A Strengthening of Hopf's Inequality

Wendi Han
The University of Hong Kong
email: wendyhan@connect.hku.hk

Guangyue Han
The University of Hong Kong
email: ghan@hku.hk

August 27, 2019

Abstract

Hopf's inequality for positive matrix yields a strengthening of Perron's theorem. We give in this paper a strengthening of Hopf's inequality using a complex extension of the Hilbert metric.

Index terms: Perron's theorem, Perron-Frobenius theorem, Hopf's inequality, positive matrix, Hilbert metric, Birkhoff contraction coefficient.

1 Introduction

Let n be an integer greater than or equal to 2. Let $A = (a_{ij})$ be an $n \times n$ positive matrix, i.e., $a_{ij} > 0$ for all i, j. By Perron's theorem [19], the largest eigenvalue (in modulus) of A, denoted by $\rho(A)$, is unique, real and positive, and therefore, the *spectral ratio* $\kappa(A)$ of A, defined as

$$\kappa(A) \triangleq \max\{|\lambda| : \lambda \text{ is an eigenvalue of } A, \lambda \neq \rho(A)\}/\rho(A),$$

is strictly less than 1. Ostrowski [17] strengthened this result and showed that

$$\kappa(A) \leq \frac{M^2 - m^2}{M^2 + m^2}, \tag{1}$$

where $m = \min_{i,j} a_{ij}$ and $M = \max_{i,j} a_{ij}$. Inspired by Ostrowski's theorem, Hopf [11] further strengthened Perron's theorem and showed that

$$\kappa(A) \leq \frac{M - m}{M + m}. \tag{2}$$

It has been observed [18] that Hopf's strengthening is tight in the sense that there are examples of A for which (2) holds with equality.

Before continuing, we digress a bit to briefly mention other related results (of less relevance though) in the literature. Frobenius [7, 8] generalized Perron's theorem to non-negative matrices, which is popularly known as the Perron-Frobenius theorem. This result is the key pillar of the theory of non-negative matrices, which has a wide range of applications in multiple disciplines; see, e.g., [22, 15, 2, 1, 12]. Accordingly, there are numerous results characterizing the isolation of the largest eigenvalue of non-negative matrices, most of them in the forms of upper bounds on the modulus of the second largest eigenvalue; see, e.g., [20] and the references therein. And it is worthwhile to note that for certain special families of symmetric non-negative matrices (such as adjacency matrices of a regular graph and transition probabilities matrices of a reversible stationary Markov chain), numerous Cheeger-type inequalities, which are in the forms of bounds on the difference between the largest and second largest eigenvalue, have been established; see, e.g. [5, 4, 16, 13] and references therein.

Although it often shows up in the literature, the exact expression as in (2) actually does not appear in [11] and only follows from Theorem 4 therein, stated for more general positive linear operators. As a matter of fact, a careful examination of the proof of Theorem 4 reveals that it yields a bound stronger than (2).

To precisely state this stronger result, we need to introduce some notation and terminologies. Let W denote the standard simplex in the n-dimensional Euclidean space:

$$W = \left\{ w = (w_1, w_2, ..., w_n) \in \mathbb{R}^n : \sum_{i=1}^{n} w_i = 1, \;\; w_i \geq 0 \text{ for all } i \right\}, \tag{3}$$

and let W° denote its interior, consisting of all the positive vectors in W. Let d_H denote the *Hilbert metric* on W°, which is defined [1] by

$$d_H(v, w) \triangleq \max_{i,j} \log \left(\frac{w_i/w_j}{v_i/v_j} \right), \text{ for any two vectors } v, w \in W^\circ. \tag{4}$$

For any positive vector $w = (w_1, w_2, \ldots, w_n) \in \mathbb{R}^n$, we define its normalized version $\mathcal{N}(w)$ as

$$\mathcal{N}(w) = \frac{(w_1, w_2, \ldots, w_n)}{w_1 + w_2 + \cdots + w_n}, \tag{5}$$

which obviously belongs to W°. Apparently, the matrix A induces a mapping $f_A : W^\circ \to W^\circ$, defined by

$$f_A(w) = \mathcal{N}(Aw), \text{ for any vector } w \in W^\circ. \tag{6}$$

It is well known that f_A is a contraction mapping under the Hilbert metric and the contraction coefficient $\tau(A)$, defined by

$$\tau(A) \triangleq \sup_{v \neq w \in W^\circ} \frac{d_H(Av, Aw)}{d_H(v, w)}$$

and often referred to as the *Birkhoff contraction coefficient*, can be explicitly computed as

$$\tau(A) = \frac{1 - \sqrt{\phi(A)}}{1 + \sqrt{\phi(A)}}, \tag{7}$$

where

$$\phi(A) = \min_{i,j,k,l} \frac{a_{ik} a_{jl}}{a_{jk} a_{il}}. \tag{8}$$

[1]The Hilbert metric is often defined on a projective space (see, e.g., [22, 12]), which is equivalent to the definition in this paper up to a usual normalization.

Moreover, we define

$$\theta(A) = \frac{\max_{i,j} a_{i,j}}{\min_{i,j} a_{i,j}}$$

and

$$\eta(A) = \frac{\min_{i,j} a_{i,j}}{\rho(A) - \max\{\min_i \sum_j a_{ij}, \min_j \sum_i a_{ij}\} + \min_{i,j} a_{i,j}}.$$

We are now ready to state the aforementioned stronger result in Part a) of the following theorem and a further strengthening in Part b).

Theorem 1.1. *a) For an $n \times n$ positive matrix A, we have*

$$\kappa(A) \leq \tau(A). \tag{9}$$

b) Furthermore, the inequality (9) can be strengthened as follows:

$$\kappa(A) \leq \max_z \left\{ \frac{1}{1+\phi(A)z} - \frac{1}{1+z} \right\}, \tag{10}$$

where the maximum is over all z with $0 \leq z \leq \frac{(n-1)\theta(A)}{\eta(A)}$.

Remark 1.1. To see that (10) is a strengthening of (9), we note that

$$\tau(A) = \max_z \left\{ \frac{1}{1+\phi(A)z} - \frac{1}{1+z} \right\},$$

where the maximum is over all z with $z \geq 0$.

As mentioned before, the inequality (9) follows from Theorem 4 in [11], which is a contraction result with respect to the Hopf oscillation. Ostrowski [18] modified Birkhoff's argument in [3] and gave an alternative proof of Theorem 1.1, which however still used the Hopf oscillation. In this work, using a complex extension of the Hilbert metric in lieu of the Hopf oscillation, we will give a new proof of (9) and give a further strengthening of it as in (10). Here, we remark that the complex Hilbert metric can be applied elsewhere; more specifically, it has been used [10] to establish the analyticity of entropy rate of hidden Markov chains and specify the corresponding domain of analyticity.

2 A Complex Hilbert Metric

Let $W_{\mathbb{C}} = \{w = (w_1, w_2, \ldots, w_n) \in \mathbb{C}^n : \sum_{i=1}^n w_i = 1\}$ and let $W_{\mathbb{C}}^+ = \{w = (w_1, w_2, \ldots, w_n) \in W_{\mathbb{C}} : \mathcal{R}(w_i/w_j) > 0 \text{ for all } i, j\}$. The following complex extension of the Hilbert metric has been proposed in [10]:

$$d_H(v, w) = \max_{i,j} \left| \log \left(\frac{w_i/w_j}{v_i/v_j} \right) \right|, \text{ for any } v, w \in W_{\mathbb{C}}^+, \tag{11}$$

where $\log(\cdot)$ is taken as the principal branch of the complex $\log(\cdot)$ function. Here we remark that there are other complex extensions of the Hilbert metric; see, e.g., [21, 6].

Our treatment however only uses the extension in (11), which will henceforth be referred to as the complex Hilbert metric. For any $\varepsilon > 0$, we define

$$W^{\circ}_{\mathbb{C}}(\varepsilon) \triangleq \{w = (w_1, w_2, \cdots, w_n) \in W_{\mathbb{C}} : \exists\, v \in W^{\circ} \text{ such that } |w_i - v_i| \leq \varepsilon v_i \text{ for all } i\}. \tag{12}$$

It can be easily verified that for ε small enough, $W^{\circ}_{\mathbb{C}}(\varepsilon) \subset W^{+}_{\mathbb{C}}$ and thereby the complex Hilbert metric is well-defined on $W^{\circ}_{\mathbb{C}}(\varepsilon)$.

Extending the definition in (5), for any complex vector $w = (w_1, w_2, \ldots, w_n)$ with $w_1 + w_2 + \cdots + w_n \neq 0$, we define its normalized version $\mathcal{N}(w)$ as

$$\mathcal{N}(w) = \frac{(w_1, w_2, \ldots, w_n)}{w_1 + w_2 + \cdots + w_n},$$

which obviously belongs to $W_{\mathbb{C}}$. And furthermore, for any $\varepsilon > 0$, extending the definition in (6), we define $f_A : W^{\circ}_{\mathbb{C}}(\varepsilon) \to W^{\circ}_{\mathbb{C}}(\varepsilon)$ by:

$$f_A(w) = \mathcal{N}(Aw), \text{ for any vector } w \in W^{\circ}_{\mathbb{C}}(\varepsilon), \tag{13}$$

which is well-defined if ε is small enough.

The following lemma has been implicitly established in [10]. We outline its proof for completeness and clarity. An interested reader may refer to the proofs of Theorem 2.4 in [10] and relevant lemmas for more technical details.

Lemma 2.1. Consider an $n \times n$ positive square matrix A.

a) For any small enough $\varepsilon > 0$, there exists $0 < \tau_{\varepsilon}(A) < 1$ such that for any $v, w \in W^{\circ}_{\mathbb{C}}(\varepsilon)$,

$$d_H(f_A(v), f_A(w)) \leq \tau_{\varepsilon}(A) d_H(v, w), \tag{14}$$

and moreover, $\tau_{\varepsilon}(A)$ tends to $\tau(A)$ as ε tends to 0.

b) For any small enough $\varepsilon > 0$ and any sequence $\{w_{(k)} = (w_{(k),1}, w_{(k),2}, \ldots, w_{(k),n}) : w_{(k)} \in W^{\circ}_{\mathbb{C}}(\varepsilon)\}$ with $\lim_{k\to\infty} w_{(k)} = v \in W^{\circ}$, there exists $0 < \tau_k(A) < 1$ for any k such that

$$d_H(f_A(v), f_A(w_{(k)})) \leq \tau_k(A) d_H(v, w_{(k)}), \tag{15}$$

and moreover, $\tau_{(k)}(A)$ tends to

$$\max_{i,j} \sum_{l=1}^{n} \max\left\{ \frac{a_{li} v_l}{\sum_{m=1}^{n} a_{mi} v_m} - \frac{a_{lj} v_l}{\sum_{m=1}^{n} a_{mj} v_m}, 0 \right\}$$

as k tends to infinity.

Proof. a) First of all, we note, by the definition in (11), that for any $v, w \in W^{\circ}_{\mathbb{C}}(\varepsilon)$,

$$\frac{d_H(f_A(v), f_A(w))}{d_H(v, w)} = \frac{d_H(\mathcal{N}(Av), \mathcal{N}(Aw))}{d_H(v, w)} = \max_{i,j} |L_{i,j}|, \tag{16}$$

where

$$L_{i,j} = \frac{\log\left(\sum_m a_{im} v_m / \sum_m a_{jm} v_m\right) - \log\left(\sum_m a_{im} w_m / \sum_m a_{jm} w_m\right)}{\max_{k,l} |\log(v_k/w_k) - \log(v_l/w_l)|}. \tag{17}$$

Letting $c_i = \log(v_i/w_i)$ for all i and choosing p, q such that $|c_p - c_q| = \max_{k,l} |c_k - c_l|$, we note that $L_{i,j}$ can be rewritten as

$$L_{i,j} = \frac{\log\left(\sum_m e^{c_m - c_q} a_{im} w_m / \sum_m e^{c_m - c_q} a_{jm} w_m\right) - \log\left(\sum_m a_{im} w_m / \sum_m a_{jm} w_m\right)}{|c_p - c_q|}.$$

An application of the mean value theorem then yields that there exists $\xi_{i,j} \in [0,1]$ such that

$$|L_{i,j}| \leq \sum_l \frac{c_l - c_q}{|c_p - c_q|} \left(\frac{e^{(c_l - c_q)\xi_{i,j}} a_{il} y_l}{\sum_m e^{(c_m - c_q)\xi_{i,j}} a_{im} y_m} - \frac{e^{(c_l - c_q)\xi_{i,j}} a_{jl} y_l}{\sum_m e^{(c_m - c_q)\xi_{i,j}} a_{jm} y_m} \right). \tag{18}$$

By the definition of $W^\circ_{\mathbb{C}}(\varepsilon)$, there exist $v^\circ, w^\circ \in W^\circ$ such that for some constant $C_1 > 0$,

$$|v_k - v_k^\circ| \leq C_1 \varepsilon v_k^\circ, \quad |w_k - w_k^\circ| \leq C_1 \varepsilon w_k^\circ \text{ for all } k.$$

Now, let

$$D_l = \frac{e^{(c_l - c_q)\xi_{i,j}} a_{il} w_l}{\sum_m e^{(c_m - c_q)\xi_{i,j}} a_{im} w_m} - \frac{e^{(c_l - c_q)\xi_{i,j}} a_{jl} w_l}{\sum_m e^{(c_m - c_q)\xi_{i,j}} a_{jm} w_m},$$

and

$$D_l^\circ = \frac{e^{(c_l^\circ - c_q^\circ)\xi_{i,j}} a_{il} w_l^\circ}{\sum_m e^{(c_m^\circ - c_q^\circ)\xi_{i,j}} a_{im} w_m^\circ} - \frac{e^{(c_l^\circ - c_q^\circ)\xi_{i,j}} a_{jl} w_l^\circ}{\sum_m e^{(c_m^\circ - c_q^\circ)\xi_{i,j}} a_{jm} w_m^\circ},$$

where we have, similarly as above, defined $c_i^\circ = \log(v_i^\circ / w_i^\circ)$ for all i. It then follows from the established facts that for some constant $C_2 > 0$,

$$\left| \sum_l \frac{c_l - c_q}{|c_p - c_q|} D_l - \sum_l \frac{c_l - c_q}{|c_p - c_q|} D_l^\circ \right| < C_2 C_1 \varepsilon,$$

and

$$\left| \sum_l \frac{c_l - c_q}{|c_p - c_q|} D_l^\circ \right| \leq \sum_{l=1}^n \max\{D_l^\circ, 0\} \leq \tau(A)$$

that

$$\left| \sum_l \frac{c_l - c_q}{|c_p - c_q|} D_l \right| \leq C_2 C_1 \varepsilon + \tau(A),$$

which immediately implies that

$$\frac{d_H(f_A(v), f_A(w))}{d_H(v, w)} \leq C_2 C_1 \varepsilon + \tau(A).$$

Setting $\tau_\varepsilon(A) = C_2 C_1 \varepsilon + \tau(A)$ and noting that ε can be chosen arbitrarily small, we establish (14) and conclude that $\tau_\varepsilon(A)$ tends to $\tau(A)$ as ε tends to 0.

b) By the proof of a), one can, for each k, choose (a) $w^\circ_{(k),i}$, $i = 1, 2, \ldots, n$, such that $\lim_{k\to\infty} w^\circ_{(k),i} = v$; (b) $0 < \xi_{(k),i,j} < 1$, $i, j = 1, 2, \ldots, n$; (c) $\varepsilon_{(k)} > 0$ such that $\lim_{k\to\infty} \varepsilon_{(k)} = 0$, and set

$$\tau_{(k)}(A) = \max_{i,j} \sum_{l=1}^n \max\left\{ \frac{a_{li} e^{\log(w^\circ_{(k),l}/v_l)\xi_{(k),i,j}} v_l}{\sum_{m=1}^n a_{mi} e^{\log(w^\circ_{(k),m}/v_m)\xi_{(k),i,j}} v_m} - \frac{a_{lj} e^{\log(w^\circ_{(k),l}/v_l)\xi_{(k),i,j}} v_l}{\sum_{m=1}^n a_{mj} e^{\log(w^\circ_{(k),m}/v_m)\xi_{(k),i,j}} v_m}, 0 \right\} + \varepsilon_{(k)},$$

which converges to

$$\max_{i,j} \sum_{l=1}^{n} \max \left\{ \frac{a_{li} v_l}{\sum_{m=1}^{n} a_{mi} v_m} - \frac{a_{lj} v_l}{\sum_{m=1}^{n} a_{mj} v_m}, 0 \right\}$$

as k tends to infinity. $\square$

3 Proof of Theorem 1.1

For a subset S of W°, we generalize the definition in (12) and define

$$S_{\mathbb{C}}(\varepsilon) \triangleq \{ w = (w_1, w_2, \cdots, w_n) \in W_{\mathbb{C}} : \exists\, v \in S \text{ such that } |w_i - v_i| \leq \varepsilon v_i \text{ for all } i \}.$$

We will need the following lemma, which, roughly speaking, asserts the equivalence between the Euclidean metric (denoted by d_E) and the Hilbert metric on a complex neighborhood of a compact subset of W°

Lemma 3.1. For any compact subset S of W°, there exists $\varepsilon_0 > 0$ such that there exist constants $G_1, G_2 > 0$ such that for all $0 < \varepsilon < \varepsilon_0$ and for all $v, w \in S_{\mathbb{C}}(\varepsilon)$,

$$G_1 d_H(v, w) < d_E(v, w) < G_2 d_H(v, w).$$

Proof. The lemma follows from some straightforward arguments underpinned by the mean value theorem and the compactness of S, which are completely parallel to those in the proof of Proposition 2.1 in [9] (a real version of this lemma). $\square$

We are now ready for the proof of Theorem 1.1.

Proof. Consider an $n \times n$ positive square matrix A. Let $x = (x_1, x_2, \ldots, x_n)$ be the eigenvector corresponding to $\rho(A)$. By Perron's theorem, we can choose x to be a positive vector with $x_1 + x_2 + \cdots + x_n = 1$, i.e., $x \in W^\circ$. Let λ be an eigenvalue of A that is different from $\rho(A)$ and let y be a corresponding eigenvector. Here we remark that while $\rho(A)$ and x are real, λ and y can be complex.

Now, consider a compact subset S of W° that contains x. It can be easily verified that for any $\varepsilon > 0$, there exists $n_0 \in \mathbb{N}$ such that for any $n \geq n_0$,

$$\mathcal{N}(A^n(x + y)) = \mathcal{N}(\rho^n(A)x + \lambda^n y) \in S_{\mathbb{C}}(\varepsilon).$$

Henceforth, we let $v = \rho(A)^{n_0} x$ and $w = \lambda^{n_0} y$. For any $m \in \mathbb{N}$, it can be verified that

$$\begin{aligned} d_H(\mathcal{N}(A^m v), \mathcal{N}(A^m(v + w))) &= d_H(\mathcal{N}(\rho(A)^m v), \mathcal{N}(\rho(A)^m v + \lambda^m w)) \\ &= d_H(\mathcal{N}(v), \mathcal{N}(v + \tilde{\lambda}^m w)), \end{aligned}$$

where we have written $\lambda/\rho(A)$ as $\tilde{\lambda}$ for notational simplicity. Now, using the definition of the complex Hilbert metric, we continue

$$
\begin{aligned}
d_H(\mathcal{N}(A^m v), \mathcal{N}(A^m(v+w))) &= \max_{i,j=1,2,\ldots,n} \left| \log \frac{(v_i + \tilde{\lambda}^m w_i)/(v_j + \tilde{\lambda}^m w_j)}{v_i/v_j} \right| \\
&= \max_{i,j=1,2,\ldots,n} \left| \log \frac{1 + \tilde{\lambda}^m (w_i/v_i)}{1 + \tilde{\lambda}^m (w_j/v_j)} \right| \\
&= \max_{i,j=1,2,\ldots,n} \left| \log \left(1 + \frac{\tilde{\lambda}^m (w_i/v_i) - (w_j/v_j)}{1 + \tilde{\lambda}^m (w_j/v_j)} \right) \right| \\
&= \max_{i,j=1,2,\ldots,n} \left| \log \left(1 + \frac{(w_i/v_i) - (w_j/v_j)}{(1/\tilde{\lambda}^m) + (w_j/v_j)} \right) \right| \\
&= \left| \log \left(1 + \frac{(w_{i_0}/v_{i_0}) - (w_{j_0}/v_{j_0})}{(1/\tilde{\lambda}^m) + (w_{j_0}/v_{j_0})} \right) \right|, \qquad (19)
\end{aligned}
$$

where we have assumed i_0, j_0 achieve the maxima in (19). We note that $w_{i_0}/v_{i_0} \neq w_{j_0}/v_{j_0}$, since otherwise it would mean $d_H(\mathcal{N}(A^m v), \mathcal{N}(A^m(v+w))) = 0$ and therefore w would be a scaled version of v, contradicting the fact that λ is different from $\rho(A)$.

It follows from the fact that $0 < \tilde{\lambda} < 1$ that there exists a constant $C_1 > 0$ such that for all m,

$$
d_H(\mathcal{N}(A^m v), \mathcal{N}(A^m(v+w))) = \left| \log \left(1 + \frac{(w_{i_0}/v_{i_0}) - (w_{j_0}/v_{j_0})}{(1/\tilde{\lambda}^m) + (w_{j_0}/v_{j_0})} \right) \right| \geq C_1 \left| \frac{(w_{i_0}/v_{i_0}) - (w_{j_0}/v_{j_0})}{(1/\tilde{\lambda}^m) + (w_{j_0}/v_{j_0})} \right|.
$$

And by Lemmas 2.1 and 3.1, there exist $0 < \tau_\varepsilon(A) < 1$ and a constant $C_2 > 0$ such that

$$
d_H(\mathcal{N}(A^m v), \mathcal{N}(A^m(v+w))) \leq C_2 \tau_\varepsilon^m(A) d_E(\mathcal{N}(v), \mathcal{N}(v+w)),
$$

which immediately implies that

$$
C_1 \left| \frac{1}{(1/\tilde{\lambda}^m) + (w_{j_0}/v_{j_0})} \right| \leq C_2 \tau_\varepsilon^m(A) \frac{d_E(\mathcal{N}(v), \mathcal{N}(v+w))}{|(w_{i_0}/v_{i_0}) - (w_{j_0}/v_{j_0})|}.
$$

One then verifies that there exists a constant $C_3 > 0$ (which depends only on x, y) such that

$$
\frac{d_E(\mathcal{N}(v), \mathcal{N}(v+w))}{|(w_{i_0}/v_{i_0}) - (w_{j_0}/v_{j_0})|} < C_3,
$$

and furthermore, there exists a constant $C_4 > 0$ such that for all m,

$$
\left| \frac{1}{(1/\tilde{\lambda}^m) + (w_{j_0}/v_{j_0})} \right| \geq C_4 \tilde{\lambda}^m.
$$

It then follows that after choosing ε small enough and then n_0 large enough, we have

$$
C_1 C_4 \tilde{\lambda}^m \leq C_2 C_3 \tau_\varepsilon^m(A),
$$

which, upon letting m tend to infinity, yields $\tilde{\lambda} \leq \tau_\varepsilon(A)$, where we have used the fact that all the constants C_1, C_2, C_3, C_4 can be chosen independent of ε. Moreover, using the fact that ε can be chosen arbitrarily small, we apply Lemma 2.1 a) to obtain $\tilde{\lambda} \leq \tau(A)$, which immediately leads to $\kappa(A) \leq \tau(A)$ and completes the proof of Part a).

We next prove Part b). First of all, noting that $\mathcal{N}(A^m(v+w))$ converges to x as m tends to infinity, a parallel argument as in the proof of Part a), supplemented by Lemma 2.1 b) (rather than Lemma 2.1 b)), yields that

$$\kappa(A) \leq \max_{i,j} \sum_{l=1}^{n} \max\left\{\frac{a_{li}x_l}{\sum_{m=1}^n a_{mi}x_m} - \frac{a_{lj}x_l}{\sum_{m=1}^n a_{mj}x_m}, 0\right\}.$$

Now, setting

$$\mathcal{T}_0 = \left\{l : \frac{a_{li}x_l}{\sum_{m=1}^n a_{mi}x_m} - \frac{a_{lj}x_l}{\sum_{m=1}^n a_{mj}x_m} \geq 0\right\} \text{ and } \mathcal{T}_1 = \left\{l : \frac{a_{li}x_l}{\sum_{m=1}^n a_{mi}x_m} - \frac{a_{lj}x_l}{\sum_{m=1}^n a_{mj}x_m} < 0\right\},$$

we have

$$\begin{aligned}
&\sum_{l=1}^{n} \max\left\{\frac{a_{li}x_l}{\sum_{m=1}^n a_{mi}x_m} - \frac{a_{lj}x_l}{\sum_{m=1}^n a_{mj}x_m}, 0\right\} \\
&\quad = \sum_{l\in\mathcal{T}_0}\left(\frac{a_{li}x_l}{\sum_{m\in\mathcal{T}_0} a_{mi}x_m + \sum_{m\in\mathcal{T}_1} a_{mi}x_m} - \frac{a_{lj}x_l}{\sum_{m\in\mathcal{T}_0} a_{mj}x_m + \sum_{m\in\mathcal{T}_1} a_{mj}x_m}\right) \\
&\quad \leq \sum_{l\in\mathcal{T}_0}\left(\frac{\beta_{i,j}a_{lj}x_l}{\beta_{i,j}\sum_{m\in\mathcal{T}_0} a_{mj}x_m + \alpha_{i,j}\sum_{m\in\mathcal{T}_1} a_{mj}x_m} - \frac{a_{lj}x_l}{\sum_{m\in\mathcal{T}_0} a_{mj}x_m + \sum_{m\in\mathcal{T}_1} a_{mj}x_m}\right),
\end{aligned}$$

where $\beta_{i,j} = \max_l\{a_{li}/a_{lj}\}$ and $\alpha_{i,j} = \min_l\{a_{li}/a_{lj}\}$. It then follows that

$$\kappa(A) \leq \max_{i,j}\left\{\frac{1}{1+(\alpha_{i,j}/\beta_{i,j})z} - \frac{1}{1+z}\right\} \leq \frac{1}{1+\phi(A)z} - \frac{1}{1+z},$$

where

$$z = \frac{\sum_{m\in\mathcal{T}_1} a_{mj}x_m}{\sum_{m\in\mathcal{T}_0} a_{mj}x_m}.$$

It has been established in [14] that

$$\frac{\min_m x_m}{\max_m x_m} \geq \eta(A),$$

which immediately implies

$$0 \leq z \leq \frac{(n-1)\theta(A)\max_m x_m}{\min_m x_m} \leq \frac{(n-1)\theta(A)}{\eta(A)},$$

completing the proof of Part b). □

Acknowledgement. The second author would like to thank Venkat Anantharam and Navin Kashyap for insightful discussions and pointing out relevant references. This work is supported by the Research Grants Council of the Hong Kong Special Administrative Region, China, under Project 17301017 and by the National Natural Science Foundation of China, under Project 61871343.

References

[1] R. Bapat and T. RagHavan. *Nonnegative Matrices and Applications*, New York: Cambridge University Press, 1997.

[2] A. Berman and R. Plemmons. *Nonnegative Matrices in the Mathematical Sciences*, Philadephia, Pa.: Society for Industrial and Applied Mathematics, 1994.

[3] G. Birkhoff. Extensions of Jentzsch's Theorem. *Transactions of the American Mathematical Society*, vol. 85, no. 1, pp. 219-227, 1957.

[4] A. Brouwer and W. Haemers. *Spectra of graphs*, Springer, New York, 2012.

[5] F. Chung. *Spectral graph theory*, Providence, R.I.: Published for the Conference Board of the mathematical sciences by the American Mathematical Society, 1997.

[6] L. Dubois. Projective metrics and contraction principles for complex cones. *Journal of the London Mathematical Society*, vol. 79, no. 3, pp. 719-727, 2009.

[7] G. Frobenius. Über matrizen aus positiven elementen. *Sitzungsberichte Preussische Akademie der Wissenschaft*, Berlin, pp. 471-476, 514-518, 1908, 1909.

[8] G. Frobenius. Über matrizen aus nicht negativen elementen. *Sitzungsberichte Preussische Akademie der Wissenschaft*, Berlin, pp. 456-477, 1912.

[9] G. Han and B. Marcus. Analyticity of entropy rate of hidden Markov chains. *IEEE Trans. Info. Theory*, vol. 52, no. 12, pp. 5251-5266, 2006.

[10] G. Han, B. Marcus and Y. Peres. A note on a complex Hilbert metric with application to domain of analyticity for entropy rate of hidden Markov processes. *Entropy of Hidden Markov Processes and Connections to Dynamical Systems*, London Mathematical Society Lecture Note Series, vol. 385, pp. 98-116, 2011.

[11] E. Hopf. An inequality for positive linear integral operators. *J. Math. Mech.*, vol. 12, no. 5, pp. 683-692, 1963.

[12] B. Lemmens and R. Nussbaum. *Nonlinear Perron-Frobenius Theory*, Cambridge University Press, 2012.

[13] D. Levin and Y. Peres. *Markov Chains and Mixing Times*, American Mathematical Society, 2nd Revised Edition, 2017.

[14] M. Lynn and W. Timlake. Bounds for Perron eigenvectors and subdominant eigenvalues of positive matrices. *Linear Algebra and its Applications*, vol. 2, no. 2, pp. 143-152, 1969.

[15] H. Minc. *Nonnegative Matrices*, New York: Wiley, 1988.

[16] R. Montenegro and P. Tetali. *Mathematical Aspects of Mixing Times in Markov chains*, Foundations and Trends in Theoretical Computer Science, Now Publishers, 2006.

[17] A. Ostrowski. On positive matrices. *Math. Ann.*, vol. 150, no. 3, pp. 276-284, 1963.

[18] A. Ostrowski. Positive matrices and functional analysis. *Recent Advances in Matrix Theory*, Madison: Univ. of Wisconsin Press, 1964.

[19] O. Perron. Grundlagen für eine theorie des Jacobischen Kettenbruchalgorithmus. *Math. Ann.*, vol. 64, pp. 11-76, 1907.

[20] U. Rothblum and C. Tan. Upper bounds on the maximum modulus of subdominant eigenvalues of nonnegative matrices. *Linear Algebra Appl*, vol. 66, pp. 45-86, 1985.

[21] H. Rugh. Cones and gauges in complex spaces: Spectral gaps and complex Perron-Frobenius theory. *Annals of Mathematics*, vol. 171, no. 3, 2010.

[22] E. Seneta. *Non-negative Matrices and Markov Chains*, Springer Series in Statistics, Springer-Verlag, New York Heidelberg Berlin, 1980.