Products of random Max-plus matrices

Hook, J

2011

MIMS EPrint: **2011.108**

ISSN 1749-9097

Products of random Max-plus matrices **DRAFT**

James Hook
School of Mathematics, University of Manchester

Abstract

Max-plus stochastic linear systems describe a wide variety of non-linear queueing processes. The dynamics of these systems are dominated by a Max-plus analogue of the Lyupanov exponent the value of which depends on the structure of the underlying support graphs as well as the properties of the waiting-time distributions. For matrices whose associated weighted graphs have identically distributed edge weights (componentwise homogeneity) we are able to decouple these two effects and provide a sandwich of bounds for the Max-plus Lyupanov exponent relating it to some classical properties of the support graph and some extreme value expectations of the waiting-time distributions. This sandwich inequality is then applied to products of componentwise exponential, Gaussian and uniform matrices.

1 Introduction

The Max-plus algebra gives us an alternative way to look at a number of interesting classically non-linear phenomena [1-8]. 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 nth 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].

In this paper we consider products of random Max-plus matrices. The Petri-net formalism provides a modeling language in which we can describe a stochastic queueing system as a type of graph then systematically obtain a Max-plus linear system describing its evolution [1, 2, 8].

1.1 Application - Highly parallel asynchronous computing

Although the Petri-net formalism provides the scope to describe many different systems in a Max-plus linear way we are able to model the following less general example directly.

Suppose that we implement some iterative algorithm, for solving a PDE say, on a large network of processors. We can model this network as a graph $G = \langle V, E \rangle$ where vertices represent processors and an edge (i, j) indicates that processor i requires information from processor j in order to perform its computation. We now assign each processor i with some task which it performs in time $c_i(1)$ then broadcast its result to its successors, all $j \in V$ such that $(j, i) \in E$, each communication taking time $s_{j,i}(1)$. Once i has performed its first computation and received the results of its predecessors first computations, all $j \in V$ such that $(i, j) \in E$, it begins its second computation taking time $c_i(2)$ which again it broadcasts to its successors taking time $s_{j,i}(2)$ and so on. Therefore if $x_i(n)$ is the time at which processor i begins its nth computation we have

$$x(n+1)_i = \max\{x(n)_i + c_i(n), \max_{j \in V:(i,j)\in E} x(n)_j + c_j(n) + s_{i,j}(n)\} \tag{2}$$

$$= A(n) \otimes x(n) = [A(n) \otimes A(n-1) \otimes ... \otimes A(1) \otimes \underline{x}(1)]_i \tag{3}$$

where $A(n)$ is the Max-plus matrix given by

$$A(n)_{i,j} = \begin{cases} c_j(n) + s_{i,j}(n) & \text{if } (i,j) \in E \\ c_i(n) & \text{if } i = j \text{ and } (i,i) \in \overline{E} \\ -\infty & \text{otherwise} \end{cases} \tag{4}$$

And if we model the waiting times $c_i(n)$ and $s_{i,j}(n)$ as random variables then our system is described by a product of random Max-plus matrices.

1.2 Technical restrictions

In this paper we consider products of i.i.d. fixed support, componentwise homogeneous, associated, out independent, $N \times N$ Max-plus matrices.

A sequence of fixed support Max-plus matrices can be defined in terms of an *associated graph* as follows. Let $G = \langle \mathbb{N}_N, E \rangle$ be a directed graph, which may contain multiple (parallel) edges between the same two vertices in the same direction. Now to each edge in $e \in E$ assign a sequence of weights $[t(e)_n]_{n=1}^{\infty}$. The matrix sequence defined by this associated graph is then given by

$$A(n)_{i,j} = \max_{e \in E: j \mapsto i} t(e)_n \tag{5}$$

The i.i.d. property requires that the edge weights are i.i.d. w.r.t. n i.e. that $[\{t(e)_n : e \in E\}]_{n=1}^{\infty}$ is an i.i.d. sequence.

Componentwise homogeneity requires the $t(e)_n$ to have identical distribution for all $n \in \mathbb{N}$, $e \in E$.

Componentwise associativity requires the r.v's $\{t(e)_n : e \in E\}$ to be are associated for each n (see section 3 and not that independent $\Rightarrow$ associated).

Out independence requires that r.v's $\{t(e)_n : e \in E : j \mapsto V\}$ be independent for each n and $j \in V$.

We further define an *associated adjacency matrix* $\mathcal{A}$ to this graph with $\mathcal{A}_{i,j}$ equal to the number of edges from j to i in G.

For the example system outlined above to fit these restrictions we would require that either $s_{i,j} = 0$ for all $(i,j) \in E$ or $(i,i) \in E$ for all i, this ensures that each edge weight in the associated graph has the form $c_j(n)$ or each edge weight has the form $c_j(n) + s_{i,j}(n)$. There are a number of ways to satisfy the remaining conditions, the simplest being that the $c_j(n)$ are i.i.d. as well as the $s_{i,j}(n)$. We then consider the transpose matrices $[A^T(n)]_{n=1}^{\infty}$ as these will be out independent - unlike the original $[A(n)]_{n=1}^{\infty}$ where the for fixed i the $A_{j,i}$ are associated and not independent. Which reconciles with the original system since

$$[\bigotimes_{n=1}^{k} A^T(n)]^T = \bigotimes_{n=k}^{1} A(n) \tag{6}$$

is identically distributed to $\bigotimes_{n=1}^{k} A(n)$ by the i.i.d. property.

1.3 Outline

The remainder of this paper is outlined as follows. In section 2 we introduce the Max-plus Lyupanov exponent and give a new proof of its existence for products of i.i.d matrices then extend this result to the reducible periodic case. In section 3 we prove a sandwich inequality relating the Max-plus Lyupanov exponent of a product of i.i.d. fixed-support, componentwise-homogenous Max-plus matrices to some extreme value expectations of the associated waiting time distributions and some classical properties of the associated adjacency matrices. In section 4 we apply this theorem to three important examples, our results can be summarized as follows

- For componentwise exponential mean-μ matrices $\lambda \asymp \mu \log \mathcal{X}$
- For componentwise Gaussian-(μ, σ) matrices $\lambda \asymp \mu + \sigma\sqrt{\log \mathcal{X}}$
- For componentwise uniform-$[a,b]$ matrices $\lambda \asymp b - \frac{b-a}{\mathcal{X}}$

Where $\mathcal{X}$ is the largest (classical) eigenvalue of the associated adjacency matrix.

2 Max-plus Lyupanov exponent

Products of Max-plus matrices of the form

$$P(n) = \bigotimes_{k=1}^{n} A(k) \tag{7}$$

can be interpreted as follows. $P(n)_{i,j}$ is given by the weight of the maximally weighted path σ of length n through the associated graphs from j to i which accumulates weight on its k'th step according to $A(k)$ so that

$$P(n)_{i,j} = \max_{k_1,k_2,\ldots,k_{n-1}} A(1)_{k_1,j} + A(2)_{k_2,k_1} + \ldots + A(n)_{i,k_{n-1}} = \max_{\sigma} W(\sigma) \tag{8}$$

Where the second maximum is taken over all paths of length n from j to i and

$$W(\sigma) = \sum_{k=1}^{n} A(k)_{\sigma(k+1),\sigma(k)} \tag{9}$$

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 Let $[A(n)]_{n=1}^{\infty}$ be an i.i.d. sequence of matrices with fixed support. Provided the associated graph is irreducible and aperiodic then with probability-1 the limit

$$\lambda = \lim_{n\to\infty} \frac{1}{n} P(n) = \lim_{n\to\infty} \frac{1}{n} \bigotimes_{k=1}^{n} A(k) \tag{10}$$

exists and is a matrix with each element equal to the same constant, the *max-plus Lyapunov exponent* which we shall also denote λ.

Remark Note that since conventional multiplication acts like taking powers in Max-plus we are justified in calling this an exponent. However since the Max-plus exponent acts in an arithmetic as opposed to the classical geometric way it tells us nothing about the stability of the system, instead it gives us the average rate of growth which is common to all solutions

$$\lim_{n\to\infty} \frac{X(n) = [\bigotimes_{k=1}^{n} A(k)] \otimes X(0)}{n} = \lambda \tag{11}$$

We only call λ a Max-plus *Lyupanov* exponent because it emerges from this Max-plus multiplicative ergodic theorem in an analogous way.

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

Claim 1 $\lambda_n = \frac{1}{n}\mathbb{E}\{\max_{i,j}[\bigotimes_{k=1}^{n} A(k)]_{i,j}\}$ is a decreasing sequence bounded below 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 nth 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

$$\begin{aligned}(n+m)\lambda_{n+m} = \mathbb{E}\{\max_{i,j}[\bigotimes_{k=1}^{n+m} A(k)]_{i,j}\} &\leq \mathbb{E}\{\max_{i,j,k,l}[\bigotimes_{k=1}^{n} A(k)]_{i,k} + [\bigotimes_{k=n+1}^{n+m} A(k)]_{l,j}\} \\ &= n\lambda_n + m\lambda_m\end{aligned} \tag{12}$$

and the claim follows from

Fekete's Subadditive Lemma A sequence $(a_n)_{n=1}^{\infty}$ is subbaditive iff $a_{n+m} \leq a_n + a_m$ for all n, m and for such a sequence

$$\lim_{n\to\infty} \frac{a_n}{n} = \min_n \frac{a_n}{n} \tag{13}$$

exists.

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

proof Clearly the weight of the path on the right hand side must be greater than that on the left. Consider the path σ that attains the maximum weight in the right hand side of the equation.

Claim 3 Provided the graph associated with M is irreducible and aperiodic there exists a path ς from i to j that coincides with σ from the $N^2 + N$th step to the $n - N^2 - N$th step.

Proof It suffices to show that for all $n > N^2 + N$ any vertex is reachable from any other in exactly n steps. Suppose we start at vertex j which must lie on at least two cycles of co-prime length p, q say. Therefore for all $n > pq < N^2$ every vertex is reachable from itself. Finally for each i there is a path from j to i of length $< N$.

C is just the difference in weight between the two paths whose expectation can easily be bounded by $2\lambda_1(N^2 + N)$.

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

$$\lim_{n,m\to\infty} \frac{1}{nm}[\bigotimes_{k=1}^{nm} A(k)]_{i,j} = \lim_{n,m\to\infty} \frac{1}{m}(\frac{1}{n}[\bigotimes_{k=1}^{n} A(k)]_{i,\sigma(n)} + \tag{14}$$

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

We can now use claim 2 to replace

$$\frac{1}{n}[\bigotimes_{k=tn+1}^{n(t+1)} A(k)]_{\sigma(tn+1),\sigma(n(t+1)} \tag{15}$$

with

$$\frac{1}{n}\max_{i',j'}[\bigotimes_{k=tn+1}^{n(t+1)} A(k)]_{i',j'} - \frac{C}{n} \tag{16}$$

Taking the limit in m the first part of each term gives us λ_n by the law of large numbers. Then taking the limit in n gives us λ 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

$$\lambda = \lim_{n\to\infty} \frac{1}{n} A^{\otimes n} \tag{17}$$

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 fewer than N (the number of vertices) 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.

2.1 Irreducible Periodic case

Suppose now that the graph associated with our matrices is periodic, clearly the previous result can not hold here. For instance it may be impossible to reach i from j in exactly n steps so that $P(n)_{i,j} = -\infty$. However we can factor the graph into periods and use the aperiodic case to prove the following.

Theorem $[A(n)]_{n=1}^{\infty}$ be an i.i.d. sequence of matrices with fixed support and that the graph associated with A is irreducible and periodic with period-p. The limit

$$\lambda = \lim_{m\to\infty} \frac{1}{mp - p(j) + p(i)} P[mp - p(j) + p(i)]_{i,j} \tag{18}$$

exists and is independent of i, j. Where we have factored G into a cyclic graph on p subsets $M_1, M_2, ..., M_p$ such that $i \in M_{p(i)}$ and $j \in M_{p(j)}$. Furthermore

$$P(n)_{i,j} = -\infty \tag{19}$$

whenever n is not congruent to $-p(j) + p(i)$ modulo p.

Proof We can factorize $P[mp - p(j) + p(i)]$ as follows

$$A[mp-p(j)+p(i)] \otimes ... \otimes A[(m-1)p+1] \otimes \underbrace{A[(m-1)p] \otimes ... \otimes A[(m-2)p+1]} \tag{20}$$

$$\otimes ... \otimes \underbrace{A(p) \otimes ... \otimes A(1)}$$

$$= D \otimes B(m-1) \otimes B(m-2) \otimes ... \otimes B(1) \tag{21}$$

Where $[B(k)]_{k=1}^{\infty}$ is a sequence of i.i.d. Max-plus matrices whose associated graph consists of p connected components each of which is irreducible and aperiodic with vertices from the M_i. Therefore

$$\lim_{n\to\infty} \frac{P(n)_{i,j}}{n} = \lim_{m\to\infty} \max_{k\in M_{p(j)}: D_{i,k}\neq-\infty} \frac{[B(m-1) \otimes ...B(1)]_{i,k}}{p(m-1)} \tag{22}$$

where $n = mp - p(j) + p(i)$ and from our previous theorem this limit equals $\frac{\lambda[p(j)]}{p}$ where $\lambda[p(j)]$ is the exponent of the $[B(k)]_{k=1}^{\infty}$ restricted to $M_{p(j)}$.

Claim $\lambda(i) \leq \lambda(j)$ for all i, j.

Proof Consider $\sigma(j)$ the maximally weighted path of length mp starting in M_j at step 0, and $\widehat{\sigma(i)}$ the maximally weighted path of length mp starting in M_i but at step $p - p(j) + p(i)$. We can construct a path ς that starting in M_j joins $\widehat{\sigma(i)}$ in fewer than p steps and shadows it for at least $mp - 2p$ steps. By the i.i.d. property $\widehat{\sigma(i)}$ has identically distributed weight to $\sigma(i)$ so that in the limit $m \to \infty$ we have the claim.

Therefore $\lambda(i) = \lambda$ is independent from i and the proof of the theorem is complete.

2.2 Reducible periodic case

If the graph associated with $[A(n)]_{n=1}^{\infty}$ is not irreducible then there are further complications in the asymptotic behavior it may for instance be impossible to reach i from j so that $P(n)_{i,j} = -\infty$ for all n. However since a reducible graph G can be represented as an acyclic graph H on G's maximal irreducible components we are able to use the result in the irreducible case to prove the following.

Theorem With probability-1 the limit

$$\lambda_{i,j} = \limsup_{n\to\infty} \underbrace{\frac{1}{n}P(n)_{i,j}} \tag{23}$$

exists and is equal to

$$\max\{-\infty, \max_{h\in H: h(i)\preceq h \preceq h(j)} \lambda(h)\} \tag{24}$$

where $h(i)$ is the maximal irreducible component of G containing i, $\preceq$ is the partial order on G's maximal irreducible components induced by reachability and $\lambda(h)$ is the Max-plus exponent of the submatrices obtained by restricting the $[A(k)]_{k=1}^{\infty}$ to h.

Remark This is only a partial result as we are taking the supremum limit. As $n \to \infty$ the underbraced term in (24) will converge to some periodic sequence of exponents each one realized as the exponent of some irreducible component of G. The exact behavior of this sequence can be extremely complicated and is beyond the scope of this paper, we simply find the maximum exponent in the sequence.

Proof For a given sequence of matrices $[A(k)]_{k=1}^{n}$ the maximally weighted path from j to i of length n through the associated weighted graph is $\sigma|_n$ and the maximally weighted path of length n restricted to h with the same periodic phase as j is $\sigma(h)|_n$.

As in the irreducible case it is possible to move from j to h and join $\sigma(h)|_n$ in fewer than $N^2 + N$ steps, a path from j to i of length n can then shadow this locally maximally weighted path for at least $n - 2N^2 - 2N$ steps before moving to i. The weight of this path $\sigma(h)|_n^*$ satisfies

$$\lim_{n\to\infty} \frac{1}{n} W(\sigma(h)|_n^*) = \lim_{n\to\infty} \frac{1}{n} W(\sigma(h)|_n) = \lambda(h) \tag{25}$$

Therefore $\lambda_{i,j} \geq \max_{h\in H:h(i)\preceq h\preceq h(j)} \lambda(h)$

Note that such paths do not necessarily exist for all n but for some non-empty modulo $p(h)$ congruency class.

Now consider $\sigma|_{n\times m}$ the maximally weighted path of length $n \times m$ from j to i. We can divide $\sigma|_{n\times m}$ up into m paths of length n

$$\sigma|_{n\times m} = (\sigma|_n^1, \sigma|_n^2, ..., \sigma|_n^m) \tag{26}$$

Up to L (the number of maximal irreducible components $h : h(i) \preceq h \preceq h(j)$) of these m path segments are spread over more than one irreducible component, they contribute $n \times L$ edge weights and we can therefore ignore them when we take the step average and the limit in m.

$$\lambda_{i,j} = \lim_{n,m\to\infty} \frac{1}{n \times m} W(\sigma|_{n\times m}) = \lim_{n,m\to\infty} \frac{1}{n \times m} \sum_{k=1}^{m} W(\sigma|_n^k) \tag{27}$$

$$\leq \lim_{n,m\to\infty} \sum_{k=1}^{m} \max_{h\in H:h(i)\preceq h\preceq h(j)} W[\sigma(h)|_n^k]$$

Where $\sigma(h)|_n^k$ is the maximally weighted path restricted to h for the edge weight sequence $[A(k)]_{k=n\times(k-1)+1}^{n\times m}$.

$$= \lim_{m\to\infty} \frac{1}{m} \sum_{k=1}^{m} \max_{h\in H:h(i)\preceq h\preceq h(j)} \lim_{n\to\infty} \frac{1}{n} W[\sigma(h)|_n^k] \tag{28}$$

$$= \max_{h\in H:h(i)\preceq h\preceq h(j)} \lambda(h)$$

which completes the proof.

The dynamics of a Max-plus system is therefore dominated by its exponents. The remainder of this paper is devoted to developing a theory to link them to the classical principal eigenvalue of the associated adjacency matrix for three examples, componentwise exponential, Gaussian and uniform.

3 Decoupling effects of graph structure and waiting-time distribution

Let $[A(n)]_{n=1}^{\infty}$ be a sequence of i.i.d. fixed-support, componentwise-homogeneous, associated, out-independent, $N \times N$ Max-plus matrices with ρ-distributed edge weights.

Let $\mathcal{X}$ be the principal eigenvalue of the associated adjacency matric $\mathcal{A}$, and π the unique invariant measure of the Markov chain on the associated graph G defined by the transition probability matrix $Q_{i,j} = \frac{\mathcal{A}_{i,j}}{d_j}$ where d_j is the out degree of j in G.

Now let $[t_{i,j}]_{i,j=1}^{\infty}$ be an array of i.i.d. ρ-distributed r.v's and define

$$Y_n^i = \sum_{j=1}^n t_{i,j} \quad X_n^m = \max_{i=1}^m Y_n^i \quad E_n^m = \mathbb{E}[X_n^m] \tag{29}$$

Theorem Provided G is irreducible the Max-plus Lyupanov exponent

$$\lambda = [\lim_{n\to\infty} \frac{1}{n} \bigotimes_{k=1}^{n} A(k)]_{i,j} \tag{30}$$

which is independent of i, j satisfies

$$\sum_{i=1}^{N} \pi_i E_1^{d_i} \leq \lambda \leq \lim_{n\to\infty} \frac{1}{n} E_n^{\mathcal{X}^n} \tag{31}$$

Remark Note that G needn't be aperiodic, if G is periodic then we define $\mathcal{X}$ to be the largest real positive eigenvalue of $\mathcal{A}$ which uniquely exists.

Lower bound As outlined

$$\gamma_1^n = \max_i [\bigotimes_{k=1}^{n} A(k)]_{i,1} \tag{32}$$

is the weight of the maximally weighted path of length n through G starting at vertex 1 which accumulates weights from the $[A(k)]_{k=1}^n$. We can therefore bound this quantity from below with the weight of any other path of length n which also starts at vertex 1 and accumulates weight in the same way, in particular we can formulate a strategy for choosing a highly weighted path and use this path's expected weight as a bound for $\mathbb{E}[\gamma_1^n]$ which in turn gives us a lower bound for λ.

Greedy strategy We will construct our highly weighted path σ as follows. First set $\sigma(1) = 1$ now choose $\sigma(2)$ such that

$$A(1)_{\sigma(2),\sigma(1)} = \max_i A(1)_{i,\sigma(1)} \tag{33}$$

and if this maximum is attained by more than one possible i then choose $\sigma(2)$ uniformly randomly from this set. Continue in this way always choosing the maximally weighted available edge at each stage so that

$$A(n)_{\sigma(n+1),\sigma(n)} = \max_i A(n)_{i,\sigma(n)} \tag{34}$$

Now since the $[A(k)]_{k=1}^{\infty}$ are i..i.d. this procedure gives rise to a Markov chain on G and since the edge weights are homogeneous and out independent we have

$$\mathbb{P}[\sigma(n+1) = i | \sigma(n) = j] = \frac{\mathcal{A}_{i,j}}{d_j} = Q_{i,j} \tag{35}$$

This follows since the probability of transition (i, j) is equal to the probability that one of the $\mathcal{A}_{i,j}$ edge weights from j to i attain the maximum of the d_j edge weights leaving j. By componentwise homogeneity and out independence this is simply the fraction of edges from j that go to i.

The step averaged weight of the first n edges of σ is therefore given by

$$\frac{1}{n}W_n(\sigma) = \frac{1}{n}\sum_{k=1}^{n} A(k)_{\sigma(k+1),\sigma(k)} \tag{36}$$

and we can rearrange the sum to obtain

$$\frac{1}{n}W_n(\sigma) = \sum_{i=1}^{N} \frac{1}{n} \sum_{j=1}^{T_n(i)} A[I(j)]_{\sigma(I(j)+1),i} \tag{37}$$

Where $I(j)$ is the jth k s.t. $\sigma(k) = i$ and $T_n(i)$ is the total number of such $k \leq n$. Noting that $A[I(j)]_{\sigma(I(j)+1),i}$ is exactly the maximum of d_i i.i.d. ρ-distributed r.v.'s then applying the law of large numbers to the $\frac{T_n(i)}{n}$ we have

$$\lim_{n\to\infty} \frac{1}{n}W_n(\sigma) = \lim_{n\to\infty} \sum_{i=1}^{N} \frac{1}{n} \sum_{j=1}^{T_n(i)} X_1^{d_i} = \lim_{n\to\infty} \sum_{i=1}^{N} \frac{\pi_i}{T_n(i)} \sum_{j=1}^{T_n(i)} X_1^{d_i} \tag{38}$$

finally applying the law of large numbers to the $X_1^{d_i}$ we obtain

$$\lim_{n\to\infty} \frac{1}{n}W_n(\sigma) = \sum_{i=1}^{N} \pi_i E_1^{d_i} \tag{39}$$

which is our lower bound

Upper bound λ_n^1 the expected step averaged weight of the maximally weighted path of length n that stats at vertex 1 can, of course, be expressed in terms of path weights

$$\lambda_n^1 = \frac{1}{n}\mathbb{E}\max_i [\bigotimes_{k=1}^{n} A(k)]_{i,1} = \frac{1}{n}\mathbb{E}\max_{\sigma|_n} W_n(\sigma) \tag{40}$$

Where the maximum is taken over all paths of length n starting at 1 and the weight function W_n defines a path's weight using the components of the matrices $[A(k)]_{k=1}^n$ in the usual way. It is important that in evaluating this expression the same waiting times are used for each different path.

The weight of each path of length-n is a sum of n i.i.d. ρ-distributed r.v's 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 associated 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 λ from above by taking the maximum in (34) while ignoring the dependence and treating each path weight as an i.i.d. sum of i.i.d. ρ-distributed r.v's.

Claim 1 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 $J \in \{1, 2, .., N\}/\{i\}$ 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 equals

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

which simplifies to

$$\frac{P(f, g)}{P(g)} \geq P(f) \tag{45}$$

so that

$$\mathbb{P}[x_i < T | x_j < T, j \in J] \geq \mathbb{P}[x_i < T] \tag{46}$$

Now consider

$$\mathbb{P}[\max_{i=1}^{N} x_i < t] = \prod_{i=1}^{N} \mathbb{P}[x_i < t | x_k < t; k = 1, 2, ..., i-1] \tag{47}$$

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

$$\mathbb{P}[\max_{i=1}^{N} x_i < t] \geq \prod_{i=1}^{N} \mathbb{P}[x_i < t] \tag{48}$$

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

$$\mathbb{E} \max_{i=1}^{N} x_i = \int_0^\infty t\rho_x(t)dt = \int_0^\infty \int_z^\infty \rho_x(t)dtdz \tag{49}$$

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

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

$$\lambda \leq \lim_{n \to \infty} \frac{1}{n} E_n^{D(n)} \tag{50}$$

Where $D^1(n)$ is the number of paths of length n starting at 1 through G.

Claim 2 $D_n^1 \leq \mathcal{X}^N$

Proof Just define the vector $\underline{D}^1(n)$ whose ith component is the number of paths of length-n which start at vertex 1 and end at vertex i then

$$\underline{D}(n) = \mathcal{A}\underline{D}(n-1) = \mathcal{A}^n \underline{D}^1(0) \tag{51}$$

where $\underline{D}^1(0) = \underline{e}_1$. So that

$$D^1(n) = |\underline{D}(n)|_{l_1} \leq \mathcal{X}^n |\underline{D}^1(0)|_{l_0} = \mathcal{X}^n \tag{52}$$

We therefore have the upper bound

$$\lambda \leq \lim_{n \to \infty} \frac{1}{n} E_n^{\mathcal{X}^N} \tag{53}$$

4 Examples

We can now substitute any distribution and graph that we like into (25) to obtain a sandwich inequality for the Max-plus Lyupanov exponent. The lower bound $\sum_{i=1}^{N} \pi_i E_1^{d_i}$ is difficult to treat for a general directed graph and general distribution as the invariant measure π can't be expressed in a closed form. However for our three example distributions we can obtain a good lower bound on this bound using the following results.

Lemma For an irreducible, directed graph

$$\sum_{i=1}^{N} \pi_i \log d_i = \log \mathcal{X} \tag{54}$$

Proof Define a co-cycle on the random walk σ with transition probabilities $Q_{i,j}$ and initial state 1 by

$$C(\sigma, n) = \frac{C(\sigma, n-1)}{d_{\sigma(n)}} \tag{55}$$

So that $C(\sigma, n)$ is the probability that starting at 1 a random walk follows σ for the first n steps. By the central limit theorem for Markov chains there exists μ and ν such that

$$\lim_{n\to\infty} \frac{\log C(\sigma, n) - \mu n}{\nu\sqrt{n}} = N(\sigma) \tag{56}$$

converges in distribution to a standard $(0, 1)$-Gaussian. Now since $C(\sigma, n)$ represents a probability

$$\sum_{\sigma|_n} C(\sigma, n) = 1 \tag{57}$$

where the sum is taken over all paths of length n that begin at 1. Therefore taking logs and dividing by n gives

$$\lim_{n\to\infty} \frac{1}{n} \log[\sum_{\sigma|_n} C(\sigma, n)] = \lim_{n\to\infty} \frac{1}{n} \log[\sum_{\sigma|_n} e^{\mu n + \nu\sqrt{n}N(\sigma)}] = 0 \tag{58}$$

Taking out the constant terms and applying the law of large numbers to the log normal we have

$$= \mu + \lim_{n\to\infty} \frac{1}{n} \log(D_n^1 \mathbb{E}[e^{\nu\sqrt{n}N}]) = \mu + \log \mathcal{X} + \lim_{n\to\infty} \frac{\nu}{\sqrt{2n}} = 0 \tag{59}$$

Where D_n^1 is the number of paths of length n that start at 1. Therefore we have $\mu = -\log \mathcal{X}$. Finally note that with probability-1

$$\lim_{n\to\infty} \frac{1}{n} \log C(\sigma) = \mu = -\log \mathcal{X} \tag{60}$$

$$= \lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} \log \frac{1}{d_{\sigma(k)}} = \lim_{n\to\infty} -\sum_{i=1}^{N} \frac{T_n(i)}{n} \log d_i = -\sum_{i=1}^{N} \pi_i \log d_i$$

where $T_n(i)$ is greatest $k \le n$ such that $\sigma(k) = i$.

Jensen's inequality If f is convex then $f[\mathbb{E}(X)] \leq \mathbb{E}[f(X)]$

The upper bound $\lambda \leq \lim_{n\to\infty} \frac{1}{n} E_n^{N\mathcal{X}^N}$ requires a novel approach since we are unable to apply standard techniques. The central limit theorem does not apply to the sum since taking the maximum over so many i.i.d. distributions forces us to look at sets with very small measures and the convergence of the central limit theorem is not sufficiently uniform for this. We are also unable to apply standard extreme value theory to the maximum as again the other operation - adding further i.i.d. r.v's - interferes with the convergence.

4.1 Componentwise exponential

The following results apply to mean-1 exponentials. For the mean-μ case multiply all bounds by μ.

Corollary Let $[A(n)]_{n=1}^{\infty}$ be a sequence of i.i.d. fixed-support componentwise-mean-1-exponential Max-plus matrices. Then the Max-plus exponent λ satisfies

$$\log \mathcal{X} \leq \lambda \leq \alpha \tag{61}$$

where α is the solution to

$$\alpha = \log \mathcal{X} + \log(1 + e\alpha) \tag{62}$$

and e is the constant 2.71...

Remark Our upper bound does not have a closed form, however it is easy to show that the ratio $r = \frac{\alpha}{\log \mathcal{X}}$ is decreasing in $\mathcal{X}$ and tends to one in the limit $\mathcal{X} \to \infty$.

Lower bound The lower bound follows more or less directly from our lemma.

Claim The expectation of the maximum of d i.i.d. mean-1 exponentials is given by $\sum_{k=1}^{d} \frac{1}{k}$

Proof t_1 the minimum of d i.i.d. mean-1 exponentials is a mean $\frac{1}{d}$ exponential. We can apply the Markov property at this time so that $t_2 - t_1$ the difference between the minimum and second smallest of d i.i.d. mean-1 exponentials is a mean $\frac{1}{d-1}$ exponential independent of t_1. Continuing in this way the expectation of the maximum of the times is given by

$$\mathbb{E}[t_d] = \mathbb{E}[t_1] + \mathbb{E}[t_2 - t_1] + ... + \mathbb{E}[t_d - t_{d-1}] = \sum_{k=1}^{d} \frac{1}{k} \tag{63}$$

So that

$$\lambda \geq \sum_{i=1}^{N} \pi_i \sum_{k=1}^{d_i} \frac{1}{k} > \sum_{i=1}^{N} \pi_i \log d_i = \log \mathcal{X} \tag{64}$$

Upper bound 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 $[X_n^{\mathcal{X}^n}]_{n=1}^{\infty}$ by

$$\mathbb{P}[X_n^{\mathcal{X}^n} < \mu_n(p_n)] = \mathbb{P}[Y_n < \mu_n(p_n)]^{\mathcal{X}^n} = p_n \tag{65}$$

And using the identity

$$e^x = \lim_{m\to\infty}(1 + \frac{x}{m})^m \tag{66}$$

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

$$\mathbb{P}[Y_n < \mu_n(p_n)] = 1 + \frac{\log p_n}{\mathcal{X}^n} \tag{67}$$

$$\mathbb{P}[Y_n > \mu_n(p_n)] = -\frac{\log p_n}{\mathcal{X}^n}$$

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

$$= \int_{\mu_n(p_n)}^{\infty} \frac{x^{n-1}e^{-x}}{(n-1)!}dx = e^{-\mu_n(p_n)} \underbrace{\sum_{k=0}^{n-1} \frac{\mu_n(p_n)^k}{k!}} \tag{68}$$

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

$$(1 + \frac{x}{n})^n \leq \sum_{k=0}^{n} \frac{x^k}{k!} \leq (1 + \frac{ex}{n})^n \tag{69}$$

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

$$\frac{1}{k!}\frac{n!}{(n-k)!n^k} \leq \frac{1}{k!} \leq \frac{1}{k!}\frac{n!e^k}{(n-k)!n^k} \tag{70}$$

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

$$e^{-\mu_n(p_n)}(1 + \frac{\mu_n(p_n)}{n-1})^{n-1} \leq \frac{-\log p_n}{\mathcal{X}^n} \leq e^{-\mu_n(p_n)}(1 + \frac{e\mu_n(p_n)}{n-1})^{n-1} \tag{71}$$

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

$$\lim_{n\to\infty} \frac{\log\log p_n}{n} = 0 \tag{72}$$

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

$$\log \mathcal{X} \leq \frac{\mu_n}{n} \leq \log \mathcal{X} + \log(1 + \frac{e\mu_n}{n}) \leq C \log \mathcal{X} + D \tag{73}$$

for some bounded C D that do not depend on $\mathcal{X}$.

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 $C \log \mathcal{X} + D$ follows from the fact that

$$F(\alpha) = \log \mathcal{X} + \log(1 + e\alpha) \tag{74}$$

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

$$\frac{\delta\alpha^*}{\delta \log \mathcal{X}} = \frac{1 + e\alpha^*}{1 + e(\alpha^* - 1)} \tag{75}$$

So that for any fixed finite $\mathcal{X}'$ we have

$$\alpha^*(\mathcal{X}) \leq \alpha^*(\mathcal{X}') + \frac{\delta\alpha^*}{\delta \log \mathcal{X}}|_{\alpha^*}(\mathcal{X}') \log \mathcal{X} \tag{76}$$

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

$$\lim_{n\to\infty} \frac{\mu_n(e^{\frac{-1}{n}}) - \mu_n(e^{-n})}{n} = 0 \tag{77}$$

Proof From equation (62) we have

$$\frac{\mu_n(e^{\frac{-1}{n}})}{n} = \log \mathcal{X} + \frac{\log n}{n} + \log(\sum_{k=0}^{n-1} \frac{\mu_n(e^{\frac{-1}{n}})^k}{k!}) \tag{78}$$

and

$$\frac{\mu_n(e^{-n})}{n} = \log \mathcal{X} - \frac{\log n}{n} + \log(\sum_{k=0}^{n-1} \frac{\mu_n(e^{-n})^k}{k!}) \tag{79}$$

So that

$$\frac{\mu_n(e^{\frac{-1}{n}}) - \mu_n(e^{-n})}{n} \leq \frac{2 \log n}{n} \tag{80}$$

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

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

$$\max_{\log \mathcal{X} \leq \alpha \leq C \log \mathcal{X} + D} 1 - \frac{\frac{(\alpha n)^{n-1}}{(n-1)!}}{\sum_{k=0}^{n-1} \frac{(\alpha n)^k}{k!}} \tag{81}$$

which gives

$$\frac{\mu_n(e^{\frac{-1}{n}}) - \mu_n(e^{-n})}{n} \leq \frac{2 \log n}{n} \underbrace{\max_{\log \mathcal{X} \leq \alpha \leq C \log \mathcal{X} + D} \frac{(n-1)!}{(\alpha n)^{n-1}} \sum_{k=0}^{n-1} \frac{(\alpha n)^k}{k!}} \tag{82}$$

where the underbraced term can be bounded above by

$$\max_{\log \mathcal{X} \le \alpha \le C \log \mathcal{X} + D} \sum_{k=0}^{n-1} \frac{1}{\alpha^{n-k-1}} \le \max_{\log \mathcal{X} \le \alpha \le C \log \mathcal{X} + D} (1 - \frac{1}{\alpha})^{-1} \tag{83}$$

So that

$$\frac{\mu_n(e^{\frac{-1}{n}}) - \mu_n(e^{-n})}{n} < \frac{2 \log n}{n} (1 - \frac{1}{\log \mathcal{X}})^{-1} \to 0 \tag{84}$$

as required.

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

$$E_n^{\mathcal{X}^n} = \int_0^{\mu_n(e^{-n})} x\rho(x)dx + \underbrace{\int_{\mu_n(e^{-n})}^{\mu_n(e^{\frac{-1}{n}})} x\rho(x)dx}$$

$$+ \int_{\mu_n(e^{\frac{-1}{n}})}^{\mu_n(e^{\frac{-1}{m^n}})} x\rho(x)dx + \sum_{m=2}^{\infty} \int_{\mu_n(e^{\frac{-1}{m^n}})}^{\mu_n(e^{\frac{-1}{(m+1)^n}})} x\rho(x)dx \tag{85}$$

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

$$\int_0^{\mu_n(e^{-n})} x\rho(x)dx \le e^{-n}\mu_n(e^{-n}) \le e^{-n} n(C \log \mathcal{X} + D) \tag{86}$$

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 (65) we obtain

$$\frac{\mu_n(e^{\frac{-1}{m^n}})}{n} < \log \mathcal{X} - \frac{1}{n} \log \log e^{\frac{1}{m^n}} + \log(1 + e\frac{\mu_n(e^{\frac{-1}{m^n}})}{n}) \tag{87}$$

So that

$$\frac{\mu_n(e^{\frac{-1}{m^n}})}{n} < \log m\mathcal{X} + \log(1 + e\frac{\mu_n(e^{\frac{-1}{m^n}})}{n}) \tag{88}$$

and as in (67) we have

$$\frac{\mu_n(e^{\frac{-1}{m^n}})}{n} < C \log m\mathcal{X} + D \tag{89}$$

So the sum of the remaining integrals is therefore bounded by

$$n(e^{\frac{-1}{2^n}} - e^{\frac{-1}{n}})(C \log 2\mathcal{X} + D) + n \sum_{m=3}^{\infty} (e^{\frac{-1}{m^n}} - e^{\frac{-1}{(m-1)^n}})[C \log mK(n) + D] \tag{90}$$

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

$$(1 - e^{-n})n(C \log 2\mathcal{X} + D) + n \sum_{m=3}^{\infty} \frac{C \log m}{(m-1)^n} \tag{91}$$

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

Therefore the mean really does converge in ratio to the median and we can use the bound obtained for the median on the mean.

4.2 Componentwise Gaussian

The following results apply to (0,1) Gaussians. For the(μ, σ) case take $\lambda \times \sigma + \mu$.

Corollary Let $[A(n)]_{n=1}^{\infty}$ be a sequence of i.i.d. fixed-support componentwise-(0,1)-Gaussian Max-plus matrices. Then the Max-plus exponent λ satisfies

$$K\sqrt{\log \mathcal{X}} \leq \lambda \leq \sqrt{2\log \mathcal{X}} \tag{92}$$

where $K > 0$ is a constant independent of G.

Lower bound From [9] we have

Lemma $E_1^d > \mu_d : \mathbb{P}[T > \mu_d] = \frac{1}{d}$ where T is a (0,1)-Gaussian.

Proof Consider the r.v's $(Z_i = \frac{1}{\Psi(T_i)})_{i=1}^d$ where the $(T_i)_{i=1}^d$ are a sequence of i.i.d. (0,1)-Gaussians and Ψ is the Gaussian probability distribution function. Z has distribution $f(x) = z^{-2}$ and

$$\mathbb{E} \min_{i=1}^{d} Z_i = \frac{N}{N-1} \tag{93}$$

Since $\frac{1}{\Psi}$ is convex Jensen's inequality gives

$$\frac{N}{N-1} = \mathbb{E} \min_{i=1}^{d} \frac{1}{\Psi(X_i)} = \mathbb{E}\Psi[\max_{i=1}^{d} T_i] \leq \Psi(E_1^d) \tag{94}$$

Corollary $E_1^d > K\sqrt{\log d}$ where

$$K > \min\{\sqrt{2} - \frac{\log 2\pi}{\log 100}, \min_{n=1}^{100} \frac{E_1^n}{\sqrt{\log n}}\} \tag{95}$$

Proof Borrel's inequality gives $\mu_d > t$ where

$$\frac{1}{\sqrt{2\pi}} \frac{t^2-1}{t^3} e^{\frac{-t^2}{2}} < \frac{1}{d} \tag{96}$$

so that

$$(\frac{t}{\sqrt{\log d}})^2 > 2 + \frac{2\log \frac{t^2-1}{t^3}}{\log d} - \frac{\log 2\pi}{\log d} \tag{97}$$

and provided $t > 1$ which can be guaranteed if a real solution to

$$\frac{1}{\sqrt{2\pi}} \frac{t^2-1}{t^3} e^{\frac{-t^2}{2}} = \frac{1}{d} \tag{98}$$

exists, which is guaranteed so long as

$$\frac{1}{d} < \frac{1}{100} < \max_{t'} \frac{1}{\sqrt{2\pi}} \frac{t'^2 - 1}{t'^3} e^{\frac{-t'^2}{2}} \tag{99}$$

then we have

$$(\frac{t}{\sqrt{\log d}})^2 > 2 - \frac{\log 2\pi}{\log d} > 2 - \frac{\log 2\pi}{\log 100} \tag{100}$$

The expectation of the maximum of d_i i.i.d. Gaussians is therefore bounded below by $K\sqrt{\log d_i}$ and

$$\lambda \geq \sum_{i=1}^{N} \pi_i K \sqrt{\log d_i} = K \sum_{i=1}^{N} \pi_i f(\log d_i) \tag{101}$$

Where $f(x) = \sqrt{x}$ is convex so that Jensen's inequality gives

$$\lambda \geq K \sqrt{\sum_{i=1}^{N} \pi_i \log_{d_i}} = K\sqrt{\log \mathcal{X}} \tag{102}$$

Upper bound Since the sum of n $(0,1)$-Gaussians is a $(0,\sqrt{n})$-Gaussian we have

$$\lambda \leq \lim_{n\to\infty} \frac{\sqrt{n}}{n} \mathbb{E} \max_{i=1}^{\mathcal{X}^n} N_i \tag{103}$$

where $[N_i]_{i=1}^{\infty}$ is a sequence of i.i.d. $(0,1)$-Gaussians. And using the main result from [10] we have

$$\lambda = \lim_{n\to\infty} \frac{1}{\sqrt{n}} \sqrt{2\log \mathcal{X}^n} = \sqrt{2\log \mathcal{X}} \tag{104}$$

4.3 Componentwise uniform

The following results apply to [0,1] Uniform distributions. For the $[a,b]$ case take $(\lambda - \frac{1}{2}) \times (b-a) + \frac{a+b}{2}$.

Corollary Let $[A(n)]_{n=1}^{\infty}$ be a sequence of i.i.d. fixed-support componentwise-uniform-$[0,1]$ Max-plus matrices. Then the Max-plus exponent λ satisfies

$$1 - \frac{1}{\mathcal{X}+1} \leq \lambda \leq 1 - \frac{1}{\mathcal{X}e} \tag{105}$$

Where e is the constant 2.71...

Lower bound The expectation of the maximum of d_i i.i.d. uniforms is $\frac{d_i}{d_i+1}$. Therefore we have

$$\lambda \geq \sum_{i=1}^{N} \pi_i \frac{d_i}{d_i+1} = K \sum_{i=1}^{N} \pi_i f(\log d_i) \tag{106}$$

Where $f(x) = \frac{e^x}{e^x+1}$ is convex so that Jensen's inequality gives

$$\lambda \geq f[\sum_{i=1}^{N} \pi_i \log_{d_i}] = f(\log \mathcal{X}) = 1 - \frac{1}{\mathcal{X}+1} \tag{107}$$

Upper bound Consider the sequence of generalized medians $[\mu_n]_{n=1}^{\infty}$ such that

$$\mathbb{P}[X_n^{\mathcal{X}^n} < \mu_n] = 1 - \frac{1}{n} \tag{108}$$

Then we have $\lambda \leq \lim_{n\to\infty} \frac{1}{n}[\mu_n(1-\frac{1}{n}) + \frac{n}{n}]$. Now

$$\mathbb{P}[Y_n < \mu_n]^{\mathcal{X}^n} = 1 - \frac{1}{n} \tag{109}$$

so that

$$\mathbb{P}[Y_n < n - t] = 1 - (1 - \frac{1}{n})^{\frac{1}{\mathcal{X}^n}} \tag{110}$$

and we can bound this probability above as follows

$$\mathbb{P}[Y_n < n - t] = \int_{\Delta^*(n-t)} d\underline{u} \leq \int_{\Delta(n-t)} d\underline{u} = \frac{(n-t)^n}{n!} \tag{111}$$

where

$$\Delta^*(n-t) = \{\underline{u} \in \mathbb{R}^n : \sum_{i=1}^{n} u_i = n - t : 0 \leq u_i \leq 1\} \tag{112}$$

and

$$\Delta(n-t) = \{\underline{u} \in \mathbb{R}^n : \sum_{i=1}^{n} u_i = n - t : 0 \leq u_i\} \tag{113}$$

Therefore

$$\lambda \leq \lim_{n\to\infty} \frac{\mu_n}{n} = \lim_{n\to\infty} 1 - \frac{(n![1-(1-\frac{1}{n})^{\frac{1}{\mathcal{X}^n}}])^{\frac{1}{n}}}{n} \tag{114}$$

Finally using $\lim_{n\to\infty}(1-\frac{x}{n})^n = e^x$ and Stirling's formula we have

$$\lambda \leq \lim_{n\to\infty} 1 - \frac{1}{e\mathcal{X}}(\frac{\sqrt{2\pi n}}{n})^{\frac{1}{n}} = 1 - \frac{1}{e\mathcal{X}} \tag{115}$$

5 Conclusion

Using our main theorem which relates the Max-plus exponent to some more tractable bounds in terms of the Markov chain on G and a novel extreme value expectation we have demonstrated an explicit link between the graph structure and the Max-plus exponent for our three examples. Increasing the number of connections between vertices will increase $\mathcal{X}$ which we have shown increases the exponent. This is due to there being more paths of a fixed length through a more connected graph and therefore a greater number of possible paths to take the maximum over.

We have also seen how the tail characteristics of the edge weight distributions can affect the exponent, the heavily tailed exponential giving rise to far larger Max-plus exponents than the Gaussian case.

The bounds give precise information about the asymptotic behavior of large systems. There are many example systems in which this could be a useful result especially if the theorem could be extended to heterogenous systems where the edge weights perhaps have distributions from the same family just with different parameters or a case in which the edge weights are draw from one of a small fixed number of distributions.

In further work I hope to apply these results/techniques to physically realistic models from distributed computing and biology.

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] C.A. Brackley, D. Broomhead, M. C. Romano and M. Thiel. (2011). A Max-Plus Model of Ribosome Dynamics During mRNA Translation. *In Review.*

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

[9] Robert J. Adler. (1990). An introduction to continuity, extrema, and related topics for general Gaussian processes. *IMS Lecture Notes - Monograph Series.* Vol 12

[10] Yash Mittal and Donald Ylvisaker. (1967). Strong Laws for the Maxima of Stationary Gaussian Processes. *Annals of probability* Volume 4, Number 3 (1976), 357-371.