# *Stochastic production trees as products of i.i.d. componentwise exponential max-plus matrices*

Hook, James and Broomhead, Dave

2011

MIMS EPrint: **2011.69**

Manchester Institute for Mathematical Sciences
School of Mathematics

The University of Manchester

Reports available from: http://eprints.maths.manchester.ac.uk/
And by contacting: The MIMS Secretary
School of Mathematics
The University of Manchester
Manchester, M13 9PL, UK

ISSN 1749-9097

# Stochastic production trees as products of i.i.d. componentwise exponential max-plus matrices **DRAFT**

James Hook, Dave Broomhead
School of Mathematics, University of Manchester

## Abstract

We introduce a class of stochastic production tree model, based on Petri nets, which admit a random matrix product description in the Max-plus algebra. With a kind of combinatorial change of variables we are able to simplify the form of the matrices arising from these models. For this class of *Componentwise exponential* matrix we prove a new result relating the (Max-plus) spectrum of the product to the principal (classical) eigenvalue of an associated adjacency matrix by means of a sandwich inequality. This theorem highlights several important theoretical factors in the dynamics of Max-plus linear systems generally and gives us some neat insight into the different production tree models.

## Introduction

The Max-plus algebra gives us an alternative way to look at a number of interesting classically non-linear phenomena [1, 2, 3, 4, 5, 6]. In particular there are many quite different dynamical systems/mathematical constructions which are quite intractable with standard algebra but become linear or in some other way simpler when expressed in Max-plus.

Dynamical systems whose variables are the starting times of different interacting events fit quite naturally into the Max-plus linear systems framework. If some event $i$ can only reoccur for the $(n+1)$th time $M_{i,j}$ seconds after event $j$ has occurred for the $n$th time then $x_i(n+1)$, the time $i$ occurs for the $(n+1)$th time, satisfies

$$x_i(n+1) = \max_j M_{i,j} + x_j(n) = [M \otimes \underline{x}(n)]_i \tag{1}$$

where $\otimes$ stands for Max-plus multiplication. The theory for Max-plus matrix algebra follows in analogy to the classical case; there are eigenvalues and eigenvectors, determinants, the Caley-Hamilton theorem holds and much more besides [6]. For an introduction to Max-plus algebra see [1].

This paper focuses on stochastic queuing systems which can be thought of as very high dimensional continuous time Markov chains. Although they are classically linear

their dimension makes them intractable, and investigating any classical structure in these models seems to be beyond the scope of standard Markov chain methods. We are however to describe them completely with much lower dimensional Max-plus linear models whose underlying structure can be studied more easily.

We show that products of i.i.d. Max-plus matrices are dominated by a Max-plus exponent whose value can be attributed to the weight of a path through the vertex set of the matrices. By taking this path-centric viewpoint we are able to give a new neater proof of the Max-plus multiplicative ergodic theorem and prove a new result relating the Max-plus exponent to the classical principal eigenvalue of an associated adjacency matrix.

In section 1 we introduce petri-nets and explain how stochastically timed event graphs can be described by products of i.i.d. max-plus matrices. In 1.1 we introduce the main production line example. In section 2 we prove the max-plus multiplicative ergodic theorem then introduce componentwise exponential max-plus matrices, which are arguably the simplest possible non-trivial random max-plus matrices with arbitrary graph topology, and prove my main result; a sandwich of bounds for the max-plus exponent based on the principal eigenvalue of an associated graph. In section 3 we generalize the first example to a production tree, show how these systems can be recast in a componentwise exponential form and then look at their exponents the theorem allows us to rigourously investigate the asymptotic behavior of the production tree systems.

# 1 Stochastic timed event graphs

Petri-nets are a modeling language that can be used to describe a wide variety of distributed systems [7]. Formally a petri-net is a directed bipartite graph in which the vertices represent transitions (signified by bars, which represent some events that can occur) and places (signified by circles, which represent some conditions or processes). Places can contain one or more tokens which should be thought of as units of information or of some product. Each token can be in a waiting state (signified by a grey dot in the place) or ready state (signified by a black dot in the place). Waiting tokens become ready after some waiting-time which is particular to the place that the token occupies and can be a fixed non-negative number or random variable. The dynamics of the petri-net move tokens from place to place.

Each transition that succeeds some place vertices can become enabled. A transition is enabled if and only if each place it is a successor of contains at least one ready token. When a transition is enabled it is able to fire. When a transition fires a ready token is removed from each place it succeeds and a waiting token is added to each place it is a predecessor of.

Note that since it is possible to have two enabled transitions only one of which is able to fire, the evolution of a petri net is not necessarily deterministic even if the place waiting-times are constant. See fig 1.

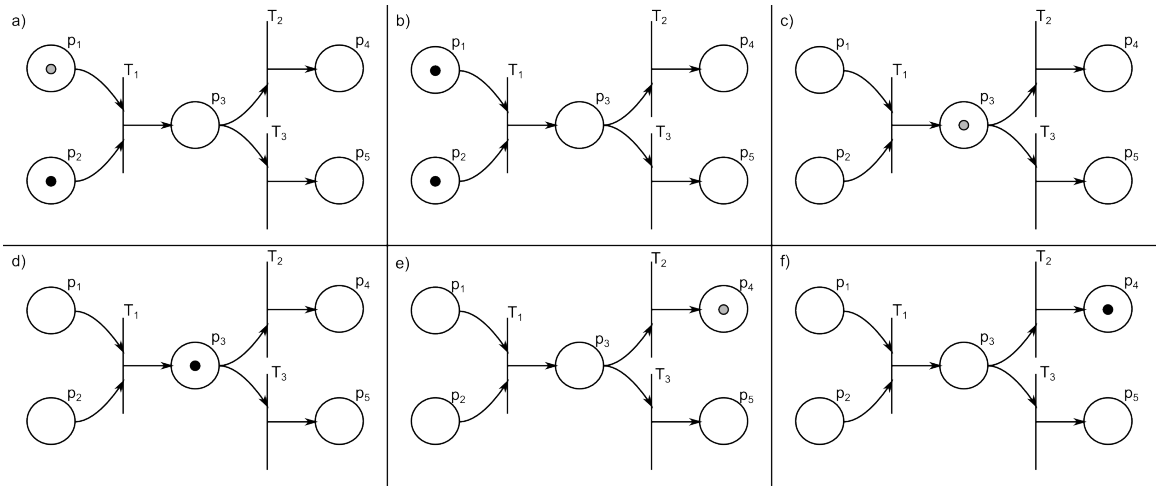


Figure 1: Evolution of a simple petri-net. a) The token in $p_1$ must wait some time to become ready, b) Now that there are ready tokens in $p_1$ and $p_2$ the transition $T_1$ is enabled, c) On firing $T_1$ takes the ready tokens from $p_1$ and $p_2$ and adds a waiting token to $p_3$, d) Now that the token in $p_3$ is ready both transitions $T_2$ and $T_3$ are enabled, e) In this example we choose $T_2$ to fire taking the token from $p_3$ and adding a waiting token to $p_4$, f) The token in $p_4$ is now ready.

**Definition** A *stochastic timed event graph* is a petri net with the following properties

- Each place has exactly one predecessor and one successor transition.
- The successive waiting times at a place $i$ are an i.i.d. sequence $[t_i(n)]_{n=1}^{\infty}$ independent of the waiting times at all other places.

The first condition guarantees that the only non-determinism in the system comes from the random waiting times and not from some 'higher' controller as in the example. The second condition enables us to describe the evolution of the dynamics with a product of i.i.d max-plus matrices.

## 1.1 Example - Asynchronous production line

The following stochastic timed event graph can be used to model a fairly general asynchronous production line. Consider a plant where we have a production line consisting of $N$ sites where each item being produced must pass from one site to the next undergoing a different process at each site, each taking a random period of time. Only one item may occupy a site at a time so that if one site is taking a long time to process an item a queue may build up behind it.

Therefore whenever an item arrives at a site the processing begins. When it is processed the item is either instantaneously moved for processing at the next site, or if the next site is occupied, must wait until the next site is unoccupied when it is instantaneously moved on.

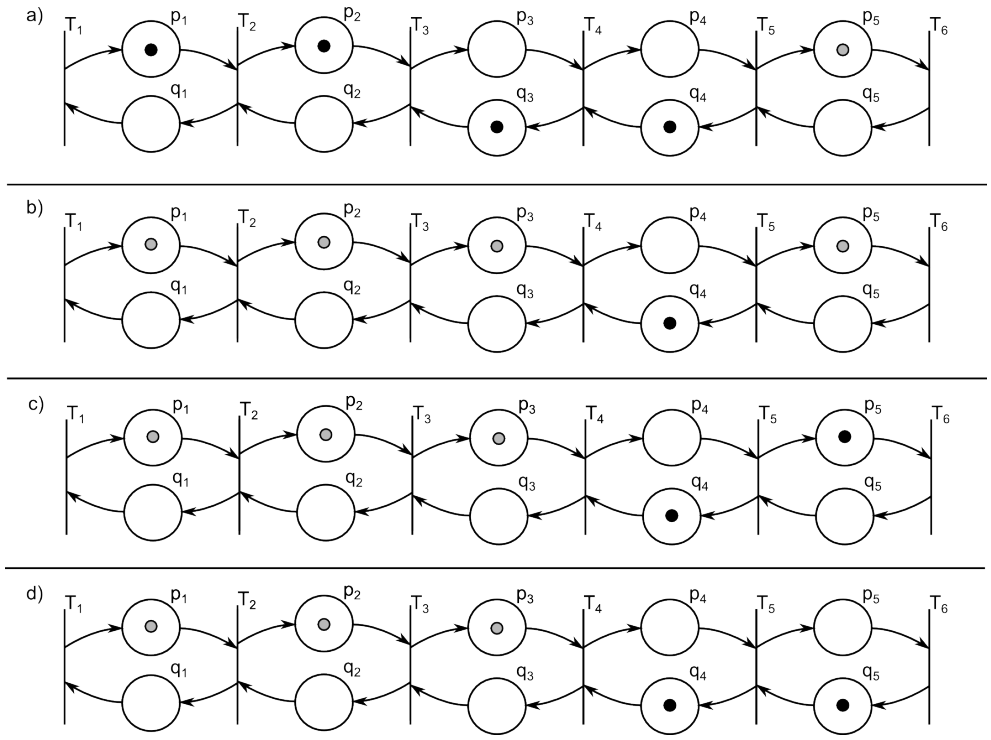


Figure 2: Evolution of asynchronous production line for $N = 5$. a) Transition $T_3$ is enabled, b) Transition $T_3$ fires instantaneously enabling $T_2$ which fires instantaneously enabling $T_1$ which also fires, c) Transition $T_6$ is enabled, d) Transition $T_6$ fires.

To model this queue with a timed event graph we divide the places up into two groups $\{p_i\}$ and $\{q_i\}$ with $i = 1, 2, ..., N$. See fig 2. The $\{p_i\}$ represent the sites in the production line so that a waiting token in place $p_i$ should be thought of as an item undergoing a process at site $i$, likewise a ready token at place $i$ should be thought of as a processed item at site $i$. The waiting times at $p_i$ are a sequence of i.i.d. random variables $[t_i(n)]_{n=1}^{\infty}$ with distribution given by the time that the $i$th process takes.

The $\{q_i\}$ are used to make sure that only one token occupies each $p_i$ at a time, a token in place $q_i$ should be thought of as the $i$th site being empty and ready to receive a new item for processing. The waiting times at $q_i$ are all zero so that all tokens in $\{q_i\}$ are instantaneously ready.

Transitions $T_i$ with $i = 2, 3, ..., N-1$ represent a processed item a site $i-1$ being transferred to site $i$. The initiating transition $T_1$ represents a new item being brought into the line and the terminal transition $T_N$ represents the completion of a finished product.

This system could be used to model all sorts of different production lines, most interestingly it gives a very complete description of a bio-molecular process called mRNA transcription, the process by which ribosomes build proteins coded for in mRNA. For a thorough analysis of the deterministically timed version of the model for this application see [8].

## 1.2 Max-plus formulation

Given an initial distribution of tokens in a stochastic timed event graph (which we call a marking) the evolution of the petri-net's state can be completely determined by the sequence of transition firing times. This statement is trivial to prove, suppose that we start with some initial marking $m_0$ and arrive after some time at a new marking $m'$. If we are given the transition firing sequence but not the sequence of states that the system moves through we can simply apply these transitions to $m_0$ until we arrive at $m'$.

The state variable for our max-plus model at stage $n$ is then given by the firing time vector $X(n) = [x_i(n)]$ where $x_i(n)$ is the time at which transition $i$ fires for the $n$'th time.

All stochastic timed event graph satisfy a max-plus linear equation which can easily be obtained [1, 2]. Since all the systems we will consider have the special property that each place only ever hold zero or one tokens the formulation is slightly simpler. At each stage the firing time vectors satisfy

$$X(n) = A(n) \otimes X(n) \oplus B(n-1) \otimes X(n-1) \tag{2}$$

Where $A(n)$ and $B(n-1)$ are max-plus matrices whose components are determined by the random waiting times associated with the places in the petri net and $M_o$ the initial marking of tokens.

$$A(n)_{i,j} = \bigoplus_{\{k:T_j \to p_k \to T_i, m_0(k)=0\}} t_k(n) \tag{3}$$

$$B(n)_{i,j} = \bigoplus_{\{k:T_j \to p_k \to T_i, m_0(k)=1\}} t_k(n) \tag{4}$$

So that $A(n)_{i,j}$ is a random variable given by the maximum of a set of the $n$th place waiting times. This maximum is taken over all the places that join $T_j$ to $T_i$ and contain no tokens in the initial marking. Likewise $B(n)_{i,j}$ is a maximum but taken over places that contain one token in the initial marking.

Since when we fix the random waiting times our system's evolution is deterministic, it follows that $X(n)$ exists and is unique. It can also be shown that (1) has a unique solution [6]. Define the Kleene star of $A(n)$ by

$$A(n)^* = I \oplus A(n) \oplus A(n)^{\otimes 2} \oplus A(n)^{\otimes 3} \oplus ... \tag{5}$$

So that $A(n)^*_{i,j}$ should be thought of as the weight of the maximally weighted path from $j$ to $i$ of any length through $A(n)$'s associated graph. Unless our petri-net's initial marking contains a circuit with no tokens in any of its places $A(n)^*$ will exist since $A^{\otimes K} = \epsilon$ for all $K > N$. Any petri-net/initial marking with such a circuit will have trivial dynamics since none of the transitions will be able to fire. We can use the Kleene star to construct a solution to (1) by

$$X(n) = A(n)^* \otimes B(n-1) \otimes X(n-1) \tag{6}$$

which on substitution to the RHS yields

$$A(n) \otimes A(n)^* \otimes B(n-1) \otimes X(n-1) \oplus B(n-1) \otimes X(n-1) \tag{7}$$

$$= [A(n) \otimes A(n)^* \oplus I] \otimes B(n-1) \otimes X(n-1) = A(n)^* \otimes B(n-1) \otimes X(n-1)$$

Therefore this is the unique solution to (1) and provides the unique evolution of our system. The dynamics of our stochastic timed event graph are governed by the max-plus linear system

$$X(n) = A(n)^* \otimes B(n-1) \otimes X(n-1) = [\bigotimes_{k=1}^{n} \underbrace{A(k)^* \otimes B(k-1)}] \otimes X(0) \qquad (8)$$

And since the components of $A(k)$ and $B(k-1)$ are drawn from disjoint sets of edge weights the underbraced terms form a sequence of i.i.d. random max-plus matrices.

## 1.3 Example

We can now construct the max-plus linear system associated with our asynchronous production line model. We will use the initial making $m_0$ where there are no tokens in any of the $p_i$ and one token in each $q_i$, this corresponds to starting the plant with no items currently in processing. As outlined we choose to represent max-plus matrices with their associated weighted graphs rather than as an array of numbers, see fig 3.

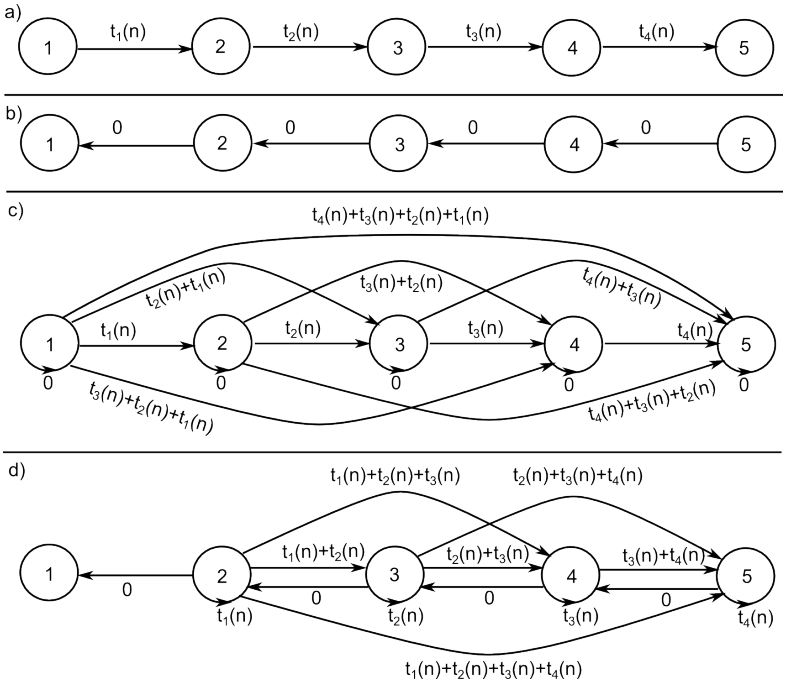


Figure 3: Graphs associated with $A(n)$,$B(n-1)$,$A(n)^*$ and $A(n)^*B(n-1)$ respectively for $N = 4$

We can now simulate our system by choosing an initial condition $X(0)$ then randomly generating a sequence of matrices $[A(k)^*B(k-1)]_{k=1}^{n}$ and max-plus multiplying the state with them to obtain $X(n)$, see section 2.1.

## 2 Max-plus Lyapunov exponent

Our max-plus linear systems are evolved according to

$$X(n) = M(n) \otimes X(n-1) = [\underbrace{\bigotimes_{k=1}^{n} M(k)}] \otimes X(0) \tag{9}$$

where the $[M(k)]_{k=1}^{\infty}$ are a sequence of i.i.d. random max-plus matrices. Since the associated graphs for these matrices have the same vertex and edge set each time, just different edge weights, we can interpret the underbraced matrix product which we denote $P(n)$ as follows. $P(n)_{i,j}$ is given by the weight of the maximally weighted path $\sigma$ of length $n$ through the associated graphs from $j$ to $i$ which accumulates weight on its $k$'th step according to $M(k)$ so that $\sigma$'s total weight is given by

$$W(\sigma) = \sum_{k=1}^{n} M(k)_{\sigma(k-1),\sigma(k)} \tag{10}$$

This maximally weighted path perspective is very natural to use in max-plus linear algebra and is essential to the new theory presented in this paper. It also enables us to give a new, simpler proof of the max-plus multiplicative ergodic theorem.

**Theorem** With probability-1 the limit

$$\lambda = \lim_{n\to\infty} \frac{1}{n} P(n) = \lim_{n\to\infty} \frac{1}{n} \bigotimes_{k=1}^{n} M(k) \tag{11}$$

exists and is a matrix with each element equal to the same constant, the *max-plus Lyapunov exponent* which we shall also denote $\lambda$. Note that since conventional multiplication acts like taking powers in max plus we are justified in calling this an exponent.

**Proof** This proof is a slightly different to the standard treatment which follows from Kingman's sub-additive ergodic theorem [2].

**Claim 1** $\lambda_n = \frac{1}{n}\mathbb{E}\{\max_{i,j}[\bigotimes_{k=1}^{n} M(k)]_{i,j}\}$ is a decreasing sequence bounded above by zero and therefore $\lim_{n\to\infty} \lambda_n = \lambda$ exists.

**Proof** Consider $(n+m)\lambda_{n+m}$ the expected weight of the maximally weighted path of length $n+m$. Now suppose that to further maximize this weight we are able to jump to any vertex in the graph between the $n$th and $(n+1)$th stages - since we are able to do nothing if we so choose the weight of this new path must be at least equal to the previous maximum so that

$$(n+m)\lambda_{n+m} \leq n\lambda_n + m\lambda_m \tag{12}$$

and the claim follows from this subadditivity.

**Claim 2** $[\bigotimes_{k=1}^{n} M(k)]_{i,j} \geq \max_{i',j'}[\bigotimes_{k=1}^{n} M(k)]_{i',j'} - C$ for some $C$ with bounded expectation.

**proof** Consider the path $\sigma$ that attains the maximum weight in the right hand side of the inequality. Provided the graph associated with $M$ is irreducible and aperiodic there exists a path $\varsigma$ from $i$ to $j$ that coincides with $\sigma$ from the $N$th step to the $n-N$th step. $C$ is just the difference in weight between the two paths whose expectation can easily be bounded by $2N\lambda_1$.

**Corollary** Our theorem follows by considering a path of length $n \times m$ from $i$ to $j$. Suppose that the maximally weighted path is $\sigma$ then

$$\lim_{n,m\to\infty} \frac{1}{nm}[\bigotimes_{k=1}^{nm} M(k)]_{i,j} = \lim_{n,m\to\infty} \frac{1}{m}(\frac{1}{n}[\bigotimes_{k=1}^{n} M(k)]_{i,\sigma(n)} + \tag{13}$$

$$\frac{1}{n}[\bigotimes_{k=n+1}^{2n} M(k)]_{\sigma(n+1),\sigma(2n)} + ... + \frac{1}{n}[\bigotimes_{k=n(m+1)+1}^{nm} M(k)]_{\sigma(n(m-1)+1),j})$$

We can now use claim 2 to replace

$$\frac{1}{n}[\bigotimes_{k=tn+1}^{n(t+1)} M(k)]_{\sigma(tn+1),\sigma(n(t+1)} \tag{14}$$

with

$$\frac{1}{n}\max_{i',j'}[\bigotimes_{k=tn+1}^{n(t+1)} M(k)]_{i',j'} + \frac{C}{n} \tag{15}$$

Taking the limit in $m$ the first part of each term gives us $\lambda_n$ by the law of large numbers. Then taking the limit in $n$ gives us $\lambda$ and takes the $C$ terms to zero. If we can not factorize our path length in this way simply include a remainder which when divided by $n$ goes to zero.

**Remark** The proof for the special case of deterministic sequences is also simpler from this maximally weighted path perspective. Consider

$$\lambda = \lim_{n\to\infty} \frac{1}{n} A^{\otimes n} \tag{16}$$

where $A$ is any irreducible aperiodic max-plus matrix. Now take the cycle $C = c(1), ..., c(m)$ with maximum average weight and consider the component $A^{\otimes n}_{i,j}$ which is the weight of the maximally weighted path of length $n$ from $i$ to $j$. To construct a path attaining this weight (for large $n$) simply move from $i$ to $c(1)$ in less than $N$ steps, complete the cycle as many times as possible and finally return to $j$ in less than m+N steps. The average weight of this and any other maximally weighted path will converge to $C$'s average weight.

The dynamics of our max-plus system are therefore dominated by $\lambda$. The remainder of this paper is devoted to calculating and bounding the exponent for some examples and developing a theory to link it with the classical eigenvalues of the adjacency matrix of $M$'s associated graph.

## 2.1 Example

Since the max-plus exponent acts in an arithmetic way as opposed to a classical Lyapunov exponent's geometric action it is very easy to approximate directly in a numerically stable way. Returning to our asynchronous production line example we choose all the waiting times to be i.i.d. mean-1 exponentials and simulate the max-plus system as outlined in 1.3. The Max-plus exponent $\lambda$ corresponds to the average time between successive completions of finished products in the production line, it is the reciprocal of the throughput.

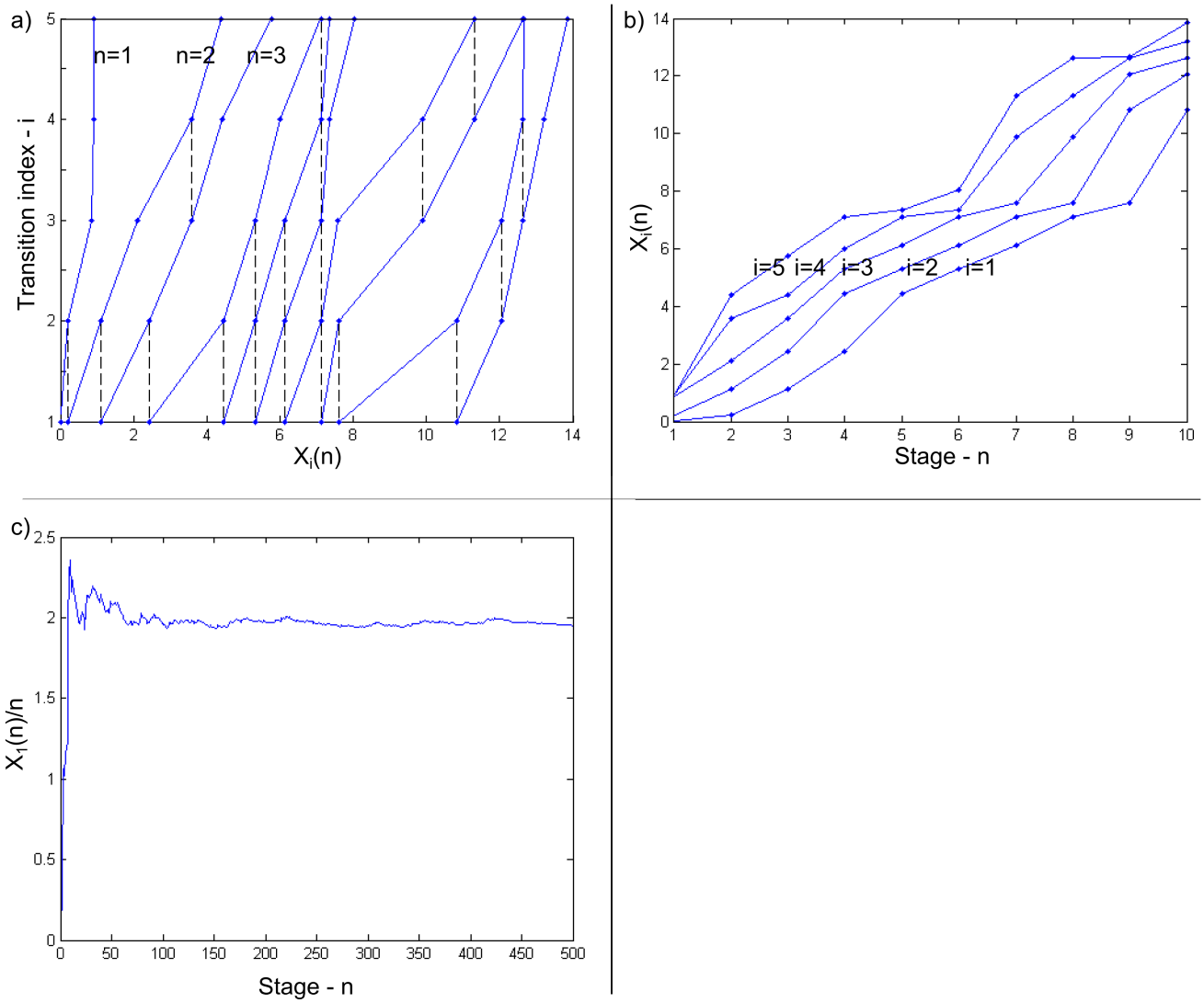


Figure 4: Simulation of asynchronous production line for $N = 4$. a) Progression of individual items through system, dashed black lines indicate jamming, b) Evolution of $X_i(n)$ variable, c) Convergence of $X_1(n)/n$ to max-plus Lyapunov exponent $\lambda$.

## 2.2 Relation with principal adjacency eigenvalue

The max-plus exponent $\lambda$ depends on both the waiting-time distributions and the topology of the matrices associated graphs. By restricting our attention to matrices whose waiting times are all i.i.d. mean-1 exponentials we can better explore the relationship between $\lambda$ and the graph's topology.

**Definition** we say that an $\mathbb{R}_{\max}^{N\times N}$ valued random variable $M$ is *componentwise exponential* iff

$$M_{i,j} = \bigoplus_{k\in K(i,j)} t_k \tag{17}$$

Where $(t_k)_{k=1}^{L}$ is a sequence of i.i.d. mean-1 exponentials and $K(i,j) \subset \{1,2,...,L\}$. We will also require the second condition that $(t_{K(i,j)})_{j=1}^{N}$ are independent for each $i$.

**Remark** The graph associated with a componentwise exponential max-plus matrix is therefore a directed weighted graph on $N$ vertices with possibly multiple edges between any two nodes. The weight of each edge is given by a mean-1 exponential random variable and the different edge's weights are *either* independent *or* identical. The second condition guarantees that all edges out of a particular vertex have independent weights.

**Theorem** Suppose that $[M(n)]_{n=1}^{\infty}$ is a sequence of i.i.d. componentwise exponential max-plus matrices. The max-plus exponent

$$\lambda = \lim_{n\to\infty} \frac{1}{n} \bigotimes_{k=1}^{n} M(k) \tag{18}$$

satisfies

$$\log \Lambda \leq \lambda < \alpha^* \tag{19}$$

where $\alpha^*$ is the unique solution to

$$\alpha = \log \Lambda + \log(1 + e\alpha) \tag{20}$$

and $\Lambda$ is the principal eigenvalue of $\mathcal{A}(M)$, the adjacency matrix of $M$'s associated graph defied by $\mathcal{A}(M)_{i,j} = k$ iff there are $k$ edges from $j$ to $i$ in $M$'s graph.

**Note** The adjacency eigenvalue $\Lambda$ tells us the rate at which the number of paths of length $n$ through the graph grows with $n$. In particular if $\chi_i(n)$ is the number of paths of length $n$ that end at vertex $i$ then

$$\chi(n+1) = \mathcal{A}(M)\chi(n) = \mathcal{A}(M)^{n+1}\chi(0) \tag{21}$$

and provided the adjacency matrix is irreducible and aperiodic there exists a unique maximal eigenvalue $\Lambda$ corresponding to a eigenvector $u$ with non-zero weight on all components so that

$$\lim_{n\to\infty} \frac{1}{n} \log \|\chi(n)\| = \log \Lambda \tag{22}$$

**Lower Bound** Recall that

$$\lambda_n = \frac{1}{n}\mathbb{E}\{\max_{i,j}[\bigotimes_{k=1}^{n} M(k)]_{i,j}\} \tag{23}$$

is the expectation of the step averaged weight of the maximally weighted path of length $n$ through the graphs associated with $[M(n)]_{n=1}^{\infty}$ and that $\lambda = \lim_{\to\infty} \lambda_n$. The max-plus exponent is therefore the step averaged weight of the maximally weighted path and we can bound it below with the step averaged weight of any other path we like. In particular we will devise a method for constructing a highly weighted path and use the expectation of its step averaged weight as a lower bound.

Our strategy constructs a path $\sigma$ by choosing to move from one vertex to the next at each stage by considering only the edge weights at that stage. Suppose that at stage $n$ our path is at vertex $i$ so $\sigma(n) = i$. The edge weight to be accumulated on the next step will be taken from the matrix $M(n)$ and since we must move from vertex $i$ it will be one of $M(n)_{i,j}$ for $j = 1, 2, ..., N$. We will always greedily choose to move along the maximally weighted edge accumulating the maximum weight so that $\sigma(n+1) = k$ where $M(n)_{i,k} \geq M(n)_{i,j}$ for all $j = 1, 2, ..., N$.

Since the edges out of each vertex are all i.i.d. the probability that any particular edge has the maximum weight is the uniform fraction $1/d_i$ where $d_i$ is the out degree of the vertex $i$. The sequence of vertices in the path $\sigma$ is therefore a markov chain with

$$\mathbb{P}[\sigma(n+1) = j | \sigma(n) = i] = \frac{A_{i,j}}{d_i} \tag{24}$$

and stationary distribution $\pi$.

Now consider the weight of the first $n$ edges in the path $\sigma$

$$W|_n(\sigma) = \max_{k\in K[\sigma(1)]} t_k(1) + \max_{k\in K[\sigma(2)]} t_k(2) + ... + \max_{k\in K[\sigma(n)]} t_k(n) \tag{25}$$

where $K(i) = \cup_{j\in\{1,2,...,N\}} K(j,i)$ are the indices for all waiting-times associated with edges leaving vertex $i$. By the law or large numbers we have

$$\lim_{n\to\infty} \frac{1}{n} W|_n(\sigma) = \sum_{i=1}^{N} \pi_i \underbrace{\mathbb{E}[\max_{k\in K(i)} t_k]} \tag{26}$$

where the underbraced term is the expectation of the maximum of $d_i$ i.i.d. mean-1 exponentials and is given by $\sum_{k=1}^{d_i} \frac{1}{k}$. The step averaged weight of the path is therefore given by

$$\sum_{i=1}^{N} \pi_i \sum_{k=1}^{d_i} \frac{1}{k} \geq \sum_{i=1}^{N} \pi_i \log d_i \tag{27}$$

where we will use the log approximation as our lower bound.

We now consider a cocycle on the markov chain $\sigma$ defined by

$$C|_n(\sigma) = [\prod_{k=0}^{n} d_{\sigma(k)}]^{-1} \tag{28}$$

and by the multiplicative ergodic theorem, with probability one, we have

$$\lim_{n\to\infty} \frac{1}{n} \log C|_n(\sigma) = -\sum_{i=1}^{N} \pi_i \log d_i \tag{29}$$

The $n$th stage entropy of the cocyle is given by

$$H_n = \frac{-1}{n} \sum_{\{\sigma|_n : \sigma(0)=1\}} C|_n \log C|_n \tag{30}$$

where the sum is taken over all paths of length $n$ that begin at vertex 1. Note that $C|_n(\sigma)$ is the probability of following exactly the path $\sigma$ from $\sigma(0)$ to $\sigma(n)$ taking the appropriate edge $[\sigma(k+1), \sigma(k)]$ whenever their are multiple edges, we therefore say that our cocycle is stochastic and

$$\lim_{n\to\infty} H_n = -\sum_{\{\sigma|_n : \sigma(0)=1\}} C|_n \sum_{i=1}^{N} \pi_i \log d_i = -\sum_{i=1}^{N} \pi_i \log d_i \tag{31}$$

The entropy of a probability measure on a set of $k$ objects is maximized when the measure is uniform. Since $C|_n$ can be thought of as a probability measure on $\{\sigma|_n : \sigma(0) = 1\}$ and in the limit $n \to \infty$ $C_n$ becomes, essentially, uniform we have

$$\lim_{n\to\infty} H_n - \frac{1}{n}\widehat{H}_n = 0 \tag{32}$$

where $\widehat{H}(n)$ is the entropy of the uniform probability distribution on $\{\sigma|_n : \sigma(0) = 1\}$ which is given by

$$\widehat{H}(n) = -\sum_{\{\sigma|_n : \sigma(0)=1\}} \frac{1}{D_n} \log D_n \tag{33}$$

where $D_n = \|\{\sigma|_n : \sigma(0) = 1\}\|$ is the number of paths of length $n$ that start at vertex 1. Finally we have

$$\sum_{i=1}^{N} \pi_i \log d_i = \lim_{n\to\infty} \frac{1}{n} \log D_n = \log \Lambda \tag{34}$$

which completes the proof of our lower bound

$$\log \Lambda < \lambda \tag{35}$$

**Remark** The use of the approximation

$$\sum_{k=1}^{N} \frac{1}{k} \approx \log N \tag{36}$$

is the reason that our bound is not sharp. This is in the sense that we can construct a sequence of i.i.d. matrices whose maximally weighted path is the same path chosen by our greedy strategy. If we set

$$M(n)_{i,j} = t_i(n) \tag{37}$$

for $i, j = 1, 2, ..., N$ then the choice of edge weights at each vertex are the same for all the different vertices at each stage and the optimal strategy for choosing a maximally weighted path is simply to choose the best available edge at each stage, as in our greedy strategy. Therefore in this system

$$\lambda = \sum_{k=1}^{N} \frac{1}{k} \tag{38}$$

and we can express this in terms of $\Lambda = N$ in the limit by

$$\lim_{N\to\infty} \lambda = \log \Lambda + \gamma + \epsilon_N \tag{39}$$

where $\gamma$ is the EulerMascheroni constant and $\epsilon_N \to 0$ like $1/N$.

**Upper bound** The expected weight of the maximally weighted path of length $n$ can, of course, be expressed in terms of path weights

$$\lambda_n = \frac{1}{n}\mathbb{E}\max_{i,j}[\bigotimes_{k=1}^{n} M(k)]_{i,j} = \frac{1}{n}\mathbb{E}\max_{\sigma|_n} \underbrace{W|_n(\sigma, [t_i(k)]_{i=1,k=1}^{i=N,k=n})}_{} \tag{40}$$

where the underbraced term is equal to the weight of the length-$n$ path $\sigma$ with edge weights determined by the i.i.d. mean-1 exponentials $[t_i(k)]_{i=1,k=1}^{i=N,k=n}$. It is important that in evaluating this expression the same waiting times are used for each different path. The maximum is then taken over all paths of length $n$.

The weight of each path of length-$n$ is a sum of $n$ i.i.d. mean-1 exponentials so that all path weights are identically distributed. However different paths can share edges which they traverse at the same step, accumulating the same waiting time. Therefore the path weights are not independent. Since their dependence arises from edge weight sharing it will only tend to make them more correlated. We say that the edge weights are *associated* random variables.

**Definition** A sequence of random variables $(x_i)_{i=1}^N$ are said to be *associated* if for all $f, g : \mathbb{R}^N \mapsto \mathbb{R}$ non-decreasing in each component we have

$$Cov[f(x1, x2, ..., x_N), g(x1, x2, ..., x_N)] \geq 0 \tag{41}$$

To see that the path weights are associated note that each $W|_n(\sigma)$ is a non-decreasing function of the waiting-times. Thus if $f$ and $g$ are non-decreasing

functions of the path weights they are non-decreasing functions of the waiting-times which are all independent so the inequality holds.

Since this association means the random variables are positively correlated it reduces their standard deviation and the expectation of their maximum. We can therefore bound $\lambda$ from above by taking the maximum in (39) while ignoring the dependence and treating each path weight as an i.i.d. sum of i.i.d. mean-1 exponentials.

**Claim** Suppose that $(x_i)_{i=1}^N$ and $(y_i)_{i=1}^N$ are identically distributed random variables and that $(x_i)_{i=1}^N$ are associated but $(y_i)_{i=1}^N$ are independent.

$$\mathbb{E}\max_{i=1}^{N} x_i \leq \mathbb{E}\max_{i=1}^{N} y_i \tag{42}$$

**Proof** For some $i \in \{1, 2, ..., N\}$, a subset $J \in \{1, 2, .., N\}$ and any positive real number $t$ define the non-decreasing functions $f$ and $g$ by

$$f(x_1, x_2..., x_N) = \begin{cases} -1 & \text{if } x_i < t \\ 0 & \text{otherwise} \end{cases} \tag{43}$$

$$g(x_1, x_2..., x_N) = \begin{cases} -1 & \text{if } \max_{j\in J} x_j < t \\ 0 & \text{otherwise} \end{cases}$$

Now to calculate the covariance of these two functions we need to consider four events

- $x_i < t$ and $\max_{j\in J} x_j < t$ occurs with probability $P(f, g) = \mathbb{P}[x_i < t, x_j < t, j \in J]$
- $x_i < t$ but $\max_{j\in J} x_j \geq t$ occurs with probability $P(f) - P(f, g)$ where $P(f) = \mathbb{P}[x_i < t]$
- $\max_{j\in J} x_j < t$ but $x_i \geq t$ occurs with probability $P(g) - P(f, g)$ where $P(g) = \mathbb{P}[x_j < t, j \in J]$
- $x_i \geq t$ and $\max_{j\in J} x_j \geq t$ occurs with probability $1 - P(f) - P(g) + P(f, g)$

The covariance of $f, g$ is then

$$Cov[f, g] = P(f, g)[P(f)P(g) - P(f) - P(g) + 1] + [P(f) - P(f, g)][P(f) - 1]P(g) \tag{44}$$
$$+[P(g) - P(f, g)][P(g) - 1]P(f) + [1 - P(f) - P(g) + P(f, g)]P(f)P(g) \geq 0$$

which simplifies to $P(f, g) \geq P(f)P(g)$ so that

$$\mathbb{P}[x_i < T | x_j < T, j \in J] \geq \mathbb{P}[x_i < T] \tag{45}$$

Now consider

$$\mathbb{P}[\max_{i=1}^{N} x_i < t] = \prod_{i=1}^{N} \mathbb{P}[x_i < t | x_k < t; k = 1, 2, ..., i - 1] \tag{46}$$

and if we set $J = \{1, 2, ..., i-1\}$ then our previous result tells us that

$$\mathbb{P}[\max_{i=1}^{N} x_i < t] \geq \prod_{i=1}^{N} \mathbb{P}[x_i < t] \tag{47}$$

which is exactly equal to the probability that the maximum of the $y_i$ is less than $t$. Finally

$$\mathbb{E} \max_{i=1}^{N} x_i = \int_0^\infty t\rho_x(t)dt = \int_0^\infty \int_z^\infty \rho_x(t)dtdz \tag{48}$$

$$= \int_0^\infty 1 - \mathbb{P}[\max_{i=1}^{N} x_i < z]dz \leq \int_0^\infty 1 - \mathbb{P}[\max_{i=1}^{N} y_i < z]dz = \mathbb{E} \max_{i=1}^{N} y_i$$

so that the expectation of the maximum of the associated variables is less than or equal to the expectation of the maximum of the independent variables as claimed.

Therefore we have the upper bound

$$\lambda \leq \lim_{n\to\infty} \frac{1}{n} \mathbb{E} \max_{i=1}^{D_n} X_i(n) \tag{49}$$

where $[X_i(n)]_{i=1}^{D_n}$ is a sequence of i.i.d. sums of $n$ i.i.d. mean-1 exponentials and $D_n$ is the number of paths of length $n$ through our graph. Unfortunately we can't approximate the $X_i(n)$ as Gaussians because the convergence in the central limit theorem is not sufficiently uniform. Instead we will show directly that as $n \to \infty$ the ratio of the median to the mean of the maximum tends to 1 so we can calculate the expectation by calculating the median, which is naturally easier than the mean for the maximum of several random variables.

Given a sequence of probabilities $(p_n)_{n=1}^\infty$ we define the sequence of generalized medians $[\mu_n(p_n)]_{n=1}^\infty$ for our sequence of random variables $[\max_{i=1}^{D_n} X_i(n)]_{n=1}^\infty$ by

$$\mathbb{P}[\max_{i=1}^{D_n} X_i(n) < \mu_n(p_n)] = \mathbb{P}[X_1(n) < \mu_n(p_n)]^{D_n} = p_n \tag{50}$$

And using the identity

$$e^x = \lim_{m\to\infty} (1 + \frac{x}{m})^m \tag{51}$$

with $x = \log p_n$ and $m = D_n$ we have, approximately but verifiably in the limit $n \to \infty$ for all sequences $(p_n)_{n=1}^\infty$ we consider

$$\mathbb{P}[X_1(n) < \mu_n(p_n)] = 1 + \frac{\log p_n}{D_n} \tag{52}$$

$$\mathbb{P}[X_1(n) > \mu_n(p_n)] = -\frac{\log p_n}{D_n}$$

And since we know the distribution of the sum of $n$ i.i.d. mean-1 exponentials we can write down an integral for the LHS which we can then integrate by parts

$$= \int_{\mu_n(p_n)}^\infty \frac{x^{n-1}e^{-x}}{(n-1)!}dx = e^{-\mu_n(p_n)} \underbrace{\sum_{k=0}^{n-1} \frac{\mu_n(p_n)^k}{k!}} \tag{53}$$

where the underbraced term is a truncation of $e^x$'s power series evaluated at $x = \mu_n(p_n)$. We can bound this sum with the inequality

$$(1 + \frac{x}{n})^n \leq \sum_{k=0}^{n} \frac{x^k}{k!} \leq (1 + \frac{ex}{n})^n \tag{54}$$

which we prove by looking at the coefficient of $x^k$

$$\frac{1}{k!}\frac{n!}{(n-k)!n^k} \leq \frac{1}{k!} \leq \frac{1}{k!}\frac{n!e^k}{(n-k)!n^k} \tag{55}$$

Therefore in the limit $n \to \infty$ we have the inequality

$$e^{-\mu_n(p_n)}(1 + \frac{\mu_n(p_n)}{n-1})^{n-1} \leq \frac{-\log p_n}{D_n} \leq e^{-\mu_n(p_n)}(1 + \frac{e\mu_n(p_n)}{n-1})^{n-1} \tag{56}$$

**Claim** Any sequence of medians $\mu_n(p_n)$ satisfying

$$\lim_{n\to\infty} \frac{\log\log p_n}{n} = 0 \tag{57}$$

including the sequence of proper medians $[\mu_n = \mu_n(\frac{1}{2})]_{n=1}^{\infty}$ satisfy, in the limit $n \to \infty$

$$\log\Lambda \leq \frac{\mu_n}{n} \leq \log\Lambda + \log(1 + \frac{e\mu_n}{n}) \leq C\log\Lambda + D \tag{58}$$

for some bounded $C$ $D$ that do not depend on $\Lambda$.

**Proof** substituting $p_n = \frac{1}{2}$ into inequality (55) then taking logs and dividing by $n$ gives the first sandwich inequality. That everything is less than or equal to $C\log\Lambda + D$ follows from the fact that

$$F(\alpha) = \log\Lambda + \log(1 + e\alpha) \tag{59}$$

is a contraction mapping whose unique fixed point $\alpha^*$ bounds $\lim_{n\to\infty} \frac{\mu_n}{n}$ from above. Now

$$\frac{\delta\alpha^*}{\delta\log\Lambda} = \frac{1 + e\alpha^*}{1 + e(\alpha^* - 1)} \tag{60}$$

So that for any fixed finite $\Lambda'$ we have

$$\alpha^*(\Lambda) \leq \alpha^*(\Lambda') + \frac{\delta\alpha^*}{\delta\log\Lambda}|_{\alpha^*}(\Lambda')\log\Lambda \tag{61}$$

**Claim** For large $n$ almost all the probability mass is close the median in the sense that

$$\lim_{n\to\infty} \frac{\mu_n(e^{\frac{-1}{n}}) - \mu_n(e^{-n})}{n} = 0 \tag{62}$$

**Proof** From equation (52) we have

$$\frac{\mu_n(e^{\frac{-1}{n}})}{n} = \log\Lambda + \frac{\log n}{n} + \log(\sum_{k=0}^{n-1} \frac{\mu_n(e^{\frac{-1}{n}})^k}{k!}) \tag{63}$$

and

$$\frac{\mu_n(e^{-n})}{n} = \log\Lambda - \frac{\log n}{n} + \log(\sum_{k=0}^{n-1} \frac{\mu_n(e^{-n})^k}{k!}) \tag{64}$$

So that

$$\frac{\mu_n(e^{\frac{-1}{n}}) - \mu_n(e^{-n})}{n} \leq \frac{2\log n}{n} \tag{65}$$

$$+\frac{\mu_n(e^{\frac{-1}{n}}) - \mu_n(e^{-n})}{n} \underbrace{\max_{\mu_n(e^{-n}) \leq x \leq \mu_n(e^{\frac{-1}{n}})} \frac{d}{dx} \log(\sum_{k=0}^{n-1} \frac{x^k}{k!})}$$

using the result of the previous claim we bound the underbraced term by

$$\max_{\log\Lambda \leq \alpha \leq C\log\Lambda + D} 1 - \frac{\frac{(\alpha n)^{n-1}}{(n-1)!}}{\sum_{k=0}^{n-1} \frac{(\alpha n)^k}{k!}} \tag{66}$$

which gives

$$\frac{\mu_n(e^{\frac{-1}{n}}) - \mu_n(e^{-n})}{n} \leq \frac{2\log n}{n} \underbrace{\max_{\log\Lambda \leq \alpha \leq C\log\Lambda + D} \frac{(n-1)!}{(\alpha n)^{n-1}} \sum_{k=0}^{n-1} \frac{(\alpha n)^k}{k!}} \tag{67}$$

where the underbraced term can be bounded above by

$$\max_{\log\Lambda \leq \alpha \leq C\log\Lambda + D} \sum_{k=0}^{n-1} \frac{1}{\alpha^{n-k-1}} \leq \max_{\log\Lambda \leq \alpha \leq C\log\Lambda + D} (1 - \frac{1}{\alpha})^{-1} \tag{68}$$

So that

$$\frac{\mu_n(e^{\frac{-1}{n}}) - \mu_n(e^{-n})}{n} < \frac{2\log n}{n}(1 - \frac{1}{\log\Lambda})^{-1} \to 0 \tag{69}$$

as required.

We can now express the mean as an integral over a series of intervals bounded by generalized medians

$$\mathbb{E}\max_{i=1}^{D_n} X_i = \int_0^{\mu_n(e^{-n})} x\rho(x)dx + \underbrace{\int_{\mu_n(e^{-n})}^{\mu_n(e^{\frac{-1}{n}})} x\rho(x)dx} \tag{70}$$

$$+ \int_{\mu_n(e^{\frac{-1}{n}})}^{\mu_n(e^{\frac{-1}{m^n}})} x\rho(x)dx + \sum_{m=2}^{\infty} \int_{\mu_n(e^{\frac{-1}{m^n}})}^{\mu_n(e^{\frac{-1}{(m+1)^n}})} x\rho(x)dx$$

Where the underbraced term is equal to $\mu_n(\frac{1}{2}) + o(n)$. All that remains is to show that the sum of the remaining terms grows slower than $n$.

The first integral is easy

$$\int_0^{\mu_n(e^{-n})} x\rho(x)dx \leq e^{-n}\mu_n(e^{-n}) \leq e^{-n}n(C\log\Lambda + D) \tag{71}$$

We treat the remaining integrals in the same way, bounding them above by the probability associated with their interval multiplied by an upper bound on the value of their upper boundary. Using the upper bound in inequality (55) we obtain

$$\frac{\mu_n(e^{\frac{-1}{m^n}})}{n} < \log\Lambda - \frac{1}{n}\log\log e^{\frac{1}{m^n}} + \log(1 + e\frac{\mu_n(e^{\frac{-1}{m^n}})}{n}) \tag{72}$$

So that

$$\frac{\mu_n(e^{\frac{-1}{m^n}})}{n} < \log m\Lambda + \log(1 + e\frac{\mu_n(e^{\frac{-1}{m^n}})}{n}) \tag{73}$$

and as in (59) we have

$$\frac{\mu_n(e^{\frac{-1}{m^n}})}{n} < C\log m\Lambda + D \tag{74}$$

So the sum of the remaining integrals is bounded by

$$n(e^{\frac{-1}{2^n}} - e^{\frac{-1}{n}})(C\log 2\Lambda + D) + n\sum_{m=3}^{\infty}(e^{\frac{-1}{m^n}} - e^{\frac{-1}{(m-1)^n}})[C\log mK(n) + D] \tag{75}$$

And in the limit$n \to \infty$ this expression is bounded above by

$$(1 - e^{-n})n(C\log 2\Lambda + D) + n\sum_{m=3}^{\infty}\frac{C\log m}{(m-1)^n} \tag{76}$$

$$\leq Cn\sum_{m=3}^{\infty}\frac{1}{(m-1)^{n-1}} \to 0$$

Therefore the mean really does look like the median and we can use the bound obtained for the median on the mean.

**Corollary** Suppose that $[M_N(n)]_{n=1}^{\infty}$ is a sequence of i.i.d. componentwise exponential max-plus matrices parameterized by $N \in \mathbb{N}$.

- If $\max_N \Lambda_N = \overline{\Lambda}$ exists then $\max_N \lambda_N = \overline{\Lambda}$ exists.
- If on the other hand $\Lambda_N \to \infty$ then $\lim_{N\to\infty}\frac{\lambda_N}{\log\Lambda_N} = 1$
- Moreover if, as in all our examples, $M_N(n)$ can be realized as submatrix of $M_{N+K}(n)$ for all $N, K \geq 0$ then either $\lim_{N\to\infty}\Lambda_N$ exists and so does $\lim_{N\to\infty}\lambda_N$ or $\lim_{N\to\infty}\frac{\lambda_N}{\log\Lambda_N} = 1$

**Proof** The final item comes from the fact that if $M_N$ is a submatrix of $M_{N+K}$ then any path through its associated graph can also be found in the graph of $M_{N+K}$ so that $\Lambda_N \leq \Lambda_{N+K}$ and $\lambda_N \leq \lambda_{N+K}$. Everything else follows directly from the main theorem.

# 3 Examples

Our result relating the max-plus exponent to the adjacency matrix eigenvalue is restrictive in the sense that it only applies to componentwise exponential matrices. The max-plus matrix product systems associated with stochastic event graphs are not typically in this form. For example in 1.3 the graph associated with $A^*(n)B(n-1)$ has edges with zero weight as well as edges whose weight are conventional sums of more than one i.i.d. mean-1 exponential.

However we shall see that this and other matrices associated with a generalization of our original production line model can be modified in such a way that preserves the max-plus exponent and provides us with a componentwise exponential form. Thus we can apply our *edge weight redistribution* procedure to obtain a new system which has the same asymptotic properties as the original stochastic event graph but is amenable to our new theory.

## 3.1 Asynchronous production trees

The production line model outlined in 1.1 requires the $N$ different process to be carried out in the total order $1 > 2 > ... > N$ to produce a finished product. We can generalize our production model to allow for a class of partial orders on these processes.

**Definition** A *tree order* $\preceq, \succeq$ is a partial order on $\{1, 2, ..., N\}$ with a unique maximal element such that $\{k \preceq i\} \cap \{k \preceq j\} = \emptyset$ whenever $\{i, j\}$ is an anti-chain. We associate the graph $G(\preceq, \succeq)$ with vertex set $\{1, 2, ..., N\}$ and an edge $(i, j)$ whenever $j$ is a maximal element of $\{k \preceq i\}$. The graph associated with a tree metric is therefore a rooted tree and for each rooted tree on $\{1, 2, ..., N\}$ there exists a unique partial order associated with it as such.

The requirement that incomparable processes do not share predecessors is necessary to obtain a stochastic even graph description as otherwise some places will have more than one succeeding transition. It is also a fairly reasonable assumption for a generic asynchronous production line, if we interpret each process as taking the products of its direct predecessors and amalgamating them in some way then there is no reason why parallel processes should be dependent in any way on each others predecessors.

We construct the stochastic event graph associated with a tree order $\preceq, \succeq$ as follows. For each $i \in \{1, 2, ..., N\}$ we include a pair of places $p_i$ and $q_i$, and a transition $T_i$. As before a token in $p_i$ represents as an item being processed at site $i$ with i.i.d. mean-1 exponential waiting times $[t_i(n)]_{n=1}^{\infty}$. A token in $q_i$ represents site $i$ being unoccupied with zero waiting times. Transition $T_i$ represents the products of process $i$'s predecessors being completed and moved into site $i$ for processing.

We include an edge from $q_i$ to $T_i$ and from $T_i$ to $p_i$, in addition we include an edge from $p_i$ to $T_j$ where $j$ is $i$'s unique direct successor in $G(\preceq, \succeq)$ and an edge from $T_j$ to $q_i$.

Finally we include a terminal transition $T_{N+1}$ which represents the completion of a finished product, we include a edge from $T_{N+1}$ to $q_k$ where $k$ is the root of $G(\preceq, \succeq)$ and an edge from $p_k$ to $T_{N+1}$. See example 3.4.

## 3.2 Edge weight redistribution

The graphs associated with $A(n)$ and $B(n-1)$ for the stochastic event graph of an asynchronous production line with partial order $G(\preceq, \succeq)$ are as follows.

- $A(n)$ is a graph on $\{1, 2, ..., N+1\}$ with an edge $(i, j)$ of weight $t_j(n)$ whenever $(i, j)$ is an edge in $G(\preceq, \succeq)$, in addition there is an edge $(N+1, k)$ of weight $t_k(n)$ where $k$ is the maximal element.
- $B(n-1)$ has the same vertex set but with the opposite edge set, so that $(i, j)$ is an edge of weight zero in $B(n-1)$ whenever $(j, i)$ is an edge in $A(n)$.
- $A^*(n)B(n-1)$ is a graph on $\{1, 2, ..., N+1\}$ where for each $i$ there are self loops $(i, i)$ of weight $t_j(n)$ and zero weight edges $(j, i)$ for each $j$ that is directly succeeded by $i$ and for each $k \succeq i$ there is an edge of weight $t_j(n) + W_{k,i}$ where the second term is the weight of the unique path from $i$ to $k$.

**Definition** The componentwise exponential matrix $\widehat{A}(n)$ whose graph on $\{1, 2, ..., N+1\}$ contains for each $i$ edges $(j, i)$ and $(i, k)$ of weight $t_j(n)$ for each $j$ a direct predecessor of $i$ and any $k \succeq i$. The matrices components are given by

$$\widehat{A}(n)_{i,j} = \begin{cases} t_i & \text{if } j \text{ is a direct predecessor of } i \\ \bigoplus_{k \in P(i)} t_k & \text{if } j \succeq i \\ \epsilon & \text{otherwise} \end{cases} \tag{77}$$

where $P(i)$ is the set of $i$'s direct predecessors.

**Claim** The max-plus exponent of the edge weight redistributed system

$$\widehat{\lambda} = \lim_{n \to \infty} \frac{1}{n} [\bigotimes_{k=1}^{n} \widehat{A}(n)]_{N+1,N+1} \tag{78}$$

is equal to that of the original stochastic event graph system

$$\lambda = \lim_{n \to \infty} \frac{1}{n} [\bigotimes_{k=1}^{n} A^*(n)B(n-1)]_{N+1,N+1} \tag{79}$$

Note that we will be calculating these exponents by looking just at the $(N+1, N+1)$th component of the product which should be thought of as the averaged weight of the maximally weighted path of length $n$ from the maximal vertex to itself.

**Proof** Again we can express the exponent in terms of path weights.

$$\widehat{\lambda}_n = \frac{1}{n} \mathbb{E} \max_{i,j} [\bigotimes_{k=1}^{n} \widehat{A}(k)]_{i,j} = \frac{1}{n} \mathbb{E} \max_{\widehat{\sigma}|_n} \underbrace{W|_n(\widehat{\sigma}, [\widehat{t_i}(k)]_{i=1,k=1}^{i=N,k=n})}_{} \tag{80}$$

and since the $\widehat{t_i}(n)$ and $t_j(m)$ are all i.i.d. we can obtain a statistically valid description of this system by setting $\widehat{t_i}(n) = t_i[n + d(i)]$ where $d(i)$ is the depth

of $i$ in the partial order, i.e. the length of the longest chain from the maximal element to $i$. We shall now show that there is a map between the sets of paths such that any sequence of edge weights in the original system can be accumulated by a different path in the redistributed system and that therefore the two exponents are the same.

Paths from $N+1$ to $N+1$ in the original system must step down the tree one vertex at a time and are then able to jump back up. In the redistributed system paths jump down the tree then step back up. We define a map $R$ from the paths in the original system to paths in the redistributed system by reading $\sigma$ and simultaneously writing $R(\sigma)$. Let $i$ be the latest step to be read from $\sigma$ and $j$ the latest step written to $R(\sigma)$.

- Start with $R(\sigma)_1 = \sigma_1 = N+1$
- Whenever $\sigma$ steps down the tree so that $d(\sigma_i) < d(\sigma_{i+1}) < ... < d(\sigma_{i+k}) \geq d(\sigma_{i+k+1})$ we set $R(\sigma)_j = \sigma_i$ $R(\sigma)_{j+1} = \sigma_{i+k}$
- whenever $\sigma$ jumps back up the tree (including when it stays at the same level) so that $d(\sigma_i) \geq d(\sigma_{i+1})$ we set $R(\sigma)_j = \sigma_i$, $R(\sigma)_{j+1} = \sigma_i^{(1)}$, ...,$R(\sigma_{j+k}) = \sigma_i^{(k)} = \sigma_{i+1}$ where $\sigma_i^{(t)}$ is $\sigma_i$'s unique $t$th successor.

Therefore if $\sigma$ is a length $n$ path from $N+1$ to $N+1$ in the original graph $R(\sigma)$ is a length $n$ path from $N+1$ to $N+1$ in the redistributed graph, and

$$\widehat{W}R(\sigma) = W(\sigma) \tag{81}$$

So that we have a measure preserving map between the $\widehat{t}_i(n)$ and $t_j(m)$ and a bijection between the paths in the maximum. Therefore when we take the expectation we can integrate over the $t_j(m)$ and have an identical integrand for each exponent. Thus the exponents are the same as claimed.

## 3.3 Asynchronous production line

We have already calculated $A(n), B(n-1)$ and $A^*(n)B(n-1)$, see section 1.3. Following the definition in 3.2 we construct the redistributed componentwise exponential matrix $\widehat{A}$ for this system

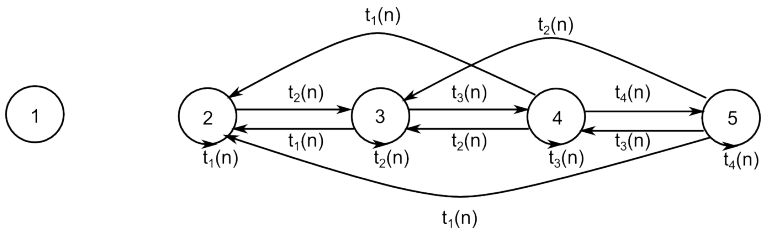


Figure 5: Graph associated with $\widehat{A}(n)$ for $N = 4$

Ignoring the minimal vertex with no connections the adjacency matrix for this graph is given by

$$\mathcal{A}_{i,j} = \begin{cases} 1 & \text{if } i \leq j+1 \\ 0 & \text{otherwise} \end{cases} \tag{82}$$

where we have identified each vertex with its depth for the index in the adjacency matrix. Now $\mathcal{A}$ is irreducible and aperiodic so $\Lambda$ exists and is unique. Also there exists $u$ with $u_i > 0$ for all $i$ such that $\mathcal{A}u = \Lambda u$. Examining the first row of this equation gives $u_1 + u_2 = \Lambda u_1$ so that $u_2 = \Lambda - 1$ the $n$th row says

$$\sum_{j=1}^{n+1} u_j = \Lambda u_{n-1} + u_{n+1} = \Lambda u_n \tag{83}$$

Which gives us a second-order, liner, constant-coefficient recurrence relation on the $u_n$ that we use to obtain

$$u_n = A_+ \mu_+^{n-1} + A_- \mu_-^{n-1} \tag{84}$$

where

$$\mu_\pm = \frac{\Lambda \pm \sqrt{\Lambda^2 - 4\Lambda}}{2} \tag{85}$$

and

$$A_\pm = \frac{1}{2} \pm \frac{\Lambda - 2}{2\sqrt{\Lambda^2 - 4\Lambda}} \tag{86}$$

We now require $u_N = u_{N-1}$ in order for the eigenvalue/vector pair to be valid. Suppose that $\Lambda > 4$ then both exponents are real and positive with $\mu_+ > \mu_-$ also $A_+ > |A_-|$ so that $u_n$ is an increasing function of $n$ and it is impossible to satisfy the final condition. Therefore $\Lambda \leq 4$ which proves that there is a finite limit in the exponent as the length of the production line grows $N \to \infty$. In fact $\Lambda \to 4$ very quickly and this can be seen in the convergence of the upper and lower bounds with $N$.

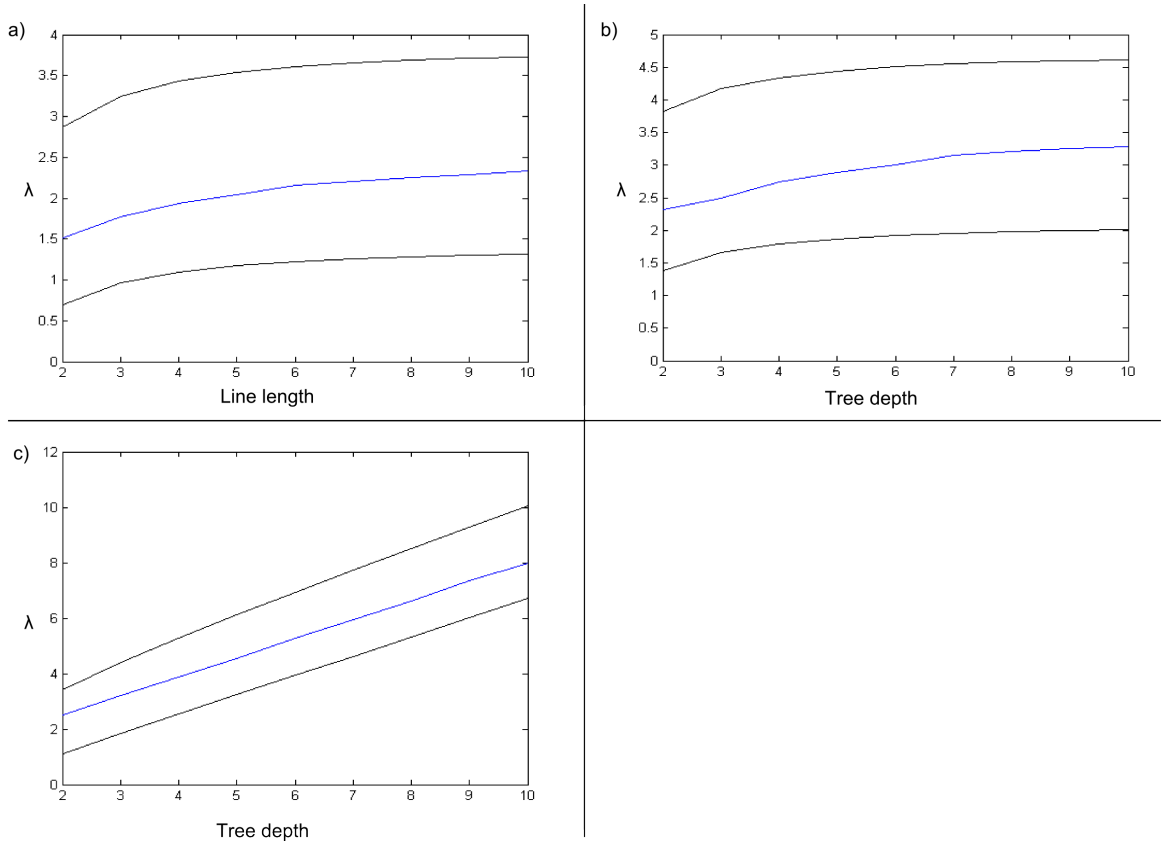


Figure 6: Max-plus exponent of production trees (blue) with bounds (black). a)Asynchronous production line, b)Asynchronous production binary tree, c)Partially synchronous production tree

## 3.4 Asynchronous production binary tree

We can construct an asynchronous production tree around a partial order derived from a binary tree.

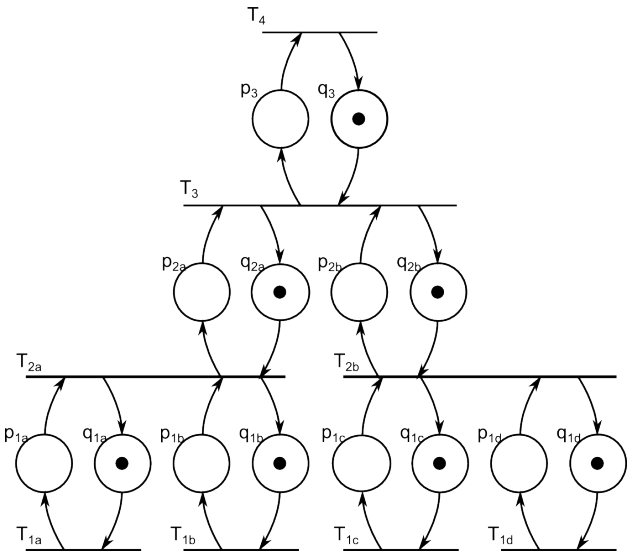


Figure 7: Petri-net and initial marking for asynchronous production binary tree for $N = 3$.

Calculation of the matrix product system and redistributed componentwise exponential matrix is as outlined previously.

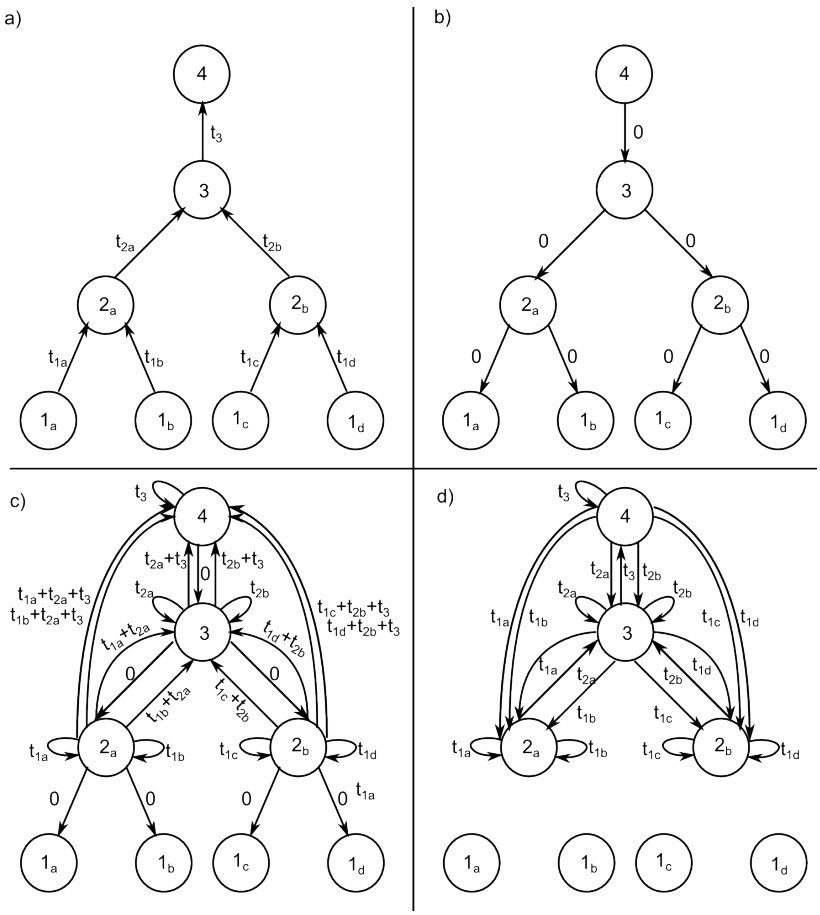


Figure 8: Graphs associated with $A$, $B$, $A^*B$ and $\widehat{A}$ respectively for $N = 3$.

To compute the adjacency eigenvalue of the graph for $\widehat{A}$ we first quotient the vertex set so that we consider sets of depth $i$ vertices. The adjacency matrix for this quotient is then given by

$$\mathcal{A}_{i,j} = \begin{cases} 1 & \text{if } i = 1,\ j = 1, 2 \\ 2 & \text{if } j \leq i + 1 \geq 3 \\ 0 & \text{otherwise} \end{cases} \tag{87}$$

which will have the same principal eigenvalue as the whole adjacency matrix. Note that this adjacency matrix is less than (componentwise) the adjacency matrix associated with the production line of length $N$ multiplied by 2. Therefore the apparent asymptotes in the bound and the exponent definitely do exist.

## 3.5 Partially synchronous production binary tree

We can generalize our production line model further by requiring some anti-chains of processes to be completed synchronously. In terms of the production line model this means demanding that some processes that do not share any precursors must start at exactly the same time, so that one process may have to wait for the other

to be ready in order for it to begin. Formally we take a partition of $\{1, 2, ..., N\}$ into disjoint subsets then define a tree order on this set of subsets.

We consider a binary tree system in which all the processes of depth $d$ are synchronized. This is equivalent to defining a total order on the partition taken by grouping together processes of the same depth in the binary tree.

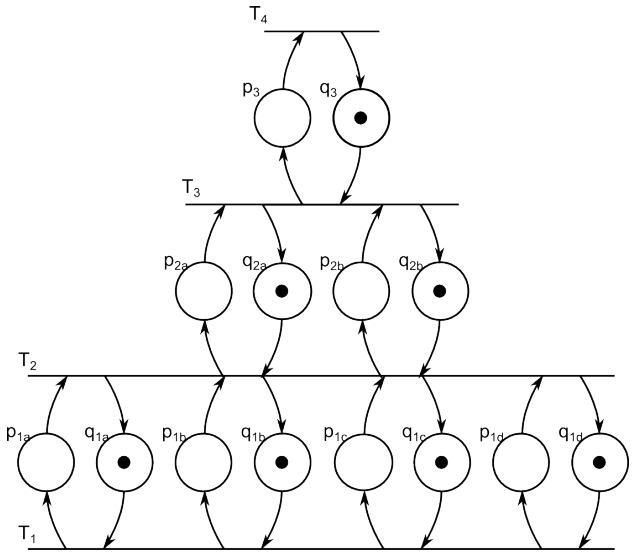


Figure 9: Petri net and initial marking for partially synchronous production binary tree for N=3.

Within this slightly more general framework the edge weight redistribution process still works in the same way.

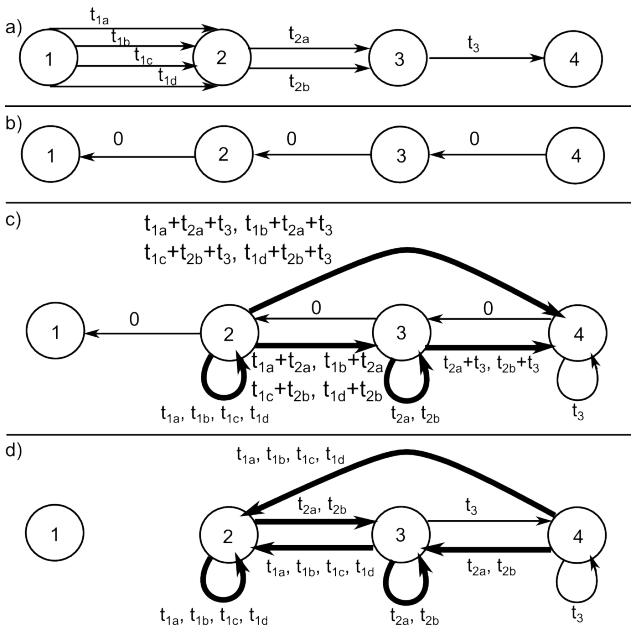


Figure 10: Graphs associated with $A$, $B$, $A^*B$ and $\widehat{A}$ respectively for $N = 3$, bold edges represent sets of more than one edge whose weights are listed in the figure.

The adjacency matrix associated with the redistributed system is given by

$$\mathcal{A}_{i,j} = \begin{cases} 2^{i-1} & \text{for } j \leq i+1 \\ 0 & \text{otherwise} \end{cases} \tag{88}$$

and using

$$\min_{j=1}^{N} \sum_{i=1}^{N} \mathcal{A}_{i,j} \leq \Lambda \leq \max_{j=1}^{N} \sum_{i=1}^{N} \mathcal{A}_{i,j} \tag{89}$$

we have the inequality

$$2^{N-2} + 2^{N-1} = 3 \times 2^{N-2} \leq \Lambda \leq \sum_{j=1}^{N} 2^{j-1} = 2^N - 1 \tag{90}$$

so that

$$\lim_{N\to\infty} \frac{\lambda}{N} = \log 2 \tag{91}$$

# Conclusion

## 3.6 Examples

Our simulations and bounds confirm that the asynchronous production tree has slightly lower throughput than the production line but that individual products are produced much faster. Enforced synchrony reduces throughput catastrophically but individual items are produced in a reasonable amount of time. We summarize our

results in the following table which shows throughput and production time in the limit $N \to \infty$, so for very large production trees. $[\lambda_1 \approx 2.5] < [\lambda_2 \approx 3.5]$.

| Production model | Throughput | Production time |
|---|---|---|
| Asynchronous Line | $\lambda_1^{-1}$ | $\lambda_1 N$ |
| Asynchronous tree | $\lambda_2^{-1}$ | $\lambda_2 \log_2 N$ |
| Partially synchronous tree | $(\log N)^{-1}$ | $\log N \log_2 N$ |

## 3.7 Theory

For smaller values of $\Lambda$ our bounds are not too sharp which is why we are unable to use them to approximate $\lambda$ with much accuracy, however they still prove essential for proving the existence of the asymptote in $\lambda$ as $N \to \infty$. For the partially synchronous case where $\Lambda \to \infty$ the bounds gives us an approximation of $\lambda$ that converges in ratio as $N \to \infty$. The availability of such results is formalized in the corollary to our main theorem which we recall here.

**Corollary** Suppose that $[M_N(n)]_{n=1}^{\infty}$ is a sequence of i.i.d. componentwise exponential max-plus matrices parameterized by $N \in \mathbb{N}$.

- If $\max_N \Lambda_N = \overline{\Lambda}$ exists then $\max_N \lambda_N = \overline{\Lambda}$ exists.
- If on the other hand $\Lambda_N \to \infty$ then $\lim_{N\to\infty} \frac{\lambda_N}{\log \Lambda_N} = 1$
- Moreover if, as in all our examples, $M_N(n)$ can be realized as submatrix of $M_{N+K}(n)$ for all $N, K \geq 0$ then either $\lim_{N\to\infty} \Lambda_N$ exists and so does $\lim_{N\to\infty} \lambda_N$ or $\lim_{N\to\infty} \frac{\lambda_N}{\log \Lambda_N} = 1$

The ideas used in the proof of the bounds rest heavily on the maximally weighted path persecutive. In the proof of the lower bound we use a Markovian strategy to obtain a highly weighted path and in the proof of the lower bound we show exploit the special *Association* between different paths weights. These ideas could form the basis of a similar result for further classes of Max-plus matrices, indeed a similar upper bound can easily be obtained for *Componentwise Gaussian* matrices where we have

$$\lambda \asymp \sqrt{\log \Lambda} \tag{92}$$

The classical theory of graph eigenvalues gives us a useful toolbox for investigating the value of $\Lambda$ in our examples which made its calculation fairly easy. In less structured systems it could still be obtained numerically with great accuracy much faster than any Max-plus exponent approximation.

## 3.8 Scope

The Max-plus formulation of a timed even graph will work for any such system with any sort of waiting time distribution. Our main theorem relied on a *Componentwise exponential* form which we obtained through a sort of combinatorial change of variables, this transformation will clearly not work for any timed event graph but similar techniques will no doubt be of use in other examples.

### 3.9 Further work

In further work we hope to generalize our result to a more general class of matrices, clearly the tail characteristics of the waiting time distributions play an important role here so something like the Perato distribution whose tail weight can be parameterized would be particularly interesting. It would also be interesting to investigate what else can be said about the dynamics of a timed event graph from the maximally weighted path perspective, in particular the importance of the maximally weighted path's path, that is the sequence of vertices it visits. A statistical result would be especially desirable here.

# References


[1] Bernd Heidergott, Geert Jan Olsder, Jacob van der Woude. (2005). Max Plus at Work. *Princeton Series in Applied Mathematics.*

[2] Heidergott, Bernd F. (2007). Max-Plus Linear Stochastic Systems and Perturbation Analysis. *Springer.*

[3] Kolokoltsov, Vassili N., Maslov, Victor P. (1997). Idempotent Analysis and Its Applications *Springer.*

[4] David Speyer, Bernd Sturmfels. (2004). The tropical Grassmannian *Advances in Geometry.* Vol. 4, No. 3

[5] Itenberg, Ilia, Mikhalkin, Grigory, Shustin, Eugenii I. (2009). Tropical Algebraic Geometry. *Springer.*

[6] P. Butkovic. (2010). Max-linear Systems: Theory and Algorithms *Springer.*

[7] T Muranta .(1989). Petri nets: Properties, analysis and applications. *Proceedings of the IEEE* Vol. 77, No. 4.

[8] C.A. Brackley, D. Broomhead, M. C. Romano and M. Thiel. (2011). A Max-Plus Model of Ribosome Dynamics During mRNA Translation. *In preparation.* arXiv:1105.3580v1