Entropy **2011**, *13*, 2036-2048; doi:10.3390/e13122036

OPEN ACCESS

entropy

ISSN 1099-4300

www.mdpi.com/journal/entropy

Article

Eigenvalue Estimates Using the Kolmogorov-Sinai Entropy

Shih-Feng Shieh

Department of Mathematics, National Taiwan Normal University, 88 SEC. 4, Ting Chou Road, Taipei 11677, Taiwan; E-Mail: sfshieh@ntnu.edu.tw

Received: 31 October 2011; in revised form: 28 November 2011 / Accepted: 12 December 2011 / Published: 20 December 2011

Abstract: The scope of this paper is twofold. First, we use the Kolmogorov-Sinai Entropy to estimate lower bounds for dominant eigenvalues of nonnegative matrices. The lower bound is better than the Rayleigh quotient. Second, we use this estimate to give a nontrivial lower bound for the gaps of dominant eigenvalues of $\mathbf{A}$ and $\mathbf{A} + \mathbf{V}$.

Keywords: Kolmogorov-Sinai entropy; Parry's theorem; Eigenvalue estimates

---

1. Introduction

The main concern of this paper is to relate eigenvalue estimates to the Kolmogorov-Sinai entropy for Markov shifts. We shall begin with the definition of the Kolmogorov-Sinai entropy. Let $\mathbf{A} = (a_{ij}) \in \mathbb{R}^{N\times N}$ be an irreducible nonnegative matrix. By an irreducible matrix $\mathbf{A}$, we mean for each $1 \leq i, j \leq N$, there exists positive integer k such that $(\mathbf{A}^k)_{ij} \neq 0$. A matrix $\mathbf{P} = (p_{ij}) \in \mathbb{R}^{N\times N}$ is said to be a *stochastic matrix compatible with* $\mathbf{A}$, if $\mathbf{P}$ satisfies

1. $0 < p_{ij} \leq 1$ if $a_{ij} > 0$,
2. $p_{ij} = 0$ if $a_{ij} = 0$,
3. $\displaystyle\sum_{j=1}^{N} p_{ij} = 1$, for all $i = 1, \ldots, N$.

We denote by $\mathcal{P}_{\mathbf{A}}$ the set of all stochastic matrices compatible with $\mathbf{A}$. By Perron-Frobenius Theorem, it is easily seen that every stochastic matrix $\mathbf{P}$ has a unique left eigenvector $\mathbf{q} > 0$ corresponding to eigenvalue 1 with $\sum_{i=1}^{N} q_i = 1$. Here we say $\mathbf{q}$ is the *stationary probability vector* associated with $\mathbf{P}$.

For a transition matrix $\mathbf{A}$, *i.e.*, $a_{ij} = 1$ or 0 for each $1 \le i, j \le N$, the *subshift of finite type* generated by $\mathbf{A}$ is defined by

$$\Sigma_{\mathbf{A}} = \{\mathbf{i} = (i_0, i_1, \ldots) |\ i_j \in \{1, \ldots, N\},\ a_{i_j, i_{j+1}} = 1,\ j = 0, 1, 2, \ldots\}$$

and the shift map on $\Sigma_{\mathbf{A}}$ is defined by $\sigma_{\mathbf{A}}(i_0, i_1, \ldots) = (i_1, i_2, \ldots)$. A cylinder of $\Sigma_{\mathbf{A}}$ is the set

$$C_{j_0, j_1, \ldots, j_n} = \{\mathbf{i} \in \Sigma_{\mathbf{A}} | i_0 = j_0, \ldots, i_n = j_n\}$$

for any $n \ge 0$. Disjoint unions of cylinders form an algebra which generates the Borel σ-algebra of $\Sigma_{\mathbf{A}}$. For any $\mathbf{P} \in \mathcal{P}_{\mathbf{A}}$ and its associated stationary probability vector $\mathbf{q}$, the Markov measure of a cylinder may then be defined by

$$\mu_{\mathbf{P},\mathbf{q}}(C_{j_0, j_1, \ldots, j_n}) = q_{j_0} p_{j_0, j_1} \cdots p_{j_{n-1}, j_n}$$

Here $\mu_{\mathbf{P},\mathbf{q}}$ is an invariant measure under the shift map $\sigma_{\mathbf{A}}$ (see e.g., [8]). The *Kolmogorov-Sinai entropy* (or called the *measure theoretic entropy*) of $\sigma_{\mathbf{A}}$ under the invariant measure $\mu_{\mathbf{P},\mathbf{q}}$ is defined by

$$h_{\mathbf{P},\mathbf{q}}(\sigma_{\mathbf{A}}) = \lim_{n \to \infty} \frac{1}{n} \sum_{j_0, j_1, \ldots, j_n} H(\mu_{\mathbf{P},\mathbf{q}}(C_{j_0, j_1, \ldots, j_n}))$$

where $H(x) = -x \log x$ and the convention $0 \log 0 = 0$ is adopted. The notion of the Kolmogorov-Sinai entropy was first studied by Kolmogorov in 1958 on the problems arising from information theory and dimension of functional spaces, that measures the uncertainty of the dynamical systems (see e.g., [6,7]). It is shown in [8] (p. 221) that

$$h_{\mathbf{P},\mathbf{q}}(\sigma_{\mathbf{A}}) = -\sum_{ij} q_i p_{ij} \log p_{ij} \tag{1}$$

where the summation in (1) is taken over all i, j with $a_{ij} = 1$. On the other hand, it is shown by Parry [9] (Theorems 6 and 7) that the Kolmogorov-Sinai entropy of $\sigma_{\mathbf{A}}$ has an upper bound $\log \lambda_N(\mathbf{A})$.

Theorem 1.1 (**Parry's Theorem**)**.** *Let* $\mathbf{A}$ *be an* $N \times N$ *irreducible transition matrix. Then for any* $\mathbf{P} \in \mathcal{P}_{\mathbf{A}}$ *and its associated stationary probability vector* $\mathbf{q}$*, we have*

$$h_{\mathbf{P},\mathbf{q}}(\sigma_{\mathbf{A}}) \le \log \lambda_N(\mathbf{A}) \tag{2}$$

where $\lambda_N(\mathbf{A})$ *denotes the dominant eigenvalue of* $\mathbf{A}$*. Moreover, if* $\mathbf{A}$ *is regular (*$\mathbf{A}^n > 0$ *for some* $n > 0$*), the equality in* (2) *holds for some unique* $\mathbf{P} \in \mathcal{P}_{\mathbf{A}}$ *and* $\mathbf{q}$ *the stationary probability vector associated with* $\mathbf{P}$*.*

Parry's Theorem shows the Kolmogorov-Sinai entropy for a Markov shift is less than or equal to its topological entropy (that is, $\log \lambda_N(\mathbf{A})$) and exactly one of the Markov measures on $\Sigma_{\mathbf{A}}$ maximizes the Kolmogorov-Sinai entropy of $\sigma_{\mathbf{A}}$ provided it is topological mixing. This is also a crucial lemma for showing the Variational Property of Entropy [8] (Proposition 8.1) in the ergodic theory. However, from the viewpoint of eigenvalue problems, combination of (1) and (2) gives a lower bound for the dominant eigenvalue of the transition matrix $\mathbf{A}$. In this paper, we generalize Parry's Theorem to general $N \times N$

irreducible nonnegative matrices. Toward this end, we extend the entropy of irreducible nonnegative matrices by

$$h_{\mathbf{P},\mathbf{q},\mathbf{A}} = -\sum_{ij} q_i p_{ij} \log \frac{p_{ij}}{a_{ij}}$$

It is easily seen that $h_{\mathbf{P},\mathbf{q},\mathbf{A}} = h_{\mathbf{P},\mathbf{q}}(\sigma_{\mathbf{A}})$.

Theorem 1.2 (**Main Result 1: The Generalized Parry's Theorem**). *Let* $\mathbf{A} \in \mathbb{R}^{N\times N}$ *an irreducible nonnegative matrix. Let* $\mathbf{P} \in \mathcal{P}_{\mathbf{A}}$ *and* $\mathbf{q}$ *be a stationary probability vector associated with* $\mathbf{P}$*, then we have*

$$h_{\mathbf{P},\mathbf{q},\mathbf{A}} \leq \log \lambda_N(\mathbf{A}) \tag{3}$$

where the summation is taken over all i, j *with* $a_{ij} > 0$*. Moreover, the equality in* (3) *holds when*

$$\mathbf{P} = \frac{1}{\lambda_N(\mathbf{A})} \mathrm{diag}(\mathbf{x})^{-1} \mathbf{A} \mathrm{diag}(\mathbf{x})$$

and

$$\mathbf{q} = \frac{\mathbf{y} \circ \mathbf{x}}{\mathbf{y}^\top \mathbf{x}}$$

where $\mathbf{x} > 0$ *and* $\mathbf{y} > 0$ *are, respectively, the right and left eigenvectors of* $\mathbf{A}$ *corresponding to the eigenvalue* $\lambda_N(\mathbf{A})$*. Here, diag*$(\mathbf{x})$ *denotes the diagonal matrix with* $\mathbf{x}$ *on its diagonal,* $\mathbf{y} \circ \mathbf{x}$ *denotes the vector* $(y_1x_1, \ldots, y_Nx_N)$*, and* $\mathbf{y}^\top$ *denotes the transpose of the column vector* $\mathbf{y}$*.*

Lower bound estimates for the dominant eigenvalue of a symmetric irreducible nonnegative matrix play an important role in various fields, e.g., the complexity of a symbolic dynamical system [5], synchronization problem of coupled systems [10], or the ground state estimates of Schrödinger operators [2]. A usual way to estimate the lower bound for $\lambda_N(\mathbf{A})$ is the Rayleigh quotient

$$\lambda_N(\mathbf{A}) \geq \frac{\mathbf{x}^\top \mathbf{A} \mathbf{x}}{\mathbf{x}^\top \mathbf{x}}$$

It is also well-known that (see e.g., [4] (Theorem 8.1.26)),

$$\min_{1\leq i\leq N} \frac{1}{x_i} \sum_{j=1}^{N} a_{ij}x_j \leq \lambda_N(\mathbf{A}) \leq \max_{1\leq i\leq N} \frac{1}{x_i} \sum_{j=1}^{N} a_{ij}x_j \tag{4}$$

provided that $\mathbf{A} \in \mathbb{R}^{N\times N}$ is nonnegative and $\mathbf{x} \in \mathbb{R}^N$ is positive. Comparing the lower bound estimate (3) with (4) as well as with the Rayleigh quotient, we have the following result.

Corollary 1.3. *Let* $\mathbf{A} \in \mathbb{R}^{N\times N}$ *be a symmetric, irreducible nonnegative matrix. Suppose* $\mathbf{x} \in \mathbb{R}^N$ *be positive. Then the matrix* $\mathbf{P} = \mathrm{diag}(\mathbf{A}\mathbf{x})^{-1}\mathbf{A}\mathrm{diag}(\mathbf{x})$ *is in* $\mathcal{P}_{\mathbf{A}}$ *and* $\mathbf{q} = \dfrac{\mathbf{x} \circ (\mathbf{A}\mathbf{x})}{\mathbf{x}^\top \mathbf{A}\mathbf{x}}$ *is the stationary probability vector associated with* $\mathbf{P}$*. In addition,*

$$h_{\mathbf{P},\mathbf{q},\mathbf{A}} \geq \log \left(\min_{1\leq i\leq N} \frac{1}{x_i} \sum_{j=1}^{N} a_{ij}x_j \right)$$

and

$$h_{\mathbf{P},\mathbf{q},\mathbf{A}} \geq \log \left(\frac{\mathbf{x}^\top \mathbf{A}\mathbf{x}}{\mathbf{x}^\top \mathbf{x}} \right)$$

Here, each equality holds if and only if $\mathbf{x}$ *is the eigenvector of* $\mathbf{A}$ *corresponding to the eigenvalue* $\lambda_N(\mathbf{A})$*.*

Here we remark that for any arbitrary irreducible nonnegative matrix $\mathbf{A}$, the entropy $h_{\mathbf{P},\mathbf{q},\mathbf{A}}$ involves the left eigenvector $\mathbf{q}$ of $\mathbf{P}$. Hence, the lower bound estimate (3) is merely a formal expression. However, for a symmetric irreducible nonnegative matrix $\mathbf{A}$ and $\mathbf{P}$ chosen as in Corollary 1.3, the vector $\mathbf{q}$ can be explicitly expressed. Therefore, $h_{\mathbf{P},\mathbf{q},\mathbf{A}}$ can be written in an explicit form. We shall further show in Proposition 2.6 that $h_{\mathbf{P},\mathbf{q},\mathbf{A}} = \frac{-1}{\mathbf{x}^\top \mathbf{y}} \sum_{i=1}^{N} x_i y_i \log \frac{x_i}{y_i}$ where $\mathbf{y} = \mathbf{A}\mathbf{x}$.

Considering symmetric nonnegative $\mathbf{A}$ and its perturbation $\mathbf{A}+\mathbf{V}$, it is easily seen that $\lambda_N(\mathbf{A}+\mathbf{V}) - \lambda_N(\mathbf{A}) \geq \mathbf{x}^\top \mathbf{V}\mathbf{x}$, where $\mathbf{x}$ is the normalized eigenvector of $\mathbf{A}$ corresponding to $\lambda_N(\mathbf{A})$. This gives a trivial lower bound for the gap of $\lambda_N(\mathbf{A}+\mathbf{V})$ and $\lambda_N(\mathbf{A})$. Upper bound estimates for the gap are well studied in the perturbation theory [4,11]. By considering $\mathbf{A}+\mathbf{V}$ as a low rank perturbation of $\mathbf{A}$, the interlace structure of eigenvalues of $\mathbf{A}+\mathbf{V}$ and of $\mathbf{A}$ is studied by [1,3]. In the second result of this paper, we give a nontrivial lower bound for $\lambda_N(\mathbf{A}+\mathbf{V}) - \lambda_N(\mathbf{A})$.

Theorem 1.4 (**Main Result 2**). *Let* $\mathbf{A} \in \mathbb{R}^{N\times N}$ *be an irreducible nonnegative matrix and* $\mathbf{x} > 0$ *be the eigenvector of* $\mathbf{A}$ *corresponding to* $\lambda_N(\mathbf{A})$ *with* $\|\mathbf{x}\|_2 = 1$. *Suppose* $\mathbf{A}$ *is symmetric. Then for any nonnegative* $\mathbf{V} = \mathrm{diag}(v_1, \ldots, v_N)$, *we have*

$$\lambda_N(\mathbf{A}+\mathbf{V}) - \lambda_N(\mathbf{A}) \geq \frac{f(1/\lambda_N(\mathbf{A})) - 1}{1/\lambda_N(\mathbf{A})} \tag{5}$$

where

$$f(z) = \prod_{i=1}^{N} (1+v_i z)^{\frac{(1+v_i z)x_i^2}{\sum_{j=1}^{N}(1+v_j z)x_i^2}}$$

Here $(f(1/\lambda_N(\mathbf{A})) - 1)\lambda_N(\mathbf{A}) \geq \mathbf{x}^\top \mathbf{V}\mathbf{x}$. *Furthermore, the equality in* (5) *holds if and only if* $v_1 = \cdots = v_N$.

This paper is organized as follows. In Section 2, we prove the generalized Parry's Theorem in three steps. First, we prove the case in which the matrix $\mathbf{A}$ has only integer entries. Next we show that Theorem 1.2 is true for nonnegative matrices with rational entries. Finally we show that it holds true for all irreducible nonnegative matrices. The proof of Corollary 1.3 is given at the end of this section. In Section 3, we give the proof of Theorem 1.4. We conclude this paper in Section 4.

Throughout this paper, we use the boldface alphabet (or symbols) to denote matrices (or vectors). For $\mathbf{u}, \mathbf{v} \in \mathbb{R}^N$, the Hadamard product of $\mathbf{u}$ and $\mathbf{v}$ is their elementwise product which is denoted by $\mathbf{u} \circ \mathbf{v} = (u_i v_i)_{1\leq i\leq N}$. The notation $\mathrm{diag}(\mathbf{u})$ denotes the $N \times N$ diagonal matrix with $\mathbf{u}$ on its diagonal. A matrix $\mathbf{A} = (a_{ij}) \in \mathbb{R}^{N\times N}$ is said to be a transition matrix if $a_{ij} = 1$ or 0 for all $1 \leq i, j \leq N$. $\lambda_1(\mathbf{A}) \leq \cdots \leq \lambda_N(\mathbf{A})$ denotes the dominant eigenvalue of a nonnegative matrix $\mathbf{A}$.

2. Proof of the Generalized Parry's Theorem and Corollary 1.3

In this section, we shall prove the generalized Parry's Theorem and Corollary 1.3. To prove inequality (3), we proceed in three steps.

Step 1: Inequality (3) is true for all irreducible nonnegative matrices with integer entries.

Let $\mathbf{A}$ be an irreducible nonnegative matrix with integer entries. To adopt Parry's Theorem, we shall construct a transition matrix $\bar{\mathbf{A}}$ corresponding to $\mathbf{A}$ for which $\lambda_N(\bar{\mathbf{A}}) = \lambda_N(\mathbf{A})^{1/2}$. To this end, we define the sets of indexes:

$$\mathcal{I} = \{1, \ldots, N\}$$
$$\mathcal{E} = \{\overrightarrow{ij}^{(k)} |\, a_{ij} \neq 0,\ 1 \leq k \leq a_{ij}\}$$

Let $\tilde{N} = \sum_{i,j=1}^{N} a_{ij} = \#\mathcal{E}$ and $\bar{N} = N + \tilde{N}$. The transition matrix $\bar{\mathbf{A}} \in \mathbb{R}^{\bar{N}\times\bar{N}}$ corresponding to $\mathbf{A}$ with index set $\mathcal{I} \cup \mathcal{E}$ is defined as follows

$$(1)\ \bar{a}_{i,\overrightarrow{ij}^{(k)}} = 1,\ \text{for all } 1 \leq k \leq a_{ij} \text{ if } a_{ij} \neq 0, \tag{6a}$$
$$(2)\ \bar{a}_{\overrightarrow{ij}^{(k)},j} = 1,\ \text{for all } 1 \leq k \leq a_{ij} \text{ if } a_{ij} \neq 0, \tag{6b}$$
$$(3)\ \text{the rest entries are set to be zero} \tag{6c}$$

It is easily seen that $\bar{\mathbf{A}}$ can be written in the block form:

$$\bar{\mathbf{A}} = \begin{bmatrix} \mathbf{0}_{N\times N} & \bar{\mathbf{A}}_{\mathcal{I}\mathcal{E}} \\ \bar{\mathbf{A}}_{\mathcal{E}\mathcal{I}} & \mathbf{0}_{\tilde{N}\times\tilde{N}} \end{bmatrix} \tag{7}$$

where $\mathbf{0}_{N\times N}$ and $\mathbf{0}_{\tilde{N}\times\tilde{N}}$ are, respectively, the zero matrices in $\mathbb{R}^{N\times N}$ and $\mathbb{R}^{\tilde{N}\times\tilde{N}}$, $\bar{\mathbf{A}}_{\mathcal{I}\mathcal{E}} \in \mathbb{R}^{N\times\tilde{N}}$ and $\bar{\mathbf{A}}_{\mathcal{E}\mathcal{I}} \in \mathbb{R}^{\tilde{N}\times N}$.

Proposition 2.1. $\lambda_{\bar{N}}(\bar{\mathbf{A}}) = \lambda_N(\mathbf{A})^{1/2}$.

Proof. From (7), we see that

$$\bar{\mathbf{A}}^2 = \left[\begin{array}{c|c} \bar{\mathbf{A}}_{\mathcal{I}\mathcal{E}}\bar{\mathbf{A}}_{\mathcal{E}\mathcal{I}} & \mathbf{0}_{N\times\tilde{N}} \\ \hline \mathbf{0}_{\tilde{N}\times N} & \bar{\mathbf{A}}_{\mathcal{E}\mathcal{I}}\bar{\mathbf{A}}_{\mathcal{I}\mathcal{E}} \end{array}\right]$$

From (6a) and (6b), for each i, j with $a_{ij} \neq 0$, we have

$$\sum_{k=1}^{a_{ij}} \bar{a}_{i,\overrightarrow{ij}^{(k)}} \bar{a}_{\overrightarrow{ij}^{(k)},j} = a_{ij} \tag{8}$$

Using (8), together with (6c), we have

$$\begin{aligned} (\bar{\mathbf{A}}_{\mathcal{I}\mathcal{E}}\bar{\mathbf{A}}_{\mathcal{E}\mathcal{I}})_{ij} &= \sum_{\alpha\in\mathcal{E}} \bar{a}_{i\alpha}\bar{a}_{\alpha j} \\ &= \begin{cases} \sum_k \bar{a}_{i,\overrightarrow{ij}^{(k)}} \bar{a}_{\overrightarrow{ij}^{(k)},j} = a_{ij} & \text{if } a_{ij} \neq 0 \\ 0 = a_{ij} & \text{if } a_{ij} = 0 \end{cases} \end{aligned} \tag{9}$$

From (9) we see that $\bar{\mathbf{A}}_{\mathcal{I}\mathcal{E}}\bar{\mathbf{A}}_{\mathcal{E}\mathcal{I}} = \mathbf{A}$. Hence $\lambda_{\bar{N}}(\bar{\mathbf{A}}^2) = \lambda_N(\bar{\mathbf{A}}_{\mathcal{I}\mathcal{E}}\bar{\mathbf{A}}_{\mathcal{E}\mathcal{I}}) = \lambda_{\tilde{N}}(\bar{\mathbf{A}}_{\mathcal{E}\mathcal{I}}\bar{\mathbf{A}}_{\mathcal{I}\mathcal{E}}) = \lambda_N(\mathbf{A})$. On the other hand, $\bar{\mathbf{A}}$ is a nonnegative matrix. From Perron-Frobenius Theorem, its dominant eigenvalue is nonnegative. The assertion follows. □

Remark 2.1. In the language of graph theory, a_{ij} represents the number of directed edges from vertex i to vertex j. Hence $\sum_{ij}(\mathbf{A}^n)_{ij}$ equals to the number of all possible routes of length $n+1$, *i.e.*,

$$\#\{\text{all possible routes of length } n+1\} = \sum_{ij}(\mathbf{A}^n)_{ij} = O(\lambda_N(\mathbf{A})^n)$$

For the construction of $\bar{\mathbf{A}}$, we add an additional vertex on every edge from vertex i to vertex j (See Figure 2.1 for the illustration). Hence, each route that obeys the rule defined by $\mathbf{A}$,

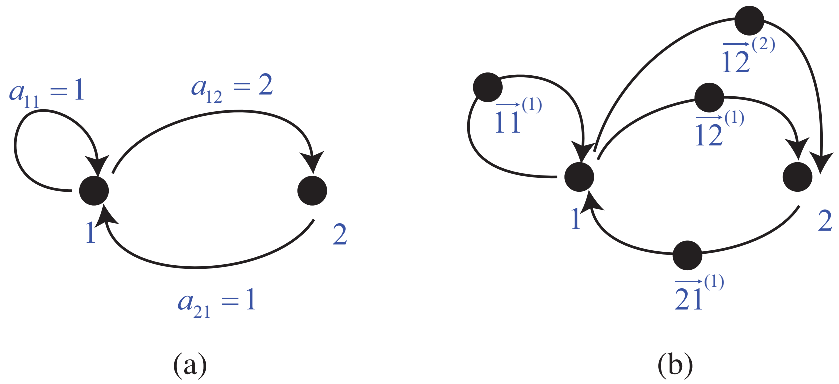
$$(i_1, i_2, \ldots, i_j, i_{j+1}, \ldots, i_{n-1}, i_n), \text{ provided } a_{i_j i_{j+1}} > 0 \text{ for all } j = 1, \ldots, n-1 \tag{10}$$

now becomes one of the following routes according to the rule defined by $\bar{\mathbf{A}}$:

$$(i_1, \overrightarrow{i_1 i_2}^{(k_1)}, i_2, \ldots i_j, \overrightarrow{i_j i_{j+1}}^{(k_j)} i_{j+1}, \ldots, i_{n-1}, \overrightarrow{i_{n-1} i_n}^{(k_{n-1})}, i_n) \tag{11}$$

where $1 \le k_j \le a_{i_j i_{j+1}}$, $j = 1, \ldots, n-1$. However, a route of the form in (11) is equivalent to the form in (10) but its length is doubled. Hence $O(\lambda_N(\bar{\mathbf{A}})^{2n}) = O(\lambda_N(\mathbf{A})^n)$.

Figure 1. Illustration for Remark 2.1 with the example $\mathbf{A} = \begin{bmatrix} 1 & 2 \\ 1 & 0 \end{bmatrix}$.



Now, let $\mathbf{P} \in \mathcal{P}_{\mathbf{A}}$ be given and $\mathbf{q}$ be its associated stationary probability vector. We shall accordingly define a stochastic matrix $\bar{\mathbf{P}} \in \mathcal{P}_{\bar{\mathbf{A}}}$ and its associated stationary probability vector $\bar{\mathbf{q}}$. The stochastic matrix $\bar{\mathbf{P}}$ is defined as follows:

$$(1)\ \bar{p}_{i,\overrightarrow{ij}^{(k)}} = \frac{p_{ij}}{a_{ij}} \text{ for all } 1 \le k \le a_{ij} \text{ provided } a_{ij} > 0 \tag{12a}$$

$$(2)\ \bar{p}_{\overrightarrow{ij}^{(k)},j} = 1 \text{ for all } 1 \le k \le a_{ij} \text{ provided } a_{ij} > 0 \tag{12b}$$

$$(3) \text{ the rest entries are set to zero} \tag{12c}$$

From (6) and (12), it is easily seen that $\bar{\mathbf{P}}$ is a stochastic matrix compatible with $\bar{\mathbf{A}}$. Let the vector $\bar{\mathbf{q}} \in \mathbb{R}^{N+\tilde{N}}$ be defined by

$$\bar{q}_i = \frac{q_i}{2},\ 1 \le i \le N \tag{13a}$$

and

$$\bar{q}_{\overrightarrow{ij}^{(k)}} = \frac{q_i p_{ij}}{2a_{ij}}, \text{ for all } 1 \le k \le a_{ij} \text{ with } a_{ij} > 0 \tag{13b}$$

Proposition 2.2. $\bar{\mathbf{q}}$ *is the stationary probability vector associated with* $\bar{\mathbf{P}}$.

Proof. We first show that $\bar{\mathbf{q}}$ is a left eigenvector of $\bar{\mathbf{P}}$ with the corresponding eigenvalue 1. For any $1 \leq j \leq N$, using (12b), (13b), and the fact that $\mathbf{q}^\top \mathbf{P} = \mathbf{q}^\top$, we have

$$\begin{aligned}(\bar{\mathbf{q}}^\top \bar{\mathbf{P}})_j = \sum_{i,k} \bar{q}_{\overrightarrow{ij}^{(k)}} \bar{p}_{\overrightarrow{ij}^{(k)},j} &= \sum_{i,a_{ij}>0} \sum_{k=1}^{a_{ij}} \frac{1}{2} q_i \frac{p_{ij}}{a_{ij}} \cdot 1 \\ &= \sum_i \frac{1}{2} q_i p_{ij} = \frac{1}{2} q_j = \bar{q}_j \end{aligned} \tag{14a}$$

On the other hand, using (12a) and (13a), for all $\overrightarrow{ij}^{(k)}$ with $a_{ij} > 0$ and $1 \leq k \leq a_{ij}$, we have

$$\begin{aligned}(\bar{\mathbf{q}}^\top \bar{\mathbf{P}})_{\overrightarrow{ij}^{(k)}} &= \bar{q}_i \bar{p}_{i,\overrightarrow{ij}^{(k)}} \\ &= \frac{1}{2} q_i \frac{p_{ij}}{a_{ij}} = \bar{q}_{\overrightarrow{ij}^{(k)}}\end{aligned} \tag{14b}$$

In (14), we have proved $\bar{\mathbf{q}}^\top \bar{\mathbf{P}} = \bar{\mathbf{q}}^\top$. Now we show that the total sum of entries of $\bar{\mathbf{q}}$ is 1. Using the fact

$$\begin{aligned}\sum_{ij} \sum_{k=1}^{a_{ij}} \bar{q}_{\overrightarrow{ij}^{(k)}} &= \sum_{ij} \sum_{k=1}^{a_{ij}} \frac{q_i p_{ij}}{2a_{ij}} \\ &= \sum_{ij} \frac{1}{2} q_i p_{ij} = \frac{1}{2} \sum_i q_i\end{aligned}$$

we conclude that

$$\begin{aligned}\sum_{\alpha \in \mathcal{I} \cup \mathcal{E}} (\mathbf{q})_\alpha &= \sum_i \bar{q}_i + \sum_{ij} \sum_{k=1}^{a_{ij}} \bar{q}_{\overrightarrow{ij}^{(k)}} \\ &= \frac{1}{2} \sum q_i + \frac{1}{2} \sum_i q_i = 1\end{aligned}$$

The proof is complete. □

From the construction of the transition matrix $\bar{\mathbf{A}}$, it is easily seen that $\bar{\mathbf{A}}$ is irreducible. In (12) and Proposition 2.2, we show that $\bar{\mathbf{P}} \in \mathcal{P}_{\bar{\mathbf{A}}}$ and the vector $\bar{\mathbf{q}}$ defined by (13) is its associated stationary probability vector. Hence the Kolmogorov-Sinai entropy $h_{\bar{\mathbf{P}},\bar{\mathbf{q}}}(\sigma_{\bar{\mathbf{A}}})$ is well-defined. Now we give the relationship between the quantities $h_{\bar{\mathbf{P}},\bar{\mathbf{q}}}(\sigma_{\bar{\mathbf{A}}})$ and $h_{\mathbf{P},\mathbf{q},\mathbf{A}}$ defined in Equation (3).

Proposition 2.3.

$$h_{\bar{\mathbf{P}},\bar{\mathbf{q}}}(\sigma_{\bar{\mathbf{A}}}) = \frac{1}{2} h_{\mathbf{P},\mathbf{q},\mathbf{A}}$$

Proof. We note that by (12b), $\log \bar{p}_{\vec{ij}^{(k)},j} = 0$ if $a_{ij} > 0$. Using the definition of $\bar{\mathbf{P}}$ and $\bar{\mathbf{q}}$ in (12) and (13), as well as the entropy formula (1), we have

$$\begin{aligned} h_{\bar{\mathbf{P}},\bar{\mathbf{q}}}(\sigma_{\bar{\mathbf{A}}}) &= -\sum_{ij,a_{ij}>0} \sum_{k=1}^{a_{ij}} \bar{q}_i \bar{p}_{i,\vec{ij}^{(k)}} \log \bar{p}_{i,\vec{ij}^{(k)}} \\ &= -\sum_{ij,a_{ij}>0} \sum_{k=1}^{a_{ij}} \frac{1}{2} q_i \frac{p_{ij}}{a_{ij}} \log \frac{p_{ij}}{a_{ij}} \\ &= -\sum_{ij,a_{ij}>0} \frac{1}{2} q_i p_{ij} \log \frac{p_{ij}}{a_{ij}} \\ &= \frac{1}{2} h_{\mathbf{P},\mathbf{q},\mathbf{A}} \end{aligned}$$

The proof is complete. □

Using Proposition 2.3, 2.1, and Parry's Theorem 1.1, it follows that

$$\begin{aligned} \frac{1}{2} h_{\mathbf{P},\mathbf{q},\mathbf{A}} &= h_{\bar{\mathbf{P}},\bar{\mathbf{q}}}(\sigma_{\bar{\mathbf{A}}}) \\ &\leq \log \lambda_N(\bar{\mathbf{A}}) = \frac{1}{2} \log \lambda_N(\mathbf{A}) \end{aligned} \tag{15}$$

Step 2: Inequality (3) *is true for all irreducible nonnegative matrices with rational entries.*

Any $N \times N$ nonnegative matrix with all entries that are rational can be written as $\mathbf{A}/n$ where $\mathbf{A}$ is a nonnegative matrix with integer entries and n is an positive integer. Suppose $\mathbf{A}$ is irreducible and $\mathbf{P} \in \mathcal{P}_{\mathbf{A}/n}$. Note that $\mathcal{P}_{\mathbf{A}/n} = \mathcal{P}_{\mathbf{A}}$. Letting $\mathbf{q}$ be a stationary probability vector associated with $\mathbf{P}$, inequality (3) for $\mathbf{A}/n$ follows from the following proposition.

Proposition 2.4.

$$h_{\mathbf{P},\mathbf{q},\mathbf{A}/n} \leq \log \lambda_N(\mathbf{A}/n)$$

Proof. From the definition of $h_{\mathbf{P},\mathbf{q},\mathbf{A}/n}$, we see that

$$\begin{aligned} h_{\mathbf{P},\mathbf{q},\mathbf{A}/n} = -\sum_{ij,a_{ij}>0} q_i p_{ij} \log \frac{p_{ij} n}{a_{ij}} &= -\sum_{ij,a_{ij}>0} q_i p_{ij} \log \frac{p_{ij}}{a_{ij}} - \sum_{ij} q_i p_{ij} \log n \\ &= h_{\mathbf{P},\mathbf{q},\mathbf{A}} - \sum_{ij} q_i p_{ij} \log n \end{aligned} \tag{16}$$

On the other hand, since $\mathbf{q}^\top \mathbf{P} = \mathbf{q}^\top$ and $\sum q_i = 1$, we have

$$\sum_{ij} q_i p_{ij} \log n = \log n \tag{17}$$

Substituting (17) into (16) and using the result (15) in **Step 1**, we have

$$h_{\mathbf{P},\mathbf{q},\mathbf{A}/n} = h_{\mathbf{P},\mathbf{q},\mathbf{A}} - \log n \leq \log \lambda_N(\mathbf{A}) - \log n = \log \lambda_N(\mathbf{A}/n)$$

□

Step 3: Inequality (3) is true for all irreducible nonnegative matrices.

It remains to show (3) holds for all nonnegative $\mathbf{A}$ with irrational entries. The assertion follows from **Step 2** and the continuous dependence of eigenvalues with respect to the matrix.

Now, we give the proof of the second assertion of Theorem 1.2.

Proposition 2.5. *The equality in (3) holds when one chooses*

$$\mathbf{P} = \frac{1}{\lambda_N(\mathbf{A})}\mathrm{diag}(\mathbf{x})^{-1}\mathbf{A}\mathrm{diag}(\mathbf{x})$$

and

$$\mathbf{q} = \frac{\mathbf{y} \circ \mathbf{x}}{\mathbf{y}^\top \mathbf{x}}$$

where $\mathbf{x} > 0$ and $\mathbf{y} > 0$ are, respectively, the right and left eigenvectors of $\mathbf{A}$ corresponding to eigenvalue $\lambda_N(\mathbf{A})$.

Proof. By setting $\mathbf{y}^\top \mathbf{x} = 1$, we may write

$$p_{ij} = \frac{a_{ij}x_j}{\lambda_N(\mathbf{A})x_i} \text{ and } q_i = x_i y_i$$

To ease the notation, set $\lambda_N = \lambda_N(\mathbf{A})$. Hence, we have

$$\begin{aligned}
h_{\mathbf{P},\mathbf{q},\mathbf{A}} &= -\sum_{ij} x_i y_i \frac{a_{ij}x_j}{\lambda_N x_i} \log \frac{x_j}{\lambda_N x_i} \\
&= \sum_{ij} \frac{y_i}{\lambda_N}(a_{ij}x_j)\log(\lambda_N x_i) - \sum_{ij} \frac{x_j}{\lambda_N}(y_i a_{ij}) \log x_j \\
&= \sum_i y_i x_i \log(\lambda_N x_i) - \sum_j x_j y_j \log x_j \left(\text{Use the facts} \sum_j a_{ij}x_j = \lambda_N x_i \text{ and} \sum_i y_i a_{ij} = \lambda_N y_j \right) \\
&= \sum_i x_i y_i \log \lambda_N \\
&= \log \lambda_N
\end{aligned}$$

The proof of Theorem 1.2 is complete. □

In the following, we give the proof of Corollary 1.3. We first prove the following useful proposition. It will be used in Section 3 as well.

Proposition 2.6. *Let $\mathbf{A} \in \mathbb{R}^{N\times N}$ be an irreducible nonnegative matrix. Suppose $\mathbf{A}$ is symmetric and $\mathbf{x} \in \mathbb{R}^N$ be positive. If $\mathbf{P} = \mathrm{diag}(\mathbf{A}\mathbf{x})^{-1}\mathbf{A}\mathrm{diag}(\mathbf{x})$ and $\mathbf{q} = \frac{\mathbf{x}\circ\mathbf{y}}{\mathbf{x}^\top\mathbf{y}}$, where $\mathbf{y} = \mathbf{A}\mathbf{x}$, then*

$$h_{\mathbf{P},\mathbf{q},\mathbf{A}} = \frac{-1}{\mathbf{x}^\top \mathbf{y}} \sum_{i=1}^N x_i y_i \log \frac{x_i}{y_i}$$

From Proposition 2.5, we see that the matrix $\mathbf{P}$ in Proposition 2.5 is a stochastic matrix compatible with $\mathbf{A}$ and $\mathbf{q}$ is its associated stationary probability vector. Hence, the entropy $h_{\mathbf{P},\mathbf{q},\mathbf{A}}$ is well defined. Now, we give the proof of this Proposition.

Proof. Since $\mathbf{A} \geq 0$ is irreducible and $\mathbf{x} > 0$, it follows $\mathbf{Ax} > 0$, and hence, $\mathrm{diag}(\mathbf{Ax})^{-1}$ is well-defined. It is easily seen that $p_{ij} = 0$ if and only if $a_{ij} = 0$. However, $\mathbf{Pe} = \mathrm{diag}(\mathbf{Ax})^{-1}(\mathbf{Ax}) = \mathbf{e}$. This shows that $\mathbf{P} \in \mathcal{P}_{\mathbf{A}}$. On the other hand, since $\mathbf{A}$ is symmetric, we see that $\mathbf{y}^\top = \mathbf{x}^\top\mathbf{A}$. Hence

$$\mathbf{q}^\top\mathbf{P} = (\mathbf{x} \circ (\mathbf{Ax}))^\top \mathrm{diag}(\mathbf{Ax})^{-1}\mathbf{A}\mathrm{diag}(\mathbf{x})/\mathbf{x}^\top\mathbf{Ax} = \mathbf{q}^\top$$

We have proved the first assertion of this proposition. By the definition of $h_{\mathbf{P},\mathbf{q},\mathbf{A}}$ in (3), we have

$$\begin{aligned} h_{\mathbf{P},\mathbf{q},\mathbf{A}} &= -\sum_{ij} \frac{a_{ij}x_ix_j}{\mathbf{x}^\top\mathbf{y}} \log\frac{x_j}{y_i} = \frac{1}{\mathbf{x}^\top\mathbf{y}}\left[\sum_{i=1}^N x_iy_i \log y_i - \sum_{i=1}^N x_jy_j \log x_j\right] \\ &= \frac{-1}{\mathbf{x}^\top\mathbf{y}}\sum_{i=1}^N x_iy_i \log\frac{x_i}{y_i} \end{aligned}$$

This completes the proof. □

Now, we are in a proposition to give the proof of Corollary 1.3.

Proof of Corollary 1.3. For convenience, we let $\mathbf{y} = \mathbf{Ax}$. Hence $\mathbf{q} = \dfrac{\mathbf{x}\circ\mathbf{y}}{\mathbf{x}^\top\mathbf{y}}$ and $p_{ij} = \dfrac{a_{ij}x_j}{y_i}$. Using Proposition 2.6, we have

$$\begin{aligned} h_{\mathbf{P},\mathbf{q},\mathbf{A}} &= \frac{-1}{\mathbf{x}^\top\mathbf{y}}\sum_{i=1}^N x_iy_i \log\frac{x_i}{y_i} && (18) \\ &\geq -\log\frac{\mathbf{x}^\top\mathbf{x}}{\mathbf{x}^\top\mathbf{y}} && (19) \\ &= \log\frac{\mathbf{x}^\top\mathbf{Ax}}{\mathbf{x}^\top\mathbf{x}} \end{aligned}$$

Here inequality (19) follows from Jensen's inequality (see e.g., [12] (Theorem 7.35)) for $-\log$ and the fact that $\sum_{i=1}^N \frac{1}{\mathbf{x}^\top\mathbf{y}}x_iy_i = 1$. Similarly, using Proposition 2.6 and the monotonicity of $\log$, we also see that

$$\begin{aligned} h_{\mathbf{P},\mathbf{q},\mathbf{A}} &= \frac{-1}{\mathbf{x}^\top\mathbf{y}}\sum_{i=1}^N x_iy_i \log\frac{x_i}{y_i} \\ &\geq \frac{1}{\mathbf{x}^\top\mathbf{y}}\sum_{i=1}^N x_iy_i \log\left(\min_{1\leq i\leq N}\frac{y_i}{x_i}\right) && (20) \\ &= \log\left(\min_{1\leq i\leq N}\frac{y_i}{x_i}\right) \end{aligned}$$

This proves the first assertion of Corollary 1.3. It is easily seen that if $\mathbf{x}$ is an eigenvector corresponding to $\lambda_N(\mathbf{A})$, then both equalities in (19) and (20) hold. From the assumption that $\mathbf{A} \geq 0$ is irreducible and $\mathbf{x} > 0$, it follows that $\mathbf{y} > 0$ also. This implies there are N terms in (18). Hence equality in (19) or in (20) holds only if $\frac{x_i}{y_i}$, for all $i = 1, \ldots, N$, are constant. That is, $\mathbf{y} = \mathbf{Ax} = \lambda\mathbf{x}$. Here λ is some eigenvalue of $\mathbf{A}$. However, $\mathbf{x} > 0$. From Perron-Frobenius Theorem it follows $\lambda = \lambda_N(\mathbf{A})$. The proof is complete. □

3. Proof of Theorem 1.4

In this section, we shall give the proof of Theorem 1.4. We first prove (5).

Proposition 3.1. *Let* $\mathbf{A}$, $\mathbf{V}$ *and* $\mathbf{x}$ *be as defined in Theorem 1.4. Then we have*

$$\lambda_N(\mathbf{A}+\mathbf{V}) - \lambda_N(\mathbf{A}) \geq \frac{f(1/\lambda_N(\mathbf{A}))-1}{1/\lambda_N(\mathbf{A})} \tag{21}$$

where

$$f(z) = \prod_{i=1}^{N}(1+v_i z)^{\frac{(1+v_i z)x_i^2}{\sum_{j=1}^{N}(1+v_j z)x_i^2}}$$

The equality holds in (21) *if and only if* $v_1 = \cdots = v_N$.

Proof. To ease the notation, we shall denote $\lambda = \lambda_N(\mathbf{A})$. Let $\mathbf{y} = (\mathbf{A}+\mathbf{V})\mathbf{x} = \lambda\mathbf{x} + \mathbf{V}\mathbf{x}$, $\mathbf{q} = \frac{\mathbf{x}\circ\mathbf{y}}{\mathbf{x}^\top\mathbf{y}}$, and $\mathbf{P} = \mathrm{diag}(\mathbf{y})^{-1}(\mathbf{A}+\mathbf{V})\mathrm{diag}(\mathbf{x}) \in \mathcal{P}_{\mathbf{A}+\mathbf{V}}$. From Theorem 1.2 and Proposition 2.6, we have

$$\begin{aligned}\log\lambda_N(\mathbf{A}+\mathbf{V}) &\geq h_{\mathbf{P},\mathbf{q},\mathbf{A}+\mathbf{V}} \\ &= \frac{1}{\mathbf{x}^\top(\mathbf{A}+\mathbf{V})\mathbf{x}}\sum_{i=1}^{N}(\lambda+v_i)x_i^2\log(\lambda+v_i)\end{aligned} \tag{22}$$

We note that

$$\log\lambda_N(\mathbf{A}) = \frac{1}{\mathbf{x}^\top(\mathbf{A}+\mathbf{V})\mathbf{x}}\sum_{i=1}^{N}(\lambda+v_i)x_i^2\log\lambda \tag{23}$$

Subtracting (23) from (22), we have

$$\log\frac{\lambda_N(\mathbf{A}+\mathbf{V})}{\lambda_N(\mathbf{A})} \geq \frac{1}{\sum_{i=1}^{N}(1+v_i/\lambda)x_i^2}\sum_{i=1}^{N}(1+v_i/\lambda)x_i^2\log(1+v_i/\lambda)$$

and hence,

$$\frac{\lambda_N(\mathbf{A}+\mathbf{V})-\lambda_N(\mathbf{A})}{\lambda_N(\mathbf{A})} \geq f(1/\lambda_N(\mathbf{A}))-1$$

This proves (21). Now we prove the second assertion of this proposition. It is easily seen that $v_1 = \cdots = v_N$ implies the equality in (21) holds. Conversely, suppose the equality in (21) holds. It is equivalent to the equality in (22) holds. Now, we write (22) in an alternative form

$$\begin{aligned}\frac{1}{\mathbf{x}^\top(\mathbf{A}+\mathbf{V})\mathbf{x}}\sum_{i=1}^{N}(\lambda+v_i)x_i^2\log(\lambda+v_i) &\leq \log\left(\frac{1}{\mathbf{x}^\top(\mathbf{A}+\mathbf{V})\mathbf{x}}\sum_{i=1}^{N}(\lambda+v_i)^2x_i^2\right) \qquad (24)\\ &= \log\frac{\mathbf{x}^\top(\mathbf{A}+\mathbf{V})^2\mathbf{x}}{\mathbf{x}^\top(\mathbf{A}+\mathbf{V})\mathbf{x}} \\ &\leq \log\lambda_N(\mathbf{A}+\mathbf{V}) \qquad (25)\end{aligned}$$

Here (24) follows from the convexity of log and Jensen's inequality. Hence, if the equality in (22) holds, then the equality in (25) also holds. This means $\mathbf{x}$ is also an eigenvector of $\mathbf{A}+\mathbf{V}$. However, since $\mathbf{x} > 0$ is the eigenvector of $\mathbf{A}$ corresponding to $\lambda_N(\mathbf{A})$, we conclude that $v_1 = \cdots = v_N$. This completes the proof. □

The following proposition can be obtained from a standard calculation.

Proposition 3.2. *Let f be the real-valued function in Proposition 3.1. Then we have*

$$f'(z) = \left(\frac{b}{1+bz} + \frac{g(z)}{(1+bz)^2}\right) f(z) \tag{26a}$$

$$f''(z) = \left(\frac{g'(z)}{(1+bz)^2} + \frac{g(z)^2}{(1+bz)^4}\right) f(z) \tag{26b}$$

where $b = \sum_{i=1}^N x_i^2 v_i$ and

$$g(z) = \sum_{i=1}^N x_i^2 \sum_{j=1}^N x_j^2 (v_i - v_j) \log(1+v_i z), \tag{27a}$$

$$g'(z) = \frac{1}{2} \sum_{i,j=1}^N x_i^2 x_j^2 (v_i - v_j)^2 \frac{1}{(1+v_i z)(1+v_j z)}, \tag{27b}$$

In the following, we show that the lower bound estimate (5) for $\lambda_N(\mathbf{A}+\mathbf{V}) - \lambda_N(\mathbf{A})$ is greater than $\mathbf{x}^\top \mathbf{V}\mathbf{x}$.

Proposition 3.3. *Let f be the real-valued function in Proposition 3.1. Then we have*

$$\frac{f(1/\lambda_N(\mathbf{A})) - 1}{1/\lambda_N(\mathbf{A})} \geq \mathbf{x}^\top \mathbf{V}\mathbf{x}$$

Proof. It is easily seen from the definition of $f(z)$ that $f(0) = 1$. Hence, using the Mean Value Theorem follows that there exists a $\zeta \in (0, 1/\lambda_N(\mathbf{A}))$ such that

$$\frac{f(1/\lambda_N(\mathbf{A})) - 1}{1/\lambda_N(\mathbf{A})} = f'(\zeta). \tag{28}$$

From (26a) and (27a), we see that $f'(0) = b = \mathbf{x}^\top \mathbf{V}\mathbf{x}$. From (26b), (27a) and (27b), we also see that $f''(z) \geq 0$ for all $z \geq 0$. This implies

$$f'(\zeta) \geq f'(0) = \mathbf{x}^\top \mathbf{V}\mathbf{x} \tag{29}$$

The assertion of this proposition follows from (28) and (29) directly. □

4. Conclusions

In this paper, we first generalize Parry's Theorem to general nonnegative matrices. This can be treated as an estimation for the lower bound for a nonnegative matrix. Second, we use the generalized Parry's Theorem to estimate a nontrivial lower bound of $\lambda_N(\mathbf{A}+\mathbf{V}) - \lambda_N(\mathbf{A})$, provided that $\mathbf{A} \geq 0$ is symmetric and $\mathbf{V} \geq 0$ is a diagonal matrix. The bound is optimal but implicit that can be applied when $\lambda_N(A)$ and its corresponding eigenvector are known. As an interesting topic to be explored in the future, rather than a nonnegative matrix eigenvalue problem, one may wish to derive a similar inequality to (3) for a general square matrix or for a generalized eigenvalue problem $\mathbf{Ax} = \lambda\mathbf{Bx}$.

References

1. Arbenz, P.; Golub, G.H. On the spectral decomposition of Hermitian matrices modified by low rank perturbations with applications. *SIAM J. Matrix Anal. Appl.* **1988**, *9*, 40–58.
2. Chang, S.-M.; Lin, W.-W.; Shieh, S.-F. Gauss-Seidel-type methods for energy states of a multi-component Bose-Einstein condensate. *J. Comp. Phys.* **2005**, *22*, 367–390.
3. Golub, G.H. Some modified matrix eigenvalue problems. *SIAM Rev.* **1973**, *15*, 318–334.
4. Horn, R.A.; Johnson, C.R. *Matrix Analysis*; Cambridge University Press: Cambridge, UK, 1985.
5. Juang, J.; Shieh, S.-F.; Turyn, L. Cellular neural networks: Space-dependent template, mosaic patterns and spatial chaos. *Internat. J. Bifur. Chaos Appl. Sci. Engrg.* **2002**, *12*, 1717–1730.
6. Kolmogorov, A.N. A new metric invariant of transitive dynamical systems and automorphisms of Lebesgue spaces. *Dokl. Akad. Nauk SSSR* **1958**, *119*, 861–864.
7. Kolmogorov, A.N. On the entropy per time unit as a metric invariant of auto-morphisms. *Dokl. Akad. Nauk SSSR* **1958**, *21*, 754–755.
8. Mañé, R. *Ergodic Theory and Differentiable Dynamics*; Springer-Verlag: Berlin, Germany, 1987.
9. Parry, W. Intrinsic Markov chains. *Trans. Amer. Math. Soc.* **1964**, *112*, 55–66.
10. Shieh, S.F.; Wang, Y.Q.; Wei, G.W.; Lai, C.-H. Mathematical analysis of the wavelet method of chaos control. *J. Math. Phys.* **2006**, *47*, 082701.
11. Stewart, G.W.; Sun, J.-G. *Matrix Perturbation Theory*; Academic Press: Boston, MA, USA, 1990.
12. Wheeden, R.L.; Zygmund, A. Measure and integral: An introduction to real analysis. In *Monographs and Textbooks in Pure and Applied Mathematics*; Marcel Dekker: New York, NY, USA, 1977.