Criteria stochastic filtering of max-plus discrete event systems with bounded random variables ⋆

Guilherme Espindola-Winck * **Laurent Hardouin** *
Mehdi Lhommeau * **Rafael Santos-Mendes** **

* *Laboratoire Angevin de Recherche en Ingénierie des Systèmes, Université d'Angers, LARIS, Polytech Angers, 49000 Angers, France (e-mails: {guilherme.espindolawinck;laurent.hardouin; mehdi.lhommeau}@univ-angers.fr)*
** *Department of Computer Engineering and Industrial Automation - DCA, School of Electrical and Computer Engineering, State University of Campinas - UNICAMP - 13083-970 Campinas, SP, Brazil (e-mail: rafael@dca.fee.unicamp.br).*

Abstract: A wide range of Discrete Event Systems (DES) such as manufacturing systems, telecommunications networks, transportation networks, and parallel computing, etc., can be modelled as max-plus systems in which maximization and addition are the main operations. In this paper, we use a nonlinear approach to deal with the error-estimation of nondeterministic max-plus systems with bounded random variables. This estimation is carried out following the ideas of the stochastic filtering theory for classical time-driven dynamic systems. The probability densities for the entries of the system matrices are assumed to be known, and a prediction-correction filtering scheme is used to compute the estimated state. The filtering algorithm is based on the minimization of a criterion, which is capable to evaluate the estimation-error, and to deal with the trade-off between prediction and measurement.

Keywords: Timed event graphs, Max-plus systems, Nonlinear stochastic filtering, Constraint Satisfaction Problem.

1. INTRODUCTION

This paper proposes a nonlinear approach to deal with the estimation of a subclass of Discrete Event Systems (DES). DES are a subclass of dynamic systems whose dynamics are event-driven, i.e., the state evolution depends entirely on the occurrence of asynchronous discrete events over time. These systems are used to model, e.g., manufacturing systems, telecommunications networks, transportation networks, and parallel computing, i.e., systems that are very difficult to describe in terms of physical equations, that are a large set of mathematical tools for control engineers. Hence, to describe their behavior, more relevant theoretical setting are considered, among them the following can be cited: languages and automata, Markov chain and Petri nets, the reader is invited to consult Cassandras and Lafortune (1999) for an overview.

Among DES, a particular class involving synchronization and delay phenomena can be modelled as max-plus systems. These systems are essentially nonlinear, but appear to be linear over idempotent semirings (often called dioids). This class of DES can be represented graphically, depicted by Timed Event Graphs (TEG). A TEG is a timed Petri net in which each place admits only one upstream transition and one downstream transition. Taking advantage of the linearity property over dioids, several authors have developed methods to estimate the system states (Hardouin et al., 2010; Loreto et al., 2010; Gonçalves et al., 2019). However, if the system is with uncertain parameters, some alternative methods can be considered in order to take advantage of the knowledge about the characteristics of this uncertainty (Cândido et al., 2013; Farahani et al., 2017; Cândido et al., 2013; Cândido et al., 2018; Mendes et al., 2019; Candido et al., 2020).

Recently, Mendes et al. (2019) have proposed a stochastic filtering algorithm for max-plus systems with bounded random variables. The algorithm proposed is two-fold: a prediction phase based on calculation of mathematical expectation w.r.t. the transition equation and a correction phase with a procedure based on Constraint Satisfaction Problem (Jaulin et al., 2001) that takes into account the measurement to correct the prediction estimation.

Contribution: The aim of this paper is to propose an improvement in the correction phase of the Mendes et al., 2019, Algorithm 3. The new algorithm that we propose is based on the minimization of a criterion that allows evaluating the estimation error of the prediction phase and to adjust the importance to be given between the prediction estimation and the corrected estimation, as the quadratic criterion does in Kalman filter for linear continuous time systems.

⋆ This work was supported by the RFI Atlanstic 2020.

This paper is organized as follows. In section 2 we present some mathematical preliminaries on max-plus systems, Timed Event Graphs, interval arithmetic and stochastic computations over max-plus systems. Section 3 recalls the filtering scheme in a prediction-correction set of procedures and presents the new filtering algorithm. In section 4 some simulation results are presented. Finally, section 5 concludes the work and gives some ideas for future improvements.

2. MATHEMATICAL PRELIMINARIES

This section is intended to serve as a review of mathematical concepts to be used throughout this paper.

2.1 Max-plus systems

Let us begin with some notations which are used through this paper. Let $\mathbb{R}_\varepsilon := \mathbb{R} \cup \{\varepsilon\}$ with $\varepsilon = -\infty$ be the set of all real numbers. The set $\mathbb{R}_\varepsilon^n := (\mathbb{R} \cup \{\varepsilon\})^n$ is the n-th fold Cartesian product of $\mathbb{R}_\varepsilon$. Its elements can be thought as points of an affine space, or as vectors which are denoted by bold symbols, for instance $\mathbf{x} = (x_1, \ldots, x_n)^{\mathrm{t}}$ and $\mathbf{0} = (0, \ldots, 0)^{\mathrm{t}}$. The notation $\mathbf{x} \leq \mathbf{y}$ denotes the usual partial order on $\mathbb{R}^n$, i.e., for $\mathbf{x}, \mathbf{y} \in \mathbb{R}_\varepsilon^n$ we have $\mathbf{x} \leq \mathbf{y} \Leftrightarrow x_j \leq y_j,\ \forall j \in \{1, \ldots, n\}$. For $A \in \mathbb{R}_\varepsilon^{n\times p}$, and $B \in \mathbb{R}_\varepsilon^{p\times q}$, we define each element of $A \otimes B \in \mathbb{R}_\varepsilon^{n\times q}$ as $(A \otimes B)_{ij} = \max_{k=1}^{p}\{a_{ik} + b_{kj}\}$.

Max-plus system is mathematically defined as [1]

$$\begin{cases} \mathbf{x}(k) = (A, B) \otimes \mathbf{y}(k),\ \mathbf{y}(k) = (\mathbf{x}^{\mathrm{t}}(k-1), \mathbf{u}^{\mathrm{t}}(k))^{\mathrm{t}}, \\ \mathbf{z}(k) = C \otimes \mathbf{x}(k),\ k \in \mathbb{N}, \end{cases} \tag{1}$$

where $\mathbf{x}(0) \in \mathbb{R}_\varepsilon^n$, $(A, B) \in \mathbb{R}_\varepsilon^{n\times(n+m)}$, $C \in \mathbb{R}_\varepsilon^{q\times(n+m)}$, $\mathbf{x}(k) \in \mathbb{R}_\varepsilon^n$, $\mathbf{z}(k) \in \mathbb{R}_\varepsilon^q$ and $\mathbf{u}(k) \in \mathbb{R}_\varepsilon^m$ are the initial state, the transition-control matrix, the observation matrix, the current state, the control input and the measurement output, respectively.

2.2 Interval arithmetic over $\mathbb{R}_\varepsilon$

Interval arithmetic is presented in Moore and Bierbaum (1979). A (closed) interval $[x]$ is a subset of $\mathbb{R}_\varepsilon$, i.e., $[x] \subset \mathbb{R}_\varepsilon$ of the form $[x] = [\underline{x}, \overline{x}] = \{x \in \mathbb{R}_\varepsilon : \underline{x} \leq x \leq \overline{x}\}$. We denote $\mathbb{IR}_\varepsilon$ the set of intervals of $\mathbb{R}_\varepsilon$. The width of an interval $[x]$ is defined as $\mathtt{w}([x]) = \overline{x} - \underline{x}$.

An interval $[x] \subseteq [y]$ if and only if $\underline{y} \leq \underline{x} \leq \overline{x} \leq \overline{y}$. Similarly, $[x] = [y]$ if and only if $\underline{x} = \underline{y}$ and $\overline{x} = \overline{y}$. Any $x \in \mathbb{R}_\varepsilon$ can be represented by the deprecated interval $[x, x] \in \mathbb{IR}_\varepsilon$ with $\mathtt{w}([x]) = 0$.

The max and + operations over $\mathbb{IR}_\varepsilon$ are defined as: $\max\{[\underline{x}, \overline{x}], [\underline{y}, \overline{y}]\} = [\max\{\underline{x}, \underline{y}\}, \max\{\overline{x}, \overline{y}\}]$ and $[\underline{x}, \overline{x}] + [\underline{y}, \overline{y}] = [\underline{x} + \underline{y}, \overline{x} + \overline{y}]$.

Let $[A] = [\underline{A}, \overline{A}] \in \mathbb{IR}_\varepsilon^{n\times p}$, and $[B] = [\underline{B}, \overline{B}] \in \mathbb{IR}_\varepsilon^{p\times q}$. For instance, the elements of $[A]$ are $[\underline{a}_{ij}, \overline{a}_{ij}]_{1\leq i\leq n, 1\leq j\leq p} \in \mathbb{IR}_\varepsilon$. Thus, we define each element of $[A] \otimes [B] \in \mathbb{IR}_\varepsilon^{n\times q}$ as $([A] \otimes [B])_{ij} = \max_{k=1}^{p}\{[a_{ik}] + [b_{kj}]\}$.

[1] (A, B) is seen as the horizontal concatenation of matrices A and B with same number of rows.

2.3 Timed Event Graphs (TEG)

Timed Event Graphs (TEG) correspond to the subclass of timed Petri nets, in which each place admits only one upstream transition and one downstream transition and all arcs have weight 1. The number of tokens in a place is interpreted as the number of available resources. For instance, the number of tokens at the place p_2 (Figure 1) corresponds to one piece that is waiting to be processed by the machine M_1. A token at p_3 indicates that a resource is being processed by the machine and is no longer waiting at p_2. In addition, each place can be associated with a time delay (p_2 has delay of 3 time unit), making it possible to model the delay of a token, i.e., the sojourn time of this token before being processed by the downstream transition. There is another TEG model in which the delay is associated with each transition. Nevertheless, it is straightforward to show that it is always possible to reduce to the case where only places are delayed (see Murata (1989)) and hereafter we consider this as an assumption.

Example 1. Consider the following TEG

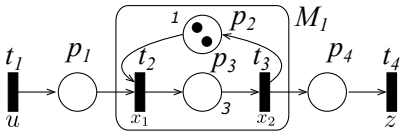

Fig. 1. Modelling of a manufacturing system composed of a single machine

First, we associate each transition with a *dater* $t_i(k) : \mathbb{N} \to \mathbb{R}_\varepsilon$, where $t_i(k)$ denotes the date at which the k-th firing of the transition i occurs. TEG have three distinguished transitions: input transitions that are not affected by the firing of other transitions; output transitions that do not affect the firing of other transitions; and internal transitions that are neither input nor output transitions. Additionally, it is considered that the TEG is operating under the *earliest firing rule*, i.e., every internal and output transition fires as soon as it is enabled. The above TEG is described by the following equations:

$$x_1(k) = \max\{u(k), x_2(k-2) + 1\},\ x_2(k) = x_1(k) + 3,\ z(k) = x_2(k).$$

By substituting $x_1(k)$ into $x_2(k)$ and introducing a new variable $x_3(k) = x_2(k-1)$, the previous equations can be rewritten in shorthand notation as the max-plus system given in (1) as shown below:

$$\begin{pmatrix} x_1(k) \\ x_2(k) \\ x_3(k) \end{pmatrix} = \begin{pmatrix} \varepsilon & \varepsilon & 1 & 0 \\ \varepsilon & \varepsilon & 4 & 3 \\ \varepsilon & 0 & \varepsilon & \varepsilon \end{pmatrix} \otimes \begin{pmatrix} x_1(k-1) \\ x_2(k-1) \\ x_3(k-1) \\ u(k) \end{pmatrix},\ z(k) = \begin{pmatrix} \varepsilon & 0 & \varepsilon \end{pmatrix} \otimes \begin{pmatrix} x_1(k) \\ x_2(k) \\ x_3(k) \end{pmatrix}.$$

In practice, it is necessary to take into account the uncertainties in the modelling. Ideally, one must consider that the delays are nondeterministic variables, i.e., for instance, the delay of p_3 is uncertain and only its probability density function is known.

In view of the stochastic behaviour of TEG, the definition of uncertain max-plus systems arise quite naturally (Cândido et al. (2018); Candido et al. (2020)), considering that the random variables are bounded:

$$\begin{cases} \mathbf{x}(k) = (A(k), B(k)) \otimes \mathbf{y}(k), \\ \mathbf{z}(k) = C(k) \otimes \mathbf{x}(k), \end{cases} \tag{2}$$

where $(A(k), B(k)) \in [(\underline{A}, \underline{B}), (\overline{A}, \overline{B})] \in \mathbb{IR}_{\varepsilon}^{n\times(n+m)}$ and $C(k) \in [\underline{C}, \overline{C}] \in \mathbb{IR}_{\varepsilon}^{q\times n}$ are matrices of independent random variables with finite support and whose entries are mutually independent. For instance, matrices $\underline{C}$ and $\overline{C}$ are respectively the lower and upper bounds of $[C]$, such that $c_{ij} \in [\underline{c}_{ij}, \overline{c}_{ij}] \in \mathbb{IR}_{\varepsilon}$. The same reasoning is applied to the lower and upper bounds of $[(A, B)]$. Moreover, the uncertain systems studied in this work are assumed to be Markovian in the sense that the current state vector depends entirely on event $k-1$ and control inputs at k, and not on the events $k-2, k-3, \ldots$ (the so-called Markov property).

2.4 Stochastic computations over max-plus systems

On the mathematical conditional expectation calculation Consider the generic max-affine equation $z_i = (A \otimes \mathbf{x})_i$ that is similar to those used in (2), precisely

$$z_i = \max_{j=1}^{n}\{a_{ij} + x_j\}, \tag{3}$$

where $a_{ij} \in [\underline{a}_{ij}, \overline{a}_{ij}] \in \mathbb{IR}_{\varepsilon}$, with $i \in \{1, \ldots, q\}$ and $j \in \{1, \ldots, n\}$ and $\mathbf{x} \in \mathbb{R}_{\varepsilon}^{n}$. In the following, a_{ij} will be considered as independent and uniformly distributed random variables. According to these assumptions, the cumulative distribution function (c.d.f.) of a_{ij}, denoted by $F_{a_{ij}}(t)$, is given by:

$$F_{a_{ij}}(t) = \begin{cases} 0 & \text{if } t \leq \underline{a}_{ij}, \\ \dfrac{t - \underline{a}_{ij}}{\overline{a}_{ij} - \underline{a}_{ij}} & \text{if } \underline{a}_{ij} < t \leq \overline{a}_{ij}, \\ 1 & \text{otherwise.} \end{cases} \tag{4}$$

such that $\mathbb{E}[a_{ij}] = (\underline{a}_{ij} + \overline{a}_{ij})/2$.

Let $v = \max_{j=1}^{n}\{w_j\}$ be the value of $(A \otimes \mathbf{0})_i$. If all w_j are independent random variables, then $F_v(t) = P[v \leq t] = P[w_1 \leq t \text{ and } w_2 \leq t \text{ and } w_n \leq t] = \prod_{j=1}^{n} P[w_j \leq t] = \prod_{j=1}^{n} F_{w_j}(t)$. As a corollary, if each w_j is shifted by a constant b_j, i.e., $v = \max_{j=1}^{n}\{w_j + b_j\}$, then $F_v(t) = \prod_{j=1}^{n} F_{w_j}(t - b_j)$.

Summing-up, by analogy, we can calculate the c.d.f. of $(A \otimes \mathbf{x})_i$, denoted by $\mathcal{F}(t)$, as $\mathcal{F}(t) = \prod_{j=1}^{n} F_{a_{ij}}(t - x_j)$, for all $i \in \{1, \ldots, q\}$ with the term $F_{a_{ij}}(t - x_j)$ related to (4). $\mathcal{F}(t)$ is a piece-wise polynomial function [2] of order bounded by n and defined on $[\underline{z}_i, \overline{z}_i] = ([A] \otimes \mathbf{x})_i$ according to interval arithmetic over max and + operations. Hence, the mathematical conditional expectation of z_i given $(A \otimes \mathbf{x})_i$, i.e., $\mathbb{E}[z_i | (A \otimes \mathbf{x})_i] \coloneqq \mathbb{E}[(A \otimes \mathbf{x})_i]$, is given by:

$$\begin{aligned} \mathbb{E}[(A \otimes \mathbf{x})_i] &= z_{i_0} + \int_{z_{i_0}}^{+\infty} (1 - \mathcal{F}(t))\mathrm{d}t, \\ &= \overline{z}_i - \int_{\underline{z}_i}^{\overline{z}_i} \mathcal{F}(t)\mathrm{d}t, \text{ for all } i \in \{1, \ldots, q\}. \end{aligned} \tag{5}$$

Since the piece-wise polynomial c.d.f. $F_{a_{ij}}(t)$ is given for the uniform law, (5) is a continuous and isotonic function on $\mathbf{x}$. Any other technique to calculate $\mathbb{E}[(A \otimes \mathbf{x})_i]$ can be alternatively considered as long as it keeps the properties of continuity and isotony w.r.t. $\mathbf{x}$. We can mention the results proposed in van den Boom and De Schutter (2014) and Farahani et al. (2017) which are possible alternative methods.

The calculations of mathematical conditional expectation of $A \otimes \mathbf{x}$ can be thought as element-wise operations of each $(A \otimes \mathbf{x})_i$.

On the inverse of the mathematical conditional expectation (Mendes et al., 2019, Sec. II) Let $\mathbf{z} = A \otimes \mathbf{x}$ be the max-affine equation that was considered in (3). Consider the following problem: given $\mathbf{z}$, find $\mathbf{x}$ such that $\mathbf{z} = \mathbb{E}[\mathbf{z}|A \otimes \mathbf{x}] \coloneqq \mathbb{E}[A \otimes \mathbf{x}]$, i.e., which state $\mathbf{x}$ leads to $\mathbf{z}$. More formally, we seek to characterize, the set $\chi = \{\mathbf{x} \in \mathbb{R}_{\varepsilon}^{n} : \mathbb{E}[A \otimes \mathbf{x}]\} \subset \mathbb{R}_{\varepsilon}^{n}$.

Contractors (see Jaulin et al. (2001)) are powerful tools to efficiently solve the problem of characterization of the set χ. The operator $\mathcal{C}_{\chi}$ is a contractor for χ if it satisfies $\forall [\mathbf{x}] \in \mathbb{IR}^{n}$ the following properties

$$\begin{aligned} \mathcal{C}_{\chi}([\mathbf{x}]) &\subset [\mathbf{x}] \quad (\textit{contractance}) \text{ and} \\ \mathcal{C}_{\chi}([\mathbf{x}]) \cap \chi &= [\mathbf{x}] \cap \chi \ (\textit{completeness}). \end{aligned}$$

A contractor is said to be minimal if $[\mathbf{x}] \cap \chi = \mathcal{C}_{\chi}([\mathbf{x}])$. In the following, it is assumed that: H1) χ is not empty, H2) $\chi \subset [\mathbf{x}]$, H3) $\mathbb{E}[A \otimes \underline{\mathbf{x}}] \leq \mathbf{z}$, H4) for all $j \in \{1, \ldots, n\}$, $\mathbb{E}[A \otimes (\underline{x}_1, \underline{x}_2, \ldots, \overline{x}_j, \ldots, \underline{x}_{n-1}, \underline{x}_n)^{\mathrm{t}})] \geq \mathbf{z}$. The composition of the following two operations summarize the contractor Ω_{χ} to contract the set χ.

Δ_{ij}^{L} *operator:*

$$\Delta_{ij}^{L}([\underline{\mathbf{x}}, \overline{\mathbf{x}}]) = [\mathbf{x}', \overline{\mathbf{x}}], \tag{6}$$

with

$$\begin{aligned} \mathbf{x}' &= (\underline{x}_1, \underline{x}_2, \ldots, x'_j, \ldots, \underline{x}_{n-1}, \underline{x}_n)^{\mathrm{t}} \\ x'_j &= \sup\{x_j \in [\underline{x}_j, \overline{x}_j]\} \text{ s.t.: } \mathbb{E}[A \otimes \mathbf{x}''] < z_i \\ \mathbf{x}'' &= (\overline{x}_1, \overline{x}_2, \ldots, x_j, \ldots, \overline{x}_{n-1}, \overline{x}_n)^{\mathrm{t}} \end{aligned}$$

Δ_{ij}^{U} *operator:*

$$\Delta_{ij}^{U}([\underline{\mathbf{x}}, \overline{\mathbf{x}}]) = [\underline{\mathbf{x}}, \mathbf{x}'], \tag{7}$$

with

$$\begin{aligned} \mathbf{x}' &= (\overline{x}_1, \overline{x}_2, \ldots, x'_j, \ldots, \overline{x}_{n-1}, \overline{x}_n)^{\mathrm{t}} \\ x'_j &= \inf\{x_j \in [\underline{x}_j, \overline{x}_j]\} \text{ s.t.: } \mathbb{E}[A \otimes \mathbf{x}''] > z_i \\ \mathbf{x}'' &= (\underline{x}_1, \underline{x}_2, \ldots, x_j, \ldots, \underline{x}_{n-1}, \underline{x}_n)^{\mathrm{t}} \end{aligned}$$

The calculation of x'_j in (6) and (7) is a one-dimensional search that can be efficiently performed by the dichotomy method (see Wilde (1964)) as follows: at each step, the search interval initialized with $[\underline{x}_j, \overline{x}_j]$ is divided into two equal intervals. The half containing the solution will be the search interval at the next step. The algorithm stops when the search interval is sufficiently small. We denote Ω_{χ} the contractor obtained from the iterated composition of $2qn$ operators defined above, i.e., $\Omega_{\chi}([\mathbf{x}]) = (\Delta_{11}^{L} \circ \Delta_{11}^{U} \circ \ldots \Delta_{1n}^{L} \circ \Delta_{1n}^{U} \circ \ldots \circ \Delta_{q1}^{L} \circ \Delta_{q1}^{U} \circ \ldots \circ \Delta_{qn}^{L} \circ \Delta_{qn}^{U})([\mathbf{x}])$. It is worth to mention that, if $\mathbb{E}[A \otimes \mathbf{x}''] < z_i$ is not satisfied, then $\Delta_{ij}^{L}([\mathbf{x}])$ is mathematically seen as the identity operator, i.e., $\Delta_{ij}^{L}([\mathbf{x}]) = \mathrm{Id}^{L}([\mathbf{x}]) = [\mathbf{x}]$. Similarly, if $\mathbb{E}[A \otimes \mathbf{x}''] > z_i$ is not satisfied, then $\Delta_{ij}^{U}([\mathbf{x}]) = \mathrm{Id}^{U}([\mathbf{x}]) = [\mathbf{x}]$.

Remark 1. The contractor Ω_{χ} satisfies the contractance, completeness and monotonic properties (see Mendes et al., 2019, Lem. 2 for proofs).

[2] The expression for $\mathcal{F}(t)$ is famous and was originally given in Frechét (1928). Moreover, it is also applicable for all cases where a_{ij} is distributed according to any piece-wise polynomial c.d.f. $F_{a_{ij}}(\bullet)$ (e.g., triangular distribution).

Remark 2. $\Omega_\chi([\mathbf{x}])$ converges to a fixed point, i.e., to an interval $\mathcal{I}$ such that $\Omega_\chi(\mathcal{I}) = \mathcal{I}$. Moreover, $\mathcal{I}$ contains χ (see Mendes et al., 2019, Lem. 3 for proofs).

Generally, the interval $\mathcal{I} = [\underline{\mathbf{x}}^{opt}, \overline{\mathbf{x}}^{opt}]$ that is obtained with Ω_χ contractor is not deprecated and to properly obtain the inverse of the mathematical conditional expectation, one must consider the deprecation procedure described below.

Procedure 1. (Mendes et al., 2019, Sec. IV) For an arbitrary $j \in \{1, \ldots, n\}$, let $a \in [\underline{x}_j, \overline{x}_j]$. Moreover, let $\chi' = \chi \cap \{x_j = a\}$, with $\{x_j = a\}$ a hyperplane in $\mathbb{R}^n_\varepsilon$, and let $\mathcal{I}' = \mathcal{I} \cap \{x_j = a\}$ Thus, $\chi' \subset \mathcal{I}'$ and, thanks to Mendes et al., 2019, Lemma 3, χ' is not empty. In general, $\mathcal{I}'$ is not minimal (Remark 2) and the contraction algorithm can be called once again in order to obtain the minimal interval containing χ'.

The previous procedure can be recursively repeated until the minimal interval is reduced to one point (all components are fixed) that necessarily belongs to χ. The remaining question is: which component should be fixed at each step and to which value? To answer this, consider now that one holds a *guess value* (obtained somehow) for a particular $\mathbf{x}$, formally $\mathbf{x}^0$. We shall look for a point $\mathbf{x} \in \chi$ that is the closest to this value, i.e., $\mathbf{x} = \arg\min_{\mathbf{x}\in\chi} \|\mathbf{x} - \mathbf{x}^0\|_\infty$. In general, $\mathbf{x}$ is not unique, i.e., multiple solutions yield the same minimum, and an optimal value for this problem cannot be guaranteed. However, following Mendes et al., 2019, Section IV, a suboptimal heuristic procedure, based on the deprecation method described above, is proposed to solve an alternative optimization problem, stated as follows. As already pointed-out, $\mathcal{I}$, such that $\chi \subset \mathcal{I}$, is the interval resulting from the Ω_χ contractor. Then, we consider the alternative minimization $\mathbf{x}' = \arg\min_{\mathbf{x}\in\mathcal{I}} \|\mathbf{x} - \mathbf{x}^0\|_\infty$, whose optimal solution [3] is given by (8) of Algorithm 1.

Remark 3. The vector $\mathbf{x}'$ not necessarily belongs to χ, but it is useful to determine at each step which component must be deprecated and to which value. Given $\chi \subset \mathcal{I}$ then the following statement holds $\min_{\mathbf{x}\in\chi} \|\mathbf{x} - \mathbf{x}^0\|_\infty \geq \min_{\mathbf{x}\in\mathcal{I}} \|\mathbf{x} - \mathbf{x}^0\|_\infty$.

The Algorithm 1 summarizes this procedure, and it should be noticed the generation of the initial interval $[\mathbf{x}]$ must satisfy H2, i.e., it must contain at least one solution of the problem characterized by the set χ. A simple rule to guarantee this is to chose $\underline{\mathbf{x}}$ such that $\overline{A} \otimes \underline{\mathbf{x}} < \mathbf{z}$ and $\overline{\mathbf{x}}$ such that $\underline{A} \otimes \overline{\mathbf{x}} > \mathbf{z}$ are respected [4].

For the "proof of concept", this inversion procedure has been naively implemented in MATLAB.

3. RECURSIVE FILTERING ALGORITHM REVISITED

3.1 Original method

In Mendes et al. (2019) an alternative Bayesian method is proposed to compute a state estimation in (2), it is based on two-fold recursive equations, where $\hat{\mathbf{x}}(k|k) \coloneqq \hat{\mathbf{x}}_{corr}(k)$ and $\hat{\mathbf{x}}(k|k-1) \coloneqq \hat{\mathbf{x}}_{pred}(k)$, as shown below:

[3] As for $\mathbf{x}$, there exist multiple solutions for $\mathbf{x}'$, but $\min_{\mathbf{x}\in\mathcal{I}} \|\mathbf{x} - \mathbf{x}^0\|_\infty$ is unique.

[4] H4 implies $\mathbb{E}[A \otimes \overline{\mathbf{x}}] \geq \mathbf{z}$ and it must always be respected, otherwise $\overline{\mathbf{x}}$ must be properly modified.

Algorithm 1: Suboptimal solver

Data: $F(A \in [A] \in \mathbb{IR}^{q\times n}_\varepsilon)$ (c.d.f. of each $a_{ij} \in [\underline{a}_{ij}, \overline{a}_{ij}] \in \mathbb{IR}_\varepsilon$), $\mathbf{z} \in \mathbb{R}^q_\varepsilon$ and $\mathbf{x}^0 \in \mathbb{R}^n_\varepsilon$.
Result: $\mathbf{x} = \texttt{Inv}(\mathbf{z}, \mathbf{x}^0)$
generate $[\underline{\mathbf{x}}, \overline{\mathbf{x}}]$; $bool \leftarrow true$;
while *bool* **do**
  $[\underline{\mathbf{x}}, \overline{\mathbf{x}}] \leftarrow \Omega_\chi([\underline{\mathbf{x}}, \overline{\mathbf{x}}])$; // Contractor
  $bool \leftarrow \underline{\mathbf{x}} \neq \overline{\mathbf{x}}$;
  if *bool* **then**
    /* Heuristic fixation of the $j-$th coordinate of $[\underline{\mathbf{x}}, \overline{\mathbf{x}}]$ based on $\mathbf{x}^0$. */
    foreach $j \in \{1, \ldots, n\}$ **do**

$$x'_j \leftarrow \begin{cases} \underline{x}_j & \text{if } x^0_j \leq \underline{x}_j \\ x^0_j & \text{if } \underline{x}_j < x^0_j < \overline{x}_j \\ \overline{x}_j & \text{otherwise.} \end{cases} \tag{8}$$

      /* $\mathbf{x}' = \arg\min_{\mathbf{x}\in[\underline{\mathbf{x}},\overline{\mathbf{x}}]} \left\|\mathbf{x} - \mathbf{x}^0\right\|_\infty$ */
    end
    $j \leftarrow \arg\max_{j\in\{1,\ldots,n\}} |x'_j - x^0_j|$;
    /* Deprecation */
    $\underline{x}_j \leftarrow x'_j$; $\overline{x}_j \leftarrow x'_j$
  end
end
return $\underline{\mathbf{x}}$; // Notice that: $\underline{\mathbf{x}} = \overline{\mathbf{x}}$

Prediction phase ($\hat{\mathbf{x}}_{pred}(k) = \mathbb{E}[\mathbf{x}(k)|\hat{\mathbf{x}}_{corr}(k-1)]$)

$$\hat{\mathbf{x}}_{pred}(k) = \mathbb{E}[(A(k), B(k)) \otimes (\hat{\mathbf{x}}^{\mathrm{t}}_{corr}(k-1), \mathbf{u}^{\mathrm{t}}(k))^{\mathrm{t}}]. \tag{9}$$

Correction phase ($\mathbf{z}(k) = \mathbb{E}[\mathbf{z}(k)|\hat{\mathbf{x}}_{corr}(k)]$)

$$\hat{\mathbf{x}}_{corr}(k) = \arg\min_{\mathbf{x}} \|\mathbf{x} - \hat{\mathbf{x}}_{pred}(k)\|_\infty , \tag{10a}$$

$$\text{s.t. } \mathbf{z}(k) = \mathbb{E}[C(k) \otimes \mathbf{x}]. \tag{10b}$$

This method is not based on the explicit calculation of the posterior probability density function (p.d.f.) of the state given the measurements, since it is an alternative to the Bayesian filter. Nevertheless, it is claimed by the authors the analogy with the classical maximum likelihood estimator and the fact that it inverts, in some sense, the direct estimation of the measurement given the state. Furthermore, the above equations are implemented by Algorithm 2, but unfortunately with over-optimism, since the corrected estimation $\hat{\mathbf{x}}_{corr}(k)$ must respect the condition that $\mathbf{z}(k) = \mathbb{E}[C(k) \otimes \hat{\mathbf{x}}_{corr}(k)]$, which is not necessary in Bayesian theory.

Algorithm 2: Filtering algorithm for state estimation in (2)

Data: $F((A(k), B(k)) \in [(A, B)] \in \mathbb{IR}^{n\times(n+m)}_\varepsilon)$, $F(C(k) \in [C] \in \mathbb{IR}^{q\times n}_\varepsilon)$, and $\mathbf{z}(k) \in \mathbb{R}^q_\varepsilon$.
Result: $\hat{\mathbf{x}}_{corr}(k) = \texttt{Filter}(\mathbf{z}(k), \hat{\mathbf{x}}_{corr}(k-1))$
$\hat{\mathbf{x}}_{pred}(k) \leftarrow \mathbb{E}[(A(k), B(k)) \otimes (\hat{\mathbf{x}}^{\mathrm{t}}_{corr}(k-1), \mathbf{u}^{\mathrm{t}}(k))^{\mathrm{t}}]$; // (5)
$\hat{\mathbf{x}}_{corr}(k) \leftarrow \texttt{Inv}(\mathbf{z}(k), \hat{\mathbf{x}}_{pred}(k))$; // Algorithm 1
return $\hat{\mathbf{x}}_{corr}(k)$

3.2 Contribution

We recall that in Kalman filter theory, the gain is the weight given to the measurements and current-state estimate, and can be "tuned" to achieve a particular performance. If the Kalman gain is large that means error in the measurement is small which means that new data put in can now very quickly get us to the true value, and therefore we will reduce the error in the estimate and vice versa. Then, we propose to improve the original filtering algorithm for max-plus systems (Algorithm 2) by introducing a criterion in the spirit of the Kalman filtering gain. The optimal gain matrix of the Kalman filter is often derived by minimizing the trace of the posterior covariance matrix (Jazwinski, 1970). By analogy, we will define a criterion J to be minimized,

allowing to take into account the trade-off between the noise in the prediction estimation and in the measurement.

The criterion J
From (9) and (10) the following information is known:

- The state corrected estimation of $\mathbf{x}(k)$ at $k-1$, formally $\hat{\mathbf{x}}_{corr}(k-1)$ (prior knowledge of state);
- The current prediction estimation for the state $\mathbf{x}(k)$ from the available state estimation at $k-1$, formally $\hat{\mathbf{x}}_{pred}(k) = \mathbb{E}[(A(k), B(k)) \otimes (\hat{\mathbf{x}}^{\mathrm{t}}_{corr}(k-1), \mathbf{u}^{\mathrm{t}}(k))^{\mathrm{t}}]$;
- The measurement output $\mathbf{z}(k)$ obtained at k.

For instance, let $\mathbf{z}_0$ be defined as the output prediction estimation, i.e., $\mathbf{z}_0 = \mathbb{E}[C(k) \otimes \hat{\mathbf{x}}_{pred}(k)]$. Consider that, $\mathbf{z}_0$ will be used as input of the inversion procedure, formally $\mathbf{x}_0 = \mathtt{Inv}(\mathbf{z}_0, \hat{\mathbf{x}}_{pred}(k))$. Based on this, it is straightforward to see that $\mathbf{x}_0 = \hat{\mathbf{x}}_{pred}(k)$. However, if we define $\mathbf{z}_1$ as a convex combination between $\mathbf{z}(k)$ and $\mathbf{z}_0$, formally $\mathbf{z}_1 = \beta\mathbf{z}(k) + (1-\beta)\mathbf{z}_0$ where $\beta \in [0,1]$, then this artificial measurement should be used also as input of the same inversion procedure, precisely $\mathbf{x}_1 = \mathtt{Inv}(\mathbf{z}_1, \hat{\mathbf{x}}_{pred}(k))$.

The pair $(\mathbf{x}_1, \mathbf{z}(k))$ as defined before is the input of the following heuristic criterion that aims at evaluating for each k the trade-off between the noise in the prediction estimation and in the measurement:

$$\begin{aligned}\mathsf{J}(\mathbf{x}_1, \mathbf{z}(k)) = \max\{&\|P^{-1}_{dyn}(\mathbf{x}_1 - \hat{\mathbf{x}}_{pred}(k))\|_\infty, \\ &\|P^{-1}_{obs}(\mathbf{z}(k) - \mathbb{E}[C(k) \otimes \mathbf{x}_1])\|_\infty\},\end{aligned} \tag{11}$$

where $P_{dyn} = \mathrm{diag}(\alpha_1, \ldots, \alpha_n)$ and $P_{obs} = \mathrm{diag}(\gamma_1, \ldots, \gamma_q)$ are weighting matrices. The first part of the criterion reflects how the estimation impacts the prediction estimation obtained in (9). The second part reflects how the estimation impacts the measurement.

A key point is to ensure that the criterion above is verified. For this reason, the parameters $\alpha_1, \ldots, \alpha_n$ must be associated with the variability of the components of the vector $\hat{\mathbf{x}}_{pred}(k)$. The inverse of these parameters is therefore an indicator of the reliability of $\hat{\mathbf{x}}_{pred}(k)$. A rule of thumb for tuning the parameter α_i for all $i \in \{1, \ldots, n\}$ is given by $\alpha_i = \frac{1}{n+m}\sum_{j=1}^{n+m} \mathtt{w}([h_{ij}])$ where $[H] = [(A, B)]$. Similarly, the parameters $\gamma_1, \ldots, \gamma_q$ must be associated with the reliability of the components of $\mathbb{E}[C(k) \otimes \mathbf{x}_1]$. The variability of these variables depends on the vector $\mathbf{x}_1$, and in principle, the parameter γ_i must depend on $\mathbf{x}_1$. In order to avoid this dependence, although it is approximate, a possible rule of thumb for γ_i for all $i \in \{1, \ldots, q\}$ is given by $\gamma_i = \frac{1}{n}\sum_{j=1}^{n} \mathtt{w}([c_{ij}])$. Hence, P_{dyn} and P_{obs} are calculated in an offline phase and remain the same throughout the filtering loop. Nevertheless, it is possible to define other suitable values for α_i and γ_i, possibly online, but it is not the subject of the present work.

It should be noted that:

- if $\mathbf{z}_1 = \mathbf{z}_0$, then the first part of the criterion, precisely $P^{-1}_{dyn}(\mathbf{x}_1 - \hat{\mathbf{x}}_{pred}(k))$ is null.
- if $\mathbf{z}_1 = \mathbf{z}(k)$ (this is the case of (10)), then the second part of the criterion, precisely, $P^{-1}_{obs}(\mathbf{z}(k) - \mathbb{E}[C(k) \otimes \mathbf{x}_1])$ is null.

The modified correction equation is also a constrained optimization problem, as shown below, and the main idea is to choose $\hat{\mathbf{x}}_{pred}(k)$ as the value of $\mathbf{x}_1$ that minimizes (11).

$$\begin{aligned}&\hat{\mathbf{x}}_{corr}(k) = \arg\min_{\mathbf{x}_1} \mathsf{J}(\mathbf{x}_1, \mathbf{z}(k)) \\ &\text{s.t. } \mathbf{z}_1 = \mathbb{E}[C(k) \otimes \mathbf{x}_1]\end{aligned} \tag{12}$$

Based on this, the search for $\mathbf{z}_1$ will be constrained to the convex combination of $\mathbf{z}(k)$ and $\mathbf{z}_0$. Hence, $\mathsf{J}(\mathbf{x}_1, \mathbf{z}(k))$ is reinterpreted by $\mathsf{J}(\beta)$ where $\beta \in [0,1]$. Any one-dimensional search method can be used to find a local minimum value of the objective function above. We summarize (9) together with (12) given by the following filtering Algorithm 3.

Algorithm 3: Criteria filtering algorithm for state estimation in (2)

Data: $F((A(k), B(k)) \in [(A, B)] \in \mathbb{IR}^{n\times(n+m)}_{\varepsilon})$,
$F(C(k) \in [C] \in \mathbb{IR}^{q\times n}_{\varepsilon})$, $P_{dyn} \in \mathbb{R}^{n\times n}_{\varepsilon}$, $P_{obs} \in \mathbb{R}^{q\times q}_{\varepsilon}$ and
$\mathbf{z}(k) \in \mathbb{R}^{q}_{\varepsilon}$.
Result: $\hat{\mathbf{x}}_{corr}(k) = \mathtt{Filter}(\mathbf{z}(k), \hat{\mathbf{x}}_{corr}(k-1))$
$\hat{\mathbf{x}}_{pred}(k) \leftarrow \mathbb{E}[(A(k), B(k)) \otimes (\hat{\mathbf{x}}^{\mathrm{t}}_{corr}(k-1), \mathbf{u}^{\mathrm{t}}(k))^{\mathrm{t}}]$; // (5)
$\mathbf{z}_0 \leftarrow \mathbb{E}[C(k) \otimes \hat{\mathbf{x}}_{pred}(k)]$; // (5)
/* Objective function. */
$$\mathsf{J}(\beta) = \begin{cases} \mathbf{z}_1 \leftarrow \beta\mathbf{z}(k) + (1-\beta)\mathbf{z}_0 \\ \mathbf{x}_1 \leftarrow \mathtt{Inv}(\mathbf{z}_1, \hat{\mathbf{x}}_{pred}(k)) \\ \mathsf{J}(\mathbf{x}_1, \mathbf{z}(k)) \leftarrow \max\{\|P^{-1}_{dyn}(\mathbf{x}_1 - \hat{\mathbf{x}}_{pred}(k))\|_\infty, \\ \qquad \|P^{-1}_{obs}(\mathbf{z}(k) - \mathbb{E}[C(k) \otimes \mathbf{x}_1])\|_\infty\} \end{cases};$$
/* One-dim minimizer for $\mathsf{J}(\beta)$ with β between 0 and 1. */
$\beta^{opt} \leftarrow \mathtt{1dMinimizer}(\mathsf{J}(\beta), 0, 1)$; // $\arg\min_{\beta} \mathsf{J}(\beta)$ s.t. $\beta \in [0,1]$
$\mathbf{z}_1^{opt} \leftarrow \beta^{opt}\mathbf{z}(k) + (1-\beta^{opt})\mathbf{z}_0$;
$\mathbf{x}_1^{opt} \leftarrow \mathtt{Inv}(\mathbf{z}_1^{opt}, \hat{\mathbf{x}}_{pred}(k))$;
return $\hat{\mathbf{x}}_{corr}(k) = \mathbf{x}_1^{opt}$

The resulting one-dimensional search in the Algorithm 3 is represented by $\beta^{opt} = \mathtt{1dMinimizer}(func, 0, 1)$ and is solved in the simulation section of this work using the `fminbnd` function, which is implemented in MATLAB. Furthermore, it is worth to mention that if the elements of the main diagonal of P^{-1}_{dyn} and P^{-1}_{obs} are 1/0 then `fminbnd` will not run as desired, and thus we replace 1/0 with a sufficient finite value, for instance 10^4.

4. NUMERICAL SIMULATIONS

In this section, two different systems are studied by means of simulations up to the occurrence of $N = 400$ event-firings.

Example 2. Third-Order System: Consider the third-order uncertain max-plus system given by (2), with $(A(k), B(k)) \in [(A, B)] \in \mathbb{IR}^{3\times4}_{\varepsilon}$ and $C(k) \in [C] \in \mathbb{IR}^{1\times3}_{\varepsilon}$:

$$[(A, B)] = \begin{pmatrix} [0,8] & [0,8] & [3,11] & [\varepsilon,\varepsilon] \\ [2,10] & [0,8] & [0,8] & [\varepsilon,\varepsilon] \\ [1,9] & [1,9] & [0,8] & [\varepsilon,\varepsilon] \end{pmatrix}, [C] = \begin{pmatrix} [0,1] \\ [0,1] \\ [\varepsilon,\varepsilon] \end{pmatrix}^{\mathrm{t}}.$$

For instance, the prior knowledge of state $\hat{\mathbf{x}}_{corr}(k-1)$ is initialized with the initial state $\mathbf{x}(0) = \mathbf{0}$.

The analysis of the root-mean-square-error [5] (RMSE) between the estimation provided by Algorithm 2 and the true value of the state is almost equal to the RMSE between the estimation provided by Algorithm 3 and the true state. Table 1 shows the obtained results.

It is important to mention that, for this example, β^{opt} asymptotically converges to 1 for most event-firings, such that $\mathbf{z}_1 = \mathbf{z}(k)$, clearly leading to the same result found using the Algorithm 2.

Example 3. Ninth-Order Flow Shop System: Consider the Flow Shop system modified from Loreto et al. (2010), modelled as a ninth order uncertain max-plus system with three

[5] Notation: $\mathtt{RMSE}(a, b) = \sqrt{\frac{1}{N}\sum_{j=1}^{N} (a(j) - b(j))^2}$.

Table 1. Comparison between the Algorithms 2 and 3 - Example 2.

State i	Alg. 2 $\texttt{RMSE}(\{\mathbf{x}(k)\},\{\hat{\mathbf{x}}^{F1}_{corr}(k)\})_i$	Alg. 3 $\texttt{RMSE}(\{\mathbf{x}(k)\},\{\hat{\mathbf{x}}^{F2}_{pred}(k)\})_i$
1	1.9195	1.7046
2	2.0124	1.9171
3	1.8908	1.8745

directly measured states x_3, x_6, x_8 and no control input (autonomous). Unfortunately, the measurement occurs under high uncertainty. The model for this system is given by (2) where $(A(k), B(k)) \in [(A, B)] \in \mathbb{IR}_{\varepsilon}^{9\times 10} = (A, B) \in \mathbb{R}_{\varepsilon}^{9\times 10}$ and $C(k) \in [C] \in \mathbb{IR}_{\varepsilon}^{3\times 9}$:

$$(A,B) = \begin{pmatrix} \varepsilon & \varepsilon & 4 & \varepsilon & \varepsilon & \varepsilon & 2 & \varepsilon & \varepsilon \\ 1 & \varepsilon & \varepsilon & \varepsilon & \varepsilon & \varepsilon & \varepsilon & 3 & \varepsilon & \varepsilon \\ \varepsilon & 5 & \varepsilon & \varepsilon & \varepsilon & \varepsilon & \varepsilon & \varepsilon & 1 & \varepsilon \\ 4 & \varepsilon & \varepsilon & \varepsilon & \varepsilon & 3 & \varepsilon & \varepsilon & \varepsilon & \varepsilon \\ \varepsilon & 3 & \varepsilon & 1 & \varepsilon & \varepsilon & \varepsilon & \varepsilon & \varepsilon & \varepsilon \\ \varepsilon & \varepsilon & 5 & \varepsilon & 4 & \varepsilon & \varepsilon & \varepsilon & \varepsilon & \varepsilon \\ \varepsilon & \varepsilon & \varepsilon & 4 & \varepsilon & \varepsilon & \varepsilon & \varepsilon & 3 & \varepsilon \\ \varepsilon & \varepsilon & \varepsilon & \varepsilon & 3 & \varepsilon & 5 & \varepsilon & \varepsilon & \varepsilon \\ \varepsilon & \varepsilon & \varepsilon & \varepsilon & \varepsilon & 2 & \varepsilon & 4 & \varepsilon & \varepsilon \end{pmatrix}, [C] = \begin{pmatrix} [\varepsilon,\varepsilon] & [\varepsilon,\varepsilon] & [\varepsilon,\varepsilon] \\ [\varepsilon,\varepsilon] & [\varepsilon,\varepsilon] & [\varepsilon,\varepsilon] \\ [\varepsilon,\varepsilon] & [\varepsilon,\varepsilon] & [0,6] \\ [\varepsilon,\varepsilon] & [\varepsilon,\varepsilon] & [\varepsilon,\varepsilon] \\ [\varepsilon,\varepsilon] & [\varepsilon,\varepsilon] & [\varepsilon,\varepsilon] \\ [\varepsilon,\varepsilon] & [0,6] & [\varepsilon,\varepsilon] \\ [\varepsilon,\varepsilon] & [\varepsilon,\varepsilon] & [\varepsilon,\varepsilon] \\ [0,6] & [\varepsilon,\varepsilon] & [\varepsilon,\varepsilon] \\ [\varepsilon,\varepsilon] & [\varepsilon,\varepsilon] & [\varepsilon,\varepsilon] \end{pmatrix}^{\mathrm{t}}.$$

For instance, the prior knowledge of state $\hat{\mathbf{x}}_{corr}(k-1)$ is initialized with the initial state $\mathbf{x}(0) = \mathbf{0}$.

The model is not disturbed by any noise, i.e., it is under deterministic behaviour. However, the measurement, represented by the interval matrix $[C]$, is under high uncertainty, making all states become noisy. Algorithm 2 will then return a poor corrected estimation of the states, whereas Algorithm 3 is able to retrieve the prediction estimation instead of the corrected one as the reliable estimation of the states (β^{opt} asymptotically converges to 0 for most event-firings).

The analysis of the RMSE between the estimation provided by Algorithm 2 and the true value of the state is greater than the RMSE between the estimation provided by Algorithm 3 and the true state. Table 2 shows the obtained results for the noisy states x_3, x_6 and x_8 only.

Table 2. Comparison between the Algorithms 2 and 3 - Example 3.

State i	Alg. 2 $\texttt{RMSE}(\{\mathbf{x}(k)\},\{\hat{\mathbf{x}}^{F1}_{corr}(k)\})_i$	Alg. 3 $\texttt{RMSE}(\{\mathbf{x}(k)\},\{\hat{\mathbf{x}}^{F2}_{pred}(k)\})_i$
3	1.7230	0.0065
6	1.6432	0.0065
8	1.5082	0.0065

5. CONCLUSIONS

This paper proposes an improvement of the filtering algorithm presented in Mendes et al. (2019). The new filtering strategy for the correction equation uses the same inversion procedure but using an artificial variable generated by a convex combination between output prediction and output measurement, rather than only the output measurement. This algorithm is able to deal with the trade-off between the noise in the measurement and in the prediction. The approach has been illustrated by two examples with an important improvement for the second one because all system noise is restricted to the observation equation, i.e., the prediction estimation is more reliable than the corrected estimation. Future work could focus on the definition of other objective functions that improve the efficiency of the proposed filtering algorithm, and on defining max-plus stochastic computations that exploit the linearity of the max operator over the max-plus semiring, i.e., leading to a max-plus linear stochastic filtering theory that resembles Kalman filter.

REFERENCES

Candido, R.M.F., Hardouin, L., Lhommeau, M., and Santos-Mendes, R. (2020). An algorithm to compute the inverse image of a point with respect to a nondeterministic max plus linear system. *IEEE Transactions on Automatic Control*, 1–1. doi:10.1109/TAC.2020.2998726.

Cândido, R.M.F., Santos-Mendes, R., Hardouin, L., and Maia, C. (2013). Particle filter for max-plus systems. *European Control Conference, ECC 2013*.

Cassandras, C.G. and Lafortune, S. (1999). *Introduction to Discrete Event Systems*. Kluwer Academic Publishers.

Cândido, R.M.F., Hardouin, L., Lhommeau, M., and Santos Mendes, R. (2018). Conditional reachability of uncertain max plus linear systems. *Automatica*, 94, 426 – 435. doi:https://doi.org/10.1016/j.automatica.2017.11.030. URL `http://www.sciencedirect.com/science/article/pii/S0005109817305721`.

Farahani, S., van den Boom, T., and De Schutter, B. (2017). On optimization of stochastic max-min-plus-scaling systems – An approximation approach. *Automatica*, 83, 20–27. doi:10.1016/j.automatica.2017.05.001.

Frechét, M. (1928). Sur la loi de probabilité de l'écart maximum.

Gonçalves, V.M., Maia, C.A., and Hardouin, L. (2019). On max-plus linear dynamical system theory: The observation problem. *Automatica*, 107, 103–111. doi:https://doi.org/10.1016/j.automatica.2019.05.026. URL `https://www.sciencedirect.com/science/article/pii/S0005109819302420`.

Hardouin, L., Maia, C.A., Cottenceau, B., and Lhommeau, M. (2010). Observer design for (max,+) linear systems. *IEEE Trans. on Automatic Control*, 55 - 2, 538 – 543.

Jaulin, L., Kieffer, M., Didrit, O., and Walter, E. (2001). *Applied Interval Analysis*. Springer-Verlag, London.

Jazwinski, A. (1970). *Stochastic Processes and Filtering Theory*, volume 63 of *Mathematics in Science and Engineering*. Academic Press.

Loreto, M.D., Gaubert, S., Katz, R.D., and Loiseau, J. (2010). Duality between invariant spaces for max-plus linear discrete event systems. *SIAM J. on Control and Optimaztion*.

Mendes, R.S., Hardouin, L., and Lhommeau, M. (2019). Stochastic filtering of max-plus linear systems with bounded disturbances. *IEEE Transactions on Automatic Control*, 64(9), 3706–3715. doi:10.1109/TAC.2018.2887353.

Moore, R.E. and Bierbaum, F. (1979). *Methods and Applications of Interval Analysis (SIAM Studies in Applied and Numerical Mathematics) (Siam Studies in Applied Mathematics, 2.)*. Soc for Industrial & Applied Math.

Murata, T. (1989). Petri nets: Properties, analysis and applications. *Proceedings of the IEEE*, 77(4), 541–580. doi:10.1109/5.24143.

van den Boom, T.J. and De Schutter, B. (2014). Analytic expressions in stochastic max-plus-linear algebra. In *53rd IEEE Conference on Decision and Control*, 1608–1613. doi:10.1109/CDC.2014.7039629.

Wilde, D. (1964). *Optimum seeking methods*. Prentice-Hall.