DOCTORAT BRETAGNE LOIRE / MATHSTIC

université angers

THÈSE DE DOCTORAT DE

L'UNIVERSITÉ D'ANGERS

ÉCOLE DOCTORALE N° 601
Mathématiques et Sciences et Technologies de l'Information et de la Communication
Spécialité : *Automatique, Productique et Robotique*

Par

Guilherme ESPINDOLA-WINCK

On the stochastic filtering of max-plus linear systems

Sur le filtrage stochastique de systèmes max-plus linéaires

Thèse présentée et soutenue à Angers, le 9 décembre 2022
Unité de recherche : Laboratoire Angevin de Recherche en Ingénierie des Systèmes (LARIS)
Thèse N° : 226600

Rapporteurs avant soutenance :

Stéphane GAUBERT Directeur de Recherche, INRIA Palaiseau
Bruno GAUJAL Directeur de Recherche, INRIA Grenoble

Composition du Jury :

Président :	Isabel DEMONGODIN	Professeur des Universités, Aix-Marseille Université
Examinateur :	Jean-Jacques LOISEAU,	Directeur de Recherche CNRS, Ecole Centrale de Nantes,
Dir. de thèse :	Laurent HARDOUIN	Professeur des universités, Université d'Angers
Co-dir. de thèse :	Mehdi LHOMMEAU	Maître de Conférences (HDR), Université d'Angers

Acknowledgement

Throughout this thesis, several people, to whom I am deeply grateful, have supported me.

Foremost, I would like to express my gratitude to my supervisors, **Prof. Laurent Hardouin** and **Dr. Mehdi Lhommeau**, for their help when needed. I am sure that the results of this thesis have been accomplished thanks to their actions, especially through scientific writing, advice and revision.

I would also like to acknowledge all the members of the evaluation committee: **Prof. Stéphane Gaubert**, **Prof. Bruno Gaujal**, **Prof. Isabel Demongoding** and **Prof. Jean-Jacques Loiseau**. Their remarks and suggestions were very important for the improvement of this manuscript.

I am grateful to all LARIS members for a pleasant environment. I recognize every precious moment among my colleagues in the doctoral office.

I would like to thank my parents as well as all my family and friends who, from near and far, helped and encouraged me at the right time.

Last but not least, a thousand thanks to Mylena.

TABLE OF CONTENTS

List of Figures

List of Tables

INTRODUCTION

Discrete Event (Dynamic) Systems [1] (DES) are discrete-state systems whose dynamics are entirely driven by the occurrence of asynchronous events over discrete time instants [1, Sec. 1.3.2].

The study of DES is of great interest in analysis, control and estimation of systems that are often of human design. Examples of DES include computer systems [2], telecommunication networks [3, 4], manufacturing lines and transportation systems. Unlike classical continuous systems, they usually cannot be described by differential equations because of the nature of the involved phenomena. These systems are then often represented by state-transition models: finite state automata and languages [5] which are used to represent simple deterministic systems; Markov chains for stochastic systems [6]; and Petri nets for more complex systems that include synchronization, conflict and parallelism [7].

Among the DES, a particular class involving only synchronization and time-delay phenomena has been the subject of a dedicated algebraic development, generally called Max-Plus linear algebra [8]. This class of DES can be represented graphically, depicted by Timed Event Graphs (TEG) and is typical for the modelling of systems arising in scheduling applications involving allocation of resources.

The fascinating mathematical theory of Max-Plus algebra is due to the fact it provides techniques for solving nonlinear problems that can be given in the form of linear problems, when arithmetical addition is replaced by the operation of *maximization* (denoted $\oplus$) and arithmetical multiplication is replaced by *addition* (denoted $\otimes$). The Max-Plus algebra setting, which is an idempotent semiring that is properly presented later in this thesis, is suitable to describe the behavior of TEG thanks to linear state equations which are very analogous to those found in classical linear system theory, *i.e.*, the behavior of a Max-Plus Linear system (MPL) can be depicted thanks to matrices defined in this algebra. The states of MPL systems represent the *dates* of occurrence of the events in a given event-horizon and can be simulated using the Max-Plus Toolbox [9] for `Scilab` and `ScicosLab` with built-in property analysis tools that mainly concern periodic behavior regime, *i.e.*,

1. The term Discrete Event Systems can be also used if we keep in mind that they are dynamic by definition.

no verification tool is available.

These linear state equations are useful to deal with performance analysis [10] and control problems addressed similarly to the classical control theory. Among the problems solved we can cite the optimal control [11, 12], the model-predictive control [13, 14], the robust controller design [15, 16] and the control strategies allowing the state to stay in a specific state subspace or semi-module [17, 18, 19, 16, 20].

These models are also useful to deal with state estimation [21, 22, 23, 24] which is a fundamental problem to address applications such as fault detection and diagnosis [25] and state feedback control [26]. However, if the system is with uncertain parameters [2], some alternative methods can be considered in order to take advantage of the knowledge about the characteristics of this uncertainty [23, 24, 23, 27, 28, 29]. It is worth mentioning that if we do not take these uncertainties into account, we may observe tracking error or instability in closed loop control strategies [30]. The nondeterministic MPL systems (or Stochastic Max-Plus Linear (SMPL) systems if probabilistic information is available) are defined as MPL systems whose matrix entries are characterized by random variables [31, 32, 30, 33, 34].

Objective of the thesis

In this work, we are interested in stochastic filtering of Uncertain Max-Plus Linear (uMPL) systems, a subclass of nondeterministic systems, with entries that are mutually independent bounded random variables with support in a real interval (bounded domain). The objective of the stochastic filtering is to be able to estimate the state of the system given a sequence of observations (measurements). Relying on the theory of estimation, it is well known that all the information for estimating the state of the system is contained in the *posterior* probability density function (PDF) of this state. However, as presented later, the solution of this problem is for linear (or nonlinear with good linear approximations) systems with Gaussian noise only (this solution is known as Kalman filter), which is not the case of nondeterministic MPL systems.

2. The noise in MPL systems is Max-Plus multiplicative and appears as uncertainties in the Max-Plus model parameters. As a result, the system matrices are uncertain.

Preliminary works

To meet probabilistic objectives, we cite some preliminary works which are the basis of this thesis:

— the *conditional* reachability analysis (RA) corresponds to the support calculation of the posterior PDF of the uMPL system states [35, 27, 29]. However, it should be noted that the *conditional* (RA) techniques are not stochastic since it does not rely on any probability measurement. Nevertheless, this approach is useful to improve the efficiency of Particle filtering algorithms. Particle filters, or Sequential Monte Carlo methods, are suboptimal Bayesian algorithms based on weighted-particle approximation of probability densities [36, 37], and in general these approaches are limited by the numerical difficulties due to the generation of the particles and by the fact that the lower dimensionality of the measurements with respect to the state, introduces an imprecise generation of particles in the state-space. Particle filters applied to MPL systems have been studied in [23, 29];

— in [28], an alternative stochastic filtering algorithm for uMPL systems has been studied. The algorithm is depicted in a two-fold scheme, as it is the case of the Bayes filter (basis of Kalman filter): a prediction phase based on the calculation of the conditional mathematical expectation $\mathbb{E}[\text{observation}|\text{state}]$; and a correction phase using available measurements and the prediction into account in order to compute a state estimate, based on a constraint satisfaction problem [38].

Contributions

In this thesis, the following problems have been analyzed:

— we present a reinterpretation of *conditional* RA for uMPL systems using Max-Plus polyhedra [39]. We show that the forward reach set, *i.e.*, the set of all states that can be reached from a previous set via dynamics, is computed in strongly polynomial-time and are characterized by a single Max-Plus polyhedron whereas in [35, 29] the complexity is exponential, and also being characterized as a collection of Difference-Bound Matrices (DBM) [40, Sec 4.1] to represent the same set (explosion in the number of DBM as the dimension evolves). Conversely, we show that the backward reach set, *i.e.*, the set of all states that lead to a measurement point via observation, is computed with exponential complexity in worst-case scenarios using either the polyhedral or the DBM approaches. Nevertheless, since the intersection

of the forward and backward reach sets is the objective of stochastic filtering, then we claim that we have obtained a significant improvement. Moreover, polyhedral-sets allow us to promptly verify inclusion by using a procedure with polynomial complexity in the dimension;

— we present conditions to compute (exactly or approximately) the forward and backward reach set using algebraic operations with quadratic complexity;

— we introduce a criterion to the alternative stochastic filtering proposed in [28] that makes it possible to define an improved filtering algorithm. The new proposed algorithm is based on the minimization of this criterion which allows evaluating the estimation error of the prediction phase and to adjust the importance to be given between the prediction estimation and the corrected estimation, as the quadratic criterion does in Kalman filter for linear continuous time-driven systems.

— we extend the stochastic filtering algorithms, that are essentially devoted to TEG whose places have at least a single token, to systems with token-free places. Briefly, filtering algorithms assume that the stochastic entries of the system matrices are independent. However, with token-free places, the entries of system matrices are sums and/or max of processing times, in which various processing times appear in multiple entries. Therefore, the entries of the matrices are not independent, and these procedures are not directly adapted.

List of Publications (International comunications)

1. G. Espindola-Winck, L. Hardouin, M. Lhommeau, and R. Santos-Mendes, *"Stochastic filtering scheme of implicit forms of uncertain max-plus linear systems,"* IEEE Transactions on Automatic Control, vol. 67, no. 8, pp. 4370–4376, 2022.

2. G. Espindola-Winck, L. Hardouin, and M. Lhommeau, *"Max-plus polyhedra-based state characterization for uMPL systems,"* in 2022 European Control Conference ECC 22, pp. 1037–1042, 2022.

3. G. Espindola-Winck, R. Ferreira Candido, L. Hardouin, and M. Lhommeau, *"Efficient state-estimation of uncertain max-plus linear systems with high observation noise,"* 16th IFAC Workshop on Discrete Event Systems WODES 22, Prague, September 2022.

4. G. Espindola-Winck, L. Hardouin, M. Lhommeau, and R. Santos Mendes, *"Criteria stochastic filtering of max-plus discrete event systems with bounded random*

variables," 1st IFAC Workshop on Control of Complex Systems Cosy 22, Bologna, November 2022.

Ongoing writing: "Forward Reachability Computation of Uncertain Max-Plus Linear Systems: a Polyhedral Approach" to be summited to *Automatica*.

List of Publications (National comunications in ***french***)

1. G. Espindola-Winck, L. Hardouin and M. Lhommeau. *"Analyse d'atteignabilité des systèmes (max,+)-linéaires à l'aide des polyèdres tropicaux,"* MSR 19 Modélisation des Systèmes Réactifs, Angers, 2019.
2. G. Espindola-Winck, L. Hardouin and M. Lhommeau. *"Sur l'estimation d'état des systèmes max-plus,"* MSR 21 Modélisation des Systèmes Réactifs, Paris, 2021.

Organization of the document

This work is organized as follows: Chapter 1 recalls the preliminaries on DES, Max-Plus algebra, Residuation theory over complete dioids and MPL systems. Chapter 2 gives an overview of the methods for RA analysis of uMPL systems. The main contribution in this Chapter is related to the use of polyhedral-sets to perform RA instead of using DBM. Chapter 3 recalls the stochastic filtering of uMPL systems and presents some contributions: the definition of criteria "à la Kalman"; an extension to TEG with token-free places; consistency with the support of the posterior PDF.

Chapter 1

Prerequisites

This chapter covers a brief introduction to Discrete Event Systems (DES), Max-Plus algebra, Residuation theory over complete dioids. At the end of this Chapter, we introduce Max-Plus Linear (MPL) systems.

1.1 Discrete Event Systems

In systems theory, systems that describe physical phenomena are controlled by considering model-based strategies using, *e.g.*, ordinary differential equations. Some properties must be satisfied in order to properly use these models:

1. the state-space must be continuous;
2. the presence of time-driven state transition mechanism.

The state-space being continuous means that it can be defined through continuous variables that can take any arbitrary value within a real (or complex) interval. Some common physical variables such as position and temperature are examples of time-driven systems. Hence, any continuous variable, as long as one can naturally define its derivatives, is suitable to be handled by this class of systems.

The second property is due to the fact that the states generally evolve as a function of time. Figure 1.1 shows a typical trajectory of such systems.

Furthermore, control engineers are also interested in the control of complex systems, *e.g.*,

— manufacturing systems;
— transportation systems;
— telecommunication networks;

To describe the behavior of these systems, the mathematical tools of time-driven systems are not suitable, hence more relevant theoretical setting are considered, among them

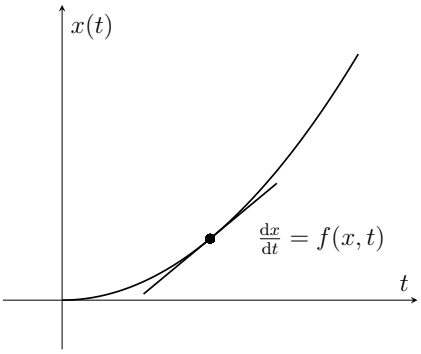


Figure 1.1 – Typical trajectory of a time-driven system

the following can be cited: languages and automata, Markov chains and Petri nets. The reader is invited to consult [1] for an overview.

As it was done for time-driven systems, we state the properties of these systems:

1. the state-space is discrete;
2. the presence of event-driven state transition mechanism.

Definition 1.1 (Discrete Event System (DES) [1])**.** *Discrete Event Systems (DES) are systems whose dynamics are event-driven, i.e., the state evolution depends entirely on the occurrence of asynchronous discrete events over time.*

Discrete state-spaces are present in several systems, the majority of them having a technological application. The event-driven property of DES is due to the fact that the states can only change over time at discrete instants, which physically coincides to the definition of asynchronous occurrences of discrete events. The following implication naturally arises:

> If it is possible to identify any sort of "event" that can trigger a state transition $\implies$ time no longer serves to drive such a system.

Example 1.1. *A machine has the following set of states that describe its behavior:*

$$x \in \{\texttt{ON}, \texttt{OFF}, \texttt{Processing}, \texttt{Idle}\}$$

Figure 1.2 depicts an arbitrary trajectory of the states of this machine. In this trajectory, events are represented by the Greek letters α, β, γ. It can be seen that a same event has different effects in the state evolution, depending on the state in which it occurs. For

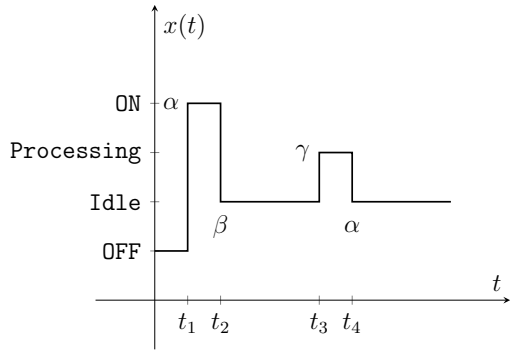


Figure 1.2 – Trajectory of the states of the machine of Example 1.1

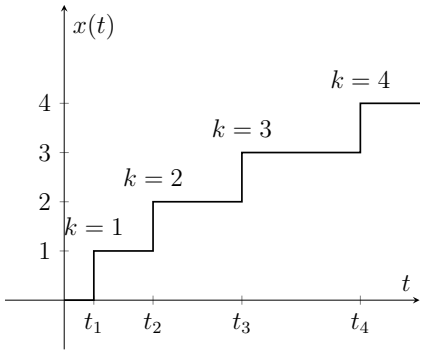


Figure 1.3 – Typical trajectory of a counter of Example 1.2

instance, in the figure: if the system is in state `OFF` *and an event* α *occurs then the next state will be* `ON`*; if* α *occurs at* `Processing` *then the next state will be* `Idle`*. The trajectory can continue indefinitely. Hence, the trajectory is given by the (untimed or logical) finite list of events* $\{\alpha, \beta, \gamma, \alpha, \dots\}$*. If the timing information is explicitly included, then we consider the timed list* $\{(\alpha, t_1), (\beta, t_2), (\gamma, t_3), (\alpha, t_4), \dots\}$*, otherwise, we consider the previous logical list of events.*

Example 1.2. *Counters are naturally discrete, i.e.,* $x = 1, 2, 3, \dots$ *counts, for instance, the number of products that have been processed by a machine. Figure 1.3 depicts this behavior. The term* $k \in \{1, 2, 3, 4\}$ *denotes the numbering of the events that increase the value of* x *(monotonic increasing function) while the set* $\{t_1, t_2, t_3, t_4\}$ *denotes the date of occurrences of such events.*

As already mentioned, it is possible to use different modeling approaches for DES, without any of them being considered as "universal". Indeed, depending on the system to be modeled, a specific approach may simplify the modeling and ease, *e.g.*, analysis and control design. For instance, state machines, a subclass of Petri nets, cannot model synchronization phenomena but are perfectly suitable to handle problems involving conflicts and/or decisions [41]. Another subclass of Petri nets, called marked graphs or event graphs, cannot model conflicts but are perfectly adapted to model synchronization phenomena [42], for instance in a manufacturing system subject to synchronization: the assembly of a product waits for n resources to become available. In this subclass, timing information is generally considered in the modelling, and we consider *Timed* Event Graphs (TEG), which explicitly contain timing information for the different events. TEG can be described in terms of state-space equations using *maximization* and *addition* operations to describe

delay and *synchronization* phenomena.

In the sequel, we present the basic concepts of Petri nets and Timed Event Graphs (TEG).

1.1.1 Petri Nets (PN)

Definition 1.2 ([1])**.** *A Petri net (PN) is a bipartite-directed graph given by the following* 5*-tuple*

$$N = (P, T, A, w, x),$$

composed of:

— *Two types of nodes:*

1. *places $p \in P$, where P is the set of places, are passive nodes depicted by circles, and refer to conditions or resources;*
2. *transitions $t \in T$, where T is the set of transitions, are active nodes depicted by bars and refer to activities, i.e., events that can change the state of the resources.*

— *Tokens, indicated as dots or numbers within places $p \in P$, are discrete variable elements, denoted $x(p) \in \mathbb{N}$, which define the state (or marking) of the PN;*

— *Directed arcs interconnecting places and transitions*[1]*:*

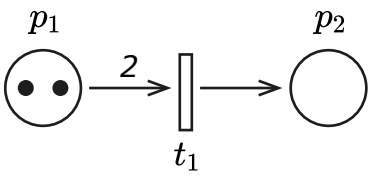
1. *places $p \in P$ exclusively connected to transitions $t \in T$;*
2. *transitions $t \in T$ exclusively connected to places $p \in P$.*

The set of arcs is denoted $A \subseteq (P \times T) \cup (T \times P)$. Furthermore, arcs are inactive elements and are depicted by arrows. They specify the causal relationships between transitions and places and indicate how the marking is changed by the firing of a transition;

— *Weight function: each arc is related to an arc weight $w : (P \times T) \cup (T \times P) \to \mathbb{N}$. The arc weight defines the number of tokens that are consumed or produced by a transition.*

Example 1.3. *Consider the following PN, denoted $N = (P, T, A, w, x)$,*

1. Directed arcs connect only nodes of different types (definition of bipartite graph).



Then the following elements of N are promptly obtained

$$P = \{p_1, p_2\}, \quad T = \{t_1\}, \quad A = \{(p_1, t_1), (t_1, p_2)\}$$
$$w(p_1, t_1) = 2, \quad w(t_1, p_2) = 1, \quad (x(p_1), x(p_2))^{\mathrm{t}} = (2, 0)^{\mathrm{t}}$$

Elementary structures

PN are very powerful, because they are capable of representing a large amount of features, such as:

— *Sequence:* Figure 1.4a depicts an example of this structure, in which a token in place p_0 enables the transition t_1. The firing of transition t_1 enables the transition t_2, since p_1 is now marked;

— *Fork:* Figure 1.4b shows an example of a fork structure. This structure is the basic feature of the creation of parallel processes;

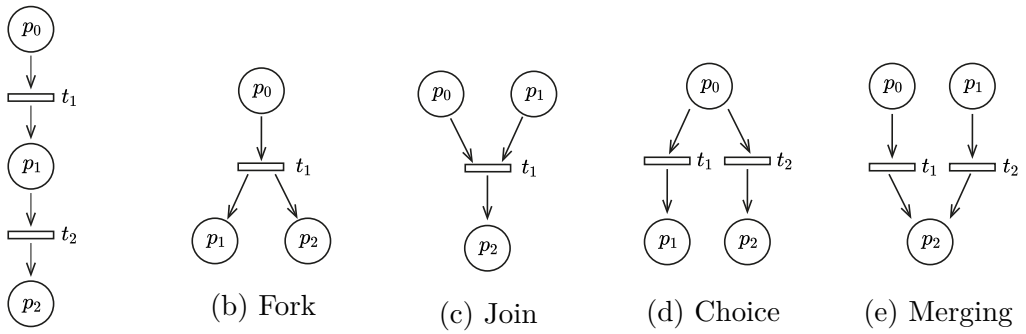
— *Join (synchronization):* Figure 1.4c combines two concurrent processes, allowing that another process continues its execution only after the end of predecessor processes (represented by the firing of transition t_1), *i.e.*, these processes need to synchronize with each other;

— *Choice:* Figure 1.4d depicts the process of making choices, *i.e.*, the firing of transition t_1 disables the transition t_2 and the firing of t_2 disables t_1;

— *Merging:* Figure 1.4e shows a net with two independent transitions t_1 and t_2 that have an output place p_2 in common. Therefore, the firing of one of these two transitions, enables the firing of subsequent transitions since p_2 becomes marked.

Special modelling features

— *Parallel processes:* A model for parallel processes may be obtained by composing the model for each individual process with fork and synchronization structures. Figure 1.5a depicts an example of a parallel process, where I and II represent parallel activities. Since p_0 is marked, then transition t_1 is enabled. When transition t_1 fires, p_1^I and p_1^{II} become marked, representing a concurrency, *i.e.*, internal



(a) Sequence (b) Fork (c) Join (d) Choice (e) Merging

Figure 1.4 – Elementary PN structures

transitions in I are independent of II. The firing of t_2 depends on p_n^I and p_n^{II}, implying that both process I and II are finished and p_0 is once again marked, ready to restart the process by enabling the firing of t_1;

— *Mutual exclusion:* The modelling of resource sharing is addressed in a mutual exclusive way. Figure 1.5b depicts a single resource (token) in place R. This place is seen as a pre-condition for sub-systems S_1 and S_2 that need such resource. After the use of this resource by S_1 or S_2, then it must be released and ready to be used again.

Dynamic evolution in PN with initial marking

Definition 1.3 ([1])**.** *A transition $t_j \in T$ in a PN is said to be* ***enabled*** *if $x(p_i) \geq w(p_i, t_j)$ for all $p_i \in I(t_j) = \{p_i \in P \mid (p_i, t_j) \in A\}$ (the set of all input places for transition t_j).*

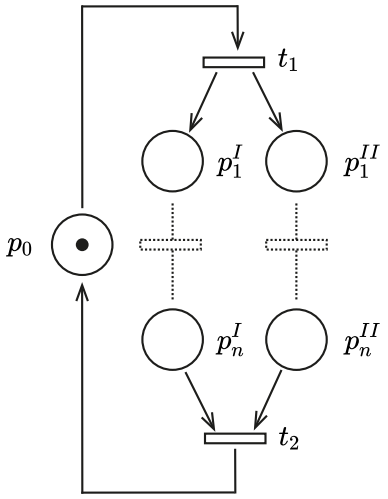
Remark 1.1. *If there is no arc from place $p_i \in P$ to transition $t_j \in T$, then $w(p_i, t_j) = 0$. Similarly, if there is no arc from t_j to p_i, then $w(t_j, p_i) = 0$.*

Definition 1.4 (Transition equation)**.** *The map $f : \mathbb{N}^{|P|} \times T \to \mathbb{N}^{|P|}$ of the PN given by $N = (P, T, A, w, x)$ is defined by the transition $t_j \in T$ if and only if*

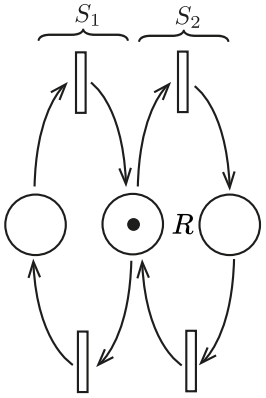
$$p_i \text{ } \textbf{\textit{is enabled}} \; \forall p_i \in I(t_j).$$

Then the following equation defines $f((x(p_1), \ldots, x(p_{|P|}))^{\mathrm{t}}, t_j)$:

$$x'(p_i) = x(p_i) - w(p_i, t_j) + w(t_j, p_i), \;\; i = 1, 2, \ldots, n$$



(a) Parallel processes



(b) Mutual exclusion

Figure 1.5 – Special PN modelling features

where $(x(p_1), \ldots, x(p_{|P|}))^{\mathrm{t}}$ *and* $(x'(p_1), \ldots, x'(p_{|P|}))^{\mathrm{t}}$ *represent the numbers of tokens in all* $p_i \in P$ *before and after the firing of* t_j*, respectively.*

In this way, the next marking (state), explicitly depends on the input and output places of a transition and the weights of the arcs that connect these places to the transition. It should be noted that the number of tokens does not necessarily need to be conserved after the firing of a transition in a PN. In general, it is possible that after several transition firings, the resulting state is $(x(p_1), \ldots, x(p_{|P|}))^{\mathrm{t}} = (0, \ldots, 0)^{\mathrm{t}}$, or that the number of tokens in one or more places is arbitrarily large.

Remark 1.2. *If a PN has transitions with choice structure, then we say that they are in conflict. In this thesis, however, the focus is on non-conflicting PN.*

Definition 1.5 (Event graph)**.** *A non-conflicting PN is called an event graph, if each place has exactly one upstream and one downstream transition and if all arcs have weight* 1*. Event graphs are perfectly suitable to model synchronization phenomenon.*

Example 1.4. *Let* $N = (P, T, A, w, x)$ *be an event graph, with initial marking depicted*

by Figure 1.6a. Then the following elements of N are promptly obtained,

$$P = \{p_1, p_2, p_3\}, \quad T = \{t_1, t_2, t_3\}, \quad A = \{(t_1, p_1), (p_1, t_2), (p_3, t_2), (t_2, p_2), (p_2, t_3), (t_3, p_3)\}$$

$$w(t_1, p_1) = w(p_1, t_2) = w(p_3, t_2) = w(t_2, p_2) = w(p_2, t_3) = w(t_3, p_3) = 1, \quad \begin{pmatrix} x(p_1) \\ x(p_2) \\ x(p_3) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

Below we list the cyclic behavior of N:

1. *Figure 1.6a: t_1 is the only transition that is enabled since there is no precedence condition. The firing of t_1 inserts a token in place p_1, since*

$$x'(p_1) = 0 - w(p_1, t_1) + w(t_1, p_1) = 0 - 0 + 1 = 1$$
$$x'(p_2) = 0 - w(p_2, t_1) + w(t_1, p_2) = 0 - 0 + 0 = 0$$
$$x'(p_3) = 1 - w(p_3, t_1) + w(t_1, p_3) = 1 - 0 + 0 = 0$$

2. *Figure 1.6b: t_1 and t_2 are enabled because:*
   — *t_1: there is no precedence condition;*
   — *t_2: for the set $I(t_2) = \{p_1, p_3\}$ we have $x(p_1) = 1 \geq w(p_1, t_2) = 1$ and $x(p_3) = 1 \geq w(p_3, t_2) = 1$.*

   Two options are possible:

   (a) The firing of t_1 inserts again a token in place p_1;

   (b) The firing of t_2 inserts a token in p_2 and eliminates a token in p_1 and another in p_3, since

$$x'(p_1) = 1 - w(p_1, t_2) + w(t_2, p_1) = 1 - 1 + 0 = 0$$
$$x'(p_2) = 0 - w(p_2, t_2) + w(t_2, p_2) = 0 - 0 + 1 = 1$$
$$x'(p_3) = 1 - w(p_3, t_2) + w(t_2, p_3) = 1 - 1 + 0 = 0$$

   We choose the second option;

3. *Figure 1.6c: t_1 and t_3 are enabled because:*
   — *t_1: there is no precedence condition;*
   — *t_2: for the set $I(t_3) = \{p_2\}$ we have $x(p_2) = 1 \geq w(p_2, t_3) = 1$. Two options are possible:*

   (a) The firing of t_1 inserts again a token in place p_1;

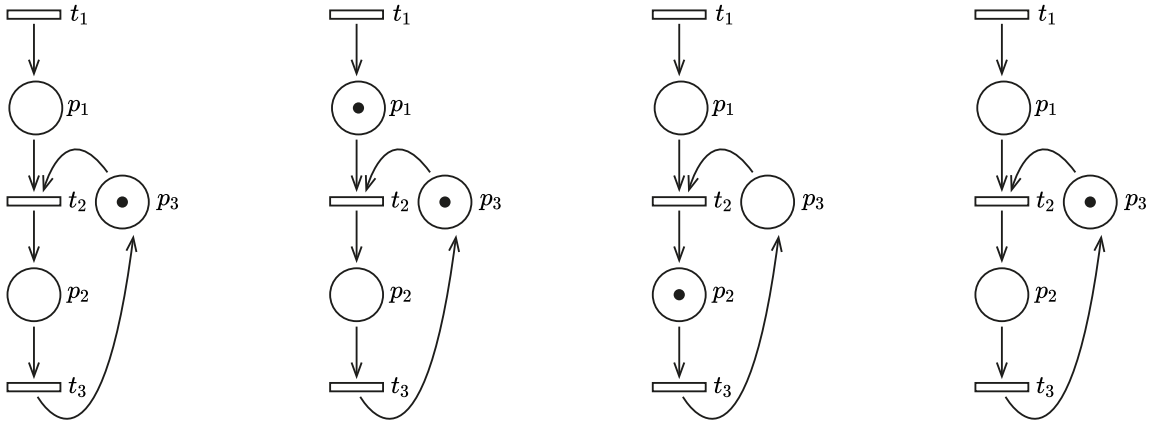


Figure 1.6 – Dynamic evolution of the markings of the PN of Example 1.4

(b) The firing of t_3 inserts a token in p_3 and eliminates a token in p_2, since

$$x'(p_1) = 0 - w(p_1, t_3) + w(t_3, p_1) = 0 - 0 + 0 = 0$$
$$x'(p_2) = 1 - w(p_2, t_3) + w(t_3, p_2) = 1 - 1 + 0 = 0$$
$$x'(p_3) = 0 - w(p_3, t_3) + w(t_3, p_3) = 0 - 0 + 1 = 1$$

We choose the second option;

4. *Note that after the considered firings, the PN is back to its initial condition. Analogously, a new cycle can be started, through the firing of transition t_1.*

1.1.2 Timed PN and Timed Event Graphs (TEG)

In a standard PN, the quantification of time does not appear. Indeed, it only models the possible ordering of firings of transitions, but not the actual firing times, since all actions are assumed to be instantaneous. In order to take into account the timing information, there are two approaches, precisely:

— *T-time PN*: The timing information is associated with the transitions of a PN. When the transition is activated, the triggering is not immediate, being triggered only after its delay (*transition delay*). When the transition delay is equal to zero, the transition will fire as soon as the transition is enabled.

— *P-time PN:* The timing information is associated with the places of a PN. After being inserted in a timed place, a token is not available to activate a next transition

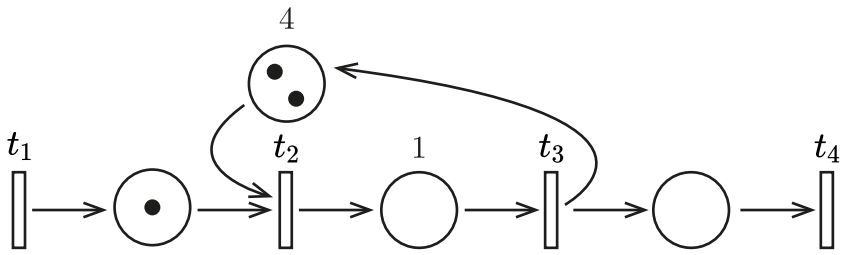


Figure 1.7 – Example of a TEG

until the minimum time (*holding time*) required by the corresponding place has elapsed.

Definition 1.6 (Timed event graph (TEG))**.** *An event graph equipped with either transition delays or holding times as timing information defines Timed Event Graph (TEG).*

Remark 1.3 ([7])**.** *If every transition in a TEG, associated with a transition delay, has at least one input place, then the transition delays can always be converted into holding times (by simply shifting each transition delay to all input places of the corresponding transition). However, in general, it is not possible to convert every TEG with holding times into a TEG with transition delays. Therefore, we will only consider TEG with timing information in the places, i.e., for each place $p \in P$ we associate a holding time $\theta(p) \in \mathbb{R}_+$, where $\mathbb{R}_+$ is the set of the non-negative real numbers.*

Figure 1.7 is an example of a TEG, and we have three distinguished transitions that appear in any TEG:

1. input transitions that are not affected by the firing of other transitions (t_1);
2. output transitions that do not affect the firing of other transitions (t_4);
3. and internal transitions that are neither input nor output transitions (t_2, t_3).

Temporal behavior of TEG

The dynamic behavior of a TEG can algebraically be represented in different ways. On top of this, we know that TEG is a subclass of timed PN, and then it is possible to consider the evolution of the marking of the places as the state recursion (see Definition 1.4). Nevertheless, in order to be able to discuss the performance of the TEG, *i.e.*, to determine its transient regime and/or its steady state, then a suitable modeling is to not consider the state of the marking but the *dates* of occurrence of the transitions (events).

In the following, we will detail how to obtain the mathematical description of the dynamics of a TEG. First, we associate each transition with a *dater* $x_{t_i}(k) : \mathbb{N} \to \mathbb{R}_+$, where $x_{t_i}(k)$ is an increasing map that denotes the date at which the firing of number $k \in \mathbb{N}$ of the transition t_i occurs. By convention, the transition firings will be numbered from 0, *i.e.*, firing number 0 is the first firing of a transition. Moreover, $x_{t_i}(k) = +\infty$ means, by convention, that the events numbered greater than or equal k, never took place. It is also assumed that the tokens have been present in the TEG since *the beginning of time*, *i.e.*, since the date $-\infty$, and that at an initial instant (which can be taken equal to 0), no transition has been fired.

Roughly speaking, the dynamic evolution of a TEG is associated to the following general synchronization structure (join structure of the PN of Figure 1.4c) depicted in Figure 1.8 for two input places, being simple to extend to multiple input places, one input place being trivial. In this example, we have three transitions t_1, t_2 and t_3 and two places p_1 and p_2 with corresponding holding times $\theta(p_1)$ and $\theta(p_2)$. In Figure 1.8a, no token is initially present in places p_1 and p_2, hence, the earliest date of the k-th firing of transition t_3 is conditioned by the dates of the k-th firings of t_1 and t_2, shifted by $\theta(p_1)$ and $\theta(p_2)$, respectively. Mathematically,

$$\begin{aligned} & x_3(k) \geq x_1(k) + \theta(p_1) \quad \text{and} \quad x_3(k) \geq x_2(k) + \theta(p_2) \\ \iff & x_3(k) \geq \max\left(x_1(k) + \theta(p_1), x_2(k) + \theta(p_2)\right) \end{aligned}$$

Hence, the *maximization* operation is responsible by the synchronization, *i.e.*, for instance, if $x_1(k) + \theta(p_1) > x_2(k) + \theta(p_2)$ then t_3 is triggered at earliest at $x_1(k) + \theta(p_1)$.

Figure 1.8b represents a more general situation where the initial marking is with tokens for some places. In this case, the reasoning is as follows: the k-th firing of the transition t_3 is conditioned by

— $(k - M(p_1))$-th firing of the transition t_1 since p_1 has initially $M(p_1)$ tokens;

— $(k - M(p_2))$-th firing of the transition t_2 since p_1 has initially $M(p_2)$ tokens;

Thus

$$\begin{aligned} & x_3(k) \geq x_1(k - M(p_1)) + \theta(p_1) \quad \text{and} \quad x_3(k) \geq x_2(k - M(p_2)) + \theta(p_2) \\ \iff & x_3(k) \geq \max\left(x_1(k - M(p_1)) + \theta(p_1), x_2(k - M(p_2)) + \theta(p_2)\right) \end{aligned}$$

By applying this procedure, the fundamental principle of each dater x_{t_i} of the corre-

(a) Free token case (b)

Figure 1.8 – General synchronization structure for two input places of a TEG

sponding transition t_i of a TEG is given in the sequel.

Fundamental principle of TEG

$$x_{t_i}(k) \geq \max_{t_j \in T \backslash \{t_i\}} \left(x_{t_j}(k - M(p_j)) + \theta(p_j) \right), \; t_i \in T$$

where $k \in \mathbb{N}$ is the event-number, t_i is the output transition, t_j is the input transition,

$$p_j \in P \backslash \{p_i\} \text{ and } p_i \in P$$

are the output places immediately after t_j and the output places immediately after t_i, respectively. Additionally,

$$M(p_j) \in \mathbb{N}$$

is the number of tokens in p_j and $\theta(p_j)$ the holding time [a] of p_j.

a. If there is no arc between t_i and p_j then $\theta(p_j) = -\infty$ and thus $\mathbf{x} = (x_{t_1}, \ldots, x_{t_n})^{\mathrm{t}} \in (\mathbb{R} \cup \{-\infty\})^n$.

Example 1.5. *Consider the TEG of Figure 1.7. By applying the fundamental principle of TEG, we obtain the following system of inequalities:*

$$\begin{cases} x_{t_1}(k) & \textit{is a control input} & \\ x_{t_2}(k) & \geq & \max(x_{t_1}(k-1), x_{t_3}(k-2)+4) \\ x_{t_3}(k) & \geq & x_{t_2}(k)+1 \\ x_{t_4}(k) & \geq & x_{t_3}(k) \end{cases}$$

Another possible way to describe TEG is by considering the so-called *counters* denoted $\mathbf{x}(\tau) \in \mathbb{N}$. Counters determine the number of events that have occurred up to time t, for instance, $x_i(\tau)$ refers to the number of occurrences of transition i up to time τ. It is assumed that time is discrete in this representation, *i.e.*, $\tau \in \mathbb{N} \mapsto \mathbf{x}(\tau) \in \mathbb{N}$.

Example 1.6. *Consider again the TEG of Figure 1.7. The time-domain model of this TEG is given by the following system of inequalities:*

$$\begin{cases} x_{t_1}(\tau) & \textit{is a control input} & \\ x_{t_2}(\tau) & \leq & \min(1 + x_{t_1}(\tau), x_{t_3}(\tau - 4) + 2) \\ x_{t_3}(\tau) & \leq & x_{t_2}(\tau - 1) \\ x_{t_4}(\tau) & \leq & x_{t_3}(\tau) \end{cases}$$

Consequently, it is possible to model the dynamic behavior of a TEG either in the event-domain (daters) or in the time-domain (counters). In this work, we consider the model dynamics in the event-domain only.

Remark 1.4. *Considering the earliest firing rule, i.e., every internal and output transition fires as soon as it is enabled, the system of inequalities $\geq$ of Example 1.5 becomes a system of equations $=$.*

In view of the recursive equations for the transition firing times of Example 1.5, it is easy to recognize that addition and the maximization operations are necessary to determine the desired timetable. Due to the maximization operation, these equations are nonlinear in conventional algebra. Nevertheless, in the sequel we introduce a suitable mathematical framework to properly deal with these systems.

1.2 Max-Plus algebra: an idempotent semiring

The Max-Plus algebra has been an important area of study since the 1970s, with special attention to the modeling and control of DES involving allocation of resources. In the sequel, we shall use [43] and [10].

1.2.1 Basic concepts

Let us begin with some notations. Let $\mathbb{R}_{\max} \stackrel{\text{def}}{=} (\mathbb{R} \cup \{\varepsilon\}, \oplus, \otimes)$ be the union of the set of all real numbers with $\varepsilon = -\infty$. This set is an idempotent semiring or dioid [2], (*i.e.*, $\forall a \in \mathbb{R}_{\max}, a \oplus a = a$), endowed with max as the addition operator $\oplus$ and $+$ as the product operator $\otimes$. These notations are consistent with respect to the usual properties:

Associativity of $\oplus$ and $\otimes$:

$$\begin{aligned} \forall a, b, c \in \mathbb{R}_{\max},\ (a \oplus b) \oplus c &= a \oplus (b \oplus c) \\ \forall a, b, c \in \mathbb{R}_{\max},\ (a \otimes b) \otimes c &= a \otimes (b \otimes c) \end{aligned}$$

Commutativity $\oplus$ and $\otimes$:

$$\begin{aligned} \forall a, b \in \mathbb{R}_{\max},\ a \oplus b &= b \oplus a \\ \forall a, b \in \mathbb{R}_{\max},\ a \otimes b &= b \otimes a \end{aligned}$$

Distributivity of $\otimes$ over $\oplus$:

$$\forall a, b, c \in \mathbb{R}_{\max},\ (a \oplus b) \otimes c = (a \otimes c) \oplus (b \otimes c)$$

Neutral element for $\oplus$ and the absorbing element for $\otimes$:

$$\begin{aligned} \forall a \in \mathbb{R}_{\max},\ a \oplus \varepsilon &= \varepsilon \oplus a = a \\ \forall a \in \mathbb{R}_{\max},\ a \otimes \varepsilon &= \varepsilon \otimes a = \varepsilon \end{aligned}$$

Neutral element for $\otimes$:

$$\forall a \in \mathbb{R}_{\max},\ a \otimes e = e \otimes a = e, \text{ with } e = 0$$

> Hereafter, as it is done in usual algebra, the operator $\otimes$ will usually be omitted in expressions whenever it is clear.

The n-th power of $a \in \mathbb{R}_{\max}$ is naturally introduced in Max-Plus algebra, by using the associative property as shown below

$$a^{\otimes n} = \underbrace{a \otimes a \otimes \cdots \otimes a}_{n \text{ times}}$$

2. Dioid is an algebraic structure that has all properties of a ring except the additive inverse.

for all $n \in \mathbb{N} \setminus \{0\}$. For $n = 0$, we have $a^{\otimes 0} = e$. Furthermore, if $a \in \mathbb{R}$, then $a^{\otimes -1}$ stands for $-a$.

Proposition 1.1. *The inequality $x \geq ax \oplus b$ of $\mathbb{R}_{\max}$ has $\hat{x} = a^\star b$ as smallest solution, where $a^\star$ (Kleene star) is defined by:*

$$a^\star = \bigoplus_{k \in \mathbb{N}} a^{\otimes k}.$$

Moreover, $\hat{x}$ satisfies the implicit equation $x = ax \oplus b$.

1.2.2 Matrices and Vectors in Max-Plus algebra

The previous results are extended to matrices and vectors. The set of matrices of shape $n \times m$ is denoted $\mathbb{R}_{\max}^{n \times m}$. An entry of the matrix $A \in \mathbb{R}_{\max}^{n \times m}$ is denoted as a_{ij} for all $i \in \{1, \ldots, n\}$ and $j \in \{1, \ldots, m\}$.

The sum of matrices $A, B \in \mathbb{R}_{\max}^{n \times m}$ is defined as

$$(A \oplus B)_{ij} = a_{ij} \oplus b_{ij}, \text{ for all } i \in \{1, \ldots, n\} \text{ and } j \in \{1, \ldots, m\}.$$

Furthermore, the commutativity property holds for the sum of $A, B \in \mathbb{R}_{\max}^{n \times m}$, *i.e.*, $A \oplus B = B \oplus A$.

Example 1.7.

$$\begin{pmatrix} e & \varepsilon \\ 3 & 2 \end{pmatrix} \oplus \begin{pmatrix} -5 & 11 \\ 1 & \varepsilon \end{pmatrix} = \begin{pmatrix} e \oplus -5 & \varepsilon \oplus 11 \\ 3 \oplus 1 & 2 \oplus \varepsilon \end{pmatrix} = \begin{pmatrix} e & 11 \\ 3 & 2 \end{pmatrix}.$$

The product of a matrix $A \in \mathbb{R}_{\max}^{n \times m}$ by a scalar $\alpha \in \mathbb{R}_{\max}$ is defined as

$$(\alpha \otimes A)_{ij} = \alpha \otimes a_{ij}, \text{ for all } i \in \{1, \ldots, n\} \text{ and } j \in \{1, \ldots, m\}.$$

The product of matrices $A \in \mathbb{R}_{\max}^{n \times l}$ and $B \in \mathbb{R}_{\max}^{l \times m}$ is defined similarly to the usual algebra, by replacing $+$ by $\oplus$ and $\times$ by $\otimes$. Formally

$$(A \otimes B)_{ik} = \bigoplus_{j=1}^{l} a_{ij} \otimes b_{jk}, \text{ for all } i \in \{1, \ldots, n\} \text{ and } k \in \{1, \ldots, m\}.$$

Example 1.8.

$$\begin{pmatrix} e & \varepsilon \\ 3 & 2 \end{pmatrix} \oplus \begin{pmatrix} -5 & 11 \\ 1 & \varepsilon \end{pmatrix} = \begin{pmatrix} (e \otimes -5) \oplus (\varepsilon \otimes 1) & (e \otimes 11) \oplus (\varepsilon \otimes \varepsilon) \\ (3 \otimes -5) \oplus (2 \otimes 1) & (3 \otimes 11) \oplus (2 \otimes \varepsilon) \end{pmatrix} = \begin{pmatrix} -5 & 11 \\ 3 & 14 \end{pmatrix}$$

Furthermore, the commutativity property does not hold for the product of $A \in \mathbb{R}_{\max}^{n \times l}$ and $B \in \mathbb{R}_{\max}^{l \times m}$, *i.e.*, $A \otimes B \neq B \otimes A$.

The operations $\oplus$ and $\otimes$ of matrices have neutral elements similar to the neutral elements of scalars. Let $\mathcal{E}$ be the zero matrix, *i.e.*, a rectangular matrix whose entries are ε. Let $\mathtt{diag}(a_1, \ldots, a_d)$ be the diagonal matrix, *i.e.*, a square matrix with its diagonal elements $a_1, \ldots, a_d \in \mathbb{R}_{\max}$ and its off diagonal entries ε. A trivial example of a diagonal matrix is the Max-Plus algebraic identity matrix of appropriate dimension $I = \mathtt{diag}(e, \ldots, e)$. The following properties are easily verified:

$$\begin{aligned} \forall A \in \mathbb{R}_{\max}^{n \times m},\ A \oplus \mathcal{E} &= \mathcal{E} \oplus A = A \\ \forall A \in \mathbb{R}_{\max}^{n \times m},\ A \otimes \mathcal{E} &= \mathcal{E} \otimes A = \mathcal{E} \\ \forall A \in \mathbb{R}_{\max}^{n \times m},\ A \otimes I &= I \otimes A = A \end{aligned}$$

The transpose of $A \in \mathbb{R}_{\max}^{n \times m}$, denoted A^{t}, is defined as the usual way, *i.e.*, $(A^{\mathrm{t}})_{ij} = a_{ji}$ for all $i \in \{1, \ldots, n\}$ and $j \in \{1, \ldots, m\}$.

The k-th power of $A \in \mathbb{R}_{\max}^{n \times n}$ is defined as

$$A^{\otimes k} = \underbrace{A \otimes A \otimes \cdots \otimes A}_{k \text{ times}}$$

where $A^{\otimes 0} = I$.

The set $\mathbb{R}_{\max}^{n} = \mathbb{R}_{\max}^{n \times 1}$ refers to the n-th fold Cartesian product of $\mathbb{R}_{\max}$. Its elements can be thought of as points of an affine space, or as vectors. They are denoted by bold symbols, for instance $\mathbf{x} = (x_1, \ldots, x_n)^{\mathrm{t}}$. The unit vector is denoted $\mathbf{u} = (e, \ldots, e)^{\mathrm{t}}$ and $\forall \alpha \in \mathbb{R}_{\max},\ \alpha \otimes \mathbf{u}$ denotes a vector whose entries are α. Furthermore, the j-th column of the identity matrix I is called the j-th basis of $\mathbb{R}_{\max}^{n}$.

Example 1.9. *The vectors* $\mathbf{v}_1 = (e, \varepsilon)^{\mathrm{t}}$ *and* $\mathbf{v}_2 = (\varepsilon, e)^{\mathrm{t}}$ *are at the same time the columns of*

$$I = \begin{pmatrix} e & \varepsilon \\ \varepsilon & e \end{pmatrix}$$

and the basis of $\mathbb{R}_{\max}^{2}$. *The latter statement is due to the fact that any* $\mathbf{v} \in \mathbb{R}_{\max}^{2}$ *may be*

uniquely written as

$$\mathbf{v} = \alpha_1 \mathbf{v}_1 \oplus \alpha_2 \mathbf{v}_2.$$

Any other pair of linearly independent vectors of $\mathbb{R}^2_{\max}$*, such as* $\mathbf{v}_1 = (5, \varepsilon)^{\mathrm{t}}$ *and* $\mathbf{v}_2 = (\varepsilon, 3)^{\mathrm{t}}$*, forms also a basis of* $\mathbb{R}^2_{\max}$*.*

Equation $\mathrm{x} = A\mathrm{x} \oplus \mathrm{b}$ of $\mathbb{R}^n_{\max}$

Let $A \in \mathbb{R}^{n\times n}_{\max}$ be a square matrix, and let $\mathbf{b} \in \mathbb{R}^n_{\max}$ and $\mathbf{x} \in \mathbb{R}^n_{\max}$ be two vectors. According to Proposition 1.1, the equation $\mathbf{x} = A\mathbf{x} \oplus \mathbf{b}$ admits $\mathbf{x} = A^\star \mathbf{b}$ as smallest solution, where

$$A^\star = \bigoplus_{k\in\mathbb{N}} A^{\otimes k}.$$

If A is a lower triangular matrix, *i.e.*, $a_{ij} = \varepsilon$ for all $i \leq j$ with $i, j \in \{1, \ldots, n\}$, then $A^{\otimes k} = \mathcal{E}$ for a sufficient large k (not greater than the matrix dimension).

1.2.3 Completion of $\mathbb{R}_{\max}$ and residuation theory over complete dioids

In a dioid $(\mathcal{D}, \oplus, \otimes)$ with ε and e their zero and unit elements, respectively, the following equivalence is considered

$$\forall a, b \in \mathcal{D},\ a = a \oplus b \iff \exists c \in \mathcal{D}, a = b \oplus c.$$

This equivalence defines the natural order relation $\succeq$ as follows:

$$a \succeq b \iff a = a \oplus b. \tag{1.1}$$

The set $\mathbb{R}_{\max}$ is linearly ordered with respect to $\oplus$ and the order $\succeq$ in this set coincides with the usual linear order $\geq$.

This order relation makes $\mathcal{D}$ to be a partially ordered set such that each pair of elements (a, b) admits the lowest upper bound $\sup\{a, b\}$, usually denoted $a \vee b$, which coincides with $a \oplus b$. Hence, a dioid is in particular a *sup-semilattice*.

The sum and the left and right products preserve this relation:

$$\forall a, b, c \in \mathcal{D},\ a \succeq b \implies \begin{cases} a \oplus c \succeq b \oplus c, & \text{(sum)} \\ a \otimes c \succeq b \otimes c, & \text{(left product)} \\ c \otimes a \succeq c \otimes b, & \text{(right product)} \end{cases}$$

Definition 1.7 (Complete dioid)**.** *A dioid $\mathcal{D}$ is complete if it is closed for infinite sums and the left and right distributivity of the product extend to infinite sums.*

For a complete dioid $\mathcal{D}$, the top element, denoted $\top$, exists and is equal to the sum of all elements of $\mathcal{D}$, *i.e.*, $\top = \bigoplus_{a \in \mathcal{D}} a$. This element respects the absorbing rule, *i.e.*, $\varepsilon \otimes \top = \varepsilon$ and $\top \otimes \varepsilon = \varepsilon$

Complete dioid $\overline{\mathbb{R}}_{\max}$

The dioid $\mathbb{R}_{\max}$ is not complete but it can be completed by adding a top element $\top = +\infty$. Hereafter, rather than using $\mathbb{R}_{\max}$ we use $\overline{\mathbb{R}}_{\max} \stackrel{\text{def}}{=} (\mathbb{R}_{\max} \cup \{\top\}, \oplus, \otimes)$ because we are able to perform computations with infinite values without a careful handling.

For a complete dioid, an inner operation representing the lower bound of the operands, denoted by $\wedge$, automatically exists. The partial order relation presented in Equation (1.1) can be expressed as

$$a \succeq b \iff a = a \oplus b \iff b = a \wedge b. \tag{1.2}$$

where $a \wedge b = \inf\{a, b\}$ is the greatest lower bound of (a, b).

In $\overline{\mathbb{R}}_{\max}$, the operation $a \wedge b$ for the scalars $a, b \in \overline{\mathbb{R}}_{\max}$ coincides with $\min(a, b)$.

For matrices with entries in $\mathcal{D}$, the order $\succeq$ on $\mathcal{D}$ induces a (partial) order on the set $\mathcal{D}^{n \times m}$, *i.e.*, for any $A, B \in \mathcal{D}^{n \times m}$ the following equivalence holds:

$$A \succeq B \iff a_{ij} \succeq b_{ij}, \text{ for all } i \in \{1, \ldots, n\} \text{ and } j \in \{1, \ldots, m\}.$$

Remark 1.5. *The dioid $\mathcal{D}^{n\times n}$ is complete whenever $\mathcal{D}$ is complete because the sum of matrices simply involves the sum of similar entries. Notice that for two matrices $A, B \in \mathcal{D}^{n\times n}$ the following holds:*

$$A \succeq B \iff A = A \oplus B \iff B = A \wedge B,$$

where

$$(A \wedge B)_{ij} = a_{ij} \wedge b_{ij}, \text{ for all } i, j \in \{1, \dots, n\},$$

is consistent, as it is for scalars in $\mathcal{D}$.

Regarding the order structure of dioids, we are interested in system of equations of form $f(x) = y$ where f is an isotone mapping. However, in dioids, the absence of additive inverses increases the difficulty of finding the solutions of systems of linear equations. Hence, we must weaken the notion of solution, denoting that a sub-solution of $f(x) = y$ is a x which satisfies $f(x) \preceq y$. Nevertheless, as in usual algebra the solution of $f(x) = y$ does not always exist and if it does, it is not necessary unique.

Below, we present the greatest solution of $f(x) \preceq y$ using residuation theory.

Definition 1.8 (Residual and residuated mapping)**.** *Let $\mathcal{D}$ and $\mathcal{C}$ be two complete dioids and $f : \mathcal{D} \to \mathcal{C}$ be an isotone mapping, i.e., order preserving. Then, f is residuable if, for all $y \in \mathcal{C}$, there exists the greatest x to $f(x) \preceq y$, hereafter denoted $f^\sharp(y)$. The monotone mapping $f^\sharp : \mathcal{C} \to \mathcal{D}, y \mapsto \bigoplus\{x \in \mathcal{D} \mid f(x) \preceq y\}$ is called residual of f. Moreover, $f^\sharp$ is the unique mapping that verifies the following statements $f \circ f^\sharp \preceq Id_{\mathcal{C}}$ et $f^\sharp \circ f \succeq Id_{\mathcal{D}}$ where $Id_{\mathcal{D}}$ and $Id_{\mathcal{C}}$ are the identity mappings on $\mathcal{D}$ and $\mathcal{C}$, respectively.*

Example 1.10. *Mappings $L_a : \mathcal{D} \to \mathcal{D}, x \mapsto ax$ and $R_a : \mathcal{D} \to \mathcal{D}, x \mapsto xa$ for any $a \in \mathcal{D}$ are both residuated. Their residuals are isotone mappings, denoted $L_a^\sharp : \mathcal{D} \to \mathcal{D}, y \mapsto a\backslash y$ ("left residuation by a") and $R_a^\sharp : \mathcal{D} \to \mathcal{D}, y \mapsto y/a$ ("right residuation by a"), respectively. Hence, the greatest solutions of $ax \preceq b$ and $xa \preceq b$ are $a\backslash b$ and b/a, respectively.*

In Appendix A.1, we collect some useful properties of left and right multiplication and their residuals.

> In $\overline{\mathbb{R}}_{\max}$, the product is commutative, hence $xa = ax \leq y$ admits $x = a\backslash y = y/a$ as greatest solution where operators $\backslash$ and $/$ are the classical subtraction $-$, *i.e.*, $x = y - a$ with the convention that $-\infty + \infty = +\infty$.

One can extend the residuation computation to matrices of dioids. Matrix inequalities of form $f(X) \preceq Y$, with X a rectangular matrix of any arbitrary dimension, are also residuable. Then

$$L_A : \mathcal{D}^{p\times m} \to \mathcal{D}^{n\times m}, X \mapsto AX$$
$$R_C : \mathcal{D}^{n\times p} \to \mathcal{D}^{n\times m}, X \mapsto XC$$

with $A \in \mathcal{D}^{n\times p}$ and $C \in \mathcal{D}^{p\times m}$, are both residuated and their residuals are given by

$$L_A^\sharp : \mathcal{D}^{n\times m} \to \mathcal{D}^{p\times m}, Y \mapsto A\backslash Y$$
$$R_A^\sharp : \mathcal{D}^{n\times m} \to \mathcal{D}^{n\times p}, Y \mapsto Y/C.$$

Hence, the greatest solutions of $AX \preceq B$ and $XC \preceq B$, where $B \in \mathcal{D}^{n\times m}$, are $A\backslash B$ and B/C, respectively. Computationally,

$$(A\backslash B)_{ij} = \bigwedge_{k=1}^{n} a_{ki}\backslash b_{kj}, \text{for all } i \in \{1,\dots,p\} \text{ and } j \in \{1,\dots,m\}$$

and

$$(B/C)_{ij} = \bigwedge_{k=1}^{m} b_{ik}/c_{jk}, \text{for all } i \in \{1,\dots,n\} \text{ and } j \in \{1,\dots,p\}$$

In $\overline{\mathbb{R}}_{\max}$, $A\backslash B = -A^{\mathrm{t}} \otimes' B = -(A^{\mathrm{t}} \otimes (-B))$ and $B/C = B \otimes' (-C^{\mathrm{t}}) = -((-B) \otimes C^{\mathrm{t}})$ where A, B and C are matrices of appropriate dimension and $\otimes'$ is the $\otimes$-product of matrices but by replacing $\oplus$ by $\min$ [a].

a. The max and min operations are dual, *i.e.*, for $x, y \in \overline{\mathbb{R}}_{\max}$, the following equalities hold: $\min(x, y) = -\max(-x, -y)$ and $\max(x, y) = -\min(-x, -y)$.

The following example deals with the solution of inequalities of form $f(\mathbf{x}) \preceq \mathbf{y}$, which is a specific case of the previous calculations that are useful in the sequel of this work.

Example 1.11. *Given $A \in \mathcal{D}^{n\times p}$, the mapping $L_A : \mathcal{D}^p \to \mathcal{D}^n, \mathbf{x} \mapsto A\mathbf{x}$ is residuable. For all $\mathbf{y} \in \mathcal{D}^n$, the residual of L_A is $L_A^\sharp(\mathbf{y}) = A\backslash\mathbf{y}$ which is given by $(A\backslash y)_i = \bigwedge_{j=1}^{n} a_{ji}\backslash y_j$, for all $i \in \{1,\dots,p\}$ where $a_{ji}\backslash y_j$ is the greatest solution of $a_{ji}x \preceq y_j$. Furthermore, due to isotonicity, $A\mathbf{x} \preceq \mathbf{y}$ implies $A(A\backslash\mathbf{y}) \preceq \mathbf{y}$, and it follows immediately that $A\mathbf{x} = \mathbf{b}$ has solution if and only if $A(A\backslash\mathbf{b}) = \mathbf{b}$.*

In $\overline{\mathbb{R}}_{\max}$, $A\mathbf{x} \leq \mathbf{y}$, with $A \in \overline{\mathbb{R}}_{\max}^{n\times p}$, $\mathbf{x} \in \overline{\mathbb{R}}_{\max}^{p}$ and $\mathbf{y} \in \overline{\mathbb{R}}_{\max}^{n}$, admits $A\backslash\mathbf{y}$ as its greatest solution, where

$$(A\backslash\mathbf{y})_i = \min_{1\leq j\leq n} (-a_{ji} + y_j) \text{ for all } i \in \{1,\ldots,p\}. \tag{1.3}$$

Example 1.12. *Given* $A = \begin{pmatrix} 2 & \varepsilon \\ \top & 3 \end{pmatrix}$ *and* $\mathbf{y} = \begin{pmatrix} \varepsilon \\ 1 \end{pmatrix}$ *then*

$$\mathbf{x} = A\backslash\mathbf{y} = \begin{pmatrix} \min(-2+\varepsilon, -\top+1) = \min(\varepsilon,\varepsilon) \\ \min(-\top+\varepsilon, -3+1) = \min(\top,-2) \end{pmatrix} = \begin{pmatrix} \varepsilon \\ -2 \end{pmatrix}$$

which is the greatest vector $\mathbf{x}$ *such that* $A\mathbf{x} \leq \mathbf{y}$. *Furthermore, in this case* $A\mathbf{x} = \mathbf{y}$ *holds.*

Theorem 1.1. *Let* $\mathcal{D}$ *and* $\mathcal{C}$ *be two complete dioids and* $f : \mathcal{D} \to \mathcal{C}$ *be a residuated mapping. Then*

$$f \circ f^\sharp \circ f = f \text{ and } f^\sharp \circ f \circ f^\sharp = f^\sharp.$$

Proposition 1.2. *Let* $\mathcal{D}$ *and* $\mathcal{C}$ *be two complete dioids and* $f, g : \mathcal{D} \to \mathcal{C}$ *be two residuated mappings. The greatest solution of* $f(x) = g(x)$ *is equal to the greatest fixed-point of the isotone mapping* $\Pi : \mathcal{D} \to \mathcal{D}$, $\Pi(x) = x \wedge g^\sharp(f(x)) \wedge f^\sharp(g(x))$.

Proof. The following equivalences hold:

$$\begin{aligned} f(x) = g(x) &\iff f(x) \succeq g(x) \text{ and } g(x) \succeq f(x) \\ &\iff g^\sharp(f(x)) \succeq x \text{ and } f^\sharp(g(x)) \succeq x \\ &\iff g^\sharp(f(x)) \wedge f^\sharp(g(x)) \succeq x \\ &\iff g^\sharp(f(x)) \wedge f^\sharp(g(x)) \wedge x = x \end{aligned}$$

Hence, the greatest fixed-point of $\Pi(x) = x \wedge g^\sharp(f(x)) \wedge f^\sharp(g(x))$ is the greatest solution of $f(x) = g(x)$. Moreover, $f, f^\sharp, g$ and $g^\sharp$ are isotone mappings, and thus the mapping Π is also isotone. □

As a consequence of the previous Proposition 1.2, if the greatest solution of $f(x) = g(x)$ exists, then it can be obtained by solving $x_k = \Pi(x_{k-1})$ with convergence (*i.e.*, $x_k = x_{k-1}$)

in a finite number of iterations k. Thus, starting from an arbitrary x_0, we obtain x_k less than or equal to x_0 because $x_k \preceq x_{k-1}$ for all k.

Remark 1.6. *In [44], it is shown that the complexity of computing this problem is pseudo-polynomial if it converges (it exists a finite solution), i.e., the convergence rate is polynomial according to the distance between the greatest fixed-point x_k and the initial value x_0. Nevertheless, if $x_k = \Pi(x_{k-1})$ is not solvable, i.e., $f(x) = g(x)$ does not possess a finite solution, then $x_k = \Pi(x_{k-1})$ is likely to run infinitely and some other techniques shall be considered in order to avoid this bad behavior.*

Fixed-point of mappings in dioids of matrices are a natural extension of Proposition 1.2 for scalars. Hence, it is possible to compute the greatest solution $\mathbf{x} \in \mathcal{D}^n$ of the vertical concatenation of $s \geq 1$ two-sided equations of form $\mathbf{a}^{\mathrm{t}}\mathbf{x} = \mathbf{b}^{\mathrm{t}}\mathbf{x}$, where $\mathbf{a}, \mathbf{b} \in \mathcal{D}^n$, as shown below.

On the two-sided equation $A\mathbf{x} = B\mathbf{x}$ in $\overline{\mathbb{R}}_{\max}^n$

Consider the matrices $A, B \in \overline{\mathbb{R}}_{\max}^{p\times n}$ and the vector $\mathbf{x} \in \overline{\mathbb{R}}_{\max}^n$, then Proposition 1.2 implies that the greatest solution of $A\mathbf{x} = B\mathbf{x}$ is obtained by considering the greatest fixed-point of the mapping

$$\Pi(\mathbf{x}) = \mathbf{x} \wedge A \backslash (B\mathbf{x}) \wedge B \backslash (A\mathbf{x}) \tag{1.4}$$

Previously, it has been shown how residuation theory can be applied to determine the greatest solution of inequalities $f(x) \preceq y$ in complete dioids, denoted $f^\sharp(y)$. Dually, it is of course also possible to determine the least solution x of inequalities such as $y \preceq f(x)$, denoted $f^\flat(y)$. Hence, $f^\flat(y)$ is called the dual residual of f.

Example 1.13. *The mapping $T_a(x) : \mathcal{D} \to \mathcal{D}, x \mapsto x \oplus a$, is dually residuated, and its residual is denoted by $T_a^\flat(y) : \mathcal{D} \to \mathcal{D} : y \mapsto y \oplus a$. It is worth to mention that*

$$y \oplus a = \varepsilon \iff a \succeq y.$$

Definition 1.9 ([45])**.** *A matrix in a dioid $\mathcal{D}^{n\times p}$ is said to be row (column) G-astic if it has at least one non-ε element in each row (column). Formally, for $A \in \mathcal{D}^{n\times p}$ we have*

that

$$\begin{aligned} A \text{ is row G-astic} &: \text{if } \bigoplus_{j=1}^{p} a_{ij} \neq \varepsilon,\ i = 1, 2, \ldots, n \\ A \text{ is column G-astic} &: \text{if } \bigoplus_{i=1}^{n} a_{ij} \neq \varepsilon,\ j = 1, 2, \ldots, p. \end{aligned}$$

Furthermore, A is said to be doubly G-astic if it is, at the same time, row and column G-astic [46].

Proposition 1.3. *In $\overline{\mathbb{R}}_{\max}$, $\mathbf{y} \leq A\mathbf{x}$ with $A \in \overline{\mathbb{R}}_{\max}^{n\times p}$, $\mathbf{x} \in \overline{\mathbb{R}}_{\max}^{p}$ and $\mathbf{y} \in \overline{\mathbb{R}}_{\max}^{n}$ admits a finite least solution if and only A is doubly G-astic with only one element different of ε per row and column.*

Proof. First, if A is doubly G-astic with only one element different of ε per row and column, then $(A\mathbf{x})_i = a_{ij(i)} \otimes x_{j(i)}$ with $j(i) \in \{1, \ldots, p\}$ for $i = 1, 2, \ldots, n$ such that $j(1) \neq j(2) \neq \ldots \neq j(n)$. Thus,

$$y_i \leq (A\mathbf{x})_i \iff y_i \leq a_{ij(i)} \otimes x_{j(i)} \iff x_{j(i)} \geq a_{ij(i)} \backslash y_i,$$

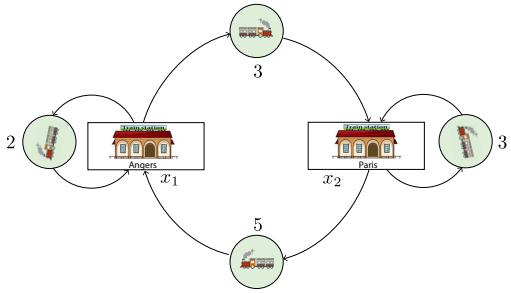
which is the element-wise finite least solution of $\mathbf{y} \leq A\mathbf{x}$. □

A general solution for this problem is given in Appendix B.2.

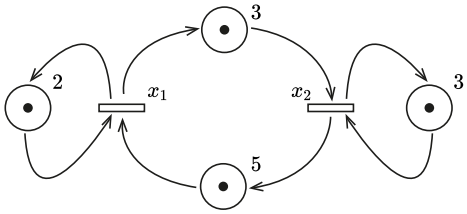
1.3 Modelling of TEG in $\overline{\mathbb{R}}_{\max}$: the Max-Plus Linear (MPL) systems

1.3.1 Motivating example

Example 1.14 ([1])**.** *Imagine a railway network consisting of two stations ($Angers \coloneqq x_1$ and $Paris \coloneqq x_2$) and three lines: one inner loop with two rail tracks each (one for $x_1 \to x_2$ and other for $x_2 \to x_1$); two outer loops with one rail track each. The basic structure of this railway network is depicted in Figure 1.9a. It is assumed that the train company operates one train on each track initially; the travel times are fixed as indicated on the arcs; trains scheduled to depart must wait for all arriving trains before departing because passengers are able to change line at the stations (to change from the inner to an outer loop or vice versa); and departures occur as soon as possible (refer to Remark 1.4). As mentioned*



(a) Precedence graph



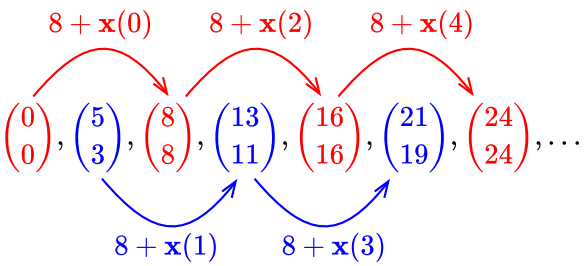
(b) TEG model

Figure 1.9 – Railway network model

before, TEG are suitable to model synchronization phenomena. Figure 1.9b depicts this model and the following equations are obtained:

$$\begin{pmatrix} x_1(k) \\ x_2(k) \end{pmatrix} = \begin{pmatrix} \max(2 + x_1(k-1), 5 + x_2(k-1)) \\ \max(3 + x_1(k-1), 3 + x_2(k-1)) \end{pmatrix}$$

where $(x_1(0), x_2(0))^{\mathrm{t}} \in \mathbb{R}^2$ *is supposed to be known. For instance, for* $(x_1(0), x_2(0))^{\mathrm{t}} = (0, 0)^{\mathrm{t}}$*, the following timetable can be achieved*



such that calculating the differences between adjacent elements of the timetable, i.e., $\forall k,\ \mathbf{x}(k+1) - \mathbf{x}(k)$ *we obtain*

$$\begin{pmatrix} 5 \\ 3 \end{pmatrix}, \begin{pmatrix} 3 \\ 5 \end{pmatrix}, \begin{pmatrix} 5 \\ 3 \end{pmatrix}, \begin{pmatrix} 3 \\ 5 \end{pmatrix}, \begin{pmatrix} 5 \\ 3 \end{pmatrix}, \begin{pmatrix} 3 \\ 5 \end{pmatrix}, \ldots,$$

i.e., on average a train leaves station x_1 *or* x_2 *at every* 4 *unit of time, but the simulated timetable is considered to be* 2*-periodic, since it is not hard to see that* $\forall k \geq 0,\ \mathbf{x}(k+2) = 8 + \mathbf{x}(k)$*.*

Furthermore, this system is nonlinear in the conventional algebra, however it can be

expressed as the following linear system in $\overline{\mathbb{R}}^2_{\max}$*:*

$$\begin{pmatrix} x_1(k) \\ x_2(k) \end{pmatrix} = \begin{pmatrix} 2 \otimes x_1(k-1) \oplus 5 \otimes x_2(k-1) \\ 3 \otimes x_1(k-1) \oplus 3 \otimes x_2(k-1) \end{pmatrix} = \begin{pmatrix} 2 & 5 \\ 3 & 3 \end{pmatrix} \otimes \begin{pmatrix} x_1(k-1) \\ x_2(k-1) \end{pmatrix}$$

This motivating example suggests that a TEG appears to have a state-space characterization similar to the one found in classical time-driven systems.

1.3.2 State-space equations

Example 1.15. *The TEG of Figure 1.10 models a simple manufacturing system. The linear dynamical system in Max-Plus algebra, with* $x_i(k)$*,* $u_j(k)$*,* $z(k)$ *being the time instants that the transitions* x_i*,* u_j *and* z *fire at* k*, is*

$$\begin{cases} x_1(k) & \geq & 1 \otimes x_2(k-2) \oplus 5 \otimes u_1(k-1) \\ x_2(k) & \geq & x_1(k) \oplus 3 \otimes x_3(k-3) \\ x_3(k) & \geq & 5 \otimes u_2(k) \\ z(k) & \geq & 2 \otimes x_2(k) \oplus 3 \otimes x_3(k-1) \end{cases}$$

Rewriting the system in matrix form, one obtains

$$\begin{cases} \mathbf{x}(k) & \geq & A_0\mathbf{x}(k) \oplus A_1\mathbf{x}(k-1) \oplus A_2\mathbf{x}(k-2) \oplus A_3\mathbf{x}(k-3) \oplus B_0\mathbf{u}(k) \oplus B_1\mathbf{u}(k-1), \\ z(k) & \geq & C_0\mathbf{x}(k) \oplus C_1\mathbf{x}(k-1), \end{cases}$$

with $A_0 = \begin{pmatrix} \varepsilon & \varepsilon & \varepsilon \\ e & \varepsilon & \varepsilon \\ \varepsilon & \varepsilon & \varepsilon \end{pmatrix}$, $A_1 = \mathcal{E}$, $A_2 = \begin{pmatrix} \varepsilon & 1 & \varepsilon \\ \varepsilon & \varepsilon & \varepsilon \\ \varepsilon & \varepsilon & \varepsilon \end{pmatrix}$, $A_3 = \begin{pmatrix} \varepsilon & \varepsilon & \varepsilon \\ \varepsilon & \varepsilon & 3 \\ \varepsilon & \varepsilon & \varepsilon \end{pmatrix}$, $B_0 = \begin{pmatrix} \varepsilon & \varepsilon \\ \varepsilon & \varepsilon \\ 5 & \varepsilon \end{pmatrix}$,

$B_1 = \begin{pmatrix} 5 & \varepsilon \\ \varepsilon & \varepsilon \\ \varepsilon & \varepsilon \end{pmatrix}$, $C_0 = \begin{pmatrix} \varepsilon & 2 & \varepsilon \end{pmatrix}$ *and* $C_1 = \begin{pmatrix} \varepsilon & \varepsilon & 3 \end{pmatrix}$.

Generally speaking, for any TEG (cf. Definition 1.6), one obtains the following inequalities:

$$\begin{cases} \mathbf{x}(k) \geq \bigoplus_{i=0}^{M} A_i\mathbf{x}(k-i) \oplus \bigoplus_{j=0}^{M} B_j\mathbf{u}(k-j), & \text{(1.5a)} \\ \mathbf{z}(k) \geq \bigoplus_{l=0}^{M} C_l\mathbf{x}(k-l), & \text{(1.5b)} \end{cases}$$

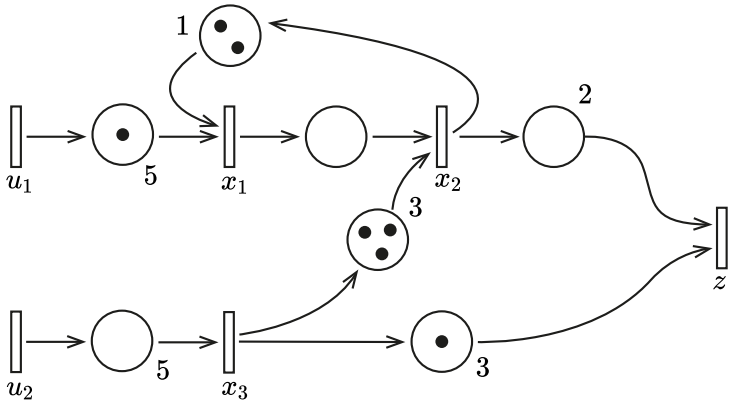


Figure 1.10 – TEG of a simple manufacturing system

where M is the maximal initial marking, $\mathbf{x} \in \overline{\mathbb{R}}_{\max}^{n}$, $\mathbf{u} \in \overline{\mathbb{R}}_{\max}^{p}$, $\mathbf{z} \in \overline{\mathbb{R}}_{\max}^{q}$, $A_0, A_1, \ldots, A_M \in \overline{\mathbb{R}}_{\max}^{n\times n}$, $B_0, B_1, \ldots, B_M \in \overline{\mathbb{R}}_{\max}^{n\times p}$ and $C_0, C_1, \ldots, C_M \in \overline{\mathbb{R}}_{\max}^{q\times n}$.

Inequality (1.5) is rewritten as

$$\mathbf{x}(k) \geq A_0\mathbf{x}(k) \oplus \left(\bigoplus_{i=1}^{M} A_i\mathbf{x}(k-i) \oplus \bigoplus_{j=0}^{M} B_j\mathbf{u}(k-j) \right),$$

and according to Proposition 1.1 its smallest solution is

$$\hat{\mathbf{x}}(k) = \bigoplus_{i=1}^{M} A_0^{\star}A_i\mathbf{x}(k-i) \oplus \bigoplus_{j=0}^{M} A_0^{\star}B_j\mathbf{u}(k-j).$$

Furthermore, $\hat{\mathbf{x}}(k)$ is also the smallest solution of the implicit equation

$$\mathbf{x}(k) = A_0\mathbf{x}(k) \oplus \left(\bigoplus_{i=1}^{M} A_i\mathbf{x}(k-i) \oplus \bigoplus_{j=0}^{M} B_j\mathbf{u}(k-j) \right), \tag{1.6}$$

and thus

$$\begin{cases} \mathbf{x}(k) = \bigoplus_{i=1}^{M} \tilde{A}_i\mathbf{x}(k-i) \oplus \bigoplus_{j=0}^{M} \tilde{B}_j\mathbf{u}(k-j), \text{ with } \tilde{A}_i = A_0^{\star}A_i \text{ and } \tilde{B}_j = A_0^{\star}B_j, & \text{(1.7a)} \\ \mathbf{z}(k) = \bigoplus_{l=0}^{M} C_l\mathbf{x}(k-l), & \text{(1.7b)} \end{cases}$$

represents the behavior of the associated TEG without implicit terms.

The non-null entries of A_0 correspond to holding times of the token-free places in

the initial marking. In practice, TEG are assumed to be live (otherwise some transitions would be frozen), this implies that there is an appropriate permutation of the transitions numbering such that A_0 can be written in strictly lower triangular form, *i.e.*, the infinite sum that defines $A_0^\star$ becomes finite.

Remark 1.7. *Rather than using $A_0^\star$, the implicit part of Equation* (1.6) *can also be eliminated by successive substitutions of scalar variables, with careful attention to the order of these substitutions.*

Eliminating the implicit part of Inequality (1.5) supposes the *earliest firing rule* (see Remark 1.4), but also that the tokens of the initial marking are immediately available [a], *i.e.*, the *initial conditions* $\{\mathbf{x}(k)\}_{k<0}$ took place on date $-\infty$. Other nonzero initial conditions can be circumvented using input transitions of the TEG.

a. It is for this reason that in the state representation, inequalities (signs $\geq$) become equations (signs $=$) (see Remark 1.4).

We can transform any M-order system to a first order system by considering an augmented version of the state-space. For the TEG this means:

— any place between two internal transitions must contain exactly one token, *i.e.*, for $i \neq 1$ the matrices A_i must be null;

— any place between an input transition and an internal transition must be without token, *i.e.*, for $j > 0$ the matrices B_j must be null;

— any place between an internal transition and an output transition must be without token, *i.e.*, for $l > 0$ the matrices C_l must be null.

Max-Plus Linear (MPL) system

Hence, from the implicit Equation (1.7), it is always possible to obtain the following canonical form:

$$\begin{cases} \mathbf{x}(k) = A\mathbf{x}(k-1) \oplus B\mathbf{u}(k), & \text{(1.8a)} \\ \mathbf{z}(k) = C\mathbf{x}(k), & \text{(1.8b)} \end{cases}$$

where $\mathbf{x}$, $\mathbf{u}$, $\mathbf{z}$, A, B and C are of appropriate dimension.

By analogy with the theory of time-driven linear systems, Equation (1.8a) is the state equation and represents the internal dynamics, Equation (1.8b) is the observation (or output) equation, $\mathbf{x}$ is the state vector, $\mathbf{u}$ is the input or control vector and $\mathbf{z}$ is the output or observation vector.

Nevertheless, a generic MPL system equation is represented by:

$$\mathbf{y}(k) = A\mathbf{x}(k-1), \tag{1.9}$$

where $A \in \overline{\mathbb{R}}_{\max}^{n\times p}$, $\mathbf{y}$ and $\mathbf{x}$ are vectors of appropriate dimension.

Remark 1.8. *Equation* (1.9) *is generic because it can represent either an autonomous MPL system (i.e.,* $B\mathbf{u}(k) = \boldsymbol{\varepsilon}$*) or a non-autonomous MPL system (i.e.,* $B\mathbf{u}(k) \neq \boldsymbol{\varepsilon}$*), both given by Equation* (1.8a)*. It can also represent Equation* (1.8b)*.*

Example 1.16. *Consider the non-autonomous MPL system given by Equation* (1.8)*:*

$$\mathbf{x}(k) = \begin{pmatrix} 3 & 2 & 2 \\ e & 1 & 3 \\ 2 & 1 & e \end{pmatrix} \mathbf{x}(k-1) \oplus \begin{pmatrix} e & \varepsilon \\ \varepsilon & e \\ \varepsilon & \varepsilon \end{pmatrix} \mathbf{u}(k),$$

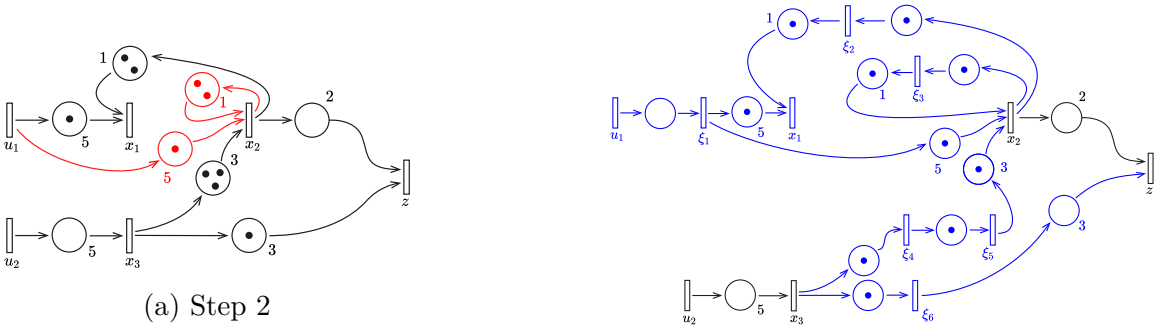
where $\mathbf{x}(k) \in \overline{\mathbb{R}}_{\max}^{3}$ *and* $\mathbf{u}(k) \in \overline{\mathbb{R}}_{\max}^{2}$*. The equivalent generic system is given by (see Equation* (1.9)*):*

$$\mathbf{x}(k) = \begin{pmatrix} 3 & 2 & 2 & e & \varepsilon \\ e & 1 & 3 & \varepsilon & e \\ 2 & 1 & e & \varepsilon & e \end{pmatrix} \tilde{\mathbf{x}}(k-1),$$

with $\tilde{\mathbf{x}}(k-1) = (\mathbf{x}^{\mathrm{t}}(k-1), \mathbf{u}^{\mathrm{t}}(k))^{\mathrm{t}} \in \overline{\mathbb{R}}_{\max}^{5}$*.*

Example 1.17. *We recall Example 1.15. In order to obtain a representation of the TEG of Figure 1.10, we consider the following steps:*

1. *Compute* $A_0^\star$*:* $A_0^\star = \begin{pmatrix} e & \varepsilon & \varepsilon \\ e & e & \varepsilon \\ \varepsilon & \varepsilon & e \end{pmatrix}$
2. *Compute* $\tilde{A}_1 = A_0^\star A_1$*,* $\tilde{A}_2 = A_0^\star A_2$*,* $\tilde{A}_3 = A_0^\star A_3$*,* $\tilde{B}_0 = A_0^\star B_0$ *and* $\tilde{B}_1 = A_0^\star B_1$*:* $\tilde{A}_1 = \mathcal{E}$*,* $\tilde{A}_2 = \begin{pmatrix} \varepsilon & 1 & \varepsilon \\ \varepsilon & 1 & \varepsilon \\ \varepsilon & \varepsilon & \varepsilon \end{pmatrix}$*,* $\tilde{A}_3 = \begin{pmatrix} \varepsilon & \varepsilon & \varepsilon \\ \varepsilon & \varepsilon & 3 \\ \varepsilon & \varepsilon & \varepsilon \end{pmatrix}$*,* $\tilde{B}_0 = \begin{pmatrix} \varepsilon & \varepsilon \\ \varepsilon & \varepsilon \\ 5 & \varepsilon \end{pmatrix}$ *and* $\tilde{B}_1 = \begin{pmatrix} 5 & \varepsilon \\ 5 & \varepsilon \\ \varepsilon & \varepsilon \end{pmatrix}$



(a) Step 2

(b) Step 3

Figure 1.11 – Canonical form of the TEG of Example 1.17

Alternative to 2. Substitute $x_1(k)$ into $x_2(k)$

$$\begin{cases} x_1(k) &= 1 \otimes x_2(k-2) \oplus 5 \otimes u_1(k-1) \\ x_2(k) &= x_1(k) \oplus 3 \otimes x_3(k-3) \\ &= 1 \otimes x_2(k-2) \oplus 3 \otimes x_3(k-3) \oplus 5 \otimes u_1(k-1) \\ x_3(k) &= 5 \otimes u_2(k) \\ z(k) &= 2 \otimes x_2(k) \oplus 3 \otimes x_3(k-1) \end{cases}$$

This step is represented by Figure 1.11a.

3. *Introduce new variables $\xi_1, \xi_2, \ldots, \xi_6$ such that:*

$$\begin{cases} x_1(k) &= 1 \otimes \xi_2(k-1) \oplus 5 \otimes \xi_1(k-1) \\ x_2(k) &= 1 \otimes \xi_3(k-1) \oplus 3 \otimes \xi_5(k-1) \oplus 5 \otimes \xi_1(k-1) \\ x_3(k) &= 5 \otimes u_2(k) \\ \xi_1(k) &= u_1(k) \\ \xi_2(k) &= x_2(k-1) \\ \xi_3(k) &= x_2(k-1) \\ \xi_4(k) &= x_3(k-1) \\ \xi_5(k) &= \xi_4(k-1) \\ \xi_6(k) &= x_3(k-1) \\ z(k) &= 2 \otimes x_2(k) \oplus 3 \otimes \xi_6(k) \end{cases}$$

which is clearly a first order system that is depicted by Figure 1.11b and charac-

terized by Equation (1.8)*, i.e.,*

$$\begin{cases} \begin{pmatrix} x_1(k) \\ x_2(k) \\ x_3(k) \\ \xi_1(k) \\ \xi_2(k) \\ \xi_3(k) \\ \xi_4(k) \\ \xi_5(k) \\ \xi_6(k) \end{pmatrix} = \begin{pmatrix} \varepsilon & \varepsilon & \varepsilon & 5 & 1 & \varepsilon & \varepsilon & \varepsilon & \varepsilon \\ \varepsilon & \varepsilon & \varepsilon & 5 & \varepsilon & 1 & \varepsilon & 3 & \varepsilon \\ \varepsilon & \varepsilon & \varepsilon & \varepsilon & \varepsilon & \varepsilon & \varepsilon & \varepsilon & \varepsilon \\ \varepsilon & \varepsilon & \varepsilon & \varepsilon & \varepsilon & \varepsilon & \varepsilon & \varepsilon & \varepsilon \\ \varepsilon & e & \varepsilon & \varepsilon & \varepsilon & \varepsilon & \varepsilon & \varepsilon & \varepsilon \\ \varepsilon & e & \varepsilon & \varepsilon & \varepsilon & \varepsilon & \varepsilon & \varepsilon & \varepsilon \\ \varepsilon & \varepsilon & e & \varepsilon & \varepsilon & \varepsilon & \varepsilon & \varepsilon & \varepsilon \\ \varepsilon & \varepsilon & \varepsilon & \varepsilon & \varepsilon & \varepsilon & e & \varepsilon & \varepsilon \\ \varepsilon & \varepsilon & e & \varepsilon & \varepsilon & \varepsilon & \varepsilon & \varepsilon & \varepsilon \end{pmatrix} \begin{pmatrix} x_1(k-1) \\ x_2(k-1) \\ x_3(k-1) \\ \xi_1(k-1) \\ \xi_2(k-1) \\ \xi_3(k-1) \\ \xi_4(k-1) \\ \xi_5(k-1) \\ \xi_6(k-1) \end{pmatrix} \oplus \begin{pmatrix} \varepsilon & \varepsilon \\ \varepsilon & \varepsilon \\ \varepsilon & 5 \\ e & \varepsilon \\ \varepsilon & \varepsilon \\ \varepsilon & \varepsilon \\ \varepsilon & \varepsilon \\ \varepsilon & \varepsilon \\ \varepsilon & \varepsilon \end{pmatrix} \begin{pmatrix} u_1(k) \\ u_2(k) \end{pmatrix}, \\ z(k) = \begin{pmatrix} \varepsilon & 2 & \varepsilon & \varepsilon & \varepsilon & \varepsilon & \varepsilon & \varepsilon & 3 \end{pmatrix} \begin{pmatrix} x_1(k) \\ x_2(k) \\ x_3(k) \\ \xi_1(k) \\ \xi_2(k) \\ \xi_3(k) \\ \xi_4(k) \\ \xi_5(k) \\ \xi_6(k) \end{pmatrix}. \end{cases}$$

It is also possible to consider a different approach to eliminate the implicit part of Inequality (1.5). First, suppose also that the corresponding TEG is under the *earliest fire rule*, then $\geq$ becomes $=$. Considering an augmented version of the state-space, the M-order system can be rewritten as

$$\begin{cases} \mathbf{x}(k) = A_0\mathbf{x}(k) \oplus A_1\mathbf{x}(k-1) \oplus B\mathbf{u}(k), & \text{(1.10a)} \\ \mathbf{z}(k) = C\mathbf{x}(k), & \text{(1.10b)} \end{cases}$$

where $\mathbf{x}$, $\mathbf{u}$, $\mathbf{z}$, A_0, A_1, B and C are of appropriate dimension. Since the TEG is considered to be live, then

$$\begin{cases} \mathbf{x}(k) = A_0^\star A_1\mathbf{x}(k-1) \oplus A_0^\star B\mathbf{u}(k), & \text{(1.11a)} \\ \mathbf{z}(k) = C\mathbf{x}(k), & \text{(1.11b)} \end{cases}$$

Triangular form of Equation 1.10

Alternatively, Equation (1.10) can be rewritten as:

$$\begin{array}{rcl} x_1(k) & = & y_1(k-1), \\ x_2(k) & = & a^0_{21}x_1(k) \oplus y_2(k-1), \\ x_3(k) & = & a^0_{31}x_1(k) \oplus a^0_{32}x_2(k-1) \oplus y_3(k-1), \\ & \vdots & \\ x_n(k) & = & \left(\bigoplus_{l=1}^{n-1} a^0_{nl}x_l(k)\right) \oplus y_n(k-1), \\ \mathbf{z}(k) & = & C\mathbf{x}(k), \end{array} \tag{1.12}$$

where $\mathbf{y}(k-1) = A_1\mathbf{x}(k-1) \oplus B\mathbf{u}(k)$ and $a^0_{ij} \in A_0$ for all $i, j \in \{1, \ldots, n\}$.

1.3.3 MPL systems as Piece-Wise Affine (PWA) systems

MPL systems are linear over $\overline{\mathbb{R}}_{\max}$. Nevertheless, in [47], it has been shown that it was always possible to represent the nonlinear dynamics of MPL systems in $\mathbb{R} \cup \{-\infty\}$ as linear Piece-Wise Affine (PWA) systems, *i.e.*, as the union of affine partitions of the nonlinear dynamics.

Equation (1.9) represents a generic MPL system, which is also represented by

$$y_i(k) = \max_{1 \le i \le p}(a_{ij} + x_j(k-1)),\ i = 1, 2, \ldots, n, \tag{1.13}$$

where $y_i, a_{ij}, x_j \in \mathbb{R} \cup \{-\infty\}$.

For each $i \in \{1, \ldots, n\}$, $\exists\, g_i \in \{1, \ldots, p\}$ such that $a_{ig_i} + x_{g_i}(k-1) \ge a_{ij} + x_j(k-1)$ and then $y_i(k) = a_{ig_i} + x_{g_i}(k-1)$. Summing up, $\forall i$ we define $\mathbf{g} = (g_1, g_2, \ldots, g_n) \in \{1, \ldots, p\}^n$, and then

$$\mathbf{y}(k) = P_{\mathbf{g}}\mathbf{x}(k-1) + \mathbf{\Xi}_{\mathbf{g}},\ \text{for } \mathbf{x}(k-1) \in R_{\mathbf{g}} \tag{1.14}$$

where $\forall (i, j) \in \{1, \ldots, n\} \times \{1, \ldots, p\}$ we have $p^{ij}_{\mathbf{g}} \in P_{\mathbf{g}} \in \{0, 1\}^{n \times p}$, as defined below

$$p^{ij}_{\mathbf{g}} = \begin{cases} 1 & \text{if } j = g_i, \\ 0 & \text{otherwise,} \end{cases}$$

the vector of constants $\boldsymbol{\Xi}_{\mathbf{g}}$ is given by

$$\boldsymbol{\Xi}_{\mathbf{g}} = \begin{pmatrix} a_{1g_1} & a_{2g_2} & \dots & a_{ng_n} \end{pmatrix}^{\mathrm{t}}$$

and each one of the p^n affine dynamics is strictly valid in the region $R_{\mathbf{g}}$, which is defined by

$$R_{\mathbf{g}} = \bigcap_{i=1}^{n} \bigcap_{\substack{j=1 \\ j \neq g_i}}^{p} \left\{ \mathbf{x} \in (\mathbb{R} \cup \{-\infty\})^p \mid x_j - x_{g_i} \leq a_{ig_i} - a_{ij} \right\}, \tag{1.15}$$

where $\mathbf{x} \equiv \mathbf{x}(k-1)$ for the evaluation of $\mathbf{x}(k-1) \in R_{\mathbf{g}}$.

In [48, Sec. 2.2], the authors have shown that it is possible to reinterpret Equation (1.15), leading to a recursive computation. First, define the partial region

$$R_{(g_1,\dots,g_k)} = \bigcap_{i=1}^{k} \bigcap_{\substack{j=1 \\ j \neq g_i}}^{p} \left\{ \mathbf{x} \in (\mathbb{R} \cup \{-\infty\})^p \mid x_j - x_{g_i} \leq a_{ig_i} - a_{ij} \right\}, \tag{1.16}$$

where $\forall k \in \{1, ..., n\}$, $(g_1, ..., g_k)$ are its partial coefficients. Note that, for $k > 1$, the partial regions (1.16) can be computed as

$$R_{(g_1,\dots,g_k)} = R_{(g_1,\dots,g_{k-1})} \cap \bigcap_{\substack{j=1 \\ j \neq g_i}}^{p} \left\{ \mathbf{x} \in (\mathbb{R} \cup \{-\infty\})^p \mid x_j - x_{g_k} \leq a_{kg_k} - a_{kj} \right\}, \tag{1.17}$$

i.e., if for some $k > 1$ the partial region $R_{(g_1,\dots,g_k)}$ is empty, then, we label these coefficients as $(g_1^{\emptyset}, ..., g_k^{\emptyset})$ and the regions $R_{(g_1,\dots,g_n)}$ are also empty if and only if $\forall i \in \{1, \dots, k\}\ g_i = g_i^{\emptyset}$.

Remark 1.9. *A procedure, with worst-case complexity $\mathcal{O}\left(p^n(np + p^3)\right)$, is described in [48, Alg. 1] and creates a PWA system from a generic MPL system*[3]*. Furthermore, it is worth to mention that a more efficient procedure (with the same worst-case complexity), using the algebraic operations $\oplus$ and $\otimes$ in the $\overline{\mathbb{R}}_{\max}$, is described in [49, Alg. 1].*

Example 1.18. *Consider the generic MPL system given by (see Equation* (1.9)*):*

$$\mathbf{y}(k) = \begin{pmatrix} 8 & 5 \\ 4 & 3 \\ 1 & 2 \end{pmatrix} \otimes \mathbf{x}(k). \tag{1.18}$$

3. In [48], it is shown that each row of A has $0 \leq p_i' \leq p$ finite elements, and then the number of regions is at most N such that $N \leq \Pi_{i=1}^{n} p_i' \leq p^n$, *i.e.*, the sparsity of A increases the performance of the procedure in the worst-case scenario.

Following Equation (1.16)*, we compute* $R_{(g_1)}$ $\forall g_1 \in \{1,2\}$*, as*

$$R_{(1)} = \left\{\mathbf{x} \in (\mathbb{R} \cup \{-\infty\})^2 \mid x_2 - x_1 \leq 3\right\}, \quad R_{(2)} = \left\{\mathbf{x} \in (\mathbb{R} \cup \{-\infty\})^2 \mid x_1 - x_2 \leq -3\right\}.$$

Now, since $k > 1$*, we use Equation* (1.17) *to compute the subsequent partial regions. Thus, to compute* $R_{(g_1,g_2)}$ $\forall (g_1,g_2) \in \{1,2\}^2 = \{(1,1),(1,2),(2,1),(2,2)\}$ *we use* $R_{(g_1)}$ *previously calculated as follows*

$$\begin{aligned}
R_{(1,1)} &= R_{(1)} \cap \{\mathbf{x} \mid x_2 - x_1 \leq 1\} = \{\mathbf{x} \mid x_2 - x_1 \leq 1\}, \\
R_{(1,2)} &= R_{(1)} \cap \{\mathbf{x} \mid x_1 - x_2 \leq -1\} = \{\mathbf{x} \mid 1 \leq x_2 - x_1 \leq 3\}, \\
R_{(2,1)} &= R_{(2)} \cap \{\mathbf{x} \mid x_2 - x_1 \leq -1\} = \emptyset, \\
R_{(2,2)} &= R_{(2)} \cap \{\mathbf{x} \mid x_1 - x_2 \leq -1\} = \{\mathbf{x} \mid x_1 - x_2 \leq -3\},
\end{aligned}$$

now,

$$\begin{aligned}
&\forall (g_1,g_2,g_3) \in \{1,2\}^3 = \\
&= \{(1,1,1),(1,1,2),(1,2,1),(1,2,2),\overbrace{(2,1,1),(2,1,2)}^{R_{(2,1)}=\emptyset},(2,2,1),(2,2,2)\},
\end{aligned}$$

and we finally compute $R_{\mathbf{g}=(g_1,g_2,g_3)}$ *using the nonempty partial regions* $R_{(g_1,g_2)}$ *previously calculated as follows*

$$\begin{aligned}
R_{(1,1,1)} &= R_{(1,1)} \cap \{\mathbf{x} \mid x_2 - x_1 \leq -1\} = \{\mathbf{x} \mid x_2 - x_1 \leq -1\}, \\
R_{(1,1,2)} &= R_{(1,1)} \cap \{\mathbf{x} \mid x_1 - x_2 \leq 1\} = \{\mathbf{x} \mid -1 \leq x_2 - x_1 \leq 1\}, \\
R_{(1,2,1)} &= R_{(1,2)} \cap \{\mathbf{x} \mid x_2 - x_1 \leq -1\} = \emptyset, \\
R_{(1,2,2)} &= R_{(1,2)} \cap \{\mathbf{x} \mid x_1 - x_2 \leq 1\} = \{\mathbf{x} \in| \ 1 \leq x_2 - x_1 \leq 3\}, \\
R_{(2,1,1)} &= R_{(2,1,2)} = \emptyset \quad \textit{because } R_{(2,1)} = \emptyset, \\
R_{(2,2,1)} &= R_{(2,2)} \cap \{\mathbf{x} \mid x_2 - x_1 \leq -1\} = \emptyset, \\
R_{(2,2,2)} &= R_{(2,2)} \cap \{\mathbf{x} \mid x_1 - x_2 \leq 1\} = \{\mathbf{x} \mid x_1 - x_2 \leq -3\},
\end{aligned}$$

Thus, according to Equation (1.14)*, the corresponding PWA system that encodes Equation*

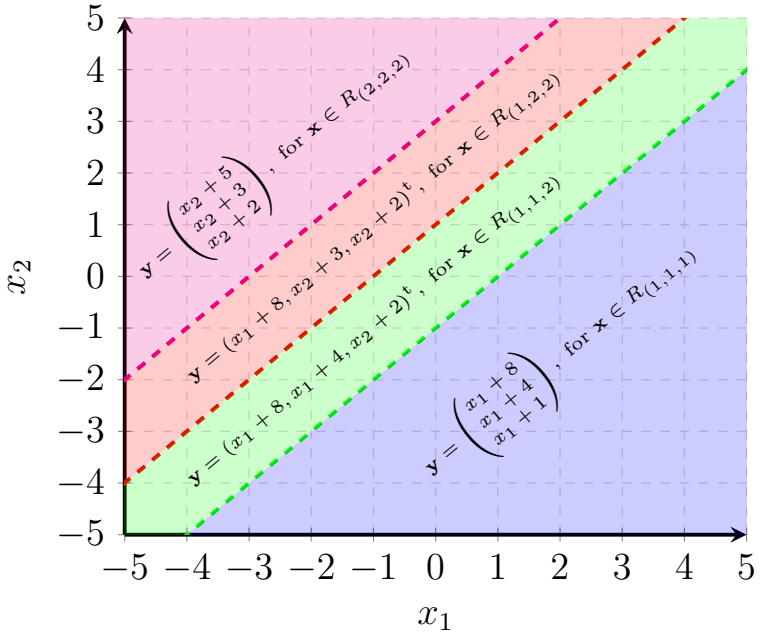


Figure 1.12 – PWA system that encodes the MPL system of Example 1.18

(1.18) *is given by:*

$$\mathbf{y}(k) = \begin{cases} \begin{pmatrix} 1 & 0 \\ 1 & 0 \\ 1 & 0 \end{pmatrix} \mathbf{x}(k-1) + \begin{pmatrix} 8 \\ 4 \\ 1 \end{pmatrix} & \textit{if} \quad \mathbf{x}(k-1) \in R_{(1,1,1)}, \\ \begin{pmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \end{pmatrix} \mathbf{x}(k-1) + \begin{pmatrix} 8 \\ 4 \\ 2 \end{pmatrix} & \textit{if} \quad \mathbf{x}(k-1) \in R_{(1,1,2)}, \\ \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{pmatrix} \mathbf{x}(k-1) + \begin{pmatrix} 8 \\ 3 \\ 2 \end{pmatrix} & \textit{if} \quad \mathbf{x}(k-1) \in R_{(1,2,2)}, \\ \begin{pmatrix} 0 & 1 \\ 0 & 1 \\ 0 & 1 \end{pmatrix} \mathbf{x}(k-1) + \begin{pmatrix} 5 \\ 3 \\ 2 \end{pmatrix} & \textit{if} \quad \mathbf{x}(k-1) \in R_{(2,2,2)}. \end{cases}$$

Figure 1.12 depicts the PWA system generated by matrix A, where, for the sake of brevity, $x_1(k-1) \equiv x_1$, $x_2(k-1) \equiv x_2$ *and* $\mathbf{y}(k) \equiv \mathbf{y}$.

1.3.4 Some notions on spectral theory of matrices in $\overline{\mathbb{R}}_{\max}$

We briefly recall the elementary facts from spectral theory of matrices in $\overline{\mathbb{R}}_{\max}^{n\times n}$. We shall use [43] and [50].

Example 1.19. *Recall Example 1.14. The timetable of the railway network simulation for the initial state* $\mathbf{x}(0) = (e, e)^{\mathrm{t}}$ *highlights that the associated model dynamics is* 2*-periodic, however it is possible to determine* $\mathbf{x}(0)$ *to ensure that the system evolves in a* 1*-periodic manner, i.e.,*

$$x_i(k) - x_i(k-1) = \lambda,\ \lambda \in \mathbb{R},\ k = 0, 1, 2, \ldots,\ i \in \{1, 2\},$$

interpreted in the classical algebra.

Example 1.19 presents a requirement that is rewritten in matrix form as follows:

$$\mathbf{x}(k) = \lambda\mathbf{x}(k-1),\ k = 0, 1, 2, \ldots \tag{1.19}$$

with $\mathbf{x} \in \overline{\mathbb{R}}_{\max}^{n}$ and $\lambda \in \overline{\mathbb{R}}_{\max}$. This yields the Max-Plus eigenproblem as stated in the sequel.

Problem 1.1. *Let*

$$\mathbf{x}(k) = A\mathbf{x}(k-1) = A^{\otimes k}\mathbf{x}(0),$$

where $\mathbf{x} \in \overline{\mathbb{R}}_{\max}^{n}$, $A \in \overline{\mathbb{R}}_{\max}^{n\times n}$ *and* $k \in \mathbb{N}$*, be an autonomous MPL system, i.e.,* $B\mathbf{u}(k) = \boldsymbol{\varepsilon}$*. We consider the problem of existence of eigenvalues* $\lambda \in \overline{\mathbb{R}}_{\max}$ *and eigenvectors* $\boldsymbol{\xi} \in \overline{\mathbb{R}}_{\max}^{n}$ *such that:*

$$A\boldsymbol{\xi} = \lambda\boldsymbol{\xi}.$$

For the evolution of the state of the system, if λ *is the eigenvalue of* A *and if* $\mathbf{x}(0)$ *is an eigenvector of* A *then*

$$\mathbf{x}(k) = \lambda\mathbf{x}(k-1) = \lambda^{\otimes k}\mathbf{x}(0),$$

or equivalently in the classical algebra as

$$\mathbf{x}(k) = \lambda + \mathbf{x}(k-1) = k\lambda + \mathbf{x}(0),$$

i.e., all state variables are incremented by λ *as* k *evolves and the steady state is reached, without going through a transient state phase. This behavior is called* 1*-periodic with a period length of* λ*.*

Theorem 1.2. *An irreducible*[4] *matrix* $A \in \overline{\mathbb{R}}_{\max}^{n\times n}$ *admits a unique eigenvalue* $\lambda \in \overline{\mathbb{R}}_{\max}$ *equal to*

$$\lambda = \bigoplus_{j=1}^{n} \Big(\mathtt{trace}(A^{\otimes j}) \Big)^{\otimes 1/j}, \tag{1.20}$$

where $\mathtt{trace}(A^{\otimes j})$ *denotes the trace of* $A^{\otimes j}$*, i.e., the* $\oplus$*-sum of its diagonal elements, and* $(a^{\otimes j})^{\otimes 1/j} = a$*.*

Remark 1.10. *Non-irreducible (or reducible) matrices may have more than one eigenvalue whereas, as given by Theorem 1.2, irreducible matrices have a unique eigenvalue but may possess several linearly independent eigenvectors, i.e., the set of all eigenvectors corresponding to the eigenvalue* λ *is the eigenspace* $E(A) = \{\mathbf{x} \in \overline{\mathbb{R}}_{\max}^{n} \mid A\mathbf{x} = \lambda\mathbf{x}\}$*.*

Remark 1.11. *The simplest method to obtain* λ *consists in evaluating Equation* (1.20)*. Nevertheless, there are several algorithms to more efficiently solve the Max-Plus eigenproblem. For more details, the reader is invited to refer to [50].*

Theorem 1.3. *Let* $A \in \overline{\mathbb{R}}_{\max}^{n\times n}$ *be an irreducible matrix with eigenvalue* $\lambda \in \overline{\mathbb{R}}_{\max}$*, then it exists* $K, c \in \mathbb{N} \setminus \{0\}$ *such that*

$$\forall k \geq K, \quad A^{k+c} = \lambda^{\otimes c} A^{\otimes k},$$

and c *is called the cyclicity of* A*.*

Theorem 1.3 means that after a transient state phase of length K, the system reaches a periodic behavior with cyclicity c, *i.e.*,

$$x_i(k+c) = x_i(k) + c\lambda, \quad i = 1, 2, \ldots, n,$$

interpreted in the classical algebra.

Example 1.20. *Recall Example 1.14. For* $\mathbf{x}(0) = (2, e)^{\mathrm{t}}$*, we obtain the following timetable*

$$\begin{pmatrix} 2 \\ e \end{pmatrix}, \begin{pmatrix} 5 \\ 5 \end{pmatrix}, \begin{pmatrix} 10 \\ 8 \end{pmatrix}, \begin{pmatrix} 13 \\ 13 \end{pmatrix}, \begin{pmatrix} 18 \\ 16 \end{pmatrix}, \begin{pmatrix} 21 \\ 21 \end{pmatrix}, \ldots,$$

4. A matrix $A \in \overline{\mathbb{R}}_{\max}^{n\times n}$ is said to be irreducible if $a_{ij} \neq \varepsilon$ $\forall i, j \in \{1, \ldots, n\}$, *i.e.*, there exists an arc from place i to place j in the corresponding TEG.

which is 2*-periodic (ciclicity* $c = 2$*) and* $\forall k \geq 1$ $\mathbf{x}(k+2) = 8 + \mathbf{x}(k)$*. On the other hand, for* $\mathbf{x}(0) = (1, e)^{\mathrm{t}}$ *we obtain the following timetable*

$$\begin{pmatrix} 1 \\ e \end{pmatrix}, \begin{pmatrix} 5 \\ 4 \end{pmatrix}, \begin{pmatrix} 9 \\ 8 \end{pmatrix}, \begin{pmatrix} 13 \\ 12 \end{pmatrix}, \begin{pmatrix} 17 \\ 16 \end{pmatrix}, \begin{pmatrix} 21 \\ 20 \end{pmatrix}, \ldots,$$

which is 1*-periodic (ciclicity* $c = 1$*) and* $\forall k \geq 0$ $\mathbf{x}(k+1) = 4 + \mathbf{x}(k)$ *with* 4 *an eigenvalue of* A *because* $\mathbf{x}(k) = 4k + \mathbf{x}(0)$*. Hence,* $\mathbf{x}(0) = (1, e)^{\mathrm{t}}$ *is an eigenvector of* A*.*

Chapter 2

Nondeterministic systems: set-membership estimation

Recall the generic MPL system Equation (1.9), precisely $\mathbf{y}(k) = A\mathbf{x}(k)$. Clearly, this system is deterministic because the output $\mathbf{y}$ is exactly predicted if one knows the values of the system matrix entries a_{ij} and the state $\mathbf{x}$. Conversely, if each a_{ij} assumes a different value at each k, then the system is called *nondeterministic* because $\mathbf{y}$ cannot be exactly predicted.

In this work, each a_{ij} is assumed to be subject to *bounded noise*, *i.e.*, disturbances and/or modelling errors, which should be taken into account in order to avoid tracking error or closed-loop instability [51, 52]. In general, these perturbations are Max-Plus multiplicative and appear as uncertainties in each a_{ij}. Stochastic MPL (SMPL) systems [53, 54, 55, 52, 22, 28] are defined as MPL systems, in which each a_{ij} is an independent-distributed random variable, and its probability density function (PDF) is supposed to be known. In this Chapter, the only probabilistic aspect of the uncertainties that are taken into account is the support of PDF, *i.e.*, the random variables a_{ij} are supported on a bounded real interval/domain (closed or open). For instance, $a_{ij} \sim \mathcal{U}(\underline{a}_{ij}, \overline{a}_{ij})$, means that a_{ij} is uniformly distributed between its lower and upper bounds $\underline{a}_{ij}$ and $\overline{a}_{ij}$, respectively.

2.1 Uncertain MPL (uMPL) systems: a bounded representation

Uncertain MPL (uMPL) systems are assumed to have each a_{ij} independently distributed within a closed specified range (bounded random variable). Thus, an autonomous uMPL system equation is given by

$$\mathbf{x}(k) = A(k)\mathbf{x}(k-1),\ \forall k\ \underline{A} \leq A(k) \leq \overline{A}, \tag{2.1}$$

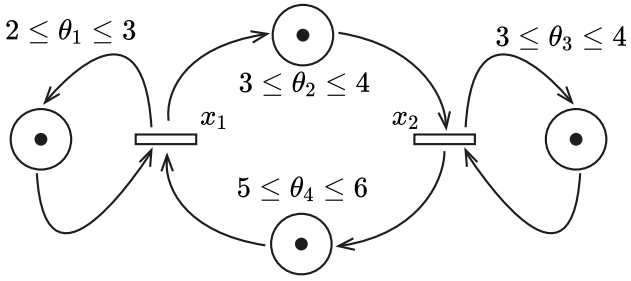


Figure 2.1 – TEG model of the uMPL system representing a railway network

where $A(k) \in \overline{\mathbb{R}}_{\max}^{n\times n}$ and $\mathbf{x} \in \overline{\mathbb{R}}_{\max}^{n}$. Then, the entries of $A(k)$ are within a closed range for each k, *i.e.*, $\underline{a}_{ij} \leq a_{ij}(k) \leq \overline{a}_{ij}$ for all $i, j \in \{1, \ldots, n\}$, *i.e.*, $\underline{A}$ and $\overline{A}$ are, respectively, the lower and upper bounds of the domain of A.

Remark 2.1. *The system in Equation 2.1 is assumed to be FIFO (first in, first out). In view of this assumption, it is always true that* $\mathbf{x}(k) \geq \mathbf{x}(k-1)$*, such that the elements of the main diagonal of* $A(k)$ *can be assumed to be greater or equal to* e *at each* k*.*

Example 2.1. *Recall Example 1.14. The travelling times were assumed to be fixed. However, in practice, these times are uncertain and are bounded within a range, as depicted in Figure 2.1. Hence, the travelling times (states) of the system are represented by:*

$$\mathbf{x}(k) \in \left\{\mathbf{x} \mid \underline{A}\mathbf{x}(k-1) \leq \mathbf{x} \leq \overline{A}\mathbf{x}(k-1)\right\},$$

where $\underline{A} = \begin{pmatrix} 2 & 5 \\ 3 & 3 \end{pmatrix}$, $\overline{A} = \begin{pmatrix} 3 & 6 \\ 4 & 4 \end{pmatrix}$ *and* $\mathbf{x}(k-1)$ *are supposed to be known at each* k*.*

Supposing that only the initial state is known, for instance $\mathbf{x}(0) = (1, e)^{\mathrm{t}}$*, then the following bounds for the timetable can be achieved:*

$$\underline{\mathbf{x}}(k) = \begin{pmatrix} 1 \\ e \end{pmatrix}, \begin{pmatrix} 5 \\ 4 \end{pmatrix}, \begin{pmatrix} 9 \\ 8 \end{pmatrix}, \begin{pmatrix} 13 \\ 12 \end{pmatrix}, \begin{pmatrix} 17 \\ 16 \end{pmatrix}, \begin{pmatrix} 21 \\ 20 \end{pmatrix}, \ldots, \qquad k = 0, 1, \ldots,$$

and

$$\overline{\mathbf{x}}(k) = \begin{pmatrix} 1 \\ e \end{pmatrix}, \begin{pmatrix} 6 \\ 5 \end{pmatrix}, \begin{pmatrix} 11 \\ 10 \end{pmatrix}, \begin{pmatrix} 16 \\ 15 \end{pmatrix}, \begin{pmatrix} 21 \\ 20 \end{pmatrix}, \begin{pmatrix} 26 \\ 25 \end{pmatrix}, \ldots, \qquad k = 0, 1, \ldots.$$

Clearly, $\underline{A}$ *and* $\overline{A}$ *share the eigenvector* $(1, e)^{\mathrm{t}}$ *and the system is* 1*-periodic with* $\underline{\lambda} = 4$ *and* $\overline{\lambda} = 5$ *(refer to Section 1.3.4), i.e.,* $\underline{\mathbf{x}}(k+1) = 4 + \underline{\mathbf{x}}(k)$ *and* $\overline{\mathbf{x}}(k+1) = 5 + \overline{\mathbf{x}}(k)$*.*

2.2 Notions on interval analysis

Interval arithmetic is presented in [56] and extended to $\overline{\mathbb{R}}_{\max}$ in [57, 58, 15, 59].

2.2.1 Basic concepts

Definition 2.1 (Closed interval over $\overline{\mathbb{R}}_{\max}$)**.** *A closed interval, denoted* $[x] = [\underline{x}, \overline{x}]$*, is a subset of* $\overline{\mathbb{R}}_{\max}$*, i.e.,* $[x] \subseteq \overline{\mathbb{R}}_{\max}$ *satisfying*

$$[x] \;=\; \left\{ x \in \overline{\mathbb{R}}_{\max} \mid \underline{x} \leq x \leq \overline{x} \right\},$$

where $\underline{x}, \overline{x} \in \overline{\mathbb{R}}_{\max}$ *(*$\underline{x} \leq \overline{x}$*) are the lower and upper bounds of* $[x]$*, respectively.*

We denote $\overline{\mathbb{IR}}_{\max}$, the set of intervals of $\overline{\mathbb{R}}_{\max}$, as $\overline{\mathbb{IR}}_{\max} = \left\{ [\underline{x}, \overline{x}] \mid \underline{x} \leq \overline{x},\ \underline{x}, \overline{x} \in \overline{\mathbb{R}}_{\max} \right\}$ such that $[x] \in \overline{\mathbb{IR}}_{\max}$ is equivalent to $[x] \subset \overline{\mathbb{R}}_{\max}$.

Definition 2.2. *The width of an interval* $[x] \in \overline{\mathbb{IR}}_{\max}$ *is defined, using conventional algebra, as* $\mathtt{w}([x]) = \overline{x} - \underline{x}$*.*

Remark 2.2 (Empty interval)**.** *From the Definitions 2.1 and 2.2 above, an interval* $[x]$ *is empty if* $\underline{x} > \overline{x}$*, i.e., if* $\mathtt{w}([x]) < 0$*.*

Definition 2.3 (Midpoint of intervals)**.** *The midpoint of an interval* $[x] \in \overline{\mathbb{IR}}_{\max}$ *is defined, using conventional algebra, as* $\mathtt{mid}([x]) = \frac{\underline{x} + \overline{x}}{2}$*.*

Remark 2.3 (Deprecated interval)**.** *If an interval* $[x]$ *has its bounds such that* $\underline{x} = \overline{x}$*, i.e.,* $\mathtt{w}([x]) = 0$*, then,* $[x]$ *is said to be deprecated. Deprecated intervals are suitable to represent elements without uncertainties, i.e., deterministic points. In this case* $[x] = [x, x] = x$*.*

The $\oplus$ and $\otimes$ operations of scalars in $\overline{\mathbb{R}}_{\max}$ are naturally extended to intervals in $\overline{\mathbb{R}}_{\max}$:

$$\begin{aligned} [x] \oplus [y] \;&=\; \{ x \oplus y \mid x \in [x] \text{ and } y \in [y] \} = \left[\underline{x} \oplus \underline{y}, \overline{x} \oplus \overline{y} \right], \\ [x] \otimes [y] \;&=\; \{ x \otimes y \mid x \in [x] \text{ and } y \in [y] \} = \left[\underline{x} \otimes \underline{y}, \overline{x} \otimes \overline{y} \right]. \end{aligned}$$

Example 2.2. *Let* $[x] = [4, 8]$ *and* $[y] = [3, 5]$*. Then*

$$\begin{aligned} [x] \oplus [y] \;&=\; [4 \oplus 3, 8 \oplus 5] = [4, 8], \\ [x] \otimes [y] \;&=\; [4 \otimes 5, 8 \otimes 3] = [9, 11]. \end{aligned}$$

Remark 2.4 (Order relation of intervals in $\overline{\mathbb{R}}_{\max}$)**.** *The* $\oplus$ *induces an order relation over intervals, similar to Equation* (1.1) *defined for scalars in* $\overline{\mathbb{R}}_{\max}$*:*

$$[x] \oplus [y] = [y] \iff [x] \leq [y] \iff \underline{x} \leq \underline{y} \text{ and } \overline{x} \leq \overline{y},$$

As an immediate consequence of the previous Remark 2.4, the following equality of intervals is defined

$$[x] = [y] \iff \underline{x} = \underline{y} \text{ and } \overline{x} = \overline{y}. \tag{2.2}$$

Below, we collect two important set-theoretic operations to properly handle intervals. First, let $[x] = [\underline{x}, \overline{x}]$ and $[y] = [\underline{y}, \overline{y}]$ be two non-disjoint intervals, then their intersection is nonempty and is defined as the set $\mathcal{Z} = \{z \in \overline{\mathbb{R}}_{\max} \mid z \in [x] \text{ and } z \in [y]\}$, which coincides with $[z] = [x] \cap [y]$, *i.e.*, $\mathcal{Z} = [z]$, where

$$[z] = [\max(\underline{x}, \underline{y}), \min(\overline{x}, \overline{y})]. \tag{2.3}$$

Secondly, let $[x] = [\underline{x}, \overline{x}]$ and $[y] = [\underline{y}, \overline{y}]$ be two intervals, then their union is defined as the set $[x] \cup [y] = \{z \in \overline{\mathbb{R}}_{\max} \mid z \in [x] \text{ or } z \in [y]\}$ which is a subset of $\overline{\mathbb{R}}_{\max}$ but not necessarily an interval [1]. Hence, in order to make the set of intervals be closed with respect to this operation, we define the *interval union*, *i.e.*, the *interval hull*[2] of $[x] \cup [y]$ as:

$$[z] = [x] \sqcup [y] = [\min(\underline{x}, \underline{y}), \max(\overline{x}, \overline{y})] \supseteq [x] \cup [y]. \tag{2.4}$$

Example 2.3. *Consider the intervals* $[x] = [e, 4]$, $[y] = [2, 5]$ *and* $[z] = [5, 7]$*. Then,*

$$\begin{aligned} [x] \cap [y] &= [\max(e, 2), \min(4, 5)] = [2, 4], \\ [x] \cap [z] &= [\max(e, 5), \min(4, 7)] = [5, 4] = \emptyset, \\ [y] \cap [z] &= [\max(2, 5), \min(5, 7)] = [5, 5]. \end{aligned}$$

Since $[x] \cap [y]$ *and* $[y] \cap [z]$ *are not empty we have that:*

$$\begin{aligned} [x] \sqcup [y] &= [\min(e, 2), \max(4, 5)] = [e, 5] = [x] \cup [y], \\ [y] \sqcup [z] &= [\min(2, 5),\ \max(5, 7)] = [2, 7] = [x] \cup [y]. \end{aligned}$$

1. The union of two intervals is an interval if and only if they are non-disjoint.
2. The interval hull of a set $\mathbb{X} \subseteq \mathbb{R}$ is the smallest interval $[\mathbb{X}]$ such that $\mathbb{X} \subseteq [\mathbb{X}]$.

Note that the intersection of $[x]$ *and* $[z]$ *is empty, and therefore* $[x] \sqcup [z]$ *is the smallest interval that encloses the disjoint union between these intervals, formally*

$$[x] \sqcup [z] = [\min(e,5), \max(4,7)] = [e,7] \supset [x] \cup [y] = \{w \mid w \in [e,4] \text{ or } w \in [5,7]\}.$$

Remark 2.5. *Any* $\diamondsuit \in \{\oplus, \otimes, \cap, \sqcup\}$ *operation over two intervals is also extended to a finite number of intervals, for instance for* $\diamondsuit = \oplus$ *we have*

$$\bigoplus_{i=1}^{n} [x_i] = \left\{ \bigoplus_{i=1}^{n} x_i \mid x_i \in [x_i] \right\} = \left[\bigoplus_{i=1}^{n} \underline{x}_i, \bigoplus_{i=1}^{n} \overline{x}_i \right]. \tag{2.5}$$

2.2.2 Matrices and Vectors

For the vector-valued case, an interval vector (also called *interval box*) is denoted as the Cartesian product of scalar intervals or more simplify according to the stacked notation

$$[\mathbf{x}] = [x_1] \times \cdots \times [x_n] \equiv ([x_1], \ldots, [x_n])^{\mathrm{t}},$$

where the set of axis-aligned interval vectors in $\overline{\mathbb{R}}_{\max}^{n}$ is denoted as $\overline{\mathbb{IR}}_{\max}^{n}$.

Remark 2.6. *The notation* $[A]$ *denotes an interval matrix, i.e., a matrix with its entries composed of intervals. Formally*

$$[A] = [\underline{A}, \overline{A}] = \left([\underline{a}, \overline{a}]_{ij} \right)_{\substack{1 \le i \le n \\ 1 \le j \le p}}$$

Remark 2.7. *Any matrix* $A \in \overline{\mathbb{R}}_{\max}^{n \times p}$ *can be represented by a deprecated interval matrix* $[A] \in \overline{\mathbb{IR}}_{\max}^{n \times p}$*, in which* $\underline{a}_{ij} = \overline{a}_{ij}$ *for all* $i \in \{1, \ldots, n\}$ *and* $j \in \{1, \ldots, p\}$*.*

The $\oplus$ and $\otimes$ are extended to interval matrices as follows: if $[A]$, $[B]$ and $[C]$ are, respectively, $(n \times p)$, $(n \times p)$ and $(p \times q)$-dimensional interval matrices, then

$$\begin{aligned}
([A] \oplus [B])_{ij} &= [a_{ij}] \oplus [b_{ij}] = [\underline{a}_{ij} \oplus \underline{b}_{ij}, \overline{a}_{ij} \oplus \overline{b}_{ij}], \\
([A] \otimes [C])_{ij} &= \bigoplus_{k=1}^{p} [a_{ik}] \otimes [c_{kj}] = \bigoplus_{k=1}^{p} \left\{ [\underline{a}_{ik} \otimes \underline{c}_{kj}, \overline{a}_{ik} \otimes \overline{c}_{kj}] \right\} \\
&= \left[\bigoplus_{k=1}^{p} \left\{ \underline{a}_{ik} \otimes \underline{c}_{kj} \right\}, \bigoplus_{k=1}^{p} \{ \overline{a}_{ik} \otimes \overline{c}_{kj} \} \right].
\end{aligned}$$

Thus, the k-th power and the Kleene star operation of a matrix of intervals are given by:

$$[A]^{\otimes k} = [\underline{A}^{\otimes k}, \overline{A}^{\otimes k}] \text{ and } [A]^{\star} = [\underline{A}^{\star}, \overline{A}^{\star}].$$

Furthermore, the $\cap$ (or $\sqcup$) operation of two interval matrices can be computed as the element-wise $\cap$ (or $\sqcup$) operation of the corresponding entries.

Example 2.4. *Given* $[A] = \begin{pmatrix} [2,7] & [4,5] \\ [4,6] & [2,6] \end{pmatrix}$ *and* $[\mathbf{v}] = \begin{pmatrix} e \\ [e,1] \end{pmatrix}$ *then* $[\mathbf{x}] = [A][\mathbf{v}]$ *is given by*

$$\begin{pmatrix} [x_1] \\ [x_2] \end{pmatrix} = \begin{pmatrix} [(2 \otimes e) \oplus (4 \otimes e), (7 \otimes e) \oplus (5 \otimes 1)] \\ [(4 \otimes e) \oplus (2 \otimes e), (6 \otimes e) \oplus (6 \otimes 1)] \end{pmatrix} = \begin{pmatrix} [4,7] \\ [4,7] \end{pmatrix}.$$

2.2.3 Mappings and inverse mappings of points

Let $f : \overline{\mathbb{R}}^n_{\max} \to \overline{\mathbb{R}}^m_{\max}, \mathbf{x} \mapsto A\mathbf{x}$ be a linear mapping and

$$F(\mathbf{x}) = [A]\mathbf{x} = \{\mathbf{y} \in \overline{\mathbb{R}}^m_{\max} \mid \underline{A}\mathbf{x} \leq \mathbf{y} \leq \overline{A}\mathbf{x}\}, \tag{2.7}$$

be a set, where $A \in [A] = [\underline{A}, \overline{A}] \in \overline{\mathbb{IR}}^{m \times n}_{\max}$ and $\mathbf{x} \in \overline{\mathbb{R}}^n_{\max}$.

In other words, $\forall A \in [A]$ we have that $f(\mathbf{x}) = A\mathbf{x} \in F(\mathbf{x}) = [A]\mathbf{x}$.

Remark 2.8. *In Lemma C.1 of Appendix C, it is demonstrated that given* $\mathbf{x}$, $\mathbf{y}$ *and* $[A] = [\underline{A}, \overline{A}]$ *such that* $\underline{A}\mathbf{x} \leq \mathbf{y} \leq \overline{A}\mathbf{x}$, *then* $\exists A \in [A]$ *such that* $A\mathbf{x} = \mathbf{y}$.

Example 2.5. *Given* $\mathbf{x} = (e, e)^{\mathrm{t}}$ *and* $y \in \begin{pmatrix} [1,4] & [2,3] \end{pmatrix} \mathbf{x}$ *then* $y \in [2,4]$. *Suppose that we take* $\mathtt{mid}([2,4]) = 3$ *to be a point* $\tilde{y}$ *in* $[2,4]$. *Thus, thanks to Remark 2.8, it exists a matrix* A *in* $[\underline{A}, \overline{A}] = \begin{pmatrix} [1,4] & [2,3] \end{pmatrix}$ *such that* $3 = A \otimes (e, e)^{\mathrm{t}}$. *This matrix is given by Lemma C.1 of Appendix C and is equal to* $A = (\tilde{y} \not\backslash \mathbf{x}) \wedge \overline{A} = \begin{pmatrix} 3 & 3 \end{pmatrix} \wedge \begin{pmatrix} 4 & 3 \end{pmatrix} = \begin{pmatrix} 3 & 3 \end{pmatrix}$, *which is promptly verified by* $\otimes$*-multiplying it by* $\mathbf{x}$.

Let $f : \overline{\mathbb{R}}^n_{\max} \to \overline{\mathbb{R}}^m_{\max}, \mathbf{x} \mapsto A\mathbf{x}$ be a linear mapping and

$$F^{-1}(\mathbf{y}) = \{\mathbf{x} \in \overline{\mathbb{R}}^n_{\max} \mid \exists A \in [A], A\mathbf{x} = \mathbf{y}\}$$

be a set, where $A \in [A] = [\underline{A}, \overline{A}] \in \overline{\mathbb{IR}}^{m \times n}_{\max}$, $\mathbf{x} \in \overline{\mathbb{R}}^m_{\max}$ and $\mathbf{y} \in \overline{\mathbb{R}}^m_{\max}$. Then, $\mathbf{x} \in F^{-1}(\mathbf{y}) \iff \underline{A}\mathbf{x} \leq \mathbf{y} \leq \overline{A}\mathbf{x}$, and thus

$$F^{-1}(\mathbf{y}) = \{\mathbf{x} \in \overline{\mathbb{R}}^n_{\max} \mid \underline{A}\mathbf{x} \leq \mathbf{y} \leq \overline{A}\mathbf{x}\}. \tag{2.8}$$

In other words, $F^{-1}(\mathbf{y})$ is the set in $\mathbf{x}$ of all $\mathbf{y} = A\mathbf{x}$, given $\mathbf{y}$ and with $A \in [A]$.

In view of $\underline{A}\mathbf{x} \leq \mathbf{y} \leq \overline{A}\mathbf{x}$, given $\mathbf{y}$, there is a unique greatest $\mathbf{x}$ given by Equation (1.3) such that $\underline{A}\mathbf{x} \leq \mathbf{y}$ holds, but not a unique least $\mathbf{x}$ such that $\overline{A}\mathbf{x} \geq \mathbf{y}$ also holds. Hence, we split $F^{-1}(\mathbf{y})$ into two sets L and U such that $F^{-1}(\mathbf{y}) = L \cap U$. In details,

$$L = \{\mathbf{x} \in \overline{\mathbb{R}}^n_{\max} \mid \mathbf{y} \leq \overline{A}\mathbf{x}\} \qquad (2.9) \qquad U = \{\mathbf{x} \in \overline{\mathbb{R}}^n_{\max} \mid \underline{A}\mathbf{x} \leq \mathbf{y}\}, \qquad (2.10)$$

Proposition 2.1 ([29])**.** *The sets L and U of Equations* (2.9) *and* (2.10)*, respectively, are equivalent to:*

$$L = \bigcup_{\mathbf{g} \in G} \bigcap_{i=1}^{m} set^i_{g_i} \text{ and } U = \{\mathbf{x} \in \overline{\mathbb{R}}^n_{\max} \mid \mathbf{x} \leq \underline{A} \backslash \mathbf{y}\},$$

where

$$G = \{1, \ldots, n\}^m, \mathbf{g} \in G, \mathbf{g} = (g_1, g_2, \ldots, g_m), g_i \in \{1, \ldots, n\}$$

and

$$set^i_j = \left(\bigcap_{k=1}^{j-1} \{\mathbf{x} \in \overline{\mathbb{R}}^n_{\max} \mid x_k < \overline{a}_{ik} \backslash y_i\} \right) \cap \{\mathbf{x} \in \overline{\mathbb{R}}^n_{\max} \mid x_j \geq \overline{a}_{ij} \backslash y_i\}.$$

Proof. The proof for U is straightforwardly done by using residuation theory, while the proof for L follows the proof given in Section B.2 of Appendix B. □

Thus,

$$F^{-1}(\mathbf{y}) = L \cap U = \left(\bigcup_{\mathbf{g} \in G} \bigcap_{i=1}^{m} set^i_{g_i} \right) \cap U = \bigcup_{\mathbf{g} \in G} SET^{\mathbf{g}}, \text{ with } SET^{\mathbf{g}} = \bigcap_{i=1}^{m} set^i_{g_i} \cap U, \quad (2.11)$$

which is a set composed of at most n^m pairwise disjoint sets (see Lemma B.1 and Remark B.2 of Appendix B for more details).

Example 2.6. *Consider the first event of an autonomous uMPL system, described by* $\mathbf{x}(1) \in [A]\mathbf{x}(0)$*, where* $[A] = \begin{pmatrix} [1,4] & [2,3] \\ [1,2] & [e,4] \end{pmatrix}$ *and* $\mathbf{x}(0) = (e,e)^{\mathrm{t}}$*. Then, the set of all* $\mathbf{x}(1) \in [A]\mathbf{x}(0)$ *given by Equation* (2.7) *is calculated below*

$$\mathbf{x}(1) \in \begin{pmatrix} [1,4] & [2,3] \\ [1,2] & [e,4] \end{pmatrix} \begin{pmatrix} e \\ e \end{pmatrix} = [\mathbf{x}](1) = \begin{pmatrix} [2,4] \\ [1,4] \end{pmatrix},$$

and depicted in Figure 2.2. If one takes a point in $[\mathbf{x}](1)$*, for instance* $\tilde{\mathbf{x}}(1) = \mathtt{mid}([\mathbf{x}](1)) = (3, 2.5)^{\mathrm{t}}$*, then thanks to Equation* (2.11) *it is possible to compute the set of all* $\mathbf{x}(0)$ *that* $\otimes$*-multiplied by a matrix* $A \in [A]$ *yield* $\mathbf{x}(1)$*. This computation is performed by expressing sets* L *and* U *according to Proposition 2.1. Then,*

$$U = \{\mathbf{x} \mid \mathbf{x} \leq \underline{A} \backslash \tilde{\mathbf{x}}(1)\}, \text{ where } \underline{A} \backslash \tilde{\mathbf{x}}(1) = \begin{pmatrix} 1 & 2 \\ 1 & e \end{pmatrix} \backslash \begin{pmatrix} 3 \\ 2.5 \end{pmatrix} = \begin{pmatrix} 1.5 \\ 1 \end{pmatrix}.$$

Calculating L *is done by defining first* set_j^i *for all* $(i, j) \in \{1, 2\} \times \{1, 2\}$*, as follows:*

$$\begin{aligned}
set_1^1 &= \{\mathbf{x} \mid x_1 \geq \overline{a}_{11} \backslash \tilde{x}_1 = 4 \not/ 3 = -1\} = \{\mathbf{x} \mid x_1 \geq -1\} \\
set_2^1 &= \{\mathbf{x} \mid x_2 \geq \overline{a}_{12} \backslash \tilde{x}_1 = 3 \not/ 3 = e\} \cap \{\mathbf{x} \mid x_1 < \overline{a}_{11} \backslash \tilde{x}_1 = 4 \not/ 3 = -1\} \\
&= \{\mathbf{x} \mid x_1 < -1, x_2 \geq e\} \\
set_1^2 &= \{\mathbf{x} \mid x_1 \geq \overline{a}_{21} \backslash \tilde{x}_2 = 2 \not/ 2.5 = 0.5\} = \{\mathbf{x} \mid x_1 \geq 0.5\} \\
set_2^2 &= \{\mathbf{x} \mid x_2 \geq \overline{a}_{22} \backslash \tilde{x}_2 = 4 \not/ 2.5 = -1.5\} \cap \{\mathbf{x} \mid x_1 < \overline{a}_{21} \backslash \tilde{x}_2 = 2 \not/ 2.5 = 0.5\} \\
&= \{\mathbf{x} \mid x_1 < 0.5, x_2 \geq -1.5\}.
\end{aligned}$$

Now, we compute $L \cap U$ *as* $\bigcup_{\mathbf{g} \in G} SET^{\mathbf{g}} = \bigcap_{i=1}^{n} set_{g_i} \cap U$*, with* $G = \{(1, 1), (1, 2), (2, 1), (2, 2)\}$ *as shown below:*

$$\begin{aligned}
SET^{(1,1)} &= set_{g_1=1}^1 \cap set_{g_2=1}^2 \cap U \\
&= \{\mathbf{x} \mid x_1 \geq -1\} \cap \{\mathbf{x} \mid x_1 \geq 0.5\} \cap \{\mathbf{x} \mid x_1 \leq 1.5, x_2 \leq 1\} \\
&= \{\mathbf{x} \mid 0.5 \leq x_1 \leq 1.5, x_2 \leq 1\}, \\
SET^{(1,2)} &= set_{g_1=1}^1 \cap set_{g_2=2}^2 \cap U \\
&= \{\mathbf{x} \mid x_1 \geq -1\} \cap \{\mathbf{x} \mid x_1 < 0.5, x_2 \geq -1.5\} \cap \{\mathbf{x} \mid x_1 \leq 1.5, x_2 \leq 1\} \\
&= \{\mathbf{x} \mid -1 \leq x_1 < 0.5, -1.5 \leq x_2 \leq 1\} \\
SET^{(2,1)} &= set_{g_1=2}^1 \cap set_{g_2=1}^2 \cap U \\
&= \{\mathbf{x} \mid x_1 < -1, x_2 \geq e\} \cap \{\mathbf{x} \mid x_1 \geq 0.5\} \cap \{\mathbf{x} \mid x_1 \leq 1.5, x_2 \leq 1\} \\
&= \emptyset, \\
SET^{(2,2)} &= set_{g_1=2}^1 \cap set_{g_2=2}^2 \cap U \\
&= \{\mathbf{x} \mid x_1 < -1, x_2 \geq e\} \cap \{\mathbf{x} \mid x_1 < 0.5, x_2 \geq -1.5\} \cap \{\mathbf{x} \mid x_1 \leq 1.5, x_2 \leq 1\} \\
&= \{\mathbf{x} \mid x_1 < -1, e \leq x_2 \leq 1\},
\end{aligned}$$

such that

$$L \cap U = \{\mathbf{x} \mid 0.5 \leq x_1 \leq 1.5, \varepsilon < x_2 \leq 1\} \cup \{\mathbf{x} \mid -1 \leq x_1 < 0.5, -1.5 \leq x_2 \leq 1\} \cup \{\mathbf{x} \mid \varepsilon < x_1 < -1, e \leq x_2 \leq 1\},$$

as depicted in Figure 2.2.

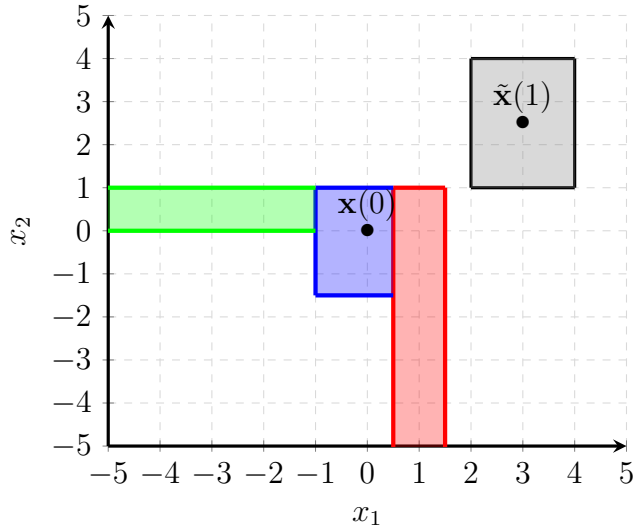


Figure 2.2 – $[\mathbf{x}](1)$ as the set of all $\mathbf{x}(1) \in [A]\mathbf{x}(0)$ (in gray), $\tilde{\mathbf{x}}(1)$ as the middle point of $[\mathbf{x}](1)$ (black dot) and set $L \cap U$ (union of red, green and blue boxes) of Example 2.6

2.3 Models of uMPL systems

2.3.1 Implicit and explicit forms of uMPL systems

The following equation defines the model of an uMPL system with implicit part

$$\begin{cases} \mathbf{x}(k) = A_0(k)\mathbf{x}(k) \oplus A_1(k)\mathbf{x}(k-1) \oplus B(k)\mathbf{u}(k), & \text{(2.12a)} \\ \mathbf{z}(k) = C(k)\mathbf{x}(k), & \text{(2.12b)} \end{cases}$$

where $A_0(k) \in [A_0] \in \overline{\mathbb{IR}}_{\max}^{n\times n}$, $A_1(k) \in [A_1] \in \overline{\mathbb{IR}}_{\max}^{n\times n}$, $B(k) \in [B] \in \overline{\mathbb{IR}}_{\max}^{n\times p}$, $C(k) \in [C] \in \overline{\mathbb{IR}}_{\max}^{q\times n}$ and $\mathbf{x}$, $\mathbf{u}$, $\mathbf{z}$ are vectors of appropriate dimension (refer to Equation (1.10)).

Since $[A_0]$ is considered to be written in strictly lower triangular form (the associated

TEG is live) then

$$\begin{cases} x_i(k) = \left(\bigoplus_{l=1}^{i-1} a_{il}^0(k) x_l(k)\right) \oplus y_i(k-1), \ i = 1, 2, \ldots, n & \text{(2.13a)} \\ \mathbf{z}(k) = C(k)\mathbf{x}(k), & \text{(2.13b)} \end{cases}$$

where

$$\mathbf{y}(k-1) = A_1(k)\mathbf{x}(k-1) \oplus B(k)\mathbf{u}(k)$$

and $a_{ij}^0(k) \in [a_{ij}^0(k)] = [\underline{a}_{ij}^0(k), \overline{a}_{ij}^0(k)]$ for all $i, j \in \{1, \ldots, n\}$.

Alternatively, using Kleene star operation, we may address Equation (2.12) as the following explicit form

$$\begin{cases} \mathbf{x}(k) = A(k)\mathbf{x}(k-1) \oplus \tilde{B}(k)\mathbf{u}(k), & \text{(2.14a)} \\ \mathbf{z}(k) = C(k)\mathbf{x}(k), & \text{(2.14b)} \end{cases}$$

where $A(k) \in [A_0^\star][A_1]$, $\tilde{B}(k) \in [A_0^\star][B]$ and $[A_0^\star] = [\underline{A}_0^\star, \overline{A}_0^\star]$.

Remark 2.9. *The entries of matrices $A(k) = A_0^\star(k)A_1(k)$ and $\tilde{B}(k) = A_0^\star(k)B(k)$ in Equation* (2.14a) *are coupled, even though the entries of $A_0(k), A_1(k)$ and $B(k)$ being considered mutually independent (refer to Equation* (2.1)*). Thus, Equation* (2.14a) *is a conservative (over-approximation) representation for the real system given by Equation* (2.12a)*, and it may lead to unfeasible states (false-positive states).*

Example 2.7. *Consider the autonomous uMPL system ($B(k)\mathbf{u}(k) = \boldsymbol{\varepsilon}$) with implicit part as given by Equation* (2.12a)*, with*

$$A_0(k) \in [A_0] = \begin{pmatrix} \varepsilon & \varepsilon \\ [1,2] & \varepsilon \end{pmatrix}, A_1(k) \in [A_1] = \begin{pmatrix} [4,6] & [3,5] \\ [3,7] & [4,5] \end{pmatrix}.$$

It is possible to obtain $\mathbf{x}(1) = (5, 9)^{\mathrm{t}}$ using Equation (2.14a)*. In details,*

$$A(1) = \begin{pmatrix} 4 & 5 \\ 8 & 7 \end{pmatrix} \in [A] = [A_0^\star][A_1] = \begin{pmatrix} [4,6] & [3,5] \\ [5,8] & [4,7] \end{pmatrix}$$

and $\mathbf{x}(0) = (1, e)^{\mathrm{t}}$ yield $\mathbf{x}(1) = A(1)\mathbf{x}(0) = (5, 9)^{\mathrm{t}}$. However, this state is unfeasible if one

considers Equation (2.13)*:*

$$x_1(1) \in [4,6]x_1(0) \oplus [3,5]x_2(0) \overset{\mathbf{x}(0)=(1,e)^{\mathrm{t}}}{\Longrightarrow} 5 \in [5,7],$$
$$x_2(2) \in [1,2]x_1(1) \oplus ([3,7]x_1(0) \oplus [4,5]x_2(0)) \overset{x_1(1)=5,\mathbf{x}(0)=(1,e)^{\mathrm{t}}}{\Longrightarrow} 9 \notin [5,8].$$

Hence, this Chapter will be focused on uMPL systems modeled exclusively by Equation (2.14) with

$$A(k) \in [A_1], \tilde{B}(k) \in [B] \text{ and } C(k) \in [C],$$

i.e., systems without implicit parts.

2.3.2 Partitioned uMPL systems: PWA-uMPL systems

In this Section, we shall use the results of [35, 27].

Equation (2.14a) can be rewritten as the following uMPL system equation (refer to Equation (1.9)):

$$\mathbf{y}(k) = A(k)\mathbf{x}(k-1), \tag{2.15}$$

where $A \in [A] \in \overline{\mathbb{IR}}_{\max}^{n\times p}$, $\mathbf{y} \in \overline{\mathbb{R}}_{\max}^{n}$ and $\mathbf{x} \in \overline{\mathbb{R}}_{\max}^{p}$. Thus,

$$y_i(k) = \max_{1\le j\le p}(a_{ij}(k) + x_j(k-1)), a_{ij}(k) \in [a_{ij}(k)] = [\underline{a}_{ij}, \overline{a}_{ij}], \quad i = 1,2,\ldots,n, \tag{2.16}$$

with $y_i, x_j \in \mathbb{R}\cup\{-\infty\}$ and $[a_{ij}(k)] \subset \mathbb{R}\cup\{-\infty\}$, represents the dynamics in the conventional algebra. Furthermore,

$$\begin{aligned} y_i(k) \in [y_i(k)] &= \left[\max_{1\le j\le p}(\underline{a}_{ij} + x_j(k-1)), \max_{1\le j\le p}(\overline{a}_{ij} + x_j(k-1))\right], \\ &\iff y_i(k) \in \left\{y_i \mid \max_{1\le j\le p}(\underline{a}_{ij} + x_j(k-1)) \le y_i \le \max_{1\le j\le p}(\overline{a}_{ij} + x_j(k-1))\right\} \\ &\iff y_i(k) \in \left\{y_i \mid \max_{1\le j\le p}(\underline{a}_{ij} + x_j(k-1)) \le y_i\right\} \cap \left\{y_i \mid y_i \le \max_{1\le j\le p}(\overline{a}_{ij} + x_j(k-1))\right\} \end{aligned}$$

Regarding the lower bound we have

$$\max_{1\le j\le p}(\underline{a}_{ij} + x_j(k-1)) \le y_i \iff \begin{cases} \underline{a}_{i1} + x_1(k-1) \le y_i \\ \vdots \\ \underline{a}_{ip} + x_p(k-1) \le y_i \end{cases},$$

and thus

$$\left\{y_i \mid \max_{1\leq j\leq p}(\underline{a}_{ij}+x_j(k-1))\leq y_i\right\}\equiv\bigcap_{j=1}^{p}\{y_i\mid \underline{a}_{ij}+x_j(k-1)\leq y_i\}.$$

On the other hand, regarding the upper bound we have (refer to Section B.2 of Appendix B for more details):

$$\left\{y_i \mid y_i\leq \max_{1\leq j\leq p}(\overline{a}_{ij}+x_j(k-1))\right\}\equiv\bigcup_{j=1}^{p}\{y_i\mid y_i\leq \overline{a}_{ij}+x_j(k-1)\}.$$

Hence,

$$y_i(k)\in\bigcap_{j=1}^{p}\{y_i\mid \underline{a}_{ij}+x_j(k-1)\leq y_i\}\cap\bigcup_{j=1}^{p}\{y_i\mid y_i\leq \overline{a}_{ij}+x_j(k-1)\},$$

i.e., $\exists g_i\in\{1,\dots,p\}$ for all $i\in\{1,\dots,n\}$, such that

$$\begin{aligned} y_i(k)&\in\bigcap_{j=1}^{p}\{y_i\mid \underline{a}_{ij}+x_j(k-1)\leq y_i\}\cap\{y_i\mid y_i\leq \overline{a}_{ig_i}+x_{g_i}(k-1)\} \\ &\iff y_i(k)\in\left[\max_{1\leq j\leq p}(\underline{a}_{ij}+x_j(k-1)),\overline{a}_{ig_i}+x_{g_i}(k-1)\right], \end{aligned} \tag{2.17}$$

with $\overline{a}_{ig_i}+x_{g_i}(k-1)\geq\overline{a}_{ij}+x_j(k-1)\iff x_j(k-1)-x_{g_i}\leq\overline{a}_{ig_i}-\overline{a}_{ij},\forall j$.

Given $\mathbf{g}=(g_1,g_2,\dots,g_n)\in\{1,\dots,p\}^n$, the affine dynamics of Equation (2.15) are therefore given by

$$\mathbf{x}(k-1)\in R_{\mathbf{g}}\implies y_i(k)\in\left[\max_{1\leq j\leq p}(\underline{a}_{ij}+x_j(k-1)),\overline{a}_{ig_i}+x_{g_i}(k-1)\right],\ i=1,2,\dots,n, \tag{2.18}$$

where

$$R_{\mathbf{g}}=\bigcap_{i=1}^{n}\bigcap_{\substack{j=1\\ j\neq g_i}}^{p}\left\{\mathbf{x}\in(\mathbb{R}\cup\{-\infty\})^p\mid x_j-x_{g_i}\leq\overline{a}_{ig_i}-\overline{a}_{ij}\right\}, \tag{2.19}$$

with the notation $\mathbf{x}\equiv\mathbf{x}(k-1)$ for the evaluation of $\mathbf{x}(k-1)\in R_{\mathbf{g}}$.

Remark 2.10. *The procedure for generating a PWA system from an uMPL system has the same complexity as the one presented in Remark 1.9. As it was considered for MPL systems (refer to Equation* (1.16)*) it is possible to compute Equation* (2.19) *in a recursive way.*

Example 2.8. *Consider the following uMPL system*

$$\mathbf{x}(k) = A(k)\mathbf{x}(k-1),\ A(k) \in \begin{pmatrix} [4,6] & [3,5] \\ [3,7] & [4,5] \end{pmatrix}. \tag{2.20}$$

In order to compute its partitioned representation, first, compute each $R_{\mathbf{g}}$*, with* $\mathbf{g} \in \{1,2\}^2 = \{(1,1),(1,2),(2,1),(2,2)\}$*, as given by:*

$$\begin{aligned}
R_{(1,1)} &= \{\mathbf{x} \in (\mathbb{R} \cup \{-\infty\})^2 \mid x_2 - x_1 \leq 1\} \cap \{\mathbf{x} \in (\mathbb{R} \cup \{-\infty\})^2 \mid x_2 - x_1 \leq 2\} \\
&= \{\mathbf{x} \in (\mathbb{R} \cup \{-\infty\})^2 \mid x_2 - x_1 \leq 1\} \\
R_{(1,2)} &= \{\mathbf{x} \in (\mathbb{R} \cup \{-\infty\})^2 \mid x_2 - x_1 \leq 1\} \cap \{\mathbf{x} \in (\mathbb{R} \cup \{-\infty\})^2 \mid x_1 - x_2 \leq -2\} \\
&= \emptyset, \\
R_{(2,1)} &= \{\mathbf{x} \in (\mathbb{R} \cup \{-\infty\})^2 \mid x_1 - x_2 \leq -1\} \cap \{\mathbf{x} \in (\mathbb{R} \cup \{-\infty\})^2 \mid x_2 - x_1 \leq 2\} \\
&= \{\mathbf{x} \in (\mathbb{R} \cup \{-\infty\})^2 \mid 1 \leq x_2 - x_1 \leq 2\}, \\
R_{(2,2)} &= \{\mathbf{x} \in (\mathbb{R} \cup \{-\infty\})^2 \mid x_1 - x_2 \leq -1\} \cap \{\mathbf{x} \in (\mathbb{R} \cup \{-\infty\})^2 \mid x_1 - x_2 \leq -2\} \\
&= \{\mathbf{x} \in (\mathbb{R} \cup \{-\infty\})^2 \mid x_1 - x_2 \leq -2\}.
\end{aligned}$$

Thus, from Equation (2.18)*, the PWA system that represents the corresponding uMPL system, is given by*

$$\mathbf{x}(k) \in \begin{cases} \begin{pmatrix} [\max(4 + x_1(k-1), 3 + x_2(k-1)), 6 + x_1(k-1)]) \\ [\max(3 + x_1(k-1), 4 + x_2(k-1)), 7 + x_1(k-1)]) \end{pmatrix} & \text{if } \mathbf{x}(k-1) \in R_{(1,1)} \\ \begin{pmatrix} [\max(4 + x_1(k-1), 3 + x_2(k-1)), 5 + x_2(k-1)]) \\ [\max(3 + x_1(k-1), 4 + x_2(k-1)), 7 + x_1(k-1)]) \end{pmatrix} & \text{if } \mathbf{x}(k-1) \in R_{(2,1)} \\ \begin{pmatrix} [\max(4 + x_1(k-1), 3 + x_2(k-1)), 5 + x_2(k-1)]) \\ [\max(3 + x_1(k-1), 4 + x_2(k-1)), 5 + x_1(k-1)]) \end{pmatrix} & \text{if } \mathbf{x}(k-1) \in R_{(2,2)} \end{cases}$$

Figure 2.3 depicts the PWA system generated by the interval matrix $[A]$*. Notation:* $X_1 \equiv \max(4 + x_1(k-1), 3 + x_2(k-1))$*,* $X_2 \equiv \max(3 + x_1(k-1), 4 + x_2(k-1))$*,* $\mathbf{x}(k) \equiv \mathbf{x}'$ *and* $\mathbf{x}(k-1) \equiv \mathbf{x}$*.*

Equation (2.17) is simply expressed $\forall i \in \{1, \ldots, n\}$ as the following subset of $(\mathbb{R} \cup \{-\infty\})^n$:

$$\mathbf{y}(k) \in \bigcap_{i=1}^{n} \bigcap_{j=1}^{p} \{\mathbf{y} \mid \underline{a}_{ij} + x_j(k-1) \leq y_i\} \cap \bigcap_{i=1}^{n} \{\mathbf{y} \mid y_i \leq \overline{a}_{ig_i} + x_{g_i}(k-1)\}. \tag{2.21}$$

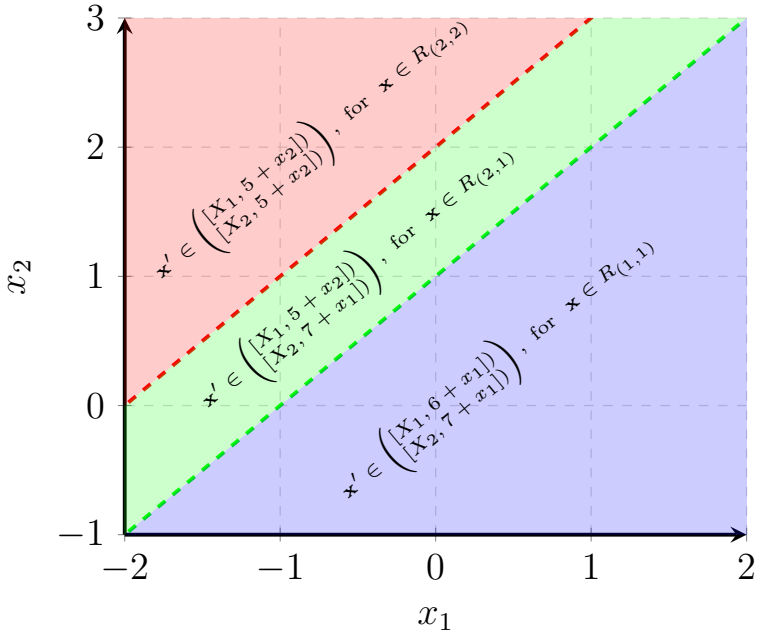


Figure 2.3 – PWA system that encodes the uMPL system of Example 2.8

2.3.3 Difference-Bound Matrices (DBM) representation of PWA-uMPL systems

Difference-Bound Matrices (DBM)

Difference-Bound Matrices (DBM) are used to represent regions (or zones as initially mentioned in [60]) in the state-space defined by a finite number of linear inequalities (or *difference constraints*) in an efficient and concise way. For instance, in [61, 62], DBM were used on timed automata. It is also important to mention the use of DBM for the analysis of computer programs, as presented in [63, 64].

Definition 2.4 (Difference-Bound Matrices (DBM))**.** *Let $X = \{x_1, x_2, \ldots, x_n\}$ be a set of states in a given state-space and $\tilde{X} = X \cup \{x_0\}$ be its augmented form, where $x_0 = 0$. A DBM is a square matrix D with its rows and columns labeled by $\tilde{X}$ (i.e., of dimension $(n+1) \times (n+1)$) such that its entries d_{ij} are in $\mathbb{R} \cup \{+\infty\}$ and represent for all $x_i, x_j \in \tilde{X}$ the difference constraints[3] (linear inequalities) $x_i - x_j \leq d_{ij}$ with $i, j \in \{1, \ldots, n+1\}$. Moreover: the artificial state x_0 allows expressing $x_i \leq d_{i0}$ and $-x_j \leq d_{0j}$; if an entry d_{ij} is equal to $+\infty$, then $x_i - x_j$ is not upper bounded, i.e., $x_i - x_j < +\infty$; the difference*

3. DBM allow expressing strict and non-strict difference constraints, for instance $x_i - x_j < c$ and $x_i - x_j \leq c$, respectively. However, in general, we are interested on non-strict constraints.

constraint $x_j - x_i \geq c$ with $c \in \mathbb{R} \cup \{+\infty\}$ is reinterpreted as $x_i - x_j \leq d_{ij} = -c$; the main diagonal elements d_{ii}, are equal to $+0$ since $x_i - x_i = 0 \leq d_{ij} = +0$.

In the sequel, we use the notation $\mathcal{R}(D)$ to represent the set of solutions of the system of inequalities (intersection of inequalities) encoded by a DBM D.

Remark 2.11 (Canonical form and emptiness verification)**.** *In general, the same region can be represented by different DBM, for instance D_1 and D_2. Nevertheless, each DBM admits a single representation in its canonical form, denoted for instance for D_1 and D_2 as $\mathtt{cf}(D_1) = \mathtt{cf}(D_2)$ (see [61, Th.2]). The algorithm of Floyd-Warshall, with complexity $\mathcal{O}(n^3)$, is responsible for computing $\mathtt{cf}(D_1)$ and $\mathtt{cf}(D_2)$. Moreover, as presented in [35, Alg. 2.1], this algorithm allows signalizing if the corresponding DBM represents an empty region, and it stops if this is the case.*

Example 2.9. *Let $X_1 = \{(x_1, x_2)^{\mathrm{t}} \in \mathbb{R}^2 \mid x_1 \geq 3, x_2 \leq 5, x_1 - x_2 \leq 4\}$ and $X_2 = \{(x_1, x_2)^{\mathrm{t}} \in \mathbb{R}^2 \mid x_1 \geq 11, x_2 \leq 5, x_1 - x_2 \leq 4\}$ be two subsets of $\mathbb{R}^2$, representing two different zones. These sets are represented by the DBM D_1 and D_2 of $(\mathbb{R} \cup \{+\infty\})^3$, as given below:*

$$D_1 = \begin{array}{c} \\ x_0 \\ x_1 \\ x_2 \end{array} \begin{array}{c} \begin{array}{ccc} x_0 & x_1 & x_2 \end{array} \\ \begin{pmatrix} +0 & -3 & +\infty \\ +\infty & +0 & 4 \\ 5 & +\infty & +0 \end{pmatrix} \end{array} \iff \begin{cases} x_1 \geq 3 & \iff & x_0 - x_1 \leq -3 \\ x_2 \leq 5 & \iff & x_2 - x_0 \leq 5 \\ & & x_1 - x_2 \leq 4 \end{cases}$$

and

$$D_2 = \begin{array}{c} \\ x_0 \\ x_1 \\ x_2 \end{array} \begin{array}{c} \begin{array}{ccc} x_0 & x_1 & x_2 \end{array} \\ \begin{pmatrix} +0 & -11 & +\infty \\ +\infty & +0 & 4 \\ 5 & +\infty & +0 \end{pmatrix} \end{array} \iff \begin{cases} x_1 \geq 11 & \iff & x_0 - x_1 \leq -11 \\ x_2 \leq 5 & \iff & x_2 - x_0 \leq 5 \\ & & x_1 - x_2 \leq 4 \end{cases}$$

Computing their canonical forms using [35, Alg. 2.1], we obtain[4]*.*

$$D_1 = \begin{array}{c} \\ x_0 \\ x_1 \\ x_2 \end{array} \begin{array}{c} \begin{array}{ccc} x_0 & x_1 & x_2 \end{array} \\ \begin{pmatrix} +0 & -3 & 1 \\ 9 & +0 & 4 \\ 5 & 2 & +0 \end{pmatrix} \end{array} \implies \mathcal{R}(D_1) \neq \emptyset \text{ and } \mathcal{R}(D_2) = \emptyset.$$

4. $\mathcal{R}(D_2) = \emptyset$ is also easily verified by considering that $x_2 \leq 5 \implies x_1 \leq 9$ when intersecting with $x_1 - x_2 \leq 4$ and thus there is no x_1 that satisfies both inequalities $x_1 \leq 9$ and $x_1 \geq 13$.

Figure 2.4 below depicts these regions.

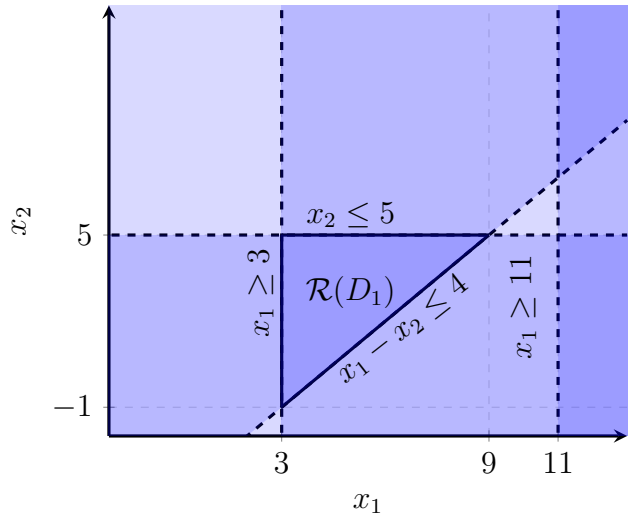


Figure 2.4 – Representation of region $\mathcal{R}(D_1)$ and inequality $x_1 \geq 11$ of Example 2.9

Remark 2.12 (Orthogonal projection of DBM)**.** *Let D be a DBM, in its canonical form, that represents the region $\mathcal{R}(D)$ generated by the variables $x_i \in X = \{x_0, x_1, x_2, \ldots, x_n\}$ and their differences $x_i - x_j, \forall i \neq j$ and $i, j \in \{0, \ldots, n\}$. Given $\widehat{X} = \{x_{i_1}, \ldots, x_{i_p}\} \subseteq X$, then the orthogonal projection of D onto $\widehat{X}$, denoted $D\lceil_{\widehat{X}}$, is obtained by deleting the rows and columns corresponding to $X \backslash \widehat{X}$, i.e., the variables $x_i \notin \widehat{X}$ (see [61, Sec. 4.1]). Furthermore, the region represented by $D\lceil_{\widehat{X}}$ is given by*

$$\mathcal{R}(D\lceil_{\widehat{X}}) = \{(x_{i_1}, \ldots, x_{i_p})^{\mathrm{t}} \in \mathbb{R}^p \mid (x_0, x_1, x_2, \ldots, x_n)^{\mathrm{t}} \in \mathcal{R}(D)\}.$$

Remark 2.13 (Cartesian product of DBM)**.** *Let D^X et D^Y be two DBM of dimensions $(p+1) \times (p+1)$ and $(n+1) \times (n+1)$, respectively. Then, the Cartesian product of their regions $\mathcal{R}(D^X)$ and $\mathcal{R}(D^Y)$ is computed as*

$$\mathcal{R}(D^X) \times \mathcal{R}(D^Y) = \{(\mathbf{x}^{\mathrm{t}}, \mathbf{y}^{\mathrm{t}})^{\mathrm{t}} \in \mathbb{R}^{p+n} \mid \mathbf{x} \in \mathcal{R}(D^X), \mathbf{y} \in \mathcal{R}(D^Y)\}.$$

Alternatively, it is possible to compute $D^{X \times Y} = D^X \times D^Y$ as an augmented DBM of dimension $(p+n+1) \times (p+n+1)$ such that

$$\mathcal{R}\left(D^{X \times Y}\right) = \mathcal{R}\left(D^X\right) \times \mathcal{R}\left(D^Y\right).$$

Computationally, calculating $D^{X \times Y}$ is done with complexity $\mathcal{O}(\max(p^2, n^2))$.

Example 2.10. *Let*

$$D^{\mathbb{R}^2} = \begin{array}{c} \\ x_0 \\ x_1 \\ x_2 \end{array} \begin{array}{c} \begin{array}{ccc} x_0 & x_1 & x_2 \end{array} \\ \begin{pmatrix} +0 & +\infty & +\infty \\ +\infty & +0 & +\infty \\ +\infty & +\infty & +0 \end{pmatrix} \end{array} \text{ and } D^Y = \begin{array}{c} \\ x_0 \\ y_1 \end{array} \begin{array}{c} \begin{array}{cc} x_0 & y_1 \end{array} \\ \begin{pmatrix} +0 & 0 \\ 1 & +0 \end{pmatrix} \end{array}$$

be the DBM that represents $\mathbb{R}^2$ *and an arbitrary DBM, respectively. Then,* $D^{\mathbb{R}^2 \times X}$ *is given by*

$$D^{\mathbb{R}^2 \times X} = D^{\mathbb{R}^2} \times D^X = \begin{array}{c} \\ x_0 \\ x_1 \\ x_2 \\ y_1 \end{array} \begin{array}{c} \begin{array}{cccc} x_0 & x_1 & x_2 & y_1 \end{array} \\ \begin{pmatrix} +0 & +\infty & +\infty & 0 \\ +\infty & +0 & +\infty & +\infty \\ +\infty & +\infty & +0 & +\infty \\ 1 & +\infty & +\infty & +0 \end{pmatrix} \end{array}.$$

Clearly, $D^{\mathbb{R}^2 \times X}$ *is in its canonical form, i.e.,* $\mathtt{cf}(D^{\mathbb{R}^2 \times X}) = D^{\mathbb{R}^2 \times X}$*. As proof of correctness, the orthogonal projection of* $\mathtt{cf}(D^{\mathbb{R}^2 \times X})$ *over* $\mathbb{R}^2$ *is* $\mathtt{cf}(D^{\mathbb{R}^2 \times X})\lceil_{\mathbb{R}^2} = D^{\mathbb{R}^2}$*, and the orthogonal projection of* $\mathtt{cf}(D^{\mathbb{R}^2 \times X})$ *over* X *is* $\mathtt{cf}(D^{\mathbb{R}^2 \times X})\lceil_X = D^X$*.*

Remark 2.14 (Representing interval vectors as DBM)**.** *An interval vector* $[\mathbf{x}] \in \overline{\mathbb{IR}}^n_{\max}$ *is such that* $x_0 - x_i \leq -\underline{x}_i$ *and* $x_i - x_0 \leq \overline{x}_i$ *for* $i = 1, 2, \ldots, n$ *in the conventional algebra. Hence,*

$$D^{[\mathbf{x}]} = \begin{array}{c} \\ x_0 \\ x_1 \\ \vdots \\ x_n \end{array} \begin{array}{c} \begin{array}{c:ccc} x_0 & x_1 & \cdots & x_n \end{array} \\ \begin{pmatrix} +0 & -\underline{x}_1 & \cdots & -\underline{x}_n \\ \hdashline \overline{x}_1 & & & \\ \vdots & & Id_n & \\ \overline{x}_n & & & \end{pmatrix} \end{array},$$

where Id_n *is a matrix of dimension* $n \times n$ *with* $+0$ *on the main diagonal and* $+\infty$ *elsewhere, is the DBM representation of* $[\mathbf{x}]$*.*

Remark 2.15 (Partial order on DBM)**.** *Let* D *and* E *be two DBM of same dimension* $n \times n$*. The partial order on* D *and* E *is defined as:* $D \preceq E \iff d_{ij} \leq e_{ij}$ *where* $d_{ij} \in D$ *and* $e_{ij} \in E$ *for all* $i, j \in \{1, \ldots, n\}$*. Then,*

$$D \preceq E \text{ implies that } \mathcal{R}(D) \subseteq \mathcal{R}(E).$$

Remark 2.16 (Set-theoretic operations of DBM)**.** *Let* D *and* E *be two DBM of same*

dimension $n \times n$ such that their region representations $\mathcal{R}(D)$ and $\mathcal{R}(E)$ are non-empty. Thus, $\min(D, E)$*, computed element-wise, is such that*

$$\mathcal{R}(D) \cap \mathcal{R}(E) = \mathcal{R}(\min(D, E)).$$

Similarly, $\max(D, E)$*, also computed element-wise, is such that*

$$\mathcal{R}(D) \cup \mathcal{R}(E) \subseteq \mathcal{R}(\max(D, E)).$$

Example 2.11. *Given two interval vectors*

$$[\mathbf{x}] = \begin{pmatrix} [1,4] \\ [1,3] \end{pmatrix} \text{ and } [\mathbf{y}] = \begin{pmatrix} [2,5] \\ [2,4] \end{pmatrix},$$

then their DBM representations (see Remark 2.14) is given below

$$D^{[\mathbf{x}]} = \begin{array}{c} \\ x_0 \\ x_1 \\ x_2 \end{array} \begin{array}{c} \begin{array}{ccc} x_0 & x_1 & x_2 \end{array} \\ \begin{pmatrix} +0 & -1 & -1 \\ 4 & +0 & +\infty \\ 3 & +\infty & +0 \end{pmatrix} \end{array} \text{ and } D^{[\mathbf{y}]} = \begin{array}{c} \\ x_0 \\ x_1 \\ x_2 \end{array} \begin{array}{c} \begin{array}{ccc} x_0 & x_1 & x_2 \end{array} \\ \begin{pmatrix} +0 & -2 & -2 \\ 5 & +0 & +\infty \\ 4 & +\infty & +0 \end{pmatrix} \end{array}.$$

Computing

$$[\mathbf{x}] \cap [\mathbf{y}] = \begin{pmatrix} [2,4] \\ [2,3] \end{pmatrix}$$

can alternatively be done by considering $\min(D^{[\mathbf{x}]}, D^{[\mathbf{y}]})$*:*

$$\min\left(\begin{pmatrix} 0 & -1 & -1 \\ 4 & 0 & \infty \\ 3 & \infty & 0 \end{pmatrix}, \begin{pmatrix} 0 & -2 & -2 \\ 5 & 0 & \infty \\ 4 & \infty & 0 \end{pmatrix}\right) = \begin{pmatrix} \min(0,0) & \min(-1,-2) & \min(-1,-2) \\ \min(4,5) & \min(0,0) & \min(\infty,\infty) \\ \min(3,4) & \min(\infty,\infty) & \min(0,0) \end{pmatrix}$$
$$= \begin{pmatrix} 0 & -2 & -2 \\ 4 & 0 & \infty \\ 3 & \infty & 0 \end{pmatrix}.$$

Similarly

$$[\mathbf{x}] \sqcup [\mathbf{y}] = \begin{pmatrix} [1,5] \\ [1,4] \end{pmatrix}$$

can alternatively be done by considering $\max(D^{[\mathbf{x}]}, D^{[\mathbf{y}]})$*:*

$$\max\left(\begin{pmatrix} 0 & -1 & -1 \\ 4 & 0 & 3 \\ 3 & 2 & 0 \end{pmatrix}, \begin{pmatrix} 0 & -2 & -2 \\ 5 & 0 & \infty \\ 4 & \infty & 0 \end{pmatrix}\right) = \begin{pmatrix} \max(0,0) & \max(-1,-2) & \max(-1,-2) \\ \max(4,5) & \max(0,0) & \max(\infty,\infty) \\ \max(3,4) & \max(\infty,\infty) & \max(0,0) \end{pmatrix}$$
$$= \begin{pmatrix} 0 & -1 & -1 \\ 5 & 0 & \infty \\ 4 & \infty & 0 \end{pmatrix}.$$

Furthermore,

$$\min(D^{[\mathbf{x}]}, D^{[\mathbf{y}]}) \preceq \max(D^{[\mathbf{x}]}, D^{[\mathbf{y}]}) \implies \mathcal{R}(\min(D^{[\mathbf{x}]}, D^{[\mathbf{y}]})) \subseteq \mathcal{R}(\max(D^{[\mathbf{x}]}, D^{[\mathbf{y}]})).$$

Figure 2.5 shows these computations.

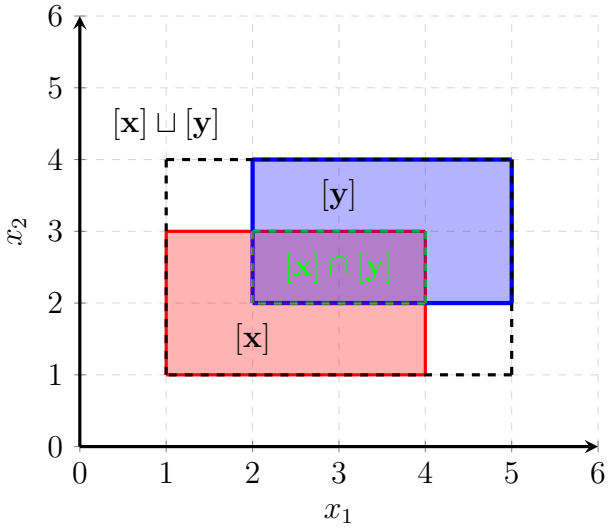


Figure 2.5 – Representations of $[\mathbf{x}]$ (or $\mathcal{R}(D^{[\mathbf{x}]})$), $[\mathbf{y}]$ (or $\mathcal{R}(D^{[\mathbf{y}]})$), $[\mathbf{x}] \cap [\mathbf{y}]$ (or $\mathcal{R}(\min(D^{[\mathbf{x}]}, D^{[\mathbf{y}]})$ and $[\mathbf{x}] \sqcup [\mathbf{y}]$ (or $\mathcal{R}(\max(D^{[\mathbf{x}]}, D^{[\mathbf{y}]})$) of Example 2.11

PWA-uMPL systems represented by DBM

In [48], the authors proposed a method to represent partitioned MPL systems as DBM (refer to Section 1.3.3). Algorithmically, it receives as input the matrix of the MPL system, and it returns as output the collection of DBM, in which each DBM represents an affine dynamics and its corresponding active region. Thus, the main advantage of using DBM to

represent partitioned MPL systems is the existence of powerful and efficient manipulation procedures (Cartesian product, orthogonal projection, canonical form, inclusion test and set-theoretic operations). Nevertheless, it is worth to be mentioned that this representation substantially increases the dimension of the problem, *i.e.*, instead of having a compact state-space equation, we obtain finitely many DBM.

Straightforwardly, in [35, 27], this representation was extended to uMPL systems and for in-depth information, please refer to these works. Briefly, recall that Equation (2.15) is alternatively expressed by Equation (2.21) and is valid for $\mathbf{x}(k-1)$ in an active region $R_{\mathbf{g}}$ given by Equation (2.19). After trivial algebraic manipulations in the conventional algebra, $\mathbf{y}(k) \in L \cap U^{\mathbf{g}}$, where

$$L = \overbrace{\bigcap_{i=1}^{n} \bigcap_{j=1}^{p} \{\mathbf{y} \in (\mathbb{R} \cup \{-\infty\})^n \mid x_j - y_i \leq -\underline{a}_{ij}\}}^{\text{"lower" dynamics}}, \tag{2.22}$$

$$U^{\mathbf{g}} = \underbrace{\bigcap_{i=1}^{n} \{\mathbf{y} \in (\mathbb{R} \cup \{-\infty\})^n \mid y_i - x_{g_i} \leq \overline{a}_{ig_i}\}}_{\text{"upper" dynamics}}, \tag{2.23}$$

for $\mathbf{x} \equiv \mathbf{x}(k-1) \in R_{\mathbf{g}}$, which is clearly the intersection of $n \times (p+1)$ difference constraints. Hence, it is possible to write all these inequalities using a single DBM, as shown below

$$D^{\mathbf{g}} = \begin{array}{c} \\ x_0 \\ y_1(k) \\ \vdots \\ y_n(k) \\ x_1(k-1) \\ \vdots \\ x_p(k-1) \end{array} \begin{array}{c} \begin{array}{ccc} x_0 & y_1(k) \cdots y_n(k) & x_1(k-1) \cdots x_p(k-1) \end{array} \\ \left(\begin{array}{c|c|c} +0 & +\infty \cdots +\infty & +\infty \cdots +\infty \\ \hline +\infty & & \\ \vdots & \mathrm{Id}_n & U^{\mathbf{g}} \text{ from Eq (2.23)} \\ +\infty & & \\ \hline +\infty & & \\ \vdots & L \text{ from Eq (2.22)} & R_{\mathbf{g}} \text{ from Eq (2.19)} \\ +\infty & & \end{array}\right) \end{array}, \tag{2.24}$$

which is a matrix of dimension $(n+p+1) \times (n+p+1)$.

Definition 2.5 (Subsystem of a PWA-uMPL system)**.** *A subsystem of a PWA-uMPL system is represented by the DBM $D^{\mathbf{g}}$ with $\mathbf{g} \in \{1, \ldots, p\}^n$.*

Example 2.12. *Let us recall Example 2.8. It was obtained* 3 *regions, precisely $R_{(1,1)}$,*

$R_{(2,1)}$ *and* $R_{(2,2)}$. *For*

$$R_{(1,1)} = \left\{ \mathbf{x}(k-1) \mid \underbrace{x_2(k-1) - x_1(k-1) \leq 1}_{d_{54}^{(1,1)}} \right\},$$

the active dynamics is given by Equations (2.22) *and* (2.23)*, as follows*

$$\mathbf{x}(k) \in L \cap \left\{ \mathbf{x}(k) \mid \underbrace{x_1(k) - x_1(k-1) \leq 6}_{d_{24}^{(1,1)}} \right\} \cap \left\{ \mathbf{x}(k) \mid \underbrace{x_2(k) - x_1(k-1) \leq 7}_{d_{34}^{(1,1)}} \right\},$$

where

$$L = \left\{ \mathbf{x}(k) \mid \underbrace{x_1(k-1) - x_1(k) \leq -4}_{d_{42}^{(1,1)}} \right\} \cap \left\{ \mathbf{x}(k) \mid \underbrace{x_2(k-1) - x_1(k) \leq -3}_{d_{52}^{(1,1)}} \right\}$$
$$\cap \left\{ \mathbf{x}(k) \mid \underbrace{x_1(k-1) - x_2(k) \leq -3}_{d_{43}^{(1,1)}} \right\} \cap \left\{ \mathbf{x}(k) \mid \underbrace{x_2(k-1) - x_2(k) \leq -4}_{d_{53}^{(1,1)}} \right\},$$

and thus

$$D^{(1,1)} = \begin{array}{r} \\ x_0 \\ x_1(k) \\ x_2(k) \\ x_1(k-1) \\ x_2(k-1) \end{array} \begin{array}{c} \begin{array}{ccccc} x_0 & x_1(k) & x_2(k) & x_1(k-1) & x_2(k-1) \end{array} \\ \begin{pmatrix} +0 & +\infty & +\infty & +\infty & +\infty \\ +\infty & +0 & +\infty & 6 & +\infty \\ +\infty & +\infty & +0 & 7 & +\infty \\ +\infty & -4 & -3 & +0 & +\infty \\ +\infty & -3 & -4 & 1 & +0 \end{pmatrix} \end{array}.$$

Now, for

$$R_{(2,1)} = \{\mathbf{x}(k-1) \mid 1 \leq x_2(k-1) - x_1(k-1) \leq 2\}$$
$$= \left\{\mathbf{x}(k-1) \mid \underbrace{x_1(k-1) - x_2(k-1) \leq -1}_{d_{45}^{(2,1)}}\right\} \cap \left\{\mathbf{x}(k-1) \mid \underbrace{x_2(k-1) - x_1(k-1) \leq 2}_{d_{54}^{(2,1)}}\right\},$$

we have the following active dynamics

$$\mathbf{x}(k) \in L$$
$$\cap \left\{\mathbf{x}(k) \mid \underbrace{x_1(k) - x_2(k-1) \leq 5}_{d_{25}^{(2,1)}}\right\} \cap \left\{\mathbf{x}(k) \mid \underbrace{x_2(k) - x_1(k-1) \leq 7}_{d_{34}^{(2,1)}}\right\},$$

and thus

$$D^{(2,1)} = \begin{array}{r} \\ x_0 \\ x_1(k) \\ x_2(k) \\ x_1(k-1) \\ x_2(k-1) \end{array} \begin{array}{c} \begin{array}{ccccc} x_0 & x_1(k) & x_2(k) & x_1(k-1) & x_2(k-1) \end{array} \\ \begin{pmatrix} +0 & +\infty & +\infty & +\infty & +\infty \\ +\infty & +0 & +\infty & +\infty & 5 \\ +\infty & +\infty & +0 & 7 & +\infty \\ +\infty & -4 & -3 & +0 & -1 \\ +\infty & -3 & -4 & 2 & +0 \end{pmatrix} \end{array}.$$

Following the same procedure, for $R_{(2,2)} = \{\mathbf{x}(k-1) \mid x_1(k-1) - x_2(k-1) \leq -2\}$ *we obtain*

$$D^{(2,2)} = \begin{array}{r} \\ x_0 \\ x_1(k) \\ x_2(k) \\ x_1(k-1) \\ x_2(k-1) \end{array} \begin{array}{c} \begin{array}{ccccc} x_0 & x_1(k) & x_2(k) & x_1(k-1) & x_2(k-1) \end{array} \\ \begin{pmatrix} +0 & +\infty & +\infty & +\infty & +\infty \\ +\infty & +0 & +\infty & +\infty & 5 \\ +\infty & +\infty & +0 & +\infty & 5 \\ +\infty & -4 & -3 & +0 & -2 \\ +\infty & -3 & -4 & +\infty & +0 \end{pmatrix} \end{array}.$$

2.4 The set-estimation problem: *conditional* reachability analysis (RA) techniques

In stochastic estimation approach, the uncertain state vector $\mathbf{x}$ is characterized by probability density functions (PDF). Conversely, in set-membership estimation (or set-estimation for short) approach, $\mathbf{x}$ is characterized by a set X such that $\mathbf{x} \in X$. Both approaches are related by the fact that X represents the support of the PDF that represent $\mathbf{x}$.

Handling sets is easier than computing PDF (more details will be presented in the next Chapter), however PDF provide more accuracy rather than simply obtaining their support. Hence, set-membership estimation is considered to be *conservative*.

In this section, we consider the model of Equation (2.14) with $A_0(k) = \mathcal{E}$ for all $1 \leq k \leq N$. We also consider that the measurement vector $\mathbf{z}(k)$ is included in a subset of the measurement-space, precisely $\mathbf{z}(k) \in Z(k) \subset \overline{\mathbb{R}}^q_{\max}$.

Definition 2.6 (Conditional reach set)**.** *Let $X_0 \subset \overline{\mathbb{R}}^n_{\max}$ be a set of initial conditions such that $\mathbf{x}(0) \in X_0$. Given that $\mathbf{z}(k)$, from event 1 up to N, is included in $Z(k)$, then $X_{k|k}$ corresponds to the conditional reach set from $X_{k-1|k-1}$ (assuming that $X_{0|0} = X_0$) obtained by Equation* (2.14a)*, which leads to $Z(k)$ via Equation* (2.14b)*.*

The interpretation of Definition 2.6 is the following two-fold procedure:

$$\mathbf{x}(k) \in \underbrace{X_{k|k}}_{\text{posterior estimation}} = \underbrace{X_{k|k-1}}_{\text{prior estimation}} \bigcap \underbrace{\tilde{X}_{k|k}}_{\text{set-inversion problem}}, \ \forall 1 \leq k \leq N, \tag{2.25}$$

where

- for the non-autonomous case: $X_{k|k-1} = \{(A\ \tilde{B})\mathbf{y}(k-1) \mid \mathbf{y}(k-1) \in X_{k-1|k-1} \times U_k, A \in [A_1], \tilde{B} \in [B]\},\ U_k = \{\mathbf{u}(k)\}$;
- for the autonomous case: $X_{k|k-1} = \{A\mathbf{x}(k-1) \mid \mathbf{x}(k-1) \in X_{k-1|k-1}, A \in [A_1]\}$;
- for $\mathbf{z}(k) \in Z(k)$, where $Z(k)$ is not a singleton: $\tilde{X}_{k|k} = \{\mathbf{x}(k) \mid \exists C \in [C], C\mathbf{x}(k) \in Z(k)\}$;
- for $\mathbf{z}(k) \in Z(k)$, where $Z(k) = \{\mathbf{z}(k)\}$, *i.e.*, $Z(k)$ is a singleton: $\tilde{X}_{k|k} = \{\mathbf{x}(k) \mid \exists C \in [C], C\mathbf{x}(k) = \mathbf{z}(k)\}$.

Characterizing $\tilde{X}_{k|k}$ corresponds to a set-inversion problem [38] in the general case and to the inverse mapping of a point problem if $Z(k)$ is a singleton (refer to Section 2.2.3). Hence, $\tilde{X}_{k|k}$ is the set of all states $\mathbf{x}(k)$ that may lead to $\mathbf{z}(k) \in Z(k)$.

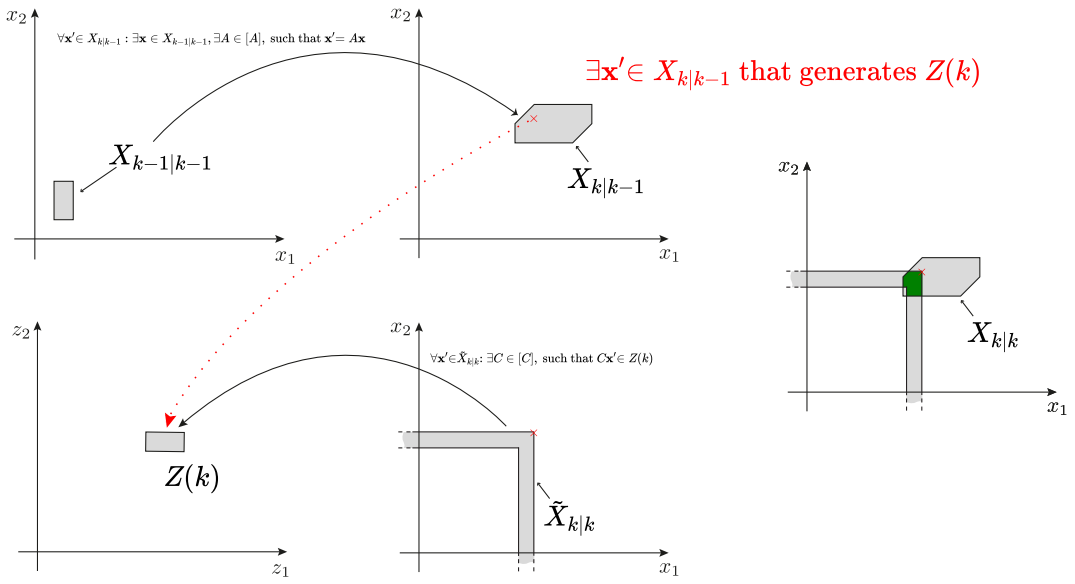


Figure 2.6 – Graphical scheme of the two-fold computation of $X_{k|k}$: from $X_{k-1|k-1}$ we compute $X_{k|k-1}$; it exists a $\mathbf{x}(k) \in X_{k|k-1}$ that generates $Z(k)$ which in turn generates $\tilde{X}_{k|k}$; finally, $X_{k|k} = X_{k|k-1} \cap \tilde{X}_{k|k}$

Remark 2.17. *If $X_{k-1|k-1} = \{\mathbf{x}(k-1)\}$ for any k then $X_{k|k-1}$ is easily computed by Equation (2.7) as $\mathbf{x}(k) \in X_{k|k-1} = \{\mathbf{x} \mid \underline{A}_1\mathbf{x}(k-1) \leq \mathbf{x} \leq \overline{A}_1\mathbf{x}(k-1)\}$, for autonomous systems. The same is considered for non-autonomous systems by considering that $\underline{A}_1\mathbf{x}(k-1) \oplus \underline{\tilde{B}}\mathbf{u}(k) \leq \mathbf{x}(k) \leq \overline{A}_1\mathbf{x}(k-1) \oplus \overline{\tilde{B}}\mathbf{u}(k)$. This computation (matrix-vector multiplication) has bilinear complexity $\mathcal{O}(n(n+p))$ for non-autonomous systems and quadratic complexity $\mathcal{O}(n^2)$ for autonomous systems.*

The Figure 2.6 depicts an example of the steps necessary to compute $X_{k|k}$.

2.4.1 DBM-RA

In [35, 27, 29], the computation of $X_{k|k}$ is addressed using forward and backward reachability analysis (RA), as an extension of the results found in [65, 66, 48].

Remark 2.18. *The image and the inverse image of a set represented by a DBM with respect to a subsystem (cf. Definition 2.5) of a PWA-uMPL system is a set that can be represented by a DBM (see [35, Prop. 5.1]).*

Remark 2.19. *The image and the inverse image of a set represented by the union of finitely many DBM with respect to a PWA-uMPL system can be represented by the union of finitely many DBM (see [35, Cor. 5.2]).*

In the sequel, we consider Remarks 2.18 and 2.19 to compute $X_{k|k-1}$ and $\tilde{X}_{k|k}$.

Prior estimation

First, we aim at computing

$$X_{k|k-1} = \{F\mathbf{y} \mid \mathbf{y} \in X_{k-1|k-1} \times U_k, F \in [F]\}, \tag{2.26}$$

with $[F] = ([A]\ [\tilde{B}])$ and $U_k = \{\mathbf{u}(k)\}$.

It is assumed that the set $X_{k-1|k-1}$ is represented by the union/collection of a finite number of DBM. Furthermore, $\mathbf{x}(k) = F(k)\mathbf{y}(k)$ with $F(k) \in [F]$ and $\mathbf{y}(k) = (\mathbf{x}^{\mathrm{t}}(k-1), \mathbf{u}^{\mathrm{t}}(k))^{\mathrm{t}}$ is also represented by the union of a finite number of DBM (see Section 2.3.3).

Image of a DBM with respect to a subsystem of a PWA-uMPL system

Let $D^{\mathbf{g}} \in (\mathbb{R} \cup \{+\infty\})^{n+m+1 \times n+m+1}$ be the DBM representation of a subsystem of a PWA-uMPL system. Given $D^{Y_k} = D^{X_{k-1}} \times D^{U_k} \in (\mathbb{R} \cup \{+\infty\})^{m+1 \times m+1}$ representing the region $\mathcal{R}\left(D^{Y_k}\right) = Y_k = X_{k-1} \times U_k \subseteq (\mathbb{R} \cup \{+\infty\})^m$, then its image with respect to $D^{\mathbf{g}}$ is given below:

1. Compute the Cartesian product between $D^{\mathbb{R}^n}$ and D^{Y_k}, mathematically $D^{\mathbb{R}^n \times Y_k} = D^{\mathbb{R}^n} \times D^{Y_k}$ (see Remark 2.13);
2. Compute the intersection of $D^{\mathbb{R}^n \times Y_k}$ with $D^{\mathbf{g}}$, mathematically $D^{\widehat{X}_k} = \min(D^{\mathbb{R}^n \times Y_k}, D^{\mathbf{g}})$ (see Remark 2.16);
3. Compute the canonical form of $D^{\widehat{X}_k}$, mathematically, $D_{cf}^{\widehat{X}_k} = \mathtt{cf}(D^{\widehat{X}_k})$ (see Remark 2.11);
4. Compute the orthogonal projection of $D_{cf}^{\widehat{X}_k}$ over $\mathbf{x}(k)$, mathematically $D^{X_k} = D_{cf}^{\widehat{X}_k} \lceil_{\{x_1(k),\ldots,x_n(k)\}}$ (see Remark 2.12);
5. D^{X_k} is the image of D^{Y_k}.

The complexity of this procedure depends essentially on the canonical form computation (step 3), which is cubic in the dimension of $D^{\widehat{X}_k}$. The term m refers to an autonomous ($m = n$) or non-autonomous system ($m = n + p$).

If $X_{k-1|k-1} \times U_k$ is represented as the union of N DBM then its image is computed with worst-case complexity $\mathcal{O}(Nm^n(n+m)^3)$ where $m = n + p$ if $[\tilde{B}] \neq \mathcal{E}$ in $[F] = ([A_1]\ [\tilde{B}])$ (*i.e.*, $U_k \neq \emptyset$) or $m = n$ otherwise. The term m^n refers to the worst-case number of subsystems of the PWA-uMPL system, and the term $(m+n)^3$ refers to the complexity of computing the image of a single DBM with respect to a subsystem of the PWA-uMPL system. Assuming $m = n$, then the worst-case becomes $\mathcal{O}(Nn^{n+3})$, *i.e.*, exponential in dimension n, which is clearly cumbersome if we want to handle large systems.

Example 2.13. *Recall Example 2.8, which is represented by* 3 *subsystems of the PWA-uMPL system that represents a uMPL system. Precisely, in Example 2.12 we have computed* $\mathbb{D} = \{D^{(1,1)}, D^{(2,1)}, D^{(2,2)}\}$ *as its subsystems. Given*

$$X_{0|0} = \{\mathbf{x}(0) \mid 0 \leq x_1(0) \leq 1, 1 \leq x_2(0) \leq 3\}, \quad D^{X_{0|0}} = \begin{array}{c} \\ x_0 \\ x_1(0) \\ x_2(0) \end{array} \begin{array}{c} \begin{array}{ccc} x_0 & x_1(0) & x_2(0) \end{array} \\ \begin{pmatrix} +0 & 0 & -1 \\ 1 & +0 & +\infty \\ 3 & +\infty & +0 \end{pmatrix} \end{array}$$

which is an interval vector in $\overline{\mathbb{IR}}^n_{\max}$*, represented by the corresponding DBM* $D^{X_{0|0}}$*, then* $X_{1|0}$ *is the collection of the images of* $X_{0|0}$ *with respect to each* $D \in \mathbb{D}$*.*

Following the procedure, we compute

$$D^{\mathbb{R}^2 \times X_{0|0}} = D^{\mathbb{R}^2} \times D^{X_{0|0}} = \begin{array}{c} \\ x_0 \\ x_1(1) \\ x_2(1) \\ x_1(0) \\ x_2(0) \end{array} \begin{array}{c} \begin{array}{ccccc} x_0 & x_1(1) & x_2(1) & x_1(0) & x_2(0) \end{array} \\ \begin{pmatrix} +0 & \infty & +\infty & 0 & -1 \\ +\infty & +0 & +\infty & +\infty & +\infty \\ +\infty & \infty & +0 & +\infty & +\infty \\ 1 & +\infty & +\infty & +0 & +\infty \\ 3 & +\infty & +\infty & +\infty & +0 \end{pmatrix} \end{array}.$$

Then, for $D^{(1,1)}$ *we compute*

$$D^{\widehat{X}_{1|\mathbf{g}=(1,1)}} = \min(D^{\mathbb{R}^2 \times X_{0|0}}, D^{(1,1)})$$

$$= \min\left(\begin{pmatrix} 0 & \infty & \infty & 0 & -1 \\ \infty & 0 & \infty & \infty & \infty \\ \infty & \infty & 0 & \infty & \infty \\ 1 & \infty & \infty & 0 & \infty \\ 3 & \infty & \infty & \infty & 0 \end{pmatrix}, \begin{pmatrix} 0 & \infty & \infty & \infty & \infty \\ \infty & 0 & \infty & 6 & \infty \\ \infty & \infty & 0 & 7 & \infty \\ \infty & -4 & -3 & 0 & \infty \\ \infty & -3 & -4 & 1 & +0 \end{pmatrix}\right) = \begin{pmatrix} 0 & \infty & \infty & 0 & -1 \\ \infty & 0 & \infty & 6 & \infty \\ \infty & \infty & 0 & 7 & \infty \\ 1 & -4 & -3 & 0 & \infty \\ 3 & -3 & -4 & 1 & 0 \end{pmatrix},$$

and thus, the canonical form of $D^{\widehat{X}_{1|\mathbf{g}=(1,1)}}$

$$D_{cf}^{\widehat{X}_{1|\mathbf{g}=(1,1)}} = \mathtt{cf}(D^{\widehat{X}_{1|\mathbf{g}=(1,1)}}) = \begin{array}{c} \\ x_0 \\ x_1(1) \\ x_2(1) \\ x_1(0) \\ x_2(0) \end{array} \begin{array}{c} \begin{array}{ccccc} x_0 & x_1(1) & x_2(1) & x_1(0) & x_2(0) \end{array} \\ \begin{pmatrix} +0 & -4 & -5 & 0 & -1 \\ 7 & +0 & 2 & 6 & 6 \\ 8 & 3 & +0 & 7 & 7 \\ 1 & -4 & -4 & +0 & 0 \\ 2 & -3 & -4 & 1 & +0 \end{pmatrix} \end{array},$$

and finally the orthogonal projection of $D_{cf}^{\widehat{X}_{1|\mathbf{g}=(1,1)}}$ *over* $\mathbf{x}(k)$

$$D^{X_{1|\mathbf{g}=(1,1)}} = D_{cf}^{\widehat{X}_{1|\mathbf{g}=(1,1)}} \Big\lceil_{\{x_1(1),x_2(1)\}}$$

$$\begin{array}{c} \\ x_0 \\ x_1(1) \\ = x_2(1) \\ x_1(0) \\ x_2(0) \end{array} \begin{array}{c} \begin{array}{ccccc} x_0 & x_1(1) & x_2(1) & x_1(0) & x_2(0) \end{array} \\ \begin{pmatrix} +0 & -4 & -5 & 0 & -1 \\ 7 & +0 & 2 & 6 & 6 \\ 8 & 3 & +0 & 7 & 7 \\ 1 & -4 & -4 & +0 & 0 \\ 2 & -3 & -4 & 1 & +0 \end{pmatrix} \end{array} \Bigg|_{\{x_1(1),x_2(1)\}} = \begin{array}{c} \\ x_0 \\ = x_1(1) \\ x_2(1) \end{array} \begin{array}{c} \begin{array}{ccc} x_0 & x_1(1) & x_2(1) \end{array} \\ \begin{pmatrix} +0 & -4 & -5 \\ 7 & +0 & 2 \\ 8 & 3 & +0 \end{pmatrix} \end{array},$$

which corresponds to

$$\begin{aligned} \mathcal{R}\left(D^{X_{1|\mathbf{g}=(1,1)}}\right) &= X_{1|\mathbf{g}=(1,1)} \\ &= \{\mathbf{x}(1) \mid 4 \leq x_1(1) \leq 7, 5 \leq x_2(1) \leq 8, -2 \leq x_2(1) - x_1(1) \leq 3\}. \end{aligned}$$

Repeating the procedure for $D^{(2,1)}$ *and* $D^{(2,2)}$*, we obtain*

$$D^{X_{1|\mathbf{g}=(2,1)}} = \begin{array}{c} \\ x_0 \\ x_1(1) \\ x_2(1) \end{array} \begin{array}{c} \begin{array}{ccc} x_0 & x_1(1) & x_2(1) \end{array} \\ \begin{pmatrix} +0 & -4 & -5 \\ 8 & +0 & 1 \\ 8 & 3 & +0 \end{pmatrix} \end{array} \quad \textit{and} \quad D^{X_{1|\mathbf{g}=(2,2)}} = \begin{array}{c} \\ x_0 \\ x_1(1) \\ x_2(1) \end{array} \begin{array}{c} \begin{array}{ccc} x_0 & x_1(1) & x_2(1) \end{array} \\ \begin{pmatrix} +0 & -5 & -6 \\ 8 & +0 & 1 \\ 8 & 2 & +0 \end{pmatrix} \end{array}.$$

Hence, $X_{1|0}$ *is represented by*

$$D^{X_{1|0}} = \left\{D^{X_{1|\mathbf{g}=(1,1)}}, D^{X_{1|\mathbf{g}=(2,1)}}, D^{X_{1|\mathbf{g}=(2,2)}}\right\}.$$

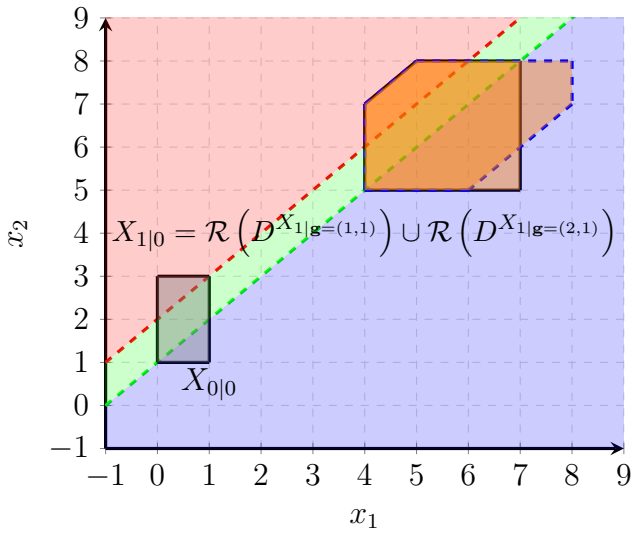


Figure 2.7 – Computation of $X_{1|0}$ of Example 2.13

Notice that $D^{X_{1|\mathbf{g}=(2,2)}} \preceq D^{X_{1|\mathbf{g}=(2,1)}}$ *then* $\mathcal{R}\left(D^{X_{1|\mathbf{g}=(2,2)}}\right) \subseteq \mathcal{R}\left(D^{X_{1|\mathbf{g}=(2,1)}}\right)$ *(cf.Remark 2.15) and then we have simply*

$$D^{X_{1|0}} = \left\{D^{X_{1|\mathbf{g}=(1,1)}}, D^{X_{1|\mathbf{g}=(2,1)}}\right\},$$

which corresponds to

$$\begin{aligned} X_{1|0} &= \mathcal{R}\left(D^{X_{1|\mathbf{g}=(1,1)}}\right) \cup \mathcal{R}\left(D^{X_{1|\mathbf{g}=(2,1)}}\right) \\ &= \{\mathbf{x}(1) \mid 4 \leq x_1(1) \leq 7, 5 \leq x_2(1) \leq 8, -2 \leq x_2(1) - x_1(1) \leq 3\} \\ &\cup \{\mathbf{x}(1) \mid 4 \leq x_1(1) \leq 8, 5 \leq x_2(1) \leq 8, -1 \leq x_2(1) - x_1(1) \leq 3\}. \end{aligned}$$

Figure 2.7 represents $X_{0|0}$ *and* $X_{1|0}$.

Set-inversion estimation

We aim at computing

$$\tilde{X}_{k|k} = \{\mathbf{x}(k) \mid \exists C \in [C], C\mathbf{x}(k) \in Z(k)\}. \tag{2.27}$$

It is assumed that $Z(k)$ is represented by the union of a finite number of DBM. Similarly to the prediction phase, the observer $\mathbf{z}(k) = C(k)\mathbf{x}(k)$ is also represented by the union of a finite number of DBM.

Inverse image of a DBM with respect to a subsystem of a PWA-uMPL observer

Let $D^{\mathbf{g}} \in (\mathbb{R} \cup \{+\infty\})^{q+n+1 \times q+n+1}$ be the DBM representation of a subsystem of a PWA-uMPL observer. Given $D^{Z(k)} \in (\mathbb{R} \cup \{+\infty\})^{q+1 \times q+1}$ representing the region $\mathcal{R}\left(D^{Z(k)}\right) = Z(k) \subseteq (\mathbb{R} \cup \{+\infty\})^q$, then its image with respect to $D^{\mathbf{g}}$ is given below:

1. Compute the Cartesian product between $D^{Z(k)}$ and $D^{\mathbb{R}^n}$, mathematically $D^{Z(k)\times\mathbb{R}^n} = D^{Z(k)} \times D^{\mathbb{R}^n}$ (see Remark 2.13);
2. Compute the intersection of $D^{Z(k)\times\mathbb{R}^n}$ with $D^{\mathbf{g}}$, mathematically $D^{\widehat{X}_k} = \min(D^{Z(k)\times\mathbb{R}^n}, D^{\mathbf{g}})$ (see Remark 2.16);
3. Compute the canonical form of $D^{\widehat{X}_k}$, mathematically, $D^{\widehat{X}_k}_{cf} = \mathtt{cf}(D^{\widehat{X}_k})$ (see Remark 2.11);
4. Compute the orthogonal projection of $D^{\widehat{X}_k}_{cf}$ over $\mathbf{x}(k)$, mathematically $D^{X_k} = D^{\widehat{X}_k}_{cf} \lceil_{\{x_1(k),\ldots,x_n(k)\}}$ (see Remark 2.12);
5. D^{X_k} is the inverse image of $D^{Z(k)}$.

The complexity of this procedure depends essentially on the canonical form computation (step 3), which is cubic in the dimension of $D^{\widehat{X}_k}$.

If $Z(k)$ is represented as the union of N_{-1} DBM then its inverse image is computed with worst-case complexity $\mathcal{O}(N_{-1}n^q(q+n)^3)$, which is also exponential in dimension n for $q = n$. The term n^q refers to the worst-case number of subsystems of the PWA-uMPL observer, and the term $(q+n)^3$ refers to the complexity of computing the image of a single DBM with respect to a subsystem of the PWA-uMPL observer.

Example 2.14. *Consider the observer $z(k) = C(k)\mathbf{x}(k)$, with $C(k) \in \begin{pmatrix}[1,3] & [e,2]\end{pmatrix}$ for all k. This observer is also represented as a PWA-uMPL system, depicted by the collection of DBM $\mathbb{D} = \{D^{(1)}, D^{(2)}\}$, given by Equation (2.24), where*

$$D^{(1)} = \begin{array}{c} \\ x_0 \\ z(k) \\ x_1(k) \\ x_2(k) \end{array} \begin{array}{c} \begin{array}{cccc} x_0 & z(k) & x_1(k) & x_2(k) \end{array} \\ \begin{pmatrix} +0 & +\infty & +\infty & +\infty \\ +\infty & +0 & 3 & +\infty \\ +\infty & -1 & +0 & +\infty \\ +\infty & 0 & 1 & +0 \end{pmatrix} \end{array} \quad \textit{and} \quad D^{(2)} = \begin{array}{c} \\ x_0 \\ z(k) \\ x_1(k) \\ x_2(k) \end{array} \begin{array}{c} \begin{array}{cccc} x_0 & z(k) & x_1(k) & x_2(k) \end{array} \\ \begin{pmatrix} +0 & +\infty & +\infty & +\infty \\ +\infty & +0 & +\infty & 2 \\ +\infty & -1 & +0 & -1 \\ +\infty & 0 & +\infty & +0 \end{pmatrix} \end{array}.$$

Suppose that $\mathbf{x}(1) = (5,7)^{\mathrm{t}}$ *originates a subset of the measurement-space, denoted* $Z(1)$*, which is, in this case, an interval vector, mathematically*

$$z(1) \in Z(1) \equiv [z](1) = [7,9] \subset \overline{\mathbb{R}}_{\max},$$

which is represented by the following DBM

$$D^{Z(1)} = \begin{array}{c} \\ x_0 \\ z(1) \end{array} \begin{array}{c} \begin{array}{cc} x_0 & z(1) \end{array} \\ \begin{pmatrix} +0 & -7 \\ 9 & +0 \end{pmatrix} \end{array}.$$

Then, $\tilde{X}_{1|1}$ *is the collection of the inverse images of* $Z(1)$ *with respect to each* $D \in \mathbb{D}$*.*

Following the procedure, we compute

$$D^{Z(1)\times\mathbb{R}^2} = D^{Z(1)} \times D^{\mathbb{R}^2} = \begin{array}{c} \\ x_0 \\ x_1(1) \\ x_1(1) \\ x_2(1) \end{array} \begin{array}{c} \begin{array}{cccc} x_0 & z(1) & x_1(1) & x_2(1) \end{array} \\ \begin{pmatrix} +0 & -7 & +\infty & +\infty \\ 9 & +0 & +\infty & +\infty \\ +\infty & +\infty & +0 & +\infty \\ +\infty & +\infty & +\infty & +0 \end{pmatrix} \end{array}.$$

Then, for $D^{(1)}$ *we compute*

$$\begin{aligned} D^{\widehat{X}_{1|\mathbf{g}=(1)}} &= \min(D^{Z(1)\times\mathbb{R}^2}, D^{(1)}) \\ &= \min\left(\begin{pmatrix} 0 & -7 & \infty & \infty \\ 9 & 0 & \infty & \infty \\ \infty & \infty & 0 & \infty \\ \infty & \infty & \infty & 0 \end{pmatrix}, \begin{pmatrix} 0 & \infty & \infty & \infty \\ \infty & 0 & 3 & \infty \\ \infty & -1 & 0 & \infty \\ \infty & 0 & 1 & 0 \end{pmatrix}\right) = \begin{pmatrix} 0 & -7 & \infty & \infty \\ 9 & 0 & 3 & \infty \\ \infty & -1 & 0 & \infty \\ \infty & 0 & 1 & 0 \end{pmatrix}, \end{aligned}$$

and thus, the canonical form of $D^{\widehat{X}_{1|\mathbf{g}=(1)}}$

$$D_{cf}^{\widehat{X}_{1|\mathbf{g}=(1)}} = \mathtt{cf}(D^{\widehat{X}_{1|\mathbf{g}=(1)}}) = \begin{array}{c} \\ x_0 \\ z(1) \\ x_1(1) \\ x_2(1) \end{array} \begin{array}{c} \begin{array}{cccc} x_0 & z(1) & x_1(1) & x_2(1) \end{array} \\ \begin{pmatrix} +0 & -7 & -4 & +\infty \\ 9 & +0 & 3 & +\infty \\ 8 & -1 & +0 & +\infty \\ 9 & 0 & 1 & +0 \end{pmatrix} \end{array},$$

and finally the orthogonal projection of $D_{cf}^{\widehat{X}_{1|\mathbf{g}=(1)}}$ *over* $\mathbf{x}(k)$

$$D^{X_{1|\mathbf{g}=(1)}} = D_{cf}^{\widehat{X}_{1|\mathbf{g}=(1)}} \Big\lceil_{\{x_1(1),x_2(1)\}}$$

$$= \begin{array}{c} \\ x_0 \\ z(1) \\ x_1(1) \\ x_2(1) \end{array} \begin{array}{c} \begin{array}{cccc} x_0 & z(1) & x_1(1) & x_2(1) \end{array} \\ \begin{pmatrix} +0 & -7 & -4 & +\infty \\ 9 & +0 & 3 & +\infty \\ 8 & -1 & +0 & +\infty \\ 9 & 0 & 1 & +0 \end{pmatrix} \end{array} \Bigg|_{\{x_1(1),x_2(1)\}}$$

$$= x_1(1) \begin{array}{c} \\ x_0 \\ x_1(1) \\ x_2(1) \end{array} \begin{array}{c} \begin{array}{ccc} x_0 & x_1(1) & x_2(1) \end{array} \\ \begin{pmatrix} +0 & -4 & +\infty \\ 8 & +0 & +\infty \\ 9 & 1 & +0 \end{pmatrix} \end{array},$$

which corresponds to

$$\begin{aligned} \mathcal{R}\left(D^{X_{1|\mathbf{g}=(1)}}\right) &= X_{1|\mathbf{g}=(1)} \\ &= \{\mathbf{x}(1) \mid 4 \leq x_1(1) \leq 8, -\infty \leq x_2(1) \leq 9, x_2(1) - x_1(1) \leq 1\}. \end{aligned}$$

Repeating the procedure for $D^{(2)}$*, we obtain*

$$D^{X_{1|\mathbf{g}=(2)}} = \begin{array}{c} \\ x_0 \\ x_1(1) \\ x_2(1) \end{array} \begin{array}{c} \begin{array}{ccc} x_0 & x_1(1) & x_2(1) \end{array} \\ \begin{pmatrix} +0 & +\infty & -5 \\ 8 & +0 & -1 \\ 9 & +\infty & +0 \end{pmatrix} \end{array},$$

which corresponds to

$$\begin{aligned} \mathcal{R}\left(D^{X_{1|\mathbf{g}=(2)}}\right) &= X_{1|\mathbf{g}=(2)} \\ &= \{\mathbf{x}(1) \mid -\infty \leq x_1(1) \leq 8, 5 \leq x_2(1) \leq 9, 1 \leq x_2(1) - x_1(1)\}. \end{aligned}$$

Hence,

$$\begin{aligned} \tilde{X}_{1|1} &= \mathcal{R}\left(D^{X_{1|\mathbf{g}=(1)}}\right) \cup \mathcal{R}\left(D^{X_{1|\mathbf{g}=(2)}}\right) \\ &= \{\mathbf{x}(1) \mid 4 \leq x_1(1) \leq 8, -\infty \leq x_2(1) \leq 9, x_2(1) - x_1(1) \leq 1\} \\ &\cup \{\mathbf{x}(1) \mid -\infty \leq x_1(1) \leq 8, 5 \leq x_2(1) \leq 9, 1 \leq x_2(1) - x_1(1)\}. \end{aligned}$$

Figure 2.8 depicts $\mathcal{R}\left(D^{X_{1|\mathbf{g}=(1)}}\right)$*,* $\mathcal{R}\left(D^{X_{1|\mathbf{g}=(2)}}\right)$ *and* $\tilde{X}_{1|1}$*.*

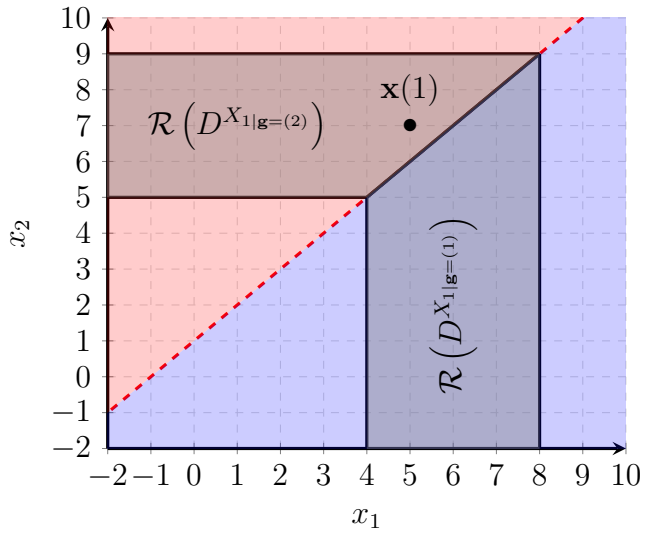


Figure 2.8 – Computation of $\tilde{X}_{1|1}$ of Example 2.14

Correction phase

In the correction phase, we correct the prior estimation set $X_{k|k-1}$ with the set $\tilde{X}_{k|k}$, obtained thanks to the observation, leading to the posterior estimation set $X_{k|k}$ (see Equation (2.25)).

Remark 2.20. *The intersection of two sets represented by the union of finitely many DBM is again a union of finitely many DBM.*

Let N_1 and N_2 be the number of DBM to represent $X_{k|k-1}$ and $\tilde{X}_{k|k}$, respectively. Then, $X_{k|k}$ is the collection of the canonical form representation of the intersection of each DBM that represents $X_{k|k-1}$ with each DBM that represents $\tilde{X}_{k|k}$. Thus, the overall complexity of computing $X_{k|k}$ is $\mathcal{O}(N_1 N_2 n^3)$, where n^3 corresponds to canonical form computation (Floyd–Warshall algorithm).

Example 2.15. *Recall the uMPL system of Examples 2.8 and 2.13. Let us consider that the associated observer is given by Example 2.14. The point* $\mathbf{x}(1) = (5, 6)^{\mathrm{t}}$ *is in* $X_{1|0}$ *and it generates* $Z(1)$ *that was used to compute* $\tilde{X}_{1|1}$ *(see Examples 2.13 and 2.14), hence* $X_{1|1} = X_{1|0} \cap \tilde{X}_{1|1}$ *is expected to be non-empty. Figure 2.9 depicts* $X_{1|1} = X_{1|0}$.

$Z(k) = \{\mathbf{z}(k)\}$ - set-inversion problem becomes an inverse mapping problem

In the vast majority of cases, more particularly in stochastic filtering problems, $\mathbf{z}(k) \in Z(k) = \{\mathbf{z}(k)\}$, *i.e.*, a sequence of measurements $\{\mathbf{z}(1), \ldots, \mathbf{z}(k)\}$ is obtained via the

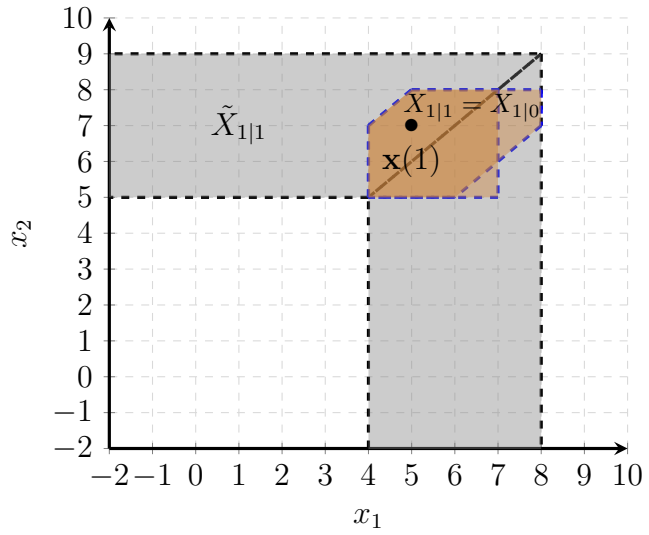


Figure 2.9 – Computation of $X_{1|1}$ of Example 2.15

observer Equation (2.14b). Then, Equation (2.27) is alternatively replaced by

$$\tilde{X}_{k|k} = \{\mathbf{x}(k) \mid \exists C \in [C], C\mathbf{x}(k) = \mathbf{z}(k)\}, \tag{2.28}$$

which is equivalent to

$$\{\mathbf{x}(k) \mid \underline{C}\mathbf{x}(k) \leq \mathbf{z}(k) \leq \overline{C}\mathbf{x}\},$$

with the solution given by Proposition 2.1. Hence, let us consider [5]

$$\tilde{X}_{k|k} = \bigcup_{\mathbf{g}=(g_1,\ldots,g_q)\in\{1,\ldots,n\}^q} SET^{\mathbf{g}} \cap U, \text{ with } SET^{\mathbf{g}} = \bigcap_{i=1}^{q} set^i_{g_i} \text{ and } U = \{\mathbf{x}(k) \mid \mathbf{x}(k) \leq \overline{\mathbf{X}}\}, \tag{2.29}$$

where

$$set^i_j = \left(\bigcap_{w=1}^{j-1} \{\mathbf{x}(k) \mid x_w(k) < z_i(k) - \overline{c}_{ik}\} \right) \cap \{\mathbf{x}(k) \mid -x_j(k) \leq \overline{c}_{ij} - z_i(k)\}.$$

The sets U and set^i_j are clearly represented by DBM with strictness sign $\bowtie \in \{<, \leq\}$,

5. $\overline{\mathbf{X}} = \underline{C} \diamond \mathbf{z}(k)$ converted in conventional algebra.

in contrast with Definition 2.4, precisely

$$D^U = \begin{array}{c} \\ x_0 \\ x_1(k) \\ \vdots \\ x_n(k) \end{array} \begin{array}{c} \begin{array}{cccc} x_0 & x_1(k) & \cdots & x_n(k) \end{array} \\ \left(\begin{array}{c|ccc} (+0,\leq) & (+\infty,<) & \cdots & (+\infty,<) \\ \hline (\overline{X}_1,\leq) & & & \\ \vdots & & \mathrm{Id}_n^{\bowtie} & \\ (\overline{X}_n,\leq) & & & \end{array}\right) \end{array} \tag{2.30}$$

where $\mathrm{Id}_n^{\bowtie}$ is a matrix of dimension $n \times n$ with $(+0, \leq)$ on the main diagonal and $(+\infty, <)$ elsewhere and

$$D^{set_j^i} = \begin{array}{c} \\ x_0 \\ x_1(k) \\ \vdots \\ x_{j-1}(k) \\ x_j(k) \\ x_{j+1}(k) \\ \vdots \\ x_n(k) \end{array} \begin{array}{c} \begin{array}{ccccc} x_0 & x_1(k) \cdots x_{j-1}(k) & x_j(k) & x_{j+1}(k) \cdots x_n(k) \end{array} \\ \left(\begin{array}{c|ccc|c|ccc} (+0,\leq) & (+\infty,<) & \cdots & (+\infty,<) & (-\underline{X}_j^{(i)},\leq) & (+\infty,<) & \cdots & (+\infty,<) \\ \hline (\underline{X}_1^{(i)},<) & & & & & & & \\ \vdots & & & & & & & \\ (\underline{X}_{j-1}^{(i)},<) & & & & & & & \\ (+\infty,<) & & & & \mathrm{Id}_n^{\bowtie} & & & \\ (+\infty,<) & & & & & & & \\ \vdots & & & & & & & \\ (+\infty,<) & & & & & & & \end{array}\right) \end{array} \tag{2.31}$$

where $\underline{X}_j^{(i)} = z_i(k) - \overline{c}_{ij}$.

Remark 2.21 (Partial order on DBM with strictness sign)**.** *Let* $(a', \bowtie')$ *and* $(a'', \bowtie'')$ *be two typical entries of DBM with strictness sign where* $\bowtie'$, $\bowtie'' \in \{<, \leq\}$ *and* $a', a'' \in \mathbb{R} \cup \{+\infty\}$*. These pairs are ordered, i.e.,* $(a', \bowtie') \preceq (a'', \bowtie'')$ *if* $a' < a''$ *or*[6] $a' = a''$ *and* $\bowtie' \preceq \bowtie''$*. (see [29, Sec. II]). Furthermore, this partial order is also applied to the element-wise operations* max *and* min *between two DBM, which are fundamental operations in order to handle intersection and union of DBM (see Remark 2.16).*

6. The symbols $<$ and $\leq$ are assumed to be partially ordered with $<$ strictly less than $\leq$, *i.e.*, $< \preceq \leq$.

Then, $SET^{\mathbf{g}}$ is represented as

$$\mathcal{R}\left(D^{SET^{\mathbf{g}}}\right) = \bigcap_{i=1}^{q} \mathcal{R}\left(D^{set^i_{g_i}}\right) = \mathcal{R}\left(\min_{i=1}^{q}\left(D^{set^i_{g_i}}\right)\right), \tag{2.32}$$

and the set $SET^{\mathbf{g}} \cap U$ is represented as

$$\mathcal{R}\left(D^{SET^{\mathbf{g}}\cap U}\right) = \mathcal{R}\left(D^{SET^{\mathbf{g}}}\right) \cap \mathcal{R}\left(D^{U}\right) = \mathcal{R}\left(\min\left(D^{SET^{\mathbf{g}}}, D^{U}\right)\right). \tag{2.33}$$

Remark 2.22 (Easy emptiness check)**.** *Let D be a DBM with strictness sign of dimension $(n+1)\times(n+1)$. If all entries different of $(+0, \leq)$ or $(+\infty, <)$ are in its first row/column, then $x_i - x_0 \bowtie' c'$ and $x_0 - x_i \bowtie'' c''$ where $\bowtie', \bowtie'' \in \{<, \leq\}$ and $c', c'' \in \mathbb{R} \cup \{+\infty\}$ for all $i \in \{1, \ldots, n\}$. Hence, rather than checking for emptiness of D using the canonical form computation, which has cubic complexity in dimension n (see Remark 2.11), it is alternatively possible to verify if it exists an empty interval $\{x_i \mid -c'' \bowtie'' x_i \bowtie' c'\}$ (note that this definition of interval differs from Definition 2.1), which is done for D with worst-case linear complexity in dimension n, i.e., $\mathcal{O}(n)$.*

It is worth to be mentioned that $D^{SET^{\mathbf{g}}\cap U}$ is computed with complexity $\mathcal{O}(q)$, however it is likely that this DBM is empty for some $\mathbf{g}$ (see Remark 2.22), hence the overall complexity of computing $D^{SET^{\mathbf{g}}\cap U}$ amounts to $\mathcal{O}(qn)$ in the worst-case scenario.

Finally, $\tilde{X}_{k|k}$ is represented as

$$\tilde{X}_{k|k} = \bigcup_{\mathbf{g}=(g_1,\ldots,g_q)\in\{1,\ldots,n\}^q} \mathcal{R}\left(D^{SET^{\mathbf{g}}\cap U}\right), \tag{2.34}$$

i.e., $\tilde{X}_{k|k}$ is represented as the collection of DBM $\{D^{SET^{\mathbf{g}=(1,\ldots,1)}\cap U}, \ldots, D^{SET^{\mathbf{g}=(n,\ldots,n)}\cap U}\}$, with cardinality n^q in worst-case scenarios. Thus, computing $\tilde{X}_{k|k}$ is done with worst-case complexity $\mathcal{O}(qn^{q+1})$ [29, Alg. 1], whereas using the set-inversion estimation procedure, the worst-case complexity is $\mathcal{O}(n^q(q+n)^3)$. Hence, the performance [7] of computing $\tilde{X}_{k|k}$ is improved. As another advantage, $\tilde{X}_{k|k}$ is represented by Equation (2.34) as a disjoint union of DBM, while it is also represented using the set-inversion procedure, but probably with overlapping DBM (see Example 2.13).

Remark 2.23 (Reinterpretation of $X_{k|k-1}$ as a collection of DBM with strictness sign)**.** *In order to compute $X_{k|k}$ using the improved computation of $\tilde{X}_{k|k}$ of Equation* (2.34)*, it is*

7. If $q = n$ then $\tilde{X}_{k|k}$ is computed with complexities $\mathcal{O}(n^{n+2})$ and $\mathcal{O}(n^{n+3})$, using inverse mapping of Equation (2.34) and using set-inversion procedure, respectively.

necessary to reinterpret the collection of DBM that represent $X_{k|k-1}$. Basically: $+0$ entries are replaced by $(+0, \leq)$; $+\infty$ entries are replaced by $(+\infty, <)$; other entries $c \in \mathbb{R}\backslash\{+0\}$ are replaced by $(c, \leq)$.

The complexity of computing $X_{k|k}$ with the new representation of $\tilde{X}_{k|k}$ given by Equation (2.34) and the reinterpretation of $X_{k|k-1}$ is still $\mathcal{O}(N_1 N_2 n^3)$ where N_1 and N_2 are the number of DBM necessary to represent $X_{k|k-1}$ and $\tilde{X}_{k|k}$, respectively. Furthermore, it is necessary to compute the canonical form representation of the intersection of each DBM that represents $X_{k|k-1}$ with each DBM that represents $\tilde{X}_{k|k}$, with a sligth modified version of the Floyd–Warshall algorithm, by considering the strictness signs of the involved DBM [67, `VeriSiMPL`]. It is also worth to note that if $X_{k|k-1}$ is given by $\{\mathbf{x}(k) \mid \underline{\mathbf{x}}(k|k-1) \leq \mathbf{x}(k) \leq \overline{\mathbf{x}}(k|k-1)\}$ (see Remark 2.17), then $X_{k|k}$ is computed with complexity $\mathcal{O}(N_2 n^2)$ since its calculation does not use the Floyd-Warshall algorithm, but the Remark 2.22.

Example 2.16. *Let us recall the uMPL system of Examples 2.8 and 2.13. Now, consider that the associated observer is the same given by Example 2.14 but with*

$$z(1) = C(1)\mathbf{x}(1) \in Z(1) = \{z(1)\},$$

where $C(1) = \begin{pmatrix} 1 & 1 \end{pmatrix}$ and $\mathbf{x}(1) = (5, 7)^{\mathrm{t}}$, yielding $z(1) = 8$, which is represented by the following DBM

$$D^{z(1)} = \begin{array}{c} \\ x_0 \\ z(1) \end{array} \begin{array}{c} \begin{array}{cc} x_0 & z(1) \end{array} \\ \begin{pmatrix} +0 & -8 \\ 8 & +0 \end{pmatrix} \end{array}.$$

Following the same procedure used in Example 2.14 we compute

$$\begin{aligned} \tilde{X}_{1|1} = &\{\mathbf{x} \mid 5 \leq x_1(1) \leq 7, -\infty \leq x_2(1) \leq 8, x_2(1) - x_1(1) \leq 1\} \\ &\cup \{\mathbf{x} \mid -\infty \leq x_1(1) \leq 7, 6 \leq x_2(1) \leq 8, 1 \leq x_2(1) - x_1(1)\}. \end{aligned}$$

However, let us consider the procedure of Equation (2.34) to alternatively compute $\tilde{X}_{1|1}$.

First, we compute D^U *using Equation* (2.30) *as*

$$D^U = \begin{array}{c} \\ x_0 \\ x_1(1) \\ x_2(1) \end{array} \begin{array}{c} \begin{array}{ccc} x_0 & x_1(1) & x_2(1) \end{array} \\ \begin{pmatrix} (+0,\leq) & (+\infty,<) & (+\infty,<) \\ (7,\leq) & (+0,\leq) & (+\infty,<) \\ (8,\leq) & (+\infty,<) & (+0,\leq) \end{pmatrix} \end{array}.$$

Then, for $\mathbf{g} = (g_1) \in \{1,2\}$ *we compute* $D^{set^i_{g_i}}$ *for* $i = 1$ *using Equation* (2.31) *as*

$$D^{set^1_{g_1=1}} = \begin{array}{c} \\ x_0 \\ x_1(1) \\ x_2(1) \end{array} \begin{array}{c} \begin{array}{ccc} x_0 & x_1(1) & x_2(1) \end{array} \\ \begin{pmatrix} (+0,\leq) & (-5,\leq) & (+\infty,<) \\ (+\infty,<) & (+0,\leq) & (+\infty,<) \\ (+\infty,<) & (+\infty,<) & (+0,\leq) \end{pmatrix} \end{array}$$

and

$$D^{set^1_{g_1=2}} = \begin{array}{c} \\ x_0 \\ x_1(1) \\ x_2(1) \end{array} \begin{array}{c} \begin{array}{ccc} x_0 & x_1(1) & x_2(1) \end{array} \\ \begin{pmatrix} (+0,\leq) & (+\infty,<) & (-6,\leq) \\ (5,<) & (+0,\leq) & (+\infty,<) \\ (+\infty,<) & (+\infty,<) & (+0,\leq) \end{pmatrix} \end{array}.$$

Thus, using Equation 2.32 we obtain that

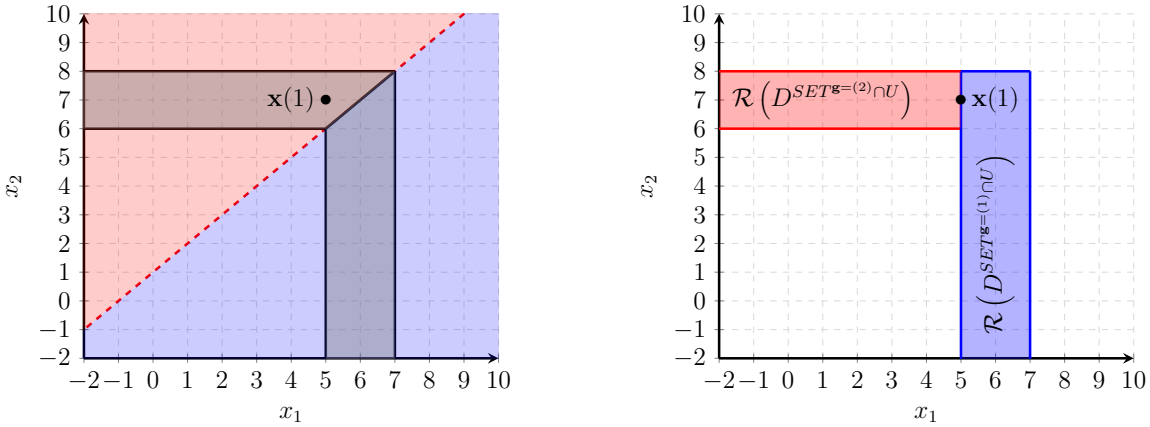
$$D^{SET^{\mathbf{g}=(1)}} = D^{set^1_{g_1=1}} \text{ and } D^{SET^{\mathbf{g}=(2)}} = D^{set^1_{g_1=2}}$$

and using Equation (2.33) *we finally obtain*

$$D^{SET^{\mathbf{g}=(1)}\cap U} = \begin{array}{c} \\ x_0 \\ x_1(1) \\ x_2(1) \end{array} \begin{array}{c} \begin{array}{ccc} x_0 & x_1(1) & x_2(1) \end{array} \\ \begin{pmatrix} (+0,\leq) & (-5,\leq) & (+\infty,<) \\ (7,\leq) & (+0,\leq) & (+\infty,<) \\ (8,\leq) & (+\infty,<) & (+0,\leq) \end{pmatrix} \end{array}$$

and

$$D^{SET^{\mathbf{g}=(2)}\cap U} = \begin{array}{c} \\ x_0 \\ x_1(1) \\ x_2(1) \end{array} \begin{array}{c} \begin{array}{ccc} x_0 & x_1(1) & x_2(1) \end{array} \\ \begin{pmatrix} (+0,\leq) & (+\infty,<) & (-6,\leq) \\ (5,<) & (+0,\leq) & (+\infty,<) \\ (8,\leq) & (+\infty,<) & (+0,\leq) \end{pmatrix} \end{array}.$$



(a) Computation using set-inversion procedure

(b) Computation using inverse mapping

Figure 2.10 – Both representations of $\tilde{X}_{1|1}$ of Example 2.16

Hence,

$$\begin{aligned}\tilde{X}_{1|1} &= \mathcal{R}\left(D^{SET^{\mathbf{g}=(1)}\cap U}\right) \cup \mathcal{R}\left(D^{SET^{\mathbf{g}=(2)}\cap U}\right) \\ &= \{\mathbf{x} \mid 5 \leq x_1(1) \leq 7, -\infty < x_2(1) \leq 8\} \cup \{\mathbf{x} \mid -\infty < x_1(1) < 5, 6 \leq x_2(1) \leq 8\}.\end{aligned}$$

Both representations of $\tilde{X}_{1|1}$ are depicted in Figure 2.10. Finally, the computation of $X_{1|1} = X_{1|0} \cap \tilde{X}_{1|1}$ is represented in Figure 2.11.

2.4.2 Polyhedral-RA

In this Section, we tackle the set-estimation problem given by the Equation (2.25) and presented in the Definition 2.6 by using Max-Plus polyhedral-sets, which will be defined in the sequel. We shall use [68, 69] to recall the results concerning Max-Plus polyhedra, which allow us to solve the set-estimation problem with algebraic operations in $\overline{\mathbb{R}}_{\max}$ only, *i.e.*, without using any kind of affine partition of the state-space.

Max-Plus polyhedra

Conventional convex polyhedra and (convex) Max-Plus (or *Tropical*) polyhedra admit *external* descriptions, in terms of half spaces (affine constraints), and *internal* descriptions, in terms of convex hulls of *generators* (*vertices* and *rays*). Nevertheless, the underlying algebras used by the two kinds of polyhedra are different, therefore the mathematical tools and definitions are also different.

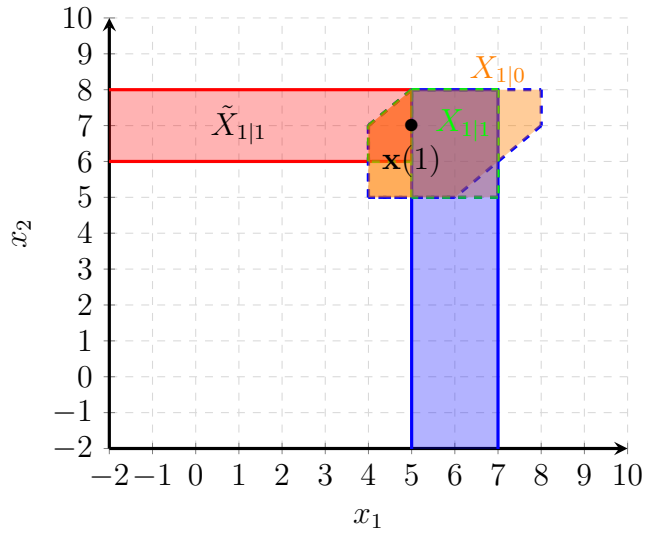


Figure 2.11 – Computation of $X_{1|1}$ of Example 2.16

Definition 2.7 (Max-Plus line segment)**.** *A subset $L_{\mathbf{v},\mathbf{w}}$ of $\overline{\mathbb{R}}_{\max}^n$ is a Max-Plus line segment if it is given by*

$$L_{\mathbf{v},\mathbf{w}} = \{\lambda\mathbf{v} \oplus \mu\mathbf{w} \mid \lambda \oplus \mu = e\}$$

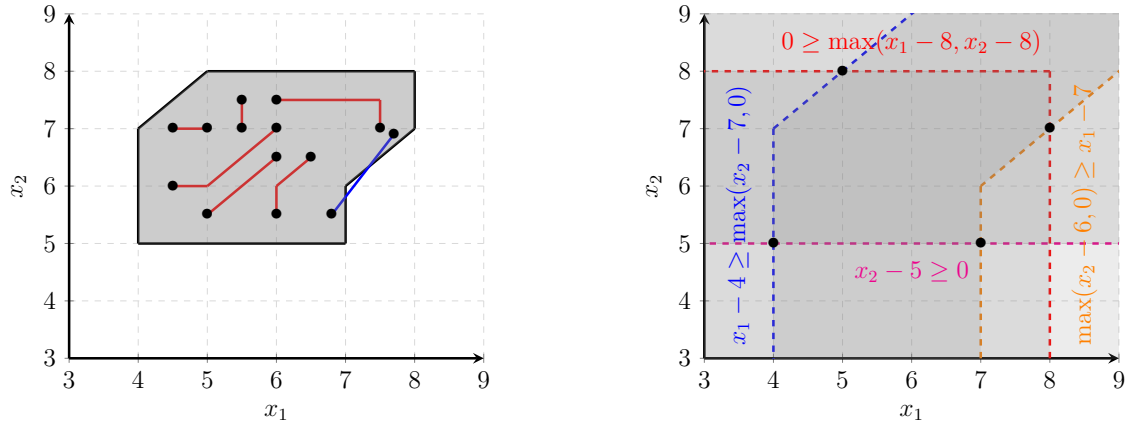
for some vectors $\mathbf{v}, \mathbf{w} \in \overline{\mathbb{R}}_{\max}^n$. Furthermore, $\mathbf{v}$ and $\mathbf{w}$ are the endpoints of L with $(\lambda = e, \mu = \varepsilon)$ and $(\lambda = \varepsilon, \mu = e)$, respectively.

Remark 2.24. *Contrary to conventional line segments, there are no positivity constraints on λ and μ for Max-Plus analogues because for all scalars $a \in \overline{\mathbb{R}}_{\max}$, we have that $a \geq \varepsilon = -\infty$, i.e., being "positive" by construction, which is omitted.*

Remark 2.25. *Max-Plus line segments of $\overline{\mathbb{R}}_{\max}^n$ are depicted as the concatenation of at most n conventional closed line segments $\mathcal{L}_{\mathbf{x},\mathbf{y}} = \{t\mathbf{x} + (1-t)\mathbf{y} \mid 0 \leq t \leq 1\}$ of slope 0,1 or $+\infty$, with endpoints $\mathbf{x}$ and $\mathbf{y}$.*

Definition 2.8 (Max-Plus convex sets)**.** *A subset $\mathcal{C}$ of $\overline{\mathbb{R}}_{\max}^n$ is Max-Plus convex if, for any $\mathbf{v}$ and $\mathbf{w}$ in $\mathcal{C}$, the Max-Plus line segment connecting $\mathbf{v}$ and $\mathbf{w}$ also lies in $\mathcal{C}$, i.e., $L_{\mathbf{v},\mathbf{w}} \in \mathcal{C}$.*

In general, Max-Plus convex sets are not convex in the conventional sense. Figure 2.12a depicts an example of a convex set in $\overline{\mathbb{R}}_{\max}^2$ and it can be seen that it contains all Max-Plus line segments between two points in itself (there are 6 kinds of Max-Plus line segments in $\overline{\mathbb{R}}_{\max}^2$ as represented by the red lines in this figure) but it does not contain the conventional line (in blue).



(a) Convexity through Max-Plus line segments

(b) Intersection of half spaces

Figure 2.12 – An example of a Max-Plus convex set in $\overline{\mathbb{R}}^2_{\max}$

Definition 2.9. *A Max-Plus half space is similar to a classical half space, and is defined as the set of points* $\mathbf{x} \in \overline{\mathbb{R}}^n_{\max}$ *satisfying*

$$\left(\bigoplus_{j=1}^{n} a_j \otimes x_j\right) \oplus b \geq \left(\bigoplus_{j=1}^{n} c_j \otimes x_j\right) \oplus d,$$

where $a_j, b, c_j, d \in \overline{\mathbb{R}}_{\max}$.

Definition 2.10. *A Max-Plus polyhedron of* $\overline{\mathbb{R}}^n_{\max}$ *is (externally) defined as the intersection of finitely many Max-Plus half spaces, i.e., as the set of points* $\mathbf{x} \in \overline{\mathbb{R}}^n_{\max}$ *satisfying* $s \geq 1$ *inequalities as given in Definition 2.9. This can be summarized as*

$$\mathcal{P} = \{\mathbf{x} \mid A\mathbf{x} \oplus \mathbf{b} \geq C\mathbf{x} \oplus \mathbf{d}\},$$

where $A, C \in \overline{\mathbb{R}}^{s \times n}_{\max}$ *and* $\mathbf{b}, \mathbf{d} \in \overline{\mathbb{R}}^s_{\max}$.

The convex set depicted in Figure 2.12a can be defined as the intersection of 4 Max-Plus half spaces, as precisely depicted in Figure 2.12b.

Definition 2.11. *The homogenization of a Max-Plus polyhedron* $\mathcal{P} = \{\mathbf{x} \mid A\mathbf{x} \oplus \mathbf{b} \geq C\mathbf{x} \oplus \mathbf{d}\} \subseteq \overline{\mathbb{R}}^n_{\max}$ *is defined as* $\widehat{\mathcal{P}} \subseteq \overline{\mathbb{R}}^{n+1}_{\max}$, *which is (externally) expressed as* $\widehat{\mathcal{P}} = \{\mathbf{z} \mid E\mathbf{z} \geq F\mathbf{z}\}$ *where* $E = (A\ \mathbf{b})$, $F = (C\ \mathbf{d})$. *When* $\mathbf{x} \in \overline{\mathbb{R}}^n_{\max}$ *and* $\alpha \in \overline{\mathbb{R}}_{\max}$, *the term* $(\mathbf{x}^{\mathrm{t}}, \alpha)^{\mathrm{t}}$ *refers to the vector* $\mathbf{z} \in \overline{\mathbb{R}}^{n+1}_{\max}$ *whose first* n *coordinates coincide with* $\mathbf{x}$ *and the latter is* α *(refers to the affine component in* $\mathcal{P}$*). Furthermore, thanks to Equation* (1.2)*, the following equivalent definition holds:* $\widehat{\mathcal{P}} = \{\mathbf{z} \mid (E \oplus F)\mathbf{z} = E\mathbf{z}\}$.

By considering the Max-Plus Minkowski-Weyl Theorem [70, Th 1], in the sequel we recall the internal representation of Max-Plus polyhedra. First, we consider the following sets.

Definition 2.12. *Given a subset* $W = \{\mathbf{w}^1, \ldots, \mathbf{w}^m\} \subset \overline{\mathbb{R}}_{\max}^n$*, the Max-Plus cone generated by* W *is defined as* $\mathtt{cone}(W) = \left\{\bigoplus_{i=1}^m \lambda_i \mathbf{w}^i \mid \mathbf{w}^i \in W, \lambda_i \in \overline{\mathbb{R}}_{\max}\right\}$.

This definition coincides with the *span* of W. Furthermore, a Max-Plus cone generated by a non-null vector is a *ray*. Such a vector is a *representative* of the ray that it generates.

Definition 2.13. *Given a subset* $V = \{\mathbf{v}^1, \ldots, \mathbf{v}^p\} \subset \overline{\mathbb{R}}_{\max}^n$*, the Max-Plus convex hull of* V *is denoted* $\mathtt{co}(V)$ *and is* $\mathtt{cone}(V)$ *with the additional constraint* $\bigoplus_{i=1}^p \lambda_i = e$.

Theorem 2.1. *A Max-Plus polyhedron* $\mathcal{P} \subseteq \overline{\mathbb{R}}_{\max}^n$ *is (internally) expressed as*

$$\begin{aligned}\mathcal{P} &= \mathtt{co}(V) \oplus \mathtt{cone}(W), \ \textit{(Minkowski sum)}\\ &= \{\mathbf{v} \oplus \mathbf{w} \mid \mathbf{v} \in \mathtt{co}(V), \mathbf{w} \in \mathtt{cone}(W)\},\end{aligned}$$

for the finite subsets $V, W \subset \overline{\mathbb{R}}_{\max}^n$ *(see [46]).*

The subsets V and W represent the generators of $\mathcal{P}$, *i.e.*, are sets of vertices and rays, respectively.

Corollary 2.1. *Let* $\mathcal{P} = \mathtt{co}(V) \oplus \mathtt{cone}(W) \subseteq \overline{\mathbb{R}}_{\max}^n$ *be a Max-Plus polyhedron, then its homogenization yields* $\widehat{\mathcal{P}} \subseteq \overline{\mathbb{R}}_{\max}^{n+1}$*, which is (internally) defined as*

$$\begin{aligned}&\widehat{\mathcal{P}} = \mathtt{cone}(\widehat{V} \cup \widehat{W}),\\ &\widehat{V} = V \times \{e\} = \{(\mathbf{v}^{\mathrm{t}}, e)^{\mathrm{t}} \mid \mathbf{v} \in V\},\ \widehat{W} = W \times \{\varepsilon\} = \{(\mathbf{w}^{\mathrm{t}}, \varepsilon)^{\mathrm{t}} \mid \mathbf{w} \in W\}.\end{aligned}$$

Thus, the internal representation of the homogenous Max-Plus polyhedron of Definition 2.11 is given by the Max-Plus cone of Corollary 2.1. Therefore, the representation of Max-Plus polyhedra is reduced to represent finitely generated Max-Plus cones.

Remark 2.26. *If* $\widehat{W} = \emptyset$ *in* $\widehat{V} \cup \widehat{W}$*, that defines* $\widehat{\mathcal{P}} = \mathtt{cone}(\widehat{V} \cup \widehat{W})$*, then* $\widehat{\mathcal{P}}$ *does not possess non-null rays according to Definitions 2.12 and 2.13. Hence,* $\widehat{\mathcal{P}}$ *is the homogenous form of* $\mathcal{P} = \mathtt{co}(V) \oplus \mathtt{cone}(\{\boldsymbol{\varepsilon}\})$ *which is equivalently equal to* $\mathcal{P} = \mathtt{co}(V) \oplus \mathtt{cone}(\emptyset) = \mathtt{co}(V)$*, where* $V = \{\mathbf{v} \mid (\mathbf{v}^{\mathrm{t}}, e)^{\mathrm{t}} \in \widehat{V}\}$*. This Max-Plus polyhedron is called compact and will be referred to as* `c-Polyhedron` *in the sequel.*

Double description

In the literature, we find some methods to convert Max-Plus cones from their external representation (cf. Definition 2.11) to their internal representation [46, 39, 69], *i.e.*, to find its generating set. It has been proved that the generating set is finitely generated [46] and is equivalent to mean payoff games [71], which are known to be in NP ∩ coNP [72], *i.e.*, the existence of a strongly polynomial procedure for solving it remains an open problem. The double description method [69] is practically efficient and simple to be used in `Polymake` [73] that implements `TPLib` [74] (even if its complexity is exponential in worst-case scenarios, as precisely described in Remark 2.27), and can be used to translate an external description into an internal description and vice-versa.

Remark 2.27 (Complexity of external-to-internal procedure)**.** *For a Max-Plus cone of* $\overline{\mathbb{R}}_{\max}^n$*, generated by the intersection of* s *half spaces, the time complexity to translate its external description into its internal description is* $\mathcal{O}(N^2 s^2 n)$*. The term* G *is related to the maximal number of generators of the* s *intermediate half spaces, and the calculation of its bounds falls back on a combinatorial problem that is hard to be determined beforehand [75]. Nevertheless, in [75, Th. 1], the authors use McMullen-type bounds to show that* G *is always upper bounded, but a general formula (i.e., for any* s *and* n*) for the lower bound is still an open problem. Precisely (see [76, Sec. 5.3.4] for more details):*

- $N \leq \mathcal{O}\left((s+n)^{\lfloor \frac{n-1}{2} \rfloor}\right)$;
- $N \geq \mathcal{O}\left((s-2n)2^{n-2}\right)$ *if* $s \geq 2n$; $N \geq \mathcal{O}\left(n^{\lfloor \frac{n-p-1}{2} \rfloor}\right)$ *if* $n \geq 2s+1$.

Hence, its upper bound is tight when $n \to +\infty$ *for a fixed* s*. In many practical applications, the exact value is much smaller than this upper bound, i.e.,* N *is a non-exponential function in the dimension* n *(i.e., the number of generators of the* s *intermediate half spaces is also non-exponential).*

Although costly, the double description is suitable to easily handle set-theoretic operations of union and intersection.

Example 2.17. *In Figure 2.12b, 4 half spaces generated a Max-Plus convex set, precisely*

$$\begin{cases} x_1 - 4 \geq \max(x_2 - 7, 0), \\ 0 \geq \max(x_1 - 8, x_2 - 8), \\ x_2 - 5 \geq 0, \\ \max(x_2 - 6, 0) \geq x_1 - 7, \end{cases} \quad or \quad \begin{pmatrix} -4 & \varepsilon \\ \varepsilon & \varepsilon \\ \varepsilon & -5 \\ \varepsilon & -6 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \oplus \begin{pmatrix} \varepsilon \\ e \\ \varepsilon \\ e \end{pmatrix} \geq \begin{pmatrix} \varepsilon & -7 \\ -8 & -8 \\ \varepsilon & \varepsilon \\ -7 & \varepsilon \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \oplus \begin{pmatrix} e \\ \varepsilon \\ e \\ \varepsilon \end{pmatrix}.$$

Hence, it is possible to obtain a Max-Plus polyhedron $\mathcal{P}$, externally represented as the intersection of these half spaces or internally as

$$\mathcal{P} = \mathtt{co}\left(\left\{\begin{pmatrix} 8 \\ 7 \end{pmatrix}, \begin{pmatrix} 7 \\ 5 \end{pmatrix}, \begin{pmatrix} 5 \\ 8 \end{pmatrix}, \begin{pmatrix} 4 \\ 5 \end{pmatrix}\right\}\right) \oplus \mathtt{cone}(\emptyset),$$

depicted by black dots in Figure 2.12b.

Membership to Max-Plus cones

Let $G = \{\mathbf{g}^1, \ldots, \mathbf{g}^k\}$ be a finite subset of $\overline{\mathbb{R}}^n_{\max}$, $\mathcal{C} = \mathtt{cone}(G) \subseteq \overline{\mathbb{R}}^n_{\max}$ be a Max-Plus cone and $\mathbf{v} \in \overline{\mathbb{R}}^n_{\max}$ be a vector. In [77], it has been proven that testing whether $\mathbf{v} \in \mathcal{C}$ is equivalent to compute the following test, with complexity $\mathcal{O}(nk)$:

$$\mathbf{v} \in \mathcal{C} \iff \bigcup_{i=1}^{k} \arg\min_{j=1}^{n}(-g_j^i + v_j) = \{1, \ldots, n\}, \tag{2.35}$$

with g_j^i the j-th component of the vector $\mathbf{g}^i$. Thus, if the test holds then $\exists \lambda_1 \oplus \cdots \oplus \lambda_k = e$ such that $\mathbf{v} = \bigoplus_{i=1}^{k} \lambda_i \mathbf{g}^i$.

Remark 2.28 (Partial order on Max-Plus cones)**.** *Given two finite subsets of $\overline{\mathbb{R}}^n_{\max}$, precisely $G = \{\mathbf{g}^i\}_{i=1}^k = \{\mathbf{g}^1, \ldots, \mathbf{g}^k\}$ and $H = \{\mathbf{h}^i\}_{i=1}^w = \{\mathbf{h}^1, \ldots, \mathbf{h}^w\}$ then a partial order can be defined on $\mathcal{G} = \mathtt{cone}(G) \subseteq \overline{\mathbb{R}}^n_{\max}$ and $\mathcal{H} = \mathtt{cone}(H) \subseteq \overline{\mathbb{R}}^n_{\max}$, as follows: $\mathcal{H} \subseteq \mathcal{G} \iff \forall \mathbf{h}^i \in H, \mathbf{h}^i \in \mathcal{G}$. The complexity of this test is $\mathcal{O}(nwk)$.*

Example 2.18. *Consider the internal representation of the Max-Plus polyhedron $\mathcal{P}$ given*

in Example 2.17. Its homogenization yields

$$\widehat{\mathcal{P}} = \mathtt{cone}\left(\left\{\underbrace{\begin{pmatrix}8\\7\\e\end{pmatrix}}_{\mathbf{g}^1}, \underbrace{\begin{pmatrix}7\\5\\e\end{pmatrix}}_{\mathbf{g}^2}, \underbrace{\begin{pmatrix}5\\8\\e\end{pmatrix}}_{\mathbf{g}^3}, \underbrace{\begin{pmatrix}4\\5\\e\end{pmatrix}}_{\mathbf{g}^4}\right\}\right),$$

which is a Max-Plus cone. Let $\mathbf{v}^1 = (8,5)^{\mathrm{t}}$ *and* $\mathbf{v}^2 = (6,7)^{\mathrm{t}}$ *be two vectors. Using Equation* (2.35) *we obtain*[8]*:*

— *for* $\widehat{\mathbf{v}}^1 = (8,5,e)^{\mathrm{t}}$*:*

$$\begin{aligned}&\left(\min_{idx}(-8+8,-7+5,0+0)\right) \cup \left(\min_{idx}(-7+8,-5+5,0+0)\right)\\ &\cup \left(\min_{idx}(-5+8,-8+5,0+0)\right) \cup \left(\min_{idx}(-4+8,-5+5,0+0)\right) = \{2,3\} \neq \{1,2,3\}\end{aligned}$$

— *for* $\widehat{\mathbf{v}}^2 = (6,7,e)^{\mathrm{t}}$*:*

$$\begin{aligned}&\left(\min_{idx}(-8+6,-7+7,0+0)\right) \cup \left(\min_{idx}(-7+6,-5+7,0+0)\right)\\ &\cup \left(\min_{idx}(-5+6,-8+7,0+0)\right) \cup \left(\min_{idx}(-4+6,-5+7,0+0)\right) = \{1,2,3\} = \{1,2,3\}\end{aligned}$$

and thus, $\widehat{\mathbf{v}}^1 \notin \widehat{\mathcal{P}} \iff \mathbf{v}^1 \notin \mathcal{P}$ *and*

$$\begin{aligned}\widehat{\mathbf{v}}^2 \in \widehat{\mathcal{P}} &\iff \widehat{\mathbf{v}}^2 = \begin{pmatrix}6\\7\\e\end{pmatrix} = -2\begin{pmatrix}8\\7\\e\end{pmatrix} \oplus -1\begin{pmatrix}7\\5\\e\end{pmatrix} \oplus -1\begin{pmatrix}5\\8\\e\end{pmatrix} \oplus e\begin{pmatrix}4\\5\\e\end{pmatrix}\\ &\iff \mathbf{v}^2 \in \mathcal{P} \iff \begin{pmatrix}6\\7\end{pmatrix} = -2\begin{pmatrix}8\\7\end{pmatrix} \oplus -1\begin{pmatrix}7\\5\end{pmatrix} \oplus -1\begin{pmatrix}5\\8\end{pmatrix} \oplus e\begin{pmatrix}4\\5\end{pmatrix}.\end{aligned}$$

Definition 2.14. *A vector* $\mathbf{v} \in \overline{\mathbb{R}}^n_{\max}$ *is said to be an extreme generator of a Max-Plus cone* $\mathcal{C} \subseteq \overline{\mathbb{R}}^n_{\max}$ *if* $\mathbf{v} \in \mathcal{C}$ *and if the following property is satisfied*

$$\mathbf{v} = \mathbf{y} \oplus \mathbf{z},\ \mathbf{y},\mathbf{z} \in \mathcal{C} \implies \mathbf{v} = \mathbf{y} \text{ or } \mathbf{v} = \mathbf{z},$$

8. The operator $\min_{idx}$ returns the index into the operating dimension that corresponds to the minimum value.

i.e., $\mathbf{v}$ *cannot be written as the* $\oplus$*-sum of two vectors of* $\mathcal{C}$ *that are both different from it.*

Remark 2.29. *If* $\mathbf{v}$ *is an extreme generator of* $\mathcal{C}$*, then the points in* $\{\lambda\mathbf{v} \mid \lambda \in \overline{\mathbb{R}}_{\max}\}$ *form an extremal ray of* $\mathcal{C}$*, i.e., are also extreme generators of* $\mathcal{C}$*.*

It follows that a Max-Plus cone $\mathcal{C} = \mathtt{cone}(G) \subseteq \overline{\mathbb{R}}^n_{\max}$ is finitely generated and its generating set $G = \{\mathbf{g}^1, \ldots, \mathbf{g}^k\} \subseteq \overline{\mathbb{R}}^n_{\max}$ satisfies its external representation of Definition 2.11. Thus, it is possible that some elements of G are *redundant*, *i.e.*, *linearly* dependent on the other ones, and in order to represent G as a non-redundant, also unique up to normalization, generating set, one must select a representative in each extreme ray of $\mathcal{C}$ (refer to [78]). Hence, $\mathcal{C}$ admits a minimal representation (or canonical form), denoted $\mathtt{cf}(\mathcal{C})$, which is generated by $G_{\mathtt{cf}} = \{\mathbf{g}^1_{\mathtt{cf}}, \ldots, \mathbf{g}^{k'}_{\mathtt{cf}}\} \subseteq \overline{\mathbb{R}}^n_{\max}$, where $k' \leq k$ and whose elements form a minimal basis of G, *i.e.*, $\mathtt{cf}(\mathcal{C}) = \mathtt{cf}(\mathtt{cone}(G)) = \mathtt{cone}(G_{\mathtt{cf}})$. The computation of $G_{\mathtt{cf}}$ is done with complexity $\mathcal{O}(nk^2)$, as summarized in the following algorithm.

Algorithm 2.1: Removing redundant generators of Max-Plus cones

Data: a finite set $G = \{\mathbf{g}^1, \ldots, \mathbf{g}^k\} \subseteq \overline{\mathbb{R}}^n_{\max}$, assuming its elements are not proportional to each other;
Result: a set $G_{\mathtt{cf}} = \{\mathbf{g}^1, \ldots, \mathbf{g}^{|\mathcal{J}|}\}$, with $|\mathcal{J}| \leq k$.
`/* Goal: Eliminate the elements of` G `which are a linear combination of the other ones */`

foreach $i \in \{1, \ldots, k\}$ **do**
  $\mathcal{J} \leftarrow \emptyset$, $\mathbf{v} \leftarrow \mathbf{g}^j$ and $H \leftarrow \{\mathbf{g}^i\}_{i \in \{1,\ldots,k\} \setminus j}$
  if $\mathbf{v} \in \mathtt{cone}(H)$ *(i.e.,* $\mathbf{v}$ *is redundant cf. Equation* (2.35)*)* **then**
    $\mathbf{g}^j \leftarrow \boldsymbol{\varepsilon}$
  else
    $\mathcal{J} \leftarrow \mathcal{J} \cup \{j\}$
  end
  $G_{\mathtt{cf}} = \{\mathbf{g}^j\}_{j \in \mathcal{J}}$
end
return $G_{\mathtt{cf}} = \{\mathbf{g}^1, \ldots, \mathbf{g}^{|\mathcal{J}|}\}$

Clearly, the operator $\mathtt{cf}(\mathcal{C})$ is idempotent, *i.e.*, $\mathtt{cf}(\mathtt{cf}(\mathcal{C})) = \mathtt{cf}(\mathcal{C})$ for an arbitrary Max-Plus cone $\mathcal{C}$.

Remark 2.30. *Algorithm 2.1 is exclusively for cones of the form* $\mathcal{C} = \mathtt{cone}(G) \subseteq \overline{\mathbb{R}}^n_{\max}$ *with* $G = \{\mathbf{g}^1, \ldots, \mathbf{g}^k\} \subseteq \overline{\mathbb{R}}^n_{\max}$ *such that*

$$\forall \mathbf{x} \in \mathcal{C} \iff \mathbf{x} = \bigoplus_{i=1}^{k} \lambda_i \mathbf{g}^i.$$

Conversely, $\mathcal{C}' = \mathtt{co}(G) \subseteq \overline{\mathbb{R}}_{\max}^n$ *with the same* G *of* $\mathcal{C}$ *has the following that holds*

$$\forall \mathbf{x} \in \mathcal{C}' \iff \mathbf{x} = \bigoplus_{i=1}^{k} \lambda_i \mathbf{g}^i, \ \ s.t. \ \bigoplus_{i=1}^{k} \lambda_i = e,$$

which is different of $\mathcal{C}$ *and is not taken into account by this algorithm. According to Corollary 2.1,* $\mathcal{C}'$ *is a Max-Plus polyhedron, i.e.,* $\mathtt{co}(G) \oplus \mathtt{cone}(\emptyset) = \mathtt{co}(G) \subseteq \overline{\mathbb{R}}_{\max}^n$*, and its homogenous cone is given by* $\widehat{\mathcal{C}}' = \mathtt{cone}(\widehat{G}) \subseteq \overline{\mathbb{R}}_{\max}^{n+1}$*, where* $\widehat{G} = G \times \{e\}$*, which is a* ***c-Polyhedron*** *according to Remark 2.26. Summarizing, computing* $\mathtt{cf}(\mathcal{C}')$ *is done by considering* $\mathtt{cf}(\widehat{C}')$ *with complexity* $\mathcal{O}((n+1)k^2)$*, resulting in set* $\widehat{G}_{\mathtt{cf}}$ *and thus in set*

$$G_{\mathtt{cf}} = \left\{ \widehat{\mathbf{g}}^i_{\mathtt{cf}} \mid \begin{pmatrix} \widehat{\mathbf{g}}^i_{\mathtt{cf}} \\ e \end{pmatrix} \in \widehat{G}_{\mathtt{cf}}, 1 \leq i \leq k \right\}.$$

The external-to-internal translation algorithm (`compute_ext_rays` of `TPLib` [74]) of the double description method computes a *minimal* generating set of a polyhedral (cone) set, which is externally represented by its homogenous form. Briefly, given a polyhedral set defined by a system of s half spaces in homogenous form, it computes by induction on $k = 1, 2, \ldots, s$ a generating set $G_k \subset \overline{\mathbb{R}}_{\max}^{n+1}$ of the intermediate homogenous cone (cf. Corollary 2.1) defined by the first k half spaces. Passing from the set G_k to the set G_{k+1} relies on: given a homogenous cone $\mathcal{C}$ and a homogenous half space $\mathcal{H} = \{\mathbf{z} \mid \mathbf{a}^{\mathrm{t}}\mathbf{z} \geq \mathbf{b}^{\mathrm{t}}\mathbf{z}\} \subseteq \overline{\mathbb{R}}_{\max}^{n+1}$, allows building a (minimal) generating set G' of $\mathcal{C} \cap \mathcal{H}$ from a generating set G of $\mathcal{C}$. The initial step is with $G_0 = \{\boldsymbol{\epsilon}^i\}_{1 \leq i \leq n+1}$ representing $\overline{\mathbb{R}}_{\max}^{n+1}$, with $\boldsymbol{\epsilon}^i \in \overline{\mathbb{R}}_{\max}^{n+1}$ a vector whose i-th coordinate is equal to e, and the others to ε. At the end, $k = s$ and $\mathtt{cf}(\mathtt{cone}(G_s)) = \mathtt{cone}(G_s)$ (idempotence property).

Remark 2.31 (Replacing the initial generating set: intersection)**.** *Based on the external-to-internal translation algorithm of the double description method, it is assumed that* $G_0 = \{\boldsymbol{\epsilon}^i\}_{1 \leq i \leq n}$ *represents* $\overline{\mathbb{R}}_{\max}^n$*. However, considering that* G_0 *represents an arbitrary Max-Plus cone, in its internal form allows computing the intersection of* G_0 *with another Max-Plus cone, in its external form, also of* $\overline{\mathbb{R}}_{\max}^n$*. Summing-up, replacing* G_0 *by an arbitrary cone in the initial step of the procedure, exactly yields a minimal system of generators of this intersection.*

As mentioned previously, the external representation of Max-Plus cones encodes a single half space in each row. Consequently, the intersection between two Max-Plus cones can be performed simply by concatenating the rows of them, as precisely described in the sequel.

Remark 2.32 (Intersection of Max-Plus cones)**.** *Let* $\widehat{\mathcal{P}} = \{\mathbf{z} \mid A\mathbf{z} \geq B\mathbf{z}\}$ *and* $\widehat{\mathcal{P}}' = \{\mathbf{z} \mid C\mathbf{z} \geq D\mathbf{z}\}$ *be two Max-Plus cones of* $\overline{\mathbb{R}}_{\max}^{n+1}$ *where* $A, B, C, D \in \overline{\mathbb{R}}_{\max}^{s\times(n+1)}$*. Then,*

$$\widehat{\mathcal{P}} \cap \widehat{\mathcal{P}}' = \left\{\mathbf{z} \mid \begin{pmatrix} A \\ C \end{pmatrix} \mathbf{z} \geq \begin{pmatrix} B \\ D \end{pmatrix} \mathbf{z}\right\} \equiv \{\mathbf{z} \mid \mathbf{z} \in \widehat{\mathcal{P}} \text{ and } \mathbf{z} \in \widehat{\mathcal{P}}'\}.$$

As already pointed out, translating external representations of Max-Plus cones to internal representation and vice-versa is cumbersome, and it would be interesting to efficiently compute the intersection of these cones when they are internally represented by generating sets. Thus, we propose the following intersection of `c-Polyhedra` in internal form (see Remark 2.26), adapted from [76, Sec. 8.2.1].

Remark 2.33 (Finite subsets as matrices)**.** *A finite subset* $V = \{\mathbf{v}^1, \ldots, \mathbf{v}^k\}$ *of* $\overline{\mathbb{R}}_{\max}^{n}$ *admits a reinterpretation as a matrix*

$$\mathtt{mat}(V) = (\mathbf{v}^1, \ldots, \mathbf{v}^k) \in \overline{\mathbb{R}}_{\max}^{n\times k},$$

i.e., as an "array" of column vectors, with the same expressiveness.

Proposition 2.2 (Intersection of `c-Polyhedra`)**.** *Let* $\mathcal{G} = \mathtt{co}(G)$ *and* $\mathcal{H} = \mathtt{co}(H)$ *be two* `c-Polyhedra` *of* $\overline{\mathbb{R}}_{\max}^{n}$ *where* $G = \{\mathbf{g}^1, \ldots, \mathbf{g}^{n_G}\}$ *and* $H = \{\mathbf{h}^1, \ldots, \mathbf{h}^{n_H}\}$ *are finite subsets of* $\overline{\mathbb{R}}_{\max}^{n}$*. Then* $\mathcal{I} = \mathcal{G} \cap \mathcal{H}$ *is represented by the Max-Plus cone* $\widehat{\mathcal{I}} = \mathtt{cone}(I)$ *where the elements of its generating set* I *are the columns of the matrix* $\widehat{G}\mathtt{mat}(\Gamma)$ *where* $\Gamma = \{\boldsymbol{\gamma} \mid \widehat{G}\boldsymbol{\gamma} = \widehat{H}\boldsymbol{\gamma}\}$*,* $\widehat{G} = (\mathtt{mat}(G \times \{e\})\ \mathcal{E})$*,* $\widehat{H} = (\mathcal{E}\ \mathtt{mat}(H \times \{e\}))$ *and* $\boldsymbol{\gamma} = (\boldsymbol{\lambda}^{\mathrm{t}}, \boldsymbol{\beta}^{\mathrm{t}})^{\mathrm{t}}$*.*

Proof. First, thanks to Remark 2.26 we have

$$\forall \mathbf{v} \in \mathcal{I} \iff (\mathbf{v}^{\mathrm{t}}, e)^{\mathrm{t}} \in \widehat{\mathcal{I}}$$
$$\begin{pmatrix} \mathbf{v} \\ e \end{pmatrix} = \mathtt{mat}(G \times \{e\})\boldsymbol{\lambda} = \mathtt{mat}(H \times \{e\})\boldsymbol{\beta} \Rightarrow \mathtt{mat}(G \times \{e\})\boldsymbol{\lambda} = \mathtt{mat}(H \times \{e\})\boldsymbol{\beta}$$

Thus, in order to obtain a two-sided equation in the same variable, we consider $\Gamma = \{\boldsymbol{\gamma} \mid \widehat{G}\boldsymbol{\gamma} = \widehat{H}\boldsymbol{\gamma}\}$. Hence, the elements of the generating set I of $\widehat{\mathcal{I}}$ are the columns of the

matrix $\widehat{G}\mathtt{mat}(\Gamma)$. Moreover, it is particularly interesting to compute the minimal canonical representation of $\widehat{\mathcal{I}}$, *i.e.*, $\mathtt{cf}(\widehat{\mathcal{I}}) = \mathtt{cf}(\mathtt{cone}(I)) = \mathtt{cone}(I_{\mathtt{cf}}))$. □

Rather than considering this costly computation, for instance using the function `mpsolve` that implements, in the Max-Plus toolbox [9] of `Scilab` and `ScicosLab`, the *elimination method* (refined in [39]), with worst-case complexity that is exponential in dimension $n_G + n_H$, we define the following fixed-point equation problem in order to check if this intersection is non-empty.

Procedure 2.1. *Let $\widehat{G}\boldsymbol{\gamma} = \widehat{H}\boldsymbol{\gamma}$ be the two-sided equation raised in Proposition 2.2. Then, define*

$$\Pi(\boldsymbol{\gamma}) = \widehat{H} \backslash (\widehat{G}\boldsymbol{\gamma}) \wedge \widehat{G} \backslash (\widehat{H}\boldsymbol{\gamma}) \wedge \boldsymbol{\gamma} \tag{2.36}$$

as a fixed-point equation problem (see Equation (1.4)) with $\boldsymbol{\gamma}_0 = \mathbf{e}$ since $\bigoplus_{k=1}^{n_G+n_H} \gamma_{0_k} = e$. If the greatest solution $\boldsymbol{\gamma}$ of $\widehat{G}\boldsymbol{\gamma} = \widehat{H}\boldsymbol{\gamma}$ exists, then it is given by the fixed-point equation $\mathcal{F} : \boldsymbol{\gamma}_l = \Pi(\boldsymbol{\gamma}_{l-1})$ (with convergence), which is smaller than or equal to $\boldsymbol{\gamma}_0$.

Remark 2.34. *Clearly, $\widehat{G}\boldsymbol{\gamma} = \widehat{H}\boldsymbol{\gamma}$ is solvable if and only if the fixed-point equation problem $\boldsymbol{\gamma}_l = \Pi(\boldsymbol{\gamma}_{l-1})$ converges, and in this case $\mathcal{I} = \mathcal{G} \cap \mathcal{H} \neq \emptyset$. Nevertheless, if the former problem is not solvable, i.e., $\widehat{G}\boldsymbol{\gamma} = \widehat{H}\boldsymbol{\gamma}$ does not possess a finite solution, then it is likely to run infinitely, with $\boldsymbol{\gamma}_l$ converging to $\boldsymbol{\varepsilon}$ (refer to Remark 1.6). In order to avoid this bad behavior, we consider the following stop condition to the fixed-point procedure: evaluate at each iteration l if at least one element of $\boldsymbol{\gamma}_l$ is equal to e, i.e., if $\bigoplus_{k=1}^{n_G+n_H} \gamma_{l_k} = e$ is respected, and if this statement holds then the procedure keeps running, else it stops and $\mathcal{I} = \emptyset$ (see [44] for more details).*

In the following, we consider the join operation between two Max-Plus polyhedra in their homogeneous forms, which yields the smallest over-approximation polyhedron of their exact union. As a matter of fact, neither the Max-Plus polyhedra nor DBM are generally able to calculate exactly the union of the regions they represent. However, we observe, in practice, that the over-approximation using polyhedra has lower risk of introducing false-positive points than using DBM [79, Cor. 2], *i.e.*, approximations using Max-Plus polyhedra are tighter than DBM approximations as it is illustrated in Figure 2.13. In the sequel, we provide the technical details concerning this operation using Max-Plus cones.

Remark 2.35 (Join operation of Max-Plus cones)**.** *Given two finite subsets of $\overline{\mathbb{R}}_{\max}^n$, precisely $G = \{\mathbf{g}^1, \ldots, \mathbf{g}^k\}$ and $H = \{\mathbf{h}^1, \ldots, \mathbf{h}^w\}$ then the join operation (or abstract*

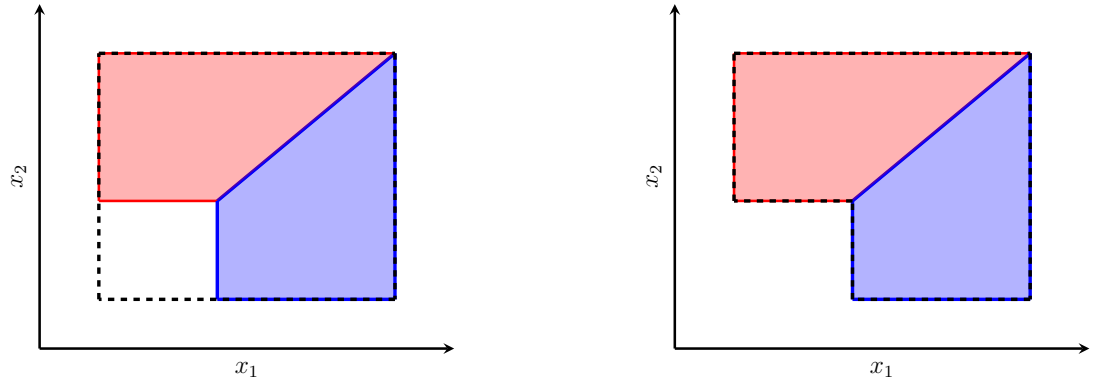
union) $\sqcup$ *between* $\mathcal{G} = \texttt{cone}(G) \subseteq \overline{\mathbb{R}}^n_{\max}$ *and* $\mathcal{H} = \texttt{cone}(H) \subseteq \overline{\mathbb{R}}^n_{\max}$*, is given by*

$$\mathcal{G} \sqcup \mathcal{H} = \texttt{cf}(\texttt{cone}(G \cup H)),$$

which has complexity $\mathcal{O}(n(k+w)^2)$*. Furthermore,*

$$\mathcal{G} \cup \mathcal{H} = \{\mathbf{v} \mid \mathbf{v} \in \mathcal{G} \text{ or } \mathbf{v} \in \mathcal{H}\} \subseteq \mathcal{G} \sqcup \mathcal{H}.$$

Figure 2.13 shows an interpretation of the tightness in the over-approximation of the union of two regions using DBM and Max-Plus polyhedra. Note that, in this case, the join operation using Max-Plus polyhedra yields the exact union.



(a) Using DBM (b) Using Max-Plus polyhedra

Figure 2.13 – Over-approximation of the union of two regions using DBM and Max-Plus polyhedra

Polytropes: a special subclass of `c-Polyhedra`

Definition 2.15. *A polytrope*[9] *is a* ***`c-Polyhedron`*** $\mathcal{P} = \texttt{co}(V) \oplus \texttt{cone}(\emptyset) = \texttt{co}(V) \subseteq \overline{\mathbb{R}}^n_{\max}$ *(see Remark 2.26), with the finite subset* $V \subseteq \overline{\mathbb{R}}^n_{\max}$*, being also convex in the ordinary sense.*

Remark 2.36. *Polytropes of* $\overline{\mathbb{R}}^n_{\max}$ *are externally defined by the intersection of a finite number of classical half spaces of the form* $x_i - x_j \leq c_{ij}$ *for all* $i, j \in \{0, \ldots, n\}$ *where* $x_i, x_j, c_{ij} \in \mathbb{R}$ *and* $x_0 = 0$*, i.e., as DBM. Hence, the external representation of polytrope is similar to the one used in DBM with non-strict constraints (see Remark 2.36), and we*

9. A polytope is the convex hull of a nonempty finite set. A polytrope is the Max-Plus analogue of a polytope.

can always represent these DBM by using Max-Plus convex polyhedral sets [80, 81, 82]. Furthermore, it is worth to be mentioned that a single DBM can only represent Max-Plus polyhedral sets that are also convex in the classical context (cf. Definition 2.15), i.e., polytrope, otherwise a collection of DBM must be considered. Hence, a Max-Plus polyhedron is represented by a collection of DBM.

Example 2.19. *Figure 2.12b depicts a Max-Plus polyhedron. This polyhedron is also represented by*

$$X = \overbrace{\{\mathbf{x} \mid 4 \leq x_1 \leq 7, 5 \leq x_2 \leq 8, -2 \leq x_2 - x_1 \leq 3\}}^{=G} \cup \underbrace{\{\mathbf{x} \mid 4 \leq x_1 \leq 8, 5 \leq x_2 \leq 8, -1 \leq x_2 - x_1 \leq 3\}}_{=H},$$

a collection DBM as given in Example 2.13. The sets G and H are equivalent to the external representation of the polytropes

$$\mathcal{G} = \left\{ \mathbf{x} \mid \begin{pmatrix} e & \varepsilon \\ \varepsilon & \varepsilon \\ \varepsilon & e \\ \varepsilon & \varepsilon \\ \varepsilon & e \\ 3 & \varepsilon \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \oplus \begin{pmatrix} \varepsilon \\ 7 \\ \varepsilon \\ 8 \\ \varepsilon \\ \varepsilon \end{pmatrix} \geq \begin{pmatrix} \varepsilon & \varepsilon \\ e & \varepsilon \\ \varepsilon & \varepsilon \\ \varepsilon & e \\ -2 & \varepsilon \\ \varepsilon & e \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \oplus \begin{pmatrix} 4 \\ \varepsilon \\ 5 \\ \varepsilon \\ \varepsilon \\ \varepsilon \end{pmatrix} \right\},$$

$$\mathcal{H} = \left\{ \mathbf{x} \mid \begin{pmatrix} e & \varepsilon \\ \varepsilon & \varepsilon \\ \varepsilon & e \\ \varepsilon & \varepsilon \\ \varepsilon & e \\ 3 & \varepsilon \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \oplus \begin{pmatrix} \varepsilon \\ 8 \\ \varepsilon \\ 8 \\ \varepsilon \\ \varepsilon \end{pmatrix} \geq \begin{pmatrix} \varepsilon & \varepsilon \\ e & \varepsilon \\ \varepsilon & \varepsilon \\ \varepsilon & e \\ -1 & \varepsilon \\ \varepsilon & e \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \oplus \begin{pmatrix} 4 \\ \varepsilon \\ 5 \\ \varepsilon \\ \varepsilon \\ \varepsilon \end{pmatrix} \right\},$$

which also admit the internal representations

$$\mathcal{G} = \mathsf{co}\left(\left\{ \begin{pmatrix} 5 \\ 8 \end{pmatrix}, \begin{pmatrix} 7 \\ 5 \end{pmatrix}, \begin{pmatrix} 4 \\ 5 \end{pmatrix} \right\} \right),$$

$$\mathcal{H} = \mathsf{co}\left(\left\{ \begin{pmatrix} 4 \\ 7 \end{pmatrix}, \begin{pmatrix} 5 \\ 8 \end{pmatrix}, \begin{pmatrix} 7 \\ 5 \end{pmatrix}, \begin{pmatrix} 8 \\ 7 \end{pmatrix} \right\} \right),$$

such that (see Remark 2.35)

$$\begin{aligned}\mathcal{G} \sqcup \mathcal{H} &= \mathtt{cf}\left(\mathtt{co}\left(\left\{\begin{pmatrix}5\\8\end{pmatrix},\begin{pmatrix}7\\5\end{pmatrix},\begin{pmatrix}4\\5\end{pmatrix},\begin{pmatrix}4\\7\end{pmatrix},\begin{pmatrix}8\\7\end{pmatrix}\right\}\right)\right),\\ &= \mathtt{co}\left(\left\{\begin{pmatrix}8\\7\end{pmatrix},\begin{pmatrix}7\\5\end{pmatrix},\begin{pmatrix}5\\8\end{pmatrix},\begin{pmatrix}4\\5\end{pmatrix}\right\}\right),\end{aligned}$$

which is equivalent to the Max-Plus polyhedron computed in Example 2.17. Moreover, in this case $\mathcal{G} \cup \mathcal{H} = \mathcal{G} \sqcup \mathcal{H}$.

Remark 2.37. *It is known that a n-dimensional polytrope is the Max-Plus convex hull of at most* $n+1$ *points or equivalently as the intersection of* $n+1$ *Max-Plus half spaces in* $\overline{\mathbb{R}}^n_{\max}$ *(refer to [68, Th 3.2] for further details and proofs).*

Proposition 2.3. *The polytrope expressed by* $\mathcal{P}_{[\mathbf{h}]} = \mathtt{co}(H) \subseteq \overline{\mathbb{R}}^n_{\max}$, *with*

$$H = \left\{\underline{\mathbf{h}}, \mathbf{v}^1, \ldots, \mathbf{v}^n\right\} \subseteq \overline{\mathbb{R}}^n_{\max}, \text{ where } \mathbf{v}^i = (\underline{h}_1, \underline{h}_2, \ldots, \overline{h}_i, \underline{h}_{i+1}, \ldots, \underline{h}_n)^{\mathrm{t}}, \tag{2.37}$$

represents the non-empty Max-Plus hypercube defined as $[\mathbf{h}] = [\underline{\mathbf{h}}, \overline{\mathbf{h}}] = \{\mathbf{h} \mid \underline{\mathbf{h}} \leq \mathbf{h} \leq \overline{\mathbf{h}}\} \subseteq \overline{\mathbb{R}}^n_{\max}$ *(see Definition 2.1).*

Proof. A simple sketch of the proof is given by considering that $[\mathbf{h}]$ is also represented by the following set $[\mathbf{h}] = S_1 \cap S_2$, where

$$\begin{aligned}S_1 &= \bigcap_{j=1}^{n} \{\mathbf{h} \mid h_j \geq \underline{h}_j\} = \bigcap_{j=1}^{n} \{\mathbf{h} \mid \underline{h}_j^{-1} \otimes h_j \geq e\} \text{ and}\\ S_2 &= \bigcap_{j=1}^{n} \{\mathbf{h} \mid h_j \leq \overline{h}_j\} = \bigcap_{j=1}^{n} \{\mathbf{h} \mid \overline{h}_j^{-1} \otimes h_j \leq e\}.\end{aligned}$$

Thus, $[\mathbf{h}]$ is the intersection of $2n$ conventional half spaces. Nevertheless, S_2 is also equivalent to

$$\tilde{S}_2 = \left\{\mathbf{h} \mid \bigoplus_{j=1}^{n} \overline{h}_j^{-1} \otimes h_j \leq e\right\},$$

representing a single Max-Plus half space. Hence, we can conclude that $[\mathbf{h}] = S_1 \cap \tilde{S}_2$, *i.e.*, $[\mathbf{h}]$ is the intersection of $n+1$ Max-Plus half spaces [10]. For instance, Figure 2.14 depicts these Max-Plus half spaces in $\overline{\mathbb{R}}^2_{\max}$. Note that all Max-Plus half spaces in S_1 intersect

10. It is worth to recall that every conventional half space is a Max-Plus half space, but the converse does not hold.

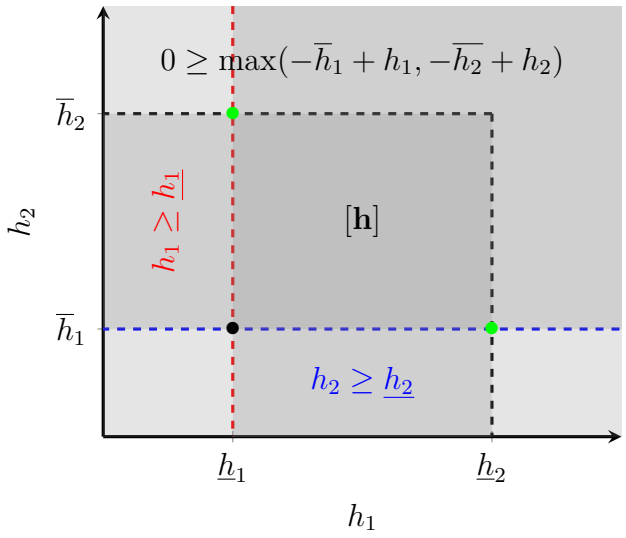


Figure 2.14 – $[\mathbf{h}]$ and its polytrope representation $\mathcal{P}_{[\mathbf{h}]} = \mathtt{co}(H)$ in $\overline{\mathbb{R}}^2_{\max}$ with $H = \left\{ \begin{pmatrix} \underline{h}_1 \\ \underline{h}_2 \end{pmatrix}, \begin{pmatrix} \overline{h}_1 \\ \underline{h}_2 \end{pmatrix}, \begin{pmatrix} \underline{h}_1 \\ \overline{h}_2 \end{pmatrix} \right\}$

at $\underline{\mathbf{h}}$, *i.e.*, being a common point (black dot); each Max-Plus half space in S_2 is oriented toward a direction j and intersects with the single Max-Plus half space in $\tilde{S}_2$ at a point $\mathbf{v}^{(j)}$ (green dots) whose i-th coordinate is equal to $\underline{h}_i$ for all $i \in \{1, \ldots, n\} \backslash j$, and the j-th coordinate is equal to $\overline{h}_j$. At the end, we have $n+1$ unique points of intersection which coincide with the elements of H, and thus Remark 2.37 holds. □

Furthermore, this result is also verifiable by considering Remark 2.30, *i.e.*, by checking that the following equality holds: $\mathtt{cf}(\mathcal{P}_{[\mathbf{h}]}) = \mathcal{P}_{[\mathbf{h}]}$ (idempotence property).

Corollary 2.2. *The polytrope $\mathcal{P}_{[\mathbf{h}]}$ of Proposition 2.3 is externally represented by*

$$\mathcal{P}_{[\mathbf{h}]} = \left\{ \mathbf{h} \mid \begin{pmatrix} \mathtt{diag}(\underline{\mathbf{h}}^{-1}) \\ \boldsymbol{\varepsilon}^{\mathrm{t}} \end{pmatrix} \mathbf{h} \oplus \begin{pmatrix} \boldsymbol{\varepsilon} \\ e \end{pmatrix} \geq \begin{pmatrix} \mathcal{E} \\ (\overline{\mathbf{h}}^{-1})^{\mathrm{t}} \end{pmatrix} \mathbf{h} \oplus \begin{pmatrix} \boldsymbol{e} \\ \varepsilon \end{pmatrix} \right\}$$

and its homogenous form $\widehat{\mathcal{P}}_{[\mathbf{h}]}$ is finitely generated by $\widehat{H} = H \times \{e\}$, with H given by Equation (2.37). In other words, all elements in $\widehat{H}$ are vertices (i.e., their $(n+1)$-th coordinate is e).

Proof. Let us suppose that $\widehat{\mathcal{P}}_{[\mathbf{h}]}$ is generated by a ray (*i.e.*, its $(n+1)$-th coordinate is ε),

then the following inequality must be respected

$$\left(\boldsymbol{\varepsilon}^{\mathrm{t}}\ e\right)\begin{pmatrix}\mathbf{h}\\ \varepsilon\end{pmatrix} \geq \left((\overline{\mathbf{h}}^{-1})^{\mathrm{t}}\ \varepsilon\right)\begin{pmatrix}\mathbf{h}\\ \varepsilon\end{pmatrix} \implies \varepsilon \geq (\overline{\mathbf{h}}^{-1})^{\mathrm{t}}\mathbf{h},$$

which is not the case. Hence, $\widehat{\mathcal{P}}_{[\mathbf{h}]}$ is only generated by vertices. □

The following equivalences always hold true:

$$\begin{aligned}\forall \mathbf{h} \in [\mathbf{h}] \iff \mathbf{h} \in \mathcal{P}_{[\mathbf{h}]} = \mathtt{co}(H) &= \left\{ \bigoplus_{j=1}^{n+1} \beta_j \mathbf{h}^j \mid \mathbf{h}^j \in H, \bigoplus_{j=1}^{n+1} \beta_j = e \right\} \\ &\iff \mathbf{h} = \bigoplus_{j=1}^{n+1} \beta_j \mathbf{h}^j, \text{ s.t. } \bigoplus_{j=1}^{n+1} \beta_j = e.\end{aligned}$$

Set-estimation using Polyhedral-RA: Prior estimation

This Section presents an approach to compute the image of a `c-Polyhedron` $\mathcal{P}_{k|k-1} = \mathtt{co}(G_{k|k-1}) \subseteq \overline{\mathbb{R}}^n_{\max}$, with $G_{k|k-1} = \{\mathbf{g}^1_{k|k-1}, \dots \mathbf{g}^N_{k|k-1}\} \subseteq \overline{\mathbb{R}}^n_{\max}$, *i.e.*, representing $\mathbf{x}(k) \in X_{k|k-1}$ of Equation (2.26). Furthermore, it is assumed that $X_{0|0}$ is known and is represented by $\mathcal{P}_{0|0} = \mathtt{co}(G_{0|0}) \subseteq \overline{\mathbb{R}}^n_{\max}$ with $G_{0|0} = \{\mathbf{g}^1_{0|0}, \dots \mathbf{g}^k_{0|0}\} \subseteq \overline{\mathbb{R}}^n_{\max}$.

In order to take the control input $\mathbf{u}(k)$ into account, we define the following extended polyhedron

$$\tilde{\mathcal{P}}_{k-1|k-1} = \mathcal{P}_{k-1|k-1} \times \{\mathbf{u}(k)\} = \mathtt{co}(G_{k-1|k-1}) \times \{\mathbf{u}(k)\} = \mathtt{co}(\tilde{G}_{k-1|k-1}) \subseteq \overline{\mathbb{R}}^{n+p}_{\max}, \quad (2.38)$$

where $\tilde{G}_{k-1|k-1} = G_{k-1|k-1} \times \{\mathbf{u}(k)\} \subseteq \overline{\mathbb{R}}^{n+p}_{\max}$. Clearly, this extended polyhedron represents the set $X_{k-1|k-1} \times U_k$ of Equation (2.26).

Definition 2.16 (Bounded linear maps)**.** *Let $f : \overline{\mathbb{R}}^n_{\max} \to \overline{\mathbb{R}}^n_{\max}$ be a linear map, such that $\underline{f} \preceq f \preceq \overline{f}$, and we denote $[f] = [\underline{f}, \overline{f}]$, then $[f](\mathbf{v}) = [\underline{f}(\mathbf{v}), \overline{f}(\mathbf{v})]$ and the following equivalences hold*

1. $\alpha[f](\mathbf{v}) = [f](\alpha\mathbf{v}) \iff \alpha f(\mathbf{v}) = f(\alpha\mathbf{v}),\ \forall f \in [f]$
2. $[f](\mathbf{v}) \oplus [f](\mathbf{w}) = [f](\mathbf{v} \oplus \mathbf{w}) \iff f(\mathbf{v}) \oplus f(\mathbf{w}) = f(\mathbf{v} \oplus \mathbf{w}),\ \forall f \in [f]$

for any $\alpha \in \overline{\mathbb{R}}_{\max}$ *and* $\mathbf{v}$, $\mathbf{w} \in \overline{\mathbb{R}}_{\max}^n$. *Let* $V = \{\mathbf{v}^1, \ldots, \mathbf{v}^p\}$ *be a finite subset of* $\overline{\mathbb{R}}_{\max}^n$. *The image of* V *under* $[f]$ *is given by*

$$[f](V) = \left\{[f](\mathbf{v}^i) \mid \mathbf{v}^i \in V, 1 \leq i \leq p\right\} \subseteq \overline{\mathbb{R}}_{\max}^n,$$

which is interpreted as

$$\{[f](\mathbf{v}^1), \ldots, [f](\mathbf{v}^p)\}.$$

The following linear map depicts the uMPL dynamics

$$f : \overline{\mathbb{R}}_{\max}^{n+p} \to \overline{\mathbb{R}}_{\max}^n \ : \ \mathbf{y}(k) \mapsto F\mathbf{y}(k), \tag{2.39}$$

where $\underline{f} \preceq f \preceq \overline{f}$, $F = (A_1 \ \tilde{B}) \in [F] = ([A] \ [\tilde{B}])$ and $\mathbf{y}(k) = (\mathbf{x}^{\mathrm{t}}(k-1), \mathbf{u}^{\mathrm{t}}(k))^{\mathrm{t}}$.

Lemma 2.1. *Let* $f : \overline{\mathbb{R}}_{\max}^n \to \overline{\mathbb{R}}_{\max}^n$ *be a linear map, where* $f \in [f] = [\underline{f}, \overline{f}]$, *as proposed in Definition 2.16 and* $\mathcal{P} = \mathsf{co}(V) = \{\bigoplus_{i=1}^p \lambda_i \mathbf{v}^i \mid \mathbf{v}^i \in V, \bigoplus_{i=1}^p \lambda_i = e\} \subseteq \overline{\mathbb{R}}_{\max}^n$ *be a* `c-Polyhedron` *for the finite subset* $V = \{\mathbf{v}^1, \ldots, \mathbf{v}^p\} \subseteq \overline{\mathbb{R}}_{\max}^n$. *Then, the direct image of* $\mathcal{P}$ *with respect to all* $f \in [f]$, *denoted*

$$\chi_f(\mathcal{P}) = \{f(\mathcal{P}) \mid f \in [f]\},$$

is equivalently expressed as

$$\chi_f(\mathcal{P}) = \mathsf{co}(\mathsf{co}(\{H_1, \ldots, H_p\}),$$

where $H_i \subseteq \overline{\mathbb{R}}_{\max}^n$ *is the generating set of the polytrope that represents each* $[f](\mathbf{v}^i) \in [f](V)$ *(cf. Definition 2.16).*

Proof. First,

$$\forall \mathbf{x} \in \mathcal{P} \iff \mathbf{x} = \bigoplus_{i=1}^p \lambda_i \mathbf{v}^i, \text{ s.t. } \bigoplus_{i=1}^p \lambda_i = e.$$

Then, $\forall f \in [f]$, we have

$$f(\mathbf{x}) \in f(\mathcal{P}) \iff f(\mathbf{x}) = f\left(\bigoplus_{i=1}^p \lambda_i \mathbf{v}^i\right), \text{ s.t. } \bigoplus_{i=1}^p \lambda_i = e.$$

with

$$f\left(\bigoplus_{i=1}^p \lambda_i \mathbf{v}^i\right) = \bigoplus_{i=1}^p f(\lambda_i \mathbf{v}^i) = \bigoplus_{i=1}^p \lambda_i f(\mathbf{v}^i),$$

where $f(\mathbf{v}^i) \in [f](\mathbf{v}^i)$. The term $[f](\mathbf{v}^i)$ is a Max-Plus hypercube represented by the polytrope $\mathtt{co}(H_i) \subseteq \overline{\mathbb{R}}^n_{\max}$ with $H_i \subseteq \overline{\mathbb{R}}^n_{\max}$ given by Equation (2.37) of Proposition 2.3. Hence, $\forall f \in [f]$, we have

$$f(\mathbf{v}^i) \in [f](\mathbf{v}^i) \iff f(\mathbf{v}^i) \in \mathtt{co}(H_i) = \left\{ \bigoplus_{j=1}^{n+1} \beta_j^i \mathbf{h}_i^j \mid \mathbf{h}_i^j \in H_i, \bigoplus_{j=1}^{n+1} \beta_j^i = e \right\}$$
$$\iff f(\mathbf{v}^i) = \bigoplus_{j=1}^{n+1} \beta_j^i \mathbf{h}_i^j, \text{ s.t. } \bigoplus_{j=1}^{n+1} \beta_j^i = e$$

and thus, with regard to $\bigoplus_{i=1}^{p} \lambda_i f(\mathbf{v}^i)$, we have

$$\bigoplus_{i=1}^{p} \lambda_i f(\mathbf{v}^i) = \bigoplus_{i=1}^{p} \lambda_i \bigoplus_{j=1}^{n+1} \beta_j^i \mathbf{h}_i^j = \bigoplus_{w=1}^{p(n+1)} \alpha_w \mathbf{g}^w,$$

where $\mathbf{g}^w \in G = \{H_1, \ldots, H_p\}$ and $\bigoplus_{w=1}^{p(n+1)} \alpha_w = e$ with $\alpha_w = \lambda_i \beta_j^i$. Finally, $\forall f \in [f]$ we have $f(\mathcal{P}) = \mathtt{co}(\mathtt{co}(G))$, which represents $\chi_f(\mathcal{P}) = \{f(\mathcal{P}) \mid f \in [f]\} \equiv \mathtt{co}(\mathtt{co}(G))$. □

The image of $\tilde{\mathcal{P}}_{k-1|k-1}$, denoted by $\mathcal{P}_{k|k-1} = [f](\tilde{\mathcal{P}}_{k-1|k-1})$, is given thanks to Lemma 2.1. Thus, consider the generating set H_i of the polytrope that represents each $[f](\tilde{\mathbf{g}}^i_{k-1|k-1}) = ([A]\ [\tilde{B}])\tilde{\mathbf{g}}^i_{k-1|k-1}$, for all $i \in \{1, \ldots, N\}$. We recall that H_i is given by Equation (2.37). Then,

$$\mathcal{P}_{k|k-1} = [f](\tilde{\mathcal{P}}_{k-1|k-1}) = \mathtt{co}(G_{k|k-1}), \text{ with } G_{k|k-1} = \{H_1, \ldots, H_N\}. \tag{2.40}$$

Furthermore, $|G_{k|k-1}| = N(n+1)$.

In order to obtain $\mathtt{cf}(\mathcal{P}_{k|k-1})$, *i.e.*, the canonical form of the `c-Polyhedron` $\mathcal{P}_{k|k-1}$, one must consider its homogenous form (see Remark 2.30). The homogenization of $\mathcal{P}_{k|k-1}$ is expressed as $\widehat{\mathcal{P}}_{k|k-1} = \mathtt{cone}(\widehat{G}_{k|k-1}) \subseteq \overline{\mathbb{R}}^{n+1}_{\max}$, where $\widehat{G}_{k|k-1} = G_{k|k-1} \times \{e\}$. Then, $\mathtt{cf}(\widehat{\mathcal{P}}_{k|k-1}) = \mathtt{cone}(\widehat{G}_{\mathtt{cf}_{k|k-1}})$ is computed thanks to Algorithm 2.1 (note that the elements of $\widehat{G}_{k|k-1}$ are not proportional to each other and line 1 of this algorithm is satisfied), and $\mathtt{cf}(\mathcal{P}_{k|k-1}) = \mathtt{co}(G_{\mathtt{cf}_{k|k-1}})$ where

$$G_{\mathtt{cf}_{k|k-1}} = \left\{ \widehat{\mathbf{g}}^i_{\mathtt{cf}_{k|k-1}} \mid \begin{pmatrix} \widehat{\mathbf{g}}^i_{\mathtt{cf}_{k|k-1}} \\ e \end{pmatrix} \in \widehat{G}_{\mathtt{cf}_{k|k-1}}, 1 \leq i \leq |\widehat{G}_{\mathtt{cf}_{k|k-1}}| \right\}.$$

The procedure is summarized by Algorithm 2.2.

Algorithm 2.2: Computing the generating set of the image of the `c-Polyhedron` $\tilde{\mathcal{P}}_{k-1|k-1}$

Data: a finite set $\tilde{G}_{k-1|k-1} = G_{k-1|k-1} \times \{\mathbf{u}(k)\} \subseteq \overline{\mathbb{R}}_{\max}^{n+p}$ and a linear map $f : \overline{\mathbb{R}}_{\max}^{n+p} \to \overline{\mathbb{R}}_{\max}^{n}$ as given by Equation (2.39);

Result: a set $G_{\mathtt{cf}_{k|k-1}} = \left\{\widehat{\mathbf{g}}^1_{\mathtt{cf}_{k|k-1}}, \ldots, \widehat{\mathbf{g}}^{|G_{\mathtt{cf}_{k|k-1}}|}_{\mathtt{cf}_{k|k-1}}\right\}$.

`/* Goal: Compute the minimal generating set of the image of the c-Polyhedron` $\tilde{\mathcal{P}}_{k-1|k-1}$ `*/`

foreach $i \in \{1, \ldots, N\}$ **do**
  Compute the hypercube $[f](\tilde{\mathbf{g}}^i_{k-1|k-1}) \leftarrow ([A]\ [\tilde{B}])\tilde{\mathbf{g}}^i_{k-1|k-1}$
  Compute H_i thanks to Equation (2.37) over $([A]\ [\tilde{B}])\tilde{\mathbf{g}}^i_{k-1|k-1}$
end
Define $G_{k|k-1} \leftarrow \{H_1, \ldots, H_N\}$ and $\widehat{G}_{k|k-1} = G_{k|k-1} \times \{e\}$
Compute $\widehat{G}_{\mathtt{cf}_{k|k-1}}$ using Algorithm 2.1
Define $G_{\mathtt{cf}_{k|k-1}} = \left\{\widehat{\mathbf{g}}^i_{\mathtt{cf}_{k|k-1}} \mid \begin{pmatrix} \widehat{\mathbf{g}}^i_{\mathtt{cf}_{k|k-1}} \\ e \end{pmatrix} \in \widehat{G}_{\mathtt{cf}_{k|k-1}}, 1 \le i \le |\widehat{G}_{\mathtt{cf}_{k|k-1}}|\right\}$
return $G_{\mathtt{cf}_{k|k-1}} = \{\widehat{\mathbf{g}}^1_{\mathtt{cf}_{k|k-1}}, \ldots, \widehat{\mathbf{g}}^{|G_{\mathtt{cf}_{k|k-1}}|}_{\mathtt{cf}_{k|k-1}}\}$

The complexity to calculate $G_{\mathtt{cf}_{k|k-1}}$ amounts to $\mathcal{O}(n^3N^2)$, since $\widehat{G}_{k|k-1}$ is a subset of $\overline{\mathbb{R}}_{\max}^{r=n+1}$ with cardinality $c = N(n+1)$, and it takes rc^2 steps to compute $\widehat{G}_{\mathtt{cf}_{k|k-1}}$ using Algorithm 2.1. Thus, $rc^2 = (n+1)(N(n+1))^2 = (n+1)^3N^2$. If $f = \underline{f} = \overline{f}$, *i.e.*, the uMPL system matrices are deprecated (representing a MPL system), then the complexity is $\mathcal{O}(nN^2)$ since the cardinality of $\widehat{G}_{k|k-1}$ is, in this case, equal to N.

Example 2.20. *Recall the autonomous system of Example 2.8, precisely*

$$\mathbf{x}(1) = A(1)\mathbf{x}(0) \text{ where } A(1) \in [A] = \begin{pmatrix} [4,6] & [3,5] \\ [3,7] & [4,5] \end{pmatrix}, \text{ for } k = 1$$

Let $\mathcal{P}_{0|0} = \mathtt{co}(G_{0|0})$ *with*

$$G_{0|0} = \left\{ \begin{pmatrix} e \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} e \\ 3 \end{pmatrix} \right\}$$

be the initial `c-Polyhedron`*, such that* $\mathbf{x}(0) \in \mathcal{P}_{0|0}$*. It is straightforward to see that this polyhedral set is a polytrope, more specifically a Max-Plus hypercube* $[\mathbf{x}](0) = ([e,1],[1,3])^{\mathrm{t}}$*,*

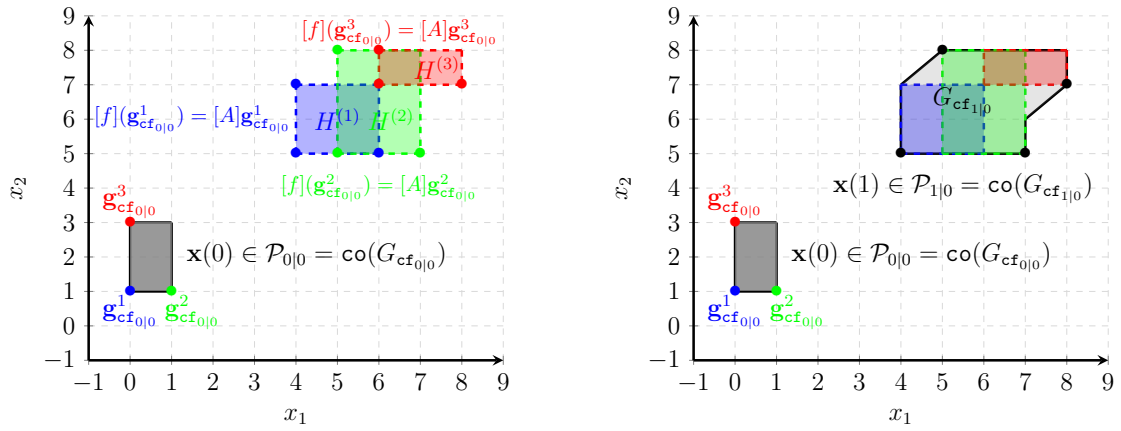


Figure 2.15 – Prior estimation set $X_{1|0}$ represented by the `c-Polyhedron` $\mathcal{P}_{1|0}$ of Example 2.20, which is computed as the minimal generating set resulting from the concatenation of $H^{(1)}$, $H^{(2)}$ and $H^{(3)}$

such that $\forall \mathbf{x}(0) \in [\mathbf{x}](0) \iff \mathbf{x}(0) \in \mathcal{P}_{0|0}$. *We define* $\mathcal{P}_{1|0}$ *as the image of* $\mathcal{P}_{0|0}$ *with respect to the linear map* $f : \mathbf{x}(0) \mapsto A(1)\mathbf{x}(0)$, *with* $\underline{f} \preceq f \preceq \overline{f}$. *Following the procedure of Algorithm 2.2, for each* $\mathbf{g}^i_{0|0}$ *of* $G_{0|0}$ *with* $i \in \{1, 2, 3\}$ *we compute* $[f](\mathbf{g}^i_{0|0}) = [A]\mathbf{g}^i_{0|0}$, *which is a Max-Plus hypercube represented by the polytrope* $\mathtt{co}(H_i)$, *where*

$$H^{(1)} = \left\{ \begin{pmatrix} 4 \\ 5 \end{pmatrix}, \begin{pmatrix} 6 \\ 5 \end{pmatrix}, \begin{pmatrix} 4 \\ 7 \end{pmatrix} \right\}, H^{(2)} = \left\{ \begin{pmatrix} 5 \\ 5 \end{pmatrix}, \begin{pmatrix} 7 \\ 5 \end{pmatrix}, \begin{pmatrix} 5 \\ 8 \end{pmatrix} \right\}, H^{(3)} = \left\{ \begin{pmatrix} 6 \\ 7 \end{pmatrix}, \begin{pmatrix} 8 \\ 7 \end{pmatrix}, \begin{pmatrix} 6 \\ 8 \end{pmatrix} \right\}.$$

Then, $G_{1|0} = \{H^{(1)}, H^{(2)}, H^{(3)}\}$ *and* $\widehat{\mathcal{P}}_{1|0} = \mathtt{cone}(\widehat{G}_{1|0})$ *where* $\widehat{G}_{1|0} = G_{1|0} \times \{\mathbf{e}\}$. *Thus,* $\mathtt{cf}(\widehat{\mathcal{P}}_{1|0})$ *yields* $\mathtt{cone}(\widehat{G}_{\mathtt{cf}_{1|0}})$ *where*

$$\widehat{G}_{\mathtt{cf}_{1|0}} = \left\{ \begin{pmatrix} 8 \\ 7 \\ e \end{pmatrix}, \begin{pmatrix} 4 \\ 5 \\ e \end{pmatrix}, \begin{pmatrix} 5 \\ 8 \\ e \end{pmatrix}, \begin{pmatrix} 7 \\ 5 \\ e \end{pmatrix} \right\}$$

(see Algorithm 2.1). Finally, $\mathcal{P}_{1|0} = \mathtt{co}(G_{\mathtt{cf}_{1|0}})$ *where*

$$G_{\mathtt{cf}_{1|0}} = \left\{ \begin{pmatrix} 8 \\ 7 \end{pmatrix}, \begin{pmatrix} 4 \\ 5 \end{pmatrix}, \begin{pmatrix} 5 \\ 8 \end{pmatrix}, \begin{pmatrix} 7 \\ 5 \end{pmatrix} \right\},$$

such that its columns are denoted by dots in Figure 2.15. As a matter of correctness, this region representation of $X_{1|0}$ *is equivalent to the DBM representation of Example 2.13.*

Set-estimation using Polyhedral-RA: Set-inversion estimation

This Section is taken from [83], and we present more details concerning the technicalities. We reinterpret, using a polyhedral approach, the computation of the inverse image of a given point (measurement), depicted by vector $\mathbf{z}(k)$, which is represented by set $\tilde{X}_{k|k}$ as proposed in Equation (2.28). First, let us recall that

$$\mathbf{x}(k) \in \tilde{X}_{k|k} \iff \underline{C}\mathbf{x}(k) \leq \mathbf{z}(k) \leq \overline{C}\mathbf{x}(k),$$

where $[C] = [\underline{C}, \overline{C}] \in \overline{\mathbb{IR}}_{\max}^{q\times n}$. Thus,

$$\tilde{X}_{k|k} = L \cap U, \text{ (cf. Equation (2.8))}, \tag{2.41}$$

where L is defined in Equation (2.9) and is equal to $\bigcap_{i=1}^{q} L_i$, with

$$L_i = \{\mathbf{x} \mid z_i(k) \leq (\overline{C}\mathbf{x})_i\} \equiv \bigcup_{j=1}^{n} \{\mathbf{x} \mid x_j \geq \overline{c}_{ij} \backslash z_i(k)\} \tag{2.42}$$

and

$$U = \{\mathbf{x} \mid \mathbf{x} \leq \underline{C} \backslash \mathbf{z}(k)\}, \tag{2.43}$$

also defined in Equation (2.10). Then,

$$\tilde{X}_{k|k} = L \cap U = \left(\bigcap_{i=1}^{q} L_i\right) \cap U = \bigcap_{i=1}^{q} L_i \cap U.$$

Lemma 2.2. *The following Max-Plus half space directly encodes L_i of Equation* (2.42) *with the same expressiveness*

$$L_i = \{\mathbf{x} \mid (z_i^{-1} \otimes \overline{c}_{i1}, \ldots, z_i^{-1} \otimes \overline{c}_{in})\mathbf{x} \geq e\}. \tag{2.44}$$

Proof. $\otimes$-multiply both sides of $z_i \leq (\overline{C}\mathbf{x})_i$ by z_i^{-1}, which yields $(z_i^{-1}\otimes\overline{c}_{i1}, \ldots, z_i^{-1}\otimes\overline{c}_{in})\mathbf{x} \geq e$. The proof is complete. □

Lemma 2.3. *The following Max-Plus half space directly encodes U with the same expressiveness*

$$U = \{\mathbf{x} \mid e \geq ((\underline{C}\backslash\mathbf{z}(k))_1^{-1}, \ldots, (\underline{C}\backslash\mathbf{z}(k))_n^{-1})\mathbf{x}\}. \tag{2.45}$$

Proof. It follows from $\{\mathbf{x} \mid \mathbf{x} \leq \underline{C}\backslash\mathbf{z}(k)\} = \bigcap_{j=1}^{n}\{\mathbf{x} \mid x_j \leq (\underline{C}\backslash\mathbf{z}(k))_j\}$, that by $\otimes$-

multiplying both sides of $x_j \leq (\underline{C} \backslash \mathbf{z}(k))_j$ by $(\underline{C} \backslash \mathbf{z}(k))_j^{-1}$ yields

$$\bigcap_{j=1}^{n} \{\mathbf{x} \mid (\underline{C} \backslash \mathbf{z}(k))_j^{-1} \otimes x_j \leq e\} \equiv \{\mathbf{x} \mid ((\underline{C} \backslash \mathbf{z}(k))_1^{-1}, \ldots, (\underline{C} \backslash \mathbf{z}(k))_n^{-1})\mathbf{x} \leq e\}.$$

The proof is complete. □

Proposition 2.4. *Let $\tilde{\mathcal{P}}_{k|k} \subseteq \overline{\mathbb{R}}_{\max}^n$ be a Max-Plus polyhedron that represents $\tilde{X}_{k|k}$. Then, $\tilde{\mathcal{P}}_{k|k}$, in its external form, is given by*

$$\tilde{\mathcal{P}}_{k|k} = \left\{ \mathbf{x} \mid \begin{pmatrix} \mathtt{diag}(\mathbf{z}^{-1}(k))\overline{C} \\ \boldsymbol{\varepsilon}^{\mathrm{t}} \end{pmatrix} \mathbf{x} \oplus \begin{pmatrix} \boldsymbol{\varepsilon} \\ e \end{pmatrix} \geq \begin{pmatrix} \mathcal{E} \\ ((\underline{C} \backslash \mathbf{z}(k))^{-1})^{\mathrm{t}} \end{pmatrix} \mathbf{x} \oplus \begin{pmatrix} \boldsymbol{e} \\ \varepsilon \end{pmatrix} \right\}, \tag{2.46}$$

(representing $s = q + 1$ single half spaces).

Proof. The proof is straightforward mixing Lemma 2.2 for $i = 1, 2, \ldots, q$ and Lemma 2.3 and does not require further details. □

Lemma 2.4. *$\tilde{\mathcal{P}}_{k|k}$ of Proposition 2.4 is a* `c-Polyhedron`*.*

Proof. The proof follows the same insight used in Corollary 2.2. First, we consider the homogenous form of $\tilde{\mathcal{P}}_{k|k}$, which is finitely generated by $\widehat{G}$. Supposing that an element of $\widehat{G}$ is a ray, then its $(n+1)$-th coordinate equal to ε and hence

$$\begin{pmatrix} \boldsymbol{\varepsilon}^{\mathrm{t}} & e \end{pmatrix} \begin{pmatrix} \mathbf{x} \\ \varepsilon \end{pmatrix} \geq \begin{pmatrix} ((\underline{C} \backslash \mathbf{z}(k))^{-1})^{\mathrm{t}} & \varepsilon \end{pmatrix} \begin{pmatrix} \mathbf{h} \\ \varepsilon \end{pmatrix} \implies \varepsilon \geq ((\underline{C} \backslash \mathbf{z}(k))^{-1})^{\mathrm{t}} \mathbf{x},$$

which is an absurd. As a consequence, all elements of $\widehat{G}$ are vertices, *i.e.*, their $(n+1)$-th coordinate is equal to e and the proof is complete. □

The external-to-internal translation of $\tilde{\mathcal{P}}_{k|k}$ is done with complexity $\mathcal{O}(N^2 s^2 n)$ where $s = q+1$ and $N \leq \mathcal{O}\left((s+n)^{\lfloor \frac{n-1}{2} \rfloor}\right)$ is the maximal number of generators of the s intermediate Max-Plus cones. If $q = n$ and $n > 0$ is an even integer, then the internal representation of $\tilde{\mathcal{P}}_{k|k}$ is computed with worst-case complexity $\mathcal{O}(n^{n+3})$. In many applications, $q \ll n$ and then the tightness condition of Remark 2.27 (precisely $n \geq 2s+1$) is likely to occur and thus, it is preferable to use the external representation of $\tilde{\mathcal{P}}_{k|k}$ for large systems.

Example 2.21. *Consider* $\mathbf{z}(1) = (5,4)^{\mathrm{t}}$ *and* $[C] = \begin{pmatrix} [1,4] & [2,3] \\ [1,2] & [e,4] \end{pmatrix}$. *Thus,* $\tilde{\mathcal{P}}_{1|1}$ *is internally represented (refer to Proposition 2.4) by* $\mathtt{co}(S) \subset \overline{\mathbb{R}}^2_{\max}$ *where* $S = \begin{pmatrix} \varepsilon & 3 & 1 & \varepsilon & 2 \\ 3 & \varepsilon & e & 2 & \varepsilon \end{pmatrix}$ *is obtained by using the double description method over the following matrix inequality*

$$\begin{pmatrix} -1 & -2 \\ -2 & e \\ \varepsilon & \varepsilon \end{pmatrix} \mathbf{x}(1) \oplus \begin{pmatrix} \varepsilon \\ \varepsilon \\ e \end{pmatrix} \geq \begin{pmatrix} \varepsilon & \varepsilon \\ \varepsilon & \varepsilon \\ -3 & -3 \end{pmatrix} \mathbf{x}(1) \oplus \begin{pmatrix} e \\ e \\ \varepsilon \end{pmatrix}.$$

The reader is invited to verify the consistency that $\forall \mathbf{x}(1) \in \tilde{\mathcal{P}}_{1|1}$ *we have* $\underline{C}\mathbf{x}(1) \leq \mathbf{z}(1) \leq \overline{C}\mathbf{x}(1)$, *as it is shown in Figure 2.16.*

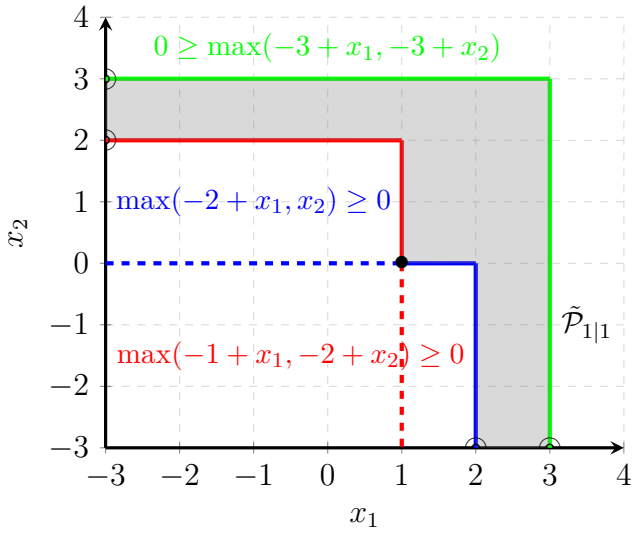


Figure 2.16 – $\tilde{\mathcal{P}}_{1|1}$ of Example 2.21

Set-estimation using Polyhedral-RA: Correction phase estimation

The set $X_{k|k}$ of Equation (2.25) is alternatively represented by the intersection of the Max-Plus polyhedra $\mathcal{P}_{k|k-1}$ (exclusively in its internal form) and $\tilde{\mathcal{P}}_{k|k}$ (exclusively in its external form), denoted by the `c-Polyhedron`[11] $\mathcal{P}_{k|k}$ that must be represented in its internal form to be used in the subsequent computation of $\mathcal{P}_{k+1|k}$ (back-shift operation:

11. The intersection of two `c-Polyhedra` is a `c-Polyhedron`.

$\mathcal{P}_{k-1|k-1} \leftarrow \mathcal{P}_{k|k}$ in the event-domain). Hence,

$$\mathbf{x}(k) \in X_{k|k} \iff \mathbf{x}(k) \in \mathcal{P}_{k|k} = \mathcal{P}_{k|k-1} \cap \tilde{\mathcal{P}}_{k|k}, \tag{2.47}$$

is computed using Remark 2.31.

> The external representation of $\mathcal{P}_{k|k-1}$ is given by $\mathcal{O}(n)$ Max-Plus half spaces. Thus, $s = \mathcal{O}(n+q)$ is the number of Max-Plus half spaces that represent $\mathcal{P}_{k|k}$ of Equation (2.47). Hence, the number of generators N of $\mathcal{P}_{k|k}$ is upper bounded by $\mathcal{O}\left((s+n)^{\lfloor \frac{n-1}{2} \rfloor}\right)$ and lower bounded by $\mathcal{O}\left((s-2n)2^{n-2}\right)$ if $s \geq 2n$ (see Remark 2.27). Assuming $q = n$, then the complexity to obtain the internal representation of $\mathcal{P}_{k|k}$ is $\mathcal{O}\left(\sigma n^3\right)$ with σ equal to
> — (best-case): 2^{2n-4} ;
> — (worst-case): n^{n-1} for n an odd integer or n^n for n an even integer

Example 2.22. *Recall the generating set of the* `c-Polyhedron` $\mathcal{P}_{1|0}$ *of Example 2.20, represented by the generating set*

$$G = \left\{ \begin{pmatrix} 8 \\ 7 \end{pmatrix}, \begin{pmatrix} 7 \\ 5 \end{pmatrix}, \begin{pmatrix} 5 \\ 8 \end{pmatrix}, \begin{pmatrix} 4 \\ 5 \end{pmatrix} \right\}.$$

Now consider the observer of Example 2.21 *but with* $\mathbf{z}(1) = C(1)\mathbf{x}(1) = (9,8)^{\mathrm{t}}$ *where* $C(1) = \begin{pmatrix} 2 & 3 \\ 2 & e \end{pmatrix}$ *and* $\mathbf{x}(1) = (6,6)^{\mathrm{t}} \in \mathcal{P}_{1|0}$*. Thus,* $\tilde{\mathcal{P}}_{1|1}$ *is externally represented by*

$$\tilde{\mathcal{P}}_{1|1} = \left\{ \mathbf{x} \mid \begin{pmatrix} -5 & -6 \\ -6 & -4 \\ \varepsilon & \varepsilon \end{pmatrix} \mathbf{x} \oplus \begin{pmatrix} \varepsilon \\ \varepsilon \\ e \end{pmatrix} \geq \begin{pmatrix} \varepsilon & \varepsilon \\ \varepsilon & \varepsilon \\ -7 & -7 \end{pmatrix} \mathbf{x} \oplus \begin{pmatrix} e \\ e \\ \varepsilon \end{pmatrix} \right\}.$$

Finally, $\mathcal{P}_{1|1}$ *is internally expressed by* $\mathcal{P}_{1|1} = \mathcal{P}_{1|0} \cap \tilde{\mathcal{P}}_{1|1}$ *using Remark 2.31. Precisely,*

$$\mathcal{P}_{1|1} = \mathtt{co}\left\{ \begin{pmatrix} 5 \\ 5 \end{pmatrix}, \begin{pmatrix} 4 \\ 6 \end{pmatrix}, \begin{pmatrix} 7 \\ 5 \end{pmatrix}, \begin{pmatrix} 4 \\ 7 \end{pmatrix} \right\}$$

as depicted in Figure 2.17.

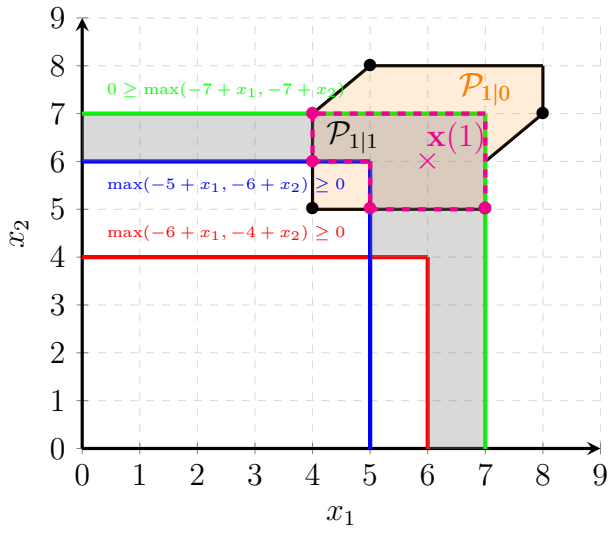


Figure 2.17 – $\mathcal{P}_{1|0}$ (orange region), $\tilde{P}_{1|1}$ (intersection of the Max-Plus half spaces in green, blue and red) and $\mathcal{P}_{1|1}$ (intersection of $\mathcal{P}_{1|0}$ and $\tilde{P}_{1|1}$ in bounded by magenta dashed lines) of Example 2.22

Set-estimation using Polyhedral-RA: Correction phase estimation with $X_{k-1|k-1} = \{\mathbf{x}(k-1)\}$

This Section is also taken from [83]. The basic idea is to assume that $X_{k-1|k-1}$ is as a singleton, precisely $X_{k-1|k-1} = \{\mathbf{x}(k-1)\}$. Thus, based on Remark 2.17, we know that $X_{k|k-1}$ is therefore a Max-Plus hypercube $[\mathbf{x}](k|k-1)$ that is also represented by the polytrope generated by the intersection of $n+1$ half spaces or equivalently by the convex hull of $n+1$ points (vertices).

Proposition 2.5. *Let $\mathcal{P}_{k|k} \subseteq \overline{\mathbb{R}}^n_{\max}$ be a* `c-Polyhedron` *that represents $X_{k|k}$. Then, $\mathcal{P}_{k|k}$, in its external form, is given by*

$$\mathcal{P}_{k|k} = \left\{ \mathbf{x} \mid \begin{pmatrix} \mathtt{diag}(\underline{\mathbf{x}}^{-1}(k|k-1)) \\ \mathtt{diag}(\mathbf{z}^{-1}(k))\overline{C} \\ \boldsymbol{\varepsilon}^{\mathrm{t}} \end{pmatrix} \mathbf{x} \oplus \begin{pmatrix} \boldsymbol{\varepsilon} \\ \boldsymbol{\varepsilon} \\ e \end{pmatrix} \geq \begin{pmatrix} \mathcal{E} \\ \mathcal{E} \\ ((\overline{\mathbf{x}}(k|k-1) \wedge \underline{C} \backslash \mathbf{z}(k))^{-1})^{\mathrm{t}} \end{pmatrix} \mathbf{x} \oplus \begin{pmatrix} \boldsymbol{e} \\ \boldsymbol{e} \\ \varepsilon \end{pmatrix} \right\}, \tag{2.48}$$

(representing $s = n + q + 1$ single half spaces).

Proof. The proof is straightforward using the proof of Proposition 2.3 and Lemmas 2.2

and 2.3. The term $\overline{\mathbf{x}}(k|k-1) \wedge \underline{C} \backslash \mathbf{z}(k)$ refers to

$$e \geq (\overline{\mathbf{x}}^{-1}(k|k-1))^{\mathrm{t}}\mathbf{x} \text{ and } e \geq ((\underline{C} \backslash \mathbf{z}(k))^{-1})^{\mathrm{t}}\mathbf{x} \Leftrightarrow e \geq \left((\overline{\mathbf{x}}^{-1}(k|k-1))^{\mathrm{t}} \oplus ((\underline{C} \backslash \mathbf{z}(k))^{-1})^{\mathrm{t}}\right)\mathbf{x}$$
$$e \geq ((\overline{\mathbf{x}}(k|k-1) \wedge \underline{C} \backslash \mathbf{z}(k))^{-1})^{\mathrm{t}}\mathbf{x}$$

□

Corollary 2.3. *$\mathcal{P}_{k|k}$ of Proposition 2.5 is a `c-Polyhedron`.*

Proof. The proof is similar to the proof of Lemma 2.4 and does not require further details. □

In this case, the external-to-internal translation of $\mathcal{P}_{k|k}$ has also exponential complexity in worst-case scenarios but rarely expected to be attained since the condition for tightness in Remark 2.27, precisely $s = n+q+1 \leq \frac{n-1}{2}$, never holds. Moreover, usually $q < n$ and then the condition $s = n + q + 1 \geq 2n$ of the same remark to obtain a best-case complexity (also costly) holds if and only if $n - q \leq 1$.

Example 2.23. *Consider the uMPL system given below:*

$$\begin{aligned} \mathbf{x}(1) &= A(1)\mathbf{x}(0),\ A(1) \in \begin{pmatrix} [1,3] & [e,4] \\ [2,4] & 2.5 \end{pmatrix}, \\ \mathbf{z}(1) &= C(1)\mathbf{x}(1),\ C(1) \in \begin{pmatrix} 0.5 & [e,1] \end{pmatrix}, \end{aligned}$$

with $\mathbf{x}(0) = (e,e)^{\mathrm{t}}$, $\mathbf{x}(1) = (3.3, 3.3)^{\mathrm{t}}$, $\mathbf{z}(1) = (0.5\ 0.5)\mathbf{x}(1) = 3.8$. *Then,* $\mathcal{P}_{1|1}$ *is externally represented by*

$$\mathcal{P}_{1|1} = \left\{ \mathbf{x} \mid \begin{pmatrix} -1 & \varepsilon \\ \varepsilon & -2.5 \\ -3.3 & -2.8 \\ \varepsilon & \varepsilon \end{pmatrix} \mathbf{x} \oplus \begin{pmatrix} \varepsilon \\ \varepsilon \\ \varepsilon \\ e \end{pmatrix} \geq \begin{pmatrix} \varepsilon & \varepsilon \\ \varepsilon & \varepsilon \\ \varepsilon & \varepsilon \\ -3.3 & -3.8 \end{pmatrix} \mathbf{x} \oplus \begin{pmatrix} e \\ e \\ e \\ \varepsilon \end{pmatrix} \right\},$$

or equivalently, internally represented by

$$\mathcal{P}_{1|1} = \mathsf{co}\left\{ \begin{pmatrix} 1 \\ 3.8 \end{pmatrix}, \begin{pmatrix} 3.3 \\ 2.5 \end{pmatrix}, \begin{pmatrix} 1 \\ 2.8 \end{pmatrix} \right\},$$

as depicted in Figure 2.18.

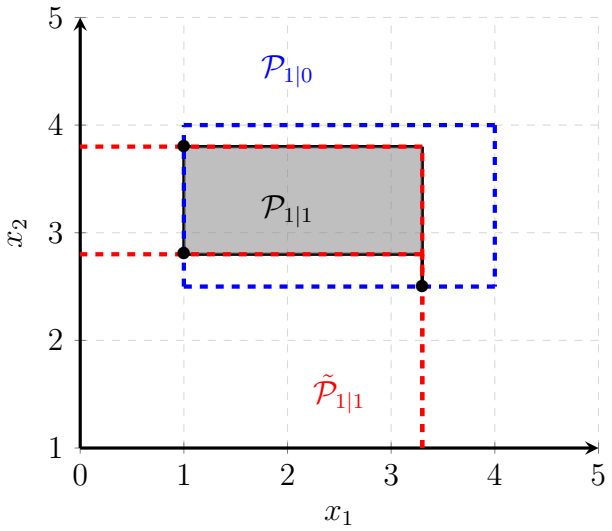


Figure 2.18 – $\mathcal{P}_{1|1}$ of Example 2.23.

2.4.3 Interval-RA

This Section is taken from [84]. The goal is to propose an approach to compute over-approximations of the image of an interval vector with respect to Equation (2.14a) and of the inverse image of a measurement point $\mathbf{z}(k)$ with respect to Equation (2.14b).

On $X_{k-1|k-1}$

Let us assume, without loss of generality, that the studied systems in this Section are autonomous, *i.e.*, $\mathbf{x}(k) = A(k)\mathbf{x}(k-1)$ with $A(k) \in [A] \in \overline{\mathbb{IR}}_{\max}^{n\times n}$. Given $X_{k-1|k-1}$ a set contained in $\overline{\mathbb{R}}_{\max}^{n}$ such that $\mathbf{x}(k-1) \in X_{k-1|k-1}$, then its exact image with respect to $[A]$ is given by Equation (2.26) as

$$X_{k|k-1} = \{A\mathbf{x} \in \overline{\mathbb{R}}_{\max}^{n} \mid \mathbf{x} \in X_{k-1|k-1}, A \in [A]\}.$$

As we already know, this computation is done either in exponential-time (worst-case scenario) using DBM approach or in cubic-time using polyhedral approach. Hence, in order to avoid this computational effort, we consider $[\mathbf{x}](k-1|k-1) \in \overline{\mathbb{IR}}_{\max}^{n}$ such that $X_{k-1|k-1} \subseteq [\mathbf{x}](k-1|k-1)$, *i.e.*, the smallest interval vector that encloses $X_{k-1|k-1}$. Thus,

set $X_{k|k-1}$ is over-approximated by the following interval vector

$$[\mathbf{x}](k|k-1) = \{\mathbf{x}(k) \in \overline{\mathbb{R}}^n_{\max} \mid \underline{A}\ \underline{\mathbf{x}}(k-1|k-1) \leq \mathbf{x}(k) \leq \overline{A}\overline{\mathbf{x}}(k-1|k-1)\}, \tag{2.50}$$

such that $X_{k|k-1} \subseteq [\mathbf{x}](k|k-1)$. Thus, this computation is done with complexity $\mathcal{O}(n^2)$ (see Remark 2.17).

On $\tilde{X}_{k|k}$

Let $\mathbf{z}(k) = C(k)\mathbf{x}(k)$ be the observation equation with $C(k) \in [C] \in \overline{\mathbb{IR}}^{q\times n}_{\max}$. Given $\mathbf{z}(k)$ a measurement point, then its exact inverse image with respect to $[C]$ is given by Equation (2.27) as

$$\tilde{X}_{k|k} = \{\mathbf{x}(k) \mid \underline{C}\mathbf{x}(k) \leq \mathbf{z}(k) \leq \overline{C}\mathbf{x}(k)\}.$$

This computation is also known to be done in exponential-time (worst-case scenarios) using either DBM or polyhedral approaches.

In general, the smallest interval vector $[\tilde{\mathbf{x}}](k|k)$ that encloses $\tilde{X}_{k|k}$ is given by

$$[\tilde{\mathbf{x}}](k|k) = \begin{pmatrix} [\varepsilon, (\underline{C}\backslash\!\backslash \mathbf{z}(k))_1] \\ \vdots \\ [\varepsilon, (\underline{C}\backslash\!\backslash \mathbf{z}(k))_n] \end{pmatrix}, \text{ such that } \tilde{X}_{k|k} \subseteq [\tilde{\mathbf{x}}](k|k)$$

However, if $\overline{C}$ (as consequence $\underline{C}$ as well) has only one element different of ε per row for the subset $\mathcal{I} \subseteq \{1, \ldots, q\}$ (refer to Definition 1.9) then, based on Proposition 1.3, consider the following procedure:

Procedure 2.2.

$\mathcal{S}^1 = \emptyset,\ \mathcal{S}^2 = \emptyset, \ldots, \mathcal{S}^n = \emptyset,$
for each $i \in \mathcal{I}, \exists j(i) \in \{1, \ldots, n\}$ *such that* $x_{j(i)}(k) \geq \overline{c}_{ij(i)} \backslash\!\backslash z_i(k),\ \mathcal{S}^{j(i)} = \mathcal{S}^{j(i)} \cup \{j(i)\},$

such that

$$[\tilde{\mathbf{x}}](k|k) = \begin{pmatrix} [\hat{\underline{x}}_1, (\underline{C}\backslash\!\backslash \mathbf{z}(k))_1] \\ \vdots \\ [\hat{\underline{x}}_n, (\underline{C}\backslash\!\backslash \mathbf{z}(k))_n] \end{pmatrix}, \text{ where } \hat{\underline{x}}_j = \begin{cases} \bigoplus_{w\in\mathcal{S}^j} \overline{c}_{iw} \backslash\!\backslash z_i(k) & \text{if } \mathcal{S}^j \neq \emptyset, \\ \varepsilon & \text{otherwise.} \end{cases} \tag{2.51}$$

Hence, this procedure is computed in bilinear/quadratic-time (matrix-vector multipli-

cation only).

Remark 2.38. *If $\overline{C}$ is doubly G-astic with only one element different of ε per row and column, then $\tilde{X}_{k|k}$ is exactly equal to $[\tilde{\mathbf{x}}](k|k)$ given in Equation* (2.51) *of Procedure 2.2.*

On $X_{k|k}$

The set $X_{k|k}$ is computed as the intersection of $X_{k|k-1}$ and $\tilde{X}_{k|k}$. Consider the alternative problem: intersection of $[\mathbf{x}](k|k-1)$ of Equation (2.50) and

$$\tilde{X}_{k|k} = \{\mathbf{x}(k) \mid \mathbf{x}(k) \leq \underline{C} \backslash \mathbf{z}(k)\} \cap \{\mathbf{x}(k) \mid \mathbf{z}(k) \leq \overline{C}\mathbf{x}(k)\},$$

such that

$$X_{k|k} \subseteq \chi = [\mathbf{x}](k|k-1) \cap \tilde{X}_{k|k} \tag{2.52}$$

$$\chi = \{\mathbf{x}(k) \mid \mathbf{z}(k) \leq \overline{C}\mathbf{x}(k)\} \cap \{\mathbf{x}(k) \mid \underline{\mathbf{x}}(k|k-1) \leq \mathbf{x}(k)\} \cap \{\mathbf{x}(k) \mid \mathbf{x}(k) \leq \overline{\Omega}\} \tag{2.53}$$

where $\overline{\Omega} = \overline{\mathbf{x}}(k|k-1) \wedge \underline{C} \backslash \mathbf{z}(k)$. Thus,

$\chi = \{\mathbf{x}(k) \mid \underline{\mathbf{x}}(k|k-1) \leq \mathbf{x}(k)\} \cap S$ where $S = \{\mathbf{x}(k) \mid \mathbf{z}(k) \leq \overline{C}\mathbf{x}(k)\} \cap \{\mathbf{x}(k) \mid \mathbf{x}(k) \leq \overline{\Omega}\}$

Lemma 2.5. *The term $\overline{\Omega} = \overline{\mathbf{x}}(k|k-1) \wedge \underline{C} \backslash \mathbf{z}(k)$ is also given by $\overline{\Omega} = \underline{\hat{C}} \backslash \mathbf{z}(k)$ where $\underline{\hat{C}} = \mathbf{z}(k) \not/ \overline{\Omega}$.*

Proof. From Lemma A.1 of Appendix A, the following holds: $(\mathbf{z}(k) \not/ \overline{\Omega}) \backslash \mathbf{z}(k) = \overline{\Omega}$, hence $\overline{\Omega} = \underline{\hat{C}} \backslash \mathbf{z}(k)$. □

Remark 2.39. *Based on Lemma 2.5, the set $S = \{\mathbf{x}(k) \mid \mathbf{z}(k) \leq \overline{C}\mathbf{x}(k)\} \cap \{\mathbf{x}(k) \mid \mathbf{x}(k) \leq \overline{\Omega}\}$ is equivalent to $\{\mathbf{x}(k) \mid \underline{\hat{C}}\mathbf{x}(k) \leq \mathbf{z}(k) \leq \overline{C}\mathbf{x}(k)\}$.*

Proposition 2.6. *Set S can be expressed equivalently as: $S = \{\mathbf{x}(k) \mid \underline{\hat{C}}\mathbf{x}(k) \leq \mathbf{z}(k) \leq \overline{\hat{C}}\mathbf{x}(k)\}$ with $\overline{\hat{C}}$ defined as*

$$\hat{\overline{c}}_{ij} = \begin{cases} \varepsilon & \text{if } \underline{\hat{c}}_{ij} > \overline{c}_{ij}, \\ \overline{c}_{ij} & \text{otherwise.} \end{cases} \tag{2.54}$$

for all $i \in \{1, \ldots, q\}$ and all $j \in \{1, \ldots, n\}$.

Proof. For the sake of brevity, in the proof we use the notation: $\mathbf{x} \equiv \mathbf{x}(k)$ and $\mathbf{z} \equiv \mathbf{z}(k)$.

First, we consider $(\hat{\underline{C}}\mathbf{x})_i \leq z_i \leq (\overline{C}\mathbf{x})_i$ for all $i \in \{1,\ldots,q\}$, this implies $(\hat{\underline{C}}\mathbf{x})_i \leq (\overline{C}\mathbf{x})_i$, *i.e.*,

$$\bigoplus_{k=1}^{n} \hat{\underline{c}}_{ik} \otimes x_k = \hat{\underline{c}}_{i1} \otimes x_1 \oplus \cdots \oplus \hat{\underline{c}}_{in} \otimes x_n \leq \overline{c}_{i1} \otimes x_1 \oplus \cdots \oplus \overline{c}_{in} \otimes x_n = \bigoplus_{k=1}^{n} \overline{c}_{ik} \otimes x_k.$$

Let us define $a = \hat{\underline{c}}_{ij} \otimes x_j$, $c = \overline{c}_{ij} \otimes x_j$,

$$b = \bigoplus_{\substack{k=1,\\ k\neq j}}^{n} \hat{\underline{c}}_{ik} \otimes x_k \text{ and } d = \bigoplus_{\substack{k=1,\\ k\neq j}}^{n} \overline{c}_{ik} \otimes x_k,$$

then it is straightforward to apply Lemma A.2 of Appendix A if $c = \overline{c}_{ij} \otimes x_j < a = \hat{\underline{c}}_{ij} \otimes x_j$ as shown below:

$$a \oplus b = \hat{\underline{c}}_{ij} \otimes x_j \oplus \bigoplus_{\substack{k=1,\\ k\neq j}}^{n} \hat{\underline{c}}_{ik} \otimes x_k = \bigoplus_{k=1}^{n} \hat{\underline{c}}_{ik} \otimes x_k \leq c \oplus d = \bigoplus_{k=1}^{n} \overline{c}_{ik} \otimes x_k = \bigoplus_{\substack{k=1,\\ k\neq j}}^{n} \overline{c}_{ik} \otimes x_k = \varepsilon \oplus d.$$

Furthermore, the following equivalence holds $\forall x_j,\ \overline{c}_{ij} \otimes x_j < \hat{\underline{c}}_{ij} \otimes x_j \Leftrightarrow \overline{c}_{ij} < \hat{\underline{c}}_{ij}$, hence, in S definition, matrix $\overline{C}$ can be replaced by matrix $\hat{\overline{C}}$. □

Remark 2.40. *In view of the previous Proposition 2.6, the observation part of the corresponding TEG is potentially simplified when evaluating its upper holding time bounds, i.e., some places can be neglected, without loss of information.*

> Using Procedure 2.2 for $S = \{\mathbf{x}(k) \mid \hat{\underline{C}}\mathbf{x}(k) \leq \mathbf{z}(k) \leq \hat{\overline{C}}\mathbf{x}(k)\}$ and assuming that $\hat{\overline{C}}$ has only one element different of ε for a subset of rows of $\{1,\ldots,q\}$, then it is straightforward to compute $[S]$ such that $S \subseteq [S]$. It is worth to mention that if $\hat{\overline{C}}$ is not row G-astic then $[S] = \emptyset$ and then $S = \emptyset$.

Finally, if $[S] \neq \emptyset$ then

$$[\chi] = [\max\left(\underline{\mathbf{x}}(k|k-1), \underline{S}\right), \overline{S} = \overline{\Omega}]. \tag{2.55}$$

It is easy to interpret that if $\hat{\underline{c}}_{ij} > \overline{c}_{ij}$ then $\overline{c}_{ij}$ can be neglected in the analysis of the TEG's behavior.

Example 2.24. *Let* $[\mathbf{x}](1|0) = ([0,4],[-2,1])^{\mathrm{t}}$, $\mathbf{z}(1) = (5,4)^{\mathrm{t}}$ *and* $[C] = \begin{pmatrix} [1,4] & [2,3] \\ [1,2] & [e,4] \end{pmatrix}$. *Then*

$$\begin{pmatrix} 1 & 2 \\ 1 & e \end{pmatrix} \mathbf{x}(1) \leq \mathbf{z}(1) \leq \begin{pmatrix} 4 & 3 \\ 2 & 4 \end{pmatrix} \mathbf{x}(1).$$

First, compute

$$\overline{\Omega} = \begin{pmatrix} 4 \\ 1 \end{pmatrix} \wedge \begin{pmatrix} 1 & 2 \\ 1 & e \end{pmatrix} \backslash \begin{pmatrix} 5 \\ 4 \end{pmatrix} = \begin{pmatrix} 3 \\ 1 \end{pmatrix}$$

and we obtain $U = \{\mathbf{x}(1) \mid \mathbf{x}(1) \leq (3,1)^{\mathrm{t}}\}$.

In order to compute the smallest interval that encloses χ *of Equation* (2.53), *we compute* $[S]$ *using Procedure 2.2 that makes it possible to replace* $\overline{C} = \begin{pmatrix} 4 & 3 \\ 2 & 4 \end{pmatrix}$ *with* $\hat{\overline{C}} = \begin{pmatrix} 4 & \varepsilon \\ 2 & 4 \end{pmatrix}$ *because* $\underline{\hat{C}} = \mathbf{z}(1) \not/ \overline{\Omega} = \begin{pmatrix} 5 \\ 4 \end{pmatrix} \not/ \begin{pmatrix} 3 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 & 4 \\ 1 & 3 \end{pmatrix}$ *is such that* $\underline{\hat{c}}_{12} > \overline{c}_{12}$. *Hence,* $\hat{\overline{C}}$ *is row G-astic with only one element different of* ε *for the subset of rows* $\{1\} \subset \{1,2\}$ *and then* $[S] = \begin{pmatrix} [1,3] \\ [\varepsilon,1] \end{pmatrix}$. *Finally,* $\chi \subseteq [\chi]$ *using Equation* (2.55), *precisely* $[\chi] = \begin{pmatrix} [1,3] \\ [-2,1] \end{pmatrix}$. *Figure 2.19 depicts* χ *and* $[\chi]$.

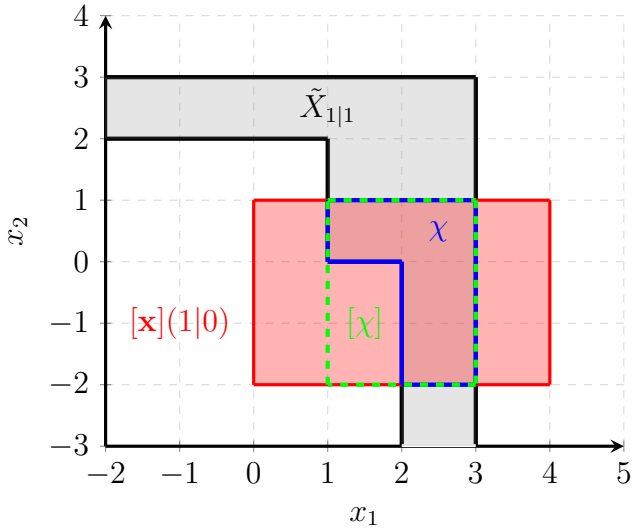


Figure 2.19 – χ and $[\chi]$ of Example 2.24 corresponds to the exact intersection between $[\mathbf{x}](1|0)$ and $\tilde{X}_{1|1}$ while $[\chi]$ is the smallest envelope such that $\chi \subseteq [\chi]$

2.5 Trajectories $\mathbf{x}(k) \in X_{k|k}$ within event-horizon

This Section summarizes the approaches studied to compute $X_{k|k}$ and their complexities in order to choose the best approach, considering the trade-off between efficiency and accuracy (risk of false-positives). An illustrative representation of the trajectories of $\mathbf{x}(k) \in X_{k|k}$ within a finite event-horizon is depicted in Figure 2.20.

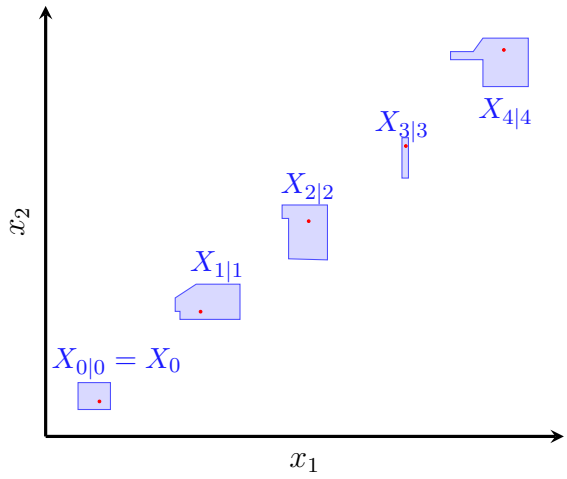


Figure 2.20 – Calculation of $X_{k|k}$ for $k = 1, 2, 3, 4$ (blue regions). The red dots represent the true state-vector $\mathbf{x}(k)$

2.5.1 Comparison table of computational complexity

The following Table 2.1 summarizes the worst-case complexity analysis of the procedures that compute $X_{k|k-1}$, $\tilde{X}_{k|k}$ and $X_{k|k}$ of Sections 2.4.1 and 2.4.2 and their over-approximations of Section 2.4.3, given a measurement point $\mathbf{z}(k)$ and $X_{k-1|k-1}$. In this analysis, we assume, without loss of generality, that the systems are autonomous and that the measurement-space is equal to the state-space ($q = n$). As a notation, we have:

— N as the number of DBM to represent $X_{k-1|k-1}$;
— N_1 and N_2 as the number of DBM to represent $X_{k|k-1}$ and $\tilde{X}_{k|k}$, respectively;
— G the number of vertices that generate the Max-Plus polyhedron $\mathcal{P}_{k-1|k-1}$ that represents $X_{k-1|k-1}$;
— α is either 2 for n an odd integer or 3 for n an even integer;
— "int. form" (internal representation) and "ext. form" (external representation) refers to the fact that int. form is computed from ext. form (which is considered as given in $\mathcal{O}(1)$ since it is only a matter of writing the inequalities);

— "approx." and "exact" refer to the fact that the Interval-RA generally yields over-approximations.

Procedure	$X_{k\|k-1}$ (Prior)	$\tilde{X}_{k\|k}$ (Measurement info.)	$X_{k\|k}$ (Posterior)
DBM-RA	$\mathcal{O}(Nn^{n+3})$	$\mathcal{O}(n^{n+2})$	$\mathcal{O}(N_1N_2n^3)$
Poly.-RA	$\mathcal{O}(G^2n^3)$ int. form	$\mathcal{O}(1)$ ext. form $\mathcal{O}(n^{n+\alpha})$ int. form	$\mathcal{O}(n^{n+\alpha})$ int. form
Int.-RA	$\mathcal{O}(n^2)$ approx.	$\mathcal{O}(n^2)$ approx. or exact	$\mathcal{O}(n^2)$ approx.

Table 2.1 – Comparison of the complexity

The calculation of $X_{k|k-1}$ using the polyhedral approach is both efficient and accurate (no risk of false-positives) while the calculation of $\tilde{X}_{k|k}$ (and therefore $X_{k|k}$) are either efficient-and-inaccurate (using interval approach) or inefficient-and-accurate (using DBM and polyhedral approaches). Nevertheless, as proposed in Remark 2.38, it is possible to compute $\tilde{X}_{k|k}$ both efficiently and accurately using the interval approach. In this case, $\tilde{X}_{k|k}$ is an interval vector or equivalently a Max-Plus hypercube generated by $n+1$ vertices (see Remark 2.37 and Proposition 2.3). Therefore, $X_{k|k}$ is yet computed in an inefficient-and-accurate way using polyhedral approach of Proposition 2.2.

In summary, the best approach combining the trade-off between efficiency and precision is the polyhedral one.

Computing a point in the estimated set $X_{k|k}$

In a closed-loop system relying on state-estimation [85], an observer-based controller is expecting an estimated vector $\mathbf{x}(k)$ and not a set $X_{k|k}$ such that $\mathbf{x}(k) \in X_{k|k}$ at each iteration k. Thus, we have to select one point in $X_{k|k}$. For instance, the estimated state can be chosen as the center [a] or as the greatest value of $X_{k|k}$. For all types of choices, the trajectory estimated by the set-estimation approach is in general less accurate than those considering the probabilistic aspects [23, 28].

a. $X_{k|k}$ is likely to be a non-convex object in conventional algebra and thus, the centroid is often not a part of the object itself

Chapter 3

Nondeterministic systems: stochastic filtering

In Chapter 2, different approaches were studied in order to be able to compute the *conditional* reach set. In this Chapter, we will show that these reach sets correspond to the support of the *posterior* conditional probability density function (PDF) $p(\mathbf{x}(k)|\{\mathbf{z}(1),\ldots,\mathbf{z}(k)\})$ in the context of Bayesian filtering and that, due to the strong nonlinearity of MPL systems, it is usually not possible to obtain an analytical form of this PDF, despite calculating its support, which makes it very difficult to define an estimator for the state-vector $\mathbf{x}(k)$.

Nevertheless, in [28, 86, 87], an original approach to alternatively solve this problem was proposed. The main idea is to drawn inspiration from the maximum likelihood estimator, but rather than looking for the value of the state $\mathbf{x}(k)$ which maximizes the likelihood function $p(\mathbf{z}(k),\mathbf{x}(k))$, we will seek, given a measurement point $\mathbf{z}(k)$, to invert the conditional mathematical expectation $\mathbb{E}[\mathbf{z}(k)|\mathbf{x}(k)]$ in order to obtain a state which "gives" the measurement, *i.e.*, a state consistent with the measurement point.

As a promising application of the stochastic filtering algorithms that will be presented in the sequel, we can cite the approach of MPL systems of queuing networks [88], the so-called *stochastic event graphs* [89], which is handled, for instance, by an algebra of unbiased gradient estimators (see [90] for details).

3.1 Nonlinear Bayesian Estimation

Bayesian methods provide a rigorous and general probabilistic framework for dynamic state estimation problems. The main idea is to recursively estimate the state-vector $\mathbf{x}(k)$ over time (k here represents the discrete-time) using the mathematical model of the dynamics and correcting it using measurement information.

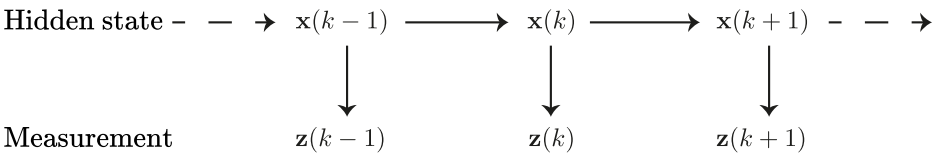


Figure 3.1 – A sequence of hidden states $\{\mathbf{x}(k)\}_{0\leq k\leq+\infty}$ is indirectly observed by a noisy sequence of measurements $\{\mathbf{z}(k)\}_{1\leq k\leq+\infty}$ of a Bayesian Network

Let

$$\begin{cases} \mathbf{x}(k) = \mathbf{f}\left(\mathbf{x}(k-1), \mathbf{w}(k)\right), & \text{(3.1a)} \\ \mathbf{z}(k) = \mathbf{h}\left(\mathbf{x}(k), \mathbf{v}(k)\right), & \text{(3.1b)} \end{cases}$$

be a *discrete nonlinear system* with $\mathbf{x} \in \mathbb{R}^n$ the state-vector and $\mathbf{z} \in \mathbb{R}^q$ the measurement-vector. We also assume that $\mathbf{w} \in \mathbb{R}^m$ and $\mathbf{v} \in \mathbb{R}^r$ are mutually independent and identically distributed (iid) noises. The nonlinear functions $\mathbf{f} : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$ and $\mathbf{h} : \mathbb{R}^n \times \mathbb{R}^r \to \mathbb{R}^q$ are the transition and observation functions, respectively.

The true state $\mathbf{x}(k)$ is assumed to be an unobserved Markov process, and the measurements points $\mathbf{z}(k)$ are the observations of a Hidden Markov Model (HMM). Thus, the term *optimal filtering*[1] is mathematically seen as a *statistical inverse problem* (or an inference or estimation problem), where the unknown quantity is a vector valued time series $\{\mathbf{x}(k)\}_{0\leq k\leq+\infty} = \{\mathbf{x}(0), \mathbf{x}(1), \mathbf{x}(2), \dots\}$, which is observed through a set of noisy measurements $\{\mathbf{z}(k)\}_{1\leq k\leq+\infty} = \{\mathbf{z}(1), \mathbf{z}(2), \dots\}$ as depicted in Figure 3.1 by the Bayesian Network.

We seek to stochastically understand Equations (3.1a) and (3.1b). For this purpose, we consider the random generation of vectors $\mathbf{x}(k)$ and $\mathbf{z}(k)$. First, let $\mathbf{x}(0)$ be a random variable distributed according to $p(\mathbf{x}(0))$, *i.e.*, $\mathbf{x}(0) \sim p(\mathbf{x}(0))$. Thus, $\mathbf{x}(0) \mapsto \mathbf{f}\left(\mathbf{x}(0), \mathbf{w}(1)\right)$ denotes how the state evolves with respect to $\mathbf{x}(0)$ and the noise $\mathbf{w}(1)$ (*stochastic diffusion*) through the nonlinear function $\mathbf{f}$, *i.e.*, $\mathbf{x}(1) \sim p(\mathbf{x}(1)|\mathbf{x}(0))$ is a new random vector conditioned to $\mathbf{x}(0)$. Extending the reasoning for the subsequent iterations and using the fact that Markov process are under the Markov property[2], one could conclude that $\mathbf{x}(k) \sim p(\mathbf{x}(k)|\mathbf{x}(k-1))$ for all k. Straightforwardly, $\mathbf{z}(k)$ is also a random vector conditioned to $\mathbf{x}(k)$, *i.e.*, $\mathbf{z}(k) \sim p(\mathbf{z}(k)|\mathbf{x}(k))$. In practice, it is possible to measure $\mathbf{z}(k)$ at each instant k and this measurement can ultimately be considered as a *sample* drawn from

1. The term optimal in this context refers to statistical optimality. Additionally, Bayesian filtering refers to the Bayesian way of formulating optimal filtering.

2. The Markov property states that the likelihood of a current state depends on the immediately previous state only.

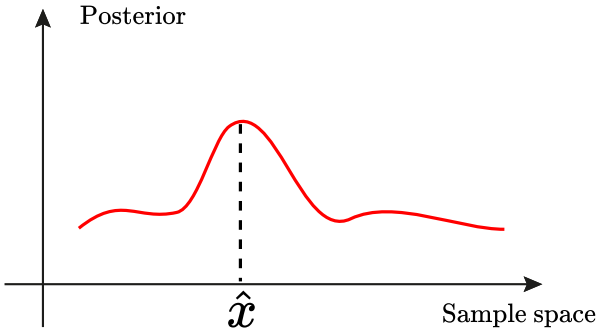


Figure 3.2 – An arbitrary posterior PDF and its estimation, which is, in this case, the maximum a-*posteriori* $\hat{x}(k|k) = \arg\ \max_{x(k)} p(x(k)|\{z(1), \ldots, z(k)\})$

$p(\mathbf{z}(k)|\mathbf{x}(k))$. Hence, the following *general probabilistic state-space model* is considered:

$$\begin{cases} \mathbf{x}(k) \sim p(\mathbf{x}(k)|\mathbf{x}(k-1)), \text{ with } \mathbf{x}(0) \sim p(\mathbf{x}(0)), \text{ (Markov process)} & \text{(3.2a)} \\ \mathbf{z}(k) \sim p(\mathbf{z}(k)|\mathbf{x}(k)), \text{ (Conditionally independent given the state).} & \text{(3.2b)} \end{cases}$$

The goal of the Bayesian approach is to compute an estimation of $\mathbf{x}(k)$ based on all available information (sequence of measurements $\mathcal{Z}(k) = \{\mathbf{z}(1), \ldots, \mathbf{z}(k)\}$) for all k. To this end, it is necessary to construct the *posterior* PDF $p(\mathbf{x}(k)|\mathcal{Z}(k))$ as depicted in Figure 3.2. Briefly, all the information needed to properly compute an optimal estimate, based on different criteria, for $\mathbf{x}(k)$ is contained in this PDF. Some examples are

— mean (expectation), mode, median, etc.

Minimum-variance unbiased estimator

A classical criterion to obtain $\mathbf{x}(k)$ at time k, using $p(\mathbf{x}(k)|\mathcal{Z}(k))$, is to consider the *minimum-variance unbiased estimator*, which gives that the optimal estimate of $\mathbf{x}(k)$ given $\mathcal{Z}(k)$, denoted as $\hat{\mathbf{x}}(k|k)$, is the mathematical expectation of the random variable $\mathbf{x}(k)$ conditioned to $\mathcal{Z}(k)$, *i.e.*,

$$\hat{\mathbf{x}}(k|k) = \mathbb{E}[\mathbf{x}(k)|\mathcal{Z}(k)] = \int \mathbf{x}(k) p(\mathbf{x}(k)|\mathcal{Z}(k)) d\mathbf{x}(k), \qquad (3.3)$$

(see [91, Th. 3.1] for more details).

The Recursive Bayesian filter allows calculating the posterior PDF and then an optimal

estimate to $\mathbf{x}(k)$. In the sequel, we briefly summarize this framework. Given:

— The probabilistic form of the dynamic system given by Equations (3.2a) and (3.2b), with known statistics of $\mathbf{w}(k)$ and $\mathbf{v}(k)$;

— Initial *prior* PDF $p(\mathbf{x}(0)|\mathbf{z}(0))$;

— Sequence of measurement up to k, formally $\mathcal{Z}(k) = \{\mathbf{z}(1), \ldots, \mathbf{z}(k)\}$.

Compute in an *on-line* two-fold manner:

1. Prediction step (a-*priori*): $p\left(\mathbf{x}(k-1)|\mathcal{Z}(k-1)\right) \longrightarrow p\left(\mathbf{x}(k)|\mathcal{Z}(k-1)\right)$
   — by using system dynamics; inflates the support of $p\left(\mathbf{x}(k-1)|\mathcal{Z}(k-1)\right)$ due to random noise;
2. Correction (or update) step (a-*posteriori*) $p\left(\mathbf{x}(k)|\mathcal{Z}(k-1)\right) \longrightarrow p\left(\mathbf{x}(k)|\mathcal{Z}(k)\right)$
   — corrects the prediction by using the new measurement $\mathbf{z}(k)$; tightens the support of $p\left(\mathbf{x}(k)|\mathcal{Z}(k-1)\right)$.

Mathematically, step 1 is computed using the *Chapman-Kolmogorov* equation under the Markov property [3] [92, 93] as follows:

$$p(\mathbf{x}(k)|\mathcal{Z}(k-1)) = \int \underbrace{p(\mathbf{x}(k)|\mathbf{x}(k-1))}_{\text{dynamics}} \underbrace{p(\mathbf{x}(k-1)|\mathcal{Z}(k-1))}_{\text{previous posterior}} \mathrm{d}\mathbf{x}(k-1), \tag{3.4}$$

assuming $p(\mathbf{x}(k-1)|\mathcal{Z}(k-1))$ is given at time $k-1$. Step 2 uses the famous Bayes rule, which effectively allows us to obtain an *optimal* solution to this problem. Precisely

$$\begin{aligned} p(\mathbf{x}(k)|\mathcal{Z}(k)) &= p(\mathbf{x}(k)|\mathbf{z}(k), \mathcal{Z}(k-1)) \\ \boxed{p(A|B,C) = \frac{p(B|A,C)p(A|C)}{p(B|C)}} &= \frac{\overbrace{p(\mathbf{z}(k)|\mathbf{x}(k), \mathcal{Z}(k-1))}^{\mathbf{z}(k) \text{ is independent of } \mathcal{Z}(k-1)} p(\mathbf{x}(k)|\mathcal{Z}(k-1))}{p(\mathbf{z}(k)|\mathcal{Z}(k-1))} \\ \boxed{\frac{\text{"likelihood"} \times \text{"prior"}}{\text{"evidence"}}} &= \frac{\overbrace{p(\mathbf{z}(k)|\mathbf{x}(k))}^{\text{measurement model}} \; p(\mathbf{x}(k)|\mathcal{Z}(k-1))}{\underbrace{p(\mathbf{z}(k)|\mathcal{Z}(k-1))}_{\text{normalization constant}}}, \\ &\propto p(\mathbf{z}(k)|\mathbf{x}(k)) p(\mathbf{x}(k)|\mathcal{Z}(k-1)) \end{aligned} \tag{3.5}$$

where $p(\mathbf{z}(k)|\mathcal{Z}(k-1)) = \int p(\mathbf{z}(k)|\mathbf{x}(k)) p(\mathbf{x}(k)|\mathcal{Z}(k-1)) \mathrm{d}\mathbf{x}(k)$.

3. The joint distribution of $\mathbf{x}(k)$ and $\mathbf{x}(k-1)$ given $\mathcal{Z}(k-1)$ can be computed by using the Markov property as follows: $p(\mathbf{x}(k), \mathbf{x}(k-1)|\mathcal{Z}(k-1)) = p(\mathbf{x}(k)|\mathbf{x}(k-1), \mathcal{Z}(k-1)) p(\mathbf{x}(k-1)|\mathcal{Z}(k-1)) = p(\mathbf{x}(k)|\mathbf{x}(k-1)) p(\mathbf{x}(k-1)|\mathcal{Z}(k-1))$.

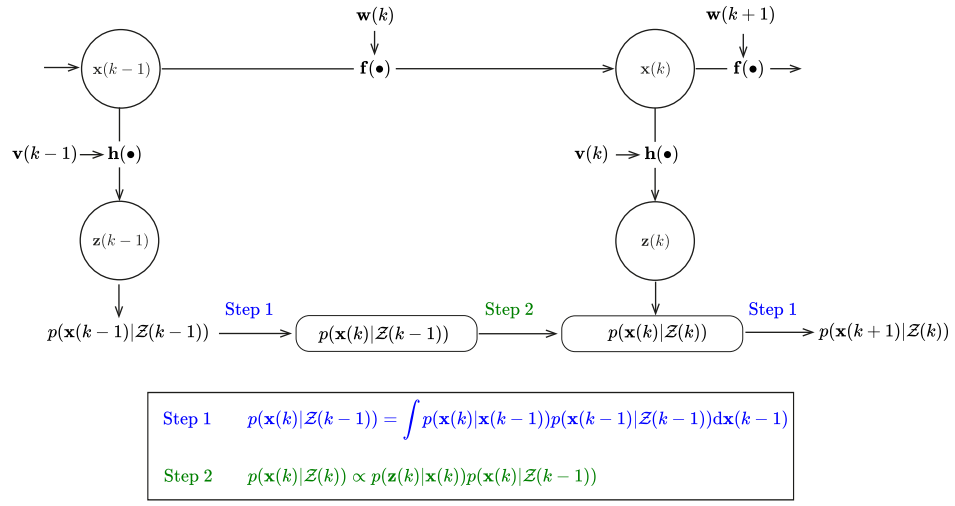


Figure 3.3 – The nonlinear system propagates the state $\mathbf{x}(k)$ through time and produces output measurements $\mathbf{z}(k)$ at each instant k. In view of this principle, the Bayesian filter tracks the PDF of $\mathbf{x}(k)$ given the measurement set $\mathcal{Z}(k)$.

Figure 3.3, adapted from [94], summarizes these two steps as the so-called Bayesian *principle*.

Although conceptually correct, the Bayesian principle is not practical since the associated integrals are intractable in most cases. Nevertheless, an optimal solution does exist for restrictive cases:

— in the case of linear systems with additive *Gaussian* noise the exact recursive solution of the problem is the very well known Kalman Filter [92, 91];

— for nonlinear problems for which the linearized model is a good approximation, the Extended Kalman filter [95] or the Unscented Kalman filter [96] can be considered.

3.2 Nondeterministic MPL systems and the Bayesian filtering

In this Section, we will formalize an "alternative" Bayesian filtering approach to handle nondeterministic MPL systems with bounded random variables (uMPL systems). First, let us study the modelling of TEG with stochastic holding times.

Example 3.1. *Consider the TEG depicted in Figure 3.4 and its model with* $x_1, x_2, x_3, a, b, c \in \mathbb{R} \cup \{-\infty\}$ *and* $k \in \mathbb{N}$.

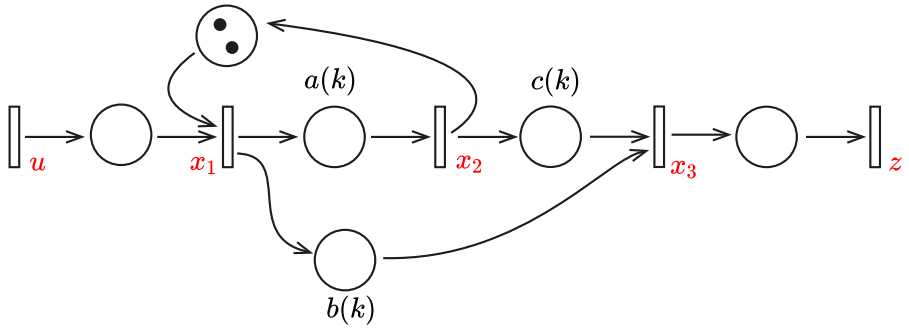


Figure 3.4 – A simple TEG with correlations in its canonical form representation

$$\begin{cases} x_1(k) &= \max\left(x_2(k-2), u(k)\right), \\ x_2(k) &= a(k) + x_1(k), \\ x_3(k) &= \max\left(b(k) + x_1(k), c(k) + x_2(k)\right), \\ z(k) &= x_3(k). \end{cases}$$

As we can see, these dynamics are nonlinear but are not of the form $\mathbf{f}(\mathbf{x}(k-1), \mathbf{w}(k))$ *and therefore the system is not a Markov process. However, by introducing a new variable* $x_4(k) = x_2(k-1)$ *and by naively substituting variables we obtain (refer to Section 1.3.2):*

$$\begin{cases} x_1(k) &= \max\left(x_4(k-1), u(k)\right) \\ x_2(k) &= a(k) + x_1(k) \\ &= \max\left(a(k) + x_4(k-1), a(k) + u(k)\right) \\ x_3(k) &= \max\left(b(k) + x_1(k), c(k) + x_2(k)\right) \\ &= \max\left(d(k) + x_4(k-1), d(k) + u(k)\right) \\ d(k) &= \max\left(b(k), c(k) + a(k)\right) \\ x_4(k) &= x_2(k-1) \\ z(k) &= x_3(k) \end{cases}$$

which is clearly of the desired form. Even though we consider that $a(k), b(k)$ *and* $c(k)$ *are mutually independent random variables distributed according to the same probability distribution over different supports, we observe, on the other hand, that* $d(k)$ *is coupled (or correlated) to* $a(k), b(k)$ *and* $c(k)$ *and it would be necessary to take into account its joint distribution with respect to these other random variables. This means that the assumption of the mutual independence of the stochastic holding times does not hold anymore. A possible conservative solution is to consider that* $d(k)$ *is a "new" independent random*

variable, i.e., uncoupled to $a(k), b(k)$ and $c(k)$.

Consider the canonical form of nondeterministic autonomous MPL systems, given by

$$\begin{cases} \mathbf{x}(k) = A(k)\mathbf{x}(k-1), \quad A(k) \in [A] \in \overline{\mathbb{IR}}_{\max}^{n\times n}, & \text{(3.6a)} \\ \mathbf{z}(k) = C(k)\mathbf{x}(k), \quad C(k) \in [C] \in \overline{\mathbb{IR}}_{\max}^{q\times n}, & \text{(3.6b)} \end{cases}$$

with vectors $\mathbf{x}$ and $\mathbf{z}$ of dimension n and q, respectively; entries of $A(k)$ and $B(k)$ are mutually independent bounded random variables distributed according to known piecewise polynomial cumulative distribution functions (CDF) [4], denoted by $F(a_{ij})$ and $F(c_{ij})$ or more simply by $F(A)$ and $F(C)$. Therefore, this canonical form is said to be Markovian (for non-autonomous systems, this statement is also valid).

However, as already pointed out, for MPL systems, the Bayesian formulation does not rely on the Kalman filter (or its extensions) due to the strong nonlinearity of the operator $\oplus = \max$ in classical algebra (lack of linearization). Then, even though we cannot directly apply the Bayesian filtering to obtain an estimate of the state $\mathbf{x}(k)$ we can still calculate the support of the posterior PDF $p(\mathbf{x}(k)|\mathcal{Z}(k))$ using the results of Chapter 2 as shown in Figure 3.5.

> In the probabilistic point of view, the disadvantage of Bayesian set-membership filtering is obvious: no estimate is computed but a region in which the state $\mathbf{x}(k)$ and its estimate $\hat{\mathbf{x}}(k|k)$ lie.

In order to overcome these unfeasible computations, an alternative filtering problem was proposed in [28, Sec. IV] and can be stated as follows: after event k, given a sequence of measurements $\mathcal{Z}(k) = \{\mathbf{z}(1), \ldots, \mathbf{z}(k)\}$, determine an estimate for $\mathbf{x}(k)$, denoted by $\hat{\mathbf{x}}(k)$, supposing that an estimate $\hat{\mathbf{x}}(0)$ is known at $k = 0$. For the unknown state trajectory $\mathcal{X}(k) = \{\mathbf{x}(0), \ldots, \mathbf{x}(k)\}$, $\hat{\mathbf{x}}(k)$ is computed as

$$\hat{\mathbf{x}}(k) = \mathbb{E}[\mathbf{x}(k)|\hat{\mathbf{x}}(k-1)]. \tag{3.7}$$

4. The CDF F_x of a random variable x is a monotonic increasing mapping $F_x : \mathbb{R} \to [0,1], t \mapsto F_x(t)$ satisfying $\lim_{t\to-\infty} F_x(t) = 0$ and $\lim_{t\to+\infty} F_x(t) = 1$. It is also equivalent to the area under the PDF $p_x(t)$ from $-\infty$ to t, *i.e.*, $F_x(t) = \int_{-\infty}^{\mathrm{t}} p_x(\tau)\mathrm{d}\tau$. Additionally, it is possible to obtain $p_x(t)$ as $p_x(t) = \frac{\mathrm{d}F_X(t)}{\mathrm{d}t}$, provided that this derivative exists.

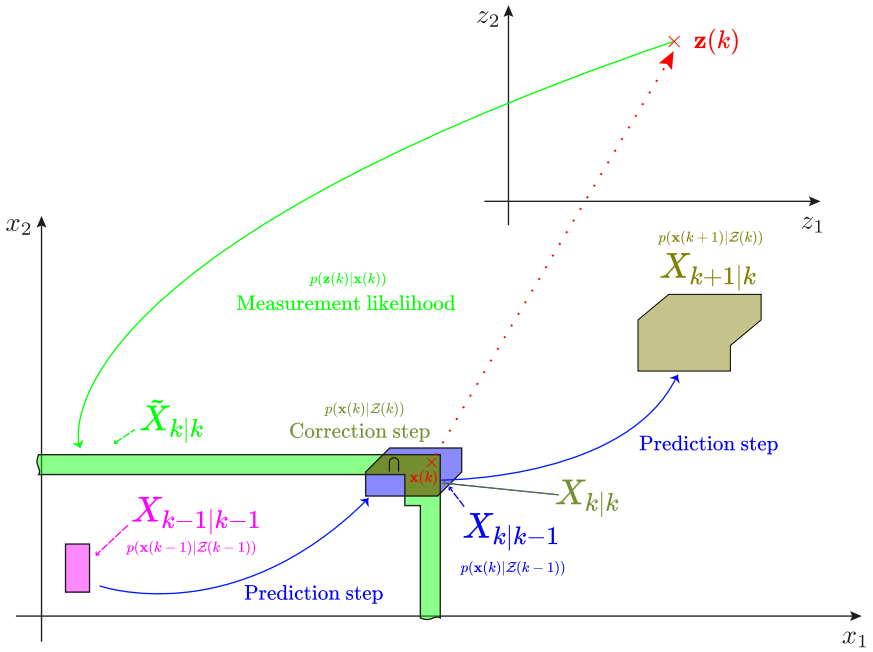


Figure 3.5 – Concatenation of the set-estimation with Bayesian principles, hereafter referred to as Bayesian set-membership filtering: from the support of $p(\mathbf{x}(k-1)|\mathcal{Z}(k-1))$, denoted $X_{k-1|k-1}$, we compute the support of the prior distribution $p(\mathbf{x}(k)|\mathcal{Z}(k-1))$, denoted $X_{k|k-1}$; the correction phase is computed by correcting $X_{k|k-1}$ thanks to the support of the measurement likelihood $p(\mathbf{z}(k)|\mathbf{x}(k))$, denoted $\tilde{X}_{k|k}$, in other words, the support of the posterior distribution $p(\mathbf{x}(k)|\mathcal{Z}(k))$ is $X_{k|k} = X_{k|k-1} \cap \tilde{X}_{k|k}$

Similarly, from $\mathcal{Z}(k)$, we will look for the state $\hat{\mathbf{x}}(k)$ such that it leads to

$$\mathbf{z}(k) = \mathbb{E}[\mathbf{z}(k)|\hat{\mathbf{x}}(k)]. \tag{3.8}$$

A sequence $\hat{\mathcal{X}}(k) = \{\hat{\mathbf{x}}(0), \ldots, \hat{\mathbf{x}}(k)\}$, ideally tracking $\mathcal{X}(k)$, is obtained if it satisfies the Equations (3.7) and (3.8). This sequence can be seen as the classical maximum likelihood estimator in the sense that it is based on the likelihood function $p(\mathbf{z}(k)|\mathbf{x}(k))$. The main difference is that instead of considering the maximum of this function, with respect to $\mathbf{x}(k)$, the estimate $\hat{\mathbf{x}}(k)$ (see Equation (3.8)) chooses the value of $\mathbf{x}(k)$ such that $\mathbf{z}(k)$ is equal to $\mathbb{E}[\mathbf{z}(k)|\hat{\mathbf{x}}(k)]$.

It is important to note that the strategy summarized by Equations (3.7) and (3.8) effectively takes into account the prior data [5], since $\hat{\mathbf{x}}(k)$ appears in both equations. Although conceptually useful, these equations are not adequate for the direct implementation of a filter, because there is no unique solution $\hat{\mathbf{x}}(k)$ guaranteed due to the nature of the Max-Plus operations.

Alternatively, we can define $\mathbf{x}(k)$ as a value which is constrained by $\mathbf{z}(k) = \mathbb{E}[\mathbf{z}(k)|\mathbf{x}(k)]$ (even though in classical filtering this is not mandatory) and is the closest to $\hat{\mathbf{x}}(k) = \mathbb{E}[\mathbf{x}(k)|\hat{\mathbf{x}}(k-1)]$ (prior estimation). To retain this value, we use a suboptimal solver based on interval contraction that will be presented in the sequel.

Alternative filtering

In short, the approach is also two-fold and can be resumed by the following equations:

$$\textbf{Prediction} \quad \hat{\mathbf{x}}(k|k-1) = \mathbb{E}[\mathbf{x}(k)|\hat{\mathbf{x}}(k-1|k-1)]. \tag{3.9}$$

$$\begin{aligned} \textbf{Update} \quad \hat{\mathbf{x}}(k|k) &= \arg\min_{\mathbf{x}} \|\mathbf{x} - \hat{\mathbf{x}}(k|k-1)\|_\infty, && \text{(3.10a)} \\ \text{s.t. } \mathbf{z}(k) &= \mathbb{E}[\mathbf{z}(k)|\mathbf{x}], && \text{(3.10b)} \end{aligned}$$

supposing that $\hat{\mathbf{x}}(k-1|k-1)$ is known (previous posterior estimation).

5. Estimators based on $p(\mathbf{x}(k)|\mathbf{z}(k))$ are Bayesian because they consider the probability density $p(\mathbf{x}(k)|\mathbf{x}(k-1))$ (see Equations (3.4) and (3.5) and note that the dependency on past measurements $\mathcal{Z}(k-1)$ has been omitted for simplicity). On the other hand, estimators purely based on the measurement likelihood $p(\mathbf{z}(k)|\mathbf{x}(k))$ are non-Bayesian.

In the sequel, we will introduce the mathematical tools necessary to properly compute Equations (3.9) and (3.10).

3.2.1 On the mathematical conditional expectation calculation: exact and approximation approaches

Clearly, a special ingredient is the computation of expected values of max-affine expressions, for instance, the expected value of

$$z = \max_{1 \leq i \leq n} (a_i + x_i),$$

where $a_1, \ldots, a_n$ are mutually independent bounded random variables with piecewise polynomial CDF $F(a_i)$, is denoted by $\mathbb{E}[z|x_1, \ldots, x_n]$.

Let us start with the simple case of the calculation of $\mathbb{E}[v]$, where $v = \max(a_1, \ldots, a_n)$. First, we compute the CDF of v as [6]

$$F_v(t) = P[v \leq t] = P[a_1 \leq t \text{ and } a_2 \leq t \text{ and } a_n \leq t] = \prod_{i=1}^{n} P[a_i \leq t] = \prod_{i=1}^{n} F_{a_i}(t).$$

As a corollary, if each a_i is shifted by a constant x_i, *i.e.*, $v = \max(a_1 + x_1, \ldots, a_n + x_n)$, then

$$F_v(t) = \prod_{i=1}^{n} F_{a_i}(t - x_i), \tag{3.11}$$

which is a piecewise polynomial function. This expression is famous and can be found in [97].

Remark 3.1. *The computation of Equation* (3.11) *can become complex for some bounded piecewise polynomial distributions, since it is necessary to consider several possible combinations of ranges. Therefore, throughout this Chapter, we consider that a_i are uniformly distributed on $[\underline{a}_i, \overline{a}_i]$, i.e., $a_i \sim \mathcal{U}(\underline{a}_i, \overline{a}_i)$ and*

$$F_{a_i}(t) = \begin{cases} 0 & \text{if } t \leq \underline{a}_i, \\ \frac{t - \underline{a}_i}{\overline{a}_i - \underline{a}_i} & \text{if } \underline{a}_i < t \leq \overline{a}_i, \\ 1 & \text{otherwise,} \end{cases}$$

6. $P[x \leq t]$ represents the probability that the random variable x takes on a value less than or equal to t.

which allows us to directly implement this computation in a procedural way.

Let us derive the expression of the mathematical expectation. Let X be a random variable with non-negative support and with its CDF $F_X(x)$. If $\forall x \leq x_0$ we have $F_X(x) \equiv 0$ and thus

$$\mathbb{E}[X] = x_0 + \int_{x_0}^{+\infty} (1 - F_X(x))\,\mathrm{d}x. \tag{3.12}$$

A simple way to understand this result is to consider the following:
First,

$$1 - F_X(x) = P[X > x] = \int_x^{+\infty} p_x(t)\mathrm{d}t$$

and

$$\int_{x_0}^{+\infty} (1 - F_X(x))\,\mathrm{d}x = \int_{x_0}^{+\infty} P[X > t]\mathrm{d}x = \int_{x_0}^{+\infty}\int_t^{+\infty} p_X(t)\mathrm{d}t\mathrm{d}x,$$

then changing the order of integration and also the integration limits, *i.e.*,

$$\{x \geq x_0, t \geq x \iff \{t \geq x_0, x_0 \leq x \leq t,$$

we obtain

$$\int_{x_0}^{+\infty}\int_{x_0}^{\mathrm{t}} p_X(t)\mathrm{d}x\mathrm{d}t = \int_{x_0}^{+\infty} [x p_X(t)]_{x_0}^{\mathrm{t}}\,\mathrm{d}t = \int_{x_0}^{+\infty} (t - x_0) p_X(t)\mathrm{d}t,$$

which is equal to

$$\int_{x_0}^{+\infty} (t - x_0) p_X(t)\mathrm{d}t = \int_{x_0}^{+\infty} t p_X(t)\mathrm{d}t - x_0 \underbrace{\int_{x_0}^{+\infty} p_X(t)\mathrm{d}t}_{=1-F_X(x_0)=1} = \int_{x_0}^{+\infty} t p_X(t)\mathrm{d}t - x_0,$$

the term $\int_{x_0}^{+\infty} t p_X(t)\mathrm{d}t$ is the usual definition of $\mathbb{E}[X]$, *i.e.*,

$$\mathbb{E}[X] = \int_{-\infty}^{+\infty} t p_X(t)\mathrm{d}t = \int_{x_0}^{+\infty} t p_X(t)\mathrm{d}t.$$

Hence,

$$\int_{x_0}^{+\infty} (1 - F_X(x))\,\mathrm{d}x = \mathbb{E}[X] - x_0.$$

For the special case of bounded random variables, *i.e.*, $a \leq X \leq b$, Equation (3.12) is rewritten as

$$\mathbb{E}[X] = b - \int_a^b F_X(x)\mathrm{d}x. \tag{3.13}$$

Conditional expectation of $\mathbf{z} = A \otimes \mathbf{x}$ given $\mathbf{x}$

Consider the Max-Plus mapping

$$\mathbf{z} = A \otimes \mathbf{x} \iff z_i = (A \otimes \mathbf{x})_i = \max_{j=1}^{n}(a_{ij} + x_j),\ i = 1, \ldots, q$$

with vectors $\mathbf{x}$ and $\mathbf{z}$ of dimension n and q, respectively; entries of $A \in \overline{\mathbb{R}}_{\max}^{q \times n}$ are mutually independent random variables distributed according to known piecewise polynomial CDF $F_{a_{ij}}$. Supposing that $\mathbf{x}$ is a deterministic vector and $\mathbf{z}$ is a random vector then the conditional expectation of $\mathbf{z}$ given $\mathbf{x}$ is denoted $\mathbb{E}[\mathbf{z}|\mathbf{x}]$ and is exprressed as

$$\mathbb{E}[\mathbf{z}|\mathbf{x}] = \begin{pmatrix} \overline{z}_1 - \int_{\underline{z}_1}^{\overline{z}_1} \prod_{j=1}^{n} F_{a_{1j}}(t - x_j)\mathrm{d}t \\ \vdots \\ \overline{z}_q - \int_{\underline{z}_q}^{\overline{z}_q} \prod_{j=1}^{n} F_{a_{qj}}(t - x_j)\mathrm{d}t \end{pmatrix} \tag{3.14}$$

where $\underline{\mathbf{z}} = \underline{A} \otimes \mathbf{x}$ and $\overline{\mathbf{z}} = \overline{A} \otimes \mathbf{x}$. Moreover, $\mathbb{E}[z_i|\mathbf{x}]$ is clearly a continuous and isotonic function of $\mathbf{x}$: if $\mathbf{x}^a \geq \mathbf{x}^b$ then

$$\prod_{j=1}^{n} F_{a_{ij}}(t - x_j^a) \leq \prod_{j=1}^{n} F_{a_{ij}}(t - x_j^b),$$

since $F_{a_{ij}}(t - x_j^a) \leq F_{a_{ij}}(t - x_j^b)$ for $j = 1, \ldots, n$, leading to $\mathbb{E}[z_i|\mathbf{x}^a] \geq \mathbb{E}[z_i|\mathbf{x}^b]$ (see [28, Lem.1]).

Remark 3.2. *A large class of random variables can properly be approximated by piecewise polynomial CDF, including those that have no upper bound (infinite support). For MPL systems, the requirement of a finite lower bound is generally not restrictive.*

Remark 3.3 (Continuity and isotony)**.** *Any other technique to calculate $\mathbb{E}[z_i|\mathbf{x}]$, and therefore, $\mathbb{E}[\mathbf{z}|\mathbf{x}]$, can alternatively be considered in the alternative filtering approach, given by Equations* (3.9) *and* (3.10)*, as long as it keeps the properties of continuity and isotony with respect to $\mathbf{x}$ (it is the essence of the success of the next part). Particularly, we can mention the results proposed in [98, 24], which are possible alternative methods.*

In [24], it was considered an approximation based on the moments of a random variable by using the fact that

$$\max(x_1,\ldots,x_n) \leq \overbrace{\max(|x_1|,\ldots,|x_n|)}^{\|(x_1,\ldots,x_n)^{\mathrm{t}}\|_\infty} \leq \overbrace{(x_1^p+\ldots x_n^p)^{1/p}}^{\|(x_1,\ldots,x_n)^{\mathrm{t}}\|_p},$$

with equality if $p \to \infty$. Briefly, the focus is on decreasing the computational burden of the calculation of $z = \mathbb{E}[\max(x_1,\ldots,x_n)]$ where $x_1,\ldots,x_n$ are random variables. Then

$$\texttt{lower}(z) \leq z \leq \texttt{upper}(z),$$

where the lower and upper bounds [24, Sec. 3] are given below, for the special case where $x_1,\ldots,x_n$ have bounded domains $\texttt{dom}(x_1),\ldots,\texttt{dom}(x_n)$:

$$\begin{aligned}\texttt{lower}(z) &= \max(\mathbb{E}[x_1],\ldots,\mathbb{E}[x_1])\\ \texttt{upper}(z) &= \left(\sum_{i=1}^{n}\mathbb{E}\left[|x_i-L|^p\right]\right)^{\frac{1}{p}}+L,\end{aligned}$$

with

$$L = \min_{i=1}^{n}(\min(\texttt{dom}(x_i),0))$$

and exploiting the Jensen's inequalities for convex and concave functions, respectively. The approximation with respect to the upper bound tightens according to a suitable tuning of the parameter p. A proper discussion on this adjustment is detailed in [24]. For comparison purposes, the p-value can also be adjusted experimentally by increasing it while the result is improved, indeed for a larger p-value the result could deviate from the exact solution. As suggested in the work, it is recommended that p be an even integer since in this case

$$\mathbb{E}[|x_i|^p] = \mathbb{E}[x_i^p].$$

Approximation of the condition expectation

Equation (3.14) exactly computes the expectation of each $\mathbb{E}[z_i|\mathbf{x}]$. By using the approximation method of [24], with a suitable tuning of the even integer p and of L, the following holds for $i = 1, \ldots, q$:

$$\mathbb{E}[z_i|\mathbf{x}] = \epsilon + \underbrace{\left(\sum_{j=1}^{n} \mathbb{E}\left[(a_{ij} + x_j - L)^p\right]\right)^{\frac{1}{p}} + L}_{\mathbb{E}^{\approx}[z_i|\mathbf{x}]} \tag{3.15}$$

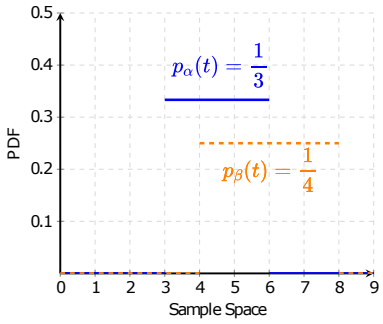
The approximation error ϵ tends to 0 if we made a good choice for the parameters. Thus, $\mathbb{E}^{\approx}[z_i|\mathbf{x}]$ is an alternative to replace $\mathbb{E}[z_i|\mathbf{x}]$ as long as it keeps properties of continuity and isotony with respect to $\mathbf{x}$. Continuity is trivial since $\phi(\boldsymbol{\nu}) = (\sum_i \mathbb{E}\left[\nu_i^p\right])^{1/p}$ with $p \neq 0$ is clearly a continuous function since $\mathbb{E}\left[\nu_i^p\right] = \int_{\mathtt{dom}(\nu_i)} (\nu_i)^p p_{\nu_i}(t)\mathrm{d}t$ is assumed to be continuous and positive. Isotony is proved in Lemma C.2 of Appendix C to show that if $x_j^a \geq x_j^b$ then $\mathbb{E}\left[(a_{ij} + x_j^a - L)^p\right] \geq \mathbb{E}\left[(a_{ij} + x_j^b - L)^p\right]$ for $j = 1, \ldots, n$ and consequently $\mathbb{E}^{\approx}[z_i|\mathbf{x}^a] \geq \mathbb{E}^{\approx}[z_i|\mathbf{x}^b]$.

Remark 3.4. *There are other upper approximations to* $\max(x_1, \ldots, x_n)$*, such as the normalized exponential function (softmax) and the LogSumExp (also seen as the sum operator of log semiring in the field of tropical analysis), which exploit Jensens' inequality and then, in theory, it is possible to get good approximations.*

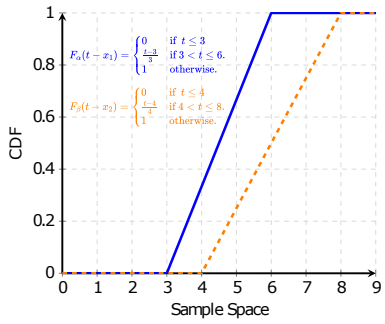
Example 3.2. *Consider a state* $x_a(k)$ *in a TEG that is the result of synchronization and delay of two other states, i.e.,* $x_a(k) = \alpha \otimes x_1(k-1) \oplus \beta \otimes x_2(k-1)$ *such that its expected value given* $x_1(k-1) = x_2(k-1) = e$ *is denoted* $\mathbb{E}[x_a(k)|(x_1(k-1), x_2(k-1))^{\mathrm{t}}]$ *and is given by*

$$\mathbb{E}[\max(\alpha, \beta)], \qquad \textit{where } \alpha \sim \mathcal{U}(3, 6) \textit{ and } \beta \sim \mathcal{U}(4, 8).$$

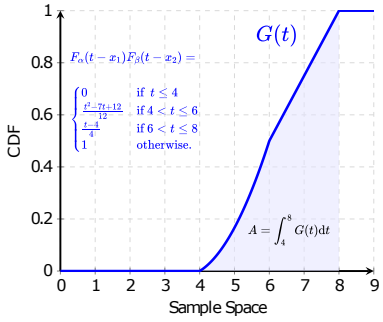
The analytical computation using Equation (3.14) *is depicted in the Figure 3.6. It should be noticed that the PDF obtained by* $\max(\alpha, \beta)$ *is not uniform, i.e., the* max *operator does not preserve the distribution (see Figure 3.7).*



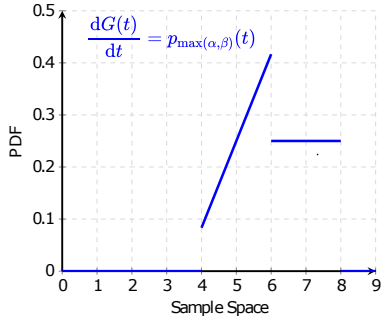
(a) PDF of α and β



(b) CDF of $\alpha + x_1$ and $\beta + x_2$



(c) Product of CDF representing $\max(\alpha+x_1, \beta+x_2)$ and the area below the resulting CDF



(d) For the sake of curiosity: the PDF obtained by differentiating the resulting PDF $G(t)$

Figure 3.6 – Computing the conditional expectation of Example 3.2 using the analytical approach - $\mathbb{E}[x_a(k)|(x_1(k-1), x_2(k-1))^t] = 8 - A = 8 - \frac{17}{9} = \frac{55}{9}$

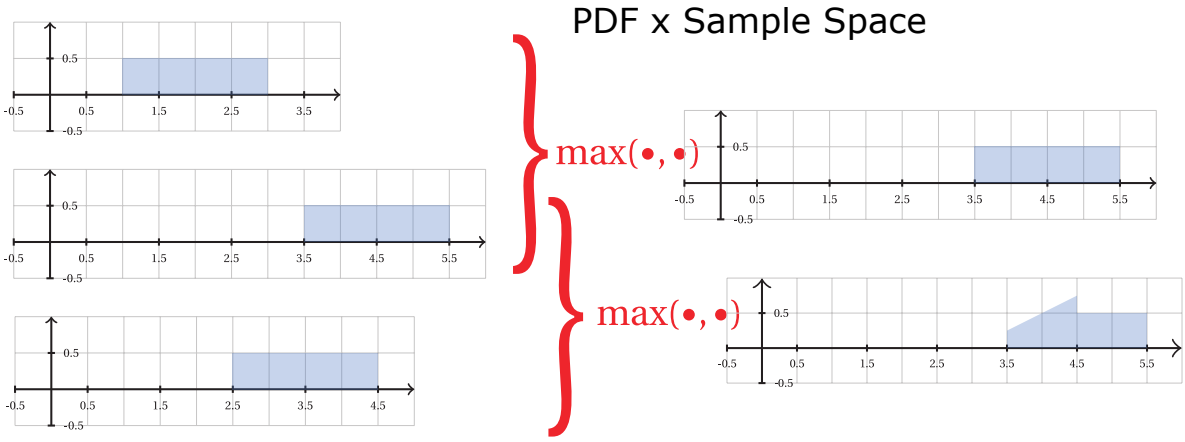


Figure 3.7 – The max operation does not preserve the distribution of its inputs, and usually $\mathbb{E}[\max(X_1, X_2)] \neq \max(\mathbb{E}[X_1], \mathbb{E}[X_2])$

This result could also be computed using Equation (3.15)*, as follows:*

$$\begin{aligned}\mathbb{E}[\max(\alpha,\beta)] &\approx (\mathbb{E}[(\alpha-L)^p]+\mathbb{E}[(\beta-L)^p])^{\frac{1}{p}}+L\\ &\approx \left(\int_{\underline{\alpha}}^{\overline{\alpha}}(t-L)^p p_\alpha(t)\,dt+\int_{\underline{\beta}}^{\overline{\beta}}(t-L)^p p_\beta(t)\,dt\right)^{\frac{1}{p}}+L\\ &\approx \left(\frac{1}{\overline{\alpha}-\underline{\alpha}}\int_{\underline{\alpha}}^{\overline{\alpha}}(t-L)^p\,dt+\frac{1}{\overline{\beta}-\underline{\beta}}\int_{\underline{\beta}}^{\overline{\beta}}(t-L)^p\,dt\right)^{\frac{1}{p}}+L\\ &\approx \left(\frac{1}{\overline{\alpha}-\underline{\alpha}}\left[\frac{(t-L)^{p+1}}{p+1}\right]_{\underline{\alpha}}^{\overline{\alpha}}+\frac{1}{\overline{\beta}-\underline{\beta}}\left[\frac{(t-L)^{p+1}}{p+1}\right]_{\underline{\beta}}^{\overline{\beta}}\right)^{\frac{1}{p}}+L,\end{aligned}$$

by arbitrarily choosing $L=3$ *and* $p=10$ *we obtain* $\mathbb{E}[\max(\alpha,\beta)]\approx 7.0247$ *to be compared with the exact value* $\frac{55}{9}$ *(other choices for the parameter* p *could potentially reduce the approximation error).*

3.2.2 On the inverse of continuous and isotonic conditional expectation functions

We shall use [28, Sec. II] in the sequel. It is assumed that the method to compute $\mathbb{E}[\mathbf{z}|\mathbf{x}]$ keeps the property of continuity and isotony (refer to [28, Lem. 3]).

> Given a q-dimensional vector $\widehat{\mathbf{z}}$, find $\mathbf{x}\in\mathbb{R}^n$ such that $\widehat{\mathbf{z}}=\mathbb{E}[\mathbf{z}|\mathbf{x}]=\mathbb{E}[C\otimes\mathbf{x}]$, with $C\in[C]\in\overline{\mathbb{IR}}_{\max}^{q\times n}$, is respected. This is also seen a *statistical inverse problem.*

Due to the multiplicity of solutions, we seek to characterize the set

$$\chi=\{\mathbf{x}\mid\widehat{\mathbf{z}}=\mathbb{E}[\mathbf{z}|\mathbf{x}]\}\subseteq(\mathbb{R}\cup\{-\infty\})^n$$

inside a given box

$$[\mathbf{x}]=\{\mathbf{x}\mid\underline{\mathbf{x}}\leq\mathbf{x}\leq\overline{\mathbf{x}}\}\subseteq(\mathbb{R}\cup\{-\infty\})^n,$$

i.e., $\chi\subseteq[\mathbf{x}]$, as close as possible to an arbitrary point $\mathbf{x}^0$ (obtained somehow).

Contractors

Contractors (see [38]) are powerful tools to efficiently solve the problem of characterization of the set χ. The operator $\mathcal{C}_\chi$ is a contractor for χ if it satisfies $\forall [\mathbf{x}] \subseteq (\mathbb{R} \cup \{-\infty\})^n$ the following properties

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

The first property states that the contraction of $[\mathbf{x}]$ always yields a contracted $[\mathbf{x}]$. The second property, combined with the definition $[\mathbf{x}] \cap \chi = \chi$, asserts that the box $\mathcal{C}_\chi([\mathbf{x}])$ does not remove a single point of χ. Moreover, a contractor is said to be *minimal* if $[\mathbf{x}] \cap \chi = \mathcal{C}_\chi([\mathbf{x}])$.

In the following, it is assumed that:
— H1) χ is not empty;
— H2) $\mathbb{E}[\mathbf{z}|\underline{\mathbf{x}}] \leq \widehat{\mathbf{z}}$;
— H3) $\forall j \in \{1, \ldots, n\}$, $\mathbb{E}[\mathbf{z}|(\underline{x}_1, \underline{x}_2, \ldots, \overline{x}_j, \ldots, \underline{x}_{n-1}, \underline{x}_n)^{\mathrm{t}})] \geq \widehat{\mathbf{z}}$.
Let us define some useful conditions.

Condition L `true` if $\mathbb{E}[z_i|\xi] < \widehat{z}_i$ with $\xi = \left(\overline{x}_1, \overline{x}_2, \ldots, \underline{x}_j, \ldots, \overline{x}_{n-1}, \overline{x}_n\right)^{\mathrm{t}}$. (3.16)

If the condition L holds, then there exists at least a point $\mathbf{x}'' = \left(\overline{x}_1, \overline{x}_2, \ldots, x''_j, \ldots, \overline{x}_{n-1}, \overline{x}_n\right)^{\mathrm{t}}$ with $\underline{x}_j \leq x''_j \leq \overline{x}_j$, such that $\mathbb{E}[z_i|\mathbf{x}''] = \widehat{z}_i$.

Similarly,

Condition U `true` if $\mathbb{E}[z_i|\eta] > \widehat{z}_i$ with $\eta = (\underline{x}_1, \underline{x}_2, \ldots, \overline{x}_j, \ldots, \underline{x}_{n-1}, \underline{x}_n)^{\mathrm{t}}$. (3.17)

If the condition L holds, then there exists at least a point $\mathbf{x}'' = \left(\underline{x}_1, \underline{x}_2, \ldots, x''_j, \ldots, \underline{x}_{n-1}, \underline{x}_n\right)^{\mathrm{t}}$, with $\underline{x}_j \leq x''_j \leq \overline{x}_j$, such that $\mathbb{E}[z_i|\mathbf{x}''] = \widehat{z}_i$.

These two conditions/consequences define two operations that both combined summarize the contractor Ω_χ on $[\mathbf{x}] = [\underline{\mathbf{x}}, \overline{\mathbf{x}}]$.

Δ_{ij}^L *operator:*

$$\Delta_{ij}^L([\underline{\mathbf{x}}, \overline{\mathbf{x}}]) = [\mathbf{x}', \overline{\mathbf{x}}] \ , \tag{3.18}$$

with

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

Δ_{ij}^U *operator:*

$$\Delta_{ij}^U([\underline{\mathbf{x}}, \overline{\mathbf{x}}]) = [\underline{\mathbf{x}}, \mathbf{x}'] \ , \tag{3.19}$$

with

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

The calculation of x'_j in Equations (3.18) and (3.19) is a one-dimensional search that can be efficiently performed by the dichotomy method (see [99]) as follows: at each step, the search interval initialized with $[\underline{x}_j, \overline{x}_j]$ is divided into two equal intervals. The half containing the solution will be the search interval at the next step. Algorithm 3.1 summarizes this search for each row/column pair (i, j) and stops when the search interval is sufficiently small.

> Let us understand a case of this procedure: for instance, in the case of the Δ_{ij}^L operator, y_j is the middle point of the search interval $[\underline{x}_j, \overline{x}_j]$ with $\mathbf{y}$ accordingly, *i.e.*, $\mathbf{y} = (\overline{x}_1, \overline{x}_2, \ldots, y_j, \ldots, \overline{x}_{n-1}, \overline{x}_n)^{\mathrm{t}}$. If $\mathbb{E}[z_i|\mathbf{x}''] \geq \widehat{z}_i$ then the solution is on the lower half (this is done by replacing the upper bound $\overline{x}_j$ of the search interval by the middle point y_j).

The composition of Δ_{ij}^L and Δ_{ij}^U for $i = 1, \ldots, q$ and $j = 1, \ldots, n$

We denote Ω_χ the contractor obtained from the iterated composition of $2qn$ operators

Algorithm 3.1: One-dimensional search

Data: $[\mathbf{x}]$, (i,j), $\widehat{z}_i$, `type` L or U, *tol* (tolerance)
Result: $[\mathbf{w},\overline{\mathbf{x}}] = \Delta_{ij}^{L}([\mathbf{x}])$ or $[\underline{\mathbf{x}},\mathbf{w}] = \Delta_{ij}^{U}([\mathbf{x}])$

$(\mathbf{y},\mathbf{w}) \leftarrow \begin{cases} (\overline{\mathbf{x}},\underline{\mathbf{x}}) & \text{if } L; \\ (\underline{\mathbf{x}},\overline{\mathbf{x}}) & \text{if U}; \end{cases}$
while $|\overline{x}_j - \underline{x}_j| > tol$ **do**
  $y_j \leftarrow (\underline{x}_j + \overline{x}_j)/2$;
  if $|\mathbb{E}[z_i|\mathbf{y}] - \widehat{z}_i| > tol$ **then**
    if $\mathbb{E}[z_i|\mathbf{y}] > \widehat{z}_i$ **then**
      $\overline{x}_j \leftarrow y_j$;
    else
      $\underline{x}_j \leftarrow y_j$;
    end
  else
    $\begin{cases} \overline{x}_j \leftarrow y_j & \text{if } L; \\ \underline{x}_j \leftarrow y_j & \text{if U}; \end{cases}$
  end
end
$w_j \leftarrow y_j$;
return $\mathbf{w}$;

defined above, *i.e.*,

$$\Omega_\chi([\mathbf{x}]) = (\Delta_{11}^{L} \circ \Delta_{11}^{U} \circ \ldots \Delta_{1n}^{L} \circ \Delta_{1n}^{U} \circ \ldots \circ \Delta_{q1}^{L} \circ \Delta_{q1}^{U} \circ \ldots \circ \Delta_{qn}^{L} \circ \Delta_{qn}^{U})([\mathbf{x}]).$$

It is worth to mention that, if the condition L is not satisfied, then $\Delta_{ij}^{L}([\mathbf{x}])$ is mathematically seen as the identity operator, *i.e.*, $\Delta_{ij}^{L}([\mathbf{x}]) = \mathrm{Id}^{L}([\mathbf{x}]) = [\mathbf{x}]$. Similarly, if the condition U is not satisfied, then $\Delta_{ij}^{U}([\mathbf{x}]) = \mathrm{Id}^{U}([\mathbf{x}]) = [\mathbf{x}]$.

Remark 3.5. *The contractor Ω_χ satisfies the contractance, completeness and monotonic*[7] *properties (see [28, Lem. 2] for proofs).*

The Algorithm 3.2 summarizes the contractor Ω_χ.

Remark 3.6. $\Omega_\chi([\mathbf{x}])$ *converges to a fixed point, i.e., to an interval $\mathcal{I}$ such that*

$$\Omega_\chi(\mathcal{I}) = \mathcal{I}.$$

Moreover, $\mathcal{I}$ contains χ (see [28, Lem. 3] for proofs).

Deprecation of $\mathcal{I}$, the suboptimal solver `Inv`:

7. Monotonic contractors verify: if $\mathbb{E}X \subset \mathbb{E}Y$ then $\mathcal{C}(\mathbb{E}X) \subset \mathcal{C}(\mathbb{E}Y)$.

Algorithm 3.2: Interval contractor

Data: $[\mathbf{x}]$, $\widehat{\mathbf{z}}$
Result: $\mathcal{I} = \Omega_\chi([\mathbf{x}])$
while $\underline{\mathbf{x}}' \neq \underline{\mathbf{x}}$ *and* $\overline{\mathbf{x}}' \neq \overline{\mathbf{x}}$ **do**
  $\underline{\mathbf{x}}' \leftarrow \underline{\mathbf{x}}$; $\overline{\mathbf{x}}' \leftarrow \overline{\mathbf{x}}$;
  foreach $i \in \{1, \ldots, q\}$ **do**
    foreach $j \in \{1, \ldots, n\}$ **do**
      `/* Lower dichotomy */`
      $\mathbf{x} \leftarrow \overline{\mathbf{x}}$; $x_j \leftarrow \underline{x}_j$;
      if $\mathbb{E}[z_i|\mathbf{x}] < \widehat{z}_i$ **then**
        $\underline{\mathbf{x}} \leftarrow \Delta^L_{ij}([\underline{\mathbf{x}}, \overline{\mathbf{x}}])$; `// a new` $\underline{x}_j$ `is computed`
      end
      `/* Upper dichotomy */`
      $\mathbf{x} \leftarrow \underline{\mathbf{x}}$; $x_j \leftarrow \overline{x}_j$;
      if $\mathbb{E}[z_i|\mathbf{x}] > \widehat{z}_i$ **then**
        $\overline{\mathbf{x}} \leftarrow \Delta^U_{ij}([\underline{\mathbf{x}}, \overline{\mathbf{x}}])$; `// a new` $\overline{x}_j$ `is computed`
      end
    end
  end
end
return $[\underline{\mathbf{x}}, \overline{\mathbf{x}}]$;

Generally, the interval $\mathcal{I} = [\underline{\mathbf{x}}^{opt}, \overline{\mathbf{x}}^{opt}]$ is not deprecated after the contraction procedure of the initial interval $[\mathbf{x}]$, then in order to obtain a point in $\mathcal{I}$ we consider the deprecation procedure described below (see [28, Sec. IV] for more details).

Procedure 3.1. *For an arbitrary $j \in \{1, \ldots, n\}$, let $a \in [\underline{x}_j, \overline{x}_j]$. Moreover, let $\chi' = \chi \cap \{x_j = a\}$, with $\{x_j = a\}$ a hyperplane in $\mathbb{R}^n$, and let $\mathcal{I}' = \mathcal{I} \cap \{x_j = a\}$. In general, $\mathcal{I}'$ is not minimal (see Remark 3.6) and the contraction algorithm must be run again to obtain the minimal interval containing χ'.*

The procedure above can be iteratively repeated until the minimal interval is reduced to one point (all components are fixed) that necessarily belongs to χ. The remaining question is: which component should be fixed at each step and to which value? To answer this, consider now that one holds a *guess value* (obtained somehow), formally $\mathbf{x}^0$. We shall look for a point $\mathbf{x} \in \chi$ that is the closest to this value, *i.e.*, $\mathbf{x}^{opt} = \arg\min_{\mathbf{x}\in\chi} \|\mathbf{x} - \mathbf{x}^0\|_\infty$. In general, $\mathbf{x}^{opt}$ is not unique, *i.e.*, multiple solutions yield the same minimum, and mostly important: an optimal value for this problem cannot be guaranteed because χ is unknown. However, following [28, Sec. IV], a suboptimal heuristic procedure, based on the deprecation method described above, is proposed to solve an alternative optimization problem,

stated as follows: as we already know, $\mathcal{I}$, such that $\chi \subset \mathcal{I}$, is the interval resulting from the Algorithm 3.2. Then, we consider the alternative minimization

$$\mathbf{x}^{subopt} = \arg\min_{\mathbf{x}\in\mathcal{I}} \|\mathbf{x} - \mathbf{x}^0\|_\infty,$$

whose optimal solution [8] is given by the line 7 of Algorithm 3.3.

Remark 3.7. *The vector $\mathbf{x}^{subopt}$ not necessarily belongs to χ, but it is useful to determine at each step which component must be deprecated and to which value. Given that, $\chi \subset \mathcal{I}$ then the following statement holds:*

$$\min_{\mathbf{x}\in\chi} \|\mathbf{x} - \mathbf{x}^0\|_\infty \geq \min_{\mathbf{x}\in\mathcal{I}} \|\mathbf{x} - \mathbf{x}^0\|_\infty.$$

The Algorithm 3.3 summarizes this procedure, and it should be noticed the generation of the initial interval $[\mathbf{x}]$ must contain at least one solution of the problem characterized by the set χ. A simple rule to guarantee this is to choose $\underline{\mathbf{x}}$ such that $\overline{C}\underline{\mathbf{x}} < \widehat{\mathbf{z}}$ and $\overline{\mathbf{x}}$ such that $\underline{C}\overline{\mathbf{x}} > \widehat{\mathbf{z}}$ are respected [9].

Let us interpret the line 9 of Algorithm 3.3. This line means that $\exists j \in \{1, \ldots, n\}$ such that $|x_j^{subopt} - x_j^0| = \left\|\mathbf{x}^{subopt} - \mathbf{x}^0\right\|_\infty$, this j is denoted $j^\sharp$. The goal of the iterative suboptimization procedure that was designed is to minimize $\left\|\mathbf{x}^{subopt} - \mathbf{x}^0\right\|_\infty$. At each iteration k we obtain $j^\sharp$, denoted $j^\sharp(k)$, and thus if we do not choose the $j^\sharp(k)$-th component of $[\mathbf{x}]$ to be deprecated then in the next iteration $k+1$ we will obtain that $j^\sharp(k+1) = j^\sharp(k)$, *i.e.*, without minimizing $\left\|\mathbf{x}^{subopt} - \mathbf{x}^0\right\|_\infty$, which is not desired, and hence, at each iteration k the index of the component that must be deprecated is $j^\sharp$.

Remark 3.8. *In [28], a proper discussion of optimality is presented. In short, it is based on analyzing $J = \|\mathbf{x}^{subopt} - \mathbf{x}^0\|_\infty$ at each iteration in Algorithm 3.3. If J remains unchanged, then an optimal value is obtained. In practice, this can be summarized by evaluating whether $j^\sharp$ is kept after each iteration since $J = |x_{j^\sharp}^{subopt} - x_{j^\sharp}^0|$, i.e., if after some iterations $j^\sharp$ no longer changes, then an optimal value is found. On the other hand, if $j^\sharp$ is not preserved then, although an optimal value could be found, we cannot assert optimality.*

8. As for $\mathbf{x}^{opt}$, there exist multiple solutions for $\mathbf{x}^{subopt}$, but $\min_{\mathbf{x}\in\mathcal{I}} \|\mathbf{x} - \mathbf{x}^0\|_\infty$ is unique.
9. H3 implies $\mathbb{E}[\mathbf{z}|\overline{\mathbf{x}}] \geq \widehat{\mathbf{z}}$ and it must always be respected, otherwise $\overline{\mathbf{x}}$ must be properly modified.

Algorithm 3.3: Suboptimal solver

Data: $\widehat{\mathbf{z}}$, $\mathbf{x}^0$
Result: $\mathbf{x} = \texttt{Inv}(\widehat{\mathbf{z}}, \mathbf{x}^0)$

generate $[\underline{\mathbf{x}}, \overline{\mathbf{x}}]$; $bool \leftarrow true$;
while *bool* **do**
  $[\underline{\mathbf{x}}, \overline{\mathbf{x}}] \leftarrow \Omega_\chi([\underline{\mathbf{x}}, \overline{\mathbf{x}}])$; // Algorithm 3.2
  $bool \leftarrow \underline{\mathbf{x}} \neq \overline{\mathbf{x}}$;
  if *bool* **then**
    foreach $j \in \{1, \ldots, n\}$ **do**

$$x_j^{subopt} \leftarrow \begin{cases} \underline{x}_j & \text{if } x_j^0 \leq \underline{x}_j \\ x_j^0 & \text{if } \underline{x}_j < x_j^0 < \overline{x}_j \\ \overline{x}_j & \text{otherwise.} \end{cases} \quad ; \quad // \ \mathbf{x}^{subopt} = \texttt{arg}\ \min_{\mathbf{x} \in [\underline{\mathbf{x}}, \overline{\mathbf{x}}]} \left\| \mathbf{x} - \mathbf{x}^0 \right\|_\infty$$

    end
    $j^\sharp \leftarrow \arg\max_{j \in \{1,\ldots,n\}} |x_j^{subopt} - x_j^0|$;
    /* Deprecation */
    $\underline{x}_{j^\sharp} \leftarrow x'_{j^\sharp}$; $\overline{x}_{j^\sharp} \leftarrow x'_{j^\sharp}$;
  end
end
return $\underline{\mathbf{x}}$ *or* $\overline{\mathbf{x}}$; // Notice that: $\underline{\mathbf{x}} = \overline{\mathbf{x}}$

3.2.3 Computational point of view of the alternative filtering

From the computational point of view, the Equations (3.9) and (3.10) can be summarized by Algorithm 3.4.

Algorithm 3.4: Alternative filter

Data: $\hat{\mathbf{x}}(k-1|k-1)$ and $\mathbf{z}(k)$
Result: $\hat{\mathbf{x}}(k|k) = \texttt{Filter}(\mathbf{z}(k), \hat{\mathbf{x}}(k-1|k-1))$

$\hat{\mathbf{x}}(k|k-1) = \mathbb{E}[\mathbf{x}(k)|\hat{\mathbf{x}}(k-1|k-1)]$; // Equation (3.14)
$\hat{\mathbf{x}}(k|k) \leftarrow \texttt{Inv}(\mathbf{z}(k), \hat{\mathbf{x}}(k|k-1))$; // Algorithm 3.3
return $\hat{\mathbf{x}}(k|k)$

Let us consider a simple graphical interpretation of Algorithm 3.4, as depicted in Figure 3.8.

It is worth to mention that this alternative filtering computes an estimate $\hat{\mathbf{x}}(k|k)$ for $\mathbf{x}(k) \in X_{k|k}$ (support of the posterior PDF) that not necessarily belongs to $X_{k|k}$.

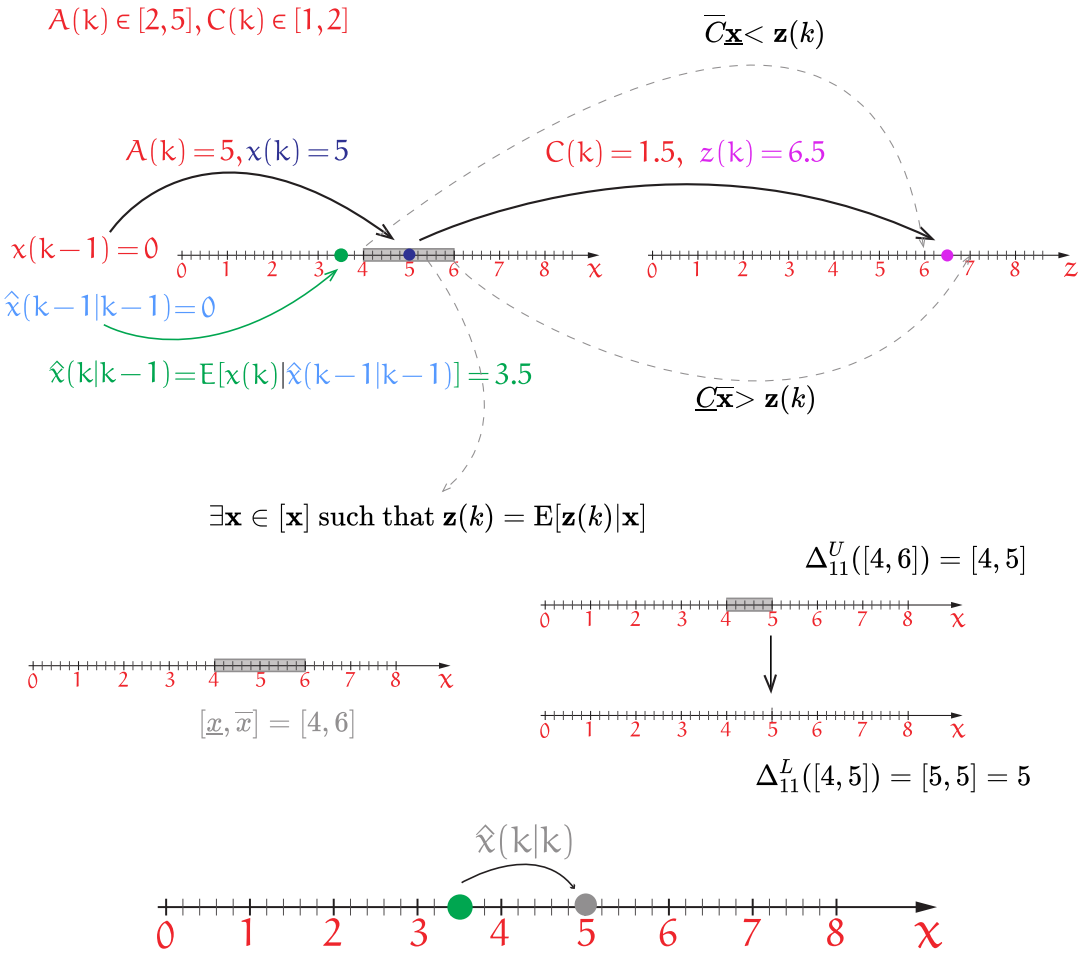


Figure 3.8 – Execution steps of Algorithm 3.4: first, a prediction value is computed using the conditional expectation calculation; the prediction value is corrected in a second step by using interval contraction of an initial interval that contains at least a $\mathbf{x}$ such that $\mathbf{z}(k) = \mathbb{E}[\mathbf{z}(k)|\mathbf{x}]$

Example 3.3. *Consider the nondeterministic MPL system with mutually independent and uniformly distributed processing times, described by*

$$\mathbf{x}(k) = A(k)\mathbf{x}(k-1), \quad \mathbf{z}(k) = C(k)\mathbf{x}(k)$$

with

$$A(k) \in [A] = \begin{pmatrix} [1,3] & [3,4] \\ [2,3] & [2,4] \end{pmatrix} \text{ and } C(k) \in [C] = \begin{pmatrix} [1,3] & [0,5] \end{pmatrix}$$

and the following initial state $\mathbf{x}(0) = (e,e)^{\mathrm{t}} \in X_0 = \{\mathbf{x} \mid 0 \leq x_1 \leq 1, 0 \leq x_2 \leq 1\}$. *Assuming that* $\hat{\mathbf{x}}(0|0) = \mathbf{x}(0) \in X_{0|0} = X_0$ *and considering the following simulated state and measurement sequences given by Table 3.1, we compute using the polyhedral set-*

	$k=0$	$k=1$	$k=2$
$\mathbf{x}(k)$	$\begin{pmatrix} e \\ e \end{pmatrix}$	$\begin{pmatrix} 3.956 \\ 2.441 \end{pmatrix}$	$\begin{pmatrix} 6.097 \\ 6.924 \end{pmatrix}$
$\mathbf{z}(k)$	–	6.537	11.158

Table 3.1 – Simulated state and measurement sequences of Example 3.3

membership method and Algorithm 3.4 the following results depicted in Figure 3.9.

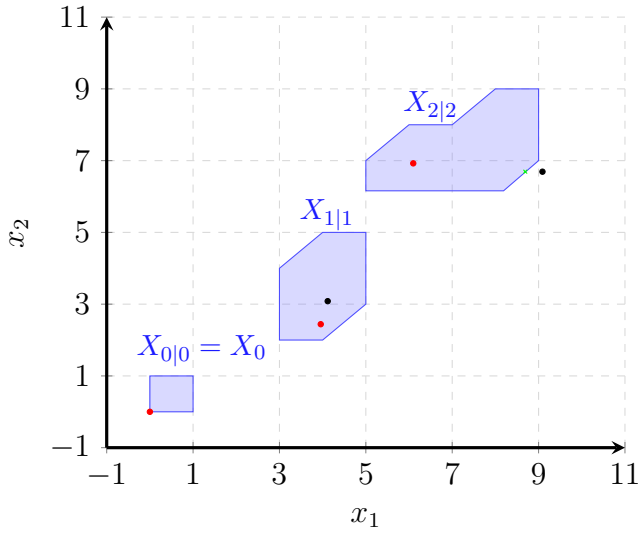


Figure 3.9 – Calculation of $X_{k|k}$ for $k = 1, 2$ (blue regions) of Example 3.3. The red dots represent the true state-vector $\mathbf{x}(k)$ and the black dots represent the estimate $\hat{\mathbf{x}}(k|k)$ using Algorithm 3.4

Thus, clearly $\hat{\mathbf{x}}(2|2)$ *does not belong to* $X_{2|2}$ *(even though* $\mathbf{z}(2) = \mathbb{E}[\mathbf{z}(2)|\mathbf{x}(2|2)]$*) thanks to Equation* (2.35)*, which denotes the membership test to the Max-Plus polyhedron that represents* $X_{2|2}$*. Briefly,* $\hat{\mathbf{x}}(2|2)$ *is not consistent with the support of the posterior PDF for* $k = 2$ *(a probabilistic issue), represented by* $X_{2|2}$*. The simple way to consistently obtain* $\hat{\mathbf{x}}(2|2)$ *belonging to* $X_{2|2}$ *is to compute its projection*[10] *over* $X_{2|2}$*, precisely* $\hat{\mathbf{x}}_{proj}(2|2) = (8.692, 6.692)^{\mathrm{t}}$ *represented by the green dot in Figure 3.9.*

3.2.4 Filtering with error criteria

The aim of this Section is to propose an improvement in the correction phase of the Algorithm 3.4. The new algorithm, originally presented in [87], is based on the minimization of a criterion that allows evaluating the estimation error of the prediction phase and to adjust the importance to be given between the prediction estimation and the corrected estimation, as the quadratic criterion does in Kalman filter for linear continuous time-driven systems.

> In Kalman filter theory, the gain is the weight given to the measurements and current-state estimate, and can be "tuned" to achieve a particular performance. If the Kalman gain is large that means error in the measurement is small which means that new data put in can now very quickly get us to the true value, and therefore we will reduce the error in the estimate and vice versa. Then, we propose to improve the original filtering algorithm for Max-Plus systems (Algorithm 3.4) by introducing a criterion in the spirit of the Kalman filtering gain. The optimal gain matrix of the Kalman filter is often derived by minimizing the trace of the posterior covariance matrix [100]. By analogy, we will define a criterion J to be minimized, allowing to take into account the trade-off between the noise in the prediction estimation and in the measurement.

The criterion J

From Equations (3.9) and (3.10) the following information is known:

10. If a homogenous point $\widehat{p} = (p, e)^{\mathrm{t}}$ does not belong to the homogenous form of a `c-Polyhedron` represented by $\widehat{\mathcal{P}} = \mathtt{co}(\widehat{G})$, where $\widehat{G}$ is its minimal generating set, then its projection, given by $\widehat{p}_{\mathrm{proj}} = \mathtt{mat}(G) \otimes (\mathtt{mat}(G) \backslash \widehat{p})$, necessarily belongs to $\widehat{\mathcal{P}}$ (see [68] for more details).

— The estimation of $\mathbf{x}(k)$ at $k-1$, formally $\hat{\mathbf{x}}(k-1|k-1)$ (prior knowledge of state);
— The current prediction estimation for the state $\mathbf{x}(k)$ from the available state estimation at $k-1$, formally $\hat{\mathbf{x}}(k|k-1) = \mathbb{E}[\mathbf{x}(k)|\hat{\mathbf{x}}(k-1|k-1)]$;
— The measurement output $\mathbf{z}(k)$ obtained at k.

For instance, let $\mathbf{z}_0$ be defined as the *output prediction estimation*, *i.e.*,

$$\mathbf{z}_0 = \mathbb{E}[\mathbf{z}(k)|\hat{\mathbf{x}}(k|k-1)].$$

Consider that, $\mathbf{z}_0$ will be used as input of the inversion procedure, formally

$$\mathbf{x}_0 = \texttt{Inv}(\mathbf{z}_0, \hat{\mathbf{x}}(k|k-1)).$$

Based on this, it is straightforward to see that $\mathbf{x}_0 = \hat{\mathbf{x}}(k|k-1)$. However, if we define $\mathbf{z}_1$ as a classical convex combination between $\mathbf{z}(k)$ and $\mathbf{z}_0$, formally

$$\mathbf{z}_1 = \beta\mathbf{z}(k) + (1-\beta)\mathbf{z}_0 \text{ where } \beta \in [0,1],$$

then this artificial measurement should be used also as input of the same inversion procedure, precisely

$$\mathbf{x}_1 = \texttt{Inv}(\mathbf{z}_1, \hat{\mathbf{x}}(k|k-1)).$$

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

$$\mathsf{J}(\mathbf{x}_1, \mathbf{z}(k)) = \max\{\|P_{dyn}^{-1}(\mathbf{x}_1 - \hat{\mathbf{x}}(k|k-1))\|_\infty, \|P_{obs}^{-1}(\mathbf{z}(k) - \mathbb{E}[\mathbf{z}(k)|\mathbf{x}_1])\|_\infty\}, \tag{3.20}$$

where $P_{dyn} = \text{diag}(\alpha_1, \ldots, \alpha_n)$ and $P_{obs} = \text{diag}(\gamma_1, \ldots, \gamma_q)$ are weighting matrices in the usual algebra. The first part of the criterion reflects how the estimation impacts the prediction estimation obtained in Equation (3.9). The second part reflects how the estimation impacts the measurement.

A key point is to ensure that the criterion above is practical. For this reason, the parameters $\alpha_1, \ldots, \alpha_n$ must be associated with the variability of the components of the vector $\hat{\mathbf{x}}(k|k-1)$. The inverse of these parameters is therefore an indicator of the reliability of $\hat{\mathbf{x}}(k|k-1)$. A rule of thumb for tuning the parameter α_i for all $i \in \{1, \ldots, n\}$ is given by considering the average width of the elements of the i-th row of the corresponding interval

matrix $[A]$. Similarly, the parameters $\gamma_1, \ldots, \gamma_q$ must be associated with the reliability of the components of $\mathbb{E}[\mathbf{z}(k)|\mathbf{x}_1]$. The variability of these variables depends on the vector $\mathbf{x}_1$, and in principle, the parameter γ_i must depend on $\mathbf{x}_1$. In order to avoid this dependence, although it is approximate, a possible rule of thumb for γ_i for all $i \in \{1, \ldots, q\}$ is given by considering the average width of the elements of the i-th row of the corresponding interval matrix $[C]$. Hence, P_{dyn} and P_{obs} are calculated in an offline phase and remain the same throughout the filtering loop. Nevertheless, it is possible to define other suitable values for α_i and γ_i, possibly online, *i.e.*, the weighting matrices could be calculated at each iteration k in order to obtain better results.

It should be noted that:

— if $\mathbf{z}_1 = \mathbf{z}_0$, then the first part of the criterion, precisely $P_{dyn}^{-1}(\mathbf{x}_1 - \hat{\mathbf{x}}(k|k-1))$ is null in the classical context;

— if $\mathbf{z}_1 = \mathbf{z}(k)$ (this is the case of Equation (3.10)), then the second part of the criterion, precisely, $P_{obs}^{-1}(\mathbf{z}(k) - \mathbb{E}[\mathbf{z}(k)|\mathbf{x}_1])$ is null, also in the classical context.

The modified correction equation is also a constrained optimization problem, as shown below, and the main idea is to choose $\hat{\mathbf{x}}(k|k)$ as the value of $\mathbf{x}_1$ that minimizes Equation (3.20).

$$\begin{aligned} &\hat{\mathbf{x}}(k|k) = \arg\min_{\mathbf{x}_1} \mathsf{J}(\mathbf{x}_1, \mathbf{z}(k)) \\ &\text{s.t. } \mathbf{z}_1 = \mathbb{E}[\mathbf{z}(k)|\mathbf{x}_1] \end{aligned} \tag{3.21}$$

Based on this, the search for $\mathbf{z}_1$ will be constrained to the convex combination of $\mathbf{z}(k)$ and $\mathbf{z}_0$. Hence, $\mathsf{J}(\mathbf{x}_1, \mathbf{z}(k))$ is reinterpreted by $\mathsf{J}(\beta)$ where $\beta \in [0, 1]$. Any one-dimensional search method can be used to find a local minimum value of the objective function above. We summarize Equation (3.9) together with (3.21) given by the following filtering Algorithm 3.5.

The resulting one-dimensional search in the Algorithm 3.5 is represented by $\beta^{opt} = \texttt{1dMinimizer}(func, 0, 1)$ and is can be solved using the `fminbnd` function, which is implemented in MATLAB. Furthermore, it is worth to mention that if the elements of the main diagonal of P_{dyn}^{-1} and P_{obs}^{-1} are $1/0$ then `fminbnd` will not run as desired, and thus we replace $1/0$ with a sufficient finite value, for instance 10^4.

Example 3.4. *Let us recall Example 3.3 with the same initial conditions and sequence of state and measurements of Table 3.1. However, instead of applying Algorithm 3.4, let us consider the application of Algorithm 3.5 but by considering the following weighting*

Algorithm 3.5: Criteria filtering algorithm for state estimation

Data: $\hat{\mathbf{x}}(k-1|k-1)$, $\mathbf{z}(k)$, P_{dyn} and P_{obs}.
Result: $\hat{\mathbf{x}}(k|k) = \texttt{Filter}(\mathbf{z}(k), \hat{\mathbf{x}}(k-1|k-1))$

1 $\hat{\mathbf{x}}(k|k-1) \leftarrow \mathbb{E}[\mathbf{x}(k)|\hat{\mathbf{x}}(k-1|k-1)]$; // Equation (3.14)

2 $\mathbf{z}_0 \leftarrow \mathbb{E}[\mathbf{z}(k)|\hat{\mathbf{x}}(k|k-1)]$; // Equation (3.14)

/* Objective function. */

3 $$\mathtt{J}(\beta) = \begin{cases} \mathbf{z}_1 \leftarrow \beta\mathbf{z}(k) + (1-\beta)\mathbf{z}_0 \\ \mathbf{x}_1 \leftarrow \mathtt{Inv}(\mathbf{z}_1, \hat{\mathbf{x}}(k|k-1)) \\ \mathtt{J}(\mathbf{x}_1, \mathbf{z}(k)) \leftarrow \max\{\|P_{dyn}^{-1}(\mathbf{x}_1 - \hat{\mathbf{x}}(k|k-1))\|_\infty, \|P_{obs}^{-1}(\mathbf{z}(k) - \mathbb{E}[\mathbf{z}(k)|\mathbf{x}_1)\|_\infty\} \end{cases} ;$$

/* One-dimensional minimizer for $\mathtt{J}(\beta)$ with β between 0 and 1 */

4 $\beta^{opt} \leftarrow \mathtt{1dMinimizer}(\mathtt{J}(\beta), 0, 1)$; // $\beta^{opt} = \arg\min_{\beta\in[0,1]} \mathtt{J}(\beta)$

5 $\mathbf{z}_1^{opt} \leftarrow \beta^{opt}\mathbf{z}(k) + (1-\beta^{opt})\mathbf{z}_0$;

6 $\mathbf{x}_1^{opt} \leftarrow \mathtt{Inv}(\mathbf{z}_1^{opt}, \hat{\mathbf{x}}(k|k-1))$;

7 **return** $\hat{\mathbf{x}}(k|k) = \mathbf{x}_1^{opt}$

matrices

$$P_{dyn}^{-1} = \begin{pmatrix} \frac{1}{1.5} & 0 \\ 0 & \frac{1}{1.5} \end{pmatrix} \text{ and } P_{dyn}^{-1} = \frac{1}{3.5}.$$

The estimated states are represented in Figure 3.10.

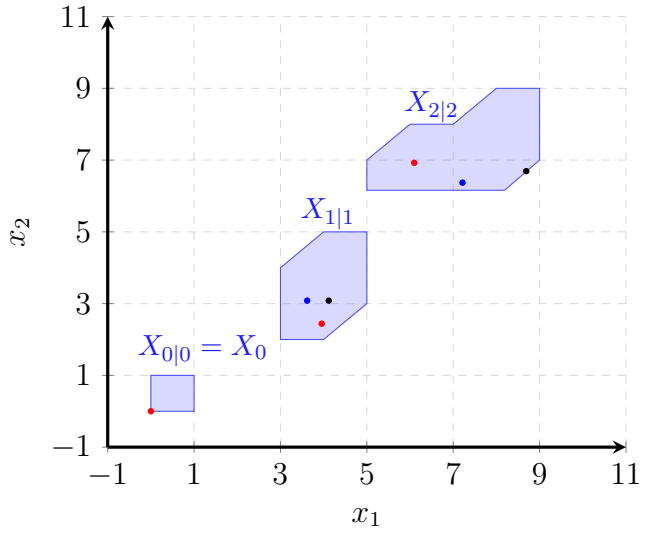


Figure 3.10 – Calculation of $X_{k|k}$ for $k = 1, 2$ (blue regions) of Examples 3.3 and 3.4. The red dots represent the true state-vector $\mathbf{x}(k)$, the black dots represent the estimate $\hat{\mathbf{x}}(k|k)$ using Algorithm 3.4 and the blue dots represent a different estimate $\hat{\mathbf{x}}'(k|k)$ using Algorithm 3.5

Thus, clearly $\hat{\mathbf{x}}'(k|k) \in X_{k|k}$ *for* $k = 1, 2$ *and all estimates are consistent with the support of the posterior PDF (it is not always the case, i.e., no guarantee). In addition, the following values of* β^{opt} *are obtained:* $k = 1,\ \beta^{opt} = 0.1753$ *and* $k = 2,\ \beta^{opt} = 0.1784$, *which means that more importance was automatically given to noise in the prediction rather than noise in the measurement (too noisy). Finally, let us analyze the average distance*

$$\frac{1}{2}\sum_{k=1}^{2} d(\hat{\mathbf{x}}(k|k), \mathbf{x}(k)) = \frac{1}{2}\sum_{k=1}^{2} \sqrt{(\hat{x}_1(k|k) - x_1(k|k))^2 + (\hat{x}_2(k|k) - x_2(k|k))^2} \approx 1.6328$$

between the estimation $\hat{\mathbf{x}}(k|k)$ *provided by Algorithm 3.4 and the true value of the state* $\mathbf{x}(k)$ *and the distance*

$$\frac{1}{2}\sum_{k=1}^{2} d(\hat{\mathbf{x}}'(k|k), \mathbf{x}(k)) = \frac{1}{2}\sum_{k=1}^{2} \sqrt{(\hat{x}'_1(k|k) - x_1(k|k))^2 + (\hat{x}'_2(k|k) - x_2(k|k))^2} \approx 0.9854$$

between the estimation provided by Algorithm 3.5 and the true state. This analysis shows that Algorithm 3.5 has slightly better results for this example.

Let us consider some other examples in the following.

Example 3.5 (Third-Order System)**.** *Consider the third-order nondeterministic MPL system with mutually independent and uniformly distributed processing times, given by*

$$\mathbf{x}(k) = A(k)\mathbf{x}(k-1), \quad \mathbf{z}(k) = C(k)\mathbf{x}(k)$$

with

$$A(k) \in [A] = \begin{pmatrix} [1,3] & [3,4] \\ [2,3] & [2,4] \end{pmatrix} \text{ and } C(k) \in [C] = \begin{pmatrix} [1,3] & [e,5] \end{pmatrix},$$

with $A(k) \in [A]$ *and* $C(k) \in [C]$ *where*

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

For instance, let us assume that $\hat{\mathbf{x}}(0|0)$ *is initialized with the initial state* $\mathbf{x}(0) = \mathbf{e}$.

The analysis of the root-mean-square-error[11] *(RMSE) between the estimation provided*

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

by Algorithm 3.4 and the true value of the state is almost equal to the RMSE between the estimation provided by Algorithm 3.5 and the true state. Table 3.2 shows the obtained results.

It is important to mention that, for this example, β^{opt} *asymptotically converges to* 1 *for most event-firings, such that* $\mathbf{z}_1 = \mathbf{z}(k)$*, clearly leading to the same result found using the Algorithm 3.4.*

State i	F_1 : Alg. 3.4 $RMSE(\{\mathbf{x}(k)\}, \{\hat{\mathbf{x}}^{F_1}(k\|k)\})_i$	F_2: Alg. 3.5 $RMSE(\{\mathbf{x}(k)\}, \{\hat{\mathbf{x}}^{F_2}(k\|k)\})_i$
1	1.9195	1.7046
2	2.0124	1.9171
3	1.8908	1.8745

Table 3.2 – Comparison between the Algorithms 3.4 and 3.5 of Example 3.5

Example 3.6 (Ninth-Order Flow Shop System)**.** *Consider the Flow Shop system modified from [22], modelled as a ninth order nondeterministic MPL system with mutually independent and uniformly distributed processing times, with three directly measured states* x_3*,* x_6*,* x_8 *and no control input (autonomous). Unfortunately, the measurement occurs under high uncertainty. The model for this system is given by*

$$\mathbf{x}(k) = A(k)\mathbf{x}(k-1), \quad \mathbf{z}(k) = C(k)\mathbf{x}(k)$$

with $A(k) = A$ *and* $C(k) \in [C]$*:*

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

For instance, let us assume that $\hat{\mathbf{x}}(0|0)$ *is initialized with the initial state* $\mathbf{x}(0) = \mathbf{e}$*.*

The model is not disturbed by any noise, i.e., it is under deterministic behavior. However, the measurement, represented by the interval matrix $[C]$*, is under high uncertainty, making all states become noisy. Algorithm 3.4 will then return a poor corrected estimation of the states, whereas Algorithm 3.5 is able to retrieve the prediction estimation instead of the corrected one as the reliable estimation of the states (*β^{opt} *asymptotically converges to* 0 *for most event-firings).*

The analysis of the RMSE between the estimation provided by Algorithm 3.4 and the true value of the state is greater than the RMSE between the estimation provided by Algorithm 3.5 and the true state. Table 3.3 shows the obtained results for the noisy states x_3*,* x_6 *and* x_8 *only.*

State i	F_1: Alg. 3.4 $RMSE(\{\mathbf{x}(k)\}, \{\hat{\mathbf{x}}^{F_1}(k\|k)\})_i$	F_2: Alg. 3.5 $RMSE(\{\mathbf{x}(k)\}, \{\hat{\mathbf{x}}^{F_2}(k\|k)\})_i$
3	1.7230	0.0065
6	1.6432	0.0065
8	1.5082	0.0065

Table 3.3 – Comparison between the Algorithms 3.4 and 3.5 of Example 3.6

Analysing the contribution: the new filtering algorithm

The new filtering strategy for the correction equation uses the same inversion procedure but using an artificial variable generated by a convex combination between output prediction and output measurement, rather than only the output measurement. This algorithm is able to deal with the trade-off between the noise in the measurement and in the prediction. Other objective functions which improve the efficiency of the proposed filtering algorithm can be considered, as well as the weighting matrices which are of utmost importance for the practical performance of the algorithm.

3.2.5 Filtering of live TEG with implicit form

The filtering Algorithms 3.4 and 3.5 assume that the stochastic entries of the system matrices are independent. Unfortunately, in most applications the associated TEG have

token-free places and due to algebraic manipulations (Kleene star), the entries of the system matrices are Max-Plus summations of processing times (for instance, $a(k) = b(k) \oplus c(k)d(k)$), in which various processing times appear in multiple entries. Therefore, the entries of the matrices are not mutual independent, and these algorithms are not directly adapted. In the sequel, we use the results originally proposed in [86].

Let us recall the model of the TEG of Example 3.1, using Equation (2.13), *i.e.*, we have $\mathbf{x}(k) = A_0(k)\mathbf{x}(k) \oplus \mathbf{y}(k-1)$, where $A_0(k)\mathbf{x}(k)$ represents the implicit part and $\mathbf{y}(k-1) = A_1(k)\mathbf{x}(k-1)$ represents the explicit part. This equation can be componentwise depicted as follows:

$$\begin{array}{rcl} x_1(k) & = & y_1(k-1), \\ x_2(k) & = & a_0^{21}(k)x_1(k) \oplus y_2(k-1), \\ x_3(k) & = & a_0^{31}(k)x_1(k) \oplus a_0^{32}(k)x_2(k) \oplus y_3(k-1), \\ & \vdots & \\ x_n(k) & = & \left(\bigoplus_{l=1}^{n-1} a_0^{nl}(k)x_l(k)\right) \oplus y_n(k-1), \end{array} \tag{3.22}$$

where each y_i is $\bigoplus_{j=1}^{n} a_1^{ij}(k) \otimes x_j(k-1)$, where each $a_1^{ij}(k)$ is assumed to be an independent random variable, and each $x_j(k-1)$ is fixed. Hence, the entries of the vector $\mathbf{y}(k)$ are assumed to be independent. Moreover, the first part of the right-hand side of Equation (3.22) is $\bigoplus_{l=1}^{i-1} a_0^{il}(k) \otimes x_l(k)$ for all $i \in \{1, \ldots, n\}$ and depends on $x_l(k)$ from $l = 1$ up to $i-1$. By assuming that the computation of Equation (3.22) is done from $i = 1$ up to n and the terms $x_l(k)$ are fixed and known, we will be able to ensure the independence of the components of the vector $\mathbf{x}(k)$, which is worth of interest in the main contribution of this Section [12].

Nevertheless, the stochastic systems described by the equation (3.22) have dependencies between their matrix entries. Hence, to properly calculate the expectation of the i-th component of $\mathbf{x}(k)$ it would be necessary to take into account its joint distribution with respect to the other components, which seems to be an intractable problem for most of the cases. From this fact, the stochastic filtering schemes that were proposed in Algorithms 3.4 and 3.5 are no longer allowed to be straightforwardly used if we are not interested in conservative results [13].

12. We consider that each $x_i(k)$ is recursively known at each subsequent row.

13. It is always possible to consider a conservative result by taking into account the explicit form of $\mathbf{x}(k) = A_0(k)\mathbf{x}(k) \oplus A_1(k)\mathbf{x}(k-1)$, which is given by either: $\mathbf{x}(k) = \mathcal{H}(k)\mathbf{r}(k)$, where $\mathcal{H}(k) \in ([A_0]\ [A_1])$ and $\mathbf{r}(k) = (\mathbf{x}(k)\ \mathbf{x}^{\mathrm{t}}(k-1))^{\mathrm{t}}$; or $\mathbf{x}(k) = \mathcal{H}(k)\mathbf{x}(k-1)$, where $\mathcal{H}(k) \in [A_0^{\star}][A_1]$.

Following Equation (1.12), we can write

$$\begin{aligned} x_i(k) &= \left(\bigoplus_{j=1}^{i-1} a_0^{ij}(k) x_j(k) \right) \oplus y_i(\mathbf{x}(k-1)), \\ \mathbf{z}(k) &= C(k)\mathbf{x}(k), \end{aligned} \tag{3.23}$$

with $y_i(\mathbf{x}(k-1))$ the i-th entry of $\mathbf{y}(\mathbf{x}(k-1)) = A_1(k)\mathbf{x}(k-1)$ (assuming the components of $\mathbf{x}(k-1)$ are obtained from the previous iteration). It is worth to mention that the $(n \times n)$-dimensional matrices $A_0(k) \in [A_0]$ and $A_1(k) \in [A_1]$, and the $(q \times n)$-dimensional matrix $C(k) \in [C]$ have entries (for instance, $a_0^{ij}(k) \in [\underline{a}_0^{ij}, \overline{a}_0^{ij}]$) that are independent random variables and uniformly distributed according to the CDF: $F(A_0(k) \in [A_0])$, $F(A_1(k) \in [A_1])$ and $F(C(k) \in [C])$.

By considering Equation (3.14), the prediction equation for all $i \in \{1, \ldots, n\}$ is given as follows:

$$\hat{x}_i(k|k-1) = \mathbb{E}\left[\left(\bigoplus_{j=1}^{i-1} a_0^{ij}(k) x_j(k) \right) \oplus y_i(\mathbf{x}(k-1)) \right]. \tag{3.24}$$

Starting from the equation

$$\hat{x}_1(k|k-1) = \mathbb{E}\left[y_1(\mathbf{x}(k-1)) \right]$$

and assuming that the values $\mathbf{x}(k-1)$ are fixed, then it is trivial to compute $\hat{x}_1(k|k-1)$. Conversely, for

$$\hat{x}_2(k|k-1) = \mathbb{E}\left[a_0^{21}(k) x_1(k) \oplus y_2(\mathbf{x}(k-1)) \right],$$

a similar trivial computation is not possible since it depends on $x_1(k)$. To calculate this and the subsequent mathematical expectations, the joint distribution of the vector $\mathbf{x}(k)$ should be obtained, which is a very difficult problem to solve. To avoid it, let us consider that a reliable value for $x_1(k)$ is given by the fixed value $\psi(x_1(k))$ (this is discussed in the following). Then, it renders $a_0^{21}(k)x_1(k) = a_0^{21}(k)\psi(x_1(k))$ to be of the same kind as $y_2(\mathbf{x}(k-1))$ and thus being easily computed. To compute

$$\hat{x}_3(k|k-1) = \mathbb{E}\left[a_0^{31}(k) x_1(k) \oplus a_0^{32}(k) x_2(k) \oplus y_3(\mathbf{x}(k-1)) \right],$$

we replace the random variables $x_1(k)$ and $x_2(k)$ by the fixed values $\psi(x_1(k))$ and $\psi(x_2(k))$, respectively. All the other subsequent computations are similar.

The question is: how to define $\psi(x_i(k))$?

We recall the method given in [28, Sec. III], which defines $\psi(x_i(k)) = \alpha_i \cdot \hat{x}_i(k|k-1)$ (multiplication in the classical context). It corresponds to the strategy for systems with periodic behaviour [a].

a. The set of parameters $\{\alpha_1, \cdots, \alpha_{n-1}\}$ is computed only once (offline phase) using Monte Carlo techniques. This set is robust under periodic steady-state regime, but if exogenous inputs modify the system behavior, the parameters should be tuned online, by evaluating the quality of the approximation through classical statistical parameters (as for instance the mean and the standard deviation of the difference between both means). Furthemore, thanks to the triangular structure of the matrix $A_0(k)$, it is possible to perform an uncoupled search for each α_i.

Let us assume that $\psi(x_i(k)) = \hat{x}_i(k|k)$ and in order to properly obtain $\hat{\mathbf{x}}(k|k)$, one must call n-times the procedure `Inv` (Algorithm 3.4). Clearly, this does not refer to a classical two-fold filter scheme because we update separately the prediction for each i-th component of the state vector.

Summing-up, the filtering algorithm of implicit forms, is given by Algorithm 3.6.

Algorithm 3.6: Filtering algorithm of implicit forms

Data: $\hat{\mathbf{x}}(k-1|k-1)$ and $\mathbf{z}(k)$
Result: $\hat{\mathbf{x}}(k|k) = \texttt{Filter}(\mathbf{z}(k), \hat{\mathbf{x}}(k-1|k-1))$
1 $\mathbf{y} = A_1(k)\hat{\mathbf{x}}(k-1|k-1)$;
2 $\hat{\mathbf{x}}(k|k-1) \leftarrow (\varepsilon, \ldots, \varepsilon)^{\mathrm{t}}$; `// initialize`
3 **foreach** $i \in \{1, \ldots, n\}$ **do**
4 $\quad \hat{x}_i(k|k-1) = \mathbb{E}[(\bigoplus_{j=1}^{i-1} a_0^{ij}(k)\hat{x}_j(k|k)) \oplus y_i]$; `// Equation (3.24)`
5 $\quad \hat{\mathbf{x}}(k|k) \leftarrow \texttt{Inv}(\mathbf{z}(k), \hat{\mathbf{x}}(k|k-1))$; `// Algorithm 3.3`
6 **end**
7 **return** $\hat{\mathbf{x}}(k|k)$;

The Algorithm 3.6 uses n-times the Algorithm 3.3. This Algorithm 3.3 returns an estimate $\hat{\mathbf{x}}(k|k)$ for all components of the state vector $\mathbf{x}(k)$, which is the solution of the constrained minimization problem (see Equations (3.10a) and (3.10b)). The input $\hat{\mathbf{x}}(k|k-1)$ of the Algorithm 3.3 is updated at each step i, with $\hat{x}_i(k|k-1)$ which is the corresponding entry of the prediction vector. At this step i, we must note that entries $j \in \{i+1, \ldots, n\}$ of $\hat{\mathbf{x}}(k|k-1)$ are still equal to ε, hence these entries j of the estimate $\hat{\mathbf{x}}(k|k)$ given by Algorithm 3.3 are equal to $\underline{x}_j$ since $x_j^0 \leq \underline{x}_j$ (see line 7).

Example 3.7. *Consider the third-order nondeterministic MPL system with mutually independent and uniformly distributed processing times, governed by the equation below*

$$\mathbf{x}(k) = A_0(k)\mathbf{x}(k) \oplus A_1(k)\mathbf{x}(k-1) \oplus B\mathbf{u}(k),$$

with $A_0(k) \in \begin{pmatrix} \varepsilon & \varepsilon & \varepsilon \\ [1,2] & \varepsilon & \varepsilon \\ [6,10] & [7,11] & \varepsilon \end{pmatrix}$, $A_1(k) \in \begin{pmatrix} [7,11] & \varepsilon & [2,9] \\ \varepsilon & [6,12] & [4,8] \\ \varepsilon & \varepsilon & [6,8] \end{pmatrix}$ *and* $B = \begin{pmatrix} e & \varepsilon \\ \varepsilon & \varepsilon \\ \varepsilon & e \end{pmatrix}$.

The output measurement is defined by

$$\mathbf{z}(k) = C(k)\mathbf{x}(k),$$

with $C(k) \in ([1,3]\ [2,4]\ \varepsilon)$.

For instance, let us assume that $\hat{\mathbf{x}}(0|0)$ *is initialized with the initial state* $\mathbf{x}(0) = (2\ 5\ 1)$.

Also, it should be mentioned that every calculation of $\hat{x}_i(k|k-1) = \mathbb{E}[x_i(k)|(\hat{\mathbf{x}}(k-1|k-1)^{\mathrm{t}}\ \mathbf{u}(k)^{\mathrm{t}})^{\mathrm{t}}]$ *takes into account a pseudo-random value* $\mathbf{u}(k)$ *drawn from the standard uniform distribution on the closed interval vector between* $(2,2)^{\mathrm{t}}$ *and* $(40,40)^{\mathrm{t}}$.

For simulation purposes, let us compare two different filter strategies for this MPL system in implicit form:

— *Filter* F_1 *uses the original filtering method given by Algorithm 3.6. In order to take the input action into account, the line 1 is replaced by* $\mathbf{y} = A_1(k)\hat{\mathbf{x}}(k-1|k-1) \oplus B\mathbf{u}(k)$.

— *Filter* F_2 *considers Algorithm 3.4 with line* 1 *given by:*

$$\hat{x}_i(k|k-1) = \mathbb{E}[x_i(k)|(\hat{\mathbf{x}}^{\mathrm{t}}(k-1|k-1)\ \mathbf{u}^{\mathrm{t}}(k))^{\mathrm{t}}],$$

with the right-hand side calculated by

$$\mathbb{E}\left[\bigoplus_{j=1}^{i-1} a_0^{ij}(k)\,(\alpha_j \cdot \hat{x}_j(k|k-1)) \oplus \bigoplus_{j=1}^{n} a_1^{ij}(k)\hat{x}_j(k-1|k-1) \oplus \bigoplus_{j=1}^{p} b^{ij}u_j(k)\right],$$

for all $i \in \{1,\ldots,n\}$ *and with* $\cdot$ *corresponding to the scalar multiplication in regular algebra. It corresponds to the strategy given in [28, Sec. III] for systems with periodic behavior, as previously presented. Let us assume that* $\alpha_j = 1$ *for all* $i \in \{1,\ldots,n\}$.

Table 3.4 shows the obtained results for simulations up to the occurrence of 4000 *events. Each position of the table corresponds to root-mean-square-error* $RMSE(x_i(k), \hat{x}_i(k|k))$ *with the usage of the corresponding filter.*

State i	F_1 $RMSE(\{\mathbf{x}(k)\}, \{\hat{\mathbf{x}}^{F_1}(k\|k)\})_i$	F_2 $RMSE(\{\mathbf{x}(k)\}, \{\hat{\mathbf{x}}^{F_2}(k\|k)\})_i$
1	2.8883	4.5451
2	1.1299	1.7783
3	1.7233	16.7846

Table 3.4 – Comparison between F_1 and F_2 of Example 3.7

The analysis of Table 3.4 indicates that the $RMSE$ *between the estimation provided by the* F_2 *method and the true value of the state is higher than the* F_1 *method with special attention to* $i = 3$ *(unobserved state). Indeed, as* $x_3(k)$ *is not observed and is the noisiest state, then this noise propagates throughout the TEG thanks to* $A_1(k)$ *and it penalizes the estimate of* $x_1(k)$ *and* $x_2(k)$*.*

CONCLUSION

Stochastic filtering of Max-Plus Linear (MPL) systems is still a work in progress, mainly due to the difficulty of using frameworks that reinterpret the basic notions of probability and stochasticity in a tropical semiring [101, 102, 103, 104].

On Chapter 2

Although the computation of the posterior probability function (PDF), the key of Bayesian formulation, cannot be computed in practice for MPL systems with stochastic processing times, we are able to compute its support with an improvement in the complexity of the forward reach set. We have shown in Chapter 2 that this computation is exact and cumbersome (worst-case scenarios) or approximate and fast, and is ultimately applied to systems without token-free place in the corresponding Timed Event Graph (TEG), otherwise we compute a conservative support. In this thesis, we did not study the use of this support to improve Particle filter algorithms, but it seems to be helpful in future works. This support is also useful to validate the alternative stochastic filtering scheme proposed in [28]. As future work, we aim to obtain a polynomial procedure to compute the backward reach set using the polyhedral approach (successful in the forward reach set computation), and then efficiently compute the *posterior* PDF support. It would also be interesting to take care of the exact calculation of this support for systems with token-free places, since we currently only have conservative solutions.

On Chapter 3

In Chapter 3, we have introduced an improvement to the alternative stochastic filtering algorithm proposed in [28]. We have presented a significant contribution to the extent that we minimize a trade-off criterion that considers noise in state and measurement. Nevertheless, in future works, we aim to further analyze and discuss the results and to also propose other trade-off criteria, since this is premature work in progress.

Also in Chapter 3, we have extended the filtering techniques to TEG with implicit part (token-free matrices), which are encountered in real applications rather than systems with

initially one token on each place. The corresponding TEG of these systems are generally live, and then some approximate solutions are discussed: first, based on an *offline* Monte Carlo simulation, which is robust only if the system works in its periodic behavior (the filter works fine); our contribution is a new *online* strategy that is efficient even if the system works in its transient behavior, without performing a single and imprecise *offline* Monte Carlo simulation, since an *adaptive tuning* occurs in the loop.

In general

As presented at the beginning of this chapter, the absence of a complete stochastic framework for MPL systems makes all known results very difficult to use (for example the well-known Kalman filter). Indeed, it seems necessary to reinterpret the basic notions of probability, ultimately changing the complete understanding of probabilistic measurement, giving rise to a *tropical* Kalman-type stochastic filtering theory [105]. For instance, in [106], the max-plus analog of the mathematical expectation is discussed. A promising idea is to use Possibility theory for dealing with certain types of uncertainty, being an alternative to probability theory [107, 108].

Appendix A

Residuation

A.1 Formulæ involving division

The following list provides some basic relations of left and right division, for the proofs and a more detailed list the reader is invited to consult [43, Chap. 4].

For a complete dioid $\mathcal{D}$ with $a, b, c \in \mathcal{D}$, we have:

$$a(a\backslash b) \succeq b \tag{A.1}$$
$$a\backslash a \succeq e \tag{A.2}$$
$$a(a\backslash a) = a \tag{A.3}$$
$$e\backslash a = a \tag{A.4}$$
$$\varepsilon\backslash a = \top \tag{A.5}$$
$$(a\backslash b)\ c \preceq a\backslash(bc) \tag{A.6}$$
$$a\backslash(b\backslash c) = (ba)\backslash c \tag{A.7}$$
$$(a\backslash b) \oplus (a\backslash c) \preceq a\backslash(b \oplus c) \tag{A.8}$$
$$(a\backslash b) \oplus (c\backslash b) \preceq (a \wedge c)\backslash b \tag{A.9}$$
$$(a\backslash b) \wedge (c\backslash b) = (a \oplus c)\backslash b \tag{A.10}$$
$$(a\backslash b) \wedge (a\backslash c) = a\backslash(b \wedge c) \tag{A.11}$$

A.2 Lemmas

Lemma A.1. *Given* $\mathbf{x} \in \overline{\mathbb{R}}_{\max}^{p}$ *and* $\mathbf{y} \in \overline{\mathbb{R}}_{\max}^{n}$*, the following equality holds* $(\mathbf{y}\not\backslash\mathbf{x})\backslash\mathbf{y} = \mathbf{x}$.

Proof. For all $i \in \{1, \ldots, n\}$ and $j \in \{1, \ldots, p\}$:

$$(\mathbf{y}\not\backslash\mathbf{x})_{ij} = y_i \not\backslash x_j = y_i - x_j,$$

and for all $j \in \{1, \ldots, p\}$:

$$((\mathbf{y}\not\backslash\mathbf{x})\backslash\mathbf{y}))_j = \bigwedge_{i=1}^{n} (\mathbf{y}\not\backslash\mathbf{x})_{ij}\backslash y_i = \bigwedge_{i=1}^{n} y_i - (y_i - x_j) = x_j.$$

□

Lemma A.2. *Let $a, b, c, d \in \overline{\mathbb{R}}_{\max}$. If $c \prec a$ then the following equivalence holds*

$$a \oplus b \leq c \oplus d \iff a \oplus b \leq d.$$

Proof. First, $a \oplus b \leq c \oplus d \Rightarrow a \oplus b \leq d$ since by assumption $c \prec a$ which implies $c \prec a \oplus b$, hence $c \prec a \oplus b \leq c \oplus d$ and $c \prec c \oplus d \iff c \prec d \iff c \oplus d = d$. Similarly, $c \prec a$ and $a \oplus b \leq d$ imply $c \prec a \oplus b \leq d \Rightarrow c \oplus c \prec a \oplus b \oplus c = a \oplus b \leq c \oplus d$, *i.e.*, we have $a \oplus b \leq d \Rightarrow a \oplus b \leq c \oplus d$ which concludes the proof. □

Appendix B

Set operations

B.1 Manipulations over sets

We recall that the complement of the intersection and the complement of the union of q sets are obtained by applying DeMorgan's Laws as shown below:

$$\left(\bigcap_{i=1}^{q}\mathcal{A}_j\right)^{\complement}=\bigcup_{i=1}^{q}\mathcal{A}_j^{\complement} \qquad \text{(B.1)} \qquad\qquad \left(\bigcup_{i=1}^{q}\mathcal{A}_j\right)^{\complement}=\bigcap_{i=1}^{q}\mathcal{A}_j^{\complement}. \qquad \text{(B.2)}$$

An alternative way to represent Equation (B.1) is related, for instance, to the following example of classical manipulation of sets:

$$\mathcal{A}_1^{\complement}\cup\mathcal{A}_1^{\complement}\cup\mathcal{A}_3^{\complement}=\mathcal{A}_1^{\complement}\cup[\mathcal{A}_1\cap\mathcal{A}_2^{\complement}]\cup[\mathcal{A}_1\cap\mathcal{A}_2\cap\mathcal{A}_3^{\complement}]$$

where $\cup$ stands for disjoint union. From this fact, the following always holds:

$$\left(\bigcap_{i=1}^{q}\mathcal{A}_j\right)^{\complement}=\bigcup_{i=1}^{q}\left[\left(\bigcap_{k=1}^{i-1}\mathcal{A}_k\right)\cap\mathcal{A}_i^{\complement}\right], \qquad \text{(B.3)}$$

with $\bigcap_{k=1}^{0}\mathcal{A}_k=\mathcal{U}$, where $\mathcal{U}$ stands for universal set. This equation represents the union of pairwise disjoint sets. To understand it better, for instance, suppose $q=2$ then Equation (B.3) yields

$$\mathcal{A}_1^{\complement}\cup[\mathcal{A}_1\cap\mathcal{A}_2^{\complement}],$$

and we aim at evaluating $\mathcal{A}_1^{\complement}\cap[\mathcal{A}_1\cap\mathcal{A}_2^{\complement}]$. Thanks to the associative law over the intersection operation, we obtain

$$\mathcal{A}_1^{\complement}\cap[\mathcal{A}_1\cap\mathcal{A}_2^{\complement}]=\underbrace{[\mathcal{A}_1^{\complement}\cap\mathcal{A}_1]}_{=\emptyset}\cap\mathcal{A}_2^{\complement}=\emptyset,$$

i.e., $\mathcal{A}_1^{\complement}$ and $[\mathcal{A}_1\cap\mathcal{A}_2^{\complement}]$ are two disjoint sets.

Another interesting property concerns, for instance, the set

$$(\mathcal{A}_1 \cup \mathcal{A}_2) \cap (\mathcal{A}_3 \cup \mathcal{A}_4)$$

which is equal to

$$(\mathcal{A}_1 \cap \mathcal{A}_3) \cup (\mathcal{A}_1 \cap \mathcal{A}_4) \cup (\mathcal{A}_2 \cap \mathcal{A}_3) \cup (\mathcal{A}_2 \cap \mathcal{A}_4)$$

thanks the distributive law over the intersection. Hence, consider the sets $\mathcal{A}_j^i$ and the following operation:

$$\begin{aligned}\bigcap_{i=1}^{q}\bigcup_{j=1}^{n}\mathcal{A}_j^i &= (\mathcal{A}_1^1 \cup \cdots \cup \mathcal{A}_n^1) \cap (\mathcal{A}_1^2 \cup \cdots \cup \mathcal{A}_n^2) \cap (\mathcal{A}_1^q \cup \cdots \cup \mathcal{A}_n^q),\\ &= (\mathcal{A}_1^1 \cap \mathcal{A}_1^2 \cap \cdots \cap \mathcal{A}_1^n) \cup (\mathcal{A}_1^1 \cap \mathcal{A}_1^2 \cap \cdots \cap \mathcal{A}_2^q) \cup \cdots \cup (\mathcal{A}_n^1 \cap \mathcal{A}_n^2 \cap \cdots \cap \mathcal{A}_n^q), \end{aligned} \tag{B.4}$$

This equation is not in a compact way. However, by considering $\mathcal{G} = \{1, \ldots, n\}^q$, $g_i \in \{1, \ldots, n\}$ and $\mathbf{g} = (g_1, g_2, \ldots, g_q) \in \mathcal{G}$, then:

$$\bigcap_{i=1}^{q}\bigcup_{j=1}^{n}\mathcal{A}_j^i = \bigcup_{\mathbf{g}\in\mathcal{G}}\bigcap_{i=1}^{q}\mathcal{A}_{g_i}^i. \tag{B.5}$$

B.2 Solution set to $\mathbf{y} \leq A\mathbf{x}$ in the Max-Plus algebra

In the sequel, we recall the results of [35] and [29].

Let $\mathbf{y} \leq A\mathbf{x}$ be an inequality, with $A \in \overline{\mathbb{R}}_{\max}^{n\times p}$, $\mathbf{x} \in \overline{\mathbb{R}}_{\max}^{p}$ and $\mathbf{y} \in \overline{\mathbb{R}}_{\max}^{n}$. The set

$$\chi = \{\mathbf{x} \in \overline{\mathbb{R}}_{\max}^{p} \mid \mathbf{y} \leq A\mathbf{x}\} = \bigcap_{i=1}^{n}\left\{\mathbf{x} \in \overline{\mathbb{R}}_{\max}^{p} \mid y_i \leq \bigoplus_{j=1}^{p} a_{ij} \otimes x_j\right\}$$

represents all $\mathbf{x}$ that satisfies the previous inequality.

Thanks to Equation (B.1) we calculate $\chi^{\complement}$ as follows

$$\chi^{\complement} = \bigcup_{i=1}^{n}\left(\left\{\mathbf{x} \in \overline{\mathbb{R}}_{\max}^{p} \mid y_i \leq \bigoplus_{j=1}^{p} a_{ij} \otimes x_j\right\}\right)^{\complement}$$

with $(\{\mathbf{x} \in \overline{\mathbb{R}}_{\max}^{p} \mid y_i \leq \bigoplus_{j=1}^{p} a_{ij} \otimes x_j\})^{\complement} = \{\mathbf{x} \in \overline{\mathbb{R}}_{\max}^{p} \mid y_i > \bigoplus_{j=1}^{p} \overline{a}_{ij} \otimes x_j\}$ because

$y_i \in \overline{\mathbb{R}}_{\max}$. Thus,

$$\chi^{\complement} = \bigcup_{i=1}^{n} \left\{ \mathbf{x} \in \overline{\mathbb{R}}_{\max}^{p} \mid y_i > \bigoplus_{j=1}^{p} a_{ij} \otimes x_j \right\}.$$

Taking into account

$$\bigoplus_{j=1}^{p} a_{ij} \otimes x_j < y_i \iff \begin{cases} a_{i1} \otimes x_1 < y_i \iff x_1 < a_{i1} \backslash y_i \text{ and} \\ a_{i2} \otimes x_2 < y_i \iff x_2 < a_{i2} \backslash y_i \text{ and} \\ \quad \vdots \\ a_{in} \otimes x_n < y_i \iff x_n < a_{in} \backslash y_i, \end{cases}$$

then

$$\chi^{\complement} = \bigcup_{i=1}^{n} \left(\bigcap_{j=1}^{p} \{ \mathbf{x} \in \overline{\mathbb{R}}_{\max}^{p} \mid x_j < \sigma_{ij} \} \right), \text{ where } \sigma_{ij} = a_{ij} \backslash y_i.$$

Using Equation (B.2), χ is re-obtained as $\chi = (\chi^{\complement})^{\complement}$, formally

$$\begin{aligned} \chi &= \left(\bigcup_{i=1}^{n} \left(\bigcap_{j=1}^{p} \{ \mathbf{x} \in \overline{\mathbb{R}}_{\max}^{p} \mid x_j < \sigma_{ij} \} \right) \right)^{\complement} = \bigcap_{i=1}^{n} \left(\bigcap_{j=1}^{p} \{ \mathbf{x} \in \overline{\mathbb{R}}_{\max}^{n} \mid x_j < \sigma_{ij} \} \right)^{\complement} \\ &= \bigcap_{i=1}^{n} \left(\bigcup_{j=1}^{p} \{ \mathbf{x} \in \overline{\mathbb{R}}_{\max}^{p} \mid x_j \geq \sigma_{ij} \} \right) \end{aligned}$$

However, it seems possible to reinterpret χ. First, we recall Equation (B.3), then

$$\begin{aligned} &\left(\bigcap_{j=1}^{p} \{ \mathbf{x} \in \overline{\mathbb{R}}_{\max}^{p} \mid x_j < \sigma_{ij} \} \right)^{\complement} = \bigcup_{j=1}^{p} \left[\left(\bigcap_{k=1}^{j-1} \{ \mathbf{x} \in \overline{\mathbb{R}}_{\max}^{p} \mid x_k < \sigma_{ij} \} \right) \cap \{ \mathbf{x} \in \overline{\mathbb{R}}_{\max}^{p} \mid x_j < \sigma_{ij} \}^{\complement} \right] \\ &= \bigcup_{j=1}^{p} \left[\left(\bigcap_{k=1}^{j-1} \{ \mathbf{x} \in \overline{\mathbb{R}}_{\max}^{p} \mid x_k < \sigma_{ij} \} \right) \cap \{ \mathbf{x} \in \overline{\mathbb{R}}_{\max}^{p} \mid x_j \geq \sigma_{ij} \} \right] \end{aligned}$$

with $\bigcap_{k=1}^{0} \{ \mathbf{x} \in \overline{\mathbb{R}}_{\max}^{p} \mid x_k < \sigma_{ij} \} = \overline{\mathbb{R}}_{\max}^{n}$, and thus

$$\chi = \bigcap_{i=1}^{n} \left(\bigcup_{j=1}^{p} set_j^i \right), \text{where } set_j^i = \left(\bigcap_{k=1}^{j-1} \{ \mathbf{x} \in \overline{\mathbb{R}}_{\max}^{p} \mid x_k < \sigma_{ij} \} \right) \cap \{ \mathbf{x} \in \overline{\mathbb{R}}_{\max}^{p} \mid x_j \geq \sigma_{ij} \}.$$

Remark B.1. *$set_{j_1}^i \cap set_{j_2}^i = \emptyset$ holds for all $i \in \{1, \ldots, n\}$ and $j_1 \neq j_2$ where $j_1, j_2 \in \{1, \ldots, p\}$, i.e., $set_{j_1}^i$ and $set_{j_2}^i$ are pairwise disjoint sets.*

Finally, Equation (B.5) allows expressing

$$\chi = \bigcap_{i=1}^{n} \left(\bigcup_{j=1}^{p} set_j^i \right) = \bigcup_{\mathbf{g} \in \mathcal{G}} SET^{\mathbf{g}}, \quad SET^{\mathbf{g}} = \bigcap_{i=1}^{n} set_{g_i}^i \tag{B.6}$$

with $\mathcal{G} = \{1, \ldots, p\}^n$, $g_i \in \{1, \ldots, p\}$ and $\mathbf{g} = (g_1, g_2, \ldots, g_n) \in \mathcal{G}$.

Lemma B.1 ([29])**.** *The sets $SET^{\mathbf{g}}$ for all $\mathbf{g} \in \mathcal{G}$ are pairwise disjoint from each other.*

Proof. Consider $\mathbf{g}, \mathbf{g}' \in \mathcal{G}$ with $\mathbf{g} \neq \mathbf{g}'$. Then, $\exists\ k \in \{1, \ldots n\}$ such that $g_k = g'_k$. $SET^{\mathbf{g}}$ and $SET^{\mathbf{g}'}$ are disjoint if and only if $SET^{\mathbf{g}} \cap SET^{\mathbf{g}'}$. Then

$$\begin{aligned} SET^{\mathbf{g}} \cap SET^{\mathbf{g}'} &= \left(\bigcap_{i=1}^{n} set_{g_i}^i \right) \cap \left(\bigcap_{i=1}^{n} set_{g'_i}^i \right) \\ &= \left(\bigcap_{i=1}^{n} set_{g_i}^i \right) \cap \left(\bigcap_{i=1}^{n} set_{g'_i}^i \right) \\ &= \left[\bigcap_{i \in \{1,\ldots,n\} \setminus k} (set_{g_i}^i \cap set_{g'_i}^i) \right] \cap \underbrace{(set_{g_k}^i \cap set_{g'_k}^i)}_{=\emptyset \text{ (see Remark B.1)}} = \emptyset \end{aligned}$$

□

Remark B.2. *In the worst-case scenario the set χ in Equation* (B.6) *can be represented by the union of p^n pairwise disjoint sets $SET^{\mathbf{g}}$. However, some $SET^{\mathbf{g}}$ may be empty. Hence, in general, the set χ can be represented by $N \leq p^n$ pairwise disjoint sets.*

Appendix C

Proofs

C.1 Proofs of Section 2.2.3

Lemma C.1 ([29]). *Given* $\underline{A}\mathbf{x} \leq \mathbf{y} \leq \overline{A}\mathbf{x}$ *where* $\underline{A}, \overline{A} \in \overline{\mathbb{R}}_{\max}^{m\times n}$, $\mathbf{x} \in \overline{\mathbb{R}}_{\max}^{n}$ *and* $\mathbf{y} \in \overline{\mathbb{R}}_{\max}^{m}$ *are assumed to be known. Then* $\exists A \in [A]$, *i.e.,* $\underline{A} \leq A \leq \overline{A}$ *such that* $A\mathbf{x} = \mathbf{y}$.

Proof. From $\underline{A}\mathbf{x} \leq \mathbf{y} \leq \overline{A}\mathbf{x}$ where $\underline{A}, \overline{A} \in \overline{\mathbb{R}}_{\max}^{m\times n}$, $\mathbf{x} \in \overline{\mathbb{R}}_{\max}^{n}$ and $\mathbf{y} \in \overline{\mathbb{R}}_{\max}^{m}$ then the following holds

$$\underline{A}\mathbf{x} \leq \mathbf{y} \iff \underline{A} \leq \hat{A} = \mathbf{y} \not{\backslash} \mathbf{x}$$

where $(\mathbf{y} \not{\backslash} \mathbf{x})_{ij} = y_i - x_j$ for all $i \in \{1, \ldots, m\}$ and $j \in \{1, \ldots, n\}$. From $\underline{A} \leq \hat{A}$ then

$$\underline{A} \leq \hat{A} \wedge \overline{A} \leq \overline{A}.$$

Let $\mathbf{y}' = (\hat{A} \wedge \overline{A})\mathbf{x}$ where $\hat{A} \wedge \overline{A} \in [A]$, thus for all $i \in \{1, \ldots m\}$

$$\begin{aligned} y_i &= \max_{1\leq j\leq n} (\min(y_i - x_i, \overline{a}_{ij}) + x_j) = \max_{1\leq j\leq n} (\min(y_i, \overline{a}_{ij} + x_j)) \\ &= \min\left(y_i, \max_{1\leq j\leq n} (\overline{a}_{ij} + x_j)\right) \end{aligned}$$

Hence,

$$\mathbf{y}' = (\hat{A} \wedge \overline{A})\mathbf{x} = \mathbf{y} \wedge \overline{A}\mathbf{x} = \mathbf{y} \iff \mathbf{y} \leq \overline{A}\mathbf{x} \text{ (Assumption)}$$

Finally, $(\hat{A} \wedge \overline{A})\mathbf{x} = \mathbf{z}$ and the proof is complete. □

Lemma C.2. *Let* X, α *and* β *be a real-valued random variable with bounded domain* $\mathtt{dom}(X)$ *and two constants of* $\mathbb{R}$, *respectively. Given that the* k*-th moment of* X *is finite, i.e.,* $\mathbb{E}[X^k] = \int_{\mathtt{dom}(X)} t^k p_X(t)dt$ *is finite, then if* $\alpha \geq \beta$ *we have* $\mathbb{E}[(X+\alpha)^p] \geq \mathbb{E}[(X+\beta)^p]$ *with* p *an integer.*

Proof. Let us explore the expansion of $(X+c)^p$ using the well known Binomial theorem

$$(X+c)^p = \sum_{k=0}^{p} \binom{p}{k} X^{p-k} c^k = \binom{p}{0} X^p + \binom{p}{1} X^{p-1} c^1 + \cdots + \binom{p}{p-1} X^1 c^{p-1} + \binom{p}{p} c^p.$$

Thus

$$\mathbb{E}[(X+c)^p] = \int_{\mathtt{dom}(X)} (t+c)^p p_X(t) \mathrm{d}t = \sum_{k=0}^{p} \underbrace{\binom{p}{k}}_{C(p,k)\ \text{positive integer}} c^k \underbrace{\int_{\mathtt{dom}(X)} t^{p-k} p_X(t) \mathrm{d}t}_{\mathbb{E}[X^k]\ \text{finite}}$$

and clearly $\mathbb{E}[(X+\alpha)^p] = \sum_{k=0}^{p} \alpha^k C(p,k) \mathbb{E}[X^k] \geq \mathbb{E}[(X+\beta)^p] = \sum_{k=0}^{p} \beta^k C(p,k) \mathbb{E}[X^k]$. $\square$

BIBLIOGRAPHY

[1] C. G. Cassandras and S. Lafortune, *Introduction to Discrete Event Systems*. Kluwer Academic Publishers, 1999.

[2] J. Mao and C. G. Cassandras, "Optimal control of multilayer discrete event systems with real-time constraint guarantees," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 44, no. 10, pp. 1425–1434, 2014.

[3] E. Altman, B. Gaujal, and A. Hordijk, "Discrete-event control of stochastic networks : multimodularity and regularity / eitan altman, bruno gaujal, arie hordijk," *SERBIULA (sistema Librum 2.0)*, vol. 1829, 01 2003.

[4] J.-Y. Le Boudec and P. Thiran, *Network Calculus: A Theory of Deterministic Queuing Systems for the Internet*. Berlin, Heidelberg: Springer-Verlag, 2001.

[5] J. E. Hopcroft, R. Motwani, and J. D. Ullman, *Introduction to Automata Theory, Languages, and Computation*. Addison Wesley, 3rd ed., 2006.

[6] B. Gaujal, A. Girault, and S. Plassart, "Dynamic speed scaling minimizing expected energy consumption for real-time tasks," *Journal of Scheduling*, vol. 23, 10 2020.

[7] T. Murata, "Petri nets: Properties, analysis and applications," *Proceedings of the IEEE*, vol. 77, no. 4, pp. 541–580, 1989.

[8] F. Baccelli and T. Konstantopoulos, "Estimates of cycle times in stochastic Petri nets," in *Proceedings of Workshop on Stochastic Analysis* (I. Karatzas, ed.), (Rutgers University), Springer, Berlin, 1991.

[9] G. Cohen, S. Gaubert, M. McGettrick, and J.-P. Quadrat, "Maxplus toolbox of scilab," 2009.

[10] 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*. No. v. 13 in Max Plus at work: modeling and analysis of synchronized systems : a course on Max-Plus algebra and its applications, Princeton University Press, 2006.

[11] G. Cohen, S. Gaubert, and J.-P. Quadrat, "Max-plus algebra and system theory: Where we are and where to go now," *Annual Reviews in Control*, vol. 23, pp. 207 – 219, 1999.

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

[13] T. van den Boom, J. Xu, and B. De Schutter, "Corrections to "Model predictive control for stochastic max-plus linear systems with chance constraints", [IEEE Trans. on Aut. Control, 64(1): 337–342, 2019]," *IEEE Transactions on Automatic Control*, vol. 65, pp. 905–906, Feb. 2020.

[14] 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, pp. 329–354, sep 2007.

[15] M. Lhommeau, L. Hardouin, J.-L. Ferrier, and I. Ouerghi, "Interval analysis in dioid: Application to robust open-loop control for timed event graphs," in *Decision and Control, 2005 and 2005 European Control Conference. CDC-ECC '05. 44th IEEE Conference on*, pp. 7744–7749, 2005.

[16] 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, pp. 218–242, Jan. 2009.

[17] S. Amari, I. Demongodin, J. J. Loiseau, and C. Martinez, "Max-plus control design for temporal constraints meeting in timed event graphs," *IEEE Transactions on Automatic Control*, vol. 57, pp. 462–467, Feb 2012.

[18] R. D. Katz, "Max-plus (a,b)-invariant spaces and control of timed discrete-event systems," *IEEE Transactions on Automatic Control*, vol. 52, pp. 229–241, Feb 2007.

[19] C. A. Maia, L. Hardouin, R. Santos-Mendes, and J. J. Loiseau, "A super-eigenvector approach to control constrained max-plus linear systems," in *2011 50th IEEE Conference on Decision and Control and European Control Conference*, pp. 1136–1141, Dec 2011.

[20] V. M. Gonçalves, C. A. Maia, and L. Hardouin, "On the steady-state control of timed event graphs with firing date constraints," *IEEE Transactions on Automatic Control*, vol. 61, pp. 2187–2202, Aug 2016.

[21] 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.

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

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

[24] S. Farahani, T. van den Boom, and B. De Schutter, "On optimization of stochastic max-min-plus-scaling systems – An approximation approach," *Automatica*, vol. 83, pp. 20–27, Sept. 2017.

[25] C. Paya, E. Le Corronc, Y. Pencolé, and P. Vialletelle, "Observer-based detection of time shift failures in (max,+)-linear systems," in *The 31st International Workshop on Principles of Diagnosis (DX-2020)*, (Nashville, United States), Sept. 2020.

[26] L. Hardouin, B. Cottenceau, Y. Shang, and J. Raisch, *Control and State Estimation for Max-Plus Linear Systems.* Now Foundations and Trends, 2018.

[27] R. M. F. Cândido, L. Hardouin, M. Lhommeau, and R. Santos Mendes, "Conditional reachability of uncertain max plus linear systems," *Automatica*, vol. 94, pp. 426 – 435, 2018.

[28] R. S. Mendes, L. Hardouin, and M. Lhommeau, "Stochastic filtering of max-plus linear systems with bounded disturbances," *IEEE Transactions on Automatic Control*, vol. 64, pp. 3706–3715, Sep. 2019.

[29] R. M. F. Candido, L. Hardouin, M. Lhommeau, and R. Santos-Mendes, "An algorithm to compute the inverse image of a point with respect to a nondeterministic max plus linear system," *IEEE Transactions on Automatic Control*, pp. 1–1, 2020.

[30] T. van den Boom and B. De Schutter, "Model predictive control for perturbed max-plus-linear systems," *Systems & Control Letters*, vol. 45, pp. 21–33, Jan. 2002.

[31] 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, pp. 299–302, Mar 1990.

[32] B. F. Heidergott, *Max-Plus Linear Stochastic Systems and Perturbation Analysis (The International Series on Discrete Event Dynamic Systems).* Secaucus, NJ, USA: Springer-Verlag New York, Inc., 2006.

[33] 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, pp. 5606–5628, Dec. 2010.

[34] L. Hardouin, C. A. Maia, B. Cottenceau, and R. Santos-Mendes, "Max-plus Linear Observer: Application to manufacturing Systems," in *p10, WODES, WODES'10*, (Berlin), pp. 171–176, sep 2010.

[35] R. M. Ferreira Cândido, *Reachability Analysis of Uncertain Max Plus Linear Systems*. Theses, Université d'Angers ; Universidade estadual de Campinas (Brésil), June 2017.

[36] 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, pp. 174 –188, feb 2002.

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

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

[39] X. Allamigeon, S. Gaubert, and E. Goubault, "Inferring min and max invariants using max-plus polyhedra," in *Proceedings of the 15th International Symposium on Static Analysis*, SAS'08, (Berlin, Heidelberg), pp. 189–204, Springer-Verlag, 2008.

[40] D. L. Dill, "Timing assumptions and verification of finite-state concurrent systems," in *Automatic Verification Methods for Finite State Systems* (J. Sifakis, ed.), (Berlin, Heidelberg), pp. 197–212, Springer Berlin Heidelberg, 1990.

[41] J. L. Peterson, "Petri nets," vol. 9, pp. 223–252, Sept. 1977.

[42] F. Commoner, A. Holt, S. Even, and A. Pnueli, "Marked directed graphs," *Journal of Computer and System Sciences*, vol. 5, no. 5, pp. 511–523, 1971.

[43] F. Baccelli, G. Cohen, G. Olsder, and J. Quadrat, *Synchronization and Linearity : An Algebra for Discrete Event Systems*. Wiley and Sons, 1992.

[44] P. Butkovic, *Max-linear Systems: Theory and Algorithms*. Springer, Jan. 2010.

[45] R. A. Cuninghame-Green and P. Butkovic, "The equation a⊗x=b⊗xover (max, +)," *Theor. Comput. Sci.*, vol. 293, pp. 3–12, 2003.

[46] P. Butkovič and G. Hegedus, "An elimination method for finding all solutions of the system of linear equations over an extremal algebra," *Mat. Obzor 20*, pp. 203–215, 1984.

[47] W. Heemels, B. De Schutter, and A. Bemporad, "Equivalence of hybrid dynamical models," *Automatica*, vol. 37, pp. 1085–1091, July 2001.

[48] D. Adzkiya, B. De Schutter, and A. Abate, "Computational techniques for reachability analysis of max-plus-linear systems," *Automatica*, vol. 53, no. 3, pp. 293–302, 2015.

[49] M. S. Mufid, D. Adzkiya, and A. Abate, "Tropical abstractions of max-plus linear systems," in *Formal Modeling and Analysis of Timed Systems* (D. N. Jansen and P. Prabhakar, eds.), (Cham), pp. 271–287, Springer International Publishing, 2018.

[50] B. Heidergott, G. J. 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*. Princeton Univerity Press, 2005.

[51] T. van den Boom and B. De Schutter, "Model predictive control for perturbed max-plus-linear systems," *Systems & Control Letters*, vol. 45, no. 1, pp. 21–33, 2002.

[52] J. Xu, T. van den Boom, and B. De Schutter, "Model predictive control for stochastic max-plus linear systems with chance constraints," *IEEE Transactions on Automatic Control*, vol. 64, no. 1, pp. 337–342, 2019.

[53] G. Olsder, J. Resing, R. de Vries, M. Keane, and G. Hooghiemstra, "Discrete event systems with stochastic processing times," *IEEE Transactions on Automatic Control*, pp. 299–302, 1990.

[54] J. Resing, R. de Vries, G. Hooghiemstra, M. Keane, and G. Olsder, "Asymptotic behavior of random discrete event systems," *Stochastic Processes and their Applications*, vol. 36, no. 2, pp. 195–216, 1990.

[55] B. F. Heidergott, *Max-Plus Linear Stochastic Systems and Perturbation Analysis (The International Series on Discrete Event Dynamic Systems)*. Berlin, Heidelberg: Springer-Verlag, 2006.

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

[57] T. Brunsch, L. Hardouin, C. A. Maia, and J. Raisch, "Duality and interval analysis over idempotent semirings," *Linear Algebra and its Applications*, vol. 437, pp. 2436–2454, 2012.

[58] L. Hardouin, B. Cottenceau, M. Lhommeau, and E. Le Corronc, "Interval systems over idempotent semiring," *Linear Algebra and its Applications*, vol. 431, pp. 855–862, Aug. 2009. doi:10.1016/j.LAA.2009.03.039.

[59] G. L. Litvinov and A. N. Sobolevskiĭ, "Idempotent interval analysis and optimization problems," *Reliable Computing*, vol. 7, no. 5, pp. 353–377, 2001.

[60] B. Berthomieu and M. Menasche, "An enumerative approach for analyzing time petri nets," in *Proceedings IFIP*, pp. 41–46, Elsevier Science Publishers, 1983.

[61] D. L. Dill, "Timing assumptions and verification of finite-state concurrent systems," in *Proceedings of the International Workshop on Automatic Verification Methods for Finite State Systems (AVMFSS'89)* (J. Sifakis, ed.), vol. 407 of *Lecture Notes in Computer Science*, pp. 197–212, Springer-Verlag, 1990.

[62] S. Yovine, "Model checking timed automata," in *European Educational Forum: School on Embedded Systems*, 1996.

[63] A. Miné, *Domaines numériques abstraits faiblement relationnels*. PhD thesis, Palaiseau, Ecole polytechnique, 2004.

[64] A. Miné, "A new numerical abstract domain based on difference-bound matrices," in *Symposium on Program as Data Objects*, pp. 155–172, Springer, 2001.

[65] D. Adzkiya, B. De Schutter, and A. Abate, "Forward reachability computation for autonomous max-plus-linear systems," in *Proceedings of the 20th International Conference on Tools and Algorithms for the Construction and Analysis of Systems (TACAS 2014)*, (Grenoble, France), pp. 248–262, Apr. 2014.

[66] D. Adzkiya, B. De Schutter, and A. Abate, "Backward reachability of autonomous max-plus-linear systems," in *Proceedings of the 12th International Workshop on Discrete Event Systems* (J. Faure, J. Lesage, and O. Gonzalez, eds.), pp. 117–122, IEEE Society, 2014. NEO; WODES 2014, Paris, France ; Conference date: 14-05-2014 Through 16-05-2014.

[67] D. Adzkiya, Y. Zhang, and A. Abate, "Verisimpl 2: An open-source software for the verification of max-plus-linear systems," *Discrete Events Dynamical Systems*, 2015.

[68] S. Gaubert and R. D. Katz, "The minkowski theorem for max-plus convex sets," *Linear Algebra and its Applications*, vol. 421, no. 2-3, pp. 356–369, 2007. Special Issue in honor of Miroslav Fiedler.

[69] X. Allamigeon, S. Gaubert, and E. Goubault, "Computing the Vertices of Tropical Polyhedra using Directed Hypergraphs," *Discrete and Computational Geometry*, vol. 49, pp. 247–279, Feb. 2013.

[70] S. Gaubert and R. Katz, "Minimal half-spaces and external representation of tropical polyhedra," *Journal of Algebraic Combinatorics*, vol. 33, no. 3, pp. 325–348, 2011.

[71] M. Akian, S. Gaubert, and A. Guterman, "Tropical polyhedra are equivalent to mean payoff games," *International Journal of Algebra and Computation*, vol. 22, no. 1, pp. 1250001, 43, 2012. See also arXiv:0912.2462.

[72] U. Zwick and M. Paterson, "The complexity of mean payoff games on graphs," *Theoretical Computer Science*, vol. 158, no. 1, pp. 343–359, 1996.

[73] E. Gawrilow and M. Joswig, *polymake: a Framework for Analyzing Convex Polytopes*, pp. 43–73. Basel: Birkhäuser Basel, 2000.

[74] X. Allamigeon, "Tropical polyhedra library," 2009.

[75] X. Allamigeon, S. Gaubert, and R. Katz, "The number of extreme points of tropical polyhedra," *Journal of Combinatorial Theory, Series A*, vol. 118, no. 1, pp. 162 – 189, 2011.

[76] X. Allamigeon, *Static analysis of memory manipulations by abstract interpretation - Algorithmics of tropical polyhedra, and application to abstract interpretation*. Theses, Ecole Polytechnique X, Nov. 2009.

[77] D. Butkovič, "A condition for the strong regularity of matrices in the minimax algebra," *Discrete Appl. Math.*, vol. 11, pp. 209–222, 1985.

[78] P. Butkovič, H. Schneider, and others, "Generators, extremals and bases of max cones," *Linear algebra and its applications*, vol. 421, no. 2-3, pp. 394–406, 2007.

[79] J. Dyhrberg, Q. Lu, M. Madsen, and S. Ravn, "Computations on zones using max-plus algebra," 11 2010.

[80] Q. Lu, M. Madsen, M. Milata, S. Ravn, U. Fahrenberg, and K. G. Larsen, "Reachability analysis for timed automata using max-plus algebra," *The Journal of Logic and Algebraic Programming*, vol. 81, no. 3, pp. 298 – 313, 2012.

[81] J. Dyhrberg, Q. Lu, M. Madsen, and S. Ravn, "Computations on zones using max-plus algebra," in *Proc. 22nd Nordic Workshop in Programming Theory (NWPT10). URL http://vbn. aau. dk/ws/files/61077428/1274952619. pdf*, 2010.

[82] X. Allamigeon, U. Fahrenberg, S. Gaubert, R. D. Katz, and A. Legay, "Tropical fourier-motzkin elimination, with an application to real-time verification," *International Journal of Algebra and Computation*, vol. 24, no. 05, pp. 569–607, 2014.

[83] G. Espindola-Winck, L. Hardouin, and M. Lhommeau, "Max-plus polyhedra-based state characterization for umpl systems," in *2022 European Control Conference (ECC)*, pp. 1037–1042, 2022.

[84] G. Espindola-Winck, R. M. F. Cândido, L. Hardouin, and M. Lhommeau, "Efficient state-estimation of uncertain max-plus linear systems with high observation noise," *IFAC-PapersOnLine*, vol. 55, no. 28, pp. 228–235, 2022. 16th IFAC Workshop on Discrete Event Systems WODES 2022.

[85] L. Hardouin, Y. Shang, C. A. Maia, and B. Cottenceau, "Observer-based controllers for max-plus linear systems," *IEEE Transactions on Automatic Control*, vol. 62, no. 5, pp. 2153–2165, 2017.

[86] G. Espindola-Winck, L. Hardouin, M. Lhommeau, and R. Santos-Mendes, "Stochastic filtering scheme of implicit forms of uncertain max-plus linear systems," *IEEE Transactions on Automatic Control*, vol. 67, no. 8, pp. 4370–4376, 2022.

[87] G. Espindola-Winck, L. Hardouin, M. Lhommeau, and R. Santos Mendes, "Criteria stochastic filtering of max-plus discrete event systems with bounded random variables," *to appear in Cosy 22, Bologna*, November 2022.

[88] F. Baccelli and V. Schmidt, "Taylor Expansions for Poisson Driven (max, +)-Linear Systems," Tech. Rep. RR-2494, INRIA, Feb. 1995.

[89] E. Altman, B. Gaujal, and A. Hordijk, "Admission control in stochastic event graphs," *IEEE Transactions on Automatic Control*, vol. 45, no. 5, pp. 854–867, 2000.

[90] N. Krivulin, "Unbiased estimates for gradients of stochastic network performance measures," *Acta Applicandae Mathematicae*, vol. 33, pp. 21–43, oct 1993.

[91] B. Anderson and J. Moore, *Optimal Filtering*. Englewood Cliffs, NJ: Prentice-Hall, 1979.

[92] A. Jazwinski, *Stochastic processes and filtering theory*. No. 64 in Mathematics in science and engineering, New York, NY [u.a.]: Acad. Press, 1970.

[93] B. Ristic, S. Arulampalam, and N. Gordon, *Beyond the Kalman Filter: Particle Filters for Tracking Applications*. Artech House radar library, Artech House, 2004.

[94] H. Fang, N. Tian, Y. Wang, M. Zhou, and M. A. Haile, "Nonlinear bayesian estimation: from kalman filtering to a broader horizon," *IEEE/CAA Journal of Automatica Sinica*, vol. 5, no. 2, pp. 401–417, 2018.

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

[96] S. Julier, J. Uhlmann, and H. 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, 2000.

[97] M. Frechét, "Sur la loi de probabilité de l'écart maximum," 1928.

[98] T. J. van den Boom and B. De Schutter, "Analytic expressions in stochastic max-plus-linear algebra," in *53rd IEEE Conference on Decision and Control*, pp. 1608–1613, 2014.

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

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

[101] P. Del Moral, J.-C. Noyer, and G. Salut, "Maslov optimisation theory: stochastic interpretation, particle resolution," in *11th International Conference on Analysis and Optimization of Systems Discrete Event Systems*, pp. 312–318, Springer, 1994.

[102] P. Del Moral and M. Doisy, "On applications of maslov optimization theory," *Mathematical Notes*, vol. 69, pp. 232–244, 01 2001.

[103] M. Akian, "Densities of idempotent measures and large deviations," *Transactions of the American Mathematical Society*, vol. 351, no. 11, pp. 4515–4543, 1999.

[104] A. Puhalskii, *Large deviations and idempotent probability*. Chapman and Hall/CRC, 2001.

[105] N. M. Tran, "Tropical gaussians: a brief survey," *Algebraic Statistics*, vol. 11, pp. 155–168, dec 2020.

[106] M. Akian, J.-P. Quadrat, and M. Viot, "Bellman processes," 1994.

[107] A. Mazeika, L. Jaulin, and C. Osswald, "A new approach for computing with fuzzy sets using interval analysis," in *2007 10th International Conference on Information Fusion*, pp. 1–8, 2007.

[108] D. Dubois and H. Prade, *Possibility Theory*, pp. 6927–6939. New York, NY: Springer New York, 2009.

DOCTORAT BRETAGNE LOIRE / MATHSTIC

université angers

Title: On the stochastic filtering of max-plus linear systems

Keywords: Max-Plus Linear Systems, Stochastic Filtering, Set-membership Estimation, Max-Plus Polyhedra

Abstract: A wide range of man-made systems, such as •telecommunications networks; •manufacturing systems; •computer systems; in which evolution is governed by events, typically a signal arrival or the completion of a task, have been called since the early 1970s *"Discrete Event Systems"* (DES) . Among DES, a particular class of systems with synchronization and delay phenomena can be modeled by linear equations in (max,+)- type algebras. This property motivated the development of the so-called Max-Plus Linear (MPL) systems theory. This theory has many analogies with the conventional theory of continuous linear systems and allows in particular to address control problems. The control problem studied in this thesis concerns the state estimation of MPL dynamical systems. Based on recent work on stochastic filtering of MPL systems and inspired by Bayesian filtering, we propose a new approach that is less expensive, but equally efficient. As for classical Bayesian filtering, our approach is two-fold: a prediction step and a correction step. The prediction step developed in this thesis is based on Max-Plus polyhedra. This new approach allows, during the prediction step, to compute the support of the *prior* probability density function (PDF) with a strongly polynomial complexity. We also show that, under certain conditions related to the observability matrix, the correction step, which corresponds to take into account measurement information in order to shrink the support of the *prior* PDF, is computed with quadratic complexity. We are then looking for a solution to improve the performance of the filter. The objective is to evaluate the quality of the prediction with respect to the errors of the measurements induced by the noise. This leads us to introduce a criterion to minimize the estimation error. Finally, we devote the last part to the study of a filtering algorithm for MPL systems in implicit form. We show that this algorithm makes it possible to take into account the dependency problems between the inputs of the system matrices of MPL systems. These dependencies between state variables being intrinsic to MPL systems.

Titre : Sur le filtrage stochastique de systèmes max-plus linéaires

Mot clés : Systèmes Max-Plus Linéaires, Filtrage Stochastique, Estimation Ensembliste, Polyèdres Max-Plus

Résumé : Dans un certain nombre de systèmes conçus par l'homme, tels que •les réseaux de télécommunication ; •les systèmes de production ; •les systèmes informatiques ; dans lesquels l'évolution est régie par des événements ponctuels, typiquement l'arrivée d'un signal ou l'achèvement d'une tâche, sont appelés depuis le début des années 70 *"Systèmes à Événements Discrets"* (SED). Parmi les SED, une classe particulière de systèmes mettant en œuvre des phénomènes de synchronisation et de retards peut être modélisée par des équations linéaires dans les algèbres de type (max,+). Cette propriété a motivé l'élaboration de ce que l'on appelle communément la théorie des systèmes Max-Plus Linéaires (MPL). Cette théorie présente de nombreuses analogies avec la théorie conventionnelle des systèmes linéaires continus et permet notamment d'aborder des problèmes de commandes. Le problème de commande étudié dans cette thèse porte sur l'estimation d'état des systèmes dynamiques MPL. Partant de travaux récents sur le filtrage stochastique des systèmes MPL et en s'inspirant du filtrage Bayésien, nous proposons une nouvelle approche moins coûteuse, mais tout aussi performante. Comme pour le filtrage Bayésien classique, notre approche est composée de deux étapes : une étape de prédiction et une étape de correction. L'étape de prédiction développée dans cette thèse est basée sur les polyèdres Max-Plus. Cette nouvelle approche permet, lors de l'étape de prédiction, de calculer le support de la densité de probabilité (PDF) *a priori* avec une complexité polynomiale. Nous montrons également que, sous certaines conditions liées à la matrice d'observabilité, l'étape de correction, qui correspond à la prise en compte de l'information apportée par la mesure afin de raffiner le support de la PDF *a priori*, est calculable avec une complexité quadratique. Nous cherchons ensuite une solution permettant d'améliorer les performances du filtre. L'objectif est d'évaluer la qualité de la prédiction par rapport aux erreurs des mesures induites par le bruit. Cela nous amène à introduire un critère permettant de minimiser l'erreur d'estimation. Enfin, nous consacrons la dernière partie à l'étude d'un algorithme de filtrage pour les systèmes MPL sous forme implicite. Nous montrons que cet algorithme permet de prendre en compte les problèmes de dépendance entre les entrées des matrices d'état d'un système MPL. Ces dépendances entre les variables d'états étant intrinsèques aux systèmes MPL.