Approximating the spectral radius of sets of matrices in the max-algebra is NP-hard

Vincent D. Blondel, Stéphane Gaubert, John N. Tsitsiklis

Abstract—The lower and average spectral radii measure respectively the minimal and average growth rates of long products of matrices taken from a finite set. The logarithm of the average spectral radius is traditionally called Lyapunov exponent. When one performs these products in the max-algebra, we obtain quantities that measure the performance of Discrete Event Systems. We show that approximating the lower and average max-algebraic spectral radii is NP-hard.

I. INTRODUCTION

For all positive real numbers \( p \), the semiring \( \mathbb{R}_p \) is the set of real nonnegative numbers, \( \mathbb{R}^+ \), equipped with the addition

\[
a +_p b = \max\{a^p + b^p, a, b\},
\]
together with the usual multiplication (of course, when \( p \) is an odd integer, \( \mathbb{R}_p \) can be embedded in the field \( \mathbb{R}^+_p \times \), but for the decision issues studied here, the specialization to nonnegative elements is essential). This family of semirings was introduced independently by Maslov and Pap (see e.g. [20], [22] and the references therein). It has the following remarkable property: all the semirings \( \mathbb{R}_p \) are isomorphic to the ordinary semiring \( \mathbb{R}_1 \) of real nonnegative numbers equipped with the usual operations. Letting \( p \) tend to \( \infty \) in (1), we obtain:

\[
a + \infty = \max\{a, b\}.
\]

The corresponding semiring \( \mathbb{R}_\infty \) (the set \( \mathbb{R}^+ \), equipped with \( +_\infty \) and the usual multiplication) is the max-times semiring or “max-algebra”, whose role in dynamic programming, discrete event system theory, optimal control, and asymptotic analysis is well known (see e.g. [1], [21], [20], [16], [19]). In contrast to the semirings \( \mathbb{R}_p \) for finite \( p \), this semiring is not isomorphic to \( \mathbb{R}_1 \). In discrete event systems applications, the max-algebra more frequently appears in an isomorphic additive form, the semiring \( \mathbb{R}_{max} \), which is the set \( \mathbb{R} \cup \{-\infty\} \), equipped with \( + \) as addition, and \( * \) as multiplication. The isomorphism is given by \( x \mapsto \log x : \mathbb{R}_{\infty} \rightarrow \mathbb{R}_{max} \). To emphasize the parallel with existing results, we will state all our results in terms of \( \mathbb{R}_\infty \) (see Table I).

In the sequel, we will use the familiar algebraic notation in the context of the semiring \( \mathbb{R}_p \), without further comments: e.g., if \( A \in \mathbb{R}_p^{m \times n} \) and \( B \in \mathbb{R}_p^{n \times p} \), \( AB \) is the \( r \times t \) matrix with entries \( A_{ij} = A_{i1}B_{1j} + p \cdot A_{i2}B_{2j} + \ldots + p \cdot A_{it}B_{tj} \). Let \( \| \cdot \| \) denote a (conventional) norm on \( \mathbb{R}^{r \times t} \). To a finite set of matrices \( \{A_1, \ldots, A_t\} \subset \mathbb{R}_p^{m \times n} \), we associate:

\[
\rho_{\max}(A_1, \ldots, A_t) \overset{\text{def}}{=} \lim_{k \to \infty} \max_{i_1, \ldots, i_t} \|A_{i_1} \cdots A_{i_t}\|^{\frac{1}{k}}, \quad (2a)
\]

\[
\rho_{\min}(A_1, \ldots, A_t) \overset{\text{def}}{=} \lim_{k \to \infty} \min_{i_1, \ldots, i_t} \|A_{i_1} \cdots A_{i_t}\|^{\frac{1}{k}}, \quad (2b)
\]

\[
\rho_E(A_1, \ldots, A_t) \overset{\text{def}}{=} \text{a.s.} \lim_{k \to \infty} \|A_{i_1} \cdots A_{i_t}\|^{\frac{1}{k}}, \quad (2c)
\]

where in (2c), \( i_1, i_2, \ldots \) is sequence of independently, identically distributed, random variables with values in \( \{1, \ldots, t\} \), drawn with the uniform distribution, and where “a.s. limit” means that the limit exists almost surely. The existence and values of all the limits in (2) are clearly independent of the choice of the norm. In particular, we may take the norm \( \|A\| = \max_{i<j} \|A_{ij} + p \cdot A_{i2} + \ldots + p \cdot A_{ir}\| \) which satisfies \( \|AB\| \leq \|A\|\|B\| \). Then, by a classical argument, the existence of the limit (2a) follows easily from the fact that the sequence \( w_k = \max_{i_1, \ldots, i_t} \|A_{i_1} \cdots A_{i_t}\| \) is submultiplicative, i.e., \( w_{k+l} \leq w_k w_l \). The existence of \( \rho_{\min} \) is proved by the same argument. As shown in [7] and [1], Chap. 7], the existence of \( \rho_E \) follows from Kingman’s subadditive ergodic theorem. We will call \( \rho_{\max}, \rho_{\min} \) and \( \rho_E \) the upper, lower, and average spectral radius of \( \{A_1, \ldots, A_t\} \), respectively. The logarithm of \( \rho_E \) is traditionally called the Lyapunov exponent or Lyapunov indicator. We note that, trivially,

\[
\rho_{\min} \leq \rho_E \leq \rho_{\max}.
\]

When \( p = 1 \), both the upper and average spectral radius are much studied quantities which are notoriously difficult to compute or approximate in practice. In [4, Th. 1 and 2], it was shown that even in the case of two matrices \( A_0, A_1 \) with entries in \( \{0,1\} \), approximating \( \rho_{\max}, \rho_{\min} \) and \( \rho_E \) is NP-hard.

Using the fact that all the semirings \( \mathbb{R}_p \) with finite \( p \) are isomorphic, it follows that analogous results hold for all semirings \( \mathbb{R}_p \). In [4, Th. 1 and 2] it was also shown that, if we allow the entries of \( A_0, A_1 \) to be in \( \mathbb{Z} \), there is no algorithm that can distinguish between instances with \( \rho_{\min} = 0 \) from instances with \( \rho_{\min} = 1 \), and similarly for \( \rho_E \). In particular, \( \rho_{\min} \) and \( \rho_E \) cannot be approximated algorithmically, and the problem of deciding whether they are zero is undecidable. The situation for \( \rho_{\max} \) is different.

Using the inequalities derived in [8], it is immediate to see

\footnote{A problem \( A \) is NP-hard if it is at least as hard as some NP-complete problem \( B \), in the sense that \( B \) can be reduced to \( A \) in polynomial time. A polynomial time algorithm for an NP-hard problem would provide polynomial time algorithms for all NP-complete problems, and would imply that the conjecture \( P \neq NP \) is false; see [11] for more details. This conjecture is widely believed to be true.}
that $\rho_{\text{max}}$ can be approximated algorithmically to arbitrary precision. One such algorithm is given in [18].

When $p = \infty$, the quantities $\rho_{\text{max}}, \rho_{\text{min}},$ and $\rho_E$ have been much studied by the Discrete Event Systems community. As shown in [1], the Lyapunov exponent $\log \rho_E$ measures the cycle time (inverse of the throughput) of random max-plus linear discrete event systems. The most intuitive particular interpretation of $\rho_E$ is probably the following: if you "play" a Tetris game of infinite height, without applying any control, just letting pieces fall down randomly, you will see, asymptotically, the heap of pieces grow at a certain mean speed: this speed is precisely $\log \rho_E$ (see [12], [6], [14], [9], and also [9] below for details). The problem of computing $\rho_E$ also arises in Statistical Physics, in the study of disordered systems ($\log \rho_E$ yields the free energy per site, at zero temperature, for some random one-dimensional Ising models [10]).

The logarithm of $\rho_{\text{max}}$ was called worst case Lyapunov exponent in [12], for it measures the worst case cycle time of certain max-plus linear discrete event systems. For a dual reason, the logarithm of $\rho_{\text{min}}$ was called optimal case Lyapunov exponent.

Since the maximization operation which is involved in the definition of $\rho_{\text{max}}$ is somehow compatible with the structure laws of the max-algebra, $\rho_{\text{max}}$ can be computed quite easily: as shown in [12], it coincides with the spectral radius of the single matrix $A = A_1 + \cdots + A_J$, which can be computed in polynomial time. So far, the basic general technique to compute $\rho_{\text{min}}$ and $\rho_E$ consists of using an "induced Markov chain" construction in the max-algebraic projective space [1, §8.4],[12, §VII]: when this chain is finite, both $\rho_{\text{min}}$ and $\rho_E$ can be computed with a number of arithmetic operations which is polynomial in the number of states of the chain. In some other special cases, $\rho_E$ can also be computed via generating series techniques [17], or, as illustrated in [6], by finding a closed form expression for the invariant measure of the above mentioned Markov chain, which is denumerable, in general. A different approach was used in [2]: we can define more generally $\rho_E$ in (2c) by taking a sequence of independent, identically distributed, random variables $i_1, \ldots, i_k$, drawn from $\{1, \ldots, J\}$ with a non uniform distribution $\pi = (\pi_1, \pi_2, \ldots, \pi_J)$, where $\pi_j$ is the probability of $\{i_1 = j\}$.

Under some technical restrictions, $\rho_E$ is an analytic function of $\pi_1, \ldots, \pi_J$ near $\pi = (1/J, \ldots, 1/J)$, and the coefficients of its power series expansion can be effectively computed. When this series is still convergent at $\pi = (1/J, \ldots, 1/J)$, this gives a way of approximating the average spectral radius.

The purpose of this paper is to analyze the complexity of computing $\rho_{\text{min}}$ and $\rho_E$ when $p = \infty$.

In a first section (§II below) we show that, when $p = \infty$, approximating $\rho_{\text{min}}$ or $\rho_E$ is NP-hard. Our proof of this result is based on a reworking of the argument given in [4, Proof of Th. 1]. We build an automaton whose number of accepting paths measures the number of satisfied clauses in a given instance of the satisfiability problem SAT.

Our proof then follows from the fact that the satisfiability problem SAT is known to be NP-complete (see the problem L01 in [11]) and that the number of accepting paths in this special automaton determines the spectral radius of an associated set of matrices.

This argument does not work when $p = \infty$: since $+\infty$ is idempotent (i.e. $a + \infty = a$), several paths count as one. However, a variant of the reduction of [5, Proof of Th. 2] can be used to prove that approximating $\rho_{\text{min}}$ and $\rho_E$ is NP-hard.

In a second section (§III below) we give a simple, independent, geometrical argument that shows that computing $\rho_{\text{min}}$ is NP-hard. The argument is based on an intuitive interpretation of products of matrices in terms of the height of a heap of pieces. In [13], [14], it was shown that the total height of a Tetris-like heap of $k$ pieces is equal to $\log \|A_{i_1} \cdots A_{i_k}\|$, where $A_{i_1}, \ldots, A_{i_k}$ are matrices associated to the pieces, and $\|A\| = \max_{ij} A_{ij}$. When all the pieces are of height 1, $\log \rho_{\text{min}}$ coincides with the inverse of the largest number of mutually disjoint pieces. NP-hardness of computing $\rho_{\text{min}}$ then follows from the fact that computing the largest number of mutually disjoint pieces is a problem that is known to be NP-hard.

II. REDUCTION FROM SAT

In the remaining part of the paper, we will assume that $p = \infty$ and we will use the matrix norm $\|A\| = \max_{ij} A_{ij}$.

Let $\Sigma \mapsto \rho(\Sigma)$ be a non-negative function that we wish to compute. We say that $\rho$ is polynomial-time approximable if there exists an algorithm which, for every rational numbers $\epsilon, \epsilon' > 0$ and every $\Sigma$, returns an approximation $\rho^*(\Sigma, \epsilon, \epsilon')$ such that $|\rho^* - \rho| \leq \epsilon \rho + \epsilon'$, in time polynomial in the description size of $\epsilon, \epsilon'$ and $\Sigma$. This allows for both an absolute and a relative error.

**Theorem 1.** Unless P=NP, the lower and average spectral radii of pairs of matrices with entries in $\{0,1\}$ are not polynomial-time approximable.

**Proof:** Let $A_1, A_2$ be square matrices with entries in $\{0,1\}$. We claim that

$$\rho_{\text{min}}(A_1, A_2) = \rho_E(A_1, A_2) \in \{0, 1\}.$$  

Indeed, in the max-algebra, any product of matrices with entries in $\{0,1\}$ gives a matrix with entries in $\{0,1\}$. A fortiori, $\|A_{i_1} \cdots A_{i_k}\| \in \{0,1\}$ for all $i_1, \ldots, i_k$. Hence, if none of the products $A_{i_1} \cdots A_{i_k}$ is 0, $\rho_{\text{min}}(A_1, A_2) = \rho_E(A_1, A_2) = 1$. But if one of these products is 0, then $\rho_{\text{min}}(A_1, A_2) = 0$ and the product that gives 0 will appear almost surely as a factor of any infinite product $A_{j_1} A_{j_2} \ldots$ of independent, identically distributed, random matrices, drawn from $\{A_1, A_2\}$ with the uniform distribution. This implies that $\rho_E(A_1, A_2) = 0$.

Due to (4), it suffices to establish the theorem for $\rho_{\text{min}}$. Any polynomial time approximation algorithm for $\rho_{\text{min}}$ gives a polynomial time algorithm for distinguishing the cases $\rho_{\text{min}} = 0$ and $\rho_{\text{min}} = 1$. Thus, in order to establish
the theorem, it suffices to show that the problem of determining whether \( \rho_{\min}(A_1, A_2) = 0 \) is NP-hard, even for the case of binary matrices. The proof is by reduction from SAT and is inspired by [5, Proof of Th. 2].

Consider an instance of SAT [11], with \( n \) variables \( x_1, \ldots, x_n \) and \( m \) clauses \( C_1, \ldots, C_m \). We can write each clause \( C_i \) as \( C_i = C_{i,1} \lor \cdots \lor C_{i,n} \), where \( C_{i,j} \) is either \( x_j \), or \( \neg x_j \), or the Boolean constant \( \text{false} \).

Let \( C = C_1 \land \cdots \land C_m \). For any \( y \in \{\text{true, false}\} \) and \( k \in \{1, \ldots, n\} \), let \( M_k(y) \) denote the diagonal \( m \times m \)

<table>
<thead>
<tr>
<th>Boolean matrix</th>
<th>( \rho_{\max} )</th>
<th>( \rho_{\min} ) and ( \rho_{\min} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( (\mathbb{R}, +, \times) )</td>
<td>Approximation algorithm [8]</td>
<td>No approximation algorithm [4]</td>
</tr>
<tr>
<td>( \mathbb{R}_p = (\mathbb{R}^+, +p, \times) ) (finite ( p ))</td>
<td>Approximation is NP-hard [4]</td>
<td>Approximation is NP-hard [4]</td>
</tr>
<tr>
<td>( \mathbb{R}_\infty = (\mathbb{R}^+, +\infty, \times) )</td>
<td>Exact polynomial time algorithm [12]</td>
<td>Approximation is NP-hard [this paper]</td>
</tr>
</tbody>
</table>

### Table I

**SUMMARY OF COMPLEXITY RESULTS AVAILABLE FOR \( \rho_{\max}, \rho_{\min}, \rho_{ \min} \).**


Then, for all \( x \in \{\text{true, false}\}^n \),

\[
M_1(x_1) \cdots M_n(x_n) = 0 \text{ if and only if } x \text{ satisfies } C. \tag{5}
\]

Let \( U \) denote the \( m \times m \) matrix whose entries are all equal to one. We now consider the \( nm \times nm \) matrices \( A_1 = M(\text{false}), A_2 = M(\text{true}) \), where

\[
M(y) = \begin{pmatrix}
U M_1(y) & M_2(y) & \cdots & M_n(y) \\
M_1(y) U & \cdots & M_n(y) U \\
M_n(y) U & \cdots & M_1(y) U
\end{pmatrix}
\]

( the blocks which are not shown are zero). We claim that

- \( C \) is not satisfiable \( \iff \rho_{\min}(A_1, A_2) = 1 \),
- \( C \) is satisfiable \( \iff \rho_{\min}(A_1, A_2) = 0 \).

In order to establish our claim, note first that for all \( k \) and for all Boolean sequences \( y \) of length \( kn \), \( M(y_1) \cdots M(y_{kn}) \) is a block diagonal matrix with diagonal blocks:

\[
B_{1,k} = U M_1(y_1) \cdots M_n(y_n) U U M_1(y_{kn+1}) \cdots M_n(y_{kn}) U, \tag{6a}
\]

\[
B_{2,k} = M_2(y_1) \cdots U M_1(y_1) M_2(y_{kn+1}) \cdots U M_1(y_{kn}), \tag{6b}
\]

\[
\vdots
\]

\[
B_{n,k} = M_n(y_1) U \cdots M_{n-1}(y_n) M_n(y_{kn+1}) \cdots M_{n-1}(y_{kn}). \tag{6c}
\]

Assume that \( C \) is not satisfiable. Using (5), we get that \( B_{1,k} \) is 1 for all possible Boolean sequences \( y \) of length \( kn \). This implies that \( \rho_{\min}(A_1, A_2) = 1 \).

Next, assume that \( C \) is satisfied by the Boolean sequence \( x_1 \ldots x_n \), and consider the infinite sequence of period \( n+1 \):

\[
y = x_1 \cdots x_n x_1 \cdots x_n x_1 \cdots x_n \cdots
\]

where \( \xi \) can take an arbitrary Boolean value. For \( k = n+1 \), each of the \( n \) products that give \( B_{1,k}, \ldots, B_{n,k} \) in (6) contains a factor of the form \( M_1(x_1) \cdots M_n(x_n) \). Since \( M_1(x_1) \cdots M_n(x_n) = 0 \) we conclude that \( \rho_{\min}(A_1, A_2) = 0 \).

**Remark.** It is not known whether the statement of the theorem remains valid if we require that the matrices have positive entries, or have a fixed, large enough, dimension.

### III. Reduction from SET PACKING

In Discrete Event Systems applications, the quantity of interest is the logarithm of \( \rho_{\min} \), rather than \( \rho_{\min} \). In this section we show that the following problem is NP-hard.

**Problem (Computing \( \rho_{\min} \)).**

**Instance:** Matrices \( A_1, \ldots, A_t \in \{0, 1, 2\}^{m \times n} \), a rational number \( q \).

**Question:** Does \( \log_2 \rho_{\min}(A_1, \ldots, A_t) < q \)?

**Theorem 2.** Computing \( \rho_{\min} \) is NP-hard.

**Proof:** The proof is based on a simple geometrical argument that involves a Tetris-like heap of pieces.

Consider a horizontal axis with \( n \geq 1 \) slots labelled \( \{1, 2, \ldots, n\} \). A piece is a solid, possibly disconnected, block of height one occupying some of the slots. Consider now a set of pieces \( A = \{a_1, \ldots, a_t\} \) each piece \( a_i \) being defined by the subset \( R(a_i) \subseteq \{1, 2, \ldots, n\} \) of slots it occupies. To an ordered sequence of pieces \( w = a_{i_1} \cdots a_{i_k} \) we associate a heap by piling up the pieces in the given order on a horizontal ground. Pieces are only subject to vertical translations and occupy the lowest possible position that is above the ground and above the pieces previously piled up. The height of a heap \( w \) on slot \( i \) is denoted by \( h_i(w) \).

The height \( h(w) \) of a heap \( w = a_{i_1} \cdots a_{i_k} \) is the largest of the heights on all slots. For instance, when \( n = 3 \), \( A = \{a_1, a_2, a_3\} \), \( R(a_1) = \{1, 2\}, R(a_2) = \{3\}, R(a_3) = \{1, 3\} \), and \( w = a_1 a_3 a_2 a_1 a_3 \), we obtain the heap with height \( h(w) = 4 \) depicted at the right of Fig. 1.

To \( k \geq 1 \), we associate the lowest possible height of a heap of \( k \) pieces taken from \( A \)

\[
\lambda_k = \min \{ h(a_{i_1} \cdots a_{i_k}) \mid a_{i_1}, \ldots, a_{i_k} \in A \},
\]

We claim that the limit

\[
\lambda = \lim_{k \to +\infty} \frac{\lambda_k}{k}
\]

is equal to \( 1/M \), where \( M \) is equal to the maximal number of pieces in a heap of height one. Indeed, a heap \( w \)
IV. Conclusion

Of course, the interest of the NP-hardness results of this paper is mostly theoretical: Theorems 1 and 2 show that there is little hope to find a polynomial algorithm to compute $p_\text{e}$ or $\rho_{\text{min}}$. But the situation seems much simpler in the case of the max-algebra, $\mathbb{R}_\infty$, than in the case of the usual algebra $(\mathbb{R}, +, \times)$. For instance, as summarized in Table 1 above, the problem of approximating $\rho_{\text{max}}$, which is NP-hard in $(\mathbb{R}, +, \times)$ becomes polynomially solvable in $\mathbb{R}_\infty$. Moreover, in this paper, we only proved that in the semiring $\mathbb{R}_\infty$, approximating $\rho_{\text{min}}$ or $p_\text{e}$ is NP-hard; this is a weak “impossibility” result, by comparison to the fact that the corresponding problems in $(\mathbb{R}, +, \times)$ are undecidable. Indeed, unlike in the usual algebra $(\mathbb{R}, +, \times)$, in the max-algebra, $\rho_{\text{min}}$ and $p_\text{e}$ can be approximated (with an exponential execution time), at least in some important special cases [17], [12], [15], [2]. Improving and generalizing these algorithms, as well as identifying new examples of exactly solved models, is certainly an interesting research direction.

REFERENCES