On computing evidential centroid through conjunctive combination: an impossibility theorem

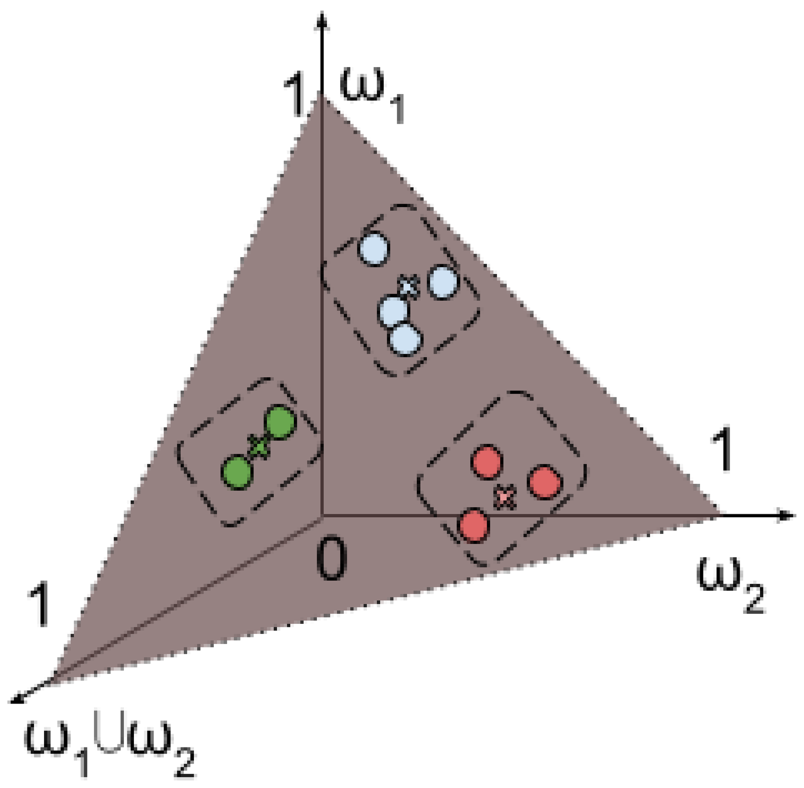
Yiru Zhang[1], Sébastien Destercke[2], Zuowei Zhang[3], Arnaud Martin[3], Tassadit Bouadi[3]

[1]CY Cergy Paris Universités, ENSEA, CNRS, UMR 8051 ETIS, Cergy, France
[2]Université de technologie de Compiègne, CNRS, UMR 7253 Heudiasyc, Compiègne, France
[3]Université Rennes, CNRS, UMR 6074 IRISA, Lannion, France

Problematic

In the theory of belief function:

- Combination operation: fuse information by combining multiple evidence corpus, resulting in a summary of the different pieces of evidence.
- Distances: measure dissimilarities between pieces of evidences, which may in turn be used to evaluate source reliability to perform a fusion or combination.



Combination of evidence corpus are used in decision making, while classification/clustering the evidence corpus can be used in grouping the uncertain/imprecise objects.
How the two applications interplay? A common utility between the two is the calculation of centroid.
One question:

Can we use combination rules to compute the centroid of a group of evidence corpus?

Properties for combination/centroid calculation

Centroid of evidence corpus:

Computing the centroid is then the problem of finding the mass $\overline{m}$, defined from an arbitrary set $\{m_1, \dots, m_\ell\}$ of mass functions as:

$$\overline{m} = \arg \min_{m \in \mathcal{E}(\Omega)} \sum_{i=1}^{\ell} d(m_i, m).$$

Usually, to guarantee the convergence of inertia in the iteration of centroid-based clustering, the combination rule must be consistent with a **metric**.

- **Property 1 Metric consistency:** *The combination rule $\odot$ and the dissimilarity measure d must be consistent. Formally, this means that the combined mass function $\overline{m}$ calculated by:*

$$\overline{m} = \bigodot_{i=1}^{\ell} m_i,$$

should be the centroid, i.e.,

$$\sum_{i=1}^{\ell} d(m_i, m') > \sum_{i=1}^{\ell} d(m_i, \overline{m}), \ \forall m' \in \mathcal{E}(\Omega), m' \neq \overline{m}.$$

Least commitment principle (LCP) in TBF:

When several belief functions are compatible with a set of constraints, the least informative according to some informational ordering (if it exists) should be selected.
LCP is applied in many seminal combination rules, it required the property of **ignorance neutrality**: defined as:

- **Property 2 Ignorance neutrality:** *Vacuous mass functions representing ignorance are neutral elements of the combination rule $\odot$, i.e.,*

$$m \odot \Omega^1 = m, \quad \forall m \in \mathcal{E}(\Omega),$$

where Ω^1 denotes a vacuous mass function.

Another property for cautiousness in an information fusion process.

- **Property 3 Idempotence:** *The combination rule $\odot$ is idempotent if and only if:*

$$m \odot m = m, \quad \forall m \in \mathcal{E}(\Omega).$$

Indeed, Idempotence is also a **necessary condition** of "metric consistency".

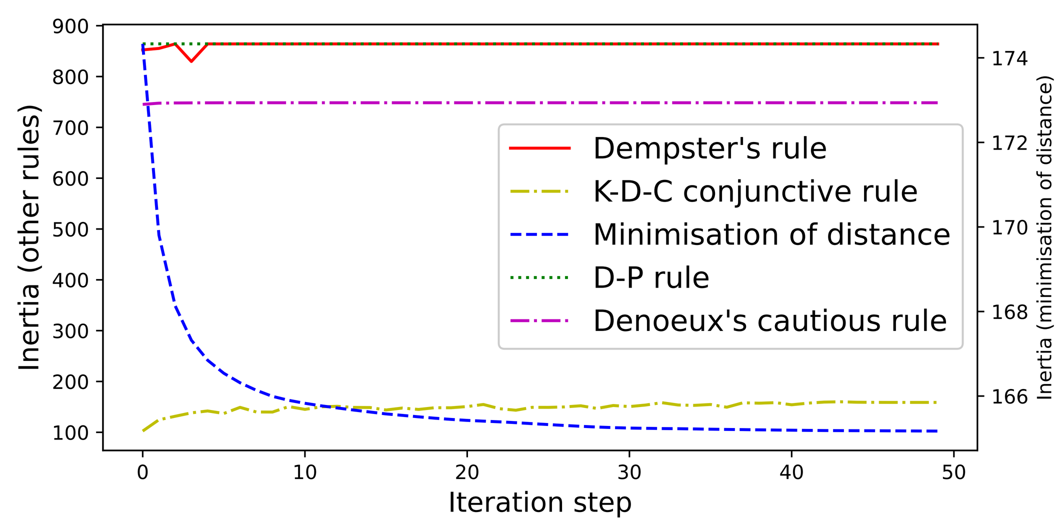
The impossibility theorem

Given a dissimilarity measure $d : \mathcal{E}(\Omega) \times \mathcal{E}(\Omega) \to \mathbb{R}_{\geq 0}$ on the set of evidence corpus, there is no combination rule $\odot$ that satisfies the properties of ***metric consistency*** and ***ignorance neutrality*** simultaneously.

Case study of combination rules

Properties owned by combination rules:

Combination rules	I.N.	idem.	M.C.
Dempster's rule	T	F	F
Dubois-Prade rule	T	F	F
Denœux's cautious rule	F	T	F
K-D-C conjunctive rule	T	T	F
Minimisation of distance	F	T	T



Inertia of clustering with different combination rules

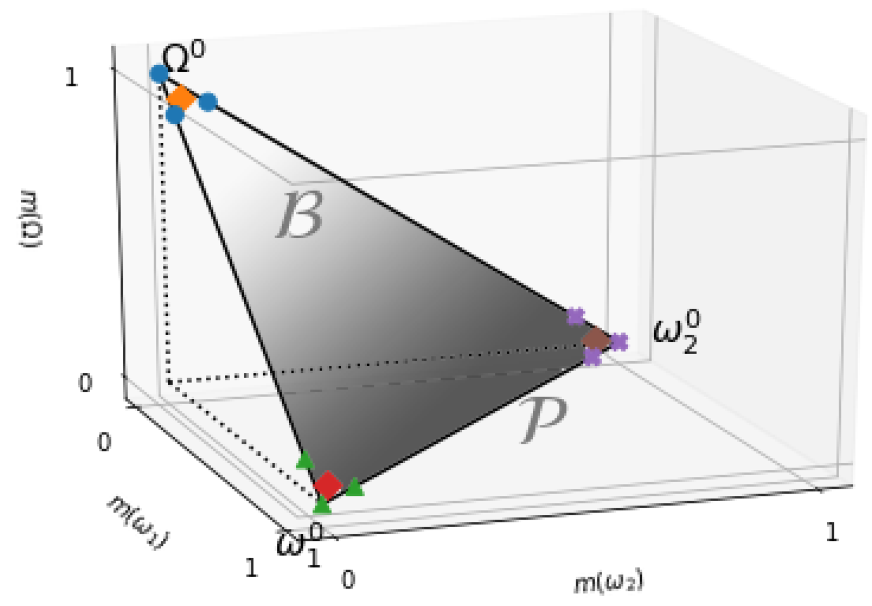
Different views on clustering evidence corpus

Illustrative data in FoD $\Omega = \{\omega_1, \omega_2\}$

	Group 1			Group 2			Group 3		
	m_1	m_2	m_3	m_4	m_5	m_6	m_7	m_8	m_9
$m(\omega_1)$	0	0.1	0	0	0.1	0	0.9	0.9	1
$m(\omega_2)$	0.1	0	0	0.9	0.9	1	0	0.1	0
$m(\Omega)$	0.9	0.9	1	0.1	0	0	0.1	0	0

A metric view

Each evidential corpus is precisely projected into a manifold consistent with a metric. (e.g. Euclidean)



A conflict-based view

Conflict measure $\kappa(m_1, m_2)$ on combination rule:

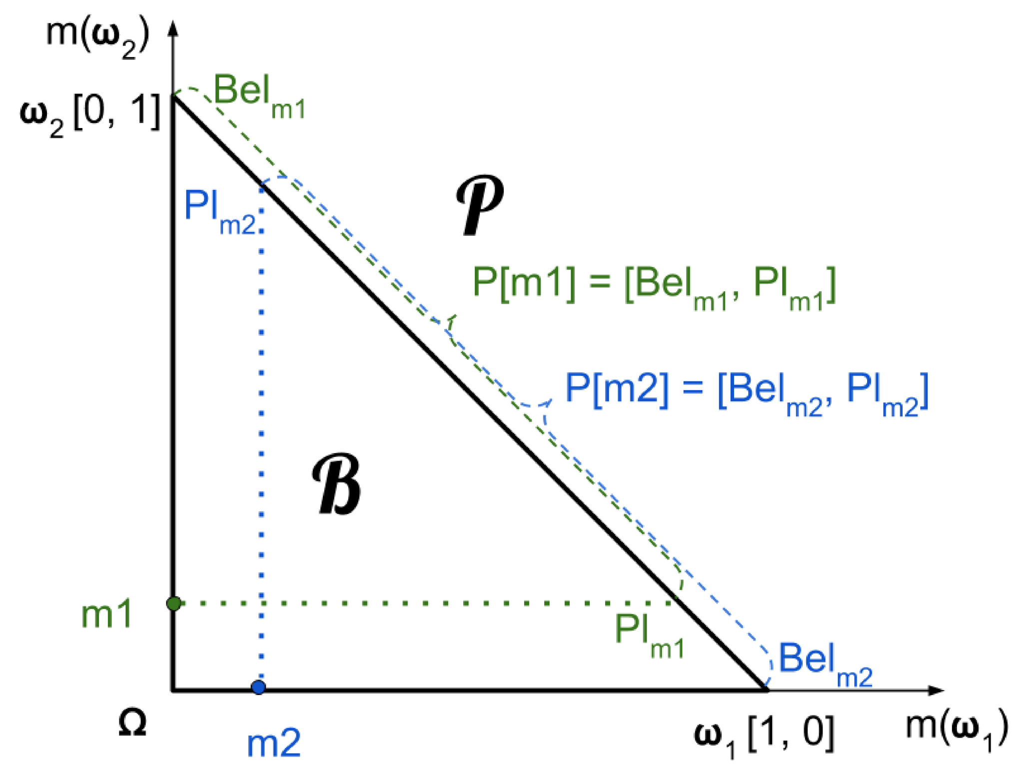
$$\kappa(m_1, m_2) = (m_1 \odot m_2)(\emptyset).$$

κ	m_4	m_5	m_6	m_7	m_8	m_9
m_1	0	0.01	0	0.09	0.09	0.1
m_2	0.09	0.09	0.1	0	0.01	0
m_3	0	0	0	0	0	0

Remark: m_1 is close to $m_4 \sim m_6$; m_2 is close to $m_7 \sim m_9$; m_3 is close to $m_4 \sim m_9$.

An imprecise probabilistic view

Each mass function is projected as interval probabilistic values, with bounds obtained from belief transform Bel and plausibility transform Pl.



Remark: This geometric representation allows to express the uncertainty (by probability) and the imprecision (by interval) simultaneously.

Conclusion and significance

- Combination rules with LCP are not compatible with any metric;
- Distance-based classification/clustering over evidential corpus is dubious;
- The interpretation of the mass functions must be clarified when selecting a distance.
- Imprecise probabilistic view is appropriate with the original interpretation.

Questions as perspectives:

- Since no metric is consistent with LCP, how do we theoretically evaluate their effectiveness/correctness?
- How can we relax the properties of metric to make the impossibility possible?
- How to properly classify/cluster objects with uncertainty/imprecision?

This paper is published in *IEEE Transactions on Artificial Intelligence*, 2022, doi: 10.1109/TAI.2022.3180973.