Preference Elicitation with Uncertainty: Extending Regret Based Methods with Belief Functions

Pierre-Louis Guillot[1] and Sebastien Destercke[1]

Heudiasyc laboratory, Compiègne 60200, France

Abstract. Preference elicitation is a key element of any multi-criteria decision analysis (MCDA) problem, and more generally of individual user preference learning. Existing efficient elicitation procedures in the literature mostly use either robust or Bayesian approaches. In this paper, we are interested in extending the former ones by allowing the user to express uncertainty in addition of her preferential information and by modelling it through belief functions. We show that doing this, we preserve the strong guarantees of robust approaches, while overcoming some of their drawbacks. In particular, our approach allows the user to contradict herself, therefore allowing us to detect inconsistencies or ill-chosen model, something that is impossible with more classical robust methods.

Keywords: Belief Functions · Preference Elicitation · Multicriteria Decision.

1 Introduction

Preference elicitation, the process through which we collect preference from a user, is an important step whenever we want to model her preferences. It is a key element of domains such as *multi-criteria decision analysis* (MCDA) or preference learning [7], where one wants to build a ranking model on multivariate alternatives (characterised by criteria, features, ...). Our contribution is more specific to MCDA, as it focuses on getting preferences from a single user, and not a population of them.

Note that within this setting, preference modelling or learning can be associated with various decision problems. Such problems most commonly include the **ranking** problem that consists in ranking alternatives from best to worst, the **sorting** problem that consists in classifying alternatives into ordered classes, and finally the **choice** problem that consists in picking a single best candidate among available alternatives. This article only deals with the choice problem but can be extended towards the ranking problem in a quite straightforward manner – commonly known as the iterative choice prodecure – by considering a ranking as a series of consecutive choices [1].

In order for the expert to make a recommendation in MCDA, she must first restrict her search to a set of plausible MCDA models. This is often done accordingly to *a priori* assumptions on the decision making process, possibly constrained by computational considerations.

In this paper, we will assume that alternatives are characterised by q real values, i.e. are represented by a vector in $\mathbb{R}^q$, and that preferences over them can be modelled by a value function $f : \mathbb{R}^q \to \mathbb{R}$ such that $a \succ b$ iff $f(a) > f(b)$. More specifically, we will look at weighted averages. The example below illustrate this setting. Our results can straightforwardly be extended to other evaluations functions (Choquet integrals, GAI, ...) in theory, but would face additional computational issues that would need to be solved.

Example 1 (choosing the best course). Consider a problem in which the DM is a student wanting to find the best possible course in a large set of courses, each of which has been previously associated a grade from 0 to 10 – 0 being the least preferred and 10 being the most preferred – according to 3 criteria: *usefulness*, *pedagogy* and *interest*. The expert makes the assumption that the DM evaluates each course according to a score computed by a weighted sum of its 3 grades. This is a strong assumption as it means for example that an increase of 0.5 in *usefulness* will have the same impact on the score regardless of the grades in *pedagogy* and *interest*. In such a set of models, a particular model is equivalent to a set of weights in $\mathbb{R}^3$. Assume that the DM preferences follow the model given by the weights $(0.1, 0.8, 0.1)$, meaning that she considers *pedagogy* to be eight time as important as *usefulness* and *interest* which are of equal importance. Given the grades reported in Table 1, she would prefer the *Optimization* course over the *Machine learning* course, as the former would have a 5.45 value, and the later a 3.2 value.

Machine learning:

usefulness	*pedagogy*	*interest*
8.5	1.5	10

Optimization:

usefulness	*pedagogy*	*interest*
3	5.5	2

Linear algebra:

usefulness	*pedagogy*	*interest*
7	5	5.5

Graph theory:

usefulness	*pedagogy*	*interest*
1	2	6

Table 1. Grades of courses

Beyond the choice of a model, the expert also needs to collect or elicit preferences that are specific to the DM, and that she could not have guessed according to *a priori* assumptions. Information regarding preferences that are specific to the DM can be collected by asking them to answer questions in several form such as the ranking of a subset of alternatives from best to worst or the choice of a preferred candidate among a subset of alternatives.

Example 2 (choosing the best course (continued)). In our example, directly asking for weights would make little sense (as our model may be wrong, and as the user cannot be expected to be an expert of the chosen model). A way to get this information from her would therefore be to ask her to pick her favorite out of two courses. Let's assume that when asked to choose between *Optimization* and *Graph theory*, she prefers the *Optimisation* course. The latter being better than the former in *pedagogy* and worse in *interest*, her answer is compatible with weights $(0.05, 0.9, 0.05)$ (strong preference for pedagogy over other criteria) but not with $(0.05, 0.05, 0.9)$ (strong preference for interest over other criteria). Her answer has therefore given the expert additional information on the preferential model underlying her decision. We will see later that this generates a linear constraint over the possible weights.

Provided we have made some preference model assumptions (our case here), it is possible to look for *efficient* elicitation methods, in the sense that they solve the decision problem we want to solve in a small enough, if not minimal number of questions. A lot of work has been specifically directed towards active elicitation methods, in which the set of questions to ask the DM is not given in advance but determined on the fly. In robust methods, this preferential information is assumed to be given with full certainty which leads to at least two issues. The first one is that elicitation methods thus do not account for the fact that the DM might doubt her own answers, and that they might not reflect her actual preferences. The second one, that is somehow implied by the first one, is that most robust active elicitation methods will never put the DM in a position where she could contradict either herself or assumptions made by the expert, as new questions will be built on the basis that previous answers are correct and hence should not be doubted. This is especially problematic when inaccurate preferences are given early on, or when the preference model is based on wrong assumptions.

This paper presents an extension of the Current Solution Strategy [3] that includes uncertainty in the answers of the DM by using the framework based on belief functions presented in [5]. Section 2 will present necessary preliminaries on both robust preference elicitation based on regret and uncertainty management based on belief functions. Section 3 will present our extension and some of the associated theoretical results and guarantees. Finally Section 4 will present some first numerical experiments that were made in order to test the method and its properties in simulations.

2 Preliminaries

2.1 Formalization

Alternatives and models: We will denote $\mathcal{X}$ the space of possible alternatives, and $\mathbb{X} \subseteq \mathcal{X}$ the subset of **available** alternatives at the disposal of our DM and about which a recommendation needs to be made. In this paper we will consider alternatives summarised by q real values corresponding to criteria, hence $\mathcal{X} \subseteq \mathbb{R}^q$. For any $x \in \mathcal{X}$ and $1 \leq i \leq q$, we denote by $x^i \in \mathbb{R}$ the evaluation of alternative x according to criterion i. We also assume that for any $x, y \in \mathcal{X}$ such that $x^i > y^i$ for some $i \in \{1, \dots, q\}$ and $x^l \geq y^l, \forall l \in \{1, \dots, q\} \setminus \{i\}$, x will always be strictly preferred to y – meaning that preferences respect *ceteris paribus* monotonicity, and we assume that criteria utility scale is given.

$\mathbb{X}$ is a finite set of k alternative such that $\mathbb{X} = \{x_1, x_2, \ldots, x_k\}$ with x_j the j-th alternative of $\mathbb{X}$. Let $\mathcal{P}(\mathbb{X})$ be a preference relation over $\mathbb{X}$, and $x, y \in \mathbb{X}$ be two alternatives to compare. We will state that $x \succ_{\mathcal{P}} y$ if and only if x is **strictly** preferred to y in the corresponding relation, $x \simeq_{\mathcal{P}} y$ if and only if x and y are **equally** preferred in the corresponding relation, and $x \succeq_{\mathcal{P}} y$ if and only if either x is strictly preferred to y or x and y are equally preferred.

Preference modelling and weighted sums: In this work, we focus on the case where the hypothesis set Ω of preference models is the set of weighted sum models[1]. A singular model ω will be represented by its vector of weights in $\mathbb{R}^q$, and ω will be used to describe indifferently the decision model and the corresponding weight vector. Ω can therefore be described as:

$$\Omega = \left\{ \omega \in \mathbb{R}^q \ : \ \omega^i \geq 0 \text{ and } \sum_{i=1}^{q} \omega^i = 1 \right\}.$$

Each model ω is associated to the corresponding aggregating evaluation function

$$f_\omega(x) = \sum_{i=1}^{q} \omega^i x^i,$$

and any two potential alternatives x, y in $\mathcal{X}$ can then be compared by comparing their aggregated evaluation:

$$x \succeq_\omega y \iff f_\omega(x) \geq f_\omega(y) \tag{1}$$

which means that if the model ω is known, $\mathcal{P}_\omega(\mathbb{X})$ is a **total** preorder over $\mathbb{X}$, the set of existing alternatives. Note that $\mathcal{P}_\omega(\mathbb{X})$ can be determined using pairwise relations $\succeq_\omega$. Weighted averages are a key model of preference learning whose linearity usually allows the development of efficient methods, especially in regret-based elicitation [2]. It is therefore an ideal starting point to explore other more complex functions, such as those that are linear in their parameters once alternatives are known (i.e., Choquet integrals, Ordered weighted averages).

2.2 Robust preference elicitation

In theory, obtaining a unique true preference model requires both unlimited time and unbounded cognitive abilities. This means that in practice, the best we can do is to collect information identifying a subset Ω' of possible models, and act accordingly. Rather than choosing a unique model within Ω', robust methods usually look at the inferences that hold for every model in Ω'. Let Ω' be the subset of models compatible with all the given preferential information, then we can define $\mathcal{P}_{\Omega'}(\mathbb{X})$, a **partial** preorder of robust preferences over $\mathbb{X}$, as follows:

$$x \succeq_{\Omega'} y \iff \forall \omega \in \Omega' \ f_\omega(x) \geq f_\omega(y). \tag{2}$$

The research question we address here is to find elicitation strategies that reduce Ω' as quickly as possible, obtaining at the limit an order $\mathcal{P}_{\Omega'}(\mathbb{X})$ having only one maximal element[2]. In practice, one may have to stop collecting information before that point, explaining the need for heuristic indicators of the fitness of competing alternatives as potential choices.

Regret based elicitation: Regret is a common way to assess the potential loss of recommending a given alternative under incomplete knowledge. It can help both the problem of making a recommendation and finding an efficient question. Regret methods use various indicators, such as the **regret** $\mathrm{R}_\omega(x, y)$ of choosing x over y according to model ω, defined as

$$\mathrm{R}_\omega(x, y) = f_\omega(y) - f_\omega(x). \tag{3}$$

[1] In principle, our methods apply to any value function with the same properties, but may have to solve computational issue that depends on the specific chosen hypothesis.

[2] Or in some cases a maximal set $\{x_1, \ldots, x_p\}$ of *equally preferred* elements s.t. $x_1 \simeq \ldots \simeq x_p$

From this regret and a set Ω' of possible models, we can then define the **pairwise max regret** as

$$\mathrm{PMR}(x, y, \Omega') = \max_{\omega \in \Omega'} \mathrm{R}_\omega(x, y) = \max_{\omega \in \Omega'} \left(f_\omega(y) - f_\omega(x)\right) \tag{4}$$

that corresponds to the maximum possible regret of choosing x over y for any model in Ω'. The **max regret** for an alternative x defined as

$$\mathrm{MR}(x, \Omega') = \max_{y \in \mathbb{X}} \mathrm{PMR}(x, y, \Omega') = \max_{y \in \mathbb{X}} \max_{\omega \in \Omega'} \left(f_\omega(y) - f_\omega(x)\right) \tag{5}$$

then corresponds to the worst possible regret one can have when choosing x. Finally the **min max regret** over a subset of models Ω' is

$$\mathrm{mMR}(\Omega') = \min_{x \in \mathbb{X}} \mathrm{MR}(x, \Omega') = \min_{x \in \mathbb{X}} \max_{y \in \mathbb{X}} \max_{\omega \in \Omega'} \left(f_\omega(y) - f_\omega(x)\right) \tag{6}$$

Picking as choice $x^* = \arg\min \mathrm{mMR}(\Omega')$ is then a robust choice, in the sense that it is the one giving the minimal regret in a worst-case scenario (the one leading to max regret).

Example 3 (choosing the best course (continued)). Let $\mathcal{X} = [0, 10]^3$ be the set of valid alternatives composed of 3 grades from 0 to 10 in respectively *pedagogy*, *usefulness* and *interest*. Let $\mathbb{X} = \{x_1, x_2, x_3, x_4\}$ be the set of available alternatives in which x_1 corresponds to the *Machine learning* course, x_2 corresponds to the *Optimization* course, x_3 corresponds to the *Linear algebra* course and x_4 corresponds to the *Graph theory* course, as reported in table 1. Let $x, y \in \mathbb{X}$ be two alternatives and Ω the set of weighted sum models, $\mathrm{PMR}(x, y, \Omega)$ can be computed by optimizing $\max_{\omega \in \Omega} \left[\omega^1(x^1 - y^1) + \omega^2(x^2 - y^2) + \omega^3(x^3 - y^3)\right]$. As this linear function of ω is optimized over a convex polytope Ω, it can easily be solved exactly using linear programming (LP). Results of $\mathrm{PMR}(x, y, \Omega)$ and $\mathrm{MR}(x, \Omega)$ are shown in Table 2. In this example, x_1 is the alternative with minimum max regret, and the most conservative candidate to answer the choice problem according to regret.

$x \backslash y$	x_1	x_2	x_3	x_4
x_1	0	**4**	3.5	0.5
x_2	**8**	0	4	4
x_3	**4.5**	0.5	0	0.5
x_4	**7.5**	3.5	6	0

x	MR
x_1	**4**
x_2	8
x_3	4.5
x_4	7.5

Table 2. Values of $\mathrm{PMR}(x, y, \Omega)$ (left) and $\mathrm{MR}(x, \Omega)$ (right)

Regret indicators are also helpful for making the elicitation strategy *efficient* and helping the expert ask relevant questions to the DM. Let Ω' and Ω'' be two sets of models such that $\mathrm{mMR}(\Omega') < \mathrm{mMR}(\Omega'')$. In the worst case, we are certain that $x^*_{\Omega'}$ the optimal choice for Ω' is less regretted than $x^*_{\Omega''}$ the optimal choice for Ω'', which means that we would rather have Ω' be our set of models than Ω''. Let $\mathcal{I}, \mathcal{I}'$ be two pieces of preferential information and $\Omega^{\mathcal{I}}$, $\Omega^{\mathcal{I}'}$ the sets obtained by integrating this information. Finding which of the two is the most helpful statement in the progress towards a robust choice can therefore be done by comparing $\mathrm{mMR}(\Omega^{\mathcal{I}})$ and $\mathrm{mMR}(\Omega^{\mathcal{I}'})$. An optimal elicitation process (w.r.t. minimax regret) would then choose the question for which the **worst** possible answer gives us a restriction on Ω that is the **most** helpful in providing a robust choice. However, computing such a question can be difficult, and the heuristic we present next aims at picking a nearly optimal question in an efficient and tractable way.

The Current Solution Strategy: Let's assume that Ω' is the subset of decision models that is consistent with every information available so far to the expert. Let's restrict ourselves to questions that consist in comparing pairs x, y of alternatives in $\mathbb{X}$. The DM can only answer with $\mathcal{I}_1 = x \succeq y$ or $\mathcal{I}_2 = x \preceq y$. A pair helpful in finding a robust solution as fast as possible can be computed as a solution to the following optimization problem that consists in finding the pair minimizing the **worst-case min max regret** :

$$\min_{(x,y) \in \mathbb{X}^2} \mathrm{WmMR}(\{x, y\}) = \min_{(x,y) \in \mathbb{X}^2} \max \left\{\mathrm{mMR}(\Omega' \cap \Omega^{x \succeq y}), \mathrm{mMR}(\Omega' \cap \Omega^{x \preceq y})\right\} \tag{7}$$

The current solution strategy (referred to as CSS) is a heuristic answer to this problem that has proved to be efficient in practice [3]. It consists in asking the DM to compare $x^* \in \arg \mathrm{mMR}(\Omega')$ the least regretted alternative to $y^* = \arg\max_{y \in \mathbb{X}} \mathrm{PMR}(x^*, y, \Omega')$ the one it could be the most regretted to (its "worst opponent"). CSS is efficient in the sense that it requires the computation of only one value of min max regret, instead of the $\mathcal{O}(q^2)$ required to solve (7).

Example 4 (Choosing the best course (continued).). Using the same example, according to table 2, we have $\mathrm{mMR}(\Omega) = \mathrm{MR}(x_1, \Omega) = \mathrm{PMR}(x_1, x_2, \Omega)$, meaning that x_1 is the least regretted alternative in the worst case and x_2 is the one it is most regretted to. The CSS heuristic consists in asking the DM to compare x_1 and x_2, respectively the *Machine learning* course and the *Optimization* course.

2.3 Uncertain preferential information

Two key assumptions behind the methods we just described are that (1) the initial chosen set Ω of models can perfectly describe the DM's choices and (2) the DM is an oracle, in the sense that any answer she provides truly reflects her preferences, no matter how difficult the question. This certainly makes CSS an efficient strategy, but also an unrealistic one. This means in particular that if the DM makes a mistake, we will just pursue with this mistake all along the process and will never question what was said before, possibly ending up with sub-optimal recommendations.

Example 5 (choosing the best course (continued)). Let's assume similarly to Example 2 that the expert has not gathered any preference from the DM yet, and that this time she asks her to compare alternatives x_1 and x_2 – respectively the *Machine learning* course and the *Optimization* course. Let's also assume similarly to Example 1 that the DM makes decisions according to a weighted sum model with weights $\omega^\star = (0.1, 0.8, 0.1)$. $f_{\omega^\star}(x_2) = 5.45 > 3.2 = f_{\omega^\star}(x_1)$, which means that she should prefer the *Optimization* course over the *Machine learning* course. However for some reason – such as her being unfocused or unsure about her preference – assume the DM's answer is inconsistent with $\omega^\star$ and she states that $x_1 \succeq x_2$ rather than $x_2 \succeq x_1$.

Then Ω' the set of model consistent with available preferential information is such that $\Omega' = \Omega^{x_1 \succeq x_2} = \{\omega \in \Omega \ : \ \sum_{i=1}^{3} \omega^i \left(x_1^i - x_2^i\right) \geq 0\} = \{\omega \in \Omega \ : \ \omega^2 \leq \frac{2}{3} - \frac{5}{24}\omega^1\}$, as represented in Figure 1. It is clear that $\omega^\star \notin \Omega'$: subsequent questions will only ever restrict Ω' and the expert will never get quite close to modelling $\omega^\star$.

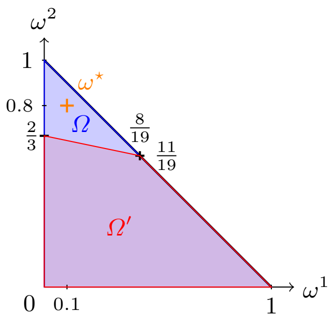

Fig. 1. Graphical representation of Ω, Ω' and $\omega^\star$

A similar point could be made if ω^*, the model according to which the DM makes her decision, does not even belong to Ω the set of weighted sums models that the expert chose.

As we shall see, one way to adapt min-max regret approaches to circumvent the two above difficulties can be to include a simple measure of how uncertain an answer is.

The belief function framework In classical CSS, the response to a query by the DM always implies a set of consistent models Ω' such that $\Omega' \subseteq \Omega$. Here, we allow the DM to give alongside her answer a confidence level $\alpha \in [0, 1]$, interpreted as how confident she is that this particular answer matches her preferences. In the framework developed in [5], such information is represented by a mass function on Ω', referred to as $m_\alpha^{\Omega'}$ and defined as :

$$m_\alpha^{\Omega'}(\Omega) = 1 - \alpha, \qquad m_\alpha^{\Omega'}(\Omega') = \alpha.$$

Such mass assignments are usually called simple support [13] and represent elementary pieces of uncertain information. A confidence level of 0 will correspond to a vacuous knowledge about the true model ω^*, and will in no way imply that the answer is wrong (as would have been the case in a purely probabilistic framework). A confidence level of 1 will correspond to the case of certainty putting a hard constraint on the subset of models to consider.

Remark 1. Note that values of α do not necessarily need to come from the DM, but can just be chosen by the analyst (in the simplest case as a constant) to weaken the assumptions of classical models. We will see in the experiments of Section 4 that such a strategy may indeed lead to interesting behaviours, without necessitating the DM to provide confidence degrees if she thinks the task is too difficult, or if the analyst thinks such self-assessed confidence is meaningless.

Dempster's rule Pieces of information corresponding to each answer will be combined through non-normalized Dempster's rule $+_\cap$. At step k, m_k the mass function capturing the current belief about the DM's decision model can thus be defined recursively as :

$$m_0 = m_1^\Omega \qquad \dots \qquad m_k = m_{k-1} +_\cap m_{\alpha_k}^{\Omega_k}. \tag{8}$$

This rule, also known as TBM conjunctive rule, is meant to combine distinct pieces of information. It is central to the Transferable Belief Model, that intends to justify belief functions without using probabilistic arguments [14].

Note that an information fusion setting and the interpretation of the TBM fit our problem particularly well as it assumes the existence of a unique true model ω^* underlying the DM's decision process, that might or might not be in our predefined set of models Ω. Allowing for an open world is a key feature of the framework. Let us nevertheless recall that non-normalized Dempster's rule $+_\cap$ can also be justified without resorting to the TBM [11, 8, 9].

In our case this independence of sources associated with two mass assignments $m_{\alpha_i}^{\Omega_i}$ and $m_{\alpha_j}^{\Omega_j}$ means that even though both preferential information account for preferences of the same DM, the answer a DM gives to the ith question does not directly impact the answer she gives to the jth question : she would have answered the same thing had their ith answer been different for some reason. This seems reasonable, as we do not expect the DM to have a clear intuition about the consequences of her answers over the set of models, nor to even be aware that such a set – or axioms underlying it – exists. One must however be careful to not ask the exact same question twice in short time range.

Since combined masses are all possibility distributions, an alternative to assuming independence would be to assume complete dependence, simply using the minimum rule [6] which among other consequences would imply a loss of expressivity [3] but a gain in computation [4].

As said before, one of the key interest of using this rule (rather than its normalised version) is to allow $m(\emptyset) > 0$, notably to detect either mistakes in the DM's answer (considered as an unreliable source) or a bad choice of model (under an open world assumption). Determining where the conflict mainly comes from and acting upon it will be the topic of future works. Note that in the specific case of simple support functions, we have the following result:

Proposition 1. *If $m_{\alpha_k}^{\Omega_k}$ are simple support functions combined through Dempster's rule, then*

$$m(\emptyset) = 0 \Leftrightarrow \exists \omega, Pl(\{\omega\}) = 1$$

with $Pl(\{\omega\}) = \sum_{E \subseteq \Omega, \omega \in E} m(E)$ the plausibility measure of model ω.

Proof. (Sketch) The $\Leftarrow$ part is obvious given the properties of Plausiblity measure. The $\Rightarrow$ part follows from the fact that if $m(\emptyset) = 0$, then all focal elements are supersets of $\bigcap_{i \in \{1,\dots,k\}} \Omega_i$, hence all contains at least one common element.

This in particular shows that $m(\emptyset)$ can, in this specific case, be used as an estimate of the logical consistency of the provided information pieces.

[3] For instance, no new values of confidence would be created when using a finite set $\{\alpha_1, \dots, \alpha_M\}$ for elicitation

[4] The number of focal sets increasing only linearly with the number of information pieces.

Consistency with robust, set-based methods : when an information $\mathcal{I}$ is given with full certainty $\alpha = 1$, we retrieve a so-called categorical mass $m_k\left(\Omega^{\mathcal{I}}\right) = 1$. Combining a set $\mathcal{I}_1, ..., \mathcal{I}_k$ of such certain information will end up in the combined mass

$$m_k\left(\bigcap_{i\in\{1,\ldots,k\}} \Omega_i\right) = 1$$

which is simply the intersection of all provided constraints, that may turn up either empty or non-empty, meaning that **inconsistency** will be a Boolean notion, i.e.,

$$m_k\left(\emptyset\right) = \begin{cases} 1 & \text{if } \bigcap_{i\in\{1,\ldots,k\}} \Omega_i = \emptyset \\ 0 & \text{otherwise} \end{cases}$$

Recall that in the usual CSS or minimax regret strategies, such a situation can never happen.

3 Extending CSS within the belief function framework

We now present our proposed extension of the Current Solution Strategy integrating confidence degrees and uncertain answers. Note that in the two first Sections 3.1 and 3.2, we assume that the mass on the empty set is null in order to parallel our approach with the usual one not including uncertainties. We will then consider the problem of conflict in Section 3.3.

3.1 Extending regret notions

Extending PMR: when uncertainty over possible models is defined through a mass function $2^{\Omega} \to [0, 1]$, subsets of Ω known as focal sets are associated to a value $m(\Omega')$ that correspond to the knowledge we have that ω belongs to Ω' and nothing more. The extension we propose averages the value of PMR on focal sets weighted by their corresponding mass :

$$\text{EPMR}(x, y, m) = \sum_{\Omega' \subseteq \Omega} m(\Omega').\text{PMR}(x, y, \Omega') \tag{9}$$

and we can easily see that in the case of certain answers ($\alpha = 1$), we do have

$$\text{EPMR}(x, y, m_k) = \text{PMR}\left(x, y, \left(\bigcap_{i\in\{1,\ldots,k\}} \Omega_i\right)\right) \tag{10}$$

hence formally extending Equation (4). When interpreting $m(\Omega')$ as the probability that ω belongs to Ω', EPMR could be seen as an expectation of PMR when randomly picking a set in 2^{Ω}.

Extending EMR: Similarly, we propose a weighted extension of maximum regret

$$\text{EMR}(x, m) = \sum_{\Omega' \subseteq \Omega} m(\Omega').\text{MR}(x, \Omega') = \sum_{\Omega' \subseteq \Omega} m(\Omega').\max_{y\in\mathbb{X}} \left\{\text{PMR}(x, y, \Omega')\right\}. \tag{11}$$

EMR is the **expectation** of the **maximal** pairwise max regret taken each time between x and $y \in \mathbb{X}$ its worst adversary – as opposed to a **maximum** considering each $y \in \mathbb{X}$ of the **expected** pairwise max regret between x and the given y, described by $\text{MER}(x, m) = \max_y \text{EPMR}(x, y, m)$. Both approaches would be equivalent to MR in the certain case, meaning that if $\alpha_i = 1$ then

$$\text{EMR}(x, m_k) = \text{MER}(x, m_k) = \text{MR}\left(x, \left(\bigcap_{i\in\{1,\ldots,k\}} \Omega_i\right)\right). \tag{12}$$

However EMR seems to be a better option to assess the max regret of an alternative, as under the assumption that the true model ω^* is within the focal set Ω', it makes more sense to compare x to its worst opponent within Ω', which may well be different for two different focal sets. Indeed, if ω^* the true model does in fact belong to Ω', decision x is only as bad as how big the regret can get for **any** adversarial counterpart $y_{\Omega'} \in \mathbb{X}$.

Extending mMR: we propose to extend it as

$$\mathrm{mEMR}(m) = \min_{x \in \mathbb{X}} \mathrm{EMR}(x, m). \tag{13}$$

mEMR **minimizes** for each $x \in \mathbb{X}$ the **expectation** of max regret and is different from the **expectation** of the **minimal** max regret for whichever alternative x is optimal, described by $\mathrm{EmMR}(m) = \sum_{\Omega'} m(\Omega') \min_{x \in \mathbb{X}} \mathrm{MR}(x, \Omega')$. Again, these two options with certain answers boil down to mMR as we have

$$\mathrm{mEMR}(m) = \mathrm{EmMR}(m) = \mathrm{mMR}\left(\bigcap_{i \in \{1,\ldots,k\}} \Omega_i \right). \tag{14}$$

The problem with EmMR is that it would allow for multiple possible best alternatives, leaving us with an unclear answer as to what is the best choice option, (arg min EmMR) not being defined. It indicates how robust in the sense of regret we expect **any** best answer to the choice problem to be, assuming there can be an optimal alternative for each focal set. In contrast, mEMR minimizes the max regret while restricting the optimal alternative x to be the same in all of them, hence providing a unique argument and allowing our recommendation system and elicitation strategy to give an optimal recommendation.

Extending CSS: our Evidential Current Solution Strategy (ECSS) then amounts, at step k with mass function m_k, to perform the following sequence of operations:

- Find $x^* = \arg \mathrm{mEMR}(m_k) = \arg\min_{x \in \mathbb{X}} \mathrm{EMR}(x, m_k)$;
- Find $y^* = \arg\max_{y \in \mathbb{X}} \mathrm{EPMR}(x^*, y, m_k)$;
- Ask the DM to compare x^*, y^* and provide α_k, obtaining $m_{\alpha_k}^{\Omega^k}$;
- Compute $m_{k+1} := m_k +_\cap m_{\alpha_k}^{\Omega^k}$
- Repeat until conflict is too high (red flag), budget of questions is exhausted, or $\mathrm{mEMR}(m_k)$ is sufficiently low

Finally, recommend $x^* = \arg \mathrm{mEMR}(m_k)$. Thanks to Equations (10), (12) and (14), it is easy to see that we retrieve CSS as the special case in which all answers are completely certain.

Example 6. Starting with intial mass function m_0 such that $m_0(\Omega) = 1$, the choice of CSS coincides with the choice of ECSS (all evidence we have is commited to $\omega \in \Omega$). With the values of PMR reported in Table 2 the alternatives the DM is asked to compare are x_1 the least regretted alternative and x_2 its most regretted counterpart. In accordance with her true preference model $\omega^* = (0.1, 0.8, 0.1)$, the DM states that $x_2 \succeq x_1$, i.e., she prefers the *Optimization* course over the *Machine learning* course, with confidence degree $\alpha = 0.7$. Let Ω_1 be the set of WS models in which x_2 can be preferred to x_1, which in symmetry with example 5 can be defined as $\Omega_1 = \{\omega \in \Omega : \omega^2 \geq \frac{2}{3} - \frac{5}{24}\omega^1\}$, as represented in Figure 2.

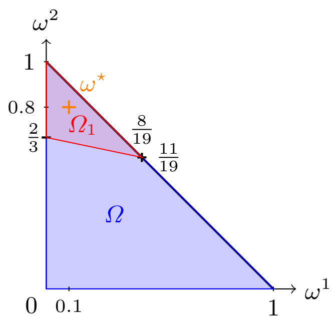

Fig. 2. Graphical representation of Ω, Ω_1 and $\omega^\star$

Available information on her decision model after step 1 is represented by mass function m_1 with

$$m_1(\ \Omega\ ) = 0.3$$
$$m_1(\ \Omega_1\ ) = 0.7$$

$x \backslash y$	x_1	x_2	x_3	x_4
x_1	0	**4**	3.5	0.5
x_2	0	0	$\frac{53}{38} \simeq \mathbf{1.39}$	-1
x_3	$-\frac{5}{6} \simeq -0.83$	**0.5**	0	$-\frac{11}{6} \simeq -1.83$
x_4	$\frac{109}{38} \simeq 2.87$	3.5	$\frac{81}{19} \simeq \mathbf{4.26}$	0

x	MR
x_1	4
x_2	$\frac{53}{38} \simeq 1.39$
x_3	**0.5**
x_4	$\frac{81}{19} \simeq 4.26$

Table 3. Values of PMR(x, y, Ω_1) (left) and MR(x, Ω_1) (right)

$x \backslash y$	x_1	x_2	x_3	x_4
x_1	0	**4**	3.5	0.5
x_2	**2.4**	0	$\frac{827}{390} \simeq 2.18$	0.5
x_3	$\frac{23}{30} \simeq \mathbf{0.77}$	0.5	0	$-\frac{68}{60} \simeq -1.13$
x_4	$\frac{809}{190} \simeq 4.26$	3.5	$\frac{909}{190} \simeq \mathbf{4.78}$	0

x	MR
x_1	4
x_2	$\frac{1283}{380} \simeq 3.38$
x_3	$\frac{23}{30} \simeq \mathbf{0.7}$
x_4	$\frac{1389}{380} \simeq 5.23$

Table 4. Values of EPMR(x, y, m_1) (left) and EMR(x, m_1) (right)

The values of PMR and MR on Ω_1 can be computed using LP as reported in table 3. Values of PMR and MR and have been previously computed on Ω as reported in Table 2. Values of EPMR and EMR can be then be deduced by combining them according to Equations (9) and (11), as reported in Table 4. In this example x_3 minimizes both MR and EMR and our extension agrees with the robust version as to which recommendation is to be made. However the most regretted counterpart to which the DM has to compare x_3 in the next step differs, as ECSS would require that she compares x_3 and x_1 rather than x_3 and x_2 for CSS.

3.2 Preserving the properties of CSS

This section discusses to what extent is ECSS consistent with three key properties of CSS:

1. CSS is monotonic, in the sense that the minmax regret mMR reduces at each iteration.
2. CSS provides strong guarantees, in the sense that the felt regret of the recommendation is ensured to be at least as bad as the computed mMR.
3. CSS produces questions that are non-conflicting (whatever the answer) with previous answers.

We would like to keep the first two properties at least in the absence of conflicting information, as they ensure respectively that the method will converge and will provide robust recommendations. However, we would like our strategy to raise questions possibly contradicting some previous answers, so as to raise the previously mentioned red flags in case of problems (unreliable DM or bad choice of model assumption). As shows the next property, our method also converges.

Proposition 2. *Let m_{k-1} and $m_{\alpha_k}^{\Omega^k}$ be two mass functions on Ω issued from ECSS such that $m_k(\emptyset) = \left[m_{k-1} +_\cap m_{\alpha_k}^{\Omega^k}\right](\emptyset) = 0$, then*

1. *$EPMR(x, y, m_k) \leq EPMR(x, y, m_{k-1})$*
2. *$EMR(x, m_k) \leq EMR(x, m_{k-1})$*
3. *$mEMR(m_k) \leq mEMR(m_{k-1})$*

Proof. (sketch) The two first items are simply due to the combined facts that on one hand we know [15] that applying $+_\cap$ means that m_k is a specialisation of m_{k-1}, and on the other hand that for any $\Omega'' \subseteq \Omega'$ we have $f(x, y, \Omega'') \leq f(x, y, \Omega')$ for any $f \in \{\text{PMR}, \text{MR}\}$. The third item is implied by the second as it consists in taking a minimum over a set of values of EMR that are all smaller.

Note that the above argument applies to any combination rule producing a specialisation of the two combined masses, including possibilistic minimum rule [6], Denoeux's family of w-based rules [4], etc. We can also show that the evidential approach, if we provide it with questions computed through CSS, is actually more cautious than CSS:

Proposition 3. *Consider the subsets of models $\Omega_1, \ldots, \Omega_k$ issued from the answers of the CSS strategy, and some values $\alpha_1, \ldots, \alpha_k$ provided a posteriori by the DM. Let m_{k-1} and $m^{\Omega^k}_{\alpha_k}$ be two mass functions issued from ECSS on Ω such that $m_k(\emptyset) = 0$. Then we have*

1. *$EPMR(x, y, m_k) \geq PMR(x, y, \left(\bigcap_{i\in\{1,\ldots,k\}} \Omega_i\right))$*
2. *$EMR(x, m_k) \geq MR(x, \left(\bigcap_{i\in\{1,\ldots,k\}} \Omega_i\right))$*
3. *$mEMR(m_k) \geq mMR(\left(\bigcap_{i\in\{1,\ldots,k\}} \Omega_i\right))$*

Proof. (sketch) The first two items are due to the combined facts that on one hand all focal elements are supersets of $\left(\bigcap_{i\in\{1,\ldots,k\}} \Omega_i\right)$ and on the other hand that for any $\Omega'' \subseteq \Omega'$ we have $f(x, y, \Omega'') \leq f(x, y, \Omega')$ for any $f \in \{\text{PMR}, \text{MR}\}$. Any value of EPMR or EMR is a weighted average over terms all greater than their robust counterpart on $\left(\bigcap_{i\in\{1,\ldots,k\}} \Omega_i\right)$, and is therefore greater itself. The third item is implied by the second as the biggest value of EMR is thus necessarily bigger than all the values of MR.

This simply shows that, if anything, our method is even more cautious than CSS. It is in that sense probably slightly too cautious in an idealized scenario – especially as unlike robust indicators our evidential extensions will never reach 0 – but provides guarantees that are at least as strong.

While we find the two first properties appealing, one goal of including uncertainties in the DM answers is to relax the third property, whose underlying assumptions (perfectness of the DM and of the chosen model) are quite strong. In Sections 3.3 and 4, we show that ECSS indeed satisfies this requirement, respectively on an example and in experiments.

3.3 Evidential CSS and conflict

The following example simply demonstrates that, in practice, ECSS can lead to questions that are possibly conflicting with each others, a feature CSS does not have. This conflict is only a possibility: no conflict will appear should the DM provide answers completely consistent with the set of models and what she previously stated, and in that case at least one model will be fully plausible[5] (see Proposition 1).

Example 7 (Choosing the best course (continued)). At step 2 of our example the DM is asked to compare x_1 to x_3 in accordance with Table 4. Even though it conflicts with ω^* the model underlying her decision the DM **has the option** to state that $x_1 \succeq x_3$ with confidence degree $\alpha > 0$, putting weight on Ω_2 the set of consistent model defined by $\Omega_2 = \{\omega \in \Omega \,:\, \sum_{i=1}^{q} \omega^i \left(x_1^i - x_3^i\right) \geq 0\} = \{\omega \in \Omega \,:\, \omega^2 \leq \frac{9}{16} - \frac{3}{8}\omega^1\}$. However as represented in figure 3, $\Omega_1 \cap \Omega_2 = \emptyset$.

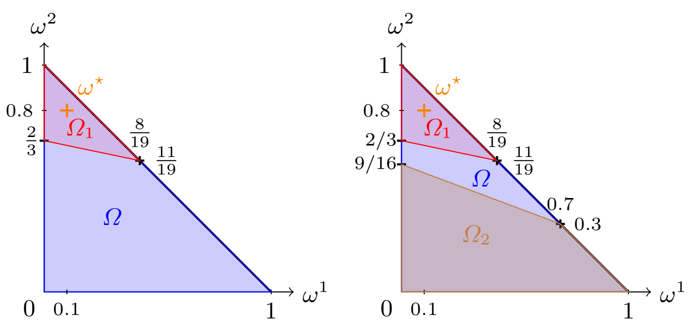

Fig. 3. Graphical representation of Ω, Ω_1, Ω_2 and $\omega^\star$

This means that $x_1 \succeq x_2$ and $x_3 \succeq x_1$ are not compatible preferences assuming the DM acts

[5] This contrasts with a Bayesian/probabilistic approach, where no model would receive full support in non-degenerate cases.

according to a weighted sum model. This can be either because she actually does not follow such a model, or because one of her answers did not reflect her actual preference (which would be the case here). Assuming she does state that $x_1 \succeq x_3$ with confidence $\alpha = 0.6$, information about the preference model at step 2 is captured through mass function m_2 defined as :

$$m_2 : \begin{array}{ll} \Omega \rightarrow 0.12 & \Omega_2 \rightarrow 0.18 \\ \Omega_1 \rightarrow 0.28 & \emptyset \rightarrow 0.42 \end{array}$$

Meaning that ECSS detects a degree of inconsistency equal to $m_2(\emptyset) = 0.42$.

This illustrating example is of course not representative of real situations, having only four alternatives, but clearly shows that within ECSS, conflict may appear as an effect of the strategy. In other words, we do not have to modify it to detect consistency issues, it will automatically seek out for such cases if $\alpha_i < 1$. In our opinion, this is clearly a desirable property showing that we depart from the assumptions of CSS.

However, the inclusion of conflict as a focal element in the study raises new issues, the first one being how to extend the various indicators of regret to this situation. In other words, how can we compute $\mathrm{PMR}(x, y, \emptyset)$, $\mathrm{MR}(x, \emptyset)$ or $\mathrm{MMR}(\emptyset)$? This question does not have one answer and requires careful thinking, however the most straightforward extension is to propose a way to compute $\mathrm{PMR}(x, y, \emptyset)$, and then to plug in the different estimates of ECSS. Two possibilities immediately come to mind:

- $\mathrm{PMR}(x, y, \emptyset) = \max_{\omega \in \Omega} \mathrm{R}_\omega(x, y)$, which takes the highest regret among all models, and would therefore be equivalent to consider conflict as ignorance. This amounts to consider Yager's rule [16] in the setting of belief functions. This rule would make the regret increases when conflict appears, therefore providing alerts, but this clearly means that the monotonicity of Proposition 3 would not longer hold. Yet, one could discuss whether such a property is desirable or not in case of conflicting opinions. It would also mean that elicitation methods are likely to try to avoid conflict, as it will induce a regret increase.
- $\mathrm{PMR}(x, y, \emptyset) = \min_{\omega \in \Omega} \max(0, \mathrm{R}_\omega(x, y))$, considering $\emptyset$ as the limit intersection of smaller and smaller sets consistent with the DM answers. Such a choice would allow us to recover monotonicity, and would clearly privilege conflict as a good source of regret reduction.

Such distinctions expand to other indicators, but we leave such a discussion for future works.

3.4 On computational tractability

ECSS requires, in principle, to compute PMR values for every possible focal elements, which could lead to an exponential explosion of the computational burden. We can however show that in the case of weighted sums and more generally of linear constraints, where PMR has to be solved through a LP program, we can improve upon this worst-case bound. We introduce two simplifications that lead to more efficient methods providing exact answers:

Using the polynomial number of elementary subsets. The computational cost can be reduced by using the fact that if $P_j = \{\Omega_{i_1}, \dots, \Omega_{i_k}\}$ is a partition of Ω_j, then :

$$\mathrm{PMR}(x, y, \Omega_j) = \max_{l \in \{i_1, \dots, i_k\}} \mathrm{PMR}(x, y, \Omega_l)$$

Hence, computing PMR on the partition is sufficient to retrieve the global PMR through a simple max. Let us now show that, in our case, the size of this partition only increases polynomially. Let $\Omega_1, \dots, \Omega_n$ be the set of models consistent with respectively the first to the nth answer, and $\Omega_i^C, \dots, \Omega_n^C$ their respective complement in Ω.

Due to the nature of the conjunctive rule $+_\cap$, every focal set Ω' of $m_k = m_1^\Omega +_\cap m_{\alpha_1}^{\Omega_1} +_\cap \dots +_\cap m_{\alpha_n}^{\Omega_n}$ is the union of elements of the partition $P_{\Omega'} = \{\tilde{\Omega}_1, \dots, \tilde{\Omega}_s\}$, with :

$$\tilde{\Omega}_k = \Omega \bigcap_{i \in U_k} \Omega_i \bigcap_{i \in \{1,\dots,n\} \setminus U_k} \Omega_i^C \ , U_k \subseteq \{1, \dots, n\}$$

Which means that for each Ω''s PMR can be computed using the PMR of its corresponding partition. This still does not help much, as there is a total of 2^n possible value of $\tilde{\Omega}_k$. Yet, in the case of convex domains cut by linear constraints, which holds for the weighted sum, the following theorem shows that the total number of elementary subset in Ω only increases polynomially.

Theorem 1. *[12, P 39] Let E be a convex bounded subset of F an euclidean space of dimension q, and $H = \{\eta_1, \dots, \eta_n\}$ a set of n hyperplanes in F such that $\forall i \in \{1, \dots, n\}$, η_i separates F into two subsets $F_0^{\eta_i}$ and $F_1^{\eta_i}$.*
To each of the 2^n possible $U \subseteq \{1, \dots, n\}$ a subset $F_U^H = F \bigcap_{i \in U} F_1^{\eta_i} \bigcap_{i \in \{1,\dots,n\} \setminus U} F_0^{\eta_i}$ can be associated.
Let $\Theta_H = \{U \subseteq \{1, \dots, n\} \ : \ F_U^H \cap E \neq \emptyset\}$ and $B_H = |\Theta_H|$, then

$$B_H \leq \Lambda_q^n = 1 + n + \binom{n}{2} + \cdots + \binom{n}{q}$$

Meaning that at most Λ_q^n of the F_U^H subsets have a non empty intersection with E.

In the above theorem (the proof of which can be found in [12], or in [10] for the specific case of $E \subset \mathbb{R}^3$), the subsets B_H are equivalent to $\tilde{\Omega}_k$, whose size only grow according to a polynomial whose power increases with q.

Using the polynomial number of extreme points in the simplex problem. Since we work with LP, we also know that optimal values will be obtained at extreme points. Optimization on focal sets can therefore ALL be done by maxing points at the intersection of q hyperplanes. This set of extreme point is

$$\mathcal{E} = \{\omega = \eta_i \cap \cdots \cap \eta_q \ : \ \{i_1, \dots, i_q\} \in \{1, \dots, n\}\} \tag{15}$$

with η_i the hyper-planes corresponding to the questions. We have $|\mathcal{E}| = \binom{n}{q} \in \mathcal{O}(n^q)$ which is reasonable whenever q is small enough (typically the case in MCDA). The computation of the coordinate of extreme points related to constraints of each subset can be done in advance for each $\omega \in \mathcal{E}$ and not once per subset and pair of alternatives, since $\mathcal{E}_{\Omega'}$ the set of extreme points of Ω' will always be such that $\mathcal{E}_{\Omega'} \subseteq \mathcal{E}$. The computation of the dot products necessary to compute $\mathrm{R}_\omega(x, y)$ for all $\omega \in \mathcal{E}$, $x, y \in \mathbb{X}$ can also be done once for each $\omega \in \mathcal{E}$, and not be repeated in each subset Ω' s.t. $\omega \in \mathcal{E}_{\Omega'}$. Those results indicate us that when q (the model-space dimension) is reasonably low and questions correspond to cutting hyper-planes over a convex set, ECSS can be performed efficiently. This will be the case for several models such as OWA or k-additive Choquet integrals with low k, but not for others such as full Choquet integrals, whose dimension if we have k criteria is $2^k - 2$. In these cases, it seems inevitable that one would resort to approximations having a low numerical impact (e.g., merging or forgetting focal elements having a very low mass value).

4 Experiments

To test our strategy and its properties, we proceeded to simulated experiments, in which the confidence degree was always constant. Such experiments therefore also show what would happen if we did not ask confidence degrees to the DM, but nevertheless assumed that she could make mistakes with a very simple noise model.

The first experiment reported in Figure 4 compares the extra cautiousness of EMR when compared to MR. To do so, simulations were made for several fixed degrees of confidence – including 1 in which case EMR coincides with MR – in which a virtual DM states her preferences with the given degree of confidence, and the value of EMR at each step is divided by the initial value so as to observe its evolution. Those EMR ratios were then averaged over 100 simulations for each degree. Results show that while high confidence degrees will have a limited impact, low confidence degrees (< 0.7) may greatly slow down the convergence.

The second experiment reported in Figure 5 aims at finding if ECSS and CSS truly generate different question strategies. To do so, we monitored the two strategies for a given confidence degree, and identify the first step k for which the two questions are different. Those values were averaged over 300 simulations for several confidence degrees. Results show that even for a high confidence degree ($\alpha = 0.9$) it takes in average only 3 question to see a difference. This shows that the methods are truly different in practice.

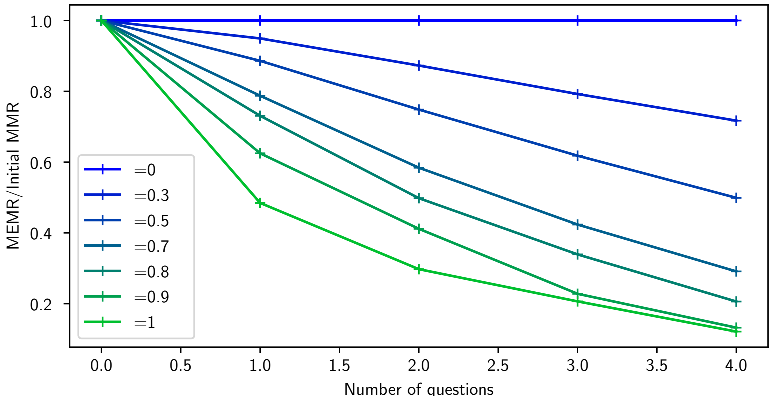

Fig. 4. Average evolution of min max regret with various degrees of confidence

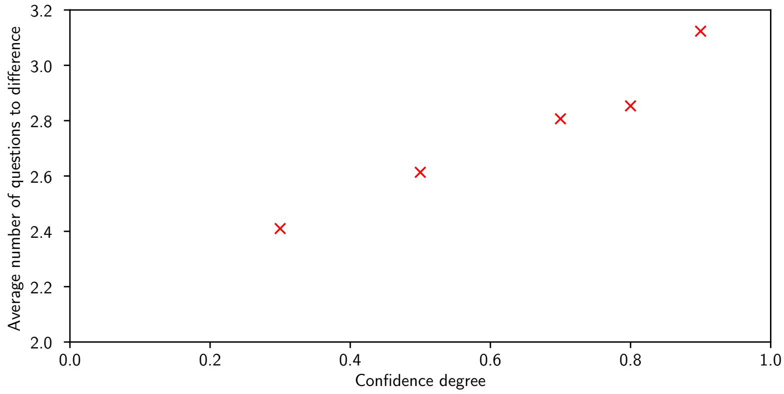

Fig. 5. Average position of the first different question in the elicitation process / degrees of confidence

The third experiment reported in Figure 6 is meant to observe how good $m(\emptyset)$ our measure of inconsistency is in practice as an indicator that something is wrong with the answers given by a DM. In order to do so simulations were made in which one of two virtual DMs answers with a fixed confidence degree and the value of $m(\emptyset)$ is recorded at each step. They were then averaged over 100 simulations for each confidence degree. The two virtual DMs behaved respectively completely randomly (RAND) and in accordance with a fixed weighted sum model (WS) with probability α and randomly with a probability $1-\alpha$. So the first one is highly inconsistent with our model assumption, while the second is consistent with this assumptions but makes mistakes.

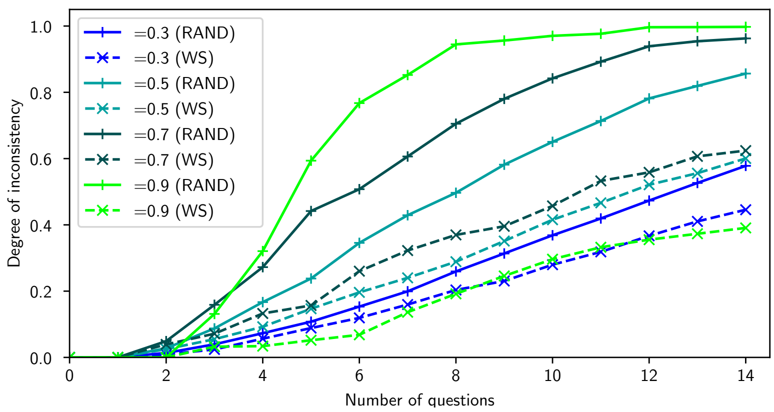

Fig. 6. Evolution of average inconsistency with a DM fitting the WS model and a randomly choosing DM

Results are quite encouraging: the inconsistency of the random DM with the model assumption is quickly identified, especially for high confidence degrees. For the DM that follows our model assumptions but makes mistakes, the results are similar, except for the fact that the conflict increase is not especially higher for lower confidence degrees. This can easily be explained that in case of low confidence degrees, we have more mistakes but those are assigned a lower weight, while in case of high confidence degrees the occasional mistake is quite impactful, as it has a high weight.

5 Conclusion

In this paper, we have proposed an evidential extension of the CSS strategy, used in robust elicitation of preferences.

We have studied its properties, notably comparing them to those of CSS, and have performed first experiments to demonstrate the utility of including confidence degrees in robust preference elicitation. Those latter experiments confirm the interest of our proposal, in the sense that it quickly identifies inconsistencies between the DM answer and model assumptions. It remains to check whether, in presence of mistakes from the DM, the real-regret (and not the computed one) obtained for ECSS is better than the one obtained for CSS.

As future works, we would like to work on the next step, i.e., identify the sources of inconsistency (whether it comes from bad model assumption or an unreliable DM) and propose correction strategies. We would also like to perform more experiments, and extend our approach to other decision models (Choquet integrals and OWA operators being the first candidates).

References

1. Benabbou, N., Gonzales, C., Perny, P., Viappiani, P.: Incremental elicitation of choquet capacities for multicriteria choice, ranking and sorting problems. Artificial Intelligence **246**, 152–180
2. Benabbou, N., Gonzales, C., Perny, P., Viappiani, P.: Minimax regret approaches for preference elicitation with rank-dependent aggregators. EURO Journal on Decision Processes **3**(1-2), 29–64
3. Boutilier, C., Patrascu, R., Poupart, P., Schuurmans, D.: Constraint-based optimization and utility elicitation using the minimax decision criterion. Artificial Intelligence **170**(8-9), 686–713
4. Denœux, T.: Conjunctive and disjunctive combination of belief functions induced by nondistinct bodies of evidence. Artificial Intelligence **172**(2-3), 234–264 (2008)
5. Destercke, S.: A generic framework to include belief functions in preference handling and multi-criteria decision. International Journal of Approximate Reasoning **98**, 62–77 (2018)
6. Destercke, S., Dubois, D.: Idempotent conjunctive combination of belief functions: Extending the minimum rule of possibility theory. Information Sciences **181**(18), 3925–3945
7. Fürnkranz, J., Hüllermeier, E.: Preference learning. Springer (2010)
8. Klawonn, F., Schweke, E.: On the axiomatic justification of dempsters rule of combination. International Journal of Intelligent Systems **7**(5), 469–478
9. Klawonn, F., Smets, P.: The dynamic of belief in the transferable belief model and specialization-generalization matrices. Proceedings of the 8th Conference on Uncertainty in Artificial Intelligence
10. Orlik, P., Terao, H.: Arrangements of hyperplanes, vol. 300. Springer Science & Business Media (2013)
11. Pichon, F., Denoeux, T.: The unnormalized dempsters rule of combination: A new justification from the least commitment principle and some extensions. Journal of Automated Reasoning **45**(1), 61–87
12. Schläfli, L., Wild, H.: Theorie der vielfachen Kontinuität, vol. 38. Springer-Verlag (2013)
13. Shafer, G.: A mathematical theory of evidence. Princeton University Press (1976)
14. Smets, P.: The combination of evidence in the transferable belief model. IEEE Transactions on Pattern Analysis and Machine Intelligence **12**(5), 447–458
15. Smets, P.: The application of the matrix calculus to belief functions. International Journal of Approximate Reasoning **31**(1-2), 1–30 (2002)
16. Yager, R.R.: On the dempster-shafer framework and new combination rules. Information sciences **41**(2), 93–137 (1987)