Centroids of the core of exact capacities: a comparative study

Enrique Miranda[1] · **Ignacio Montes[1]**

Abstract
Capacities are a common tool in decision making. Each capacity determines a core, which is a polytope formed by additive measures. The problem of eliciting a single probability from the core is interesting in a number of fields: in coalitional game theory for selecting a fair way of splitting the wealth between the players, in the transferable belief model from evidence theory or for transforming a second order into a first order model. In this paper, we study this problem when the goal is to determine the centroid of the core of a capacity, and we compare four approaches: the Shapley value, the average of the extreme points, the incenter with respect to the total variation distance and the limit of a procedure of uniform contraction. We show that these four centroids do not coincide in general, we give some sufficient conditions for their equality, and we analyse their axiomatic properties. We also discuss how to define a notion of centrality measure indicating the degree of centrality of an additive measure in the core. Finally, we also analyse these four centroids in the more general context of imprecise probabilities.

Keywords Capacities · Game solutions · Cores · Supermodularity · Imprecise probabilities

1 Introduction

A problem that naturally arises in many branches of operations research, such as decision making (Huntley & Troffaes, 2012; Keith & Ahner, 2021; Troffaes, 2007) or expected utility theory (Gilboa & Schmeidler, 1989; Klibanoff et al., 2005; Sarin & Wakker, 1992), is that of determining the probability measure modelling the underlying uncertainty. Due to a number of factors (missing data, conflicting sources of information, etc.), it is sometimes difficult, or even impossible, to elicit such probability measure with minimal guarantees. A possible approach in those cases is to consider instead a *capacity* (Grabisch, 2016), that in turns deter-

---

Enrique Miranda and I. Montes have contributed equally to this work.

✉ Enrique Miranda
mirandaenrique@uniovi.es

Ignacio Montes
imontes@uniovi.es

[1] Department of Statistics and O.R., University of Oviedo, Federico García Lorca, 33007 Oviedo, Asturias, Spain

mines a polytope of associated additive measures: its *core*. When the capacity is normalised, the core is formed by probability measures.

Capacites appear in many different contexts and with different interpretations: (i) in *decision making*, the values of the capacity measure the beliefs supporting the occurrence of each event; (ii) in *coalitional game theory*, each event represents a coalition of players, the value of the capacity is interpreted as the minimum reward guaranteed by the coalition, and the core contains the distributions of the rewards compatible with these constraints; and (iii) in *imprecise probability theory* (Augustin et al., 2014), a capacity gives lower bounds for the real but unknown values of the probability measure underlying the experiment.[1] The core, also called credal set (Levi, 1980), is formed then by the probability measures that are candidates for being the unknown probability measure. Among the many applications of capacities, we refer for instance to Grabisch (2013) and Shapley (1953) for some applications in game theory, and to Angilella et al. (2016) and Destercke (2017) in ordinal classification.

In any of these contexts, the problem of selecting a probability measure from the core of the normalised capacity is quite common. For instance, within coalitional game theory we may consider the *solution* of a game as a way to fairly divide the wealth it represents among the players; within imprecise probability theory, we may also consider transformations between imprecise and precise probability models (Klir & Parviz, 1992; Smets, 2005); in addition, we may also look for the element of the core maximising the entropy (Abellán & Moral, 2003; Jaffray, 1995) or establish procedures for assigning relevance degrees to the different features in machine learning problems (Kumar et al., 2020; Lundberg & Lee, 2017).

Our goal here is also to select a probability measure from the core of a normalised capacity, but the interpretation of the output of the process is different to the cases mentioned above: we seek to determine the *center* of the core of the capacity, similarly to the point with greatest data depth (Cascos, 2009; Tukey, 1975) in a data cloud; thus, our final result should be an element in the interior of the core, whenever the latter is non-empty. This already rules out methods based on maximising the entropy or minimising the Kullback Leibler divergence. Pursuing this objective, in this paper we analyse four centers: (i) the Shapley value, that first appeared in coalitional game theory (Shapley, 1953); (ii) the average of the extreme points of the core; (iii) the incenter with respect to a distance (in our case, the total variation distance, due to its good properties within the imprecise probability framework (Montes et al., 2020b)); and a center that is obtained following a procedure of uniform contraction of the core.

The investigation of the notion of center of a core leads naturally to that of a measure of centrality with respect to a given set of probability measures. We propose an axiomatic definition and analyse several examples, both based on a choice of a centroid or not.

The remainder of the paper is organised as follows. After introducing some preliminary notions in Sect. 2, in Sect. 3 we discuss the four possible notions of centroids mentioned above and study the relationships between them. In Sect. 4, we make a further comparison in terms of the axiomatic properties they satisfy and in Sect. 5 we discuss the notion of centrality measure. Finally, in Sect. 6 we show that the centroids can be defined in the more general context of coherent lower previsions. Some additional comments are given in Sect. 7. To ease the reading, proofs have been relegated to the Appendix.

A preliminary version of this paper was presented at ECSQARU 2021 Conference (Miranda & Montes, 2021). This extended version contains a deeper discussion of the four centroids, additional results, proofs and other comments stemming from the discussions carried out at the conference.

---

[1] For this reason, capacites are called *lower probabilities* in that framework.

2 Preliminary concepts

Let us introduce the main concepts we shall use in this paper. We refer to Grabisch (2016) for more details.

Consider a finite possibility space $\mathcal{X} = \{x_1, \ldots, x_n\}$, let $\mathbb{P}(\mathcal{X})$ be the set of all the probability measures on $\mathcal{X}$ and let $\mathbb{P}^*(\mathcal{X})$ be the probability measures assigning strictly positive probabilities to all the non-empty events.

A *capacity* is a function $\mu : \mathcal{P}(\mathcal{X}) \to \mathbb{R}$ that is *monotone*: $A \subseteq B$ implies $\mu(A) \le \mu(B)$ and satisfies $\mu(\emptyset) = 0$. A capacity is *normalised* when in addition $\mu(\mathcal{X}) = 1$. Throughout this paper, we will always consider normalised capacities. The conjugate of a capacity μ, denoted by $\bar{\mu}$, is defined by $\bar{\mu}(A) = 1 - \mu(A^c)$ for any $A \subseteq \mathcal{X}$. Also, a capacity μ determines the *core*, defined by:

$$\text{core}(\mu) = \{P \in \mathbb{P}(\mathcal{X}) \mid P(A) \ge \mu(A) \, \forall A \subseteq \mathcal{X}\}.$$

When the core is non-empty, the capacity is called *balanced* and it holds that $\mu \le \bar{\mu}$. A balanced capacity is called *exact* when $\mu(A) = \min_{P \in \text{core}(\mu)} P(A)$ for every $A \subseteq \mathcal{X}$.

As we mentioned in the introduction, capacities can be interpreted in many different ways. In this paper, we will focus mostly on the decision making or game theoretic interpretations (cases (i) and (ii) from Sect. 1). The interpretation as imprecise probability models (case (iii)) will be explored in more detail in Sect. 6.

The core of a capacity is a closed and convex subset of $\mathbb{P}(\mathcal{X})$, and since it is determined by a finite number of restrictions, it is a polytope that can be characterised by a finite number of *extreme points*. Recall that $P \in \text{core}(\mu)$ is an extreme point of the core if $P = \lambda P_1 + (1 - \lambda) P_2$ for some $\lambda \in (0, 1)$ and $P_1, P_2 \in \text{core}(\mu)$, then $P_1 = P_2 = P$. The set of extreme points shall be denoted by $\text{ext}(\text{core}(\mu))$.

A property that an exact capacity may satisfy is that of *supermodularity*, also called *convexity* in coalitional game theory, which means that:

$$\mu(A \cup B) + \mu(A \cap B) \ge \mu(A) + \mu(B) \quad \forall A, B \subseteq \mathcal{X}.$$

When the capacity is supermodular, the set of extreme points of the core is given by $\{P_\sigma \mid \sigma \in S_n\}$ (Shapley, 1971), where S_n denotes the set of permutations of $\{1, \ldots, n\}$, and given $\sigma \in S_n$, P_σ is determined by:

$$P_\sigma\left(\{x_{\sigma(1)}, \ldots, x_{\sigma(i)}\}\right) = \mu\left(\{x_{\sigma(1)}, \ldots, x_{\sigma(i)}\}\right) \quad \forall i = 1, \ldots, n. \tag{1}$$

A capacity can be equivalently represented in terms of its *Möbius inverse*, which is the function $m : \mathcal{P}(\mathcal{X}) \to \mathbb{R}$ given by:

$$m(A) = \sum_{B \subseteq A} (-1)^{|A \setminus B|} \mu(B) \quad \forall A \subseteq \mathcal{X}.$$

From the Möbius inverse we can retrieve the capacity by means of:

$$\mu(A) = \sum_{B \subseteq A} m(B) \quad \forall A \subseteq \mathcal{X}.$$

The Möbius inverse m takes values in $\mathbb{R}$. When it is non-negative for every event, the capacity μ is called *belief function* and its conjugate $\bar{\mu}$ is called *plausibility function*, creating a bridge with Evidence Theory (Shafer, 1976). Any belief function is a supermodular capacity.

3 Center points of an exact capacity

Next, we introduce the different notions of centroid of the core of a capacity we shall compare in this paper. We shall consider four possibilities: the Shapley value, the average of the extreme points, the incenter with respect to the total variation distance and the contraction centroid. In general, we shall use the notation Φ^μ to denote a centroid of a capacity μ. Moreover, we shall assume throughout that the capacity μ is exact.

3.1 The Shapley value

One of the most popular notions of centroid of a capacity is the *Shapley value*. It was introduced by Shapley (1953, 1971) in the framework of coalitional game theory, as a "fair" procedure to distribute some wealth between the players. Later on, it was rediscovered in the context of non-additive measures (Dubois & Prade, 1980) and popularised by Smets as the *pignistic transformation* of a belief function (Smets & Kennes, 1994).

Definition 1 Given an exact capacity μ, its Shapley value is defined as the probability measure associated with the following distribution:

$$\Phi_1^\mu(\{x\}) = \sum_{A\,|\,x\notin A} \frac{|\,A\,|!(n-|\,A\,|-1)!}{n!}\big(\mu(A\cup\{x\}) - \mu(A)\big) \quad \forall x\in\mathcal{X}. \tag{2}$$

When μ is a belief function with Möbius inverse m, it was proven in Smets (2005) that Φ_1^μ can be equivalently computed as

$$\Phi_1^\mu(\{x\}) = \sum_{A\,|\,x\in A} \frac{m(A)}{|\,A\,|} \quad \forall x\in\mathcal{X}. \tag{3}$$

More generally, when μ is supermodular (and in particular when it is a belief function), the extreme points of $\text{core}(\mu)$ are given by Eq. (1), and the Shapley value can be computed as

$$\Phi_1^\mu(\{x\}) = \frac{1}{n!}\sum_{\sigma\in S_n} P_\sigma(\{x\}) \quad \forall x\in\mathcal{X}. \tag{4}$$

Even if the expressions in Eqs. (3) and (4) were only established for belief functions and supermodular capacities, respectively, it follows from basic combinatorial analysis that they can be extended to arbitrary capacities:[2]

Proposition 1 *Let μ be an exact capacity with Möbius inverse m, and let Φ_1^μ be given by Eq. (2). Then*

$$\Phi_1^\mu(\{x\}) = \sum_{A\,|\,x\in A} \frac{m(A)}{|\,A\,|} = \sum_{\sigma\in S_n} \frac{P_\sigma(\{x\})}{n!} \quad \forall x\in\mathcal{X}, \tag{5}$$

where P_σ is given by Eq. (1).

It is worth remarking that, even if $m(A)$ may be negative in some events A, Proposition 1 implies that the aggregation by means of Eq. (5) always produces a non-negative value.

While the Shapley value seems like a reasonable choice as a central point, it has one important drawback: it is only guaranteed to belong to the core of μ (i.e., we can only assure

---

[2] See also (Grabisch 2016, Ex.3.6.1) for some comments in this respect using *interaction indexes* from game theory.

that $\Phi_1^\mu \geq \mu$) when the capacity μ is supermodular. In those cases, it follows immediately from Eq. (4) and the fact that P_σ dominates μ for every permutation $\sigma \in S_n$.

More generally, when the exact capacity μ is not supermodular, we can only assure that Φ_1^μ dominates μ for small cardinalities, as showed by Baroni and Vicig (2005, Prop.5). We refer to Miranda and Montes (2018) for a study of the connection between the Shapley value and the core of the capacity.

3.2 Average of the extreme points

The second possibility we consider in this paper is the average of the extreme points of the core of the capacity:

Definition 2 Let μ be an exact capacity, and denote by $P_1, \ldots, P_k$ the extreme points of its core. The *average of the extreme points*, also called *vertex centroid* (Elbassioni & Tiway, 2012), is defined as

$$\Phi_2^\mu(\{x\}) = \frac{1}{k} \sum_{i=1}^{k} P_i(\{x\}) \quad \forall x \in \mathcal{X}. \tag{6}$$

It follows from Definition 2 that Φ_2^μ always belongs to core(μ); considering the comments in the previous section, this implies that it need not coincide with the Shapley value when supermodularity is not satisfied. As we shall see later on, they need not coincide either even if the capacity is supermodular: while the set of extreme points of the core is $\{P_\sigma \mid \sigma \in S_n\}$, a key difference is that in the computation of Shapley value in Eq. (4) we are allowing for repetitions of the same extreme point, while in Definition 2 we do not.

It is also important to clarify that the average of the extreme points does not generally coincide with the *center of gravity* of the core, defined as the expectation of the set over a uniform probability distribution.

Example 1 Consider a 3-element possibility space and the exact capacity μ given by:

A	$\{x_1\}$	$\{x_2\}$	$\{x_3\}$	$\{x_1, x_2\}$	$\{x_1, x_3\}$	$\{x_2, x_3\}$	$\mathcal{X}$
$\mu(A)$	0	0	0	0	0	0.5	1

This capacity is supermodular[3] and the extreme points of core(μ) are given by:

	x_1	x_2	x_3
P_1	0	0	1
P_2	0	1	0

	x_1	x_2	x_3
P_3	1/2	1/2	0
P_4	1/2	0	1/2

The average of the extreme points is given by:

$$\Phi_2^\mu(\{x_1\}) = 0.25, \quad \Phi_2^\mu(\{x_2\}) = \Phi_2^\mu(\{x_3\}) = 0.375,$$

while the expectation over the core with respect to the uniform distribution E is:

$$E(\{x_1\}) = 2/9, \quad E(\{x_2\}) = E(\{x_3\}) = 7/18.$$

Therefore, both concepts do not coincide in general. Intuitively, if we assume that the mass is uniformly distributed over core(μ), we can see in Fig. 1 that there is more mass for values of x_1 closer to 0 than closer to 1/2, whence $E(\{x_1\})$ must be smaller than $\Phi_2^\mu(\{x_1\}) = 1/4$.

---

[3] It is known that any exact capacity in a 3-element possibility space is supermodular too.
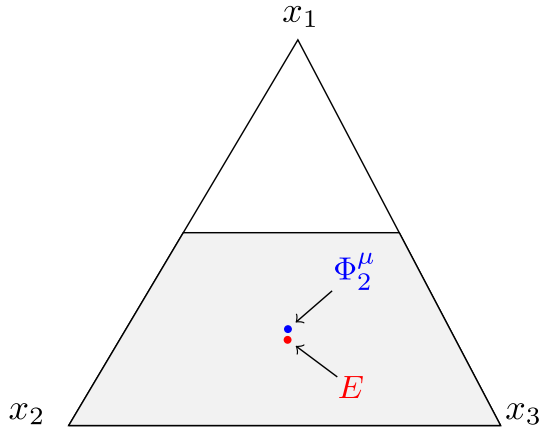
In fact, this example also shows that the center of gravity does not coincide with the Shapley value either, even in this case where μ is supermodular, because:

$$\Phi_1^\mu(\{x_1\}) = 1/6, \quad \Phi_1^\mu(\{x_2\}) = \Phi_1^\mu(\{x_3\}) = 5/12. \quad \blacklozenge$$

While the center of gravity has the advantage of being applicable over any closed and convex set in $\mathbb{P}(\mathcal{X})$, and not only on polytopes, it also has the drawback of being computationally more expensive (see for example (Elbassioni & Tiway, 2012)). For this reason, we have left this approach out of our study.

3.3 Incenter

Next we consider the *incenter*, that corresponds to the center (or centers) of the largest balls included in the interior of the core of the capacity.[4] To make this notion precise, we must specify the distance under which the balls are defined. In this respect, there are several possibilities, such as the Euclidean or the L_1 distances for example, or even the Kullback-Leibler divergence. We have considered in this paper the *total variation distance* (Levin et al., 2009), which is the one associated with the supremum norm:

$$d_{TV}(P, Q) = \max_{A \subseteq \mathcal{X}} \mid P(A) - Q(A) \mid \qquad \forall P, Q \in \mathbb{P}(\mathcal{X}).$$

Our choice of this distance is due to the fact that the closed balls it induces are always polytopes, unlike the case of the Euclidean distance or the Kullback-Leibler divergence, and those closed balls correspond to exact capacities satisfying supermodularity (Montes et al. 2020b, Sec.2). Although the L_1 distance also induces a polytope, it does not correspond to the core of an exact capacity and our analysis in Destercke et al. (2022), Montes et al. (2020b) shows that its use is rather complex. In what follows, for the sake of notational simplicity, the total variation distance will be simply denoted by d. Moreover, we shall denote

$$B_c^\alpha(P_0) = \{P \in \mathbb{P}(\mathcal{X}) \mid d(P, P_0) \le \alpha\}, \quad B_o^\alpha(P_0) = \{P \in \mathbb{P}(\mathcal{X}) \mid d(P, P_0) < \alpha\}$$

the closed and open balls centered on P_0 and with radius α, respectively.

---

[4] Here we are identifying any element of the core with its associated probability mass function; then $\mathbb{P}(\mathcal{X})$ can be regarded as a subset of the $|\mathcal{X}|$-dimensional Euclidean space $\mathbb{R}^{|\mathcal{X}|}$ and the interior of core(μ) is understood with respect to the restriction of the usual topology to $\mathbb{P}(\mathcal{X})$.

This leads us to the following definition:

Definition 3 Let μ be an exact capacity. The *incenter radius* of core(μ) is defined as

$$\alpha_I = \sup \left\{ \alpha > 0 \mid \exists P_0 \in \text{core}(\mu) \text{ such that } B_o^\alpha(P_0) \subseteq \text{core}(\mu) \cap \mathbb{P}^*(\mathcal{X}) \right\}. \qquad (7)$$

Any $P_0 \in \text{core}(\mu)$ such that $B_o^{\alpha_I}(P_0) \subseteq \text{core}(\mu)$ is called an *incenter* of μ.

It may be surprising to see that in Eq. (7) we are requiring the inclusion of the open ball $B_o^\alpha(P_0)$ into the intersection core(μ) $\cap \mathbb{P}^*(\mathcal{X})$. The reason is that if we simply require $B_o^\alpha(P_0) \subseteq \text{core}(\mu)$ then we may obtain centers that are in the boundary of the core, something in our view counterintuitive. This is illustrated in the following example.

Example 2 Let $\mathcal{X} = \{x_1, x_2, x_3\}$ and consider the exact capacity μ given by:

A	$\{x_1\}$	$\{x_2\}$	$\{x_3\}$	$\{x_1, x_2\}$	$\{x_1, x_3\}$	$\{x_2, x_3\}$	$\mathcal{X}$
$\mu(A)$	0.5	0.1	0	0.75	0.5	0.1	1

The value α_I determined by Eq. (7) is $\alpha_I = 0.125$, and any convex combination of the probability measures Q_1 and Q_2 given by:

$$Q_1(\{x_1\}) = 0.65, \quad Q_1(\{x_2\}) = 0.225, \quad Q_1(\{x_3\}) = 0.125,$$
$$Q_2(\{x_1\}) = 0.625, \quad Q_2(\{x_2\}) = 0.25, \quad Q_2(\{x_3\}) = 0.125.$$

is an incenter of core(μ).

Besides, if we do not require the open ball to be included in $\mathbb{P}^*(\mathcal{X})$, we obtain:

$$\sup\{\alpha > 0 \mid \exists P_0 \in \text{core}(\mu) \text{ such that } B_o^\alpha(P_0) \subseteq \text{core}(\mu)\} = 0.2,$$

and the only $P_0 \in \text{core}(\mu)$ satisfying $B_o^{0.2}(P_0) \subseteq \text{core}(\mu)$ is given by:

$$P_0(\{x_1\}) = 0.7, \quad P_0(\{x_2\}) = 0.3, \quad P_0(\{x_3\}) = 0.$$

This probability measure belongs to the boundary of core(μ), which leads us to believe that it does not adequately represent the idea underlying the notion of incenter.

This example also shows the necessity of taking the open ball rather than the closed in the definition of the incenter radius in Eq. (7). The graphical representation of P_0, Q_1 and Q_2 can be seen in Fig. 2. ♦

On the other hand, when the core of the capacity is included in the interior of the simplex, then we immediately have that $B_o^\alpha(P) \subseteq \text{core}(\mu) \cap \mathbb{P}^*(\mathcal{X})$ if and only if $B_c^\alpha(P) \subseteq \text{core}(\mu)$, whence Eq. (7) can be rewritten as

$$\alpha_I = \sup \left\{ \alpha > 0 \mid \exists P_0 \in \text{core}(\mu) \text{ such that } B_c^\alpha(P_0) \subseteq \text{core}(\mu) \right\}. \qquad (8)$$

A first question arising naturally from the definition of incenter is whether it always exists. Our next result shows that this is indeed the case.

Proposition 2 *Consider an exact capacity μ such that core(μ) has a non-empty interior. Then the value α_I given by Eq. (7) is a maximum. As a consequence, the incenter of μ always exists.*

On the other hand, a capacity may have more than one incenter; while this was already showed by Example 2, we next give another example where the core of the capacity is included in $\mathbb{P}^*(\mathcal{X})$:

Fig. 2 Graphical representation of the core of core(μ) in Example 2, the incenters Q_1 and Q_2 (in blue) and the probability P_0 (in red)
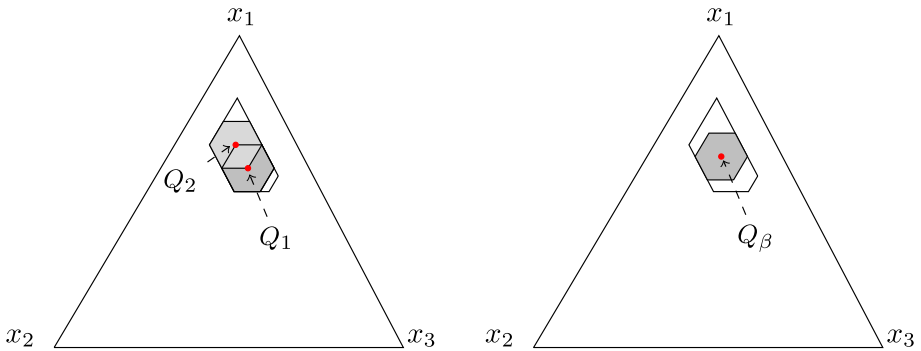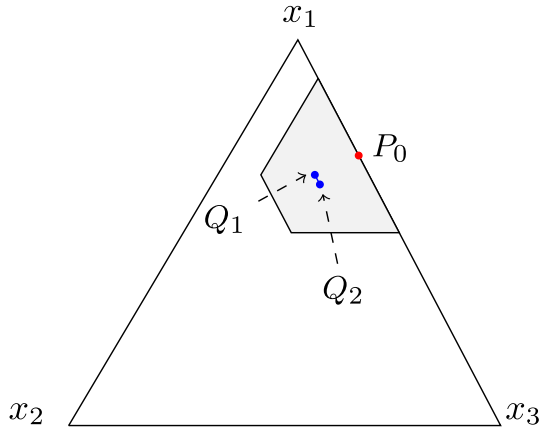

Fig. 3 Graphical representation of the core of μ, the incenters Q_1 and Q_2 with the ball they induce (left hand-side figure) and the incenter Q_β for $\beta = 0.5$ with the ball it induces (right hand-side figure)

Example 3 Let $\mathcal{X} = \{x_1, x_2, x_3\}$, and consider the capacity μ given by:

A	$\{x_1\}$	$\{x_2\}$	$\{x_3\}$	$\{x_1, x_2\}$	$\{x_1, x_3\}$	$\{x_2, x_3\}$	$\mathcal{X}$
$\mu(A)$	0.5	0.1	0.1	0.65	0.75	0.2	1

The value α_I determined in Eq. (7) is given by:

$$\alpha_I = \max\left\{\alpha \mid \exists P_0 \in \text{core}(\mu) \text{ such that } B_o^\alpha(P_0) \subseteq \text{core}(\mu) \cap \mathbb{P}^*(\mathcal{X})\right\} = 0.075.$$

To see that there is more than one $P_0 \in \text{core}(\mu)$ such that $B_o^{\alpha_I}(P_0) \subseteq \text{core}(\mu) \cap \mathbb{P}^*(\mathcal{X})$, note that the set of such P_0 is given by the probability measures Q_1, Q_2 defined below as well as any convex combination $Q_\beta = \beta Q_1 + (1 - \beta) Q_2$ for $\beta \in [0, 1]$:

$$Q_1(\{x_1\}) = 0.575, \quad Q_1(\{x_2\}) = 0.175, \quad Q_1(\{x_3\}) = 0.25.$$
$$Q_2(\{x_1\}) = 0.65, \quad Q_2(\{x_2\}) = 0.175, \quad Q_2(\{x_3\}) = 0.175.$$

Figure 3 gives the graphical representation of the core of μ as well as the balls $B_c^{\alpha_I}(Q_1)$, $B_c^{\alpha_I}(Q_2)$ and $B_c^{\alpha_I}(Q_\beta)$ for $\beta = 0.5$. ♦

Due to the lack of uniqueness, we shall denote by Ψ_3^μ the non-empty set of incenters of the capacity μ.

Fig. 4 Graphical representation of the core of core(μ) (in gray) in Example 4, as well as the ball $B_c^{\alpha'_C}(P_0)$
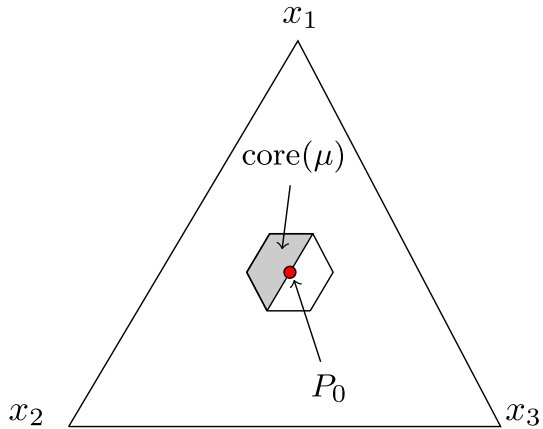
A dual approach to the one above is to consider the *circumcenter* of the capacity, which is the center (or centers) of the smallest balls that include the core. Formally, we may consider

$$\alpha_C = \inf \left\{ \alpha > 0 \mid \exists P_0 \in \mathbb{P}(\mathcal{X}) \text{ such that } B_c^{\alpha}(P_0) \supseteq \text{core}(\mu) \right\},$$

and then consider the set of those P such that $B_c^{\alpha_C}(P) \supseteq \text{core}(\mu)$. However, this approach has the two drawbacks we have discussed so far: not only it need not lead to a unique solution, but also it may produce values outside the core, as showed in Bader et al. (2012); this last issue may be overcome by considering instead

$$\alpha'_C = \inf \left\{ \alpha > 0 \mid \exists P_0 \in \text{core}(\mu) \text{ such that } B_c^{\alpha}(P_0) \supseteq \text{core}(\mu) \right\},$$

and by calling circumcenter those probabilities $P \in \text{core}(\mu)$ such that $B_d^{\alpha'_C}(P) \supseteq \text{core}(\mu)$. However, this second approach does not prevent us from obtaining circumcenters that lie in the boundary of the core, leading to the same counterintuitive situation we discussed in Example 2.

Example 4 Consider $\mathcal{X} = \{x_1, x_2, x_3\}$ and the capacity μ given by:

A	$\{x_1\}$	$\{x_2\}$	$\{x_3\}$	$\{x_1, x_2\}$	$\{x_1, x_3\}$	$\{x_2, x_3\}$	$\mathcal{X}$
$\mu(A)$	0.3	0.2	0.2	0.7	0.6	0.5	1

This capacity is exact and the extreme points of its core are given by:

	x_1	x_2	x_3		x_1	x_2	x_3
P_1	0.3	0.4	0.3	P_3	0.4	0.4	0.2
P_2	0.5	0.2	0.3	P_4	0.5	0.3	0.2

It follows that $\alpha'_C = 0.1$ and that the only circumcenter is P_0 given by $P_0(\{x_1\}) = 0.4$, $P_0(\{x_2\}) = P_0(\{x_3\}) = 0.3$. The graphical representation of the core of μ and the ball $B_c^{\alpha'_C}(P_0)$ can be seen in Fig. 4. ♦

3.4 Contraction centroid

Our fourth and last approach is motivated by the lack of uniqueness that has been illustrated in the case of the incenter.

Given an exact capacity μ with conjugate $\bar{\mu}$, we can split the events in $\mathcal{X}$ into $\mathcal{L}^=$ and $\mathcal{L}^>$ such that:

$$\mu(A) = \bar{\mu}(A) \quad \forall A \in \mathcal{L}^= \quad \text{and} \quad \mu(A) < \bar{\mu}(A) \quad \forall A \in \mathcal{L}^>.$$

Using this notation, the core of μ can be expressed as:

$$\text{core}(\mu) = \{P \in \mathbb{P}(\mathcal{X}) \mid P(A) = \mu(A) \, \forall A \in \mathcal{L}^=, \ P(A) \geq \mu(A) \, \forall A \in \mathcal{L}^>\}. \quad (9)$$

Note that when $\mathcal{L}^>$ is empty we obtain that $\mu(A) = \bar{\mu}(A)$ for any $A \subseteq \mathcal{X}$, meaning that μ is additive and that its core contains one single element.

The idea in this approach is to contract the core in a uniform manner as long as we can, and then proceed in the same way by reducing the number of constraints. More specifically, we increase the value of the capacity in a constant amount α in all the events $A \in \mathcal{L}^>$. In this respect, we wonder whether there is some value α small enough such that this approach gives rise to a non-empty core, and also which is the maximum/supremum we can consider. Our next result gives the answer to both questions.

Proposition 3 *Consider an exact capacity μ, and let us express its core as in Eq. (9) using the sets $\mathcal{L}^=$ and $\mathcal{L}^>$. For a given $\alpha > 0$, let us define:*

$$core(\mu)_\alpha = \{P \in \mathbb{P}(\mathcal{X}) \mid P(A) = \mu(A) \, \forall A \in \mathcal{L}^=, \ P(A) \geq \mu(A) + \alpha \, \forall A \in \mathcal{L}^>\},$$

and let $\Lambda = \{\alpha > 0 \mid core(\mu)_\alpha \neq \emptyset\}$. It holds that:

(a) $\Lambda \neq \emptyset$.
(b) *The set Λ has a maximum α_S.*
(c) *There exists some $A \in \mathcal{L}^>$ such that $P(A)$ is constant for any $P \in core(\mu)_{\alpha_S}$.*

This result assures that we can uniformly increase the capacity in the events whose value is imprecise (i.e., such that $\mu(A) < \bar{\mu}(A)$) and that when the process stops the size of $\mathcal{L}^>$ has decreased in the sense that, given the exact capacity μ_{α_S} that is the lower envelope of the set

$$\text{core}(\mu)_{\alpha_S} = \{P \in \mathbb{P}(\mathcal{X}) \mid P(A) \geq \mu(A) + \alpha_S \, \forall A \in \mathcal{L}^>\}$$

and its conjugate $\bar{\mu}_{\alpha_S}$, then $\{A \mid \mu_{\alpha_S}(A) < \bar{\mu}_{\alpha_S}(A)\} \subsetneq \mathcal{L}^>$.

This means that we may apply the same procedure to the capacity μ_{α_S}, and after iterating it a finite number of times, we obtain a set formed by a single element that we shall call the contraction centroid. In other words, the procedure leads to the values $\alpha_{S_1} = \max \Lambda_1, \ldots, \alpha_{S_l} = \max \Lambda_l$ and the chain of nested sets[5]

$$\text{core}(\mu) \supset \text{core}(\mu)_{\alpha_{S_1}} \supset \ldots \supset \text{core}(\mu)_{\alpha_{S_l}} = \{\Phi_4^\mu\}. \quad (10)$$

Φ_4^μ is the *contraction centroid* of the capacity μ.

Let us illustrate this procedure with an example.

Example 5 Consider again the exact capacity from Example 3. There, the set $\mathcal{L}^=$ only contains the trivial events $\emptyset$ and $\mathcal{X}$, while $\mathcal{L}^>$ contains the six non-trivial events:

A	$\{x_1\}$	$\{x_2\}$	$\{x_3\}$	$\{x_1, x_2\}$	$\{x_1, x_3\}$	$\{x_2, x_3\}$	$\mathcal{X}$
$\mu(A)$	0.5	0.1	0.1	0.65	0.75	0.2	1
$\bar{\mu}(A)$	0.8	0.25	0.35	0.9	0.9	0.5	1

---

[5] This is an abuse of notation, since to be correct we should write $core(\mu_{\alpha_{S_1}})_{\alpha_{S_2}}$ instead of $core(\mu)_{\alpha_{S_2}}$, and similarly for the subsequent iterations of the procedure; we are using the one in Eq. (10) so as to alleviate the notation.
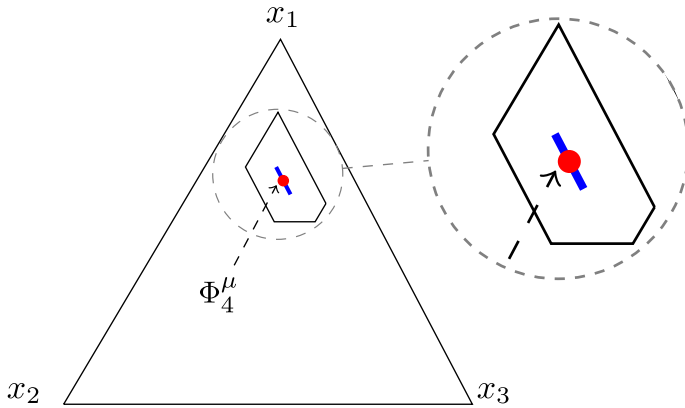
Fig. 5 Graphical representation of the core of μ, as well as the sets $\text{core}(\mu)_{\alpha_{S_1}}$ (in blue) and $\text{core}(\mu)_{\alpha_{S_2}}$ (in red)

Let us see that $\alpha_{S_1} = \max \Lambda = 0.075$. On the one hand, $\text{core}(\mu)_{\alpha_{S_1}}$ is non-empty because it includes for instance the probability measure P given by:

$$P(\{x_1\}) = 0.625, \quad P(\{x_2\}) = 0.175, \quad P(\{x_3\}) = 0.2.$$

To see on the other hand that this is the maximum value of Λ, note that if we increase μ in $\alpha > 0$, to keep exactness it should be that:

$$1 \geq \big(\mu(\{x_2\}) + \alpha\big) + \big(\mu(\{x_1, x_3\}) + \alpha\big) = 0.85 + 2\alpha,$$

whence $\alpha \leq 0.075$. Therefore, $\alpha_{S_1} = 0.075$, and this gives rise to the following core:

$$\text{core}(\mu)_{\alpha_{S_1}} = \big\{P \in \mathbb{P}(\mathcal{X}) \mid P(A) \geq \mu(A) + \alpha_{S_1} \ \forall A \neq \emptyset, \mathcal{X}\big\}.$$

The exact capacity μ_1 determined as the lower envelope of $\text{core}(\mu)_{\alpha_{S_1}}$ and its conjugate $\overline{\mu}_1$ are given by:

A	$\{x_1\}$	$\{x_2\}$	$\{x_3\}$	$\{x_1, x_2\}$	$\{x_1, x_3\}$	$\{x_2, x_3\}$	$\mathcal{X}$
$\mu_1(A)$	0.575	0.175	0.175	0.75	0.825	0.35	1
$\overline{\mu}_1(A)$	0.65	0.175	0.25	0.825	0.825	0.425	1

Let us apply now the procedure to this capacity. The sets $\mathcal{L}_1^>$ and $\mathcal{L}_1^=$ are given by:

$$\mathcal{L}_1^= = \big\{\{x_2\}, \{x_1, x_3\}, \emptyset, \mathcal{X}\big\}, \quad \text{and} \quad \mathcal{L}_1^> = \big\{\{x_1\}, \{x_3\}, \{x_1, x_2\}, \{x_2, x_3\}\big\},$$

i.e., there are two non trivial events whose value is now fixed, $\{x_2\}$ and $\{x_1, x_3\}$.

Repeating the same steps, we obtain $\alpha_{S_2} = 0.0375$, and in this case $\text{core}(\mu)_{\alpha_{S_2}}$ is formed by a single probability measure, that is therefore the contraction centroid Φ_4^{μ}. It is given by:

$$\Phi_4^{\mu}(\{x_1\}) = 0.6125, \quad \Phi_4^{\mu}(\{x_2\}) = 0.175, \quad \Phi_4^{\mu}(\{x_1\}) = 0.2125. \tag{11}$$

In Fig. 5 we have depicted the sets $\text{core}(\mu)_{\alpha_{S_1}}$ (in blue) and $\text{core}(\mu)_{\alpha_{S_2}}$ (in red), as well as the initial core of μ. ♦

In the above procedure, it is worth mentioning that the lower envelope of the set core$(\mu)_\alpha$ does not necessarily coincide with the capacity μ' given by

$$\mu'(A) = \begin{cases} \mu(A) + \alpha & \text{if } A \in \mathcal{L}^> \\ \mu(A) & \text{otherwise.} \end{cases}$$

While by construction it dominates this capacity, they may not agree on some events because μ' need not be exact. This can be seen in Example 5, where given $A = \{x_1, x_2\}$, it holds that $\mu'(A) = \mu(A) + \alpha_{S_1} = 0.65 + 0.075 < 0.75 = \mu_1(A)$.

Proposition 3 assures that there exists a maximum value $\alpha_S = \max \Lambda$ giving rise to a non-empty set. This naturally leads us to the problem of computing more efficiently the value of α_S. Our next result gives an explicit formula for α_S for a particular case of exact capacities which only coincide with their conjugate for the trivial cases of $\emptyset, \mathcal{X}$.

Definition 4 Let μ be an exact capacity with conjugate $\bar{\mu}$. We shall call μ *maximally imprecise* when $\mu(A) < \bar{\mu}(A)$ for every $A \neq \emptyset, \mathcal{X}$.

Let us define

$$\mathbb{A}(\mathcal{X}) = \left\{ \mathcal{A} = \left(A_i\right)_{i=1,\dots,k} \mid \exists \beta_\mathcal{A} \in \mathbb{N} \text{ such that } \sum_{i=1}^k I_{A_i} = \beta_\mathcal{A} \right\}.$$

In other words, $\mathbb{A}(\mathcal{X})$ is the class of all finite families of subsets of $\mathcal{X}$ such that every $x \in \mathcal{X}$ belongs to the same number of elements in the family. Note that in each of these families $\mathcal{A}$ there may be repeated elements, i.e., we may consider for instance the family $\mathcal{A} = \left(\{x_1\}, \{x_1\}, \{x_2\}, \{x_2, x_3\}, \{x_3\}\right)$ on $\mathcal{X} = \{x_1, x_2, x_3\}$. We consider on $\mathbb{A}(\mathcal{X})$ the partial order determined by the inclusion, i.e., we say that $\mathcal{A}_1 \subseteq \mathcal{A}_2$ when each element in the family $\mathcal{A}_1$ also belongs to the family $\mathcal{A}_2$.

Theorem 4 *Let μ be a maximally imprecise exact capacity. Then*

$$\alpha_S = \min_{\mathcal{A} \in \mathbb{A}(\mathcal{X})} h_\mathcal{A}, \text{ where } h_\mathcal{A} := \frac{1}{|\mathcal{A}|} \left(\beta_\mathcal{A} - \sum_{A \in \mathcal{A}} \mu(A) \right). \tag{12}$$

Let us return to the running Example 5. In that case, α_S satisfies Eq. (12) for $\mathcal{A} = \left(\{x_2\}, \{x_1, x_3\}\right)$ and $\beta_\mathcal{A} = 1$, giving that

$$\frac{1}{|\mathcal{A}|} \left(\beta_\mathcal{A} - \mu(\{x_2\}) - \mu(\{x_1, x_3\}) \right) = \frac{1}{2}(1 - 0.1 - 0.75) = 0.075,$$

which is indeed the value we obtained in Example 5.

The computation of α_S in Eq. (12) requires the computation of the value $h_\mathcal{A}$ for all the families $\mathcal{A}$. Our next result shows a more tractable expression when the capacity is supermodular. For this aim, we denote by $\mathbb{A}^*(\mathcal{X})$ the subclass of $\mathbb{A}(\mathcal{X})$ formed by the partitions of $\mathcal{X}$.

Theorem 5 *Let μ be a maximally imprecise supermodular capacity with conjugate $\bar{\mu}$. Then:*

$$\alpha_S = \min_{\mathcal{A} \in \mathbb{A}^*(\mathcal{X})} \left\{ \frac{1 - \sum_{A \in \mathcal{A}} \mu(A)}{|\mathcal{A}|}, \frac{\sum_{A \in \mathcal{A}} \bar{\mu}(A) - 1}{|\mathcal{A}|} \right\}. \tag{13}$$

This means that, under supermodularity, it suffices to focus on partitions of $\mathcal{X}$, which simplifies considerably the computation of Eq. (12).

Example 6 Consider again our running Example 3. The capacity in that example is supermodular because any exact capacity in a 3-element space is, so Theorem 5 is applicable. Next table summarises the values associated with each partition in Eq. (13):

$\mathcal{A} \in \mathbb{A}^*(\mathcal{X})$	$\|\mathcal{A}\|$	$(1-\sum_{A \in \mathcal{A}} \mu(A))/\|\mathcal{A}\|$	$(\sum_{A \in \mathcal{A}} \bar{\mu}(A)-1)/\|\mathcal{A}\|$
$\{x_1\}, \{x_2\}, \{x_3\}$	3	$(1-(0.5+0.1+0.1))/3 = 0.1$	$(0.8+0.25+0.35-1)/3 = 0.1\overline{3}$
$\{x_1, x_2\}, \{x_3\}$	2	$(1-(0.65+0.1))/2 = 0.125$	$((0.9+0.35)-1)/2 = 0.125$
$\{x_1, x_3\}, \{x_2\}$	2	$(1-(0.75+0.1))/2 = 0.075$	$((0.9+0.25)-1)/2 = 0.075$
$\{x_2, x_3\}, \{x_1\}$	2	$(1-(0.2+0.5))/2 = 0.15$	$((0.5+0.8)-1)/2 = 0.15$

The minimum of these values is $\alpha_S = 0.075$ and it is attained in the partition $\mathcal{A}$ formed by $\{x_1, x_3\}$ and $\{x_2\}$, both for the capacity and its conjugate. This is in line with our comments in Example 5. ♦

Let us show that the result in Theorem 5 does not generalise to the case where μ is an exact capacity but not supermodular:

Example 7 Consider $\mathcal{X} = \{x_1, x_2, x_3, x_4\}$ and the capacity μ, with conjugate $\bar{\mu}$, given by:

A	$\mu(A)$	$\bar{\mu}(A)$		A	$\mu(A)$	$\bar{\mu}(A)$
$\{x_1\}$	0	0.1		$\{x_2, x_3\}$	0.6	0.8
$\{x_2\}$	0.1	0.2		$\{x_2, x_4\}$	0.4	0.5
$\{x_3\}$	0.5	0.6		$\{x_3, x_4\}$	0.7	0.8
$\{x_4\}$	0.2	0.3		$\{x_1, x_2, x_3\}$	0.7	0.8
$\{x_1, x_2\}$	0.2	0.3		$\{x_1, x_2, x_4\}$	0.4	0.5
$\{x_1, x_3\}$	0.5	0.6		$\{x_1, x_3, x_4\}$	0.8	0.9
$\{x_1, x_4\}$	0.2	0.4		$\{x_2, x_3, x_4\}$	0.9	1

The extreme points of its core are given by:

	x_1	x_2	x_3	x_4			x_1	x_2	x_3	x_4
P_1	0.1	0.2	0.5	0.2		P_3	0.1	0.1	0.5	0.3
P_2	0	0.2	0.6	0.2		P_4	0	0.2	0.5	0.3

It follows that μ is an exact capacity, and that it is maximally imprecise too. Note that μ is not supermodular, since

$$0.6 = \mu(\{x_1, x_2\}) + \mu(\{x_2, x_4\}) \not\leq \mu(\{x_1, x_2, x_4\}) + \mu(\{x_2\}) = 0.5.$$

The value of α_S is attained with $\mathcal{A} = (\{x_3\}, \{x_2, x_4\}, \{x_1, x_2\}, \{x_1, x_3, x_4\})$, that produces:

$$\alpha_S = \frac{1}{4}\left(2 - \mu(\{x_3\}) - \mu(\{x_2, x_4\}) - \mu(\{x_1, x_2\}) - \mu(\{x_1, x_3, x_4\})\right) = 0.025.$$

To see this, note any $P \in \text{core}(\mu)_{0.025}$ satisfies:

$$2 = \beta_{\mathcal{A}} \geq \left(\mu(\{x_3\}) + 0.025\right) + \left(\mu(\{x_2, x_4\}) + 0.025\right)$$
$$+ \left(\mu(\{x_1, x_2\}) + 0.025\right) + \left(\mu(\{x_1, x_3, x_4\}) + 0.025\right) = 2,$$

which implies that $P(A) = \mu(A) + 0.025$ for every $A \in \mathcal{A}$, and also that $\text{core}(\mu)_\alpha = \emptyset$ for every $\alpha > 0.025$. Moreover, the above restrictions imply that

$$P(\{x_3\}) = 0.525.$$
$$P(\{x_2, x_4\}) = 0.425 \Rightarrow P(\{x_1, x_3\}) = 0.575, \text{ whence } P(\{x_1\}) = 0.05.$$
$$P(\{x_1, x_2\}) = 0.225 \Rightarrow P(\{x_2\}) = 0.175.$$

$$P(\{x_2, x_4\}) = 0.425 \Rightarrow P(\{x_4\}) = 0.25.$$

Thus, $\text{core}(\mu)_{0.025}$ only includes the probability mass function $(0.05, 0.175, 0.525, 0.25)$. It is easy now to verify that this probability measure satisfies $P(A) \geq \mu(A) + 0.025$ for any non-trivial event A.

Moreover, the bounds determined by the partitions in $\mathbb{A}^*(\mathcal{X})$ are the following:

Partition $\mathcal{A}$	$|\mathcal{A}|$	$(1 - \sum_{A \in \mathcal{A}} \mu(A))/|\mathcal{A}|$	$(\sum_{A \in \mathcal{A}} \bar{\mu}(A) - 1)/|\mathcal{A}|$
$\{x_1\}, \{x_2\}, \{x_3\}, \{x_4\}$	4	$(1-0.8)/4$	$(1.2-1)/4$
$\{x_1, x_2\}, \{x_3\}, \{x_4\}$	3	$(1-0.9)/3$	$(1.2-1)/3$
$\{x_1, x_3\}, \{x_2\}, \{x_4\}$	3	$(1-0.8)/3$	$(1.1-1)/3$
$\{x_1, x_4\}, \{x_2\}, \{x_3\}$	3	$(1-0.8)/3$	$(1.2-1)/3$
$\{x_2, x_3\}, \{x_1\}, \{x_4\}$	3	$(1-0.8)/3$	$(1.2-1)/3$
$\{x_2, x_4\}, \{x_1\}, \{x_3\}$	3	$(1-0.9)/3$	$(1.2-1)/3$
$\{x_3, x_4\}, \{x_1\}, \{x_2\}$	3	$(1-0.8)/3$	$(1.1-1)/3$
$\{x_1, x_2\}, \{x_3, x_4\}$	2	$(1-0.9)/2$	$(1.1-1)/2$
$\{x_1, x_3\}, \{x_2, x_4\}$	2	$(1-0.9)/2$	$(1.1-1)/2$
$\{x_1, x_4\}, \{x_2, x_3\}$	2	$(1-0.8)/2$	$(1.2-1)/2$
$\{x_1\}, \{x_2, x_3, x_4\}$	2	$(1-0.9)/2$	$(1.1-1)/2$
$\{x_2\}, \{x_1, x_3, x_4\}$	2	$(1-0.9)/2$	$(1.1-1)/2$
$\{x_3\}, \{x_1, x_2, x_4\}$	2	$(1-0.9)/2$	$(1.1-1)/2$
$\{x_4\}, \{x_1, x_2, x_3\}$	2	$(1-0.9)/2$	$(1.1-1)/2$

We obtain the minimum value $\frac{0.1}{3} = 0.0\bar{3}$, strictly greater than $\alpha_S = 0.025$. We conclude that Theorem 5 does not hold without the hypothesis of supermodularity. $\blacklozenge$

3.5 Relationships between the centroids

So far, we have introduced four different notions of the center of an exact capacity. Let us begin by showing that these four notions are indeed different:

Example 8 Consider the capacity defined in Example 3; there, we gave the set of incenters Ψ_3^μ, while the contraction centroid Φ_4^μ was given in Example 5. The extreme points of the core of μ are given by:

	x_1	x_2	x_3
P_1	0.55	0.1	0.35
P_2	0.5	0.15	0.35
P_3	0.5	0.25	0.25

	x_1	x_2	x_3
P_4	0.65	0.25	0.1
P_5	0.8	0.1	0.1

From this, we conclude that the average of the extreme points of $\text{core}(\mu)$ is the probability measure with mass function

$$\Phi_2^\mu(\{x_1\}) = 0.6, \quad \Phi_2^\mu(\{x_2\}) = 0.17, \quad \Phi_2^\mu(\{x_3\}) = 0.23. \tag{14}$$

On the other hand, the Shapley value is given by

$$\Phi_1^\mu(\{x_1\}) = 0.6\bar{3}, \quad \Phi_1^\mu(\{x_2\}) = 0.158\bar{3}, \quad \Phi_1^\mu(\{x_3\}) = 0.208\bar{3}.$$
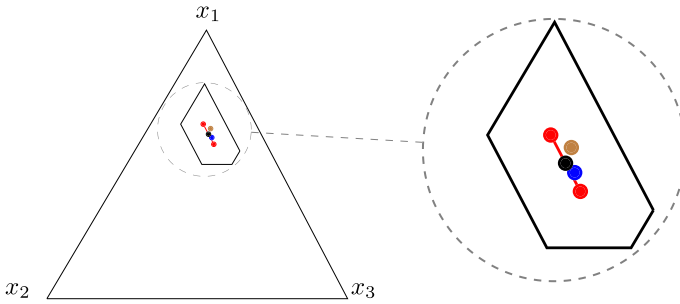
Fig. 6 Graphical representation of the centroids of the exact capacity in Example 3: the average of the extreme points (in blue), the Shapley value (in brown), the set of incenters (in red) and the contraction centroid (in black)

This can be derived also using Proposition 1, noting that the permutations lead to the following extreme points:

		x_1	x_2	x_3
$\sigma_1 = (1, 2, 3)$	P_{σ_1}	0.5	0.15	0.35
$\sigma_2 = (1, 3, 2)$	P_{σ_2}	0.5	0.25	0.25
$\sigma_3 = (2, 1, 3)$	P_{σ_3}	0.55	0.1	0.35
$\sigma_4 = (2, 3, 1)$	P_{σ_4}	0.8	0.1	0.1
$\sigma_5 = (3, 1, 2)$	P_{σ_5}	0.65	0.25	0.1
$\sigma_6 = (3, 2, 1)$	P_{σ_6}	0.8	0.1	0.1

Here, $P_{\sigma_4} = P_{\sigma_6}$, so this extreme point gets twice the weight of the others in the computation of Shapley value, hence the difference with Φ_2^μ. These centroids are represented in Fig. 6. ♦

Even if the four approaches do not lead to the same solution in general, in Examples 5 and 8 we have seen an example where the contraction centroid Φ_4^μ belongs to the set of incenters Ψ_3^μ. This leads us to investigate whether there is a connection between these two approaches. Our next result shows that, under some conditions, the set we obtain in the first step of the contraction approach coincides with the set of incenters.

Proposition 6 *Let μ be a maximally imprecise exact capacity with conjugate $\bar{\mu}$ satisfying $\mu(A) > 0$ for every $A \neq \emptyset$, and let α_S be the coefficient defined in Proposition 3. Given $P_0 \in core(\mu)$ and $\alpha \leq \alpha_S$,*

$$P_0 \in core(\mu)_\alpha \Leftrightarrow B_c^\alpha(P_0) \subseteq core(\mu).$$

As a consequence, $core(\mu)_{\alpha_S} = \Psi_3^\mu$.

On the other hand, an advantage of using the contraction procedure is that the existence of Φ_4^μ is guaranteed even if the interior of $core(\mu)$ is empty.

4 Properties of the centroids

Next we compare the different centroids in terms of the axiomatic properties they satisfy. There exist several axiomatic characterisations of Shapley value in the context of coalitional game theory; arguably the most important one is that as the unique additive measure Φ^μ satisfying the following axioms:

Efficiency $\sum_{i=1}^{n} \Phi^\mu(\{x_i\}) = \mu(\mathcal{X})$.

Symmetry $\mu(A \cup \{x_i\}) = \mu(A \cup \{x_j\})$ for any $A \subseteq \mathcal{X} \setminus \{x_i, x_j\}$ implies that $\Phi^\mu(\{x_i\}) = \Phi^\mu(\{x_j\})$.

Linearity $\Phi^{\lambda_1 \mu_1 + \lambda_2 \mu_2} = \lambda_1 \Phi^{\mu_1} + \lambda_2 \Phi^{\mu_2}$ for any $\lambda_1, \lambda_2 \in \mathbb{R}$ and every μ_1, μ_2.

Null player $\mu(A \cup \{x_i\}) = \mu(A)$ for any $A \subseteq \mathcal{X} \setminus \{x_i\}$ implies $\Phi^\mu(\{x_i\}) = 0$.

Throughout this paper, we are restricting ourselves to normalised capacities, which implies that (i) the efficiency axiom simplifies to $\sum_{i=1}^{n} \Phi^\mu(\{x_i\}) = 1$; and (ii) in order to guarantee that $\lambda_1 \mu_1 + \lambda_2 \mu_2$ is again a normalised capacity, we must have $\lambda_2 = 1 - \lambda_1$, and therefore the linearity axiom implies that $\Phi^{\lambda \mu_1 + (1-\lambda) \mu_2} = \lambda \Phi^{\mu_1} + (1-\lambda) \Phi^{\mu_2}$ for any $\lambda \in \mathbb{R}$ and any normalised capacities μ_1, μ_2 whenever $\lambda \mu_1 + (1-\lambda) \mu_2$ is a normalised capacity too.

Next we investigate to which extent these properties are satisfied by the other centroids proposed in this paper. In this respect, note that in the framework of this paper, any center of a capacity shall be a probability measure, whence the efficiency property is trivially satisfied. Note also that when analysing the behaviour of the set of incenters, we shall say that it satisfies a property if and only if any of its elements does.

As we have already mentioned, the Shapley centroid does not necessarily satisfy *feasibility*, meaning that Φ_1^μ may not belong to the core of μ. By construction, the vertex, incenter and contraction centroids do satisfy feasibility. With respect to the other properties, it is not difficult to establish the following:

Proposition 7 Φ_2^μ, *any* $\Phi_3^\mu \in \Psi_3^\mu$ *and* Φ_4^μ *satisfy the symmetry and null-player properties.*

To see that they do not satisfy linearity in general, whence their difference with Shapley value, consider the following example:

Example 9 Consider $\mathcal{X} = \{x_1, x_2, x_3\}$, the exact capacities μ_1, μ_2 and their average $\mu := 0.5\mu_1 + 0.5\mu_2$, given in the following table:

A	$\{x_1\}$	$\{x_2\}$	$\{x_3\}$	$\{x_1, x_2\}$	$\{x_1, x_3\}$	$\{x_2, x_3\}$	$\mathcal{X}$
$\mu_1(A)$	0.1	0.2	0.3	0.35	0.55	0.75	1
$\mu_1(A)$	0.3	0.2	0.1	0.75	0.55	0.35	1
$\mu(A)$	0.2	0.2	0.2	0.55	0.55	0.55	1

Because of the symmetry property, it is easy to see that for μ we obtain that

$$\Phi_2^\mu = \Phi_3^\mu = \Phi_4^\mu = (1/3, 1/3, 1/3).$$

On the other hand, $\Phi_2^{\mu_1} = (0.17, 0.31, 0.52)$ and $\Phi_2^{\mu_2} = (0.52, 0.31, 0.17)$, whence $0.5 \cdot \Phi_2^{\mu_1} + 0.5 \cdot \Phi_2^{\mu_2} = (0.345, 0.31, 0.345)$.

With respect to the contraction centroid, in this case $\Phi_4^{\mu_1} = (0.175, 0.325, 0.5)$ and $\Phi_4^{\mu_2} = (0.5, 0.325, 0.175)$, whence

$$0.5\Phi_4^{\mu_1} + 0.5\Phi_4^{\mu_2} = (0.3375, 0.325, 0.3375).$$

Finally, with respect to the incenters, $\Psi_3^{\mu_1}$ is the set of convex combinations of $\{(0.175, 0.375, 0.45), (0.175, 0.275, 0.55)\}$ and $\Psi_3^{\mu_2}$ is the set of convex combinations of $\{(0.45, 0.375, 0.175), (0.55, 0.275, 0.175)\}$. Thus, Ψ_3^μ does not coincide with the set $\{0.5Q_1 + 0.5Q_2 \mid Q_1 \in \Psi_3^{\mu_1}, Q_2 \in \Psi_3^{\mu_2}\}$, because this latter set includes for instance $(0.3375, 0.325, 0.3375)$. As a consequence, none of three centroids satisfies linearity. ♦

Next we consider other desirable properties of a centroid.

Definition 5 Let Φ^μ be a centroid of an exact capacity μ. We say that it satisfies:

- *Continuity* if for any $\varepsilon > 0$, there exists $\delta > 0$ such that $d(\mu_1, \mu_2) := \max_{A \subseteq \mathcal{X}} \mid \mu_1(A) - \mu_2(A) \mid < \delta$ implies $d(\Phi^{\mu_1}, \Phi^{\mu_2}) < \varepsilon$.
- *Ignorance preservation* if $\text{core}(\mu) = \mathbb{P}(\mathcal{X})$, then Φ^μ is the uniform distribution.

When dealing with the incenter, the previous properties should be slightly rewritten due to its lack of uniqueness: the incenter satisfies ignorance preservation if $\text{core}(\mu) = \mathbb{P}(\mathcal{X})$ implies that Ψ_3^μ only contains the uniform distribution; and it satisfies continuity when for any $\varepsilon > 0$ there exists some $\delta > 0$ such that $d(\mu_1, \mu_2) < \delta$ implies that $d(\nu_1, \nu_2) < \varepsilon$, where ν_1 and ν_2 are the lower envelopes of $\Psi_3^{\mu_1}$ and $\Psi_3^{\mu_2}$.

Proposition 8 (a) Φ_1^μ *satisfies continuity and ignorance preservation.*
(b) Φ_2^μ *satisfies ignorance preservation.*
(c) Ψ_3^μ, Φ_4^μ *satisfy continuity and ignorance preservation.*

To see that the average of the extreme points does not satisfy continuity, consider the following example:

Example 10 Consider $\mathcal{X} = \{x_1, x_2, x_3\}$, $\epsilon \in (0, 0.05)$ and the exact capacity μ given by

A	$\{x_1\}$	$\{x_2\}$	$\{x_3\}$	$\{x_1, x_2\}$	$\{x_1, x_3\}$	$\{x_2, x_3\}$	$\mathcal{X}$
$\mu(A)$	0.1	0.2	0.3	0.35	$0.55 - 2\epsilon$	$0.75 + \epsilon$	1

Then μ is supermodular, and the extreme points of $\text{core}(\mu)$ are given by:

		x_1	x_2	x_3
$\sigma_1 = (1, 2, 3)$	P_{σ_1}	0.1	0.25	0.65
$\sigma_2 = (1, 3, 2)$	P_{σ_2}	0.1	$0.45 + 2\varepsilon$	$0.45 - 2\varepsilon$
$\sigma_3 = (2, 1, 3)$	P_{σ_3}	0.15	0.2	0.65
$\sigma_4 = (2, 3, 1)$	P_{σ_4}	$0.25 - \varepsilon$	0.2	$0.55 + \varepsilon$
$\sigma_5 = (3, 1, 2)$	P_{σ_5}	$0.25 - 2\varepsilon$	$0.45 + 2\varepsilon$	0.3
$\sigma_6 = (3, 2, 1)$	P_{σ_6}	$0.25 - \varepsilon$	$0.45 + \varepsilon$	0.3

All these extreme points are different, and their average is given by

$$\Phi_2^\mu = \left(\frac{1.1 - 4\varepsilon}{6}, \frac{2 + 5\varepsilon}{6}, \frac{2.9 - \varepsilon}{6} \right).$$

On the other hand, when $\varepsilon = 0$ we obtain that $P_{\sigma_5} = P_{\sigma_6}$, and, as we have seen in Example 9, the average of the extreme points becomes $\Phi_2^\mu = (0.17, 0.31, 0.52) \neq \lim_{\varepsilon \to 0}(\frac{1.1 - 4\varepsilon}{6}, \frac{2 + 5\varepsilon}{6}, \frac{2.9 - \varepsilon}{6})$. ♦

Another desirable property would be that the centroid preserves the same preferences as μ, in the sense that $\mu(A) \geq \mu(B) \Rightarrow \Phi^\mu(A) \geq \Phi^\mu(B)$. Since Φ^μ is an additive model, we shall only require this property on the singletons: otherwise the capacity should satisfy

$$\mu(A) \geq \mu(B) \Rightarrow \mu(A \cup C) \geq \mu(B \cup C) \quad \forall C \subseteq (A \cup B)^c,$$

which need not hold. Unfortunately, none of the centroids considered in this paper satisfies the above property, as the following example shows:

Example 11 Consider $\mathcal{X} = \{x_1, x_2, x_3\}$ and let μ be the exact capacity, with associated Möbius inverse m, given by

A	$\{x_1\}$	$\{x_2\}$	$\{x_3\}$	$\{x_1, x_2\}$	$\{x_1, x_3\}$	$\{x_2, x_3\}$	$\mathcal{X}$
$\mu(A)$	0.18	0.2	0.22	0.54	0.45	0.6	1
$m(A)$	0.18	0.2	0.22	0.16	0.05	0.18	0.01

Fig. 7 Graphical representation of the core of μ in Example 3
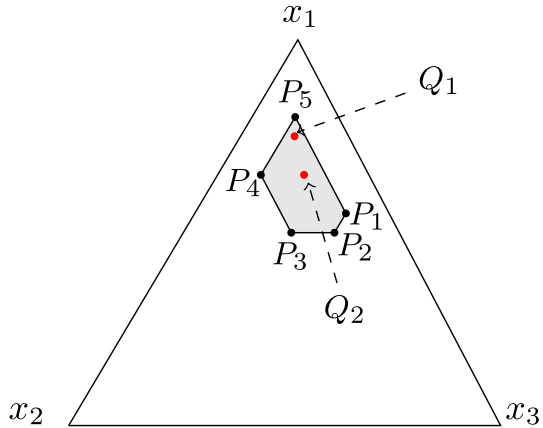
that satisfies $\mu(\{x_3\}) > \mu(\{x_2\}) > \mu(\{x_1\})$. Using Eq. (3), we obtain that $\Phi_1^\mu = (0.288\overline{3}, 0.37\overline{3}, 0.338\overline{3})$, for which $\Phi_1^\mu(\{x_2\}) > \Phi_1^\mu(\{x_3\}) > \Phi_1^\mu(\{x_1\})$. Moreover, in this case all permutations of $\mathcal{X}$ produce a different extreme point, whence $\Phi_2^\mu = \Phi_1^\mu$.

With respect to the incenter, it can be checked that the largest value α such that $\mathrm{core}(\mu)_\alpha \neq \emptyset$ is $\alpha = 0.11$, from which we deduce that Ψ_3^μ is given by all the convex combinations of $(0.29, 0.36, 0.35)$ and $(0.29, 0.38, 0.33)$, all of which produce the same order between x_1, x_2 and x_3 as Φ_1^μ. This also implies that $\Phi_4^\mu = (0.29, 0.37, 0.34)$. Hence, none of the centroids keeps the same preferences as the original capacity. ♦

5 Centrality measures

More generally, instead of determining which element of the core can be considered its center, we may define a *centrality measure*, that allows us to quantify how far inside the core an element is.

Consider for instance the same capacity as in Example 3, whose core is depicted in Fig. 7. Intuitively, given the probability measures Q_1 and Q_2 defined as:

$$Q_1(\{x_1\}) = 0.75, \; Q_1(\{x_2\}) = Q_1(\{x_3\}) = 0.125,$$
$$Q_2(\{x_1\}) = 0.65, \; Q_2(\{x_2\}) = 0.15, \; Q_3(\{x_3\}) = 0.2, \tag{15}$$

and emphasised in red in Fig. 7, Q_2 should have a greater *centrality degree* than Q_1.

This simple example suggests the following definition of centrality measure.

Definition 6 Given an exact capacity μ whose core satisfies $|\,\mathrm{core}(\mu)\,| > 1$, a *centrality measure* is a function $\varphi : \mathbb{P}(\mathcal{X}) \to [0, 1]$ satisfying the following properties:

CM1 $\varphi(P) = 0$ for every $P \notin \mathrm{core}(\mu)$.
CM2   If $P \in \mathrm{ext}\big(\mathrm{core}(\mu)\big)$, then $\varphi(P) = 0$.
CM3   There exists a unique $P_0 \in \mathrm{core}(\mu)$ satisfying $\varphi(P_0) = 1$. Such P_0 is called *central point* in $\mathrm{core}(\mu)$ with respect to φ.
CM4   Consider $P \in \mathrm{ext}\big(\mathrm{core}(\mu)\big)$, P_0 the probability given in the previous item and $\lambda, \beta \in [0, 1]$ such that $\lambda \geq \beta$. Given $P_1 = \lambda P + (1 - \lambda)P_0$ and $P_2 = \beta P + (1 - \beta)P_0$, it holds that $\varphi(P_1) \leq \varphi(P_2)$.

The idea underlying these properties is the following: **CM1** tells us that an element outside the core should have degree of centrality zero; from **CM2**, the same should hold for the extreme points of the core; **CM3** means that there is a unique probability P_0 with degree of centrality 1; finally, property **CM4** represents the idea that the closer a probability is to P_0, the greater its degree of centrality is. We should mention that in Definition 6, and also in the remainder of this section, we are not considering the case where the core is a singleton, $\text{core}(\mu) = \{P_0\}$, because in that case we can trivially assign a centrality degree 1 to P_0 and 0 to any other probability measure.

We next discuss two possible strategies for defining a centrality measure. The first one consists in considering a centroid in the interior of the core, and to measure the distance with respect to it. It requires to specify both the centroid and the distance. Out of the options considered in the previous section, we would reject Φ_1^{μ} because it may not belong to the core and Φ_3^{μ} because of non-uniqueness. With respect to the distance, we consider here the total variation, although as argued before it would also be possible to consider other options such as the L_1 or the Euclidean distances, or even the Kullback Leibler divergence.

In this sense, if we let Φ^{μ} be our centroid of choice and take[6]

$$\beta = \min \left\{ d(\Phi^{\mu}, P_i) \mid P_i \in \text{ext}(\text{core}(\mu)) \right\}, \tag{16}$$

then we can define

$$\varphi_1(P) = 1 - \min \left\{ \frac{d(P, \Phi^{\mu})}{\beta}, 1 \right\} \quad \forall P \in \text{core}(\mu) \tag{17}$$

and $\varphi_1(P) = 0$ otherwise.

A second approach would consist in defining directly a chain $\{\text{core}(\mu)_\alpha\}_{\alpha \in [0,1]}$ of sets such that $\text{core}(\mu)_0 = \text{core}(\mu)$, $\text{core}(\mu)_1$ is a singleton determining the centroid Φ and where $\text{core}(\mu)_\alpha$ is included in the interior of $\text{core}(\mu)$ for any $\alpha > 0$, and letting

$$\varphi_2(P) = \sup \left\{ \alpha \in [0, 1] \mid P \in \text{core}(\mu)_\alpha \right\}. \tag{18}$$

The chain $\{\text{core}(\mu)_\alpha\}_{\alpha \in [0,1]}$ of sets could be defined, for example, as:

$$\text{core}(\mu)_\alpha = CH\left(\left\{ (1 - \alpha)P + \alpha\Phi^{\mu} \mid P \in \text{ext}(\text{core}(\mu)) \right\}\right) \quad \forall \alpha \in [0, 1], \tag{19}$$

where CH denotes the *convex hull*. Let us show that both approaches lead to a centrality measure.

Proposition 9 *Let μ be an exact capacity, and let φ_1, φ_2 be given by Eqs. (17) and (18). Then φ_1, φ_2 satisfy conditions (CM1)–(CM4).*

Example 12 Consider again our running Example 3. The extreme points of $\text{core}(\mu)$ were given in Example 8. Taking as centroid the average of the extreme points Φ_2^{μ}, given in Eq. (14), we obtain the following distances:

$d(P_i, \Phi_2^{\mu})$	P_1	P_2	P_3	P_4	P_5
	0.12	0.12	0.1	0.13	0.2

---

[6] Recall that we are dealing with exact capacities and therefore the core has a finite number of extreme points. This guarantees that the value β in Eq. (16) is a minimum and also that $\beta > 0$, since Φ^{μ} does not coincide with any extreme point of $\text{core}(\mu)$.
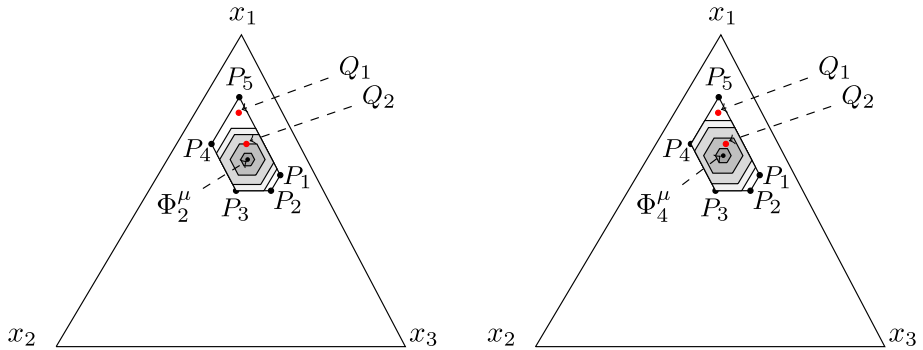
Fig. 8 Centrality measure φ_1 defined by Eq. (17) using Φ_2^μ (left hand side) and Φ_4^μ (right hand side) as centroids, with the curves with centrality degree 0, 0.2, 0.5 and 0.8

Eq. (16) gives $\beta = 0.1$, whence $\varphi_1(P) = 1 - \min\{10d(P, \varphi_2^\mu), 1\}$. Considering the probabilities Q_1 and Q_2 in Eqs. (15), (17) produces the following centrality degrees:

$$\varphi_1(Q_1) = 1 - \min\{10d(Q_1, \varphi_2^\mu), 1\} = 1 - 1 = 0,$$

$$\varphi_1(Q_2) = 1 - \min\{10d(Q_2, \varphi_2^\mu), 1\} = 1 - 0.5 = 0.5,$$

using that $d(Q_1, \varphi_2^\mu) = 0.15$ and $d(Q_2, \varphi_2^\mu) = 0.05$.

If we choose instead as centroid Φ_4^μ, given in Eq. (11), we obtain the following distances to the extreme points:

	P_1	P_2	P_3	P_4	P_5
$d(P_i, \Phi_4^\mu)$	0.1375	0.1375	0.1125	0.1125	0.1875

Thus, in this case $\beta = 0.1125$ and $\varphi_1(P) = 1 - \min\left\{\frac{d(P, \varphi_4^\mu)}{0.1125}, 1\right\}$. Considering again the probability measures Q_1 and Q_2 from Eq. (15) we obtain:

$$\varphi_1(Q_1) = 1 - \min\left\{\frac{d(Q_1, \varphi_4^\mu)}{0.1125}, 1\right\} = 1 - 1 = 0.$$

$$\varphi_1(Q_2) = 1 - \min\left\{\frac{d(Q_2, \varphi_4^\mu)}{0.1125}, 1\right\} = 1 - \frac{0.0375}{0.1125} = \frac{2}{3}.$$

We can see that the centrality degree of Q_1 is zero in both cases (for the centroids Φ_2^μ and Φ_4^μ) but for Q_2, the centrality degree is slightly greater when considering the contraction centroid Φ_4^μ.

Figure 8 shows the curves with centrality degree 0, 0.2, 0.5 and 0.8 for φ_1 when considering as centroid φ_2^μ (left hand side figure) and φ_4^μ (right hand side figure). We can also consider the centrality measure φ_2 defined using the chain of sets in Eq. (19) defined using the extreme points. In that case, taking the average of the extreme points Φ^μ as centroid, the centrality degrees of the probability measures Q_1 and Q_2 are $\varphi_2(Q_1) \approx 0.1923$ and $\varphi_2(Q_2) \approx 0.7142$. In contrast, if we use the contraction centroid, we obtain $\varphi_2(Q_1) = 2/9$ and $\varphi_2(Q_2) = 2/3$.

It is worth mentioning here that for this second approach, any $P \in \text{int}(\text{core}(\mu))$ has a strictly positive centrality degree. This is in contrast with the centrality measure φ_1, that
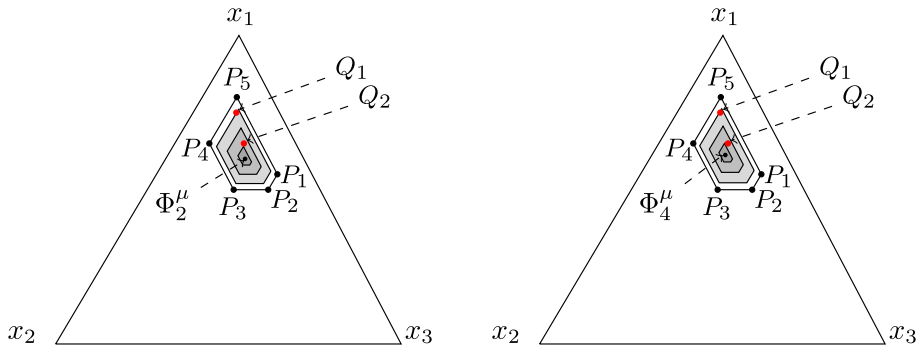
Fig. 9 Centrality measure φ_2 from Eq. (19) using Φ_2^μ (left hand side) and Φ_4^μ (right hand side) as centroids, with the curves with centrality degree 0, 0.2, 0.5 and 0.8

assigns zero centrality degree for some probability measures in the interior of the core of μ, as for example for Q_1.

Figure 9 shows the curves of centrality degree 0.2, 0.5 and 0.8 for φ_2 when considering as centroid Φ_2^μ (left hand side) and Φ_4^μ (right hand side). ♦

It is also possible to define a centrality measure by considering the chain of sets from Eq. (10). For this, note that for each $P \in \mathrm{core}(\mu)$ there is $j \in \{1, \ldots, l\}$ such that $P \in \mathrm{core}(\mu)_{\alpha_{S_{j-1}}} \setminus \mathrm{core}(\mu)_{\alpha_{S_j}}$, where $\mathrm{core}(\mu)_{\alpha_0} = \mathrm{core}(\mu)$. Also, there is $\alpha \in \Lambda_{j-1}$ such that $P \in \big(\mathrm{core}(\mu)_{j-1}\big)_\alpha$, but $P \notin \big(\mathrm{core}(\mu)_{j-1}\big)_{\alpha+\varepsilon}$ for any $\varepsilon > 0$. Then we let:

$$\varphi_3(P) = \frac{\alpha_{S_1} + \cdots + \alpha_{S_{j-1}} + \alpha}{\alpha_{S_1} + \cdots + \alpha_{S_l}} \text{ if } P \in \mathrm{core}(\mu), \tag{20}$$

and $\varphi_3(P) = 0$ if $P \notin \mathrm{core}(\mu)$.

Proposition 10 *The function φ_3 defined in Eq. (20) satisfies conditions (CM1)–(CM4).*

Example 13 In our running Example 3, and in particular in Example 5, we have seen that $\alpha_{S_1} = 0.075$ and $\alpha_{S_2} = 0.0375$, so in this case we have $\mathrm{core}(\mu)_{\alpha_{S_1}}$ and $\mathrm{core}(\mu)_{\alpha_{S_2}}$, where the latter is a singleton formed by Φ_4^μ. It holds that $Q_1 \in \big(\mathrm{core}(\mu)_{\alpha_{S_0}}\big)_{0.025}$ and $Q_1 \notin \big(\mathrm{core}(\mu)_{\alpha_{S_0}}\big)_{0.025+\varepsilon}$, whence:

$$\varphi_3(Q_1) = \frac{0.025}{\alpha_{S_1} + \alpha_{S_2}} = \frac{0.025}{0.075 + 0.0375} = \frac{2}{9}.$$

On the other hand, $Q_2 \in \big(\mathrm{core}(\mu)_{\alpha_{S_0}}\big)_{0.05}$ but $Q_2 \notin \big(\mathrm{core}(\mu)_{\alpha_{S_0}}\big)_{0.05+\varepsilon}$, whence:

$$\varphi_3(Q_2) = \frac{0.05}{\alpha_{S_1} + \alpha_{S_2}} = \frac{0.05}{0.075 + 0.0375} = \frac{4}{9}.$$

Figure 10 depicts the curves of centrality degree 0.2, 0.5 and 0.8 for φ_3. ♦

6 Centroids from the perspective of imprecise probabilities

We conclude this paper by considering the centroid problem within the framework of imprecise probabilities, that include capacities as a particular case and that are capable of modelling

Fig. 10 Centrality measure φ_3 from Eq. (20) with the curves with centrality degree 0, 0.2, 0.5 and 0.8
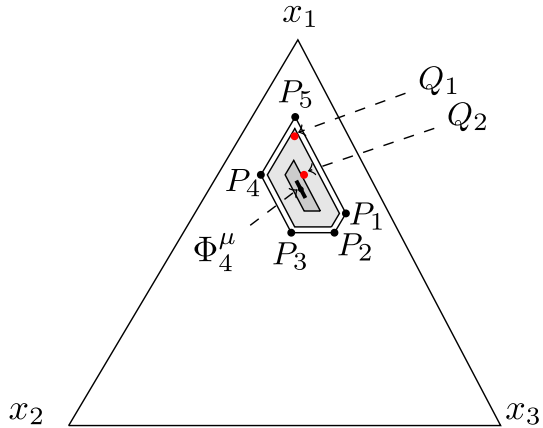

Table 1 Summary of the connection between the terminologies employed in coalitional game theory, decision making and imprecise probability theory

Coalitional game theory	Decision making	Imprecise probabilities
Normalised game (μ)	Normalised capacity (μ)	Lower probability ($\underline{P}$)
Conjugate ($\bar{\mu}$)	Conjugate ($\bar{\mu}$)	Conjugate upper probability ($\overline{P}$)
Balanced game	Balanced capacity	Lower probability avoiding sure loss
Exact game	Exact capacity	Coherent lower probability
Convex game	Supermodular capacity	2-monotone lower probability
Core (core(μ))	Core (core(μ))	Credal set ($\mathcal{M}(\underline{P})$)

more general scenarios. Before we do this, we make a number of clarifications about the terminology.

Within imprecise probability theory, a (exact) capacity μ is usually denoted by $\underline{P}$, and it is called *(coherent) lower probability* while its conjugate function, called *(coherent) upper probability*, is denoted by $\overline{P}$. $\underline{P}$ and $\overline{P}$ may be understood as functions giving lower and upper bounds to a real but unknown probability P_0, meaning that all we know about P_0 is that $\underline{P}(A) \leq P_0(A) \leq \overline{P}(A)$ for any event $A \subseteq \mathcal{X}$. Following this interpretation, the core of a lower probability $\underline{P}$ is called *credal set* (Levi, 1980) and it is denoted by $\mathcal{M}(\underline{P})$; it may be interpreted as the set of probability measures that are compatible with the information given by the lower probability. Finally, in this context the property of supermodularity is usually called 2-monotonicity.

The correspondence[7] between the terminology used in decision making, game theory and imprecise probabilities can be seen in Table 1.

In this section, we shall recall the basics from the more general theory of (coherent) lower previsions, and show that the four centroids analysed before can be also considered in this context.

---

[7] We refer to Miranda and Montes (2018) and also Grabisch (2016) for some comments on the connections between game theory and imprecise probabilities.

6.1 (Coherent) lower previsions

In the theory of imprecise probabilities from Walley (1991), rather than giving lower bounds for the values taken by some unknown probability measure on events, we give lower bounds for the values taken by its expectation operator. This is done by means of a *lower prevision*, a function $\underline{P} : \mathcal{L}(\mathcal{X}) \to \mathbb{R}$, where $\mathcal{L}(\mathcal{X})$ denotes the set of random variables, or *gambles*, defined on $\mathcal{X}$. Its conjugate *upper prevision* is defined by $\overline{P}(f) = -\underline{P}(-f)$ for any $f \in \mathcal{L}(\mathcal{X})$. One underlying interpretation is that there exists a probability measure P_0 modeling our uncertainty, and all we know about it is that $\underline{P}(f) \leq E_{P_0}(f) \leq \overline{P}(f)$ for any $f \in \mathcal{L}(\mathcal{X})$. Associated with a lower prevision we can define a *credal set* by:

$$\mathcal{M}(\underline{P}) = \{P \in \mathbb{P}(\mathcal{X}) \mid P(f) \geq \underline{P}(f) \ \ \forall f \in \mathcal{L}(\mathcal{X})\}, \tag{21}$$

using for simplicity the same symbol P to denote a probability measure and its associated expectation operator: $P(f) = E_P(f)$. The lower prevision $\underline{P}$ is called *coherent* when $\underline{P}(f) = \min_{P \in \mathcal{M}(\underline{P})} P(f)$ for any $f \in \mathcal{L}(\mathcal{X})$. The credal set associated with a coherent lower prevision is a closed and convex subset of $\mathbb{P}(\mathcal{X})$ but it may not be a polytope. In fact, there is a one-to-one correspondence between coherent lower previsions and closed and convex subsets of $\mathbb{P}(\mathcal{X})$. This allows us to understand the extent of the generality of this theory: while coherent lower probabilities (or exact capacities) give rise to credal sets (or cores) that are polytopes, coherent lower previsions induce closed and convex sets of probabilities that need not be a polytope.

One particular situation where the credal set in Eq. (21) is a polytope is when the coherent lower prevision $\underline{P}$ satisfies 2-monotonicity (Walley, 1981):

$$\underline{P}(f \vee g) + \underline{P}(f \wedge g) \geq \underline{P}(f) + \underline{P}(g) \ \ \forall f, g \in \mathcal{L}(\mathcal{X}),$$

where $\vee$ and $\wedge$ denote the pointwise maximum and minimum. When this property is satisfied, the 2-monotone lower prevision determines the same credal set as its restriction to events, which is a 2-monotone lower probability $\underline{P}'$ defined by $\underline{P}'(A) := \underline{P}(I_A)$; and the latter determines the values of the coherent lower prevision $\underline{P}$ by means of the Choquet integral (Choquet, 1953).

6.2 Centroids for coherent lower previsions

Let us discuss how the different notions of centroids we have analysed in this paper may be applied on arbitrary credal sets or, equivalently, on coherent lower previsions.

Shapley value
The Shapley value can be straightforwardly defined by considering the coherent lower probability associated to the lower prevision (its restriction to indicators of events) and applying any of the equivalent representations of the Shapley value given in Sect. 3.1.

Nevertheless, an important drawback here is that two different coherent lower previsions with the same restriction to events will have the same Shapley value, and so will be indistinguishable in this respect. This is illustrated in our next example:

Example 14 Consider $\mathcal{X} = \{x_1, x_2, x_3\}$ and the coherent lower previsions $\underline{P}_1$ and $\underline{P}_2$ inducing the following credal sets:[8]

$$\mathcal{M}(\underline{P}_1) = \big\{ P \in \mathbb{P}(\mathcal{X}) \mid 0.5 \geq P(\{x_2\}) \geq P(\{x_3\}) \big\}.$$
$$\mathcal{M}(\underline{P}_2) = \big\{ P \in \mathbb{P}(\mathcal{X}) \mid 0.5 \geq P(\{x_3\}) \geq P(\{x_2\}) \big\}.$$

By taking lower envelopes, we obtain that both induce the same coherent lower probability $\underline{P}$:

A	$\{x_1\}$	$\{x_2\}$	$\{x_3\}$	$\{x_1, x_2\}$	$\{x_1, x_3\}$	$\{x_2, x_3\}$	$\mathcal{X}$
$\underline{P}(A)$	0	0	0	0.5	0.5	0	1

Hence, the Shapley value is the same for both $\underline{P}_1$ and $\underline{P}_2$, and it is given by $\Phi_1^{P_1} = \Phi_1^{P_2} = (0.5, 0.25, 0.25)$. Note that the Shapley value does not belong to the interior of the credal sets for neither $\underline{P}_1$ nor $\underline{P}_2$. ♦

Average of the extreme points
Whenever the credal set $\mathcal{M}(\underline{P})$ is a polytope (i.e., it has a finite number of extreme points), the average of the extreme points can be computed using Eq. (6). While this definition imposes a restriction on the credal set, it is applicable for those associated with a lower probability $\underline{P}$ that is coherent (Wallner, 2007), and therefore also in the particular cases of 2-monotonicity, belief functions or p-boxes (Montes & Destercke, 2017). In addition, it is also applicable to some models of coherent lower previsions that are not determined by their restrictions to events, such as those associated with comparative probabilities (Miranda & Destercke, 2015).[9]

Incenter
As we did in Sect. 3.3, we can find the (set of) incenters. In this case, the definition of incenter can be straightforwardly given: the incenter radius of $\mathcal{M}(\underline{P})$ is given by

$$\alpha_I = \sup \big\{ \alpha > 0 \mid \exists P_0 \in \mathcal{M}(\underline{P}) \text{ such that } B_o^\alpha(P_0) \subseteq \mathcal{M}(\underline{P}) \cap \mathbb{P}^*(\mathcal{X}) \big\},$$

and any $P_0 \in \mathcal{M}(\underline{P})$ such that $B_0^{\alpha_I}(P_0) \subseteq \mathcal{M}(\underline{P})$ is called incenter of $\underline{P}$.

Proposition 11 *Let $\underline{P}$ be a coherent lower prevision whose credal set has non-empty interior. Then the value α_I is a maximum. As a consequence, the incenter of $\underline{P}$ always exists.*

Contraction centroid
The only centroid whose extension to coherent lower previsions is not straightforward is the contraction centroid. Assuming again that $\mathcal{M}(\underline{P})$ is a polytope, we know that it is determined by a finite number of constraints. This means that there are two (disjoint) set of gambles $\mathcal{L}^>$ and $\mathcal{L}^=$ such that the coherent lower prevision $\underline{P}$ and its conjugate $\overline{P}$ satisfy:

$$\underline{P}(f) = \overline{P}(f) \ \ \forall f \in \mathcal{L}^= \text{ and } \underline{P}(f) < \overline{P}(f) \ \ \forall f \in \mathcal{L}^>,$$

and that the credal set can be expressed as:

$$\mathcal{M}(\underline{P}) = \big\{ P \in \mathbb{P}(\mathcal{X}) \mid P(f) = \underline{P}(f) \, \forall f \in \mathcal{L}^=, \ P(f) \geq \underline{P}(f) \, \forall f \in \mathcal{L}^> \big\}. \qquad (22)$$

---

[8] As we have explained, any coherent lower prevision is in one-to-one correspondence with a closed and convex subset of $\mathbb{P}(\mathcal{X})$.

[9] One possibility for credal sets that are not polytopes would be to consider the average with respect to a uniform distribution over the infinite family of extreme points; this uniform distribution might be defined letting go of countable additivity, considering the comments in Walley (1991, Sec.4 2.9). A deeper study of this matter is left as future work.

Note that we may assume without loss of generality that these constraints include the indicator functions of the proper events: $\{I_A \mid \emptyset \neq A \subset \mathcal{X}\} \subseteq \mathcal{L}^> \cup \mathcal{L}^=$. In that case, when $\mathcal{L}^>$ is empty we obtain that $\underline{P}(I_A) = \overline{P}(I_A)$ for any $A \subseteq \mathcal{X}$ (or $\underline{P}(A) = \overline{P}(A)$, if we use this abuse of notation), meaning that $\mathcal{M}(\underline{P})$ contains one single probability measure. Moreover, using the properties of coherent lower and upper previsions, we can also assume without loss of generality that $0 \leq \min f < \max f = 1$ for every $f \in \mathcal{L}^> \cup \mathcal{L}^=$.

The idea in this approach is the same explained in Sect. 3.4: we contract the credal set in a uniform manner as long as possible, increasing the value of the lower prevision in a constant amount α in all the gambles $f \in \mathcal{L}^>$, and then proceed in the same way by reducing the number of constraints. As we showed in Proposition 3, there exists a value α small enough such that this approach produces a non-empty credal set and there is a maximum value satisfying this property.

Proposition 12 *Let $\underline{P}$ be a coherent lower prevision whose credal set is a polytope that can be expressed as in Eq. (22). For a given $\alpha > 0$, let:*

$$\mathcal{M}(\underline{P})_\alpha = \left\{ P \in \mathbb{P}(\mathcal{X}) \mid P(f) = \underline{P}(f) \; \forall f \in \mathcal{L}^=, \; P(f) \geq \underline{P}(f) + \alpha \; \forall f \in \mathcal{L}^> \right\}. \quad (23)$$

Consider also the set $\Lambda = \{\alpha > 0 \mid \mathcal{M}(\underline{P})_\alpha \neq \emptyset\}$. It holds that:

(a) $\Lambda \neq \emptyset$.
(b) *The set Λ has a maximum α_S.*
(c) *Given the set $\mathcal{M}(\underline{P})_{\alpha_S}$, there exists some $f \in \mathcal{L}^>$ such that $P(f)$ is constant for any $P \in \mathcal{M}(\underline{P})_{\alpha_S}$.*

Moreover, as we explained before, when the coherent lower prevision satisfies 2-monotoninicty, it is determined by its restriction to events. Hence, Theorem 4 also applies in this context, where we simply need to understand $\underline{P}(A)$ and $\overline{P}(A)$ as the lower and upper previsions of the indicator I_A.

Finally, the connection between the set of incenters and the first step of the process determining the contraction centroid also holds for coherent lower previsions.

Proposition 13 *Let $\underline{P}$ be a coherent lower prevision whose associated credal set $\mathcal{M}(\underline{P})$ is included in $\mathbb{P}^*(\mathcal{X})$ and such that $\mathcal{L}^= = \emptyset$. If α_S is the incenter radius, then for any $P_0 \in \mathcal{M}(\underline{P})$ and any $\alpha \leq \alpha_S$:*

$$P_0 \in \mathcal{M}_\alpha(\underline{P}) \Leftrightarrow B_c^\alpha(P_0) \subseteq \mathcal{M}(\underline{P}).$$

6.3 Particular cases

We have seen that the four centroids can be defined to coherent lower previsions. In this subsection we analyse them for some particular families of imprecise models. We start with *probability intervals*.

A probability interval (de Campos et al., 1994) is an uncertainty model $\mathcal{I}$ that gives lower and upper bounds to the probability of the singletons:

$$\mathcal{I} = \left\{ [l_i, u_i] \mid l_i \leq u_i \; \forall i = 1, \ldots, n \right\}.$$

It determines a credal set given by:

$$\mathcal{M}(\mathcal{I}) = \left\{ P \in \mathbb{P}(\mathcal{X}) \mid l_i \leq P(\{x_i\}) \leq u_i \; \forall i = 1, \ldots, n \right\}.$$

This credal set is non-empty if and only if $\sum_{i=1}^n l_i \leq 1 \leq \sum_{i=1}^n u_i$, and we say that the probability interval *avoids sure loss*. Then, taking lower and upper envelopes of $\mathcal{M}(\mathcal{I})$ we obtain a lower and an upper probability, and we say that the probability interval is *coherent* when $\underline{P}(\{x_i\}) = l_i$ and $\overline{P}(\{x_i\}) = u_i$ for every $i = 1, \ldots, n$. In that case, $\underline{P}$ is 2-monotone and the values of $\underline{P}$ and $\overline{P}$ for any event $A \subseteq \mathcal{X}$ can be computed as de Campos et al. (1994):

$$\underline{P}(A) = \max\left\{\sum_{x_i \in A} l_i, 1 - \sum_{x_i \notin A} u_i\right\}, \quad \overline{P}(A) = \min\left\{\sum_{x_i \in A} u_i, 1 - \sum_{x_i \notin A} l_i\right\}.$$

For the particular case of coherent probability intervals, we can give an explicit formula for the value α_S in the contraction method.

Proposition 14 *Let $\underline{P}$ and $\overline{P}$ be the coherent lower and upper probability determined by a coherent probability interval $\mathcal{I} = \{[l_i, u_i] \mid \forall i = 1, \ldots, n\}$. Consider:*

$$\mathcal{I}^> = \left\{i \in \{1, \ldots, n\} \mid l_i < u_i\right\}, \quad \mathcal{I}^= = \left\{i \in \{1, \ldots, n\} \mid l_i = u_i\right\}.$$

Then:

(a) *The value $\alpha_S = \max \Lambda$ is given by:*

$$\alpha_S = \min\left\{\frac{1}{|\mathcal{I}^>|}\left(1 - \sum_{i=1}^n l_i\right), \frac{1}{|\mathcal{I}^>|}\left(\sum_{i=1}^n u_i - 1\right), \frac{1}{2}\min_{i \in \mathcal{I}^>}(u_i - l_i)\right\}. \quad (24)$$

(b) *The credal set $\mathcal{M}(\underline{P})_{\alpha_S}$ determined by means of Eq. (23) is a probability interval avoiding sure loss.*

(c) *If $\alpha_S = \frac{1}{|\mathcal{I}^>|}\left(1 - \sum_{i=1,\ldots,n} l_i\right)$ or $\alpha_S = \frac{1}{|\mathcal{I}^>|}\left(\sum_{i=1,\ldots,n} u_i - 1\right)$, then $\mathcal{M}(\underline{P})_{\alpha_S} = \left\{\Phi_4^{\underline{P}}\right\}$.*

The value α_S obtained in Eq. (24) is consistent with that from Theorem 5 for the particular case of maximally imprecise probability intervals. This shows also that in that case Eq. (13) can be simplified, in the sense that we do not need to consider all partitions of $\mathcal{X}$, but only the partitions $\{x_1\}, \ldots, \{x_n\}$ and $\{x_i\}, \mathcal{X} \setminus \{x_i\}$ for every $i = 1, \ldots, n$.

From Proposition 14 we can also deduce an explicit formula for the value α_S for the *Linear Vacuous* (LV) and the *Pari Mutuel Model* (PMM), which constitute particular instances of *distortion models* (Montes et al., 2020a, b) or *nearly linear models* (Corsato et al., 2019). The PMM (Montes et al., 2019; Pelessoni et al., 2010; Walley, 1991) is determined by the coherent lower probability:

$$\underline{P}_{PMM}(A) = \max\left\{(1 + \delta)P_0(A) - \delta, 0\right\} \quad \forall A \subseteq \mathcal{X},$$

where $P_0 \in \mathbb{P}(\mathcal{X})$ is a given probability measure and $\delta > 0$. Similarly, the LV (Walley, 1991) is defined by the coherent lower probability

$$\underline{P}_{LV}(A) = (1 - \delta)P_0(A) \quad \forall A \subset \mathcal{X} \quad \text{and} \quad \underline{P}_{LV}(\mathcal{X}) = 1,$$

where $P_0 \in \mathbb{P}(\mathcal{X})$ and $\delta \in (0, 1)$.

Both the PMM and the LV are instances of probability intervals, where:

$$\mathcal{I}_{PMM} = \left\{\left[\max\{(1 + \delta)P_0(\{x\}) - \delta, 0\}, \min\{(1 + \delta)P_0(\{x\}), 1\}\right] \mid x \in \mathcal{X}\right\},$$

$$\mathcal{I}_{LV} = \left\{\left[(1 - \delta)P_0(\{x\}), \min\{(1 - \delta)P_0(\{x\}) + \delta, 1\}\right] \mid x \in \mathcal{X}\right\}.$$

This means that we can apply Proposition 14 for computing the value $\alpha_S = \max \Lambda$. In fact, when both P_0 and the lower probability only take the values 0 and 1 for trivial events

(the impossible and sure events), the formula for $\alpha_S = \max \Lambda$ can be simplified and the procedure of contracting the credal set finishes in only one step.

Corollary 15 *Consider a PMM $\underline{P}_{PMM}$ or a LV $\underline{P}_{LV}$ determined by a probability measure $P_0 \in \mathbb{P}(\mathcal{X})$ and a distortion parameter δ. Assume that $\underline{P}_{PMM}(A)$ and $\underline{P}_{LV}(A)$ belong to $(0, 1)$ for every $A \neq \emptyset, \mathcal{X}$. Then:*

(a) *For the PMM, $\alpha_S = \max \Lambda = \frac{\delta}{n}$ and $\mathcal{M}_\alpha = \left\{\Phi_4^{P_{PMM}}\right\}$ where $\Phi_4^{P_{PMM}}(\{x_i\}) = (1 + \delta)P_0(\{x_i\}) - \frac{\delta}{n}$ for any $i = 1, \ldots, n$.*

(b) *For the LV, $\alpha_S = \max \Lambda = \frac{\delta}{n}$ and $\mathcal{M}_\alpha = \left\{\Phi_4^{P_{LV}}\right\}$ where $\Phi_4^{P_{LV}}(\{x_i\}) = (1-\delta)P_0(\{x_i\}) + \frac{\delta}{n}$ for any $i = 1, \ldots, n$.*

(c) *For both the PMM and LV there is a unique incenter and $\Phi_1 = \Phi_2 = \Phi_3 = \Phi_4$.*

In this respect, it is worth remarking that (i) the good behaviour of these two distortion models is in line of other desirable properties they possess, as discussed in Destercke et al. (2022), Montes et al. (2020a) and Montes et al. (2020b); and (ii) the centroid of a distortion model originated by a probability measure P_0 does not coincide with P_0, because the distortion is not done uniformly in all directions of the simplex. This was already observed in Miranda and Montes (2018 for the particular case of the Shapley value.

6.4 Properties

We consider now the properties of the centroids considered in Sect. 4.

Proposition 16 *Let $\underline{P}$ be a coherent lower prevision. Then, the properties in Proposition 7 and 8 still hold.*

We have already mentioned that two different lower previsions may have the same restriction to events. For this reason, in addition to the aforementioned properties, it would be desirable that the center of a coherent lower prevision $\underline{P}$ does not necessarily coincide with the center of its restriction to events. In this respect, it is not difficult to show that $\Psi_3^{\frac{P}{}}, \Phi_4^{\frac{P}{}}$ are capable of distinguishing between lower previsions and lower probabilities, and so does $\Phi_2^{\frac{P}{}}$ (when $\mathcal{M}(\underline{P})$ is a polytope). On the other hand, Shapley value is only defined via the lower probability, so it does not distinguish between lower probabilities and lower previsions as we showed in Example 14.

6.5 Centrality measures

Finally, we may try to define centrality measures for the credal set determined by a coherent lower prevision. For this aim, we can consider exactly the definition of centrality measure as in Definition 6, and the centrality measures φ_1, φ_2 and φ_3 defined in Eqs. (17), (18) and (20), respectively. Note that in the case of φ_1 and φ_3 we need to restrict ourselves to coherent lower previsions whose credal set is a polytope to assure that the minimum in Eq. (16) is strictly positive and that the contraction approach finishes in a finite number of steps, respectively.

Proposition 17 *Given a coherent lower prevision $\underline{P}$, the function φ_2 is a centrality measure. Moreover, if $\mathcal{M}(\underline{P})$ is a polytope, φ_1 and φ_3 are centrality measures too.*

Table 2 Summary of the properties satisfied by the four centroids

Property	Shapley value	Average of extremes	Incenter (d_{TV})	Contraction center
Efficiency	Yes	Yes	Yes	Yes
Symmetry	Yes	Yes	Yes	Yes
Linearity	Yes	No	No	No
Null player	Yes	Yes	Yes	Yes
Feasibility	No	Yes	Yes	Yes
Continuity	Yes	No	Yes	Yes
Ignorance preservation	Yes	Yes	Yes	Yes
Preference preservation	No	No	No	No
Distinguish lower probabilities and previsions	No	Yes	Yes	Yes

7 Conclusions

We have performed a comparative analysis of four alternatives for defining a center of an exact capacity: the Shapley value, the average of the extreme points, the incenter for the total variation distance and the limit of uniform contractions. Our results show that these four approaches may lead to different results, and also illustrate some of the properties each of them satisfies: a summary can be seen in Table 2. Note that our goal with this paper is not to take the stance that one of them is *better* than the others. Instead, we intend to provide some assistance to the practitioner in her choice of a centroid: if for instance she considers that linearity is an essential property, she must pick the Shapley value; if she wants to ensure that the center belongs to the interior of the core of the capacity, she must select one of the others; and so on. Also, we have seen that these centroids can also be applied in the more general framework of coherent lower previsions. Since coherent lower previsions are in one-to-one correspondence with closed and convex set of probabilities, they include the particular case of polytopes. We have seen that the results we have proved for exact capacities can also be extended to coherent lower previsions.

Let us recall that the center in this paper is understood as a probability measure that is in the interior of the core of the capacity and that can play the role of its representative. For this reason, we have left out of our study other approaches based, for example, on maximising the entropy or minimising the Kullback Leibler divergence, that in our context may produce counterintuitive results.

In addition to the comparison performed between the four centroids, some comments regarding their computation in practice must be done:

- First of all, the computation of the Shapley value is known to be a hard problem that exponentially increases with $| \mathcal{X} |$, since it requires the computation of $\mu(A)$ for the $2^n - 1$ non-trivial events.
- Secondly, the average of the extreme points is simple as long as these are known. Under the assumption of supermodularity, the extreme points coincide with the probability measures P_σ defined in Eq. (1). Even more, there are particular situations where their computation is even simpler: for example, when μ is minitive there are at most 2^{n-1} (Miranda et al., 2003) and when μ corresponds to a p-box the maximal number of extreme points is at most the n-th Pell number (Montes & Destercke, 2017). The problem is more challenging

when the capacity is not supermodular: even if the number of extreme points is at most $n!$ (Wallner, 2007), there is no simple procedure for their computation.

- Thirdly, computing the contraction centroid may be an extremely complicated problem. Even if Theorem 4 gives a formula for computing the value α_S under fairly general conditions, it requires the computation of all the families $\mathcal{A}$ in $\mathbb{A}(\mathcal{X})$, whose number is extremely large. The complexity is significantly reduced for supermodular capacities, because Theorem 5 gives a simpler expression for α_S that depends on the partitions of $\mathcal{X}$. Still, the problem is rather complex because the number of partitions in a n element possibility space coincides with the n-th Bell number. Nevertheless, we have seen that for particular models that may arise in practice such as probability intervals or some distortion models, the computation of the four centroids is much simpler.
- Finally, under fairly general conditions the set of incenters coincide with the first step of the contraction approach (Proposition 6), hence both approaches are equivalent from the computational viewpoint.

While our results give some overview of the properties of the centroids of a capacity, there is still much work to be done in order to have a full picture of this problem. First and foremost, it would be interesting to extend our approaches to infinite possibility spaces. While this seems immediate in the case of the incenter or the vertex centroid (see also footnote 4), in the case of the Shapley value we should consider the generalisations carried out in Neyman (2002), and in the case of the contraction centroid we should verify that the process stabilises in a finite number of steps. In addition, we could consider other possibilities in the context of game solutions, such as the Banzhaf value (Banzhaf, 1965) or more generally probabilistic solutions (Weber, 1988), or other alternatives to the total variation distance, such as the Euclidean distance or the L_1 distance. It would also be interesting to obtain further conditions for the equality between some of these centroids. And finally, a deeper study of centrality measures and their axiomatic properties would be of interest.

Declarations

Appendix A: Proofs

Proof of Proposition 1 From Eq. (2), using the definition of the Möbius inverse, we obtain that

$$\Phi_1^\mu(\{x\}) = \sum_{A|x\notin A} \frac{|A|!(n-|A|-1)!}{n!}\big(\mu(A\cup\{x\})-\mu(A)\big)$$

$$= \sum_{A|x\notin A} \frac{|A|!(n-|A|-1)!}{n!}\left(\sum_{B\subseteq A\cup\{x\}} m(B)-\sum_{B\subseteq A} m(B)\right)$$

$$= \sum_{A|x\notin A} \frac{|A|!(n-|A|-1)!}{n!}\left(\sum_{x\in B\subseteq A\cup\{x\}} m(B)\right).$$

Now, in the sum above any event B that includes x appears associated with sets A that include $B\setminus\{x\}$ and do not include x, and which are thus of the form $(B\setminus\{x\})\cup C$ for some $C\subseteq B^c$. This means that in the sum above $m(B)$ is multiplied by

$$\left(\sum_{k=0}^{n-|B|} \frac{(|B|-1+k)!(n-(|B|-1+k)-1)!}{n!}\right)\binom{n-|B|}{k}$$

$$= \frac{(n-|B|)!}{n!}\sum_{k=0}^{n-|B|} \frac{(|B|-1+k)!}{k!} = \frac{(n-|B|)!(|B|-1)!}{n!}\sum_{k=0}^{n-|B|}\binom{|B|-1+k}{k}$$

$$= \frac{(n-|B|)!(|B|-1)!}{n!}\binom{n-|B|+|B|-1+1}{n-|B|}$$

$$= \frac{(n-|B|)!(|B|-1)!}{n!}\binom{n}{n-|B|} = \frac{1}{|B|},$$

where the third equality follows from the hockey-stick identity. This shows that the Shapley value coincides with the pignistic transformation.

To see the equality with the average of the P_σ, note that for a fixed permutation,

$$P_\sigma(\{x_{\sigma(i)}\}) = \mu(\{x_{\sigma(1)},\ldots,x_{\sigma(i)}\}) - \mu(\{x_{\sigma(1)},\ldots,x_{\sigma(i-1)}\})$$

$$= \sum_{x_{\sigma(i)}\in B\subseteq\{x_{\sigma(1)},\ldots,x_{\sigma(i)}\}} m(B),$$

whence $m(B)$ is included in the sum that determines the probability of $x_{\sigma(i)}$ only in those permutations when B is a subset of $\{x_{\sigma(1)},\ldots,x_{\sigma(i)}\}$ that includes $x_{\sigma(i)}$. Reasoning as above, the proportion of such permutations is $\frac{1}{|B|}$. Thus, the average of the P_σ coincides with the pignistic transformation, and as a consequence also with the Shapley value. $\square$

Proof of Propositions 2 and 11 Since capacities are particular cases of coherent lower previsions, we prove the statement in Proposition 11, from which we immediately deduce Proposition 2.

Consider a coherent lower prevision $\underline{P}$. If the topological interior of $\mathcal{M}(\underline{P})$ is non-empty, we consider the set

$$\Lambda_1 = \big\{\alpha\in[0,1] \mid \exists P\in\mathcal{M}(\underline{P}) \text{ such that } B_o^\alpha(P)\subseteq\mathcal{M}(\underline{P})\cap\mathbb{P}^*(\mathcal{X})\big\}.$$

This set is non-empty because

$$\text{int}\big(\mathcal{M}(\underline{P})\big) \neq \emptyset \Leftrightarrow \text{int}\big(\mathcal{M}(\underline{P})\big) \cap \mathbb{P}^*(\mathcal{X}) \neq \emptyset$$

and given $P \in \text{int}\big(\mathcal{M}(\underline{P})\big) \cap \mathbb{P}^*(\mathcal{X})$ there exists some $\alpha > 0$ such that $B_o^\alpha(P) \subseteq \mathcal{M}(\underline{P}) \cap \mathbb{P}^*(\mathcal{X})$. Besides being non-empty, the set Λ_1 is directed ($\alpha \in \Lambda_1 \Rightarrow \alpha' \in \Lambda_1 \ \forall \alpha' \in [0, \alpha]$). Let $\alpha_1 := \sup \Lambda_1$. This means that for every $n \in \mathbb{N}$ there exists some $P_n \in \mathcal{M}(\underline{P})$ such that $B_o^{\alpha_1 - 1/n}(P_n) \subseteq \mathcal{M}(\underline{P}) \cap \mathbb{P}^*(\mathcal{X})$. The sequence $(P_n)_n$ is included in the compact set $\mathcal{M}(\underline{P})$, that is as a consequence also sequentially compact. Therefore, there exists a subsequence $(P_{n'})_{n'}$ that converges to some P. Note that for this subsequence we also have that $B_o^{\alpha_1 - 1/n'}(P_{n'}) \subseteq \mathcal{M}(\underline{P}) \cap \mathbb{P}^*(\mathcal{X})$. Since $\mathcal{M}(\underline{P})$ is closed, $P \in \mathcal{M}(\underline{P})$. Moreover, for any $\epsilon > 0$ there is some n_ϵ such that $d(P_{n'}, P) < \epsilon$ for every $n' \geq n_\epsilon$. Take $\epsilon = \frac{1}{n}$ for a fixed n. Then for any $Q \in B_o^{\alpha_1 - 2/n}(P)$ and any $m \geq n_\epsilon$ it holds that

$$d(Q, P_{m'}) \leq d(Q, P) + d(P, P_{m'}) < \alpha_1 - \frac{1}{n}.$$

Since we can take m arbitrarily large, we deduce that

$$B_o^{\alpha_1 - 2/n}(P) \subseteq B_o^{\alpha_1 - 1/n}(P_{m'}) \subseteq B_o^{\alpha_1 - 1/m'}(P_{m'}) \subseteq \mathcal{M}(\underline{P}) \cap \mathbb{P}^*(\mathcal{X}).$$

Since this holds for every n, we deduce that $B_o^{\alpha_1}(P) = \cup_n B_o^{\alpha_1 - \frac{1}{n'}}(P) \subseteq \mathcal{M}(\underline{P}) \cap \mathbb{P}^*(\mathcal{X})$. Therefore, α_1 belongs to Λ_1. $\qquad \square$

Proof of Propositions 3 and 12 Again, we prove Proposition 12 and, since exact capacities are particular cases of coherent lower previsions, Proposition 3 immediately follows.

(a) Let us denote $\mathcal{L}^> = \{f_1, \ldots, f_k\}$. For each $i = 1, \ldots, k$, it follows by coherence that there is some $P_i \in \mathcal{M}(\underline{P})$ such that $P_i(f_i) = \overline{P}(f_i) > \underline{P}(f_i)$. If we now let $P := \frac{P_1 + \cdots + P_k}{k}$, it belongs to the convex set $\mathcal{M}(\underline{P})$ and by construction it satisfies $P(f_i) > \underline{P}(f_i)$ for all $i = 1, \ldots, k$. Given $\alpha = \min_{i=1,\ldots,k}\big(P(f_i) - \underline{P}(f_i)\big) > 0$, we conclude that $P \in \mathcal{M}(\underline{P})_{\alpha_i}$ and therefore Λ is non-empty.

(b) Since $\mathcal{X}$ is finite, the topology of $\mathcal{M}(\underline{P})$ is equivalent to the topology associated with the Euclidean distance. The definition of $\mathcal{M}(\underline{P})_\alpha$ implies then that, for every $\alpha \in \Lambda$, the set $\mathcal{M}(\underline{P})_\alpha$ is a closed subset of $\mathcal{M}(\underline{P})$, and it is therefore compact. Thus, $(\mathcal{M}(\underline{P})_\alpha)_{\alpha \in \Lambda}$ is a decreasing sequence of compact subsets of $\mathcal{M}(\underline{P})$, and as a consequence their intersection $\mathcal{M}^*$ is non-empty. But this intersection $\mathcal{M}^*$ must coincide with $\mathcal{M}(\underline{P})_{\alpha_S}$ for $\alpha_S = \sup \Lambda$, and this implies that this supremum is a maximum.

(c) Let $\underline{P}_{\alpha_S}, \overline{P}_{\alpha_S}$ denote the lower and upper envelopes of $\mathcal{M}(\underline{P})_{\alpha_S}$. Assume ex-absurdo that $\underline{P}_{\alpha_S}(f_i) < \overline{P}_{\alpha_S}(f_i)$ for every $f_i \in \mathcal{L}^>$. Then reasoning as in the first statement we can find some $P \in \mathcal{M}(\underline{P})_{\alpha_S}$ such that $P(f_i) > \underline{P}_{\alpha_S}(f_i)$, and this contradicts that α_S is the maximum value of Λ.

$\qquad \square$

Proof of Theorem 4 We are looking for the maximum α such that the set:

$$\text{core}(\mu)_\alpha = \big\{P \in \mathbb{P}(\mathcal{X}) \mid P(A) \geq \mu(A) + \alpha \ \forall A \neq \emptyset, \mathcal{X}\big\}$$

is non-empty. This is equivalent to requiring that the capacity ν given by:

$$\nu(A) = \begin{cases} 0, & \text{for } A = \emptyset \\ \mu(A) + \alpha, & \text{for } A \neq \emptyset, \mathcal{X} \\ 1, & \text{for } A = \mathcal{X} \end{cases}$$

is balanced. Since the possibility space $\mathcal{X}$ is finite, this means (Walley 1991, Lemma 2.4.4) that for any $l \in \mathbb{N}$ and $A_1, \ldots, A_l$, we should have

$$\max_{x \in \mathcal{X}} \sum_{i=1}^{l} \left(I_{A_i}(x) - \nu(A_i) \right) \geq 0. \tag{A1}$$

Note that we can assume that the events $A_1, \ldots, A_l$ are proper subsets of $\mathcal{X}$, given that $I_{\mathcal{X}} - \nu(\mathcal{X}) = 1 - 1 = 0$ and $I_{\emptyset} - \nu(\emptyset) = 0 - 0 = 0$. This allows us to rewrite Eq. (A1) as

$$\max_{x \in \mathcal{X}} \sum_{i=1}^{l} \left(I_{A_i}(x) - \mu(A_i) - \alpha \right) \geq 0$$

for any $l \in \mathbb{N}$ and any proper subsets $A_1, \ldots, A_l$ of $\mathcal{X}$.

Now, given such sets $A_1, \ldots, A_l$, the gamble $f := I_{A_1} + \ldots + I_{A_l}$ can be rewritten as $f := a_1 I_{B_1} + \ldots + a_m I_{B_m}$ for non-negative integers $a_1 > \ldots > a_{m-1} > a_m \geq 0$ and where $\{B_1, \ldots, B_m\}$ is a partition of $\mathcal{X}$. Let us define the sets $C_i^j := B_i$ for $i = 2, \ldots, m$ and for $j = 1, \ldots, a_1 - a_i$. It holds that

$$\mathcal{A} = \left(A_1, \ldots, A_l, C_2^1, \ldots, C_2^{a_1 - a_2}, \ldots, C_m^1, \ldots, C_m^{a_1 - a_m} \right)$$

belongs to the class $\mathbb{A}(\mathcal{X})$ and $\beta_{\mathcal{A}} = a_1$. Then:

$$\max_{x \in \mathcal{X}} \sum_{i=1}^{l} \left(I_{A_i}(x) - \mu(A_i) - \alpha \right) = \max_{x \in \mathcal{X}} \left(\sum_{i=1}^{l} \left(I_{A_i}(x) - \mu(A_i) \right) \right.$$

$$+ \sum_{i=2}^{m} \sum_{j=1}^{a_1 - a_i} \left(I_{C_i^j}(x) - \mu(C_i^j) \right) - \sum_{i=2}^{m} \sum_{j=1}^{a_1 - a_i} \left(I_{C_i^j}(x) - \mu(C_i^j) \right) - l\alpha \left. \right)$$

$$= \beta_{\mathcal{A}} - \sum_{i=1}^{l} \mu(A_i) - \sum_{i=2}^{m} \sum_{j=1}^{a_1 - a_i} \mu(C_i^j) - \min_{x \in \mathcal{X}} \left(\sum_{i=2}^{m} \sum_{j=1}^{a_1 - a_i} \left(I_{C_i^j}(x) - \mu(C_i^j) \right) \right) - l\alpha.$$

This means that

$$\max_{x \in \mathcal{X}} \sum_{i=1}^{l} \left(I_{A_i}(x) - \mu(A_i) - \alpha \right) \geq 0$$

if and only if

$$\alpha \leq \frac{1}{l} \left(\beta_{\mathcal{A}} - \sum_{i=1}^{l} \mu(A_i) - \sum_{i=2}^{m} \sum_{j=1}^{a_1 - a_i} \mu(C_i^j) \right) - \min_{x \in \mathcal{X}} \frac{1}{l} \sum_{i=2}^{m} \sum_{j=1}^{a_1 - a_i} \left(I_{C_i^j}(x) - \mu(C_i^j) \right). \tag{A2}$$

Now, instead of considering the events $A_1, \ldots, A_l$, consider $A_1, \ldots, A_l, C_2^1$. We get:

$$\frac{1}{l+1} \left(\beta_{\mathcal{A}} - \sum_{i=1}^{l} \mu(A_i) - \sum_{i=2}^{m} \sum_{j=1}^{a_1 - a_i} \mu(C_i^j) \right)$$

$$- \min_{x \in \mathcal{X}} \left[\frac{1}{l+1} \left(\sum_{i=2}^{m} \sum_{j=1}^{a_1 - a_i} (I_{C_j}(x) - \mu(C_j)) - (I_{C_2^1}(x) - \mu(C_2^1)) \right) \right]$$

$$\leq \frac{1}{l} \left(\beta_{\mathcal{A}} - \sum_{i=1}^{l} \mu(A_i) - \sum_{i=2}^{m} \sum_{j=1}^{a_1 - a_i} \mu(C_i^j) \right) - \min_{x \in \mathcal{X}} \frac{1}{l} \sum_{i=2}^{m} \sum_{j=1}^{a_1 - a_i} \left(I_{C_i^j}(x) - \mu(C_i^j) \right).$$

In fact, observe that $\beta_{\mathcal{A}} - \sum_{i=1}^{l} \mu(A_i) - \sum_{i=2}^{m} \sum_{j=1}^{a_1-a_i} \mu(C_i^j) \geq 0$ since μ is an exact capacity, and that the minimum in each equation is given by $- \sum_{i=2}^{m} \sum_{j=1}^{a_1-a_i} \mu(C_i^j) + \mu(C_2^1)$, and $- \sum_{i=2}^{m} \sum_{j=1}^{a_1-a_i} \mu(C_i^j)$ respectively, in both cases considering $x \in B_1$.

Iterating the procedure, the minimum value in the right-hand side part of Eq. (A2) is attained for $A_1, \ldots, A_l, C_2^1, \ldots, C_2^{a_1-a_2}, \ldots, C_m^1, \ldots, C_m^{a_1-a_m}$. With this, we deduce that $\mathcal{M}_{\alpha}(\underline{P})$ is non-empty if and only if $\alpha \leq \min_{\mathcal{A} \in \mathbb{A}} \frac{1}{|\mathcal{A}|} \left(\beta_{\mathcal{A}} - \sum_{i=1}^{l} \mu(A_i) \right)$, and the proof is complete. $\qquad \square$

In order to prove Theorem 5, we must establish first a couple of auxiliary lemmas.

Lemma 18 *Let μ be a maximally imprecise exact capacity, and consider $\mathcal{A} \in \mathbb{A}(\mathcal{X})$ such that there exists $\mathcal{A}_1 \in \mathbb{A}(\mathcal{X})$ with $\mathcal{A}_1 \subset \mathcal{A}$. Then $\mathcal{A}_2 = \mathcal{A} \backslash \mathcal{A}_1 \in \mathbb{A}(\mathcal{X})$, and $\min\{h_{\mathcal{A}_1}, h_{\mathcal{A}_2}\} \leq h_{\mathcal{A}}$.*

Proof Let $\mathcal{A} \in \mathbb{A}$ and assume that there exists $\mathcal{A}_1 \in \mathbb{A}$ such that $\mathcal{A}_1 \subset \mathcal{A}$. This means that $\sum_{A \in \mathcal{A}_1} I_A = \beta_{\mathcal{A}_1}$, with $\beta_{\mathcal{A}_1} < \beta_{\mathcal{A}}$. Take $\mathcal{A}_2 = \mathcal{A} \backslash \mathcal{A}_1 \subset \mathcal{A}$. It holds that:

$$\sum_{A \in \mathcal{A}_2} I_A = \sum_{A \in \mathcal{A}} I_A - \sum_{A \in \mathcal{A}_1} I_A = \beta_{\mathcal{A}} - \beta_{\mathcal{A}_1} \in \mathbb{N},$$

which implies that $\mathcal{A}_2 \in \mathbb{A}$ and also that $\beta_{\mathcal{A}} = \beta_{\mathcal{A}_1} + \beta_{\mathcal{A}_2}$ and $|\mathcal{A}| = |\mathcal{A}_1| + |\mathcal{A}_2|$. Hence

$$\frac{1}{|\mathcal{A}|} \left(\beta_{\mathcal{A}} - \sum_{A \in \mathcal{A}} \mu(A) \right)$$

$$= \frac{1}{|\mathcal{A}_1| + |\mathcal{A}_2|} \left(\left(\beta_{\mathcal{A}_1} - \sum_{A \in \mathcal{A}_1} \mu(A) \right) + \left(\beta_{\mathcal{A}_2} - \sum_{A \in \mathcal{A}_2} \mu(A) \right) \right)$$

$$= \frac{|\mathcal{A}_1|}{|\mathcal{A}_1| + |\mathcal{A}_2|} \frac{1}{|\mathcal{A}_1|} \left(\beta_{\mathcal{A}_1} - \sum_{A \in \mathcal{A}_1} \mu(A) \right)$$

$$+ \frac{|\mathcal{A}_2|}{|\mathcal{A}_1| + |\mathcal{A}_2|} \frac{1}{|\mathcal{A}_2|} \left(\beta_{\mathcal{A}_2} - \sum_{A \in \mathcal{A}_2} \mu(A) \right).$$

Thus, $h_{\mathcal{A}}$ is a convex combination of $h_{\mathcal{A}_1}$ and $h_{\mathcal{A}_2}$, and as a consequence $\min\{h_{\mathcal{A}_1}, h_{\mathcal{A}_2}\} \leq h_{\mathcal{A}}$. $\qquad \square$

Lemma 19 *Let μ be a maximally imprecise exact capacity, and let $\mathcal{A} = (A_i)_{i \in I} \in \mathbb{A}(\mathcal{X})$ be a family where the minimum in Eq. (12) is attained.*

(a) *If $\beta_{\mathcal{A}} = 1$, then $\mathcal{A}$ is a partition of $\mathcal{X}$.*
(b) *If $\beta_{\mathcal{A}} = |\mathcal{A}| - 1$, then $\mathcal{A}^c = (A_i^c)_{i \in I}$ is a partition of $\mathcal{X}$.*
(c) *If $1 < \beta_{\mathcal{A}} < |\mathcal{A}| - 1$ and for every $A, B \in \mathcal{A}$ at least one of $A \cap B$, $A \backslash B$ and $B \backslash A$ is empty, then there exists $\mathcal{A}_1 \in \mathbb{A}^*(\mathcal{X})$ such that $\mathcal{A}_1 \subset \mathcal{A}$.*

Proof (a) First of all, assume that $\beta_{\mathcal{A}} = \sum_{A \in \mathcal{A}} I_A = 1$. This means that each $x \in \mathcal{X}$ belongs to one and only one event $A \in \mathcal{A}$, hence $\mathcal{A}$ is a partition of $\mathcal{X}$.

(b) Secondly, assume that $\beta_{\mathcal{A}} = \sum_{A \in \mathcal{A}} I_A = |\mathcal{A}| - 1$. This means that

$$|\mathcal{A}| - 1 = \sum_{A \in \mathcal{A}} I_A = \sum_{A \in \mathcal{A}} [1 - (1 - I_A)] = |\mathcal{A}| - \sum_{A \in \mathcal{A}} (1 - I_A) = |\mathcal{A}| - \sum_{A \in \mathcal{A}} I_{A^c},$$

whence $\sum_{A \in \mathcal{A}} I_{A^c} = 1$ and applying (a), $\mathcal{A}^c = (A_i^c)_{i \in I}$ is a partition of $\mathcal{X}$.

(c) From the condition, given two different elements $A, B \in \mathcal{A}$ either they are disjoint or one of them is included in the other. If we then consider the partial order $\preceq$ on $\mathcal{A}$ given by set inclusion, then we can find a subfamily $\mathcal{A}_1 \subset \mathcal{A}$ of maximal elements in the order that are pairwise disjoint. Since any element of $\mathcal{X}$ must be included in some maximal element, it follows that the subfamily $\mathcal{A}_1$ is a partition of $\mathcal{X}$.

Proof of Theorem 5 Let $\mathcal{A} = (A_i)_{i=1,\dots,k}$ be the element in $\mathbb{A}(\mathcal{X})$ where the minimum in Eq. (12) is attained. According to Lemma 19, if $\beta_{\mathcal{A}} = 1$, then $\mathcal{A}$ is partition so $\mathcal{A} \in \mathbb{A}^*(\mathcal{X})$ hence Eq. (13) holds. If $\beta_{\mathcal{A}} = |\mathcal{A}| - 1$, then $\mathcal{A}^c = (A_i^c)_{i=1,\dots,k}$ is a partition, so it belongs to $\mathbb{A}^*(\mathcal{X})$. Also:

$$
\begin{aligned}
\alpha_S &= \frac{1}{|\mathcal{A}|}\left((|\mathcal{A}| - 1) - \sum_{A \in \mathcal{A}} \mu(A) \right) = \frac{1}{|\mathcal{A}|}\left((|\mathcal{A}| - 1) + \sum_{A \in \mathcal{A}} (1 - \mu(A) - 1) \right) \\
&= \frac{1}{|\mathcal{A}|}\left((|\mathcal{A}| - 1 - |\mathcal{A}|) + \sum_{A \in \mathcal{A}} \bar{\mu}(A^c) \right) = \frac{1}{|\mathcal{A}|}\left(\sum_{A \in \mathcal{A}} \bar{\mu}(A^c) - 1 \right) \\
&= \frac{1}{|\mathcal{A}|}\left(\sum_{A \in \mathcal{A}^c} \bar{\mu}(A) - 1 \right),
\end{aligned}
$$

hence Eq. (13) holds.

Assume now that $1 < \beta_{\mathcal{A}} < |\mathcal{A}| - 1$, and let us prove that it is possible to find another $\mathcal{A}^* \in \mathbb{A}(\mathcal{X})$ such that $\beta_{\mathcal{A}^*} < \beta_{\mathcal{A}}$ and where α_S is attained.

From item (c) in Lemma 19 we deduce that either there is $\mathcal{A}^* \in \mathbb{A}(\mathcal{X})$ with $\mathcal{A}^* \subset \mathcal{A}$ or there are two different $A_i, A_j \in \mathcal{A}$ with $A_i \cap A_j \neq \emptyset$, $A_i \setminus A_j \neq \emptyset$ and $A_j \setminus A_i \neq \emptyset$. In this second case, applying 2-monotonicity with the sets A_i and A_j above we deduce that:

$$
\begin{aligned}
h_{\mathcal{A}} &= \frac{1}{|\mathcal{A}|}\left(\beta_{\mathcal{A}} - \sum_{A \in \mathcal{A}} \mu(A) \right) = \frac{1}{|\mathcal{A}|}\left(\beta_{\mathcal{A}} - \sum_{A \in \mathcal{A} \setminus \{A_i, A_j\}} \mu(A) - \mu(A_i) - \mu(A_j) \right) \\
&\geq \frac{1}{|\mathcal{A}|}\left(\beta_{\mathcal{A}} - \sum_{A \in \mathcal{A} \setminus \{A_i, A_j\}} \mu(A) - \mu(A_i \cap A_j) - \mu(A_i \cup A_j) \right) = h_{\mathcal{A}_1},
\end{aligned}
$$

where $\mathcal{A}_1 = (\mathcal{A} \setminus \{A_i, A_j\}) \cup (A_i \cap A_j, A_i \cup A_j)$, using that $\beta_{\mathcal{A}_1} = \beta_{\mathcal{A}}$. Thus, $\alpha_S = h_{\mathcal{A}_1}$.

Now, if in $\mathcal{A}_1$ it is possible to find two different events B_i, B_j with $B_i \cap B_j \neq \emptyset$, $B_i \setminus B_j \neq \emptyset$ and $B_j \setminus B_i \neq \emptyset$ a similar reasoning shows that $\mathcal{A}_2 = \mathcal{A}_1 \cup (B_i \cup B_j, B_i \cap B_j) \setminus (B_i, B_j)$ also satisfies $\beta_{\mathcal{A}_2} = \beta_{\mathcal{A}_1}$ and $h_{\mathcal{A}_2} = h_{\mathcal{A}_1} = \alpha_S$. Iterating the procedure, we find after a finite number of steps that there are no different events C and D in the family $\mathcal{A}_k$ such that $C \cap D \neq \emptyset$, $C \setminus D \neq \emptyset$ and $D \setminus C \neq \emptyset$. But then, applying Lemma 19 we deduce that there is $\mathcal{A}^* \in \mathbb{A}^*(\mathcal{X})$ with $\mathcal{A}^* \subset \mathcal{A}_k$. Applying Lemma 18 that either $h_{\mathcal{A}_k} = h_{\mathcal{A}^*}$ or $h_{\mathcal{A}_k} = h_{\mathcal{A}_k \setminus \mathcal{A}^*}$. Since both $\beta_{\mathcal{A}^*}$ and $\beta_{\mathcal{A}_k \setminus \mathcal{A}^*}$ are strictly smaller than $\beta_{\mathcal{A}_k}$, we deduce that we can find another element of $\mathbb{A}(\mathcal{X})$ where the value α_S is attained and with a smaller value of $\beta_{\mathcal{A}}$. If we repeat this process we end up with a family $\mathcal{A}' \in \mathbb{A}(\mathcal{X})$ such that $\beta_{\mathcal{A}'} = 1$, and where α_S is attained, at which point we apply the first part of the proof.

Proof of Propositions 6 and 13 We start proving Proposition 13. From it, Proposition 6 trivially follows just noting that, by hypothesis, $\mathcal{L}^=$ would be empty and $\mathcal{L}^>$ would be formed by all the proper events of $\mathcal{X}$ (because μ is by hypothesis maximally imprecise), and this allows to apply Proposition 13 to a lower prevision were $\mathcal{L}^>$ contains the indicator functions of the proper events.

($\Rightarrow$) Assume that $P_0(f) \geq \underline{P}(f) + \alpha$ for every $f \in \mathcal{L}^>$, and consider $Q \in B_c^\alpha(P_0)$. Since we can assume without loss of generality that $0 = \min f < \max f = 1$ for every $f \in \mathcal{L}^>$, it follows that $\mid Q(f) - P_0(f) \mid \leq \alpha$, whence $Q(f) \geq P_0(f) - \alpha \geq \underline{P}(f)$ for every $f \in \mathcal{L}^>$. Since by assumption $\mathcal{L}^= = \emptyset$, this implies that $Q \in \mathcal{M}(\underline{P})$.

($\Leftarrow$) Consider a probability measure P_0 such that $B_c^\alpha(P_0) \subseteq \mathcal{M}(\underline{P})$, and let us prove that $P_0(f) \geq \underline{P}(f) + \alpha$ for every $f \in \mathcal{L}^>$. Assume ex-absurdo that $P_0(f) - \alpha < \underline{P}(f)$ for some gamble $f \in \mathcal{L}^>$, and let us show that there exists some $Q \in B_c^\alpha(P_0)$ such that $Q(f) < \underline{P}(f)$. To see that this is indeed the case, note that since $B_c^\alpha(P_0)$ is included in $\mathbb{P}^*(\mathcal{X})$, it must be $P_0(\{x_i\}) > \alpha$ for every $x_i \in \mathcal{X}$. If we now take $x_m, x_M \in \mathcal{X}$ such that:

$$0 = \min f = f(x_m), \quad 1 = \max f = f(x_M),$$

and define Q by means of the mass function

$$Q(\{x_m\}) = P_0(\{x_m\}) + \alpha, \quad Q(\{x_M\}) = P_0(\{x_M\}) - \alpha,$$

and $Q(\{x\}) = P_0(\{x\})$ for any other $x \neq x_m, x_M$, it follows that $\mid Q(B) - P_0(B) \mid \leq \alpha$ for every $B \subseteq \mathcal{X}$, whence $Q \in B_c^\alpha(P_0) \subseteq \mathcal{M}(\underline{P})$. However:

$$Q(f) = \sum_{x \in \mathcal{X}} Q(\{x\})f(x) = \sum_{x \neq x_m, x_M} Q(\{x\})f(x) + Q(\{x_m\})f(x_m) + Q(\{x_M\})f(x_M)$$

$$= \sum_{x \neq x_m, x_M} P_0(\{x\})f(x) + \big(P_0(\{x_m\}) + \alpha\big)f(x_m) + \big(P_0(\{x_M\}) - \alpha\big)f(x_M)$$

$$= P_0(f) - \alpha\big(f(x_M) - f(x_m)\big) = P_0(f) - \alpha < \underline{P}(f),$$

hence $Q \notin \mathcal{M}(\underline{P})$, we obtain a contradiction.

The second part of Proposition 6 is an immediate consequence of the first once we realise that, since $\mathcal{M} \subseteq \mathbb{P}^*(\mathcal{X})$, we can compute α_I by means of Eq. (8). □

Proof of Propositions 7, 8 and 16 We start proving the properties of the centroids for coherent lower previsions (Proposition 16). Since exact capacities are particular cases of coherent lower previsions, Propositions 7 and 8 can be regarded as a corollary. Throughout this proof, for simplicity we use the notation $\underline{P}(A) := \underline{P}(I_A)$ and $\overline{P}(A) = \overline{P}(I_A)$.

It x_i is a null-player, $\underline{P}(\mathcal{X}) = 1 = \underline{P}(\mathcal{X} \setminus \{x_i\})$, whence $\overline{P}(\{x_i\}) = 0$. Since $\Phi_2^{\underline{P}}, \Phi_3^{\underline{P}}, \Phi_4^{\underline{P}}$ belong to the credal set, any of these centroids gives probability zero to x_i.

With respect to symmetry, let $\sigma_{i,j}$ denote the permutation of $\mathcal{X}$ that exchanges x_i and x_j and leaves the other elements fixed, and let $\underline{P}_{\sigma_{i,j}}$ be given by $\underline{P}_{\sigma_{i,j}}(f) = \underline{P}(f \circ \sigma_{i,j})$.

If P is an extreme point of $\mathcal{M}(\underline{P})$, then $P_{\sigma_{i,j}}$ is an extreme point of $\mathcal{M}(\underline{P}_{\sigma_{i,j}})$. If we now assume that by symmetry $\underline{P}_{\sigma_{i,j}} = \underline{P}$, it follows that $\Phi_2^{\underline{P}} = \Phi_2^{\underline{P}_{\sigma_{i,j}}}$. But on the other hand it must be $\Phi_2^{\underline{P}}(\{x_i\}) = \Phi_2^{\underline{P}_{\sigma_{i,j}}}(\{x_j\})$, and as a consequence $\Phi_2^{\underline{P}}(\{x_i\}) = \Phi_2^{\underline{P}}(\{x_j\})$.

Concerning the set of incenters, if P satisfies that $B_o^\alpha(P) \subseteq \mathcal{M}(\underline{P}) \cap \mathbb{P}^*(\mathcal{X})$, then $B_o^\alpha(P_{\sigma_{i,j}}) \subseteq \mathcal{M}(\underline{P}_{\sigma_{i,j}}) \cap \mathbb{P}^*(\mathcal{X}) = \mathcal{M}(\underline{P}) \cap \mathbb{P}^*(\mathcal{X})$. Therefore,

$$\left\{ P(\{x_i\}) \mid P \in \Psi_3^{\underline{P}} \right\} = \left\{ P(\{x_i\}) \mid P \in \Psi_3^{\underline{P}_{\sigma_{i,j}}} \right\} = \left\{ P(\{x_j\}) \mid P \in \Psi_3^{\underline{P}} \right\}.$$

Finally, for the contraction centroid note that, for any gamble $f \in \mathcal{L}^>$ and any $\alpha > 0$, it holds that

$$P(f) \geq \underline{P}(f) + \alpha \Leftrightarrow P_{\sigma_{i,j}}(f) \geq \underline{P}_{\sigma_{i,j}}(f) + \alpha = \underline{P}(f) + \alpha,$$

whence $\mathcal{M}(\underline{P})_\alpha$ is invariant under $\sigma_{i,j}$. From here we deduce that $\Phi_4^{\frac{P}{4}}(\{x_i\}) = \Phi_4^{\frac{P}{4}}(\{x_j\})$.

With respect to ignorance preservation, in the case of $\Phi_1^{\frac{P}{1}}$ it follows immediately from Eq. (3) because the Möbius inverse of $\underline{P}$ satisfies $m(\mathcal{X}) = 1$ and zero for any other $A \neq \mathcal{X}$. In the case of $\Phi_2^{\frac{P}{2}}$, it suffices to note that the extreme points of $\mathcal{M}(\underline{P})$ are the degenerate distributions; and for $\Phi_3^{\frac{P}{3}}$, $\Phi_4^{\frac{P}{4}}$ it suffices to use that the uniform distribution P_0 is the only one in $\mathcal{M}(\underline{P})$ for which $B_o^{\frac{1}{|\mathcal{X}|}}(P_0)$ is included in $\mathcal{M}(\underline{P}) \cap \mathbb{P}^*(\mathcal{X})$.

Finally, with respect to continuity, in the case of Φ_1^μ it follows directly from Eq. (2); while in the case of Φ_3^μ, Φ_4^μ it follows trivially because we are using in their definition the topology of the total variation. $\qquad\square$

Proof of Propositions 9, 10 and 17 We first prove that φ_1, φ_2 and φ_3 are centrality measures for a coherent lower prevision $\underline{P}$ (assuming that $\mathcal{M}(\underline{P})$ is a polytope for φ_1 and φ_3), proving hence Proposition 17. Since any exact capacity is a particular type of coherent lower prevision whose core is a polytope, Propositions 9 and 10 trivially follow as a corollary.

Let us begin by showing that φ_1 is a centrality measure:

CM1 This holds by definition.
CM2 By construction, if $P \in \text{ext}(\mathcal{M}(\underline{P}))$ then $d(P, \Phi^{\underline{P}}) \geq \beta$, hence $\varphi(P) = 1 - 1 = 0$.
CM3 By definition $\varphi_1(P) = 1$ iff $d(P, \Phi^{\underline{P}}) = 0$, and since d is a distance this holds iff $P = \Phi^{\underline{P}}$.
CM4 This follows if and only if for any $P \in \text{ext}(\mathcal{M}(\underline{P}))$ and any $\lambda, \beta \in [0, 1]$ such that $\lambda \geq \beta$ it holds that

$$d(\lambda P + (1 - \lambda)\Phi^{\underline{P}}, \Phi^{\underline{P}}) \leq d(\beta P + (1 - \beta)\Phi^{\underline{P}}, \Phi^{\underline{P}}). \tag{A3}$$

If we denote $P_1 := \lambda P + (1 - \lambda)\Phi^{\underline{P}}$ and $P_2 := \beta P + (1 - \beta)\Phi^{\underline{P}}$, there is some $a \in (0, 1)$ such that $P_2 = a P_1 + (1 - a)\Phi^{\underline{P}}$. As a consequence, for any event A it holds that

$$| P_2(A) - \Phi^{\underline{P}}(A) | \leq a | P_1(A) - \Phi^{\underline{P}}(A) | + (1 - a) | \Phi^{\underline{P}}(A) - \Phi^{\underline{P}}(A) |$$
$$= a | P_1(A) - \Phi^{\underline{P}}(A) |,$$

from which Eq. (A3) follows.

Let us prove now that φ_2 is a centrality measure.

CM1,CM3 These follow immediately from the definition of $\mathcal{M}(\underline{P})_0$, $\mathcal{M}(\underline{P})_1$.
CM2 This holds because any extreme point P of $\mathcal{M}(\underline{P})$ does not belong to $\mathcal{M}(\underline{P})_\alpha$ for any $\alpha > 0$.
CM4 Consider $P_1 = \lambda P + (1 - \lambda)P_0$ and $P_2 = \beta P + (1 - \beta)P_0$ for $\lambda \geq \beta$, where P is an extreme point of $\mathcal{M}(\underline{P})$ and P_0 is the unique element of $\mathcal{M}(\underline{P})_1$. Then (CM4) holds if an only if for any $\gamma \in (0, 1)$ such that $P_1 \in \mathcal{M}(\underline{P})_\gamma$, also $P_2 \in \mathcal{M}(\underline{P})_\gamma$. This is a consequence of the convexity of the set $\mathcal{M}(\underline{P})_\gamma$, given that $P_2 = a P_1 + (1 - a)P_0$ for some $a \in [0, 1]$.

Finally, let us see that φ_3 satisfies the four properties in Definition 6.

CM1 For any $P \notin \mathcal{M}(\underline{P})$, $P \notin \mathcal{M}(\underline{P})_\alpha$ for any $\alpha > 0$, whence $\varphi_3(P) = 0$.
CM2 If $P \in \text{ext}(\mathcal{M}(\underline{P}))$, then $P(f) = \underline{P}(f)$ for some gamble f. This implies that $P(f) \not\geq \underline{P}(f) + \alpha$, and therefore $P \notin \mathcal{M}(\underline{P})_\alpha$ for any $0 < \alpha \leq \alpha_{S_1}$, which implies that $\varphi_3(P) = 0$.

CM3 By definition of the procedure, $P_0 = \Phi_4^{\underline{P}}$ is the unique probability in $\mathcal{M}_{\alpha_{S_l}}(\underline{P})$, so it is the unique probability with centrality degree 1.

CM4 Take $\lambda \geq \beta$, $P \in \text{ext}(\mathcal{M}(\underline{P}))$ and define $P_1 = \lambda P + (1 - \lambda)P_0$ and $P_2 = \beta P + (1 - \beta)P_0$. Let us see that if $P_1(f) \geq \underline{P}(f) + \alpha$, then $P_2(f) \geq \underline{P}(f) + \alpha$ too, for every $f \in \mathcal{L}^>$. Consider two cases:

1. Case 1: $P(f) \geq P_0(f)$. In that case, it holds that $P(f) \geq P_0(f) \geq \underline{P}(f) + \alpha$, hence:

$$P_2(f) = \beta P(f) + (1 - \beta)P_0(f) \geq \beta P_0(f) + (1 - \beta)P_0(f) = P_0(f) \geq \underline{P}(f) + \alpha.$$

2. Case 2: $P(f) < P_0(f)$. Then $P_1(f) \geq \underline{P}(f) + \alpha$ is equivalent to:

$$\lambda P(f) + (1 - \lambda)P_0(f) \geq \underline{P}(f) + \alpha,$$

which is equivalent to:

$$\lambda \big(P(f) - P_0(f)\big) \geq \underline{P}(f) + \alpha - P_0(f).$$

Using that $P(f) - P_0(f) < 0$, the previous inequality is equivalent to:

$$\lambda \leq \frac{\underline{P}(f) + \alpha - P_0(f)}{P(f) - P_0(f)}.$$

Since $\beta \leq \lambda$, following the reversed steps, if follows that:

$$P_2(f) = \beta P(f) + (1 - \beta)P_0(f) \geq \underline{P}(f) + \alpha.$$

With this, we conclude that if $P_1(f) \geq \underline{P}(f) + \alpha$, $P_2(f) \geq \underline{P}(f) + \alpha$ too, for every $f \in \mathcal{L}^>$. Hence, if the centrality degree of P_1 is $\varphi_3^{\underline{P}}(P_1)$, the centrality degree of P_2 should be at least $\varphi_3^{\underline{P}}(P_1)$.   □

Proof of Proposition 14 Since $\underline{P}$ is associated with a probability interval,

$$\mathcal{M}(\underline{P}) = \{P \mid l_i \leq P(\{x_i\}) \leq u_i \ \forall i = 1, \ldots, n\},$$

meaning that, while we can assume without loss of generality that $\mathcal{L}^= \cup \mathcal{L}^>$ include the indicators of proper events, in this case only the indicators of the singletons are necessary, and $\mathcal{L}^=$ and $\mathcal{L}^>$ reduce to $\mathcal{I}^=$ and $\mathcal{I}^>$, respectively.

Let α be the minimum in Eq. (24), and let us see that $\alpha = \alpha_S$. We consider the following cases:

1. Assume that $\alpha = \frac{1}{|\mathcal{I}^>|}\left(1 - \sum_{i=1}^n l_i\right)$. Define P_0 by:

$$P_0(\{x_i\}) = \begin{cases} l_i + \alpha, & \text{if } i \in \mathcal{I}^>. \\ l_i, & \text{if } i \in \mathcal{I}^=. \end{cases}$$

P_0 satisfies the following properties:

(i) It is a probability measure because it is non-negative and

$$\sum_{i=1}^n P_0(\{x_i\}) = \sum_{i \in \mathcal{I}^>}(l_i + \alpha) + \sum_{i \in \mathcal{I}^=} l_i$$

$$= \sum_{i \in \mathcal{I}^>} l_i + \mathcal{I}^> \alpha + \sum_{i \in \mathcal{I}^=} l_i = \sum_{i=1}^n l_i + \left(1 - \sum_{i=1}^n l_i\right) = 1.$$

(ii) $P_0(\{x_i\}) \in [l_i, u_i]$. On the one hand, if $i \in \mathcal{I}^=$, $P_0(\{x_i\}) = l_i$. On the other hand, if $i \in \mathcal{I}^>$, $P_0(\{x_i\}) = l_i + \alpha > l_i$. Also, by definition of α it holds that $\alpha \le \frac{1}{2}(u_i - l_i)$, hence:

$$P_0(\{x_i\}) = l_i + \alpha < l_i + \frac{1}{2}(u_i - l_i) = \frac{1}{2}(l_i + u_i) < u_i.$$

(iii) To see that $\mathcal{M}(\underline{P})_\alpha = \{P_0\}$, and as a consequence that $\alpha = \alpha_S$, note that, by construction,

$$\mathcal{M}(\underline{P})_\alpha = \{P \mid P(\{x_i\}) \in [\underline{P}(\{x_i\}) + \alpha, \overline{P}(\{x_i\}) - \alpha] \; \forall i \in \mathcal{I}^>,$$
$$P(\{x_i\}) = \underline{P}(\{x_i\}) = \overline{P}(\{x_i\}) \; \forall i \in \mathcal{I}^=\}$$
$$= \{P \mid P(\{x_i\}) \in [l_i + \alpha, u_i - \alpha] \; \forall i \in \mathcal{I}^>, P(\{x_i\}) = l_i = u_i \; \forall i \in \mathcal{I}^=\}.$$

It follows then that by definition $P_0 \in \mathcal{M}(\underline{P})_\alpha$. Since moreover

$$\sum_{i=1}^n l_i + \mid \mathcal{I}^> \mid \alpha = 1,$$

we deduce that any $Q \in \mathcal{M}_\alpha(\underline{P})$ must coincide with P_0 and that $\mathcal{M}_{\alpha'}(\underline{P}) = \emptyset$ for any $\alpha' > \alpha$.

This implies that $\mathcal{M}(\underline{P})_\alpha = \{P_0\}$ is formed by a unique probability measure P_0, which coincides with Φ_4^P.

2. Assume that $\alpha = \frac{1}{|\mathcal{I}^>|}\left(\sum_{i=1}^n u_i - 1\right)$. Define P_0 by:

$$P_0(\{x_i\}) = \begin{cases} u_i - \alpha, & \text{if } i \in \mathcal{I}^>. \\ u_i, & \text{if } i \in \mathcal{I}^=. \end{cases}$$

Following the same steps as in the previous case, we obtain that $\mathcal{M}(\underline{P})_\alpha = \{P_0\}$, where $P_0 = \Phi_4^P$.

3. Assume that $\alpha = \frac{1}{2}(u_i - l_i)$ for some $i \in \{1, \dots, n\}$. We define a new probability interval $\mathcal{I}^*$ given by:

$$[l_i^*, u_i^*] = \begin{cases} [l_i, u_i] & \text{if } i \in \mathcal{I}^=. \\ [l_i + \alpha, u_i - \alpha] & \text{if } i \in \mathcal{I}^>. \end{cases}$$

Let us prove some interesting properties of this probability interval:

(i) $\mathcal{M}(\mathcal{I}^*) \ne \emptyset$: on the one hand,

$$\sum_{i=1}^n l_i = \sum_{i \in \mathcal{I}^>}(l_i + \alpha) + \sum_{i \in \mathcal{I}^=} l_i = \sum_{i=1}^n l_i + \mid \mathcal{I}^> \mid \alpha \le 1,$$

since by hypothesis $\alpha \le \frac{1}{|\mathcal{I}|}\left(1 - \sum_{i=1}^n l_i\right)$ and, since $\mathcal{I}$ avoids sure loss, $\sum_{i=1}^n l_i \le 1$. Similarly:

$$\sum_{i=1}^n u_i = \sum_{i \in \mathcal{I}^>}(u_i - \alpha) + \sum_{i \in \mathcal{I}^=} u_i = \sum_{i=1}^n u_i - \alpha \mid \mathcal{I}^> \mid \ge 1,$$

since by hypothesis $\alpha \le \frac{1}{|\mathcal{I}|}\left(\sum_{i=1}^n u_i - 1\right)$ and, since $\mathcal{I}$ avoids sure loss, $\sum_{i=1}^n u_i \ge 1$. We therefore conclude that $\mathcal{I}^*$ avoids sure loss.

(ii) By construction, $\mathcal{M}(\mathcal{I}^*) = \mathcal{M}(\underline{P})_\alpha$. Moreover, given $\alpha' > \alpha$, if follows that there is some $i = 1, \ldots, n$ such that $l_i + \alpha' > u_i - \alpha'$, because $\alpha' > \min_i \frac{u_i - l_i}{2}$, and therefore $\mathcal{M}(\underline{P})_{\alpha'} = \emptyset$. Thus, $\alpha = \alpha_S$.

It only remains to see that $\alpha = \alpha_S$, but this straightforwardly follows from the fact that the chosen α saturates the probability of at least one of the events.

$\square$

Proof of Corollary 15 Consider first the PMM. Since we are assuming that $\underline{P}_{PMM}(A) \in (0, 1)$ for non-trivial events A, it corresponds to the probability interval $\mathcal{I}_{PMM}$ given by:

$$l_i = (1 + \delta) P_0(\{x_i\}) - \delta, \quad u_i = (1 + \delta) P_0(\{x_i\}) \quad \forall i = 1, \ldots, n.$$

Applying Eq. (24) to this probability interval, we obtain:

$$\alpha_S = \min \left\{ \frac{1}{n} \left(1 - \sum_{i=1}^{n} l_i \right), \frac{1}{n} \left(\sum_{i=1}^{n} u_i - 1 \right), \frac{1}{2} \min_{i=1,\ldots,n} (u_i - l_i) \right\}$$
$$= \min \left\{ \frac{1}{n} (1 - (1 + \delta - n\delta)), \frac{1}{n} \delta, \frac{1}{2} \delta \right\} = \frac{\delta}{n}.$$

Following the steps in the previous proposition, we deduce that $\mathcal{M}(\underline{P}_{PMM})_{\alpha_S} = \{ \Phi_4^{P_{PMM}} \}$, where $\Phi_4^{P_{PMM}}$ is given by:

$$\Phi_4^{P_{PMM}} = u_i - \alpha_S = (1 + \delta) P_0(\{x_i\}) - \frac{\delta}{n} \quad \forall i = 1, \ldots, n.$$

Similarly, for the LV we are assuming that $\underline{P}_{LV}(A) \in (0, 1)$ for non-trivial events A, so the LV corresponds to a probability interval $\mathcal{I}_{LV}$ given by:

$$l_i = (1 - \delta) P_0(\{x_i\}), \quad u_i = (1 - \delta) P_0(\{x_i\}) + \delta \quad \forall i = 1, \ldots, n.$$

Applying again Eq. (24):

$$\alpha_S = \min \left\{ \frac{1}{n} \left(1 - \sum_{i=1}^{n} l_i \right), \frac{1}{n} \left(\sum_{i=1}^{n} u_i - 1 \right), \frac{1}{2} \min_{i=1,\ldots,n} (u_i - l_i) \right\}$$
$$= \min \left\{ \frac{\delta}{n}, \frac{1}{n} (-\delta + n\delta), \frac{\delta}{2} \right\} = \frac{\delta}{n}.$$

From Proposition 14, we obtain $\mathcal{M}(\underline{P}_{LV})_{\alpha_S} = \{ \Phi_4^{P_{LV}} \}$, where $\Phi_4^{P_{LV}}$ is given by:

$$\Phi_4^{P_{LV}}(\{x_i\}) = l_i + \alpha_S = (1 - \delta) P_0(\{x_i\}) + \frac{\delta}{n}.$$

It follows from the results in Miranda and Montes (2018) that the values obtained for the PMM and the LV coincide with the Shapley value. Also, we have seen that $\mathcal{M}(\underline{P})_{\alpha_S}$ coincides with the set of incenters with respect to the total variation distance (see Proposition 13). Since in this case $\mathcal{M}(\underline{P})_{\alpha_S}$ is a singleton (for both the PMM and the LV), the incenter is unique and it coincides with the contraction center. Finally, to see that these centers also coincide with the average of the extreme points, we just need to note (see Montes et al. 2019, Sec.3.1 and Montes et al. 2020a, Sec.5.1) that using the approach based on the permutations, each extreme point appears in the same number of permutations, so the average of the extreme points coincides with the Shapley value. $\square$

References

Abellán, J., & Moral, S. (2003). Maximum entropy for credal sets. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems, 11*, 587–597.

Angilella, S., Bottero, M., Corrente, S., Ferretti, V., Greco, S., & Lami, I. M. (2016). Non additive robust ordinal regression for urban and territorial planning: An application for sitting in urban waste landfill. *Annals of Operations Research, 245*, 427–456.

Augustin, T., Coolen, F., de Cooman, G., & Troffaes, M. (Eds). (2014). *Introduction to Imprecise Probabilities*. Wiley Series in Probability and Statistics. Wiley.

Bader, U., Gelander, T., & Monod, N. (2012). A fixed point theorem for L^1 spaces. *Inventiones mathematicae, 189*, 143–148.

Banzhaf, J. F. (1965). Weighted voting does not work: A mathematical analysis. *Rutgers Law Review, 19*, 317–343.

Baroni, P., & Vicig, P. (2005). An uncertainty interchange format with imprecise probabilities. *International Journal of Approximate Reasoning, 40*, 147–180.

Cascos, I. (2009). Data depth: Multivariate statistics and geometry. In W. S. Kendall & I. Molchanov (Eds.), *New perspectives in stochastic geometry*. Oxford Scholarship Online.

Choquet, G. (1953). Theory of capacities. *Annales de l'Institut Fourier, 5*, 131–295.

Corsato, C., Pelessoni, R., & Vicig, P. (2019). Nearly-linear uncertainty measures. *International Journal of Approximate Reasoning, 114*, 1–28.

de Campos, L. M., Huete, J. F., & Moral, S. (1994). Probability intervals: A tool for uncertain reasoning. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems, 2*, 167–196.

Destercke, S. (2017). On the median in imprecise ordinal problems. *Annals of Operations Research, 256*, 375–392.

Destercke, S., Montes, I., & Miranda, E. (2022). Processing multiple distortion models: A comparative study. *International Journal of Approximate Reasoning, 145*(C), 91–120.

Dubois, D., & Prade, H. (1980). *Fuzzy sets and systems. Theory and applications*. Academic Press.

Elbassioni, K., & Tiway, H. R. (2012). Complexity of approximating the vertex centroid of a polyhedron. *Theoretical Computer Science, 421*, 56–61.

Gilboa, I., & Schmeidler, D. (1989). Maxmin expected utility with a non-unique prior. *Journal of Mathematical Economics, 18*, 141–153.

Grabisch, M. (2013). The core of games on ordered structures and graphs. *Annals of Operations Research, 204*, 33–64.

Grabisch, M. (2016). *Set functions, games and capacities in decision making*. Springer.

Huntley, N., & Troffaes, M. (2012). Normal form backward induction for decision trees with coherent lower previsions. *Annals of Operations Research, 195*, 111–134.

Jaffray, J. (1995). On the maximum-entropy probability which is consistent with a convex capacity. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems, 3*(1), 27–33.

Keith, A., & Ahner, D. (2021). A survey of decision making and optimization under uncertainty. *Annals of Operations Research, 300*, 319–353.

Klibanoff, P., Marinacci, M., & Mukerji, S. (2005). A smooth model of decision making under ambiguity. *Econometrica, 73*(6), 1849–1892.

Klir, G. J., & Parviz, B. (1992). Probability-possibility transformations: A comparison. *International Journal of General Systems, 21*, 291–310.

Kumar, I. E., Venkatasubramanian, S., Scheidegger, C., & Friedler, S. A. (2020). Problems with Shapley-value-based explanations as feature importance measures. arXiv:2002.11097.

Levi, I. (1980). *The enterprise of knowledge*. MIT Press.

Levin, D. A., Peres, Y., & Wilmer, E. L. (2009). *Markov chains and mixing times*. American Mathematical Society.

Lundberg, S., & Lee, S. I. (2017). A unified approach to interpreting model predictions. arXiv:1705.07874.

Miranda, E., Couso, I., & Gil, P. (2003). Extreme points of the credal sets generated by 2-alternating capacities. *International Journal of Approximate Reasoning, 33*(1), 95–115.

Miranda, E., & Destercke, S. (2015). Extreme points of the credal sets generated by comparative probabilities. *Journal of Mathematical Psychology, 64*(65), 44–57.

Miranda, E., & Montes, I. (2018). Shapley and Banzhaf values as probability transformations. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems, 26*(6), 917–947.

Miranda, E., & Montes, I. (2021). Centroids of credal sets: a comparative study. In J. Vejnarová, & N. Wilson (Eds.), *Symbolic and quantitative approaches to reasoning with uncertainty. ECSQARU 2021, Volume 12897 of Lecture Notes in Artificial Intelligence* (pp. 427–441). Springer.

Montes, I., & Destercke, S. (2017). Extreme points of p-boxes and belief functions. *Annals of Mathematics and Artificial Intelligence, 81*(3), 405–428.

Montes, I., Miranda, E., & Destercke, S. (2019). Pari-mutuel probabilities as an uncertainty model. *Information Sciences, 481*, 550–573.

Montes, I., Miranda, E., & Destercke, S. (2020a). Unifying neighbourhood and distortion models: Part I—New results on old models. *International Journal of General Systems, 49*(6), 602–635.

Montes, I., Miranda, E., & Destercke, S. (2020b). Unifying neighbourhood and distortion models: Part II—New models and synthesis. *International Journal of General Systems, 49*(6), 636–674.

Neyman, A. (2002). Values of games with infinitely many players. *Handbook of Game Theory with Economic Applications, 3*, 2121–2167.

Pelessoni, R., Vicig, P., & Zaffalon, M. (2010). Inference and risk measurement with the pari-mutuel model. *International Journal of Approximate Reasoning, 51*, 1145–1158.

Sarin, R., & Wakker, P. (1992). A simple axiomatization of nonadditive expected utility. *Econometrica, 60*(6), 1255–1272.

Shafer, G. (1976). *A mathematical theory of evidence*. Princeton University Press.

Shapley, L. S. (1953). A value for n-person game. *Annals of Mathematics Studies, 28*, 307–317.

Shapley, L. S. (1971). Cores of convex games. *International Journal of Game Theory, 1*, 11–26.

Smets, P. (2005). Decision making in the TBM: The necessity of the pignistic transformation. *International Journal of Approximate Reasoning, 38*, 133–147.

Smets, P., & Kennes, R. (1994). The transferable belief model. *Artificial Intelligence, 66*(2), 191–234.

Troffaes, M. C. M. (2007). Decision making under uncertainty using imprecise probabilities. *International Journal of Approximate Reasoning, 45*(1), 17–29.

Tukey, J. (1975). Mathematics and the picturing of data. In *Proceedings of the international congress of mathematicians* (Vol. 2, pp. 523–531).

Walley, P. (1981). *Coherent lower (and upper) probabilities*. Statistics research report.

Walley, P. (1991). *Statistical reasoning with imprecise probabilities*. Chapman and Hall.

Wallner, A. (2007). Extreme points of coherent probabilities in finite spaces. *International Journal of Approximate Reasoning, 44*(3), 339–357.

Weber, R. J. (1988). Probabilistic values for games. In A. E. Roth (Ed.), *The Shapley value. Essays in honour of L.S. Shapley* (pp. 101–119). Cambridge University Press.