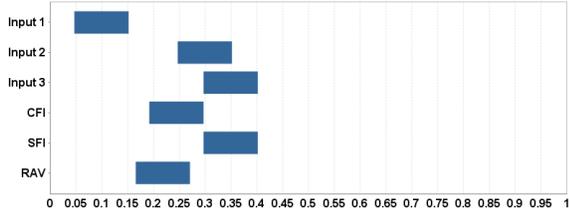
Second, the resulting fusion may not follow intuition. That is, while the fusion based on the FI and a given FM is expected to provide a fused result that considers the *worth* of all sources and their combinations; in effect, this is not the case. As an example, changing the *worth* of source  $x_2$  to 0, i.e.,  $g_1(\{x_2\}) = 0$  in the example in Section IV-A1 does not affect the final output of the SFI and CFI. For comparison, the results of the RAV are changed to  $RAV_{g_1}(H_a) = 0.191$  and  $RAV_{g_1}(H_b) = 0.809$ .

3) *Intuitive source selection*: Consider three sources  $X = \{x_1, x_2, x_3\}$  and an associated FM  $g_2$ . The worths of all combinations of sources within the FM are 1, except the worths of the singletons which are  $g_2(\{x_1\}) = 0.99, g_2(\{x_2\}) = 0.99,$  and  $g_2(\{x_3\}) = 1.0$ . In other words, the FM expresses that all sources are of very high worth, with sources  $g_2(\{x_1\})$  and  $g_2(\{x_2\})$  fractionally lower than  $g_2(\{x_3\})$ . Consider the following evidence set  $H_c$  provided by the sources:  $H_c = \{h_c(\{x_1\}) = 0.1, h_c(\{x_2\}) = 0.2, h_c(\{x_3\}) = 0.9\}$ . The results for fusing the evidence  $H_c$  in respect to the FM  $g_2$  using the SFI, CFI, and RAV are given in Table I.

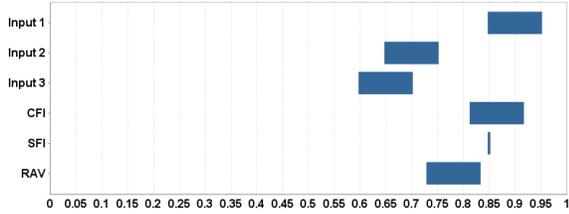
The results in Table I for both the SFI and CFI show how both of these fusion operators are swayed severely by a very

TABLE II: Interval results (in respect to FM  $g_1$ , see Fig. 2)

	SFI	CFI	RAV
$\overline{H}_a$	[0.3,0.4]	[0.195,0.295]	[0.169,0.269]
$\overline{H}_b$	[0.85,0.85]	[0.815,0.915]	[0.731,0.831]



(a) Standard interval inputs (evidence set  $\overline{H}_a$ ), and fused outputs



(b) Mirrored interval inputs (evidence set  $\overline{H}_b$ ), and fused outputs

Fig. 3: Synthetic data example for mirroring interval inputs.

small difference in worth, exposing an instability in both FIs in relation to small variations within the FM and resulting in a potentially non-intuitive result. The RAV result follows intuition, i.e. since all sources are essentially equally worthy, the outcome should be very near to the average of the three sources.

### B. Examples for Synthetic Interval-Valued Data

As the RAV is an arithmetic operator, its extension from numeric to interval-valued data is straightforward by employing interval arithmetic. Similarly, it can be extended for application to fuzzy-number valued data by using an  $\alpha$ -cut decomposition of fuzzy numbers. Because of space constraints, here, we only provide examples for interval-valued evidence and numeric worth, leaving further detail to future publications.

As in Section IV-A1, consider three sources  $X = \{x_1, x_2, x_3\}$  and the FM  $g_1$  as depicted in Fig. 2. For continuity, we generate interval-valued evidence by perturbing the evidence from Section IV-A1 with  $\pm 0.5$ . Thus, consider the following two sets of interval-valued evidence provided by each source:  $\overline{H}_a = \{h_a(\{x_1\}) = [0.05, 0.15], h_a(\{x_2\}) = [0.25, 0.35], h_a(\{x_3\}) = [0.3, 0.4]\}$  and  $\overline{H}_b = \{h_b(\{x_1\}) = [0.85, 0.95], h_b(\{x_2\}) = [0.65, 0.75], h_b(\{x_3\}) = [0.6, 0.7]\}$ . Note that the evidence in  $\overline{H}_b$  is still a mirror image of that in  $\overline{H}_a$ . The results of the fusion for the SFI, CFI, and the RAV are provided numerically in Table II and visualized in Fig. 3.

As for the numeric case in Section IV-A1, the interval-valued outputs in Table II and more visually, Fig. 3 highlight that for the FIs, the outputs for symmetrically mirrored inputs are not symmetrically mirrored themselves as would be intuitive. The outputs for the RAV are however symmetric, as shown for the numeric case in Section IV-A1.

### C. Real-world data examples

As laid out in Section I, the motivation for the introduction of the RAV is largely to provide useful and intuitive fusion while leveraging the potential of the FM. Thus, we would ideally like to be able to take a FM which encapsulates external knowledge on a problem—e.g., the worth of combinations of different experts—to fuse evidence arising from those experts. In the current literature, it is challenging to find examples of externally specified FMs. Most effort has been channeled towards learning and optimizing FMs *in respect to* a FI, which effectively means that the FM is tuned to result in 'optimal' fusion for the given FI, rather than providing an independently 'optimal' source of information.

A key interdisciplinary area of recent work where externally specified FMs have been used is the estimation of age-at-death from skeletal remains via applying multiple aging techniques to different parts of the human skeleton [9]. The confidences in the different techniques are subsequently captured as the densities of a FM, while the rest of the FM is completed using the Sugeno  $\lambda$ -FM. The Sugeno  $\lambda$ -FM does not in fact add external information to the FM, but rather completes the FM's lattice in a mathematically correct form, allowing application of FIs for fusion. However, as the resulting FM has been externally specified (i.e. not optimized in respect to a FI), it provides a good basis for demonstration of the RAV on real-world data, and its comparison to the FI.

We replicate the results for two sets of skeletal remains as shown in [9], providing the aggregation outputs for the SFI, CFI and the RAV, as well as the chronological (ground truth) age in Fig. 4. For details on the raw data, please consult [9].

From the plots, it is clear that—as expected for identical inputs—all data fusion operators provide comparable results. It is however interesting to note how the RAV produces higher confidence around the correct age in Fig. 4a and lower confidence—where the age estimate is incorrect—in Fig. 4b. This behavior of the RAV would be highly valuable in the generation of meaningful decision support outputs (e.g.: “What was the age of this person at death and how sure are we about it?”) and we will explore it further in future work.

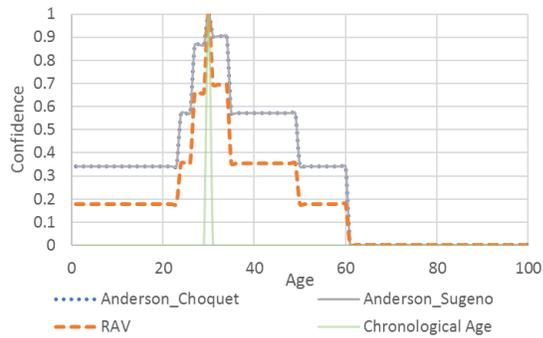
Finally, before concluding this paper, we briefly review the computational complexity of the RAV in the next section.

### V. COMPUTATIONAL COMPLEXITY

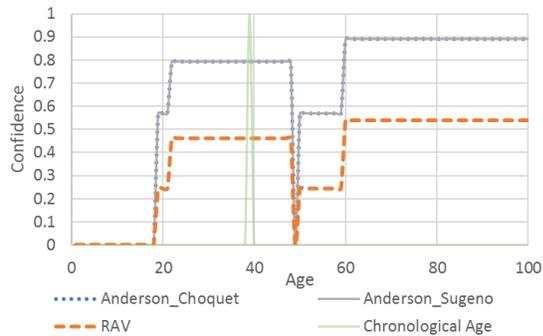
The computational complexity  $C_{\text{RAV}}$  of the arithmetic RAV operator can be captured at the level of the mathematical operations required for its computation for a given number of sources,

$$C_{\text{RAV}} = \binom{n}{k} (((k-1) * 2 * x) + (k * y) + z), \quad (6)$$

where  $n$  is the number of sources,  $k \in \{2, \dots, n\}$ ,  $x$  is the number of additions,  $y$  is the number of product operations, and  $z$  is the number of divisions. Using standard complexity notation,  $C_{\text{RAV}}$  can be captured as  $O(n^2 \log n)$ , compared to the  $O(n \log n)$  complexity of the FI. The complexity of the former is driven by its recursive nature, resulting in the



(a) Skeleton 208 (Fig. 3 in [9])



(b) Skeleton 779R (Fig. 6 in [9])

Fig. 4: Comparative fusion for anthropological data from [9].

factorial term in (6), while that of the latter is driven by the sorting operation.

Theoretically, the RAV is more computationally expensive than the FI, but comparing the complexity of the RAV operator to that of the CFI or SFI in practice is non-trivial. Clearly, the computation of the RAV is in principle more computationally expensive, requiring substantially more, and more computationally complex, (i.e. *division* rather than just *sum* and *product*) arithmetic operations. However, we note that the RAV 1) provides substantially more detailed output than the FI, resulting in a lattice of sub-aggregation results at each node, rather than just an overall aggregation results as for the FI, and 2) enables the partial re-computation of the lattice in cases where only evidence from a subset of sources has changed.

## VI. DISCUSSION AND CONCLUSIONS

In this paper we proposed a new family of recursive average (RAV) operators (i.e. the recursive weighted power mean), expanding in particular on the arithmetic RAV. The motivation for the new operator family is the desire to provide an intuitive path to aggregating data (evidence) in conjunction with the powerful structure of a fuzzy measure. We showed how the proposed RAV avoids potential challenges, such as the non-symmetry of outputs for symmetrically flipped inputs that can occur when employing a FI in conjunction with a FM, and thus how the arithmetic RAV can provide a useful alternative to FIs. We provided both synthetic and real world data fusion

examples for numeric and interval data, comparing the fusion performance of the Sugeno and Choquet FIs with that of the arithmetic RAV. Finally, we briefly reviewed the RAV's computational complexity.

In future work, we are looking to further explore the wider family of RAVs while also focusing on the knowledge that can be extracted from a FM when it has been learned in respect to a RAV.

## REFERENCES

- [1] D. T. Anderson, J. M. Keller, and T. C. Havens, "Learning fuzzy-valued fuzzy measures for the fuzzy-valued sugeno fuzzy integral," in *International conference on information processing and management of uncertainty*, 2010, pp. 502–511.
- [2] M. Grabisch, T. Murofushi, and M. Sugeno, *Fuzzy Measures and Integrals: Theory and Applications*. Physica-Verlag, Heidelberg, 2000.
- [3] M. Grabisch, H. Nguyen, and E. Walker, *Fundamentals of uncertainty calculi, with applications to fuzzy inference*. Kluwer Academic, Dordrecht, 1995.
- [4] D. Dubois and H. Prade, *Fuzzy Sets and Systems: Theory and Applications*, ser. Mathematics in science and engineering. Academic Press, 1980.
- [5] C. Wagner and D. T. Anderson, "Extracting meta-measures from data for fuzzy aggregation of crowd sourced information," in *IEEE Int. Conf. Fuzzy Systems*, June 2012, pp. 1–8.
- [6] T. Havens, D. Anderson, and C. Wagner, "Data-informed fuzzy measures for fuzzy integration of intervals and fuzzy numbers," *Fuzzy Systems, IEEE Transactions on*, vol. 23, no. 5, pp. 1851–1875, 2015.
- [7] A. J. Pinar, J. Rice, L. Hu, D. T. Anderson, and T. C. Havens, "Efficient multiple kernel classification using feature and decision level fusion," *IEEE Transactions on Fuzzy Systems*, vol. PP, no. 99, pp. 1–1, 2016.
- [8] M. A. Islam, D. T. Anderson, and T. C. Havens, "Multi-criteria based learning of the choquet integral using goal programming," in *2015 Annual Conference of the North American Fuzzy Information Processing Society (NAFIPS) held jointly with 2015 5th World Conference on Soft Computing (WConSC)*, Aug 2015, pp. 1–6.
- [9] M. Anderson, D. T. Anderson, and D. Wescott, "Estimation of adult skeletal age-at-death using the sugeno fuzzy integral," *American Journal of Physical Anthropology*, vol. 142, pp. 30–41, 2010.
- [10] S. Auephanwiriyaikul, J. Keller, and P. Gader, "Generalized choquet fuzzy integral fusion," *Information Fusion*, vol. 3, pp. 69–85, 2002.
- [11] H. Tahani and J. Keller, "Information fusion in computer vision using the fuzzy integral," *Systems, Man and Cybernetics, IEEE Transactions on*, vol. 20, no. 3, pp. 733–741, May 1990.
- [12] G. Choquet, "Theory of capacities," *Annales de l'Institut Fourier*, vol. 5, pp. 131–295, 1953.
- [13] T. Murofushi and M. Sugeno, "An interpretation of fuzzy measures and the choquet integral as an integral with respect to a fuzzy measure," *Fuzzy Sets and Systems*, vol. 29, no. 2, pp. 201–227, 1989.
- [14] M. Sugeno, "Theory of fuzzy integrals and its applications," *Ph.D. thesis*, vol. Tokyo Institute of Technology, 1974.
- [15] M. Grabisch, "Fuzzy integral in multicriteria decision making," *Fuzzy Sets Syst.*, vol. 69, no. 3, pp. 279–298, 1995.
- [16] D. Zhang and Z. Wang, "Fuzzy integrals of fuzzy-valued functions," *Fuzzy Sets and Systems*, vol. 54, pp. 63–67, 1993.
- [17] R. Yang, Z. Wang, P.-A. Heng, and K.-S. Leung, "Fuzzified choquet integral with a fuzzy-valued integrand and its application on temperature prediction," *Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on*, vol. 38, no. 2, pp. 367–380, April 2008.
- [18] T. C. Havens, D. T. Anderson, and J. M. Keller, "A fuzzy choquet integral with an interval type-2 fuzzy number-valued integrand," in *IEEE Int. Conf. Fuzzy Systems*, July 2010, pp. 1–8.
- [19] D. T. Anderson, T. C. Havens, C. Wagner, J. M. Keller, M. F. Anderson, and D. J. Wescott, "Sugeno fuzzy integral generalizations for sub-normal fuzzy set-valued inputs," in *IEEE International Conference on Fuzzy Systems*, June 2012, pp. 1–8.
- [20] D. T. Anderson, S. R. Price, and T. C. Havens, "Regularization-based learning of the choquet integral," in *2014 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)*, July 2014, pp. 2519–2526.