Technical report 04-001

Modelling and control of discrete event systems using switching max-plus-linear systems

T.J.J. van den Boom and B. De Schutter

If you want to cite this report, please use the following reference instead:
MODELLING AND CONTROL OF DISCRETE EVENT SYSTEMS
USING SWITCHING MAX-PLUS-LINEAR SYSTEMS

T.J.J. van den Boom¹ and B. De Schutter¹,
Delft Center for Systems and Control, Delft University of Technology,
Mekelweg 2, 2628 CD Delft, The Netherlands,

Abstract: In this paper we consider the modelling and control of discrete event systems using switching max-plus-linear systems. In switching max-plus-linear systems we can switch between different modes of operation. In each mode the discrete event system is described by a max-plus-linear state equation with different system matrices for each mode. The switching allows us to change the structure of the system, to break synchronization and to change the order of events. We will give some examples of this type of system. We define the model predictive control design problem for this type of discrete event system, and we show that solving this problem in general leads to a mixed integer optimization problem.

1 INTRODUCTION

The class of discrete event systems essentially consists of man-made systems that contain a finite number of resources (such as machines, communications channels, or processors) that are shared by several users (such as product types, information packets, or jobs) all of which contribute to the achievement of some common goal (the assembly of products, the end-to-end transmission of a set of information packets, or a parallel computation) [1]. In general, models that describe the behavior of a discrete event system are nonlinear in conventional algebra. However, there is a class of discrete event systems — the max-plus-linear discrete event systems — that can be described by a model that is “linear” in the max-plus algebra [1, 4], which has maximization and addition as its basic operations. The max-plus-linear discrete event systems can be characterized as the class of discrete event systems in which only synchronization and no concurrency or choice occurs.

In this paper we will consider discrete event systems that can switch between different modes of operation. In each mode the system is described by a max-plus-linear state equation with different system matrices for each mode. The switching changes the structure of the system, and so allows us to break synchronization and to change the order of events. We define a control design problem for such a system to optimize the system’s behavior. In general this will lead to a mixed integer optimization problem.

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

2.1 Max-plus algebra

In this section we give the basic definition of the max-plus algebra and we present some results on a class of max-plus functions.

Define \( \varepsilon = -\infty \) and \( \mathbb{R}_\varepsilon = \mathbb{R} \cup \{ \varepsilon \} \). The max-plus-algebraic addition \((\oplus)\) and multiplication \((\otimes)\) are defined as follows [1, 4]:

\[
x \oplus y = \max(x, y) \quad x \otimes y = x + y
\]

for numbers \( x, y \in \mathbb{R}_\varepsilon \), and

\[
[A \oplus B]_{ij} = a_{ij} \oplus b_{ij} = \max(a_{ij}, b_{ij})
\]

\[
[A \otimes C]_{ij} = \bigoplus_{k=1}^{n} a_{ik} \otimes c_{kj} = \max_{k=1, \ldots, n} (a_{ik} + c_{kj})
\]

for matrices \( A, B \in \mathbb{R}_{\varepsilon}^{m \times n} \) and \( C \in \mathbb{R}_{\varepsilon}^{n \times p} \).

2.2 Max-plus-linear systems

In [1, 4] it has been shown that discrete event systems in which there is synchronization but no concurrency can be described by a model of the form

\[
x(k) = A(k) \otimes x(k-1) \oplus B(k) \otimes u(k).
\]

Systems that can be described by this model will be called max-plus-linear systems. The index \( k \) is called the event counter. For discrete event systems the state \( x(k) \) typically contains the time instants at which the internal events occur for the \( k \)th time, the input \( u(k) \) contains the time instants at which the input events occur for the \( k \)th time, and the output \( y(k) \) contains the time instants at which the output events occur for the \( k \)th time.
2.3 Switching max-plus-linear systems

In this paper we will consider discrete event systems that can switch between different modes of operation. In each different mode $\ell = 1, \ldots, n_m$, the system is described by a max-plus-linear state equation

$$x(k) = A^{(\ell)}(k) \otimes x(k-1) \otimes B^{(\ell)}(k) \otimes u(k) \quad (2)$$

in which the matrices $A^{(\ell)}$, $B^{(\ell)}$ are the system matrices for the $\ell$-th mode. The switching allows us to change the structure of the system, to break synchronization and to change the order of events. Note that each mode $\ell$ corresponds to a set of required synchronizations and an event order schedule, which leads to a model (2) with system matrices $(A^{(\ell)}, B^{(\ell)})$ for the $\ell$-th model.

The moments of switching are determined by a switching mechanism. We define the switching variable $z(k)$, which may depend on the previous state $x(k-1)$, previous mode $\ell(k-1)$, the input variable $u(k)$ and an (additional) control variable $v(k)$:

$$z(k) = \Phi(x(k-1), \ell(k-1), u(k), v(k)) \in \mathbb{R}^{n_z} . \quad (3)$$

We partition $\mathbb{R}^{n_z}$ in $n_m$ subsets $Z^{(i)}$, $i = 1, \ldots, n_m$. The mode $\ell(k)$ is now obtained by determining in which set $z(k)$ is for event $k$. So if $z(k) \in Z^{(i)}$, then $\ell(k) = i$. In some systems the switching mechanism will completely depend on the state $x(k-1)$ and input $u(k)$, in other examples $z(k)$ will be equal to $v(k)$ and so we can control the switching by choosing an appropriate $v(k)$.

3 EXAMPLES OF SWITCHING MAX-PLUS-LINEAR SYSTEMS

3.1 Production system

Consider the production system of Figure 1. This system consists of five machines $M_1$, $M_2$, $M_3$, $M_4$ and $M_5$. The raw material is fed to machine $M_1$ and $M_2$, where preprocessing is done. Both products now have to be finished in either unit $M_3$ and $M_4$, which basically perform the same task, but the processing time of $M_3$ is longer than $M_4$. Therefore, the products coming from machine $M_1$ and $M_2$ are directed to a switching device $S_w$, that feeds the first product in the $k$th cycle to the slower machine $M_3$ and the second product to the faster machine $M_4$. Finally, the products are assembled (instantaneously) in machine $M_5$ and become available. We assume that each machine starts working as soon as possible on each batch, i.e., as soon as the raw material or the required intermediate products are available, and as soon as the machine is idle (i.e., the previous batch has been finished and has left the machine). We define $u(k)$ as the time instant at which the system is fed for the $k$th time, $y(k) = x_5(k)$ as time instant at which the $k$th product leaves the system and $x_7(k)$ as the time instant at which machine $i$ starts for the $k$th time. The variable $t_j$ for $j = 1, \ldots, 4$ is the transportation time, and $d_i$ for $i = 1, \ldots, 5$ is the processing time on machine $i$. The system equations for $x_1$ and $x_2$ are given by

$$x_1(k) = \max(x_1(k-1) + d_1, u(k) + t_1)$$
$$x_2(k) = \max(x_2(k-1) + d_2, u(k) + t_2)$$

If $x_1(k) + d_1 \leq x_2(k) + d_2$, the product of $M_1$ will be directed to $M_3$ and the product of $M_2$ will be directed to $M_4$, and we obtain:

$$x_3(k) = \max(x_3(k) + u(k) + d_1, x_3(k-1) + d_3)$$
$$= \max(x_3(k) + u(k) + d_1 + t_1, x_3(k-1) + d_3, u(k) + d_1 + t_1)$$
$$x_4(k) = \max(x_4(k) + u(k) + d_2, x_4(k-1) + d_4)$$
$$= \max(x_4(k) + u(k) + d_2 + t_2, x_4(k-1) + d_4, u(k) + d_2 + t_2)$$

For this first mode $(x_1(k) + d_1 \leq x_2(k) + d_2) \quad \text{we obtain the system matrices:}$

$$A^{(1)} = \begin{bmatrix}
1 & \epsilon & \epsilon & \epsilon & \epsilon \\
\epsilon & 3 & \epsilon & \epsilon & \epsilon \\
\epsilon & \epsilon & 4 & \epsilon & \epsilon \\
8 & 11 & 12 & 9 & \epsilon \\
\end{bmatrix}, \quad B^{(1)} = \begin{bmatrix}
4 \\
1 \\
4 \\
11 \\
\end{bmatrix}$$

Similarly, for this second mode $(x_1(k) + d_1 > x_2(k) + d_2) \quad \text{we obtain the system matrices:}$

$$A^{(2)} = \begin{bmatrix}
1 & \epsilon & \epsilon & \epsilon & \epsilon \\
\epsilon & 3 & \epsilon & \epsilon & \epsilon \\
\epsilon & \epsilon & 6 & \epsilon & \epsilon \\
7 & 12 & 12 & 9 & \epsilon \\
\end{bmatrix}, \quad B^{(2)} = \begin{bmatrix}
4 \\
1 \\
4 \\
10 \\
\end{bmatrix}$$

\footnote{I.e. with a negligible processing time.}
To decide the switching mechanism, we define the switching variable
\[
\begin{bmatrix}
z_1(k) \\
z_2(k)
\end{bmatrix} = \begin{bmatrix}
x_1(k) + d_1 \\
x_2(k) + d_2
\end{bmatrix}
= \begin{bmatrix}
\max(x_1(k-1) + 2d_1, u(k) + d_1 + t_1) \\
\max(x_2(k-1) + 2d_2, u(k) + d_2 + t_2)
\end{bmatrix}
= \begin{bmatrix}
\max(x_1(k-1) + 2, u(k) + 5) \\
\max(x_2(k-1) + 6, u(k) + 4)
\end{bmatrix}
\]
and the sets
\[
\mathcal{Z}^{(1)} = \{ z \in \mathbb{R}^2 | z_1 \leq z_2 \},
\mathcal{Z}^{(2)} = \{ z \in \mathbb{R}^2 | z_1 > z_2 \}.
\]

Now the state space equation of our system is given by (2).

3.2 Railway network
Consider the railroad network of Figure 2 (see also [11]). There are 4 stations in this railroad network (A, B, C and D) that are connected by 5 single tracks (1/7, 2/4, 3, 5, 9) and one double track (tracks 6 & 8). There are three trains available. The first train follows the route D → A → B → D, the second train follows the route A → B → C → A, and the third train follows the route D → A → C → D. We assume that there exists a periodic timetable that schedules the earliest departure times of the trains. The period of the timetable is \( T = 60 \) minutes. So if a departure of a train from station B is scheduled at 5:30 a.m., then there is also scheduled a departure of a train from station B at 6:30 a.m., 7:30 a.m., and so on.

Each track of the railway network has a number and a train allocated to it. For the sake of simplicity we will say "(virtual) train \( j \)" to denote the (physical) train on a specific track. The number of tracks in the network is equal to 7, the number of physical trains in the network is equal to 3, and the number of virtual trains in the network is equal to 9. Let \( x_j(k) \), \( j = 1, \ldots, 9 \) be the time instant at which train \( j \) departs from its station for the \( k \)th time. Let \( d_j(k) \) be the departure time for this train according to the time schedule, and let \( u_j(k) \) be the transportation time for this train \( j \).

Table 1 summarizes the information in connection with the nominal travelling times and the departure times. All the times are measured in minutes. The indicated departure times are the earliest departure times in the initial station of the track expressed in minutes after the hour. The first period starts at time \( t = 0 \). At the beginning of the first period the first train is in station A and the second train is in station B.

The nominal travelling times and the departure times for the railroad network.

The continuity constraints are that the trains on tracks 1, 2 and 3 are physically the same train, and the same holds for the trains on tracks 4, 5 and 6 and for the trains on tracks 7, 8 and 9.

Connection constraints are introduced to allow the passengers to change trains. In this network, train 1 has to wait for train 9 in the previous cycle with minimum connection time \( c_{\text{min}} = 3 \). In the same way, train 2 waits for train 6 in the previous cycle, train 4 wait for train 7, and train 9 waits for train 5. The minimum stopping time of train \( j \) at station \( j \) to allow passenger to get off or on the train is fixed at \( s_{\text{min}} = 1 \).

Follow constraints are introduced to guarantee sufficient separation time between two trains on the same track (moving in the same direction). In this network, train 4 is scheduled behind train 2 (train 4 follows train 2) with a minimum separation time.

\[
\begin{array}{|c|c|c|c|c|}
\hline
\text{train} & \text{from/to} & \text{travel time} & \text{dep/arr} & \text{constraints} \\
\hline
1 & D-A & 12 & 00-12 & \text{same train as } 3 \text{ connects to } 9 \text{ follows } 7 \text{ in the previous cycle} \\
2 & A-B & 12 & 15-27 & \text{same train as } 1 \text{ connects to } 6 \text{ follows } 4 \text{ in the previous cycle} \\
3 & B-D & 20 & 30-50 & \text{same train as } 2 \text{ follows } 2 \text{ connects to } 7 \\
4 & A-B & 12 & 19-31 & \text{same train as } 6 \text{ follows } 2 \text{ connects to } 7 \\
5 & B-C & 10 & 34-44 & \text{same train as } 4 \text{ connects to } 5 \\
6 & C-A & 25 & 47-12 & \text{same train as } 9 \text{ follows } 1 \text{ connects to } 5 \\
7 & D-A & 12 & 04-16 & \text{same train as } 7 \text{ connects to } 5 \\
8 & A-C & 25 & 19-44 & \text{same train as } 8 \text{ connects to } 5 \\
9 & C-D & 10 & 47-57 & \text{same train as } 3 \text{ connects to } 5 \\
\hline
\end{array}
\]

Note: 3⁻ denotes train 3 in the previous cycle.
$f_{\text{min}} = 4$. In the same way, train 2 follows train 4 in the previous cycle, train 7 follows train 1, and train 1 follows train 7 in the previous cycle.

Each train departs as soon as all the connections are guaranteed (except for a connection when it is broken), the passengers have gotten the opportunity to change over and the earliest departure time indicated in the timetable has passed. We assume that in the first period all the trains depart according to schedule. The $j$-th state $x_j(k)$ is the time instant at which the train on track $j$ departs from the initial station of the track for the $k$th time.

Now we write down the equations that describe the evolution of the $x_j(k)$’s. First we consider the train on track 1 and we determine $x_1(k)$, the time instant at which this train departs from station A for the $k$th time. The train has to wait at least until the train has arrived in station A for the $(k-1)$th time and the passengers have got the time to get out of the train so we have $x_1(k) \geq x_3(k-1) + a_3(k-1) + 1$. Furthermore, the train on track 1 has to wait for the passengers of the train on track 9 in the $(k-1)$th cycle, which arrives in station B at time instant $x_9(k-1)+a_9(k-1)$. The passengers have $c_{\text{min}} = 3$ minutes to change trains. Further the train on track 1 has to follow the train on track 7 in the previous cycle with a minimum separation time $f_{\text{min}} = 4$. According to the timetable the train on track 1 can only depart after time instant $00 + 60$. Hence, we have

$$x_1(k) = \max(x_3(k-1)+a_3(k-1)+s_{\text{min}},$$

$$x_7(k-1)+f_{\text{min}},$$

$$x_9(k-1)+a_9(k-1)+c_{\text{min}}, d_1(k))$$

$$= \max(x_3(k-1)+21, x_7(k-1)+4, x_9(k-1)+13, 60)$$

for $k = 1, 2, \ldots$ with $x_3(0) = x_9(0) = -\infty$.

Using a similar reasoning, we find that the other departure times are given by

$$x_2(k) = \max(x_1(k)+13, x_4(k-1)+4, x_6(k-1)+28, 15, 60),$$

$$x_3(k) = \max(x_2(k)+13, 30, 60),$$

$$x_4(k) = \max(x_2(k)+4, x_6(k-1)+26),$$

$$x_5(k) = \max(x_4(k)+13, 34, 60),$$

$$x_6(k) = \max(x_5(k)+11, 47, 60),$$

$$x_7(k) = \max(x_9(k-1)+11, x_1(k)+4, 4+60),$$

$$x_8(k) = \max(x_7(k)+13, 19, 60),$$

$$x_9(k) = \max(x_8(k)+26, x_5(k)+13, 47, 60)$$

for $k = 1, 2, \ldots$ with $x_j(0) = -\infty$ for $j = 1, 2, \ldots, 9$.

Using successive substitution we can eliminate all right-hand terms with index $k$. By defining the appropriate matrix $A^{(1)}$ and by using the $(\otimes, \oplus)$-notation, we can rewrite the state equations as:

$$x(k) = A^{(1)}(k) \otimes x(k-1) \oplus d(k). \quad (4)$$

In the nominal operation we have assumed that some trains should give pre-defined connections to other trains, and the order of trains on the same track is fixed. However, if one of the preceding trains has a too large delay, then it is sometimes better — from a global performance viewpoint — to let a connecting train depart anyway or to change the departure order on a specific track. This is done in order to prevent an accumulation of delays in the network. In this paper we consider the switching between different operation modes, where each mode corresponds to a different set of pre-defined or broken connections and a specific order of train departures. We allow the system to switch between different modes, allowing us to break train connections and to change the order of trains. Note that any broken connection or change of train order leads to a new model, similar to the nominal equation (4), but now with adapted system matrix $A^{(\ell)}$ for the $\ell$-th model. We have the following system equation for the perturbed operation for $\ell = 2, \ldots, n_m$:

$$x(k) = A^{(\ell)}(k) \otimes x(k-1) \oplus d(k). \quad (5)$$

In this railway network the switching variable $x(k)$ is equal to the control vector $v(k)$, and each entry of $v(k)$ corresponds to a specific control action, so a specific (scheduled) synchronization or specific (scheduled) event order. We assume $v(k)$ to be binary, where $v_1(k) = 0$ corresponds to the nominal case, and $v_1(k) = 1$ to the perturbed case (the synchronization is broken or the order of two events is switched). Each combination $v_1(k), \ldots, v_m(k)$ corresponds to a fixed routing schedule with a specific train order and specific connections.

4 THE MODEL PREDICTIVE CONTROL PROBLEM

Consider switching max-plus-linear model (2)–(3). We have two possible input signals, $v(k)$ and $u(k)$.
Let $\mathcal{V}(k)$ and $\mathcal{U}(k)$ be the sets of possible future control actions $v(k)$ and $u(k)$, respectively. Sometimes values are predefined (e.g. in the railway system $u(k) = d(k)$) or not applicable (e.g. in the production system $v(k)$ is not used). Often $v(k)$ is assumed to be binary, and each entry corresponds to a specific control action (e.g. a specific scheduled synchronization or specific scheduled event order).

Just as in conventional Model Predictive Control (MPC) [10] we define a cost criterion $J$ and we aim at computing the optimal input sequences $u(k), \ldots, u(k + N_p - 1), v(k), \ldots, v(k + N_p - 1)$ that minimize a cost criterion $J(k)$, possibly subject to linear constraints on the inputs and states, where $N_p$ is the prediction horizon. The cost criterion reflects the input and output cost functions (e.g. for the production system the due date error is given by $\Delta u(k + j) = u(k + j) - u(k + j - 1) = 0$, for $j = N_c, \ldots, N_p - 1$. Now the MPC control problem for event step $k$ can be defined as:

$$
\min_{u(k) \in \mathcal{U}, v(k) \in \mathcal{V}(k)} J(k)
$$

subject to

$$
x(k + j) = A((k))x(k + j - 1) + B((k))u(k + j) + \Phi(x(k - 1), \ell(k - 1), u(k), v(k)) \in Z((k))$$

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

$$
\Delta v(k + m) = 0
$$

$$
A_c(k)u(k) + B_c(k)u(k) \leq c_c(k)
$$

for $j = 0, \ldots, N_p - 1$, $m = N_c, \ldots, N_p - 1$

where (13) may represent additional linear constraints on the inputs and the outputs.

MPC uses a receding horizon principle. This means that after computation of the optimal control sequences $u(k), \ldots, u(k + N_c - 1)$ and $v(k), \ldots, v(k + N_c - 1)$, only the first control samples $u(k)$ and $v(k)$ will be implemented, subsequently the horizon is shifted one sample, and the optimization is restarted with new information of the measurements.

In principle we have all elements to solve the receding horizon control problem (7)–(13). In general we will have a mixed integer optimal control problem with both real parameters and binary parameters. Sometimes (e.g. the production system) the problem can be recast as a Extended Linear Complementary Problem (ELCP) and can be solved efficiently [7]. If the optimization is over a binary valued $v(k)$ (e.g. the railway problem) we obtain an integer optimization problem, which can be solved using genetic algorithms [5], tabu search [9], or a branch-and-bound method [3]. In some particular cases the problem can be recast as a Mixed Integer Linear Programming (MILP) or a Mixed Integer Quadratic Programming (MIQP), for which reliable algorithms are available [2, 8].
The application of the derived controller design method to a railway network is given in [11].

5 DISCUSSION

We have presented a new way to model a class of discrete event systems — the max-plus-linear discrete event systems — that can operate in different modes, in which the dynamics can be described by a model that is “linear” in the max-plus algebra. We have discussed two examples, a production system and a railway network. An MPC controller design technique has been derived for this type of systems. In general the resulting optimization problem requires a mixed integer optimization algorithm.

In future research we will study on the characterization of all discrete event systems that can be recast as a switching max-plus-linear system. Furthermore, we will try to find out what conditions are needed on $J, \Phi$ and $Z^{(\ell(k))}$ to obtain particular optimization problems (ELCP, MILP, MIQP).

Acknowledgments

Research partially funded by the Dutch Technology Foundation STW project “Model predictive control for hybrid systems” (DMR.5675), by the European IST project “Modelling, Simulation and Control of Nonsmooth Dynamical Systems (SICONOS)” (IST-2001-37172), and by the NWO/STW VIDI project “Multi-agent control of large-scale hybrid systems” (DWV.6188).

References


