Available online at www.sciencedirect.com
ScienceDirect
Linear Algebra and its Applications 418 (2006) 900–912
LINEAR ALGEBRA AND ITS APPLICATIONS
www.elsevier.com/locate/laa

On dispersal and population growth for multistate matrix models

Chi-Kwong Li *, Sebastian J. Schreiber

Department of Mathematics, The College of William and Mary, P.O. Box 8795, Williamsburg, VA 23187, United States

Received 27 July 2005; accepted 17 March 2006
Available online 22 May 2006
Submitted by R.A. Brualdi

Abstract

To describe the dynamics of stage-structured populations with m stages living in n patches, we consider matrix models of the form **SD**, where **S** is a block diagonal matrix with $n \times n$ column substochastic matrices $S_1, \ldots, S_m$ along the diagonal and **D** is a block matrix whose blocks are $n \times n$ nonnegative diagonal matrices. The matrix **S** describes movement between patches and the matrix **D** describes growth and reproduction within the patches. Consider the multiple arc directed graph G consisting of the directed graphs corresponding to the matrices $S_1, \ldots, S_m$, where each directed graph is drawn in a different color. We say G has a *polychromatic cycle* if G has a directed cycle that includes arcs of more than one color. We prove that $\rho(\mathbf{SD}) \leqslant \rho(\mathbf{D})$ for all block matrices **D** with nonnegative diagonal blocks if and only if G has no polychromatic cycle. Applications to ecological models are presented.

AMS classification: 15A51; 15A18; 92D25

Keywords: Stochastic matrices; Multistate matrix models; Population growth

1. Introduction

Denote by $\rho(A)$ the spectral radius of a square matrix A. The column sum norm of $A = (a_{ij})_{1 \leqslant i,j \leqslant n}$ is defined by

* Corresponding author. Tel.: +1 757 221 2042; fax: +1 757 221 7400.
E-mail addresses: ckli@math.wm.edu (C.-K. Li), sjs@math.wm.edu (S.J. Schreiber).

0024-3795/$ - see front matter © 2006 Elsevier Inc. All rights reserved.
doi:10.1016/j.laa.2006.03.025

$$\|A\| = \max\left\{\sum_{i=1}^{n} |a_{ij}| : 1 \leqslant j \leqslant n\right\}.$$

An $n \times n$ nonnegative matrix S is *column substochastic* (respectively, *column stochastic*) if all the column sums are bounded by one (respectively, equal to one). The well-known fact that $\rho(A) \leqslant \|A\|$ immediately implies the following proposition.

Proposition 1.1. *Let D be a diagonal nonnegative matrices and S a column substochastic matrix. Then*

$$\rho(SD) \leqslant \rho(D).$$

For biological models that account for spatial structure, this fact has some important implications. For example, consider a population (e.g. viruses, animals, plants, molecules) residing in an environment consisting of n spatial locations or patches. If this population did not disperse across the environment, $x = (x_1, \dots, x_n)^{\mathrm{t}}$ denotes the vector of abundances (e.g. density, concentration, or expected size) and d_i is the per-capita growth rate of the population in the ith patch, then a matrix model of this population is given by

$$x(t+1) = Dx(t),$$

where $x(t)$ denotes the vector of abundances in the tth time step and $D = \mathrm{diag}(d_1, \dots, d_n)$. Since $x(t) = D^{\mathrm{t}}x(0)$, population i grows asymptotically at a geometric rate if and only if $d_i > 1$. In particular, the entire population $x_1(t) + \cdots + x_n(t)$ grows asymptotically at a geometric rate if and only if $\rho(D) = \max_i d_i > 1$. Now assume the population disperses across the environment after the reproductive or growth phase. More specifically, a fraction of individuals s_{ij} successfully moves from location j to location i. Then the matrix $S = (s_{ij})$ is column substochastic, and the population dynamics become

$$x(t+1) = SDx(t).$$

This dispersing population exhibits growth if and only if $\rho(SD) > 1$. Since $\rho(SD) < \rho(D)$ if not all d_i's are equal and S is irreducible, one can conclude that generically dispersal decreases the asymptotic population growth rate. Moreover, since $\rho(SD) = \rho(DS)$, this conclusion holds whether growth occurs before dispersal, as we have assumed, or growth occurs after dispersal.

Many biological models not only account for spatial structure but also account for stage structure. For example, in ecological models, the population may consist of individuals in different age classes (e.g. juveniles, sub-adults, adults) living in different spatial locations [1]. Similarly, epidemiological models often account for different classes of individuals (e.g. susceptible, exposed, infected, removed) as well as spatial structure [3]. For these multistate models, one can ask the following question:

Question 1.2. Under what conditions does dispersal decrease the asymptotic growth rate of a population?

In other words, when does the analog of Proposition 1.1 hold for these models. To address this question, consider a population with m life stages living in n spatial locations. Let $x_j^i \in [0, \infty)$ denote the abundance of stage i individuals in location j. Then

$$x^i = \begin{pmatrix} x_1^i \\ \vdots \\ x_n^i \end{pmatrix}, \quad x_j = \begin{pmatrix} x_j^1 \\ \vdots \\ x_j^m \end{pmatrix}, \quad \text{and} \quad x = \begin{pmatrix} x^1 \\ \vdots \\ x^m \end{pmatrix}$$

are the vector of abundances in stage i, the vector of abundances in location j, and the vector of all abundances, respectively. Let us assume that the population first goes through a growth phase in which individuals within a location survive, reproduce, and change stages. For each spatial location j, let A_j be an $m \times m$ nonnegative matrix representing the growth dynamics in location j. In the absence of spatial considerations, the population dynamics are given by

$$x_j(t+1) = A_j x_j(t), \quad j = 1, \ldots, n.$$

If P is the permutation matrix such that

$$Px = P \begin{pmatrix} x^1 \\ \vdots \\ x^m \end{pmatrix} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix},$$

then

$$x(t+1) = \mathbf{D}x(t) \quad \text{with } \mathbf{D} = P^{\mathrm{t}}(A_1 \oplus \cdots \oplus A_n)P,$$

where $A_1 \oplus \cdots \oplus A_n$ denotes the block diagonal matrix with diagonal blocks $A_1, \ldots, A_n$. Let $\{E_{11}, E_{12}, \ldots, E_{nn}\}$ be the standard basis for the linear space of $n \times n$ matrices, and let $A \otimes B = (a_{ij}) \otimes B = (a_{ij}B)$ be the Kronecker product of the two matrices $A = (a_{ij})$ and B. Then $\mathbf{D}$ is a block matrix whose blocks are $n \times n$ nonnegative diagonal matrices

$$\mathbf{D} = \sum_{j=1}^{n} A_j \otimes E_{jj}.$$

To account for movement between locations after the growth phase, let $S_1, \ldots, S_m$ be column substochastic $n \times n$ matrices whose (i, j)th entry corresponds to fraction of individuals in a given stage that move successfully from patch j to patch i. Including these spatial movements by setting $\mathbf{S} = S_1 \oplus \cdots \oplus S_m$, we see that the population dynamics become

$$x(t+1) = \mathbf{SD}x(t).$$

Hunter and Caswell [6] discuss alternative representations of the same model. The population exhibits asymptotically geometric growth if and only if $\rho(\mathbf{SD}) > 1$. Unlike the purely spatially structured model, the following example illustrates that the inclusion of spatial movement into a stage-structured population can enhance the asymptotic growth rate of the population.

Example 1.3. Consider a population of juveniles and reproductively mature adults living in two spatial locations (e.g. salmon where juveniles develop in fresh water and adults become reproductively mature in the ocean). For illustrative purposes, let us assume that in location 1 (i.e. a freshwater river), all adults produce two juveniles before dying but juveniles can not become reproductively mature adults. In location 2 (i.e. the ocean), all juveniles become reproductively mature adults but progeny produced by the adults in location 2 can not survive (i.e. salmon fry can not develop in salt water). In other words, $m = n = 2$, $A_1 = \begin{pmatrix} 0 & 2 \\ 0 & 0 \end{pmatrix}$, and $A_2 = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$. In which case,

$$\rho(\mathbf{D}) = \rho\left(\begin{pmatrix} 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}\right) = 0$$

as $\mathbf{D}^2 = 0$: without movement between the patches, the population goes extinct in two time steps. Alternatively, if all juveniles move to patch 2 and all adults move to patch 1, then $S_1 = \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix}$ and $S_2 = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}$. In which case,

$$\rho(\mathbf{SD}) = \rho\left(\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}\right) = \sqrt{2}$$

and the population grows asymptotically at a geometric rate.

Proposition 1.1 and Example 1.3 suggest the following general question:

Question 1.4. Given $\mathbf{S}$ and $\mathbf{D}$ is there a practical way of determining whether or not $\rho(\mathbf{SD}) \leqslant \rho(\mathbf{D})$?

As a step to understanding this question, we provide in Section 2 an affirmative answer to the following question:

Question 1.5. Given $\mathbf{S}$ is there a practical way of determining whether $\rho(\mathbf{SD}) \leqslant \rho(\mathbf{D})$ for all block matrices $\mathbf{D}$ whose blocks are $m \times m$ nonnegative diagonal matrices?

In particular, if one does not have much information about $\mathbf{D}$, and yet one would like to control or change $\mathbf{S}$ to ensure that $\rho(\mathbf{SD}) \leqslant \rho(\mathbf{D})$, our result provides useful information.

Our paper is organized as follows. In Section 2, we present the statement of our main theorem (Theorem 2.1) answering Question 1.5, and illustrate applications of the result to biological models studied by other authors. In Section 3, we give a proof of Theorem 2.1.

2. Statement of theorem and applications

To state our theorem, we need to introduce one definition. We say the $n \times n$ nonnegative matrices $S_1, \ldots, S_m$ admit a *polychromatic cycle* if there exist nonzero entries of $S_1 + \cdots + S_m$ at the $(i_1, i_2), (i_2, i_3) \ldots, (i_{k-1}, i_k), (i_k, i_1)$ positions for some distinct $i_1, \ldots, i_k \in \{1, \ldots, n\}$ and these nonzero entries do not come from a single matrix S_j. One can think of this definition as follows. Let G_j be the directed graph corresponding to the matrix S_j for $j = 1, \ldots, m$. In other words, G_j has vertex set $V(G_j) = \{1, \ldots, n\}$, and the arc set $E(G_j)$ consists of arcs (r, s) from vertex r to vertex s is the (r, s) entry of S_j is nonzero. Here we do not consider self loops. Consider the multiple arc directed graph G consisting of the directed graphs $G_1, \ldots, G_m$, where each directed graph is drawn in a different color. A polychromatic cycle is a (directed) cycle in G that includes arcs of more than one color. For instance, in Example 1.3, the entries $(2, 1)$ and $(1, 2)$ of S_1 and S_2, respectively, define a polychromatic cycle for the matrices S_1 and S_2.

Theorem 2.1. *Suppose* $\mathbf{S} = S_1 \oplus \cdots \oplus S_m$ *so that* $S_1, \ldots, S_m$ *are* $n \times n$ *column substochastic matrices. The following conditions are equivalent*:

(a) *For every block matrix* $\mathbf{D} = (D_{ij})_{1 \leqslant i,j \leqslant m}$, *where each* $D_{ij} \in M_n$ *is a diagonal matrix with nonnegative diagonal entries*, *we have* $\rho(\mathbf{SD}) \leqslant \rho(\mathbf{D})$.
(b) $S_1, \ldots, S_m$ *do not admit a polychromatic cycle*.

The proof of Theorem 2.1 will be given in Section 3. Here we provide two applications to multistate matrix models studied by other researchers.

Example 2.2 (*Patch development models*). Many species live in environments where the patches change state stochastically in time. For instance, a patch of land may be recently disturbed by a fire or have been undisturbed for an extended period of time. Alternatively, a patch of land may have recently experienced a rainfall or be going through a dry spell. Following Horvitz and Schemske [5], Caswell [1, Example 4.3.1.5 on p. 70] describes matrix models that account for changes in the state of a patch. These models assume that each patch exhibits transitions between n different states and individuals living in these patches can be in one of m stages, where $m \geqslant 2$. Let x^i_j denote the abundance of individuals of state j living in a patch in stage i. Let S be a column stochastic matrix that represents the transition probabilities between patch states i.e. s_{ij} is the probability that the patch goes from state j to state i in one time step. Let $A_1, \ldots, A_n$ be $m \times m$ nonnegative matrices that represent the population dynamics (i.e. transitions between life stages and reproduction) for a patch in states $1, \ldots, n$. If we define

$$\mathbf{S} = I_m \otimes S \quad \text{and} \quad \mathbf{D} = \sum_{j=1}^{n} A_j \otimes E_{jj}$$

then the population dynamics are given by $x(t+1) = \mathbf{SD}x(t)$. Since $\mathbf{S}$ admits a polychromatic cycle if and only if S has a cycle, Theorem 2.1 implies that $\rho(\mathbf{SD}) \leqslant \rho(\mathbf{D})$ for all $\mathbf{D}$ if and only if S admits no cycle.

Example 2.3 (*Planktonic dynamics*). Many species have a single life stage that disperses through the environment [6–8]. For matrix models of these species, Theorem 2.1 typically implies that dispersal decreases the asymptotic population growth rate. Sometimes this application of Theorem 2.1 requires augmenting the matrix model by additional state variables. For instance, Caswell [1, Example 4.3.1.3 on p. 68] describes a stage structured and spatially structured model of Davis [2] for planktonic species (e.g. copepods, water fleas, etc.) dispersing in ocean currents during their larval stage. This model assumes that the growth dynamics are independent of spatial location and determined by two nonnegative matrices: a nonnegative $m \times m$ matrix $T = (t_{ij})$ that represents the transitions between life stages and a nonnegative $m \times m$ matrix $F = (f_{ij})$ that represents reproduction. Since reproduction only contributes to first life stage (i.e. the larval stage), F has all zero entries except in the first row. To describe dispersal between locations by the larvae, let S be an $n \times n$ column substochastic matrix whose (i, j)th entry s_{ij} corresponds to the likelihood that a newly born larva from location j ends up in location i. Then, the model for the planktonic dynamics is given by

$$x_i(t+1) = Tx_i(t) + \sum_{j=1}^{n} s_{ij} F x_j(t), \quad i = 1, \ldots, n,$$

where the first term corresponds to transitions between life stages and the second term corresponds to new larvae dispersing to an ith patch. Equivalently,

$$x(t+1) = \mathbf{A}x(t), \quad \text{where } \mathbf{A} = T \otimes I_n + (S \oplus (I_{m-1} \otimes I_n))(F \otimes I_n). \tag{2.1}$$

In particular, if the planktonic larvae do not disperse between spatial locations, then the matrices S and $\mathbf{A}$ reduce to I_n and $(T + F) \otimes I_n$, respectively. We claim that

$$\rho(T \otimes I_n + (S \oplus (I_{m-1} \otimes I_n))(F \otimes I)) \leqslant \rho((T + F) \otimes I_n) \tag{2.2}$$

for any column substochastic matrix S. In other words, dispersal of the larva reduces the asymptotic growth rate of the population.

To prove this claim using Theorem 2.1, we need to introduce an extra variable $\tilde{x}_j^0$ that keeps track of the newly born larval stage. Let $\tilde{x}_j^i$ for $i = 0, 1, \ldots, m$ and $j = 1, \ldots, n$ correspond to the abundance of life stage i in location j, $\tilde{x}^i = \begin{pmatrix} \tilde{x}_1^i \\ \vdots \\ \tilde{x}_n^i \end{pmatrix}$, and $\tilde{x} = \begin{pmatrix} \tilde{x}^0 \\ \vdots \\ \tilde{x}^m \end{pmatrix}$. Moreover, define

$$\mathbf{D} = \begin{pmatrix} f_{11} & f_{11} & f_{12} & \cdots & f_{1m} \\ t_{11} & t_{11} & t_{12} & \cdots & t_{1m} \\ t_{21} & t_{21} & t_{22} & \cdots & t_{2m} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ t_{m1} & t_{m1} & t_{m2} & \cdots & t_{mm} \end{pmatrix} \otimes I_n \quad \text{and} \quad \mathbf{S} = S \oplus (I_m \otimes I_n).$$

With the inclusion of this additional variable, the planktonic model in (2.1) becomes

$$\tilde{x}(t + 1) = \mathbf{SD}\tilde{x}(t).$$

Furthermore, we have

$$\begin{pmatrix} I_n & I_n & 0_{n,n(m-1)} \\ 0_{n(m-1),n} & 0_{n(m-1),n} & I_{m-1} \otimes I_n \end{pmatrix} \mathbf{SD} = \mathbf{A} \begin{pmatrix} I_n & I_n & 0_{n,n(m-1)} \\ 0_{n(m-1),n} & 0_{n(m-1),n} & I_{m-1} \otimes I_n \end{pmatrix}.$$

Hence, if $\tilde{x} = \begin{pmatrix} \tilde{x}^0 \\ \tilde{x}^1 \\ \vdots \\ \tilde{x}^m \end{pmatrix}$ is a nonnegative right eigenvector of $\mathbf{SD}$, then $x = \begin{pmatrix} \tilde{x}^0 + \tilde{x}^1 \\ \tilde{x}^2 \\ \vdots \\ \tilde{x}^m \end{pmatrix} \neq 0$ is a nonnegative right eigenvector of $\mathbf{A}$ with the same eigenvalue. Conversely, if $y = (y_1, \ldots, y_n)$ is a left eigenvector of $\mathbf{A}$, then $\widetilde{y} = (y_1, y_1, y_2, \ldots, y_n) \neq 0$ is a left eigenvector of $\mathbf{SD}$ with the same eigenvalue. It follows that that $\rho(\mathbf{SD}) = \rho(\mathbf{A})$. Evidently, $\mathbf{S}$ does not admit a polychromatic cycle. By Theorem 2.1, we have $\rho(\mathbf{A}) = \rho(\mathbf{SD}) \leqslant \rho(\mathbf{D})$ and thus (2.2) holds for any column substochastic matrix S.

3. Proof of Theorem 2.1

In this section, we give a proof of Theorem 2.1. Because our proof contains some intricate combinatorial arguments and constructions, we give several examples at the end of this section to illustrate the ideas and constructions in our proofs. In particular, Example 3.1 illustrates the idea and construction in the proof of (a) ⇒ (b), and the other examples illustrate the ideas and constructions in the proof of (b) ⇒ (a). Readers may study the examples along with the proofs to gain better insight.

Throughout this section, we will assume that P is the permutation matrix such that

$$Px = P\begin{pmatrix} x^1 \\ \vdots \\ x^m \end{pmatrix} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}. \tag{3.1}$$

We often use the notation $P\mathbf{S}P^{\mathrm{t}} = \widetilde{\mathbf{S}}$ and $P\mathbf{D}P^{\mathrm{t}} = \widetilde{\mathbf{D}}$. Note that $\widetilde{\mathbf{D}}$ will be a direct sum of n matrices of order $m \times m$; and $\widetilde{\mathbf{S}}$ will be an $n \times n$ block matrix such that each block is an $m \times m$ diagonal matrix.

Also, we will continue to use the graph theory notation introduced at the beginning of Section 2. Note that if G_j is the directed graph of S_j, and if we relabel the vertices of G_j, it is the same as replacing S_j by $Q^{\mathrm{t}}S_jQ$ for a suitable $n \times n$ permutation matrix Q. Note that replacing S_j by $Q^{\mathrm{t}}S_jQ$ for all $j = 1, \ldots, m$, is the same as replacing $\mathbf{S}$ by $(I_m \otimes Q)^{\mathrm{t}}\mathbf{S}(I_m \otimes Q)$. If we also replace each $\mathbf{D}$ by $(I_m \otimes Q)^{\mathrm{t}}\mathbf{D}(I_m \otimes Q)$, conditions (a) and (b) will not be affected.

We will write $X \geqslant Y$ if $X - Y$ is nonnegative. Clearly, if X, Y and Z are nonnegative matrices satisfying $X \geqslant Y$, then $XZ \geqslant YZ$ and hence $\rho(XZ) \geqslant \rho(YZ)$; see [4, Theorem 8.4.5].

The proof of (a) ⇒ (b): We prove the contrapositive. Suppose $\neg(b)$ holds, i.e., $S_1, \ldots, S_m$ admit a polychromatic cycle. We will construct an $m \times m$ block matrix $\mathbf{D}$ such that $\rho(\mathbf{D}) > \rho(\mathbf{SD})$ and each block is a nonnegative diagonal matrix.

Assume $S_1, \ldots, S_m$ admit a polychromatic cycle $(i_1, \ldots, i_k)$. Replacing $\mathbf{S}$ and $\mathbf{D}$ by $(I \otimes Q)^{\mathrm{t}}\mathbf{S}(I \otimes Q)$ and $(I_m \otimes Q)^{\mathrm{t}}\mathbf{D}(I_m \otimes Q)$ with a suitable permutation matrix Q, we may assume that $(i_1, \ldots, i_k) = (1, \ldots, k)$. Then G has a polychromatic cycle with arcs $(1, 2), (2, 3), \ldots, (k, 1)$. For $i = 1, \ldots, m$, construct the $n \times n$ zero-one matrix F_i from S_i by changing those entries of S_i which contribute to the arcs $(1, 2), (2, 3), \ldots, (k, 1)$ of the polychromatic cycle to one, and changing all other entries to zero. Let $\mathbf{F} = d(F_1 \oplus \cdots \oplus F_m)$, where

$$d = \min\{s : s \text{ is a positive entry of } \mathbf{S}\}.$$

Then $\mathbf{F}$ has k nonzero entries all equal to d and $\mathbf{S} \geqslant \mathbf{F}$. We will construct $\mathbf{D}$ so that

$$\rho(\mathbf{FD}) = \rho(\mathbf{DF}) > \rho(\mathbf{D}). \tag{3.2}$$

Since $\mathbf{S} \geqslant \mathbf{F}$, it will then follow that $\rho(\mathbf{SD}) \geqslant \rho(\mathbf{FD}) > \rho(\mathbf{D})$.

Let P be the permutation matrix satisfying (3.1), and let $\widetilde{\mathbf{F}} = P\mathbf{F}P^{\mathrm{t}} = d(\widetilde{F}_{ij})_{1 \leqslant i,j \leqslant n}$, where $\widetilde{F}_{ij}$ are $m \times m$ diagonal matrices. Then $\widetilde{F}_{ij} = 0$ if $(i, j) \notin \{(1, 2), (2, 3), \ldots, (k - 1, k), (k, 1)\}$. Since the nonzero entries in $F_1 + \cdots + F_m$ do not come from a single matrix F_r, the matrices $\widetilde{F}_{12}, \ldots, \widetilde{F}_{k-1,k}, \widetilde{F}_{k,1}$ cannot be identical and each of them has only one nonzero entry equal to 1 on the diagonal.

Let $\{e_1, \ldots, e_m\}$ be the standard basis for $\mathbb{R}^m$, and let $\mu > \max\{d^{-k}, 1\}$. Suppose

$$\widetilde{\mathbf{D}} = \widetilde{D}_1 \oplus \cdots \oplus \widetilde{D}_n,$$

where

(i) $\widetilde{D}_j = \left[e_r + \mu \sum_{s \neq r} e_s\right] e_r^{\mathrm{t}}$ if the (r, r) entry of $\widetilde{F}_{j,j+1}$ is nonzero for $j \in \{1, \ldots, k - 1\}$,
(ii) $\widetilde{D}_k = \left[e_r + \mu \sum_{s \neq r} e_s\right] e_r^{\mathrm{t}}$ if the (r, r) entry of $\widetilde{F}_{k,1}$ is nonzero,
(iii) $\widetilde{D}_j = 0_m$ for $j \in \{k + 1, \ldots, n\}$.

Then $\rho(\widetilde{\mathbf{D}}) = 1$, and $\widetilde{\mathbf{D}}\widetilde{\mathbf{F}} = (R_{ij})$, where

$$(R_{12}, R_{23}, \ldots, R_{k-1,k}, R_{k,1}) = (d\widetilde{D}_1, \cdots, d\widetilde{D}_k)$$

and $R_{ij} = 0$ for all other (i, j). Thus,

$$(\widetilde{\mathbf{D}}\widetilde{\mathbf{F}})^k = T_1 \oplus \cdots \oplus T_k \oplus 0_{m(n-k)},$$

where each T_j is a cyclic product of $d\widetilde{D}_1, \ldots, d\widetilde{D}_k$ so that all T_j have the same eigenvalues. Let $\widetilde{D}_i = \left[e_{r_i} + \mu \sum_{s \neq r_i} e_s\right] e^{\mathrm{t}}_{r_i}$ for $i = 1, \ldots, k$. Then

$$\widetilde{D}_1 \cdots \widetilde{D}_k = (\nu_1 \cdots \nu_{n-1}) \left[e_{r_1} + \mu \sum_{s \neq r_1} e_s \right] e^{\mathrm{t}}_{r_n},$$

where for $j = 1, \ldots, k-1$,

$$\nu_j = e^{\mathrm{t}}_{r_j} \left[e_{r_{j+1}} + \mu \sum_{s \neq r_{j+1}} e_s \right] = \begin{cases} 1 & \text{if } r_j = r_{j+1}, \\ \mu & \text{if } r_j \neq r_{j+1}. \end{cases}$$

Since $\widetilde{F}_{12}, \ldots, \widetilde{F}_{k1}$ are not identical, the matrices $\widetilde{D}_1, \ldots, \widetilde{D}_k$ are not identical. So, there is $j \in \{1, \ldots, k-1\}$ such that $\nu_j = \mu$. As a result, the matrix $T_1 = d^k(\widetilde{D}_1 \cdots \widetilde{D}_k)$ has exactly one nonzero column which contains a diagonal entry of the form $d^k \mu^s$ for some $s \geqslant 1$. It follows that $\rho((\widetilde{\mathbf{D}}\widetilde{\mathbf{F}})^k) = \rho(T_1) = d^k \mu^s > 1$ because $\mu > \max\{d^{-k}, 1\}$ and $s \geqslant 1$. Let $\mathbf{D} = P^{\mathrm{t}} \widetilde{\mathbf{D}} P$. Then

$$\rho(\mathbf{DF}) = \rho(\widetilde{\mathbf{D}}\widetilde{\mathbf{F}}) > 1 = \rho(\widetilde{\mathbf{D}}) = \rho(\mathbf{D}).$$

The proof of (b) ⇒ (a): Suppose $\mathbf{S}$ satisfy condition (b). Using the graph theory description of condition (b) before Theorem 2.1, we see that if G_i is the directed graph associated with S_i for $i = 1, \ldots, m$, then in the multiple arc directed graph G with vertex set $V(G) = \{1, \ldots, n\}$ and multiple arc set $E(G) = \cup_{j=1}^m E(G_j)$ every cycle is monochromatic. We will show that

$$\rho(\mathbf{SD}) \leqslant \rho(\mathbf{D}) \tag{3.3}$$

for all $m \times m$ block matrices $\mathbf{D}$ in which each block is a diagonal matrix with positive diagonal entries. By continuity, the inequality (3.3) will also hold for $m \times m$ block matrices $\mathbf{D}$ such that each block is a nonnegative diagonal matrix.

To prove inequality (3.3), we can impose some additional assumptions on the matrix $\mathbf{D}$ as follows. Suppose $P\mathbf{D}P^{\mathrm{t}} = \widetilde{\mathbf{D}} = D_1 \oplus \cdots \oplus D_n$. We can replace $\mathbf{D}$ by $\mathbf{D}/\rho(\mathbf{D})$ on both sides of (3.3) and assume that $\rho(\mathbf{D}) = 1$. Furthermore, we can replace D_j by $D_j/\rho(D_j)$ for each $j = 1, \ldots, n$, and assume that $\rho(D_1) = \cdots = \rho(D_n) = 1$. Note that after such a replacement, $\rho(\mathbf{D})$ will stay the same, but $\rho(\mathbf{SD})$ may increase.

Now, suppose $\mathbf{D}$ is an $m \times m$ block matrix such that each block is an $n \times n$ diagonal matrix with positive diagonal entries, and $P\mathbf{D}P^{\mathrm{t}} = \widetilde{\mathbf{D}} = D_1 \oplus \cdots \oplus D_n$ with $\rho(D_j) = 1$ for each $j = 1, \ldots, n$. Our strategy is to show that there exists a diagonal matrix $\mathbf{V}$ with positive diagonal entries such that $\mathbf{VDV}^{-1}$ is column stochastic and $\mathbf{VSV}^{-1}$ has column sum norm at most one. It will then follow that:

$$\rho(\mathbf{SD}) = \rho(\mathbf{VSV}^{-1}\mathbf{VDV}^{-1}) \leqslant \|\mathbf{VSV}^{-1}\mathbf{VDV}^{-1}\| \leqslant \|\mathbf{VSV}^{-1}\|\|\mathbf{VDV}^{-1}\| \leqslant 1 = \rho(\mathbf{D}).$$

To achieve our goal, let Q be an $n \times n$ permutation matrix such that

$$Q^{\mathrm{t}}(S_1 + \cdots + S_m)Q = (T_{ij})_{1 \leqslant i,j \leqslant k} \tag{3.4}$$

is in block upper triangular form (Frobenius normal form) and each diagonal block T_{jj} is an irreducible square matrix. We may assume that $Q = I_n$; otherwise, replace $\mathbf{S}$ by $(I_m \otimes Q)^{\mathrm{t}}\mathbf{S}(I_m \otimes Q)$, and replace $\mathbf{D}$ by $(I_m \otimes Q)^{\mathrm{t}}\mathbf{D}(I_m \otimes Q)$, accordingly.

Suppose $C_1, \ldots, C_k$ are the strongly connected components of the directed graph G with multiple arcs corresponding to the matrices $T_{11}, \ldots, T_{kk}$. For $r \in \{1, \ldots, m\}$, if (p, q) is an arc of G_r in C_i, then (p, q) belongs to a cycle in G_r lying entirely in C_i. Otherwise, the arc (p, q) will lie in a polychromatic cycle in the strongly connected component C_i. So, each arc of G_r in C_i must belong to a strongly connected component of G_r lying entirely in C_i. Consequently, every C_i is a directed graph consisting of strongly connected *sub-components* $C_{i,1}, \ldots, C_{i,p_i}$ from the directed graphs $G_1, \ldots, G_m$.

To construct the desired diagonal matrix $\mathbf{V}$, we need to relabel the vertices of G. We do this in two steps.

Step 1: We relabel the sub-components $C_{i,1}, \ldots, C_{i,p_i}$ of C_i for each i as follows.

If $p_i = 1$, then no relabeling is needed. If $p_i > 1$ then $C_{i,1}$ has a common vertex with one of directed graphs $C_{i,2}, \ldots, C_{i,p_i}$. Since C_1 is strongly connected, permuting the subscripts $2, \ldots, p_i$ if necessary, we may assume that $C_{i,1}$ and $C_{i,2}$ have common vertices. Observe that $C_{i,1}$ and $C_{i,2}$ cannot have two or more common vertices; otherwise, there will be a two-color cycle in C_i. If $p_i > 2$, then $C_{i,1} \cup C_{i,2}$ will have a common vertex with one of the sub-components $C_{i,3}, \ldots, C_{i,p_i}$, say, $C_{i,3}$. Here the union of two directed graphs means the union of vertex sets as well as the arc sets of the two directed graphs. Again, $C_{i,1} \cup C_{i,2}$ and $C_{i,3}$ cannot have two common vertices; otherwise, there will be a polychromatic cycle in C_i. We can repeat this argument until we are done with C_{i,p_i}.

Step 2: We relabel the vertex set $V(G) = \{1, \ldots, n\}$ as follows.

Assume that C_1 is a union of $C_{1,1}, C_{1,2}, \ldots, C_{1,p_1}$, and C_1 has q_1 vertices. Relabel these vertices by the indices $1, \ldots, q_1$ as follows. Arrange the vertices of $C_{1,1}$ in a sequence (in any order); then continue with the vertices in $C_{1,2}$, till we get to the vertices of C_{1,p_1}. Assume C_2 is a union of $C_{2,1}, \ldots, C_{2,p_2}$ and has q_2 vertices. Then use a similar procedure to relabel the vertices of C_2 by the indices $q_1 + 1, \ldots, q_1 + q_2$. Continue this process till we are done with C_k.

Suppose Q is the $n \times n$ permutation matrix corresponding to the relabeling of vertices of G in Step 2. Again, we may assume that $Q = I_n$. Otherwise, replace $\mathbf{S}$ by $(I_m \otimes Q)^{\mathrm{t}}\mathbf{S}(I_m \otimes Q)$, and replace $\mathbf{D}$ by $(I_m \otimes Q)^{\mathrm{t}}\mathbf{D}(I_m \otimes Q)$ accordingly.

We are now ready to construct the desired diagonal matrix $\mathbf{V}$. Since D_j is a positive matrix and $\rho(D_j) = 1$, there is a left Perron vector $\mathbf{v}_j = (v_{j1}, \ldots, v_{jm})$ with positive entries such that $v_{j1} = 1$ and $\mathbf{v}_j D_j = \mathbf{v}_j$ for $j = 1, \ldots, n$; see [4, Theorem 8.2.11]. If V_j is the diagonal matrix with diagonal entries $v_{j1}, \ldots, v_{jm}$, then $V_j D_j V_j^{-1}$ is column stochastic. Let $\mu_1, \ldots, \mu_n$ be positive numbers, and $\mathbf{V}$ be such that $P\mathbf{V}P^{\mathrm{t}} = \mu_1\ V_1 \oplus \cdots \oplus \mu_n V_n$. Then

$$P(\mathbf{V}\mathbf{D}\mathbf{V}^{-1})P^{\mathrm{t}} = (V_1 D_1 V_1^{-1} \oplus \cdots \oplus V_n D_n V_n^{-1})$$

is column stochastic. Consider the diagonal entries of $\mathbf{V}$ arranged as follows:

$$\begin{array}{cccc} \mu_1 v_{11} & \mu_1 v_{12} & \cdots & \mu_1 v_{1m} \\ \mu_2 v_{21} & \mu_2 v_{22} & \cdots & \mu_2 v_{2m} \\ \vdots & \vdots & \cdots & \vdots \\ \mu_n v_{n1} & \mu_n v_{n2} & \cdots & \mu_n v_{nm}. \end{array} \tag{3.5}$$

If $U_i = \mathrm{diag}(\mu_1 v_{1i}, \mu_2 v_{2i}, \ldots, \mu_n v_{ni})$ for $i = 1, \ldots, m$, then $\mathbf{V} = U_1 \oplus \cdots \oplus U_m$. We will select positive numbers $\mu_1, \ldots, \mu_n$ so that the column sum norm of $U_i S_i U_i^{-1}$ is at most one for each i. Then $\mathbf{V}$ will satisfy the desired property.

If C_1 has only one vertex, set $\mu_1 = 1$. Otherwise, suppose the sub-component $C_{1,1}$ of C_1 is a strongly connected component of G_r and has α_1 vertices. Then in the matrix $\mathbf{S} = S_1 \oplus \cdots \oplus S_m$, only S_r has a leading $\alpha_1 \times \alpha_1$ irreducible principal submatrix, and all other S_j will have a diagonal leading $\alpha_1 \times \alpha_1$ principal submatrix. For $j = 1, \ldots, \alpha_1$, select μ_j so that $\mu_j v_{jr} = 1$. Then for any choices of other μ_j for $j > \alpha_1$, if $\mathbf{VSV}^{-1} = \widehat{S}_1 \oplus \widehat{S}_2 \oplus \cdots \oplus \widehat{S}_m$ then the $\alpha_1 \times \alpha_1$ leading submatrix of $\widehat{S}_j$ will be the same as that of S_j for $j = 1, \ldots, m$. Note that at this point, we have selected the first α_1 rows in list (3.5).

Next, consider the sub-component $C_{1,2}$ in C_1. By the discussion in the Step 1 relabeling procedure, $C_{1,1}$ and $C_{1,2}$ has exactly one common vertex say, $s \in \{1, \ldots, \alpha_1\}$. Suppose $C_{1,2}$ comes from $G_t \neq G_r$ and has $\alpha_2 - \alpha_1 + 1$ vertices (including s). Then $\mu_s v_{st}$ is determined in the preceding paragraph, and we can choose $\mu_{\alpha_1+1}, \ldots, \mu_{\alpha_2}$ so that $\mu_j v_{jt} = \mu_s v_{st}$ for $j = \alpha_1 + 1, \ldots, \alpha_2$. Then for any choices of other μ_j for $j > \alpha_2$, if $\mathbf{VSV}^{-1} = \widehat{S}_1 \oplus \widehat{S}_2 \oplus \cdots \oplus \widehat{S}_m$ then the $\alpha_2 \times \alpha_2$ leading submatrix of $\widehat{S}_j$ will be the same as that of S_j for $j = 1, \ldots, m$. Note that at this point, we have selected the first α_2 rows in list (3.5).

In the labeling of the sub-components of C_1 in the Step 1 relabeling procedure, every additional sub-component would have exactly one common vertex with the union of the previously labeled sub-components. So, we can repeat the above process to select μ_j until we are done with all the sub-components of C_1. Since C_1 has q_1 vertices, for any choices of other μ_j with $j > q_1$, if $\mathbf{VSV}^{-1} = \widehat{S}_1 \oplus \widehat{S}_2 \oplus \cdots \oplus \widehat{S}_m$ then the $q_1 \times q_1$ leading submatrix of $\widehat{S}_j$ will be the same as that of S_j for $j = 1, \ldots, m$. Note that at this point, we have selected the first q_1 rows in list (3.5).

Now we move to the second strongly connected component C_2 of G. Let us identify the smallest constant η such that

$$\mu_i v_{i\ell} \leqslant \eta v_{j\ell} \quad \text{if } j > q_1, \text{ and } \ell \in \{1, \ldots, n\} \tag{3.6}$$

for each number $\mu_i v_{i\ell}$ in the first q_1 rows of list (3.5). In the future selection of μ_j for $j > q_1$, we will insist that $\mu_j \geqslant \eta$. Then

$$(\mu_i v_{i\ell})(\mu_j v_{j\ell})^{-1} \leqslant 1 \quad \text{for all } i \leqslant q_1 < j, \text{ and } \ell \in \{1, \ldots, n\}. \tag{3.7}$$

In other words, all diagonal entries of $(\mu_i V_i)(\mu_j V_j)^{-1}$ are less than or equal to one for $i \leqslant q_1 \leqslant j$. If C_2 has only one vertex, set μ_{q_1+1} to be any number larger than η. Otherwise, consider the first sub-component $C_{2,1}$ in C_2. Assume $C_{2,1}$ belongs to G_ℓ and has β_1 vertices. Then for $j = q_1 + 1, \ldots, q_1 + \beta_1$, choose $\mu_j \geqslant \eta$ so that $\mu_j v_{j\ell}$ are all equal. Then for any choices of other μ_j with $j > q_1 + \beta_2$, if $\mathbf{VSV}^{-1} = \widehat{S}_1 \oplus \widehat{S}_2 \oplus \cdots \oplus \widehat{S}_m$ the $q_1 \times q_1$ leading submatrix and the following $\beta_1 \times \beta_1$ principal submatrix of $\widehat{S}_j$ will be the same as that of S_j for $j = 1, \ldots, m$. Note that at this point, we have selected the first $q_1 + \beta_1$ rows in list (3.5).

Applying similar arguments as those to C_1 with the precaution that $\mu_j \geqslant \eta$ for $j = q_1 + 1, \ldots, q_1 + q_2$, we can select the first $q_1 + q_2$ rows of list (3.5) so that for any choices of other μ_j with $j > q_1 + q_2$, if $\mathbf{VSV}^{-1} = \widehat{S}_1 \oplus \widehat{S}_2 \oplus \cdots \oplus \widehat{S}_m$ then the $q_1 \times q_1$ leading submatrix and the following $q_2 \times q_2$ principal submatrix of $\widehat{S}_j$ will be the same as that of S_j for $j = 1, \ldots, m$.

Note that each S_j has an upper triangular block form according to the Frobenius normal form (T_{ij}) in (3.4). If $\mathbf{VSV}^{-1} = \widehat{S}_1 \oplus \widehat{S}_2 \oplus \cdots \oplus \widehat{S}_m$ and there is a nonzero (i, j) entry ξ in S_ℓ such that $1 \leqslant i \leqslant q_1 < j \leqslant q_1 + q_2$, then it is obtained from the original entry by multiplying the quantity $(\mu_i v_{i\ell})(\mu_j v_{j\ell})^{-1}$, and hence it is not larger than the original entry by (3.7). So, the first $q_1 + q_2$ columns of $\mathbf{VSV}^{-1}$ have column sums bounded above by one.

Now, update η in (3.6) so that

$$\mu_i v_{i\ell} \leqslant \eta v_{j\ell} \quad \text{if } i \leqslant q_1 + q_2 < j, \ \ell \in \{1, \ldots, n\}$$

for all numbers $\mu_i v_{i\ell}$ in the first $q_1 + q_2$ rows of list (3.5). Then we can proceed to consider the strongly connected component C_3 of G and determine μ_j for $j = q_1 + q_2 + 1, \ldots, q_1 + q_2 + q_3$.

Repeating the above argument, we can determine $\mu_1, \ldots, \mu_n$ and construct the diagonal matrix $\mathbf{V} = U_1 \oplus \cdots \oplus U_m$ so that $U_i S_i U_i^{-1}$ is still in upper block triangular form, whose diagonal blocks are the same as those of S_i, and $S_i - U_i S_i U_i^{-1}$ is nonnegative. Hence,

$$\|\mathbf{V}\mathbf{S}\mathbf{V}^{-1}\| \leqslant \|\mathbf{S}\| \leqslant 1. \qquad \square$$

The following example illustrates the construction in our proof of the implication (a) ⇒ (b).

Example 3.1. Let $\mathbf{S} = S_1 \oplus S_2$, where

$$S_1 = \frac{1}{2}\begin{pmatrix} 2 & 1 & 0 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{pmatrix} \quad \text{and} \quad S_2 = \frac{1}{2}\begin{pmatrix} 0 & 0 & 2 \\ 0 & 2 & 0 \\ 2 & 0 & 0 \end{pmatrix}.$$

Then the directed graph G of $S_1 + S_2$ admits a polychromatic cycle with arcs $(1, 2), (2, 3), (3, 1)$. Let $\mathbf{F} = F_1 \oplus F_2$ with

$$F_1 = \begin{pmatrix} 0 & 1/2 & 0 \\ 0 & 0 & 1/2 \\ 0 & 0 & 0 \end{pmatrix} \quad \text{and} \quad F_2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1/2 & 0 & 0 \end{pmatrix}.$$

Then $\mathbf{S} \geqslant \mathbf{F}$ and $\widetilde{\mathbf{F}} = P\mathbf{F}P^{\mathrm{t}} = (\widetilde{F}_{ij})_{1 \leqslant i,j \leqslant 3}$, where

$$\widetilde{F}_{12} = \widetilde{F}_{23} = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad \widetilde{F}_{31} = \frac{1}{2}\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \quad \text{and} \quad \widetilde{F}_{ij} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \text{ otherwise.}$$

Suppose $\mu > 0$ satisfies $\mu^2 > 8$. Construct $\mathbf{D}$ so that $\widetilde{\mathbf{D}} = P\mathbf{D}P^{\mathrm{t}} = \widetilde{D}_1 \oplus \widetilde{D}_2 \oplus \widetilde{D}_3$ with

$$\widetilde{D}_1 = \widetilde{D}_2 = \begin{pmatrix} 1 & 0 \\ \mu & 0 \end{pmatrix} \quad \text{and} \quad \widetilde{D}_3 = \begin{pmatrix} 0 & \mu \\ 0 & 1 \end{pmatrix}.$$

Then

$$\widetilde{\mathbf{D}}\widetilde{\mathbf{F}} = \frac{1}{2}\begin{pmatrix} 0_2 & \widetilde{D}_1 & 0_2 \\ 0_2 & 0_2 & \widetilde{D}_2 \\ \widetilde{D}_3 & 0_2 & 0_2 \end{pmatrix}$$

and

$$(\widetilde{\mathbf{D}}\widetilde{\mathbf{F}})^3 = \frac{1}{8}\left\{\widetilde{D}_1\widetilde{D}_2\widetilde{D}_3 \oplus \widetilde{D}_2\widetilde{D}_3\widetilde{D}_1 \oplus \widetilde{D}_3\widetilde{D}_1\widetilde{D}_2\right\} = \frac{1}{8}\left\{\begin{pmatrix} 0 & \mu \\ 0 & \mu^2 \end{pmatrix} \oplus \begin{pmatrix} \mu^2 & 0 \\ \mu^3 & 0 \end{pmatrix} \oplus \begin{pmatrix} \mu^2 & 0 \\ \mu & 0 \end{pmatrix}\right\}$$

has spectral radius $\mu^2/8 > 1$, and hence $\rho(\widetilde{\mathbf{D}}\widetilde{\mathbf{F}}) > 1$. Since $\mathbf{S} \geqslant \mathbf{F}$, we have

$$\rho(\mathbf{SD}) \geqslant \rho(\mathbf{FD}) = \rho(\mathbf{DF}) = \rho(\widetilde{\mathbf{D}}\widetilde{\mathbf{F}}) > 1 = \rho(\mathbf{D}).$$

The next three examples illustrate the idea and construction in our proof for implication (b) ⇒ (a).

Example 3.2. Suppose $\mathbf{S} = S_1 \oplus S_2 \oplus S_3$, and Q is a permutation matrix such that $Q^{\mathrm{t}} S_1 Q$, $Q^{\mathrm{t}} S_2 Q$ and $Q^{\mathrm{t}} S_3 Q$, are the column stochastic matrices

$$\frac{1}{3}\begin{pmatrix}3&1&1&1&1\\0&1&1&0&0\\0&1&1&0&0\\0&0&0&2&0\\0&0&0&0&2\end{pmatrix},\quad \frac{1}{3}\begin{pmatrix}3&1&1&1&1\\0&1&0&1&0\\0&0&2&0&0\\0&1&0&1&0\\0&0&0&0&2\end{pmatrix},\quad \frac{1}{3}\begin{pmatrix}3&1&1&1&1\\0&1&0&0&1\\0&0&2&0&0\\0&0&0&2&0\\0&1&0&0&1\end{pmatrix}.$$

We may relabel the vertices of G and assume that $Q = I_5$. Then G has two strongly connected components C_1 and C_2, where C_1 has vertex set $\{1\}$ and C_2 has vertex set $\{2, 3, 4, 5\}$. Furthermore, C_2 has three sub-components C_{21}, C_{22}, C_{23} with vertex sets $\{2, 3\}$, $\{2, 4\}$, $\{2, 5\}$, respectively. Now, suppose $\mathbf{D}$ is given such that $P^{\mathrm{t}}\mathbf{D}P = D_1 \oplus \cdots \oplus D_5$ with

$$D_1 = \frac{1}{12}\begin{pmatrix}1&3&8\\1&3&8\\1&3&8\end{pmatrix},\quad D_2 = D_3 = \frac{1}{6}\begin{pmatrix}1&2&3\\1&2&3\\1&2&3\end{pmatrix},\quad \text{and}\quad D_4 = D_5 = \frac{1}{4}\begin{pmatrix}1&1&2\\1&1&2\\1&1&2\end{pmatrix}.$$

Then $\rho(D_j) = 1$ for each j; $(1, 3, 8)D_1 = (1, 3, 8)$, $(1, 2, 3)D_j = (1, 2, 3)$ for $j = 2, 3$, $(1, 1, 2)D_j = (1, 1, 2)$ for $j = 4, 5$. Consider the list

$$\begin{array}{lll}\mu_1 1 & \mu_1 3 & \mu_1 8\\ \mu_2 1 & \mu_2 2 & \mu_2 3\\ \mu_3 1 & \mu_3 2 & \mu_3 3\\ \mu_4 1 & \mu_4 1 & \mu_4 2\\ \mu_5 1 & \mu_5 1 & \mu_5 2.\end{array}$$

We follow the construction in the proof and choose $\mu_1 = 1$, and insist that:

$$\mu_j \geqslant \eta = \max\{v_{1j}/v_{\ell j} : 1 \leqslant j \leqslant 3; 2 \leqslant \ell \leqslant 5\} = 4, \quad j = 2, 3, 4, 5.$$

Now, C_{21} comes from G_1 and has vertex set $\{2, 3\}$; so, we choose $\mu_2 = \mu_3 = 4$. Next, C_{22} comes from S_2 and has vertex set $\{2, 4\}$; so, we choose μ_4 such that $\mu_2 v_{22} = \mu_4 v_{42}$, i.e., $\mu_4 = 8$. Finally, C_{23} comes from S_3 and has vertex set $\{2, 5\}$; so, we choose μ_5 such that $\mu_2 v_{32} = \mu_5 v_{35}$, i.e., $\mu_5 = 6$. Consequently, we have

$$\mathbf{V} = \mathrm{diag}(1, 4, 4, 8, 6) \oplus \mathrm{diag}(3, 8, 8, 8, 6) \oplus \mathrm{diag}(8, 12, 12, 16, 12).$$

Then $\mathbf{VDV}^{-1}$ is column stochastic, and $\|\mathbf{VSV}^{-1}\| \leqslant 1$.

Example 3.3. Suppose $\mathbf{S}$ is such that $S_2 = \cdots = S_m = I_n$. Let Q be a permutation matrix such that $Q^{\mathrm{t}}S_1Q$ is in the block upper triangular form. Then $Q^{\mathrm{t}}(S_1 + \cdots + S_m)Q = (T_{ij})_{1\leqslant i,j\leqslant k}$ is in block upper triangular form (3.4) as in our proof. We can relabel the vertices of G and assume that $Q = I_n$. Now, G has k connected components, and each comes from G_1. Suppose T_{jj} is $n_j \times n_j$ for $j = 1, \ldots, k$. Given any $\mathbf{D}$, we can follow the construction in our proof and obtain the desired $\mathbf{V} = U_1 \oplus \cdots \oplus U_m$ such that $U_1 = \mu_1 I_{n_1} \oplus \cdots \oplus \mu_k I_{n_k}$ with $\mu_1 < \cdots < \mu_k$. Then $\mathbf{VDV}^{-1}$ is column stochastic, $U_1S_1U_1^{-1}$ is in block upper triangular form such that $S_1 \geqslant U_1S_1U_1^{-1}$, and $U_iS_iU_i^{-1} = I_n$ for $i = 2, \ldots, m$; thus, $\|\mathbf{VSV}^{-1}\| \leqslant 1$.

Example 3.4. Suppose there is an $n \times n$ permutation matrix Q such that $Q^{\mathrm{t}}(S_1 + \cdots + S_n)Q$ is in upper triangular form. Thus, each connected component of G has only one vertex. We may relabel the vertices of G and assume that $Q = I_n$. Given any $\mathbf{D}$, we can follow the construction in our proof and obtain the desired $\mathbf{V} = U_1 \oplus \cdots \oplus U_m$ such that each U_i is a diagonal matrix with diagonal entries arranged in ascending order. Since each S_i is in upper triangular form, we see that $S_i - U_i S_i U_i^{-1}$ is a nonnegative matrix in strictly upper triangular form. So, $\mathbf{VDV}^{-1}$ is column stochastic and $\|\mathbf{VSV}^{-1}\| \leqslant 1$.

Acknowledgments

Research of the authors were partially supported by NSF. The first author was also supported by a HK RCG grant. The authors would like to thank Professor Eva Czabarka for her helpful comments. They are also indebted to the referee for the thorough reviews.

References

[1] H. Caswell, Matrix Population Models, Sinauer, Sunderland, Massachuesetts, 2001.
[2] C.S. Davis, Interaction of a copepod population with the mean circulation on Georges Bank, J. Marine Res. 42 (1984) 573–590.
[3] O. Diekmann, J.A.P. Heesterbeek, Mathematical epidemiology of infectious diseases, Wiley Series in Mathematical and Computational Biology, John Wiley & Sons Ltd., Chichester, 2000.
[4] R.A. Horn, C.R. Johnson, Matrix Analysis, Cambridge University Press, Cambridge, 1985.
[5] C.C. Horvitz, D.W. Schemske, Chapter: Seed dispersal and environmental heterogeneity in a neotropical herb: a model of population and patch dynamics, Frugivores and seed dispersal, Dr. W. Junk Publishers, Dordrecht, Netherlands, 1986.
[6] C.M. Hunter, H. Caswell, The use of the vec-permutation matrix in spatial matrix population models, Ecol. Model. 188 (2005) 15–21.
[7] J.D. Lebreton, Demographic models for subdivided populations: the renewal equation approach, Theoret. Popul. Biol. 49 (1996) 291–313.
[8] J.T. Wootton, D.A. Bell, A metapopulation model of the peregrine falcon in California: viability and management strategies, Ecol. Appl. 2 (1992) 307–321.