2013 European Control Conference (ECC)
July 17-19, 2013, Zürich, Switzerland.

Particle Filters For Max Plus Systems*

Renato Markele Ferreira Cândido[1], Rafael Santos Mendes[1], Laurent Hardouin[2] and Carlos Andrey Maia[3]

Abstract— **The main objective of this article is to synthesize a particle filter algorithm for max-plus systems. It is presented a brief introduction to the max-plus approach for Discrete Event Systems. Next, the fundamentals of the particle filter and the way in which they can be applied to max-plus systems are presented. It leads to the algorithm for particles filtering. Lastly, some examples are given. The results shows the accuracy of the method and the improvements in comparison with the deterministic observer.**

I. Introduction

The filtering problem consists in the estimation of an unknown current state of a dynamical system based on noisy observations of its outputs. In this paper one attempts to estimate sequentially the state of a Discrete Event System (DES) in a noise context, where the state is not fully observable, by using a Bayesian filter. The main idea in the Bayesian approach to dynamic state estimation is to construct the *posterior* probability density function of the states, based on all available information, including the sequence of received measurements [1]. The optimal algorithms for recursive Bayesian state estimation are restricted to linear systems affected by Gaussian noise (e.g. The Kalman filter) [2], to systems where the state space is finite and discrete-valued (e.g. Grid-Based Methods) and to certain subclasses of nonlinear problems, discovered by Beneš [3] and Daum [4], [5]. The systems covered in this work do not fit into any of these classes, thus, a sub-optimal alternative has been considered: The *Particle Filters* (PFs). These filters performs a *Sequential Monte Carlo* estimation to produce an approximation of a probability distribution based on a set of samples (particles) with associated weights.

The Discrete Event (Dynamic) Systems (DEDS) are *discrete-state* systems whose state evolution depends entirely on the occurrence of asynchronous discrete events over time [6]. Among the existing models for treating these systems stand out the approach based in the Max-Plus Algebra [7], which is the model adopted in this work to obtain the equation that describes the evolution of the states and the equation that relates the states with the output.

In a previous work [8], it was considered a particular case in which some components of the state vector are directly observable. In this work it is assumed the general case where this characteristic is not required. This generalization has been achieved due to a step of *reconditioning* of the particles discussed in section III.

The paper is organized as follows. In section II, we present a brief review of Petri nets and max-plus algebra, in this same section, we also presents the structure of the system discussed in this paper as well as the scene of the disturbances that affect it. In section III, is described the operating principle of particle filter and its application to max-plus systems. Also in section III, the proposed algorithm for particles filtering of max plus systems is synthesized. In section IV, we present some simulations and the conclusions are presented in section V.

II. Petri Nets and Max-Plus Algebra

Petri Nets describe DEDS pictorially and can be viewed as a bipartite graph. The systems described in these networks are dynamic, which is a feature not observed in conventional graphs. An important equivalence can be made between the DEDS without concurrency and a subclass of Petri nets called *event graphs* [7]. For any timed event graph is possible to obtain mathematical models, that are similar to conventional linear dynamic systems, in terms of recursive and linear equations in an appropriate algebra called Max-Plus Algebra. A timed event graph (TEG) is a Petri net in which all places have exactly one upstream and one downstream transition. In this paper, the p-timed Petri nets are used. In these networks each place is associated with a minimum time of permanency for the tokens.

Fig. 1 shows an example of a timed event graph. In this figure, transition u has no predecessor(s) so, it is an input of the network. In this type of transition *firing* are due to external decisions. The transitions z_1 and z_2 do not have successor, these transitions are the outputs of the network. The transitions x_1 and x_2 are called state transitions. In the context of filtering developed in the next section, it is assumed that only the output transitions are observed.

For each transition T_i, one can associate a non-decreasing sequence formed by the variables $x_i(k),\ k = 0, 1, 2, ...,$ called ***daters***, containing the k-th instant of firing of the transition T_i. Assuming that the sequences of firing associated with the input transitions are known it is possible to

*The authors are grateful to the Brazilian Agencies CAPES, CNPQ and FAPEMIG for the financial support

[1]R. M. F. Cândido and R. S. Mendes are with the Department of Computer Engineering and Industrial Automation, School of Electrical and Computer Engineering, State University of Campinas - UNICAMP, 13083-970 Campinas - SP - Brazil. email: renatomarkele@hotmail.com; rafael@dca.fee.unicamp.br

[2]L. Hardouin is with Laboratoire d'Ingénierie des Systèmes Automatisés - ISTIA, Université d'Angers, Av. Notre Dame du Lac 62, 49000 Angers , France. email: Laurent.hardouin@istia.univ-angers.fr

[3]C. A. Maia is with Department of Electrical Engineering, Federal University of Minas Gerais - UFMG, Av. Antônio Carlos 6627, Pampulha, 31270-010 Belo Horizonte - MG, Brazil. email: maia@cpdee.ufmg.br

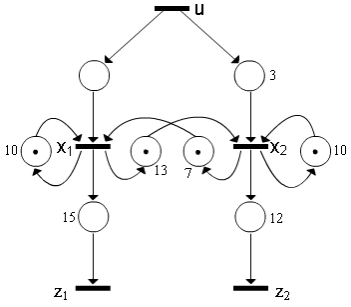

Fig. 1. A timed events graph

determine the sequences of firing associated with all transitions of the TEG. Considering again Fig. 1 the following relationships are true:

$$x_1(k) = max\{10 + x_1(k-1); 7 + x_2(k-1); u(k)\} \quad (1a)$$
$$x_2(k) = max\{13 + x_1(k-1); 10 + x_2(k-1); 3 + u(k)\} \quad (1b)$$
$$z_1(k) = 15 + x_1(k) \quad (1c)$$
$$z_2(k) = 12 + x_2(k) \quad (1d)$$

Considering now the set: $\mathbb{R} \cup \{-\infty\} \cup \{\infty\}$ and the operations: $sum \equiv \oplus \equiv max$ and $product \equiv \otimes \equiv +$, one can obtain:

$$x_1(k) = 10 \otimes x_1(k-1) \oplus 7 \otimes x_2(k-1) \oplus u(k) \quad (2a)$$
$$x_2(k) = 13 \otimes x_1(k-1) \oplus 10 \otimes x_2(k-1) \oplus 3 \otimes u(k) \quad (2b)$$
$$z_1(k) = 15 \otimes x_1(k) \quad (2c)$$
$$z_2(k) = 12 \otimes x_2(k) \quad (2d)$$

which, in matrix form, is:

$$x_k = A \otimes x_{k-1} \oplus B \otimes u_k \quad (3a)$$
$$z_k = C \otimes x_k \quad (3b)$$

where:

$$A = \begin{pmatrix} 10 & 7 \\ 13 & 10 \end{pmatrix}; \; B = \begin{pmatrix} 0 \\ 3 \end{pmatrix}; \; C = \begin{pmatrix} 15 & -\infty \\ -\infty & 12 \end{pmatrix}$$
$$x_k = \begin{pmatrix} x_1(k) \\ x_2(k) \end{pmatrix}; \; u_k = u(k) \;\; \text{e} \;\; z_k = \begin{pmatrix} z_1(k) \\ z_2(k) \end{pmatrix}$$

Therefore, all transitions of the timed event graph were described by a recursive linear system of equations.

A set D, endowed with two internal operations: $sum(\oplus)$ and $product(\otimes)$ is a **dioid** or **idempotent semiring** if the sum is associative, commutative and idempotent (i.e. $a \oplus a = a$) and the product is associative and left and right distributive with respect to the sum[1]. It is also necessary the existence of null (or zero) element (i.e. $\exists \epsilon \in D : \forall a \in D, a \oplus \epsilon = 0$) and the identity element (i.e. $\exists e \in D : \forall a \in D, a \otimes e = e \otimes a = a$) so that the zero element is absorbing for the $\otimes$ operation (i.e. $\forall a \in D, a \otimes \epsilon = \epsilon \otimes a = \epsilon$) [7]. Given these conditions, it is easy to see that the set $\mathbb{R} \cup \{-\infty\} \cup \{\infty\}$ and the operations $\oplus \equiv max$ and $\otimes \equiv +$ with $\epsilon = -\infty$ and $e = 0$ is a dioid. Moreover, it can be stated that this is a **complete dioid** because it is closed for infinite sums and the left and right distributivity of the product extends to infinite sums. This set is called **Max-Plus** and noted by $\overline{\mathbb{R}}_{max}$.

This paper considers discrete event systems that can be described by equations (3) where $x_k \in \mathbb{R}^{n \times 1}$, $u_k \in \mathbb{R}^{p \times 1}$ and $z_k \in \mathbb{R}^{q \times 1}$. It is assumed that the matrices A, B and C have suitable dimensions and that its elements a_{ij}, b_{ij} and c_{ij} respectively, are deterministic or independent random variables with uniform or exponential known distributions.

III. Particles Filters Applied to Max Plus Systems

The filtering problem considered in this work, consists in estimate sequentially the states x_k based on a known sequence of measurements z_k and inputs u_k. The Particle Filters are suboptimal filters that perform a Sequential Monte Carlo (SMC) estimation based on particle representation of probability densities. The following is a brief description of the Monte-Carlo method [1].

Let $I = \int g(\phi) d\phi$ be an integral that we wish to solve and suppose $g(\phi)$ can be factored as $g(\phi) = f(\phi) \cdot \pi(\phi)$ in a such way that $\pi(\phi)$ is a probability density. It is generally impossible to sample from $\pi(\phi)$ directly. So, assuming that it is possible to draw $N \gg 1$ samples $\{\phi^i, i = 1, ..., N\}$ distributed according to $q(\phi)$, similar[2] to $\pi(\phi)$, the Monte Carlo estimate integral is $I_N = \sum_{i=1}^{N} f(\phi^i) \cdot \omega(\phi^i)$, where $\omega(\phi^i) = \frac{\tilde{\omega}(\phi^i)}{\frac{1}{N}\sum_{j=1}^{N} \tilde{\omega}(\phi^j)}$ and:

$$\tilde{\omega}(\phi) = \frac{\pi(\phi)}{q(\phi)}. \quad (4)$$

This process is called *importance sampling*.

The importance sampling is then a technique to approximate a target distribution $\pi(\phi)$ using samples drawn from a importance distribution $q(\phi)$ by means of a process of weighting of the samples.

Now, let $\mathbf{X}_k = \{x_j, j = 0, ..., k\}$ be the sequence of the first $k+1$ firing times of all states of a system described by equations (3). The joint posterior density at k is denoted by $p(\mathbf{X}_k|\mathbf{Z}_k)$ and its marginal is $p(x_k|\mathbf{Z}_k)$. Assuming the existence of a set of particles (or samples) denoted by $\{\mathbf{X}_k^i, i = 1, ...N\}$ and their respective weights $\{\omega_k^i, i = 1, ..., N\}$, the joint posterior at k can be approximated as:

$$p(\mathbf{X}_k|\mathbf{Z}_k) \cong \sum_{i=1}^{N} \omega_k^i \delta(\mathbf{X}_k - \mathbf{X}_k^i), \quad (5)$$

where $\delta(\cdot)$ is the impulse function.

If the particles $\mathbf{X}_k^i$ were drawn from an importance density $q(\mathbf{X}_k|\mathbf{Z}_k)$, then according to (4):

$$\omega_k^i \propto \frac{p(\mathbf{X}_k^i|\mathbf{Z}_k)}{q(\mathbf{X}_k^i|\mathbf{Z}_k)}. \quad (6)$$

[1] the product is not necessarily commutative

[2] A probability density $q(\phi)$ (usually referred to as importance density) is similar to $\pi(\phi)$ if $\forall \phi : \; \pi(\phi) > 0 \Rightarrow q(\phi) > 0$.

To determine a recursive filtering procedure, suppose that at $k-1$ are available a set of particles $\{\mathbf{X}_{k-1}^i, i=1,...N\}$, drawn according to the density $q(\mathbf{X}_{k-1}|\mathbf{Z}_{k-1})$, and a set of weights $\{\omega_{k-1}^i, i=1,...,N\}$ such that the following approximation is true:

$$p(\mathbf{X}_{k-1}|\mathbf{Z}_{k-1}) \cong \sum_{i=1}^{N} \omega_{k-1}^i \delta(\mathbf{X}_{k-1} - \mathbf{X}_{k-1}^i).$$

Suppose now that the density of importance q be such that:

$$q(\mathbf{X}_k|\mathbf{Z}_k) = q(x_k|x_{k-1}, z_k) q(\mathbf{X}_{k-1}|\mathbf{Z}_{k-1}). \quad (7)$$

If each particle $\mathbf{X}_{k-1}^i$ is augmented to $\mathbf{X}_k^i$ using the new state $x_k^i \sim q(x_k|x_{k-1}, z_k)$ ($\sim$ stands for distributed as) then the following approximation is true:

$$p(\mathbf{X}_k|\mathbf{X}_{k-1}, \mathbf{Z}_{k-1}) \cong \sum_{i=1}^{N} \omega_{k-1}^i \delta(\mathbf{X}_k - \mathbf{X}_k^i).$$

As soon as the measure z_k is available, [1] shows that the updating of each weight ω_{k-1}^i is achieved by using:

$$\omega_k^i \propto \omega_{k-1}^i \frac{p(z_k|x_k^i) p(x_k^i|x_{k-1}^i)}{q(x_k^i|x_{k-1}^i, z_k)}, \quad (8)$$

and an ulterior normalization, which leads to the approximation (5), concluding the recursive procedure.

The particle filter consists then in recursive propagation of the importance weights ω_k^i and states x_k^i as each measurement is received. The choice of the importance is one of the most critical issues in the design of a particle filter [1] although this can be chosen freely since the condition of similarity is respected. It has been shown in [9] that the optimal importance density function that minimizes the variance of important weights, conditioned upon x_k^i and z_k is given by:

$$q(x_k|x_{k-1}^i, z_k)_{opt} = p(x_k|x_{k-1}^i, z_k).$$

However, in this paper the following suboptimal choice has been adopted:

$$q(x_k|x_{k-1}^i, z_k) = p(x_k|x_{k-1}^i). \quad (9)$$

This choice allows the use of the equation (3a) to augment each existing particle $\mathbf{X}_{k-1}^i$ to $\mathbf{X}_k^i$. This step is similar, in Kalman filtering, to the procedure of prediction of the state.

By substituting (9) in (8) one obtains:

$$\omega_k^i = \omega_{k-1}^i p(z_k|x_k^i). \quad (10)$$

This equation is used to update the weights of the particles at the time when a new measurement becomes available. This is similar, in Kalman filtering, to the update step. The measure z_k, the particle x_k^i and the input u_k are known at the time of this calculation, thus, the update can be done by:

$$\omega_k^i \quad = \quad \omega_{k-1}^i \cdot V(x_k, z_k), \quad (11)$$

the likelihood function $V(x,z)$ is derived in Appendix.

The estimation of x_k is therefore given by:

$$\hat{x}_k = \sum_{i=1}^{N} x_k^i w_k^i. \quad (12)$$

In [10] is shown that the importance weights are a *Martingale* sequence in k so, its variance is an increasing function of k. This feature of the importance weights is harmful to the performance of the filter and leads to the degeneracy phenomenon that occurs when, after a certain number of recursive steps, many particles get very low weights, which leads to a negligible contribution to the approximation of $p(x_k|\mathbf{Z}_k)$. The *effective sample size* $\hat{N}_{eff} = \frac{1}{\sum_{i=1}^{N} \omega_k^i}$ is used as measure of degeneracy of the algorithm. This phenomenon is impossible to avoid but there is a strategy to overcome degeneracy of samples. Whenever the value of $\hat{N}_{eff}$ is below a certain threshold N_{thr} a process of *resampling* is used. In this process the particles with low weights are eliminated and the particles with non-negligible weights are cloned proportionally to their respective weights [1].

In general, the update of the weights is done by the likelihood function. However a problem occurs when there is a deterministic relationship between the output and the state. For example, assume that in a given system the relationship between the state and the output is given by (13).

$$z(k) \quad = \begin{bmatrix} 2 & \mu(2,3) \end{bmatrix} \otimes \begin{bmatrix} x_1(k) \\ x_2(k) \end{bmatrix}, \quad (13)$$

where $\mu(2,3)$ represents a uniform random variable distributed between 2 and 3.

The equation (13) can be seen as $z(k) = Cx(k)$ with $C = \begin{bmatrix} 2 & \mu(2,3) \end{bmatrix}$. The cumulative probability and probability density functions of element c_{12} are given in the appendix (see Eq. (A.1) and (A.2)). For the element c_{11} these equations are given respectively by:

$$F_{11}(\tau) = \begin{cases} 0 & \text{if } \tau < 2 \\ 1 & \text{if } \tau \geq 2 \end{cases} \quad (14)$$

$$p_{11}(\tau) = \delta(\tau - 2) \quad (15)$$

Consider that at a certain iteration k, the measurement $z(k) = 5$ is acquired. Consider also that after the process of predicting the state, the following set of particles approximating the *prior* density $p(x_k|x_{k-1})$ was available:

$$\mathbf{x}_k = \begin{Bmatrix} 2.05 & 0.60 & 2.80 & 1.80 & 2.95 & 3.05 & 2.95 \\ 1.80 & 2.80 & 1.70 & 2.20 & 2.80 & 1.20 & 0.30 \end{Bmatrix}$$

where each column represents a particle $x_k^i = \begin{bmatrix} x_1^i(k) & x_2^i(k) \end{bmatrix}^T$. Using (A.5), the cumulative probability and probability density functions given by (14) and (15), respectively, one can calculate the likelihood of each particle

x_k^i:

$$V(x_k^1, z) = \prod_{i=1}^{1}(\sum_{j=1}^{2} p_{ij}(z_i - x_j) \prod_{\substack{k=1,\\ k\neq j}}^{2} F_{ik}(z_i - x_k))$$
$$= p_{11}(5-2.05)F_{12}(5-1.8)$$
$$+ p_{12}(5-1.8)F_{11}(5-2.05) = 0\cdot 1 + 0\cdot 1 = 0.$$

$$V(x_k^2, z) = p_{11}(4.4)F_{12}(2.2) + p_{12}(2.2)F_{11}(4.4)$$
$$= 0\cdot 1 + 1\cdot 1 = 1.$$

Similarly: $V(x_k^3, z) = 0,\ V(x_k^4, z) = 1,\ V(x_k^5, z) = 1,$ $V(x_k^6, z) = 0$ and $V(x_k^7, z) = 0$.

For this simple example, is easy to conclude that the *posterior* possible values for the state $[x_1(k)\ x_2(k)]$, given the measure $z(k) = 5$, belong to region $R_1 \cup R_2$, where $R_1 = \{[x_1(k)\ x_2(k)] \in \mathbb{R}^2 : x_1(k) = 3\ and\ x_2(k) < 2\}$ and $R_2 = \{[x_1(k)\ x_2(k)] \in \mathbb{R}^2 : x_1(k) \leq 3\ and\ 2 \leq x_2(k) \leq 3\}$, depicted in Fig. 2. The region R_1 is a half line that corresponds to an event with null probability measure in the *prior* probability space, i.e. the probability of a particle be within R_1 is zero. In the particular case where all the elements of C are deterministic this would lead to a step of updating of the weights (see Eq. (11)) where all weights would become zero.

It should be noted however that, if the measure $z(k) = 5$ occurs then the state belongs to the feasible region $R_1 \cup R_2$. So, the original distribution of the particles must be rearranged in order to undertake this fact. This rearrangement of the particles can be done by a step of *reconditioning* of the particles before calculating the likelihood function. At this step, the following considerations are done: (1) Particles sufficiently close to the feasible region can be modified in order to belong to the feasible region if such changes interfere only in states related deterministically with the output. e.g.: Assume that in a given system the output z_i is related with the state x_j as follows: $z_i = c_{ij} \otimes x_j$, where c_{ij} is deterministic. To determine if a particle X^n is " sufficiently close to the feasible region" it is evaluated (using classic algebra!) if $z_i - (c_{ij} + x_j^n)$ is smaller than a threshold K_{thr}. If so, then the state x_j^n is modified to $x_j^n = z_i - c_{ij}$. If there are more than one output deterministically related to a same state x_j, then this state will be modified to $x_j = \min_i(z_i - c_{ij})$ if and only if $\min_i(z_i - (c_{ij} + x_j)) \leq K_{thr}$; (2) The states x_j^n that have been modified must not interfere in the likelihood calculus of the particle X^n since this would involve accounts with the impulse function $\delta(\tau)$ which would lead to infinite likelihoods or even indeterminacies. Reviewing (A.5) it can be observed that the calculation of the likelihood function involves products, thus, a way to get these states not interfering in the likelihood calculus is to consider $p_{ij}(\tau) = 1\ if\ \tau = 0$ and $p_{ij}(\tau) = 0\ if\ \tau \neq 0$ as the probability density function and maintain the cumulative probability function given in (14) for the deterministic elements c_{ij}; (3) Possible changes in the importance density $q(\mathbf{X}_k|\mathbf{Z}_k)$ that these changes cause can be minimized by redistributing the particles that were not changed to the feasible region in order to maintain the original mean.

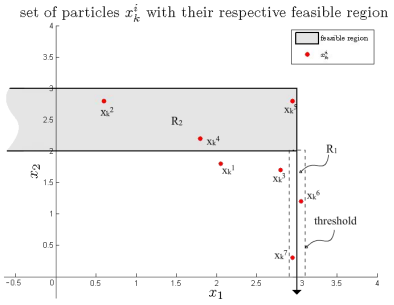

Fig. 2. Feasible region

Thereby, after the step of reconditioning, the particles x_k^6 and x_k^7 in Fig. 2 become $[3.00\ \ 1.20]^T$ and $[3.00\ \ 0.30]^T$ respectively, and its likelihoods are both 1.

Thus, the algorithm for particles filtering of max-plus systems can be synthesized as follows:

1) $k = 0$;
2) Initialize N particles, X_0^i, $i = 1, \ldots, N$ and its respective weights;
3) $k \leftarrow k + 1$;
4) Augment each existing particle $\mathbf{X}_{k-1}^i$ to $\mathbf{X}_k^i$ (see Eq. (3a));
5) Read the measurement z_k;
6) Recondition;
7) Update the weights of X_k^i (see Eq. (11));
8) If necessary, resample;
9) Estimate x_k (see Eq. (12));
10) Go to item 3);

IV. Results

Before presenting the results, a particular notation used in this paper is introduced. Consider the matrix $\mathbf{A}$ given by:

$$\mathbf{A} = \begin{bmatrix} a_{11} & [\underline{a}_{12},\ \hat{a}_{12}] \\ [\underline{a}_{21},\ \hat{a}_{21}] & a_{22} \end{bmatrix}$$

The entries of $\mathbf{A}$, noted by $[\underline{a}_{ij},\ \hat{a}_{ij}]$, are continuous random variables with lowest value equals to $\underline{a}_{ij}$ and mean $\underline{a}_{ij} + \hat{a}_{ij}$. This notation will be used as well for uniform distribution as for exponential distribution. The entries noted simply by a_{ij} are deterministic.

Two different systems were considered. For both of them, inputs are considered to be null (autonomous system). The first system is described by the following matrices:

$$\mathbf{A} = \begin{bmatrix} \varepsilon & \varepsilon & \varepsilon & \varepsilon \\ [1,\ 1.5] & 1 & \varepsilon & \varepsilon \\ 2 & [2,\ 1] & [2,\ 2] & \varepsilon \\ \varepsilon & [0,\ 1] & \varepsilon & \varepsilon \end{bmatrix}, \mathbf{C}^T = \begin{bmatrix} \varepsilon \\ \varepsilon \\ 3 \\ [0,\ 3] \end{bmatrix}.$$

In the first simulation, it was considered that the random variables of the system are uniformly distributed. In Tab.I are a estimation for the maximum and the least square error of the particle filter. Note that only the state x_4 has presented nonzero errors. In Fig. 3 one can observe the estimated firing

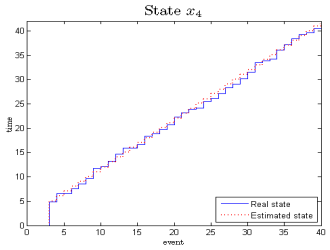

Fig. 3. Firing sequence of transition x_4

sequence of the state x_4 and its real firing sequence. It can be noted that the estimates for this state were very close to the real state.

TABLE I
LEAST SQUARE AND MAXIMUM ERRORS OF FILTER

State	Least Square	Maximum
x_1	0.00	0.00
x_2	0.00	0.00
x_3	0.00	0.00
x_4	0.55	1.04

In the second simulation, it was considered that the random variables of the system are exponentially distributed. Fig. 4 presents the estimated firing sequences and the real firing sequences of the states x_3 and x_4. In Tab. II are a estimation for the maximum and the least square error of the particle filter.

TABLE II
LEAST SQUARE AND MAXIMUM ERRORS OF FILTER

State	Least Square	Maximum
x_1	0.00	0.00
x_2	0.03	0.23
x_3	1.83	13.98
x_4	1.25	3.95

Comparing to the previous simulation, in this simulation, in which were used exponentially distributed random variables, the estimated errors are higher. However, by observing Fig. 4, can be seen that the estimations of the states still close to the real states.

The second system considered, proposed in [11], is described by the following matrices:

$$\mathbf{A} = \begin{bmatrix} \varepsilon & \varepsilon & 4 & \varepsilon & \varepsilon & \varepsilon & 2 & \varepsilon & \varepsilon \\ [1,\ 3] & \varepsilon & \varepsilon & \varepsilon & \varepsilon & \varepsilon & \varepsilon & 3 & \varepsilon \\ \varepsilon & 5 & \varepsilon & \varepsilon & \varepsilon & \varepsilon & \varepsilon & \varepsilon & 1 \\ 4 & \varepsilon & \varepsilon & \varepsilon & \varepsilon & 3 & \varepsilon & \varepsilon & \varepsilon \\ \varepsilon & [3,\ 1] & \varepsilon & [1,\ 1] & \varepsilon & \varepsilon & \varepsilon & \varepsilon & \varepsilon \\ \varepsilon & \varepsilon & 5 & \varepsilon & 4 & \varepsilon & \varepsilon & \varepsilon & \varepsilon \\ \varepsilon & \varepsilon & \varepsilon & 4 & \varepsilon & \varepsilon & \varepsilon & \varepsilon & 3 \\ \varepsilon & \varepsilon & \varepsilon & \varepsilon & 3 & \varepsilon & 5 & \varepsilon & \varepsilon \\ \varepsilon & \varepsilon & \varepsilon & \varepsilon & \varepsilon & 2 & \varepsilon & 4 & \varepsilon \end{bmatrix},$$

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

Considering that the random variables of this system are uniformly distributed, it was applied the particle filter and the results were compared to the results of the *Observer* [12], which determines a lower bound for each state variable. In Fig. 5 can be found the estimated firing sequences for states x_2 and x_5 produced by the particle filter and by the observer as well as the real firing sequences for these states. In Tab. III the estimation errors of the Particle Filter (P.F.) and of the Observer (Obs.) are compared.

TABLE III
COMPARING THE PARTICLE FILTER AND THE OBSERVER

Least Square Error			**Maximum Error**		
Est.	P.F.	Obs.	Est.	P.F.	Obs.
x_1	0.00	0.00	x_1	0.00	0.00
x_2	1.26	2.23	x_2	2.83	5.54
x_3	0.00	0.00	x_3	0.00	0.00
x_4	0.00	0.00	x_4	0.00	0.00
x_5	0.57	2.93	x_5	1.21	6.28
x_6	0.00	0.00	x_6	0.00	0.00
x_7	0.00	0.00	x_7	0.00	0.00
x_8	0.00	0.00	x_8	0.00	0.00
x_9	0.00	0.00	x_9	0.00	0.00

The presented results shows that the estimations produced by the particle filter are, in general, closer to the real state than its lower bound calculated by the Observer.

V. CONCLUSIONS

We have developed a particle filter algorithm for max-plus systems composed of three key steps which are: (1) the sequential importance sampling, (2) resampling and (3)

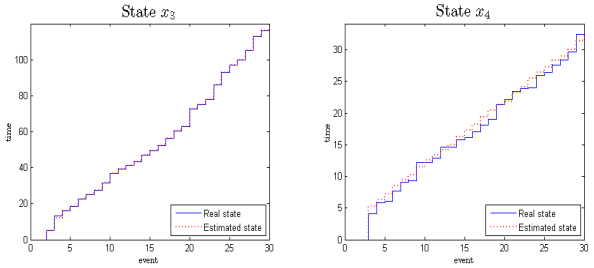

Fig. 4. Firing sequence of transitions x_3 and x_4

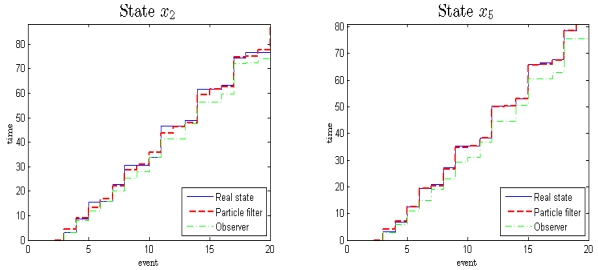

Fig. 5. Firing sequence of transitions x_2 and x_5

reconditioning. The choice of importance density adopted in this paper yielded simplicity to the algorithm since the set of particles for a given value of k can be obtained using the equation of dynamic state (see Eq. (3a)) and the set of particles existing in $k-1$. Although this is a suboptimal choice for the importance density, the results presented shows that the particle filter developed produces estimates close to the real values of the system state. The step of reconditioning the particles allowed us to estimate states associated deterministically with the output, moreover, this step can cause small changes in the density of importance. We are seeking a reconditioning algorithm that interferes even less in the importance density.

APPENDIX

A. Derivation of the likelihood function

Let us consider the max-plus equation $z_k = C \otimes x_k$, where $x \in \mathbb{R}^{n\times 1}$, $z \in \mathbb{R}^{q\times 1}$ and $C \in \mathbb{R}^{q\times n}$. It is assumed that the matrix elements of C, noted by c_{ij}, are independent random variables uniformly distributed between $\underline{c_{ij}}$ and $\overline{c_{ij}}$ or independent random variables distributed exponentially starting from $\underline{c_{ij}}$ with mean $\underline{c_{ij}} + \frac{1}{\lambda}$. The cumulative probability and probability density functions of a random variables with uniform distribution are given respectively by:

$$F_{ij}(\tau) = \begin{cases} 0 & \text{if } \tau \le \underline{c_{ij}} \\ \frac{\tau - \underline{c_{ij}}}{\overline{c_{ij}} - \underline{c_{ij}}} & \text{if } \underline{c_{ij}} < \tau \le \overline{c_{ij}} \\ 1 & \text{if } \tau > \overline{c_{ij}} \end{cases} \quad \text{(A.1)}$$

$$p_{ij}(\tau) = \begin{cases} \frac{1}{\overline{c_{ij}} - \underline{c_{ij}}} & \text{if } \underline{c_{ij}} < \tau \le \overline{c_{ij}} \\ 0 \quad \text{if } \tau \le \underline{c_{ij}} \text{ or } \tau > \overline{c_{ij}} \end{cases} \quad \text{(A.2)}$$

and for a random variables with exponential distribution, the cumulative probability and probability density functions are respectively:

$$F_{ij}(\tau) = \begin{cases} 0 & \text{if } \tau \le \underline{c_{ij}} \\ 1 - e^{-\lambda(\tau - \underline{c_{ij}})} & \text{if } \tau > \underline{c_{ij}} \end{cases} \quad \text{(A.3)}$$

$$p_{ij}(\tau) = \begin{cases} 0 & \text{if } \tau \le \underline{c_{ij}} \\ \lambda e^{-\lambda(\tau - \underline{c_{ij}})} & \text{if } \tau > \underline{c_{ij}} \end{cases} \quad \text{(A.4)}$$

The probability density function of the random variable z conditioned upon the vector x, that is, $p_z(t|x)$, where $t = [t_1 \dots t_q]^T \in \mathbb{R}^{q\times 1}$ is determined in the following [8].

Due to the independence of the elements of C it follows that:

$$P[z \le t] = P[z_1 \le t_1 \ \& \ \dots \ \& \ z_q \le t_q] = \prod_{i=1}^{q} P[z_i \le t_i]$$

Nevertheless:

$$\begin{aligned} P[z_i \le t_i] &= P[\max_j (c_{ij} + x_j) \le t_i] \\ &= P[c_{i1} \le t_i - x_1 \ \& \ \dots \ \& \ c_{in} \le t_i - x_n] \\ &= \prod_{j=1}^{n} F_{ij}(t_i - x_j) \end{aligned}$$

Therefore the cumulative probability function of the random variable z conditioned upon the vector x, is given by:

$$F_z(t|x) = \prod_{i=1}^{q}\prod_{j=1}^{n} F_{ij}(t_i - x_j)$$

By differentiating successively with respect to $t_1 \dots t_q$ one obtains the wanted probability density function:

$$\begin{aligned} p_z(t|x) &= \prod_{i=1}^{q} \frac{\partial}{\partial t_i} (\prod_{j=1}^{n} F_{ij}(t_i - x_j)) \\ &= \prod_{i=1}^{q} (\sum_{j=1}^{n} \frac{\partial}{\partial t_i}(F_{ij}(t_i - x_j)) \prod_{\substack{k=1,\\ k\neq j}}^{n} F_{ik}(t_i - x_k)) \\ &= \prod_{i=1}^{q} (\sum_{j=1}^{n} p_{ij}(t_i - x_j) \prod_{\substack{k=1,\\ k\neq j}}^{n} F_{ik}(t_i - x_k)) \end{aligned}$$

If the vector z is known, this function is called likelihood function of x. In this case, by substituting t by the known value of z, one obtains that:

$$V(x,z) = \prod_{i=1}^{q} (\sum_{j=1}^{n} p_{ij}(z_i - x_j) \prod_{\substack{k=1,\\ k\neq j}}^{n} F_{ik}(z_i - x_k)) \quad \text{(A.5)}$$

REFERENCES

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

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

[3] V. E. Beneš, "Exact finite dimensional filters for certain diffusions with nonlinear drift," *Stochastics*, vol. 5, 1981.

[4] F. Daum, "Exact finite-dimensional nonlinear filters," in *Automatic Control, IEEE Transactions on*, vol. 31, no. 7, jul 1986, pp. 616 – 622.

[5] ——, "Beyond kalman filters: practical design of nonlinear filters," *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, vol. 2561, pp. 252–262, sep 1995.

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

[7] F. Baccelli, G. Cohen, G. J. Olsder, and J.-P. Quadrat, *Synchronization and Linearity: An Algebra for Discrete Event Systems*, ser. Wiley series in probability and mathematical statistics: Probability and mathematical statistics. Wiley, 1992.

[8] D. F. e Silva, R. S. Mendes, L. Hardouin, C. A. Maia, and B. Cottenceau, "Filtragem estocástica aplicada a sistemas max-plus lineares," *Sociedade Brasileira de Automática*, 2011.

[9] A. Doucet, S. Godsill, and C. Andrieu, "On sequential monte carlo sampling methods for bayesain filtering," *Statistics and Computing*, vol. 10, 2000.

[10] A. Kong, J. S. Liu, and W. H. Wong, "Sequential imputations and bayesian missing data problems," *Journal of the American Statistical Association*, vol. 89, no. 425, Mar. 1994.

[11] M. D. Loreto, S. Gaubert, R. D. Katz, and J. Loiseau, "Duality between invariant spaces for max-plus linear discret event systems," *SIAM J. on Control and Optimaztion*, 2010.

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