Stochastic Filtering of Max-plus Linear Systems with Bounded Disturbances

Rafael Santos-Mendes, Laurent Hardouin, Mehdi Lhommeau

Stochastic Filtering of Max-plus Linear Systems with Bounded Disturbances

Rafael Santos Mendes, Laurent Hardouin, and Mehdi Lhommeau,

***Abstract*—The objective of this work is to propose a filtering strategy for max-plus linear systems with bounded disturbances without the direct calculation of the *a posteriori* state probability. The strategy is based on the inversion of the expectation of the measure with respect to the state variable. Among the possible solutions, the closest to the prediction is chosen. An algorithm based on Interval Propagation is proposed to solve this problem. Simulations are performed to show the consistence of the proposed methodology with other approaches in the literature.**

***Index Terms*—Discrete Event Dynamics Systems, Idempotent Semirings, Max-Plus Algebra, Dioid, Observer, Stochastic Filtering, State Estimation.**

I. Introduction

Discrete Event Dynamic Systems (DEDS) [1] constitute a class of systems whose dynamics are event driven $i.e.$ the state is modified exclusively by the occurrence of an event. The study of this class of system is important in many engineering areas like computer networks, transport systems, logistic planning, manufacturing systems and many others. Many typical control problems that arise in the theory of Continuous Dynamic Systems have an analogous statement in DEDS theory, like optimal control synthesis [2], [3], model predictive control [4], [5], robustness [6], [7], observer design [8], [9], stochastic filtering [10], [11] etc..

Among the many models developed for the study of DEDS, the approach based on idempotent semirings is the one adopted in this work. Idempotent semiring (or dioids) are algebraic structures whose fundamental operations are appropriate for the modeling of crucial aspects of DEDS, leading to simpler expressions if compared to those obtained by the traditional algebra ($i.e.$ linear algebra and the field of real numbers). The max-plus algebra is an instance of this algebraic structure and gives rise to the class of DEDS known as max-plus linear (MPL) sytems[1] that are the focus of this work. In section II this approach is further detailed.

For every dynamic system, the knowledge of the state (in real time or not) is of paramount importance for its analysis and control. For this reason, the problems related to the design of state observers and of stochastic filters (in the presence of stochastic perturbations) are particularly relevant, as well for continuous system as for DEDS. In the case of MPL systems, the works of DiLoreto et al. [9] and of Hardouin et al. [8], [12] introduce alternatives for the observer design. In both cases, uncertainty with respect to model parameters are admitted in the form of intervals. The former is based on the study of semimodules[2] and the authors consider a duality principle and the characterization of invariant subsemimodule to compute the observer matrices. The disturbances are given thanks to an implicit system depicting the delay assumed to be in a known interval. The estimated state can then be computed thanks to the available measure. An example introduced in this work and revisited in [12] is analyzed in section V. The observer introduced in [8] is founded on the residuation theory [13] with strategies very similar to those of the classical Luenberger observer [14] for systems described by differential equations and leads to an estimation of the state as close as possible, from below, to the real state. The result is the greatest lower bound based exclusively on structural assumptions, $i.e.$ without taking into account the statistical properties of the perturbations. For this reason, this approach is used as reference in section V of this paper.

The filtering problem arises when the dynamic relations within the system ($i.e.$ its state equations) and/or the directly observed outputs of the system are influenced by random variables. The obtention of estimates for the state constitutes the problem of stochastic filtering. In general, the solution for the stochastic filtering problem is the mathematical expectation of the *a posteriori* probability density of the system state given the measures. In the case of linear systems with additive gaussian noise the exact recursive solution of the problem is the very well known Kalman Filter [15], [16]. For non-linear problems for which the linearized model is a good approximation, the Extended Kalman Filter [17] or the Unscented Kalman Filter [18] can be considered. This is not the case for the MPL systems, due to their discrete character and consequent discontinuities.

The Particle Filter [19], [20] uses a particle representation of the probability density of the system state to perform a Monte-Carlo sequential estimation[3] of the state. A particle representation is a set of samples of the variable to be estimated, sampled according to an "importance density" $i.e.$ according to a density similar to $f(x_k|\, x_{k-1})$ (for more details, see [19]). The approach proposed in [21] and [10] uses this filtering technique to produce state estimates for MPL systems. This approach is limited by the numerical difficulties due

R. Santos Mendes is with DCA/FEEC/University of Campinas UNICAMP, SP BRAZIL e-mail: rafael@dca.fee.unicamp.br.

L. Hardouin and M. Lhommeau are with LARIS/ISTIA/University of Angers, FRANCE e-mails: laurent.hardouin@univ-angers.fr and mehdi.lhommeau@univ-angers.fr.

This work has been supported by Fundação de Amparo à Pesquisa no Estado de São Paulo - FAPESP, project number 2015/19384-9.

[1]MPL systems can also be described in other instances of the idempotent semirings.

[2]The semimodules are a generalization of linear spaces for idempotent semirings.

[3]A Monte-Carlo sequential estimation is a technique for the implementation of a Bayesian filter through Monte-Carlo simulations.

to the generation of the particles and by the fact that the lower dimensionality of the measures with respect to the state, introduces an imprecise generation of particles in the state space.

The approach based on stochastic Max-min-plus-scaling (MMPS) systems [22] constitutes a general framework, that encompasses MPL systems, in which the dynamics of discrete event and hybrid systems can be properly described. In [11] this approach is described and it is shown that a number of optimization stochastic problems can be treated by this methodology, including a particular stochastic filtering problem. The problem of determining the mathematical expectation of MMPS functions is central in this approach and in [23] an approximation method is proposed to determine the mathematical expectation of max-affine functions (that are particular cases of MMPS functions). The functions treated in the appendix of this work are max-affine functions and, as will be stated in the Conclusions, the results proposed in [23] can be combined with those presented in this work to obtain more general filtering schemas.

The classical approach for the stochastic filtering problem is briefly reviewed in section II, where the intrinsic difficulties related to the calculation of the *a posteriori* state probability given the measures are stressed. The non-linearities, resulting from the use of the $\max$ operator, turn the recursive calculation of the involved probabilities an intractable problem. The objective of this work is therefore to propose a filtering strategy for MPL systems with bounded disturbances without the direct calculation of the *a posteriori* state probability. As detailed in the next sections, this approach is based on the inversion of of the function $E[z|\, x]$ with respect to x, and an algorithm for this is proposed in section III. The filtering proposition is then presented in section IV followed by some simulation results in section V. In section VI some conclusions are drawn.

II. Max-plus Linear Systems and Filtering Theory

An idempotent semiring $\mathcal{S}$ is an algebraic structure with two internal operations denoted by $\oplus$ and $\otimes$. The operation $\oplus$ is associative, commutative and idempotent, $i.e.$ $a \oplus a = a$. The operation $\otimes$ is associative (but not necessarily commutative) and is right and left distributive with respect to $\oplus$. The neutral elements of $\oplus$ and $\otimes$ are represented by ε and e respectively, and ε is absorbing with respect to $\otimes$ ($\forall a \in \mathcal{S}, \varepsilon \otimes a = a \otimes \varepsilon = \varepsilon$). As in classical algebra, the operator $\otimes$ will be usually omitted in expressions, $a^i = a \otimes a^{i-1}$ and $a^0 = e$. In this algebraic structure, a partial ordering is defined by $a \succeq b \Leftrightarrow a = a \oplus b \Leftrightarrow \quad b = a \wedge b$ (where $a \wedge b$ is the greatest lower bound for a and b). Therefore an idempotent semiring $\mathcal{S}$ is a partially ordered set (see [24], [25] for an exhaustive approach). An idempotent semiring $\mathcal{S}$ is complete if it is closed with respect to the addition of an infinite number of elements and distributive with respect to the addition of an infinite number of elements. Particularly, $\top = \bigoplus_{x \in \mathcal{S}} x$ is the greatest element of $\mathcal{S}$ ($\top$ is called top of $\mathcal{S}$). The set $\overline{\mathbb{R}}_{\max} = \mathbb{R} \cup \{-\infty, +\infty\}$ equipped with the $\max$ operator as addition and the usual addition ($+$) as product is a complete idempotent semiring denoted by $\overline{\mathbb{R}}_{\max}$, with $\varepsilon = -\infty$ and $e = 0$. For matrices A and $B \in \overline{\mathbb{R}}_{\max}^{n \times m}$ it is usual to define: $(A \oplus B)_{ij} = a_{ij} \oplus b_{ij}$ and for matrices $A \in \overline{\mathbb{R}}_{\max}^{n \times m}$ and $B \in \overline{\mathbb{R}}_{\max}^{m \times p}$: $(A \otimes B)_{ij} = \oplus_{k=1}^{m} (a_{ik} \otimes b_{kj})$.

Typically, DEDS present synchronization and concurrence aspects and according to [24] "*synchronization requires the availability of several resources or users at the same time, whereas concurrency appears for instance when, at a certain time, some user must choose among several resources*". From this definition, it is clear that the $\max$ operator arises naturally in synchronization modeling. Taking into account the instant of events occurrence leads to "timed models" whereas "logical models" ([26]) are concerned exclusively with the possible sequences of events and with the conditions that may give rise to them. Hence, idempotent semirings are useful to model the class of Discrete Event Dynamic Systems (DEDS) in which synchronization and delay phenomena are involved.

Consider now the class of timed models in which only synchronization aspects appear (no concurrence aspects). A system trajectory is defined as the sequence of the time instants of the events occurrences. Let an event be labeled as $i \in \{1, \ldots, n\}$, and let $x_i(k)$ represent the instant of the k-th occurrence of this event. Through the appropriate algebraic manipulation and transformation it is possible to model a significant class of DEDS (referred to as max-plus linear systems) as follows:

$$\begin{aligned} x(k) &= Ax(k-1) \oplus Bu(k) \\ z(k) &= Cx(k), \end{aligned} \tag{1}$$

where $u \in (\overline{\mathbb{R}}_{\max})^p$, $z \in (\overline{\mathbb{R}}_{\max})^q$ and $x \in (\overline{\mathbb{R}}_{\max})^n$ are respectively the input (or control variable), the output (or measure) and the state vector. The involved matrices have dimensions $A \in (\overline{\mathbb{R}}_{\max})^{n \times n}$, $B \in (\overline{\mathbb{R}}_{\max})^{n \times p}$, $C \in (\overline{\mathbb{R}}_{\max})^{q \times n}$, and all matricial operations are in $\overline{\mathbb{R}}_{\max}$. Applications of Equation (1) to the modeling and controling of DEDS are found in many engineering areas like manufacturing systems [24], transport systems [27], computer networks [28] and many others.

In many applications the entries of matrices A, B and C are associated to processing or activity times and in a more general framework they must be considered as random variables. It should be remarked that if one does not take this into account, significant tracking error or even unstable behavior of the system may occur [29]. In this work the entries of matrices A, B and C are taken as independent bounded random variables distributed according to piecewise polynomial cumulative distribution functions respectively given by $F(A)$, $F(B)$ and $F(C)$. These distributions are considered known.

The objective of this paper is to propose an algorithm for the stochastic filtering of the perturbed linear max-plus systems described above. The stochastic filtering problem is conceptually very similar to the one concerning continuous dynamic systems [15], [30], [17] and can be formulated as follows. Given a sequence of observations (or measures) $Z_k = \{z(1), \ldots, z(k)\}$ an estimate for the state variable $x(k)$ is desired (the sequence $x(0), \ldots, x(k)$ is not directly measured). In the classical approach the estimate is the mathematical expectation of the random variable $x(k)$ conditioned to the

observations Z_k *i.e.* $\hat{x}(k|k) = E[x_k|\ Z_k]$. This estimate can be recursively calculated ([15], section 6.6) as follows. The probability density $f(x_{k-1}|\ Z_{k-1})$ is assumed to be known and $f(x_k|\ Z_k)$ is obtained from it. This calculation is performed in two phases. The first one, known as "prediction" calculates $f(x_k|\ Z_{k-1})$ from $f(x_{k-1}|\ Z_{k-1})$.

$$f(x_k|\ Z_{k-1}) = \int f(x_k|\ x_{k-1})\ f(x_{k-1}|\ Z_{k-1})\ dx_{k-1} \quad (2)$$

The second phase, referred to as "update" uses the Bayes formula to calculate the probability density $f(x_k|\ Z_k)$ from $f(x_k|\ Z_{k-1})$.

$$\begin{aligned} f(x_k|\ Z_k) &= f(x_k|\ z_k, Z_{k-1}) \\ &= \frac{f(z_k|\ x_k, Z_{k-1})\ f(x_k|\ Z_{k-1})}{f(z_k|\ Z_{k-1})} \\ &= \frac{f(z_k|\ x_k)\ f(x_k|\ Z_{k-1})}{\int f(z_k|\ \xi)\ f(\xi|\ Z_{k-1})\ d\xi} \end{aligned} \quad (3)$$

The above recursions depend essentially on the knowledge of the conditional probability densities $f(x_k|\ x_{k-1})$ and $f(z_k|\ x_k)$ that depend on the dynamical model of the system (Equation (1)) and on the probability densities $f(A)$, $f(B)$ and $f(C)$. Briefly, given x', u, $f(A)$, $f(B)$ and the relation $x = Ax' \oplus Bu$ one important step of the problem is to obtain $f(x)$ ($f(x)$ corresponds to $f(x_k|\ x_{k-1})$ in Equation 2). Similarly, given x, $f(C)$ and the relation $z = Cx$ another important step is the calculation of $f(z)$ ($f(z)$ corresponds to $f(z_k|\ x_k)$ in Equation 3). Although these calculations are numerically feasible, to obtain an analytical form for these probability densities is not trivial, because they involve a very large number of possibilities. This fact turns the calculation of the integrals in Equations (2) and (3) intractable because the mathematical form of the involved probabilities does not remain the same as k evolves. Moreover the results are very sensible to the form of initial probability density $f(x_0)$. In the remaining sections of this paper an alternative method for the state estimation of perturbed max-plus linear systems is proposed, based on the inversion, with respect to the state, of the mathematical expectation of the measure. The methodology is presented in section IV; in section III important preliminary results are developed.

III. Preliminary Results

Consider the function $z = Cx$ written in max-plus algebra, with $C \in \overline{\mathbb{R}}_{\max}^{q\times n}$, $x \in \overline{\mathbb{R}}_{\max}^{n}$ and $z \in \overline{\mathbb{R}}_{\max}^{q}$, where C is a matrix of independent random variables with finite support. Each random variable c_{ij} is assumed to be distributed according to a piecewise polynomial cumulative distribution function (c.d.f.) $F_{ij}(\cdot)$ and matrices $\underline{C}$ and $\overline{C}$ denote respectively their lower and upper bounds. In the following, the expectation $E[z|\ x]$ is calculated.

Recall first that, if X is a random variable with cumulative distribution function (c.d.f.) $F_X(x) \equiv 0$ for all $x \leq x_0$ then[4] $E[X] = x_0 + \int_{x_0}^{\infty}(1 - F_X(x))\,dx$. Based on this, it is straightforward that for two random variables X_1 and X_2 such that, for all x, $F_{X_1}(x) \leq F_{X_2}(x)$, $E[X_1] \geq E[X_2]$. Moreover, if $Z = \max_{j=1}^{n}\{X_j\}$ and X_j are independent random variables, then $F_Z(z) = P[Z \leq z] = P[X_1 \leq z,\ \text{and } X_2 \leq z,\ \text{and } \ldots,\ \text{and } X_n \leq z] = \prod_{j=1}^{n} P[X_j \leq z] = \prod_{j=1}^{n} F_{X_j}(z)$.

For a given x, let z_i, $i = 1, \ldots, q$ be a component of $z = Cx$. Then, $z_i = \max_{j=1}^{n}\{y_{ij}\}$, where $y_{ij} = c_{ij} + x_j$ and c_{ij} is distributed according to $F_{ij}(\cdot)$. The c.d.f. of each y_{ij} is therefore given by $H_{ij}(z) = F_{ij}(z - x_j)$. In the particular case in which c_{ij} is uniformly distributed:

$$H_{ij}(z) = \begin{cases} 0 & \text{if } \ z \leq \underline{c}_{ij} + x_j \\ \frac{z - x_j - \underline{c}_{ij}}{\overline{c}_{ij} - \underline{c}_{ij}} & \text{if } \ \underline{c}_{ij} + x_j < z \leq \overline{c}_{ij} + x_j \\ 1 & \text{if } \ z > \overline{c}_{ij} + x_j. \end{cases} \quad (4)$$

The generalization of the above expression for the case in which c_{ij} is distributed according to any other piecewise polynomial function $F_{ij}(\cdot)$ is straightforward. Therefore, in view of the previous results, the c.d.f of z_i is:

$$G_i(z) = \prod_{j=1}^{n} H_{ij}(z), \quad (5)$$

which is also a piecewise polynomial function that can be directly calculated and integrated resulting in:

$$E[z_i|\,x] = z_{i0} + \int_{z_{i0}}^{\infty}(1 - G_i(z))\,dz, \quad (6)$$

z_{i0} being the inferior bound for z_i.

In the following, two important properties related to $E[z|\,x]$ are proved.

Lemma 1: $E[z|\,x]$ is a continuous and isotonic function of x.

Proof: Continuity with respect to x is immediately deduced from Equations (4), (5) and (6). To prove isotony, if $x^1 \geq x^2$ then for all $j = 1, \ldots, n$, $x_j^1 \geq x_j^2$, therefore in view of Equations (4) and (5), for all $i \in \{1, \ldots, p\}$, $H_{ij}^1 \leq H_{ij}^2$ and consequently $G_i^1(z) \leq G_i^2(z)$ leading to $E[z_i|\,x^1] \geq E[z_i|\,x^2]$. ■

Remarks:

- A large class of random variables can be properly approximated by random variables with piecewise polynomial c.d.f.'s, including those without an upper bound (infinite support). For the max-plus applications, the exigence of a finite lower bound is not in general restrictive.

- The hypothesis of statistical independence between the entries of the matrices A, B and C (Equation 1) is applicable to practical problems in which each one of these random variables typically represent the delay time of a single and independent process. This is the case in many problems in the areas of planning, production,

[4] if $f_X(x)$ is the p.d.f. of X, then $\int_{x_0}^{\infty}(1 - F_X(x))\,dx = \int_{x_0}^{\infty} P[X > x]\,dx = \int_{x_0}^{\infty}\int_{x}^{\infty} f_X(t)\,dt\,dx = \int_{x_0}^{\infty}\int_{x_0}^{t} f_X(t)\,dx\,dt = \int_{x_0}^{\infty} f_X(t)(t - x_0)\,dt = E[X] - x_0$

communication and traffic systems (see for example section 1.2 of [24]). In many other applications, however, to obtain the model expressed by Equation (1), algebraic operations must be performed over the delay times of the independent processes, resulting in matrices A, B and C whose entries are not mutually independent. This question is adressed in the Appendix of this article, where an approximative method to obtain $E[z|\,x]$ is proposed.

- The results presented in the remaining of this paper depend only on the properties proved in Lemma 1. Therefore any other technique to calculate $E[z|\,x]$ in a different statistical scenario, but keeping the properties of continuity and isotony with respect to x, can be adapted to the methodology hereafter presented. Particularly, the results presented in [23] allow the approximation of $E[z|\,x]$ in the case in which the random variables involved are max-affine functions of a set of independent random variables whose distributions are not constrained to be bounded.

The inverse of the function $E[z|\,x]$

Consider the problem: given z^*, find x^* such that $z^* = E[z|\,x^*]$. Let $\chi \subset \overline{\mathbb{R}}^n_{\max}$ be the set of all solutions of this problem inside a given region $[\underline{x}\ \overline{x}] = \{x \in \overline{\mathbb{R}}^n_{\max} / \ \underline{x} \leq x \leq \overline{x}\}$. It is assumed that: *H1)* χ is not empty; *H2)* $E[z|\,\underline{x}] \leq z^*$; *H3)* for all $j \in \{1,\ldots,n\}$, $E[z|\,(\underline{x}_1, \underline{x}_2, \ldots, \overline{x}_j, \ldots, \underline{x}_{n-1}, \underline{x}_n)^T] \geq z^*$. Of course[5], *H3* implies $E[z|\,\overline{x}] \geq z^*$. Given that the solution for the above problem is not in general unique, we will look for the one that is the closest to a given point x^0. The meaning of the point x^0 will be clarified in the next section.

To approach this problem, the technique known as Interval Propagation [31] is used. According to this theory, given an interval containing a set of solutions of the problem, a "contraction operator" $\mathcal{C}$ must satisfy the following properties.

Contractance Property:

$$\mathcal{C}([\underline{x}\ \overline{x}]) \subset [\underline{x}\ \overline{x}]$$

Completeness Property:

$$\mathcal{C}([\underline{x}\ \overline{x}]) \cap \chi = [\underline{x}\ \overline{x}] \cap \chi,$$

with χ defined as before. In particular, by definition, $[\underline{x}\ \overline{x}] \cap \chi = \chi$.

Two contraction operators are proposed for the inversion of the function $E[z|\,x]$, respectively to contract the lower and the upper bound of an interval. The above properties are proved in the sequel. Consider first the following conditions:

Condition L: $\exists$ i,j such that $E[z_i|\,\xi] < z_i^*$, with $\xi = (\overline{x}_1, \overline{x}_2, \ldots, \underline{x}_j, \ldots, \overline{x}_{n-1}, \overline{x}_n)^T$

Condition U: $\exists$ i,j such that $E[z_i|\,\eta] > z_i^*$, with $\eta = (\underline{x}_1, \underline{x}_2, \ldots, \overline{x}_j, \ldots, \underline{x}_{n-1}, \underline{x}_n)^T$

[5]Hypothesis *H3* is necessary to prove lemma 3. It should be noted that, if a solution exists it is straightforward to increase each x_j such that *H3* be satisfied.

If "condition L" is satisfied then there exists at least one point $x'' = (\overline{x}_1, \overline{x}_2, \ldots, x''_j, \ldots, \overline{x}_{n-1}, \overline{x}_n)^T$, with $\underline{x}_j \leq x''_j \leq \overline{x}_j$, such that $E[z|\,x''] = z_i^*$, therefore this condition is sufficient to perform the contraction defined by the following operator:

Lower Contractor:

$$\mathcal{C}^L_{ij}([\underline{x}\ \overline{x}]) = [x'\ \overline{x}] \tag{7}$$

with:

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

Analogously, if "condition U" is satisfied then there exists at least one point $x'' = (\underline{x}_1, \underline{x}_2, \ldots, x''_j, \ldots, \underline{x}_{n-1}, \underline{x}_n)^T$, with $\underline{x}_j \leq x''_j \leq \overline{x}_j$, such that $E[z|\,x''] = z_i^*$, therefore this condition is sufficient to perform the contraction defined by the following operator:

Upper Contractor:

$$\mathcal{C}^U_{ij}([\underline{x}\ \overline{x}]) = [\underline{x}\ x'] \tag{8}$$

with:

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

The calculation of x'_j in Equations (7) and (8) is an one-dimensional search that can be efficiently performed by the dichotomy method [32] as follows. At each step, the search interval (initialized with $[\underline{x}_j\ \overline{x}_j]$) is divided into two equal intervals (producing a dichotomy). The half containing the solution will be the search interval at the next step[6]. The algorithm stops when the search interval is sufficiently small.

Lemma 2: The operators defined by Equations (7) and (8) satisfy the contractance and completeness properties. Moreover they are monotonic, $i.e.$:

$$\text{If } \ [\underline{x}_1\ \overline{x}_1] \subset [\underline{x}_2\ \overline{x}_2] \ \text{ then } \ \mathcal{C}([\underline{x}_1\ \overline{x}_1]) \subset \mathcal{C}([\underline{x}_2\ \overline{x}_2]).$$

Proof: The contractance property is a direct consequence of the definitions. To prove completeness, consider x'', as defined in Equation (7), with $x_j = x'_j$. Clearly $x' \leq x''$ and therefore, from lemma 1, $E[z|\,x'] \leq E[z|\,x'']$. But $E[z|\,x''] < z_i^*$, then $x' \notin \chi$ and, since $E[z|\,x]$ is a continuous function of x (lemma 1), $[\underline{x}\ x'] \cap \chi = \varnothing$. Therefore $[x'\ \overline{x}] \cap \chi = [\underline{x}\ \overline{x}] \cap \chi$. A completely symmetrical argument shows that in Equation (8), $[x'\ \overline{x}] \cap \chi = \varnothing$ and therefore, $[\underline{x}\ x'] \cap \chi = [\underline{x}\ \overline{x}] \cap \chi$. To prove monotonicity, consider the intervals $[\underline{x}^B\ \overline{x}^B] \subset [\underline{x}^A\ \overline{x}^A]$, $i.e.$ $\underline{x}^A \leq \underline{x}^B \leq \overline{x}^B \leq \overline{x}^A$. Let $x''^A = (\overline{x}^A_1, \overline{x}^A_2, \ldots, x'^A_j, \ldots, \overline{x}^A_{n-1}, \overline{x}^A_n)^T$ and $x''^B = (\overline{x}^B_1, \overline{x}^B_2, \ldots, x'^B_j, \ldots, \overline{x}^B_{n-1}, \overline{x}^B_n)^T$ be, as defined in Equation 7, such that $E[z_i|\,x''^A] = E[z_i|\,x''^B] = 0$. Let $x''^C = (\overline{x}^B_1, \overline{x}^B_2, \ldots, x'^A_j, \ldots, \overline{x}^B_{n-1}, \overline{x}^B_n)^T$. Since $x''^A \geq x''^C$,

[6]In the case of the lower contractor, let x_j be the middle point of the search interval and let x'' be defined as in Equation (7). If $E[z_i|\,x''] \geq z_i^*$ then the solution is on the lower half. For the upper contractor the procedure is symmetrical.

$E[z_i|\, x''^B] = E[z_i|\, x''^A] = 0 \geq E[z_i|\, x''^C]$ and therefore $x''^B \geq x''^C$, leading to $x_j'^B \geq x_j'^A$ and $x'^B \geq x'^A$. This proves that $\mathcal{C}_{ij}^L([\underline{x}^B\ \overline{x}^B]) = [x'^B\ \overline{x}^B] \subset \mathcal{C}_{ij}^L([\underline{x}^A\ \overline{x}^A]) = [x'^A\ \overline{x}^A]$. A completely symmetrical argument proves also that $\mathcal{C}_{ij}^U([\underline{x}^B\ \overline{x}^B]) = [\underline{x}^B\ x'^B] \subset \mathcal{C}_{ij}^U([\underline{x}^A\ \overline{x}^A]) = [\underline{x}^A\ x'^A]$. ■

Consider now a generalization of the preceding operators as follows. If conditions L or U are not satisfied, operators $\mathcal{C}_{ij}^L$ and $\mathcal{C}_{ij}^U$ are simply taken as the identity operator (Id$([\underline{x}\ \overline{x}]) = [\underline{x}\ \overline{x}]$). Define the operator:

$$\Omega([\underline{x}\ \overline{x}]) = \mathcal{C}_{11}^L \circ \mathcal{C}_{11}^U \circ \mathcal{C}_{12}^L \circ \mathcal{C}_{12}^U \circ \ldots \mathcal{C}_{ij}^L \circ \mathcal{C}_{ij}^U \circ \ldots \mathcal{C}_{qn}^L \circ \mathcal{C}_{qn}^U([\underline{x}\ \overline{x}]) \tag{9}$$

This operator is the composition of $2 \times n \times q$ previously defined operators and clearly preserves the contractance, completeness and monotonicity properties. Let $\Im$ be the set of all boxes contained by the initial interval $[\underline{x}\ \overline{x}]$ and containing χ, *i.e.* , if $I \in \Im$ then $\chi \subset I \subset [\underline{x}\ \overline{x}]$. Thanks to the properties derived above, $\Im$ constitutes a complete lattice and $\Omega : \Im \to \Im$ is an order preserving function. The Knaster-Tarski theorem [33] can therefore be applied to conclude that the iterative application of operator Ω will converge to a fixed point, that is, to an interval I^* such that $\Omega(I^*) = I^*$. In the sequel it is proved that I^* is the minimal element of $\Im$.

Lemma 3: $I^* = \Omega(I^*)$ is the minimal interval that contains χ.

Proof: From Equation (9), the interval I^* is such that for all i and j the conditions L and U are not satisfied. Let $I^* = [\underline{x}^*\ \overline{x}^*]$. From hypothesis *H3*, the initial interval is such that, for all j, $E[z|\,(\underline{x}_1, \underline{x}_2, \ldots, \overline{x}_j, \ldots, \underline{x}_{n-1}, \underline{x}_n)^T] \geq z^*$. Then, it is true that for all i and j, $E[z_i|\,(\underline{x}_1, \underline{x}_2, \ldots, \overline{x}_j, \ldots, \underline{x}_{n-1}, \underline{x}_n)^T] \geq z_i^*$. From Equation (8) it is clear that this condition is preserved in all upper contractions, therefore, for all i and j, $E[z_i|\,(\underline{x}_1^*, \underline{x}_2^*, \ldots, \overline{x}_j^*, \ldots, \underline{x}_{n-1}^*, \underline{x}_n^*)^T] \geq z_i^*$. However, since condition U is not satisfied for the interval I^*, it is true that for all i and j, $E[z_i|\,(\underline{x}_1^*, \underline{x}_2^*, \ldots, \overline{x}_j^*, \ldots, \underline{x}_{n-1}^*, \underline{x}_n^*)^T] \leq z_i^*$ (see the definition of the condition U). Therefore, for all i and j, $E[z_i|\,(\underline{x}_1^*, \underline{x}_2^*, \ldots, \overline{x}_j^*, \ldots, \underline{x}_{n-1}^*, \underline{x}_n^*)^T] = z_i^*$, *i.e.* , for all j, $(\underline{x}_1^*, \underline{x}_2^*, \ldots, \overline{x}_j^*, \ldots, \underline{x}_{n-1}^*, \underline{x}_n^*)^T \in \chi$. Consider now, two different indices j and $k \in \{1, \ldots, n\}$. Given that χ is a continuous variety in $\overline{\mathbb{R}}_{\max}^n$,
$(\underline{x}_1^*, \underline{x}_2^*, \ldots, \overline{x}_j^*, \ldots, \underline{x}_k^*, \ldots, \underline{x}_{n-1}^*, \underline{x}_n^*)^T \in \chi$, and
$(\underline{x}_1^*, \underline{x}_2^*, \ldots, \underline{x}_j^*, \ldots, \overline{x}_k^*, \ldots, \underline{x}_{n-1}^*, \underline{x}_n^*)^T \in \chi$, then for an arbitrary $x_j \in [\underline{x}_j^*\ \overline{x}_j^*]$ it must exist a value $x_k \in [\underline{x}_k^*\ \overline{x}_k^*]$ such that $(\underline{x}_1^*, \underline{x}_2^*, \ldots, x_j, \ldots, x_k, \ldots, \underline{x}_{n-1}^*, \underline{x}_n^*)^T \in \chi$, i.e, every arbitrary hyperplane $x_j = constant$ intersects χ. This fact shows that if $\Omega(I^*) = I^*$ and I_1 and I_2 are disjoint intervals such that $I_1 \cup I_2 = I^*$ then, $I_1 \cap \chi \neq \varnothing$ and $I_2 \cap \chi \neq \varnothing$. Therefore I^* is minimal. ■

Summing up, Algorithm 1 presented below can be used to determine the minimal interval that contains all the solutions (set χ) contained in the initial interval $[\underline{x}\ \overline{x}]$.

Algorithm 1: Contraction Algorithm

Data: $F(C)$ (c.d.f. of matrix C), z^*
Result: $[\underline{x}\ \overline{x}]$ = Contract($\underline{x}$, $\overline{x}$)

```
continue ← 1 ;
while continue do
    x̲' ← x̲;
    x̄' ← x̄;
    for i = 1 : q do
        for j = 1 : n do
            if cond L then
                x̲_j ← C_ij^L([x̲ x̄]) (Equation 7);
            end
            if cond U then
                x̄_j ← C_ij^U([x̲ x̄]) (Equation 8);
            end
        end
    end
    continue ← (x̲ ≠ x̲' or x̄ ≠ x̄');
end
```

As discussed before, among all the possible solutions $x \in \chi$, we are looking for the one that is the closest to a given point x^0, that is $x^* = \arg\min_{x \in \chi} \|x - x^0\|_\infty$. In the following, a suboptimal procedure, based on the iterative fixation of the components of x is proposed. To approach this, it must be observed that, thanks to lemma 3, if conditions U and L can not be verified, then for all $j \in \{1, \ldots, n\}$ the intersection between the hyperplane $x_j = constant$ ($\in [\underline{x}_j\ \overline{x}_j]$) and χ is not empty. Let $\chi' \subset \chi$ be this intersection. Once fixed the value of x_j (making $\underline{x}_j = \overline{x}_j = x_j$), the preceding algorithm can be restarted and, after convergence of the new run, the smallest interval containing χ' will be obtained. This procedure can be repeated until all the components of x are fixed. The remaining question is: which component should be fixed and to which value. To answer this, let $I^* = [\underline{x}_j^*\ \overline{x}_j^*]$ be the interval resulting from the "contraction algorithm" previously presented and define x' as follows:

$$x' = \arg\min_{x \in I^*} \|x - x^0\|_\infty.$$

A solution[7] for x' is:

$$x'_j = \begin{cases} \underline{x}_j^* & \text{if } x_j^0 < \underline{x}_j^* \\ x_j^0 & \text{if } \underline{x}_j^* \leq x_j^0 \leq \overline{x}_j^* \\ \overline{x}_j^* & \text{if } x_j^0 > \overline{x}_j^* \end{cases} \tag{10}$$

Given that $\chi \subset I^*$, it is clear that:

$$\begin{aligned} \min_{x \in \chi} \|x - x^0\|_\infty &\geq \min_{x \in I^*} \|x - x^0\|_\infty \\ \|x^* - x^0\|_\infty &\geq \|x' - x^0\|_\infty \end{aligned}$$

After each run of the contraction algorithm, in general, both sides of the preceding inequality increase. At the end of the procedure, $\chi = I^*$ because I^* converges to a point, therefore equality is verified. As a consequence, if after each run the right hand side of the above equation remains unchanged, then

[7]There exist multiple solutions for x', but $\min_{x \in I^*} \|x - x^0\|_\infty$ is unique.

an optimal solution is obtained.

Let j^* be such that $|x'_{j^*} - x^0_{j^*}| = \|x' - x^0\|_\infty$. If the j^{th} component is fixed with $j \neq j^*$, after the next run of the contraction algorithm the component x'_{j^*} will be in general modified by the algorithm and this will lead to a value $x_{j^*} \neq x'_{j^*}$ and therefore will increase $\|x' - x^0\|_\infty$. For this reason, after each run of the contraction algorithm, j^* is determined and x_{j^*} is fixed to x'_{j^*}. The contraction algorithm is then run again and so on until a single value of x^* is obtained. As remarked before, if j^* is the same after each run of the contraction algorithm then $\|x' - x^0\|_\infty$ is also the same ($i.e.$ it does not increase) and as a result the obtained x^* is optimal. If j^* is not preserved, the strategy is in general suboptimal. The initial interval must contain at least one solution of the problem. A simple rule to obtain this is to chose $\underline{x}$ such that $\overline{C}\underline{x} < z^*$ and $\overline{x}$ such that $\underline{C}\overline{x} > z^*$. If *H3* is not satisfied, $\overline{x}$ must be properly modified. The proposed procedure is summarized in Algorithm 2.

Algorithm 2: Inversion Algorithm

Data: $F(C)$ (c.d.f. of matrix C)
Result: $x^* = \text{Inv}(z^*, x^0)$
initialize $\underline{x}, \overline{x}$;
$continue \leftarrow 1$;
while *continue* **do**
  $[\underline{x}\ \overline{x}] = \text{Contract}(\underline{x}, \overline{x})$;
  $continue \leftarrow (\underline{x} \neq \overline{x})$;
  if *continue* **then**
    $x' \leftarrow$ (Equation 10);
    $j^* = \arg\max_{j\in\{1,\dots,n\}} |x'_j - x^0_j|$;
    $\underline{x}_{j^*} \leftarrow x'_{j^*}$;
    $\overline{x}_{j^*} \leftarrow x'_{j^*}$;
  end
end
$x^* \leftarrow \underline{x}$;

IV. Stochastic Filtering

In this section a Stochastic Filter for linear max-plus systems, based on the results of the preceding section, is proposed. Consider the dynamic equations proposed in section II rewritten below:

$$x(k+1) = Ax(k) \oplus Bu(k) \quad (11)$$
$$z(k) = Cx(k) \quad (12)$$

As stated before, the independent variable k is the event counter and each of the system variables (x, z and u) are time instances of the event occurrences. Vectors x, z and u are respectively n, q and $p-$dimensional; A, B and C are respectively $n \times n$, $n \times p$ and $q \times n$ random matrices such that its entries are independent random variables distributed according to known piecewise polynomial cumulative distribution functions: $F(A)$, $F(B)$ and $F(C)$.

The filtering problem can be stated as follows. After event k, given a sequence of measured values for z, $Z(k) = \{z^*(1)\dots z^*(k)\}$ determine an estimate for $x(k)$ noted by $\hat{x}(k)$, supposing that an estimate $\hat{x}(0)$ is known at $k = 0$. Traditional methods to obtain $\hat{x}(k)$, however, do not lead to computationally feasible solutions for this problem. The determination of the density of probability $f(x(k)|\, Z(k))$, the basis of the bayesian approach, is very difficult to perform in this case, due to the multiplicity of conditions that must be considered. In the following, an alternative approach is proposed.

For the state trajectory $X(k) = \{x(0), \dots, x(k)\}$ and a given measure sequence $Z(k) = \{z^*(1)\dots z^*(k)\}$, consider first the sequence $\hat{x}(k)$ satisfying:

$$\hat{x}(k) = E[x(k)|\, \hat{x}(k-1)] \quad (13)$$
$$z^*(k) = E[z(k)|\, \hat{x}(k)] \quad (14)$$

On one hand, if matrices A, B and C are deterministic, the preceding equations are also deterministic and if the initial conditions are exact then, trivially, the estimate will coincide with the real trajectory. On the other hand, the sequence $\{\hat{x}(k)\}$ is similar to the classical maximum likelihood estimator in the sense that it is based on the likelihood function $l(x_k) = f(z_k|\, x_k)$. The main difference is that instead of taking the maximum of this function with respect to the state x_k, the estimate $\hat{x}(k)$ (see Equation 14) chooses the value of x_k such that the mean value of the measure z_k given x_k is the actual measure. According to [15] estimators based exclusively on the conditional probability density function $f(z_k|\, x_k)$ are non-bayesian since they do not take into account any prior information about the state. Estimators based on the pdf $f(x_k|\, z_k) = (1/c)\, f(z_k|\, x_k)\, f(x_k)$ (c is a constant, see Equation 3)[8] clearly consider the prior probability density $f(x_k)$. It is important to note that the schema summarized above effectively takes into account the prior data, since it demands the value of $\hat{x}(k)$ to be the same in both equations.

Although conceptually useful, Equations (13) and (14) are not adequate for the direct implementation of a filter, because the existence of an unique solution $\hat{x}(k)$ for both equations is not guaranteed. In the following, an alternative solution that approaches the preceding one is proposed. As usual ([17], [30], [15]), the filtering procedure is divided into two parts: prediction and measure update.

Prediction Equation:

$$\hat{x}(k|\, k-1) = E[x(k)|\, \hat{x}(k-1|\, k-1)] \quad (15)$$

Update Equation:

$$\hat{x}(k|\, k) = \arg\min_x ||\ x - \hat{x}(k|\, k-1)\ ||_\infty \quad (16)$$

$$s.t.\ z^*(k) = E[z(k)|\ x\ ]$$

It can be noted that if $\hat{x}(k|\, k-1) = \hat{x}(k|\, k)$, the schema proposed by Equations (15) and (16) is the same as the one proposed by Equations (13) and (14). From the computational point of view, the prediction $\hat{x}(k|\, k-1)$ can be obtained directly from Equation (6) and the update equation can be solved by the use of Algorithm 2. At each iteration, the

[8]The dependence on the past mesures Z_{k-1} has been ommited for simplicity

filtering procedure can be summarized as in Agorithm 3 below.

Algorithm 3: Filtering Algorithm

Data: $F(A)$, $F(B)$, $F(C)$ (c.d.f.'s of syst. matrices)
Result: $\hat{x}(k|\,k)$ = Filt($z^*(k)$, $\hat{x}(k-1|\,k-1)$)
1 $\hat{x}(k|\,k-1) \leftarrow E[x(k)|\,\hat{x}(k-1|\,k-1)]$ (Equation 6);
2 $\hat{x}(k|\,k) \leftarrow$ Inv($z^*(k)$, $\hat{x}(k|\,k-1)$) (Algorithm 2);

In the next section some simulations are presented.

V. Simulation Results

In this section two different systems are studied by means of simulations. The first one is a third order system with high level noise and such that its dynamic matrix A has no null element (for all i and j, $a_{ij} \neq \varepsilon$). As analyzed later, the main point in this simulation is the comparison between the predicted state and the state estimate. The second simulation concerns a ninth order Flow Shop system, formerly considered by [9]. The main point in this simulation is the comparison between the herein proposed state estimate and the observer proposed by [8]. In both cases only uniformly distributed random variables and only autonomous[9] systems ($B = \varepsilon$) are considered.

Example 1 - Third Order System: Consider the third order autonomous linear system given by Equations (11) and (12) with the following matrices:

$$A = \begin{bmatrix} [e\ 8] & [e\ 8] & [3\ 11] \\ [2\ 10] & [e\ 8] & [5\ 13] \\ [1\ 9] & [1\ 9] & [e\ 8] \end{bmatrix}; B = \begin{bmatrix} \varepsilon \\ \varepsilon \\ \varepsilon \end{bmatrix}; C = \begin{bmatrix} [e\ 1] \\ [e\ 1] \\ \varepsilon \end{bmatrix}' \quad (17)$$

Figure 1 presents a realization of this system obtained by simulation up to the occurrence of 15 events, starting with the exact estimate of the initial state.

The analysis of realizations with a larger amount of events indicate that the root mean square error (RMSE)[10] between the predicted and the true value of the state is significantly greater than the root mean square error between the estimate and the true state. Table I shows the obtained results for simulations up to the occurrence of 400 events.

i	RMSE$(x_i(k), \hat{x}_i(k\|\,k))$	RMSE$(x_i(k), \hat{x}_i(k\|\,k-1))$
1	3.6495	3.9115
2	2.8686	4.0529
3	3.9488	3.9488

Table I
Comparison between predictions and estimates

It is important to observe that for $i = 3$ the predicted and estimated state are the same because the element c_{13} of matrix C is null ($c_{13} = \varepsilon$). Indeed, in this case, the measure $z^*(k)$ does not bring any new information to the estimate $\hat{x}_3(k|\,k)$ since trivially, from Equations (12) and (17), $x_3(k)$ has no influence on $z^*(k)$. In general, this fact constitutes a difference between the approach herein presented and the classical approach based on the *a posteriori* probability density $f\,(x|\,z^*)$. As well, it must be noted that the third column of matrix A is greater than the other two columns, making $x_1(k)$ and $x_2(k)$ dependent on $x_3(k-1)$ which is not directly observed. However, although expressive, the RMSE values observed for $\hat{x}_1(k|\,k)$ and $\hat{x}_2(k|\,k)$ are significantly reduced if the combined range of variation of the entries of the matrices A and C are taken into account.

Example 2 - Ninth Order Flow Shop System: Consider now the Flow Shop system presented in [9] and also analysed in [12], modeled as an autonomous ninth order linear max-plus system with three directly measured states (x_3, x_6 and x_8). Excepted the elements a_{21}, a_{52} and a_{54} all inputs of matrix A are deterministic. The model for this system is given by Equations (11) and (12) with the matrices:

$$A = \begin{bmatrix} \varepsilon & \varepsilon & 4 & \varepsilon & \varepsilon & \varepsilon & 2 & \varepsilon & \varepsilon \\ [1,\ 7] & \varepsilon & \varepsilon & \varepsilon & \varepsilon & \varepsilon & \varepsilon & 3 & \varepsilon \\ \varepsilon & 5 & \varepsilon & \varepsilon & \varepsilon & \varepsilon & \varepsilon & \varepsilon & 1 \\ 4 & \varepsilon & \varepsilon & \varepsilon & \varepsilon & 3 & \varepsilon & \varepsilon & \varepsilon \\ \varepsilon & [3,\ 5] & \varepsilon & [1,\ 3] & \varepsilon & \varepsilon & \varepsilon & \varepsilon & \varepsilon \\ \varepsilon & \varepsilon & 5 & \varepsilon & 4 & \varepsilon & \varepsilon & \varepsilon & \varepsilon \\ \varepsilon & \varepsilon & \varepsilon & 4 & \varepsilon & \varepsilon & \varepsilon & \varepsilon & 3 \\ \varepsilon & \varepsilon & \varepsilon & \varepsilon & 3 & \varepsilon & 5 & \varepsilon & \varepsilon \\ \varepsilon & \varepsilon & \varepsilon & \varepsilon & \varepsilon & 2 & \varepsilon & 4 & \varepsilon \end{bmatrix} \quad (18)$$

$$B = \begin{bmatrix} \varepsilon & \varepsilon & \varepsilon & \varepsilon & \varepsilon & \varepsilon & \varepsilon & \varepsilon & \varepsilon \end{bmatrix}';$$

$$C = \begin{bmatrix} \varepsilon & \varepsilon & e & \varepsilon & \varepsilon & \varepsilon & \varepsilon & \varepsilon & \varepsilon \\ \varepsilon & \varepsilon & \varepsilon & \varepsilon & \varepsilon & e & \varepsilon & \varepsilon & \varepsilon \\ \varepsilon & \varepsilon & \varepsilon & \varepsilon & \varepsilon & \varepsilon & \varepsilon & e & \varepsilon \end{bmatrix}.$$

The goal of this example is to compare the outcomes of the herein proposed filter with those of the observer proposed in [8] for linear max-plus systems. It is based on the Luenberger observer for continuous linear systems and determines, for each k, the greatest lower bound for the state allowed by the current measures. The observer is based strictly on structural considerations (inferior and superior bounds for the matrices) and does not take into account the statistical properties of the uncertainties, $i.e.$, it is a non-deterministic non-stochastic approach.

A realization of this system has been obtained by simulation up to the occurrence of 15 events, starting with the exact estimate of the initial state. The observer proposed in [8] and the filter proposed herein have been simulated and respectively generated the signals $x_i^{obs}(k)$ and $\hat{x}_i(k|\,k)$. The states x_1, x_3, x_4, x_6, x_7, x_8 and x_9, have been perfectly recovered by both the observer and the filter, since they are not perturbed by any noise. The noisy states x_2, and x_5 are depicted in the Figure 2.

The analysis of realizations with a larger amount of events indicate that the mean square error (RMSE) between the observed and the true value of the state is significantly greater than the mean square error between the filter estimate and the true state. Table II shows the obtained results for simulations

[9] The presence of the term Bu in Equation (11) does not change the nature of the problem and has a small effect on the involved calculations.

[10] Notation: RMSE$(x, y) = \sqrt{\frac{1}{N}\sum_{k=1}^{N}(x(k) - y(k))^2}$.

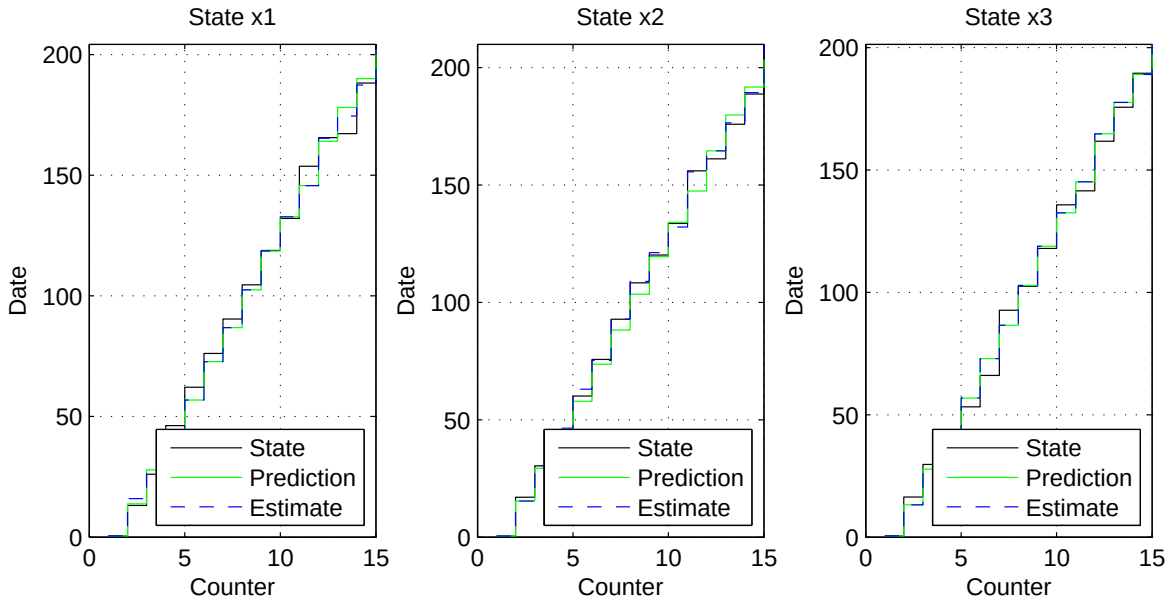

Figure 1. State trajectory for Example 1

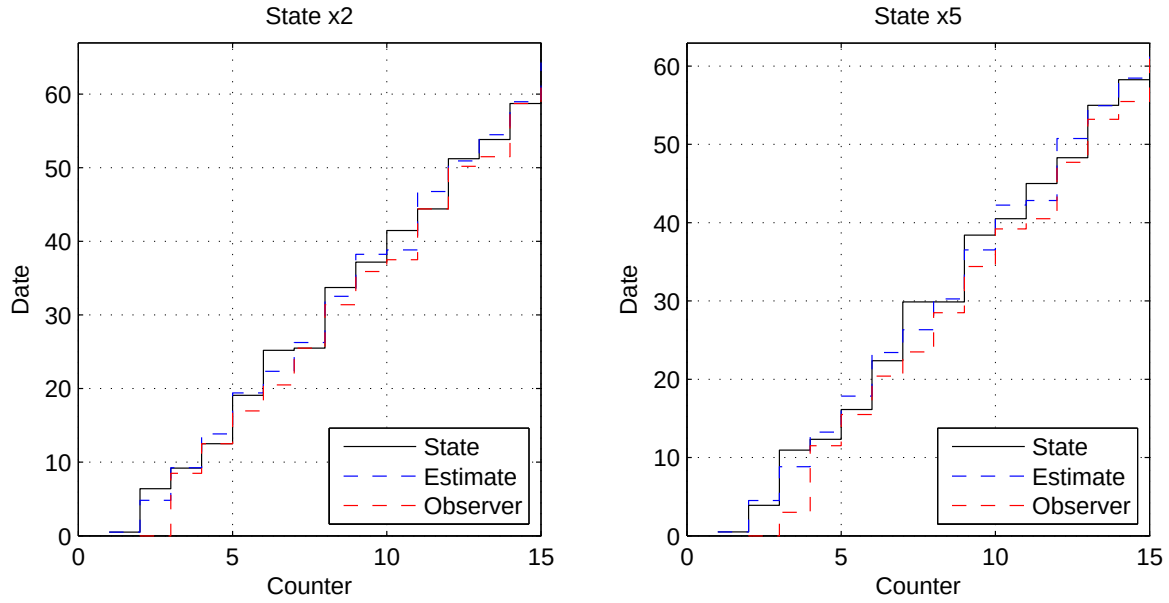

Figure 2. State trajectory for Example 2

up to the occurrence of 400 events.

	RMSE$(x_i(k), \hat{x_i}(k\vert\, k))$	RMSE$(x_i(k), x_i^{obs}(k))$
$i = 2$	1.4317	2.5027
$i = 5$	1.5348	3.2494

Table II
COMPARISON BETWEEN THE OBSERVER AND THE FILTER

This comparison allows us therefore to evaluate how the filter, based on the knowledge of the noise statistics, can improve the state estimation beyond the structural considerations.

VI. CONCLUSIONS

This paper presents an algorithm for the stochastic filtering of max-plus linear systems. The basis of the proposal are, as usual in filtering theory, an equation for the prediction of the next state given an actual estimate (Equation 15) and an equation for the update of the estimate given a new measure (Equation 16). The prediction, given by $E[x(k)|\, \hat{x}(k-1|\, k-1)]$, can be obtained by Equations (4) to (6) through direct calculation. The update is performed by Algorithms 1 and 2, whose convergence is guaranteed by lemmas 2 and 3. It can be seen as the inverse of $E[z^*|\, x]$ with respect to x. Since the solution may not to be unique, the closest to the prediction is

chosen.

The recursive calculation of the *a posteriori* probability of the state given the measures for linear max-plus systems is not known up to this time. This is due to multiplicity of cases and the lack of regularity concerning the mathematical form of the probability densities that may occur. Although the proposed method is not based on the explicit calculation of the *a posteriori* probability of the state given the measures, it is argued its similarity with the classical maximum likelihood estimator and the fact that it inverts, in some sense, the direct estimation of the measure given the state.

The performed simulations show the consistency of the proposed method in the sense that the calculation of $\hat{x}(k|\,k)$ takes into account the prior information $\hat{x}(k|\,k-1)$ and the measure $z^*(k)$. Besides, a comparison with the observer proposed in [8] show how the consideration of the statistical aspects of the noise can improve the estimates if compared with the lower bound calculated by the observer, based only on structural assumptions.

The filtering schema herein proposed can be further developed by the consideration of more general probability densities for the entries of the matrices A, B and C than the bounded disturbances considered here. Particularly the work [23] seems to be promising in this sense. Moreover other optimization strategies can be considered for the update equation (Equation 16), taking into account the trade off between the noise in the measure versus the noise in the prediction.

APPENDIX
SYSTEMS WITH INDEPENDENT TIME DELAYS

Consider the following linear equation, written in max-plus algebra, typically obtained from a Timed Event Graph[11] ([24],[25]):

$$\begin{aligned} x(k) &= (\oplus_{j=0}^{m} A_j x(k-j)) \oplus (\oplus_{j=0}^{r} B_j u(k-j)) \\ z(k) &= \oplus_{j=0}^{s} C_j x(k-j), \end{aligned} \tag{19}$$

where the dimensions of all matrices and vectors are the same as in Equation (1). In a large variety of applications, the entries of matrices A_j $(j = 0,\ldots,m)$, B_j $(j = 0,\ldots,r)$ and C_j $(j = 0,\ldots,s)$ correspond to time delays of single and independent processes and, in what follows, they are assumed to be independent random variables with finite support and piecewise polynomial cumulative distribution functions.

Equation (19) can be transformed into Equation (1) by taking into account the solution of the equation $x = A_0 x \oplus b$ in max-plus algebra, given by[12] $x = A_0^* b$. Equation (19) then becomes:

$$\begin{aligned} x(k) &= A_0^*(\oplus_{j=1}^{m} A_j x(k-j)) \oplus A_0^*(\oplus_{j=0}^{r} B_j u(k-j)) \\ z(k) &= \oplus_{j=0}^{s} C_j x(k-j). \end{aligned} \tag{20}$$

Defining $\tilde{x}(k) = [x(k)',\ldots,x(k-m+1)',u(k)',\ldots,u(k-r+1)']' \in (\overline{\mathbb{R}}_{\max})^{nm+pr}$ and, accordingly, matrices $\tilde{A}$, $\tilde{B}$ and $\tilde{C}$, the preceding equations result in $\tilde{x}(k) = \tilde{A}\tilde{x}(k-1) \oplus \tilde{B}u(k)$ and $z(k) = \tilde{C}\tilde{x}(k)$.

It should be noted that the entries of matrix $\tilde{A}$ are not mutually independent and therefore Equations (4), (5) and (6) can not be straightforwardly used to obtain $E[\tilde{x}(k)|\,\tilde{x}(k-1)]$. In the following, an approximate method for the calculation of these mathematical expectations is proposed, based in one property of the matrix A_0 (Equation 19). As shown in [24], section 2.5.3, if the TEG associated to Equation (19) is live then, by a convenient permutation of the coordinates, A_0 can always be put in a strictly lower triangular form. As a consequence, without loss of generality, the first expression in Equation (19) can be written as:

$$\begin{aligned} x_1(k) &= y_1(k) \\ x_2(k) &= a_{2,1}^0 x_1(k) \oplus y_2(k) \\ x_3(k) &= a_{3,1}^0 x_1(k) \oplus a_{3,2}^0 x_2(k) \oplus y_3(k) \\ &\vdots \\ x_n(k) &= \oplus_{l=1}^{n-1}(a_{n,l}^0 x_l(k)) \oplus y_n(k) \end{aligned} \tag{21}$$

where $a_{i,l}^0$ are the entries of the matrix A_0 and $y_i(k)$ is the i^{th} component of the vector $(\oplus_{j=1}^{m} A_j x(k-j)) \oplus (\oplus_{j=0}^{r} B_j u(k-j))$. Given that the objective is the calculation of the conditional expectation of $x_i(k)$, it should be noted that each $y_i(k)$ is the max of terms of the form $a_{il}^j + x_l(k-j)$ or $b_{il}^j + u_l(k-j))$, where $x_l(k-j)$ and $u_l(k-j))$ are fixed and a_{il}^j and b_{il}^j are independent random variables.

Starting from the equation $x_1(k) = y_1(k)$ and assuming that the values $x(k-1),\ldots,x(k-m)$ and $u(k-1),\ldots,u(k-r)$ are fixed, it is possible the direct calculation of $E[x_1(k)|\,x(k-1),\ldots,x(k-m),u(k-1),\ldots,u(k-r)]$ = $E[x_1(k)|\,\tilde{x}(k-1)]$ using Equations (4), (5) and (6). The calculation of $E[x_2(k)|\,\tilde{x}(k-1)]$ can not be performed in the same way, given that it depends on $x_1(k)$. To calculate this and the subsequent mathematical expectations, the joint distribution of the vector $x(k)$ should be obtained which is an intractable problem. To avoid it, the following strategy is proposed. To calculate the expectation of the second component, the random variable $x(1)$ in the right hand of Equations (21) is simply replaced by the deterministic term $\alpha_1 E[x_1(k)|\,\tilde{x}(k-1)]$ (the parameter α_1 is discussed in the following) rendering the second equation to be of the same kind as $y_i(k)$ and thus computable by Equations (4), (5) and (6). To calculate $E[x_3(k)|\,\tilde{x}(k-1)]$, the random variables $x_1(k)$ and $x_2(k)$ are respectively replaced by $\alpha_1 E[x_1(k)|\,\tilde{x}(k-1)]$ and $\alpha_2 E[x_2(k)|\,\tilde{x}(k-1)]$, previously calculated. The calculation of all subsequent expectations can be performed by a similar procedure, thanks to the triangular structure of matrix A_0. This leads to the complete and computationally efficient evaluation of $E[\tilde{x}(k)|\,\tilde{x}(k-1)]$.

To obtain a proper set of parameters $[\alpha_1,\ldots,\alpha_{n-1}]$ and to evaluate the quality of this approximation, Monte Carlo simulation techniques can be used as follows. It should be noted first that only a limited region for the variable $\tilde{x}(k-1)$, denoted by Υ, is statistically relevant for evaluation purposes thanks to the periodic behavior of max-plus linear systems. Therefore given the matrices A_0, ..., A_m and a set of pa-

[11] A Timed Event Graph (TEG) is a Petri Net whose places have only one transition upstream and only one transition downstream. TEG's are often used as a first step in the modeling of a Discrete Event Dynamic System, because max-plus equations are naturally derived from them.

[12] The operator $*$ is known as the Kleene operator, defined by $A_0^* = E \oplus A_0 \oplus A_0^2 \oplus \ldots$, where E is the identity matrix.

rameters $\theta = [\alpha_1, \ldots, \alpha_{n-1}]$, it is possible to compare, within the region Υ, the mathematical expectations obtained by the proposed approximation and those obtained by Monte Carlo Simulation, and evaluate the quality of the approximation through classical statistical parameters (as for instance the mean and the standard deviation of the difference between both means). Once more, thanks to the triangular structure of the matrix A_0, it is possible to perform an uncoupled search for each α_i, aiming the minimization of the difference between both means. The mean, together with the standard deviation associated to the optimal θ allow the evaluation of the obtained approximation. Given the matrices A_0, ..., A_m this optimization procedure must be performed only once (off line) and therefore it does not increase the processing time of the filtering algorithm previously described.

REFERENCES

[1] C. Cassandras and S. Lafortune, *Introduction to Discrete Event Systems*, ser. SpringerLink Engineering. Springer, 2007.

[2] C. A. Maia, L. Hardouin, R. S. Mendes, and B. Cottenceau, "Optimal Closed-Loop Control for Timed Event Graphs in Dioid," *IEEE Transactions on Automatic Control*, vol. 48, no. 12, pp. 2284–2287, 2003.

[3] R. Lüders and R. Santos-Mendes, "Generalized Multivariable Control of Discrete Event Systems in Dioid," in *6th International Workshop on Discrete Event Systems (WODES'02)*, Zaragoza, Spain, 2002.

[4] B. De Schutter and T. van den Boom, "Model predictive control for max-plus-linear discrete event systems," *Automatica*, vol. 37, no. 7, pp. 1049–1056, jul 2001.

[5] I. Necoara, B. De Schutter, T. van den Boom, and H. Hellendoorn, "Stable model predictive control for constrained max-plus-linear systems," *Discrete Event Dynamic Systems: Theory and Applications*, vol. 17, no. 3, pp. 329–354, sep 2007.

[6] M. Lhommeau, L. Hardouin, B. Cottenceau, and C. A. Maia, "Robusness analysis of control laws for (max,+)-linear systems," in *International Workshop on Max-algebra (IWMA)*, Birmingham, Inglaterra, Jun. 2003.

[7] I. Necoara, B. De Schutter, T. van den Boom, and H. Hellendoorn, "Robust control of constrained max-plus-linear systems," *Int. J. of Robust and Nonlinear Control*, vol. 19, no. 2, pp. 218–242, Jan. 2009.

[8] L. Hardouin, C. A. Maia, B. Cottenceau, and M. Lhommeau, "Observer Design for (max,plus) Linear Systems," *IEEE Transactions on Automatic Control*, vol. 55, no. 2, feb 2010.

[9] M. DiLoreto, S. Gaubert, R. Katz, and J.-J. Loiseau, "Duality between invariant spaces for max-plus linear discrete event systems," *SIAM Journal on Control and Optimization*, vol. 48, no. 8, pp. 5606–5628, Dec. 2010.

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

[11] S. Farahani, T. van den Bom, and B. D. Schutter, "On optimization of stochastic max-min-plus-scaling systems - an approximation approach," *Automatica*, vol. 83, pp. 20–27, 2017.

[12] L. Hardouin, C. A. Maia, B. Cottenceau, and R. Santos-Mendes, "Max-plus Linear Observer: Application to manufacturing Systems," in *Proceedings of the 10th International Workshop on Discrete Event Systems, WODES'10*, Berlin, sep 2010, pp. 171–176.

[13] T. Blyth and M. Janowitz, *Residuation Theory*. Oxford: Pergamon Press, 1972.

[14] D. Luenberger, "An introduction to observers," *IEEE Transactions on Automatic Control*, vol. 16, no. 6, pp. 596–602, December 1971.

[15] A. Jazwinski, *Stochastic Processes and Filtering Theory*, ser. Mathematics in Science and Engineering. Academic Press, 1970, vol. 63.

[16] B. Anderson and J. Moore, *Optimal Filtering*, ser. Dover Books on Electrical Engineering Series. Dover Publications, 2005.

[17] Y. Bar-Shalom, X. Li, and T. Kirubarajan, *Estimation with Applications to Tracking and Navigation*. Norwood: Artech House, 2001.

[18] S. Julier, J. Uhlmann, and H. F. Durrant-Whyte, "A new method for the nonlinear transformation of means and covariances in filters and estimators," *IEEE Transactions on Automatic Control*, vol. 45, no. 3, pp. 477–482, March 2000.

[19] B. Ristic, S. Arulampalam, and N. Gordon, *Beyond the Kalman Filter*. Artech House, 2004.

[20] M. Arulampalam, S. Maskell, N. Gordon, and T. Clapp, "A tutorial on particle filters for online nonlinear/non-gaussian bayesian tracking," *Signal Processing, IEEE Transactions on*, vol. 50, no. 2, pp. 174 –188, feb 2002.

[21] D. F. e. Silva, R. S. Mendes, L. Hardouin, C. A. Maia, and B. Cottenceau, "Filtragem estocástica aplicada a sistemas max-plus lineares," in *Anais do X Simpósio Brasileiro de Automação Inteligente*, 2011, pp. 1388–1393.

[22] G. J. Olsder, J. A. C. Resing, R. E. D. Vries, M. S. Keane, and G. Hooghiemstra, "Discrete event systems with stochastic processing times," *IEEE Transactions on Automatic Control*, vol. 35, no. 3, pp. 299–302, Mar 1990.

[23] S. Farahani, T. van den Bom, H. van den Weide, and B. D. Schutter, "An approximation method for computing the expected vale of max-affine expressions," *European Journal of Control*, vol. 27, pp. 17–27, 2016.

[24] F. Baccelli, G. Cohen, G. Olsder, and J. Quadrat, *Synchronisation and Linearity: An Algebra for Discrete Event Systems*. New York: John Wiley and Sons, 1992.

[25] B. Heidergott, G. Olsder, and J. van der Woude, *Max Plus at Work: Modeling and Analysis of Synchronized Systems : a Course on Max-Plus Algebra and Its Applications*, ser. Max Plus at work: modeling and analysis of synchronized systems : a course on Max-Plus algebra and its applications. Princeton University Press, 2006, no. v. 13.

[26] P. Ramadge and W. Wonham, "The control of discrete event systems," *IEEE TAC*, vol. 77, pp. 81–98, 1989.

[27] R. D. Katz, "Max-plus (a,b)-invariant spaces and control of timed discrete-event systems," *IEEE Transactions on Automatic Control*, vol. 52, no. 2, pp. 229–241, Feb 2007.

[28] M. Cavalcante and R. S. Mendes, "Modular methodology for the netwrok calculus in a time-varying context," *IEEE Transactions on Information Theory*, vol. 59, no. 10, pp. 6342–6356, October 2013.

[29] T. van den Boom and B. De Schutter, "Model predictive control for perturbed max-plus-linear systems: A stochastic approach," *International Journal of Control*, vol. 77, no. 3, pp. 302–309, feb 2004.

[30] B. Anderson and J. Moore, *Optimal Filtering*, ser. Dover Books on Electrical Engineering. Dover, 2005.

[31] L. Jaulin, M. Kieffer, O. Didrit, and E. Walter, *Applied Interval Analysis*. London: Springer-Verlag, 2001.

[32] D. Wilde, *Optimum seeking methods*. Prentice-Hall, 1964.

[33] A. Tarski, "A lattice-theoretical fixpoint theorem and its applications," *Pacific Journal of Mathematics*, vol. 5, no. 2, pp. 285–309, 1955.

PLACE PHOTO HERE

Rafael Santos Mendes received the Electrical Engineering degree from the State University of Campinas, Brazil (UNICAMP, 1981), the M.S. degree from the State University of Campinas (UNICAMP, 1984) and the Docteur de l'Université degree from the Paul Sabatier University/LAAS-CNRS (Toulouse, France, 1988). Currently, he is an Associate Professor in the School of Electrical and Computer Engineering, UNICAMP. His research interests include Discrete Event Dynamic Systems, Dioid Algebra, Modal Logic and Stochastic Filtering.

PLACE PHOTO HERE

Laurent Hardouin received the degrees of Master and Ph.D. in Acoustic and Control from the University of Poitiers, France, in 1990 and 1993, respectively. He received the Habilitation Diriger des Recherches from the University of Angers, France, in 2004. He is currently a Full Professor of ISTIA Engineering School at the University of Angers, France. He specializes in discrete-event systems, max-plus algebra, and interval analysis with applications to computer networks, manufacturing systems, transportation systems, and robotics

PLACE
PHOTO
HERE

Mehdi Lhommeau was born in 1975. He received the Ph.D. degree from the University of Angers, France, in 2003. He is currently an Associate Professor at the University of Angers. His main research interests include modeling and control of timed discrete event systems in max-plus algebra with applications in manufacturing systems and computer networks.