Author's Accepted Manuscript Intl. J. Intelligent Systems DOI http://dx.doi.org/10.1002/int.22083

Maximum Entropy Ordered Weighted Averaging in the binomial decomposition framework

Silvia Bortot[1,*], Ricardo Alberto Marques Pereira[1], Anastasia Stamatopoulou[2]
[1] *Department of Economics and Management, University of Trento, Italy*
[2] *Department of Industrial Engineering, University of Trento, Italy*

We consider the maximum entropy constrained optimization problem associated with Ordered Weighted Averaging (OWA) in the binomial decomposition framework.

We begin by reviewing the analytic solution of the maximum entropy method proposed by Filev and Yager in 1995, and later by Fullér and Majlender in 2001. Next we briefly review the binomial decomposition framework, which allows for an alternative parametric description of the OWA functions. The values of the binomial coefficients α_j, $j = 1, \ldots, n$ are uniquely determined by the weighting structure of the OWA function.

We observe that for low orness values $\Omega \in [0, 0.5]$ the optimal weights are decreasing, whereas they are increasing for high orness values $\Omega \in [0.5, 1]$. Moreover, we notice that the optimal values of the first and last weights have a wide range in $[0, 1]$, while the values of the other weights have more restricted ranges.

As for the optimal α_j, $j = 1, \ldots, n$ coefficients we find that their behavior with respect to orness values $\Omega \in [0, 1]$ is very different for low/high orness. We illustrate graphically the optimal α_j, $j = 1, \ldots, n$ coefficients in two parts, first for low orness values $\Omega \in [0, 0.5]$ and then for high orness values $\Omega \in [0.5, 1]$.

We observe that the optimal α_j, $j = 1, \ldots, n$ for low orness values $\Omega \in [0, 0.5]$ are all non negative and take values in the unit interval, independently of the dimension n. On the contrary, the optimal values of the α_j, $j = 1, \ldots, n$ coefficients for high orness values $\Omega \in [0.5, 1]$ depend strongly on the dimension n, both in the complexity of their distribution and in the amplitude of their scale.

1. INTRODUCTION

Multiple criteria decision making (MDCM) is a major branch of Operational Research which focuses on solving decision problems under various criteria and constraints. In the context of multiple criteria decision models, there is a variety of important constrained optimization problems involving ordered weighting averaging within the aggregation process.

[*]Author to whom all corrispondence should be addressed; e-mail silvia.bortot@unitn.it

In the binomial decomposition framework, due to Calvo and De Baets [8] and later discussed by Bortot and Marques Pereira [4], Bortot et al. [5],Bortot et al. [6] in the context of generalized Gini welfare functions, any OWA function with weights w_i with $i = 1, \ldots, n$ can be represented as a linear combination of binomial OWA functions C_j with $j = 1, \ldots, n$ whose coefficients α_j with $j = 1, \ldots, n$ are subject to boundary and monotonicity conditions. The coefficients α_j with $j = 1, \ldots, n$ are uniquely determined by the weighting structure of the OWA function, and vice-versa. Therefore an OWA function can be equivalently described either by its weights w_i with $i = 1, \ldots, n$ or by its coefficients α_j with $j = 1, \ldots, n$ in the binomial decomposition.

The maximum entropy approach to weight elicitation in ordered weighted averaging (OWA) was introduced by O'Hagan [47], [48], Filev and Yager [17], [18], and later discussed by Fullér and Majlender [21], Liu and Chen [39], Cheng and Chang [13], Yager [54], and Ahn [1]. Comprehensive reviews of the maximum entropy method can be found in Xu [52], Wu et al. [51], Liu [38], and Carlsson and Fullér [10].

In the maximum entropy approach, a unique OWA weighting structure is associated with every given orness value $\Omega \in [0, 1]$. The maximum entropy method is based on the solution of a constrained optimization problem: maximize entropy $S(\boldsymbol{w})$ given an orness value $\Omega \in [0, 1]$.

Applications of the maximal entropy methods approach can be found in various fields, see for instance Chang et al. [11], Liaw et al. [37], Yusoff and Merigó-Lindahl [57], Chuu [15], He et al. [34], and Kang et al. [35]. Two very recent applications of the maximum entropy method are described in Kim and Ahn [36], Brunelli and Fedrizzi [7]. The former studies the elicitation of orness values and its application in the choice of investment alternatives, while the latter suggests to use optimal entropy weights for the construction of inconsistency indices of pairwise comparison matrices in the AHP framework.

The report is organized as follows. In Section 2 we briefly present the basic concepts of OWA functions, entropy, and orness. We describe the maximum entropy constrained optimization problem and we discuss the analytic solution of the problem introduced by Filev and Yager in 1995 [17], and later by Fullér and Majlender in 2001 [21]. In Section 3 we briefly review the binomial decomposition of OWA functions and we describe the maximum entropy method in the context of the binomial decomposition. In Section 4 we describe the numerical results obtained with Matlab regarding the maximum entropy constrained optimization problem, firstly in terms of the OWA weights w_i with $i = 1, \ldots, n$ and secondly in terms of the binomial decomposition coefficients α_j with $j = 1, \ldots, n$ in dimensions $n = 3, 5, 7, 9$. The final Section contains some concluding remarks.

2. OWA FUNCTIONS AND THE MAXIMUM ENTROPY METHOD

We begin by introducing some basic notations and definitions regarding OWA functions, orness, and entropy in the context of averaging functions. We describe the maximum entropy approach, a method to obtain the OWA weights as the optimal solution of a constrained optimization problem.

We consider the standard framework of averaging functions on the $\mathbb{D}^n$ domain, with $\mathbb{D} \subseteq \mathbb{R}$ real interval and $n \geq 2$. Points in $\mathbb{D}^n$ are denoted $\boldsymbol{x} = (x_1, \ldots, x_n)$, with $\mathbf{1} = (1, \ldots, 1)$, $\mathbf{0} = (0, \ldots, 0)$. For every $x \in \mathbb{D}$, we have $x \cdot \mathbf{1} = (x, \ldots, x)$. Given $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{D}^n$, $\boldsymbol{x} \geq \boldsymbol{y}$ indicates that $x_i \geq y_i$ for every $i = 1, \ldots, n$, and by $\boldsymbol{x} > \boldsymbol{y}$ we mean $\boldsymbol{x} \geq \boldsymbol{y}$ and $\boldsymbol{x} \neq \boldsymbol{y}$. Given $\boldsymbol{x} \in \mathbb{D}^n$, the increasing reorderings of $\boldsymbol{x}$ are denoted as $x_{(1)} \leq \cdots \leq x_{(n)}$ and the decreasing ones as $x_{[1]} \geq \cdots \geq x_{[n]}$. In other words we can say that, $x_{(1)} = \min\{x_1, \ldots, x_n\} = x_{[n]}$ and $x_{(n)} = \max\{x_1, \ldots, x_n\} = x_{[1]}$.

We begin by defining averaging functions on the $\mathbb{D}^n$ domain. Complete reviews on averaging functions can be found in Fodor and Roubens [20], Calvo et al. [9], Torra et al. [50], Beliakov et al. [2], Grabisch et al. [32], Mesiar et al. [43], Grabisch et al. [33], and Beliakov et al. [3].

Definition 1 *A function* $A : \mathbb{D}^n \longrightarrow \mathbb{D}$ *is an* averaging function *if it is monotonic and idempotent. An averaging function is said to be* strict *if it is strictly monotonic. Notice that monotonicity and idempotency imply that* $\min(\boldsymbol{x}) \leq A(\boldsymbol{x}) \leq \max(\boldsymbol{x})$*, for all* $\boldsymbol{x} \in \mathbb{D}^n$*.*

We consider a particular instance of averaging functions, the ordered weighted averaging (OWA) function introduced by Yager [53]. The OWA function has been widely used to combine multiple objectives into a single objective function. A fundamental aspect of OWA function is the reordering step in the collection of the variables. This fact implies that the weights are associated with a particular ordered position and not with a specific variable.

Definition 2 *Given a weighting vector* $\boldsymbol{w} = (w_1, \ldots, w_n) \in [0,1]^n$*, with* $\sum_{i=1}^n w_i = 1$*, the* Ordered Weighted Averaging (OWA) function *associated with* $\boldsymbol{w}$ *is the averaging function* $A : \mathbb{D}^n \longrightarrow \mathbb{D}$ *defined as*

$$A(\boldsymbol{x}) = \sum_{i=1}^{n} w_i \, x_{(i)}. \tag{1}$$

The traditional form of OWA functions as introduced in [53] is $A(\boldsymbol{x}) = \sum_{i=1}^n \tilde{w}_i \, x_{[i]}$, where $\tilde{w}_i = w_{n-i+1}$. In [55, 56] the theory and applications of OWA functions are discussed in detail.

The OWA function allows to model various averaging functions from the minimum $\boldsymbol{w} = (1, 0, \ldots, 0, 0)$ through the arithmetic mean $\boldsymbol{w} = (1/n, 1/n, \ldots, 1/n, 1/n)$ to the maximum $\boldsymbol{w} = (0, 0, \ldots, 0, 1)$.

Yager has introduced two measures which characterize OWA functions, the orness and the entropy. Orness is the degree to which the OWA function resembles the maximum based on its weighting vector.

Definition 3 *Given an OWA function associated with a weighting vector* $\boldsymbol{w} = (w_1, \ldots, w_n) \in [0,1]^n$*, with* $\sum_{i=1}^n w_i = 1$*, the* orness *of an OWA function is defined as*

$$\Omega(\boldsymbol{w}) = \frac{1}{n-1} \sum_{i=1}^{n} (i-1) w_i \, . \tag{2}$$

The orness measure takes values in the unit interval. When orness is equal to 1, the OWA is the maximum. Conversely, when orness is equal to 0, then the OWA is the minimum. Values between 0 and 1 represent trade-offs between minimum and maximum.

The entropy of an OWA function provides a measure of weight dispersion.

DEFINITION 4 *Given an OWA function associated with a weighting vector* $\boldsymbol{w} = (w_1, \ldots, w_n) \in [0,1]^n$, *with* $\sum_{i=1}^n w_i = 1$, *the* entropy *of an OWA function is defined as*

$$S(\boldsymbol{w}) = -\sum_{i=1}^{n} w_i \, ln(w_i) \tag{3}$$

where we assume, as usual, the continuous extension $w \ln w = 0$ *when* $w = 0$.

The entropy measure reaches its maximum value when all the weights are equal to $1/n$. Consequently, we can say that the entropy measures the degree to which we use all the objectives equally.

We now define a set of linear orness preserving weight transformations which will be useful in the maximal entropy framework to be discussed later. These orness preserving weight transformations require $n \geq 3$, since weight transformations in $n = 2$ necessarily changes orness .

Given $n \geq 3$ and $\boldsymbol{w} = (w_1, \ldots, w_n) \in (0,1)^n$ with $w_1 + \ldots + w_n = 1$, we consider three weights $w_i, w_j, w_k \in (0,1)$ with $1 \leq i \leq j \leq k \leq n$.

An orness preserving linear weight transformation, which is defined as

$$w_i(t) = w_i + s_i \, t \,, \qquad w_j(t) = w_j + s_j \, t \,, \qquad w_k(t) = w_k + s_k \, t \tag{4}$$

for t in a right neighborhood of $t = 0$, satisfies

$$s_i + s_j + s_k = 0 \,, \qquad (i-1)\, s_i + (j-1)\, s_j + (k-1)\, s_k = 0 \tag{5}$$

where s_i, s_j, s_k are the slopes of the corresponding weight transformations. Therefore, combining the two equations in (5), we obtain

$$(i-k)\, s_i = -(j-k)\, s_j \,, \quad (j-i)\, s_j = -(k-i)\, s_k \,, \quad (k-j)\, s_k = -(i-j)\, s_i \tag{6}$$

which implies that only two *sign configurations* are possible, either

$$(s_i \oplus \quad s_j \ominus \quad s_k \oplus) \qquad \textit{or} \qquad (s_i \ominus \quad s_j \oplus \quad s_k \ominus) \,. \tag{7}$$

Consider for instance the case $n = 5$ and the uniform weight distribution $\boldsymbol{w} = (w_1 = 1/5, w_2 = 1/5, w_3 = 1/5, w_4 = 1/5, w_5 = 1/5)$. Considering $i = 2$, $j = 3$, $k = 5$, an orness preserving weight transformation in this case could be associated with $\boldsymbol{s} = (s_1 = 0, s_2 = 2/3, s_3 = -1, s_4 = 0, s_5 = 1/3)$.

2.1. Maximun entropy method

A fundamental issue in the theory of OWA functions is the determination of the associated weights. In the literature several methodologies have been introduced in order to obtain appropriate OWA weights, as discussed by Filev and Yager [18].

One of the primal methods of deriving the associated weights of the OWA function is proposed by O'Hagan [47], [48]. This approach is based on the constrained optimization method in which a predefined degree of orness is assumed, and the weights are computed by maximizing the entropy. O'Hagan called the optimal weights as maximum entropy OWA (MEOWA) weights.

Filev and Yager [17] introduced a method to generate the MEOWA weights without solving the maximum entropy constraint optimization problem. Fullér and Majlender [21] transformed the maximum entropy method into a polynomial equation method and Liu and Chen [39] introduced a general form of the MEOWA function using a parametric geometric approach. Comprehensive reviews of the maximum entropy method can be found in Xu [52], Wu et al. [51], Liu [38], and Carlsson and Fullér [10].

The maximum entropy method is based on the solution of the following constrained optimization problem,

$$\max_{\boldsymbol{w}} \ S(\boldsymbol{w}) = -\sum_{i=1}^{n} w_i \, ln \, w_i \tag{8}$$

subject to

$$\sum_{i=1}^{n} w_i = 1\,, \qquad \frac{1}{n-1}\sum_{i=1}^{n}(i-1)\,w_i = \Omega \tag{9}$$

where the weights w_i are subject to conditions $w_i \geq 0$ for $i = 1, \ldots, n$ and $\Omega \in [0,1]$ is a given degree of orness.

The two extreme orness cases $\Omega \in [0,1]$ are associated with unique weighting vectors: $\boldsymbol{w} = (w_1, w_2, \ldots, w_{n-1}, w_n) = (1, 0, \ldots, 0, 0)$ in the case $\Omega = 0$, and $\boldsymbol{w} = (w_1, w_2, \ldots, w_{n-1}, w_n) = (0, 0, \ldots, 0, 1)$ in the case $\Omega = 1$. In both cases the entropy $S(\boldsymbol{w})$ is null.

On the other hand, if orness lies in the open interval $\Omega \in (0,1)$, it can be shown that the optimal weights are always positive, which means that the constrained optimization problem (8)-(9) is a Lagrange problem with two equality constraints within the open domain $\boldsymbol{w} \in (0,1)^n$.

PROPOSITION 1 *Consider OWA functions associated with weighting vectors* $\boldsymbol{w} = (w_1, \ldots, w_n) \in [0,1]^n$, *with* $\sum_{i=1}^{n} w_i = 1$. *Given an orness value* $\Omega \in (0,1)$, *the associated maximum entropy weights are necessarily positive, that is,* $\boldsymbol{w} = (w_1, \ldots, w_n) \in (0,1)^n$.

Proof: Given $n \geq 2$ and orness values $\Omega \in (0,1)$, consider weighting vectors $\boldsymbol{w} = (w_1, \ldots, w_n) \in (0,1)^n$ with $w_1 + \ldots + w_n = 1$, and the entropy function

$$S(\boldsymbol{w}) = -\sum_{i=1}^{n} w_i \, ln \, w_i \ \in \ (0, ln n]\,. \tag{10}$$

In the $n = 2$ case, all the weighting vectors associated with orness values $\Omega \in (0, 1)$ are positive, and thus the proof is trivial. In the $n \geq 3$ cases, a general orness preserving weight transformation as in (4)-(5) defines a function

$$S(\boldsymbol{w}(t)) = -\sum_{i=1}^{n} w_i(t)\, ln\, w_i(t) \qquad (11)$$

whose derivative is

$$S'(\boldsymbol{w}(t)) = -\sum_{i=1}^{n} w_i'(t)\,(ln\, w_i(t) + 1) = -\sum_{i=1}^{n} s_i\, ln\,(w_i + s_i\, t) \qquad (12)$$

since $\sum_{i=1}^{n} s_i = 0$. In particular, an orness preserving transformation on three weights w_i, w_j, w_k with $1 \leq i \leq j \leq k \leq n$ leads to

$$S'(\boldsymbol{w}(t)) = -ln\,[(w_i + s_i\, t)^{s_i}\,(w_j + s_j\, t)^{s_j}\,(w_k + s_k\, t)^{s_k}] \qquad (13)$$

where s_i, s_j, s_k satisfy conditions (5).
For the first of the two possible sign configurations as in (7), we obtain

$$S'(\boldsymbol{w}(t)) = -ln\Big[\frac{(w_i + s_i\, t)^{s_i}\,(w_k + s_k\, t)^{s_k}}{(w_j - \,|s_j|\; t)^{|s_j|}}\Big] \qquad (14)$$

within a right neighborhood of $t = 0$ if the condition $w_i^{s_i}\, w_k^{s_k} < w_j^{|s_j|}$ holds. This condition in fact applies if either w_i and w_k are sufficiently small.

Analogously, for the second of the two possible sign configurations as in (7), we get

$$S'(\boldsymbol{w}(t)) = -ln\Big[\frac{(w_j + s_j\, t)^{s_j}}{(w_i - \,|s_i|\; t)^{|s_i|}\;(w_k - \,|s_k|\; t)^{|s_k|}}\Big] \qquad (15)$$

within a right neighborhood of $t = 0$ if the condition $w_j^{s_j} < w_i^{|s_i|}\, w_k^{|s_k|}$ holds. This condition in fact applies if w_j is sufficiently small.

Let us now consider the general case $n \geq 3$ and $\boldsymbol{w} = (w_1, \ldots, w_n) \in [0, 1]^n$ with $w_1 + \ldots + w_n = 1$. Weighting vectors $\boldsymbol{w}$ with $n-1$ null weights have zero entropy and are thus not maximal. Maximum entropy weighting vectors must necessarily have at least two positive weights, both in the $(0, 1)$ interval due to unit normalization.

Consider now the existence of a null weight, which is either on the left, center, or right of the two positive weights. In the first and third cases, we can show that such weighting vector is not maximal using the orness preserving weight transformation of sign configuration ($s_i \oplus \; s_j \ominus \; s_k \oplus$). As we have seen before, in such case the condition $w_i^{s_i}\, w_k^{s_k} < w_j^{|s_j|}$ after equation (14) holds and therefore the entropy increases as the null weight increases from 0.

Analogously, in the second case, when the null weight is between the two positive weights, we can show that such weighting vector is not maximal using the orness preserving weight transformation of sign configuration ($s_i \ominus \; s_j \oplus \; s_k \ominus$). In such case the condition $w_j^{s_j} < w_i^{|s_i|}\, w_k^{|s_k|}$ after equation (15) holds and therefore the entropy increases as the null weight increases from 0.

We have thus shown that weighting vectors with one or more null weights do not correspond to maximum entropy OWA functions with orness $\Omega \in (0, 1)$.

□

The Lagrange function of problem (8)-(9) can be denoted as

$$L(\boldsymbol{w}, \lambda_1, \lambda_2) = -\sum_{i=1}^{n} w_i\, ln(w_i) + \lambda_1\big(\sum_{i=1}^{n} w_i - 1\big) + \lambda_2\big(\frac{1}{n-1}\sum_{i=1}^{n}(i-1)\, w_i - \Omega\big) \quad (16)$$

where λ_1, λ_2 are the Lagrange multipliers. The partial derivatives of the Lagrange function can be computed as

$$\frac{\partial L}{\partial w_i} = -ln(w_i) - 1 + \lambda_1 + \lambda_2 \frac{i-1}{n-1} = 0 \qquad i = 1, \ldots, n \quad (17)$$

$$\frac{\partial L}{\partial \lambda_1} = \sum_{i=1}^{n} w_i - 1 = 0 \quad (18)$$

$$\frac{\partial L}{\partial \lambda_2} = \frac{1}{n-1}\sum_{i=1}^{n}(i-1)w_i - \Omega = 0 \quad (19)$$

where $\Omega \in (0, 1)$ is a given degree of orness.

In the following subsections we describe the analytic solution of the maximum entropy method introduced by Filev and Yager in 1995 [17], and by Fullér and Majlender in 2001 [21].

2.2. Filev and Yager's analytic construction of MEOWA weights

Filev and Yager in 1995 [17] introduced a method to obtain the optimal weights thus solving the maximum entropy constraint optimization problem (8)-(9). Using the Lagrange multiplier method they derived an exponential equation which determines the optimal weighting vector.

Consider the constrained optimization problem described in (8)-(9). Expression (17) can be written as

$$ln(w_i) + 1 = \lambda_1 + \lambda_2 \frac{i-1}{n-1}$$

$$w_i = e^{\lambda_1 + \lambda_2 \frac{i-1}{n-1} - 1} \quad (20)$$

$$\sum_{i=1}^{n} w_i = \sum_{i=1}^{n} e^{\lambda_1 + \lambda_2 \frac{i-1}{n-1} - 1}. \quad (21)$$

Using expression (18) we obtain

$$e^{\lambda_1 - 1} \sum_{i=1}^{n} e^{\lambda_2 \frac{i-1}{n-1}} = 1\,. \quad (22)$$

Dividing expression (20) by expression (22) we obtain the expression for the optimal weights w_i, $i = 1, \ldots, n$,

$$w_i = \frac{e^{\lambda_1 - 1}\, e^{\lambda_2 \frac{i-1}{n-1}}}{e^{\lambda_1 - 1} \sum_{j=1}^{n} e^{\lambda_2 \frac{j-1}{n-1}}} = \frac{e^{\lambda_2 \dfrac{i-1}{n-1}}}{\sum_{j=1}^{n} e^{\lambda_2 \frac{j-1}{n-1}}}\,. \quad (23)$$

We can substitute the weight expression (23) into (19) to obtain a non-linear equation connecting the given degree of orness Ω and the Lagrange multiplier λ_2 in the following way,

$$\frac{1}{n-1}\sum_{i=1}^{n}(i-1)\frac{e^{\lambda_2\frac{i-1}{n-1}}}{\sum_{j=1}^{n} e^{\lambda_2\frac{j-1}{n-1}}} = \Omega \tag{24}$$

or alternatively,

$$\sum_{i=1}^{n}\left(\frac{i-1}{n-1}-\Omega\right) e^{\lambda_2\frac{i-1}{n-1}} = 0\,. \tag{25}$$

Clearly, using expressions (23) and (25) we can obtain directly the optimal weights given a specified degree of orness Ω by first computing the Lagrange multiplier λ_2 from equation (25) and then by computing the Lagrange multiplier λ_1 from (22), thus solving the constrained optimization problem in (8)-(9).

2.3 Fullér and Majlender's analytic construction of MEOWA weights

Fullér and Majlender in 2001 [21] also introduced a method to obtain the optimal weights thus solving the maximum entropy constraint optimization problem (8)-(9). Using the Lagrange multiplier method they derived a polynomial equation whose solution can determine the optimal weighting vector.

Consider the constrained optimization problem described in (8)-(9). For $i = 1$ equation (17) becomes

$$-ln(w_1) - 1 + \lambda_1 = 0 \Rightarrow \lambda_1 = ln(w_1) + 1\,. \tag{26}$$

For $i = n$ equation (17) becomes

$$-ln(w_n) - 1 + \lambda_1 + \lambda_2 = 0\,. \tag{27}$$

Substituting expression (26) in (27) we obtain

$$\lambda_2 = ln(w_n) - ln(w_1)\,. \tag{28}$$

For $1 \le i \le n$, using (26) and (28), equation (17) turns into

$$ln(w_i) = \frac{n-i}{n-1}\,ln(w_1) + \frac{i-1}{n-1}\,ln(w_n) \tag{29}$$

or equivalently,

$$w_i = \sqrt[n-1]{w_1^{n-i}\, w_n^{i-1}}. \tag{30}$$

In the case that $w_1 = w_n$ then from (30) we have that

$$w_1 = w_2 = \ldots = w_n = \frac{1}{n}$$

which corresponds to the optimal solution of the maximum entropy constraint optimization problem (8) for orness $\Omega = 0.5$. In fact, when $w_1 = w_2 = \ldots = w_n = \frac{1}{n}$

Author's Accepted Manuscript Intl. J. Intelligent Systems DOI http://dx.doi.org/10.1002/int.22083

then $S(\boldsymbol{w}) = ln(n)$ which is the global optimal value of dispersion for any OWA function.

Now consider the case $w_1 \neq w_n$. Introducing the notation

$$u_1 = w_1^{1/(n-1)} \quad \textit{and} \quad u_n = w_n^{1/(n-1)} \tag{31}$$

in expression (30) we get

$$w_i = u_1^{n-i}\, u_n^{i-1}\,, \quad 1 \leq i \leq n. \tag{32}$$

In what follows, we will determine the optimal weight distribution by means of the implicit solution of a polynomial equation on w_n. For this purpose, we make use of the following general results.

PROPOSITION 2 *Given $a, b \in \mathbb{R}$ with $a \neq b$ we have that*

$$\sum_{i=1}^{n} a^{n-i}\, b^{i-1} = \frac{a^n - b^n}{a - b}$$

or equivalently,

$$(a-b) \sum_{i=1}^{n} a^{n-i}\, b^{i-1} = a^n - b^n. \tag{33}$$

Proof: Let us use the following notation: $\sum_{i=1}^{n} a^{n-i}\, b^{i-1} = (*)$. We have the following,

$$\begin{aligned}
(a-b)\,(*) &= a \sum_{i=1}^{n} a^{n-i}\, b^{i-1} - b \sum_{i=1}^{n} a^{n-i}\, b^{i-1} \\
&= \sum_{i=1}^{n} a^{n-i+1}\, b^{i-1} - \sum_{i=1}^{n} a^{n-i}\, b^{i} \\
&= \left(a^n b^0 + a^{n-1} b^1 + \ldots + a^2 b^{n-2} + a^1 b^{n-1}\right) - \\
&\quad - \left(a^{n-1} b^1 + a^{n-2} b^2 + \ldots + a^1 b^{n-1} + a^0 b^n\right) \\
&= a^n - b^n.
\end{aligned}$$

□

PROPOSITION 3 *Given $a, b \in \mathbb{R}$ with $a \neq b$ we have that*

$$\sum_{i=1}^{n} i\, a^{n-i}\, b^{i-1} = \frac{1}{a-b}\left[a \sum_{i=1}^{n} a^{n-i}\, b^{i-1} - n b^n\right] \tag{34}$$

Author's Accepted Manuscript Intl. J. Intelligent Systems DOI http://dx.doi.org/10.1002/int.22083

Proof: Let us use the following notation: $\sum_{i=1}^{n} i\, a^{n-i}\, b^{i-1} = (**)$. We have the following,

$$
\begin{aligned}
a\,(**) &= \sum_{i=1}^{n} i\, a^{n-i+1}\, b^{i-1} \\
b\,(**) &= \sum_{i=1}^{n} i\, a^{n-i}\, b^{i} = \sum_{j=2}^{n} (j-1)\, a^{n-j+1}\, b^{j-1} \\
&= \sum_{i=1}^{n} (i-1)\, a^{n-i+1}\, b^{i-1} + nb^{n}\,. \\
a\,(**) - b\,(**) &= \sum_{i=1}^{n} a^{n-i+1}\, b^{i-1} - nb^{n} = a \sum_{i=1}^{n} a^{n-i}\, b^{i-1} - nb^{n} \\
(**) &= \frac{1}{a-b}\left[\, a\,(*) - nb^{n}\,\right].
\end{aligned}
$$

□

These two general results will now be instrumental to obtain the basic equation which determines the maximum entropy OWA weights, as described below.

From equation (18) and Proposition 2 we have

$$
\begin{aligned}
\sum_{i=1}^{n} u_1^{n-i}\, u_n^{i-1} = 1 \quad &\Leftrightarrow \quad \frac{u_1^n - u_n^n}{u_1 - u_n} = 1 \\
&\Leftrightarrow \quad u_1^n - u_n^n = u_1 - u_n. \qquad (35)
\end{aligned}
$$

We recall equation (19). By substituting (32) in (19) we obtain

$$
\sum_{i=1}^{n} (i-1)\, u_1^{n-i}\, u_n^{i-1} = (n-1)\,\Omega\,. \qquad (36)
$$

The left hand side of equation (36) from can be expressed as

$$
\sum_{i=1}^{n} (i-1)\, u_1^{n-i}\, u_n^{i-1} = \sum_{i=1}^{n} i\, u_1^{n-i}\, u_n^{i-1} - \sum_{i=1}^{n} u_1^{n-i}\, u_n^{i-1} \qquad (37)
$$

and, using Proposition 3 we have

$$
\begin{aligned}
\sum_{i=1}^{n} (i-1)\, u_1^{n-i}\, u_n^{i-1} &= \frac{1}{u_1 - u_n}\Big[u_1 \sum_{i=1}^{n} u_1^{n-i}\, u_n^{i-1} - nu_n^n - (u_1 - u_n) \sum_{i=1}^{n} u_1^{n-i}\, u_n^{i-1} \Big] \\
&= \frac{1}{u_1 - u_n}\Big[u_n \sum_{i=1}^{n} u_1^{n-i}\, u_n^{i-1} - nu_n^n \Big] \qquad (38)
\end{aligned}
$$

thus obtaining

$$
\begin{aligned}
\sum_{i=1}^{n} (i-1)\, u_1^{n-i}\, u_n^{i-1} &= \frac{1}{u_1 - u_n}\Big[u_n \big(\frac{u_1^n - u_n^n}{u_1 - u_n}\big) - nu_n^n \Big] \\
&= \frac{1}{u_1 - u_n}\big(u_n - nu_n^n \big)\,. \qquad (39)
\end{aligned}
$$

Therefore, equation (36) can be expressed as

$$\begin{aligned}
\frac{1}{u_1 - u_n}\Big(u_n - n u_n^n\Big) &= (n-1)\,\Omega \\
u_n(1 - n u_n^{n-1}) &= u_1(n-1)\,\Omega - u_n(n-1)\,\Omega \\
u_n\big(1 - n u_n^{n-1} + (n-1)\,\Omega\big) &= u_1(n-1)\,\Omega \\
\frac{u_1}{u_n} &= \frac{(n-1)\,\Omega + 1 - n w_n}{(n-1)\,\Omega}\,.
\end{aligned} \tag{40}$$

Rewriting equation (35) as

$$\begin{aligned}
u_1^n - u_n^n = u_1 - u_n \quad &\Leftrightarrow \quad \frac{u_1^n}{u_n} - u_n^{n+1} = \frac{u_1}{u_n} - 1 \\
&\Leftrightarrow \quad u_1^{n-1}\frac{u_1}{u_n} - u_n^{n-1} = \frac{u_1}{u_n} - 1
\end{aligned} \tag{41}$$

and, using equation (40), we obtain

$$\begin{aligned}
u_1^{n-1}\Big(\frac{(n-1)\,\Omega + 1 - n w_n}{(n-1)\,\Omega}\Big) - u_n^{n-1} &= \frac{1 - n w_n}{(n-1)\,\Omega} \\
w_1((n-1)\,\Omega + 1 - n w_n) - w_n(n-1)\,\Omega &= 1 - n w_n \\
w_1 &= \frac{((n-1)\,\Omega - n)w_n + 1}{(n-1)\,\Omega + 1 - n w_n}\,.
\end{aligned} \tag{42}$$

Equation (35) can also be rewritten as

$$\begin{aligned}
u_1^n - u_n^n = u_1 - u_n \quad &\Leftrightarrow \quad u_1(w_1 - 1) = u_n(w_n - 1) \\
&\Leftrightarrow \quad w_1(w_1 - 1)^{n-1} = w_n(w_n - 1)^{n-1}
\end{aligned} \tag{43}$$

and using equation (42), we obtain

$$w_n(w_n - 1)^{n-1} = \frac{\big((n-1)\,\Omega - n\big)w_n + 1}{(n-1)\,\Omega + 1 - n w_n}\Big[\frac{(n-1)\,\Omega(w_n - 1)}{(n-1)\,\Omega + 1 - n w_n}\Big]^{n-1}$$

leading to the fundamental polynomial equation for the weight w_n in relation with a given orness value $\Omega \in (0, 1)$,

$$w_n\big[(n-1)\,\Omega + 1 - n w_n\big]^n = \big((n-1)\,\Omega\big)^{n-1}\big[\big((n-1)\,\Omega - n\big)w_n + 1\big]\,. \tag{44}$$

Thus, the optimal value of w_n should satisfy equation (44). Fullér and Majlender in [21] proved that there exists a unique meaningful $w_n \in (0, 1)$ for this equation, given an orness value $\Omega \in (0, 1)$. Once the weight w_n is obtained, the weight w_1 is determined by equation (42). The other weights are obtained using equation (30).

2.4. Gajdos' generating function framework

Author's Accepted Manuscript Intl. J. Intelligent Systems DOI http://dx.doi.org/10.1002/int.22083

The Ordered Weighted Averaging (OWA) functions in the general framework introduced by Gajdos [22] are of the form

$$A_f(\boldsymbol{x}) = \sum_{i=1}^{n} \left[f\left(\frac{n-i+1}{n}\right) - f\left(\frac{n-i}{n}\right) \right] x_{(i)} \tag{45}$$

where f is a continuous and increasing function on the unit interval, with $f(0) = 0$ and $f(1) = 1$. The general graphical representation of Gajdos' weights with generating function f is illustrated in Fig. 1. An analogous and equivalent weight generation mechanism, based on the notion of quantifier, has been discussed by Yager in [54].

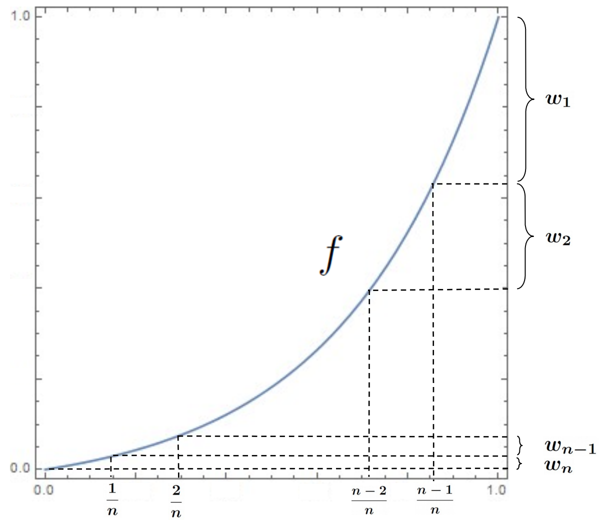


Figure 1: General representation of Gajdos' weights.

Concerning the generating function f, the integer parametric choices $f(t) = t^k$, with $k = 1, \ldots, n$, can be seen in relation with the k-additivity of the OWA function, as discussed in Gajdos [22]. In what follows, however, we wish to discuss the choice of an exponential form for the weight generating function.

PROPOSITION 4 *The maximum entropy OWA functions can be written in the general framework (45) introduced by Gajdos with the following weight generating function* $f : [0, 1] \to [0, 1]$,

$$f(x) = \frac{e^{\alpha x} - 1}{e^{\alpha} - 1} \quad \alpha \neq 0\,, \quad f(x) = x \quad \alpha = 0\,, \qquad x \in [0, 1] \tag{46}$$

which is continuous and increasing on the unit interval, with $f(0) = 0$ *and* $f(1) = 1$.

Proof: Consider the exponential function $f : [0, 1] \to [0, 1]$,

$$f(x) = \frac{e^{\alpha x} - 1}{e^{\alpha} - 1} \quad \alpha \neq 0\,, \quad f(x) = x \quad \alpha = 0\,, \qquad x \in [0, 1]\,. \tag{47}$$

Clearly, the function f is continuous and increasing on the unit interval, with $f(0) = 0$ and $f(1) = 1$.

The weighting structure of the OWA function $A_f(\boldsymbol{x}) = w_i\, x_{(i)}$ associated to function f is of the form

$$w_i = f\Big(\frac{n-i+1}{n}\Big) - f\Big(\frac{n-i}{n}\Big) = \frac{e^{\alpha \frac{(n-i+1)}{n}} - e^{\alpha \frac{(n-i)}{n}}}{e^{\alpha} - 1} \qquad i = 1, \dots, n \tag{48}$$

where we assume $\alpha \neq 0$, otherwise $w_i = 1/n$ for $i = 1, \dots, n$. Normalization follows immediately,

$$\begin{aligned}\sum_{i=1}^{n} w_i &= \frac{1}{e^{\alpha}-1}\Big[e^{\alpha} - e^{\alpha \frac{(n-1)}{n}} + e^{\alpha \frac{(n-1)}{n}} - e^{\alpha \frac{(n-2)}{n}} + \dots \\ &\qquad \dots + e^{\alpha \frac{2}{n}} - e^{\alpha \frac{1}{n}} + e^{\alpha \frac{1}{n}} - 1\Big] \\ &= \frac{1}{e^{\alpha}-1}\big(e^{\alpha} - 1\big) = 1\,. \end{aligned} \tag{49}$$

We will now study the relation between consecutive weights w_i, w_{i+1}, for $i = 1, \dots, n-1$,

$$\begin{aligned} w_i - w_{i+1} &= \frac{e^{\alpha \frac{(n-i+1)}{n}} - 1}{e^{\alpha} - 1} - 2\Big(\frac{e^{\alpha \frac{(n-1)}{n}} - 1}{e^{\alpha}-1}\Big) + \frac{e^{\alpha \frac{(n-i-1)}{n}} - 1}{e^{\alpha}-1} \\ &= \frac{1}{e^{\alpha}-1}\Big[e^{\alpha \frac{(n-i+1)}{n}} - 2e^{\alpha \frac{(n-1)}{n}} + e^{\alpha \frac{(n-i-1)}{n}}\Big] \\ &= \frac{1}{e^{\alpha}-1}\, e^{\alpha \frac{(n-i-1)}{n}} \Big(e^{\alpha \frac{2}{n}} - 2e^{\alpha \frac{1}{n}} + 1\Big). \end{aligned} \tag{50}$$

Introducing the notation $t = e^{\alpha \frac{1}{n}}$, the expression $(e^{\alpha \frac{2}{n}} - 2e^{\alpha \frac{1}{n}} + 1)$ becomes $(t^2 - 2t + 1)$, which is equal to zero if and only if $t = 1$, otherwise it is positive. The sign of the weight differences $w_i - w_{i+1}$ for $i = 1, \dots, n-1$, is therefore uniform along the weighting vector, and depends exclusively on the sign of the parameter α, through the denominator $e^{\alpha} - 1$.

By substituting $\alpha = -\beta\, n/(n-1)$ in expression (48) we obtain

$$\begin{aligned} w_i &= \frac{e^{-\beta \frac{(n-i+1)}{n-1}} - e^{-\beta \frac{(n-i)}{n-1}}}{e^{-\beta \frac{n}{n-1}} - 1} = \frac{e^{-\beta \frac{n}{n-1}} \Big(e^{-\beta \frac{1}{n-1}} - 1\Big)\, e^{\beta \frac{i}{n-1}}}{e^{-\beta \frac{n}{n-1}} - 1} \\ &= e^{\beta \frac{i}{n-1}} \frac{1 - e^{-\beta \frac{1}{n-1}}}{e^{\beta \frac{n}{n-1}} - 1} = e^{\beta \frac{i}{n-1}} \frac{1}{\sum_{j=1}^{n} e^{\beta \frac{j}{n-1}}} \end{aligned} \tag{51}$$

where we have used that

$$\sum_{j=1}^{n} e^{\beta \frac{j}{n-1}} = e^{\beta \frac{1}{n-1}} \frac{e^{\beta \frac{n}{n-1}} - 1}{e^{\beta \frac{1}{n-1}} - 1} = \frac{e^{\beta \frac{n}{n-1}} - 1}{1 - e^{-\beta \frac{1}{n-1}}} \tag{52}$$

which follows from the formula for the geometric series

$$c + cr + cr^2 + \ldots + cr^n = c\,\frac{r^{n-1}}{r-1} \tag{53}$$

with $c = e^{\beta/(n-1)} = r \neq 1$ for any $\alpha, \beta \neq 0$.

The form obtained in (51) for the exponential Gajdos' weights coincides with that of the MEOWA weights as obtained by Filev and Yager in [17],

$$w_i \;=\; \frac{e^{\beta\,\frac{(i-1)}{n-1}}}{\sum_{j=1}^{n} e^{\beta\,\frac{(j-1)}{n-1}}} = \frac{e^{\beta\,\frac{i}{n-1}}}{\sum_{j=1}^{n} e^{\beta\,\frac{j}{n-1}}}\,. \tag{54}$$

□

In relation with the sign of the parameter α, which is opposite to that of the parameter β, we obtain decreasing weights for $\alpha > 0$, as illustrated in Figure 1, constant weights for $\alpha = 0$, and increasing weights for $\alpha < 0$, as explained after equation (50).

3. THE BINOMIAL DECOMPOSITION OF OWA FUNCTIONS

In this section we present a brief review of the basic facts on Choquet integration, focusing on the Möbius representation framework. For recent reviews of Choquet integration see Grabisch and Labreuche [29], [30], [31], and Grabisch et al. [28] for the general case, Mayag et al. [41], [42] for the 2-additive case in particular.

We then briefly review the concept of binomial decomposition in terms of the binomial OWA functions C_j and, finally, we describe the maximum entropy approach in the context of the binomial decomposition.

Consider a finite set of interacting elements $N = \{1, 2, \ldots, n\}$. Any subsets S, $T \subseteq N$ with cardinalities $0 \leq s, t \leq n$ are usually called coalitions. The concepts of capacity and Choquet integral in the definitions below are due to Choquet [14], Sugeno [49], Chateauneuf and Jaffray [12], Murofushi and Sugeno [46], Denneberg [16], Grabisch [23], [24], and Marichal [40].

Definition 5 *A* capacity *on the set N is a set function $\mu : 2^N \longrightarrow [0,1]$ satisfying*

(i) $\mu(\emptyset) = 0$, $\mu(N) = 1$ *(boundary conditions)*

(ii) $S \subseteq T \subseteq N \quad \Rightarrow \quad \mu(S) \leq \mu(T)$ *(monotonicity conditions).*

Definition 6 *Let μ be a capacity on N. The* Choquet integral $\mathcal{C}_\mu : \mathbb{D}^n \longrightarrow \mathbb{D}$ *with respect to μ is defined as*

$$\mathcal{C}_\mu(\boldsymbol{x}) = \sum_{i=1}^{n} \left[\mu(A_{(i)}) - \mu(A_{(i+1)})\right] x_{(i)} \qquad \boldsymbol{x} = (x_1, \ldots, x_n) \in \mathbb{D}^n \tag{55}$$

where $(\cdot)$ indicates a permutation on N such that $x_{(1)} \leq x_{(2)} \leq \cdots \leq x_{(n)}$. Moreover, $A_{(i)} = \{(i), \ldots, (n)\}$ and $A_{(n+1)} = \emptyset$.

Definition 7 *Let μ be a capacity on the set N. The* Möbius transform $m_\mu : 2^N \longrightarrow \mathbb{R}$ *associated with the capacity μ is defined as*

$$m_\mu(T) = \sum_{S \subseteq T} (-1)^{t-s} \mu(S) \qquad T \subseteq N \tag{56}$$

where s and t denote the cardinality of the coalitions S and T, respectively.

Conversely, given the Möbius transform m_μ, the associated capacity μ is obtained as

$$\mu(T) = \sum_{S \subseteq T} m_\mu(S) \qquad T \subseteq N\,. \tag{57}$$

In the Möbius representation, the boundary conditions take the form

$$m_\mu(\emptyset) = 0 \qquad \sum_{T \subseteq N} m_\mu(T) = 1 \tag{58}$$

and the monotonicity conditions can be expressed as follows: for each $i = 1, \ldots, n$ and each coalition $T \subseteq N \setminus \{i\}$, the monotonicity condition is written as

$$\sum_{S \subseteq T} m_\mu(S \cup \{i\}) \geq 0 \qquad T \subseteq N \setminus \{i\} \qquad i = 1, \ldots, n\,. \tag{59}$$

This form of the monotonicity conditions derives from the original monotonicity conditions in Definition 5, expressed as $\mu(T \cup \{i\}) - \mu(T) \geq 0$ for each $i \in N$ and $T \subseteq N \setminus \{i\}$.

Defining a capacity μ on a set N of n elements requires $2^n - 2$ real coefficients, corresponding to the capacity values $\mu(T)$ for $T \subseteq N$. In order to control exponential complexity, Grabisch [25] introduced the concept of k-additive capacities, which has been further discussed in Grabisch [26], Miranda and Grabisch [44], and Miranda et al. [45].

Definition 8 *A capacity μ on the set N is said to be* k-additive *if its Möbius transform satisfies $m_\mu(T) = 0$ for all $T \subseteq N$ with $t > k$, and there exists at least one coalition $T \subseteq N$ with $t = k$ such that $m_\mu(T) \neq 0$.*

In the k-additive case, with $k = 1, \ldots, n$, the capacity μ is expressed as follows in terms of the Möbius transform m_μ,

$$\mu(T) = \sum_{S \subseteq T,\, s \leq k} m_\mu(S) \qquad T \subseteq N \tag{60}$$

and the boundary and monotonicity conditions (58) and (59) take the form

$$m_\mu(\emptyset) = 0 \qquad \sum_{T \subseteq N,\, t \leq k} m_\mu(T) = 1 \tag{61}$$

$$\sum_{S \subseteq T,\, s \leq k-1} m_\mu(S \cup \{i\}) \geq 0 \qquad T \subseteq N \setminus \{i\} \qquad i = 1, \ldots, n\,. \tag{62}$$

Finally, we examine the particular case of symmetric capacities and Choquet integrals, which play a crucial role in this paper.

Author's Accepted Manuscript Intl. J. Intelligent Systems DOI http://dx.doi.org/10.1002/int.22083

DEFINITION 9 *A capacity* μ *is said to be* symmetric *if it depends only on the cardinality of the coalition considered, in which case we use the simplified notation*

$$\mu(T) = \mu(t) \qquad \textit{where} \qquad t = |T| \,. \tag{63}$$

Accordingly, for the Möbius transform m_μ associated with a symmetric capacity μ we use the notation

$$m_\mu(T) = m_\mu(t) \qquad \text{where} \qquad t = |T| \,. \tag{64}$$

In the symmetric case, the expression (57) for the capacity μ in terms of the Möbius transform m_μ reduces to

$$\mu(t) = \sum_{s=1}^{t} \binom{t}{s} m_\mu(s) \qquad\qquad t = 1, \ldots, n \tag{65}$$

and the boundary and monotonicity conditions (58) and (59) take the form

$$m_\mu(0) = 0 \qquad\qquad \sum_{s=1}^{n} \binom{n}{s} m_\mu(s) = 1 \tag{66}$$

$$\sum_{s=1}^{t} \binom{t-1}{s-1} m_\mu(s) \geq 0 \qquad\qquad t = 1, \ldots, n \,. \tag{67}$$

The monotonicity conditions correspond to $\mu(t) - \mu(t-1) \geq 0$ for $t = 1, \ldots, n$.

The Choquet integral (55) with respect to a symmetric capacity μ reduces to an Ordered Weighted Averaging (OWA) function, see Fodor et al. [19], and Yager [53],

$$\mathcal{C}_\mu(\boldsymbol{x}) = \sum_{i=1}^{n} [\mu(n-i+1) - \mu(n-i)] x_{(i)} = \sum_{i=1}^{n} w_i \, x_{(i)} = A(\boldsymbol{x}) \tag{68}$$

where the weights $w_i = \mu(n-i+1) - \mu(n-i)$ satisfy $w_i \geq 0$ for $i = 1, \ldots, n$ due to the monotonicity of the capacity μ, and $\sum_{i=1}^{n} w_i = 1$ due to the boundary conditions $\mu(0) = 0$ and $\mu(n) = 1$. Comprehensive reviews of OWA functions can be found in Yager and Kacprzyk [55] and Yager et al. [56].

The weighting structure of the OWA function (68) is thus of the general form

$$w_i = f\left(\frac{n-i+1}{n}\right) - f\left(\frac{n-i}{n}\right) \tag{69}$$

where f is a continuous and increasing function on the unit interval, with $f(0) = 0$ and $f(1) = 1$. As mentioned before, Gajdos [22] has shown that the OWA function A is associated with a k-additive capacity μ, with $k = 1, \ldots, n$, if and only if f is polynomial of order k. In fact, in (65), the k-additive case is obtained simply by taking

$$m_\mu(k+1) = \cdots = m_\mu(n) = 0 \,, \qquad\qquad m_\mu(k) = (t-1) \ldots (t-k+1)/k! \tag{70}$$

Author's Accepted Manuscript Intl. J. Intelligent Systems DOI http://dx.doi.org/10.1002/int.22083

which is polynomial of order k in the coalition cardinality t.

We now consider OWA functions $A : \mathbb{D}^n \longrightarrow \mathbb{D}$ and we recall the binomial decomposition of OWA functions due to Calvo and De Baets [8], with the addition of a uniqueness result, see also Bortot and Marques Pereira [4].

We begin by introducing the convenient notation

$$\alpha_j = \binom{n}{j} m_\mu(j) \qquad j = 1, \ldots, n\,. \tag{71}$$

In this notation, which has no relation with that of the parameter α in Section 2.4, the upper boundary condition (66) reduces to

$$\sum_{j=1}^{n} \alpha_j = 1 \tag{72}$$

and the monotonicity conditions (67) take the form

$$\sum_{j=1}^{i} \frac{\binom{i-1}{j-1}}{\binom{n}{j}} \alpha_j \geq 0 \qquad\qquad i = 1, \ldots, n\,. \tag{73}$$

Definition 10 *The* binomial OWA functions $C_j : \mathbb{D}^n \longrightarrow \mathbb{D}$*, with* $j = 1, \ldots, n$*, are defined as*

$$C_j(\boldsymbol{x}) = \sum_{i=1}^{n} w_{ji}\, x_{(i)} = \sum_{i=1}^{n} \frac{\binom{n-i}{j-1}}{\binom{n}{j}} x_{(i)} \qquad j = 1, \ldots, n \tag{74}$$

where the binomial weights w_{ji}*,* $i, j = 1, \ldots, n$ *are null when* $i + j > n + 1$ *according to the usual convention that* $\binom{p}{q} = 0$ *when* $p < q$*, with* $p, q = 0, 1, \ldots$

Except for $C_1(\boldsymbol{x}) = \bar{x}$, the binomial OWA functions C_j, $j = 2, \ldots, n$ have an increasing number of null weights, in correspondence with $x_{(n-j+2)}, \ldots, x_{(n)}$. The weight normalization of the binomial OWA functions, $\sum_{i=1}^{n} w_{ji} = 1$ for $j = 1, \ldots, n$, is due to the column-sum property of binomial coefficients,

$$\sum_{i=1}^{n} \binom{n-i}{j-1} = \sum_{i=0}^{n-1} \binom{i}{j-1} = \binom{n}{j} \qquad j = 1, \ldots, n\,. \tag{75}$$

Proposition 5 [Binomial decomposition] *Any OWA function* $A : \mathbb{D}^n \longrightarrow \mathbb{D}$ *can be written uniquely as*

$$A(\boldsymbol{x}) = \alpha_1 C_1(\boldsymbol{x}) + \alpha_2 C_2(\boldsymbol{x}) + \cdots + \alpha_n C_n(\boldsymbol{x}) \tag{76}$$

where the coefficients α_j*,* $j = 1, \ldots, n$ *are subject to conditions (72) and (73). In the binomial decomposition (76) the* k*-additive case, with* $k = 1, \ldots, n$*, is obtained simply by taking* $\alpha_{k+1} = \cdots = \alpha_n = 0$*.*

Author's Accepted Manuscript Intl. J. Intelligent Systems DOI http://dx.doi.org/10.1002/int.22083

Consider an OWA function $A : \mathbb{D}^n \longrightarrow \mathbb{D}$ associated with a symmetric capacity μ as in (68),

$$A(\boldsymbol{x}) = \sum_{i=1}^{n} w_i\, x_{(i)} \qquad w_i = \mu(n-i+1) - \mu(n-i) \tag{77}$$

where the weights w_i, $i = 1, \ldots, n$ can be expressed, using (76), in the form

$$\begin{aligned} w_i &= \sum_{j=1}^{n} w_{ji}\, \alpha_j = \sum_{j=1}^{n-i+1} w_{ji}\, \alpha_j \qquad i = 1, \ldots, n \\ &= w_{1i}\, \alpha_1 + w_{2i}\, \alpha_2 + \ldots + w_{n-i,i}\, \alpha_{n-i} + w_{n-i+1,i}\, \alpha_{n-i+1} \end{aligned} \tag{78}$$

and the coefficients α_j, $j = 1, \ldots, n$ are subject to conditions (72) and (73). Notice that the binomial weights w_{ji}, $i, j = 1, \ldots, n$ are null when $i+j > n+1$ as explained in (74). More explicitly, the weights w_i with $i = 1, \ldots, n$ can be written as

$$\begin{cases} w_1 = w_{11}\alpha_1 + w_{21}\alpha_2 + \cdots + w_{n-1,1}\alpha_{n-1} + w_{n1}\alpha_n \\ w_2 = w_{12}\alpha_1 + w_{22}\alpha_2 + \cdots + w_{n-1,2}\alpha_{n-1} \\ \qquad \ldots \\ w_{n-1} = w_{1,n-1}\alpha_1 + w_{2,n-1}\alpha_2 \\ w_n = w_{1n}\alpha_1 \end{cases} \tag{79}$$

In the case $n = 3$, for instance, we obtain

$$\begin{cases} w_1 = \dfrac{1}{3}\,\alpha_1 + \dfrac{2}{3}\,\alpha_2 + \alpha_3 \\ w_2 = \dfrac{1}{3}\,\alpha_1 + \dfrac{1}{3}\,\alpha_2 \\ w_3 = \dfrac{1}{3}\,\alpha_1\,. \end{cases}$$

It is clear from expression (79) that the values of the coefficients α_j, $j = 1, \ldots, n$ are uniquely determined by the weighting structure of the ordered weighted averaging function A: w_n determines α_1, then w_{n-1} determines α_2, and so on.

4. MAXIMUM ENTROPY METHOD IN THE CONTEXT OF THE BINOMIAL DECOMPOSITION

Let us now consider the maximum entropy method in the context of the binomial decomposition using the coefficients α_j, with $j = 1, \ldots, n$, to describe the OWA function instead of the weights w_i, with $i = 1, \ldots, n$. As we have seen, the coefficients α_j, with $j = 1, \ldots, n$, are uniquely determined by the weighting structure of the OWA function, and vice-versa as indicated in expression (79). This means that we can express the measures of orness (2) and entropy (3) in terms of the vector $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_n)$ containing the coefficients α_j, with $j = 1, \ldots, n$, in the following way

Author's Accepted Manuscript Intl. J. Intelligent Systems DOI http://dx.doi.org/10.1002/int.22083

PROPOSITION 6 *Given an OWA function associated with the binomial decomposition coefficients α_j, with $j = 1, \ldots, n$, the* orness *of an OWA function can be written as*

$$\Omega(\boldsymbol{\alpha}) = \frac{1}{n-1} \sum_{i=1}^{n} \sum_{j=1}^{n-i+1} (i-1) \frac{\binom{n-i}{j-1}}{\binom{n}{j}} \alpha_j \tag{80}$$

where the coefficients α_j, with $j = 1, \ldots, n$ are subject to conditions (72) and (73).

Proof. The expression of orness (80) is obtained directly from (2) by substituting (78). In the case $n = 3$, for instance, we obtain

$$\Omega(\alpha_1, \alpha_2, \alpha_3) = \frac{1}{2} (w_2 + 2w_3) .$$

□

PROPOSITION 7 *Given an OWA function associated with the binomial decomposition coefficients α_j, with $j = 1, \ldots, n$, the* entropy *of an OWA function can be written as*

$$S(\boldsymbol{\alpha}) = -\sum_{i=1}^{n} \Big[\Big(\sum_{j=1}^{n-i+1} \frac{\binom{n-i}{j-1}}{\binom{n}{j}} \alpha_j \Big) \, ln \Big(\sum_{j=1}^{n-i+1} \frac{\binom{n-i}{j-1}}{\binom{n}{j}} \alpha_j \Big) \Big] \tag{81}$$

where the coefficients α_j, with $j = 1, \ldots, n$ are subject to conditions (72) and (73).

Proof. The expression of entropy (81) is obtained directly from (3) by substituting (78). In the case $n = 3$, for instance, we obtain

$$\begin{aligned} S(\alpha_1, \alpha_2, \alpha_3) \;\; = \;\; & -[(\frac{1}{3}\alpha_1 + \frac{2}{3}\alpha_2 + \alpha_3) ln(\frac{1}{3}\alpha_1 + \frac{2}{3}\alpha_2 + \alpha_3) + \\ & (\frac{1}{3}\alpha_1 + \frac{1}{3}\alpha_2) ln(\frac{1}{3}\alpha_1 + \frac{1}{3}\alpha_2) + \frac{1}{3}\alpha_1 ln(\frac{1}{3}\alpha_1)] . \end{aligned}$$

□

The constrained optimization problem described in (8)-(9) can be rewritten in terms of the coefficients α_j, with $j = 1, \ldots, n$, as follows

$$\max_{\boldsymbol{\alpha}} \;\; S(\boldsymbol{\alpha}) \tag{82}$$

subject to

$$\sum_{j=1}^{n} \alpha_j = 1 \, , \qquad \frac{1}{n-1} \sum_{i=1}^{n} \sum_{j=1}^{n-i+1} (i-1) \frac{\binom{n-i}{j-1}}{\binom{n}{j}} \alpha_j = \Omega \tag{83}$$

where $\Omega \in [0, 1]$ is a given degree of orness and the alpha coefficients are subject to conditions

$$\sum_{j=1}^{i} \frac{\binom{i-1}{j-1}}{\binom{n}{j}} \alpha_j \geq 0 \;\; for \;\; each \;\; i = 1, \ldots, n . \tag{84}$$

In analogy with what has been discussed earlier in relation with the constrained optimization problem (8)-(9), the two extreme orness cases $\Omega = 0,\, 1$ are associated with unique alpha coefficient vectors: $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \ldots, \alpha_{n-1}, \alpha_n) = (0, 0, \ldots, 0, 1)$ corresponding to $\boldsymbol{w} = (w_1, w_2, \ldots, w_{n-1}, w_n) = (1, 0, \ldots, 0, 0)$ in the case $\Omega = 0$, and a more complex $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \ldots, \alpha_{n-1}, \alpha_n)$, solution of the linear system (79), corresponding to $\boldsymbol{w} = (w_1, w_2, \ldots, w_{n-1}, w_n) = (0, 0, \ldots, 0, 1)$ in the case $\Omega = 1$. In both cases the value of entropy $S(\boldsymbol{\alpha})$ is zero.

On the other hand, if orness is in the open interval $\Omega \in (0, 1)$, it has been shown in Proposition 1 that the optimal weights are always positive, which means that the constraints (84) are strict. In other words, the constrained optimization problem (82)-(83) is a Lagrange problem with two equality constraints within an open domain in the space of vectors $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \ldots, \alpha_{n-1}, \alpha_n)$ corresponding to $\boldsymbol{w} \in (0, 1)^n$.

The Lagrange function of the constrained optimization problem (82)-(83) can be written as

$$\begin{aligned} L(\boldsymbol{\alpha}, \lambda_1, \lambda_2) \; = \; & -\sum_{i=1}^{n} \Big[\Big(\sum_{j=1}^{n-i+1} \frac{\binom{n-i}{j-1}}{\binom{n}{j}}\, \alpha_j \Big)\, ln \Big(\sum_{j=1}^{n-i+1} \frac{\binom{n-i}{j-1}}{\binom{n}{j}}\, \alpha_j \Big) \Big] \; + \\ & + \lambda_1 \Big[\sum_{j=1}^{n} \alpha_j - 1 \Big] + \lambda_2 \Big[\frac{1}{n-1} \sum_{i=1}^{n} \sum_{j=1}^{n-i+1} (i-1)\, \frac{\binom{n-i}{j-1}}{\binom{n}{j}}\, \alpha_j - \Omega \Big] \end{aligned} \tag{85}$$

where λ_1 and λ_2 are Lagrange multipliers.
Using the Lagrange function (85), the Lagrange conditions can be summarized with respect to $\alpha_j, \lambda_1,\, \lambda_2$, with $i, j = 1, \ldots, n$, as

1. The partial derivatives of the Lagrange function with respect to the free variables are equal to zero

$$\frac{\partial L}{\partial \alpha_j} = 0 \qquad j = 1, \ldots, n. \tag{86}$$

2. The partial derivatives of the Lagrange function with respect to the Lagrange multipliers corresponding to equality constraints are equal to zero

$$\frac{\partial L}{\partial \lambda_1} = 0\,, \quad \frac{\partial L}{\partial \lambda_2} = 0. \tag{87}$$

Given a predefined degree of orness Ω the solution of the constrained optimization problem in (82) with respect to the constraints (83)and (84) will give us the optimal values of the coefficients α_j, with $j = 1, \ldots, n$ which maximize the entropy.

We will now examine the numerical solutions of the constrained optimization problem related to the maximum entropy method in dimension $n = 3$. More specifically, we describe the numerical results of the maximum entropy optimization problem, firstly in terms the OWA weights w_i, $i = 1, \ldots, n$, and secondly in terms of the binomial decomposition coefficients α_j, $j = 1, \ldots, n$ for the cases $n = 3, 5, 7, 9$. To do that, since we are dealing with non-linear constrained optimization problems, we use the non-linear programming solver $fmincon$ in the MATLAB environment.

Author's Accepted Manuscript Intl. J. Intelligent Systems DOI http://dx.doi.org/10.1002/int.22083

4.1. Maximum entropy method in the case $n = 3$: weights

In this subsection we consider the maximum entropy method in the case $n = 3$. The constrained optimization problem described in (8)-(9) for $n = 3$ reduces to

$$\max_{\boldsymbol{w}} S(\boldsymbol{w}) = -\big(w_1\, lnw_1 + w_2\, lnw_2 + w_3\, lnw_3\big) \tag{88}$$

subject to

$$w_1 + w_2 + w_3 = 1\,, \qquad \frac{1}{2}\big(w_2 + 2w_3\big) \;=\; \Omega \tag{89}$$

where $\Omega \in (0, 1)$ is a given degree of orness. We know from Proposition 1 that the maximum entropy weights satisfy the positivity constraints

$$w_1 > 0, \quad w_2 > 0, \quad w_3 > 0\,. \tag{90}$$

To solve the above nonlinear constrained optimization problem we use the Lagrange multiplier method. The Lagrange function of problem (88)-(89) can be written as

$$\begin{aligned} L(\boldsymbol{w}, \lambda_1, \lambda_2) \;&=\; -\big(w_1\, ln\, w_1 + w_2\, ln\, w_2 + w_3\, ln\, w_3\big) + \lambda_1\big(w_1 + w_2 + w_3 - 1\big) + \\ &\quad +\lambda_2\Big[\frac{1}{2}(w_2 + 2w_3) - \Omega\Big] \end{aligned} \tag{91}$$

where λ_1, λ_2 are the Lagrange multipliers. Using the Lagrange function (91) we need to solve the following system of equations

1. The partial derivatives of the Lagrange function with respect to the weight variables are equal to zero,

$$\frac{\partial L}{\partial w_1} = 0, \quad \frac{\partial L}{\partial w_2} = 0, \quad \frac{\partial L}{\partial w_3} = 0. \tag{92}$$

   Using (91) we obtain

$$\begin{cases} -ln\, w_1 - 1 + \lambda_1 = 0 \\ -ln\, w_2 - 1 + \lambda_1 + \dfrac{1}{2}\,\lambda_2 = 0 \\ -ln\, w_3 - 1 + \lambda_1 + \lambda_2 = 0 \end{cases} \tag{93}$$

2. The partial derivatives of the Lagrange function with respect to the Lagrange multipliers corresponding to equality constraints are equal to zero,

$$\frac{\partial L}{\partial \lambda_1} = 0, \quad \frac{\partial L}{\partial \lambda_2} = 0. \tag{94}$$

   Using (91) we obtain

$$w_1 + w_2 + w_3 - 1 = 0\,, \quad \frac{1}{2}(w_2 + 2w_3) - \Omega = 0. \tag{95}$$

Author's Accepted Manuscript Intl. J. Intelligent Systems DOI http://dx.doi.org/10.1002/int.22083

Using equations (95) we can express the weights w_1, w_2 and w_3 in the following way

$$w_1 = w\,, \quad w_2 = 2 - 2\Omega - 2w_1\,, \quad w_3 = 2\Omega + w_1 - 1\,. \tag{96}$$

Using expressions (96) the partial derivatives of the Lagrange function with respect to the weight variables (93) can be rewritten as

$$\begin{cases} ln\, w + 1 - \lambda_1 = 0 \\ ln\,(2 - 2\Omega - 2w) + 1 - \lambda_1 - \dfrac{1}{2}\lambda_2 = 0 \\ ln\,(2\Omega + w - 1) + 1 - \lambda_1 - \lambda_2 = 0 \end{cases} \tag{97}$$

We exponenciate

$$\begin{cases} w = e^{\lambda_1 - 1} \\ 2 - 2\Omega - 2w = e^{\lambda_1 + \frac{1}{2}\lambda_2 - 1} \\ 2\Omega + w - 1 = e^{\lambda_1 + \lambda_2 - 1} \end{cases} \tag{98}$$

and by substituting the first of these equations into the second and third, we obtain

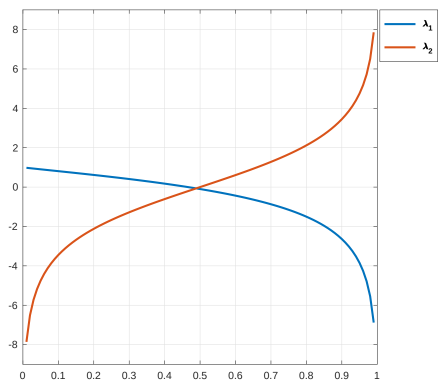
$$2 - 2e^{\lambda_1 - 1} - e^{\lambda_1 + \frac{1}{2}\lambda_2 - 1} \;=\; 2\Omega \;=\; 1 + e^{\lambda_1 + \lambda_2 - 1} - e^{\lambda_1 - 1}\,. \tag{99}$$

In Fig. 2 we represent the values of the Lagrange multipliers λ_1, λ_2 with respect to the optimal solutions of the constrained optimization problem in (88)-(89) with respect to the orness values $\Omega \in [0, 1]$.

In the table in Fig. 2 we present the values of the Lagrange multipliers λ_1, λ_2 obtained from equations (113) for orness values $\Omega \in [0, 1]$ in the case $n = 3$.

Orness	λ_1	λ_2
0.0	1	$-\infty$
0.1	0.809198	-3.44751
0.2	0.617079	-2.11974
0.3	0.409359	-1.28034
0.4	0.175311	-0.609237
0.5	-0.0986123	0
0.6	-0.433925	0.609237
0.7	-0.870983	1.28034
0.8	-1.50266	2.11974
0.9	-2.63832	3.44751
1.0	$-\infty$	∞

(a)

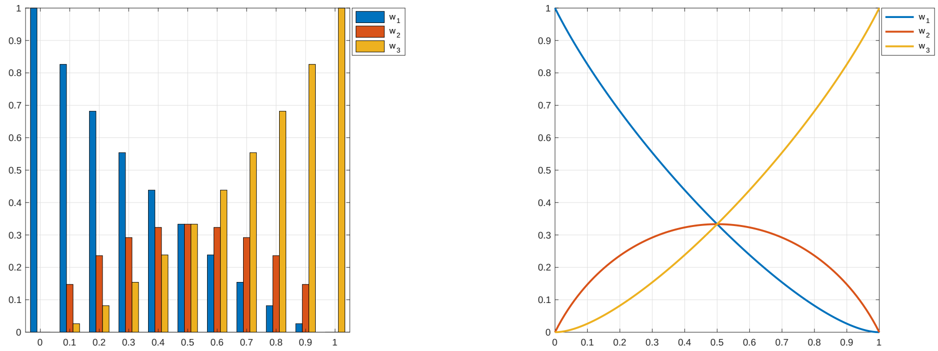


(b)

Figure 2: The Lagrange multipliers λ_1, λ_2 for orness $\Omega \in [0, 1]$, with n=3.

We can see that as the value of the orness varies from zero to one the Lagrange multiplier λ_1 is decreasing while λ_2 is increasing. Moreover we can observe that the Lagrange multiplier λ_2 is equal to zero when orness is equal to 0.5.

In Fig. 3 we can see the optimal values of the OWA weights that maximize entropy subject to the constraints (90) associated with the orness value $\Omega \in [0, 1]$, in the case $n = 3$.



(a) (b)

Figure 3: The optimal values of the OWA weights for orness $\Omega \in [0, 1]$, with n=3.

We observe that w_1 is decreasing, whereas w_3 is increasing with respect to $\Omega \in [0, 1]$. On the other hand, w_2 is increasing in the first half of the unit interval and decreasing thereafter. We see that when orness is equal to 0.5 the weights w_1, w_2 and w_3 are equal to 1/3, which corresponds to the weighting structure of the arithmetic mean in the case $n = 3$. Our results agree with those discussed by Filev and Yager in [17]. As expected, for the extreme orness values $\Omega = 0, 1$ the associated weights are uniquely defined as ($w_1 = 1$, $w_2 = 0$, $w_3 = 0$) and ($w_1 = 0$, $w_2 = 0$, $w_3 = 1$), respectively. In both cases the entropy value is null.

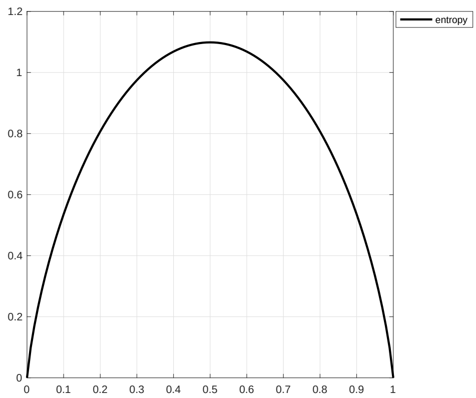


Figure 4: The optimal values of the objective function for the orness $\Omega \in [0, 1]$, with n=3.

In Fig. 4 we illustrate the values of the entropy function with respect to the optimal solutions of the constrained optimization problem in (88)-(89) with respect to the orness values $\Omega \in [0,1]$. We observe that entropy is increasing as the orness increases from zero to 0.5, reaching its maximum value $\ln 3$ when the orness is equal to 0.5. For orness values greater than 0.5 the entropy is decreasing. The minimum value of entropy is obtained for the extreme values of orness $\Omega = 0,\, 1$.

4.2. Maximum entropy method in the case $n = 3$: alphas

In this subsection we consider the maximum entropy method in the context of the binomial decomposition in the case $n = 3$. The constrained optimization problem (82)- (83) for $n = 3$ reduces to

$$\max_{\boldsymbol{\alpha}} S(\boldsymbol{\alpha}) = -\big[\big(\frac{1}{3}\alpha_1 + \frac{2}{3}\alpha_2 + \alpha_3\big) ln\big(\frac{1}{3}\alpha_1 + \frac{2}{3}\alpha_2 + \alpha_3\big) + \\ + \big(\frac{1}{3}\alpha_1 + \frac{1}{3}\alpha_2\big) ln\big(\frac{1}{3}\alpha_1 + \frac{1}{3}\alpha_2\big) + \frac{1}{3}\alpha_1 ln\big(\frac{1}{3}\alpha_1\big)\big] \tag{100}$$

subject to

$$\alpha_1 + \alpha_2 + \alpha_3 = 1\,, \qquad \frac{1}{2}\alpha_1 + \frac{1}{6}\,\alpha_2 = \Omega \tag{101}$$

where $\Omega \in (0,1)$ is a given degree of orness. We know from Proposition 1 and (90) that the maximum entropy alpha coefficients satisfy the positivity constraints

$$\frac{1}{3}\alpha_1 > 0, \quad \frac{1}{3}\alpha_1 + \frac{1}{3}\alpha_2 > 0, \quad \frac{1}{3}\alpha_1 + \frac{2}{3}\alpha_2 + \alpha_3 > 0\,. \tag{102}$$

To solve the above nonlinear constrained optimization problem we use the Lagrange multiplier method given that the solution has to satisfy the inequality constraints. The Lagrange function of problem (100)-(101) can be written as

$$\begin{aligned} L(\boldsymbol{\alpha},\lambda_1,\lambda_2) \;=\; & -\big[\big(\frac{1}{3}\alpha_1 + \frac{2}{3}\alpha_2 + \alpha_3\big) ln\big(\frac{1}{3}\alpha_1 + \frac{2}{3}\alpha_2 + \alpha_3\big) + \\ & + \big(\frac{1}{3}\alpha_1 + \frac{1}{3}\alpha_2\big) ln\big(\frac{1}{3}\alpha_1 + \frac{1}{3}\alpha_2\big) + \frac{1}{3}\alpha_1 ln\big(\frac{1}{3}\alpha_1\big)\big] + \\ & + \lambda_1\big(\alpha_1 + \alpha_2 + \alpha_3 - 1\big) + \lambda_2\big(\frac{1}{2}\alpha_1 + \frac{1}{6}\,\alpha_2 - \Omega\big) \end{aligned} \tag{103}$$

where λ_1, λ_2 are the Lagrange multipliers. Using the Lagrange function (103) we need to solve the following system of equations

1. The partial derivatives of the Lagrange function with respect to the free variables are equal to zero

$$\frac{\partial L}{\partial \alpha_1} = 0, \quad \frac{\partial L}{\partial \alpha_2} = 0, \quad \frac{\partial L}{\partial \alpha_3} = 0 \tag{104}$$

or equivalently,

$$\begin{cases} -\frac{1}{3}\Big[\ln(\frac{1}{3}\alpha_1+\frac{2}{3}\alpha_2+\alpha_3)+\ln(\frac{1}{3}\alpha_1+\frac{1}{3}\alpha_2)+\ln(\frac{1}{3}\alpha_1)\Big]-1+\lambda_1+\frac{1}{2}\lambda_2=0 \\ -\frac{1}{3}\Big[2\ln(\frac{1}{3}\alpha_1+\frac{2}{3}\alpha_2+\alpha_3)+\ln(\frac{1}{3}\alpha_1+\frac{1}{3}\alpha_2)\Big]-1+\lambda_1+\frac{1}{6}\lambda_2=0 \\ -\frac{1}{3}\Big[3\ln\big(\frac{1}{3}\alpha_1+\frac{2}{3}\alpha_2+\alpha_3\big)\Big]-1+\lambda_1=0\,. \end{cases} \tag{105}$$

2. The partial derivatives of the Lagrange function with respect to the Lagrange multipliers corresponding to equality constraints are equal to zero

$$\frac{\partial L}{\partial \lambda_1}=0, \quad \frac{\partial L}{\partial \lambda_2}=0 \tag{106}$$

or equivalently,

$$\alpha_1+\alpha_2+\alpha_3-1=0\,, \quad 3\alpha_1+\alpha_2-6\Omega=0 \tag{107}$$

Using expressions (107) we can write the coefficients α_2 and α_3 in terms of the coefficient α_1 in the following way

$$\alpha_2=6\Omega-3\alpha_1\,, \quad \alpha_3=1-6\Omega+2\alpha_1\,. \tag{108}$$

We introduce the simplified notation $\alpha_1=\alpha$. Using expressions (108) the partial derivatives of the Lagrange function with respect to the free variables (105) can be rewritten as

$$\begin{cases} -\frac{1}{3}\big[\ln(1-2\Omega+\frac{1}{3}\alpha)+\ln(2\Omega-\frac{2}{3}\alpha)+\ln(\frac{1}{3}\alpha)\big]-1+\lambda_1+\frac{1}{2}\lambda_2=0 \\ -\frac{1}{3}\big[2\ln(1-2\Omega+\frac{1}{3}\alpha)+\ln(2\Omega-\frac{2}{3}\alpha)\big]-1+\lambda_1+\frac{1}{6}\lambda_2=0 \\ -\frac{1}{3}\big[3\ln(1-2\Omega+\frac{1}{3}\alpha)\big]-1+\lambda_1=0 \end{cases} \tag{109}$$

Applying Gauss-Jordan elimination we obtain

$$\begin{cases} \ln(2\Omega-\frac{2}{3}\alpha)+\ln(\frac{1}{3}\alpha)+2-2\lambda_1-\frac{3}{2}\lambda_2=0 \\ \ln(2\Omega-\frac{2}{3}\alpha)+1-\lambda_1-\frac{1}{2}\lambda_2=0 \\ \ln(1-2\Omega+\frac{1}{3}\alpha)+1-\lambda_1=0\,. \end{cases} \tag{110}$$

We apply again Gauss-Jordan elimination

$$\begin{cases} \ln(\frac{1}{3}\alpha)+1-\lambda_1-\lambda_2=0 \\ \ln(2\Omega-\frac{2}{3}\alpha)+1-\lambda_1-\frac{1}{2}\lambda_2=0 \\ \ln(1-2\Omega+\frac{1}{3}\alpha)+1-\lambda_1=0\,. \end{cases} \tag{111}$$

Author's Accepted Manuscript Intl. J. Intelligent Systems DOI http://dx.doi.org/10.1002/int.22083

We exponenciate

$$\begin{cases} \alpha = 3e^{\lambda_1+\lambda_2-1} \\ 2\Omega - \dfrac{2}{3}\alpha = e^{\lambda_1+\frac{1}{2}\lambda_2-1} \\ 1 - 2\Omega + \dfrac{1}{3}\alpha = e^{\lambda_1-1} \end{cases} \tag{112}$$

and by substituting the first of these equations into the second and third, we obtain

$$2e^{\lambda_1+\lambda_2-1} + e^{\lambda_1+\frac{1}{2}\lambda_2-1} \;=\; 2\Omega \;=\; 1 + e^{\lambda_1+\lambda_2-1} - e^{\lambda_1-1}\,. \tag{113}$$

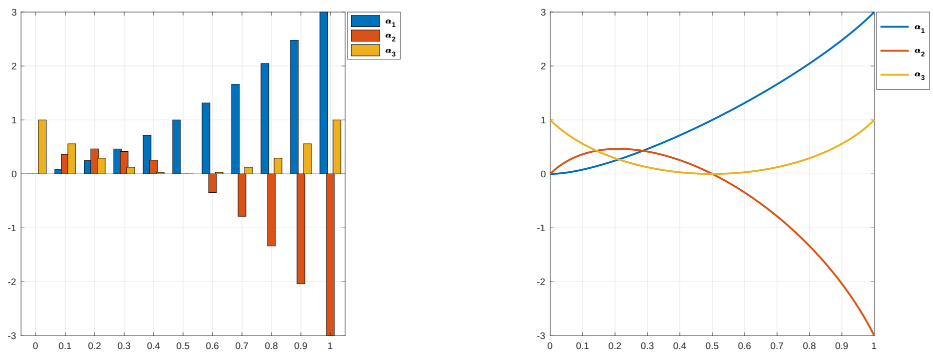
These equations are equivalent to the analogous equations (99). Using the latter, we obtain that

$$e^{\lambda_1+\lambda_2-1} = 1 - e^{\lambda_1-1} - e^{\lambda_1+\frac{1}{2}\lambda_2-1}\,. \tag{114}$$

In fact, using this result we can immediately transform the left hand side of (113) into the left hand side of (99).

The values of the Lagrange multipliers λ_1, λ_2 with respect to the optimal solutions of the constrained optimization problem in (100)-(101) are represented in Fig. 2 with respect to the orness values $\Omega \in [0,1]$.

In Fig. 5 we illustrate the optimal values of the coefficients α_1, α_2, α_3 which maximize entropy subject to the constraints (101) for $\Omega \in [0,1]$. As we can see, the coefficients α_1, α_2, α_3 follow different trends with respect to $\Omega \in [0,1]$. The coefficient α_1 is non negative and increasing across the unit interval, while the coefficient α_2 takes negative values for high orness $\Omega \in [0.5,1]$. The coefficient α_3 is also non negative with respect to $\Omega \in [0,1]$, in the first half of the unit interval it is decreasing, whereas in the second half of the unit interval it shows an increasing pattern. In the case $\Omega = 0$, we have $\boldsymbol{\alpha} = (\alpha_1,\alpha_2,\alpha_3) = (0,0,1)$, whereas in the case $\Omega = 1$, we have a more complex vector of coefficients, $\boldsymbol{\alpha} = (\alpha_1,\alpha_2,\alpha_3) = (3,-3,1)$.



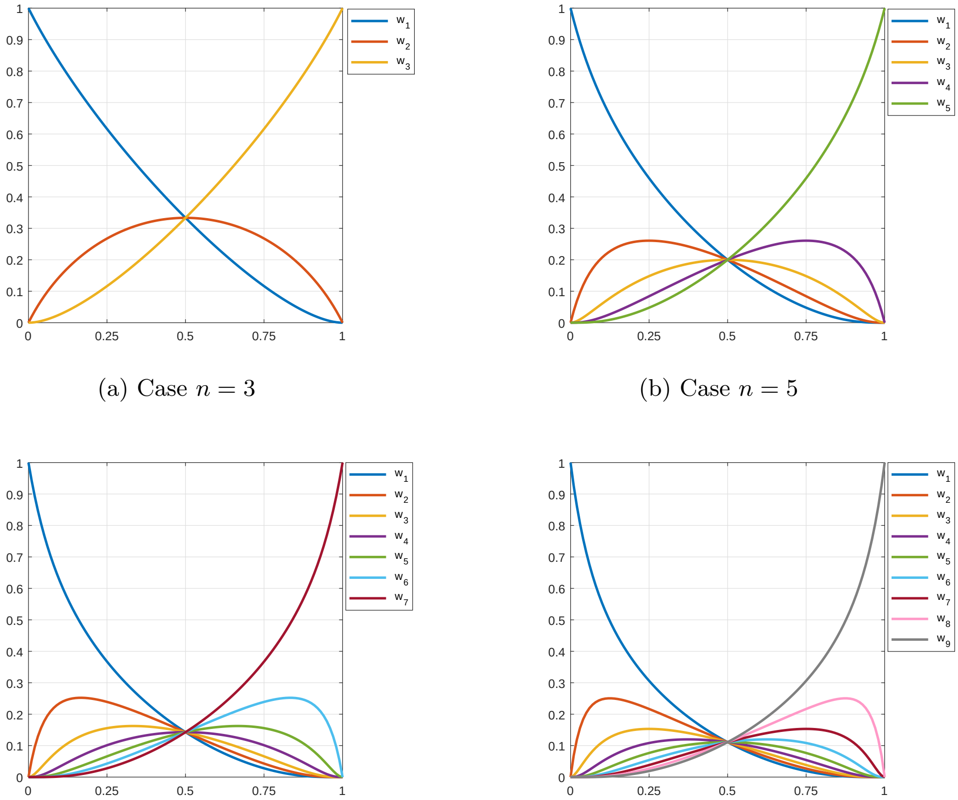
(a) (b)

Figure 5: The optimal values of the coefficients α_1, α_2, α_3 for the orness $\Omega \in [0,1]$, with n=3.

The values of the entropy function with respect to the optimal solutions of the constrained optimization problem in (100)-(101) are represented in Fig. 4 with respect to the orness values $\Omega \in [0,1]$. The optimal values of the entropy S and the Lagrange multipliers λ_1, λ_2 are, as expected, the same in the constrained optimization problems (88)-(89) and (100)-(101).

4.3. The optimal weights for orness $\Omega \in [0,1]$ in the cases $n = 3, 5, 7, 9$

We consider the cases $n = 3, 5, 7, 9$ and we compute the optimal maximum entropy weights by solving the non-linear constrained optimization problem in (8) using MATLAB's function $fmincon$ for orness values $\Omega \in [0,1]$.



(a) Case $n = 3$ (b) Case $n = 5$

(c) Case $n = 7$ (d) Case $n = 9$

Figure 6: The optimal weights for orness $\Omega \in [0,1]$ in the cases $n = 3, 5, 7, 9$.

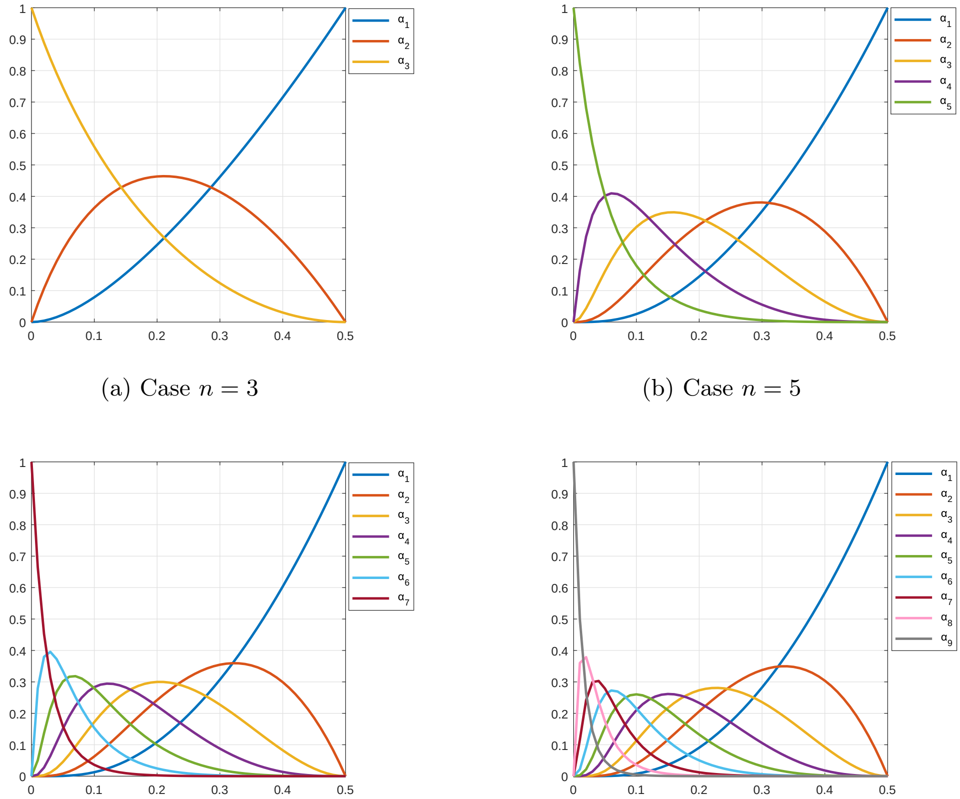
In Fig. 6 we illustrate the optimal weights for orness $\Omega \in [0,1]$ in the cases $n = 3, 5, 7, 9$. As we can see, for low orness values $\Omega \in [0, 0.5]$ the optimal weights are decreasing,whereas for high orness values $\Omega \in [0.5, 1]$ the optimal weights are increasing. As expected, for orness $\Omega = 0.5$ we have $w_i = 1/n$, for $i = 1, \ldots, n$, in all the cases $n = 3, 5, 7, 9$. Moreover, we observe that the optimal values of the first and

Author's Accepted Manuscript Intl. J. Intelligent Systems DOI http://dx.doi.org/10.1002/int.22083

last weights w_1, w_n have a wide range between zero and one, whereas on the other hand the values of the other weights do not. Finally, if for some given level of orness $\Omega \in [0,1]$ the optimal weighting vector is $\boldsymbol{w}_\Omega = (w_1, w_2, \ldots, w_n)$, then the optimal weighting vector for orness $1 - \Omega$ is its reverse, that is, $\boldsymbol{w}_{1-\Omega} = (w_n, w_{n-1} \ldots, w_1)$.

4.4. The optimal alphas for orness $\Omega \in [0,1]$ in the cases $n = 3, 5, 7, 9$

In this subsection we focus on the maximum entropy method in the context of the binomial decomposition. We consider the cases $n = 3, 5, 7, 9$ and we compute the optimal α_j with $j = 1, \ldots, n$ coefficients for orness values $\Omega \in [0,1]$. The binomial α_j with $j = 1, \ldots, n$ coefficients are determined in the following way: first we compute the optimal weights for orness $\Omega \in [0,1]$ in the cases $n = 3, 5, 7, 9$ using MATLAB's function *fmincon*. Once the weights are obtained, the coefficients α_j with $j = 1, \ldots, n$ are computed from the linear system from expression (79). Due to the fact that the behavior of the coefficients α_j with $j = 1, \ldots, n$ with respect to orness is very different for low/high orness values, we illustrate the optimal values of the coefficients α_j with $j = 1, \ldots, n$ in two parts, first for low orness values $\Omega \in [0, 0.5]$, and then for high orness values $\Omega \in [0.5, 1]$.



(a) Case $n = 3$

(b) Case $n = 5$

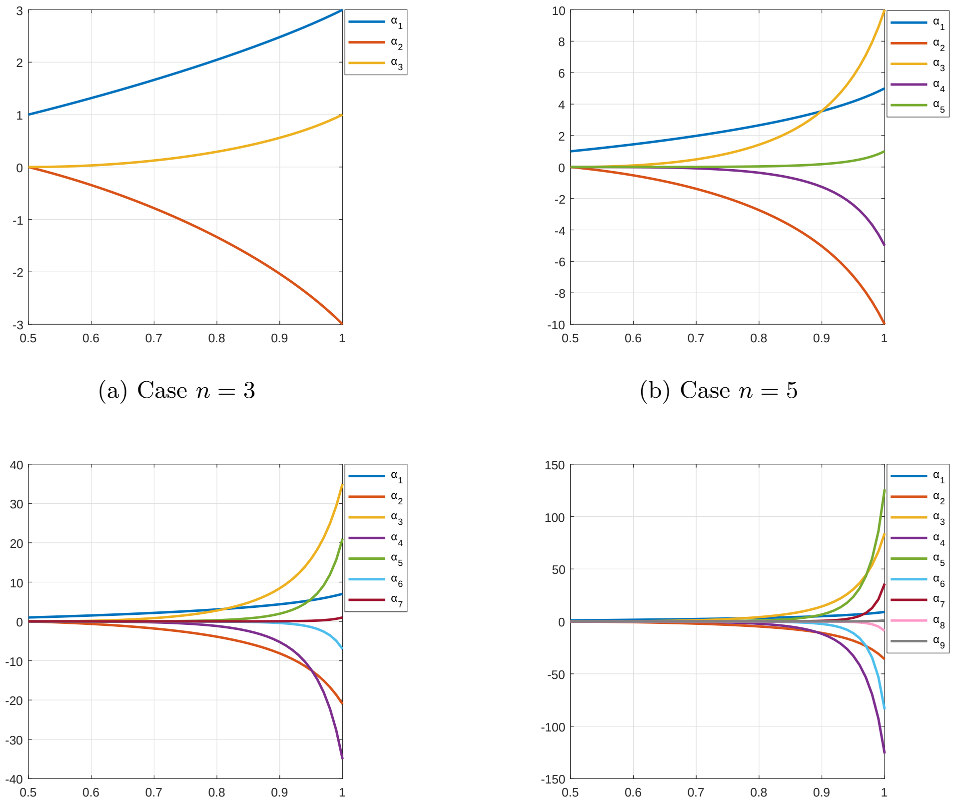
(c) Case $n = 7$

(d) Case $n = 9$

Figure 7: The optimal α_j, $j = 1, \ldots, n$, for orness $\Omega \in [0, 0.5]$ in the cases $n = 3, 5, 7, 9$.

In Fig. 7 we represent the optimal α_j coefficients, with $j = 1, \ldots, n$, for orness values $\Omega \in [0, 0.5]$ in the cases $n = 3, 5, 7, 9$. For orness values $\Omega \in [0, 0.5]$ we observe the following: The optimal coefficients α_j with $j = 1, \ldots, n$ are all non negative and take values in the unit interval, independent of the order of n. The coefficient α_1 is increasing whereas the coefficient α_n is decreasing. The intermediate coefficients $\alpha_2, \ldots, \alpha_{n-1}$ initially increase (with the higher α_j to increase faster) and then decrease as orness increases. Moreover, as the value of orness Ω increases (towards to 0.5), the number of the higher coefficients α_j that become zero or are very close to zero is increasing. Finaly, as the order of n increases, the values of the higher coefficients α_j go faster to zero (as orness increases).

Author's Accepted Manuscript Intl. J. Intelligent Systems DOI http://dx.doi.org/10.1002/int.22083



(a) Case $n = 3$

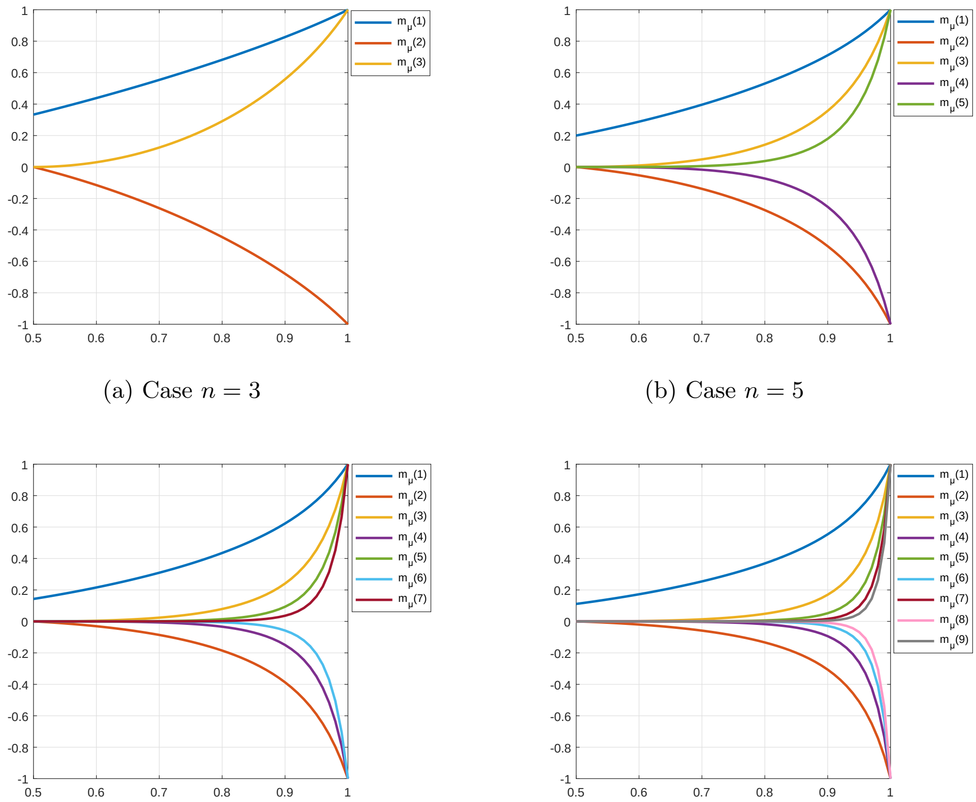
(b) Case $n = 5$

(c) Case $n = 7$

(d) Case $n = 9$

Figure 8: The optimal α_j, $j = 1, \ldots, n$, for orness $\Omega \in [0.5, 1]$ in the cases $n = 3, 5, 7, 9$.

In Fig. 8 we represent the optimal α_j coefficients, with $j = 1, \ldots, n$, for orness values $\Omega \in [0.5, 1]$ in the cases $n = 3, 5, 7, 9$. It is clear that the range of the optimal α_j values depends on the order of n for orness values $\Omega \in [0.5, 1]$. In order to have them on the same scale for the cases $n = 3, 5, 7, 9$ we examine the optimal Möbius transforms $m_\mu(j) = \alpha_j / \binom{n}{j}$, $j = 1, \ldots, n$ for orness values $\Omega \in [0.5, 1]$.



(a) Case $n = 3$ (b) Case $n = 5$

(c) Case $n = 7$ (d) Case $n = 9$

Figure 9: The optimal Möbius transforms $m_\mu(j) = \alpha_j / \binom{n}{j}$, $j = 1, \ldots, n$, for orness $\Omega \in [0.5, 1]$ in the cases $n = 3, 5, 7, 9$.

In Fig. 9 we illustrate the optimal Möbius transforms $m_\mu(j) = \alpha_j / \binom{n}{j}$, $j = 1, \ldots, n$, for orness $\Omega \in [0.5, 1]$ in the cases $n = 3, 5, 7, 9$. We notice that for orness $\Omega = 0.5$ we have $m_\mu(1) = 1$ and $m_\mu(k) = 0$ with $k = 2, \ldots, n$. Moreover, the optimal Möbius transforms $m_\mu(j)$, with j odd, are increasing and take non negative values, whereas the optimal Möbius transforms $m_\mu(j)$, with j even, are decreasing and take non positive values (as orness increases). Finally, as the order of n increases the number of the higher optimal Möbius transforms $m_\mu(j)$ that are zero increases too for orness values Ω close to 0.5.

5. CONCLUSIONS

In the context of multiple criteria decision making we have examined the maximum entropy constrained optimization problem associated with Ordered Weighted Averaging. OWA functions are of the form $A(\boldsymbol{x}) = \sum_{i=1}^{n} w_i \, x_{(i)}$, with $\sum_{i=1}^{n} w_i = 1$. We have discussed two measures which characterize OWA functions, orness and entropy. Orness is the degree to which an OWA function resembles the maximum operator, whereas entropy is a measure of weight dispersion.

We have reviewd the maximum entropy approach, a method to determine appropriate OWA weights with minimum dispersion. The maximum entropy method is based on the solution of a constrained optimization problem, maximize entropy $S(\boldsymbol{w})$ given a specified degree of orness Ω. We have briefly reviewed the analytic solution of the maximum entropy method proposed by Filev and Yager in 1995 [17], and later by Fullér and Majlender in 2001 [21].

In this context, we have discussed the numerical results concerning the constrained optimization problem of the maximum entropy method, initially in terms of the OWA weights w_i, with $i = 1, \ldots, n$, and then in terms of the binomial decomposition coefficients α_j, with $j = 1, \ldots, n$, in dimensions $n = 3, 5, 7, 9$.

Next we have briefly reviewed the framework of the binomial decomposition, which allows for an alternative parametric description of the OWA functions. The values of the binomial coefficients α_j, $j = 1, \ldots, n$ are uniquely determined by the weighting structure of the OWA function. Consequently, the weights w_i, $i = 1, \ldots, n$ of the OWA function can be written explicitly in terms of the binomial coefficients α_j, $j = 1, \ldots, n$.

We have observed that for low orness values $\Omega \in [0, 0.5]$ the optimal weights are decreasing, whereas for high orness values $\Omega \in [0.5, 1]$ the optimal weights are increasing. Moreover, we have noticed that the optimal values of the first and last weights w_1, w_n have a wide range between zero and one, while on the other hand the values of the other weights do not.

Regarding the optimal α_j coefficients, we have found that their behavior with respect to orness values $\Omega \in [0, 1]$ is very different for low/high orness values. For this reason we have illustrated graphically the optimal α_j coefficients in two parts, first for low orness values $\Omega \in [0, 0.5]$ and then for high orness values $\Omega \in [0.5, 1]$. We have observed that the optimal α_j, with $j = 1, \ldots, n$ for low orness values $\Omega \in [0, 0.5]$ are all non negative and take values in the unit interval, independently of the dimension n. On the contrary, the optimal values of the α_j coefficients for high orness values $\Omega \in [0.5, 1]$ depend strongly on the dimension n, both in the complexity of their distribution and in the amplitude of their scale.

References

[1] Ahn BS. Compatible weighting method with rank order centroid: Maximum entropy ordered weighted averaging approach. European Journal of Operational Research 2011;212(3):552–559.

[2] Beliakov G, Pradera A, Calvo T. Aggregation Functions: A Guide for Practitioners. Studies in Fuzziness and Soft Computing, Vol. 221. Heidelberg: Springer; 2007.

[3] Beliakov G, Bustince Sola H, Calvo T. A Practical Guide to Averaging Functions. Studies in Fuzziness and Soft Computing, Vol. 329. Heidelberg: Springer; 2016.

[4] Bortot S, Marques Pereira RA. The binomial Gini inequality indices and the binomial decomposition of welfare functions. Fuzzy Sets and Systems 2014;255:92-114.

[5] Bortot S, Marques Pereira RA, Nguyen TH. The binomial decomposition of OWA functions, the 2-additive and 3-additive cases in n dimensions. International Journal of Intelligent Systems 2018;33:187-212.

[6] Bortot S, Fedrizzi M, Marques Pereira RA, Nguyen TH. The binomial decomposition of generalized Gini welfare functions, the S-Gini and Lorenzen cases. Information Sciences 2018;460-461:555-577.

[7] Brunelli M, Fedrizzi M: A general formulation for some inconsistency indices of pairwise comparisons. Annals of Operations Research (online since 14 June 2018). DOI: 10.1007/s10479-018-2936-6.

[8] Calvo T, De Baets B. Aggregation operators defined by k-order additive/maxitive fuzzy measures. International Journal of Uncertainty, Fuzzyness and Knowledge-Based Systems 1998;6(6):533–550.

[9] T. Calvo, A. Kolesárova, M. Komorníková, R. Mesiar. Aggregation operators: Properties, classes and construction methods. In: Calvo T, Mayor G, Mesiar R, editors. Aggregation Operators: New Trends and Applications. Heidelberg: Physica-Verlag; 2002. pp. 3-104.

[10] Carlsson C, Fullér R. Maximal Entropy and Minimal Variability OWA Operator Weights: A Short Survey of Recent Developments. In: Collan M, Kacprzyk J, editors. Soft Computing Applications for Group Decision-making, Consensus Modeling. Studies in Fuzziness and Soft Computing, Vol. 357. Heidelberg: Springer; 2018. pp. 187–199.

[11] Chang YC, Chang KH, Liaw CS.Innovative reliability allocation using the maximal entropy ordered weighted averaging method. Computers & Industrial Engineering 2009;57(4):1274–1281.

[12] Chateauneuf A, Jaffray JY. Some characterizations of lower probabilities and other monotone capacities throught the use of Möbius inversion. Mathematical Social Sciences 1989;17(3):263–283.

[13] Cheng CH, Chang JR. MCDM aggregation model using situational ME-OWA and ME-OWGA operators. International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems 2006;14(4):421–443.

[14] Choquet G. Theory of capacities. Annales de l'Institut Fourier 1953;5:131-295.

[15] Chuu SJ. An investment evaluation of supply chain RFID technologies: A group decision-making model with multiple information source. Knowledge-Based Systems 2014;66:210–220.

[16] Denneberg D. Non-additive measure and integral. Dordrecht: Kluwer Academic Publishers; 1994.

[17] Filev D, Yager RR. Analytic properties of maximum entropy OWA operators. Information Sciences 1995;85:11–27.

[18] Filev D, Yager RR. On the issue of obtaining OWA operator weights. Fuzzy Sets and Systems 1998;94(2):157–169.

[19] Fodor J, Marichal JL, Roubens M. Characterization of the ordered weighted averaging operators. IEEE Trans. on Fuzzy Systems 1995;3(2):236–240.

[20] Fodor J, Roubens M. Fuzzy Preference Modelling and Multicriteria Decision Support. Dordrecht: Kluwer Academic Publishers; 1994.

[21] Fullér R, Majlender P. An analytic approach for obtaining maximal entropy OWA operator weights. Fuzzy Sets and Systems 2001;124(1):53–57.

[22] Gajdos T. Measuring inequalities without linearity in envy: Choquet integrals for symmetric capacities. Journal of Economic Theory 2002;106(1):190–200.

[23] Grabisch M. Fuzzy integral in multicriteria decision making. Fuzzy Sets and Systems 1995;69(3):279-298.

[24] Grabisch M. The application of fuzzy integrals in multicriteria decision making. European Journal of Operational Research 1996;89(3):445-456.

[25] Grabisch M. k-order additive discrete fuzzy measures and their representation. Fuzzy Sets and Systems 1997;92(2):167-189.

[26] Grabisch M. Alternative representations of discrete fuzzy measures for decision making. International Journal of Uncertainty, Fuzzyness and Knowledge-Based Systems 1997;5(5):587-607.

[27] Grabisch M. k-Additive measures: recent issues and challenges. In: Proc. 5th International Conference on Soft Computing and Information/Intelligent Systems, Izuka, Japan; 1998. pp.394–397.

[28] Grabisch M, Kojadinovich I, Meyer P. A review of methods for capacity identification in Choquet integral based multi-attribute utility theory: Applications of the Kappalab R package. European Journal of Operational Research 2008;186(2):766-785.

[29] Grabisch M, Labreuche C. Fuzzy measures and integrals in MCDA. In: Figueira J, Greco S, Ehrgott M, editors. Multiple Criteria Decision Analysis. Heidelberg: Springer; 2005. pp. 563–604.

[30] Grabisch M, Labreuche C. A decade of application of the Choquet and Sugeno integrals in multi-criteria decision aid. 4OR 2008;6(1):1-44.

[31] Grabisch M, Labreuche C. A decade of application of the Choquet and Sugeno integrals in multi-criteria decision aid. Annals of Operations Research 2010;175(1):247-286.

[32] Grabisch M, Marichal JL, Mesiar R, Pap E. Aggregation Functions. Encyclopedia of Mathematics and its Applications, Vol. 127. Cambridge: Cambridge University Press; 2009.

[33] Grabisch M, Marichal JL, Mesiar R, Pap E. Aggregation functions: means. Information Sciences 2011;181(1):1–22.

[34] He YL, Liu JNK, Hu YX, Wang XZ. OWA operator based link prediction ensemble for social network. Expert Systems with Applications 2015;42(1):21–50.

[35] Kang B, Deng Y, Hewage K, Sadiq R. Generating Z-number based on OWA weights using maximum entropy. International Journal of Intelligent Systems 2018;33:1745–1755.

[36] Kim, EY, Ahn BS. Implicit Elicitation of Attitudinal Character in the OWA Operator. International Journal of Intelligent Systems 2018;33:281–287.

[37] Liaw CS, Chang YC, Chang KH, Chang TY. ME-OWA based DEMATEL reliability apportionment method. Expert Systems with Applications 2011;38(8):9713–9723.

[38] Liu X. A review of the OWA determination methods: Classifcation and some extensions. In: Yager RR, Kacprzyk J, Beliakov G, editors. Recent Developments in the Ordered Weighted Averaging Operators: Theory and Practice. Studies in Fuzziness and Soft Computing, Vol. 265. Heidelberg: Springer; 2011. pp. 49–90.

[39] Liu X, Chen L. On the properties of parametric geometric OWA operator. International Journal of Approximate Reasoning 2004;35(2):163–178.

[40] Marichal JL. Aggregation operators for multicriteria decision aid. Ph.D. Thesis. University of Liège; 1998.

[41] Mayag B, Grabisch M, Labreuche C. A representation of preferences by the Choquet integral with respect to a 2-additive capacity. Theory and Decision 2011;71(3):297-324.

[42] Mayag B, Grabisch M, Labreuche C. A characterization of the 2-additive Choquet integral through cardinal information. Fuzzy Sets and Systems 2011;184(1):84-105.

[43] Mesiar R, Kolesárová A, Calvo T, Komorníková M. A Review of Aggregation Functions. In: Bustince H, Herrera F, Montero J, editors. Fuzzy Sets and Their Extensions: Representation, Aggregation and Models. Studies in Fuzziness and Soft Computing, Vol. 220. Heidelberg: Springer; 2008. pp. 121–144.

[44] Miranda P, Grabisch M. Optimization issues for fuzzy measures. International Journal of Uncertainty, Fuzzyness and Knowledge-Based Systems 1999;7(6):545–560.

[45] Miranda P, Grabisch M, Gil P. Axiomatic structure of k-additive capacities. Mathematical Social Sciences 2005;49(2):153-178.

Author's Accepted Manuscript Intl. J. Intelligent Systems DOI http://dx.doi.org/10.1002/int.22083

[46] Murofushi T, Sugeno M. Some quantities represented by the Choquet integral. Fuzzy Sets and Systems 1993;2(56):229–235.

[47] O'Hagan M. Fuzzy decision aids. In: Proc. 21st Asilomar Conference on Signals, Systems and Computers, IEEE and Maple Press, Pacific Grove CA; 1987. pp.624–628.

[48] O'Hagan M. Aggregating template or rule antecedents in real-time expert systems with fuzzy set. In: Proc. 22nd Asilomar Conference on Signals, Systems and Computers, IEEE and Maple Press, Pacific Grove CA; 1988. pp.681–689.

[49] Sugeno M. Theory of fuzzy integrals and its applications. Ph.D. Thesis. Tokyo Institut of Technology; 1974.

[50] Torra V, Narukawa Y. Modeling Decisions: Information Fusion and Aggregation Operators. Heidelberg: Springer; 2007.

[51] Wu J, Sun BL, Liang CY, Yang SL. A linear programming model for determining ordered weighted averaging operator weights with maximal Yager's entropy. Computers & Industrial Engineering 2009;57(3):742–747.

[52] Xu Z. An overview of methods for determining OWA weights. International Journal of Intelligent Systems 2005;20(8):843–865.

[53] Yager RR. On ordered weighted averaging aggregation operators in multicriteria decision making. IEEE Trans. on Systems, Man and Cybernetics 1988;18(1):183–190.

[54] Yager RR. Weighted Maximum Entropy OWA Aggregation With Applications to Decision Making Under Risk. IEEE Trans. on Systems, Man and Cybernetics 2009;39(3):555–564.

[55] Yager RR, Kacprzyk J, editors. The Ordered Weighted Averaging Operators. Theory and Applications. Dordrecht: Kluwer Academic Publisher; 1997.

[56] Yager RR, Kacprzyk J, Beliakov J, editors. Recent Developments in the Ordered Weighted Averaging Operators: Theory and Practice. Studies in Fuzziness and Soft Computing, Vol. 265. Heidelberg: Springer; 2011.

[57] Yusoff B, Merigó-Lindahl JM. Analytical Hierarchy Process under Group Decision Making with Some Induced Aggregation Operators. In: Laurent A, Strauss O, Bouchon-Meunier B, Yager RR, editors. Information Processing and Management of Uncertainty in Knowledge-Based Systems (IPMU 2014). CCIS, Vol.442. Heidelberg: Springer; 2014. pp. 476–485.