

CONSTRAINED OPTIMIZATION PROBLEMS UNDER UNCERTAINTY WITH COHERENT LOWER PREVISIONS

ERIK QUAEGBEUR, KEIVAN SHARIATMADAR, AND GERT DE COOMAN

ABSTRACT. We investigate a constrained optimization problem with uncertainty about constraint parameters. Our aim is to reformulate it as a (constrained) optimization problem without uncertainty. This is done by recasting the original problem as a decision problem under uncertainty. We give results for a number of different types of uncertainty models—linear and vacuous prevision, and possibility distributions—and for two common but different optimality criteria for such decision problems—maximinity and maximality. We compare our approach with other approaches that have appeared in the literature.

1. INTRODUCTION

Consider the following optimization problem:

$$\begin{aligned} & \text{maximize} && f(x) \\ & \text{subject to} && x \triangleleft Y, \end{aligned}$$

where f is a bounded real-valued objective function defined on a set \mathcal{X} of optimization variables x , Y is an uncertain variable taking values in a set \mathcal{Y} , and \triangleleft is a relation on $\mathcal{X} \times \mathcal{Y}$.¹ Optimization problems of this type are encountered in many fields and applications. We give some practical examples in Section 7.

To make this more concrete, consider the following toy problem. Suppose we know that Y is close to the real number c , and we are looking for the largest number x such that still $x \leq Y$. This problem could be represented mathematically as follows. The linguistic information ‘ Y is close to the real number c ’ can, for example, be modeled by a possibility distribution π on the set \mathbb{R} of real numbers with mode c ,² and we are looking for the solution of the following optimization problem:

$$\begin{aligned} & \text{maximize} && x \\ & \text{subject to} && x \leq Y, \end{aligned}$$

which can be seen as a special case of the more general version, where f is the identity function, and $\triangleleft := \leq$.³

The main difficulty with these problems is the uncertainty that appears in the constraint: because the value of Y is uncertain, it is uncertain for which values of x the constraint $x \triangleleft Y$

Key words and phrases. constrained optimization, maximinity, maximality, coherent lower prevision, linear prevision, vacuous prevision, possibility distribution.

¹The boundedness of f is not an essential technical requirement, but we make it here in order to allow us to work with the more standard imprecise probability models [25], which require boundedness. For discussions of imprecise probability models without this limitation, we refer to [23, 24].

²For a discussion of why this makes sense when modeling uncertainty with imprecise probability models, see [27].

³The technical problem that the identity function is not bounded on the reals can be overcome, albeit somewhat artificially, by letting \mathcal{X} be a bounded real interval that is large enough. Alternatively, we could extend the present discussion to include imprecise probability models that allow for unbounded gain functions, as discussed in [23, 24]. But this goes beyond the scope of this paper, which is intended as an exposition of ideas and first principles.

is satisfied, and therefore the domain over which $f(x)$ has to be maximized is not well known. The optimization problem is therefore ill-posed: it is underspecified, as there is no unique way of interpreting what is meant by maximizing a function over an uncertain domain; it might also be over-specified, as it may be infeasible in some particular case.

We will address this difficulty in Section 2 by reducing the optimization problem to a well-posed decision problem from which the uncertainties present in the description of the constraint have been eliminated. We mean by this that uncertain variables no longer appear in the problem; this is achieved by appropriately replacing the uncertain variables in the original formulation by their uncertainty models. For example, a naive way of doing this, when the uncertainty model is a classical probability distribution and $\mathcal{Y} \subseteq \mathbb{R}$, would consist of replacing the uncertain variable by its expectation.

We then investigate what results can be obtained for different types of uncertainty models for the uncertain variable Y —expectation operators or probability measures in Section 3, intervals or vacuous previsions [see, e.g., 25] in Section 4, and possibility distributions [see, e.g., 7] in Section 5—and for two different optimality criteria for decision problems—maximinity and maximality [see, e.g., 22]. After this, in Section 6, we compare our approach with related work.

The salient feature of our approach is that we use the same methodology and mathematical theory for solving all of these problems. Indeed, probability measures, intervals and (numerical) possibility measures can all be seen as special cases of a common and unifying uncertainty modeling framework: the theory of coherent lower previsions [see, e.g., 15, 25]. Since we are convinced that coherent lower previsions (and imprecise probability models in general) provide a useful interpretational framework for possibility distributions [see also our work on possibilistic lower previsions 4], we believe this paper establishes a novel approach to possibilistic optimization. It has the distinct advantage that the results we derive have a clear and useful behavioral interpretation.

In the remainder of this Introduction we provide some further background information. In Section 1.1, we give a brief introduction to coherent lower previsions and how they (mathematically) generalize both probability and possibility distributions as uncertainty models. In Section 1.2, we do the same for decision making with coherent lower previsions. We deal with some odds and ends in Section 1.3.

1.1. Coherent lower and upper previsions. Consider an uncertain variable Y that can take values in a set \mathcal{Y} . The uncertainty of an agent dealing with Y is described using an uncertainty model. The uncertainty model allows the agent to draw inferences about Y and functions thereof and make decisions in problems involving Y .

The classical uncertainty model is the probability distribution; in this context Y is usually called a random variable. Working with probability distributions or probability measures is equivalent to working with expectation operators [28]. We use the terminology of de Finetti [6] and call such expectation operators linear previsions. A linear prevision is a real functional P defined on the space $\mathcal{G}(\mathcal{Y})$ of bounded real-valued functions on \mathcal{Y} , that satisfies the following three so-called coherence conditions:

$$\text{Positivity:} \quad P(g) \geq \inf g, \quad (P1)$$

$$\text{Homogeneity:} \quad P(\lambda g) = \lambda P(g), \quad (P2)$$

$$\text{Additivity:} \quad P(g+h) = P(g) + P(h), \quad (P3)$$

for all bounded real-valued functions g, h on \mathcal{Y} and all real λ . In the interpretation of de Finetti [6], the real-valued functions $g = g(Y)$ are seen as gambles about Y and their

previsions $P(g)$ are seen as fair prices: the agent is prepared to sell a gamble $g(Y)$ for any price higher than $P(g)$ and buy it for any lower price.

Given a linear prevision P , the relationship with the corresponding (finitely additive) probability measure—also denoted by P —defined on the set of all events $2^{\mathcal{Y}}$ is given by $P(A) := P(I_A)$, where I_A is the indicator of A —it takes the value 1 on A and is 0 elsewhere. Probability measures P consequently satisfy three coherence conditions:

$$\text{Positivity: } P(A) \geq 0, \quad (P1')$$

$$\text{Unit Norm: } P(\mathcal{Y}) = 1, \quad (P2')$$

$$\text{Additivity: } P(A \cup B) = P(A) + P(B), \quad (P3')$$

where $A, B \subseteq \mathcal{Y}$ are disjoint. There is a one-to-one correspondence between probabilities and linear previsions (their expectation operators).

An important alternative—even orthogonal, one could say—uncertainty model is the possibility distribution. In this paper, we consider possibility distributions $\pi : \mathcal{Y} \rightarrow [0, 1]$ that are normal, meaning that $\sup \pi = 1$. One can equivalently work with possibility measures Π defined for events A of \mathcal{Y} , and by extension—using the Choquet integral—with possibility operators Π defined for gambles g on \mathcal{Y} [5]:

$$\Pi(A) := \sup_{y \in A} \pi(y), \quad \Pi(g) := \inf g + \int_{\inf g}^{\sup g} \Pi(\{y \in \mathcal{Y} : g(y) \geq t\}) dt. \quad (1)$$

The necessity measure N conjugate to Π is defined by $N(A) = 1 - \Pi(A^c)$ for events A , and more generally by $N(g) = -\Pi(-g)$ for gambles g . Necessity and possibility measures N and Π —and therefore also the operators derived from them—can, as a class onto themselves, be characterized by the following conditions:

$$\text{Positivity: } N(A) \geq 0, \quad \Pi(A) \geq 0, \quad (\Pi1)$$

$$\text{Normality: } N(\emptyset) = 0, \quad \Pi(\mathcal{Y}) = 1, \quad (\Pi2)$$

$$\text{Maxitivity: } N(\bigcap \mathcal{A}) = \inf_{A \in \mathcal{A}} N(A), \quad \Pi(\bigcup \mathcal{A}) = \sup_{A \in \mathcal{A}} \Pi(A), \quad (\Pi3)$$

where $\mathcal{A} \subseteq 2^{\mathcal{Y}}$, and if $\mathcal{A} = \emptyset$ the infimum is one and the supremum is zero by definition.

Both types of uncertainty model, in the incarnations described above, are special cases of the much more general coherent upper and lower previsions [see, e.g., 15, 25, for details and terminology]. These conjugate operators \underline{P} and \bar{P} , defined for all gambles g on \mathcal{Y} and related by $\bar{P}(g) = -\underline{P}(-g)$, are characterized by the following three conditions:⁴

$$\text{Positivity: } \underline{P}(g) \geq \inf g, \quad \bar{P}(g) \leq \sup g, \quad (\underline{P}1)$$

$$\text{Positive Homogeneity: } \underline{P}(\lambda g) = \lambda \underline{P}(g), \quad \bar{P}(\lambda g) = \lambda \bar{P}(g), \quad (\underline{P}2)$$

$$\text{Super/Sub-additivity: } \underline{P}(g+h) \geq \underline{P}(g) + \underline{P}(h), \quad \bar{P}(g+h) \leq \bar{P}(g) + \bar{P}(h), \quad (\underline{P}3)$$

where $g, h \in \mathcal{G}(\mathcal{Y})$ and $\lambda > 0$. Their restriction to (indicators of) events are called coherent lower and upper probabilities. In the behavioral interpretation of Walley [25], who follows the lead of de Finetti [6] in this regard, lower and upper previsions for gambles are again seen as prices: respectively the agent's supremum acceptable buying price and infimum acceptable selling price.

⁴The assumption that \underline{P} and \bar{P} are defined on the whole of $\mathcal{G}(\mathcal{Y})$ is not a trivial one in general: so-called natural extension from a partial specification requires solving a linear programming problem [25, Ch. 3]. However, for the cases looked at in this paper, natural extension just requires calculating a (Choquet) integral in the most complex case, which is far less computationally demanding.

When the lower prevision coincides with the upper prevision, they are linear previsions [see, e.g., 25]. Possibility operators are a special type of coherent upper previsions and necessity operators are similarly special coherent lower previsions [for more details about their relation, see 3, 5, 16, 26, 27]. In this paper, we use the theory of coherent lower (and upper) previsions as a unifying framework for linear previsions and possibility/necessity operators. All the uncertainty models we deal with fall into either category, but thanks to this framework, we can treat constrained optimization problems with uncertain variables described by both types in the constraint specification in a unified way and with a unified interpretation. It also leaves the door open to similarly deal with problems involving other types of uncertainty models under the coherent lower prevision umbrella.

It will also be useful to shed some light on the link between coherent lower previsions on the one hand, and sets of linear previsions (or probabilities) on the other. Given a coherent lower prevision \underline{P} , we can consider its set of dominating linear previsions:

$$\mathcal{M}(\underline{P}) := \{P: (\forall g \in \mathcal{G}(\mathcal{Y})) P(g) \geq \underline{P}(g)\}$$

This set is non-empty, convex, and closed (in the topology of point-wise convergence [see, e.g., 25]). Moreover, it turns out that \underline{P} is the lower envelope of the set $\mathcal{M}(\underline{P})$:

$$\underline{P}(g) = \min\{P(g): P \in \mathcal{M}(\underline{P})\} \text{ for all gambles } g \text{ on } \mathcal{Y}. \quad (2)$$

These correspondences tell us that, at least from a mathematical point of view, working with a coherent lower prevision is equivalent to working with a convex closed set of linear previsions (or probabilities).

We close this section with some properties that all coherent lower and upper previsions have, and which we use further on in the paper:

$$\text{Constant additivity: } \underline{P}(g + \mu) = \underline{P}(g) + \mu, \quad (\underline{P4})$$

$$\text{Mixed super-additivity: } \bar{P}(g + h) \geq \underline{P}(g) + \bar{P}(h), \quad (\underline{P5})$$

where $g, h \in \mathcal{G}(\mathcal{Y})$ and $\mu \in \mathbb{R}$.

1.2. Decision making with coherent lower previsions. We now turn to decision making, and consider a subject who may choose between a number of acts, or decisions, x in a set \mathcal{X} , the outcome of which depends on the uncertain variable Y . To be more specific, we assume that with each possible decision x there is associated a gain function g_x on \mathcal{Y} , with the following interpretation: if an agent makes decision x , then the corresponding outcome of this decision depends on the actual value y that Y assumes, and then has corresponding utility $g_x(y)$.

We will assume that each g_x is a bounded real-valued function, a gamble on \mathcal{Y} . Making decisions in the face of uncertainty about Y is in this sense taking gambles on the value of Y . We will also assume that all utilities are expressed in units of some linear utility scale [for more details, see 25].

If our agent has beliefs about Y expressed in terms of a coherent lower prevision \underline{P} —and therefore supremum buying prices for gambles about Y —we can use these beliefs to induce a binary relation on the set \mathcal{X} of all decisions. We say that he strictly prefers decision x_1 to decision x_2 , and we write $x_1 \succ x_2$, if he is willing to pay some strictly positive price in order to exchange g_{x_2} for g_{x_1} :

$$x_1 \succ x_2 \Leftrightarrow \underline{P}(g_{x_1} - g_{x_2}) > 0. \quad (3)$$

This definition gives an elegant and useful behavioral interpretation to the relation \succ , a strict partial order (irreflexive and transitive) on \mathcal{X} . Using Equation (2) and the properties

of linear prevision, we see that the equivalence

$$x_1 \succ x_2 \Leftrightarrow (\forall P \in \mathcal{M}(\underline{P})) P(g_{x_1}) > P(g_{x_2}) \quad (4)$$

also allows us to give a robustness interpretation to this strict partial order: $x_1 \succ x_2$ if and only if x_1 has a strictly higher expected utility than x_2 for all the probabilities that are compatible with the lower prevision \underline{P} .

It is important to realize that, unless \underline{P} is a linear prevision, \succ is not necessarily a total order: there may be x_1 and x_2 that are incomparable in the sense that neither $x_1 \succ x_2$ nor $x_2 \succ x_1$.

Given a set of decisions \mathcal{X} , we should aim for those decisions that are as high as possible in the order \succ ; we are looking for those decisions x that are maximal, or undominated, in that no decision z is considered better:

$$\begin{aligned} x \text{ is maximal} &\Leftrightarrow (\forall z \in \mathcal{X}) z \not\succeq x \Leftrightarrow (\forall z \in \mathcal{X}) \underline{P}(g_z - g_x) \leq 0 \\ &\Leftrightarrow (\forall z \in \mathcal{X}) \bar{P}(g_x - g_z) \geq 0 \Leftrightarrow \inf_{z \in \mathcal{X}} \bar{P}(g_x - g_z) \geq 0. \end{aligned} \quad (5)$$

When \underline{P} is a linear prevision P , we see that $x_1 \succ x_2 \Leftrightarrow P(g_{x_1}) > P(g_{x_2})$; in this case the order \succ is a total strict order. Therefore a decision x is maximal if and only if it has the highest expected gain $P(g_x)$, i.e., we maximize expected utility:

$$x \text{ is maximal} \Leftrightarrow x \in \operatorname{argmax}_{z \in \mathcal{X}} P(g_z).$$

This leads us to consider an alternative, albeit less well justified criterion for optimality of a decision under a coherent lower prevision \underline{P} :

$$x \text{ is maximin} \Leftrightarrow x \in \operatorname{argmax}_{z \in \mathcal{X}} \underline{P}(g_z), \quad (6)$$

or, in other words, if it maximizes the lower expected gain. Maximin solutions can be seen as resulting from worst-case reasoning; the solutions resulting from the best-case reasoning analogue are called maximax solutions and are obtained by replacing \underline{P} by \bar{P} in Equation (6).

Any maximin solution is automatically also maximal, since it follows from the mixed super-additivity ($\underline{P5}$) of \bar{P} that:

$$\bar{P}(g_x - g_z) \geq \underline{P}(g_x) + \bar{P}(-g_z) = \underline{P}(g_x) - \underline{P}(g_z).$$

Whether or not there are maximal and maximin solutions depends on the behavior of $\inf_{z \in \mathcal{X}} \bar{P}(g_x - g_z)$ and $\underline{P}(g_x)$ as functions of x on \mathcal{X} . For more details, properties and further interpretation of criteria for optimal decision making with imprecise probabilities, we refer to [22].

1.3. Some odds and ends. We work with general spaces \mathcal{X} and \mathcal{Y} and relations \triangleleft , and give illustrations for the particular case where $\mathcal{X} = \mathcal{Y} := \mathbb{R}$ and $\triangleleft := \leq$. In case a general treatment is too involved, we will restrict attention to this particular case.

Some useful (abuse of) notation:

$$\begin{aligned} x \triangleleft &:= \{y \in \mathcal{Y} : x \triangleleft y\}, & \overline{B \triangleleft} &:= \bigcup_{x \in B} x \triangleleft, & \underline{B \triangleleft} &:= \bigcap_{x \in B} x \triangleleft, \\ \triangleleft y &:= \{x \in \mathcal{X} : x \triangleleft y\}, & \overline{\triangleleft A} &:= \bigcup_{y \in A} \triangleleft y, & \underline{\triangleleft A} &:= \bigcap_{y \in A} \triangleleft y, \end{aligned} \quad (7)$$

where, as needed, $x \in \mathcal{X}$, $y \in \mathcal{Y}$, $B \subseteq \mathcal{X}$, and $A \subseteq \mathcal{Y}$. Similar definitions can be used for the complementary relation \triangleleft instead of \triangleleft . Note that $\underline{B \triangleleft} \subseteq \overline{B \triangleleft}$ and $\underline{\triangleleft A} \subseteq \overline{\triangleleft A}$.

2. REFORMULATION AS A DECISION PROBLEM UNDER UNCERTAINTY

Let us have a fresh look at the original optimization problem, now with the decision making approach described in Section 1.2 in mind:

$$\begin{aligned} & \text{maximize} && f(x) \\ & \text{subject to} && x \triangleleft Y. \end{aligned}$$

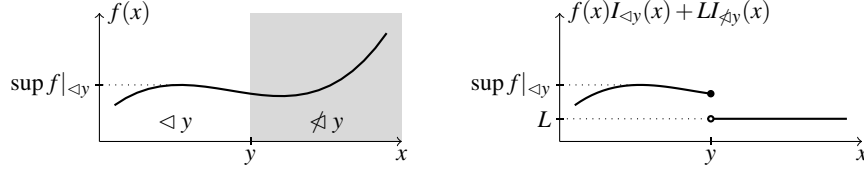
What is clear is that calling some x the solution to this problem corresponds to deciding that x is optimal in some sense. What is not yet clear is the gain function associated to each possible decision x in \mathcal{X} .

The gain functions we propose are based on the reformulation of the original optimization problem as an *unconstrained* one:

$$\text{maximize} \quad f(x)I_{\triangleleft Y}(x) + LI_{\not\triangleleft Y}(x).$$

In this expression, $I_{\triangleleft Y}$ and $I_{\not\triangleleft Y}$ are the indicator functions (cf. Section 1.1) of the sets $\triangleleft Y$ and $\not\triangleleft Y$ of optimization variables x that respectively do and do not satisfy the constraint (cf. Section 1.3); as Y is an uncertain variable, these are uncertain sets. Furthermore, L is the penalty value, a real number that should be chosen strictly smaller than $\inf f$ to make sure that breaking the constraint is penalized, i.e., leads to an objective function value that is inferior to any value that can be obtained when not breaking the constraint.

To see the intuition behind this reformulation, consider the case that $Y = y$ and the illustration below, in which $\mathcal{X} = \mathcal{Y} := \mathbb{R}$ and $\triangleleft := \leq$:



We used the notation $f|_B$ to denote restriction of f to the subset B of its domain. If—as in this illustration—there is no uncertainty about Y , then the unconstrained problem is equivalent to the original one due to the restriction placed on the penalty value.

In case there is uncertainty about Y , the objective function of the unconstrained problem provides us with a gain function $G_x(Y) := f(x)I_{\triangleleft Y}(x) + LI_{\not\triangleleft Y}(x)$ for each possible decision x in \mathcal{X} . Given that $I_{\triangleleft Y}(x) = I_{x \triangleleft}(Y)$, we can write:

$$G_x = f(x)I_{x \triangleleft} + LI_{x \not\triangleleft} = L + (f(x) - L)I_{x \triangleleft} = f(x) + (L - f(x))I_{x \not\triangleleft}. \quad (8)$$

By associating this form of gain function with the original optimization problem and by choosing a decision criterion, we give meaning to the original optimization problem under uncertainty. Next to its intuitive appeal, we think this choice of gain function is reasonable, because—together with either maximality or minimality as the decision criterion—(i) it results in the standard constrained optimization problem when there is no uncertainty about Y , (ii) the penalty value L allows us to quantify the impact of breaking the constraint (which would not be possible when just replacing Y with an expected value of some sort in the original optimization problem), and (iii) it is quite simple. More complex gain functions could also have advantages, but the complexity of the resulting decision problems may be practically prohibitive.

Now, given our choice of gain function and given that our uncertainty about Y is quantified by a coherent lower prevision \underline{P} , what becomes of the two decision criteria we are considering? By using Equation (8) and the constant additivity ($\underline{P}4$) and positive homogeneity ($\underline{P}2$)

of \underline{P} , we see—using the shorthand $\underline{P}(x \triangleleft) := \underline{P}(I_{x \triangleleft})$ (cf. Section 1.1)—that

$$\underline{P}(G_x) = \underline{P}(L + (f(x) - L)I_{x \triangleleft}) = L + (f(x) - L)\underline{P}(x \triangleleft). \quad (9)$$

Plugging this into the maximin criterion (6), we find that the set of maximin solutions is given by

$$\operatorname{argmax}_{x \in \mathcal{X}} \underline{P}(G_x) = \operatorname{argmax}_{x \in \mathcal{X}} (f(x) - L)\underline{P}(x \triangleleft), \quad (10)$$

since an additive constant has no impact on the location of the maximum. For the maximality criterion (5), no immediate simplifications are possible; plugging in the gain function from Equation (8) shows that some x in \mathcal{X} is maximal if and only if

$$\inf_{z \in \mathcal{X}} \bar{P}(G_x - G_z) = \inf_{z \in \mathcal{X}} \bar{P}\left((f(x) - L)I_{x \triangleleft} - (f(z) - L)I_{z \triangleleft}\right) \geq 0. \quad (11)$$

In the next three sections, we investigate what becomes of the gain function G_x -specific criteria (10) and (11) when considering different types of coherent lower previsions \underline{P} . So the main difficulty will be to find practical expressions for $\underline{P}(x \triangleleft)$ and $\inf_{z \in \mathcal{X}} \bar{P}(G_x - G_z)$, and using those to find sufficiently explicit expressions for the sets of maximin and maximal solutions, respectively. Before starting with that, we quickly look at the generality of the problem we consider.

2.1. Generality of the considered problem. It is good to realize that the optimization problem we consider is quite general, as problems with uncertain variables in the objective function can be reduced to it (making abstraction of possible interpretational issues):

$$\begin{aligned} & \text{maximize} && f(x, Y) \\ & \text{subject to} && x \triangleleft Y \end{aligned}$$

is formally equivalent to a problem that is of the same form as the original one,

$$\begin{aligned} & \text{maximize} && x_0 \\ & \text{subject to} && x_0 \leq f(x, Y) \text{ and } x \triangleleft Y, \end{aligned}$$

where $x_0 \in \mathbb{R}$. We have moved from $x \in \mathcal{X}$ to $(x_0, x) \in \mathbb{R} \times \mathcal{X}$ as the optimization variable.

We need to check whether this generalization is a correct one in the sense that it gives rise to the same maximin and maximal solutions for the original problem with an objective function f that does not depend on Y . The gain function—cf. Equation (8)—now becomes

$$G_{(x_0, x)} = L + (x_0 - L)I_{x_0 \leq (f(x))I_{x \triangleleft}}.$$

So the set of maximin solutions is

$$\begin{aligned} \operatorname{argmax}_{(x_0, x) \in \mathbb{R} \times \mathcal{X}} \underline{P}(G_{(x_0, x)}) &= \operatorname{argmax}_{x_0 \in \mathbb{R}, x \in \mathcal{X}} (x_0 - L)I_{x_0 \leq (f(x))I_{x \triangleleft}} \\ &= \begin{cases} \{(f(z), z) : z \in \operatorname{argmax}_{x \in \mathcal{X}} (f(x) - L)\underline{P}(x \triangleleft)\} & \text{if } \sup_{x \in \mathcal{X}} \underline{P}(x \triangleleft) > 0, \\ \mathbb{R} \times \mathcal{X} & \text{if } \sup_{x \in \mathcal{X}} \underline{P}(x \triangleleft) = 0. \end{cases} \end{aligned}$$

Ignoring the circumventable technical difference in case $\sup_{x \in \mathcal{X}} \underline{P}(x \triangleleft) = 0$, this is the same as found in Equation (10). Similarly, for the maximal solutions—cf. Equation (11)—we see

$$\inf_{(z_0, z) \in \mathbb{R} \times \mathcal{X}} \bar{P}(G_{(f(x), x)} - G_{(z_0, z)}) = \inf_{z \in \mathcal{X}} \bar{P}(G_{(f(x), x)} - G_{(f(z), z)}) = \inf_{z \in \mathcal{X}} \bar{P}(G_x - G_z).$$

Because $x_0 < f(x)$ implies $\underline{P}(G_{(f(x), x)} - G_{(x_0, x)}) = (f(x) - x_0)\underline{P}(x \triangleleft) > 0$ if $\underline{P}(x \triangleleft) > 0$ and thus $(f(x), x) \succ (x_0, x)$ by Equation (3), we can only have a (technical) difference whenever $\underline{P}(x \triangleleft) = 0$ for some x in \mathcal{X} such that $(f(x), x)$ is maximal, which can be circumvented by just positing that $x_0 < f(x) \Rightarrow (f(x), x) \succ (x_0, x)$.

3. LINEAR PREVISIONS

When the uncertainty about Y is described by a linear prevision (or expectation operator) P , for which both criteria reduce to maximizing expected utility, the set of optimal solutions is

$$\operatorname{argmax}_{x \in \mathcal{X}} P(G_x) = \operatorname{argmax}_{x \in \mathcal{X}} (f(x) - L)P(x \triangleleft). \quad (12)$$

For our prototypical particular case $\mathcal{X} = \mathcal{Y} := \mathbb{R}$ and $\triangleleft := \leq$. So for any x in \mathcal{X} we see that then $x \triangleleft = [x, +\infty)$. Define the distribution function of P by $F_P(x) := P((-\infty, x])$ for any x in \mathcal{X} . Then the set of optimal solutions for linear previsions P with continuous distribution functions is $\operatorname{argmax}_{x \in \mathcal{X}} (f(x) - L)(1 - F_P(x))$, which is straightforward to find using analytical or numerical methods.

4. VACUOUS PREVISIONS

In this section, we consider vacuous previsions, which express ignorance relative to a non-empty subset of the possible parameter values. The general case consists of a vacuous prevision relative to a subset A of \mathcal{Y} and for a given function g , so $\underline{P}(g) := \inf g|_A$ and $\overline{P}(g) := \sup g|_A$. It provides a model for the available information when we know that Y assumes a value in A but nothing more. This is corroborated by the fact that

$$\mathcal{M}(\underline{P}) = \{P : P(A) = 1\} \quad (13)$$

is the set of linear previsions (probabilities) associated with this vacuous prevision.

For our particular case, we will consider a vacuous prevision relative to an interval $[a, b] \subset \mathbb{R}$, so then $A := [a, b]$.

4.1. Maximinity. For maximinity in the general case, we combine the vacuous prevision's definition with Equation (10); the optimal x in \mathcal{X} are those that maximize

$$\underline{P}(G_x) = L + (f(x) - L) \inf I_{x \triangleleft} |_A. \quad (14)$$

So, by evaluating the expression $\inf I_{x \triangleleft} |_A$, we see that

$$\underline{P}(G_x) = \begin{cases} f(x), & A \subseteq x \triangleleft, \\ L, & \text{otherwise.} \end{cases} \quad (15)$$

The expression $A \subseteq x \triangleleft$ can be expanded to $(\forall y \in A) x \triangleleft y$ and from this and Equation (7), we can deduce it is equivalent to $x \in \triangleleft A$. So, maximizing $\underline{P}(G_x)$ is in the end the same as maximizing $f|_{\triangleleft A}$ and the set of optimal solutions is

$$\operatorname{argmax}_{x \in \triangleleft A} f(x). \quad (16)$$

Notice that this set does not depend on L .

In our particular case, as $\triangleleft [a, b] = \bigcap_{y \in [a, b]} \{x \in \mathcal{X} : x \leq y\} = (-\infty, a]$ by Equation (7), this set of solutions is $\operatorname{argmax}_{x \leq a} f(x)$. If f is the identity (or, more generally, strictly increasing), then the maximin solution is a : the largest number (in the sense of maximinity) smaller than Y , when all we know is that $a \leq Y \leq b$, is a .

4.2. Maximality. For maximality in the general case, we combine the vacuous prevision's definition with Equation (11); the optimal x in \mathcal{X} are those for which

$$\begin{aligned}\bar{P}(G_x - G_z) &= \bar{P}\left((f(x) - L)I_{x \triangleleft} - (f(z) - L)I_{z \triangleleft}\right) \\ &= \sup_{y \in A} \left((f(x) - L)I_{x \triangleleft}(y) - (f(z) - L)I_{z \triangleleft}(y) \right) \geq 0\end{aligned}\quad (17)$$

for all z in \mathcal{X} .

We look for an explicit expression for $\bar{P}(G_x - G_z)$ graphically, by considering all the possible positions A can be in relative to $x \triangleleft$ and $z \triangleleft$. For this, we first divide \mathcal{X}^2 in quadrants: We consider all positions of A relative to these quadrants. When A intersects some quadrant, it is shaded; so we need to consider all $2^4 - 1 = 15$ possible non-empty shadings and calculate $\bar{P}(G_x - G_z)$ for each of these shadings using the expression in Equation (17).

$x \triangleleft \cap z \triangleleft$	$x \triangleleft \cap z \not\triangleleft$
$x \not\triangleleft \cap z \triangleleft$	$x \not\triangleleft \cap z \not\triangleleft$

We find:

$$\begin{aligned}\bar{P}(G_x - G_z) &= \begin{cases} f(x) - L, & \text{[shaded top-left, top-right, middle-left, middle-right, bottom-left, bottom-right]} \\ 0, & \text{[shaded top-right, middle-right]} \\ \max\{0, f(x) - f(z)\}, & \text{[shaded middle-right]} \\ f(x) - f(z), & \text{[shaded middle-left, middle-right]} \\ L - f(z), & \text{[shaded bottom-left, bottom-right]} \end{cases} \\ &= \begin{cases} f(x) - L, & A \cap x \triangleleft \cap z \not\triangleleft \neq \emptyset, \\ 0, & A \cap x \triangleleft = \emptyset \wedge A \cap z \not\triangleleft \neq \emptyset, \\ \max\{0, f(x) - f(z)\}, & A \cap x \triangleleft \cap z \not\triangleleft = \emptyset \wedge A \cap x \triangleleft \neq \emptyset \wedge A \cap z \not\triangleleft \neq \emptyset, \\ f(x) - f(z), & A \cap x \triangleleft \neq \emptyset \wedge A \cap z \not\triangleleft = \emptyset, \\ L - f(z), & A \cap x \triangleleft = \emptyset \wedge A \cap z \not\triangleleft = \emptyset. \end{cases}\end{aligned}$$

The first three cases are always non-negative, the fourth one can be both positive and negative, and the last one is always negative.

Therefore, only the last two cases are important when checking the condition for an x in \mathcal{X} to be maximal, i.e., to avoid its being non-maximal:

$$\begin{aligned}\inf_{z \in \mathcal{X}} \bar{P}(G_x - G_z) &\geq 0 \\ &\Leftrightarrow (\forall z \in \mathcal{X}) \\ &\quad \left(\neg(A \cap x \triangleleft = \emptyset \wedge A \cap z \not\triangleleft = \emptyset) \wedge ((A \cap x \triangleleft \neq \emptyset \wedge A \cap z \not\triangleleft = \emptyset) \Rightarrow f(x) - f(z) \geq 0) \right) \\ &\Leftrightarrow (\forall z \in \mathcal{X}) \left((A \cap x \triangleleft \neq \emptyset \vee A \cap z \not\triangleleft \neq \emptyset) \wedge (A \cap x \triangleleft = \emptyset \vee A \cap z \not\triangleleft \neq \emptyset \vee f(x) \geq f(z)) \right) \\ &\Leftrightarrow (\forall z \in \mathcal{X}) \left(A \cap z \not\triangleleft \neq \emptyset \vee (A \cap x \triangleleft \neq \emptyset \wedge (A \cap x \triangleleft = \emptyset \vee f(x) \geq f(z))) \right) \\ &\Leftrightarrow (\forall z \in \mathcal{X}) \left(A \cap z \not\triangleleft \neq \emptyset \vee A \cap x \triangleleft \neq \emptyset \wedge (A \cap z \not\triangleleft \neq \emptyset \vee f(x) \geq f(z)) \right) \\ &\Leftrightarrow (\forall z \in \mathcal{X}) \left(A \cap z \not\triangleleft \neq \emptyset \vee A \cap x \triangleleft \neq \emptyset \right) \wedge (\forall z \in \mathcal{X}) \left(A \cap z \not\triangleleft \neq \emptyset \vee f(x) \geq f(z) \right) \\ &\Leftrightarrow (A \cap x \triangleleft \neq \emptyset \vee (\forall z \in \mathcal{X}) (A \cap z \not\triangleleft \neq \emptyset)) \wedge (\forall z \in \mathcal{X}) (A \cap z \not\triangleleft = \emptyset \Rightarrow f(x) \geq f(z)) \\ &\Leftrightarrow (x \in \overline{\triangleleft A} \vee \underline{\triangleleft A} = \emptyset) \wedge (\forall z \in \mathcal{X}) (z \in \underline{\triangleleft A} \Rightarrow f(x) \geq f(z)) \\ &\Leftrightarrow \underline{\triangleleft A} = \emptyset \vee (x \in \overline{\triangleleft A} \wedge f(x) \geq \sup f|_{\underline{\triangleleft A}}).\end{aligned}\quad (18)$$

So, if $\triangleleft A = \emptyset$, then all elements of \mathcal{X} are maximal; otherwise, only those x in $\overline{\triangleleft A}$ such that $f(x) \geq \sup f|_{\triangleleft A}$ are. Notice that the set of maximal solutions does not depend on L either.

In our particular case, as $\leq [a, b] = (-\infty, a]$ and similarly $\overline{\leq [a, b]} = (-\infty, b]$ by Equation (7), we see that those $x \leq b$ such that $f(x) \geq \sup f|_{(-\infty, a]}$ are maximal. Comparing them to those we found to be optimal under maximinity in the previous section, we see that now more elements can be optimal and that all maximin solutions are also maximal solutions. For instance, when f is the identity map, the set of all maximal elements is $[a, b]$: any element of $[a, b]$ is a largest number (in the sense of maximality) smaller than Y , when all we know is that $a \leq Y \leq b$. This suggests that the maximality criterion generally leads to a more robust solution, as was intended from the outset: see also Equation (4).

5. POSSIBILITY DISTRIBUTIONS

In this section, we consider possibility distributions. The general case consists of a possibility distribution π on \mathcal{Y} , so for a subset A of \mathcal{Y} , $\overline{P}(A) := \sup \pi|_A$ and $\underline{P}(A) := 1 - \sup \pi|_{A^c}$. For any function h on \mathcal{Y} , we can use the Choquet integral [see, e.g., 3, 17, for motivation] to associate a coherent upper prevision with this upper probability:

$$\overline{P}(h) := \inf h + \int_{\inf h}^{\sup h} \overline{P}(\{y \in \mathcal{Y} : h(y) \geq t\}) dt = \inf h + \int_{\inf h}^{\sup h} \sup \pi|_{\{y \in \mathcal{Y} : h(y) \geq t\}} dt. \quad (19)$$

For our prototypical example, we will consider a continuous fuzzy number—i.e., quasi-concave— π with leftmost mode $c \in \mathbb{R}$.

5.1. Maximinity. For maximinity in the general case, we combine the possibility distribution's definition with Equation (10); the optimal x in \mathcal{X} are those that maximize

$$\underline{P}(G_x) = L + (f(x) - L)(1 - \sup \pi|_{x \not\triangleleft}), \quad (20)$$

because $\underline{P}(x \triangleleft) = 1 - \overline{P}(x \not\triangleleft) = 1 - \sup \pi|_{x \not\triangleleft}$. So, maximizing $\underline{P}(G_x)$ corresponds to maximizing $(f(x) - L)(1 - \sup \pi|_{x \not\triangleleft})$ and the set of optimal solutions is

$$\operatorname{argmax}_{x \in \mathcal{X}} (f(x) - L)(1 - \sup \pi|_{x \not\triangleleft}). \quad (21)$$

Notice that this set generally depends on L .

In our particular case, as $\sup \pi|_{x \triangleleft} = \pi(\min\{x, c\})$, $\operatorname{argmax}(f - L)(1 - \pi)|_{< c}$ is the set of solutions. When f is the identity map and π increases linearly from 0 in a to 1 in $c > a$, the minimax solution is given by $\max\{a, \frac{L+c}{2}\}$, where $\frac{L+c}{2} < c$ because L is small with respect to f -values, i.e., a and c .

5.2. Maximality. For maximality in the general case, we combine the possibility distribution's definition with Equation (11); the optimal $x \in \mathcal{X}$ are those for which

$$\overline{P}(G_x - G_z) = \inf(G_x - G_z) + \int_{\inf(G_x - G_z)}^{\sup(G_x - G_z)} \overline{P}(\{y \in \mathcal{Y} : G_x(y) - G_z(y) \geq t\}) dt \geq 0 \quad (22)$$

for all z in \mathcal{X} .

To calculate a more explicit expression, we need to be able to determine the level set appearing as an argument of \overline{P} for every t in the range of $G_x - G_z$.

$$G_x(y) - G_z(y) = \begin{cases} \Delta_{xz} := f(x) - f(z), & y \in x \triangleleft \cap z \triangleleft, \\ \Delta_{xL} := f(x) - L, & y \in x \triangleleft \cap z \not\triangleleft, \\ -\Delta_{zL} := L - f(z), & y \in x \not\triangleleft \cap z \triangleleft, \\ 0, & y \in x \not\triangleleft \cap z \not\triangleleft. \end{cases}$$

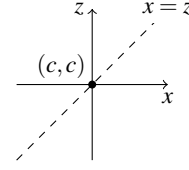
We furthermore need to know the relative order of these four possible values: Δ_{xL} is the largest, $-\Delta_{zL}$ is the smallest, and the relative order of 0 and Δ_{xz} depends on the relative order of $f(x)$ and $f(z)$. So there are two cases:

$$\{y \in \mathcal{Y} : G_x(y) - G_z(y) \geq t\} = \begin{cases} \text{if } f(x) \geq f(z) \text{ then } \begin{cases} x \triangleleft \cup z \triangleleft, & -\Delta_{zL} < t \leq 0, \\ x \triangleleft, & 0 < t \leq \Delta_{xz}, \\ x \triangleleft \cap z \triangleleft, & \Delta_{xz} < t \leq \Delta_{xL}, \end{cases} \\ \text{if } f(x) \leq f(z) \text{ then } \begin{cases} x \triangleleft \cup z \triangleleft, & -\Delta_{zL} < t \leq \Delta_{xz}, \\ z \triangleleft, & \Delta_{xz} < t \leq 0, \\ x \triangleleft \cap z \triangleleft, & 0 < t \leq \Delta_{xL}. \end{cases} \end{cases}$$

We see that in both cases the integrand is piecewise constant, and the integral is therefore a weighted sum of three terms:

$$\bar{P}(G_x - G_z) = -\Delta_{zL} + \begin{cases} \Delta_{zL}\bar{P}(x \triangleleft \cup z \triangleleft) + \Delta_{xz}\bar{P}(x \triangleleft) + \Delta_{zL}\bar{P}(x \triangleleft \cap z \triangleleft), & f(x) \geq f(z), \\ \Delta_{xL}\bar{P}(x \triangleleft \cup z \triangleleft) - \Delta_{xz}\bar{P}(z \triangleleft) + \Delta_{xL}\bar{P}(x \triangleleft \cap z \triangleleft), & f(x) \leq f(z). \end{cases} \quad (23)$$

Using the above expression to derive the maximal solutions turns out to be quite cumbersome in general. We investigate whether we can obtain some results in our particular case, for which $x \triangleleft \cap z \triangleleft = [x, z]$, $x \triangleleft = [x, +\infty)$, and $z \triangleleft = (-\infty, z]$. So the relative location of x , y , and c —the fuzzy number π 's left-most mode—will determine the values of $\bar{P}(x \triangleleft \cap z \triangleleft) = \sup \pi|_{[x, z]}$, $\bar{P}(x \triangleleft) = \sup \pi|_{[x, +\infty)}$, and $\bar{P}(z \triangleleft) = \sup \pi|_{(-\infty, z]}$. These upper probabilities are constant on the partitioning of (x, z) -space delineated by $x = z$, $x = c$, and $z = c$ shown on the right. We find



$$\bar{P}(x \triangleleft \cap z \triangleleft) = \frac{1}{\pi(z)} \begin{array}{|c|} \hline \pi(x) \\ \hline 0 \\ \hline \end{array}, \quad \bar{P}(x \triangleleft) = \frac{1}{\pi(x)} \begin{array}{|c|} \hline \pi(x) \\ \hline \pi(x) \\ \hline \end{array}, \quad \bar{P}(z \triangleleft) = \frac{1}{\pi(z)} \begin{array}{|c|} \hline 1 \\ \hline \pi(z) \\ \hline \end{array}. \quad (24)$$

Because of the maxitivity of possibility measures, we also find

$$\bar{P}(x \triangleleft \cup z \triangleleft) = \max\{\bar{P}(x \triangleleft), \bar{P}(z \triangleleft)\} = \frac{1}{\pi(x) \sim \pi(z)} \begin{array}{|c|} \hline 1 \\ \hline \pi(x) \sim \pi(z) \\ \hline \end{array}, \quad (25)$$

where $\pi(x) \sim \pi(z) := \max\{\pi(x), \pi(z)\}$.

Calculating $\bar{P}(G_x - G_z)$ for general objective functions f still results in quite complicated expressions. Therefore we further restrict ourselves to strictly increasing objective functions. After combining (23) with the expressions (24)–(25) above and some manipulations of the expressions so obtained—among them $\Delta_{xz} = \Delta_{xL} - \Delta_{zL}$ —, we find

$$\bar{P}(G_x - G_z) = \begin{array}{|c|} \hline \Delta_{xL} \\ \hline \pi(x)\Delta_{xL} \\ \hline \pi(x)\Delta_{xz} \\ \hline \end{array} \begin{array}{|c|} \hline \Delta_{xL} - \\ (1 - \pi(z))\Delta_{zL} \\ \hline \Delta_{xz} \\ \hline \pi(x)\Delta_{xL} - \\ (\pi(x) + 1 - \pi(x) \sim \pi(z))\Delta_{zL} \\ \hline \end{array}, \quad (26)$$

where the regions for which $\bar{P}(G_x - G_z)$ is not uniformly non-negative have been shaded light gray. So for strictly increasing f , we find that

$$x \leq c \text{ is maximal} \Leftrightarrow \sup_{x < z < c} (1 - \pi(z)) \Delta_{zL} \leq \Delta_{xL}, \quad (27)$$

$$x \geq c \text{ is maximal} \Leftrightarrow \sup_{z < c} (\pi(x) + 1 - \pi(x) \smile \pi(z)) \Delta_{zL} \leq \pi(x) \Delta_{xL}. \quad (28)$$

To compare this solution with the one we obtained in the vacuous case of Section 4.2, assume π has the interval $[a, b]$ as support. Then it follows from Equations (27) and (28) that no x outside $[a, b]$ can be maximal under π : $z := a$ dominates $x < a$ and $z := b$ dominates $x > b$. This should not surprise us, as the beliefs embedded in the possibility distribution π tell us at the very least that $a \leq Y \leq b$.

To illustrate that possibility distributions generally do provide more useful information, consider the case where f is the identity map and π increases linearly from 0 in a to 1 in $c > a$, which we considered at the end of Section 5.1 and for which we found a maximin solution $\max\{a, \mu\}$ with $\mu := \frac{L+c}{2}$. Equation (27)'s right-hand condition can be rewritten as $x \geq \max\{a, \frac{\mu-a}{c-a}L + \frac{c-\mu}{c-a}\mu\}$, from which it can also be verified that the maximin solution is maximal, as $L < \mu$.

6. A SELECTED OVERVIEW OF THE RELATED LITERATURE

Of course we are not the first to consider constrained optimization problems under (non-probabilistic) uncertainty, where the uncertainty is present in the constraint. Let us give a brief selected overview of the related literature and compare our approach to the ones found there. We describe results and approaches from the literature as much as possible in the language and notation of this paper, to simplify comparisons.

6.1. Vacuous previsions. Soyster [21] investigates so-called ‘inexact’ linear programming problems:

$$\begin{aligned} & \text{maximize} && c'x \\ & \text{subject to} && Yx \leq b \text{ and } x \geq 0, \end{aligned}$$

where $\mathcal{X} := \mathbb{R}^n$, $\mathcal{Y} := \mathbb{R}^{m \times n}$, $c \in \mathbb{R}^n$ (primes indicate transposition), $b \in \mathbb{R}^m$, and the lower prevision describing the uncertain matrix Y is vacuous relative to a convex subset $A := \times_{j=1}^n A_j$ of matrices, where each A_j is a convex subset of \mathbb{R}^m . The implicit decision criterion used is maximin, so this is a special case of Section 4.1 where, in Equation (16) we now have that $\underline{\Delta}A = \{x \geq 0 : \bar{y}x \leq b\}$, where \bar{y} is the matrix with components $\bar{y}_{ij} := \sup_{y \in A} y_{ij}$.

Soyster's approach can be seen as a special case of robust optimization [see, e.g., 2], in which the uncertainty about the parameters is expressed using an ‘uncertainty set’ of possible values for what we call the uncertain variable Y . The optimization problem is then transformed in a new one by requiring that the constraints are satisfied for all possible values at the same time. This is still a special case of our approach in which the uncertainty is expressed using vacuous previsions and where the optimality criterion used is maximin (cf. Section 4.1).

The work discussed above fits in the large body of literature that deals with optimization problems involving uncertainty of a kind that is formalized by vacuous previsions in this paper. Most of the work we have come across essentially uses maximin as the optimality criterion; the results we derive for the maximality case (cf. Section 4.2) seem not to have appeared before. (Interesting related work has also been done in which the notion of optimality criterion as we use it does not appear [see, e.g., 8, Ch. 3 and 4].) Vacuous

previsions are of course very simple uncertainty models. From there, a next step up in complexity are models using possibility distributions, which we deal with next.

6.2. Possibility distributions. With the increased complexity of possibility distributions comes an increased importance of the interpretation attached to the uncertainty models using them. We use the behavioral interpretation of Walley [25] discussed in Section 1.1. This in combination with the type of optimization problem that is the subject of this paper is something we have not come across yet in the literature. Nevertheless, it may be interesting to do a comparison with approaches inspired by other interpretations.

To discuss the approach based on the ideas of Bellman and Zadeh [1] concerning decision making in a fuzzy context, consider the following modification of our original problem:

$$\begin{aligned} & \text{maximize} && f(x) \\ & \text{subject to} && x \tilde{\leq} y, \end{aligned}$$

in which no uncertainty is assumed to be present in the parameters (viz. the constant $y \in \mathcal{Y}$ replacing Y), but where the constraints themselves are assumed to be fuzzy (indicated with a tilde). Mathematically, this fuzziness is expressed using a possibility distribution $\pi_{\langle y}$ on \mathcal{X} that expresses to what degree the constraints are satisfied. The maximization itself is also taken to be fuzzy, which is expressed mathematically using a possibility distribution $\pi_{f, \tilde{\leq} y}$ on \mathcal{X} that expresses to what degree the objective function or goal f is maximized; the subscript $\tilde{\leq} y$ is meant to indicate that (the support of) $\pi_{\langle y}$ first needs to be known to give an informed definition of $\pi_{f, \tilde{\leq} y}$.

The fuzzy constrained optimization problem then gives rise to a fuzzy decision that is some conjunction—e.g., minimization— $\pi_{f, \tilde{\leq} y} \wedge \pi_{\langle y}$ of the goal and constraint possibility distributions. So compared to our approach, the uncertainty or fuzziness is here not expressed on \mathcal{Y} but on \mathcal{X} , but as in our approach, the important conceptual step in solving the problem is the reformulation as a decision problem. As related by Kaymak and Sousa [12, Sec. 3], in the tradition of Zimmermann [29] an additional step is taken by removing the fuzzy nature of the solution using an extra maximization, i.e., $\operatorname{argmax}_{x \in \mathcal{X}} \pi_{f, \tilde{\leq} y}(x) \wedge \pi_{\langle y}(x)$ is called optimal. This contrasts with our approach when using the maximality criterion, where we explicitly do not try to obtain a unique optimal solution.

The Bellman and Zadeh [1]-based approach sketched here is but one end of the spectrum of optimization problems involving possibility distributions in their mathematical formulation [see, e.g., 10, 13, 20]. At the other end of the spectrum, it is not the constraints and the maximization that are fuzzified, but the parameters, i.e., Y is replaced by a fuzzy set \tilde{y} , or, mathematically, a possibility distribution π on \mathcal{Y} .

Jamison and Lodwick [11] are two of many to do this for linear programming problems; we translate the for this discussion relevant ideas of their paper to our context. As we did in Section 2, their idea is also to first move to an unconstrained optimization problem, but now with a variable penalty factor:

$$\text{maximize} \quad f(x) - L_{\tilde{y}}(x)I_{\tilde{y}}(x),$$

where $L_{\tilde{y}}$ is a real-valued function on \mathcal{X} that gives the larger a penalty the more severely the constraint is broken. Then a fuzzy number—a possibility distribution on \mathbb{R} — $\tilde{G}_x := f(x) - L_{\tilde{y}}(x)I_{\tilde{y}}(x)$ is associated with every x in \mathcal{X} using the extension principle; these can be seen as fuzzy gains. They define the solutions of the problem to be those x with a maximal mid-point average $\frac{1}{2} \int_0^1 (\max G_x(t) - \min G_x(t)) dt$, where $G_x(t) := \{r \in \mathbb{R} : \tilde{G}_x(r) \geq t\}$ is the level set at t . This optimality criterion lies qua idea in between maximinity and maximaxity, but not qua execution, as the way in which we use possibility distributions to express uncertainty differs markedly.

7. PRACTICAL EXAMPLES

The type of optimization problem we have considered is encountered in many fields and applications. In this section, we give one general practical example problem and two specific ones from the literature.

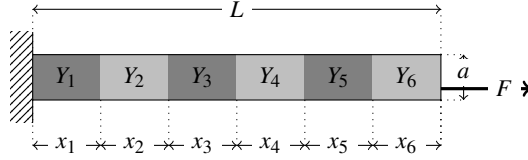
Linear programming problems. This classical optimization problem falls into our standard form in case the constraints' coefficients are parametrized:

$$\begin{aligned} & \text{maximize} && c^T x \\ & \text{subject to} && A(Y)x \leq b(Y), \end{aligned}$$

where $\mathcal{X} := \mathbb{R}^n$, $c \in \mathbb{R}^n$, $b : \mathcal{Y} \rightarrow \mathbb{R}^m$, and $A : \mathcal{Y} \rightarrow \mathbb{R}^{m \times n}$.

Finite element analysis – the loaded beam problem [see, e.g., 14]: minimize the production cost of a beam of length L and cross-section area a by maximizing the amount of less expensive material used, while guaranteeing it can still bear a prescribed load F . We give an illustration with 6 elements:

$$\begin{aligned} & \text{maximize} && x_2 + x_4 + x_6 \\ & \text{subject to} && \sum_{i=1}^6 x_i = L, \\ & && \sum_{i=1}^6 \frac{x_i}{Y_i} < \frac{Da}{FL}, \end{aligned}$$



where x is a vector of component lengths in $\mathcal{X} := (\mathbb{R}_{>0})^6$, Y is an uncertain vector of Young's moduli taking values in $\mathcal{Y} \subset (\mathbb{R}_{>0})^6$, and D is a limit imposed on the beam's elongation under maximum design load F . The less expensive material is shown in light gray.

Graph theory – the shortest path problem [see, e.g., 18]: find the shortest route between the source and destination nodes s and d in a weighted directed graph—with a set of nodes V and an adjacency matrix A —in which there is uncertainty about the weight attached to some of the edges. Formally,

$$\begin{aligned} & \text{minimize} && \sum_{(u,v) \in V^2} Y_{uv} x_{uv} \\ & \text{subject to} && x_{uv} A_{uv} = x_{uv} \text{ for all } (u,v) \text{ in } V^2, \\ & && \sum_{v \in V} x_{wv} - \sum_{u \in V} x_{uw} = \delta_{ws} - \delta_{dw} \text{ for all } w \text{ in } V, \end{aligned}$$

where the optimization variable x belongs to the set of adjacency matrices $\{0, 1\}^{n \times n}$ and Y is an uncertain matrix of edge weights taking values in $\mathcal{Y} := (\mathbb{R}_{\geq 0})^{n \times n}$. The uncertainty appears in the goal function; we can remedy this by introducing an auxiliary optimization variable component x_0 such that $\mathcal{X} := \mathbb{R}_{\leq 0} \times \{0, 1\}^{n \times n}$, leading to an equivalent reformulation in the standard form (cf. Section 2.1):

$$\begin{aligned} & \text{maximize} && x_0 \\ & \text{subject to} && x_0 \leq -\sum_{(u,v) \in V^2} Y_{uv} x_{uv} \text{ and the constraints listed above.} \end{aligned}$$

8. CONCLUSIONS

In this paper, we have looked at a specific, but nevertheless very general class of constrained optimization problems and investigated how uncertainty about parameters appearing in the constraints can be dealt with. We have shown that the way forward consists in reformulating the—due to this uncertainty—non-well-posed problem into a well-posed decision problem. We have then restricted our attention to forms of uncertainty that can be represented using coherent lower and upper previsions and presented the general form of the resulting decision problems, one for each of two typical optimality criteria—maximality and maximality—used in conjunction with such previsions.

Once these two decision problem templates were in place, the question became for which specific types of coherent lower and upper previsions we could solve them to a sufficient degree so as to obtain a constrained optimization problem without uncertainty. We showed that this is generally the case for linear previsions and vacuous previsions. It became clear that using maximinity results in less complicated mathematical problems as compared to the maximality criterion. This observation is given more weight by the fact that, when modeling uncertainty using a possibility distribution, it was straightforward to obtain a general result for maximinity but not for maximality.

We encountered similar differences in complexity in preliminary investigations of other popular types of coherent lower previsions, such as those associated with p-boxes [17], and those that are independent products [25] of vacuous and linear previsions as uncertainty models. However, the maximinity criterion can also result in quite complicated expressions, as we encountered when looking at linear-vacuous [25] previsions. Ongoing work [9, 19] consists in investigating whether general results can be found for these and other uncertainty models when we restrict the generality of the initial constrained optimization problem, e.g., to linear programming problems, and testing them on numerical examples.

ACKNOWLEDGEMENTS

During part of his research on this topic, Erik Quaeghebeur was supported by a Fellowship of the Belgian American Educational Foundation. This research was supported by the IWT SBO project 60043, “Fuzzy Finite Element Method”. We wish to thank the area editor for extensive and useful comments, suggestions, and pointers to the related literature and the reviewers for pointing out where the exposition could be clarified. We also wish to thank Nathan Huntley for alerting us to the possibility of interpretational issues arising when moving uncertainty from the objective function to the constraints.

REFERENCES

- [1] R. E. Bellman and L. A. Zadeh. Decision-making in a fuzzy environment. *Management Science*, 17(4):B-141–B-164, 1970. doi:10.1287/mnsc.17.4.B141.
- [2] Ahoren Ben-Tal and Arkadi Nemirovski. Robust optimization – methodology and applications. *Mathematical Programming*, 92(3):453–480, 2002. doi:10.1007/s101070100286.
- [3] Gert de Cooman. Integration and conditioning in numerical possibility theory. *Annals of Mathematics and Artificial Intelligence*, 32(1):87–123, 2001.
- [4] Gert de Cooman. A behavioural model for vague probability assessments. *Fuzzy Sets and Systems*, 154:305–358, 2005. doi:10.1016/j.fss.2005.01.005. With discussion.
- [5] Gert de Cooman and Dirk Aeyels. Supremum preserving upper probabilities. *Information Sciences*, 118:173–212, 1999.
- [6] Bruno de Finetti. *Theory of Probability*. Wiley, 1974–1975. 2 Volumes.
- [7] Didier Dubois and Henri Prade. *Théorie des possibilités*. Masson, 2nd edition, 1988.
- [8] M. Fiedler, Jiri Nedoma, Jaroslav Ramík, Jiri Rohn, and Karel Zimmermann. *Linear Optimization Problems with Inexact Data*. Springer, 2006.
- [9] Nathan Huntley, Rolando Quiñones, Keivan Shariatmadar, Erik Quaeghebeur, Gert de Cooman, and Etienne Kerre. Implementation of maximin and maximal solutions for linear programming problems under uncertainty. Submitted to FLINS 2012.

- [10] Masahiro Inuiguchi and Jaroslav Ramík. Possibilistic linear programming: a brief review of fuzzy mathematical programming and a comparison with stochastic programming in portfolio selection problem. *Fuzzy Sets and Systems*, 111(1):3–28, 2000. doi:10.1016/S0165-0114(98)00449-7.
- [11] K. David Jamison and Weldon A. Lodwick. Fuzzy linear programming using a penalty method. *Fuzzy Sets and Systems*, 119(1):97–110, 2001. doi:10.1016/S0165-0114(99)00082-2.
- [12] Uzay Kaymak and João M. C. Sousa. Weighting of constraints in fuzzy optimisation. *Constraints*, 8:61–78, 2003.
- [13] Weldon Lodwick and Katherine Bachman. Solving large-scale fuzzy and possibilistic optimization problems. *Fuzzy Optimization and Decision Making*, 4:257–278, 2005. doi:10.1007/s10700-005-3663-4.
- [14] Daryl L. Logan. *A first course in the finite element method*. Chris Carson, 2007.
- [15] Enrique Miranda. A survey of the theory of coherent lower previsions. *International Journal of Approximate Reasoning*, 48:628–658, 2008.
- [16] Enrique Miranda and Gert de Cooman. Epistemic independence in numerical possibility theory. *International Journal of Approximate Reasoning*, 32:23–42, 2003. doi:10.1016/S0888-613X(02)00087-7.
- [17] Enrique Miranda, Matthias C. M. Troffaes, and Sébastien Destercke. Generalised p-boxes on totally ordered spaces. *International Journal of Advances in Soft Computing*, pages 235–242, 6 2008. doi:10.1007/978-3-540-85027-4_29.
- [18] Roberto Montemanni and Luca Maria Gambardella. The robust shortest path problem with interval data via Benders decomposition. *4OR: A Quarterly Journal of Operations Research*, 3(4):315–328, 2005. doi:10.1007/s10288-005-0066-x.
- [19] Erik Quaeghebeur, Nathan Huntley, Keivan Shariatmadar, and Gert de Cooman. Maximin and maximal solutions for linear programming problems with possibilistic uncertainty. Submitted to IPMU 2012.
- [20] Heinrich Rommelfanger. Fuzzy linear programming and applications. *European Journal of Operational Research*, 92(3):512–527, 1996. doi:10.1016/0377-2217(95)00008-9.
- [21] A. L. Soyster. Convex programming with set-inclusive constraints and applications to inexact linear programming. *Operations Research*, 21(5):1154–1157, 9–10 1973. URL <http://www.jstor.org/stable/168933>.
- [22] Matthias C. M. Troffaes. Decision making under uncertainty using imprecise probabilities. *International Journal of Approximate Reasoning*, 45(1):17–29, 5 2007. doi:10.1016/j.ijar.2006.06.001.
- [23] Matthias C. M. Troffaes and Gert de Cooman. Extension of coherent lower previsions to unbounded random variables. In *Proceedings of IPMU 2002*, pages 735–742. 2002.
- [24] Matthias C. M. Troffaes and Gert de Cooman. Lower previsions for unbounded random variables. In P. Grzegorzewski, O. Hryniewicz, and M. Ángeles Gil, editors, *Soft Methods in Probability, Statistics and Data Analysis*, pages 146–155. Physica-Verlag, New York, 2002.
- [25] Peter Walley. *Statistical Reasoning with Imprecise Probabilities*. Chapman and Hall, London, 1991.
- [26] Peter Walley and Gert de Cooman. Coherence of rules for defining conditional possibility. *International Journal of Approximate Reasoning*, 21:63–107, 1999. doi:10.1016/S0888-613X(99)00007-9.

- [27] Peter Walley and Gert de Cooman. A behavioral model for linguistic uncertainty. *Information Sciences*, 134(1-4):1–37, 2001. doi:10.1016/S0020-0255(01)00090-1.
- [28] Peter Whittle. *Probability via Expectation*. Springer, 3rd edition, 1992.
- [29] Hans-Jürgen Zimmermann. Description and optimization of fuzzy systems. *International Journal of General Systems*, 2:209–215, 1975. doi:10.1080/03081077608547470.

E-mail address, Erik Quaeghebeur: Erik.Quaeghebeur@UGent.be
URL, Erik Quaeghebeur: <http://users.UGent.be/~equaeghe>

E-mail address, Keivan Shariatmadar: Keivan.Shariatmadar@UGent.be
URL, Keivan Shariatmadar: <http://users.ugent.be/~kshariat>

E-mail address, Gert de Cooman: Gert.deCooman@UGent.be
URL, Gert de Cooman: <http://users.UGent.be/~gdcooma>

(Keivan Shariatmadar, Erik Quaeghebeur & Gert de Cooman) GHENT UNIVERSITY, SYSTEMS RESEARCH GROUP, EESA DEPARTMENT, FACULTY OF ENGINEERING, TECHNOLOGIEPARK-ZWIJNAARDE 914, 9052 ZWIJNAARDE, BELGIUM