The Lyapunov exponent and joint spectral radius of pairs of matrices are hard – when not impossible – to compute and to approximate*

John N. Tsitsiklis†   Vincent D. Blondel‡

Abstract

We analyse the computability and the complexity of various definitions of spectral radii for sets of matrices. We show that the joint and generalized spectral radii of two integer matrices are not approximable in polynomial time, and that two related quantities – the lower spectral radius and the largest Lyapunov exponent – are not algorithmically approximable.

Key words: Lyapunov exponent, Lyapunov indicator, joint spectral radius, generalized spectral radius, discrete differential inclusion, computational complexity, NP-hard, algorithmic solvability.

*This work was completed while Blondel was visiting Tsitsiklis at MIT. This research was supported by the ARO under grant DAAL-03-92-G-0115.

†Laboratory for Information and Decision Systems, Massachusetts Institute of Technology, Cambridge, MA 02139, USA; email: jnt@mit.edu

‡Institute of Mathematics, University of Liège, B-4000 Liège, Belgium; email: vblondel@ulg.ac.be

1 Introduction

The *spectral radius* of a real matrix A is defined by

$$\rho(A) := \max\{|\lambda| : \lambda \text{ is an eigenvalue of } A\}.$$

This definition can be extended in various ways to sets of matrices. Due to their numerous practical applications, these possible extensions have been the object of intense attention in recent years. In this paper we analyse some of these extensions from a computational complexity point of view.

Let $\|\cdot\|$ be any matrix norm (in the sequel we always assume that matrix norms are submultiplicative, i.e., that $\|AB\| \leq \|A\|\|B\|$). The well-known identity $\rho(A) = \lim_{k\to\infty} \|A^k\|^{1/k}$ (see for example [HJ, Corollary 5.6.14]) justifies the generalizations of the concept of spectral radius to sets of matrices given next. Let Σ be a set of matrices in $R^{n\times n}$; the *joint spectral radius* $\overline{\rho}(\Sigma)$ is defined [RS] by

$$\overline{\rho}(\Sigma) = \limsup_{k\to\infty} \overline{\rho}_k(\Sigma),$$

where $\overline{\rho}_k(\Sigma) = \sup\{\|A_1A_2\cdots A_k\|^{1/k} : \text{each } A_i \in \Sigma\}$ for $k \geq 1$. It is shown in [DL] (notice that our notations are different from those used there) that $\overline{\rho}(\Sigma) \leq \overline{\rho}_k(\Sigma)$ for all $k \geq 1$, and therefore the joint spectral radius can be given in the equivalent form $\overline{\rho}(\Sigma) = \lim_{k\to\infty} \overline{\rho}_k(\Sigma)$. Similarly to $\overline{\rho}$ we define the *lower spectral radius* $\underline{\rho}(\Sigma)$ by

$$\underline{\rho}(\Sigma) = \liminf_{k\to\infty} \underline{\rho}_k(\Sigma),$$

where $\underline{\rho}_k(\Sigma) = \inf\{\|A_1A_2\cdots A_k\|^{1/k} : \text{each } A_i \in \Sigma\}$ for $k \geq 1$.

As for the single matrix case, the quantities $\overline{\rho}_k(\Sigma)$ and $\underline{\rho}_k(\Sigma)$ generally depend on the matrix norm used but the limiting values $\overline{\rho}(\Sigma)$ and $\underline{\rho}(\Sigma)$ do not. To see this, remember that any two submultiplicative norms $\|.\|_1$ and $\|.\|_2$ are related by $\alpha\|A\|_1 \leq \|A\|_2 \leq \beta\|A\|_1$ for some $0 < \alpha < \beta$. For any product $A_1A_2\cdots A_k$ one has $\alpha^{1/k}\|A_1A_2\cdots A_k\|_1^{1/k} \leq$

$\|A_1A_2\cdots A_k\|_2^{1/k} \leq \beta^{1/k}\|A_1A_2\cdots A_k\|_1^{1/k}$ and by letting k tend to infinity we conclude that the joint and lower spectral radii are well defined independently of the matrix norm used.

The joint and lower spectral radii correspond in a certain sense to two extreme cases. With the joint spectral radius we calculate the largest possible average norm that can be obtained by multiplying matrices from Σ, whereas with the lower spectral radius we calculate the lowest possible such norm. We now define an additional quantity that is intermediate between these two extreme cases. Let us assume that we have a probability distribution P over the set Σ and that we generate an infinite sequence $(A_i)_{i\geq 1}$ of elements of Σ by picking each matrix A_i randomly and independently according to the assumed probability distribution P. A probability distribution will be said *nontrivial* if nonzero probabilities are attached to all matrices of Σ. The *largest Lyapunov exponent* (also called *top Lyapunov exponent* or *asymptotic growth rate*) associated with P and Σ is defined by (see [O], see also [CKN] for a more readable account):

$$\lambda(\Sigma, P) = \lim_{k\to\infty} \frac{1}{k} E\Big[\log(\|A_1\cdots A_k\|)\Big].$$

It can be shown that this limit exists and, as for the previous cases, does not depend on the matrix norm used (see [O] for the a proof of the first of these statements). In order for our development to be uniform we transform the largest Lyapunov exponent into the *Lyapunov spectral radius* $\rho_P(\Sigma)$ by defining

$$\rho_P(\Sigma) = e^{\lambda(\Sigma,P)}.$$

Basic inequalities relating $\underline{\rho}, \rho_P$ and $\overline{\rho}$ are given by

$$\underline{\rho}(\Sigma) \leq \rho_P(\Sigma) \leq \overline{\rho}(\Sigma).$$

Moreover, since $\rho(A) = \lim_{k\to\infty} \|A^k\|^{1/k}$, the definitions of $\underline{\rho}, \rho_P$ and $\overline{\rho}$ coincide with the usual spectral radius when Σ consists of a single matrix.

Additional definitions similar to those of $\underline{\rho}$, ρ_P and $\overline{\rho}$ are possible by replacing the norm appearing in their definitions by a spectral radius. One obtains, for example, the *generalized spectral radius* $\overline{\rho}'(\Sigma)$ defined by Daubechies and Lagarias [DL] by setting

$$\overline{\rho}'(\Sigma) = \limsup_{k\to\infty} \overline{\rho}'_k(\Sigma),$$

where $\overline{\rho}'_k(\Sigma) = \sup\{(\rho(A_1 A_2 \cdots A_k))^{1/k} : \text{each } A_i \in \Sigma\}$ for $k \geq 1$. Similar definitions lead to the spectral quantities $\underline{\rho}'$ and ρ'_P. It has been conjectured in [DL] and established by Berger and Wang that the generalized spectral radius $\overline{\rho}'$ coincides with the joint spectral radius $\overline{\rho}$ when Σ is finite (see [BW, Theorem IV], or [E, Theorem 1] for an elementary proof). Gurvits has also shown [G2, Theorem B.1] that $\underline{\rho}'$ coincides with $\underline{\rho}$ when Σ is finite. In the sequel we shall always assume that the set Σ is finite and shall, for convenience, deal only with the three spectral radii defined in terms of norms.

The generalized spectral radius was introduced in Daubechies and Lagarias [DL] for studying concepts associated to Markov chains, random walks, and wavelets. The logarithm of the joint spectral radius also appears in the context of discrete linear inclusions where it is called *Lyapunov indicator*, see for example [B1]. In systems theoretic terms, the generalized spectral radius can be associated with the stability properties of time-varying systems in the worst case over all possible time variations, or with the stability of "asynchronous" [T2] or "desynchronised" [K2] systems.

The definition of the lower spectral radius is natural for formalizing control design notions associated to discrete-time systems. Instead of viewing the order of matrix multiplication as an externally imposed time variation, we view it as a control action, and we are interested in the stability properties that can be obtained by choosing control actions in the best possible way. Despite this natural interpretation, the definition of the lower spectral radius seems quite recent (the first reference seem to be [G2], see also [BT2] for connections with control concepts).

Finally, the largest Lyapunov exponent appears in the context of discrete linear differential inclusions (see [BGFB] and references therein) and is related to time-varying systems in which time variations are random. Besides systems-theoretic interpretations, Lyapunov exponents are pervasive in many areas of applied mathematics such as mathematical demography [C,R2], percolation processes [D], and Kalman filtering [B2]. Other references and descriptions of applications appear in the yearly conference proceedings [A] and in the survey [CKN].

We now briefly survey how these quantities can be computed or approximated. By letting k tend to infinity, the inequalities (with our notations)

$$\overline{\rho}'_k(\Sigma) \leq \overline{\rho}(\Sigma) \leq \overline{\rho}_k(\Sigma)$$

proved in [DL, Lemma 3.1] can be used to derive algorithms which compute arbitrarily precise approximations for $\overline{\rho}(\Sigma)$ (see for example [G1] for one such algorithm).

These approximating algorithms can in turn be used in procedures that decide, after finitely many steps, whether $\overline{\rho} > 1$ or $\overline{\rho} < 1$ (such procedures are given, e.g., by Brayton and Tong [BT3] in a system theory context and by Barabanov [B1] in the context of discrete linear inclusions). These procedures may not terminate when $\overline{\rho}$ happens to be equal to 1 The existence of algorithms for computing arbitrarily precise approximations of $\overline{\rho}$ does not rule out the possibility that the decision problem "$\overline{\rho} < 1$" is undecidable. It is so far unknown whether this problem, which was the original motivation for the research reported in this paper, is algorithmically solvable (see [LW] for a discussion of this issue and for a description of its connection with the finiteness conjecture, see also the discussion in [G2]). A negative result in this direction is given by Kozyakin who shows [K2] that the set of pairs of 2×2 matrices that have a joint spectral radius less than one is not semialgebraic.

In our first result (Theorem 1) we show that, unless $P = NP$, approximating algorithms for $\overline{\rho}$ can not possibly run in polynomial time. More precisely, we show that, unless $P = NP$, there is no algorithm that can compute $\overline{\rho}(\Sigma)$ with a relative error bounded by $\epsilon > 0$, in time polynomial in the size of Σ and ϵ (see later for more precise definitions). As a corollary we show that it is NP-hard to decide if all possible products of two given matrices are stable.

The situation for the largest Lyapunov exponent and for the lower spectral radius are somewhat different from that of the joint spectral radius. Computable upper bounds for ρ_P for the case where Σ consists of nonnegative matrices are given in Gharavi and Anantharam [GA] and analytic solutions are available for special cases (see for example [LR] for an analytic solution for the case where Σ consists of two 2×2 matrices one of which is singular). In general, no exact, or even approximate, computational methods other than simulation seem to be available for computing ρ_P or $\underline{\rho}$. The problem of computing ρ_P has been known for at least 20 years, and we quote from Kingman [K1, p. 897] (the same quotation appears in [C]): "Pride of place among the unsolved problems of subadditive ergodic theory must go to the calculation of the constant γ (a constant that is equal to the logarithm of ρ_P). In none of the applications described here is there an obvious mechanism for obtaining an exact numerical value, and indeed this usually seems to be a problem of some depth."

In our second result (Theorem 2) we show that no approximating algorithm exists for $\underline{\rho}$ and ρ_P. More precisely, let ρ be any function satisfying

$$\underline{\rho}(\Sigma) \leq \rho(\Sigma) \leq \rho_P(\Sigma)$$

for some nontrivial probability distribution P and for all Σ. We show that the problem of computing ρ exactly, or even approximately, is algorithmically undecidable. We also show that, when all the matrices in Σ are constrained to have nonnegative coefficients, then the problem of computing ρ becomes NP-hard.

If the decision problem "$\rho < 1$" was decidable for such a function ρ, then the associated decision procedure could be used to compute arbitrary precise approximations of ρ. Since ρ is not computable when $\underline{\rho} \leq \rho \leq \rho_P$, we conclude, as a corollary to Theorem 2, that "$\rho < 1$" is undecidable for the Lyapunov spectral radius, for the lower spectral radius, and for all intermediate functions between these two.

For convenience of the exposition we shall restrict our attention in the sequel to *pairs* of matrices with integer entries. Our results being negative they equally apply to sets of $k \geq 2$ matrices or to infinite sets, and to matrices with real entries.

A earlier version of this paper appears in the conference proceedings [TB].

2 Approximability of the joint spectral radius

As explained in the introduction, the joint spectral radius can be approximated to arbitrary precision. We show in this section that, unless $P = NP$, approximating algorithms cannot run in polynomial-time. Following Papadimitriou [P1], we say that a function $\rho(\Sigma)$ is *polynomial-time approximable* if there exists an algorithm $\rho^*(\Sigma, \epsilon)$, which, for every rational number $\epsilon > 0$ and every set of matrices Σ with $\rho(\Sigma) > 0$, returns an approximation of $\rho(\Sigma)$ with a relative error of at most ϵ (i.e., such that $|\rho^* - \rho| \leq \epsilon|\rho|$) in time polynomial in the size of Σ and ϵ. By "size of Σ and ϵ" we mean the description size, or "bit size", of Σ and ϵ. For example, if ϵ is the ratio of two relatively prime numbers p and q, the size of ϵ can be taken to be $\log(pq)$.

Theorem 1 Unless $P = NP$, the joint (generalized) spectral radius $\overline{\rho}$ of two matrices is not polynomial-time approximable. This is true even for the special case where Σ consists of two matrices with $\{0, 1\}$ entries.

Proof. Our proof proceeds by reduction from the classical SAT problem (see [GJ] for a definition of SAT), it is inspired from the proof of Theorem 6 in [PT] and it is similar to the proof of Theorem 2 in [BT1] (however, we were unable to deduce this theorem from Theorem 2 in [BT1].)

Starting from an instance of SAT with n variables $x_1, \dots, x_n$ and m clauses $C_1, \dots, C_m$, we construct two directed graphs G_0 and G_1. The graphs have identical nodes but have different edges. Besides the start node s, there is a node u_{ij} associated to each clause C_i and variable x_j, a 0-th node u_{0j} associated to each variable x_j, and a $(n+1)$-th node $u_{i(n+1)}$ associated to each clause C_i. Edges are constructed as follows: for $i = 1, \dots, m$ and $j = 1, \dots, n$ there is

- an edge $(u_{ij}, u_{i(j+1)})$ in both G_0 and G_1 if the variable x_j does not appear in clause C_i;
- an edge (u_{ij}, u_{0j}) in G_0 and an edge $(u_{ij}, u_{i(j+1)})$ in G_1 if the variable x_j appears in clause C_i negatively;
- an edge (u_{ij}, u_{0j}) in G_1 and an edge $(u_{ij}, u_{i(j+1)})$ in G_0 if the variable x_j appears in clause C_i positively.

For $i = 1, \dots, m$ there are edges (s, u_{i1}) in both graphs. The graphs have edges $(u_{0j}, u_{0(j+1)})$ for $j = 1, \dots, n-1$ and have an edge from u_{0n} to s. There are no edges leaving from $(u_{i(n+1)}, s)$ for $i = 1, \dots, m$.

Let r denote the total number of nodes ($r = (n+1)(m+1)$). We construct two $r \times r$ matrices A_0 and A_1. Associated to the graph G_0 (respectively, G_1) is the $r \times r$ matrix A_0 (respectively, A_1) whose (i,j)-th entry is equal to 1 if there is an edge from node j to node i in G_0 (respectively G_1), and is equal to zero otherwise.

To any given node α we associate a column-vector $x(\alpha)$ of dimension r whose entries are all zero with the exception of the entry corresponding to the node α which is equal to one.

We need two observations.

1. Let a partition of the nodes be given by $P_{n+2} = \{s\}$, $P_{n+1} = \{u_{i1} : i = 1, \dots, m\}$, $P_n = \{u_{01}, u_{i2} : i = 1, \dots, m\}, \dots, P_2 = \{u_{0(n-1)}, u_{in} : i = 1, \dots, m\}$ and $P_1 = \{u_{0n}, u_{i(n+1)} : i = 1, \dots, m\}$. We use ℓ_α to denote the index of the partition to which the node α belongs. Any edge (from G_0 or G_1) leaving from a node of partition P_h goes to a node of partition P_{h-1}. Furthermore, the unique edge leaving from partition P_1 goes back to partition P_{n+2}. Thus, any path in G_0 and G_1 that starts from node α either gets to a node $u_{i(n+1)}$, from which there is no outgoing edge, or it visits node s after ℓ_α steps. In matrix terms this implies the following. Let α be an arbitrary node and let ℓ_α be its associated partition index. If h is a positive integer equal to ℓ_α modulo $(n+2)$ and A is a product of h factors in $\{A_0, A_1\}$, then

$$Ax(\alpha) = \mu x(s)$$

for some nonnegative scalar μ.

2. Let $q_1, \dots, q_n \in \{0, 1\}$ be a truth assignment of the boolean variables x_j and consider the product $A_{q_n} \cdots A_{q_1}$. The vector $A_{q_n} \cdots A_{q_1} x(u_{i1})$ is equal to $x(u_{0n})$ if the clause C_i is satisfied and is equal to $x(u_{i(n+1)})$ otherwise. Let A_* be any of A_0 or A_1. There are no edges leaving from $u_{i(n+1)}$, there is one edge from u_{0n} to s, and there are edges from s to u_{i1} for $i = 1, \dots, m$. Thus we have $A_* x(u_{i(n+1)}) = 0$, $A_* x(u_{0n}) = x(s)$, and $A_* x(s) = \sum_{i=1}^m x(u_{i1})$. From this we conclude

$$A_* A_{q_n} \cdots A_{q_1} A_* x(s) = A_* A_{q_n} \cdots A_{q_1} \sum_{i=1}^m x(u_{i1}) = A_* \sum_{i=1}^m A_{q_n} \cdots A_{q_1} x(u_{i1}) = \lambda x(s)$$

where λ is equal to the number of clauses that are satisfied by the given truth assignment.

With these two observations we now prove the theorem.

Assume first that the instance of SAT is satisfied by the assignment $x_i = q_i$ for $q_1, \dots, q_n \in$

$\{0,1\}$ and define A by $A = A_* A_{q_n} \cdots A_{q_1} A_*$ with A_* any of A_0 or A_1. Since all m clauses are satisfied we have $Ax(s) = mx(s)$ and thus $\overline{\rho}(A_0, A_1) \geq m^{1/(n+2)}$.

Assume now that the instance of SAT is not satisfiable. Let $y_i = \sum_{\alpha \in P_i} x(\alpha)$ for $i = 1, \ldots, n+2$ and consider the vector max norm $\|\cdot\|$. Let A be a product of $n+2$ factors in $\{A_0, A_1\}$. Since the instance of SAT is not satisfiable we have $\|Ay_i\| \leq (m-1)\|y_i\| = m-1$ for $i = 1, \ldots, n+2$. Let now e denote the vector whose entries are all equal to one. Then $e = \sum_{i=1}^{n+2} y_i$ and $Ae = \sum_{i=1}^{n+2} Ay_i$. The nonzero entries of Ay_i are at the same place as the nonzero entries of y_i. Hence, $\|Ae\| = \|\sum Ay_i\| = \max_i \|Ay_i\| \leq m-1$. The entries of A are all nonnegative and so $\|A\| = \|Ae\|$ for the max row sum matrix norm. Thus we have $\|A\| \leq m-1$ whenever A is a product of $n+2$ factors in $\{A_0, A_1\}$. From this we conclude that $\overline{\rho}(A_0, A_1) \leq (m-1)^{1/(n+2)}$.

Suppose now that $\rho^*(\Sigma, \epsilon)$ is an algorithm which, for every $\epsilon > 0$ and Σ with $\rho(\Sigma) > 0$, returns an approximation of $\rho(\Sigma)$ with $|\rho^* - \rho| \leq \epsilon|\rho|$. By running this algorithm on the pair of $\{0,1\}$ matrices A_0, A_1 obtained from the instance and on a sufficiently small ϵ (e.g., we can take $\epsilon < (m/(m-1))^{1/(n+2)} - 1$), we are able to distinguish $\overline{\rho}(A_0, A_1) \geq m^{1/(n+2)}$ from $\overline{\rho}(A_0, A_1) \leq (m-1)^{1/(n+2)}$. The algorithm thus allows us to decide the instance of SAT. Since all transformation are performed in polynomial time, the algorithm cannot possibly run in time polynomial in the size of Σ and ϵ unless $P = NP$. $\square$

Remarks:

1. Since the problem remains NP-hard when the matrices have $\{0,1\}$ entries, a corollary of the theorem is the following:

Corollary 1 Unless $P = NP$, the joint (or generalized) spectral radius of two $n \times n$ matrices, with $\{0,1\}$ entries, is not approximable with relative error 10^{-k} (k positive integer) in a number of operations polynomial in n and k.

2. If a polynomial-time algorithm was available for checking the stability of all products of two given matrices, then the algorithm could be used to approximate the joint spectral radius in polynomial time. Thus we have:

Corollary 2 Consider all possible products of two given real matrices A_0 and A_1. It is NP-hard to decide if all products are stable. This is true even if the matrices have $\{0, 1\}$ entries.

3. As indicated by a reviewer it may be possible to improve the theorem by proving that, for a suitably small constant ϵ, and unless $P = NP$, no polynomial time approximation algorithm of relative error ϵ exists for $\overline{\rho}$. Such a result would have to be derived from negative results on the approximability of the MAX-SAT problem.

4. L. Gurvits has kindly communicated to us that he has also proved Corollary 2 using a different reduction (unpublished).

3 Approximability of the lower spectral radius and the Lyapunov exponent

In this section we show that the lower spectral radius and the Lyapunov spectral radius, and intermediate quantities between these two, cannot be approximated by an algorithm. Let ρ be a quantity that we wish to compute and let us fix some positive constants K and L with $L < 1$. Consider an algorithm which on input Σ outputs a number $\rho^*(\Sigma)$. We say that this algorithm is a (K, L)-*approximation algorithm* if for every Σ we have

$$|\rho^* - \rho| \leq K + L\rho.$$

This definition allows for an *absolute error* of K and a *relative error* of L. Despite the latitude allowed by this definition, we show below that (K, L)-approximation algorithms do not exist for the Lyapunov and the lower spectral radii.

In order to prove our result we shall need the following definition. We say that a set of matrices Σ is *mortal* if there exists some $k \geq 1$ and matrices $A_i \in \Sigma$ such that $A_1 A_2 \cdots A_k = 0$. The following result can be found in [BT1] (see, in particular, Theorem 1, Theorem 2, and Remark 2 after Theorem 1) and builds on an earlier result by Paterson [P2].

Proposition. Mortality of two integer matrices of size $n \times n$ is undecidable for $n = 6n_p + 6$ where n_p is any number of pairs of words for which Post's correspondence problem is undecidable. (We may take $n_p = 7$, see below.)

Post's correspondance problem is a classical undecidable problem on words (for a description of the problem and a proof of its undecidability see, e.g., Hopcroft and Ullman [H]). In a recent contribution Matiyasevich and Sénizergues [M] have shown that Post's correspondance problem is undecidable as soon as $n_p \geq 7$. Thus we can take $n_p = 7$, and mortality of pairs of 48×48 integer matrices is undecidable. We are now able to prove our theorem. The proof essentially uses the fact that any (K, L)-approximation algorithm can be used to decide mortality of matrices.

Theorem 2 Let n_p be a number of pairs of words for which Post's correspondance problem is undecidable. Fix any $K > 0$ and L with $0 \leq L < 1$. Let ρ be a function defined on pairs of matrices and assume that $\underline{\rho}(\Sigma) \leq \rho(\Sigma) \leq \rho_P(\Sigma)$ for some nontrivial probability distribution P and for all pairs Σ.

1. There exists no (K, L)-approximation algorithm for computing ρ. This is true even for the special case where Σ consists of one $(6n_p + 7) \times (6n_p + 7)$ integer matrix and one $(6n_p + 7) \times (6n_p + 7)$ integer diagonal matrix.

2. For the special cases where Σ consists of two integer matrices with $\{0, 1\}$ entries, there exists no polynomial time (K, L)-approximation algorithm for computing ρ unless

$P = NP$.

Proof. Let $K > 0$ and $0 \leq L < 1$ be given and ρ be as above. Suppose that there exists a (K, L)-approximation algorithm for ρ and let Σ be an arbitrary family of $n \times n$ integer matrices.

We claim that the (K, L)-approximation algorithm can be used to decide whether or not Σ is mortal. This will establish the theorem.

We form a family Σ' of $(n+1) \times (n+1)$ matrices as follows. For each $A \in \Sigma$, we construct $B \in \Sigma'$ by letting $B = \mathrm{diag}\{cA, d\}$, where c and d are positive constants satisfying $K + d(L+1) < (1-L)c - K$.

Suppose that Σ is mortal. Then, it is easily seen that $\underline{\rho}(\Sigma') = \rho_P(\Sigma') = d$ and thus $\rho(\Sigma') = d$. In this case, applying a (K, L)-approximation algorithm to Σ', would give a result ρ^* bounded by $\rho^* \leq K + (L+1)d$.

Suppose now that Σ is not mortal. The matrices in Σ' have integer entries that are either equal to zero, or are larger than c. Since Σ is not mortal, any product of k matrices has some entry whose magnitude is at least c^k and it follows that $\underline{\rho}(\Sigma') \geq c$ and thus $\rho(\Sigma') \geq c$. In this case, applying a (K, L)-approximation algorithm to Σ', would give a result ρ^* satisfying $\rho - \rho^* \leq L\rho + K$ or $\rho^* \geq (1-L)\rho - K \geq (1-L)c - K$.

Having chosen c and d so that $K + d < (1-L)c - K$, the result of the (K, L)-approximation algorithm applied to Σ' allows us to determine whether Σ is mortal or not.

The mortality problem is undecidable even for the case where Σ consists of two $(6n_p + 6) \times (6n_p + 6)$ integer matrices. The fact that one of the matrices may be taken diagonal follows from the observation that the Lyapunov exponent and lower spectral radius are left unchanged by similarity transformation of the matrices, combined with the fact that the matrices used in the paper [P2], to which [BT1] refers, are all diagonalisable. The first part of the theorem is therefore proved.

For proving the second part of the theorem, we invoke the same reduction and use the fact that checking whether two matrices with $\{0,1\}$ entries are mortal is an NP-complete problem. $\square$

Remarks:

1. Note that the matrices in Σ are not irreducible. It is not clear whether a similar negative result can be obtained if we restrict the set Σ to irreducible matrices.

2. If an algorithm was available for checking the presence of a stable matrix in the set of all products of two given matrices, then the algorithm could be used to approximate the lower spectral radius. Thus we have:

Corollary 3 Consider all possible products of two given real matrices A_0 and A_1. It is undecidable to decide if one of the products is stable. This is true even if the two matrices are integer, of size 49×49, and one of them is diagonal.

References

[A] L. Arnold, H. Crauel and J.-P. Eckmann (eds.), Lyapunov exponents, Proceedings of a conference held in Oberwolfach, Lecture notes in mathematics, vol. 1486, Springer-Verlag, New York, 1991.

[B1] N. E. Barabanov, Lyapunov indicators of discrete inclusions, part I, II and III, Translation from Avtomatika i Telemekhanika, **2** (1988), 40-46, **3** (1988), 24-29 and **5** (1988), 17-24.

[BGFB] S. Boyd, L. El Ghaoui, E. Feron and V. Balakrishnan, Linear matrix inequalities in system and control theory, SIAM studies in applied mathematics, vol. 15, SIAM, Philadelphia, 1994.

[BW] M. Berger and Y. Wang, Bounded semigroup of matrices, Linear Algebra Appl.,

166 (1992), 21-27.

[BT1] V. D. Blondel and J. N. Tsitsiklis, When is a pair of matrices mortal?, Technical report LIDS-P-2314, LIDS, MIT, January 1996, to appear in Information Processing Letters.

[BT2] V. D. Blondel and J. N. Tsitsiklis, Complexity of stability and controllability of elementary hybrid systems, preprint (1997).

[B2] P. Bougerol, Filtre de Kalman Bucy et exposants de Lyapounov, in [A].

[BT3] R. Brayton and C. Tong, Constructive stability and asymptotic stability of dynamical systems, IEEE Transactions on Circuits and Systems, **27** (1980), 1121-1130.

[C] J. Cohen, Subadditivity, generalized products of random matrices and operation research, SIAM Review, **30** (1988), 69-86.

[CKN] J. Cohen, H. Kesten and M. Newman (eds), Random matrices and their applications, Contemporary mathematics, vol. 50, American Mathematical Society, Providence, 1986.

[D] R. Darling, The Lyapunov exponent for product of infinite-dimensional random matrices, in [A]

[DL] I. Daubechies and J. C. Lagarias, Sets of matrices all infinite products of which converge, Linear Algebra Appl., **162** (1992), 227-263.

[E] L. Elsner, The generalized spectral radius theorem: an analytic-geometric proof, Linear Algebra Appl., **220** (1995), 151-159.

[GJ] M. R. Garey and D. S. Johnson, Computers and Intractability: A Guide to the Theory of NP-completeness, Freeman and Co., New York, 1979.

[G1] G. Gripenberg, Computing the joint spectral radius, Linear Algebra Appl., **234** (1996), 43-60.

[GA] R. Gharavi and V. Anantharam, An upper bound for the largest Lyapunov exponent of a Markovian random matrix product of nonnegative matrices, preprint (1995).

[G2] L. Gurvits, Stability of discrete linear inclusion, Linear Algebra Appl., **231** (1995), 47-85.

[H] J. E. Hopcroft and J. D. Ullman, Formal languages and their relation to automata, Addison-Wesley, 1969.

[HJ] R. A. Horn and C. R. Johnson, Matrix Analysis, Cambridge University Press, Cambridge, 1985.

[K1] J. Kingman, Subadditive ergodic theory, Ann. Probab., **1** (1976), 883-909.

[K2] V. S. Kozyakin, Algebraic unsolvability of problem of absolute stability of desynchronized systems, Translation from Avtomatika i Telemekhanika, **6** (1990), 41-47.

[LW] J. C. Lagarias and Y. Wang, The finiteness conjecture for the generalized spectral radius of a set of matrices, Linear Algebra Appl., **214** (1995), 17-42.

[LR] R. Lima and M. Rahibe, Exact Lyapunov exponent for infinite products of random matrices, J. Phys. A: Math. Gen., **27** (1994), 3427-3437.

[M] Y. Matiyasevich and G. Sénizergues, Decision problems for semi-Thue systems with a few rules, preprint, 1996.

[O] V. I. Oseledec, A multiplicative ergodic theorem. Lyapunov characteristic numbers for dynamical systems, Trans. Moscow Math. Soc., **19** (1968), 197-231.

[P1] C. H. Papadimitriou, Computational complexity, Addison-Wesley, Reading, 1994.

[PT] C. H. Papadimitriou and J. N. Tsitsiklis, The complexity of Markov decision processes, Math. Oper. Res., **12** (1987), 441-450.

[P2] M. Paterson, Unsolvability in 3×3 matrices, Studies in Applied Mathematics, **49** (1970), 105-107.

[R1] K. Ravishankar, Power law scaling of the top Lyapunov exponent of a product of random matrices, J. Statistical Physics, **54** (1989), 531-537.

[R2] J. Roerdink, The biennal life strategy in a random environment, J. Math. Biol., **26** (1988), 199-215.

[RS] G.-C. Rota and G. and Strang, A note on the joint spectral radius, Indag. Math., **22** (1960), 379-381.

[TB] J. N. Tsitsiklis and V. D. Blondel, The spectral radius of a pair of matrices is hard to compute, Proceedings of the 35th Conference on Decision and Control, Kobe, December 1996, pp. 3192-3197.

[T1] J. N. Tsitsiklis, On the control of discrete-event dynamical systems, Math. Control, Signals, and Systems, **2** (1989), 95-107.

[T2] J. N. Tsitsiklis, On the stability of asynchronous iterative processes, Math. Systems Theory, **20** (1987), 137-153.