On the Solution of Forward and Inverse Problems in Possibilistic Uncertainty Quantification for Dynamical Systems

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

July 30, 2020

Preprint submitted to 9th International Workshop on Reliable Engineering Computing on June 3, 2020.

Abstract

In this contribution, we adress an apparent lack of methods for the robust analysis of dynamical systems when neither a precise statistical nor an entirely epistemic description of the present uncertainties is possible. Relying on recent results of possibilistic calculus, we revisit standard prediction and filtering problems and show how these may be solved in a numerically exact way.

Keywords: State Estimation, Filter Design, Possibility Theory, Uncertainty Quantification, Imprecise Probabilities

1 Introduction

In the analysis and control of dynamical systems [1], the forward problem usually involves a prediction about the future system behavior, while the inverse problem aims at finding the feasible configurations of past system states that are in agreement with the system dynamics and the obtained measurements. Therein, the inclusion of non-probabilistic models of uncertainty is becoming increasingly more important.

The available techniques for robust system analysis can be divided into two approaches: Stochastic approaches, on the one hand, look back on a long tradition in the quantification of aleatory uncertainty in dynamical systems with popular results, such as the Kalman filter [2]. However, most of the available techniques share a common denominator; when considering the various forms of noise, they rely on precise probability distributions for their description. On the other hand, purely set-based approaches, such as tube-based model predictive control [3], usually do not take into account any stochastic information about the disturbances. Instead, they consider only epistemic uncertainty by finding sets to which the disturbances are confined. The simultaneous consideration of both aleatory and epistemic uncertainties is a topic which is seldomly adressed.

Nonetheless, often a mix of statistical variability and an epistemic lack of knowledge is present in models of real-world processes and there is a need for sensible uncertainty quantification frameworks which can address both these types of uncertainty. In this contribution, we demonstrate how possibility theory [4] originating from fuzzy set theory is one such framework, and we derive tools for a possibilistic analysis of linear time-invariant (LTI) systems. Possibilistic calculus is based on the idea of descriptions of imprecise probabilities, and hence possibilities may be interpreted intuitively by viewing them as upper probabilities. This allows for a more general description of measurement error and process noise due to not being restricted to unique probability distributions. Essentially, this calculus consists of the analysis of nested confidence intervals for imprecisely defined probability distributions, for which we present intuitive techniques for prediction and estimation. Special emphasis is put on the preservation of possibility-probability consistency in all calculations, which is enabled by recent theoretical results.

2 Possibility Spaces

We briefly recount the basic terminology of possibility theory as formalized in [4]. A set-valued function $\Pi : 2^\Omega \to [0,1]$ defined on the power set 2^Ω of the universe of discourse Ω may be called a *possibility measure* if it satisfies three conditions

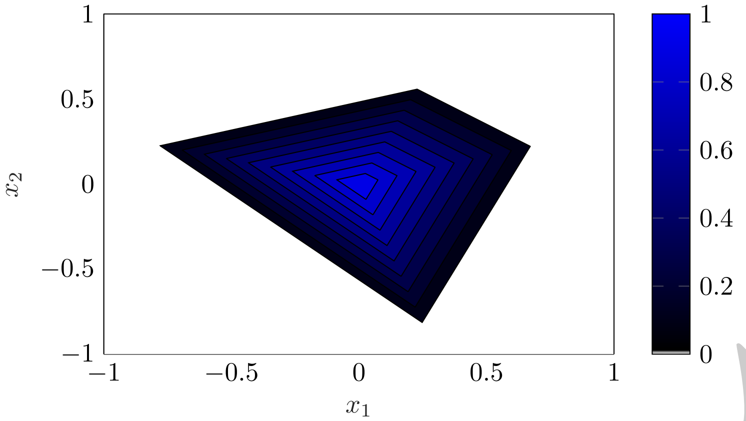

Figure 1: Convex α-cuts of a 2-dimensional convex fuzzy variable.

similar to those of a probability measure. Explicitly, $\Pi(\emptyset) = 0$, $\Pi(\Omega) = 1$ and $\Pi(\bigcup_k U_k) = \sup_k \Pi(U_k)$ for countable collections of sets $U_k \subseteq \Omega$ have to be fulfilled. The dual *necessity measure* is defined by $\mathrm{N}(U) = 1 - \Pi(\Omega \setminus U)$ for all $U \subseteq \Omega$.

Let $\widetilde{X} : \Omega \to \mathscr{X}$ be an ($\mathscr{X}$-valued) *uncertain (random, fuzzy, etc.) variable*. Below, $\mathscr{S} = \mathscr{S}(\mathscr{X})$ denotes a σ-field on $\mathscr{X}$. In practical applications, typically $\mathscr{X} \subseteq \mathbb{R}^N$, and $\mathscr{S}$ is the corresponding Borel σ-algebra. The uncertain variable $\widetilde{X}$ possesses a *possibility distribution* $\Pi_{\widetilde{X}}(U) = \Pi(\widetilde{X}^{-1}(U)) = \Pi(\{\omega \in \Omega : \widetilde{X}(\omega) \in U\})$ for all $U \in \mathscr{S}$ and a function $\pi_{\widetilde{X}} : \mathscr{X} \to [0,1]$ induces a possibility distribution via $\Pi_{\widetilde{X}}(U) = \sup_{x \in U} \pi_{\widetilde{X}}(x)$ if it is a measurable function satisfying $\sup_{x \in \mathscr{X}} \pi_{\widetilde{X}}(x) = 1$. It is then called a *possibility density*. The *superlevel sets/α-cuts* are defined as $\mathscr{C}^{\alpha}_{\Pi_{\widetilde{X}}} = \{x \in \mathscr{X} : \pi_{\widetilde{X}}(x) > \alpha\}$.

It is well-known that possibility theory offers a general framework for the analysis of *imprecise probabilities* [5], i.e. for the consideration of partially defined probability distributions on the measure space $(\mathscr{X}, \mathscr{S})$. The fundamental principle of *probability-possibility consistency* was first introduced in [6]. It states that a probability measure $\mathrm{P}_{\widetilde{X}}$ and a possibility measure $\Pi_{\widetilde{X}}$ on $(\mathscr{X}, \mathscr{S})$ are consistent if the probability of any event is bounded from above by its possibility, i.e. if

$$\mathrm{P}_{\widetilde{X}}(U) \leq \Pi_{\widetilde{X}}(U) \qquad \forall\, U \in \mathscr{S} \tag{1}$$

and consequently from below by the necessity, $\mathrm{N}_{\widetilde{X}}(U) \leq \mathrm{P}_{\widetilde{X}}(U)$. In short, we write $\mathrm{P}_{\widetilde{X}} \preceq \Pi_{\widetilde{X}}$. Notice that the necessity – and hence the lower probability – of any α-cut is always bounded according to $\mathrm{N}_{\widetilde{X}}(\mathscr{C}^{\alpha}_{\widetilde{X}}) \geq 1 - \alpha$ for all $\alpha \in [0,1]$. In order to show consistency, one does not need to check consistency for all possible events as in Eq.(1). Instead, it is shown in [7] that it suffices to check the condition

$$\mathrm{P}_{\widetilde{X}}\left(\mathscr{C}^{\alpha}_{\Pi_{\widetilde{X}}}\right) \geq 1 - \alpha \qquad \forall\, \alpha \in [0,1]. \tag{2}$$

Furthermore, consistency is not a one-to-one relationship. In particular, given a possibility measure $\Pi_{\widetilde{X}}$, the credal set

$$\mathfrak{C}\left(\Pi_{\widetilde{X}}\right) = \left\{\mathrm{P}_{\widetilde{X}} : \mathrm{P}_{\widetilde{X}} \preceq \Pi_{\widetilde{X}}\right\} \tag{3}$$

contains a potentially infinite number of elements, i.e. of consistent probability distributions. One important observation is that the credal set of a possibility distribution $\Pi^1_{\widetilde{X}}$ is included in that of a second one $\Pi^2_{\widetilde{X}}$, i.e. $\mathfrak{C}(\Pi^1_{\widetilde{X}}) \subseteq \mathfrak{C}(\Pi^2_{\widetilde{X}})$ if and only if $\pi^1_{\widetilde{X}}(x) \leq \pi^2_{\widetilde{X}}(x)$ for all $x \in \mathscr{X}$. Then, $\Pi^1_{\widetilde{X}}$ is said to be *more specific* than $\Pi^2_{\widetilde{X}}$.

For an efficient analysis, we propose a quadruple description of convex fuzzy variables $\widetilde{\mathbf{X}} = (\mathbf{A}, \mathbf{b}, \mathbf{C}, \mathbf{d})$, an N^{ext}_X-dimensional generalization of the well-known one-dimensional fuzzy numbers. Therein, the matrix $\mathbf{A} = \mathbf{A}(\alpha) \in \mathbb{R}^{M^{\text{int}}_X \times N^{\text{int}}_X}$ and the vector $\mathbf{b} = \mathbf{b}(\alpha) \in \mathbb{R}^{M^{\text{int}}_X}$ are associated with an α-dependent internal halfspace. The matrix $\mathbf{C} \in \mathbb{R}^{N^{\text{ext}}_X \times N^{\text{int}}_X}$ and the vector $\mathbf{d} \in \mathbb{R}^{N^{\text{ext}}_X}$ provide the external description. The convex α-cuts of $\widetilde{\mathbf{X}}$, i.e. the superlevel sets of its possibility density, are given by

$$\mathscr{C}^{\alpha}_{\widetilde{\mathbf{X}}} = \left\{\mathbf{x} = \mathbf{C} \cdot \xi + \mathbf{d} \;:\; \xi \in \mathbb{R}^{N^{\text{int}}_X} \wedge \mathbf{A}(\alpha) \cdot \xi \leq \mathbf{b}(\alpha)\right\} \qquad \forall\, \alpha \in [0,1]. \tag{4}$$

See Fig. 1 vor a visual representation of the α-cuts of a 2-dimensional convex fuzzy variable.

The possibility distribution may be reconstructed by the Decomposition Theorem

$$\pi_{\widetilde{\mathbf{X}}}(\mathbf{x}) = \sup\left\{\alpha \in [0,1] : \mathbf{x} \in \mathscr{C}^{\alpha}_{\widetilde{\mathbf{X}}}\right\} \qquad \forall\, \mathbf{x} \in \mathbb{R}^{N^{\text{ext}}_X} \tag{5}$$

which is e.g. found in [8]. For a variety of reasons, this representation is well suited for the uncertainty analysis of linear time-invariant systems: Firstly, through the division in an internal and an external description, it is possible to account for interdependency of multivariate uncertain variables and, thus, to avoid certain effects of overestimation. Secondly, the rules of possibilistic calculus may be evaluated very efficiently for a selection of operations. And thirdly, it naturally fits into optimization frameworks, given the convexity of the α-cuts. The benefits are also highlighted below.

3 Possibilistic Calculus

In the following, selected elements of possibilistic calculus are presented. The guiding principle behind the specific implementations of these operations is the *preservation of consistency*. That is, if a possibility distribution and a probability distribution were consistent before performing an operation, such as the the aggregation of several distributions or the propagation through a model, then the resulting distributions ought to be consistent as well. This ensures that possibilities can truly be seen as upper probabilities. Special emphasis is put on the implementation for convex fuzzy variables.

3.1 Aggregation

Often only the marginal distributions $\pi_{\widetilde{X}_1}, \ldots, \pi_{\widetilde{X}_N}$ on the distinct domains $\mathscr{X}_1, \ldots, \mathscr{X}_N$ are available. Yet, e.g. the propagation and the inversion operations discussed below require the joint distribution which has to be constructed by a suitable aggregation operator. Consistency-preserving aggregation operators are presented e.g. in [9] and, more recently, in [10]. In the latter, it is shown that if the probability distributions in the respective credal sets are assumed to be independent, then a consistency-preserving joint possibility distribution is given by

$$\pi^{\mathrm{ind}}_{\widetilde{X}_1,\ldots,\widetilde{X}_N}(x_1,\ldots,x_N) = \min_{i=1,\ldots,N} 1 - \left(1 - \pi_{\widetilde{X}_i}(x_i)\right)^N \quad \forall\, (x_1,\ldots,x_N) \in \mathscr{X}_1 \times \ldots \times \mathscr{X}_N. \tag{6}$$

From the proof of Theorem 28 in [10], we infer that Eq. (6) is equivalent to

$$\mathscr{C}^{\alpha}_{\Pi^{\mathrm{ind}}_{\widetilde{X}_1,\ldots,\widetilde{X}_n}} = \mathscr{C}^{1-\sqrt[N]{1-\alpha}}_{\Pi_{\widetilde{X}_1}} \times \ldots \times \mathscr{C}^{1-\sqrt[N]{1-\alpha}}_{\Pi_{\widetilde{X}_n}} \qquad \forall\, \alpha \in [0,1]. \tag{7}$$

Hence, if $\widetilde{X}_i = (\mathbf{A}_{\widetilde{X}_i}, \mathbf{b}_{\widetilde{X}_i}, \mathbf{C}_{\widetilde{X}_i}, \mathbf{d}_{\widetilde{X}_i})$ for $i = 1,\ldots,N$ are N given convex fuzzy variables, then the joint fuzzy vector $\widetilde{\mathbf{X}} = (\mathbf{A}_{\widetilde{\mathbf{X}}}, \mathbf{b}_{\widetilde{\mathbf{X}}}, \mathbf{C}_{\widetilde{\mathbf{X}}}, \mathbf{d}_{\widetilde{\mathbf{X}}})$ is also a convex fuzzy variable with the matrices

$$\mathbf{A}_{\widetilde{\mathbf{X}}} = \begin{bmatrix} \mathbf{A}_{\widetilde{X}_1} \circ \Gamma^N & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \mathbf{A}_{\widetilde{X}_N} \circ \Gamma^N \end{bmatrix}, \; \mathbf{b}_{\widetilde{\mathbf{X}}} = \begin{bmatrix} \mathbf{b}_{\widetilde{X}_1} \circ \Gamma^N \\ \vdots \\ \mathbf{b}_{\widetilde{X}_N} \circ \Gamma^N \end{bmatrix}, \; \mathbf{C}_{\widetilde{\mathbf{X}}} = \begin{bmatrix} \mathbf{C}_{\widetilde{X}_1} & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \mathbf{C}_{\widetilde{X}_N} \end{bmatrix}, \; \mathbf{d}_{\widetilde{\mathbf{X}}} = \begin{bmatrix} \mathbf{d}_{\widetilde{X}_1} \\ \vdots \\ \mathbf{d}_{\widetilde{X}_N} \end{bmatrix}$$

where $\Gamma^N : \alpha \mapsto 1 - \sqrt[N]{1-\alpha}$ is a (monotone) rescaling function.

3.2 Propagation

The pushforward distribution of a fuzzy variable $\widetilde{Y} = \phi(\widetilde{X})$ under a $(\mathscr{X}, \mathscr{Y})$-measurable function $\phi : \mathscr{X} \to \mathscr{Y}$ is given by the extension principle [11]

$$\pi_{\widetilde{Y}}(y) = \sup_{x \in \mathscr{X} : y = \phi(x)} \pi_{\widetilde{X}}(x) \qquad \forall\, y \in \mathscr{Y}. \tag{8}$$

The preservation of consistency is discussed e.g. in [12] and [10]. Of course, if only the marginals $\widetilde{X}_1, \ldots, \widetilde{X}_N$ are known, then they need to be aggregated into the joint vector $\widetilde{\mathbf{X}}$ beforehand. E.g. in [12] it is shown that Zadeh's non-interactive aggregation, i.e. the min-intersection of the possibility densities, which appears in the original formulation of the extension principle, does not preserve consistency.

The connection to the propagation of intervals is discussed in [8]. In particular, it can be shown that Eq. (8) is equivalent to

$$\mathscr{C}^{\alpha}_{\Pi_{\widetilde{Y}}} = \phi\left(\mathscr{C}^{\alpha}_{\Pi_{\widetilde{X}}}\right) \qquad \forall\, \alpha \in [0,1]. \tag{9}$$

Hence, if $\widetilde{\mathbf{X}} = (\mathbf{A}, \mathbf{b}, \mathbf{C}, \mathbf{d})$ is a given convex fuzzy variable and ϕ is a linear map represented by the matrix $\mathbf{M} \in \mathbb{R}^{N_{\mathscr{Y}} \times N_{\mathscr{X}}}$, then $\widetilde{\mathbf{Y}} = (\mathbf{A}, \mathbf{b}, \mathbf{MC}, \mathbf{Md})$ is also a convex fuzzy variable. E.g. the sum $\widetilde{Y} = \sum_{i=1}^{N} \widetilde{X}_i$ may then be obtained by propagation of $\widetilde{\mathbf{X}}$ under the matrix $\mathbf{M} = [1, \ldots, 1]$.

Marginalization, i.e. the computation of the marginal distributions $\pi_{\widetilde{X}_1}, \ldots, \pi_{\widetilde{X}_N}$ from the joint distribution $\pi_{\widetilde{X}_1,\ldots,\widetilde{X}_N}$ is just a special case of propagation under $\phi : \mathbf{x} \mapsto x_i$ for $i = 1, \ldots N$. Hence, it preserves consistency as well.

3.3 Inversion

The preservation of consistency for inversion under a surjective and $(\mathscr{X},\mathscr{Y})$-measurable function $\phi : \mathscr{X} \to \mathscr{Y}$ is investigated in [13]. Suppose the possibility distribution of $\widetilde{Y}$ is known and one wishes to infer the distribution of $\widetilde{X}$. Generally, there exists a (possibly infinite) number of possibility distributions $\Pi_{\widetilde{X}}$ yielding the pushforward distribution $\Pi_{\widetilde{Y}}$ under ϕ. These extensions may be gathered in the *set of inverse possibility distributions*

$$\mathscr{I}^{\phi}_{\Pi_{\widetilde{Y}}} = \left\{\Pi_{\widetilde{X}} : \Pi_{\widetilde{Y}}(V) = \Pi_{\widetilde{X}}\left(\phi^{-1}(V)\right) \; \forall V \in \mathscr{S}(\mathscr{Y})\right\}. \tag{10}$$

However, it is possible to account for this whole set by just one possibility distribution. The *minimum specific inverse possibility distribution* is given by

$$\pi^{\text{inv}}_{\widetilde{X}}(x) = \Pi_{\widetilde{Y}}(\phi(x)) \qquad \forall x \in \mathscr{X} \tag{11}$$

This distribution is the least specific possibility distribution in the corresponding set of inverse possibility distributions and therefore a suitable representation thereof. Analogously, one can define a set of inverse probability distributions $\mathscr{I}^{\phi}_{\mathrm{P}_{\widetilde{Y}}}$ given a probability distribution $\mathrm{P}_{\widetilde{Y}}$. The perhaps most important result is that $\Pi^{\text{inv}}_{\widetilde{X}}$ is consistent with all probability distributions $\mathrm{P}_{\widetilde{X}} \in \mathscr{I}^{\phi}_{\mathrm{P}_{\widetilde{Y}}}$ from the set of inverse probability distributions $\mathscr{I}^{\phi}_{\mathrm{P}_{\widetilde{Y}}}$ of all probability distributions $\mathrm{P}_{\widetilde{Y}} \in \mathscr{C}\left(\Pi_{\widetilde{Y}}\right)$ which are consistent with $\Pi_{\widetilde{Y}}$. Analogously to the forward propagation, we infer that Eq. (11) is equivalent to

$$\mathscr{C}^{\alpha}_{\Pi^{\text{inv}}_{\widetilde{X}}} = \phi^{-1}\left(\mathscr{C}^{\alpha}_{\Pi_{\widetilde{Y}}}\right) \qquad \forall \alpha \in [0,1]. \tag{12}$$

Hence, if $\widetilde{\mathbf{Y}} = \left(\mathbf{A}_{\widetilde{\mathbf{Y}}}, \mathbf{b}_{\widetilde{\mathbf{Y}}}, \mathbf{C}_{\widetilde{\mathbf{Y}}}, \mathbf{d}_{\widetilde{\mathbf{Y}}}\right)$ is a given convex fuzzy variable and ϕ is a linear map, i.e. it corresponds to a matrix $\mathbf{M} \in \mathbb{R}^{N_{\mathbf{Y}} \times N_{\mathbf{X}}}$, then one obtains the convex fuzzy variable $\widetilde{\mathbf{X}} = \left(\mathbf{A}_{\widetilde{\mathbf{X}}}, \mathbf{b}_{\widetilde{\mathbf{X}}}, \mathbf{C}_{\widetilde{\mathbf{X}}}, \mathbf{d}_{\widetilde{\mathbf{X}}}\right)$ with the matrices

$$\mathbf{A}_{\widetilde{\mathbf{X}}} = \begin{bmatrix} \mathbf{A}_{\widetilde{\mathbf{Y}}} & \mathbf{0} \\ -\mathbf{C}_{\widetilde{\mathbf{Y}}} & \mathbf{M} \\ \mathbf{C}_{\widetilde{\mathbf{Y}}} & -\mathbf{M} \end{bmatrix}, \quad \mathbf{b}_{\widetilde{\mathbf{X}}} = \begin{bmatrix} \mathbf{b}_{\widetilde{\mathbf{Y}}} \\ \mathbf{d}_{\widetilde{\mathbf{Y}}} \\ \mathbf{d}_{\widetilde{\mathbf{Y}}} \end{bmatrix}, \quad \mathbf{C}_{\widetilde{\mathbf{X}}} = \begin{bmatrix} \mathbf{0} & \mathbb{I} \end{bmatrix} \quad \text{and} \quad \mathbf{d}_{\widetilde{\mathbf{X}}} = \mathbf{0}$$

from the minimum specific inverse possibility distribution. This operation augments the vector of interior variables to $\xi_{\widetilde{\mathbf{X}}} = \begin{bmatrix} \xi^{\mathrm{T}}_{\widetilde{\mathbf{Y}}} & \mathbf{x}^{\mathrm{T}} \end{bmatrix}^{\mathrm{T}}$ where $\xi^{\mathrm{T}}_{\widetilde{\mathbf{Y}}}$ are the interior variables of $\widetilde{\mathbf{Y}}$. Obviously, the lower two rows of the internal halfspace representation correspond to the equality constraint $\mathbf{C}_{\widetilde{\mathbf{Y}}}\xi_{\widetilde{\mathbf{Y}}} + \mathbf{d}_{\widetilde{\mathbf{Y}}} = \mathbf{M}\mathbf{x}$ which might as well be considered explicitly.

3.4 Confidence Sets

For the inverse problem, it is furthermore helpful to study the relation between α-cuts and confidence sets. Assume an uncertain variable $\widetilde{Y} = \psi(\vartheta, \widetilde{X})$ be a function of the uncertain variable $\widetilde{X}$ with known possibility distribution and of an unknown parameter $\vartheta \in \Theta$. Adopting a frequentist point of view, ϑ is not an uncertain variable, and hence, probabilistic or possibilistic expressions about it require discretion. Nevertheless, with the assumed knowledge, we can construct a confidence set $\mathscr{K}^{\alpha}_{\vartheta} = \mathscr{K}^{\alpha}_{\vartheta}(\widetilde{Y}) = \mathscr{K}^{\alpha}_{\vartheta}(\phi(\vartheta, \widetilde{X}))$ of ϑ for a given confidence level $\alpha \in [0,1]$ through

$$\mathscr{K}^{\alpha}_{\vartheta} = \left\{\theta \in \Theta : \widetilde{Y} \in \psi\left(\theta, \mathscr{C}^{\alpha}_{\widetilde{X}}\right)\right\}. \tag{13}$$

Proposition 1. *It holds that* $\mathrm{P}\left(\vartheta \in \mathscr{K}^{\alpha}_{\vartheta}\right) \geq 1 - \alpha$ *for all* $\vartheta \in \Theta$.

Proof. Let $\vartheta \in \Theta$ be arbitrary but fixed. From $\widetilde{Y} = \psi\left(\vartheta, \widetilde{X}\right)$, it follows that

$$\mathscr{K}^{\alpha}_{\vartheta} = \left\{\theta \in \Theta : \psi\left(\vartheta, \widetilde{X}\right) \in \psi\left(\theta, \mathscr{C}^{\alpha}_{\widetilde{X}}\right)\right\}. \tag{14}$$

The set membership definition for $\vartheta \in \mathscr{K}^{\alpha}_{\vartheta}$, i.e. $\psi(\vartheta, \widetilde{X}) \in \psi(\vartheta, \mathscr{C}^{\alpha}_{\widetilde{X}})$, is trivially fulfilled if $\widetilde{X} \in \mathscr{C}^{\alpha}_{\widetilde{X}}$. Hence, the probability of the former being fulfilled is greater than the probability of the latter being fulfilled, yielding

$$\mathrm{P}\left(\vartheta \in \mathscr{K}^{\alpha}_{\vartheta}\right) \geq \mathrm{P}\left(\widetilde{X} \in \mathscr{C}^{\alpha}_{\widetilde{X}}\right) \geq 1 - \alpha \tag{15}$$

which concludes the proof. □

Notice that the proof of Proposition 1 requires the model ψ to be 'correct'. If ψ is not surjective, then it is possible that the confidence levels may be empty above a certain level γ^*. The actual value of γ^* can be computed by checking the existence of feasible solutions, e.g. within a bisection algorithm. Due to this result, we argue that – in possibility theory – α-cuts and confidence sets are similar concepts for uncertain variables and unknown parameters, respectively. Having observed $\widetilde{Y} = y$, we can even define a corresponding confidence density $\gamma\colon \Theta \to [0,1]$ which is computed from the possibility density of $\widetilde{X}$ through

$$\gamma_{\hat{\vartheta}}(\vartheta) = \sup_{x \in \mathscr{X} : y = \psi(\vartheta, x)} \pi_{\widetilde{X}}(x) \tag{16}$$

Furthermore, confidence sets can be computed very efficiently from convex fuzzy variables. Let $\widetilde{\mathbf{X}} = \left(\mathbf{A}_{\widetilde{\mathbf{X}}}, \mathbf{b}_{\widetilde{\mathbf{X}}}, \mathbf{C}_{\widetilde{\mathbf{X}}}, \mathbf{d}_{\widetilde{\mathbf{X}}}\right)$ be a given convex fuzzy variable and let $\psi : (\vartheta, \mathbf{x}) \mapsto \mathbf{M}\mathbf{x} + \mathbf{N}\vartheta$ be an affine map for suitably sized matrices $\mathbf{M}$ and $\mathbf{N}$. Then, an estimator of ϑ can also be expressed in quadruple form $\hat{\vartheta} = \left(\mathbf{A}_{\hat{\vartheta}}, \mathbf{b}_{\hat{\vartheta}}, \mathbf{C}_{\hat{\vartheta}}, \mathbf{d}_{\hat{\vartheta}}\right)$ where

$$\mathbf{A}_{\hat{\vartheta}} = \begin{bmatrix} \mathbf{A}_{\widetilde{\mathbf{X}}} & \mathbf{0} \\ \mathbf{M}\mathbf{C}_{\widetilde{\mathbf{X}}} & \mathbf{N} \\ -\mathbf{M}\mathbf{C}_{\widetilde{\mathbf{X}}} & -\mathbf{N} \end{bmatrix}, \quad \mathbf{b}_{\hat{\vartheta}} = \begin{bmatrix} \mathbf{b}_{\widetilde{\mathbf{X}}} \\ \mathbf{y} - \mathbf{M}\mathbf{d}_{\widetilde{\mathbf{X}}} \\ \mathbf{y} - \mathbf{M}\mathbf{d}_{\widetilde{\mathbf{X}}} \end{bmatrix}, \quad \mathbf{C}_{\hat{\vartheta}} = \begin{bmatrix} \mathbf{0} & \mathbb{I} \end{bmatrix} \quad \text{and} \quad \mathbf{d}_{\hat{\vartheta}} = \mathbf{0}$$

Therein, the lower two rows of the internal halfspace representation correspond to the equality constraint $\mathbf{M}\mathbf{C}_{\widetilde{\mathbf{X}}}\xi_{\widetilde{\mathbf{X}}} + \mathbf{M}\mathbf{d}_{\widetilde{\mathbf{X}}} + \mathbf{N}\theta = \mathbf{y}$ which might as well be considered explicitly.

Again, even though similar in appearance, a confidence density is not a possibility density and a confidence set is not an α-cut from a frequentist point of view. Likewise, even though the computation is similar, the underlying concepts used for the inversion and for the confidence sets are fundamentally different. The inversion intends to find an actual possibility distribution for an input variable from the possibility distribution of a dependent output variable, whereas the confidence distribution indicates sets in which an unknown but fixed parameter is likely to reside.

In a subjective setting, however, it seems reasonable to view $1-\alpha$ as the supremum acceptable buying price to pay for the gamble $f_\alpha : \theta \mapsto \mathscr{I}_{\mathscr{K}_{\hat{\vartheta}}^{\alpha}}(\theta)$, where $\mathscr{I}_{\mathscr{K}_{\hat{\vartheta}}^{\alpha}}$ is the indicator function of the confidence set $\mathscr{K}_{\hat{\vartheta}}^{\alpha}$ for all $\alpha \in [0,1]$ and, subsequently, to define a coherent lower prevision [5] therefrom.

4 Possibilistic System Analysis

The elements of possibilistic calculus in combination with convex fuzzy variables as proposed and described above may be used as a basis for the analysis of LTI systems with possibilistic error descriptions via convex fuzzy variables. For simplicity, we assume the following time-discrete state-space description, omitting any control input or other, which could, however, easily be included. The system dynamics is modeled by

$$\mathbf{x}_{k+1} = \mathbf{F}\mathbf{x}_k + \mathbf{v}_k \tag{17}$$

and the observation model is

$$\mathbf{y}_k = \mathbf{H}\mathbf{x}_k + \mathbf{w}_k. \tag{18}$$

Therein, the process noise $\mathbf{v}_k$ and the measurement error $\mathbf{w}_k$ are assumed to be independent and identically distributed (iid) realizations of the uncertain variables $\widetilde{\mathbf{V}}$ and $\widetilde{\mathbf{W}}$ which are modelled as convex fuzzy variables.

4.1 Prediction

For the purpose of predicting the future system behavior, we assume that information about the current system state $\mathbf{x}_k$ is present in the form of a convex fuzzy variable $\widetilde{\mathbf{X}}_k$. This knowledge facilitates the prediction of the system states $\mathbf{x}_{k+1}, \mathbf{x}_{k+2}, \ldots$ and of the system output $\mathbf{y}_k, \mathbf{y}_{k+1}, \ldots$ at future points in time.

4.1.1 State Prediction

State prediction is mostly concerned with the evaluation of the system dynamics (17) and corresponds to a simple aggregation and propagation. In order to predict $\mathbf{x}_{k+1}$ from $\widetilde{\mathbf{X}}_k$, we consider the joint distribution of the vector $\begin{bmatrix}\widetilde{\mathbf{X}}_k^{\mathrm{T}} & \widetilde{\mathbf{V}}^{\mathrm{T}}\end{bmatrix}^{\mathrm{T}}$. Propagation under $\mathbf{M} = \begin{bmatrix}\mathbf{F} & \mathbb{I}\end{bmatrix}$ yields the convex fuzzy variable $\widetilde{\mathbf{X}}_{k+1}$ with the matrices

$$\mathbf{A}_{\widetilde{\mathbf{X}}_{k+1}} = \begin{bmatrix} \mathbf{A}_{\widetilde{\mathbf{X}}_k} \circ \Gamma^2 & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_{\widetilde{\mathbf{V}}} \circ \Gamma^2 \end{bmatrix}, \qquad \mathbf{b}_{\widetilde{\mathbf{X}}_{k+1}} = \begin{bmatrix} \mathbf{b}_{\widetilde{\mathbf{X}}_k} \circ \Gamma^2 \\ \mathbf{b}_{\widetilde{\mathbf{V}}} \circ \Gamma^2 \end{bmatrix},$$

$$\mathbf{C}_{\widetilde{\mathbf{X}}_{k+1}} = \begin{bmatrix} \mathbf{F}\mathbf{C}_{\widetilde{\mathbf{X}}_k} & \mathbf{C}_{\widetilde{\mathbf{V}}} \end{bmatrix} \qquad \text{and} \qquad \mathbf{d}_{\widetilde{\mathbf{X}}_{k+1}} = \mathbf{F}\mathbf{d}_{\widetilde{\mathbf{X}}_k} + \mathbf{d}_{\widetilde{\mathbf{V}}}.$$

This step can also be applied recursively in order to predict $\widetilde{\mathbf{X}}_{k+2}$, $\widetilde{\mathbf{X}}_{k+3}$, etc.

Proposition 2. *The system state $\widetilde{\mathbf{X}}_{k+1} = \mathbf{F}\widetilde{\mathbf{X}}_k + \widetilde{\mathbf{V}}$ will fall into the α-cut $\mathscr{C}^{\alpha}_{\widetilde{\mathbf{X}}_{k+1}}$ with probability*

$$\mathrm{P}\left(\widetilde{\mathbf{X}}_{k+1} \in \mathscr{C}^{\alpha}_{\widetilde{\mathbf{X}}_{k+1}}\right) \geq 1-\alpha \qquad \forall\, \alpha \in [0,1], \tag{19}$$

for all independent and consistent probability distributions $\mathrm{P}_{\widetilde{\mathbf{X}}_k} \preceq \Pi_{\widetilde{\mathbf{X}}_k}$, and $\mathrm{P}_{\widetilde{\mathbf{V}}} \preceq \Pi_{\widetilde{\mathbf{V}}}$.

Proof. The proposition follows directly from the consistency-preserving properties of the aggregation and propagation. □

4.1.2 Output Prediction

Information about the output $\mathbf{y}_k$ may be obtained by additionally considering the measurement model (18). By propagation of the joint vector $\begin{bmatrix}\widetilde{\mathbf{X}}_k^{\mathrm{T}} & \widetilde{\mathbf{W}}^{\mathrm{T}}\end{bmatrix}^{\mathrm{T}}$ under $\mathbf{M} = \begin{bmatrix}\mathbf{H} & \mathbb{I}\end{bmatrix}$ one obtains the convex fuzzy variable $\widetilde{\mathbf{Y}}_k$ with the matrices

$$\mathbf{A}_{\widetilde{\mathbf{Y}}_k} = \begin{bmatrix}\mathbf{A}_{\widetilde{\mathbf{X}}_k} \circ \Gamma^2 & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_{\widetilde{\mathbf{W}}} \circ \Gamma^2\end{bmatrix}, \qquad \mathbf{b}_{\widetilde{\mathbf{Y}}_k} = \begin{bmatrix}\mathbf{b}_{\widetilde{\mathbf{X}}_k} \circ \Gamma^2 \\ \mathbf{b}_{\widetilde{\mathbf{W}}} \circ \Gamma^2\end{bmatrix},$$

$$\mathbf{C}_{\widetilde{\mathbf{Y}}_k} = \begin{bmatrix}\mathbf{H}\mathbf{C}_{\widetilde{\mathbf{X}}_k} & \mathbf{H}\mathbf{C}_{\widetilde{\mathbf{W}}}\end{bmatrix} \qquad \text{and} \qquad \mathbf{d}_{\widetilde{\mathbf{Y}}_k} = \mathbf{H}\left(\mathbf{d}_{\widetilde{\mathbf{X}}_k} + \mathbf{d}_{\widetilde{\mathbf{W}}}\right).$$

Proposition 3. *The measurement $\widetilde{\mathbf{Y}}_k = \mathbf{H}\widetilde{\mathbf{X}}_k + \widetilde{\mathbf{W}}$ will fall into the α-cut $\mathscr{C}^{\alpha}_{\widetilde{\mathbf{Y}}_k}$ with probability*

$$\mathrm{P}\left(\widetilde{\mathbf{Y}}_k \in \mathscr{C}^{\alpha}_{\widetilde{\mathbf{Y}}_k}\right) \geq 1-\alpha \qquad \forall\, \alpha \in [0,1] \tag{20}$$

for all independent and consistent probability distributions $\mathrm{P}_{\widetilde{\mathbf{X}}_k} \preceq \Pi_{\widetilde{\mathbf{X}}_k}$, and $\mathrm{P}_{\widetilde{\mathbf{W}}} \preceq \Pi_{\widetilde{\mathbf{W}}}$.

Proof. The proposition follows directly from the consistency-preserving properties of the aggregation and propagation. □

4.2 Estimation

The elicitation of information about the system state is performed in a similar manner by projecting the measurements and possible perturbations onto the state space. Considering the involved process noise and measurement error, it is possible to construct confidence sets of the actual system states $\mathbf{x}_k$ for further analysis. Adopting a frequentist point of view, the $\mathbf{x}_k$ are not uncertain variables as they have already been realized.

To this end, we assume that the measurements are available in the vector $\widetilde{\mathbf{P}} = \begin{bmatrix}\widetilde{\mathbf{Y}}_1^{\mathrm{T}} & \dots & \widetilde{\mathbf{Y}}_k^{\mathrm{T}}\end{bmatrix}^{\mathrm{T}}$. Evidently, these measurements only depend on the uncertain variables $\widetilde{\mathbf{Q}} = \begin{bmatrix}\widetilde{\mathbf{V}}^{\mathrm{T}} & \dots & \widetilde{\mathbf{V}}^{\mathrm{T}}\end{bmatrix}^{\mathrm{T}}$, containing $k-1$ instances of the process noise $\widetilde{\mathbf{V}}$, and $\widetilde{\mathbf{R}} = \begin{bmatrix}\widetilde{\mathbf{W}}^{\mathrm{T}} & \dots & \widetilde{\mathbf{W}}^{\mathrm{T}}\end{bmatrix}^{\mathrm{T}}$, containing k instances of the measurement error $\widetilde{\mathbf{W}}$. If we denote by $\vartheta = \begin{bmatrix}\mathbf{x}_1^{\mathrm{T}} & \dots & \mathbf{x}_k^{\mathrm{T}}\end{bmatrix}^{\mathrm{T}}$ the vector of past system states which are to be estimated, then Eqs. (17) and (18) can be written in matrix-vector form as

$$\begin{bmatrix}\mathbf{0} \\ \widetilde{\mathbf{P}}\end{bmatrix} = \underbrace{\left[\begin{array}{cccc} -\mathbf{F} & \mathbb{I} & & \mathbf{0} \\ & \ddots & \ddots & \\ \mathbf{0} & & -\mathbf{F} & \mathbb{I} \\ \hline -\mathbf{H} & & \mathbf{0} & \\ & \ddots & & \\ \mathbf{0} & & -\mathbf{H} & \end{array}\right]}_{=:\mathbf{N}} \vartheta + \begin{bmatrix}\widetilde{\mathbf{Q}} \\ \widetilde{\mathbf{R}}\end{bmatrix}. \tag{21}$$

For $\mathbf{t} = \begin{bmatrix}\mathbf{0}^{\mathrm{T}} & \mathbf{p}^{\mathrm{T}}\end{bmatrix}^{\mathrm{T}}$, where $\mathbf{p}$ is a realization of $\widetilde{\mathbf{P}}$, and the aggregated joint vector $\widetilde{\mathbf{S}} = \begin{bmatrix}\widetilde{\mathbf{Q}}^{\mathrm{T}} & \widetilde{\mathbf{R}}^{\mathrm{T}}\end{bmatrix}^{\mathrm{T}}$, it is possible to estimate ϑ by $\hat{\vartheta}$ with the matrices

$$\mathbf{A}_{\hat{\vartheta}} = \begin{bmatrix}\mathbf{A}_{\widetilde{\mathbf{S}}} & \mathbf{0} \\ \mathbf{C}_{\widetilde{\mathbf{S}}} & \mathbf{N} \\ -\mathbf{C}_{\widetilde{\mathbf{S}}} & -\mathbf{N}\end{bmatrix}, \quad \mathbf{b}_{\hat{\vartheta}} = \begin{bmatrix}\mathbf{b}_{\widetilde{\mathbf{S}}} \\ \mathbf{t} - \mathbf{d}_{\widetilde{\mathbf{S}}} \\ \mathbf{t} - \mathbf{d}_{\widetilde{\mathbf{S}}}\end{bmatrix}, \quad \mathbf{C}_{\hat{\vartheta}} = \begin{bmatrix}\mathbf{0} & \mathbb{I}\end{bmatrix} \quad \text{and} \quad \mathbf{d}_{\hat{\vartheta}} = \mathbf{0}.$$

The confidence sets $\mathscr{K}^{\alpha}_{\hat{\mathbf{x}}_i}$ of $\mathbf{x}_i$ for $i = 1, \dots, k$ can be obtained by marginalization of $\hat{\vartheta}$.

Proposition 4. *The (non-empty) α-cuts $\mathscr{K}_{\hat{\vartheta}}^{\alpha}$ are confidence sets of ϑ, i.e.*

$$\mathrm{P}\left(\vartheta \in \mathscr{K}_{\hat{\vartheta}}^{\alpha}\right) \geq 1-\alpha \qquad \forall\, \alpha \in [0, \alpha^*] \tag{22}$$

for all independent and consistent probability distributions $\mathrm{P}_{\widetilde{\mathbf{V}}} \preceq \Pi_{\widetilde{\mathbf{V}}}$, and $\mathrm{P}_{\widetilde{\mathbf{W}}} \preceq \Pi_{\widetilde{\mathbf{W}}}$.

Proof. The proposition follows from the consistency-preserving properties of the aggregation, and Proposition 1. □

5 Numerical Example

Consider the LTI system with the system and measurement matrices

$$\mathbf{F} = \begin{bmatrix} \frac{1}{6} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{1}{6} \end{bmatrix} \qquad \text{and} \qquad \mathbf{H} = \begin{bmatrix} 1 & 0 \end{bmatrix}. \tag{23}$$

The initial conditions are chosen as $\mathbf{x}_0 = \begin{bmatrix} 0 & 40 \end{bmatrix}^{\mathrm{T}}$, i.e. they are assumed to be known precisely, even though it would suffice to only know a possibility distribution. Furthermore, it is assumed that the process noise $\widetilde{V}_1$ and $\widetilde{V}_2$ each are distributed according to some probability distribution which is unimodal, symmetric about zero and has a support bounded by $|\widetilde{V}_i| \leq v^{\max} = \frac{1}{10}$ for $i \in \{1,2\}$. Dubois et al. show in [14] that a triangular fuzzy number with possibility density $\pi_{\widetilde{V}_i}(v) = 1 - \frac{|v|}{v^{\max}}$ for $|v| \leq v^{\max}$ and zero outside induces a maximally specific possibility distribution which is consistent with all admissible probability distributions of $\widetilde{V}_i$. See Figure 2 for a visualization of the α-cuts of $\widetilde{V}_1$ and $\widetilde{V}_2$. Considering their independent aggregation, their joint vector $\widetilde{\mathbf{V}}$ is described by

$$\mathbf{A}_{\widetilde{\mathbf{V}}} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ -1 & 0 \\ 0 & -1 \end{bmatrix}, \quad \mathbf{b}_{\widetilde{\mathbf{V}}} = v^{\max}\sqrt{1-\alpha} \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \quad \mathbf{C}_{\widetilde{\mathbf{V}}} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \text{and} \quad \mathbf{d}_{\widetilde{\mathbf{V}}} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \qquad \forall\, \alpha \in [0,1].$$

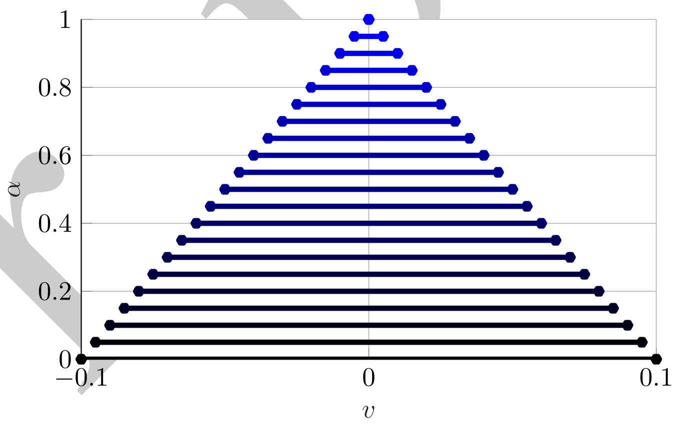

Figure 2: α-cuts of the triangular fuzzy numbers $\widetilde{V}_1$ and $\widetilde{V}_2$.

The measurement noise $\widetilde{W}$, is assumed to follow a probability distribution with mean $\mu = 0$ and variance $\sigma^2 = 1$. Further information about additional moments is not provided. In [15], it is argued that, by means of the Chebychev inequality, a suitable representation of such knowledge is given by the possibility density

$$\pi_{\widetilde{W}}(w) = \min\left(1, \frac{\sigma^2}{(\mu - w)^2}\right) \qquad \forall\, w \in \mathbb{R}. \tag{24}$$

which can be expressed by

$$\mathbf{A}_{\widetilde{W}} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \quad \mathbf{b}_{\widetilde{W}} = \frac{\sigma}{\sqrt{\alpha}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad \mathbf{C}_{\widetilde{W}} = 1 \quad \text{and} \quad \mathbf{d}_{\widetilde{W}} = \mu \qquad \forall\, \alpha \in (0,1].$$

See Figure 3 for a visualization of the α-cuts of $\widetilde{W}$. The support, i.e. the α-cut for $\alpha = 0$, is simply $\mathscr{C}_{\widetilde{W}}^{0} = \mathbb{R}$.

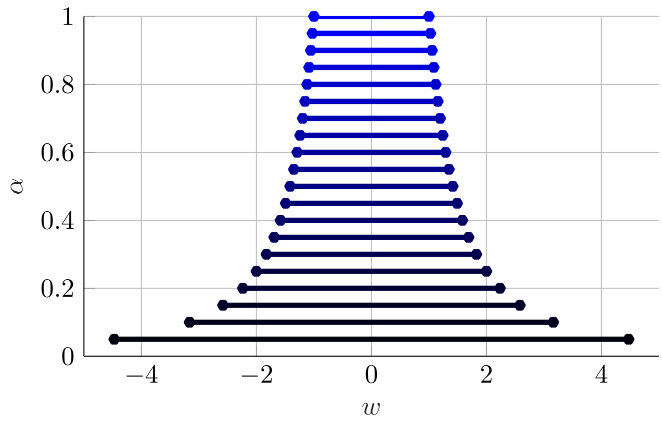

Figure 3: α-cuts of the Chebychev possibility distribution of $\widetilde{W}$.

In the reference simulation, the 'experimental' data y_k^{exp} were generated with consistent probability distributions for $\widetilde{\mathbf{V}}$ and $\widetilde{W}$. In particular, the system noise was generated from uniform probability distributions on the given support $[-\frac{1}{10}, \frac{1}{10}]$ and the measurement noise from a mixture with equal weights of two normal distributions with means $\pm\sqrt{\frac{3}{4}}$ and variance $\frac{1}{4}$ as shown in Figure 4.

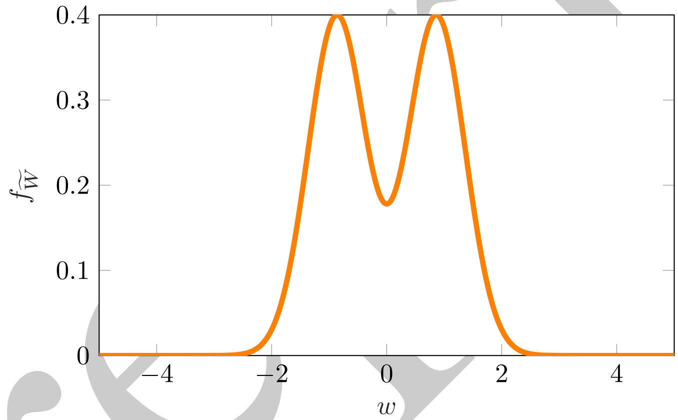

Figure 4: Probability density of $\widetilde{W}$ in the reference simulation.

An exemplary prediction of the output measurements is provided in Figure 5. Evidently, all data points lie within the support of the prediction with a tendency to regions with a higher possibility density. The obtained results can even be used to verify Proposition 3; the empirical probability of the reference measurements y_k^{exp} being contained in the α-cuts $\mathscr{C}_{\widetilde{Y}_k}^{\alpha}$ stays well above the guaranteed value of $1-\alpha$, see Figure 6.

Similarly, the descriptions of process noise and measurement error can be employed to estimate the states x_1 and x_2 from the experimental data. The reference data are estimated accurately, i.e. they tend to lie in high-membership regions, and the order of magnitude of the area of the confidence sets is reasonably small in order to enable meaningful inference. Again, the relative frequency of the confidence sets containing the actual state samples stays well above the guaranteed level of confidence as illustrated in Figure 8, adding further verification to Proposition 4.

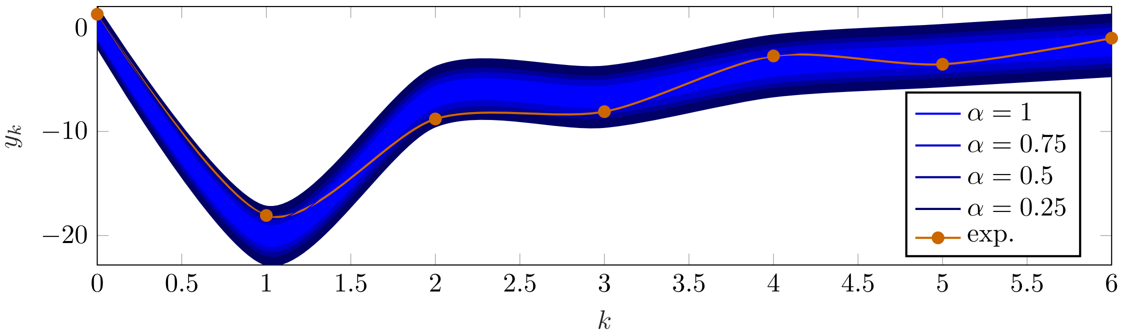

Figure 5: α-cuts of the predicted measurement and simulated reference.

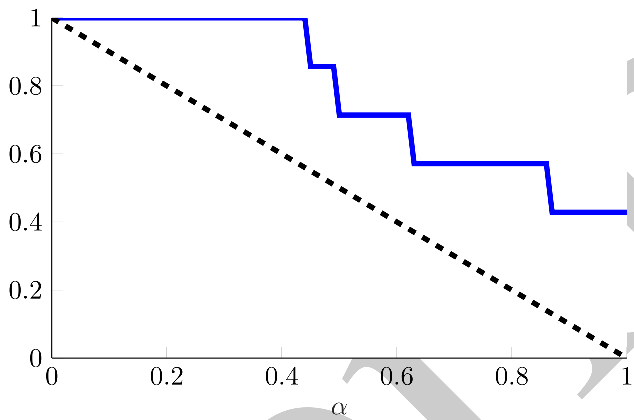

Figure 6: Empirical probabilities of the measurements y_k^{exp} being contained in the α-cuts $\mathscr{C}^{\alpha}_{\tilde{Y}_k}$ compared to the guaranteed probability $1-\alpha$.

6 Conclusions

In this contribution, computationally exact solutions to a selection of forward and inverse problems in possibilistic LTI system analysis have been provided. The representation of convex fuzzy variables by the proposed quadruple description has been shown to be very well-suited for this purpose. It is conceptually easy to implement and allows for an intuitive design of a possibilistic batch filter.

The presented approach is evidently closely related to existing filtering formulations, such as moving horizon estimation [16]. The novelty is that, while the presented approach relies on the propagation of convex sets, i.e. the α-cuts, this approach is able to also carry on statistical information in the form of the size of these sets for different values $\alpha \in [0,1]$ without requiring a precise statistical modeling. Hence, it is possible to arrive at more expressive results than purely set-based techniques, but it does not force the practitioner to specify unique, yet unwarranted, probability distributions thereon. It can, furthermore, serve as a reference solution for a more efficient (particle) filter, which has not yet been developed, with a recursive prediction and updating formulation similar to set-membership filtering approaches, refer e.g. to [17]. However, this requires the derivation of a sensible procedure for the aggregation of confidence sets.

Other perspectives with respect to this contribution, which could prove to be of interest in the future, include the extension to to non-linear system models and the consideration of the inferred information in decision-making and the connection to results concerning robust optimization and control in a possibilistic context [18].

Acknowledgements

The authors gratefully acknowledge the support provided by the German Research Foundation (DFG) in the framework of the research project no. 319924547 (HA2798/9-1).

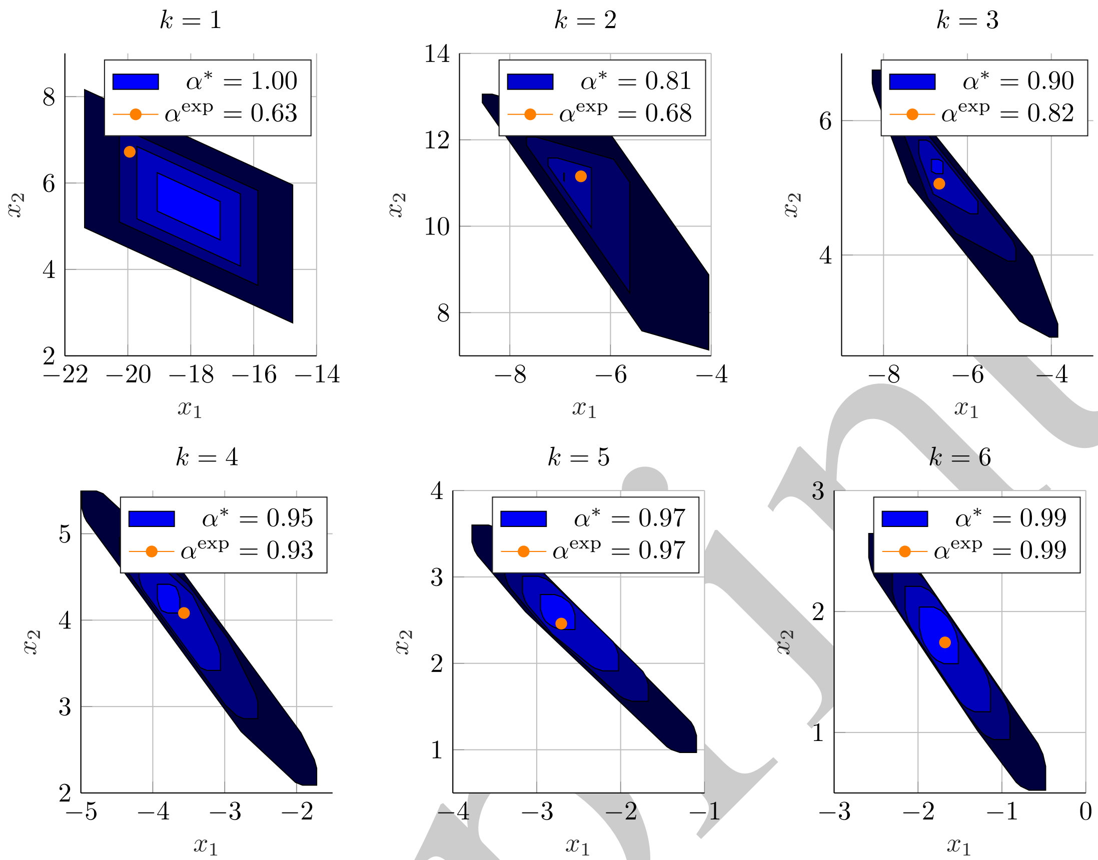

Figure 7: Confidence sets of the past system states (■) compared to the simulated reference (●).

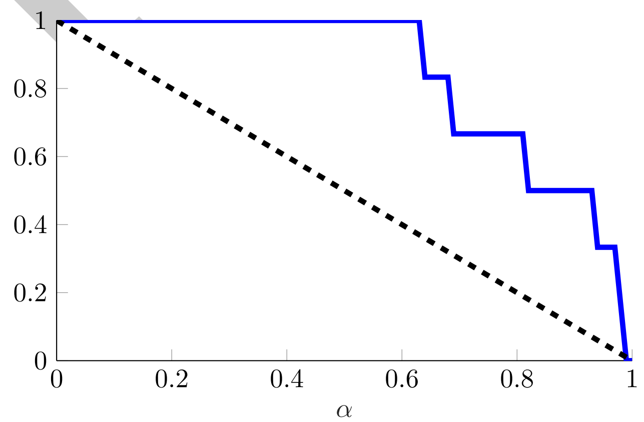

Figure 8: Relative frequency of the reference states $\mathbf{x}_k$ being contained in the α-cuts $\mathscr{C}^{\alpha}_{\tilde{X}_k}$ compared to the guaranteed confidence level.

References

[1] Sigurd Skogestad and Ian Postlethwaite. *Multivariable Feedback Control: Analysis and Design*, volume 2. Wiley New York, 2007.

[2] R. E. Kalman. A new approach to linear filtering and prediction problems. *Journal of Fluids Engineering*, 82(1):35–45, 03 1960.

[3] Wilbur Langson, Ioannis Chryssochoos, SV Raković, and David Q Mayne. Robust model predictive control using tubes. *Automatica*, 40(1):125–133, 2004.

[4] Didier Dubois and Henri Prade. *Possibility Theory: An Approach to Computerized Processing of Uncertainty*. Springer Science & Business Media, 2012.

[5] Thomas Augustin, Frank PA Coolen, Gert De Cooman, and Matthias CM Troffaes. *Introduction to Imprecise Probabilities*. John Wiley & Sons, 2014.

[6] Didier Dubois and Henri Prade. When upper probabilities are possibility measures. *Fuzzy Sets and Systems*, 49(1):65–74, 1992.

[7] Inés Couso, Susana Montes, and Pedro Gil. The necessity of the strong α-cuts of a fuzzy set. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 9(02):249–262, 2001.

[8] Michael Hanss. *Applied Fuzzy Arithmetic: An Introduction with Engineering Applications*. Springer, 2005.

[9] Sébastien Destercke, Didier Dubois, and Eric Chojnacki. A consonant approximation of the product of independent consonant random sets. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 17(06):773–792, 2009.

[10] Dominik Hose and Michael Hanss. Possibilistic calculus as a conservative counterpart to probabilistic calculus. *Mechanical Systems and Signal Processing*, 133:106290, 2019.

[11] Lotfi A Zadeh. The concept of a linguistic variable and its application to approximate reasoning - I. *Information Sciences*, 8(3):199–249, 1975.

[12] Cédric Baudrit, Didier Dubois, and Dominique Guyonnet. Joint propagation and exploitation of probabilistic and possibilistic information in risk assessment. *IEEE Transactions on Fuzzy Systems*, 14(5):593–608, 2006.

[13] Dominik Hose and Michael Hanss. Consistent inverse probability and possibility propagation. In *Proceedings of the 2019 Conference of the International Fuzzy Systems Association and the European Society for Fuzzy Logic and Technology*. Atlantis Press, 2019.

[14] Didier Dubois, Laurent Foulloy, Gilles Mauris, and Henri Prade. Probability-possibility transformations, triangular fuzzy sets, and probabilistic inequalities. *Reliable Computing*, 10(4):273–297, 2004.

[15] Cédric Baudrit and Didier Dubois. Practical representations of incomplete probabilistic knowledge. *Computational Statistics & Data Analysis*, 51(1):86–108, 2006.

[16] Christopher V Rao, James B Rawlings, and David Q Mayne. Constrained state estimation for nonlinear discrete-time systems: Stability and moving horizon approximations. *IEEE Transactions on Automatic Control*, 48(2):246–258, 2003.

[17] Pei Hua Leong and Girish N Nair. Set-membership filtering using random samples. In *2016 19th International Conference on Information Fusion (FUSION)*, pages 1087–1094. IEEE, 2016.

[18] Dominik Hose, Markus Mäck, and Michael F. Hanss. Robust Optimization in Possibility Theory. *ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part B: Mechanical Engineering*, 5(4), 06 2019.