Reduction of variables and constraints in fitting antibuoyant fuzzy measures to data using linear programming

Gleb Beliakov[a], Marek Gagolewski[a,b], Simon James[a,1]

[a] *Deakin University, School of Information Technology, Geelong, VIC 3220, Australia*
[b] *Warsaw University of Technology, Faculty of Mathematics and Information Science, ul. Koszykowa 75, 00-662 Warsaw, Poland*

Abstract

The discrete Choquet integral with respect to various types of fuzzy measures serves as an important aggregation function which accounts for mutual dependencies between the inputs. The Choquet integral can be used as an objective (or constraint) in optimisation problems, and the type of fuzzy measure used determines its complexity. This paper examines the class of antibuoyant fuzzy measures, which restrict the supermodular (convex) measures and satisfy the Pigou–Dalton progressive transfers principle. We determine subsets of extreme points of the set of antibuoyant fuzzy measures, whose convex combinations form a basis of three proposed algorithms for random generation of fuzzy measures from that class, and also for fitting fuzzy measures to empirical data or solving best approximation problems. Potential applications of the proposed methods are envisaged in social welfare, ecology, and optimisation.

Keywords: fuzzy measures, Choquet integral, supermodularity, capacities, progressive transfers

---

1. Introduction

Fuzzy measures and integrals represent powerful tools for multiple criteria decision making, single and multiobjective optimisation and other areas in which explicit models of interaction between the variables or parameters is

*Corresponding author; Email: simon.james@deakin.edu.au

important [8, 18]. Fuzzy measure values reflect relative contributions of not just individual variables but their subsets (called coalitions). Their significant modelling capacity comes at the cost of exponentially many parameters and even more relations between those parameters in the form of linear constraints.

Prominent classes of fuzzy measures include sub- and super-modular measures, belief and plausibility measures, possibility and necessity measures, k-additive, p-symmetric, maxitive and minitive fuzzy measures and many alike. In this paper we focus on buoyant and antibuoyant fuzzy measures, which narrow down the classes of sub- and super-modular fuzzy measures [5, 6] and satisfy an important economic principle of regressive (progressive) transfers, also known as the Pigou–Dalton principle [14], see also [3]. An important consequence of this principle arises in mathematical optimisation: if an (anti)buoyant fuzzy measure is used to define an optimisation objective subject to linear constraints, the optimum is guaranteed to lie within one particular canonical simplex (out of $n!$ simplices) of the simplicial partition of the domain $[0, 1]^n$ [1, 5].

Construction and identification of various classes of fuzzy measures, in particular their random generation and/or fitting to the available data, is one problem arising from applications [4, 12, 16, 19]. For this it is important to find a suitable representation of fuzzy measures of a given class, so that the number of parameters and constraints is reduced. For the classes of (anti)buoyant fuzzy measures, as well as sub- and super-modular measures, the number of linear constraints is larger than those needed for simple monotonicity, which makes the task of learning such fuzzy measures from data not readily scalable.

The $\{0, 1\}$-fuzzy measures have been studied by Combarro et al. [13, 25] as defining vertices of the convex polytope of fuzzy measures and for the purposes of random generation of fuzzy measures in learning contexts. Some subsets, such as k-additive fuzzy measures ($k > 2$) do not have vertices coinciding with vertices for general fuzzy measures and the convex polytopes are much more complex than the larger fuzzy measure set. The $\{0, 1\}$-fuzzy measures can be identified by their 'minimum sets', i.e., sets A such that $\nu(B) = 0$ for all $B \subset A$ and $\nu(B) = 1$ for all $B \supseteq A$ [25, 29].

If a particular subset of fuzzy measures is a polytope, then it can be defined as the set of all convex combinations of the vertices of that polytope. Therefore, the elements of that subset can be represented through the coefficients that correspond to the vertices, which is useful for learning the suitable elements from data or their random generation. This approach was taken in [13] in the context of k-additive fuzzy measures.

In the case of 2-additive fuzzy measures, all the vertices of that polytope are $\{0,1\}$-fuzzy measures, but the sets of 3-additive and larger fuzzy measures involve many other vertices. Nevertheless, in learning and random generation thereof, it makes sense to identify and use a restricted subset of vertices, with the purpose of reducing the number of defining parameters and constraints, i.e., by using a suitable simplification.

In this paper we follow a similar approach, relying on the fact that convex combinations of (anti)buoyant fuzzy measures remain in that class. Hence we construct various subsets of vertices of (anti)buoyant fuzzy measures and use their convex combinations as possible representations of those objects. Because of the duality between these two classes, we focus on the antibuoyant fuzzy measures. Analogously to reducing the number of parameters in regression, the reduced subsets of vertices limit the modelling capacity of the chosen subsets. On the other hand, the requirements of monotonicity and antibuoyancy are satisfied automatically and need not be enforced through (a large) number of additional constraints.

Some vertices of the antibuoyant set of fuzzy measures can be identified analogously to $\{0,1\}$-measures, by setting the values at certain subsets to 0, and then determining the consequent maximum values of the fuzzy measure subject to antibuoyancy. It turns out that such fuzzy measures are p-symmetric and defined by $(|A_1|+1)\times(|A_2|+1)\times\cdots\times(|A_p|+1)$ values, with $\{A_1,\ldots,A_p\}$ denoting the partition into *sets of indifference* (see Definition 6 below). Some of these can be generated automatically. In particular, algorithms will be presented for single minimum sets and two or more singletons, which result in there being two sets of indifference, A and $N \setminus A$.

There are three contributions in this paper. Firstly, we construct subsets of vertices of antibuoyant fuzzy measures based on their desired cardinalities. Secondly, we study the behaviour of the randomly generated fuzzy measures based on convex combinations of these vertices. We propose three methods for generation of the antibuoyant measures: a) a method based on linear extensions, b) a method based on generating vertices randomly from the minimal sets, and c) a specific combination of these methods. Lastly, we formulate and solve the antibuoyant fuzzy measure learning problem from either empirical data or as an approximation to another more general fuzzy measure.

The paper is structured as follows. In Section 2 we provide the basic definitions needed for the rest of the paper. In Section 3, we determine specific vertices of the polytope of antibuoyant fuzzy measures and their subsets. Section 4 treats in detail a method of random generation of antibuoyant fuzzy measures. Learning of antibuoyant fuzzy measures is considered in Section 5, which is followed by a discussion and conclusions.

2. Preliminaries

We focus on the learning of fuzzy measures satisfying antibuoyancy, which in turn are used to define the parameters of a Choquet integral.

2.1. Fuzzy measures and the Choquet integral

The Choquet integral [10] has received a great deal of attention in recent research, particularly for analysis and prediction tasks, e.g., see [15, 22, 23, 28, 30]. As an averaging function, it has been shown to offer similar versatility in modelling to neural networks and other machine learning techniques, while at the same time having a structure that offers both reliability and interpretability [4, 21].

The parameters of the Choquet integral are given by an associated fuzzy measure.

Definition 1. Let $N = \{1, 2, \ldots, n\}$. A discrete fuzzy measure, or capacity, is a set function $\mu : 2^N \to [0, 1]$ satisfying monotonicity with respect to set inclusion, i.e., $\mu(A) \leq \mu(B)$ whenever $A \subset B$ and with boundary conditions $\mu(\emptyset) = 0$ and $\mu(N) = 1$.

For any given input vector and fuzzy measure, the Choquet integral associates weights according to the relative ordering of the inputs. Following [8, 18], we will use the concept of the discrete derivative to describe this.

Definition 2. Let μ be a set function on N and $A \subseteq N \setminus \{i\}$. The derivative of μ at A with respect to i is

$$\Delta_i \, \mu(A \cup \{i\}) = \mu(A \cup \{i\}) - \mu(A).$$

This allows us to express the calculation of the Choquet integral of a given vector $\mathbf{x} \in \mathbb{R}^n_+$ in the following way.

Definition 3. For a given fuzzy measure μ, the discrete Choquet integral $C_\mu : \mathbb{R}^n_+ \to \mathbb{R}_+$ is given by

$$C_\mu(\mathbf{x}) = \sum_{i=1}^{n} x_{(i)} \; \Delta_{(i)} \, \mu(H_i),$$

where $x_{(1)} \leq \cdots \leq x_{(n)}$ denotes an increasing permutation of the inputs and $H_i = \{(i), (i+1), \ldots, (n)\}$ is the set of corresponding indices from (i) up to (n).

2.2. Classes of simplified fuzzy measures

As evident from the preceding definitions, one of the challenges for the Choquet integral in practice is the exponentially increasing number of parameters required to define the fuzzy measure, which is equal to $2^n - 2$ once we assume $\mu(\emptyset) = 0$ and $\mu(N) = 1$. In response to this, special classes of fuzzy measure have been introduced that reduce the number of parameters required, or simplify the fitting problem in other ways. We present the definitions for k-additive and p-symmetric fuzzy measures here, which hold particular importance for some results and concepts that we will explore in subsequent sections.

The class of k-additive fuzzy measures was introduced in [17, 24]. The most straightforward definition is based on the Möbius representation for fuzzy measures.

Definition 4. For a fuzzy measure μ, its Möbius representation is given for each set $A \subseteq N$ by,

$$M^\mu(A) = \sum_{B \subseteq A} (-1)^{|A \setminus B|} \mu(B).$$

In terms of the Möbius representation of a fuzzy measure, the Choquet integral can be calculated as a linear combination of the fuzzy measure values and a transformed dataset based on taking the minimum across all subsets, i.e., we have

$$C_\mu(\mathbf{x}) = \sum_{A \subseteq N} M^\mu(A) \min_{i \in A} x_i. \tag{1}$$

The idea of k-additivity can be simply expressed as follows.

Definition 5. A fuzzy measure is said to be k-additive when for all $A \subseteq N$ such that $|A| > k$, it holds that $M^\mu(A) = 0$.

This effectively reduces the number of unknowns from $2^n - 2$ to

$$\sum_{i=1}^{k} \frac{n!}{i!(n-i)!}.$$

When $k = 1$, we have the case of additive fuzzy measures, which in turn results in the Choquet integral being equivalent to the weighted arithmetic mean, i.e.,

$$\mathsf{WAM}(\mathbf{x}) = \sum_{i=1}^{n} w_i x_i,$$

with all $w_i \geq 0$ and $\sum\limits_{i=1}^{n} w_i = 1$. The corresponding fuzzy measure will have singleton values $\mu(\{i\}) = w_i$. With $k = n$ we recover the case of general fuzzy measures.

The class of p-symmetric fuzzy measures was introduced in [26]. These fuzzy measures rely on a partition of the inputs into distinct groupings referred to as subsets of indifference.

Definition 6. Given a subset $A \subseteq N$, we say that A is a set of indifference if and only if for all $B_1, B_2 \subset A$ with $|B_1| = |B_2|$ and every $C \subset N \backslash A$ it holds $\mu(B_1 \cup C) = \mu(B_2 \cup C)$.

From this, the definition of p-symmetric fuzzy measures can be given as follows.

Definition 7. A fuzzy measure is said to be p-symmetric if and only if the coarsest partition of the universal set into sets of indifference is $\{A_1, \ldots, A_p\}$ with $A_i \neq \emptyset$ for all $i \in \{1, \ldots, p\}$.

The case of $p = 1$ corresponds with symmetric fuzzy measures where $|A| = |B|$ implies $\mu(A) = \mu(B)$ and which coincide with the ordered weighted averaging (OWA) operators [32]. This is usually expressed as

$$\mathsf{OWA}(\mathbf{x}) = \sum_{i=1}^{n} w_i x_{(i)},$$

with the weights $\mathbf{w} = (w_1, \ldots, w_n)$ being non-negative and summing to 1 as they do for the WAM, however here $x_{(i)}$ denotes a reordering of the inputs such that $x_{(1)} \geq x_{(2)} \geq \cdots \geq x_{(n)}$ (i.e., the opposite interpretation of the notation when used for the Choquet integral).

For p-symmetric fuzzy measures, once the partition is known, the values assigned to each subset of N depend only on the number of members from each of the sets A_i and hence the number of unknowns reduces to $(|A_1| + 1) \times \ldots \times (|A_p| + 1) - 2$.

2.3. Antibuoyant fuzzy measures

The buoyancy property for fuzzy measures was proposed in [5] and applied to welfare and ecology problems in [6], and to discrete optimisation in [1]. The buoyancy definition for OWA operators [33] requires the weighting vector to be non-increasing.

Definition 8. An OWA operator is said to be a *buoyancy measure* if its weighting vector $\mathbf{w}$ satisfies the additional condition that $w_i \geq w_j$ for all $i < j$.

This results in the largest weight being applied to the largest input and so on. We refer to such weight vectors $\mathbf{w}$ as buoyant and adopt the term antibuoyant where $w_i \leq w_j$. For fuzzy measures, this property is required of all effective weighting vectors that can result from the different relative orderings of the input vector.

As in [5], we adopt the notation $\mathbf{w}^\sigma$ for these resulting vectors, i.e.,

$$\begin{aligned} \mathbf{w}^\sigma &= (\Delta_{(n)}\,\mu(H_n), \Delta_{(n-1)}\,\mu(H_{n-1}), \ldots, \Delta_{(1)}\,\mu(H_1)) \\ &= (\mu(\{(1)\}), \mu(\{(1),(2)\}) - \mu(\{(1)\}), \ldots, \mu(N) - \mu(N \setminus \{(n)\}) \end{aligned}$$

We can then define buoyant and antibuoyant fuzzy measures as follows.

Definition 9. A fuzzy measure μ is buoyant if all ordered weighting vectors $\mathbf{w}^\sigma$ associated with μ are buoyant. A fuzzy measure is antibuoyant if all ordered weighting vectors are antibuoyant.

Example 1. For $N = \{1, 2, 3\}$, the fuzzy measure with values $\mu(\{1\}) = 0.5$, $\mu(\{2\}) = 0.45$, $\mu(\{3\}) = 0.8$, $\mu(\{1, 2\}) = 0.75$, $\mu(\{1, 3\}) = \mu(\{2, 3\}) = 0.9$, and $\mu(\{1, 2, 3\}) = 1$ is buoyant. The 3! potential weighting vectors $\mathbf{w}^\sigma$ depending on the input ordering are: $(0.5, 0.25, 0.25)$, $(0.5, 0.4, 0.1)$, $(0.45, 0.3, 0.25)$, $(0.45, 0.45, 0.1)$, and $(0.8, 0.1, 0.1)$ (twice).

Note that $\mu(\{3\}) > \mu(\{1,2\})$ even though the cardinality is lower, i.e., buoyant and antibuoyant fuzzy measures are not necessarily *balanced*.

From hereon, we will be concerned with the *antibuoyant* case, where smaller weights are applied to larger inputs, since such weighting vectors are most important when considering the measures of diversity satisfying the Pigou–Dalton principle [14] often applied in economics [27] and ecology [31].

In the welfare setting, the Pigou–Dalton principle requires that any redistribution of wealth from a richer to poorer individual should increase overall measures of welfare (or decrease measures of inequality). For an input vector $x_1 \geq x_2 \geq \cdots \geq x_n$, we can say that a function f satisfies the Pigou–Dalton principle if for all $x_i - h \geq x_j + h, h > 0$ it holds that $f(x_1, \ldots, x_i, \ldots, x_j, \ldots, x_n) \leq f(x_1, \ldots, x_i - h, \ldots, x_j + h, \ldots, x_n)$. It is also referred to as the *progressive transfers* principle.

We remark that while antibuoyancy is related to the convexity/concavity of the resulting Choquet integral, it is a stricter requirement since a Choquet integral can be concave (defined by a supermodular fuzzy measure [7, 10]) without its associated fuzzy measure being antibuoyant [5]. The duals of buoyant fuzzy measures are antibuoyant and vice versa.

2.4. Learning antibuoyant fuzzy measures

The standard data-fitting problem with the Choquet integral assumes we have a dataset consisting of M input/output pairs $(\mathbf{x}^m, y^m)$ for $m = 1, \ldots, M$. We then aim to learn the values of a fuzzy measure μ such that the overall difference between the actual outputs y^m and predicted outputs $C_\mu(\mathbf{x}^m)$ is minimised. If the criteria for *closeness* is based on the least absolute deviation, we have the objective [8]

$$\text{Minimise}_\mu \sum_{m=1}^{M} |C_\mu(\mathbf{x}^m) - y^m|.$$

For implementation, among other methods, this can be expressed as a linear program by learning the Möbius values of the fuzzy measure and fitting to transformed input vectors with arguments $\min_{i \in A} x_i$ and arranged in cardinality ordering. The residuals are broken into their positive and negative components such that $C_\mu(\mathbf{x}^m) - y^m = r_m^+ - r_m^-$ and $|C_\mu(\mathbf{x}^m) - y^m| = r_m^+ + r_m^-$ (see, e.g., [8]). Linear constraints are then imposed so that the fuzzy measure is monotone and satisfies the boundary condition $\mu(N) = 1$.

For antibuoyant fuzzy measures, antibuoyancy requires constraints of the form

$$\sum_{\substack{B \subseteq A \cup \{i,j\} \\ \{i,j\} \subseteq B}} M^{\mu}(B) \geq \sum_{\substack{B \subseteq A \cup \{i\} \\ i \in B}} M^{\mu}(B) - \sum_{\substack{B \subseteq A \cup \{j\} \\ j \in B}} M^{\mu}(B), \tag{2}$$

for all $i, j \in N$ and all $A \subseteq N \setminus \{i, j\}$.

These constraints make the monotonicity constraints redundant, however many more are required.

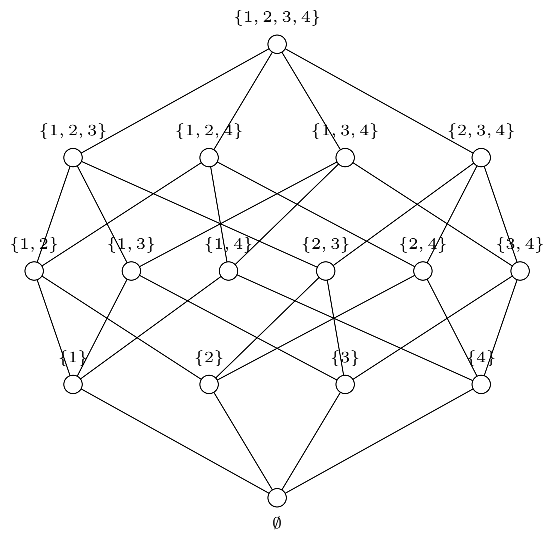

Figure 1: Hasse Diagram for a fuzzy measure with $n = 4$

For illustration, consider the Hasse diagram in Fig. 1. Edges map to the discrete derivatives and hence the potential weights associated with each variable. In the case of general fuzzy measures, we merely require that each of these weights is non-negative and so the number of edges in the Hasse diagram corresponds with the number of weights.

For antibuoyancy however, we need to look at pairs of adjacent edges between subsequent levels. At each level, there are $\binom{n}{i}$ subsets, each with $\binom{i}{i-1} = i$ edges leading to subsets below. Hence, the number of antibuoyancy constraints, which require comparison of adjacent edges at different levels, will be:

$$\sum_{i=2}^{n} \binom{n}{i} i(i-1) = \sum_{i=2}^{n} \frac{n!}{(i-2)!(n-i)!}. \tag{3}$$

For comparison, the number of monotonicity constraints is

$$\sum_{i=1}^{n} \binom{n}{i} i = \sum_{i=1}^{n} \frac{n!}{(i-1)!(n-i)!}.$$

Table 1 shows the number of variables, monotonicity constraints and antibuoyancy constraints for each $n = 2, \ldots, 10$. While the growth in the number of unknowns with increasing n is a challenge often cited, clearly the number of monotonicity constraints is also a problem for general fuzzy measures and requiring the fuzzy measure to be antibuoyant exacerbates this problem further.

Table 1: Number of variables, monotonicity and antibuoyancy constraints required when fitting fuzzy measures to data

n	variables	monotonicity	antibuoyancy
2	2	4	2
3	6	12	12
4	14	32	48
5	30	80	160
6	62	192	480
7	126	448	1344
8	254	1024	3584
9	510	2304	9216
10	1022	5120	23040

The following methods are hence targeted toward the learning of a suitably rich but simplified class of antibuoyant fuzzy measures such that the fitting problem is reduced.

3. Vertices of the antibuoyant fuzzy measures polytope

A key result that applies to general fuzzy measures as well as some special families is that they are closed under convex combinations. As observed in, e.g., [25], this can be particularly useful in learning or random generation since the result of any convex combination or weighted mean of fuzzy measures need not be checked to ensure properties such as monotonicity and boundary conditions are satisfied. Of interest then is understanding the extreme points or *vertices* of such sets, which define polytopes in a (2^n-2)-dimensional space or some subspace thereof. While previous studies [13, 25] have highlighted that the growth in the vertices presents a significant challenge, there remains the potential for particular vertices and subsets of vertices to be useful both when it comes to learning particular types of fuzzy measures as well as for random generation.

3.1. Vertices defining the polytope of general fuzzy measures

Works such as [25, 29] have established that the vertices of the polytope defining the set of general fuzzy measures correspond with the set of $\{0,1\}$-fuzzy measures, i.e., fuzzy measures whose sets are assigned a value of either 0 or 1. As soon as one subset has a value of 1, all supersets will also be fixed at 1, so it has been shown that such fuzzy measures can be identified according to their minimum sets (by themselves or in combination), defined in [25] as sets A such that

$$\begin{aligned} &\mu(A) = 1 \\ &\mu(B) = 1, \forall B \supseteq A \\ &\mu(C) = 0, \forall C \subset A. \end{aligned}$$

The set of $\{0,1\}$-fuzzy measures for $n = 3$ is shown in Table 2 along with the minimal sets defining them.

Unfortunately, the number of vertices follows the sequence of the Dedekind numbers, $1, 4, 18, 166, 7579, 7828352, 2414682040996, 5.6\times10^{22}$ [13] making it impossible to list or store the fuzzy measures even for modest values of n.

Some fuzzy measure families can be expressed as convex sets defined by subsets of the $\{0,1\}$-vertices. Some notable examples include:

- additive fuzzy measures — defined by the vertices with a singleton as the minimal set, i.e., for each $i = 1,\ldots,n$,

$$\mu(A) = \begin{cases} 1, & A \ni i \\ 0, & \text{otherwise.} \end{cases}$$

Table 2: List of all $\{0,1\}$-fuzzy measures for $n = 3$ and their minimal sets

label	$\{1\}$	$\{2\}$	$\{3\}$	$\{1,2\}$	$\{1,3\}$	$\{2,3\}$	$\{1,2,3\}$	minimal sets
u_1	0	0	0	0	0	0	1	$\{1,2,3\}$
u_2	0	0	0	0	0	1	1	$\{2,3\}$
u_3	0	0	0	0	1	0	1	$\{1,3\}$
u_4	0	0	0	0	1	1	1	$\{1,3\},\{2,3\}$
u_5	0	0	0	1	0	0	1	$\{1,2\}$
u_6	0	0	0	1	0	1	1	$\{1,2\},\{2,3\}$
u_7	0	0	0	1	1	0	1	$\{1,2\},\{1,3\}$
u_8	0	0	0	1	1	1	1	$\{1,2\},\{1,3\},\{2,3\}$
u_9	0	0	1	0	1	1	1	$\{3\}$
u_{10}	0	0	1	1	1	1	1	$\{3\},\{1,2\}$
u_{11}	0	1	0	1	0	1	1	$\{2\}$
u_{12}	0	1	0	1	1	1	1	$\{2\},\{1,3\}$
u_{13}	0	1	1	1	1	1	1	$\{2\},\{3\}$
u_{14}	1	0	0	1	1	0	1	$\{1\}$
u_{15}	1	0	0	1	1	1	1	$\{1\},\{2,3\}$
u_{16}	1	0	1	1	1	1	1	$\{1\},\{3\}$
u_{17}	1	1	0	1	1	1	1	$\{1\},\{2\}$
u_{18}	1	1	1	1	1	1	1	$\{1\},\{2\},\{3\}$

- symmetric fuzzy measures — defined by the symmetric $\{0,1\}$-fuzzy measures, i.e., for each $i = 1,\ldots,n$,

$$\mu(A) = \begin{cases} 1, & |A| \geq i \\ 0, & \text{otherwise.} \end{cases}$$

Recall that Choquet integrals defined with respect to additive fuzzy measures result in the WAM and when defined with respect to symmetric fuzzy measures, the Choquet integral is equivalent to the OWA.

From Table 2, the vertices for $n = 3$ would be u_9, u_{11}, and u_{14} for additive fuzzy measures and u_1, u_8, and u_{18} for symmetric fuzzy measures. Another example is the set of belief measures, defined by the set of vertices that only have a single subset as their minimal subsets.

Special classes such as the k-additive fuzzy measures are actually much more complex in their structure, having vertices that are not $\{0,1\}$-fuzzy measures and growing at an even faster pace than the Dedekind numbers.

However, while this is the case for $k \geq 3$, the 2-additive case does allow representation from a subset of the $\{0,1\}$-fuzzy measure vertices. The set is comprised of the singleton and pair minimal sets, as well as the pairs of singletons. For $n = 3$, this would be vertices $u_2, u_3, u_5, u_9, u_{11}, u_{13}, u_{14}, u_{16}, u_{17}$ shown in Table 2. The number of such vertices for any given n is

$$n + 2\binom{n}{2} = n + n(n-1) = n^2.$$

For purposes of learning, this is $n(n-1)/2$ higher than the number of variables that would be required in the standard approach to learning 2-additive fuzzy measures using Möbius values, however there is an advantage in that we can do away with the monotonicity constraints.

We let $\mathbf{U} = \{\mu^1, \ldots, \mu^u\}$ denote the subset of fuzzy measure vertices we wish to use, with $|\mathbf{U}| = u$. Then, any convex combination of the fuzzy measures in $\mathbf{U}$ is given by

$$\mu = \sum_{i=1}^{u} c_i \mu^i$$

with $\sum_{i=1}^{u} c_i = 1$ and $c_i \geq 0$ the set of coefficients.

Hence, when it comes to fitting, rather than our variables being the individual weights of the fuzzy measure, we transform the data such that the c_i are our decision variables, i.e., we have

$$\sum_{A \subseteq N} \left(\sum_{i=1}^{u} c_i \mu^i(A) \right) \min_{j \in A} x_j = \sum_{i=1}^{u} c_i \left(\sum_{A \subseteq N} \mu^i(A) \min_{j \in A} x_j \right),$$

which is linear in $\mathbf{c}$.

Any subset of the $\{0,1\}$-fuzzy measure vertices (or indeed any finite set of fuzzy measures) can hence be used as an approximation of the best-fitting fuzzy measure within the resulting convex set. We can arbitrarily choose any number of vectors to include in $\mathbf{U}$, obtaining access to the full set of fuzzy measures once we have all vertices.

While this approach may not be particularly advantageous when it comes to fitting general fuzzy measures, we will see that for antibuoyant fuzzy measures we can use subsets of vertices to drastically reduce the difficulty of the fitting problem, albeit for simplified classes of antibuoyant fuzzy measures.

3.2. Antibuoyant vertices based on the $\{0,1\}$ vertices

Evidently, the $\{0,1\}$-fuzzy measures are not generally antibuoyant and so the vertices of the antibuoyant set would also be expected, like the k-additive fuzzy measures, to form a much more complex structure. However, just as the concept of minimal sets is used to identify the $\{0,1\}$-fuzzy measures, we can use an approach based on zero-valued sets and p-symmetric fuzzy measures.

Given $\mu(A) = 0$ (or multiple sets $A_1, A_2, \ldots,$ etc.) and $\mu(B) > 0, \forall B \supset A$, we can set the $\mu(B)$ values to the largest values possible subject to antibuoyancy. Take the following example.

Example 2. Let $\mu(\{1\}) = 0$ with all other values non-zero. The largest possible assignment for supersets of $\{1\}$ is such that we would obtain a weighted mean of the first two arguments, hence we have $\mu(\{1,2\}) = \mu(\{1,3\}) = 1/2$ and $\mu(\{1,2,3\}) = 1$. From this, the maximum value that can be assigned to $\mu(\{2\})$ and $\mu(\{3\})$, based on being subsets of the previously determined pairs, is $1/4$. This leaves us lastly to assign the value of $\mu(\{2,3\})$, whose only superset is $\{1,2,3\}$, and whose subsets will weight $1/4$ to the largest input. The largest value it can take is hence $(1+1/4)/2 = 5/8$.

By proceeding in this way, we obtain fuzzy measures at the extremes of the antibuoyant set.

Proposition 1. *Fuzzy measures with all non-zero sets assigned the maximum values subject to the antibuoyancy condition are vertices of the antibuoyant fuzzy measures polytope.*

Proof. A fuzzy measure is a vertex if it cannot be obtained as a convex combination of two or more other vertices. We can first note that with the antibuoyancy condition, increases to any set $\mu(A)$ will never result in a decrease to permissible values of $\mu(B)$ for any $B \supset A$, as this would be equivalent to redistributing weight from smaller to larger inputs resulting in a violation of the Pigou–Dalton principle. This means we can contain our focus to local pairs $A \subset B$ with $|A| = |B| - 1$. We hence consider two cases, (i) $\mu(A) = 0 < \mu(B)$ and (ii) $0 < \mu(A) < \mu(B)$ and show that a convex combination $\mu = c_1\mu^1 + c_2\mu^2$ with μ^1 and μ^2 distinct from μ cannot exist.

Case (i). Clearly if $\mu^1(A) > 0$ or $\mu^2(A) > 0$ then $\mu(A) = 0$ cannot be obtained as a convex combination. Further, if $\mu(A) = \mu^1(A) = \mu^2(A) = 0$, then it is not possible for $\mu^1(B) > \mu(B)$ or $\mu^2(B) > \mu(B)$ because this would imply that $\mu(B)$ is not set to the maximum value subject to antibuoyancy. We can also note that non-zero subsets of B do not affect the maximum value $\mu(B)$ can take if $\mu(A) = 0$.

Case (ii). Let us assume $\mu^1(A) < \mu(A)$ and hence $\mu^2(A) > \mu(A)$. This means that subsets $C \subset A$ must also have a higher value associated with them in μ^2 than in μ, otherwise $\mu(A)$ has not been assigned the maximum allowable value, which in turn means that either there is a set $C \subset A$ such that $\mu(C) = 0$ and $\mu^2(C) > 0$, or that one of the non-zero subsets is not set to its maximum in μ. Hence, μ cannot be obtained from a convex combination that involves μ^2 unless $c_2 = 0$. □

There will be a correspondence between each of these and the $\{0, 1\}$-fuzzy measures, i.e., they will have the same zero-valued sets, however it should be noted that this will by no means capture all of the antibuoyant vertices. For instance, in the previous example, assigning $\mu(\{2, 3\}) = 1/2$ (which is the minimum and not the maximum value once other values are set) is also an extreme point of the antibuoyant set, since it cannot be obtained by convex combinations of any other vertices.

The aim is not to be able to exhaustively list all of the antibuoyant vertices, since, as is the case with k-additive and general fuzzy measures, this would be met with prohibitively high storage space required for implementation. Rather, we aim to be able to quickly identify specific vertices that can then be used as vectors in $\mathbf{U}$ as described previously for fitting. In particular, we might aim to find vertices *analogous* to those used to form the set of additive fuzzy measures, symmetric fuzzy measures or 2-additive fuzzy measures.

3.3. Algorithmic generation of antibuoyant vertices based on zero-valued sets and p-symmetric fuzzy measures

As it happens, vertices generated in the manner as described in Example 2 are particular instances of p-symmetric fuzzy measures.

Let us assume we have an antibuoyant fuzzy measure corresponding with one of the $\{0, 1\}$-fuzzy measures such that $\mu(A) = 0$ and non-zero for all subsets B such that $B \not\subseteq A$. Then, there will be two sets of indifference, A and $N \backslash A$. This means there will be $(|A| + 1)(n - |A| + 1)$ values that need to be assigned, which can be achieved according to the following algorithm.

The algorithm sets up the $(n - |A| + 1) \times (|A| + 1)$ matrix holding the values of the p-symmetric fuzzy measure. The columns correspond with the

Algorithm 1 Generation of a p-symmetric antibuoyant vertex

Input n, A ▷ A is the largest zero-set
Output μ

$a \leftarrow |A|$
$b \leftarrow n - a$
Initialise $(b+1) \times (a+1)$ matrix P, with $P_{i,j} = 0, \forall i, j$
for $i = 2, 3, \ldots, b+1$ **do**
  $P_{i,a+1} \leftarrow (i-1)/b$ ▷ Values assigned to supersets of A
end for
for $j = a, a-1, \ldots, 1$ **do**
  for $i = 2, 3, \ldots, b+1$ **do**
    $P_{i,j} \leftarrow (P_{i-1,j} + P_{i,j+1})/2$ ▷ Mean of entries to the right and above
  end for
end for
for $B \subseteq N$ **do**
  $i \leftarrow |B \setminus A| + 1$
  $j \leftarrow |B \cap A| + 1$
  $\mu(B) \leftarrow P_{i,j}$
end for

number of elements from the zero-valued set A, while rows correspond with the number of elements from the non-zero set. The entries across the first row (subsets of A) will all be zero, and then values are assigned to the supersets of A so that the weights along this simplex will be equally distributed between the smallest $(n-|A|)$ inputs. The algorithm then proceeds along the matrix, top to bottom and right to left, assigning the mean of the upper and right entries.

Here is an example of how the algorithm is applied.

Example 3. Let $A = \{1, 2, 3\}$ for $n = 5$. Then we have a $(2+1) \times (3+1)$ matrix P. Results from the following steps are shown in Figure 2.

Step 1) We know that all subsets $B \subseteq A$ will satisfy $\mu(B) = 0$.

Step 2) We then fill down the last column with an arithmetic sequence ending in 1.

Step 1)

	0	1	2	3
0	0	0	0	0
1				
2				

Step 2)

	0	1	2	3
0	0	0	0	0
1				1/2
2				1

Step 3)

	0	1	2	3
0	0	0	0	0
1			1/4	1/2
2			5/8	1

Step 4)

	0	1	2	3
0	0	0	0	0
1	1/16	1/8	1/4	1/2
2	7/32	3/8	5/8	1

Figure 2: Calculation of the entries in the matrix P from Example 3

Step 3) Starting on the second row and the second last column (the cell corresponding with 2 elements from A and 1 element not from A), we assign the average of the cell to the right and above.

Step 4) This continues until we complete the entire matrix.

After the matrix is obtained, the corresponding fuzzy measure values are determined based on how many of the elements come from A and how many come from $N \backslash A$.

Note that after the zero-valued subsets are assigned, this process ensures the subset values will be as large as possible and hence will produce a fuzzy measure that sits at the extreme of the antibuoyant set, i.e., a vertex. Taking the average of the cells above and right results in subsequent weights from the associated simplex having values $w_k = P_{i,j} - P_{i-1,j} = P_{i,j+1} - P_{i,j} = w_{k-1}$. Hence the effective weight for the larger input w_k, or the value assigned to the subset of lower cardinality, will be as large as possible. Then, due to the averaging property, we also know that these weights will be between the weights resulting from values already assigned that result from adding the elements in the reverse ordering, i.e., $P_{i-1,j+1} - P_{i-1,j} \leq w_k = w_{k-1} \leq P_{i,j+1} - P_{i-1,j+1}$. The antibuoyancy property can be seen to be satisfied since the w_k resulting from proceeding right from an entry $P_{i,j}$ will be guaranteed to be lower or equal to the next step going down from $P_{i,j+1}$, since $P_{i,j+1} - P_{i-1,j+1} \leq P_{i+1,j+1} - P_{i,j+1}$.

This process allows us to create antibuoyant fuzzy measures corresponding with the $\{0,1\}$-fuzzy measures with singleton minimum sets and which form the vertices of the additive fuzzy measures, i.e., the minimum set $\{1\}$ would correspond with the defining zero-set being $A = N\backslash\{i\}$. In fact, note for these that $n - |A| = 1$ and hence the matrix will only have two rows, with the non-zero entries simply being $1/2^{|A|+1-j}$ or $\mu(B) = 1/2^{n-|B|}$ for all $B \supseteq \{i\}$.

We emphasise that these fuzzy measures are not themselves additive, but rather give the extreme case of assigning as much value as possible to a single variable, whilst maintaining the antibuoyancy property. For a chosen $\{i\}$ being the non-zero singleton set, we can observe that the corresponding input will always be assigned as much weight as possible in weight vectors resulting from μ. We have

$$\mu(A) - \mu(A\backslash\{i\}) > \mu(A) - \mu(A\backslash\{j\}),$$

for all $A \supseteq \{i,j\}, i \neq j$, since it will always hold that $\mu(A\backslash\{i\}) = 0$ and $\mu(A\backslash\{j\}) > 0$.

It is informative to look at the Shapley values associated with such fuzzy measures. Shapley values (see, e.g., [8, 18, 20]) are often used to interpret and understand the overall behaviour of fuzzy measures, giving an idea of the average importance or weight assigned to a particular input depending on the relative order. The calculation in standard representation is

$$\phi(i) = \sum_{A \subseteq N\backslash\{i\}} \frac{(n-|A|-1)!|A|!}{n!}[\mu(A \cup \{i\}) - \mu(A)],$$

which we see averages the discrete derivatives involving i across the fuzzy measure. Table 3 gives the ratio of importance as measured by the Shapley value $\phi(i)$ for the antibuoyant vertices generated from zero-sets $N\backslash\{1\}$ for $n = 2, \ldots, 5$.

Table 3: Maximum importance allocation to a single input subject to antibuoyancy

n	Ratio of Shapley values
2	3:1
3	14:5:5
4	45:17:17:17
5	124:49:49:49:49

These ratios help show the maximum diversity that can be achieved when distributing importance amongst the variables for the class of antibuoyant fuzzy measures.

The vertices with minimal sets defined by pairs of singletons can also be found using Algorithm 1. In this case, the defining zero-set will be $A = N\backslash(\{i\} \cup \{j\})$. Example 3 above gives one such instance.

Note that for an analogue of 2-additive fuzzy measures, we require minimum sets made up from the singletons, pairs of singletons, and each of the pairs by themselves. While the singletons and pairs of singletons can be generated using Algorithm 1, generating vertices based on pairs will proceed slightly differently. We still obtain two sets of indifference, with $A = N\backslash\{i, j\}$ being our zero-set and $\{i, j\}$ corresponding with the minimum set. We then proceed in a similar fashion, except that now both of the first two rows of the matrix are pre-filled with zeros. This is because we only have value assigned once both i and j are in the set. These antibuoyant fuzzy measures are hence also fairly simple, with supersets $B \supseteq A$ having value $\mu(B) = 1/2^{n-|B|}$.

There are, of course, many other types of vertices, some of which would correspond with combinations of minimum sets and zero-sets, however we will not consider these here. With the exception of combinations of singletons, such fuzzy measures would often require more than two sets of indifference, and hence we no longer have the option of averaging the pairs of adjacent cells in the matrix P. While noting this, we will see that the p-symmetric fuzzy measures that can be easily generated using Algorithm 1 might be sufficiently useful for contributing importance to certain variables, and in combination with the symmetric antibuoyant fuzzy measures may be adequate for suitable approximations determined by the set of fuzzy measures in $\mathbf{U}$.

4. Random generation of antibuoyant fuzzy measures

In [6], some basic random generation methods for antibuoyant and buoyant fuzzy measures were proposed. One approach consisted of taking random values within each cardinality and then using those values to determine subsequent intervals as cardinality is increased. This can only produce balanced fuzzy measures, i.e., fuzzy measures such that if $|A| > |B|$, then $\mu(A) \geq \mu(B)$, and so an augmentation method was suggested whereby convex combination with extreme measures, such as those corresponding with only supersets of the singletons being non-zero, were taken.

Here we propose some further options, in particular by adapting the linear extension approach such as the one that has been used for general fuzzy measures in [2, 9, 11, 12].

4.1. Linear extensions for fuzzy measures

Fuzzy measures define a partial order over all subsets of N. There are many complete orders compatible with this partial order, and of course, any completely defined fuzzy measure (with distinct values) will be associated with one of these orderings.

The main steps of linear extension approaches to generating random fuzzy measures is to, first, determine a random linear extension compatible with the partial order, and then assign values based on a sorted uniform distribution on $[0,1]^{2^n-2}$ with the value of $\mu(N)$ set to 1, or alternatively the values can be generated on $[0,1]^{2^n-1}$ and then normalised.

An important aspect of such approaches is the probabilities with which the linear extensions are generated, so that they converge towards a uniform distribution, which in turn makes them suitable for experimentation without bias concerns.

4.2. Algorithm for generating random values

Any antibuoyant fuzzy measure will also be associated with a linear extension of the fuzzy measure, and indeed, for any linear extension, there will be a number of potential random assignments of antibuoyant fuzzy measure values.

However, while the monotonicity requirements are satisfied with any sorted vector of values assigned to the linear extension, there are further requirements when it comes to antibuoyancy. The approach we take starts with the linear extension, and then assigns values in order, storing the weights corresponding with the edges into each subset in the corresponding Hasse diagram and using this to set the minimum value. Algorithm 2 is given, with $\mathbf{L}$ denoting the linear extension, i.e., with each L_i a given set and $L_i \prec L_{i+1}$ consistent with the partial ordering. The vector $\mathbf{d}$ is intended to be calculated from a sorted randomly generated vector $\mathbf{r}$ with $d_i = r_i - r_{i-1}$ and $r_0 = 0$ by convention.

Algorithm 2 Random generation of an antibuoyant fuzzy measure

---

Input $\mathbf{L}, \mathbf{d}$ ▷ A linear extension for a given n and a vector of differences
Output μ

Initialise $\mu(A) \leftarrow 0$, edgesInto$(A) \leftarrow 0$, for all A
for $i = 1, 2, \ldots, 2^n - 1$ **do**
  $A \leftarrow L_i$
  $\text{minVal}(A) \leftarrow \max\left\{\max_{B \subseteq N} \mu(B), \max_{B \subseteq A}(\mu(B) + \text{edgesInto}(B))\right\}$
  $\mu(A) \leftarrow \text{minVal}(A) + d_i$
  if $|A| = 1$ **then**
    $\text{edgesInto}(A) \leftarrow \mu(A)$
  else
    $\text{edgesInto}(A) \leftarrow \mu(A) - \min_{B \subset A, |B| = |A| - 1} \mu(B)$
  end if
end for
$\mu(A) \leftarrow \mu(A)/\mu(N)$, for all A

As well as being capable of generating random antibuoyant fuzzy measures corresponding with linear extensions, the same algorithm can also be used to generate vertices of the antibuoyant set, by setting $\mathbf{d}$ such that all values are zero except one d_i, which is set to 1. Once normalised, this gives the fuzzy measure on the given simplex corresponding with the linear extension, with all sets $L_1, \ldots, L_{i-1}$ assigned a zero value, and all $L_{i+1}, \ldots, N$ are minimised subject to L_i.

Of course due to the closure of antibuoyant fuzzy measures, it is always possible to use a number of different random generation methods and then take a convex combination.

4.3. Balanced and unbalanced antibuoyant fuzzy measures

The random generation method described tends toward generating fuzzy measure values with minimum-like behaviour and the majority of weight allocated to higher subsets. This is due to there being more jumps required whenever a subset is preceded by a subset of higher cardinality in the linear extension.

For $n = 3$, subject to permutations or relabelling of the inputs, there are only two unique linear extensions, i.e.,

$$\{1\}, \{2\}, \{3\}, \{1,2\}, \{1,3\}, \{2,3\}, \{1,2,3\},$$

and

$$\{1\}, \{2\}, \{1,2\}, \{3\}, \{1,3\}, \{2,3\}, \{1,2,3\}.$$

The first type will always result in balanced fuzzy measures, and hence we would expect these to be distributed around the distribution of symmetric antibuoyant fuzzy measures. We make the observation that if $\mu(\{1\}) = t$ then $\mu(\{1,2\} \geq 2t$ and $\mu(\{1,2,3\}) \geq 3t$. On the other hand, if $\mu(\{1\}) = t$ in the unbalanced extension, then $\mu(\{1,2\}) \geq 2t$, $\mu(\{1,3\}) \geq 4t$ (since $\mu(\{3\}) \geq 2t$) and $\mu(\{1,2,3\}) \geq 7t$ since $\mu(\{1,3\}) - \mu(\{1\}) \geq 3t$.

This discrepancy is further exacerbated for $n = 4$. For balanced fuzzy measures, we have $\mu(\{1\}) = t \to \mu(\{1,2,3,4\}) \geq 4t$, while in the extreme case of the linear extension corresponding with the binary order of subsets, $\mu(\{1\}) = t \to \mu(\{1,2,3,4\}) \geq 100t$. Even if only one of the pairs is out of order, i.e., we have $\{3,4\}$ out of position such that

$$\{1\}, \ldots, \{2,4\}, \{1,2,3\}, \{3,4\}, \{1,2,4\}, \{1,3,4\}, \{2,3,4\}, \{1,2,3,4\},$$

then just with this change it follows that $\mu(\{1\}) = t \to \mu(\{1,2,3,4\}) \geq 8t$.

We therefore may wish to do some adjustment so that the value distribution is not so extreme. We can take convex combinations of our generated fuzzy measure with other symmetric fuzzy measures. This is appropriate if we want the distribution of our randomly generated values to be closer to resembling the distribution of symmetric antibuoyant fuzzy measures. The additive symmetric fuzzy measure is one option that should disrupt the difference in values between subsets as little as possible. Coefficients can be selected at random, or we can employ fitting so that the overall behaviour is targeted towards a specific fuzzy measure. For example, we can follow the steps in Algorithm 3, which supposes a randomly generated antibuoyant fuzzy measure μ^r and a randomly generated antibuoyant weight vector $\mathbf{w}$.

Algorithm 3 Random generation of an antibuoyant fuzzy measure with cardinality index adjustment

---

Input $\mu^r, \mathbf{w}$ ▷ A randomly generated fuzzy measure using Algorithm 2 and random weight vector
Output μ

Initialise μ^s ▷ This will hold the values of a symmetric antibuoyant fuzzy measure
Sort w in such a way that $w_1 \le w_2 \le \cdots \le w_n$
for $i = 1, 2, \ldots, n$ **do**
    $\mu(A) \leftarrow \sum_{j=1}^{i} w_j$, for all $|A| = i$
end for
$c_1, c_2 \leftarrow \arg\min_{c_1,c_2} \|\mu^s - (c_1\mu^a + c_2\mu^r)\|$
$\mu \leftarrow (c_1\mu^a + c_2\mu^r)$

The values of c_1, c_2 are found using linear fitting as will be described in the next section, with $\mathbf{U}$ consisting of a randomly generated fuzzy measure along with the additive and symmetric fuzzy measure μ^a, and the target fuzzy measure being μ^s, which is determined randomly. The aim here is toward ensuring that the antibuoyant fuzzy measures generated are uniform in their associated cardinality indices and not as extreme.

5. Learning experiments

Different choices for the set of fuzzy measures in $\mathbf{U}$ will clearly affect the modelling capability and fitting complexity. The two main aims for our experimentation here are, firstly, to compare different choices of $\mathbf{U}$ when it comes to modelling accuracy, and secondly, to assess whether or not the approach of fitting convex combinations of antibuoyant vertices provides adequate performance for applications. We hence perform two sets of experiments toward these aims.

5.1. Fitting experiments 1: fitting to random fuzzy measures

In the first set of experiments, we generate a random antibuoyant fuzzy measure and try to learn the values directly using different subsets of antibuoyant fuzzy measure vertices. In this case, given the randomly generated antibuoyant fuzzy measure μ^r and the set of antibuoyant vertices in $\mathbf{U}$, we optimise the following with respect to $\mathbf{c}$:

$$
\begin{aligned}
\text{Minimise} \quad & \sum_{A \subseteq N} \frac{(n-|A|)!|A|!}{n!} \left| \mu^r(A) - \sum_{i=1}^{u} c_i \mu^i(A) \right| \\
\text{s.t.} \quad & \sum_{i=1}^{u} c_i = 1 \\
& c_i \geq 0, \ \forall i.
\end{aligned} \tag{4}
$$

The coefficient of the absolute differences scales the absolute differences between target and fitted fuzzy measure values according to the number of subsets for each cardinality. Our main aim is to compare the relative performance of different set choices for $\mathbf{U}$, which are summarised in Table 4.

As noted previously, antibuoyant fuzzy measures randomly generated using the linear extension method in Algorithm 2 tend to have very low values for smaller subsets. To understand some of the bias associated with the different choices of $\mathbf{U}$, we generated the random fuzzy measures μ^r using two approaches with 100 experiments each. The first approach used Algorithm 2 as is, and then in the second approach we augmented each of those measures by taking a weighted average of the generated μ^r and the additive and symmetric fuzzy measure. The augmented fuzzy measures hence moderate the extreme tendency of weighting toward the subsets of larger cardinality, but will obviously favour any $\mathbf{U}$ that includes the additive symmetric fuzzy measure (i.e., sym and add+sym).

Results for the 100 randomly generated antibuoyant fuzzy measures using the standard approach are shown in Table 5 while results for the augmented fuzzy measures are shown in Table 6.

Values in the tables are rounded to three decimal places with the mean result of the 100 experiments in each case summarised along with the standard deviation. The tables are split between random and non-random methods, with bold denoting the best overall performance across all 8 choices of $\mathbf{U}$.

Table 4: Sets of antibuoyant vertices to construct $\mathbf{U}$

name	$\lvert\mathbf{U}\rvert$	description of vertices included
add	n	analogue of vertices resulting in the WAM, i.e., Algorithm 1 with $A = N \setminus \{i\}, i = 1, \ldots, n$
sym	n	analogue of vertices resulting in the OWA, i.e., $\mu(A) = 0$ if $\lvert A\rvert < i$ and $\mu(A) = (\lvert A\rvert - i + 1)/(n - i + 1)$ otherwise, $i = 1, \ldots, n$
add+sym	$2n$	analogues of both WAM and OWA included
2add	n^2	analogue of the 2-additive vertices, i.e., Algorithm 1 with $A = N \setminus \{i\}, i = 1, \ldots, n$ (as with add), $A = N \setminus \{i, j\}$ for all pairs, and vertices with only sets $A \supseteq \{i, j\}$ having a non-zero value and set to maximum (see description in Section 3.3)
rand1	n or n^2	random set of vertices generated using Algorithm 1 with the zero-set A chosen randomly each time (for $n = 3$ only the six non-empty sets $A \subset N$ are used instead of using n^2)
rand2	n or n^2	random set of vertices generated using Algorithm 2 with a randomly generated linear extension and $\mathbf{d} = (0, \ldots, 0, 1, 0, \ldots, 0)$ with $d_i = 1$ in a random position

The values themselves represent the objective given in (4), and so should be considered taking the value of n into account.

For fuzzy measures generated using the linear extension approach in Algorithm 2, the 2add set performed the best overall for lower values of n, but was outperformed by the add+sym set for $n > 6$. It was also outperformed by the rand2 approach with $\lvert\mathbf{U}\rvert = n^2$ (matching the number of variables used). A reasonable explanation as to why we see these results is the ten-

Table 5: Antibuoyant fuzzy measures generated using the linear extension method – mean (sd) weighted accuracy fitting to 100 random μ^r

Subsets of vertices defining particular classes

n	add	sym	add+sym	2add
3	0.133 (0.071)	0.105 (0.045)	0.037 (0.019)	**0.023** (0.020)
4	0.160 (0.069)	0.141 (0.048)	0.071 (0.023)	**0.045** (0.020)
5	0.181 (0.065)	0.141 (0.040)	0.107 (0.026)	**0.082** (0.032)
6	0.223 (0.066)	0.156 (0.038)	0.140 (0.034)	0.115 (0.034)
7	0.264 (0.063)	0.147 (0.035)	0.143 (0.034)	0.151 (0.033)
8	0.328 (0.075)	0.157 (0.029)	0.156 (0.029)	0.185 (0.044)

Subsets of vertices chosen randomly

	random vertices method 1		random vertices method 2	
n	$\|U\| = n$	$\|U\| = n^2$	$\|U\| = n$	$\|U\| = n^2$
3	0.123 (0.074)	0.035 (0.023)	0.176 (0.082)	0.073 (0.047)
4	0.193 (0.089)	0.061 (0.026)	0.179 (0.073)	0.093 (0.041)
5	0.292 (0.126)	0.142 (0.049)	0.190 (0.067)	0.114 (0.036)
6	0.479 (0.177)	0.265 (0.080)	0.166 (0.045)	**0.101** (0.027)
7	0.602 (0.207)	0.340 (0.090)	0.152 (0.036)	**0.097** (0.021)
8	0.784 (0.295)	0.424 (0.117)	0.147 (0.030)	**0.093** (0.019)

dency of the linear extension method to generate fuzzy measures that tend more toward minimum-like behaviour, and hence which have less variability between subsets of the same cardinality. These antibuoyant fuzzy measures are therefore closer to being symmetric, which is why both sym and add+sym exhibit good performance for larger n. The performance of the rand2 set on the other hand can be put down to the random vertices being generated in the same way as the random fuzzy measure, which makes them more likely to exhibit similar overall behaviour.

When it comes to the second round of experiments where the random fuzzy measures are augmented by mixing with the symmetric and additive fuzzy measure, the add+sym set produced the best overall results, achieving slightly better results than the sym set. This is not surprising since, as mentioned, $\mathbf{U}$ in the case of both sym and add+sym includes the additive symmetric fuzzy measure. It is worthy to observe here that 2add now has better performance compared to the rand2 method (although somewhat worse than the sym sets). This further supports the presumption with the first set of 100 experiments that the rand2 performance was due to the sim-

Table 6: Antibuoyant fuzzy measures generated using augmented method – mean (sd) weighted accuracy fitting to 100 random μ^r

Subsets of vertices defining particular classes

n	add	sym	add+sym	2add
3	0.235 (0.096)	0.079 (0.038)	**0.027** (0.014)	0.041 (0.032)
4	0.321 (0.137)	0.109 (0.043)	0.054 (0.020)	0.100 (0.057)
5	0.377 (0.189)	0.106 (0.035)	**0.079** (0.023)	0.166 (0.092)
6	0.427 (0.213)	0.113 (0.033)	**0.096** (0.027)	0.241 (0.108)
7	0.474 (0.256)	0.111 (0.033)	**0.105** (0.031)	0.306 (0.153)
8	0.633 (0.342)	0.113 (0.031)	**0.112** (0.031)	0.456 (0.222)

Subsets of vertices chosen randomly

	random vertices method 1		random vertices method 2	
n	$\|U\| = n$	$\|U\| = n^2$	$\|U\| = n$	$\|U\| = n^2$
3	0.123 (0.072)	0.045 (0.030)	0.194 (0.091)	0.070 (0.052)
4	0.145 (0.066)	**0.051** (0.019)	0.243 (0.135)	0.113 (0.064)
5	0.209 (0.089)	0.095 (0.030)	0.364 (0.161)	0.246 (0.125)
6	0.276 (0.092)	0.164 (0.046)	0.493 (0.188)	0.414 (0.174)
7	0.394 (0.130)	0.231 (0.066)	0.593 (0.285)	0.516 (0.268)
8	0.516 (0.179)	0.306 (0.078)	0.798 (0.383)	0.726 (0.366)

ilar way in which the fuzzy measures included were generated. The rand1 has improved performance in this case, generating vertices that are likely to have more deviation between subsets of the same cardinality and less-close to symmetric fuzzy measures.

Before moving to the second set of experiments, it is worth noting that even the seemingly high values for $n = 8$ for some choices of $\mathbf{U}$ should be interpreted in context. Recall that these are weighted sums across the entire fuzzy measure, and so even the worst performance of 0.798 only reflects an average difference for each cardinality to be around 0.1, or alternatively can be compared to the difference between the minimum fuzzy measure and the additive symmetric fuzzy measure, which would be 3.5.

5.2. Fitting experiments 2: fitting to random data sets

The aim of the second set of experiments is to compare the performance for each choice of $\mathbf{U}$ against the general fitting approach with constraints. Using the latter approach would clearly always be able to identify the randomly generated μ^r perfectly when fitting direct to the fuzzy measure values, and so for the comparison we use the same μ^r from the previous experiments to generate random datasets and then add noise.

First, n-dimensional vectors $\mathbf{x}^m, m = 1, \ldots, 100$, are randomly generated from a multivariate uniform distribution on $[0, 1]^n$, and then random noise δ with mean 0 and standard deviation 0.05 is added to the outputs so that $y^m = C_{\mu^r}(\mathbf{x}^m) + \delta^m$. We generate 2×2^n data observations for each dataset so that there are at least two times as many data points as variables, and when adding noise, we limit y^m so that it lies between 0 and 1. We then compare the fitting accuracy of the vertex-based methods with the general fitting subject to constraints approach. Here the optimisation problem being solved for the different $\mathbf{U}$ is

$$\begin{aligned} \text{Minimise} \quad & \frac{1}{100} \sum_{m=1}^{100} \left| y^m - \sum_{i=1}^{u} c_i C_{\mu^i}(\mathbf{x}^m) \right| \\ \text{s.t.} \quad & \sum_{i=1}^{u} c_i = 1 \\ & c_i \geq 0, \ \forall i. \end{aligned} \tag{5}$$

The general fitting approach fits with respect to the fuzzy measure values as decision variables with constraints of the form given in (2). The objective in both cases is minimising the sum of positive and negative residual components $r_m^+ + r_m^-$. The results for the same fuzzy measures (generated using the linear extension as well as the augmented version) are shown in Tables 7 and 8. The best performing of the 8 choices for $\mathbf{U}$ are bolded, and the results for the general fitting with constraints approach are reproduced as the last column in both tables as the baseline measure. Results in this case represent the sum of absolute residuals, so can be interpreted against the number of data which is 2×2^n (although note that the standard deviation is based on the distribution of the total objective sums).

In terms of comparing the different choices of $\mathbf{U}$, the results and differences for these experiments are consistent with those when fitting directly to the fuzzy measures. The 2add analogue works best for lower values of n when the linear extension method is used and add+sym gives good performance in both cases (although it has an advantage in the augmented case).

It can be observed that the best non-random methods are within about 10% of the error achieved by general fitting with constraints. Thus, depending on the reduction of variables desired, choices of $\mathbf{U}$ that either combine

Table 7: Antibuoyant fuzzy measures generated using the linear extension method – mean (sd) accuracy fitting to 100 random μ^r

	Subsets of vertices defining particular classes				
n	add	sym	add+sym	2add	general
3	0.735 (0.218)	0.698 (0.200)	0.546 (0.117)	**0.533** (0.123)	0.497 (0.121)
4	1.484 (0.291)	1.488 (0.288)	1.210 (0.202)	**1.156** (0.193)	1.064 (0.184)
5	2.837 (0.466)	2.840 (0.318)	2.515 (0.252)	**2.395** (0.261)	2.181 (0.227)
6	5.997 (0.776)	5.740 (0.452)	5.290 (0.379)	**5.042** (0.378)	4.599 (0.351)
7	12.368 (1.412)	11.205 (0.575)	10.635 (0.544)	10.421 (0.589)	9.404 (0.435)
8	25.933 (2.599)	21.909 (0.944)	21.166 (0.808)	21.563 (1.317)	19.023 (0.583)

	Subsets of vertices chosen randomly				
	random vertices method 1		random vertices method 2		general
n	$\lvert U\rvert = n$	$\lvert U\rvert = n^2$	$\lvert U\rvert = n$	$\lvert U\rvert = n^2$	
3	0.758 (0.226)	0.554 (0.123)	0.833 (0.261)	0.601 (0.155)	0.497 (0.121)
4	1.677 (0.495)	1.197 (0.214)	1.649 (0.343)	1.275 (0.221)	1.064 (0.184)
5	4.090 (1.209)	2.710 (0.376)	2.968 (0.511)	2.552 (0.296)	2.181 (0.227)
6	9.886 (3.010)	6.517 (0.976)	5.769 (0.584)	5.173 (0.406)	4.599 (0.351)
7	21.048 (6.481)	13.971 (2.108)	11.007 (0.550)	**10.311** (0.520)	9.404 (0.435)
8	47.469 (13.042)	31.053 (5.505)	21.571 (1.119)	**20.330** (0.728)	19.023 (0.583)

the sym and add vertices, or even the sym and 2add vertices, seem quite capable of giving good modelling performance. This is despite requiring far fewer variables and only requiring a single monotonicity constraint in fitting.

6. Discussion

Addressing the complexity of the fitting problem when modelling with fuzzy measures and the Choquet integral is a key challenge for practical implementation and broader uptake in the research community. A number of useful simplifications have been introduced to reduce the number of variables required and complexity of the fitting problem, however if a new concept such as antibuoyancy cannot be expressed or contained to these reduced and simplified representations, it puts us back at square one. While requiring even more fitting constraints in the general approach, in the above we have demonstrated that there are useful strategies allowing the modelling and analysis of data with antibuoyant fuzzy measures. While such fuzzy measures seem to form a relatively narrow sub-class, we know that the antibuoyancy property,

Table 8: Antibuoyant fuzzy measures generated using augmented method – mean (sd) weighted accuracy fitting to 100 random μ^r

Subsets of vertices defining particular classes

n	add	sym	add+sym	2add	general
3	0.922 (0.313)	0.641 (0.158)	**0.532** (0.112)	0.540 (0.142)	0.486 (0.122)
4	2.028 (0.628)	1.313 (0.217)	1.134 (0.156)	1.216 (0.222)	1.016 (0.156)
5	4.362 (1.476)	2.707 (0.297)	**2.444** (0.268)	2.775 (0.609)	2.192 (0.266)
6	8.568 (2.950)	5.388 (0.432)	**5.040** (0.363)	5.813 (1.119)	4.582 (0.336)
7	17.111 (6.352)	10.802 (0.629)	**10.333** (0.533)	12.292 (2.823)	9.565 (0.516)
8	38.872 (16.383)	21.410 (0.848)	**20.691** (0.726)	28.193 (9.023)	19.473 (0.728)

Subsets of vertices chosen randomly

	random vertices method 1		random vertices method 2		general
n	$\|U\| = n$	$\|U\| = n^2$	$\|U\| = n$	$\|U\| = n^2$	
3	0.697 (0.181)	0.553 (0.136)	0.834 (0.296)	0.607 (0.159)	0.486 (0.122)
4	1.437 (0.323)	**1.112** (0.158)	1.646 (0.454)	1.325 (0.255)	1.016 (0.156)
5	3.073 (0.716)	2.471 (0.296)	4.205 (1.411)	3.379 (1.047)	2.192 (0.266)
6	6.442 (1.515)	5.247 (0.520)	9.228 (3.039)	7.902 (2.492)	4.582 (0.336)
7	13.677 (3.583)	11.376 (1.709)	20.072 (7.925)	17.718 (6.573)	9.565 (0.516)
8	28.466 (9.303)	22.557 (2.659)	46.832 (19.402)	42.615 (18.205)	19.473 (0.728)

or equivalently the Pigou–Dalton principle, make these very attractive for a wide variety of applications in domains including economics, ecology, and even bibliometrics [4, 16].

The results of experiments conducted here have shown that, in terms of fitting performance, the symmetric class of antibuoyant fuzzy measures are already capable of fitting to datasets where antibuoyancy is assumed, however whether stemming from desirability or practical observation, there may be a need for non-symmetric behaviour to be incorporated. For this, we have proposed the use of vertices analogous to those that define the sets of additive and 2-additive fuzzy measures, respectively. Allowing convex combination with these vertices allows a richer behaviour where not only the size of the inputs but also the importance of the variables to which they pertain is taken into account.

Depending on how much it is desired that the size of the fitting problem be reduced, we would hence recommend to firstly incorporate the symmetric vertices, and then incorporate more of the vertices from our other choices of $\mathbf{U}$. For $n = 8$, we note that even combining the sym and 2add sets would still reduce the number of unknown variables from 254 to 72 and only require one monotonicity constraint.

7. Conclusion

We examined three problems associated with the class of antibuoyant fuzzy measures that are related to the Pigou–Dalton progressive transfers principle and which represent a subset of supermodular fuzzy measures. Firstly, we established a subset of extreme points of antibuoyant fuzzy measures, analogous to the $\{0, 1\}$-fuzzy measures, which can be used to define a sufficiently rich class of antibuoyant measures through convex combinations. Secondly, we proposed three methods for efficient random generation of antibuoyant fuzzy measures of that subclass, which will be useful in performing simulation studies involving non-additive but convex aggregation of inputs. Thirdly, we also formulated and proposed algorithms for fitting antibuoyant fuzzy measures to either a set of empirical values, or to a more general fuzzy measure. The latter serves the purpose of finding the best approximation of a given fuzzy measure by an element of a particular subclass, which provides a reasonable simplification that satisfies some stipulated conditions. In the context of supermodularity and antibuoyancy, these conditions represent types of convexity, and convexity plays an extremely important role in optimisation problems. In particular, optimising a piecewise linear objective expressed as the Choquet integral (and hence accounting for interactions between the variables), general fuzzy measures lead to a difficult multiextremal problem, supermodular fuzzy measures lead to a convex problem that can be expressed through (a large) linear programming problem, while antibuoyant fuzzy measures lead to a much smaller linear program that allows solution in polynomial time. Of course, by duality we obtain the corresponding results for buoyant fuzzy measures.

Acknowledgements

The work was supported by the Australian Research Council Discovery project DP210100227. The authors would also all like to acknowledge the collegial support and friendship of Radko Mesiar throughout our research careers. Happy 70th Birthday!

References

[1] G. Beliakov. Knapsack problems with dependencies through non-additive measures and Choquet integral. *European Journal of Operational Research*, https://doi.org/10.1016/j.ejor.2021.11.004, 2021.

[2] G. Beliakov, F.J. Cabrerizo, E. Herrera-Viedma, and J-Z. Wu. Random generation of k-interactive capacities. *Fuzzy Sets and Systems*, 430:48–55, 2022.

[3] G. Beliakov, M. Gagolewski, and S. James. Penalty-based and other representations of economic inequality. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 24(Suppl.1):1–23, 2016.

[4] G. Beliakov and S. James. Citation based journal ranks: The use of fuzzy measures. *Fuzzy Sets and Systems*, 167:101–119, 2011.

[5] G. Beliakov and S. James. Choquet integral optimisation with constraints and the buoyancy property for fuzzy measures. *Information Sciences*, 578:22–36, 2021.

[6] G. Beliakov and S. James. Choquet integral based measures of economic welfare and species diversity. *International Journal of Intelligent Systems*, 2021. (in press) http://doi.org/10.1002/int.22609.

[7] G. Beliakov, S. James, and G. Li. Learning Choquet integral-based metrics in semi-supervised classification. *IEEE Transactions on Fuzzy Systems*, 19:562–574, 2011.

[8] G. Beliakov, S. James, and J-Z. Wu. *Discrete Fuzzy Measures: Computational Aspects.* Springer, Berlin, Heidelberg, 2019.

[9] G. Beliakov and J.-Z. Wu. Random generation of capacities and its application in comprehensive decision aiding. *Information Sciences*, 577:424–435, 2021.

[10] G. Choquet. Theory of capacities. *Ann. Inst. Fourier*, 5:1953–1954, 1953.

[11] E. F. Combarro, J. Hurtado de Saracho, and I. Díaz. Minimals plus: An improved algorithm for the random generation of linear extensions of partially ordered sets. *Information Sciences*, 501:50–67, 2019.

[12] E. F. Combarro, I. Díaz, and P. Miranda. On random generation of fuzzy measures. *Fuzzy Sets and Systems*, 228:64–77, 2013.

[13] E. F. Combarro and P. Miranda. On the structure of the k-additive fuzzy measures. *Fuzzy Sets and Systems*, 161:2314–2327, 2010.

[14] H. Dalton. The measurement of the inequality of incomes. *The Economic Journal*, 30:348–361, 1920.

[15] H. E. de Oliveira, L. T. Duarte, and J. M. T. Romano. Identification of the Choquet integral parameters in the interaction index domain by means of sparse modeling. *Expert Systems with Applications*, 187:115874, 2022.

[16] M. Gagolewski and R. Mesiar. Monotone measures and universal integrals in a uniform framework for the scientific impact assessment problem. *Information Sciences*, 263:166–174, 2014.

[17] M. Grabisch. k–Order additive discrete fuzzy measures and their representation. *Fuzzy Sets and Systems*, 92:167–189, 1997.

[18] M. Grabisch. *Set Functions, Games and Capacities in Decision Making.* Springer, Berlin, New York, 2016.

[19] M. Grabisch, I. Kojadinovic, and P. Meyer. A review of methods for capacity identification in choquet integral based multi-attribute utility theory: Applications of the Kappalab R package. *European J. of Operational Research*, 186(2):766–785, 2008.

[20] M. Grabisch, J.-L. Marichal, and M. Roubens. Equivalent representations of set functions. *Mathematics of Operations Research*, 25(2):157–178, 2000.

[21] A. Honda and S. James. Parameter learning and applications of the inclusion-exclusion integral for data fusion and analysis. *Information Fusion*, 56:28–38, 2020.

[22] M. A. Islam, D. T. Anderson, A. J. Pinar, T. C. Havens, G. Scott, and J. M. Keller. Enabling explainable fusion in deep learning with fuzzy integral neural networks. *IEEE Transactions on Fuzzy Systems*, 28(7):1291–1300, 2020.

[23] G. Lucca, G. P. Dimuro, J. Fernández, H. Bustince, B. Bedregal, and J. A. Sanz. Improving the performance of fuzzy rule-based classification systems based on a nonaveraging generalization of CC-integrals named $c_{F_1F_2}$-integrals. *IEEE Transactions on Fuzzy Systems*, 27(1):124–134, 2019.

[24] R. Mesiar. Generalizations of k–order additive discrete fuzzy measures. *Fuzzy Sets and Systems*, 102:423–428, 1999.

[25] P. Miranda, E. F. Combarro, and P. Gil. Extreme points of some families of non-additive measures. *European Journal of Operational Research*, 174:1865–1884, 2006.

[26] P. Miranda, M. Grabisch, and P. Gil. p-symmetric fuzzy measure. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 10(supp01):105–123, 2002.

[27] H. Moulin. *Fair Division and Collective Welfare*. Cambridge, Massachusetts: MIT Press, 2004.

[28] B. J. Murray, M. A. Islam, A. J. Pinar, D. T. Anderson, G. J. Scott, T. C. Havens, and J. M. Keller. Explainable AI for the Choquet integral. *IEEE Transactions on Emerging Topics in Computational Intelligence*, 5(4):520 – 529, 2021.

[29] D. Radojević. The logical representation of the discrete Choquet integral. *Belgian Journal of Operations Research, Statistics and Computer Science*, 38:67–89, 1998.

[30] A. F. Tehrani. On correlated information for learning predictive models under the Choquet integral. *Expert Systems*, 38(8):e12777, 2021.

[31] H. Tuomisto. An updated consumer's guide to evenness and related indices. *Oikos*, 121:1203 – 1218, 2012.

[32] R. R. Yager. On ordered weighted averaging aggregation operators in multicriteria decision making. *IEEE Transactions on Systems, Man and Cybernetics*, 18:183–190, 1988.

[33] R. R. Yager. Families of OWA operators. *Fuzzy Sets and Systems*, 59:125–148, 1993.