

Max-Plus Singular Values

Hook, James

2014

MIMS EPrint: **2014.7**

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

MAX-PLUS SINGULAR VALUES

James Hook*

March 10, 2014

Abstract

In this paper we prove a new characterization of the max-plus singular values of a max-plus matrix, as the max-plus eigenvalues of an associated max-plus matrix pencil. This new characterization allows us to compute max-plus singular values quickly and accurately. As well as capturing the asymptotic behavior of the singular values of classical matrices whose entries are exponentially parameterized we show experimentally that max-plus singular values give order of magnitude approximations to the classical singular values of parameter independent classical matrices.

We also discuss Hungarian scaling, which is a diagonal scaling strategy for preprocessing classical linear systems. We show that Hungarian scaling can dramatically reduce the d -norm condition number and that this action can be explained using our new theory for max-plus singular values.

Introduction

Max-plus algebra concerns the semiring $\mathbb{R}_{\max} = \mathbb{R} \cup \{-\infty\}$ with addition and multiplication operations

$$a \oplus b = \max\{a, b\}, \quad a \otimes b = a + b, \quad a, b \in \mathbb{R}_{\max}.$$

More generally tropical algebra is the study of any semiring in which the addition operation is max or min, for example max-times, min-max and max-average.

Max-plus algebra naturally describes certain dynamical systems and operations research problems [1]. Max-plus algebra can also be used to approximate or bound the solutions to certain classical algebra problems, which is the topic of this paper.

An $n \times m$ max-plus matrix $G \in \mathbb{R}_{\max}^{n \times m}$ is simply an $n \times m$ array of entries from \mathbb{R}_{\max} . The max-plus Singular Value Decomposition (SVD) of a max-plus matrix was introduced by De Schutter and De Moor in [2]. They work in the max-plus algebra of pairs, which is roughly max-plus algebra with subtraction. In this setting equalities are replaced with weaker relations, which they call balances. Their main result is proving the existence of a max-plus SVD which looks exactly as the classical SVD but with max replacing sum, sum replacing times and balancing replacing equality. The max-plus SVD is useful for analyzing certain max-plus linear systems. De Schutter and De Moore also use the decomposition to introduce a definition of the rank of a max-plus matrix, which

*School of Mathematics, The University of Manchester, Manchester, M13 9PL, UK (james.hook@manchester.ac.uk). This work was supported by Engineering and Physical Sciences Research Council (EPSRC) grant EP/I005293 'Nonlinear Eigenvalue Problems: Theory and Numerics'.

is useful in max-plus linear signal processing problems. However they do not provide a polynomial time algorithm for computing the max-plus SVD of a max-plus matrix $G \in \mathbb{R}_{\max}^{n \times n}$ and the method that they describe requires one to solve a difficult classical algebra problem, namely to find the asymptotic behavior of the analytic SVD of a matrix whose entries are exponentials with exponents given by the entries of the max-plus matrix, $A(t) = (a_{ij}(t))$ with

$$a_{ij}(t) = b_{ij} \exp(g_{ij}t),$$

for generic $B = (b_{ij}) \in \mathbb{C}^{n \times n}$. In this paper we take the opposite approach! We want to use the max-plus singular values of $G = (g_{ij})$ to tell us something about the classical singular values of A , rather than the other way around. As well as enabling us to compute the asymptotics of the singular values of a matrix whose entries are exponentials, we show that max-plus singular values can be used to approximate the log of classical singular values of a fixed matrix $M \in \mathbb{C}^{n \times m}$. The theory we develop also explains the action of Hungarian scaling, which is a diagonal-scaling/balancing technique for classical linear systems.

Using our new characterization, the max-plus singular values of an $n \times m$ max-plus matrix G can be computed in a numerically stable way with $O(k\tau)$ complexity, where $k = \min\{n, m\}$ and τ is the number of non-zero elements in the matrix. We perform these computations using our own algorithm, which is loosely based on the max-plus eigensolver algorithm of Gassner and Klinz [3]. In this paper we focus on computing the max-plus singular values rather than the max-plus SVD decomposition, but it is possible to use our results to compute the singular vectors in polynomial time using our matrix pencil description of the problem, the max-plus eigensolver algorithm and through repeated use of the max-plus algebra of pairs Cramer's rule [4, Chapter 3.5].

This paper is organized as follows. In Section 1 we introduce all of the important definitions and recall some background results. In Section 2 we prove that the max-plus singular values of a max-plus matrix can be computed as the max-plus eigenvalues of an associated max-plus pencil. In section 3 we discuss valuation of classical matrices, which is a way of transforming a classical matrix into a max-plus one - so that the valuation of a classical matrix is amenable to max-plus techniques. In Section 4 we use our new theory to explain the action of Hungarian scaling, which can reduce the condition number of badly conditioned matrices. Finally in Section 5 we illustrate our theory with some examples, including one from a "real life" fluid dynamics problem.

1 Background

For $q_0, \dots, q_d \in \mathbb{R}_{\max}$, let

$$q(z) = \bigoplus_{k=0}^d z^{\otimes k} \otimes q_k = \max\{kz + q_k : k = 0, 1, \dots, d\},$$

be a *max-plus polynomial*. A max-plus polynomial is a convex, piecewise-affine function whose *max-plus roots* are the points at which it is non-differentiable. The multiplicity of a root is the change in derivative at that root. Equivalently q 's roots are the values at which the maximum expression for q is attained more than once and the multiplicity of a root is equal to the maximum difference in index between two terms that attain this maximum.

We also include $-\infty$ as a root with multiplicity k , whenever q_0, \dots, q_{k-1} are all equal to $-\infty$.

Theorem 1.1 (Ostrowski [5]) Let $p(z) = \sum_{k=0}^d p_k z^k \in \mathbb{C}[z]$ be a classical polynomial with roots $|z_1| \geq \dots \geq |z_d|$ and define $q(z)$ to be the max-plus polynomial

$$q(z) = \bigoplus_{k=0}^d z^{\otimes k} \otimes \log |p_k|,$$

with max-plus roots $r_1 \geq \dots \geq r_d$. Then

$$\frac{1}{2} \exp(r_1) < |z_1| \leq d \exp(r_1),$$

$$\left[1 - \left(\frac{1}{2}\right)^{\frac{1}{k}}\right] \exp(r_k) \leq |z_k| \leq \exp(r_k) \left[1 - \left(\frac{1}{2}\right)^{\frac{1}{d-k+1}}\right]^{-1}, \text{ for } k = 2, \dots, d-1,$$

$$\frac{1}{d} \exp(r_d) \leq |z_d| < 2 \exp(r_d).$$

These sharpness of these bounds can be improved in cases where there the max-plus roots are well separated from each other [6]. The max-plus roots of a max-plus polynomial can be computed exactly in linear time using the Graham scan algorithm [7] and the approximation $\log |z_i| \approx r_i$ can then be used as an initial guess for iterative polynomial root finders such as the Aberth Ehrlich method [8]. The max-plus roots of a max-plus polynomial can also be used to compute the exact asymptotic growth rates of the classical roots of a parameterized classical polynomial, which we explain after this supporting result.

Lemma 1.2 Let $\mathcal{R} : \mathbb{R}_{\max}^{d+1} \mapsto \mathbb{R}_{\max}^d$ be the function that maps the coefficients q_0, \dots, q_d of a max-plus polynomial $q(z) = q_0 \oplus \dots \oplus q_d \otimes z^{\otimes d}$ to its roots r_1, \dots, r_d . Then, \mathcal{R} is multiplicatively homogeneous and uniformly continuous.

Proof Let $F : \mathbb{R}_{\max}^{d+1} \mapsto \mathbb{R}_{\max}^{d+1}$ be the function that takes the coefficients (q_0, \dots, q_d) of the max-plus polynomial $q(z)$ to the coefficients $(\hat{q}_0, \dots, \hat{q}_d)$ of the max-plus polynomial $\hat{q}(z)$, with

$$\hat{q}_i = \max\{a \in \mathbb{R}_{\max} : \underbrace{q(z) \oplus a \otimes z^{\otimes i}} = q(z)\},$$

where in standard notation the underbraced term is given by

$$q(z) \oplus a \otimes z^{\otimes i} = \max\{q(z), a + iz\}.$$

Since $q(z)$ is convex $\hat{q}_i = \min\{q(z) - iz : z \in \mathbb{R}_{\max}\}$, and by construction $\hat{q}(z) = q(z)$ for all z , also for all i there exists z with $\hat{q}(z) = \hat{q}_i + iz$. Thus

$$\mathcal{R}(q_0, \dots, q_d) = \hat{\mathcal{R}} \circ F(q_0, \dots, q_d),$$

where $\hat{\mathcal{R}}$ is \mathcal{R} restricted to the image of F , and is given by

$$\hat{\mathcal{R}}(\hat{q}_0, \dots, \hat{q}_d) = (\hat{q}_d - \hat{q}_{d-1}, \dots, \hat{q}_2 - \hat{q}_1),$$

which is clearly homogeneous and uniformly continuous. It should also be clear that F is homogeneous, so all that remains is to show that F is uniformly continuous. Let $\tilde{q} = q_i + \Delta_i$ be a perturbation of q , with $|\Delta_i| \leq \epsilon$ for all i , then

$$|F(\tilde{q}_0, \dots, \tilde{q}_d) - F(q_0, \dots, q_d)|_{\infty} = \max_i |\min\{\tilde{q}(z) - iz\} - \min\{q(z) - iz\}| \leq \epsilon. \quad \square$$

Corollary 1.3 Let $p^t(z) = \sum_{k=0}^d z^k p_k(t)$ be a parameterized polynomial with roots $|z_1(t)| \geq \dots \geq |z_d(t)|$ then for each i the limit

$$r_i = \lim_{t \rightarrow \infty} \frac{1}{t} \log |z_i(t)|,$$

exists and is equal to the i th max-plus root of the max-plus polynomial

$$q(z) = \bigoplus_{k=0}^d z^{\otimes k} q_k,$$

where

$$q_k = \lim_{t \rightarrow \infty} \frac{1}{t} \log |p_k(t)|.$$

Proof Let q^t be the parameterized max-plus polynomial with

$$q^t(z) = \bigoplus_{k=0}^d z^{\otimes k} \log |p_k(t)|.$$

The roots $r_1(t), \dots, r_d(t)$ of q^t are given by

$$(r_k(t))_{k=1}^d = \mathcal{R}[(\log |p_k(t)|)_{k=0}^d],$$

so by homogeneity

$$\left(\frac{r_k(t)}{t}\right)_{k=1}^d = \mathcal{R}\left[\left(\frac{\log |p_k(t)|}{t}\right)_{k=0}^d\right],$$

and by uniform continuity

$$\lim_{t \rightarrow \infty} \left(\frac{r_k(t)}{t}\right)_{k=1}^d = \mathcal{R}\left[\lim_{t \rightarrow \infty} \left(\frac{\log |p_k(t)|}{t}\right)_{k=0}^d\right].$$

Finally each sandwich inequality in Theorem 1.1 is of the form

$$c_k r_k(t) \leq \log |z_k(t)| \leq C_k r_k(t),$$

with finite non-zero $c_k, C_k \in \mathbb{R}$, so that

$$\lim_{t \rightarrow \infty} \frac{\log |z_k(t)|}{t} = \lim_{t \rightarrow \infty} \frac{r_k(t)}{t} = r_k,$$

where r_1, \dots, r_d are the roots of q as in the statement of the Corollary. \square

We can also use max-plus polynomial roots to define the max-plus eigenvalues of a max-plus matrix. Let $G \in \mathbb{R}_{\max}^{n \times n}$ be a max-plus matrix. The *max-plus eigenvalues* $\mu_1 \geq \dots \geq \mu_n$ of G are the max-plus roots of the *max-plus characteristic polynomial*

$$\chi_G(z) = \text{perm}(G \oplus z \otimes I),$$

where

$$\text{perm}(M) = \max_{\pi \in P_n} \sum_{k=1}^n m_{\pi(k), k},$$

is called the *max-plus permanent*, where P_n is the set of all permutation on $\{1, 2, \dots, n\}$, and I is the $n \times n$ max-plus identity matrix with zeros on the diagonal and $-\infty$ off the diagonal.

Proposition 1.4 Let $\mathcal{E} : \mathbb{R}_{\max}^{n \times n} \mapsto \mathbb{R}_{\max}^n$ be the function that maps the max-plus matrix $G = (g_{ij})$ to its max-plus eigenvalues μ_1, \dots, μ_n . Then, \mathcal{E} is uniformly continuous.

Proof Let $\chi_G(z)$ be G 's max-plus characteristic polynomial. The coefficients of $\chi_G(z)$ are maximums of sums of entries in G and as such are uniformly continuous in G . Since G 's eigenvalues are the roots of the characteristic polynomial, uniform continuity of \mathcal{E} follows from Lemma 1.2. \square

Remark The *uniform* continuity of max-plus eigenvalues with respect to the matrix entries means that max-plus eigenvalues are always very well conditioned i.e. not sensitive to small perturbations to the matrix.

The max-plus eigenvalues of an $n \times n$ max-plus matrix can be computed with cost $O(n\tau)$, where τ is the number of finite coefficients in the matrix (the $-\infty$ entries play the role of zero in max-plus algebra since $a \oplus -\infty = a$ for all $a \in \mathbb{R}_{\max}$). Just like the max-plus roots of a max-plus polynomial the max-plus eigenvalues of a max-plus matrix tell us about the asymptotic behavior of an associated classical system.

Theorem 1.5 (Akian, Gaubert, Bapat [9]) Let $G = (g_{ij}) \in \mathbb{R}_{\max}^{n \times n}$ be a max-plus matrix and let $B = (b_{ij}) \in \mathbb{C}^{n \times n}$ be a complex matrix. Now let $A(t) = (a_{ij}(t))$ be the parameterized matrix with

$$a_{ij}(t) = b_{ij} \exp(g_{ij}t),$$

where by convention $\exp(-\infty) = 0$. Let $\lambda_1(t), \dots, \lambda_n(t)$ be the analytic eigenvalues of A , with $\lambda_{n-k+1}(t), \dots, \lambda_n(t) \equiv 0$. For all G and generic B , including generic symmetric B , and for $i = 1, \dots, n - k$

$$\mu_i = \lim_{t \rightarrow \infty} \frac{\log |\lambda_i(t)|}{t},$$

exists and is independent of B .

Moreover these limits are equal to G 's finite max-plus eigenvalues, while G 's full spectrum of max-plus eigenvalues is given by μ_1, \dots, μ_n , with μ_1, \dots, μ_{n-k} defined as above and $\mu_{n-k+1}, \dots, \mu_n = -\infty$.

In both Corollary 1.3 and Theorem 1.5 the asymptotic behavior of the solution to a classical problem whose coefficients/entries are exponentials is shown to be determined by the solution to an associated max-plus problem. These associated max-plus problems are called *valuations*. There are many different possible valuations of classical algebra problem; we examine a different method for valuating parameter independent classical matrices in Section 3.

A major drawback to the theory of max-plus eigenvalues is that the max-plus characteristic polynomial of the valuation of a classical matrix is not necessarily equal to the valuation of the classical characteristic polynomial of that matrix. So that while there is a very strong relationship between a classical scalar polynomial and its valuation, the relationship between a classical matrix and its valuation is not always so strong and there are degenerate cases where the max-plus valuation tells us very little about the original system. This is why we needed the genericity conditions on B in Theorem 1.5 and why, when we consider valuation of parameter independent classical matrices in Sections 3 and 4, we cannot apply the bounds of Theorem 1 to show that the classical eigenvalues of a matrix are always close to the exponentials of the max-plus eigenvalues of its valuation.

As we will illustrate in the example below, this drawback also means that we are not able to define the max-plus singular values of a max-plus matrix G in terms of the max-plus eigenvalues of $G \otimes G^T$.

Unlike max-plus eigenvalues, which are defined in a max-plus way, and then turn out to give us information about an associated classical algebra system; the established definition of the max-plus singular values of a max-plus matrix are given directly in terms of an associated classical algebra system. Theorem 2.1, which is the main result of this paper, works backwards to give a max-plus characterization of the max-plus singular values, which allows us to compute them using max-plus techniques/algorithms.

Theorem 1.6 (De Schutter, De Moor [2]) *Let $G = (g_{ij}) \in \mathbb{R}_{\max}^{n \times n}$ be a max-plus matrix and let $B = (b_{ij}) \in \mathbb{C}^{n \times n}$ be a complex matrix. Now let $A(t) = (a_{ij}(t))$ be the parameterized matrix with*

$$a(t)_{i,j} = b_{i,j} \exp(g_{ij}t),$$

where by convention $\exp(-\infty) = 0$. Let $A(t) = U(t)\Sigma(t)V(t)$ be the analytic SVD of A , with $\Sigma = \text{diag}(\sigma_1(t), \dots, \sigma_n(t))$, and suppose that $\sigma_{n-k+1}(t), \dots, \sigma_n(t) \equiv 0$. For all G , generic B and $i = 1, \dots, n - k$

$$s_i = \lim_{t \rightarrow \infty} \frac{\log \sigma_i(t)}{t},$$

exists and is independent of the choice of B .

The max-plus singular values of G are defined by s_1, \dots, s_n , with s_1, \dots, s_{n-k} defined as above and $s_{n-k+1}, \dots, s_n = -\infty$.

Like the max-plus eigenvalues, the max-plus singular values of G give the asymptotics of a related classical algebra system. However, even though Theorem 1.5 is valid for generic symmetric matrices, it is not valid for matrices of the form AA^T for generic A - as we demonstrate in the following example.

Example 1. Consider

$$A(t) = \begin{bmatrix} \alpha \exp(at) & \beta \exp(bt) \\ \gamma \exp(ct) & \delta \exp(dt) \end{bmatrix},$$

with $\alpha, \beta, \gamma, \delta \in \mathbb{C}$ and $a, b, c, d \in \mathbb{R}_{\max}$ with $b > a > d > c$ and $a + d > b + c$. We can compute the singular values $\sigma_1(t), \sigma_2(t)$ of $A(t)$ as the square roots of the eigenvalues $\lambda_1(t), \lambda_2(t)$ of

$$A(t)A^T(t) = \begin{bmatrix} \alpha^2 \exp(2at) + \beta^2 \exp(2bt) & \alpha\gamma \exp(at + ct) + \beta\delta \exp(bt + dt) \\ \alpha\gamma \exp(at + ct) + \beta\delta \exp(bt + dt) & \gamma^2 \exp(2ct) + \delta^2 \exp(2dt) \end{bmatrix},$$

which has characteristic polynomial of the form

$$\chi_{AA^T}(z) = z^2 + p_1(t)z + p_2(t),$$

with

$$p_1(t) = \alpha^2 \exp(2at) + \beta^2 \exp(2bt) + \gamma^2 \exp(2ct) + \delta^2 \exp(2dt),$$

which for $\beta \neq 0$, has asymptotic growth

$$\lim_{t \rightarrow \infty} \frac{\log |p_1(t)|}{t} = 2b,$$

and

$$\begin{aligned} p_2(t) &= [\alpha^2 \exp(2at) + \beta^2 \exp(2bt)][\gamma^2 \exp(2ct) + \delta^2 \exp(2dt)] \\ &\quad - [\alpha\gamma \exp(at + ct) + \beta\delta \exp(bt + dt)][\alpha\gamma \exp(at + ct) + \beta\delta \exp(bt + dt)] \\ &= F(t) + \alpha^2\delta^2 \exp(2at + 2dt) + \beta^2\gamma^2 \exp(2bt + 2ct) - 2\alpha\beta\gamma\delta \exp(at + bt + ct + dt), \end{aligned}$$

where

$$F(t) = \alpha^2\gamma^2 \exp(2at + 2ct) - \alpha^2\gamma^2 \exp(2at + 2ct) + \beta^2\delta^2 \exp(2bt + 2dt) - \beta^2\delta^2 \exp(2bt + 2dt) \equiv 0.$$

For $\alpha\delta \neq 0$, $p_2(t)$ has asymptotic growth

$$\lim_{t \rightarrow \infty} \frac{\log |p_2(t)|}{t} = 2a + 2d.$$

Therefore, for $\alpha, \delta, \beta \neq 0$, i.e. generic $\alpha, \beta, \gamma, \delta \in \mathbb{C}$, by Corollary 1.3 we have

$$\lim_{t \rightarrow \infty} \frac{\log |\lambda_i(t)|}{t} = r_i,$$

where r_1, r_2 are the max-plus roots of the max-plus polynomial

$$q(z) = z^{\otimes 2} \oplus 2b \otimes z \oplus 2a + 2d = \max\{2z, z + 2b, 2a + 2d\}.$$

These roots are given by $r_1 = 2b$ and $r_2 = 2a + 2d - 2b$. Finally

$$\lim_{t \rightarrow \infty} \frac{\log \sigma_i(t)}{t} = \lim_{t \rightarrow \infty} \frac{\log \sqrt{\lambda_i(t)}}{t} = \frac{r_i}{2} = \sigma_i,$$

where these limits are given by $s_1 = b$ and $s_2 = a + d - b$. These exponents are therefore the max-plus singular values of the max-plus matrix

$$G = \begin{bmatrix} a & b \\ c & d \end{bmatrix}.$$

However the max-plus eigenvalues of

$$G \otimes G^T = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \otimes \begin{bmatrix} a & c \\ b & d \end{bmatrix} = \begin{bmatrix} a \otimes a \oplus b \otimes b & a \otimes c \oplus b \otimes d \\ a \otimes c \oplus b \otimes d & c \otimes c \oplus d \otimes d \end{bmatrix} = \begin{bmatrix} 2b & b + d \\ b + d & 2d \end{bmatrix},$$

do not agree with this calculation. The max-plus eigenvalues μ_1, μ_2 of $G \otimes G^T$ are the max-plus roots of the max-plus characteristic polynomial

$$\begin{aligned} \chi_{G \otimes G^T}(z) &= \text{perm} \left(\begin{bmatrix} 2b \oplus z & b + d \\ b + d & 2d \oplus z \end{bmatrix} \right) \\ &= (2b \oplus z) \otimes (2d \oplus z) \oplus (b + d) \otimes (b + d) \\ &= z^{\otimes 2} \oplus 2b \otimes z \oplus 2b + 2d = \max\{2z, z + 2b, 2b + 2d\}. \end{aligned}$$

Therefore the eigenvalues of $G \otimes G^T$ are $\mu_1 = 2b$ and $\mu_2 = 2d$. This would suggest that the max-plus singular values of G should be $s_1 = b$ and $s_2 = d$, which does not agree with the previous calculation.

The second coefficient in the max-plus characteristic polynomial of $G \otimes G^T$ is not equal to the exponent of the highest order term in the corresponding coefficient of the classical algebra characteristic polynomial of $A \otimes A^T$ and because of this the two different the calculations for the singular values of G do not agree. This situation cannot be avoided as generically matrices of the form AA^T contain different permutations with the same weight but opposite signature. These terms cancel out in the classical algebra characteristic polynomial but not in the max-plus one. This is why the max-plus singular values can not be calculated from the max-plus eigenvalues of $G \otimes G^T$.

2 Max-Plus Singular Values

In this section we introduce our new max-plus characterization of max-plus singular values. We first need to define the max-plus eigenvalues of a max-plus pencil.

Let $G, H \in \mathbb{R}_{\max}^{n \times n}$ be max-plus matrices. The *max-plus eigenvalues* of the *max plus pencil*

$$Q(z) = G \oplus z \otimes H,$$

are the max-plus roots of the max-plus characteristic polynomial

$$\chi_Q(z) = \text{perm}(G \oplus z \otimes H).$$

Just like the max-plus eigenvalues of a matrix, the max-plus eigenvalues of a pencil or more generally a max-plus matrix polynomial can be shown to capture the asymptotic growth rates of the classical eigenvalues of an associated classical algebra system [10]. This is our main result.

Theorem 2.1 (Max-Plus Singular Values) *Let $G \in \mathbb{R}_{\max}^{n \times n}$ be a max-plus matrix. The max-plus singular values of G are given by the max-plus eigenvalues of the max-plus pencil,*

$$Q(z) = G \oplus z \otimes \underline{0},$$

where $\underline{0}$ is an $n \times n$ matrix of zeros.

Example 2 Before the proof of Theorem 2.1. we return to Example 1. Using our new characterization we calculate the max-plus singular values of G as the max-plus roots of the max-plus characteristic polynomial

$$\begin{aligned} \chi_Q(z) &= \text{perm} \left(\left[\begin{array}{cc} a & b \\ c & d \end{array} \right] \oplus z \otimes \left[\begin{array}{cc} 0 & 0 \\ 0 & 0 \end{array} \right] \right) = \text{perm} \left(\left[\begin{array}{cc} a \oplus z & b \oplus z \\ c \oplus z & d \oplus z \end{array} \right] \right) \\ &= (a \oplus z) \otimes (d \oplus z) \oplus (b \oplus z) \otimes (c \oplus z) \\ &= z^{\otimes 2} \oplus b \otimes z \oplus a + d = \max\{2z, z + b, a + d\}, \end{aligned}$$

which gives $s_1 = b$ and $s_2 = a + d - b$, which agrees with the calculation in Example 1.

Proof of Theorem 2.1 We start by finding the leading order terms in the coefficients of $A(t)A(t)^T$'s characteristic polynomial,

$$\det[A(t)A(t)^T - \lambda I] = \sum_{m=0}^n (-\lambda)^m p_{n-m}(t),$$

where

$$p_m(t) = \sum_{\mathcal{I}} \sum_{\pi \in P_m} \text{sgn}(\pi) \prod_{k=1}^m [A(t)A(t)^T]_{i_{\pi(k)}i_k},$$

where first sum is taken over all m -subsets $\mathcal{I} = \{i_1 < i_2 < \dots < i_m\} \subset \{1, 2, \dots, n\}$. Expanding the $A(t)A(t)^T$ product gives

$$\begin{aligned} p_m(t) &= \sum_{\mathcal{I}} \sum_{\pi} \text{sgn}(\pi) \prod_{k=1}^m \sum_{j=1}^n a_{i_{\pi(k)}j}(t) a_{i_k j}(t), \\ &= \sum_{\mathcal{I}} \sum_{f: \mathcal{I} \rightarrow \{1, 2, \dots, n\}} \sum_{\pi} \text{sgn}(\pi) H(\mathcal{I}, f, \pi), \end{aligned}$$

where the second sum is taken over all functions $f: \mathcal{I} \mapsto \{1, 2, \dots, n\}$ and

$$H(\mathcal{I}, f, \pi) = \prod_{k=1}^m a_{i_{\pi(k)}f(i_k)}(t) a_{i_k f(i_k)}(t).$$

Now suppose that f is such that there exist $k_1 \neq k_2$ with $f(i_{k_1}) = f(i_{k_2})$. Let g be the permutation that just switches k_1 and k_2 , then $H(\mathcal{I}, f, \pi \circ g) = H(\mathcal{I}, f, \pi)$, but $\text{sgn}(\pi \circ g) = -\text{sgn}(\pi)$. Therefore the contribution to p_m from all non-injecting f sums to zero and we need only consider injective f , which can all be expressed as

$$f(i_k) = j_{\varsigma(k)},$$

for some $\mathcal{J} = \{j_1 < j_2 < \dots < j_m\} \subset \{1, 2, \dots, n\}$ and some $\varsigma \in P_m$. Now

$$p_m(t) = \sum_{\mathcal{I}} \sum_{\mathcal{J}} \sum_{\pi} \sum_{\varsigma} \text{sgn}(\pi) W(\mathcal{I}, \mathcal{J}, \pi, \varsigma) W(\mathcal{I}, \mathcal{J}, id, \varsigma),$$

where id is the identity permutation and the *weight* terms are given by

$$W(\mathcal{I}, \mathcal{J}, \pi, \varsigma) = \prod_{k=1}^m a(t)_{i_{\pi(k)}j_{\varsigma(k)}}.$$

Since

$$W(\mathcal{I}, \mathcal{J}, \pi, \varsigma) = W(\mathcal{I}, \mathcal{J}, id, \pi^{-1} \circ \varsigma),$$

we can make the substitution $\eta = \pi^{-1} \circ \varsigma$ to obtain

$$p_m(t) = \sum_{\mathcal{I}} \sum_{\mathcal{J}} \sum_{\eta} \sum_{\varsigma} \text{sgn}(\varsigma \circ \eta^{-1}) W(\mathcal{I}, \mathcal{J}, id, \eta) W(\mathcal{I}, \mathcal{J}, id, \varsigma),$$

Now, either all of these weight terms are identically zero in which case $p_m = 0$ and we set $q_m = -\infty$, or there are some nonzero terms, which we now assume to be the case. Each weight term is of the form

$$W(\mathcal{I}, \mathcal{J}, id, \varsigma) = \left[\prod_{k=1}^m b_{i_k j_{\varsigma(k)}} \right] \exp\left[\left(\sum_{k=1}^m g_{i_k j_{\varsigma(k)}} \right) t \right] =: \Theta(\mathcal{I}, \mathcal{J}, \varsigma) \exp[R(\mathcal{I}, \mathcal{J}, \varsigma)t].$$

So the asymptotic growth rate of a nonzero weight terms is given by

$$\lim_{t \rightarrow \infty} \frac{1}{t} \log |W(\mathcal{I}, \mathcal{J}, \varsigma)| = \sum_{k=1}^m g_{i_k j_{\varsigma(k)}} = R(\mathcal{I}, \mathcal{J}, \varsigma). \quad (1)$$

For generic G all $(\mathcal{I}, \mathcal{J}, \varsigma)$ triples with nonzero weight have distinct growth rates. In particular the triple with the greatest growth rate $(\mathcal{I}^*, \mathcal{J}^*, \varsigma^*)$ is unique. Therefore

$$p_m(t) = W(\mathcal{I}^*, \mathcal{J}^*, \varsigma^*)^2 + \text{lower order terms},$$

so that

$$\lim_{t \rightarrow \infty} \frac{1}{t} \log |p_m| = 2R(\mathcal{I}^*, \mathcal{J}^*, \varsigma^*) = 2q_m,$$

which defines q_m .

Otherwise, for degenerate nongeneric G when this maximum is not unique, we have

$$p_m(t) = c_m \exp(2q_m t) + \text{lower order terms},$$

where

$$c_m = \sum_{\mathcal{I}} \sum_{\mathcal{J}} \sum_{\eta \in P_m: R(\mathcal{I}, \mathcal{J}, \eta) = q_m} \sum_{\varsigma \in P_m: R(\mathcal{I}, \mathcal{J}, \varsigma) = q_m} \text{sgn}(\eta \circ \varsigma^{-1}) \Theta(\mathcal{I}, \mathcal{J}, \eta) \Theta(\mathcal{I}, \mathcal{J}, \varsigma).$$

The coefficient c_m is therefore a polynomial in some of the entries of $B = (b_{ij})$, and each Θ term is the product of a unique m -subset of B 's entries. Choose any $(\mathcal{I}^*, \mathcal{J}^*, \varsigma^*)$ triple that appears in the expression for c_m and let $\mathcal{I}^* = \{i_1, \dots, i_m\}$ and $\mathcal{J}^* = \{j_1, \dots, j_m\}$. Treating c_m as a polynomial in the entries of B we differentiate twice with respect to each B entry from the triple $(\mathcal{I}^*, \mathcal{J}^*, \varsigma^*)$, to obtain

$$\frac{d^{2m} c_m}{d^2 b_{i_1 j_{\varsigma^*(1)}} \cdots d^2 b_{i_m j_{\varsigma^*(m)}}} = 2^m. \quad (2)$$

This is because

$$\frac{d^{2m} \Theta(\mathcal{I}, \mathcal{J}, \eta) \Theta(\mathcal{I}, \mathcal{J}, \varsigma)}{d^2 b_{i_1 j_{\varsigma^*(1)}} \cdots d^2 b_{i_m j_{\varsigma^*(m)}}} = 0,$$

unless each B entry that we differentiate with appears twice in total between the two terms $\Theta(\mathcal{I}, \mathcal{J}, \eta)$ and $\Theta(\mathcal{I}, \mathcal{J}, \varsigma)$. However no such entry can appear twice in the same Θ term because η and ς have to be permutations. Therefore the only term with non-zero derivative in (2) is

$$\Theta[\mathcal{I}^*, \mathcal{J}^*, \varsigma^*] \Theta[\mathcal{I}^*, \mathcal{J}^*, \varsigma^*] = [b_{i_1 j_{\varsigma^*(1)}} \times \cdots \times b_{i_m j_{\varsigma^*(m)}}]^2,$$

and its derivative is 2^m , and the sign preceding it in the sum is positive. Thus since c_m must either be identically zero or only zero for a lower dimensional (non-generic) subset of possible B , we can assume that c_m is non-zero and

$$\lim_{t \rightarrow \infty} \frac{1}{t} \log |p_m| = 2R(\mathcal{I}^*, \mathcal{J}^*, s) = 2q_m.$$

So by Corollary 1.3 the eigenvalues $\lambda_1(t), \dots, \lambda_n(t)$ of $A(t)A(t)^T$, satisfy

$$\lim_{t \rightarrow \infty} \frac{1}{t} \log |\lambda_i(t)| = 2r_i,$$

where $2r_1, \dots, 2r_n$ are the max-plus roots of the max-plus polynomial

$$\tilde{q}(z) = \bigoplus_{k=0}^n z^{\otimes k} \otimes 2q_{n-k}.$$

By the homogeneity result of Lemma 1.2 the singular values of $A(t)$, $\sigma_i(t) = \sqrt{|\lambda_i(t)|}$, $i = 1, \dots, n$ satisfy

$$\lim_{t \rightarrow \infty} \frac{1}{t} \log \sigma_i(t) = r_i,$$

where r_1, \dots, r_n are the max-plus roots of the max-plus polynomial

$$q(z) = \bigoplus_{k=0}^n z^{\otimes k} \otimes q_{n-k},$$

and these roots are equal to the max-plus singular values of G .

All that remains is to show that this polynomial is equal to the max-plus characteristic polynomial of the pencil in the statement of the theorem. The characteristic polynomial is

$$\chi_Q(z) = \text{perm}(G \oplus z \otimes Q) = \bigoplus_{k=0}^d z^{\otimes k} \otimes h_{n-k},$$

where the coefficients h_m are given by

$$h_m = \max_{\pi} \max_{\mathcal{I}} \sum_{k=1}^m g_{i_k, \pi(i_k)},$$

where the second maximum is taken over all m -subsets $\mathcal{I} = \{i_1 < \dots < i_m\} \subset \{1, 2, \dots, n\}$. The action of π restricted to \mathcal{I} is simply an injective function. So like before we can express it using $\pi(i_k) = j_{\varsigma(k)}$ for some $\mathcal{J} = \{j_1 < j_2 < \dots < j_m\} \subset \{1, 2, \dots, n\}$ and some $\varsigma \in P_m$. Thus

$$h_m = \max_{\mathcal{I}} \max_{\mathcal{J}} \max_{\varsigma \in P_m} \sum_{k=1}^m g_{i_k j_{\varsigma(k)}} = \max_{\mathcal{I}} \max_{\mathcal{J}} \max_{\varsigma \in P_m} R(\mathcal{I}, \mathcal{J}, \varsigma) = q_m,$$

where $R(\mathcal{I}, \mathcal{J}, \varsigma)$ is as defined in (1). These coefficients are exactly those of the max-plus polynomial $q(z)$ derived in the first part of the proof. Hence $\chi_Q(z) = q(z)$, and the max-plus eigenvalues of the matrix pencil Q are equal to the max-plus singular values of the matrix G , as required. \square

Theorem 2.2 (Rectangular Case) *Let $G \in \mathbb{R}_{\max}^{n \times m}$ be a rectangular max-plus matrix and let \tilde{G} be the $k \times k$ square max-plus matrix obtained by “padding out” G with $-\infty$ ’s, where $k = \max\{n, m\}$. Let $s_1 \geq \dots \geq s_n$ be max-plus singular values of \tilde{G} . Then,*

- if $n > m$ (tall skinny case), then G ’s max-plus singular values are given by s_1, \dots, s_m ;
- if $n < m$ (short fat case), then G ’s max-plus singular values are given by s_1, \dots, s_n .

Proof This follows from the classical case. Choose generic $B = (b_{ij}) \in \mathbb{R}^{n \times m}$ and set $\tilde{B} = (\tilde{b}_{ij}) \in \mathbb{R}^{k \times k}$ by padding B out with zeros. Consider the $n \times m$ classical parameterized matrix $A(t) = ((a_{ij}(t)))$ with $a_{ij}(t) = b_{ij} \exp(g_{ij}t)$ and also the $k \times k$ matrix $\tilde{A}(t) = ((\tilde{a}_{ij}(t)))$ with $\tilde{a}_{ij}(t) = \tilde{b}_{ij} \exp(\tilde{g}_{ij}t)$. In the tall skinny case $\tilde{A}(t)$ ’s singular values are $A(t)$ ’s as well as $n - m$ zeros. In the short fat case $\tilde{A}(t)$ ’s singular values are equal to those of $A(t)$.

The classical singular values of the matrices A and \tilde{A} therefore match up in this way and since the max-plus singular values of G and \tilde{G} are defined as the asymptotic growth rates of these classical singular values, they must also agree and we are done. \square

Theorem 2.3 (Symmetric Case) *Let $G \in \mathbb{R}_{\max}^{n \times n}$ be symmetric. Then the max-plus singular values and max-plus eigenvalues of G are equal.*

Proof Theorem 1.5 is valid for generic B and generic symmetric B but Theorem 1.6 is only valid for generic B . Therefore we can not prove this result by using the analogy with the classical case, as we would need to reason about the singular values of a parameterized matrix $A(t) = (a_{ij}(t))$ with $a_{ij}(t) = b_{ij} \exp(g_{ij}t)$, for generic symmetric B . Instead we will show directly that the max-plus eigenvalues and singular values are equal in the symmetric case. The validity of Theorem 1.6 for generic symmetric B then follows from this theorem as a corollary.

Recall that G 's max-plus eigenvalues are the roots of

$$\chi_G(z) = \text{perm}(G \oplus z \otimes I),$$

and that G 's max-plus singular values are the roots of

$$\chi_Q(z) = \text{perm}(G \oplus z \otimes Q).$$

Since $Q \geq I$ in every component, we have $\chi_Q(z) \geq \chi_G(z)$ for all z .

For fixed z , it follows from the strong duality principle for linear programming problems that

$$\min\left\{\sum_i u_i + v_i : u, v \in \mathbb{R}^n, [G \oplus z \otimes I]_{ij} - u_i - v_j \leq 0 \text{ for all } i, j\right\} = \text{perm}(G \oplus z \otimes I), \quad (3)$$

which is discussed in more detail in Section 4. Now let $(u, v) \in \mathbb{R}^n$ be optimal solutions to (3) and define

$$a_i = \frac{u_i + v_i}{2}.$$

Then

$$2 \sum_i a_i = \sum_i u_i + v_i = \text{perm}(G \oplus z \otimes I) = \chi_G(z),$$

and

$$[G \oplus z \otimes I]_{ij} - a_i - a_j = \frac{[G \oplus z \otimes I]_{ij} - u_i - v_j}{2} + \frac{[G \oplus z \otimes I]_{ji} - u_j - v_i}{2} \leq 0.$$

In particular, $[G \oplus z \otimes I]_{ii} - 2a_i \leq 0$ so that $a_i \geq z/2$. Therefore

$$[G \oplus z \otimes Q]_{ij} - a_i - a_j = \max\{[G \oplus z \otimes I]_{ij} - u_i - v_j, z - u_i - v_j\} \leq 0.$$

So (a, a) is a feasible solution to

$$\min\left\{\sum_i c_i + d_i : [G \oplus z \otimes Q]_{ij} - c_i - d_j \leq 0 \text{ for all } i, j\right\},$$

and

$$\begin{aligned} \chi_Q(z) &= \text{perm}(G \oplus z \otimes Q) = 2 \sum_i a_i \\ &\geq \min\left\{\sum_i c_i + d_i : [G \oplus z \otimes Q]_{ij} - c_i - d_j \leq 0 \text{ for all } i, j\right\} \\ &= \text{perm}(G \oplus z \otimes Q) = \chi_Q(z). \end{aligned}$$

Thus we have $\chi_Q(z) = \chi_G(z)$ for all z , which means that G 's max-plus eigenvalues and singular values must be equal as they are each the non-differentiability points of the same function. \square

3 Valuation

The results stated so far only tell us about the asymptotics of exponentially parameterized systems. It is obvious that max-plus algebra has a strong relationship with these systems, but we really want to be able to say things about parameter independent classical matrices.

The following is a *heuristic* derivation for such a technique.

Let $M = (m_{ij}) \in \mathbb{C}^{n \times n}$ have singular values $\sigma_1, \dots, \sigma_n$. If M 's entries vary a lot in magnitude then it might resemble one of our previously discussed exponentially parameterized matrices evaluated at a large value of t . Let $A(t) = (a_{ij}(t))$ with

$$a_{ij}(t) = b_{ij} \exp(g_{ij}t),$$

for some $B = (b_{ij}) \in \mathbb{C}^{n \times n}$ and for some $G = (g_{ij}) \in \mathbb{R}_{\max}^{n \times n}$ and suppose that for some large value of t , $t = t^*$ we have $M = A(t^*)$. The singular values $\sigma_1(t), \dots, \sigma_n(t)$ of $A(t)$ satisfy

$$\lim_{t \rightarrow \infty} \frac{\log \sigma_i(t)}{t} = s_i,$$

where s_1, \dots, s_n are the max-plus singular values of the max-plus matrix G . This gives us the approximation

$$\log \sigma_i \approx s_i t^*.$$

The reason that this argument is only a heuristic is that the rate of convergence of the limits in Theorems 1.5 and 1.6 are not independent of the matrix G , and as such there is no absolute scale for determining what values of t^* are actually 'large'. Indeed, we can rescale G (and correspondingly t^*) by any factor we like! In particular we can rescale G and t^* so that $t^* = 1$. This is equivalent to taking B to be the classical matrix with entries

$$b_{ij} = \begin{cases} \frac{m_{ij}}{|m_{ij}|} & \text{for } m_{ij} \neq 0, \\ 1 & \text{otherwise,} \end{cases}$$

and also setting G to be the max-plus matrix with entries

$$g_{ij} = \log |m_{ij}|, \tag{4}$$

which we call $G = \mathcal{V}(M)$, the *valuation* of M . Then the same approximation gives

$$\log \sigma_i \approx s_i,$$

where s_1, \dots, s_n are the max-plus singular values of G defined in (4).

We therefore expect the log of the singular values of M to be approximated by the max-plus singular values of $\mathcal{V}(M) = G$. Likewise we expect the log-of-absolute value of the eigenvalues of M to be approximated by the max-plus eigenvalues of G . Bounding the error in this approximation is equivalent to bounding the rate of convergence in the different limit theorems presented earlier. Upper bounds for the eigenvalues of M based on max-plus eigenvalues of G are derived in [11] but lower bounds are much harder to derive as there are degenerate and close to degenerate systems with very small or zero eigenvalues, which are not detected by the tropical eigenvalues.

Example 2. Consider the symmetric matrix M and its valuation $G = \mathcal{V}(M)$

$$M = \begin{bmatrix} 100 & 100 \\ 100 & 100 + \epsilon \end{bmatrix}, \quad G = \begin{bmatrix} \log(100) & \log(100) \\ \log(100) & \log(100 + \epsilon) \end{bmatrix}.$$

The matrix M has eigenvalues ≈ 200 and ϵ , but G has tropical eigenvalues $\log(100 + \epsilon)$ and $\log(100)$. So that for small ϵ the tropical eigenvalues do not even capture the order of magnitude of the log-of-the-absolute-value of the classical eigenvalues. However away from degenerate cases the approximation works remarkably well, as we will show in the subsequent examples.

4 Hungarian Scaling

The *optimal assignment problem* for $G \in \mathbb{R}_{\max}^{n \times n}$ is to compute

$$\text{perm}(G) = \max_{\pi \in P(n)} \sum_{i=1}^n g_{\pi(i), i},$$

which can be expressed as a Linear Programming Problem (LPP)

$$\text{perm}(G) = \max \left\{ \sum_{i,j=1}^n g_{ij} d_{ij} : D = (d_{ij}) \in \mathbb{R}_+^{n \times n} : \sum_{j=1}^n d_{ij} = \sum_{j=1}^n d_{ji} = 1, \text{ for all } i \right\},$$

this follows from the Birkhoff-von Neumann Theorem, which states that any doubly stochastic matrix can be expressed as a stochastic combination of permutation matrices [12]. In standard form this LPP is equivalent to

$$\max \{ f^T x : x \geq 0 : Cx \leq l \},$$

where $x \in \mathbb{R}^{n^2}$ is a vectorized representation of D and $f \in \mathbb{R}^{n^2}$, $l \in \mathbb{R}^{2n}$ and $C \in \mathbb{R}^{2n \times n^2}$. This LPP has symmetric dual LPP

$$\min \{ l^T y : y \geq 0 : C^T y \geq f \}.$$

With some rearranging the dual LPP can be rewritten in a more convenient form by splitting $y \in \mathbb{R}^{2n}$ into a pair of shorter vectors $u, v \in \mathbb{R}^n$

$$\min \left\{ \sum_{i=1}^n u_i + v_i : u, v \in \mathbb{R}^n : g_{ij} - u_i - v_j \leq 0 \right\}.$$

An optimal solution (u, v) to the dual LPP is called a *Hungarian pair* after the Hungarian algorithm, which is a widely used primal dual algorithm for solving the optimal assignment problem. The Hungarian algorithm is so called because its co-inventors Dénes Kőnig and Jenő Egerváry are both themselves Hungarian.

The strong duality principal (see e.g. [13, Chapter 5]) states that the optimal values of a LPP and its dual LPP are equal. Therefore for any Hungarian pair (u, v) of G we have

$$\sum_{i=1}^n u_i + v_i = \text{perm}(G),$$

which is a fact that we used in the proof of Theorem 2.3.

Hungarian pairs can be used to construct useful diagonal scalings. Suppose that we want to solve $Mx = b$ for some $M = (m_{ij}) \in \mathbb{C}^{n \times n}$. Let $G = (g_{ij}) = \mathcal{V}(M) \in \mathbb{R}_{\max}^{n \times n}$ be M 's valuation with

$g_{ij} = \log |m_{ij}|$. Let π be an optimal assignment for G and let (u, v) be a Hungarian pair for G . Now define $L, R \in \mathbb{R}^{n \times n}$ to be diagonal matrices and $P \in \mathbb{R}^{n \times n}$ to be a permutation matrix with

$$L_{ii} = \exp(-u_i), \quad R_{ii} = \exp(-v_i), \quad P_{ij} = 1 \Leftrightarrow \pi(i) = j, \quad \text{for all } i, j.$$

These matrices can then be applied to M to give

$$H = PLMR,$$

where H has entries of modulus one on the diagonal and modulus less than or equal to one off the diagonal. We call H the *Hungarian scaling and reordering* of M and without the application of P we call it the *Hungarian scaling* of M .

In some cases H is close to being diagonally dominant, as the diagonal contains entries of modulus one and the off diagonal entries are all smaller. In these cases the performance of iterative methods is dramatically improved by applying the Hungarian scaling/reordering as a preprocessing step [14]. More generally the scaling/reordering can be shown to improve the speed of sparse direct linear system solvers through improved pivoting [15]. Hungarian scaling is a technique that is already widely used and is implemented in the HSL-MC64 software package. We have found that the Hungarian scaling also tends to significantly reduce the condition number of a matrix, which is the focus of the remainder of this paper.

The *d-norm condition number* of a matrix $M \in \mathbb{C}^{n \times n}$ is given by

$$\kappa(M) = \frac{\sigma_1}{\sigma_n},$$

where $\sigma_1 \geq \dots \geq \sigma_n$ are the singular values of M . The condition number of a matrix measures the stability of the matrix inverse function at that matrix. Therefore if M has a very large condition number then the solution to $Mx = b$ will be very sensitive to small perturbations, which can lead to major numerical inaccuracies. Techniques, including diagonal scalings, aimed at reducing matrix condition number can therefore significantly improve the accuracy of subsequent numerical linear system solves (see e.g. [16]).

Let $M \in \mathbb{R}^{n \times n}$ be a classical matrix and let $G = \mathcal{V}(M) \in \mathbb{R}_{\max}^{n \times n}$ be its valuation. We have shown that the classical singular values $\sigma_1, \dots, \sigma_n$ of M can be approximated by the max-plus singular values s_1, \dots, s_n of V using

$$\log \sigma_i \approx s_i.$$

We can therefore approximate M 's condition number using

$$\log[\kappa(M)] \approx s_1 - s_n = \hat{\kappa}(G),$$

which we call the *max-plus condition number approximation*. A classical matrix is said to be perfectly conditioned if all of its singular values are equal to one, in analogy we define a *max-plus perfectly conditioned matrix* to be a max-plus matrix whose max-plus singular values are all equal to zero.

Theorem 4.1 (Max-plus conditioning of Hungarian scaled matrices) *Let $M = (m_{ij}) \in \mathbb{C}_{\max}^{n \times n}$ and let $G = (g_{ij}) = \mathcal{V}(M) \in \mathbb{R}_{\max}^{n \times n}$ be its valuation. Now let $H = (h_{ij}) \in \mathbb{C}^{n \times n}$ be a (not necessarily Hungarian) diagonal scaling of M given by*

$$H = LMR,$$

where

$$L_{ii} = \exp(-u_i), \quad R_{ii} = \exp(-v_i),$$

for some $u, v \in \mathbb{R}^n$. Now let $W = (w_{ij}) = \mathcal{V}(H) \in \mathbb{R}_{\max}^{n \times n}$ be the valuation of the scaled matrix H . The matrix W is max-plus perfectly conditioned if and only if (u, v) is a Hungarian pair of G .

Proof The entries of the rescaled matrix H are given by

$$h_{ij} = m_{ij} \exp(-u_i - v_j),$$

so that

$$w_{ij} = \log |m_{ij}| - u_i - v_j = g_{ij} - u_i - v_j.$$

Suppose that W is a max-plus perfectly conditioned matrix, then the characteristic polynomial

$$\chi_Q(z) = \text{perm}(W \oplus z \otimes \underline{Q}),$$

is differentiable everywhere except at $z = 0$, so it must be given by

$$\chi_Q(z) = \begin{cases} nz & \text{for } z \geq 0, \\ 0 & \text{otherwise.} \end{cases}$$

Since $\chi_Q(0) = 0$ we have

$$g_{ij} - u_i - v_j = w_{ij} \leq w_{ij} \oplus 0 \leq 0,$$

for all i, j . Also since

$$\lim_{z \rightarrow -\infty} \chi_Q(z) = \text{perm}(W) = \max_{\pi \in P(n)} \sum_{i=1}^n g_{i, \pi(i)} + u_i + v_{\pi(i)} = 0,$$

we have

$$\sum_i^n u_i + v_j = \text{perm}(G).$$

So that (u, v) is an optimal solution to

$$\min \left\{ \sum_{i=1}^n u_i + v_i : u, v \in \mathbb{R}^n : g_{ij} - u_i - v_j \leq 0 \right\},$$

i.e. (u, v) is a Hungarian pair for G .

Conversely suppose that (u, v) is a Hungarian pair for G . Since $w_{ij} = g_{ij} - u_i - v_j \leq 0$ for all i, j , we have for all $z \geq 0$

$$G \oplus z \otimes \underline{Q} = z \otimes \underline{Q}.$$

Also

$$\lim_{z \rightarrow -\infty} \chi_Q(z) = \text{perm}(W) = \text{perm}(G) - \sum_i^n u_i - v_i = 0.$$

So that $\chi_Q(z)$ is a convex piecewise-affine function with $\lim_{z \rightarrow -\infty} \chi_Q(z) = 0$ and $\chi_Q(z) = nz$ for $z \geq 0$. It is therefore differentiable everywhere except for $z = 0$; equivalently W 's max-plus singular values are all equal to zero and it is tropically perfectly conditioned. \square

Thus Hungarian scalings are optimal at reducing the max-plus condition number of the valuation. By the hypothesis that the max-plus singular values of the valuation approximate the log of the of the classical singular values, we can also expect Hungarian scalings to reduce the order of magnitude of the classical condition number.

5 Examples

Example 3. Classical matrix with exponential components We randomly generate a 10×10 parameterized matrix $A(t) = ((a_{ij}(t)))$ with

$$a_{ij} = b_{ij} \exp(g_{ij}t),$$

where $B = (b_{ij})$ is a matrix of ones and $G = (g_{ij})$ is a randomly generated max-plus matrix sampled using

$$g_{i,j} = \begin{cases} -\infty & \text{with probability 0.5,} \\ \text{sampled from a standard Gaussian} & \text{otherwise.} \end{cases}$$

For $t = 0.1, 0.2, \dots, 10$ we compute the classical singular values $\sigma_1(t), \dots, \sigma_{10}(t)$ of $A(t)$ using MATLAB `svd.m`. We also compute the max-plus singular values s_1, \dots, s_{10} of G using our own MATLAB routine `mpsv.m`. Figure 1 is a plot of

$$\left(\frac{\log \sigma_i(t)}{t} \right)_{i=1}^{10}$$

against t . Notice that each of these quantities converges as t grows and that the different limits are given by the max-plus singular values of G , which are indicated with red lines. The tropical singular values of G have multiplicities¹ 1, 1, 2, 1, 1, 2, 2, these multiplicities also correspond to the number of different classical singular values whose log converges to that limit.

We also apply the Hungarian scaling to the same matrix for each value of t and make the same plot for the rescaled matrices. After Hungarian rescaling the max-plus singular values are all equal to zero, for all t . As before the log of the classical singular values divided by t converges to the max-plus singular values. The condition number of the original matrix A grows exponentially with t but the condition number of the rescaled matrix does not, moreover it can be shown to converge.

Example 4. Sparse unsymmetric matrix from a fluid dynamics problem We use the matrix M of the `steam3.m` problem from the University of Florida sparse matrix collection [17]. The unsymmetric 131×131 matrix M has 536 nonzero entries, which vary a lot in magnitude. Using MATLAB `svd.m` we compute M 's singular values. We then valueate M and compute $G = \mathcal{V}(M)$'s tropical eigenvalues using `mpsv.m`. We also apply the Hungarian scaling to M and compute the classical singular values of the rescaled matrix $H = LMR$. Figure 2 shows the classical and max-plus singular values of the original matrix M and the Hungarian scaled matrix H . Figure 2 also shows the magnitude of the entries in M and H .

Table 1 summarizes the results of this experiment. Notice that the tropical condition number and the log of the classical condition numbers roughly agree for the two matrices, that the condition number is significantly reduced by Hungarian scaling and that for both matrices the max-plus singular values give good order of magnitude approximations of the classical singular values.

Discussion

We have given the max-plus singular values of a max-plus matrix a new characterization as a the eigenvalues of a max-plus matrix pencil. This then enables us to compute max-plus singular values

¹The multiplicity of a max-plus eigenvalue of singular value is defined as its multiplicity as a root of the appropriate characteristic max-plus polynomial.

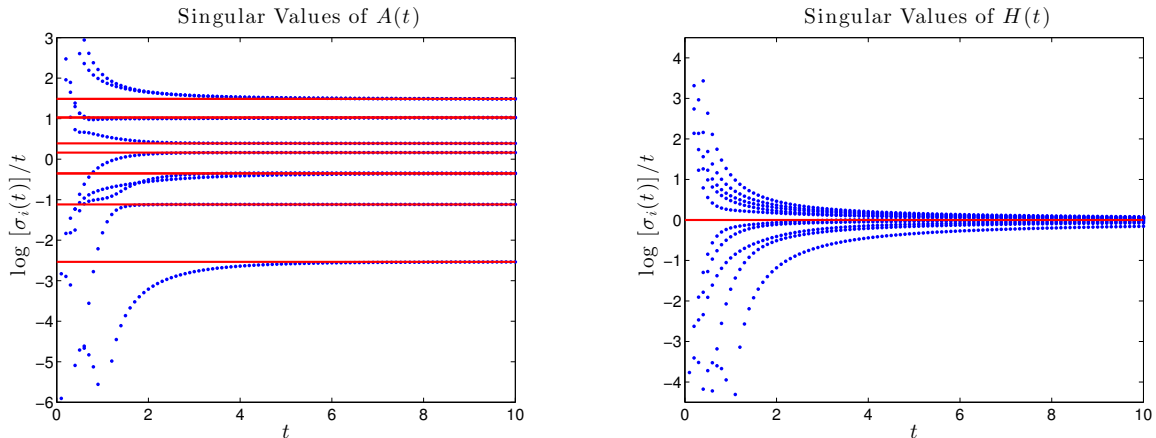


Figure 1: Log of classical singular values divided by t (blue dots) and tropical singular values (red lines). Left, original system; right, Hungarian scaled.

Table 1: Results for Example 4

	original matrix	Hungarian scaled
classical condition number	1.28×10^{15}	15.17
max-plus condition number of valuation	14.52	0
$\max\{\frac{10^{s_i}}{\sigma_i}, \frac{\sigma_i}{10^{s_i}} : i = 1, \dots, 131\}$	2.59	21.27

using fast and accurate network flow algorithms. We have demonstrated experimentally that the distribution over the log scale of the classical singular values of a classical matrix is approximated remarkably well by the max-plus singular values of the valuation. Also, we have shown that Hungarian scaling can dramatically reduce the condition number of a matrix and that this action can be explained with our new theory for max-plus singular values.

Acknowledgement

This work was supported by the Engineering and Physical Sciences Research Council (EPSRC) grant EP/I005293 ‘Nonlinear Eigenvalue Problems: Theory and Numerics’. I am very grateful to Professor Françoise Tisseur for reading previous drafts of this work and for many useful comments and suggestions.

References

- [1] Bernd Heidergott, Geert J. Olsder, and Jacob van der Woude. *Max Plus at Work: Modeling and Analysis of Synchronized Systems: A Course on Max-Plus Algebra and Its Applications*. Princeton University Press, November 2005.

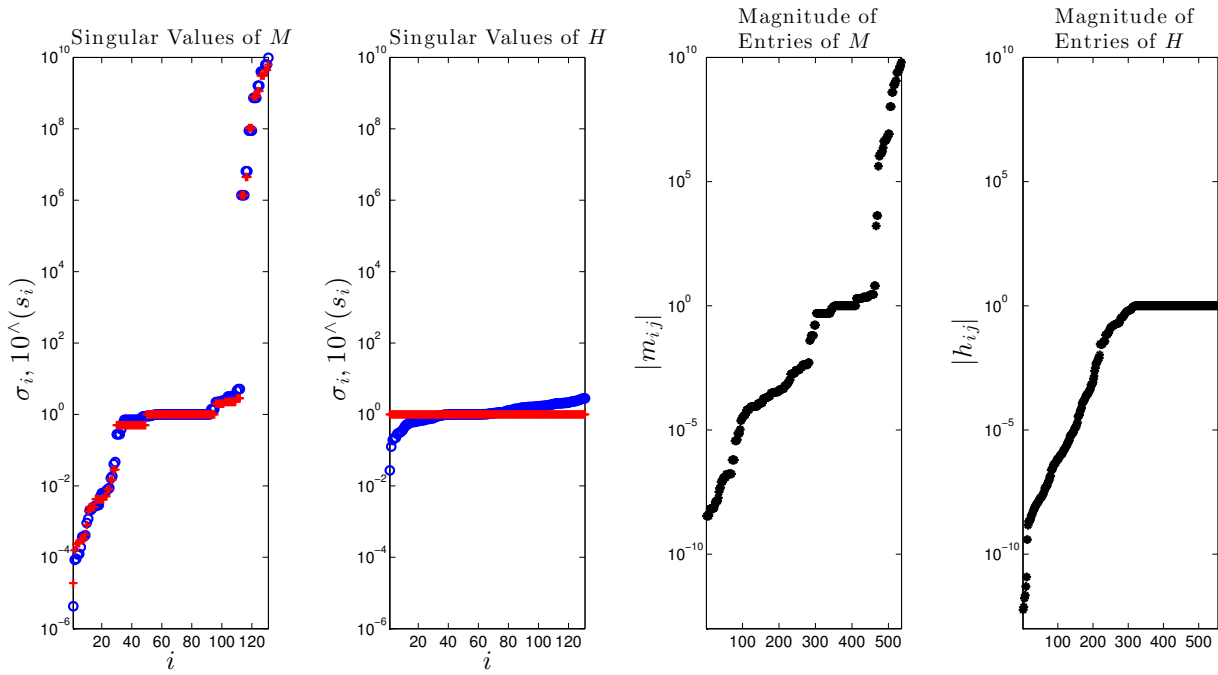


Figure 2: Left and middle left: classical singular values (blue circles) and exponential of max-plus singular values of valuation (red crosses) for M and H respectively. Middle right and right: magnitude of entries (black diamond) of M and H respectively.

- [2] Bart De Schutter and Bart De Moor. The QR decomposition and the singular value decomposition in the symmetrized max-plus algebra revisited. *SIAM Review*, 44(3):pp. 417–454, 2002.
- [3] Elisabeth Gassner and Bettina Klinz. A fast parametric assignment algorithm with applications in max-algebra. *Networks*, 55(2):61–77, 2010.
- [4] G. Olsder F. Baccelli, G. Cohen and J. Quadrat. *Synchronisation and linearity: an algebra for discrete event systems*. Wiley, 1992.
- [5] Alexandre Ostrowski. Recherches sur la methode de Graee et les zeros des polynomes et des series de Laurent. *Acta Math*, 72:99–155, 1940.
- [6] Meisam Sharify. *Scaling Algorithms and Tropical Methods in Numerical Matrix Analysis*. PhD thesis, 2011.
- [7] Ronald L. Graham. An efficient algorithm for determining the convex hull of a finite planar set. *Inf. Process. Lett.*, 1(4):132–133, 1972.
- [8] S. Ghaderpanah and S. Klasa. Polynomial scaling. *SIAM J. Numer. Anal.*, 27(1):117–135, February 1990.
- [9] Stephane Gaubert Marianne Akian, Ravindra Bapat. Generic asymptotics of eigenvalues using min-plus algebra. In *Proceedings of the Satellite Workshop on Max-Plus Algebras, IFAC SSSC01*. Elsevier, 2001.

- [10] Stephane Gaubert Marianne Akian, Ravindra Bapat. Perturbation of eigenvalues of matrix pencils and optimal assignment problem. *arXiv:math/0402438*, 2004.
- [11] M. Akian, S. Gaubert, and A. Marchesini. Tropical bounds for eigenvalues of matrices. Technical Report arXiv:1309.7319, 2013.
- [12] G. Birkhoff. Three observations on linear algebra. *Univ. Nac. Tucuman.Rev.Ser.*, 5:147–151, 1946.
- [13] R. Vanderbel. *Linear Programming*. Springer London, 2014.
- [14] Michele Benzi, John C. Haws, and Miroslav Tuma. Preconditioning highly indefinite and nonsymmetric matrices. *22(4):1333–1353*, 2000.
- [15] J.D. Hogg and J.A. Scott. The effects of scalings on the performance of a sparse symmetric indefinite solver. Technical report, RAL-TR-2008-007, Rutherford Appleton Laboratory, 2008.
- [16] Nicholas J. Higham. *Accuracy and Stability of Numerical Algorithms*. Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, second edition, 2002.
- [17] Timothy A. Davis. The university of florida sparse matrix collection. *NA DIGEST*, 92, 1994.