

Consistent Inverse Probability and Possibility Propagation

Dominik Hose and Michael Hanss

Institute of Engineering and Computational Mechanics, University of Stuttgart,
Pfaffenwaldring 9, 70569 Stuttgart, Germany,
{dominik.hose, michael.hanss}@itm.uni-stuttgart.de

Abstract

Given a probability distribution of an output quantity of a model, it is generally not possible to infer a unique probability distribution of the uncertain input quantity. In this contribution, it is shown that by reverting back to the coarser framework of possibility theory this problem possesses a conceptually straightforward solution with some powerful properties in the view of imprecise probability descriptions.

Keywords: Inverse Problems, Probability-Possibility Consistency, Imprecise Probabilities, Uncertainty Propagation, Possibility Theory.

Prologue

A classical example for the shortcomings of probability theory is the paradox of partial ignorance which has been employed by many scholars, such as Thierry Denœux or Didier Dubois:

Suppose, you have the same knowledge about a certain positive variable $x > 0$, e.g. the proportion of alcohol in a liquid, and about its reciprocal $y = \frac{1}{x}$. Then, $x \in [a, b]$ is equivalent to $y \in [\frac{1}{b}, \frac{1}{a}]$, but a uniform distribution on $[a, b]$ is not compatible with a uniform distribution on $[\frac{1}{b}, \frac{1}{a}]$. Therefore, uniform probability distributions cannot represent total ignorance since invariance under transformations cannot be guaranteed.¹

The conclusion, thus, has to be that the Principle of Maximum Entropy [11], leading to the uniform distributions, is not universally applicable.

¹Didier Dubois at the 8th International Workshop on Reliable Engineering Computing 2018, <https://stream.liv.ac.uk/s/cskc5b62> (at 8 min 40 sec)

1 Introduction

In various disciplines, e.g. in statistical mechanics [11], it is necessary to infer an input (a-priori) distribution of an unknown quantity, e.g. the position or velocity distribution of a particle swarm, from the a-posteriori distribution of a measurable output quantity, e.g. the energy. These inverse problems generally possess an infinite number of solutions, and regularizing assumptions, such as maximum entropy of the input distribution, are needed in order to find a unique solution.

In this contribution, an alternative to the Principle of Maximum Entropy in inverse problems is provided.

The solution of fuzzy equations [5], where the parameters are of possibilistic rather than probabilistic nature, has received attention in fuzzy linear systems [15, 20] and for the solution of fuzzy-valued differential equations [18]. Furthermore, Mauris presents a possibilistic moment matching technique in [16] employing arguments of minimum specificity which are also considered in this contribution. Yet, a general theory for this type of problem remains to be established.

The paper is organized as follows: Section 2 contains the mathematical preliminaries about probability and possibility theory required to understand the main results. Section 3 is divided into three parts. The first part contains known results about consistency in the forward uncertainty propagation within the context of imprecise probabilities. The second part then derives similar results in the inverse propagation case, where Theorem 2 constitutes the core of this contribution. The proofs of Theorems 1 and 2 are given in the last part which may be skipped for improved readability. Finally, in Section 4, some concluding remarks are given.

Illustrative examples are provided throughout the text in order to clarify the implications of the presented results.

2 Preliminaries

The following section provides the necessary definitions to understand the main results. It follows standard notation, e.g. as in [19] for probability theory, or in [14] for possibility theory.

2.1 Probability and Possibility Spaces

Consider a set Ω , the universal set, and a σ -algebra Σ on Ω . The tuple (Ω, Σ) is called a measurable space. A probability measure on the measurable space (Ω, Σ) is a function $P : \Sigma \rightarrow [0, 1]$ satisfying

1. $P(\emptyset) = 0$,
2. $P(\Omega) = 1$ and
3. if $\{U_n \in \Sigma\}_{n \in \mathbb{N}}$ is a countable collection of pairwise disjoint events with union U , then $P(U) = \sum_{n \in \mathbb{N}} P(U_n)$.

The set of all probability measures on (Ω, Σ) is denoted by $\mathbb{M}^p(\Omega, \Sigma)$.

A possibility measure on Ω is a function $\Pi : 2^\Omega \rightarrow [0, 1]$ with

1. $\Pi(\emptyset) = 0$,
2. $\Pi(\Omega) = 1$ and
3. if $\{U_n \subseteq \Omega\}_{n \in \mathbb{N}}$ is a countable collection of pairwise disjoint subsets with union U , then $\Pi(U) = \sup_{n \in \mathbb{N}} \Pi(U_n)$.

The set of all possibility measures on Ω is denoted by $\mathbb{M}^\pi(\Omega)$. The corresponding necessity measure N is given by

$$N(U) = 1 - \Pi(\Omega \setminus U) \quad \forall U \subseteq \Omega. \quad (1)$$

2.2 Possibilities as Upper Probabilities

As stated in [12], the Dempster-Shafer Theory of Evidence [3] can serve as a general framework for descriptions of uncertainty, unifying probability theory, possibility theory, p-boxes, etc. [4].

An important consequence is that possibilities can serve as upper probabilities, providing a coarse description of incomplete probabilistic knowledge. In this context, the concept of *consistency* is of fundamental importance. Here, the definition by Dubois and Prade [7], viewing a possibility measure as an upper probability measure, is employed. More precisely, a probability measure $P \in \mathbb{M}^p(\Omega, \Sigma)$ and a possibility

measure $\Pi \in \mathbb{M}^\pi(\Omega)$ are called consistent if the probability of an event $U \in \Sigma$ is dominated by its possibility

$$P(U) \leq \Pi(U). \quad (2)$$

From the upper bound provided by the possibility measure, it follows from $P(\Omega \setminus U) \leq \Pi(\Omega \setminus U)$ that the probability measure is also bounded from below by the necessity measure, since

$$N(U) = 1 - \Pi(\Omega \setminus U) \leq 1 - P(\Omega \setminus U) = P(U). \quad (3)$$

Consequently, any possibility measure induces a credal set of consistent probability distributions

$$\mathcal{P}_\Pi(\Omega, \Sigma) = \{P \in \mathbb{M}^p(\Omega, \Sigma) : P(U) \leq \Pi(U) \quad \forall U \in \Sigma\}. \quad (4)$$

For practical methods for the construction of possibility distributions, e.g. from probability distributions, refer e.g. to [8]. Illustrative results about which probability distributions are actually contained in \mathcal{P}_Π may be found in [2].

2.3 Specificity

The specificity is a useful tool for analyzing a possibility measure as it allows to assess the uncertainty contained in the corresponding credal sets.

Given two possibility measures $\Pi', \Pi'' \in \mathbb{M}^\pi(\Omega)$, Π' is called more specific than Π'' , denoted by $\Pi' \preceq \Pi''$, if

$$\Pi'(U) \leq \Pi''(U) \quad \forall U \subseteq \Omega. \quad (5)$$

The following lemma by Dubois and Prade [7] illustrates how a comparison of the specificity of two possibility distributions allows for the comparison of their induced credal sets.

Lemma 1. *Let (Ω, Σ) be a measurable space. Given two possibility measures $\Pi', \Pi'' \in \mathbb{M}^\pi(\Omega)$, if $\Pi' \preceq \Pi''$, then their induced credal sets satisfy*

$$\mathcal{P}_{\Pi'}(\Omega, \Sigma) \subseteq \mathcal{P}_{\Pi''}(\Omega, \Sigma). \quad (6)$$

Proof. Let $P \in \mathcal{P}_{\Pi'}(\Omega, \Sigma)$. Then $P(U) \leq \Pi'(U) \leq \Pi''(U)$ for all $U \in \Sigma$ and therefore $P \in \mathcal{P}_{\Pi''}(\Omega, \Sigma)$. \square

Thus, specificity can be interpreted as a measure for the imprecision in the credal set as motivated by the following two examples.

Example 1 (Minimum Specific Possibility Measure). Let (Ω, Σ) be a measurable space. The uniform possibility measure $\Pi^\vee \in \mathbb{M}^\pi(\Omega)$ defined by $\Pi^\vee(U) = 1$ for all $U \in \Sigma$ is the least specific possibility measure. It is consistent with all probability measures on (Ω, Σ) , i.e. $\mathcal{P}_{\Pi^\vee}(\Omega, \Sigma) = \mathbb{M}^p(\Omega, \Sigma)$, and therefore truly represents total ignorance.

Example 2 (Maximum Specific Possibility Measure). Let (Ω, Σ) be a measurable space. The singleton possibility measure $\Pi^\omega \in \mathbb{M}^\pi(\Omega)$ for a fixed $\omega \in \Omega$ with $\{\omega\} \in \Sigma$ is given by

$$\Pi^\omega(U) = \begin{cases} 1 & \text{if } \omega \in U, \\ 0 & \text{else} \end{cases} \quad \forall U \subseteq \Omega. \quad (7)$$

It is maximally specific as its induced credal set consists only of the probability measure that assigns all probability to the singleton ω , i.e. $\mathcal{P}_{\Pi^\omega}(\Omega, \Sigma) = \{P^\omega\}$ where $P^\omega = \Pi^\omega$. Both measures describe that the elementary event ω is absolutely certain.

Especially Dubois and Prade strongly advocate the Principle of *Minimum Specificity* [8], which states that without further knowledge one ought to choose the least specific possibility distribution that describes a state of information. This will be of importance below in the definition of the minimum specific inverse possibility distribution.

2.4 Uncertain Variables

Let (Ω, Σ) and (Ω_X, Σ_X) be measurable spaces. The function $X : \Omega \rightarrow \Omega_X$ is called a Σ - Σ_X -measurable function if its induced σ -algebra on Ω given by

$$\sigma(X, \Sigma_X) = \sigma(\{X^{-1}(U_X) \mid U_X \in \Sigma_X\}) \quad (8)$$

satisfies $\sigma(X, \Sigma_X) \subseteq \Sigma$. In this contribution, X will generally be called an (Ω_X, Σ_X) -valued uncertain variable. The pushforward of $P \in \mathbb{M}^p(\Omega, \Sigma)$ by X is a probability measure

$$P_X : \Sigma_X \rightarrow [0, 1], \quad (9)$$

$$U_X \mapsto P(X^{-1}(U_X))$$

on (Ω_X, Σ_X) , which is also referred to as the probability distribution of X . Similarly, the pushforward of $\Pi \in \mathbb{M}^\pi(\Omega_X)$ by X is a possibility measure

$$\Pi_X : 2^{\Omega_X} \rightarrow [0, 1], \quad (10)$$

$$U_X \mapsto \Pi(X^{-1}(U_X))$$

on (Ω_X) called the the possibility distribution of X .

Lemma 2 (Pushforward Consistency). *Let (Ω, Σ) and (Ω_X, Σ_X) be measurable spaces and let $X : \Omega \rightarrow \Omega_X$ be an (Ω_X, Σ_X) -valued uncertain variable. If $\Pi \in \mathbb{M}^\pi(\Omega)$ and $P \in \mathcal{P}_\Pi$, then their pushforwards by X also satisfy $P_X \in \mathcal{P}_{\Pi_X}$.*

Proof. Let $U_X \in \Sigma_X$. Since X is a Σ - Σ_X -measurable function, the pre-image $U = X^{-1}(U_X)$ is measurable. Immediately, it follows that

$$P_X(U_X) = P(U) \leq \Pi(U) = \Pi_X(U_X) \quad (11)$$

yielding consistency. \square

A more general version of Lemma 2 may be found in [6].

Below, the existence of the underlying measurable space (Ω, Σ) will always be assumed without explicitly mentioning it. Furthermore, only uncertain variables $\mathbf{X} : \Omega \rightarrow \mathbb{R}^M$ on the Borel space $(\mathbb{R}^M, \mathcal{B}(\mathbb{R}^M))$, where the Borel σ -algebra $\mathcal{B}(\mathbb{R}^M)$ is the σ -algebra generated by all open sets in \mathbb{R}^M , will be considered, i.e. \mathbf{X} is a Σ - $\mathcal{B}(\mathbb{R}^M)$ -measurable function. All of the above is implied when defining \mathbf{X} to be an \mathbb{R}^M -valued uncertain variable with probability distribution $P_{\mathbf{X}} \in \mathbb{M}^p(\mathbb{R}^M, \mathcal{B}(\mathbb{R}^M))$ or possibility distribution $\Pi_{\mathbf{X}} \in \mathbb{M}^\pi(\mathbb{R}^M)$.

A probability density function $p_{\mathbf{X}} : \mathbb{R}^M \rightarrow \mathbb{R}_0^+$ of \mathbf{X} is any Borel measurable function satisfying

$$P_{\mathbf{X}}(U_X) = \int_{U_X} p_{\mathbf{X}} d\lambda \quad \forall U_X \in \mathcal{B}(\mathbb{R}^M) \quad (12)$$

where λ is the Lebesgue measure, and a possibility density function $\pi_{\mathbf{X}} : \mathbb{R}^M \rightarrow [0, 1]$ of \mathbf{X} is any function satisfying

$$\Pi_{\mathbf{X}}(U_X) = \sup \pi_{\mathbf{X}}(U_X) \quad \forall U_X \subseteq \mathbb{R}^M. \quad (13)$$

The possibility density function can also be referred to as the fuzzy set membership function from classical fuzzy set theory and is given by

$$\pi_{\mathbf{X}}(\mathbf{x}) = \sup \pi_{\mathbf{X}}(\{\mathbf{x}\}) = \Pi_{\mathbf{X}}(\{\mathbf{x}\}) \quad \forall \mathbf{x} \in \mathbb{R}^M. \quad (14)$$

Furthermore, Eqs. (12) and (13) can also be interpreted in the sense that a valid density function induces the corresponding distribution.

From the definition of consistency, it follows that the cumulative probability distribution function of $P_{\mathbf{X}}$ is necessarily bounded by the cumulative possibility and necessity distribution functions of $\Pi_{\mathbf{X}}$. Notice, however, that the boundedness of the cumulative distribution functions does not necessarily imply consistency in the general case.

Example 3 (Consistent Probability and Possibility Distributions on X). Let X be an \mathbb{R} -valued uncertain variable. The probability density function p_X (a normalized combination of two normal distributions with means 0.33 and 0.00 and standard deviations 0.17 and 0.12, respectively) shown in Fig. 1a induces a probability distribution P_X on X .

The possibility density function π_X shown in Fig. 1b is the optimal transformation [8] of p_X and its induced possibility distribution on X is – by construction – consistent with P_X .

Consequently, the cumulative probability distribution function $P_X((-\infty, x])$ is bounded from above

by the cumulative possibility distribution function $\Pi_X((-\infty, x])$ and from below by the cumulative necessity distribution function $N_X((-\infty, x])$, all shown in Fig. 1c.

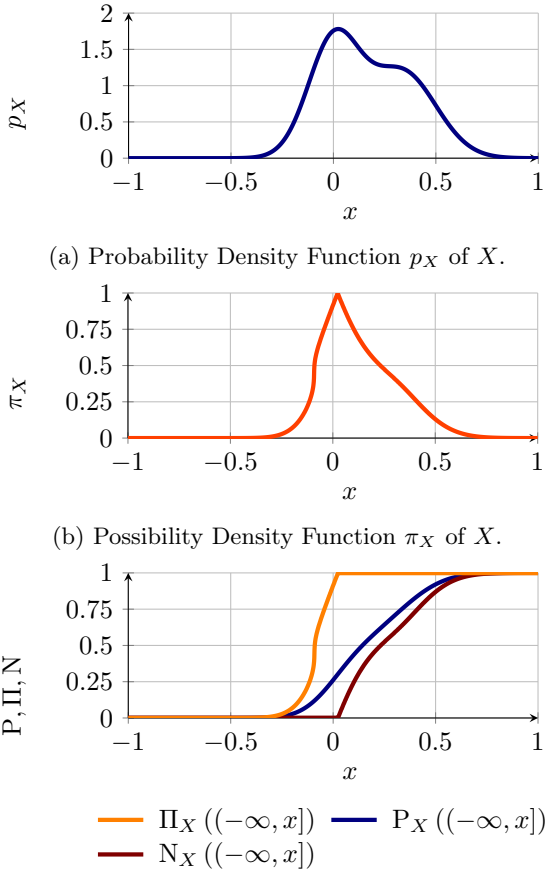


Figure 1: Distributions in Ex. 3.

Finally, two possibility density functions can be compared in terms of their specificity by the following lemma.

Lemma 3 (Possibility Density Specificity). *Let \mathbf{X} be an \mathbb{R}^M -valued uncertain variable with possibility distributions Π'_X and Π''_X . If their possibility density functions satisfy*

$$\pi_{\mathbf{X}_1}(\mathbf{x}) \leq \pi_{\mathbf{X}_2}(\mathbf{x}) \quad \forall \mathbf{x} \in \mathbb{R}^M \quad (15)$$

then $\Pi_{\mathbf{X}_1} \preceq \Pi_{\mathbf{X}_2}$.

Proof. Let $U_X \subseteq \mathbb{R}^M$. Then $\Pi_{\mathbf{X}_1}(U_X) = \sup \pi_{\mathbf{X}_1}(U_X) \leq \sup \pi_{\mathbf{X}_2}(U_X) = \Pi_{\mathbf{X}_2}(U_X)$. \square

3 Results

This section deals with the propagation of probabilistic and possibilistic uncertainties through a model [1]. In

particular, given the \mathbb{R}^M -valued uncertain variable \mathbf{X} and the $\mathcal{B}(\mathbb{R}^M)$ - $\mathcal{B}(\mathbb{R}^N)$ -measurable function $\phi: \mathbb{R}^M \rightarrow \mathbb{R}^N$ (the model), then their concatenation $\mathbf{Y} = \phi \circ \mathbf{X}$, which is written $\mathbf{Y} = \phi(\mathbf{X})$ for readability, is an \mathbb{R}^N -valued uncertain variable, too, since the concatenation of two measurable functions is also measurable.

3.1 Forward Uncertainty Propagation

If \mathbf{X} possesses the probability distribution $P_{\mathbf{X}} \in \mathbb{M}^p(\mathbb{R}^M, \mathcal{B}(\mathbb{R}^M))$, then ϕ induces the probability distribution

$$P_{\mathbf{Y}}(U_Y) = P_{\mathbf{X}}(\phi^{-1}(U_Y)) \quad \forall U_Y \in \mathcal{B}(\mathbb{R}^N) \quad (16)$$

on \mathbf{Y} , the pushforward of $P_{\mathbf{X}}$ by ϕ . In particular, if $p_{\mathbf{X}}$ and $p_{\mathbf{Y}}$ are their respective probability density functions, then it holds that

$$\int_{U_Y} p_{\mathbf{Y}} d\lambda = \int_{\phi^{-1}(U_Y)} p_{\mathbf{X}} d\lambda \quad \forall U_Y \in \mathcal{B}(\mathbb{R}^N). \quad (17)$$

In the same manner, the pushforward of $\Pi_{\mathbf{X}} \in \mathbb{M}^\pi(\mathbb{R}^M)$ by ϕ is given by

$$\Pi_{\mathbf{Y}}(U_Y) = \Pi_{\mathbf{X}}(\phi^{-1}(U_Y)) \quad \forall U_Y \in \mathcal{B}(\mathbb{R}^N). \quad (18)$$

On an elementary basis, if \mathbf{X} and \mathbf{Y} possess the possibility density functions $\pi_{\mathbf{X}}$ and $\pi_{\mathbf{Y}}$, then it follows from $U_Y = \{\mathbf{y}\}$ that

$$\pi_{\mathbf{Y}}(\mathbf{y}) = \sup \pi_{\mathbf{X}}(\phi^{-1}(\{\mathbf{y}\})) \quad \forall \mathbf{y} \in \mathbb{R}^N \quad (19)$$

where the supremum of the empty set is zero. This expression is known as the extension principle originally formulated by Zadeh. For further reading regarding forward possibility propagation, i.e. fuzzy arithmetic, refer to [9].

The following lemma states that the forward propagation of consistent possibility and probability distributions yields consistent possibility and probability distributions.

Lemma 4 (Forward Uncertainty Propagation Consistency). *Given an \mathbb{R}^M -valued uncertain variable \mathbf{X} with possibility distribution $\Pi_{\mathbf{X}}$ and probability distribution $P_{\mathbf{X}} \in \mathcal{P}_{\Pi_{\mathbf{X}}}$ and $\mathbf{Y} = \phi(\mathbf{X})$ for a Borel measurable function $\phi: \mathbb{R}^M \rightarrow \mathbb{R}^N$, then $P_{\mathbf{Y}} \in \mathcal{P}_{\Pi_{\mathbf{Y}}}$.*

Proof. The proof is similar to the one of Lemma 2. Let $P_{\mathbf{X}} \in \mathcal{P}_{\Pi_{\mathbf{X}}}$ and $U_Y \in \mathcal{B}(\mathbb{R}^N)$. Since ϕ is a Borel measurable function, the pre-image $U_X = \phi^{-1}(U_Y)$ is measurable. Immediately, it follows that

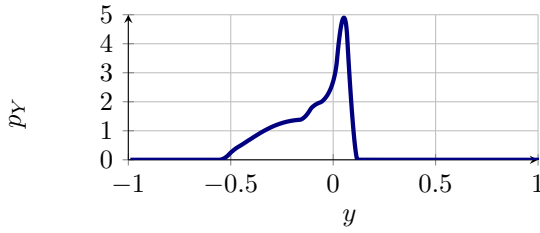
$$\begin{aligned} P_{\mathbf{Y}}(U_Y) &= P_{\mathbf{X}}(\phi^{-1}(U_Y)) \\ &\leq \Pi_{\mathbf{X}}(\phi^{-1}(U_Y)) = \Pi_{\mathbf{Y}}(U_Y) \end{aligned} \quad (20)$$

yielding consistency. \square

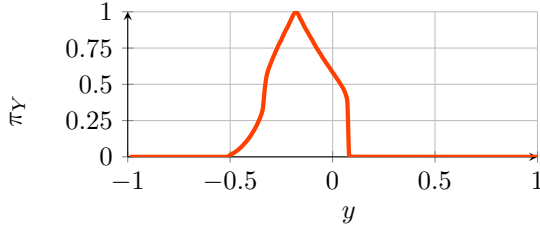
The intent of this contribution is to show comparable results in a setting of inverse uncertainty propagation.

Example 4 (Consistent Probability and Possibility Distributions on Y). The function $\phi : x \mapsto 4.20x^5 + 0.34x^4 - 4.51x^3 - 0.12x^2 + 1.31x^1 - 0.21$ is measurable and produces an \mathbb{R} -valued uncertain variable $Y = \phi(X)$, where X is the uncertain variable with probability and possibility distributions given in Ex. 3. The pushforward probability distribution P_Y of Y is induced by the probability density function p_Y in Fig. 2a.

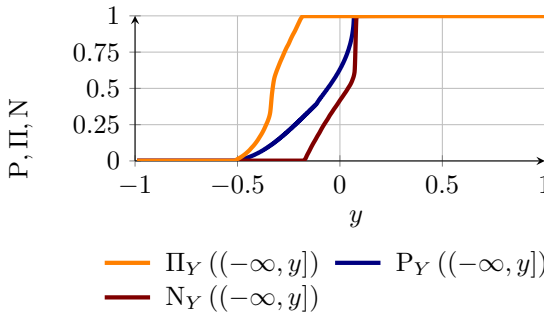
The pushforward possibility distribution Π_Y of Y is induced by the possibility density function π_Y in Fig. 2b. From Lemma 2 it follows immediately that P_Y and Π_Y are consistent, too. Consequently, the cumulative probability distribution function $P_Y((-\infty, y])$ is bounded by the cumulative possibility distribution function $\Pi_Y((-\infty, y])$ and by the cumulative necessity distribution function $N_Y((-\infty, y])$, all shown in Fig. 2c.



(a) Pushforward Probability Density Function p_Y of Y .



(b) Pushforward Possibility Density Function π_Y of Y .



(c) Cumulative Distribution Functions of Y .

Figure 2: Distributions in Ex. 4.

3.2 Inverse Uncertainty Propagation

Let \mathbf{X} be an \mathbb{R}^M -valued uncertain variable, $\phi : \mathbb{R}^M \rightarrow \mathbb{R}^N$ a Borel measurable function and $\mathbf{Y} = \phi(\mathbf{X})$ an \mathbb{R}^N -valued uncertain variable. Now, contrary to the previous section, suppose that only the probability distribution P_Y of \mathbf{Y} is given. Then, all admissible probability distributions of \mathbf{X} that yield P_Y as their pushforward by ϕ have to satisfy Eq. (16). They are gathered in the credal set

$$\begin{aligned} \mathcal{I}_{P_Y}^\phi &= \{P'_{\phi^{-1}(\mathbf{Y})} \in \mathbb{M}^p(\mathbb{R}^M, \mathcal{B}(\mathbb{R}^M)) : \\ P_Y(U_Y) &= P'_{\phi^{-1}(\mathbf{Y})}(\phi^{-1}(U_Y)) \\ &\forall U_Y \in \mathcal{B}(\mathbb{R}^N)\}. \end{aligned} \quad (21)$$

The question which of all those distributions to choose cannot be answered unanimously without further regularizing assumptions. Yet, prior knowledge for the application of Bayes' Theorem may not always be available and the Principle of Maximum Entropy [11] which states that among all admissible solutions $P'_{\phi^{-1}(\mathbf{Y})} \in \mathcal{I}_{P_Y}^\phi$ one ought to choose the one induced by the probability density function $p_{\phi^{-1}(\mathbf{Y})}$ that maximizes the entropy is deficient as explained above. A different, yet equally arbitrary, choice is the pignistic transform [17] of $\mathcal{I}_{P_Y}^\phi$, i.e. its center of gravity. For further reading, refer e.g. to [12].

In this contribution, it will be shown that it is not necessary to – somewhat randomly – settle on one particular probability distribution if one is willing to revert to the coarser framework of possibility theory which allows one to represent all of them by just one possibility distribution.

In order to do so, it is necessary to define the set of admissible inverse possibility distributions for \mathbf{X} if only the possibility distribution Π_Y of $\mathbf{Y} = \phi(\mathbf{X})$ is given. Similarly to Eq. 21, the only requirement for its members is to satisfy Eq. (18), i.e. it is given by

$$\begin{aligned} \mathcal{I}_{\Pi_Y}^\phi &= \{\Pi'_{\phi^{-1}(\mathbf{Y})} \in \mathbb{M}^\pi(\mathbb{R}^M) : \\ \Pi_Y(U_Y) &= \Pi'_{\phi^{-1}(\mathbf{Y})}(\phi^{-1}(U_Y)) \\ &\forall U_Y \subseteq \mathbb{R}^M\}. \end{aligned} \quad (22)$$

Applying the Principle of Minimum Specificity, the least specific possibility distribution in $\mathcal{I}_{\Pi_Y}^\phi$ is found by the following theorem.

Theorem 1 (Minimum Specific Inverse Possibility Distribution). *Let \mathbf{X} be an \mathbb{R}^M -valued uncertain variable, $\phi : \mathbb{R}^M \rightarrow \mathbb{R}^N$ a Borel measurable and surjective function and $\mathbf{Y} = \phi(\mathbf{X})$ an \mathbb{R}^N -valued uncertain variable with a given possibility distribution Π_Y . Furthermore, define*

$$\pi_{\phi^{-1}(\mathbf{Y})}^*(\mathbf{x}) = \pi_Y(\phi(\mathbf{x})) \quad \forall \mathbf{x} \in \mathbb{R}^M. \quad (23)$$

Then,

- $\pi_{\phi^{-1}(\mathbf{Y})}^*(\mathbf{x})$ is a possibility density function which induces a possibility distribution $\Pi_{\phi^{-1}(\mathbf{Y})}^*$,
- $\Pi_{\phi^{-1}(\mathbf{Y})}^*$ is an admissible inverse possibility distribution, i.e. $\Pi_{\phi^{-1}(\mathbf{Y})}^* \in \mathcal{I}_{\Pi_{\mathbf{Y}}}^\phi$ and
- $\Pi_{\phi^{-1}(\mathbf{Y})}^*$ is less specific than all other admissible inverse possibility distributions, i.e. for all $\Pi'_{\phi^{-1}(\mathbf{Y})} \in \mathcal{I}_{\Pi_{\mathbf{Y}}}^\phi$ it holds that $\Pi'_{\phi^{-1}(\mathbf{Y})} \preceq \Pi_{\phi^{-1}(\mathbf{Y})}^*$.

Proof. See Section 3.3.1. \square

Therefore, $\Pi_{\phi^{-1}(\mathbf{Y})}^*$ may be called the Minimum Specific Inverse Possibility Distribution, which was first proposed in [10] where it was termed 'Most General Inverse Solution'.

Theorem 1 requires ϕ to be surjective. However, this is merely a matter of further restricting $(\mathbb{R}^N, \mathcal{B}(\mathbb{R}^N))$. The authors assume that the presented results should hold up for any combination of topological space and their Borel σ -algebras. If surjectivity cannot be guaranteed, $\Pi_{\phi^{-1}(\mathbf{Y})}^*$ may degenerate to a subnormalized possibility distribution indicating incompatibility between the output distributions and the model, see [10].

The following theorem is the main result of this contribution. It states that the set of all admissible probability distributions is actually contained in the credal set of the minimum specific inverse possibility distribution.

Theorem 2 (Inverse Uncertainty Propagation Consistency). *Let \mathbf{X} be an \mathbb{R}^M -valued uncertain variable, $\phi : \mathbb{R}^M \rightarrow \mathbb{R}^N$ a Borel measurable and surjective function and $\mathbf{Y} = \phi(\mathbf{X})$ an \mathbb{R}^N -valued uncertain variable with a given probability distribution $\mathbf{P}_{\mathbf{Y}} \in \mathbb{MP}(\mathbb{R}^N, \mathcal{B}(\mathbb{R}^N))$ and a given possibility distribution $\Pi_{\mathbf{Y}} \in \mathbb{M}^\pi(\mathbb{R}^M)$ induced by a Borel measurable possibility density function $\pi_{\mathbf{Y}}$. Then, if $\mathbf{P}_{\mathbf{Y}} \in \mathcal{P}_{\Pi_{\mathbf{Y}}}$, it follows that $\mathcal{I}_{\mathbf{P}_{\mathbf{Y}}}^\phi \subseteq \mathcal{P}_{\Pi_{\phi^{-1}(\mathbf{Y})}^*}$.*

Proof. See Section 3.3.2. \square

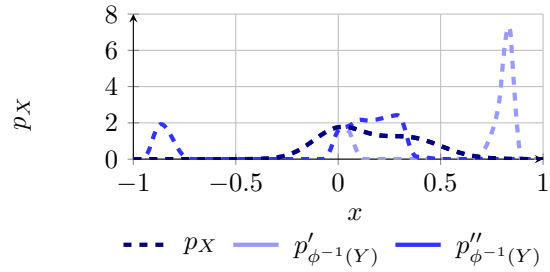
The Borel measurability of $\pi_{\mathbf{Y}}$ is not a serious limitation in Theorem 2 since most possibility density functions meet this requirement. For example, every possibility distribution obtained by the optimal transformation [8] of a probability distribution possesses a Borel measurable density function.

Example 5 (Consistency of Inverse Distributions on X). Let X and Y be the uncertain variables with their respective probability and possibility distributions given in Exs. 3 and 4. Alongside p_X , the probability density functions $p'_{\phi^{-1}(Y)}$ and $p''_{\phi^{-1}(Y)}$ shown

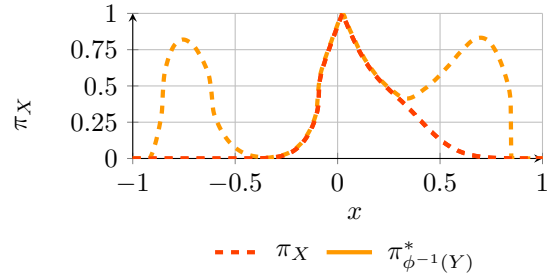
in Fig. 3a also induce admissible probability distributions \mathbf{P}_X , $\mathbf{P}'_{\phi^{-1}(Y)}$ and $\mathbf{P}''_{\phi^{-1}(Y)}$ of X that yield \mathbf{P}_Y as their pushforward by ϕ .

Alongside π_X , the minimum specific inverse possibility density function $\pi_{\phi^{-1}(Y)}^*$ shown in Fig. 3b induces the minimum specific inverse possibility distribution of X yielding Π_Y as its pushforward by ϕ .

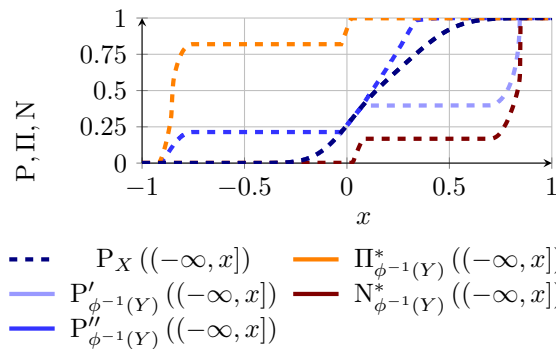
From Theorem 2 it follows immediately, that the probability distributions \mathbf{P}_X , $\mathbf{P}'_{\phi^{-1}(Y)}$ and $\mathbf{P}''_{\phi^{-1}(Y)}$ are consistent with the minimum specific inverse possibility distribution $\Pi_{\phi^{-1}(Y)}^*$. Consequently, their cumulative probability distribution functions are bounded by the cumulative possibility distribution function $\Pi_{\phi^{-1}(Y)}^*((-\infty, x])$ and by the cumulative necessity distribution function $N_{\phi^{-1}(Y)}^*((-\infty, x])$, all shown in Fig. 3c.



(a) Admissible Inverse Probability Density Functions of X in $\mathcal{I}_{\mathbf{P}_{\mathbf{Y}}}^\phi$.



(b) Admissible Inverse Possibility Density Functions of X in $\mathcal{I}_{\Pi_{\mathbf{Y}}}^\phi$.



(c) Admissible Inverse Cumulative Distribution Functions of X .

Figure 3: Distributions in Ex. 5.

3.3 Proofs

Below, the proofs of Theorems 1 and 2 are given. Readers not interested in the mathematical details may skip this part and continue in Section 4.

3.3.1 Proof of Theorem 1

In the following proof, it is useful to notice that

$$\begin{aligned}\Pi_{\phi^{-1}(\mathbf{Y})}^*(U_X) &= \sup \pi_{\phi^{-1}(\mathbf{Y})}^*(U_X) \\ &= \sup \pi_{\mathbf{Y}}(\phi(U_X)) \\ &= \Pi_{\mathbf{Y}}(\phi(U_X))\end{aligned}\quad (24)$$

for all $U_X \subseteq \mathbb{R}^M$.

a. To show that $\Pi_{\phi^{-1}(\mathbf{Y})}^*$ is a possibility measure on \mathbb{R}^M , consider

1. $\Pi_{\phi^{-1}(\mathbf{Y})}^*(\emptyset) = \Pi_{\mathbf{Y}}(\phi(\emptyset)) = \Pi_{\mathbf{Y}}(\emptyset) = 0$,
2. $\Pi_{\phi^{-1}(\mathbf{Y})}^*(\mathbb{R}^M) = \Pi_{\mathbf{Y}}(\phi(\mathbb{R}^M)) \stackrel{\phi \text{ surjective}}{=} \Pi_{\mathbf{Y}}(\mathbb{R}^N) = 1$ and
3. if $\{U_{X,n} \subseteq \mathbb{R}^M\}_{n \in \mathbb{N}}$ is a countable collection of pairwise disjoint subsets with union U_X , then

$$\begin{aligned}\Pi_{\phi^{-1}(\mathbf{Y})}^*(U_X) &= \Pi_{\mathbf{Y}}(\phi(U_X)) \\ &= \Pi_{\mathbf{Y}}\left(\phi\left(\bigcup_{n \in \mathbb{N}} U_{X,n}\right)\right) \\ &= \Pi_{\mathbf{Y}}\left(\bigcup_{n \in \mathbb{N}} \phi(U_{X,n})\right) \\ &= \sup_{n \in \mathbb{N}} \Pi_{\mathbf{Y}}(\phi(U_{X,n})) \\ &= \sup_{n \in \mathbb{N}} \Pi_{\phi^{-1}(\mathbf{Y})}^*(U_{X,n})\end{aligned}$$

b. Let $U_Y \in \mathcal{B}(\mathbb{R}^N)$. Then,

$$\begin{aligned}\Pi_{\phi^{-1}(\mathbf{Y})}^*(\phi^{-1}(U_Y)) &= \Pi_{\mathbf{Y}}(\phi(\phi^{-1}(U_Y))) \\ &= \Pi_{\mathbf{Y}}(U_Y)\end{aligned}\quad (25)$$

and, thus, $\Pi_{\phi^{-1}(\mathbf{Y})}^* \in \mathcal{I}_{\Pi_{\mathbf{Y}}}^{\phi}$.

c. Let $\Pi'_{\phi^{-1}(\mathbf{Y})} \in \mathcal{I}_{\Pi_{\mathbf{Y}}}^{\phi}$ and $\mathbf{x} \in \mathbb{R}^M$. Then,

$$\begin{aligned}\pi'_{\phi^{-1}(\mathbf{Y})}(\mathbf{x}) &\leq \sup \pi'_{\phi^{-1}(\mathbf{Y})}(\phi^{-1}(\phi(\{\mathbf{x}\}))) \\ &= \Pi'_{\phi^{-1}(\mathbf{Y})}\left(\phi^{-1}\left(\underbrace{\phi(\{\mathbf{x}\})}_{=\{\phi(\mathbf{x})\} \in \mathcal{B}(\mathbb{R}^N)}\right)\right) \\ &= \Pi_{\mathbf{Y}}(\phi(\{\mathbf{x}\})) \\ &= \pi_{\phi^{-1}(\mathbf{Y})}^*(\mathbf{x}),\end{aligned}\quad (26)$$

and with Lemma 3 it follows that $\Pi'_{\phi^{-1}(\mathbf{Y})} \leq \Pi_{\phi^{-1}(\mathbf{Y})}^*$.

3.3.2 Proof of Theorem 2

Let $P_{\mathbf{Y}} \in \mathcal{P}_{\Pi_{\mathbf{Y}}}$ and $P'_{\mathbf{X}} \in \mathcal{I}_{P_{\mathbf{Y}}}^{\phi}$. To show that $P'_{\mathbf{X}} \in \mathcal{P}_{\Pi_{\phi^{-1}(\mathbf{Y})}^*}$, let $U_X \in \mathcal{B}(\mathbb{R}^M)$ and define

$$\alpha = \Pi_{\phi^{-1}(\mathbf{Y})}^*(U_X). \quad (27)$$

Since ϕ is a Borel measurable function, the pre-image of the measurable sub-level set

$$S^{\alpha} = \{\mathbf{y} \in \mathbb{R}^N : \pi_{\mathbf{Y}}(\mathbf{y}) \leq \alpha\} \in \mathcal{B}(\mathbb{R}^N) \quad (28)$$

is a measurable set since $\pi_{\mathbf{Y}}$ is a measurable function, i.e.

$$\phi^{-1}(S^{\alpha}) \in \mathcal{B}(\mathbb{R}^M). \quad (29)$$

Trivially, S^{α} also satisfies

$$\Pi_{\mathbf{Y}}(S^{\alpha}) = \sup \pi_{\mathbf{Y}}(S^{\alpha}) \leq \alpha. \quad (30)$$

Furthermore, let $\mathbf{x} \in U_X$. The evaluation of

$$\begin{aligned}\pi_{\mathbf{Y}}(\phi(\mathbf{x})) &\stackrel{\text{Thm. 1}}{=} \pi_{\phi^{-1}(\mathbf{Y})}^*(\mathbf{x}) \\ &\stackrel{\mathbf{x} \in U_X}{\leq} \sup \pi_{\phi^{-1}(\mathbf{Y})}^*(U_X) \\ &= \Pi_{\phi^{-1}(\mathbf{Y})}^*(U_X) \\ &\stackrel{\text{Eq. (27)}}{=} \alpha\end{aligned}\quad (31)$$

yields $\phi(\mathbf{x}) \in S^{\alpha}$ and therefore $\phi(U_X) \subseteq S^{\alpha}$. Ultimately, one obtains

$$U_X \subseteq \phi^{-1}(\phi(U_X)) \subseteq \phi^{-1}(S^{\alpha}). \quad (32)$$

Then, it is straightforward to show that

$$\begin{aligned}P'_{\mathbf{X}}(U_X) &\stackrel{\text{Eq. (32)}}{\leq} P'_{\mathbf{X}}\left(\underbrace{\phi^{-1}(S^{\alpha})}_{\substack{\text{Eq. (29)} \\ \in \mathcal{B}(\mathbb{R}^M)}}\right) \\ &\stackrel{P'_{\mathbf{X}} \in \mathcal{I}_{P_{\mathbf{Y}}}^{\phi}}{=} P_{\mathbf{Y}}(S^{\alpha}) \\ &\stackrel{P_{\mathbf{Y}} \in \mathcal{P}_{\Pi_{\mathbf{Y}}}}{\leq} \Pi_{\mathbf{Y}}(S^{\alpha}) \\ &\stackrel{\text{Eq. (30)}}{\leq} \alpha \\ &\stackrel{\text{Eq. (27)}}{=} \Pi_{\phi^{-1}(\mathbf{Y})}^*(U_X)\end{aligned}\quad (33)$$

proving the theorem.

4 Discussion

Theorem 2 is a powerful result since it enables one to capture the entire set of admissible inverse probability distributions to the inverse probability propagation problem in just one possibility distribution. This minimum specific inverse possibility distribution can then be used for further analysis of the system under uncertainty, and bounds for the underlying probability distribution can be guaranteed nonetheless.

Due to the limited length of this contribution, corollary results about the upper and lower previsions induced by the various credal sets cannot be included here. However – in a nutshell – the expected values of all admissible inverse probability distributions on X can easily be bounded through the upper and lower expected values of the minimum specific inverse possibility distribution.

In the past, possibility theory has often been criticized by statisticians for being an unnecessary theory of uncertainty that would only present solutions to problems already solved by probability theory [13]. In this contribution, the authors hope to have made another case for the usefulness of possibility theory, showing that the stated inverse problem, which is generally hard to solve within probability theory, can be handled very efficiently by possibility theory.

Acknowledgement

The authors gratefully acknowledge the support provided by the German Research Foundation (DFG) in the framework of the research project no. 319924547. They would also like to thank the anonymous referees as well as Dennis Grunert and Markus Mäck for their valuable remarks during the preparation of this paper.

References

- [1] C. Baudrit, I. Couso, D. Dubois, et al., Joint propagation of probability and possibility in risk analysis: Towards a formal framework, *International Journal of Approximate Reasoning* 45 (1) (2007) 82–105.
- [2] C. Baudrit, D. Dubois, Practical representations of incomplete probabilistic knowledge, *Computational Statistics & Data Analysis* 51 (1) (2006) 86–108.
- [3] A. P. Dempster, Upper and lower probabilities induced by a multivalued mapping, in: *Classic Works of the Dempster-Shafer Theory of Belief Functions*, Springer, 2008, 57–72.
- [4] S. Destercke, D. Dubois, E. Chojnacki, Unifying practical uncertainty representations. Part I: Generalized p-boxes, *International Journal of Approximate Reasoning* 49 (3) (2008) 649–663.
- [5] D. Dubois, E. Kerre, R. Mesiar, H. Prade, Fuzzy interval analysis, in: *Fundamentals of Fuzzy Sets*, Springer, 2000, 483–581.
- [6] D. Dubois, H. Prade, Random sets and fuzzy interval analysis, *Fuzzy Sets and Systems* 42 (1) (1991) 87–101.
- [7] D. Dubois, H. Prade, When upper probabilities are possibility measures, *Fuzzy Sets and Systems* 49 (1) (1992) 65–74.
- [8] D. Dubois, H. Prade, Practical methods for constructing possibility distributions, *International Journal of Intelligent Systems* 31 (3) (2016) 215–239.
- [9] M. Hanss, *Applied Fuzzy Arithmetic – An Introduction with Engineering Applications*, Springer, Berlin, 2005.
- [10] D. Hose, M. Hanss, On inverse fuzzy arithmetical problems in uncertainty analysis, in: *Proceedings of the 7th International Conference on Uncertainties in Structural Dynamics*, Leuven (Belgium), 2018.
- [11] E. T. Jaynes, Information theory and statistical mechanics, *Physical Review* 106 (4) (1957) 620–630.
- [12] G. Klir, M. Wierman, *Uncertainty-Based Information: Elements of Generalized Information Theory*, Springer Science & Business Media, 1999.
- [13] M. Laviolette, J. W. Seaman, J. D. Barrett, W. H. Woodall, A probabilistic and statistical view of fuzzy methods, *Technometrics* 37 (3) (1995) 249–261.
- [14] B. Liu, *Theory and Practice of Uncertain Programming*, Springer, 2002.
- [15] W. A. Lodwick, D. Dubois, Interval linear systems as a necessary step in fuzzy linear systems, *Fuzzy Sets and Systems* 281 (2015) 227–251.
- [16] G. Mauris, The principle of possibility maximum specificity as a basis for measurement uncertainty expression, in: *IEEE International Workshop on Advanced Methods for Uncertainty Estimation in Measurement, AMUEM 2009, IEEE, 2009, 5–9*.
- [17] P. Smets, Constructing the pignistic probability function in a context of uncertainty, in: *Uncertainty in Artificial Intelligence, Vol. 10 of Machine Intelligence and Pattern Recognition*, North-Holland, 1990, 29–39.
- [18] L. Stefanini, A generalization of Hukuhara difference and division for interval and fuzzy arithmetic, *Fuzzy Sets and Systems* 161 (11) (2010) 1564–1584.
- [19] T. J. Sullivan, *Introduction to Uncertainty Quantification*, Springer, 2015.
- [20] H. Tanaka, Fuzzy data analysis by possibilistic linear models, *Fuzzy Sets and Systems* 24 (3) (1987) 363–375.