

POLYNOMIAL-TIME COMPUTATION OF THE JOINT SPECTRAL RADIUS FOR SOME SETS OF NONNEGATIVE MATRICES*

VINCENT D. BLONDEL[†] AND YURII NESTEROV[‡]

Abstract. We propose two simple upper bounds for the joint spectral radius of sets of nonnegative matrices. These bounds, the *joint column radius* and the *joint row radius*, can be computed in polynomial time as solutions of convex optimization problems. We show that these bounds are within a factor $1/n$ of the exact value, where n is the size of the matrices. Moreover, for sets of matrices with independent column uncertainties or with independent row uncertainties, the corresponding bounds coincide with the joint spectral radius. In these cases, the joint spectral radius is also given by the largest spectral radius of the matrices in the set. As a by-product of these results, we propose a polynomial-time technique for solving Boolean optimization problems related to the spectral radius. We also describe economics and engineering applications of our results.

Key words. joint spectral radius, joint column radius, joint row radius, nonnegative matrices, asynchronous systems, convex optimization, Leontief model

AMS subject classifications. 15A18, 15-04, 15A60, 90C22, 93-04

DOI. 10.1137/080723764

1. Introduction. A discrete-time switching linear system generates trajectories of points

$$(1.1) \quad x_{k+1} = A_k x_k, \quad x_0 \in R^n,$$

with the matrices A_k taken in some *uncertainty set* $\mathcal{M} \subset R^{n \times n}$. The worst-case growth rate of these trajectories can be characterized by a joint spectral radius. The *joint spectral radius* (JSR) of the set of matrices \mathcal{M} is the smallest value $\rho \geq 0$ such that for every trajectory there is some constant C for which

$$\|x_k\| \leq C \rho^k$$

for all k . This optimal ρ provides valuable information about the switching linear system. In particular, the trajectories of the switching system all converge to the origin if and only if $\rho < 1$. In [22] the JSR of \mathcal{M} is defined in the following equivalent form:

$$(1.2) \quad \rho(\mathcal{M}) = \limsup_{k \rightarrow \infty} \max\{\|A\|^{1/k} : A \text{ is a product of length } k \text{ of matrices in } \mathcal{M}\}.$$

During the last decade, the JSR has proved useful in a number of application contexts, including wavelets [9, 11], capacity of codes [6, 17], switched and hybrid systems [14, 23], sensor networks [10, 16], combinatorics on words [8, 15], autoregressive

*Received by the editors May 9, 2008; accepted for publication (in revised form) by Y. C. Eldar April 29, 2009; published electronically August 5, 2009. The research results presented in this paper were supported by grant “Action de recherche concertée ARC 04/09-315” from the “Direction de la recherche scientifique - Communauté française de Belgique” and by the Belgian Programme on Interuniversity Attraction Poles initiated by the Belgian Federal Science Policy Office. The scientific responsibility rests with its authors.

<http://www.siam.org/journals/simax/31-3/72376.html>

[†]Division of Applied Mathematics, Université catholique de Louvain, 4 avenue Georges Lemaitre, B-1348 Louvain-la-Neuve, Belgium (vincent.blondel@uclouvain.be).

[‡]CORE, Université catholique de Louvain, 34 voie du Roman Pays, 1348 Louvain-la-Neuve, Belgium (yurii.nesterov@uclouvain.be).

models, Markov chains, etc. Unfortunately, the JSR of a set of matrices is notoriously difficult to compute and to approximate. In fact, even for the case of two matrices of dimension 47×47 that have nonnegative rational entries, the problem of checking the inequality $\rho \leq 1$ is algorithmically undecidable (see [5, 7]), and it is still unknown if the problem $\rho < 1$ is algorithmically decidable. Today, the list of matrix sets \mathcal{M} that have polynomial-time computable JSR is desperately small: the list includes the case where \mathcal{M} contains only triangular matrices of identical orientation and the case of a symmetric set $A \in \mathcal{M} \implies A^T \in \mathcal{M}$. In this last case the JSR is given by the largest singular value of the matrices in the set [20].

In recent years most research efforts have concentrated on finding reasonable approximations for the JSR (see, e.g., [3, 4, 19, 21, 25]). But all these approximations suffer from an intrinsic limitation since it is known that the problem of computing the JSR of two matrices with binaries entries is NP-hard and that, unless $P = NP$, there is no approximation algorithm that is polynomial with respect to the accuracy [24]. In all known approximation schemes, even for sets with only two matrices, the complexity of the computations grows exponentially with respect to the required accuracy.

In this paper, we propose two upper bounds for the JSR of arbitrary sets of nonnegative matrices, which are both within a factor $1/n$ of the exact value. These bounds, the *joint column radius* (JCR) and the *joint row radius* (JRR), can be computed in polynomial time as solutions to convex optimization problems.

We also consider the special case for which the set of matrices \mathcal{M} has independent column, or row, uncertainties. In terms of the switching systems (1.1) the row independent uncertainty situation corresponds to systems for which at every iteration all entries of the state vector x_k are updated, and the i th entry is updated by choosing one of the elements in $\{q^T x_t : q \in Q_i\}$. In particular, this includes the situation of asynchronous linear systems for which at every iteration only some of the state entries are updated according to a linear transformation and the others are kept unchanged.

It appears that in this special case, the JRR coincides with the JSR and with the largest spectral radius of the matrices in the set. As a by-product of this result, we are able to solve in polynomial time some Boolean optimization problems related to the spectral radius. Another interesting consequence of our results is the quasi-convexity of the spectral radius of a matrix with nonnegative matrices in each column when all other columns are fixed.

This paper is organized as follows. In the next section, we define the JCR (resp., JRR). We show that this value is the solution of a convex optimization problem, and we establish an upper and a lower bound for its quality as an approximation to the JSR. Then, in section 3, we show that for sets of matrices that have independent column (or row) uncertainty sets, the JCR (resp., JRR) coincides with the JSR. Moreover, it appears that in the uncertainty set there is always a critical matrix whose spectral radius coincides with the JSR.

In the last section, we discuss several applications of our results.

Notation. The entries of a (column) vector $x \in R^n$ are denoted by $x = (x^{(1)}, \dots, x^{(n)})^T$. For x and y in R^n we denote by $\langle x, y \rangle$ their inner product:

$$\langle x, y \rangle = \sum_{i=1}^n x^{(i)} y^{(i)} = x^T y.$$

The set of square matrices is denoted by M_n , and the set of square matrices with nonnegative entries is denoted by M_n^+ . For a vector $x \in R^n$, we denote by $D(x) \in M_n$ the diagonal matrix with the vector x as its diagonal, and by $e^x \in R^n$ we denote the

vector with coordinates $e^{x^{(i)}}$. Further, $\mathbf{1}$ denotes the vector of all ones, $\mathbf{0}$ denotes the vector of all zeros, and e_i denotes the i th coordinate vector in R^n . Finally, Δ_n denotes the standard simplex in R^n , and $|\mathcal{M}|$ the cardinality of the set \mathcal{M} .

2. Joint column (row) radius. Let $\rho(A)$ be the spectral radius of the matrix A , i.e., the largest magnitude of its eigenvalues. According to the Perron–Frobenius theorem, an irreducible¹ nonnegative matrix $A \in M_n^+$ has a unique eigenvector u (up to scalar multiplication) such that

$$A^T u = \rho(A) u,$$

and all components of the vector u are then positive. Let the column decomposition of A be given by $A = (a_1, \dots, a_n)$. It is well known (see, e.g., [13]) that the spectral radius of a nonnegative matrix admits the following representation:

$$(2.1) \quad \rho(A) = \inf_{u > 0} \max_{1 \leq i \leq n} \frac{\langle a_i, u \rangle}{u^{(i)}}.$$

Changing the variables $u = e^x$, we obtain

$$(2.2) \quad \rho(A) = \inf_{x \in R^n} \left[\phi_A(x) \stackrel{\text{def}}{=} \max_{1 \leq i \leq n} \langle a_i, e^x \rangle \cdot e^{-x^{(i)}} \right].$$

Note that the objective function $\phi_A(x)$ in this problem is convex. If A is irreducible, then the problem (2.2) has a unique solution ray spanned by direction $\mathbf{1} \in R^n$. We denote by $x(A)$ the point of the optimal ray satisfying the equation

$$\langle \mathbf{1}, x(A) \rangle = 0,$$

and we denote $u(A) \stackrel{\text{def}}{=} e^{x(A)} > \mathbf{0}$. Note that $A^T u(A) = \rho(A) \cdot u(A)$.

Representation (2.1) explains the role of the spectral radius in estimating the rate of growth/decrease of the powers of positive matrices. Indeed, for an arbitrary point $x_0 \in R_+^n$ define the sequence

$$x_k = A^k x_0, \quad k \geq 1.$$

Then

$$(2.3) \quad \begin{aligned} \langle u(A), x_{k+1} \rangle &= \langle u(A), Ax_k \rangle = \langle A^T u(A), x_k \rangle \\ &= \langle D^{-1}(u(A))A^T u(A), D(u(A))x_k \rangle \\ &\stackrel{(2.1)}{\leq} \langle \rho(A) \cdot \mathbf{1}, D(u(A))x_k \rangle = \rho(A) \cdot \langle u(A), x_k \rangle. \end{aligned}$$

It is interesting to note that exactly the same reasoning can be used in the analysis of switching systems.

¹A matrix A is reducible if there is a permutation matrix P for which

$$PAP^T = \begin{pmatrix} F & 0 \\ G & H \end{pmatrix}.$$

The matrix is irreducible if no such permutation exists.

Consider a compact set \mathcal{M} of nonnegative matrices. We define the *joint column radius* (JCR) of \mathcal{M} as follows:

$$(2.4) \quad \sigma(\mathcal{M}) = \inf_{x \in \mathbb{R}^n} \max_{A \in \mathcal{M}} \phi_A(x).$$

In this expression, the functions $\phi_A(x)$ are convex in x and so the function $\max_{A \in \mathcal{M}} \phi_A(x)$ is also convex. Since the JCR is a solution of the convex minimization problem (2.4), it can be computed in polynomial time by standard convex optimization algorithms.

We can provide another interesting interpretation of the JCR. Denote $\widehat{\mathcal{M}} = \text{Conv}(\mathcal{M})$ and consider the following optimization problem:

$$(2.5) \quad \max_{A \in \widehat{\mathcal{M}}} \rho(A).$$

The inequality

$$(2.6) \quad \max_{A \in \widehat{\mathcal{M}}} \rho(A) \leq \rho(\mathcal{M})$$

was proved as Proposition 15 in [4] for a finite collection \mathcal{M} . By the Carathéodory theorem this inequality can be easily extended to arbitrary convex uncertainty sets.

Note that

$$\begin{aligned} \max_{A \in \widehat{\mathcal{M}}} \rho(A) &\stackrel{(2.2)}{=} \max_{A \in \widehat{\mathcal{M}}} \inf_{x \in \mathbb{R}^n} \phi_A(x) \leq \inf_{x \in \mathbb{R}^n} \max_{A \in \widehat{\mathcal{M}}} \phi_A(x) \\ &= \inf_{x \in \mathbb{R}^n} \max_{A \in \mathcal{M}} \phi_A(x) = \sigma(\mathcal{M}). \end{aligned}$$

Thus, we can treat $\sigma(\mathcal{M})$ as a value of the usual *Lagrangian relaxation* of the non-convex optimization problem (2.5). Note that

$$\max_{A \in \mathcal{M}} \rho(A) \leq \max_{A \in \widehat{\mathcal{M}}} \rho(A) \leq \sigma(\mathcal{M}).$$

In section 3 we will discuss nontrivial situations when these inequalities can be replaced by equalities.

A quantity similar to the JCR can be introduced for the set of transposed matrices,

$$\mathcal{M}^T \stackrel{\text{def}}{=} \{A^T : A \in \mathcal{M}\}.$$

Namely, we can define the *joint row radius* (JRR) of the set \mathcal{M} as follows:

$$(2.7) \quad \sigma_T(\mathcal{M}) = \sigma(\mathcal{M}^T).$$

Since $\rho(\mathcal{M}) = \rho(\mathcal{M}^T)$, the above discussion also applies to the JRR. However, in general we have $\sigma_T(\mathcal{M}) \neq \sigma(\mathcal{M})$. In the remaining parts of the paper we will work mainly with the JCR, without mentioning that all our results can also be applied to the JRR.

Let us now prove that the JCR provides good upper and lower bounds for the JSR. First of all, we need to prove two technical lemmas.

LEMMA 1. *Consider the following perturbation of the uncertainty set \mathcal{M} :*

$$\mathcal{M}_\epsilon \stackrel{\text{def}}{=} \{A + \epsilon \mathbf{1} \mathbf{1}^T, A \in \mathcal{M}\}, \quad \epsilon \geq 0.$$

Then $\lim_{\epsilon \downarrow 0} \sigma(\mathcal{M}_\epsilon) = \sigma(\mathcal{M})$.

Proof. Indeed,

$$\begin{aligned} \sigma(\mathcal{M}) \leq \sigma(\mathcal{M}_\epsilon) &= \inf_{u > \mathbf{0}} \max_{A \in \mathcal{M}} \max_{1 \leq i \leq n} \left[\frac{1}{u^{(i)}} \cdot \langle a_i + \epsilon \mathbf{1}, u \rangle \right] \\ &\leq \inf_{\substack{\langle \mathbf{1}, u \rangle = 1, \\ u > \mathbf{0}}} \max_{A \in \mathcal{M}} \max_{1 \leq i \leq n} \left[\frac{1}{u^{(i)}} \cdot \langle a_i, u \rangle + \frac{\epsilon}{u^{(i)}} \right] \\ &\leq \inf_{\substack{\langle \mathbf{1}, u \rangle = 1, \\ u > \mathbf{0}}} [\xi_{\mathcal{M}}(u) + \epsilon F(u)] \stackrel{\text{def}}{=} \tau(\epsilon), \end{aligned}$$

where $\xi_{\mathcal{M}}(u) = \max_{A \in \mathcal{M}} \max_{1 \leq i \leq n} \left[\frac{1}{u^{(i)}} \cdot \langle a_i, u \rangle \right]$, and $F(u) = \max_{1 \leq i \leq n} \frac{1}{u^{(i)}}$ is a penalty function for the positive orthant. Since $F(u)$ is below bounded on its feasible set, by Theorem 1.3.2 in [18] we have²

$$\lim_{\epsilon \downarrow 0} \tau(\epsilon) = \inf_{\substack{\langle \mathbf{1}, u \rangle = 1, \\ u > \mathbf{0}}} \xi_{\mathcal{M}}(u) = \sigma(\mathcal{M}). \quad \square$$

LEMMA 2. *Let the elements of all matrices in \mathcal{M} be positive. Then there exists a matrix $A_* = (A_1 e_1, \dots, A_n e_n)$ formed by some matrices $A_i \in \widehat{\mathcal{M}}$, $i = 1, \dots, n$, such that*

$$(2.8) \quad \rho(A_*) = \sigma(\mathcal{M}).$$

Proof. Definition (2.4) of JCR can be written in the following form:

$$\begin{aligned} \sigma(\mathcal{M}) &= \inf_{x \in \mathbb{R}^n} \max_{A \in \mathcal{M}} \max_{1 \leq i \leq n} \left[e^{-x^{(i)}} \langle A e_i, e^x \rangle \right] \\ (2.9) \quad &= \inf_{x \in \mathbb{R}^n} \max_{1 \leq i \leq n} \left[\psi_i(x) \stackrel{\text{def}}{=} \max_{A \in \mathcal{M}} \left(e^{-x^{(i)}} \langle A e_i, e^x \rangle \right) \right]. \end{aligned}$$

Under conditions of the lemma, the infimum of the upper level minimization problem is attained at some point $x^* \in \mathbb{R}^n$. Therefore, there exists a vector of optimal dual multipliers $\lambda_* \in \Delta_n$ such that

$$0 = \sum_{i=1}^n \lambda_*^{(i)} g_i^*, \quad g_i \in \partial \psi_i(x_*),$$

$$\partial \psi_i(x_*) = \left\{ g = \nabla \left(e^{-x^{(i)}} \langle A e_i, e^x \rangle \right), A_i \in \widehat{\mathcal{M}} : \langle A e_i, e^{x_*} \rangle = e^{x_*^{(i)}} \psi_i(x_*) \right\}.$$

Moreover, there exist matrices $A_i^* \in \widehat{\mathcal{M}}$ such that

$$x^* = \arg \min_{x \in \mathbb{R}^n} \max_{1 \leq i \leq n} \left[e^{-x^{(i)}} \langle A_i^* e_i, e^x \rangle \right].$$

Hence, in view of (2.2), e^{x^*} is an eigenvector of the matrix $A_* = (A_1^* e_1, \dots, A_n^* e_n)$, and $\rho(A_*) = \sigma(\mathcal{M})$. \square

²In this theorem, the optimization problems are written in a min-form. However, all arguments work for the inf-form also.

We can now prove our main result.

THEOREM 1. *Let \mathcal{M} be a compact set of nonnegative matrices. Then*

$$(2.10) \quad \frac{1}{p} \cdot \sigma(\mathcal{M}) \leq \rho(\mathcal{M}) \leq \sigma(\mathcal{M}),$$

where $p = \min\{n, |\mathcal{M}|\}$.

Proof. Let us fix an arbitrary $\epsilon > 0$. Then, by Lemma 2 there exist a matrix

$$A_*^\epsilon = (A_1^\epsilon e_1, \dots, A_n^\epsilon e_n), \quad A_i^\epsilon \in \widehat{\mathcal{M}}_\epsilon, \quad i = 1, \dots, n,$$

which ensures equality $\rho(A_*^\epsilon) = \sigma(\mathcal{M}_\epsilon)$. Denote by m the number of different matrices in the representation of A_*^ϵ :

$$\{A_i^\epsilon\}_{i=1}^n = \{B_j^\epsilon\}_{j=1}^m.$$

Then, $\frac{1}{m} A_*^\epsilon \leq B_\epsilon \stackrel{\text{def}}{=} \frac{1}{m} \sum_{j=1}^m B_j^\epsilon \in \widehat{\mathcal{M}}_\epsilon$. Hence,

$$\frac{1}{m} \sigma(\mathcal{M}) \leq \frac{1}{m} \sigma(\mathcal{M}_\epsilon) \stackrel{(2.8)}{=} \frac{1}{m} \rho(A_*^\epsilon) \leq \rho(B_\epsilon) \stackrel{(2.6)}{\leq} \rho(\mathcal{M}_\epsilon).$$

Note that $m \leq p$. Since the JSR is a continuous function of the elements of the compact set \mathcal{M} (see [1] and an alternative proof in [12]), we obtain the first inequality in (2.10).

In order to prove the second inequality, assume that \mathcal{M} contains an irreducible matrix A_0 . Then the minimizer $x(A_0)$ of function $\phi_{A_0}(x)$ at the hyperplane $\mathcal{L} \stackrel{\text{def}}{=} \{x \in R^n : \langle \mathbf{1}, x \rangle = 0\}$ is unique. Therefore, the corresponding restrictions of the level sets of this function are bounded. Hence, the objective function of problem (2.4) also has bounded restrictions of its level sets. Consequently, there exists at least one optimal solution $x(\mathcal{M})$ of problem (2.4) belonging to the hyperplane \mathcal{L} . Note that all components of $u(\mathcal{M}) \stackrel{\text{def}}{=} e^{x(\mathcal{M})}$ are positive. Therefore,

$$(2.11) \quad A^T u(\mathcal{M}) \leq \sigma(\mathcal{M}) \cdot u(\mathcal{M}), \quad A \in \mathcal{M}.$$

Define now the vector norm

$$\|x\| = \sum_{i=1}^n u^{(i)}(\mathcal{M}) \cdot |x^{(i)}|.$$

Then, for any $x \in R^n$ and any $A \in \mathcal{M}$, we have

$$\|Ax\| \leq \|A|x|\| = \langle u(\mathcal{M}), A|x| \rangle \stackrel{(2.11)}{\leq} \sigma(\mathcal{M}) \cdot \|x\|.$$

This means that $\rho(\mathcal{M}) \leq \sigma(\mathcal{M})$.

If \mathcal{M} does not contain an irreducible matrix, we can consider the sets \mathcal{M}_ϵ with $\epsilon > 0$. By the above reasoning, we have justified that

$$\rho(\mathcal{M}) \leq \rho(\mathcal{M}_\epsilon) \leq \sigma(\mathcal{M}_\epsilon).$$

It remains to apply Lemma 1. \square

Let us now look at two examples.

Example 1. Consider $\mathcal{M} = \{A_1, \dots, A_n\}$, where the matrix $A_i = \mathbf{1} \cdot e_i^T$. Note that

$$A_i \cdot A_j = A_j$$

for any $i, j = 1, \dots, n$. Therefore, $\rho(\mathcal{M}) = \rho(A_i) = 1$. On the other hand,

$$\begin{aligned} \sigma(\mathcal{M}) &= \inf_{u > \mathbf{0}} \max_{1 \leq i \leq n} \max_{1 \leq j \leq n} \frac{1}{u^{(j)}} \langle A_i e_j, u \rangle \\ &= \inf_{u > \mathbf{0}} \max_{1 \leq j \leq n} \frac{1}{u^{(j)}} \langle \mathbf{1}, u \rangle \stackrel{(2.1)}{=} \rho(\mathbf{1} \cdot \mathbf{1}^T) = n. \end{aligned}$$

Hence, the lower bound in inequality (2.10) cannot be improved. It is interesting that in this example the bound provided by $\sigma_T(\mathcal{M})$ is exact:

$$\begin{aligned} 1 &= \rho(\mathcal{M}) \stackrel{(2.10)}{\leq} \sigma_T(\mathcal{M}) \\ &\stackrel{(2.7)}{=} \inf_{u > \mathbf{0}} \max_{1 \leq i \leq n} \max_{1 \leq j \leq n} \frac{1}{u^{(j)}} \langle A_i^T e_j, u \rangle \\ &= \inf_{u > \mathbf{0}} \max_{1 \leq i \leq n} \max_{1 \leq j \leq n} \frac{u^{(i)}}{u^{(j)}} \leq 1. \end{aligned}$$

In dimension 2 the JCR can be characterized in a very transparent way.

Example 2. Let \mathcal{M} be a finite set of nonnegative two-by-two matrices:

$$\mathcal{M} = \{A_k = (a_k, b_k), k = 1, \dots, m\}.$$

Then

$$\begin{aligned} \sigma(\mathcal{M}) &= \inf_{x \in \mathbb{R}^2} \max_{1 \leq k \leq m} \phi_{A_k}(x) = \inf_{u > \mathbf{0}} \max_{1 \leq k \leq m} \max \left\{ \frac{\langle a_k, u \rangle}{u^{(1)}}, \frac{\langle b_k, u \rangle}{u^{(2)}} \right\} \\ &= \inf_{u > \mathbf{0}} \max \left\{ \max_{1 \leq k \leq m} \frac{\langle a_k, u \rangle}{u^{(1)}}, \max_{1 \leq k \leq m} \frac{\langle b_k, u \rangle}{u^{(2)}} \right\}, \\ \left(\tau \stackrel{\text{def}}{=} \frac{u^{(2)}}{u^{(1)}} \right) &= \min_{\tau > 0} \max \left\{ \max_{1 \leq k \leq m} (a_k^{(1)} + a_k^{(2)} \cdot \tau), \max_{1 \leq k \leq m} \left(b_k^{(1)} \cdot \frac{1}{\tau} + b_k^{(2)} \right) \right\}. \end{aligned}$$

Thus, the value of $\sigma(\mathcal{M})$ is easy to find after an appropriate sorting of the coefficients. Note that this solution is a nontrivial function of the initial data.

It is interesting that sometimes, for a richer set of variants in the switching system, it is possible to ensure that the JCR and the JSR coincide. We give a nontrivial example of such a situation in the next section.

3. Sets with independent column uncertainty. Combining the result of Lemma 2 with the upper bound (2.10), and the observation (2.6), we can see that the inclusion $A_* \in \widehat{\mathcal{M}}$ guarantees that the JSR and the JCR coincide. The simplest way to ensure this inclusion is to assume that the set \mathcal{M} has independent column uncertainties:

$$(3.1) \quad \mathcal{M} = \{(a_1, \dots, a_n) : a_i \in Q_i, i = 1, \dots, n\},$$

where all sets $Q_i \subset \mathbb{R}^n, i = 1, \dots, n$, are closed and bounded.

THEOREM 2. For a set \mathcal{M} satisfying condition (3.1), we have

$$(3.2) \quad \max_{A \in \widehat{\mathcal{M}}} \rho(A) = \rho(\mathcal{M}) = \sigma(\mathcal{M}).$$

If the solution of problem (2.4) exists, then $\sigma(\mathcal{M}) = \rho(A_*)$ for some extreme point A_* of the set \mathcal{M} . Therefore,

$$(3.3) \quad \max_{A \in \widehat{\mathcal{M}}} \rho(A) = \rho(\mathcal{M}) = \sigma(\mathcal{M}) = \rho(A_*).$$

Proof. Let us fix some $\epsilon > 0$. Under assumption (3.1), the matrix A_*^ϵ in (2.8) belongs to the set $\widehat{\mathcal{M}}_\epsilon$. Therefore,

$$\rho(\mathcal{M}_\epsilon) \stackrel{(2.10)}{\leq} \sigma(\mathcal{M}_\epsilon) \stackrel{(2.8)}{=} \rho(A_*^\epsilon) \stackrel{(2.6)}{\leq} \rho(\mathcal{M}_\epsilon).$$

It remains to use the continuity arguments.

Thus, we have proved that

$$\max_{A \in \widehat{\mathcal{M}}_\epsilon} \rho(A) = \rho(\mathcal{M}_\epsilon) = \sigma(\mathcal{M}_\epsilon) = \rho(A_*^\epsilon).$$

Let us show that it is always possible to choose A_*^ϵ as an extreme point of the set \mathcal{M}_ϵ . For this, let us look more carefully at the structure of the optimality condition for problem (2.9) written for the set \mathcal{M}_ϵ . Denote $\widehat{Q}_i^\epsilon \stackrel{\text{def}}{=} \text{Conv}(Q_i^\epsilon)$, $i = 1, \dots, n$. Then the KKT conditions are as follows:

$$(3.4) \quad \begin{aligned} \mathbf{0} &= \sum_{i=1}^n \lambda_*^{(i)} g_i^*, \quad \lambda_* \in \Delta_n, \\ g_i^* \in \partial \psi_i(x_*) &= \left\{ g = \nabla(e^{-x^{(i)}} \langle a, e^{x^*} \rangle) : a \in B_i^*(x_*) \right\}, \\ B_i^*(x_*) &= \text{Arg max}_{a \in \widehat{Q}_i^\epsilon} \langle a, e^{x^*} \rangle, \quad i = 1, \dots, n. \end{aligned}$$

It is clear that each $B_i^*(x_*)$ contains at least one extreme point from Q_i^ϵ .

Let us show first that $\lambda_* > \mathbf{0}$. Note that g_i^* can be represented as

$$g_i^* = e^{-x_*^{(i)}} D(a_i^*) \cdot e^{x_*} - e^{-x_*^{(i)}} \cdot \langle a_i^*, e^{x_*} \rangle \cdot e_i$$

with certain $a_i^* \in \widehat{Q}_i^\epsilon$. Therefore,

$$\mathbf{0} = D \left(\sum_{i=1}^n \lambda_*^{(i)} e^{-x_*^{(i)}} a_i^* \right) \cdot e^{x_*} - \left(\sum_{i=1}^n \lambda_*^{(i)} e^{-x_*^{(i)}} \cdot e_i \cdot (a_i^*)^T \right) e^{x_*}.$$

Denote $A_* = (a_1^*, \dots, a_n^*) \in \widehat{\mathcal{M}}_\epsilon$ and $\hat{A}_* = D(e^{x_*}) A_* D(e^{-x_*})$. Then, the above first-order optimality conditions can be written as follows:

$$\begin{aligned} \mathbf{0} &= (D(e^{x_*}) A_* D(e^{-x_*}) D(\lambda_*) - D(\lambda_*) D(e^{-x_*}) A_*^T D(e^{x_*})) \cdot \mathbf{1} \\ &= (\hat{A}_* D(\lambda_*) - D(\lambda_*) \hat{A}_*^T) \cdot \mathbf{1}. \end{aligned}$$

Since $A_* > \mathbf{0}$, all elements of the matrix \hat{A}_* are also positive. Therefore, the above equation is possible only with $\lambda_* > \mathbf{0}$.

Let us fix an arbitrary number $k, 1 \leq k \leq n$. Without loss of generality, assume that $k = 1$. Let us fix all $g_i^* \in \partial\psi_i(x_*)$, $i = 2, \dots, n$. It is convenient to represent them in matrix form:

$$(g_2^*, \dots, g_n^*) \stackrel{\text{def}}{=} G = - \begin{pmatrix} b^T \\ C \end{pmatrix},$$

where $G \in R^{n \times (n-1)}$ and $C \in R^{(n-1) \times (n-1)}$. Note that each $g_i \in \partial\psi_i(x_*)$ can be represented as follows:

$$g_i = \sum_{j=1}^n a^{(j)} e^{x_*^{(j)} - x_*^{(i)}} (e_j - e_i), \quad a \in B_i^*(x^*) \subset \text{int } R_+^n.$$

Note that $\langle \mathbf{1}, g_i \rangle = 0$, and $g_i^{(j)} > 0$ for $j \neq i$ with $g_i^{(i)} < 0$. Therefore, matrix C is strictly diagonally dominant and has negative off-diagonal entries. Hence, the matrix C^{-1} exists and all its elements are positive. Therefore, the (overdetermined) linear system

$$g_1 = G\lambda$$

has a positive solution for any $g_1 \in \partial\psi_1(x^*)$ (the first equation of this system is a linear consequence of the others). This implies that in the optimality conditions (3.4) we can choose g_1^* using an arbitrary $a_1 \in B_1^*(x^*)$. In particular, we can choose it as an extreme point of Q_1^ϵ .

This reasoning can be repeated for the remaining indices $k = 2, \dots, n$. Then, we end up with an extreme point $A_*^\epsilon \in \mathcal{M}_\epsilon$. Note that the latter set is obtained from \mathcal{M} by adding the same ϵ to all entries of the matrices. Hence, when ϵ goes to zero, the shape of Q_1^ϵ remains unchanged. Therefore, in view of the continuity of the function $\rho(A)$, any limit point of extreme points A_*^ϵ is an extreme point of the set \mathcal{M} . \square

Besides its direct applications, Theorem 2 has an interesting algebraic consequence.

COROLLARY 1. *Consider the function $\rho(a_1, \dots, a_n)$ with $a_i \in R_+^n, i = 1, \dots, n$. Then this function is quasi-convex in each a_i when all other columns are fixed.*

Proof. Let us fix arbitrary positive vectors x, y, a_2, \dots, a_n . Define the uncertainty set

$$\mathcal{M} = \{A(\alpha) = (\alpha x + (1 - \alpha)y, a_2, \dots, a_n), \alpha \in [0, 1]\}.$$

Then, by Theorem 2, $\rho(\mathcal{M}) = \max\{\rho(A(0)), \rho(A(1))\}$. If the above vectors are non-negative, we can justify the result by continuity arguments. \square

Let us note that independent row (or column) uncertain sets arise in a very natural way. One such situation is in the context of asynchronous linear systems.

Example 3. Let A be a given nonnegative matrix, and consider the row independent uncertainty set

$$\mathcal{A} = \{M : \text{the } i\text{th row of } M \text{ is given by the } i\text{th row of } A \text{ or by the } i\text{th row of } I\}.$$

Thus, the switched linear system $x_{k+1} = A_k x_k$ with $M_k \in \mathcal{A}$ corresponds to asynchronous updates of the state entries. We claim that $\rho(\mathcal{A}) = \max\{1, \rho(A)\}$.

This result easily follows from our construction. Indeed,

$$\begin{aligned}
 \rho(\mathcal{A}) &\leq \min_{x \leq 0} \max_i \max \left\{ 1, \frac{1}{x_i} \langle a_i, x \rangle \right\} \\
 &= \min_{x \leq 0} \max \left\{ 1, \max_i \frac{1}{x_i} \langle a_i, x \rangle \right\} \\
 &= \max \left\{ 1, \min_{x \leq 0} \max_i \frac{1}{x_i} \langle a_i, x \rangle \right\} \\
 &= \max\{1, \rho(A)\}.
 \end{aligned}$$

4. Applications. Let us now consider two applications of our results.

4.1. Leontief model with uncertain data. In the input-output static Leontief model, we have n industries with production levels

$$p^{(i)} \geq 0, \quad i = 1, \dots, n.$$

In order to produce one unit of the product of industry i , we need to spend $A^{(i,j)}$ units of the product of industry j . Thus, the structure of production dependencies is given by the nonnegative *consumption matrix* $A \in M_n^+$. Further, given the demand vector $d \in R_+^n$, we can find the necessary level of production as a solution to the system of linear equations:

$$(4.1) \quad p = Ap + d.$$

The economics is called *productive* if (4.1) has a nonnegative solution for any demand vector d . The standard condition for that is, of course,

$$(4.2) \quad \rho(A) < 1.$$

Clearly, the smaller $\rho(A)$ is, the more flexible is the economics, and the smaller is the production level that is necessary for satisfying the current demand pattern. However, in practice the spectral radius $\rho(A)$ is difficult to estimate. Indeed, from the statistics of the previous years, we can get only sets of *different* consumption matrices

$$\mathcal{M} = \{A_1, \dots, A_k\}.$$

Hence, the JSR $\rho(\mathcal{M})$ is a good and robust measure of economic flexibility. However, in general this value is difficult to compute even for small dimensions.

Now we have an alternative way to address this problem. Indeed, for each industry i , using the data of the previous years, we can get examples of the *distribution patterns*.³ Denoting the convex hull of these vectors by Q_i , we can form the uncertainty set by (3.1). In this case, by Theorem 2, $\rho(\mathcal{M}) = \sigma(\mathcal{M})$, and it can be efficiently computed.

4.2. Optimal growth of a linear system. In some applications we need to find a matrix from a certain set, which has maximal spectral radius. If the number of matrices in the set is large, or infinite, such a problem looks completely intractable.

³Alternatively, we can use statistics on the *consumption patterns*, which corresponds to the uncertain rows of matrix A .

However, it appears that in the case of independent column (or row) uncertainties we can solve the problem by applying Theorem 2. For example, the Boolean problem

$$\max_{x \in \{0,1\}^n} \rho(AD(x) + BD(\mathbf{1} - x)),$$

where the coefficients of matrices A and B are nonnegative, can be solved by convex programming techniques. Note that for each column we can consider several variants. A typical application of this type is the design of a network structure with several independent variants of connections for each node. Multiplication of an input vector by such a matrix represents the output of the system. In this case, the best design corresponds to the matrix with the largest possible spectral radius.

Acknowledgments. We wish to thank John Tsitsiklis and Vladimir Protasov for interesting discussions. In particular, they have independently provided alternative elementary proofs for the observation made in Example 3. We also wish to thank the anonymous referee for interesting comments and for pointing out useful references.

REFERENCES

- [1] N. E. BARABANOV, On the *Lyapunov exponent of discrete inclusions*. I–III, *Automat. Remote Control*, 49 (1988), pp. 152–157, 283–287, 558–565.
- [2] M. A. BERGER AND Y. WANG, *Bounded semigroups of matrices*, *Linear Algebra Appl.*, 166 (1992), pp. 21–27.
- [3] V. D. BLONDEL AND Y. NESTEROV, *Computationally efficient approximations of the joint spectral radius*, *SIAM J. Matrix Anal. Appl.*, 27 (2005), pp. 256–272.
- [4] V. D. BLONDEL, Y. NESTEROV, AND J. THEYS, *On the accuracy of the ellipsoid norm approximation of the joint spectral radius*, *Linear Algebra Appl.*, 394 (2005), pp. 91–107.
- [5] V. D. BLONDEL AND V. CANTERINI, *Undecidable problems for probabilistic automata of fixed dimension*, *Theory Comput. Syst.*, 36 (2003), pp. 231–245.
- [6] V. D. BLONDEL, R. JUNGERS, AND V. PROTASOV, *On the complexity of computing the capacity of codes that avoid forbidden difference patterns*, *IEEE Trans. Inform. Theory*, 52 (2006), pp. 5122–5127.
- [7] V. D. BLONDEL AND J. N. TSITSIKLIS, *The boundedness of all products of a pair of matrices is undecidable*, *Systems Control Lett.*, 41 (2000), pp. 135–140.
- [8] V. D. BLONDEL, J. CASSAGNE, AND R. M. JUNGERS, *On the number of α -power-free binary words for $2 < \alpha \leq 7/3$* , *Theoret. Comput. Sci.*, 410 (2009), pp. 2823–2833.
- [9] D. COLLELA AND C. HEIL, *Characterization of scaling functions: Continuous solutions*, *SIAM J. Matrix Anal. Appl.*, 15 (1994), pp. 496–518.
- [10] V. CRESPI, G. V. CYBENKO, AND G. JIANG, *The Theory of Trackability with Applications to Sensor Networks*, Technical report TR2005-555, Department of Computer Science, Dartmouth College, Hanover, NH, 2005. Available online at <http://www.ists.dartmouth.edu/library/157.pdf>
- [11] I. DAUBECHIES AND J. C. LAGARIAS, *Two-scale difference equations. II. Local regularity, infinite products of matrices and fractals*, *SIAM J. Math. Anal.*, 23 (1992), pp. 1031–1079.
- [12] C. HEIL AND G. STRANG, *Continuity of the joint spectral radius: Applications to wavelets*, in *Linear Algebra for Signal Processing*, IMA Vol. Math. Appl. 69, Springer-Verlag, New York, 1995, pp. 51–61.
- [13] R. HORN AND C. JOHNSON, *Matrix Analysis*, Cambridge University Press, Cambridge, UK, 1985.
- [14] R. M. JUNGERS, *Infinite Matrix Products. From the Joint Spectral Radius to Combinatorics*, Ph.D. thesis, Université catholique de Louvain, Louvain-la-Neuve, Belgium, 2008.
- [15] R. M. JUNGERS, V. Y. PROTASOV, AND V. D. BLONDEL, *Overlap-free words and spectra of matrices*, *Theoret. Comput. Sci.*, to appear.
- [16] R. JUNGERS, V. PROTASOV, AND V. BLONDEL, *Efficient algorithms for deciding the type of growth of products of integer matrices*, *Linear Algebra Appl.*, 428 (2008), pp. 2296–2311.
- [17] B. E. MOISION, A. ORLITSKY, AND P. H. SIEGEL, *On codes that avoid specified differences*, *IEEE Trans. Inform. Theory*, 47 (2001), pp. 433–442.
- [18] YU. NESTEROV, *Introductory Lectures on Convex Optimization*, Kluwer, Boston, 2004.

- [19] P. A. PARRILO AND A. JADBABAIE, *Approximation of the joint spectral radius using sum of squares*, *Linear Algebra Appl.*, 428 (2008), pp. 2385–2402.
- [20] E. PLISCHKE AND F. WIRTH, *Duality results for the joint spectral radius and transient behavior*, *Linear Algebra Appl.*, 428 (2008), pp. 2368–2384.
- [21] V. Y. PROTASOV, *A generalized joint spectral radius: A geometric approach*, *Izv. Math.*, 61 (1997), pp. 995–1030.
- [22] G.-C. ROTA AND W. G. STRANG, *A note on the joint spectral radius*, *Indag. Math.*, 22 (1960), pp. 379–381.
- [23] R. SHORTEN, F. WIRTH, O. MASON, K. WULFF, AND C. KING, *Stability criteria for switched and hybrid systems*, *SIAM Rev.*, 49 (2007), pp. 545–592.
- [24] J. N. TSITSIKLIS AND V. D. BLONDEL, *The Lyapunov exponent and joint spectral radius of pairs of matrices are hard—when not impossible—to compute and to approximate*, *Math. Control Signals Systems*, 10 (1997), pp. 31–40.
- [25] F. WIRTH, *The generalized spectral radius and extremal norms*, *Linear Algebra Appl.*, 342 (2002), pp. 17–40.