Delft University of Technology

Delft Center for Systems and Control

Technical report 12-025

Model predictive control for stochastic switching max-plus-linear systems*

S. van Loenhout, T. van den Boom, S. Farahani, and B. De Schutter

If you want to cite this report, please use the following reference instead:
S. van Loenhout, T. van den Boom, S. Farahani, and B. De Schutter, "Model predictive control for stochastic switching max-plus-linear systems," *Proceedings of the 11th International Workshop on Discrete Event Systems (WODES 2012)* (A. Ramírez Treviño, E. López Mellado, J.J. Lesage, and M. Silva, eds.), Guadalajara, Mexico, pp. 79–84, Oct. 2012.

Delft Center for Systems and Control
Delft University of Technology
Mekelweg 2, 2628 CD Delft
The Netherlands
phone: +31-15-278.51.19 (secretary)
fax: +31-15-278.66.79
URL: http://www.dcsc.tudelft.nl

Model Predictive Control for Stochastic Switching Max-Plus-Linear Systems

Stefan van Loenhout * **Ton van den Boom** * **Samira Farahani** *
Bart De Schutter *

* *Delft Center for Systems and Control, Delft University of Technology, Delft, The Netherlands (e-mail: {a.j.j.vandenboom,s.safaeifarahani,b.deschutter}@tudelft.nl)*

Abstract: A switching max-plus-linear system can operate in different modes. In each mode the system is described by a max-plus linear system equation. The switching may depend on the previous mode, on the state, and on the input. *Stochastic* switching max-plus-linear systems may include two types of stochastic uncertainty, namely stochastic parametric uncertainty and stochastic mode switching uncertainty. For both types of uncertainty results have appeared in the literature. In this paper we will consider stochastic switching max-plus-linear systems with parametric uncertainty and mode switching uncertainty in one single unified framework. First we derive a (general) model that includes both types of stochastic uncertainty. Next a model predictive control method is used to control the system, and we distinguish between the case where the two types of uncertainty are dependent or independent.

Keywords: max-plus systems, stochastic systems, mode switching, model predictive control

1. INTRODUCTION

Recently, perturbed max-plus-linear (MPL) systems have drawn some considerable attention. MPL systems are a subclass of discrete event systems (DES). In general DES are man-made systems that contain a finite number of resources (such as machines, communication channels, or processors) that are shared by several users (such as product types, information packages, or jobs) all of which contribute to some common goal (the assembly of products, the end-to-end transmission of a set of information packages, or a parallel computation) (Baccelli et al., 2001). The class of MPL systems consists of DES with synchronization but no choice. DES are in general nonlinear in conventional algebra. However, the subclass of MPL systems can be written in a linear form in the max-plus algebra (Cuninghame-Green, 1979). In (van den Boom and De Schutter, 2006) the class of MPL systems has been extended to the class of switching MPL systems. This class consists of DES that can switch between different modes of operation. In each mode the system is described by a max-plus-linear state equation and a max-plus-linear output equation, with different system matrices for each mode.

In contrast to conventional linear systems, where noise is often considered to be additive, the influence of noise and disturbances on MPL systems is multiplicative (in the max-plus-algebraic sense). This results in perturbed system matrices, and as a consequence the system properties may change over the events. For perturbed MPL systems, the perturbations are in general related to processing and transportation times, which in this paper are assumed to be stochastic quantities. This allows us to use more information about the uncertainties.

For uncertain switching MPL systems (van den Boom and De Schutter, 2011) the mode switching depends on a switching mechanism that is modeled with probabilities that may depend on (a combination of) the following four variables: the previous mode, the previous state, the input signal, or an (auxiliary) control signal.

In (De Schutter and van den Boom, 2001; van den Boom and De Schutter, 2006), the methodology from model predictive control (MPC) (Rawlings and Mayne, 2009; Maciejowski, 2002) has been applied and translated to (switching) MPL systems. MPC depends on the availability of a prediction model. The MPC controller solves an optimization problem at each event step, which results in the optimal control action. Current literature has already investigated several subclasses of perturbed MPL systems. In fact, we can distinguish two main research directions. In the first one, we assume that some parameters (usually time-based parameters) vary over the events (Farahani et al., 2010; van den Boom and De Schutter, 2004). In the other direction, we consider a system that can switch between several modes of operation (van den Boom and De Schutter, 2006, 2010, 2011). In this paper we incorporate both the parametric uncertainty (van den Boom and De Schutter, 2004) as the mode switching uncertainty (van den Boom and De Schutter, 2011) simultaneously.

This paper proposes a unified setting for stochastic switching MPL systems, viz. a setting that includes both types of uncertainty. This approach allows us to use more information about the uncertainties, resulting in a less conservative controller (since it no longer assumes the worst-case scenario at every event step, but rather the average-case scenario). First, we define an extended MPL model that includes both uncertainties in a stochastic setting. Then, we will analyze the design of the model predictive controller.

The main computational issue with this approach will be computing the cost function. Since the model contains two types of uncertainty, we need to determine the joint density function. We propose several classifications for which the computation of the expectation of the output cost function can be simplified. However, even for the simplified optimization problem, the computation time is still a limiting factor (since we need to perform numerical integration over all the uncertainties). Therefore, we will use some approximation methods that reduce the computation time significantly (at the cost of introducing a small error in obtaining the optimal control action).

This paper is organized as follows. In Section 2, we give a concise introduction to stochastic MPL systems. In Section 3, we analyze how we can transform the principles of MPC to stochastic switching MPL systems, and the main difficulties that are encountered. In Section 4, we elaborate on the computation of the expectation of the cost function.

2. MAX-PLUS ALGEBRA AND MAX-PLUS-LINEAR SYSTEMS

2.1 Max-Plus Algebra

The mathematical framework behind MPL systems is the max-plus algebra (Cuninghame-Green, 1979; Baccelli et al., 2001; Heidergott et al., 2006). Define $\varepsilon = -\infty$ and $e = 0$ as the max-plus algebraic zero and identity element respectively. Define the set $\mathbb{R}_{\max} = \mathbb{R} \cup \{\varepsilon\}$. Then, for $a, b \in \mathbb{R}_{\max}$, the operators are defined as:

$$a \oplus b = \max(a, b)$$
$$a \otimes b = a + b$$

The corresponding matrix operators are defined as:

$$[A \oplus B]_{ij} = a_{ij} \oplus b_{ij} = \max(a_{ij}, b_{ij})$$
$$[A \otimes C]_{ij} = \bigoplus_{k=1}^{n} \left(a_{ik} \otimes b_{kj}\right) = \max_{k=1,\ldots,n} (a_{ik} + b_{kj})$$

for $A, B \in \mathbb{R}_{\max}^{m \times n}$ and $C \in \mathbb{R}_{\max}^{n \times p}$.

2.2 MPL Systems and Uncertainty

DES with synchronization (which requires the availability of several resources at the same time) but no choice (which appears when some user must choose among several resources) can be modeled as (Cuninghame-Green, 1979; Baccelli et al., 2001; Heidergott et al., 2006):

$$x(k) = A \otimes x(k-1) \oplus B \otimes u(k) \tag{1a}$$
$$y(k) = C \otimes x(k) \tag{1b}$$

where the index k is called the event counter. Here, $x(k)$ contains time information about the internal events, while $u(k)$ and $y(k)$ contain time information about the input and output events respectively.

A switching MPL model was introduced in (van den Boom and De Schutter, 2011):

$$x(k) = A_{\ell(k)} \otimes x(k-1) \oplus B_{\ell(k)} \otimes u(k) \tag{2a}$$
$$y(k) = C_{\ell(k)} \otimes x(k) \tag{2b}$$

where we denote the mode of the system at the k-th event as $\ell(k) \in \{1, \ldots, n_L\}$, where n_L is the number of modes. This means that for each mode ℓ, the system behavior is described by an MPL state equation (2a) and an MPL output equation (2b) with system matrices A_ℓ, B_ℓ, and C_ℓ.

2.3 Stochastic Switching MPL Systems

In stochastic switching MPL systems we consider two stochastic phenomena:

(1) uncertainty in the switching behavior (van den Boom and De Schutter, 2011),
(2) parametric uncertainty in the system parameters (van den Boom and De Schutter, 2004).

Up to now we have always considered the two types of uncertainty separately. The most obvious path to include both uncertainties will be to enhance (1) with the mode switching uncertainty first, and then to include the parametric uncertainty for each mode independently, resulting in a model of the form

$$x(k) = A_{\ell(k)}(e(k)) \otimes x(k-1) \oplus B_{\ell(k)}(e(k)) \otimes u(k) \tag{3a}$$
$$y(k) = C_{\ell(k)}(e(k)) \otimes x(k) \tag{3b}$$

where $A_{\ell(k)}(e(k))$, $B_{\ell(k)}(e(k))$, and $C_{\ell(k)}(e(k))$ are the system matrices that correspond to mode $\ell(k)$ and where the stochastic random vector $e(k)$ represents the stochastic parametric uncertainty at the k-th event step. Since the model description (1) is written in the max-plus algebra, the noise and modeling errors are max-plus multiplicative[1].

The switching between the different modes can be modeled with so called switching probabilities (van den Boom and De Schutter, 2011):

$$P\big[L = \ell(k) \mid \ell(k-1), x(k-1), u(k), v(k), e(k)\big]$$

where L is a stochastic random variable, $\ell(k)$ is its value (in fact the mode that the model will switch to), $\ell(k-1)$ is the previous mode (the mode that the system will switch from), $x(k-1)$ is the previous state, $u(k)$ is the current input, and $v(k)$ is an auxiliary control signal that directly affects the switching behavior. Note that the switching probability may depend on the parametric uncertainty vector $e(k)$.

3. MODEL PREDICTIVE CONTROL

In (De Schutter and van den Boom, 2001), the MPC framework had been extended to deterministic MPL models. Assume we have a model as in (3), with system matrices $A_{\ell(k)}(e(k)) \in \mathbb{R}_{\max}^{n \times n}$, $B_{\ell(k)}(e(k)) \in \mathbb{R}_{\max}^{n \times n_u}$, and $C_{\ell(k)}(e(k)) \in \mathbb{R}_{\max}^{n_y \times n}$, where n is equal to the internal state dimension, while n_y and n_u are equal to the number of outputs and inputs respectively.

[1] In contrast to conventional linear systems, where noise and disturbances are usually modeled by including an extra term in the system equations (i.e., the noise is considered to be additive), the influence of noise and disturbances in MPL DES appear as an additional term to the system parameters, i.e. the entries of the system matrices. Addition in conventional algebra means multiplication in max-plus algebra. Hence, noise and disturbances in MPL systems are not max-plus-additive, but max-plus-multiplicative.

3.1 Cost Function

Since we will consider a stochastic system, the cost criterion is given by an expectation:

$$\mathbb{E}\big[J(k)\big] = \sum_{j=0}^{N_\mathrm{p}-1} \sum_{i=1}^{n_y} \mathbb{E}\big[\kappa_i(k+j)\big] - \lambda \sum_{j=0}^{N_\mathrm{p}-1} \sum_{l=1}^{n_u} u_l(k+j) \tag{4}$$

where $\mathbb{E}\big[\ \cdot\ \big]$ denotes the expectation of some random variables and $\kappa_i = \max(y_i(k) - r_i(k), 0)$ is the tardiness error for the i-th output at the k-th event step. This tardiness error penalizes the late (but not the early) deliveries for the i-th output at the k-th event step, where $r(k)$ is a given vector of reference (due date) signals. The second term in cost function (4) maximizes the input instants, which results in minimum input buffer levels.

3.2 Prediction Model

As observed in the previous subsection, we need to express the future output $y(k+j)$ as a function of the future input $u(k+j)$, for $j = 0, \ldots, N_\mathrm{p} - 1$. Since the model (3) is available, it is quite straightforward to compute the future output signal by successive substitution. Note that switching MPL systems are different from conventional time-driven systems in the sense that the event counter k is not directly related to a specific time. So far we have assumed that at event step k the state $x(k)$ is available (recall that $x(k)$ contains the time instants at which the internal activities or processes of the system start for the k-th cycle). Let t be the time instant when a prediction has to be done. In this paper we can define the initial cycle k as follows:

$$k = \arg\max \Big\{ l | x_i(l-1) \leq t\ ,\ \ \forall i \in \{1, 2, \ldots, n\} \Big\}$$

This means that state $x(k-1)$ is completely known. Now consider the following vectors:

$$\begin{aligned}
\tilde{x}(k) &= \big[x^T(k)\ \ldots\ x^T(k+N_\mathrm{p}-1)\big]^T \\
\tilde{u}(k) &= \big[u^T(k)\ \ldots\ u^T(k+N_\mathrm{p}-1)\big]^T \\
\tilde{y}(k) &= \big[y^T(k)\ \ldots\ y^T(k+N_\mathrm{p}-1)\big]^T \\
\tilde{e}(k) &= \big[e^T(k)\ \ldots\ e^T(k+N_\mathrm{p}-1)\big]^T \\
\tilde{\ell}(k) &= \left[\ell(k)\ \ldots\ \ell(k+N_\mathrm{p}-1)\right]^T
\end{aligned}$$

Define the matrices:

$$\tilde{C}(\tilde{\ell}(k), \tilde{e}(k)) = \begin{bmatrix} \tilde{C}_1(\tilde{\ell}(k), \tilde{e}(k)) \\ \vdots \\ \tilde{C}_{N_\mathrm{p}}(\tilde{\ell}(k), \tilde{e}(k)) \end{bmatrix}$$

$$\tilde{D}(\tilde{\ell}(k), \tilde{e}(k)) = \begin{bmatrix} \tilde{D}_{1,1}(\tilde{\ell}(k), \tilde{e}(k)) \cdots & \mathcal{E} \\ \vdots \quad \ddots & \vdots \\ \tilde{D}_{N_\mathrm{p},1}(\tilde{\ell}(k), \tilde{e}(k)) \cdots \tilde{D}_{N_\mathrm{p},N_\mathrm{p}}(\tilde{\ell}(k), \tilde{e}(k)) \end{bmatrix}$$

where

$$\begin{aligned}
\tilde{C}_m(\tilde{\ell}(k), \tilde{e}(k)) &= C_{\ell(k+m-1)}(e(k+m-1)) \\
&\otimes A_{\ell(k+m-1)}(e(k+m-1)) \otimes \ldots \otimes A_{\ell(k)}(e(k)) \\
\tilde{D}_{m,n}(\tilde{\ell}(k), \tilde{e}(k)) &= C_{\ell(k+m-1)}(e(k+m-1)) \\
&\otimes A_{\ell(k+m-1)}(e(k+m-1)) \otimes \ldots \otimes A_{\ell(k+n)}(e(k+n)) \\
&\otimes A_{\ell(e(k+n-1))}(e(k+n-1))
\end{aligned}$$

Using the recursion from (3), we can rewrite the future output $y(k+j)$ as a function of $u(k+i)$ for $i = 0, \ldots, j$ and $x(k-1)$:

$$\tilde{y}(k) = \tilde{C}(\tilde{\ell}(k), \tilde{e}(k)) \otimes x(k-1) \oplus \tilde{D}(\tilde{\ell}(k), \tilde{e}(k)) \otimes \tilde{u}(k)$$

3.3 Constraints

Since MPC solves an optimization problem, it becomes fairly easy to incorporate constraints on the (predicted) model. Since the inputs $u(k+j)$ corresponds to consecutive event times, we should have:

$$\Delta u(k+j) = u(k+j) - u(k+j-1) \geqslant 0$$

for $j = 0, \ldots, N_\mathrm{p} - 1$. Furthermore, in order to reduce the number of decision variables (and thus the computational complexity), we introduce a control horizon N_c ($\leqslant N_\mathrm{p}$). We impose that the input rate should be constant from the point $k + N_\mathrm{c} - 1$ on:

$$\Delta^2 u(k+j) = \Delta u(k+j) - \Delta u(k+j-1) = 0$$

for $j = N_\mathrm{c}, \ldots, N_\mathrm{p} - 1$. Additionally, we can add constraints on the future input and output event times as well.

3.4 Optimization Problem

The aim of the optimization problem is to compute an optimal control signal $\tilde{u}(k)$ that minimizes $\mathbb{E}\left[J(k)\right]$ subject to linear constraints defined in the previous subsection. The final optimization problem at event step k can now be defined as:

$$\min_{\tilde{u}(k)} \sum_{i=1}^{N_\mathrm{p} n_y} \mathbb{E}\left[\max(\tilde{y}_i(k) - \tilde{r}_i(k), 0)\right] - \lambda \sum_{i=1}^{N_\mathrm{p} n_u} \tilde{u}_i(k)$$

subject to

$$\begin{aligned}
&\tilde{y}(k) = \tilde{C}(\tilde{\ell}(k), \tilde{e}(k)) \otimes x(k-1) \oplus \tilde{D}(\tilde{\ell}(k), \tilde{e}(k)) \otimes \tilde{u}(k) \\
&\Delta u(k+j) \geqslant 0 \quad \text{for} \quad j = 0, \ldots, N_\mathrm{c} - 1 \\
&\Delta^2 u(k+j) = 0 \quad \text{for} \quad j = N_\mathrm{c}, \ldots, N_\mathrm{p} - 1 \\
&A_\mathrm{con}\tilde{u}(k) + B_\mathrm{con}\mathbb{E}[\tilde{y}(k)] \leqslant c_\mathrm{con}(k)
\end{aligned}$$

This optimization problem can be solved in various ways. The computational complexity of this optimization problem depends on the complexity of the computation of expectation of the cost function. This will be discussed in the next section.

4. EXPECTATION OF THE COST FUNCTION

The most extensive computation problem in the previous section is the computation of the expectation of the cost function. Since we have a finite number of modes (n_L), there is a finite number of possible mode switching sequences over the whole prediction horizon (in fact we have $(n_L)^{N_\mathrm{p}}$ possibilities), and thus the mode switching uncertainty is a discrete random variable. On the other hand, the parametric uncertainty is related to processing and transportation times, which means that the parametric uncertainty is a continuous random variable. This means that for the computation of the expectation of the cost function we need to perform a combination of integration over the stochastic variable $\tilde{e}(k)$ and summation over the stochastic variable $\tilde{\ell}(k)$.

4.1 Joint Probability Density Function

So, since we will need to compute the expectation of a function ($J(k)$) which is a function of two types of random variables, we need to find an expression for the joint probability density function (jdf). However, this jdf is not available directly. Therefore, we can use conditional probability theory to write it in a more usable way. The jdf is denoted by $f_{\mathcal{L},\mathcal{E}}$, where $\mathcal{L}$ is the (discrete) sample space of all the mode switching sequences $\tilde{\ell}(k)$, and $\mathcal{E}$ is the (continuous) sample space of the parametric uncertainty $\tilde{e}(k)$.

We assume that the pdf of the uncertainty $\tilde{e}(k)$ does not depend on the discrete mode sequence $\tilde{\ell}(k)$ and the continuous probability density function (pdf) of $\tilde{e}(k)$ is given by $f_{\mathcal{E}}(\tilde{e}(k))$. With use of conditional probability theory (Kingman and Taylor, 1966), we now have

$$f_{\mathcal{L},\mathcal{E}}(\tilde{\ell}(k),\tilde{e}(k)) = f_{\mathcal{E}}(\tilde{e}(k))\tilde{P}\big[L=\tilde{\ell}(k) \mid E=\tilde{e}(k)\big]$$

where $\tilde{P}\big[L=\tilde{\ell}(k) \mid E=\tilde{e}(k)\big]$ is the probability that we have mode switching sequence $\tilde{\ell}(k)$, given the parametric uncertainty $\tilde{e}(k)$.

When both the continuous pdf $f_{\mathcal{E}}(\tilde{e}(k))$ and the mode switching probability $\tilde{P}\big[L=\tilde{\ell}(k) \mid E=\tilde{e}(k)\big]$ are known, we can define an expression for the expectation of the cost function:

$$\mathbb{E}\big[J(k)\big] = \sum_{\tilde{\ell}(k)\in\mathcal{L}} \Big[\int_{\mathcal{E}} J(\tilde{\ell}(k),\tilde{e}(k)) f_{\mathcal{E}}(\tilde{e}(k)) \; \tilde{P}\big[L=\tilde{\ell}(k) \mid E=\tilde{e}(k)\big]\,\mathrm{d}\tilde{e}(k)\Big] \quad (5)$$

4.2 Switching Probability

However, we still need to obtain an expression for the mode switching probability, denoted here by $\tilde{P}$. Recall that the mode switching probability for a single mode may depend on the previous mode $\ell(k-1)$, the previous mode $x(k-1)$, the current input $u(k)$, and an auxiliary control signal $v(k)$. Hence,

$$\tilde{P}\big[L=\tilde{\ell}(k) \mid \ell(k-1), x(k-1), \tilde{u}(k), \tilde{v}(k), \tilde{e}(k)\big] = \prod_{j=0}^{N_{\mathrm{p}}-1} P\big[L=\ell(k+j) \mid \ell(k+j-1), x(k+j-1), u(k+j), v(k+j), e(k+j)\big]$$

We are interested in computing the expected value $\mathbb{E}\big[J(k)\big]$ in the optimization phase of the MPC algorithm. In every optimization step we may assume the values $\ell(k-1)$, $x(k-1)$ to be known, just as the values of $\tilde{u}(k)$ and $\tilde{v}(k)$, which are proposed by the optimizer. Note that the dependence on these "known" values is sometimes dropped in the definition of the mode switching probability, just for brevity, and we write

$$\tilde{P}\big[L=\tilde{\ell}(k) \mid \ell(k-1), x(k-1), \tilde{u}(k), \tilde{v}(k), \tilde{e}(k)\big] = \tilde{P}\big[L=\tilde{\ell}(k) \mid E=\tilde{e}(k)\big]$$

4.3 Case 1: Completely Independent Random Variables

In this case we assume that the mode switching uncertainty $\tilde{\ell}(k)$ does not depend on the parametric uncertainty $\tilde{e}(k)$, so $\tilde{P}\big[L=\tilde{\ell}(k) \mid E=\tilde{e}(k)\big] = \tilde{P}\big[L=\tilde{\ell}(k)\big]$. Represent the set $\mathcal{L}$ of all possible consecutive mode switching vectors as $\mathcal{L} = \{\tilde{\ell}^1, \tilde{\ell}^2, \ldots, \tilde{\ell}^M\}$ for $M = (n_L)^{N_{\mathrm{p}}}$. Now, we can compute the cost function in a more efficient way.

Theorem 1. If the parametric uncertainty $\tilde{\ell}(k)$ and the mode switching uncertainty $\tilde{e}(k)$ are independent, we can rewrite the expectation of the cost function as:

$$\begin{aligned}\mathbb{E}\big[J(k)\big] &= \mathbb{E}_{\tilde{e},\tilde{\ell}}\big[J(k)\big] \\ &= \sum_{\tilde{\ell}(k)\in\mathcal{L}} \Big[\tilde{P}\big[L=\tilde{\ell}(k)\big]\mathbb{E}_{\tilde{e}}\big[J(\tilde{\ell}(k),\tilde{e}(k))\big]\Big] \\ &= \sum_{m=1}^{M} \Big[\tilde{P}\big[L=\tilde{\ell}^m(k)\big]\mathbb{E}_{\tilde{e}}\big[J(\tilde{\ell}^m(k),\tilde{e}(k))\big]\Big]\end{aligned} \quad (6)$$

where $\mathbb{E}_{\tilde{e}}\big[J(\tilde{\ell}^m(k),\tilde{e}(k)\big]$ is the expectation of the cost function after substitution of a given mode switching sequence $\tilde{\ell}^m(k) \in \mathcal{L}$.

Proof. Since $\tilde{\ell}(k)$ does not depend on $\tilde{e}(k)$, we have:

$$\tilde{P}\big[L=\tilde{\ell}^m(k) \mid E=\tilde{e}(k)\big] = \tilde{P}\big[L=\tilde{\ell}^m(k)\big]$$

Thus, after substitution, we obtain:

$$\mathbb{E}\big[J(k)\big] = \sum_{m=1}^{M} \Big[\int_{\mathcal{E}} J(\tilde{\ell}^m(k),\tilde{e}(k)) f_{\mathcal{E}}(\tilde{e}(k)) \; \tilde{P}\big[L=\tilde{\ell}^m(k)\big]\,\mathrm{d}\tilde{e}(k)\Big]$$

Since the mode switching probability $\tilde{P}$ does not depend on $\tilde{e}$, we can write it outside the integral.

$$\mathbb{E}\big[J(k)\big] = \sum_{m=1}^{M} \Big[\tilde{P}\big[L=\tilde{\ell}^m(k)\big] \int_{\mathcal{E}} J(\tilde{\ell}^m(k),\tilde{e}(k)) f_{\mathcal{E}}(\tilde{e}(k))\,\mathrm{d}\tilde{e}(k)\Big]$$

where the integral is equal to $\mathbb{E}_{\tilde{e}}\big[J(\tilde{\ell}^m(k),\tilde{e}(k)\big]$. $\square$

We can observe that we have split the problem up in two sub-problems, involving respectively an expected value computation over a discrete stochastic variable and an expected value computation over a continuous stochastic variable. The above derivations show how we can deal with the mode switching uncertainty in the first place. Secondly, we have to compute the expectation over the parametric uncertainty only. This approach simplifies the computation of the expectation of the cost function significantly. However, we are still facing the computational complexity of both sub-problems. Since there are $(n_L)^{N_{\mathrm{p}}}$ possible mode switching sequences, the size of the set $\mathcal{L}$ grows very fast as we increase either n_L or N_{p}. On the other hand, we need to (numerically) integrate over all the uncertainties $\mathrm{d}\tilde{e}(k) \in \mathcal{E}$. This is very time-consuming as well. However, since by Theorem 1 we have split the computation of $\mathbb{E}\big[J(k)\big]$ into two sub-problems, we can apply approximations to each of these sub-problems separately.

At first, we can reduce the computational complexity of the mode switching uncertainty. In general, we may know that some mode switching sequences are more likely to occur than others. Therefore, we can choose to neglect some mode switching sequences that are not likely to occur, resulting in a reduced set $\mathcal{L}^{\text{red}}$ that should be used in (6) instead of $\mathcal{L}$. In (van den Boom and De Schutter, 2010) this approximation approach is presented. In fact, in this way we can often significantly reduce the number of terms in the sum in (6), and thus the computational complexity as well, while still maintaining an adequate cumulative probability of these mode switching sequences to occur.

On the other hand, we can approximate the numerical integration over all the uncertainties with the use of vector norms. Since within $\mathbb{E}_{\tilde{e}}\big[J(\tilde{\ell}^m(k), \tilde{e}(k)\big]$ we will be computing the expectation of the maximum of some random variables (cf. (4)), we can use the following approximation (Farahani et al., 2010):

$$\begin{aligned}\mathbb{E}\big[\max(x_1, \ldots, x_n)\big] &\leq \mathbb{E}\big[\max(|x_1|, \ldots, |x_n|)\big] \\ &\leq \mathbb{E}\big[(|x_1|^p + \ldots + |x_n|^p)^{1/p}\big] \\ &\leq \left(\sum_{j=1}^{n} \mathbb{E}\big[\,|x_j|^p\,\big]\right)^{1/p}\end{aligned}$$

for some positive integer p. If we drop the absolute value sign, assuming p to be an even number, the term $\mathbb{E}\big[\,|x_j|^p\,\big]$ is equal to the p-th moment of a random variable. There are several probability distributions for which the random variables have finite moments and a closed-form expression of these moments exists (such as the uniform distribution, the normal distribution, the Beta distribution, etc.). More details, and the exact derivation can be found in (Farahani et al., 2010). In (Farahani et al., 2010; van den Boom and De Schutter, 2004), it has been proved that both the full problem as well as the approximation are convex optimization problems [2].

Case 2: Dependent Random Variables

In case the mode switching uncertainty $\tilde{\ell}(k)$ depends on the parametric uncertainty $\tilde{e}(k)$, the problem becomes much more difficult since then we have to use the expression given in (5), which is repeated here for easy reference:

$$\begin{aligned}\mathbb{E}\big[J(k)\big] = \sum_{m=1}^{M} \Big[\int_{\mathcal{E}} & J(\tilde{\ell}^m(k), \tilde{e}(k)) f_{\mathcal{E}}(\tilde{e}(k)) \\ & \tilde{P}\big[L = \tilde{\ell}^m(k) \mid E = \tilde{e}(k)\big]\, \mathrm{d}\tilde{e}(k)\Big]\end{aligned}$$

The approximation theory from the previous section cannot be applied now. However, if we assume that the conditional mode switching probabilities P as well as the probability density function $f_{\mathcal{E}}$ are modeled or *approximated* [3] as multi-variable piecewise polynomial functions, possibly multiplied by an exponential, that are defined on polyhedral regions, then we can proceed as follows. Recall that J is a summation of maximum operators and linear terms (cf. (4)). Hence, J is a piecewise affine function (defined on polyhedral regions) of $\tilde{e}(k)$ and as a consequence, it is thus of the same (but more simple) form as P and $f_{\mathcal{E}}$. So if we combine the regions of P, $f_{\mathcal{E}}$, and J and multiply the piecewise polynomial and piecewise affine functions and the exponentials, we find that there exists a polyhedral partition [4] $\{\mathcal{R}_i\}_{i=1}^{n_{\mathcal{R}}}$ of $\mathcal{E}$ such that

$$\begin{aligned}J(\tilde{\ell}^m(k), \tilde{e}(k))\, f_{\mathcal{E}}(\tilde{e}(k))\, \tilde{P}\big[L = \tilde{\ell}^m(k) \mid E = \tilde{e}(k)\big] \\ = \zeta_{im}(\tilde{e}(k)) \quad \text{when } \tilde{e}(k) \in \mathcal{R}_i\end{aligned}$$

where ζ_{im} is a function of the form

$$\zeta_{im}(\tilde{e}) = \sum_{j=1}^{n_{im}} \xi_{imj} \left(\prod_{l=1}^{n_{\tilde{e}}} \tilde{e}_l^{k_{imjl}} \exp(\nu_{imjl}\tilde{e}_l)\right)$$

for some real-valued constants ξ_{imj} and ν_{imjl}, some positive integers n_{im}, and some non-negative integers $k_{imj,}$, and where $n_{\tilde{e}}$ is equal to the number of components of $\tilde{e}$. As a result we find

$$\mathbb{E}\big[J(k)\big] = \sum_{m=1}^{M} \sum_{i=1}^{n_{\mathcal{R}}} \int_{\mathcal{R}_i} \zeta_{im}(\tilde{e}(k))\mathrm{d}\tilde{e}(k)$$

Since each $\mathcal{R}_i$ is a polyhedron, we can do a substitution of variables by expressing an arbitrary point $\tilde{e}(k) \in \mathcal{R}_i$ as a linear, non-negative, and/or convex combination of the central generators [5], extreme rays, and finite vertices of $\mathcal{R}_i$ (which can be computed using, e.g., the double-description introduced by Motzkin et al. (1953)). As a result

$$\int_{\mathcal{R}_i} \zeta_{im}(\tilde{e}(k))\mathrm{d}\tilde{e}(k)$$

reduces to the repeated integration of a polynomial function, possibly multiplied by an exponential. This integration can be done analytically.

5. DISCUSSION

In this paper we have studied switching max-plus-linear systems with both stochastic parameter uncertainty and stochastic switching uncertainty in a unified framework. To describe the stochastic behavior of the system we have considered the joint probability distribution with both continuous and discrete random variables. For the computation of the expectation of the cost function of the model predictive control problem we can distinguish two cases. In the first case we assume the two random variables to be completely independent. In this case we can split the joint probability function into two parts, and we can use previously derived approximation methods to compute the expectation of the cost function. In the second case the switching probability depends on the parametric random variable and the computation becomes more complicated. If the switching probability is a piecewise polynomial function, possibly multiplied with an exponential, and the same holds for the probability density function of the parametric random variable, we can determine an expression for the computation of the expected value cost function using repeated analytic integration.

ACKNOWLEDGMENTS

Research partially funded by the Dutch Technology Foundation STW project "Model-predictive railway traffic management – A framework for closed-loop control of large-scale railway systems" and by the European Union

[2] Since the sum-operator and the expectation are additive operators, the convexity is preserved.

[3] Note that by considering enough regions we can in general approximate the real probabilities or the real probability density function arbitrarily close.

[4] I.e., the regions $\mathcal{R}_i$ are non-empty and mutually disjoint, and their union equals $\mathcal{E}$.

[5] I.e., basis vectors for the lineality space associated to the polyhedron (Schrijver, 1986).

Seventh Framework Programme [FP7/2007-2013] under grant agreement no. 257462 HYCON2 Network of Excellence.

REFERENCES

Baccelli, F., Cohen, G., Olsder, G., and Quadrat, J. (2001). *Synchronization and Linearity*. John Wiley & Sons.

Cuninghame-Green, R. (1979). *Minimax Algebra*, volume 166 of *Lecture Notes in Economics and Mathematical Systems*. Springer-Verlag, Berlin, Germany.

De Schutter, B. and van den Boom, T. (2001). Model predictive control for max-plus-linear discrete event systems. *Automatica*, 37(7), 1049–1056.

Farahani, S., van den Boom, T., van der Weide, H., and De Schutter, B. (2010). An approximation approach for model predictive control of stochastic max-plus linear systems. In *Proceedings of the 10th International Workshop on Discrete Event Systems (WODES 2010)*, 386–391. Berlin, Germany.

Heidergott, B., Olsder, G., and van der Woude, J. (2006). *Max Plus at Work*. Princeton University Press, Princeton, New Jersey.

Kingman, J. and Taylor, S. (1966). *Introduction to Measure and Probability*. Cambridge University Press.

Maciejowski, J. (2002). *Predictive Control with Constraints*. Prentice Hall, Harlow, England.

Motzkin, T., Raiffa, H., Thompson, G., and Thrall, R. (1953). The double description method. In H. Kuhn and A. Tucker (eds.), *Contributions to the Theory of Games*, number 28 in *Annals of Mathematics Studies*, 51–73. Princeton University Press, Princeton, New Jersey.

Rawlings, J. and Mayne, D. (2009). *Model Predictive Control: Theory and Design*. Nob Hill Publishing, Madison, Wisconsin.

Schrijver, A. (1986). *Theory of Linear and Integer Programming*. John Wiley & Sons, Chichester, UK.

van den Boom, T. and De Schutter, B. (2004). Model predictive control for perturbed max-plus-linear systems: A stochastic approach. *International Journal of Control*, 77(3), 302–309.

van den Boom, T. and De Schutter, B. (2006). Modelling and control of discrete event systems using switching max-plus-linear systems. *Control Engineering Practice*, 14(10), 1199–1211.

van den Boom, T. and De Schutter, B. (2010). Model predictive control for randomly switching max-plus-linear systems using a scenario-based algorithm. In *Proceedings of the 49th IEEE Conference on Decision and Control*, 2298–2303. Atlanta, Georgia.

van den Boom, T. and De Schutter, B. (2011). Modeling and control of switching max-plus-linear systems with random and deterministic switching. *Discrete Event Dynamic Systems*.