A NOTE ON TIME AND SPACE METHODS IN NETWORK CALCULUS

Jean-Yves Le Boudec and Patrick Thiran
Institute for Computer Communications and their Applications
Swiss Federal Institute of Technology at Lausanne, CH-1015 Lausanne

Abstract

We model some lossless queuing systems arising in guaranteed service networks as non-linear min-plus systems that can be bounded by linear systems. We introduce the distinction between space and time methods for such systems. Space methods are related to min-plus system theory, while time methods borrow from dynamic programming and exploit the causality of the operators. We apply this to the window flow control problem previously studied by Cruz and Okino; Chang; Agrawal and Rajan. We explain the existing bounds by the combination of the space and time methods, and derive another bound that can be tighter.

1 Introduction

A number of recent papers [1, 2, 3, 4] has brought together a set of calculus rules, called network calculus, for networks with guaranteed quality of services. These results extend the original theory of service curves introduced by Cruz in [5] by placing it the general context of min-plus algebra [6]. The results give bounds for such quantities as delays, backlogs in networks which offer guaranteed service, with or without flow control.

The starting point in this paper is the main result in [2], which characterizes the service curve offered by a network with flow control. Similar, but not identical results are in [3, 4]. In this paper, we place the results in [2] in a general min-plus framework, and show how its derivation calls for *space* and *time* methods. Space methods are essentially iterative methods that manipulate complete trajectories; the method we describe here is contained in [6]. In contrast, time methods consider one trajectory viewed over a varying time window. We describe a space method which is quite general and applies to some non-linear models. The time method we describe here is an application of the dynamic programming ideas, or of the Dijkstra algorithm. It applies only to linear and causal models, but it provides finer results. The analysis not only explains in a systematic way the results that were obtained in an ad-hoc fashion in [2], but it also provides another bound, which can be tighter.

The paper starts with some notation and background on recent results on window flow control (Section 2). We consider two examples. Example 1 is from [3] or [4], while Example 2 is from [2]. Both give lower bounds on service guarantees for a network with window flow control, with the result in Example 2 being more detailed. Then we introduce our system modelling concepts and apply them to the two examples (Section 3). We describe the space method (Section 4) and the time method (Section 5), and show that that the result in Example 1 can be explained by the space method, whereas we need a time method to explain, and improve the result in Example 2.

2 Notation and Background

We consider a discrete time system, described by an infinite sequence $\vec{x} = (\vec{x}(0), \vec{x}(1), \ldots, \vec{x}(t), \ldots)$ where t is the time. We assume that for all t the values $\vec{x}(t)$ are in $(\mathbb{R}^+ \cup \{\infty\})^J$, where J is a fixed, finite integer.

For $\vec{z}, \vec{z'} \in (\mathbb{R}^+ \cup \{\infty\})^J$, we define $\vec{z} \wedge \vec{z'}$ as the coordinate-wise minimum of $\vec{z}$ and $\vec{z}'$, and similarly for the $+$ operator. We write $\vec{z} \leq \vec{z'}$ with the meaning that $\vec{z}_j \leq \vec{z'}_j$ for $j = 1 \ldots J$. Note that the comparison so defined is not a total order, namely, we cannot guarantee that either $\vec{z} \leq \vec{z'}$ or $\vec{z'} \leq \vec{z}$ holds. For a constant K, we note $\vec{z} + K$ the vector defined by adding K to *all* elements of $\vec{z}$.

For sequences, we note similarly $(\vec{x} \wedge \vec{y})(t) = \vec{x}(t) \wedge \vec{y}(t)$ and $(\vec{x} + K)(t) = \vec{x}(t) + K$ for all $t \in \mathbb{N}$, and write $\vec{x} \leq \vec{y}$ meaning that $\vec{x}(t) \leq \vec{y}(t)$ for all t.

We further call $\mathcal{F}_J$ the set of sequences $\vec{x}$ that are wide-sense increasing and non-negative, namely, for which $0 \leq x_j(0) \leq x_j(1) \leq x_j(2) \leq \ldots$ for all $j = 1 \ldots J$.

We denote vectors with the arrow symbol as in $\vec{x}$ or with a non-arrowed greek letter.

The min-plus convolution operation, which we note $\otimes$ is defined as follows [6, 1, 4, 3].

$$(\vec{x} \otimes \vec{y})(t) = \inf_{u \text{ such that } 0 \leq u \leq t} \{\vec{x}(u) + \vec{y}(t-u)\}.$$

Note that $(\vec{x} \otimes \vec{y}) + K = (\vec{x} + K) \otimes \vec{y} = \vec{x} \otimes (\vec{y} + K)$.

In network calculus, when the dimension $J = 1$, $\vec{x}(t) = x(t)$ is for example the number of bits, or

ATM cells, counted from time 0 to t at a given observation point. We say that a network element guarantees to a flow $x(t)$ a service curve β if the output $y(t)$ satisfies $y \geq \beta \otimes x$. For example, an ideal constant bit rate server with rate c offers a service curve defined by $\beta(t) = ct$. More realistically, the IETF assumes that guaranteed service nodes offer a service curve $\beta(t) = R(t-T)^+$ [1]. Similarly, we say that a flow, described by $x(t)$, is constrained by α if $x \leq \alpha \otimes x$ [3, 2, 1, 4]. We say that α is an arrival curve for the flow. The concept of arrival curve generalizes that of leaky buckets. Service and arrival curves are the key concept for modelling guaranteed service schedulers and computing useful bounds.

We consider here the following two examples.

Example 1 This example is found independently in [3] and [4]. A data flow $a(t)$ is fed via a window flow controller to a network offering a service curve of β. The window flow controller limits the amount of data admitted into the network in such a way that the total backlog is less than or equal to K, where K (the window size) is a fixed number (Figure 1).

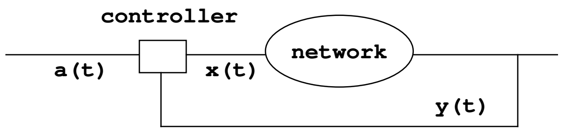

Figure 1: Example 1, from [3] or [4]

Call $x(t)$ the flow admitted to the network, and $y(t)$ the output. The definition of the controller means that $x(t)$ is the maximum solution to

$$\begin{cases} x \leq a \\ x \leq y + K \end{cases} \tag{1}$$

which implies that $x(t) = a \wedge (y+K)(t)$. Note that we do not know the mapping $x(t) \to y(t)$, but we do know that

$$y \geq \beta \otimes x \tag{2}$$

In [3], this used to derive that

$$x \geq \overline{(\beta + K)} \otimes a \tag{3}$$

In the formula, $\overline{(\beta + K)}$ is the sub-additive closure of $\beta + K$ [3, 1, 6]. The sub-additive closure $\overline{\alpha}$ of a vector or function α is defined by

$$\overline{\alpha}(t) = \inf\{\delta, \alpha, \alpha^{(2)}, \alpha^{(3)}, \dots\}$$

with $\alpha^{(i)} = \alpha \otimes \dots \otimes \alpha$ (i times) and δ is the fixed function defined by $\delta(t) = \infty$ for $t \geq 1$ and $\delta(0) = 0$.

Equation (3) means that the complete system offers a service curve equal to $\overline{(\beta + K)}$.

We show in this paper that this result is indeed obtained by a space method.

Example 2 This example is a more detailed representation of window-flow control. Compared to example 1, the additional modelling assumptions in [2] can be re-formulated as follows (see Figure 2 for the notation):

- the output of the window flow controller (marked 1 on the figure) is constrained by an arrival curve β.
- the output of network elements 1 and 2 is constrained by a maximum service rate . More precisely, the number of bits output at station i ($i = 1, 2$) during time interval $(s, t]$ is bounded by $M_i(t) - M_i(s)$ for some fixed functions M_1 and M_2. This models the fact that the server is busy serving other flows.

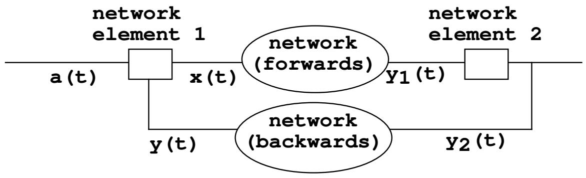

Figure 2: Example 2, from [2]

As with Example 1, the controller guarantees that a maximum of K bits are backlogged in the loop. Network elements f and b are assumed to offer service curves S_f and S_b.

With these assumptions, the admitted flow x is the maximum solution to the system

$$\begin{cases} x \leq a \\ x \leq \beta \otimes x \\ x(t) \leq \inf_{0 \leq u \leq t} \{x(u) + M_1(t) - M_1(u)\} \\ x \leq y + K \end{cases} \tag{4}$$

As with Example 1, we do not know the exact mapping $x(t) \to y(t)$, but we do know that

$$\begin{cases} y_1 \leq x \\ y_1 \geq S_f \otimes x \\ y_2 \leq y_1 \\ y_2(t) \leq \inf_{0 \leq u \leq t} \{y_2(u) + M_2(t) - M_2(u)\} \\ y \leq y_2 \\ y \geq S_b \otimes y_2 \end{cases} \tag{5}$$

and that $\vec{x}$, $\vec{y_2}$ is the maximum couple of functions such that (4) and (5) hold.

Additional assumptions made in [2] are

$$\begin{cases} \beta(u) \leq C_1 u \\ \hat{S}_i(t-s) \leq M_i(t) - M_i(s) \leq (t-s)C_i, \end{cases} \tag{6}$$

with $i = 1, 2$. In (6), C_1 and C_2 represent the maximum line rates at network elements 1 and 2, while functions $\hat{S}_1$ and $\hat{S}_2$ give minimum guarantees on the service rates. With these assumptions, Cruz and Okino derive in [2] a service curve S_1 for network element 1, namely, they find an S_1 such that $x \geq S_1 \otimes a$.

We give more detail about S_1 in Section 4 and show that it can be obtained by a combination of time and space methods. Further, we are able to give another bound S_e.

3 System Modelling

3.1 General Model

In this Section we show how to use concepts from min-plus algebra in order to model problems such as in the two examples.

For operators $\Pi : \mathcal{F}_J \to \mathcal{F}_J$ we define the following properties, which are direct applications of [6]:

Definition 1 ([6]) • Π *is* isotone *if* $\vec{x}(t) \le \vec{y}(t)$ *for all* t *always implies* $\Pi(\vec{x})(t) \le \Pi(\vec{y})(t)$ *for all* t.

- Π *is* causal *if for all* t, $\Pi(\vec{x})(t)$ *depends only on* $\vec{x}(0), \dots, \vec{x}(t)$
- Π *is* upper-semi-continuous *if for any decreasing sequence of sequences* $(\vec{x}^i(t))_i$ *we have* $\inf_i \Pi(\vec{x}^i) = \Pi(\inf_i \vec{x}^i)$.

  In extension, the condition means that whenever we have $\vec{x}_j^{i+1}(t) \le \vec{x}_j^i(t)$ *for all* i, j, t, *then we must have* $\Pi(\vec{x_j^*})(t) = \inf_i \Pi(\vec{x}_j^i)(t)$ *for all coordinate* j *and time* t, *where* $\vec{x^*}$ *is defined by* $\vec{x^*}_j(t) = \inf_i \vec{x}_j^i(t)$.
- Π *is* time-invariant *if* $\vec{y}(t) = \Pi(\vec{x})(t)$ *for all* t *and* $\vec{x}'(t) = \vec{x}(t+s)$ *for some* s *always implies that for all* t $\Pi(\vec{x}')(t) = \vec{y}(t+s)$.

We propose to model network elements as *isotone, causal, upper-semicontinuous* operators. The first two properties are intuitive. We discuss now the third one. It is a technical assumption required for the theorems in Section 4 to hold. However if the operator Π is causal and if, for all time t, there exists a *finite* set E_t such that $\Pi(\vec{x})(s) \in E_t$ for all $s \le t$, then Π is necessarily upper-semi-continuous. The latter assumption is true in practice since we count bits or cells and time is discrete. Therefore, our modelling assumptions are not a practical restriction.

For Π and Π' we note $\Pi \circ \Pi'$ the compound operator, defined by $(\Pi \circ \Pi')(\vec{x}) = \Pi[\Pi'(\vec{x})]$. We also note $\Pi^{(i)} = \Pi \circ \ldots \circ \Pi$ (i times, $i \ge 1$).

We will use the following definition in Section 4.

Definition 2 ([6]) *The* closure $\overline{\Pi}$ *of the operator* Π *is defined by*

$$\overline{\Pi}(\vec{x}) = \inf \left\{ \vec{x}, \Pi(\vec{x}), \Pi^{(2)}(\vec{x}), \dots, \Pi^{(i)}(\vec{x}), \dots \right\}$$

Lastly, we write $\Pi \le \Pi'$ to express that $\Pi(\vec{x}) \le \Pi'(\vec{x})$ for all $\vec{x} \in \mathcal{F}_J$.

Proposition 1 *If* Π *and* Π' *are isotone and* $\Pi \le \Pi'$ *then* $\overline{\Pi} \le \overline{\Pi'}$

The proof is simple and left to the reader.

3.2 Min-Plus linear operators

We also define *min-plus linear* operators:

Definition 3 ([6]) *Operator* Π *is* min-plus linear *if it is upper-semi-continuous and* $\Pi(\vec{x}+K) = \Pi(\vec{x})+K$ *for all constant* K.

Min-plus operators are the equivalent in min-plus algebra of traditional linear system theory. In particular, it is shown in [6] that an operator is min-plus linear if and only if it can be represented under the form

$$\Pi(\vec{x})(t) = \inf_u \{H(t,u) + \vec{x}(u)\}$$

which we denote as with standard matrix theory $\Pi(\vec{x}) = H \cdot \vec{x}$. H is called the matrix representation of the linear operator Π. Note that for any fixed t, s, $H(t,s)$ has J coordinates $H_j(t,s)$.

In general, a network element cannot be assumed to be a min-plus linear operator on its input. A notable exception is the case of shapers. Here $J = 1$ again and a shaper, with shaping curve σ, is a system that forces an input flow $a(t)$ flow to have an output $x(t)$ which has σ as arrival curve, at the expense of possibly delaying bits in a buffer. A shaper is linear and time invariant, with $x = \sigma \otimes a$ [3, 4]. We will see later that linear operators can be used to obtain bounds, even when the system is not linear.

The composition of operators translates into min-plus matrix multiplication, namely, if Π and Π' are linear, with matrices H and H', then the compound operator $\Pi \circ \Pi'$ is also linear, with matrix $H \cdot H'$, defined by

$$(H \cdot H')(t,s) = \inf_u \{H(t,u) + H'(u,s)\}.$$

For a linear operator with matrix H, being causal is equivalent to $H_j(t,s) = \infty$ for $s > t$ and for all coordinates j.

As with standard system theory, if Π is time invariant, causal and min-plus linear, then there exists some $\beta \in \mathcal{F}_J$ such that $\Pi(\vec{x}) = \beta \otimes \vec{x}$ [6, 3, 4]. We say that Π is the convolution by β operator and note $\Pi = \mathcal{C}_\beta$. Note that in that case the closure $\overline{\Pi}$ is also time invariant, causal and min-plus linear, with $\overline{\Pi}(\vec{x}) = \overline{\beta} \otimes \vec{x}$. In the formula, $\overline{\beta}$ is the sub-additive closure of β [1, 3].

We introduce the following linear, causal but non-time invariant family of operators, which we will use to model Examples 1 and 2.

Definition 4 *For a given* $\vec{\alpha} \in \mathcal{F}_J$, *with* $\alpha_j(0) = 0$ *for all coordinates* j, *define the min-plus linear operator* $h_{\vec{\alpha}}$ *by*

$$h_{\vec{\alpha}}(\vec{x})(t) = \inf_{u \text{ such that } 0 \le u \le t} \{\vec{\alpha}(t) - \vec{\alpha}(u) + \vec{x}(u)\}$$

In other words, $h_{\vec{\alpha}}$ is defined by the matrix H with

$$\begin{cases} H(t,s) = \vec{\alpha}(t) - \vec{\alpha}(s) & \text{if} \quad s \le t \\ H_j(t,s) = \infty & \text{if} \quad s > t,\ j = 1 \ldots J \end{cases}$$

It can easily be shown that the $h_{\vec{\alpha}}$ operators are idempotent, namely:

$$h_{\vec{\alpha}} \circ h_{\vec{\alpha}} = h_{\vec{\alpha}} \tag{7}$$

and since $\vec{\alpha}(0) = \vec{0}$, we have $h_{\vec{\alpha}}(\vec{x}) \leq \vec{x}$; thus

$$\overline{h_{\vec{\alpha}}} = h_{\vec{\alpha}} \tag{8}$$

3.3 Application to Examples 1 and 2

Consider first Example 1. Here $J = 1$. Define Π as the operator that maps $x(t)$ to $y(t)$. From Equation (1), we derive that $x(t)$ is the maximum solution to

$$x \leq a \wedge (\Pi(x) + K) \tag{9}$$

The operator Π can be assumed to be isotone, causal and upper-semi-continuous, but not necessarily linear. However, we know that $\Pi \geq \mathcal{C}_\beta$. We will exploit this formulation in Section 4.

Consider now Example 2. Define Π_f as the one-dimensional operator that maps $x(t)$ to $y_1(t)$ and Π_b the one-dimensional operator that maps $y_2(t)$ to $y(t)$. From Equation (4), we derive that $(x(t), y_2(t))$ is the maximum sequence such that

$$\left\{ \begin{array}{lcl} x & \leq & a \wedge (\beta \otimes x) \wedge h_{M_1}(x) \wedge (\Pi_b(y_2) + K) \\ y_2 & \leq & \Pi_f(x) \wedge h_{M_2}(y_2). \end{array} \right. \tag{10}$$

In Section 4 we will show the existence of such a maximum. Here we have thus $J = 2$. Let $\vec{z} = (x, y_2)$ and define the non-linear operator $\vec{\Pi}(\vec{z}) = (\Pi_1(\vec{z}), \Pi_2(\vec{z}))$ by

$$\left\{ \begin{array}{lcl} \Pi_1(\vec{z}) & = & (\beta \otimes x) \wedge h_{M_1}(x) \wedge (\Pi_b(y_2) + K) \\ \Pi_2(\vec{z}) & = & \Pi_f(x) \wedge h_{M_2}(y_2). \end{array} \right.$$

The problem in Example 2 is thus equivalent to finding the maximum solution to the problem $\vec{z} \leq \vec{\Pi}(\vec{z}) \wedge (a, \infty)$.

4 Space Method

In this Section we apply results from [6] to the problems formulated in the previous section.

Theorem 1 ([6], Theorem 4.70, item 6) *Let Π be an operator $\mathcal{F}_J \to \mathcal{F}_J$, and assume it is isotone and upper-semi-continuous. For any fixed function $\vec{a} \in \mathcal{F}_J$, the problem*

$$\vec{x} \leq \Pi(\vec{x}) \wedge \vec{a} \tag{11}$$

has one maximum solution, given by $\vec{x} = \overline{\Pi}(\vec{a})$

The theorem is proven in [6], though with some large amount of notation, using the fixed point method. It can also easily be proven directly as follows. Consider the sequence of decreasing sequences defined by $\vec{x}^0 = \vec{a}$ and $\vec{x}^{n+1} = \Pi(\vec{x}^n) \wedge \vec{x}^n, n \geq 0$. Then $\vec{x}^* = \inf_n \vec{x}^n$ is a solution of (11) because Π is upper-semi-continuous. Conversely, if $\vec{x}$ is a solution, then $\vec{x} \leq \vec{x}^n$ for all n because Π is isotone and thus $\vec{x} \leq \vec{x}^*$.

We call the application of this theorem the space method, because it is based on an iteration on complete sequences $x(t)$. The space method does not require the operators to be causal.

Let us apply the theorem to Example 1. We know now that (9) has one maximum solution and that it is given by

$$x(t) = \overline{(\Pi + K)}(a)(t)$$

Now from (2) we have $\Pi(x) + K \geq \beta \otimes x + K$. From Proposition 1, we have:

$$x \geq \overline{(\beta + K)} \otimes a$$

which is Equation (3).

Now let us consider Example 2. It follows easily from the problem formulation in (10) and the isotony of the operators that if $(x(t), y_2(t))$ is the maximum solution to (10), then $y_2(t)$ is also the maximum the solution to

$$y_2 \leq \Pi_f(x) \wedge h_{M_2}(y_2)$$

and thus, by application of Theorem 1,

$$y_2 = \overline{h_{M_2}} \circ \Pi_f(x)$$

and thus from (8)

$$y_2 = h_{M_2} \circ \Pi_f(x)$$

and thus $x(t)$ is a solution to the one-dimensional problem

$$x \leq a \wedge (\beta \otimes x) \wedge h_{M_1}(x) \wedge (\Pi_b \circ h_{M_2} \circ \Pi_f(x) + K). \tag{12}$$

Thus we can conclude again from Theorem 1 that

$$x = \overline{Q}(a)$$

with the operator Q defined by

$$Q(x) = (\beta \otimes x) \wedge h_{M_1}(x) \wedge (\Pi_b \circ h_{M_2} \circ \Pi_f(x) + K).$$

From (5) and (6) we can bound Q from below by

$$Q(x) \geq (G \otimes x) \wedge h_{M_1}(x) \tag{13}$$

with G defined as follows [2]. First let $G_0 = (S_b \otimes \hat{S}_2 \otimes S_f + K) \wedge \beta$. Then define

$$\Delta = \inf\{t \geq 0 : G_0(t) \leq C_1 t\}$$

and let $G(t) = G_0(t)$ if $t \geq \Delta$ and $G(t) = G_0(\Delta)$ otherwise. Equation (13) can be re-written $Q \geq \mathcal{C}_G \wedge h_{M_1}$. Thus, from Proposition 1:

$$\overline{Q} \geq \overline{\mathcal{C}_G \wedge h_{M_1}} \tag{14}$$

We can further exploit (14) by using the bound

$$h_{M_1}(x) \geq \hat{S}_1 \otimes x$$

and we obtain from (14):

$$\overline{Q} \geq \overline{\mathcal{C}_G \wedge \mathcal{C}_{\hat{S}_1}} = \overline{\mathcal{C}_{G\wedge\hat{S}_1}} = \mathcal{C}_{\overline{G\wedge\hat{S}_1}}. \qquad (15)$$

This shows that a service curve for network element 1 on Figure 2 is given by

$$S_1'(t) = \overline{G \wedge \hat{S}_1}(t). \qquad (16)$$

However, the service curve in (16) is weaker than the service curve S_1 found in [2]. Indeed, S_1 is given by

$$S_1 = \inf\left\{\hat{S}_1, \hat{S}_1^{(2)} \otimes G, \hat{S}_1^{(3)} \otimes G^{(2)}, \dots\right\} \qquad (17)$$

whereas S_1' can be expanded as

$$S_1' = \inf_{m,n\geq 1}\left\{\hat{S}_1^{(m)} \otimes G^{(n)}\right\}.$$

For example, if $\hat{S}_1(t) = R(t-T)^+$, then $S_1' = 0$ whereas $S_1 \neq 0$. We show in the next Section how to obtain S_1 with the time method.

5 Time Method

The results in the previous section apply to the general case of a multi-dimensional, non-linear isotone, upper-semi-continuous operator. This has allowed us to bound our network model by a linear system. The method we propose in this section goes one step further, but applies only to linear, causal operators.

5.1 The Time Method Theorem

Theorem 2 (Time Method) *Consider a min-plus linear and causal operator $\mathcal{F}_J \to \mathcal{F}_J$, with matrix representation $H \geq 0$. The maximum solution $\vec{x}^*$ to the problem*

$$\vec{x} \leq \Pi(\vec{x}) \wedge \vec{a} \qquad (18)$$

is given by

$$\begin{array}{l} \vec{x}^*(0) = \vec{a}(0) \\ \vec{x}^*(t) = \vec{a}(t) \wedge \inf_{0\leq u\leq t-1}\{H(t,u) + \vec{x}^*(u)\}. \end{array} \qquad (19)$$

Proof: Note that the existence of a maximum solution is given by Theorem 1. By considering each of the J coordinates of $\vec{x}$, we see that it is sufficient to consider the case $J=1$, which we do now (and consequently drop the arrows). Define x^* by the recursion in the Theorem. From $H \geq 0$ it follows easily by induction that x^* is a solution to problem (18). Conversely, for any solution x, $x(0) \leq a(0) = x^*(0)$ and if $x(u) \leq x^*(u)$ for all $0 \leq u \leq t-1$, it follows that $x(t) \leq x^*(t)$ which shows that x^* is the maximal solution. □

Corollary 1 *Consider a min-plus linear and causal operator $\mathcal{F}_J \to \mathcal{F}_J$, with matrix representation $H \geq 0$. The closure $\overline{H}$ of H satisfies*

$$\text{if } s<t \text{ then } \overline{H}(t,s) = \inf_{s\leq u\leq t-1}\{H(t,u) + \overline{H}(u,s)\} \qquad (20)$$

Note that $\overline{H} = H \cdot \overline{H}$, as shown later in the following section. The contribution of the corollary is that the inf in formula (20) is for all $u \leq t-1$ instead of $u \leq t$.

Proof: From Theorem 1, for a fixed s, $\overline{H}(.,s)$ is the maximum solution to problem (18), with $\vec{a}$ defined for all coordinate j by

$$\begin{cases} \vec{a}_j(t) = \infty & \text{if } t \neq s \\ \vec{a}_j(s) = 0 \end{cases}$$

The corollary follows now immediately from Theorem 2. □

In the case $J=1$, we can interpret $\overline{H}(t,s)$ as the distance between vertices s and t on a graph with edge costs given by matrix H. The corollary is then a simple recursion, similar to the ones used in dynamic programming or in Dijkstra's shortest path algorithm.

5.2 Application to the closure $\overline{H_1 \wedge H_2}$

Now let us move on to an application that underlies Example 2. From now on $J=1$. We start from Eq. (14), which gives a lower bound on the service offered by station 1. The main fact which is implicitely exploited in [2] is that $h_{M_1} \circ h_{M_1} = h_{M_1}$. This fact can be exploited in general via the following result.

Theorem 3 *Consider two min-plus linear and causal operators $\mathcal{F}_1 \to \mathcal{F}_1$, with matrix representations $H_1 \geq 0$ and $H_2 \geq 0$, and let $H = H_1 \wedge H_2$.*

Then, for all $s<t$, there exists an integer $k \geq 1$, a sequence of positive *integers $n_1, \dots, n_k$, a sequence of times $u_0 = t > u_1 > \dots > u_{k-1} > u_k = s$ and a sequence $i(k)$ of alternating 1s and 2s such that*

$$\begin{aligned} \overline{H}(t,s) \quad \geq \quad & H_{i(1)}^{(n_1)}(t,u_1) \\ & + H_{i(2)}^{(n_2)}(u_1,u_2) + \dots + H_{i(k)}^{(n_k)}(u_{k-1},s). \end{aligned}$$

The theorem is for the dimension $J=1$. It gives a decomposition of $\overline{H}(t,s)$ along the time axis as an alternating sequence of powers of H_1 and H_2. The sequence $i(m)$, $1 \leq m \leq n$ in the theorem is such that $i(2m+1) = i(1)$ and $i(2m) = 3 - i(1)$ for integer values of m.

Proof: From Corollary 1, there exists v_1, with $s \leq v_1 < t$ such that

$$\overline{H}(t,s) = H(t,v_1) + \overline{H}(v_1,s)$$

Now we always have that $\overline{H}(s,s)=0$, so that if $v_1 = s$ then the theorem is proven. Otherwise, we apply the same procedure to v_1 and finally get a sequence $v_n = s < v_{n-1} < \ldots < v_1 < t$ such that

$$\overline{H}(t,s) = H(t,v_1) + H(v_1,v_2) + \ldots + H(v_{n-1},s)$$

Now $H(v_m,v_{m+1}) = H_1(v_m,v_{m+1})$ or $H(v_m,v_{m+1}) = H_2(v_m,v_{m+1})$. By grouping together the time instants v_m that for which the former condition (respectively the latter) is true, we obtain the theorem. $\square$

5.3 Application to Example 2

We now show how the time method is able to explain, and give another bound that can improve the best bound S_1 in [2] which we already mentioned in Equation (17).

From Equation (14) we can lower-bound the non-linear operator Q that defines network element 1 by the closure of H, where $H = \mathcal{C}_G \wedge h_{M_1}$. Remember that $\mathcal{C}_G$ is the convolution operator by G while h_{M_1} is defined by $h_{M_1}(x)(t) = \inf_{0\le u\le t}\{M_1(t) - M_1(u) + x(u)\}$.

We can now apply Theorem 3 with $H_1 = h_{M_1}$ and $H_2 = \mathcal{C}_G$. We have $H(t,t)=0$ because this is true for h_M. Now the main fact us $h_{M_1}^{(n)} = h_{M_1}$. In addition to this, note that $x \ge h_{M_1}(x)$ for all x so we can combine the various cases for sequence $i(n)$ in Theorem 3 into:

$$\overline{H}(t,s) \ge h_{M_1} \circ \mathcal{C}_{G^{(n_{2p})}} \circ h_{M_1} \ldots \circ \mathcal{C}_{G^{(n_2)}} \circ h_{M_1} \tag{21}$$

for some integer $p \ge 0$. Now $h_{M_1} \ge \mathcal{C}_{\hat{S}_1}$ thus

$$\overline{H}(t,s) \ge \hat{S}_1^{(p+1)} \otimes G^{(n_2+\ldots+n_{2p})}(t-s) \tag{22}$$

We have thus obtained a first service curve formulation. Now $\hat{S}_1^{(i)} \ge \hat{S}_1^{(i')}$ if $i \le i'$ because $\hat{S}_1(0)=0$. Let $m = n_2 + \ldots + n_{2p}$; from Equation (22) we get

$$\overline{H}(t,s) \ge \hat{S}_1^{(m+1)} \otimes G^{(m)}(t-s) \ge S_1$$

where S_1 is the service curve from [2], given in Equation (17). This proves how we can obtain the bound in [2] from the time method.

We are now also able to derive another bound. Equation (21) can be rewritten as

$$\begin{aligned}\overline{H}(t,s) \;\ge\; & M_1(t) - M_1(u_{2p}) + G^{(n_{2p})}(u_{2p} - u_{2p-1}) \\ & + M_1(u_{2p-1}) - M_1(u_{2p-2}) \\ & + G^{(n_{2p-2})}(u_{2p-2} - u_{2p-3}) + \ldots \\ & + M_1(u_3) - M_1(u_2) + G^{(n_2)}(u_2 - u_1) \\ & + M_1(u_1) - M_1(s) \end{aligned} \tag{23}$$

Now from (6) we have

$$\begin{aligned} & M_1(t) - M_1(u_{2p}) + M_1(u_{2p-1}) - M_1(u_{2p-2}) + \ldots \\ & \quad + M_1(u_1) - M_1(s) \ge M_1(t) - M_1(s) - C_1 u \end{aligned}$$

where $u = u_{2p} - u_{2p-1} + \ldots + u_2 - u_1$. This shows that

$$\overline{H}(t,s) \ge \hat{S}_1(t-s) + \inf_{n\ge 0, n\le u\le t-s}\{G^{(n)}(u) - C_1 u\}$$

Define

$$S_e(v) = \hat{S}_1(v) + \inf_{n\ge 0, 0\le u\le v \text{ and } n\le u}\{G^{(n)}(u) - C_1 u\} \tag{24}$$

We have shown that a service curve for network element 1 is S_e. In general, S_e is better if the delay introduced by the feedback loop in Figure 2 is large compared to the delay parameter of $\hat{S}_1$. Figure 3 shows the values of S_e and S_1 for one example.

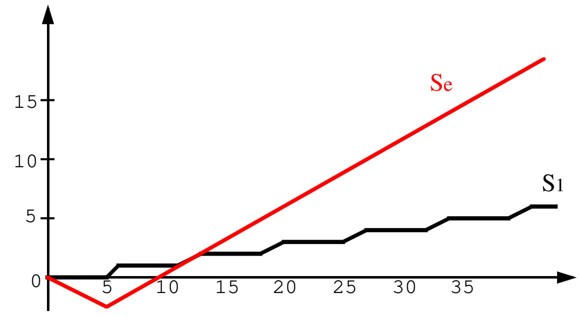

Figure 3: The service curves S_e and S_1 on one example. Here $K=\infty$, $\beta(t) = \inf(C_1 t, b + rt)$, $\hat{S}_1(t) = (C_1 t - T_1)^+$, with $T_1 = 5$, $C_1 = 1$, $r = 0.5$ and $b = 0.5$.

References

[1] J.-Y. Le Boudec. 'Network calculus made easy', Technical report 96/218, EPFL-DI, Dec. 1996, http://lrcwww.epfl.ch/PS_files/d4paper.ps, to appear in *IEEE TIT*.

[2] R.L. Cruz and C.M. Okino. 'Service guarantees for a flow control', Preprint, a first version also in 34th Allerton Conf. on Comm., Cont., and Comp. Monticello, IL, Oct. 1996.

[3] C.S. Chang. 'On deterministic traffic regulation and service guarantee: A systematic approach by filtering', in *Proc. Infocom'97*, Kobe, Japan, April 1997.

[4] R. Aggrawal and R. Rajan. 'Performance bounds for guaranteed and adaptive services', Technical report RC 20649, IBM, Dec. 1996.

[5] R. L. Cruz. 'Quality of service guarantees in virtual circuit switched networks', *IEEE JSAC*, pp. 1048–1056, August 1995.

[6] F. Baccelli, G. Cohen, G. J. Olsder and J.-P. Quadrat. *Synchronization and Linearity, An Algebra for Discrete Event Systems*, John Wiley and Sons, August 1992.