Linear Algebra and its Applications 434 (2011) 1409–1424

Contents lists available at ScienceDirect

Linear Algebra and its Applications

journal homepage: www.elsevier.com/locate/laa

Nonnegative matrix semigroups with finite diagonals

Alexey I. Popov [a,*], Heydar Radjavi [b], Peter Williamson [b]

[a] *Department of Mathematical and Statistical Sciences, University of Alberta, Edmonton, AB, Canada T6G 2G1*
[b] *Department of Pure Mathematics, University of Waterloo, Waterloo, ON, Canada N2L 3G1*

ARTICLE INFO

Article history:
Received 24 March 2010
Accepted 28 June 2010
Available online 4 August 2010

Submitted by P. Šemrl

AMS classification:
15A48
20M20

Keywords:
Nonnegative matrices
Semigroups of matrices
Diagonal entries of matrices
Indecomposability

ABSTRACT

Let $\mathcal{S}$ be a multiplicative semigroup of matrices with nonnegative entries. Assume that the diagonal entries of the members of $\mathcal{S}$ form a finite set. This paper is concerned with the following question: Under what circumstances can we deduce that $\mathcal{S}$ itself is finite?

1. Introduction

It is a simple classical result, found in every book on the representation theory of groups, that if $\mathcal{G}$ is an irreducible group of complex matrices (or matrices over any algebraically closed field) and if $\mathcal{G}$ has finite trace, that is, if $\{\operatorname{tr} A : A \in \mathcal{G}\}$ is a finite set, then $\mathcal{G}$ itself is finite. Here "irreducible" means that there is no nontrivial subspace simultaneously invariant under all the members of $\mathcal{G}$ (viewed as linear operators). This result was extended to semigroups of matrices (i.e., sets of matrices closed under multiplication) in [6]. It was shown in [8] that the result holds if trace is replaced by any linear functional. There are other results available in which the finiteness is replaced by boundedness in the assumption as well as the conclusion [8].

We are interested in semigroups of matrices with nonnegative entries (called nonnegative matrices for short). A natural analogue of irreducibility here is indecomposability, which is a much weaker

* Corresponding author. Tel.: +1 780 432 2064; fax: +1 780 492 6826.
E-mail addresses: apopov@math.ualberta.ca (A.I. Popov), hradjavi@uwaterloo.ca (H. Radjavi), pwilliam@uvic.ca (P. Williamson).

0024-3795/$ - see front matter
doi:10.1016/j.laa.2010.06.049

assumption. A set $\mathcal{S}$ of nonnegative matrices is said to be decomposable if there is a permutation matrix P such that all the members of $\{P^{-1}SP : S \in \mathcal{S}\}$ are in simultaneous block form $\begin{bmatrix} X & Y \\ 0 & Z \end{bmatrix}$. (If we prefer to view the matrices as operators represented relative to the basis $\{e_j\}$, then a set is indecomposable if it has no simultaneous invariant subspace spanned be a nonempty, proper subset of $\{e_j\}$. An equivalent definition of indecomposability for a nonnegative semigroup $\mathcal{S}$ is: for each pair i and j, the (i, j) entry of some member of $\mathcal{S}$ is positive.)

As anybody who has used, say, the Perron–Frobenius theorem, knows, indecomposability is a very useful condition to have, especially in dealing with nonnegative matrices. It would be interesting to get boundedness and finiteness results of the type mentioned above with this weaker condition. One of the results proved in the recent paper [2] is that if $\mathcal{S}$ is an indecomposable semigroup of nonnegative matrices and if for some nonzero positive linear functional ϕ the set $\{\phi(S) : S \in \mathcal{S}\}$ is bounded, then $\mathcal{S}$ itself is bounded.

We are mainly interested in the finiteness analogue of this assumption. It has been shown recently in [4] that if all the diagonal entries of an indecomposable nonnegative semigroup consist of zeros and ones, then the semigroup is finite (and furthermore, all entries are in $\{0, 1\}$ after a suitable diagonal similarity). The indecomposability condition is clearly necessary for this result. For example, the semigroup of all upper-triangular nonnegative matrices whose diagonal elements are all equal to 1 is by no means finite.

In this paper, we ask the question: if all diagonal entries in an indecomposable nonnegative semigroup come from a fixed finite set, is the semigroup itself finite? The following example shows that in general the answer is negative.

Example 1.1. Let

$$\mathcal{S} = \left\{ \begin{bmatrix} E & 0 \\ 0 & E \end{bmatrix}, \begin{bmatrix} 0 & E \\ E & 0 \end{bmatrix}, \begin{bmatrix} 0 & S \\ E & 0 \end{bmatrix} \right\},$$

where $E = \begin{bmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{bmatrix}$ and S runs over the set of all matrices of form $\begin{bmatrix} p & q \\ q & p \end{bmatrix}$, where $p, q \geqslant 0$, $p + q = 1$.

The semigroup in Example 1.1 is indecomposable and is not very far from having only zeros and ones on the diagonals: the set of all the diagonal entries of matrices in $\mathcal{S}$ is $\{0, 1/2\}$. However, $\mathcal{S}$ is not finite, and incidentally, consists of doubly stochastic matrices. (Recall that a nonnegative matrix is said to be *row* (*column*) *stochastic* if each of its rows (columns) sums to 1. A matrix is *doubly stochastic* if it is both row and column stochastic.)

Although the answer in general is negative, we get affirmative results in two significant cases: that of a self-adjoint semigroup (that is a semigroup $\mathcal{S}$ such that $S \in \mathcal{S}$ implies $S^* \in \mathcal{S}$), and that of constant-rank semigroups. We also obtain results about the non-diagonal entries in certain cases.

All semigroups in this paper consist of nonnegative matrices. A semigroup $\mathcal{S}$ of nonnegative matrices will be called a *semigroup with finite diagonals* if all the diagonal entries of all the matrices in $\mathcal{S}$ come from a finite set. We will call $\mathcal{S}$ a *semigroup with finite trace* if the set $\{\mathrm{tr}(S) : S \in \mathcal{S}\}$ is finite.

In Section 2, some useful properties of semigroups with finite diagonals are collected. The main result of Section 3 is Theorem 3.4 stating that if a semigroup with finite diagonals is self-adjoint then it is finite. Our methods also reveal the structure of such semigroups. In Section 4, we show that if all the nonzero matrices in a semigroup with finite diagonals have the same rank then the semigroup is finite. Finally, in Section 5 we characterize possible sets of values for the diagonal elements of semigroups with finite diagonals.

Throughout the paper, the following result, which can be found in [1, Section 3.3] (see also [7, Lemma 5.1.9]) will be used without additional references:

Theorem 1.2. *Let E be a nonnegative idempotent of rank r:*

(i) *If E has no zero rows or columns then there exists a permutation matrix P such that $P^{-1}EP$ has the block-diagonal form*

$$E_1 \oplus \cdots \oplus E_r,$$

where each E_i is an idempotent of rank one whose entries are all positive.

(ii) *In general, there exists a permutation matrix P such that $P^{-1}EP$ has the block-triangular form*

$$E = \begin{bmatrix} 0 & XF & XFY \\ 0 & F & FY \\ 0 & 0 & 0 \end{bmatrix},$$

with square diagonal blocks, where $F = F_1 \oplus \cdots \oplus F_r$ is an idempotent without zero rows or columns as in (i) and X and Y are two nonnegative matrices.

As mentioned above, we will use this result repeated and so for clarity, we call the (2, 2) block, F, of E the *rigid part* of E.

2. Preliminary results

In this section, we will state some auxiliary lemmas that will be important later in the paper and collect some partial solutions of the main problem.

The next two lemmas reveal certain useful properties of the members of semigroups with finite traces.

Lemma 2.1. *Let $\mathcal{S}$ be a semigroup with finite trace. If $S \in \mathcal{S}$ then all the nonzero eigenvalues of S are roots of unity of degree at most the size of matrices in $\mathcal{S}$. In particular, $\rho(S) \leqslant 1$ for all $S \in \mathcal{S}$.*

Proof. By Ref. [5, Proposition 2.2], $\rho(S) \leqslant 1$. Let n be the size of S and $(\lambda_i)_{i=1}^n$ be the sequence of the eigenvalues of S (with multiplicities), ordered by

$$1 = |\lambda_1| = \cdots = |\lambda_k| > |\lambda_{k+1}| \geqslant \cdots \geqslant |\lambda_n| \geqslant 0,$$

where $0 \leqslant k \leqslant n$. By the Perron–Frobenius theorem, the modulus-one eigenvalues of S are roots of unity of degree at most k. It is left to show that $\lambda_{k+1} = \cdots = \lambda_n = 0$.

Since $|\lambda_i| < 1$ for all $i = k+1, \ldots, n$, for each $\varepsilon > 0$ there is $N \in \mathbb{N}$ such that for all $j \geqslant N$

$$\varepsilon > \sum_{i=k+1}^{n} |\lambda_i|^j \geqslant \left| \sum_{i=k+1}^{n} \lambda_i^j \right| \geqslant 0.$$

Therefore, the sequence $\left(\left|\sum_{i=k+1}^{n} \lambda_i^j\right|\right)_{j=1}^{\infty}$ either has a strictly decreasing subsequence or a constant zero tail. If the former were true, the set $\left\{\sum_{i=k+1}^{n} \lambda_i^j : j \in \mathbb{N}\right\}$ would be infinite. However, this set cannot be infinite because $\left\{\sum_{i=1}^{k} \lambda_i^j : j \in \mathbb{N}\right\}$ and $\left\{\sum_{i=1}^{n} \lambda_i^j : j \in \mathbb{N}\right\}$ are both finite. Thus, for some $r \in \mathbb{N}$ we have

$$\sum_{i=k+1}^{n} \lambda_i^{rj} = 0, \quad j \in \mathbb{N}.$$

By Ref. [3] (see also [7, Lemma 2.1.15(ii)]) this implies $\lambda_i^r = 0$, and hence $\lambda_i = 0$ for all $i = k+1, \ldots, n$. □

Lemma 2.2. *Let $\mathcal{S}$ be an indecomposable semigroup with finite trace. Then each $S \in \mathcal{S}$ is similar to a matrix of the form*

$$\begin{bmatrix} U & 0 \\ 0 & N \end{bmatrix},$$

where U is a unitary diagonal matrix and N is a nilpotent matrix.

Proof. Let $S \in \mathcal{S}$ and let

$$\begin{bmatrix} J & 0 \\ 0 & N \end{bmatrix}$$

be the Jordan form of S, where J is an invertible matrix and N is a nilpotent matrix. Write $J = D + M$, where D is a diagonal matrix and the only possible positions of nonzero elements of M are on the super-diagonal.

We claim that $M = 0$. Indeed, suppose $M \neq 0$. By Ref. [2, Proposition 8] the semigroup $\mathcal{S}$ is bounded. Hence so is the set $\{J^m : m \in \mathbb{N}\}$. Let k be such that $M^k \neq 0$ and $M^{k+1} = 0$. Since $DM = MD$ we get

$$J^m = D^m + \binom{m}{1} D^{m-1}M + \cdots + \binom{m}{k} D^{m-k}M^k$$

for all $m \geqslant k$. By Lemma 2.1 all the diagonal entries of D are of absolute value 1, hence $\|D^m\| = 1$ for all m. This implies $\|J^m\| \to \infty$ as $m \to \infty$. Therefore, $M = 0$.

This shows that S is similar to

$$\begin{bmatrix} D & 0 \\ 0 & N \end{bmatrix}.$$

By Lemma 2.1, D is unitary. $\square$

Corollary 2.3. *Let $\mathcal{S}$ be an indecomposable semigroup with finite trace. Then there exists $m \in \mathbb{N}$ such that S^m is an idempotent for each $S \in \mathcal{S}$.*

Proof. Let n be the size of matrices in $\mathcal{S}$. Put $m = n!$. Let $S \in \mathcal{S}$. By Lemma 2.2, S is similar to

$$\begin{bmatrix} U & 0 \\ 0 & N \end{bmatrix},$$

where U is a unitary diagonal matrix and N is a nilpotent matrix. By Lemma 2.1 every diagonal entry of U is a root of unity of degree at most n. Hence S^m is similar to

$$\begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix},$$

where I is an identity matrix. Therefore, S^m is an idempotent. $\square$

In the next lemma which will be used in Section 4 we establish a useful property of idempotents without zero rows and zero columns in semigroups with finite trace. Recall that a collection of matrices is *block-monomial* if each member has only one nonzero block in each block row and block column under a given block structure.

Lemma 2.4. *Let $\mathcal{S}$ be an indecomposable semigroup with finite trace and E an idempotent in $\mathcal{S}$ with no zero rows or columns. Then, after a permutation of the basis which makes E block-diagonal with each diagonal block being a rank-one idempotent matrix with strictly positive entries, the set $\mathcal{S}_E = \{A \in \mathcal{S} : \operatorname{rank}(A) = \operatorname{rank}(E)$ and $EAE = A\}$ is a finite block-monomial group relative to the block structure inherited from E.*

Proof. First, we will show that $\mathcal{S}_E$ is a group with identity E. Indeed, let $A, B \in \mathcal{S}_E$. Then clearly $EABE = AB$. Also, in some basis, E can be represented as $E = \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix}$. Since $EAE = A$ and $EBE = B$, and $\operatorname{rank}(A) = \operatorname{rank}(B) = \operatorname{rank}(E)$, in this basis A and B will be represented as $A = \begin{bmatrix} A_0 & 0 \\ 0 & 0 \end{bmatrix}$ and $B = \begin{bmatrix} B_0 & 0 \\ 0 & 0 \end{bmatrix}$, where A_0 and B_0 are invertible matrices. Then the representation of AB is $AB = \begin{bmatrix} A_0B_0 & 0 \\ 0 & 0 \end{bmatrix}$, and A_0B_0 is again invertible, so that $\operatorname{rank}(AB) = \operatorname{rank}(E)$. This shows that $\mathcal{S}_E$ is a semigroup.

Let us show that $\mathcal{S}_E$ is closed under inverses. Let $A \in \mathcal{S}_E$ be arbitrary. By Corollary 2.3 there is $m \in \mathbb{N}$ such that A^m is an idempotent which we will denote by F. Since $\mathcal{S}_E$ is a semigroup, $F \in \mathcal{S}_E$. In particular,

$\operatorname{rank}(F) = \operatorname{rank}(E)$ and $EFE = F$. Therefore, $E = F$. Thus the matrix A^{m-1} is the inverse of A in $\mathcal{S}_E$, and hence $\mathcal{S}_E$ is a group.

Let $r = \operatorname{rank}(E)$. Applying a permutation, we can write $E = E_1 \oplus \cdots \oplus E_r$, where each E_i is a rank-one idempotent without zero rows or zero columns. Applying a suitable diagonal similarity to $\mathcal{S}$, we can also assume without loss of generality that E is row stochastic. In particular, since the blocks of E have rank one, each block of E is a strictly positive matrix having all rows equal to each other.

Let $K(r, s)$ stand for the $r \times s$ matrix having the value $1/\sqrt{rs}$ at each entry. A straightforward calculation shows that $K(r, s)K(s, t) = K(r, t)$ for all r, s, and $t \in \mathbb{N}$.

For each $i = 1, \dots, r$, denote the size of E_i by r_i. Since each E_i is row stochastic, we have $E_iK(r_i, r_j) = K(r_i, r_j)$ for all $i, j \in \{1, \dots, r\}$. Let $L_{ij} = E_iK(r_i, r_j)E_j$=$K(r_i, r_j)E_j$. Then

$$L_{ij}L_{jk} = E_iK(r_i, r_j)E_jE_jK(r_j, r_k)E_k = E_iK(r_i, r_j)K(r_j, r_k)E_k = E_iK(r_i, r_k)E_k = L_{ik}. \tag{1}$$

Let $A \in \mathcal{S}_E$ be arbitrary. Write A in the block form inherited from E:

$$A = \begin{bmatrix} A_{11} & \dots & A_{1r} \\ \vdots & & \vdots \\ A_{r1} & \dots & A_{rr} \end{bmatrix}.$$

Since $EAE = A$, we get $A_{ij} = E_iA_{ij}E_j$ for all $i, j \in \{1, \dots, r\}$. The ranks of E_i and E_j are equal to 1; thus for each $i, j \in \{1, \dots, r\}$ there exists a nonnegative λ_{ij} such that $A_{ij} = \lambda_{ij}L_{ij}$.

This shows that every matrix $A \in \mathcal{S}_E$ can be represented as a numerical matrix $\widetilde{A} = (\lambda_{ij})_{i,j=1}^{r}$. By formula (1) we also conclude that $\widetilde{AB} = \widetilde{A}\widetilde{B}$. Observe also that $\widetilde{E}$ is the $r \times r$ identity matrix. Therefore, the set $\mathcal{G} = \{\widetilde{A} : A \in \mathcal{S}_E\}$ is a group of nonnegative invertible matrices.

Since $\mathcal{S}$ is an indecomposable semigroup with bounded trace, by [2, Proposition 8] $\mathcal{S}$ itself is bounded. In particular, $\mathcal{S}_E$ is bounded, and hence $\mathcal{G}$ is bounded. Therefore, by Ref. [7, Lemma 5.1.11] $\mathcal{G}$ is a finite monomial group. Hence $\mathcal{S}_E$ is finite and block-monomial. □

The next lemma is a technical statement that allows us to work with the north-west corners of matrices in a semigroup.

Lemma 2.5. *Let $\mathcal{S}$ be an indecomposable semigroup of $N \times N$ matrices. Let $k \in \{1, \dots, N\}$ and*

$$J_k = \{S \in \mathcal{S} : \text{rows } k+1 \text{ through } N \text{ of } S \text{ are zero}\}.$$

Put $\mathcal{S}_k = \{A : A$ is the north-west $k \times k$ corner of some $S \in J_k\}$. If $\mathcal{S}_k$ has no permanent zero rows, that is, if for each $i \in \{1, \dots, k\}$ there is a matrix $A \in \mathcal{S}_k$ such that the ith row of A is not zero, then $\mathcal{S}_k$ is an indecomposable semigroup.

Proof. A straightforward calculation shows that $\mathcal{S}_k$ is a semigroup for each k. We now establish the indecomposability statement.

We need to show that for each $i, j \in \{1, \dots, k\}$ there is a matrix $A \in \mathcal{S}_k$ such that the (i, j) entry of A is different from zero. Pick a matrix $U \in \mathcal{S}_k$ whose ith row is not zero, say $(U)_{im} \neq 0$ for some $m \in \{1, \dots, k\}$. There is a matrix V such that $T := \begin{bmatrix} U & V \\ 0 & 0 \end{bmatrix} \in J_k$. Since $\mathcal{S}$ is indecomposable, there is $S \in \mathcal{S}$ such that $(S)_{mj} \neq 0$. Then $(TS)_{ij} \neq 0$. Also, $TS \in J_k$. Clearly, the north-west $k \times k$ corner of TS has a nonzero (i, j) entry. □

The next lemma is the same statement about the south-east corners of matrices in a semigroup. Its proof is analogous to that of Lemma 2.5, so we omit it.

Lemma 2.6. *Let $\mathcal{S}$ be an indecomposable semigroup of $N \times N$ matrices. Let $k \in \{1, \dots, N\}$ and*

$$J'_k = \{S \in \mathcal{S} : \text{ columns } 1 \text{ through } k \text{ of } S \text{ are zero}\}.$$

Put $\mathcal{S}'_k = \{A : A$ is the south-east $(N-k) \times (N-k)$ corner of some $S \in J'_k\}$. If $\mathcal{S}'_k$ has no permanent zero columns then $\mathcal{S}'_k$ is an indecomposable semigroup.

In the rest of this section we record some simple partial results regarding the main problem.

Theorem 2.7. *Let $\mathcal{S}$ be a commutative finitely generated indecomposable semigroup with finite trace. Then $\mathcal{S}$ is finite.*

Proof. Let $\{A_i\}_{i=1}^n$ be the set of generators of $\mathcal{S}$. By commutativity, each $S \in \mathcal{S}$ can be written as $S = \prod_{i=1}^n A_i^{k_i}$ for some $k_i \geqslant 0$. By Corollary 2.3, there is $m \in \mathbb{N}$ such that $E_i := A_i^m$ is an idempotent for each $i = 1, \ldots, n$. Thus $S = \prod_{i=1}^n A_i^{k_i}$, where $k_i \in \{0, 1, \ldots, 2m-1\}$ and therefore $\mathcal{S}$ is finite. □

Theorem 2.8. *Let $\mathcal{S}$ be an indecomposable semigroup of invertible matrices with finite trace. Then $\mathcal{S}$ is finite and after a diagonal similarity, is in fact a permutation group.*

Proof. First, we prove that $\mathcal{S}$ is actually a group of matrices. Indeed, clearly the only idempotent in $\mathcal{S}$ is the identity matrix. If $S \in \mathcal{S}$ then by Corollary 2.3 there is $m \in \mathbb{N}$ such that $S^m = I$. Then $S^{-1} = S^{m-1} \in \mathcal{S}$.

By Ref. [2, Proposition 8], $\mathcal{S}$ is bounded. By Ref. [7, Lemma 5.1.11], after a diagonal similarity, $\mathcal{S}$ is a permutation group and is thus finite. □

It should be noted that if we replace the condition about the trace in the last theorem with the condition of finiteness of the diagonal entries of the members of $\mathcal{S}$ then Theorem 2.8 becomes a special case of Theorem 4.8 which will be proved in the next section. The following simple example shows that the finiteness of the trace does not in general imply the finiteness of all the diagonal entries.

Example 2.9

$$\mathcal{S} = \left\{ \begin{bmatrix} p & q \\ p & q \end{bmatrix} : p + q = 1,\ p \geqslant 0,\ q \geqslant 0 \right\}.$$

Then $\{\operatorname{tr}(S) : S \in \mathcal{S}\} = \{1\}$, but the diagonal entries of members of $\mathcal{S}$ take all values in $[0, 1]$.

Before introducing our general results, we record a theorem for the case of matrices of very small size.

Theorem 2.10. *Let $\mathcal{S}$ be an indecomposable semigroup with finite diagonals consisting of 2×2 or 3×3 matrices. Then $\mathcal{S}$ is finite.*

Proof. For two indices i and j and a subset X of $\mathcal{S}$, put $X_{ij} = \{S_{ij} : S \in X\}$, where S_{ij} stands for the (i, j) entry of S.

Assume $\mathcal{S}$ consists of 2×2 matrices. Suppose $\mathcal{S}$ is infinite. Then without loss of generality we can assume that the set $\mathcal{S}_{12}$ is infinite. Fix $A \in \mathcal{S}$ such that $A_{21} \neq 0$. By the hypothesis, $(\mathcal{S}A)_{11} = \{S_{11}A_{11} + S_{12}A_{21} : S \in \mathcal{S}\}$ should be finite, which is impossible.

Now assume $\mathcal{S}$ consists of 3×3 matrices. Suppose $\mathcal{S}$ is infinite. Again, we can assume that $\mathcal{S}_{12}$ is infinite. Fix $A \in \mathcal{S}$ such that $A_{21} \neq 0$. Since $(\mathcal{S}A)_{11} = \{S_{11}A_{11} + S_{12}A_{21} + S_{13}A_{31}\}$ is finite, the set $\mathcal{S}_{13}$ is necessarily infinite. By considering $(B\mathcal{S})_{33}$, where $B_{31} \neq 0$, we see that $\mathcal{S}_{23}$ is infinite. Analogously, $\mathcal{S}_{21}$ is infinite.

Let $\mathcal{F} = \{S_{ii} : S \in \mathcal{S},\ 1 \leqslant i \leqslant 3\}$ and $\mathcal{F}_1 = \{a - bc : a, b, c \in \mathcal{F}\}$. Since $(ST)_{11} \in \mathcal{F}$ and $(ST)_{11} - S_{11}T_{11} = S_{12}T_{21} + S_{13}T_{31}$ for all $T, S \in \mathcal{S}$, we have $S_{12}T_{21} + S_{13}T_{31} \in \mathcal{F}_1$ for all $T, S \in \mathcal{S}$. Since $\mathcal{S}_{21}$ is infinite and $\mathcal{F}_1$ is finite, by the Pigeon Hole principle, there exist $T', T'' \in \mathcal{S}$ and a number $a \in \mathcal{F}_1$ such that $T'_{21} \neq T''_{21}$ and

$$\begin{aligned} S_{12}T'_{21} + S_{13}T'_{31} &= a, \\ S_{12}T''_{21} + S_{13}T''_{31} &= a \end{aligned} \tag{2}$$

for infinitely many $S \in \mathcal{S}$. Moreover, since $S_{12}T_{21} + S_{13}T_{31}$ is equal to zero only when $S_{12} = 0$ or $T_{21} = 0$, we can assume that $a \neq 0$. Since (2) has more than one solution, the matrix

$$\begin{bmatrix} T'_{21} & T'_{31} \\ T''_{21} & T''_{31} \end{bmatrix}$$

is not invertible. Hence the second line of (2) is in fact a scalar multiple of the first line. Since $a \neq 0$, this implies that $T'_{21} = T''_{21}$, a contradiction. □

3. Self-adjoint semigroups

In this section, we show that if a semigroup with finite diagonals is self-adjoint then it is finite. Moreover, our argument reveals the structure of such semigroups. In contrast with most statements in the other sections, it should be noted that the semigroups in the present section are not assumed to be indecomposable.

Definition 3.1. A collection $\mathcal{C}$ of matrices is called *self-adjoint* if for each $S \in \mathcal{C}$ we have $S^* \in \mathcal{C}$. Note that for our purposes, S^* is just the transpose of S.

We start with two nice properties of self-adjoint semigroups with finite trace.

Lemma 3.2. *Let $\mathcal{S}$ be a self-adjoint semigroup with finite trace. Then for each $S \in \mathcal{S}$ the matrix SS^* is an idempotent.*

Proof. By Lemma 2.1, every eigenvalue of SS^* is either zero or a root of unity. Since SS^* is self-adjoint, $\sigma(SS^*) \subseteq \{0, 1\}$. Since SS^* is also diagonalizable, the Lemma follows. □

Lemma 3.3. *If $\mathcal{S}$ is a self-adjoint semigroup with finite trace then each idempotent in $\mathcal{S}$ is self-adjoint.*

Proof. Let $E = E^2 \in \mathcal{S}$. Then E is unitarily similar to the matrix in the block form $\begin{bmatrix} I & X \\ 0 & 0 \end{bmatrix}$, where I is an identity matrix. With the same similarity, E^* is similar to $\begin{bmatrix} I & 0 \\ X^* & 0 \end{bmatrix}$. Then EE^* is similar to $\begin{bmatrix} I + XX^* & 0 \\ 0 & 0 \end{bmatrix}$. By Lemma 3.2, EE^* is an idempotent, hence $(I + XX^*)^2 = (I + XX^*)$. This, however, can only happen when $X = 0$. □

The next theorem is the main result of this section.

Theorem 3.4. *Let $\mathcal{S}$ be a (not necessarily indecomposable) semigroup with finite diagonals. If $\mathcal{S}$ is self-adjoint then $\mathcal{S}$ is finite. Moreover, all the entries of all matrices in $\mathcal{S}$ are of the form $\sqrt{\xi\eta}$, where ξ and η are either diagonal values of some matrices in $\mathcal{S}$ or zero.*

Remark 3.5. The statement in Theorem 3.4 can be abbreviated as follows. Let $\mathcal{S}$ be a self-adjoint semigroup of $N \times N$ matrices with finite diagonals. If $\mathcal{F} = \{S_{ii} : S \in \mathcal{S},\ i = 1, \dots, N\} \cup \{0\}$ and $\widehat{\mathcal{F}} = \{S_{ij} : S \in \mathcal{S},\ i, j = 1, \dots, N\}$ then

$$\widehat{\mathcal{F}} \subseteq \sqrt{\mathcal{F} \cdot \mathcal{F}}.$$

Proof of Theorem 3.4. Let $\mathcal{F} = \{S_{ii} : S \in \mathcal{S},\ i = 1, \dots, N\}$. We will prove that every $S \in \mathcal{S}$ can be written in the block form

$$S = \Delta_1 \begin{bmatrix} u_1 v_1^* & 0 & \dots & 0 \\ 0 & u_2 v_2^* & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \dots & u_k v_k^* \end{bmatrix} \Delta_2^*, \tag{3}$$

where Δ_1 and Δ_2 are each permutations and u_i, v_i are vectors whose entries are either of the form $\sqrt{\xi}$, with $\xi \in \mathcal{F}$ or are all zero (with no restrictions on the size of u_i and v_i; that is, the blocks $u_i v_i^*$ are in general rectangular).

Fix $S \in \mathcal{S}$. Set $P = SS^*$ and $Q = S^*S$. By Lemma 3.2, both P and Q are self-adjoint idempotents. Choose two permutations Γ_1 and Γ_2 such that the matrices $P_1 = \Gamma_1 P \Gamma_1^*$ and $Q_1 = \Gamma_2 Q \Gamma_2^*$ are block-diagonal with self-adjoint blocks of rank one or zero. Since $\operatorname{rank}(P) = \operatorname{rank}(Q) = \operatorname{rank}(S)$, we deduce that P_1 and Q_1 have the same number of nonzero blocks. Denote this number by r. That is, $P_1 = (P_1)_1 \oplus \dots \oplus (P_1)_r \oplus 0$ and $Q_1 = (Q_1)_1 \oplus \dots \oplus (Q_1)_r \oplus 0$, where either of the last zero entries could be absent.

Put $T = \Gamma_1 S \Gamma_2^*$. Then clearly $TT^* = P_1$ and $T^*T = Q_1$. Write T in the rectangular block form

$$T = \begin{bmatrix} T_{11} & \dots & T_{1r} & T_{1r+1} \\ \vdots & & \vdots & \\ T_{r1} & \dots & T_{rr} & T_{rr+1} \\ T_{r+11} & \dots & T_{r+1r} & T_{r+1r+1} \end{bmatrix},$$

where the vertical sizes of blocks are those of the blocks of P_1 and the horizontal sizes are those of the blocks of Q_1, and the $(r+1)$th row or $(r+1)$th column, or both could be void.

Since $P_1 = TT^*$ has the same range as T, we get $P_1 T = T$. Analogously, $TQ_1 = T$. Therefore, $P_1 T Q_1 = T$. Observe that in fact T is a partial isometry with corresponding projections P_1 and Q_1.

We claim that each block row and each block column of T has at most one nonzero block. Indeed, since TT^* is block-diagonal, we get $\sum_{k=1}^{r+1} T_{ik} T_{jk}^* = 0$ for all $i \neq j$. Hence for each k and $i \neq j$ we have $T_{ik} T_{jk}^* = 0$. This implies that if for some n and m the (n, m) entry of T_{ik} is not zero then the mth column of each T_{jk} is zero for all $j \neq i$. Since $P_1 T Q_1 = T$ and the diagonal entries of P_1 and Q_1 are strictly positive or zero, the entries of all T_{ij} are either all zero or are all nonzero simultaneously. It follows that each block column of T can contain at most one nonzero block. Considering T^*T, we get the same conclusion about the block rows.

Changing the order of blocks in Q_1 (by changing Γ_2), if necessary, we can assume that T is block-diagonal with rectangular diagonal blocks:

$$T = \begin{bmatrix} T_1 & \dots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \dots & T_r & 0 \\ 0 & \dots & 0 & 0 \end{bmatrix},$$

where $T_i = (P_1)_i T_i (Q_1)_i$ for all $i = 1, \dots, r$. Also, $T_i T_i^* = (P_1)_i$ and $T_i^* T_i = (Q_1)_i$.

Recalling that every $(P_1)_i$ and $(Q_1)_i$ is a rank-one projection, write $(P_1)_i = x_i x_i^*$ and $(Q_1)_i = y_i y_i^*$ for some vectors x_i and y_i satisfying $\|x_i\| = \|y_i\| = x_i^* x_i = y_i^* y_i = 1$ $(i = 1, \dots, r)$. Clearly, $\operatorname{rank}(T_i) = 1$ for all $i = 1, \dots, k$. Hence for each i there exist vectors u_i and v_i such that $T_i = u_i v_i^*$.

Fix i and denote for simplicity of notation $x = x_i$, $y = y_i$, $u = u_i$, and $v = v_i$. Since $P_1 T = T$ and $TQ_1 = T$, we get $xx^*uv^* = uv^*$ and $uv^*yy^* = uv^*$. Let $\alpha = x^*u$ and $\beta = v^*y$. Then $uv^* = \alpha xv^* = \beta uy^*$. This is only possible when $u = \alpha x$ and $v = \beta y$.

This shows that there is a scalar γ such that $uv^* = \gamma xy^*$. We claim that $\gamma = 1$. Indeed, from the the equality $TT^* = P_1$, we obtain $\gamma^2 (xy^*)(xy^*)^* = \gamma^2 xy^*yx^* = \gamma^2 xx^*$ is equal to xx^*. Since $\gamma \geqslant 0$, we get $\gamma = 1$.

We have shown that $T_i = x_i y_i^*$ for each $i = 1, \dots, r$. To establish formula (3), it is left to note that since for all i and j the numbers $(x_i)_j^2$ and $(y_i)_j^2$ are some diagonal entries of P_1 and Q_1, respectively, the entries of x_i and y_i are all of the form $\sqrt{\xi}$ with $\xi \in \mathcal{F}$. $\square$

Remark 3.6. The representation (3) in the proof of the above theorem will still be valid if we replace finiteness of the diagonal entries in the hypothesis of Theorem 3.4 with the finiteness of the trace.

4. Constant-rank semigroups

In this section, we will prove that if all nonzero matrices in an indecomposable semigroup with finite diagonals have the same rank, then the semigroup must be finite. The key step in obtaining this result is proving that the idempotent matrices in such a semigroup form a finite set (Theorem 4.3). We will need a series of lemmas to prove this.

Recall that if E is a nonnegative idempotent matrix then, after a permutation, E can be written as

$$E = \begin{bmatrix} 0 & XF & XFY \\ 0 & F & FY \\ 0 & 0 & 0 \end{bmatrix}, \tag{4}$$

where F is a nonnegative idempotent without zero rows or zero columns and X, Y are two nonnegative matrices. Furthermore, once E is in the form (4), then the $(2, 2)$ block, F, of E is called the *rigid part* of E.

The next lemma is the first step in establishing the finiteness of the set of idempotents of a semigroup with finite diagonals. Note that it requires neither indecomposability nor constancy of rank.

Lemma 4.1. *Let $\mathcal{S}$ be a semigroup with finite diagonals. Then the set*

$$\{F : F \text{ is the rigid part of some } E = E^2 \in \mathcal{S}\}$$

is finite.

Proof. Let N be the size of matrices in $\mathcal{S}$. Fix three numbers $m, n, k \geqslant 0$ such that $m + n + k = N$. We will prove that the set

$$\mathcal{F} = \{F : F \text{ is the rigid part of some } E = E^2 \in \mathcal{S} \\ \text{whose diagonal blocks are of size } m, n, \text{ and } k, \text{ respectively}\}$$

is finite. For each $F \in \mathcal{F}$ there exists a permutation matrix P such that $P^{-1}FP = E_1 \oplus \cdots \oplus E_r$, where each E_i is an idempotent of rank one whose entries are all positive. There are only finitely many choices for the permutation P, the number of blocks, r, and the sizes of each block in this representation. Therefore, it suffices to show that, after a fixed permutation P, there are only finitely many members in $\mathcal{F}$ having the same sequence of block sizes.

Let $F', F'' \in \mathcal{F}$ and a permutation P be such that $P^{-1}F'P = E'_1 \oplus \cdots \oplus E'_r, P^{-1}F''P = E''_1 \oplus \cdots \oplus E''_r$ and the sizes of E'_i and E''_i are the same for all $i = 1, \ldots, r$. Fix $i \in \{1, \ldots r\}$. We will prove that if the sequences of the diagonal entries of E'_i and E''_i are the same (that is, if $(E'_i)_{jj} = (E''_i)_{jj}$ for all j) then $E'_i = E''_i$. Since there are only finitely many choices for such diagonal sequences, the conclusion will follow.

Relabel for convenience $E'_i = Q, E''_i = R$. If Q and R have size 1, we are done. Hence we can assume that the size is at least 2. Since Q and R are both positive rank-one matrices with equal diagonals, there is a positive diagonal matrix D such that $R = DQD^{-1}$. Also, since Q and R are both strictly positive, RQ is again of rank one. Thus, $\sigma(RQ) = \{\text{tr}(RQ), 0\}$. Let $Q = (q_{ij}), D = \text{diag}(d_j)$:

$$\begin{aligned} \text{tr}(RQ) - 1 &= \text{tr}(DQD^{-1}Q) - \text{tr}(Q^2) = \sum_{i,j} d_i d_j^{-1} q_{ij} q_{ji} - \sum_{i,j} q_{ij} q_{ji} \\ &= \sum_{i,j} (d_i d_j^{-1} - 1) q_{ij} q_{ji} = \sum_{i<j} (d_i d_j^{-1} + d_j d_i^{-1} - 2) q_{ij} q_{ji}. \end{aligned}$$

We will be done if we prove that D is a multiple of the identity. Assume otherwise. Fix $i < j$ such that $d_i \neq d_j$. Observe that for $a > 0$ we have $a + a^{-1} \geqslant 2$ and the equality holds if and only if $a = 1$. Hence using $a = d_i d_j^{-1}$, we get $(d_i d_j^{-1} + d_j d_i^{-1} - 2) q_{ij} q_{ji} > 0$, by strict positivity of elements of Q.

Thus $\operatorname{tr}(RQ) > 1$, and therefore the spectral radius of RQ, $\rho(RQ) > 1$, so that $\rho(F''F') > 1$. This is impossible by Lemma 2.1. □

In the following lemma, we establish finiteness of the set of idempotents of a special kind in semigroups with finite diagonals having constant rank.

Lemma 4.2. *Let $\mathcal{S}$ be an indecomposable semigroup with finite diagonals such that all nonzero members of $\mathcal{S}$ have the same rank and*

$$\mathcal{E} = \left\{ E = E^2 \in \mathcal{S} : E = \begin{bmatrix} F & X \\ 0 & 0 \end{bmatrix} \text{ for some block } X \right\},$$

where F is a fixed idempotent matrix without zero rows and columns. Then $\mathcal{E}$ is finite.

Proof. Denote by r the rank of all nonzero members of $\mathcal{S}$. Applying a suitable permutation to $\mathcal{S}$ we can assume that F is of the form $F = F_1 \oplus \cdots \oplus F_r$, where each F_i is an idempotent of rank one whose entries are all positive. Furthermore, applying a diagonal similarity, we can assume that F is row stochastic.

Let k be the size of F. Define J_k and $\mathcal{S}_k$ as in Lemma 2.5. Clearly, $\mathcal{E} \subseteq J_k$. We shall show that every nonzero member of $\mathcal{S}_k$ has rank r. Indeed, pick any nonzero $A \in \mathcal{S}_k$. Then there is a matrix T in $\mathcal{S}$ of the form $T = \begin{bmatrix} A & B \\ 0 & 0 \end{bmatrix}$ for some nonnegative matrix B. Pick any $E = \begin{bmatrix} F & X \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} F & FX \\ 0 & 0 \end{bmatrix} \in \mathcal{E}$. Then $ETE = \begin{bmatrix} FAF & FAFX \\ 0 & 0 \end{bmatrix}$. Since $A \neq 0$ and F is block-diagonal with diagonal blocks having no zero entries, $FAF \neq 0$. Therefore, $ETE \neq 0$, and thus $\operatorname{rank}(ETE) = r$. Since each column of $FAFX$ is a linear combination of columns of FAF, we get $\operatorname{rank}(ETE) = \operatorname{rank}(FAF) = r$. Hence $r = \operatorname{rank}(FAF) \leqslant \operatorname{rank}(A) \leqslant \operatorname{rank}(T) = r$, and thus $\operatorname{rank}(A) = r$. So, in view of Lemma 2.5, we conclude that $\mathcal{S}_k$ is an indecomposable semigroup with finite diagonals such that every nonzero member of $\mathcal{S}_k$ has rank r. Then clearly F is a nonzero idempotent in $\mathcal{S}_k$. Define

$$\mathcal{S}_0 = F\mathcal{S}_k F.$$

By Lemma 2.4 we deduce that $\mathcal{S}_0$ is a finite group that is block-monomial relative to the block structure inherited from F.

Consider the set

$$\mathcal{X} = \left\{ X : X = FX \text{ and } \begin{bmatrix} F & X \\ 0 & 0 \end{bmatrix} \in \mathcal{S} \right\}.$$

To prove the lemma, we need to show that $\mathcal{X}$ is finite. Write every $X \in \mathcal{X}$ in a block form compatible with the block form of F:

$$X = \begin{bmatrix} X_1 \\ \vdots \\ X_r \end{bmatrix} = \begin{bmatrix} F_1 X_1 \\ \vdots \\ F_r X_r \end{bmatrix}.$$

Since the blocks of F are row stochastic and have rank one, all rows of each F_i and of each $X_i = F_i X_i$ are the same ($i = 1, \ldots, r$). This in particular implies that any given entry of X can be moved into the $(1, 1)$ position by applying a suitable permutation to $\mathcal{S}$ that keeps F in block-diagonal form with the same diagonal blocks (the order of blocks can change). Therefore, it is enough to prove that the $(1, 1)$ entry of X can only take finitely many values as X runs over $\mathcal{X}$. Denote the $(1, 1)$ entry of X by a_X. Put $\mathcal{X}_1 = \{X \in \mathcal{X} : a_X \neq 0\}$. To prove the lemma, we need to show that $\{a_X : X \in \mathcal{X}_1\}$ is finite.

Since $\mathcal{S}$ is indecomposable, there exists a matrix $S = \begin{bmatrix} H & K \\ R & Q \end{bmatrix}$ in $\mathcal{S}$ such that the $(1, 1)$ entry of R, R_{11}, is nonzero. For each $X \in \mathcal{X}_1$ the north-west block of the product

$$\begin{bmatrix} F & X \\ 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} H & K \\ R & Q \end{bmatrix} \cdot \begin{bmatrix} F & X \\ 0 & 0 \end{bmatrix}$$

belongs to $\mathcal{S}_0$ and is equal to $F(H+XR)F$. Since $\mathcal{S}_0$ is block monomial with respect to the block structure of F, so is the set $\{F(H+XR)F : X \in \mathcal{X}_1\}$. However, H is fixed and all matrices in this expression are nonnegative. Therefore, the set $\mathcal{Y}_1 = \{FXRF : X \in \mathcal{X}_1\}$ is finite and has the property that every row of blocks in each matrix in $\mathcal{Y}_1$ has at most one nonzero block.

Write RF in a block form, conforming to the block columns of F:

$$RF = \begin{bmatrix} L_1 & \dots & L_r \end{bmatrix}.$$

For each $X \in \mathcal{X}_1$ we have

$$FXRF = \begin{bmatrix} X_1L_1 & \dots & X_1L_r \\ \vdots & & \vdots \\ X_rL_1 & \dots & X_rL_r \end{bmatrix}.$$

Since $a_X \neq 0$ for all $X \in \mathcal{X}_1$, the block $X_1L_1 \neq 0$. Therefore, $X_1L_i = 0$ for all $i \in \{2, \dots, r\}$. Again, by $a_X \neq 0$ this implies that the first row of each L_i is equal to zero ($i = 2, \dots, r$). Since the first entry in every row of $X_1 = F_1X_1$ is equal to a_X, the leading entry of $RX = L_1X_1 + \dots + L_rX_r$ is equal to $s \cdot a_X$, where s is the sum of elements from the first row of L_1. Observe that $s \neq 0$ by the choice of R. Also, RX is the south-east block of the product

$$\begin{bmatrix} H & K \\ R & Q \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} F & X \\ 0 & 0 \end{bmatrix},$$

which belongs to $\mathcal{S}$. Therefore, there are only a finite number of values for $s \cdot a_X$. Since s is independent of X, the set $\{a_X : X \in \mathcal{X}_1\}$ is finite which completes the proof. $\square$

Theorem 4.3. *Let $\mathcal{S}$ be an indecomposable semigroup with finite diagonals. If all nonzero elements of $\mathcal{S}$ have the same rank then the set of idempotents in $\mathcal{S}$ is finite.*

Proof. Each idempotent in $\mathcal{S}$ is in the form of (4) after a suitable permutation. Since the number of possible permutations is finite, it is enough to prove that for each permutation P, the indecomposable semigroup $P^{-1}\mathcal{S}P$ contains finitely many idempotents in the form of (4).

Relabeling, if necessary, we can assume that the permutation P has already been applied to $\mathcal{S}$. For a fixed nonnegative idempotent F without zero rows or zero columns, define

$$\mathcal{E}_F = \{E = E^2 \in \mathcal{S} : \text{ the } (2,2) \text{ block of } E \text{ in the block form of (4) is } F\},$$
$$\mathcal{X}_F = \{XF : XF \text{ is the } (1,2) \text{ block in the form of (4) for some } E \in \mathcal{E}_F\},$$
$$\mathcal{Y}_F = \{FY : FY \text{ is the } (2,3) \text{ block in the form of (4) for some } E \in \mathcal{E}_F\}.$$

Fix the $(2,2)$-block F. By Lemma 4.1, it suffices to show that $\mathcal{X}_F$ and $\mathcal{Y}_F$ are finite.

Denote by k (by n, respectively) the number of rows in the $(2,1)$ block (in the $(2,3)$ block, respectively) of the representation (4). Let $i \in \{0, \dots, n\}$. We will prove that the set $\mathcal{Y}_{F,i} = \{FY \in \mathcal{Y}_F : Y \text{ has exactly } i \text{ zero columns}\}$ is finite. Suppose that $i = 0$. Define J'_k and $\mathcal{S}'_k$ as in Lemma 2.6. By Lemma 2.6, $\mathcal{S}'_k$ is an indecomposable semigroup with finite diagonals. Therefore, by Lemma 4.2 the set of all idempotents in $\mathcal{S}'_k$ of the form $\begin{bmatrix} F & FY \\ 0 & 0 \end{bmatrix}$ is finite. This shows that $\mathcal{Y}_{F,0}$ is finite.

Suppose $i > 0$. Then there is a permutation Q which turns idempotents of the form $\begin{bmatrix} F & FY \\ 0 & 0 \end{bmatrix}$ in $\mathcal{S}'_k$ into the idempotents of the form $\begin{bmatrix} 0 & 0 \\ 0 & E_1 \end{bmatrix}$, where E_1 is of the form $\begin{bmatrix} F & FY_1 \\ 0 & 0 \end{bmatrix}$ and Y_1 has no zero columns. Now the finiteness of $\mathcal{Y}_{F,i}$ follows from the argument in the previous paragraph applied to the semigroup $Q^{-1}\mathcal{S}Q$.

The finiteness of each $\mathcal{X}_F$ is established by applying an analogous argument to $\mathcal{S}^*$. $\square$

The following example shows that the condition on the rank is important in Theorem 4.3.

Example 4.4. An indecomposable semigroup with finite diagonals having infinitely many idempotents:

$$\mathcal{S} = \left\{ \begin{bmatrix} I & S \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} E & 0 \\ 0 & E \end{bmatrix}, \begin{bmatrix} 0 & E \\ E & 0 \end{bmatrix}, \begin{bmatrix} E & E \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ E & E \end{bmatrix} \right\},$$

where $E = \begin{bmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{bmatrix}$ and S runs over all matrices of form $\begin{bmatrix} p & q \\ q & p \end{bmatrix}$, with $p + q = 1, p, q \geqslant 0$.

Lemma 4.5. *Let N be a nonnegative $n \times n$ matrix such that $N^2 = 0$. Then there exists a permutation of the basis vectors such that N can be written as $N = \begin{bmatrix} 0 & A \\ 0 & 0 \end{bmatrix}$ (with square diagonal blocks). Moreover, if N is nonzero then A can be chosen to contain no zero columns or (alternatively) no zero rows.*

Proof. Let $\mathcal{F} = \{i : Ne_i = 0\}$, where (e_i) is the standard unit vector basis. We will first show that $\mathcal{F}$ cannot be empty. Suppose otherwise. Then $Ne_1 = (a_1 e_1 + \cdots + a_n e_n)$ for some nonnegative a_i, where at least one, say a_k, is positive. Then, by the nonnegativity of N and since $Na_k e_k \neq 0$, $\|N^2 e_1\| \geqslant \|N(a_k e_k)\| > 0$, which is a contradiction. Therefore, applying a suitable permutation, we can assume that $\mathcal{F} = \{1, \ldots, k\}$ for some k. Since $N^2 = 0$, for each $i \in \{k+1, \ldots, n\}$ we have $Ne_i = \sum_{j \in \mathcal{F}} a_{ij} e_j$ for some nonnegative a_{ij}. This shows that N can be represented in the desired form with A having no zero columns (provided $N \neq 0$). If A has zero rows then, applying a permutation and partitioning the first diagonal block into two diagonal subblocks, we obtain a new A with no zero rows (but some zero columns). $\square$

Before we can state the main result of this section, we need another lemma.

Lemma 4.6. *Let $\mathcal{S}$ be an indecomposable semigroup with finite diagonals. If all nonzero members of $\mathcal{S}$ have the same rank, then the set $\{N \in \mathcal{S} : N$ is nilpotent$\}$ is finite.*

Proof. Denote by r the rank of the nonzero elements in $\mathcal{S}$. The proof is by induction on the size n of matrices in $\mathcal{S}$. If $n = 1$ then there are no nonzero nilpotent matrices in $\mathcal{S}$. Let $n > 1$.

Clearly, since the rank of all nonzero elements of $\mathcal{S}$ is the same, if $N \in \mathcal{S}$ is nilpotent then $N^2 = 0$. By Lemma 4.5, after a permutation of the basis, we can write $N = \begin{bmatrix} 0 & A \\ 0 & 0 \end{bmatrix}$ for some nonnegative matrix A without zero rows. Since the number of possible permutations is finite, it is enough, as in Theorem 4.3, to show that $\mathcal{S}$ contains only finitely many nilpotent matrices in this block form.

Define

$$\mathcal{N}_k = \left\{ \begin{bmatrix} 0 & A \\ 0 & 0 \end{bmatrix} \in \mathcal{S} : A \text{ has } k \text{ nonzero rows and no zero rows} \right\}.$$

(Note that we have to allow A to have zero columns in the definition above, because the diagonal blocks have to be square.) For a matrix $N \in \mathcal{N}_k$, we will denote by a_N the leading entry, A_{11}, of the block A. As in the Proof of Theorem 4.3, it is enough to show that the set $\{a_N \neq 0 : N \in \mathcal{N}_k\}$ is finite.

Pick any matrix $M = \begin{bmatrix} H & L \\ J & K \end{bmatrix} \in \mathcal{S}$ such that the leading entry of J is different from zero. If $a_N \neq 0$ then NM is not nilpotent, and hence a power of NM is a nonzero idempotent by Corollary 2.3. Denote this idempotent by E_N. Since N and E_N have the same range, $E_N N = N$. In particular, the zero rows of E_N and N are the same. Hence in the block form inherited from N we get $E_N = \begin{bmatrix} Q & Z \\ 0 & 0 \end{bmatrix}$. Clearly, $Q = Q^2$ and $Z = QZ$, so that Q has no zero rows.

Case 1. Suppose that E_N and N have common zero columns. After a suitable permutation the matrices E_N and N can be written in the block form

$$\begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & XF & XFY \\ 0 & 0 & F & FY \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & B \\ 0 & 0 & 0 & C \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

respectively, where F has no zero columns and the fourth block column in each of the two matrices has no common zero columns. Since $E_N N = N$, we get $B = XFC$ and $C = FC$. In particular FY and C have no common zero columns. Let j be the number of zero columns in the first two block columns. Define $\mathcal{S}_j'$ as in Lemma 2.6. Then $\mathcal{S}_j'$ is an indecomposable semigroup. We will show now that the rank of nonzero elements in $\mathcal{S}_j'$ is equal to r.

Let $\widetilde{F} = \begin{bmatrix} F & FY \\ 0 & 0 \end{bmatrix}$, $\widetilde{X} = \begin{bmatrix} 0 & 0 \\ X & 0 \end{bmatrix}$, and $\widetilde{C} = \begin{bmatrix} 0 & C \\ 0 & 0 \end{bmatrix}$, then $E_N = \begin{bmatrix} 0 & \widetilde{X}\widetilde{F} \\ 0 & \widetilde{F} \end{bmatrix}$ and $N = \begin{bmatrix} 0 & \widetilde{X}\widetilde{C} \\ 0 & \widetilde{C} \end{bmatrix}$. Let $V \in \mathcal{S}_j'$ be nonzero. Then there exists $T = \begin{bmatrix} 0 & U \\ 0 & V \end{bmatrix} \in \mathcal{S}$. Consider the products $E_N T = \begin{bmatrix} 0 & \widetilde{X}\widetilde{F}V \\ 0 & \widetilde{F}V \end{bmatrix}$ and $NT = \begin{bmatrix} 0 & \widetilde{X}\widetilde{C}V \\ 0 & \widetilde{C}V \end{bmatrix}$. Since $V \neq 0$ and the matrices $\widetilde{F}$ and $\widetilde{C}$ have no common zero columns, one of the matrices $E_N T$ or NT is different from zero and hence has rank r. It is left to note that $\operatorname{rank}(E_N T) = \operatorname{rank}(\widetilde{F}V)$, $\operatorname{rank}(NT) = \operatorname{rank}(\widetilde{C}V)$, and $r = \operatorname{rank}(T) \geqslant \operatorname{rank}(V) \geqslant \operatorname{rank}(\widetilde{F}V) \vee \operatorname{rank}(\widetilde{C}V) = \operatorname{rank}(E_N T) \vee \operatorname{rank}(NT) = r$.

So, the semigroup $\mathcal{S}_j'$ is an indecomposable semigroup with finite diagonals whose nonzero elements have constant rank. Also, the size of matrices in $\mathcal{S}_j'$ is smaller than n. Thus, by the induction hypothesis, there are finitely many nilpotent matrices in $\mathcal{S}_j'$. Therefore, the matrix $\widetilde{C}$ comes from a finite set. By Theorem 4.3, there are finitely many idempotents in $\mathcal{S}$, hence the matrix $\widetilde{X}$ also comes from a finite set. Hence so does the matrix N.

Case 2. Suppose E_N and N have no common zero columns. Then in particular Q is an idempotent without zero rows and zero columns.

Write $Q = Q_1 \oplus \cdots \oplus Q_r$, where each Q_i is a rank-one idempotent without zero entries. In this block structure, write

$$A = \begin{bmatrix} A_1 \\ \vdots \\ A_r \end{bmatrix} = \begin{bmatrix} Q_1 A_1 \\ \vdots \\ Q_r A_r \end{bmatrix} \quad \text{and} \quad J = \begin{bmatrix} J_1 & \ldots & J_r \end{bmatrix}.$$

Applying a suitable diagonal similarity (note that these diagonal similarities come from a finite set since they depend on E_N only, and the set of idempotents in $\mathcal{S}$ is finite by Theorem 4.3), we can assume that Q is row stochastic. Then the rows of A_1 are all the same. Write $NM = \begin{bmatrix} AJ & AK \\ 0 & 0 \end{bmatrix}$. Clearly, $Q(AJ) = (AJ)Q$ and, since $E_N N = N$, $Q(AJ) = AJ$. The size of Q is $n - k$. Let $\mathcal{S}_{n-k}$ be as in Lemma 2.5. Then $\mathcal{S}_{n-k}$ is indecomposable. Therefore, the matrix AJ is block monomial by Lemma 2.4.

We have

$$AJ = \begin{bmatrix} A_1 J_1 & \ldots & A_1 J_r \\ \vdots & & \vdots \\ A_r J_1 & \ldots & A_r J_r \end{bmatrix}.$$

The leading block of AJ is different from zero. Hence $A_1 J_i = 0$ for all $i \in \{2, \ldots, r\}$. The leading entry of A_1 is nonzero. Hence the first row of each J_i $(i \in \{2, \ldots, r\})$ is zero. Denote the sum of elements in the first row of J_1 by s. By analyzing the product of $\begin{bmatrix} H & L \\ J & K \end{bmatrix}$ and $\begin{bmatrix} 0 & A \\ 0 & 0 \end{bmatrix}$, we get: the value $s a_N$ is on the diagonal of this product, and hence can only take finitely many values. Since s is independent of N and is different from zero, this shows that a_N can only take finitely many values, too. $\square$

Lemma 4.7. *Let $\mathcal{S}$ be an indecomposable semigroup with finite diagonals such that all nonzero members of $\mathcal{S}$ have the same rank. Let $E \in \mathcal{S}$ be a nonzero idempotent. Then the set $\mathcal{S}_E = \{S \in \mathcal{S} : ESE = S\}$ is a finite group with unit E.*

Proof. By Lemma 2.1, $\rho(T) = 1$ for all $T \in \mathcal{S}_E$. So, the statement follows from [7, 5.2.2(iv)]. The condition in [7, 5.2.2(iv)] that $\mathcal{S} = \overline{\mathbb{R}^+\mathcal{S}}$ is not essential since it is only used to establish that $\mathcal{S}_E$ is bounded (which follows from [2, Proposition 8]) and that for each $S \in \mathcal{S}_E$ a sequence of powers of S converges to an idempotent in $\mathcal{S}_E$ (which follows from Lemma 2.2). □

Theorem 4.8. *Let $\mathcal{S}$ be an indecomposable semigroup with finite diagonals. If all nonzero members of $\mathcal{S}$ have the same rank, then $\mathcal{S}$ is finite.*

Proof. Let $\mathcal{E}$ be the set of all nonzero idempotents in $\mathcal{S}$. For each $E \in \mathcal{E}$, denote $\mathcal{S}_E = \{S \in \mathcal{S} : ESE = S\}$.

By Lemma 4.7, $\mathcal{S}_E$ is a finite group with unit E. We claim that each non-nilpotent member of $\mathcal{S}$ belongs to $\cup_{E\in\mathcal{E}}\mathcal{S}_E$. Indeed, by Lemma 2.2, each $S \in \mathcal{S}$ is represented in some basis as $\begin{bmatrix} U & 0 \\ 0 & N \end{bmatrix}$, where U is a unitary diagonal matrix and N is a nilpotent matrix. If S is not nilpotent then $N = 0$ because the rank of all nonzero elements of $\mathcal{S}$ is the same. Therefore, a power S^m of any non-nilpotent $S \in \mathcal{S}$ is a nonzero idempotent E such that $ESE = S$.

Since the set $\mathcal{E}$ is finite by Theorem 4.3, this shows that the set of non-nilpotent matrices in $\mathcal{S}$ is finite. The finiteness of nilpotent elements in $\mathcal{S}$ is shown in Lemma 4.6. □

The natural (in view of Theorem 2.8) question whether the finiteness of diagonal entries in the statement of Theorem 4.8 can be replaced with finiteness of the trace has a negative answer, as Example 2.9 in Section 2 shows. In fact, the semigroup in that example consists of idempotents only, so that the corresponding question asked about Theorem 4.3 would already have a negative answer.

5. Admissible diagonal values

In this section, we analyze what values there could be on the diagonal positions of a semigroup with finite diagonals.

Theorem 5.1. *Let $\mathcal{S}$ be an indecomposable semigroup with finite diagonals. Then for each $S \in \mathcal{S}$ the sequence (S_{ii}) can be partitioned into disjoint subsequences each of which either adds up to 1 or consists of zeros.*

Proof. Let $S \in \mathcal{S}$ be fixed. By Lemma 2.1, the possible eigenvalues of S are roots of unity and zero. After a permutation, S can be decomposed into a block triangular form whose diagonal blocks are indecomposable matrices $S_1, \dots, S_k$. Pick any $i \in \{1, \dots, k\}$ and denote for convenience $T = S_i$. It is enough to prove that the statement of theorem is valid for T.

Since T is indecomposable, T is not nilpotent. Let $r \geqslant 1$ be the number of nonzero eigenvalues (counting multiplicities) of T. Then $r = \operatorname{rank}(T)$. By Corollary 2.3, the minimal rank of nonzero matrices in the norm closed semigroup generated by T is r. Hence by the Perron–Frobenius theorem [7, Corollary 5.2.13], after a permutation, T can be written in the block form

$$T = \begin{bmatrix} 0 & \dots & 0 & T_r \\ T_1 & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & T_{r-1} & 0 \end{bmatrix}.$$

If $r > 1$ then all the diagonal elements are zero, since permutations only change the order of diagonal elements. If $r = 1$ then zero has multiplicity $n - 1$ (where n is the size of T). Since $1 \in \sigma(T)$, we get $\operatorname{tr}(T) = 1$, hence the sum of diagonal elements of T is 1. □

Definition 5.2. A finite set $\mathcal{F} \subseteq \mathbb{R}_+$ is called *admissible* if $\mathcal{F}$ can be written as a (not necessarily disjoint) union $\mathcal{F} = \mathcal{F}_1 \cup \cdots \cup \mathcal{F}_n$ where each $\mathcal{F}_k = \{x_1, \ldots, x_{i_k}\}$ satisfies the condition that

$$\sum_{j=1}^{i_k} m_j x_j = 1$$

for some $m_j \in \mathbb{N}$ $(j = 1, \ldots, i_k)$.

Example 5.3. The set $\left\{\frac{1}{5}, \frac{1}{3}, \frac{2}{9}, \frac{2}{3}\right\}$ is admissible since $5 \cdot \frac{1}{5} = 1$, $\frac{1}{3} + 3 \cdot \frac{2}{9} = 1$, and $\frac{2}{3} + \frac{1}{3} = 1$. The sets $\{0\}$ and $\left\{\frac{3}{7}, \frac{2}{5}\right\}$ are not admissible.

The following lemma is obvious.

Lemma 5.4. *A finite union of admissible sets is admissible.*

Theorem 5.5. *Let $\mathcal{F} \subseteq \mathbb{R}$ be such that $0 \in \mathcal{F}$. Then $\mathcal{F}$ is admissible if and only if there exists an indecomposable semigroup $\mathcal{S}$ with finite diagonals such that the set of diagonal values of all the matrices in $\mathcal{S}$ is equal to $\mathcal{F}$.*

Proof. If $\mathcal{S}$ is an indecomposable semigroup with finite diagonals and $S \in \mathcal{S}$ then the set $\mathcal{F}_S$ of all the diagonal entries of S is admissible by Theorem 5.1. Since $\mathcal{S}$ is a semigroup with finite diagonals, there are only finitely many choices for the set $\mathcal{F}_S$. Therefore, $\mathcal{F} = \cup_{S \in \mathcal{S}} \mathcal{F}_S$ is admissible by Lemma 5.4.

Let $\mathcal{F}$ be admissible. Write $\mathcal{F} = \mathcal{F}_1 \cup \cdots \cup \mathcal{F}_n$ as in the definition of an admissible set. We will show that there exists a semigroup $\mathcal{S}$ as in the statement of the theorem.

For each $k \in \{1, \ldots, n\}$, write $\mathcal{F}_k = \{x_1^{(k)}, \ldots, x_{i_k}^{(k)}\}$ and fix $m_1^{(k)}, \ldots, m_{i_k}^{(k)}$ such that $\sum_{j=1}^{i_k} m_j^{(k)} x_j^{(k)} = 1$. Put $N_k = \sum_{j=1}^{i_k} m_j^{(k)}$ and define the vector $y^{(k)} = (y_i^{(k)})_{i=1}^{N_k} \in \mathbb{R}^{N_k}$ by putting

$$y_i^{(k)} = x_j^{(k)} \text{ for all } i \in \left[\sum_{p=1}^{j-1} m_p^{(k)} + 1, \sum_{p=1}^{j} m_p^{(k)}\right] \cap \mathbb{N} \quad (j = 1, \ldots, i_k).$$

That is, $y^{(k)}$ has exactly $m_j^{(k)}$ coordinates equal to $x_j^{(k)}$. Also $\sum_{i=1}^{N_k} y_i^{(k)} = 1$. For each $i, j \in \{1, \ldots, n\}$, define the rank-one $N_j \times N_i$ matrix

$$T_{ij} = \begin{bmatrix} y_1^{(i)} & \ldots & y_{N_i}^{(i)} \\ \vdots & & \vdots \\ y_1^{(i)} & \ldots & y_{N_i}^{(i)} \end{bmatrix}.$$

Since each T_{ij} is row stochastic, a routine check shows that for all $i, j, k \in \{1, \ldots, n\}$ we have $T_{ij}T_{jk} = T_{ik}$.

Now let E_{ij} be the block matrix with n vertical and n horizontal blocks such that the (k, l) block of E_{ij} is equal to the $N_k \times N_l$ zero matrix if $k \neq i$ or $l \neq j$ and is equal to T_{ij} if $k = i$ and $l = j$. Define

$$\mathcal{S} = \{E_{ij} : 1 \leqslant i, j \leqslant n\} \cup \{\mathbf{0}\}.$$

Then clearly $\mathcal{S}$ is an indecomposable semigroup whose set of diagonal elements is $\mathcal{F}$. $\square$

The last statement to be proved in this paper is the assertion that if an admissible set $\mathcal{F} \subseteq \mathbb{R}_+$ does not contain zero, then there may not be an indecomposable semigroup of matrices whose diagonal entries form a set which is exactly $\mathcal{F}$. It will need an auxiliary lemma which may be of some independent interest.

Lemma 5.6. *Let $\mathcal{S}$ be a semigroup with finite diagonals such that no member of $\mathcal{S}$ has zero on the diagonal. If the minimal rank $m_{\mathcal{S}}$ of nonzero elements in $\mathcal{S}$ is not one, then $\mathcal{S}$ is decomposable.*

Proof. Suppose $\mathcal{S}$ is indecomposable and $m_{\mathcal{S}} \geqslant 2$. Fix a minimal idempotent $E \in \mathcal{S}$. Since E has no zeros on the diagonal, $E = E_1 \oplus \cdots \oplus E_{m_{\mathcal{S}}}$, where each E_i is a strictly positive idempotent.

Let $S \in \mathcal{S}$ be an arbitrary matrix. By Corollary 2.3, there is $m \in \mathbb{N}$ such that $(ESE)^m$ is an idempotent which we will denote by F. Clearly, $EF = FE = F$. Since the diagonal values of matrices in $\mathcal{S}$ do not admit zeros, $E = F$ by minimality of E.

We claim that up to a permutation similarity, S is block-diagonal relative to the block-structure inherited from E. Indeed, let us first show that ESE is block-diagonal. Suppose that ESE is not block-diagonal, that is, ESE has a nonzero, non-diagonal block. Without loss of generality, we can assume that the $(1, 2)$ block of ESE is not zero:

$$ESE = \begin{bmatrix} E_1 & X & \dots & * \\ * & E_2 & \dots & * \\ \vdots & \vdots & \ddots & \vdots \\ * & \dots & * & E_{m_{\mathcal{S}}} \end{bmatrix},$$

where $X \neq 0$. Since the diagonal blocks of E are strictly positive, it is easy to see that the $(1, 2)$ block of $(ESE)^m$ would be different from zero, too, which is a contradiction. Therefore, ESE is block-diagonal. Again, since the diagonal blocks of E are strictly positive, this is only possible if S is block-diagonal itself. $\square$

Proposition 5.7. *If $\mathcal{F} = \left\{\frac{1}{2}, \frac{1}{3}\right\}$ then there is no indecomposable semigroup, $\mathcal{S}$, such that the set of diagonal entries of matrices in $\mathcal{S}$ is equal to $\mathcal{F}$.*

Proof. Suppose such a semigroup, $\mathcal{S}$, exists. By Lemma 5.6, $\mathcal{S}$ contains an idempotent E of rank one. Since E cannot have zeros on the diagonal, E must be strictly positive. Since also $\operatorname{tr}(E) = 1$, there are, up to a diagonal similarity, only two choices for E:

$$\text{either} \quad E = \begin{bmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{bmatrix} \quad \text{or} \quad E = \begin{bmatrix} 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 1/3 \end{bmatrix}.$$

That is, $\mathcal{S}$ consists of either 2×2 matrices or 3×3 matrices. We will consider these two cases separately.

Assume the size of matrices in $\mathcal{S}$ is 2. Let A be a matrix having $1/3$ on the diagonal. That is, up to a permutation, $A = \begin{bmatrix} 1/3 & a \\ b & c \end{bmatrix}$ for some a, b, and c. By Lemma 2.1, the eigenvalues of A are either zero or roots of unity of degree at most 2. Also, $\operatorname{tr}(A) \geqslant 0$. Therefore, the only possible values for $\operatorname{tr}(A)$ are 0, 1, and 2. In either case, c cannot belong to $\mathcal{F}$.

Now let the size of matrices in $\mathcal{S}$ be 3. Again, fix a matrix A with $1/2$ on the diagonal. Denote the two other diagonal entries of A by a and b. Observe that in this case, the only possible values for $\operatorname{tr}(A)$ are 0, 1, 2, and 3, none of which can be achieved by choosing a and b in $\mathcal{F}$. $\square$

References

[1] A. Berman, R. Plemmons, Nonnegative Matrices in the Mathematical Sciences, Academic Press, New York, 1979.

[2] H. Gessesse, A.I. Popov, H. Radjavi, E. Spinu, A. Tcaciuc, V.G. Troitsky, Bounded indecomposable semigroups of non-negative matrices, Positivity, in press, doi:10.1007/s11117-009-0024-5.

[3] I. Kaplansky, The Engel–Kolchin Theorem Revisited, Contributions to Algebra (Collection of Papers Dedicated to Ellis Kolchin), Academic Press, New York, 1977.

[4] L. Livshits, G. MacDonald, H. Radjavi, Positive matrix semigroups with binary diagonals, preprint.

[5] W.E. Longstaff, H. Radjavi, On permutability and submultiplicativity of spectral radius, Canad. J. Math. 47 (5) (1995) 1007–1022.

[6] J. Okniński, Semigroups of Matrices, World Scientific, Singapore, 1998.

[7] H. Radjavi, P. Rosenthal, Simultaneous Triangularization, Springer-Verlag, New York, 2000.

[8] H. Radjavi, P. Rosenthal, Limitations on the size of semigroups of matrices, Semigroup Forum 76 (2008) 25–31.